

# Fondements de la biologie moléculaire

Mp : trans porcary Ce cours a des références

## LA TRANSCRIPTION CHEZ LES PROCARYOTES

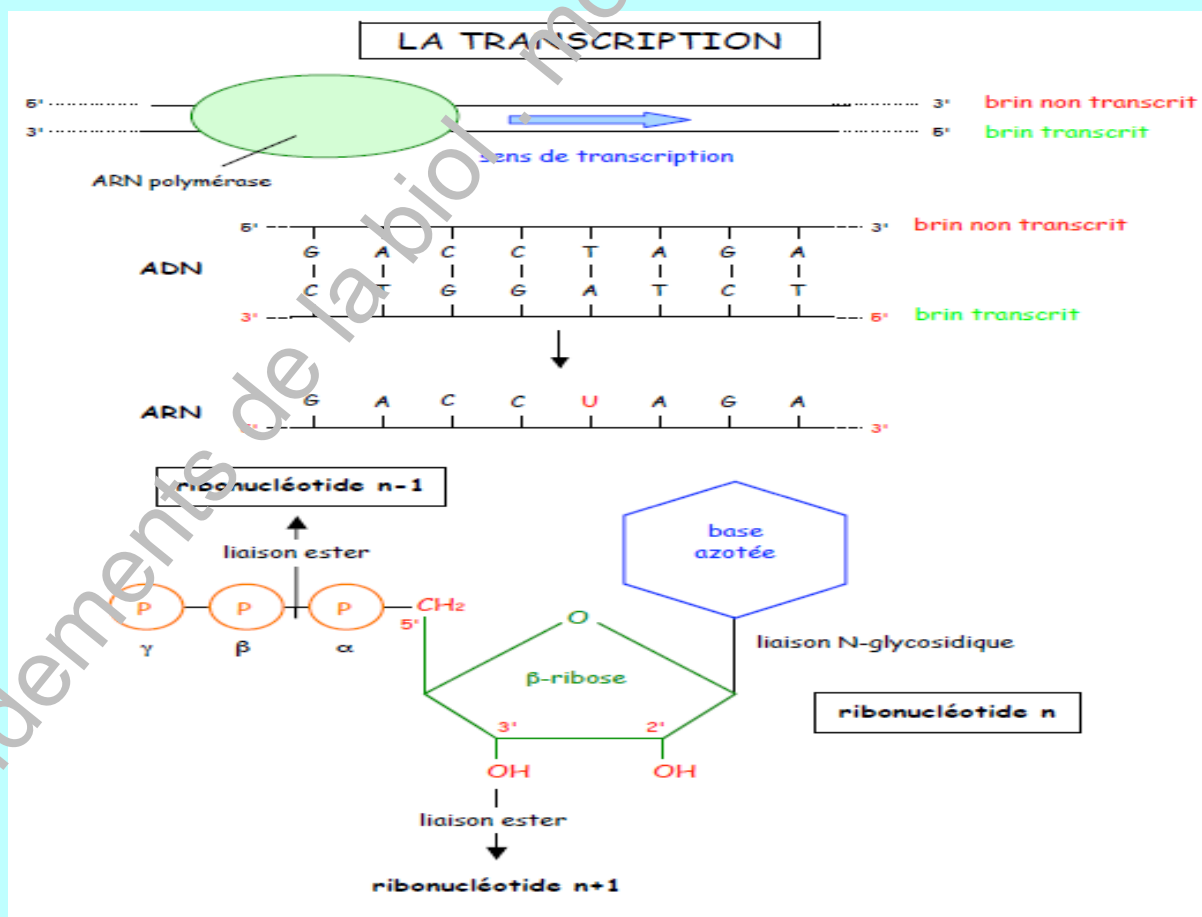
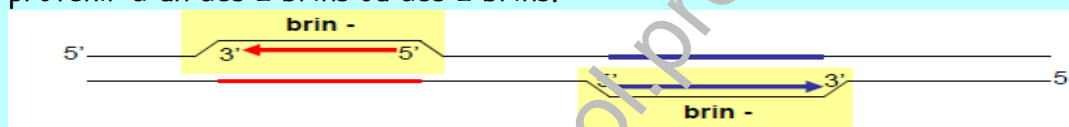
Comme pour la réplication :

- L'ADN sert de matrice
- La synthèse d'ARN se fait de 5' → 3'
- Se passe en 3 étapes : initiation, élongation, terminaison,
- L'initiation se fait au niveau d'une région particulière (promoteur),
- La synthèse nécessite l'ouverture de l'ADN,
- La terminaison se fait au niveau d'une région particulière (terminateur).
- L'ARN polymérase **recopie le brin « + »** de l'ADN en ARN, elle reconnaît une séquence localisée en 5' du gène : le **promoteur** qui contient des motifs particuliers

5' ---CGTTAACGTAGTCATCGT--- brin + non matrice  
3' ---GCAATTGCATCAGTAGCA--- brin - brin matrice

5' ---CGUUAACGUAGUCAUCGU--- ARN = séquence du brin +

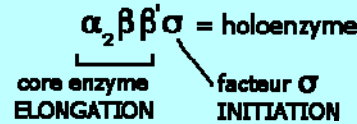
Seul un des 2 brins d'ADN est transcrit à la fois mais les produits de transcription peuvent provenir d'un des 2 brins ou des 2 brins.



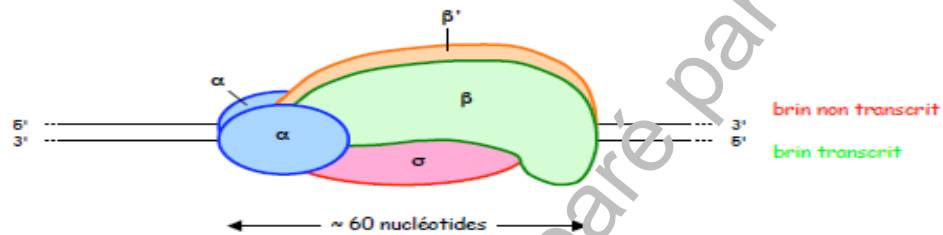
# Fondements de la biologie moléculaire

Le brin d'ARN synthétisé a donc la même séquence (avec des U à la place des T) que le brin d'ADN 5'-3' dont la séquence code pour le gène, puisque la matrice d'ADN 3'-5' servant à la synthèse de l'ARN en est la séquence complémentaire. Une molécule d'ADN donne de nombreuses molécules d'ARN lorsque la transcription est activée, plusieurs ARN polymérase pouvant travailler à la file sur la molécule d'ADN. C'est une première étape d'amplification de l'information allant donner la protéine.

- Les procaryotes possèdent **1 seule RNA pol**, constituée de plusieurs sous unités:

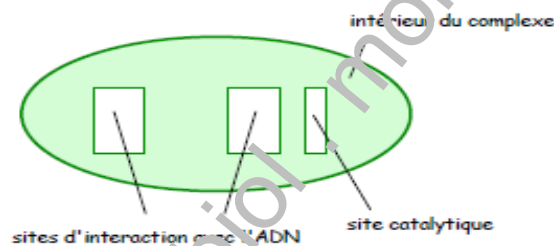


L'ARN polymérase bactérienne ou holoenzyme (500 kDa) est une enzyme multimérique composée de 5 sous-unités  $\alpha_2\beta\beta'\sigma$ :



Ces 5 sous-unités s'organisent pour former 3 domaines fonctionnels de l'enzyme :

- 2 domaines d'interactions avec l'ADN
- 1 site catalytique pour la formation des liaisons phosphodiester.



