

Université de  
Biskra

Département  
des sciences  
de la nature  
et de la vie

# Expression et régulation des gènes des eucaryotes

M1 Biochimie appliquée

Dr. TRABSA HAYAT

---

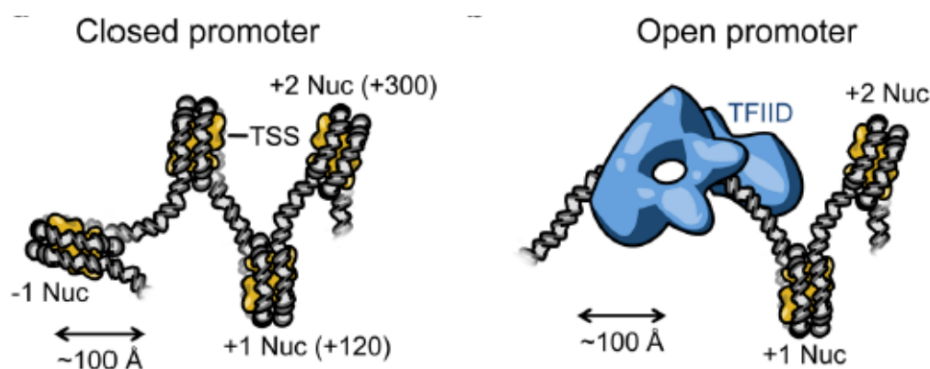
# **Expression des gènes classe II chez les eucaryotes (Transcription)**

## Expression des gènes classe II chez les eucaryotes (Transcription)

### 1. Initiation de la transcription

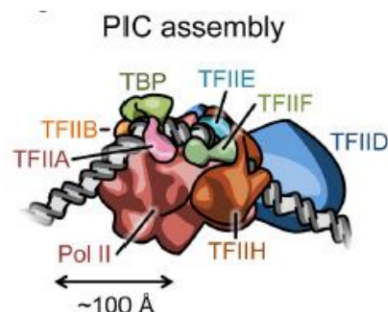
Les promoteurs minimaux sont classés en deux types principaux sur la base de la présence ou de l'absence d'une boîte TATA. Les promoteurs ne présentant pas de boîte TATA possèdent par contre un élément INR (Initiator Element) et/ou un élément DPE (Downstream Promoter Element).

La formation du complexe de pré-initiation (Figure 1), ou PIC, est initiée par la liaison de la sous-unité TBP (TATA binding protein) du complexe TFIID à la boîte TATA (Promoteurs avec TATA Box) située environ 25 nucléotides en amont du site d'initiation de la transcription. Cette liaison est stabilisée par TFIIA. La fixation de TFIID provoque une importante distorsion de l'ADN au niveau de la boîte TATA qui aide au recrutement de TFIIB (**Figure 1**).



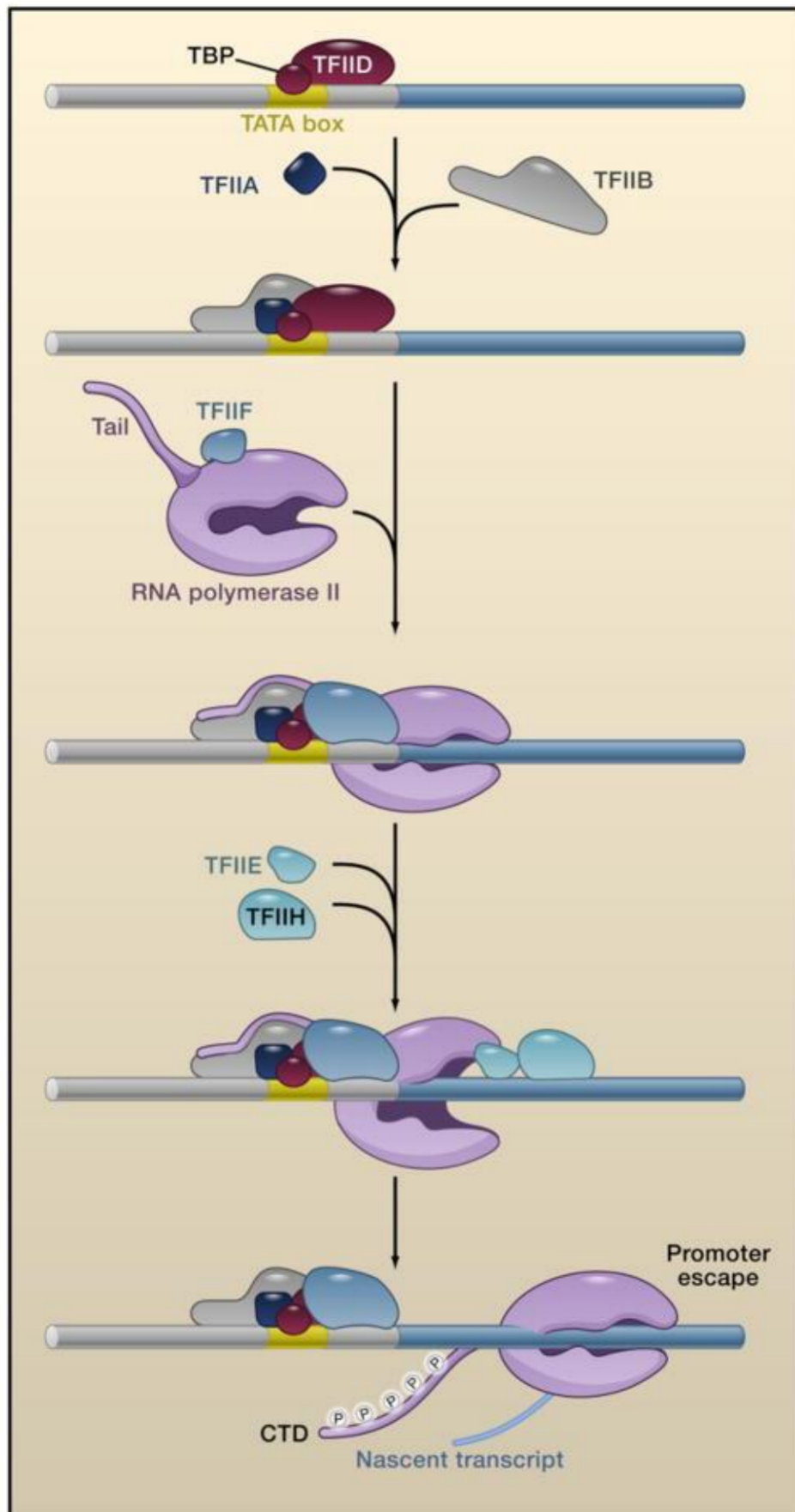
**Figure 1.** Fixation de complexe TFIID.

Le complexe TFIIB se lie à la protéine TBP ainsi qu'à une séquence située en aval de la boîte TATA appelée BRE (pour TFIIB-recognition element). TFIIB définit le site d'initiation de la transcription et favorise le recrutement de l'ARN pol II et de TFIIF. A cette étape même si l'ARN pol II est associée au complexe de manière stable, complexe de pré-initiation intermédiaire, elle ne peut initier la transcription en l'absence de TFIIE et TFIIH. TFIIE est d'abord recruté par interaction directe avec l'ARN pol II. Une fois incorporé dans le PIC, TFIIE recrute le facteur TFIIH (**Figure 2**).



**Figure 2.** Formation de PIC.

La liaison des facteurs généraux de transcription et la formation du PIC (**Figure 3**) peuvent être contrôlées par d'autres facteurs de transcription, **activateurs** ou **répresseurs**. De nombreux facteurs de transcription gène-spécifiques peuvent être recrutés au niveau des promoteurs.



**Figure 3.** Etapes de l'initiation de la transcription.

Le complexe ainsi formé est initialement instable étant donné que le promoteur se trouve dans une conformation « fermée » en double hélice. L'étape suivante de l'initiation est donc « l'ouverture » du promoteur provoquée par une dénaturation locale de l'ADN. Cette étape est assurée par les complexes TFIIE et TFIIH. TFIIE stabilise le complexe de pré-initiation intermédiaire via des interactions avec la quasi totalité des facteurs généraux de transcription au sein du PIC ainsi qu'avec l'ARN pol II (**Figure 2**). Quant le TFIIH se lie en aval de l'ARN pol II et ouvre l'ADN double brin grâce à son activité hélicase ATP dépendante, ensuite l'ouverture de la double hélice d'ADN entre les positions -9 et +2 de part et d'autre du site d'initiation de la transcription (site +1) pour former la **bulle transcriptionnelle (Figure 4)**. Ceci rendra possible l'appariement des ribonucléotides (rNTP) libres au brin matrice et donc l'initiation de la transcription. La catalyse nécessite la présence de deux ions  $Mg^{2+}$  qui se retrouvent au site actif de la polymérase. L'ARN pol II peut alors accéder au brin matrice et initier la synthèse de l'ARN au niveau du point +1 d'initiation de la transcription. Au fur et à mesure que la transcription progresse, la bulle s'étend en aval d'une paire de base de plus que la position du dernier rNTP ajouté.

## 2. Transition initiation-élongation

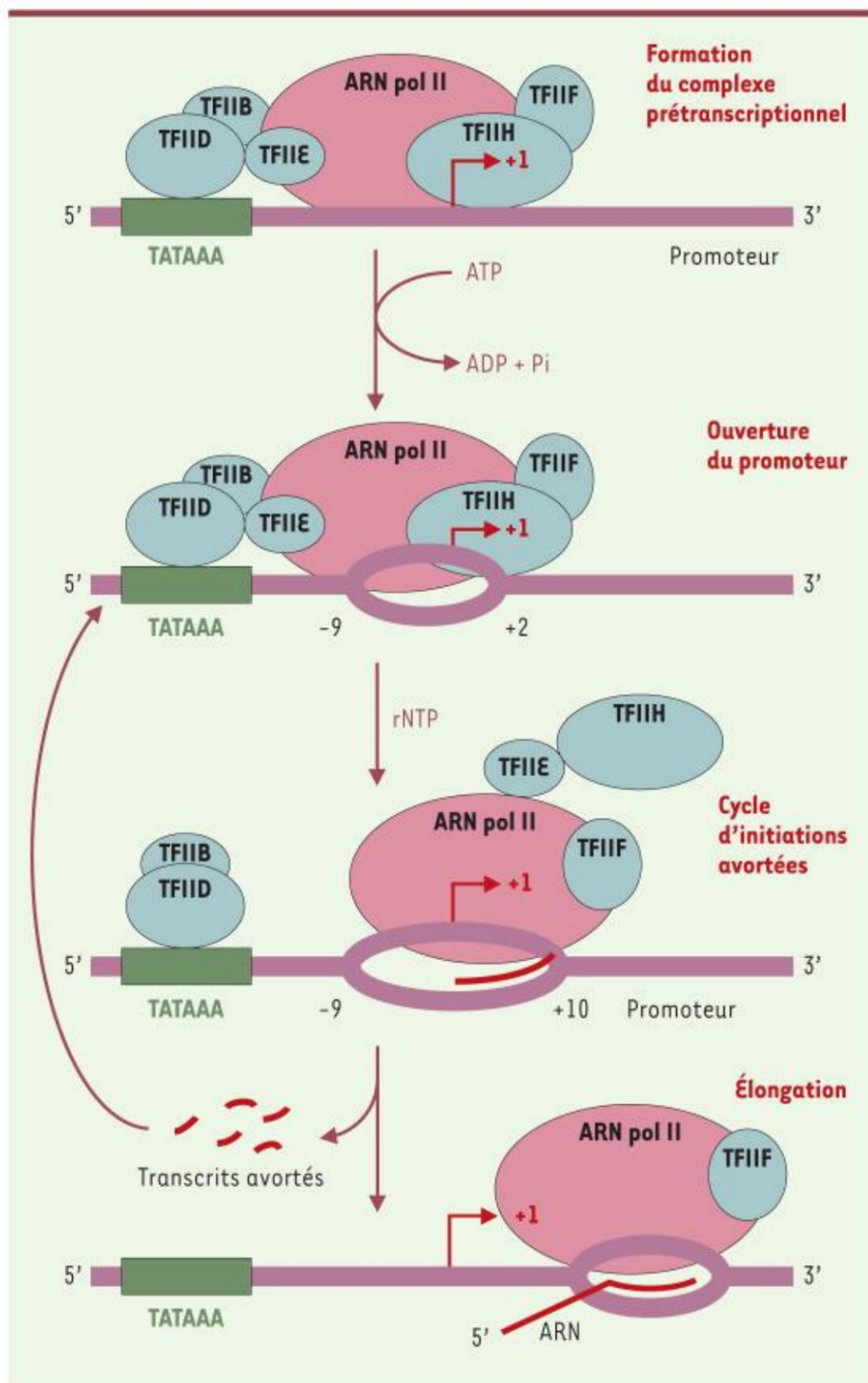
L'ARN pol II initiant la transcription doit subir une maturation structurale et fonctionnelle. La libération du promoteur est la première étape de ce processus de maturation, durant laquelle la polymérase rompt ses contacts avec les séquences promotrices et avec certains facteurs liés au promoteur et resserre simultanément son étreinte sur l'ARN naissant. La libération du promoteur est considérée comme terminée lorsque l'ARN naissant est stablement associé au complexe de transcription. Une première étape est franchie lorsque la taille de l'ARN synthétisé atteint une longueur de **quatre nucléotides**.

Dans le complexe d'initiation, le domaine amino-terminal de TFIIB s'insère dans le site catalytique de l'ARN pol II et semble stabiliser l'association de l'ARN naissant avec l'ARN pol II. Tandis que cette interaction aide à stabiliser le court transcrit ARN, après l'addition du **cinquième nucléotide**, le transcrit naissant commence à se heurter au domaine amino-terminal de TFIIB qui doit alors être déplacé (**Figure 4**). En conséquence, l'ARN croissant aide à diriger les changements conformationnels et augmente la stabilité du complexe d'**élongation**. Enfin, lorsque le transcrit naissant a une taille de **8 à 9 nucléotides**, il atteint la longueur de l'hybride ARN-ADN dans le complexe d'élongation. La formation de cet hybride ARN-ADN stable diminue la probabilité d'un relargage de l'ARN.

La partie amont de la bulle de transcription qui s'est allongée se referme à cette étape, stabilisant encore le complexe et relargant l'énergie accumulée, ce qui pourrait aider la libération du promoteur. TFIIH joue également un rôle important à cette étape, non seulement en ouvrant un peu plus le double brin d'ADN en aval mais aussi en phosphorylant les résidus sérine 5 du CTD de l'ARN pol II. Cette phosphorylation pourrait en effet aider à rompre les contacts avec certains facteurs liés au promoteur.

Des analyses de la transition de l'initiation à l'élongation ont montré que tous les facteurs généraux coexistent dans les complexes d'initiation matures, mais que, suite à l'ajout de nucléotides, ce complexe est déstabilisé. Durant cette transition, TFIID, TFIIA, TFIIH, TFIIE et le complexe médiateur restent

associés au promoteur tandis que TFIIB et TFIIF sont relargués. TFIIF est relargué du complexe d'initiation quelque part après la formation de la **dixième liaison phospho-diester**. Après son relargage, TFIIF peut se réassocier à l'ARN pol II en pause, ce qui rend ce facteur unique en son genre puisqu'il est le seul des facteurs de transcription généraux détecté dans les complexes d'élongation (Figure 4).

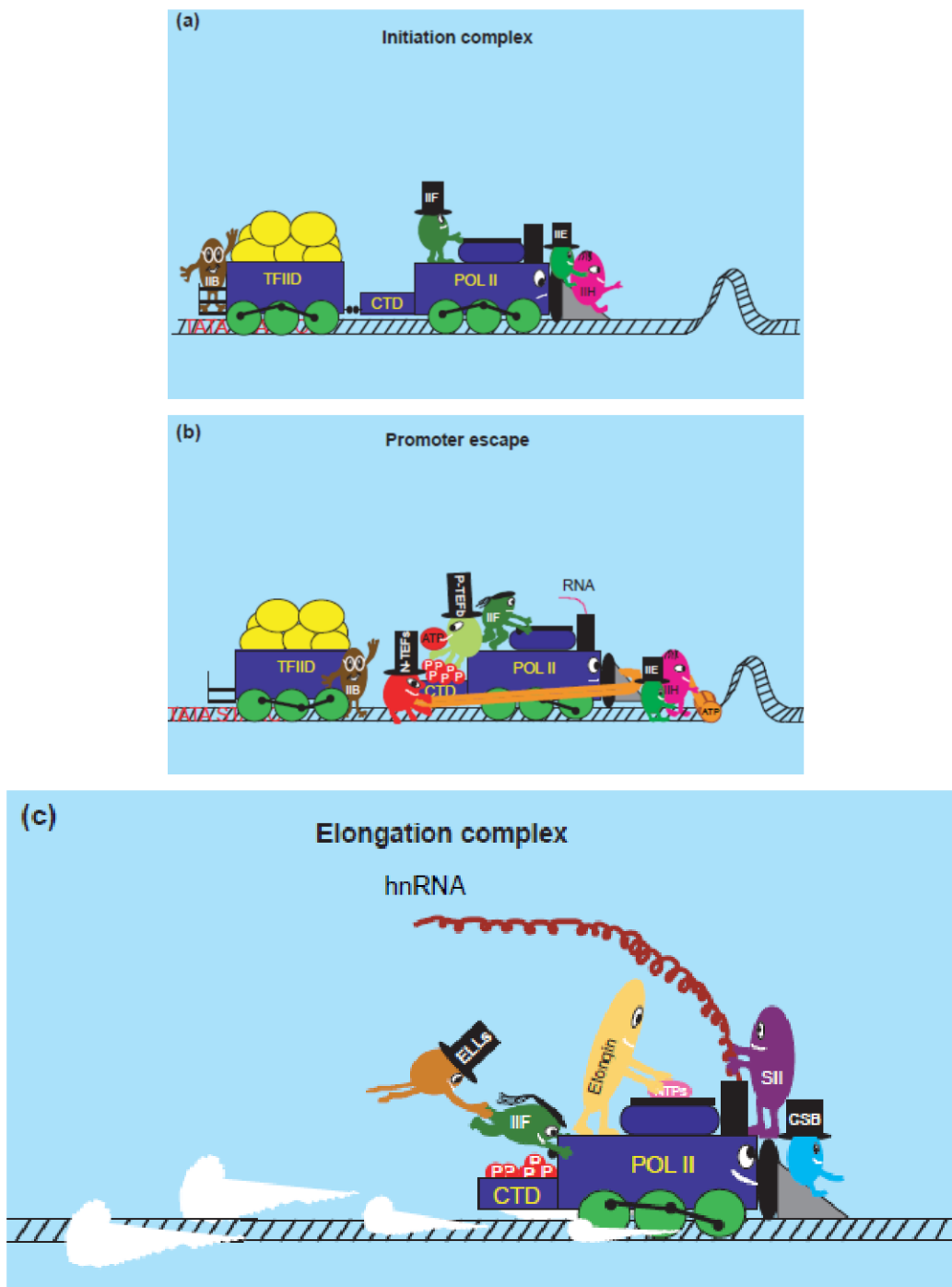


**Figure 4.** Etapes de Transition initiation-élongation d'ARN pol II.

À la position +11, le complexe entre alors dans le mode de **l'élongation productive**.

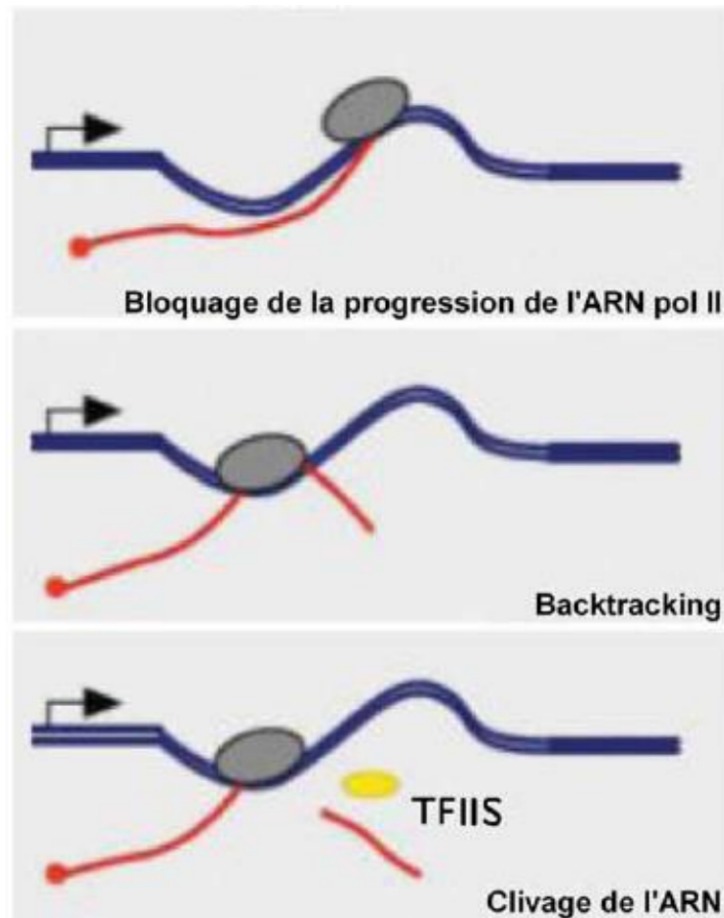
### 3. Elongation de la transcription

L'élongation de la transcription est un processus discontinu qui est fréquemment interrompu par divers facteurs. En effet, l'ARN pol II est très sensible aux signaux intrinsèques tels que des **séquences** et/ou des **structures de l'ADN** ou du transcrit naissant particulières et aux signaux extrinsèques tels que l'association de facteurs régulateurs. Ces signaux peuvent entraîner un **blocage** de la progression de la polymérase, qui marque alors une pause transcriptionnelle. Ces pauses ont été identifiées pour chacun des trois types d'ARN pol eucaryotes, ainsi que pour les ARN pol virales et procaryotes. Divers facteurs, appelés **facteurs d'élongation** (**Figure 5 c**), ont pour rôle de contrebalancer ou de limiter ces blocages, modulant ainsi les pauses transcriptionnelles et donc la vitesse d'élongation.



**Figure 5.** Transition complexe d'initiation-complexe d'élongation d'ARN pol II.

Lorsque l'ARN pol II glisse de quelques nucléotides vers l'arrière sur la matrice ADN dans un mouvement appelé « **backtracking** », elle se déplace par rapport à l'ARN naissant de sorte que l'extrémité 3'-OH de cet ARN est déplacée par rapport au site actif de la polymérase. Cette ARN pol II arrêtée ne peut pas reprendre la transcription en l'absence de facteurs **stimulant le clivage** de l'ARN. En effet, la formation d'une nouvelle extrémité 3'-OH de l'ARN correctement alignée avec le site actif de la polymérase est essentielle pour la reprise de l'élongation. Le facteur accessoire **TFIIS** (également appelé **SII**) réactive les complexes de transcription ayant subi un « backtracking » en activant le clivage du transcrit par l'ARN pol II (**Figure 6**). Cela permet le réalignement du site actif de la polymérase avec la nouvelle extrémité de l'ARN afin que la transcription puisse reprendre son cours.



**Figure 6.** Mode d'action de TFIIS.

Le facteur général de transcription **TFIIF** diminue le temps des pauses marquées par l'ARN pol II et stimule la vitesse d'élongation. Il a été proposé que l'association transitoire de TFIIF avec l'ARN pol II en pause induit un changement conformationnel dans la polymérase nécessaire pour une élongation optimale. TFIIF serait ensuite relargué du complexe d'élongation.

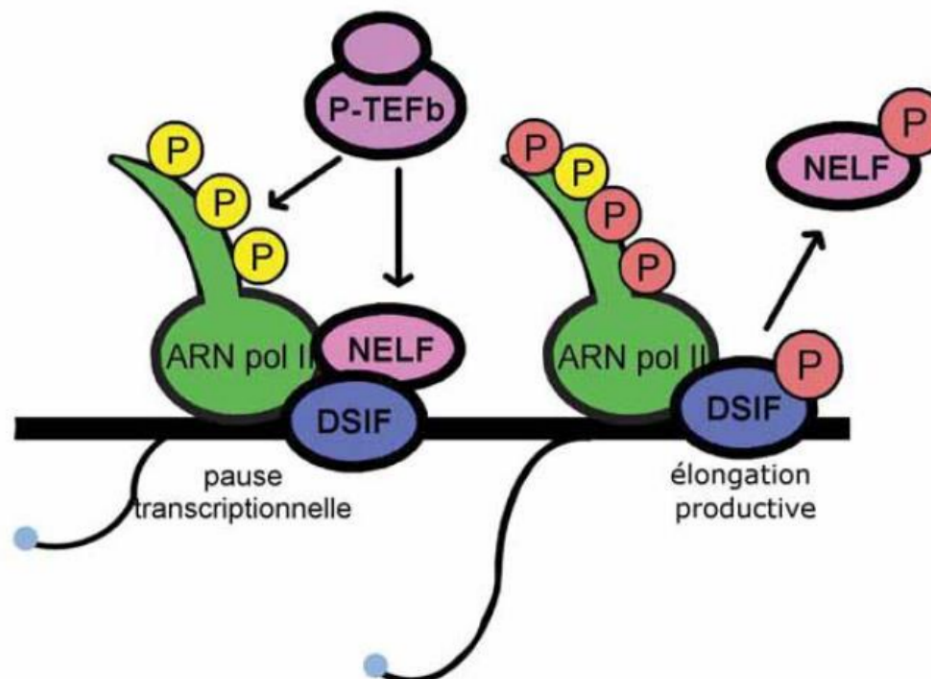
Un autre facteur de transcription participe dans l'élongation c'est le complexe **Elongin** (ou **SIII**) est constitué de trois sous-unités, l'Elongin A est la sous-unité transcriptionnellement active tandis que les sous-unités B et C agissent comme des sous-unités régulatrices. L'Elongin A stimule fortement le taux d'élongation par l'ARN pol II, il a été proposé que ce complexe pourrait supprimer les pauses transitoires de l'ARN pol II en aidant à maintenir l'extrémité 3'-OH de l'ARN naissant dans la bonne position dans le

site actif de la polymérase. Il a été montré que le complexe Elongin n'augmente l'élongation par l'ARN pol II qu'en l'absence du facteur TFIIF dans le complexe d'élongation. Il est donc possible que le complexe Elongin serve de facteur complémentaire pour supprimer les pauses transcriptionnelles au cas où TFIIF échouerait à se réassocier avec le complexe d'élongation en pause.

Le facteur d'élongation **PAF1c** (Polymerase Associated Factor complex) augmente les taux d'élongation de la transcription tant sur de l'ADN nu que sur la chromatine, et fonctionne de concert avec nombre d'autres facteurs modulant la structure de la chromatine et la maturation des ARN.

La famille des **ELL** (eleven-nineteen lysinerich Leukemia) c'est des analogue fonctionnelle de TFIIF et Elongin ; Il ont montré (2012) qu'une diminution de l'expression de ELL, dans des cellules humaines en culture, entraînait une déstabilisation du PIC et une perturbation de l'élongation précoce. Ces résultats suggèrent que ELL pourrait également jouer un rôle beaucoup plus précoce au cours du cycle de transcription par l'ARN pol II.

