

Le Code Génétique & La Traduction

Hubert Becker

IPCB, UMR7156 4^{ème} étage

<http://gmgm.unistra.fr/index.php?id=3634>

Le Code Génétique

I. Le code génétique

II. L'aminocyl-ARNt

III. La synthèse de l'aminocyl-ARNt

IV. L'identité d'un ARNt

V. La reprogrammation

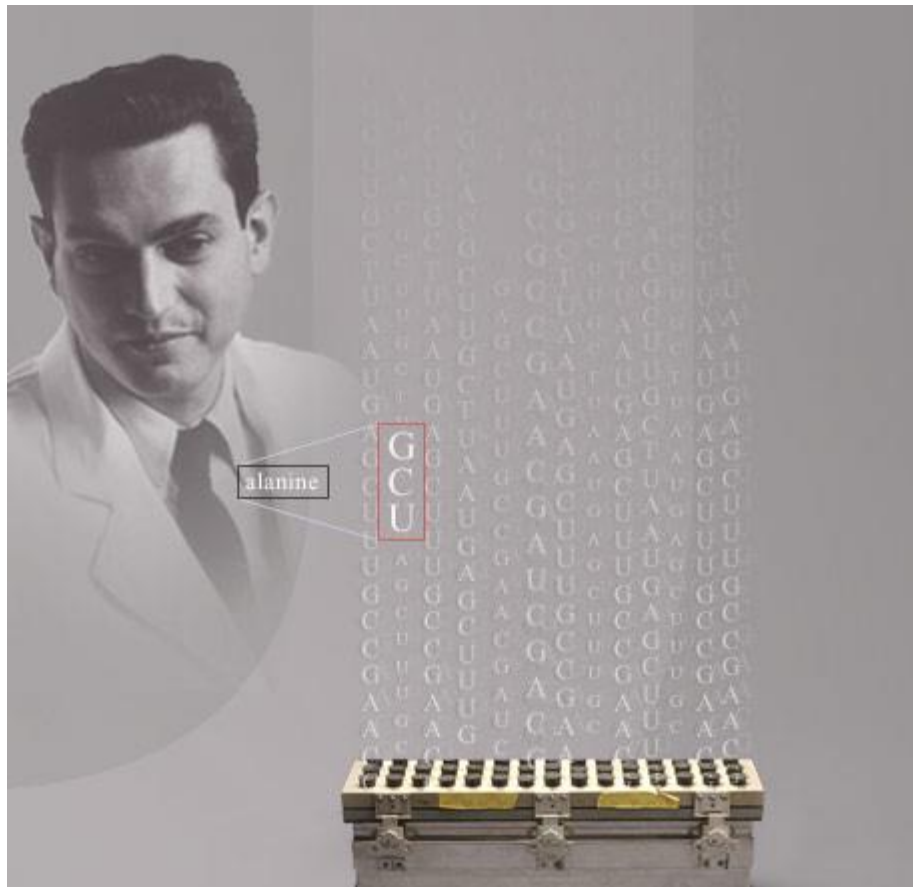
VI. Universalité

VII. Usage des codons

I. Le Code Génétique

Déchiffrage

http://www.history.nih.gov/exhibits/nirenberg/HS5_cracked.htm



I. Le Code Génétique

Déchiffrage

Taille unité codante c'est savoir
Combien de combinaisons

4 Nucléotides \longrightarrow 20 acides aminés

I. Le Code Génétique

Déchiffrage

CHARACTERISTICS AND COMPOSITION OF RNA CODING UNITS*
BY J. HEINRICH MATTHAEI, OLIVER W. JONES, ROBERT G. MARTIN, AND)
MARSHALL W. NIRENBERG
NATIONAL INSTITUTE OF ARTHRITIS AND METABOLIC DISEASES, BETHESDA
Communicated by Richard Roberts, February 27, 1962
Proc. Natl. Acad. Sci. U.S.A. **48**, 666-677

Cette expérience nécessitait du Mg²⁺, de l'ATP, des ribosomes, un RNA synthétique, des aa dont un marqué radioactivement [¹⁴C]aa.



Le 27 Mai 1961 3:00 AM

I. Le Code Génétique

Déchiffrage

Extrait brut d'*E.coli* (ribosomes actifs + enzymes chargeant l'ARN 4S)

+

5'UUU...UUU3'

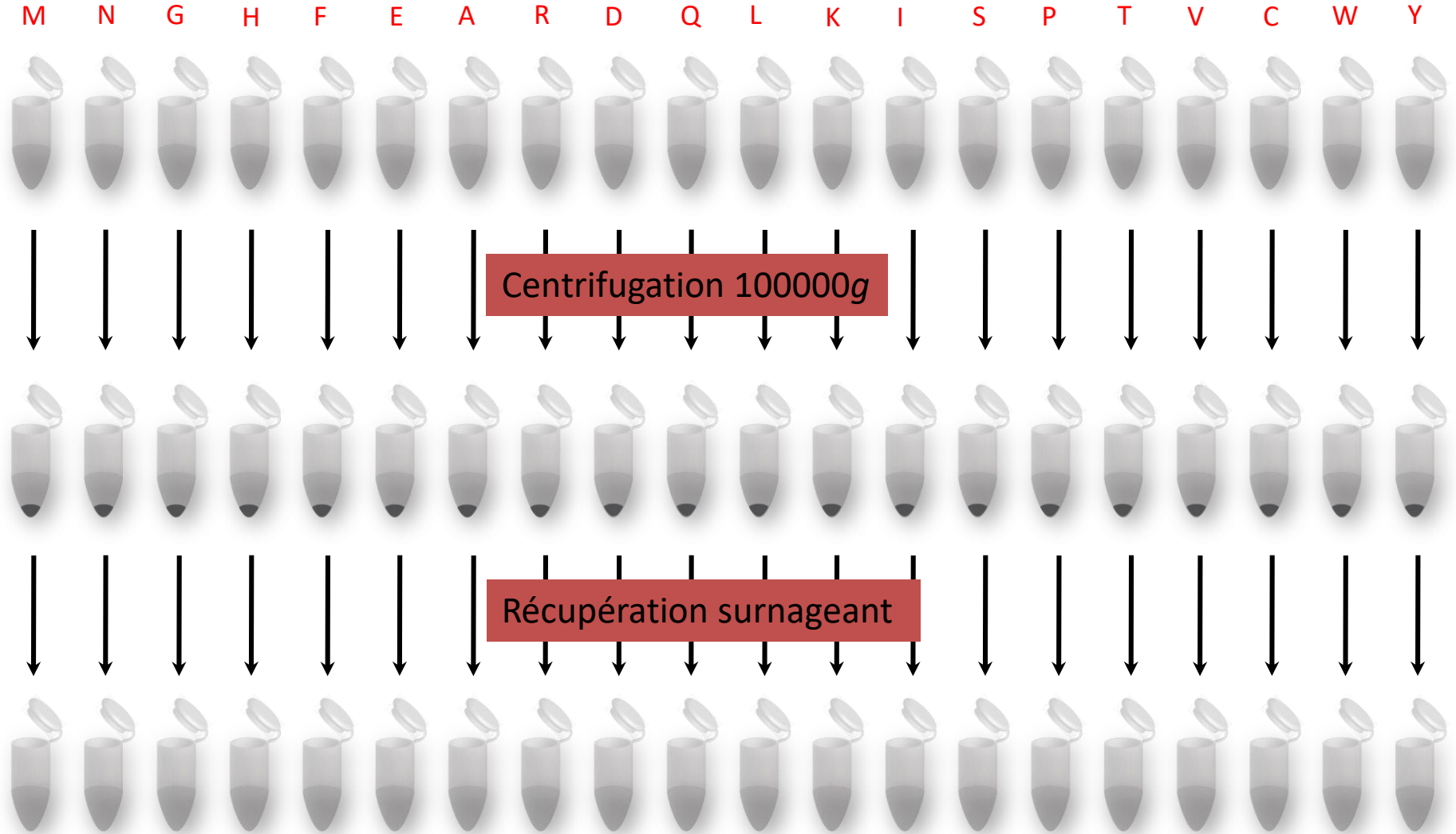
+

M	N	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
N	G	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
G	H	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
H	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
A	R	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
R	D	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
D	Q	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
Q	L	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
L	K	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
K	I	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K
I	S	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
S	P	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
P	T	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
T	V	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
V	C	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
C	W	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
W	Y	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
Y		Y																	



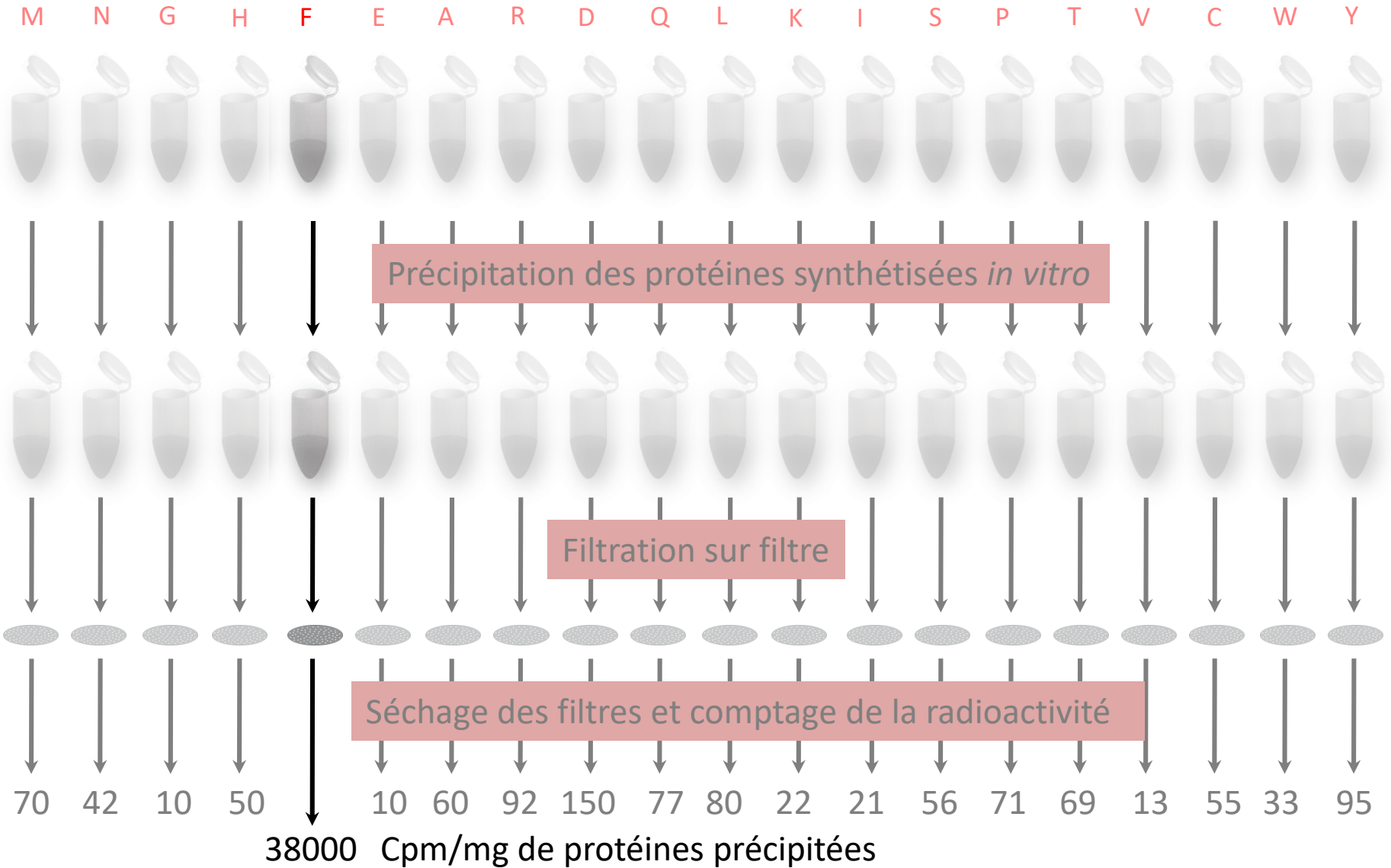
I. Le Code Génétique

Déchiffrage



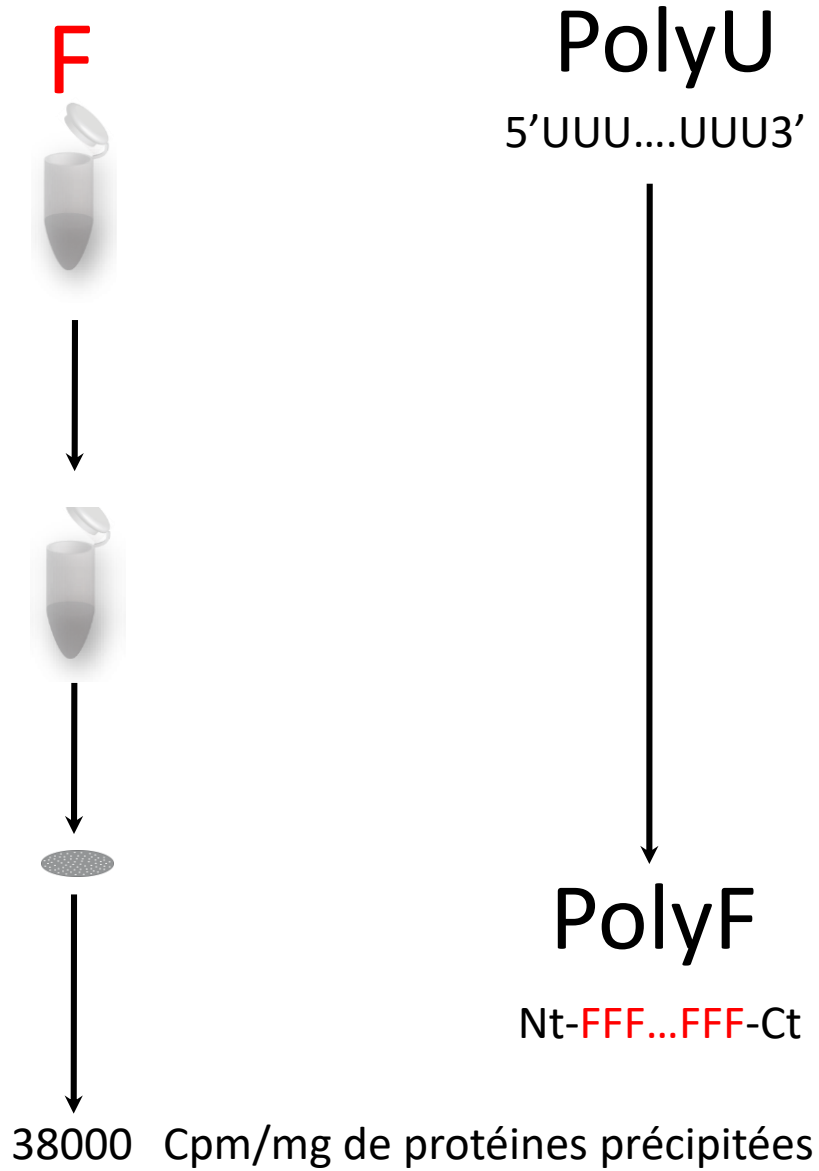
I. Le Code Génétique

Déchiffrage



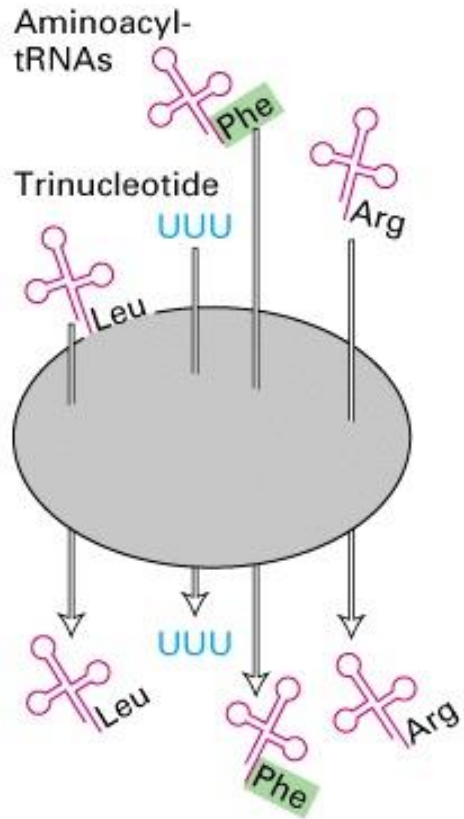
I. Le Code Génétique

Déchiffrage

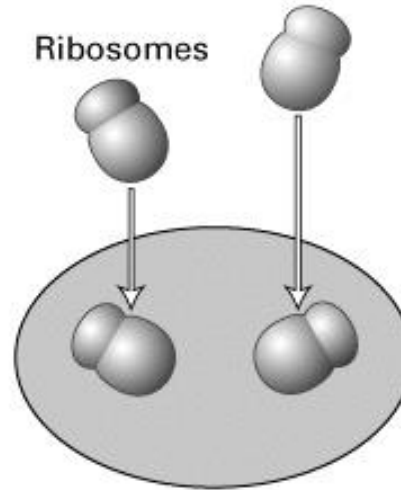


I. Le Code Génétique

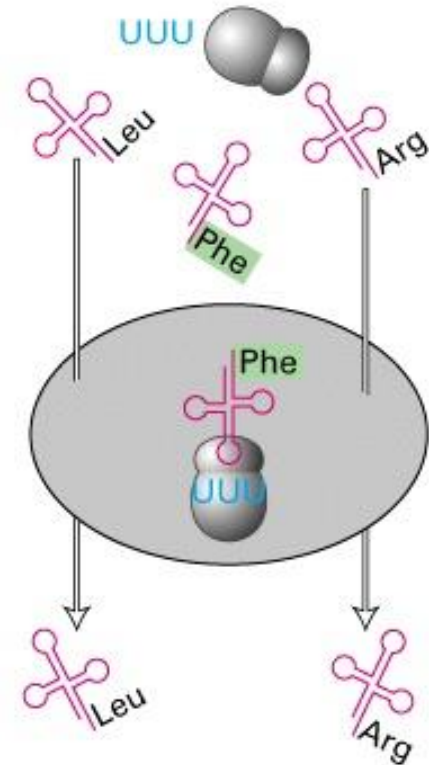
Déchiffrage



Trinucleotide and all tRNAs pass through filter



Ribosomes stick to filter



Complex of ribosome, UUU, and Phe-tRNA sticks to filter

U
UU
UUU
UUUU
UUUUU
UUUUUU

I. Le Code Génétique

Déchiffrage

Trinucléotide	[¹⁴ C]Aminoacyl-ARNt lié au ribosome		
	Phe-ARNt ^{Phe}	Lys-ARNt ^{Lys}	Pro-ARNt ^{Pro}
UUU	4,6	0	0
AAA	0	7,7	0
CCC	0	0	3,1

Source: Modified from Nirenberg, M. & Leder, P. (1964) RNA code words and protein synthesis. *Science* **145**, 1399.

UUU code F

I. Le Code Génétique

Déchiffrage

The genetic code has seven main characteristics:

1. It is made up of codons, which are triplets of bases. Each codon specifies a specific amino acid.
2. The codons do not overlap; that is, the sequence GCCAC contains two triplets, "GCC" and "CAC" not counting the "CCC" and other subsequent three-letter sequences.
3. The code includes punctuation in the form of three "stop" codons that do not code for an amino acid: UAA, UAG, and UGA.
4. The genetic code is known as a "degenerate" code. This means that each amino acid is triggered by between one and six codons. (There are only 20 amino acids and 64 possible codon triplets).
5. To read each gene and glean the necessary information to form proteins, cells begin at a fixed and particular starting point on the mRNA strand. The initiation codon is AUG (methionine).
6. The mRNA strand is read from the 5' to the 3' end.
7. If there are mutations or errors in the DNA, the message may be changed and incorrect protein formation results.

I. Le Code Génétique

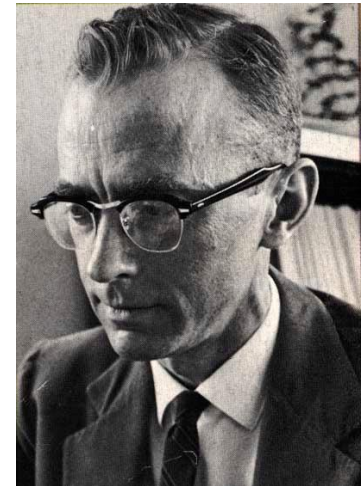
Déchiffrage



Prix Nobel de Physiologie et de Médecine
en 1968



[Har Gobind Khorana](#)
Synthèse RNA

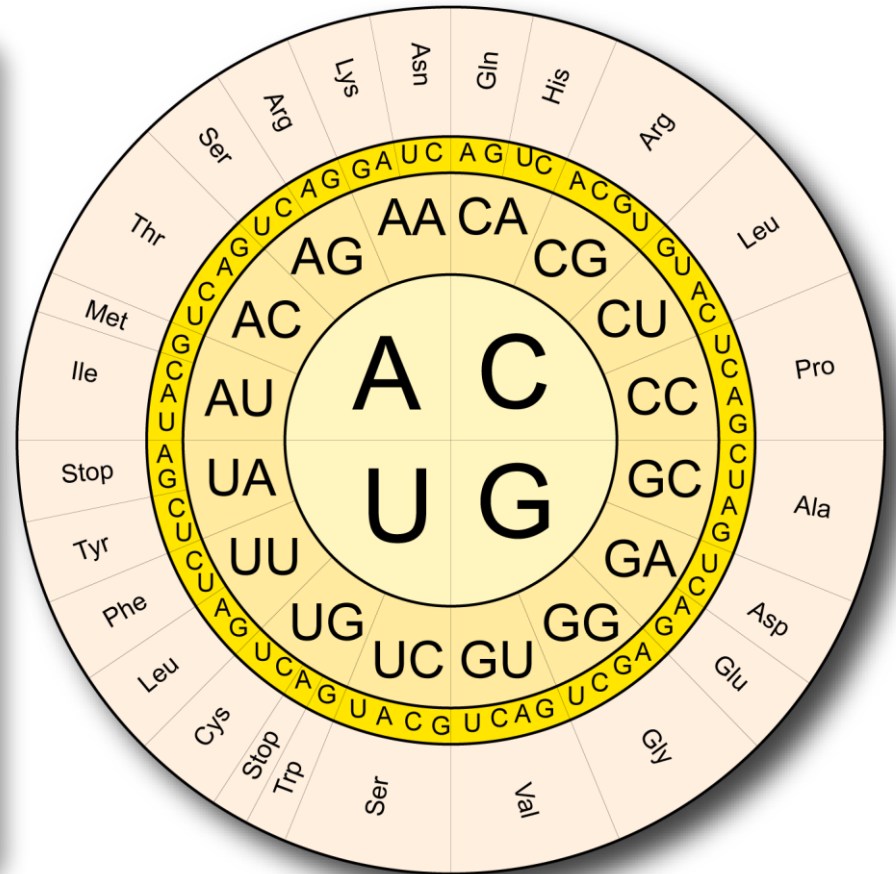


[Robert W. Holley](#)
Structure ARNt

I. Le Code Génétique

Déchiffrage

		Deuxième lettre				
		U	C	A	G	
U	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
	C	Leu	Ser	Stop	Stop	A
		Leu	Ser	Stop	Trp	G
A	U	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
	C	Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
G	Met	Thr	Lys	Arg	A	
	Val	Ala	Lys	Arg	G	
	Val	Ala	Asp	Gly	U	
	Val	Ala	Asp	Gly	C	
Val	Ala	Glu	Gly	A		
Val	Ala	Glu	Gly	G		



(Lettres pas utilisées pour le code à 1 lettre des aa: BJOUXZ)

II. Le Code Génétique

Dégénérescence

		Deuxième lettre				
		U	C	A	G	
Première lettre	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	Stop	Stop	A
		Leu	Ser	Stop	Trp	G
	C	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
	G	Met	Thr	Lys	Arg	G
Val		Ala	Asp	Gly	U	
Val		Ala	Asp	Gly	C	
Val		Ala	Glu	Gly	A	
Val	Ala	Glu	Gly	G		

 Un aa peut être codé par plusieurs codons

Il existe 61 codons pour coder 20 aa naturels

La plupart des aa sont codés par plus d'1 codon

Seuls M et W sont codés par 1 seul codon

Qd plusieurs codons codent pour un même aa les 2 premiers Nt sont invariants et c'est le 3^{ème} Nt qui varie

 Sauf qd aa codé par 6 codons: L, R et S

I est le seul aa codé par 3 codons

Il existe 3 codons dits non codants ou STOP

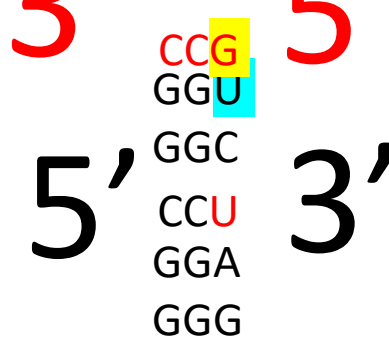
UAA (ochre)

UAG (ambre)

UGA (Opale)

II. Le Code Génétique

Dégénérescence



		Deuxième lettre				
		U	C	A	G	
Première lettre	U	Phe	Ser	Tyr	Cys	Troisième lettre
		Phe	Ser	Tyr	Cys	
	Leu	Ser	Stop	Stop		
	Leu	Ser	Stop	Trp		
	C	Leu	Pro	His	Arg	
		Leu	Pro	His	Arg	
		Leu	Pro	Gln	Arg	
		Leu	Pro	Gln	Arg	
	A	Ile	Thr	Asn	Ser	
		Ile	Thr	Asn	Ser	
		Ile	Thr	Lys	Arg	
	Met	Thr	Lys	Arg		
G	Val	Ala	Asp	Gly		
	Val	Ala	Asp	Gly		
	Val	Ala	Glu	Gly		
	Val	Ala	Glu	Gly		
	Val	Ala	Glu	Gly		

Si vous regardez ce code l'ordre des bases est toujours le même pour que les codons soient bien ordonnés:

- 1^{ère} lettre du codon **UCAG** du haut vers le bas
- 2^{ème} lettre du codon **UCAG** de la gauche vers la droite
- 3^{ème} lettre du codon **UCAG, UCAG UCAG et UCAG** du haut vers le bas.

Pour tout aa* codé par plus d'un codon, c'est toujours la 3^{ème} lettre du codon qui est dégénérée

Dans le cas de la **Leucine** on a aussi la 1^{ère} lettre qui est dégénérée: codons 5'**UUA**3' et 5'**UUG**3' mais aussi 5'**CUN**3' (N=UCAG)

Les codons **Arginine** ont également une 1^{ère} lettre dégénérée:

5'**CGN**3' ou 5'**AGA**3' et 5'**AGG**3'

Les codons les plus dégénérés sont les codons **Sérine** dont la 1^{ère}, 2^{ème} et 3^{ème} lettre sont dégénérée:

5' **UCN**' ou 5'**AGU**3' et 5'**AGC**3'

aa* : acide aminé

II. Le Code Génétique

Dégénérescence

		Deuxième lettre				
		U	C	A	G	
Première lettre	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	Stop	Stop	A
		Leu	Ser	Stop	Trp	G
	C	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
	G	Met	Thr	Lys	Arg	G
Val		Ala	Asp	Gly	U	
Val		Ala	Asp	Gly	C	
Val		Ala	Glu	Gly	A	
Val	Ala	Glu	Gly	G		

aa* : acide aminé

Si je prend la 3^{ème} colonne du code elle se compose d'aa codés par 2 codons et si vous regardez attentivement la dégénérescence de la 3^{ème} lettre est pour chaque aa toujours soit UC soit AG.