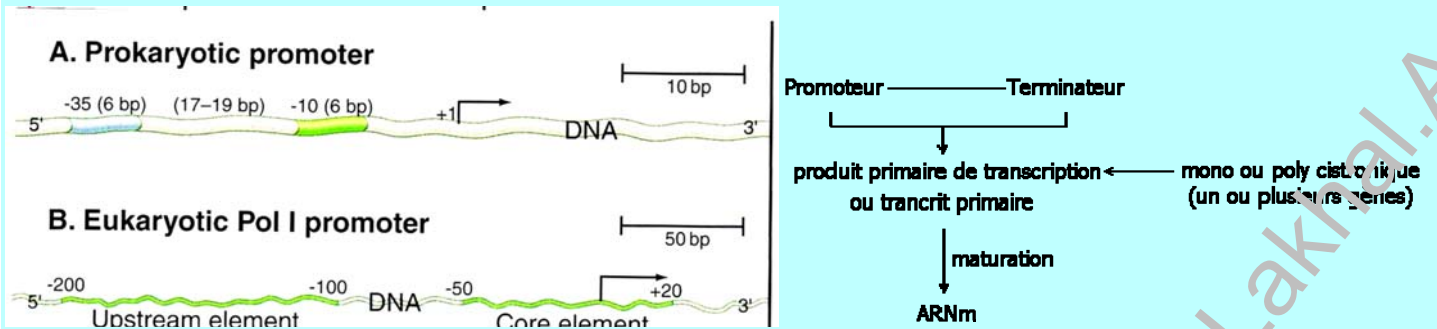
SOUS UNITÉ	FONCTION
$\beta$	se charge de la fixation de nucléosides tri-phosphates
$\beta'$	se charge de la fixation de la matrice
alpha	reconnaissance probable des promoteurs
sigma	reconnaît les promoteurs "forts"

Ce holoenzyme se charge de la synthèse d'ARNt, r ou m indifféremment. Le core enzyme se lie à l'ADN à n'importe quel niveau et n'entraîne pas l'ouverture de la double hélice d'ADN. L'holoenzyme se lie à l'ADN sur des sites de fixation spécifiques (promoteurs) et entraîne l'ouverture de l'ADN à ce niveau. La synthèse de l'ARN avec comme matrice un des 2 brins d'ADN peut commencer. Très rapidement sigma se détache, car il a une forte affinité pour ce site, si il reste l'élongation ne se fait pas.

## 1. L'initiation

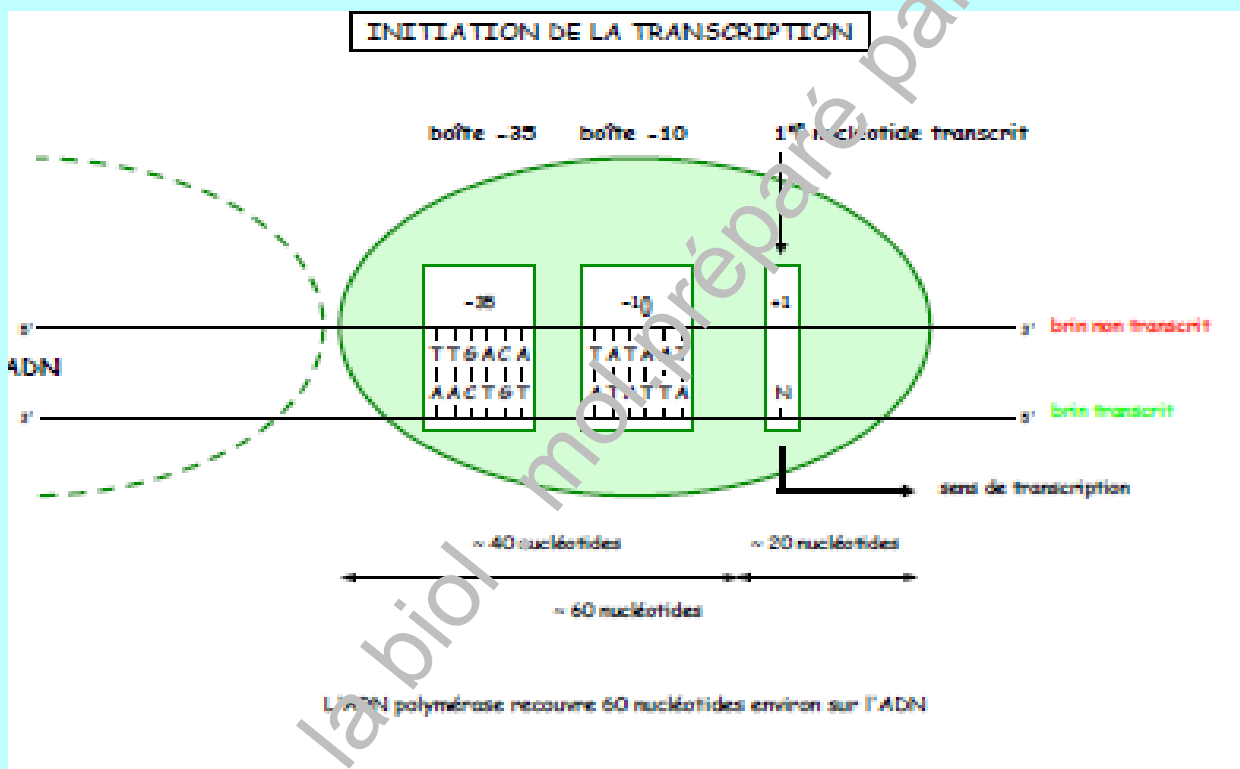
La reconnaissance du promoteur chez les procaryotes se fait par : les facteurs  $\sigma$

# Fondements de la biologie moléculaire



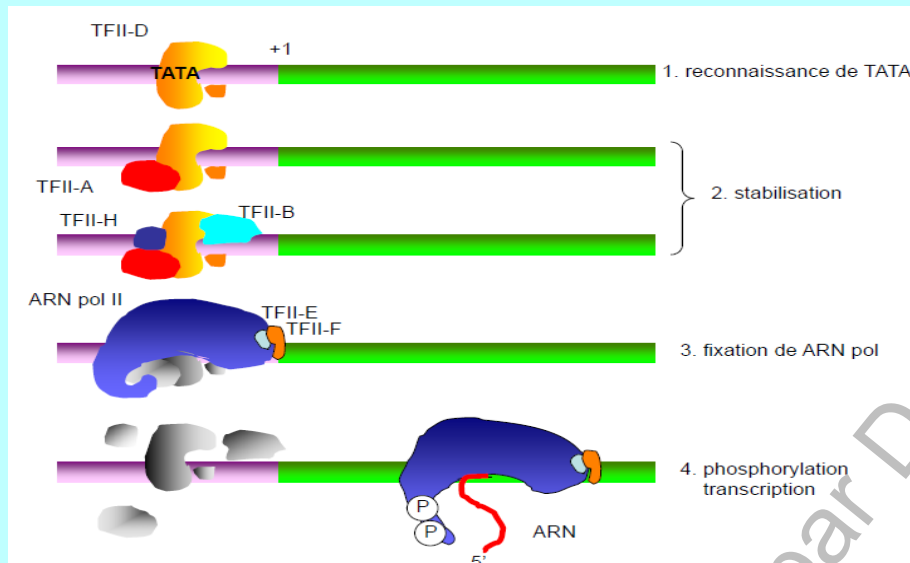
Il existe 2 séquences consensus en amont du point de départ :

- séquence **-10** ou encore **TATA box** (car riche en A et en T)
- séquence **-35**. Ces 2 séquences sont séparées par 17 p b.



