

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
FACULTE DES SCIENCES DE LA NATURE ET DE LA VIE



Département des Sciences Agronomiques
et Biotechnologies

Biologie Moléculaire et Génie Génétique

Niveau : Licence 3 en Biotechnologie
Microbienne et végétale

Dr. Nehal Fatima

Année : 2017-2018

Module : Biologie Moléculaire et Génie génétique

- Introduction

Partie I : Biologie moléculaire

1. L'ADN, Support de l'information génétique
2. Structure des acides nucléiques
3. Propriétés physicochimiques de l'ADN
 - 3.1 La solubilité
 - 3.2 La densité
 - 3.3 La charge
 - 3.4 Propriétés spectrales
4. La Réplication de L'ADN
5. Expression de l'information génétique
 - 5.1 La transcription
 - 5.1.1 La transcription chez les procaryotes
 - 5.1.2 La transcription chez les eucaryotes
 - 5.1.3 Maturation des ARNm
 - 5.2 La traduction
 - 5.2.1 Le code génétique
 - 5.2.2 Les ribosomes
 - 5.2.3 Rôle des ARNt dans la traduction
 - 5.2.4 La traduction chez les procaryotes
 - 5.2.5 La traduction chez les eucaryotes
 - 5.2.6 Modifications post-traductionnelles
 - 5.3 Régulation de l'expression génétique
 - 5.3.1 Régulation de la transcription
 - a. Contrôle négatif de la transcription (Répression et induction)
 - b. Contrôle positif de la transcription
 - c. Contrôle de l'expression génétique par atténuation
 - 5.3.2 Régulation traductionnelle
6. Techniques de base de biologie moléculaire
 - 6.1 Extraction et purification des acides nucléiques

- 6.2 Dosage des acides nucléiques
- 6.3 Electrophorèse
- 6.4 Marquage des acides nucléiques
 - 6.4.1 Marquage des sondes
 - a. La nick translation
 - b. Le multiamorçage au hasard ou multirandom priming
 - c. Marquage d'oligonucleotides de synthèse
- 6.5 Hybridation des acides nucléiques
 - 6.5.1 Types d'hybridation
 - a. Hybridation d'ADN sur support solide : Southern Blot
 - b. Hybridation in situ
 - c. Hybridation sur chromosome
 - d. Hybridation sur colonies de bactéries, sur plage de lyse
 - e. Polymorphisme de Longueur des Fragments de Restriction de l'ADN (RFLP).
- 6.6 Réaction de polymérisation en chaînes (PCR)
- 6.7 Séquençage
 - 6.7.1 Méthode de Sanger
 - 6.7.2 Technique chimique de Maxam et Gilbert

Partie II : Génie génétique

- 1. Outils enzymatique du génie génétique
 - 1.1 Les nucléases
 - a. Les endonucléases
 - 1.2 Les polymérases
 - a. ADN polymérase I et fragment de klenow
 - b. ADN Taq polymérase et autres ADN polymérases thermostables
 - c. Transcriptase reverse ou inverse ou retrotranscriptase RT
 - d. ARN polymérase
 - 1.3 Ligases
 - 1.4 Phosphatases
 - 1.5 Kinases
- 2. Vecteurs de clonage et analyse des banques
 - 2.1 Vecteurs de clonage
 - 2.1.1 Propriétés des vecteurs de clonage

- 2.1.2 Différents vecteurs de clonage
 - a. Les plasmides
 - b. Les phages
 - c. Les cosmides
 - d. Les chromosomes artificiels bactériens
 - e. Les chromosomes artificiels de levures
 - f. Vecteurs navettes et d'expression
- 2.2 Types de banques et source d'ADN
 - 2.2.1 Banques ADN
 - 2.2.2 Banques d'ADNc
- 3. Criblage des banques génomiques
 - 3.1 Criblage par antibiotique
 - 3.2 La méthode de sélection blanc-bleue
 - 3.3 Criblage par hybridation
 - 3.4 Criblage par complémentarité
 - 3.5 La méthode de coupure avec des enzymes de restriction
- 4. Production d'insuline par génie génétique.

Références bibliographiques

Introduction

La biologie moléculaire parfois abrégé Bio mol ou BM est une discipline scientifique au croisement de la génétique, de la biochimie et de la physique dont l'objet est la compréhension des mécanismes de fonctionnement de la cellule au niveau moléculaire qui permettent la conservation et la perpétuation de la structure vivante au niveau du génotype. Le terme Bio mol utilisé pour la première fois en 1938 par Warren Weaver, désigne également l'ensemble des techniques de manipulation d'acides nucléiques (ADN, ARN), appelée aussi techniques de génie génétique. Pour cela, les techniques d'étude et de modification des gènes et de leur expression font partie intégrante de la biologie moléculaire. La biologie moléculaire est une discipline scientifique qui décrit la manière dont l'information génétique est **conservée, transmise et exprimée**.

Le génie génétique est un «ensemble de techniques permettant d'identifier et d'isoler, de modifier et de transférer de façon contrôlée du matériel génétique». Il s'agit donc d'un outil aux applications extrêmement variées et qui permet en particulier d'intervenir avec précision sur le patrimoine génétique des êtres vivants.

Pratiquement, cette méthodologie permet d'identifier un gène spécifique parmi les nombreux gènes d'un organisme. D'abord amplifié afin qu'il soit plus facile d'accès, le gène peut ensuite être découpé et isolé des autres molécules d'ADN. Il peut enfin être réinséré dans une molécule d'ADN d'origine différente, permettant ainsi de transférer de l'information génétique d'une cellule vers une autre. Le résultat obtenu est un ADN recombiné. Le génie génétique permet également d'apporter des modifications à des gènes (mutagenèse dirigée) qui, par conséquent, produiront des protéines modifiées.

Partie I : Biologie moléculaire

1. L'ADN, Support de l'information génétique

L'information génétique est transmise sous deux formes d'une génération à l'autre:

- Soit sous la forme d'un œuf fécondé (reproduction sexuée) qui reçoit un exemplaire de chaque gène parental.

- Soit sous la forme d'une cellule fille (reproduction asexuée) qui reproduit à l'identique la cellule-mère.

Le travail pionnier de Fred Griffith, en 1928 sur le transfert de la virulence du pathogène *Streptococcus pneumoniae*, communément appelé pneumocoque est l'expérience qui prouva que l'ADN était bien le matériel génétique : Griffith découvrit que s'il faisait bouillir des bactéries virulentes et les injectait à des souris, celles-ci n'étaient pas infectées et qu'on ne pouvait retrouver aucun pneumocoque chez les animaux. Quand il injectait un mélange de bactéries virulentes tuées et de bactéries non virulentes tuées et de bactéries non virulentes vivantes, les souris mouraient. De plus, on pouvait isoler des bactéries virulentes de ces souris mortes. Griffith donna à ce changement de bactéries non virulentes en pathogène virulents, le nom de **transformation** (Figure 1).

- Ceci suggère donc qu'il existe chez les bactéries un "facteur ou principe transformant",
- Probablement résistant et libéré par la chaleur,
- Susceptible d'être intégré par d'autres bactéries comme les bactéries R
- Et qui leur confère de façon héréditaire de nouvelles propriétés génétiques (comme la virulence).
- Nature de ce matériel?

Oswald Avery et ses collègues déterminèrent par la suite quel constituant des pneumocoques tués par la chaleur était responsable de la transformation de Griffith. Ces chercheurs détruisirent sélectivement les constituants d'extraits purifiés de Pneumocoques virulents (cellules S), au moyen d'enzymes hydrolysant l'ADN, l'ARN ou les protéines. Ils exposèrent par la suite des souches non virulentes (Souches R) aux extraits traités. La transformation des bactéries non virulentes n'avait plus lieu que si l'ADN avait été hydrolysé, ce qui suggérait que l'ADN portait l'information requise pour la transformation (Figure 2).

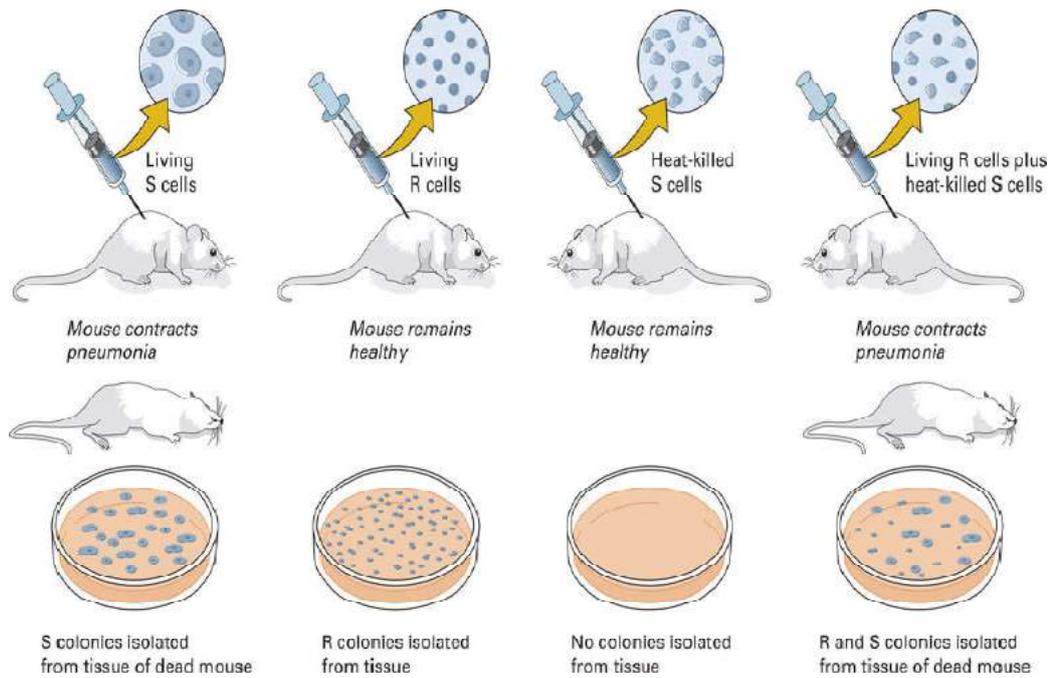


Figure 1: Expérience de Griffith (1928).

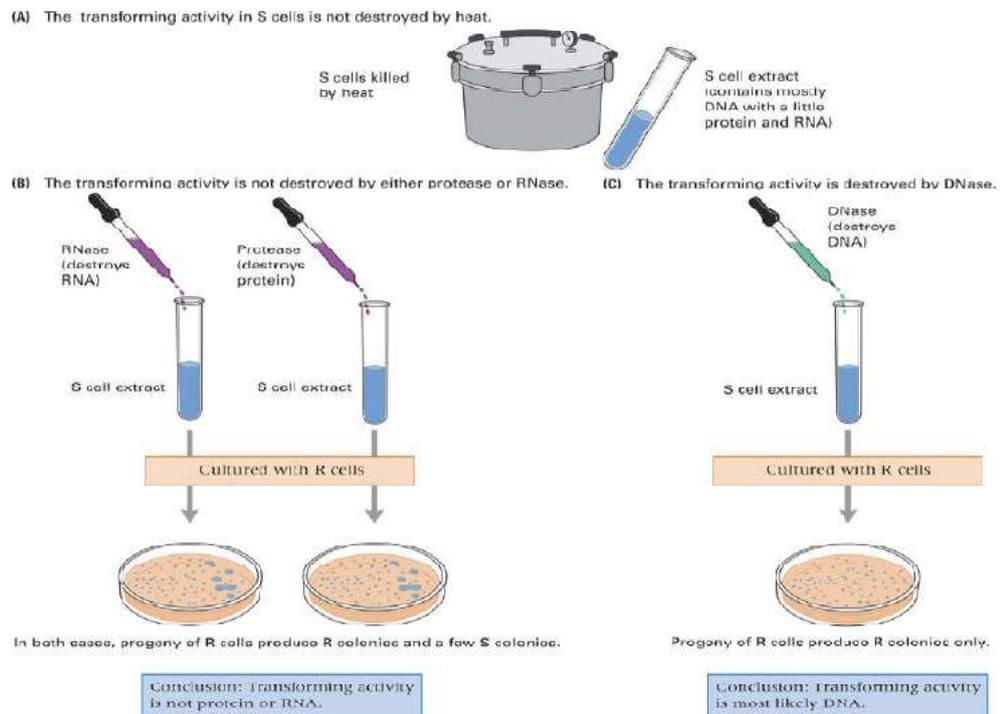


Figure 2: Expérience de Avery McLeod et McCarty (1944).

Les expériences d'Hershey et Chase (1952) réalisées indiquant que l'ADN était le matériel génétique d'un virus bactérien. D'Hershey et Chase marquèrent l'ADN du virus à l'aide du ^{32}P et les protéines de la capside à l'aide de ^{35}S . Ils mélangèrent le bactériophage radioactif avec *E. coli* et incubèrent le mélange pendant quelques minutes. La suspension fut alors agitée violemment dans un mixeur pour détacher toutes les particules du phage adsorbées. Après centrifugation, la radioactivité dans le surnageant (où se trouvaient les virus) et dans le culot fut mesurée. On retrouva les protéines radioactives en grande partie dans le surnageant, tandis que l'ADN marqué au ^{32}P était resté dans les bactéries. Puisqu'il y avait injection de matériel génétique et production de particules T2, l'ADN était forcément la molécule porteuse de l'information génétique de T2.

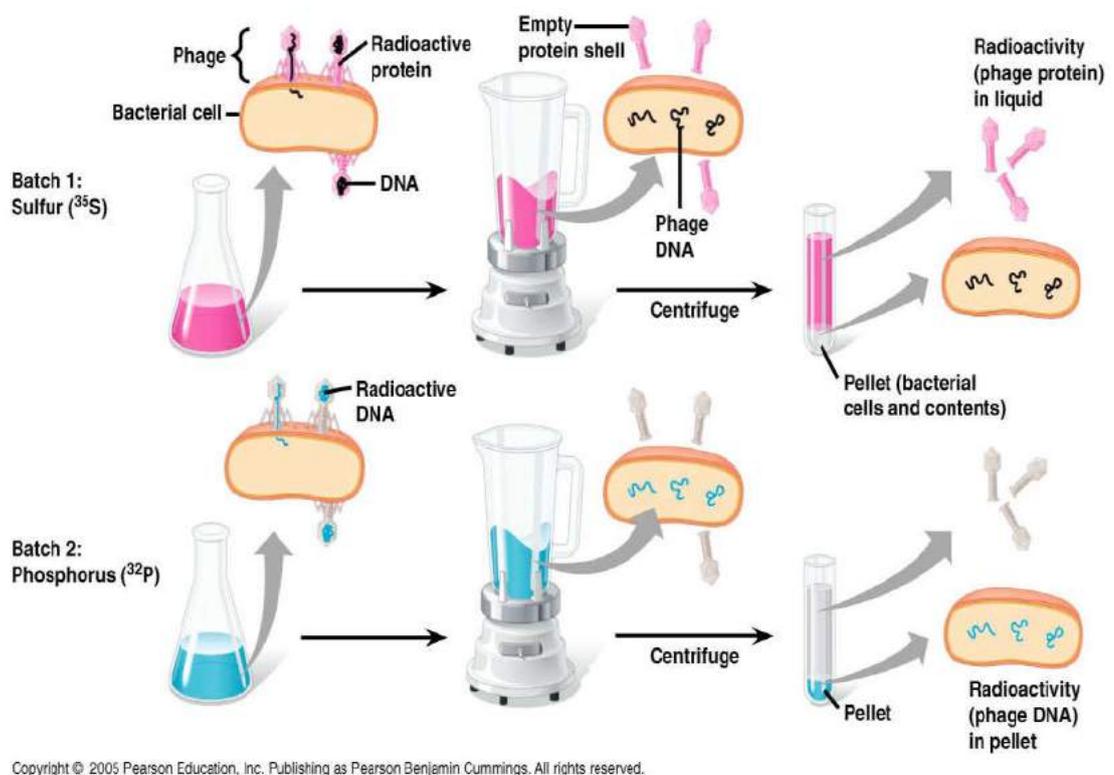


Figure 3: Expérience d'Hershey et Chase (1952).

Les bactéries infectées par des virus aux protéines marquées (^{32}P) ne deviennent pas radioactives mais celles infectées par des virus à l'ADN marqué (^{35}S) le deviennent.

2. Structure des acides nucléiques

Les travaux de Kornberg révélèrent que les précurseurs de l'ADN étaient des nucléotides riches en énergie (dNTP, dCTP, dTTP, dGTP). Chaque nucléotide est composé d'acide phosphoriques (3), base azotée (A, T, C, G) et un pentose (2'-désoxy-β-D-Ribose). La liaison formée entre deux phosphates donne un anhydride avec une liaison très riche en énergie, et la liaison entre le groupement acide du phosphate et l'hydroxyle du sucre donne une liaison ester. La liaison formée entre le sucre et la base est une liaison osidique de type β-N-glycosidique.

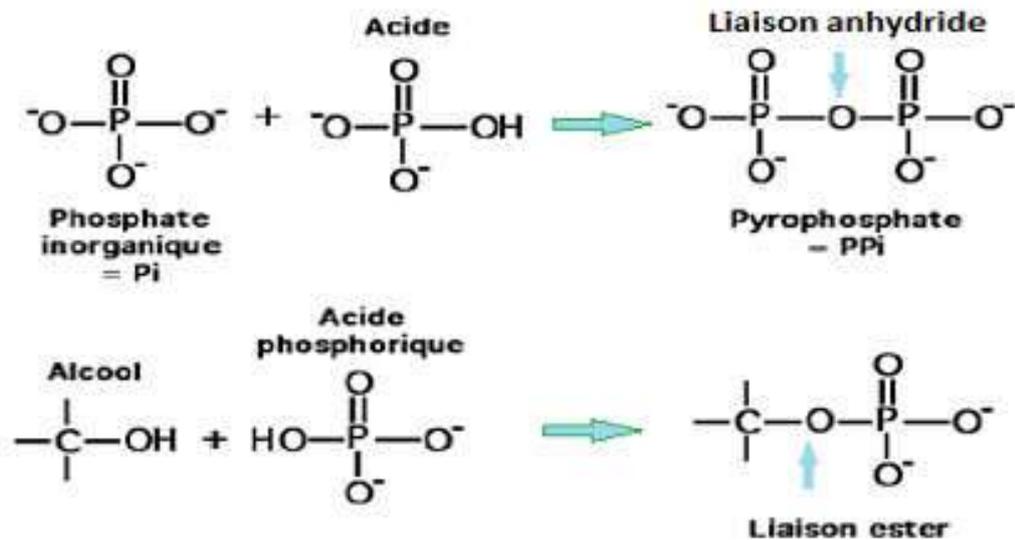


Figure 4: Formes du phosphate et les liaisons formées.

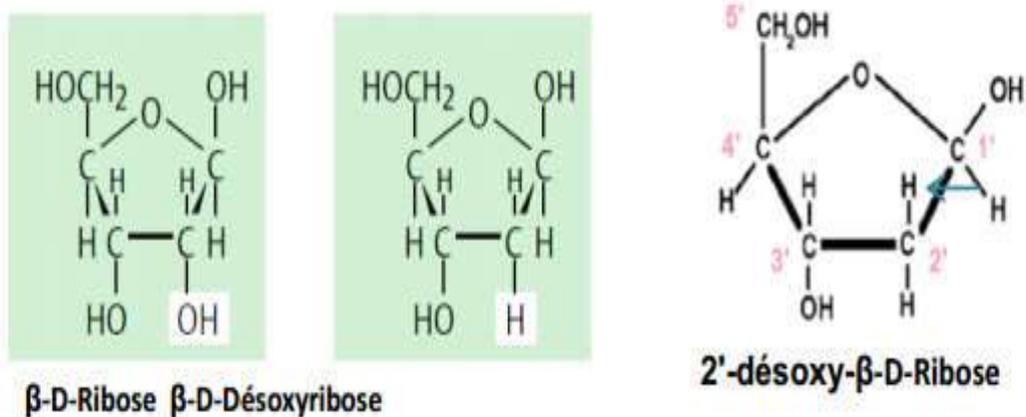


Figure 5 : A- Les pentoses qui rentrent dans la composition des acides nucléiques (ADN, ARN). B- Numérotation d'un sucre, le prime (') est ajouté dans la numérotation pour la distinguer à celle des bases.

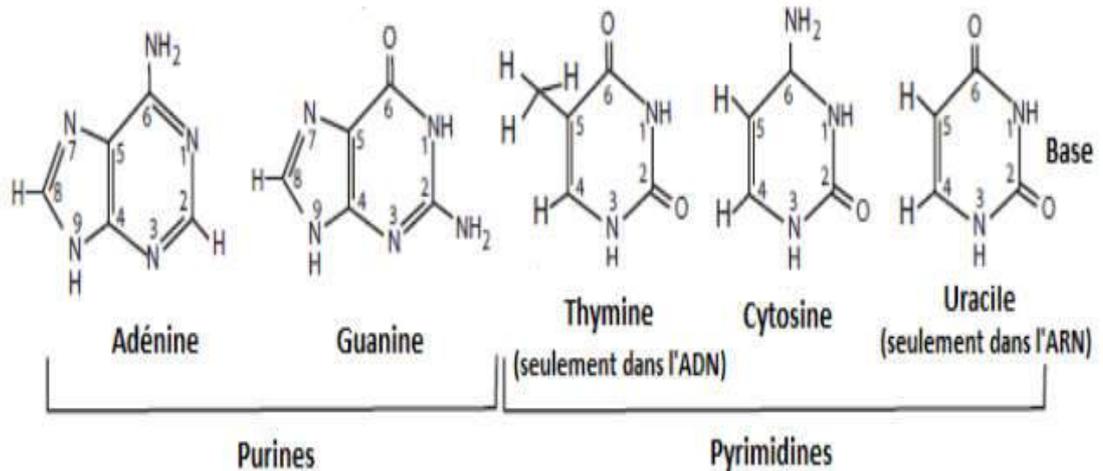


Figure 6 : Les cinq bases azotées de l'ADN et l'ARN (Purines [A et G], Pyrimidines [C, T et U]).

Il existe par ailleurs des **bases puriques modifiées** :

- Dans l'ADN, on trouve par exemple, la **N⁶-méthyladénine**
- Dans les ARN (en particulier les ARN_t), on trouve par exemple, la **N⁷-méthylguanine** et l'**hypoxanthine**.

Et des **bases pyrimidiques modifiées** :

- Dans l'ADN, la plus commune est la **N⁵-méthylcétosine**.
- Dans les ARN, on trouve, entre autre bases peu communes, la **pseudouracile** et la **dihydrouracile**.

Les acides désoxyribonucléiques sont de très grandes molécules composées de deux chaînes polynucléotidiques enroulées l'une autour de l'autre pour former une double hélice régulière (La forme B a un diamètre de 2,47 nm).

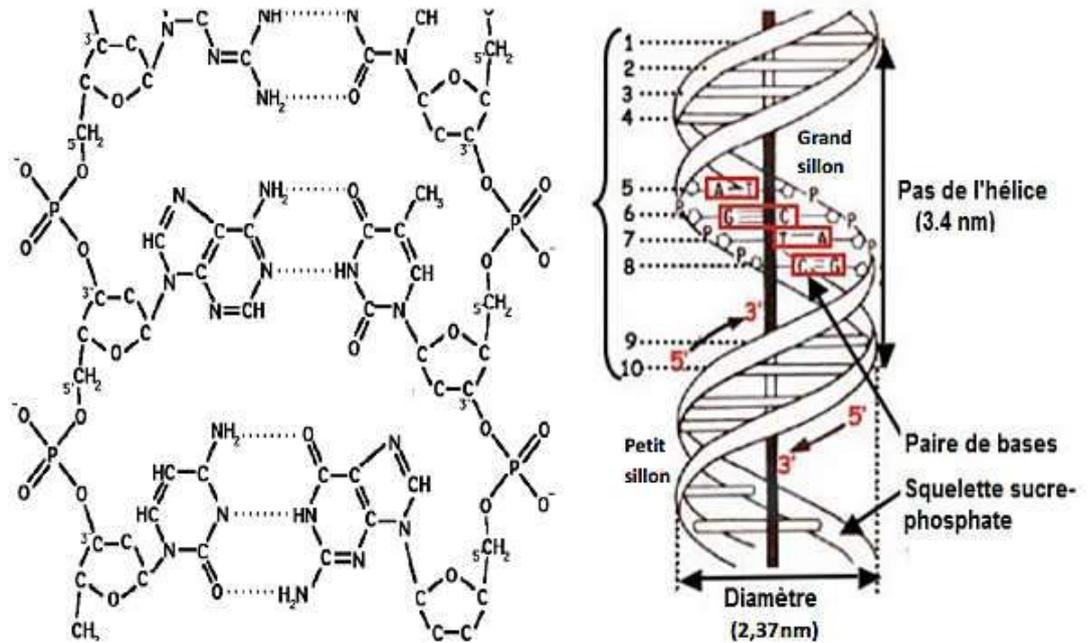


Figure 7: Structure détaillée de la double hélice de l'ADN (forme B). Un tour d'hélice recouvre environ 10.5 pb (34 Å). Le squelette de chaque brin est formé de sucre et de phosphate. Les bases azotées sont projetées vers l'intérieur de la molécule mais restent accessibles au solvant par les deux sillons (petit et grand sillon).

La stabilisation de la double hélice s'effectue par des liaisons chimiques faibles et non covalentes à la fois internes et externes:

- 1- Liaisons hydrogène entre les bases.
- 2- Interactions hydrophobes et électroniques entre les bases
- 3- Interactions des groupements sucre et phosphate avec le milieu aqueux (interactions hydrophiles).

Cette stabilité n'entraîne pas une rigidité de la molécule d'ADN et celle-ci peut adopter des conformations différentes selon les régions. Plusieurs conformations correspondant à des sens d'enroulement différents ou des pas différents ont été trouvées, les principales étant :

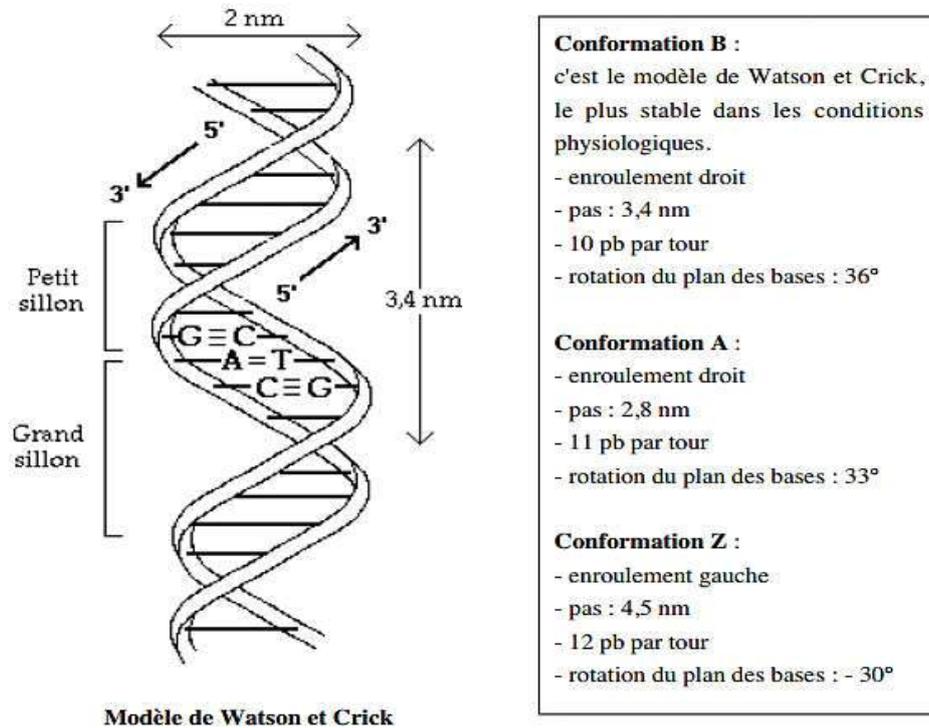


Figure 8 : Structure détaillée de la double hélice de l'ADN (forme B). Un tour d'hélice recouvre environ 10.5 pb (34 Å). Le squelette de chaque brin est formé de sucre et de phosphate. Les bases azotées sont projetées vers l'intérieur de la molécule mais restent accessibles au solvant par les deux sillons (petit et grand sillon).

3. Propriétés physico-chimiques de l'ADN

Les liaisons hydrogène et les interactions hydrophobes qui maintiennent la structure en double hélice sont des forces faibles et des quantités relativement petites d'énergie peuvent séparer les deux brins, un processus appelé dénaturation.

3.1 La Solubilité

L'ADN est un polyanion dont les sels de sodium sont solubles dans l'eau en formant des solutions à viscosité élevée. Les alcools et en particulier l'éthanol, précipitent les molécules d'ADN sous forme agglomérats en longues fibres.

3.2 La densité

La densité des molécules d'ADN est telle qu'on peut les séparer par ultracentrifugation dans des gradients de densité (chlorure de césium).

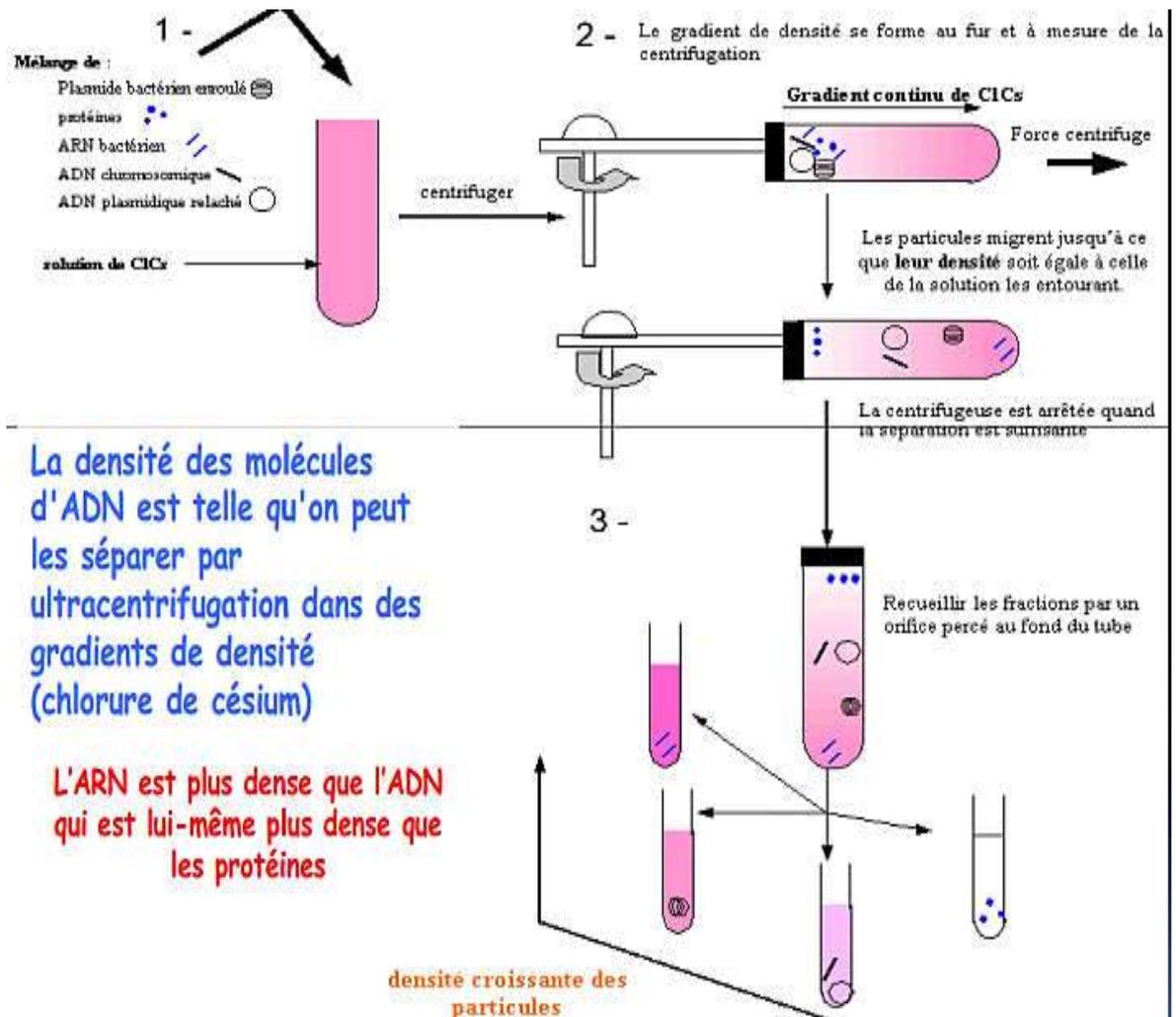


Figure 9 : Détermination de la densité de l'ADN par ultracentrifugation dans un gradient de densité.

3.3 La charge

La charge de ces molécules à pH physiologique est négative et directement proportionnelle à leur longueur (nb de nucléotides). Charge dont la contribution est uniquement due au groupement phosphates (à ce pH les bases ne portent aucune charge). Cette propriété est utilisée pour les séparer par électrophorèse.

3.4 Propriétés spectrales

- Le spectre d'absorption de l'ADN natif n'est pas identique à celui du même ADN dénaturé par la chaleur (chauffage à 100°C) ou par l'urée ou encore à pH très alcalin. L'ADN dénaturé a une absorption à 260 nm plus élevée que l'ADN natif, d'un facteur 1,6. Cette propriété est appelée *l'effet hyperchrome* ou *hyperchromicité*.

- L'absorption de l'ADN natif, à 260 nm, en fonction de la température (courbe de fusion), présente l'allure d'une sigmoïde : le point d'inflexion de cette courbe, qui correspond à la demi-variation d'absorbance, est la température de fusion de la molécule, notée T_m .

- Lors d'un refroidissement lent, l'absorption suit la courbe de fusion en sens inverse. Lors d'un refroidissement rapide, l'absorption ne suit pas la courbe de fusion en sens inverse mais une autre courbe qui n'aboutit à la même valeur originale de l'absorption, mais à une valeur plus élevée : c'est le phénomène d'*hystérésis*.

- Il faut noter que la température de fusion T_m est dépendante de la force ionique du milieu et qu'elle diminue lorsque cette dernière augmente (dans des milieux où $[NaCl] > 1M$). - La valeur de la température de fusion T_m est d'autant plus élevée que le pourcentage de bases G + C est grand.

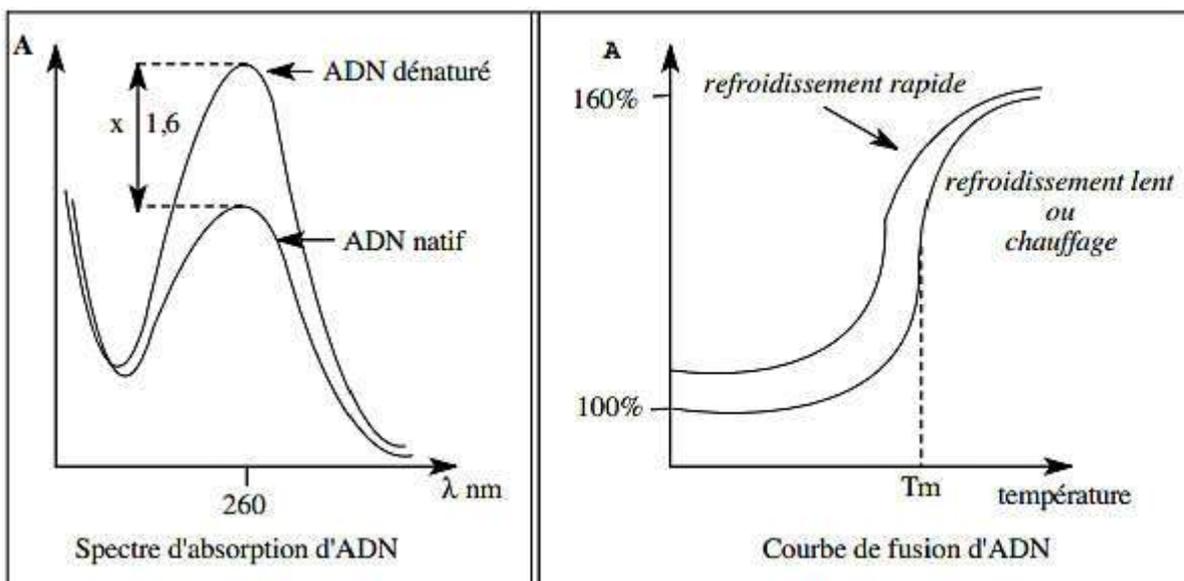


Figure 10 : Propriétés spectrales de l'ADN natif et dénaturé.

4. Réplication de l'ADN

La réplication de l'ADN est un mécanisme très important elle permet **la duplication** du patrimoine génétique pour obtenir 2 exemplaires, chacun transmis à une des deux cellules filles lors de la division cellulaire. La réplication a donc lieu **avant la division cellulaire**.

Ce mécanisme doit être le plus **fidèle** possible, pour que le patrimoine génétique qui va être transmis aux cellules filles soit le plus parfaitement identique à celui de la cellule mère. En effet, la survie à court terme de la cellule et de l'organisme dépend de la prévention de la modification de son ADN, puisque les modifications peuvent conduire à des événements extrêmement délétères : mutations, associées à certaines pathologies (cancer). Cependant, à long terme, il peut y avoir des mutations qui entraîneront l'évolution de l'espèce.

La réplication de l'ADN est un mécanisme cellulaire **semi conservatif** : sur la double hélice d'ADN de la cellule fille, un des deux bras provient du brin parental et l'autre est synthétisé lors de la réplication par polymérisation (brin fils). Chaque brin parental est lu dans le sens 3' → 5' et sert de matrice à la synthèse d'un brin complémentaire et anti-parallèle, appelé brin fils ou brin néosynthétisé, synthétisé dans le sens 5' → 3'. Pour cela les deux brins parentaux doivent se séparer, une rupture des liaisons hydrogènes est donc nécessaire.

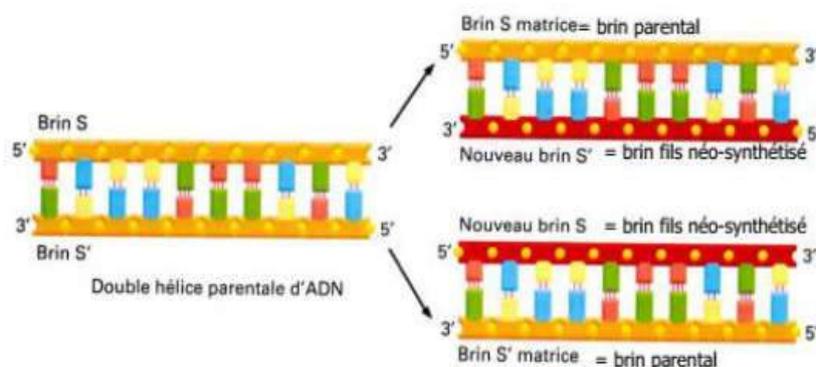


Figure 11 : Schéma récapitulatif de la réplication

4.1 REPLICATION CHEZ LES PROCARYOTES (*E. COLI*)

-L'ADN parental, où chacun des deux bras sert de matrice pour synthétiser le brin anti-parallèle. Séparation des deux brins (dénaturation) : pour passer à un état simple brin et le rendre accessible aux enzymes de la polymérisation. Il est aussi appelé ADN matrice.