En effet, l'ARN pol II en pause semble empêcher le positionnement des nucléosomes au niveau du promoteur en entrant en compétition avec ces derniers pour l'occupation du promoteur. Ce qui implique deux types de facteurs de régulations des pauses d'élongation (**Figure 7**) ; *facteurs imposant la pause transcriptionnelle à proximité du promoteur (N-TEF)* et *facteurs permettant la levée de la pause transcriptionnelle ((P-TEFb).*



**Figure 7.** P-TEFb lève la pause transcriptionnelle marquée par l'ARN pol II à proximité du promoteur. Peu après l'initiation de la transcription la progression de l'ARN pol II est bloquée par l'action concertée des deux facteurs négatifs d'élongation (N-TEF), DSIF et NELF. Pour contrecarrer l'action de N-TEF, le facteur d'élongation de la transcription P-TEFb phosphoryle DSIF, NELF et les résidus Ser2 du CTD de l'ARN pol II. Ces événements de phosphorylation facilitent la reprise d'une élongation productive en causant la dissociation de NELF de l'ARN pol II et en convertissant DSIF en activateur de la transcription.

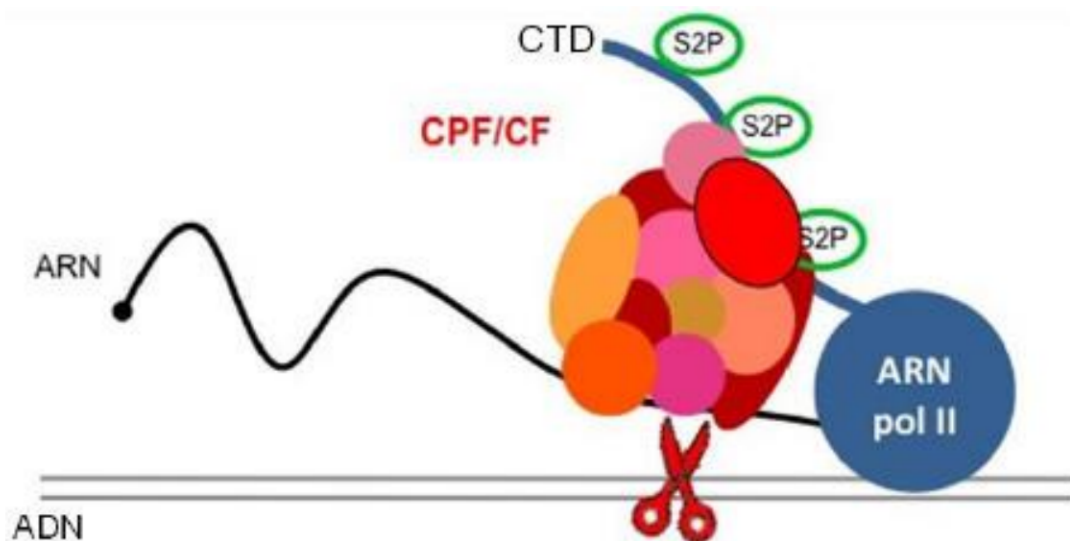
#### 4. Terminaison de la transcription

La terminaison se produit lorsque l'ARN pol II cesse la synthèse de l'ARN et que la polymérase et l'ARN naissant se détachent de la matrice ADN. Des études récentes ont également montré l'importance de la terminaison pour la synthèse de l'ARNm. Ainsi, une terminaison défective à l'extrémité 3' d'un gène transcrit par l'ARN pol II entraîne une diminution de l'épissage et une augmentation de la dégradation de l'ARN, ainsi qu'une diminution de l'initiation de la transcription sur ce gène. Cela pourrait servir de mécanisme de contrôle, diminuant la transcription lorsque la terminaison échoue. La terminaison de la transcription par l'ARN pol II peut se faire par différents moyens en fonction des signaux de maturation en 3' de l'ARN et des facteurs de terminaison présents à l'extrémité du gène.

La terminaison de la transcription par l'ARN pol II suit deux principales voies alternatives. Pour chacune d'elle, la terminaison est provoquée par le recrutement de complexes de terminaison qui se fixent à la fois à des séquences spécifiques sur le transcrit naissant et à l'ARN pol II. Les deux complexes de terminaison alternativement responsables de la terminaison sont le complexe de clivage et de polyadénylation (CPF-CF) et le complexe Nrd1- Nab3-Sen1 (NNS).

Le CPF-CF reconnaît un ensemble de séquences spécifiques sur l'ARN naissant puis il clive cet ARN et ajoute une queue poly(A) à l'extrémité 3' ainsi produite. Cette queue poly(A) est constituée d'environ 70 adénosines (250 chez les mammifères) et elle est importante pour la stabilité, l'export et la traduction des ARN (**Figure 8**).

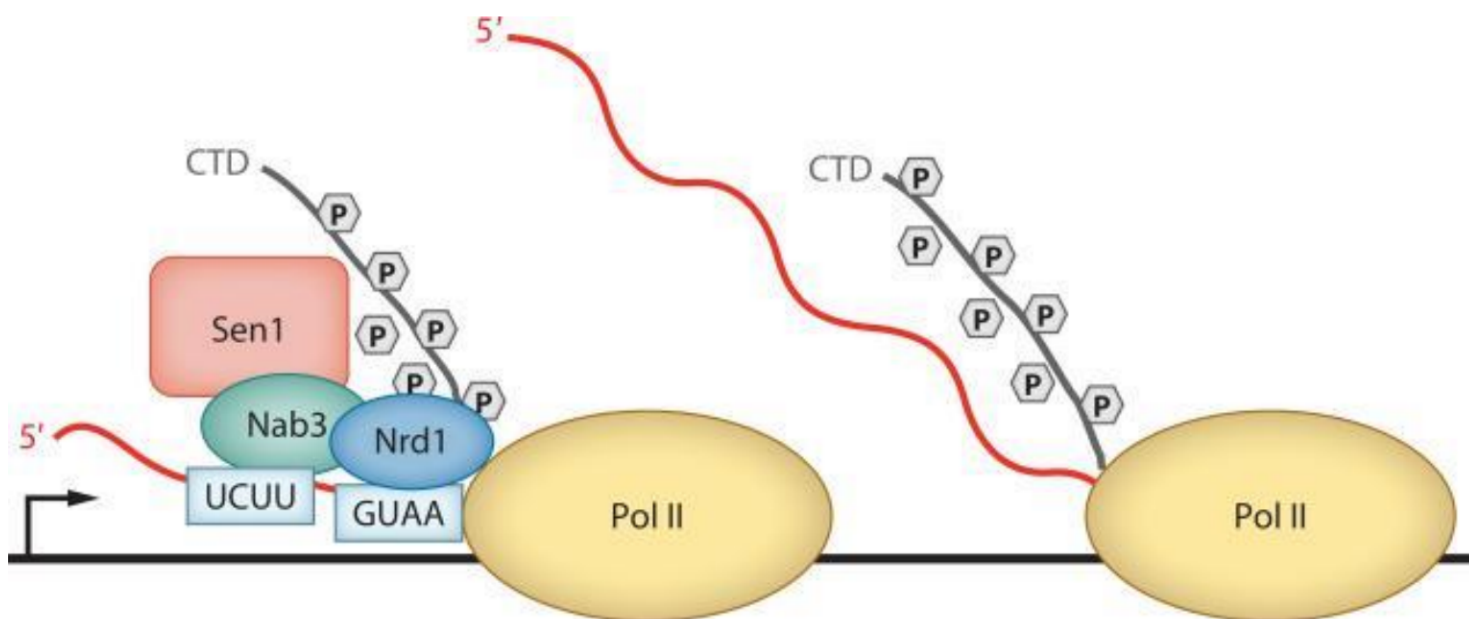
Chez *S. cerevisiae*, les signaux de recrutement du complexe CPF-CF sont composés de cinq éléments situés dans la région du site de clivage et de polyadénylation. Ces séquences sont appelées « élément d'efficacité », « élément de positionnement », « élément riche en U en amont », « élément riche en U en aval » et « site de clivage et de polyadénylation ». L'ensemble de ces séquences sont liées par des sous-unités du CPF-CF (20 Sous-unités) qui participent au positionnement du complexe sur l'ARN.



**Figure 8.** Terminaison *via* CPF-CF.

Le complexe NNS est composé de 3 protéines essentielles : Nrd1 (nuclear pre-mRNA down-regulation), Nab3 (nuclear polyadenylated RNA-binding) et Sen1 (splicing endonuclease), qui se lient coopérativement à l'ARN (**Figure 9**).

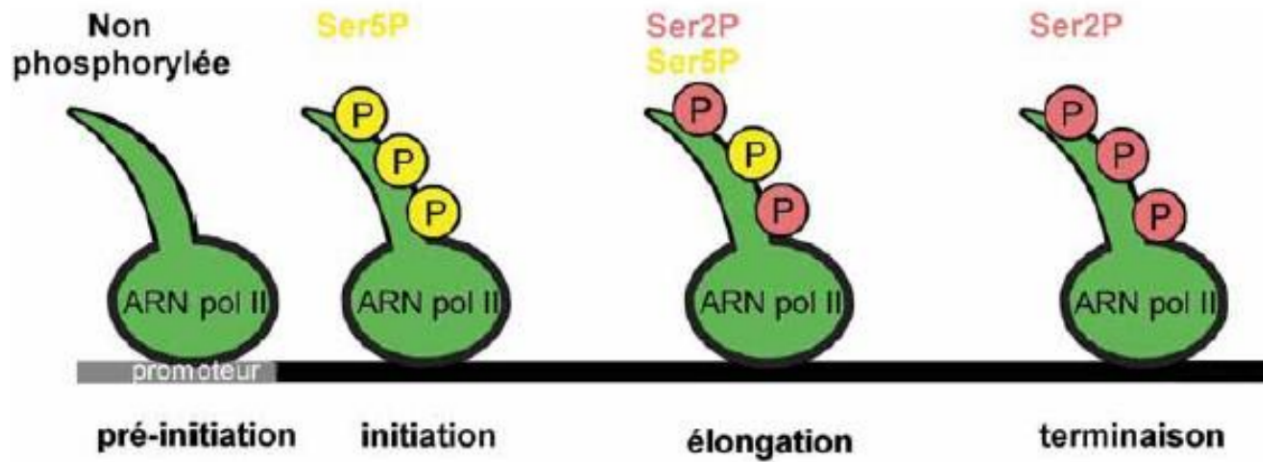
Nrd1 est une protéine de 68kDa, elle est au coeur du mécanisme de terminaison par le NNS. Elle coordonne les étapes de ciblage de l'ARN naissant, de sa libération, de sa polyadénylation et du recrutement de l'exosome responsable de la dégradation du transcrit. Un motif de reconnaissance de l'ARN situé dans sa région centrale lui permet de se fixer à l'ARN de manière spécifique. Nrd1 interagit avec les deux autres sous-unités du NNS, Sen1 et Nab3. Ce dernier, protéine de 90kDa, lie l'ARN grâce à un motif de reconnaissance des ARN dans sa partie centrale. Sen1, est une hélicase de 252kDa ARN/ADN. Avec une capacité à interagir spécifiquement avec le CTD de l'ARN pol II, ce qui lui donne un rôle très important dans la terminaison.



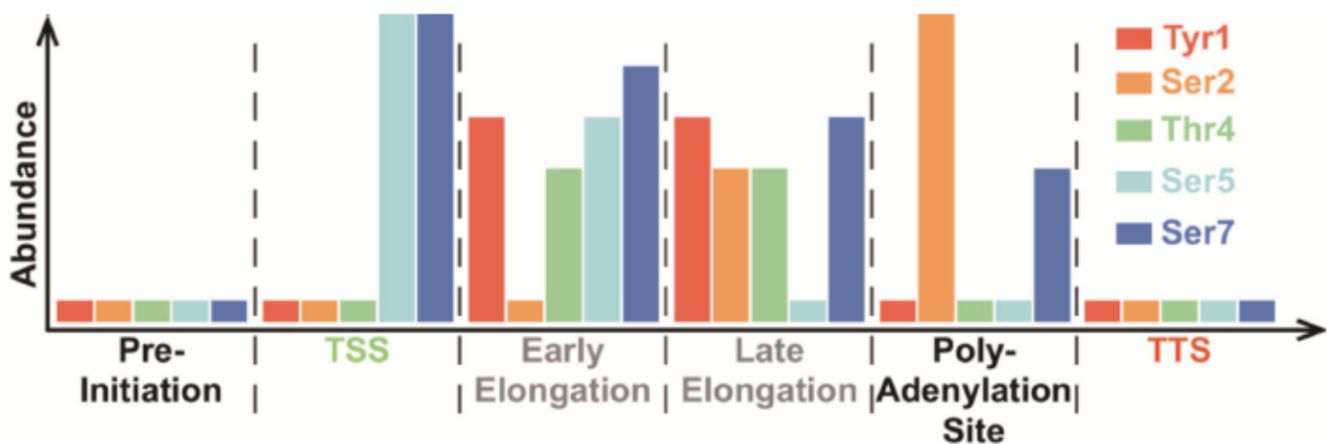
**Figure 9.** Terminaison via le complexe NNS.

### 5. Phosphorylation de CTD d'ARN pol II

Le domaine carboxy-terminal (CTD) de la plus grande des sous-unités, Rpb1, présente un motif heptapeptidique Tyr-Ser-Pro-Thr-Ser-Pro-Ser répété en tandem. Tandis que les trois ARN polymérases ont la même origine évolutive et ont une composition en sous-unités similaire, le CTD est unique à l'ARN pol II et sa longueur semble corrélée à la complexité du génome. Le CTD est le site majeur de phosphorylations réversibles de l'ARN pol II. En effet, cinq des sept résidus consensus peuvent être phosphorylés. Les résidus sérines 2 et 5 (Ser2 et Ser5) ont été identifiés comme des sites de phosphorylation majeurs (**Figure 10 ; 11**).



**Figure 10.** L'ARN pol II est représentée à quatre positions différentes le long d'un gène et à chacune de ces positions, son domaine carboxy-terminal (CTD) présente un état de phosphorylation particulier. En jaune sont représentés les résidus Ser5 phosphorylés et en rose sont représentés les résidus Ser2 phosphorylés.



**Figure 11.** Variation de CTD d'ARN pol II de *S. cerevisiae* (TSS = transcription start site. TTS = transcription termination site).

D'une manière générale, l'activation des gènes dans les cellules eucaryotes (dans lesquelles la cible de la machinerie transcriptionnelle est la chromatine) débute par deux étapes majeures. Premièrement, la barrière engendrée par la présence de nucléosomes, qui constituent les unités de base de la chromatine, est levée grâce à l'action combinée d'un ensemble de cofacteurs qui induisent des changements locaux et transitoires de la structure chromatinienne à proximité du promoteur. Dans un deuxième temps, le complexe de pré-initiation (PIC) constitué de l'ARN pol II et des facteurs généraux de transcription s'assemble et initie la transcription.

# **Expression des gènes classe I chez les eucaryotes (Transcription)**

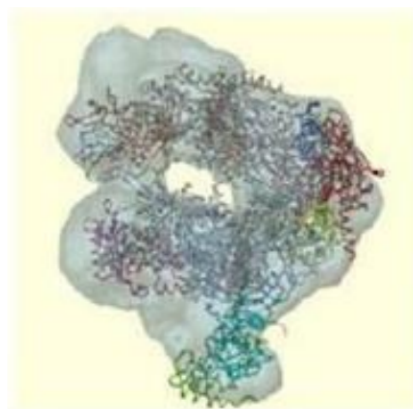
## Expression des gènes classe I chez les eucaryotes (Transcription)

### 1. ARN polymérase I

La transcription de l'ADN ribosomique codant les ARNr 18S, 25S et 5.8S est assurée par l'ARN polymérase I. L'ARN polymérase I (pol I) est entièrement dédiée à la transcription des ADNr qui représente à elle seule 50% des efforts transcriptionnels de la cellule. La Pol I est un complexe enzymatique constitué de 14 unités protéiques. 12 de ces unités forment un coeur protéique retrouvé dans les trois ARN polymérases. Ce coeur est composé de cinq protéines qui sont communes aux trois polymérases : Rpb5, Rpb6, Rpb8, Rpb10 et Rpb12 ; et de sept paralogues présents dans les différentes ARN polymérases. Deux autres protéines sont spécifiques de la Pol I (**Figure 1**).

**Tableau 1.** Composition d'ARN pol I et II.

Sous unité	ARN pol I	ARN pol II
<b>Protéines communes</b>	<b>Rpb5, Rpb6, Rpb8, Rpb10, Rpb12</b>	
<b>Protéines paralogues</b>	<b>Rpa190</b>	<b>Rpb1</b>
	<b>Rpa135</b>	<b>Rpb2</b>
	<b>Rpc40</b>	<b>Rpb3</b>
	<b>Rpc19</b>	<b>Rpb11</b>
<b>Protéines spécifiques</b>	<b>Rpa12</b>	<b>Rpb7</b>
	<b>Rpa14</b>	<b>Rpb9</b>
	<b>Rpa43</b>	<b>Rpb4</b>
	<b>Rpa34</b>	
	<b>Rpa49</b>	

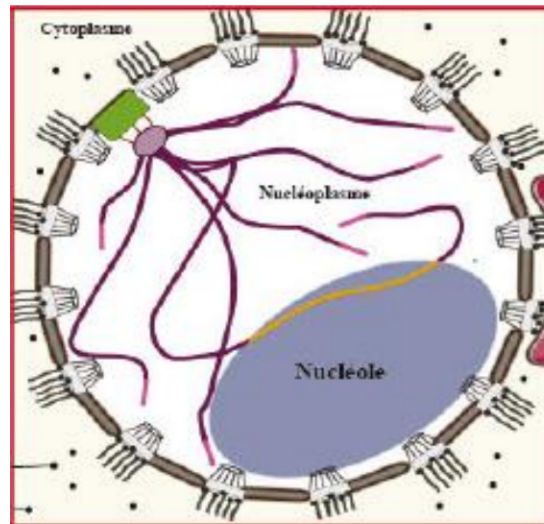


**Figure 1.** Comparaison de l'enveloppe de l'ARN polymérase I avec la structure cristalline de l'ARN polymérase II.

La transcription par l'ARN pol I a lieu dans une région spécifique du noyau, le nucléole.

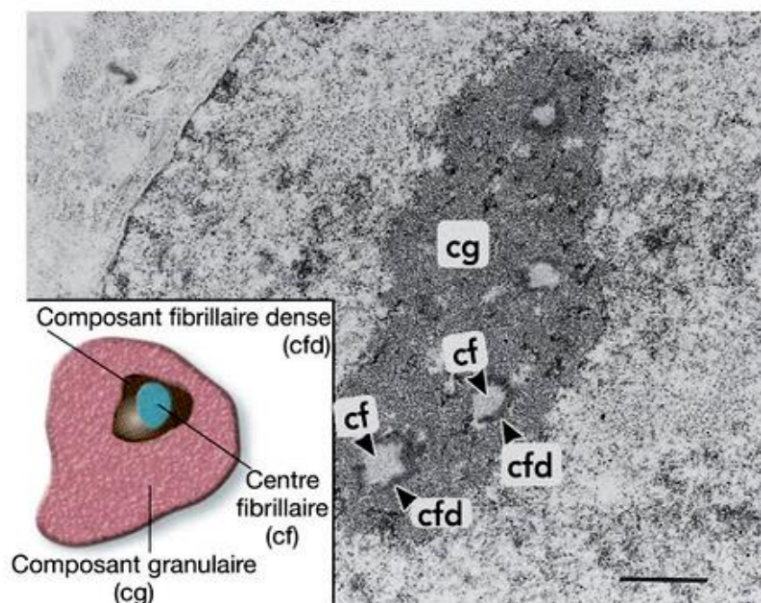
## 2. Nucléole

On trouve au sein du noyau un territoire facilement identifiable, le nucléole. Ce n'est pas un organelle, car il n'est pas entouré d'une membrane, mais un territoire dynamique à forte concentration en protéines et en ARN, d'où une haute densité photonique et électronique, qui permet de le visualiser aisément.



**Figure 2.** Noyau et nucléole de *S. cerevisiae*.

L'observation électronique du nucléole de mammifères met en évidence un empilement concentrique de trois domaines. La masse de nucléole est formée d'un composant granulaire (cg) qui contient les sous-unités ribosomiques à divers stades d'assemblage. Les centres fibrillaires (cf) entourés d'un composant fibrillaire plus dense (cfd) sont englobés parmi les granules. Les cf contiennent l'ADNr et le cfd contient les transcrits pré-ARNr naissants et les protéines associées. La transcription de précurseur pré-ARNr se fait à la limite entre les cf et cfd (**Figure 3**).

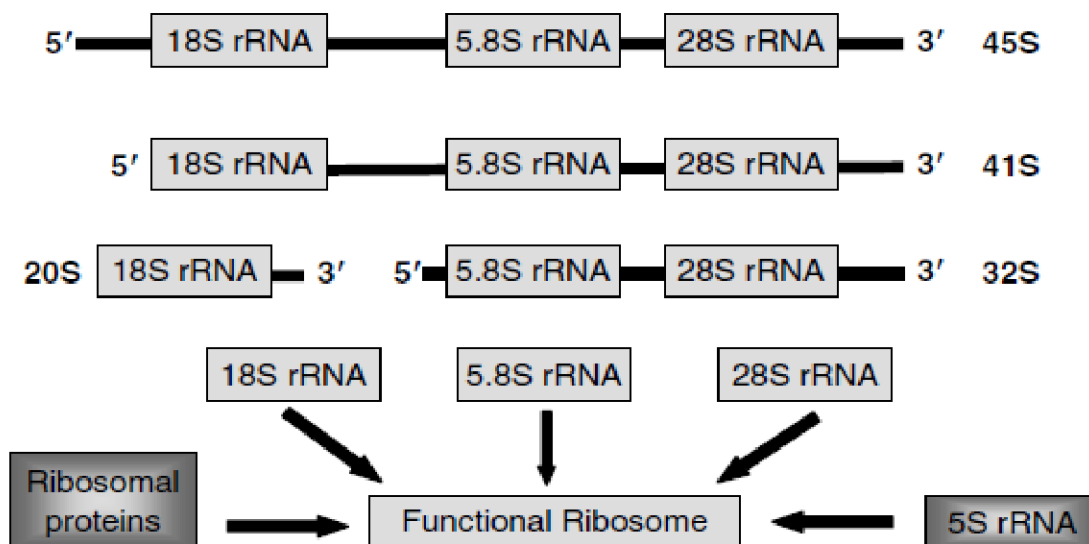


**Figure 3.** Microphotographie électronique d'un nucléole.

On trouve de un à cinq nucléoles dans les cellules de mammifères, alors qu'un seul est présent chez la levure *S.cerevisiae*. Ceci est dû au fait que le locus d'ADNr n'est porté que par un seul chromosome chez *S. cerevisiae* (le chromosome XII), alors qu'il est réparti sur plusieurs chez l'homme par exemple.

### 3. Transcription d'ARNr

Chez les eucaryotes la biogenèse des ribosomes est un processus largement conservé. Elle débute dans le nucléole, où le précurseur des grands ARNr, appelé pré-ARNr (pré-ARN 47S chez les mammifères) est synthétisé par l'ARN pol I. Ce précurseur subit par la suite un ensemble de modifications post-transcriptionnelles, faisant intervenir des petits ARN guides: les "small nucleolar ARN" (snoARN). En plus de ces snoARN, une série d'exo et d'endonucléases interviennent dans la maturation et permettent la production des formes matures des ARNr 18S, 5,8S et 25S (Figure 4).



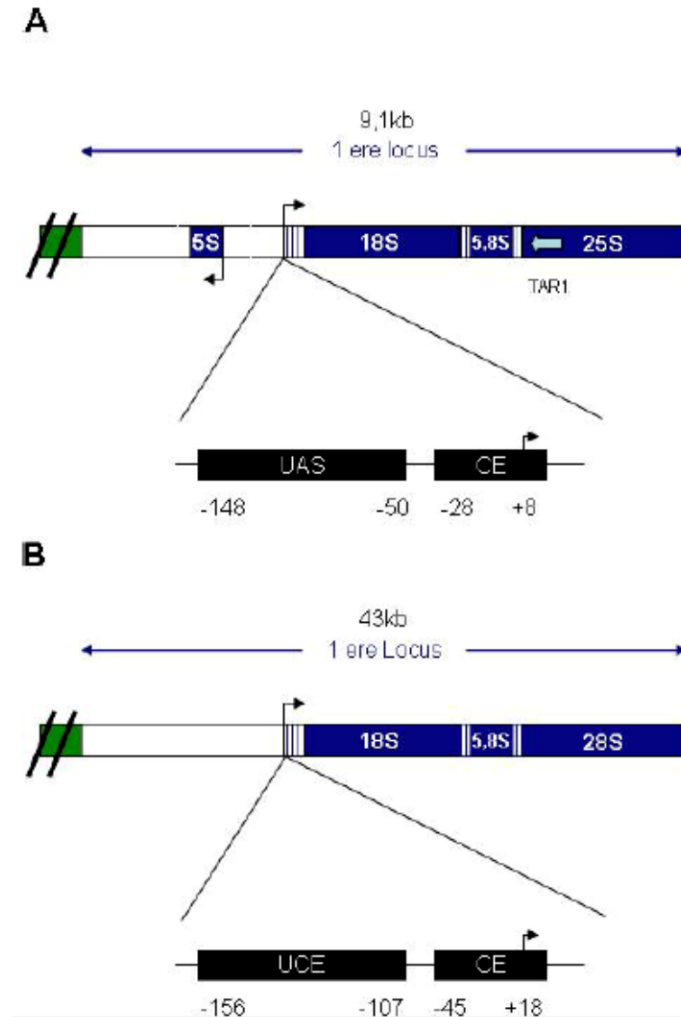
**Figure 4.** Etapes de biogénèse d'ARNr des eucaryotes.

### 4. Unité de transcription de l'ARN polymérase I

Comme nous l'avons vu ci-dessus, la synthèse des quatre ARNr d'eucaryotes est obtenue par la maturation de deux transcrits primaires, les pré-ARNr, synthétisés respectivement par les ARN polymérase I (ARNr 25S, 28S et 5,8S) et III (ARNr 5S). Leur taille (exprimée en Svedbergs) n'est pas la même chez tous les eucaryotes, d'où une certaine confusion dans la nomenclature puisqu'on parle ainsi de pré-ARNr 35S chez la levure et pré-ARNr 47S chez les mammifères.

L'ADNr a une grande diversité d'organisation chez les eucaryotes. L'ARNr de levure et de mammifères qui diffèrent par la taille du transcrit et de la région intergénique. Chez *S. cerevisiae*, l'unité de base de l'ADNr a une taille de 9,1 kb (**Figure 5**) avec un transcrit de 6,7 kb (l'ARNr 35S) et une séquence intergénique de 2,4 kb. Chez les mammifères, l'ADNr est de 43 kb, avec une séquence codante de 13,7 kb et une séquence intergénique de 29,3 kb.

