

Cours de Biologie Moléculaire et Génie Génétique
Pour les étudiants en Licence Microbiologie

Pr. Mohamed Sifour
Université de Jijel
2024-2025

Contenu de la matière

Partie I : Biologie moléculaire :

1. Expression de l'information génétique: synthèse protéique (Transcription, Traduction).

2. Régulation de l'expression génique : Régulation transcriptionnelle, Régulation traductionnelle.

3. Techniques de base de la biologie moléculaire :

- préparation des acides nucléiques (extraction et purification)
- séparations des acides nucléiques (électrophorèse sur gel d'agarose, en champ pulsé,.....).
- détection, caractérisation et identification des acides nucléiques (transfert sur membrane, marquage, hybridation...).
- Le séquençage de l'ADN.
- amplification in vitro des acides nucléiques (PCR, RT (reverse-transcriptase)-PCR ...).

Partie II : génie génétique : **1. clonage in vivo :** **1.1. Éléments nécessaires au clonage :** l'ADN à cloner, enzymes de restriction, enzymes de ligation, les vecteurs de clonage, leur construction et leurs caractéristiques, les cellules hôte. **1.2. Les étapes du clonage :** construction du vecteur, insertion de l'ADN à cloner, transformation des bactéries, sélection des recombinants, analyse des recombinants. **2. Technologie de l'ADN recombinant :** Synthèse de protéines recombinantes, ADNc et vecteurs d'expression. Exemple de production de protéines par *E. coli* et par *Saccharomyces cerevisiae*.

Sommaire

Partie I : Biologie moléculaire

1- Rappel

1-1- Structure et fonction des gènes

1-1-1- Structure des acides nucléiques (ADN et ARN)

1-1-2- Propriétés physico-chimiques des acides nucléiques

1-1-3- L'organisation de l'ADN dans les cellules

1-1-4- Structure de gènes

1-2- La réplication de l'ADN

1-2-1- La réplication chez les procaryotes

1-2-2- La réplication chez les eucaryotes

1-3- Réparation de l'ADN

2- Expression de l'information génétique

2-1- Le code génétique

2-2- La transcription

2-2-1- La transcription chez les procaryotes

2-2-2- La transcription chez les eucaryotes

2-3- La traduction

2-3-1- ARN de transfert

2-3-2- Les ribosomes et ARN ribosomaux

2-3-3- Les étapes de la traduction

2-3-4- Les modifications post-traductionnelles

3- Régulation de l'expression des gènes

3-1- Régulation chez les procaryotes

3-2- Régulation chez les eucaryotes

4. Techniques de base de la biologie moléculaire :

3-1 Outils et techniques de la biologie moléculaire

3-2- Synthèse d'oligonucléotides, Les enzymes de la biologie moléculaire, vecteurs de clonage ...

3-3- Electrophorèse, PCR, Hybridation et sondes nucléiques, Séquençage de l'ADN ...

Partie II : génie génétique

1- Génie génétique (définition, historique...)

2- Les banques d'ADN

3- Clonage (transformation, criblage, clonage dans les cellules eucaryotes, expression des gènes dans les cellules procaryotes et eucaryotes

4- les applications du génie génétique

Partie I : Biologie moléculaire :

1- Rappel

1-1- Structure et fonction des gènes

1-1-1- Structure des acides nucléiques (ADN et ARN)

Structure de l'ADN

L'acide désoxyribonucléique (ADN) est un polymère contenant des monomères appelés nucléotides (un polynucléotides).

Un nucléotide se compose de :

- **Un sucre** : Un pentose (à cinq carbones) appelé 2'-désoxyribose car le groupement OH sur le deuxième carbone du ribose est remplacé par un atome d'hydrogène.

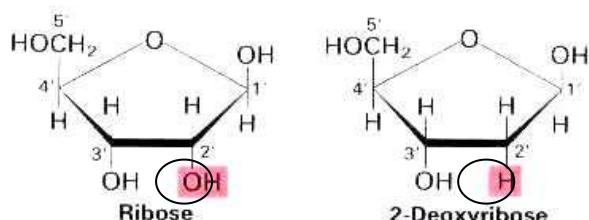


Figure 1 : Ribose et désoxyribose

- **Une base azotée**

Une base purique (l'adénine et la guanine) ou pyrimidique (la cytosine et la thymine)

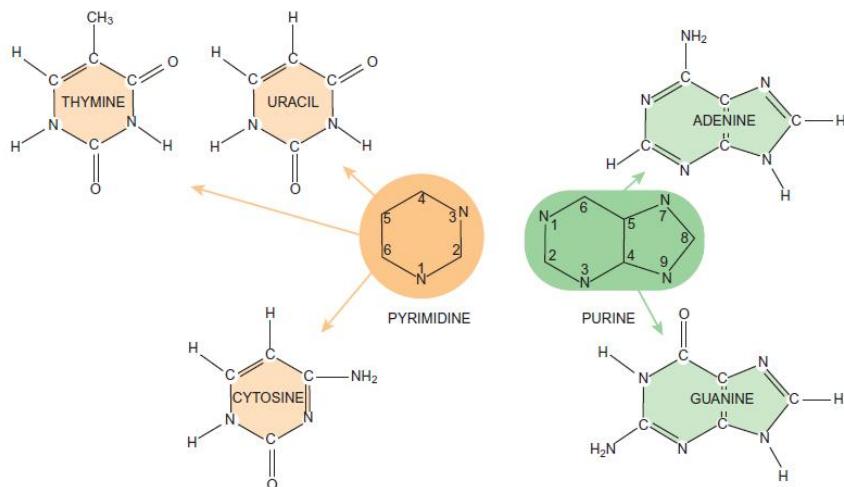


Figure 2: Principales bases azotées. Les principales bases azotées entrant dans la composition des acides nucléiques se répartissent en deux familles chimiques: bases puriques et pyrimidiques. L'uracile est une base pyrimidique spécifiquement trouvée dans l'ARN.

The Bases of the Nucleic Acids: The four bases of DNA are adenine, guanine, cytosine, and thymine. In RNA, uracil replaces thymine. Pyrimidine bases contain one-ring structures, whereas purine bases contain two-ring structures (Clark et al., 2019).

- **Un groupement phosphate (PO_4)**

- Les bases sont attachées au sucre par une liaison entre :

- Le carbone 1' du sucre et l'azote 9 des bases puriques
- Le carbone 1' du sucre et l'azote 1 des bases pyrimidiques

- L'association d'un sucre et d'une base est appelée nucléoside

Un nucléotide contient un nucléoside et une ou plusieurs molécules d'acide phosphorique. Les groupements phosphate sont liés au carbone 5' du sucre. Dans la cellule, on rencontre les nucléotides à

l'état libre (ils jouent un rôle important de transporteurs de l'énergie pour les réactions enzymatiques) ou polymérisés sous forme d'ADN ou d'ARN).

Les acides désoxyribonucléotides sont de très grandes molécules généralement composées de deux chaînes polynucléotidiques enroulées l'une autour de l'autre pour former une double hélice d'un diamètre de 2 nm.

Le phosphate en 5' d'un nucléotide forme une liaison avec le carbone en 3' du nucléotide suivant ; la réaction conduit à l'élimination d'un groupement OH au niveau du carbone en 3'. La liaison est appelée 3'-5' phosphodiester. La chaîne polynucléotidique possède un 5' phosphate libre à une extrémité et un groupement 3' hydroxyle libre à l'autre extrémité.

La partie sucre et phosphate de la molécule d'ADN constitue sa colonne vertébrale et est située à l'extérieur de l'hélice. Les bases azotées se font face au centre de l'hélice et sont empilées les unes sur les autres.

La double hélice effectue un tour toutes les 10 paires de bases environ. Le pas de l'hélice est de 34 Å (3.4 nm).

La double hélice a une structure antiparallèle : l'un des brins est orienté dans le sens 5' → 3' et l'autre dans le sens 3' → 5'.

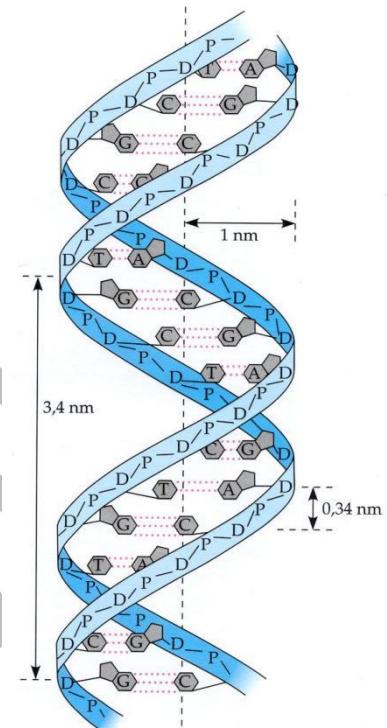
- L'adénine (A) est toujours associée à la thymine (T) par deux liaisons hydrogène
- La guanine (G) est toujours associée à la cytosine (C) par trois liaisons hydrogène

Cette appariement de bases AT et GC implique que les deux brins dans une double hélice d'ADN sont complémentaires.

