**IV- La transcription**

1. **Généralités**

* **Notions de gène et de génome**
* **Le génome** est l’ensemble du matériel génétique des individus compris dans une cellule.
* **Un gène** est une séquence (fragment) d’ADN avec une structure nécessaire à la **synthèse** **d’un produit fonctionnel** qui peut être sous la forme d’ARN ou de polypeptide.
* **Les classes d’ARN**
* L’**ARN messager** (ARNm) transporte, de l’ADN au ribosome, les instructions codées pour la synthèse des chaînes polypeptidiques.
* L’**ARN ribosomal** (ARNr) associé à des protéines, constitue les sous-unités du ribosome, le siège de l’assemblage des protéines.
* L’**ARN de transfert** (ARNt) fait le lien entre la séquence nucléotidique codante de l’ARNm et la séquence d’acides aminés d’une chaîne polypeptidique.

**Remarque**

D’autres classes de molécules d’ARN sont présentes dans le noyau des cellules (ARNpn, ARNpno, ARNmi, ARNpi).

1. **Présentation de la transcription**

La transcription correspond à la **synthèse d’une molécule d’ARN** à partir d’une matrice d’ADN. Les gènes sont transcrits seulement quand leurs produits sont nécessaires.

1. **Eléments nécessaires à la transcription**

La transcription nécessite trois composants essentiels :

* Une **matrice d’ADN**
* Les matériaux de base (**les substrats**) pour produire une nouvelle molécule d’ARN
* Le **système de transcription**, comprenant les protéines nécessaires pour catalyser la synthèse d’ARN
* **La matrice de transcription (le brin transcrit)**

La matrice pour la synthèse d’ARN est un des brins de la double hélice d’ADN. Le brin utilisé pour la transcription est appelé **brin matrice**. L’autre brin n’est normalement pas transcrit. La transcription produit une molécule d’ARN qui est complémentaire du brin matrice et de polarité opposée. Le transcrit d’ARN a donc la même polarité et la même séquence de bases que le brin qui n’a pas servi de matrice, sauf que l’ARN contient U à la place de T. Ainsi, quand on écrit des séquences nucléotidiques, on ne donne généralement que la séquence du brin d’ADN qui n’a pas servi de matrice, parce qu’elle sera la même que la séquence du transcrit d’ARN. Pour cette raison, ce brin est souvent appelé « brin sens ».

* **Le système de transcription**

L’action de **l’ARN polymérase** est assistée par plusieurs protéines auxiliaires qui s’associent de façon transitoire à la polymérase à différentes étapes du processus

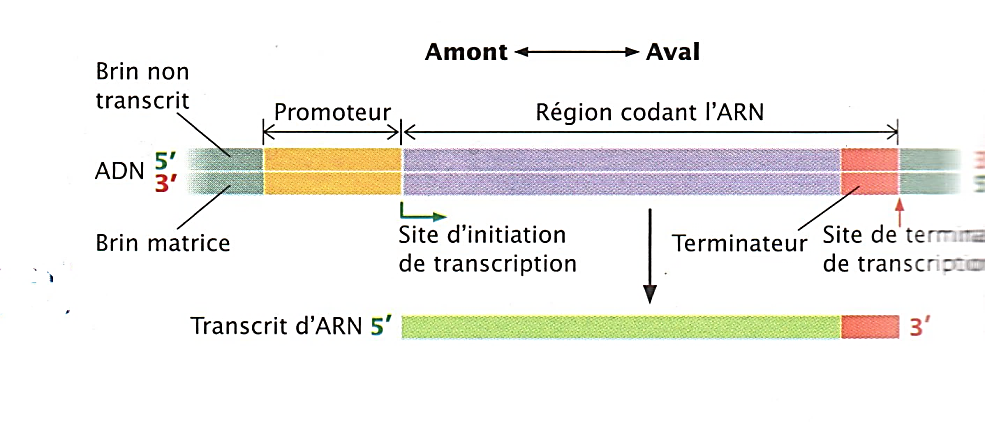
* **Les substrats de la transcription**

L’ARN est synthétisé à partir de **ribonucléosides triphosphates** (rNTP). Les nucléotides sont ajoutés un à un à l’extrémité 3’-OH de la chaîne en formation. Deux groupements phosphate sont clivés du rNTP entrant, et le groupement phosphate restant forme une liaison phosphodiester qui attache le nucléotide à la molécule d’ARN en croissance.

