

TRANSCRIPTION CHEZ LES EUKARYOTES

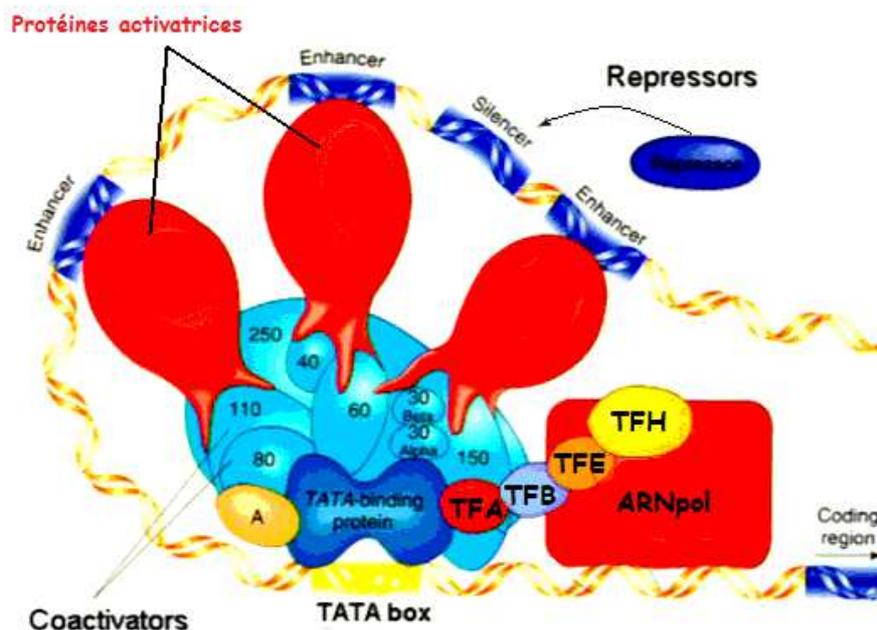
I) Introduction

A) Vue générale.

Contrairement aux procaryotes l'ARNpol **ne se fixe pas directement sur le gène**
→ ARNpol est **recrutée** avec de nombreuses protéines dont **TATA box**.

Pour que la transcription se déroule, il faut

- **Des facteurs transcriptionnels (TF)**
- **Des séquences très en amont du site d'initiation** qui se fixent (**enhancer**) ou qui ne se fixent pas (**silencer**) à des **protéines (activateurs ou répresseur)** qui forment un pont entre le lieu de la transcription et les séquences activatrices.



Essentiel : L'ADN n'est pas une molécule linéaire, elle a des formes rondes. Il génère des courbures qui mettent en contact des sites initialement éloignés. Ca permet la régulation des gènes.

Il y a trois types d'ARN polymérase.

- La I n'est pas sensible à l'amanitine.
- La II est inhibée par l'amanitine.
- La III n'est sensible qu'à forte dose.

B) Catégories majeures d'ARN chez les êtres humains contribuant à l'expression des gènes

Les gènes qui codent pour les protéines sont transcrits par l'ARN polymérase II.

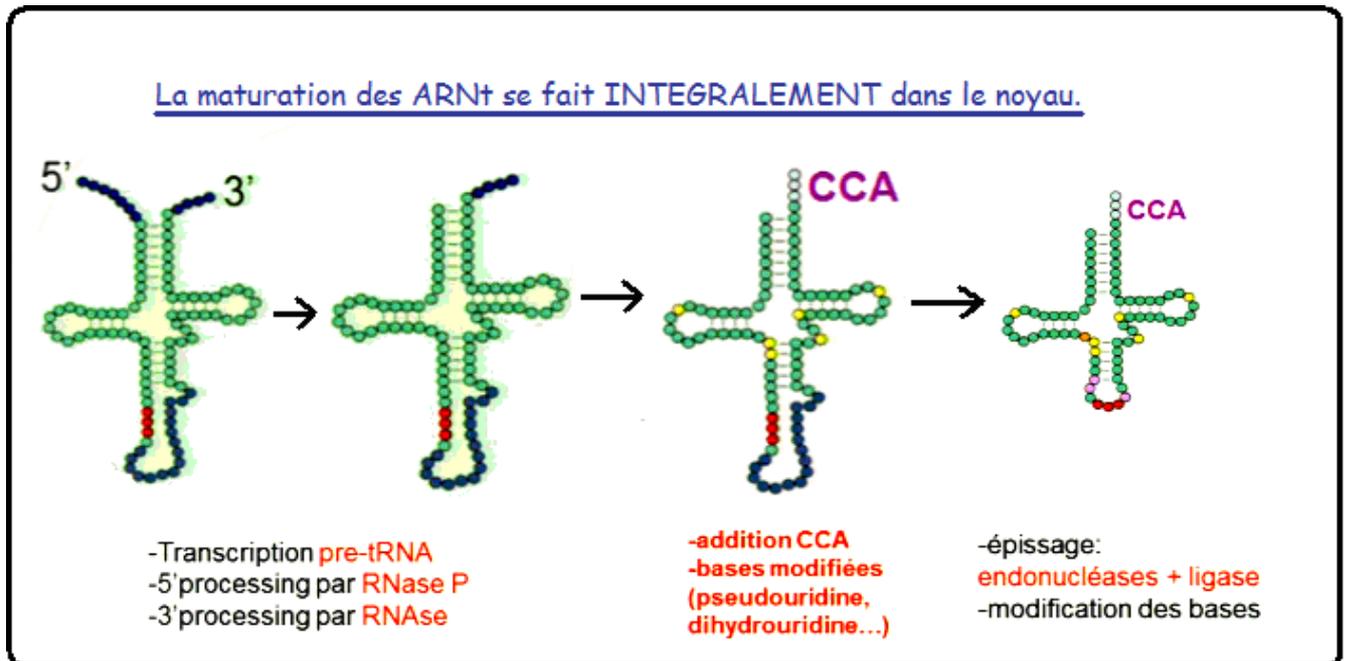
ARN	fonction	Types	Quantité	Stabilité	ARN pol
Ribosomal (ARNr)	-ribosome	28S, 18S 5.8S, 5S	80%	stable	I sauf 5S=III
petits ARN nucléaires snRNA= [U1,U2, U4, U5], U6...	Épissage des pré-ARNm	30	<1%	stable	II ± III
Transfert (ARNt)	Adaptateur entre ARNm et ribosomes	> 40	15%	stable	III
snRNA	Maturation des ARNr	>100	< 1%	stable	II ± III
Messagers (ARNm)	Codent les protéines	10 ⁵	2-5%	stabilité variable	II

Tous ces ARN sont des ARN non codants.

L'ARN ce n'est pas qu'un support pour la synthèse des protéines.
Certains ont des fonctions catalytiques, d'autres sont messagers

On va les passer en revue (bon courage, et bisous !)

1. La maturation des ARN de transfert, transcrits par ARNpol III.



- L'ajout de CCA se fait grâce à une *nucléotidine transférase*.
- La modification des bases a un rôle fonctionnel
 - o exemple : isomérisation d'uridine en pseudo-uridine :
La pseudo-uridine peut engager deux liaisons hydrogène alors que l'uridine n'en engage qu'une seule → La pseudouridine stabilise des ARNt
- **Épissage.**
 - Pour l'ARNt, l'intron est retiré
 - o sans l'intervention d'un splicéosome
 - o sans l'intervention des snRNA
 - o mais grâce à des **endonucléases** qui coupent les liaisons phosphodiester, et retirent les introns. Resoudure des extrémités de l'ARN grâce à des **ligases**.

La protéine « La » : Pour les ARNt elle est excisée.

Au départ : attachée à l'extrémité 3', mais elle est clivée lors de la maturation.

Rôle : Facilite la formation des ribonucléoprotéides. Assure une protection contre les nucléases,

Elle envoie un signal de séquestration dans le noyau tant que les ARNt ne sont pas matures.

Protéine LA = auto-antigène dans le lupus érythémateux et le syndrome de Gougerot-Sjörgen.

2. Les gènes des snoRNA

• Répartis en 2 classes

- Classe 1 : gènes standard avec une région codante et un promoteur en amont
- Classe 2 : majorité. Localisés **dans les introns** qui codent les protéines ribosomales.

Un ribosome est constitué de protéines et d'ARN ribosomales ce qui fait que pour que les snoRNA soient synthétisés, il faut qu'au préalable les protéines ribosomales le soient.

→ La mission essentielle snoRNA est d'assurer la maturation des ARN ribosomales.

C'est un système qui équilibre la production des ARN et des P ribosomales, et qui optimise la fabrication des ribosomes.

Les gènes des snoRNA peuvent être en copie unique ou en clusters (=bcp de gènes au même endroit et très proches)

• Exemple : Cluster SNURF.

→ Regroupe notamment les gènes qui codent les snoRNA.

En principe, chaque gène a un allèle paternel, et un allèle maternel.

Les deux sont fonctionnels, c'est-à-dire que la transcription peut se faire à partir de l'un ou de l'autre.

Mais certains n'expriment que l'allèle paternel ou maternel : c'est ce qu'on appelle *l'empreinte génomique*. C'est le cas des gènes du cluster SNURF.

