

VII. Expression des gènes (SYNTHÈSE DES PROTÉINES)

7.1. Introduction

L'expression des gènes, encore appelée expression génique ou expression génétique, désigne **l'ensemble des processus biochimiques par lesquels l'information héréditaire** stockée dans un gène est lue pour aboutir à la fabrication de molécules qui auront un rôle actif dans le fonctionnement cellulaire, comme les protéines ou les ARN.

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

7.2. La transcription

- Au cours de la transcription, un brin d'ADN qui compose un gène, appelé brin non codant, sert de modèle pour la synthèse d'un brin d'ARN correspondant (complémentaire) par une enzyme appelée **l'ARN polymérase**. Ce brin d'ARN se nomme le transcrit primaire.
- La transcription est le processus de copie du matériel génétique (ADN ou ARN) en ARN. Chez les procaryotes une seule ARN-polymérase effectue la transcription pour tous les types d'ARN, 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).
- La transcription correspond à la synthèse d'une molécule d'ARN à partir d'une matrice d'ADN. Les gènes sont transcrits seulement quand leurs produits sont nécessaires pour la cellule. Ces produits correspondent soit à une chaîne polypeptidique soit à un ARN fonctionnel.

La transcription nécessite trois composants essentiels :

7.2.1. La matrice de transcription (le brin transcrit)

Le brin utilisé pour la transcription est appelé brin matrice. L'autre brin n'est normalement pas transcrit. La transcription produit une molécule d'ARN qui a la même polarité et la même séquence de bases que le brin qui n'a pas servi de matrice, sauf que l'ARN contient U à la place de T. Pour cette raison, le brin non transcrit est appelé « brin sens ».

7.2.2. Le système de transcription

La transcription est effectuée par une ARN polymérase dont l'action est assistée par plusieurs protéines auxiliaires qui s'associent à la polymérase à différentes étapes du processus. Il s'agit principalement des facteurs de transcription (TFI, TFII...).

7.2.3. Les substrats de la transcription

L'ARN est synthétisé à partir de ribonucléosides triphosphates (rNTP) qui sont ajoutés un à un à l'extrémité 3'-OH de la chaîne en formation. Deux groupements phosphate sont clivés du rNTP entrant, et le groupement phosphate restant forme une liaison phosphodiester qui attache le nucléotide à la molécule d'ARN en croissance. Les nucléotides sont toujours ajoutés à l'extrémité 3' de la molécule, donc le sens de transcription est de 5' vers 3'.

7.3. Les mécanismes de la transcription ADN

Les étapes de la transcription, la transcription se déroule en trois étapes :

- L'initiation ;
- L'élongation ;
- La terminaison.

7.3.1. Transcription chez les procaryotes

Le processus de transcription de l'ADN peut être divisé en 3 étapes principales : initiation, élongation et terminaison. Ces étapes sont également impliquées dans la [réplication](#) de l'ADN.

La transcription est catalysée par l'enzyme ARN polymérase, qui se fixe et se déplace le long de la molécule d'ADN jusqu'à ce qu'elle reconnaisse une [séquence promotrice](#). Cette zone d'ADN indique le point de départ de la transcription, et il peut y avoir plusieurs séquences de [promoteur](#) dans une molécule d'ADN. Les facteurs de transcription sont des [protéines](#) qui contrôlent le taux de transcription; elles aussi se lient aux séquences promotrices avec l'ARN polymérase.

Une fois liée à la séquence promotrice, l'ARN polymérase déroule une partie de la double [hélice](#) d'ADN, exposant les [bases](#) sur chacun des deux brins d'ADN.

A. 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.
- 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

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.

- 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.
- 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.