Les nucléotides sont toujours ajoutés à l’extrémité 3’ de la molécule, donc le **sens de transcription est de 5’ vers 3’**.

1. **L’unité de transcription**

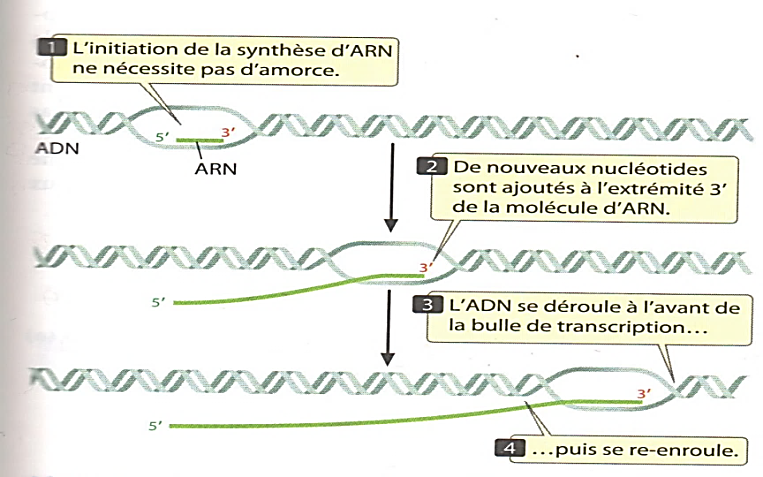
Une unité de transcription comprend trois régions importantes : un promoteur, une séquence codante et un terminateur.



* Le **promoteur**: une séquence d’ADN que le système de transcription reconnaît et à laquelle il se lie. Il indique le brin d’ADN qui doit être transcrit et le sens de la transcription. Il détermine aussi les sites d’initiation de la transcription, le premier nucléotide qui sera transcrit en ARN. Dans la plupart des unités de transcription, le promoteur est localisé à côté du site d’initiation de transcription, mais n’est pas, lui-même, transcrit.
* La **région codant l’ARN :** la séquence d’ADN qui est effectivement transcrite en molécule d’ARN.
* Le **terminateur** : une séquence nucléotidique qui signale l’endroit où doit s’arrêter la transcription. Les terminateurs font généralement partie de la séquence codante

1. Conventionnellement, la séquence est écrite de la gauche vers la droite, en commençant par l’extrémité 5’. Le premier nucléotide transcrit (le site d’initiation) est indiqué par +1, les nucléotides en aval du site d’initiation sont numérotés positivement, et les nucléotides en amont du site d’initiation reçoivent des numéros négatifs. Par exemple, le nucléotide +34 serait 34 nucléotides en aval du point de départ, tandis que le nucléotide -75 se trouverait 75 nucléotides en amont. Il n’y a pas de nucléotide 0.**La synthèse d’ARN n’a pas besoin d’amorce**

Contrairement à la réplication de l’ADN, l’initiation de la synthèse de l’ARN **n’a pas besoin d’une amorce**



1. **La transcription des procaryotes**

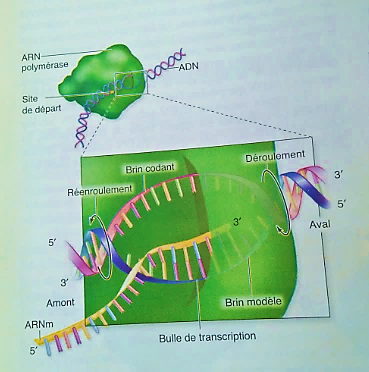
Les procaryotes ont **une seule ARN polymérase.**

* **Initiation**

L’ARN polymérase des procaryotes se fixe directement sur le promoteur. L’union de l’ARN polymérase au promoteur est la première étape de la transcription. L’ARN polymérase commence à dérouler l’hélice d’ADN.

