

## Biologie moléculaire

### D- La synthèse des protéines – Première partie.

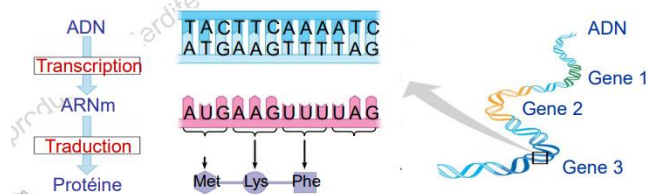
#### Généralités

Le matériel génétique ou génome contient les gènes.

L'information des **gènes codants** sert à la **synthèse des protéines**.

Leur séquence est transcrite dans le noyau en ARNm. L'ARNm rejoint le cytosol et sa séquence est traduite en acides aminés.

L'information des **gènes non codants** sert à la **synthèse des autres ARNs** (ARN ribosomiaux, de transfert, petits ARN nucléaires ou nucléolaires...) et est **uniquement transcrite dans le noyau (non traduite++)**.

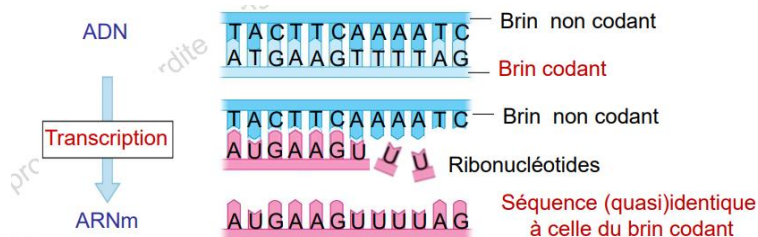


#### L'expression d'un gène codant débute par sa transcription en ARNm.

Un gène est une séquence d'ADN double brin.

L'un des brins contient l'information du gène à retranscrire dans l'ARNm (**brin codant**), l'autre brin ne contenant pas d'information (**brin non codant**).

Comme la transcription repose sur le **principe de complémentarité**, le brin non codant sert de matrice pour former l'ARNm à partir de ribonucléotides.



#### L'expression d'un gène codant s'achève par la traduction de l'ARNm.

L'ARNm rejoint le cytosol où la séquence de ribonucléotides est **traduite** en une **séquence d'acides aminés** pour former la **protéine**.

Cette étape de **traduction** repose sur un code à trois lettres appelé code génétique qui indique à quel acide aminé correspond chaque triplet de nucléotides (codons) de l'ARNm.

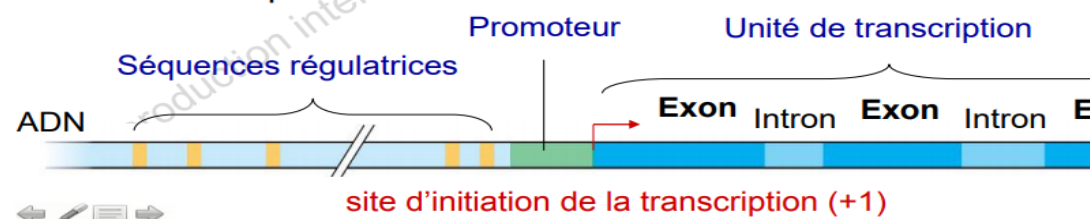
#### Un gène codant eucaryote comprend deux régions :

-Une **région destinée à être transcrite (unité de transcription)**. C'est une succession de séquences codantes (exons) et non codantes (introns), transcrite du nucléotide (+1) au signal de terminaison (signal Poly-A) .

-Et des **régions en amont et non transcrites** :

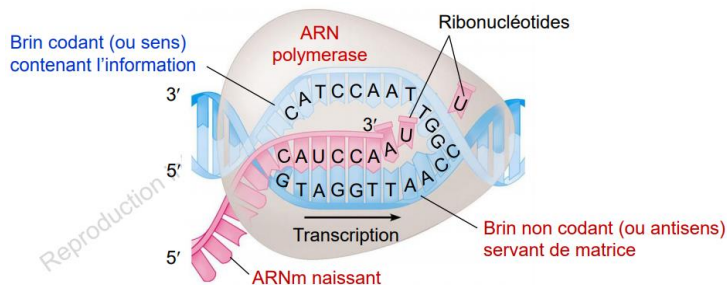
Le **promoteur**, situé près du site d'initiation de la transcription, contient dans la majorité des gènes codant la séquence TATAA appelée TATA box,

**Des séquences proximales et distales** plus éloignées assurent la régulation de la transcription



### L'ARNm eucaryote est synthétisé par l'ARN polymérase II.

Elle se fixe au promoteur du gène et recopie l'unité de transcription. Elle relie entre eux les rNTPs complémentaires du brin non codant dans le sens 5' → 3', du site d'initiation de la transcription au signal Poly-A. Sa liaison au promoteur et son activation requièrent d'autres protéines.

