

Chapitre 3 :
Expression des gènes
(SYNTHESE DES PROTEINES)

C'est l'acte par lequel une cellule assemble une chaîne protéique en combinant des acides aminés isolés présents dans son cytoplasme, guidé par l'information, contenue dans l'ADN.

Elle se déroule en deux étapes :

- **la transcription de l'ADN en ARN messenger**
- **La traduction de l'ARN messenger en une protéine.**

3.1. La transcription :

La transcription est le processus de copie du matériel génétique en ARN. Chez **les procaryotes** une seule ARN-polymérase effectue la transcription pour tous les types d'ARN, constituée de plusieurs sous unités : α , β , β' et σ . Cette enzyme a deux dénominations : **enzyme cœur** ou **holoenzyme**, selon sa constitution.

Le complexe holoenzyme correspond à l'association des sous unité, α , β , β' et σ . Il initie la transcription. Puis le facteur σ est relargie pour former l'enzyme cœur (constituée des sous unités α , β et β'), qui synthétise tous Il existe :

- 2 α : Assemblage de l'enzyme, assure la liaison au promoteur.
- 1 β : Assure la liaison des nucléotides.
- 1 β' : Assure la liaison à la matrice d'ADN.
- 1 σ : Assure la reconnaissance du promoteur, initiation de la transcription.

Tandis que chez **les eucaryotes** trois ARN-polymérases différentes interviennent selon qu'il s'agit de produire un ARN ribosomique, un ARN messenger ou un petit ARN (ARN de transfert par exemple).

- **L'ARN polymérase I** localisée dans le Nucléole synthétise les ARNr(5.8S, 18S, 28 S)leur activité cellulaire est de 60 à 70 %
- **L'ARN polymérase II** localisée dans le Nucléoplasmesynthétise les ARNm et certains des ARN nucléaires (ARNsn) leur activité cellulaire est de 10 à 30 %
- **L'ARN polymérase III** localisée dans le Nucléoplasmesynthétise la plupart des ARN ARNt, ARNr 5S et certainsARNsn, leur activité cellulaire est 10 % environ

C. Mécanisme de la transcription :

3.1.1. Transcription chez les procaryotes :

• **Phase d'initiation :**

L'ARN polymérase reconnaît et se fixe sur de l'ADN, localisées sur une séquence localisée en 5' du gène avant le point d'initiation de la transcription. Ces séquences sont appelées **promoteurs** ou régions promotrices. Pour un gène donné : convention +1 au lieu où la transcription est initiée. Donc le nt +1 est le premier nt transcrit sur l'ARN, et donc le ribonucléotide incorporé en 5' du futur transcrit. Avant ce site, toutes les paires de bases sont numérotées de manière négative (figure 22).

Pour les gènes **procaryotes**, on retrouve dans les régions promotrices des **séquences consensus** (conservées au cours de l'évolution) localisées en position **-12** et **-35**. Ces séquences sont des séquences **hexamériques** (6nt) **riches en T et en A**, surtout celle positionnée en -12 (figure 22).

Région en -35 (3' TTGACA 5') : lieu de fixation de l'ARN polymérase.

Région en -12 (3' TATAAT 5') = **boîte de Pribnow**. Région très importante, composée principalement de thymine et d'adénine.

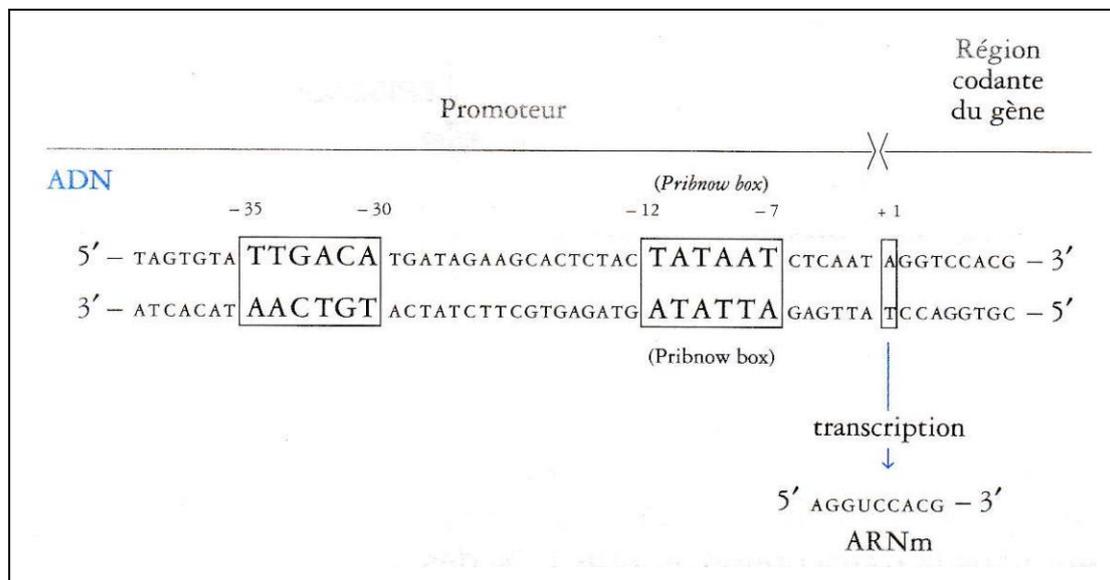


Figure 22 : séquence d'un promoteur des procaryotes (Étienne et al ; 1999)

La fixation du complexe ARN polymérase sur la séquence -35 de l'ADN génomique entraîne l'ouverture de la double hélice, pour permettre la lecture de l'ADN par l'ARN polymérase (figure 21). L'ADN et l'ARN sont associés par liaisons hydrogènes) qui s'établit sur **12 bases sur une boucle qui contient** (zone dénaturée).

L'ARN polymérase se déplace le long de l'ADN et lit le brin qui doit être transcrit, c'est-à-dire qu'il y a incorporation de ribonucléotides (polymérisation) et que l'ARN polymérase favorise la création des liaisons phosphodiester.

L'ATP, GTP, UTP et CTP rentrent sous forme triphosphate dans la réaction, mais dans la molécule d'ADN, ils sont sous forme monophosphate. Il y a donc élimination du pyrophosphate, ce qui apporte l'énergie nécessaire à la réaction de polymérisation. Au bout d'une dizaine de ribonucléotides incorporés, le facteur σ est libéré. L'ARN polymérase adopte donc sa conformation cœur, pour poursuivre la synthèse de l'ARN.