Rapport de Chargaff (Chargaff's rules): égalité des concentrations de A et de B, d'un côté, de G et de C, de l'autre, d'où $A/T=G/C=1$ et $(A+G)/(T+C)=1$.

Le pourcentage GC est caractéristique de l'espèce.

Il existe un grand nombre de formes d'ADN, A, B, Z... la forme présente de façon majoritaire dans les cellules est appelée ADN B.



Structure de l'ARN

Il existe plusieurs différences importantes entre l'ADN et l'ARN :

- Dans l'ARN, le ribose remplace le 2'-désoxyribose de l'ADN (figure 1).
- La thymine est remplacée par l'uracile (U) (aussi complémentaire de l'adénine).
- L'ARN est monocaténaire (simple brin). Un brin d'ARN peut se replier sur lui-même par appariement des bases complémentaires pour former une structure en forme d'épingle à cheveux.
- Il existe trois principaux types d'ARN : l'ARN messager, l'ARN ribosomal et l'ARN de transfert.

1-1-2- Propriétés physico-chimiques des acides nucléiques (voir TD)

1-1-3- L'organisation de l'ADN dans les cellules

L'ADN semble être beaucoup plus organisé dans la chromatine des eucaryotes et est associée avec plusieurs protéines, les plus importantes étant les histones. Les histones sont de petites protéines basiques

riches en acides aminés, lysine et/ou arginine (chargés positivement). Les charges positives leur permettent de s'associer fortement à l'ADN chargé négativement (les charges négatives du phosphore de l'ADN). Les cinq types d'histoires chez presque toutes les cellules eucaryotes: H1, H2A, H2B, H3 et H4. Le nucléosome est composé de huit molécules d'histones (deux de H2A, de H2B, de H3, et de H4) associés à environ 200 pb d'ADN.

L'association de l'ADN et de protéines qui constitue les chromosomes est appelée chromatine. La molécule d'histone H1 participe à la compaction d'ADN. Les protéines non histones (acides) sont mises en jeu dans plusieurs aspects du fonctionnement de la chromatine, notamment la régulation de l'expression des gènes.

Bien que l'ADN existe en double hélice à la fois chez les procaryotes et les eucaryotes, son organisation est différente dans les deux types cellulaires. L'ADN est organisé en cercle fermé chez presque tous les procaryotes (le chromosome du *Borrelia* est fait d'une molécule d'ADN linéaire). Cette double hélice circulaire est de plus enroulée en superhélice et est associée à des protéines basiques et non pas avec les histones complexées à presque tous les ADN eucaryotes. Ces protéines qui ressemblent à des histones, ont apparemment pour fonction d'organiser l'ADN bactérien en une structure enroulée de type chromatine.

1-1-4- Structure de gènes

L'information biologique est codée par la séquence des bases azotées de l'ADN, elle est organisée en un grand nombre de gène dont chacun contient les informations nécessaires à la synthèse d'un peptide.

Un gène est une unité d'information et il correspond à un segment d'ADN qui code la séquence des acides aminés d'une protéine.

Un gène peut être aussi défini comme une séquence polynucléotidique codant pour un polypeptide, un ARNt ou un ARNr. Les gènes ont des tailles très variables : de 100 paires de bases à plusieurs millions de paires de bases.

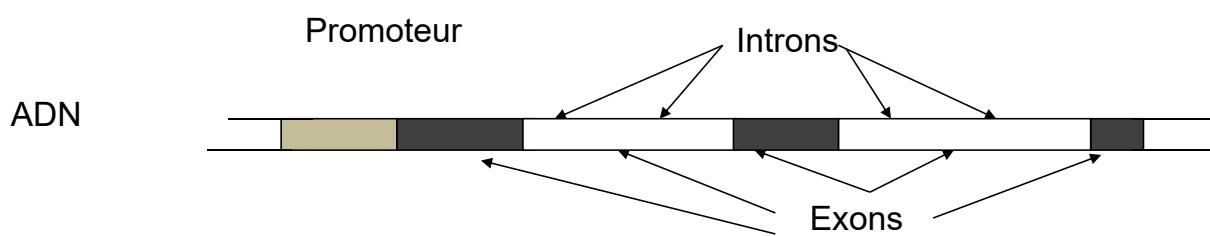
Introns et exons

Les gènes des procaryotes ont une structure très différente de celle des gènes des eucaryotes. Dans les gènes bactériens et viraux, l'information codante est habituellement continue (il y a des gènes bactériens qui contiennent des introns) alors que chez les eucaryotes, les gènes contiennent des séquences codantes (exons) interrompues régulièrement par des séquences non codantes (introns). [Une exception intéressante à cette règle concerne les gènes eucaryotes des histones qui ne possèdent pas d'introns].

Les promoteurs des gènes

L'expression des gènes est réglée par un fragment de la séquence d'ADN présent en amont de la séquence codante: le promoteur

Des séquences conservées d'ADN du promoteur sont reconnus par l'ARN polymérase qui s'y fixe, ainsi que des protéines associées appelées facteurs de transcription et qui permettent la transcription du gène en ARN.



Structure d'un gène

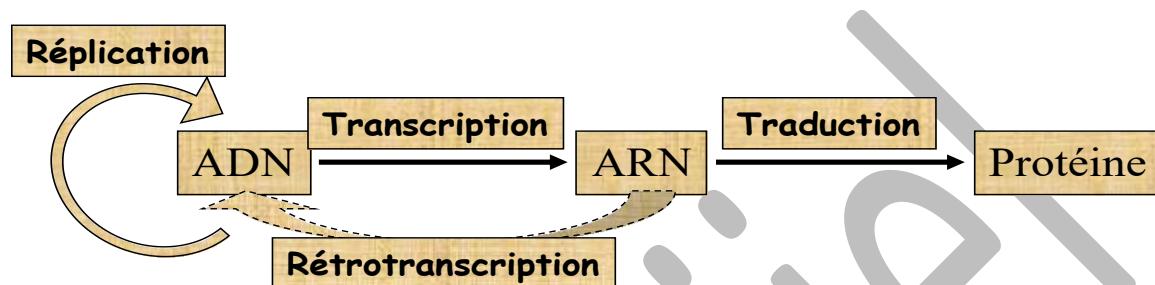
Expression des gènes

L'information biologique codée dans les gènes est rendue disponible lors de l'expression des gènes. Au cours de ce processus, une copie ARN du gène est synthétisée et elle dirige ensuite la synthèse d'un polypeptide.

L'utilisation de l'information est décrite par le « **dogme central** ». Il pose que le transfert d'information est unidirectionnel : dans le sens ADN → ARN → Protéine

Exception : les rétrovirus possèdent une enzyme appelée transcriptase inverse (reverse transcriptase) capable de synthétiser une copie d'ADN d'un fragment d'ARN].

Le fonctionnement d'une cellule dépend de l'activité coordonnée de nombreuses protéines. L'expression des gènes assure que les protéines sont synthétisées au bon endroit et au bon moment.



1-2- La réplication

A- La réplication est semi-conservative

La réplication est le processus par lequel une cellule copie son ADN. La réplication est nécessaire pour que l'information génétique présente dans les cellules puisse être transmise aux cellules filles après la division cellulaire. L'ADN est copié dans le sens 5' → 3' par l'ADN polymérase. Ces enzymes travaillent sur de l'ADN monobrin en synthétisent un nouveau brin complémentaire du brin original.

La réplication est dite semi-conservative car, dans chaque molécule fille d'ADN, un brin provient de la molécule mère, et un autre brin est nouvellement synthétisé. La réplication fournit deux molécules d'ADN identiques entre elles et identiques à la molécule mère.

