Genetic code

The set of correspondences between nucleotide pair triplets in DNA and amino acids in protein.

A group of three following bases (nucleotides) is called **triplet (codon)**. A triplet is always read from the 5´ end of the mRNA, so for example the triplet for tryptophane is: 5´– **U G G** – 3´

There are 64 types of codons, from these only 61 code amino acids. The first, **Start triplet**, in mRNA is always **AUG**, which codes for methionine. Three of these triplets don't code for any amino acid and have an important role during the termination of synthesis of the polypeptide chain. Those are the so called termination codons or **stop codons**. They are: **UAA, UAG** and **UGA**.

Number of letters in the code

In reading an mRNA molecule from one particular end, only one of four different bases, A, U, G, or C, can be found at each position. Thus, if the words were one letter long, only four words would be possible. This vocabulary cannot be the genetic code, because we must have a word for each of the 20 amino acids commonly found in cellular proteins. If the words were two letters long, then $4^2 = 16$ words would be possible; for example, AU, CU, or CC. This vocabulary is still not large enough.

If the words are three letters long, then $4^3 = 64$ words are possible; for example, AUU, GCG, or UGC. This vocabulary provides more than enough words to describe the amino acids. We can conclude that the code word must consist of at least three nucleotide pairs. However, if all words are "triplets," then we have a considerable excess of possible words over the 20 needed to name the common amino acids.

The genetic code is:

1-Tripleted – three nitrogen bases of nucleotides following one after another in the mRNA (or in the DNA,) coding for one amino acid (except the stop triplets).

2- Universal – present in all organisms.

3-not overlapping – triplets follow one after the other in the DNA (in a linear setting) without interruption.

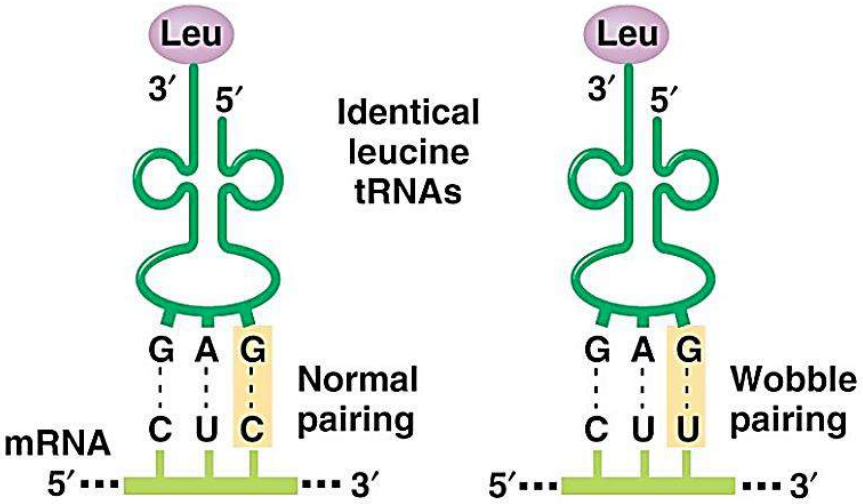
4- Degenerated- It simply means that each of the 64 triplets must have some meaning within the code; so at least some amino acids must be specified by two or more different triplets.