-Les dNTP (dATP, dCTP, dTTP et dGTP) rentrent dans la réaction sous forme triphosphate, et sont intégrés dans le brin sous forme monophosphate : le relargage du pyrophosphate apporte l'énergie nécessaire à la réplication.

-Les ions Mg^{2+} stabilisent les dNTP en les protégeant d'une hydrolyse enzymatique et sont également importants pour l'ADN polymérase.

-Les enzymes suivantes :

- Les hélicases, à l'initiation de la réplication
- Les topoisomérases
- Les ADN polymérase
- Les primases

Les enzymes utilisées :

- **Les hélicases**

Elles permettent aux deux brins de l'ADN parental de se séparer en supprimant les liaisons hydrogène entre les bases qui se font face. Le déplacement le long de l'ADN parental nécessite une énergie (hydrolyse des molécules d'ATP en ADP et P_i).Après elles, des protéines se positionnent : les protéines SSB, qui se lient à des simples brins, pour éviter la renaturation spontanée de l'ADN.

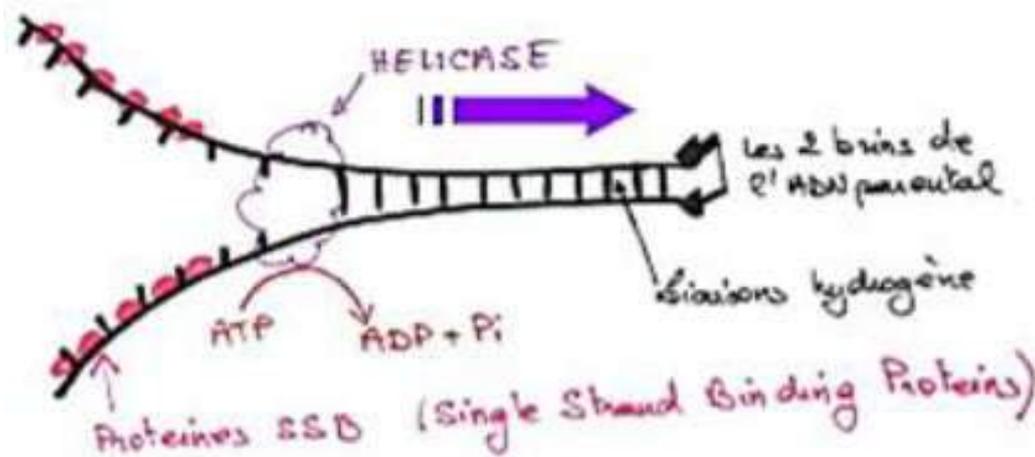


Figure 12: Enzyme intervenant pendant la réplication

- **Les topoisomérases**

Les topoisomérases modifient l'état d'enroulement (+ ou -) de la molécule d'ADN. Chez la bactérie, la topoisomérase est la **gyrase bactérienne** (TOP2).

Plusieurs étapes :

- 1) Coupure transitoire de l'ADN
- 2) Libre rotation de l'ADN
- 3) Ressoudure de l'ADN (= ligation), toujours par la topoisomérase

Il existe deux types de topoisomérases :

- Les topoisomérases I n'agissent que sur un seul brin de l'ADN et ne nécessitent donc pas d'ATP
- Les topoisomérases II agissent sur les deux brins de l'ADN, et nécessitent de l'ATP pour maintenir les deux fragments d'ADN. Chez les procaryotes, la gyrase bactérienne est une topoisomérase II.

Les topoisomérases agissent derrière l'action de l'hélicase, qui induit des surenroulements positifs (en avançant) qui vont super vriller la double hélice d'ADN. La gyrase va supprimer ces surenroulements positifs et introduire des surenroulements négatifs. Cette action est simultanée aux hélicases et permet une bonne accessibilité aux enzymes de la réplication à la molécule d'ADN.

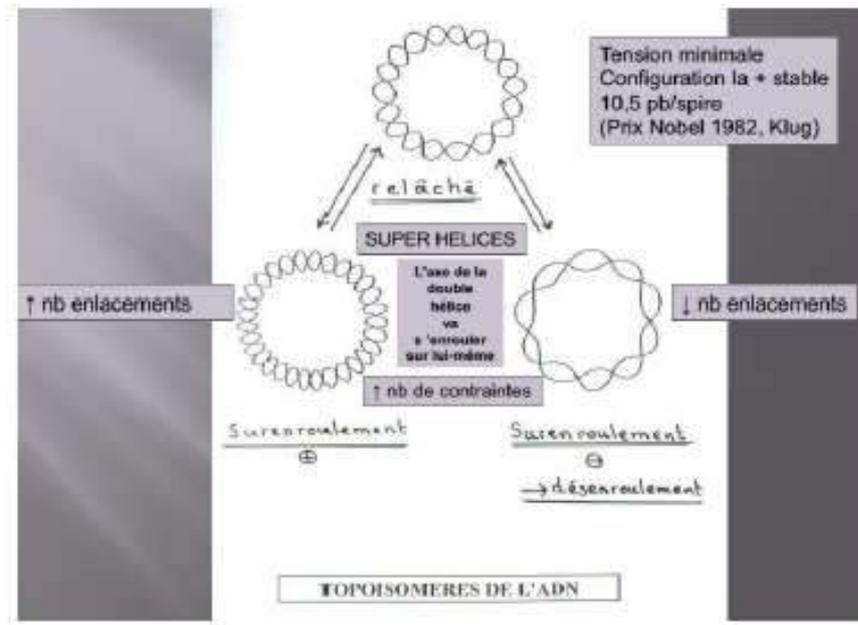


Figure 13 : Les différents états de l'ADN

• **Les ADN polymérases**

Elles permettent la synthèse du brin fils par polymérisation en utilisant un brin de l'ADN parental comme matrice. L'activité ADN-polymérasique est la synthèse de l'ADN dans le sens 5'→3' dans la chaîne du brin en cours de synthèse. Cette activité nécessite une initiation : démarrage par une amorce de nucléotides. Pour la réplication, l'amorce est un ARN synthétisée par polymérisation par une primase (ARN polymérase). L'ADN-polymérase se positionne sur l'extrémité 3'OH pour polymériser de l'ADN dans le sens 5'→3' après incorporation successive de nucléotides. La synthèse se fait de manière complémentaire et anti parallèle à l'ADN parental. Lors de la formation de la liaison phosphodiester, il y a libération de deux pyrophosphates qui s'hydrolysent en deux phosphates inorganiques.

La survie de la cellule à court terme dépend de **la fidélité** de la réplication. Cette haute-fidélité nécessite plusieurs **mécanismes de vérification** des nucléotides :

- **Avant addition** covalente du nucléotide : Le dNTP entrant est celui qui possède la meilleure affinité et un état énergétique optimal. A ce moment il y a une **transconformation** de la polymérase : elle change de conformation pour vérifier la bonne géométrie de la paire de base.

Après addition covalente du nucléotide : activité **exonucléasique** (= **fonction d'édition**) de la polymérase : correction des erreurs d'appariement en 3'↔5' par clivage entre deux nucléotides adjacents en bout de chaîne.

L'ADN polymérase a une activité exonucléasique dans le sens 3'↔5'. C'est une fonction primordiale dans la réplication, qui est possédée par toutes les ADN polymérases dans le sens 3'-5'. C'est un mécanisme qui concourt à l'augmentation de la fidélité de la réplication. L'enzyme se modifie pour retirer la mauvaise liaison, l'emmène au niveau de son site catalytique P (site de polymérisation), et la détruit.

Chez *E. Coli*, on retrouve trois ADN polymérases : **les ADN polymérases I II, III.**

L'ADN polymérase III ne fonctionne pas seule mais est associée à un complexe multimérique appelé holoenzyme. Holoenzyme ADN pol III est l'enzyme majeure, qui réalise la synthèse de l'ADN (des brins fils).

La 2^{ème} enzyme qui intervient dans la réplication chez la bactérie est l'ADN POL I, qui possède en plus une activité exonucléasique 5'-3' permettant la dégradation des amorces d'ARN.

L'ADN pol III a une activité ADN polymérasique dans le sens 5'-3' et une activité de correction exonucléasique dans le sens 3'-5'.

L'ADN pol I a trois activités : deux identiques à l'ADN pol III et une activité exonucléasique dans le sens 5'-3'.

- **primases**

Les primases sont des ARN polymérases ADN-dépendante : elles synthétisent les amorces d'ARN et sont dites ADN-dépendantes car elles se servent de l'ADN comme matrice. L'amorce d'ARN se fait dans le sens 5' → 3' du sens de la chaîne en cours de synthèse. Les primases sont capables de synthétiser de courtes séquences d'ARN (de 4 à 12 ribonucléotides) qui sont ensuite utilisées comme amorces par l'ADN polymérase répliquative.

Chez *E. Coli*, la primase et l'hélicase forment un seul complexe appelé **le primosome**.

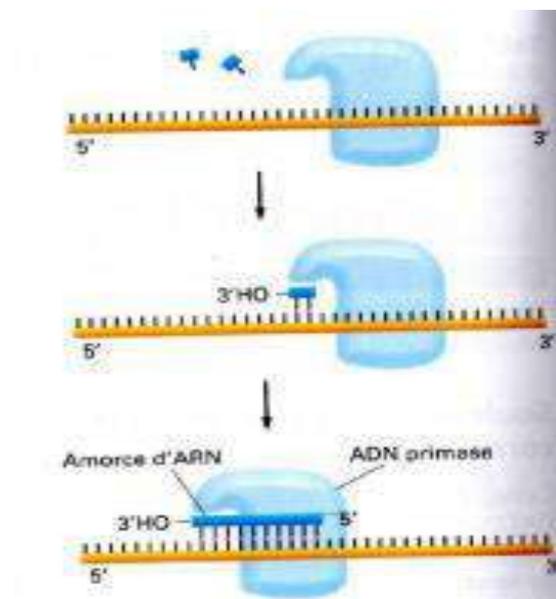


Figure 14 : Schéma représentant le primosome

4.2 MECANISME DE REPLICATION

1. *Initiation de la réplication*

Chez *E. Coli*, il n'y a qu'un seul chromosome, et environ 4,6 .10⁶ paires de bases.

Il existe une seule origine de réplication (appelée oriC : environ 245 pb), contenant des séquences répétées riches en A et T. C'est ici que se fixent les protéines d'amorçage, qui initient la réplication par l'ouverture locale de la double hélice d'ADN. Localement on va avoir le passage à la forme simple brin, c'est l'oeil de réplication.

A partir de là se propage la réplication. C'est une propagation bidirectionnelle. On a deux fourches de réplifications : une à droite et l'autre à gauche du point d'initiation.

Les deux fourches progressent jusqu'à avoir deux molécules d'ADN double brin. On obtient donc un chromosome pour chaque cellule filles (donc deux au total), le plus fidèle possible au chromosome bactérien de la cellule mère. Chacune des deux cellules filles possède un chromosome à deux brins, qui correspondent à un brin synthétisé et à un brin parental.

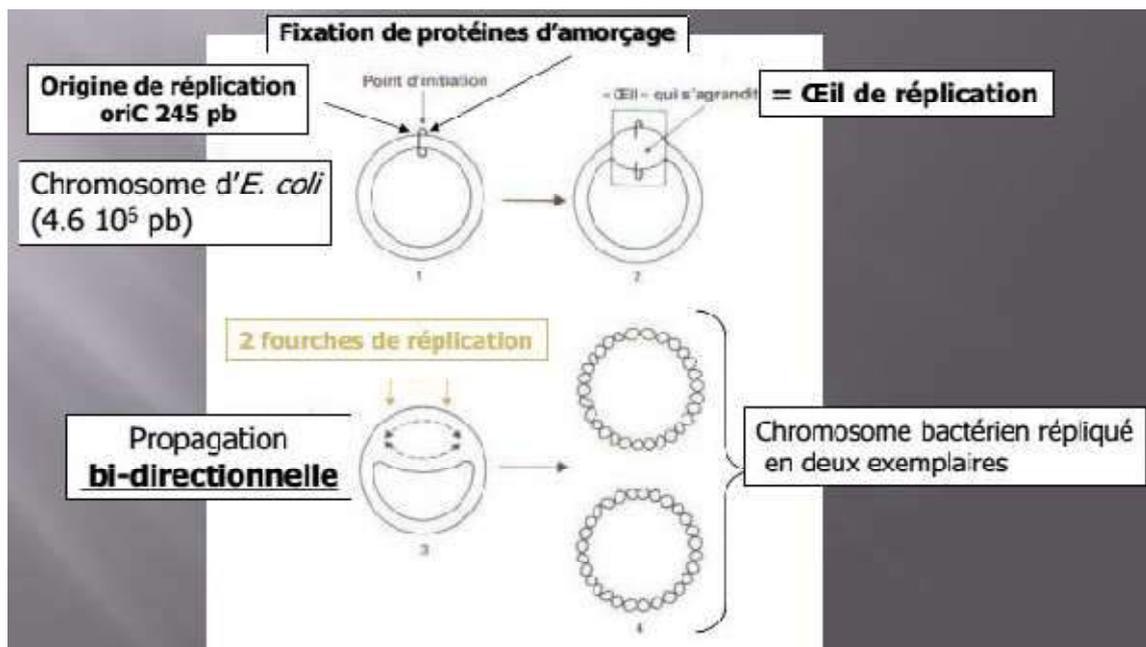


Figure 15: Mécanisme de réplication bidirectionnelle

RAPPEL DEROULEMENT DE LA REPLICATION:

- Hélicases
- Gyrase
- Protéine SSB pour éviter la renaturation et maintien de la structure sous une forme facile d'accès aux enzymes. Inhibent les structures en épingle à cheveux.
- amorces d'ARN : primase + hélicase = primosome
- synthétisées par la primase en 5'-3'

2. Réplication des 2 brins d'ADN parentaux

La synthèse des brins fils se fait par l'ADN polymérase III. L'activité enzymatique se fait toujours dans le sens 5'-3'.

Le sens de propagation de la réplication doit s'effectuer dans le sens d'ouverture de la boucle d'ADN, c'est-à-dire dans le sens 5'-3'. La réplication est donc discontinue sur un des deux brins.

On distingue un brin dit avancé et un brin dit retardé :

- Pour le **brin avancé**, la synthèse de l'ADN se fait dans le sens d'ouverture de la molécule d'ADN, donc dans le sens de propagation de la fourche de réplication. C'est le brin **continu**.

- Pour le **brin retardé** ou **discontinu**, la synthèse d'ADN se fait dans le sens contraire de l'ouverture la molécule d'ADN : l'ADN polymérase n'exerce son activité polymérasique seulement dans le sens 5'-3' de la chaîne en cours de synthèse, or sur ce brin d'ADN, le sens de propagation de la fourche de réplication est 3'-5'. La synthèse de ce brin est donc discontinue. Elle va s'effectuer de manière **rétrograde**, à partir de multiples fragments d'ARN. On obtient donc de multiples fragments d'ADN (fragments de 1000 à 2000 nt), appelés **fragments d'Okazaki**.

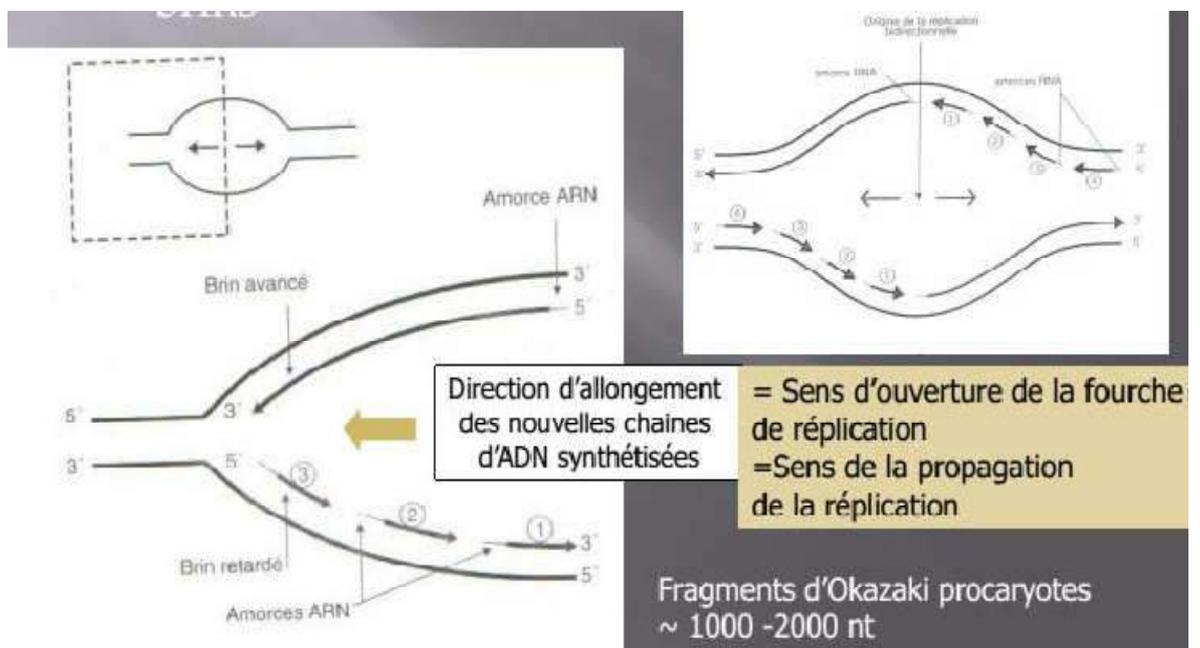


Figure 16: Mécanisme de réplication des deux brins chez les procaryotes

Comment s'effectue la synthèse simultanée des deux brins ?

Il existe un modèle démontré chez la bactérie et qui semble aussi être démontré pour les cellules eucaryotes.

Chez la bactérie : mise en jeu de deux molécules d'ADN polymérase III qui travaille en même temps au niveau de la fourche de réplication : une qui synthétise de manière continue le brin avancé et l'autre qui synthétise de manière discontinue le brin retardé.

Le brin d'ADN parental, qui sert de matrice au brin d'ADN retardé, forme une boucle pour repositionner l'amorce de façon à ce que l'allongement du fragment d'Okazaki par l'ADN polymérase se fasse dans le sens 5'-3' mais aussi dans le sens de propagation de la fourche de réplication.

Processivité de l'enzyme :

La processivité est donc la capacité de polymériser sur une longue distance. L'ADN polymérase III est naturellement peu processive : elle a tendance à synthétiser spontanément de l'ADN sur une courte distance et à se dissocier aussi spontanément de l'ADN parental. Cela est en adéquation avec la synthèse du brin discontinu : synthèse d'un fragment, puis détachement et synthèse d'un autre fragment. Mais sur le brin continu, l'ADN polymérase III doit synthétiser sur une longue distance. L'ADN POL III est donc maintenue au brin d'ADN parental par un collier coulissant (clamp). Pour *E. Coli*, c'est la sous-unité β du complexe holoenzyme qui joue le rôle de clamp. Cela permet à l'ADN POL III de rester associé au brin parental.

La vitesse de réplication procaryote est de **500 et 1000 nt/s**, ce qui permet à la réplication du chromosome bactérien de se faire en 40min.

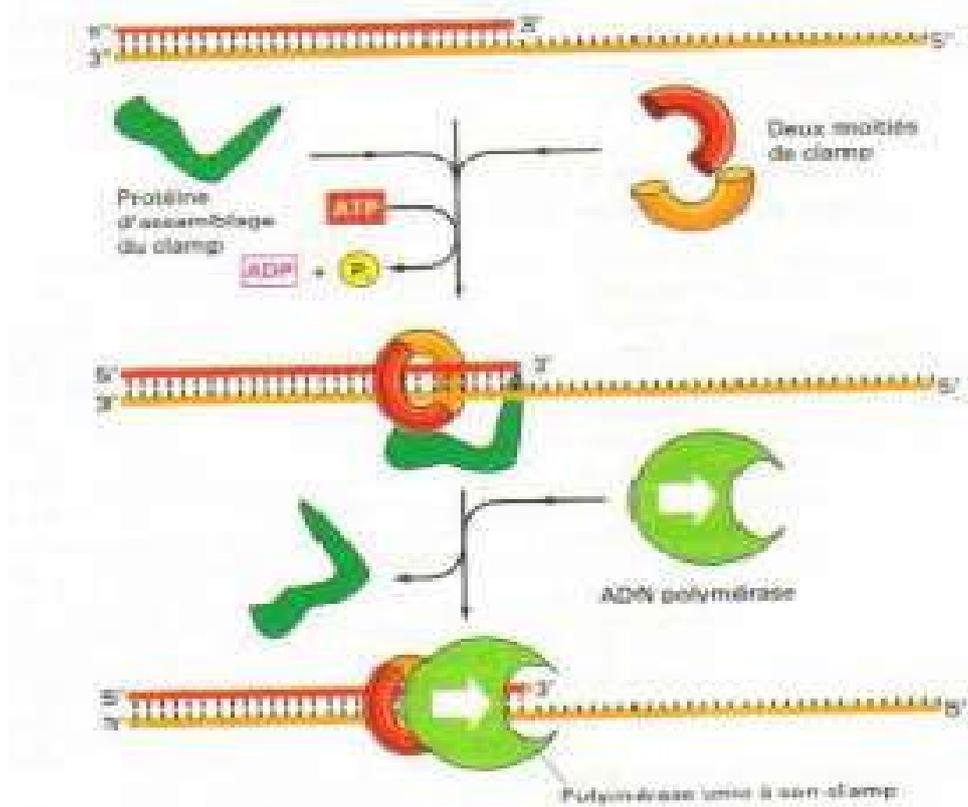


Figure 17 : Phénomène de processivité

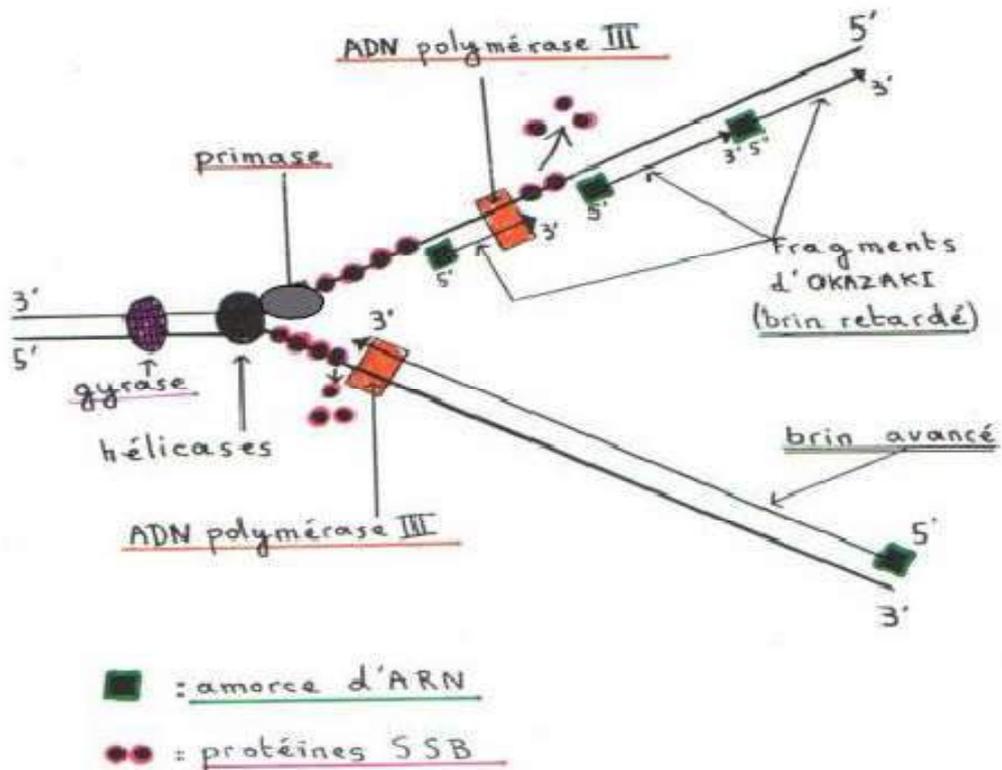


Figure 18 : Schéma récapitulatif de réplication chez les procaryotes

Finition des brins d'ADN « fils » :

Pour finir, il faut :

- Eliminer l'amorce ARN
- Remplacer les amorces d'ARN par de l'ADN
- Transformer le brin discontinu en un brin continu

Pour avoir un fragment d'ADN continu, l'ADN POL I se positionne au niveau 3'OH d'un fragment d'Okasaki et comble la lacune qui le sépare de l'amorce d'ARN suivante.

Dès qu'elle va rencontrer l'amorce d'ARN, de par son activité exonucléasique 5'-3', elle hydrolyse cette amorce d'ARN et continue, de par son activité ADN-polymérasique, à synthétiser de l'ADN à partir du fragment d'Okasaki précédent. Elle remplace ainsi l'amorce d'ARN en ADN.

La dernière étape consiste à relier ces fragments d'ADN grâce à une enzyme, appelée ADN ligase, qui permet la création d'une liaison phosphodiester entre 2 nucléotides adjacents.

On obtient alors un fragment continu.

Quand la synthèse d'ADN est terminée, les 2 fourches de réplication vont se rencontrer (car l'ADN est circulaire. Il y a alors séparation des 2 doubles hélices par la gyrase.



Figure 19 : Finition des brins d'ADN

5. Expression de l'information génétique

5.1 Transcription

La transcription est le premier processus de régulation utilisé par les cellules, tissus et organismes pour faciliter et contrôler les programmes complexes de l'expression génétique, le métabolisme cellulaire, et le développement des tissus et les organes. On appelle transcription la synthèse de brin d'ARN dont la séquence est dictée par celle de la molécule d'ADN correspondante.

Les trois principaux types d'ARN connus sont:

- ARN messager (**ARNm**) : Spécifie les codons.
- ARN de transfert (**ARNt**): Spécifie les anticodons.
- ARN ribosomique (**ARNr**): Constituant de ribosomes et impliqué dans la synthèse des protéines.

Chez les procaryotes ces ARN sont synthétisées par une seule ARN polymérase. Chez les eucaryotes elles sont synthétisées par trois types de polymérases, **ARN pol I**, **ARN pol II**, et **ARN pol III** (Tableau 1). La réaction nécessite un brin d'ADN qui sert de modèle, les nucléosides trois phosphate (ATP, GTP, CTP, UTP) et Mg^{++} .

Tableau 1 : Les ARN polymérases chez les cellules eucaryotes

Enzyme	Location	Produit
ARN polymérase I	Nucléole	rRNA (5.8S, 18S, 28S)
ARN polymérase II	Chromatine, matrice nucléaire	ARNm
ARN polymérase III	Chromatine, matrice nucléaire	ARNt, 5S, ARNr

Gene : Le gène est l'unité fonctionnelle de l'information génétique. Il est constitué d'un ensemble de nucléotides qui contient toutes les informations nécessaires pour transcrire un ARNm

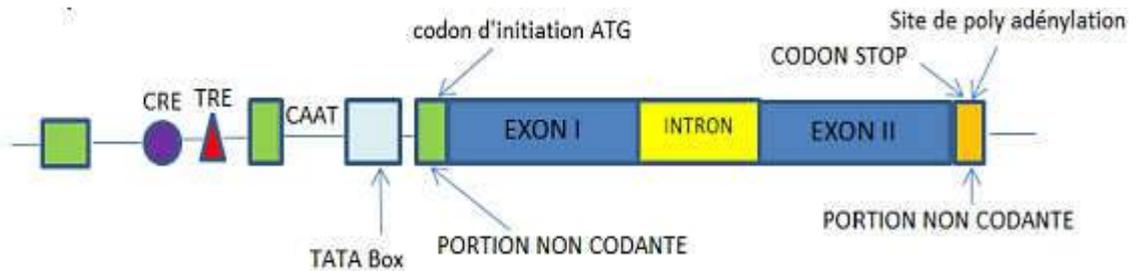


Figure 20 : Structure général d'un gène chez les eucaryotes. (CRE [Cyclic AMP Responsive Elément], TRE [12- O-Tetradacanoylphorbol 13-acetate Responsive Element]: des éléments de réponse).

Promoteur: Site sur l'ADN d'une centaine de bases auquel l'ARN polymérase (ainsi que les facteurs d'initiation requis) se fixe pour commencer la transcription. Chez *E. coli* il ya environ 2000 sites promoteurs (4.8 x10⁶ pb), elle sont qualifiés d'éléments *cis*.

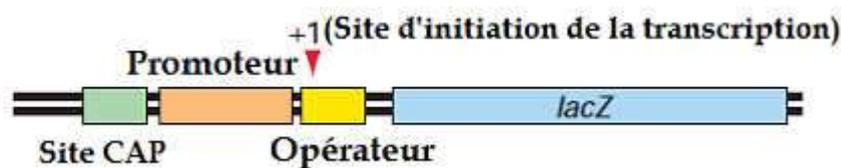


Figure 21: Structure partielle de l'opéron lactose d'*E. coli*

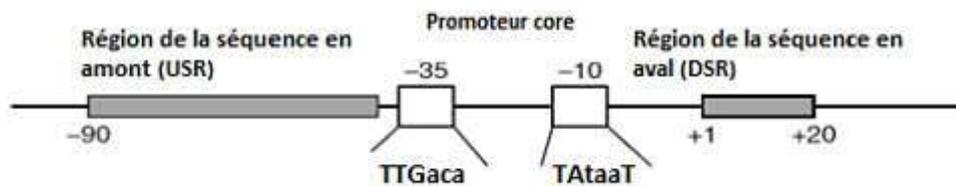


Figure 22 : Structure du promoteur bactérien.

Les bactéries possèdent une seule ARN polymérase (l'enzyme minimale) capable à elle seule de synthétiser de l'ARN, elle est composée de quatre sous unités, chacune codée par un gène différent. La forme active de l'enzyme contient les sous-unités $\alpha 2 \beta \beta' \omega - \sigma$ (PM \approx 500000 Da). Parmi ces sous unités les polypeptides β, β' sont à l'origine de l'activité catalytique et constituent le site actif de la transcription, ce site ressemble à celui de l'ADN polymérase par le fait que sous sa forme active, il contient deux ions métalliques. La sous unité σ aide à trouver un site promoteur ou doit commencer la transcription.

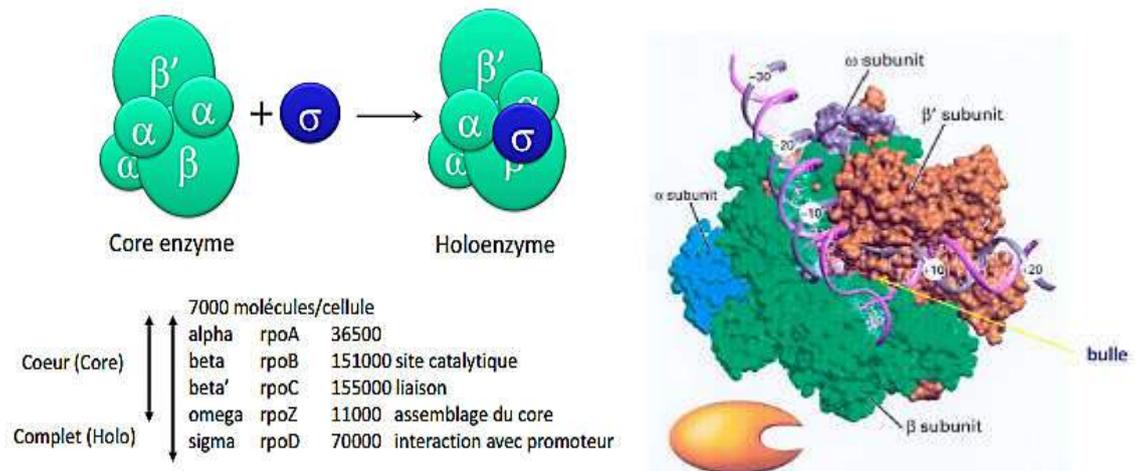


Figure 23 : Structure de l'ARN polymérase bactérienne.

Chez les bactéries, il existe plusieurs formes alternatives de l'ARN polymérase, chaque forme contient une sous-unité σ particulière (σ^{32} , σ^{54} , σ^{70} , σ^S et σ^E). La sous-unité σ^{70} est la plus grande forme (70 kDa). Ces éléments *trans* reconnaissent les séquences consensus de différents promoteurs et permettent une spécificité de l'initiation de la transcription. Cela permet de contrôler leur expression.

Tableau 2 : Exemples de facteurs sigma chez *E. coli*

Gène	Facteur	Rôle	Promoteur -35 espaceur -10
rpoD	σ^{70}	Général	TTGACA 15-18pb TATAAT
rpoH	σ^{32}	Choc thermique	CCCTTGAA 13-15pb CCCGATNT
rpoE	σ^E	Choc thermique	
rpoN	σ^{54}	Carence en azote	CTGGNA 6pb TTGCA
fliA	σ^F	flagelles	CTAAA 15pb GCCGATAA

5.1.1 La transcription chez les procaryotes

La synthèse de l'ARN, comme presque toutes les réactions biologiques de polymérisation comprend trois étapes: l'initiation, l'élongation, et la terminaison.

A. Initiation

L'ARN polymérase se lie au promoteur du gène. Les deux brins d'ADN se déroulent à cet endroit (formation de bulles de transcription) et la transcription commence

Ces structures secondaires suivies d'une série d'uridines sont des terminateurs efficaces de la transcription. Des séquences riches en GC suivie par autres riche en AT sont aussi des sites spécifiques de terminaison. Ces régions qui ne nécessitent pas de facteurs additionnels sont nommées "**terminateurs intrinsèques**".

D'autres types de terminateurs nécessitent des facteurs protéiques spécifiques. Chez *E. coli* la protéine *Rho* se fixe fortement à l'ARN (mais pas ni à l'enzyme ni à l'ADN), elle parcourt alors la chaîne jusqu'au complexe ARN polymérase-ADN. Dès que l'ARN polymérase s'arrête à un site de terminaison *Rho* dépendant. Le facteur *Rho* favorise la libération de l'ARN et de l'enzyme positionnée sur l'ADN, terminant ainsi la transcription.

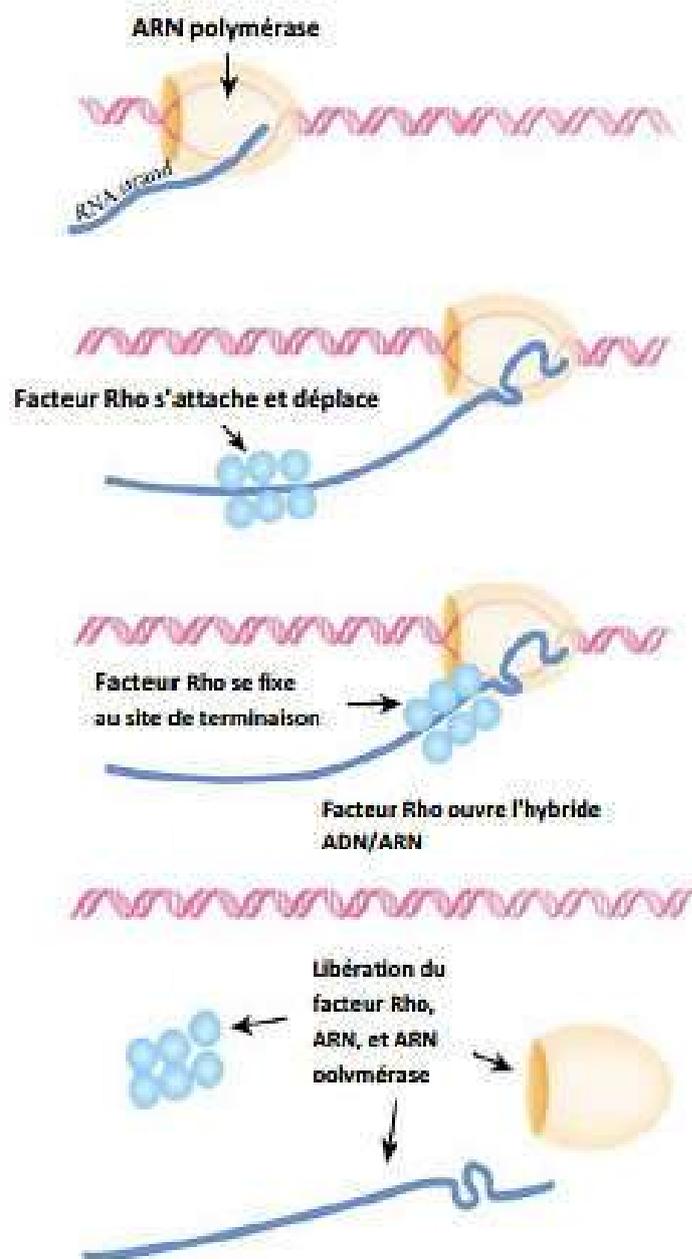


Figure 28 : Terminaison de la transcription chez les bactéries par le facteur *Rho*

5.1.2 La transcription chez les eucaryotes

La transcription chez les eucaryotes est prise en charge par des ARN polymérase très apparentées à celles présentes chez les procaryotes. Mais c'est un processus beaucoup plus complexe et il existe plusieurs différences notables:

- La transcription chez les eucaryotes est prise en charge par 3 ARN polymérases (Pol I, II, et III) très apparentées à celles présentes chez les procaryotes.
- La transcription au contraire de la traduction se fait au niveau du noyau.

- Elle nécessite plusieurs facteurs d'initiation (facteurs généraux de transcription (FGT: TFII A, TFII B, TFII D (TBP, TAF), TFII E, TFIIH).

- Remodelage de la chromatine par des protéines régulatrices et des enzymes de modification de la chromatine ----> : changement de conformation lors de l'ouverture de la double hélice de l'ADN (ADN inclus dans les nucléosomes).

- En plus des éléments cis (ex: TATA Box) il existe en plus du promoteur des autres séquences ou unités de contrôle :

- Activateurs : **Enhancers**

- Extincteurs: **Silencers.**

Ces séquences peuvent être localisées dans la région régulatrice en 5', en amont du point d'initiation, mais aussi parfois à l'intérieur du gène ou même dans la région 3' en aval de la séquence codante

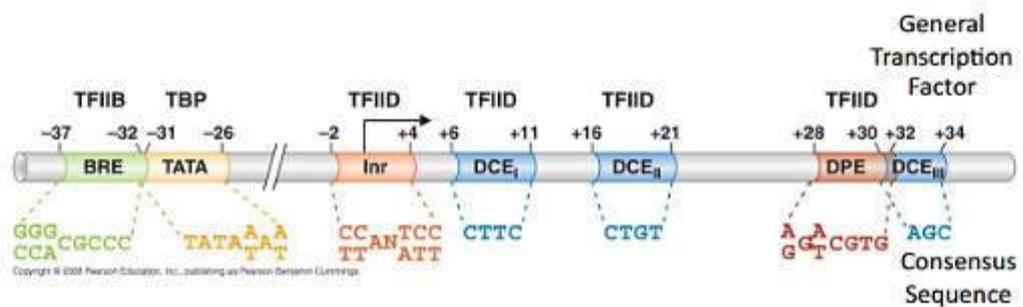


Figure 29 : Les composants d'un promoteur eucaryote.