**Dans le cluster SNURF
seuls les allèles paternels des gènes des snoRNA sont exprimés.**

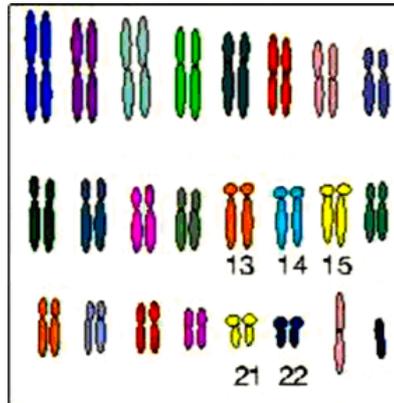
En cas de dysrégulation de cette empreinte génomique

→ syndromes malformatifs : *Prader Willi / Angelman*.

Ces gènes qui codent les snoRNA ont des séquences conservées appelées boîtes (= identiques d'une espèce à l'autre, par exemple les pharmanes ont les mêmes séquences que nous.)

- **snoRNA HBOX ACA :**
→ pseudo-uridylation
- **snoRNA CD BOX :**
→ Méthylation sur le carbone 2' du ribose qui possède un groupement OH

3. Transcription des ADN ribosomaux.



Les ADN ribosomaux sont regroupés en clusters sur les chromosomes acrocentriques (13, 14, 15, 21, et 22)

Notion de transcrit primaire polygénique.

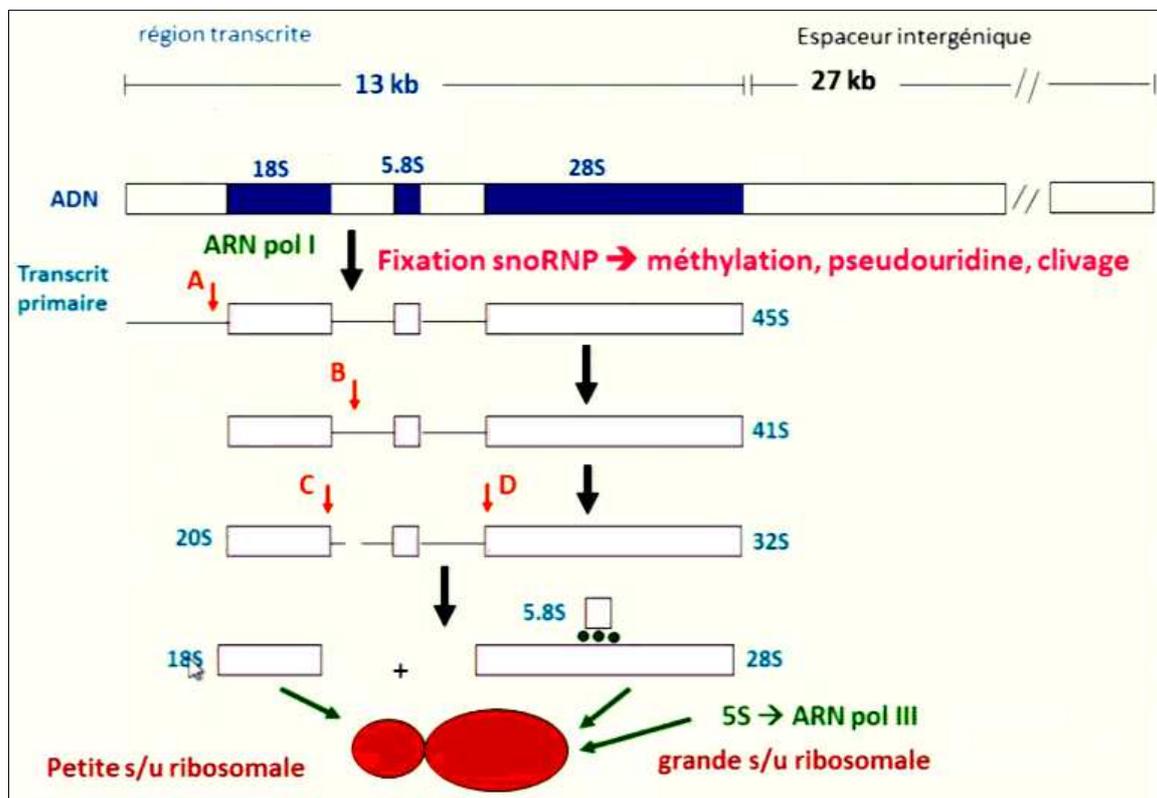
D'abord on a 1 Unité de transcription = 40kb

- une région transcrite de 13kb
- un espaceur intergénique de 27kb
- répétées en tandem 30 à 40x → rythme de transcription soutenu.

Action de l'ARNpol I : donne un *transcrit primaire*,

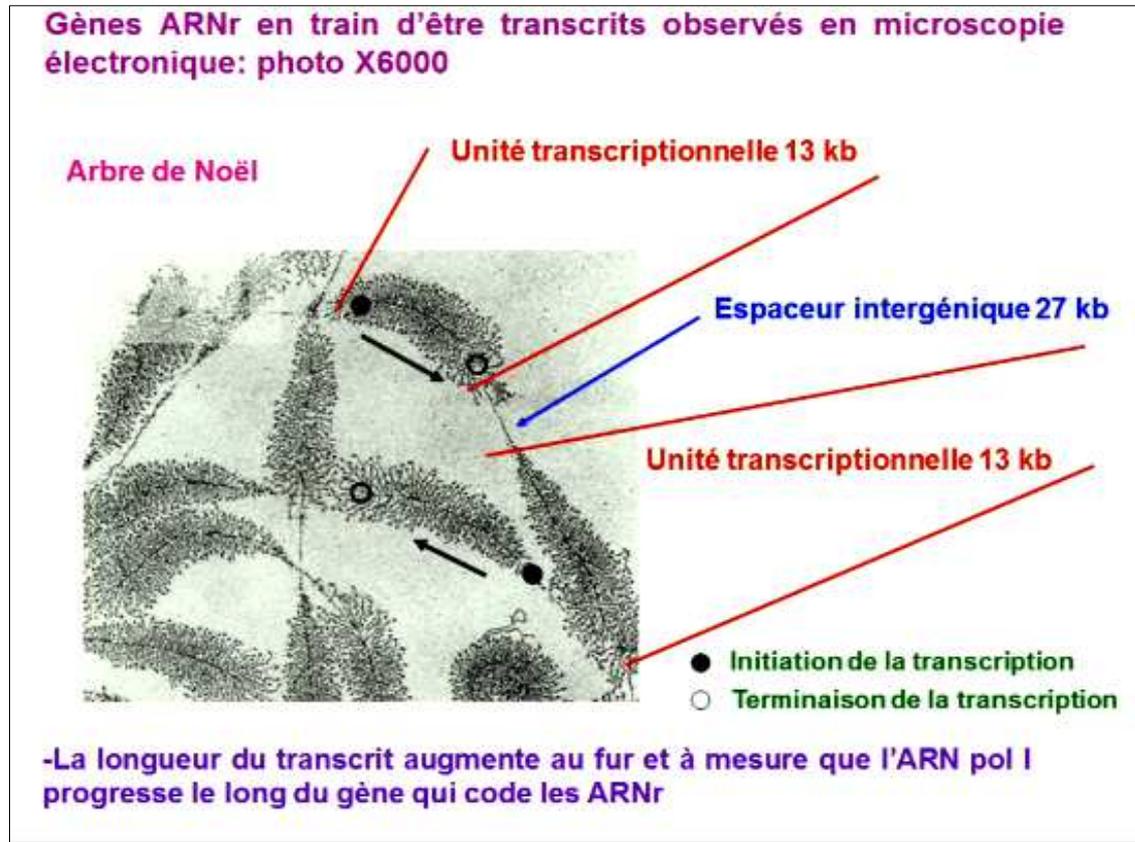
Action des *snoRNP* : modification des bases et fixation aux points de clivage

Action des *endonucléases A, B, C, et D* : clivage



18s : petite sous unité ribosomale

28s 5s et 5.8s : grande sous unité ribosomale



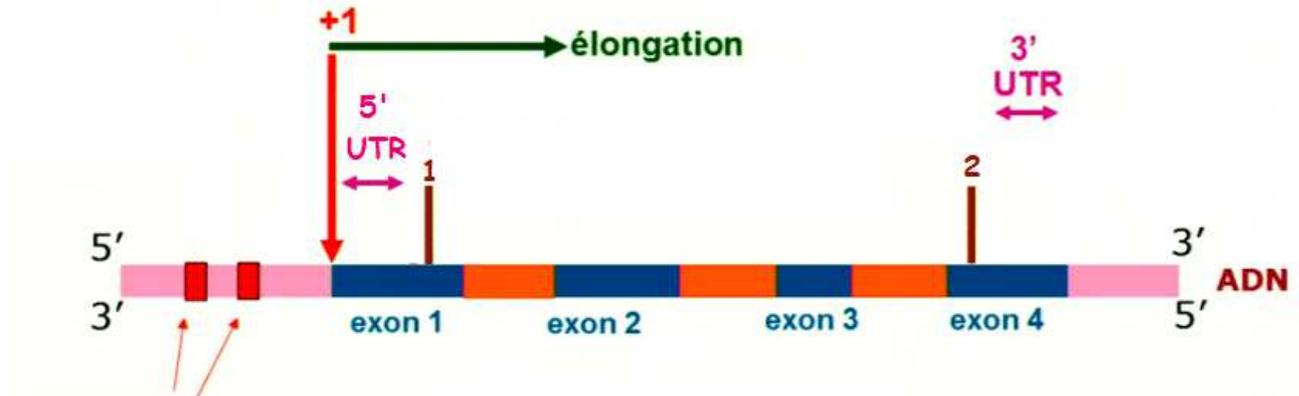
Les filaments perpendiculaires à la ligne sombre ce sont les transcrits

Le côté où les filaments sont courts : initiation de la transcription

Le côté où les filaments sont longs : terminaison de la transcription.