La raison est toute simple: Grace au fait que dans l'ARN on peut faire une pdb G-U, vous allez pouvoir décoder les codons se terminant (en 3') par U ou C avec un seul ARNt qui commence avec un G (en 5')

De la même manière, vous allez pouvoir décoder les codons se terminant (en 3') par A ou G avec un seul ARNt qui commence avec un U (en 5')

Le décodage de la dégénérescence de 3^{ème} lettre du codon par la 1^{ère} lettre de l'anticodon de l'ARNt s'appelle:

WOBBLE PAIRING

I. Le Code Génétique

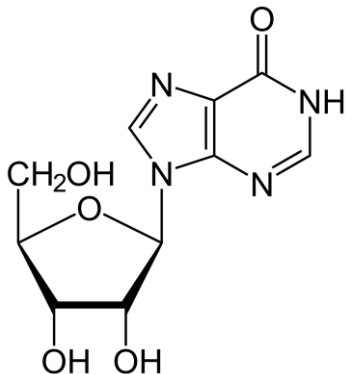
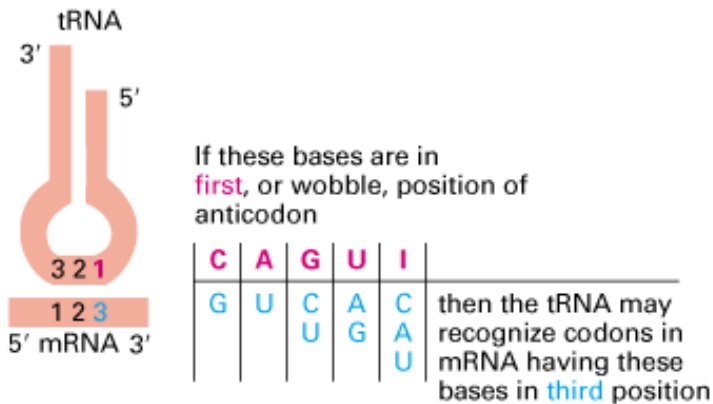
2. Le Wobble pairing



Il y a l'explication dans le volet commentaire du PPT



Puisqu'un aa peut être codé par plusieurs codons, y a-t-il autant de tRNA que de codons ? **NON.** Chez les procaryotes il existe 30 espèces de tRNA pour décoder 61 codons et chez les eukaryotes 50 (48 chez l'homme). Il n'y a donc pas assez de tRNA. Donc pour certains tRNA un même tRNA doit pouvoir lire plusieurs codons. Ceci est rendu possible grâce au **Wobble pairing** (« appariement oscillant »)



L'inosine est le nucléotide dont la base est l'hypoxanthine qui résulte de la dédarboxylation oxydative de l'adénine: du coup au lieu d'un NH2 en position 6 on a un C=O.

L'association **codon/@codon** se fait de manière anti-parallèle

Un même tRNA peut lire plusieurs codons et peuvent décrypter la dégénérescence du codon (au max 1 tRNA peut lire 3 codons ≠)

Certains aa sont codés par plusieurs codons, il existe donc jusqu'à 4 tRNA ≠ qui décodent le même aa on parle d'**ARNt isoaccepteurs**

I. Le Code Génétique

2. Le Wobble pairing



Il y a l'explication dans le volet commentaire du PPT



Puisqu'un aa peut être codé par plusieurs codons, y a-t-il autant de tRNA que de codons ? **NON.** Chez les procaryotes il existe 30 espèces de tRNA pour décoder 61 codons et chez les eukaryotes 50 (48 chez l'homme). Il n'y a donc pas assez de tRNA. Donc pour certains tRNA un même tRNA doit pouvoir lire plusieurs codons. Ceci est rendu possible grâce au **Wobble pairing** (« appariement oscillant »)



If these bases are in **first**, or wobble, position of anticodon

C	A	G	U	I
G	U	C	A	C
		U	G	A
				U

then the tRNA may recognize codons in mRNA having these bases in **third** position

L'association **codon/@codon** se fait de manière anti-parallèle

Un même tRNA peut lire plusieurs codons et peuvent décrypter la dégénérescence du codon (au max 1 tRNA peut lire 3 codons ≠)

Certains aa sont codés par plusieurs codons, il existe donc jusqu'à 4 tRNA ≠ qui décodent le même aa on parle d'**ARNt isoaccepteurs**



If these bases are in **third**, or wobble, position of codon of an mRNA

C	A	G	U
G	U	C	A
I	I	U	G
			I

then the codon may be recognized by a tRNA having these bases in **first** position of anticodon

Chez l'homme 16 tRNA utilisent le wobble alors que 32 restants sont spécifiques d'un seul triplet

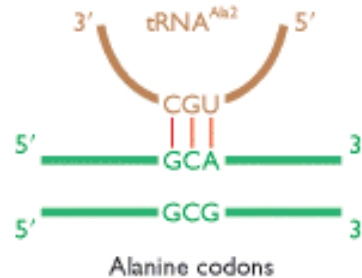
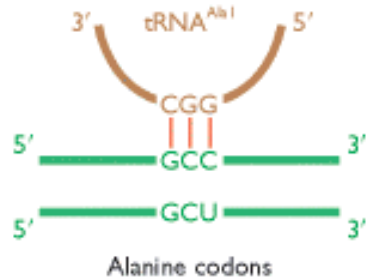
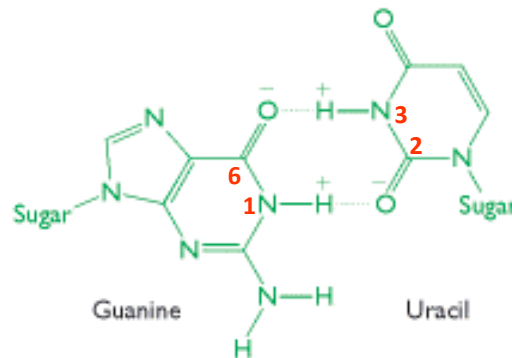
I. Le Code Génétique

2. Le Wobble pairing

Un exemple de Wobble pairing: L'Ala est codée par 4 codons ≠: GC(U,C,A,G) et décodé par 2 ARNt^{Ala}:

- 1 ARNt^{Ala} avec un @codon 5' **G**GC3' capable de reconnaître 5'GG**C**3' et l'autre 5'GC**U**3'
- 1 ARNt^{Ala} avec un @codon 5' **U**GC3' capable de reconnaître 5'GC**A**3' et l'autre 5'GC**G**3'

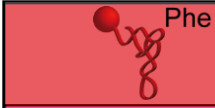

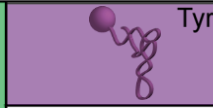
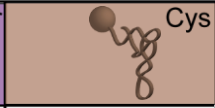









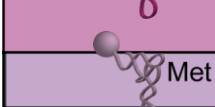









Les Nucléotides du Wobble pairing sont en rouge sur fond jaune



II. L'aminocyl-ARNt (aa-ARNt)

C'est un ARNt sur lequel est attaché l'aa correspondant. Comme il y a 20 aa à décoder il y a 20 familles d'aminocyl-ARNt isoaccepteurs.

Second Codon Letter

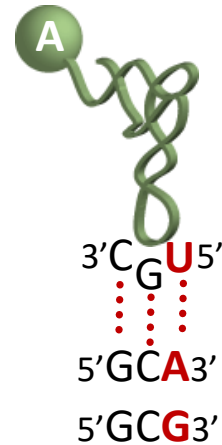
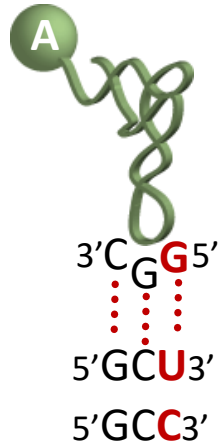
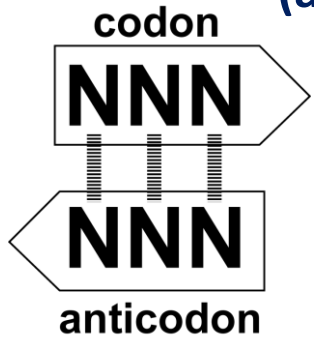
		U	C	A	G	
First Codon Letter	U	 Phe	 Ser	 Tyr	 Cys	U
	C	 Leu	 Pro	Stop	Stop	A
		 Leu		Stop	Trp	G
	A	 Ile	 Thr	 His	 Arg	U
		 Met		 Gln		C
	G	 Val	 Ala	 Asn	 Ser	U
 Asp				 Arg	C	
 Val	 Ala	 Glu	 Gly	A		
				G		

Third Codon Letter

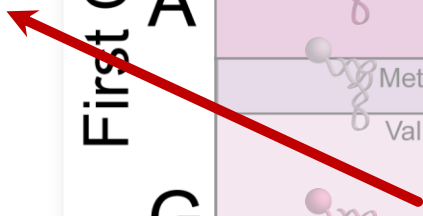
N'oubliez pas, 4 codons nécessitent 2 aa-ARNt pour être décodés. Ces 2 ARNt^{Ala} ont une séquence différentes mais un point commun ils portent le même aa: Ala du coup ont a 2 Ala-ARNt^{Ala} isoaccepteurs de la racine grecque iso – égal (identique)

II. L'aminoacyl-ARNt (aa-ARNt)

C'est un ARNt sur lequel est attaché l'aa correspondant. Comme il y a 20 aa à décoder il y a 20 familles d'aa-ARNt isoaccepteurs.



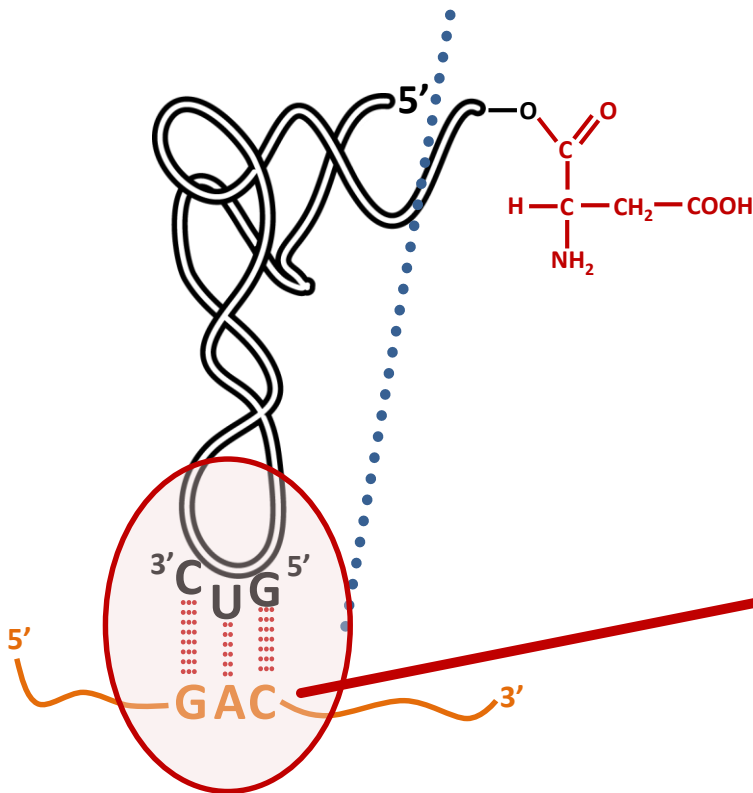
		Second Codon Letter				
		U	C	A	G	
First Codon Letter	U	Phe	Ser	Tyr	Cys	U
	C	Leu	Pro	His	Arg	C
	A	Ile	Thr	Asn	Ser	A
	G	Met	Val	Lys	Arg	G
						U
						C
						A
						G
						U
						C
						A
						G
						U
						C
						A
						G
						U
						C
						A
						G



5'GCU3'
5'GCC3'
5'GCA3'
5'GCG3'

II. L'aminocyl-ARNt (aa-ARNt)

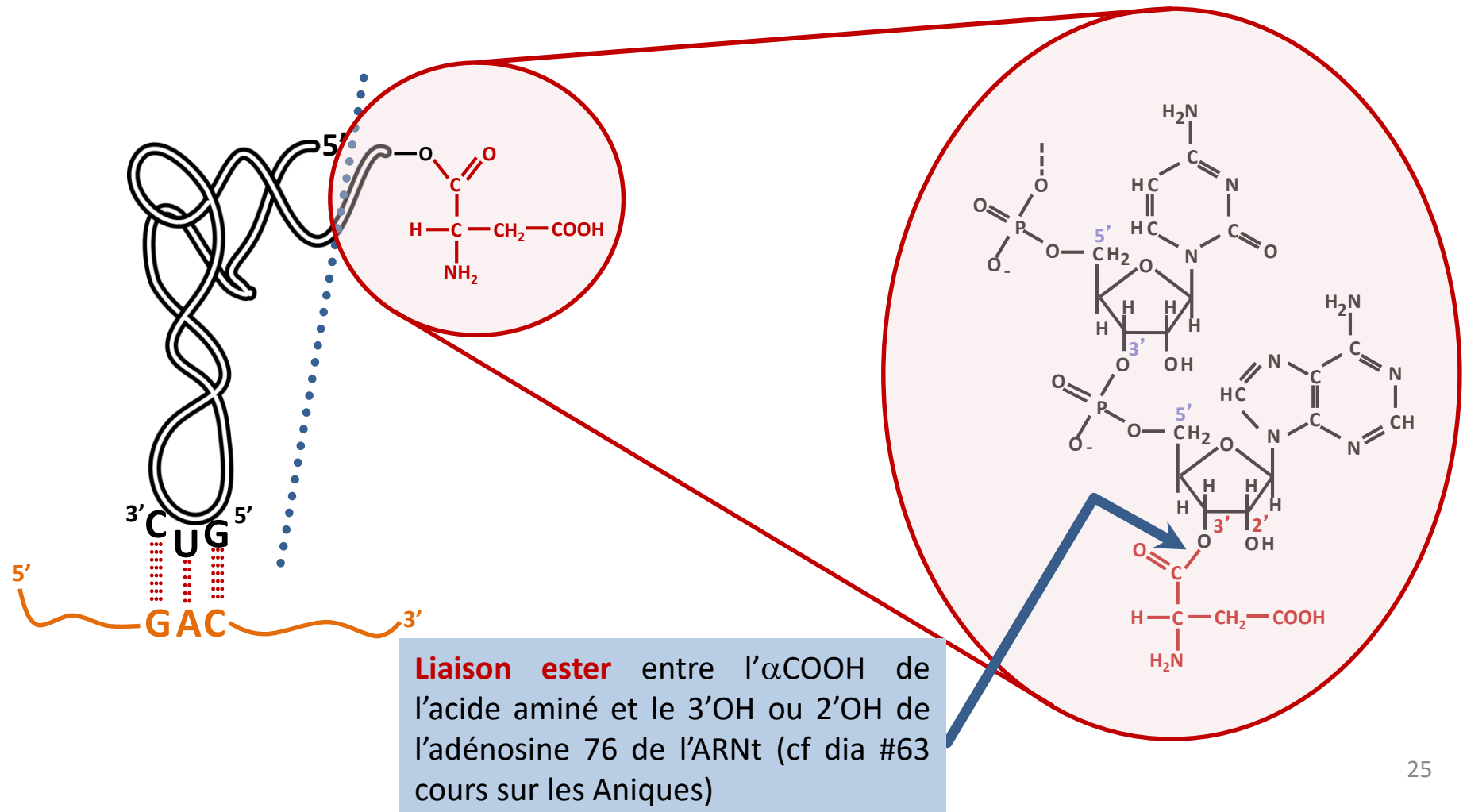
L'aa-ARNt a 2 identités, **1 identité génétique** (sa séquence nucléotidique): si je regarde son anticodon = 5'GUC3' donc décode 5'GAC3' et si je regarde dans le code.... Il s'agit donc de l'aspartyl-ARNt^{Asp}



		Deuxième lettre				
		U	C	A	G	
U	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
	C	Leu	Ser	Stop	Stop	A
		Leu	Ser	Stop	Trp	G
C	U	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
	C	Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
A	U	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
	C	Ile	Thr	Lys	Arg	A
		Met	Thr	Lys	Arg	G
G	U	Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
	C	Val	Ala	Glu	Gly	A
		Val	Ala	Glu	Gly	G

II. L'aminocyl-ARNt (aa-ARNt)

Et 1 identité "peptidique" l'acide aminé auquel il est lié et cette identité est le résultat d'une réaction enzymatique catalysée par une **aminocyl-ARNt synthétase** qui reconnaît spécifiquement son ARNt ou ses ARNt isoaccepteur et l'acide aminé correspondant ou homologue et aminoacyle l'ARNt avec l'aa.... Il s'agit de **l'acide aspartique** donc il s'agit de l'Asp-ARNt^{Asp}



II. L' aminoacyl-ARNt

Pour synthétiser ces 20 familles aminoacyl-ARNt il faut 20 aminoacyl-ARNt synthétases – une pour chaque couple aa/ARNt isoaccepteur

		Second Codon Letter				
		U	C	A	G	
First Codon Letter	U	Phe	Ser	Tyr	Cys	U
		Leu		Stop	Stop	A
	C		Pro	His	Arg	G
				Gln		U
	A	Ile	Thr	Asn	Ser	C
		Met		Lys	Arg	A
G	Val	Ala	Asp	Gly	G	
			Glu		U	
					C	
					A	
					G	
					U	
					C	
					A	
					G	
					U	
					C	
					A	
					G	

N'oubliez pas, 4 codons nécessitent 2 aa-ARNt pour être décodés. Ces 2 ARNt^{Ala} ont une séquence différentes mais un point commun ils portent le même aa: Ala du coup ont a 2 Ala-ARNt^{Ala} isoaccepteurs de la racine grecque iso – égal (identique)

II. L' aminoacyl-ARNt

		Second Codon Letter				
		U	C	A	G	
First Codon Letter	U	Phe	Ser	Tyr	Cys	U
	C	Leu	Pro	Stop	Stop	A
				Trp	G	
			Arg		U	
	A	Ile	Thr	His	Arg	C
		Met	Ala	Gln	Ser	A
Val	Asn	Arg		G		
		Lys	Gly	U		
G			Asp		C	
			Glu		A	
					G	

L' aminoacyl-ARNt synthétase spécifique de l'Asp et de l'ARNt^{Asp} s'appelle: l'Aspartyl-ARNt synthétase et est abrégée: AspRS

II. L' aminoacyl-ARNt

Second Codon Letter

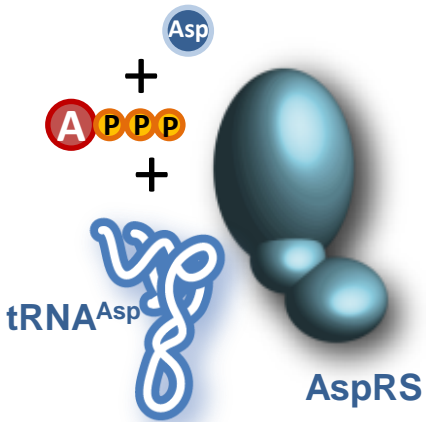
		U	C	A	G	
First Codon Letter	U	Phe	Ser	Tyr	Cys	U
	C	Leu	Pro	Stop	Stop	A
				Trp	G	
	A	Ile	Thr	His	Arg	U
		Met	Ala	Gln		C
	G	Val			Asn	Ser
				Lys	Arg	G
			Asp	Gly	U	
			Glu		C	
					A	
					G	

Third Codon Letter

L' aminoacyl-ARNt synthétase spécifique de l'Asp et de l'ARNt^{Asp} s'appelle: l'Aspartyl-ARNt synthétase et est abrégée: AspRS; de la même manière l' aminoacyl-ARNt synthétase spécifique de la Gly et de l'ARNt^{Gly} s'appelle: la Glycyl-ARNt synthétase et est abrégée: GlyRS.... Etc...

III. La synthèse de l' aminoacyl-ARNt

La réaction d' aminoacylation



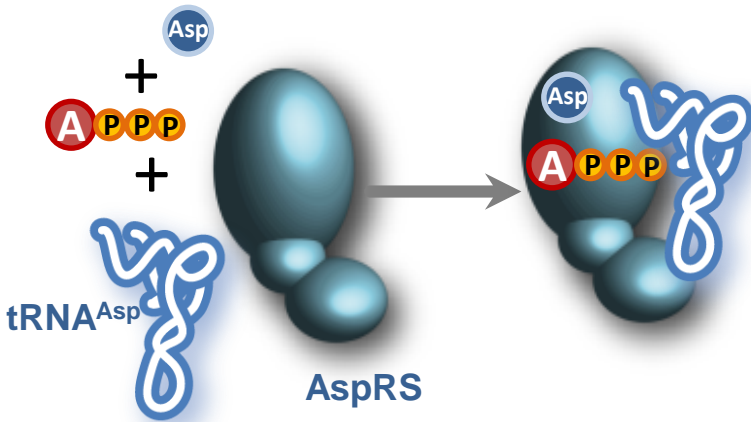
La réaction d' aminoacylation se déroule toujours en 2 étapes: **1^{ère} étape**: activation de l'acide aminé puis **2^{ème} étape** transfert de l'acide aminé activé sur l'ARNt. Comme exemple: la réaction d' aminoacylation de l'ARNt^{Asp} par l'Asp catalysée par l'AspRS.



L'AspRS a 3 substrats: l'Asp + l'ATP et l'ARNt^{Asp} et toutes les aminoacyl-ARNt synthétases ont toujours 3 substrats: leur aa homologue, l'ATP et leur ARNt homologue

III. La synthèse de l' aminoacyl-ARNt

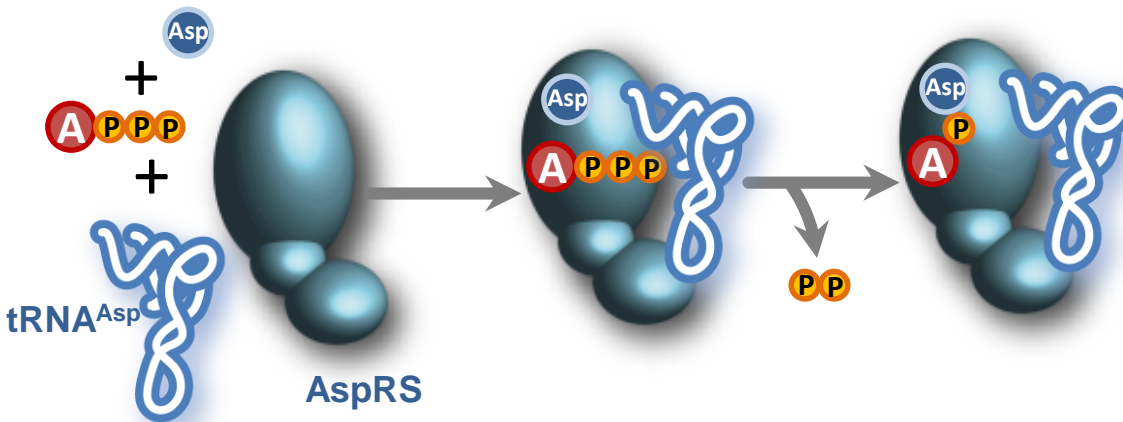
La réaction d' aminoacylation



L'AspRS va d'abord lier tous ses substrats: **ATP**, **Asp** et **ARNt^{Asp}**. Il n'y a pas d'ordre dans l'association des substrats.

III. La synthèse de l' aminoacyl-ARNt

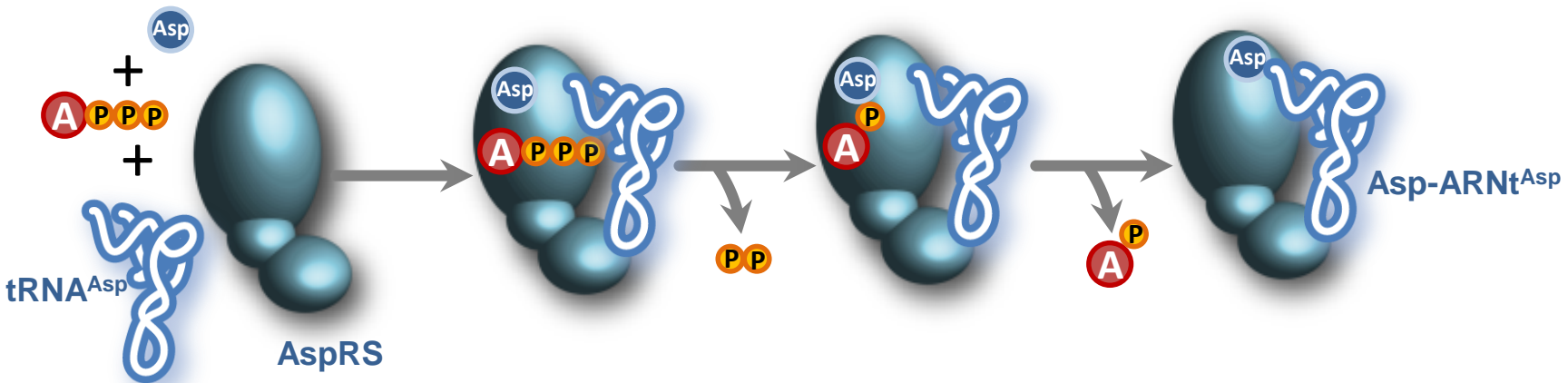
La réaction d' aminoacylation



La 1^{ère} étape catalysée par l'AspRS est l'étape d'activation qui consiste à activer l'Asp en liant l'aCOOH de l'Asp à l'AMP de l'ATP et en libérant du pyrophosphate (PP_i) = 2 phosphates de l'ATP qui sont liés entre-eux. La réaction d'activation nécessite du Mg²⁺. L'aa en l'occurrence l'Asp est activé car la liaison entre l'aCOOH de l'Asp au P de l'AMP est une liaison **anhydride mixte (acyl-phosphate)** qui est une **liaison riche en énergie** symbolisée par un ~.