La région promotrice de l'ADNr est composée de deux éléments principaux : le coeur du promoteur ou bien Contrôle Element (CP ou CE), essentiel à l'initiation de la transcription ; et la séquence activatrice UE (Upstream Element), requise pour une forte transcription (Figure 5). L'espacement de ses deux éléments est crucial à leur fonction *in vivo* mais la région CP est suffisante pour démarrer la transcription *in vitro*.



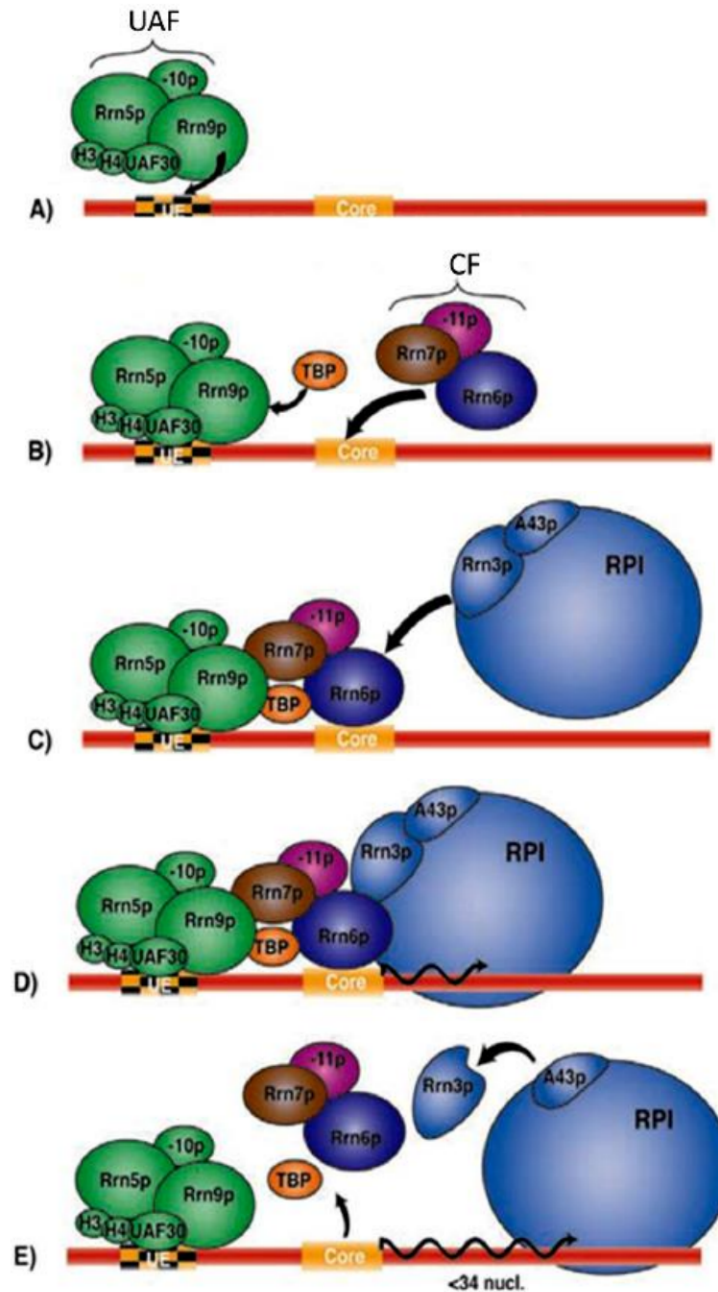
**Figure 5.** Structure du locus d'ADNr : Les boîtes noires représentent les régions promotrices. Les petites flèches indiquent le site d'initiation de la transcription. UAS = Upstream Activating Sequence, CE = Contrôle Element, UCE = Upstream Control Element. **(A)** Organisation du locus d'ADNr chez la levure *S. cerevisiae*. Le gène codant pour ARNr 5S est localisé entre deux gènes codant pour ARNr 35S. **(B)** Organisation du locus d'ADNr de cellules humaines. La longueur des séquences est indiquée en kb. La flèche bleu clair dans l'ADNr de *S. cerevisiae* symbolise le gène TAR1 transcrit par l'ARN polymérase II.

## 5. Transcription du pré-ARNr par l'ARN Pol. I

### A. Initiation

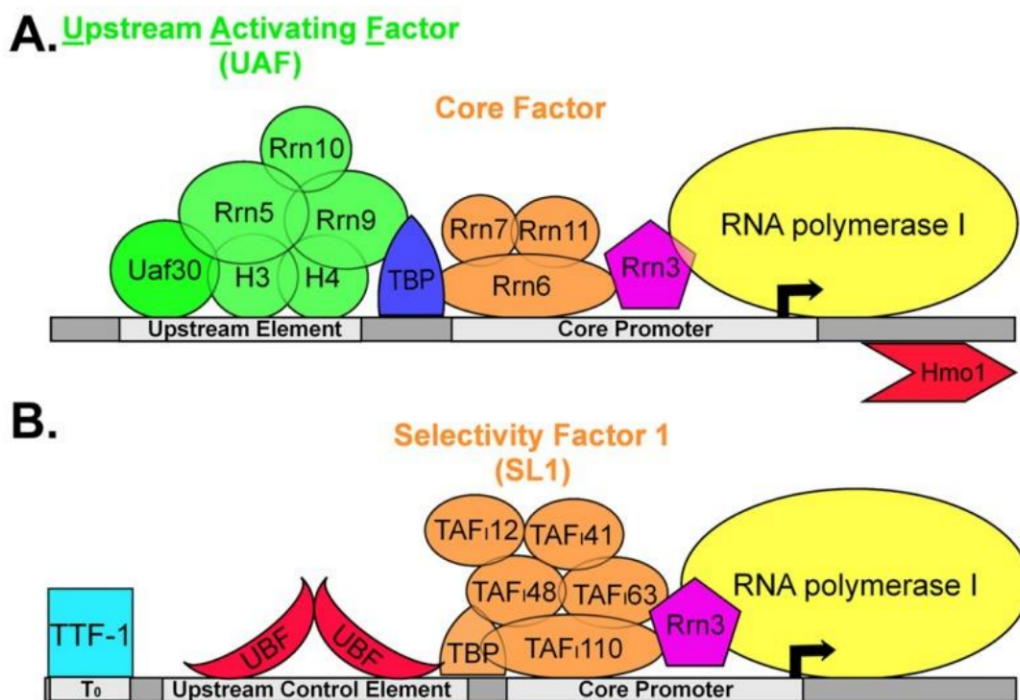
L'initiation de la transcription de l'ADNr par l'ARN Pol. I de *S. cerevisiae*. requiert l'utilisation de 11 facteurs additionnels, répartis en 4 complexes distincts. Elle débute lorsque le complexe multiprotéique **UAF (Upstream Activating Factor)** s'associe de manière très stable à la séquence activatrice **UE** (Figure

6). Ce complexe est composé de 6 protéines différentes dont les histones **H3** et **H4**, impliquées dans l'affinité du complexe avec l'ADN ; le facteur **Uaf30p**, dédié à la spécificité de séquence ; et le facteur **Rrn9p**, qui interagit avec la protéine de liaison à une boîte TATA (**TBP** : TATA Binding Protein). La liaison entre la protéine TBP et Rrn9p va permettre le recrutement du complexe **CF (Core Factor)**, composé par **Rrn6p**, **Rrn7p** et **Rrn11p**, qui va renforcer l'association de la **polymérase I** avec la région promotrice et positionner correctement l'ARN Pol. I au site d'initiation de la transcription. L'interaction entre Rrn6p et Rrn3p permet l'association de cette dernière avec Rpa43p, sous-unité de l'ARN Pol. I, qui rendra la polymérase compétente pour la transcription.



**Figure 6.** Recrutement de l'ARNr Pol. I et initiation de la transcription de l'ADNr chez *S. cerevisiae*.

Le recrutement d'ARN pol I chez souris et chez l'humain nécessite deux facteurs, UBF (Upstream Binding Factor), et TIF-IB (souris) ou SL1 (humain). qu'il existe une certaine conservation de l'organisation du promoteur de l'ADNr entre la levure et les mammifères ce qui laissait supposer une homologie fonctionnelle entre leurs facteurs de pré-initiation (Figure 7).



**Figure 7.** Comparaison entre PIC de levure et humain.

SL1 est un complexe de 300 kDa, composé de la TBP et de trois sous-unités distinctes appelées hTAFI110, hTAFI63 et hTAFI48. Un système *in vitro* constitué des formes recombinantes des quatre protéines du complexe SL1 était insuffisant pour induire la transcription. Cela suggère que d'autres facteurs et/ou des modifications supplémentaires sont requis pour une activité optimale de SL1. En effet, à partir de cellules humaines, une quatrième TAFI de 41 kDa (TAFI41).

Le fait que UBF ait été isolé comme un facteur se liant spécifiquement à l'UCE a suggéré que UBF serait uniquement un facteur d'initiation qui co opérait avec SL1 pour organiser un complexe pré-initiation pour le recrutement de l'ARN polymérase I. Des recherches ont montré que UBF agirait également après le recrutement, en activant la transcription au niveau du « promoteur escape ».

### B. Elongation

Le recrutement de la Pol I par le CF permet de la positionner directement au point +1 de la transcription. Il est généralement admis qu'à la différence de la Pol II, la transcription Pol I ne nécessite pas d'hydrolyse de nucléotide triphosphate (NTP). L'initiation de la transcription nécessite l'association de la Pol I avec Rrn3 chez les mammifères et la levure. Juste après l'initiation de la transcription, la protéine Rrn3p, la TBP et le CF se dissocient de la Pol I, alors que l'UAF semble rester associé à l'UE, ce qui pourrait permettre un deuxième cycle de transcription plus rapide.

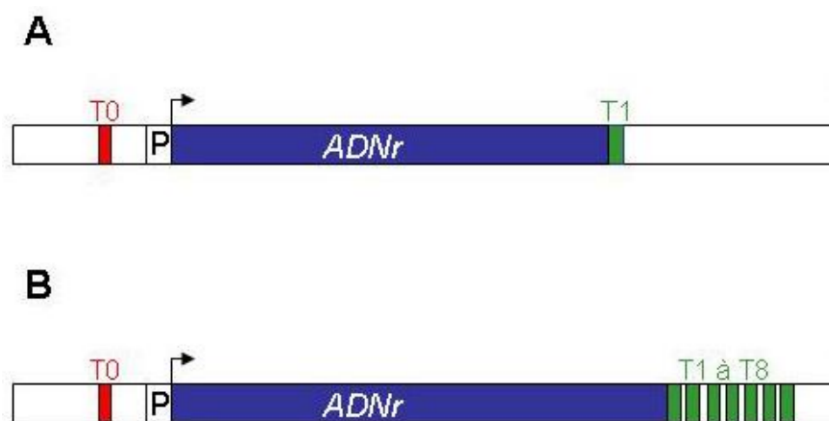
La dissociation de Rrn3p n'est pas nécessaire pour la poursuite de l'élongation dans la mesure où la fusion traductionnelle de Rrn3p avec la sous-unité A43 de la pol I n'affecte en rien la viabilité ou la croissance de la levure.

Les mécanismes contrôlant l'élongation par l'ARN polymérase I sont encore mal connus. Des résultats récents, encore que spéculatifs, suggèrent que certaines des protéines intervenant dans l'élongation par l'ARN polymérase II opèrent également avec l'ARN polymérase I. Il s'agit spécifiquement des facteurs Spt4, Spt5, Spt6, TFIIH et CSB/Rad26.

Le rôle précis de TFIIH dans la transcription reste assez mal connu. La proposition de participation du facteur TFIIH dans la transcription par l'ARN polymérase I se base sur le fait que TFIIH est capable de stimuler *in vitro* la transcription par l'ARN polymérase I. De plus, il a été localisé dans les nucléoles de cellules humaines. Néanmoins, on ne sait pas si la présence de TFIIH n'est pas due à une éventuelle présence d'ARN polymérase II dans le nucléole.

### C. Terminaison

Ce processus est déclenché par des séquences spécifiques, les terminateurs. Les facteurs spécifiques, Reb1 chez la levure et TTF-I (Transcription release Factor for Pol I) chez la souris, se fixent sur les terminateurs et agissent comme des barrières qui bloquent la translocation de l'ARN polymérase I. Chez *S. cerevisiae*, le terminateur (T1), située 13 nucléotides après la région transcrite par l'ARN polymérase I. Chez les mammifères, l'organisation est similaire, le terminateur est composé d'une séquence répétée 8 fois (T1 à T8), qui est située 21 nucléotides en aval de la séquence transcrite (**Figure 8**).

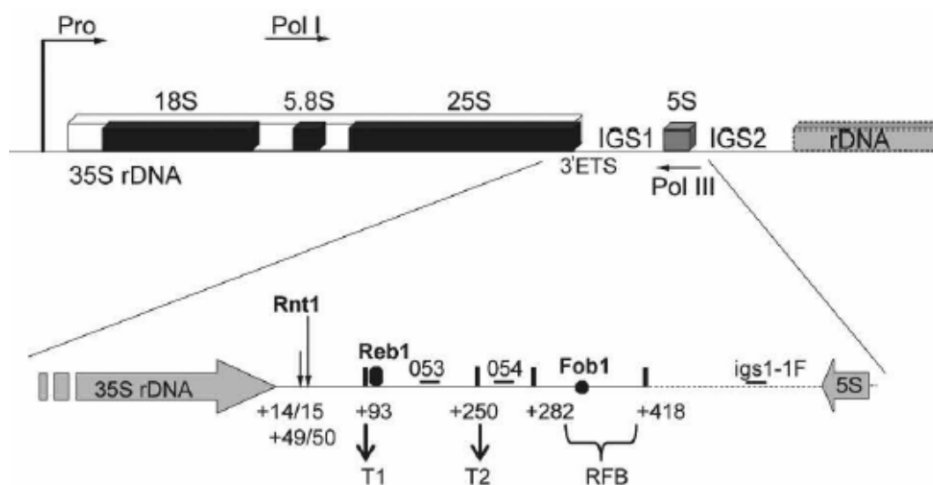


**Figure 8. Localisation des éléments terminateurs au sein de l'unité d'ADNr. (A)** Représentation schématique de deux unités successives d'ADN ribosomique de *S. cerevisiae*. La séquence transcrite par l'ARN polymérase I est indiquée par une boîte bleue. Le promoteur est noté par un « P ». L'élément de pause T1 est indiqué en vert, le terminateur proximal T0 est indiqué en rouge. **(B)** Représentation schématique de deux unités successives d'ADN ribosomique de mammifère. La séquence transcrite par l'ARN polymérase I est indiquée par une boîte bleue. Le promoteur est schématisé par un « P ». Les éléments de pause T1 à T8 sont indiqués en vert, le terminateur proximal T0 est indiqué en rouge.

En plus d'interrompre la transcription par l'ARN polymérase I, les protéines Reb1 et TTF-I semblent activer la transcription des copies des ADNr adjacentes. Ces protéines se fixent non seulement sur le terminateur mais aussi sur une séquence d'environ 200 pb située en amont du promoteur, et nommée terminateur proximal T0. En effet, chez la souris et chez *S. cerevisiae* cette séquence semble être importante pour l'initiation de la transcription par l'ARN polymérase I.

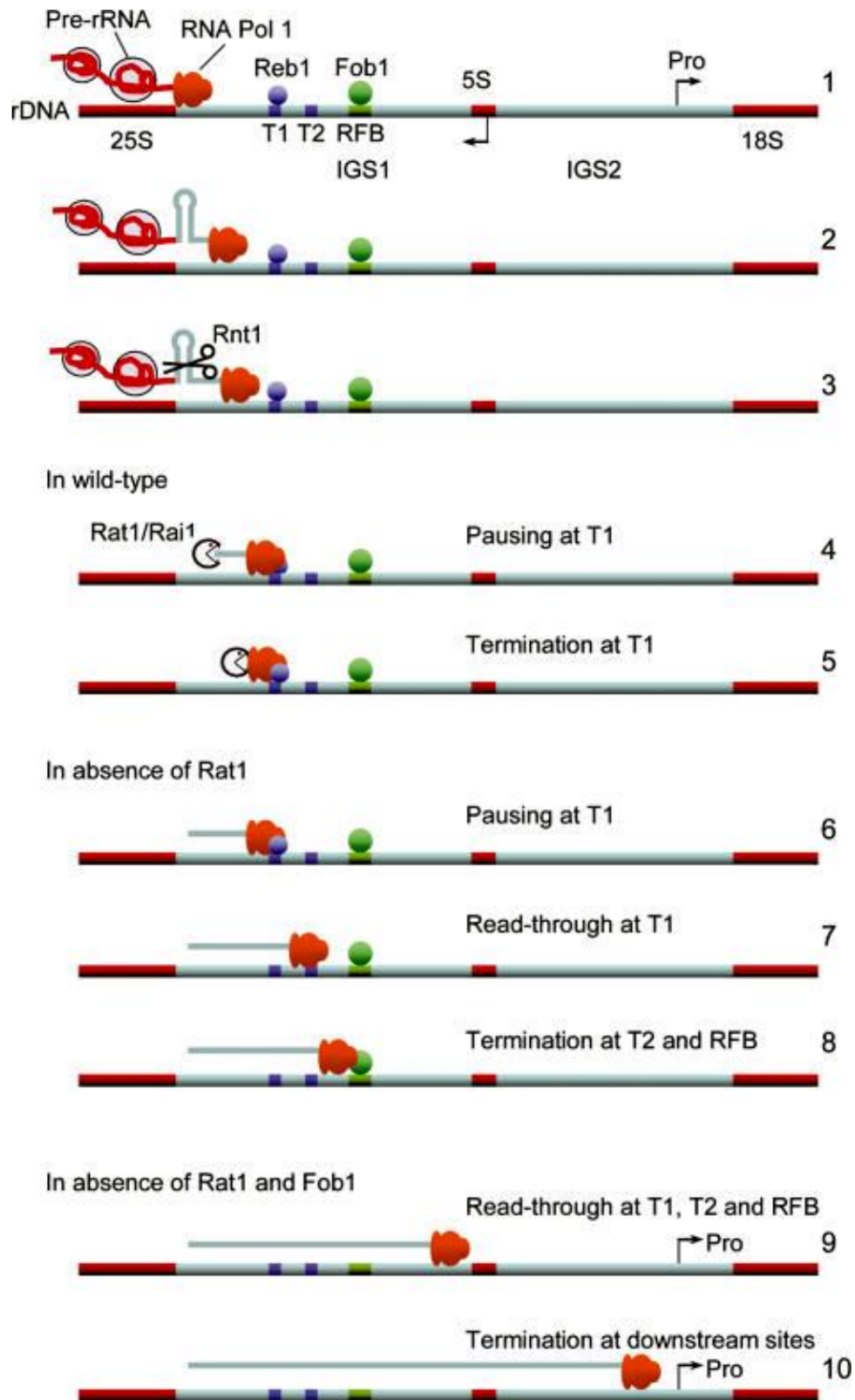
L'endonucléase Rnt1p est responsable de ce clivage co-transcriptionnel qui génère le pré-ARNr de 35S, avec une extrémité 3' variable. Après ce clivage, l'ARN Pol. I poursuit la transcription de l'ADNr (Figure 9).

Chez *S. cerevisiae*, le terminateur contient trois séquences notées (**Figure 9**) ; Reb1-dependent terminator (T1), fail-safe terminator (T2) et replication fork barrier (RFB) (RFB-binding protein Fob1).



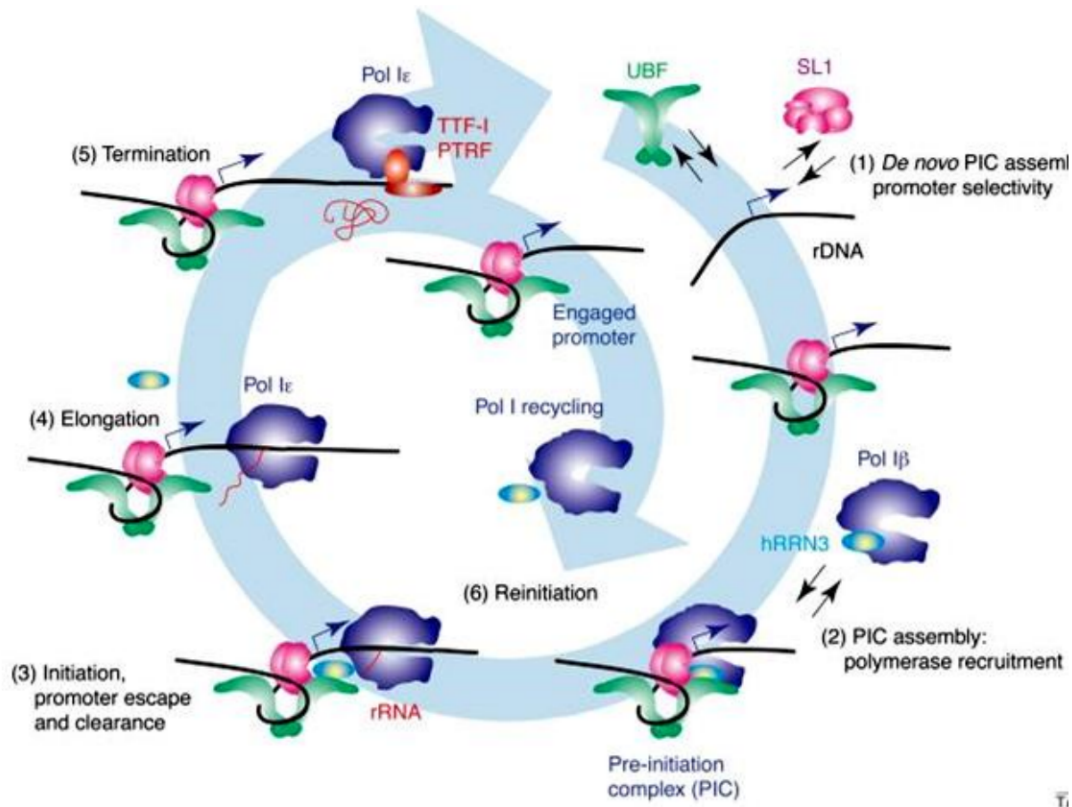
**Figure 9.** Les sites intergéniques entre ARN 25S et 5S (intergenic spacer regions IGS1) contiennent une séquence RFB (replication fork barrier) qui se positionne de +282 à +418 après la séquence codante d'ARN 25S. Le site majeur d'arrêt de la transcription est situé approximativement 93 nucléotides en aval de l'extrémité de gène l'ARNr de 25S (Site T1). Le second site de terminaison se trouve à 250 nt de l'extrémité de l'ARNr de 25S (Site T2). Reb1p s'associe à l'ADN aux positions T1 et T2.

Chaque unité de transcription formerait une boucle. Le transcrit naissant peut être clivé co-transcriptionnellement 14 nucléotides (nt) à 49nt en aval de l'extrémité 3' de l'ARNr de 25S (site de clivage de Rnt1p). L'endonucléase Rnt1p est responsable de ce clivage co-transcriptionnel qui génère le pré-ARNr de 35S, avec une extrémité 3' variable. Après ce clivage, l'ARN Pol. I poursuit la transcription de l'ADNr, l'ARN naissant dégradé par le complexe Rat1/Rai1 (deux sous-unités de Rat1p qui est une exonucléase 5'-3', requise pour la maturation de différents ARN non codants). L'ARN pol I pause au niveau de T1 à cause de la fixation du Reb1p. En absence de Reb1, l'arrêt de la transcription serait inefficace au site T, le T2 et RFB présentent des sites de secours (**Figure 10**).



**Figure 10.** Mode de terminaison d'ARN pol I chez *S. cerevisiae*.

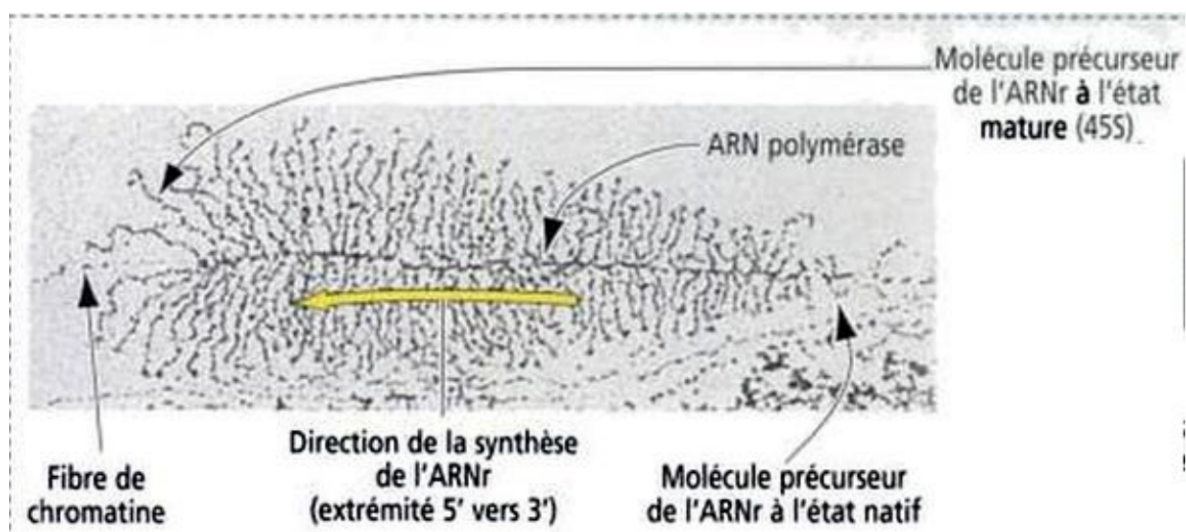
Les molécules d'ARN polymérase I terminant la transcription d'une unité d'ADNr pourraient alors être recrutées directement sur le promoteur de l'unité suivante (**Figure 11**).



