

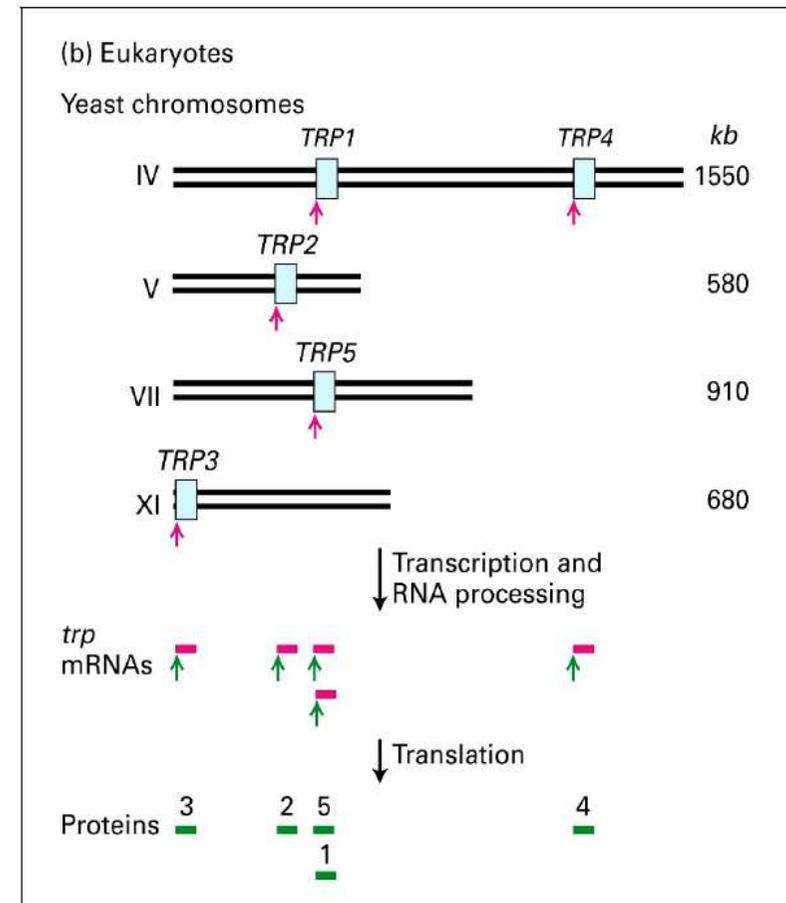
## **2. LA TRANSCRIPTION EUCARYOTE**

# LES GENES EUCARYOTES

## Particularités des gènes eucaryotes

Le génome eucaryote est plus complexe que le génome procaryote :

- Chaque gène est transcrit à partir de **son propre promoteur**, pour produire un seul ARNm (dit **monocistronique**) afin de produire une seule protéine.
- Les gènes codant des protéines intervenant dans une même voie métabolique sont en général situés **sur des chromosomes différents**.
- Certains gènes existent parfois en plusieurs copies (Histones, ARNr, ARNt...).
- **Les gènes eucaryotes sont morcelés** : certaines parties du gène (ADN) ne sont pas retrouvées dans l'ARNm mature. Ces segments uniquement présents dans l'ADN ont été appelés **introns** ; les parties présentes à la fois dans l'ADN et dans l'ARNm (les parties codant pour la protéine) ont été appelées **exons**.



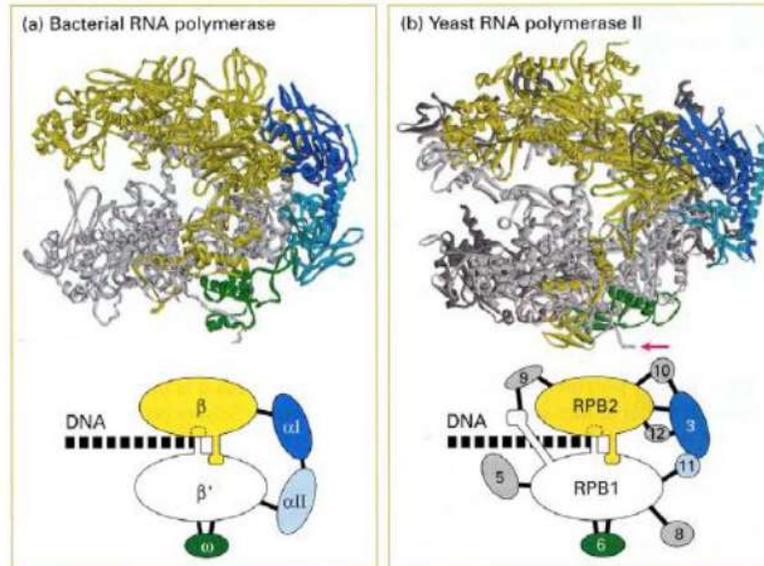
# LES ARN POLYMERASES EUCARYOTES

Il existe trois types de promoteurs eucaryotes, pour les trois types d'ARN polymérase les reconnaissant : Pol-I, Pol-II et Pol-III.

- Les promoteurs **Pol-I** se trouvent face aux ARNr 5,8 18S et 28S. **ARN polymérase I**
- Les promoteurs **Pol-II** se trouvent face aux ARNm et quelques snARN. **ARN polymérase II**
- Les promoteurs **Pol-III** se trouvent face aux ARNt, ARNr 5S et l'ARN 7SL **ARN polymérase III**

## ARN pol II

CTD  
=  
queue C-terminale



Lodish, H., et al. (2004) *Molecular Cell Biology*.

2 grandes sous-unités  
équivalentes à  $\beta'$  et  $\beta$   
2 sous-unités équivalentes à  $\alpha$   
1 sous-unité équivalente à  $\alpha'$   
  
4 autres sous-unités communes

5 sous-unités ARN pol I  
3 sous-unités ARN pol II  
7 sous-unités ARN pol III

**CTD de l'ARN polymérase II** : 26 à 52 copies YSPTSPS → Phosphorylation

# LES PROMOTEURS EUCARYOTES

Contrairement à l'enzyme bactérienne  
les ARN polymérases eucaryotes ne sont pas capables de reconnaître un promoteur

Le promoteur est reconnu par des **facteurs généraux de transcription (GTF)**  
qui à leur tour recrutent l'ARN polymérase  
et favorisent l'initiation de la transcription

## **Cœur du promoteur**

Autour du site de démarrage de la transcription

Séquence la plus courte permettant  
la fixation des facteurs généraux de transcription  
le recrutement de l'ARN polymérase  
et l'initiation de la transcription

## **Eléments de contrôle agissant à proximité du cœur du promoteur**

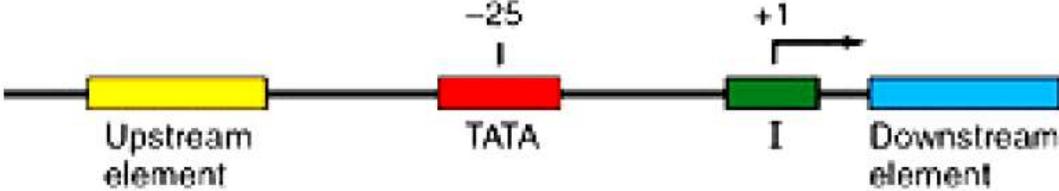
Localisés dans les 100 à 200 pb en amont du cœur du promoteur  
Séquences reconnues par d'autres facteurs de transcription  
ubiquitaires ou spécifiques du type cellulaire  
Augmentent l'efficacité de la transcription

# LES PROMOTEURS EUCARYOTES

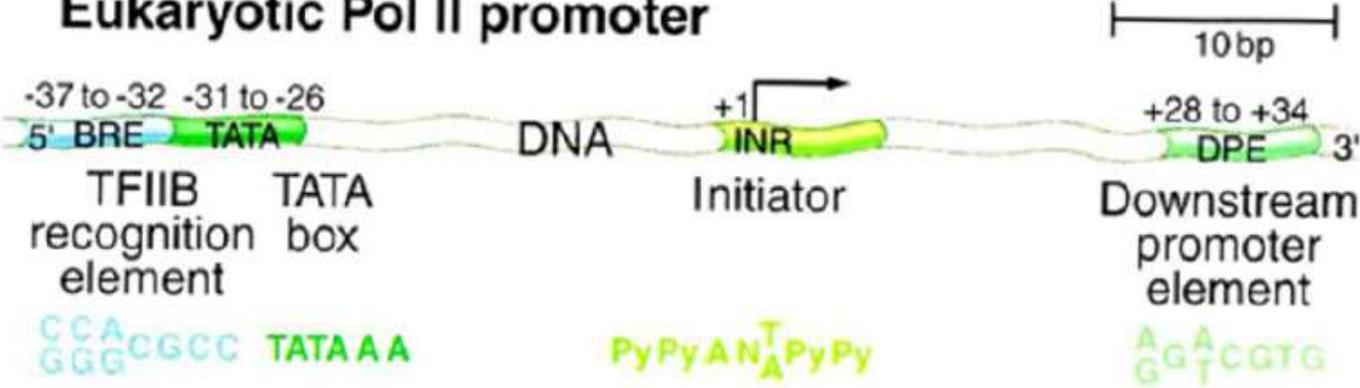
### Promoteurs spécifiques de l'ARN polymérase II

Ces promoteurs contiennent une boîte TATA et au moins une autre séquence importante en amont.

