

BIOLOGIE MOLECULAIRE

MODULE 2

INTRO

Selon la théorie du **dogme central de la biologie moléculaire** énoncée par **Francis Crick en 1955** :

- Le flux de l'information génétique dans la cellule est unidirectionnel
- L'ADN (acide désoxyribonucléique) est le substrat biochimique de l'hérédité : il contient toutes les informations nécessaires à la cellule et il est capable de s'auto-répliquer pour assurer la transmission du patrimoine héréditaire.

L'information génétique contenue dans l'ADN va ensuite être transcrite dans un ARN (acide ribonucléique).

L'ARN messager va jouer le rôle d'intermédiaire entre le noyau et les ribosomes auxquels il va délivrer son information.

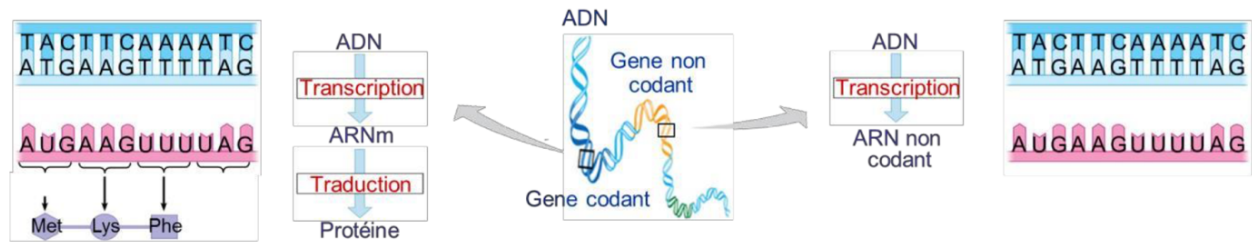
Exceptions! Cas spéciaux de transferts : par exemple des transferts rétrograde d'information de l'ARN vers l'ADN comme on peut l'observer chez les rétrovirus, ou encore de la capacité qu'ont certaines plantes ou certains virus à produire de l'ADN à partir d'ARN.



I - PRINCIPES GÉNÉRAUX DE L'EXPRESSION D'UN GÈNE : TRANSCRIPTION /TRADUCTION

Le matériel génétique ou génome contient les gènes → un gène contient une information.

Gène = enchaînement linéaire de nucléotides formant une séquence d'ADN délimitée par un signal de début "START" et par un signal de fin "STOP"



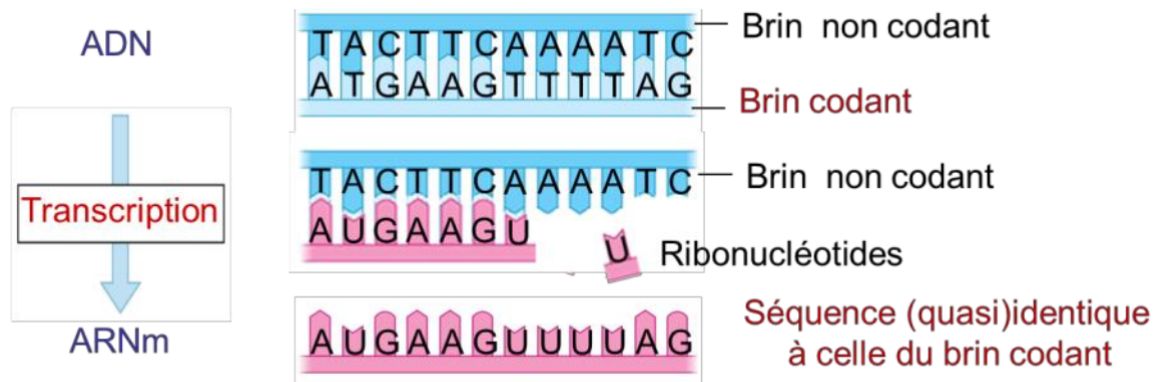
1^{ère} étape : La transcription

Il s'agit de retranscrire la séquence de désoxyribonucléotides du gène en une séquence de **ribonucléotides** qui sera retrouvée dans l'**ARN messenger**.

Le **brin codant** contient l'**information** qui doit être retranscrite dans l'ARN messenger.

Le **brin non codant** sert de **matrice** pour transcrire quasiment à l'identique l'information du brin codant dans l'ARN messenger.