**Figure 11.** Cycle de transcription du pré-ARNr par l'ARN pol II.

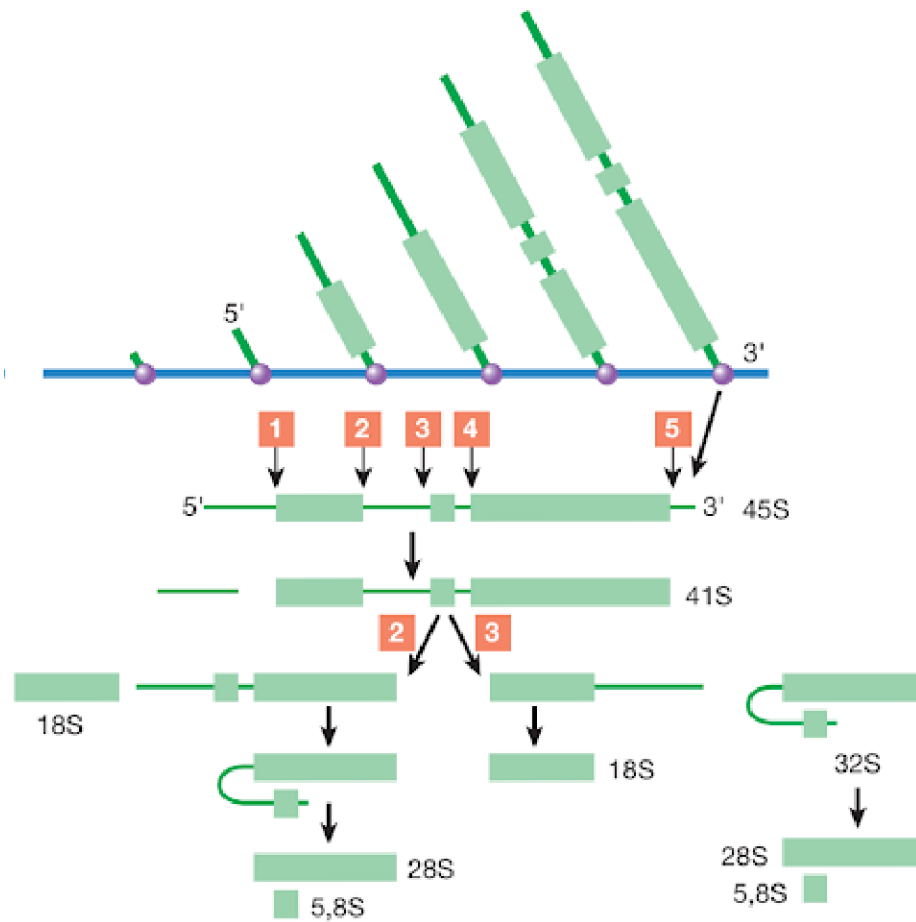
## 6. Assemblage co-transcriptionnel de la machinerie de maturation des ARNr

En 1969 Miller et Beatty ont observé chez le Xénope que les pré-ARNr en cours de transcription sont décorés en 5' de boules évoquant un arbre de Noël. Sur cette base, il a été proposé que les boules terminales représentent la machinerie de maturation qui assemble le pré ARN ribosomique en cours de transcription. La levure présente également ces arbres de Noël, et il a été récemment montré (Figure 12). La nature des molécules de ces boules terminales n'a été déterminée que 30 ans plus tard, renommées alors "SSU processome".



**Figure 12.** Processus de synthèse d'ARNr.

Les processome contiennent les snoARN, qui sont essentiels pour les étapes de clivage de pré-ARN (Figure 13).



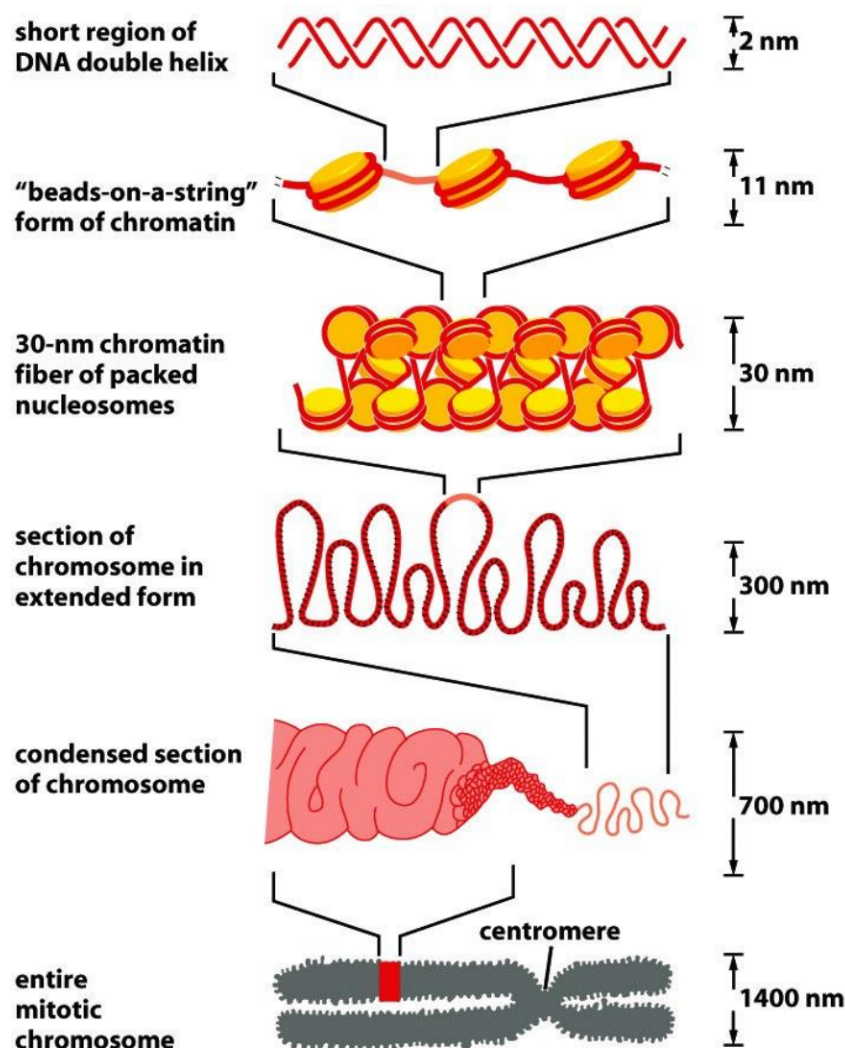
**Figure 13.** Représentation proposée pour la maturation du pré-ARNr.

# **Chapitre 3 : Régulation de remodelage de la chromatine**

## Régulation de remodelage de la chromatine

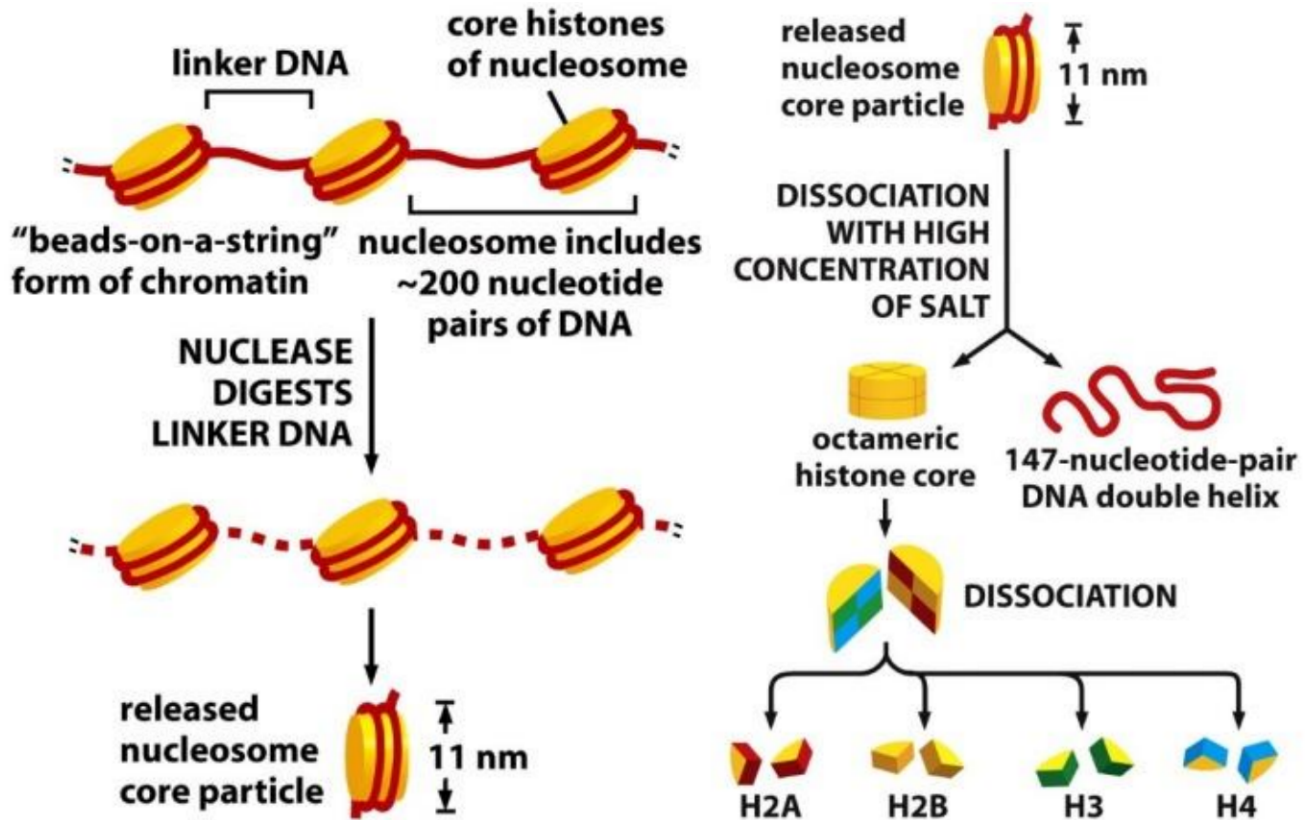
### 1. Rappel sur la chromatine

Le génome humain est constitué de  $3.3 \times 10^9$  bp (paires de bases), ce qui correspond à une longueur de 1.1 m. Cet ADN doit donc être replié et compacté pour qu'il puisse entrer dans l'espace très petit qui constitue le noyau cellulaire dont le diamètre vaut  $10\mu\text{m}$ . Ce repliement résulte de l'association de l'ADN avec des protéines structurales, les histones, donnant ainsi naissance à une structure hiérarchique appelée chromatine (Figure 1). Le premier niveau de repliement résulte de l'enroulement de l'ADN autour d'un cœur protéique pour former un nucléosome, qui est l'unité répétitive fondamentale de la chromatine. Cela raccourcit d'environ sept fois l'ADN nu. Dans un second temps, le cordon de nucléosomes se replie en un filament plus court et plus épais appelé fibre de 30 nm, environ 40 fois plus court que l'ADN nu. Cet élément est, à son tour, replié en une fibre de 100 à 300 nm de diamètre, qui forme vraisemblablement des domaines en boucles de 15000 à 100000 paires de bases, qui vont interagir avec les protéines qui constituent la matrice nucléaire.



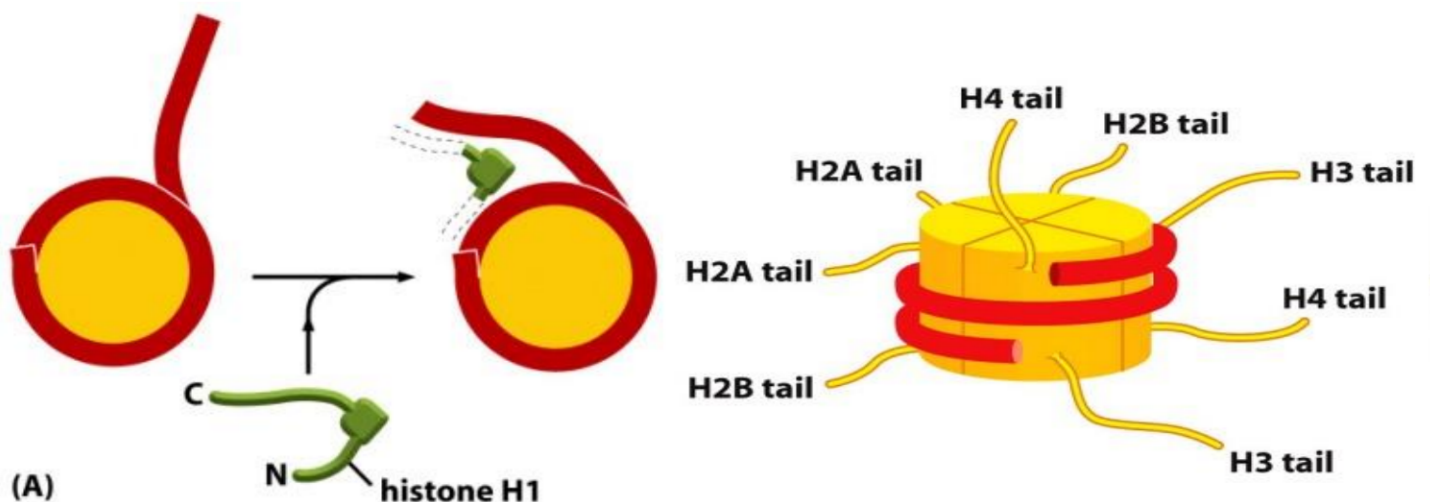
**Figure 1.** Condensation de génome des eucaryotes.

Les unités fondamentales de la chromatine, créent une barrière aux mécanismes nucléaires, tels que la transcription, en empêchant l'association de facteurs essentiels et de régulateurs gène-spécifiques au niveau de leurs séquences cibles sur les promoteurs. Les unités de base constituant la chromatine sont les nucléosomes, qui consistent en une séquence d'ADN génomique de 146 paires de bases enroulée autour de deux copies de chacune des protéines histone coeur H2A, H2B, H3 et H4. Les régions d'ADN situées entre deux nucléosomes adjacents sont occupées par une protéine histone linker H1.



**Figure 2.** Les unités de base constituant la chromatine

Les queues N-terminales des histones sont projetées vers l'extérieur de la structure où elles peuvent interagir avec des protéines régulatrices, avec l'ADN de liaison (ADN linker), situé entre deux nucléosomes, ainsi qu'avec l'ADN des nucléosomes avoisinants.

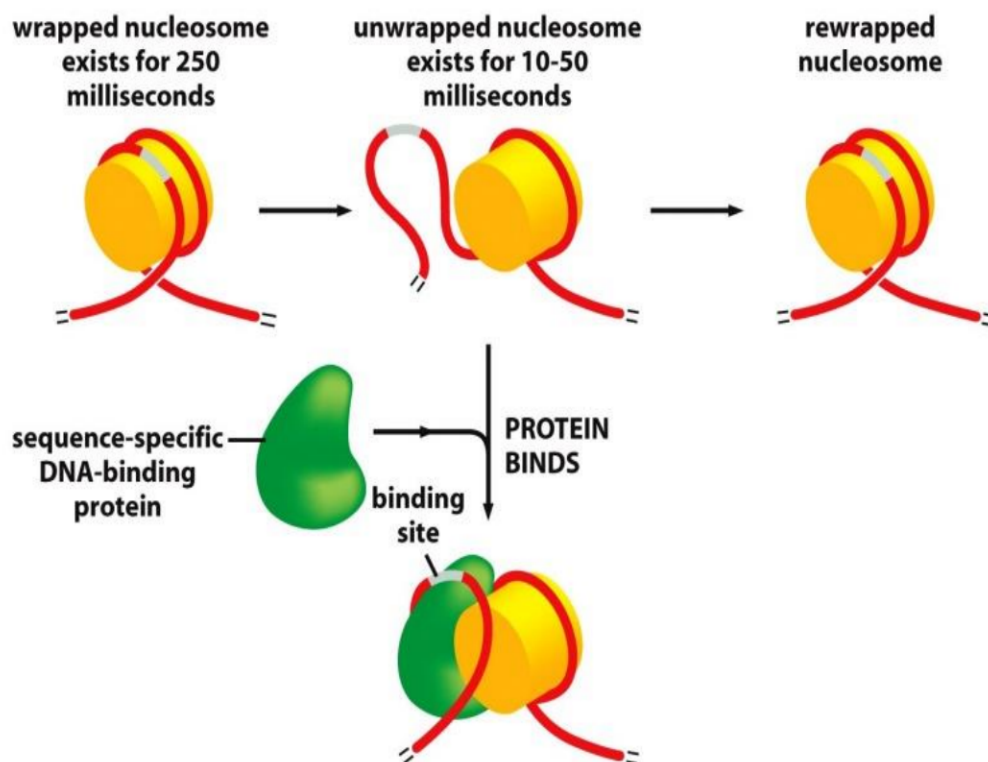


**Figure 3.** Structure de nucléosome

## 2. Régulation de remodelage de la chromatine

La régulation transcriptionnelle passe donc par la mise en place d'un équilibre dynamique entre l'empaquetage des séquences régulatrices au sein de la chromatine et l'accès à ces séquences par les régulateurs transcriptionnels. L'accès est donc restreint par les nucléosomes, mais ceux-ci peuvent être repositionnés, éjectés ou modifiés par des enzymes connues sous le nom de **complexes de remodelage ou de modification de la chromatine**.

Deux classes majeures de complexes régulent l'accessibilité des facteurs de transcription à leurs séquences cibles. *Les complexes de modification des histones* peuvent reconfigurer la structure des nucléosomes en modifiant de façon covalente les queues des histones pour réduire les interactions ADNprotéines ; tandis que *les complexes de remodelage de la chromatine ATP-dépendants* peuvent déplacer les nucléosomes en altérant les contacts ADN-histones, ce qui expose des séquences ADN cibles des facteurs de transcription.



**Figure 4.** Principe d'intervention des complexes de remodelage ou de modification de la chromatine.

## 3. Complexes des modifications covalentes des histones

Le domaine N-terminal ainsi que la région C-terminale des histones peuvent être le site de nombreuses modifications covalentes au niveau de certains de leurs résidus (**Figure 5**). Les histones peuvent être acétylées (lysines), méthylées (lysines, arginines), phosphorylées (sérines, thréonines), ubiquitinées (lysines), sumoylées (lysines) ou ADP-ribosylées (acides glutamiques). L'acétylation des histones par les histones acétyltransférases (HAT) est la mieux caractérisée de ces modifications.

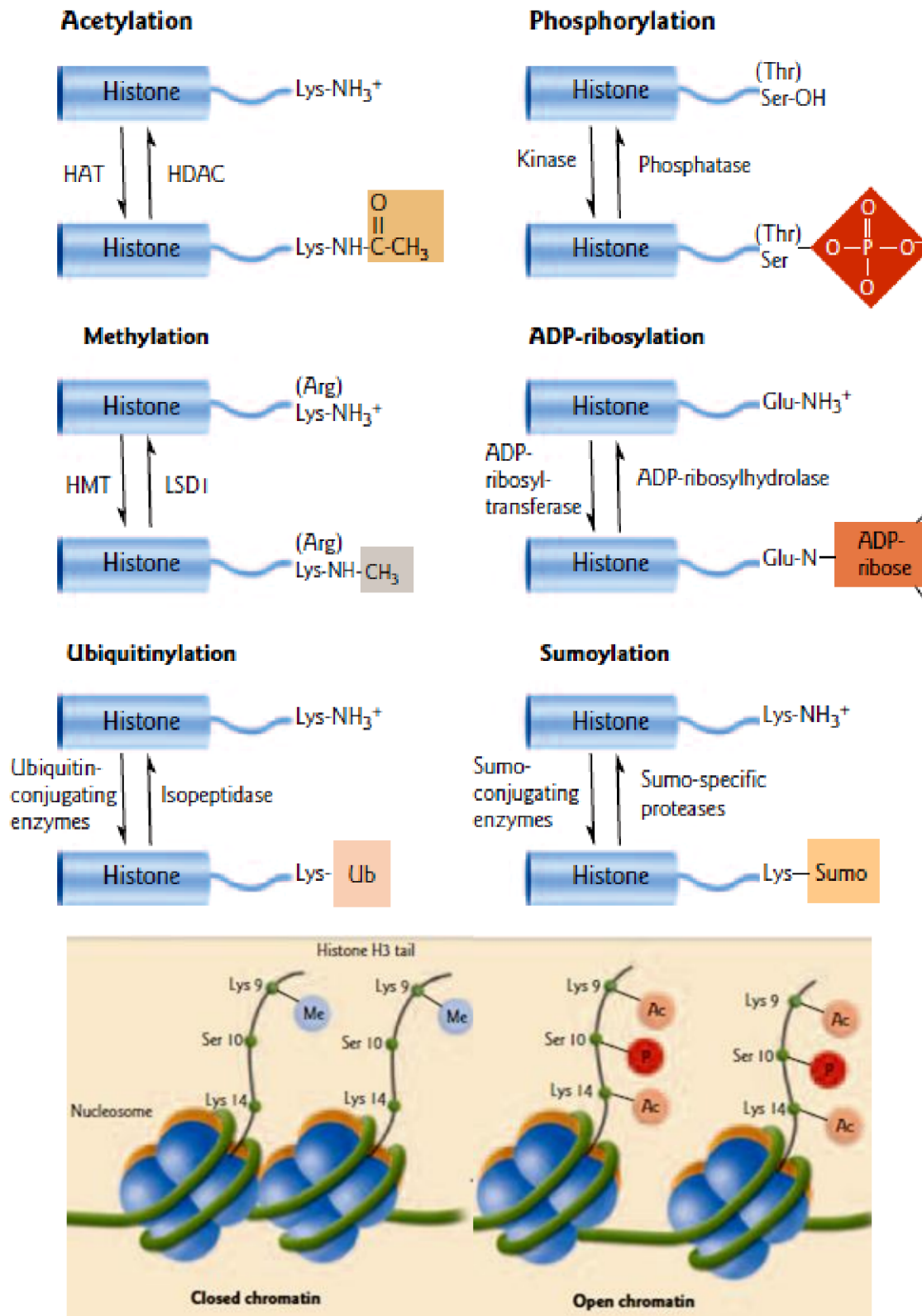


Figure 5. Modifications covalentes des histones.

### 3.1. Acétylation des histones

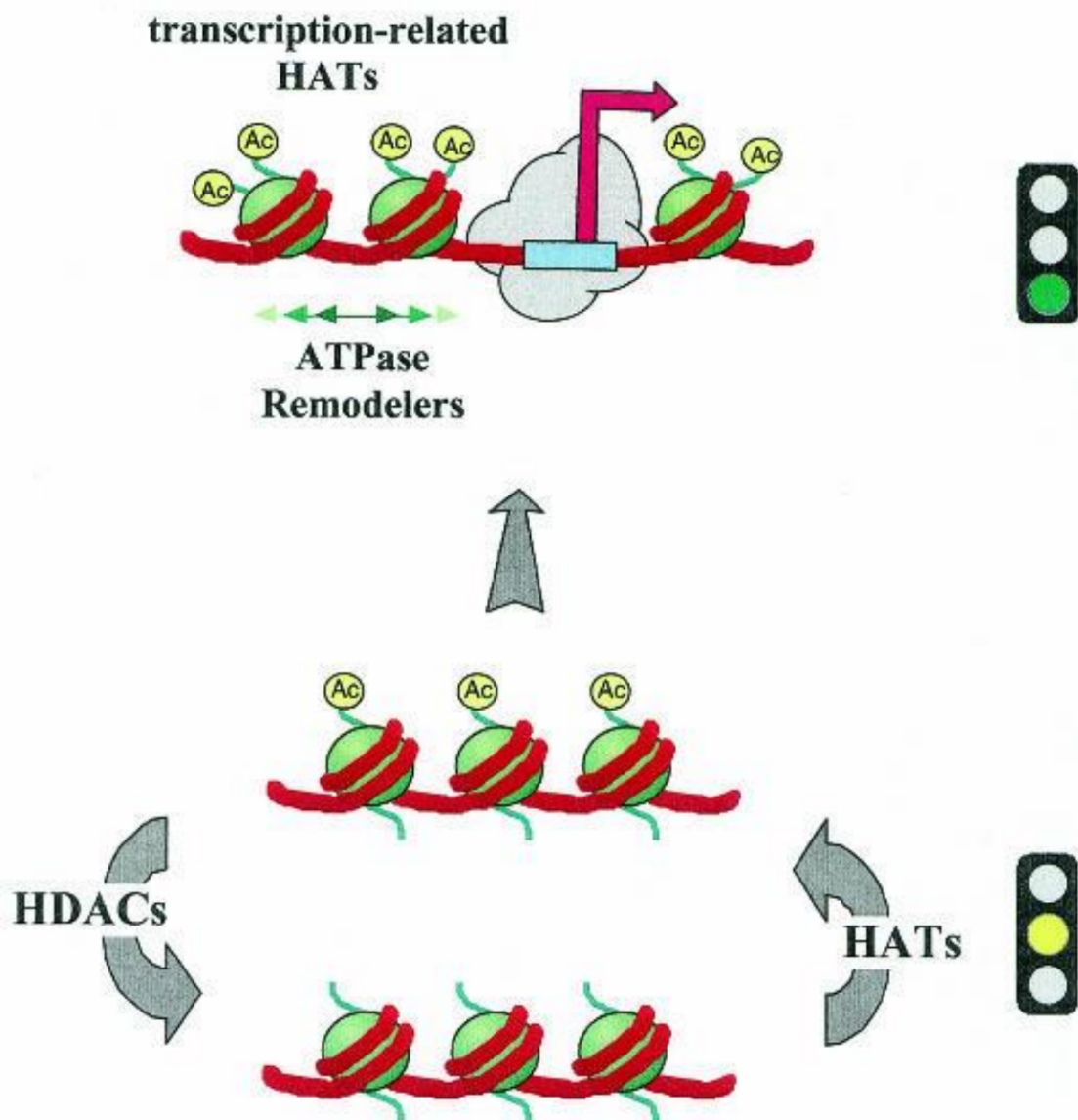
L'acétylation est considérée comme un mécanisme d'activation transcriptionnelle. Cependant, si l'on regarde plus en détail la constitution de la chromatine, on voit qu'il existe déjà un niveau basal d'acétylation, et donc l'acétylation ciblée envers certains résidus se fait dans un contexte global d'**acétylation** et **dé-acétylation** déjà présentes. En plus, l'acétylation n'est pas toujours synonyme d'activation de la transcription, vu que l'on trouve aussi des histones acétylés dans l'hétérochromatine. Réciproquement, la dé-acétylation d'un résidu ne provoque pas toujours la répression des gènes

concernés. Ceci suggère donc que l'effet sur la transcription dépend de la combinaison des différentes modifications effectuées au niveau de la chromatine.

L'acétylation est le transfert d'un groupement acétyle depuis l'acétyl coenzyme-A vers un résidu lysine de la protéine cible. Cette modification provoque une sensibilité accrue des nucléosomes à la digestion par des nucléases, ce qui indique que l'ADN est plus accessible. L'acétylation des histones réalisée par les histones acétyltransférases (HAT). Les HAT de type B ont une localisation cytoplasmique et acétylent les histones H4 avant leur entrée dans le noyau et leur dépôt sur l'ADN néosynthétisé. Les HAT de type A sont nucléaires et acétylent les histones déjà incorporées au sein de nucléosomes, ceci probablement en liaison avec les processus transcriptionnels.

Les HAT sont également capables d'acétyler d'autres substrats dont des protéines qui se fixent sur la chromatine et des facteurs de la transcription. L'acétylation de ces protéines régule leur activité. Ainsi, les HAT participent à la régulation de transcription à plusieurs niveaux.