L'extrémité libre est ici 5'.

II) La transcription a trois étapes chez les eucaryotes.



éléments cis :

- promoteur (initiation)
- régulateurs

+1 La première base transcrite =
Site de début de la transcription

5'UTR : région non traduite

1 : triplet de début de la traduction

3' UTR : région non traduite

2 : triplet de terminaison de la traduction

Éléments cis : initiation et régulation. Donnent les instructions à l'ARNpol

Régions codantes : exons

ARNpol se fixe sur les régions promotrices, et commence à transcrire.

Le promoteur n'est pas copié dans l'ARNm.

A) Etape 1 : l'initiation

1) Les séquences de type 1 :

Notion de fidélité = JE TRANSCRIS TOUJOURS AU MEME ENDROIT.

Elles indiquent où commence la transcription = et c'est toujours au même endroit → fidélité.

Plusieurs motifs peuvent composer le promoteur :

Boîte TATA reconnue par facteur transcriptionnel TBP

- **T** (82% des cas) **A** (97%) **T** (93%) **A** (85%)
- En amont en position **-30**.
- Provoque une asymétrie dans le gène qui signale le début de la transcription.

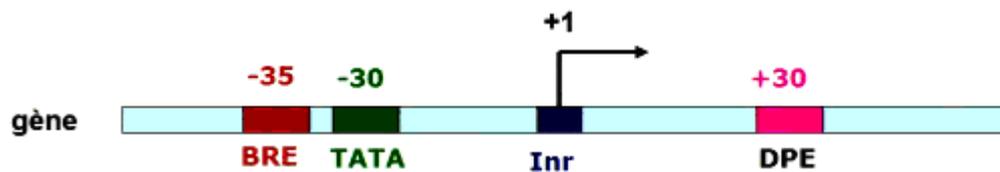
Et aussi :

Inr : encadre site de début de transcription

BRE : en amont à -35

DPE : en aval à + 30

Ces séquences n'existent pas impérativement toutes les quatre dans tous les gènes. Un peut suffire, le plus souvent : 2



séquence :

TATA 5' T₈₂ A₉₇ T₉₃ A₈₅ A₆₃ A₈₈ A₅₀ 3'

reconnue par facteur transcriptionnel :

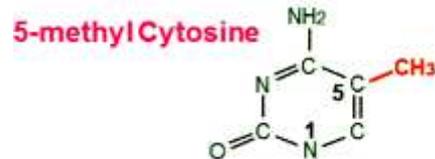
TBP

Propriété du promoteur : règles pour que ça fonctionne.

- Il peut n'y avoir qu'une séquence pour initier la transcription.
- Il faut qu'elle soit reconnue par une protéine sur un ADN double brin.
- Il faut qu'elle soit sur le brin codant (orienté 5' → 3')
- Il faut un respect strict des séquences et des distances.
- Interactions avec des protéines régulatrices qui recrutent l'ARN polymérase.

TATA n'est pas indispensable : exemple des gènes domestiques.

- Ils sont ubiquitaires
- Codent pour des enzymes de la glycolyse
- Taux de transcription faible et continu
- **Ne possèdent ni boîte TATA ni même de séquence initiatrice.**
- Riches en doublets CG : rôle dans l'expression des gènes.
 - o Doublets **hypométhylés** : activation
 - o Doublets **hyperméthylés** : inhibition



Les doublets CG sont rassemblés en **ilots HTF** (500 à 2000 pb) principalement *dans la région du promoteur ou du premier exon*.

Hpa II est une enzyme de restriction méthylation sensitive :

→ Elle clive les doublets CG **uniquement si les cytosines ne sont pas méthylées**.

Les ilots CPG sont des séquences régulatrices sont reconnus par des facteurs transcriptionnels :

- Facteur Sp1 se fixe sur boîte GC : GGGGCGGG
- Sp1 active la transcription

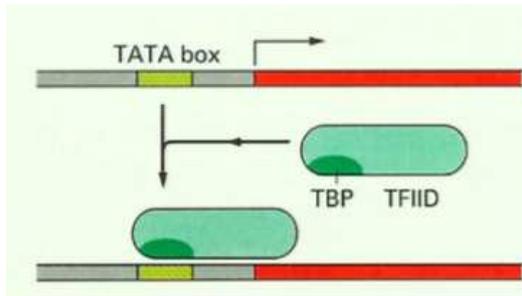
Chez les gènes domestiques, les ilots CPG sont des séquences régulatrices qui recrutent des facteurs transcriptionnels modulant l'expression des gènes.

Exceptions :

- Certains gènes domestiques possèdent TATA
- Les caractéristiques des gènes domestiques peuvent être rencontrées dans d'autres gènes.

Comment la boîte TATA initie la transcription ?

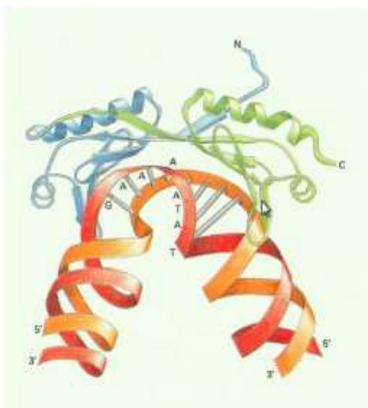
Le 1^{er} évènement est la reconnaissance de la boîte TATA par le TBP qui fait partie du facteur TFIID :



Les résidus phénylalanine Hphobe de TBP introduisent une flexibilité dans la double hélice au niveau de TATA en élargissant le petit sillon.

A cet endroit là, l'ADN a tendance à être relâché. La courbure est le signal de début de transcription.

Rappel : TATA => fidélité de la transcription



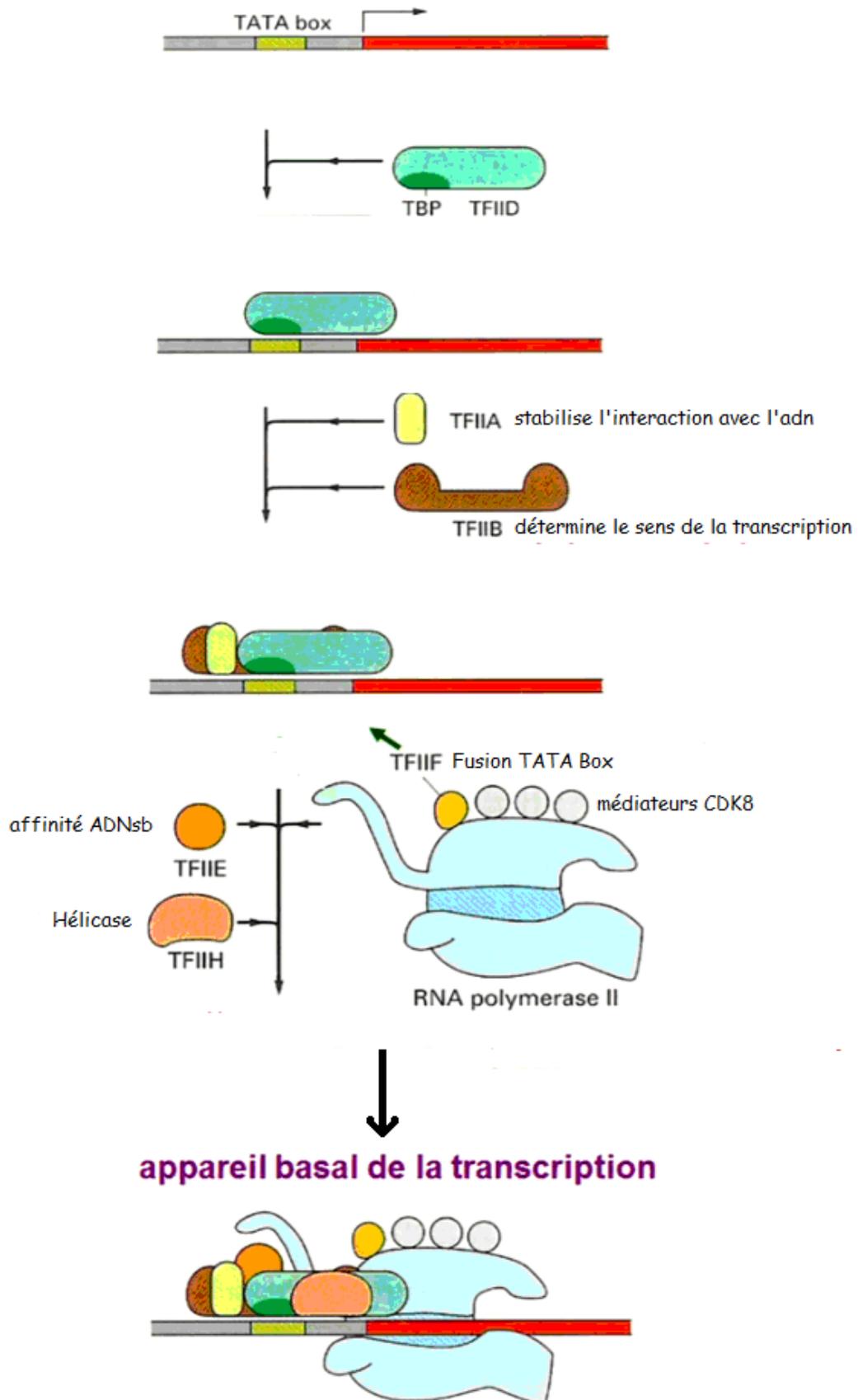
On dit que TBP est posé sur l'ADN comme la selle sur un cheval



