

Chapitre 2 : l'expression du patrimoine génétique chez les eucaryotes

1. Une relation entre gènes et protéines

A. Les protéines sont des molécules séquencées

Les **protéines** sont des molécules organiques essentielles dans un organisme car elles remplissent diverses fonctions fondamentales (accélérateurs de réactions chimiques = les enzymes – rôle dans la vision : les pigments rétiens – rôle de soutien de la peau = le collagène – défense immunitaire = les anticorps – transport de dioxygène = l'hémoglobine etc). Elles sont formées par un **enchaînement d'acides aminés**.

De la structure primaire à la structure quaternaire

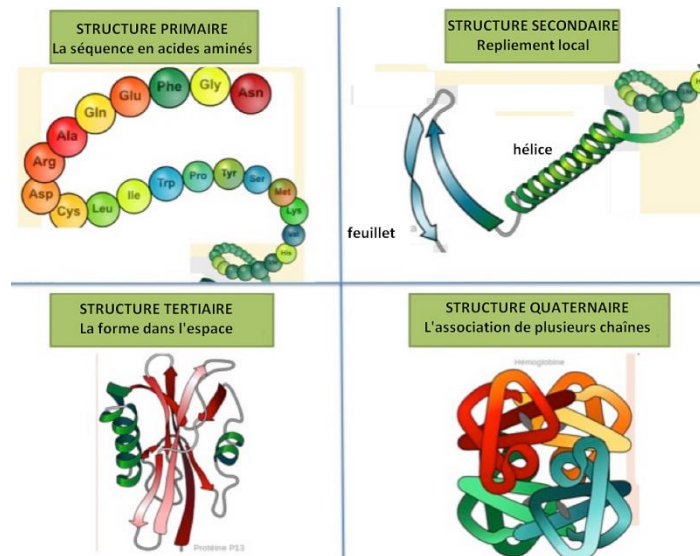
On appelle **structure primaire** : le nombre, la nature (il existe **20 sortes d'acides aminés**) et l'ordre d'enchaînement de ces acides aminés. Ces acides aminés créent des liaisons chimiques avec les acides aminés voisins si bien que la protéine se replie dans l'espace adoptant une **configuration spatiale** (c'est la **structure tertiaire**).

Parfois, deux ou plusieurs polypeptides s'associent entre eux pour former une « supra-protéine » (c'est la **structure quaternaire**).

Ex. : cas de l'hémoglobine : 4 chaînes polypeptidiques semblables 2 à 2 : 2 chaînes alpha, 2 chaînes bêta.

Cas de l'insuline : 2 chaînes polypeptidiques

La forme d'une protéine est capitale car c'est elle qui conditionne sa fonction.



B. Le cas d'une maladie génétique : la Drépanocytose

Activité 5 : relation ADN – protéine : exemple de la drépanocytose

	Phénotype macroscopique (clinique)	Phénotype cellulaire	Phénotype moléculaire	Génotype
Individu sain	Aucun symptôme	GR : disque biconcave – souple, déformable	HbA : 4 chaînes protéiques : - 2 chaînes α (142 aa) - 2 chaînes β (147 aa)	Gène de 444 paires de nucléotides Avec en 20 ^e position : A
Individu malade	Maux de tête – essoufflement – vertiges – problèmes osseux, oculaires, rénaux... mort	GR : en forme de faucille (=falciforme) – indéformable, à durée de vie plus courte	HbS : Même structure IV ^{aire} mais sur les chaînes β en 6 ^e position : VAL6 au lieu de GLU6	Gène de 444 paires de nucléotides Avec en 20 ^e position : T

Pour information : Le remplacement en 6^e position dans la chaîne protéique d'une VAL (valine) à la place de la GLU (acide glutamique) entraîne la polymérisation des HbS qui forment de longues fibres à l'intérieur du cytoplasme des GR. Ce sont ces fibres qui expliquent la forme, la rigidité et la faible durée de vie des globules rouges.

L'expression du génotype conduit à la synthèse d'une protéine, c'est le phénotype moléculaire et c'est lui qui est déterminant pour la réalisation du phénotype cellulaire et macroscopique.

L'ordre dans lequel sont enchaînés les acides aminés (séquence primaire) dépend de l'ordre dans lequel sont enchaînés les nucléotides (séquence nucléotidique), on dit qu'UN GÈNE CODE POUR UNE PROTÉINE.