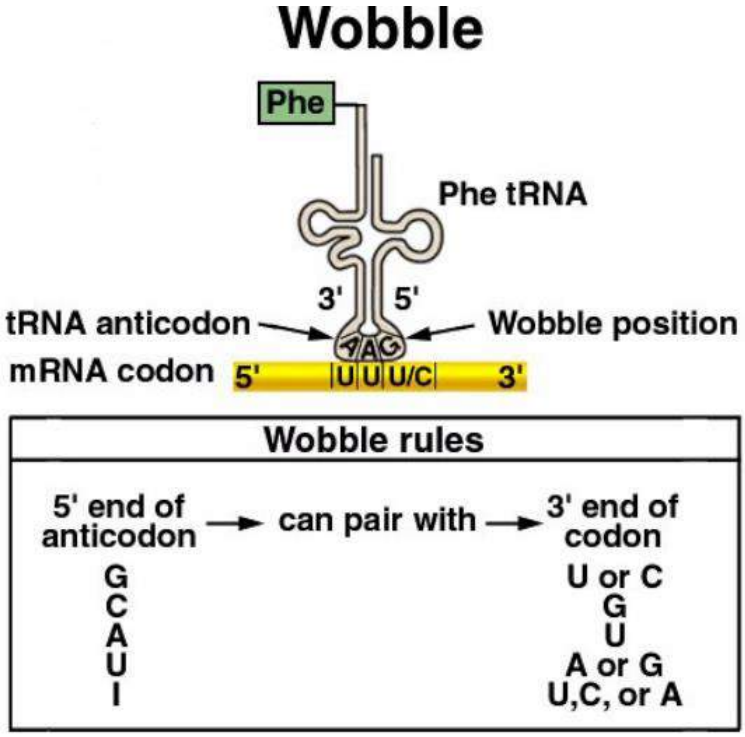
Multiple codons for a single amino acid

The number of codons for a single amino acid varies, ranging from one (tryptophan = UGG) to as many as six (serine = UCU or UCC or UCA or UCG or AGU or AGC). Why? The answer can be divided into two parts:

1. Certain amino acids can be brought to the ribosome by several alternative tRNA types having different anticodons, whereas certain other amino acids are brought to the ribosome by only one tRNA.
2. Certain tRNA species can bring their specific amino acids in response to several codons, not just one, through a loose kind of base pairing at one end of the codon and anticodon. This sloppy pairing is called **wobble.**

Wobble is caused by the third nucleotide of an anticodon (at the 5′ end) that is not quite aligned. This out-of-line nucleotide can sometimes form hydrogen bonds not only with its normal complementary nucleotide in the third position of the codon, but also with a different nucleotide in that position (see the figures below).





I (inosine) is one of the rare bases found in tRNA, often in the anticodon.

Figure: Wobble pairing

Translation:

The translation process is the rendering of the nitrogen bases sequence in the mRNA into the sequence of amino acids in the protein.

Components of Translation

• **mRNA:**

– Eukaryotes: made in the nucleus, transported to the cytoplasm.

– Prokaryotes: transcription and translation occur concurrently.

• **tRNA:** Adaptor molecules that mediate the transfer of information from nucleic acids to protein.

• **Ribosomes**: manufacturing units of a cell; located in the cytoplasm. Contain ribosomal RNA (rRNA) and proteins.
Ribosomes consist of two subunits: large and small.
– Prokaryotes: 50S + 30S = 70S
– eukaryotes: 60S + 40S = 80S.

• **Enzymes:** required for the attachment of amino acids to the correct tRNA molecule (Aminoacyl-tRNA synthetases) and for peptide bond formation between amino acids (Peptidyl Transferase).

• **Proteins:** soluble factors necessary for proper initiation, elongation and termination of translation.

Translation includes:

1- The formation of an initiation complex.

2- The activation of an amino acid and its bond on the tRNA.

3- The lengthening of the peptide chain.

4- The termination of the synthesis of the polypeptide.

The formation of the initiation complex

In prokaryotes, messenger RNA and formylmethionyl-tRNA must be brought to the ribosome for protein synthesis to begin. How is this accomplished?

Three protein **initiation factors** (IF1, IF2, and IF3) are essential. The 30S ribosomal subunit first forms a complex with IF1 and IF3. The binding of these factors to the 30S subunit prevents it from prematurely joining the 50S subunit to form a 70S complex, devoid of mRNA and fMet-tRNA. The small ribosomal subunit identifies start codons through interaction of the small (16S) rRNA with the Shine-Dalgarno sequence . This sequence, located near the AUG start codon, base-pairs to a sequence at or very near the 3′ end of 16S rRNA, thereby binding the mRNA and small ribosomal subunit to each other. Initiation factor 2 binds GTP, and the concomitant conformational change enables IF_2 to associate with formylmethionyl-tRNA. The IF2-GTP-initiator tRNA complex binds with mRNA and the 30S subunit to form the 30S initiation complex. After the hydrolysis of GTP bound to IF2, all three initiation factors release and the large (50S) ribosomal subunit binds, forming the 70S **initiation complex**. When the 70S initiation complex has been formed, the

ribosome is ready for the elongation phase of protein synthesis. The fMet-tRNA molecule occupies the P site on the ribosome. The other two sites for tRNA molecules, the A site and the E site, are empty. Formylmethionyl-tRNA is positioned so that its anticodon pairs with the initiating AUG codon on mRNA.

