

CHAPITRE IV : LE CODE GENETIQUE

Ce sont les bases azotées qui constituent la seule partie variable d'un mRNA. Par conséquent, ce sont les bases qui sont impliquées dans le code génétique. Mais il y a seulement 4 bases différentes (A, U, C et G) et il existe 20 acides aminés différents ! Comment 4 bases peuvent-elles donc coder 20 acides aminés ?

La correspondance n'est pas un nucléotide pour un acide aminé puisqu'il n'existe que 4 nucléotides différents pour 20 acides aminés différents ($4^1 = 4 \ll 20$).

Le code génétique repose donc sur une combinaison de nucléotides.

Une combinaison de 2 parmi 4 nucléotides possibles ne peut suffire ($4^2 = 16 \text{ doublets} < 20$).

Une combinaison de 3 parmi 4 nucléotides est plausible puisqu'elle offre $4^3 = 64 \text{ triplets}$ possibles. Ce système peut coder 64 acides aminés, ce qui est cette fois largement suffisant. Cette très belle hypothèse de départ s'est trouvée effectivement confirmée. Un code à 3 lettres, cela veut donc dire que 3 nucléotides (triplet ou « **codon** ») portés sur le mRNA seront traduits pour positionner un acide aminé. On dispose ainsi de 64 codons.

3 codons (UAA, UAG et UGA) sont des « codons non sens » qui ne peuvent pas être traduits en acides aminés. Ces codons sont en fait des signaux de fin de lecture. On les appelle « codons stop ». En fin de compte, il reste 61 codons pour 20 acides aminés. Mis à part la méthionine et le tryptophane, codés par un seul codon, les 18 autres acides aminés sont codés par plusieurs codons : de 2 à 6 (6 codons pour la leucine par exemple).

DECHIFFRAGE DU CODE GENETIQUE

Pour déchiffrer le code génétique, deux techniques très astucieuses ont été utilisées :

La première stratégie a consisté à fabriquer au laboratoire un mRNA très particulier : le « poly U », ne comportant que des nucléotides à uracile. Mis en contact dans un système *in vitro* avec des ribosomes, des acides aminés et autres facteurs nécessaires à la traduction, le poly U a permis la synthèse d'une chaîne peptidique formée uniquement de phénylalanine. Cela signifiait donc que le codon UUU code la phénylalanine. On peut répéter cette expérience avec le poly A et trouver ainsi que le codon AAA code la lysine etc.

Des polynucléotides mixtes ayant 2 nucléotides différents répétés ont également été utilisés. Le poly AC par exemple (ACACACACACAC&) code ainsi (Thr-His-Thr-His)_n. Ce résultat ne permet cependant pas de savoir si Thr est codée ACA et His par CAC et vice versa.

L'expérience a donc été répétée avec un poly AAC. Celui-ci donne 3 sortes de chaînes polypeptidiques selon que la lecture commence en AAC (poly Asn), en ACA (poly Thr) ou CAA (poly Gln). Le seul codon en commun avec l'expérience précédente est ACA et le seul acide aminé en commun dans les polypeptides obtenus est Thr. On en déduit donc que ACA code Thr (et CAC code His).

La deuxième stratégie utilise différents « aminoacyl tRNA » (aa~tRNA), des ribosomes et des trinucéotides synthétiques différents dans chaque expérience. Les trinucéotides peuvent