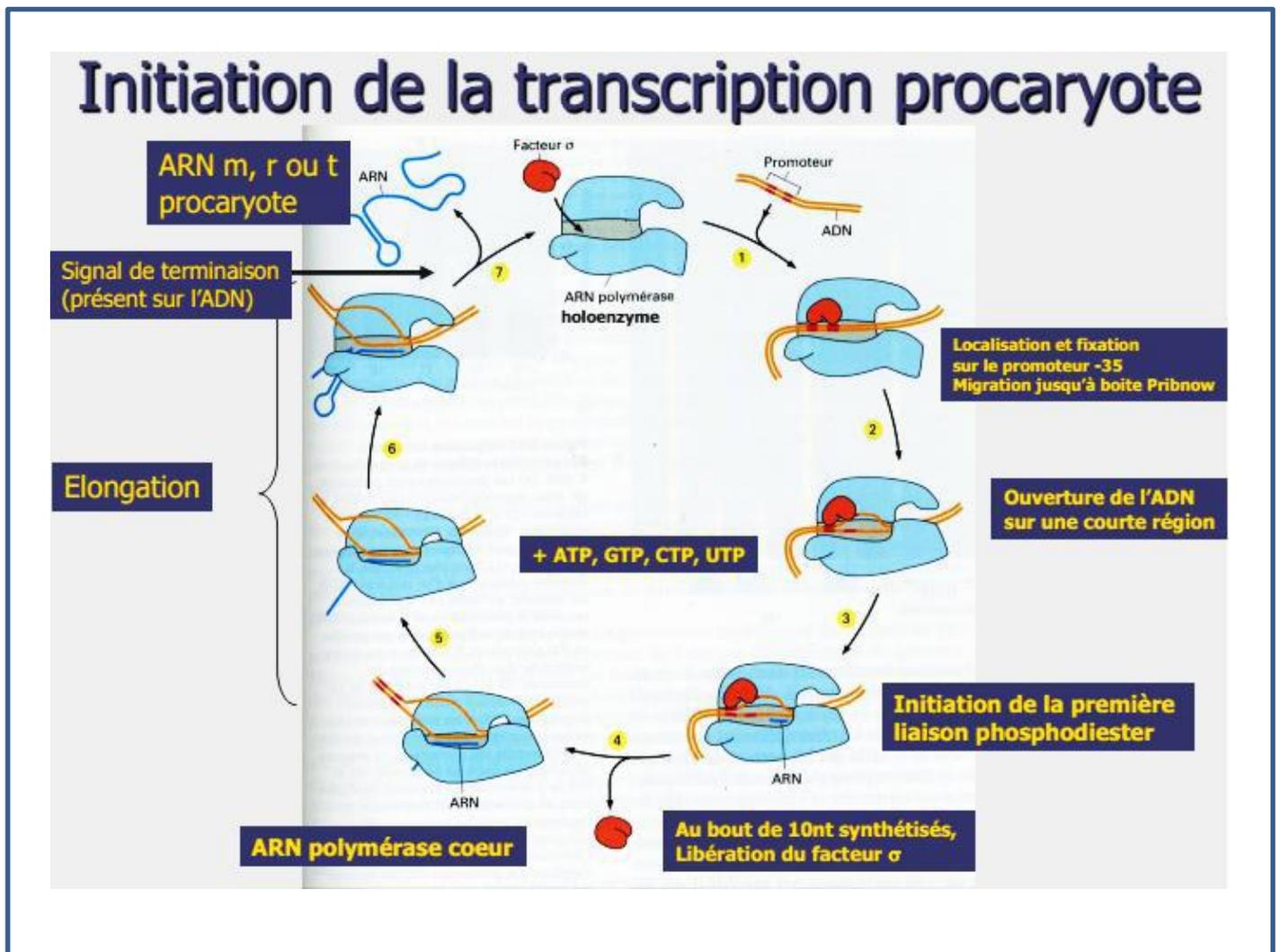


Figure 23: initiation et élongation de la transcription chez les procaryotes (P. Cohen, 2017)

- **Elongation**

La polymérisation se fait bien dans le sens 5'→ 3' et elle s'effectue en l'absence de toute amorce préexistante. Le premier ribonucléotide est sous forme triphosphate, les autres sont sous forme monophosphate. Puis création de la liaison phosphodiester au niveau du 3'OH avec un autre ribonucléotide qui va libérer son pyrophosphate, ce qui libère l'énergie nécessaire à la réaction. Ensuite **progression de l'ARN polymérase** et de la boucle de transcription en supprimant les liaisons hydrogènes. Une fois la séquence transcrite, **l'ADN se renature** spontanément

- **Terminaison**

Quand l'ARN polymérase reconnaît un dernier signal sur l'ADN : **le site de terminaison**, la synthèse s'achève et il y aura dissociation du complexe ADN-ARN polymérase suivie par la libération de la polymérase et de la chaîne d'ARN qui vient d'être transcrite.

3.1.2. Transcription chez les eucaryotes :

- **Initiation**

Contrairement à ce qui se passe chez les procaryotes, l'ARNpol II des eucaryotes ne reconnaît pas seule le promoteur proximal. Elle effectue ce travail en compagnie de nombreux co-facteurs protéiques qui se recrutent les uns les autres et qui forment avec elle un complexe d'initiation. Ces facteurs sont notés TFIIA, TFIIB, etc. pour Transcription Factor for RNA polymerase II. Ils correspondent aux facteurs généraux de la transcription, car ils s'assemblent sur tous les promoteurs utilisés par l'ARNpol II. La séquence d'assemblage du complexe d'initiation est décrite sur la figure 24.

La TBP est la première protéine qui reconnaît une séquence spécifique de l'ADN initiatrice de la transcription (la boîte TATA). Le facteur TFIIB semble impliqué dans la sélection précise du site d'initiation (nucléotide à partir duquel se déroule la transcription). Le facteur TFIIF comporte plusieurs activités enzymatiques dont une activité hélicase permettant l'ouverture de la double hélice d'ADN au niveau du promoteur, et une activité kinase responsable de la phosphorylation de la queue C-terminale de l'ARN polymérase II. Cette phosphorylation provoque une modification de la structure tridimensionnelle de l'ARN polymérase qui entraîne la dissociation du complexe d'initiation et le début de la transcription (Figure 24).

TFII = *Transcription Factor for RNAPolymerase II* (facteurs généraux de la transcription pour l'ARN polymérase II).

TBP = *TATA box-Binding Protein* (protéine de liaison à la boîte TATA).