The assembled structure of the ribosome creates three pockets for the binding of two molecules of tRNA:

- The far left pocket is the Exit site or E site: it binds the deacylated tRNA (no amino acid attached)
- The one in the middle is known as the peptidyl or the P site: it binds to the tRNA holding the growing chain of polypeptide.
- The site on the right is termed the amino acyl, or the A site: it binds to the incoming tRNA molecule.

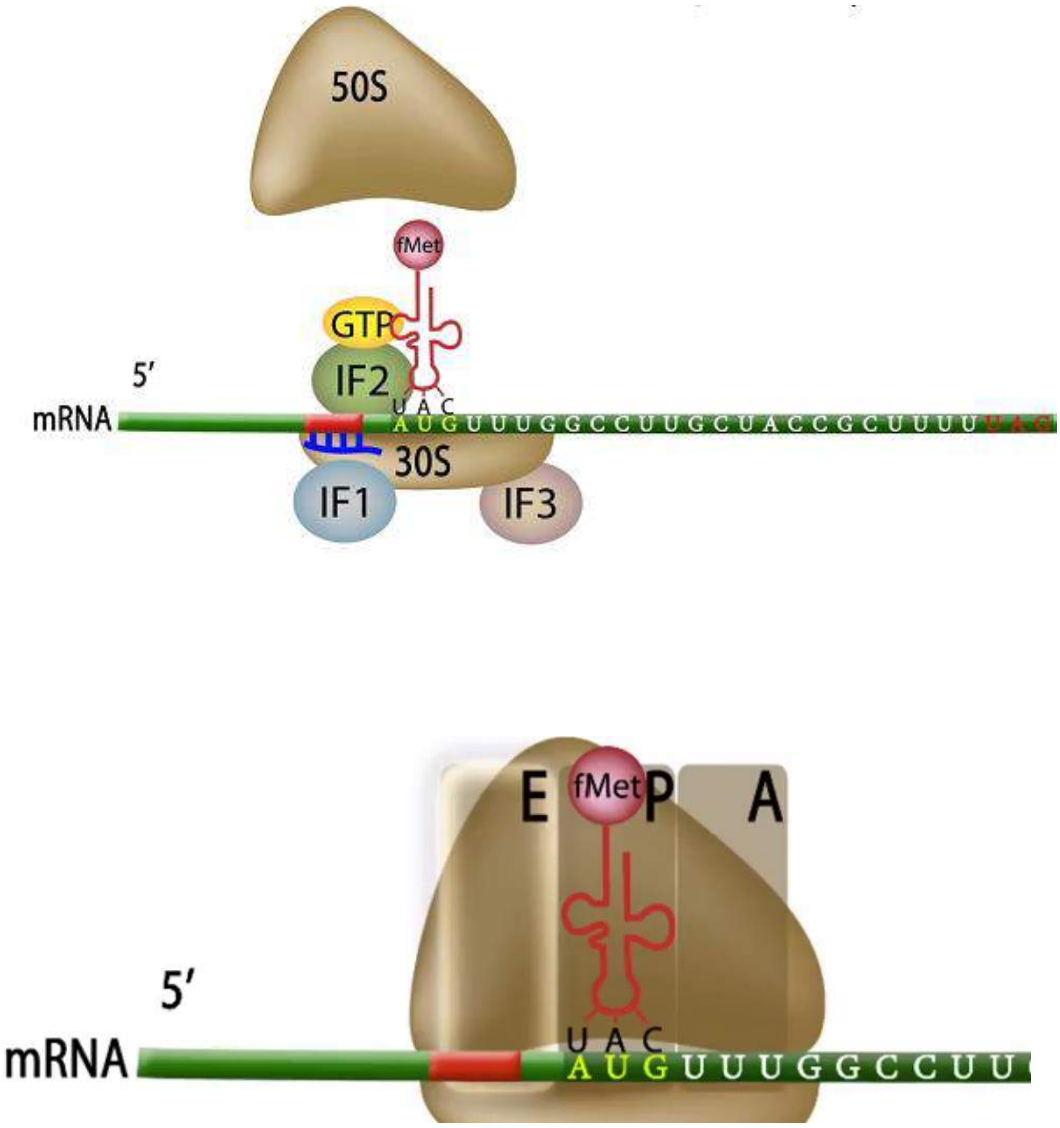


Figure: The formation of the initiation complex in prokaryotes.