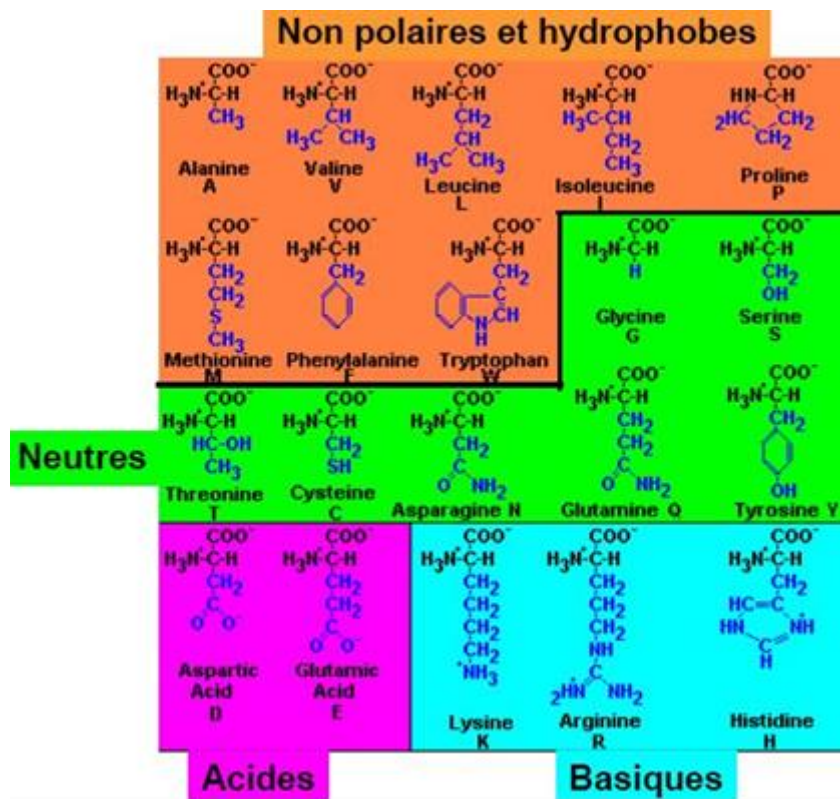
Cette dégénérescence du code génétique n'est pas l'effet du hasard. Mis à part le cas de

L'arginine, la leucine et la sérine représentés par (4 + 2 = 6) codons et les non sens_ représentés par (2 + 1 = 3) codons, tous les codons « **synonymes** » pour un acide aminé donné ne diffèrent que par **la troisième base du triplet**.

De plus, cette troisième base est soit une pyrimidine (U ou C), soit une purine (A ou G) soit l'une des quatre (U, C, A ou G). Cette règle s'applique également pour les triplets codant l'arginine, la leucine ou la sérine pris par groupes indépendants de 4 et de 2 codons.

Très remarquable aussi le fait que lorsque la seconde position du codon est occupée par une pyrimidine, l'acide aminé déterminé est généralement hydrophile (polaire ou chargé) : c'est le cas de Tyr, His, Gln, Asn, Lys, Asp, Glu, Cys, Arg, Ser et Gly.

Si la seconde position est une purine, l'acide aminé est alors généralement polaire hydrophobe : c'est le cas de Phe, Leu, Ile, Met, Val, Pro et Ala.

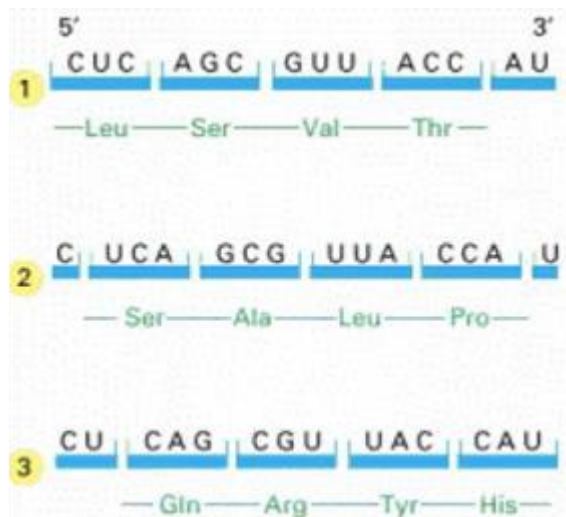


3-LE CODE GENETIQUE EST NON CHEVAUCHANT

Les codons sont toujours lus sans chevauchement. La partie exprimée d'un gène est transcrite puis traduite régulièrement triplet par triplet. Les codons sont lus **en série** depuis un point de départ jusqu'à un point de terminaison **sans recouvrement** mutuel entre les codons.

Le message génétique du mRNA doit toujours être lu dans le bon ordre et selon le bon groupement. On appelle « **cadre de lecture** » ou « **ORF** » (de l'anglais : _Open Reading Frame_) cette façon d'ordonner les nucléotides d'un mRNA en triplets successifs dans le bon sens et selon les bons groupements.