2. La synthèse des protéines

A. La transcription

Activité 6 : la transcription

La transcription (ou synthèse de l'ARN) est le processus par lequel l'information contenue dans la séquence des nucléotides de l'ADN est transférée en « langage » ARN.

a) Localisation

La transcription a lieu dans le noyau. En effet, l'information génétique est portée par des chromosomes (molécules d'ADN) dont la taille interdit le passage par les pores nucléaires. Or la synthèse protéique c'est-à-dire l'enchaînement des acides aminés, a lieu dans le cytoplasme. Il faut donc admettre le transfert de l'information du noyau vers le cytoplasme par des molécules « messagers » : ce sont les ARNm (m = messenger), l'acide ribonucléique.

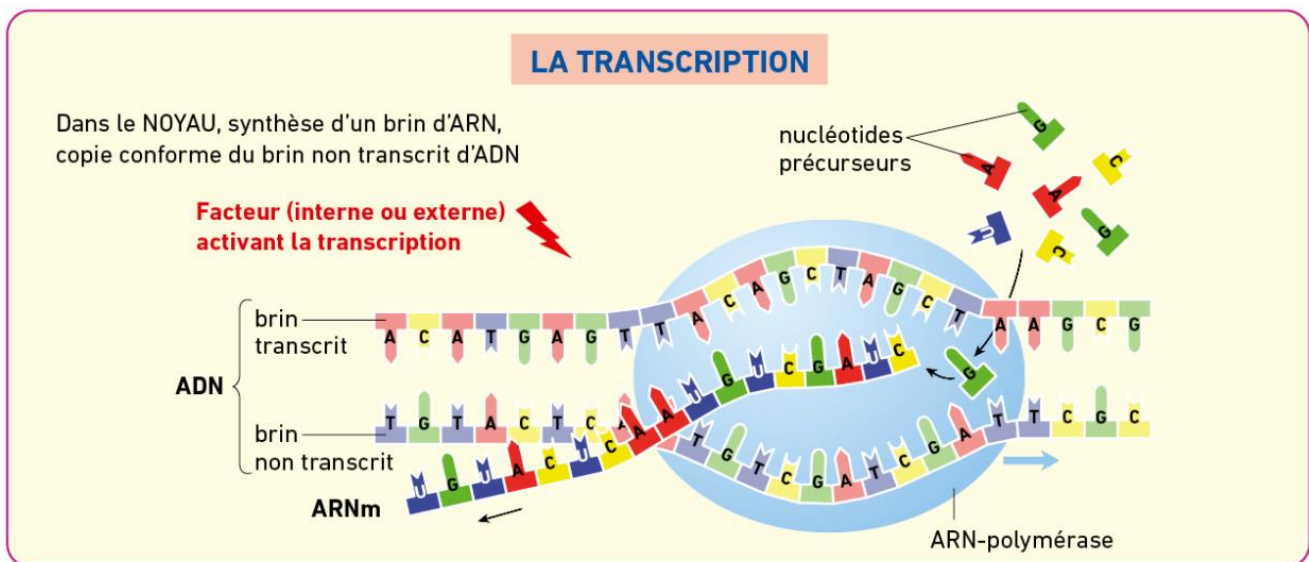
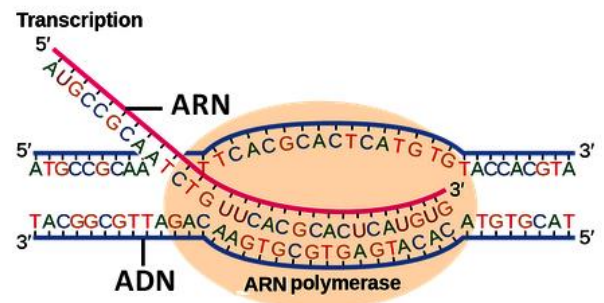
Un complexe enzymatique constitué en particulier d'ARN polymérase permet de transcrire le « langage ADN » en « langage ARN » mais ayant pour particularité :

- d'être **monocaténaire** → ce qui suppose qu'un seul brin de l'ADN sert de modèle à la transcription.
- de correspondre à un enchaînement de nucléotides MAIS dans lequel l'**uracile** a remplacé la thymine.
- Ribose remplace le désoxyribose.

b) Mécanisme

Après ouverture et le déroulement d'une petite portion de la double hélice d'ADN correspond au gène à transcrire, une enzyme, l'ARN polymérase, se déplace le long d'un des brins de l'ADN qui sert de modèle, c'est le **brin transcrit** (ou **brin non codant** / l'autre brin étant le **brin non transcrit** ou **codant** (dans le sens qui donne le code (traduction anglaise prêtant à confusion)).