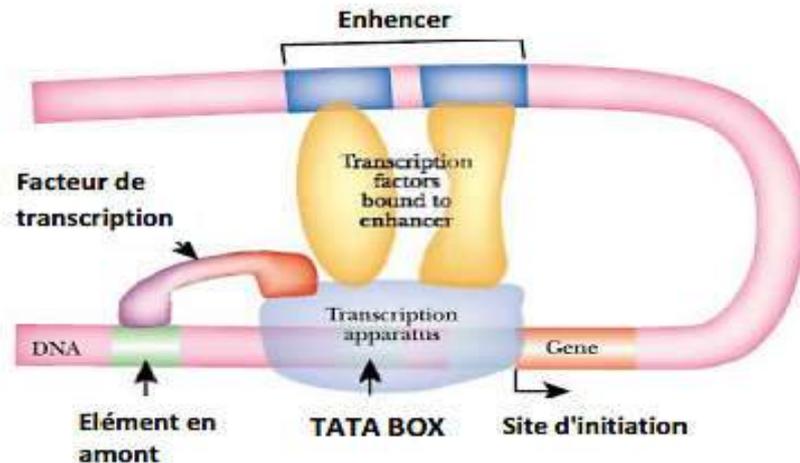


Figure 30 : Contrôle de la transcription par un enhancer localisé en aval (downstream) du site de l'initiation de la transcription. L'ADN forme une loop pour faciliter le contact de l'enhancer avec le complexe de l'initiation de la transcription par l'intermédiaire des facteurs de transcription.

Comme chez les procaryotes, la transcription se fait en 3 étapes : initiation, élongation et terminaison.

A. Initiation

La reconnaissance du promoteur et la mise en place du complexe de pré initiation chez la plupart des pol II débute au niveau du TATA box , ce dernier est reconnu par TFIID, plus précisément par la sous unité **TBP** (TATA box **B**inding **P**rotein). D'autres sous unités de ce complexe sont appelées **TAF** (TBP associated factors) reconnaissent d'autres éléments du promoteur, tels Inr, DPE et DCE, bien que la liaison la plus forte soit celle entre TBP et TATA.

Le complexe TBP-TATA attire sur le promoteur d'autres facteurs généraux de transcription [TFIIA → liaison du Pol II au promoteur (upstream), TFIIB → liaison du Pol II au promoteur (downstream), TFIIF → accompagne Pol II dès qu'elle se fixe sur le promoteur, TFIIE → essentiel pour l'élongation et la libération du Pol II du promoteur , et TFIIH → il a un rôle dans la phosphorylation du CTD du Pol II et il a aussi un rôle dans l'élongation] et la polymérase elle même. La double hélice à ce niveau se sépare par l'intermédiaire du TFIIH en hydrolysant la liaison anhydride de l'ATP.

La grande sous unité de pol II possède sur l'extrémité C-terminale une série de séquence répétée de 7 acides aminés (**heptapeptide: Tyr-Ser-Pro-Thr-Ser-Pro-Ser**) appelée domaine

carboxy terminal (CTD) ou la queue. Le nombre de répétitions dépend des espèces: 27 fois chez les levures, 45 fois chez la mouche *Drosophila*, et 52 fois chez l'homme. Chaque motif répété contient des sites de phosphorylation par des kinases spécifiques, notamment une qui est une sous unité de TFIIF. La régulation du niveau de phosphorylation du CTD de Pol II contrôle aussi les étapes ultérieures (élongation, maturation).

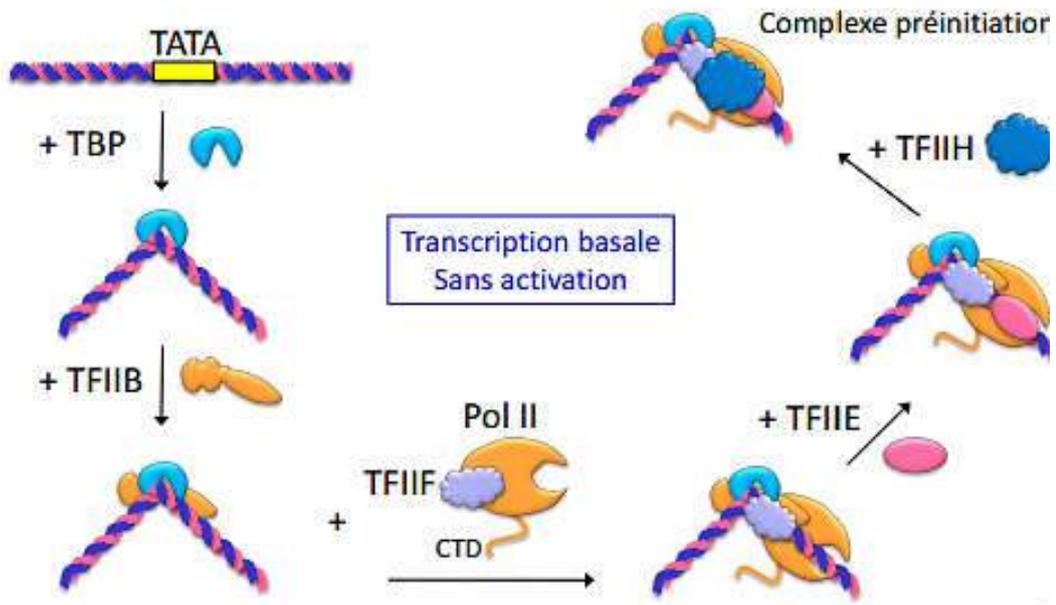
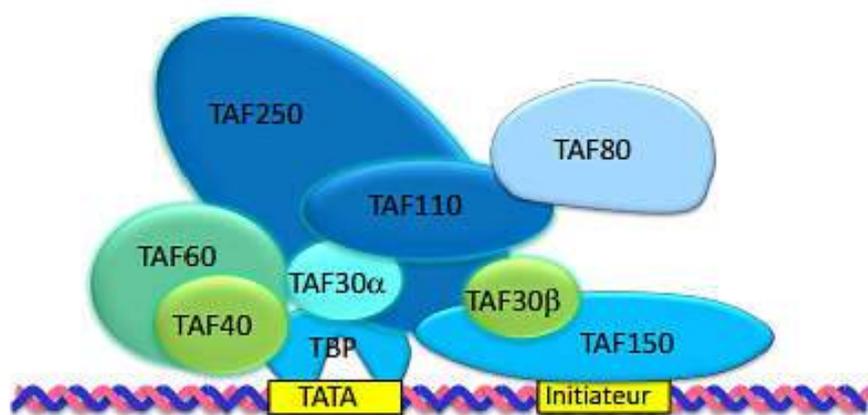


Figure 31 : Formation du complexe de démarrage Pol-II.



TAF: TBP-Associated Factors : ouvrent la chromatine, acetylent histones etc.

Figure 32 : Formation du complexe TFII-D : Activation de la transcription.

B.Elongation

Dans l'étape de l'élongation la Pol II se débarrasse de la plupart de ses facteurs d'initiation, comme les FGT et le médiateur (requis pour atteindre des niveaux élevés de transcription *in vivo*: car l'ADN est empaqueté sous forme de chromatine). De nouveaux facteurs appelés facteurs d'élongation vont les remplacer.

Différents protéines sont supposées stimuler l'élongation comme la kinase P-TEFb (phosphoryle des résidus Ser de CTD, et active le facteur d'élongation hSPT5), TAT-SF1, P-TEFb (stimule l'élongation par trois voies distinctes). Il existe aussi des facteurs d'élongation de type ELL (suppriment les arrêts temporaires de l'enzyme). Un autre facteur important dans l'élongation et n'a aucun rôle dans l'initiation est le TFIIS, ce facteur joue de rôle important le premier est de réduire les pause de la polymérase et le deuxième est de corriger les erreurs (mésappariements) par l'activité RNase.

Comme l'initiation de la transcription, l'élongation se déroule en présence d'histones, ce qui fait appel à des facteurs facilitant la transcription en présence de chromatine. Le facteur FACT (facilitates Chromatin Transcription) rend la transcription sur type de matrice chromatine beaucoup plus efficace. C'est un hétérodimère de deux protéines, la première Spt16 se lie sur le module H2A/H2B et la deuxième SSRP1 se lie sur le module H3/H4 pour démanteler les nucléosome.

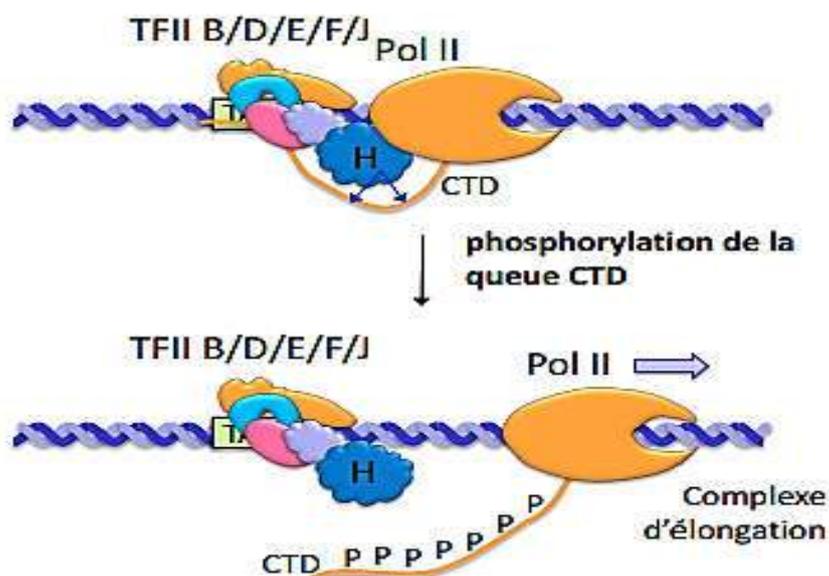


Figure 33 : Elongation de la transcription.

C. Terminaison

La transcription se termine aux alentours de la séquence de polyadénylation transcrite. La protéine CPSF (Cleavage and polyadenylation Specificity Factor) qui était associée à l'ARN polymérase s'y lie. La perte de l'interaction CPSF/ARN polymérase déstabilise le complexe de transcription.

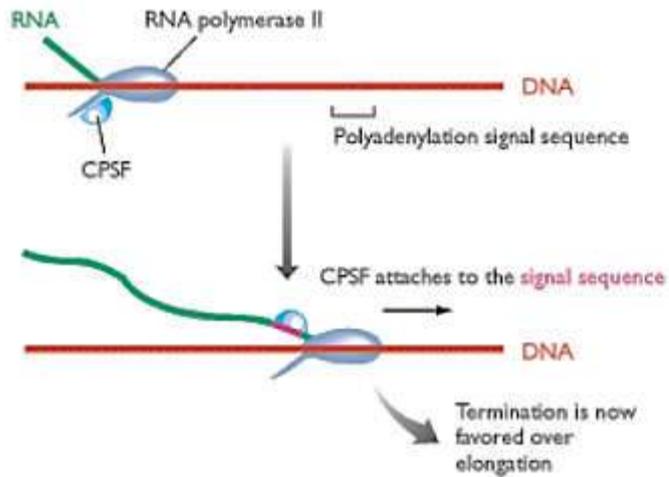


Figure 34: Terminaison de la transcription chez les eucaryotes.

5.1.3 Maturation des ARNm

La maturation des ARNm prématurés se fait au niveau du noyau et entraîne des changements dans leur structure.

-Addition d'une coiffe (7mG) en 5' et d'une queue (poly A) en 3' pour la plupart des transcrits destinés à devenir des ARNm.

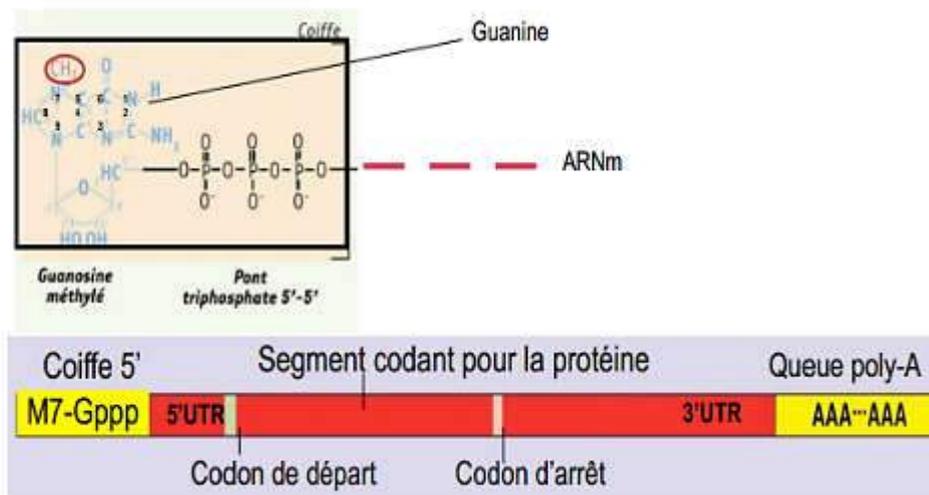


Figure 35 : Ajout d'une coiffe de guanosine méthylée.

-Ajout d'une queue poly-A à l'extrémité 3' (de 50 à 250 nucléotides d'adénine). Dès la fin de la transcription, avant même que l'ARN ne quitte le noyau. La polyadénylation facilite l'exportation des ARNm hors du noyau, les protège des dégradations une fois dans le cytosol et facilite leur traduction.

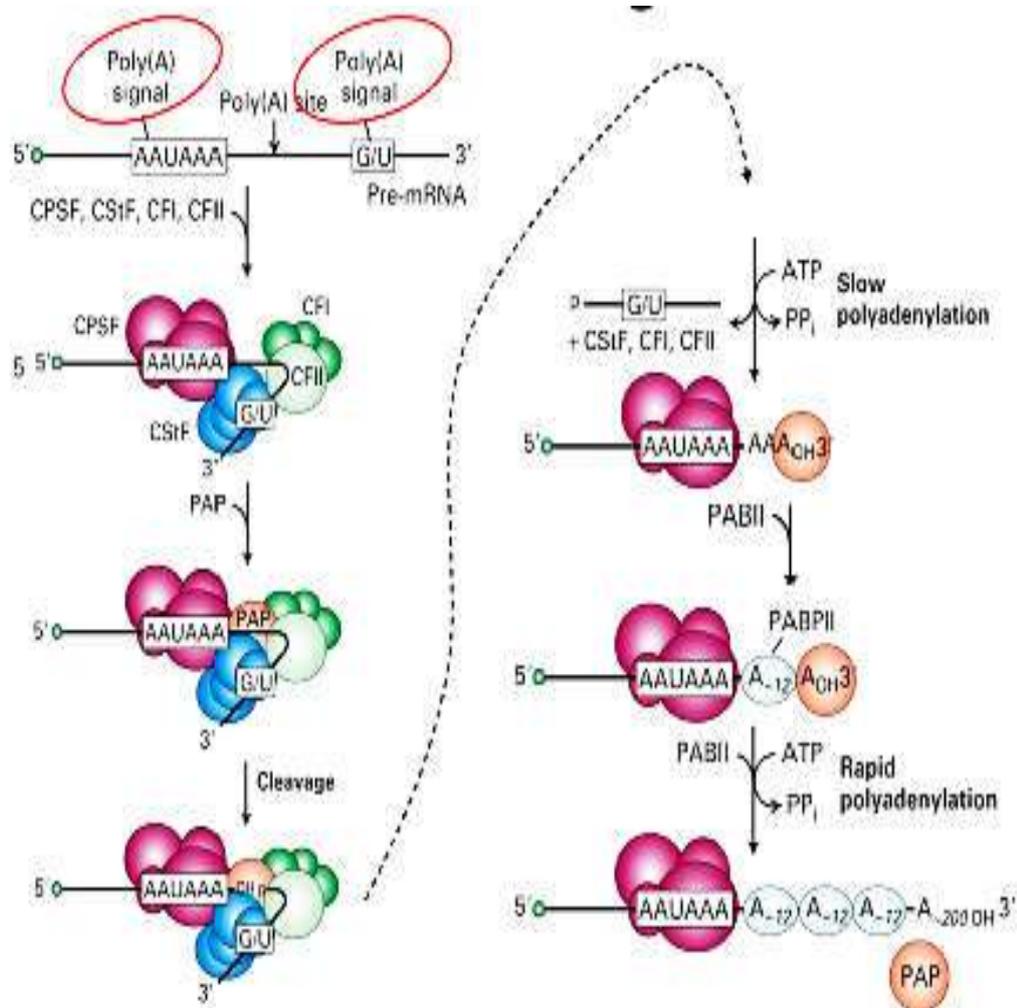


Figure 36 : Mise en place de la queue poly-A en 3' du messager.

-Le transcrit eucaryotes contiennent des sections non codantes (**les introns**) et des sections codantes (**les exons**). **L'épissage** consiste à enlever les introns puis à recoller les exons. Les introns restent à l'intérieur du noyau puis sont dégradés. L'épissage est catalysé par des **snRNP (Small nuclear ribonucleotide Particles)** plus d'autres protéines. L'ensemble constituant le **spliceosome**. Les RNP sont des structures multimoléculaires composées de protéines et de petits ARN. U1 et U2 se fixent d'abord sur le pre-messager, puis U4 et U6 viennent interagir avec U1 et U2, ce qui rapproche les deux extrémités exonique. Puis

L'activité catalytique du spliceosome permet de cliver la séquence intronique et de liguer les séquences exoniques.

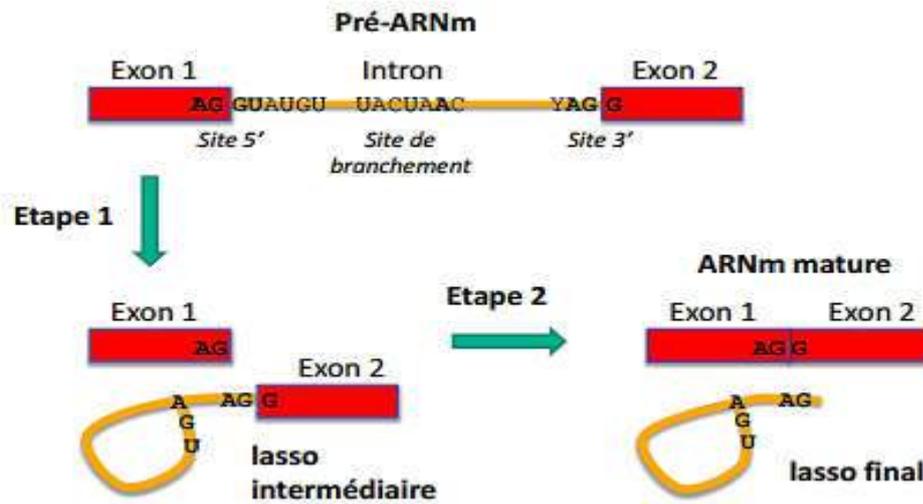


Figure 37: Les deux étapes de la réaction catalytique d'épissage.

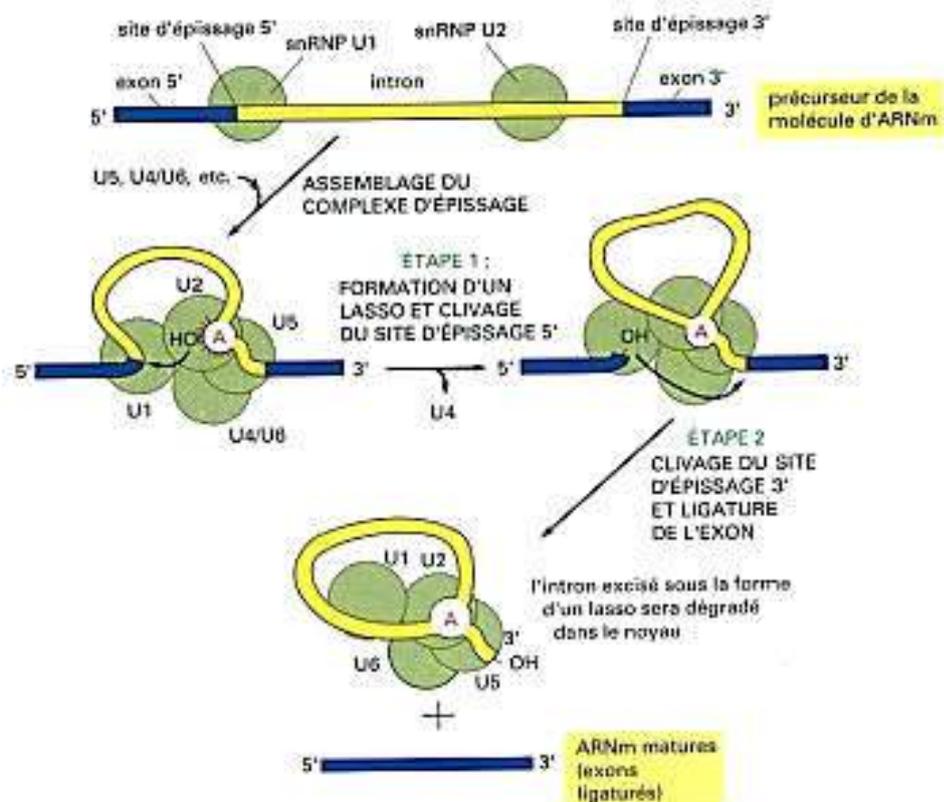


Figure 38 : Catalyse et épissage par le complexe des snRNP U1, U2,U5,U4/U6

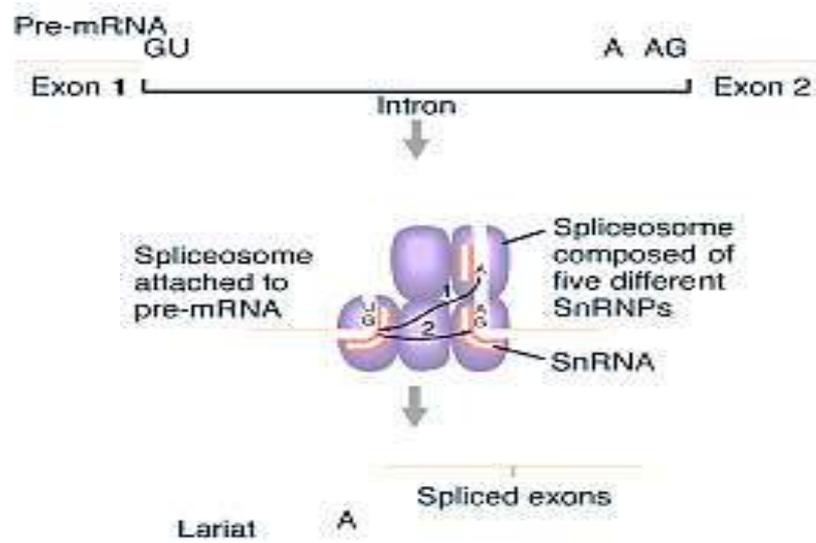


Figure 39: Les étapes d'assemblage du splicéosome.

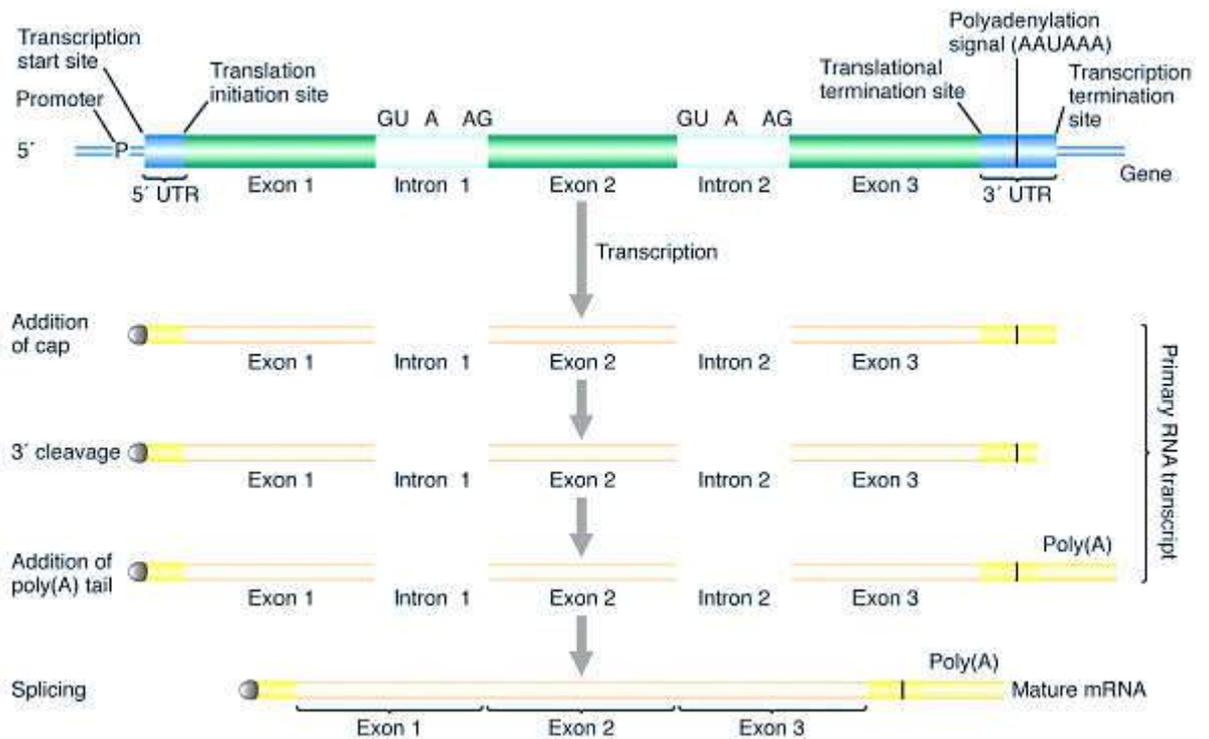


Figure 40 : Résumé des étapes de maturation de l'ARN des eucaryotes.

L'épissage alternatif permet de générer des protéines différentes à partir d'un même gène.

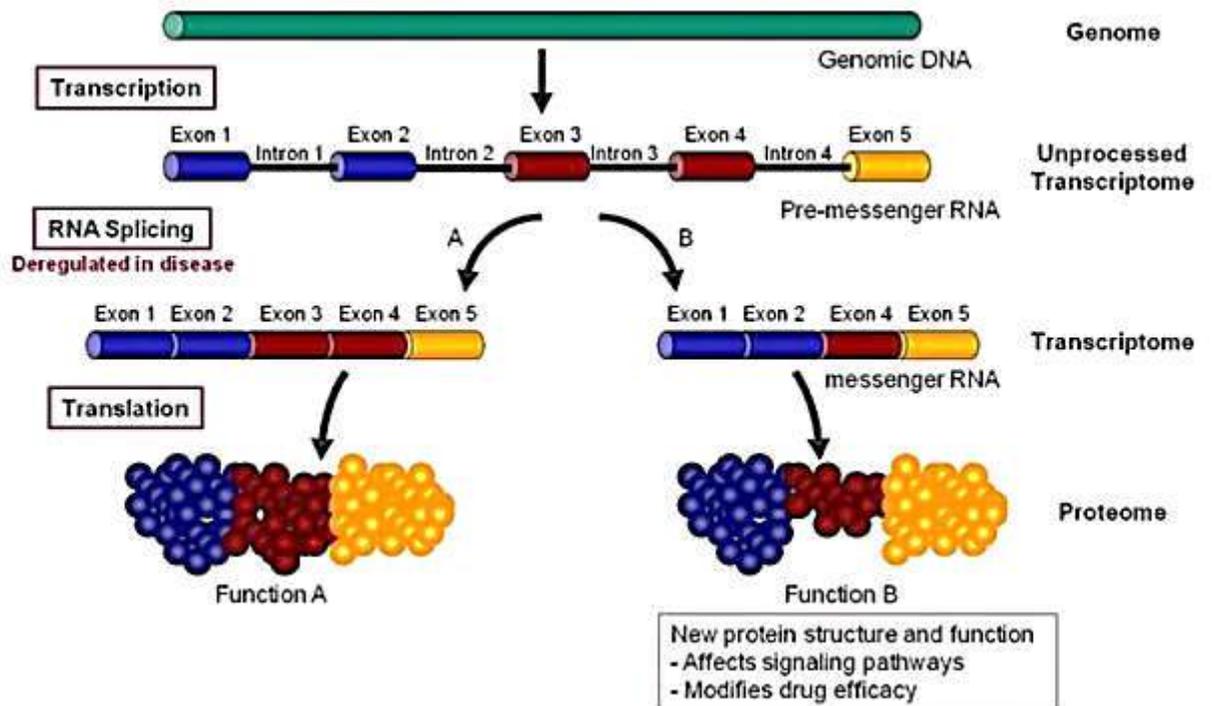


Figure 41 : Epissage alternatif.

5.2 La Traduction

La traduction est parmi les événements les plus conservés chez tous les organismes, et les plus coûteuses en énergie. Dans la réplication et la transcription le transfert de l'information génétique nécessite une matrice d'ADN pour produire un acide nucléique (ADN, ARN), mais dans la traduction la matrice est l'ARN et le produit final est une polypeptide ou protéine.

Quatre composants principaux sont impliqués dans la traduction:

- 1- **ARNm**: Cette macromolécule fournit l'information qui doit être interprétée par la machinerie de la traduction et constitue la matrice pour la traduction. La région de l'ARNm codant la protéine est constituée d'une série ordonnée d'unités de trois nucléotides appelées codons (spécifient l'ordre des acides aminés).
- 2- **ARNt**: Ils fournissent l'interface entre les acides aminés ajoutés à la chaîne peptidique en croissance et les codons dans l'ARNm.

- 3- **Aminoacyl-ARNt synthétase:** Ces enzymes couples des acides aminés avec des ARNt spécifiques qui reconnaissent le codon approprié.
- 4- **Ribosomes:** Le ribosome coordonne la reconnaissance de l'ARNm par chaque ARNt et catalyse la formation de la liaison peptidique entre la chaîne polypeptidique en croissance et l'acide aminé chargé sur l'ARNt sélectionné.

-La synthèse de la protéine est centralisée au niveau des ribosomes.

-Les ribosomes se déplacent dans le sens 5' vers 3' sur l'ARNm et synthétisent le polypeptide correspondant de l'extrémité **NH2 terminale** vers l'extrémité **COOH terminale**.

- La séquence de l'ARNm est décodée par groupe de trois nucléotides (codon) qui correspondent à un acide aminé particulier ou aux signaux d'initiation et de terminaison

4.2.1 Le code génétique

Le code génétique établit une correspondance entre la matrice d'acides nucléiques et le polypeptide produit (Un acide aminé correspond à un codon). Le code génétique est écrit en ARNm plutôt qu'avec l'ADN. L'aspect le plus intéressant du code génétique est que certains acides aminés sont codés par plusieurs codons apparentés mais différents. Parmi un total de 64 codons, 4 sont dédiés à l'initiation et à l'arrêt de la traduction.

Deuxième base

		U	C	A	G	
U	UUU	Phe	UCU	UAU	UGU	Troisième base (extrémité 3')
	UUC					
	UUA	UCA	UAA Arrêt	UGA Arrêt		
	UUG	UCG	UAG Arrêt	UGG Trp		
C	CUU	Leu	CCU	CAU	CGU	
	CUC					CCC
	CUA	CCA	CAA	CGA	Arg	
	CUG	CCG	CAG	CGG		
A	AUU	Ile	ACU	AAU	AGU	
	AUC					ACC
	AUA	ACA	AAA	AGA	Arg	
	AUG	ACG	AAG	AGG		
G	GUU	Val	GCU	GAU	GGU	
	GUC					GCC
	GUA	GCA	GAA	GGA	Gly	
	GUG	GCG	GAG	GGG		

Cheez les eucaryotes (pointing to AUG)

Figure 42 : Correspondance entre les triplets de bases de l'ARNm et les acides aminés et les non sens (codon stop).

Chez les procaryote:

>>> **Le codon AUG code pour la formyl-méthionine**

-Le codon AUG (code pour formyl-MET ou MET) = **codon d'initiation**

-Le premier nucléotide du premier codon détermine le cadre de lecture ouvert (ORF : Open Reading Frame).

-Il y a 64 possibilités de combiner les nucléotides en codons (4³). Comme il y a 20 acides aminés, il y a donc 44 codons supplémentaires :

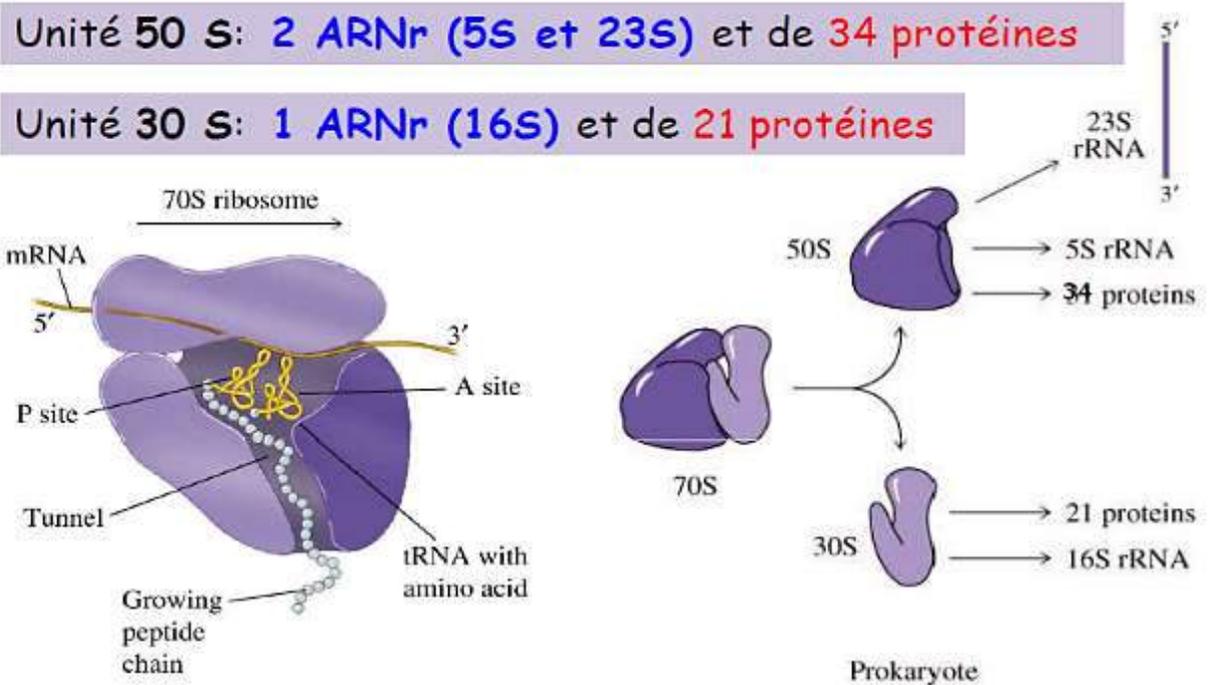
- 3 correspondent à des codons stop ou non-sens: **UAA, UAG et UGA**
- les autres sont des synonymes qui codent pour différents acides aminé

(**Dégénérescence du code génétique**)

-**Code UNIVERSEL**: c'est le même pour tous les êtres vivants sauf quelques rares exceptions. Cette caractéristique permet le transfert de gènes d'une espèce à l'autre = génie génétique.

5.2.2 Les ribosomes

Une cellule en phase active de croissance possède des milliers de ribosomes. Ce sont les sites de synthèse protéique. Chaque ribosome est composé de deux sous unités 30S (unité Svedberg: unité de coefficient de sédimentation) et 50S, l'ensemble donne des ribosomes 70S. La machinerie de polymérisation des acides aminés est composée d'au moins 3 ARN et 50 protéines différentes. Les ARNm procaryotes possèdent un site de liaison au ribosome (RBS) appelé séquence de Shine Dalgarno (S/D). Cet élément typiquement localisé 3 à 9 nucléotides en amont du codon d'initiation, il s'apparie avec la séquence complémentaire localisée sur l'ARNr 16S. La grande sous-unité contient le centre peptidyl-transférase qui est responsable de la formation des liaisons peptidiques. La petite sous-unité contient le centre de décodage dans lequel les ARNt chargés lisent ou décodent les codons de l'ARNm. Les sous-unités ribosomiques sont composées d'un ou plusieurs ARNr et de nombreuses protéines



✓ ARNm est associé à la sous unité 30S

✓ 2 sites de liaison à l' ARNt (sites P et A)/ sous unité 50S

Figure 43 : Structure générale des ribosomes procaryotes . La figure montre la composition en ARNr et en protéines des différents sous unités, de même que la longueur des ARNr et le nombre de protéines ribosomiques.

Tableau 3 : Structure comparée des ribosomes

	Taille (S)	Sous - unités (S)	ARNr (S)	Nombre de Protéines
Bactéries	70	50	23 + 5	31 (L1, L2 ..)
		30	16	21 (S1, S2...)
Eucaryotes	80	60	28 + 5,8 + 5	40
		40	18	30
Mitochondries	55	35	21 + 3	-
		25	12	

5.2.3 Rôles des ARNt dans la traduction

L'ARNt joue le rôle d'adaptateur entre les codons et les acides aminés qu'ils spécifient. Et l'ARNr en plus de son rôle structural des ribosomes en association avec des protéines, il a un rôle dans la fixation de l'ARNm (ARNr16S chez les procaryotes) et la formation des liaisons peptidique (activité peptidyl transférase de l'ARNr 23S chez les procaryotes). La liaison du tRNA à son acide aminé spécifique est catalysée par une enzyme: aminoacyl tRNA synthetase.