**Eucaryotes** : La reconnaissance du promoteur implique nombreux facteurs protéiques de transcription spécifiques des différentes ARN polymérases : ARNpol II (TFIIA-H), ARN pol I (TAFs), ARN pol III (TFIIIA-C)

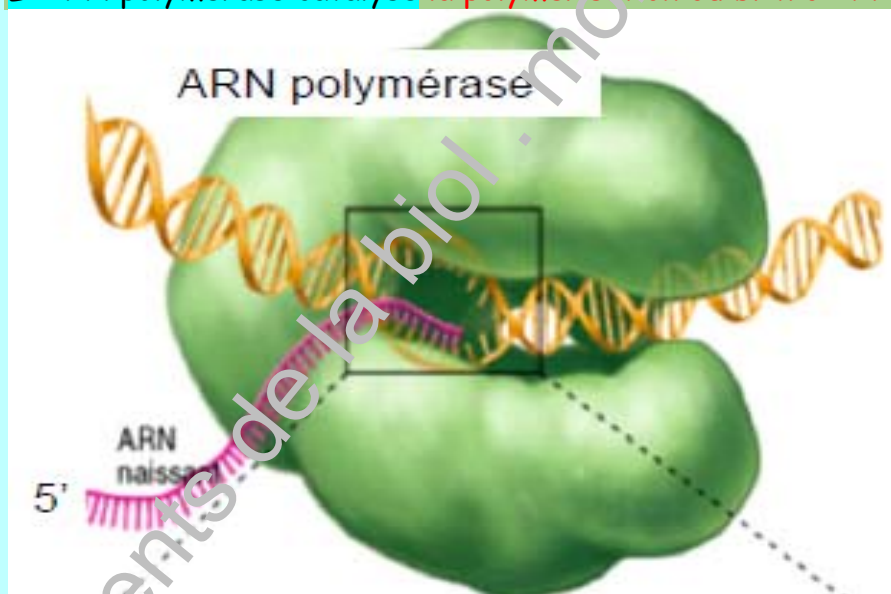
- La fixation de l'ARN polymérase est conditionnée par la séquence d'intervention des facteurs de transcription

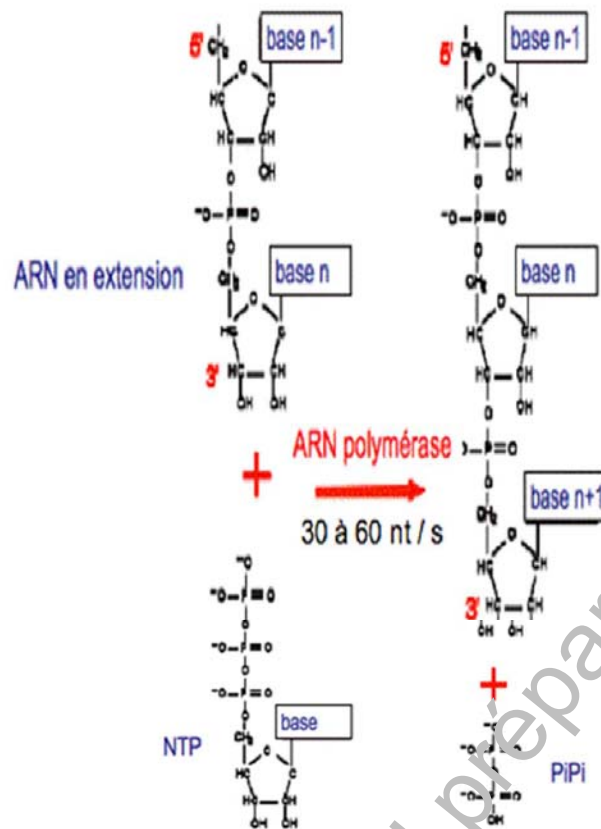


## 2. Phase d'élongation

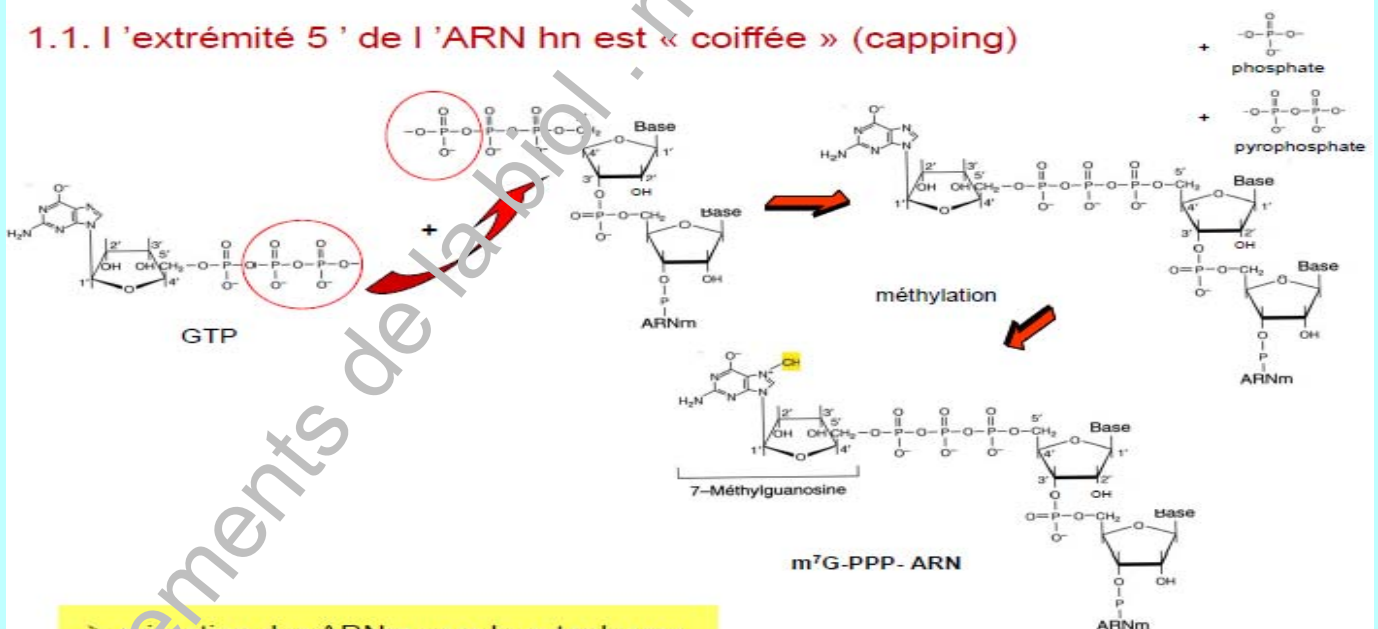
- L'élongation se fait dans le sens 5' - 3' par formation de liaisons phosphodiester
- La lecture du brin « - » par l'ARN polymérase s'accompagne de l'ouverture de la double hélice d'ADN et est suivi par le ré-appariement des 2 brins d'ADN.
- La transcription s'accompagne d'un remodelage des histones et des nucléosomes.