Au fur et à mesure de sa progression le long de l'ADN, elle incorpore des nucléotides présents dans le nucléoplasme selon la complémentarité avec les nucléotides de l'ADN (G – C ; A – U) et les enchaîne.



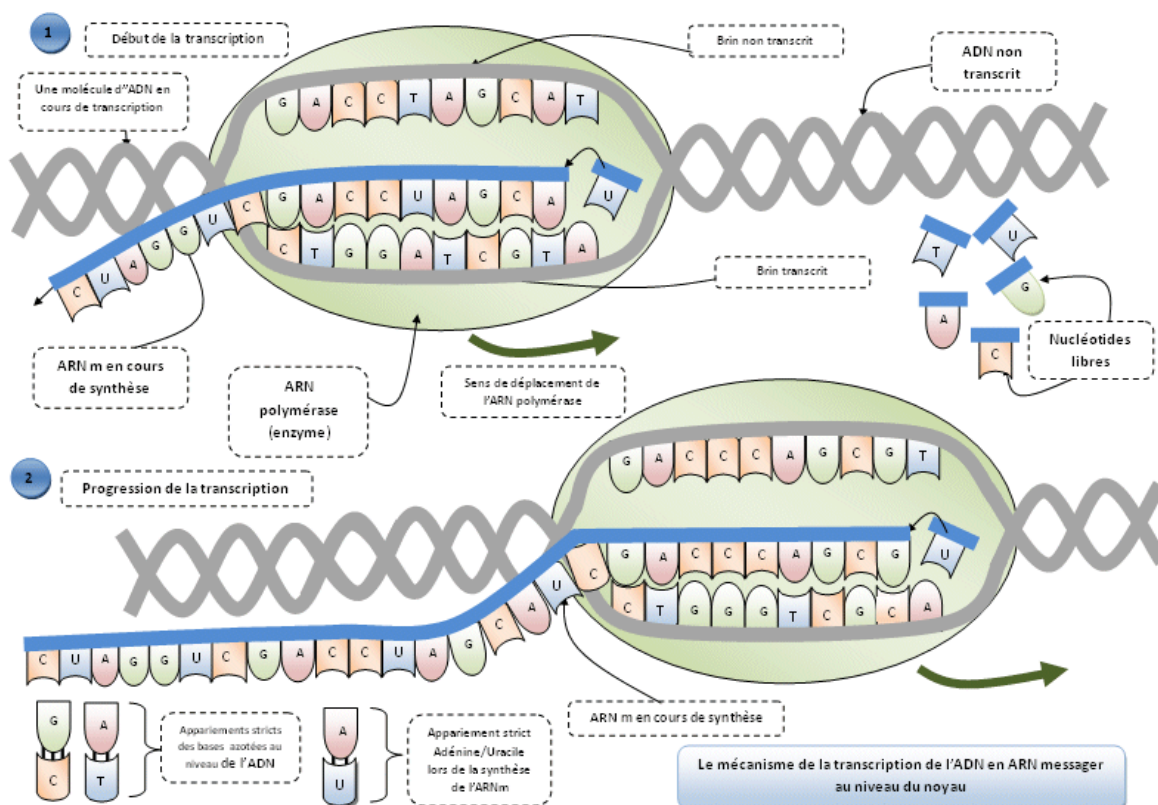
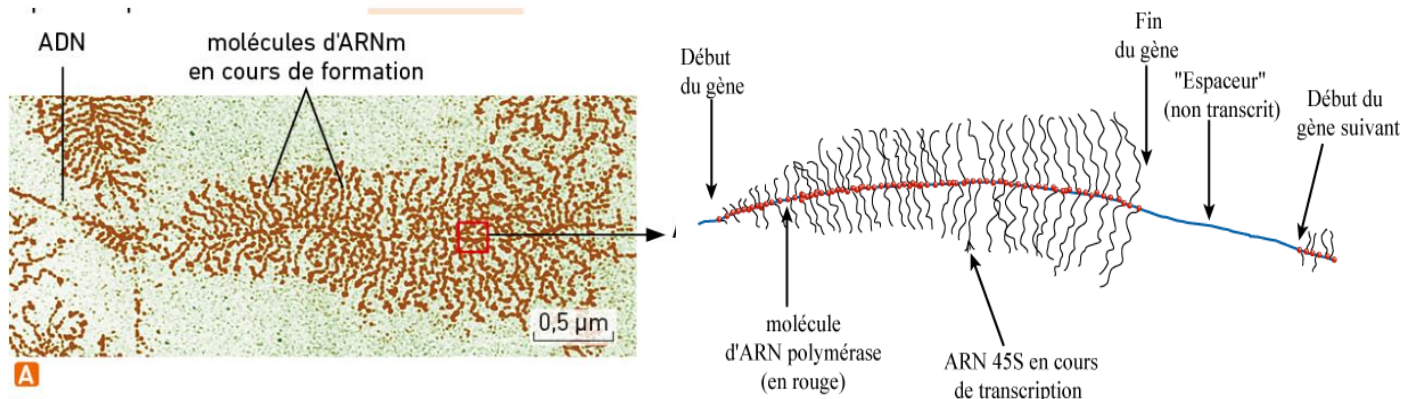
Comparons la réplication à la transcription