<p>Les gènes codants : ils servent à la synthèse des protéines et leur séquence de désoxyribonucléotides va être tout d'abord transcrite en séquence de ribonucléotides que l'on retrouvera dans l'ARNm, puis traduite en une séquence d'acides aminés pour former une protéine.</p> <p>Ils vont subir dans leurs expression deux étapes : une étape de transcription, puis une étape de traduction</p>	<p>Les gènes non codants : ils servent uniquement à la synthèse d'ARNs non codant comme les ARNs ribosomiaux, les ARNs de transfert, les petits ARNs nucléaires ou nucléolaires.</p> <p>Ils vont donc être uniquement transcrit. Il n'y aura pas dans son expression d'étape de traduction.</p>
--	--



La transcription est assurée par une **ARN polymérase** : enzyme capable de synthétiser de l'ARN à partir d'ADN.

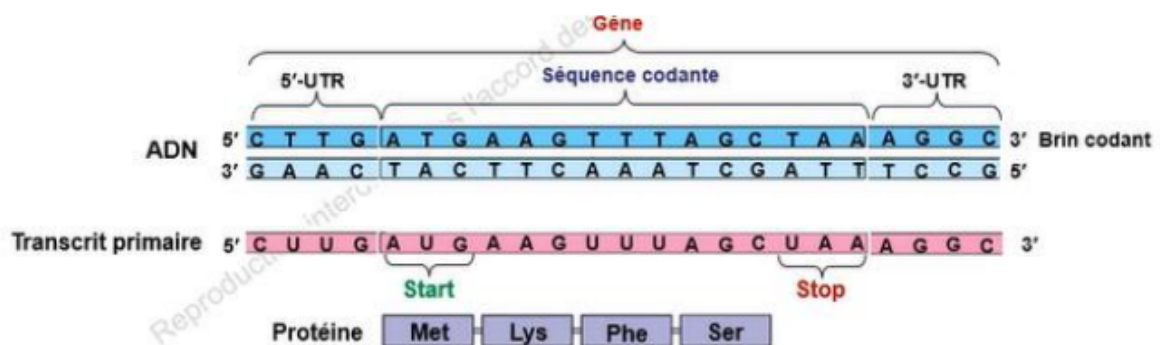
Le transcrit obtenu est appelé transcrit primaire et sera utilisé tel quel chez les procaryotes, mais devra subir des étapes de maturation chez les eucaryotes. (vous verrez tout ça plus tard)

Un gène contient aussi des séquences dites « non codantes ».

Ces séquences encadrent en 5' et en 3' la séquence codante. Cependant elles ne seront pas traduites.

Elles sont appelées « **5'-UTR et 3'-UTR (Untranslated)** ».

L'ARN polymérase débute la transcription en amont de la séquence codante et l'achève en aval : elle produit un ARN plus grand que celui qui correspond à la séquence codante du gène.



2^{ème} étape : La traduction

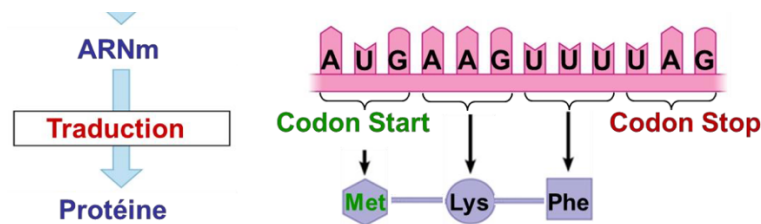
L'expression d'un gène se termine par sa traduction dans le cytoplasme.

Il s'agit de quoi?

La traduction consiste à décoder le message de l'ARNm afin de former une protéine

Les **ribonucléotides** sont lus trois par trois → chaque triplet de nucléotides forme un **codon**.

La traduction débute au niveau d'un codon appelé **codon START** pour s'achever au niveau d'un **codon STOP**.



Elle repose sur le code génétique qui indique à quel acide aminé correspond chaque codon de l'ARNm (donc 3 nucléotides codent un acide aminé)

C'est bien le code génétique qui permet de déchiffrer l'information de l'ARNm.

