

CHAPITRE 15

Les gènes et leur fonctionnement

Aperçu du chapitre

- 15.1 Nature des gènes
- 15.2 Le code génétique
- 15.3 La transcription des procaryotes
- 15.4 La transcription de l'eucaryote
- 15.5 L'épissage du pré-ARNm
- 15.6 Structure de l'ARNt et des ribosomes
- 15.7 La traduction
- 15.8 L'expression génique en résumé
- 15.9 Les mutations : des gènes modifiés



Introduction

Nous avons vu comment les gènes déterminent les caractères et comment on peut suivre ces caractères dans les croisements génétiques. Nous avons vu aussi que, dans les gènes, l'information réside dans une molécule d'ADN ; la figure ci-dessus montre tout l'ADN du chromosome dans une cellule d'*E. coli*. L'information de l'ADN est répliquée par la cellule, puis répartie également pendant la division cellulaire. Cette information ressemble beaucoup au plan d'un édifice. La construction de l'édifice utilise l'information contenue dans ce plan, mais il faut aussi des matériaux de construction, des charpentiers et d'autres travailleurs spécialisés utilisant divers outils et collaborant à la construction. De même, l'information présente dans l'ADN a besoin de nucléotides et d'acides aminés comme matériaux de construction, des multiples formes d'ARN et de nombreuses protéines fonctionnant de façon coordonnée pour construire la structure de la cellule.

Nous allons maintenant nous tourner vers la nature des gènes eux-mêmes et la manière dont les cellules tirent l'information stockée dans l'ADN dans l'expression génique. On peut considérer l'expression génique comme une transformation du génotype en phénotype.

15.1 Nature des gènes

Objectifs

1. Évaluer les arguments en faveur de l'hypothèse un gène/ un polypeptide.
2. Faire la distinction entre transcription et traduction.
3. Donner la liste des rôles joués par l'ARN dans l'expression génique.

Nous savons que l'ADN code les protéines, mais cela ne nous dit pas grand-chose sur la manière dont l'information présente dans l'ADN est capable de contrôler le fonctionnement cellulaire. Les chercheurs savaient que les mutations géniques modifiaient les protéines, en particulier les enzymes, bien avant de connaître la structure et le code de l'ADN. Dans cette section, nous analysons la preuve du lien entre gènes et enzymes.

Garrod arrive à la conclusion que les maladies héréditaires peuvent impliquer des enzymes spécifiques

En 1902, le médecin britannique Archibald Garrod remarqua que, chez ses patients, certaines maladies semblaient plus fréquentes dans certaines familles. En examinant plusieurs générations de ces familles, il constata que certaines maladies se comportaient comme si elles étaient dues à des allèles récessifs simples. Garrod en conclut que ces maladies étaient des caractères mendéliens provenant de modifications de l'information héréditaire chez un ancêtre des familles affectées.

Garrod étudia en détail les symptômes de l'alcaptonurie. L'urine des patients atteints de cette maladie contient de l'acide homogentisique (alcaptone) qui s'oxyde rapidement à l'air et devient noire. Chez les individus normaux, l'acide homogentisique est dégradé en molécules plus simples. Garrod en conclut que les patients souffrant d'alcaptonurie ne possédaient pas l'enzyme nécessaire pour catalyser cette dégradation et il pensait que beaucoup d'autres maladies héréditaires pouvaient aussi traduire des déficiences enzymatiques.

Beadle et Tatum ont montré que les gènes définissent les enzymes

En fait, Garrod était en avance sur son temps, car la relation entre les allèles mendéliens et les enzymes devait encore rester obscure pendant près de 40 ans. C'est en 1941 qu'une série d'expériences effectuées à l'Université Stanford par George Beadle et Edward Tatum prouva la liaison entre les caractères observés dans les croisements et les protéines. Ils appliquèrent alors une nouvelle technologie, en induisant, par les rayons X, une série de mutants. Leur travail est résumé à la figure 15.1, puis décrit.

Neurospora crassa, la moisissure du pain

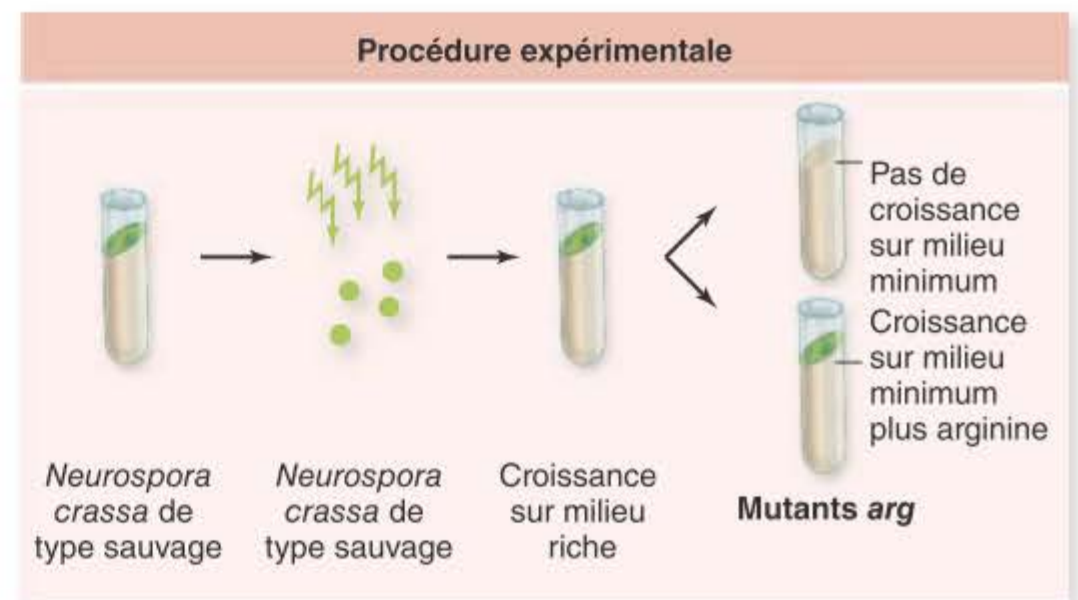
Pour leurs expériences, Beadle et Tatum ont choisi la moisissure du pain, *Neurospora crassa*. Ce champignon est facile à cultiver en laboratoire sur un milieu défini ne contenant qu'une source de carbone (le glucose), une vitamine (la biotine) et des sels inorganiques. On parle d'un milieu « minimum » parce qu'il répond aux besoins minimums nécessaires à la croissance. Toute cellule capable de se développer sur ce milieu minimum doit pouvoir synthétiser toutes les molécules biologiques nécessaires.

En exposant des spores de *Neurospora* aux rayons X, Beadle et Tatum s'attendaient à trouver des mutants incapables de se développer sur le milieu minimum. Ces **mutants auxotrophes** proviennent de dommages subis par les gènes codant des fonctions nécessaires à la synthèse de molécules biologiques importantes.

Les mutants auxotrophes

Pour identifier les mutants auxotrophes, Beadle et Tatum cultivèrent d'abord les champignons sur un milieu riche, puis ils placèrent des cellules individuelles sur le milieu minimum. Ils identifièrent toutes les cellules incapables de synthétiser les substances nécessaires à leur croissance. Ils se concentrèrent sur la possibilité de synthétiser l'arginine, acide aminé dont on connaissait la voie de biosynthèse.

Pour identifier les mutants incapables de synthétiser l'arginine, ils sélectionnèrent les cultures ne se développant que sur le milieu minimum complété par de l'arginine. Cela aboutit à une collection de mutants indépendants, tous incapables de synthétiser l'arginine, et l'on pouvait localiser les mutations à des endroits différents sur les chromosomes. On a ainsi identifié quatre gènes qui furent baptisés *argE*, *argF*, *argG* et *argH*.



Résultats

Mutation de l'enzyme	Plus Ornithine	Plus Citruline	Plus Arginosuccinate	Plus Arginine
E				
F				
G				
H				

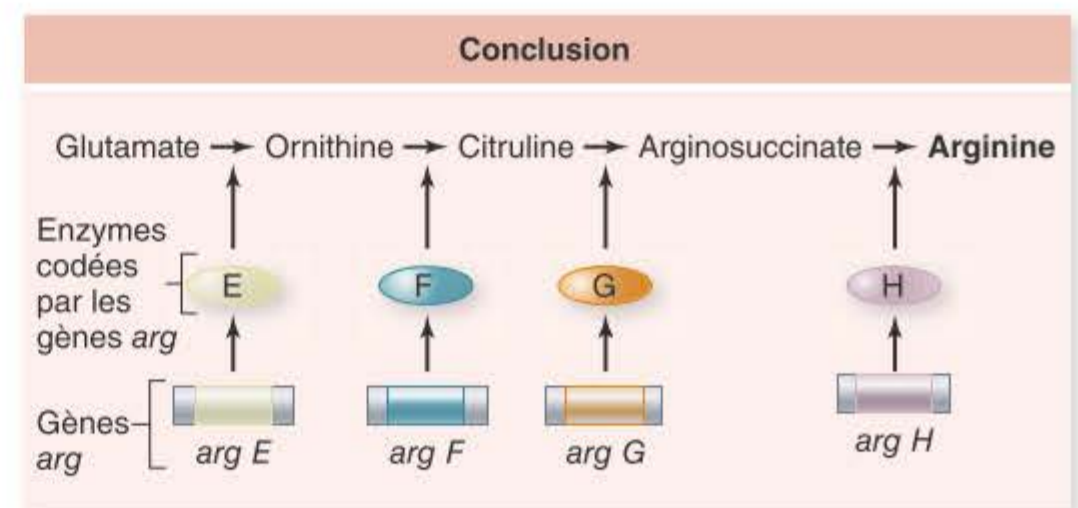


Figure 15.1 L'expérience de Beadle et Tatum. On a induit des mutations chez des *Neurospora* de type sauvage par les rayons X pour obtenir des mutants déficients pour la synthèse de l'arginine (au-dessus). Dans chaque mutant, on a identifié la déficience spécifique en le cultivant sur un milieu comprenant les intermédiaires de la voie de biosynthèse de l'arginine (centre). Un mutant ne se développera que sur le milieu contenant un intermédiaire produit après l'enzyme déficiente de la voie de chaque mutant. On peut ensuite mettre en relation les enzymes de la voie et les gènes sur le chromosome (en bas).

Un gène/un polypeptide

Avec ces outils en main, Beadle et Tatum ont pu disséquer génétiquement la biosynthèse de l'arginine. En complétant le milieu minimum par les intermédiaires de la voie biochimique, ils purent identifier une lésion spécifique pour chaque mutant. Pour chaque intermédiaire, si la mutation affecte une enzyme agissant dans la voie avant l'intermédiaire ajouté, la croissance doit être possible – mais pas si la mutation affecte une étape postérieure à l'intermédiaire ajouté (figure 15.1).

De cette façon, Beadle et Tatum purent isoler une souche mutante déficiente pour chaque enzyme de la voie de biosynthèse. Tous les mutants étudiés étaient donc déficients pour une seule enzyme à cause d'une mutation à un seul endroit d'un chromosome.

Beadle et Tatum arrivèrent à la conclusion que les gènes déterminent la structure des enzymes et que chaque gène code la structure d'une seule enzyme (voir figure 15.1). Ils appelèrent cette relation l'*hypothèse un gène/une enzyme*. Beaucoup d'enzymes étant composées de nombreuses sous-unités polypeptidiques codées par des gènes distincts, on parle aujourd'hui plus souvent de l'hypothèse un gène/un polypeptide. On décrit ainsi clairement la relation moléculaire entre le génotype et le phénotype.

À mesure que nous connaissons mieux les génomes et l'expression des gènes, nous verrons que cette relation évidente est beaucoup trop simple. Les gènes des eucaryotes sont beaucoup plus complexes que ceux des procaryotes et certaines enzymes sont composées, au moins en partie, d'ARN, lui-même intermédiaire dans la synthèse des protéines. Néanmoins, ce principe un gène/un polypeptide est un point de départ utile quand on s'intéresse à l'expression génique.

Analyse des données Si vous avez obtenu un double mutant avec *argE* et *argG*, sur quel milieu cette souche se développera-t-elle ? De façon générale, quel enseignement vous donnera un double mutant ?

Le dogme central décrit le flux d'information passant de l'ADN à l'ARN et aux protéines dans les cellules

Pour passer du génotype au phénotype, il faut que l'information emmagasinée dans l'ADN soit transformée en protéines. La nature du flux d'information dans les cellules fut d'abord décrite par Francis Crick

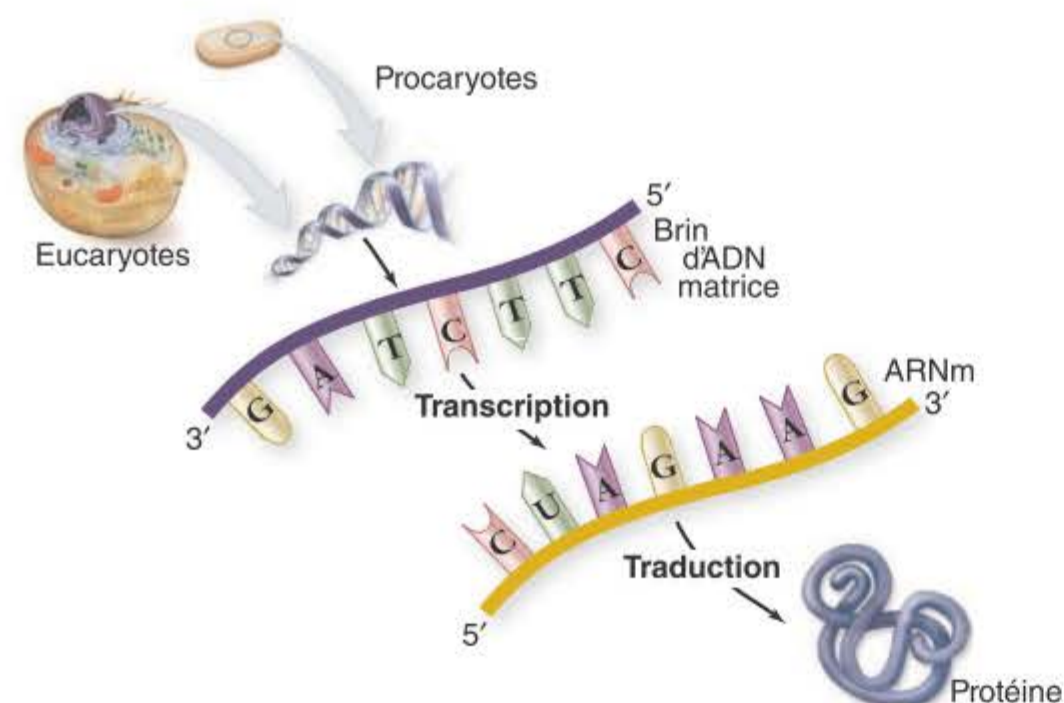


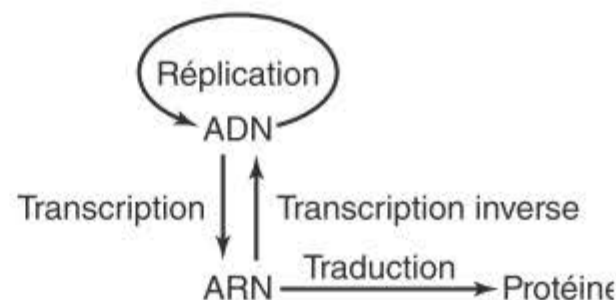
Figure 15.2 Le dogme central de la biologie moléculaire. L'ADN est transcrit en ARNm qui est traduit en protéine.

comme le **dogme central de la biologie moléculaire**. L'information passe dans un sens du gène (ADN) à une copie d'ARN du gène et cette copie contrôle l'assemblage progressif d'une chaîne d'acides aminés en protéine (figure 15.2). En bref,



Ce dogme central représente un cadre décrivant le flux d'information dans les systèmes biologiques. L'étape allant de l'ADN à l'ARN est la **transcription** : elle donne une copie exacte de l'ADN, un peu comme une transcription officielle contient les termes exacts d'un acte de justice. L'étape de l'ARN à la protéine est la **traduction** : elle suppose la traduction de la langue des acides nucléiques en la langue des protéines.

Depuis la formulation originelle du dogme central, on a découvert une catégorie de virus, les **rétrovirus**, capables de convertir leur génome d'ARN en une copie d'ADN, grâce à une enzyme virale, la **transcriptase inverse**. Cette conversion est en contradiction avec la direction du flux d'information du dogme central, et cette découverte exigeait une adaptation du flux potentiel de l'information pour inclure ce flux « inverse » de l'ARN à l'ADN.



La transcription copie l'ADN sous forme d'ARN

La transcription synthétise une copie d'ARN de l'information présente dans l'ADN. C'est donc une synthèse d'ARN dirigée par l'ADN et effectuée par une enzyme, l'ARN polymérase (figure 15.3). Ce processus est

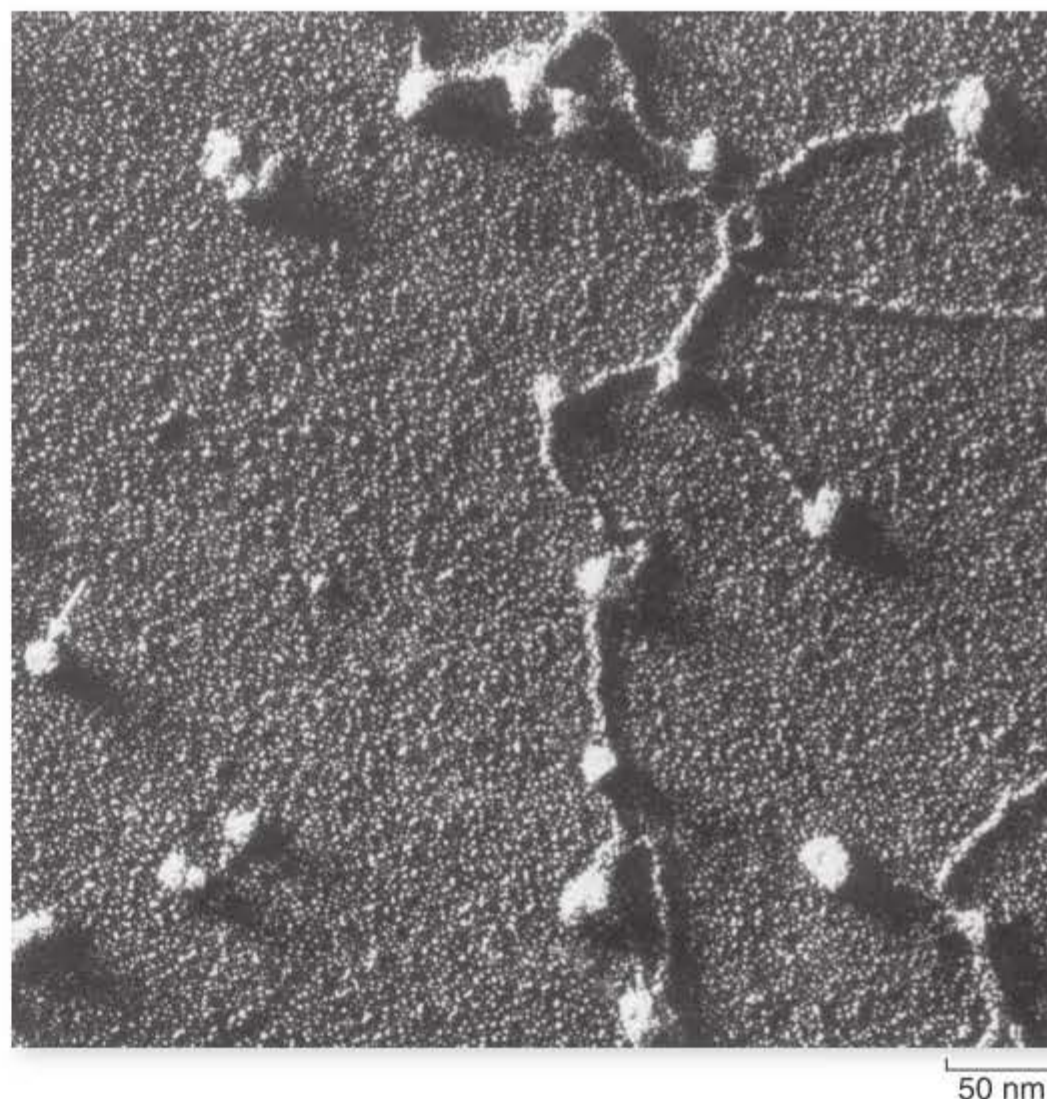
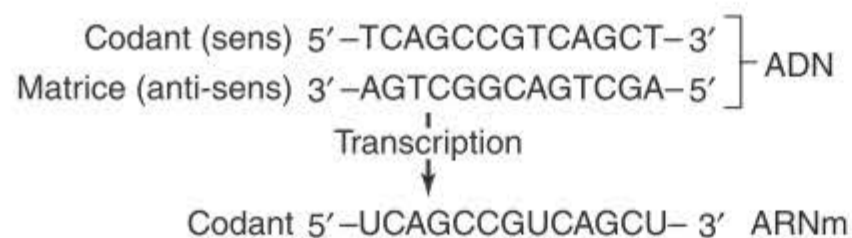


Figure 15.3 L'ARN polymérase. Microphotographie électronique avec ombrage au platine montrant des molécules d'ARN polymérase fixées aux brins d'ARN.

basé sur le principe de complémentarité, décrit au chapitre 14, pour édifier l'ARN en prenant l'ADN comme modèle.

L'ADN étant bicaténaire et l'ARN monocaténaire, un seul des deux brins d'ADN doit être copié : c'est le **brin matrice**. La séquence du brin d'ARN transcrit est complémentaire de celle du brin matrice. Le brin d'ADN qui ne sert pas de modèle est le **brin codant**. Sa séquence est la même que celle du transcrit d'ARN, sauf que l'uracile (U) de l'ARN est une thymine (T) dans le brin d'ADN codant. Une autre convention pour les deux brins d'ADN consiste à désigner le brin codant comme le brin sens, parce qu'il a le même « sens » que l'ARN. Le brin matrice serait alors le brin anti-sens.



L'ARN transcrit servant au contrôle de la synthèse des polypeptides est l'*ARN messenger* (ARNm). Son nom traduit son rôle dans la cellule dans le transfert du message d'ADN au ribosome pour son utilisation.



Question L'ARN polymérase n'est pas capable d'autocorrection. Quelle est la conséquence pour les taux d'erreur à la transcription en comparaison de la réplication de l'ADN ? Pourquoi pensez-vous que l'autocorrection est plus importante pour l'ADN polymérase que pour l'ARN polymérase ?

La traduction utilise l'information de l'ARN pour synthétiser les protéines

La traduction est nécessairement beaucoup plus complexe que la transcription. Dans ce cas, l'ARN ne peut servir directement de modèle pour la protéine parce qu'il n'y a pas de complémentarité – une séquence d'acides aminés ne peut s'aligner à un modèle d'ARN par une sorte de « correspondance chimique ». Les généticiens moléculaires ont imaginé l'existence d'une sorte de molécule adaptatrice capable d'interagir avec l'ARN et les acides aminés : on a découvert que l'*ARN de transfert* (ARNt) remplissait ce rôle. Cette nécessité d'un intermédiaire ajoute au processus un niveau de complexité qui n'existe ni à la réplication de l'ADN, ni à sa transcription en ARN.

La traduction se déroule sur le ribosome, machine cellulaire de synthèse protéique, et elle exige la participation de multiples formes d'ARN et beaucoup de protéines. Voici un aperçu de ces processus, qui seront détaillés dans le reste du chapitre.