- x Des séquences régulatrices en amont : Upstream element
- x Des séquences régulatrices en aval : Downstream element
- x Une séquence promotrice au niveau du site d'initiation de la transcription : I ou INR



### Eukaryotic Pol II promoter



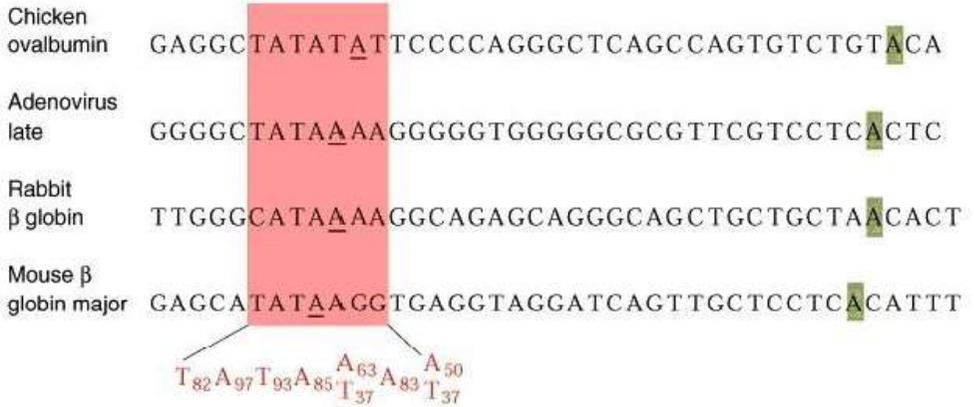
# LES PROMOTEURS EUCARYOTES

## La TATA Box :

La majorité des promoteurs de gènes protéiques eucaryotes contiennent une boîte TATA (similaire à la séquence -10 procaryote).

x La boîte TATA est trouvée à 25-30 paires de bases en amont du site de départ de transcription (TSS). Une position relativement constante dans les promoteurs eucaryotes. C'est le seul signal dans le promoteur se trouvant à une distance définie du TSS.

x La séquence d'environ 8 paires de bases contient pratiquement que des adénines et thymines, et tend à être encadrée par des séquences riches en guanine et cytosine, ces dernières pouvant participer à la fonction du promoteur.



x La boîte TATA est très semblable à la séquence TATA des procaryotes, la position mise à part (-10 chez les procaryotes).

x TATA consensus : *GTATAAAAGGCGGGG* (mais beaucoup de variations). Le consensus de la TATA box est faible et cet élément est même absent dans de nombreux promoteurs.

# LES PROMOTEURS EUCARYOTES

## L'initiateur :

- x L'initiateur (INR), se trouve près du site de début de transcription, entre les positions -3 et +5.
- x Il y a peu ou pas de similarité entre les initiateurs de différents promoteurs, toutefois, la première base du mRNA transcrit tend à être un A, souvent flanqué de pyrimidines.
- x La RNA Pol II peut parfois initier la transcription avec l'INR seul, dans des promoteurs simples sans boîte TATA.

## Autres séquences promotrices possibles :

Séquences courtes de 6-20 nt affectant généralement l'efficacité de l'initiation de la transcription.

- x **CCAAT box**
  - 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
- x **GC box :**
  - 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.
  - Présent dans le promoteur des gènes constitutifs.
- x **CRE**
- x **AP2 box...**

# LES FACTEURS GÉNÉRAUX DE LA TRANSCRIPTION

Les facteurs TFII sont les facteurs généraux de la transcription permettant l'utilisation des promoteurs de l'ARN pol II :

**TFIID** (800 kDa)  
= TBP *TATA-binding protein* (38 kDa)  
+ TAF<sub>II</sub> *TBP-associated factors* (30-250 kDa)

**TFIIA** requis *in vivo*  
Interaction TBP

**TFIIB** monomérique  
Permet la fixation de l'ARN polymérase

**TFIIF** tétramérique  
Complexe pré-existant avec l'ARN polymérase II  
Sous-unité RAP74 : ATPase ADN-dépendante (ADN hélicase)  
Sous-unité RAP38 : homologie avec  $\sigma$

**TFIIH** nonamérique  
Sous-unité XPB = ATPase ADN-dépendante (ADN hélicase)  
→ Complexe de pré-initiation ouvert

Les **facteurs généraux de la transcription** forment un complexe qui se lie au cœur du promoteur.

A pas confondre avec ce qu'on appelle les **facteurs de transcription** qui se lient aux éléments de contrôle en amont du promoteur pour moduler l'efficacité de la transcription (facteurs activateurs ou inhibiteurs) !!

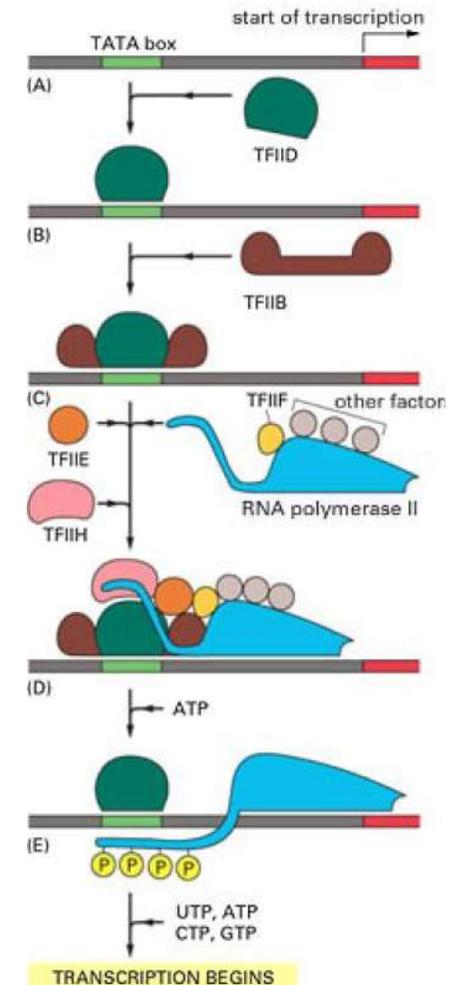
# INITIATION DE LA TRANSCRIPTION

## Fixation de l'ARN pol II sur le promoteur :

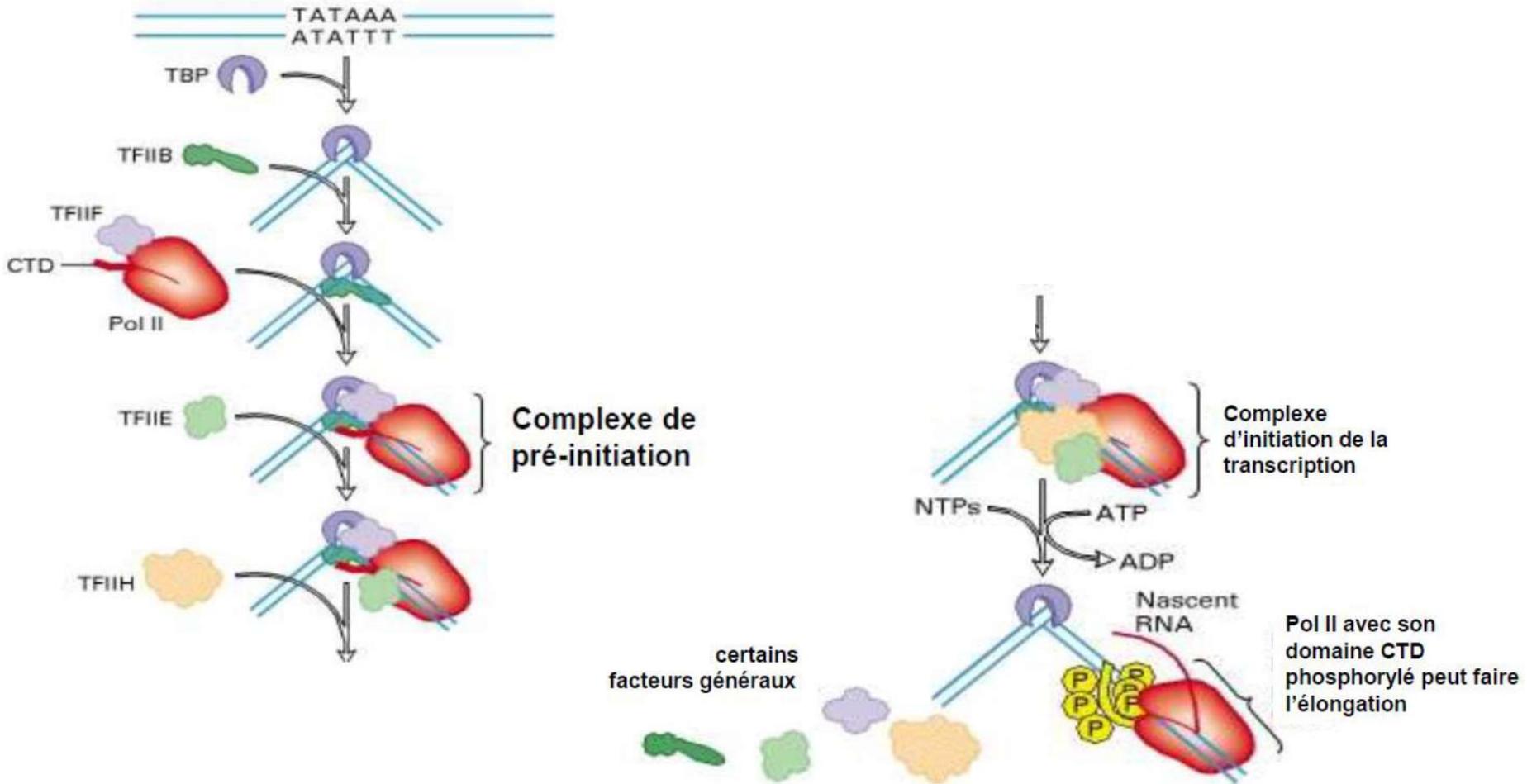
- De nombreux **facteurs de transcription** (TF II) doivent se fixer sur le promoteur afin de permettre la fixation de l'ARN polymérase. En effet, l'ARNpol II n'a qu'une faible affinité pour l'ADN.
- L'ordre d'assemblage de ces facteurs de transcription est essentiel. Le premier facteur qui se fixe est le facteur TFIID. Il est en outre composé d'une protéine appelée **TBP** (TATA box Binding Protein) qui se fixe spécifiquement à la séquence consensus du promoteur. Sa structure en « fer à cheval » lui permet une bonne fixation sur l'ADN.
- L'association de toutes ces protéines sur le site d'initiation forme le **complexe d'initiation de la transcription**.