		Second Letter				
		T	C	A	G	
First Letter	T	TTT } Phe TTC } TTA } Leu TTG }	TCT } Ser TCC } TCA } TCG }	TAT } Tyr TAC } TAA Stop TAG Stop	TGT } Cys TGC } TGA Stop TGG Trp	T C A G
	C	CTT } Leu CTC } CTA } CTG }	CCT } Pro CCC } CCA } CCG }	CAT } His CAC } CAA Gln CAG }	CGT } Arg CGC } CGA } CGG }	T C A G
	A	ATT } Ile ATC } ATA } ATG Met	ACT } Thr ACC } ACA } ACG }	AAT } Asn AAC } AAA Lys AAG }	AGT } Ser AGC } AGA Arg AGG }	T C A G
	G	GTT } Val GTC } GTA } GTG }	GCT } Ala GCC } GCA } GCG }	GAT } Asp GAC } GAA Glu GAG }	GGT } Gly GGC } GGA } GGG }	T C A G

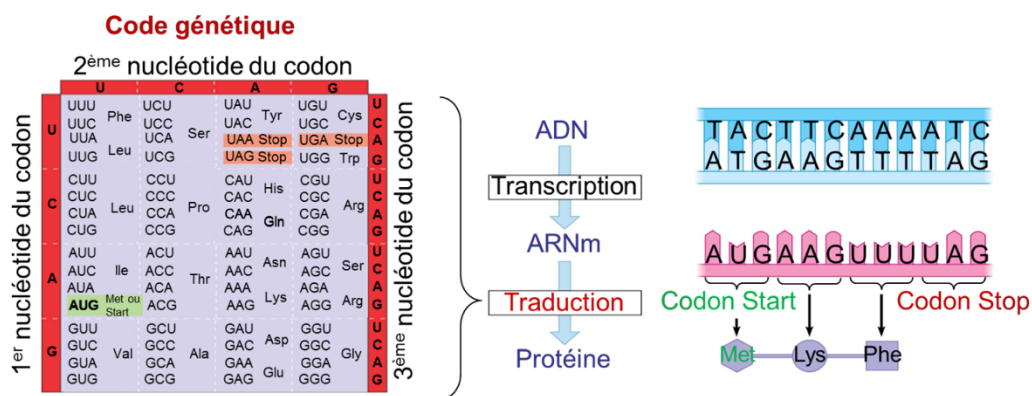
Ce tableau liste toutes les possibilités de codons différents et donne leur correspondance.

Au total, il existe $4^3 = 64$ combinaisons de trois nucléotides pouvant former un codon.

Parmi ces 64 combinaisons, quatre d'entre elles restent assez particulières :

Le **codon AUG** qui code pour la **méthionine** initie toujours la traduction et joue donc le **rôle de codon START**. Il peut également se trouver ailleurs dans la séquence d'un ARNm où il prendra alors le même sens.

Trois codons : **UAA ; UAG ; UGA** ne codent pour aucun acide aminé et indiquent la fin de la traduction de la protéine (**codon STOP**).



Le **code génétique** possède 4 caractéristiques majeures :

Quasi-universel	La plupart des espèces vivantes utilisent la même correspondance entre codons et acides aminés. Cependant, il subsiste de rares exceptions, telles que les mitochondries qui reposent sur le sens de quelques codons.
Non-chevauchant	Chaque nucléotide de l'ARNm ne peut appartenir qu'à un seul codon. L'ARNm est ainsi décodé selon un cadre de lecture fixe et précis.
Non-ambigu	Un codon donné correspond toujours au même acide aminé.
Dégénéré	Comme il existe un excès de codons par rapport au nombre d'acides aminés, la majorité des acides aminés sont spécifiés par

plusieurs codons différents SAUF pour la méthionine et le tryptophane.

CADRE DE LECTURE

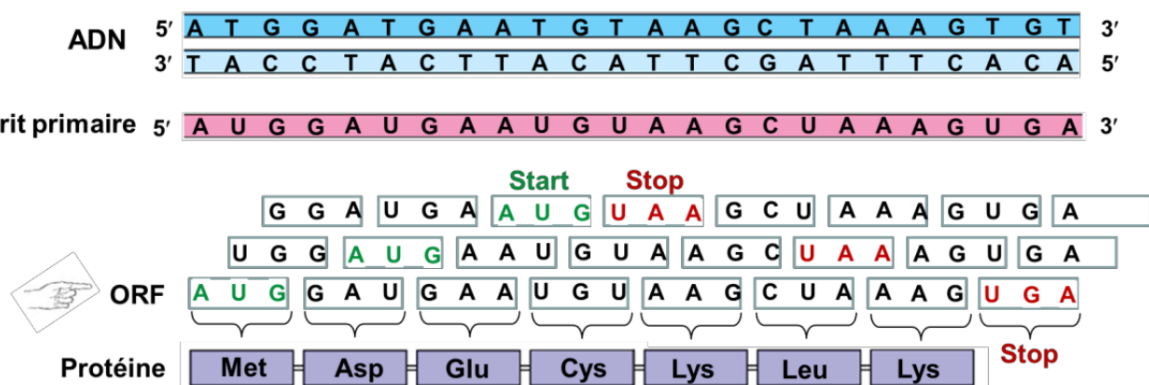
En théorie, il existe **trois cadres de lecture** pour déchiffrer la séquence de l'ARN messager. Chacun de ces cadres correspond à une lecture des nucléotides de l'ARN messager trois par trois, et ces cadres sont **décalés les uns par rapport aux autres d'un nucléotide**.