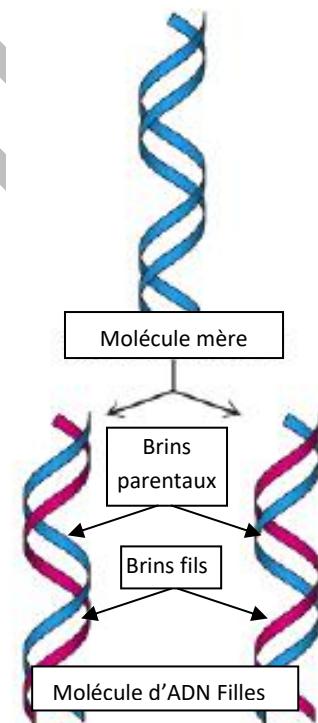


Figure 6 : la réplication semi-conservative de l'ADN

B- La topologie de l'ADN

Le super-enroulement de l'ADN est physiologiquement fondamental aussi bien sur le plan structural que fonctionnel. Les enzymes responsables des super-enroulements sont appelées les topoisomérasées. Ces enzymes sont très conservées et sont présentes chez toutes les espèces avec pour rôle principale celui de contrôler l'état topologique de la molécule d'ADN. Pour cela, ces enzymes doivent pouvoir créer des superenroulements (appelés également supertours) pour modifier le nombre d'enlacements de l'ADN, ce qui ne peut être rendu possible que si ces enzymes coupent l'ADN sur un ou deux brins. On en connaît ainsi deux types :

- **Topoisomérase de type I**: Ces enzymes se fixent au DNA, elle coupe l'un des 2 brins, elles s'attachent ensuite par l'une de leur tyrosine au groupement 5' phosphate libre de l'ADN coupé. L'ADN peut alors se dérouler, l'enzyme répare ensuite la cassure en libérant l'ADN avec des supertour en (-). Cette coupure,

puis la régénération de la liaison initiale après le déroulement de l'ADN, ne consomme pas d'ATP, puisque l'énergie de la liaison phosphodiester rompue est conservée par la liaison avec le résidu de tyrosine de l'enzyme.

Topoisomérase de type II: elles coupent l'ADN sur les 2 brins puis le ressoudent après avoir créé des superenroulement négatifs, ou après avoir supprimé les supertours positifs. Cette action nécessite l'hydrolyse d'une molécule d'ATP

Donc le rôle de ces enzymes est de superenrouler négativement l'ADN pour permettre sa compaction dans la cellule, et permettent des événements majeurs de la vie cellulaire comme la transcription et la réPLICATION (dérouler les deux brins).

Exemple : La gyrase bactérienne (inhibé par les antibiotiques).

C- Les polymérases

Ces enzymes sont capables de synthétiser le brin d'ADN complémentaire d'un brin d'ADN parental servant de matrice (template). Elles catalysent la formation des liaisons ester entre le groupement 5' phosphate d'un dNTP libre et le groupement 3' hydroxyle libre du brin d'ADN en formation. La polymérisation nécessite en outre la présence d'une amorce (primer), séquence d'ARN (10 à 40 N) hybride au brin à copier et synthétisée par une ARN polymérase (primase).

La polymérisation doit être dirigée par une matrice, en l'occurrence le brin de DNA à copier, selon les règles de l'appariement des bases de Watson-Crick (A = T, G = C). Les polymérases ajoutent des nucléotides au brin en cours de formation en se déplaçant le long de la matrice, puis elles s'en dissocient.

L'ADN polymérase, connue actuellement sous le nom de DNA polymérase I, a été isolée d'*E. coli* par A. Kornberg en 1955. Par la suite, au moins quatre autres polymérases ont été décrites, les polymérases II, III, IV et V. La polymérase III est l'enzyme essentiel pour la synthèse des chaînes polynucléotidiques chez *E. coli*, tandis que la polymérase I joue un rôle dans la synthèse du brin retardé et que les polymérases II, IV et V interviennent dans la réparation du DNA.

Les ADN polymérases possèdent une activité de polymérisation dans le sens 5'--> 3' et une activité exonucléasique dans le sens 3'--> 5'. L'ADN polymérase I possède de plus une activité exonucléasique dans le sens 5'--> 3' (dégradation des amores). La DNA polymérase III est l'enzyme responsable de la synthèse de l'ADN chez *E. coli* ; d'une masse d'environ 800 kDa, elle est constituée de 10 types différents de sous-unités.

La DNA polymérase I n'est pas l'enzyme responsable de la synthèse du DNA, mais elle effectue diverse tâches lors de la réPLICATION, de la réparation ou des recombinaisons du DNA. C'est une chaîne polypeptidique unique de 103 kDa qui possède, outre une activité polymérase et une activité exonucléase de relecture 3'-->5', une activité exonucléase 5'-->3' localisée dans un domaine qui peut être séparé de l'enzyme par un traitement protéasique limité. Lorsque ce domaine est éliminé, il demeure un fragment de 68 kDa, dit fragment de Klenow, qui retient les activités polymérase et exonucléase de relecture.

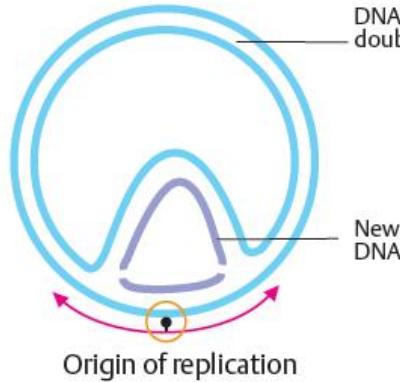
Chez les eucaryotes supérieurs (mammifères), les ADN polymérases, au nombre minimum de cinq. Alpha (α), bêta (β), delta (δ) et epsilon (ϵ) (polymérase nucléaires) assurent l'activité de polymérisation 5'-->3'. Gamma (γ) (polymérase mitochondriale). (δ) et (ϵ) ont une activité d'éxonucléase 3'-->5'.

La réPLICATION démarre à l'intérieur de la molécule d'ADN en un point appelé origine de réPLICATION. Les origines de réPLICATION contiennent généralement des séquences riches en appariements faibles A-T. la région ou' la double hélice est déroulée et du nouvel ADN synthétisé est appelée fourche de réPLICATION (replication fork)

1-2-1- Réplication chez les procaryotes

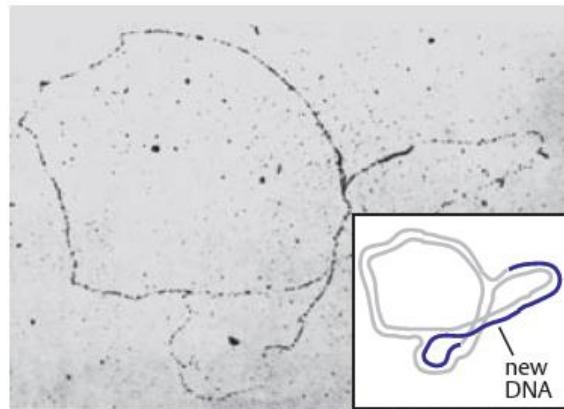
La réplication du DNA se déroule en trois phases : initiation, elongation et terminaison. Chez *E. coli*, l'origine de réplication est un site unique du DNA de 245 pb, dénommé *locus Ori C*. La protéine dnaA induit la séparation des deux brins d'ADN et définit l'origine de réplication puis des protéines dnaB et C (respectivement une hélicase et une primase) vont catalyser la formation d'une fourche de réplication. La séparation des deux fourches s'effectue dans les deux sens (2 fourches de réplication). Une topoisomérase, la DNA-gyrase, élimine les supertours formés par le déroulement du DNA. Ensuite, des protéines SSB (Single-Stranded DNA-Binding Protein) se fixent sur les brins séparés de DNA et maintiennent les deux brins sous une forme monocaténaire indispensable à leur fonction de matrices. Les protéines qui interviennent dans cette étape d'initiation forment le complexe d'initiation, encore appelé primosome.

Réplicon : la portion du génome qui contient une origine et qui se réplique en une seule unité.



1. DNA replication in the bacterial chromosome

A. Prokaryotic replication begins at one site



2. Prokaryotic replication in an autoradiogram in *E. coli* (J. Cairns)

Figure 7: une seule origine de réplication chez les procaryotes. Les flèches indiquent le sens de déplacement de la fourche de réplication (Passarge, 2007).

La formation d'un brin d'ADN complémentaire (élongation) s'opère par additions successives de désoxyribonucléotides à un oligonucléotide en élongation. La polymérisation ne s'opère que dans le sens 5' → 3'. Comme les deux brins d'ADN sont antiparallèles, un des brins en formation, le brin avancé, s'allongera de façon continue dans le sens 5' → 3', sens de déplacement de la fourche de réplication. L'autre brin, dit retardé, se forme de façon discontinue par de petits fragments d'ADN de 1000 à 2000 nucléotides (fragments d'Okazaki) dont la synthèse et à chaque fois amorcé dans le sens 5' → 3' (sens inverse du déplacement de la fourche).

L'ADN polymérase III va s'insérer dans la fourche au niveau de l'amorce d'ARN synthétisée par la primase, puis elle ajoute les désoxyribonucléotides au fur et à mesure du déroulement du DNA par l'hélicase et de l'élimination des supertours par l'ADN gyrase.