2^{ème} évènement : TFIID recrute des protéines :

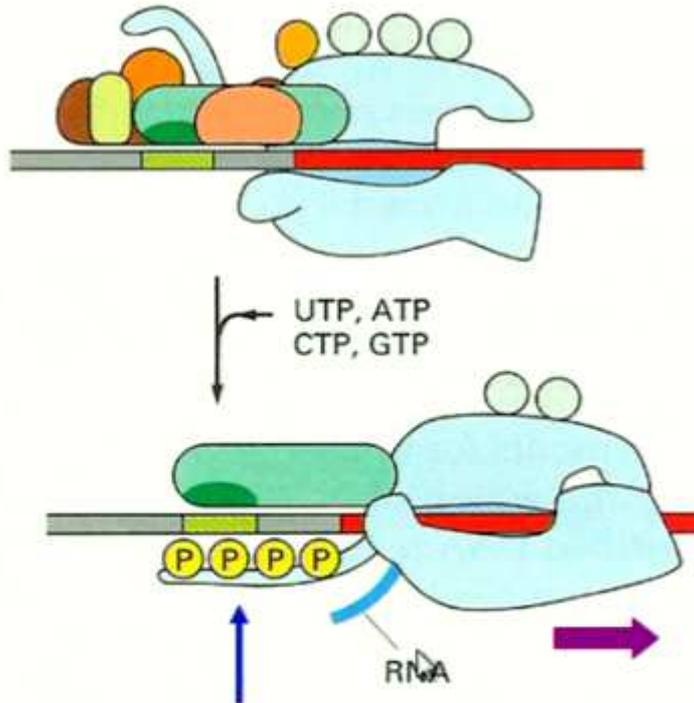
- **TFIIA** stabilise les interactions entre protéines régulatrices et ADN
- **TFIIB** (monomérique) : interagit en amont et en aval de la boîte TATA
 - o en amont avec le grand sillon
 - o en aval avec le petit sillon
 - o accentue l'asymétrie et indique le sens de la transcription.
- **L'ARN pol II** ramène son bouclier avec ses copines les protéines.
- **TFIIF** : génère des torsions dans l'ADN : favorise la fusion, facilite la configuration en simple brin.
- **TFIIE** : très grande affinité pour ADN simple brin.
- **TFIIH** : sépare l'ADN et le fait passer en simple brin car il possède une activité hélicase ATP dépendante.

Tout ça pour → **favoriser l'apparition de la fenêtre simple brin.**



Mais ça ne suffit pas ! Quand y'en a plus, y'en a encore !

ARN pol doit pouvoir s'échapper, pour passer dans la phase **d'élongation**. Comment ça marche ?



Complexe [cdk7-cycline H] associé à TFIIF
→ phosphoryle C-Terminal Domain
= 52 x (tyr-ser-pro-thr-ser-pro-ser)

La **CDK7** (ou 9) phosphoryle **TFIIF**, qui acquiert une fonction kinase,
TFIIF phosphoryle à son tour deux résidus **sérine** l'extrémité C-TER de l'ARNpol II.

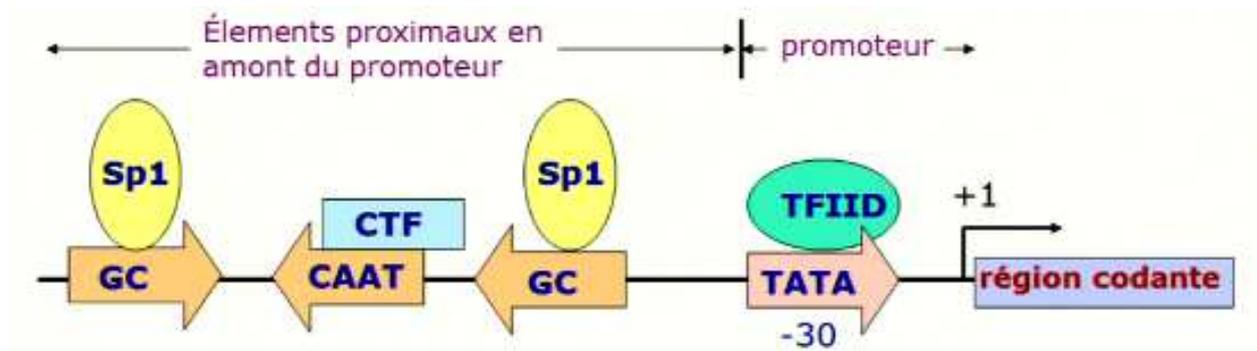
Le motif tyr-**ser**-pro-thr-**ser**-pro-ser est répété 52x
Donc :

En tout 104 charges négatives seront rajoutées à l'extrémité C-TER
→ signal de la transcription.

La phosphorylation a lieu pendant l'initiation et l'élongation

2) Les éléments proximaux du promoteur : séquences de type 2.
Notion de fréquence : COMBIEN DE FOIS JE TRANSCRIS ?

Nombreuses séquences : boîtes CAAT, GC, CACC (spécifique des gènes de la β -Globine)



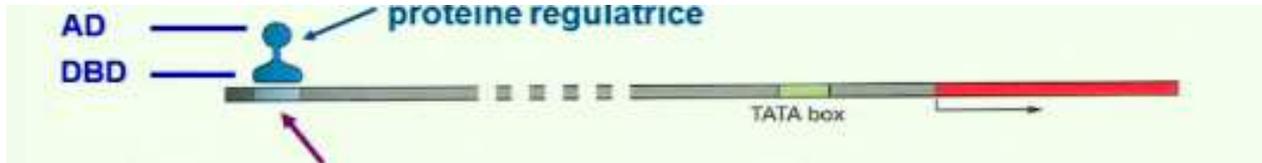
Les boîtes CAAT et GC :

- Fonction indépendante de l'orientation
- Pas d'impératif de distance.
- Recrutent les facteurs Trans

3) Promoteur = éléments de type 3 A QUELLE VITESSE JE TRANSCRIS ?

Promoteur distal activateur = enhancer.

Promoteur distal inhibiteur = repressor



Propriétés des enhancer (activation) et des repressors (inhibition)

- Fonctionne à **distance** du site d'initiation de la transcription
- Fonctionne **en amont ou en aval** du site d'initiation de la transcription
- Fonctionne **dans les 2 orientations**
- Fonctionne **en recrutant une pour plusieurs protéines (facteurs trans)**
- Fonctionne en facilitant la fixation du **complexe ARNpol** sur le promoteur.
- Fonctionne **sans impératif de séquence** très sévère
- Fonctionne sur des promoteurs **hétérologues**.
 - o Si en amont on met un enhancer viral, il sera efficace c'est-à-dire pourra activer la synthèse de la β -globine humaine.



- Sont soumis à des signaux de contrôle négatifs.
 - o Si on les met à un certain endroit du génome ils perdent leur caractéristique, ils sont inhibés.

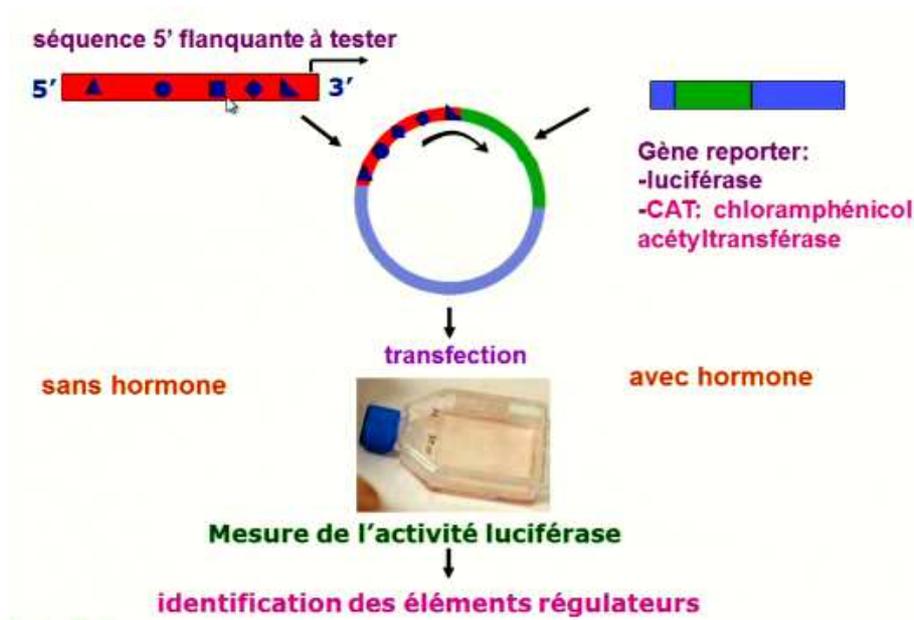
4) Utilisation d'un gène reporter pour identifier les éléments cis de régulation de l'ADN : exemple des HRE (Hormone Reponse Element)

Exemple de la luciférase.

