

## 2 L'ADN ET LES PROTÉINES : GÉNÉRALITÉS

Toutes les caractéristiques d'un organisme vivant sont déterminées par le type de **protéines** que fabrique cet être vivant. Ce qui différencie un être vivant d'un autre, un moustique d'un humain ou d'un érable ou un humain d'un autre humain, c'est **les protéines** que chacun fabrique.

De même, ce qui différencie une cellule nerveuse d'un individu d'une cellule de la peau de ce même individu, c'est les **enzymes** que chacune de ces cellules fabriquent. Plusieurs de ces enzymes sont les mêmes, cependant quelques-unes diffèrent, ce qui explique les différences de forme et de fonctions entre ces cellules.

Chaque protéine est fabriquée par la cellule qui l'utilise à partir des **acides aminés** apportés par le sang. Cependant, pour fabriquer une protéine, en plus des acides aminés, la cellule doit posséder la « **recette** » de la protéine, c'est-à-dire, l'information lui dictant l'ordre dans lequel elle doit assembler les divers acides aminés formant cette protéine. La moindre erreur dans l'assemblage des acides aminés, la moindre « **faute d'orthographe** », risque de rendre la protéine non fonctionnelle. Pour chacune des milliers de protéines fabriquées par une cellule, il y a, dans le noyau de la cellule, l'information, « **la recette** », permettant de fabriquer cette protéine. Le noyau des cellules humaines comporte entre 20'000 et 22'000 recettes différentes de protéines (projet du génome humain (HUGO)). On appelle « **gène** » chacune de ces recettes.

On s'est longtemps interrogé sur la nature de ces gènes nécessaires à la synthèse des protéines. En 1944, Oswald Avery et ses collaborateurs démontrèrent que les **gènes** étaient formés d'une **matière chimique** qu'on ne retrouve pratiquement que dans le noyau des cellules : l'**acide désoxyribonucléique** ou **ADN**.

On ignorait cependant, à cette époque, quelle était la **structure** exacte de cette molécule. C'est en **1953** que deux chercheurs, Francis **Crick** et James **Watson**, publièrent dans la revue *Nature* un des plus célèbres articles de l'histoire de la biologie. Cet article rapportait leur découverte de la structure de la molécule d'ADN.

L'ADN est une très longue molécule pelotonnée située dans le noyau des cellules ou bien enroulée librement dans la cellule bactérienne qui, elle n'a pas de noyau.

Si l'on arrivait à dérouler tout l'ADN d'une cellule humaine, la longueur du filament (invisible à l'œil nu) atteindrait 1,5 m. Quant à l'ADN de toutes les cellules de notre corps mis bout à bout, il formerait un filament d'une longueur égale à la distance Terre - Lune. L'ADN s'enroule plusieurs fois sur lui-même et donne une structure complexe de câbles tressés associés à des protéines et visibles au microscope ordinaire sous la forme de bâtonnets : ce sont les **chromosomes** (*figures 3 et 5*). Il y en a **23 paires** dans chaque cellule humaine et leur nombre n'a rien à voir avec la complexité de l'espèce.

Entre les périodes de division de la cellule, l'ADN se présente sous forme de **chromatine** (*figures 4 et 5*). Elle se présente le plus souvent sous la forme d'une matière **sans structure** particulière. A certains moments de la vie de la cellule (aux moments des multiplications (divisions)), la chromatine perd son aspect diffus et se condense en structures bien définies : les chromosomes.