Chaque fragment d'Okazaki a une amorce, quand l'ADN polymérase III arrive à l'extrémité 5' de l'amorce précédente, elle est remplacée par la polymérase I qui agit d'abord comme une exonucléase en dégradant l'ARN dans le sens 5' → 3', puis comme une polymérase en remplaçant l'amorce ARN par l'ADN.

Les différents fragments du brin retardataire sont ensuite réunis par l'enzyme ADN ligase (elle va établir une liaison phosphodiester entre une extrémité 5' monophosphate d'un fragment et l'extrémité 3' hydroxyle de l'autre fragment).

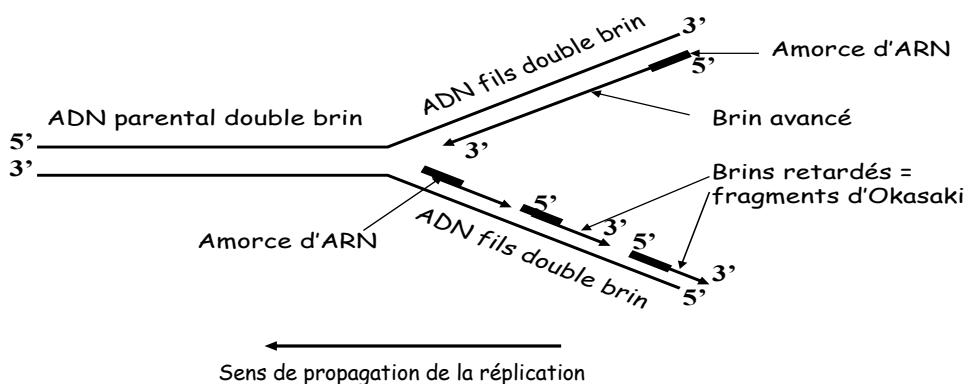


Figure 8: réPLICATION de l'ADN

Chez *E. coli*, la fourche de réPLICATION de l'ADN circulaire s'arrête lorsque le complexe de polymérisation atteint la région de terminaison (Ter).

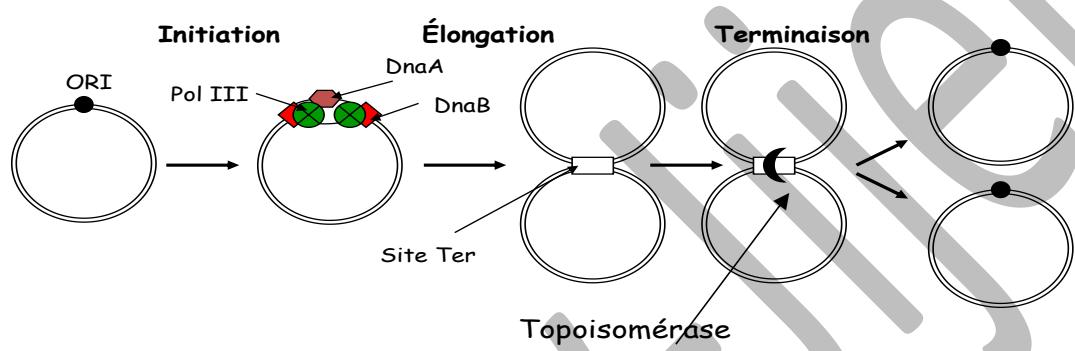


Figure 9: réPLICATION des molécules circulaires d'ADN

NB : un modèle différent de réPLICATION de l'ADN est observé durant la conjugaison chez *E. coli* et la multiplication des virus, tel le phage lambda : mécanisme du cercle roulant (rolling circle).

1-2-2- RéPLICATION chez les eucaryotes

Les grands principes de la réPLICATION sont les mêmes chez les procaryotes et les eucaryotes. La réPLICATION est bidirectionnelle, complémentaire, antiparallèle, simultanée pour chaque brin mais discontinue pour l'un des deux brins, elle s'effectue dans le sens 5'-3', avec des amorces d'ARN. Mais comme le génome est plus grand et réparti sur plusieurs chromosomes, il existe plusieurs sites d'initiation sur chaque chromosome (plusieurs origines).

L'ADN des chromosomes des eucaryotes est emballé sous forme de complexes ADN-protéines (nucléosome). Lorsque la fourche de réPLICATION progresse l'ADN doit être déroulé du nucléosome pour que la réPLICATION puisse avoir lieu. Ceci ralenti la fourche de réPLICATION (et pourrait expliquer la faible longueur des fragments d'Okasaki chez les eucaryotes (100 à 200 bases) en comparaison avec ceux des procaryotes (1000-2000). Après le passage de la fourche de réPLICATION, les nucléosomes se reforment.

Le brin avancé est synthétisé par le complexe ADN polymérase δ/protéine PCNA (proliferating cell nuclear antigen). Le brin retardé, les amorces d'ARN et les fragments d'Okasaki seraient l'œuvre d'une primase et de l'ADN polymérase α. L'élimination des amorces ARN est assurée par une Arase H (hybride), et leur remplacement par l'ADN résulte de l'activité des polymérases α et β.

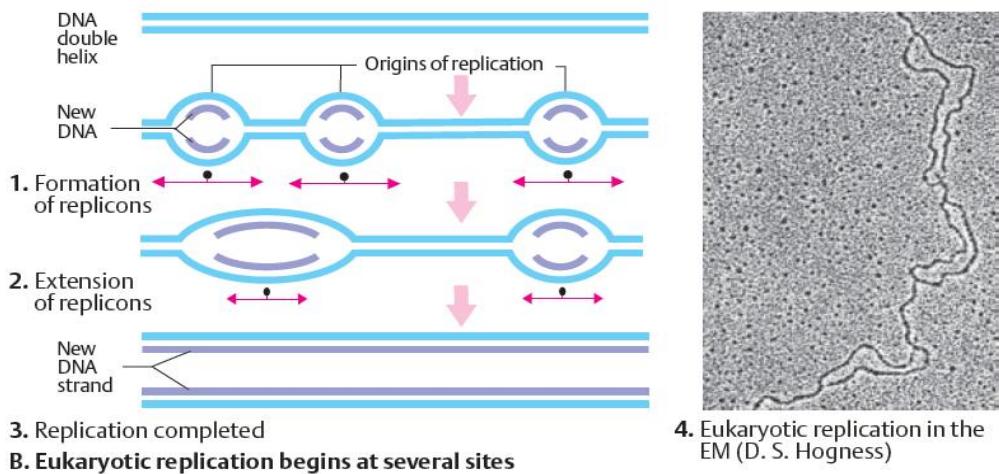


Figure 10: une seule origine de réplication chez les eucaryotes. Les flèches indiquent le sens de déplacement de la fourche de réplication (Passarge, 2007).

La réplication des chromosomes linéaires des eucaryotes pose un problème qui n'existe pas pour les chromosomes circulaires des procaryotes : les deux amorces se situent aux extrémités de chaque brin fils néosynthétisés ne pourront être remplacées. Cela conduirait à un raccourcissement des chromosomes à chaque cycle de réplication.

Le problème est résolu par l'intervention d'une télomérase. Afin de permettre la réplication de ces extrémités, la télomérase leur ajoute des séquences non codantes.

1-3- Réparation de l'ADN

La polymérisation du DNA s'effectue avec une extrême fidélité, moins d'une erreur pour 10^9 nucléotides ajoutés (1 sur 1 milliard). Cette fidélité est due non seulement à la spécifité d'appariement des Bases AT et GC mais aussi à la propriété de la correction d'épreuves de l'ADN polymérase III (au moment de la réplication), cette activité 3'-->5' exonucléase permet à l'enzyme d'éliminer tout nouveau nucléotide qui présenterait un mésappariement de bases.

A- Les mutations (TD)

Les mutations correspondent à des erreurs de copie des bases puriques ou pyrimidiques qui se produisent lors de la réplication de l'ADN. Ces mutations spontanées lors de la réplication sont très peu grâce à l'activité de correction des ADN polymérasées. Il existe cependant des agents qui vont augmenter le taux d'erreurs, appelés agents mutagènes.

Les différents types de mutation:

- Les mutations ponctuelles (des substitutions): Il s'agit du changement d'une base par une autre (point mutation)
 - . Transition: une base purique (ou pyrimidique) remplace une autre
 - . Transversion: une base purique remplace une base pyrimidique, ou inversement.
- Les mutations par insertion /délétion : délétion (des oubli), insertion (des ajouts)

B- Les agents mutagènes

Agents chimiques

Ces agents peuvent produire des modifications chimiques sur certaines bases (désamination de la cytosine en uracile $\text{NH}_2 \rightarrow \text{OH}$)

L'action de l'acide nitreux qui provoque la désamination de l'adénine, ce qui forme de l'hypoxantine (H) apparié avec la cytosine. La paire initiale A-T remplacée par une paire H-C puis GC.