→ Leur décodage aboutirait à des protéines différentes.

En pratique, un **seul cadre de lecture va être utilisé** :

Ce **cadre de lecture** est dit **ouvert** ou **ORF (Open Reading Frame)** et est celui qui va utiliser le codon initiateur, le **codon Start** qui code pour la méthionine, et ainsi on obtiendra une seule et unique protéine.

Les deux autres *cadres théoriques sont dits bloqués*, car ils contiennent généralement un codon Stop prématuré et ainsi, leur traduction aboutirait à des *protéines tronquées*.



Comment la cellule va-t-elle reconnaître ce cadre ouvert de lecture?

⇒ Par **l'intermédiaire du ribosome** qui va reconnaître une **séquence spécifique**

- chez les **procaryotes** elle est appelée **séquence de Shine-Dalgarno**
- chez les **eucaryotes** elle est appelée **séquence de Kozak**

MUTATIONS

Il existe différentes **mutations** du code génétique qui amènent diverses conséquences :

- 1- La substitution est le **remplacement** d'un nucléotide dans un codon par un autre.
Cela peut amener à 3 types de mutations :

La mutation synonyme	La mutation faux-sens	La mutation non-sens
Est dite neutre car elle ne change ni l'acide aminé codé ni la protéine synthétisée après la traduction (ex : GGT changé en GGU = Gly dans les deux cas).	Elle change le sens du codon et l'acide aminé dans la séquence de la protéine.	Elle crée un codon qui interrompt la traduction . En effet, elle remplace le codon de départ par un codon STOP prématuré. La protéine est alors dite « tronquée ».

- 2 - Il existe également des **insertions** et des **délétions** dans le code génétique : il s'agit d'ajouter ou supprimer des nucléotides.

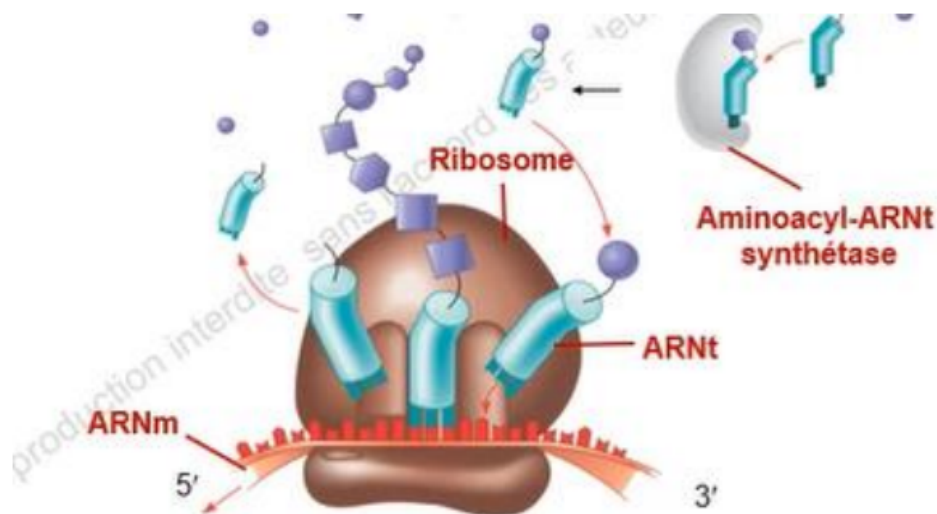
(Rappel : un codon = 3 nucléotides)

Si multiple de 3	Si non multiple de 3
Mutation non décalante : car ajout ou retrait d'un ou plusieurs codon(s) sans effet sur les codons précédant. Le cadre de lecture est respecté. La conséquence est un ajout ou un retrait d'acides aminés sur la protéine.	Le cadre de lecture de l'ARNm sera décalé d'un ou deux nucléotides. Il y aura la présence de faux sens multiples voire la modification de la position du Codon Stop ce qui pourra selon les cas aboutir à une protéine raccourcie ou allongée.

