

## I. Synthèse protéique

La synthèse protéique est l'ensemble des processus biochimiques par les quels une cellule peut produire une chaîne protéique en combinant des acides aminés isolés présents dans son cytoplasme, guidé par l'information génétique contenue dans l'ADN, elle se déroule en 2 étapes essentielles: la transcription de l'ADN en ARN messager (ARNm) et la traduction de l'ARNm en une protéine.

### I.1. Transcription

**I.1.1. Généralité :** C'est la première étape de l'expression génétique. Au cours de laquelle tous les types d'ARN (ARNm, ARNt, ARNr) sont produits à partir d'une matrice d'ADN par un ensemble protéique très complexe : l'**ARN polymérase**. Cependant seuls les ARNm renferment l'information génétique nécessaire à la synthèse des protéines. En effet, un nombre d'observations a indiqué que l'ARN est l'intermédiaire entre l'ADN et la protéine :

- Absence d'ADN dans le cytoplasme qui est le lieu de la synthèse protéique, par contre l'ARN est présent à la fois dans le cytoplasme et le noyau.
- Une corrélation existe entre la quantité d'ARN et la quantité de protéine synthétisée.

Cela a été mis en évidence par des expériences après l'extraction et la visualisation d'ADN et de l'ARN par centrifugation sur un gradient de densité. Il est important de noter que la synthèse d'ARN s'effectue:

- Seulement à partir de certaines parties de l'ADN, autrement seul l'ADN des gènes est transcrit grâce à une ARN polymérase (ADN dépendante),
- A partir de l'un des deux brins du gène, mais ce n'est pas toujours le même brin qui est copié : pour certains gènes ce sera un brin et pour d'autres ce sera l'autre brin
- Dans le sens 5' vers 3',
- De façon antiparallèle
- De façon complémentaire (Fig.01).
- En absence de toute amorce préexistante.

Au cours de la transcription, l'ARN est synthétisé par polymérisation des sous unités ribonucléotides triphosphates (ATP, CTP, GTP, UTP), ces ribonucléotides apportent à la fois les nucléotides monophosphates (AMP, CMP, GMP, UMP) constituant la molécule d'ARN et l'énergie nécessaire à l'ARN polymérase pour relier chaque nucléotide au précédent. L'ARN polymérase catalyse la polymérisation de l'ARN. Le 3'OH d'un ribonucléotides réagit avec le 5'OH d'un ribonucléotides pour former une liaison phosphodiester. L'ordre dans lequel les ribonucléotides sont ajoutés à la chaîne d'ARN en croissance est déterminé par l'ordre des bases sur l'ADN matrice. Les ribonucléotides sont additionnés à la chaîne en croissance au niveau de son extrémité 3'OH libre.

**I.1.2. Mécanisme de la transcription chez les procaryotes :** Chez les procaryotes, l'ARN est synthétisé par une **ARN polymérase unique** formée de cinq sous-unités polypeptidiques ( $\alpha_2\beta\beta'\sigma$ ). Cette forme est appelée **holoenzyme**, la sous unité  $\sigma$  peut se dissocier des autres sous unités pour laisser une forme appelée **enzyme centrale**. Ces deux formes de l'ARN polymérase ont des rôles différents lors de la transcription qui se fait en trois étapes : l'initiation, l'élongation et la terminaison.