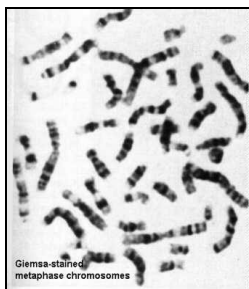


Figure 3 : Les chromosomes

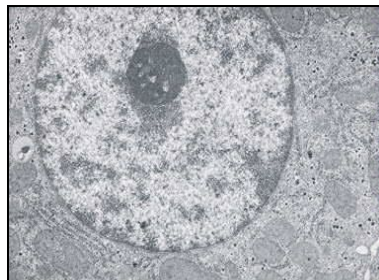


Figure 4 : La chromatine

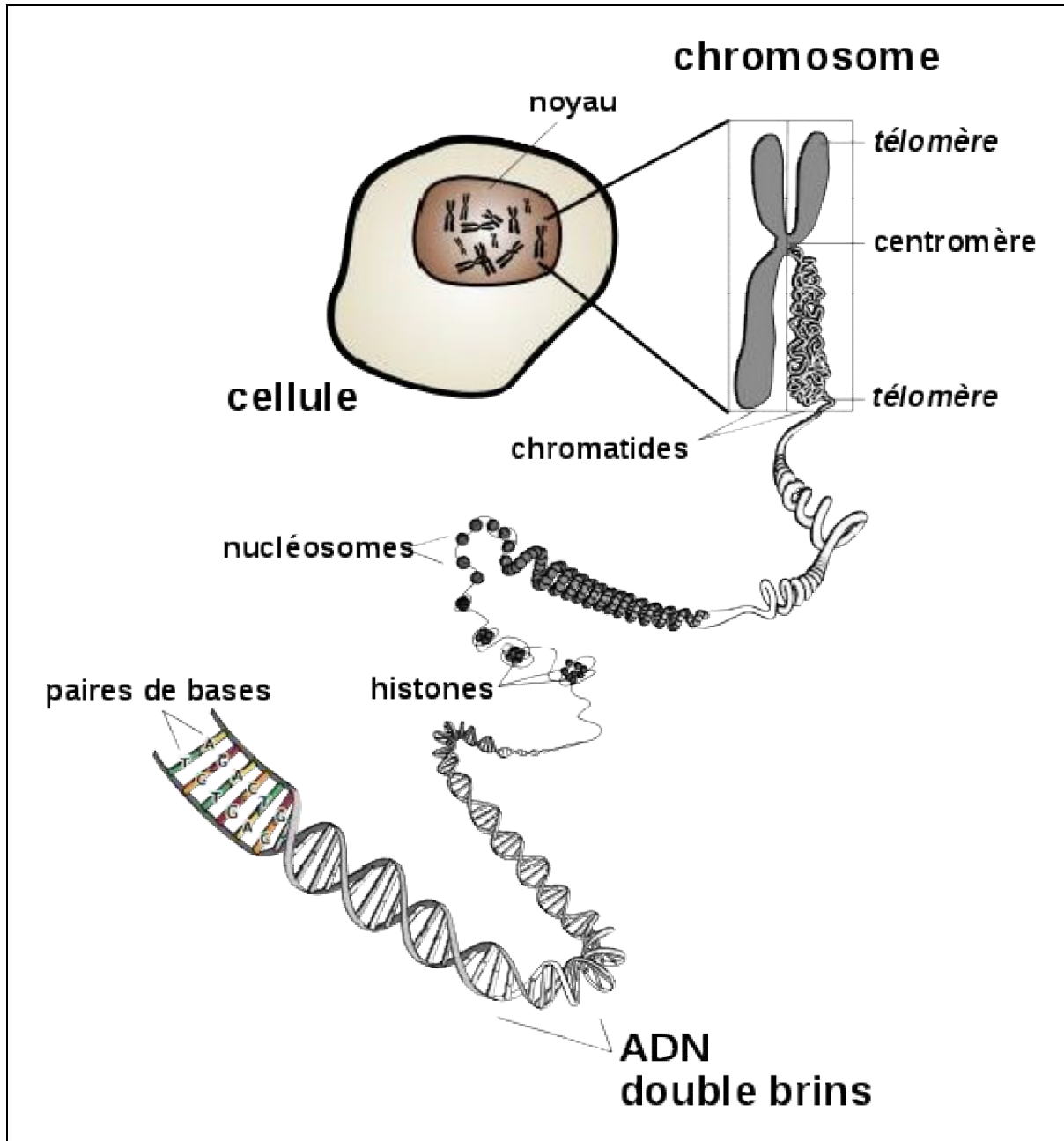


Figure 5 : Le noyau, l'ADN et les chromosomes

### 3 LA STRUCTURE DE LA MOLÉCULE D'ADN

L'ADN, l'acide désoxyribonucléique, est un **polymère** formé de molécules plus petites, les **nucléotides** (figure 6).

Un nucléotide est une molécule formée de **trois unités moléculaires** :

- ◆ Une **base azotée** (il y en a 4 différentes)
  - A = **adénine**
  - T = **thymine**
  - G = **guanine**
  - C = **cytosine**
  
- ◆ Un sucre, le **désoxyribose**
  
- ◆ Un groupement **phosphate** ( $\text{PO}_4^{3-}$ )

Les nucléotides peuvent se lier les uns aux autres par l'intermédiaire de leur groupement phosphate et de leur sucre. Ils peuvent donc former de longues chaînes :

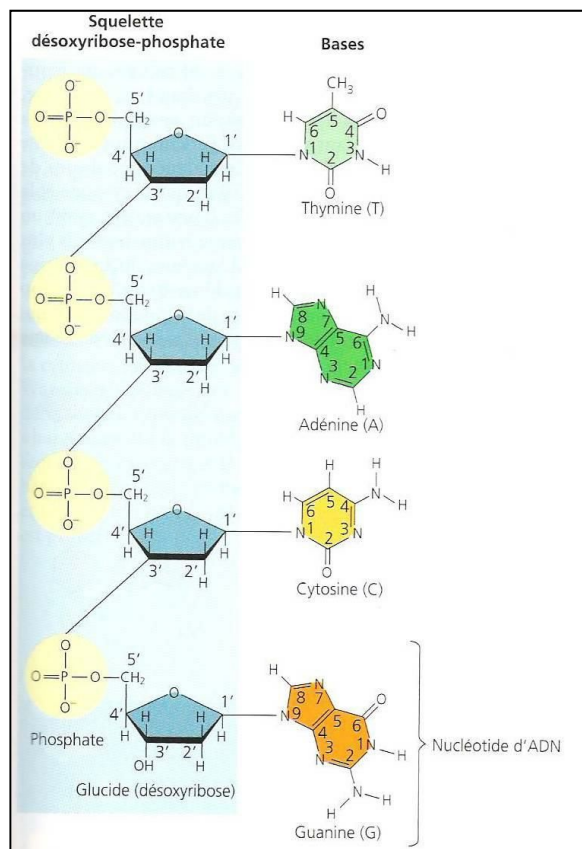


Figure 6 : La structure d'un brin d'ADN

**L'ADN est-elle donc une grande molécule formée par l'enchaînement les uns aux autres de nombreux nucléotides ?**

Pas tout à fait. L'ADN est en fait constitué de deux chaînes de nucléotides (figure 7).