Supposons qu'on analyse un segment, avec de nombreuses séquences. On veut savoir si ces séquences ont un rôle dans l'expression des gènes (ou pas)

On va le mettre en amont d'un gène dont il est possible de mesurer l'activité.

Il ne faut pas que la protéine soit exprimée dans la cellule de l'étude pour ne pas fausser les résultats en mesurant l'activité endogène.



On utilise le gène de la luciférase comme gène reporter.

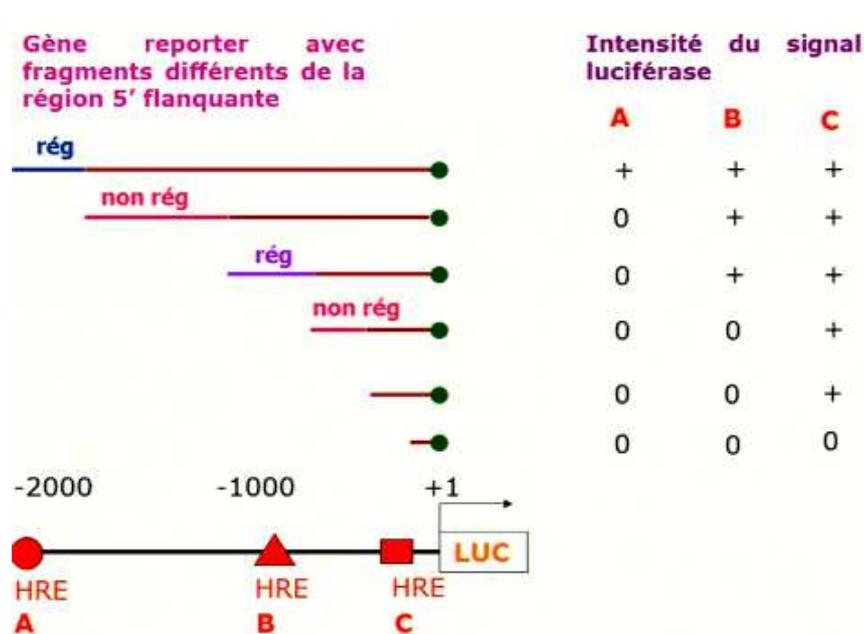
La luciférase n'est pas exprimée dans les cellules des mammifères. Repérable facilement car elle envoie une fluorescence quand elle sera activée, que l'on captera avec un luminomètre.

Si la séquence est active, la luciférase va être activée et briller.

Transfection du plasmide obtenu dans des cellules avec hormone / sans hormone

Puis mesure de l'activité de l'activité luciférase.

On coupe successivement la séquence en petits morceaux et on procède par tâtonnements pour identifier les régions régulatrices.



Petit à petit quand on retire les fragments, en comparant avec l'activité, on peut savoir quels sont les fragments qui interviennent dans la régulation du gène.

B) Etape 2 : l'élongation.

Mais visiblement on n'a rien dit là dessus, donc on va juste considérer que l'ARNpol se déplace sur l'ADN et fait son boulot c'est à dire qu'il transcrit l'ADN pour donner de l'ARN. C'est déjà pas mal. Donc on passe à l'étape 3 :

C) Etape 3 : La terminaison

Chez les eucaryotes, les signaux de terminaison de la transcription ont des mécanismes encore mal élucidés)

La terminaison de la transcription ne peut être effectuée que si la maturation de l'ARNm à l'extrémité 3' est finie :

- **1** : retrait du dernier intron
- **2** : clivage en aval du site de polyadénylation
- **3** : polyadénylation

Les mécanismes impliqués sont multiples :

- La terminaison se produit en aval de l'endroit où le transcrit sera polyadénylé
- Des séquences spécifiques induisent une pause de la transcription
- Décrochage spontané de l'ARNm

Variabilité des séquences de terminaison :

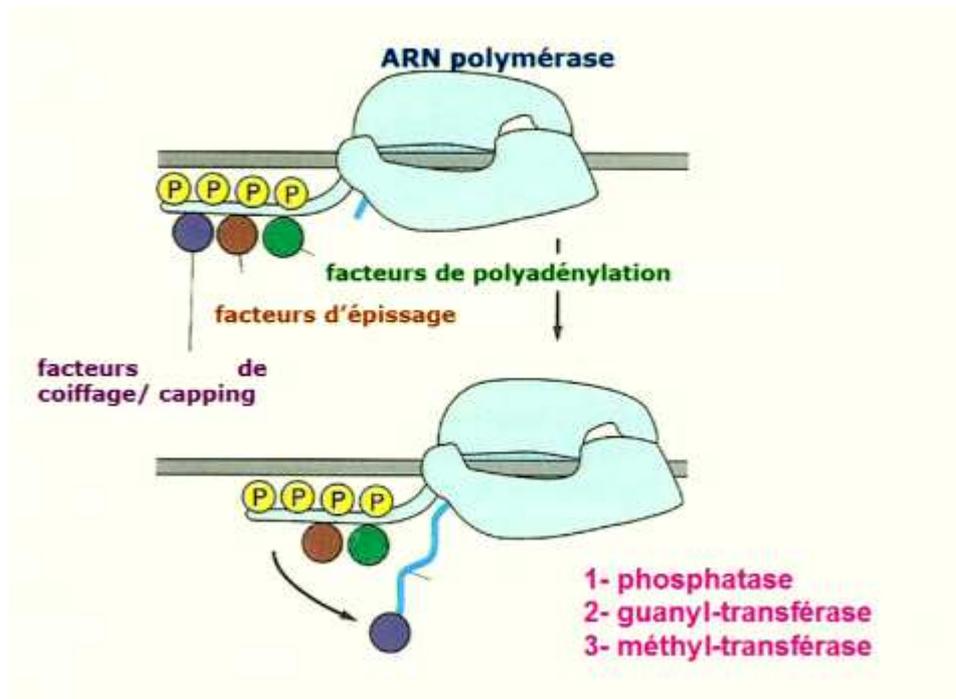
- Pour certains gènes : à proximité en aval de l'extrémité 3' du messager mature
- Pour d'autres gènes : 1600 pb en aval de l'extrémité 3'
→ chaînes β et ϵ de la globine

Caractéristiques :

- L'arrêt ne s'effectue pas au même endroit pour tous les ARNm produits à partir d'un même gène.
- Implication de l'extrémité C-TER domain de l'ARNpol et des protéines fixées.

1) 5' Capping

On n'attend pas la fin de la transcription pour commencer la maturation avec les protéines en rose. Dès que le transcrit naissant a une longueur de **25 nucléotides** les protéines qui vont intervenir dans la maturation de l'extr 5' sont transférées sur l'ARN naissant pour qu'elles puissent effectuer l'addition de la coiffe en 5'

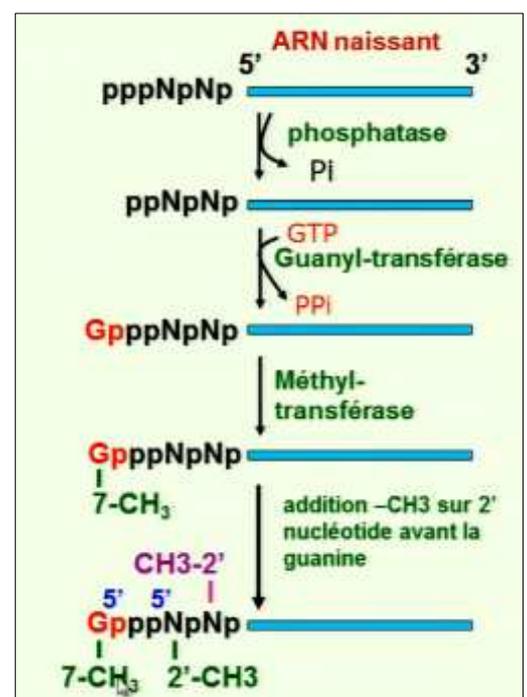


La coiffe en 5' des ARNm synthétisée par l'ARN pol II.



- 1 : une phosphatase retire un phosphate en 5'
 - 2 : une guanyl transférane rajoute une guanosine en 5'
 - 3 : une méthyl transférane rajoute un CH₃ sur l'azote de la 7-guanine.
 - 4 : Le nucléotide après la guanine est lui aussi méthylé : en 2'
- Le nucléotide encore après peut également être méthylé.

