

L'ordre et la nature des acides aminés (ou séquence) d'un polypeptide dépend de la séquence des nucléotides de l'ADN du gène qui le code. Une mutation, peut entraîner une modification de la séquence des acides aminés qui, à son tour, pourra entraîner une modification de la conformation de la protéine et, généralement, une modification de sa fonction. La séquence des nucléotides de la molécule d'ADN représente donc une information génétique qui dicte la séquence des acides aminés des protéines. Seules les protéines sont l'expression directe de l'information génétique. Ce n'est pas le cas des autres molécules du vivant (lipides et glucides) même si leur métabolisme est commandé par l'information génétique.

Chapitre 3

L'expression du patrimoine
génétique

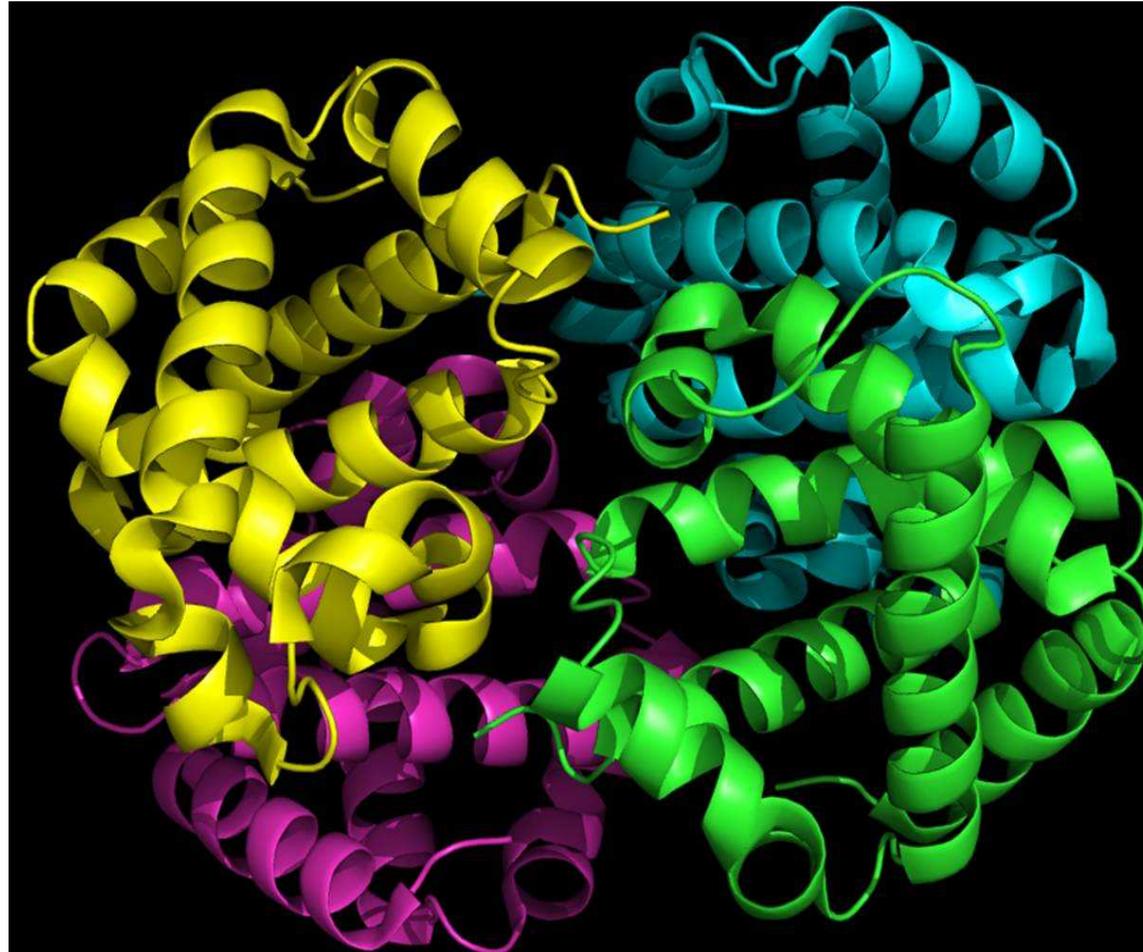
Comment produit on une
protéine à partir d'un gène?

Ou comment s'exprime l'information
génétique ?

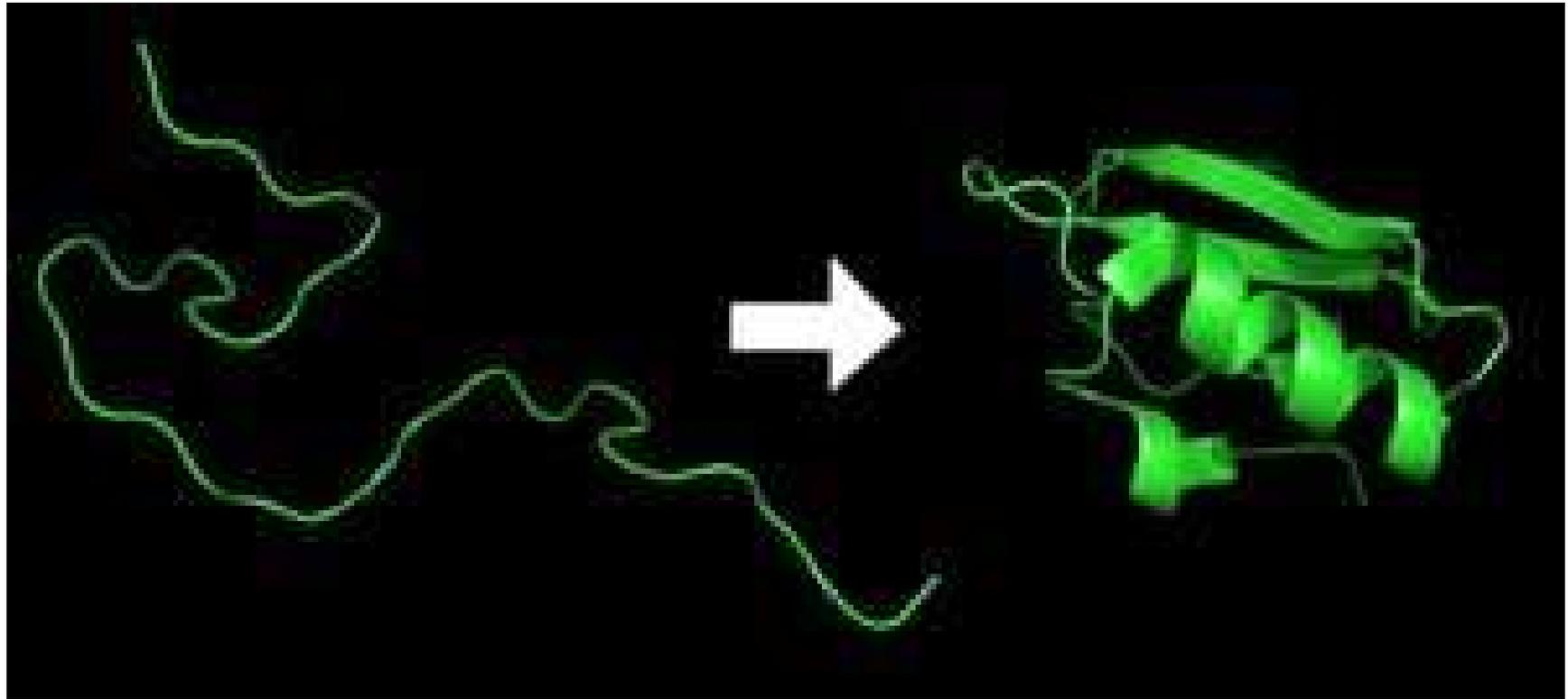
Un gène est une unité d'information, contenue dans l'ADN et qui gouverne la synthèse d'une protéine par une cellule.

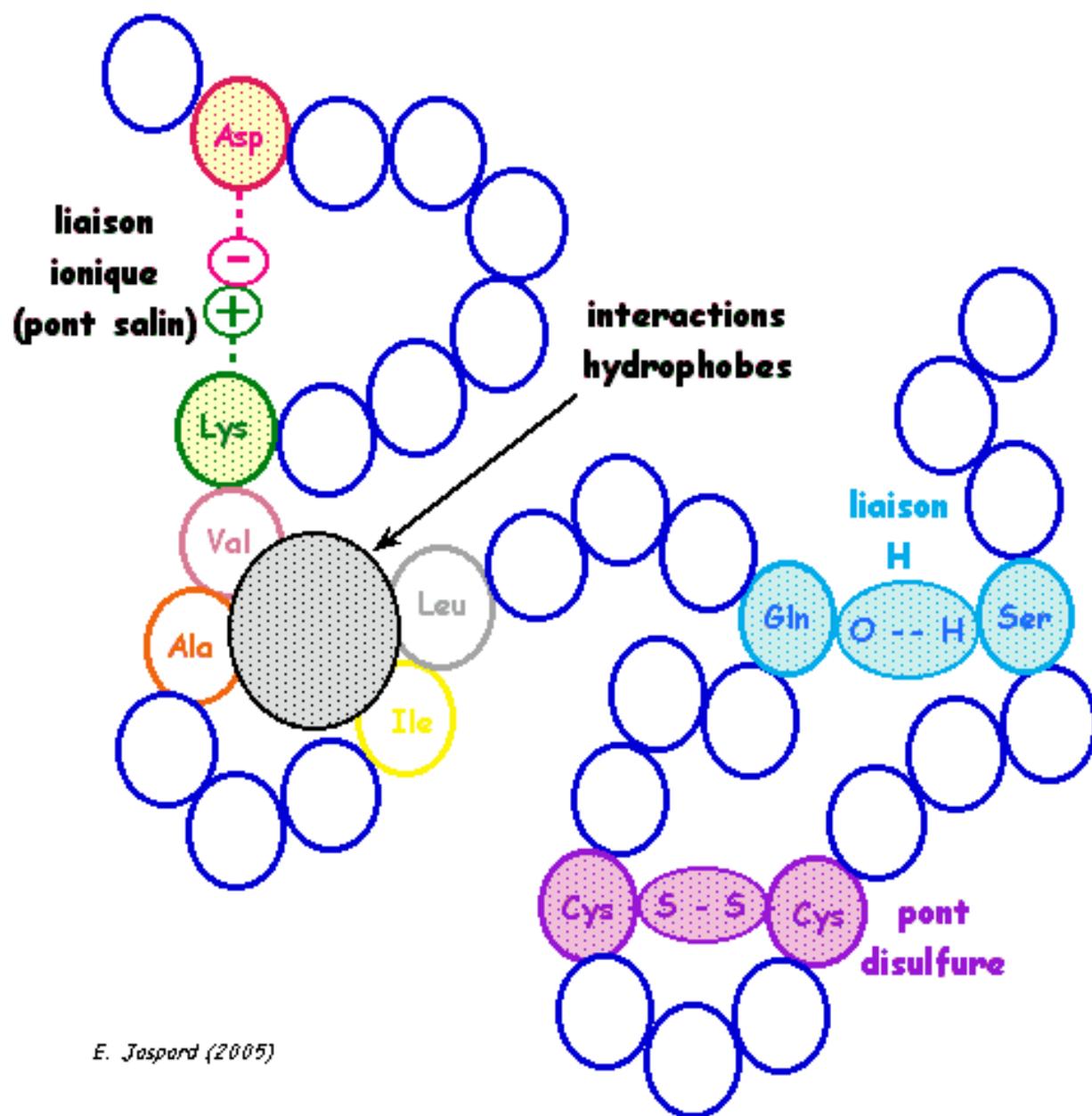
Une protéine est une molécule du vivant formée d'une ou plusieurs chaînes (ou séquences) d'acides aminés. On appelle peptide un polymère de quelques acides aminés (dipeptide formé de deux acides aminés, tripeptide de trois acides aminés, etc.) et polypeptide un polymère formé d'un grand nombre d'acides aminés (employé très souvent comme synonyme de protéine).

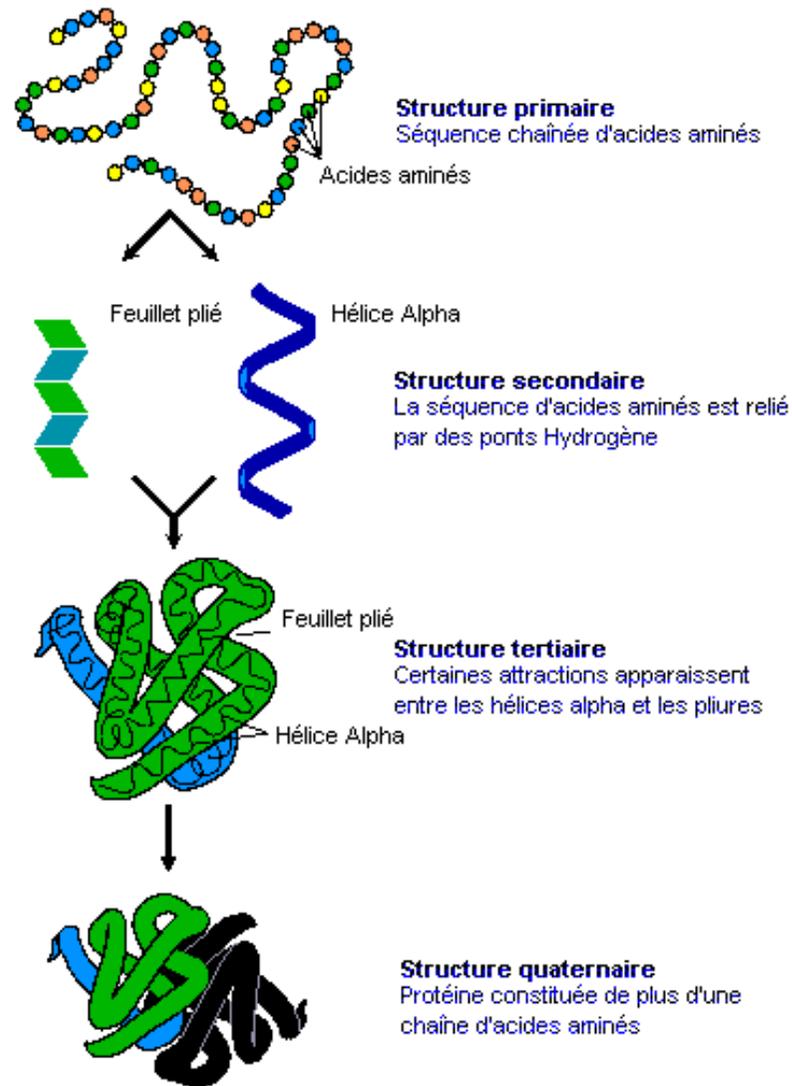
Ex : L'hémoglobine



En outre, des liaisons faibles peuvent s'établir entre des acides aminés distants sur une même chaîne, ce qui entraîne des repliements, ou appartenant à des chaînes différentes, ce qui lie les chaînes entre elles. La protéine fonctionnelle a donc une structure tridimensionnelle, ou conformation spécifique.