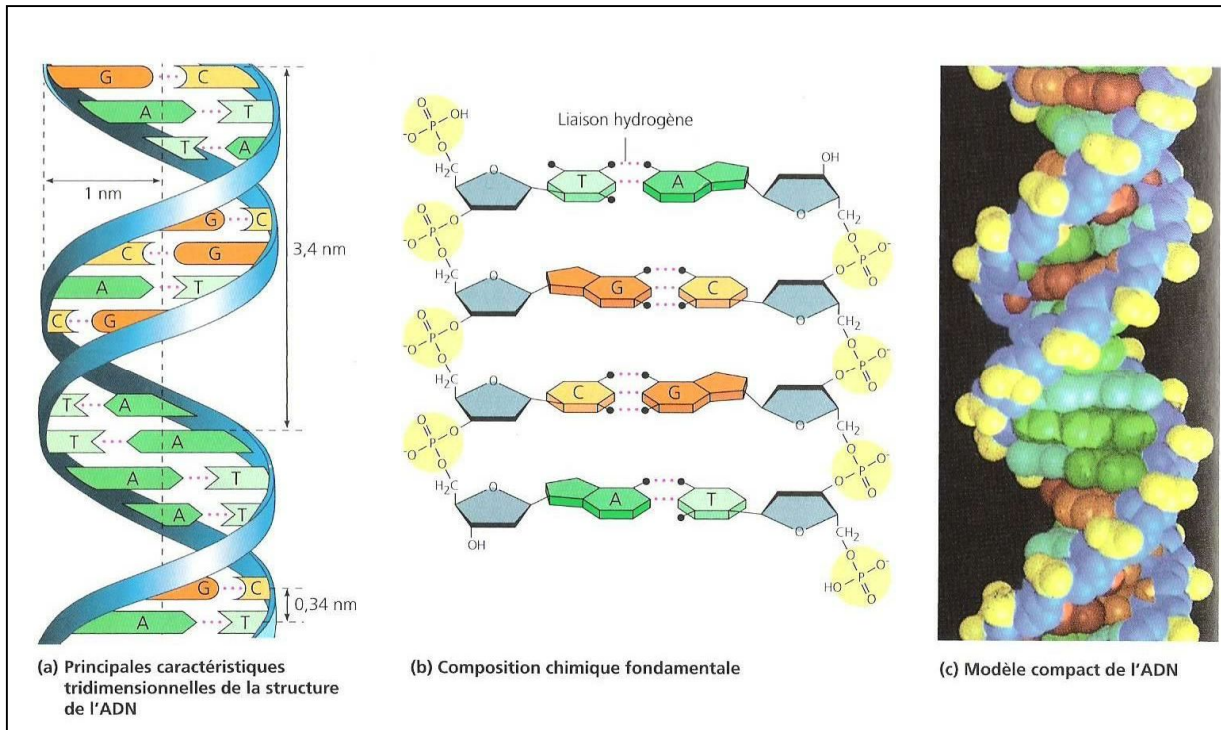


Figure 7 : La double hélice d'ADN

Lorsqu'on défait un segment d'ADN en nucléotides, on obtient toujours exactement le même nombre de nucléotides A que T et le même nombre de nucléotides C que G. Crick et Watson, en étudiant la structure des nucléotides, constatèrent qu'il pouvait se former des liaisons faibles (**liaisons hydrogènes**) entre les nucléotides A et T et entre les C et G.

Il résulte de cette propriété que deux chaînes de nucléotides peuvent se lier ensemble par leurs bases azotées si leurs bases sont **complémentaires** (A face à T et C face à G) ce qui est le cas dans la molécule d'ADN.

Compte tenu de l'angle entre les liaisons des nucléotides, la double chaîne (l'échelle) n'est pas plane mais "tordue" autour de son axe central formant ainsi une **double hélice** (comme un escalier en spirale, les liaisons sucre-phosphate (liaisons **phosphodiesters** = liaisons fortes) correspondant aux deux rampes et les liaisons entre les bases azotées correspondant aux marches).

Si l'on dénature l'ADN, ce seront d'abord les liaisons hydrogène qui vont céder.

## 4 DE L'ADN AUX PROTÉINES

On sait aujourd'hui que la « *recette* » pour la fabrication des protéines, l'information correspondant à l'ordre d'assemblage des acides aminés d'une protéine, est présente sous forme d'ADN dans le **noyau** de la cellule. L'ADN constitue pour la cellule un gigantesque livre de recette comportant toutes les recettes de protéines que la cellule peut fabriquer.

Comme nous l'avons également vu, l'ADN est une énorme molécule formée de l'union les unes aux autres de centaines de milliers de molécules plus petites appelées nucléotides.

C'est en fait l'**arrangement** des nucléotides (= séquence des bases) sur l'ADN qui va déterminer la nature de la protéine. La quantité d'information contenue dans les 46 chromosomes d'une cellule humaine correspond à peu près à l'information contenue dans une encyclopédie de 1000 volumes de 1000 pages chacun. Cette information est contenue dans un volume équivalent environ à un millième du volume d'une tête d'épingle, le volume du noyau cellulaire.

Pour lire le message, il faut évidemment connaître le code. Ce code, le « **code génétique** », a été découvert au début des années 60.

Pour être complet il faut parler de l'**ARN (l'acide ribonucléique)**. Il s'agit de la molécule qui va permettre le passage entre l'ADN et la protéine (*figure 8*).