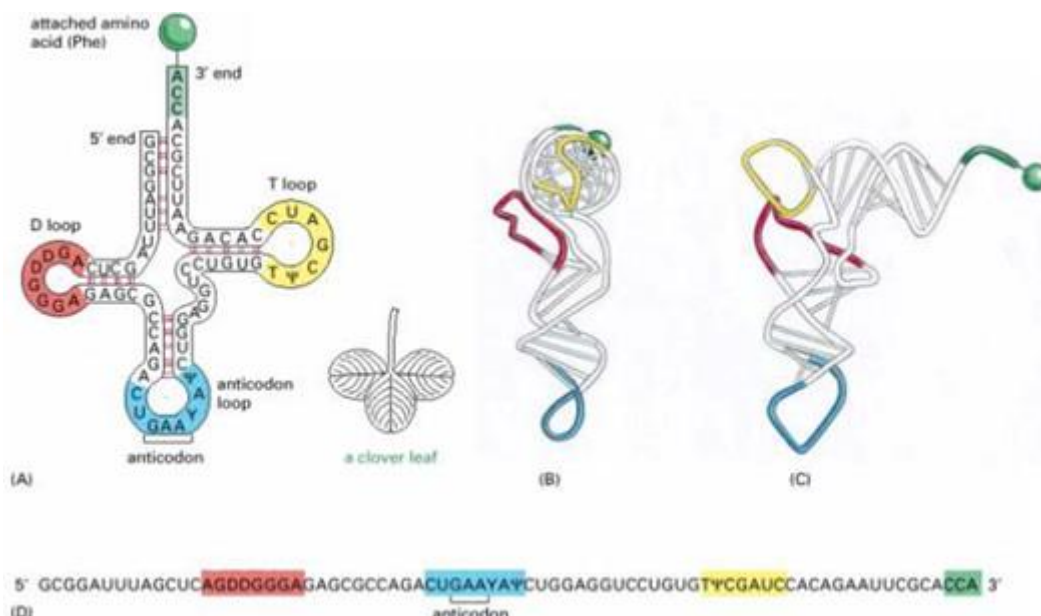
Un mRNA : 3 cadres de lecture possibles



tRNA ET « AMINOACYL-tRNA SYNTHETASES » : DEUX ADAPTATEURS DU CODE GENETIQUE

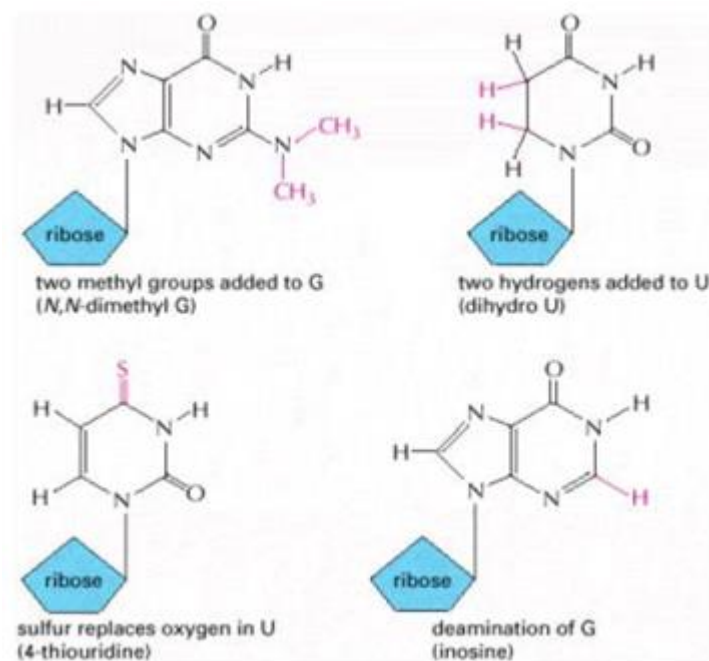
PROPRIETES DES tRNA ET NOTION DU WOBBLE

Les molécules de « **RNA de transfert** » (tRNA), comme toutes les autres sortes de RNA, résultent d'une transcription enzymatique à partir de matrice de DNA situées sur le génome procaryote ou eucaryote (nucléaire, mitochondrial ou chloroplastique) : « **gènes tRNA** ». Chaque tRNA peut servir un grand nombre de fois. Il prélève l'acide aminé correspondant dans le cytoplasme, dépose sa cargaison au niveau du ribosome, puis quitte celui-ci afin d'aller chercher une autre cargaison. La structure du tRNA supporte bien la fonction de **navette** qu'il doit assurer pour un acide aminé spécifique.