In eukaryotes, translation of mRNA into protein begins after assembly of initiator tRNA (Met-tRNA), mRNA, and separated 40S and 60S ribosomal subunits into an 80S ribosome in which Met-tRNA is positioned in the ribosomal P site at the initiation codon. This process is mediated by eukaryotic initiation factors.

In Eukaryotes, the situation is a little different from the prokaryotes. Here, the small ribosomal unit doesn't attaches directly to certain sequences in the mRNA. Instead the small ribosomal unit binds first to the methylated cap (7-methyl guanosine) at the 5’ end of the mRNA. Then it migrates to the initiation site, usually the first AUG it encounters as it scans the mRA in the 5’ to 3’ direction.

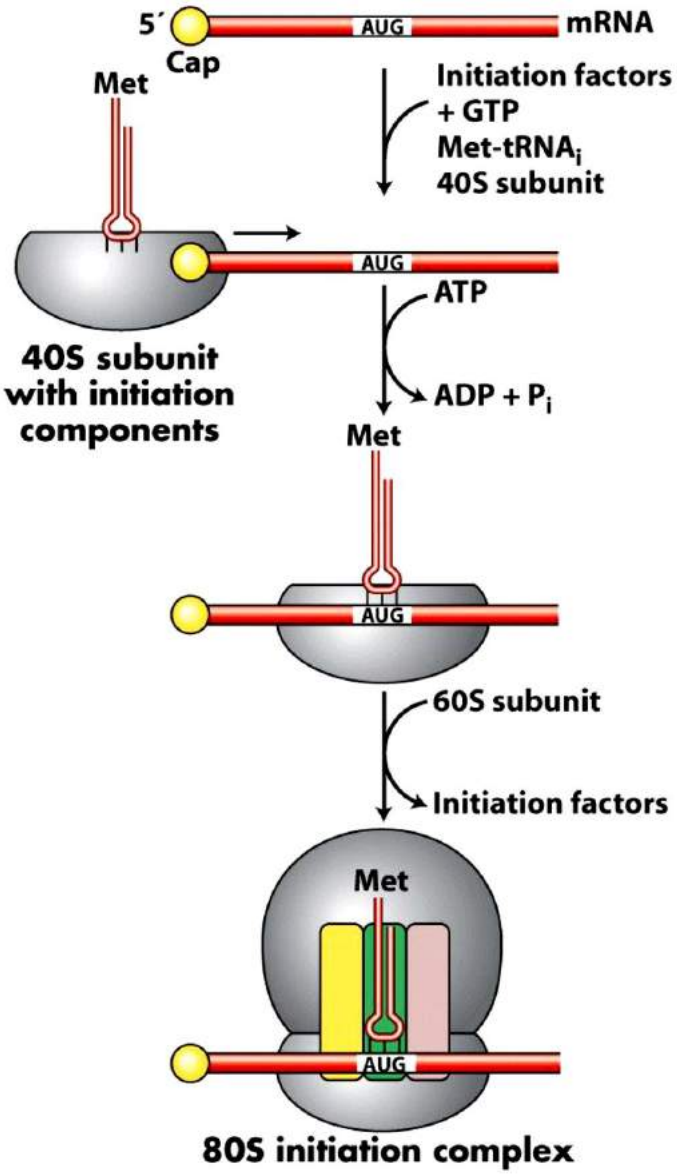


Figure: Translation Initiation in Eukaryotes.

Activation of the amino acid and its binding to the tRNA

The tRNA has an important role in reading of the code. To the adenine on its 3´ OH acceptor arm an amino acid is attached. Amino acids, which are supposed to be build-in into the protein molecule, are not able to read the information in the mRNA – by themselves. This is why the amino acid has to connect to the tRNA, which transfers it to a ribosome. This connection is carried out by **aminoacyl tRNA synthetase.**