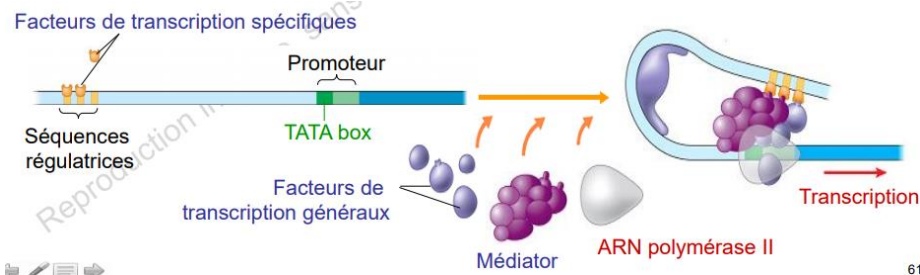


### La séquence TATA box permet d'initier la transcription .

Elle recrute la machinerie basale de transcription qui comprend :

- L'**ARN polymérase II**, régulée par phosphorylation de l'extrémité C-terminale
- Et les **facteurs généraux de transcription (TFII A, B, D, E, F et H)** qui permettent à l'ARN polymérase II de se fixer au promoteur et l'activent

Les séquences régulatrices d'amont fixent d'autres protéines. Ces protéines facilitent / répriment la transcription (enhancer/repressor)



Les séquences régulatrices des gènes codants varient . ++

**Chaque gène possède sa propre combinaison de séquences régulatrices.** ++

Une séquence donnée ne peut fixer qu'un facteur de transcription donné.

L'ensemble des séquences régulatrices d'un gène permet la fixation d'une combinaison particulière de facteurs de transcription.

**Chaque gène est régulé / une combinaison de facteurs de transcription.**

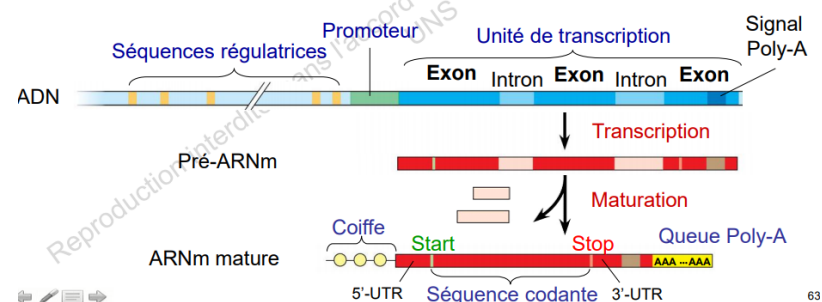
Il s'exprimera ou non selon leur présence, variable selon le type cellulaire.

### Un gène codant eucaryote est transcrit en ARN pré-messager.

Ce **transcrit primaire** subit une **maturation pour pouvoir être traduit.**

Des **modifications co-transcriptionnelles** assurent sa maturation en ARNm :

- L'**ajout d'une « coiffe » à l'extrémité 5'** et d'une **« queue » Poly-A en 3'**
- L'**excision (élimination) des introns** et l'**épissage des exons (ligation)** de telle sorte que sa séquence codante entre les signaux Start/Stop soit ininterrompue



### La structure des gènes procaryotes et eucaryotes est différente :

- Les **gènes procaryotes** sont **compacts** (absence d'introns) et **regroupés**.
- Les **mêmes séquences régulatrices** contrôlent **un ensemble de gènes** (opéron) qui est transcrit en un long ARNm (polycistron) ne nécessitant pas de maturation (cf. modèle de l'opéron lactose).
- Il n'existe pas de membrane séparant le noyau rudimentaire et le cytosol.** ++ La **transcription** et la **traduction** des **gènes procaryotes** est **simultanée**. Les **ribosomes** peuvent débiter la traduction d'un ARN messager alors que sa **synthèse n'est pas encore achevée**.

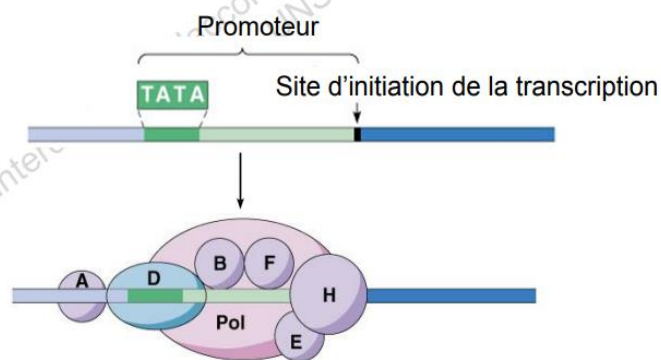
**L'initiation de la transcription nécessite plusieurs étapes.**

Elle **début**e par la fixation du **complexe TFIID** sur le **promoteur**.