Les rôles de l'ARN dans l'expression génique sont multiples

Tous les ARN sont synthétisés à partir d'un ADN matrice. L'expression génique implique la participation de nombreuses sortes d'ARN dont les rôles diffèrent dans l'ensemble du processus. Voici un bref résumé de ces rôles, qui seront détaillés par la suite.

ARN messenger. Même avant que l'on ait dévoilé les détails de l'expression génique, les généticiens savaient qu'il devait exister une forme intermédiaire d'information passant du noyau au

cytoplasme pour le fonctionnement du ribosome. C'était l'« hypothèse du messenger », et cette expression est restée dans le terme *ARN messenger* (ARNm).

ARN ribosomique. Dans les ribosomes, on trouve l'**ARN ribosomique (ARNr)**. Il en existe de nombreuses formes, et elles se trouvent dans les deux sous-unités. Cet ARNr est essentiel pour le fonctionnement du ribosome.

ARN de transfert. La molécule adaptatrice intermédiaire entre l'ARNm et les acides aminés est l'**ARN de transfert (ARNt)**. Ces molécules portent les acides aminés fixés par covalence à une extrémité et possèdent à l'autre bout un anticodon capable de s'apparier par ses bases à un codon de l'ARNm. Les ARNt interprètent l'information de l'ARNm et interviennent dans le positionnement des acides aminés sur le ribosome.

Petit ARN nucléaire. Les **petits ARN nucléaires (ARNsn)** font partie de la machine de maturation nucléaire eucaryote des « pré-ARNm ». Nous parlerons de ce type d'épissage plus loin dans ce chapitre.

ARN SRP. Chez les eucaryotes, certaines protéines sont synthétisées par les ribosomes sur le réticulum endoplasmique rugueux (RER), et la **particule de reconnaissance du signal** (*signal recognition particle*, ou **SRP**) intervient dans ce processus et sera décrite plus tard. La SRP contient ARN et protéines.

Petits ARN. Cette classe d'ARN comprend les **micro-ARN (miARN)**, ainsi que les **petits ARN d'interférence (siARN)**. Ces derniers participent au contrôle de l'expression des gènes décrit au chapitre 16.

Questions d'apprentissage 15.1

Garrod a montré que des modifications d'enzymes peuvent être une cause de désordres métaboliques. Beadle et Tatum ont prouvé que chaque gène code une seule enzyme. L'information génétique s'écoule de l'ADN (les gènes) aux protéines (enzymes) avec l'ARN comme intermédiaire. La transcription convertit l'information présente dans l'ADN en un transcrit d'ARN, et la traduction convertit cette information en protéine. Il existe plusieurs formes d'ARN avec des fonctions différentes : l'ARNm (le transcrit), l'ARNt (l'intermédiaire) et l'ARNr (dans les ribosomes), ainsi que les ARNsn, ARN SRP et les petits ARN (miARN, siARN).

- Pourquoi les cellules ont-elles besoin d'une molécule adaptatrice comme l'ARNt entre ARN et protéine ?

15.2 Le code génétique

Objectifs

1. Résumer les expériences qui ont révélé le code génétique.
2. Décrire les caractéristiques du code génétique.
3. Identifier les rapports entre codons et acides aminés.

Comment l'ordre des nucléotides d'une molécule d'ADN détermine-t-il la séquence des acides aminés d'un polypeptide ? La réponse à cette question fut donnée en 1961, grâce à une expérience menée par

Francis Crick et Sydney Brenner. Cette expérience était tellement élégante et le résultat tellement important pour la compréhension du code génétique que nous allons la décrire en détail.

Le code est lu par groupes de trois

Crick et Brenner pensaient que le code génétique consistait vraisemblablement en une série de blocs d'information, ou **codons**, correspondant chacun à un acide aminé de la protéine codée. Ils supposèrent ensuite que l'information d'un codon était probablement une séquence de trois nucléotides. Avec quatre nucléotides d'ADN (G, C, T et A), la participation de deux d'entre eux à un codon ne donnerait que 4^2 , soit 16 codons différents – trop peu pour coder 20 acides aminés. Mais trois nucléotides donnent 4^3 , soit 64 combinaisons différentes de trois, plus qu'il n'en faut.

Existe-t-il des interruptions dans le code génétique ?

En théorie, une ponctuation pourrait exister entre les codons du gène, avec entre eux des nucléotides non utilisés, comme les espaces séparant les mots de cette phrase. Inversement, les codons pourraient être contigus et former une suite continue de nucléotides.

Si l'information du message génétique est séparée par des espaces, la modification d'un mot n'affecterait pas toute la phrase. Au contraire, si tous les mots sont liés, mais lus par groupes de trois, toute modification n'impliquant pas un groupe de trois modifierait toute la phrase. Ces deux moyens d'utiliser l'information de l'ADN impliquent des modes différents de traduction en protéines.



Preuve de l'absence d'interruptions

Pour faire un choix entre ces mécanismes, Crick et ses collègues utilisèrent une substance chimique pour créer des mutations impliquant des insertions ou des délétions d'une seule base dans une molécule d'ADN viral. Ils montrèrent ainsi qu'une insertion combinée à une délétion rétablissait la fonction, alors que l'une ou l'autre de ces mutations faisait disparaître cette fonction. Dans ce cas, seule la région située entre l'insertion et la délétion devait être modifiée. En choisissant une région du gène codant une partie de la protéine non essentielle à son fonctionnement, cette petite modification ne changeait pas le phénotype.

Avec une ou deux délétions rapprochées, le message génétique était décalé et toute la séquence d'acides aminés était modifiée après la délétion. Avec trois délétions cependant, la protéine était normale après

les délétions. Ils obtenaient les mêmes résultats en ajoutant à l'ADN 1, 2 ou 3 nucléotides.

Crick et Brenner en conclurent donc que le code génétique est lu par trois nucléotides à la fois (autrement dit, c'est un code de triplets), et que la lecture est continue, sans ponctuation entre les unités de trois nucléotides.

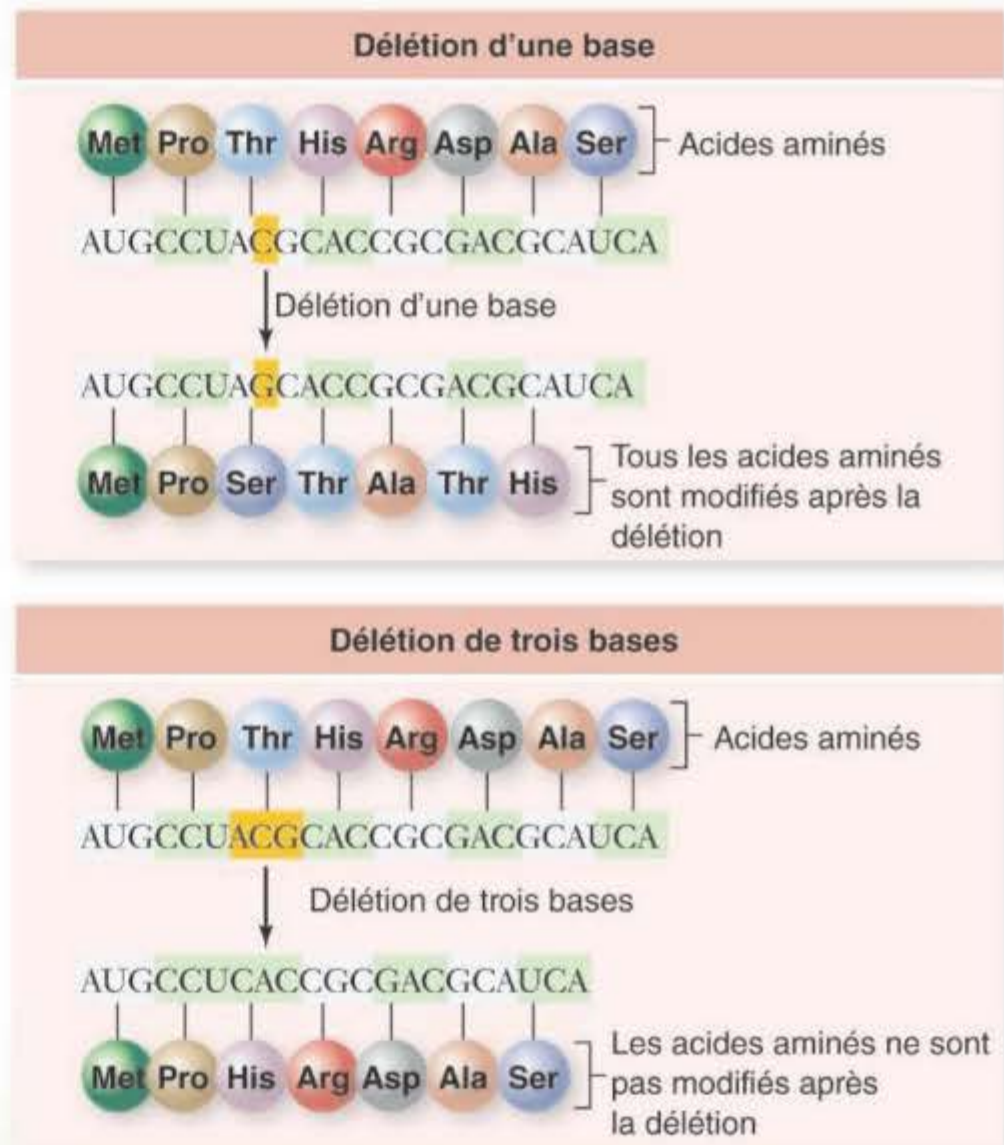
Ces expériences montrent l'importance du **cadre de lecture** pour le message génétique. En l'absence de ponctuation, le cadre de lecture établi par le premier codon de la séquence détermine la manière dont tous les codons suivants sont lus. Les mutations utilisées par Crick et Brenner sont des **mutations avec décalage de phase** : elles modifient le cadre de lecture du message génétique.

DÉMARCHE SCIENTIFIQUE

Hypothèse : le code génétique est lu par groupes de trois bases.

Prédiction : si le code génétique est lu par groupes de trois, la délétion d'une ou deux bases devrait déplacer le cadre de lecture après la délétion. Mais la délétion de trois bases donnerait une protéine avec un seul acide aminé en moins, sans modification en aval.

Test : les mutants ayant perdu une seule base sont collectés, tous avec un phénotype mutant. Trois de ces délétions dans une même région sont combinées pour vérifier les conséquences d'une délétion de trois bases.



Résultat : la combinaison de trois mutations n'a pas le même effet drastique que la perte d'une ou deux bases.

Conclusion : le code génétique est lu par groupes de trois.

Autres expériences : si vous aviez aussi des mutants avec des additions d'une base, quelles seraient les conséquences d'une combinaison d'une délétion et d'une addition ?

Figure 15.4 Le code génétique est formé de triplets.

Nirenberg et d'autres ont déchiffré le code

L'attribution des 64 codons potentiels aux différents acides aminés fut une des principales réalisations de la biochimie du 20^e siècle. Pour réussir ce déchiffrement, deux technologies apparentées étaient nécessaires : (1) des systèmes biochimiques non cellulaires permettant la synthèse des protéines à partir d'un ARN particulier et (2) la possibilité d'obtenir des ARN synthétiques définis pouvant servir dans un système acellulaire.

Pendant une période de cinq ans, de 1961 à 1966, les recherches réalisées principalement par le biochimiste américain Marshall Nirenberg aboutirent à élucider le code génétique. Le groupe de Nirenberg montra d'abord que l'addition de la molécule synthétique d'ARN polyU (une molécule d'ARN formée d'une succession d'uraciles) à leurs systèmes acellulaires produisait un polypeptide, la polyphénylalanine (une succession de phénylalanines). UUU code donc la phénylalanine.

Ils utilisèrent ensuite des enzymes pour produire des polymères d'ARN formés de plusieurs nucléotides. Grâce à ces polymères, ils purent déduire la composition de beaucoup de codons potentiels, mais pas l'ordre des bases de chaque codon.

Les chercheurs purent ensuite utiliser des enzymes pour synthétiser des séquences particulières de trois bases et les tester en liaison avec l'équipement de synthèse protéique. Ce *test de liaison des triplets* leur permit d'identifier 54 des triplets possibles.

Le chimiste organique H. Gobind Khorana apporta la dernière pièce du puzzle en produisant par synthèse organique des molécules artificielles d'ARN de séquence définie et en voyant ensuite les polypep-

tides produits dans les systèmes acellulaires. La combinaison de toutes ces méthodes a permis de spécifier toutes les séquences de trois nucléotides possibles et de déterminer l'ensemble du code génétique (tableau 15.1).

Le code est dégénéré, mais spécifique

Certaines caractéristiques sautent aux yeux dans le tableau 15.1. En premier lieu, 61 des 64 codons possibles servent à spécifier des acides aminés. Trois codons, UAA, UGA et UAG, ont une autre fonction, ce sont les **codons stop**. La seule autre forme de « ponctuation » AVG, qui est un signal de « départ » : c'est le **codon initiateur**. Ce codon a une double fonction, car il code aussi un acide aminé, la méthionine (Met).

On peut voir que les 61 codons sont plus que suffisants pour coder 20 acides aminés. Cela laisse beaucoup de codons en trop. Un moyen de gérer cette abondance serait de n'utiliser que 20 des 61 codons, mais ce n'est pas ce que font les cellules. En réalité, tous les 61 codons sont utilisés, et le code est dégénéré, ce qui signifie que certains acides aminés sont spécifiés par plusieurs codons. On n'a par contre jamais trouvé l'inverse, un même codon ne peut spécifier plusieurs acides aminés.

Cette dégénérescence n'est pas uniforme. Certains acides aminés n'ont qu'un codon, et d'autres jusqu'à six. En outre, la base dégénérée se trouve en général en troisième position dans le codon : les deux premiers nucléotides sont les mêmes et deux des quatre nucléotides situés en troisième position codent le même acide aminé (Le mode de synthèse des

TABLEAU 15.1 Le code génétique

DEUXIÈME LETTRE													
Première lettre	U			C			A			G			Troisième lettre
U	UUU	Phe Phénylalanine	UCU	Ser Sérine	UAU	Tyr Tyrosine	UGU	Cys Cystéine	U				
	UUC		UCC			UAC				UGC	C		
	UUA	Leu Leucine	UCA			UAA	"Stop"			UGA	"Stop"	A	
	UUG		UCG			UAG	"Stop"			UGG	Trp Tryptophane	G	
C	CUU	Leu Leucine	CCU	Pro Proline	CAU	His Histidine	CGU	Arg Arginine	U				
	CUC				CAC		CGC			C			
	CUA				CCA		CAA		Gln Glutamine	CGA	A		
	CUG				CCG		CAG			CGG	G		
A	AUU	Ile Isoleucine	ACU	Thr Thréonine	AAU	Asn Asparagine	AGU	Ser Sérine	U				
	AUC				AAC		AGC			C			
	AUA				AAA	Lys Lysine	AGA		Arg Arginine	A			
	AUG	Met Méthionine ; départ	ACG			AGG	G						
G	GUU	Val Valine	GCU	Ala Alanine	GAU	Asp Aspartate	GGU	Gly Glycine	U				
	GUC				GAC		GGC			C			
	GUA				GAA	Glu Glutamate	GGA			A			
	GUG				GAG		GGG			G			

Un codon comporte trois nucléotides qui sont lus dans l'ordre représenté. Par exemple, ACU code la thréonine. La première lettre, A, est dans la colonne Première lettre, la seconde lettre, C, est dans la colonne Deuxième lettre et la troisième, U, est dans la colonne Troisième lettre. Chaque ARNm est reconnu par une séquence anticodon correspondante d'une molécule d'ARNt. Beaucoup d'acides aminés sont déterminés par plusieurs codons. Par exemple, la thréonine est déterminée par quatre codons qui ne diffèrent que par le troisième nucléotide (ACU, ACC, ACA et ACG).



Figure 15.5 Porc transgénique. Le porcelet de droite est normal. Celui de gauche a été transformé pour exprimer un gène de méduse qui code la protéine à fluorescence verte. La couleur du nez de ce porcelet provient de l'expression de ce gène introduit. Ces animaux transgéniques montrent la nature universelle du code génétique.

protéines sur les ribosomes explique comment fonctionnent ces codons, et sera présenté dans la section 15.7.)

Le code est pratiquement universel, mais pas tout à fait

Le code génétique est le même chez presque tous les organismes. Son universalité est une des meilleures preuves que tous les êtres vivants se partagent un héritage évolutif commun. Le code étant universel, on peut transférer des gènes d'un organisme à un autre et ils peuvent s'exprimer dans leur nouvel hôte (figure 15.5). Cette universalité de l'expression des gènes est au cœur des nombreux progrès du génie génétique décrits au chapitre 17.

En 1979, des chercheurs ont entamé la détermination de toute la séquence des nucléotides des génomes mitochondriaux chez l'homme, la vache et les souris. Il y eut un certain choc quand ces chercheurs constatèrent que le code génétique utilisé par ces mitochondries de mammifères n'était pas tout à fait le même que le « code universel » devenu si familier pour les biologistes.

Dans les génomes mitochondriaux, UAG, qui aurait dû être un codon stop, codait le tryptophane, AUA codait la méthionine au lieu de l'isoleucine, AGA et AGG ne codaient pas l'arginine, mais étaient des codons stop. On a aussi trouvé de petites différences par rapport au code universel dans les génomes des chloroplastes et des ciliés (groupes de protistes).

Il apparaît donc que le code génétique n'est pas tout à fait universel. Il y a un certain temps, probablement après avoir entamé leur existence endosymbiotique, les mitochondries et les chloroplastes ont commencé à lire différemment le code, en particulier sa partie associée aux signaux « stop ».

Question Pourquoi pensez-vous que le code génétique est presque universel ?

Questions d'apprentissage 15.2

On a montré que le code génétique est continu et consiste en triplets de nucléotides avec deux formes de ponctuation : trois bases codent un acide aminé, et les groupes de trois sont lus dans l'ordre. Sur les 64 codons, 61 spécifient des acides aminés, l'un d'eux code aussi l'initiation et trois codons indiquent un « stop ». Certains acides aminés correspondant à plusieurs codons, on dit que le code est dégénéré. Tous les codons ne codent cependant qu'un acide aminé.

- Qu'arriverait-il si un codon spécifiait plusieurs acides aminés ?

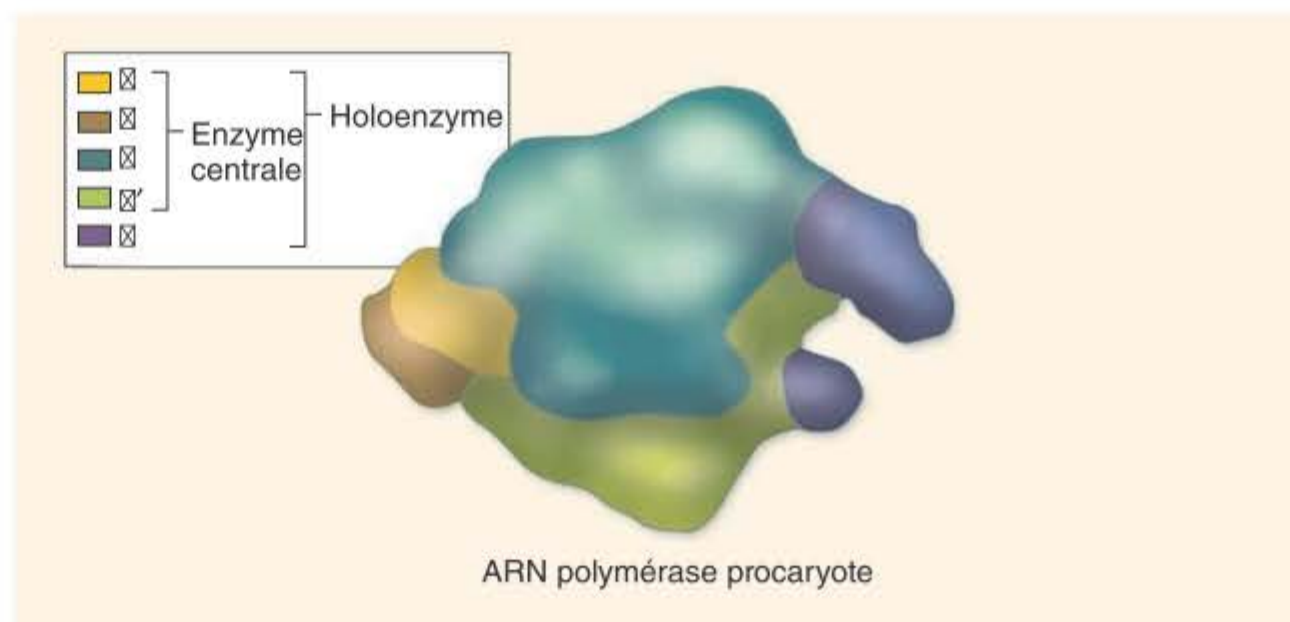
15.3 La transcription des procaryotes

Objectifs

1. Décrire la transcription chez les bactéries.
2. Montrer les différences entre initiation, élongation et terminaison de la transcription.
3. Définir les caractéristiques particulières de la transcription des procaryotes.

Nous entamons l'examen de l'expression génique par une description du processus de transcription chez les procaryotes. La transcription eucaryote sera décrite ensuite en mettant l'accent sur ses différences avec celle des procaryotes.

Figure 15.6 ARN polymérase bactérienne et transcription. *a.* Il existe deux formes de l'ARN polymérase ; la polymérase centrale et la holoenzyme. *b.* La sous-unité σ de la holoenzyme reconnaît des éléments du promoteur aux positions -35 et -10 et s'unit à l'ADN. L'hélice s'ouvre à la région -10 et la transcription commence au site de départ à +1.



a.

Les procaryotes ont une seule ARN polymérase

L'unique ARN polymérase des procaryotes existe sous deux formes : la *polymérase centrale* et la *holoenzyme*. La polymérase centrale peut synthétiser l'ARN en utilisant l'ADN matrice, mais ne peut entamer correctement la synthèse. La holoenzyme en est capable.

La polymérase centrale est composée de quatre sous-unités : deux sous-unités β identiques, les sous-unités β et β' (figure 15.6a). Les deux sous-unités α assurent la stabilité du complexe et peuvent s'unir aux molécules de régulation. Le site actif de l'enzyme est formé des deux sous-unités β et β' , qui s'unissent à l'ADN modèle et aux précurseurs de ribonucléotide triphosphates.

La holoenzyme, capable d'entamer correctement la synthèse, est formée par l'addition d'une sous-unité σ à la polymérase centrale (figure 15.6a). Étant capable de reconnaître des signaux spécifiques dans l'ADN, elle permet à l'ARN polymérase de trouver l'origine des gènes, ce qui est essentiel pour son fonctionnement. Notez que l'initiation de la synthèse de l'ARNm n'a pas besoin d'une amorce, contrairement à la réplication de l'ADN.

L'initiation se situe au niveau des promoteurs

L'initiation correcte de la transcription nécessite deux sites de l'ADN : l'un est le **promoteur**, site de reconnaissance et de liaison pour l'ARN polymérase, et le **site de départ** lui-même. La polymérase réclame aussi un signal pour terminer la transcription, c'est le **terminateur**. La région située entre le promoteur et le terminateur est l'**unité de transcription**.

On peut comparer la polymérase se déplaçant le long de l'ADN à l'eau qui s'écoule dans une rivière. Nous pouvons parler de sites situés « en amont » ou « en aval » du site de départ sur l'ADN. Nous pouvons aussi appliquer cette comparaison à un système simple de numérotation des bases de l'ADN par rapport à leur position dans l'unité de transcription. La première base transcrite est désignée comme **+1**, et cette numérotation se poursuit en aval jusqu'à la dernière base transcrite. Toutes les bases situées en amont du site de départ reçoivent des valeurs négatives, à partir de **-1**.