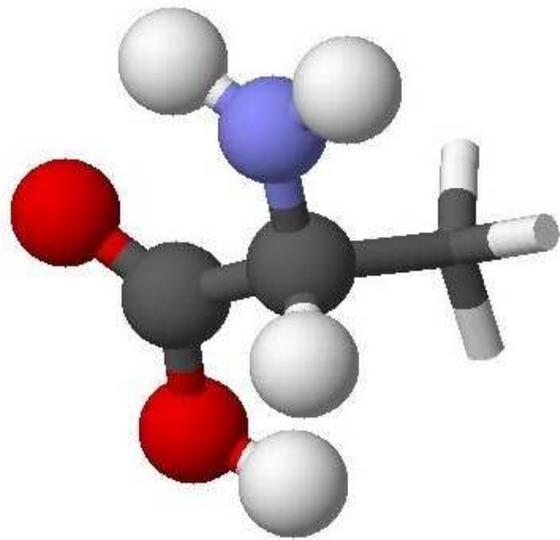


Il existe vingt acides aminés naturels qui ne diffèrent que par la nature de leur radical R.

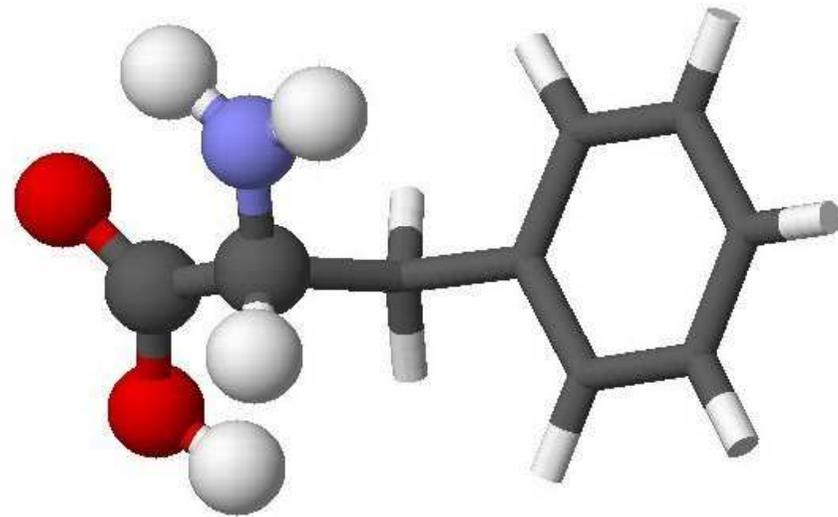
Les 20 acides aminés

Acide glutamique	Glu	E
Acide aspartique	Asp	D
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Cystéine	Cys	C
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I

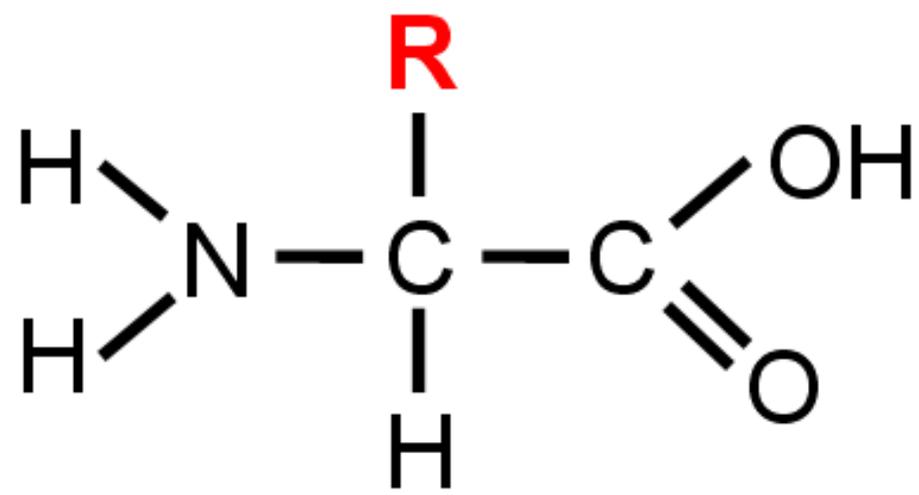
Leucine	Leu	L
Lysine	Lys	K
Méthionine	Met	M
Phénylalanine	Phe	F
Proline	Pro	P
Sérine	Ser	S
Thréonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V



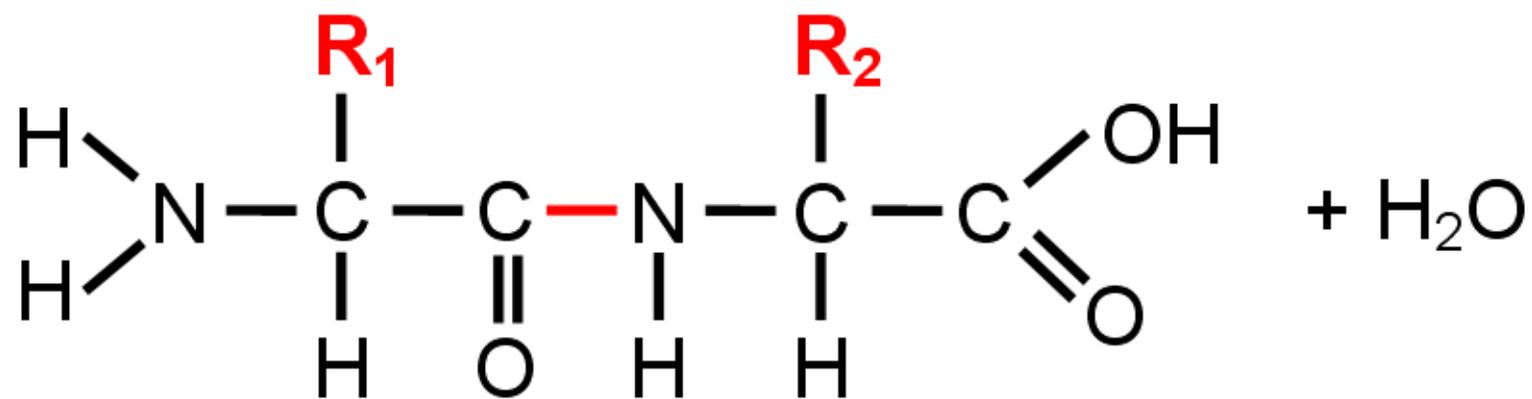
Alanine



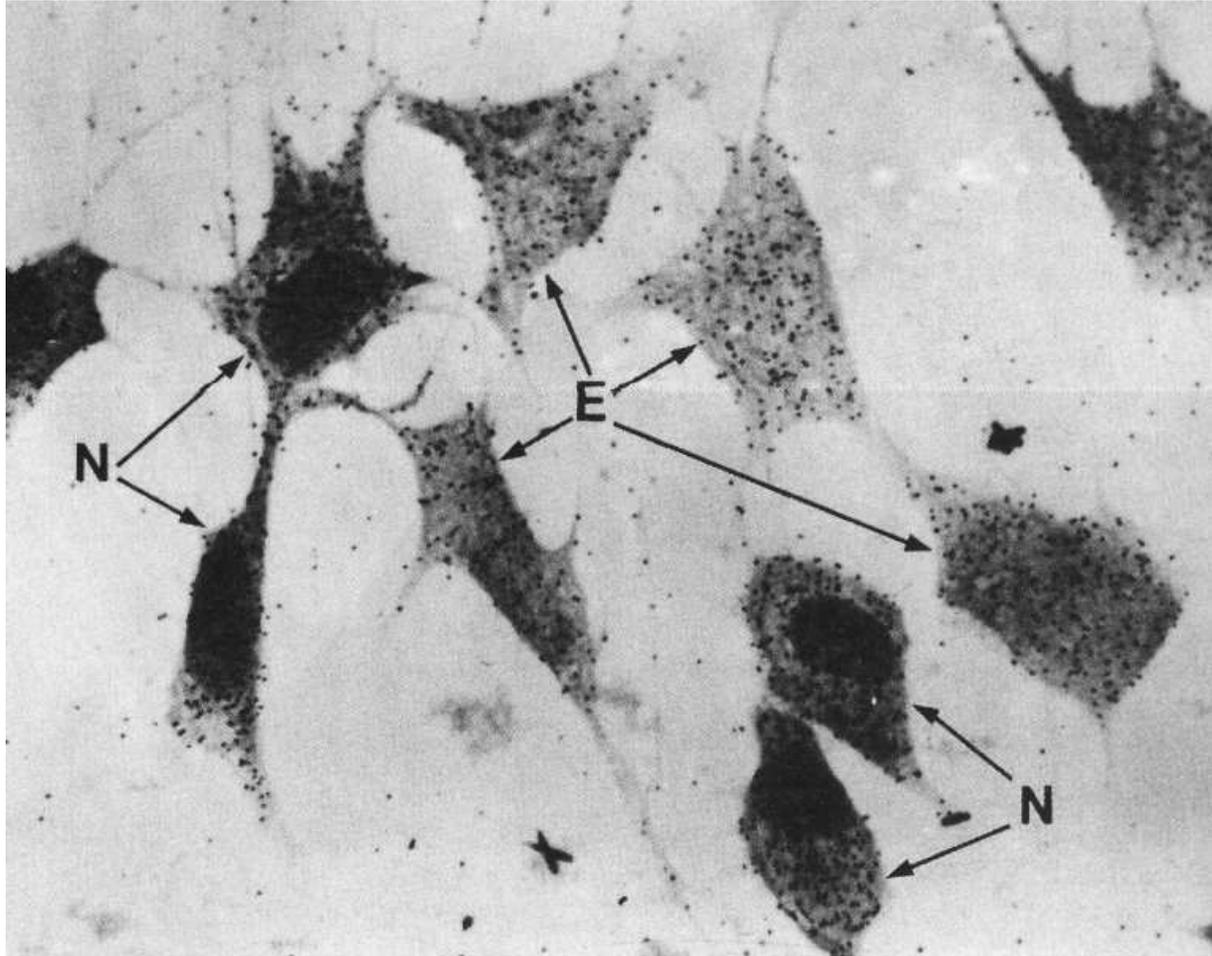
Phénylalanine

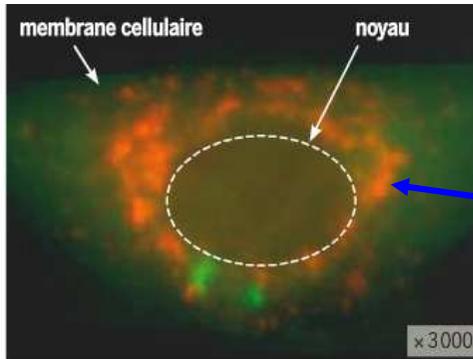


- Il existe 20 acides aminés naturels différents (1). Par commodité on les désigne par une abréviation de trois lettres ou par un symbole constitué d'une lettre majuscule. Ils possèdent tous un motif commun : CH + groupe carboxyle COOH + groupe amine NH₂ (3 et en boules sur 2) et diffèrent par un radical R (en bâtonnets sur 2). Deux acides aminés successifs sont reliés par une liaison forte : la liaison peptidique (3).



Des cellules animales sont cultivées sur un milieu contenant un **acide aminé marqué**. Le **noyau (N)** de certaines cellules a été enlevé **(E)** quelques minutes avant la mise en culture. On réalise ensuite une autoradiographie. Alors que l'information génétique se trouve dans le noyau, on constate que le noyau n'est pas indispensable à la synthèse protéique (= protéosynthèse) qui a lieu dans le cytoplasme.

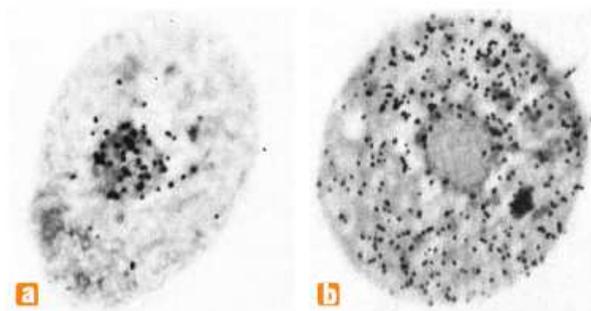




Bordas p. 54

La synthèse des protéines se déroule dans le cytoplasme...
(en rouge, protéines en cours de synthèse)

...alors que l'information génétique se trouve dans le noyau.



Bordas p. 54

Il existe un **intermédiaire** entre le noyau et le cytoplasme : l'ARN messager ou **ARNm**.