Et tout ça forme la coiffe →



La coiffe joue un rôle primordial car elle est le point d'attache pour un complexe protéique qu'on appelle le CBC

- stabilisé des ARN
- export des ARN dans le cytoplasme
- indispensable pour la traduction.
- Une théorie veut que la transcription s'arrête si y'a pas de coiffe.

Variante coiffe:

2,2,7 tri-methyl-guanosine

On la retrouve :

- sur les snARN transcrits par ARNpol II
- sur la moitié des snoARN

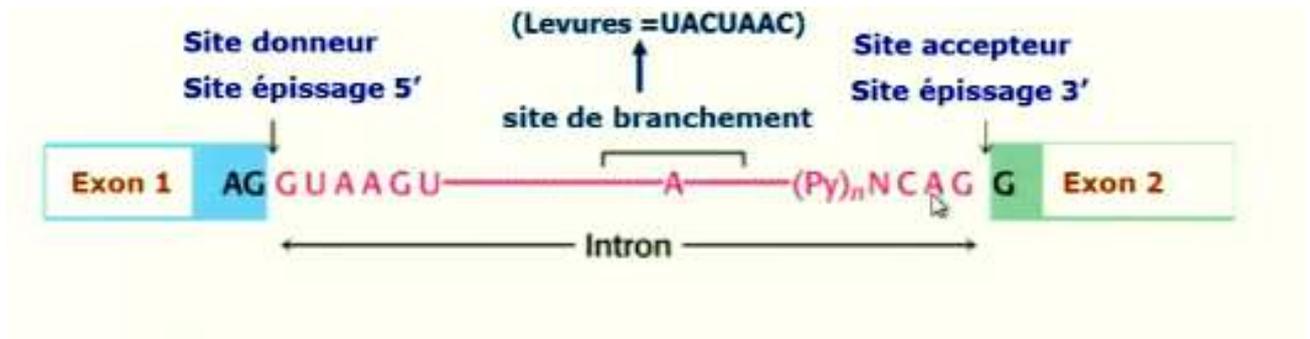
Il n'y a de coiffe sur les ARN transcrits par l'ARN pol III

Donc pas de coiffe sur les ARNt, ARNr, et ARNm

2) L'épissage (retrait des introns) est réalisé par un splicéosome.
= assemblage de snRNP et de protéines.

Se déroule dans le noyau et fait appel à des séquences consensus dont la présence est indispensable.

a. Les différents sites



Il y a un site d'épissage en 5' (donneur) et un site d'épissage en 3' (accepteur)

Les séquences flanquantes

Et un site de branchement à l'intérieur de l'intron.

Nucléotides du point de fixation

Deux nucléotides nécessaires à l'épissage et qui sont les points de fixation du splicéosome.

Nucléotide GU dans le site d'épissage en 5'

Nucléotide AG dans le site d'épissage en 3'

Ca c'est le splicéosome majeur. (Il existe aussi un splicéosome mineur mais OSEF)

La présence de ces nucléotides est nécessaire mais pas suffisante.

Séquences flanquantes

Ce qui est important **c'est les séquences flanquantes.** (voir schéma)

Site de branchement.

Chez les levures : séquence constante.

Chez les mammifères : séquence variable, **mais** il y a toujours l'adénine libre (rôle dans le déclenchement de l'épissage).

Distance site de branchement / extrémité 3' : varie entre 18 et 40 nucléotides,

Distance site de branchement / extrémité 5' : peut aller jusqu'à plusieurs milliers.

Taille moyenne d'un exon : 150 / 200 pb

Taille moyenne d'un intron : 5000 à 10000 pb

ESE et ESS:

Exonic Splice Enhancer/Silencer.

Envoie un signal à la cellule pour indiquer où se trouvent les introns et les exons, et donc s'il faut les retirer ou non.

En principe ces séquences sont dans l'exon qui est en aval de l'intron qui doit être retiré

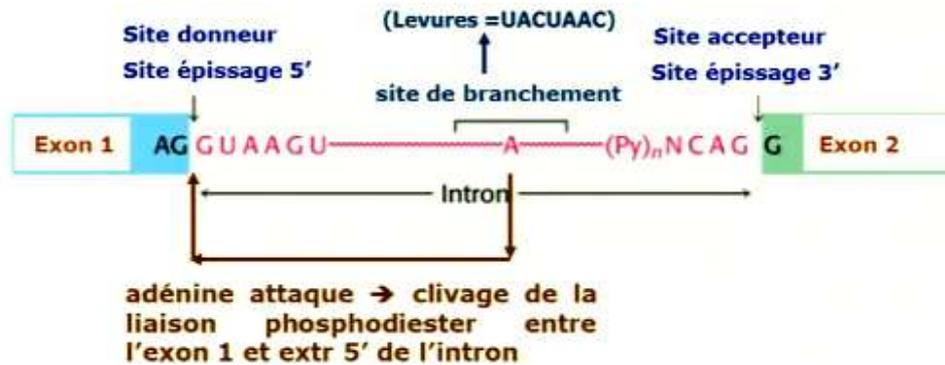
Caractéristiques :

- entre 8 et 10 nucléotides
- 10 séquences connues dans les exons humains
- Reconnue par les protéines Sr (dont l'extrémité C-TER est riche en sérine et Arginine) et les HnRNP.
 - o Ce sont les protéines Sr qui indiquent à la cellule où se trouvent les exons.
 - o HnRNP se fixent sur les introns et participent à leur condensation.
Elles camouflent aussi des sites des introns qui présentent une analogie avec les séquences flanquantes.

b. Les réactions chimiques

1^{ère} estérification 2' → 5'

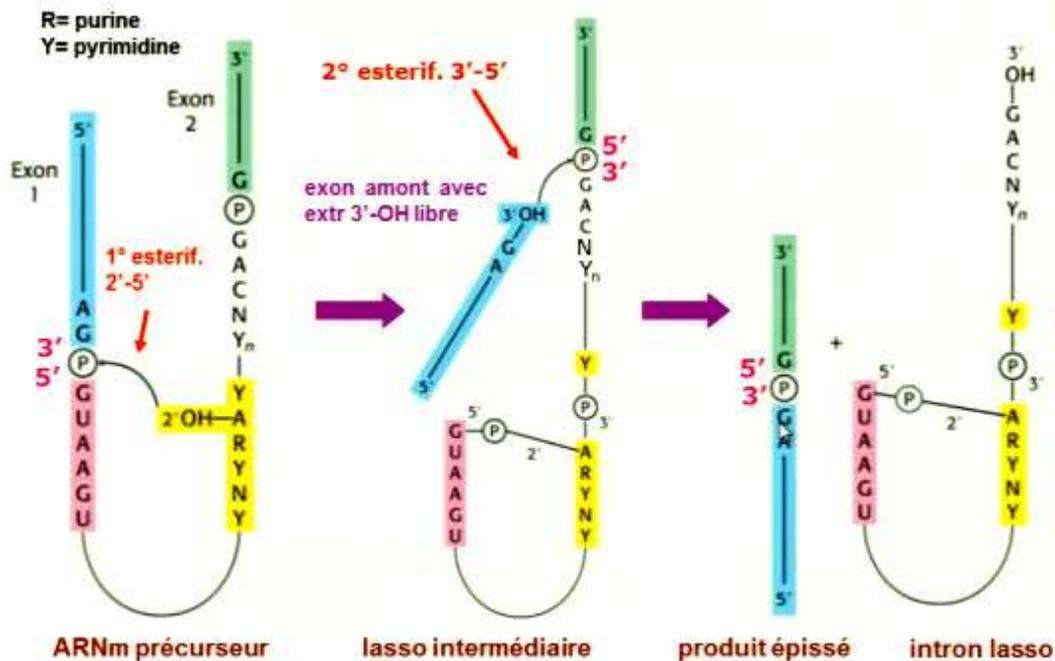
L'adénine du site de branchement rompt la liaison phosphodiester entre l'exon 1 et l'extrémité 5' de l'intron.



C'est une réaction de trans estérification. Elle forme ainsi une liaison phosphodiester 2' → 5'. On se retrouve avec une conformation en lasso intermédiaire.

2^{ème} estérification 3' → 5'

A son tour, l'exon libre attaque la liaison phosphodiester au niveau de l'exon 2.



Ce qu'on obtient au final : Un produit épissé + un intron en lasso.

Les réactions de transestérification ne consomment pas d'énergie. Par contre le positionnement des snRNP sur les sites d'épissage consomme de l'énergie.

c. Le splicéosome.

L'épissage des introns est réalisé par les snRNP.

Ils sont constitués de

- snRNA synthétisée par l'ARNpol II : possède une coiffe tri-méthylée en 5' → sa structure secondaire va changer au fur et à mesure de l'épissage. Cette modification va permettre à des snRNA différents de s'hybrider.