* **Elongation**

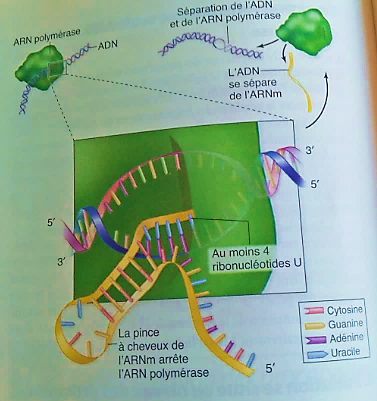
La transcription de la chaîne d’ARN débute en général par ATP ou GTP. L’un des deux devient l’extrémité 5’ de la chaîne, qui s’allonge dans le sens 5’ – 3’ par l’addition de ribonucléotides. La région contenant l’ARN polymérase, l’ADN modèle et l’ARN en croissance est appelée **bulle de transcription.** La position de l’extrémité 3’ de l’ARN interagit avec un ribonucléotide triphosphate entrant. La bulle formée par l’ARN polymérase descend le long de l’ADN à une vitesse constante d’environ 50 nucléotides par seconde. Après le passage de la bulle de transcription, l’ADN transcrit s’enroule à nouveau en la quittant



**Schéma d’une bulle de transcription**

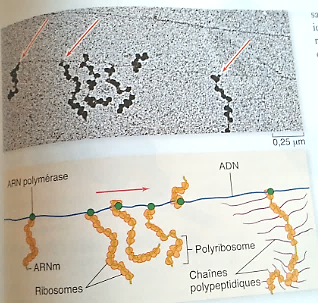
* **Terminaison**

L’extrémité de l’unité de transcription bactérienne est marquée par des séquences de terminaison qui disent « stop » à la polymérase. L’arrivée à ces séquences arrête la formation des liaisons phosphodietster, provoque la dissociation de l’ARN et l’ADN dans la bulle de transcription, la libération de l’ARN polymérase et la reconstitution de l’hélice d’ADN. Le transcrit d’ARN forme une structure bicaténaire : **une épingle à cheveux, suivie d’au moins 4 ribonucléotides Uraciles**. La formation de l’épingle à cheveux arrête l’ARN polymérase. Le brin d’ARN se sépare donc de l’ADN dans la bulle de ces terminateurs et aide à mettre fin à la transcription.



**Remarques**

* La transcription procaryote est couplée à la traduction. L’ARNm produit par la transciption commence à être traduit avant la fin de la transcription. Dès qu’une extrémité 5’ est disponible, des ribosomes s’y fixent pour entamer la traduction



* Les gènes procaryotes sont souvent organisés de façon à grouper ceux qui codent des fonctions apparentées. Ce groupement de gènes fonctionnellement apparentés est un **opéron**. Un opéron est une unité de transcription unique codant plusieurs enzymes nécessaires à une voie biochimique. Le groupement des gènes par fonction permet une régulation commune. On parle d’unité **polycistronique** dont l’exemple le plus classique est l’opéron lactose (voir chapitre de régulation de l’expression des gènes).