Le promoteur est une courte séquence située en amont du site de départ et n'est donc pas transcrite par la polymérase. Deux séquences de 6 bases sont communes aux promoteurs bactériens : l'une est située

35 nucléotides en amont du site de départ (-35) et l'autre se trouve à 10 nucléotides de ce site (-10) (figure 15.6b). Ces deux sites font que le promoteur est asymétrique, ils indiquent non seulement le site d'initiation, mais aussi le sens de la transcription.

L'union de l'ARN polymérase au promoteur est la première étape de la transcription. Cette union est contrôlée par la sous-unité σ de l'ARN polymérase holoenzyme, qui reconnaît la séquence -35 du promoteur et positionne l'ARN polymérase au site de départ correct, orientée pour transcrire dans la bonne direction. Unie au promoteur, l'ARN polymérase commence à dérouler l'hélice d'ADN au site -10 (voir figure 15.6b). La polymérase couvre une région de quelque 75 pb, mais elle n'en déroule que 12-14.

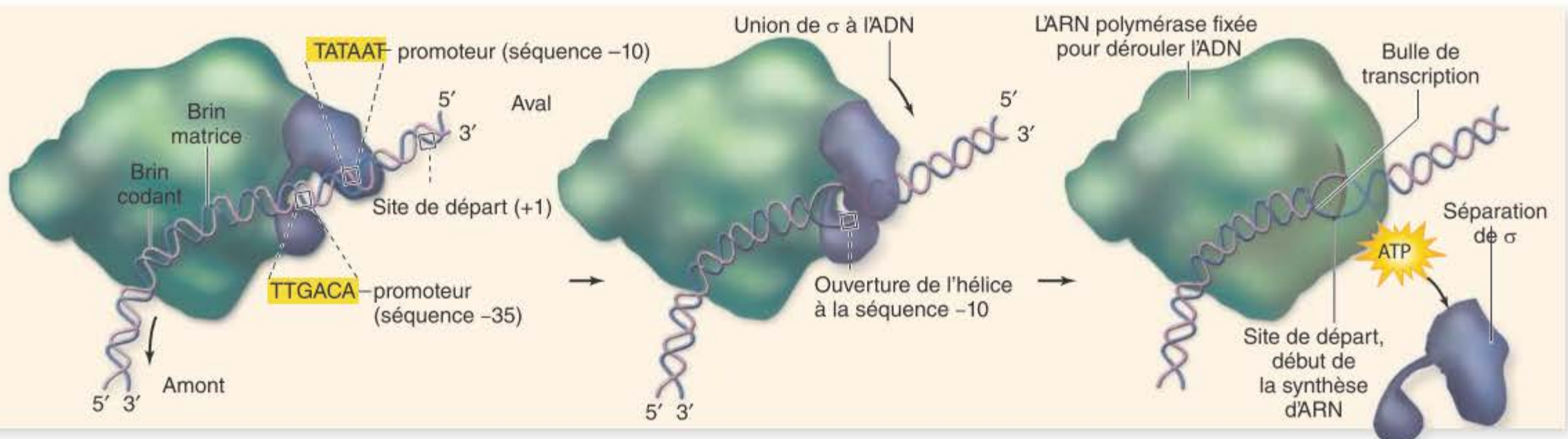
Question Le promoteur procaryote possède deux éléments distincts non identiques. Quelle en est l'importance pour l'initiation de la transcription ?

L'élongation ajoute les nucléotides successifs

Chez les procaryotes, la transcription de la chaîne d'ARN débute en général par ATP ou GTP. L'un des deux devient l'extrémité 5' de la chaîne, qui s'allonge dans le sens 5'-3' par l'addition de ribonucléotides. Quand l'ARN polymérase quitte la région promotrice, le facteur σ n'est plus nécessaire, mais il reste associé à l'enzyme.

Cet éloignement du promoteur, le *dégagement*, ou *échappée*, n'implique pas seulement une synthèse des premiers nucléotides du transcrit et une progression, car l'enzyme est en contact étroit avec l'ADN pendant l'initiation. Elle doit rompre ces contacts avec la région promotrice pour pouvoir descendre progressivement sur le modèle. L'enzyme y arrive par des changements de conformation pendant ce stade de dégagement et elle est ensuite moins étroitement liée à l'ADN que pendant l'union initiale au promoteur.

La région contenant l'ARN polymérase, l'ADN matrice et l'ARN en croissance est appelée **bulle de transcription**, en raison de la présence d'une « bulle » d'ADN localement déroulé (figure 15.7). Dans cette bulle, les 9 premières bases du jeune brin d'ARN forment temporairement une hélice avec le brin d'ADN matrice. La position de l'extrémité 3' de l'ARN est ainsi stabilisée et elle peut interagir avec un ribonucléotide triphosphate entrant. L'enzyme elle-même couvre une cinquantaine de pb d'ADN autour de cette bulle de transcription.



b.

La bulle formée par l'ARN polymérase descend le long de l'ADN à une vitesse constante d'environ 50 nucléotides par seconde, le brin d'ARN en croissance émergeant de la bulle. Après le passage de la bulle de transcription, l'ADN transcrit s'enroule à nouveau en la quittant.

La terminaison a une localisation spécifique

L'extrémité de l'unité de transcription bactérienne est marquée par des séquences de terminaison qui disent « stop » à la polymérase. L'arrivée à ces séquences arrête la formation des liaisons phosphodiester, provoque la dissociation de l'hybride ARN-ADN dans la bulle de transcription, la libération de l'ARN polymérase et la reconstitution de l'hélice d'ADN.

Les terminateurs les plus simples consistent en une série de paires de bases G-C suivie d'une série de paires A-T. Le transcrit d'ARN de cette région stop peut former une structure bicaténaire dans la région GC, une *épingle à cheveux* suivie d'au moins quatre ribonucléotides uracile (U) (figure 15.8). La formation de l'épingle à cheveux arrête l'ARN polymérase en la plaçant directement sur la série de quatre uraciles. L'appariement de U avec le A de l'ADN donne la plus faible des quatre paires de bases hybrides et n'est pas suffisant pour conserver les brins hybrides liés pendant la pause de la polymérase. Le brin d'ARN se sépare donc de l'ADN dans la bulle de ces terminateurs et aide à mettre fin à la transcription.

La transcription procaryote est couplée à la traduction

Chez les procaryotes, l'ARNm produit par la transcription commence à être traduit avant la fin de la transcription – elle lui est donc *couplée*

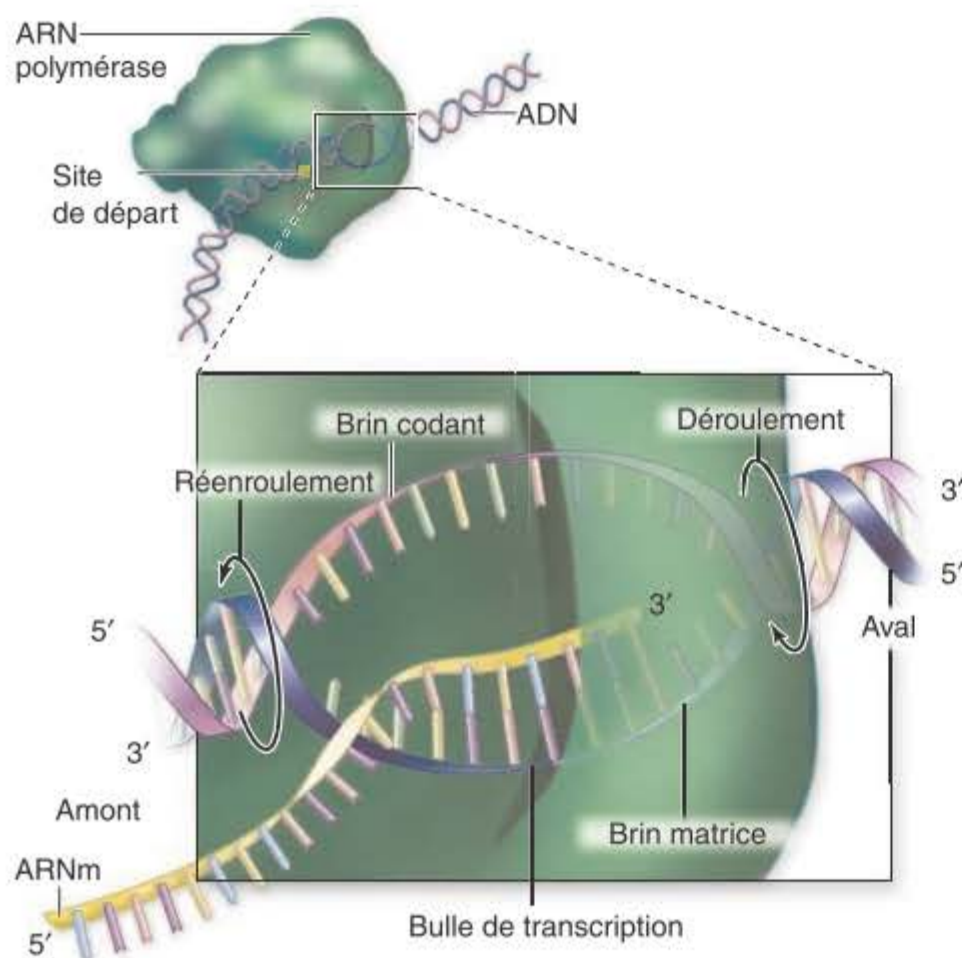


Figure 15.7 Schéma d'une bulle de transcription. Le duplex d'ADN est déroulé par le complexe de l'ARN polymérase et reforme sa spirale en le quittant. Un des brins d'ADN fonctionne comme matrice et les nucléotides libres s'ajoutent au bout de l'ARN en croissance. Il y a une courte région hybride ADN-ARN au sein de la bulle.

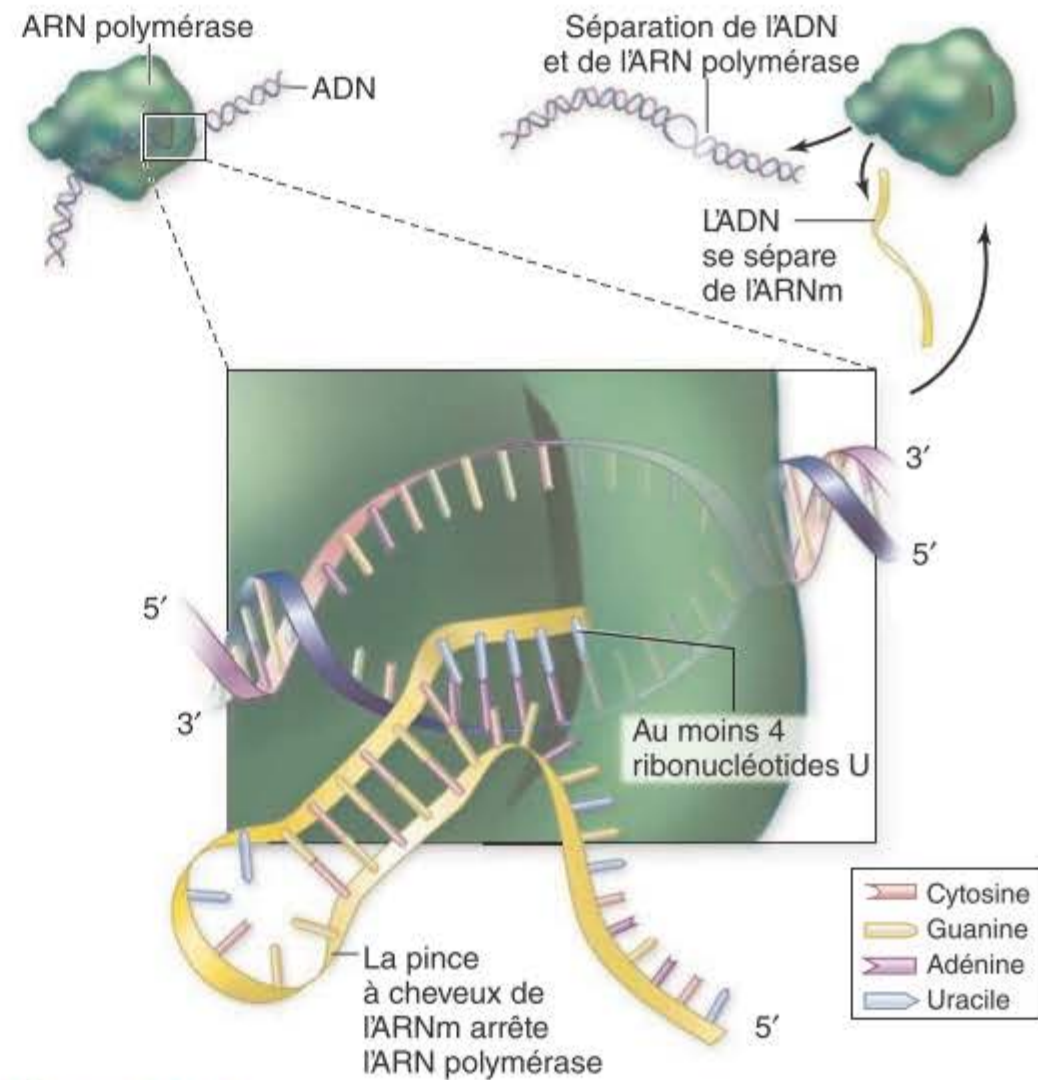


Figure 15.8 Terminateur bactérien de la transcription. La région complémentaire G-C forme une tige bicaténaire et une boucle monocaténaire, l'épingle à cheveux. La séquence de U forme un hybride ADN-ARN moins stable qui libère l'enzyme.

(figure 15.9). Dès qu'une extrémité 5' est disponible, des ribosomes s'y fixent pour entamer la traduction. (Ce couplage est impossible chez les eucaryotes parce que la transcription se déroule dans le noyau et la traduction dans le cytoplasme.)

Une autre différence entre l'expression génique des procaryotes et des eucaryotes provient du fait que les ARNm des procaryotes renferment plusieurs gènes. Les gènes procaryotes sont souvent organisés de façon à grouper ceux qui codent des fonctions apparentées. Ce groupe de gènes fonctionnellement apparentés est un **opéron**. Un opéron est une unité de transcription unique codant plusieurs enzymes nécessaires à une voie biochimique. Le groupement des gènes par fonction permet une régulation commune : nous y reviendrons dans le prochain chapitre.

Questions d'apprentissage 15.3

Chez les bactéries, la transcription est effectuée par l'ARN polymérase, dont il existe deux formes : une polymérase centrale et une holoenzyme. La forme holoenzyme effectue l'initiation, elle peut reconnaître correctement les séquences promotrices. L'élongation est une synthèse d'ARN par l'enzyme centrale, qui ajoute des nucléotides d'ARN les uns aux autres jusqu'à son arrivée à un terminateur, où la synthèse s'arrête et le transcrit est libéré. Chez les procaryotes, la traduction de l'ARN débute avant la fin de la transcription, les deux processus étant couplés.

- Les levures sont des organismes unicellulaires comme les bactéries ; pensez-vous qu'elles ont le même couplage transcription/traduction ?

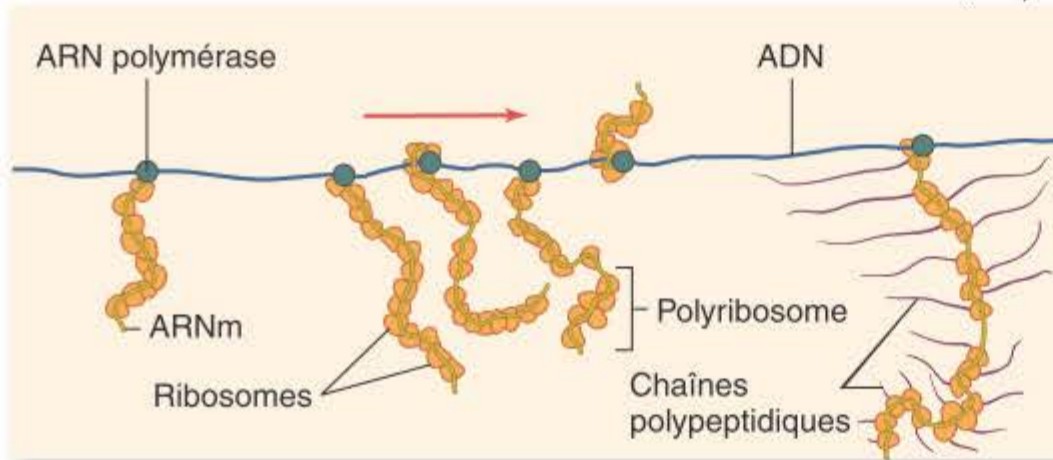
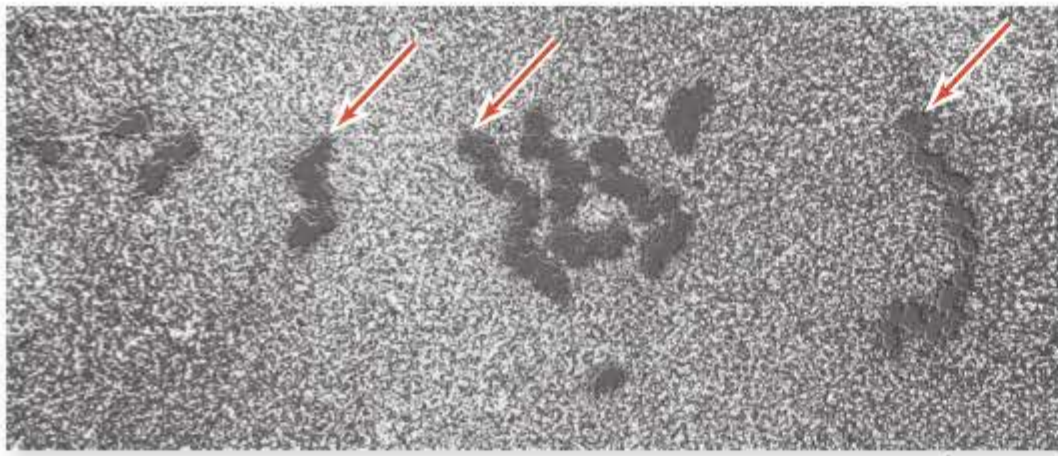


Figure 15.9 La transcription et la traduction sont couplées chez les procaryotes. Dans cette photomicrographie de l'expression génique chez *E. coli*, la traduction se déroule pendant la transcription. Les flèches indiquent les ARN polymérases, et les ribosomes sont attachés aux ARNm partant de la polymérase. Les polypeptides synthétisés par les ribosomes sont invisibles sur la photo ; ils ont été rajoutés au dernier ARNm du dessin.

15.4 La transcription de l'eucaryote

Objectifs

1. Montrer les différences entre les promoteurs des trois polymérases eucaryotes.
2. Décrire la maturation des transcrits eucaryotes.
3. Expliquer les différences entre l'élongation et la transcription chez les bactéries et les eucaryotes.

Le mécanisme de base de la transcription par l'ARN polymérase est le même dans tous les systèmes. En fait, les ARN polymérases des bactéries, des archées et des eucaryotes dérivent toutes d'une même enzyme ancestrale. Cependant, la production précise de l'ARN chez les eucaryotes est suffisamment différente pour être décrite. Une différence est particulièrement évidente : les bactéries n'ont qu'une ARN polymérase, alors que les eucaryotes ont trois ARN polymérases de structure et de fonction différentes. L'ARN polymérase I transcrit l'ARNr, l'ARN polymérase II transcrit l'ARNm et certains petits ARN nucléaires et l'ARN polymérase III transcrit l'ARNt et d'autres petits ARN. Ensemble, ces trois enzymes effectuent toutes les transcriptions dans le noyau des cellules eucaryotes.

Chaque polymérase possède son propre promoteur

Les trois ARN polymérases eucaryotes ont besoin de systèmes de contrôle différents dans l'ADN pour permettre à chacune de savoir où

entamer la transcription. Chaque polymérase reconnaît une structure promotrice différente.

Promoteurs de l'ARN polymérase I

Les promoteurs de l'ARN polymérase I ont d'abord posé un problème aux biologistes, parce qu'on ne trouvait pas de point commun entre les gènes d'ARNr chez les différentes espèces en dehors de la région codante. On pense aujourd'hui que ces promoteurs sont propres à chaque espèce, ce qui explique l'absence de séquences conservées lors de comparaisons entre espèces.

Promoteurs de l'ARN polymérase II

La plupart des premiers gènes eucaryotes clonés renfermaient une séquence, TATA, située en amont du site de départ. La similitude entre la séquence TATA et l'élément -10 des procaryotes a fait penser que c'était le promoteur. Cela s'est avéré trop simple et l'on est arrivé à l'idée d'un « promoteur central », qui peut être formé de plusieurs éléments distincts, dont la boîte TATA.

Promoteurs de l'ARN polymérase III

Les promoteurs de l'ARN polymérase III causèrent aussi la surprise chez les biologistes qui étudiaient le contrôle de l'expression génique. Pour analyser les régions régulatrices, une technique fréquente consistait à produire une succession de délétions à partir de l'extrémité 5' des gènes jusqu'à la suppression d'une transcription spécifique. Pour les gènes de l'ARNt cependant, les délétions 5' n'avaient pas d'effet sur l'expression ! On a constaté que les promoteurs se trouvaient en réalité à l'intérieur des gènes eux-mêmes. On n'a pas prouvé que c'était le cas de tous les gènes de la polymérase III, mais ce l'est pour la plupart.

Le promoteur central s'unit à des facteurs qui recrutent l'ARN Pol II

L'initiation de la transcription au niveau des promoteurs de l'ARN polymérase II correspond à l'initiation chez les procaryotes, mais, au lieu d'utiliser un seul facteur permettant l'identification du promoteur, les eucaryotes utilisent une foule de facteurs de transcription. Ces facteurs fonctionnent ensemble pour associer l'ARN polymérase II à un promoteur. Ils interagissent avec l'ARN polymérase II pour former un complexe d'initiation au niveau du promoteur (figure 15.10). Nous analyserons ce complexe en détail au chapitre 16 dans le cadre du contrôle de l'expression génique.

Quand le transcrit atteint quelque 20 nucléotides, il est modifié par l'addition de GTP au groupement PO_4^- 5', pour donner ce que l'on a appelé la **coiffe 5'** (figure 15.11). Cette coiffe est unie au transcrit par son extrémité 5' ; c'est la seule liaison 5'-5' observée dans les acides nucléiques. Le G du GTP est en outre modifié par l'addition d'un groupement méthyle, et on le désigne souvent comme une **coiffe méthyle-G**. Cette coiffe est importante pour la traduction, la stabilité de l'ARN et pour la maturation ultérieure.