I. La transcription et la maturation dans le noyau

A. Les caractéristiques de la transcription (ADN- ARNmessenger)



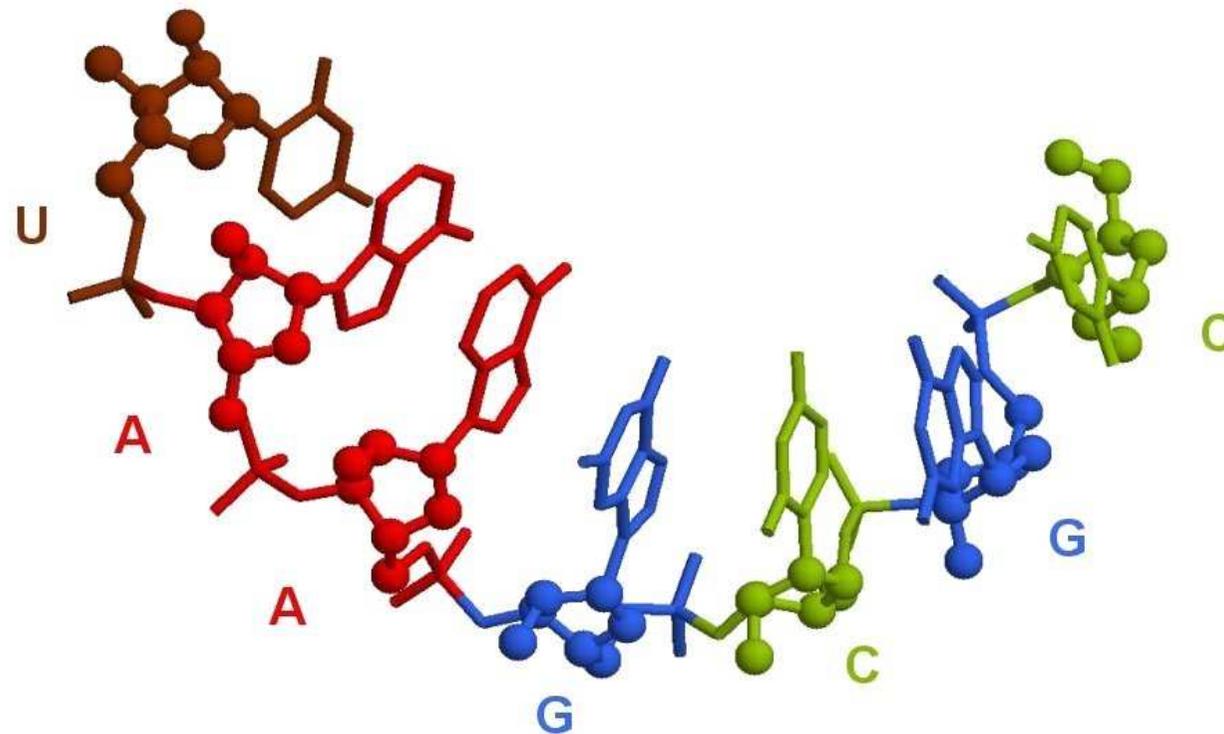
1. Deux types d'acides nucléiques

Le vert de méthyle colore l'ADN en vert et la pyronine colore l'acide ribonucléique, ou ARN, en rouge.

Contrôle.

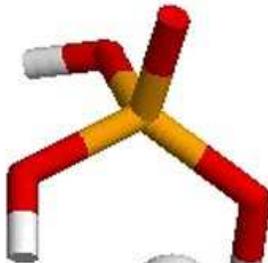
- Si on traite une coupe par la ribonucléase, qui détruit (dépolymérise) l'acide ribonucléique, la coloration vert de méthyle-pyronine colore simplement l'ADN restant dans le noyau en vert.
- Si on traite une coupe par la désoxyribonucléase, qui détruit (dépolymérise) l'acide désoxyribonucléique, la coloration vert de méthyle-pyronine colore simplement l'ARN restant dans le noyau en rouge.

L'ARN (1) est une molécule constituée d'une séquence n'excédant pas quelques milliers de ribonucléotides, ne comportant qu'un seul brin et de courte durée de « vie »



Chaque ribonucléotide (2) est formé par l'association d'un acide phosphorique (*représenté en bâtonnets*) qui relie les nucléotides entre eux, d'un ribose (glucide, *représenté en boules et bâtonnets*) et d'une base azotée (*représentée en bâtonnets*) : *guanine (G), cytosine (C), adénine (A) ou uracile (U).*

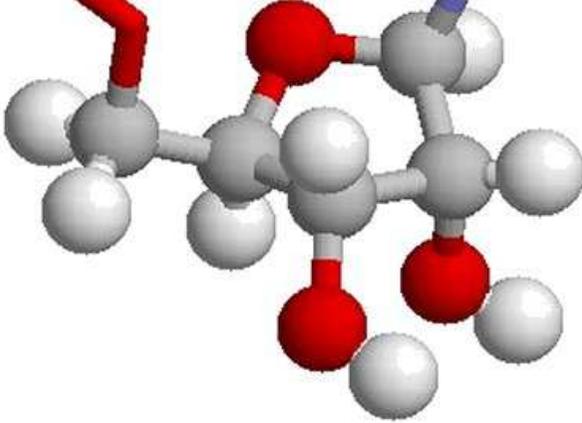
Acide phosphorique



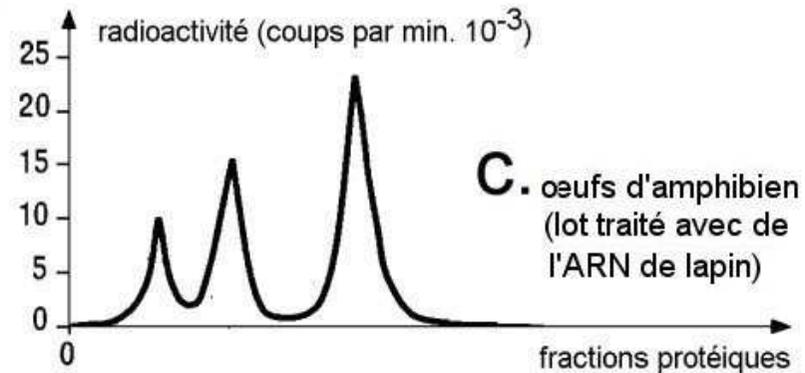
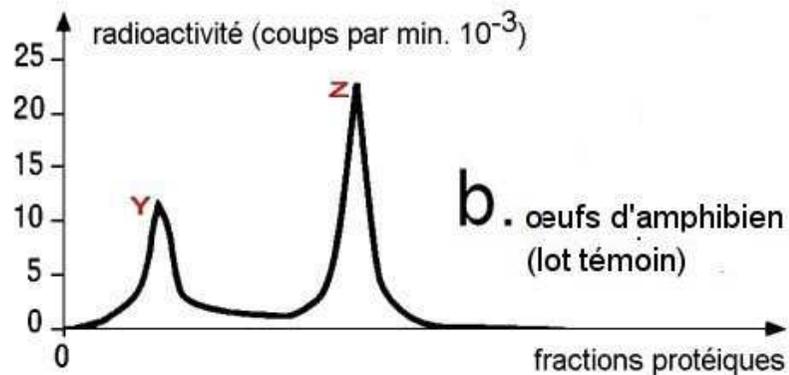
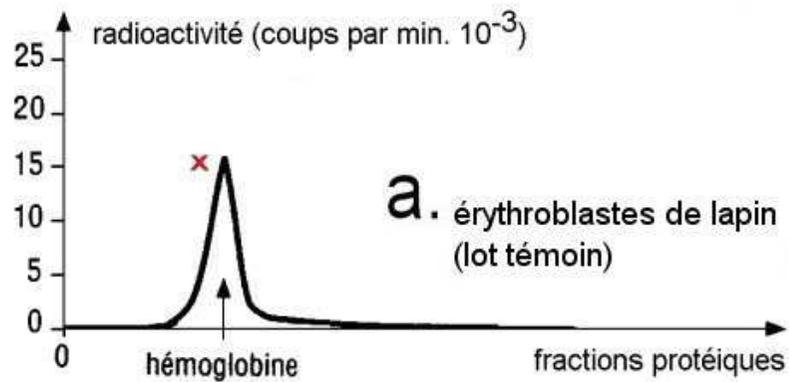
Uracile



Ribose

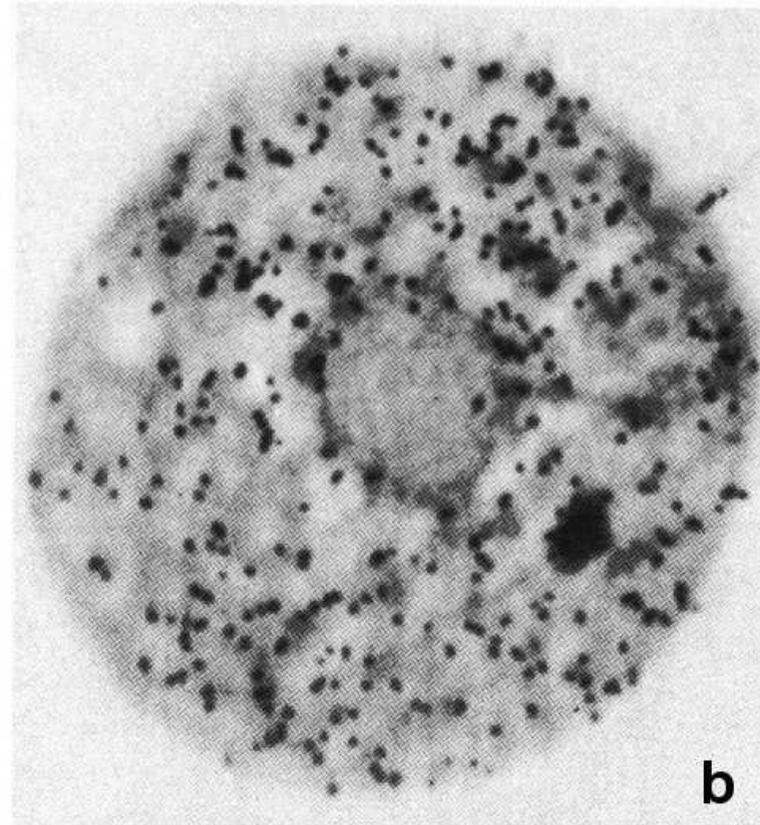
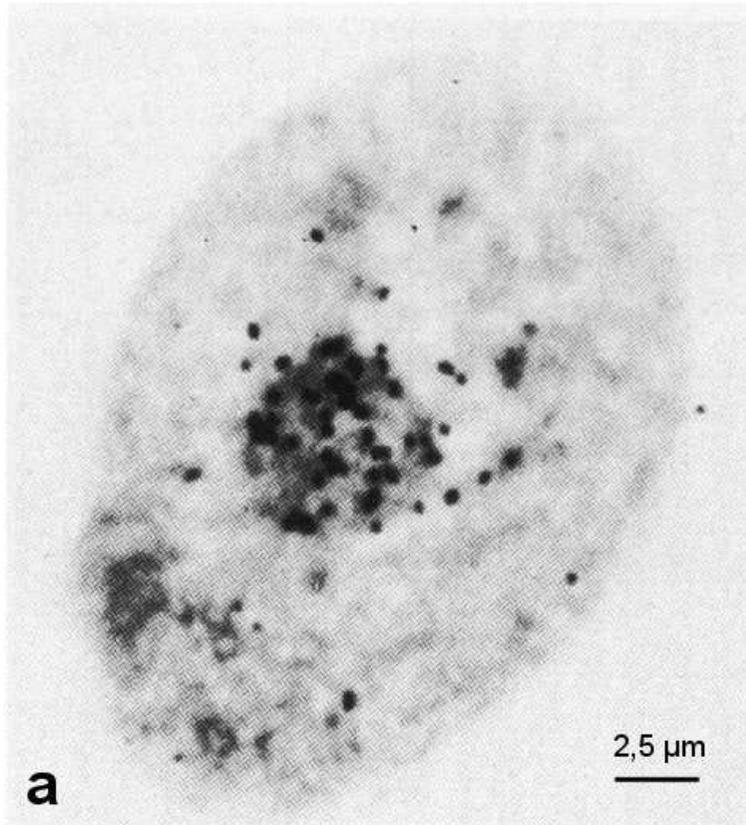


2. l'ARN messenger est un acide nucléique qui peut sortir du noyau



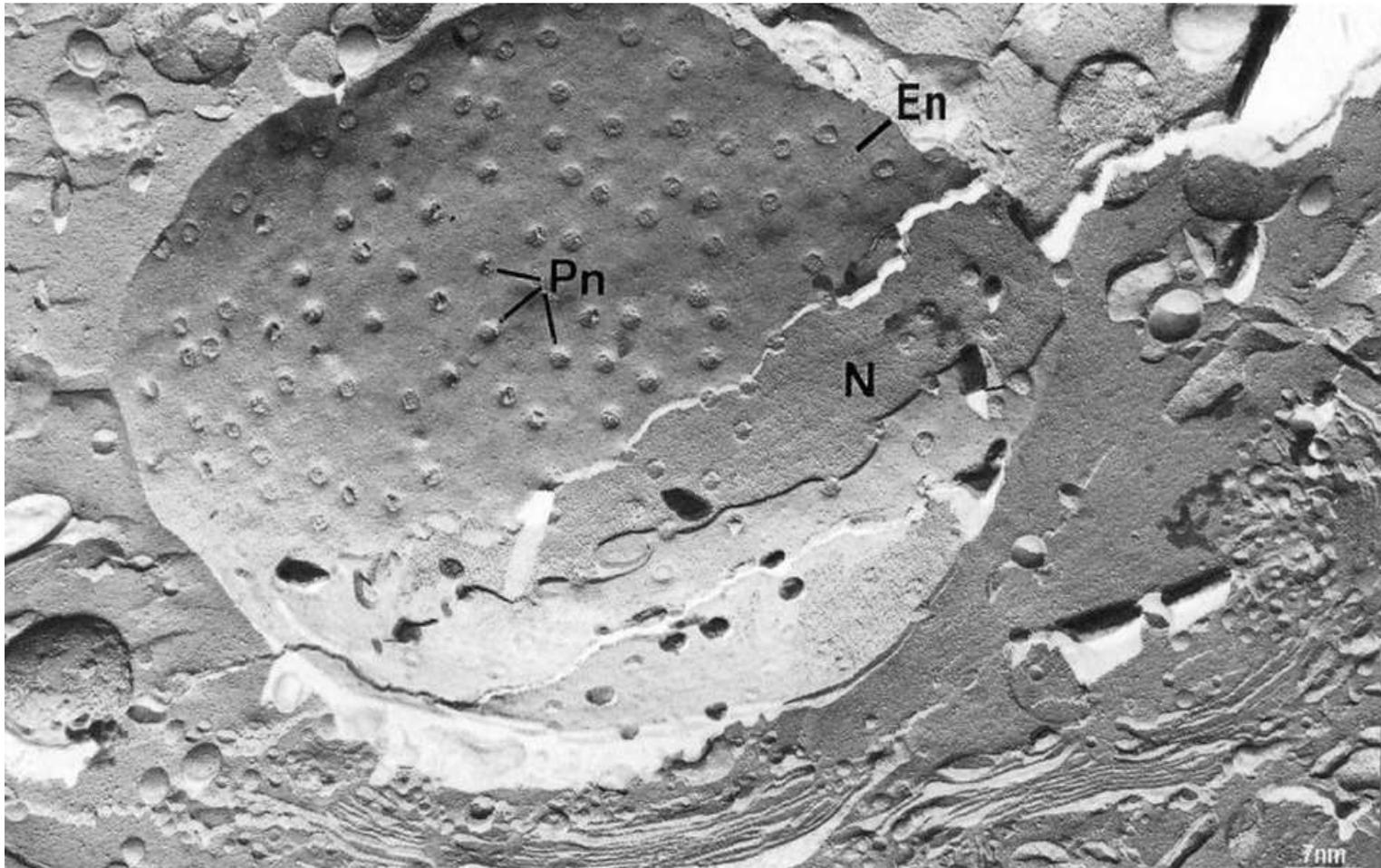
- a. Des érythroblastes de lapin sont cultivés dans un milieu contenant un acide aminé radioactif, l'histidine. Les protéines fabriquées par les érythroblastes sont ensuite extraites et soumises à une électrophorèse qui montre un pic radioactif (X) spécifique de l'hémoglobine de lapin
- b. On fait incuber des œufs d'amphibien (le xénope) dans un milieu contenant de l'histidine radioactive. L'électrophorèse montre alors deux pics protéiques (Y et Z) spécifiques des protéines produites par les œufs de xénope.
- c. On prélève de l'ARN dans le cytoplasme des érythroblastes de lapin, et on l'injecte dans des œufs de xénope que l'on fait ensuite incuber dans un milieu contenant de l'histidine radioactive. On réalise ensuite une électrophorèse. On observe que les œufs d'amphibien expriment sous forme de protéine l'information apportée par l'ARN des érythroblastes de lapin.