Elles vont s'associer à des protéines :

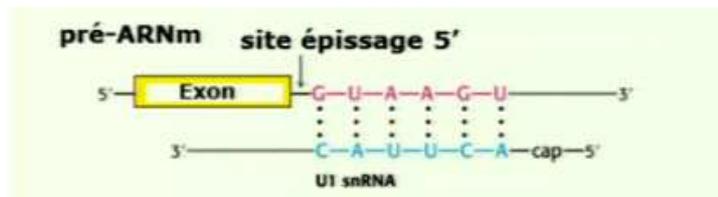
- protéines communes:
 - o **Protéines Sm.**
Fixées sur les snRNA grâce aux protéines SMN.
Les snRNA subissent une maturation qui se déroule à la fois dans le noyau et dans le cytoplasme. Les snRNA effectuent une navette entre noyau et cytoplasme.
Mutation des SMN → myotrophie spinale.
Maladies auto-immunes → anticorps anti-sm
- protéines spécifiques à chaque snRNP.

Il faut que le splicéosome s'assemble (snRNP + autres protéines + facteurs d'épissage + ARN en cours d'épissage)

1^{ère} étape : snRNP U1 reconnaît le site d'épissage en 5'.

La concentration relative de deux protéines ASF/SF2 et HNRNPA1 guide le choix du site d'épissage en 5'

Une fois que la cellule a choisi le site d'épissage, U1 va venir s'hybrider avec le site d'épissage en 5' (liaisons hydrogène)



Si la cellule se trompe de site → conséquences délétaires mais pas toujours.

2^{ème} étape : U2 se fixe sur le site de branchement.

Il est aidé par des protéines

- **U2AF** reconnaît la succession de pyrimidines en 3' de l'intron
- **SF3A / SF3B** : activation de U2. Ne se fixent sur U2 que si les riboses sont méthylées.
- **BBP + ATP** : fixation sur le site de branchement
-

3^{ème} étape :

- **U4/U6 et U5** rejoignent **U1/U2**.

Ce qui est primordial : la cellule doit savoir très exactement où se trouvent les exons

Processus de vérification des sites d'épissages par plusieurs snRNP.

- U1 et U5 reconnaissent le site d'épissage en 5' et le site d'épissage en 3'
- U6 reconnaît le site d'épissage en 5'
- U6 expulse U4 et U1 et interagit avec 5' à la place de U1 et de U2

Des hélicases ATP dépendantes permettent à U6 et à U2 de s'apparier en un complexe unique lié par des liaisons hydrogène.

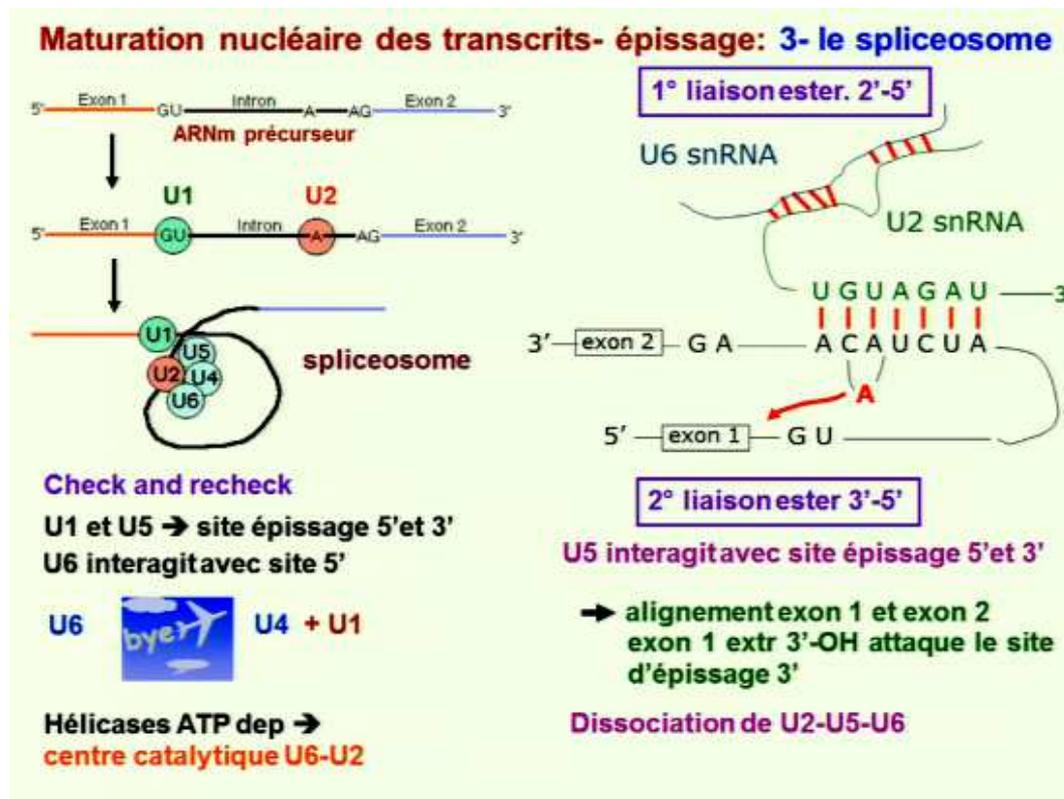
U6-U2 est le complexe catalytique de l'épissage

Il est à l'origine de la première estérification.

U2 se fixe sur le site de branchement mais laisse une base libre : l'adénine du site de branchement qui reste disponible pour déclencher l'attaque sur la liaison phosphodiester.

U5 provoque l'alignement exon 1 et exon 2 à proximité. L'extrémité 3'OH libre de l'exon1 va rompre la liaison phosphodiester entre l'intron et l'exon 2

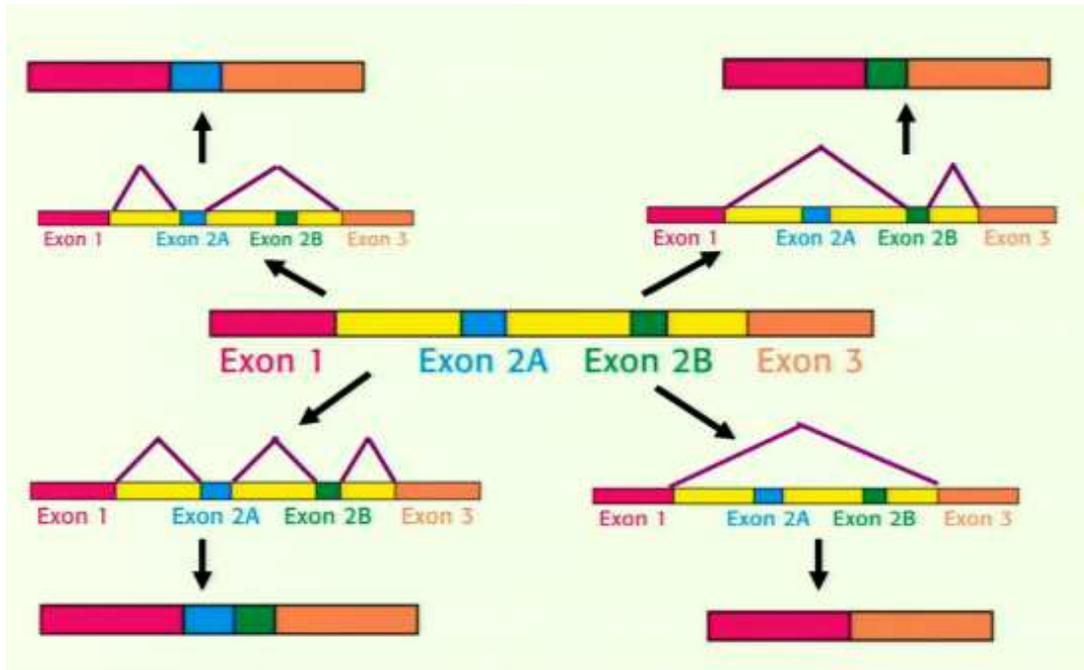
- Obtention du produit épissé
- Dégradation de l'intron
- Dissociation de U2 U5 et U6
- Redémarrage du cycle sur un autre intron.



d. l'épissage alternatif

Dans un tissu où le ratio ASF/SF2 HF-RNPA est différent il est possible que le choix des sites d'épissages soit différent.

On a dit tout à l'heure que le choix d'un autre site d'épissage pouvait engendrer des effets délétères. En fait il est possible que ce choix soit volontaire et qu'il donne lieu à des transcrits matures différents, codant pour des protéines tout à fait fonctionnelles.



On estime que 30% des gènes humains peuvent subir un épissage alternatif.

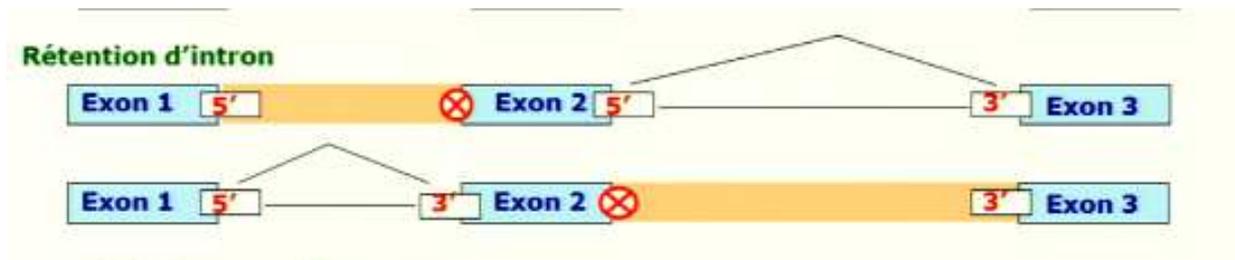
L'épissage différentiel quand il donne lieu à des protéines fonctionnels est bénéfique puisqu'il participe à la diversité génétique.

e. les erreurs d'épissage.