L'ARN polymérase catalyse la polymérisation du brin d'ARN



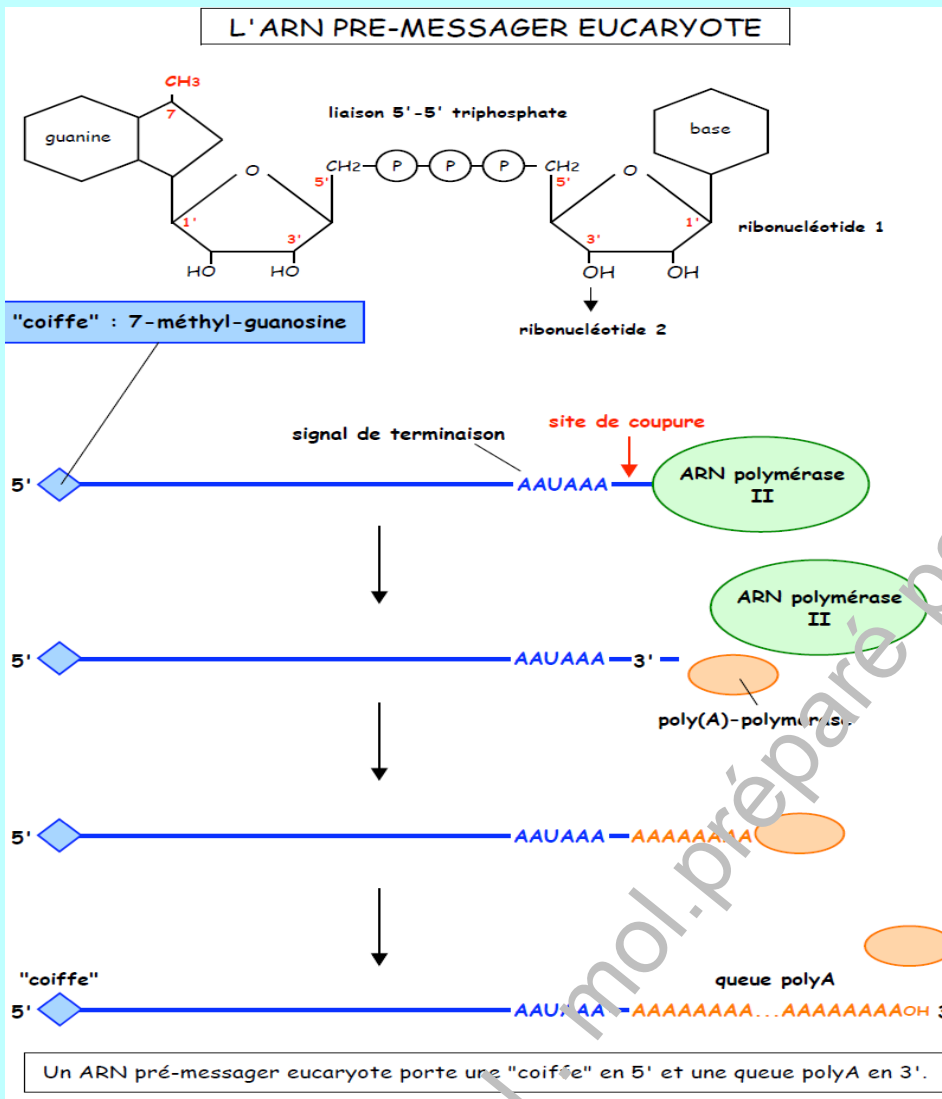


## 1.1. L'extrémité 5' de l'ARN hn est « coiffée » (capping)



- migration des ARNm vers le cytoplasme
- favorise l'initiation de la traduction
- stabilise l'extrémité 5' des ARNm

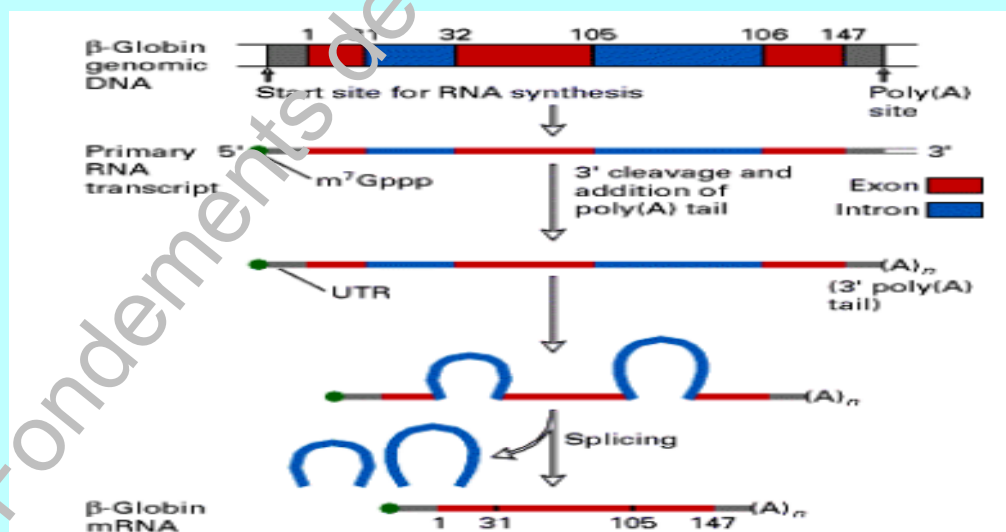
**Ajout de la coiffe du côté 5' des ARN naissants.** Cette coiffe consiste en une 7-méthylguanylate ( $m^7G$ ), c'est-à-dire une guanidine modifiée. Les deux premières réactions sont catalysées par une enzyme de capping qui s'associe à l'ARN polymérase peu après l'initiation de la transcription. Deux méthyltransférases différentes catalysent deux autres réactions. La *S*-adénylméthionine (*S*-Ado-Met) est la source du groupement méthyle ( $CH_3$ ) pour les deux dernières étapes de méthylation. Le capping est nécessaire à l'exportation des ARNm hors du noyau.



L'extrémité 3' de la plupart des ARNm eucaryotiques est polyadénylée. Une fois la transcription achevée, l'ARNm sera clivé 15-30 nucléotides après des AMP seront ajoutés par une poly(A) polymérase pour former une queue poly(A). La polyadénylation facilite l'exportation des ARNm hors du noyau, les protège des dégradations une fois dans le cytosol et facilite leur traduction.