Les radicaux libres

Ions superoxydes (H_2O_2), radicaux hydroxides...

Oxyder G en oxo-guanine qui peut s'apparier avec A ou bien C.

G-C --> Oxo-G-C, après replication : Oxo-G-A --> T-A

Agents intercalent

Ces composés peuvent se glisser entre les plans des bases de l'ADN et provoquer des délétions ou des insertions (Bromure d'ethidium, proflavine...)

Agents analogues de bases

5 bromo-uracile qui se comporte comme la T et s'apparie donc avec le A; mais l'équilibre de tautomérisation la transforme en forme énol, qui a la réplication suivante:

S'apparie avec G, puis elle-même avec C.

T-A --> C-G

Agents physiques

Les radiations (rayons X, UV qui sont absorbées par l'ADN, notamment au niveau des bases pyrimidiques)

C- Réparation

Réparation directe

Il existe des enzymes capables de reconnaître les bases modifiées, de retirer ces modifications et de revenir à des bases normales.

- La guanine peut être transformée en O⁶-methylguanine (par des agents alkylants) elle peut s'apparier par erreur avec la thymine. La réparation de cette méthylation s'effectue de manière directe (sans coupure de l'ADN) par O⁶-methylguanine méthyltransférase, qui transfère le groupement en question de la base vers un de ces acides aminés.

- La photoréactivation: qui répare les dimères pyrimidiques produit par les radiations UV. ADN photolyase induite par la lumière visible répare les dimères pyrimidiques en cassant les liens qui se forment lors de la dimérisation.

Réparation par excision

La réparation des dimères pyrimidiques (dimères de thymines) peut s'effectuer par un système de réparation spécifique (nucleotide excision repair). L'enzyme UvrABC (chez *E. coli*) peut reconnaître la déformation de l'ADN provoquée par le dimère de thymine et coupe l'ADN en amont ou en aval de celui-ci. L'ADN polymérase comble l'espace manquant selon le modèle du brin parental. La ligase relie le fragment réparé et le reste de l'ADN.

Réparation des mésappariement (Mismatch Repair)

Ce système corrige les erreurs introduites durant la réplication de l'ADN et n'ont pas été corrigés par l'activité de l'ADN polymérase. La réparation de ce défaut nécessite la reconnaissance à la fois du brin parental et du brin néoformé. Chez *E. coli*, le brin parental est facilement identifié puisqu'il est marqué avec des adénines méthylées. Le brin naissant porteur de l'erreur est ainsi reconnu puis coupé et le fragment d'ADN incorrect est ensuite éliminé. Il est ensuite remplacé grâce à l'action combinée de plusieurs protéines (Mut H, Mut S, Mut L, SSB, exonucléase I, ADN polymérase III et Ligase).

Réparation par recombinaison

Il arrive qu'au moment de la duplication de l'ADN la fourche de réplication rencontre sur un des brins matrice une lésion ancienne qui n'a pas pu être réparée (dimère de thymine...). L'information génétique est donc perdue puisque les deux brins parentaux sont déjà séparés. Chez *E. coli*, RecA est la protéine qui assure la réparation par recombinaison. Elle va utiliser la deuxième hélice d'ADN intacte pour reconstituer celle qui est lésée et combler la lacune sur le brin fils en face du dimère de thymine. Ceci crée donc une lacune située cette fois sur le brin parental de l'autre double hélice, mais qui possède un brin fils normal, à partir duquel une polymérase pourra combler la lacune. Sur l'autre double hélice, les systèmes d'excision-resynthèse pourront enlever le dimère de thymine et combler la lacune car le brin fils est redevenu normal.

Expression des gènes SOS (SOS repair)

L'arrêt de la réPLICATION (quand les erreurs sont graves) provoque l'apparition d'ADN monocaténaire reconnu par l'enzyme de réparation RecA. Celle-ci opère en assurant des recombinaisons et en induisant l'autolyse du répresseur des gènes SOS, la protéine LexA. Cette enzyme induit ainsi l'expression des gènes SOS (recA, lexA, uvrA, uvrB, et d'autres gènes).

2- Expression de l'information génétique

2-1- le code génétique

Les acides aminés sont codés par 64 triplets de bases appelés codons.

Codon = triplet de nucléotides (ex : AGC = Ser)

Chaque protéine est codée par un gène (ou plusieurs si la protéine est polymérique). L'unité de base des protéines est l'acide aminé. Chaque acide aminé est codé par un codon.

Propriétés du code génétique

➤ Dégénéré : Un acide aminé est codé par plusieurs codons

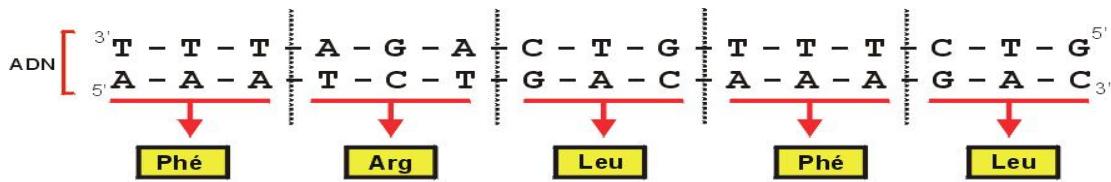
La plupart des acides aminés sont désignés par plus d'un codon (à l'exception de la méthionine et du tryptophane). La dégénérescence du code génétique aide à minimiser les effets des mutations. Le code d'initiation, qui code la méthionine, est AUG. Il y a trois codons stop : UAG, UGA et UAA.

➤ Universel : Identique chez tous les êtres vivants (exception : ADN mitochondrial et certains organismes unicellulaires)

➤ Non chevauchant: Les codons sont lus en série depuis un point de départ jusqu'au codon stop.

Tableau 1 : le code génétique standard

Second nucleotide in codon																
First nucleotide in codon (5' end)		U		C		A		G		Third nucleotide in codon (3' end)						
		U	C	A	G											
U	UUU	Phe	F	Phenylalanine	UCU	Ser	S	Serine	UAU	Tyr	Y	Tyrosine	UGU	Cys	C	Cysteine
	UUC	Phe	F	Phenylalanine	UCC	Ser	S	Serine	UAC	Tyr	Y	Tyrosine	UGC	Cys	C	Cysteine
	UUA	Leu	L	Leucine	UCA	Ser	S	Serine	UAA	Termination				UGA	Termination	
	UUG	Leu	L	Leucine	UCG	Ser	S	Serine	UAG	Termination				UGG	Trp	W
C	CUU	Leu	L	Leucine	CCU	Pro	P	Proline	CAU	His	H	Histidine	CGU	Arg	R	Arginine
	CUC	Leu	L	Leucine	CCC	Pro	P	Proline	CAC	His	H	Histidine	CGC	Arg	R	Arginine
	CUA	Leu	L	Leucine	CCA	Pro	P	Proline	CAA	Gln	Q	Glutamine	CGA	Arg	R	Arginine
	CUG	Leu	L	Leucine	CCG	Pro	P	Proline	CAG	Gln	Q	Glutamine	CGG	Arg	R	Arginine
A	AUU	Ile	I	Isoleucine	ACU	Thr	T	Threonine	AAU	Asn	N	Asparagine	AGU	Ser	S	Serine
	AUC	Ile	I	Isoleucine	ACC	Thr	T	Threonine	AAC	Asn	N	Asparagine	AGC	Ser	S	Serine
	AUA	Ile	I	Isoleucine	ACA	Thr	T	Threonine	AAA	Lys	K	Lysine	AGA	Arg	R	Arginine
	AUG	Met	M	Methionine	ACG	Thr	T	Threonine	AAG	Lys	K	Lysine	AGG	Arg	R	Arginine
G	GUU	Val	V	Valine	GCU	Ala	A	Alanine	GAU	Asp	D	Aspartic acid	GGU	Gly	G	Glycine
	GUC	Val	V	Valine	GCC	Ala	A	Alanine	GAC	Asp	D	Aspartic acid	GGC	Gly	G	Glycine
	GUU	Val	V	Valine	GCA	Ala	A	Alanine	GAA	Glu	E	Glutamic acid	GGA	Gly	G	Glycine
	GUG	Val	V	Valine	GCG	Ala	A	Alanine	GAG	Glu	E	Glutamic acid	GGG	Gly	G	Glycine



Phase de lecture

A partir d'une séquence quelconque, suivant la base choisie pour commencer un codon, il est possible de lire trois ensembles de codons. Chaque ensemble de codons est appelé une phase de lecture. C'est le codon d'initiation qui détermine la phase de lecture d'une séquence qui code une protéine. Les autres phases de lecture ont tendance à comporter des codons stop et elles ne sont pas utilisées. Une phase ouverte de lecture (Open Reading Frame : ORF) est une séquence de codons bornée par un codon start et un codon stop.