L'acétylation est un processus qui peut être inversé par l'action de protéines appelées histones déacétylases (HDACs)

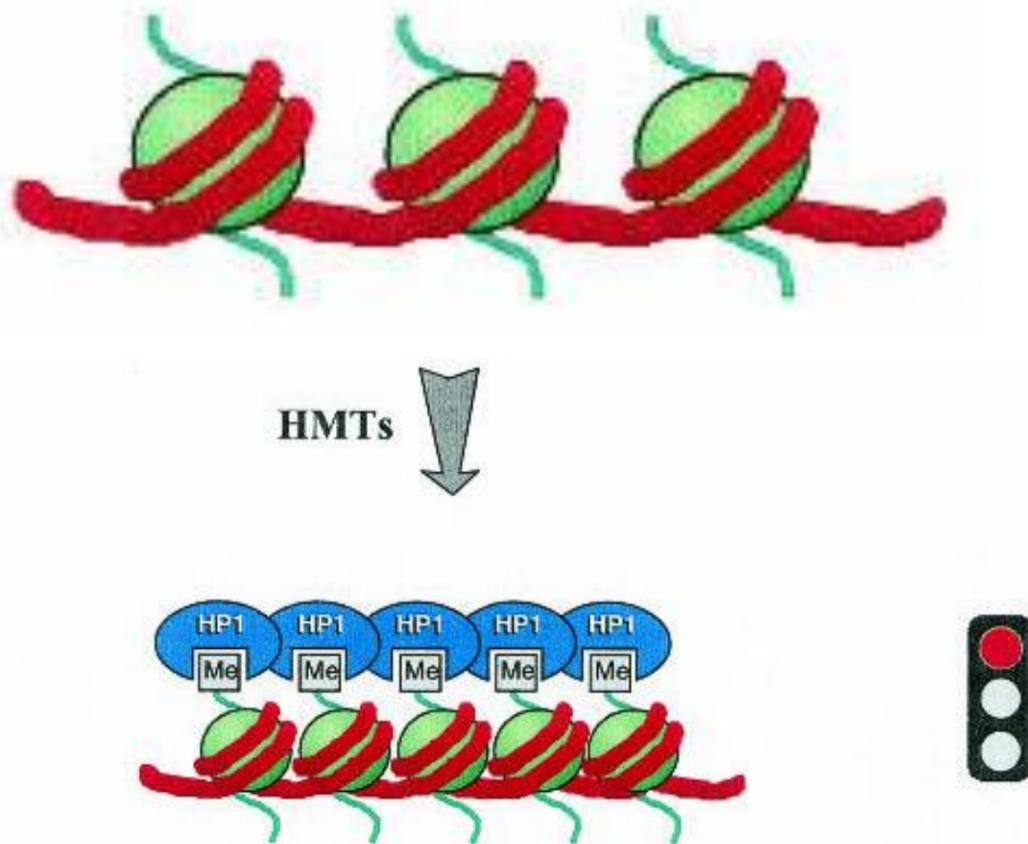


**Figure 6.** Effet d'acétylation sur la structure de chromatine.

### 3.1. Méthylation des histones

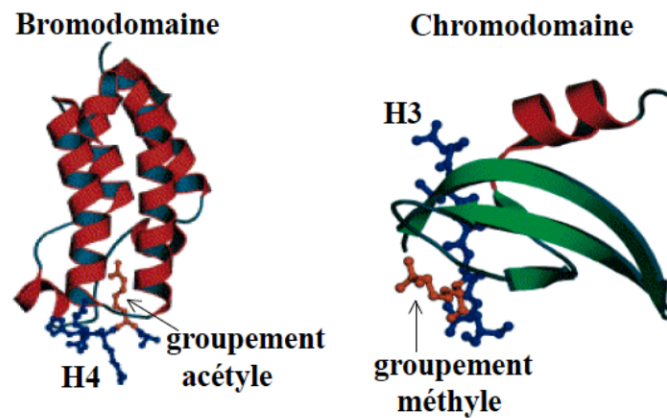
La méthylation est le transfert d'un groupement méthyle depuis la S-adenosyl méthionine vers la protéine cible. Les histones peuvent être méthylés par des histones méthyltransferases (HMT) au niveau des résidus lysine ou arginine. Contrairement à l'acétylation, la méthylation ne modifie pas la charge globale des histones. La méthylation des histones a été corrélée à la régulation de la transcription, cependant,

la méthylation, comme l'acétylation, n'a pas comme conséquence systématique une activation transcriptionnelle. Par exemple, la méthylation du résidu lysine 9 de l'histone H3 a été associée à la formation d'hétérochromatine (Figure 7). Mais par contre la méthylation des arginines dans les histones H3 et H4 a été corrélée avec un état actif de la transcription.



**Figure 7.** Exemple d'effet de méthylation du résidu lysine 9 de l'histone H3. Cette modification est par ailleurs reconnue par le répresseur transcriptionnel Hp1 chez la drosophile et les mammifères.

Les différentes modifications des histones constituent un signal de ciblage et d'ancrage de l'activité de remodelage vers certaines régions chromatinienne. L'acétylation et la méthylation des histones conduites au recrutement des motifs protéiques qui sont nommés les **bromodomaines** et les **chromodomaines**, respectivement (Figure 8).



**Figure 8.** Structure de bromodomaine et chromodomaine.

L'association des différentes modifications formerait alors un code, déchiffré par d'autres protéines, qui déterminerait la voie dans laquelle s'engage la cellule. Elle est soutenue par le fait qu'il existe des coopérations et/ou interférences entre les différentes modifications et le remodelage de la chromatine.

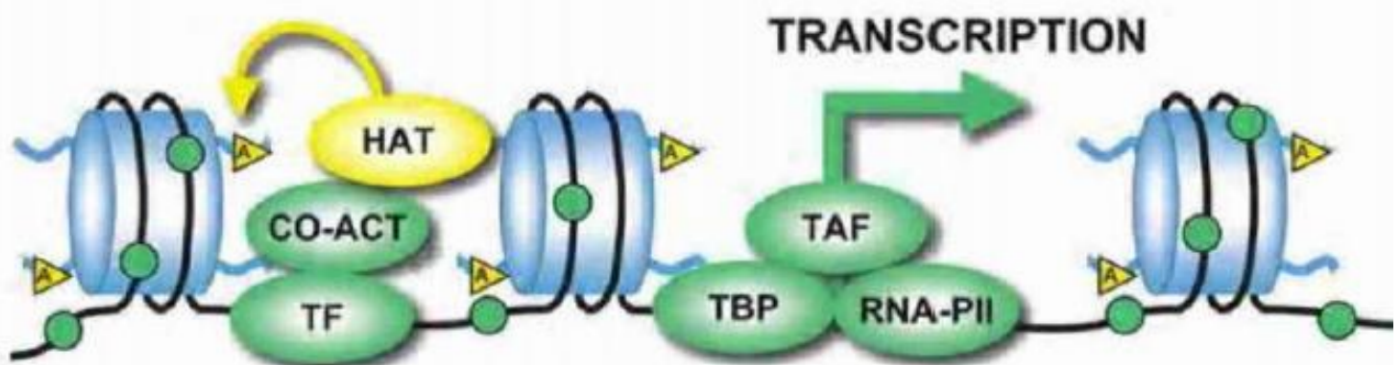
#### 4. Complexes de remodelage de la chromatine ATP-dépendants

La deuxième catégorie de mécanismes de remodelage de la structure chromatinienne met en jeu des complexes de remodelage spécialisés, qui agissent dans la réorganisation de la chromatine associée à des gènes particuliers. La structure de la chromatine réduit l'interaction de l'ADN avec les facteurs nucléaires. Par exemple, la protéine TBP (TATA box Binding Protein) peut s'accrocher à la séquence TATA de l'ADN nu, mais est incapable de s'accrocher sur la chromatine.

La création d'un ADN accessible, dans le contexte chromatinien, est un événement clé dans la régulation transcriptionnelle.

Ces complexes utilisent l'énergie d'hydrolyse de l'ATP pour induire des changements locaux dans l'arrangement des nucléosomes et dans l'enroulement de l'ADN au sein de la chromatine. Les mécanismes de modifications covalentes des histones et de remodelage des nucléosomes ne sont pas indépendants. Les gènes de l'endonucléase HO et de l'interféron- $\beta$  sont donnés à titre d'exemple pour montrer l'effet de la coopération entre les modifications covalentes des histones et les facteurs de remodelage.

Les facteurs de remodelage de la chromatine ont été étudiés généralement comme facteurs d'**activation** génétique. En réalité, ils peuvent aussi faciliter la **répression** génétique (coactivateurs et corépresseurs).



**Figure 8.** la coopération entre les modifications covalentes des histones et les facteurs de remodelage.

#### 4.1. Diversité des facteurs de remodelage

Le terme de remodelage recouvre un ensemble de changements, peu ou pas définis sur le plan moléculaire, qui affectent les nucléosomes. Ces changements d'ordre conformationnel sont catalysés par des complexes multiprotéiques appelés facteurs de remodelage. Ces facteurs ont tous en commun une sous-unité catalytique ATPase ayant un domaine hélicase.

L'activité ATPase permet au complexe de modifier la structure nucléosomique, en partie grâce à l'énergie libérée par l'hydrolyse de l'ATP. Les facteurs de remodelage ayant des sous-unités ATPase identiques ou très similaires ont été classés, ici, dans une même famille (Tableau 1).

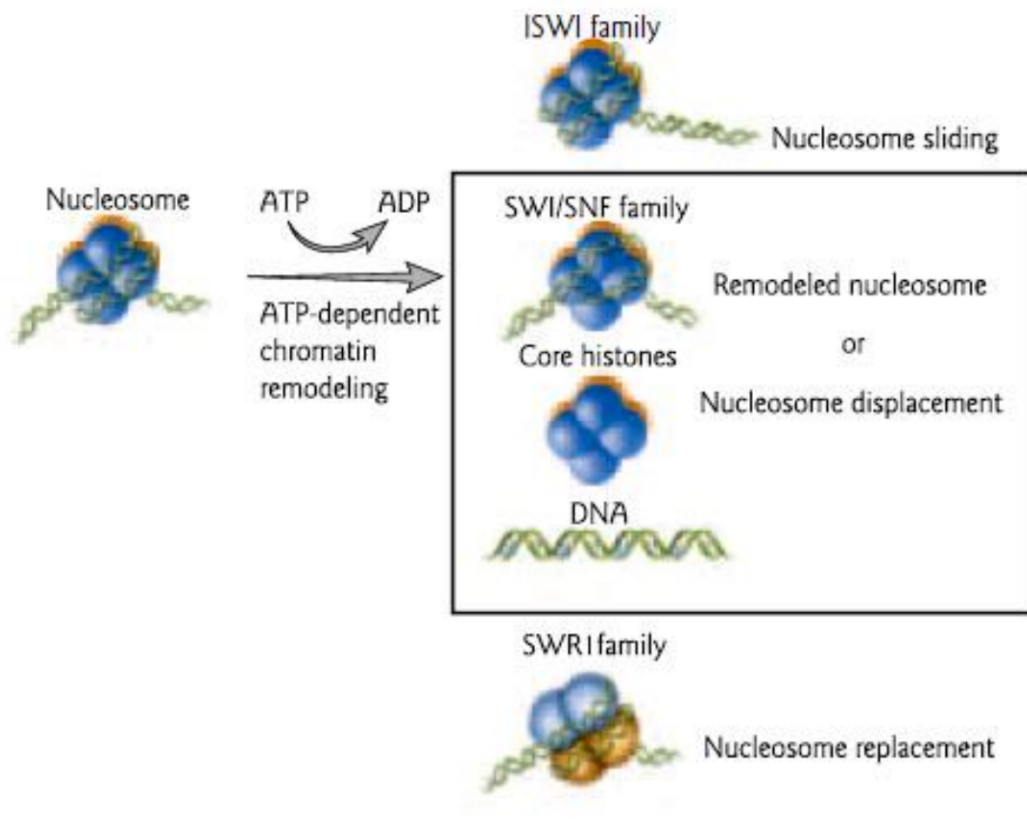
L'un des complexes de remodelage de la chromatine les mieux caractérisés est le complexe SWI/SNF. Le complexe SWI/SNF a été identifié

**Tableau 1.** Exemples des familles des complexes de remodelage de la chromatine.

Famille	Nom	Abréviation	espèce	SU
	<b>SWI/SNF</b>	<i>Mating-type switching/sucrose non-fermenting</i>	<i>S. cerevisiae</i>	11
	<b>RSC</b>	<i>Remodels the structure of chromatin</i>	<i>S. cerevisiae</i>	15
	<b>complexe Brahma</b>		<i>D. melanogaster</i>	6 à 9
<b>SWI-SNF</b>	<b>complexe BRM</b>	Brahma	<i>H. sapiens</i>	9 à 12
	<b>ISW1</b>	<i>Imitation of switch 1</i>	<i>S. cerevisiae</i>	4
	<b>ISW2</b>	<i>Imitation of switch 2</i>	<i>S. cerevisiae</i>	2
	<b>NURF</b>	<i>Nucleosome remodeling factor</i>	<i>D. melanogaster</i>	4
	<b>ACF</b>	<i>ATP-utilizing chromatin assembly and remodeling factor</i>	<i>D. melanogaster</i>	4
	<b>CHRAC I</b>	<i>Chromatin accessibility complex</i>	<i>D. melanogaster</i>	5
<b>ISWI</b>	<b>RSF</b>	<i>Remodeling and spacing factor</i>	<i>H. sapiens</i>	2
	<b>NuRD</b>	<i>Nucleosome remodeling histone deacetylase complex</i>	<i>H. sapiens</i>	7
	<b>NRD</b>	<i>Nucleosome remodeling and deacetylating</i>	<i>H. sapiens</i>	6
<b>Mi2/NuRD</b>	<b>NURD</b>	<i>Nucleosome remodeling and histone deacetylation</i>	<i>H. sapiens</i>	18

#### 4.2. Mode d'action des complexes de remodelage de la chromatine

Elles peuvent soit repositionner les nucléosomes par un **transfert** en *trans* vers une autre molécule d'ADN (par exemple de l'ADNb à l'ADNa), ou en *cis*, en faisant glisser les nucléosomes en avant ou en arrière de leur position d'origine. Le **désassemblage/réassemblage** des nucléosomes, par des facteurs, peut faciliter la transcription d'une matrice de chromatine en retirant un dimère d'histone H2A/H2B du nucléosome.



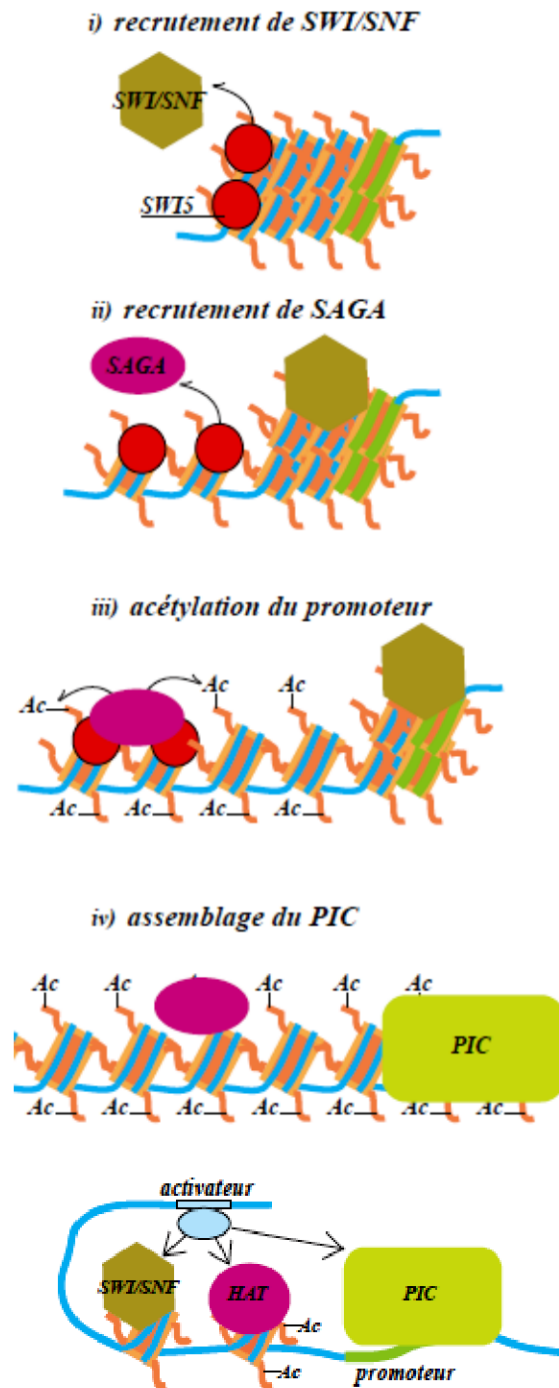
**Figure 9.** Mode d'action des complexes de remodelage de la chromatine.

## 5. Coopération entre les modifications covalentes des histones/les facteurs de remodelage/ activateurs ou répresseur

Les facteurs de remodelage de la chromatine peuvent être adressés vers des loci spécifiques par des interactions avec d'autres protéines qui régulent les gènes, ou par des épitopes situés sur des histones. On va citer deux modes d'action pour cette coopération ; activation de transcription par la décondensation de la chromatine et répression par condensation de la chromatine.

### 5.1. Promoteur d'endonucléase HO

L'activateur SWI5 recrute le complexe de remodelage SWI/SNF qui commence à remodeler en rendant plus accessible le promoteur. Ensuite, SWI5 recrute le complexe SAGA, qui acétyle les queues des histones, en rendant le promoteur plus accessible et en favorisant ainsi l'assemblage du complexe de pré-initiation de la transcription (PIC).



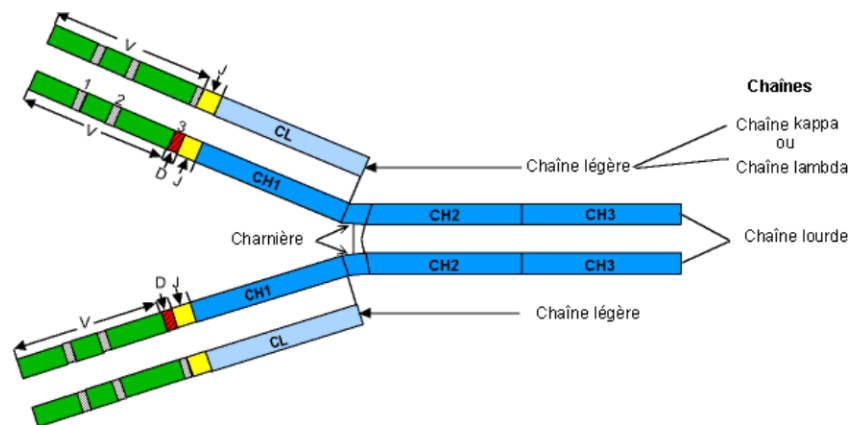
**Figure 10.** Mode de collaboration entre le complexe SWI/SNF et le complexe SAGA.

# Réarrangement des gènes des immunoglobulines (Recombinaison V(D)J)

## 1. Rappel sur les immunoglobulines

Les immunoglobulines (Ig) sont des glycoprotéines hétérodimériques constituées de deux chaînes lourdes IgH (H pour « Heavy ») identiques reliées entre elles par un ou plusieurs ponts disulfures et de deux chaînes légères IgL (L pour « Light ») identiques, chacune liée à une chaîne lourde par un pont disulfure (Figure1), l'ensemble forme un complexe protéique d'environ 150kDa.

Il existe deux types de chaînes légères, kappa ( $\kappa$ ) et *lambda* ( $\lambda$ ) et cinq classes de chaînes lourdes à savoir : *mu* ( $\mu$ ), *delta* ( $\delta$ ), *gamma* ( $\gamma$ ), *alpha* ( $\alpha$ ) et *epsilon* ( $\epsilon$ ), correspondant respectivement aux monomères d'IgM, IgD, IgG, IgA et IgE. Les Ig possèdent un domaine variable (**V**) qui assure la liaison à l'antigène (Ag) et un domaine constant (**C**) qui confère les fonctions effectrices telles que l'activation du complément ou la liaison aux récepteurs Fc. Les chaînes lourdes comportent, selon l'isotype, trois à quatre domaines CH : (CH1, CH2, CH3 et CH4).



**Figure 1. Représentation bi-dimensionnelle d'une Ig.** VH, VL: régions variables des chaînes lourdes et légères. CH, CL: régions constantes des chaînes lourdes et légères, respectivement.

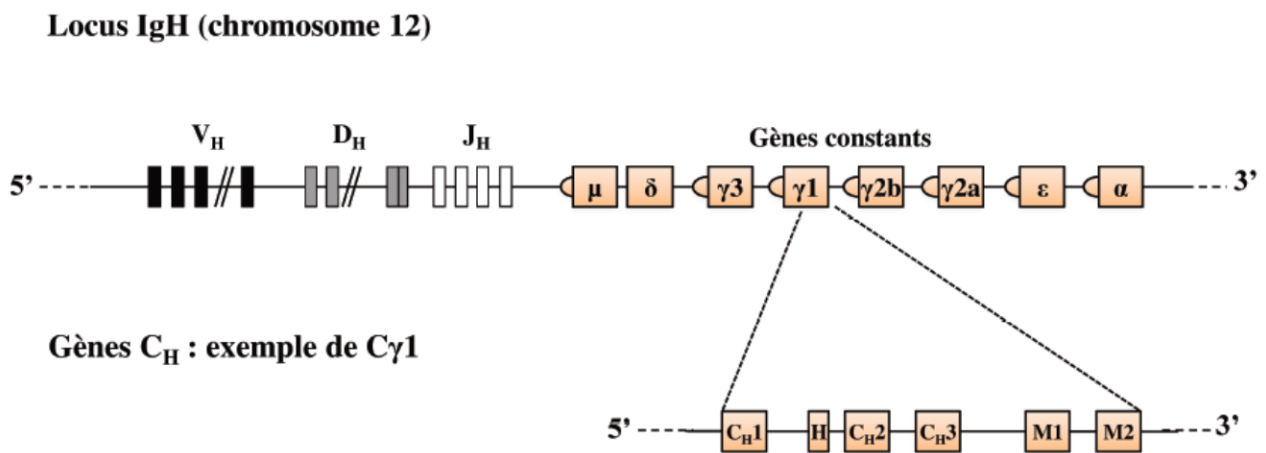
## 2. Organisation des loci d'Immunoglobulines

Chez les eucaryotes les gènes d'Ig sont regroupés en un nombre des loci distincts. A titre d'exemple, la souris comme chez l'homme, trois loci localisés sur trois chromosomes. On distingue deux loci codant les chaînes légères, le locus *Igk* et le locus *Igλ*, et un codant les chaînes lourdes, le locus *IgH*.

### a. Locus de chaînes lourdes (IgH)

Le locus des chaînes lourdes IgH murin (Figure 2) est situé sur le chromosome 12 (chromosome 14 chez l'homme), où il s'étend sur environ 2,8 Mégabases (Mb) divisé en 4 régions principales contenant les familles de gènes  $V_H$ ,  $J_H$ ,  $C_H$  et  $D_H$  (segments dits de " diversité " propres au locus de chaînes lourdes). Le nombre des segments de chaque famille est variable selon l'espèce et même la souche.

Exp. murine : plus de 200 segments  $V_H$  (dont une trentaine de pseudogènes), une douzaine de segments  $D_H$  et un groupe de quatre segments  $J_H$  tous fonctionnels. En 3' des segments  $V_DJ$ , on trouve huit gènes codant les régions constantes répartis selon l'ordre suivant  $\mu$ ,  $\delta$ ,  $\gamma 3$ ,  $\gamma 1$ ,  $\gamma 2a$ ,  $\gamma 2b$ ,  $\epsilon$  et  $\alpha$ . Chaque gène constant est composé de multiples exons codant les domaines structuraux propres à chaque chaîne lourde ainsi que les régions charnières pour les IgG. Enfin les exons de membrane (M1 et M2) codent les régions trans-membranaires et intracytoplasmiques des BCRs (B-cell receptor).



**Figure 2.** Configuration germinale de locus d'IgH murin.

### b. Locus de chaînes légères Kappa (Ig $\kappa$ )

Le locus kappa (Figure 3A) est situé sur le chromosome 6 de la souris (chromosome 2 chez l'homme). Il est légèrement plus grand que le locus IgH et s'étend sur environ 3,2 Mb. Dans sa configuration germinale, il comporte 161 segments  $V_{\kappa}$  dont 91 à 92 sont fonctionnels. Les gènes  $J_{\kappa}$  sont au nombre de 5 dont un non fonctionnel. La région constante est codée par un seul segment  $C_{\kappa}$ .

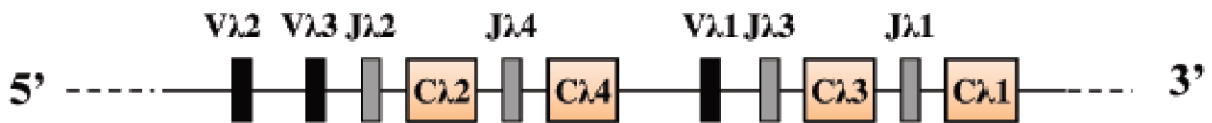
### c. Locus de chaînes légères lambda ( $\lambda$ )

Le locus lambda murin (Ig  $\lambda$ ), situé sur le chromosome 16 (chromosome 22 chez l'homme), est assez petit de 200 Kilobases (Kb) et présente une diversité limitée (Figure 3B). Il possède une structure particulière, et est constitué de trois segments  $V\lambda$  et de 4 familles comprenant chacune une paire de segment  $J\lambda$  et  $C\lambda$  (noté chacun de 1 à 4). Il est organisé en 2 clusters ordonnés :  $V\lambda 2-V\lambda 3-J\lambda 2-C\lambda 2-J\lambda 4-C\lambda 4$  et  $V\lambda 1-J\lambda 3-C\lambda 3-J\lambda 1-C\lambda 1$ .

### A. Locus $Ig\kappa$ (chromosome 6)



### B. Locus $Ig\lambda$ (chromosome 16)



**Figure 3.** Configuration germinale de locus d' $Ig\kappa$  et  $Ig\lambda$  murin.

## 6. Production et génération de la diversité des Ig

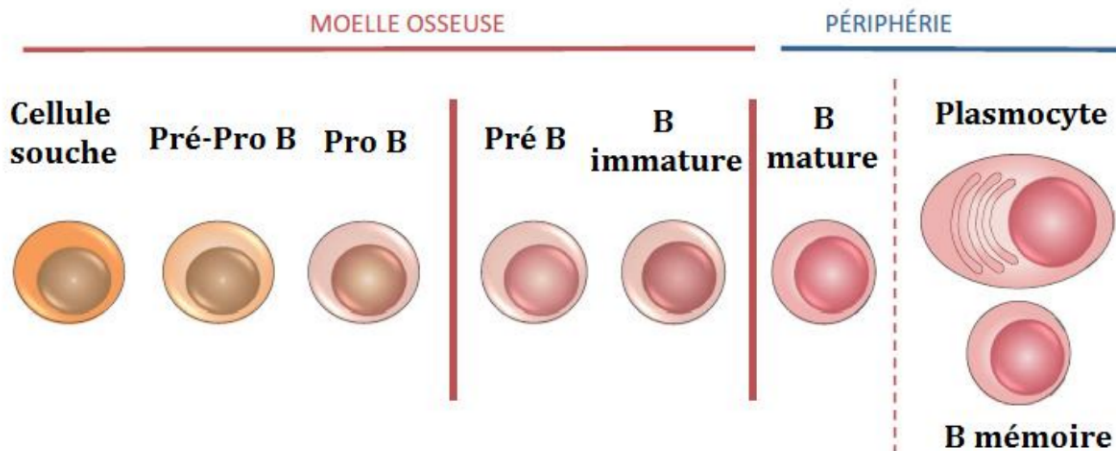
L'immunité adaptative ou acquise est une réponse spécifique de l'antigène assurée par les lymphocytes B (LB). Elle repose sur la reconnaissance d'un déterminant antigénique (épitope) par les récepteurs des cellules B (BCR). Au cours de la réponse adaptative, les LB interviennent dans l'immunité humorale. La réponse à médiation humorale implique d'abord la forme membranaire des immunoglobulines (Ig) au sein du BCR et aboutit à la sécrétion d'Ig solubles ou anticorps (Ac) par les plasmocytes (i.e. stade de différenciation terminale des cellules B). Ces molécules sont capables de lier spécifiquement un antigène (Ag) dans le but de le neutraliser ou de l'éliminer.