La spécificité du promoteur reconnu (c'est à dire le contrôle de l'expression gène spécifique) est assurée par d'autres facteurs TFII, tels :

- Sp1 qui se fixe sur la GC box
- NF1 ou CTF qui se fixent sur le CAATbox

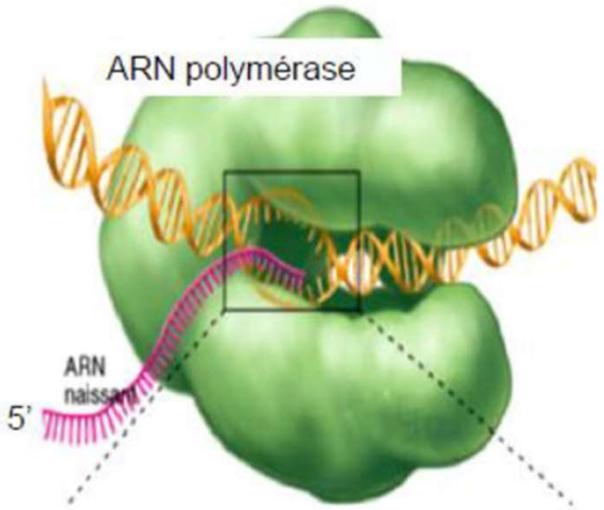
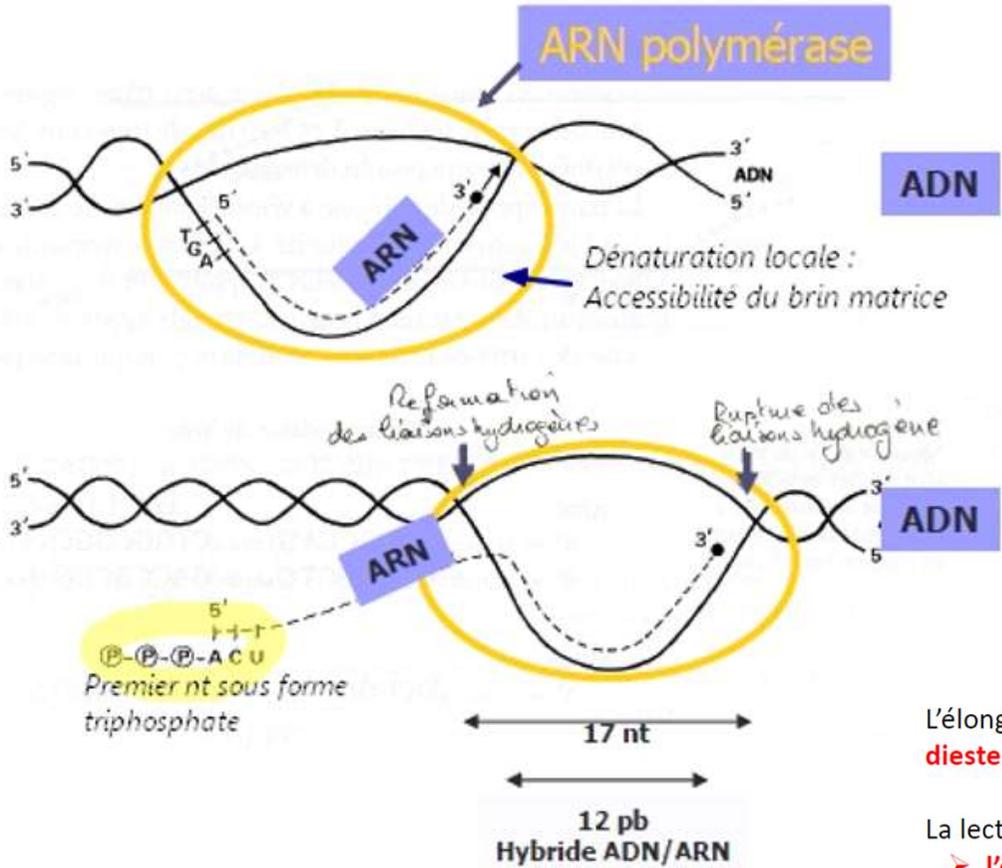


# INITIATION DE LA TRANSCRIPTION



# ELONGATION DE LA TRANSCRIPTION

La zone dénaturée et qui se déplace s'appelle « **bulle de transcription** ».



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.

# ELONGATION DE LA TRANSCRIPTION

## Intervention de facteurs d'élongation

- Protéines diminuant la probabilité de dissociation ADN/ARN,
- Protéines associées en complexes de remodelage de la chromatine (problèmes des nucléosomes),
- Problème des superenroulements créés par le déplacement de l'ARNpol sur l'ADN :
  - en amont, l'ouverture de la double hélice et l'avancée de l'ARNpol crée des superenroulement positifs,
  - en aval, on observe la formation de superenroulements positifs

Il existe des topoisomérases capables de résoudre ses superenroulements grâce à une coupure simple brin de l'ADN.

# TERMINAISON DE LA TRANSCRIPTION

Chez les eucaryotes, la phase de terminaison de la transcription diffère selon le type d'ARN polymérase impliquée.

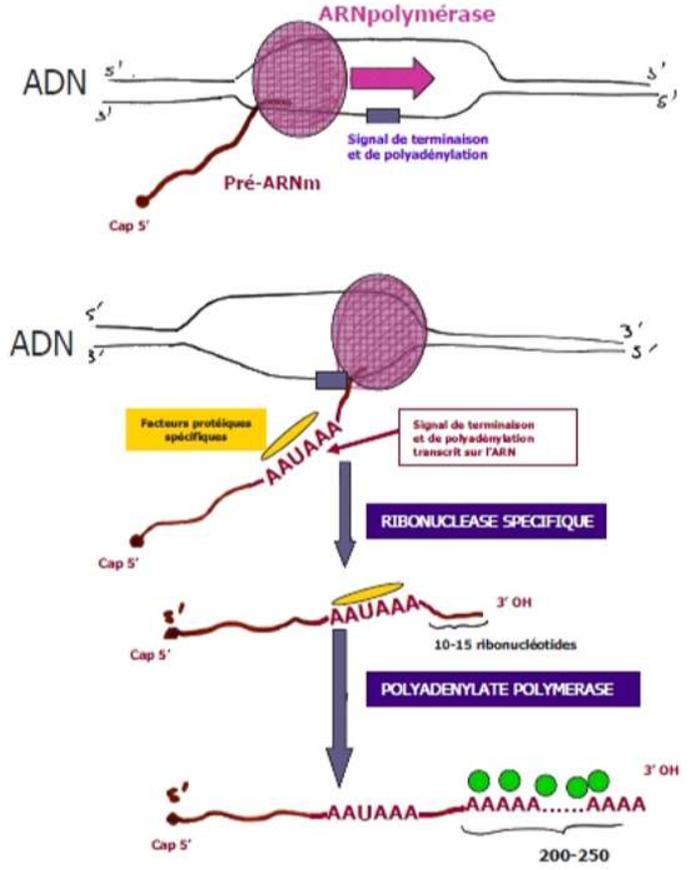
**ARN pol I** : intervention d'un facteur protéique qui bloquerait la transcription

**ARN pol II** : la terminaison est couplée à la maturation des ARN messagers (voir plus loin)

**ARN pol III** : implication de séquences riches en U dans l'ARN qui provoquent la dissociation de l'ARN pol

**Pour l'ARN pol II, la transcription s'arrête de façon stochastique 500 à 2000 nucléotides en aval (après) une séquence correspondant à un signal de maturation (signal de polyadénylation).**

