

8. La synthèse protéique

L'expression d'un gène est un processus entier qui décode l'information portée par un gène donné et la traduit en protéine. Il s'agit d'une suite de synthèses chimiques et de réactions aboutissant à la production d'un acide ribonucléique qui sera traduit en protéine.

Dans un premier temps a lieu la synthèse d'un acide ribonucléique (ARNm), dont la séquence est complémentaire d'un des deux brins du gène donc identique à celle de l'autre brin ; c'est la transcription. Dans un second temps, l'ARN messenger est « lu » par groupe de trois nucléotides (lettres) grâce à un ARN complémentaire qui porte l'acide aminé correspondant ; c'est la traduction

1. La transcription :

Chez les procaryotes, il n'y a pas vraiment de noyau, alors que chez les eucaryotes il est très clairement défini. Ainsi, les procaryotes peuvent coupler transcription et traduction, mais les eucaryotes non. La transcription chez les eucaryotes se passe dans le noyau et la traduction se fera dans le cytoplasme.

Un gène a une structure spécifique qui va permettre sa transcription. Il est généralement constitué de :

- **Zone promotrice (promoteur) du brin transcrit du gène :** cette zone est reconnue par l'enzyme catalysant la synthèse d'ARN sur laquelle elle va se fixer. Cette zone n'est pas transcrite.
- **Zone transcrite :** elle porte l'information génétique du gène.
- **Zone de terminaison :** qui porte les signaux de terminaison. Elle est transcrite (mais non traduite).

+1

Promoteur non transcrit	Zone transcrite	Zone de terminaison transcrite
-------------------------	-----------------	-----------------------------------

Chez les eucaryotes un seul gène est transcrit. Chez les procaryotes, il existe le système des opérons ou plusieurs gènes (qui fonctionnent généralement ensemble) qui peuvent être transcrit sur le même ARN.

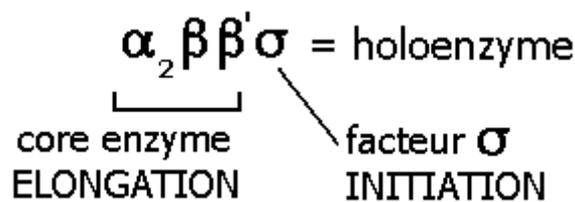
Généralement, la transcription est caractérisée par les points suivants :

- L'ADN sert de matrice (template).
- Seul un des deux brins de l'ADN est transcrit (brin non sens ou antisens).

- La synthèse d'ARN se fait dans le sens 5' → 3' grâce à une ARN polymérase ADN dépendante ou transcriptase.
- La synthèse d'ARN se passe en 3 étapes : initiation, élongation, terminaison.
- L'initiation se fait au niveau d'une région particulière appelée promoteur.
- La synthèse nécessite l'ouverture de l'ADN.
- La terminaison se fait au niveau d'une région particulière appelée terminateur.
- Les brins d'ARN ont une purine en 5' : pppA.
- Un gène est transcrit puis traduit à une chaîne peptidique.

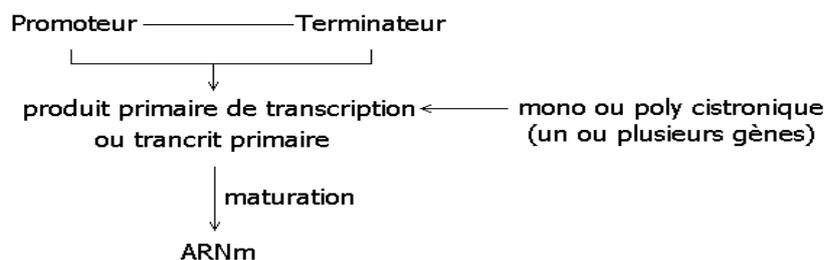
1.1. La transcription chez les procaryotes :

Les procaryotes possèdent une seule ARN polymérase, mais celle-ci est constituée de plusieurs sous-unités ; l'ensemble est appelé holoenzyme. Ce holoenzyme se charge de la synthèse d'ARNt, ARNr ou ARNm indifféremment.



Le core enzyme se lie à l'ADN à n'importe quel niveau et n'entraîne pas l'ouverture de la double hélice d'ADN. L'holoenzyme se lie à l'ADN sur des sites de fixation spécifiques (promoteurs) et entraîne l'ouverture de l'ADN à ce niveau. La synthèse de l'ARN avec comme matrice un des 2 brins d'ADN peut commencer. Très rapidement sigma se détache, car il a une forte affinité pour ce site, si il reste l'élongation ne se fait pas.