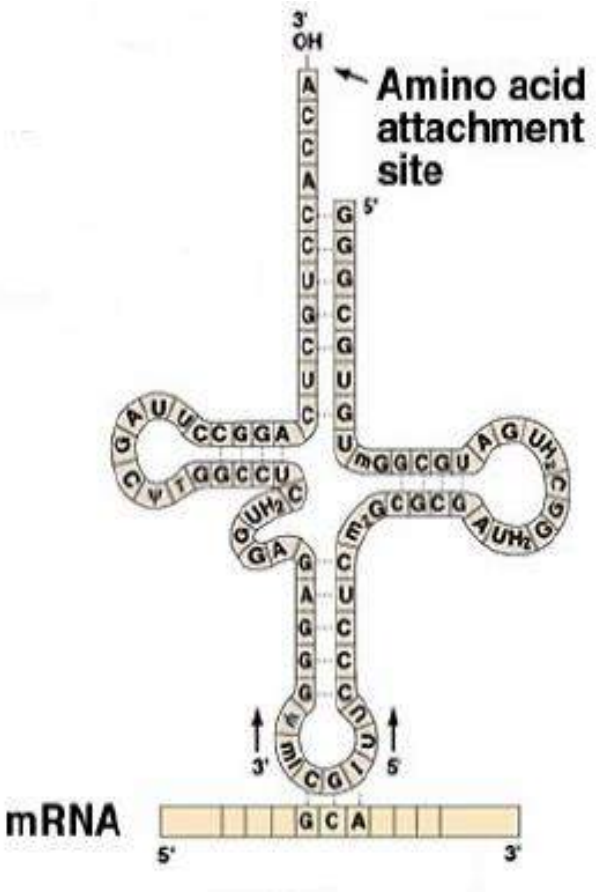


Figure: Aminoacyl tRNA

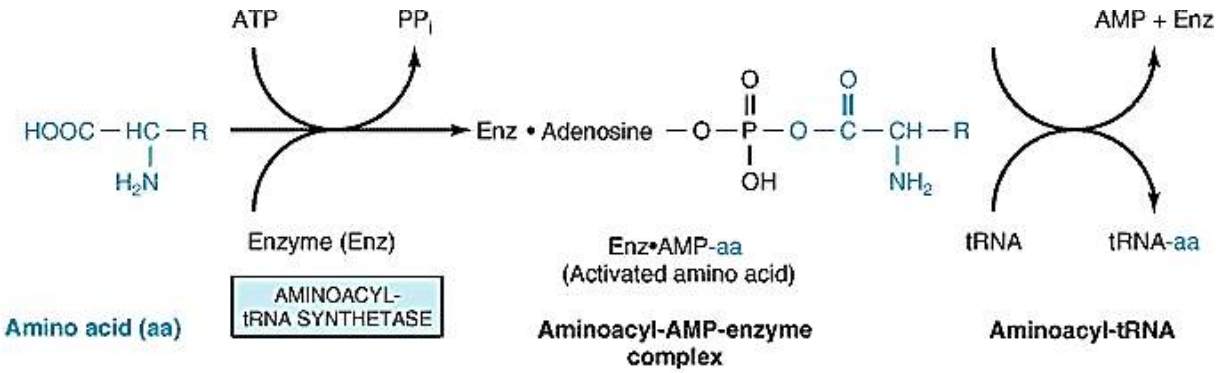


Figure: Activation of amino acids

- The activation of amino acids takes place in cytosol.
- The activation of amino acids is catalyzed by their aminoacyl tRNA synthetases.
- All the 20 amino acids are activated and bound to 3’ end of their specific tRNA in the presence of ATP and Mg++.
- The N-formylated methionine is chain initiating amino acid in bacteria whereas methionine is chain initiating amino acid in eukaryotes.
- Methionine is activated by methionyl-tRNA synthetase.
- Similarly, all 2o amino acids are activated (amino acyl-AMP enzyme complex) and then bound to their specific tRNA forming Aminoacyl tRNA.

Elongation

Elongation is carried out with the elongation factors. The three elongation factors each have their own role in translation.

EF1α (EF-Tu)

EF1α deposits the aa-tRNA on the A-site of the ribosome by forming a ternary complex with GTP and the aa-tRNA.

EF1β (EF-Ts)

EF1β is required to regenerate active EF1α. Once EF1α-GDP has detached from the ribosome, EF1β acts as a catalyst to convert it into active EF1α-GTP. The EF1α is then ready to interact with a new aa-tRNA, to begin the cycle again.

EF2 (EF-G)

EF2 assists in the translocation of tRNA and mRNA (base-paired together) through the ribosome, acting to move the mRNA-bound tRNA from the A-site to the P-site, thereby freeing the A-site ready for the next aa-tRNA to bind. At the same time, it moves the deacylated-tRNA from the P-site to the E-site, where it exits the ribosome.

NOTE: EF1α, EF1β and EF2 are elongation factors in eukaryotes. However, EF-Tu, EF-Ts and EF-G are in prokaryotes.