(Le mot stochastique est synonyme d'aléatoire, en référence au hasard et s'oppose par définition au déterminisme)

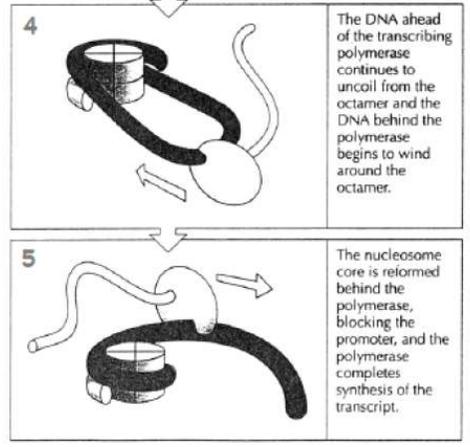
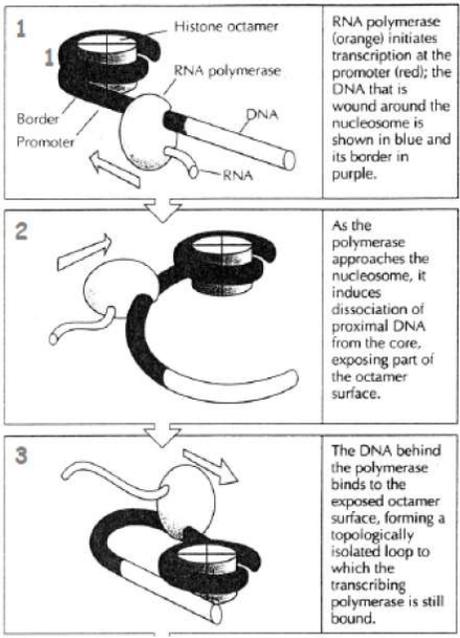
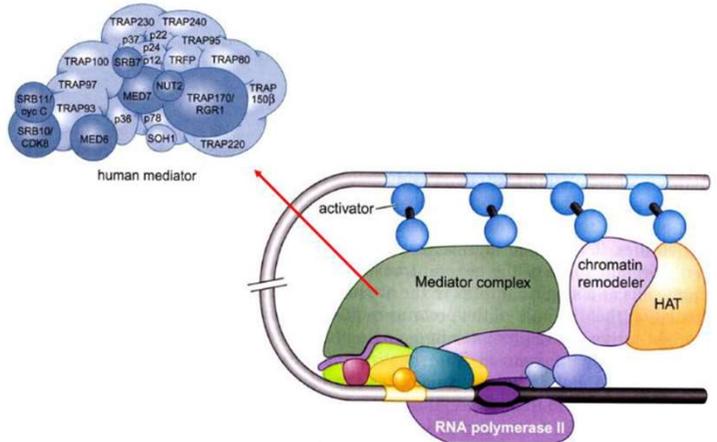


La terminaison de la transcription par l'ARN pol II est couplée à la maturation de l'ARNm

# TRANSCRIPTION EUCARYOTE ET CHROMATINE

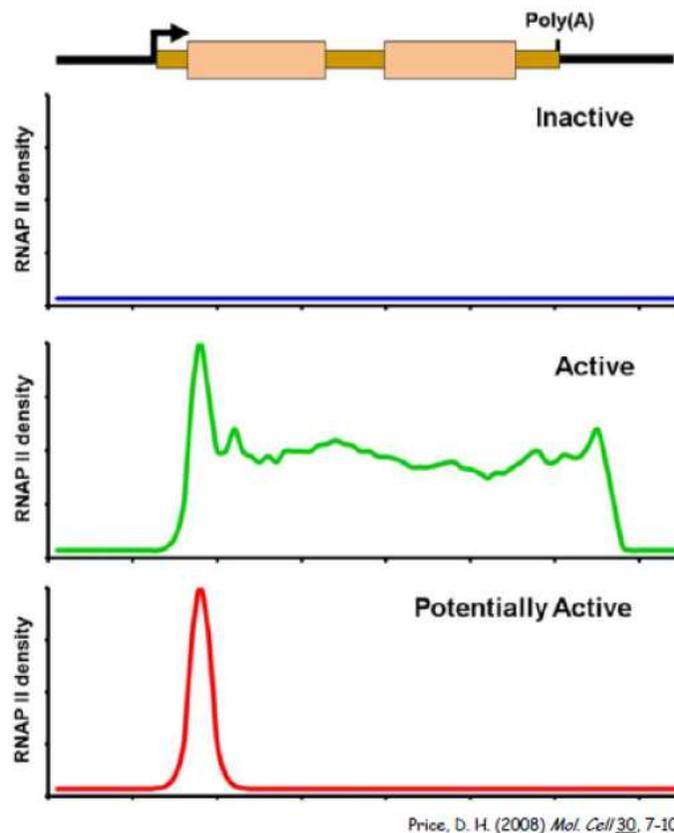
Au cours de la transcription chez les eucaryotes, la machinerie de transcription doit faire avec les octamères d'histones associés à l'ADN. Des protéines s'occupent de dissocier l'ADN des histones de manière à ne pas interrompre le processus transcriptionnel.

In vivo, la formation du complexe pré-initiateur nécessite de nombreux facteurs protéiques et s'accompagne d'une modification de la structure de la chromatine.



Des chaperons des histones facilitent le passage de l'ARN polymérase  
Dissociation/réassociation H2A-H2B  
Nucléosome → pause de l'ARN polymérase  
Intervention de TFIIS

# CARTOGRAPHIE DES INTERACTIONS ARN pol II / ADN



Dans une cellule donnée, on trouve un certain nombre de molécules d'ARN polymérase II, dont la répartition le long de l'ADN n'est pas uniforme.

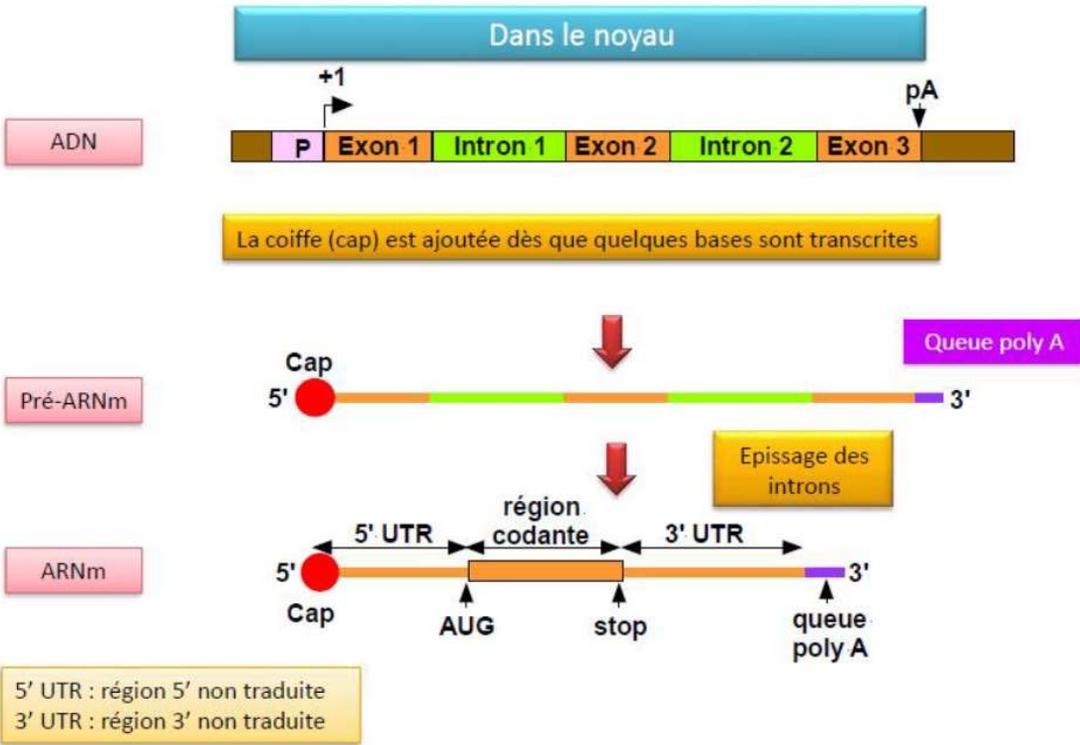
On ne trouve pas de polymérase associée aux séquences d'ADN correspondant à des gènes inactif dans cette cellule.

En revanche, lorsqu'on regarde ce qui se passe au niveau des gènes activement transcrits dans cette cellule, on retrouve des ARN pol II réparties tout le long de la séquence de ces gènes.

Enfin, si on regarde à un instant  $t$  des gènes qui ne sont pas transcrits mais dont la transcription est susceptible de s'activer en réponse à des signaux spécifiques, on trouve des molécules d'ARN pol II associées à une région restreinte en aval du promoteur de ces gènes, qui sont « en attente ».

# MATURATION DES ARNm

L'ARN polymérase II des eucaryotes produit en fait un ARN immature appelé **ARN pré-messager**. Plusieurs modifications interviennent pour permettre l'obtention d'un **ARN messenger mature** qui permettra la production d'une protéine.