Les acteurs et le déroulement de la traduction

1- Les acteurs de la traduction

L'ARNm	Contient les instructions pour la synthèse de la protéine.
Les ARNt (ARN de transfert)	Portent leur unique acide aminé spécifique, ils se fixent sur le codon de l'ARNm correspondant à ce dernier et apportent l'acide aminé au ribosome.
Les aminoacyl-ARNt synthétases	Fixent les acides aminés sur les ARNt.
Les ribosomes , formés de protéines et d'ARN ribosomaux	Constitués de deux sous-unités - accueillent les ARNt chargés de leur acide aminé. Ensuite, ils relient entre eux ces acides aminés par des liaisons peptidiques afin de former la protéine.



Maintenant plus en détail...

LES ARN DE TRANSFERT

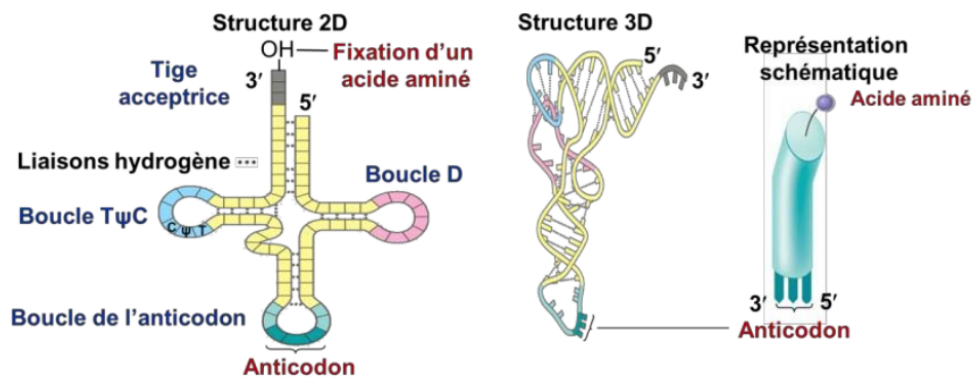
Les ARNt apportent les acides aminés au ribosome.

Ce sont des molécules qui sont formées :

- D'une **tige acceptrice** : elle peut être chargée avec un acide aminé à son extrémité 3'-OH.
- De **trois boucles** : la **boucle TψC**, la **boucle D** et la **boucle de l'anticodon**

Cette boucle de l'anticodon contient une séquence de trois nucléotides qu'on appelle l'anticodon. Cette séquence va être **spécifique de chaque ARN de transfert**.

C'est par l'intermédiaire de cette séquence que l'ARN de transfert va venir se fixer par complémentarité au codon de l'ARN messager qui spécifie l'acide aminé fixé sur l'ARN de transfert.



Les ARNt vont être produits à partir de gènes non codant, tout d'abord sous la forme de précurseurs ou pré-ARNt.

Ces pré-ARNt vont devoir subir une étape de maturation → modifications de nombreuses bases. Environ 10% à 25 % des bases l'ARNt vont être modifiées en ce qu'on appelle des bases mineures.

Ainsi, l'ARN de transfert mature va contenir des bases inhabituelles comme l'inosine, la pseudo-uridine, voire même la thymine qui normalement est spécifique de l'ADN + (c'est le cas notamment de la ribothymidine qu'on trouve dans la boucle TψC), ou encore la dihydrouridine.

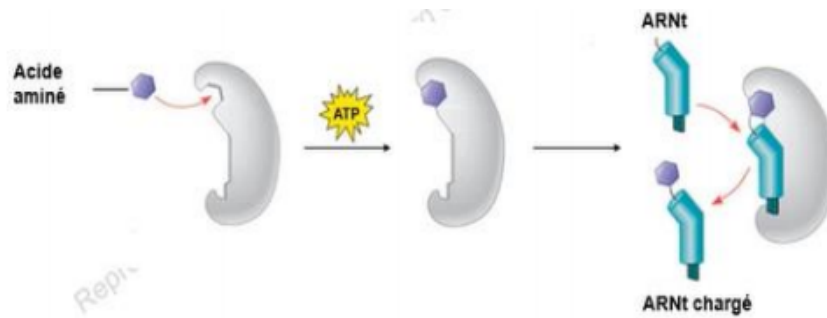
LES AMINOACYL-ARNt SYNTHÉTASES

Chaque aminoacyl-ARNt synthétase de la cellule est spécifique d'un seul et unique acide aminé

Les aminoacyl-ARNt synthétases se fixent de façon très spécifique les acides aminés aux ARNt → elle va d'abord l'activer grâce à l'ATP.