NB : Les mitochondries sont dotées d'un petit génome d'ADN qui contient environ 20 gènes pour lesquels on peut observer des déviations par rapport au code génétique standard.

2-2- La transcription

C'est la première étape de l'expression génétique et elle met en jeu la synthèse par RNA polymérase un ARN à partir d'une matrice d'ADN. L'ARN est synthétisé à partir du brin matrice (transcrit, antisens) et a donc la même séquence que le brin non matrice (codant, sens, informatif). La polymérisation se fait dans le sens 5'-->3' donc lecture du brin d'ADN dans le sens 3'-->5'.

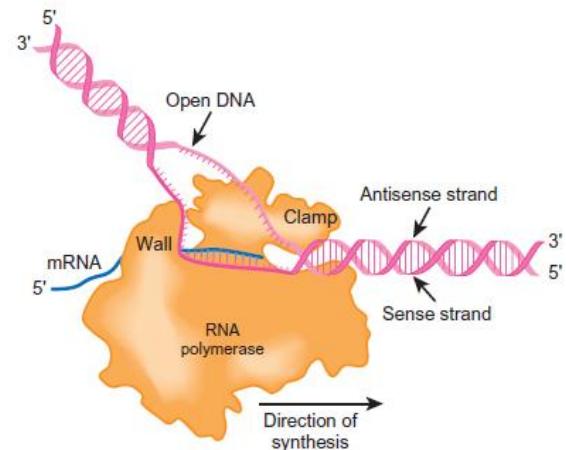


Figure 12: Naming the Basic Components Involved in Transcription. The DNA is shown in its doublehelical form.

After local separation of the strands, the new RNA is synthesized so that it base pairs with one of the DNA strands—the antisense or template strand. The other DNA strand is inactive and is called the sense or coding strand. The enzyme RNA polymerase synthesizes single-stranded RNA in the 5' to 3' direction. The sequence of bases in the RNA is the same as in the sense strand of DNA and complementary to the antisense strand of DNA (except that uracil substitutes for thymine) (Clark et al., 2019).

2-2-1- La transcription chez les procaryotes

La transcription s'effectue chez les bactéries grâce à une seul enzyme, l'RNA polymérase. C'est une enzyme de grande taille (PM 450000) contenant 5 sous-unités : 2 α , β , β' , et σ . Cette forme ($\alpha_2 \beta \beta' \sigma$) est appelée holoenzyme. La sous-unité σ peut se dissocier des autres sous-unités pour laisser une forme ($\alpha_2 \beta \beta'$) appelée enzyme centrale (le cœur) (core enzyme).

La transcription comporte trois étapes essentielles.

Initiation: Chez toutes les bactéries, le démarrage de la transcription est conditionné par la reconnaissance de séquences spécifiques de l'ADN appelées sites promoteurs. La séquence promoteur procaryote n'est pas transcrive. Dans les promoteurs d'*E. coli*, on trouve une séquence de six bases (TTGACA), la séquence -35 (35 bases en amont du site d'initiation), et aussi une séquence TATAAT boîte de Pribnow (TATA Box) qui se situe environ 10 bases en amont du site d'initiation (région -10). La sous-unité σ de l'ARN polymérase est responsable de la reconnaissance des séquences -10 et -35 et de la liaison au promoteur. Une fois l'ARN polymérase fixées à la région -35, les deux brins d'ADN se séparent temporairement autour de la région -10

autorisant le début de la transcription. Le premier nucléotide utilisé par l'ARN polymérase est souvent une purine.

L'elongation : Après l'allongement des 90 premiers nucléotides, le facteur σ se détache de l'ARN polymérase pour aller assurer un nouveau démarrage de la transcription. La polymérase poursuit l'allongement de la chaîne de l'ARN par l'addition des ribonucléotides. L'ARN en cours de formation reste lié à l'ADN sur une longueur d'environ 12 nucléotides compris dans la bulle formée par séparation des deux brins d'ADN. A mesure que l'ARN polymérase avance la partie proximale de la molécule d'ARN se détache de l'ADN permettant ainsi à la double hélice d'ADN de se reconstituer.

La terminaison : peut s'accomplir de deux manières :

- par la reconnaissance d'une séquence de l'ADN appelée terminateur. Il s'agit d'une séquence palindromique riche en G-C suivie d'une région riche en adénine. Quand cette séquence est transcrise, l'ARN forme un appariement intrabrin en épingle à cheveux qui contribue à détacher l'ARN polymérase en libérant la molécule d'ARN appelée aussi transcript primaire.
- un autre mécanisme d'arrêt de la transcription complémentaire du premier est parfois assuré par une protéine spécifique, le facteur ρ (rho) qui interagit avec l'hybride ADN/ARN et le dissocie en consommant de l'ATP.

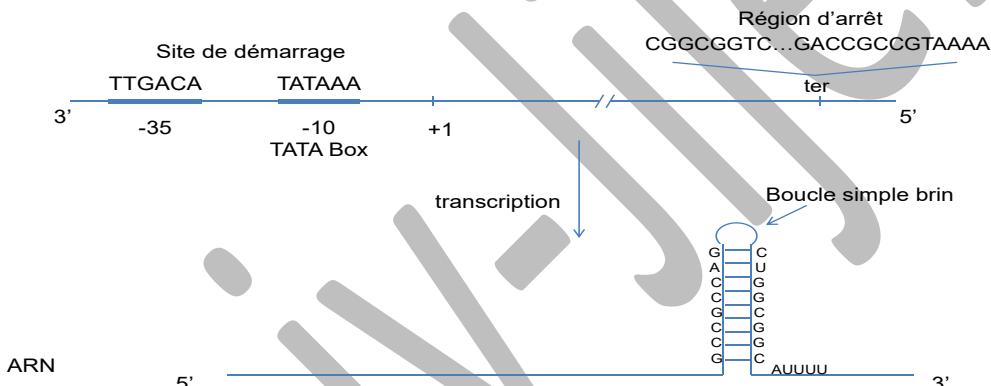


Figure 13 : Séquences consensus des promoteurs et terminateurs chez les procaryotes

Les ARN procaryotes sont dits polycistroniques car ils codent généralement plus d'une chaîne polypeptidique. Une molécule d'ARNm polycistronique peut contenir des espaces intercistroniques appelés espaces.

Exemple : l'opéron lactose

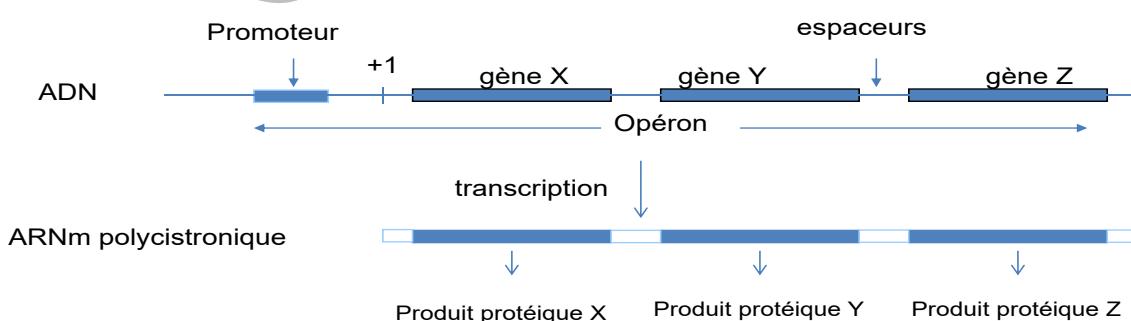


Figure 14 : ARNm polycistronique: dans un opéron bactérien, un promoteur contrôle la transcription de plusieurs séquences codantes (cistrons)

2-2-1- La transcription chez les eucaryotes

Chez les eucaryotes, il existe trois ARN polymérases I, II et III, et une ARN polymérase mitochondriale.

- ARN polymérase I : présente dans le nucléole, catalyse la synthèse des ARNr (sauf l'ARNr 5S)
 - ARN polymérase II : présente dans le nucléoplasme, responsable de la synthèse des ARNm
 - ARN polymérase III : présente dans le nucléoplasme, responsable de la synthèse des ARNt et l'ARNr 5S
- Les ARN polymérases eucaryote sont généralement d'une masse moléculaire supérieure à celle de l'ARN polymérase eucaryote (supérieure de 500000 Daltons).

Bien que la transcription reste dans son principe identique chez les eucaryote et les procaryotes, des différences, parfois importantes, existent. Elles tiennent à la structure de la matrice d'ADN, aux polymérases, et aux modifications post-transcriptionnelles que subissent les transcrits primaires de l'ARN.