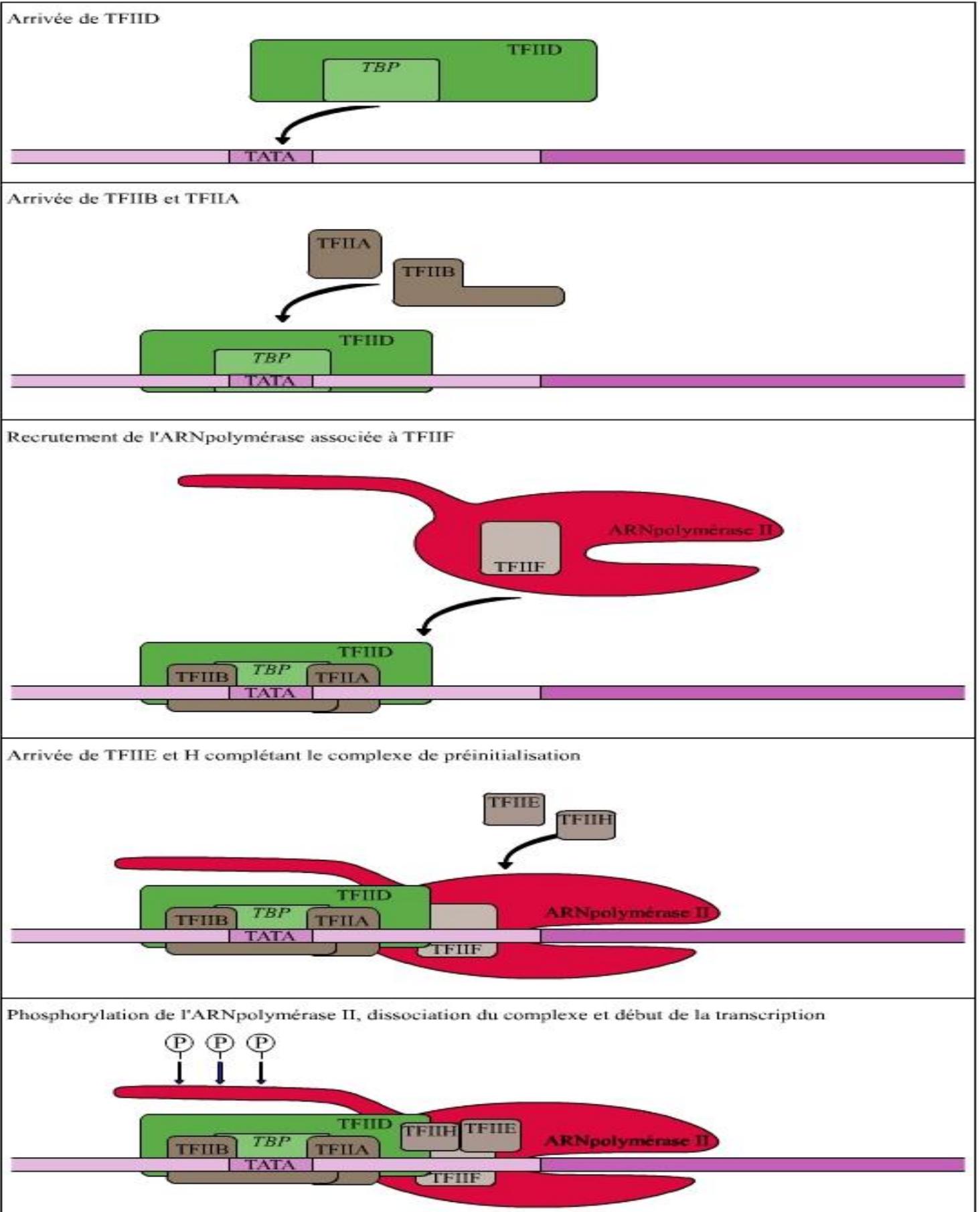


Figure 24 : Mise en place du complexe d'initiation

La liaison du complexe de transcription au promoteur proximal provoque l'ouverture et le déroulement des deux brins de son ADN, tout en indiquant le brin qui va être transcrit.

2.1.3. 2.1.3. Intervention de facteurs spécifiques de la transcription

Le complexe d'initiation composé de l'ARNpol II et des différents TFII est suffisant pour obtenir une activité transcriptionnelle in vitro mais à très faible taux. L'augmentation de cette activité basale (ou sa répression) est sous la dépendance de facteurs spécifiques qui vont interagir avec le complexe d'initiation. Ces protéines activatrices ou inhibitrices (éléments trans-régulateurs) se lient à des promoteurs distaux spécifiques (séquences cis-régulatrices) de l'ADN, appelées amplificateurs (*enhancers*) lorsqu'ils recrutent des cofacteurs activateurs, ou silencers (*silencers*) lorsqu'ils recrutent des cofacteurs inhibiteurs. Ces promoteurs distaux peuvent être situés à des milliers de nucléotides du promoteur proximal. Malgré la distance qui sépare les promoteurs proximaux des promoteurs distaux, ces derniers agissent sur le promoteur proximal par le jeu de courbures de l'ADN, des facteurs de transcription et du médiateur qui maintient liés tous ces acteurs (voir fig. 25).

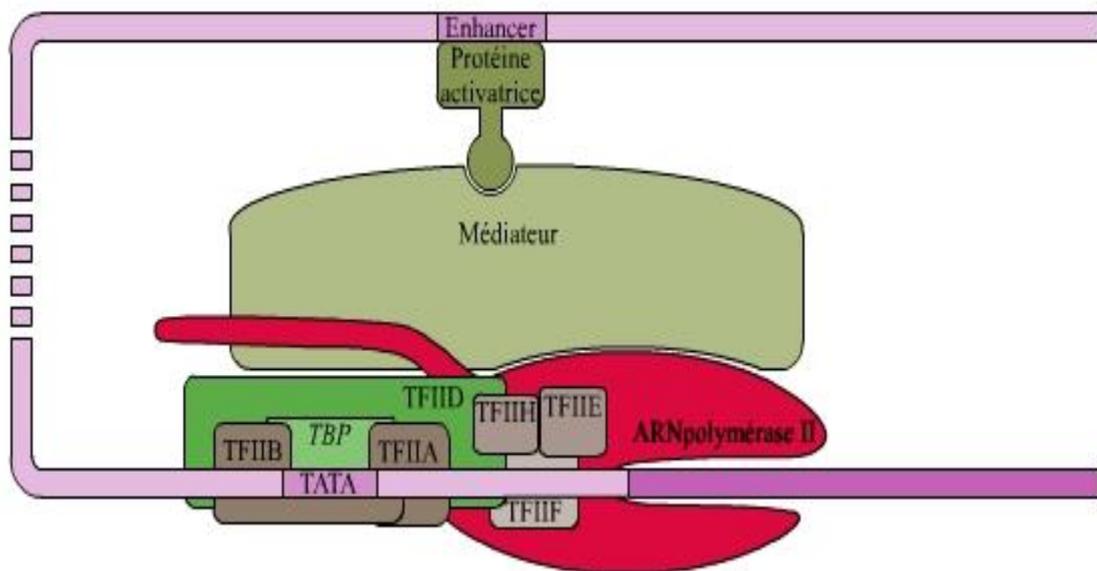


Figure 25 : La régulation spécifique de la transcription

Les amplificateurs (*enhancers*) sont des séquences situées parfois à plusieurs milliers de nucléotides d'un promoteur et qui, par un jeu d'interactions protéiques, stabilisent le complexe d'initiation, favorisant ainsi la transcription. Ces interactions sont rendues possibles par la courbure de l'ADN qui rapproche des éléments situés à de grandes distances les uns de autres.

Cette activation est spécifique du gène et utilise une multitude de facteurs de transcription spécifiques (les protéines activatrices) qui agissent généralement sous forme dimérique. Une répression spécifique peut agir selon le même type de modalité.