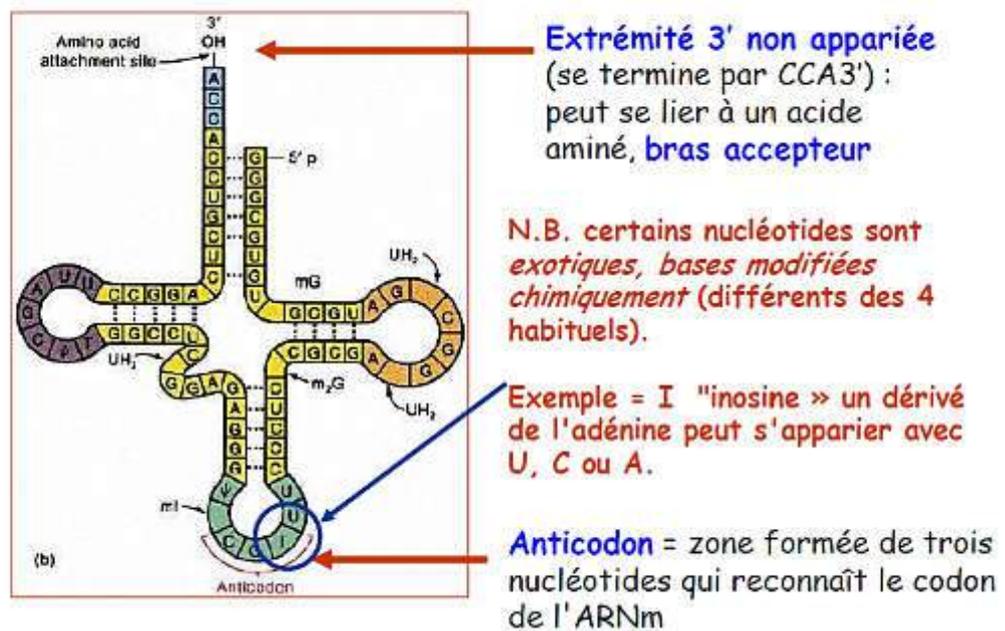
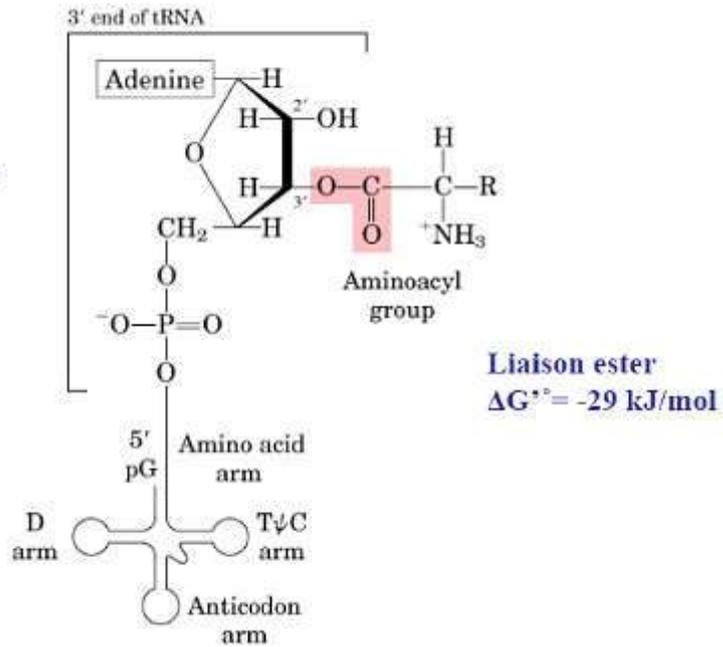
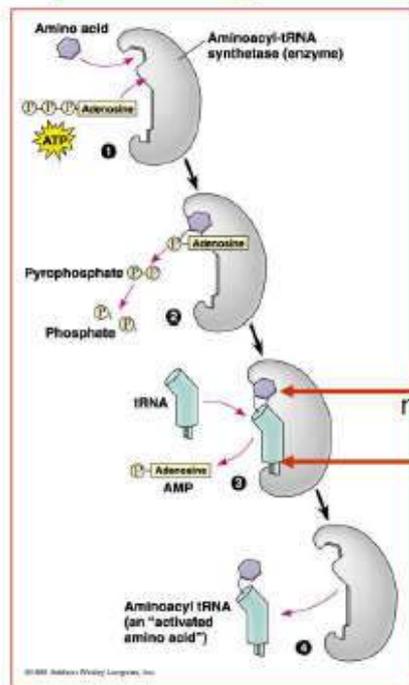


Figure 44 : Structure de l'ARNt déchargée (sans acide aminé). Cette structure dite en feuille de trèfle. L'acide aminé se fixe au niveau du ribosome du A terminal de l'extrémité acceptrice.

Ces enzymes sont considérées comme étant les réels responsables de lecture du code génétique et assurent donc la fidélité de la traduction.



❖ Aminoacyl-ARNt synthétase



Il existe 20 aminoacyl-ARNt synthétase



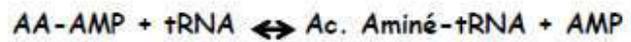
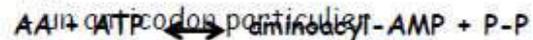
L' enzyme unit l' acide aminé à l' ARNt adéquat

Site actif de l' enzyme

reconnaît:

- un acide aminé particulier

Et



acides aminés activés et ARNt chargé

Figure 45: Réaction d' aminoacylation catalysée par une aminoacyl-ARNt synthétase.

5.2.4 Étapes de la traduction chez les procaryotes

Au delà des aminoacyl-ARNt synthétases, des ARNt, les ribosomes, et les ARNm cette synthèse nécessite plusieurs protéines appelées : facteurs d'initiation, d'élongation et de terminaison, tandis que l'énergie est fournie par la molécule de GTP. Les structures différentes dans les ARNm procaryotes et eucaryotes impliquent des voies différentes pour ces événements.

A. Initiation de la traduction

Pour que la traduction soit initiée avec succès, trois événements doivent se produire

- 1- Le ribosome doit être recruté sur l'ARNm
- 2- Un ARNt initiateur chargé par N-formyl-Met doit s'insérer dans le site P du ribosome
- 3- Le ribosome doit être positionné de manière précise sur le codon d'initiation.

Le N-formylmethionyl- ARNt initiateur se lie d'abord à la sous-unité libre 30S. Ensuite, l'ARNm se lie à la sous-unité 30S et se positionne correctement par des interactions à la fois avec l'extrémité 3' de l'ARNr 16S et avec l'anticodon fmét-ARNt. Finalement la sous-unité 50S s'attache au complexe Sous-unité 30S –ARNm. Le résultat de l'initiation est la formation d'un ribosome complet (70S) un niveau du site d'initiation de l'ARNm, avec le fMet-ARNtifMet occupant le site P, et un site A libre. Les aminoglycosides comme la streptomycine sont des antibiotiques largement utilisés contre les infections bactériennes comme la tuberculose. Ces antibiotiques inhibent la formation du complexe d'initiation 70S en bloquant l'assemblage des deux sous unités.

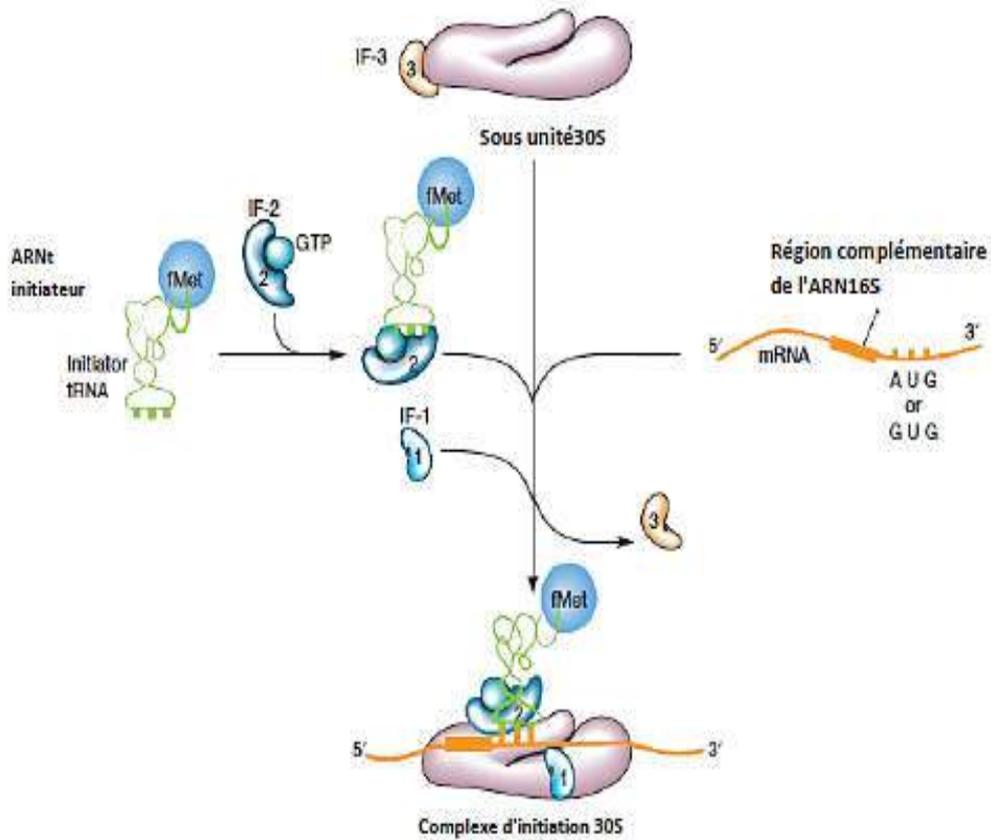


Figure 46 :Le processus de l'initiation de la traduction chez les procaryotes.

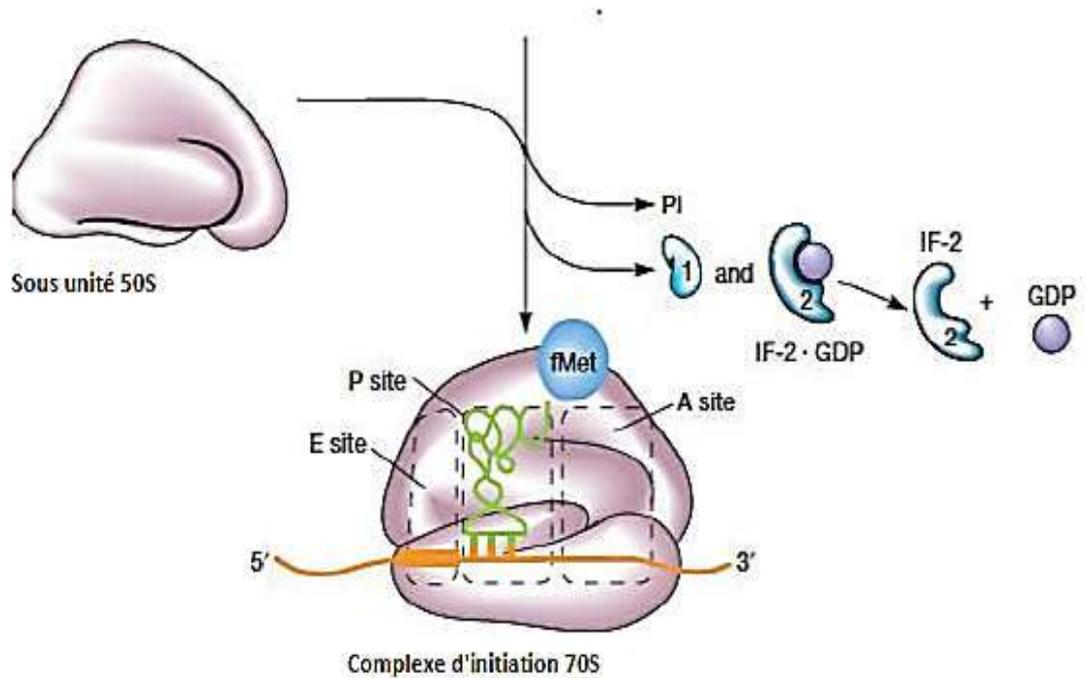


Figure 47:Formation du complexe d'initiation 70S. IF: initiation factor. A: Aminoacyl or acceptor site, P: peptidyl site, E: Exit site

Chez les procaryotes trois facteurs d'initiation dirigent l'assemblage d'un complexe d'initiation.

- 1- Facteur d'initiation 3 : IF-3 :** se lie à la petite sous-unité et l'empêche de s'associer avec la grande sous-unité libre, et favorise la liaison correcte de l'ARNm.
- 2- Facteur d'initiation 2 : IF-2** c'est une protéine liant et hydrolysant la GTP (GTPase), il lie le GTP au fMet-ARNt initiateur pour faciliter sa liaison à la sous unité 30S et empêche la fixation de d'autre ARNt chargé.
- 3- Facteur d'initiation 1: IF-1** c'est un facteur semble nécessaire à la libération du IF-2 et du GDP. Il empêche la fixation d'ARNt sur la partie de la petite sous-unité qui fera partie du site A. IF-1 peut aussi faciliter l'assemblage des deux sous-unités.

B.Élongation

Une fois le ribosome 70S assemblé, avec l'ARNt initiateur chargé au site P, la synthèse de polypeptide peut commencer. Trois événements clés doivent se produire pour l'addition correcte des acides aminés.

- 1- La liaison de l'aa-ARNt au niveau du site A en concordance avec le codon qui est exposé.
- 2- La réaction de transpeptidation.
- 3- La Translocation Vers Le Site P.

Après la translocation (libération du site A) le ribosome est prêt pour un autre cycle de polymérisation. Deux protéines auxiliaires (facteurs d'élongation) contrôlent ces événements. Les deux facteurs utilisent la GTP comme source d'énergie pour accroître la vitesse et la précision du fonctionnement du ribosome. L'aminoacyl-ARNt est ramené au site A par le facteur d'élongation EF-Tu, ce facteur est lié à une GTP. L'activité GTPase est activée juste après l'établissement d'un appariement codon-anticodon correcte, et le EF-Tu-GDP sera libéré et converti en EF-Tu-GTP de nouveau par l'intermédiaire d'un deuxième facteur d'élongation appelé EF-Ts. Pour qu'une nouvelle élongation puisse se produire l'ARNt du site P doit se déplacer vers le site E et l'ARNt du site A doit se déplacer au site P. En même temps l'ARNm doit se déplacer de trois nucléotides pour présenter le codon suivant. Cette phase finale d'élongation nécessite l'EF-G ou translocase et l'hydrolyse de GTP.

Le chloramphénicol c'est un antibiotique inhibant la synthèse des protéines chez les bactéries. Il empêche le positionnement correcte de l'Amino-acyl-ARNt dans le site A pour la réaction peptidyl transférase dans la grande sous-unité.

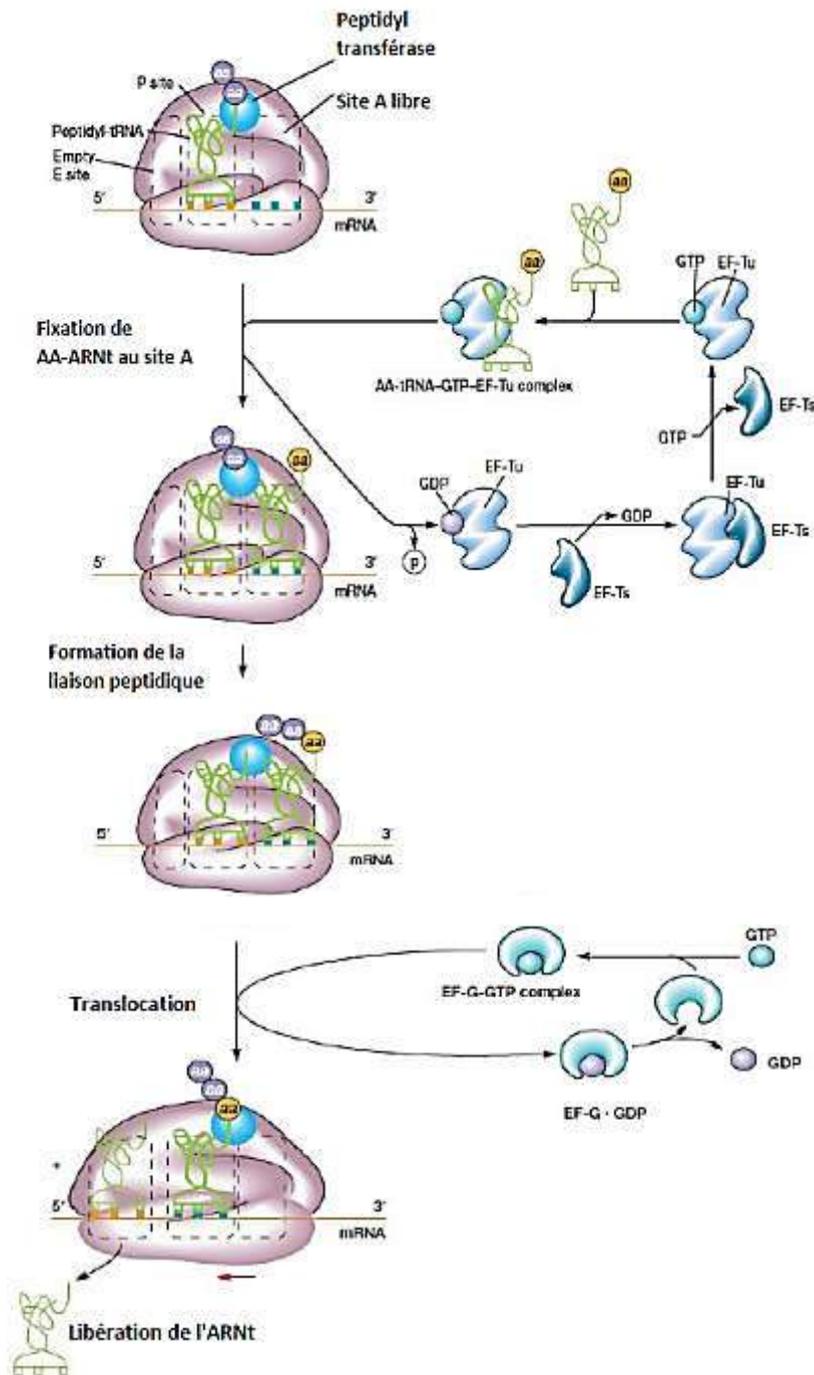


Figure 48: Etape d'élongation de la synthèse protéique.

C.Terminaison de la traduction

La terminaison de la traduction se fait quand un codon stop (non-sens: UAA, UAG, et UGA) est atteint, les ARNt ne pouvant plus s'y fixer. Trois facteurs de libération RF (RF-1, RF-2, et RF-3) aident le ribosome à reconnaître ces séquences et coupent le polypeptide attaché à l'ARNt terminal, libérant le produit fini. Par la suite les sous-unités se dissocient, elles peuvent alors former de nouveaux complexes d'initiation et recommencer le processus.

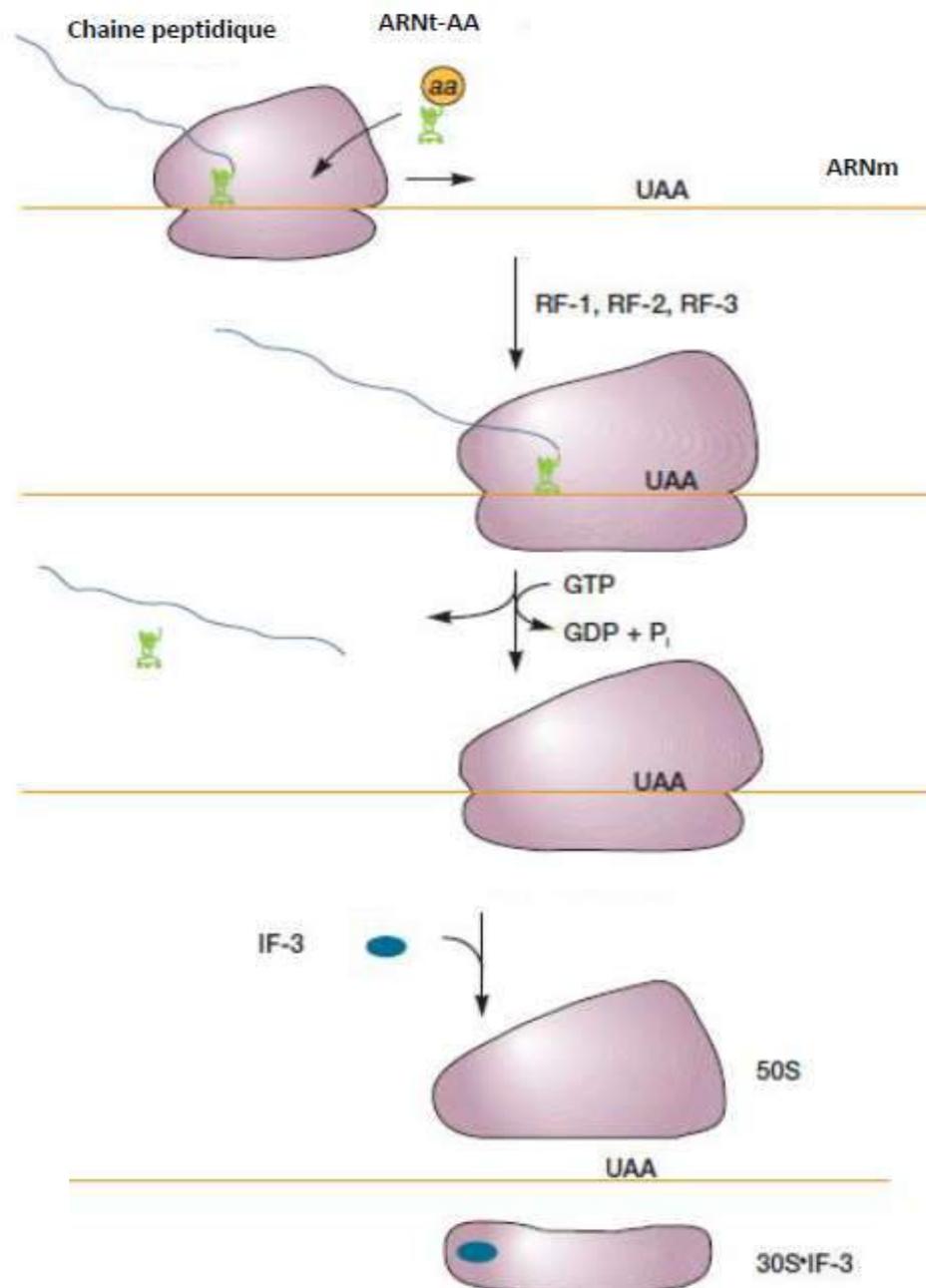


Figure 49 :Étape de terminaison de la traduction.

5.2.5 La traduction chez les eucaryotes

Comme chez les procaryotes on distingue trois étapes dans la synthèse protéique chez les eucaryotes (initiation, élongation et terminaison).

A. Initiation

L'initiation est similaire à celle chez les procaryotes par plusieurs aspects. Les deux utilisent un codon d'initiation et un ARNt initiateur, et les deux font appel à des facteurs d'initiation pour former un complexe avec la petite sous-unité ribosomique qui lie l'ARNm avant l'ajout de la grande sous-unité. Néanmoins, les eucaryotes utilisent une méthode complètement différente pour reconnaître l'ARNm et le codon d'initiation.

Chez les eucaryotes, la petite sous unité est déjà associée avec un ARNt initiateur. Au contraire des procaryotes la liaison de l'ARNt initiateur (chargé par une méthionine non modifiée) à la sous unité 40S précède toujours l'association avec l'ARNm. La reconnaissance de l'ARNm fait appel à plusieurs facteurs auxiliaires. Le ribosome recruté au niveau de la coiffe 5' de l'ARNm scanne l'ARNm dans le sens 5' → 3' jusqu'à atteindre la séquence de Kozak 5'CCAUGG3', cette séquence permet de connaître le codon d'initiation (AUG). Enfin, la grande sous unité du ribosome est recruté Après l'appariement de bases de l'ARNt initiateur avec le codon d'initiation. A la fin d'un cycle de traduction, le ribosome eucaryote se dissocie en ses sous-unités constitutives

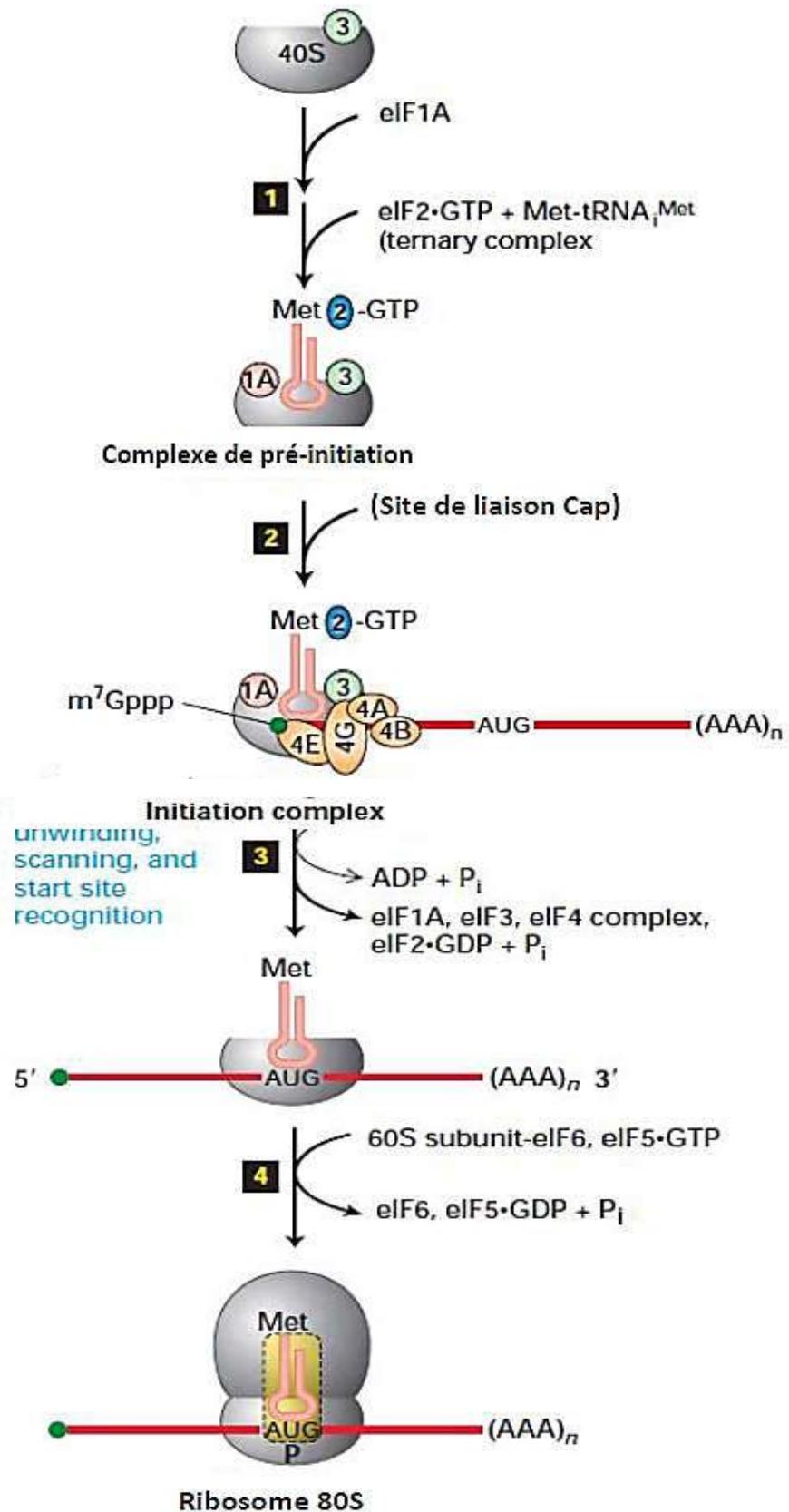


Figure 50 : Formation du complexe d'initiation chez les eucaryotes.

B.Élongation

Une fois la le ribosome assemblé, avec l'ARNt initiateur chargé au site P, la synthèse du polypeptide peut commencer.

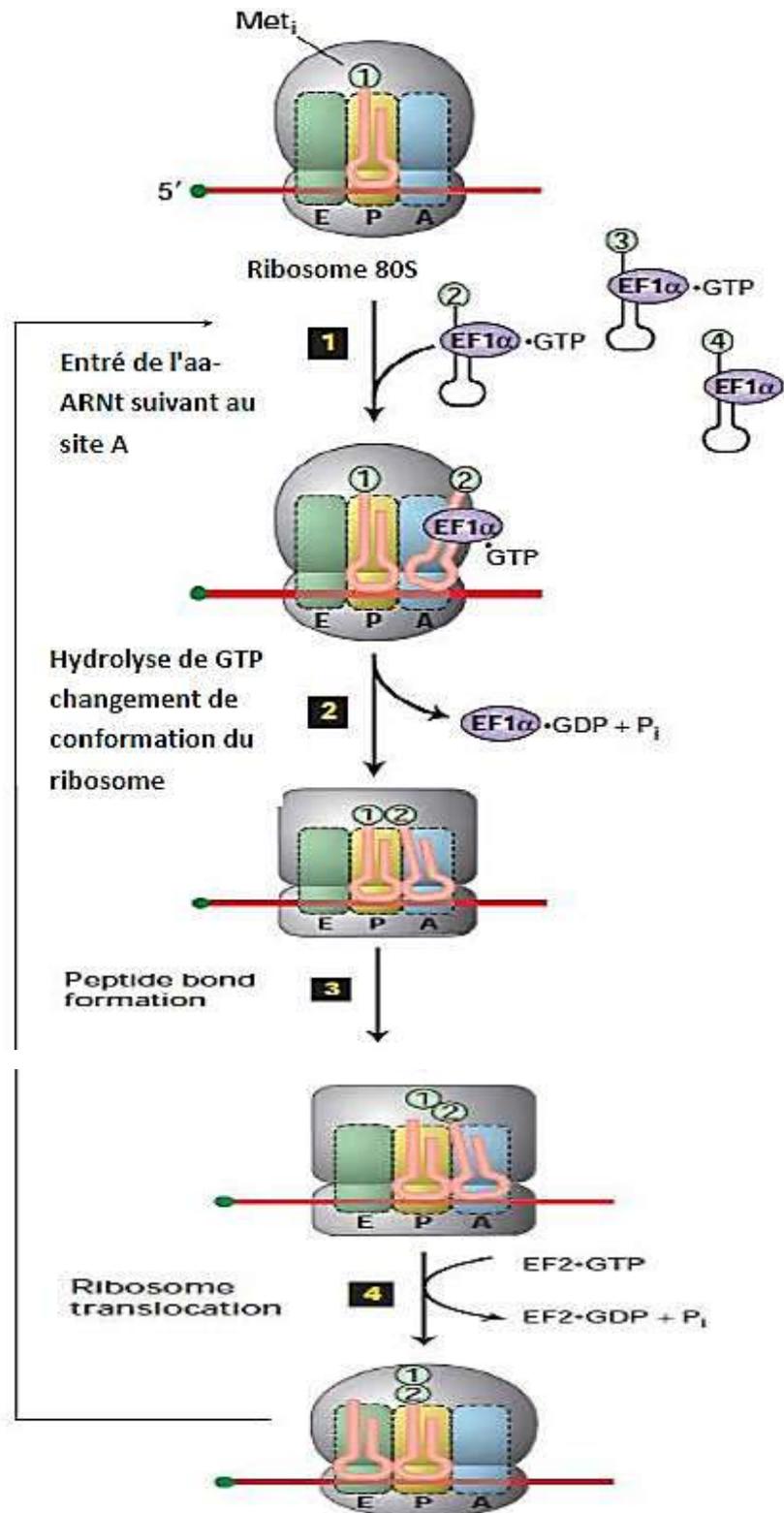


Figure 51 : Élongation de la traduction chez les eucaryotes.

C.Terminaison de la traduction

Tous comme l'initiation et l'élongation, la terminaison de la traduction est contrôlée par une série d'attachement-relâchement de facteurs dans un ordre précis.

Tableau 4 : Tableau comparatif de la traduction chez les procaryotes et les eucaryotes.

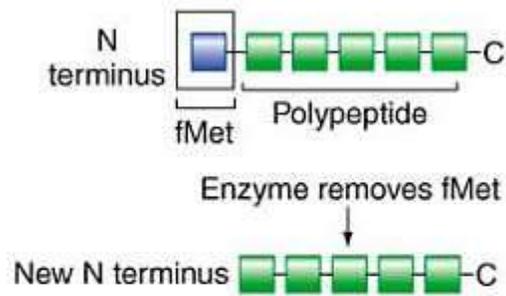
	Eucaryotes	Procaryotes
Ribosomes	80 s : 40 s + 60 s	70 s : 30 s + 50 s
ARNm	Coiffe en 5' monocistronique pas de séquence SD AUG comme codon initiateur queue polyA	pas de coiffe polycistronique séquence SD en -10 AUG ou GUG pour le codon initiateur pas de polyA en 3'
ARNt initiateur	aa non formylé	aa formylé
Initiation	eIF1 à eIF6	IF1 IF2 α pour AUG et IF2 pour GUG
Elongation	eEF1 α eEF1 β eEF2	EFTu -> liaison de l'aatRNA EFTs -> liaison de l'aatRNA EFG -> translocation
Terminaison	RF (release factor)	RF1 -> UAA et UAG RF2 -> UAA et UGA RF3

5.2.6 Modifications post-traductionnelles

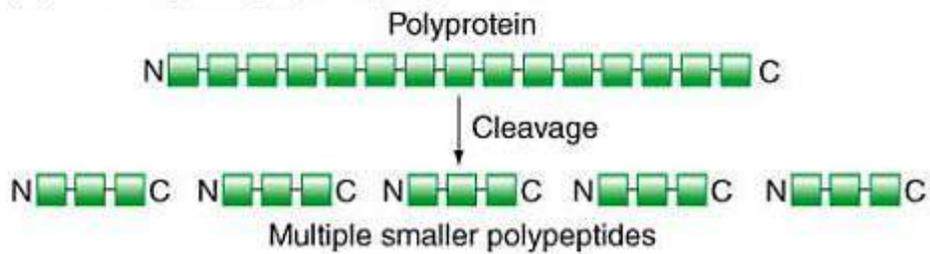
Les principales modifications post-traductionnelles sont :

- Phosphorylation (Ser, Thr, ou Tyr: important dans la transduction de signaux à travers les membranes.
- L'élimination du formyl-méthionine ou méthionine par une enzyme spécifique (Met-aminopeptidase : MAP) se fait chez plus de 50% des protéines des deux types de cellules. Le N-formyl de la première méthionine est enlevé chez *E. coli* par une deformylase.
- Clivage de chaîne polypeptidique (Ex. L'insuline)
- Glycosylation
- Addition des lipides
- Formation des ponts disulfures ou pont covalents (élastine, collagène).
- Acétylation (\approx 100 protéines cytoplasmique ont un N-terminal acétylé), la réaction est catalysée par N- α - acétyl transférase (exemple: acétylation des histones).
- Hydroxylation (hydroxylation des prolines et lysine)
- Biotinylation: Addition de la biotine [vit B8] (coenzyme de quelques enzymes de carboxylation [transfert de CO₂]).
- Élimination des séquences surnuméraires au niveau protéique, ce processus est appelé l'épissage protéique et le peptide éliminé est nommé intéine.

(a) Cleavage may remove an amino acid.



(b) Cleavage may split a polyprotein.



(c) Addition of chemical constituents may modify a protein.

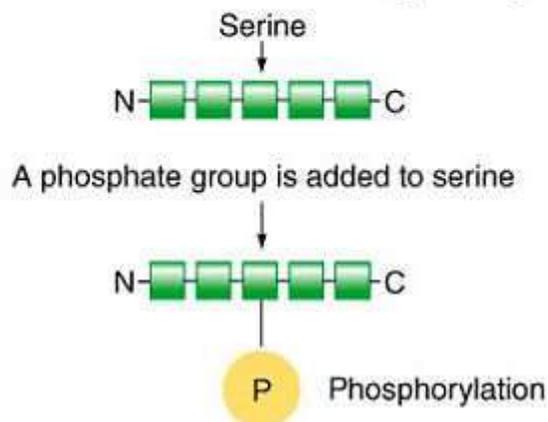


Figure 52 : Les principales modifications post-traductionnelles.

Bien que la protéine est constituée au départ d'une simple suite d'acides aminés, elle n'est fonctionnelle qu'après une phase de repliement qui lui donne sa structure tertiaire (tridimensionnelle).

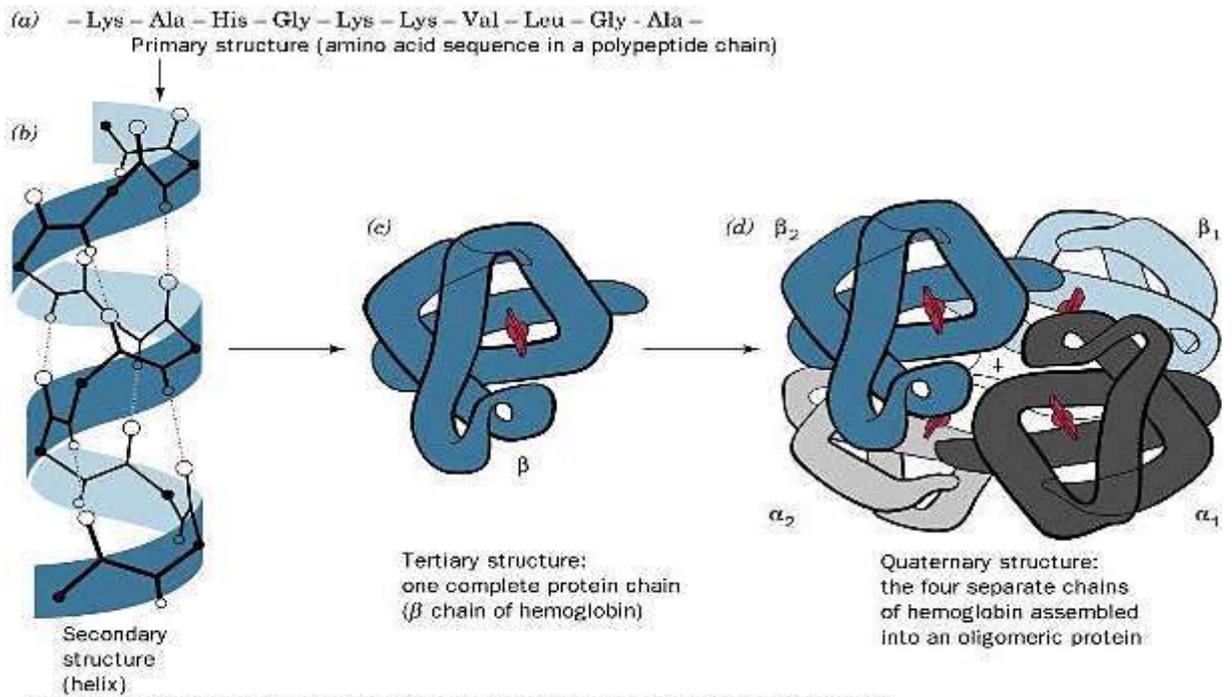


Figure 53 :Repliement proteique.

5.3 Régulation de l'expression génétique

La régulation de l'expression génétique c'est un processus d'intérêt majeur pour toutes les cellules, elle permet d'économiser l'énergie et les ressources. Les cellules microbiennes ont divers mécanismes par lesquels régulent leurs activités en réponse au changement de l'environnement externe, afin d'optimiser ses fonctions principales (nutrition, croissance, résistance...etc.).

Il y a deux modes majeurs de régulation dans les cellules:

- 1- Régulation de l'activité des enzymes présentes (post-traductionnelle).
- 2- Régulation de la production des enzymes (transcriptionnelle ou post-transcriptionnelle).