Chez les eucaryotes la transcription, qui s'opère dans le noyau, est beaucoup plus complexe car elle nécessite une altération localisé et réversible de la structure de la chromatine. Le démarrage s'effectue grâce aux séquences promotrices se trouvant autour des positions -25 et -75. Un facteur TBP (TATA Binding Protein) est responsable de la reconnaissance du site d'initiation. L'arrêt de la transcription est annoncé par une région du brin transcrit riche en base T (TTATT).

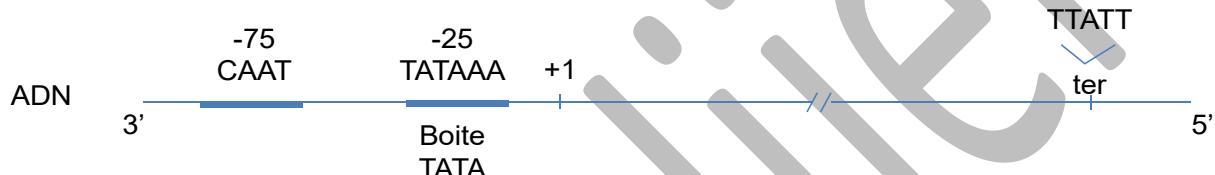


Figure 15: Promoteurs et terminateurs eucaryotes

ARN polymérase I

La synthèse de l'ARN ribosomal s'effectue dans un compartiment spécialisé du noyau, le nucléole. Les ARN 5.8S, 18S et 28 S sont synthétisés en quantités équimolaires par l'ARN polymérase I sous forme d'un ARN précurseur 45S (l'ADN contient un grand nombre de copies des gènes des ARNr).

ARN polymérase III

ARN polymérase III copie: - l'ARNr 5S - les ARNt

- certains ARN nucléaires appelés petits ARN nucléaires ou ARNsn

La transcription des ARNt nécessite les facteurs TFIIIB et TFIIIC. L'ARNt subit un processus d'épissage (élimination de l'intron) avant de passer dans le cytoplasme. Le processus d'élimination de l'intron est assez différent de celui utilisé pour l'ARNm.

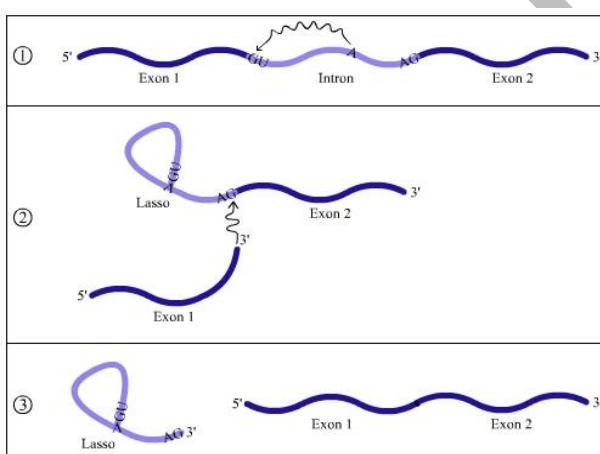
Un troisième facteur TFIIIA est nécessaire pour la transcription des gènes de l'ARNr 5S (pas de maturation pour l'ARN 5S et ARNsn).

ARN polymérase II

Les ARNm eucaryotes sont monocistroniques et possèdent une durée de vie assez longue. La liaison de l'ARN polymérase à la boîte TATA (-25) est assistée par une série de facteurs de transcription: TFIIA, TFIIB, etc. Les ARNm sont protégés à leur extrémité 5' par une coiffe (capping) composée d'une guanosine méthylée liée par un groupe triphosphate au premier ribonucléotide. Cette modification apportée à l'ARN en cours de transcription est reconnue par les ribosomes lors de l'initiation de la traduction. Une fois le transcrit primaire est libéré, l'extrémité 3' est modifiée par addition de 100 à 200 résidus d'acide adénylique (polyadénylation) grâce à l'action d'une polyA-polymérase, qui utilise des ATP comme donneurs d'adénine. La queue de poly A a pour rôle de permettre l'exportation de l'ARNm dans le cytosol, ainsi que de l'y stabiliser. (Les ARN des histones ne possèdent pas de queue poly A).

L'ARN transcrit connues sous le nom ARN nucléaires hétérogènes (ARNnh) ou ARN prémessagé. L'épissage de l'ARNm nucléaire est une étape majeure de sa maturation. Ce processus est l'élimination des séquences non codantes (introns) pour produire des ARNm matures. Plusieurs ribonucléoprotéines (snRNP : Small nuclear ribonucleoproteins, ou snurps) interviennent dans le processus d'épissage. Ce sont des complexes (splicéosome) contenant de petits ARN nucléaire (snARN), U1 à U6.

L'excision-épissage est réalisée par réaction d'un nucléotide à adénine (A) situé dans l'intron avec un nucléotide à guanine situé en 5' de l'intron. Cela entraîne la séparation de l'intron d'avec l'exon 1 (situé en amont) et la formation d'une structure en lasso interne à l'intron. Ensuite, l'extrémité 3' de l'exon 1 réagit avec l'extrémité 5' de l'exon 2 permettant l'épissage des deux exons et la libération du lasso qui sera dégradé par des ribonucléases.



<http://www.snv.jussieu.fr/vie/dossiers/transcription/transcription.htm>

La transcription chez les eucaryotes, Auteurs : Françoise Ibarrondo, Gilles Camus
Document publié le : 12 décembre 2006

2-3- La traduction

2-3-1- ARN de transfert

L'extrémité 3'OH est terminée par le triplet CCA ou' ce produira une liaison ester entre l'ARNt et l'AA. Une zone anticodon qui se lie par des liaisons faibles (hydrogène) au codon présent dans l'ARNm. La boucle de l'anticodon porte une séquence de trois bases, qui reconnaît et hybride un codon complémentaire porté par l'ARNm.

Avant de participer à la traduction de l'ARNm, les ARNt fixent l'AA correspondant à leur anticodon. Chaque AA est d'abord activé, en présence d'ATP, par une aminoacyl ARNt synthétase qui catalyse en une première étape la formation d'un aminoacyl-AMP. L'enzyme reste liée au complexe aminoacyl-AMP jusqu'à la rencontre d'un ARNt spécifique de l'AA. Dans une deuxième étape, l'AA est transféré sur l'extrémité 3'OH de l'ARNt et forme un aminoacyl-ARNt. C'est à partir de ces aminoacyl-ARNt que les AA seront assemblés pour former les protéines.

2-3-2- Les ribosomes et ARN ribosomaux

La comparaison des séquences de protéines ribosomales de plusieurs organismes révèle une très grande conservation au cours de l'évolution. Les ARNr sont aussi très bien conservés.

Les deux sous unités sont souvent libres et ne s'associent qu'au moment de la traduction. La fonction principale des ribosomes est d'orienter correctement les aminoacyl ARNt par rapport à l'ARNm et de créer les liaisons peptidiques.

Plusieurs ribosomes peuvent successivement commencer la traduction de la même molécule d'ARNm constituant un polysome ou polyribosome

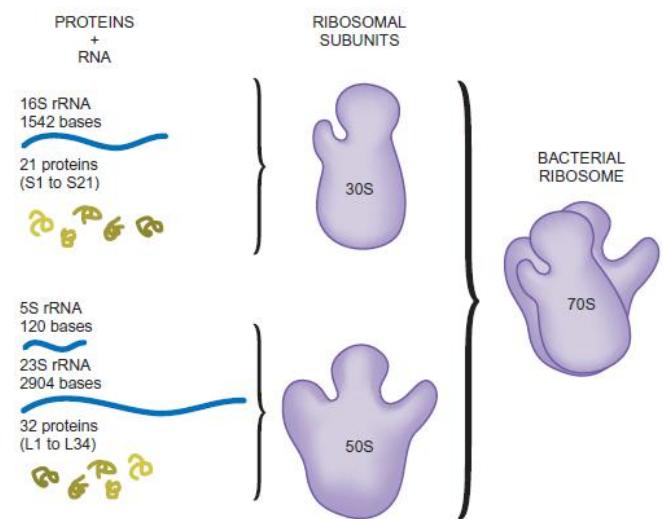


Figure 17: Components of a Bacterial Ribosome

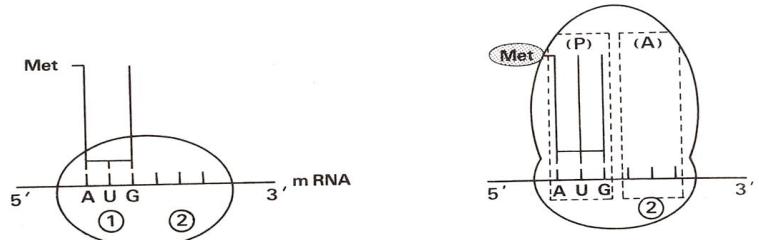
The ribosome is composed of 30S and 50S subunits. The 30S subunit consists of one 16S rRNA together with 21 proteins, and the 50S subunit has a 5S and 23S rRNA plus 34 proteins.