2.2. L'élongation

L'ARNpol II est équipée de facteurs protéiques d'élongation qui facilitent sa progression au travers d'une chromatine dont ils relâchent la structure (c'est l'un d'entre eux qui est mis hors d'état d'agir par le poison de l'amanite phalloïde...). Un ARN pré-messager complémentaire du brin matrice de l'ADN (brin antisens), donc identique au brin codant de l'ADN (brin sens), aux riboses et uraciles près, commence à être synthétisé selon la direction 5'-3' (voir fig. 26).

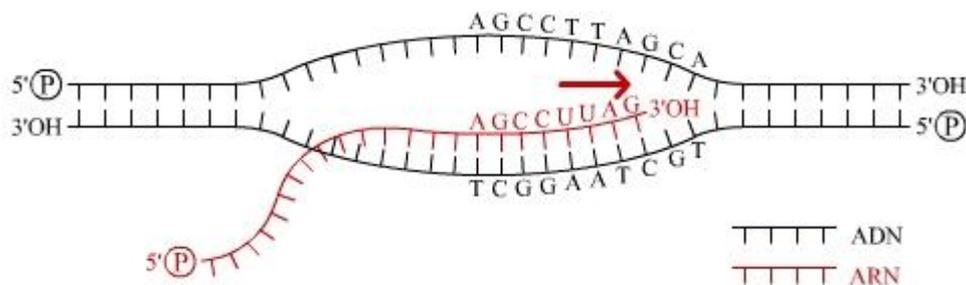


Figure 26 - La phase d'élongation de la transcription

L'ARN polymérase lit le brin patron (ou brin antisens), qui est complémentaire du brin codant (ou brin sens), depuis son extrémité 3' vers son extrémité 5'. Le brin d'ARN néosynthétisé est donc identique au brin codant d'ADN, aux uraciles et riboses près.

- **La terminaison**

L'ARNpol II est également équipée de facteurs protéiques de terminaison. Elle reconnaît ainsi un ou plusieurs signaux de terminaison portés par le brin progressivement parcouru et qui annoncent la fin de la transcription sur le brin d'ADN matrice (TTATTT par exemple, parfois aussi plus en aval ATACAAC...). Elle arrête bientôt son travail de transcription et libère l'ARNpm qu'elle vient d'assembler.

D. Maturation des ARNs :

Entre le moment où ils sont synthétisés et celui où ils jouent un rôle dans la synthèse des protéines, les ARNs subissent un certain nombre de modifications post-transcriptionnelles, on dit qu'ils subissent une maturation. Cette maturation est différente selon le type d'ARN:

* **Pour les ARNt :** Un gène d'ARNt donne un précurseur qui subira des clivages et des additions (addition de 3 nucléotides au niveau de l'extrémité 3', CCA) et d'autres modification comme des méthylations (méthylation de U en T) et des désaminations (désamination de l'A en Hypoxanthine).

* **Pour les ARNr :** Ces modifications correspondent à des clivages successifs conduisant à la perte de certains segments du transcrit initial.

* **Pour les ARNm :**

- Chez les procaryotes, il n'existe pratiquement pas de modification de l'ARNm néosynthétisé, d'ailleurs, la traduction de l'ARNm commence ne 5' avant même que la transcription ne soit achevée en 3'.

- Chez les eucaryotes, d'importantes modifications interviennent avant que l'ARNm ne puisse être traduits :

➤ **Addition du Poly (A) à l'extrémité 3' :**

Ce Poly (A) est une succession de nucléotides à adénines (A) (environ 25 nucléotides) et on pense qu'il aiderait au passage de l'ARNm du noyau vers le cytoplasme et qu'il protégerait l'ARNm au cours de la traduction.

Le site de polyadénylation est codé au niveau du gène. Le site de clivage déterminé par le dinucléotide CA est entouré par une séquence AAUAAA très conservée, située 10 à 30 nucléotides en amont du site de clivage, et par une séquence DSE (DownStream Element) riche en U ou en GU, situé une trentaine de nucléotides en aval du site de clivage. La séquence AAUAAA est reconnue spécifiquement par un complexe protéique appelé CPSF (Cleavage and Polyadenylation Specific Factor), et la séquence DSE par un complexe CstF (Cleavage Stimulation Factor). Ces deux complexes et d'autres composants, comme l'ARN polymérase II et la poly(A) polymérase (PAP), vont interagir en formant le complexe de clivage qui va cliver la molécule de préARN au niveau du site de clivage (voir fig. 27) (figure 29)..