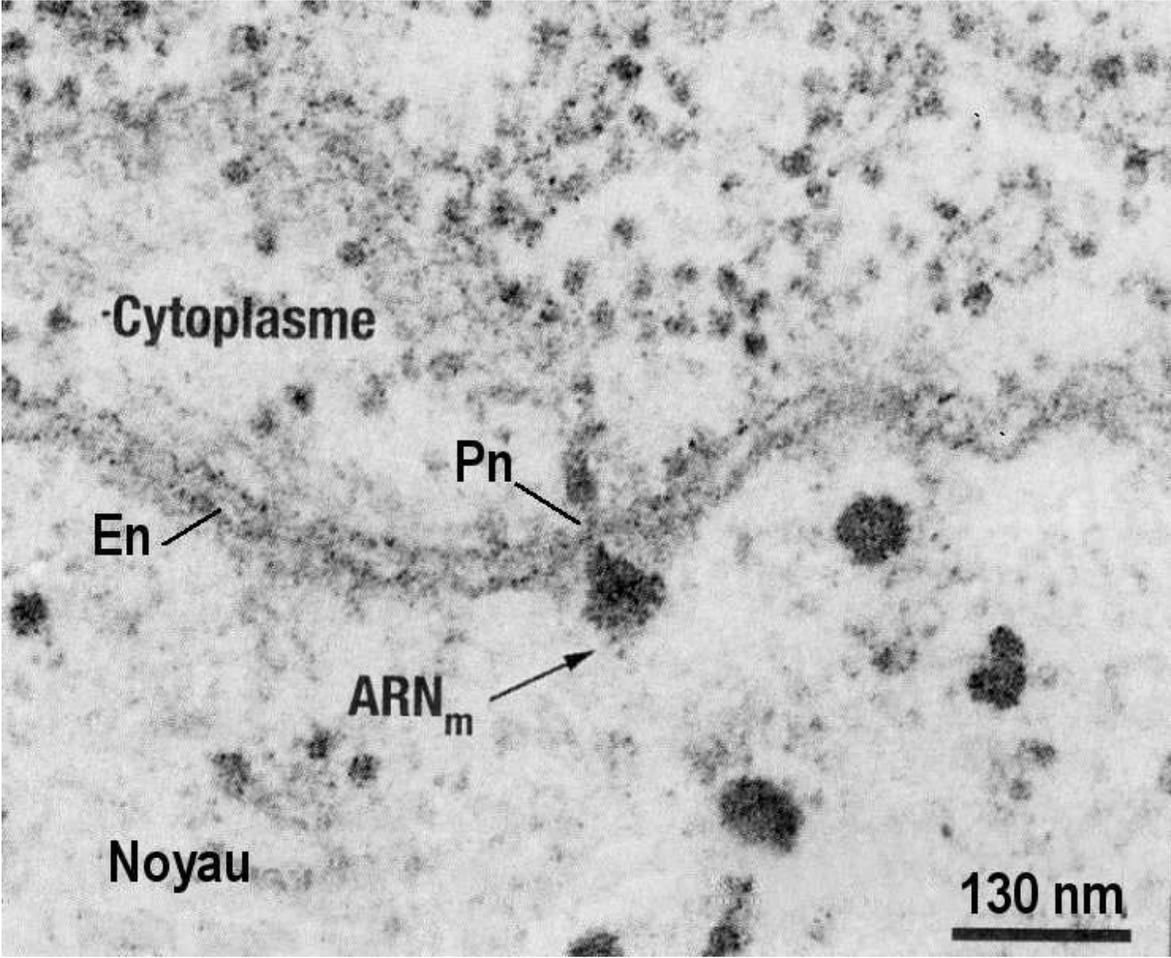
Experience d'autoradiographie (injection d'uracile radioactif)



Des cellules animales sont cultivées sur un milieu contenant de l'**uracile radioactif**.

- a. Autoradiographie après culture sur milieu radioactif pendant 15 minutes.**
- b. Autoradiographie après culture sur milieu radioactif pendant 15 minutes puis transfert sur un milieu de culture non radioactif pendant une heure et demie. L'ARN est formé dans le noyau (a) mais, contrairement à l'ADN, on le retrouve peu après dans le cytoplasme (b).**



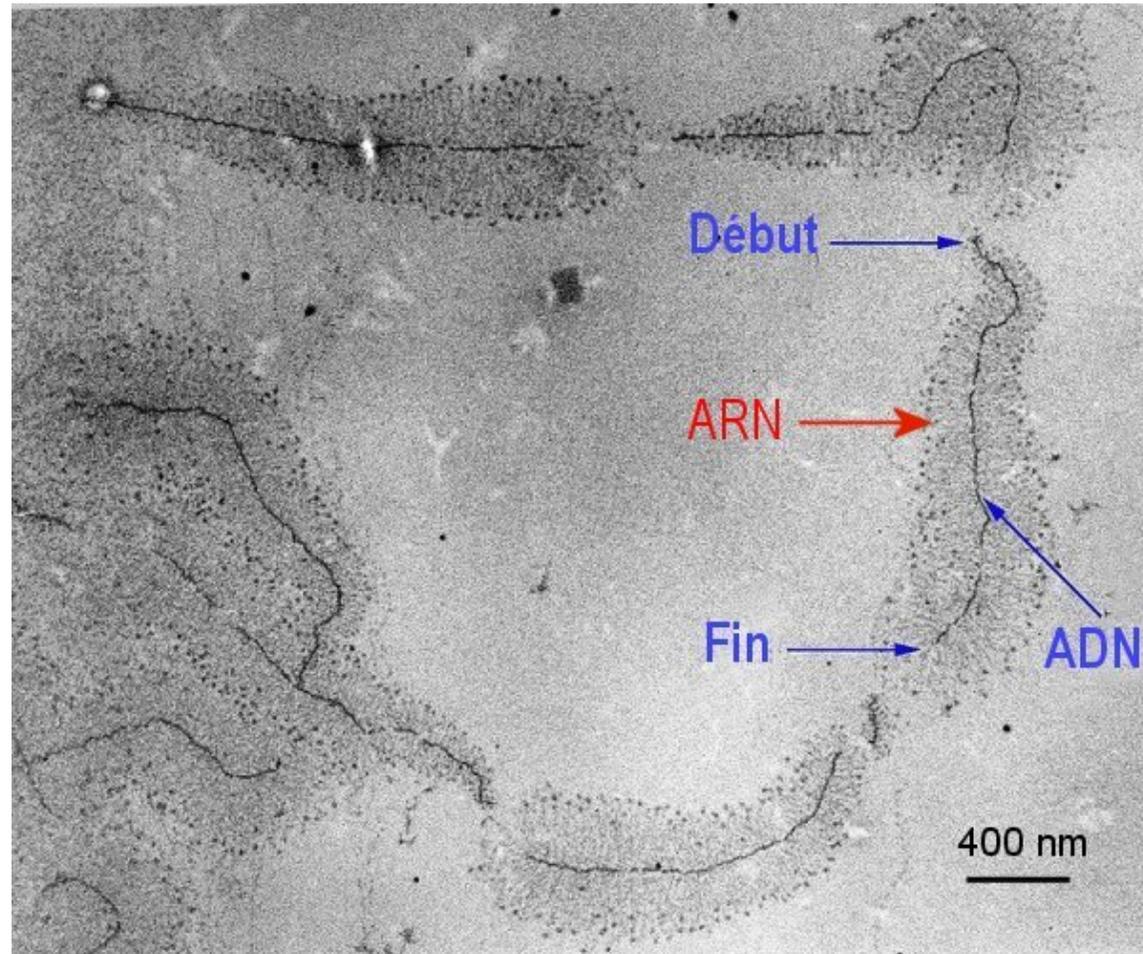


Document ADN /protéine
Correspondance

Comme l'ADN, l'acide ribonucléique messenger, ou ARNm, est un support d'information produit dans le noyau mais il peut migrer dans le cytoplasme en passant par les pores de l'enveloppe nucléaire. Il présente en outre cinq particularités :

- il est formé d'un brin unique de nucléotides ;
- il ne dépasse pas quelques milliers de nucléotides car il ne correspond qu'à un seul gène ;
- le glucide des nucléotides est du ribose (ribonucléotides) ;
- l'uracile (U) remplace la thymine (T) ;
- sa durée de « vie » est courte (quelques heures en général).

3. Les étapes de la transcription



3. Les étapes de la transcription

La transcription s'opère en trois étapes.

1. L'initiation

Sur l'ADN, chaque gène est précédé d'une séquence, ou promoteur, qui indique à la fois le brin à transcrire et le début de la zone à transcrire. Celui-ci permet également la fixation d'une enzyme : l'ARN polymérase (ARNpol).

2. L'élongation.

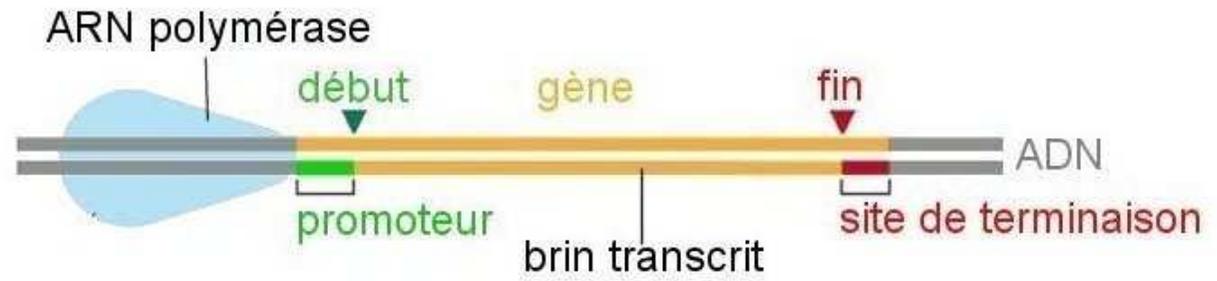
L'ARN polymérase progresse le long de l'ADN et, en respectant la complémentarité des bases, associe un ribonucléotide à chaque désoxyribonucléotide rencontré.

L'ARN obtenu est donc complémentaire du brin transcrit et identique, aux uraciles et riboses près, au brin non transcrit

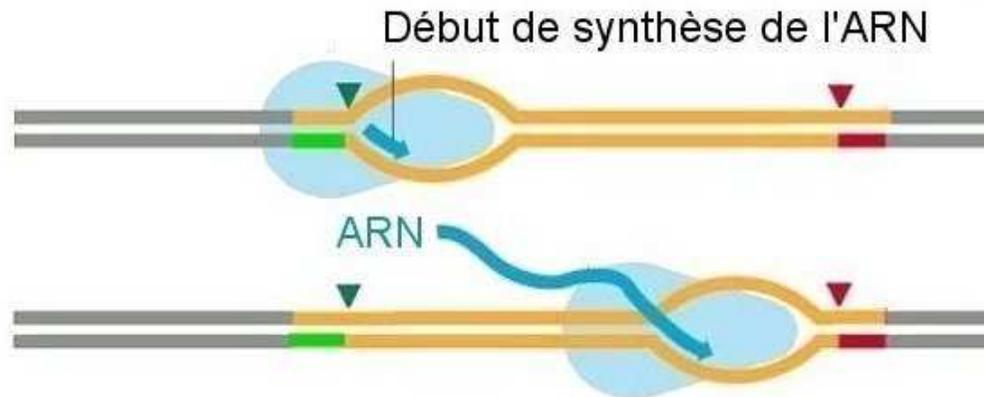
3. La terminaison.

Quand l'ARN polymérase rencontre sur l'ADN un site de terminaison il y a la libération de l'ARN qui pourra quitter le noyau en empruntant les pores nucléaires.