	Réplication	Transcription
Modèle	ADN : - les 2 chaînes - Toute la molécule	ADN - Une seule chaîne - Des portions seulement (= unité de transcription = « gènes »)
Produit	2 molécules d'ADN identiques	1 ARN (ARN-pré-messager ou « Transcrit primaire »)
Nucléotides polymérisés	A, T, G, C	A, U, G, C
Enzymes impliquées	ADN polymérase + autres	ARN polymérase + autres
Vitesse	50 à 100 nucléotides/sec.	30 nucléotides/sec.
Démarrage signalé par...	Origine de réplication (œil)	Séquence « promoteur » en début de gène
Fin signalée par...	Rien (molécule entièrement répliquée)	Signal de terminaison en fin de gène.

c) un système d'amplification

Constat : Au cours de la transcription, l'ADN paraît hérissé de multiples filaments d'ARN (chromosome en « écouvillon » (doc.A)) et plus on se rapproche de la fin du gène, plus la taille des ARNm augmente.

Explication : En effet, **plusieurs ARN polymérases** se succèdent le long du gène à transcrire et entame « à la chaîne » la transcription d'ARNm strictement identiques entre eux qui constituent autant de copies du même gène. Ce système d'amplification permettra que **plusieurs centaines de copies de la même protéine** seront ensuite fabriquées dans le cytoplasme au cours de l'étape suivante.