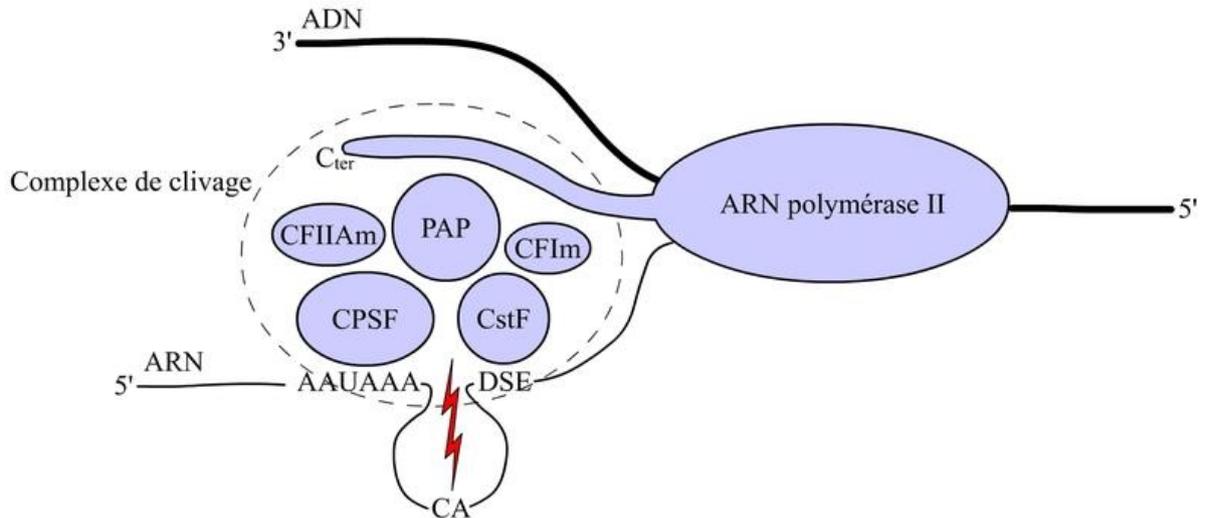


Figure 27 : Le complexe de clivage impliqué dans la maturation en 3' du préARNm

CPSF = *Cleavage and Polyadenylation Specific Factor*

CstF = *Cleavage Stimulation Factor*

PAP = *Poly(A) Polymerase*

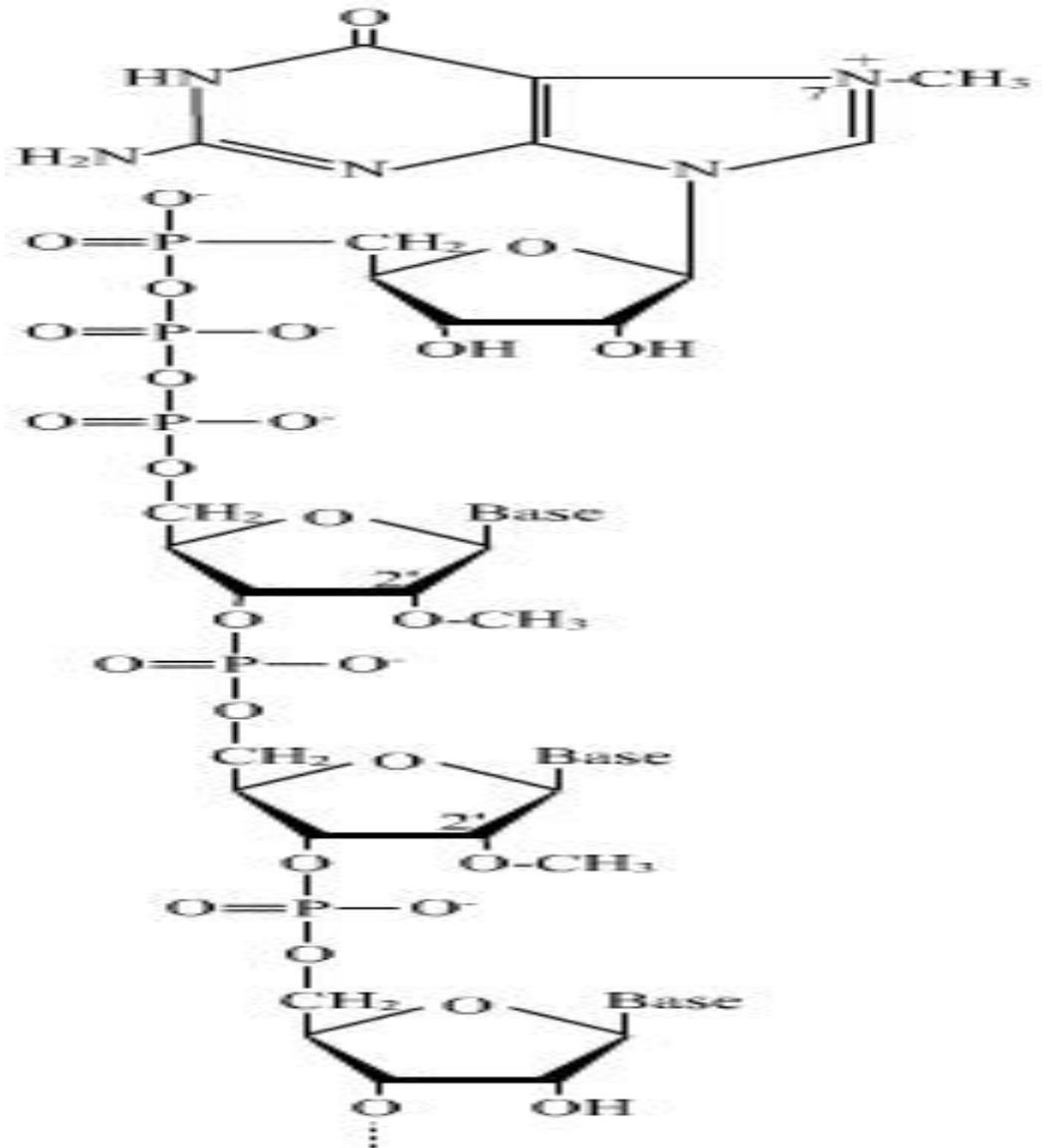
CF = *Cleavage Factor*

DSE = *Downstream Element*

➤ **Addition du CAP au niveau de l'extrémité 5' (CAP vient du terme capuchon) :**

Elle a lieu dès le début de la transcription avant que la chaîne ne compte plus de 30 nucléotides. Elle consiste en l'ajout d'un nucléotide à guanine sur l'extrémité 5' de l'ARN (L'ARNm n'aura donc pas d'extrémité 5' phosphate libre) suivi de sa méthylation sur l'azote 7 de la base, ainsi que de la méthylation en 2' du ribose du premier ou des deux premiers nucléotides du transcrit primaire. La particularité de cet ajout consiste dans le type de liaison mis en jeu : au lieu d'être reliés par une liaison ester-phosphorique entre le groupement OH porté par le carbone 3' du ribose du nucléotide à guanine et l'acide phosphorique alpha du premier nucléotide de l'ARN natif, les deux nucléotides sont reliés par une liaison anhydride d'acide entre les acides phosphoriques des deux nucléotides.

Il résulte de ces modifications que l'extrémité 5' de l'ARNm n'est pas porteuse des trois acides phosphoriques libres habituels, mais d'un GMP, ce qui limite la réactivité de cette extrémité et sa reconnaissance par les exonucléases (protection contre la dégradation). Cet



ensemble sert de coiffe protectrice à l'extrémité 5' de l'ARNm. Elle est également nécessaire à l'exportation de l'ARNm vers le cytoplasme et à la liaison de ce dernier avec la petite sous-unité du ribosome, lors de l'étape d'initiation de la traduction. Le CAP protégerait ainsi l'extrémité 5' de l'ARNm des attaques des enzymes (phosphatases et nucléases)(figure 28) (figure 29)..

Figure 28 : Addition du CAP au niveau de l'extrémité 5'

➤ **L'excision des introns**

L'autre modification observée dans de nombreux cas correspond à l'excision des **introns** (fragments d'ADN internes qui sont dites non codantes), les parties conservées sont appelées **exons** (parties codantes) (figure 29).