III. La synthèse de l' aminoacyl-ARNt

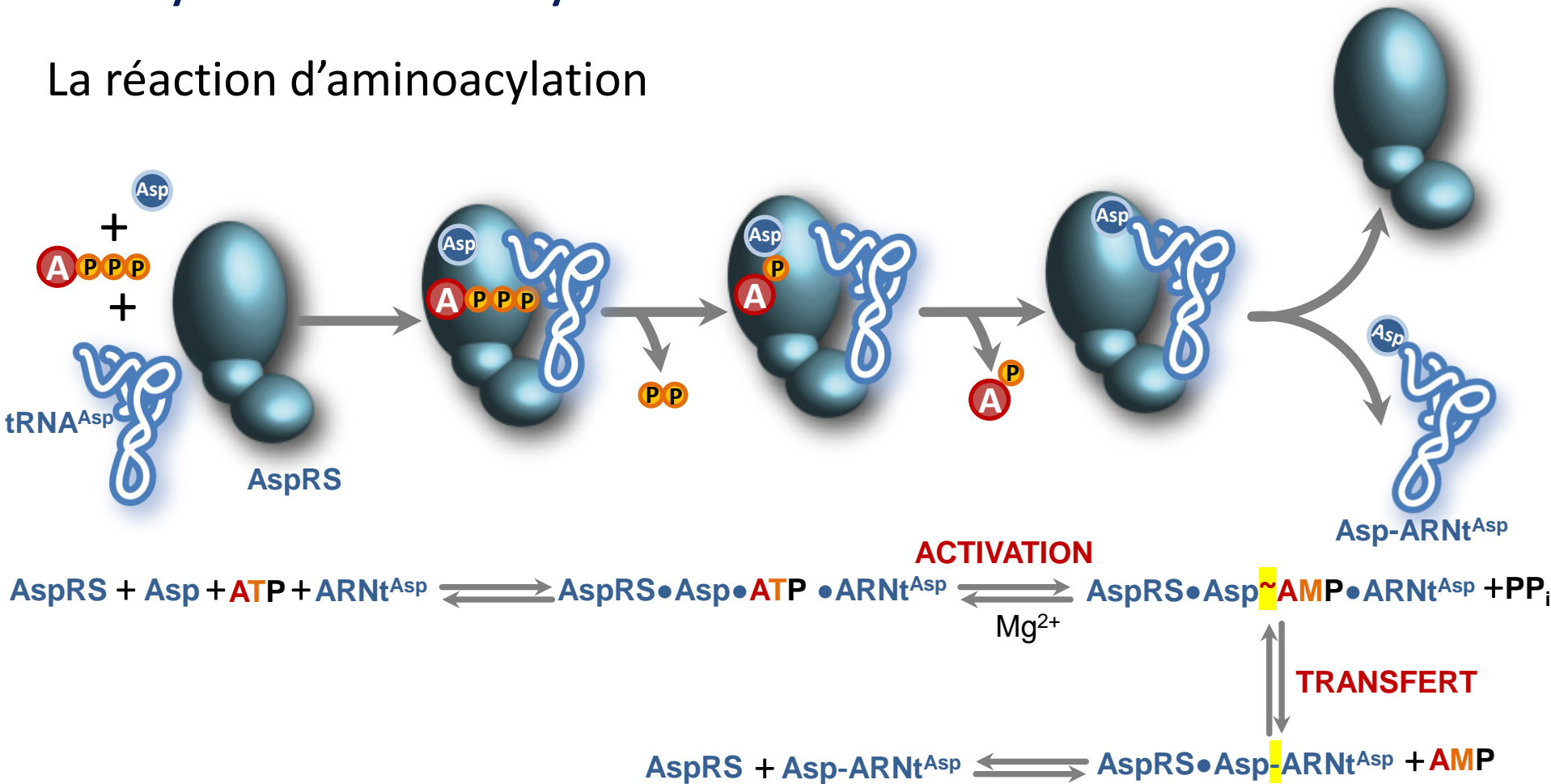
La réaction d' aminoacylation



La 2^{ème} étape catalysée par l'AspRS est l'étape de transfert de l'Asp activé sur l'ARNt^{Asp} en liant l'aCOOH de l'Asp au OH 3' ou OH 2' du ribose de l'Adénosine 76 de l'ARNt formant ainsi l'AspRS•Asp-ARNt^{Asp}. La liaison créée est une liaison ester dont la création a été possible grâce à l'énergie libérée lors du clivage de la liaison anhydride mixte riche en énergie de l'Asp-AMP. Lors du transfert l'AMP est libéré.

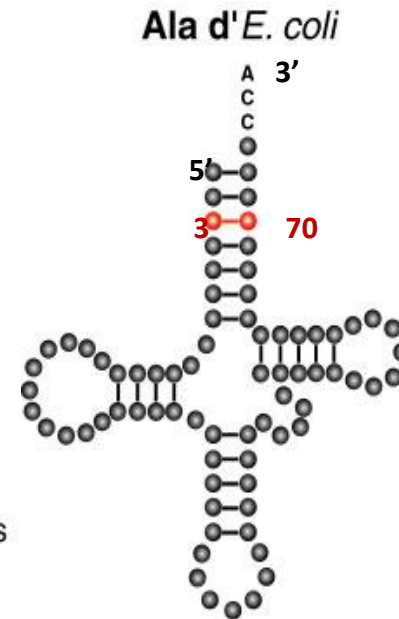
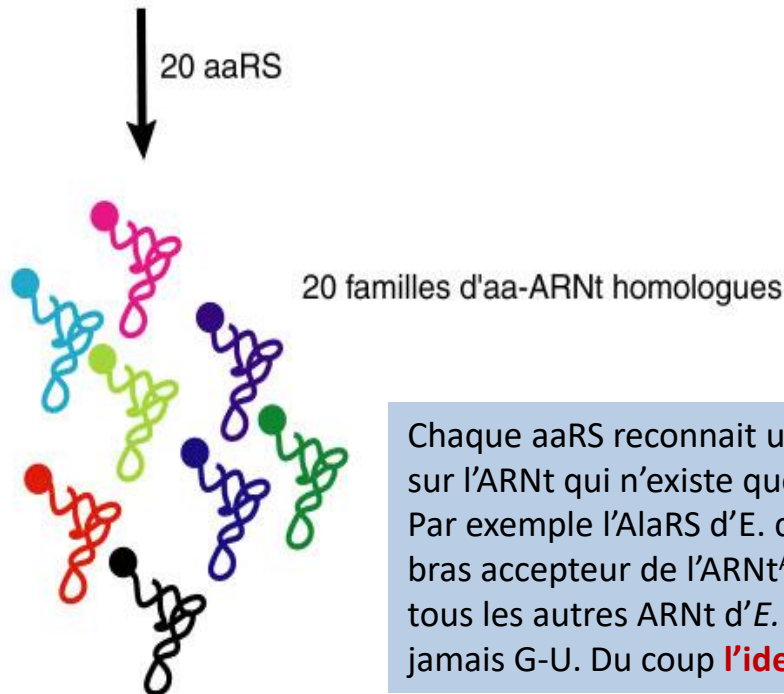
III. La synthèse de l' aminoacyl-ARNt

La réaction d' aminoacylation



Une fois le transfert effectué l'AspRS libère l'Asp-ARNt^{Asp} et peut recommencer un nouveau cycle catalytique

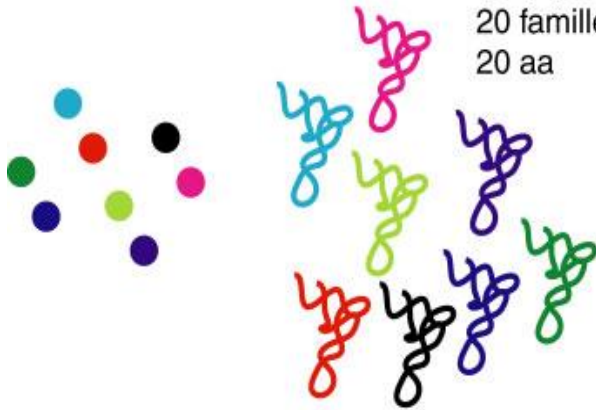
IV. L'identité d'un ARNt



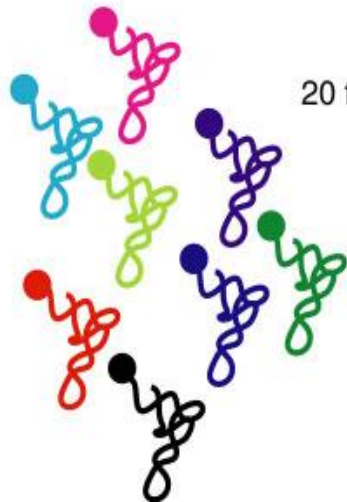
Chaque aaRS reconnaît une combinaison bien précise et unique de nucléotides sur l'ARNt qui n'existe que dans son ARNt homologue et dans aucun autre ARNt. Par exemple l'AlaRS d'*E. coli* reconnaît uniquement la 3^{ème} paire de base (pdb) du bras accepteur de l'ARNt^{Ala} qui est **G₃-U₇₀** et qui n'existe que dans l'ARNt^{Ala}. Dans tous les autres ARNt d'*E. coli* cette pdb est G-C ou C-G ou A-U ou U-A ou U-G mais jamais G-U. Du coup **l'identité Ala est G₃-U₇₀**

IV. L'identité d'un ARNt

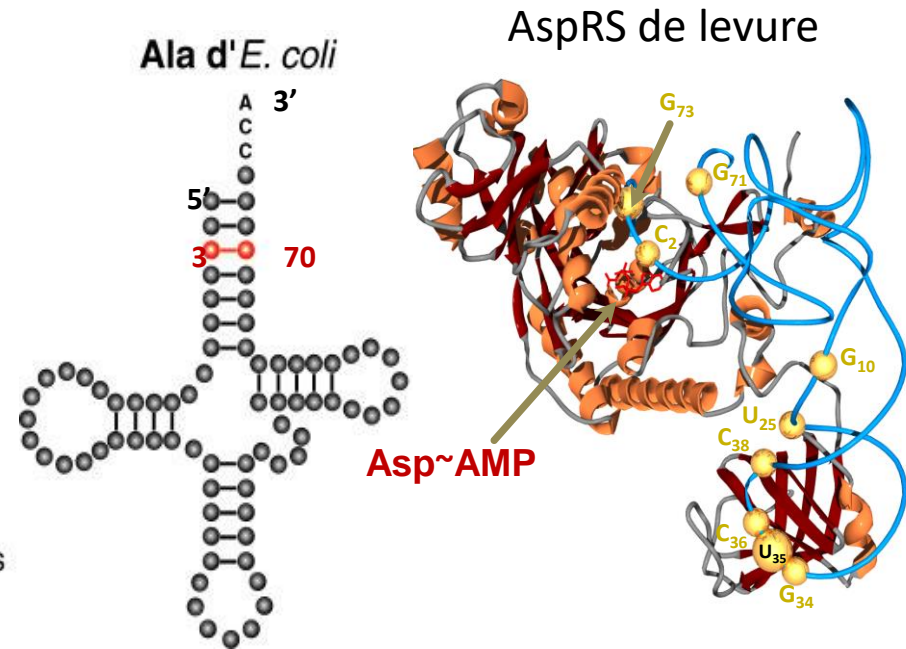
40 à 100 ARNt différents
20 familles d'ARNt isoaccepteurs
20 aa



20 aaRS



20 familles d'aa-ARNt homologues



Pour l'AspRS de levure (vous avez la structure 3D en complexe avec son ARNt et l'Asp-AMP, les déterminants d'identité sont G_{73} , C_2-G_{71} , $G_{10}-U_{25}$, $G_{34}U_{35}C_{36}$ et C_{38} (sphères dorées)

V. La Reprogrammation du code génétique

L'IG peut être modifiée à deux niveaux:

- Au niveau de l'ADN par des mutations
- Au niveau de l'ARN par des modifications post-transcriptionnelles

1. Mutations: Missens, Non sens, Suppressives

Missens

Elle **change la nature** du codon. Ex: soit le codon GGA codant pour la G une mutation se situerait sur le premier Nt du codon par ex: **G**GA→**A**GA qui code pour l'R. Ce codon ne sera plus décodé par l'**ARNt^{Gly(UCC)}** mais par l'**ARNt^{Arg(UCU)}**.

Il y aura incorporation d'**R*** à la place de **G*** → sans conséquences ou dramatique si aa du site actif d'une enzyme (perte d'activité ou ↓^{tion} de l'activité ou ↑^{tion} d'activité).

Non sens

C'est une mutation qui transforme un codon spécifiant un aa en codon **STOP** qui ne code pour aucun aa. Ex: si la 3^{ème} position du codon UAC codant pour la Y est muté en G, UAC→UAG = STOP, il y aura arrêt de la synthèse protéique (SPique*) et obtention d'une prot tronquée. Selon le lieu d'arrêt de la SPique la prot peut être inactive.

* Code à 1 lettre des aa, Spique = Synthèse Protéique

V. La Reprogrammation du code génétique

Suppressive

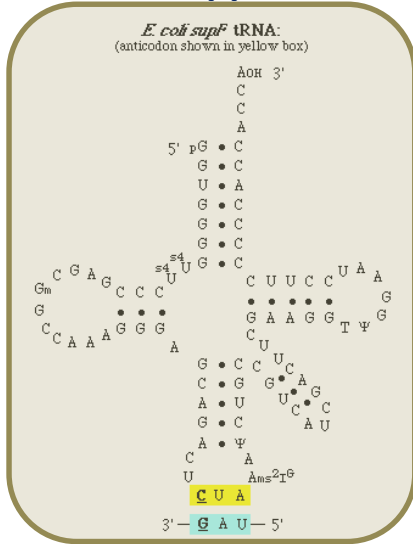
Une cellule doit se défendre contre les mutations qu'elle subit (radiations UV, ionisantes, composés chimiques...). Un des moyens pour reverser les mutations est la mutation suppressive.

- C'est une mutation **sur le tRNA** qui permet de corriger une mutation non sens par retour à la bonne lecture du codon, elle touche donc l'**@codon*** du tRNA
- Le tRNA qui a fait l'objet d'une mutation suppressive est appelé **tRNA supprimeur**
- Il existe des tRNA supprimeur naturels dont la cellule s'est dotée en prévision de ces mutation non sens, ex: le **tRNA_s^{Tyr}** dans la levure qui existe en faible quantité et qui permet de corriger la mutation non sens **UAC→UAG**.

Les tRNAs ne permettent pas de corriger tous les codons STOP car les codons STOP naturels sont placés dans un environnement de structures particulier qui permet la mise en place des étapes de terminaison de la traduction.

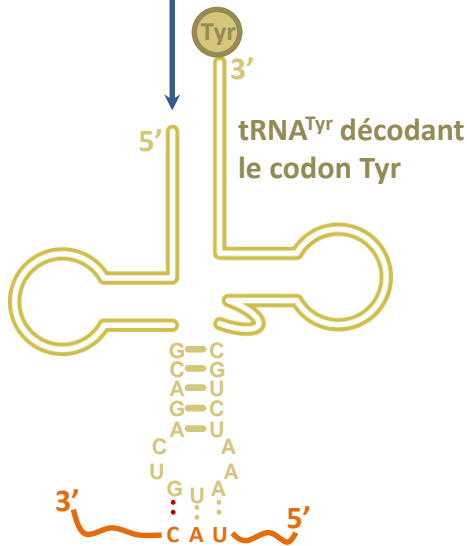
V. La Reprogrammation et dynamique du code génétique

Les tRNA suppresseurs



On peut trouver la séquence des gènes d'ARNt et du coup la séquence des ARNt en cherchant la banque de données " tRNA database " (<http://trna.bioinf.uni-leipzig.de/DataOutput/Search>) et on trouve 2 gènes.

	Organism	Amino Acid	Acc-stem	D-stem	D-loop	D-stem	Ac-stem	Ac-loop	Ac-stem	V-region	T-stem	T-loop	T-stem	Acc-stem	CCA
✓ <i>DNA</i>	<i>Escherichia coli K12, K12-MG1655 Tyr</i>	-	GGTGGGG	TT CCCG	AGC--GGCCAA	AGGG	A GCAGA	CTGTAAA	TCTGC	CGTC-----ATCG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA
✓ <i>DNA</i>	<i>Escherichia coli K12, K12-MG1655 Tyr</i>	-	GGTGGGG	TT CCCG	AGC--GGCCAA	AGGG	A GCAGA	CTGTAAA	TCTGC	CGTC-----ACAG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA

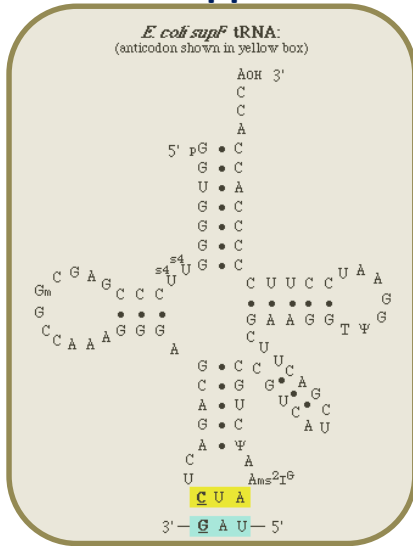


mRNA d'un gène essentiel contenant un codon Tyr UAC

:

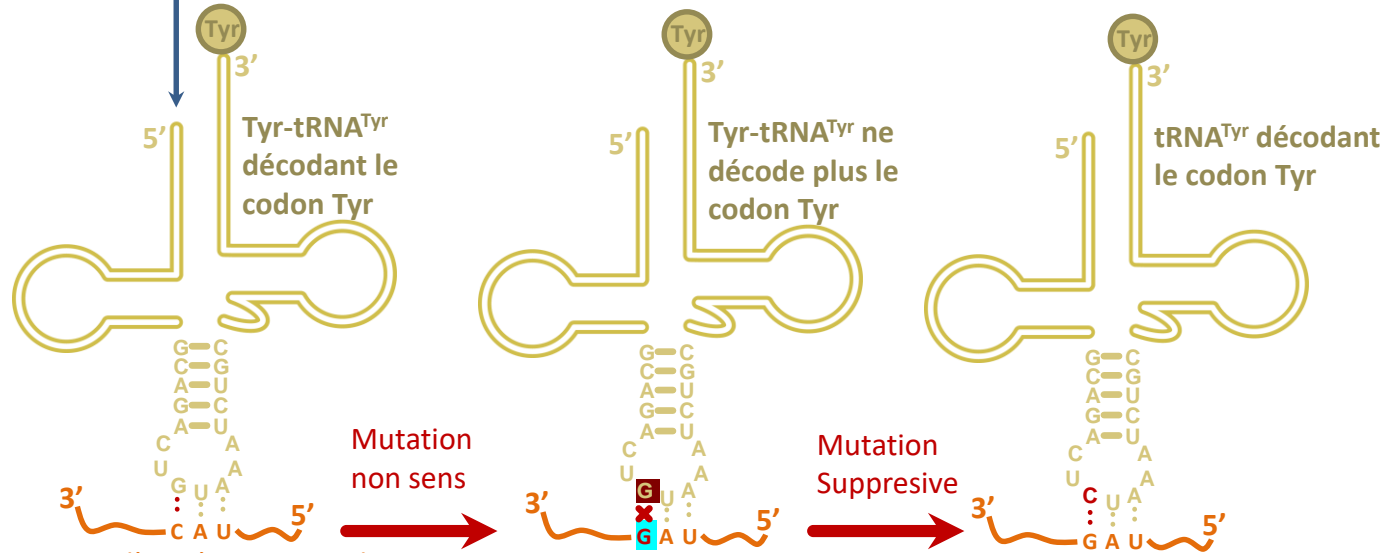
V. La Reprogrammation et dynamique du code génétique

Les tRNA supprimeurs



On peut trouver la séquence des gènes d'ARNt et du coup la séquence des ARNt en cherchant la banque de données " tRNA database " (<http://trna.bioinf.uni-leipzig.de/DataOutput/Search>) et on trouve 2 gènes.

	Organism	Amino Acid	Acc-stem	D-stem	D-loop	D-stem	Ac-stem	Ac-loop	Ac-stem	V-region	T-stem	T-loop	T-stem	Acc-stem	CCA
✓ <i>Dna</i>	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	-1 1	8 10	14	22	26 27	32	39	44	49 53	53	61	66	73 74
✓ <i>Dna</i>	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	-	GGTGGGG TT CCCG	AGC--GGCCAA AGGG	A	GCAGA	CTGTAAA	TCTGC	CGTC-----ATCG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA
✓ <i>Dna</i>	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	-	GGTGGGG TT CCCG	AGC--GGCCAA AGGG	A	GCAGA	CTGTAAA	TCTGC	CGTC-----ACAG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA



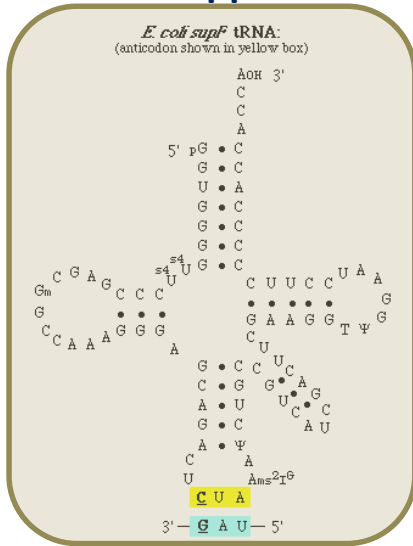
mRNA d'un gène essentiel contenant un codon Tyr UAC

Le codon Tyr CAU est devenu un codon **STOP** ambre UAG, le G en 3' du codon ne peut pas faire de pdb avec le G en 5' de l'@codon

:

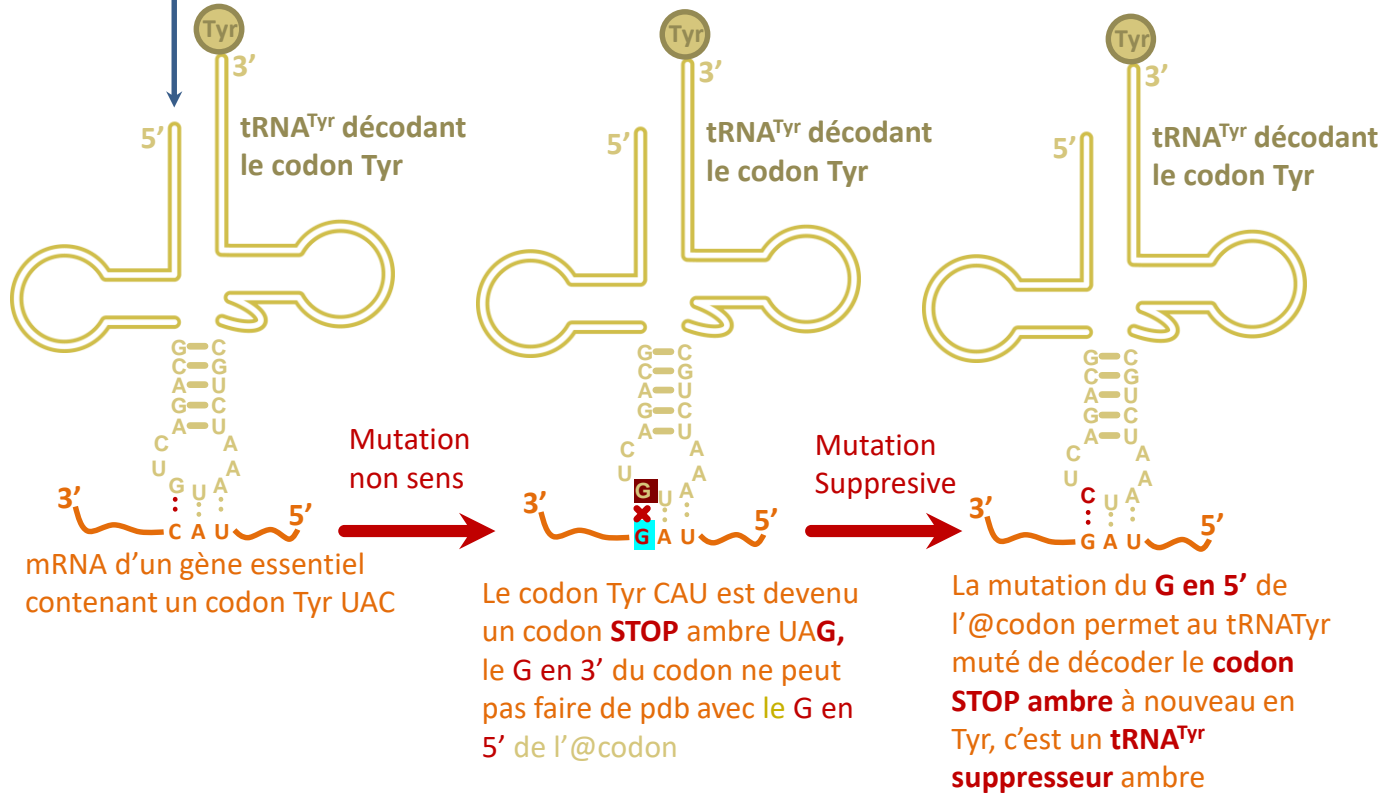
V. La Reprogrammation et dynamique du code génétique

Les tRNA suppresseurs



On peut trouver la séquence des gènes d'ARNt et du coup la séquence des ARNt en cherchant la banque de données " tRNA database " (<http://trna.bioinf.uni-leipzig.de/DataOutput/Search>) et on trouve 2 gènes.

	Organism	Amino Acid	Acc-stem	D-stem	D-loop	D-stem	Ac-stem	Ac-loop	Ac-stem	V-region	T-stem	T-loop	T-stem	Acc-stem	CCA
✓ DNA	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	-1 1	8 10	14	22	26 27	32	39	44	49 53	57	61	66	73 74
✓ DNA	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	- GGTGGGG	TT CCCG	AGC--GGCCAA	AGGG A	GCAGA	CTGTAAA	TCTGC	CGTC-----ATCG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA
✓ DNA	<i>Escherichia coli</i> K12, K12-MG1655	Tyr	- GGTGGGG	TT CCCG	AGC--GGCCAA	AGGG A	GCAGA	CTGTAAA	TCTGC	CGTC-----ACAG-----ACTTC	GAAGG	TTCGAAT	CCTTC	CCCCACC	CCA



V. La Reprogrammation et dynamique du code génétique



Il existe 22 aa naturels génétiquement codés:

21^{ème} = Sélénocystéine (Sec) = (U)

22^{ème} = Pyrrolysine (Pyl)

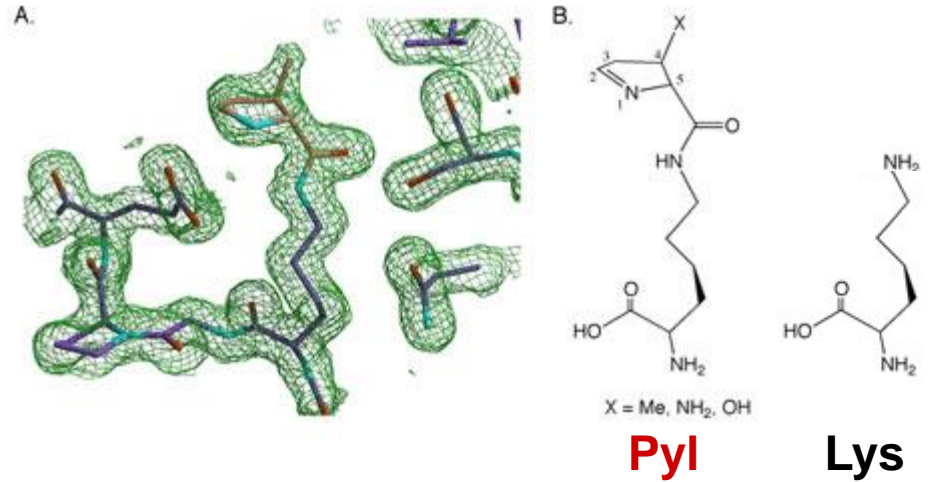
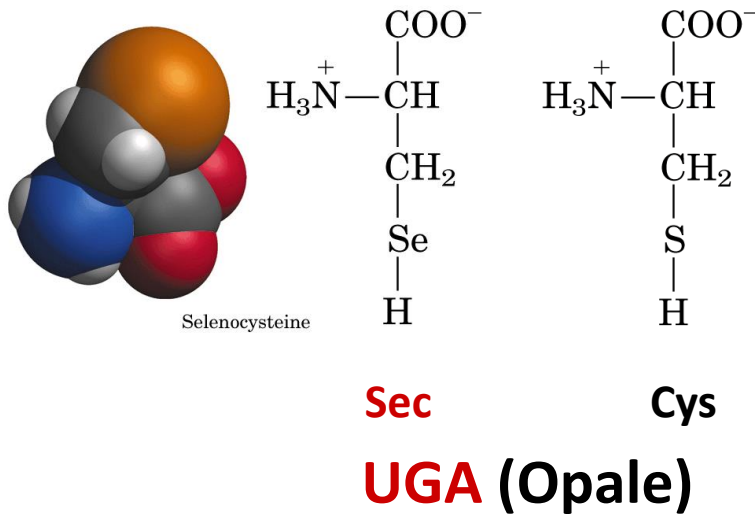


Certain codons STOP peuvent être codants, ils sont redéfinis en codons codant un aa

V. La Reprogrammation et dynamique du code génétique

2. Reprogrammation de codons STOP

A l'heure actuelle 2 aa sont codés par reprogrammation contexte-dépendante des codons STOP il s'agit du 21^{ème} aa du CG: la **sélocystéine** (Sec) et du 22^{ème} aa: la **Pyrrrolysine** (Pyl).