Elle est capable de le fixer sur plusieurs ARNs de transfert (**ARNt isoaccepteurs**).

Les **aminoacyl-ARNt synthétases** ont une activité de correction (proofreading), Cela leur permet d'éliminer un acide aminé fixé par erreur avant de libérer l'ARNt, ce qui évite son incorporation erronée et assure donc la fidélité de la traduction.



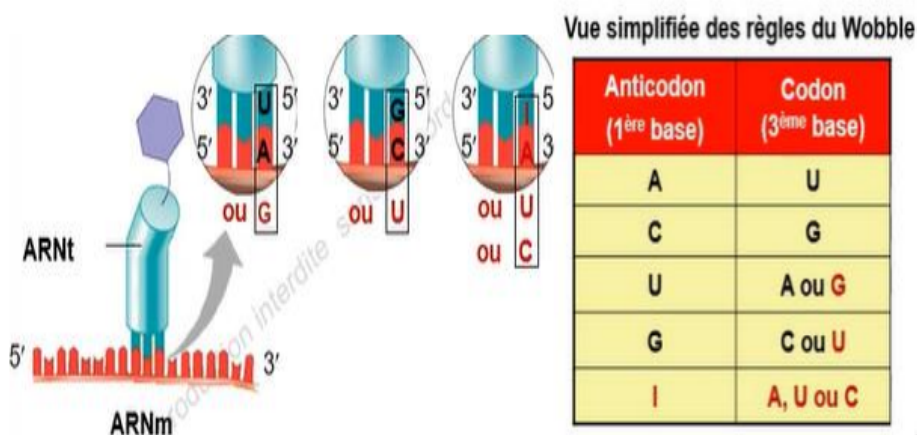
Un appariement flexible : le Wobble

Le déchiffrement du code utilise le Wobble comme un appariement flexible.

Cet appariement ne respecte pas le principe de complémentarité.

Cependant, la règle de l'appariement entre une purine et une pyrimidine est dans la majorité des cas respectée.

Le Wobble permet à l'anticodon d'un ARNt de s'apparier avec plusieurs codons qui spécifient le même acide aminé (codons synonymes) afin de réduire le nombre d'ARNt nécessaire à la traduction.



LE RIBOSOME

Structure macromoléculaire constituée de **2 sous unités** – elle va assurer la traduction.

- ♦ **La petite sous unité** : elle a pour rôle de **se fixer à l'ARNm**. C'est elle qui va décoder l'information contenue dans l'ARNm en s'assurant de la **correspondance entre codons et anticodons**.
- ♦ **La grosse sous unité** : elle se fixe à la petite sous unité et possède à la fois un rôle structural et fonctionnel.

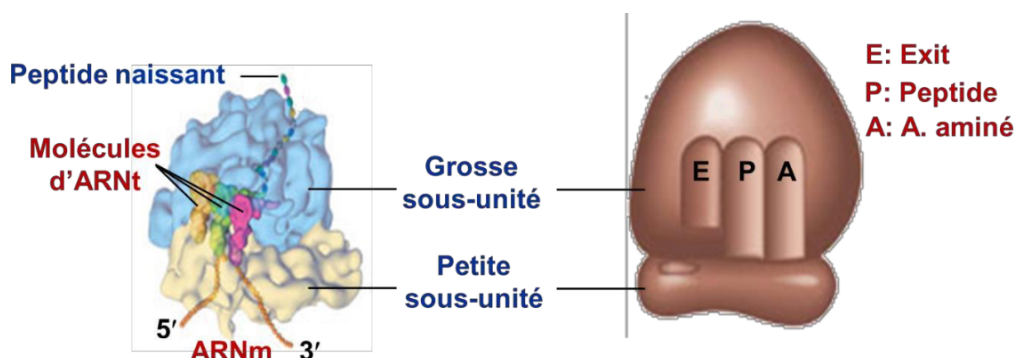
Cette grosse sous unité contient en effet **3 sites E, P et A** qui vont avoir pour but d'accueillir les ARNt.