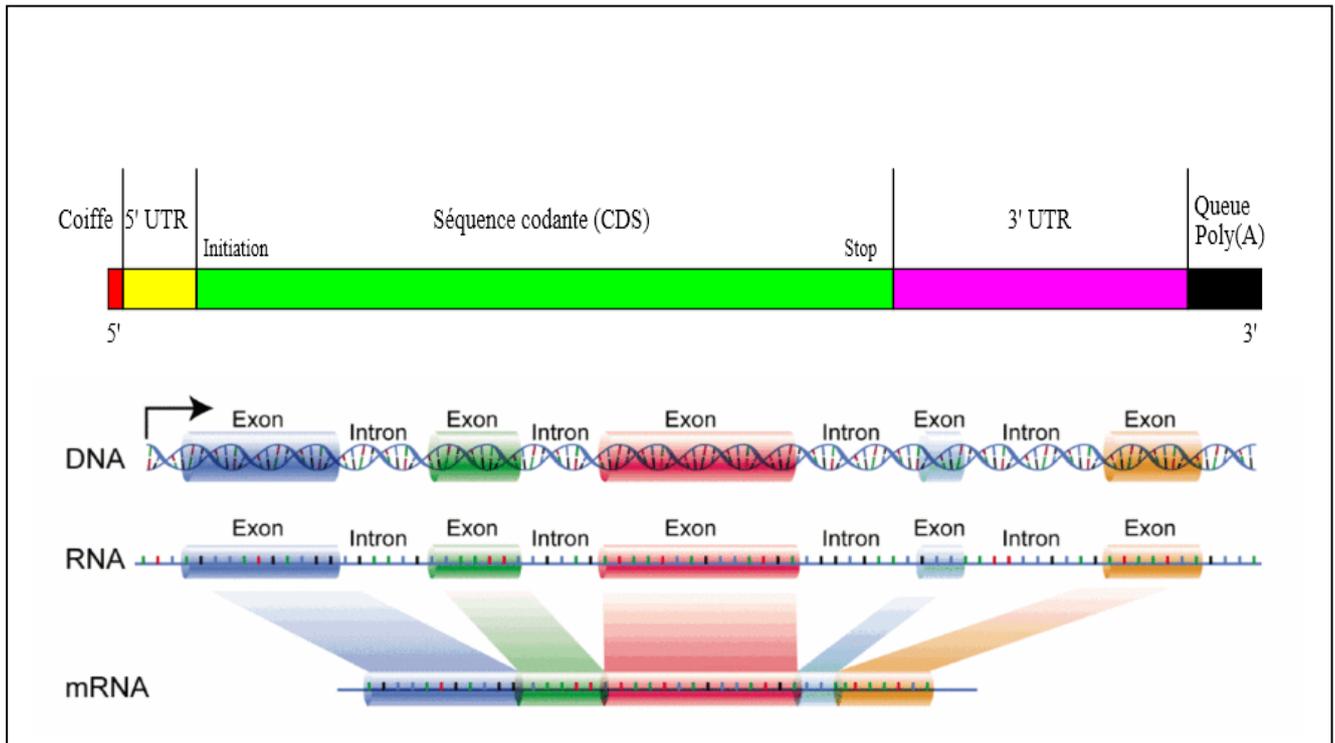


Figure 29 : Structure d'un ARN messager typique d'eucaryote, comprenant la coiffe, la région 5' non traduite, la région codante entre le codon d'initiation et le codon stop, la région 3' non traduite, la queue de poly(A) et L'ADN est transcrit en ARN qui, chez les eucaryotes, est épissé en ARN messager.

3.2. La traduction :

A. Le code génétique :

La seconde étape dans la synthèse d'un polypeptide consiste en la traduction de l'information portée par l'ARNm. Le code génétique (voir tableau du code génétique) permet de passer du langage **nucléique** élaboré par les 4 signes représentés par les 4 bases de l'ARNm, au langage **protéique** élaboré par 20 signes représentés par les 20 acides aminés susceptibles d'entrer dans la constitution d'un polypeptide.

L'unité élémentaire de ce code est le **codon**, formé par trois bases successives sur l'ARNm. Puisqu'il y a 4 nucléotides différents et que chaque codon en comporte 3, il existe au total $4^3 = 64$ codons possibles. Comme il y a 20 acides aminés il y a donc 44 codons supplémentaires. 3 codons correspondent à des codons STOP ou non-sens, les autres sont des synonymes qui codent pour différents acides aminés. Le code génétique est dit **dégénéré** puisque chaque acide aminé est codé par plusieurs codons (à l'exception du Tryptophane et de la méthionine). L'initiation des chaînes protéiques est assurée par le codon AUG qui code pour la méthionine, les triplets UAA, UAG et UGA sont des signaux qui assurent la terminaison des chaînes.

Les codons sont lus dans le sens 5'→3' le long de la chaîne de l'ARNm. A un codon correspond un **anticodon** (séquence de 3 nucléotides successifs sur un ARNt) et donc à un acide aminé spécifique

- **Initiation**

Près de l'extrémité 5' phosphate de l'ARNm se trouve un codon signal qui indique que la traduction doit débiter, on l'appelle **codon initiateur**. Ce codon est presque toujours AUG qui code pour la méthionine, donc toutes les chaînes protéiques en cours de synthèse ont la Met comme premier acide aminé, cependant cette Met sera enlevée juste après la synthèse peptidique.

Juste avant que la traduction ne commence, le ribosome n'est pas constitué, les 2 sous unités sont en effet dissociées et libres dans le cytoplasme. De nombreux éléments participent au recrutement des ribosomes (coiffe, queue poly-A, diverses protéines...).

A la phase d'initiation, la petite sous unité forme un complexe avec l'ARNm (au niveau du codon AUG) d'une part, et avec l'ARNt portant la méthionine initiale d'autre part. La

grande sous unité s'ajoute alors, le ribosome est maintenant constitué et fonctionnel. Les ribosomes ont deux sites de liaisons (figure 30).

Le site A (site acide aminé) : Où viendra se placer l'ARNt porteur de l'acide aminé.

Le site P (site peptidique) : Où viendra se placer l'ARNt porteur de la chaîne protéique en cours d'élongation.

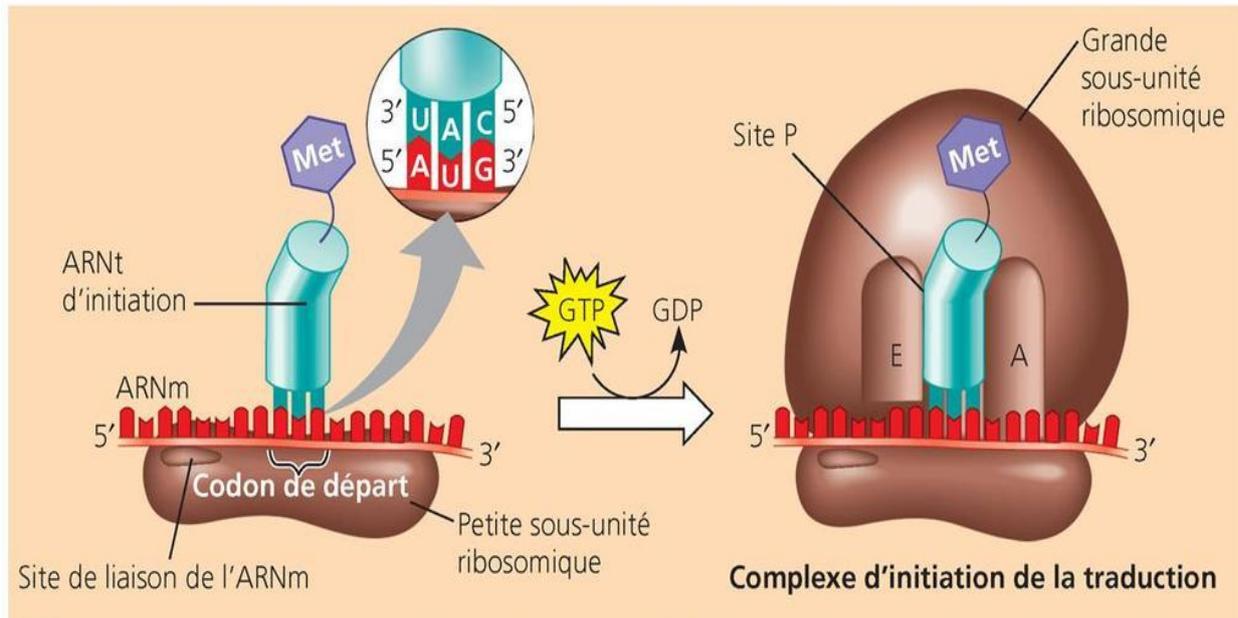


Figure 30 : Complexe d'initiation de la traduction

- **Elongation** :