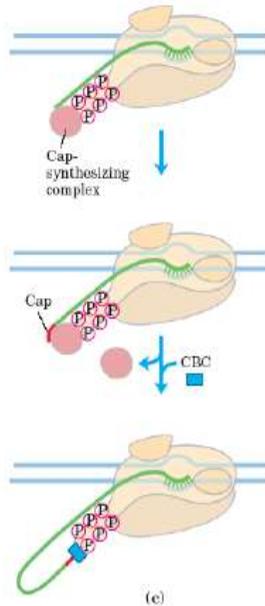
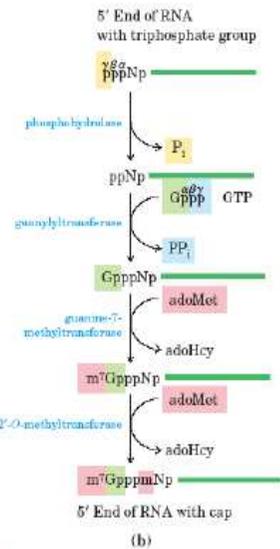
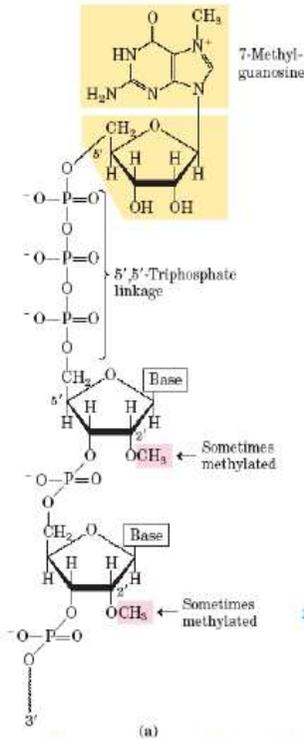


3 modifications principales ont lieu dans le noyau :

- l'ajout d'une coiffe en 5'
- l'ajout d'une queue polyA en 3'
- l'épissage des introns

# MATURATION DES ARNm

- Ajout de la coiffe** = méthyl-7-guanosine triphosphate par le CSC (Cap Synthesis Complex)
- 1. Phosphatase** modifiant l'extrémité 5'triphosphate en 5'diphosphate,
  - 2. Guanyl transférase** ajout d'un GMP en sens inverse (liaison 5'-5')
  - 3. Methyl transférase** méthylant l'azote 7 de la guanine à partir de S-adénosylmethionine



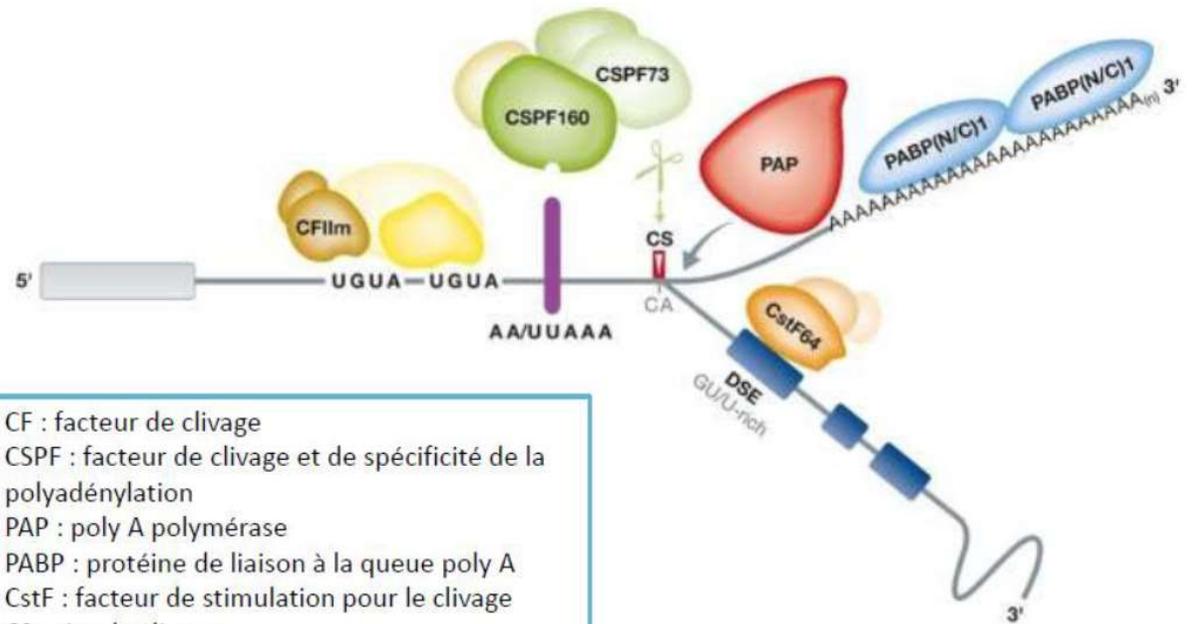
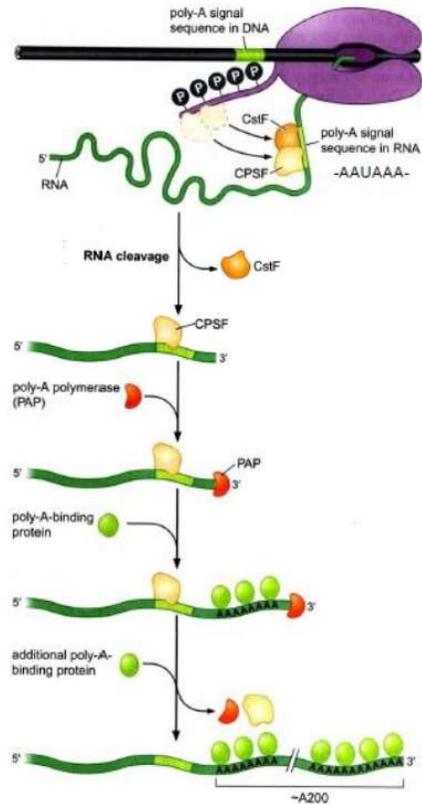
## Rôles de la coiffe :

- Protège l'ARNm de la dégradation
- Favorise l'export de cet ARNm vers le cytosol
- Dans le cytosol, favorise la traduction de l'ARNm

# MATURATION DES ARNm

La queue poly A est ajoutée en deux temps (clivage/polyadénylation) :

- 1) Le pré-ARNm est clivé (coupé) au niveau du site de clivage.
- 2) La polyA polymérase est recrutée et synthétise une queue polyA (environ 200 nt)

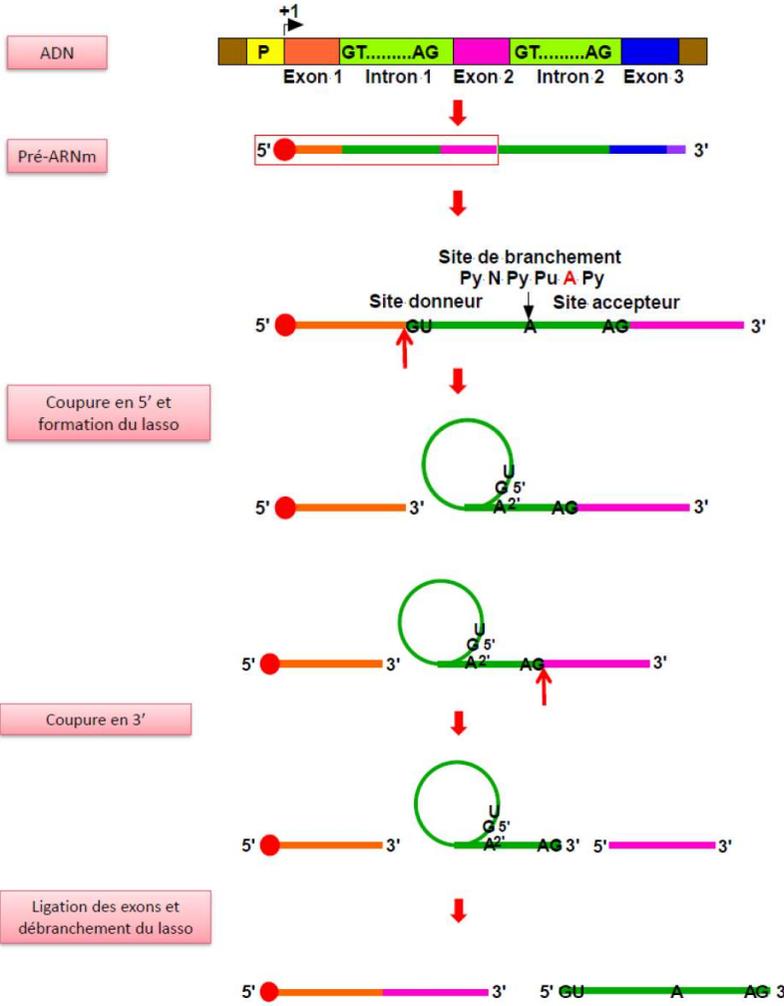


Queue polyA synthétisée par la PAP

CF : facteur de clivage  
CSPF : facteur de clivage et de spécificité de la polyadénylation  
PAP : poly A polymérase  
PABP : protéine de liaison à la queue poly A  
CstF : facteur de stimulation pour le clivage  
CS : site de clivage  
DSE : élément de séquence en aval

