

**Expression de l'information génétique**

**Objectifs spécifiques :**

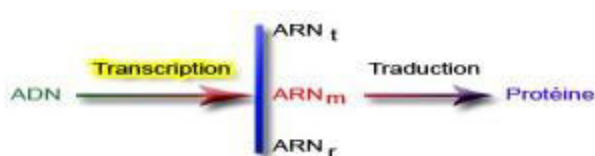
- Définir la transcription d'un gène
- Citer les caractéristiques de la transcription
- Préciser le mécanisme de la transcription de l'ADN chez les procaryotes avec ses trois étapes (initiation, élongation et terminaison)
- Reproduire la fourche de réplication de l'ADN à l'aide d'un schéma synthétique
- Schématiser la zone d'ouverture de la molécule d'ADN « œil de réplication »
- Positionner sur le schéma « œil de réplication » les différents protéines et enzymes de la réplication
- Préciser les différents étapes de maturations d'un transcrit
- Schématiser un épissage constitutif et alternatif en situant les zones consensus d'épissage
- Préciser le mécanisme de la transcription de l'ADN chez les eucaryotes
- Décrire la maturation des ARN pré-messager
- Etablir un tableau de comparaison entre la réplication chez les eucaryotes et procaryotes

**1. Définition de la transcription d'un gène :**

La transcription est un mécanisme faisant intervenir un système enzymatique, qui convertit l'information génétique d'un segment d'ADN en un brin d'ARN complémentaire (simple brin). Cette molécule est appelée transcrit.

L'information génétique, nécessaire à la synthèse des protéiques est organisée sous forme de gènes. Ainsi un gène est composé de séquences codant pour la protéine ainsi que des séquences non codantes situées en amont 5' et en aval 3' des séquences codantes nécessaires à la régulation de son expression.

Les gènes ne dirigent pas directement la synthèse des protéines pour lesquelles ils codent. L'information qu'ils contiennent est d'abord exprimée sous forme d'ARN messager (ARN<sub>m</sub>) grâce à un processus appelé « transcription » et c'est l'ARN<sub>m</sub> qui va diriger la synthèse de la protéine produit final de l'expression d'un gène au moyen d'un autre processus appelé « traduction » ceci constitue le dogme central de la biologie moléculaire figure 1.



## Figure1.dogme central de la biologie moléculaire

La transcription assure donc la synthèse de l'ARNm et la traduction assure la synthèse des protéines.

Le processus de transcription ne concerne donc pas toutes les molécules ADN d'une cellule mais uniquement la fraction de cet ADN correspondant aux gènes actifs de la cellule.

### 2. Mécanisme généraux de la transcription :

- La transcription ne concerne pas toute l'information génétique, chez les eucaryotes une grande part de cette information n'est pas exprimée.
- La transcription est assurée par l'ARN polymérase, elle possède une activité 5' → 3' de polymérisation et capable d'initier la synthèse d'une chaîne polypeptidique à partir de nucléotides libres.
- L'ARN polymérase utilise l'un des deux brins de l'ADN comme matrice pour assurer la synthèse de l'ARN aussi appelé brin antisens. Cette synthèse se fait dans le sens 5' → 3' selon les règles d'appariement complémentaire des bases.
- Le brin d'ARN est ainsi complémentaire et antiparallèle au brin d'ADN matrice transcrit.
- Il est identique exception faite du T(Thymine) remplacé par U(Uracile), au brin d'ADN non transcrit appelé brin sens car il possède la même orientation et la même séquence que l'ARNm dirigera la synthèse de la protéine.
- L'ARN possède certaines caractéristiques par rapport à l'ADN :
  - Il contient un ribose à la place du désoxyribose
  - Un uracile à la place de la thymine (les autres bases sont identiques)
  - Il est monocaténaire généralement linéaire et parfois replié sur lui-même

### 3. La transcription chez les eucaryotes :

1. **Initiation** : l'ARN polymérase se fixe en amont du gène au niveau de différents promoteurs le plus connu la boîte TATA (TATA ou une variante) à -25 début du site de la transcription. D'autres promoteurs sont également connus la boîte GC (GGGCGG) promoteur des gènes domestiques ou ménage (housekeeping gns) et la boîte CAAT (CCAAT) à -80 du début du site de la transcription. Ces promoteurs font partie des séquences dites actifs en cis.
2. **Elongation** : l'ARN polymérase se déplace ensuite le long du brin d'ADN séparant au fur et à mesure les deux brins de la double hélice et catalysant l'addition de ribonucléotides à l'extrémité 3'OH de la chaîne d'ARN en formation d'après la séquence de la matrice d'ADN. Une fois le site d'initiation de la transcription libérée une nouvelle molécule d'ARN polymérase vient s'y fixer et amorcer la synthèse d'un nouveau transcrit. Plusieurs ARN polymérase transcrivent donc simultanément l'ADN d'un gène en se déplaçant vers l'aval l'une derrière l'autre réalisant ainsi la synthèse de plusieurs molécules d'ARN à partir d'un seul gène.