Une longue série de phénomènes d'activation, de coopération cellulaire et de réarrangements intra-géniques est nécessaire avant d'aboutir à la sécrétion clonale d'un

type d'Ac, spécifique d'un antigène, par une cellule plasmocytaire. En effet, les gènes codant les Ig ne sont pas fonctionnels dans leur configuration germinale.

Les premiers réarrangements ont lieu au cours du développement B précoce (Figure 4) et cela indépendamment de la rencontre avec l'antigène. Ce processus appelé recombinaisons V(D)J va assembler les régions variables des chaînes lourdes et légères d'Ig et permettre à chaque LB d'exprimer un BCR unique.

D'autres remaniements géniques se déroulent dans les organes lymphoïdes secondaires où les LB activés par l'Ag vont pouvoir améliorer l'affinité et/ou changer les propriétés effectrices de leur BCR grâce aux mécanismes d'hypermutations somatiques (SHM pour Somatic Hypermutation) et de commutation de classe (CSR pour Class switch Recombination).



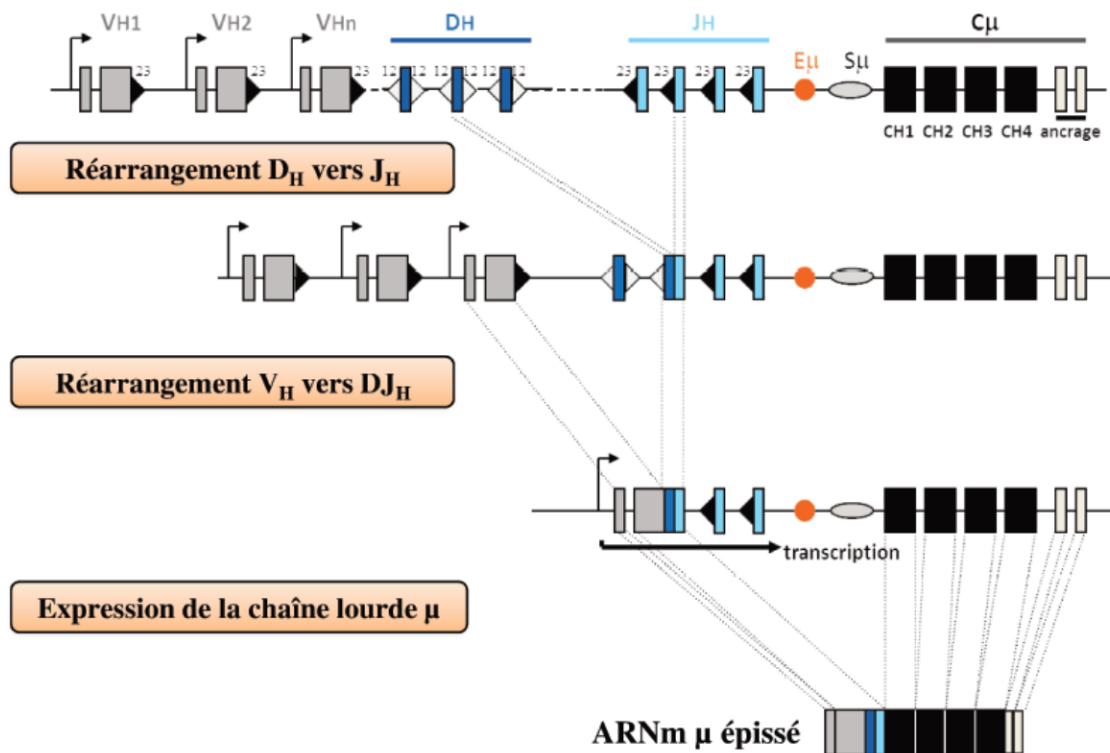
**Figure 4.** Ontogénèse des LB.

### 7. Recombinaisons V(D)J des gènes Ig

Les réarrangements géniques débutent, dans les précurseurs des cellules B, par le locus des gènes des chaînes lourdes. La synthèse d'une chaîne lourde nécessite l'intervention de quatre types de gènes : un segment VH (variable), un segment DH de diversité, un segment JH de jonction et un segment CH qui code la partie constante.

Les étapes du réarrangement des gènes codant pour un domaine variable de chaîne lourde (VH). Il se fait en trois étapes : choix d'un gène DH et d'un gène JH ; choix d'un gène VH ; génération d'un ARN pré-messager (pré-ARNm) à partir de la séquence VDJ-domaine

constant ainsi constituée sur le chromosome 14 réarrangé. La synthèse protéique d'une chaîne lourde  $\mu$  se fera après épissage de ce pré-ARNm (Figure 5).



**Figure 5.** Réarrangement des loci des chaînes lourdes d'Ig.

Les régions variables des chaînes légères sont générées par une unique étape de jonction des segments VL et JL pour former un exon VJ. Les régions d'ADN comprises entre les différents segments sont déléetées lors de ces réarrangements sous forme d'un ADN circulaire (cercle d'excision ou épisome).

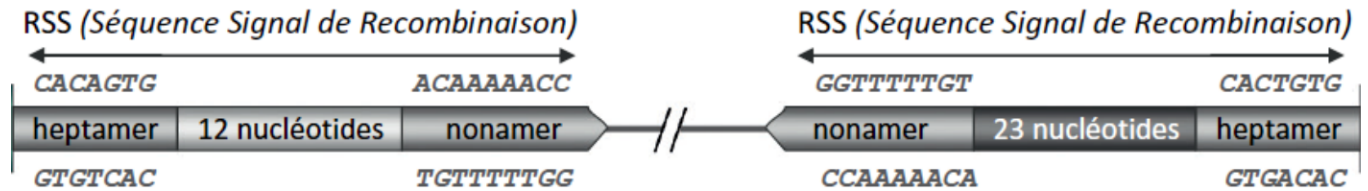
Le passage de la forme membranaire à la forme sécrétée des immunoglobulines s'effectue par épissage alternatif d'un même transcrit primaire (pré-ARNm) de chaîne lourde permettant l'élimination des séquences codant les parties transmembranaire et cytoplasmique. Le grand nombre de segments V, D et J disponibles et les multiples combinaisons possibles entre ces éléments constituent la base de la diversité combinatoire.

### 8. Mécanisme de recombinaisons V(D)J des gènes IgH

Les réarrangements entre les éléments V, D et J sont réalisés grâce à de courtes séquences cibles, appelées RSS (Recombination Signal Sequence ou Séquence Signal de

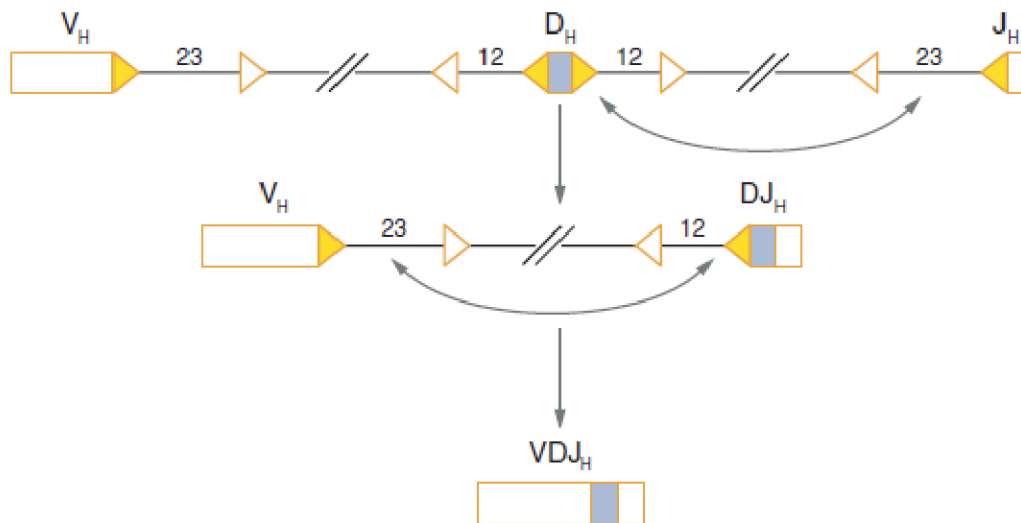
Recombinaison), situées immédiatement en 3' des segments V, en 5' et 3' des segments D, et en 5' des segments J.

Les RSS sont composées de deux motifs très conservés (un heptamère et un nonamère) séparés par un « espaceur » de 12 ou 23 nucléotides (figure 6).



**Figure 6.** RSS.

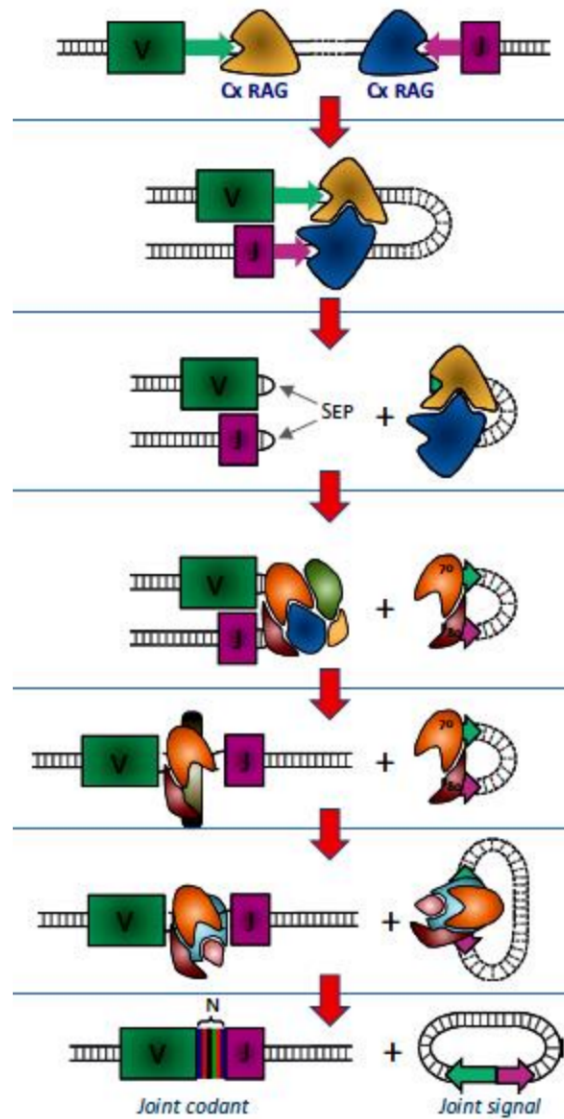
La longueur des espaceurs tient une place importante dans ce processus puisqu' un réarrangement ne peut subvenir qu' entre deux segments dont les RSS possèdent des espaceurs de longueurs différentes (i.e entre une RSS-12 et une RSS-23) selon la règle dite « 12/23 » (Figure 7).



**Figure 7.** Réarrangement V(D)J via RSS (12/23).

Les réarrangements VDJ sont provoqués par l'activité des protéines RAG-1 et RAG-2 (recombination activating gene), qui sont capables de reconnaître les SSR et d'engendrer des coupures d'ADN double brin entre les SSR et les segments de gène V, D et J. Elles permettent le clivage double brin de l'ADN au niveau de la séquence RSS, et la formation d'une structure en épingle à cheveux (SEP).

L'expression des gènes RAG-1 et RAG-2 est strictement contrôlée lors du développement lymphocytaire B permettant les réarrangements d'abord au niveau du locus IgH, puis au niveau des gènes de chaînes légères. La résolution des cassures double brin de l'ADN générées par le complexe RAG est assurée par un système ubiquitaire de réparation de l'ADN appelé NHEJ (Non-Homologous End- Joining). Ce complexe multi-enzymatique permet la jonction d'extrémités non homologues, par l'ouverture de la structure en épingle à cheveux et la ligation des segments codants et non codants de l'ADN (Figure 8).



**Figure 8.** Mécanisme d'action des RAG1 et RAG2.

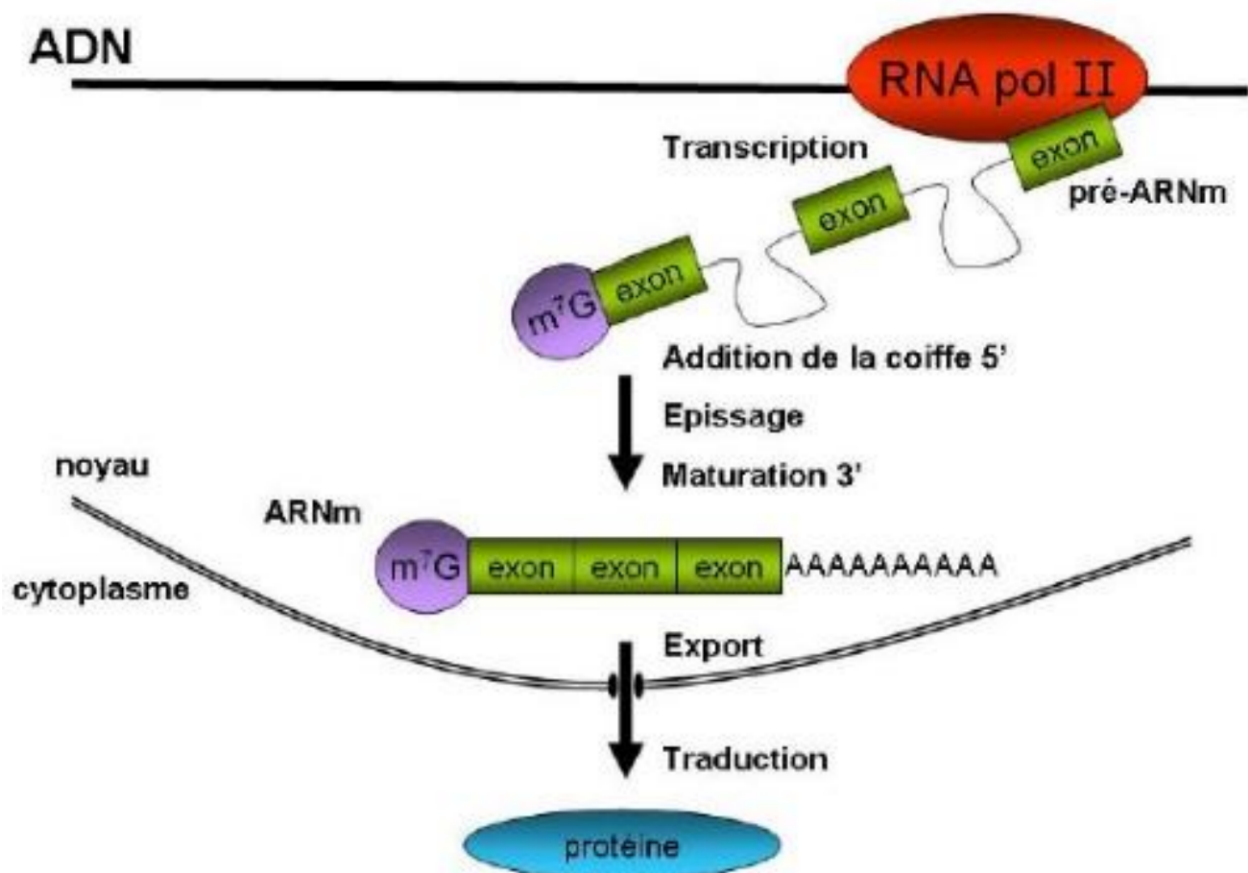
# **Chapitre 4 : Régulation Post-transcriptionnelle**

## Régulation Post-transcriptionnelle

### I. Polyadénylation d'ARNm

#### 1. Généralité sur la maturation d'ARNm

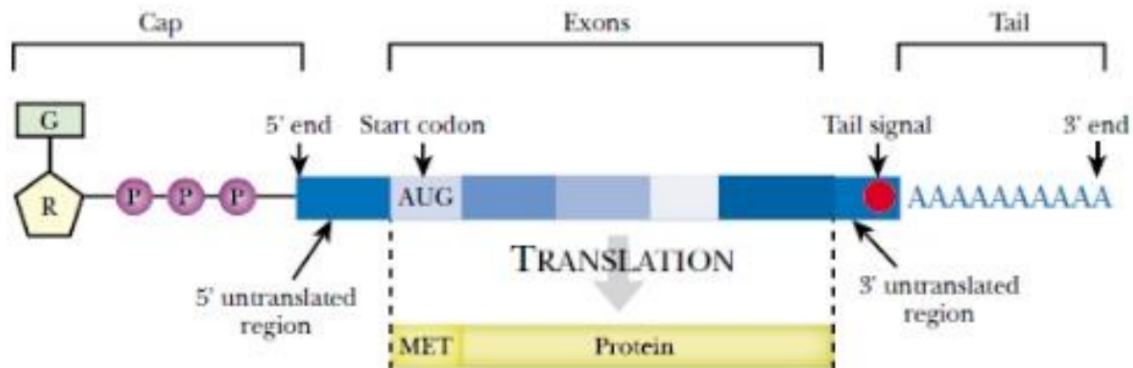
Chez les eucaryotes supérieurs, l'expression d'un gène débute par la transcription par l'ARN polymérase II d'un ARN immature appelé pré-ARN messager (pré-ARNm) ou, historiquement, hnRNA (« heterogeneous nuclear RNAs »), du fait de leur taille hétérogène et de leur localisation nucléaire. Ce pré-ARNm est maturé dans le noyau en 3 étapes **co-transcriptionnelles** (Figure 1). L'ARN néo-transcrit est déphosphorylé dans sa partie 5' par une triphosphatase, une guanine peut alors être transférée sur l'extrémité 5' du pré-ARNm. Cette guanine est méthylée par la guanine-7-méthyltransférase pour former la structure  $m^7G(5')ppp(5')N$  (N représentant un des quatre nucléotides), appelée **coiffe**.



**Figure 1. Etapes de maturation d'ARNm au cours d'expression de gène.**

Le pré-ARNm est alors épissé, ce qui permet l'excision des introns grâce à l'assemblage d'un complexe ribonucléoprotéique appelé « spliceosome ». La maturation de l'extrémité 3' (aussi

appelée polyadénylation) comprend le clivage dans la partie 3' terminale du pré-ARNm et la polyadénylation proprement dite (addition d'une queue poly(A) contenant 150 à 300 résidus d'Adénosines). L'ARNm mature est alors exporté vers le cytoplasme où il est traduit (Figure 2).



**Figure 2.** Schéma d'ARNm mature.

A l'exception des gènes des histones métazoaires dont l'expression est répléon dépendante, tous les ARNm subissent une maturation 3' qui génère une queue poly(A) à l'extrémité 3' de chaque transcrit. La queue poly(A) peut être de différentes tailles (60 à 250 nt) selon l'organisme et l'ARNm.

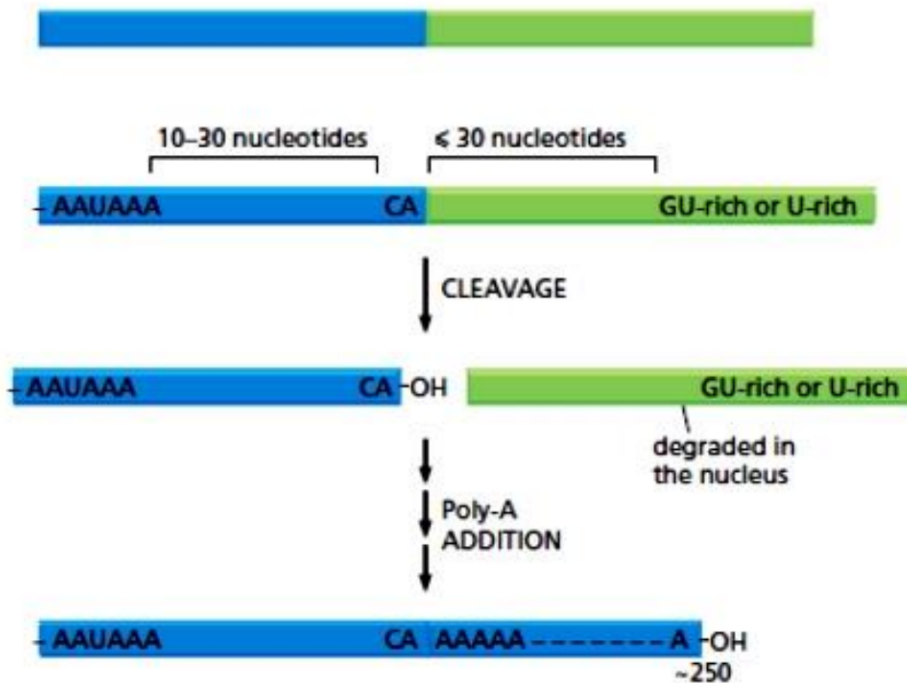
La maturation 3' promeut la terminaison de la transcription, confère sa stabilité à l'ARNm, facilite l'export et la traduction. La queue poly(A) est ajoutée par une réaction de clivage et de polyadénylation catalysée par la machinerie de maturation 3', un complexe multi-protéique qui reconnaît le signal poly(A) sur le transcrit naissant. Bien que ces réactions puissent être séparées biochimiquement *in vitro*, les événements de maturation du pré-ARNm sont interconnectés les uns avec les autres et avec la transcription dans la cellule, mais ils sont aussi liés à des événements postérieurs comme la traduction ou la dégradation de l'ARNm dans le cytoplasme.

De plus, le choix alternatif du site de la maturation en 3' permet d'accroître la diversité du transcriptome et du protéome.

## 2. Mécanisme de polyadénylation

### 2.1. Séquence permettant la polyadénylation

La maturation en 3' se fait suite à la reconnaissance de différentes séquences parmi lesquelles le signal de polyadénylation et le DSE (« Downstream Sequence Element ») qui permettent la polyadénylation en aval du site de clivage (Figure 3).



**Figure 3.** Eléments de séquence permettant la polyadénylation.

- A. Signal de polyadénylation :** c'est un hexamère ou hexanucléotide constitué de la séquence AAUAAA. Cette séquence située entre 10 et 30 nucléotides en amont du site de clivage.
- B. Downstream Sequence Element :** cet élément est localisé à plus de 30 nucléotides en aval du site de clivage. Il n'est plus présent sur l'ARN clivé et n'est donc pas impliqué dans l'étape de la polyadénylation. Le DSE est moins conservé que l'hexanucléotide. Il peut être soit riche en uraciles (UUUUU), soit riche en guanines et uraciles YGUGUUY (Y=pyrimidine).
- C. Site de clivage :** Le site de clivage se situe généralement entre 10 et 30 nucléotides en aval de l'hexamère et à 30 nucléotides, ou plus, en amont du DSE. Le site de clivage n'est pas strictement conservé. L'ordre de préférence pour le nucléotide du site de clivage est A>U>C>>G. Chez les vertébrés, 70% des sites de clivage sont localisés en 3' d'un résidu adénosine. De ce fait, le premier résidu d'adénosine de la queue poly(A)

est généralement un nucléotide transcrit. Le nucléotide précédant le site de clivage est une cytosine dans 59% des cas. Le site de clivage est donc le plus souvent constitué du dinucléotide CA et le clivage a généralement lieu après le résidu adénosine.

## 2.2. Machinerie de polyadénylation

La machinerie de maturation en 3' requiert un large complexe composé de facteurs multi-protéiques (Tableau 1).

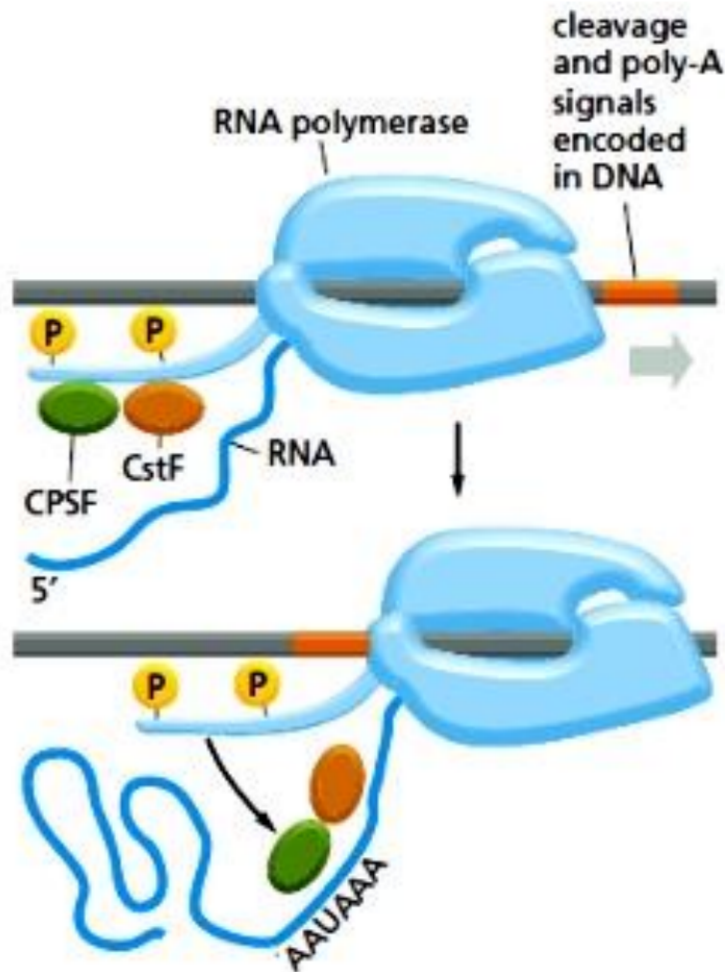
**Tableau 1.** Machinerie de polyadénylation.