Pré-ARNm synthétisé par l'ARN pol II

# MATURATION DES ARNm



Epissage des introns par le spliceosome :

Site donneur d'épissage en amont

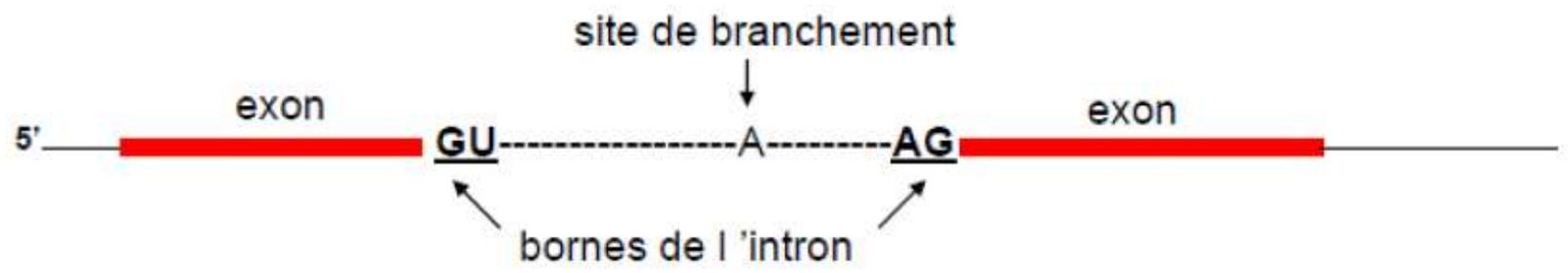
Site de branchement à l'intérieur de l'intron

Site accepteur d'épissage en aval

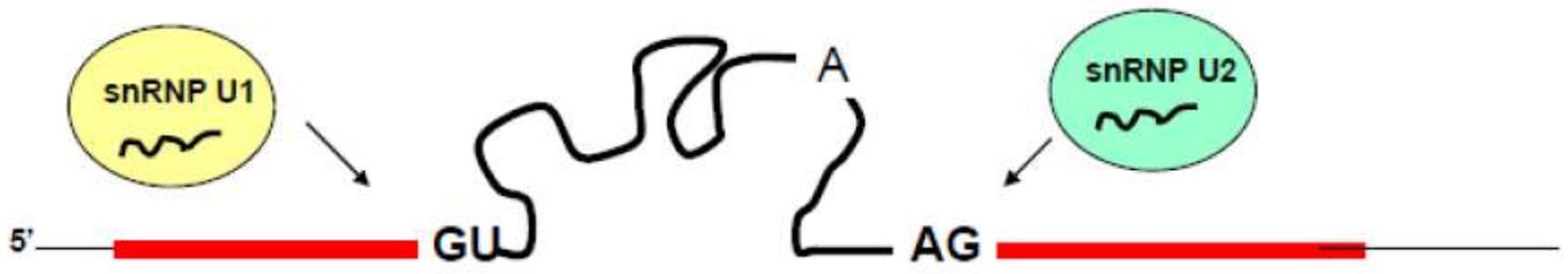
L'épissage est réalisé par le **spliceosome** constitué de protéines associées à des snRNA : **les snRNP** (U1, U2, U4, U5, U6).