⇒ La cavité **A "acide aminé"** : celle par laquelle un ARN de transfert chargé de son acide aminé va pénétrer à l'intérieur du ribosome (moi je retenais aussi A pour arrivée)

⇒ La cavité **P pour "peptide"** : celle au niveau de laquelle va être positionné le peptide en cours de synthèse

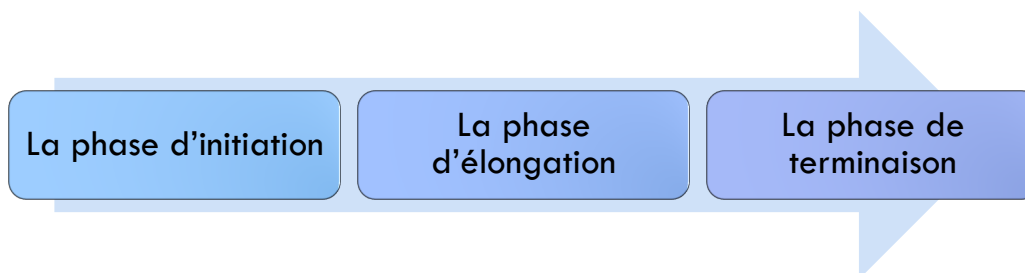
⇒ La cavité **E pour "exit"** : celle par laquelle les ARNt vont sortir du ribosome

La grosse sous unité contient un ARN ribosomal particulier qui joue le **rôle d'enzyme** formant les **liaisons peptidiques** entre les acides aminés qui sont apportés par les ARNs de transfert.



LE DEROULEMENT DE LA TRADUCTION

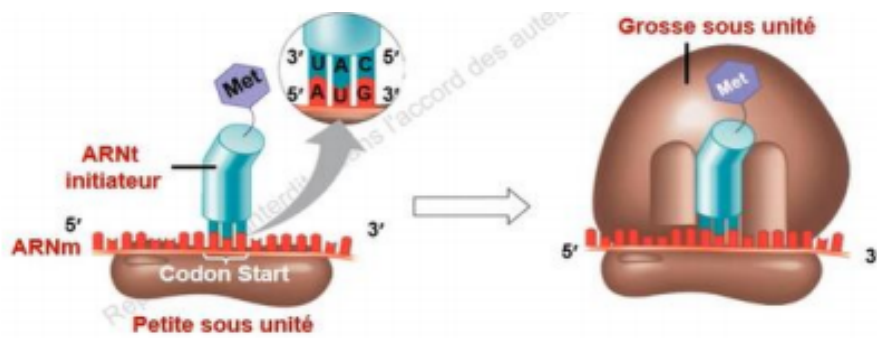
La traduction se fait en 3 phases



La phase d'initiation	La phase d'élongation	La phase de terminaison
Elle aboutit à l'assemblage du ribosome complet sur l'ARNm au niveau du Codon Start AUG qui indique le début de la séquence codante à traduire.	Elle correspond au déplacement du ribosome sur l'ARNm de codon en codon selon le cadre de lecture jusqu'au Codon Stop d'arrêt de la traduction. A chaque codon, un nouvel acide aminé apporté par un ARNt est incorporé au peptide en cours de synthèse par formation d'une liaison peptidique.	Elle correspond à la fin de la traduction avec libération de la protéine complète.

1^{ère} étape - La phase d'initiation comprend deux étapes

L'assemblage d'un complexe de pré-initiation sur l'ARNm	Il se forme au niveau du Codon Start chez les procaryotes et en amont chez les eucaryotes. Il comprend notamment la petite sous-unité et l'ARN de transfert initiateur portant la méthionine.
L'assemblage du ribosome complet au niveau du Codon Start	Il nécessite le déplacement du complexe de pré-initiation sur l'ARNm chez les eucaryotes. A ce stade, l'ARN de transfert initiateur et la méthionine sont positionnés au site P du ribosome.



2^{ème} étape - La phase d'élongation : une succession de cycles

Correspond au **déplacement du ribosome sur l'ARN messager selon le cadre de lecture** jusqu'au codon STOP d'arrêt de la traduction. Et à chacun des codons, un nouvel acide aminé va être incorporé au peptide en cours de synthèse par formation d'une liaison peptidique.

A chaque codon sur lequel va se positionner le ribosome, un ARN de transfert chargé d'un acide aminé va venir se positionner au niveau du **site A**.