Un autre événement est possible au début de l'élongation : Pol peut faire une pause. Cela se produit en général dans les 50 pb qui suivent l'initiation, et il semble maintenant qu'il s'agit d'un phénomène fréquent. On a d'abord pensé à des gènes de choc thermique, permettant l'induction rapide de la transcription après un choc thermique, mais un examen global montre que la majorité des gènes des organismes pluricellulaires peuvent héberger une polymérase à l'arrêt. Ces résultats sont difficiles à interpréter parce qu'il est impossible de distinguer les

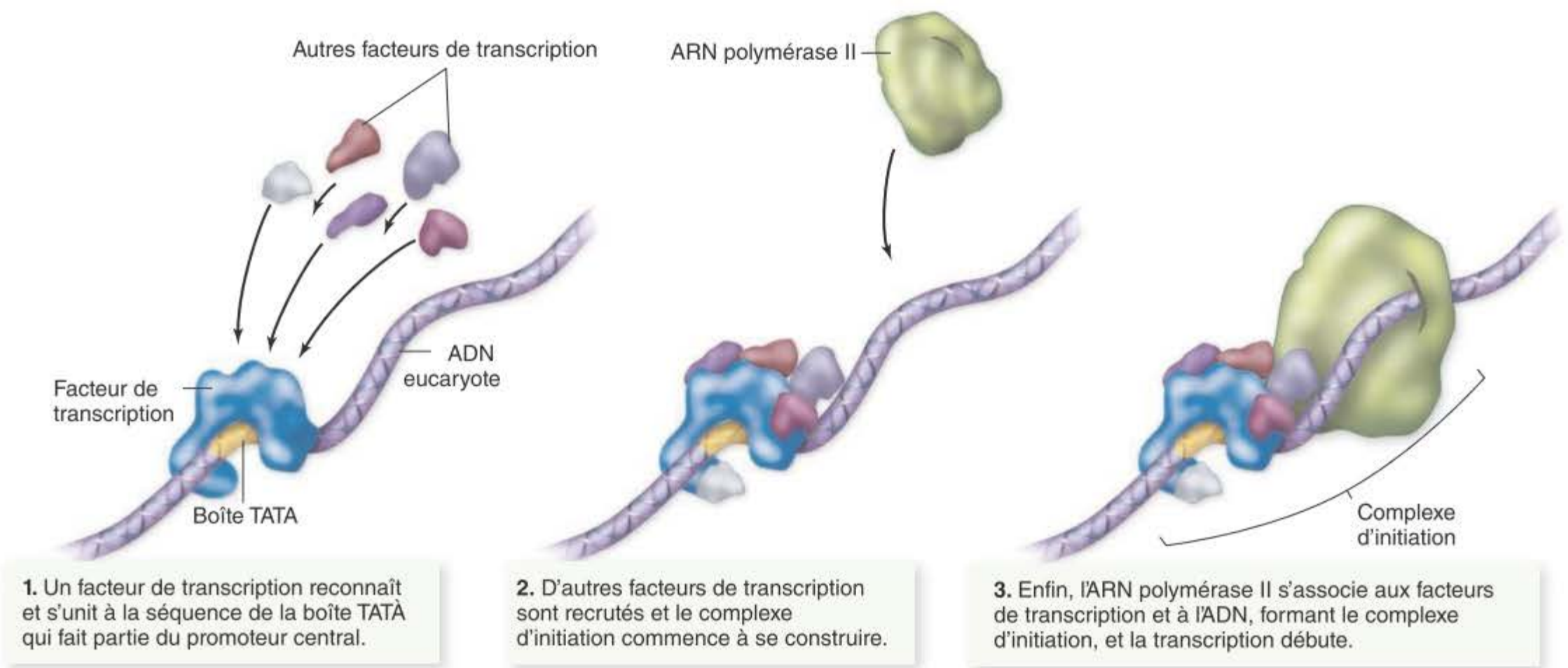


Figure 15.10 Le complexe d'initiation eucaryote. À la différence de la transcription dans les cellules procaryotes, où l'ARN polymérase reconnaît le promoteur et s'y fixe, la transcription eucaryote exige la fixation de facteurs de transcription au promoteur avant l'union de l'ARN polymérase II à l'ADN. L'association de facteurs de transcription et de l'ARN polymérase II au promoteur constitue le complexe d'initiation.

petits transcrits avortés d'une Pol à l'arrêt susceptible de repartir. Ce redémarrage implique la participation et la coordination des éléments du promoteur, des facteurs d'élongation, du nucléosome et même de l'enzyme responsable de la coiffe. Pour le moment, la relation n'est pas claire entre le contrôle de l'initiation du transcrit et la fin de l'arrêt du promoteur voisin.

Le complexe d'élongation de la transcription peut recruter d'autres facteurs

Parmi les facteurs d'élongation de la transcription, certains arrivent après l'initiation. Cela semble coordonné par un élément structural de la grande sous-unité, le domaine carboxy terminal (DCT). Le DCT est composé de 7 acides aminés répétés de nombreuses fois (52 répétitions chez les humains). On y trouve des sérines pouvant être phosphorylées, et il fonctionne comme une plateforme qui recrute des facteurs interve-

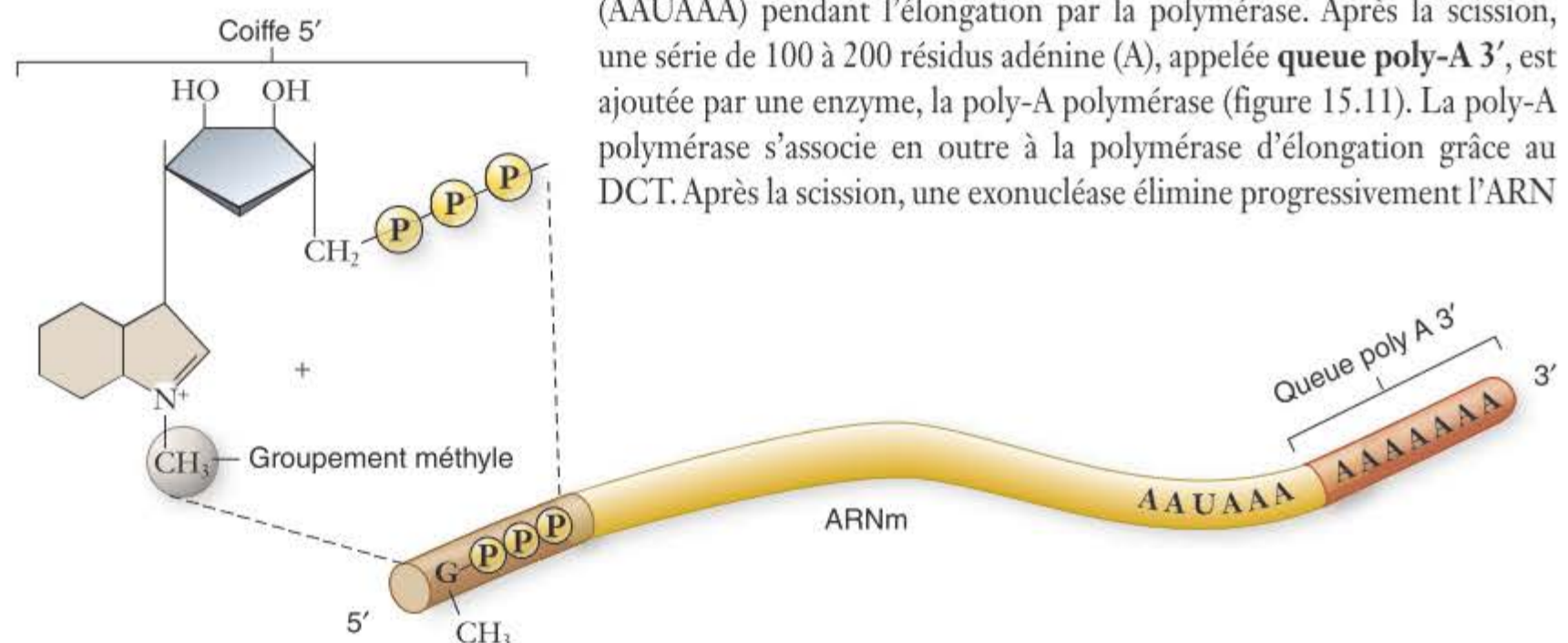
nant dans l'élongation et des enzymes modifiant l'ARN. Par exemple, l'enzyme pour la coiffe, dont il a été question, interagit avec le DCT immédiatement après l'initiation.

Certains des facteurs d'élongation qui interagissent avec le DCT interviennent dans le redémarrage et dans la progression dans les nucléosomes. Les premières études *in vitro* montraient que l'ARN Pol ne pouvait pas transcrire l'ADN condensé en chromatine. Il est clair que ce n'est pas le cas *in vivo*, et cela s'explique par le fait que les facteurs d'élongation interagissent avec Pol et avec les nucléosomes et repositionnent la polymérase. Parmi ces facteurs, on trouve le DCT (il aide à la transcription de la chromatine), des facteurs qui modifient les histones et des complexes de remodelage de la chromatine (décrits au chapitre 16).

La polyadénylation participe à la terminaison des transcrits d'ARNm

Une différence majeure entre les procaryotes et les eucaryotes est le fait que, chez les eucaryotes, l'extrémité du transcrit n'est pas celle de l'ARNm. Le transcrit eucaryote est scindé en amont à un site spécifique (AAUAAA) pendant l'élongation par la polymérase. Après la scission, une série de 100 à 200 résidus adénine (A), appelée **queue poly-A 3'**, est ajoutée par une enzyme, la poly-A polymérase (figure 15.11). La poly-A polymérase s'associe en outre à la polymérase d'élongation grâce au DCT. Après la scission, une exonucléase élimine progressivement l'ARN

Figure 15.11 Modifications après la transcription aux extrémités 5' et 3'. Les molécules d'ARNm eucaryotes sont modifiées dans le noyau par l'addition d'un GTP méthylé à l'extrémité 5' du transcrit, la coiffe 5', et d'une longue chaîne de résidus adénine à son extrémité 3', la queue poly-A.



en aval, séparant ainsi le complexe de transcription de la matrice. La queue poly-A semble jouer un rôle dans la stabilité des ARNm en les protégeant de la dégradation (voir chapitre 16). La terminaison des transcrits non polyadénylés repose sur un autre mécanisme impliquant d'autres facteurs.

Les transcrits primaires sont modifiés pour donner les ARNm

Les transcrits eucaryotes peuvent comporter des séquences non codantes qui doivent être éliminées pour aboutir à l'ARNm final. Ce processus, appelé épissage du préARNm, est effectué par un complexe dénommé **splicéosome**. Ce sujet est décrit à la section 15.5. D'après ce que nous venons de voir, il est clair que l'ARNm des eucaryotes est différent du transcrit primaire. Avec une extrémité 5' modifiée, une extrémité 3' synthétisée par une autre enzyme et l'élimination des introns, l'ARNm est une forme profondément modifiée de transcrit primaire.

À première vue, ce processus semble totalement différent de celui des procaryotes, où la transcription et la traduction sont couplées. Il est intéressant de constater que toutes les modifications du transcrit surviennent pendant la transcription, comme c'est le cas pour la traduction chez les procaryotes.

Questions d'apprentissage 15.4

Les eucaryotes ont trois ARN polymérase, les polymérase I, II et III. Chacune synthétise un ARN différent et reconnaît son propre promoteur. Le promoteur de la polymérase I est spécifique de l'espèce. Le promoteur central de la polymérase II comprend de nombreux éléments. Le promoteur de la polymérase III est situé à l'intérieur du gène. La polymérase II est responsable de la synthèse de l'ARNm. L'élongation par pol II exige différents facteurs et des enzymes induisent des modifications pendant la transcription pour produire une coiffe, polyadényler et épisser le transcrit.

- La complexité du génome eucaryote nécessite-t-il trois polymérase ?

15.5 L'épissage des pré-ARNm

Objectifs

1. Expliquer la relation entre les gènes et les protéines chez les procaryotes et les eucaryotes.
2. Décrire la réaction d'épissage des pré-ARNm.
3. Montrer comment l'épissage modifie la nature des gènes.

Les premiers gènes isolés étaient des gènes procaryotes d'*E. coli* et de ses virus. On a clairement compris la nature et une partie du contrôle de l'expression génique dans ces systèmes avant l'isolement de gènes eucaryotes. On s'attendait à une expression génique essentiellement semblable chez les eucaryotes, avec éventuellement des différences de détail. L'isolement des premiers gènes d'organismes eucaryotes a provoqué un choc dans le monde de la biologie.

Il peut y avoir des interruptions dans les gènes eucaryotes

Beaucoup de ces gènes paraissaient contenir des séquences que l'on ne retrouvait pas dans l'ARNm. Cette découverte était totalement inattendue. En biologie moléculaire, un principe fondamental découlant des travaux sur *E. coli* était la colinéarité du gène et de la protéine produite ; autrement dit, la séquence des bases du gène correspond à la séquence des bases de l'ARNm, qui correspond elle-même à la séquence des acides aminés de la protéine.

Chez les eucaryotes, les gènes sont interrompus par des séquences qui ne sont représentées ni dans l'ARNm ni dans la protéine. Le terme « gène fractionné » fut utilisé à l'époque, mais la nomenclature qui a prévalu correspond à la nature inattendue de ces séquences. On désigne comme « séquences intermédiaires », ou **introns**, les séquences non codantes de l'ADN qui interrompent la séquence du gène et **exons** les séquences codantes parce qu'elles s'expriment (figure 15.12).

Le splicéosome est l'organite d'épissage

Il reste vrai que l'ARNm eucaryote mature et la protéine correspondante sont colinéaires, mais ce n'est pas le cas pour le gène. Imaginez que vous regardez une autoroute à partir d'un satellite. Des voitures sont

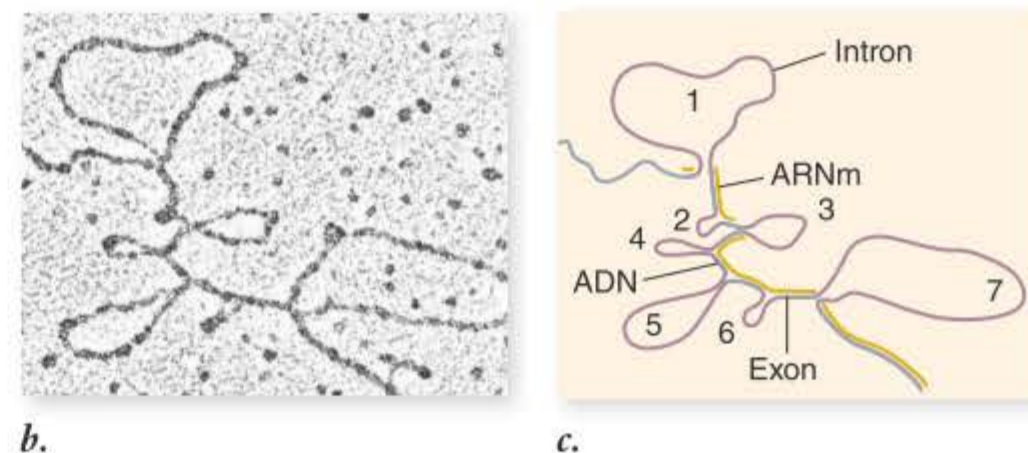
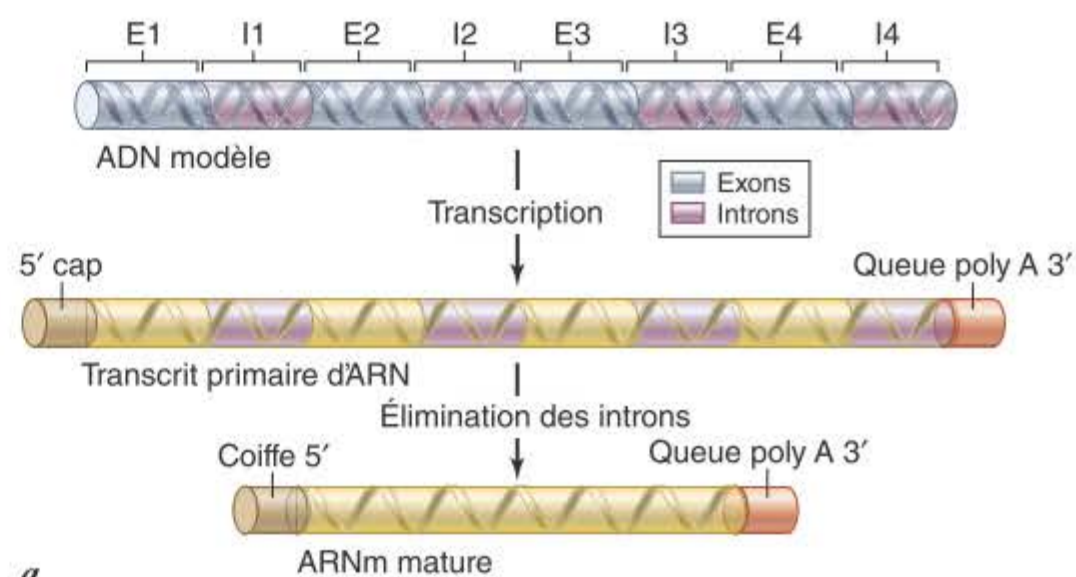


Figure 15.12 Les gènes eucaryotes contiennent des introns et des exons. *a.* Les gènes eucaryotes renferment des séquences codantes, les exons, et des séquences intercalaires, les introns. *b.* Photomicrographie électronique montrant les hybrides formés par l'ARNm et l'ADN du gène de l'ovalbumine, avec sept introns. Les introns de la séquence d'ADN n'ont pas de correspondants dans la séquence d'ARNm et forment donc sept boucles. *c.* Dessin correspondant.

Analyse des données Montrez comment le gène de la figure 15.12 pourrait coder plusieurs transcrits.

dispersées au hasard tout au long du ruban de béton, certaines roulant en groupes, d'autres isolées ; la plus grande partie de la route serait vide. C'est à cela que ressemble un gène eucaryote – des exons dispersés entourés des séquences beaucoup plus longues des introns.

Chez l'homme, 1 à 1,5 % seulement du génome est consacré aux exons qui codent des protéines ; 24 % correspondent aux introns non codants entourant ces exons.

L'épissage

Voici la bonne question : que font les cellules eucaryotes des introns non codants ? La réponse est que le transcrite primaire est découpé et transformé en ARNm mature. Ce processus est l'**épissage du pré-ARNm** ; il se fait dans le noyau avant l'exportation de l'ARNm vers le cytoplasme.

Les jonctions intron-exon sont identifiées par de **petites particules de ribonucléoprotéines nucléaires**, les **RNPsn**. Les RNPsn sont des complexes composés d'ARNsn et de protéines. Elles s'associent ensuite à d'autres protéines en un complexe plus volumineux, le **splicéosome**, responsable de l'épissage, ou élimination des introns.

Pour que l'épissage soit correct, le splicéosome doit pouvoir identifier les jonctions intron-exon. Tous les introns débutent par la même séquence de deux bases et se terminent par une autre séquence de deux bases qui les marque pour être éliminés. Il existe en outre, à l'intérieur de l'intron, un nucléotide A conservé, le *point de ramification*, important pour l'épissage (figure 15.13).

L'épissage débute par une scission de l'extrémité 5' de l'intron. Cette extrémité 5' se fixe au groupement OH 2' du point de ramification et forme une structure ramifiée appelée *lasso* pour sa ressemblance avec le lasso des cow-boys (figure 15.13). L'extrémité 3' du premier exon déplace ensuite l'extrémité 3' de l'intron, réunit les deux exons et libère l'intron comme un lasso.

L'épissage se déroule pendant la transcription. Le transcrite amène l'enzyme près du DCT, permettant le recrutement des facteurs d'épissage et des RNP du splicéosome. Il semble que la vitesse de la transcription, qui n'est pas constante, peut intervenir dans le choix du site d'épissage. Les sites d'épissage peuvent devenir inaccessibles à la suite de l'arrêt de l'enzyme et de la formation d'une structure secondaire dans le transcrite.

Répartition des introns

Aucune règle ne régit le nombre d'introns par gène, ni la taille des introns et des exons. Certains gènes n'ont pas d'introns ; d'autres peuvent en avoir 50. La taille des exons va de quelques nucléotides à 7500, et la taille des introns est tout aussi variable. La présence des introns explique en partie pourquoi les « séquences codantes » ne constituent en réalité qu'une si faible partie du génome eucaryote (voir, au chapitre 18, les résultats du Projet Génome Humain).

Pour expliquer l'existence des introns, on pense que les exons représentent les domaines fonctionnels des protéines et que les associations intron-exon trouvées dans les gènes proviennent du brassage de ces unités fonctionnelles au cours de longues périodes d'évolution. Cette hypothèse du *brassage des exons* a été proposée peu après la découverte des introns et elle a fait l'objet de nombreux débats pendant des années.

La masse récente de données génomiques a donné un nouvel éclairage à ce sujet en permettant d'analyser statistiquement la localisation des introns et la structure intron-exon. Cette analyse a conforté l'hypothèse du brassage des exons pour beaucoup de gènes ; ce n'est cependant pas vraiment universel, car toutes les protéines n'obéissent pas à ce modèle. Il est possible que l'origine des introns ne soit pas unique et qu'une seule hypothèse ne suffise pas.

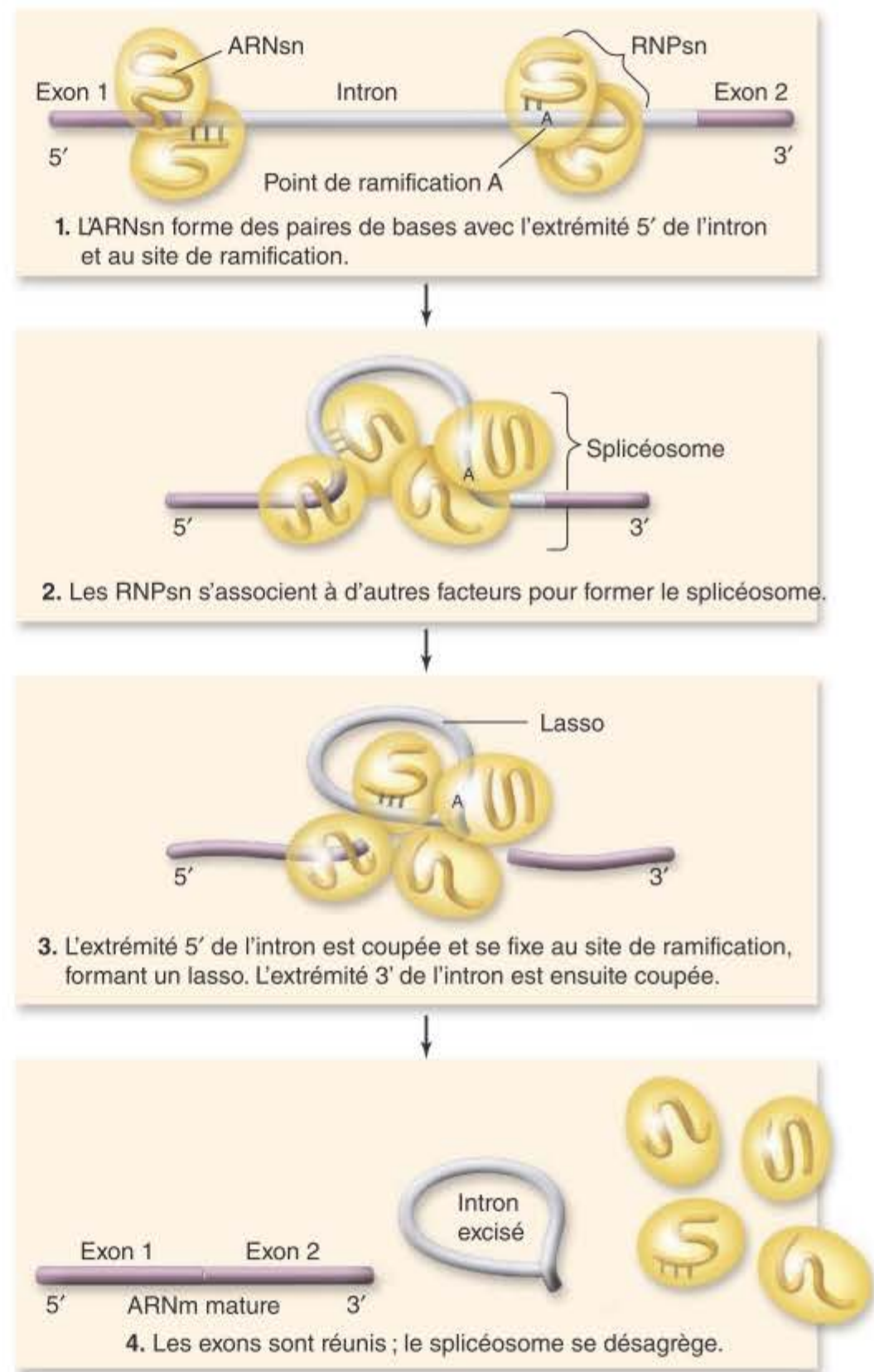


Figure 15.13 Épissage du pré-ARNm par le splicéosome. Les particules de RNPsn contiennent l'ARNsn qui réagit avec l'extrémité 5' d'un intron et avec un site de ramification interne à cet intron. Plusieurs RNPsn s'unissent à d'autres protéines pour former le splicéosome. L'intron formant une boucle, l'extrémité 5' est coupée et unie à un site proche de son extrémité 3'. L'intron forme un lasso qui est excisé et les exons sont épissés. Le splicéosome se désorganise ensuite et libère l'ARNm.