1 snRNP = 1 snRNA + 10-20 polypeptides

# MATURATION DES ARNm

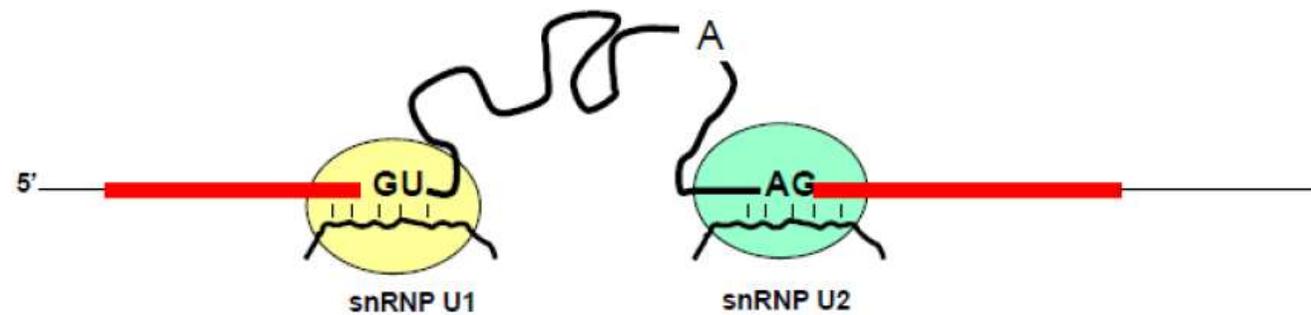


- Reconnaissance des bornes donneur et accepteur

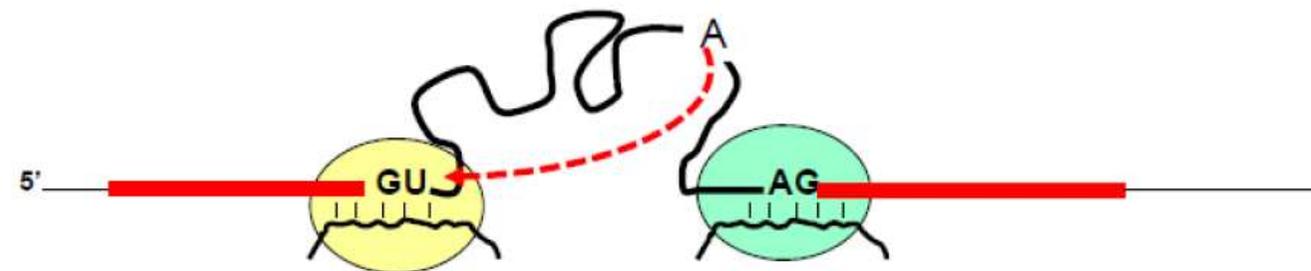


# MATURATION DES ARNm

- Fixation des snRNP U1 et U2 sur les bornes

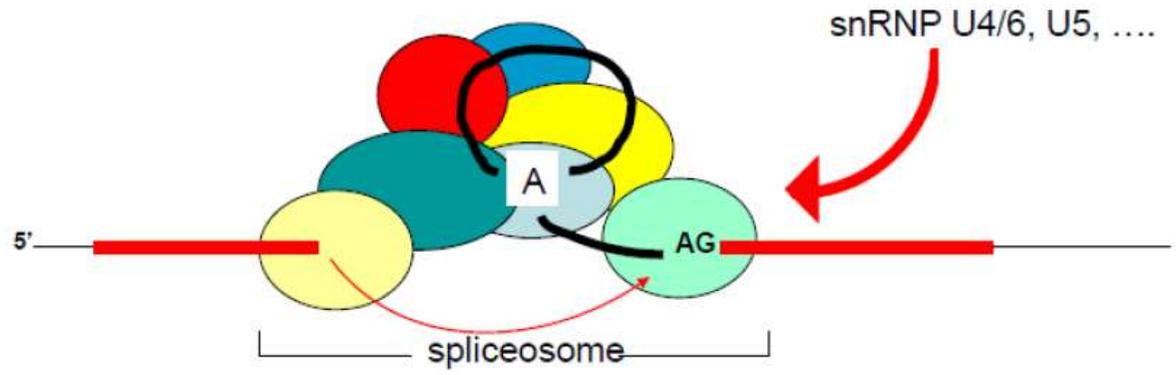


- 1ère réaction de trans-estérification

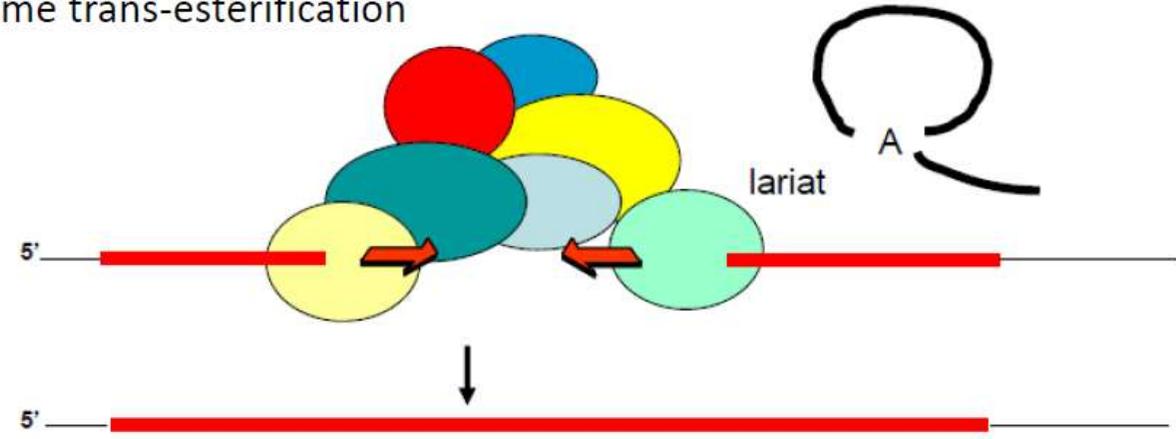


# MATURATION DES ARNm

- Recrutement du spliceosome



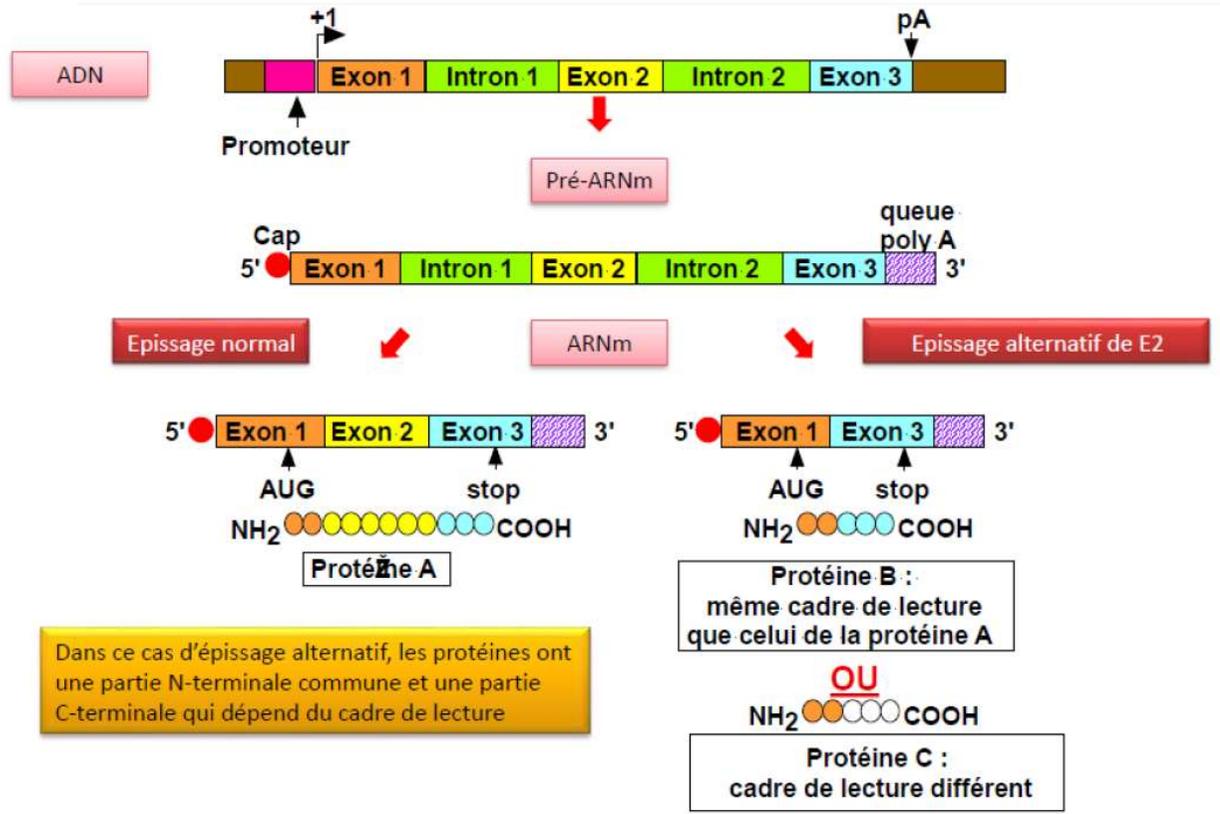
- 2ème trans-estérification



# UN GENE, PLUSIEURS PROTEINES

Pour un certain nombre de gènes, il existe des mécanismes qui permettent de produire des ARN messagers matures différents à partir d'un même gène de départ. Un certain nombre de ces mécanismes sont transcriptionnels ou post-transcriptionnels.

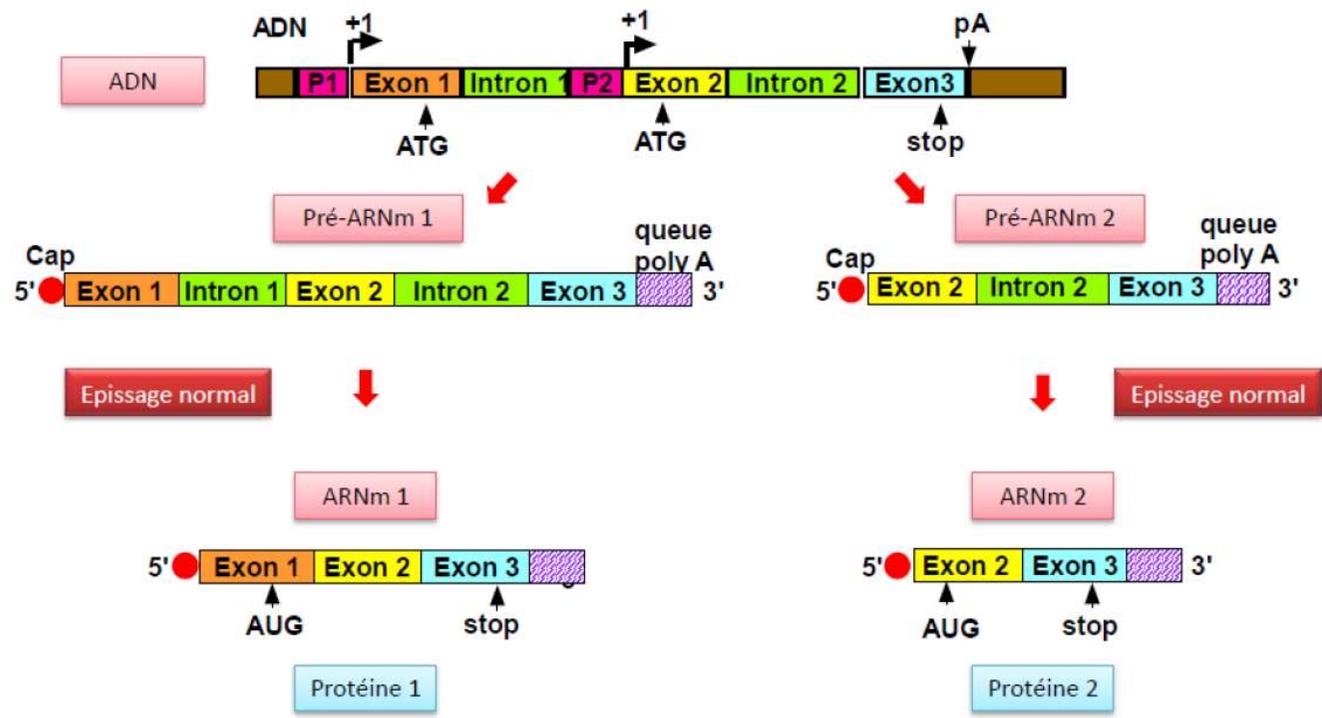
Ex. L'épissage alternatif



# UN GENE, PLUSIEURS PROTEINES

Pour un certain nombre de gènes, il existe des mécanismes qui permettent de produire des ARN messagers matures différents à partir d'un même gène de départ. Un certain nombre de ces mécanismes sont transcriptionnels ou post-transcriptionnels.

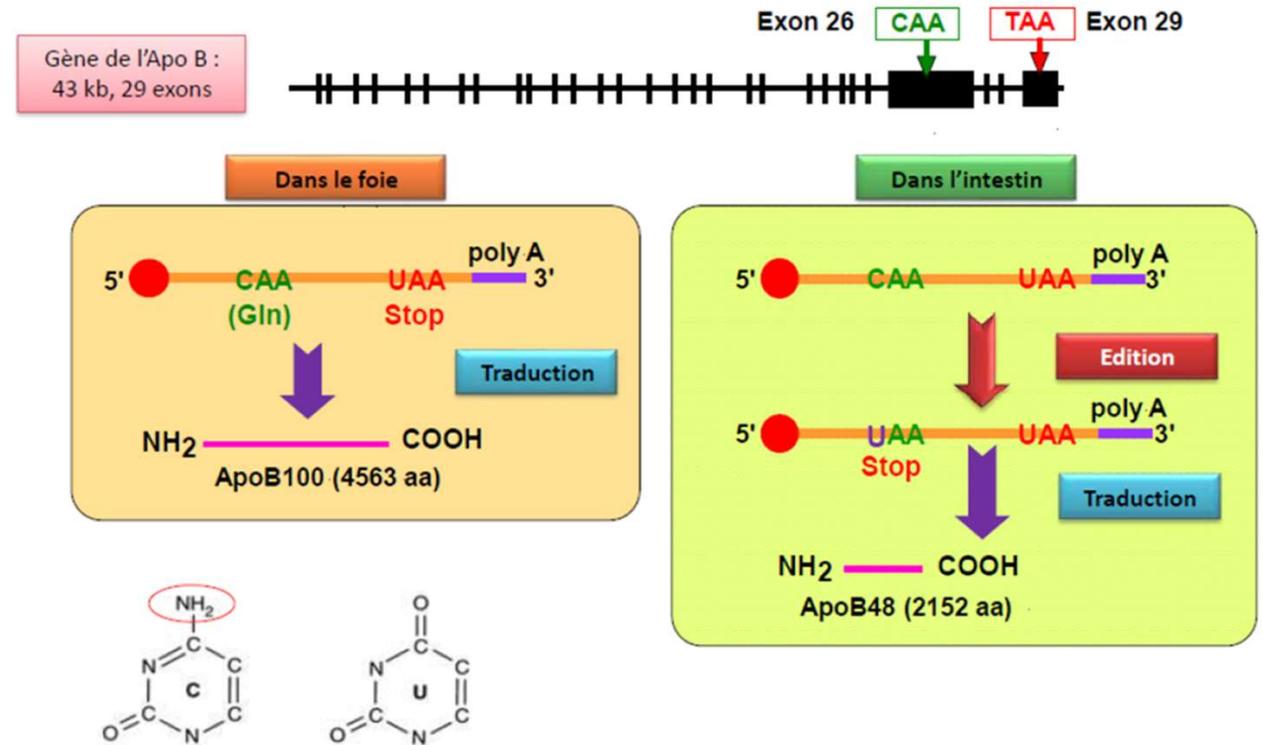
Ex. L'utilisation de promoteurs alternatifs