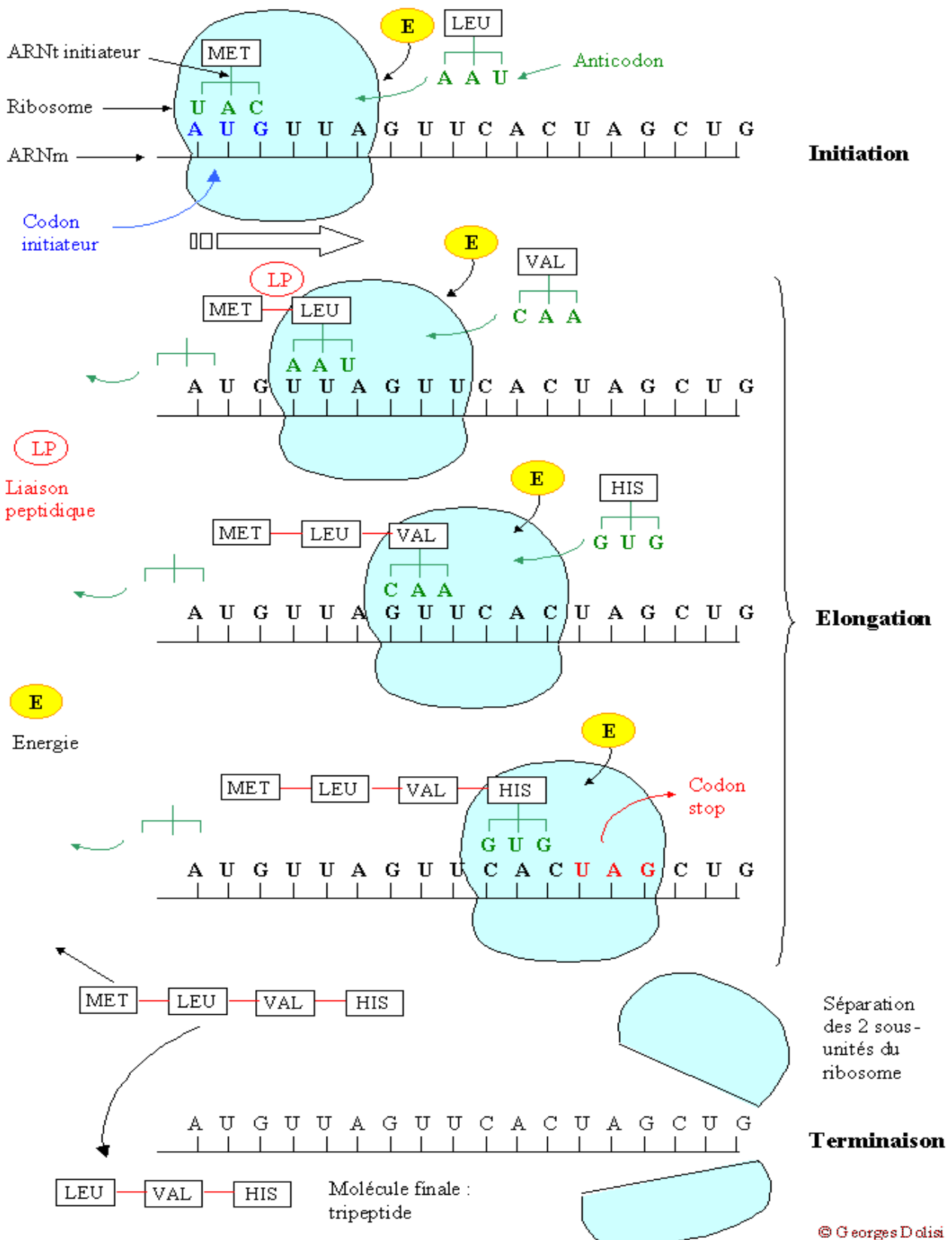


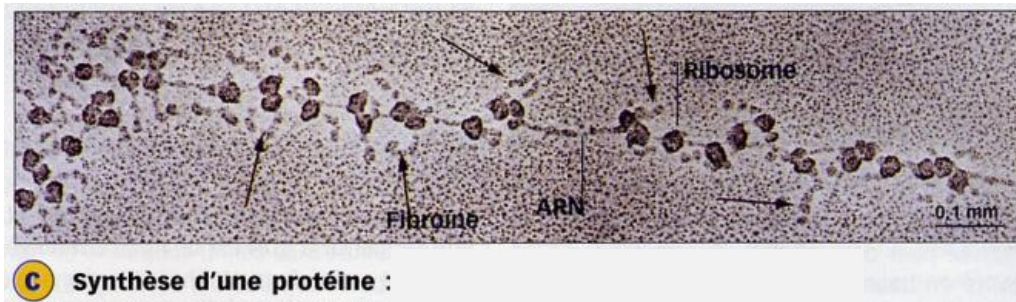
B. La traduction

Activité 7 : la traduction

La traduction a lieu dans le cytoplasme

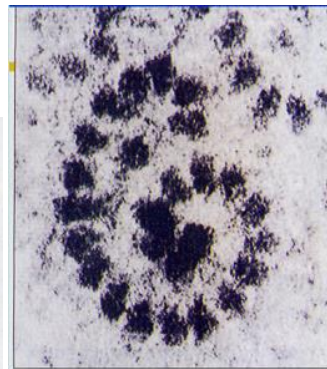
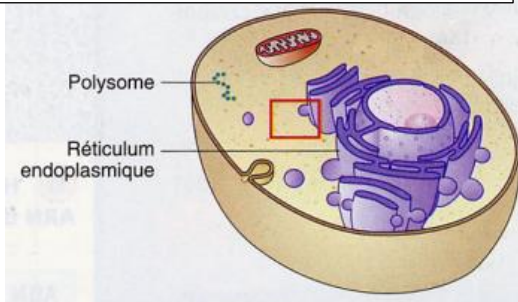
- Elle fait intervenir d'autres ARN, les **ribosomes** qui, en se fixant sur l'ARNm, associent, selon le **code à triplet**, les **acides aminés** entre eux.
- Toute traduction commence par un **codon initiateur** (AUG → MET) et se termine au niveau d'un **codon stop**.