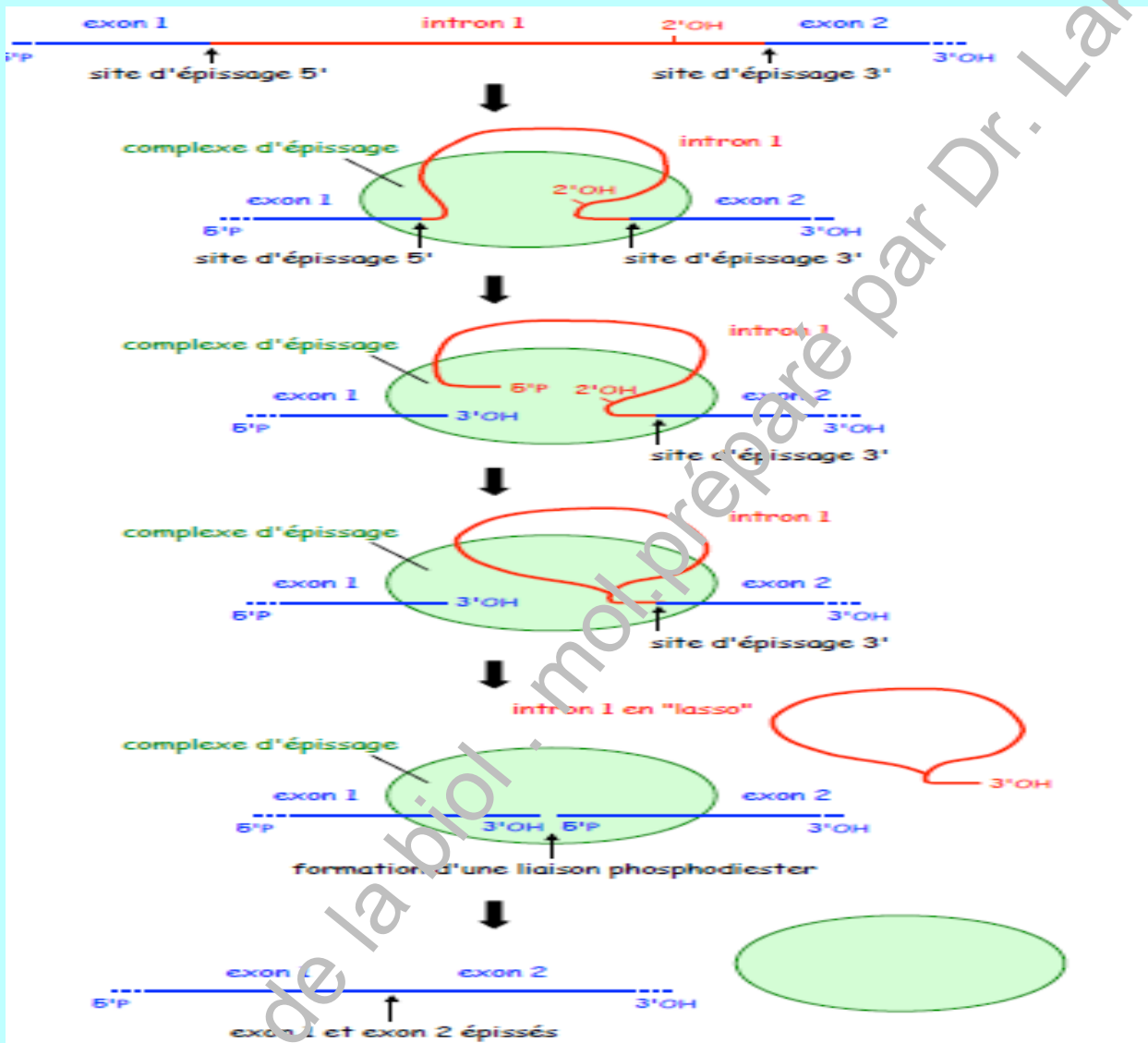
## L'épissage

La plupart des gènes eucaryotiques sont morcelés: La séquence codante est interrompue par des séquences non codantes, appelées **introns**. Par extension, les séquences codantes sont des **exons**. Les gènes sont transcrits sous forme d'un **ARN pré-messager**, lequel sera ensuite **cappé, polyadénylé**, puis les séquences introniques seront éliminées par **épissage** pour générer l'ARNm mature.

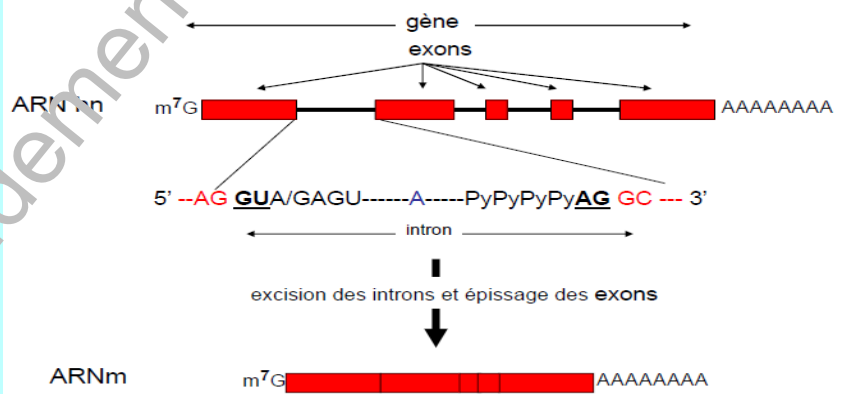


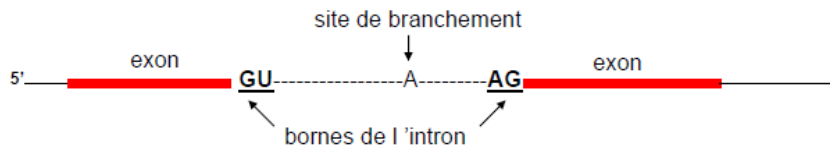
# Fondements de la biologie moléculaire

Exemple du gène  $\beta$ -globine, qui comprend 3 exons (rouge) et 2 introns (bleu). Les introns interrompent la séquence codante au niveau des codons correspondant aux acides aminés 31 et 32, et 105 et 106. La transcription commence un peu en amont de l'exon 5' et s'étend en aval de l'exon 3', produisant un ARN pré-messager contenant les séquences introniques et exoniques ainsi que deux régions non codantes supplémentaires (gris) aux deux extrémités du transcrit primaire (régions UTR). Les régions introniques seront éliminées secondairement par l'épissage. Les régions UTR ne sont pas épissées.

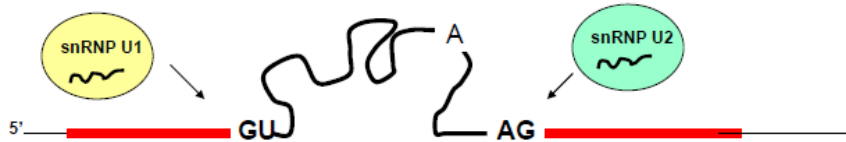


## 1.3. les ARN<sub>m</sub> sont épissés

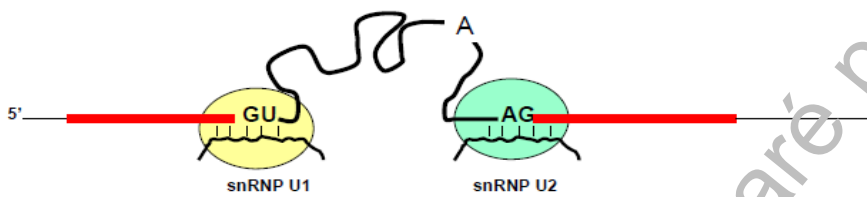




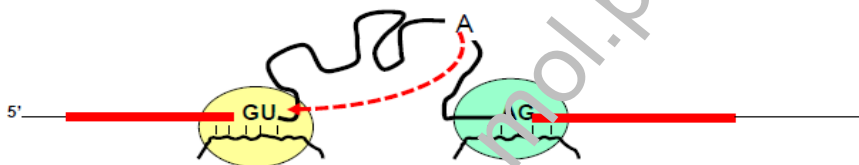
o Reconnaissance des bornes donneur et accepteur



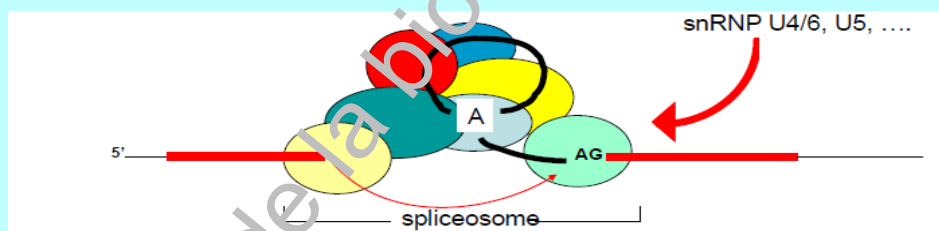
Fixation des snRNP U1 et U2 sur les bornes