3. **Terminaison** : la terminaison de la transcription chez les eucaryotes est une étape très mal connue. Elle se fait très en aval du début de la polymérisation de la molécule ARN. Il semble que l'arrêt de la transcription se fait au niveau de région en épingle à cheveux riche en GC suivie par des régions riche en AT. A la fin de la transcription la polymérisation la polymérase se détache de la matrice ADN et une molécule ARN est libérée. Cette molécule ARN est de grande taille contient aussi bien les séquences correspondant aux exons que celles correspondant aux introns. Elle est appelé HnRNA (Hetero Nuclear RNA) ou transcrit primaire figure2.

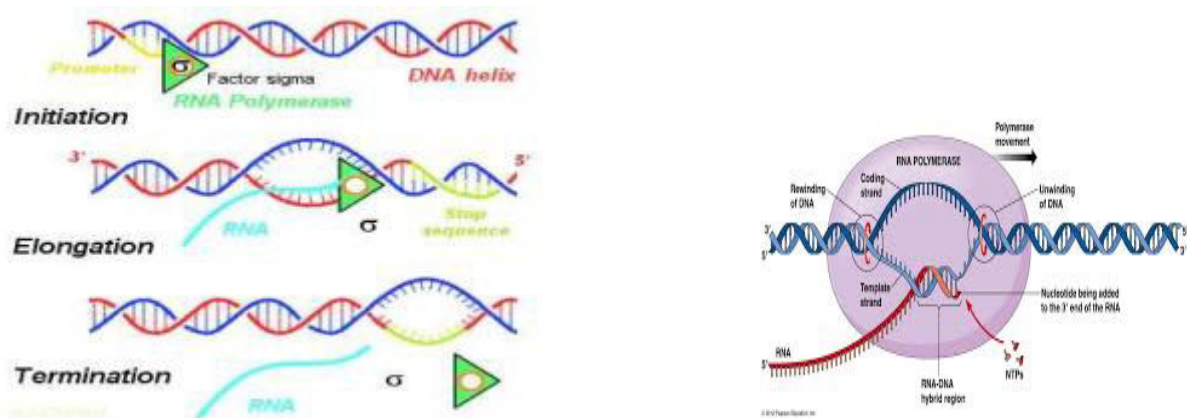


Figure2. Les étapes de la transcription

#### 4. La transcription chez les procaryotes :

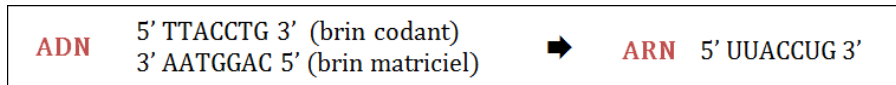
L'ARN polymérase est une protéine ADN dépendante, multimérique possédant les sous-unités  $\alpha$ ,  $\beta$ ,  $\beta'$  et  $\sigma$ . Elle est présente sous deux formes l'**enzyme-cœur** ( $\alpha 2\beta\beta'$ ) et l'**holoenzyme** ( $\alpha 2\beta\beta'\sigma$ ). Les ARN polymérases ne nécessitent pas d'amorce et ne possèdent pas d'activité exo-nucléasique et donc de correction d'erreur, le taux d'erreur est ainsi plus important que pour les ADN-polymérases, mais ce taux est supporté.

Les nucléotides triphosphates sont additionnés à l'extrémité 3' de la chaîne en cours de synthèse par complémentarité de la matrice d'ADN. L'hydrolyse de la liaison anhydride fournit l'énergie pour la synthèse de la liaison phosphodiester.

La réaction est réalisée en milieu tamponné à pH neutre, contenant du sel de  $Mg^{2+}$ , un agent de protection des groupements SH de l'enzyme (agent réducteur comme le  $\beta$ -mercaptoéthanol), les 4 ribonucléotides (rNTP) et de l'ADN bicaténaire contenant un promoteur comme matrice.

La molécule d'ADN est composée d'un **brin matriciel** (ou **brin anti-codant**) dirigé par définition de 3' vers 5' et servant comme son nom l'indique de matrice à l'ARN-polymérase,

ainsi que d'un **brin codant** dirigé par définition de 5' vers 3' et ayant une séquence identique à l'ARN transcrit mise à part le fait que la thymine est changée par l'uracile. L'ARN messager est lui monocaténaire et dirigé tout comme le brin codant de 5' vers 3'.