UAG (ambre)

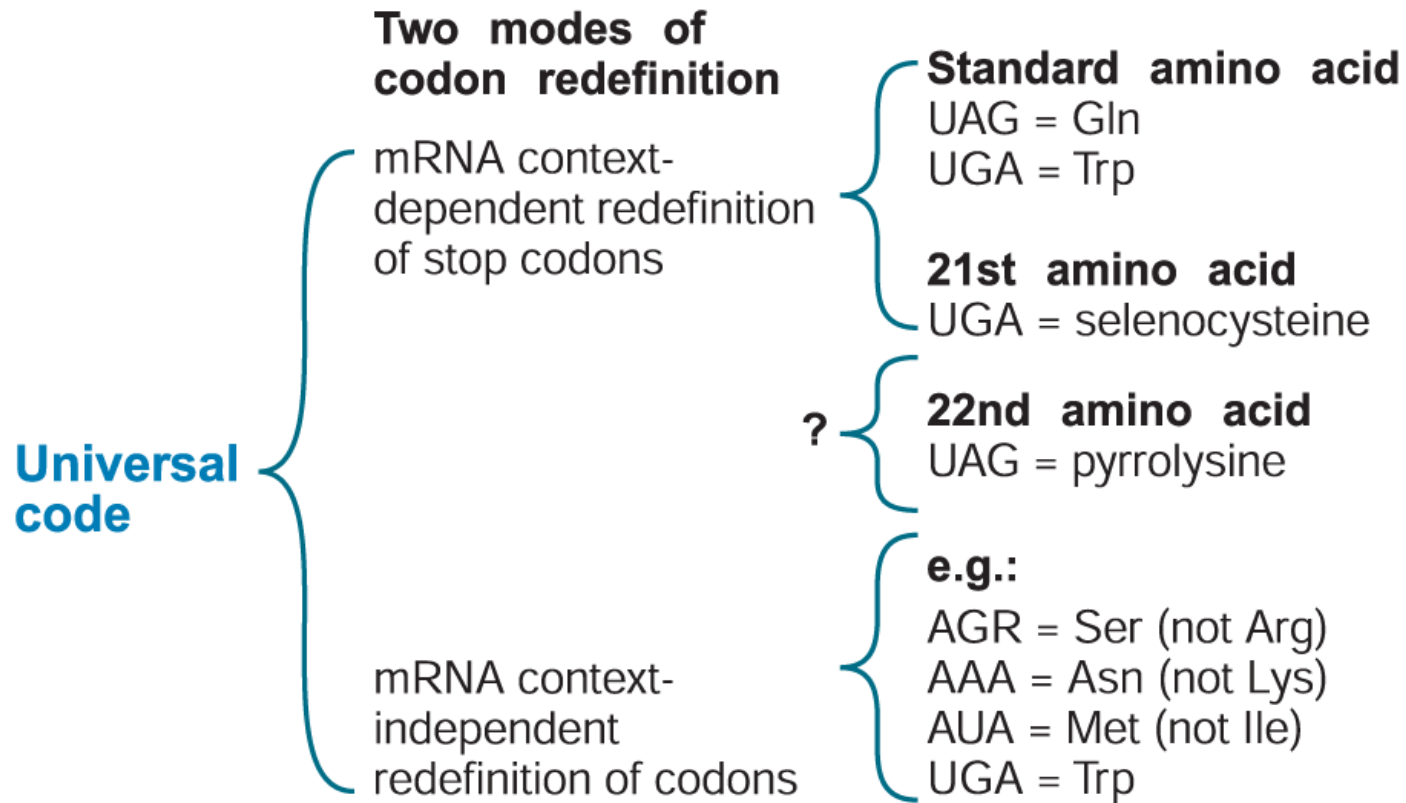


Dans ce cas, le codon STOP est toujours un STOP sauf lorsque le mRNA d'un gène présente une **structure secondaire et/ou tertiaire** indiquant au ribosome que ce codon STOP code pour un nouvel aa

V. La Reprogrammation et dynamique du code génétique

2. Reprogrammation de codons STOP

Il existe des codons STOP qui sont situés à l'intérieur de régions codantes. La synthèse protéique ne s'arrête pas à ces codons STOP par ce que ceux-ci sont décodés par un aa. Il existe 2 modes de reprogrammation de codons



VI. Universalité ou déviations du code génétique

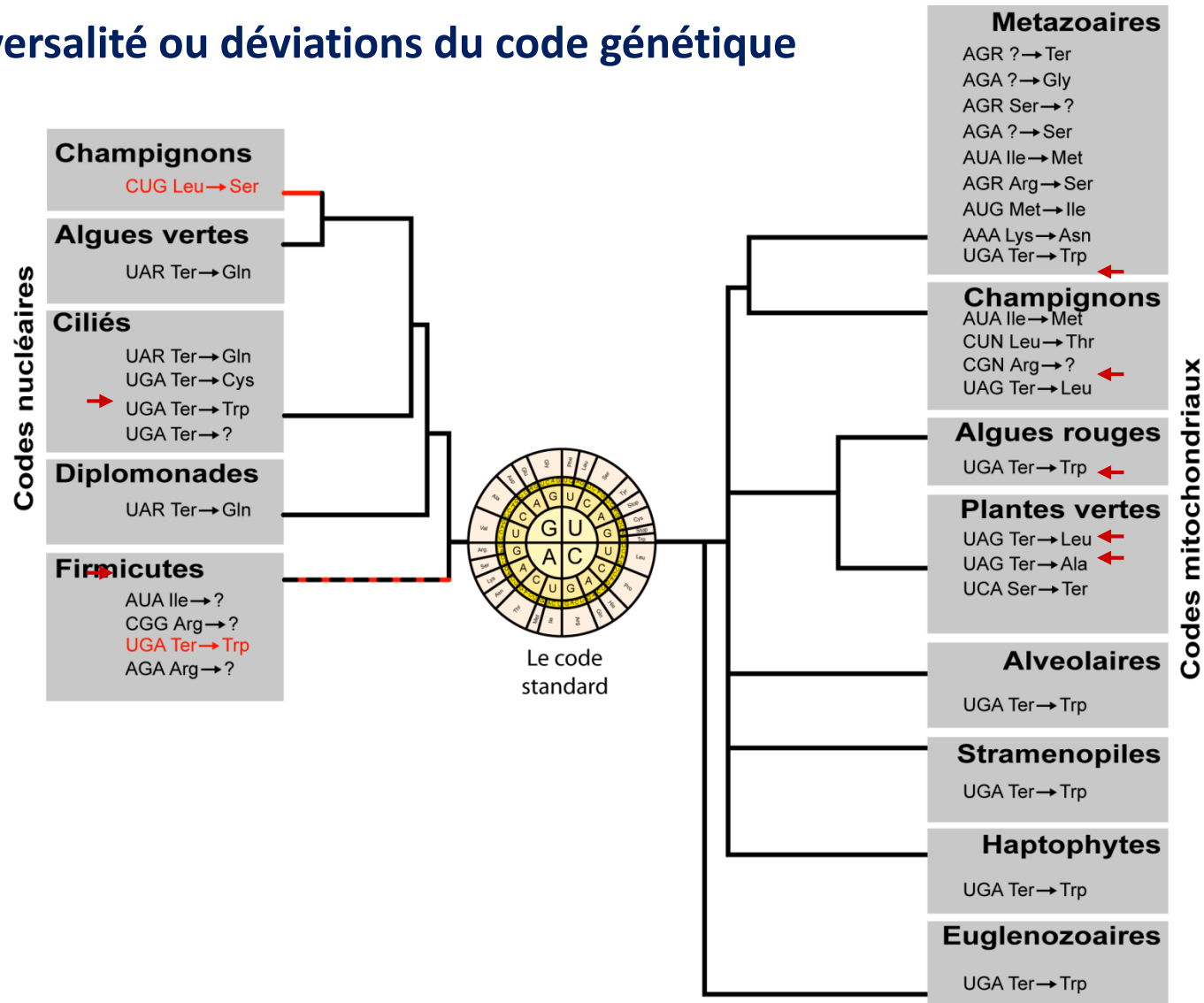
Crick, F. H. (1968). The origin of the genetic code. J. Mol. Biol. 38, 367-379.

Le code génétique découvert par Marshal W. Nirenberg (Nirenberg et Matthaei, 1961) dans les années 60 a longtemps été considéré comme un accident figé ou "frozen accident" (Crick, 1968), incapable d'évoluer. L'aspect immuable du code génétique tirait son origine de la simple hypothèse logique que tout changement de ce code conduirait à l'altération de la signification des codons et donc par là même provoquerait l'introduction d'erreurs lors de la traduction de chaque ARN messenger.



Le CG est quasiment universel les principales déviations sont observées dans les génomes mitochondriaux

VI. Universalité ou déviations du code génétique



Vous trouvez ici une compilation des principale déviations par rapport au code Standard, si vous ne deviez en retenir qu'une seule **UGA = Trp**

VII. Fréquence d'usage des codons

⚠ Peut-il y avoir 6 codons réellement codants pour un aa ? La réponse est oui, tous les codons sont toujours utilisés

⚠ Certains codons peuvent-ils ne jamais être utilisés ? La réponse est quasi toujours non, tous les codons sont généralement utilisés (mais ça peut arriver pour un STOP)

⚠ Certains codons sont-ils plus utilisés que d'autres ? La réponse est oui, c'est l'usage des codons et il est spécifique de chaque organisme

Vous allez sur le site et vous rentrez les noms d'organismes suivant: *Escherichia coli*; *Thermus thermophilus*, *Saccharomyces cerevisiae* et *Homo Sapiens* (1 à la fois)

<http://www.kazusa.or.jp/codon/>

Codon Usage Database

Data source

[NCBI-GenBank](#) Flat File Release 160.0 [June 15 2007].

Data amount

35,799 organisms
3,027,973 complete protein coding genes (CDS's)

[Announcement](#)

QUERY Box for search with Latin name of organism

Case: sensitive insensitive

Input a scientific name (or its [regular expression](#)) for an organism and press "Submit" or return key. Use **Latin name** such as "Marchantia polymorpha", "Saccharomyces cerevisiae" etc., not "liverwort", "yeast" etc.

VII. Fréquence d'usage des codons

Vous obtenez les 4 tableaux suivants:

Escherichia coli K12 [gbtct]: 14 CDS's (5122 codons)

fields: [triplet] [frequency: per thousand] [(number)]

UUU 19.7(101)	UCU 5.7(29)	UAU 16.8(86)	UGU 5.9(30)
UUC 15.0(77)	UCC 5.5(28)	UAC 14.6(75)	UGC 8.0(41)
UUA 15.2(78)	UCA 7.8(40)	UAA 1.8(9)	UGA 1.0(5)
UUG 11.9(61)	UCG 8.0(41)	UAG 0.0(0)	UGG 10.7(55)
CUU 11.9(61)	CCU 8.4(43)	CAU 15.8(81)	CGU 21.1(108)
CUC 10.5(54)	CCC 6.4(33)	CAC 13.1(67)	CGC 26.0(133)
CUA 5.3(27)	CCA 6.6(34)	CAA 12.1(62)	CGA 4.3(22)
CUG 46.9(240)	CCG 26.7(137)	CAG 27.7(142)	CGG 4.1(21)
AUU 30.5(156)	ACU 8.0(41)	AAU 21.9(112)	AGU 7.2(37)
AUC 18.2(93)	ACC 22.8(117)	AAC 24.4(125)	AGC 16.6(85)
AUA 3.7(19)	ACA 6.4(33)	AAA 33.2(170)	AGA 1.4(7)
AUG 24.8(127)	ACG 11.5(59)	AAG 12.1(62)	AGG 1.6(8)
GUU 16.8(86)	GCU 10.7(55)	GAU 37.9(194)	GGU 21.3(109)
GUC 11.7(60)	GCC 31.6(162)	GAC 20.5(105)	GGC 33.4(171)
GUA 11.5(59)	GCA 21.1(108)	GAA 43.7(224)	GGA 9.2(47)
GUG 26.4(135)	GCG 38.5(197)	GAG 18.4(94)	GGG 8.6(44)

Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%

Thermus thermophilus HB8 [gbtct]: 2238 CDS's (669747 codons)

fields: [triplet] [frequency: per thousand] [(number)]

UUU 6.7(4471)	UCU 0.7(495)	UAU 1.2(830)	UGU 0.2(125)
UUC 30.9(20692)	UCC 15.3(10233)	UAC 27.5(18407)	UGC 3.7(2493)
UUA 1.1(744)	UCA 0.3(210)	UAA 0.6(431)	UGA 1.4(970)
UUG 9.6(6428)	UCG 3.8(2561)	UAG 1.2(837)	UGG 13.8(9274)
CUU 16.9(11332)	CCU 4.5(3009)	CAU 0.9(582)	CGU 1.5(1006)
CUC 72.9(48852)	CCC 48.0(32153)	CAC 17.7(11824)	CGC 29.0(19436)
CUA 3.3(2184)	CCA 1.4(969)	CAA 3.1(2076)	CGA 1.1(766)
CUG 41.0(27438)	CCG 11.6(7757)	CAG 21.2(14228)	CGG 36.8(24675)
AUU 2.4(1594)	ACU 0.4(291)	AAU 0.4(273)	AGU 0.4(263)
AUC 22.3(14902)	ACC 27.0(18095)	AAC 15.1(10097)	AGC 13.8(9217)
AUA 1.2(789)	ACA 0.4(298)	AAA 2.9(1930)	AGA 0.8(533)
AUG 14.4(9621)	ACG 9.5(6356)	AAG 32.6(21825)	AGG 16.3(10939)
GUU 2.5(1706)	GCU 2.6(1751)	GAU 1.7(1112)	GGU 2.2(1456)
GUC 27.0(18068)	GCC 84.7(56708)	GAC 34.1(22864)	GGC 37.0(24778)
GUA 1.5(992)	GCA 1.7(1134)	GAA 10.6(7089)	GGA 5.5(3662)
GUG 50.1(33558)	GCG 26.2(17562)	GAG 75.3(50430)	GGG 48.3(32366)

Coding GC 69.67% 1st letter GC 72.19% 2nd letter GC 45.02% 3rd letter GC 91.78%

Saccharomyces cerevisiae [gbpln]: 14411 CDS's (6534504 codons)

fields: [triplet] [frequency: per thousand] [(number)]

UUU 26.1(170666)	UCU 23.5(153557)	UAU 18.8(122728)	UGU 8.1(52903)
UUC 18.4(120510)	UCC 14.2(92923)	UAC 14.8(96596)	UGC 4.8(31095)
UUA 26.2(170884)	UCA 18.7(122028)	UAA 1.1(6913)	UGA 0.7(4447)
UUG 27.2(177573)	UCG 8.6(55951)	UAG 0.5(3312)	UGG 10.4(67789)
CUU 12.3(80076)	CCU 13.5(88263)	CAU 13.6(89007)	CGU 6.4(41791)
CUC 5.4(35545)	CCC 6.8(44309)	CAC 7.8(50785)	CGC 2.6(16993)
CUA 13.4(87619)	CCA 18.3(119641)	CAA 27.3(178251)	CGA 3.0(19562)
CUG 10.5(68494)	CCG 5.3(34597)	CAG 12.1(79121)	CGG 1.7(11351)
AUU 30.1(196893)	ACU 20.3(132522)	AAU 35.7(233124)	AGU 14.2(92466)
AUC 17.2(112176)	ACC 12.7(83207)	AAC 24.8(162199)	AGC 9.8(63726)
AUA 17.8(116254)	ACA 17.8(116084)	AAA 41.9(273618)	AGA 21.3(139081)
AUG 20.9(136805)	ACG 8.0(52045)	AAG 30.8(201361)	AGG 9.2(60289)
GUU 22.1(144243)	GCU 21.2(138358)	GAU 37.6(245641)	GGU 23.9(156109)
GUC 11.8(76947)	GCC 12.6(82357)	GAC 20.2(132048)	GGC 9.8(63903)
GUA 11.8(76927)	GCA 16.2(105910)	GAA 45.6(297944)	GGA 10.9(71216)
GUG 10.8(70337)	GCG 6.2(40358)	GAG 19.2(125717)	GGG 6.0(39359)

Coding GC 39.77% 1st letter GC 44.58% 2nd letter GC 36.64% 3rd letter GC 38.10%

Homo sapiens [gbpri]: 93487 CDS's (40662582 codons)

fields: [triplet] [frequency: per thousand] [(number)]

UUU 17.6(714298)	UCU 15.2(618711)	UAU 12.2(495699)	UGU 10.6(430311)
UUC 20.3(824692)	UCC 17.7(718892)	UAC 15.3(622407)	UGC 12.6(513028)
UUA 7.7(311881)	UCA 12.2(496448)	UAA 1.0(40285)	UGA 1.6(63237)
UUG 12.9(525688)	UCG 4.4(179419)	UAG 0.8(32109)	UGG 13.2(535595)
CUU 13.2(536515)	CCU 17.5(713233)	CAU 10.9(441711)	CGU 4.5(184609)
CUC 19.6(796638)	CCC 19.8(804620)	CAC 15.1(613713)	CGC 10.4(423516)
CUA 7.2(290751)	CCA 16.9(688038)	CAA 12.3(501911)	CGA 6.2(250760)
CUG 39.6(1611801)	CCG 6.9(281570)	CAG 34.2(1391973)	CGG 11.4(464485)
AUU 16.0(650473)	ACU 13.1(533609)	AAU 17.0(689701)	AGU 12.1(493429)
AUC 20.8(846466)	ACC 18.9(768147)	AAC 19.1(776603)	AGC 19.5(791383)
AUA 7.5(304565)	ACA 15.1(614523)	AAA 24.4(993621)	AGA 12.2(494682)
AUG 22.0(896005)	ACG 6.1(246105)	AAG 31.9(1295568)	AGG 12.0(486463)
GUU 11.0(448607)	GCU 18.4(750096)	GAU 21.8(885429)	GGU 10.8(437126)
GUC 14.5(588138)	GCC 27.7(1127679)	GAC 25.1(1020595)	GGC 22.2(903565)
GUA 7.1(287712)	GCA 15.8(643471)	GAA 29.0(1177632)	GGA 16.5(669873)
GUG 28.1(1143534)	GCG 7.4(299495)	GAG 39.6(1609975)	GGG 16.5(669768)

Coding GC 52.27% 1st letter GC 55.72% 2nd letter GC 42.54% 3rd letter GC 58.55%

VII. Fréquence d'usage des codons

Si je compare 2 bactéries, 1 mésophile: *E. coli* (37 °C) et 1 thermophile *T. Thermophilus* (75 °C)

Escherichia coli K12 [gbbet]: 14 CDS's (5122 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU 19.7(101)	UCU 5.7(29)	UAU 16.8(86)	UGU 5.9(30)
UUC 15.0(77)	UCC 5.5(28)	UAC 14.6(75)	UGC 8.0(41)
UUA 15.2(78)	UCA 7.8(40)	UAA 1.8(9)	UGA 1.0(5)
UUG 11.9(61)	UCG 8.0(41)	UAG 0.0(0)	UGG 10.7(55)
CUU 11.9(61)	CCU 8.4(43)	CAU 15.8(81)	CGU 21.1(108)
CUC 10.5(54)	CCC 6.4(33)	CAC 13.1(67)	CGC 26.0(133)
CUA 5.3(27)	CCA 6.6(34)	CAA 12.1(62)	CGA 4.3(22)
CUG 46.9(240)	CCG 26.7(137)	CAG 27.7(142)	CGG 4.1(21)
AUU 30.5(156)	ACU 8.0(41)	AAU 21.9(112)	AGU 7.2(37)
AUC 18.2(93)	ACC 22.8(117)	AAC 24.4(125)	AGC 16.6(85)
AUA 3.7(19)	ACA 6.4(33)	AAA 33.2(170)	AGA 1.4(7)
AUG 24.8(127)	ACG 11.5(59)	AAG 12.1(62)	AGG 1.6(8)
GUU 16.8(86)	GCU 10.7(55)	GAU 37.9(194)	GGU 21.3(109)
GUC 11.7(60)	GCC 31.6(162)	GAC 20.5(105)	GGC 33.4(171)
GUA 11.5(59)	GCA 21.1(108)	GAA 43.7(224)	GGA 9.2(47)
GUG 26.4(135)	GCG 38.5(197)	GAG 18.4(94)	GGG 8.6(44)

Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%

Thermus thermophilus HB8 [gbbet]: 2238 CDS's (669747 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU 6.7(4471)	UCU 0.7(495)	UAU 1.2(830)	UGU 0.2(125)
UUC 30.9(20692)	UCC 15.3(10233)	UAC 27.5(18407)	UGC 3.7(2493)
UUA 1.1(744)	UCA 0.3(210)	UAA 0.6(431)	UGA 1.4(970)
UUG 9.6(6428)	UCG 3.8(2561)	UAG 1.2(837)	UGG 13.8(9274)
CUU 16.9(11332)	CCU 4.5(3009)	CAU 0.9(582)	CGU 1.5(1006)
CUC 72.9(48852)	CCC 48.0(32153)	CAC 17.7(11824)	CGC 29.0(19436)
CUA 3.3(2184)	CCA 1.4(969)	CAA 3.1(2076)	CGA 1.1(766)
CUG 41.0(27438)	CCG 11.6(7757)	CAG 21.2(14228)	CGG 36.8(24675)
AUU 2.4(1594)	ACU 0.4(291)	AAU 0.4(273)	AGU 0.4(263)
AUC 22.3(14902)	ACC 27.0(18095)	AAC 15.1(10097)	AGC 13.8(9217)
AUA 1.2(789)	ACA 0.4(298)	AAA 2.9(1930)	AGA 0.8(533)
AUG 14.4(9621)	ACG 9.5(6356)	AAG 32.6(21825)	AGG 16.3(10939)
GUU 2.5(1706)	GCU 2.6(1751)	GAU 1.7(1112)	GGU 2.2(1456)
GUC 27.0(18068)	GCC 84.7(56708)	GAC 34.1(22864)	GGC 37.0(24778)
GUA 1.5(992)	GCA 1.7(1134)	GAA 10.6(7089)	GGA 5.5(3662)
GUG 50.1(33558)	GCG 26.2(17562)	GAG 75.3(50430)	GGG 48.3(32366)

Coding GC 69.67% 1st letter GC 72.19% 2nd letter GC 45.02% 3rd letter GC 91.78%

Usage différent dû à la température physiologique

Codons moins fréquents chez *Tth* que chez *Eco*: ceux se terminant par A ou U puisque *Tth* a 75% de G-C

Comme *E. coli* est souvent utilisé pour la surexpression de protéines par transformation avec un plasmide de surexpression, il est important de savoir quels codons de *Tth* sont très nettement plus utilisés par *Tth* que par *Eco* = codons rares de *Eco* vs* *Tth*

codons rares de *Eco* vs *Tth*

Un codon rare est décodé par un ARNt isoaccepteur mineur, cela veut dire que la quantité (concentration) de cet ARNt est faible dans l'organisme.

Si je veux surexprimer une protéine de *Tth* dans *Eco* il faut faire en sorte que *Eco* ait l'usage des codons le plus proche de *Tth*.