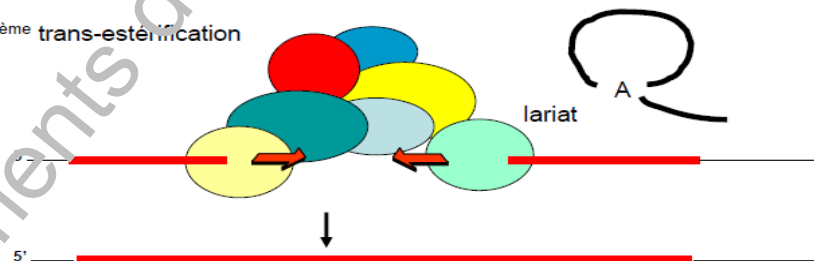
o 1<sup>ère</sup> réaction de trans-estérification



Recrutement du spliceosome



o 2<sup>ème</sup> trans-estérification



Les mutations affectant les sites d'épissage modifient la nature du transcrit et sont généralement la cause de maladies.

## Principe général du mécanisme d'épissage.

L'épissage est catalysé par des snRNP (small nuclear Ribonucleotide Particles), plus d'autres protéines, l'ensemble constituant le spliceosome. Les RNP sont des structures multimoléculaires composées de protéines et de petits ARN. U1 et U2 se fixent d'abord sur le pré-messager, puis U4 et U6 viennent interagir avec U1 et U2, ce qui rapproche les

# Fondements de la biologie moléculaire

deux extrémités **exoniques**. Puis l'activité catalytique du **spliceosome** permet de **cliver** la séquence **intronique** et de **liquer les séquences exoniques**

L'épissage est réalisé par le **spliceosome** constitué de protéines associées à des snRNA : les **snRNP** (U1, U2, U4, U5, U6)

1 snRNP = 1 snRNA + 10-20 polypeptides

## 3. Maturation des ARNr

Les ARNr sont le produit de **gènes multi-copies** (28+18+5,8s et 5s)

• les ARNr eucaryotes 5,8s, 18s et 28s sont transcrits par l'**ARN pol I**

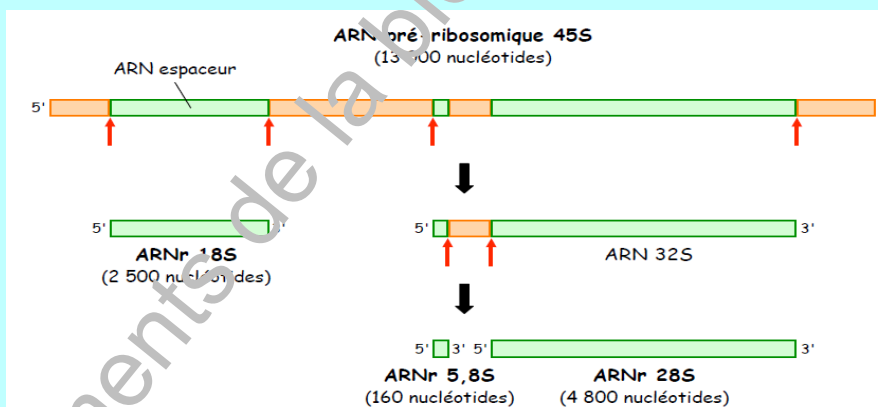
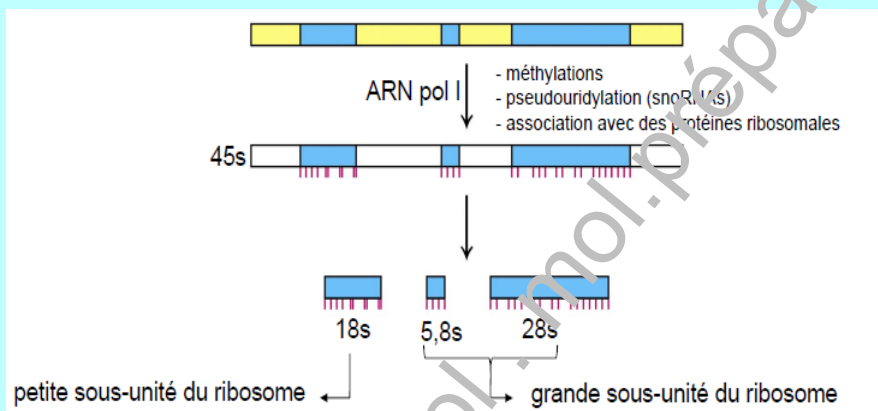
L'ARNr 5s est transcrit par l'**ARN pol III**

• les ARNr 5,8, 18 et 28s sont produits et maturés au **niveau du nucléole**

• présence d'un **5' NTP**, **absence de polyadénylation** : **protection des extrémités**

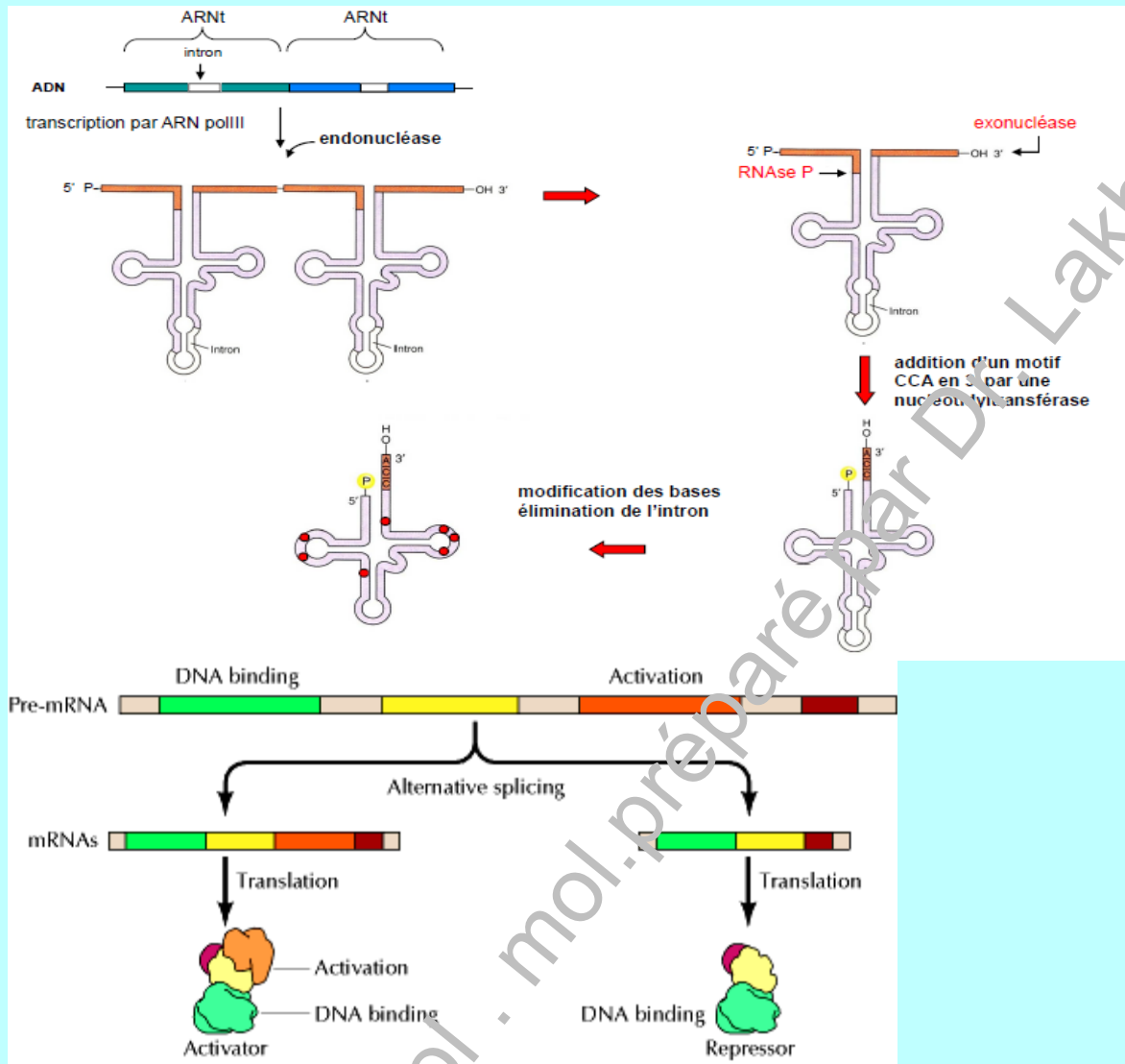
• par la formation de **structures II** et une association avec des protéines