L'épissage peut donner plusieurs transcrits d'un même gène

Une conséquence du processus d'épissage est la complexité plus grande de l'expression génique chez les eucaryotes. Un même transcrite primaire peut être épissé en ARNm différents par l'inclusion de lots différents d'exons : c'est l'**épissage alternatif**.

On a démontré l'importance du mode normal d'épissage pour le fonctionnement de l'organisme. On estime que près de la moitié des maladies génétiques connues chez l'homme sont dues à un mauvais épissage. Les mutations dans les signaux d'épissage peuvent être à l'origine de nouveaux sites d'épissage ou peuvent supprimer des sites normaux. (Au chapitre 16, nous verrons comment l'épissage alternatif peut réguler l'expression génique.)

15.6 Structure de l'ARNt et des ribosomes

Objectifs

1. Expliquer pourquoi le chargement de l'ARNt est important pour la traduction.
2. Identifier les sites de liaison dans le ribosome.

Grâce au séquençage du génome humain et à la grande efficacité des systèmes de séquençage des transcrits, on a pu comparer au génome quantité de transcrits provenant de tissus différents. En se basant sur l'analyse des transcrits de l'homme et de la souris, on estime que 95 % des gènes de mammifères donnent des transcrits alternatifs, avec une moyenne de 4 transcrits par gène. D'autres estimations du nombre moyen de transcrits par gène ont donné des valeurs supérieures ou inférieures. Un autre travail récent a abouti aux mêmes niveaux globaux d'épissage alternatif, avec jusqu'à 25 transcrits différents pour certains gènes. Cette même analyse a cependant montré que la plupart des gènes n'ont qu'un seul transcrit principal. Nous restons en présence d'une image incomplète, mais les recherches à venir devraient la clarifier.

Il est important de noter que ces analyses sont des estimations globales appliquant les dernières méthodes de séquençage à l'analyse de populations d'ARN provenant de tissus différents. Les fonctions potentielles des protéines correspondant à ces formes n'ont été étudiées que pour un faible pourcentage des gènes potentiellement épissés. Ces analyses expliquent cependant en partie comment les quelque 20 000 gènes humains peuvent coder plus de 100 000 protéines différentes que l'on estime présentes dans les cellules humaines.

Questions d'apprentissage 15.5

Chez les procaryotes, les gènes sont colinéaires avec les protéines produites. Par contre, les gènes eucaryotes renferment des exons, régions exprimées, et des introns séparant les exons. Les introns sont éliminés par le splicéosome grâce à un mécanisme qui réunit les exons. L'épissage alternatif peut donner des ARNm différents, et donc des protéines différentes, à partir d'un même gène. Il paraît répandu chez les organismes pluricellulaires.

- Quels avantages l'épissage alternatif pourrait-il conférer à un organisme ?

Le ribosome est l'organite clé de la traduction, qui demande cependant aussi la participation de l'ARNm, de l'ARNt et d'une foule d'autres facteurs. L'interaction des ribosomes avec l'ARNt et l'ARNm est essentielle pour ce processus. Pour le comprendre, nous examinons d'abord la structure de la molécule adaptatrice, l'ARNt, et celle du ribosome lui-même.

Les aminoacyl-ARNt synthétases fixent les acides aminés à l'ARNt

Pour la progression de la synthèse protéique, chaque acide aminé doit être fixé au bon anticodon. Cette liaison covalente est effectuée par des enzymes d'activation, les **aminoacyl-ARNt synthétases**. Il existe une de ces enzymes pour chacun des 20 acides aminés communs.

Structure de l'ARNt

L'ARNt a deux fonctions et doit être capable d'interagir avec l'ARNm et avec les acides aminés. Sa structure est très conservée dans tous les organismes vivants et il peut prendre une forme en feuille de trèfle grâce à des appariements de bases au sein de la molécule et à des régions à double brin. Cette structure primaire est ensuite repliée dans l'espace en une molécule en forme de L avec deux extrémités fonctionnelles : le **site accepteur** et la **boucle de l'anticodon** (figure 15.14).

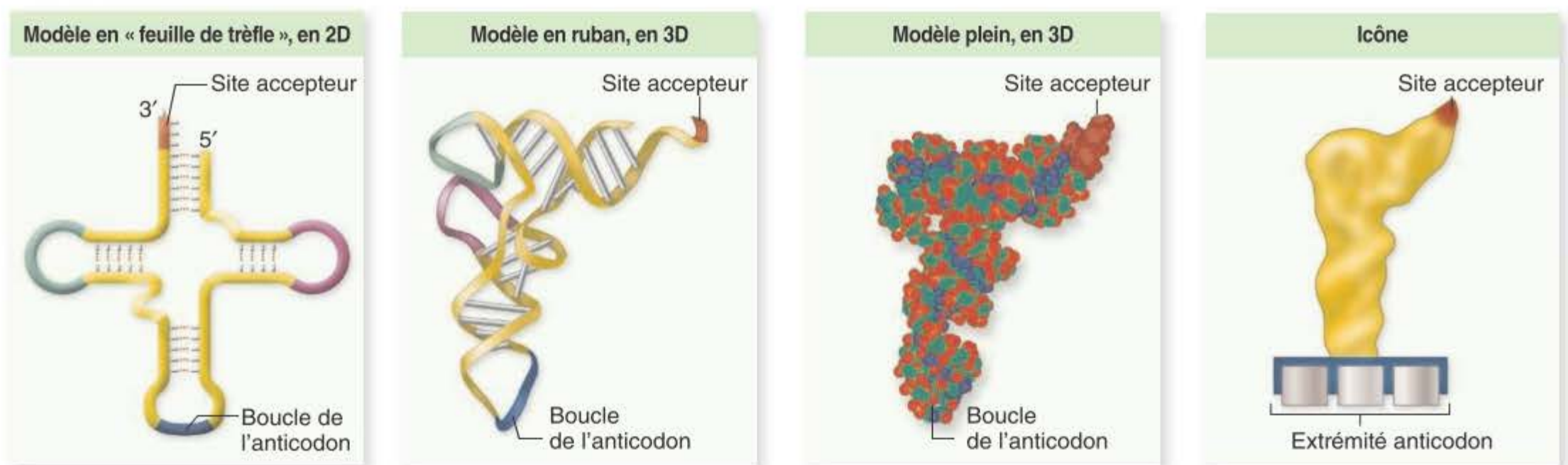


Figure 15.14 Structure de l'ARNt. L'appariement de bases au sein de la molécule donne trois structures en forme de tiges et de boucles et une forme caractéristique en feuille de trèfle. La boucle du bas contient la séquence anticodon, qui peut former des paires de bases avec des codons de l'ARNm. Les acides aminés se fixent à l'extrémité -OH monocaténaire libre de la tige réceptrice. Dans sa structure tridimensionnelle finale, les boucles d'ARNt se replient en forme de L.

Le site accepteur est l'extrémité 3' de la molécule, qui se termine toujours par 5'CCA3'. L'acide aminé se fixe à cette extrémité de la molécule. La boucle de l'anticodon est la boucle inférieure de la feuille de trèfle, elle peut former des paires de bases avec des codons de l'ARNm.

Le chargement

Les aminoacyl-ARNt synthétases doivent pouvoir reconnaître des molécules d'ARNt spécifiques aussi bien que les acides aminés qui leur correspondent. Bien que 61 codons codent des acides aminés, il n'y a en réalité pas 61 ARNt dans les cellules : leur nombre varie selon les espèces. Certaines aminoacyl-ARNt synthétases doivent donc pouvoir reconnaître plusieurs ARNt – mais chacune ne peut reconnaître qu'un seul acide aminé.

La réaction catalysée par les enzymes est la **réaction de chargement** de l'ARNt, et le produit est un acide aminé lié à un ARNt, qui devient un *ARNt chargé*. Une molécule d'ATP fournit l'énergie à cette réaction endergonique. L'ARNt chargé provenant de la réaction est un intermédiaire activé capable de former une liaison peptidique sans nouvel apport d'énergie.

La réaction de chargement unit le site accepteur à l'extrémité carboxyle d'un acide aminé (figure 15.15). Il faut absolument garder à l'esprit cette orientation, parce que toutes les liaisons peptidiques se formeront entre le groupement amine d'un acide amine et le groupement carboxyle d'un autre.

La fixation correcte des acides aminés aux ARNt est importante parce que les ribosomes ne la vérifient pas. Les ribosomes peuvent seulement assurer la régularité de l'appariement codon-anticodon. Dans une expérience élégante, on a transformé la cystéine en alanine après la réaction de chargement, quand l'acide aminé était déjà fixé à l'ARNt. Quand on a utilisé cet ARNt chargé dans un système de synthèse protéique in vitro, l'alanine a été incorporée à la place de la cystéine ; cela

montre que le ribosome ne peut pas vérifier les acides aminés attachés à l'ARNt.

En réalité, la réaction de chargement est donc la véritable étape de traduction ; les acides aminés sont incorporés à un polypeptide en fonction seulement de l'anticodon de l'ARNt et de son interaction avec l'ARNm.

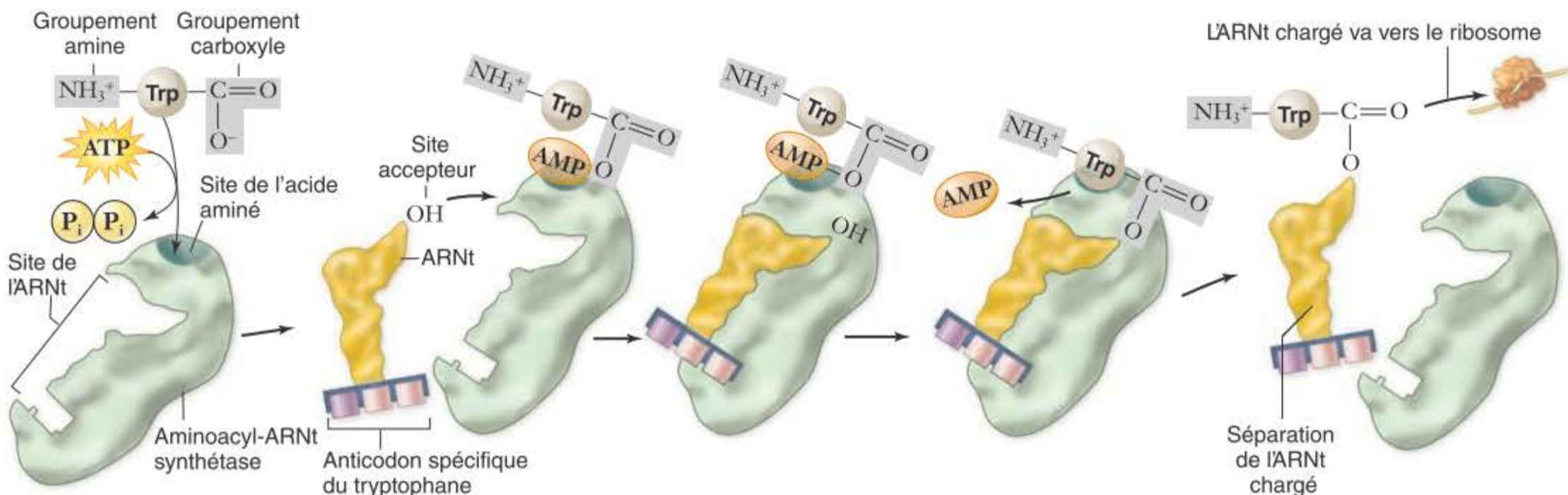
Le ribosome possède plusieurs sites de liaison à l'ARNt

On peut diviser la biosynthèse de tout biopolymère en initiation, élongation et terminaison – on a vu cette distinction pour la réplication de l'ADN comme pour la transcription. Dans le cas de la traduction, ou synthèse protéique, ces trois étapes se situent toutes sur le ribosome, assemblage macromoléculaire volumineux formé d'ARNr et protéines. Nous donnerons bientôt quelques détails concernant l'assemblage des deux sous-unités du ribosome au cours de l'initiation.

Pour fonctionner, le ribosome doit pouvoir s'unir à la fois à au moins deux ARNt chargés pour qu'une liaison peptidique puisse se former entre leurs acides aminés, comme on l'a vu à la section 15.1. Le ribosome bactérien possède trois sites de liaison, illustrés à la figure 15.16.

- Le **site P** (peptidyle) s'unit à l'ARNt fixé à la chaîne peptidique en croissance.
- Le **site A** (aminoacyle) s'unit à l'ARNt portant l'acide aminé suivant à ajouter.
- Le **site E** (exit) s'unit à l'ARNt qui a apporté l'acide aminé précédent.

Les ARN de transfert passent successivement par ces sites pendant l'élongation. Par rapport à l'ARNm, les sites sont disposés de 5' à 3' dans l'ordre E, P et A. Les ARNt chargés entrants pénètrent dans le ribosome au site A, passent par P et le quittent par le site E.



1. À la première étape de la réaction, l'acide aminé est activé. Il réagit avec l'ATP et produit un intermédiaire dont l'extrémité carboxyle de l'acide aminé est attachée à l'AMP. À cette occasion, les deux phosphates terminaux (pyrophosphates) sont enlevés de l'ATP.

2. Le complexe acide aminé-AMP reste uni à l'enzyme. L'ARNt suivant s'unit à l'enzyme.

3. La seconde étape de la réaction fait passer l'acide aminé de l'AMP à l'ARNt, donnant un ARNt chargé et l'AMP. L'ARNt chargé consiste en un acide aminé spécifique attaché à l'accepteur 3' de son ARN.

Figure 15.15 Chargement de l'ARNt. Il existe 20 aminoacyl-ARNt synthétases spécifiques des différents acides aminés, comme le tryptophane (Trp). L'enzyme doit aussi reconnaître les molécules d'ARNt et se fixer à l'ARNt possédant un anticodon pour l'acide aminé, ACC pour le tryptophane. La réaction utilise l'ATP et produit un intermédiaire activé qui n'a plus besoin d'énergie pour la formation de la liaison peptidique.

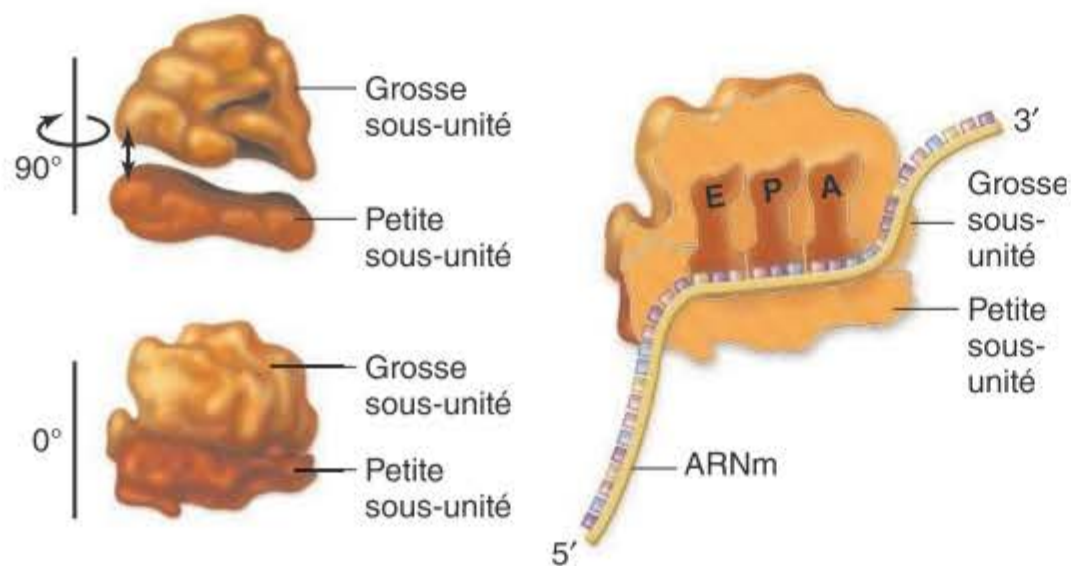


Figure 15.16 Les ribosomes ont deux sous-unités. Les sous-unités du ribosome s'assemblent et se séparent au cours d'un cycle du ribosome. La petite sous-unité s'insère dans une dépression superficielle de la grosse. Les ribosomes ont trois sites de liaison à l'ARNt : les sites aminoacyl (A), peptidyle (P) et sortie (E).

Le ribosome a des fonctions de décodage et d'enzyme

Les deux fonctions du ribosome impliquent le décodage du message transcrit et la formation des liaisons peptidiques. La fonction décodage réside principalement dans la petite sous-unité du ribosome. La formation de la liaison peptidique nécessite une enzyme, la peptidyl transférase, qui se trouve dans la grosse sous-unité.

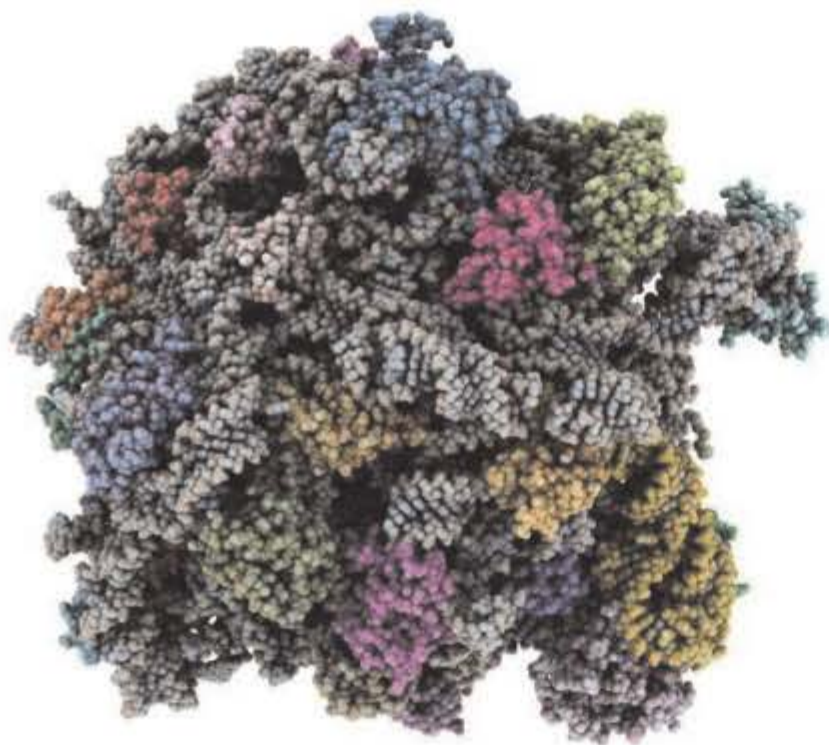


Figure 15.17 Structure en 3-D du ribosome procaryote. On voit ici la structure atomique complète de la grosse sous-unité ribosomique. L'ARN ribosomique est beige, bleu et vert pâle ; toutes les autres couleurs correspondent aux différentes protéines du ribosome. Les faces des deux sous-unités sont recouvertes par l'ARNr et leurs interactions avec les ARNt, les acides aminés et l'ARNm impliquent donc toutes l'ARNt. Les protéines sont absentes du site actif, mais abondantes partout ailleurs en surface. Elles stabilisent la structure par leurs interactions avec les brins d'ARN voisins.

Notre conception du ribosome a beaucoup changé au cours du temps. À l'origine, les biologistes moléculaires pensaient que les protéines du ribosome remplissaient cette fonction et que l'ARNr était un support structural assurant le positionnement correct des protéines. C'est plutôt l'inverse aujourd'hui ; le ribosome est considéré comme des ARNr restant à leur place grâce aux protéines. Les faces des deux sous-unités qui interagissent sont tapissées par l'ARNr, et les parties des deux sous-unités qui interagissent avec l'ARNm, l'ARNt et les acides aminés sont aussi formées surtout d'ARNr (figure 15.17). On pense aujourd'hui que l'activité de peptidyl transférase réside dans un ARNr de la grosse sous-unité.

Questions d'apprentissage 15.6

L'ARN de transfert possède deux régions fonctionnelles : l'une s'unit à un acide aminé et l'autre peut former des paires de bases avec l'ARNm. La réaction de chargement de l'ARNt unit l'extrémité carboxyle d'un acide aminé au site accepteur 3' de cet ARNt ; la traduction est impossible sans ARNt chargés. Cette réaction est catalysée par 20 aminoacyl-ARNt synthétases, une pour chaque acide aminé. Le ribosome possède trois sites différents de liaison à l'ARNt, un pour l'ARNt ajoutant un acide aminé à la chaîne en croissance (site P), un pour l'ARNt chargé suivant (site A) et un pour l'ARNt précédent, désormais sans acide aminé (site E). On peut considérer que le ribosome a deux fonctions : décodage et fonction enzymatique.

- Quelle conséquence aurait pour la traduction un ARNt mutant avec un anticodon complémentaire d'un codon stop ?

15.7 La traduction

Objectifs

1. Décrire le mécanisme d'initiation de la traduction.
2. Expliquer le cycle d'élongation.
3. Comparer la traduction sur le RER et dans le cytoplasme.

Pour les cellules, la traduction est un des processus les plus complexes et les plus exigeants en énergie. Le survol que nous en avons fait est peut-être trompeusement simple : l'ARNm passe par le ribosome, tandis que les ARNt portant les acides aminés s'unissent au ribosome, où ils interagissent avec l'ARNm à la suite d'un appariement des bases aux codons de l'ARNm. Le ribosome et les ARNt déterminent la position des acides aminés de telle sorte que les liaisons peptidiques peuvent se former entre chaque nouvel acide aminé et le polypeptide en croissance.

Des facteurs accessoires sont nécessaires à l'initiation

Comme on l'a déjà vu, le codon de départ est AUG, qui code aussi la méthionine. Le ribosome utilise généralement le premier AUG qu'il rencontre dans un brin d'ARNm pour signaler le démarrage de la traduction.