# UN GENE, PLUSIEURS PROTEINES

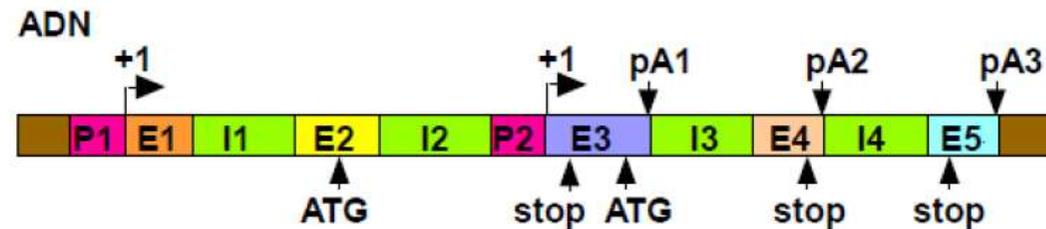
Pour un certain nombre de gènes, il existe des mécanismes qui permettent de produire des ARN messagers matures différents à partir d'un même gène de départ. Un certain nombre de ces mécanismes sont transcriptionnels ou post-transcriptionnels.

Ex. L'édition d'ARNm  
(la séquence est « réécrite » après la transcription)



# UN GENE, PLUSIEURS PROTEINES

Les différents mécanismes peuvent être combinés

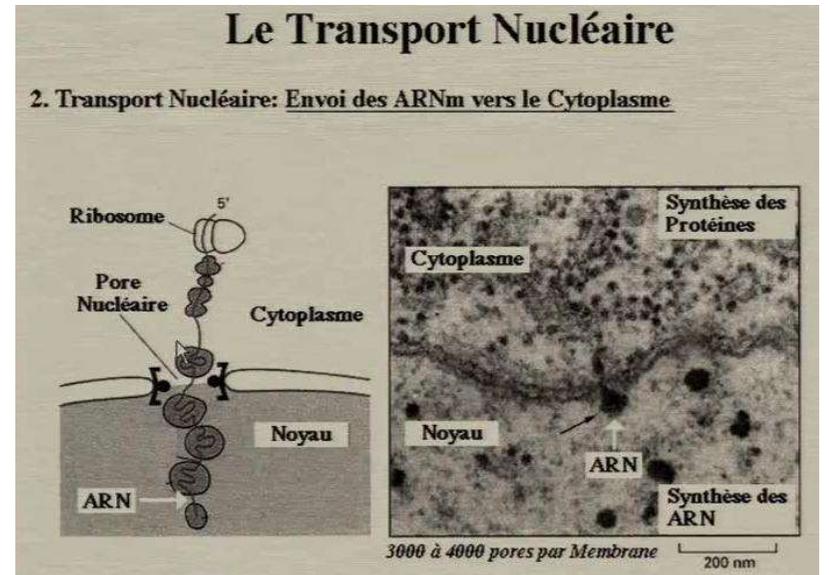
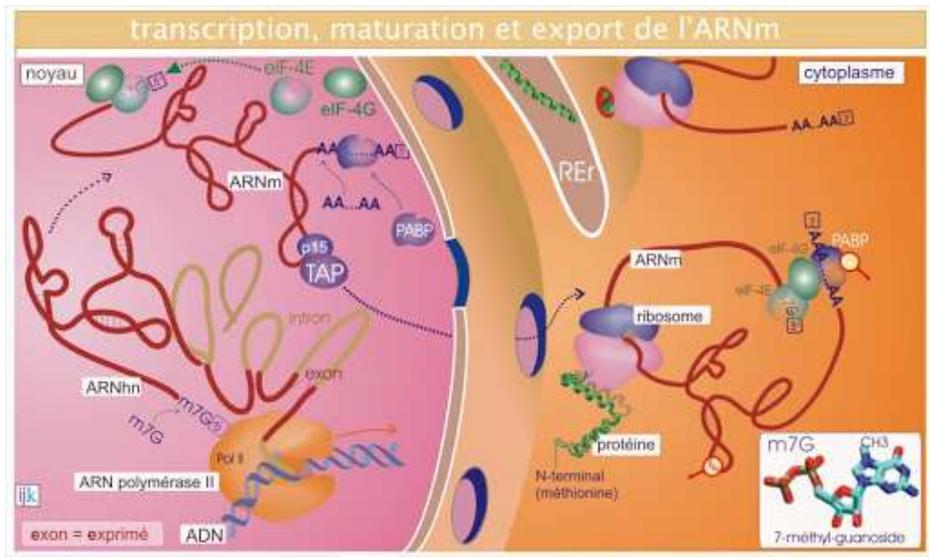


## Conclusion

- ✓ Selon sa régulation, un gène peut donner plusieurs protéines dans un ou plusieurs types cellulaires
- ✓ Un ARNm ne comprend que des exons
- ✓ Le 1<sup>er</sup> exon est toujours précédé d'un promoteur
- ✓ Le dernier exon est déterminé par le choix du site de poly-adénylation
- ✓ Les exons peuvent être codants ou non codants

# L'EXPORT DES ARNm VERS LE CYTOSOL

Les ARNm matures sont exportés dans le cytosol



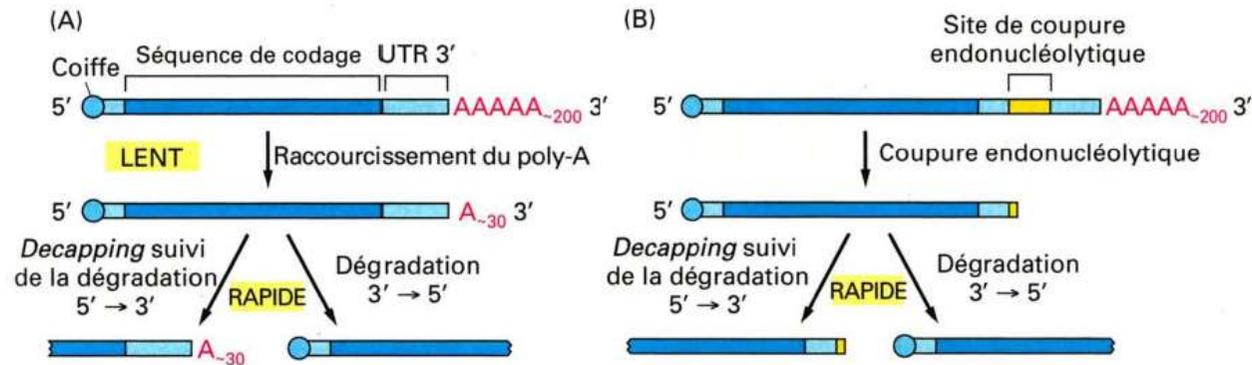
Mécanisme d'export des ARNm vers le cytosol via les pores nucléaires

# DEGRADATION DES ARNm EUCARYOTES

Les ARNm permettent la synthèse des protéines via le mécanisme de **traduction**. Dès que la cellule, a produit suffisamment de protéines pour assurer son métabolisme, elle va dégrader les ARNm.

Selon les besoins cellulaires, les ARNm ont donc une durée de vie limitée, variant de **quelques minutes à quelques heures**.

Il existe deux mécanismes de dégradation :



NB: Ces mécanismes permettent également de dégrader les ARNm non sens (ayant un codon stop prématuré)

# CONCLUSION

# COMPARAISON PROCARYOTES / EUCARYOTES

	PROCARYOTES	EUCARYOTES
<b>Lieu</b>	Transcription et traduction dans le cytoplasme : la traduction peut donc commencer avant la fin de la transcription	Transcription et maturation des ARNm dans le noyau Traduction dans le cytoplasme → transcription et traduction non couplées
<b>Maturation des ARNm</b>	Pas de coiffe	Coiffe en 5'
	Pas de queue poly A	Coiffe en 3'
	Pas d'épissage (pas d'intron)	Epissage (introns)

# COMPARAISON PROCARYOTES / EUCARYOTES