Il **se lie** à la **TATA Box** grâce à l'une de ses **sous-unités** appelée **TBP**.

Les autres complexes et l'ARN polymérase sont ensuite recrutés.

L'ensemble forme la machinerie basale de transcription encore inactive tant que **l'extrémité C-terminale de l'ARN polymérase n'est pas phosphorylée**.+++

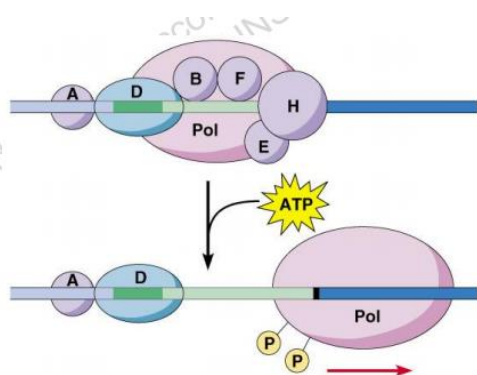
**La phosphorylation de la polymérase initie la transcription.**

Le **complexe TFIIF** phosphoryle l'extrémité **C-terminale** de l'**ARN Pol II**.

Elle est activée et la transcription de l'ARNm débute (phase d'élongation).

La maturation de l'ARN pré-messager débute simultanément.

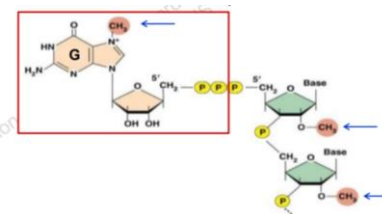
**D'autres phosphorylations de l'extrémité C-terminale** recrutent les enzymes de **maturation du pré-ARNm** (couplage élongation-maturation).

**Les modifications co-transcriptionnelles du pré-ARNm :**

**La coiffe comprend plusieurs modifications de l'extrémité 5' :**

**Ajout d'un nucléotide à guanine à l'extrémité 5'-P du transcrit** et **méthylation (-CH<sub>3</sub>) de la guanine ajoutée et du ribose des deux premiers nucléotides.**

La coiffe protégera le transcrit de la dégradation, augmentant sa durée de vie, et est nécessaire à sa reconnaissance par les ribosomes.

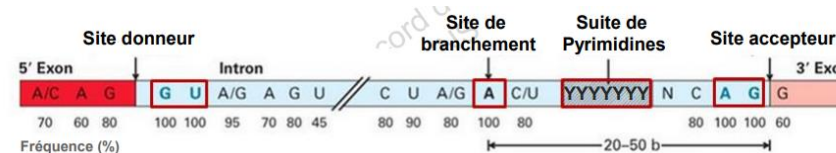


**L'épissage** fait intervenir des **séquences introniques** appelées **consensus**.

Elles sont **quasi invariables** et **retrouvées dans tous les gènes codants** :

-**Site donneur d'épissage (GU)** au **début** et **accepteur (AG)** à la **fin** de l'intron.

-**Site de branchement (A)** et **suite de pyrimidine (Y)** **avant** la fin de l'intron.

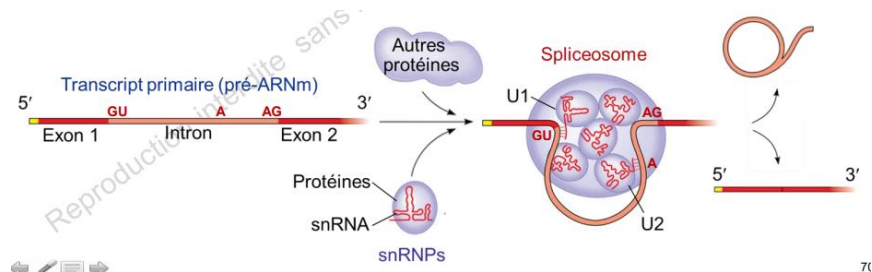
**Le spliceosome est le complexe enzymatique qui assure l'épissage.** +++

Il est formé par les ribonucléoprotéines U1, U2, U4, U5 et U6 constituées de diverses protéines et des petits ARNs nucléaires (snRNAs) correspondants.