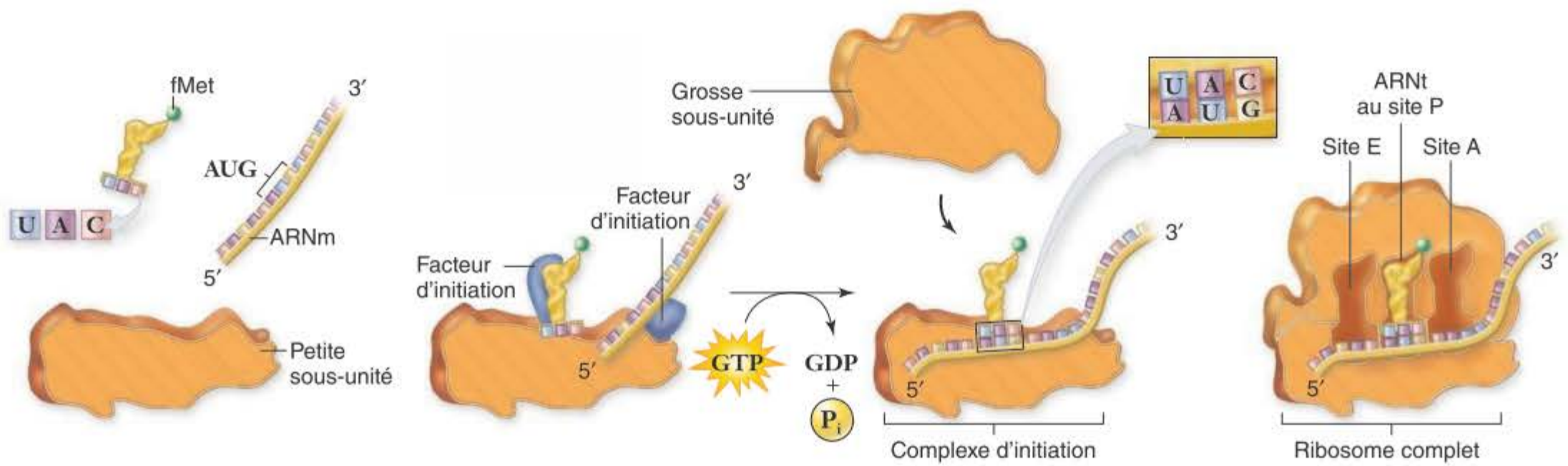


Figure 15.18 Initiation de la traduction. Chez les procaryotes, les facteurs d'initiation jouent un rôle clé dans le positionnement de la petite sous-unité ribosomique, de l'ARNt^{fMet} initiateur et de l'ARNm. Quand l'ARNt^{fMet} est placé sur le premier codon AUG de l'ARNm, la grosse sous-unité ribosomique s'y unit et forme les sites E, P et A où les molécules successives d'ARNt s'unissent au ribosome, et la synthèse du polypeptide débute. Les sous-unités ribosomiques sont représentées en coupe médiane.

L'initiation procaryote

Chez les procaryotes, le **complexe d'initiation** comprend une molécule spéciale d'ARNt **initiateur** chargée d'une méthionine chimiquement modifiée, la *N-formylméthionine*. L'ARNt initiateur est représenté par ARNt^{fMet}. Le complexe d'initiation comprend aussi la petite sous-unité ribosomique et le brin d'ARNm (figure 15.18). La petite sous-unité se place correctement sur l'ARNm grâce à une séquence conservée à l'extrémité 5' de l'ARNm, la **séquence d'union au ribosome** (*ribosome-binding sequence*, ou **RBS**), complémentaire de l'extrémité 3' de l'ARNr de la petite sous-unité.

Plusieurs facteurs d'initiation interviennent dans cette interaction entre ribosome, ARNm et ARNt^{fMet} pour donner le complexe d'initiation. Ces facteurs ne sont impliqués que dans l'initiation et ne font pas partie du ribosome.

Après la formation du complexe ARNm, ARNt initiateur et petite sous-unité ribosomique, la grosse sous-unité s'y ajoute, et la traduction peut débiter. Quand le ribosome est complet, l'ARNt initiateur est uni au site P et le site A est libre.

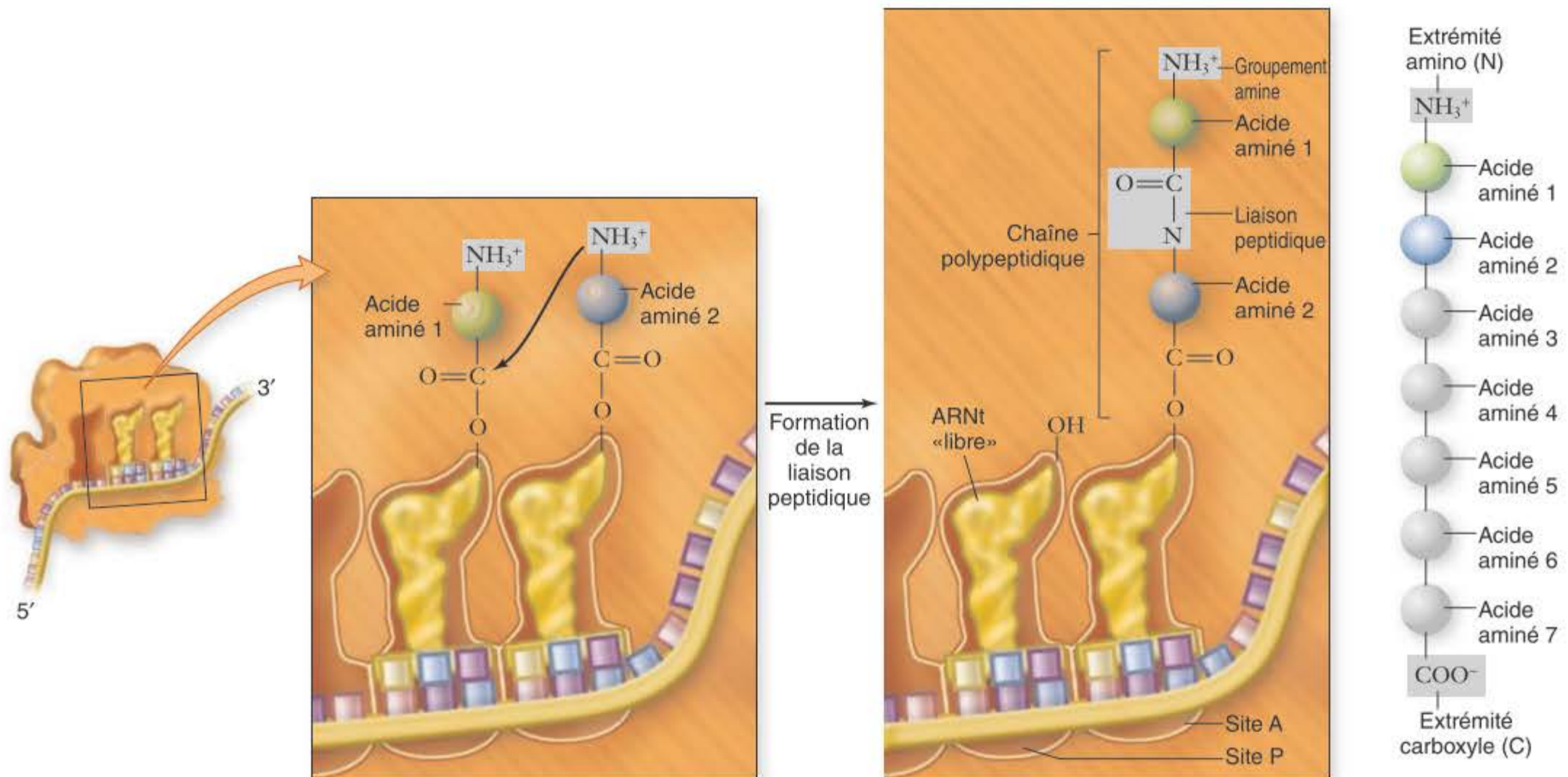


Figure 15.19 Formation de la liaison peptidique. Les liaisons peptidiques se forment entre un « nouvel » ARNt chargé au site A et la chaîne en croissance attachée à l'ARNt du site P. La liaison se forme entre le groupement amine du nouvel acide aminé et le groupement carboxyle de la chaîne. La liaison est ainsi rompue entre la chaîne en croissance et son ARNt, qui passe au site A, tandis que le nouvel acide aminé reste fixé à son ARNt.

L'initiation eucaryote

L'initiation est semblable chez les eucaryotes, mais elle en diffère par deux points importants. En premier lieu, l'acide aminé d'initiation des eucaryotes n'est pas la *N*-formylméthionine, mais la méthionine. En second lieu, le complexe d'initiation est beaucoup plus compliqué que chez les procaryotes et contient au moins neuf facteurs protéiques, dont beaucoup comptent plusieurs sous-unités. Les ARNm n'ont pas de RBS. La liaison de la petite sous-unité à l'ARNm débute au niveau de sa coiffe 5'.

L'élongation ajoute les acides aminés successifs

Quand le ribosome est complètement assemblé autour de l'ARNt initiateur et de l'ARNm, le deuxième ARNt chargé peut arriver au ribosome et se fixe au site A libre. Pour cela, il faut un **facteur d'élongation**, **EF-Tu**, qui s'unit à l'ARNt chargé et au GTP.

Une liaison peptidique peut alors se former entre l'acide aminé de l'ARNt initiateur et celui de l'ARNt chargé qui vient d'arriver au site A. La géométrie de cette liaison par rapport aux deux ARNt chargés est importante pour comprendre le processus. Rappelez-vous que l'acide aminé est

attaché à l'ARNt par son extrémité carboxyle. La liaison peptidique se forme entre l'extrémité amine de l'acide aminé entrant (au site A) et l'extrémité carboxyle de la chaîne en croissance (au site P) (figure 15.19).

L'addition des acides aminés successifs constitue une série d'événements se déroulant de façon cyclique. La figure 15.20 montre les détails du cycle d'élongation.

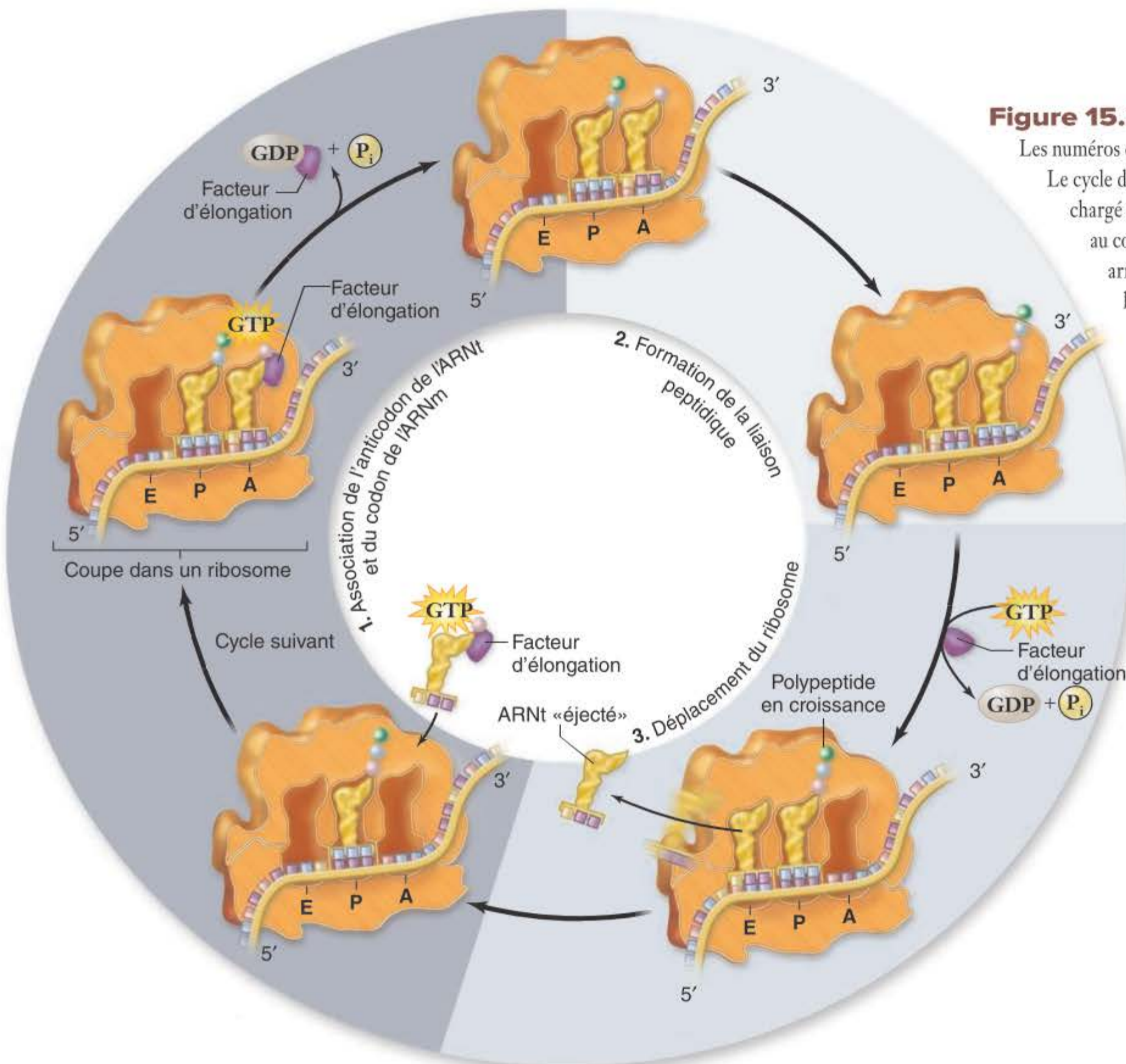
- 1. Appariement de l'anticodon de l'ARNt et du codon de l'ARNm.** Chaque nouvel ARNt chargé arrive au ribosome uni à EF-Tu et GTP. Il se fixe au site A si son anticodon est complémentaire du codon de l'ARNm à ce site.

Après la fixation, le GTP est hydrolysé et EF-Tu-GDP se sépare du ribosome, où il est remplacé par un autre facteur. On pense que cette fixation en deux temps et l'hydrolyse du GTP améliorent la précision de la traduction.

- 2. Formation de la liaison peptidique.** La peptidyl transférase, située dans la grosse sous-unité, catalyse la formation d'une liaison peptidique entre le groupement amine de l'acide aminé au site A et le groupement carboxyle de la chaîne en croissance. En même temps, la liaison entre cette chaîne et l'ARNt est rompue au site P, qui est libre (il n'est plus chargé). Le résultat final est le passage de la chaîne en croissance à l'ARNt au site A.

Figure 15.20 Cycle d'élongation.

Les numéros correspondent à ceux du texte. Le cycle débute quand un nouvel ARNt chargé avec l'anticodon correspondant au codon de l'ARNm au site A arrive avec EF-Tu. EF-Tu hydrolyse le GTP et se sépare du ribosome. Une liaison peptidique se forme entre l'acide aminé du site A et la chaîne en croissance au site P, faisant passer la chaîne au site A et laissant libre l'ARNt au site P. Le passage par le ribosome exige un autre facteur d'élongation et l'hydrolyse du GTP. L'ARNt passe ainsi du site A au site P, le codon d'ADN suivant au site A et l'ARNt libre au site E.



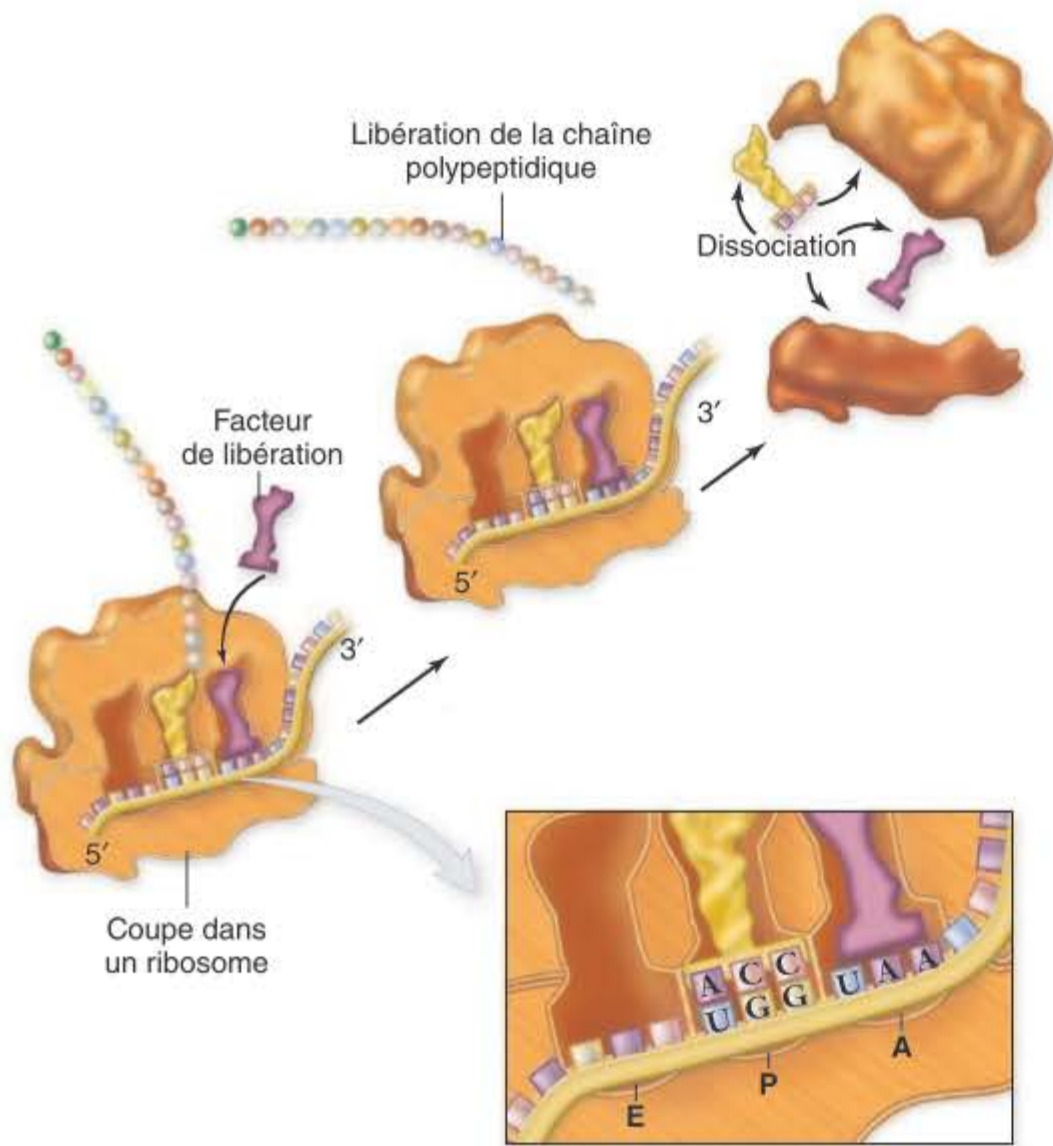


Figure 15.21 Terminaison de la synthèse protéique. Il n'existe pas d'anticodon complémentaire des trois codons de terminaison. Quand un ribosome en rencontre un, il met donc fin à la translocation. Un facteur de libération spécifique facilite la libération de la chaîne polypeptidique en brisant la liaison covalente unissant le polypeptide à l'ARNt au site P.

- 3. Translocation du ribosome.** Après la formation de la liaison peptidique, le ribosome se déplace par rapport à l'ARNm et aux ARNt. Le codon suivant de l'ARNm se glisse dans le site A et l'ARNt passe au site P avec la chaîne en croissance. L'ARNt non chargé qui se trouvait au site P est maintenant au site E, et il sera éjecté au cycle suivant. Cette étape de translocation nécessite le facteur accessoire EF-G et l'hydrolyse d'un autre GTP.
- 4.** Ce cycle d'élongation se poursuit par l'addition de nouveaux acides aminés. Le ribosome descend le long de l'ARNm dans le sens 5'-3' en lisant les codons successifs. Les ARNt traversent le ribosome dans le sens opposé, du site A au site P et finalement au site E, avant leur éjection sous la forme d'ARNt libres, susceptibles d'être chargés d'un autre acide aminé et de servir à nouveau.

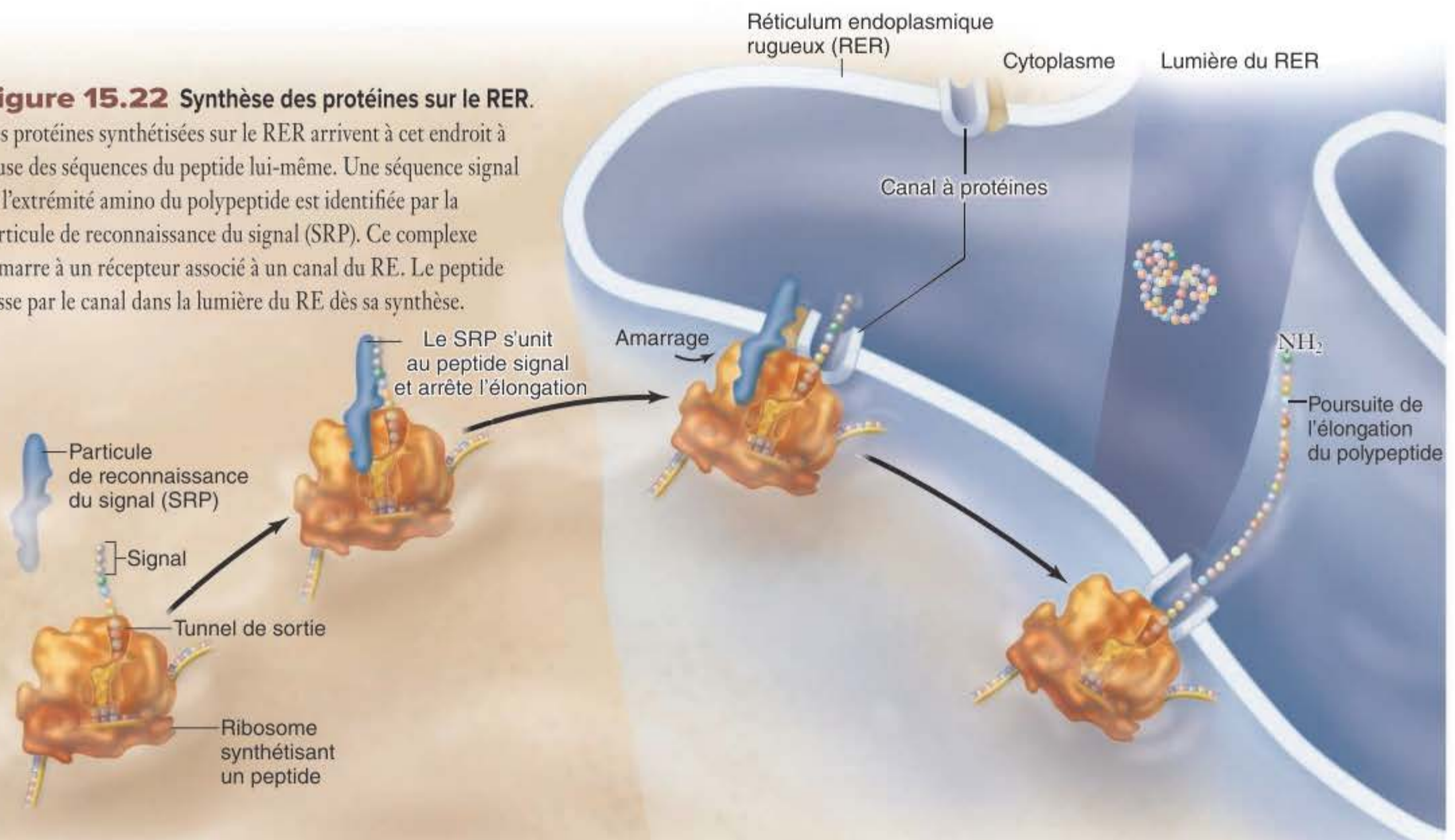
Le ballottement

On a vu qu'il y a moins d'ARNt que de codons. Cette situation est logique parce que l'appariement entre la base 3' du codon et la base 5' de l'anticodon est moins strict que les autres. Dans certains ARNt, la présence de bases modifiées avec un appariement moins précis au niveau 5' de l'anticodon augmente cette flexibilité : on parle de **ballottement**, parce que ces ARNt peuvent quelque peu « balloter » sur l'ARNm, et un même ARNt peut « lire » plusieurs codons de l'ARNm.

? **Question** Quelle est la relation entre le ballottement, le nombre d'ARNt et la dégénérescence du code génétique ?