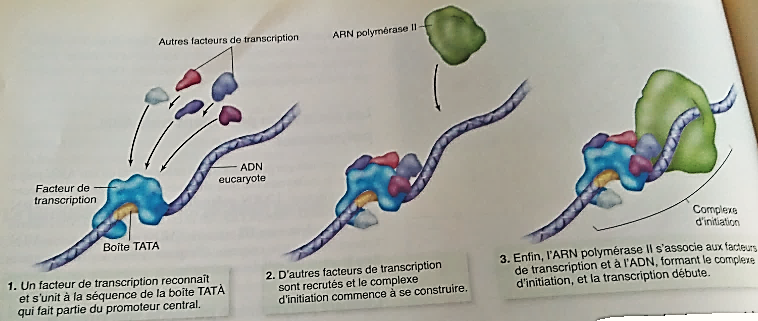
1. **La transcription des eucaryotes**

* **Les ARN polymérases des eucaryotes**

La plupart des cellules eucaryotes possèdent trois types d’ARN polymérases :

* **L’ARN polymérase** I transcrit les **ARNr** ;
* **L’ARN polymérase II** transcrit les **pré-ARNm**, les ARNpno, certains ARNmi et certains ARNpn ;
* **L’ARN polymérase III** transcrit de petites molécules d’ARN, spécifiquement les **ARNt**, les **petits ARNr**, certains ARNmi et certains ARNpn.
* **Les différentes phases de la transcription**
* **Complexe d’initiation**

L’ARN polymérase II des eucaryotes ne se fixe pas directement sur le promoteur. Elle se fixe par l’intermédiaire de facteurs de la transcription comprenant plusieurs protéines (TFIIA, TFIIB ...). Ces protéines associées à l’ARN polymérase II constituent le complexe d’initiation de la transcriptionet catalysent la formation de la première liaisonphosphodiester entre les deux premiers nucléotides de l’ARNm. Durant la phase d’élongation, le site promoteur est libéré par la progression de l’ARN polymérase II sur l’ADN et un autre complexe d’initiation peut se mettre en place.

****

* **Formation du transcrit primaire ou pré-ARNm**

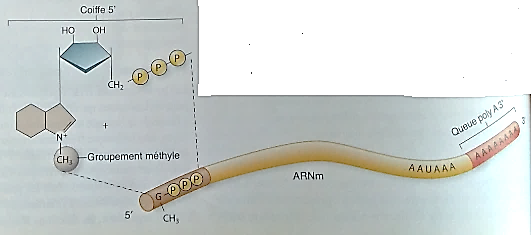
Le transcrit primaire correspond à une **copie intégrale des exons et des introns d’un gène**. Les promoteurs de l’ARN polymérase II possèdent des éléments de contrôle situés en amont du site d’initiation :

* **La boîte TATA :** située à environ -25 paires de bases de l’origine de la transcription. C’est une séquence de six nucléotides riches en A et T. La séquence dite consensus (statistiquement la plus rencontrée) est TATAAA.
* **La boîte GC** (située le plus souvent dans la région entre -110 et -40). Elle peut se présenter sous forme d’hexanucléotides : 5’-GGGCGG-3’. Le motif riche en bases G et C peut être répété plusieurs fois.
* **La boîte CCAAT** (souvent située dans la région entre -120 et -80). Cette boîte peut être située avant ou après une boîte GC ou même entre deux boîtes GC.

Les transcrits originels synthétisés par l’ARN polymérase II subissent plusieurs modifications. L’ARN synthétisé par l’ARN polymérase II est appelé **transcrit primaire**. La forme finale modifiée est désignée par l’**ARNm mature**.

* **Addition de la coiffe à l’extrémité 5’**

La première base du transcrit est généralement une Adénine (A) ou une guanine (G), et elle est ensuite modifiée par addition de GTP au groupement PO45’ pour former la coiffe. Cette coiffe protège l’extrémité 5’ de l’ARNm de l’attaque par des enzymes de dégradation et elle intervient aussi dans l’initiation de la traduction

****

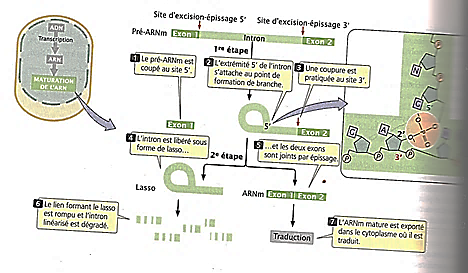
* **Addition de la queue poly(A) à l’extrémité 3’**