Chez E-coli, **une seule ARN-polymérase catalyse la synthèse de tous les ARN** de la cellule (en mettant à part l'ARN des amorces nécessaire à la réplication de l'ADN).

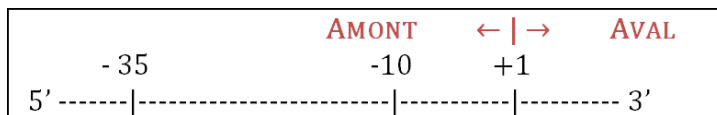
La transcription est divisée en plusieurs étapes : la pré-initiation, l'initiation, l'élongation et la terminaison.

### 1. Pré-initiation

Le promoteur est une séquence d'une centaine de nucléotides située dans la région régulatrice et désignant le début de la transcription. Il est situé en amont du site d'initiation et porte des éléments de séquence reconnus par l'ARN-polymérase et déterminant le sens de la transcription.

Le promoteur est constitué de courtes séquences conservées d'une unité de transcription à l'autre et appelées séquences consensus :

- En -10 du site d'initiation on trouve la **TATA box** : (TATAAT )
- En -35 du site d'initiation on trouve : « TTGACA »



La **force intrinsèque d'un promoteur** est définie par le nombre relatif d'initiation par unité de temps (vitesse de la transcription). Elle dépend de la proportion des paires de bases A-T par rapport aux paires de bases C-G, de la position des séquences en -35 et en -10, en effet plus les séquences consensus sont proches du site d'initiation plus le promoteur sera fort et finalement de la force d'interaction de l'ARN-polymérase avec le promoteur.

Le promoteur agit sur la transcription du segment d'ADN qui lui est adjacent sur le même chromosome, on dit que le promoteur est actif en « cis ». Le promoteur n'est pas actif sur les séquences codantes situées ailleurs sur le chromosome, dans ce cas là on parlerait alors du qualificatif « **trans** ». Le cis s'oppose au trans.

L'affinité de l'ARN-polymérase pour l'ADN dépend de la forme de l'enzyme : l'enzyme-cœur a une affinité faible et non spécifique, l'holoenzyme a une affinité très forte et spécifique pour le promoteur. On peut faire la remarque que la sous-unité sigma  $\sigma$  à l'état libre ne se fixe pas sur l'ADN. La sous-unité  $\beta'$  étant basique et l'ADN étant acide, ce sera elle qui facilitera l'interaction du complexe avec le promoteur.

La **sous-unité sigma  $\sigma$**  permet donc une reconnaissance spécifique du promoteur par l'ARN-polymérase et diminue l'affinité de l'enzyme pour les régions non promotrices. Il agit de manière cyclique, en effet après l'initiation faite, le facteur sigma se détache pour être recyclé et réutilisé pour d'autres initiations de gènes.

L'ARN-polymérase entraîne la dénaturation des deux brins d'ADN sur 14 paires de nucléotides, on parle de complexe ouvert qui augmente encore l'affinité de l'enzyme pour la double hélice.

## **2. Initiation :**

L'initiation correspond à la synthèse de la première liaison phosphodiester réalisé par la sous-unité  $\beta$  qui correspond à la sous-unité catalytique de l'ARN-polymérase.

L'interaction de cette sous-unité est inhibée par la **rifampicine** qui inhibe ainsi de manière irréversible la transcription de l'ADN, c'est le cas de la tuberculose. Une mutation dans la SU  $\beta$  induit l'apparition de souches bactériennes résistantes à la rifampicine.

Le déroulement des premières étapes de la transcription est donc :

1. liaison non spécifique de l'holoenzyme.
2. formation d'un complexe fermé au niveau du promoteur
3. formation du complexe ouvert (déroulement sur 14 nucléotides)
4. Mise en place du premier nucléotide (très souvent A ou G)
5. Allongement de 4 à 5 nucléotides
6. Détachement du facteur sigma, après la transcription des 4-5 premiers nucléotides.

## **3. Elongation :**

L'élongation correspond au déplacement de la bulle de transcription le long de la molécule d'ADN. La région désappariée est alors de 70 paires de bases. Pendant la transcription, l'ARN forme un court appariement avec le brin matriciel de l'ADN formant une hélice hybride ADN-ARN sur une dizaine de paires de bases.

L'élongation est inhibée par des aminosides ou amino-glucosides.