Figure 15.22 Synthèse des protéines sur le RER.

Les protéines synthétisées sur le RER arrivent à cet endroit à cause des séquences du peptide lui-même. Une séquence signal de l'extrémité amino du polypeptide est identifiée par la particule de reconnaissance du signal (SRP). Ce complexe s'amarrage à un récepteur associé à un canal du RE. Le peptide passe par le canal dans la lumière du RE dès sa synthèse.



La terminaison a besoin de facteurs accessoires

L'élongation se poursuit jusqu'au codon stop au bout de la chaîne (par exemple UAA à la figure 15.21). Ces codons stop ne s'unissent pas à l'ARNt : ils sont identifiés par des facteurs de libération, protéines qui libèrent le nouveau polypeptide du ribosome.

Les protéines peuvent être dirigées vers le RE

Chez les eucaryotes, la traduction est possible dans le cytoplasme ou sur le RER. Les protéines traduites sur le RER sont ciblées en fonction de leur propre séquence initiale d'acides aminés. Les ribosomes du RER ont une traduction active et ne sont pas liés de façon permanente au RE.

Un polypeptide débutant par une courte série d'acides aminés dénommée **séquence d'adressage** est spécifiquement identifié et fixé par un complexe cytoplasmique de protéines, la **particule de reconnaissance du signal** (*signal recognition particle*, ou **SRP**). Le complexe séquence d'adressage-SRP est à son tour identifié par une protéine réceptrice de la membrane du RE. L'union du récepteur du RE au complexe séquence d'adressage-SRP fixe le ribosome engagé dans la traduction de la protéine au RE : c'est l'*arrimage* (figure 15.22).

À mesure que la protéine s'assemble, elle passe par un canal formé par le complexe d'arrimage à l'intérieur du RE, dans la citerne. D'où le terme *arrimage* – le ribosome n'est pas vraiment fixé au RE lui-même, mais, grâce à la nouvelle protéine pénétrant dans le RE, le ribosome fait penser à un bateau attaché au quai par un cordage.

Le mécanisme à la base du passage de la protéine au travers des membranes par le SRP et son récepteur et le complexe de canal a persisté dans les trois types de cellules : eucaryotes, bactéries et archées. Seules les cellules eucaryotes disposant d'un système endomembranaire, cette universalité peut sembler curieuse ; cependant, les bactéries et les archées exportent des protéines à travers leur membrane plasmique, et le mécanisme est semblable à celui qui est utilisé par les eucaryotes pour faire passer les protéines dans les citernes du RE.

Dans la lumière de la citerne du RE, la nouvelle protéine peut être modifiée par l'addition de sucres (glycosylation) et transportée par des vésicules vers l'appareil de Golgi (voir chapitre 4). C'est le point de départ de la voie de transport des protéines, qui peut aboutir à d'autres cibles intracellulaires, à l'incorporation à la membrane plasmique ou à la libération au dehors de la cellule elle-même.

Questions d'apprentissage 15.7

L'initiation de la traduction implique une interaction entre la petite unité ribosomique et un ARNt initiateur chargé. Le cycle d'élongation implique l'entrée de nouveaux ARNt chargés au site A du ribosome, la formation de liaisons peptidiques entre les acides aminés et le déplacement du ribosome le long de l'ARNm. Les ARNt traversent le ribosome du site A aux sites P et E pendant ce processus. Chez les eucaryotes, les séquences d'adressage du nouveau polypeptide peuvent le diriger vers le RE avec son ribosome. Les polypeptides synthétisés sur le RER ne sont pas libérés dans le cytoplasme, mais pénètrent à l'intérieur de la citerne.

- Quelles étapes de la traduction exigent de l'énergie ?

15.8 L'expression génique en résumé

En raison de la complexité du processus de l'expression génique, il est utile de revenir en arrière pour résumer quelques points essentiels :

- L'expression génique transforme l'information du génotype en phénotype.
- La transcription produit une copie du gène sous forme d'ARNm et l'ARNm sert à diriger la synthèse d'une protéine par traduction.
- On peut diviser la transcription et la traduction en initiation, cycle d'élongation et terminaison, aboutissant à la synthèse de leurs polymères respectifs. (C'est également vrai pour la réplication de l'ADN.)
- L'expression des gènes eucaryotes est beaucoup plus complexe que celle des procaryotes.

La structure des gènes eucaryotes, avec leurs séquences codantes interrompues, complique le processus de l'expression génique et la nature de l'information génétique. Cela implique la nécessité d'une maturation entre la transcription et la traduction, et donc, un même gène peut être à l'origine de plusieurs messages. Chez les eucaryotes, la transcription s'effectue en outre dans le noyau, tandis que la traduction se déroule dans le cytoplasme. ce qui nécessite le transfert de l'ARNm vers le cytoplasme par les pores avant la traduction. Tout le processus eucaryote est résumé à la figure 15.23 et les différences entre les procaryotes et les eucaryotes sont résumées au tableau 15.2.

Questions d'apprentissage 15.8

La plus grande complexité de l'expression génique des eucaryotes est liée à l'organisation fonctionnelle de la cellule, l'ADN se trouvant dans le noyau et les ribosomes dans le cytoplasme. Les différences entre procaryotes et eucaryotes concernent principalement des détails, mais certaines différences ont une signification fonctionnelle.

15.9 Les mutations : des gènes modifiés

Objectifs

1. Décrire les conséquences de différentes mutations ponctuelles.
2. Expliquer la nature de l'expansion des triplets répétés.
3. Citer les différentes mutations chromosomiques et leurs effets.

Une mutation est une modification du matériel génétique qui peut être héritée. Une modification aussi simple que l'altération d'une seule base peut entraîner une substitution d'acide aminé susceptible d'entraîner un phénotype clinique déficient. On le voit pour l'anémie à cellules falciformes à la figure 15.24. Dans l'allèle responsable des cellules falciformes, une seule base A est remplacée par une base T, et un acide

Figure 15.23 Vue d'ensemble de l'expression génique chez les eucaryotes.

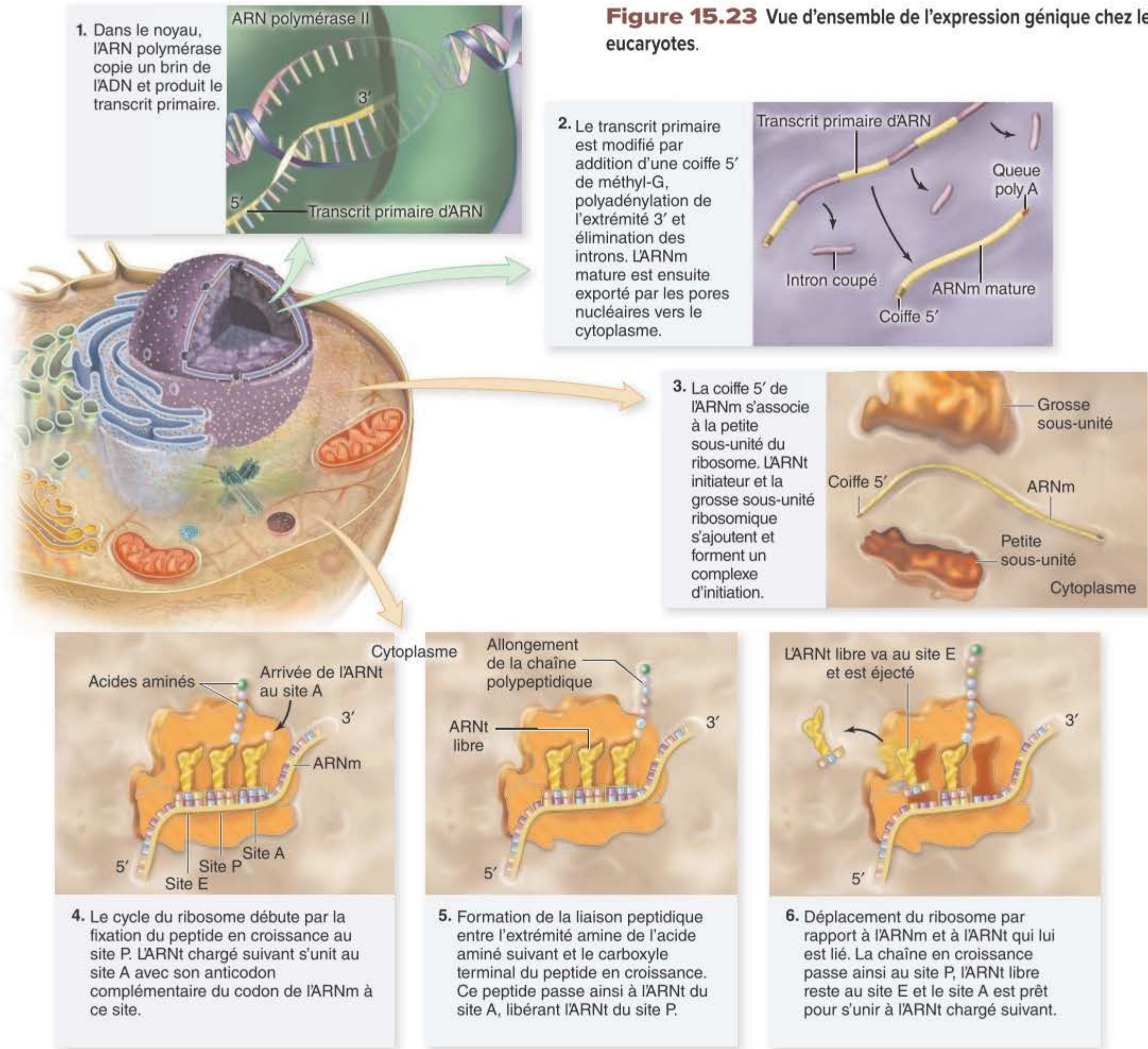


TABLEAU 15.2 Différences entre l'expression génique des procaryotes et des eucaryotes

Caractéristique	Procaryotes	Eucaryotes
Introns	Pas d'introns, sauf dans certains gènes d'archées	La plupart des gènes ont des introns
Nombre de gènes dans l'ARNm	Plusieurs gènes peuvent être transcrits en une seule molécule d'ARNm. Ils ont souvent des fonctions apparentées et forment un opéron, ce qui favorise la régulation coordonnée des voies biochimiques.	Un seul gène par molécule d'ARNm ; autres modes de régulation.
Site de transcription et de traduction	Pas de noyau délimité par une membrane, transcription et traduction sont couplées.	Transcription dans le noyau ; l'ARNm est transporté vers le cytoplasme pour traduction.
Initiation de la traduction	Débute au codon AUG précédé d'une séquence spéciale qui fixe le ribosome.	Débute au codon AUG précédé de la coiffe 5' (GTP méthylé) qui fixe le ribosome.
Modification de l'ARNm après transcription	Aucune ; la traduction commence avant la fin de la transcription. Transcription et traduction sont couplées.	Plusieurs modifications tant que l'ARNm est dans le noyau : les introns sont éliminés et les exons sont épissés ; une coiffe 5' est ajoutée, ainsi qu'une queue poly-A.

glutamique est ainsi remplacé par une valine. Cette substitution de la valine non-polaire entraîne l'agglutination des chaînes β en polymères et modifie la forme des cellules, cause de la maladie.

Nous allons aborder à présent les différentes modifications susceptibles de se produire dans les séquences d'ADN, ainsi que leurs conséquences. Nous compléterons cette discussion en tenant compte des données récentes sur la fréquence de ces événements chez les humains.

Les mutations ponctuelles concernent un seul site de l'ADN

Une mutation concernant une seule base est une **mutation ponctuelle**, ou **modification d'un seul nucléotide (MSN)**. Ce peut être la substitution d'une base par une autre, la délétion ou l'addition d'une seule base (ou d'un petit nombre de bases) (figure 15.25).

Substitution de bases

La substitution d'une paire de bases par une autre dans l'ADN est une mutation par **substitution de bases**. Le code génétique étant dégénéré, une substitution de bases ne modifie pas nécessairement l'acide aminé codé. Si le nouveau codon code encore le même acide aminé, on dit que la mutation est *silencieuse* (figure 15.25b). Si la substitution de bases modifie un acide aminé de la protéine, c'est une **mutation faux-sens** parce que le « sens » du codon sera modifié après la transcription du gène mutant (figure 15.25c). On peut avoir des *transitions* ou des *transversions*. Une transition ne change pas le type de bases de la paire – une pyrimidine est remplacée par une pyrimidine ou une purine par une purine. Par contre, une transversion modifie le type de bases de la paire – comme pyrimidine pour purine ou inversement. Diverses maladies génétiques humaines, comme l'anémie à cellules falciformes, sont dues à des substitutions de bases.

Mutations non-sens

On trouve une catégorie particulière de substitution de bases quand le codon est transformé en codon stop (voir figure 15.25d). On parle alors

de **mutation non-sens** parce que la mutation ne donne plus un « sens » à l'appareil de traduction. Le codon stop entraîne une terminaison prématurée de la traduction et aboutit à une protéine tronquée. La longueur de la protéine produite dépend de la position du codon stop introduit dans le gène.

Mutations avec décalage de phase

L'addition ou la délétion d'une seule base a des conséquences beaucoup plus graves que la substitution d'une base par une autre. Ces mutations sont des *mutations avec décalage de phase* : elles modifient le cadre de lecture de l'ARNm en aval de la mutation. Comme on l'a vu à la section 15.2, c'est le type de mutation utilisé par Crick et Brenner pour déduire la nature du code génétique.

La modification du cadre de lecture au début d'un gène, et donc de son transcrit d'ARNm, signifie que la plus grande partie de la protéine sera modifiée. Le décalage de phase peut aussi entraîner une terminaison prématurée de la traduction parce que 3 des 64 codons sont des codons stop, ce qui représente une forte probabilité dans la séquence remaniée aléatoirement par le décalage de phase.

Mutations par expansion de triplets répétés

En raison de la longue histoire de la génétique moléculaire et de la période relativement courte pendant laquelle les analyses moléculaires ont été possibles chez l'homme, la découverte d'un nouveau type de mutation chez les humains est étonnante. C'est cependant ce qui s'est produit pour un des premiers gènes isolés associé à une maladie humaine, le gène de la *maladie de Huntington*. Ce gène comporte un triplet d'ADN répété, et cette unité répétée est amplifiée en comparaison de l'allèle normal. Depuis cette première découverte, on a trouvé au moins 20 autres maladies humaines dues à ce mécanisme. La fréquence de ce type de mutation est inconnue mais, jusqu'à présent, on ne l'a observé que chez les humains et les souris ; il n'existe donc peut-être que chez les vertébrés, ou même seulement chez les mammifères. On n'a jamais trouvé une telle mutation chez les drosophiles par exemple.

On peut trouver l'expansion du triplet dans une région codante ou dans l'ADN transcrit non codant. Dans le cas de la maladie de Hun-

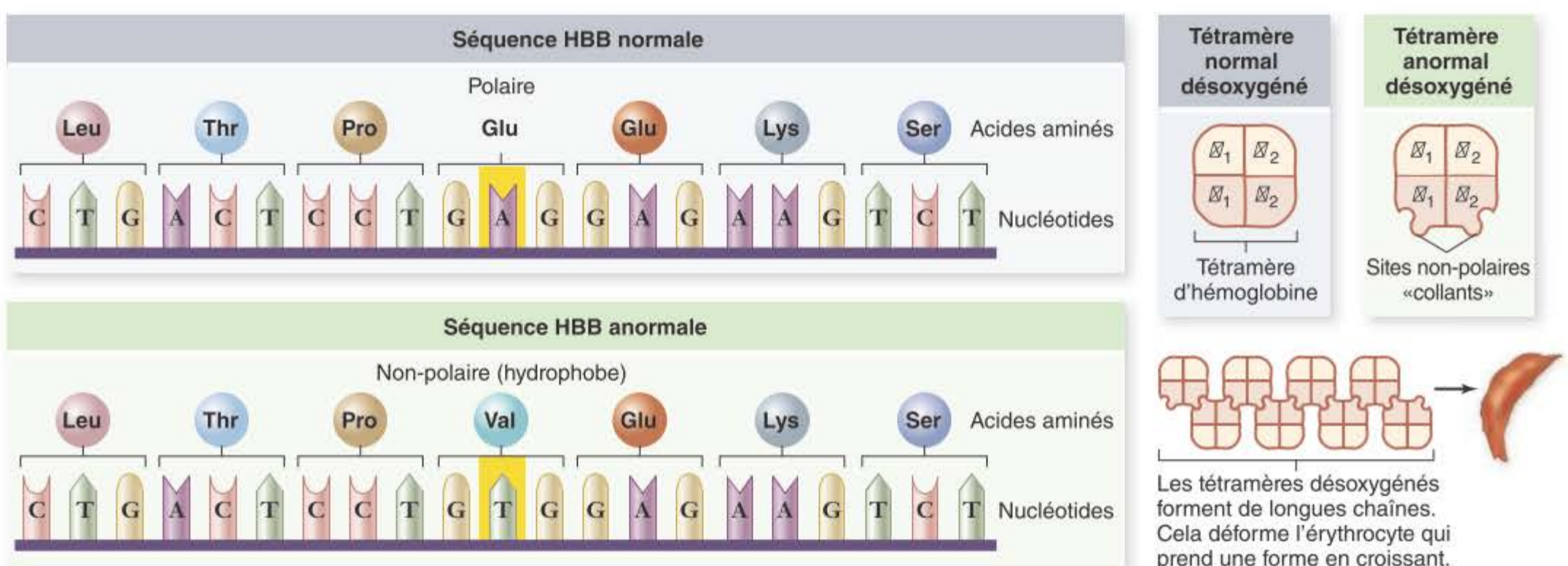


Figure 15.24 L'anémie à cellules falciformes est provoquée par l'altération d'une protéine. L'hémoglobine est un tétramère formé de deux globines α et de deux globines β . L'allèle des cellules falciformes du gène de la globine β possède une base modifiée suite au remplacement de Glu par Val. Une région hydrophobe est ainsi créée à la surface de la protéine, qui est « collante » et forme de longues chaînes qui perturbent la forme des érythrocytes.

tington, l'unité répétée se trouve en fait dans la région codante du gène, où le triplet code la glutamine, et l'expansion est à l'origine d'une région polyglutaminique dans la protéine. Plusieurs autres maladies neurodégénératives correspondent aussi à ce type de mutation. Dans le cas du syn-

drome du chromosome X fragile, forme héréditaire de retard mental, la répétition se trouve dans l'ADN non codant.

Les mutations chromosomiques modifient la structure des chromosomes

Les mutations ponctuelles affectent un seul site du chromosome, mais les modifications les plus étendues peuvent modifier la structure même du chromosome et donner des **mutations chromosomiques**. Beaucoup de cancers humains sont associés à des anomalies chromosomiques, dont l'importance clinique est donc grande. Voici, en bref, les altérations possibles de la structure chromosomique, toutes représentées à la figure 15.26.

Délétions

Une **délétion** est la perte d'un morceau de chromosome. Les décalages de phase peuvent provenir d'une ou plusieurs petites délétions, mais des régions beaucoup plus grandes d'un chromosome peuvent aussi être perdues. Si la quantité d'information perdue est trop grande, la délétion est généralement létale pour l'organisme.

Un syndrome humain dû à une délétion est le *cri-du-chat*, nom qui se réfère aux sons émis par les enfants atteints de ce syndrome. Le syndrome du cri-du-chat est dû à une grande délétion dans le bras court du chromosome 5. Il entraîne généralement une mort prématurée, bien que certains individus affectés aient une durée de vie normale. Il a des conséquences diverses, entre autres des problèmes respiratoires.

Duplications

La **duplication** d'une région d'un chromosome n'a pas nécessairement des conséquences phénotypiques. Ses effets dépendent de la localisation des « points de rupture » où s'est produite la duplication. Si la région dupliquée ne se trouve pas au sein d'un gène, il n'y a aucun effet. Si la duplication est proche de la région d'origine, on parle d'une *duplication en tandem*. Les duplications en tandem sont importantes pour l'évolution des familles de gènes apparentés, comme la famille des globines codant l'hémoglobine.

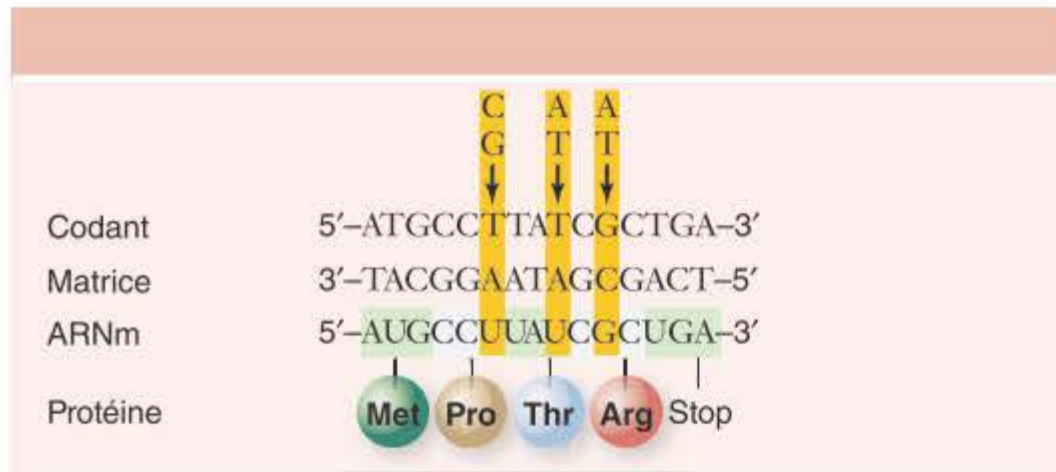
Inversions

On a une **inversion** quand un segment chromosomique est rompu à deux endroits, inversé et ressoudé. Une inversion peut être sans effet sur le phénotype si les cassures ne se sont pas produites à l'intérieur d'un gène. En fait, bien que tous les humains possèdent le « même » génome, l'ordre des gènes de tous les individus d'une population n'est pas exactement le même à cause des inversions survenant dans des lignées différentes.

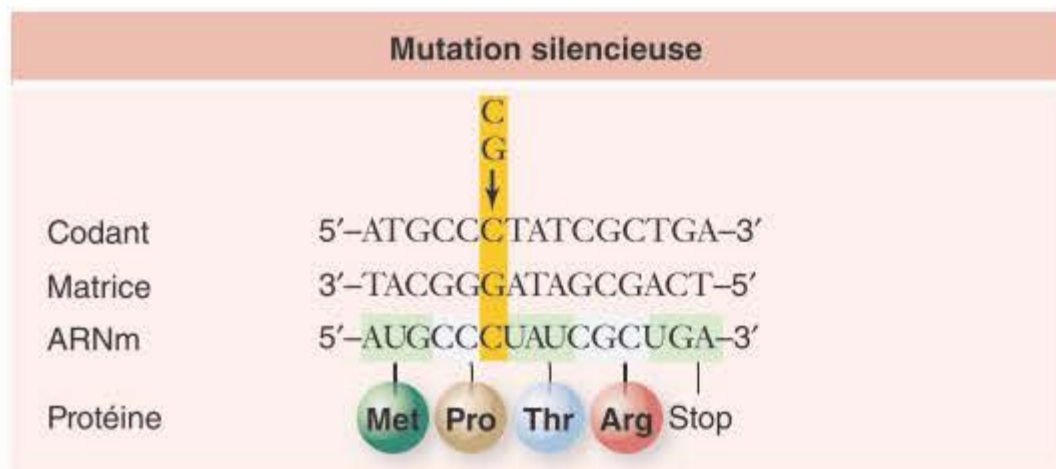
Translocations

Si un morceau d'un chromosome est excisé et rattaché à un autre chromosome, on parle d'une **translocation**. Les translocations sont complexes parce qu'elles peuvent entraîner des problèmes durant la méiose, en particulier si deux chromosomes différents tentent de s'apparier en première division méiotique.