Initiation

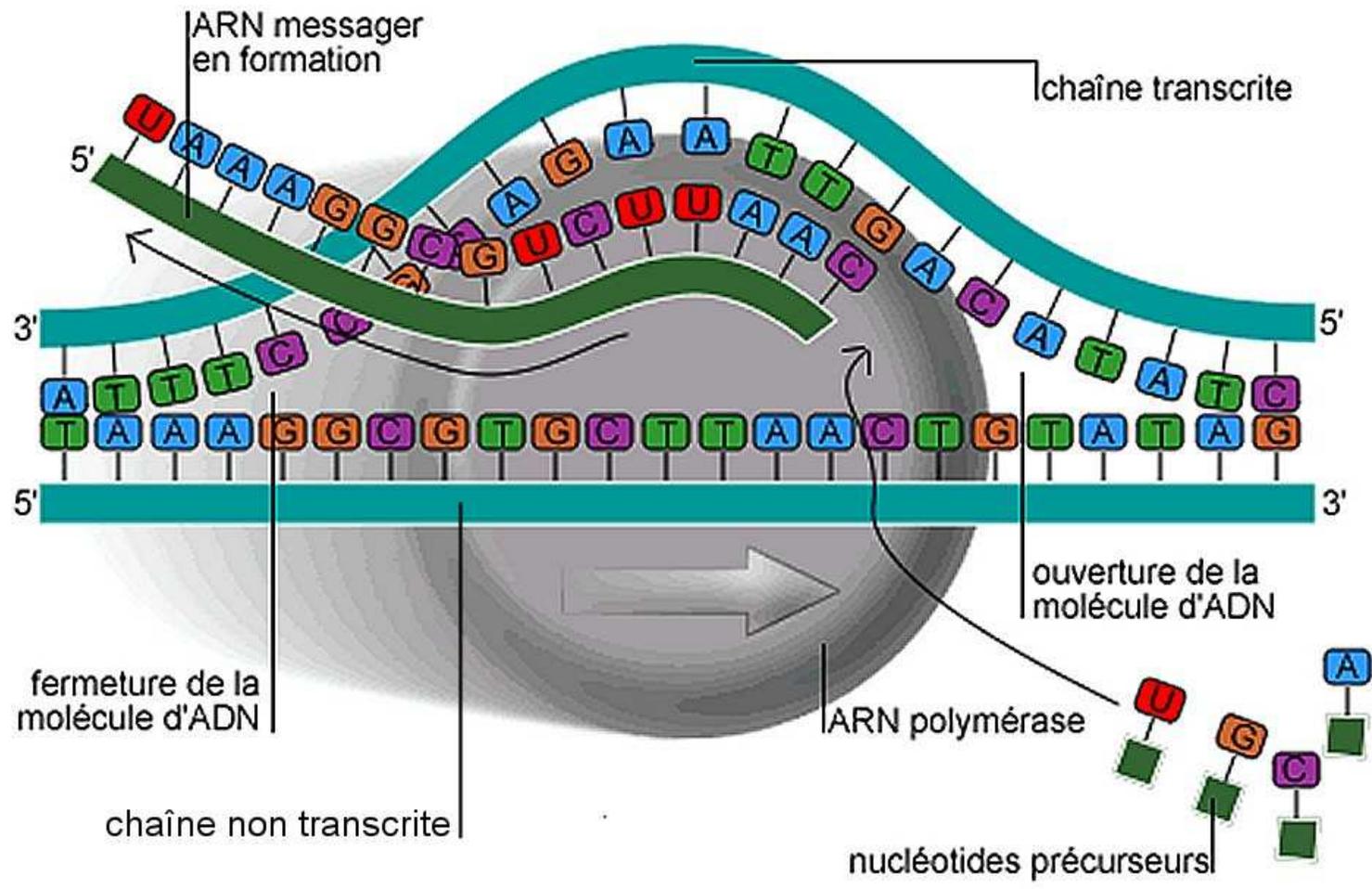


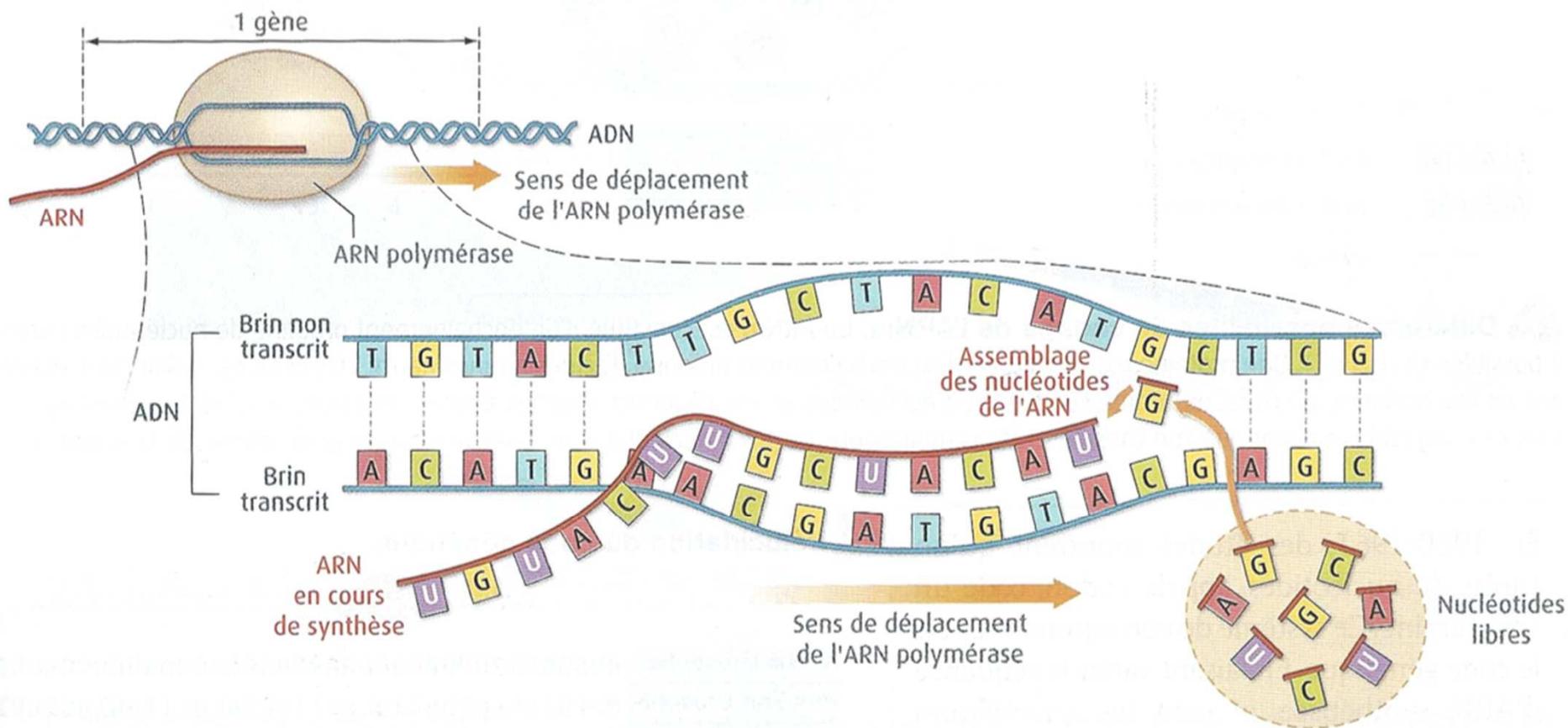
Elongation



Terminaison







I. La transcription et la maturation dans le noyau

B. La maturation

Une apparente contradiction

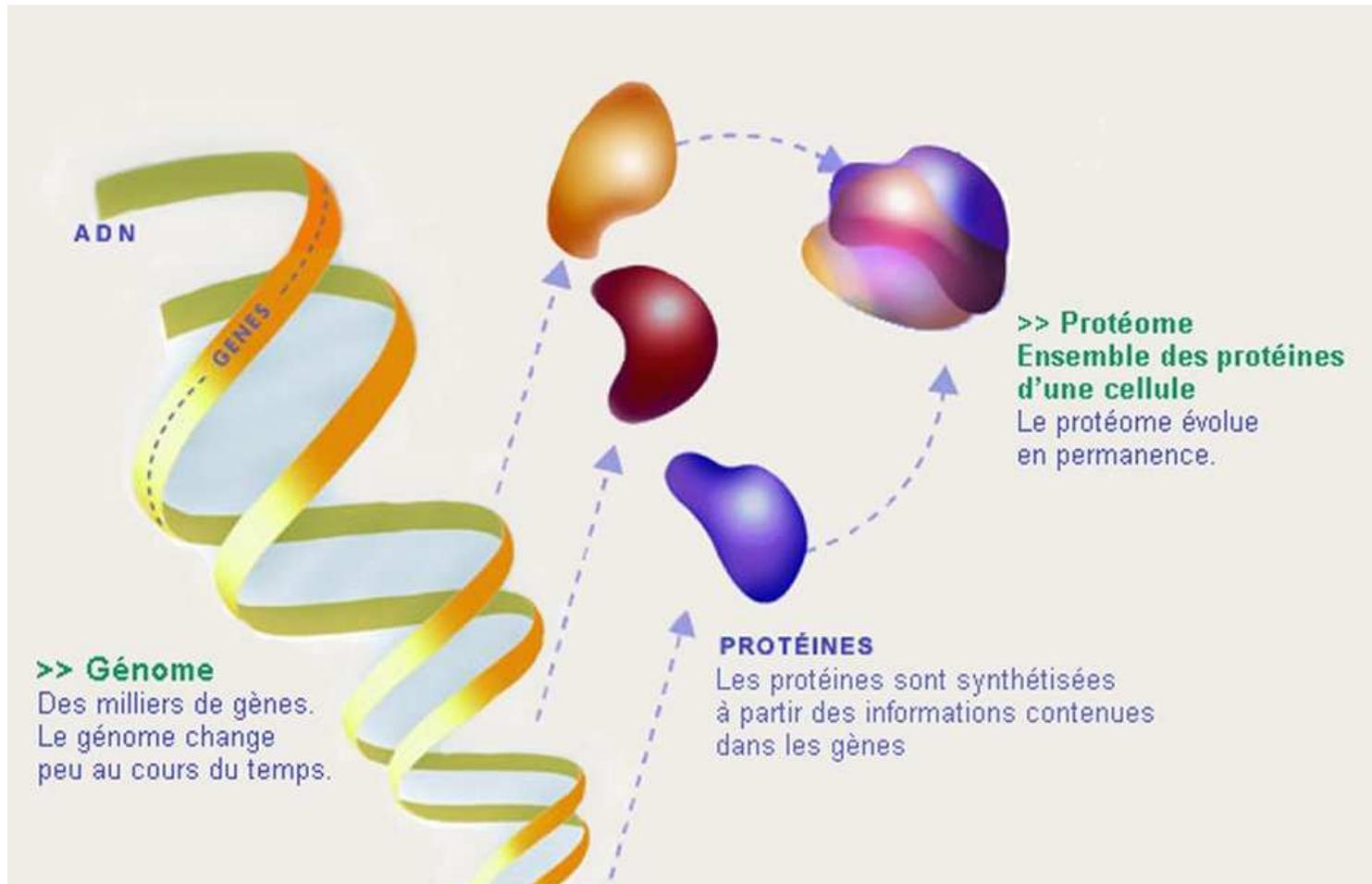
Le **génom**e est l'ensemble du matériel génétique d'un organisme. De 1990 à 2003, le projet Génome humain a établi la séquence complète des 3,2 milliards de paires de bases (= paires de nucléotides) de l'ADN qui constituent le génome humain. Il a été mené à bien grâce à l'utilisation d'un grand nombre de séquenceurs qui décryptent automatiquement l'ADN. On a ainsi pu établir que le génome humain compte environ 25 000 gènes (20 à 30 000 selon les auteurs) alors que l'on pensait jusqu'alors qu'il y avait au moins 100 000.

Une apparente contradiction

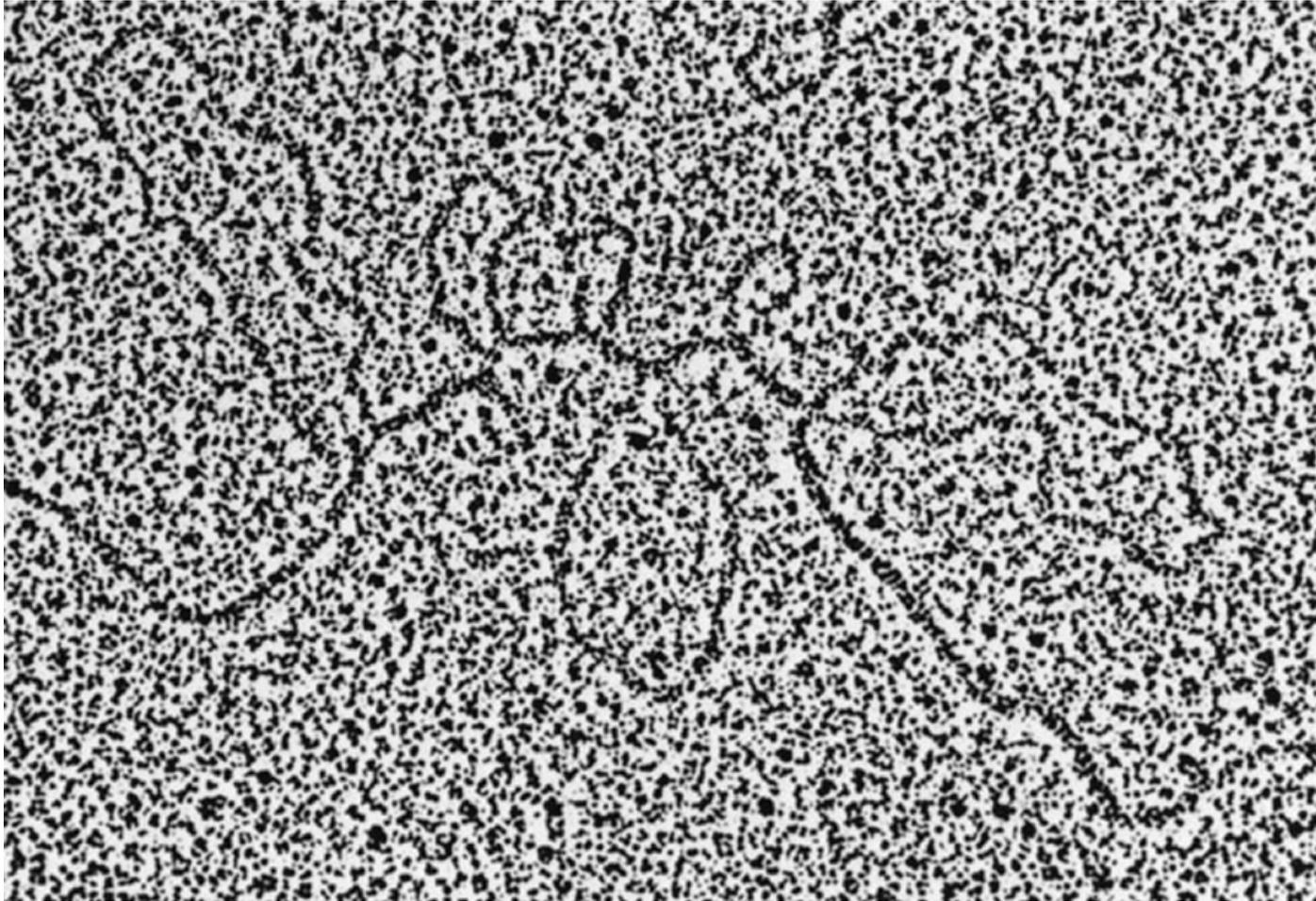
Le protéome

ensemble des protéines codées par un génome. Après le séquençage du génome humain, la communauté scientifique internationale est maintenant engagée dans l'inventaire du protéome humain (*Human Proteome Project*) qui est actuellement estimé à 100 000 protéines (1 000 000 pour certains auteurs). Dans ce cadre, la France est chargée de déterminer les protéines codées par le chromosome 14 (comme ce fut le cas pour le génome).