RAPPELS : Une molécule de tRNA se compose d'un seul brin de RNA d'une longueur d'environ 80 nucléotides. Ce brin se replie sur lui-même pour former une molécule pourvue d'une structure secondaire en feuille de trèfle. La molécule de tRNA possède en effet une structure trilobée.

Cette structure est elle-même tordue et pliée et adopte une conformation tridimensionnelle très compacte vaguement semblable à un « L » inversé. La boucle qui dépasse à une extrémité du L comprend « l'anticodon ». L'extrémité 3' dépasse à l'autre extrémité du L et représente le « site de liaison de l'acide aminé ». Cette extrémité se termine toujours par les trois nucléotides CCA : on l'appelle également « extrémité CCA ».



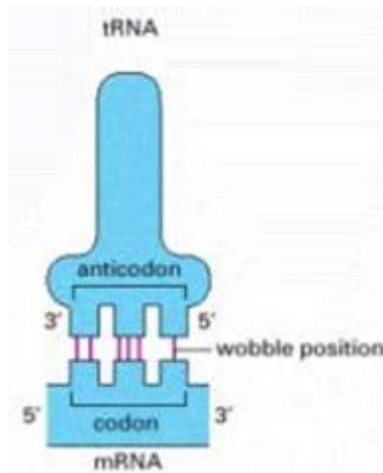
Les tRNA contiennent des **nucléotides atypiques**, inhabituels par la nature de leur base azotée. On trouve par exemple de la thymine, ainsi que d'autres bases azotées méthylées. En particulier, on trouve de « **l-hypoxanthine** » comme base azotée dont le nucléotide est « **IMP** ».

L'ensemble de ces bases atypiques n'est d'ailleurs pas incorporé tel quel au moment de la synthèse des tRNA. Ces bases sont formées par modification « **post transcriptionnelle** » d'une des 4 bases régulières (A, U, C ou G). Ainsi, IMP et TMP sont obtenus respectivement par désamination d'un AMP et méthylation d'un UMP.

Si chaque codon du code génétique correspond à un tRNA différent, on aurait trouvé 61 tRNA distincts pour assurer le mécanisme de la traduction dans les cellules. En fait, selon les cas, il n'en existe que 22 à 46 environ. Ce nombre suffit parce que les anticodons de certains tRNA peuvent reconnaître deux anticodons différents ou plus.

Cette souplesse est possible parce que les règles de l'appariement entre la troisième base d'un codon et la base correspondante de l'anticodon (**position d'oscillation de l'anticodon** : 1ère base de l'anticodon) ne sont pas très strictes (contrairement à ce que ce passe entre les bases du DNA et du mRNA).

Ce relâchement des règles d'appariement des bases est appelé « **wobble** » (mot en anglais qui signifie _avoir du jeu_ **oscillation** ou base **fluctuante** en français)



DANS LES MITOCHONDRIES DES MAMMIFERES: WOBBLE DE TYPE U/N

Ce type de wobble, décrit par Sanger et *al.* [1980], ne se rencontre que dans les mitochondries. Le nucléotide U en position wobble sur l'anticodon peut reconnaître laquelle des quatre bases (A, U, C, G).

Dans ce cas, lors de la synthèse des protéines effectuées dans les mitochondries, il suffit :

- D'un seul tRNA pour reconnaître les codons d'une **boîte à 4** , soit **8 tRNA** (Leu, Val, Ser, Pro, Thr, Ala, Arg et Gly).
- D'un seul tRNA pour reconnaître les codons d'une boîte à 2 supérieure soit tRNA (Phe, Ile, Tyr, His, Asn, Cys et Ser).
- D'un seul tRNA pour reconnaître les codons d'une boîte à 2 inférieure, soit 6 tRNA (Leu, Met, Gln, Lys, Glu et Trp).

Remarque : un seul tRNA est requis pour Met qu'il s'agisse de la **Met initiale** ou de la **Met interne**). Ceci n'est pas le cas pour la synthèse des protéines au niveau du cytoplasme.