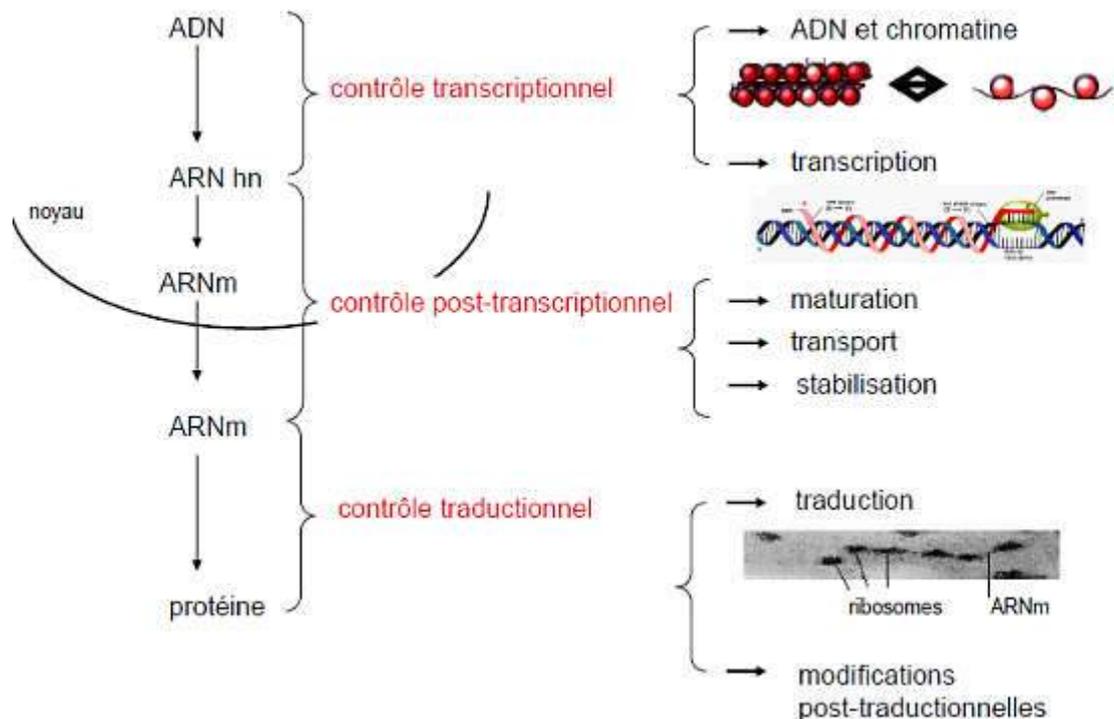


Figure 54: Niveaux de régulation de l'expression.

La régulation de l'activité enzymatique s'effectue par deux moyens principaux :

- 1- Le contrôle de la quantité d'enzyme : Des systèmes complexes modulent la quantité d'enzyme en réponse à des conditions changeantes (Exemple le cas des enzymes inductibles).

- 2- Le contrôle de l'activité enzymatique: L'activité enzymatique peut être directement modifiée par des modifications covalentes et conformationnelles, ou par des modifications non covalentes en modulant l'affinité des enzymes au substrat par des molécules ou sous-unités protéiques régulatrices.

5.3.1 Régulation de la transcription

L'expression d'un gène peut être régulée à plusieurs étapes, la plus commune étant celle de l'initiation de la transcription. Il ya deux raison qui peuvent justifier cela.

- Le coût en énergie et en ressources pour la transcription
- La régulation lors de l'initiation est plus facilement réalisable.

Il y a 2 modes distincts du contrôle de l' initiation de la transcription

- 1- Un contrôle **constitutif** qui dépend de la **structure du promoteur**
- 2- Un contrôle de **régulation** qui est sous la dépendance de **protéines régulatrices**

Il ya deux types de protéines régulatrices :

- 1- Les activateurs (régulateurs positifs: augmente la transcription d'un gène)
- 2- Les répresseurs (régulateurs négatifs: diminue ou bloque sa transcription)

A. Contrôle négatif de la transcription "Répression et induction"

Répression: Activité de synthèse d'une enzyme empêchée lorsque le produit de la réaction est présent en excès.

Induction: Production d'une enzyme seulement lorsque son substrat est présent.

Opéron : c'est un ensemble de gènes transcrit en un seul ARNm polycistronique sous la direction d'un seul promoteur (le mot cistron désigne la plus petite unité génétique servant à une seule fonction). Les gènes d'un opéron donné sont soumis à un contrôle coordonné (système de contrôle remarquablement économique).

Lorsque *E. coli* est cultivé dans un milieu de culture où le lactose est la seule source de carbone, la concentration cellulaire de trois protéines augmente.

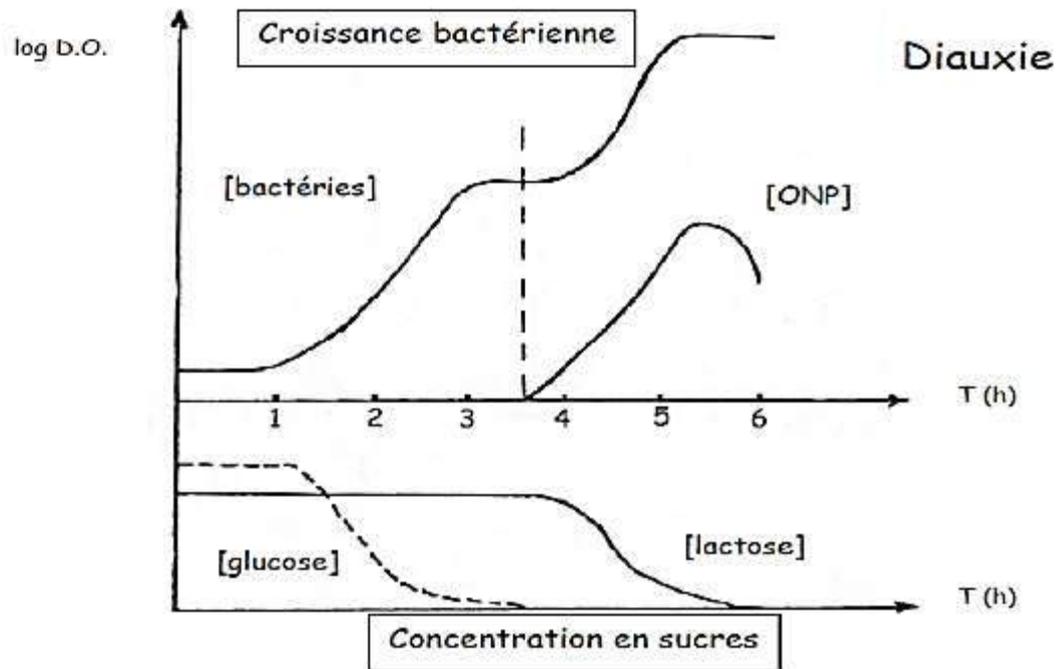


Figure 55 : Croissance bactérienne en présence de glucose et de lactose.

Les gènes de l'opéron *lac* ont un fort niveau d'expression seulement en présence du lactose, et en absence du glucose (source de carbone et d'énergie préférée). Deux protéines régulatrices sont impliquées:

- Un activateur appelé **CAP** (Catabolite Activator Protein), connue aussi par le nom de CRP (cAMP receptor protein). Le gène codant CAP au contraire du ce codant le répresseur est localisé sur un autre endroit sur le chromosome. Il se fixe sur le site CAP.
- Le répresseur *lac* (codé par le gène *lac I*), il se fixe sur l'opérateur.

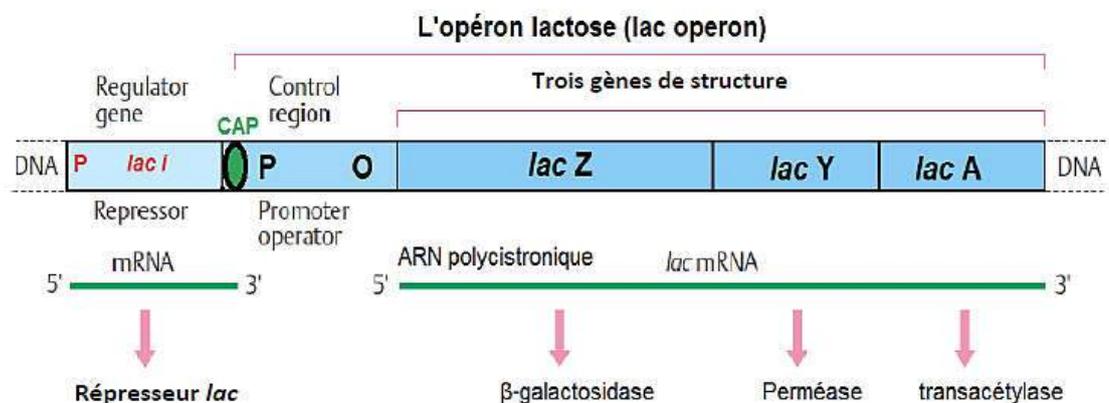


Figure 56: Organisation de l'opéron lactose d'*E. coli*.

Le blocage de l'expression de l'opéron lactose en présence du glucose ou autre source d'énergie et de carbone facilement assimilable que le lactose est appelé la répression catabolique. L'une des conséquences de cette répression est la croissance diauxique ou triauxique (selon le nombre).

L'opéron *lac* ne possède pas un promoteur fort (faible séquences consensus surtout la région -35) pour cette raison la fixation de l'ARN polymérase doivent être stimulée par une protéine d'activation spécifique (contrôle positif), appelé protéine réceptrice d'AMPc (CRP) ou protéine activatrice du catabolisme (CAP).

En absence du glucose, les niveaux d'AMPc (synthétisée à partir de l'ATP par l'adénylate cyclase) augmentent, aboutissant à la fixation de la CRP à l'AMPc. Le complexe CRP-AMPc se fixe sur le site CAP. L'ADN subit alors une inclinaison de 90°, qui stimule la fixation des éléments *trans* de l'ARN polymérase aux éléments *cis* du promoteur pour augmenter le taux de la transcription de 50 fois.

En absence du glucose et la présence du lactose, l'allolactose(β -D-galactopyranosyl-(1-6)- β -glucopyranose), une molécule inductrice synthétisée à partir du lactose (β , 1-4) par un réaction de transglycosylation se fixe sur le répresseur et l'inactive de sorte qu'il ne puisse plus se fixer à la séquence de l'opérateur. L'ARN polymérase activée par le complexe CRP-AMPc peut initier la synthèse de l'ARN avec un taux de transcription élevé.

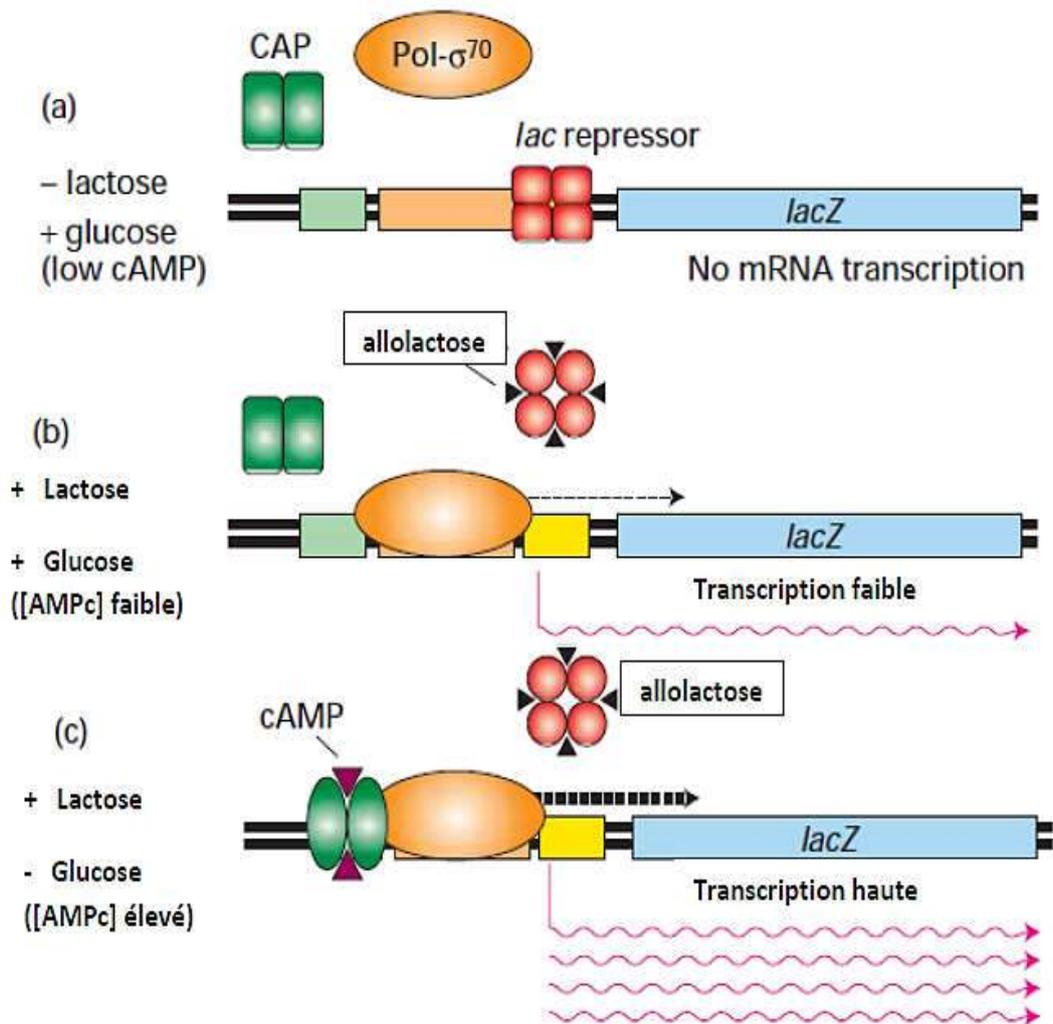
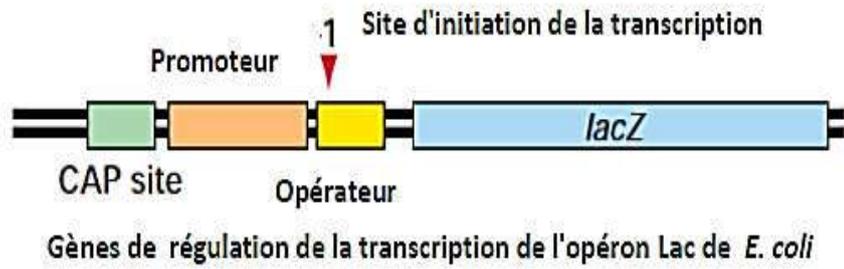


Figure 57 : Expression des gènes *lac* en présence/absence du glucose et le lactose.

B. Contrôle positif de la transcription

Dans ce type de contrôle, une protéine régulatrice active à la fixation de l'ARN polymérase d'où le terme positif. Un excellent exemple de ce type de contrôle est l'opéron *malEFG* (l'opéron du catabolisme du maltose) chez *E. coli*.

En absence d'inducteur, ni la protéine activatrice, ni l'ARN polymérase ne peuvent se lier à l'ADN. La liaison d'une molécule inductrice dans ce cas là c'est le maltose à la protéine activatrice provoque un changement de sa conformation (protéine allostérique), à son tour, se lie au site de fixation de l'activateur. Ceci permet à l'ARN polymérase de se fixer sur le promoteur et commencer la transcription

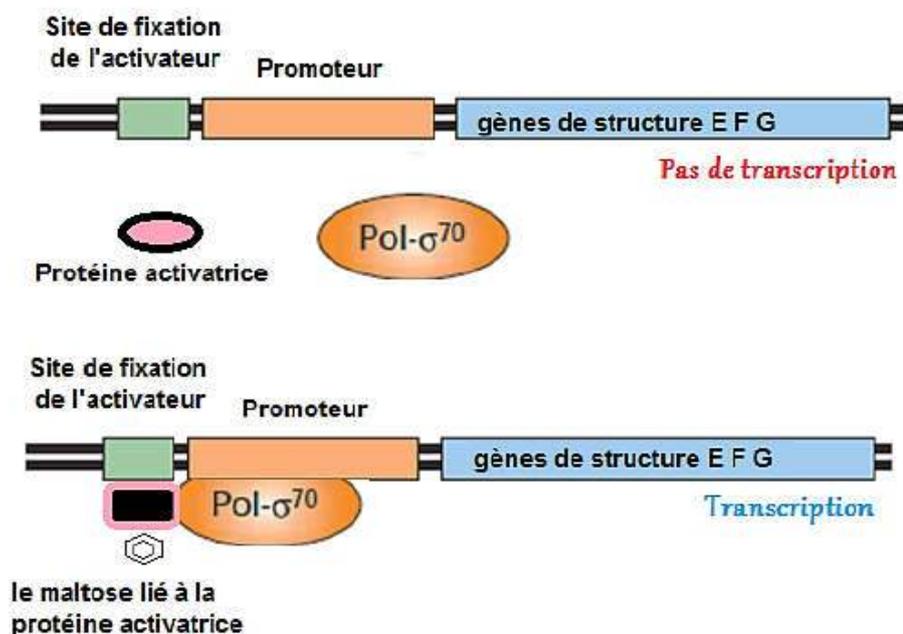


Figure 58: Mécanisme de contrôle positif de l'opéron *malEFG* chez *E. coli*.

C. Contrôle de l'expression génétique par atténuation

L'atténuation dépend du fait que la transcription et la traduction sont étroitement couplées. Chez *E. coli* les opérons de biosynthèse des acides aminés (Ex. Tryptophane) sont contrôlés par ce mécanisme. L'opéron Trp est composé de cinq gènes adjacents codant les enzymes qui synthétisent le tryptophane. Ces gènes sont contrôlés par un répresseur et ne s'expriment qu'à l'épuisement de tryptophane (l'absence du tryptophane qui arrête la répression).

Le phénomène d'atténuation contrôle la décision de transcrire un ARNm entier. Si les taux de tryptophane sont élevés, la plupart des transcrits terminent leur synthèse de façon prématurée (incomplète). Au contraire, si les taux sont insuffisants, la polymérase transcrit les gènes en entier.

Une séquence leader de 161 nucléotides localisés en amont du premier codon du gène *trpE*, elle contient 4 régions (1, 2, 3, et 4). Cette séquence contient aussi deux codons pour le tryptophane, alors s'il ya manque de tryptophane la synthèse du peptide leader sera incomplet à cause du manque des ARNt chargés par le Trp.

L'atténuation (l'arrêt de la transcription) dépend de la capacité de l'ARN à former des structures secondaires entre des régions complémentaires (2=3 ou 3=4).

Si le tryptophane est abondant, le ribosome traduira la séquence leader jusqu'au codon stop, le reste de l'ARNm leader forme une structure secondaire entre les régions 3 et 4 suivi par une séquence riche en uracile qui provoque finalement la terminaison.

Si le tryptophane est absent ou en faible quantité, la synthèse du peptide leader sera incomplet. Le blocage du ribosome permet dans ce cas là la formation d'une autre structure secondaire (entre les régions 2 et 3), cette structure n'est pas un signal de terminaison de la transcription et empêche la formation du signal de terminaison. Ceci permet à l'ARN polymérase de transcrire tous les gènes de structure.

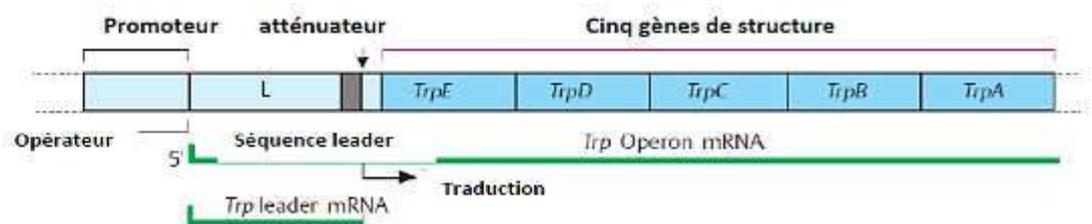


Figure 59 :L'opéron tryptophane chez *E. coli*.

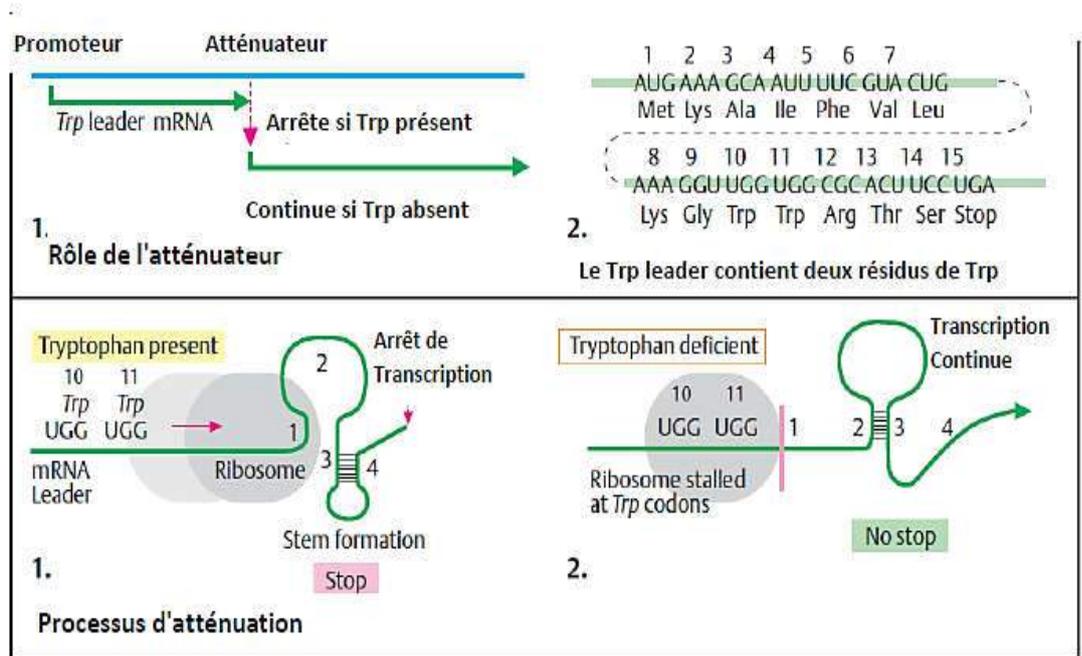


Figure 60 : Processus de l'atténuation de l'opéron tryptophane d'*E. coli*

6. Techniques de base de biologie moléculaire

6.1 Extraction et purification des acides nucléiques

Toute étude de génétique moléculaire implique la disposition d'échantillon d'acides nucléiques. Les techniques d'extraction d'acides nucléiques sont relativement simples. Il convient simplement d'éviter toute destruction enzymatique ou mécanique. En effet les acides nucléiques qui sont stables dans la cellule intacte, deviennent très vulnérables à la digestion par les nucléases endogènes une fois la cellule lysée.

Le protocole d'extraction d'ADN génomique à partir d'une culture bactérienne est basé sur le principe de la lyse alcaline. Après extraction, l'ADN est resuspendu dans 30 µl de TE et conservé à -20°C. L'ADN génomique est analysé par électrophorèse sur gel d'agarose à 1% (w/v). Le protocole d'extraction d'ADN est le suivant :

1. Transférer 1 ml d'une préculture d'une nuit à 37°C dans un tube eppendorf
2. Centrifuger 1 min à 13000 rpm
3. Eliminer le surnageant
4. Resuspendre le culot dans 500 µl d'une solution TE (10 mM Tris-HCl pH 8,1mM EDTA pH 8)
5. Ajouter 20 µl lysozyme (50 mg / ml) et incuber 30min à 37 °C
6. Ajouter 60 µl d'EDTA (0,5M pH 8)
7. Ajouter 10 µl Protéinase K (20mg/ ml) et incuber 1 h à 37°C
8. Ajouter 100 µl SDS 10% et incuber 10 min à 50°C
9. Ajouter 350 µl d'acétate de potassium 3M pH 5,2
10. Mélanger délicatement par inversion du tube en incuber 10 min dans la glace
11. Centrifuger 10 min à 10000rpm
12. Transférer le surnagent dans un tube eppendorf neuf
13. Ajouter 1 volume d'isopropanol
14. Mélanger par inversion (jusqu'à apparition de l'ADN sous forme de filament) et laisser précipiter pendant 10 min à - 20 °C
15. Centrifuger 1min à 10000 rpm
16. Eliminer le surnageant
17. Ajouter 200 µl d'éthanol 70% pour laver l'ADN, sans aucune agitation
18. Centrifuger 10 min à 10000 rpm
19. Eliminer le surnageant

20. Sécher le culot en inversant le tube sur un papier filtre pendant 10 min

21. Dissoudre le culot dans 20 μ l TE.

6.2 Dosage des acides nucléiques

Les pyrimidines et les purines absorbent fortement les UV à 260nm. Une unité de densité optique à 260 nm correspond à:

- Une solution de DNA double brin à 50 μ g/ml.

- Une solution de DNA simple brin ou RNA à 25 μ g/ml.

$$C = A_{260} * DF * 100$$

Ces valeurs s'appliquent à des acides nucléiques parfaitement purs et en solution homogène. Pour vérifier la pureté de l'ADN il faut calculer : P (pureté) = A_{260} / A_{280} Une solution d'ADN est considérée pure si : $1.7 \leq P \leq 2$

6.3 Electrophorèse

Bien que la séquence des acides nucléiques détermine ses fonctions biologiques, la longueur de ces polymères constitue leur propriété physique la plus utile. Parmi les techniques les plus utilisées pour la séparation des acides nucléiques selon leur taille est l'électrophorèse.

C'est une technique de bioséparation permettant la séparation des biomolécules selon leur charge et leur taille (ADN, ARN, protéines). Le terme électrophorèse vient de : électroce qui signifie énergie électrique et de phoresis (photo en grec) qui veut dire porter, avoir en soi.

Les acides nucléiques en solution portent une charge négative à cause de l'ionisation de leurs groupements phosphate (macromolécules polyanioniques uniformément chargées). Ils se déplacent donc vers l'électrode (+). Toutes les séquences d'acides nucléiques ont un rapport charge/masse à peu près identique, quelle que soit leur longueur, parce que tous les nucléotides sont à peu près de même masse (masse molaire moléculaire moyenne des nucléotides 330 g/mol).

Suivant les théories de l'électrophorèse la mobilité μ dans un champ électrique au sein d'un gel doué de pouvoir de filtration est:

$$\text{Log } \mu = \text{Log } \mu_0 - \text{Kr} \cdot C$$

Log μ_0 : La mobilité de la molécule en milieu liquide

C: Concentration du gel

Kr: Coefficient de retardement dû au gel (Lui même est en fonction de la masse moléculaire de la molécule).

Deux sortes de matrice de gel sont utilisées: l'agarose et le polyacrylamide. Le gel agit à la manière d'un tamis moléculaire au travers duquel les molécules d'ADN en mouvement

doivent passer; les molécules de grande taille ont plus de difficulté pour passer à travers les mailles (pores) créés par le réseau des microfibrilles du gel et vont donc migrer plus lentement que les petites molécules. La vitesse de migration de l'ADN est influencée par deux paramètres :

- La masse moléculaire (nombre de pb)
- La concentration du gel

Il faut noter que la nature et la concentration du support de l'électrophorèse sont en fonction de la taille des fragments à séparer.

Tableau 5: Taille des fragments à séparer selon la concentration du gel.

% d'agarose	Taille de fragments à séparer en kb	% d'acrylamide	Taille de fragments à séparer en pb
0.5	1 à 30	4	200 à 800
0.7	0.8 à 12	5	80 à 200
1	0.5 à 10	8	40 à 100
1.2	0.4 à 7	11	10 à 50
1.5	0.2 à 3		

Dans toutes les électrophorèses, des marqueurs de taille (ou de poids moléculaire) sont déposés et migrent parallèlement au DNA étudié. Une fois l'électrophorèse terminée, les molécules fluorophores, comme le bromure d'éthidium, qui se fixe à l'ADN en s'intercalant entre les bases. Les bandes d'ADN (regroupent toute les molécules d'ADN de taille identique) sont visualisées sous UV.

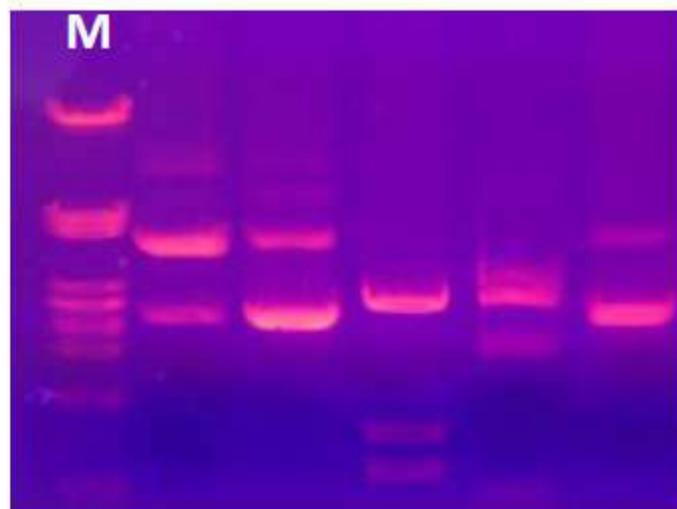


Figure 62: Visualisation des bandes d'ADN sous UV. **M:** marqueur de taille.

La mobilité relative est le rapport entre la distance de migration d'une bande et la distance de migration du front de migration. La droite $\log(\text{nombre de kb}) = f(\text{mobilité relative})$, permet de déterminer la taille d'une molécule d'ADN inconnue.

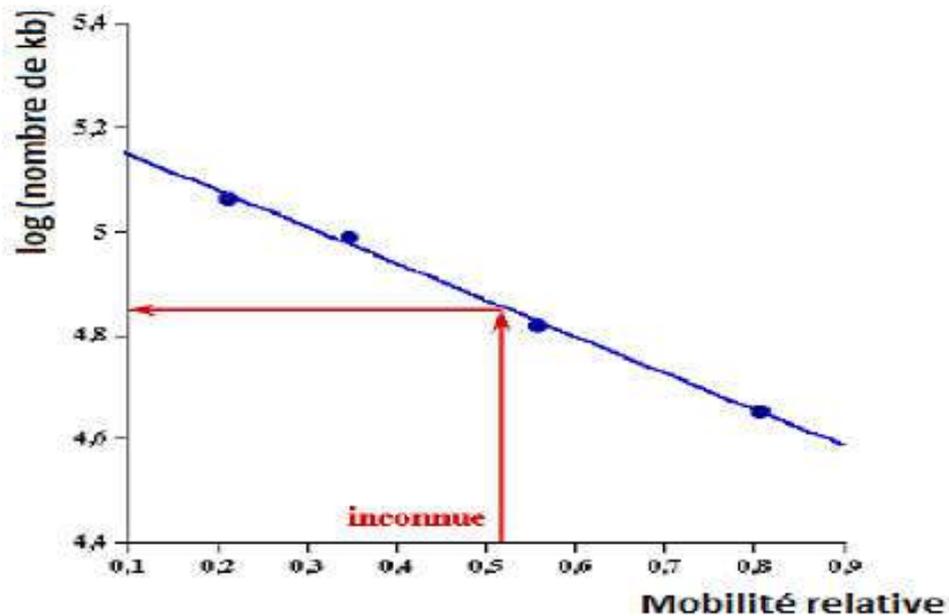


Figure 63: Distance de migration en fonction du logarithme de la taille des fragments d'ADN.

6.4 Marquages des acides nucléiques

Il existe deux grands types de marquages :

Les marquages chauds: le signal est lié à la radioactivité des sondes qui contiennent un ou plusieurs atomes radioactifs de courte période comme le phosphore 32 (P_{32}) ou le soufre 35 (S_{35}) ou le tritium (H_3).

Les marquages froids qui sont de deux types: soit la sonde est liée à un fluorochrome, soit à un ligand qui sera reconnu spécifiquement et avec une forte affinité par une molécule réceptrice.

5.4.1 Marquage des sondes : Il existe en fait plusieurs stratégies pour marquer les sondes.

On peut distinguer :

- Nick translation
- Multi-amorçage au hasard
- Marquage des oligonucléotides

NB : Une sonde nucléotidique est un segment de nucléotides qui permet de rechercher de manière spécifique un fragment d'acide nucléique que l'on désire étudier. Cette réaction

sonde-fragment correspond à une réaction d'hybridation moléculaire

a- La Nick translation : La DNase I coupe aléatoirement les brins de l'ADN bicaténaire.

Les extrémités 3' ainsi libérées deviennent un site de fixation pour l'ADN polymérase I qui va :

1. Détruire l'ADN par son activité exonucléasique (dans le sens 5'→3')
2. Resynthétiser (activité polymérasique) une nouvelle chaîne avec les nucléotides dont un au moins est radioactifs présents dans le milieu d'incubation enzymatique

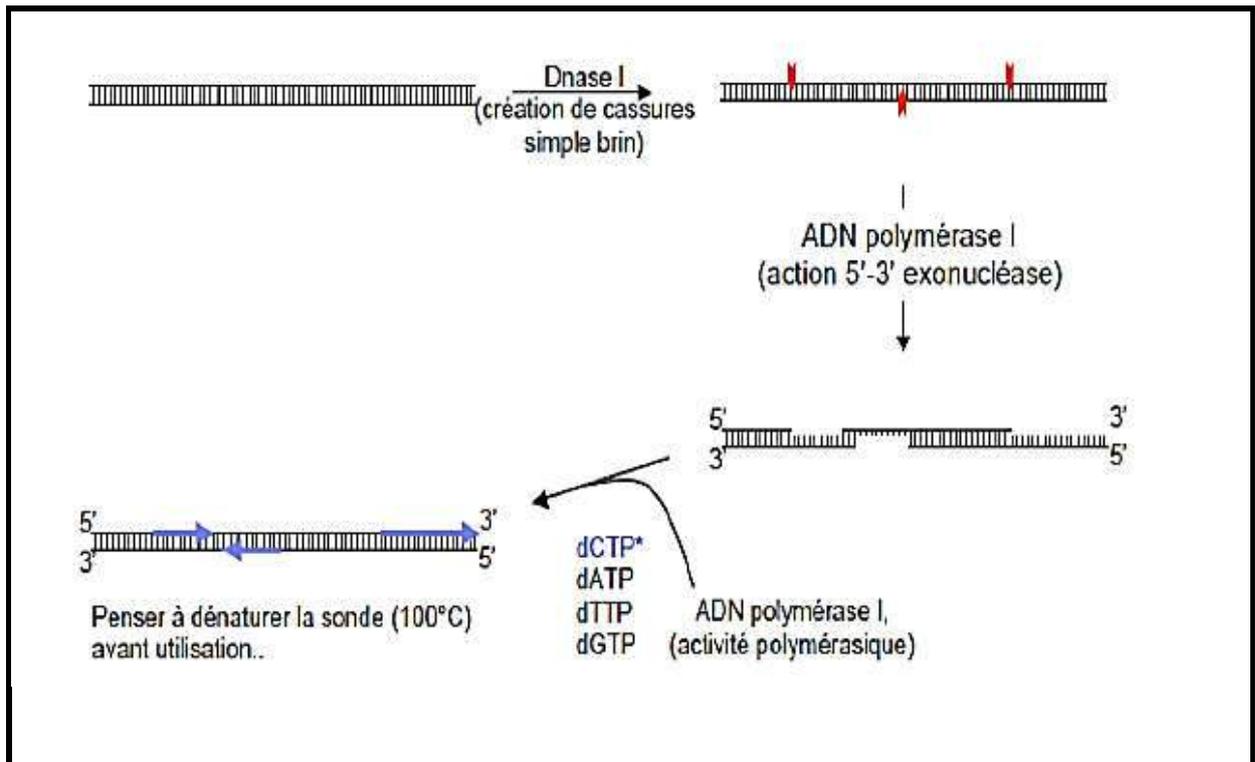


Figure 64 : Nick Translation.

b- Le multi-amorçage au hasard ou multirandom priming :

L'ADN bicaténaire est dénaturé par chauffage à 100°C afin de séparer les deux brins, puis refroidi brutalement pour empêcher que les deux brins ne se réassocient. Chacun des deux brins est hybridé à 25°C au hasard avec des petites amorces (6 nucléotides). Les séquences de ces hexanucléotides sont différentes. C'est une combinaison de 4 nucléotides parmi 6 positions différentes ; ce qui donne 1024 séquences possibles et il y aura certainement quelques hexanucléotides qui s'hybrident avec les deux d'ADN. L'enzyme ADN polymérase I, grâce à son fragment de Klenow, assure la synthèse des deux brins complémentaires dans le sens 5'→3' en utilisant les nucléotides triphosphates (ATP, CTP, TTP et GTP) dont un est marqué. Les brins d'ADN néosynthétisés sont alors marqués.

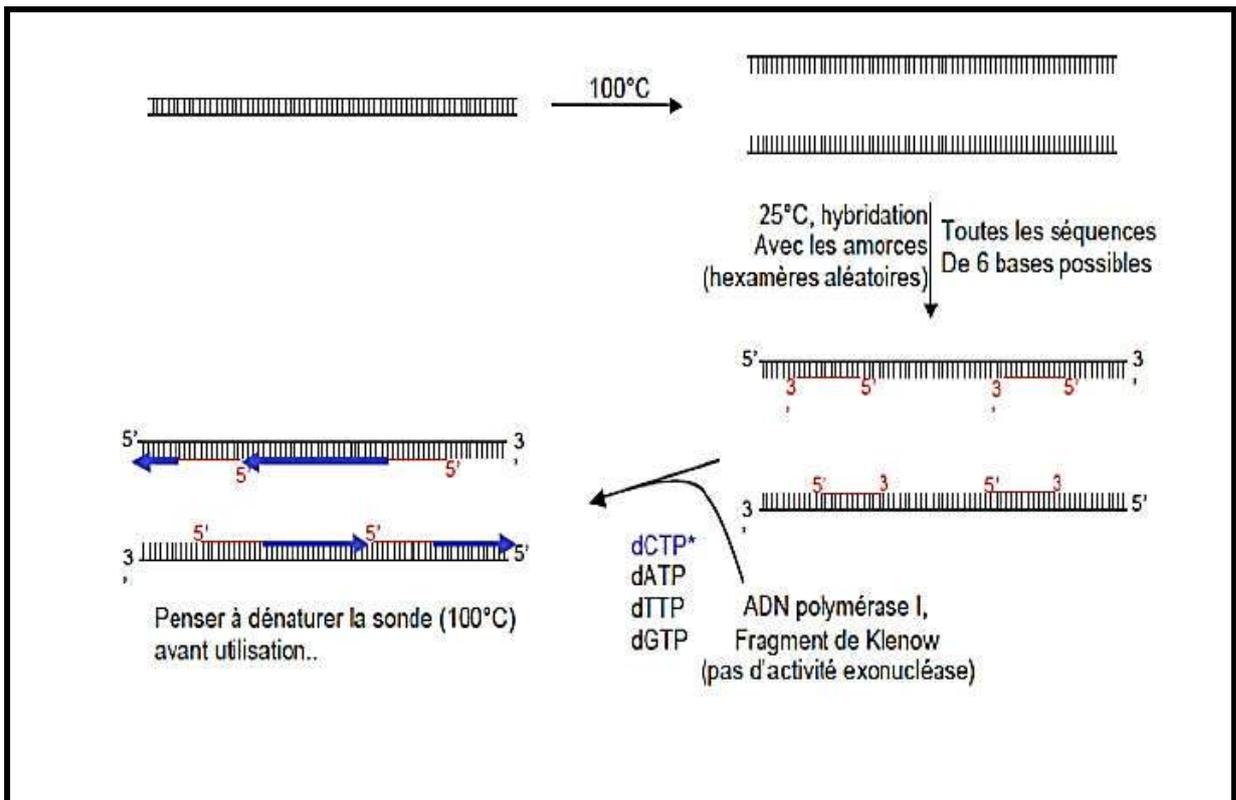


Figure 65: Amorçage au hasard.

c- Marquage des oligonucléotides

A partir d'une extrémité 5'OH on peut ajouter un phosphate marqué à l'aide d'une kinase et de λ ^{32}P ATP. Il faudra au préalable obtenir l'acide nucléique avec un 5'OH.