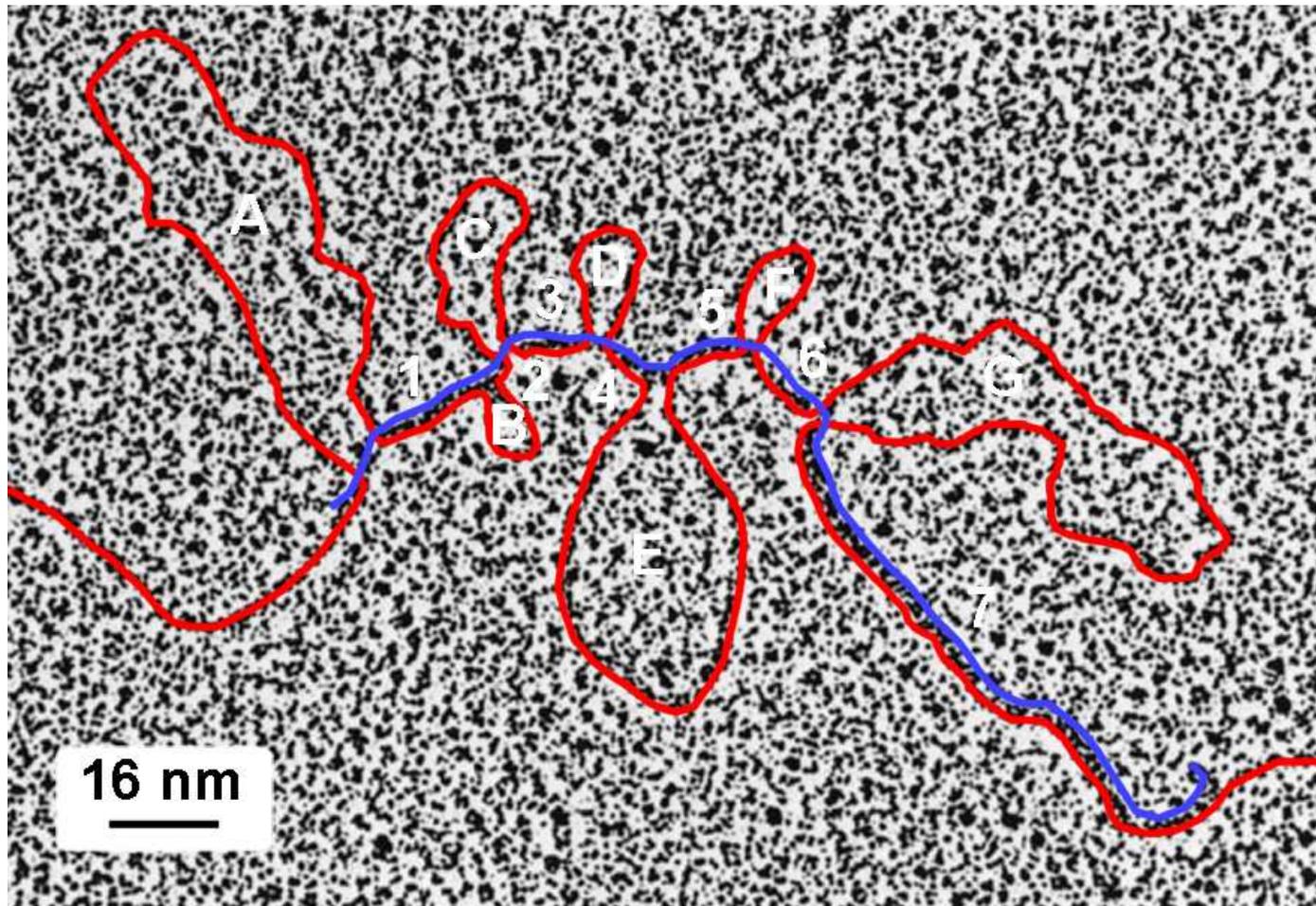
Une apparente contradiction



Comment expliquer ces boucles?

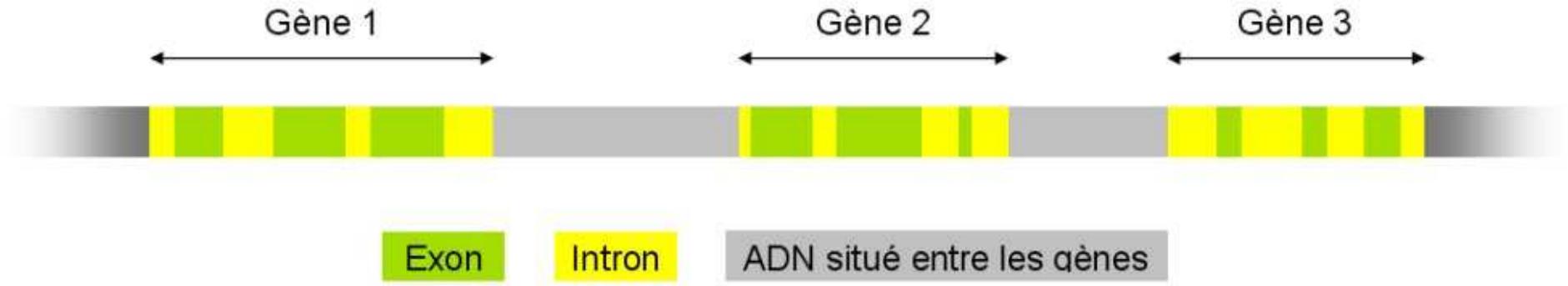


En rouge l'ADN/ en bleu l'ARN



Des gènes morcelés

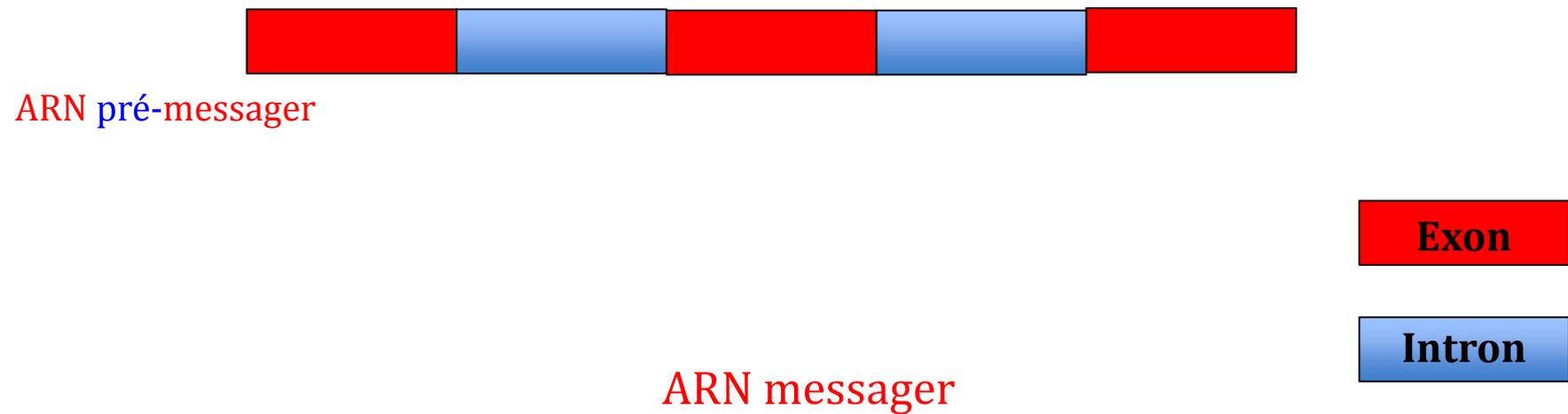
La molécule d'ADN du gène de l'ovalbumine de poule est chauffée ce qui casse les liaisons faibles et sépare les deux brins d'ADN. On ajoute alors l'ARNm correspondant à ce gène qui s'associe au brin d'ADN portant la séquence complémentaire et on obtient une hybridation ADN-ARN. On observe qu'environ 75% du gène ne se retrouve pas sur l'ARNm.



En effet l'ARNpm subit une maturation qui consiste à éliminer (excision) certaines séquences non codantes, ou introns. Les séquences codantes restantes, ou exons, sont alors liées bout à bout (**épissage**) de manière à former l'ARN messenger (ARNm) mature qui sort du noyau et participe à la traduction (entre le codon initiateur AUG et le codon stop, ici UAA). Beaucoup de gènes sont ainsi morcelés.

D'après les expériences d'hybridation ADN/ARN, on remarque que l'ARNm est plus court que l'ADN dont il est issu...

C'est en fait tout à fait normal...



A l'issue de la transcription, l'ARN pré-messager subit un **épissage** : certains fragments sont éliminés (introns), d'autres sont conservés (exons) et vont former l'ARN messager ou ARNm.

Cette maturation peut être plus complexe et aboutir à la formation de plusieurs ARNm...



OU

ARN messager

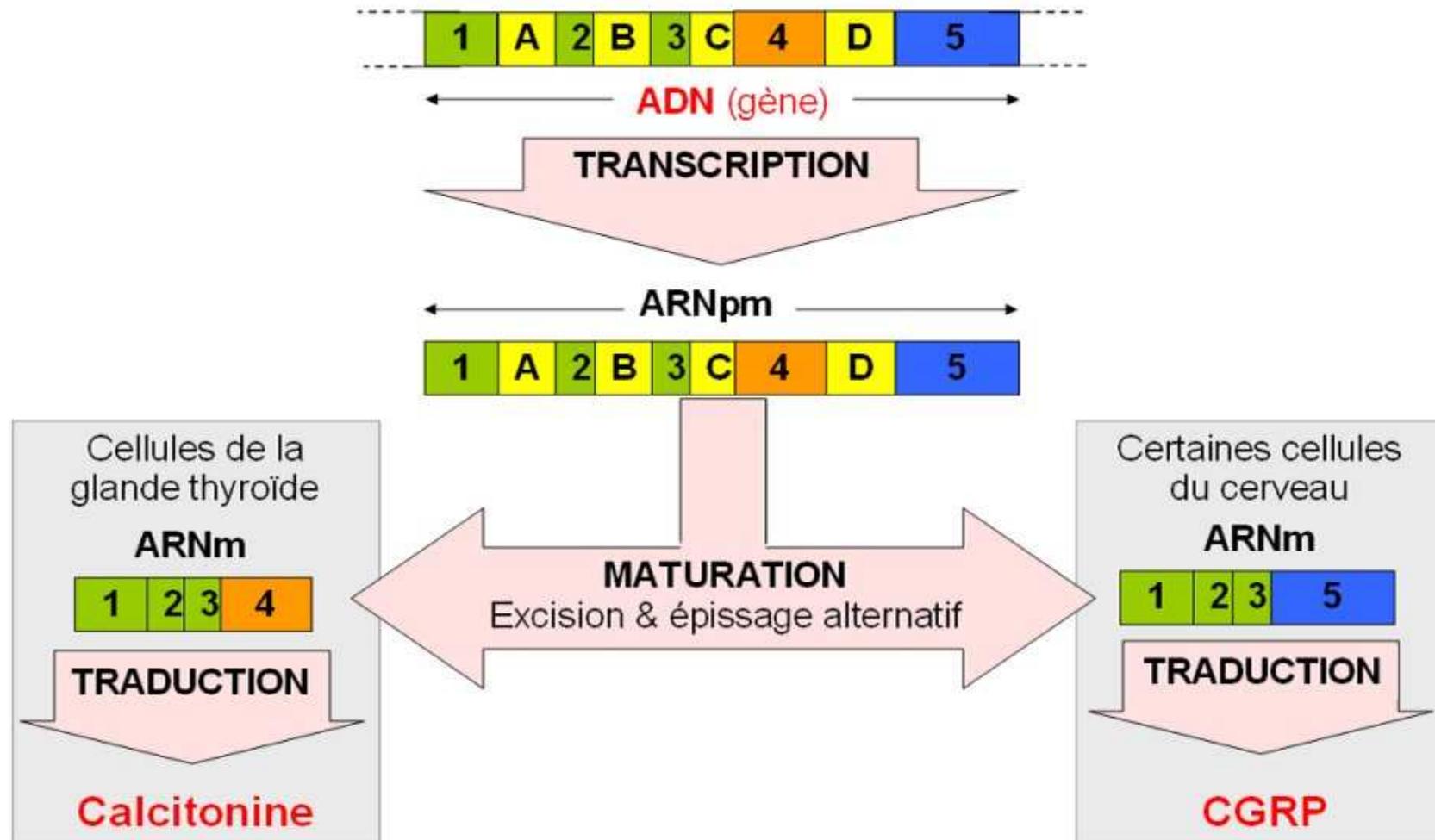
ARN messager

En effet, la maturation de l'ARN peut, à partir d'un même ARN pré-messager, être à l'origine de plusieurs ARN messagers différents. Ainsi un seul gène peut aboutir à la production de plusieurs protéines différentes : **c'est l'épissage alternatif.**