Elongation steps;

i. Binding of AA-tRNA at A-site:

- The 2nd tRNA carrying next aminoacid comes into A-site and recognizes the codon on mRNA. This binding is facilitated by EF-TU and utilizes GTP.
- After binding, GTP is hydrolysed and EF-TU-GDP is releasd
- EF-TU-GDP then and enter into EF-TS cycle.

ii. Peptide bond formation:

- The amino acid present in t-RNA of P-site ie Fmet is transferred to t-RNA of A-site forming peptide bond. This reaction is catalyzed by peptidyl transferase.
- Now, the t-RNA at P-site become uncharged

iii. Ribosome translocation:

- After peptide bond formation ribosome moves one codon ahead along 5'-3' direction on mRNA, so that dipeptide-tRNA appears on P-site and next codon appears on A-site.
- The uncharged tRNA exit from ribosome and enter to cytosol.

- The ribosomal translocation requires EF-G-GTP (translocase enzyme).
- The codon on A-site is now recognized by other aminoacyl-tRNA as in previous.
- The dipeptide on P-site is transferred to A-site forming tripeptide.
- This process continues giving long polypeptide chain of amino acids.

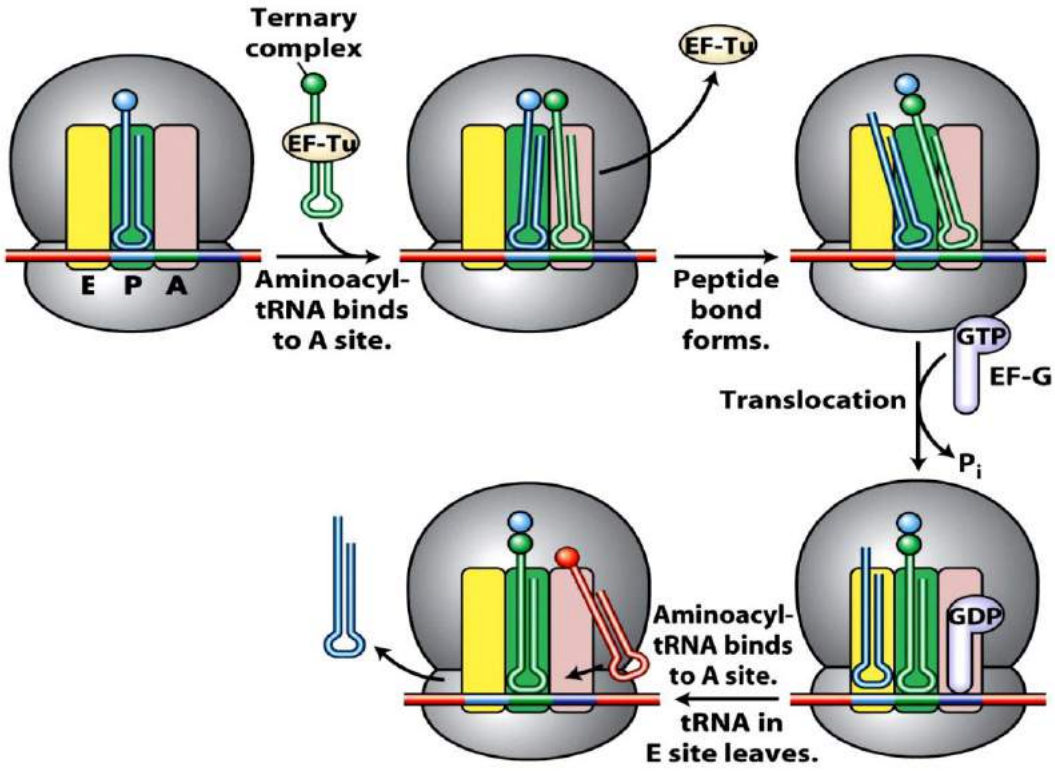


Figure: Elongation in prokaryotes

Termination:

- The peptide bond formation and elongation of polypeptide continues until stop codon appear on A-site.
- If stop codon appear on A-site it is not recognized by t-RNA carrying amin oacids because stop codon do not have anticodon on mRNA.
- The stop codon are recognized by next protein called release factor (Rf-1, RF-2 and RF-3) which hydrolyses and cause release of all component ie 30s, 50S, mRNA and polypeptide separates.
- RF-1 recognizes UAA and UAG while RF-2 recognizes UAA and UGA while RF-3 dissociate 30S and 50S subunits.
- In case of eukaryotes only one release actor eRF causes dissociation.