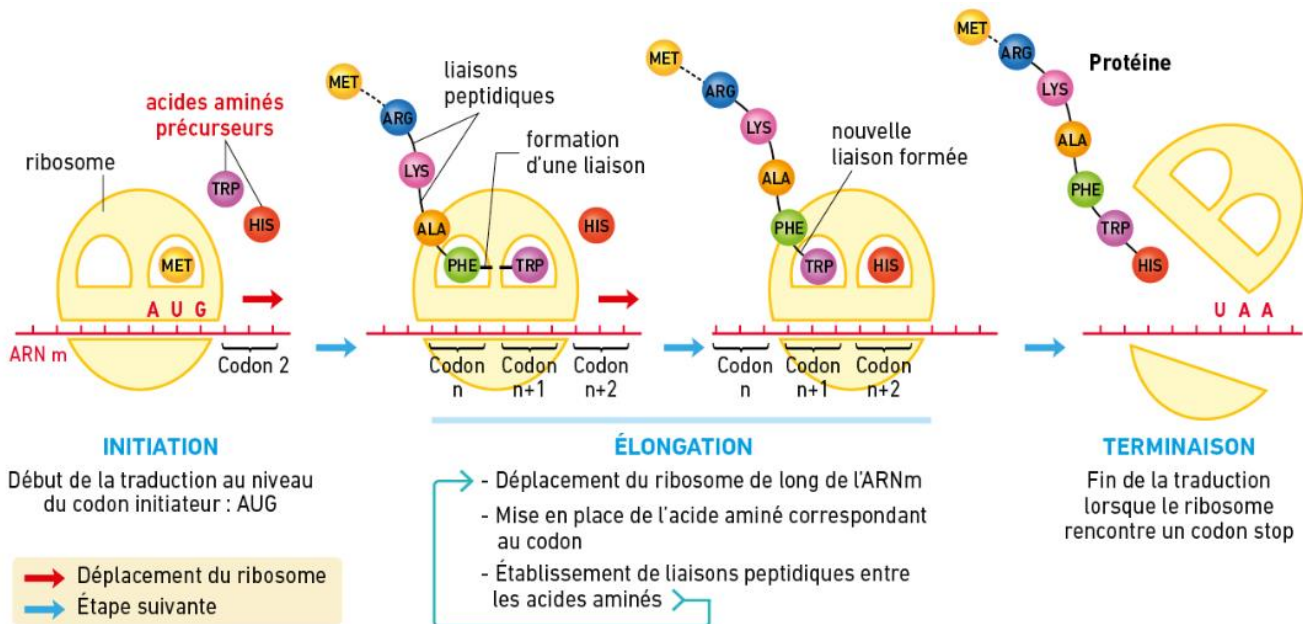
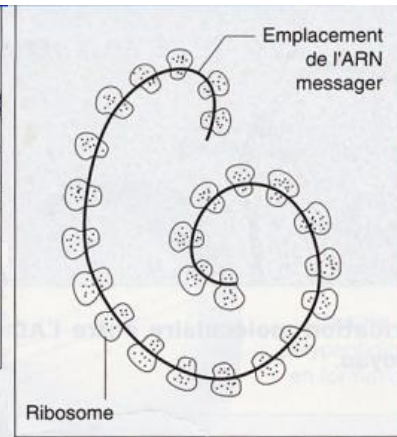


C Synthèse d'une protéine :

Polysome : Un polysome ou polyribosome est un **ensemble de ribosomes reliés entre eux par un ARN messager**. Cet ensemble a l'aspect d'un collier de perles.



b Électronographie d'un polysome (MET x 180 000).



C Le mécanisme de la traduction.

Là encore, il existe un système d'amplification : plusieurs ribosomes se déplacent le long de l'ARNm et permettent ainsi la synthèse de très nombreuses copies de la même protéine.

1 gène donne des milliers d'ARNm par la transcription
1 ARNm donne des milliers de copies de la même protéine

Le **code génétique** (voir Nirenberg [1961]) est ce système de correspondance entre **un triplet de nucléotides** ou **codon** et **un acide aminé**. Ce code est :

- **universel** (à de rares exceptions)
- **non ambiguë** (= univoque) : **un codon ne désigne qu'un acide aminé**
- **redondant** : **un acide aminé peut être désigné par plusieurs codons**.
- 3 triplets ne codent pour aucun acide aminé et sont qualifiés de **codon-stop**, ils indiquent la fin de la traduction.