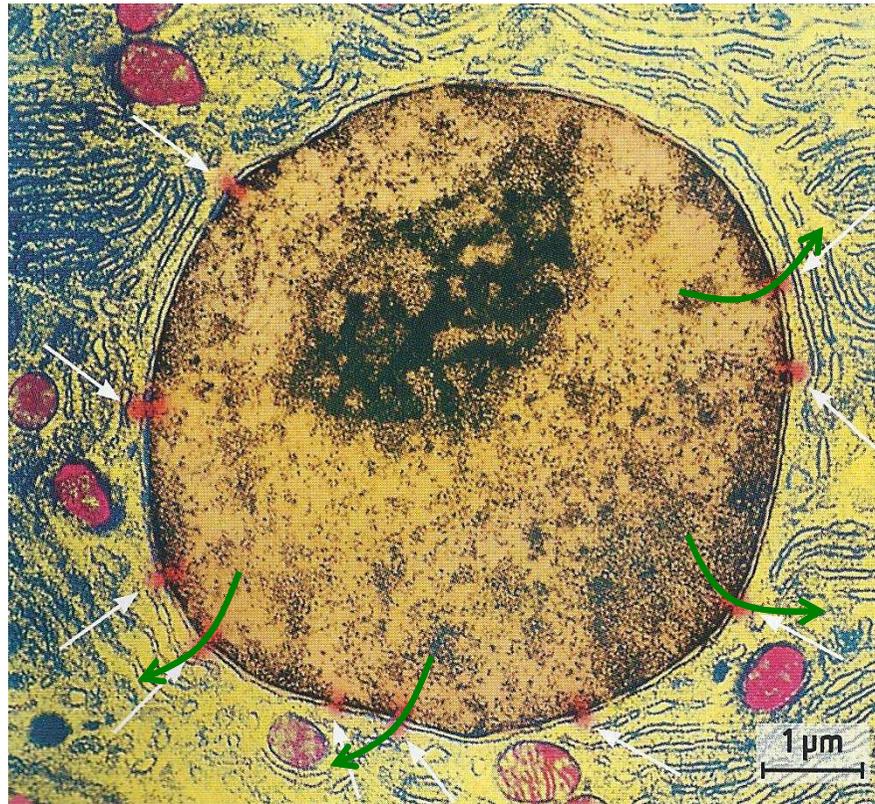
Exercice

Un gène, deux protéines...

- Dans certaines cellules de la glande **thyroïde**, le gène Calc-1 gouverne la synthèse de calcitonine, une hormone qui intervient dans la régulation de la quantité de calcium dans le sang. Dans certaines cellules du cerveau, le même gène Calc-1 commande la synthèse d'une substance permettant la communication entre neurones (*neurotransmetteur*), le *CGRP*, qui a notamment une action vasodilatatrice (*augmentation du diamètre des artères*)
- L'**épissage peut être alternatif**, c'est à dire conduire à des ARNm différents. Selon la maturation que subit l'ARN, un même gène peut coder des protéines différentes.



Une fois synthétisé l'ARNm sort du noyau par les pores nucléaires pour aller dans le cytoplasme, au niveau du réticulum endoplasmique.



b Noyau de cellule (MET, image colorisée).
Les flèches indiquent la présence de pores dans l'enveloppe nucléaire.

La séquence des nucléotides des portions codantes de l'ADN représente une information qui détermine la séquence des acides aminés d'une protéine donnée grâce à un système de correspondance, le code génétique.

La protéosynthèse fait intervenir une molécule transitoire et éphémère, l'ARNm, qui permet la sortie de l'information génétique du noyau. Chez les eucaryotes l'ADN est d'abord transcrit en ARNpm qui, après maturation (épissage alternatif), peut être à l'origine plusieurs ARNm différents traduits, selon le code génétique, en autant de protéines différentes.

II. La traduction

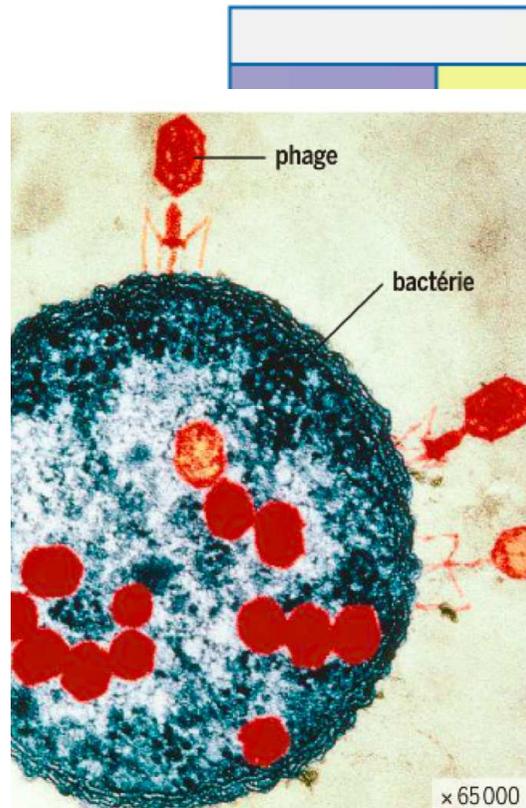
1. Le code génétique

Problème

Il existe 20 acides aminés et seulement quatre nucléotides différents.

Combien faut-il de nucléotides au minimum pour désigner un acide aminé ?

Comment a-t-il été établi ?



Les phages sont des **virus** qui infectent des bactéries et s'y multiplient (*photographie ci-contre*), ce qui aboutit à la destruction de ces dernières.

En 1961, Crick et son équipe ont obtenu, en utilisant des **agents mutagènes**, divers phages portant des mutations par addition ou délétion sur un gène impliqué dans l'infection des bactéries. Ces phages ont été classés en fonction du nombre de nucléotides supprimés ou ajoutés dans le gène. Crick a alors recherché une relation avec le caractère infectieux du phage ainsi muté.

Mutations	Virulence
addition d'un nucléotide	non infectieux
addition de deux nucléotides	non infectieux
addition de trois nucléotides	infectieux
addition de quatre nucléotides	non infectieux
addition de six nucléotides	infectieux
délétion d'un nucléotide	non infectieux
addition et délétion d'un nucléotide	infectieux
délétion de trois nucléotides	infectieux

Remarque : Crick suppose que, si une mutation ne modifie qu'un ou deux acides aminés, la protéine impliquée dans l'infection reste fonctionnelle.

- Un triplet (ou codon) de nucléotides code pour un acide aminé

Les travaux de Nirenberg :

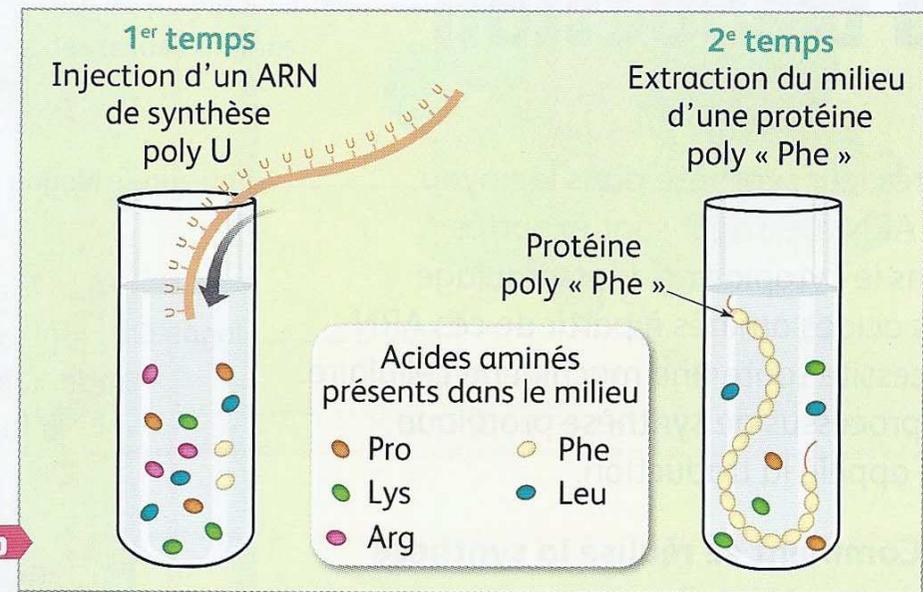
En essayant plusieurs combinaisons de nucléotides, les scientifiques ont, en l'espace de deux ans, décrypté l'intégralité du code génétique. Ces travaux ont de plus confirmé ce qu'avait prévu l'équipe de Crick : chaque acide aminé est codé par un triplet de nucléotides appelé codon ; il en existe 64 différents.

ARN messenger de synthèse	Séquence protéique obtenue
...UUUUUUUUUUUUU...	...Phe-Phe-Phe-Phe...
...AAAAAAAAAAAAA...	...Lys-Lys-Lys-Lys...
...UCUCUCUCUCUC...	...Ser-Lys-Ser-Lys...

a Résultats des expériences de Nirenberg.

Protocole suivi par l'équipe de Nirenberg afin d'élucider le code génétique.

b



- Si un nucléotide désigne un acide aminé on a alors $4^1 = 4$ acides aminés possibles.
- Si deux nucléotides désignent un acide aminé on a alors $4^2 = 16$ acides aminés possibles.
- Si trois nucléotides désignent un acide aminé on a alors $4^3 = 64$ acides aminés possibles.

Puisqu'il y a 64 combinaisons de trois nucléotides parmi quatre pour seulement 20 acides aminés, plusieurs combinaisons correspondent au même acide aminé.

		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	GAC	C	
		GUA	GCA	GAA	GGA	A	
		GUG	GCG	GAG	GGG	G	

A quelques exceptions près, le code génétique est commun à tous les êtres vivants :

CNDP-INRP Anagène

Fichier Edition Traiter Options Fenêtre Aide

Conversion

	1	5	130	145	350	355
Traitement	Gène codant la bêta					
Am-betacod.adn	AUGGUGCACCUGACUCCUG		AGGUCGCCUAUCA		IUGCCUGGCCCCACAAGUAUCACUAA	
Pro-betacod.adn	MetValHisLeuThrProG		AlaAlaTyrGl		AlaLeuAlaHisLysTyrHis	
Traitement	Gène codant le marq					
Am-acod.adn	AUGGCCGAGGUGUUGCGGA		JUCCUGAAGCUGUU		ICAUGGUGGGCCACCGUGUCCACUACUAU	GGCGGUCCGGAACCCGUGA
Pro-acod.adn	MetAlaGluValLeuArgT		heLeuLysLeuPh		ieMetValGlyHisArgValHisTyrTyr	nAlaValArgAsnPro
Traitement	Gène codant 1'hormo					
Am-fshcod.adn	AUGAAGACACUCCAGUUUU		AAGAAUAA			
Pro-fshcod.adn	MetLysThrLeuGlnPheP		_ysGlu			
Traitement	Gène codant 1'hormo					
Am-lhbcod.adn	AUGGAGAUGCUCAGGGGC		JUGACCUGUGACCA		ICCUCUAA	
Pro-lhbcod.adn	MetGluMetLeuGlnGlyL		_euThrCysAspHi		ieLeu	
Traitement	Gène codant la rhod					
Am-rhonorm.cod	AUGAAUGGCACAGAAGGCC		JUGGUGGUCCUGGC		UAAGCCCAUGAGCAACUCCGCUUCGGG	CUAA
Pro-rhonorm.cod	MetAsnGlyThrGluGlyP		_euValValLeuAl		isLysProMetSerAsnPheArgPheGly	a

Sélection : 15/15 lignes

Toutes les séquences codantes d'ARNm commencent par le même **codon AUG**, qui **code la méthionine (Met)**. C'est le **codon d'initiation**. Toutes les séquences codantes d'ARNm finissent par un **codon stop (UAA, UAG ou UGA)** qui ne code aucun acide aminé.

Le code génétique

Deuxième nucléotide

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

Le code génétique est non ambigu car **61 codons désignent chacun un seul acide aminé** alors qu'il n'y en a que 20.

Quatre codons ont un rôle particulier, de ponctuation.

Le codon AUG code l'acide aminé méthionine mais indique aussi le début d'une traduction (codon d'initiation).

Les codons UAA, UAG et UGA sont des codons stop qui indiquent la fin de la traduction. Le code génétique est redondant (= dégénéré) car plusieurs codons désignent le même acide aminé.

Pour aller plus loin...