Facteur	Nomination	Quelques Rôles
CPSF	Cleavage and Polyadenylation Specificity Factor	Reconnaît la séquence AAUAAA, ou riche U. Stabilise le complexe de maturation en 3' Endoribonucléase responsable du clivage en 3'
CstF	Cleavage stimulation Factor	Reconnaît les séquences riches en G/U Stabilise le complexe de maturation en 3'
CFIm § CFIIIm	Cleavage Factor I et II	CFIm Aide à reconnaître la séquence AAUAAA CFIIIm Dissocie les facteurs de maturation en 3' de la pol II et Permet une interaction entre CPSF et CFIm et recrute de l'ATP
PAP	Poly(A) Polymérase	Polymérise la queue poly(A) sur le pré-ARNm clivé
PABPII	Poly(A) Binding Protein	Permet une polyadénylation processive
CTD Pol II	Partie carboxy-terminale de la RNA polymérase II	Permet de recruter plusieurs facteurs de maturation en 3'

## 2.3. Mode d'action de la maturation en 3'

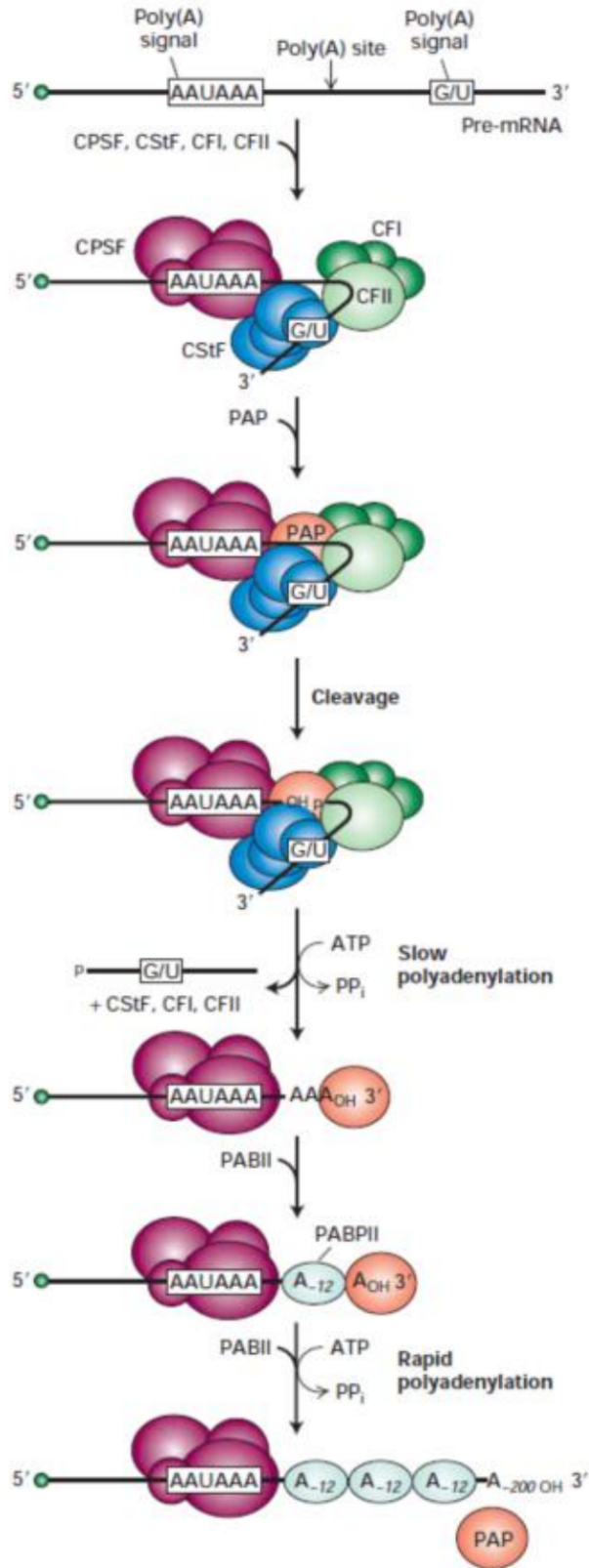
L'ordre réel de l'arrivée de chaque facteur sur le pré-ARNm n'est pas clairement établi. Néanmoins, le complexe de maturation en 3' est recruté à proximité du site de clivage grâce à son interaction avec la partie CTD pol II. Les facteurs CFIm et CPSF sont tous deux associés

précocement sur le pré-ARNm (Figure 4). L'assemblage serait ensuite rapidement complété par la fixation de CstF sur le pré-ARNm, et de CFIm sur le complexe de maturation en 3'. Le CPSF reconnaît le AAUAAA alors que CstF-64 reconnaît le DES. Les interactions entre CPSF, CstF et le pré-ARNm sont stabilisées par les autres sous-unités de ces facteurs.



**Figure 4.** Recrutement des facteurs de polyadénylation sur les séquences.

C'est l'interaction entre CPSF et CstF qui définit la position du site de clivage et permet le recrutement de la PAP. Le complexe ainsi formé est positionné de part et d'autre du site de clivage et permet à CPSF de catalyser le clivage endonucléotidique du pré-ARNm, mais la réaction de clivage nécessite le recrutement des autres facteurs de polyadénylation (Figure 5).



**Figure 5.** Mécanisme d'action de la machinerie de polyadénylation.

### 3. Polyadénylation alternatives

De nombreux signaux poly(A) aient été identifiés dans des endroits variés de la région transcrite. Les signaux poly(A) putatifs sont fréquemment retrouvés dans les régions riches en A-U telles que les introns, les régions non traduites (UTR) et les régions intergéniques. Il existe trois types de polyadénylation alternative qui sont définies en fonction de l'organisation des sites d'épissage et des signaux de maturation en 3' (Figure 6).

**Classe I :** correspond à la présence de plusieurs signaux de polyadénylation disposés en tandem sur le dernier exon en 3'. Ce type de polyadénylation alternative génère des ARNm avec des 3'UTR de différentes tailles entraînant différentes régulations post-transcriptionnelles (transport, stabilité, localisation, traduction...).

**Classe II :** correspond à des exons 3' terminaux alternatifs mutuellement exclusifs qui sont en compétition pour leur utilisation dans un ARNm. Ce type de polyadénylation alternative est couplé à un épissage alternatif et génère des ARNm contenant des 3'UTR différents et codant des protéines qui diffèrent dans leur région C-terminale. Leur sélection peut être régulée par la reconnaissance des sites 3' d'épissage et/ou des signaux de maturation en 3'.

**Classe III :** est également couplée à un épissage alternatif entraînant des signaux de maturation en 3' mutuellement exclusifs, mais un des exons possède un site 5' d'épissage interne. Cet exon appelé exon

composite interne 3' terminal peut être, comme son nom l'indique, utilisé en tant qu'exon interne quand son site 5' d'épissage est reconnu, ou comme exon terminal quand son signal de polyadénylation est sélectionné. La régulation de l'utilisation de ces exons fait donc intervenir une compétition entre le site 5' d'épissage interne et le signal de polyadénylation.

## II. Epissage d'ARNm

### 1. Généralité sur l'épissage d'ARNm

Chez les eucaryotes, les gènes nucléaires codant les protéines comportent pour la plupart des séquences codantes (exons) interrompues par des séquences non codantes (introns). Des séquences introniques sont également présentes dans des gènes codant des ARN stables (ARN de transfert, ARN ribosomiques), dans des gènes d'organites cellulaires (chloroplaste, mitochondrie), mais le mécanisme d'élimination de ces introns est différent de celui utilisé pour les gènes nucléaires eucaryotes. Les transcrits des gènes nucléaires sont produits par l'ARN polymérase II et sont soumis à des processus complexes de maturation, ceci pour

convertir les pré-ARN messagers (pré-ARNm) en ARN messagers (ARNm) fonctionnels. L'épissage correspond à l'excision des introns du transcrit primaire puis à la jonction des exons. Il est dit **constitutif** lorsqu'un exon est systématiquement inséré dans le transcrit mature. A l'inverse, l'épissage **alternatif** permet de produire des protéines différentes à partir d'un seul gène, ceci du fait de l'existence de plusieurs sites potentiels d'épissage en compétition les uns avec les autres au sein du pré-ARNm. Certains exons seront choisis et inclus dans l'ARNm, alors que d'autres ne le seront pas. Environ 75 % des événements d'épissage alternatif ont pour conséquence de modifier la séquence codante des protéines, ce qui conduit à la production de protéines distinctes et seulement 19 % conduiraient à la production de protéines tronquées. Ce processus contribue grandement à la diversité du protéome et fournit une part d'explication au différentiel entre le nombre de gènes et celui des protéines, l'épissage alternatif constitue un mode particulier de régulation de l'expression des gènes. Ainsi, un gène donné est en général capable de coder différentes protéines, en fonction du stade du développement, l'état de différenciation cellulaire ou l'état physiologique. Pour atteindre cette flexibilité, la sélection des sites d'épissage chez les eucaryotes supérieurs a évolué pour pouvoir dépendre de multiples paramètres comme la force du site d'épissage, la présence ou l'absence de régulateurs d'épissage, la structure secondaire de l'ARN, l'architecture exon/intron et le processus de synthèse du pré-ARNm.

## 2. Mécanisme d'épissage

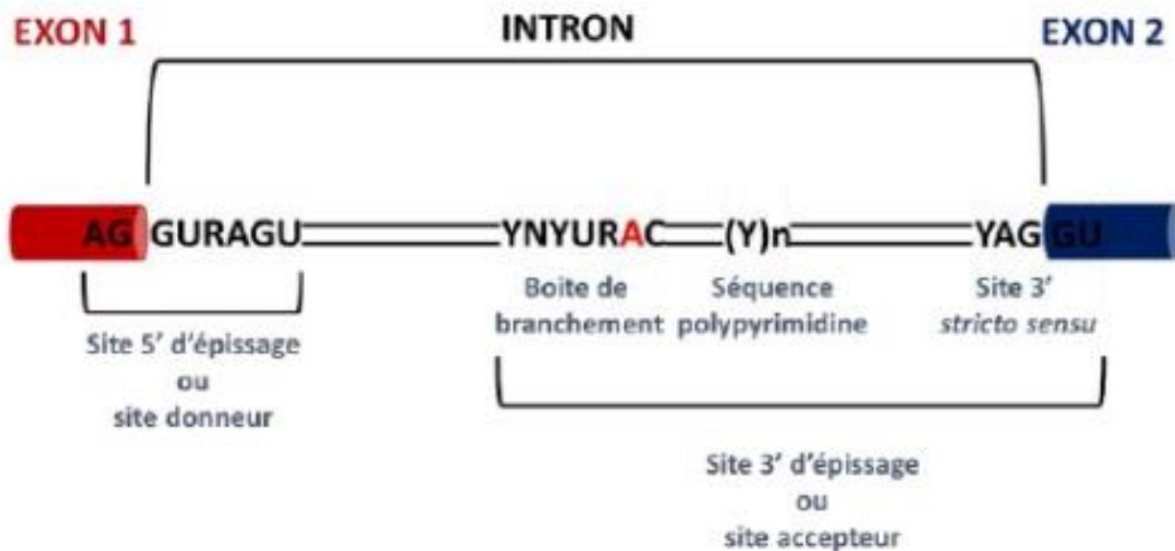
L'épissage des pré-ARNm nucléaires est un processus crucial pour l'expression des gènes. La reconnaissance des introns par le spliceosome est par conséquent un événement clé pour la synthèse d'un ARNm fonctionnel. Cette reconnaissance des séquences introniques repose sur la présence de **séquences fortement conservées** au cours de l'évolution qui sont situées aux extrémités des introns. Ces séquences sont appelées site **5' d'épissage** (ou site donneur) et site **3' d'épissage** (ou site accepteur).

L'élimination des introns est un processus très complexe qui nécessite l'assemblage d'un macrocomplexe, le spliceosome, composé de 5 petits ARN nucléaires snRNA (U1, U2, U4, U5 et U6) et d'un très grand nombre de protéines (plus de 100 protéines différentes) autour de l'intron à éliminer. Il existe en fait 2 types d'introns chez les eucaryotes supérieurs : les introns dits majeurs, éliminés par le spliceosome précédent par la suite et les introns dits mineurs, rencontrés moins fréquemment. Ces introns sont éliminés par un autre type de

spliceosome, renfermant les snRNA U11 et U12 à la place des snRNA U1 et U2, les snRNA U4atac et U6atac à la place des snRNA U4 et U6, et le snRNA U5.

### 2.1. Séquences définissant les introns

Des séquences consensus de ces sites chez les vertébrés ont pu être déterminées par comparaison de séquences (Figure 7). La séquence consensus au niveau du site 5' donneur d'épissage est AG/GURAGU, où "/" représente la jonction exon/intron et **R**, un résidu purine. Les deux nucléotides GU à l'extrémité 5' de l'intron sont les plus conservés et sont très importants pour l'efficacité d'épissage. Le site 3' d'épissage est constitué de 3 éléments, dont chacun est essentiel aux processus de reconnaissance de l'intron et d'épissage. Il comprend, tout comme le site 5', une séquence située à la jonction de l'intron et de l'exon (site d'épissage *stricto sensu*) dont la séquence consensus est YAG/GU, où "/" représente la jonction exon/intron et **Y**, un résidu pyrimidine. Enfin, le site accepteur comporte aussi une séquence polypyrimidine, localisée entre le site 3' stricto sensu et la boîte de branchement, qui renferme de long enchainements de pyrimidine (10 à 20 résidus C ou U).



**Figure 7.** Représentation schématique de l'organisation d'un intron situé entre deux exons. Les séquences conservées jouant un rôle essentiel dans la réaction d'épissage sont représentées en gras (Y pour pyrimidine, R pour purine, N pour un des 4 résidus A, C, G ou T).

Il est à noter que les introns mineurs se caractérisent également par la présence de sites 5' et 3' d'épissage, mais les séquences consensus de ces sites sont différentes. En effet, le site 5' est caractérisé par une séquence consensus /RUAUCCU, le site 3' stricto sensu par (YA(C ou G)/N) et la boîte de branchement par UCCUUAACU (où le point de branchement correspond au résidu adénosine indiqué en gras).

L'ensemble de ces séquences (site 5' et 3' d'épissage) permet la reconnaissance des introns par le spliceosome lors du processus d'épissage. L'efficacité de cette reconnaissance et, par conséquent, de la réaction d'épissage diminue en fonction de la dégradation des séquences, et est à la base de l'épissage alternatif, une reconnaissance moins efficace d'un exon alternatif, laissant place à une régulation plus fine pour favoriser l'inclusion ou l'exclusion de l'exon.

### 2.2. Réaction d'épissage

L'épissage des pré-ARNm nucléaires implique deux étapes successives de **trans-estérification** qui permettent d'éliminer les introns (Figure 8). La première étape débute par l'attaque nucléophile du **phosphate en 5'** de l'intron (site 5' donneur) par **l'hydroxyle en 2'** du résidu adénosine de la boîte de branchement (point de branchement). Cette réaction génère deux intermédiaires réactionnels : l'exon avec un hydroxyle libre en 3' et un intermédiaire intron-exon dans lequel le nucléotide du point de branchement est impliqué dans trois liaisons phosphodiester, formant ainsi une structure ayant la forme d'un lasso. Lors de la deuxième étape de trans-estérification, l'hydroxyle en 3' de l'exon, libéré lors de la première étape, réalise une attaque nucléophile sur le phosphate en 3' de l'intron (site 3' stricto sensu). Cette seconde étape conduit à la ligation des deux exons et à la libération de l'intron sous forme d'un lasso.

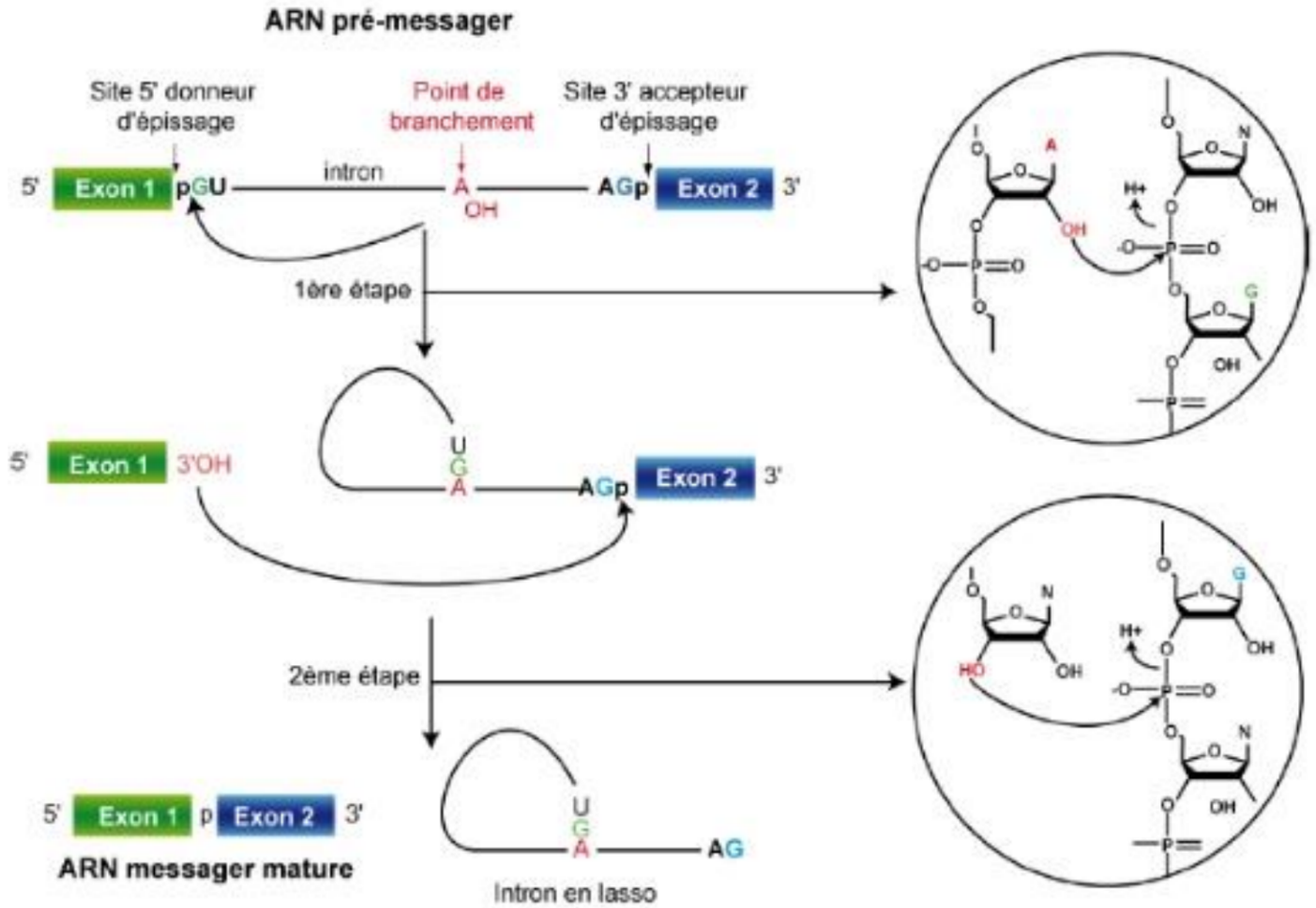


Figure 8. Réactions de *trans-estérification* d'épissage.

### 2.3. Composants et l'assemblage du spliceosome

Le spliceosome est un macrocomplexe ribonucléoprotéique (RNP) architecturé par l'assemblage de 5 particules ribonucléoprotéiques appelées UsnRNP (snRNP U1, U2, U4, U5, U6). Chaque UsnRNP est composée d'un nombre variable de protéines, ainsi que d'un ou deux petits ARN nucléaires appelés UsnRNA (pour "Uridine rich small nuclear ribonucleic acids"). Chacune de ces UsnRNP s'architecture elle-même autour d'un ou plusieurs petits ARN nucléaires riches en uridine (ou UsnRNA), et comprend un ensemble de 7 protéines Sm ou Sm-like (Lsm) ainsi que plusieurs protéines spécifiques à chaque UsnRNP. Comme décrit par la suite, les UsnRNP et les protéines associées au spliceosome reconnaissent les séquences nécessaires à l'épissage. Des interactions ARN-ARN et ARN-protéines se forment alors de manière ordonnée et transitoire autour des sites d'épissage, conduisant à la formation du spliceosome au sein duquel ont lieu les 2 réactions de trans-estérification.

Le mécanisme d'épissage a donc été déterminé à partir des complexes intermédiaires de réaction. Ainsi, le snRNP U1 est le premier facteur à entrer en jeu. Il se lie au « site donneur 5' » par l'intermédiaire de son snRNA en s'appariant à la séquence donneur 5'. Cette première interaction est indépendante de l'ATP. Parallèlement, le facteur d'épissage 1 (SF1, pour splicing factor 1) s'accroche au point de branchement, formant le complexe E'. Un second complexe appelé E est créé à la suite de l'ajout du facteur auxiliaire U2 (U2AF) un hétérodimère, composé de deux sous unités, découvert en 1989 par Grenn et Zamore. Une sous unité de U2AF se fixe sur la région riche en pyrimidines et au point de branchement tandis que la sous unité se fixe au site accepteur 3' AG. Le remplacement de SF1 par le snRNP U2 au point de branchement donne le complexe A encore appelé pré-spliceosome A. Cette réaction se fait de manière ATP-dépendant. Ensuite, le recrutement des snRNP U4, U5 et U6 permet la formation du complexe B. Les snRNPs U4 et U6, s'assemblent en un complexe contenant 2 snRNAs. L'association avec U5 crée un tri-snRNP pré-assemblé. Des changements conformationnels et des remaniements tels que la perte des sous unités U1 et U4 permet l'obtention du complexe C qui se trouve être le spliceosome actif. Ce spliceosome actif, isolé en 2008 par Bessonov induit les 2 trans-estérifications successives. Ainsi les 2 exons sont joints et la partie intronique, sous forme de lasso, est éliminée. Les snRNP sont ensuite recyclées afin d'être utilisées pour un autre épissage (Figure 9).

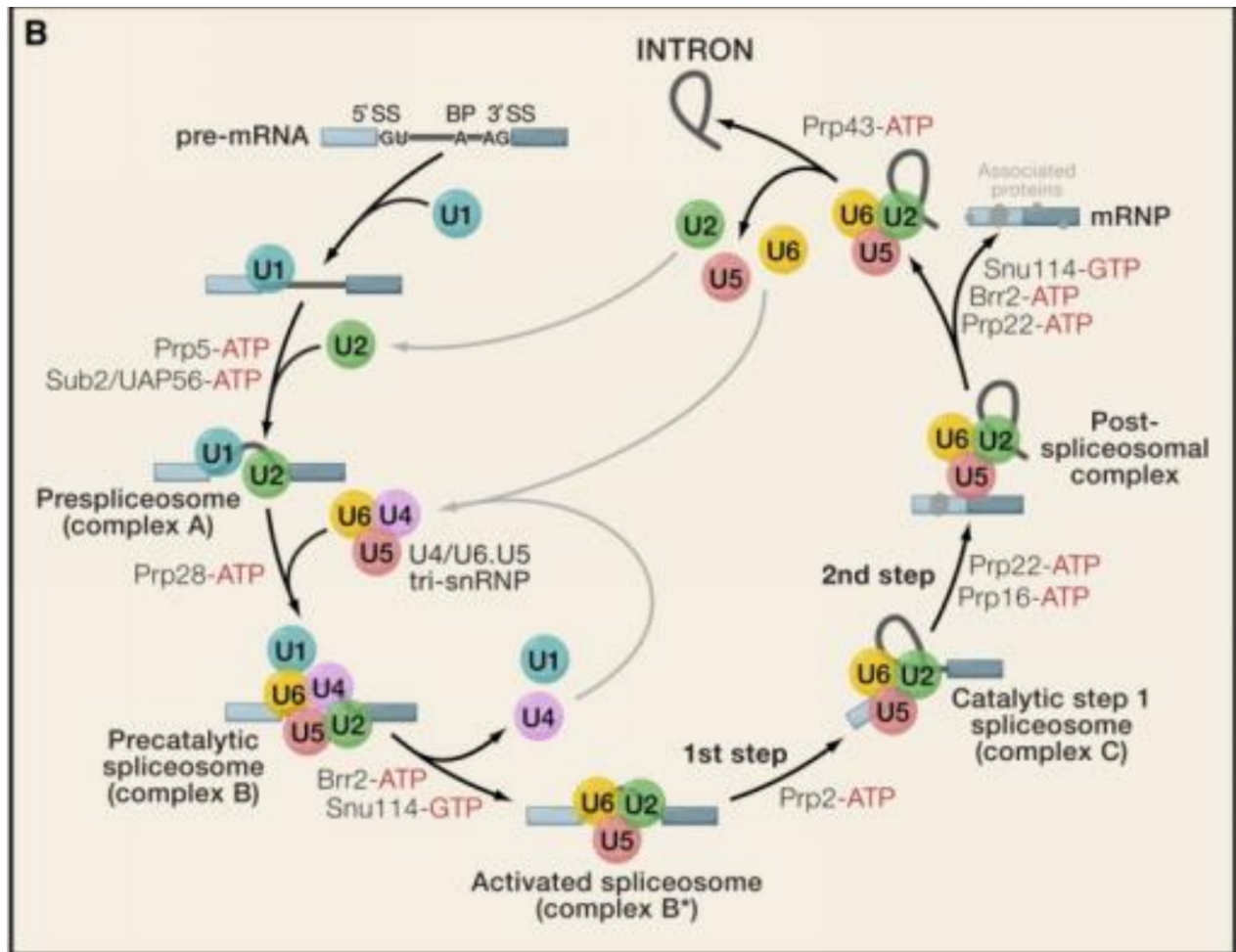


Figure 9. Assemblage de spliceosome et épissage d'ARNm.

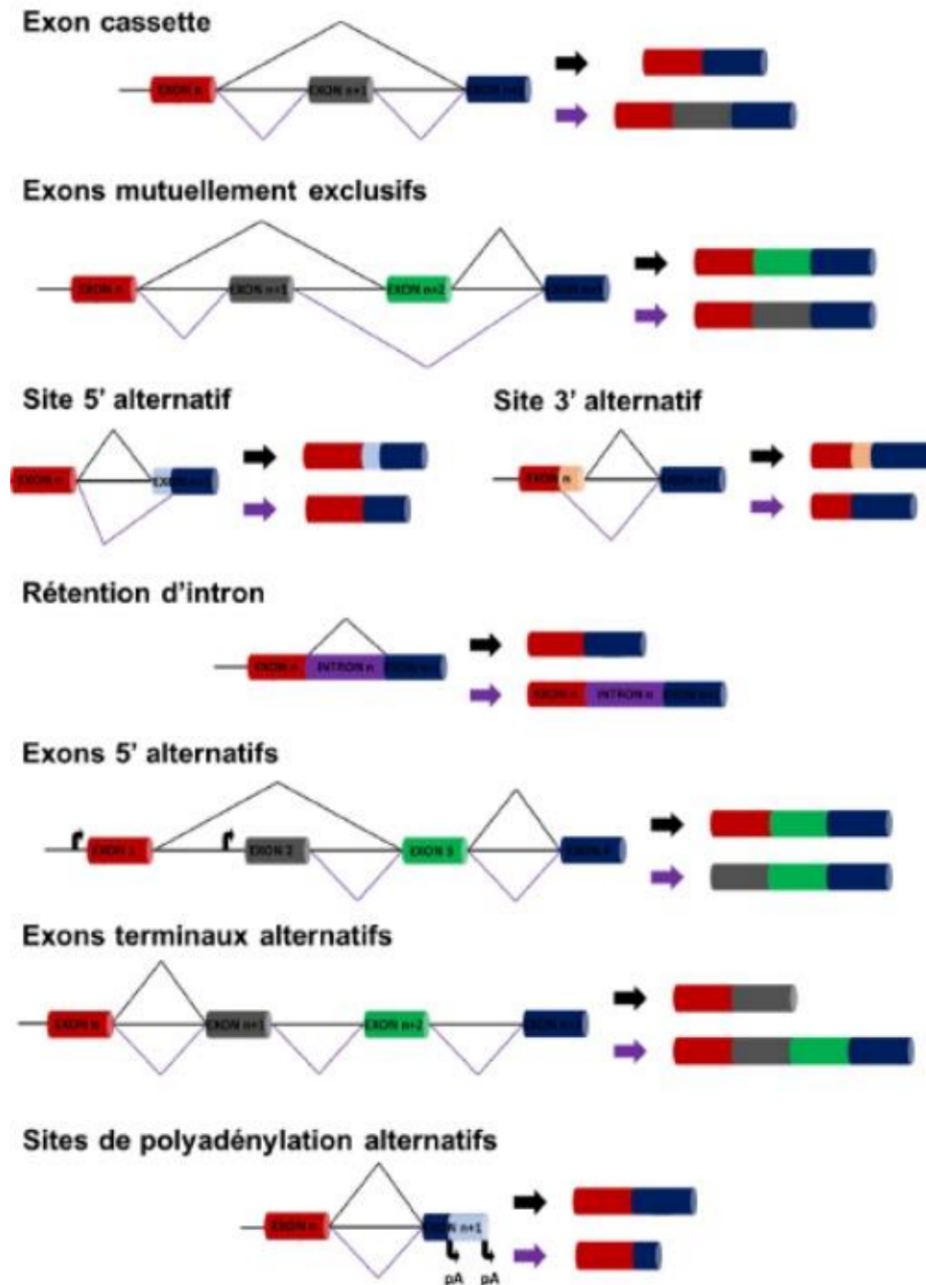
### 3. Epissage alternatif

Il existe chez l'homme de nombreux introns dont les sites 5' et 3' d'épissage sont dégénérés par rapport aux séquences consensus. Par conséquent, la seule reconnaissance de ces sites n'est pas suffisante pour la définition et l'épissage des introns, ce qui nécessite l'intervention de séquences régulatrices supplémentaires. Ces séquences régulatrices permettent des modulations du taux d'utilisation des sites d'épissage entraînant un épissage dit alternatif des exons. L'épissage alternatif est un phénomène majeur chez l'homme, qui au cours de l'évolution a permis d'augmenter significativement la capacité codante de notre génome en exprimant plusieurs protéines avec des fonctions ou des spécificités distinctes voir même opposées, à partir d'un seul gène. A l'heure actuelle, il est estimé que près de 93% des gènes humains subissent au niveau de leurs pré-ARNm un épissage alternatif qui peut varier d'un tissu à l'autre.