Après synthèse, les ARN messagers sont clivés dans leur partie 3’ une vingtaine de bases en aval d’une séquence spécifique : AAUAAA. Cette coupure est réalisée par une endonucléase. Après cette coupure, une enzyme la poly(A) polymérase en présence d’ATP additionne un nombre variable d’A. La présence de poly(A) aurait également **une fonction de protection des ARNm sur l’extrémité** **3’**. Elle stabilise de nombreux ARNm en augmentant leur durée de vie, et donc leur disponibilité pour la traduction. Elle protège l’ARNm de la dégradation par des enzymes présentes dans le cytoplasme. Elle a aussi un rôle facilitateur de l’attachement des ribosomes à l’ARNm.

* **Maturation du transcrit primaire en ARNm : excision-épissage**

Rappelons que chez les eucaryotes, les gènes sont interrompus par des séquences qui ne sont représentées ni dans l’ARNm ni dans la protéine. On désigne par **introns**, les séquences non codantes de l’ADN qui interrompent la séquence du gène et **exons**, les séquences codantes parce qu’elles s’expriment. Les introns doivent être éliminés pour aboutir à l’ARNm final. Le transcrit primaire est découpé et transformé en ARNm mature. Un mécanisme appelé **excision-épissage** permet la maturation du transcrit primaire en ARNm. Il s’agit de l’élimination des introns par **excision** suivie **d’épissage des exons**, c’est à dire la réunion bout à bout des exons.

L’épissage débute par une scission de l’extrémité 5’ de l’intron qui forme une structure ramifiée appelée « lasso ». L’extrémité 3’ du premier exon déplace ensuite l’extrémité 3’ de l’intron, réunit les deux exons et libère l’intron comme un lasso



L’excision-épissage se déroule à l’intérieur du **splicéosome,** qui est un des plus grands complexes moléculaires de la cellule. Ce processus se fait dans le noyau avant l’exportation de l’ARNm vers le cytoplasme.

* **Fin de la transcription**

Des signaux de polyadénylation (séquence-type : 5’-A-A-T-A-A-A-3’ sur le brin sens) annoncent la terminaison de la transcription d’un gène. L’ARN polymérase reconnaît ce signal sur l’ADN mais poursuit la transcription. Dans ces conditions, le transcrit primaire sera remodelé par des modifications post transcriptionnelles. En effet, chez les eucaryotes, la transcription et la traduction sont séparées dans le temps et dans l’espace. **La transcription s’effectue dans le noyau**, tandis que l’essentiel de la **traduction a lieu dans le cytoplasme** ; cette séparation permet à l’ARNm de subir des modifications importantes avant d’être traduit. Des modifications sont apportées aux extrémités 5’ et 3’ ainsi qu’à la section codant les protéines de la molécule d’ARN.

**Remarque**

Certains gènes n’ont pas d’introns. D’autres peuvent en avoir 50. La taille des exons va de quelques nucléotides à 7500 et la taille des introns est tout aussi variable.

* **Exemples de modifications post-transcriptionnelles**
* La **méthylation** : addition d’un ou plusieurs groupements méthyl sur la base ou le sucre. Par exemple la méthylation de la guanosine donne la 7-méthylguanosine ;
* La **désamination** : élimination d’un groupement aminé (-NH2) de la base. Par exemple la désamination de l’adénosine donne l’inosine ;
* La **substitution d’un S** : remplacement d’un oxygène par un soufre. Par exemple la 4-thiouridine ;
* L’**isomérisation d’une base** : changement de position des atomes sur le noyau d’une base. Par exemple l’isomérisation de l’uridine donne la pseudo-uridine. ;
* La **saturation d’une double liaison** : transforme une double liaison en simple liaison. Par exemple la conversion de l’uridine en dihydro-uridine ;
* Le **changement de nucléotides** : remplace le nucléotide existant par un nouveau nucléotide.