• **absence d'épissage**



↑ **site de clivage des endonucléases spécifiques**

• Après formation de **pré-ribosomes**, ceux-ci sont exportés dans le **cytosol** et subissent **une maturation en ribosomes 40s et 60s**



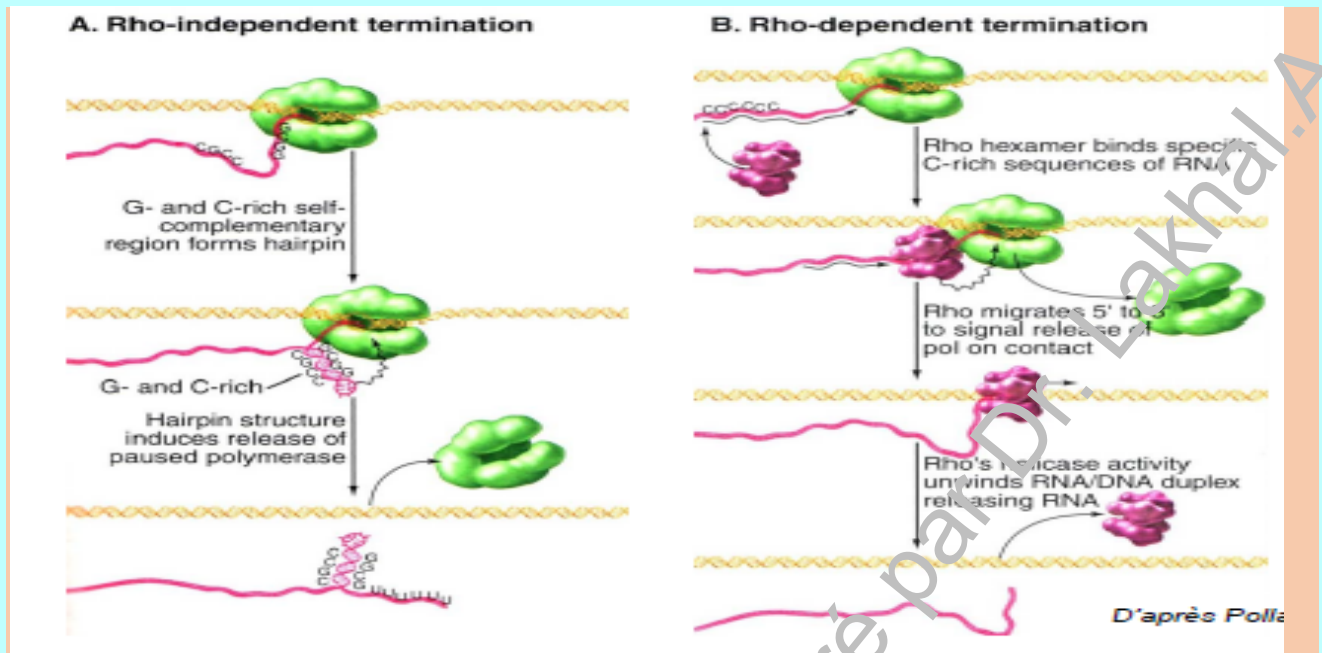
**Splicing alternatif d'un pré-mRNA qui code un facteur de transcription.** Dans cet exemple, le facteur de transcription est codé par 4 exons: le premier code le domaine de liaison à l'ADN, le deuxième un domaine de fonction inconnue, le troisième le domaine permettant l'activation de la transcription et le quatrième un domaine de fonction inconnue. Le pré-mRNA peut subir deux splicings différents. Dans le premier cas (à gauche), les 3 introns sont éliminés et les 4 exons sont ligés, générant un mRNA qui code le facteur de transcription pleine longueur. Dans le deuxième cas (à droite), l'exon 2 est ligé à l'exon 4 au lieu d'être ligé à l'exon 3, ce qui génère un facteur de transcription inactif car dépourvu du domaine d'activation de la transcription **à titre informatif pas à l'examen**

## La terminaison

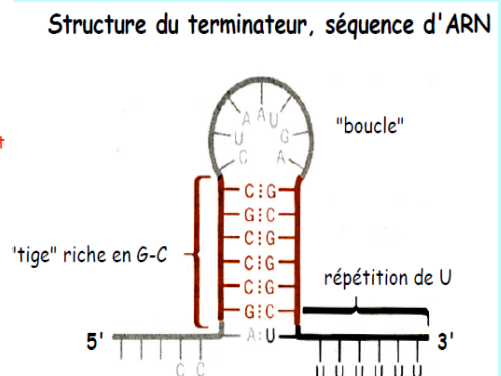
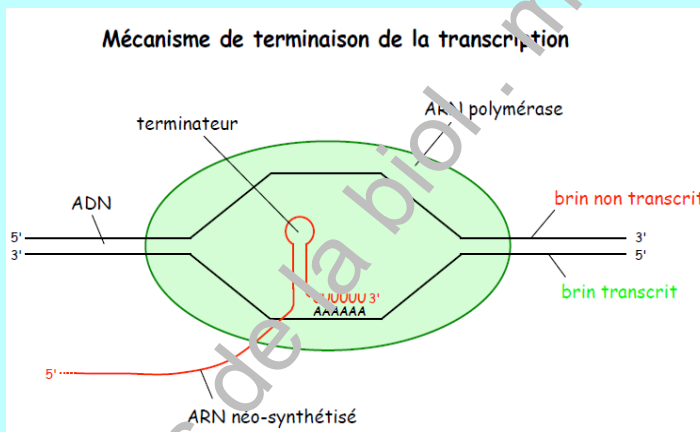
La transcription s'arrête lorsque l'ARN polymérase atteint un site de terminaison sur l'ADN et se détache. elle se fait de 2 façons:

- spontanément au niveau d'une séquence particulière de l'ADN, « **Termineur** »
- grâce à un **facteur protéique** de terminaison (**facteur rho**).