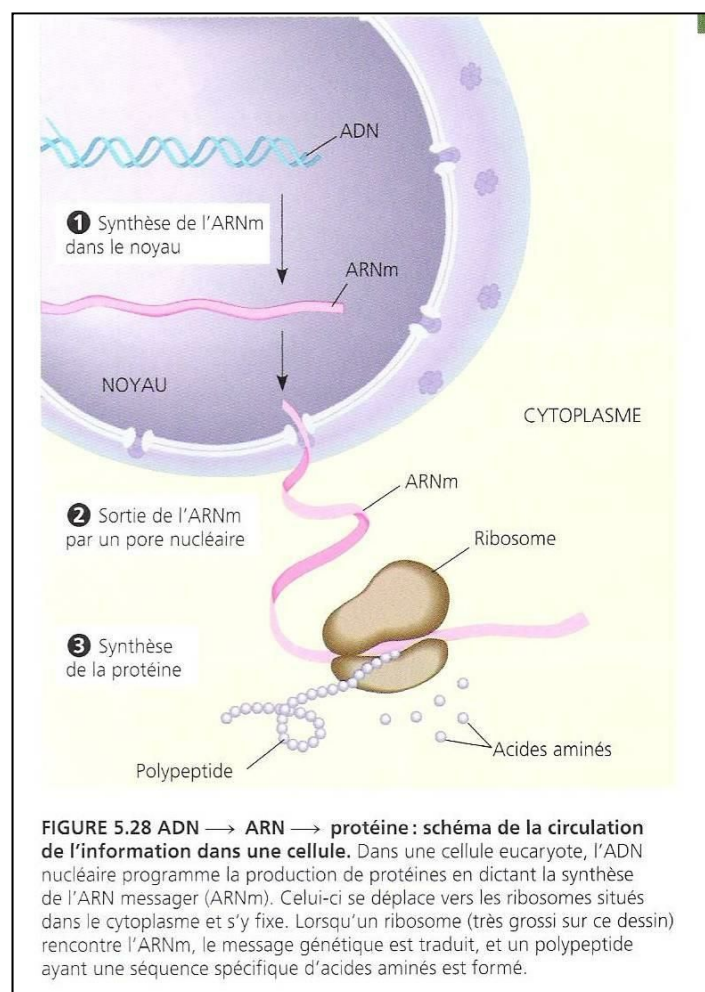


Figure 8 : De l'ADN aux protéines

## 4.1 L'acide ribonucléique (ARN)

Tout comme l'ADN, les molécules d'ARN sont fabriquées dans le **noyau** de la cellule. Toutefois, contrairement à l'ADN, l'ARN ne se limite pas au noyau. Il peut **migrer** dans d'autres parties de la cellule. Tout comme l'ADN, les molécules d'ARN sont composées de **nucléotides**. Toutefois, alors que les molécules d'ADN sont formées de deux brins parallèles de nucléotides, une molécule d'ARN n'en compte qu'un.

La structure de l'ARN est à peu de chose près la même que celle de l'ADN. Les différences se situent à différents niveaux (*tableau 1*) :

- ◆ L'**uracile** remplace la thymine
- ◆ Le degré de polymérisation est moins important
- ◆ Le sucre de l'ARN est du **ribose** et non du désoxyribose
- ◆ L'ARN est **monocaténaire**, c'est-à-dire qu'il ne possède qu'un brin au lieu de deux pour l'ADN

	<b>ADN</b>	<b>ARN</b>
<i>Fonction</i>	Gènes; contrôle de la synthèse des protéines	Aide l'ADN à synthétiser les protéines
<i>Sucre</i>	Désoxyribose	Ribose
<i>Bases</i>	Adénine, guanine, thymine, cytosine	Adénine, guanine, uracile, cytosine
<i>Structure</i>	Deux brins reliés par les bases; enroulement hélicoïdal	Un seul brin non hélicoïdal

*Tableau 1 : Comparaison ADN-ARN*

## 4.2 La transcription : synthèse de l'ARN à partir de l'ADN

*Définition :* \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**L'information génétique siège dans le noyau alors que la synthèse des protéines s'effectue dans le cytoplasme (ribosomes).**

*Comment se fait le transfert du message d'un site à l'autre ?*

L'ARN messager est transcrit à partir du brin codant d'un gène. Il transmet l'information de l'ADN aux structures cellulaires assurant la synthèse des protéines. Une enzyme appelée **ARN polymérase** écarte les deux brins d'ADN et assemble les nucléotides d'ARN au fur et à mesure que leur base s'apparie avec la matrice d'ADN (*figure 9*). Les ARN polymérases ne peuvent ajouter des nucléotides qu'à l'extrémité 3' du polymère en cours de synthèse. Par conséquent, la molécule d'ARN s'allonge dans la **direction 5' → 3'**.

Des séquences particulières d'ADN marquent le début et la fin de la transcription du gène. La séquence d'ADN à laquelle l'ARN polymérase se lie pour commencer la transcription est appelée **promoteur**. Quant à la séquence qui marque la fin de la transcription, elle est appelée **terminateur**.

Les **trois étapes** de la transcription illustrées à la *figure 9* sont :

- 1) L'**initiation**,
- 2) L'**élongation**,
- 3) La **terminaison**.

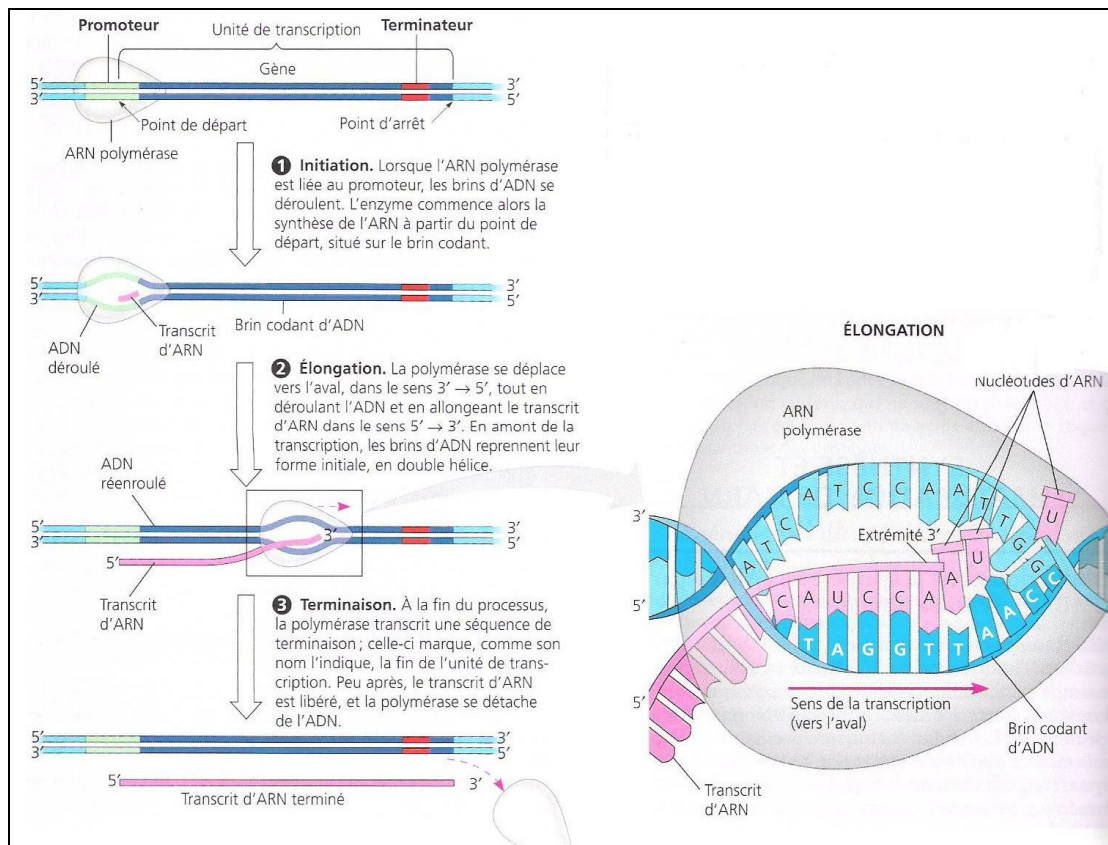


Figure 9 : Les étapes de la transcription

1) Liaison de l'ARN polymérase et **initiation** de la transcription :