Contrairement à l'ADN polymérase, l'ARN polymérase n'a pas d'activité exonucléasique (activité correctrice). Mais le manque d'activité correctrice est sans gravité pour plusieurs raisons : il n'y a pas de transmission à la descendance, le taux d'erreurs est relativement faible (10^{-4} à 10^{-5}), il y a beaucoup de molécules d'ARNm, et leur temps de demi-vie est relativement court.



Le point de départ de la transcription est le nucléotide numéro +1, c'est la paire de base qui correspond au premier nucléotide de l'ARN. Le point de départ fait partie du promoteur qui constitue le signal de début de la transcription. Dans les promoteurs procaryotes, il existe 2 séquences consensus en amont du point de départ : séquence -10 ou encore **TATA box** (car riche en A et en T), et séquence -35 (**TTGACA**).

La zone de terminaison, chez les procaryotes, est une séquence d'ADN palindromique riche en GC suivit d'A : l'ARN va donc être constituée d'une séquence palindromique (CG)_n suivit de U. Ceci forme une tige boucle stable suivit d'un poly U apparié aux poly A. L'appariement entre U et A est peu stable, l'ARN se décroche de l'ADN de façon spontanée, l'ADN se referme et l'enzyme libère l'ADN.

1.2. La transcription chez les eucaryotes :

Bien que la transcription reste dans son principe identique chez les eucaryotes et les procaryotes, des différences parfois importantes, existent. Elles tiennent à la structure de la matrice ADN, aux ARN polymérase, et aux modifications post-transcriptionnelles que subissent les transcrits primaires d'ARN.

La formation des ARN se fait grâce à 3 ARN polymérase ADN dépendantes : ARN pol I (transcription des ARNr 5'8S, 18S et 28S), ARN pol II (ARNm primaire et ARNsn) et ARN pol III (ARNt et ARN 5S).

Les ARN polymérase, chez les eucaryotes, sont incapables à elles seules, d'initier la transcription. Il faut pour cela tout un assortiment de protéines appelées facteurs de transcription généraux qui vont se lier au promoteur. Cet assemblage fournit différentes possibilités de régulation de l'initiation de la transcription.

La région promotrice comporte essentiellement les séquences suivantes :

- **TATA box** : rencontrée chez presque tous les gènes, elle est située en - 32.
- **GC box** : elle se trouve dans la zone promotrice de certains gènes à un emplacement variable (- 110, - 40).
- **CCAAT box** : présente dans de nombreux gènes, elle est située dans la zone promotrice entre - 120 et - 80.

Les signaux de terminaison de la transcription sont mal compris. Cependant, il existe une séquence consensus **AAUAAA** servant apparemment de signal pour qu'une ARN-endonucléase vienne couper le transcrit en 3'.

➤ **Maturation de l'ARNm :**

Les ARNm eucaryotes sont synthétisés sous forme d'un précurseur, ils subissent une maturation pour donner un ARNm actif. Chez les eucaryotes, un gène n'est pas une séquence

codante mais un ensemble de séquences codantes (exons) séparées par des séquences non-codantes (introns). Juste après la transcription, on a un ARNm primaire ou hnRNA (ARN hétérogène nucléaire) constitué d'exons et d'introns.

Le hnRNA subit plusieurs maturations successives :

1. L'ajout d'une coiffe en 5' : l'extrémité 5' du transcrit primaire est coiffée par addition d'une guanine méthylée par une transférase sur le n7. Cette coiffe participe à l'initiation de la traduction en augmentant l'affinité des enzymes de la traduction pour cet ARN et empêche sa dégradation de l'ARN par une exonucléase.

2. La polyadénylation en 3' : une queue polyA est ajoutée grâce à la polyA polymérase. La polyadénylation protège l'extrémité 3' de l'ARN des exonucléases.

3. L'excision-épissage (splicing) : le splicing consiste en l'élimination des introns et la ligature des exons par la suite. Il se fait grâce à des complexes ribonucléoprotéiques hnRNP ou de manière spontanée par un processus auto-catalytique. Avec un même hnRNA, si on fait un type d'épissage on aura un type de protéine, et si on change le type d'épissage, on aura une autre protéine.

2. La traduction :

C'est le mécanisme par lequel le flux d'information va passer de la forme acide nucléique (alphabet à 4 lettres) à la forme protéine (alphabet à 20 lettres) selon un code universel : le code génétique.