*comparé à

VII. Fréquence d'usage des codons

Si je compare 2 bactéries, 1 mésophile: *E. coli* (37 °C) et 1 thermophile *T. Thermophilus* (75 °C)

***Escherichia coli* K12 [gbtct]: 14 CDS's (5122 codons)**

fields: [triplet] [frequency: per thousand] ([number])

UUU 19.7(101)	UCU 5.7(29)	UAU 16.8(86)	UGU 5.9(30)
UUC 15.0(77)	UCC 5.5(28)	UAC 14.6(75)	UGC 8.0(41)
UUA 15.2(78)	UCA 7.8(40)	UAA 1.8(9)	UGA 1.0(5)
UUG 11.9(61)	UCG 8.0(41)	UAG 0.0(0)	UGG 10.7(55)
CUU 11.9(61)	CCU 8.4(43)	CAU 15.8(81)	CGU 21.1(108)
CUC 10.5(54)	CCC 6.4(33)	CAC 13.1(67)	CGC 26.0(133)
CUA 5.3(27)	CCA 6.6(34)	CAA 12.1(62)	CGA 4.3(22)
CUG 46.9(240)	CCG 26.7(137)	CAG 27.7(142)	CGG 4.1(21)
AUU 30.5(156)	ACU 8.0(41)	AAU 21.9(112)	AGU 7.2(37)
AUC 18.2(93)	ACC 22.8(117)	AAC 24.4(125)	AGC 16.6(85)
AUA 3.7(19)	ACA 6.4(33)	AAA 33.2(170)	AGA 1.4(7)
AUG 24.8(127)	ACG 11.5(59)	AAG 12.1(62)	AGG 1.6(8)
GUU 16.8(86)	GCU 10.7(55)	GAU 37.9(194)	GGU 21.3(109)
GUC 11.7(60)	GCC 31.6(162)	GAC 20.5(105)	GGC 33.4(171)
GUA 11.5(59)	GCA 21.1(108)	GAA 43.7(224)	GGA 9.2(47)
GUG 26.4(135)	GCG 38.5(197)	GAG 18.4(94)	GGG 8.6(44)

Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%

Les 3 codons les plus problématiques sont:

- STOP ambre UAG: 837 fois + dans *Tth* que *Eco*
- Leu (L) CUC: 7 fois + dans *Tth* que *Eco*
- Arg (R) AGG: 10 fois + dans *Tth* que *Eco*

Solution:

Utiliser la souche d'*E. coli* BL21-CodonPlus (DE3)-RIL

<https://www.ncbi.nlm.nih.gov/pubmed/10910733>

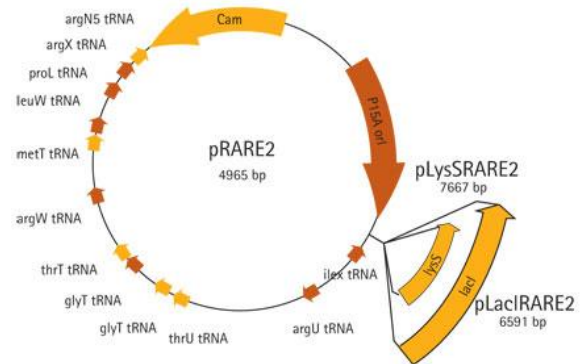
E. coli BL21-CodonPlus (DE3)-RIL contient un plasmide *pRARE* surexprimant les **ARNt^{Arg}** (R), **ARNt^{Ile}** (I) et **ARNt^{Leu}** (L) mineur qui décode les codons R I L rares d'*Eco*. Du coup **ces codons ne sont plus rares**

***Thermus thermophilus* HB8 [gbtct]: 2238 CDS's (669747 codons)**

fields: [triplet] [frequency: per thousand] ([number])

UUU 6.7(4471)	UCU 0.7(495)	UAU 1.2(830)	UGU 0.2(125)
UUC 30.9(20692)	UCC 15.3(10233)	UAC 27.5(18407)	UGC 3.7(2493)
UUA 1.1(744)	UCA 0.3(210)	UAA 0.6(431)	UGA 1.4(970)
UUG 9.6(6428)	UCG 3.8(2561)	UAG 1.2(837)	UGG 13.8(9274)
CUU 16.9(11332)	CCU 4.5(3009)	CAU 0.9(582)	CGU 1.5(1006)
CUC 72.9(48852)	CCC 48.0(32153)	CAC 17.7(11824)	CGC 29.0(19436)
CUA 3.3(2184)	CCA 1.4(969)	CAA 3.1(2076)	CGA 1.1(766)
CUG 41.0(27438)	CCG 11.6(7757)	CAG 21.2(14228)	CGG 36.8(24675)
AUU 2.4(1594)	ACU 0.4(291)	AAU 0.4(273)	AGU 0.4(263)
AUC 22.3(14902)	ACC 27.0(18095)	AAC 15.1(10097)	AGC 13.8(9217)
AUA 1.2(789)	ACA 0.4(298)	AAA 2.9(1930)	AGA 0.8(533)
AUG 14.4(9621)	ACG 9.5(6356)	AAG 32.6(21825)	AGG 16.3(10939)
GUU 2.5(1706)	GCU 2.6(1751)	GAU 1.7(1112)	GGU 2.2(1456)
GUC 27.0(18068)	GCC 84.7(56708)	GAC 34.1(22864)	GGC 37.0(24778)
GUA 1.5(992)	GCA 1.7(1134)	GAA 10.6(7089)	GGA 5.5(3662)
GUG 50.1(33558)	GCG 26.2(17562)	GAG 75.3(50430)	GGG 48.3(32366)

Coding GC 69.67% 1st letter GC 72.19% 2nd letter GC 45.02% 3rd letter GC 91.78%



https://www.merckmillipore.com/INTERSHOP/static/WFS/Merck-Site-/Merck/en_US/Freestyle/BI-Bioscience/Genomic-Analysis/pRARE2.jpg

VII. Fréquence d'usage des codons

Vous obtenez les 4 tableaux suivants:

Escherichia coli K12 [gbbet]: 14 CDS's (5122 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU 19.7(101)	UCU 5.7(29)	UAU 16.8(86)	UGU 5.9(30)
UUC 15.0(77)	UCC 5.5(28)	UAC 14.6(75)	UGC 8.0(41)
UUA 15.2(78)	UCA 7.8(40)	UAA 1.8(9)	UGA 1.0(5)
UUG 11.9(61)	UCG 8.0(41)	UAG 0.0(0)	UGG 10.7(55)
CUU 11.9(61)	CCU 8.4(43)	CAU 15.8(81)	CGU 21.1(108)
CUC 10.5(54)	CCC 6.4(33)	CAC 13.1(67)	CGC 26.0(133)
CUA 5.3(27)	CCA 6.6(34)	CAA 12.1(62)	CGA 4.3(22)
CUG 46.9(240)	CCG 26.7(137)	CAG 27.7(142)	CGG 4.1(21)
AUU 30.5(156)	ACU 8.0(41)	AAU 21.9(112)	AGU 7.2(37)
AUC 18.2(93)	ACC 22.8(117)	AAC 24.4(125)	AGC 16.6(85)
AUA 3.7(19)	ACA 6.4(33)	AAA 33.2(170)	AGA 1.4(7)
AUG 24.8(127)	ACG 11.5(59)	AAG 12.1(62)	AGG 1.6(8)
GUU 16.8(86)	GCU 10.7(55)	GAU 37.9(194)	GGU 21.3(109)
GUC 11.7(60)	GCC 31.6(162)	GAC 20.5(105)	GGC 33.4(171)
GUA 11.5(59)	GCA 21.1(108)	GAA 43.7(224)	GGA 9.2(47)
GUG 26.4(135)	GCG 38.5(197)	GAG 18.4(94)	GGG 8.6(44)

Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%

Saccharomyces cerevisiae [gbpln]: 14411 CDS's (6534504 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU 26.1(170666)	UCU 23.5(153557)	UAU 18.8(122728)	UGU 8.1(52903)
UUC 18.4(120510)	UCC 14.2(92923)	UAC 14.8(96596)	UGC 4.8(31095)
UUA 26.2(170884)	UCA 18.7(122028)	UAA 1.1(6913)	UGA 0.7(4447)
UUG 27.2(177573)	UCG 8.6(55951)	UAG 0.5(3312)	UGG 10.4(67789)
CUU 12.3(80076)	CCU 13.5(88263)	CAU 13.6(89007)	CGU 6.4(41791)
CUC 5.4(35545)	CCC 6.8(44309)	CAC 7.8(50785)	CGC 2.6(16993)
CUA 13.4(87619)	CCA 18.3(119641)	CAA 27.3(178251)	CGA 3.0(19562)
CUG 10.5(68494)	CCG 5.3(34597)	CAG 12.1(79121)	CGG 1.7(11351)
AUU 30.1(196893)	ACU 20.3(132522)	AAU 35.7(233124)	AGU 14.2(92466)
AUC 17.2(112176)	ACC 12.7(83207)	AAC 24.8(162199)	AGC 9.8(63726)
AUA 17.8(116254)	ACA 17.8(116084)	AAA 41.9(273618)	AGA 21.3(139081)
AUG 20.9(136805)	ACG 8.0(52045)	AAG 30.8(201361)	AGG 9.2(60289)
GUU 22.1(144243)	GCU 21.2(138358)	GAU 37.6(245641)	GGU 23.9(156109)
GUC 11.8(76947)	GCC 12.6(82357)	GAC 20.2(132048)	GGC 9.8(63903)
GUA 11.8(76927)	GCA 16.2(105910)	GAA 45.6(297944)	GGA 10.9(71216)
GUG 10.8(70337)	GCG 6.2(40358)	GAG 19.2(125717)	GGG 6.0(39359)

Coding GC 39.77% 1st letter GC 44.58% 2nd letter GC 36.64% 3rd letter GC 38.10%

La rareté de certains codons R I L est aussi la principale différence entre usage des codons chez les prokaryotes vs eucaryotes même lorsque les 2 espèces vivent toutes les 2 à la même température physiologique, ex: *Eco* vs *Sc* et *Hsa*.

codons rares de *Eco* vs *Sc* et *Hsa*

La version la plus efficace de BL21 RIL s'appelle: **Rosetta 2** et surexprime en plus d'autres ARNt mineurs: Met, Thr, Pro, Gly...

Homo sapiens [gbpri]: 93487 CDS's (40662582 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU 17.6(714298)	UCU 15.2(618711)	UAU 12.2(495699)	UGU 10.6(430311)
UUC 20.3(824692)	UCC 17.7(718892)	UAC 15.3(622407)	UGC 12.6(513028)
UUA 7.7(311881)	UCA 12.2(496448)	UAA 1.0(40285)	UGA 1.6(63237)
UUG 12.9(525688)	UCG 4.4(179419)	UAG 0.8(32109)	UGG 13.2(535595)
CUU 13.2(536515)	CCU 17.5(713233)	CAU 10.9(441711)	CGU 4.5(184609)
CUC 19.6(796638)	CCC 19.8(804620)	CAC 15.1(613713)	CGC 10.4(423516)
CUA 7.2(290751)	CCA 16.9(688038)	CAA 12.3(501911)	CGA 6.2(250760)
CUG 39.6(1611801)	CCG 6.9(281570)	CAG 34.2(1391973)	CGG 11.4(464485)
AUU 16.0(650473)	ACU 13.1(533609)	AAU 17.0(689701)	AGU 12.1(493429)
AUC 20.8(846466)	ACC 18.9(768147)	AAC 19.1(776603)	AGC 19.5(791383)
AUA 7.5(304565)	ACA 15.1(614523)	AAA 24.4(993621)	AGA 12.2(494682)
AUG 22.0(896005)	ACG 6.1(246105)	AAG 31.9(1295568)	AGG 12.0(486463)
GUU 11.0(448607)	GCU 18.4(750096)	GAU 21.8(885429)	GGU 10.8(437126)
GUC 14.5(588138)	GCC 27.7(1127679)	GAC 25.1(1020595)	GGC 22.2(903565)
GUA 7.1(287712)	GCA 15.8(643471)	GAA 29.0(1177632)	GGA 16.5(669873)
GUG 28.1(1143534)	GCG 7.4(299495)	GAG 39.6(1609975)	GGG 16.5(669768)

Coding GC 52.27% 1st letter GC 55.72% 2nd letter GC 42.54% 3rd letter GC 58.55%

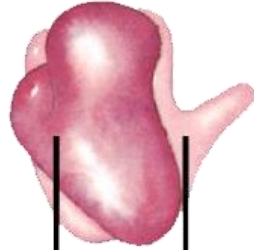
La Traduction

Nous n'allons voir dans le détail que la traduction **chez les procaryotes**. Vous avez à votre disposition les diapos qui suivent mais également **un film** qui vous montre dans le détail ce processus dynamique de la synthèse protéique.

Regarder le film: translation_movie.wmv

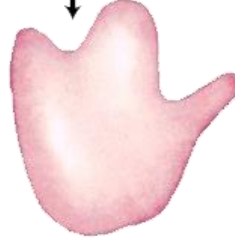
I. Le ribosome

Ribosome bactérien 70S
 $M_r = 2,7 \times 10^6$



$M_r = 1,8 \times 10^6$
ARNr 5S (120 Ntides)
ARNr 23S (3200 Ntides)
36 protéines

50S

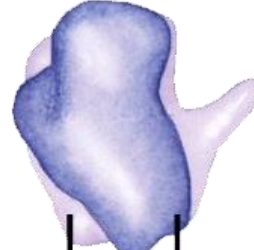


30S



$M_r = 0,9 \times 10^6$
ARNr 16S (1540 Ntides)
21 protéines

Ribosome eucaryotique 80S
 $M_r = 4,2 \times 10^6$



$M_r = 2,8 \times 10^6$
ARNr 5S (120 Ntides)
ARNr 28S (4700 Ntides)
ARNr 5,8S (160 Ntides)
49 protéines

60S



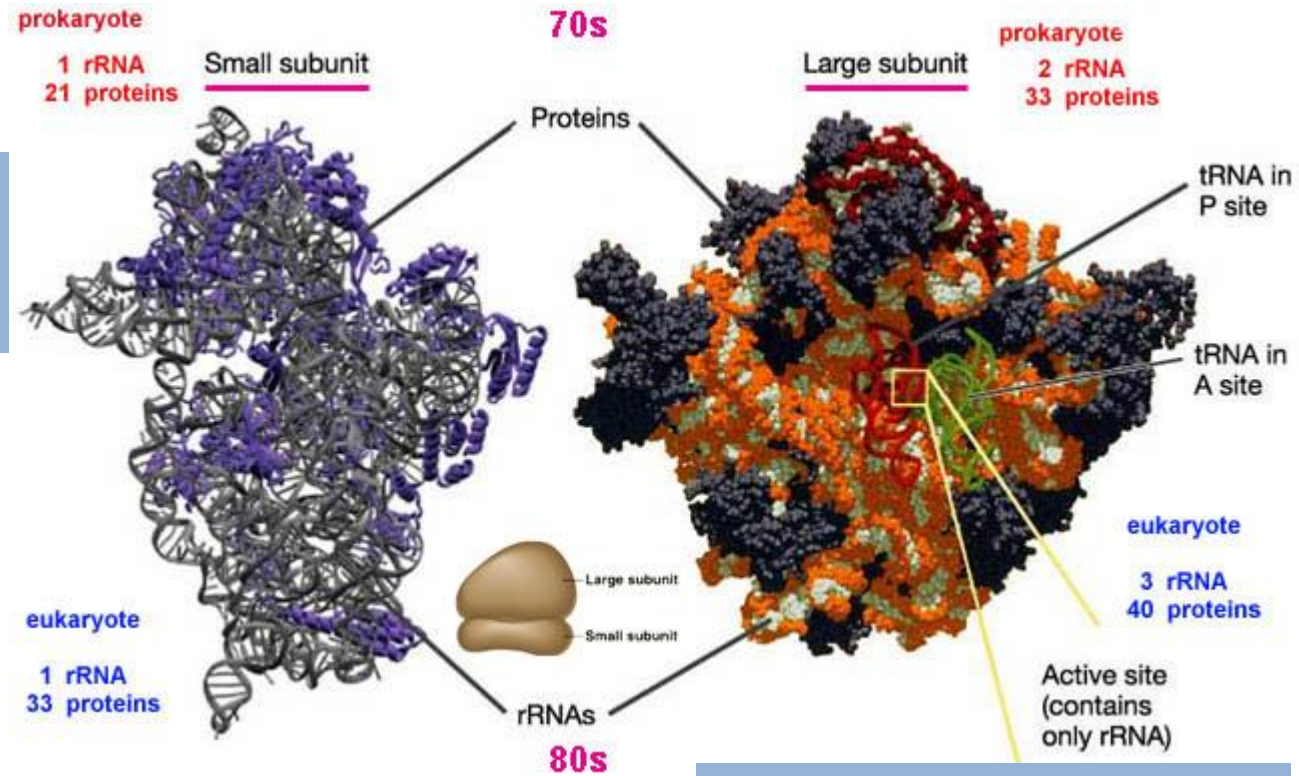
40S



$M_r = 1,4 \times 10^6$
ARNr 18S (1900 Ntides)
33 protéines

I. Le ribosome

Dans la petite su ou su 30S l'ARN 16S est gris et les protéines de la petite su S1 à S21 sont en violet foncé



Dans la grande su ou su 50S les ARN 23S et 5S sont en orange et les protéines de la grande su L1 à L33 sont en gris.

Le ribosome assemblé (70S) est principalement constitué d'ARN et est un **ribozyme** cela veut dire que la réaction qu'il catalyse est ARN-dépendante: c'est l'ARN 23S qui compose le site catalytique de formation de la liaison peptidique.

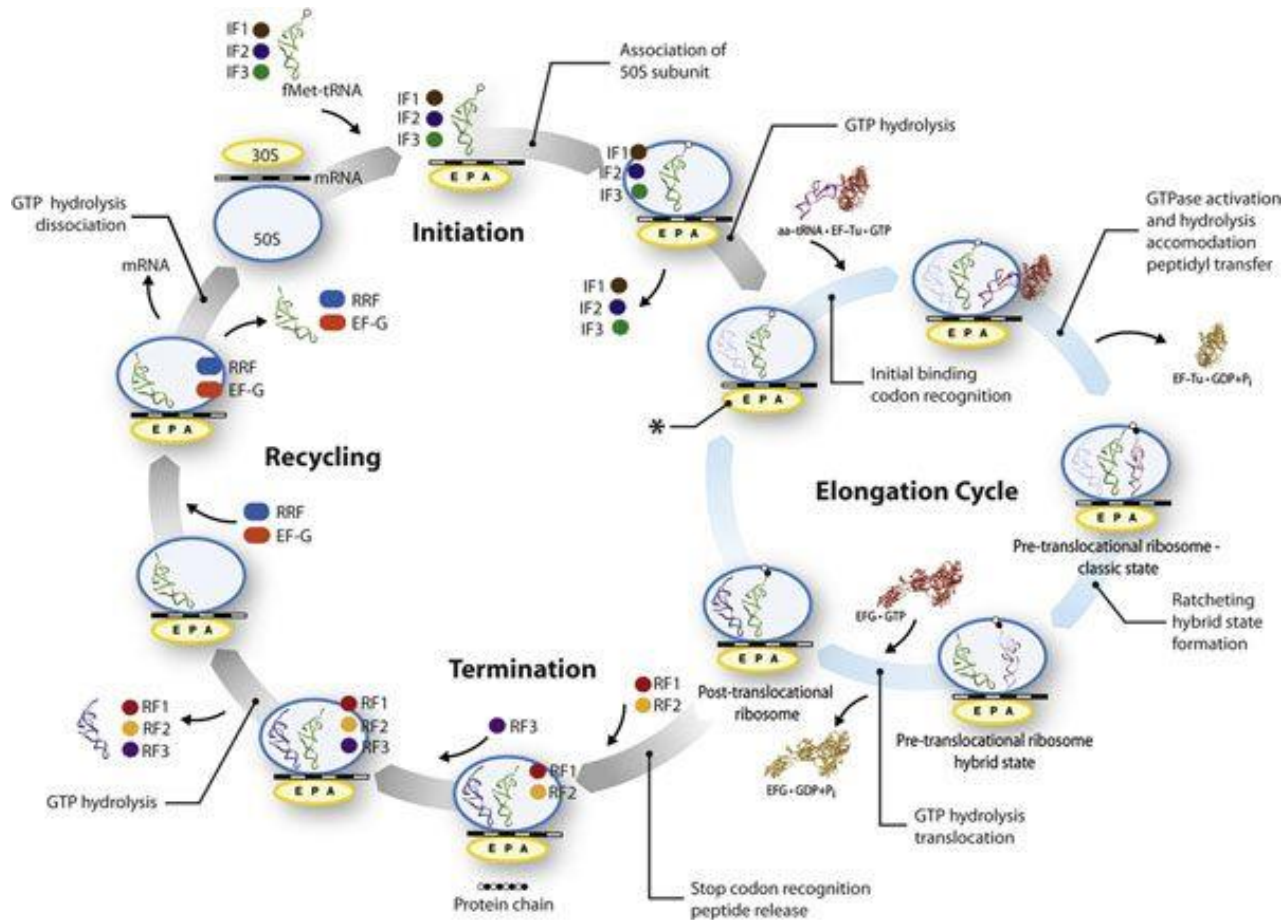
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | [Cite this article](#)



Regarder le film: [translation_movie.wmv](#)

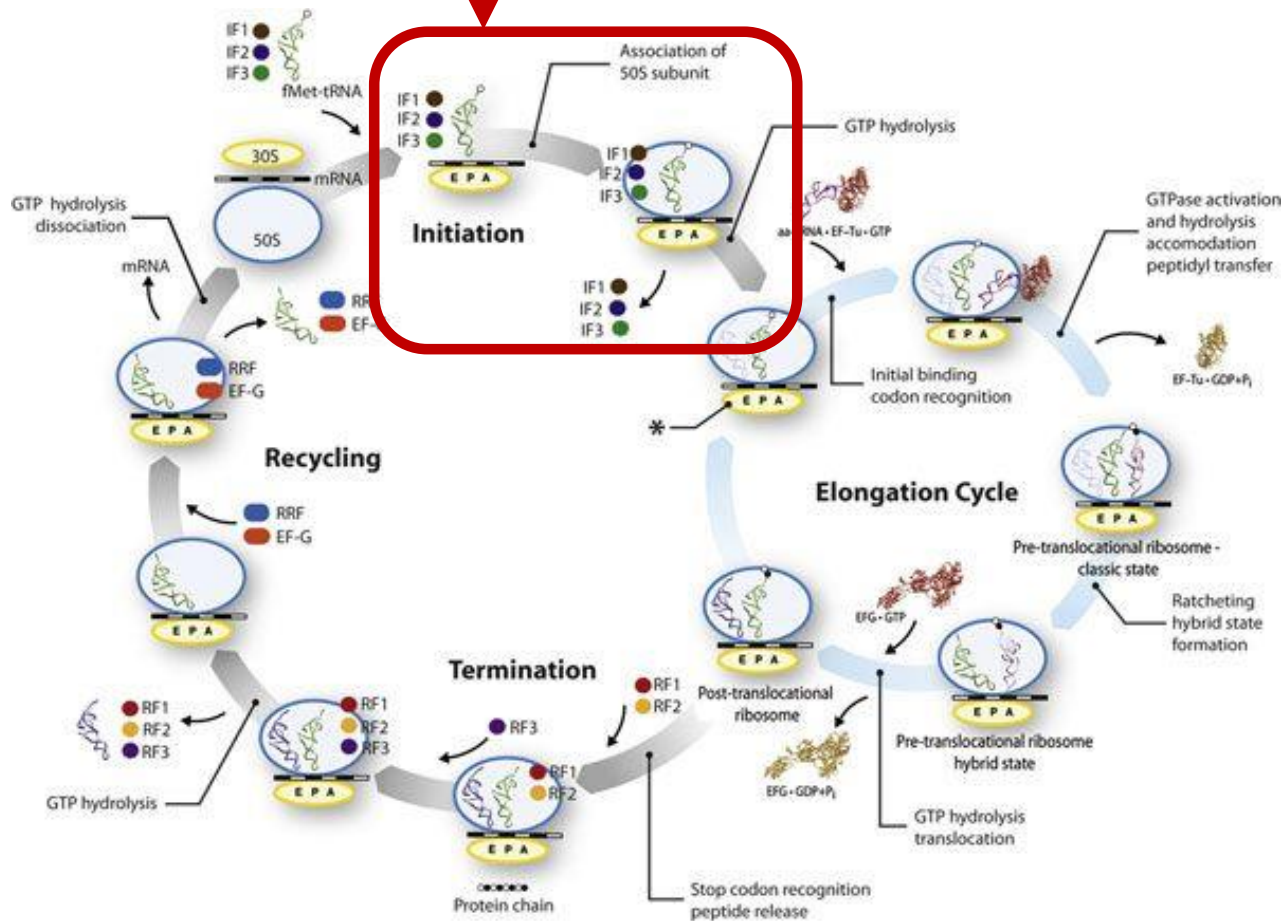
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | Cite this article



3 Phases: **INITIATION** – ELONGATION (cyclique) - TERMINAISON

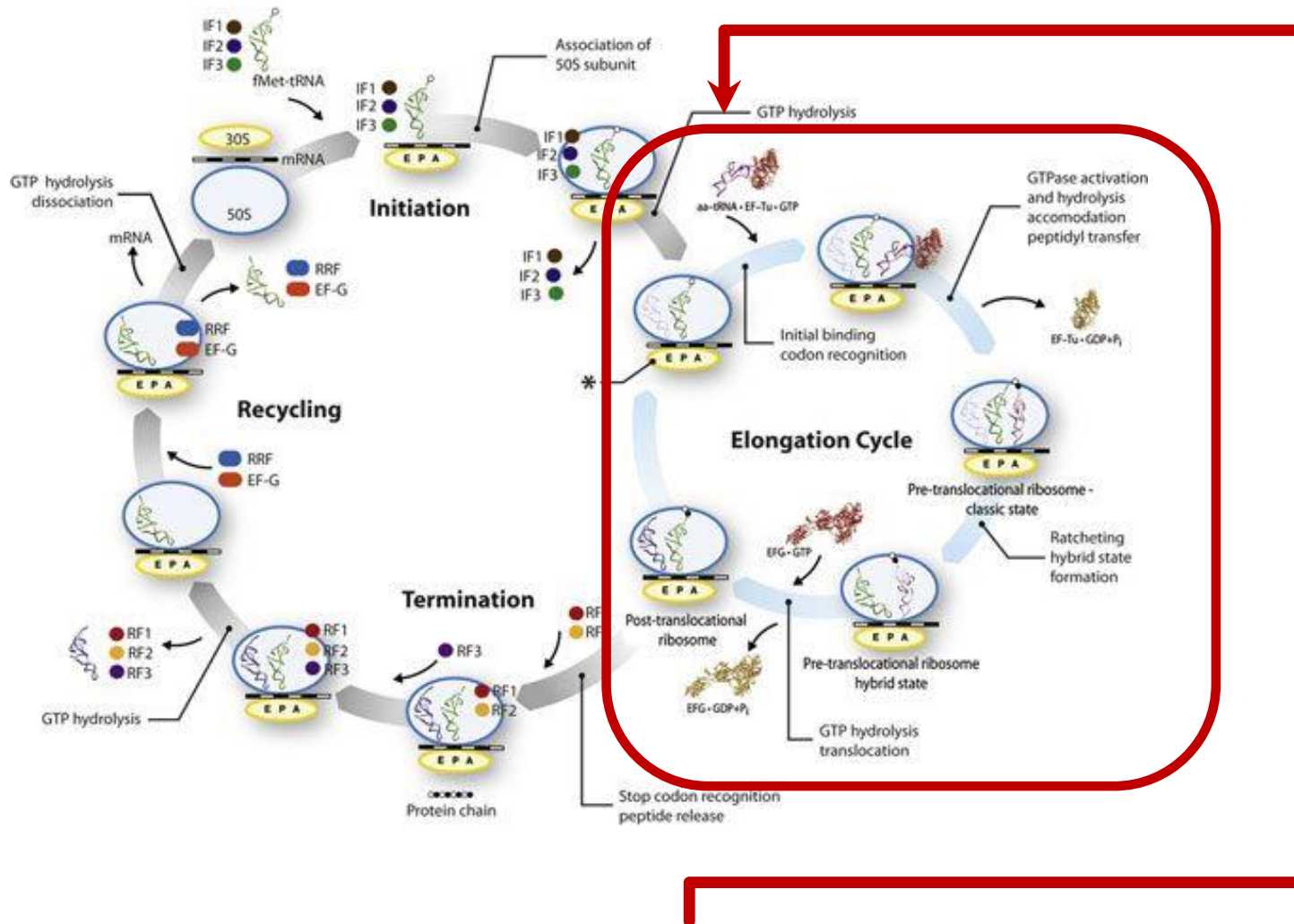
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | [Cite this article](#)