## **4. Terminaison :**

La terminaison se fait lorsque l'enzyme arrive au niveau d'une séquence spécifique appelé **terminateur**. Le terminateur se présente sous la forme d'un **palindrome** (cf. lexique) qui peut être parfait ou imparfait. Ce palindrome entraîne une complémentarité de séquence au niveau de l'ARNm qui permet la mise en place d'une structure en épingle à cheveux (ou tige-boucle) qui est un appariement intra-chaîne qui déstabilise l'ARN-polymérase jusqu'à dissociation.

Elle peut être facilitée par un **facteur rho  $\rho$**  suivant la séquence du terminateur, on met ainsi en évidence des **terminateurs rho indépendant** (environ les 2/3) et des **terminateurs rho dépendant** (environ 1/3) :

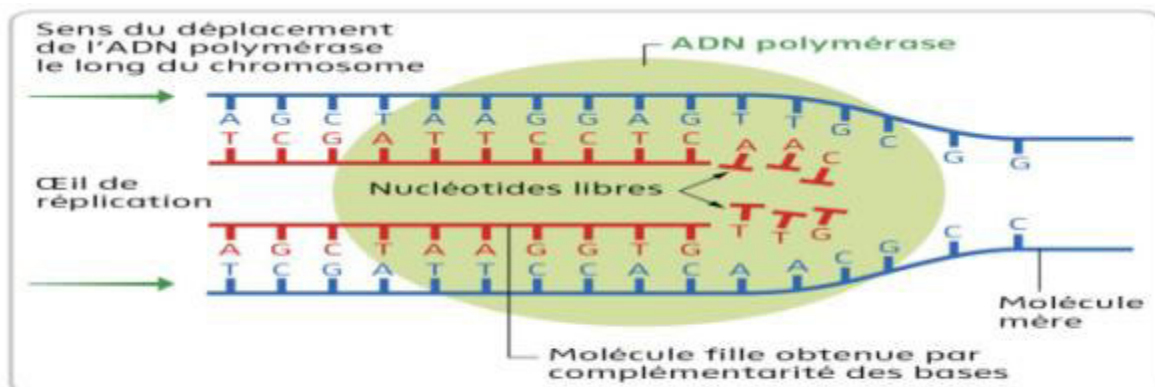
- Pour les terminateurs rho indépendant on trouve une structure en épingle à cheveux riche en paires de bases G-C, suivie d'une séquence poly-U d'environ 6 nucléotides permettant une dissociation plus facile de l'hybride ADN-ARN.
- Pour les terminateurs rho dépendant on trouve une structure en épingle à cheveux plus courte et qui n'est pas riche en paires de bases G-C et qui est non-suivie d'une séquence poly-U. Il y a donc nécessité du facteur rho qui a une affinité pour les ARN en court de synthèse, le parcourant de 5' vers 3' jusqu'à trouver l'ARN-polymérase. Le facteur rho est ATP dépendante, dont l'hydrolyse permettra la dissociation du complexe.

### 5. Maturation des transcrits primaires :

Le **transcrit primaire** correspond à l'ARN non mature qui nécessite une maturation sous forme de clivages ou de modifications de bases. Cette maturation n'est pas obligatoire. Après maturation on obtient l'ARN mature. On peut prendre quelques exemples comme la maturation du transcrit primaire donnant les ARNr (transcrit primaire 45 S) ou les ARNt par des ribonucléases, ou la maturation du transcrit primaire codant les ARNm. Généralement, il y a peu de modifications pour les ARNm ; beaucoup d'ARNm sont traduits en protéines alors qu'ils sont encore transcrits (Attention : il n'y a pas de noyau chez les procaryotes).

Le transcrit primaire code soit pour un produit, on parle d'**ARN monocistronique**, soit plusieurs produits, on parlera alors d'**ARN polycistronique** .