B. 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

C. 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.

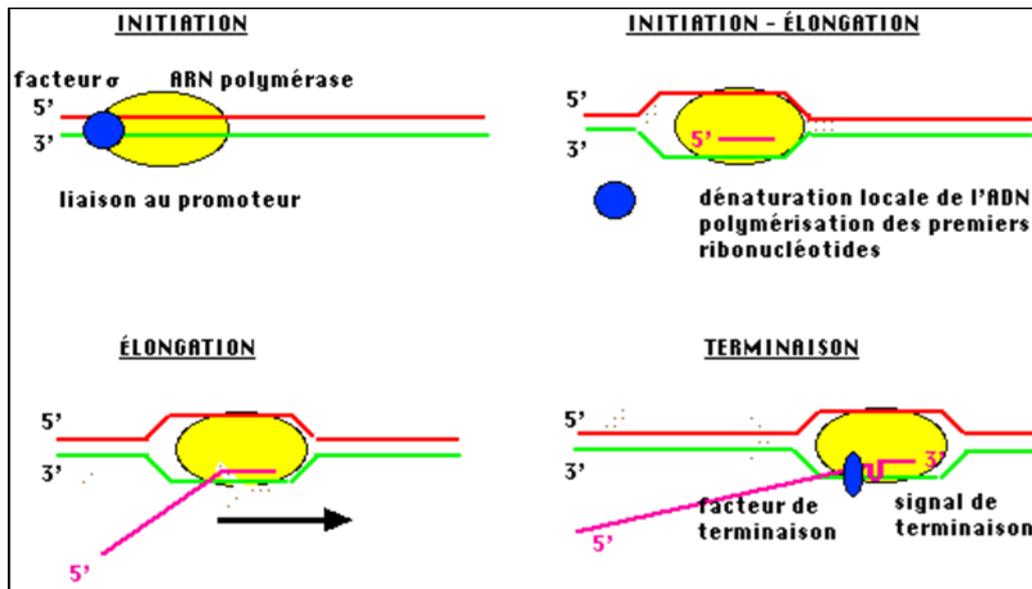


Figure 1 : La Transcription chez les procaryote

7.4. Transcription chez les eucaryotes

Dans le cas des eucaryotes, la transcription est effectuée dans le noyau, et est similaire à celle des procaryotes, mais de plus grande complexité. Différents ARNp transcrivent différents types de gènes. L'ARNp transcrit les pré-ARNm, tandis que l'ARNp et les ARNp transcrivent les ARN ribosomiaux et les ARNt, respectivement.

Les ARN transcrits sont ensuite modifiés. Le pré-ARNm, par exemple, subit un processus de maturation qui après des coupes et des épissures successives élimine certains segments de l'ARN appelés introns pour produire l'ARNm final.

A. Initiation

- Contrairement à ce qui se passe chez les procaryotes, l'ARNpol II des eucaryotes ne reconnaît pas seul le promoteur proximal.
- Elle effectue ce travail en compagnie de nombreux cofacteurs 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 .
- 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 TFIIH 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.
- 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.

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).

Intervention de facteurs spécifiques 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 des 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.

B. 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).

- 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.