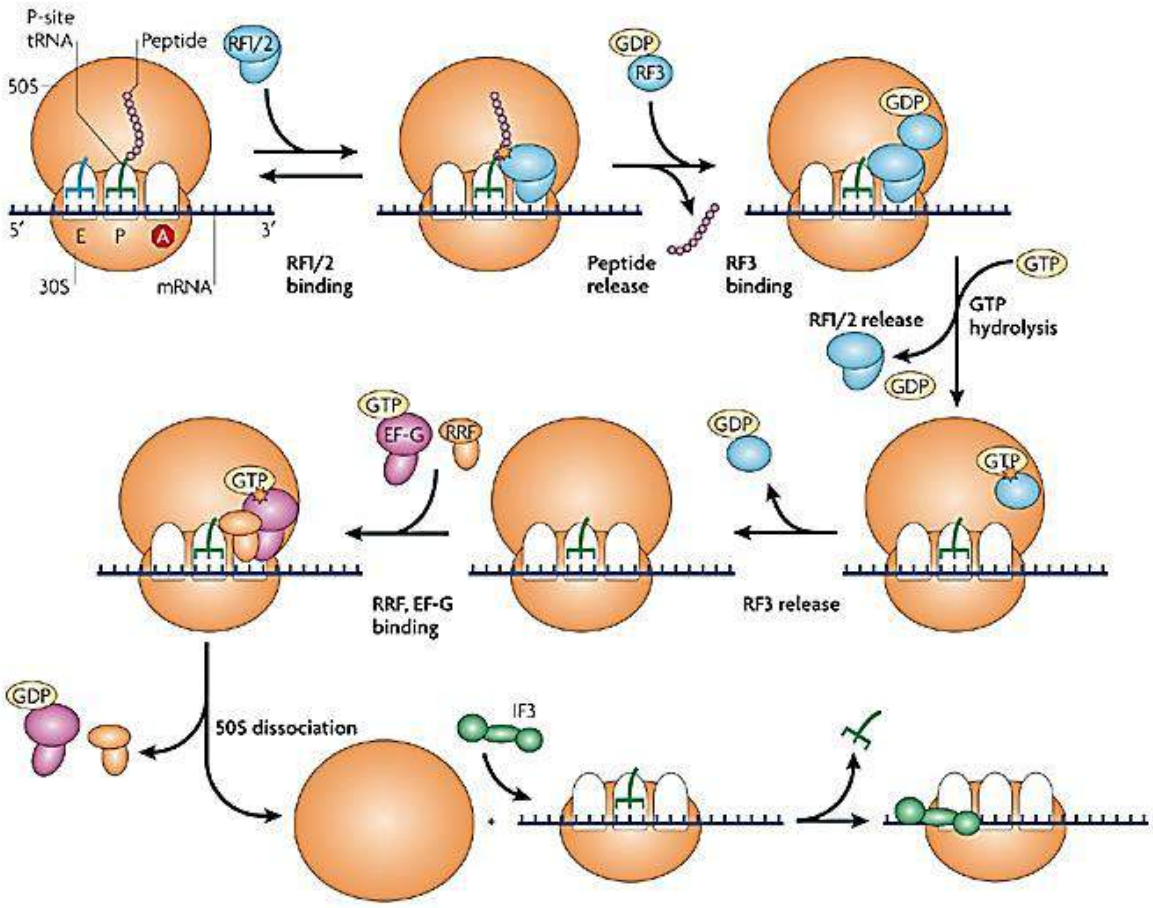


Figure: Termination in prokaryotes.

Post translation modification:

The newly formed polypeptide may not be biology functional so it undergoes several folding and processing known as post translation modification.

1. Amino terminal and carboxyl terminal modification:

- The N-formylmethionine in case of bacteria is removed from polypeptide chain and some carboxyl terminal are also removed by enzymatic action to make functional protein. In case of eukaryotic protein, amino terminal is N- acetylated.

2. Loss of signal sequences:

- In some protein the amino terminal end is cleaved by specific peptidase so that protein loss its signaling property.

3. Modification of individual aminoacids:

- The aminoacids may be phosphorylated, acetylated for modification

4. Attachment of carbohydrate side chain:

- Carbohydrate side chain is added to make protein functional. Eg, glycoprotein. Lipoprotein

5. Addition of isoprenyl group:

- In some protein, isoprenyl group is added so to make protein active.