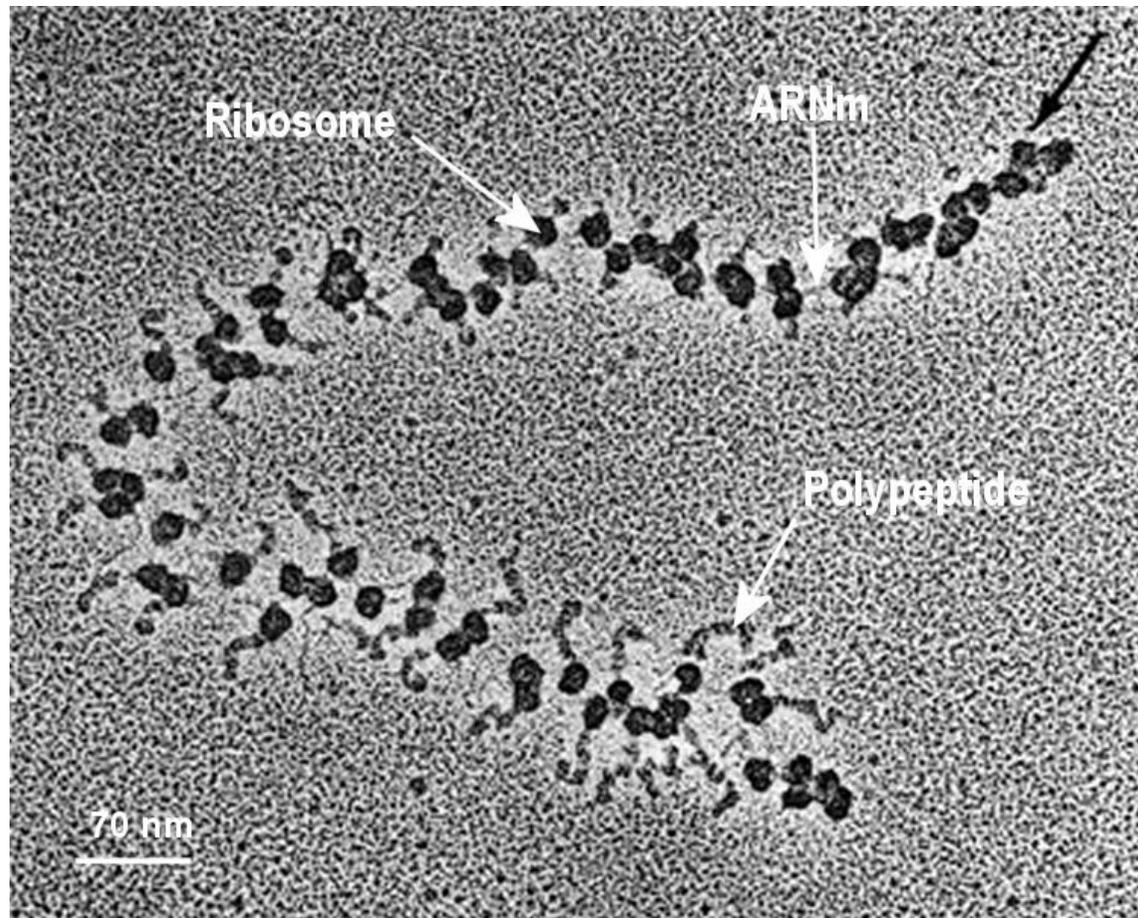
La protéine cdc2 est une molécule indispensable au bon déroulement du cycle cellulaire chez les eucaryotes. Non seulement on constate que la séquence de son ARNm est très voisine chez des espèces très différentes mais que les mêmes codons codent toujours les mêmes acide des aminés.

Bilan

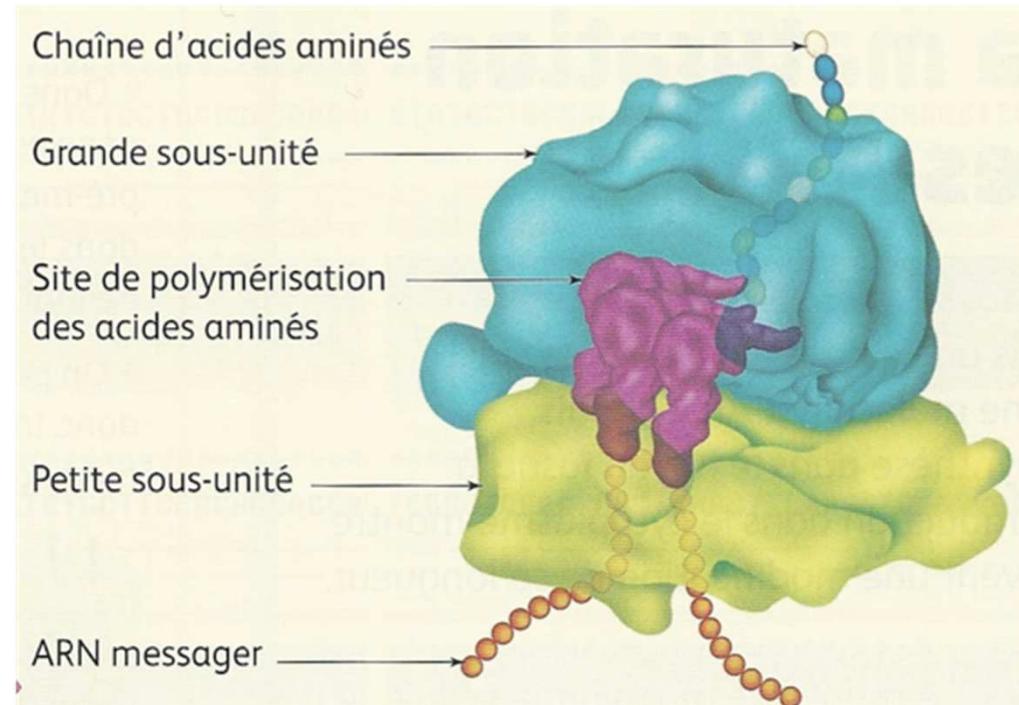
- Le code génétique fait correspondre un acide aminé à trois bases successives de l'ARNm qui forment un codon. Il a quatre particularités :
- il est non ambigu, à un codon correspond un seul acide aminé ;
 - il est dégénéré (ou redondant), plusieurs codons correspondent au même acide aminé ;
 - il est non chevauchant, un nucléotide n'appartient qu'à un seul codon ;
 - il est universel, c'est le même pour tous les êtres vivants (*les exceptions sont très rares*).

2. Les étapes de la traduction

- Conversion ARN m –Protéines



Les ribosomes sont des organites cytoplasmiques globuleux et de petite taille (20 à 30 nm de diamètre) mais visibles au ME. Ils permettent la synthèse d'un polypeptide à partir de l'information génétique portée par l'ARNm. La longueur du polypeptide est plus courte au début de la région traduite et plus longue vers la fin.



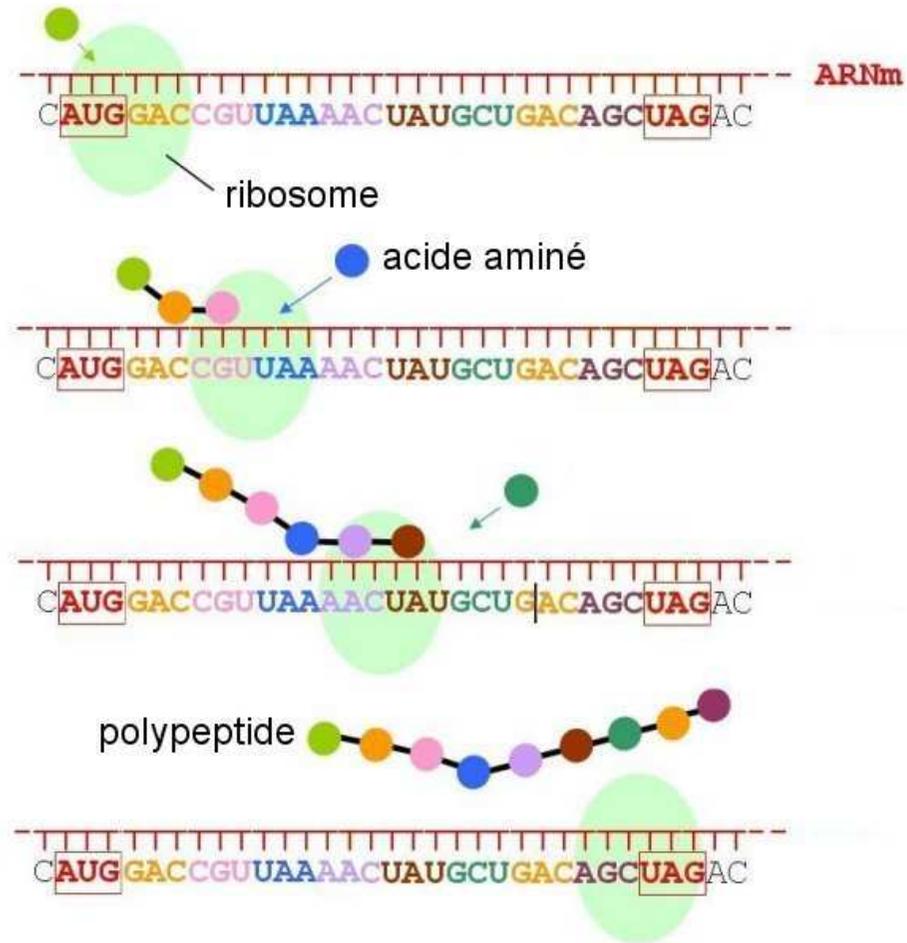
Initiation



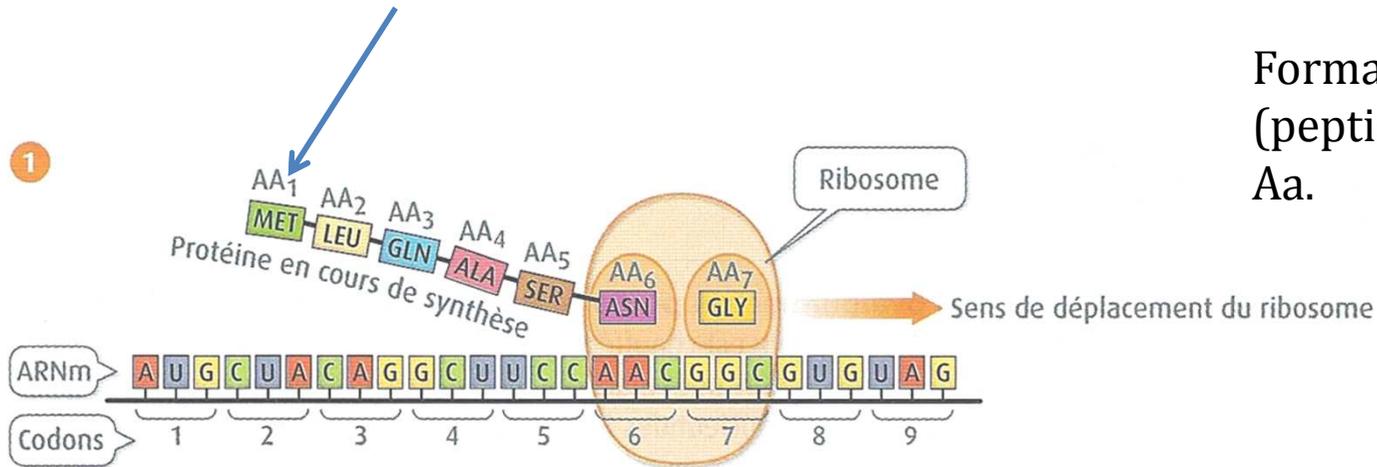
Elongation



Terminaison



Codon initiateur : AUG = méthionine



D'après Belin 2011

Formation d'une liaison
(peptidique) entre deux
Aa.

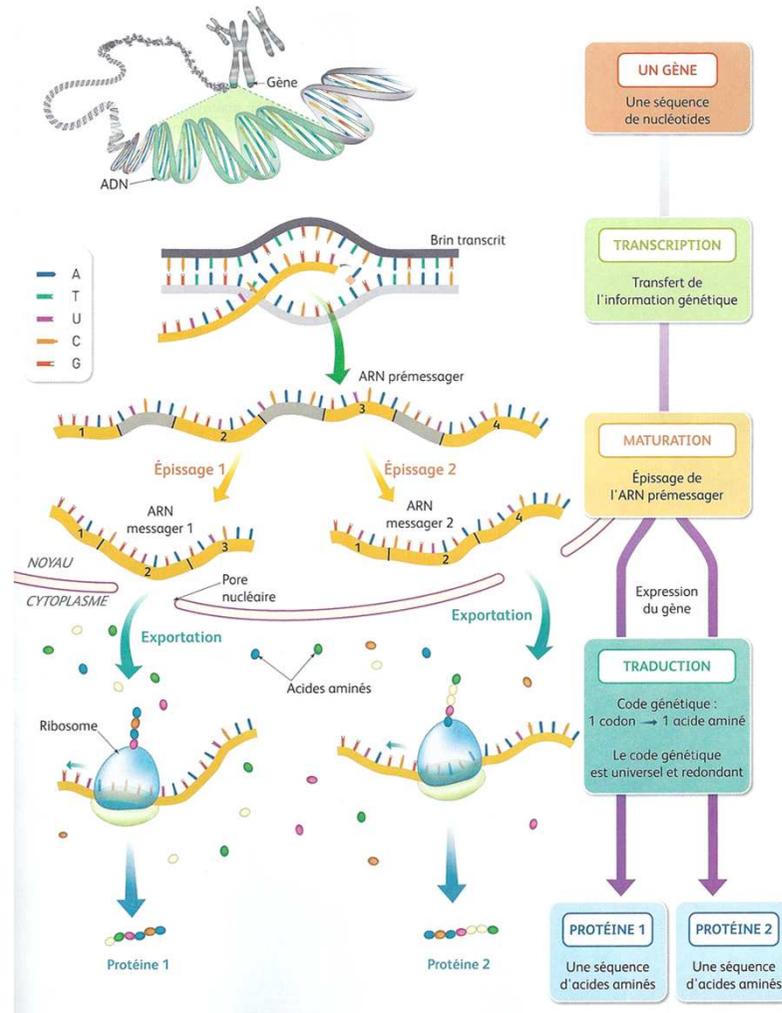
↑

Codon stop : UAG ou UAA ou UGA

Bilan

La traduction s'opère en trois étapes.

1. L'initiation. Un ribosome se fixe au niveau du codon initiateur (AUG) de l'ARNm.
2. L'élongation. Le ribosome progresse le long de l'ARNm. À chaque codon rencontré il recrute un nouvel acide aminé selon le code génétique et l'associe par liaison forte à l'acide aminé précédemment recruté. Plusieurs ribosomes peuvent se succéder en même temps sur même molécule d'ARNm, l'ensemble forme alors un polysome.
3. La terminaison. Parvenu à un codon stop le ribosome se dissocie de l'ARNm et libère le polypeptide.



Bilan:

Un gène s'exprime grâce à la succession de trois évènements (la transcription, la maturation et la traduction de l'ARNm) qui aboutit à la protéine, constituant le phénotype moléculaire.