		Second letter					
		U	C	A	G		
First letter	U	UUU } Phe UUC UUA } Leu UUG	UCU } UCC } Ser UCA UCG	UAU } Tyr UAC UAA Stop UAG Stop	UGU } Cys UGC UGA Trp UGG Trp	U C A G	Third letter
	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 } Ile AUC AUA } Met AUG	ACU } ACC } Thr ACA ACG	AAU } Asn AAC AAA } Lys AAG	AGU } Ser AGC AGA Stop AGG Stop	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	

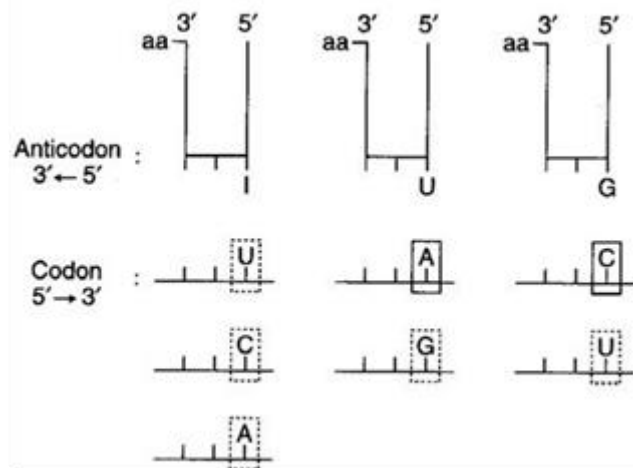
On dénombre donc **au total** seulement **22 tRNA** pour la synthèse des protéines dans les **mitochondries des mammifères**.

DANS LE CYTOPLASME : WOBBLE DE TYPE I/UCA, U/G et G/U

Ce type de wobble, suggéré par Crick [1966], réduit à 32 le nombre de tRNA nécessaires pour la synthèse des protéines au niveau du compartiment cytoplasmique.

La première base d'un anticodon détermine si un même tRNA est susceptible de reconnaître 1, 2 ou 3 codons différents :

- Lorsque (I) est situé en position wobble sur l'anticodon, il peut s'apparier aussi bien à (U), à (C) qu'à (A) [appariements wobbles]
- Lorsque (U) occupe la position wobble sur l'anticodon, (U) peut s'apparier aussi bien à (A) [appariement classique] qu'à (G) [appariement wobble].
- Lorsque (G) occupe la position wobble sur l'anticodon, (G) peut s'apparier aussi bien à (C) [appariement classique] qu'à (U) [appariement wobble].



Vérifions selon ces règles le nombre de tRNA minimum pour reconnaître les 61 codons figurant sur le code génétique universel :

1) 1 tRNA pour chaque boîte à 2 supérieure : cas de Phe, Tyr, His, Asn, Asp, Cys, Ser, soit 7 tRNA ;

2) 1 tRNA pour chaque boîte à 2 inférieure : cas de Leu, Gln, Lys, Glu et Arg, soit 5 tRNA ;

3) 2 tRNA pour chaque boîte à 4.

Un premier tRNA aura I en position wobble et reconnaîtra les 3 codons supérieurs

Un deuxième tRNA avec C en position wobble sera nécessaire pour reconnaître le codon inférieur (liaison classique C-G) : cas de Leu, Val, Ser, Pro, Thr, Ala, Arg et Gly. Soit $2 \times 8 = 16$ tRNA

4) 1 tRNA pour chaque boîte à 3 supérieure : cas de Ile, soit 1 tRNA

5) 1 tRNA pour chaque boîte à 1 inférieure (liaison classique C-G) : cas de Met et Trp. soit 2 tRNA. En fait, il faut 2 (et non 1) tRNA pour reconnaître Met (l'un pour les Met à incorporer en cours de chaîne, l'autre pour la Met initiale qui démarre toute nouvelle chaîne peptidique). Ces 2 tRNA auront le même anticodon AUG mais différeront par ailleurs.

Soit donc un total de :

32 tRNA différents.

Remarque : D'autres types de wobble existent (chez la levure, les bactéries). En fonction du type de wobble qui régit les appariements entre anticodons et codons respectifs, nous retiendrons qu'il faut un minimum de 22 tRNA (mitochondrie de mammifères), de 32tRNA (cytoplasme de mammifères), 46 tRNA chez certaines levures etc. Le nombre de tRNA nécessaire pour véhiculer les 20 acides aminés sera compris entre 22 et 61.