3 Phases: INITIATION – ELONGATION (cyclique) - TERMINAISON

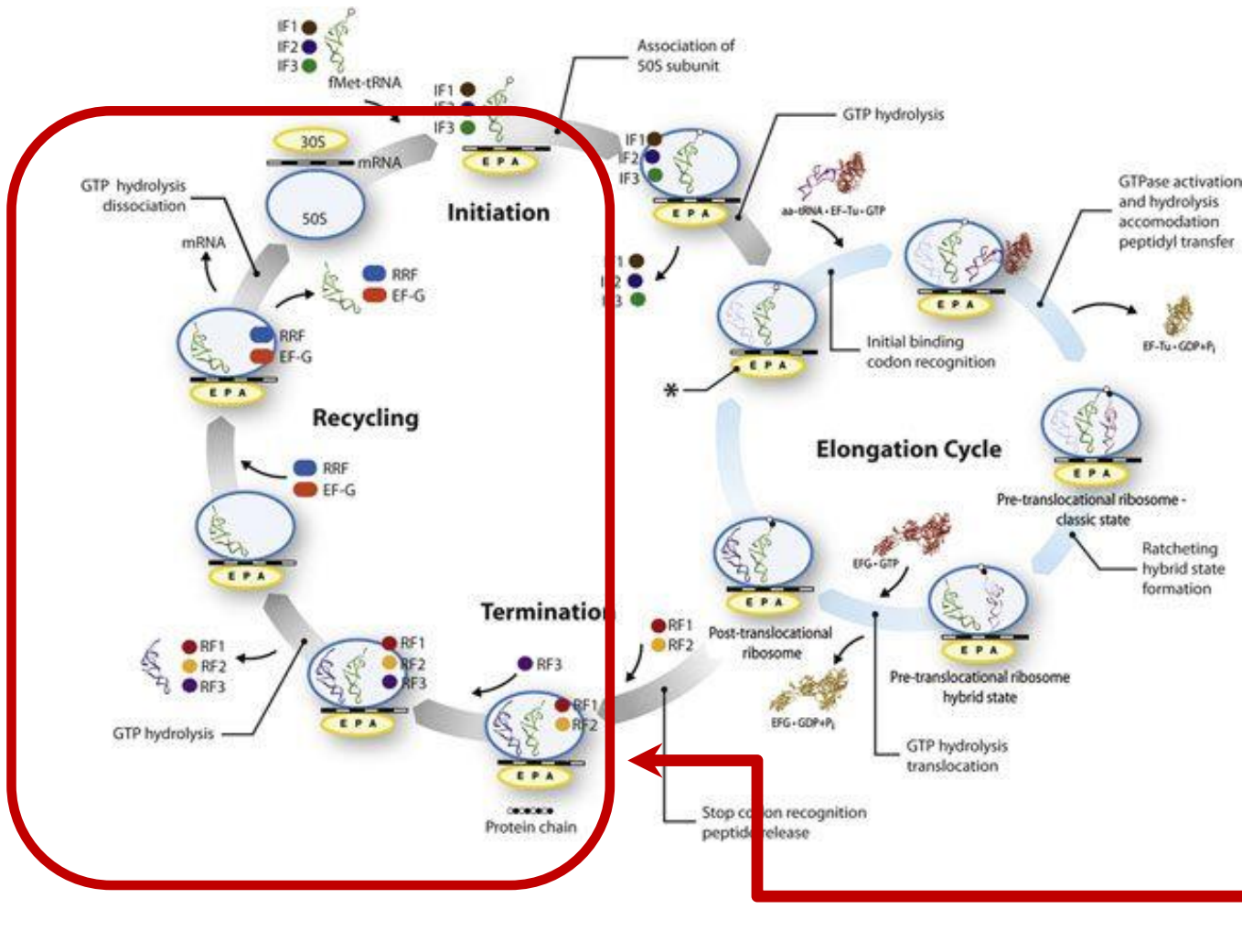
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | Cite this article



3 Phases: INITIATION – ELONGATION (cyclique) - **TERMINAISON**

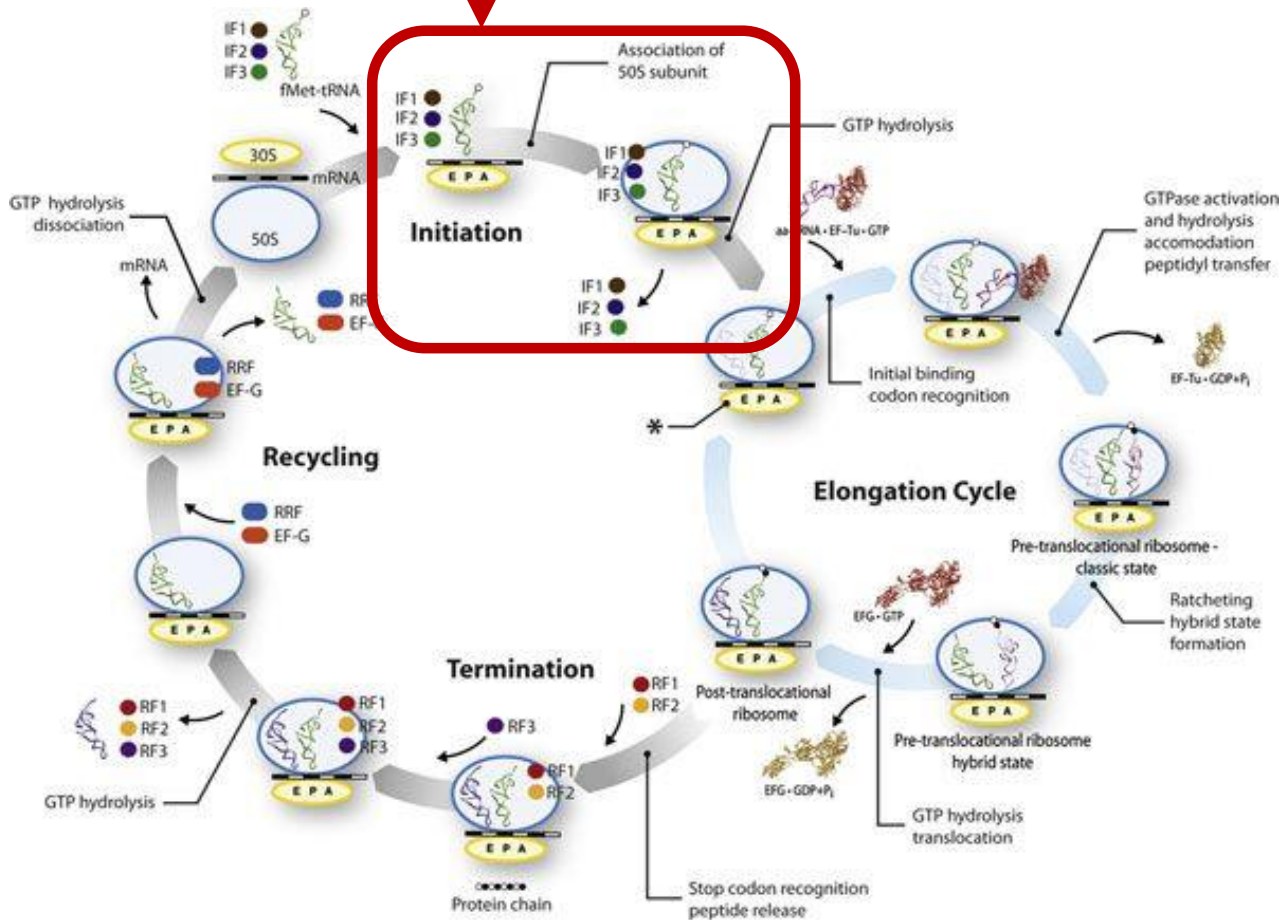
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | Cite this article



L'initiation Doit amener à la formation dun complexe ternaire $su30S \cdot tRNA_i^{Met} \cdot mRNA$

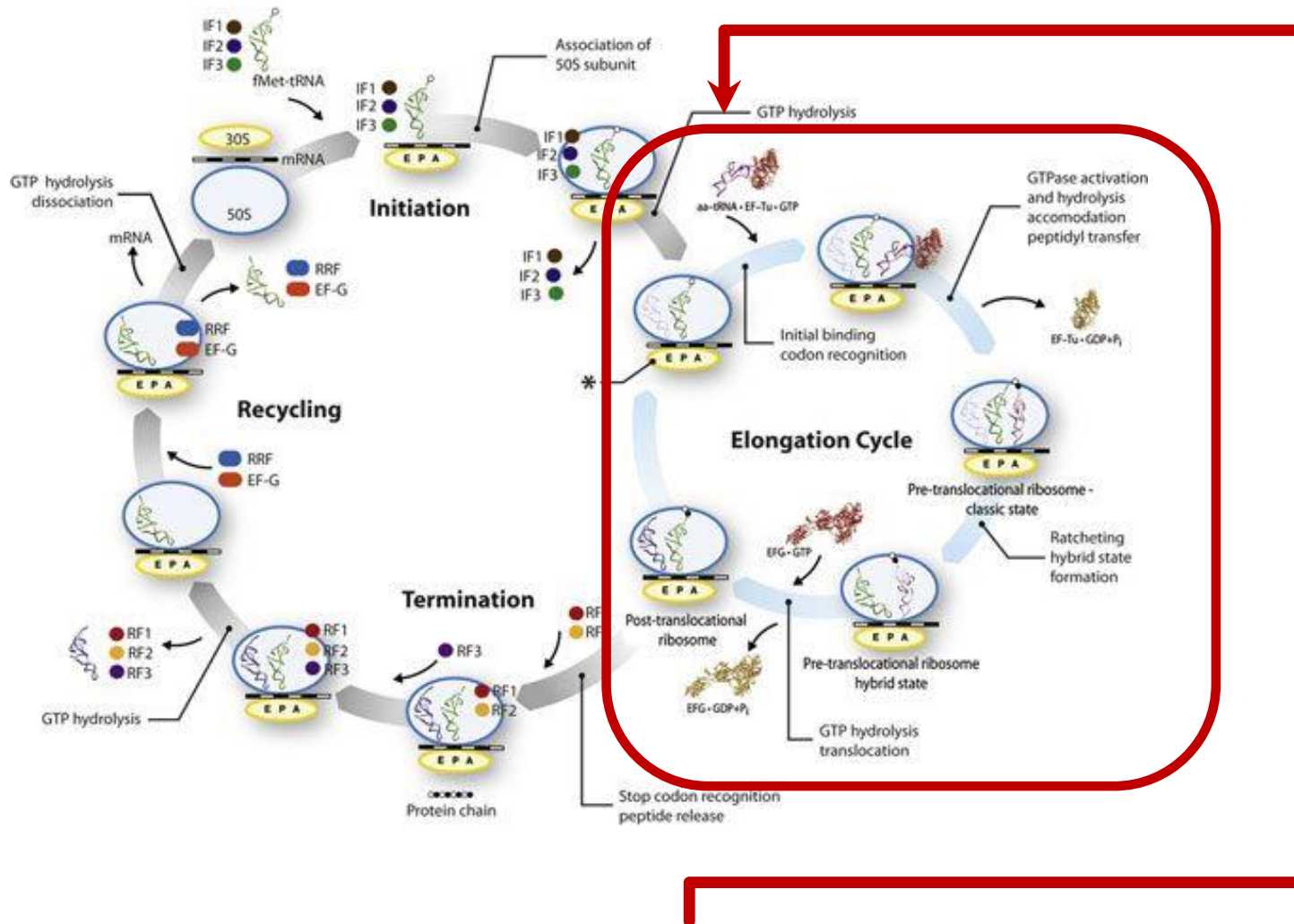
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | Cite this article



L'élargissement, c'est une répétition cyclique de plusieurs étapes: 1) Fixation de l'aa-tRNA 2) Transpeptidation c'est la formation de la liaison peptidique entre 2 aa codés par 2 codons consécutifs 3) Translocation permet au ribosome d'avancer sur le prochain codon ou autre alternative de faire reculer l'ARNm dans le ribosome

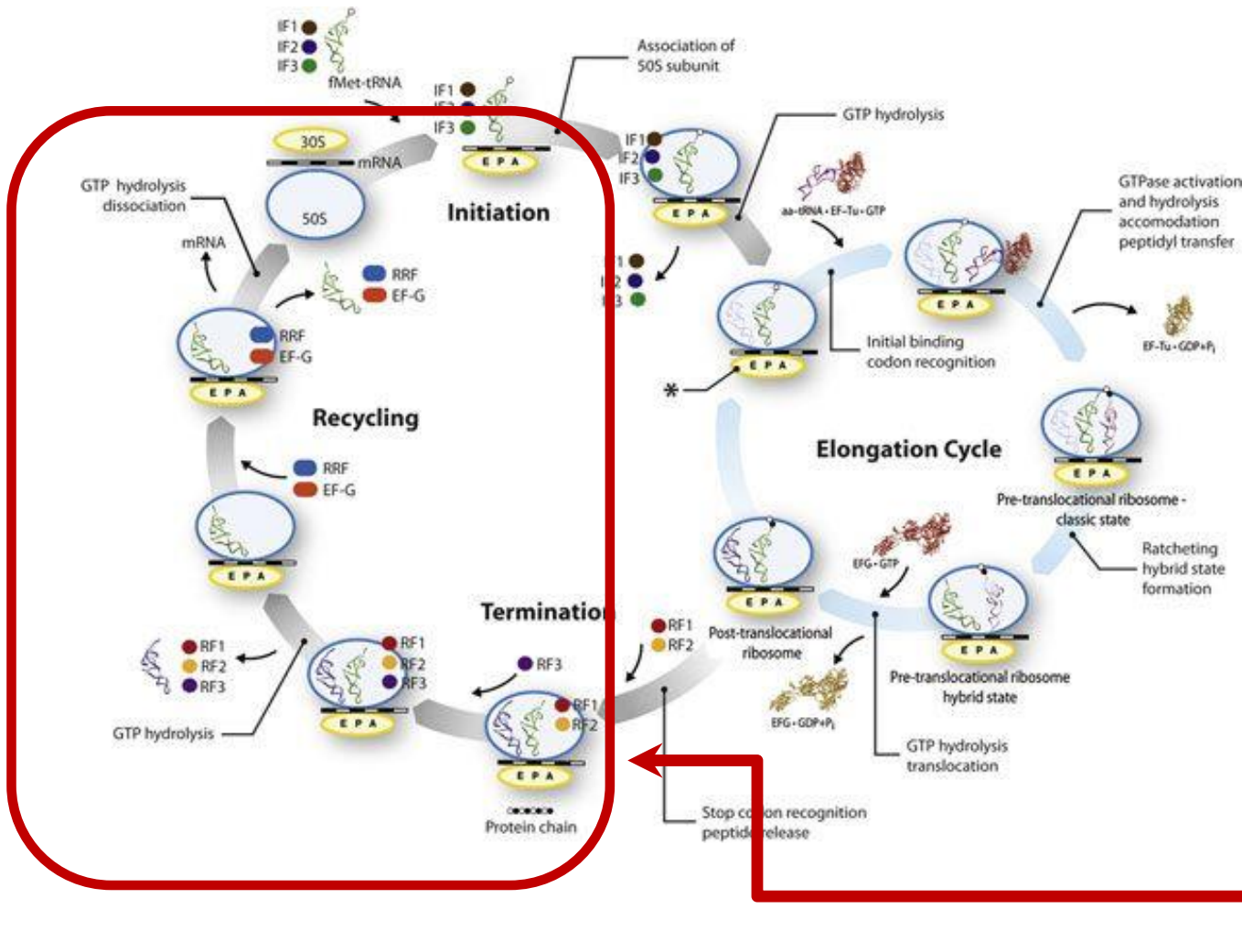
II. Vue d'ensemble de la traduction

Review | Published: 13 April 2010

From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery

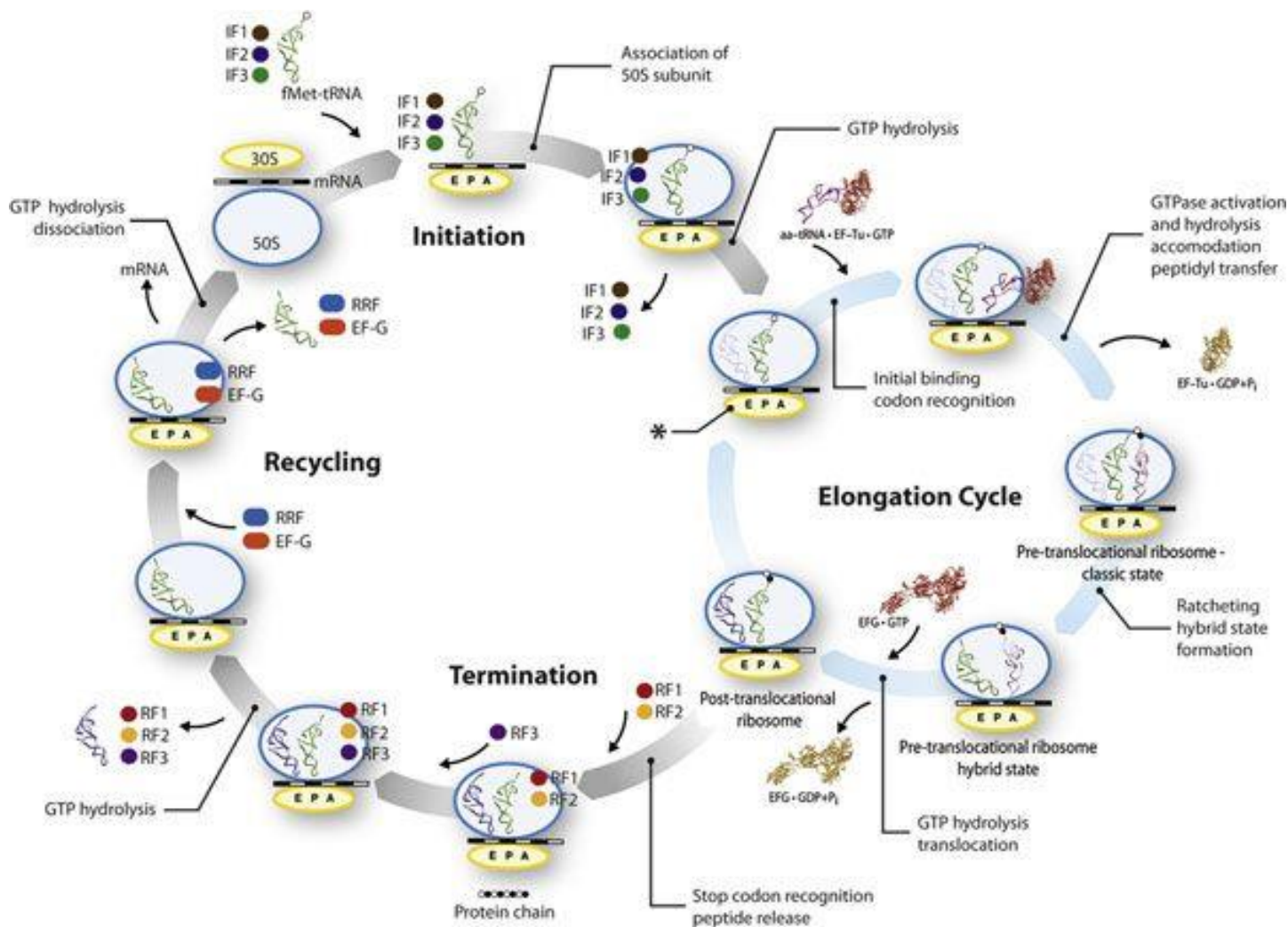
Xabier Agirrezabala & Joachim Frank

Human Genomics 4, Article number: 226 (2010) | Cite this article



La **terminaison**, c'est quand le ribosome rencontre un codon STOP et consiste en la libération de la protéine synthétisée par traduction de l'ARNm entier. Elle est suivie du recyclage du ribosome

II. Vue d'ensemble de la traduction



Le ribosome n' est pas suffisant à lui seul pour assurer cette incorporation cyclique des aa, il lui faut des Prot accessoires appelées **facteurs de traduction** qui sont de 3 types:

Facteurs d' initiation IF
(Initiation factors)

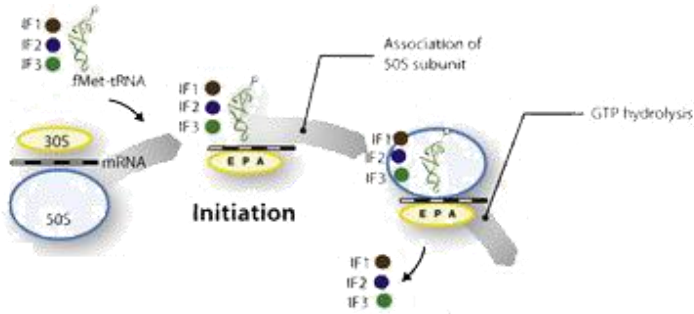
Facteurs d'élongation EF
(Elongation factors)

Facteurs de terminaison RF
(Release factors)

Le ribosome contient 3 sites de fixation de l' aa-ARNt de la droite vers la gauche : Site **A** , **P** et **E**.

A pour **Aminoacyl-ARNt** donc ce site ne contiendra que des aa-ARNt; **P** pour **Peptidyl-ARNt** donc ce site contiendra le tRNA porteur du peptide en cours d'élongation et **E** pour **Exit** c'est le site de sortie du tRNA qui n'est plus porteur d'un aa. Comme le ribosome est constitué de 2 su, il y a un ½ site A sur la 30S et un ½ site A sur la 50S, il y a un ½ site P sur la 30S et un ½ site P sur la 50S et le site E n'existe que sur la 50S

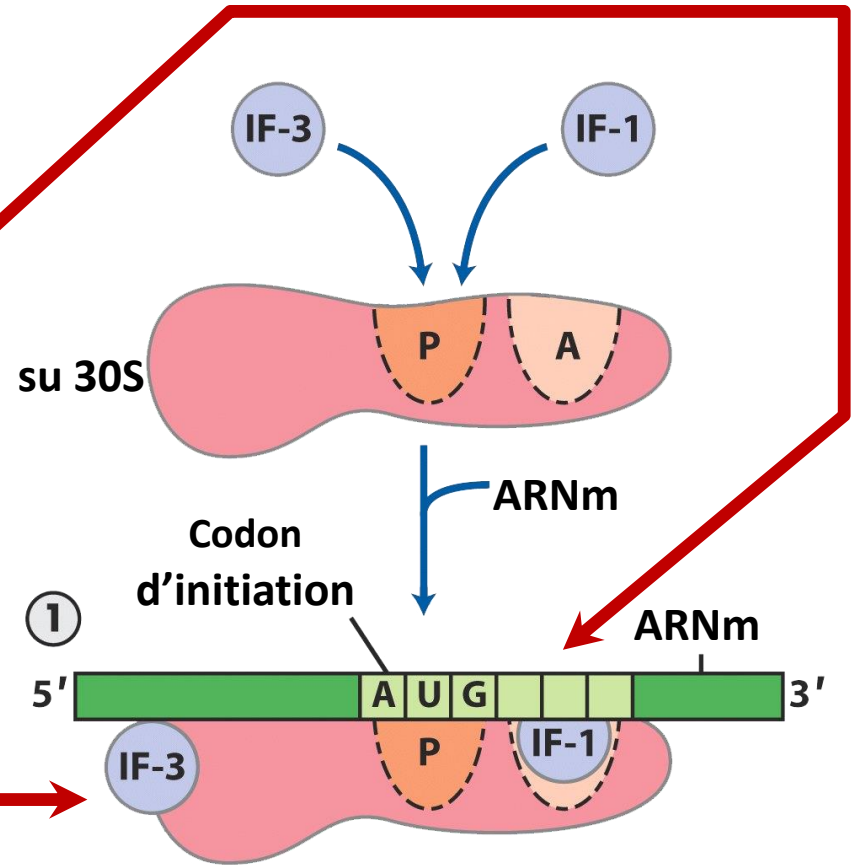
III. L'initiation



Chez les bactéries il y a 3 facteurs d'initiation: **IF1 = 8 kDa** empêche l'accès du site A du ribosome pour le fMet-tRNA_i^{Met} en se logeant dans le 1/2 site A de la 30S.

IF2 = 97,2 kDa qui sous sa forme liée au GTP a pour fonction d'apporter le fMet-tRNA_i^{Met} au niveau du 1/2 site P de la su 30S.

IF3 = 20,7 kDa qui est un anti-associant empêchant la su 50S de se fixer sur la 30S tant qu'il est fixé à la su 30S il est également impliqué dans le bon positionnement de l'ARNm sur la 30S.

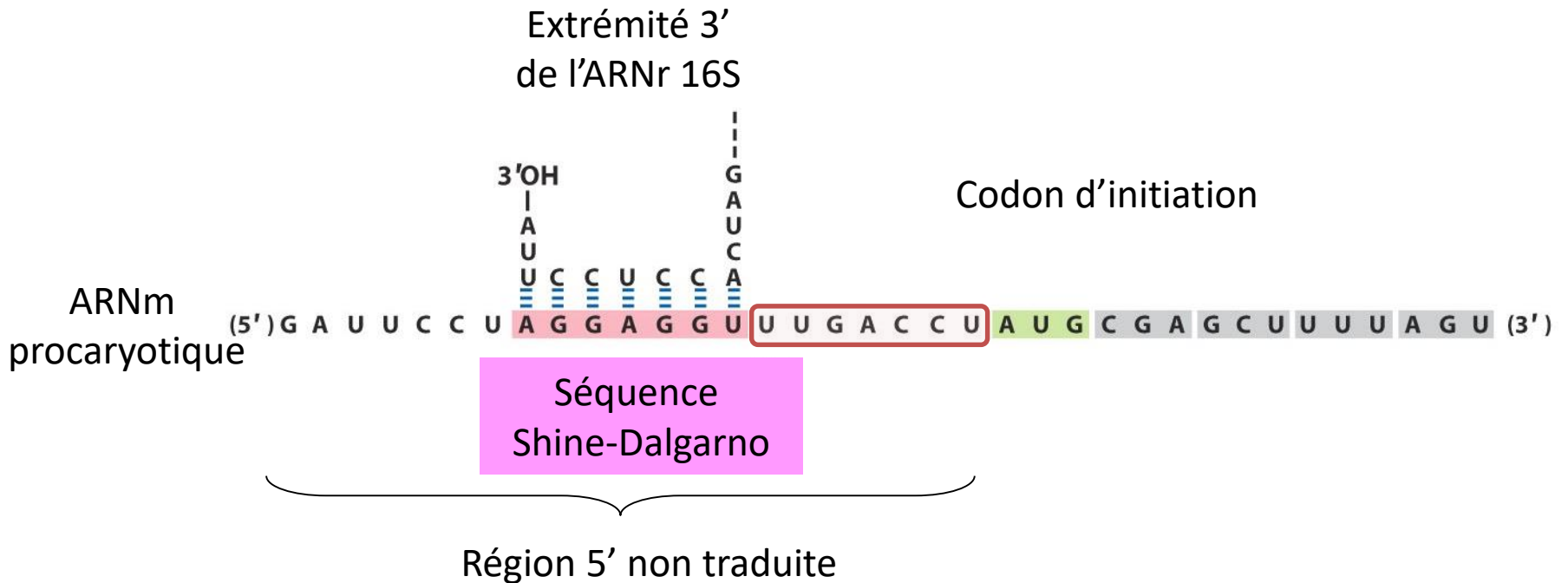


Les 3 IF ont **beaucoup d'affinité** pour la **30S libre** et donc ils s'y fixent.

Vous remarquerez que le **codon d'initiation de l'ARNm est placé dans le 1/2 site P de la 30S** ce qui veut dire que le 1^{er} aminoacyl-ARNt, le fMet-tRNA_i^{Met}, ne se fixera pas dans le site A (pour aminoacyl-ARNt) mais dans le site P et c'est logique comme cela au cours de l'élongation le site A sera libre pour y déposer le 2^{ème} aa-ARNt et faire la liaison peptidique.