Une cellule garde toujours dans son cytoplasme 20 type d'acides aminés. L'activation de ces derniers est nécessaire à la traduction, Ces derniers se lient à une enzyme l'aminoacyl ARNt synthétase. Les aminoacyl-ARNt synthétases sont capables de relier le bon acide aminé à l'ARNt porteur du bon anticodon. Il en existe autant que d'acides aminés et chacune est capable de reconnaître les différents ARNt synonymes.

2.1. Le code génétique :

La traduction consiste à faire correspondre la séquence nucléotidique de l'ARNm à la séquence d'acides aminés de la protéine. La question qui s'impose est : *comment peut-on faire correspondre les quatre nucléotides de l'information génétique aux vingt acides aminés des protéines?*

La réponse à cette question a permis de définir le code génétique comme étant un arrangement qui fait correspondre trois nucléotides à un acide aminé. Cela donne soixante quatre combinaisons possibles couvrant largement les vingt acides aminés naturels. Le triplet

qui définit un acide aminé est appelé **un codon** d'où le code génétique est formé de soixante quatre codons. Ce code possède quatre propriétés:

- **Le code génétique est dégénéré** : c'est-à-dire que la plupart des acides aminés sont définis par plus d'un seul codon.
- **Le code génétique n'est pas chevauchant** : c'est à dire qu'un nucléotide n'appartient qu'à un seul codon et la lecture se fait codon par codon.
- **Le code génétique est à quelques exceptions près universelles.**
- **Trois codons ne définissent aucun acide aminé** et sont ainsi appelés **codons stop** ou **codons non sens**. Il s'agit des codons UAA, UAG et UGA.

La machine de synthèse est formée d'acides aminés, de l'ARNm, de ribosomes et de l'ARNt.

2.2. La traduction chez les eucaryotes :

La lecture de la matrice ARNm se fait dans le sens 5' → 3' et l'ensemble se passe en 3 étapes: **initiation**, **élongation** et **terminaison**. Ces 3 étapes sont catalysées par des **facteurs IF (initiation factors), EF (élongation factor) et RF (release factor)**.

Les sous-unités du ribosome (40S et 60S) s'assemblent au niveau du codon initiateur (5'AUG3'), c'est l'initiation. Le methionine-tRNA (ARNt-met) qui ne reconnaît que ce codon initiateur se fixe au niveau du site P (Peptidyl) du ribosome.

Ensuite vient l'élongation, le ribosome se déplace vers l'extrémité 3' de l'ARNm, les aatRNA se logent dans le site A (Aminoacyl) du ribosome en respectant évidemment la séquence de l'ARNm (code génétique). Le peptide fixé sur le tRNA (site P) se détache et se fixe sur l'acide aminé qui est lié au tRNA (site A), c'est l'élongation. Le ribosome se déplace vers le 3' de l'ARNm. Le tRNA du site P est libéré, celui qui était dans le site A vient dans le site P. Le site A est alors libre et peut accueillir le prochain aatRNA.

Arrivé au codon terminateur (soit UAA, UAG ou UGA) comme il n'existe pas de tRNA correspondant, le ribosome se disloque, c'est la terminaison. Elle se passe grâce aux facteurs de terminaison RF (Release Factor) qui imitent la structure 3D d'un tRNA permettant ainsi le décrochage du peptide lié au tRNA (site P) donnant ainsi la protéine.

2.3. La traduction chez les procaryotes :

À la différence des ARNm eucaryotes, ceux procaryotes sont polycistroniques, c'est à dire qu'un ARNm code pour plusieurs protéines (structure en opérons).

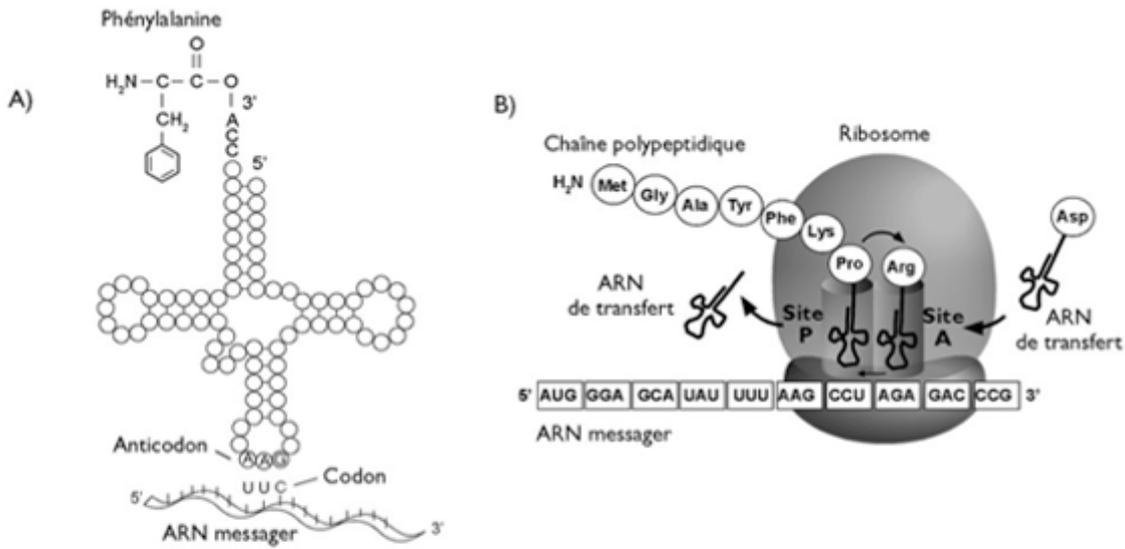
La traduction chez les procaryotes est très proche de celle des eucaryotes. Elle se passe également en 3 étapes qui sont catalysées par des facteurs IF (initiation factor), EF (élongation factor) et RF (release factor).