### Les ribonucléoprotéines snRNPs « repèrent » et définissent les introns.

Le site donneur et de branchement fixent respectivement U1 et U2 via leur snRNA respectif qui s'apparie par complémentarité avec le pré-ARNm.

Les complexes U1 et U2 recrutent les autres complexes ce qui permet de rapprocher les exons pour les réactions suivantes: l'intron est éliminé sous la forme d'un lasso (ou lariat) et les exons sont reliés entre eux.

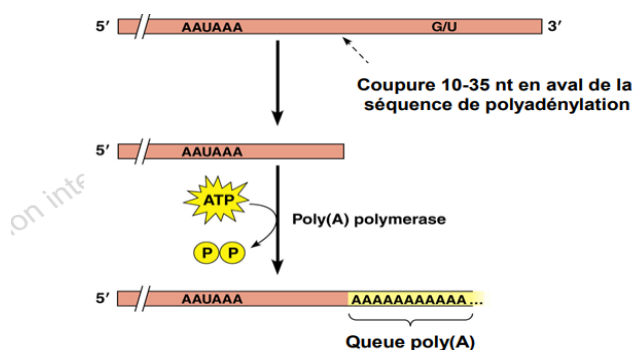


La transcription s'arrête vers la séquence de polyadénylation AAUAAA.

Des complexes coupent le pré-ARNm quelques nucléotides après.

La poly(A) polymérase ajoute une queue poly(A) d'environ 250 nucléotides.

La polyadénylation ralentit aussi la dégradation du transcrit mature.

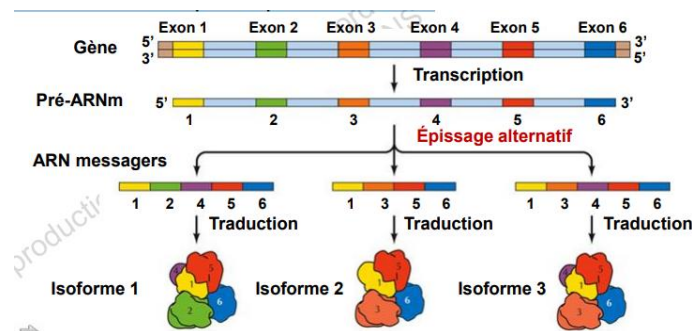


### La transcription d'un gène peut produire différents ARNm.

La plupart des gènes subissent un épissage dit alternatif.

Certains exons peuvent être ignorés lors de l'épissage (saut exons) et ne pas être inclus dans l'ARNm, des introns peuvent être retenus, etc....

A partir d'un seul gène, on peut ainsi obtenir différents ARNm qui seront traduits en isoformes protéiques différentes.



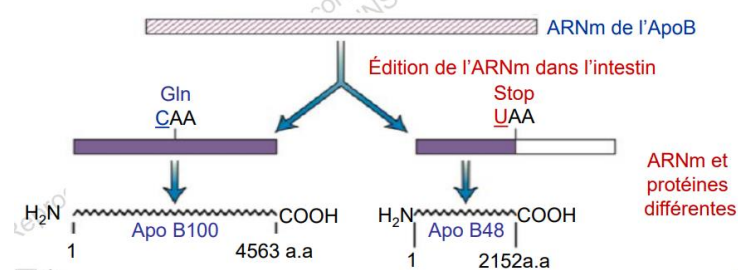
### La séquence d'un ARNm mature peut encore être changée (édition).

Exemple de l'ARNm de l'apolipoprotéine B (ApoB).

Dans l'hépatocyte, cet ARNm n'est pas modifié et est traduit en ApoB100.

Dans l'entérocyte, une cytosine de l'ARNm est désaminée en uracile.

Il y a arrêt de la traduction (codon Stop) et production d'ApoB48, tronquée.



**Points clés**

Les gènes codants eucaryotes permettent la synthèse des protéines.

Ils possèdent un promoteur et des séquences régulatrices non transcrits

Le promoteur minimal est constant dans la plupart des gènes

Il est constitué par la TATA box qui fixe la machinerie basale

Les séquences régulatrices sont variables selon les gènes et fixent des facteurs de transcription spécifiques qui régulent la machinerie basale

Leur séquence transcrite est morcelée (succession d'exons et introns)

Ces gènes sont transcrits par l'ARN polymérase II chez les eucaryotes

Elle utilise le principe de complémentarité des bases

Le transcrit primaire subit une maturation (coiffe, polyadénylation, épissage)

Plusieurs ARNm et protéines peuvent être produits à partir d'un seul gène

*Désolée pour cette partie que j'avais oublié dans ma première fiche, mais vaut mieux tard que jamais...*