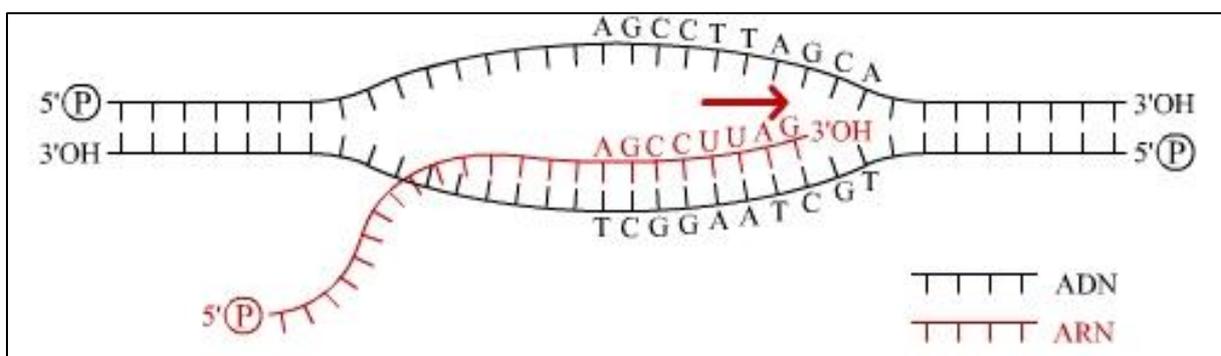


Figure 2 : L'élongation chez les eucaryotes

C. 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'ARN pm qu'elle vient d'assembler.

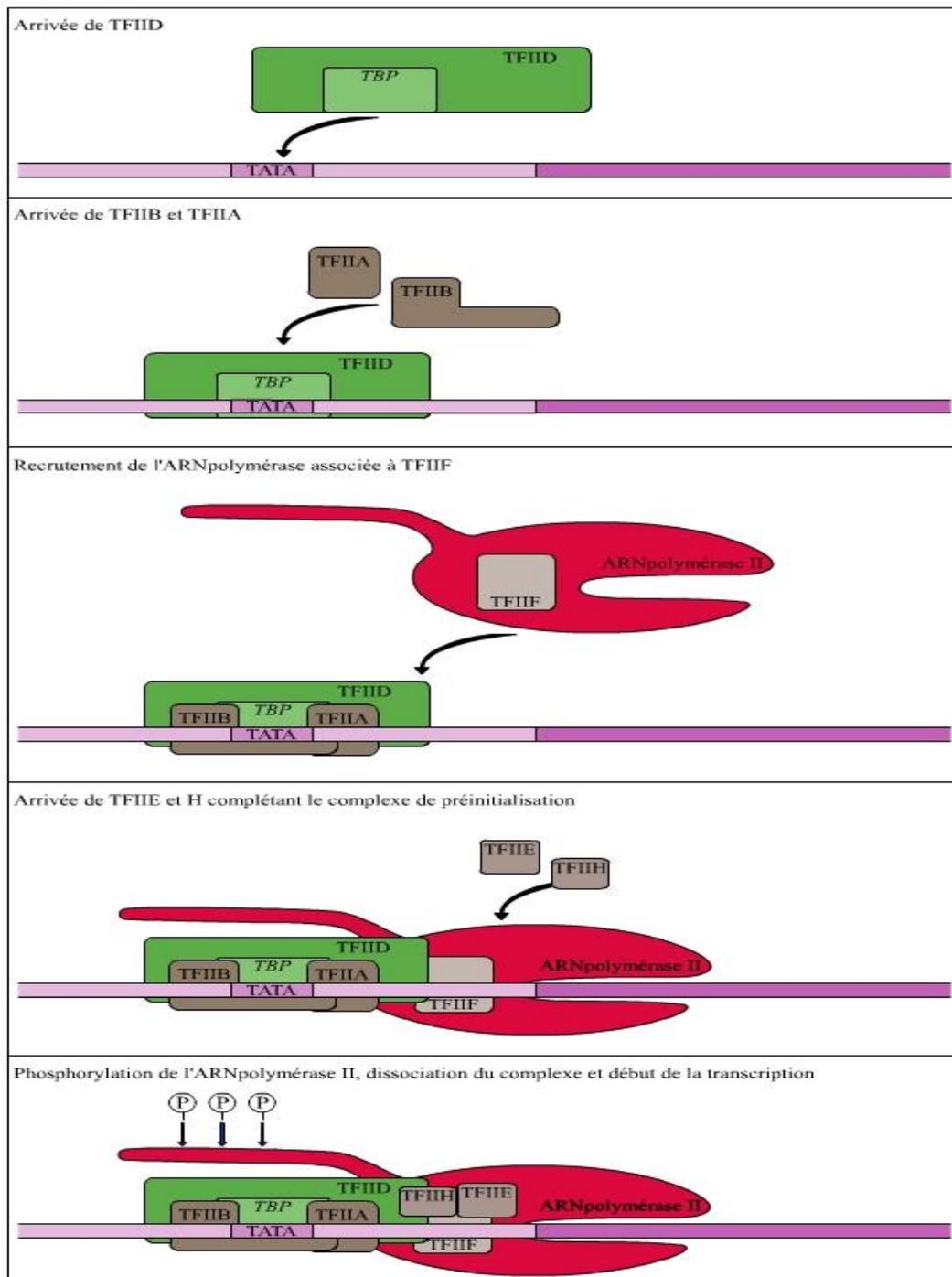


Figure 3 : La transcription chez les eucaryotes

7.5. 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 posttranscriptionnelles, 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'.