2-3-3- Les étapes de la traduction

- La traduction de l'ARN en protéines s'accomplit toujours dans le sens 5'....3'.
- Chez les procaryotes la traduction commence avant même la fin de la transcription.
- Chez les eucaryotes les ARNm sont transcrits puis maturés dans le noyau. Ils migrent ensuite dans les cytoplasmes pour y être traduits en protéines.

L'initiation :

- « codon initiateur » : AUG situé à environ 100 pb de l'extrémité 5' de l'ARNm, indiquant le début de la traduction.
- AUG=Met (chez les bactéries: N-formyl-méthionine): toutes les protéines commencent par une méthionine (qui sera clivée lors de la maturation de la protéine)
- Petite sous-unité s'associe avec l'ARNm au niveau du codon AUG avec l'ARNr portant la Met
- Grande sous-unité se fixe à ce complexe pour former le ribosome fonctionnel
- IF3 empêcherait l'association des sous-unités 30S et 50S
- IF2 lie le GTP au fMet-ARNr et dirige la liaison du fMet-ARNr à la sous-unité 30S
- La seconde étape est l'association de fMet-ARNr et de l'ARNm à l'ensemble IF-30S-GTP au niveau d'une séquence de l'ARNm (AGGAGGU) appelée site de fixation du ribosome (RBS: Ribosome Binding Site) ou séquence de Shine Dalgarno.
- Le codon AUG initiateur est défini par l'hybridation entre la séquence RBS et une séquence proche de l'extrémité 3' de l'ARN 16S.
- Chez les eucaryotes, les ribosomes reconnaissent la coiffe méthylée grâce à des facteurs appelés protéines fixatrices de la coiffe (Cap binding protein).
- IF3 et ensuite libéré permettant Initiation l'association entre les deux sous-unités (50S et 30S).
- IF-1 empêche la fixation d'ARNr sur la partie de la petite sous-unité qui fera partie du site A.

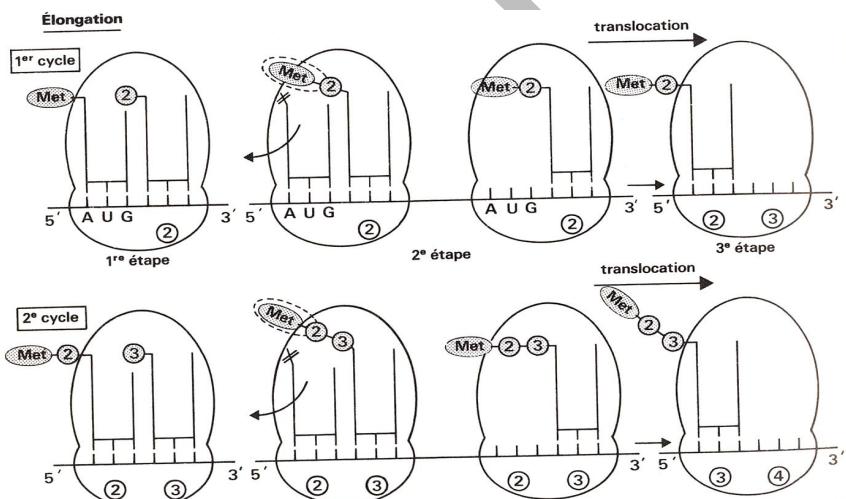


L'elongation

- Chaque ribosome possède deux sites de fixation pour l'ARNr : P (peptidyl) et A (Aminoacyl).
- Le premier site P est occupé par le fMet-ARNr^{fMet} apparié à l'AUG de l'ARNm.
- Le site A peut alors recevoir l'aminoacyl-ARNr dont l'anticodon correspond au second codon de l'ARNm. Une liaison peptidique est établie entre le groupement COOH du premier acide aminé et le groupe NH₂ du second acide aminé grâce à l'enzyme peptidyl transférase, composante de la sous-unité 50S.

- A la suite de cette liaison, l'ARNt^{fMet} déchargé quitte le site P. Le dipeptide se trouve alors lié à l'ARNt portant le deuxième acide aminé et placé au site A.
- Ce peptidyl ARNt se déplace ensuite vers le site P par mouvement relatif de l'ARNm d'une longueur de trois bases. Le site A peut alors recevoir un nouvel aminoacyl-ARNt. Une fois dans le site A, une liaison peptidique associe l'acide aminé à celui qui le précédait dans le site P. La répétition de ces mouvements engendre l'elongation du polypeptide. L'allongement de la chaîne polypeptidique nécessite l'hydrolyse du GTP.
- des facteurs d'elongation (EF : Elongation Factors) sont également indispensables.
- EF-Tu réagit avec le GTP et l'ARNt chargé pour former le complexe aminoacyl-ARNt-GTP-EF-Tu et placer, grâce à l'hydrolyse du GTP en GDP, l'aminoacyl-ARNt au site A du ribosome.
- EF-Ts participe à la génération du complexe EF-Tu-GTP.
- la translocation (le déplacement du peptidyl-ARNt du site A au site P) se fait par l'intermédiaire du facteur d'elongation EF-G qui se fixe au ribosome avec du GTP qui est ensuite hydrolysé ce qui libère de l'énergie pour la translocation.

- Les mécanismes d'allongement (élongation) de la chaîne peptidique chez les eucaryotes s'avèrent similaires avec la participation des facteurs d'elongation eEF1 et eEF2.

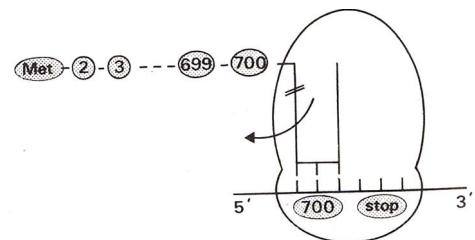


La terminaison :

S'effectue quand le ribosome rencontre un codon stop (Aucun ARNt ne se fixe en face de ce codon). L'intervention d'un facteur de terminaison (RF : Release Factors). Les RF entrent dans le site A et provoquent la libération du polypeptide achevé. Chez *E. coli*, RF1 reconnaît UAG et UAA et RF2 reconnaît UGA et UAA. Chez les eucaryotes, le facteur de terminaison appelé eRF.

N.B : un même ARNm peut servir de matrice à plusieurs ribosome en même temps = polysome

Terminaison



2-3-4- Modification post traductionnelle

Après la traduction, les polypeptides peuvent subir différentes modifications par addition de groupements chimiques aux chaînes latérales, au N ou C des acides aminés ou par clivage protéolitique. De telles modifications pourraient être indispensables à l'acquisition d'une activité fonctionnelle complète. Par exemple, le résidu d'initiation formyl-méthionine ou méthionine est souvent excisé par une enzyme spécifique (Methionyl-aminopeptidase). La protéolyse est également nécessaire pour activer les protéines qui sont synthétisées en tant que précurseurs inactifs appelés proprotéines (Insuline). Les modifications peuvent impliquer l'addition de petits groupements, comme la méthylation, la phosphorylation, l'acétylation, et l'hydroxylation ou l'addition de plus grosses structures comme des lipides, des oligosaccharides (glycosylation).

Références

- **Beaumont S.**, Biologie Moléculaire. Cours, exercices, annales et QCM corrigés, Dunod, Paris, France, 2006
- **Clark D., Pazdernik N. and McGehee M.**; . Molecular Biology, Third Edition, 2019
- **Lehninger**, Principles of Biochemistry, Fourth edition 2005
- **Lewin B.**, Gene, 8ème édition, 2005
- **Meftah A. et Julien R.**, Biologie moléculaire, 2eme edition, Dunod, Paris, France, 2003
- **Passarge E.** , Color Atlas of Genetics, Third edition, 2007, Thieme, Germany
- **Prescott et al.** Microbiologie 2003
- **Serge Weinman, Pierre Méhul.** TOUTE LA BIOCHIMIE. Dunod, Paris, 2013
- **SwyngedauwB., J-S Silvestre** , Biologie et Génétique Moléculaires Aide-Mémoire, 3e Edition, Dunod, Paris, 2008
- **Voet et al.** Fundamental of Biochemistry 1999
- **WinterP.C., Hikey G.I. et Fletcher H.L.**, L'essentiel en génétique, Port royal Livres, Paris 2000
- Site interne
- La transcription chez les eucaryotes, *Auteurs : Françoise Ibarrondo, Gilles Camus Document publié le : 12 décembre 2006.*
<http://www.snv.jussieu.fr/vie/dossiers/transcription/transcription.htm>