L'ARNm, libre dans le cytoplasme, sera fixé par une petite sous-unité du ribosome (formé de deux sous-unités 30S et 50S) tel qu'un codon initiateur (toujours le même) AUG sera exposé en premier plan ce qui va faire appel à une grande sous-unité du ribosome qui couvre l'ensemble. Cette sous-unité est formée de deux sites: site A et site P. Le site P se trouve déjà en face du codon initiateur ce qui fait appel à un ARNt initiateur portant l'acide aminé initiateur : la méthionine. La fixation de cet ensemble dans le site P déclenche la synthèse protéique.

La séquence Shine-Dalgarno, séquence consensus: 5' AGGAGG 3', est nécessaire à l'initiation. Elle est située dix nucléotides en amont de l'AUG. L'extrémité 3' du ribosome 16 S (partie de 30 S) interagit avec cette séquence Shine-Dalgarno, et permet ainsi le recrutement de la petite sous unité du ribosome.

Le site A de la grande sous-unité étant libre devant le deuxième codon, ceci fait appel à un deuxième ARNt qui porte le deuxième acide aminé. L'installation de cet ARNt dans le site A stimule la formation de la liaison peptidique entre les deux acides aminés et la rupture de la liaison entre le premier ARNt et son acide aminé ce qui rend le site P libre d'où le glissement du ribosome d'un codon. Ainsi, le site A devient libre devant le troisième codon et le site P héberge le deuxième ARNt chargé des deux premiers acides aminés ce qui stimule l'arrivée d'un troisième ARNt qui porte le troisième acide aminé et ainsi de suite. C'est l'étape d'élongation.

En dernier, en face d'un codon non-sens et comme il n'existe aucun ARNt capable de reconnaître ce codon, la synthèse s'arrête par la libération de la chaîne peptidique et le détachement des deux sous-unités du ribosome.



En A), structure en trèfle de l'ARNt. Il porte un acide aminé (phénylalanine) à son extrémité supérieure ; le site anticodon est situé à son extrémité inférieure. En B), schéma de fonctionnement de l'ARNt : porteur d'un acide aminé (à droite), il va transférer celui-ci à la protéine en cours de synthèse à l'intérieur du ribosome. (Joëlle Brodeur, Martin Toussaint, Projetbleu)

		NUCLÉOTIDE 2 ^{ème} POSITION					
		U	C	A	G		
NUCLÉOTIDE 1 ^{ère} POSITION	U	UUU } phényl- UUC } alanine UUA } leucine UUG }	UCU } UCC } sérine UCA } UCG }	UAU } tyrosine UAC } UAA } non-sens UAG }	UGU } cystéine UGC } UGA } non-sens UGG } tryptophane	U C A G	
	C	CUU } CUC } leucine CUA } CUG }	CCU } CCC } proline CCA } CCG }	CAU } histidine CAC } CAA } glutamine CAG }	CGU } CGC } arginine CGA } CGG }	U C A G	
	A	AUU } AUC } isoleucine AUA } AUG } méthionine	ACU } ACC } thréonine ACA } ACG }	AAU } asparagine AAC } AAA } lysine AAG }	AGU } sérine AGC } AGA } arginine AGG }	U C A G	
	G	GUU } GUC } valine GUA } GUG }	GCU } GCC } alanine GCA } GCG }	GAU } acide GAC } aspartique GAA } acide GAG } glutamique	GGU } GGC } glycine GGA } GGG }	U C A G	
						NUCLÉOTIDE 3 ^{ème} POSITION	

Code génétique