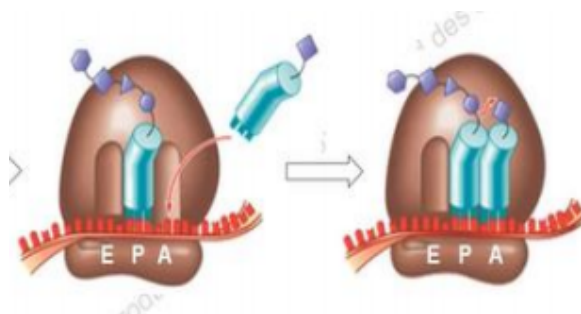
Si l'appariement codon-anticodon est correct, alors le **peptide est transféré sur l'acide aminé** par la formation d'une **liaison peptidique** \Rightarrow réalisée grâce à l'**activité peptidyltransférase** d'un ARN ribosomal de la grosse sous unité.

Le peptide va se trouver allongé d'un acide aminé, mais positionné cette fois-ci au niveau du site A du ribosome.

Puis, le ribosome se déplace d'un codon.

Le peptide allongé d'un acide aminé revient au site (P) et l'ARNt déchargé passe au site (E) et est éjecté.

Le cycle recommence ainsi par arrivée au site (A) ribosomal d'un nouvel ARNt chargé d'un AA.

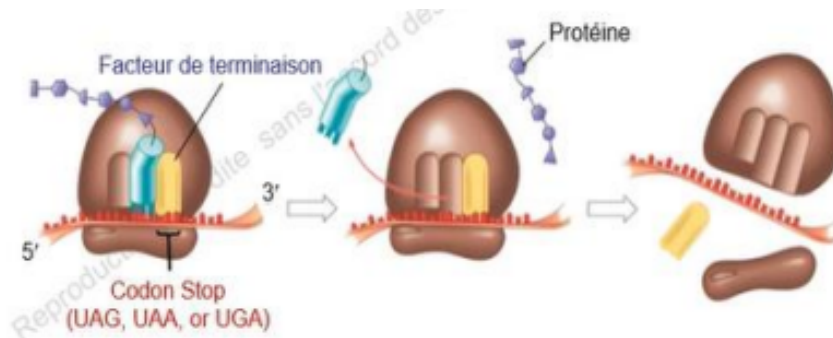


3^{ème} étape - La phase de terminaison

Elle correspond à l'arrêt de la traduction qui s'achève lorsque le ribosome rencontre un Codon Stop.

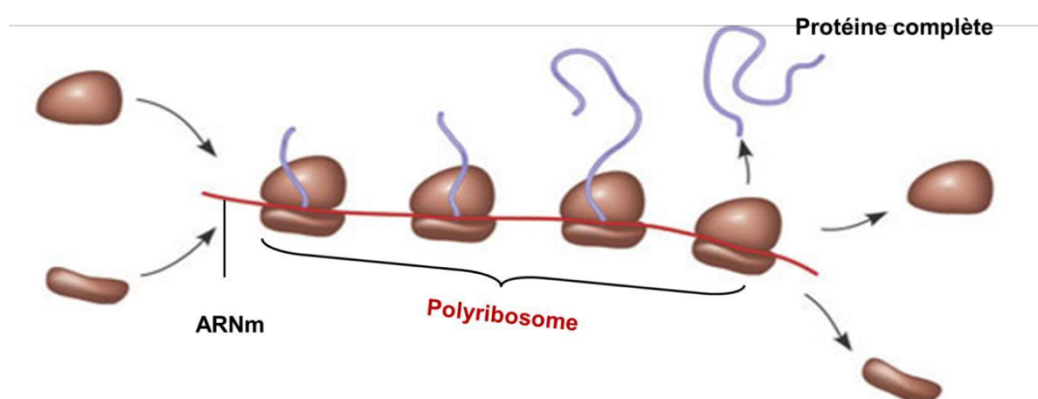
Il n'existe pas d'ARNt spécifique aux Codons Stop.

Une protéine appelée facteur de terminaison se positionne au niveau du site (A), la protéine est alors libérée et le ribosome se dissocie pour participer éventuellement à un autre cycle de traduction.



A noter : Un ARNm est traduit simultanément par plusieurs ribosomes. On a un ARNm sur lequel sont positionnés de multiples ribosomes et au niveau de chacun d'entre eux, la protéine qui est en cours de synthèse plus ou moins avancée.

L'ensemble de l'ARNm – ribosomes forme un polyribosome. Il contient en permanence de nombreuses molécules en cours de synthèse. Ainsi, l'efficacité et la rapidité de la traduction sont accrues (bravo les ribosomes).



Cette dédi va à la vie, au courage de sortir de sa confort zone, de prendre le risque, d'entamer de choses nouvelles et de pousser ses limites.