### 6. La zone d'ouverture de la molécule d'ADN « œil de réplication »



**C** Représentation simplifiée d'une fourche de réplication. Remarque : une seule ADN polymérase est ici représentée.

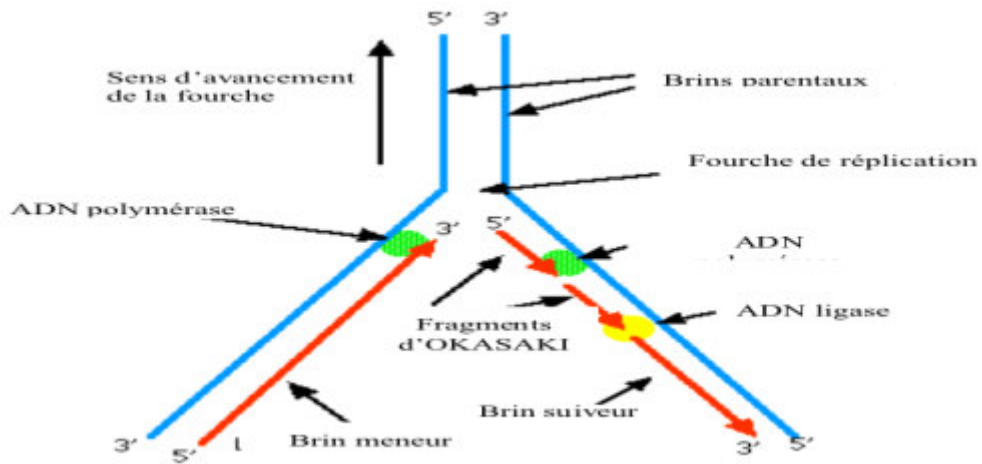


Figure3. La fourche de réplication de l'ADN

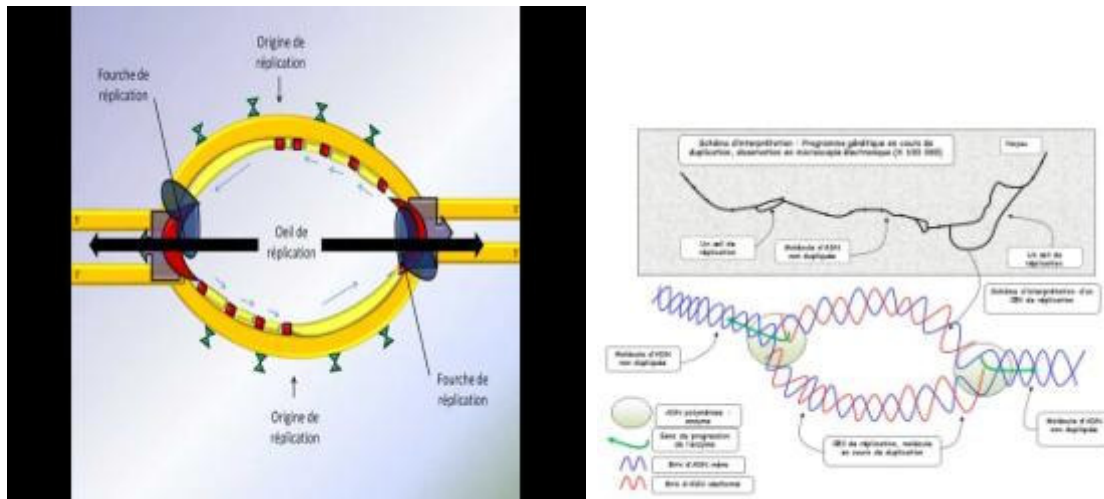




Figure4. Œil de réplication

7. Tableau comparative de la réplication chez les eucaryotes et chez les procaryotes :

**Organisation des gènes :**

Procaryotes	Eucaryotes
<p>Gènes continus</p> 	<p>Gènes discontinus</p>  <p>La plupart des gènes eucaryotes possèdent des séquences nucléotidiques qui ne codent pas pour des séquences polypeptidiques. Ce sont des "séquences intervenantes non traduites" appelées aussi .... Les gènes sont donc morcelés</p>

**2- La transcription des gènes :**

**2-1- Les séquences promoteurs:**

Procaryotes	Eucaryotes
<ul style="list-style-type: none"> <li>• Boite "Pribnows" : 5'-TataaT 3' région -10</li> <li>• Séquence région -35</li> </ul> <p>Schéma de l'unité de transcription :</p>	<ul style="list-style-type: none"> <li>• boite "Hogness" : TATA box 5'-TATA 3' région -25 à -30</li> <li>• CCAAT box région -70</li> <li>• GC box : région -80 à -160</li> </ul> <p>Schéma de l'unité de transcription :</p>

**2-2- Les séquences régulatrices :**

Procaryotes	Eucaryotes
	<p>Ces séquences se situent toujours loin des gènes dans la séquence nucléotidique. Ils peuvent être en amont ou en aval du gène ou des gènes impliqués.</p> <ul style="list-style-type: none"> <li>• Séquences "Enhancers" = séquences stimulatrices</li> <li>• Séquences "Silencers" = séquences inhibitrices</li> </ul>

**Référence :**

- Cours biologie moléculaire Professeur Boujama (1999). Faculté de médecine oran1.
- Biologie moléculaire et médecine (2° Éd.) Coll. De la biologie à la clinique. Auteurs : KAPLAN Jean Claude, Delpech Marc
- Centre national de la recherche scientifique.