Le **promoteur** d'un gène inclut le **point de départ** de la transcription (les nucléotides à partir desquels la synthèse de l'ARN commence). En plus de servir de site de liaison à l'ARN polymérase et de marquer le début de la transcription, il **détermine** lequel des deux **brins** de l'hélice d'ADN sera **codant**. Un ensemble de protéines appelées **facteurs de transcription** servent d'intermédiaires : ce sont eux qui permettent la liaison de l'ARN polymérase et le début de la transcription. Ce n'est que lorsqu'un certain nombre de facteurs de transcription ont été fixés au promoteur que l'ARN polymérase se lie à celui-ci. L'ensemble constitué par l'ARN polymérase et les facteurs de transcriptions liés au promoteur est appelé **complexe d'initiation**. Une fois que la polymérase est fermement liée au promoteur, les deux brins d'ADN se déroulent à cet endroit, et l'enzyme commence à transcrire le brin codant.

2) L'**élongation** du brin d'ARN :

Pendant qu'elle se déplace le long de l'ADN, l'ARN polymérase continue de dérouler la double hélice ; elle expose 10 à 20 bases environ à la fois. Elle permet ainsi leur appariement avec les nouveaux nucléotides d'ARN (*figure 9*). Précisons qu'elle ajoute ceux-ci à l'extrémité 3' de la molécule d'ARN en cours de synthèse, tout en avançant le long de la double hélice. La partie de la double hélice d'ADN qui se trouve derrière l'ARN polymérase se reconstitue, et la nouvelle molécule d'ARN se détache progressivement du brin codant d'ADN. Un même gène peut être transcrit simultanément par plusieurs molécules d'ARN polymérase, qui se suivent tel un convoi de camions. De chaque molécule émerge un nouveau filament d'ARN en formation : sa longueur reflète la distance parcourue par l'enzyme sur le brin codant



depuis le point de départ. La transcription simultanée d'un même gène par un grand nombre de molécule de polymérase accroît le nombre d'ARN pré messenger fabriqués ; cela permet à la cellule de produire la protéine correspondante en grande quantité.

3) La **terminaison** de la transcription :

Elle se produit lorsque l'ARN polymérase entre en contact avec la séquence **terminateur**. A ce moment-là, l'ARN polymérase se détache et libère l'ARN pré messenger synthétisé.

A la fin de la transcription, l'**ARN pré messenger** doit encore subir quelques modifications avant d'être exporté vers le cytoplasme. Ces modifications sont appelées **maturation** de l'ARN pré messenger. Tout d'abord, les deux extrémités sont modifiées, puis dans l'ARN pré messenger, il y a des régions codantes (**exons**) et des régions non-codantes (**introns**). Au cours de la maturation, les introns sont éliminés et l'ARNm mature sera constitué uniquement d'exons et des extrémités modifiées.

*Exercice :*

a) Trouver le brin complémentaire du brin d'ADN suivant :

**3' T A C G G C C T T A T A T A T G C G C A A 5'**

-----

b) Déterminer le brin codant et transcrivez-le :

-----

### 4.3 Le code génétique

L'ARNm joue le rôle de molécule messagère entre le noyau cellulaire et les ribosomes ; il sert d'intermédiaire entre le lieu de stockage et le lieu d'expression de l'information génétique.

Les instructions concernant la **structure d'une protéine** doivent être contenues dans **l'ordre des bases de l'ARNm**.

Le passage de l'ordre du gène à l'ordre de la protéine fait nécessairement intervenir un **système de correspondance** que l'on appelle le **code génétique** (tableau 2).

- ◆ L'ordre dans lequel s'enchaînent les **acides aminés** dans la **protéine** est déterminé par l'ordre dans lequel s'enchaînent les nucléotides dans le brin d'ADN.
- ◆ L'ADN contient 4 types de nucléotides qui diffèrent par leur base azotée. La protéine est formée par **20** acides aminés différents.  
Le problème est donc de savoir comment, à partir d'un alphabet à quatre lettres, nous pouvons composer un code capable de désigner les 20 objets que sont les acides aminés.
- ◆ Si un mot du code pour un acide aminé n'était constitué que d'une seule base, seuls quatre acides aminés pourraient être codés. Des mots de code à deux bases avec 16 ( $4^2$ ) combinaisons possibles ne suffisent pas non plus. Seul un code à trois ( $4^3$ ) donne suffisamment de combinaisons différentes pour permettre de coder les 20 acides aminés utilisés par les organismes.
- ◆ L'information élémentaire est donc portée par un triplet de nucléotides ou **codon**. Chaque codon (3 nucléotides) spécifie la place d'un acide aminé dans la séquence qui constitue une protéine. Mais avec quatre bases différentes, on obtient en code triplet 64 combinaisons possibles, plus qu'il n'est nécessaire pour les 20 acides aminés. Ainsi presque tous les acides aminés sont codés par plusieurs triplets.  
61 triplets codent les acides aminés, les trois autres, appelés « codons stop », sont des signaux codant la fin d'une chaîne lors de la biosynthèse des protéines. Le triplet AUG code par contre le début ou l'initiation d'une synthèse de protéines et en même temps l'acide aminé méthionine. En temps qu'acide aminé de départ la méthionine est plus tard éliminée de la protéine.