# Fondements de la biologie moléculaire



• -**Facteur rho** : rho est une protéine hexamérique qui se lie à l'ARN grâce à la protéine NusA qui est fixée au niveau d'une tige boucle de l'ARN en cours de synthèse. Cette tige boucle est due à une séquence particulière de l'ADN appelée site nut. La formation de la tige boucle ralentit l'enzyme au minimum favorisant la liaison de NusA à l'ARN. Une fois rho fixée sur l'ARN elle se dirige vers l'extrémité 3' qui est en cours de synthèse. Lorsqu'elle a rattrapée l'extrémité 3' elle décroche l'ARN. L'ADN se referme et l'enzyme libère l'ADN.

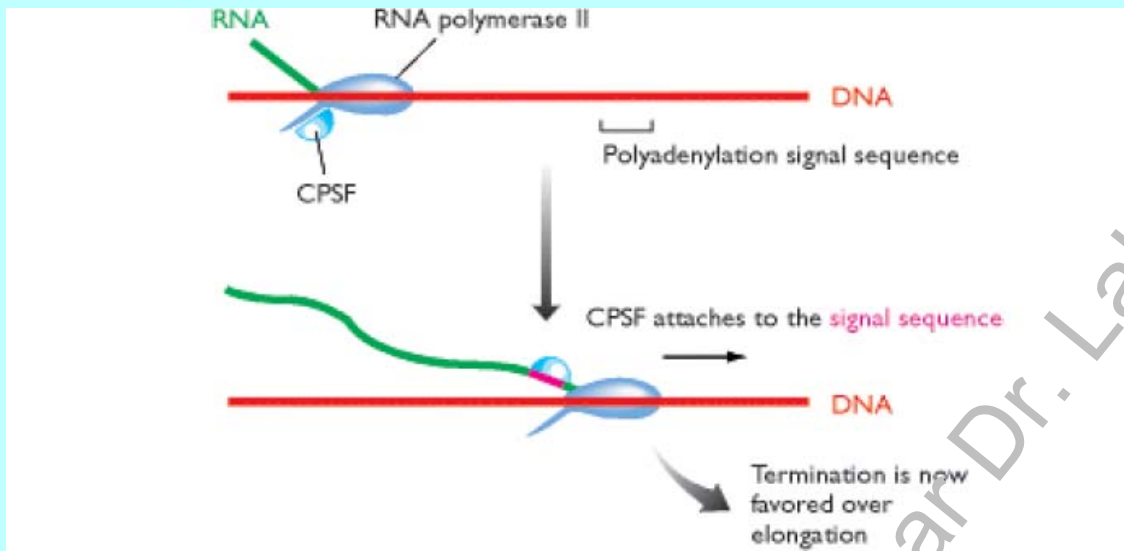


Chez les bactéries, la transcription s'achève au niveau d'une séquence palindromique inversée. La transcription de cette séquence palindromique inversée entraîne la formation d'une épingle à cheveux au niveau de l'ARN néosynthétisé, ce qui déstabilise le complexe de transcription.

Le terminateur déstabilise les liaisons faibles entre les sous-unités de l'ARN polymérase et entraîne leur séparation et l'arrêt de la transcription



# Fondements de la biologie moléculaire



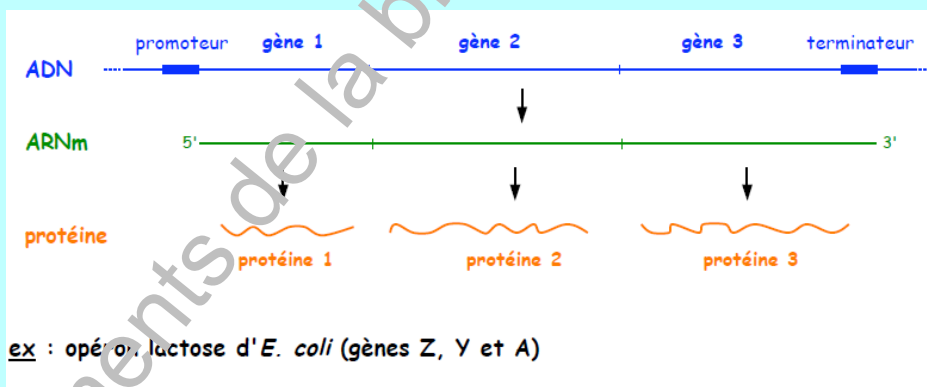
Chez les eucaryotes, la transcription se termine aux alentours de la séquence de polyadénylation (qui a aussi un rôle dans la traduction). Une fois la séquence de polyadénylation transcrite (en rose), la protéine CPSF (cleavage and polyadenylation specificity factor) qui était associée à l'ARN polymérase s'y lie. La perte de l'interaction CPSF/ARN polymérase déstabilise le complexe de transcription.

## - Initiation et régulation de la transcription chez les procaryotes

Chez les procaryotes, il y a 4 sortes de régulation qui peuvent se combiner entre elles :

### Régulation négative :

Un répresseur est fixé à l'endroit du promoteur et inactive la transcription. la présence d'un ligand peut être nécessaire pour que le promoteur s'accroche ou au contraire, un ligand peut interagir avec le répresseur et le pousser à se détacher.



### Régulation positive :

L'ARN polymérase a besoin d'un activateur pour commencer la transcription. Soit l'activateur peut être présent uniquement s'il est en interaction avec un ligand, soit au contraire, l'interaction avec un ligand l'empêche de se lier au promoteur et d'activer la transcription. Il est courant que plusieurs gènes, se succédant sur la molécule d'ADN, soient transcrits en même temps, à la suite sur une même molécule d'ARN qui est alors dite polycistronique. La transcription de ces gènes est alors régulée de la même façon, ce qui est

intéressant pour des protéines qui ont un rôle dans un processus commun. L'ensemble de ce promoteur et ces quelques gènes est un **opéron**.

## 2- Chez les eucaryotes

La régulation de la transcription est **plus complexe** que chez les procaryotes. **Des facteurs de transcription**, qui forment des **complexes protéiques** à l'endroit de la **TATA box**, une région riche en **T et A** un peu en amont du **site d'initiation**, permettent à l'ADN polymérase de reconnaître le site d'initiation, de s'y fixer et de commencer la transcription. Les complexes protéiques formés par les facteurs de transcription impliquent **de nombreuses protéines** qui interagissent entre elles : ceci permet une réponse, finement modulée, à un spectre très large **de signaux**. Des éléments **amplificateurs** situés à grande distance du gène interviennent aussi, la molécule d'ADN pouvant faire des boucles. Les protéines se fixent sur des **amplificateurs** sont responsables de la régulation de l'assemblage du complexe protéique au niveau du **promoteur**.

### Résumé.

(A) Chez les eucaryotes, la transcription a lieu dans le noyau. Le transcrit primaire, ou **ARN pré messager**, contient **les introns et les exons**, il subit des modifications à ces deux extrémités (**capping et polyadénylation**) qui contribueront à son exportation dans le **cytosol**, lieu de la traduction, et les **introns** sont éliminés par le mécanisme d'**épissage**. Le **capping et l'épissage** démarrent avant que la transcription de l'**ARN pré messager** soit **achevée**. La complexité des processus chez les eucaryotes permet une plus grande finesse de régulation, et permet de synthétiser plus d'un type de protéine par gène.

(B) Chez les bactéries, le transcrit ne subit **aucune maturation**. La traduction peut **démarrer dès que l'ARN commence à être synthétisé**. Le taux final d'une protéine dans la cellule dépend de l'efficacité de chaque étape et des taux de dégradation des ARN et des protéines.