Après initiation, le premier acide aminé alors en place, il va falloir maintenant au cours de la phase suivante, appelée élongation, former une liaison peptidique pour chaque acide aminé à accrocher, c'est-à-dire pour chaque liaison à fabriquer, un même cycle à 3 étapes est à chaque fois décrit :

a. Accrochage d'un nouvel aminoacylARNt dans le ribosome : Le deuxième ARNt vient avec l'acide aminé n°2 dans le site A de la grande sous unité, c'est le codon n°2 placé sur l'ARNm après le codon AUG qui détermine donc le choix du 2^{ème} anti-codon c'est-à-dire du 2^{ème} ARNt donc du 2^{ème} acide aminé.

b. La formation de la liaison peptidique :

Il y a rupture de la liaison entre la Met et le 1^{er} ARNt, c'est alors que se forme la liaison peptidique entre le COOH libre de la méthionine et le NH₂ de l'acide aminé n° 2 porté par l'ARNt n°2. Mais en fait, le COOH de la méthionine n'étant pas libre puisqu'il est engagé dans la liaison avec le 1^{er} ARNt, la formation de la liaison peptidique donnant le dipeptide et le détachement du 1^{er} ARNt se font simultanément (en même temps), c'est la peptidyltransferase qui intervient à ce stade. A ce moment il y a formation d'un dipeptide logé dans le site A et porté par l'ARNt n°2, l'ARNt n°1 est éjecté du site P et libéré dans le cytoplasme (figure 31).

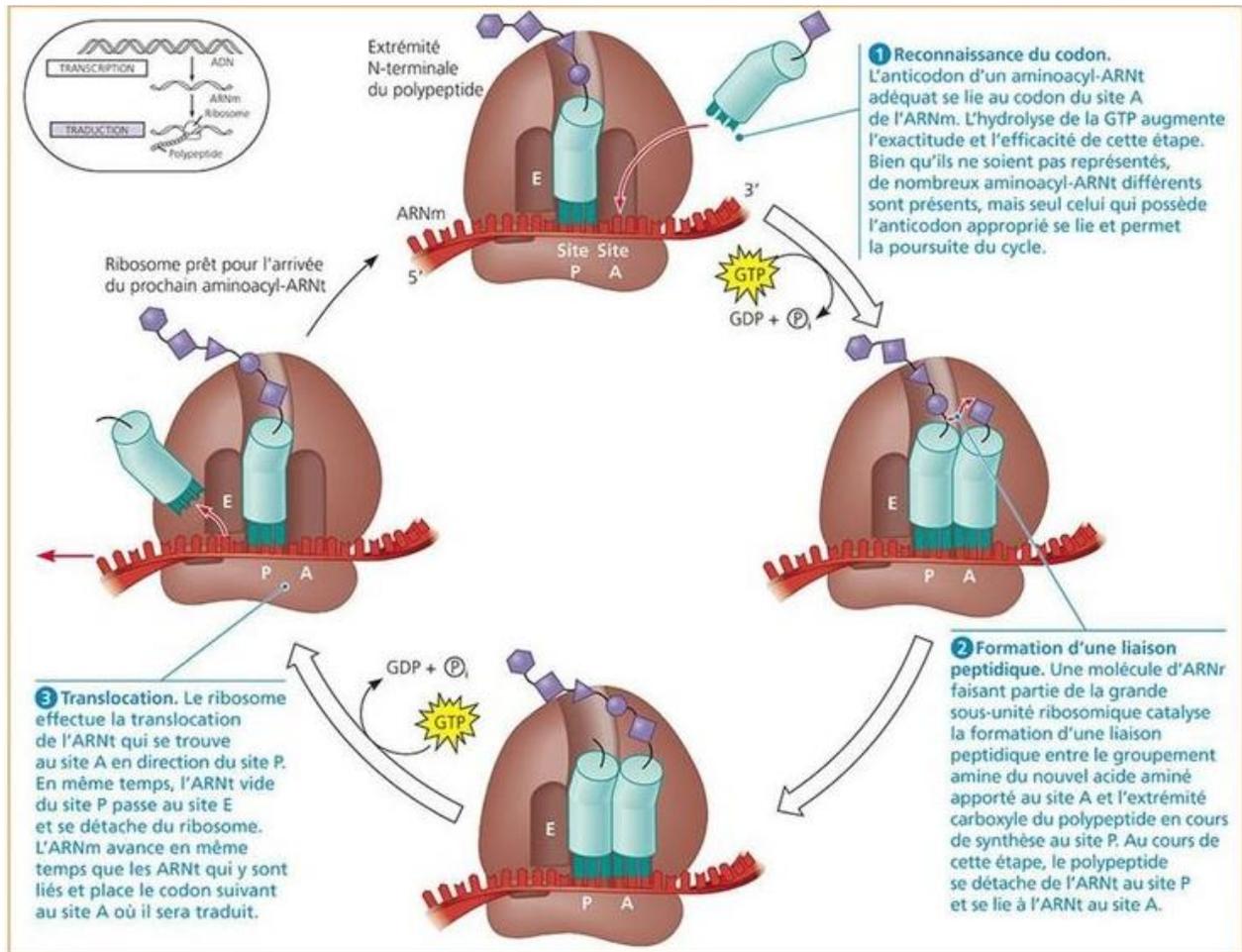


Figure 31 : Etapes d'élongation de la chaîne peptidique

c. Translocation

Le ribosome va avancer d'un cran sur l'ARNm dans la direction 5'→3'. Un cran veut dire 3 nucléotides ou codon. Un nouveau codon n°3 se trouve donc en face du site A, simultanément l'ARNt n°2 qui portait le dipeptide est passé du site A au site P, il a donc changé de loge, d'où le nom de translocation. De nombreux cycles vont se succéder avec à chaque fois les mêmes trois étapes. L'ARNm formé est lu par plusieurs ribosomes (séparés par

une centaine de nucléotides) et permet ainsi la synthèse de plusieurs molécules protéiques. Les ribosomes s'attachent à l'ARNm les uns après les autres et forment alors un polysome (= succession de ribosomes le long d'un même ARNm) (figure 32).