Le code génétique est **un code universel**, c'est-à-dire le même chez tous les organismes vivants. Ceci atteste de l'origine unique des organismes vivants.

Résumons les propriétés du code génétique :

- 1) **Universel** : tous les êtres vivants (sauf quelques exceptions) possèdent le même.
- 2) **Non ambiguë** : à un codon correspond un seul et unique acide aminé.
- 3) **Dégénéré** : à un acide aminé peuvent correspondre plusieurs codons (il existe en effet 64 possibilités de codons, et seulement 20 acides aminés).

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

Ce tableau donne les diverses combinaisons possibles des 4 nucléotides pris 3 par 3 et leur « signification ».

Tableau 2: Le code génétique

## 4.4 La traduction

**Définition :** \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Au cours du processus de la traduction, la cellule construit une protéine à partir des instructions contenues dans le message génétique. Celui-ci consiste en une série de **codons** alignés sur une molécule d'ARNm, et il est interprété par une molécule d'ARN d'un autre type, qui porte le nom d'ARN de **transfert** (ARNt). Cet ARN a pour fonction d'acheminer vers un ribosome les molécules d'acides aminés qui se trouvent dans le cytoplasme. Le **ribosome** ajoute chacun des acides aminés que l'ARNt lui apporte à l'extrémité de la chaîne de polypeptide en cours de synthèse (*figure 10*). Ceux-ci se composent de deux sous-unités de grandeur différente qui existent d'abord à l'état dissocié dans le cytoplasme.

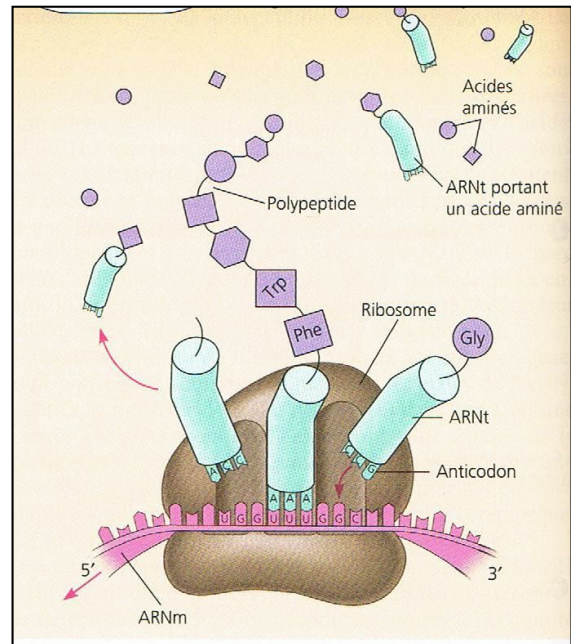


Figure 10 : structure d'un ARNt

Les molécules d'ARNt ne sont pas toutes identiques. Chacune d'elles sert à associer un certain codon d'ARNm avec un acide aminé particulier. C'est ce lien qui constitue la clé de la traduction du message génétique en une séquence d'acides aminés. Cette molécule d'ARN contient d'un côté un **acide aminé** et de l'autre une succession de trois bases, l'**anticodon**, qui peut à chaque fois s'apparier à un triplet complémentaire (**codon**) de l'ARNm.

L'ARNt un « adaptateur-décodeur » permettant la traduction de l'information du langage ARNm en langage protéine. Le "décodeur" correspond à la partie en contact direct avec l'ARNm (c'est l'anticodon). "L'adaptateur" correspond à la partie liée à un acide aminé particulier, déterminé par l'anticodon. Il y a autant d'ARNt qu'il y a de codons pour chaque acide aminé (*figure 11*).

La traduction, ou synthèse, d'un polypeptide comprend trois étapes principales, qui rappellent celles de la transcription : il s'agit de l'initiation, de l'élongation et de la terminaison. Celles-ci ne peuvent se dérouler qu'en présence de « facteurs » protéiques assistant l'ARNm, l'ARNt et les ribosomes (ARNr) pendant la traduction.

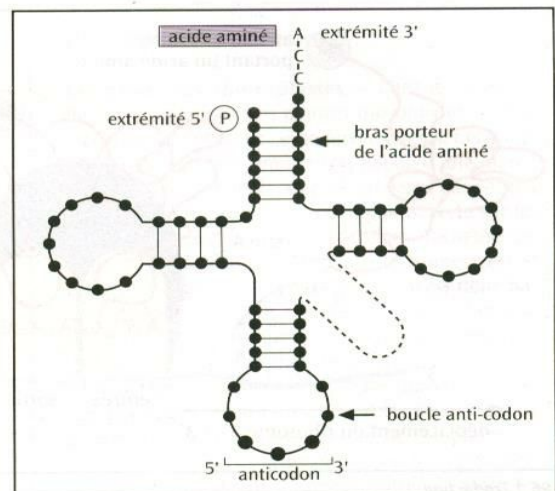


Figure 11 : structure d'un ARNt

### 1) L'initiation (*figure 12*)

Cette étape met en jeu l'ARNm, un ARNt portant le premier acide aminé du polypeptide en cours de formation et les deux-unités d'un ribosome. Les deux sous-unités ribosomales se fixent sur le brin d'ARNm. Les ARN de transfert, formés d'un acide aminé et d'un site anti-codon, sont capables de se fixer à un codon complémentaire formé de 3 nucléotides de l'ARNm. L'ARNt qui porte l'acide aminé méthionine se fixe par complémentarité de

son anticodon sur le premier codon AUG (codon « start »). Des protéines appelées **facteurs d'initiation** jouent un rôle essentiel dans l'agencement de tous ces éléments.