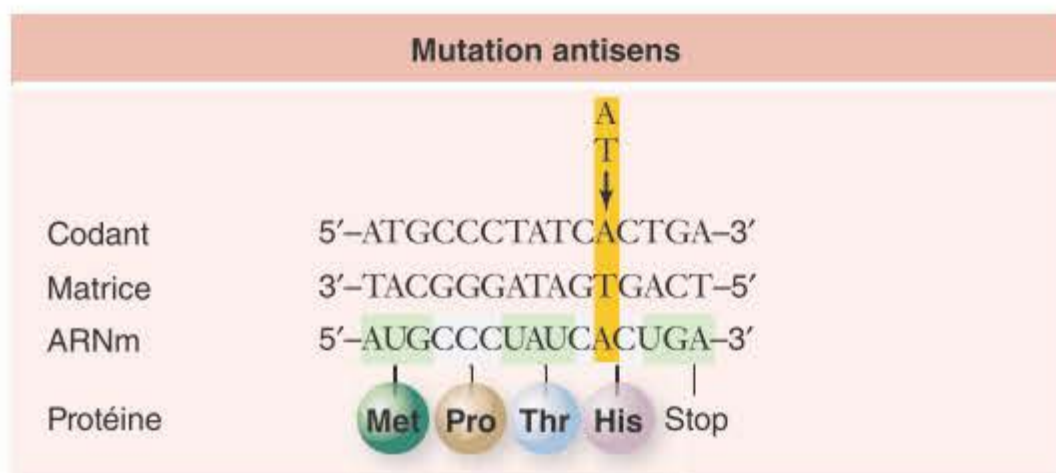
Les translocations peuvent aussi déplacer des gènes d'une région chromosomique à une autre et modifier ainsi l'expression des gènes de cette région. On a montré que deux formes de leucémie sont associées à des translocations qui déplacent des oncogènes entre régions d'un chromosome, où ils s'expriment de manière inadéquate dans les cellules sanguines.



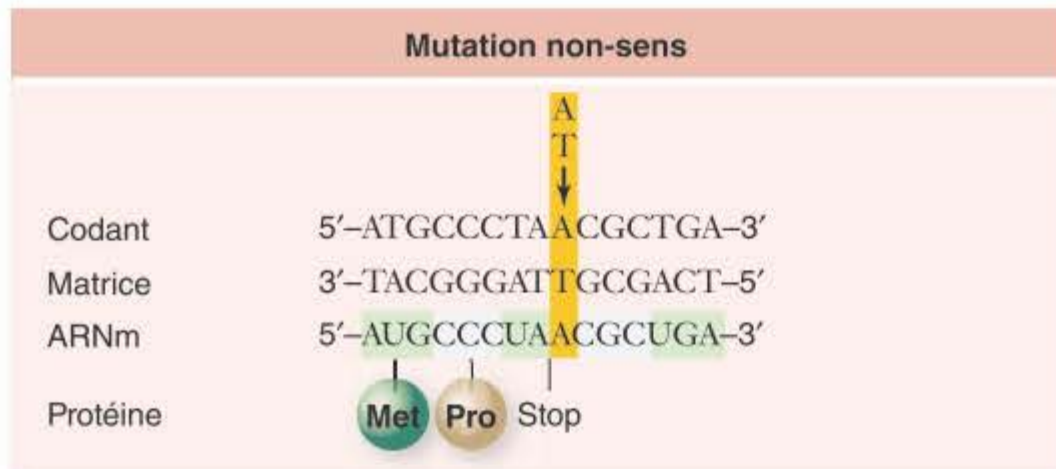
a.



b.



c.



d.

Figure 15.25 Types de mutations. a. Gène hypothétique, avec l'ARNm codé et la protéine. Les flèches au-dessus du gène indiquent les sites des mutations décrites dans le reste de la figure. b. Mutation silencieuse. Une modification en troisième position d'un codon est souvent silencieuse à cause de la dégénérescence du code génétique. Dans ce cas, la mutation de T/A à C/G ne change pas l'acide aminé codé (proline). c. Mutation faux sens. La mutation G/C en A/T remplace l'acide aminé arginine par l'histidine. d. Mutation non-sens. La mutation T/A en A/T produit un codon stop UAA dans l'ARNm.

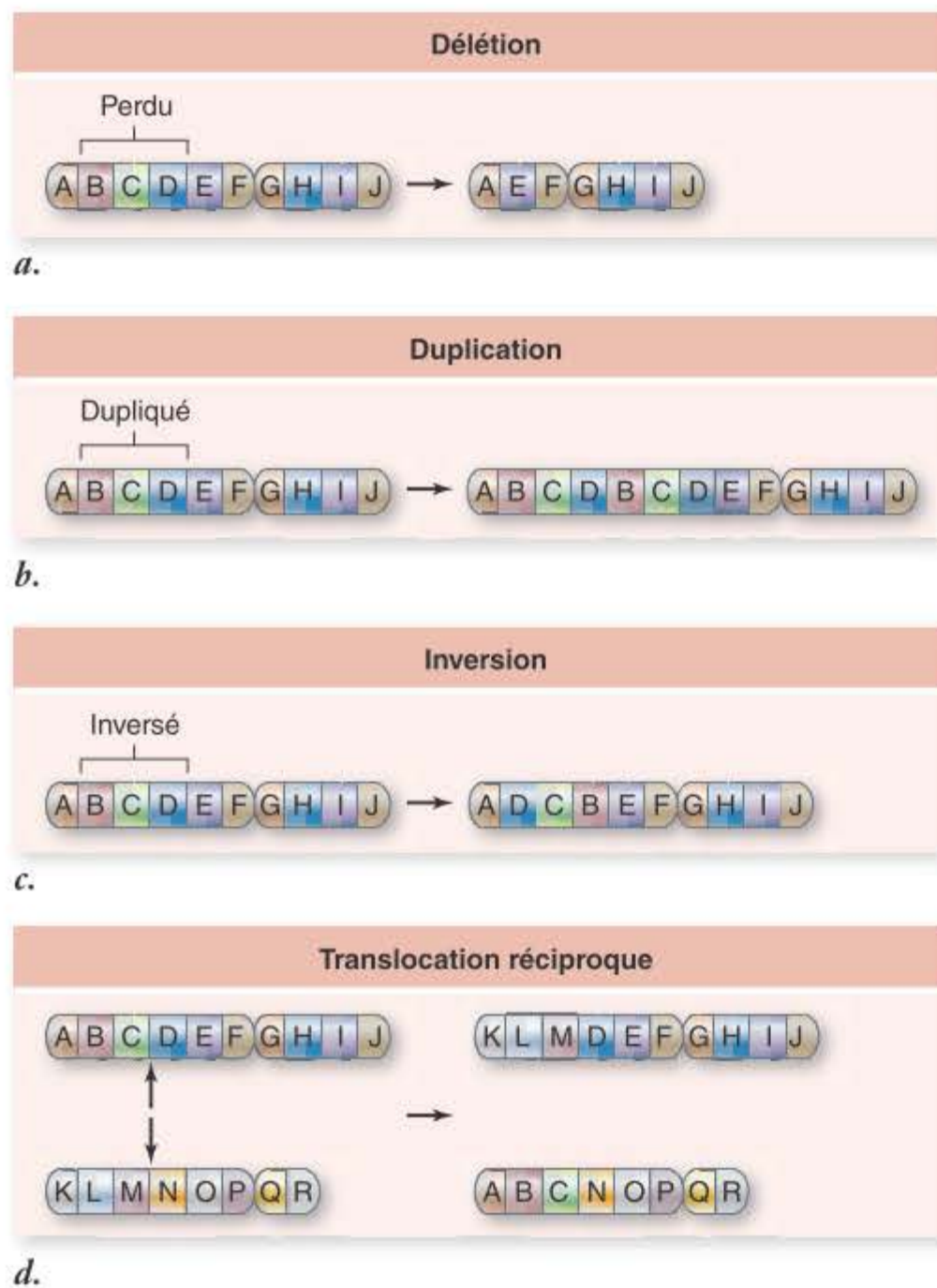


Figure 15.26 Mutations chromosomiques. Des modifications de grande ampleur des chromosomes sont aussi possibles. Du matériel peut être perdu (a), dupliqué (b) et inversé (c). On a une translocation quand un chromosome est brisé et que le morceau s'unit à un autre chromosome. C'est fréquent quand les deux chromosomes sont rompus et échangent du matériel : on parle alors d'une translocation réciproque (d).

Les mutations sont le point de départ de l'évolution

Si les génomes ne se modifiaient pas au cours du temps, l'évolution serait impossible. Trop de modification est cependant nuisible pour un individu dont le génome est profondément altéré. Il doit donc exister un équilibre délicat entre l'importance de la nouvelle variation survenant dans l'espèce et la santé de ses individus. Ce sujet sera exploré plus en détail au chapitre 20, lorsque seront abordées l'évolution et la génétique des populations.

L'altération à grande échelle des chromosomes a également été importante dans l'évolution, bien que son rôle soit mal connu. Il est clair que des familles de gènes proviennent de la duplication d'un gène ancestral, suivie d'une divergence fonctionnelle des copies dupliquées. Il est

clair aussi que, même parmi des espèces étroitement apparentées, il peut exister des différences dans le nombre et la disposition des gènes. Des réarrangements de grande ampleur peuvent aussi survenir.

On a mesuré directement les taux de mutation chez les humains

Longtemps, l'utilisation de méthodes indirectes ne nous a permis que d'estimer les taux de mutations humaines. Les premières tentatives remontent aux années 1930 et aux travaux de Haldane sur l'hémophilie. Maintenant qu'il est possible de séquencer des génomes entiers rapidement et à faible coût, nous pouvons mesurer directement les taux de mutation chez les humains en comparant les génomes des parents et des descendants.

Un travail récent portant sur 78 « trios » (mère, père, enfant) d'Islande a donné un taux de $1,2 \times 10^{-8}$ mutations MSN par paire de bases et par génération. Cela correspond à environ 70 nouvelles mutations par naissance. Ce travail montre également que 76 % de ces nouvelles mutations par MSN apparaissent dans le génome paternel et que ce taux augmente avec l'âge du père. D'autres méthodes, comme celle qu'utilisait Haldane, ont confirmé cette différence paternelle. Celle-ci traduit probablement l'accumulation des mutations en fonction de l'âge dans la population de cellules souches à l'origine des spermatozoïdes. De petites délétions et insertions de moins de 50 bases, appelées indels, surviennent avec une fréquence d'un cinquième à un dixième de celui des MSN. Les délétions et duplications beaucoup plus importantes, appelées **variations du nombre de copies (VNC)** sont beaucoup plus rares. On estime qu'environ une grande VNC survient sur 42 naissances. Certains éléments génétiques mobiles qui encombrant notre génome (voir chapitre 18) peuvent également se déplacer. Le taux d'insertion d'éléments mobiles semble être de l'ordre de 1 pour 20 naissances.

Ces nouvelles données clarifient de mieux en mieux l'évolution humaine, particulièrement dans les analyses comparatives avec les autres espèces de primates. C'est également important pour les maladies, car de nouvelles mutations paraissent maintenant être une source plus importante de maladies génétiques humaines qu'on ne le pensait.

Questions d'apprentissage 15.9

Parmi les mutations ponctuelles (changements de bases isolées, additions ou délétions), on trouve les mutations faux-sens entraînant la substitution d'un acide aminé par un autre, les mutations non-sens, qui bloquent la traduction, et les mutations avec décalage de phase qui empêchent la lecture correcte des codons. L'expansion des triplets répétés est le résultat d'une duplication anormale d'un codon à chaque cycle de division cellulaire. Les mutations affectant les chromosomes sont les délétions, les duplications, les inversions et les translocations.

- Une inversion ou une duplication doit-elle toujours avoir des conséquences sur le phénotype ?

15.1 Nature des gènes

Garrod arrive à la conclusion que les maladies héréditaires peuvent impliquer des enzymes spécifiques.

Garrod a constaté que l'alcaptonurie est due à une enzyme modifiée.

Beadle et Tatum ont montré que les gènes définissent les enzymes.

On a constaté que les mutants de *Neurospora* incapables de synthétiser l'arginine étaient dépourvus d'enzymes spécifiques. Beadle et Tatum ont proposé l'hypothèse « une enzyme/un polypeptide ».

Le dogme central décrit le flux d'information passant de l'ADN à l'ARN et aux protéines dans les cellules (figure 15.2).

Le brin d'ADN copié en ARNm est désigné comme le brin matrice (anti-sens) ; l'autre est le brin codant (sens).

La transcription copie l'ADN sous forme d'ARN.

La traduction utilise l'information de l'ARN pour synthétiser les protéines.

Une molécule adaptatrice, l'ARNt, est nécessaire pour faire la connexion entre l'information de l'ARNm et la séquence des acides aminés.

Les rôles de l'ARN dans l'expression génique sont multiples.

15.2 Le code génétique

Le code est lu par groupes de trois.

Crick et Brenner ont montré que le code n'est pas chevauchant et qu'il est lu par groupes de trois. Cette découverte est à la base du cadre de lecture.

Nirenberg et d'autres ont déchiffré le code.

Un codon consiste en 3 nucléotides, il y a donc 64 codons possibles. Trois codons sont des signaux « stop », un codon signifie le démarrage et code aussi la méthionine. Donc, 61 codons codent les 20 acides aminés.

Le code est dégénéré, mais spécifique.

Beaucoup d'acides aminés ont plusieurs codons, mais chaque codon n'est spécifique que d'un seul acide aminé.

Le code est pratiquement universel, mais pas tout à fait.

Dans certains génomes de mitochondries et de protistes, un codon STOP est lu comme acide aminé ; sinon, le code est universel.

15.3 La transcription des procaryotes

Les procaryotes ont une seule ARN polymérase.

Il existe deux formes d'ARN polymérase procaryote : la polymérase centrale, capable de synthétiser l'ARNm, et la holoenzyme, enzyme centrale plus le facteur σ , capable d'entamer correctement la synthèse (figure 15.6).

L'initiation se situe au niveau des promoteurs.

L'initiation a besoin d'un site de départ et d'un promoteur. Le promoteur se trouve en amont du point de départ et le site -35 de liaison de l'ARN polymérase holoenzyme positionne correctement la polymérase.

L'élongation ajoute les nucléotides successifs.

La transcription progresse dans le sens 5'-3'. La bulle de transcription renferme l'ARN polymérase, l'ADN matrice localement déroulé et le transcrite d'ARNm en croissance (figure 15.7).

La terminaison a une localisation spécifique.

Le terminateur consiste en séquences complémentaires formant une boucle en épingle à cheveux bicaténaire où la polymérase s'arrête (figure 15.8).

La transcription procaryote est couplée à la traduction.

La traduction débute alors que les ARNm sont encore en cours de transcription.

15.4 La transcription eucaryote

L'ARN polymérase I transcrit l'ARNr, la polymérase II transcrit l'ARNm et certains ARNsn, la polymérase III transcrit l'ARNt.

Chaque polymérase possède son propre promoteur.

Le promoteur central s'unit à des facteurs qui recrutent l'ARN Pol II.

Les facteurs de transcription généraux s'unissent au promoteur et recrutent la polymérase pour former le complexe d'initiation. Quand le transcrite est long

de quelque 20 nucléotides, il reçoit une coiffe, et la polymérase peut faire une pause. Des facteurs d'élongation sont nécessaires pour un redémarrage.

Le complexe d'élongation de la transcription peut recruter d'autres facteurs.

Certains facteurs généraux du complexe d'initiation sont remplacés par des facteurs d'élongation. Le domaine carboxyle terminal (DCT) interagit avec les facteurs d'élongation et les enzymes qui modifient le transcrite.

La polyadénylation participe à la terminaison des transcrits d'ARNm.

La poly-A polymérase coupe le transcrite et ajoute de 1 à 200 résidus adénine (A) à l'extrémité 3'.

Les transcrits eucaryotes sont transformés en ARNm.

L'ARNm possède une coiffe 5', une queue poly-A 3' et les introns sont éliminés.

15.5 L'épissage des pré-ARNm

Il peut y avoir des interruptions dans les gènes eucaryotes.

L'ADN codant (exons) est interrompu par des introns non codants. Ces introns sont éliminés par épissage (figure 15.13).

Le spliceosome est l'organite d'épissage.

Les RNTsn reconnaissent les jonctions intron-exon et recrutent des spliceosomes. Le spliceosome finit par réunir l'extrémité 3' du premier exon à l'extrémité 5' du suivant.

L'épissage peut donner plusieurs transcrits d'un même gène.

15.6 Structure de l'ARNt et des ribosomes

Les aminoacyl-ARNt synthétases fixent les acides aminés à l'ARNt.

Le chargement de l'ARNt fixe l'extrémité carboxyle d'un acide aminé à l'extrémité 3' de l'ARNt adéquat (figure 15.15).

Le ribosome possède plusieurs sites de liaison à l'ARNt (figure 15.16).

Un ARNt chargé s'unit d'abord au site A, puis il va au site P, où son acide aminé est lié à la chaîne peptidique et, finalement, sans son acide aminé, il va au site E, d'où il est libéré.

Le ribosome a des fonctions de décodage et d'enzyme.

Les ribosomes maintiennent les ARNt et l'ARNm en position pour permettre à une enzyme ribosomique de former les liaisons peptidiques.

15.7 La traduction

Des facteurs accessoires sont nécessaires à l'initiation.

Chez les procaryotes, la formation du complexe d'initiation est favorisée par la séquence de liaison au ribosome (RBS) de l'ARNm, complémentaire d'une petite sous-unité. Les eucaryotes utilisent la coiffe 5' dans le même but.

L'élongation ajoute les acides aminés successifs.

Tandis que le ribosome se déplace le long de l'ARNm, les nouveaux acides aminés d'ARNt chargés s'ajoutent au peptide en croissance (figure 15.19).

La terminaison a besoin de facteurs accessoires.

Les codons stop sont identifiés par les facteurs de terminaison.

Les protéines peuvent être dirigées vers le RE.

Chez les eucaryotes, les protéines possédant une séquence signal à leur extrémité amino s'unissent au SRP et ce complexe s'amarre au RE.

15.9 Les mutations : des gènes modifiés

Les mutations ponctuelles concernent un seul site de l'ADN.

Les substitutions de bases échangent une base pour une autre et les mutations avec décalage de phase impliquent l'addition ou la délétion d'une base. Les mutations par expansion des triplets peuvent être à l'origine de maladies génétiques.

Les mutations chromosomiques modifient la structure des chromosomes.

Les mutations chromosomiques sont des duplications, délétions, inversions ou translocations.

Les mutations sont le point de départ de l'évolution.

Les nouvelles données ont changé notre manière de voir la nature des gènes.

COMPRÉHENSION

- Les expériences de Beadle et Tatum sur mutants nutritionnels de *Neurospora* ont prouvé que
 - la moisissure du pain peut se développer en laboratoire sur milieu minimum.
 - les rayons X peuvent endommager l'ADN.
 - les cellules ont besoin d'enzymes.
 - les gènes spécifient les enzymes.
- Quel ce que le *dogme central* de la biologie moléculaire ?
 - L'ADN est le matériel génétique.
 - L'information passe directement de l'ADN à la protéine.
 - L'information passe de l'ADN à l'ARN et à la protéine.
 - Un gène code un seul polypeptide.
- Dans le code génétique, un codon
 - consiste en trois bases.
 - spécifie un seul acide aminé.
 - spécifie plusieurs acides aminés.
 - a et b sont corrects.
- La transcription eucaryote diffère de la procaryote par le fait que
 - les eucaryotes n'ont qu'une ARN polymérase.
 - les eucaryotes ont trois ARN polymérases.
 - les procaryotes ont trois ARN polymérases.
 - a et c sont corrects.
- Pourrait-on trouver un anticodon sur un des types suivants d'ARN ?
 - ARNsn (petit ARN nucléaire)
 - ARNm (ARN messenger)
 - ARNt (ARN de transfert)
 - ARNr (ARN ribosomique)
- L'ARN polymérase s'unit _____ pour entamer _____
 - à l'ARNm ; la traduction
 - au promoteur ; la transcription
 - à l'amorce ; la transcription
 - au facteur de transcription ; la traduction
- Pendant la traduction, le codon de l'ARNm est en réalité « lu » par
 - le site A du ribosome.
 - le site P du ribosome.
 - l'anticodon d'un ARNt.
 - l'anticodon d'un acide aminé.

APPLICATION

- Vous disposez de mutants qui affectent tous la même voie biochimique. Si l'addition d'un intermédiaire de cette voie permet la croissance, cela signifie que l'enzyme codée par le gène muté
 - fonctionne après l'intermédiaire utilisé.
 - fonctionne avant l'intermédiaire ajouté.
 - est nécessaire à la production de cet intermédiaire.
 - ne peut fonctionner pour produire cet intermédiaire.
- Le processus d'épissage
 - existe chez les procaryotes.
 - réunit les introns.
 - peut produire beaucoup d'ARNm à partir d'un même transcrit.
 - ne fait que réunir les exons des différents gènes d'une seule façon.

- L'enzyme responsable des liaisons peptidiques est appelée peptidyl *transférase* parce qu'elle transfère
 - un nouvel acide aminé d'un ARNt au peptide en croissance.
 - le peptide en croissance d'un ARNt à l'acide aminé suivant.
 - le peptide d'un acide aminé à un autre.
 - le peptide du ribosome à un ARNt chargé.
- Quand on compare l'expression des gènes chez les procaryotes et les eucaryotes,
 - les gènes eucaryotes peuvent produire plusieurs protéines.
 - les gènes procaryotes peuvent produire plusieurs protéines.
 - tous deux produisent des ARNm colinéaires par rapport à la protéine.
 - a et b sont corrects.
- Le codon CCA peut muter pour donner
 - une mutation silencieuse
 - un codon pour la lysine
 - un codon stop
 - a et b sont corrects
- Une inversion
 - produira nécessairement un phénotype mutant.
 - ne donnera un phénotype mutant que si les points de rupture se trouvent au sein d'un gène.
 - arrête la transcription dans la région inversée parce que le chromosome y est à l'envers.
 - interfère avec la traduction des gènes de la région inversée.
- Quelle est la relation entre les mutations et l'évolution ?
 - Les mutations améliorent les gènes.
 - Les mutations peuvent créer de nouveaux allèles.
 - Les mutations se sont produites au début de l'évolution, mais pas aujourd'hui.
 - Il n'y a pas de relation entre l'évolution et les mutations génétiques.

RÉVISION

- La séquence d'un brin d'ADN modèle est la suivante :
3'-CGTTACCCGAGCCGTACGATTAGG-5'
À partir de l'information de cette séquence, déterminez
 - la séquence de l'ARNm de ce gène.
 - la séquence des acides aminés de la protéine.
- Les mutations avec décalage de phase donnent souvent des protéines tronquées. Expliquez-le en vous basant sur le code génétique.
- Montrez les conséquences des mutations suivantes sur la protéine produite (la protéine débute par le codon START). Donnez le nom de ces mutations.
Brin modèle d'origine :
3'-CGTTACCCGAGCCGTACGATTAGG-5'
 - 3'-CGTTACCCGAGCCGTAACGATTAGG-5'
 - 3'-CGTTACCCGATCCGTACGATTAGG-5'
 - 3'-CGTTACCCGAGCCGTTTCGATTAGG-5'
- Plusieurs caractéristiques sont propres aux bactéries, et d'autres aux eucaryotes. Certaines de ces caractéristiques pourraient-elles offrir la possibilité de contrôler l'expression génique d'une manière spécifique soit aux eucaryotes, soit aux bactéries ?