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 codon 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.

		Seconde lettre				
		U	C	A	G	
Première lettre	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G

Figure 4 : Le code génétique

A. 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 .

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.

B. 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 2ème anti-codon c'est-à-dire du 2ème ARNt donc du 2ème acide aminé.

b. La formation de la liaison peptidique :

- Il y a rupture de la liaison entre la Met et le 1er 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 1er ARNt, la formation de la liaison peptidique donnant le dipeptide et le détachement du 1er 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.

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).

C. 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 viendrait 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 dissocie en 2 sous unités qui pourront recommencer de nouvelles lectures de l'ARNm
- 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).

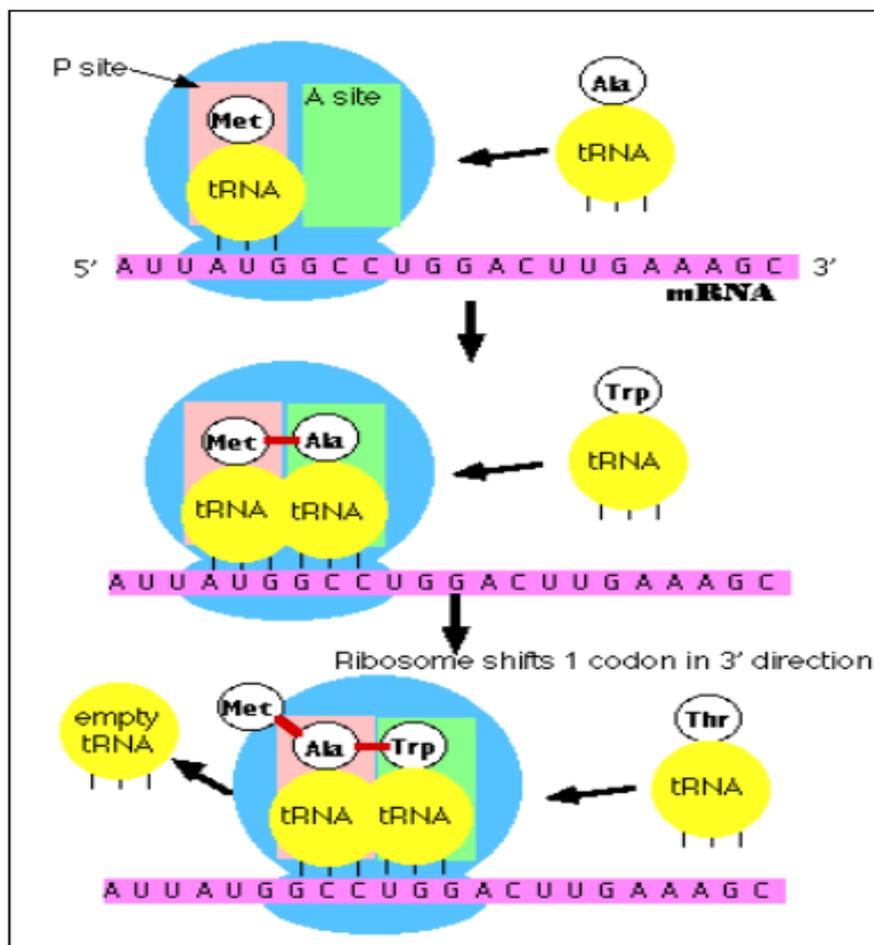


Figure 5 : La traduction