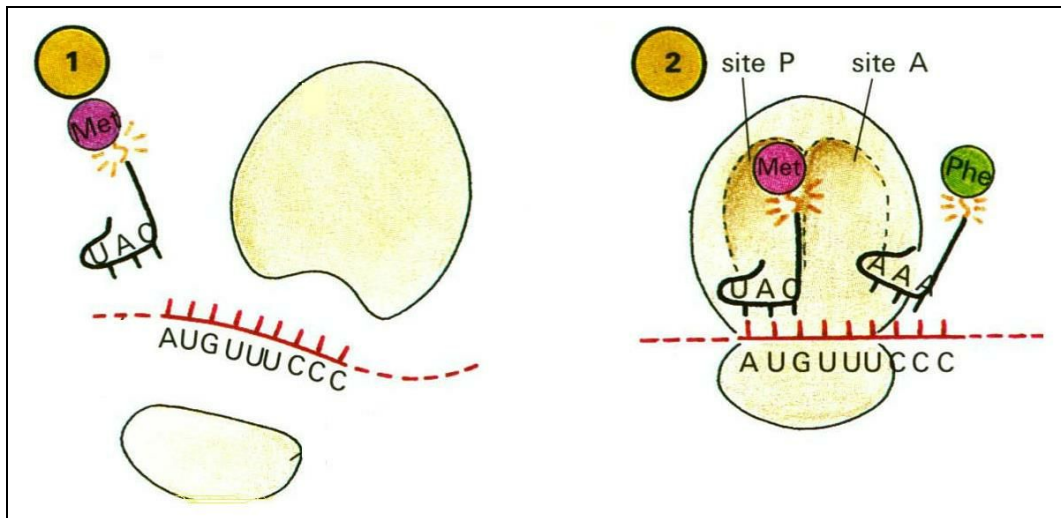


Figure 12 : L'initiation

## 2) L'élongation (figure 13)

Au cours de l'étape de la traduction appelée élongation, les acides aminés sont ajoutés un à un, à la suite l'un de l'autre. Chaque ajout suppose la participation de plusieurs protéines appelées **facteurs d'élongation** et se déroule selon un cycle comptant trois phases :

### a) Reconnaissance du codon

Le codon de l'ARNm placé dans le site A forme des liaisons hydrogène avec l'anticodon d'une molécule d'ARNt qui porte l'acide aminé correspondant. C'est un facteur d'élongation qui achemine l'ARNt jusqu'au site A.

### b) Formation d'une liaison peptidique

Une molécule d'ARNr faisant partie de la grande sous-unité ribosomique forme une liaison peptidique entre le polypeptide émergent du site P et l'acide aminé nouvellement arrivé au site A. Puis, le polypeptide se sépare de l'ARNt auquel il était attaché, et l'acide aminé situé à l'extrémité qui s'est détachée se lie à l'acide aminé apporté au site A par l'ARNt suivant.

### c) Translocation

Le ribosome effectue ensuite la translocation (déplacement) de l'ARNt qui se trouve au site A et qui porte le polypeptide. Il le déplace vers le site P. Au cours de cette translocation, l'anticodon de l'ARNt reste fixé au codon correspondant d'ARNm par des liaisons hydrogènes. L'ARNm avance donc en même temps et place le codon suivant du message génétique au site A, où il sera traduit. Pendant ce temps, l'ARNt qui se trouvait au site P est acheminé au site E (pour *exit*). De là, il quitte le ribosome. L'ARNm traverse toujours le ribosome dans la même direction, c'est-à-dire en commençant par l'extrémité 5'. Cela revient à dire que le ribosome se déplace dans le sens 5' → 3' sur l'ARNm. Il suffit de se souvenir que le ribosome et l'ARNm bougent l'un par rapport à l'autre, en sens inverse et codon par codon. Le cycle d'élongation dure moins d'un dixième de seconde et se

répète chaque fois qu'un acide aminé est ajouté à la chaîne polypeptidique, et ce, jusqu'à ce que celle-ci soit complète.

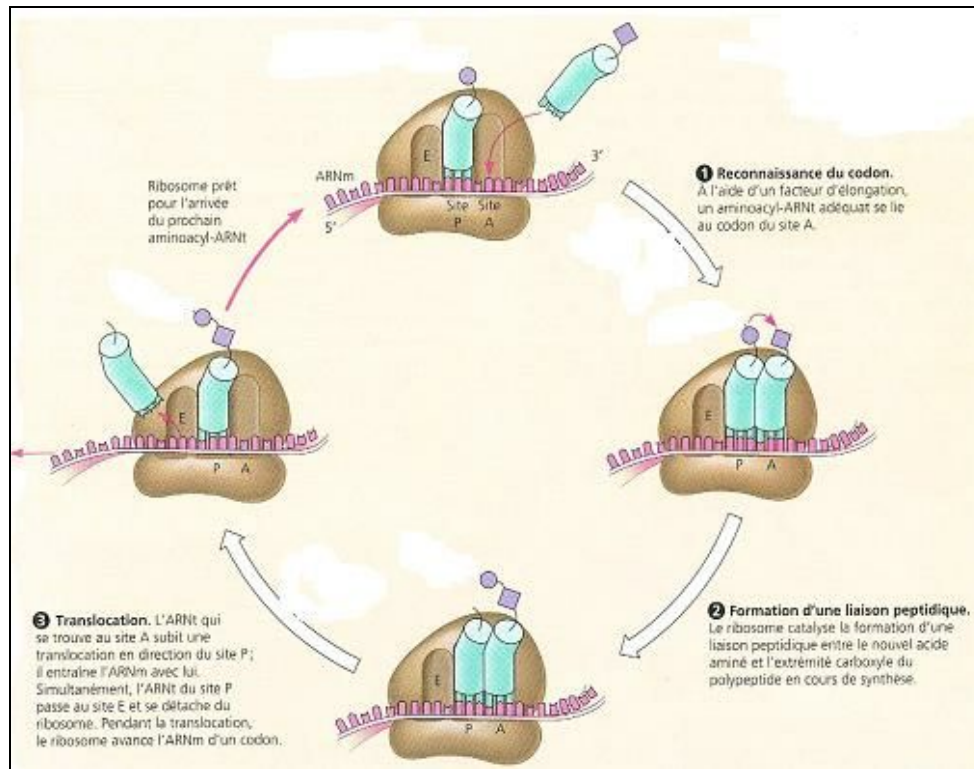


Figure 13 : Cycle d'élongation de la traduction

### 3) La terminaison (figure 14)

La dernière étape de la traduction est la terminaison. Les triplets de bases UAA, UAG et UGA de l'ARNm ne codent pas pour des acides aminés (ces codons n'ont pas d'anticodons complémentaires), mais servent de signal de fin de la traduction. L'élongation se poursuit jusqu'à ce que l'un de ces codons d'arrêt arrive au site A. Sous l'effet d'un **facteur de terminaison** les différentes structures se détachent. Le ribosome se dissocie en deux sous unités et la protéine formée peut ensuite être exportée de la cellule (réticulum endoplasmique, appareil de Golgi) ou restée dans la cellule.