Le code génétique

		2 ^e nucléotide					
		U	C	A	G		
1 ^{er} nucléotide	U	UUU	UCU	UAU	UGU	U	3 ^e nucléotide
		UUC	UCC	UAC	UGC	C	
		UUA	UCA	UAA	UGA	A	
		UUG	UCG	UAG	UGG	G	
	C	CUU	CCU	CAU	CGU	U	
		CUC	CCC	CAC	CGC	C	
		CUA	CCA	CAA	CGA	A	
		CUG	CCG	CAG	CGG	G	
	A	AUU	ACU	AAU	AGU	U	
		AUC	ACC	AAC	AGC	C	
		AUA	ACA	AAA	AGA	A	
		AUG	ACG	AAG	AGG	G	
G	GUU	GCU	GAU	GGU	U		
	GUC	GCC	GAC	GGC	C		
	GUA	GCA	GAA	GGA	A		
	GUG	GCG	GAG	GGG	G		

Le code génétique.

C. Les modifications des ARN après transcription : la maturation

Activité 8 : la maturation de l'ARN

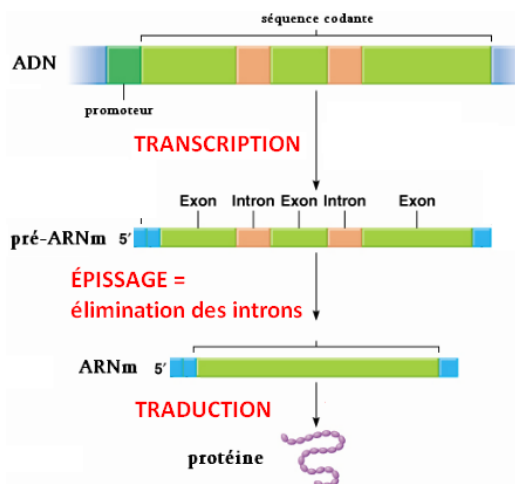
a. Les gènes des eucaryotes sont des gènes « morcelés »

On détermine la taille des ARNm rendus radioactifs produits soit dans le noyau, soit dans le cytoplasme.

On s'aperçoit que, au cours du temps, les ARNm parvenus dans le cytoplasme sont plus courts que ceux produits dans le noyau.

En effet, les gènes des eucaryotes sont pour la plupart des **gènes discontinus, morcelés, « en mosaïque »** : la portion codante d'un gène (= une unité de transcription) est constituée de régions qui vont participer à la formation de la chaîne polypeptidique, on les appelle **exons** et de régions qui seront transcrites mais non traduites, qui ne participeront donc pas à la formation de la protéine et qui seront donc **éliminées** après la transcription (**introns**).

A la suite de la transcription, l'ARN a une séquence strictement complémentaire de la portion codante du gène : c'est un **pré-ARNm**. Mais ce dernier va subir, dans le noyau, une **maturation** qui consiste à **l'élimination des introns** (c'est **l'épissage**) et **collage** bout à bout des **exons** : il se forme alors l'**ARNm** qui passe dans le cytoplasme.



Remarque 1 : Chez les vertébrés, la plupart des **exons** a une longueur comprise entre **50 et 400 bases**, certains peuvent en faire plusieurs milliers. De même, la longueur des **introns** varie entre quelques dizaines et plusieurs dizaines de milliers de bases, et peut être très variable même entre espèces proches. **Le nombre d'exons et d'introns peut varier d'une espèce à l'autre pour des protéines homologues.**

Remarque 2 : **98% du génome humain n'est pas codant.** Même en ajoutant à cela les séquences non codantes, mais importantes pour l'expression des gènes, ainsi que les séquences ribosomiques, **plus de 90% du génome n'a pas de fonction connue**, et une grande partie est probablement composée de débris d'éléments transposables (notion de

« transposons »), et de microsatellites. Les modifications se produisant dans ces régions ajoutent à la variabilité (ce sont sur ces régions qui intéressent la police scientifique lors de la réalisation des « empreintes génétiques »), mais n'ont pas d'influence sur le phénotype et la valeur sélective.