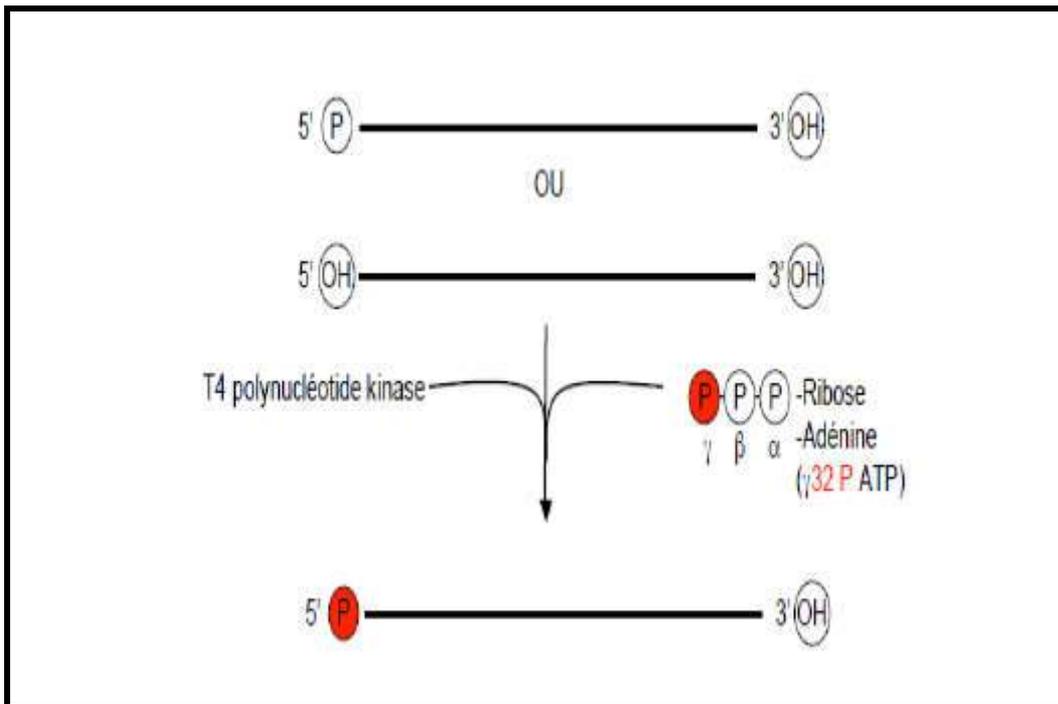


Figure 66 : Marquage d'oligonucléotide synthétique.

6.5 Hybridation des acides nucléiques

L'hybridation est une propriété fondamentale des acides nucléiques qui repose sur les règles de complémentarité. Il est possible d'apparier des brins d'ADN ou ARN avec des oligonucléotides qui reconnaissent spécifiquement des séquences sur les brins d'ADN de manière antiparallèle et complémentaire. Ces oligonucléotides sont appelés sondes nucléiques. Pour le cas de deux brins complémentaires de la même molécule d'ADN (brins homologues) on parle de renaturation (hybridation à 100%). Chaque 1% de non homologie diminue la température d'hybridation (T_h) de 1°C.

L'hybridation moléculaire est utilisée surtout dans la détection de l'homologie entre les molécules d'ADN de sources différentes. La complémentarité dépendra de de la concentration en ADN et du temps. Ces deux derniers paramètres sont importants pour la rencontre entre les deux brins.

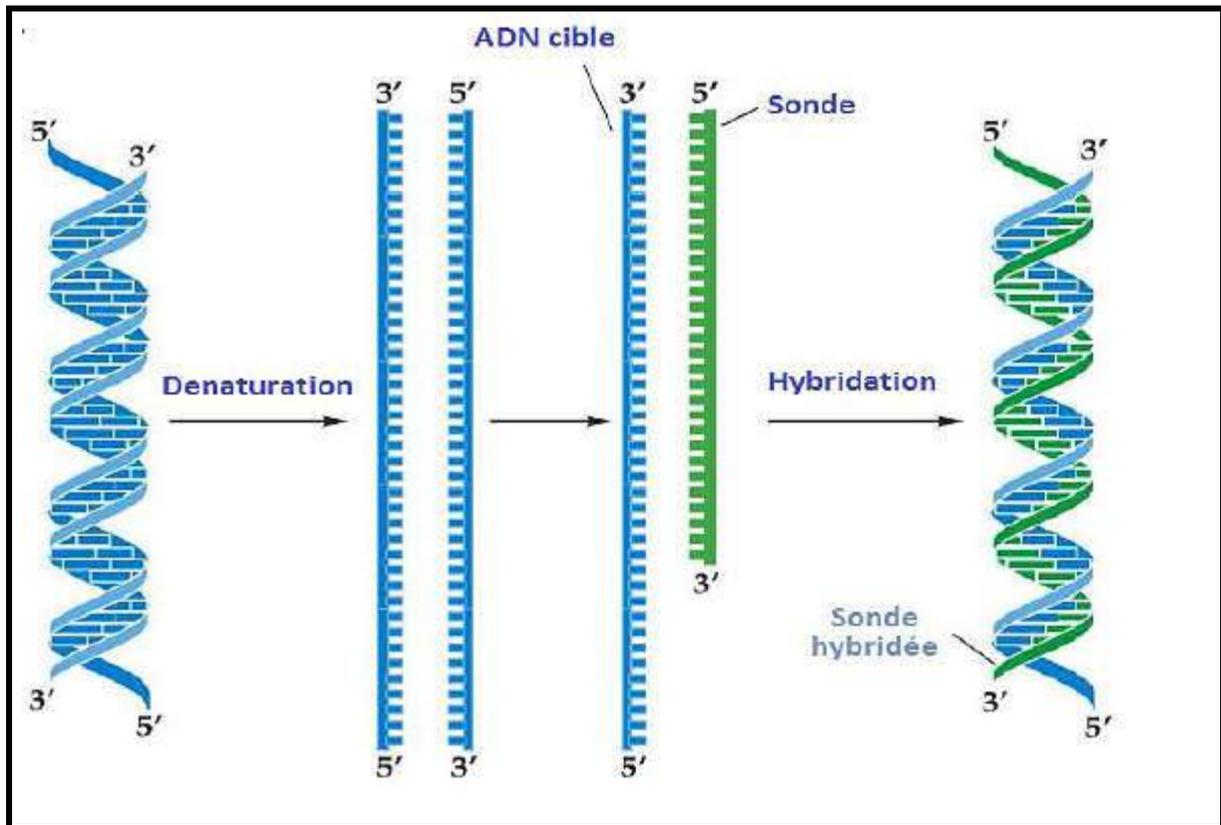


Figure 67: Hybridation des acides nucléiques.

La **T_m** est estimée en utilisant plusieurs formules, cela dépend de la:

- Longueur du fragment (nombre de nucléotides).
- Composition du milieu
- Mis-appariement
- Le contenu en C+G

La formule de Wallace permet de déduire la **T_m** des petits oligonucléotides (Ex. Les amorces).

Formule de Wallace: **$T_m = 4(C+G) + 2(A+T)$** ; **$T_h = T_m - 5$**

Pour les oligonucléotides avec $N < 100$ nucléotides on utilise la formule suivante en tenant compte la concentration du sel, les mésappariements et la concentration de formamide:

$T_m = 16.6 \log[Na^+] + 0.41(\%(C+G)) + 81.5 - (675/N) - \% \text{mismatch} - 0.65\% \text{ de formamide}$

Dans les conditions réactionnelles standards la **T_m** pour les longs séquences (>100):

$T_m = 69.3 + 0.41(\%(C+G))$

La formule globale:

$$T_m = 81.5 + 16.6 \log[\text{sel}] + 0.41[\%(\text{C}+\text{G})] - \% \text{ mis-appariement} - (500/N) - 0.65\%(\text{formamide}).$$

Facteurs influençant l'hybridation

- La concentration de l'ADN et le temps d'hybridation
- La température d'hybridation
- La force ionique
- La complexité des séquences

6.5.1 Types d'hybridation

L'objectif de l'hybridation est la détection de la présence d'un acide nucléique d'une séquence donnée par l'utilisation d'un fragment d'ADN complémentaire = sonde. Cela peut avoir lieu en solution ou sur support solide (immobilisation de la cible sur une membrane[nitrocellulose, nylon], sur verre, colonies bactériennes, chromosomes, plage de lyse ...etc.).

a. Hybridation d'ADN sur support solide : Southern blot

La procédure de ce type d'hybridation est résumée en sept étapes. On peut aussi révéler la présence des séquences spécifiques dans les ARN (Northern Blot) et des protéines (Western Blot) utilisant le même principe.

- 1– Extraction de l'ADN
- 2– Digestion enzymatique (par les enzymes de restriction)
- 3– Electrophorèse du produit de la digestion
- 4– Transfert sur membrane
- 5– Hybridation avec une sonde spécifique marquée.
- 6– Lavages
- 7– Révélation

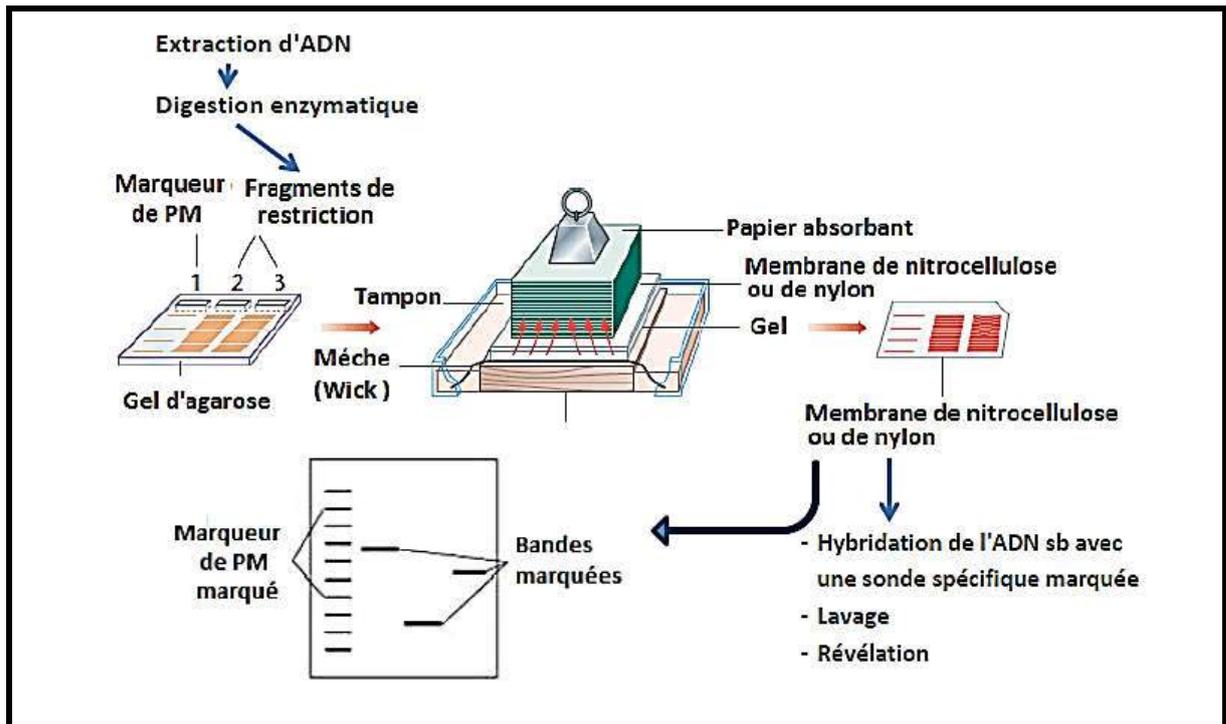


Figure 58: Southern Blot.

b. Hybridation *in situ* (HIS)

On appelle hybridation *in situ* (HIS) l'utilisation de sondes d'acides nucléiques pour mettre en évidence et localiser, dans des cellules ou des tissus, des séquences d'acides nucléiques, complémentaires de la sonde par leurs bases. L'HIS est un outil incomparable pour étudier l'expression des gènes. Elle est très proche, dans son principe, des Southern et des Northern blots et repose, comme eux, sur l'hybridation d'une sonde d'acide nucléique (ADN ou ARN) marquée avec une séquence complémentaire d'acides nucléiques que l'on cherche à identifier et à localiser. Mais les Southern et Northern blot se font sur des broyats de tissus, alors que l'HIS s'effectue sur une coupe histologique de tissu, apportant ainsi des informations précises sur la localisation des acides nucléiques étudiés.

c. Hybridation sur chromosomes (FISH)

La FISH (Fluorescence *in Situ* Hybridization) repose sur la capacité d'hybridation de deux brins d'ADN complémentaires. La région à étudier (située sur un chromosome préalablement légèrement dénaturé par traitement chimique pour le débarrasser des protéines associées) est repérée grâce à une sonde oligonucléotidique complémentaire. Certains de ces

nucléotides de cette sonde sont couplés à une molécule antigénique reconnue par un anticorps fluorescent.

En utilisant diverses sondes, greffées à des antigènes différents, on peut ainsi visualiser simultanément plusieurs séquences sur un ou plusieurs chromosomes. Technique permettant de déterminer la position d'un fragment d'ADN dans le génome ; le repérage se fait par rapport au bras du chromosome (p:bras court et q:bras long) et par rapport aux bandes (mises en évidence par la coloration Giemsa) du chromosome.

d. Hybridation sur colonie de bactéries, sur plage de lyse

Ce type d'hybridation permet la détection parmi un grand nombre de bactéries ou de phages recombinants (plage de lyse) celle ou celui qui contient le fragment d'ADN cible. La réalisation de cette technique passe par les étapes suivantes:

- Culture pure sur boîte de Petri (colonies ou plage de lyse)
- Transfert sur membrane de nylon ou de nitrocellulose
- Lyse alcaline des bactéries et dénaturation de l'ADN
- Fixation par la chaleur ou les UV
- Hybridation avec une sonde marquée
- Lavage
- Autoradiographie
- Localisation des clones d'intérêt.

e. Polymorphisme de longueur des fragments de restriction de l'ADN (RFLP)

La technique RFLP repose sur la digestion d'un DNA cible par une ou plusieurs enzymes de restriction. Ces fragments de restriction obtenus sont ensuite séparés selon leur longueur, par électrophorèse sur gel d'agarose et les fragments séparés sont hybridés avec un DNA sonde.

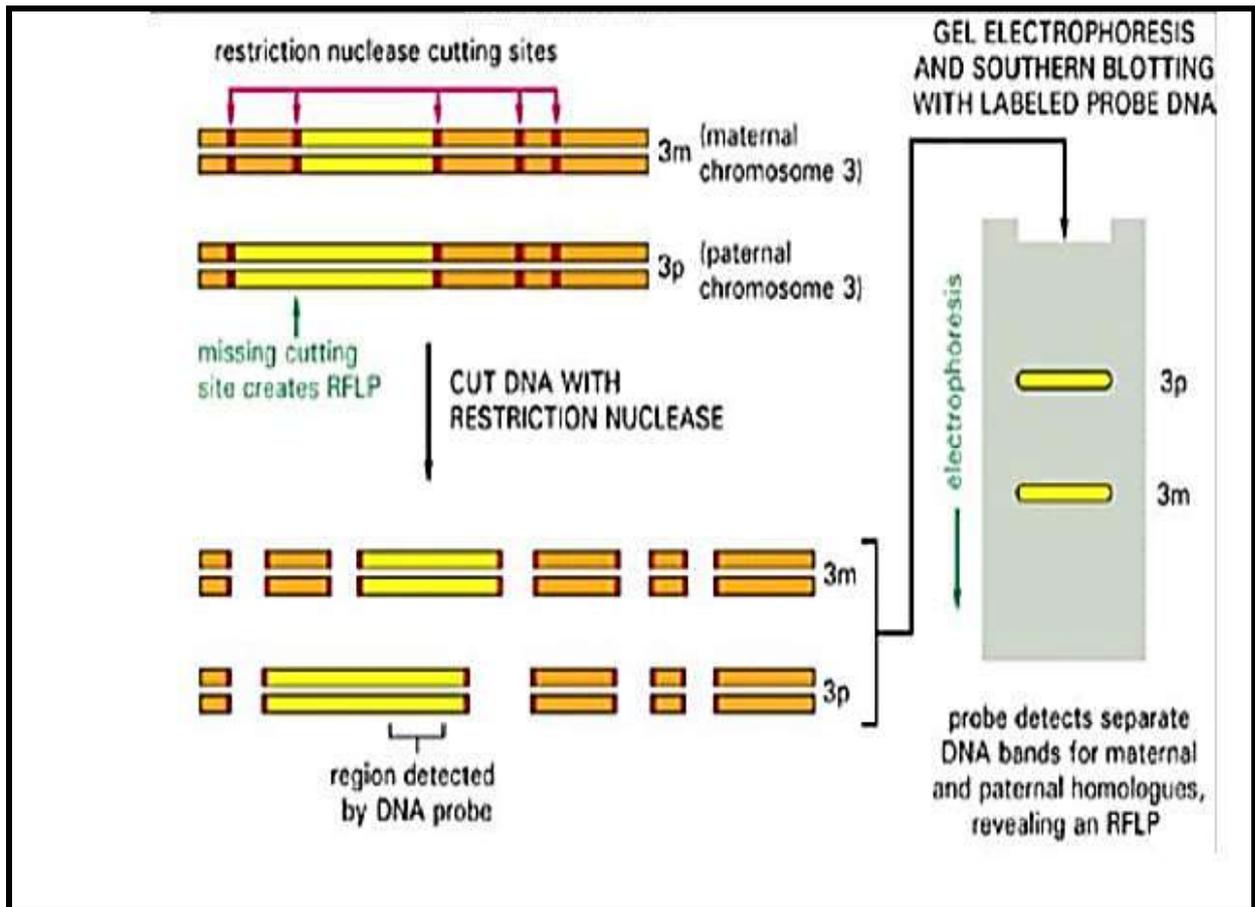


Figure 69 : La technique RFLP.

6.6 Réaction de Polymérisation en Chaîne "PCR"

La synthèse et le séquençage d'ADN ont permis l'émergence d'une méthode d'amplification de l'ADN appelée PCR (**P**olymerase **C**hain **R**eaction), à partir d'un gène (ou fragment) spécifique, de grande quantité d'ADN sont obtenues *in vitro*. L'ADN polymérase copiant jusqu'à un milliard de fois cette région cible « quantité suffisante pour être révélée ». Cette méthode de Biologie Moléculaire a été mise au point en 1985 par Kary Mullis, qui obtint pour ces travaux le prix Nobel de Chimie en 1993.

La technique comporte des cycles successifs. Chaque cycle comprend une succession de trois phases:

1-Dénaturation (autour de 95°C) : Sert à dénaturer l'ADN pour obtenir des matrices simple brin.

2- Hybridation (autour de Th): Lors du refroidissement du mélange, les amorces étant en excès, la plupart des brins d'ADN cibles se fixent à celles et non entre eux.

3- Extension (72°C) : L'ADN polymérase « prolonge les amorces : côté 3'-OH libre » en utilisant les brins cibles comme matrice.

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique. Pour effectuer ces transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil programmable : un thermocycleur. La progression géométrique de raison 2 (tab.) permet d'établir une relation pour calculer le nombre des copies totale et cibles.

Tableau 6 : Progression géométrique de raison 2 des cycles de PCR et nombre de copies.

Cycle	1	2	3	4	5	6	-----n	n : nombre de cycle
Copies totale (N)	2	4	8	16	32	64	-----	$N=2^n$
Copies parasites	2	4	6	8	10	12	-----	2n
Copies cibles	0	0	2	8	22	52	-----	2^n-2n

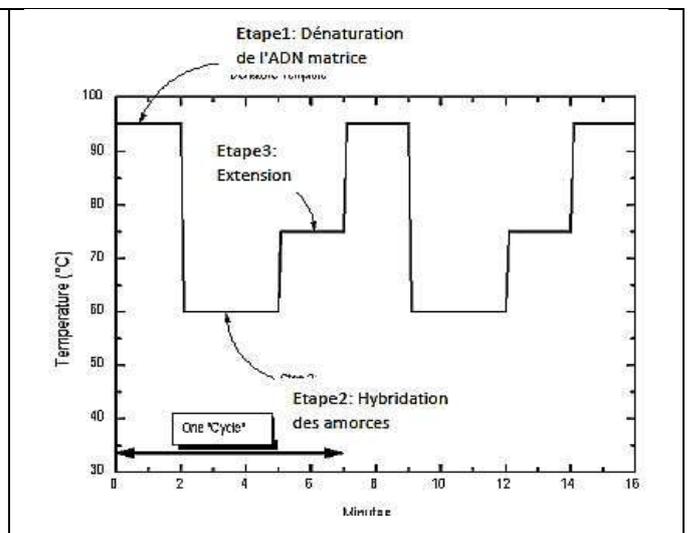
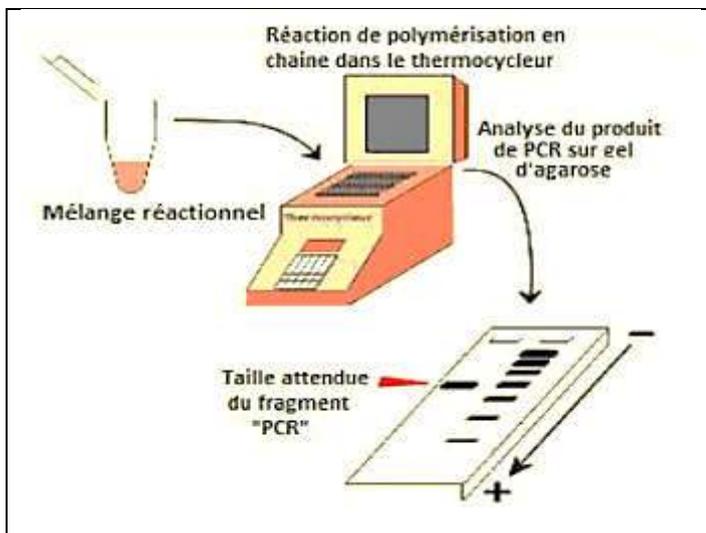


Figure 70 : Réaction d'amplification d'ADN dans le thermocycleur et visualisation du produit de PCR par électrophorèse

Figure 71:Profile de cycles de la température de la PCR

La connaissance de séquence de bordures entourant un gène d'intérêt est importante, les produits d'amplification obtenus par PCR peuvent être utilisés pour le clonage et le séquençage de ces gènes, ou pour des études comparatives ou phylogénétiques. Dans ces deux derniers cas, les amorces sont spécifiques à des régions conservées d'un gène commun d'une grande variété d'organismes. Par Exemple : le gène qui code pour l'ARNr 16S a des régions hautement conservées encadrant une région hautement variables. La PCR est utilisée aussi dans :

- La mutagenèse dirigée
- Les études paléontologiques (fossiles : momies, plantes, animaux, ...etc.)
- Le diagnostique microbiologique
- Le diagnostique moléculaire des maladies génétiques et infectieuses (surtout les cancers et les microorganismes difficilement cultivés Ex. Bartonella, virus ...etc.).

6.7Le Séquençage

Le séquençage de l'ADN constitue une méthode dont le but est de déterminer la succession linéaire des bases A, C, G et T prenant part à la structure de l'ADN. La lecture de cette séquence permet d'étudier l'information biologique contenue par celle-ci. Étant donné l'unicité et la spécificité de la structure de l'ADN chez chaque individu, la séquence de l'ADN permet de nombreuses applications dans le domaine de la médecine, comme, par exemple, le diagnostic, les études génétiques, l'étude de paternité, la criminologie, la compréhension de mécanismes physiopathologiques, la synthèse de médicaments, les enquêtes épidémiologiques.

6.7.1 Méthode de Sanger

La méthode de séquençage de Sanger (dite par terminaison de chaîne) utilise des nucléotides appelés didésoxyribonucléotides (ddNTP) qui ont un atome d'hydrogène à la place du groupement OH sur le carbone 3' du ribose.

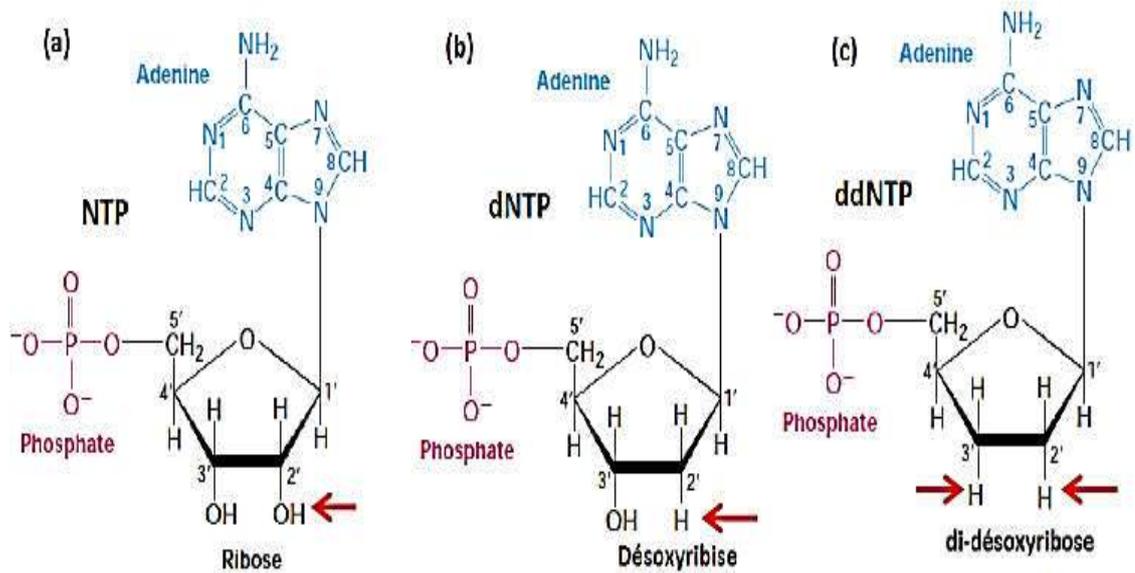


Figure 72 : Les différentes formes des nucléotides.

L'utilisation d'un ddNTP permet d'obtenir un ensemble de fragments d'ADN de différentes tailles, correspondant aux emplacements d'un nucléotide donné.
 exemple: synthèse du brin complémentaire, et arrêt aléatoire par un ddGTP quand il y a une Cytosine sur la séquence

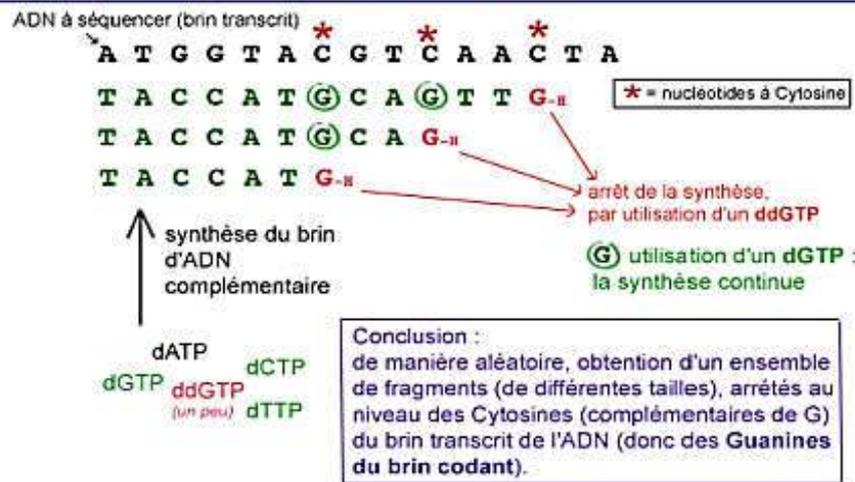


Figure 73 : Principe de Sanger.

4 mélanges sont préparés:

- le fragment qui doit être séquencé cloné dans un vecteur
- un petit morceau d'ADN dont la séquence est complémentaire à l'extrémité 3' du fragment à séquencer = amorce
- les 4 dNTP's (dCTP, dATP, dGTP, dTTP) (Un des nucléotides est marqué au ³⁵S)
- l'ADN polymérase

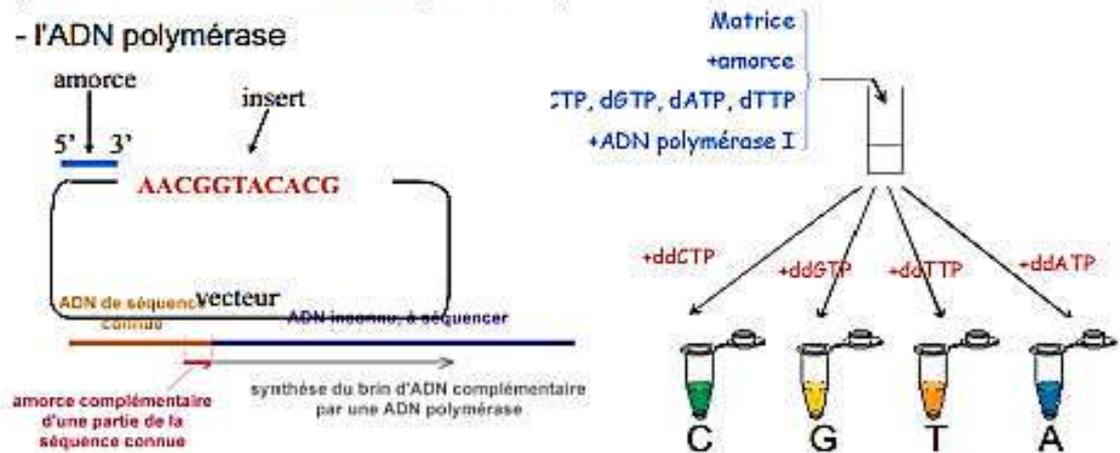


Figure 74 : Mélange réactionnel de Sanger.

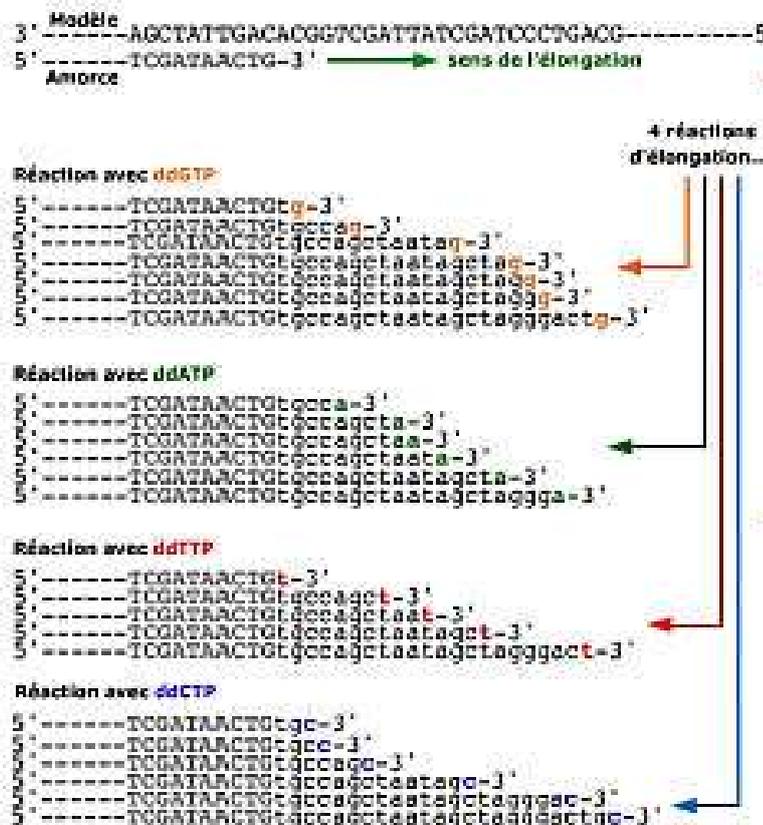


Figure 75 : Ensemble des fragments obtenus en fin de séquençage.

- Électrophorèse sur gel d'acrylamide
- Un des dNTP est marqué au ³⁵S, rendant le fragment radioactif
- Détection des fragments d'ADN, en exposant un film photographique au gel.
- Lecture de la séquence

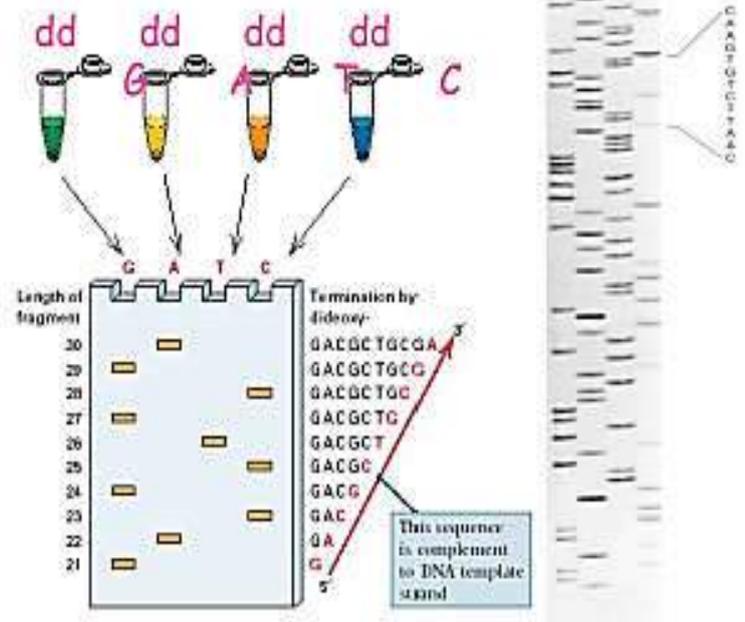


Figure 76 :Migration et visualisation des produits du séquençage.

L'Evolution de la Technique de Sanger est affectée par Deux changements:

- 1- Didésoxynucléotides marqués avec des fluorophores
- 2- Automatisation

6.7.2 Technique chimique de Maxam et de Gilbert

La méthode chimique de Maxam et de Gilbert est moins utilisée actuellement que la méthode enzymatique de Sanger. L'ADN à séquencer est tout d'abord marqué en 5' avec phosphore 32 (dATP), puis clivé après A, G, C ou T par divers réactifs chimiques. Après clivage par ces réactifs, les fragments produits sont séparés par électrophorèse en gel de polyacrylamide. L'examen de l'autoradiographie correspondante permet de connaître la séquence du brin analysé.

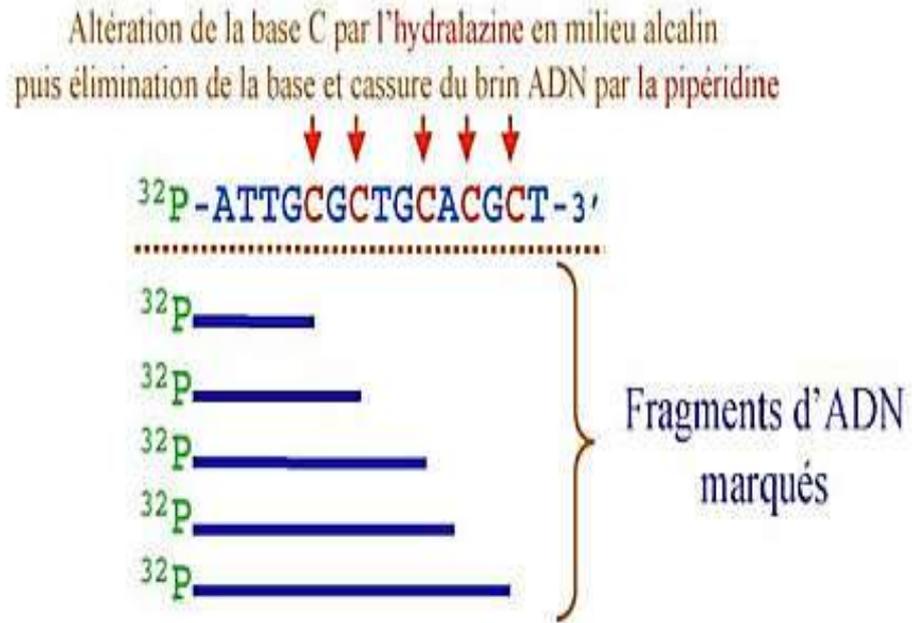


Figure 77 : Technique de Maxam-Gilbert.

Partie II : Génie génétique

1. Outils enzymatique du génie génétique

À l'heure actuelle, aucun procédé chimique ne peut atteindre la capacité de manipuler l'ADN *in vitro* d'une manière prévisible. Les enzymes jouent un rôle vital dans le processus du génie génétique. Ce sont des catalyseurs biologiques accélérant des réactions chimiques qui, sans cela, seraient trop lentes pour permettre à la cellule de fonctionner. Ces molécules remplissent leur rôle en interagissant physiquement avec les molécules de substrat, pour créer ou rompre des liaisons chimiques.

1.1 Nucléases

Les nucléases sont des enzymes qui dégradent les molécules d'ADN ou RNA en rompant les liaisons phosphodiester liant un nucléotide au suivant. Il existe plusieurs types :

- **Selon le substrat:**

- DNases desoxyribonucléases
- RNases ribonucléases

- **Selon la fonction:**

- **Exonucléases** : libèrent par hydrolyse le nucléotide situé à l'extrémité 5' ou 3'.
- **Endonucléases** : hydrolysent une liaison ester interne.

A. Les endonucleases :

Découverte par Werner Arber, Daniel Nathans et Hamilton O. Smith vers 1965, ce sont des endonucléases capables de couper l'ADN double brin à des sites spécifiques de 4 à 6 paires de bases (parfois plus) appelée site de restriction. Les séquences habituellement reconnues sont palindromiques ; c'est-à-dire que la succession des nucléotides est identique pour le brin sens (lecture 5'-3') et pour le brin antisens (toujours dans le sens de lecture 5'-3')

En effet, les enzymes de restriction peuvent couper (et ainsi conduire à la destruction) de l'ADN étranger (notamment des virus). Cela limite les infections virales chez les bactéries, d'où le terme de restriction. L'ADN bactérien lui-même est protégé de la coupure par des modifications du type méthylation. L'enzyme de modification est une ADN méthyltransférase qui méthyle spécifiquement les bases cytosine (sur le carbone 5) ou de l'adénine (sur l'azote 6) dans la séquence de reconnaissance. Leur nomenclature ne diffère de celle des endonucléases de restriction que par l'ajout d'une lettre M au début du nom : **M. EcoR I**.

Dans les années 1960, les biologistes des bactériophages découvraient la base biochimique du phénomène de restriction et de modification par l'hôte. L'aboutissement de cette recherche fut la purification de l'endonucléase de restriction d'*E.coli* K12 par Meselson et Yuan en 1968. La percée technologique tant attendue survint en 1970 avec la découverte d'une enzyme d'*Haemophilus influenzae* qui a des propriétés plus simples à exploiter. En effet l'enzyme reconnaît un motif de séquence spécifique dans une molécule d'ADN double brin et couper l'ADN dans ce motif en générant des fragments de longueur et de séquences caractéristiques.