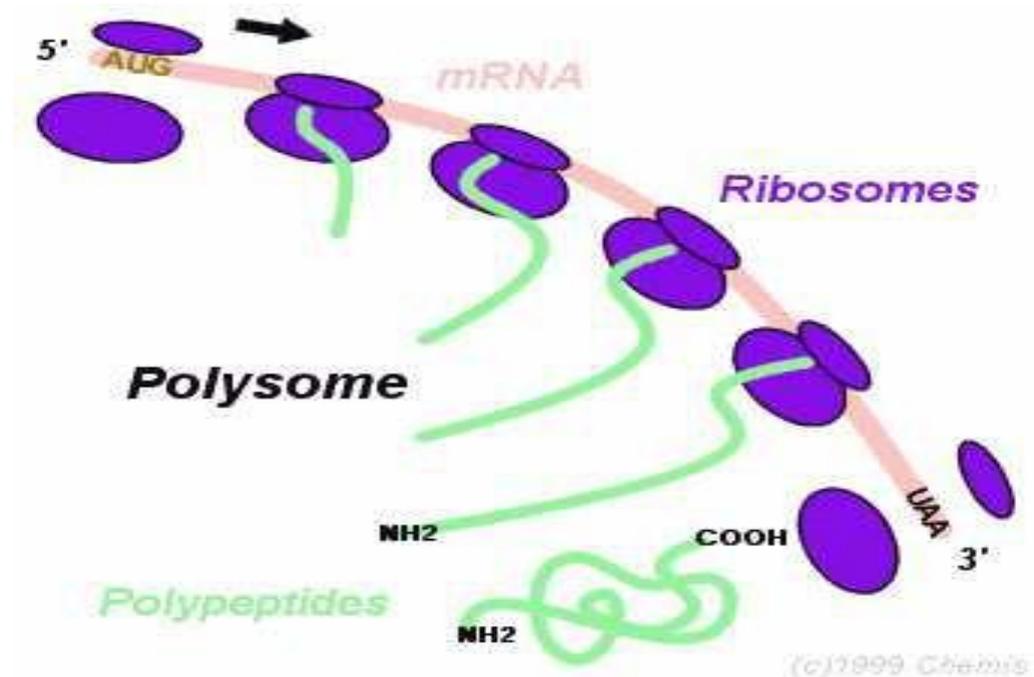


Figure 32 : formation du polysome

- **Terminaison**

La fin de la traduction se produit lorsque le ribosome en avançant d'un cran sur l'ARNm trouve un codon STOP, UAA, UAG ou UGA. Il n'existe aucun ARNt qui viendra dans le site A, il se produira alors une coupure de la liaison entre le dernier ARNt et la chaîne peptidique, la liaison qui unissait ce dernier ARNt au dernier acide aminé est hydrolysée libérant ainsi la chaîne peptidique. C'est la peptidyltransferase qui fera cette dernière coupure. Le ribosome se dissocié en 2 sous unités qui pourront recommencer de nouvelles lectures de l'ARNm (figure 33).

L'ARNm formé est lu par plusieurs ribosomes (séparés par une centaine de nucléotides) et permet ainsi la synthèse de plusieurs molécules protéiques. Les ribosomes s'attachent à l'ARNm les uns après les autres et forment alors un polysome (= succession de ribosomes le long d'un même ARNm).

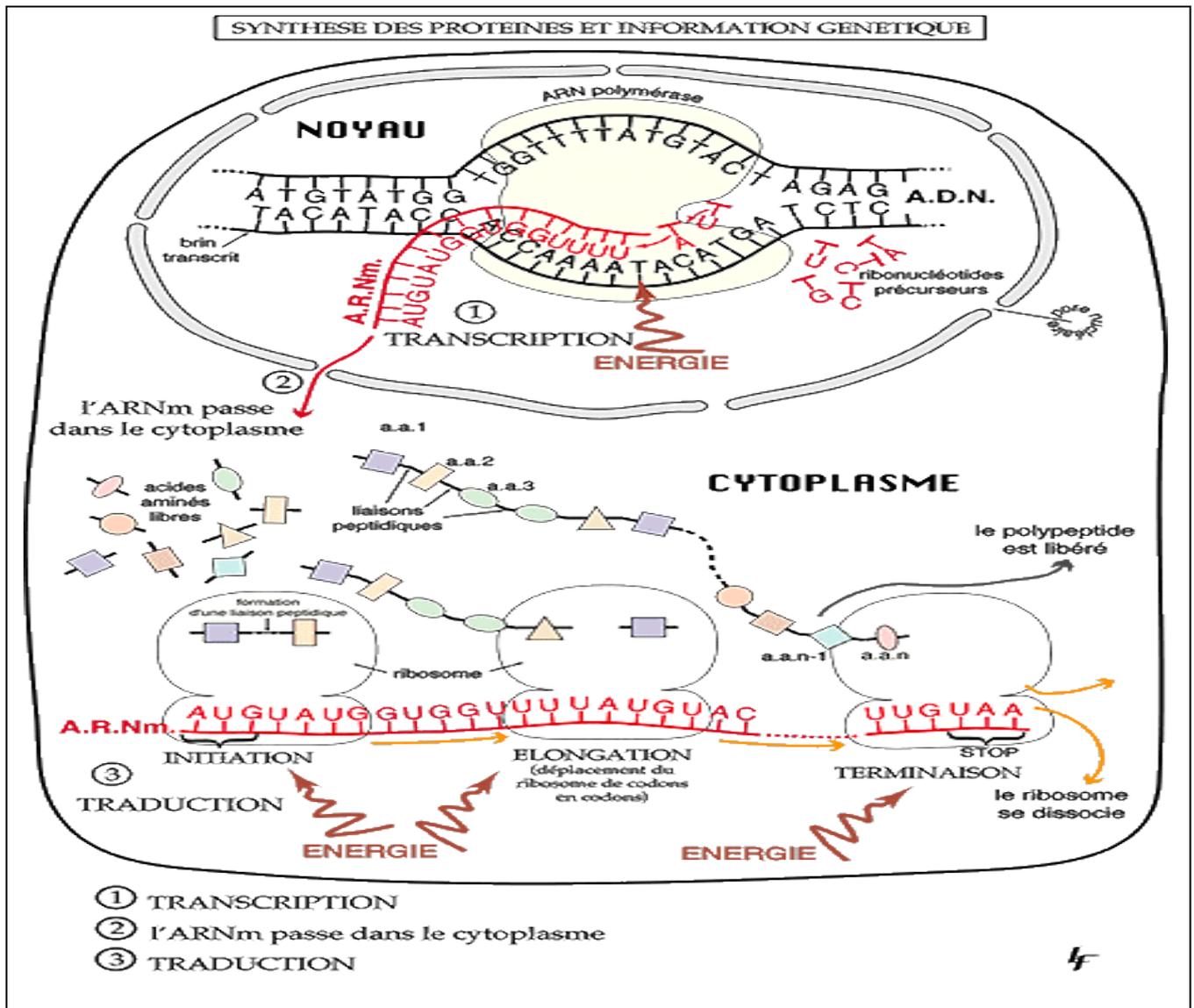


Figure 33 : Schéma résumé les étapes de la synthèse des protéines, transcription et traduction