Selon les espèces, le morcellement des gènes est plus ou moins important.

b. L'épissage alternatif

Un même **pré-ARNm** peut donner des ARNm différents au cours de la maturation et donc des **protéines différentes**. En effet, selon le contexte, certains exons peuvent ou pas être retenus pour former l'ARNm.

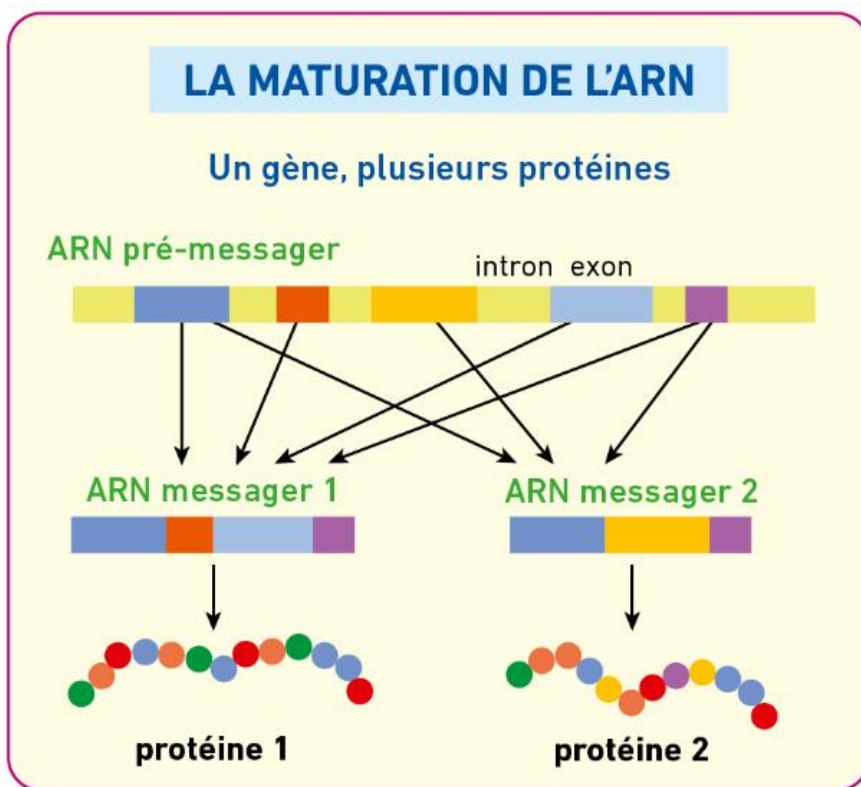
Ainsi avec un « minimum » de gènes (environ 20 000 chez l'Homme), l'organisme peut coder pour un « maximum » de protéines différentes et cette capacité diffère en fonction de plusieurs facteurs :

- du type cellulaire (cellules musculaire, neurone, fibroblaste...): dans l'exemple traité, la protéine [la tropomyosine] conserve le même rôle – celui de participer au cytosquelette – mais sa structure diffère légèrement selon le type cellulaire.
- du cycle de vie de l'individu car en effet, des gènes peuvent être actifs à un moment donné de la vie et devenir inactifs (ou inversement)

(distinction génome / protéome)

Remarque 1 : cette découverte a amené les scientifiques à reconsidérer la notion de « 1 gène – 1 protéine ».

Remarque 2 : plus que le nombre total de gènes, le processus d'épissage a dû jouer un rôle important dans l'évolution des espèces.



D. La régulation de l'expression des gènes

Une cellule n'exprime qu'une partie des gènes qu'elle possède : en exprimant des gènes différents, les cellules se **spécialisent**. L'activité des gènes d'une cellule est donc régulée sous l'influence de facteurs internes, notamment au cours du développement, ou externes, en réponse aux conditions environnementales.

Le **phénotype**, ensemble des caractéristiques d'un être vivant, peut-être défini à chacun des niveaux d'organisation du vivant, de la molécule à l'organisme. Il résulte en partie de l'expression de son **génotype**, c'est-à-dire de son patrimoine génétique, transmis de générations en générations. L'ensemble des ARN et des protéines résultant de la transcription et de la traduction de l'ADN constitue le **phénotype moléculaire**.

C'est l'activité de ces molécules produites à partir de l'ADN qui va, en interaction avec de multiples facteurs, déterminer le phénotype cellulaire, et en définitive l'ensemble du phénotype macroscopique observable à l'échelle de l'organisme.

Du génotype au phénotype