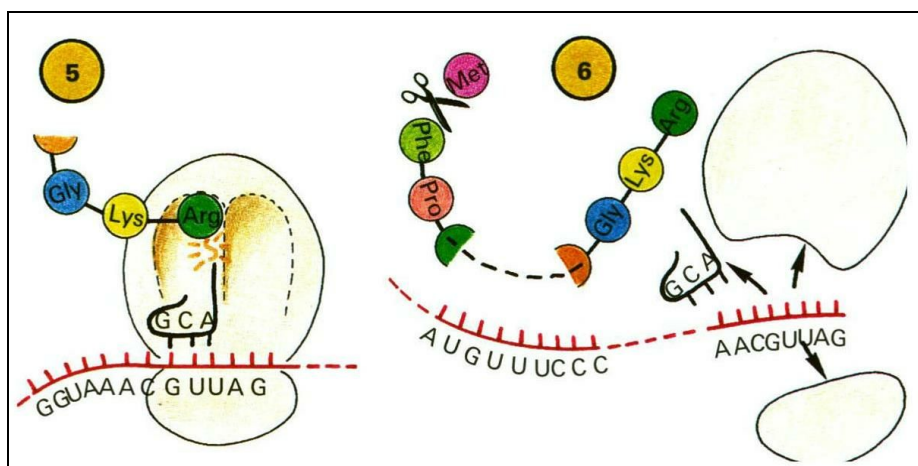


Figure 14 : La terminaison

Ces réactions sont possibles grâce aux mitochondries qui fournissent l'énergie nécessaire à l'établissement des liaisons peptidiques.

Différents ribosomes peuvent fonctionner en même temps pour traduire une seule molécule d'ARNm et produire presque simultanément plusieurs molécules de protéines. Un ribosome peut se combiner avec n'importe quel ARNm et n'importe quel ARNt et peut donc être utilisé pour la synthèse de différentes protéines. Il n'y a donc pas de ribosomes spécifiques pour chaque sorte de protéines produites par la cellule.

**Exercice :**

- a) Reprendre l'ARNm trouvé à l'exercice page 16 du dossier

-----

- b) Traduire l'ARNm en protéine à l'aide du *tableau 2* :

-----

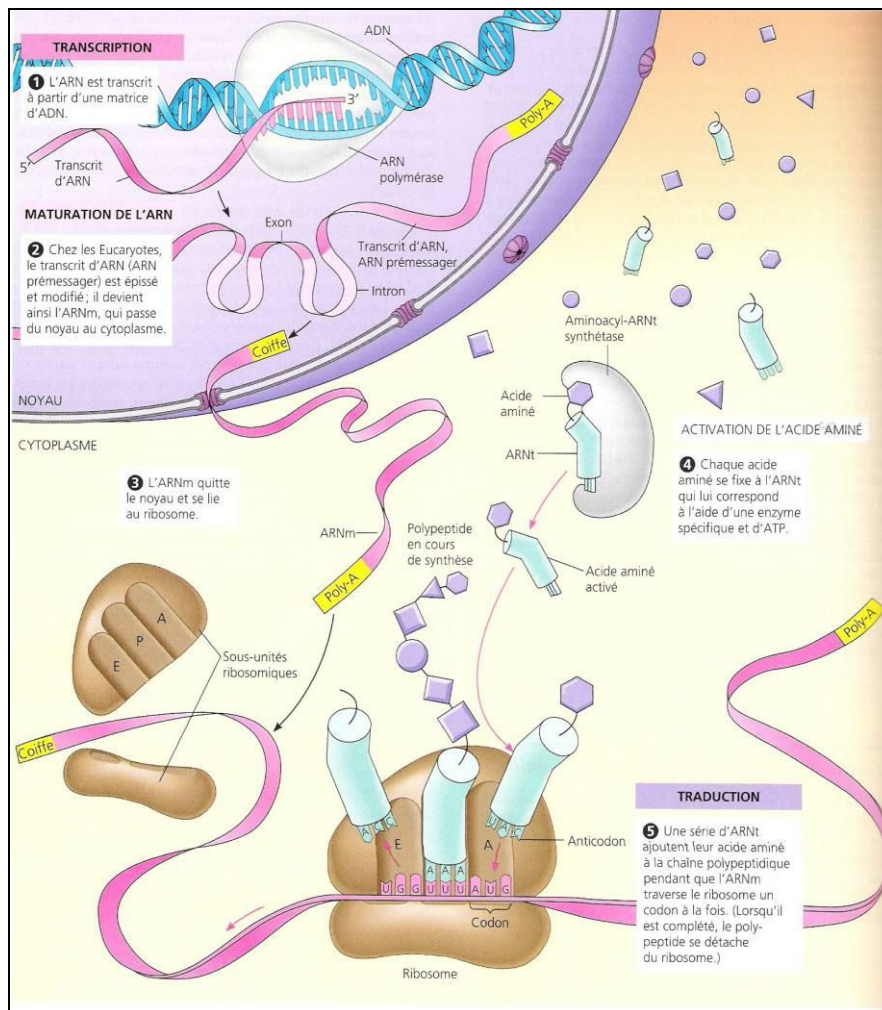


Figure 15 : Résumé transcription-traduction