#### 3.1. Différents types d'évènements d'épissage alternatif

Chez les eucaryotes supérieurs plusieurs situations et types d'épissages alternatifs ont été décrits, parmi les quelles (Figure 10);

- A. **exons cassettes** : c'est l'évènement d'épissage alternatif le plus fréquent, les exons cassettes sont soit inclus soit exclus de l'ARNm par la machinerie d'épissage.



**Figure 10.** Différents modes d'épissage alternatif des pré-ARNm.

- B. exons mutuellement exclusifs :** bien que deux exons mutuellement exclusifs aient chacun des sites 5' et 3' fonctionnels, la reconnaissance par le spliceosome d'un des couples de sites abolit la reconnaissance de l'autre. Par conséquent, ces exons ne sont jamais présents ensemble au sein du même ARNm.
- C. Sites 5' ou 3' alternatifs d'épissage :** il peut exister au niveau de certains exons plusieurs sites 5' ou 3' d'épissage fonctionnels qui entrent alors en compétition. Selon le site utilisé lors de l'épissage, les conséquences de la sélection d'un site par rapport à un autre sont la formation d'isoformes d'ARNm plus ou moins long.

### III. Dégradation et surveillance des ARNm eucaryotes

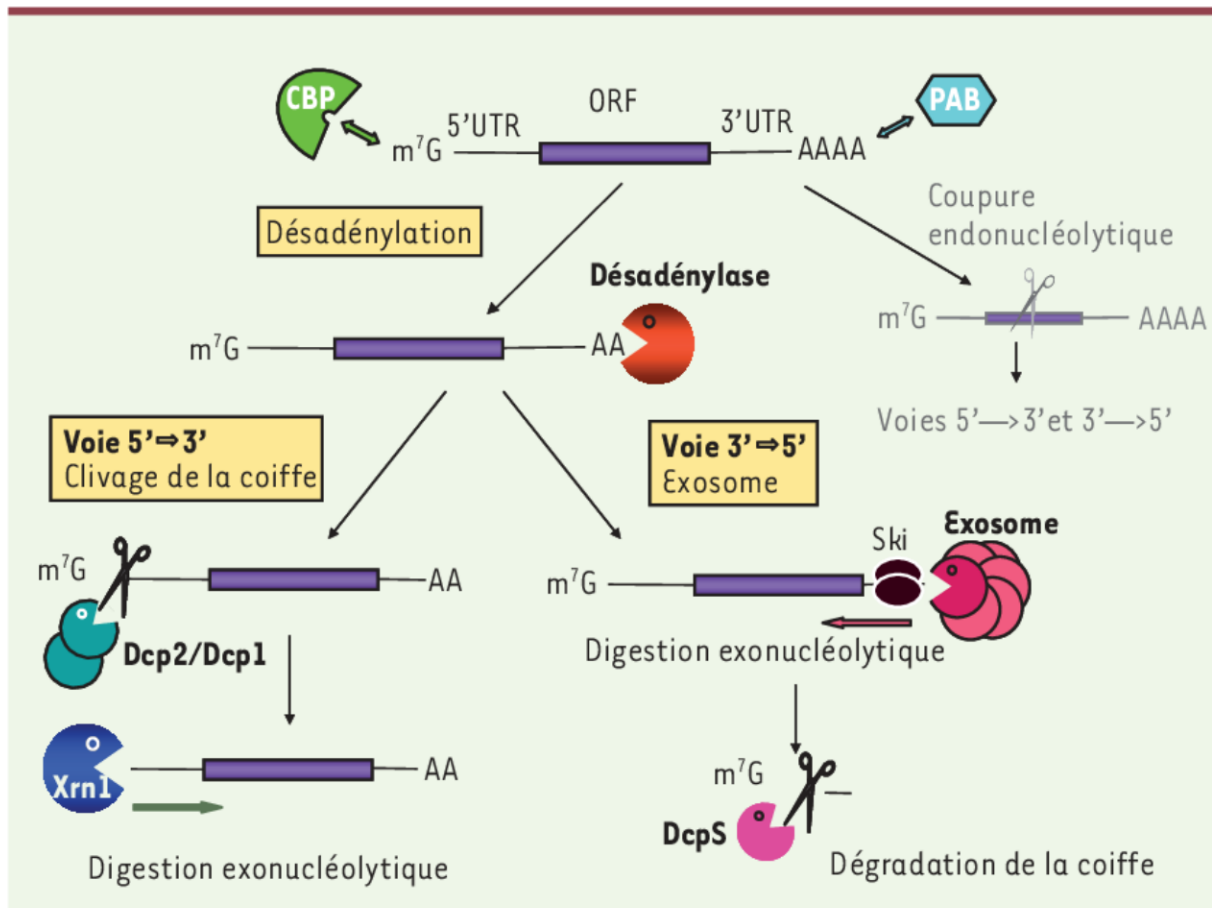
#### 1. Mécanisme général de dégradation cytoplasmique des ARNm

La dégradation des ARNm a longtemps été considérée comme une simple étape d'élimination des ARNm. Les temps de demi-vie des ARNm, révélateurs de leur vitesse de dégradation, peuvent varier énormément. Ainsi chez les mammifères, ils varient de 15 minutes à 24 heures, montrant que la dégradation des ARN est un processus spécifique et non un processus par défaut. La dégradation des ARNm est essentielle. En l'absence d'une dégradation rapide des ARNm, une régulation rapide de la production de protéines par la seule induction transcriptionnelle serait impossible.

Dans la cellule, plusieurs facteurs contribuent à la protection des ARNm contre l'action des ribonucléases. Lors de sa synthèse, l'ARNm subit des modifications à ses extrémités 5' et 3' : ajouts d'une coiffe et d'une queue poly(A). Ces marques, liées par des protéines spécifiques, les *cap-binding protein* (CBP) et *Poly(A)-binding protein* (PAB) le protègent de l'action des exoribonucléases. L'ARNm est également associé à d'autres protéines au sein de larges particules ribonucléoprotéiques (mRNP) qui notamment le protègent de coupures endonucléolytiques non spécifiques.

La dégradation d'un ARNm résulte d'un profond remaniement de la mRNP, déclenché par des signaux spécifiques et qui conduit au recrutement des enzymes et activateurs de la dégradation (*Figure 11*)

L'événement initial le plus fréquent est une désadénylation. Le corps de l'ARNm est ensuite dégradé par une des 2 voies : 5' → 3' ou 3' → 5'. Dans le cas de la voie 5' → 3', la désadénylation est suivie du clivage de la coiffe par l'enzyme Dcp2/Dcp1 puis d'une digestion exonucléolytique par Xrn1. Dans le cas de la voie 3' → 5', la désadénylation est suivie du recrutement du complexe Ski et d'une digestion exonucléolytique par l'exosome ; la coiffe est finalement dégradée par l'enzyme DcpS. En grisé est représentée une voie plus rare, amorcée par une coupure endonucléolytique de l'ARNm, suivie des voies classiques



**Figure 11.** Les principales voies cytoplasmiques de dégradation des ARNm normaux.

## 2. Contrôle de qualité des ARNm

Un rôle majeur de la dégradation est l'élimination des ARNm dont la maturation est incorrecte ou qui contiennent des erreurs. Les systèmes de surveillance de la qualité des ARNm se répartissent en 2 grandes catégories selon qu'ils font intervenir ou non la traduction des ARNm.

### Dans le noyau

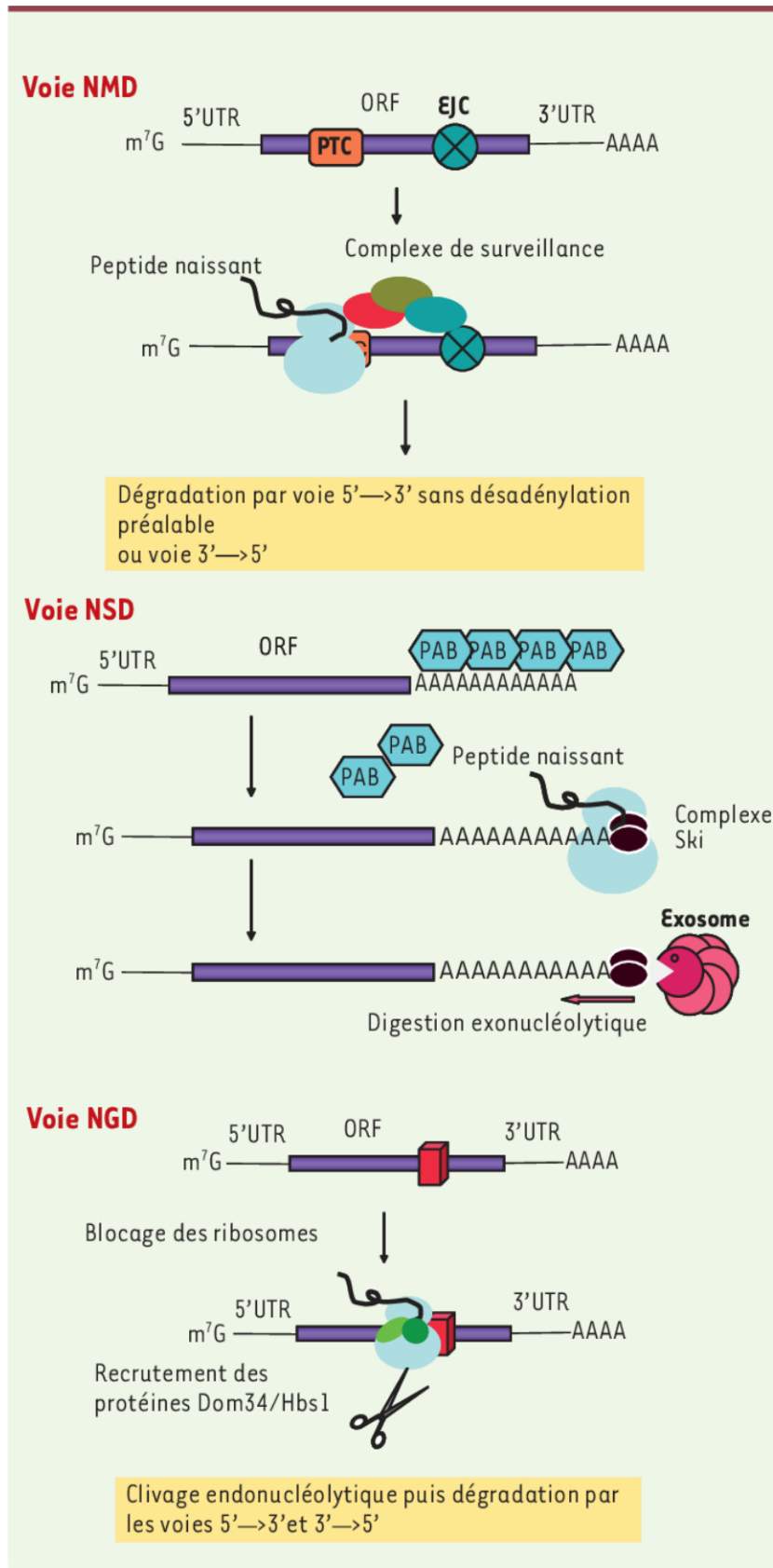
Un premier niveau de contrôle détecte, dans le noyau, la présence de défauts de maturation (absence de coiffe, erreurs d'épissage, défauts dans la queue poly(A)...), qui s'accompagnent souvent de défauts d'export vers le cytoplasme. Des études récentes chez la levure montrent que ces ARNm déficients sont dégradés par l'exosome nucléaire, à la suite de leur polyadénylation par une poly(A) polymérase spécifique (complexe TRAMP). Dans ce cas, la polyadénylation n'est plus une marque protectrice mais un signal de dégradation, comme cela est observé pour les ARN bactériens.

Le complexe TRAMP est aussi impliqué dans la dégradation de formes aberrantes d'ARN non codants et de transcrits cryptiques CUT (*cryptic unstable transcripts*) synthétisés par l'ARN polymérase II à partir de régions intergéniques du génome de levure et probablement d'autres génomes eucaryotes.

### Dans le cytoplasme

Le second niveau de contrôle est cytoplasmique. Il est connecté à la traduction et permet l'élimination d'ARNm qui conduiraient à la synthèse de protéines aberrantes, potentiellement délétères pour la cellule. Trois types d'ARNm aberrants sont concernés (*Figure 12*) : les ARNm possédant un codon stop précoce (voie NMD pour *non-sense-mediated decay*), les ARNm ne possédant pas de codon stop (voie NSD pour *non-stop decay*) et les ARNm sur lesquels les ribosomes sont bloqués, piégeant ainsi les facteurs de traduction (voie NGD pour *no-go decay*).

La voie NMD représentée est celle observée chez les mammifères, elle fait intervenir l'EJC (complexe déposé au niveau des jonctions exon-exon) dans le mécanisme de reconnaissance d'un codon stop précoce (PTC, *premature termination codon*). La reconnaissance du PTC conduit à l'assemblage du complexe de surveillance qui recrute les enzymes de dégradation. **Voies NSD et NGD.** En l'absence de codon stop (voie NSD) ce qui conduit à un blocage du ribosome à l'extrémité de l'ARNm, ou lorsque le ribosome est bloqué sur l'ARNm en cours de traduction (voie NGD, par exemple, à la suite de la présence d'une structure secondaire), les complexes Ski ou Dom34/Hbs1 se lient à l'ARNm et recrutent les enzymes de dégradation : exosome ou endonucléase ;



**Figure 12.** Les principales voies cytoplasmiques de surveillance de la qualité des ARNm. Voie NMD.

### 3. La dégradation est un phénomène contrôlé

Au cours de sa vie, un ARNm s'associe au sein de diverses mRNP avec des protéines qui vont déterminer son sort : maturation, export, localisation, stockage, traduction, dégradation.

La dégradation est, en particulier, fortement connectée à la traduction. Ainsi, les protéines qui lient la coiffe et la queue poly(A) sont à la fois nécessaires à une traduction efficace et pour protéger l'ARNm de la dégradation. Certains ARNm dont la traduction est inhibée par des protéines spécifiques se retrouvent stockés dans les *foci* de dégradation. Ce stockage est réversible, l'ARNm pouvant ensuite être de nouveau traduit, ou bien dégradé. Des études globales ont montré que la part de la dégradation dans la régulation des quantités d'ARNm peut être considérable. Par exemple, lors de l'activation des lymphocytes T (réponse immunitaire), 50 % des effets observés sur les quantités d'ARNm sont dus à une régulation de leur stabilité.

#### **Protéines régulatrices**

La plupart des protéines régulant la dégradation des ARNm se fixent sur l'ARNm lui-même *via* des séquences spécifiques situées dans les régions 5' UTR, 3' UTR ou parfois dans la région codante. Fréquemment ces protéines activent la dégradation par une accélération de la désadénylation. Le nombre de ces régulateurs est trop important pour permettre leur description de manière exhaustive. Deux exemples bien caractérisés ont été choisis pour illustrer ce propos.

#### ***Protéines Puf***

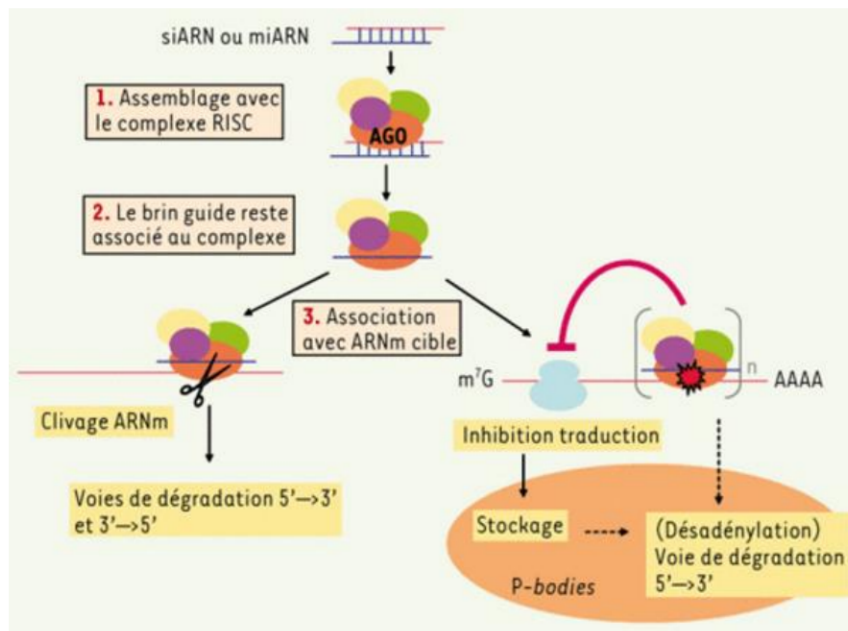
Les protéines Puf représentent une famille de protéines reconnaissant des séquences spécifiques trouvées dans certains ARNm. Ces ARNm peuvent contenir plusieurs sites de liaison identiques ou différents, augmentant ou modulant leur dégradation. Des analyses globales ont révélé qu'un type de protéine Puf reconnaît généralement une classe d'ARNm codant pour des protéines de fonction voisine ou de localisation identique (par exemple Puf3 de levure et la mitochondrie). Certaines protéines Puf accélèrent la déadénylation et le clivage de la coiffe.

#### ***Protéines se liant aux séquences ARE (AU-rich elements)***

Chez les vertébrés, certains ARNm, souvent de demi-vie courte comme les ARNm des cyclines, cytokines, facteurs de croissance, protooncogènes... possèdent des séquences riches en résidus adénine et uracile dans leur région 3' UTR. Ces séquences sont reconnues par des protéines déstabilisatrices (tristétraproline, BRF1, AUF-1, KSRP...) qui agissent en recrutant les enzymes de dégradation (désadénylases, exosome...). Dans d'autres conditions, elles peuvent lier des protéines stabilisatrices qui entrent en compétition avec les protéines déstabilisatrices (comme Hu-R de la famille ELAV) ou pourraient inhiber directement la machinerie de dégradation. Certaines de ces protéines régulent parfois aussi la traduction, renforçant encore les liens fonctionnels entre dégradation et traduction.

### Petits ARN régulateurs

La découverte de l'interférence ARN, qui valut en 2006 le prix Nobel à A. Fire et C. Mello, a révélé l'existence d'un nouveau mode de dégradation des ARN *via* l'association de petits ARN (*Figure 13*). Cette dégradation est différente de celle, non spécifique, observée chez les mammifères lors de la réponse antivirale (réponse interféron), déclenchée par de longs ARN double-brin et faisant intervenir la protéine kinase PKR et la RNase L. Elle est absente chez la levure *S. cerevisiae*.



**Figure 13.** Les différents modes d'interférence ARN cytoplasmique.



# Liste des références

- 1- [Anton Eberharter](#) , [Peter B Becker](#) 2002 Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep. 2002 Mar;3(3):224-9. doi: 10.1093/embo-reports/kvf053.
- 2- Bruce Alberts, Rebecca Heald, Alexander Johnson, David Morgan; Martin Raff, Keith Roberts and Peter Walter. Molecular Biology of the cell. Norton & Company, Inc., New York.
- 3- Carla Sawan and Zdenko Herceg. 2010 Histone Modifications and Cancer in Advances in Genetics, Vol. 70. Elsevier Inc. 57-85.
- 4- Christine Mayr 2017 Regulation by 3'-Untranslated Regions. *Annual Review of Genetics*. GE51CH09-Mayr, 22 August 2017.
- 5- Danièle Hernandez-Verdun et Émilie Louvet. 2004 Le nucléole : structure, fonctions et maladies associées The nucleolus: functions, assembly and pathology. Med Sci (Paris) Volume 20, Number 1, Janvier 2004.
- 6- David Alan Schneider 2012 RNA polymerase I activity is regulated at multiple steps in the transcription cycle: recent insights into factors that influence transcription elongation. *Gene*. 2012 February 10; 493(2): 176–184. doi:10.1016/j.gene.2011.08.006.
- 7- David S. Latchman 2008 eukaryotic transcription factors, Fifth edition. Academic Press is an imprint of Elsevier.
- 8- Estelle Nicolas, Laurence Vandell et Didier Trouche. 2010 *Quand la méthylation des histones entre en scène... médecine/sciences 2010 ; 17 : 476-9.*
- 9- Gary H. Perdew, John P. Vanden Heuvel, Jeffrey M. Peters. 2006 Regulation of Gene Expression molecular mechanisms. *Humana Press Inc.*
- 10- Glenn A. Maston, Sara K. Evans, and Michael R. Green Transcriptional Regulatory Elements in the Human Genome. *Annu. Rev. Genomics Hum. Genet.* 2006. 7:29–59.
- 11- Jackie Russell and Joost C.B.M. Zomerdijk. 2005 RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem Sci.* 2005 February ; 30(2): . doi:10.1016/j.tibs.2004.12.008.

- 12- JANET C. LINDSEY, PH.D., JENNIFER A. ANDERTON, B.SC., MERYL E. LUSHER, PH.D., AND STEVEN C. CLIFFORD, PH.D. 2005 Epigenetic events in medulloblastoma development. *Neurosurg Focus* 19 (5):E10, 2005.
- 13- Jesse R. Raab and Rohinton T. Kamakaka 2010. Insulators and promoters: closer than we think. *Nat Rev Genet.* 2010 June ; 11(6): 439–446. doi:10.1038/nrg2765.
- 14- Johannes Gräff, Isabelle M. Mansuy. 2008 Epigenetic codes in cognition and behaviour. *Behavioural Brain Research* 192 (2008) 70–87.
- 15- Kim Nasmyth. 1993 Regulating the HO endonuclease in yeast. *Current Opinion in Genetics and Development* 1993, 3:286-294.
- 16- Lizabeth A. Allison. 2007 *Fundamental Molecular Biology*. BLACKWELL PUBLISHING.
- 17- Marie-France Langelier, Vincent Trinh, and Benoit Coulombe Gros plan sur l'ARN polymérase II. *Med Sci (Paris)*. 2002 February ; 18(2): 210–216. doi:10.1051/medsci/2002182210.
- 18- Michael Carey and Stephen T. Smale 2000 *Transcriptional Regulation in Eukaryotes. Concepts, Strategies, and Techniques*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 19- Pollard Jeffrey W. and Walker John M. 1997 *Basic cell culture Protocols second edition*. Humana Press, Totowa, New Jersey.
- 20- Ralph Raply and John M Walker 2009 *Molecular Biology and Biotechnology* 5th Edition. Royal Society of Chemistry, RSC Publishing.
- 21- Sarah J. Goodfellow and Joost C. B. M. Zomerdijk. 2012 Basic Mechanisms in RNA Polymerase I Transcription of the Ribosomal RNA Genes. *Subcell Biochem.* 2012 ; 61: . doi:10.1007/978-94-007-4525-4\_10.
- 22- Yashar Sadian, Lucas Tafur, Jan Kosinski, Arjen J Jakobi, Rene Wetzels, Katarzyna Buczak, Wim JH Hagen, Martin Beck , Carsten Sachse & Christoph W Müller. 2017 Structural insights into transcription initiation by yeast RNA polymerase I. *The EMBO Journal* Vol 36, No 18.

23- Camier, S., & Séraphin, B. (2007). Détruisez ce message (ARN) après l'avoir lu !

*Destroy this message after reading it! Médecine/Sciences, 23(10), 850–856.*

Google books

<https://books.google.dz/books?id=DSJ1mJkOzzAC&printsec=frontcover&dq=promoteur+eucaryote+pdf&hl=fr&sa=X&ved=0ahUKEwjtmO2o6J3gAhXCA2MBHWsQC84Q6AEILDAB#v=onepage&q=promoteur%20&f=false>

[https://books.google.dz/books?id=x\\_vfiWmrTbkC&printsec=frontcover&dq=promoteur+eucaryote+pdf&hl=fr&sa=X&ved=0ahUKEwjtmO2o6J3gAhXCA2MBHWsQC84Q6AEIRTA#v=onepage&q=promoteur&f=false](https://books.google.dz/books?id=x_vfiWmrTbkC&printsec=frontcover&dq=promoteur+eucaryote+pdf&hl=fr&sa=X&ved=0ahUKEwjtmO2o6J3gAhXCA2MBHWsQC84Q6AEIRTA#v=onepage&q=promoteur&f=false)

[https://books.google.dz/books?id=hywtDwAAQBAJ&printsec=frontcover&dq=promoteur+eucaryote+pdf&hl=fr&sa=X&ved=0ahUKEwiTuPHC7Z3gAhUNmRQKHVC\\_B2cQ6AEIjzAA#v=onepage&q=g%C3%A8ne&f=false](https://books.google.dz/books?id=hywtDwAAQBAJ&printsec=frontcover&dq=promoteur+eucaryote+pdf&hl=fr&sa=X&ved=0ahUKEwiTuPHC7Z3gAhUNmRQKHVC_B2cQ6AEIjzAA#v=onepage&q=g%C3%A8ne&f=false)