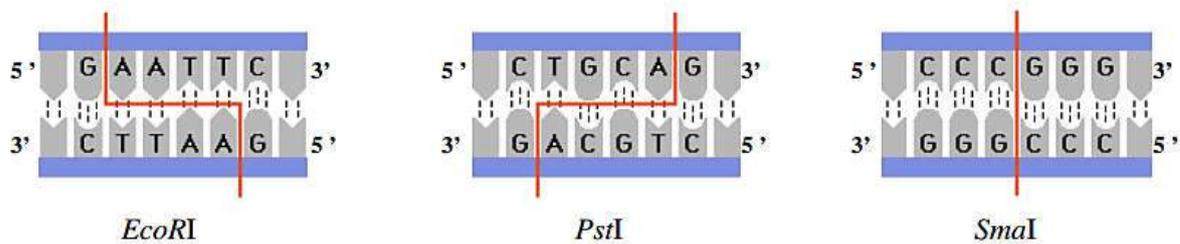


Figure 78 : Séquences palindromiques des enzymes de restriction.

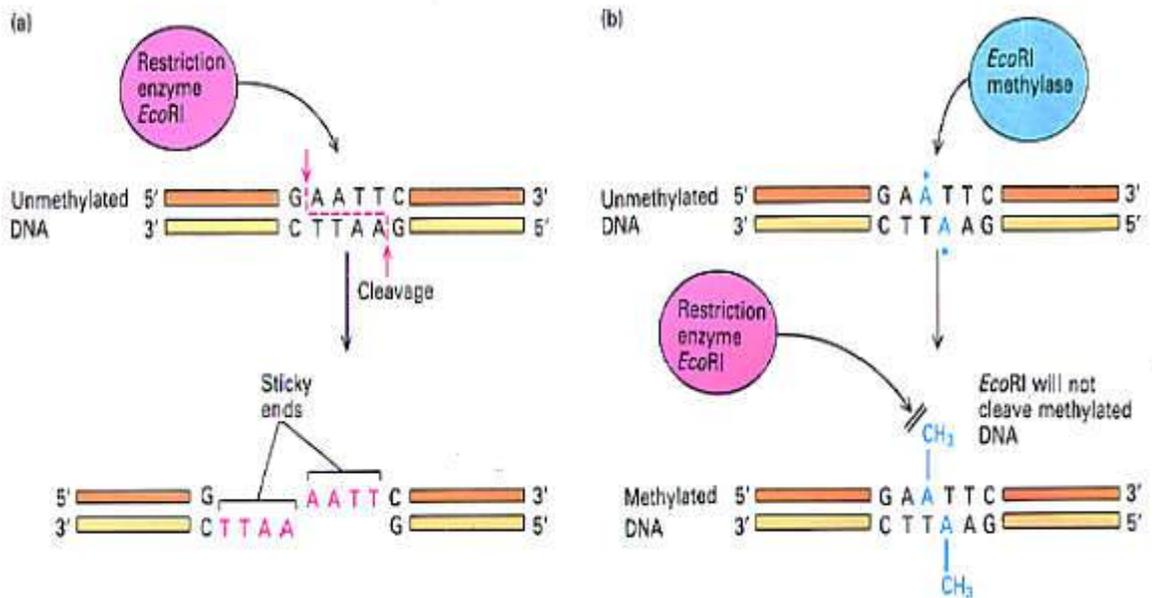


Figure 79: Système restriction /Méthylation.

Il existe trois types d'enzymes de restriction isolées des bactéries.

- 1- **Type I** : Dont l'action nécessite la présence de Mg^{++} , d'ATP comme cofacteur, et de S- adénosyle-méthionine. Leur site de coupure est éloigné de leur site de reconnaissance (jusqu'à plusieurs milliers de nucléotides, plus loin dans certain cas).
- 2- **Type II** : Pratiquement les seules utilisées en génie génétique, reconnaissent l'ADN à des sites particuliers et coupent dans ces sites ou à proximité immédiate d'eux. Ce sont des nucléases qui coupent à l'intérieur, donc ce sont des endonucléases (Type II restriction endonucléases).
- 3- **Type III** : Ressemblent à celles du type I pour la séparation des sites de reconnaissance et de coupure, mais qui s'apparentent à celles du type II par leur mode d'action.

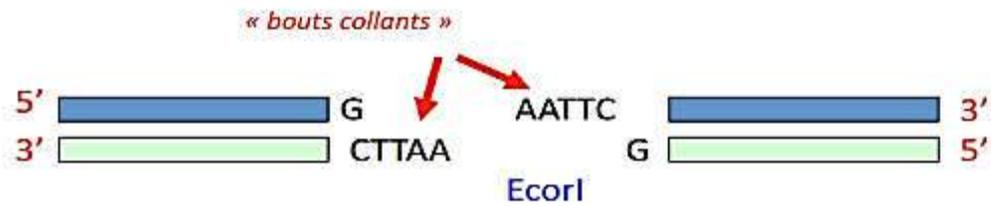
La Nomenclature des enzymes de restriction dépend de règles spécifiques qui tiennent compte de la bactérie dont a été isolée l'enzyme de restriction :

- La première lettre, en majuscule, représente l'initiale du genre bactérien.
- Les deux lettres, minuscules, qui suivent la première sont représentatives de l'espèce.
- Le chiffre romain qui suit ces trois lettres est le numéro d'ordre de découverte de l'enzyme pour la même bactérie source
- La dernière lettre majuscule n'est pas obligatoire pour toutes les endonucléases de restriction. Elle est représentative de la souche de la bactérie d'où l'enzyme a été isolée.

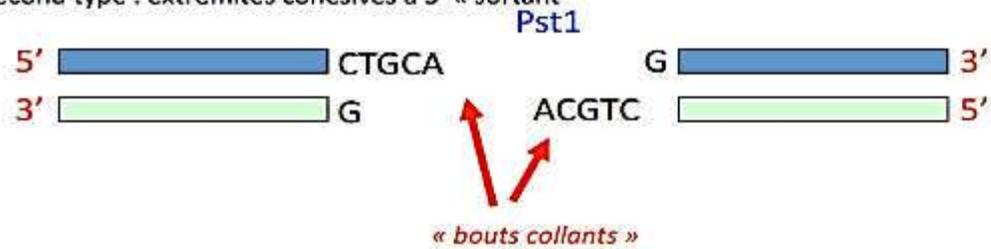


Deux types de coupures peuvent se produire en fonction des enzymes : La coupure à extrémités franches (ou **bouc francs**) et la coupure à extrémités sortantes (ou **cohésives**). Dans le premier cas, la coupure a lieu au milieu du palindrome, dans le deuxième cas, la coupure a lieu de part et d'autre du centre de symétrie.

- Premier type : extrémités cohésives à 5' « sortant »:



- Second type : extrémités cohésives à 3' « sortant



- Troisième type : coupure à bouts francs

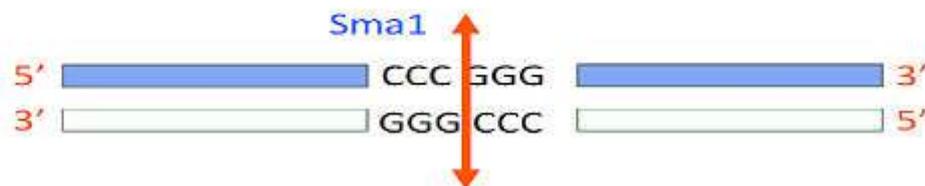


Figure 80 : Les différents types de coupures des enzymes de restriction.

Les enzymes de restriction qui reconnaissent les mêmes séquences sont appelés **isoschysomères**. Des enzymes qui reconnaissent la même séquence mais la coupent de façon différente sont parfois appelées **néoschizomères**. Dans des conditions particulières telles qu'un pH élevé ou une force ionique basse, certaines endonucléases de restriction coupent des séquences similaires mais pas identiques à leur séquence de reconnaissance normale. Cette spécificité modifiée est appelée l'activité étoile « **star activity** ».

Tableau 7 : Exemples des enzymes de restriction de type II.

Enzyme	L'organisme source	Séquence de reconnaissance
<i>HpaII</i>	<i>Haemophilus parainfluenzae</i>	C/CGG GGC/C
<i>MboI</i>	<i>Moraxella bovis</i>	/GATC CTAG/
<i>NdeI</i>	<i>Neisseria denitrificans</i>	/GATC CTAG/
<i>EcoRI</i>	<i>Escherichia coli</i> RY13	G/AATC CTAA/G
<i>EcoRII</i>	<i>Escherichia coli</i> RY13	/CC(A or T)GG GG(T or A)CC/
<i>EcoRV</i>	<i>Escherichia coli</i> J62/pGL74	GAT/ATC CTA/TAG
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	G/GATCC CCTAG/G
<i>SauI</i>	<i>Staphylococcus aureus</i>	CC/TNAGG GGANT/CC
<i>BglI</i>	<i>Bacillus globigii</i>	GCCNNNN/NGGC CGGN/NNNNCCG
<i>NotI</i>	<i>Nocardia oitidis-caviarum</i>	GC/GGCCGC CGCCGG/CG
<i>DraI</i>	<i>Deinococcus radiophilus</i>	RG/GNCCY YCCNG/GR

1.2 Les polymérase

a. ADN polymérase I et fragment de Klenow

Comme toutes les ADN polymérase, l'ADN polymérase I synthétise une chaîne d'ADN dans le sens 5'-3' à partir d'une matrice d'ADN et d'une amorce ADN ou ARN. Elle possède aussi deux activités exonucléasiques, l'une dans le sens 3'-5' (édition ou correction des erreurs), l'autre dans le sens 5'-3'. On utilise le plus souvent au laboratoire le fragment de Klenow qui a perdu (par clivage protéolytique) l'activité exonucléasique 5'-3' et a conservée l'activité polymérasique ainsi que l'activité exonucléasique 3'-5'.

b. ADN Taq polymérase et autres ADN polymérase thermostables

Il s'agit d'une ADN polymérase isolée de bactérie (*Thermophilus aquaticus*) vivant dans les sources d'eau chaude. Son avantage réside en sa très grande thermostabilité (jusqu'à 95 °C) et son fonctionnement optimal à température élevée (70°C).

- La Taq polymérase est inhibée par les ions phosphates.
- La Taq polymérase est utilisée pour la réaction de polymérisation en chaîne (Polymerase Chain Reaction = PCR), technique courante d'amplification des fragments de DNA. Parce qu'elle est dépourvue d'activités d'édition.
- La Taq polymérase est responsable de nombreux mésappariements : de l'ordre de 1 pour 100 paires de bases.

La Pfu et la Pwo DNA polymérase proviennent de *Pyrococcus furiosus*, bactérie découverte dans des sources géothermiques en Italie et de *Pyrococcus woesei*. Elles ont les mêmes séquences et ont donc des activités identiques. Activité 5'-3' polymérase et 3'-5' exonucléase mais pas d'activité 5'-3' exonucléase. L'activité 3'-5' exonucléase dite correctrice permet de diminuer le taux d'erreurs à 10^{-6} par base dupliquée.

c. Transcriptase réverse ou inverse ou rétrotranscriptase RT

C'est une ADN polymérase ARN ou ADN dépendante, comme les autres polymérases, elle a besoin d'un brin d'ADN ou d'ARN matrice. Elle possède aussi une activité de type RNaseH. Les transcriptases réverses sont produites par des cellules infectées par des rétrovirus, virus à ARN qui font synthétiser un ADNc par la cellule hôte afin de permettre leur réplication.

d. ARN polymérase

Cet enzyme transcrit l'un des brins de l'ADN double brin en ARN. La synthèse s'effectue dans le sens 5'-3' en l'absence d'amorce. Par contre, chaque type d'ARN polymérase ne fonctionne qu'en présence de son promoteur, c'est-à-dire qu'une séquence spécifique d'ADN reconnue par l'enzyme, sur laquelle ce dernier se fixe et initie la transcription. Les trois ARN polymérases utilisées en laboratoire sont SP6 (*Salmonella thyphimurium*), T3 et T7 (des bactériophages T3 et t7), elles permettent la transcription in vitro qui fournit de grandes quantités d'une séquence d'ARN.

1.3 .Ligases

Elles assurent la formation d'une liaison phosphodiester entre une extrémité 3'-OH d'un fragment d'ADN et une extrémité 5'P d'un autre fragment adjacent. L'ADN ligase agit spécifiquement sur l'ADN double brin et l'ARN ligase sur l'ARN et l'ADN simple brin.

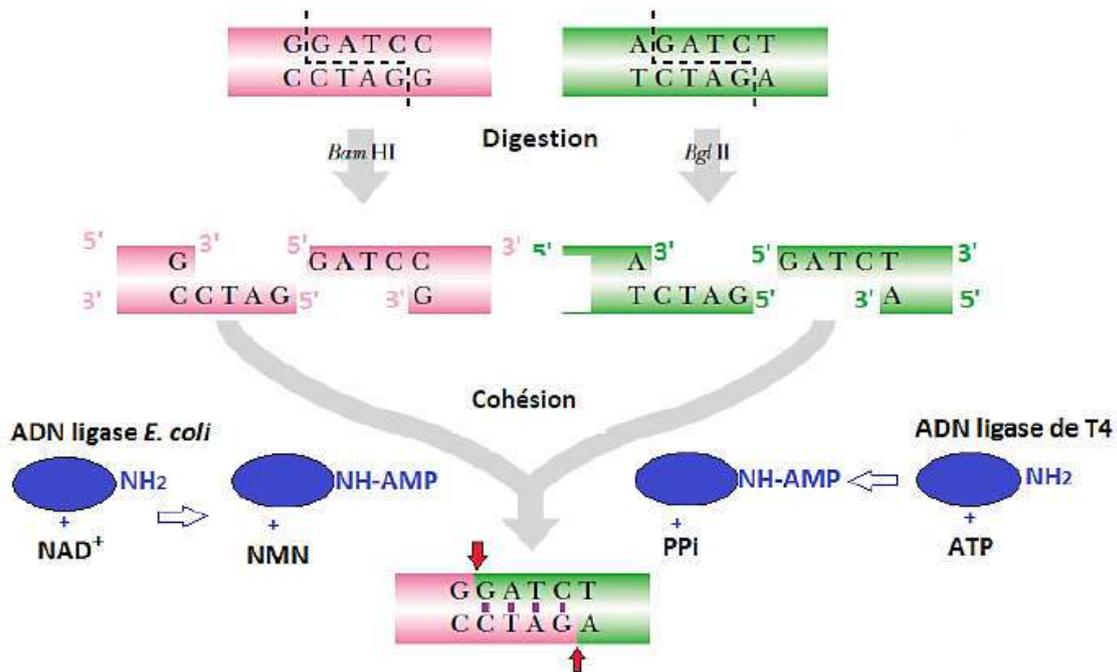


Figure 81: L'une des modalités de l'utilisation des ligases en génie génétique (ligation des extrémités cohésives des enzymes de restriction). L'enzyme ligature les extrémités 3'-OH et 5'-phosphate. L'ADN ligase est adénylylée par NAD (*E. coli*) ou par l'ATP (phage T4). Après l'adénylylation, l'extrémité 5' phosphate favorise la formation de liaison phosphoester par attaque nucléophilique ($_{:}\ddot{O}H$).

1.4 .Phosphatases

Elles catalysent l'élimination d'un phosphate en 5' d'une chaîne d'ADN. L'enzyme la plus utilisée est la phosphatase alcaline intestinale bovine (CIP : Calf intestinal phosphatase).

1.5 .Kinase

Elles permettent le transfert d'un groupement phosphate en position gamma d'une molécule d'ATP sur l'extrémité 5'-OH d'un ADN. L'enzyme la plus utilisée est la polynucléotide kinase du phage T4.

2. Vecteurs de clonage et analyse des banques

Les techniques dites de clonage moléculaire permettent de purifier les gènes sous forme de molécules caractérisées disponibles en grande quantité. Le clonage moléculaire est différent du clonage reproductif (créer un individu génétiquement identique à un autre mais d'un âge différent) ou du clonage thérapeutique (fabriquer des tissus à partir de cellules souches pour effectuer des greffes compatibles avec le receveur). Le clonage moléculaire consiste à (figure) :

- Fractionner les génomes de façon spécifique à d'enzyme de restriction
- Purifier les fragments en les associant individuellement à un petit réplicon (phage ou plasmide) appelé vecteur
- Introduire et propager séparément chacun des vecteurs recombinants dans des lignées cellulaires clonales
- Trier ces lignées cellulaires pour identifier celles qui portent les séquences du gène recherché.

On appellera donc banque d'ADN une collection de clones cellulaires ou viraux portant chacun un fragment particulier de l'ADN cloné. Les deux premières étapes impliquent une méthodologie commune dites « recombinaison in vitro ». la suite des opération varie selon le type de vecteur et selon le type de cellule utilisées pour propager le vecteur.

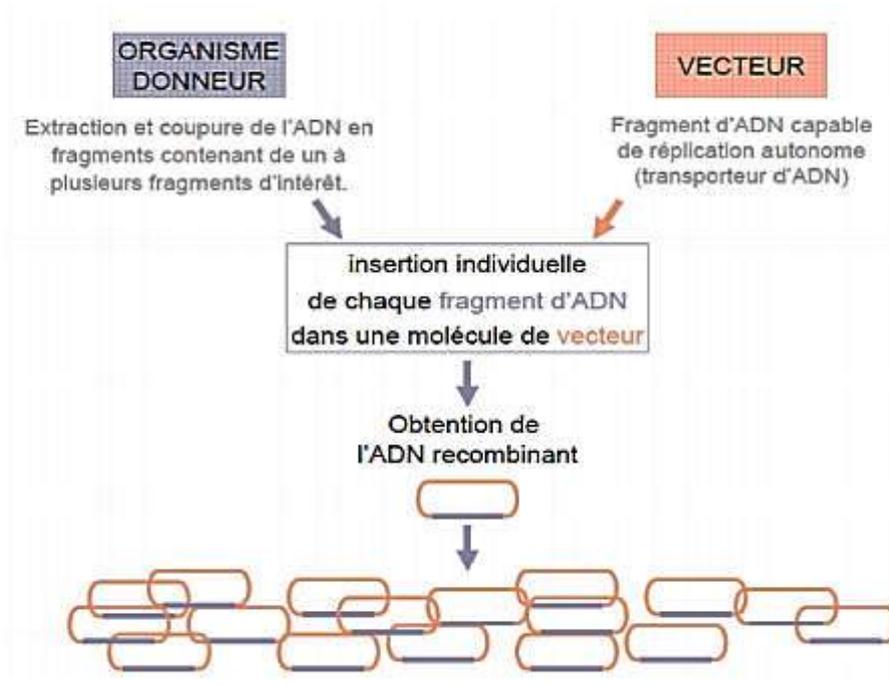


Figure 82: Stratégie de clonage.

2.1 Les Vecteurs de clonage

Un vecteur est une séquence d'ADN permettant la propagation, la sélection, la modification d'une séquence d'ADN d'intérêt. En bref, l'étude et la manipulation d'une séquence d'ADN isolée.

Il existe 2 catégories de vecteurs:

- Vecteur de clonage: Destiné à isoler physiquement un fragment d'ADN et à amplifier le nombre de copies.
- Vecteur d'expression: Destiné à transférer un gène et le faire exprimer dans une cellule hôte qui n'est pas sa cellule d'origine

Tableau 8 : Vecteurs et taille des banques.

Vecteur de clonage	Taille des inserts (kpb)	Taille des banques		
		<i>E. coli</i>	<i>S. cerevisiae</i>	<i>Homo</i>
Plasmides	< ~ 10	5 000	15 000	1×10^7
Phages λ	< 20	2 000	7 000	3×10^6
Cosmides	40	500	1 500	5×10^5
BACs	< 300	90	300	1×10^5
YACs	> 500	10	30	$< 10^4$

2.1.1 Propriétés des vecteurs de clonage

- Capables de répllication autonome (répllication épisomale) dans une cellule hôte donnée.
- Possède un polylinker ou site multiple de clonage.
- Petite taille: pour permettre l'insertion d'un fragment d'ADN plus ou moins grand.
- Présence de gènes de sélection: sélection des cellules hôtes qui ont intégré un vecteur.
- Stabilité: maintient sans modification dans la cellule hôte, quel que soit le nombre de division.

2.1.2 Différents vecteurs de clonage

A. Les plasmides

Les plasmides sont des molécules d'ADN bicaténaire, circulaires et extrachromosomique, de petite taille (5 à 4000 fois plus petit que le chromosome), se répliquant d'une manière autonome. Ils ne contiennent généralement pas de gènes indispensables à la survie de la cellule. Ils apportent des gènes pouvant présenter un avantages dans des conditions de culture particulières (antibiotique, métaux lourd, hyper salinité,...) où ils vont apporter un avantage pour la croissance de la cellule les possédant.

Caractéristiques des plasmides :

- Peuvent être présent en plusieurs centaines de copies au sein d'une même cellule.
 - bactérienne (cela dépend de la souche bactérienne).
 - Présence de la séquence ORI (qui permet la réplication autonome).
 - Petite taille: 2 à 5 kb (le chromosome d'E. coli fait 4700kb).
 - Possèdent des sites de restriction qui permettent l'introduction d'ADN étranger.
- Ces sites doivent être situés dans des gènes pour permettre la sélection des recombinants par rapport aux plasmides recircularisés

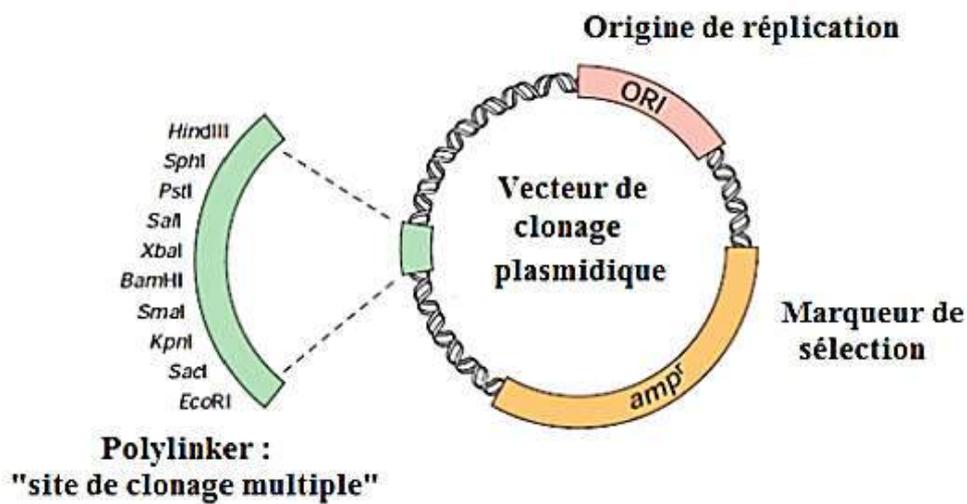


Figure 83 : Composants basiques d'un vecteur plasmidique qui peut se répliquer chez *E. coli*

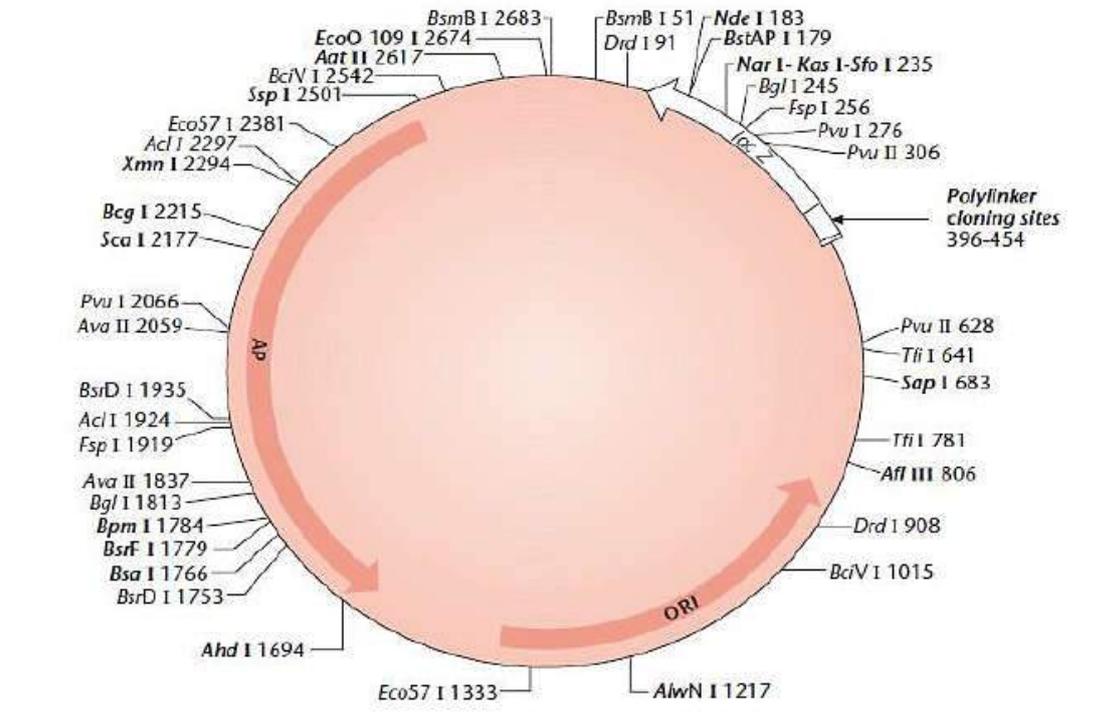


Figure 85 : La carte génétique de certain vecteur de type pUC dérivés de pBR322, le site de clonage multiple (polylinker) est introduit dans le gène *lacZ*, sans interrompre la fonction du gène.

B. Les phages : Les vecteurs phagiques présentent deux avantages essentiels par rapport aux vecteurs plasmidiques:

- Taille de l'insert plus grande (12 à 22kb).
- Infection (pénétration de l'ADN) spontanée des cellules bactériennes.

Le bactériophage lambda est un virus qui infecte les souches d'*Escherichia coli* et d'autres entérobactéries en s'adsorbant sur une porine de la membrane externe, le transporteur du maltose LamB. L'ADN du phage est une molécule linéaire de 48,5 Kpb compacté dans la capsid, il porte à ses extrémités des séquences cohésives de 12pb dites extrémités cos. Après injection dans la cellule, cet ADN se recircularise sur lui-même spontanément par les extrémités cos et la forme circulaire est stabilisée par la ligase de l'hôte. Selon les conditions physiologiques qui prévalent dans la cellule infectée, le phage sauvage choisit alors entre deux voies :

La voie lysogénique : L'ADN circulaire du phage s'intègre par recombinaison homologue entre séquences du phage (*attP*) et de la bactérie (*attB*). La réaction est catalysée par une topoisomérase de type II spécifique de ce site, codée par le phage, appelée intégrase. L'ADN du phage intégré dirige la synthèse d'un répresseur qui empêche l'expression des gènes phagiques et ce dernier reste silencieux et peut persister sous cette forme intégrée, répliquée passivement avec le chromosome, pendant de nombreuses générations. On parle de prophage et de stade lysogène pour la bactérie.

La voie lytique : Le phage se réplique selon un mécanisme bidirectionnel, à la manière d'un chromosome bactérien, puis selon un mécanisme de réplication dit en « cercle roulant ». En cercle roulant sur le brin non clivé, la polymérase allonge indéfiniment le brin clivé sous forme de brin monocaténaire (figure). Le brin monocaténaire est à son tour recopié pour générer une molécule bicaténaire portant plusieurs génomes du phage Lambda en tandem : on parle de **concatémères**. La protéine A, reconnaît les sites *cos* et se fixant aux têtes de la capsid du phage (formées par autoassemblage) de la protéine E) permet à la capsid d'aspirer l'ADN adjacent. Celui-ci se compacte dans la capsid ce qui active le dimère de la protéine A qui clive les deux sites *cos* et permet alors la fermeture de la tête et l'addition de la queue. D'autres protéines du phage (gènes de lyse) permettent la libération du phage

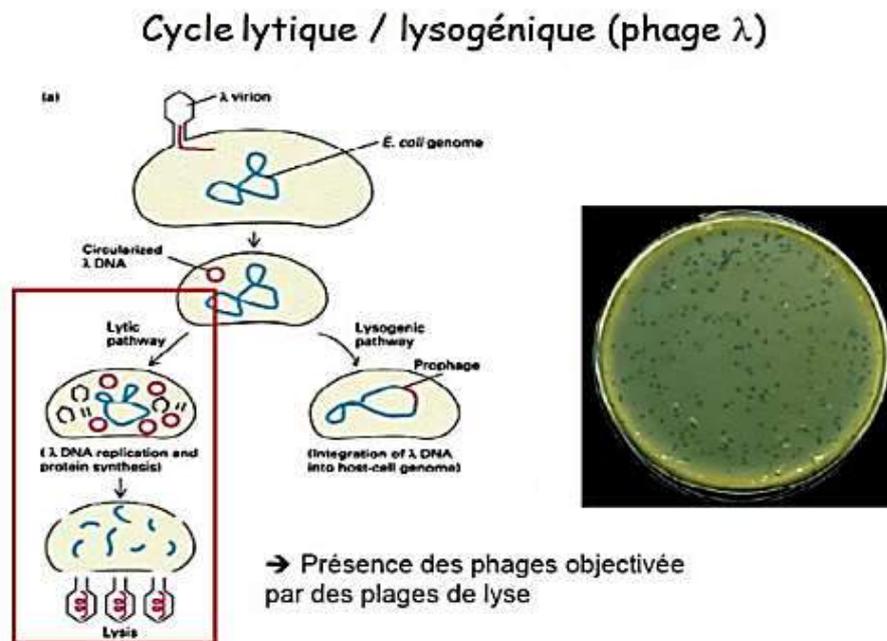


Figure 86: Cycle lytique et lysogénique du phage λ .

Il existe deux stratégies d'utilisation de vecteurs lambda:

I. Stratégie par insertion (environ 0 – 12 kb) : un site pour EcoR1.

II. Stratégie par délétion/remplacement (environ 16-26 kb)

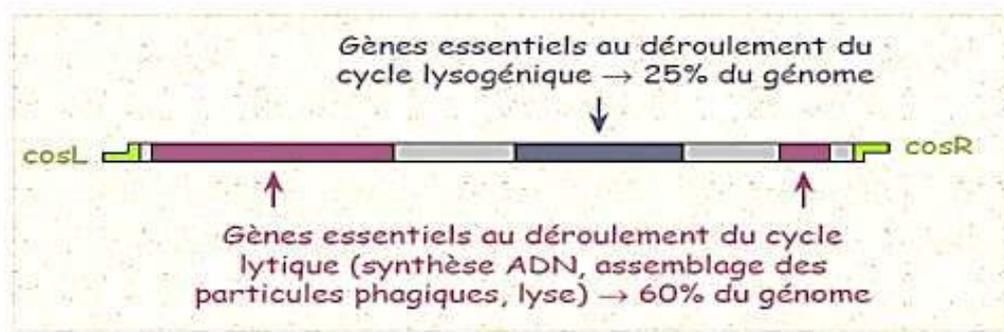
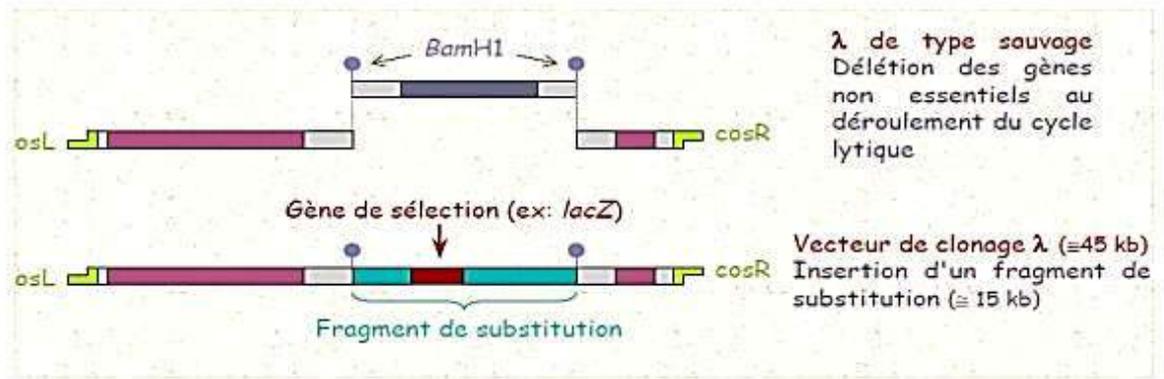


Figure 87 : Gènes essentiels au déroulement du cycle lytique et lysogénique.

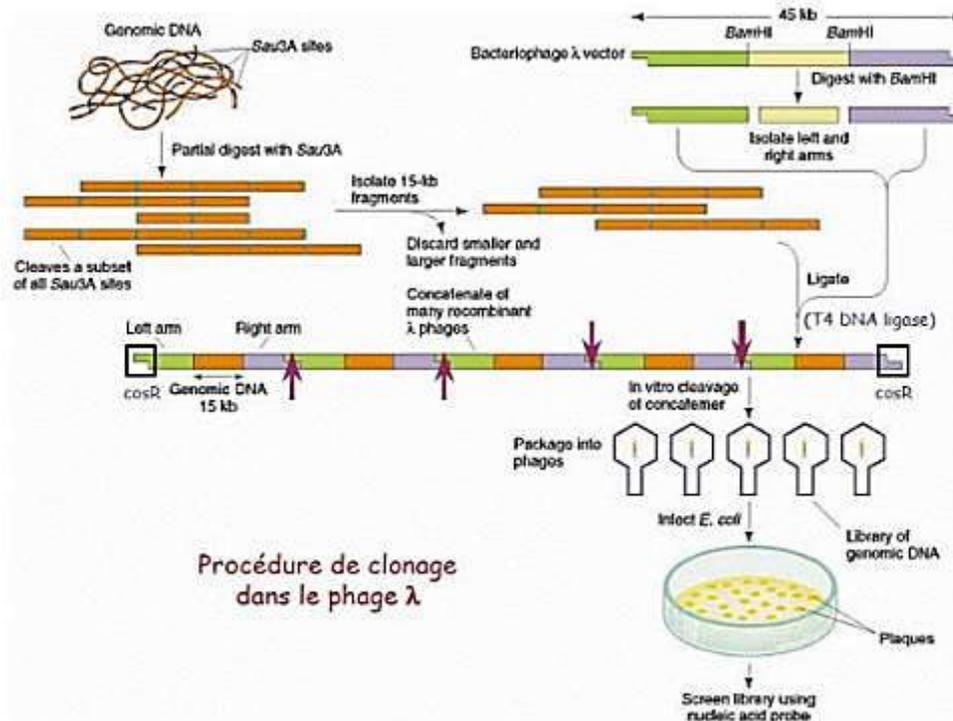


Figure 88: Procédure de clonage dans le phage λ .

c. Les cosmides

Les cosmides sont des vecteurs artificiels hybrides: phage lambda-plasmides. En fait, ils se comportent comme des plasmides avec des sites de restriction permettant l'insertion d'ADN étranger. Ils renferment également un gène de résistance aux antibiotiques (tétracycline). De plus, un site cos d'un virus lambda a été inclus dans leur ADN circulaire ce qui permettra au cosmide d'être empaqueté dans la tête d'un virus lambda. Les cosmides permettent le clonage de fragment allant jusqu'à 45 kb (capacité $\lambda \times 3$).

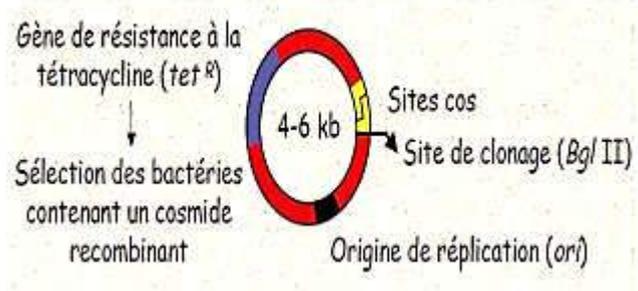


Figure 89 : Cosmide.

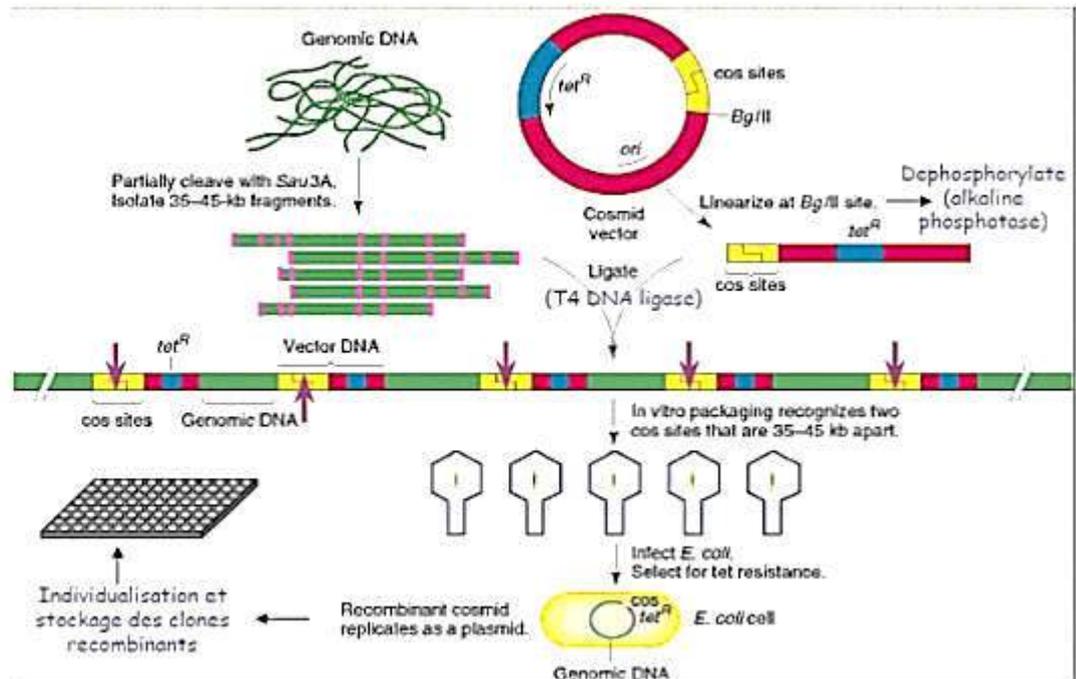


Figure 90 : Procédure de clonage dans un cosmide.

d. Les chromosomes artificiels bactériens (BAC= Bacterial Artificial Chromosome)

Le plasmide sexuel F d'*E. coli* est naturellement présent à une copie par génome. Il porte une origine de réplication à bas nombre de copie et un grand nombre de gènes impliqués dans ses fonctions de facteur sexuel (synthèse des pilus, conjugaison et transfert) qui peuvent être délétés sans affecter sa réplication. C'est un très grand plasmide, de plus de 100 kpb qui apparaissent très stable.

Un vecteur BAC (moins de 10 kb) comporte l'origine de réplication (*oriS*) du facteur F, le gène d'une protéine Rep essentielle au fonctionnement d'*oriS*, deux gènes *par* qui dirigent la partition (distribution) des copies du plasmide après réplication entre chacune des cellules filles, un marqueur de sélection et un multisite de clonage (*Bam*HI et *Hind*III). On peut y cloner des fragments de 100 à 300 kpb qui seront propagés de façon fidèle par *E. coli*.