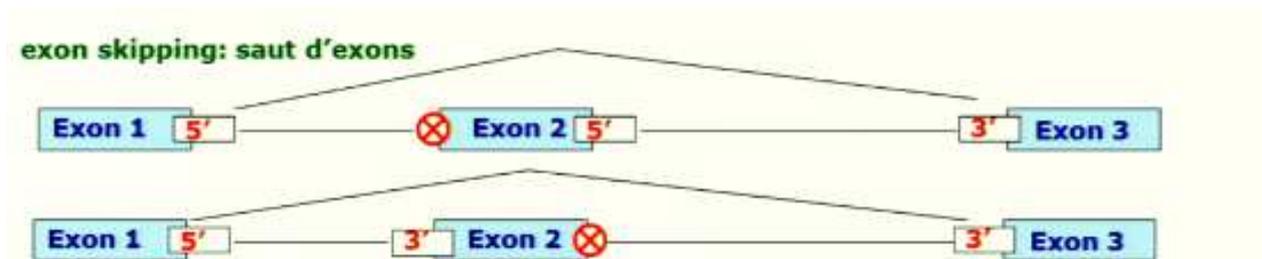
- Rétention d'intron

Si on a une mutation dans un site d'épissage, un intron peut être conservé et considéré comme une séquence codante.



- Saut d'exon (Exon skipping)

Si on a une mutation dans un site d'épissage, il peut y avoir un saut d'exon. L'exon est zappé et du coup ne code pas au moment de la traduction.



- activation des sites cryptiques

Dans l'intron on a une séquence flanquante avec un dinucléotide AG qui fait croire qu'il est le vrai site d'épissage alors que pas du tout.

- Si le leurre est dans un intron : exon rallongé
- Si le leurre est dans un exon : exon raccourci.



3) La polyadénylation en 3'

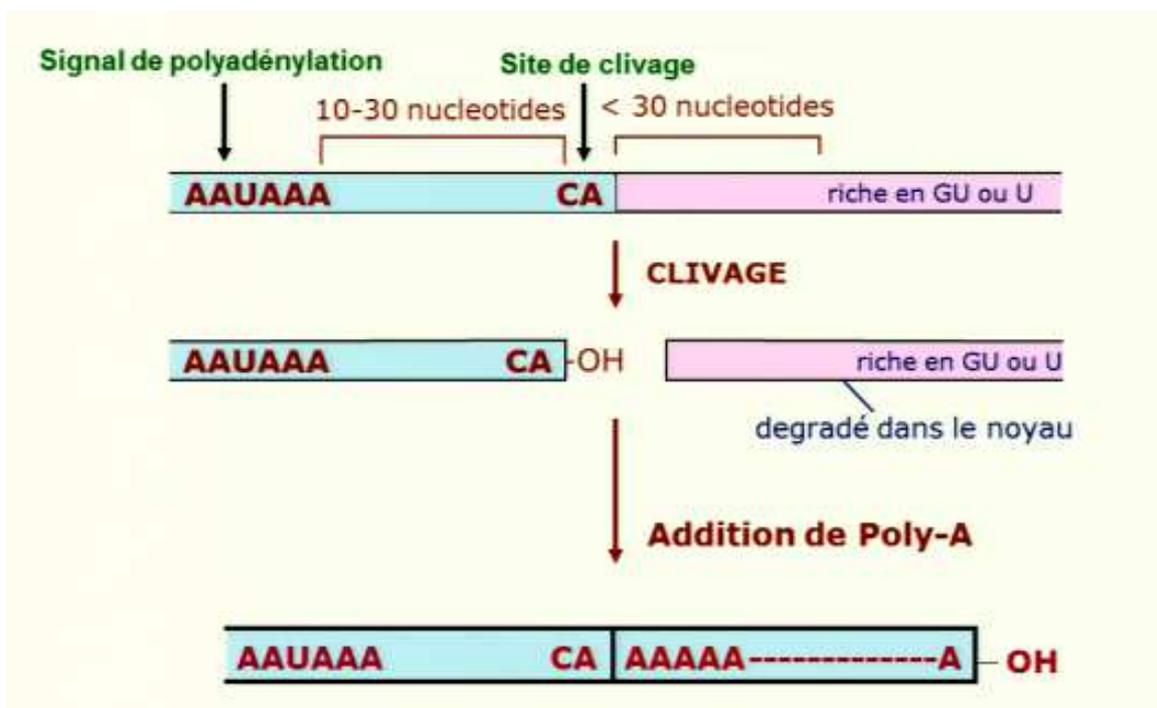
ARNpol ramène des enzymes qui sont transférées sur le transcrit.

Il y a un **signal de polyadénylation AAUAAA** indiquant le lieu de la maturation.

10-30 nucléotides en aval → le **dinucléotide CA** indique le site de clivage

Encore en aval de CA → il y a une **région très riche en G ou en U** qui va être dégradée dans le noyau.

Enfin, une enzyme va rajouter des adénosines qu'on appelle la **queue poly-A**



L'extrémité 3' est maturée grâce à 6 protéines multimériques.

A un moment donné l'ARNpol atteint le signal de polyadénylation.

- **CPSF** : reconnaît AAUAAA
- **CSTF** : reconnaît GU

Ces protéines sont transférées sur le transcrit.

Deux facteurs de clivage **CF1** et **CF2** coupent l'ARN et le CSTF se dissocie.

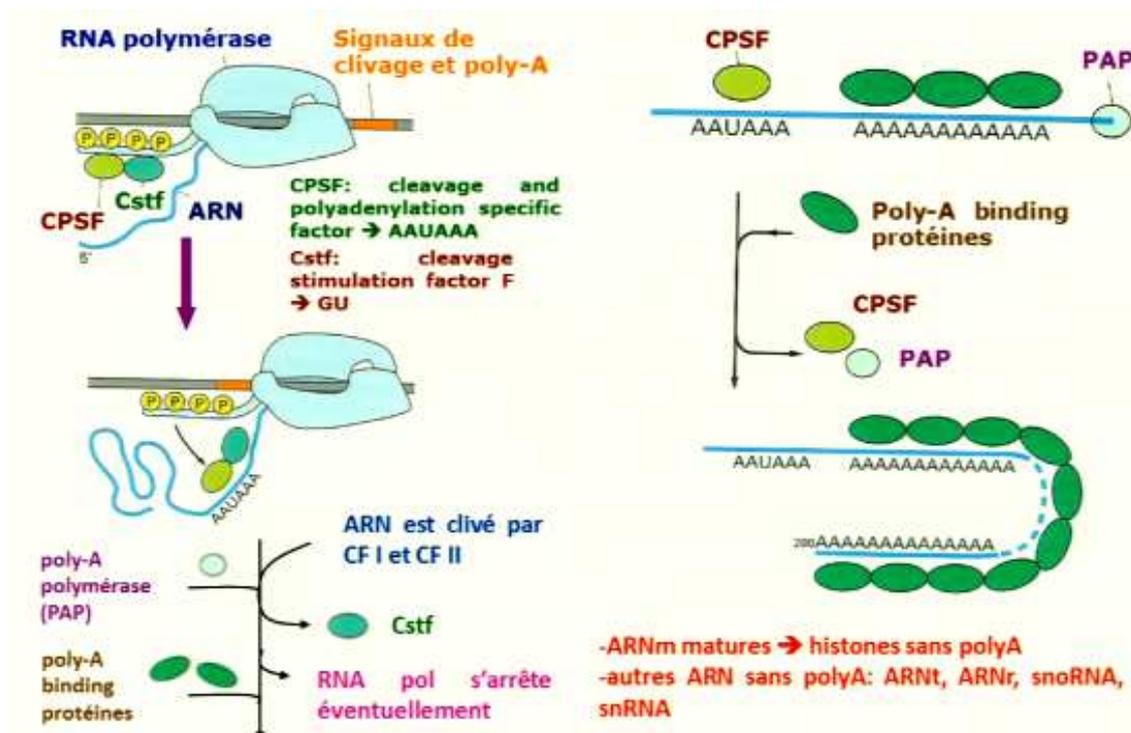
L'ARN messenger se trouve à l'état libre : ca ne veut pas dire que l'ARNpol va s'arrêter.
Elle peut continuer à transcrire mais inefficacement.

L'ARN doit être mûré

Encore deux autres protéines :

- **Poly-A polymérase (PAP)** rajoute les adénosines. Elle n'a pas besoin d'une matrice, son substrat est l'ATP, elle rajoute entre 200 et 250 adénosines.
- **Poly-binding-protein** reconnaît les adénosines.

Au bout d'un certain temps, la PAP s'arrête, le CPSF se dissocie, et voilà de quoi a l'air l'extrémité 3' polyadénylée avec la succession des adénosines.



Certains ARN n'ont pas de polyA :

- les histones,
- les ARNt,
- les ARN ribosomiaux, 7
- les snoRNA
- et les snRNA