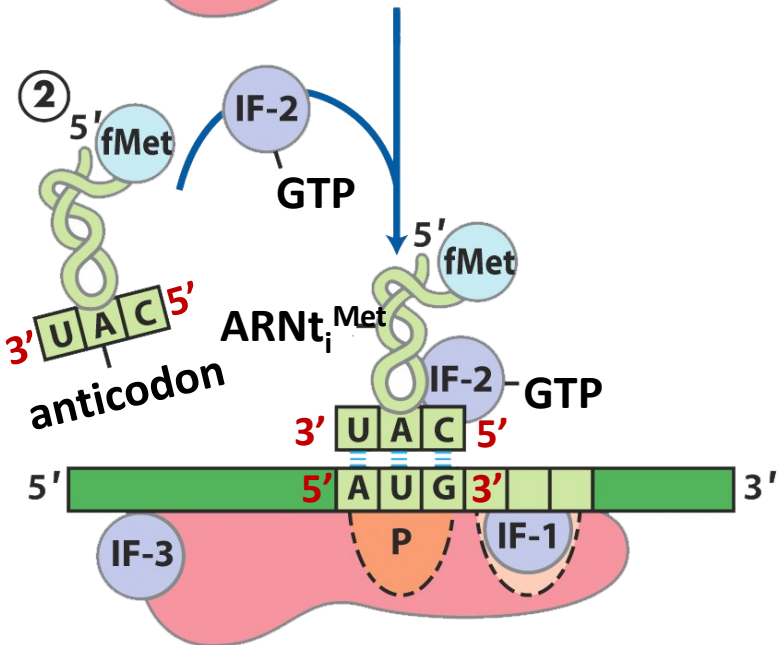
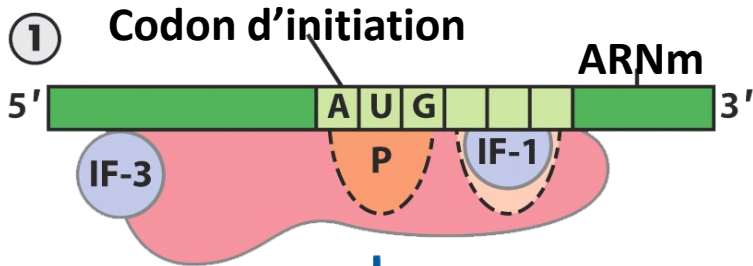
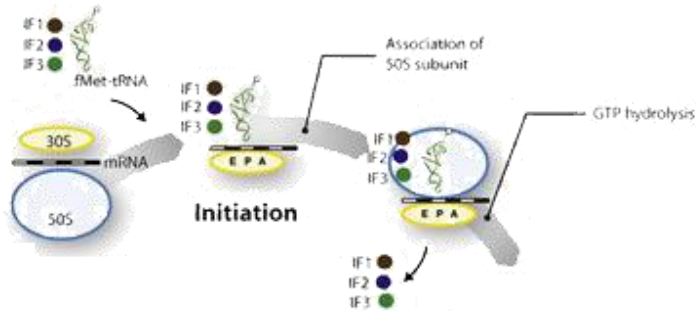
III. L'initiation

Tous les ARNm de prokaryotes ont en 5' une séquence riche en purine, la séquence **Shine Dalgarno (SD)**, dont la séquence consensus est **5'AGGAGGU3'** et l'ARNr 16S contient à son extrémité 3' une séquence riche en pyrimidine complémentaire de la séquence SD = séquence **anti-SD**. L'interaction SD•@SD permet l'ancrage de l'ARNm sur la su 30S du ribosome.



Comme la distance entre la **Shine Dalgarno** et l'**AUG initiateur** est toujours la même, du coup l'AUG est automatiquement dans le ½ site P de la 30S

III. L'initiation



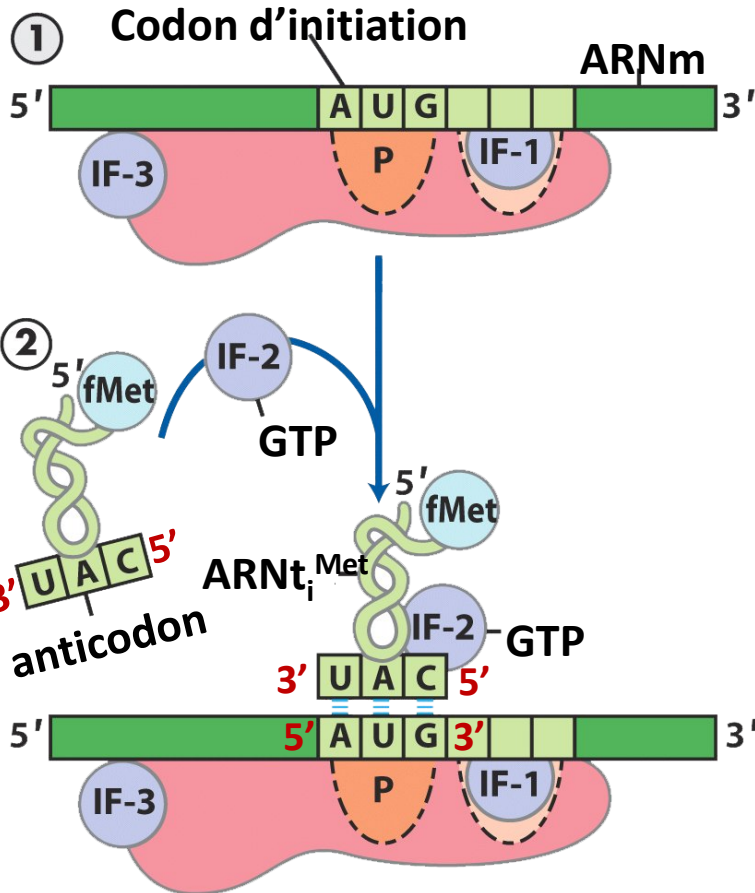
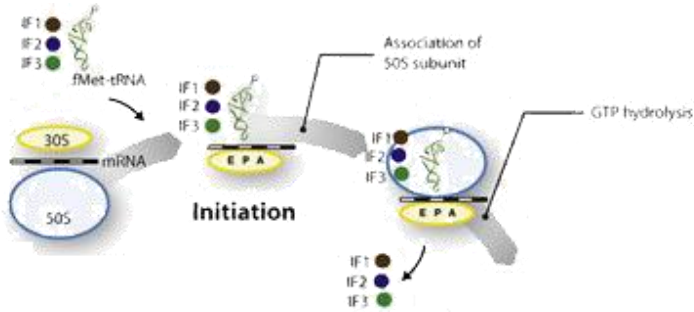
IF2 = 97,2 kDa qui sous sa forme liée au GTP a pour fonction d'apporter le fMet-tRNA_i^{Met} au niveau du ½ site P de la su 30S.

IF2 est une protéine G qui peut lier soit le GTP soit le GDP et n'a d'affinité pour le fMet-tRNA_i^{Met} que quand il est lié au GTP (IF2•GTP). IF2•GTP ne peut lier que le le fMet-tRNA_i^{Met} et aucun autre aa-tRNA (IF2•GTP•fMet-tRNA_i^{Met} = complexe ternaire)
IF2•GDP n'a plus aucune affinité pour le le fMet-tRNA_i^{Met}.

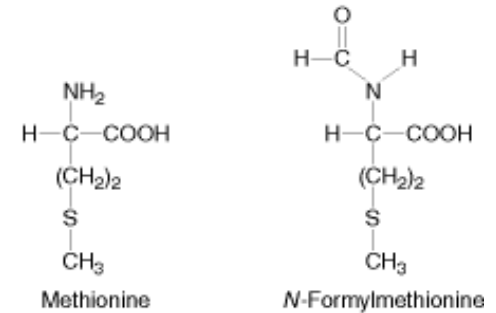
NB: IF2-GTP ne lie que le fMet-tRNA_i^{Met} et pas le tRNA_i^{Met} non aminoacylé

IF2•GTP•fMet-tRNA_i^{Met} amène donc le le fMet-tRNA_i^{Met} au niveau du ½ site P pour que celui-ci interagisse grâce à son @codon 5'CAU3' avec l'AUG dans le ½ site P de la 30S.

III. L'initiation

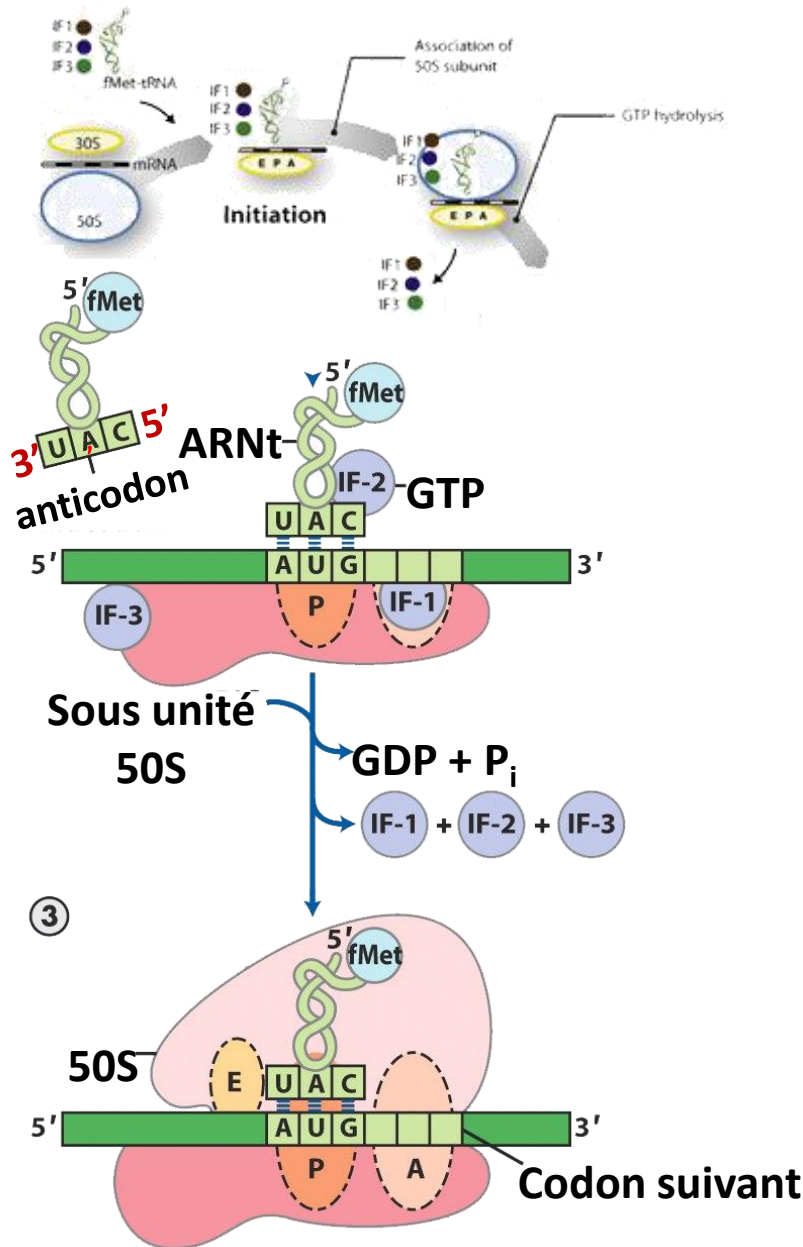


Chez les prokaryotes et uniquement chez eux, le tRNA^{Met} qui se lie sur l'AUG initiateur n'est pas aminoacylé par une Met mais par une **formylméthionine** d'où sont nom **fMet-tRNA_i^{Met}**.



Le **i** de tRNA_i^{Met} veut dire **initiateur** et il est différent du tRNA^{eMet} qui est le tRNA^{Met} **élongateur** qui décodera des codons Met au cours de l'élongation.

III. L'initiation



Lorsqu'il y a **interaction codon-@codon**, **IF2●GTP●fMet-tRNA_i^{Met}** hydrolyse son GTP et du coup devient **IF2●GDP**, car change de conformation et perd toute affinité pour le fMet-tRNA_i^{Met} et quitte le ½ site P de la su 30S.

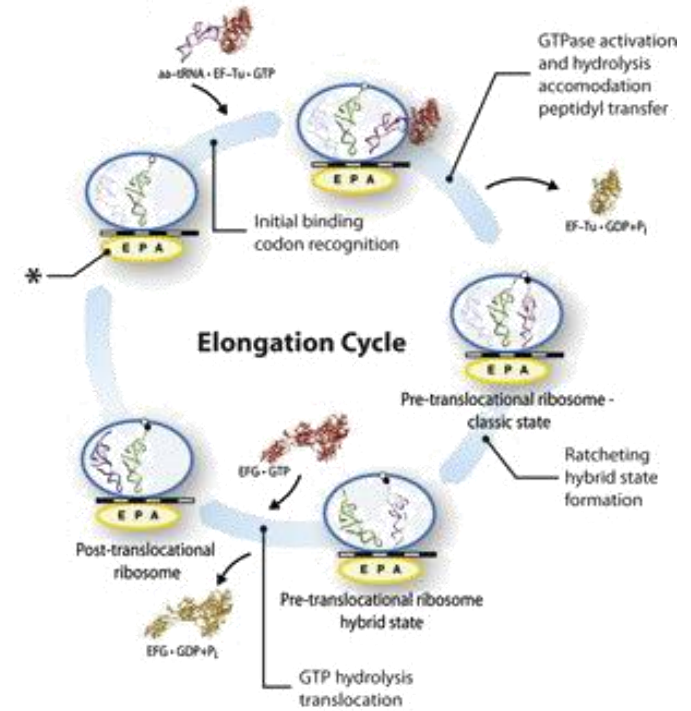
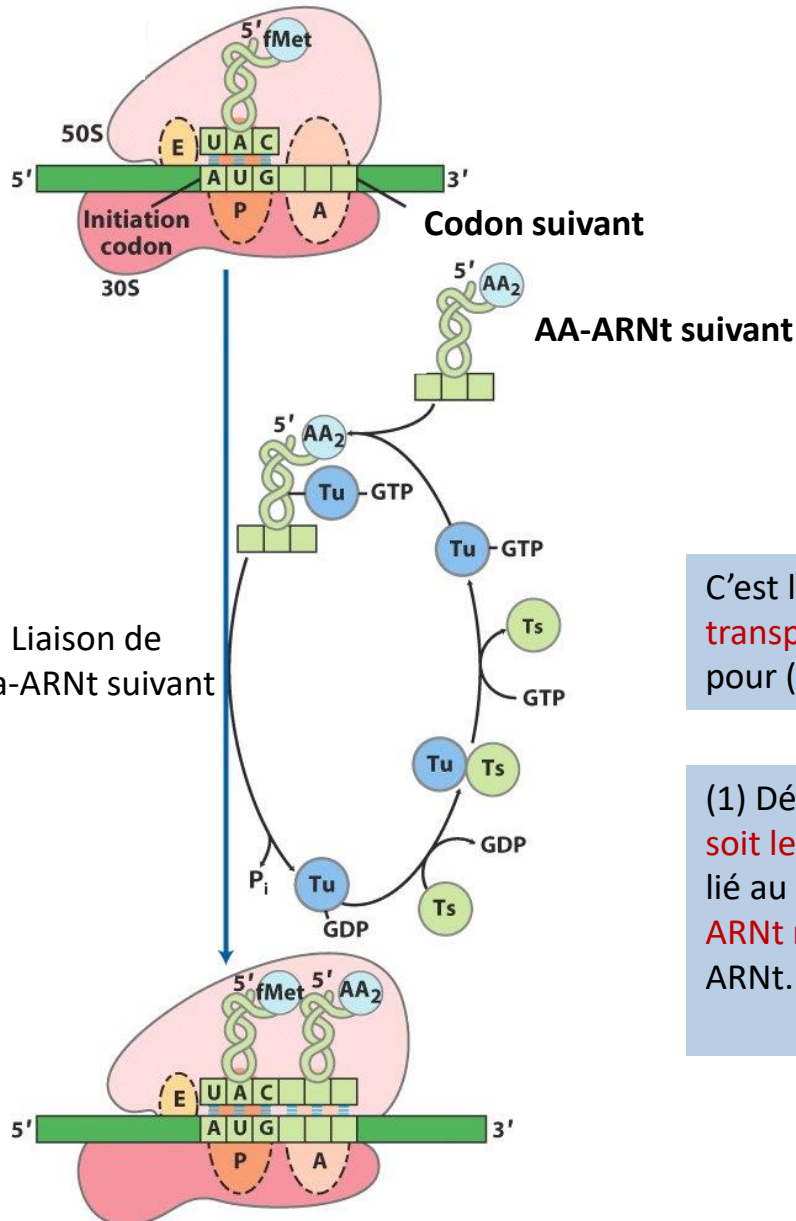
Le changement de conformation d'**IF2●GDP** se propage à la su 30S qui change de conformation et du coup **IF1 et IF3 perdent toute affinité pour la 30S et la quittent.**

Comme IF3 n'est plus là l'anti-associant n'est plus là et donc la su **50S se fixe sur la 30S**. Du coup on a 1 site A entier, 1 site P entier et la 50S ramène le site E

Comme IF1 est parti, le site A est libre et ça tombe bien on va y déposer l'aa-ARNt correspondant au codon 2

IV. L'élongation

(1) Dépôt de l'aa-ARNt

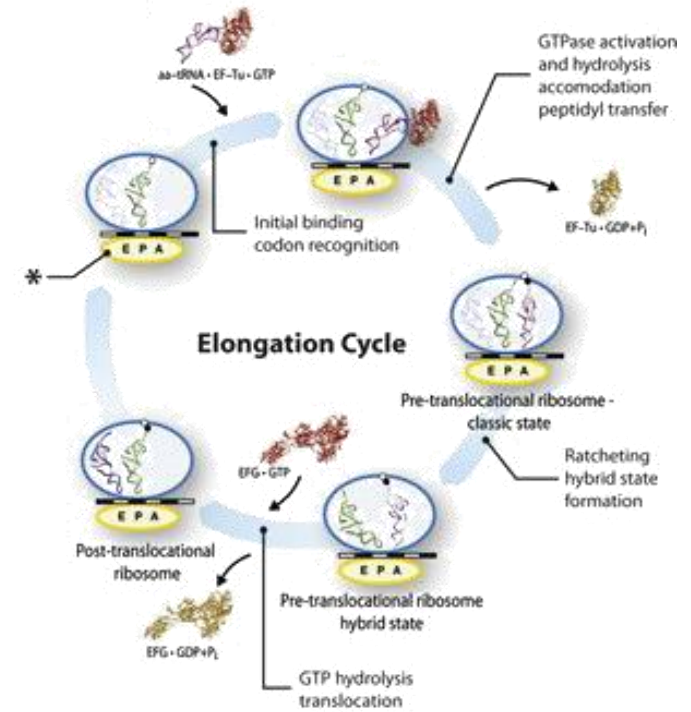
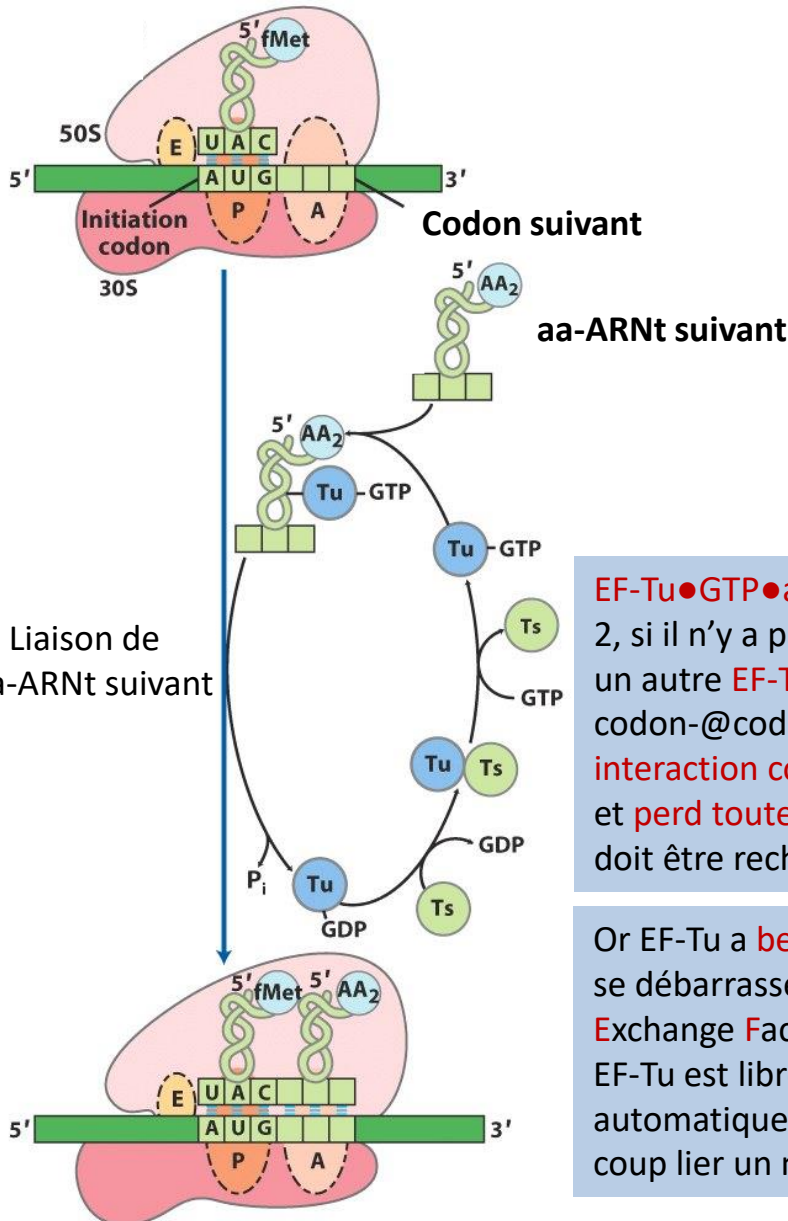


C'est la répétition cyclique de: (1) **dépôt** d'un aa-ARNt dans le site A, (2) **transpeptidation**, (3) **translocation**. Il y a 2 facteurs d'élongation (EF) EF-Tu pour (1) et EF-G pour (3) il n'y a pas d'EF pour (2).

(1) Dépôt: effectué par **EF-Tu•GTP**. EF-Tu est une protéine G qui peut lier **soit le GTP** soit le GDP et n'a d'affinité pour les aa-ARNt que quand il est lié au GTP (**EF-Tu•GTP**). **EF-Tu•GTP ne peut lier que les aa-tRNA et pas les ARNt non aminoacylés**. **EF-Tu•GDP n'a plus aucune affinité** pour les aa-ARNt.

IV. L'élongation

(1) Dépôt de l'aa-ARNt

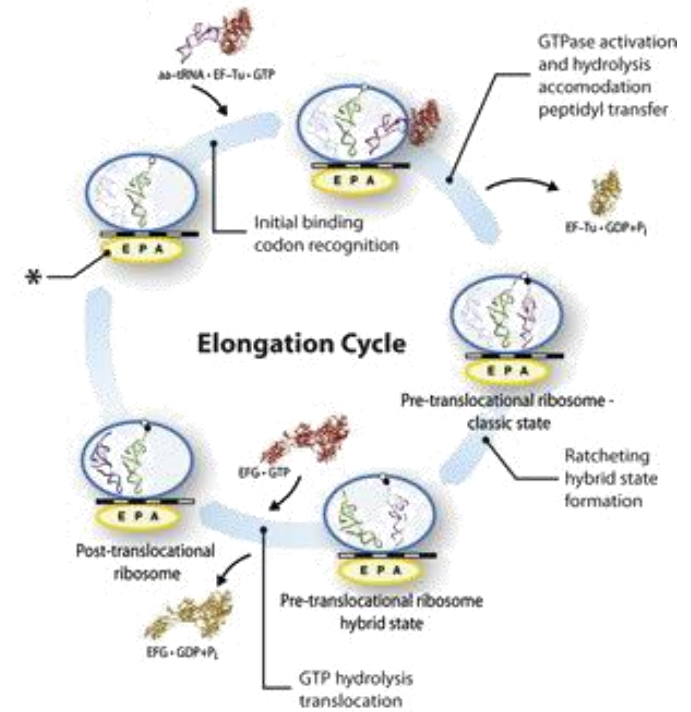
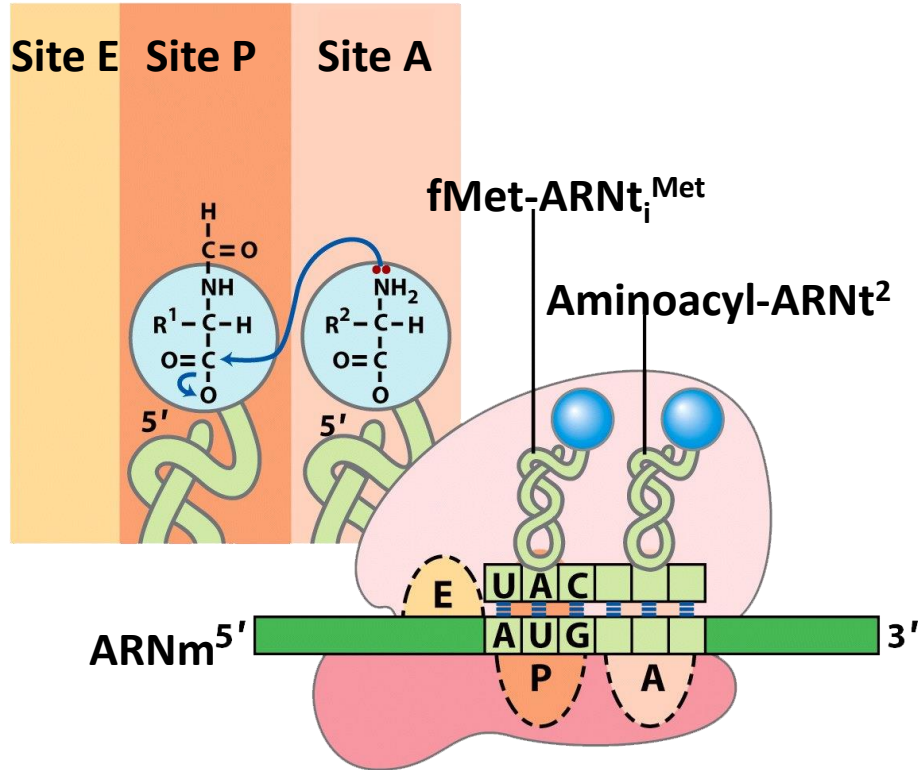


EF-Tu•GTP•aa-ARNt amène un aa-ARNt au niveau du site A qui contient le codon 2, si il n'y a pas interaction codon-@codon EF-Tu•GTP•aa-ARNt quitte le site A et un autre EF-Tu•GTP•aa-ARNt prend sa place et essaye de faire une interaction codon-@codon. Lorsque EF-Tu•GTP•aa₂-ARNt^{AA2} se fixe dans le site A, il y a interaction codon-@codon et du coup EF-Tu•GTP•aa₂-ARNt^{AA2} hydrolyse son GTP et perd toute affinité pour l'aa₂-ARNt^{AA2}. EF-Tu•GDP ne peut plus lier d'aa-ARNt il doit être recharger en GTP.

Or EF-Tu a beaucoup plus d'affinité pour le GDP que le GTP, du coup il ne peut pas se débarrasser du GDP tout seul il a besoin de l'aide de son GEF = Guanine Exchange Factor qui est EF-Ts. EF-Ts se lie à EF-Tu•GDP et enlève le GDP, du coup EF-Tu est libre de tout nucléotide guanylé (à base de guanine) et se recharge automatiquement and GTP car dans la cellule [GTP]≫[GDP]. EF-Tu•GTP peut du coup lier un nouvel aa-ARNt pour l'amener au ribosome....