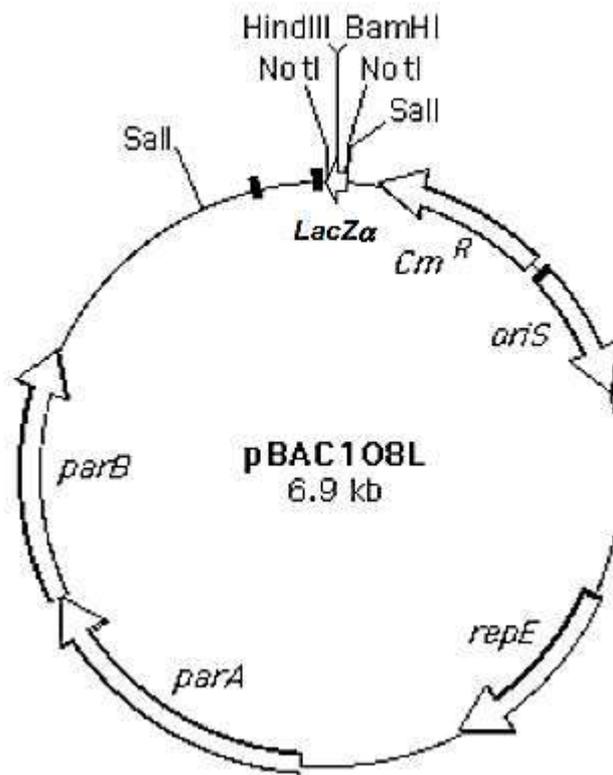


Figure 91: Chromosome artificiel bactérien (BAC).

e.Chromosomes artificiels de levure ou YAC (Yeast Artificial Chromosomes)

Le clonage de fragments plus longs que 40-50 kb a longtemps été difficile chez *Escherichia coli*, à la fois faute de vecteurs appropriés et pour les raisons d'instabilité qui viennent d'être invoquées. L'utilisation d'un nouvel hôte, l'eucaryote unicellulaire *Saccharomyces cerevisiae*, a permis de franchir une nouvelle étape dans la course au clonage des grands fragments qui peuvent atteindre jusqu'à 2000 kb. Un vecteur YAC comporte des gènes marqueurs de sélection (URA3 et TRP1), une origine de réplication (ARS1), un centromère (CEN) et deux télomères (TEL).

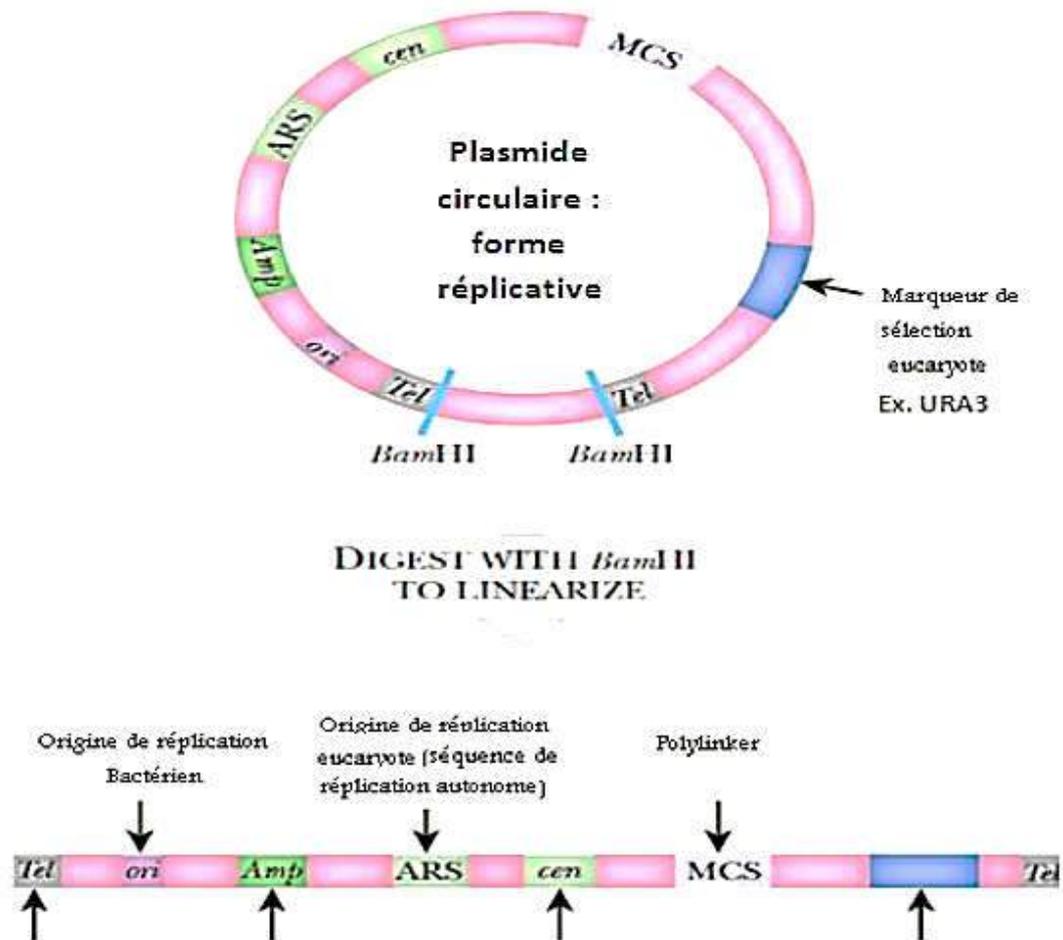


Figure 92 : Diagramme d'un YAC contenant des éléments essentiels pour la répllication chez les bactéries (Ori, gène de résistance (*Amp*), forme circulaire), et chez les levures (forme linéaire obtenu après digestion par *Bam*HI).

f. Vecteurs navettes et d'expression

Pour répondre aux besoins de la biotechnologie, surtout si c'est le but du clonage est d'obtenir un niveau d'expression élevé d'un gène dans un hôte approprié, d'autres vecteurs spécialisés ont été produits.

Les vecteurs navettes peuvent transporter un ADN cloné entre deux organismes différents et se répliquer de manière stable dans chacun d'eux. Il faut noter que les organismes possèdent des systèmes de régulation complexe qui sont des obstacles à l'expression des gènes étrangers. Pour cette raison les vecteurs d'expression ont été élaborés. Un vecteur d'expression permet de cloner un gène, mais contient aussi les séquences de régulation nécessaire à son expression.

Dans la plupart des cas on utilise avec le promoteur l'opérateur responsable sur son contrôle. Comme sa on peut régler l'expression par l'ajout de l'inducteur ou du répresseur.

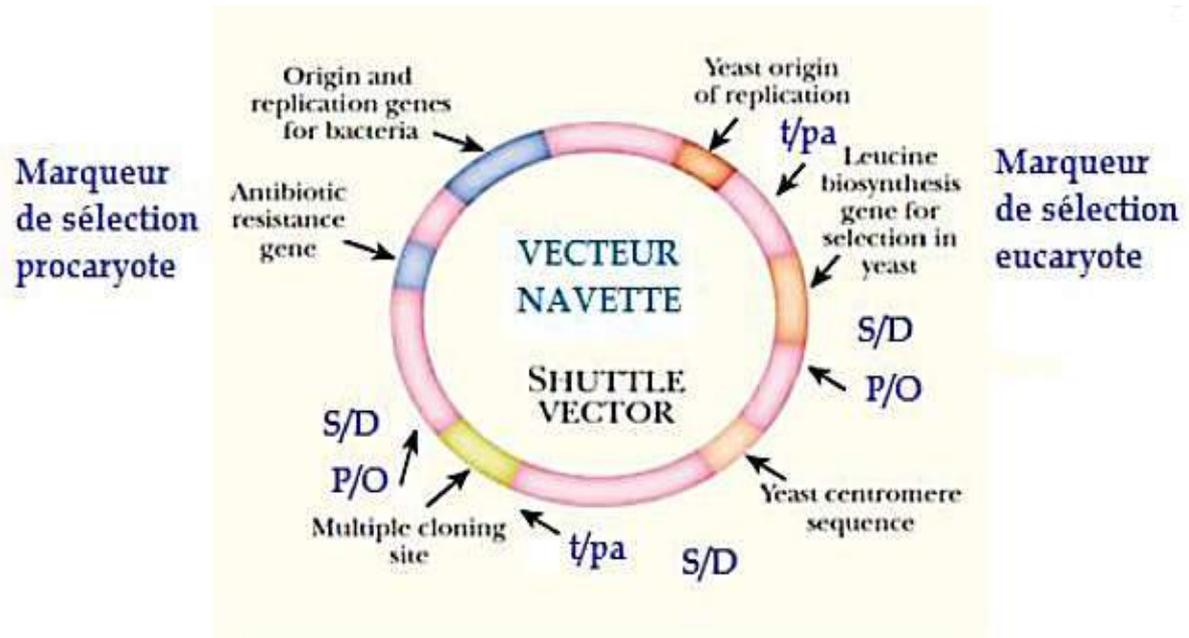


Figure 93 : Caractéristique d'un vecteur navette d'expression (shuttle expression vector). t/pa: signaux de terminaison de la transcription/ poly adénylation. T1, T2: terminateurs de la transcription. P/O : promoteur –Opérateur, S/D: séquence de Shine Dalgarno.

2.2 Types de banques et sources d'ADN

Une banque est une population de vecteurs contenant des fragments d'ADN divers représentatif d'un type cellulaire, d'un génome, d'un stade particulier de développement. Une banque est complète si collectivement l'ensemble de ces inserts recouvre totalement le génome de l'organisme de départ. La formule de Clark et Carbon permet d'évaluer le nombre de n de clones nécessaires.

Soit un génome de taille G , cloné en fragments de longueur moyenne l . La probabilité P qu'une séquence quelconque soit représentée dans la banque, étant 1 moins la probabilité qu'elle ne se trouve sur aucune des clones de la banque, est :

$$P = 1 - (1 - l/G)^n$$

Il en suit que, pour trouver un clone portant la séquence recherchée avec une probabilité P choisie (en générale 0,99), il faudra cribler n clones, où :

$$n = \frac{\log(1 - P)}{\log(1 - l/G)}$$

Vous verrez aussi la formule équivalent :

$$P = 1 - e^{-n/G} \text{ ou } 1 - e^{-a}$$

où a est la somme des longueurs des séquences clonées divisée par la taille du génome

Figure 94 : La formule de Clark et Carbon.

Une Banque est caractérisée par :

- le type de vecteur
- la nature des inserts (ADN génomique, ADNc), leur origine
- le pourcentage de vecteurs possédant un insert
- la taille moyenne des inserts
- la taille de la banque c'est dire la quantité de clones / ml de banque

2.2.1 Banques génomiques : On peut utiliser l'ADN génomique total d'un organisme. Cet ADN sera clivé en fragment dont les tailles et les extrémités seront compatibles avec celles du vecteur utilisé. De plus, il est souhaitable que la banque contienne des fragments chevauchants afin d'éviter que la séquence recherchée ne soit malencontreusement coupée en plusieurs morceaux ; l'une de façon de répondre à toutes ces contraintes et de réaliser une digestion partielle du génome à cloner avec une enzyme de restriction.

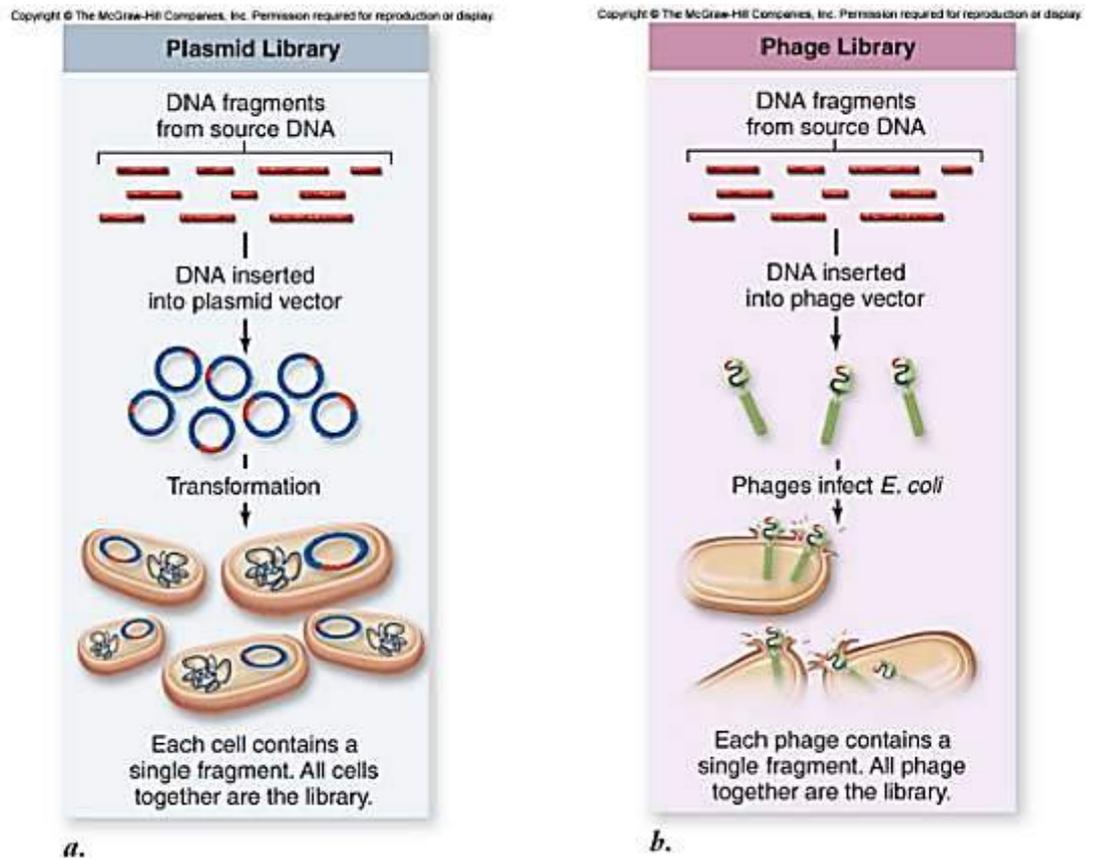
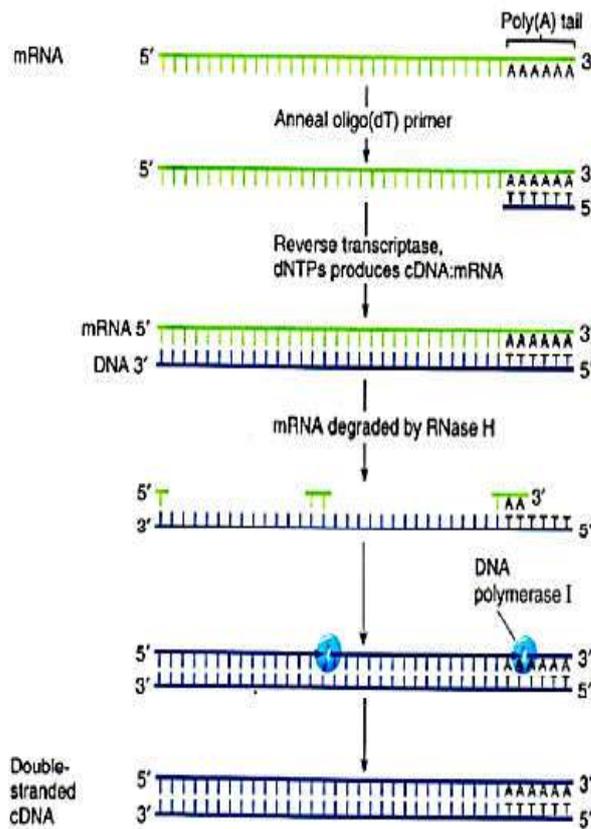


Figure 95 : Banques génomiques : A Bactériennes ; B Phagiques

2.2.2 Banques d'ADNc : On peut n'être intéressé que par le clonage de séquences effectivement traduites en protéines. Le principe de constitution d'une banque d'ADNc est représenté sur la figure. L'ARNm est d'abord recopié en ADN par la transcriptase inverse. Un oligonucléotide synthétique (olig-dT) s'appariera à tous les ARNm au niveau de leurs queues poly-A et fournira l'amorce requise par la transcriptase inverse. Cette enzyme allonge l'amorce, recopie l'ARNm jusqu'à son extrémité 5' et forme une boucle en épingle à cheveux à l'extrémité de l'ADNc néosynthétisé. L'ARNm est alors éliminé ou bien hydrolysé par la soude ou par une ARNase et l'ADN polymérase I peut alors utiliser la boucle en épingle à cheveux comme amorce pour initier la synthèse du brin complémentaire. Cette boucle est digérée par une nucléase SI. Les ADNc peuvent à ce stade être purifiés sur gel d'agarose et peuvent être clonés directement dans un vecteur coupé par une enzyme de restriction.



Comment convertir des ARNm en ADNc ?

- hybridation d'un oligoT sur la queue polyA de l'ARNm

- synthèse d'un brin d'ADN complémentaire par la transcriptase inverse

- dégradation de l'ARNm (par Rnase H ou traitement alcalin)

- synthèse du brin complémentaire d'ADN (par

- obtention d'un ADNc

Figure 96 : Les étapes de la synthèse d'ADNc.

3. Criblage des banques génomiques

On peut espérer que le gène cloné s'exprimera dans les cellules transformées et leur confèrera un phénotype particulier. En pratique cette approche est sévèrement limitée par les problèmes de compatibilité des systèmes d'expression des organismes donneurs et des cellules hôtes de la banque. Si on veut détecter un gène particulier, il faut repérer et identifier la colonie bactérienne contenant le gène cible. Cette opération s'appelle criblage de banques génomiques. Pour cela, il existe différentes techniques parmi lesquelles :

1. Le criblage par antibiotique
2. Le criblage Bleu-Blanc
3. La méthode de sélection par hybridation moléculaire
4. La méthode de criblage par Complémentation
5. La méthode de coupure avec de enzymes de restriction.

3.1 Criblage par antibiotique

Le taux de transformation des bactéries est généralement très faible. Il est fondamental de disposer de méthodes permettant d'obtenir des bactéries transformées par les plasmides et seulement celles ci. Pour cela, on utilise la résistance à un antibiotique (ampicilline). La transformation d'une bactérie sensible à un antibiotique aboutit à l'apparition d'une résistance uniquement pour les bactéries ayant incorporé les plasmides.

On utilise pour cela une résistance à un second antibiotique. (tétracycline). L'insertion du fragment d'ADN dans le plasmide doit inactiver le gène de résistance au deuxième antibiotique. Les bactéries transformées par les plasmides recombinants sont donc résistants au premier antibiotique et sensible au deuxième. Les bactéries non transformées sont sensibles au deux antibiotiques.

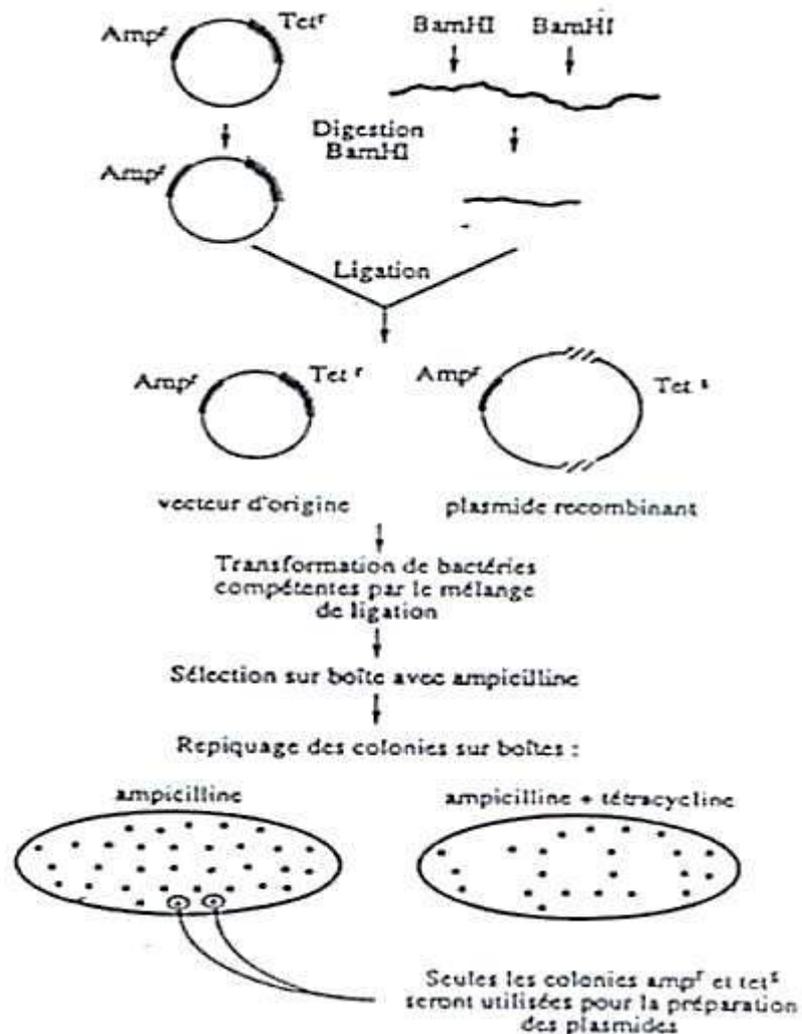


Figure 97 : Criblage par antibiotique.

3.2 La méthode de sélection blanc-bleue

L'inactivation insertionnelle du gène lac Z* peut être utilisé pour détecter les recombinants sur une plaque contenant de l'IPTG et du X-Gal. Le X-Gal est converti en produit bleu si le gène lac Z' est intact et induit par le IPTG, ainsi les recombinants forment des colonies blanches. Les colonies blanches ne possèdent pas de β-galactosidase et contiennent donc probablement le fragment ciblé inséré. Les colonies bleues contiennent probablement le vecteur religaturé.

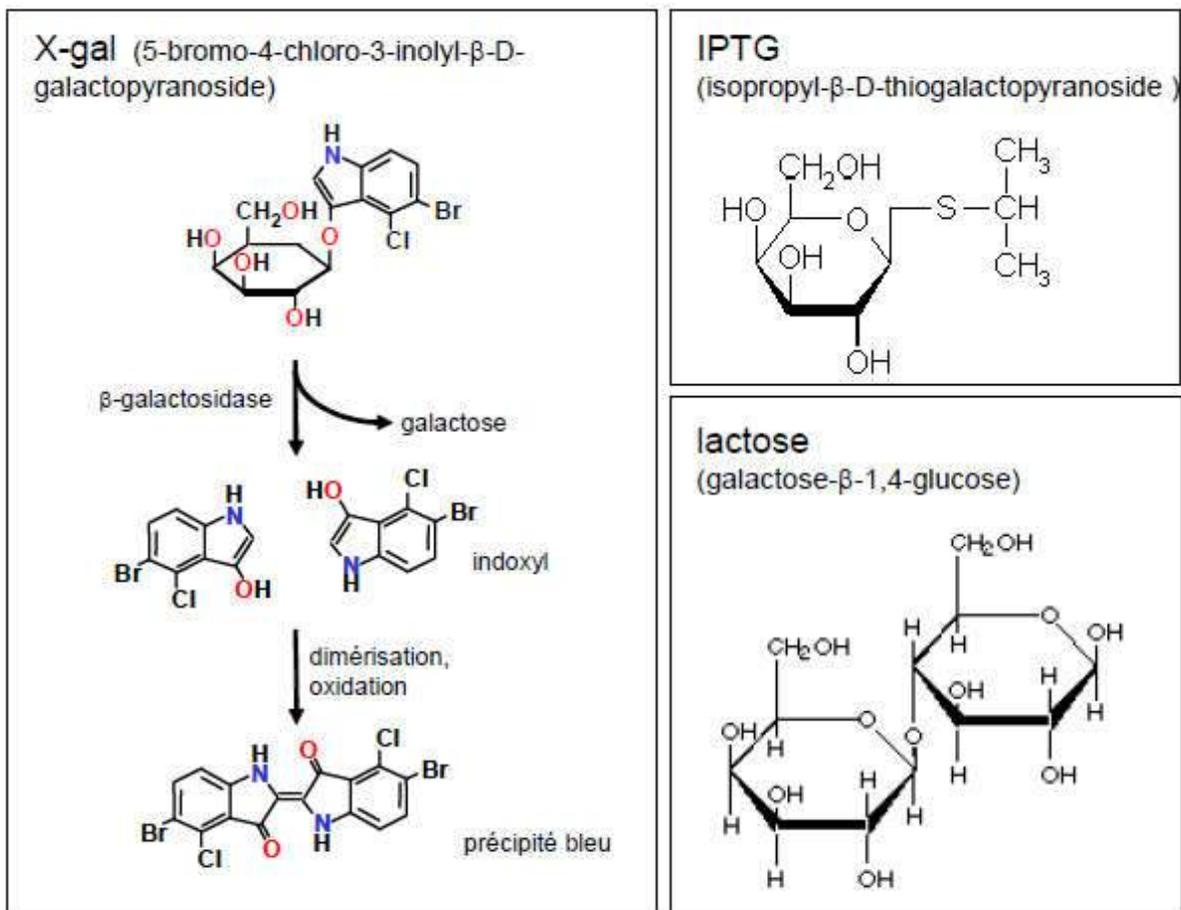


Figure 98 : X-Gal et IPTG : analogues du lactose.

3.3 Criblage par hybridation

A la suite d'une transformation, il est possible de rechercher les colonies ayant intégré un plasmide recombinant par hybridation avec une sonde nucléotidique homologue du gène d'intérêt. Une réplique des colonies sur boîte est effectuée sur un filtre de nylon ou une membrane de nitrocellulose. Après dénaturation, une hybridation avec une sonde nucléotidique radioactive complémentaire d'une séquence de l'ADN recherché est réalisée. L'autoradiographie révèle les colonies possédant l'ADN recherché. Elles peuvent être récupérées et exploitées.

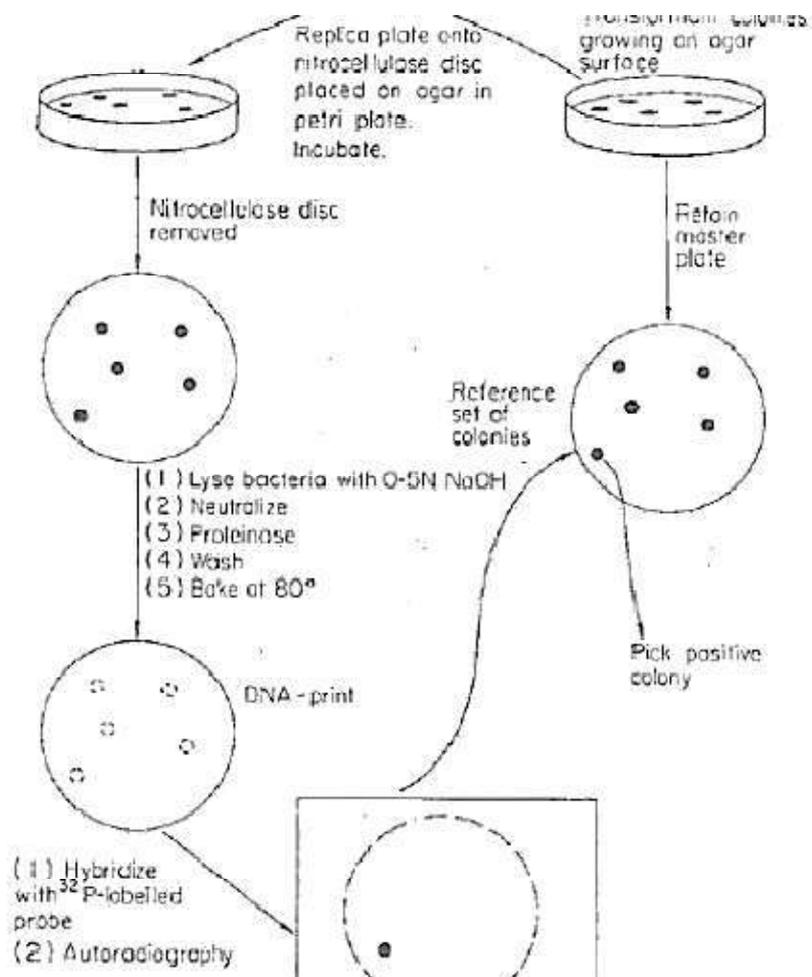


Figure 99 : Criblage par hybridation sur colonies.

3.4 Criblage par complémentation

Le principe consiste à isoler un mutant de levure portant une déficience intéressante. On transforme ensuite ce mutant à l'aide de la séquence d'intérêt. Des répliques sur membrane de nylon sont ensuite réalisées et utilisées pour sélectionner les colonies contenant

le gène de type sauvage permettant de restaurer l'activité déficiente. Exemple : clonage du gène **ARG1** chez la levure.

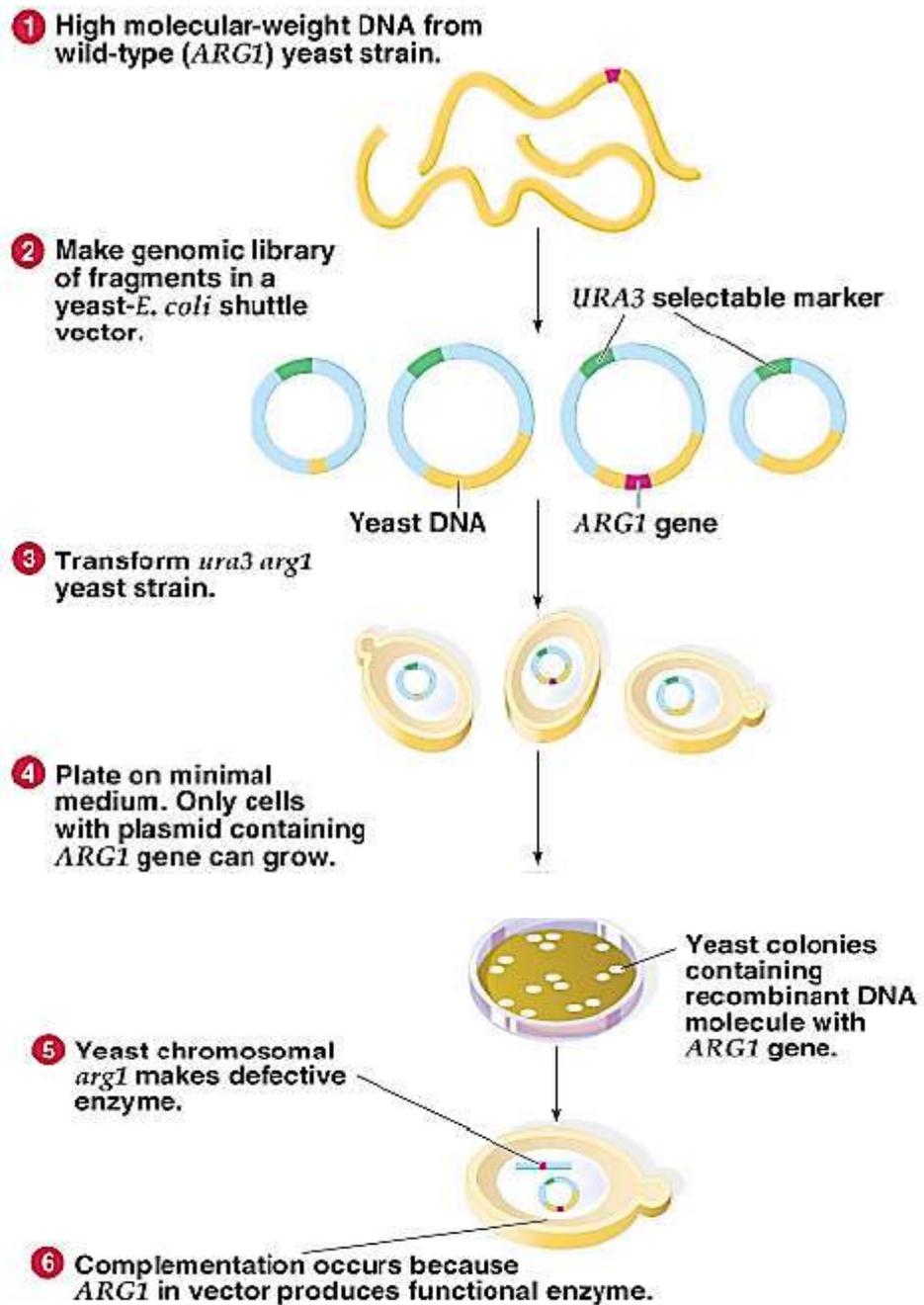


Figure 100 : Criblage par complémentation

3.5 La méthode de coupure avec de enzymes de restriction.

On peut essayer d'ordonner directement l'ensemble des clones d'une banque en comparant leurs profils de restriction après digestion par une enzyme donnée et séparation des fragments sur gel.

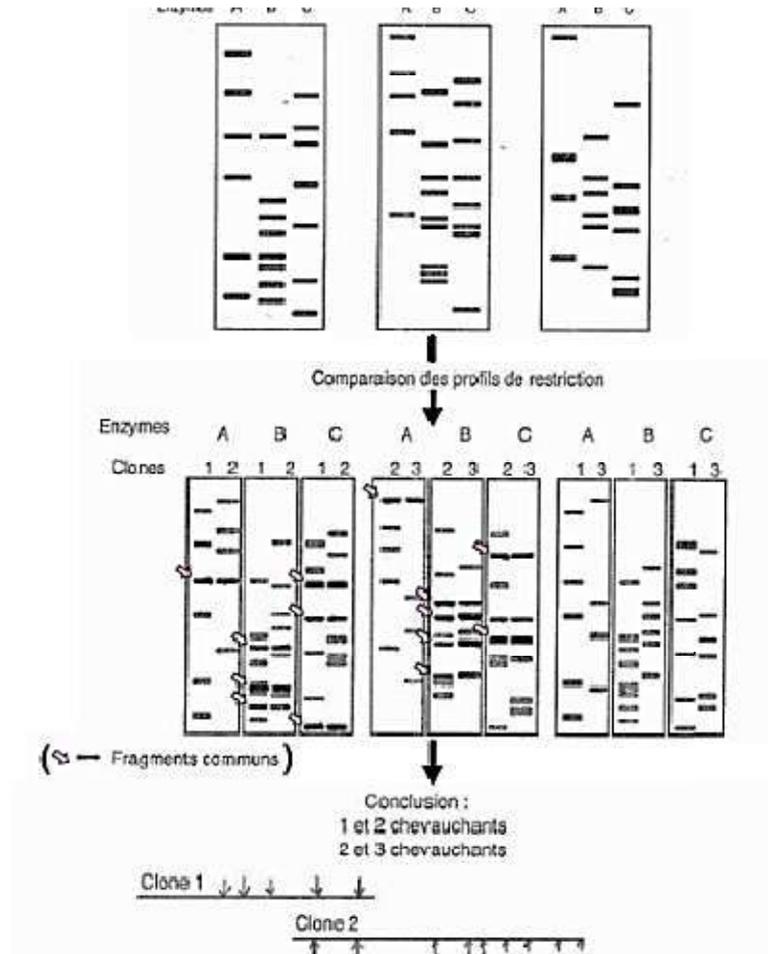


Figure 101 : Criblage à l'aide d'enzymes de restriction.

4. Production d'insuline par génie génétique

La production des protéines humaines est le secteur des biotechnologies le plus rentable.

Dans les cellules de mammifères la quantité de protéine de grand intérêt pharmaceutique comme l'insuline est très faible. Ce qui rend leur extraction et purification extrêmement coûteuse (onéreuse). De plus, même si ces protéines peuvent être produites par des cultures cellulaires; c'est une voie beaucoup plus coûteuse et difficile que la production par une culture microbienne à fort rendement. En conséquence l'industrie de la biotechnologie a mis au point des microorganismes génétiquement modifiés pour produire ces protéines.

La forme active de l'insuline est composée de deux polypeptides (A et B) reliés par des ponts disulfures, ces parties sont codées par des parties séparées du gène de l'insuline.

Préproinsuline: Polypeptide contenant:

- Une séquence signal impliquée dans l'excrétion de la protéine.
- Les polypeptides A et B de l'insuline active.
- Un polypeptide de connexion absent dans l'insuline active

Pour produire l'insuline utilisant *E. coli* comme cellule hôte. Le gène entier a pu être synthétisée, car cette hormone est une protéine relativement petite, cela permet d'éviter l'obstacle des codons utilisés par des espèces différents (ARNt différents). Le gène obtenu est Préproinsuline inséré en aval d'un promoteur fort d'*E. coli* de telle sorte que le fragment d'insuline soit synthétisé comme partie d'une protéine de fusion.

Pour séparer l'insuline de la protéine de fusion, le gène codant l'insuline est séparé à celui codant la protéine de fusion par le codon ATG (MET), car l'insuline ne contient pas de méthionine. Cela permet de séparer les deux protéines utilisant le bromure de cyanogène, sans endommager la proinsuline.

L'insuline est obtenue à partir de la proinsuline par la formation des ponts disulfure. La proinsuline se replie naturellement en amenant face à face les résiduels cystéiniques impliqués dans la formation de ces ponts. Puis l'élimination du peptide de connexion par deux protéases (la trypsine et la carboxy peptidase B) n'ayant aucun effet sur le produit final.

Références bibliographiques

Ameziane N, Bogard M, Lamoril J. (2005). Principes de biologie moléculaire en biologie clinique. Collection Campus Référence Elsevier.

Brown T. A. (2010). Gene Cloning and DNA Analysis: An Introduction. 6th Ed, Wiley Blackwell. London, UK. 782p.

Carter M, Shieh J. (2015). Molecular Cloning and Recombinant DNA Technology. Guide to Research Techniques in Neuroscience (Second Edition), 219-237.

Clark D. (2005). Molecular Biology: Understanding the Genetic Revolution. Southern Illinois University. USA Elsevier Academic Press. 784p.

Dorado D, Besnard G, Unver T, Hernández P. (2015). Polymerase Chain Reaction (PCR). *Reference Module in Biomedical Sciences*.

Gubta P. K. (2005). Microbiology, Cell Physiology and Biotechnology. 2nd Ed. National Offset Printers, Meerut, India. 311p.

Housset C. et A. Raisonnier. (2006). Biologie Moléculaire. Biochimie PCEM1 Université Paris-VI. 204p.

Lamoril J, Ameziane N, Deybach JC, Bouizegarène P, Bogard M. (2008). Les techniques de séquençage de l'ADN : une révolution en marche. Première partie DNA sequencing technologies: A revolution in motion. Part one. Immuno-analyse et biologie spécialisée : 23, 260—279.

Perrin-Schmitt F. (2011). Biotechnologies et applications (génie génétique), ed. Ellipses Marketing.

Primrose S, Twyman R-M, Old R. (2004). Principes de génie génétique, ed. De Boeck.

Prescott M. L., J. P. Harley, et D. A. Klein. (2007). Microbiologie. 2eEd. Edition De Boek Université. Paris. 1137p.

Sambrook, J., and D. Russell. (2001).Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory. New York. USA. 2344p.

Schmitt RD. (2005). Atlas de poche de Biotechnologie et de Génie Génétique, ed. Médecine Sciences Publications.

Robert O. (2010). Clonage et OGM, ed. Larousse.

WatsonJ., T. Baker, S. Bell, A. Gann, M. Levine, R. Losick. (2009). Biologie Moléculaire du gène. 6eEd. Pearson Education- Paris-France. 688p.