IV. L'élongation

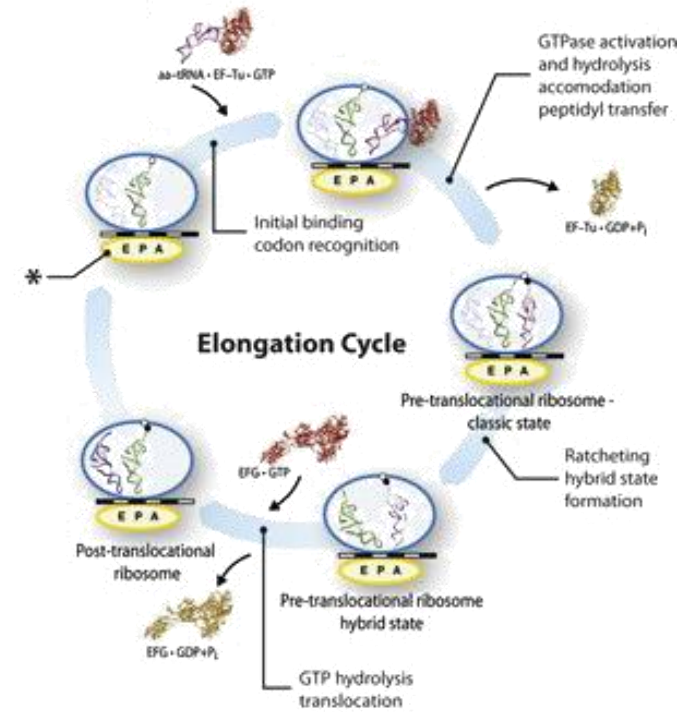
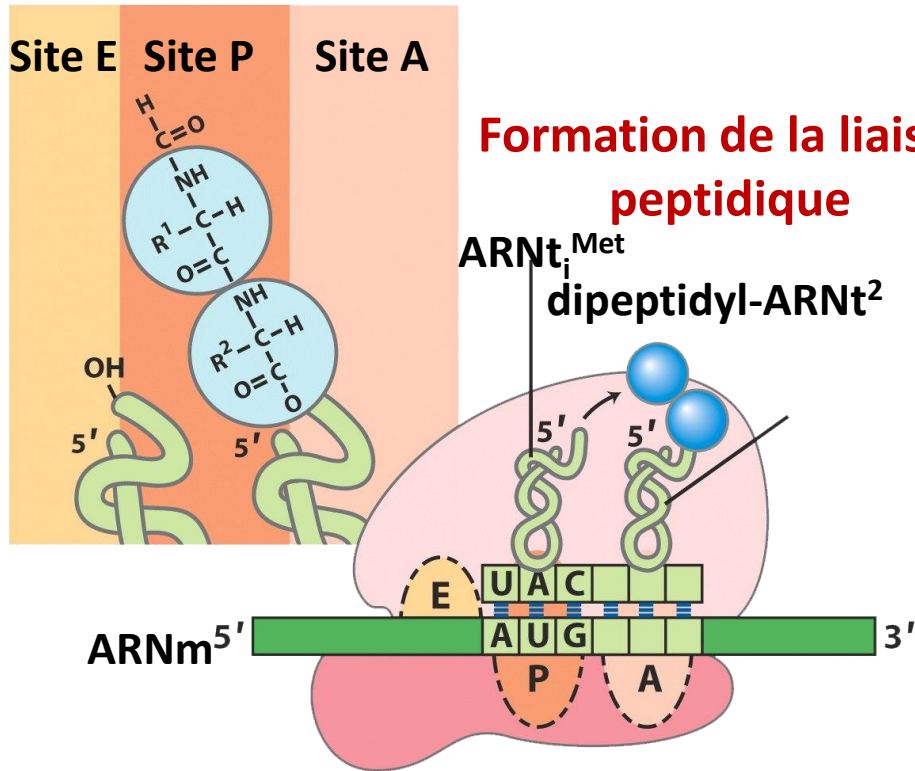
(2) Transpeptidation



L'ARN 23S de la su 50S rend le NH_2 de l'aa porté par l'ARNt qui est dans le site A nucléophile et donc capable d'attaquer la liaison ester entre la fMet et le tRNA_i^{Met} qui est dans le site P.

IV. L'élongation

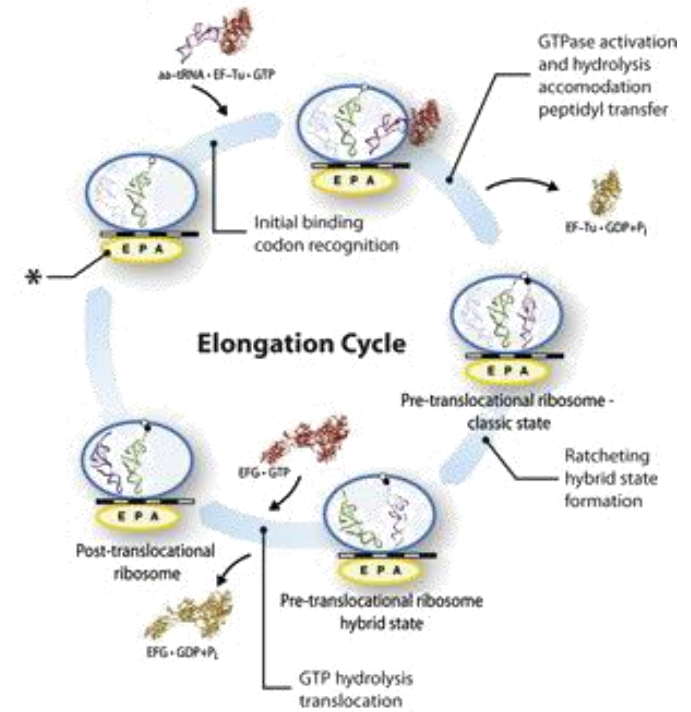
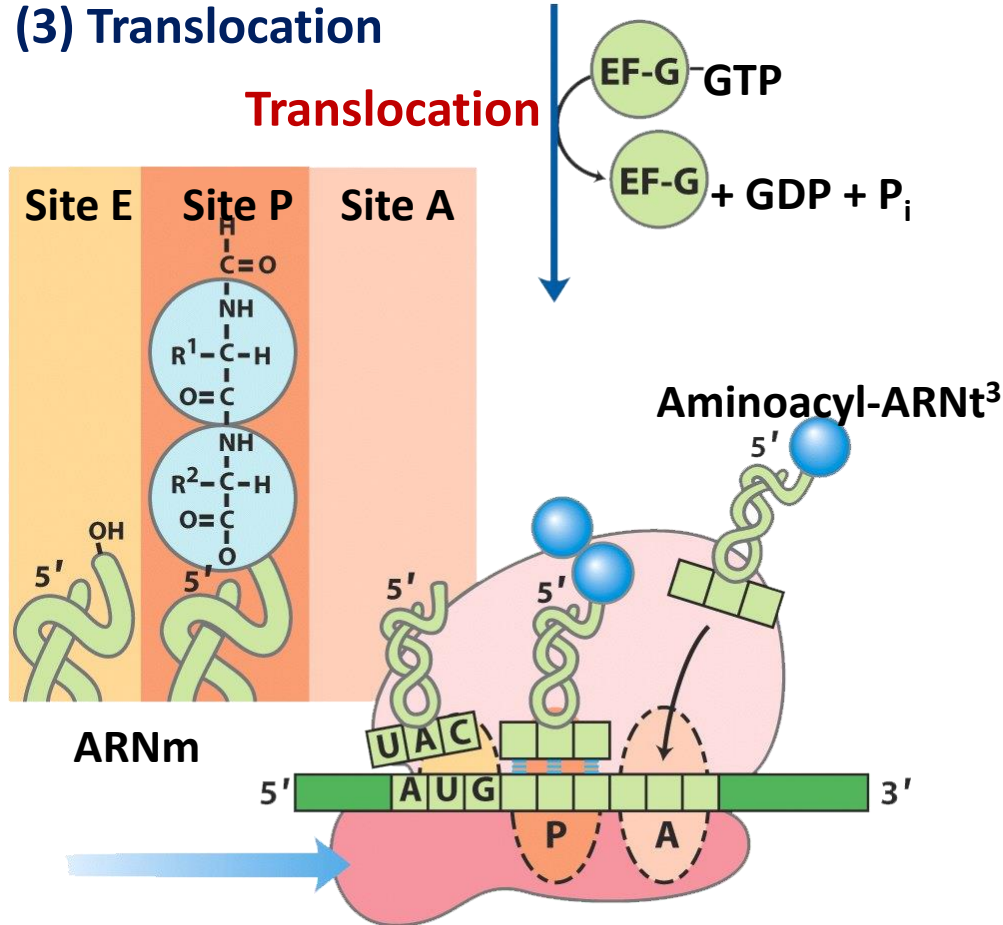
(2) Transpeptidation



Il y a du coup formation de la **liaison peptidique** entre de **COOH** de la **fMet** et le **NH₂** de l'**aa₂** porté par le tRNA^{AA2} qui est dans le site A, ce qui veut dire que l'**aa₁** a été transféré sur l'**aa₂** et du coup le tRNA^{AA2} porte un didpeptide: **fMet-aa₂-tRNA^{AA2}**.

IV. L'élongation

(3) Translocation

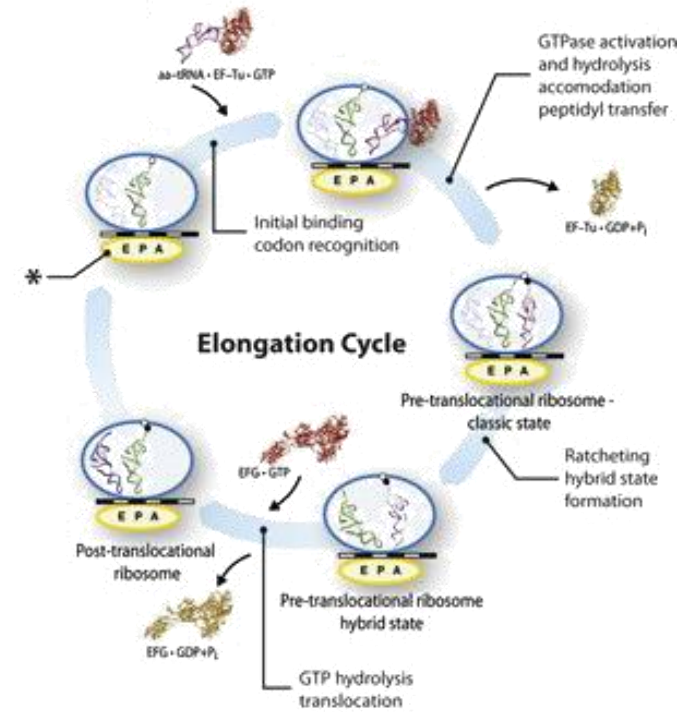
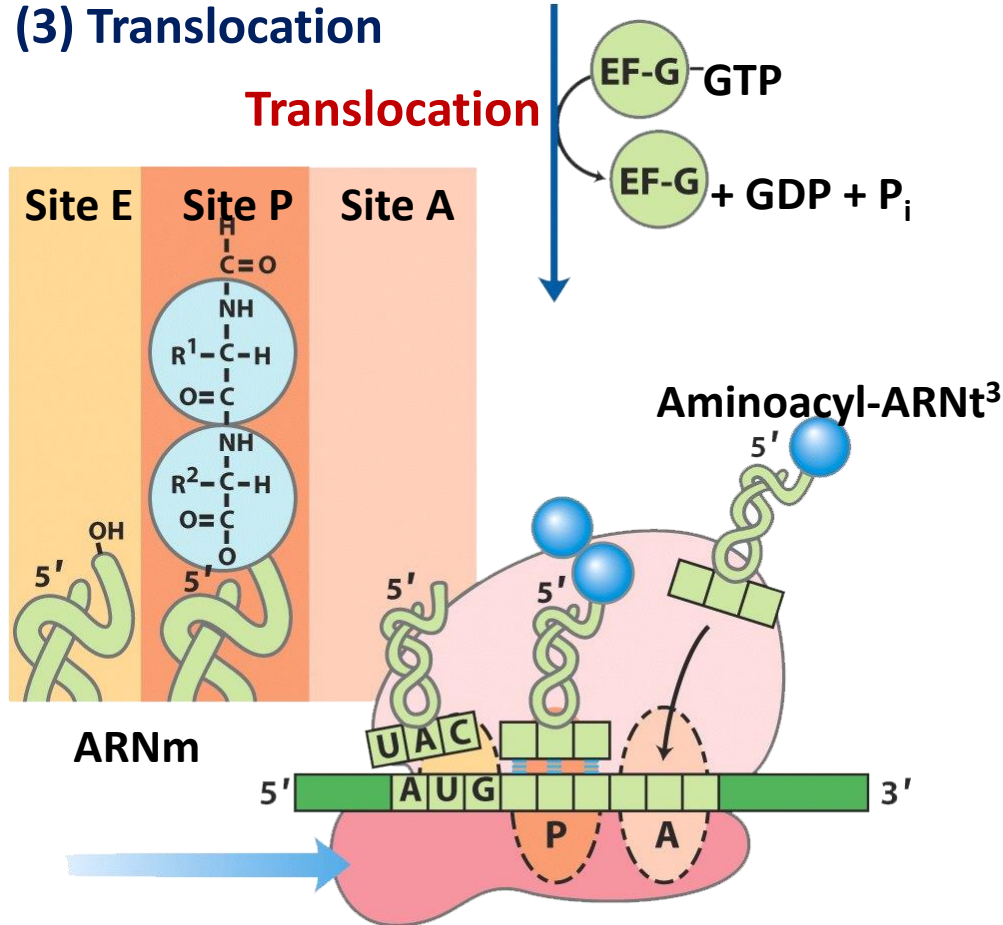


Sens de déplacement du ribosome

Du coup il y a un peptidyl-ARNt ($f\text{Met-aa}_2\text{-tRNA}^{\text{AA}_2}$) dans le site A, il faut le faire migrer dans le site P qui est le site dédié aux peptidyl-ARNt. Ceci nécessite un EF = EF-G (la translocase). EF-G est une **protéine G** qui peut lier **soit le GTP** soit le GDP ne peut provoquer la translocation que sous forme **EF-G • GTP** et fait glisser **fMet-aa₂-tRNA^{AA₂}** dans le site P. **EF-G • GDP** quitte le ribosome et nécessitera un GEF pour être rechargé en GTP. Le tRNA_i^{Met} a été transloqué dans le site E et quittera le ribosome lors de la fixation d'un nouvel aa-ARNt dans le site A.

IV. L'élongation

(3) Translocation

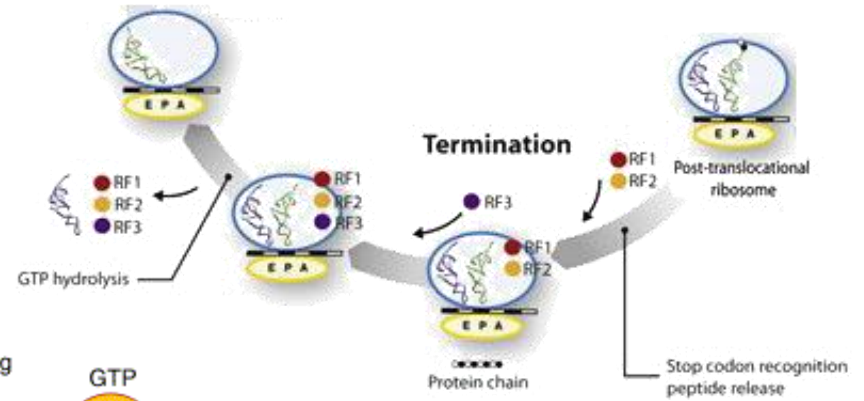


EF-Tu • GTP • aa₃-ARNt^{AA3} dépose l'aa₃-ARNt^{AA3} dans le site A et hydrolyse son GTP et quitte le site A. Le NH₂ de l'aa₃ porté par l'ARNt^{AA3} qui est dans le site A attaque la liaison ester entre l'aa₂ du dipeptide fMet-aa₂ et le tRNA^{AA2} qui est dans le site P et du coup le tRNA^{AA3} porte un tripeptide: fMet-aa₂-aa₃-tRNA^{AA3}. Puis translocation et on recommence....

IV. La terminaison

Trends in Biochemical Sciences

Volume 28, Issue 2, February 2003, Pages 99-105



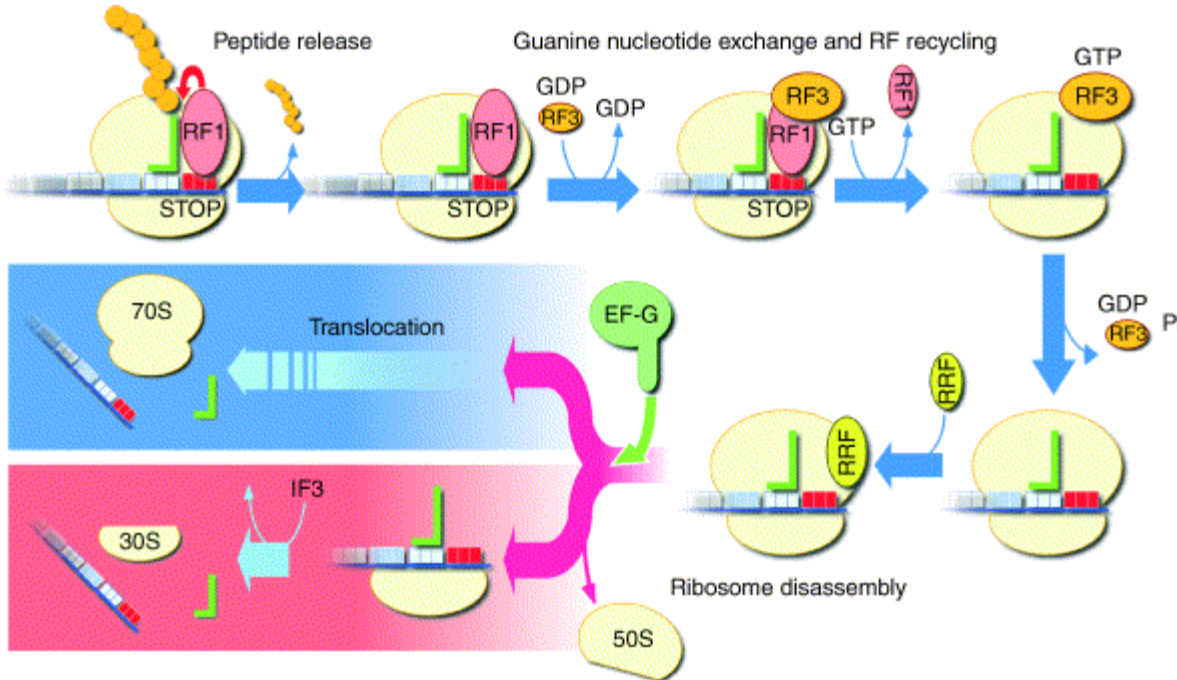
La terminaison nécessite 3 facteurs de terminaison ou **Release Factors (RF)**: **RF1**, **RF2** et **RF3**. La terminaison se produit lorsque le ribosome a un codon STOP dans le site A.

Si le site A contient **UAA** ou **UAG** alors c'est **RF1** qui se fixe dans le site A

Si le site A contient **UAA** ou **UGA** alors c'est **RF2** qui se fixe dans le site A

RF1 (ou RF2) clivent la liaison ester entre le tRNA et la protéine

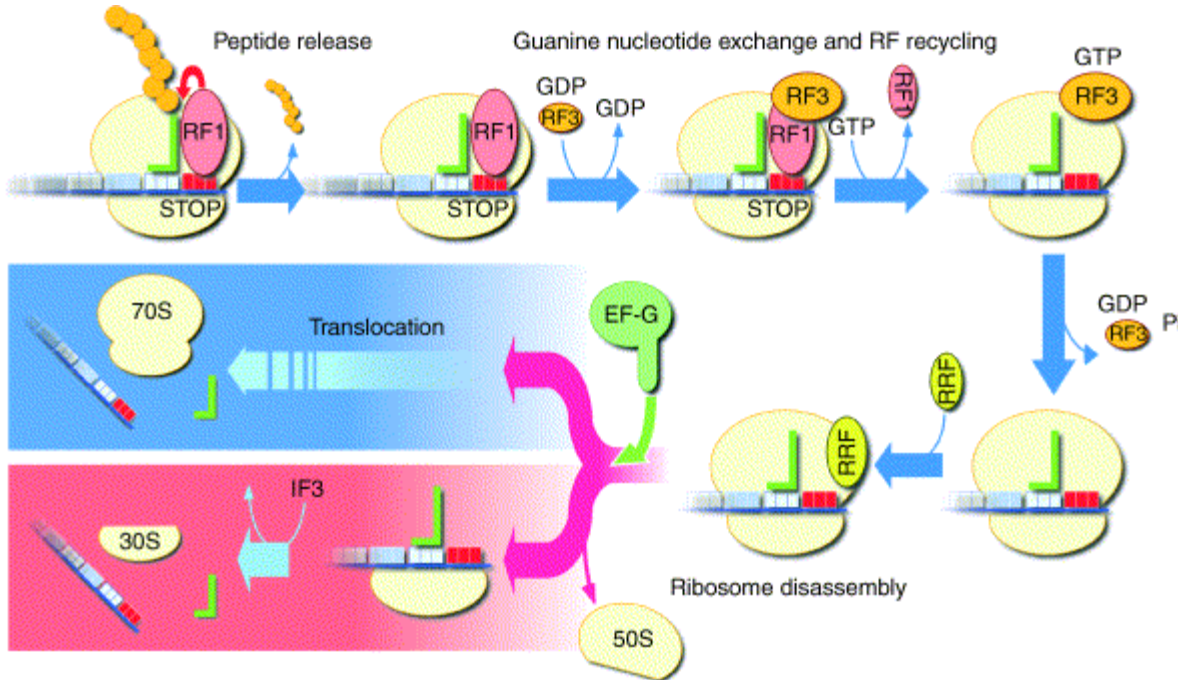
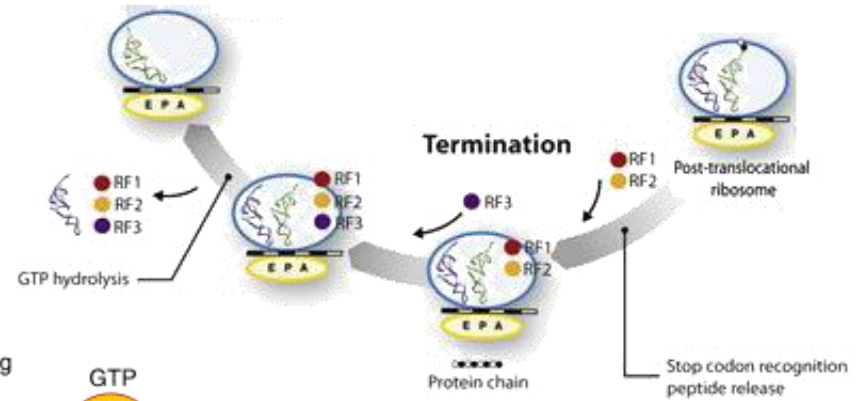
T/BS



IV. La terminaison

Trends in Biochemical Sciences

Volume 28, Issue 2, February 2003, Pages 99-105



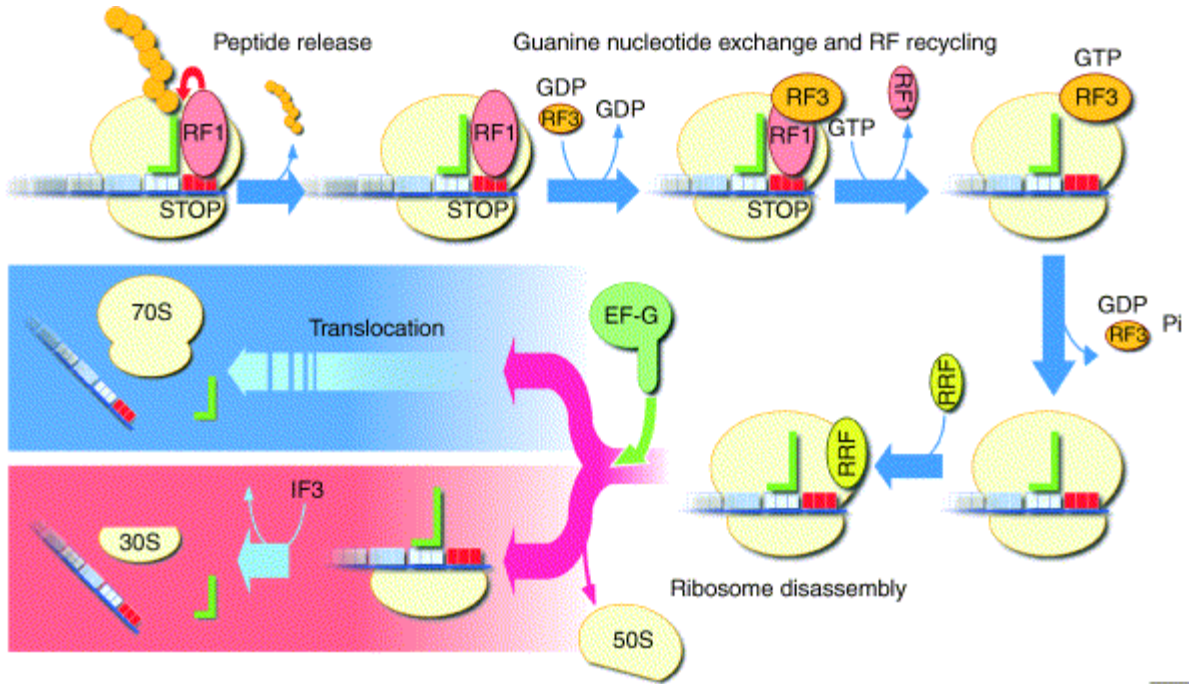
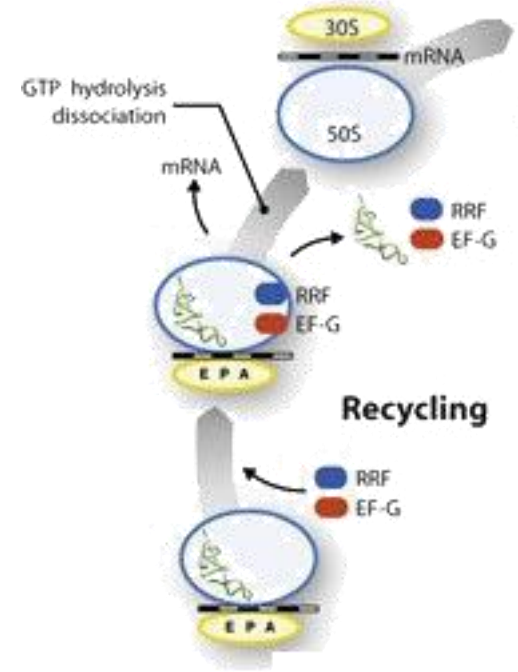
RF3 catalyse la dissociation de RF1 (ou RF2) du site A:

- 1) Sous forme libre RF3 est lié au GDP
- 2) En complexe avec le ribosomes, RF1 (ou RF2) agit comme un GEF Guanine nucleotide Exchange Factor et lors du relargage de la protéine, provoque la liaison du GTP à RF3.
- 3) RF3 change de conformation avec forte affinité pour le ribosome et provoque la dissociation de RF1 (ou RF2).
- 4) Puis RF3 se dissocie du ribosome après hydrolyse du GTP.
- 5) RRF et EF-G se lient au 70S et dans une réaction GTP-dépendante provoque la dissociation des 2 su.

IV. Recyclage

[Trends in Biochemical Sciences](#)

Volume 28, Issue 2, February 2003, Pages 99-105



T/BS

RRF et EF-G se lient au 70S et dans une réaction GTP-dépendante provoque la dissociation des 2 su.

Regarder à nouveau le film: [translation_movie.wmv](#)