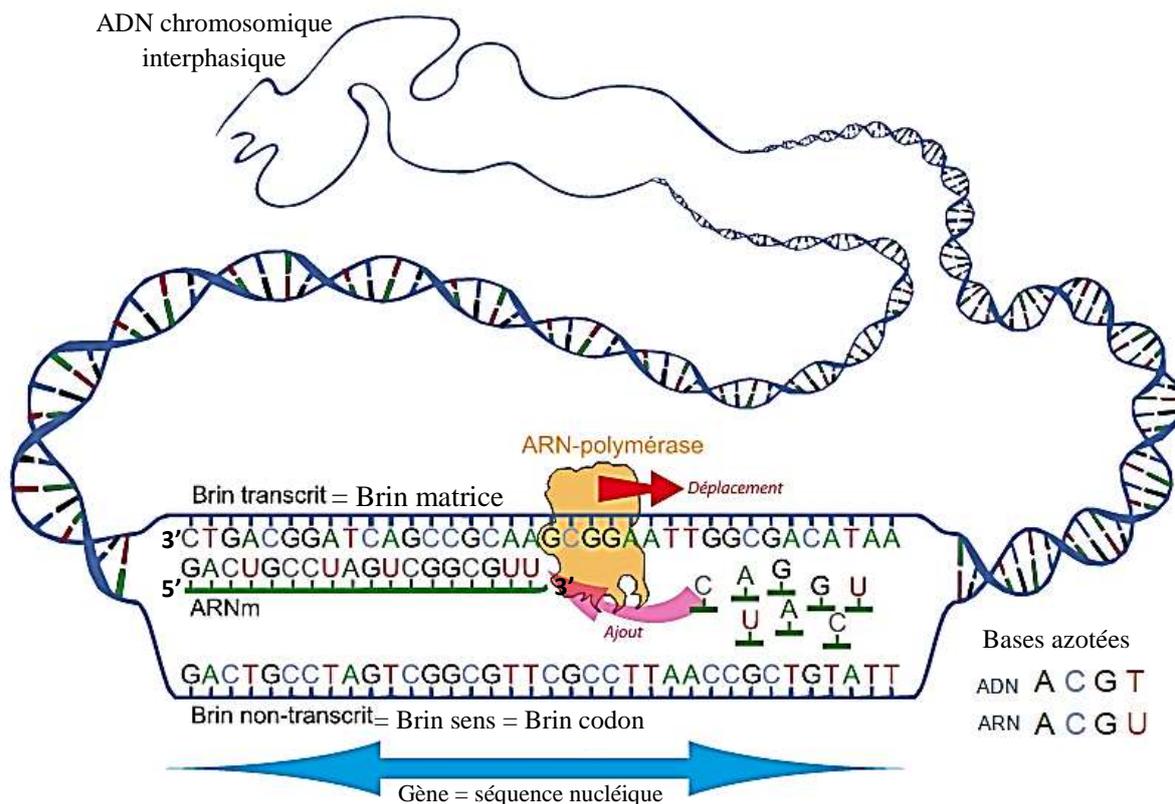


Fig.01 : Transcription d'ADN en ARN

**A. Initiation :** L'initiation correspond à la synthèse de la première liaison phosphodiester catalysée par l'ARN polymérase. La transcription ne peut démarrer de façon aléatoire, elle commence au niveau d'une région d'ADN appelée **promoteur**, définie comme étant une séquence d'une centaine de nucléotides située dans la région régulatrice et désignant le début de la transcription, la séquence de promoteur est située juste avant le début de la région où démarrera la transcription. Il est important de noter que chez les procaryotes, plusieurs gènes peuvent être transcrits à partir d'un **promoteur unique**. Le promoteur contient des séquences d'ADN spécifiques qui sont des points d'attache pour l'ARN polymérase. Chez *E. coli* deux éléments de séquences reconnus par l'ARN polymérase sont présents au sein des promoteurs : la **séquence -10** et la **séquence -35** (par référence au premier nucléotide d'ARN incorporé). La sous unité  $\sigma$  de l'ARN polymérase est responsable de sa reconnaissance et de sa liaison au promoteur (diminue l'affinité de l'enzyme pour les régions non promotrices). Une fois l'ARN polymérase mise en place, la sous unité  $\sigma$  se détache du complexe et  $\beta'$  assure la liaison à l'ADN. Lorsque l'enzyme se fixe au promoteur, elle commence par former un **complexe promoteur fermé** au sein duquel l'ADN est sous forme de double hélice. L'ARN-polymérase entraîne par la suite la dénaturation (dissociation) des deux brins d'ADN sur 14 paires de nucléotides conduisant à l'ouverture de la double hélice qui se dissocie partiellement au niveau de la boîte -10 (riche en liaisons faibles A-T) formant : **un complexe promoteur ouvert**. La sous unité  $\sigma$  se dissocie ensuite du complexe promoteur ouvert pour laisser l'enzyme centrale. Cette dernière commence à ajouter les ribonucléotides correspondants (Fig. 02).

**B. Elongation :** L'élongation correspond au déplacement de la bulle de transcription le long de la molécule d'ADN dans le sens 5' 3'. L'ARN polymérase ajoute des ribonucléotides à l'extrémité 3'OH de la molécule d'ARN dans l'ordre spécifié par le brin d'ADN matricier. La polymérisation d'ARN

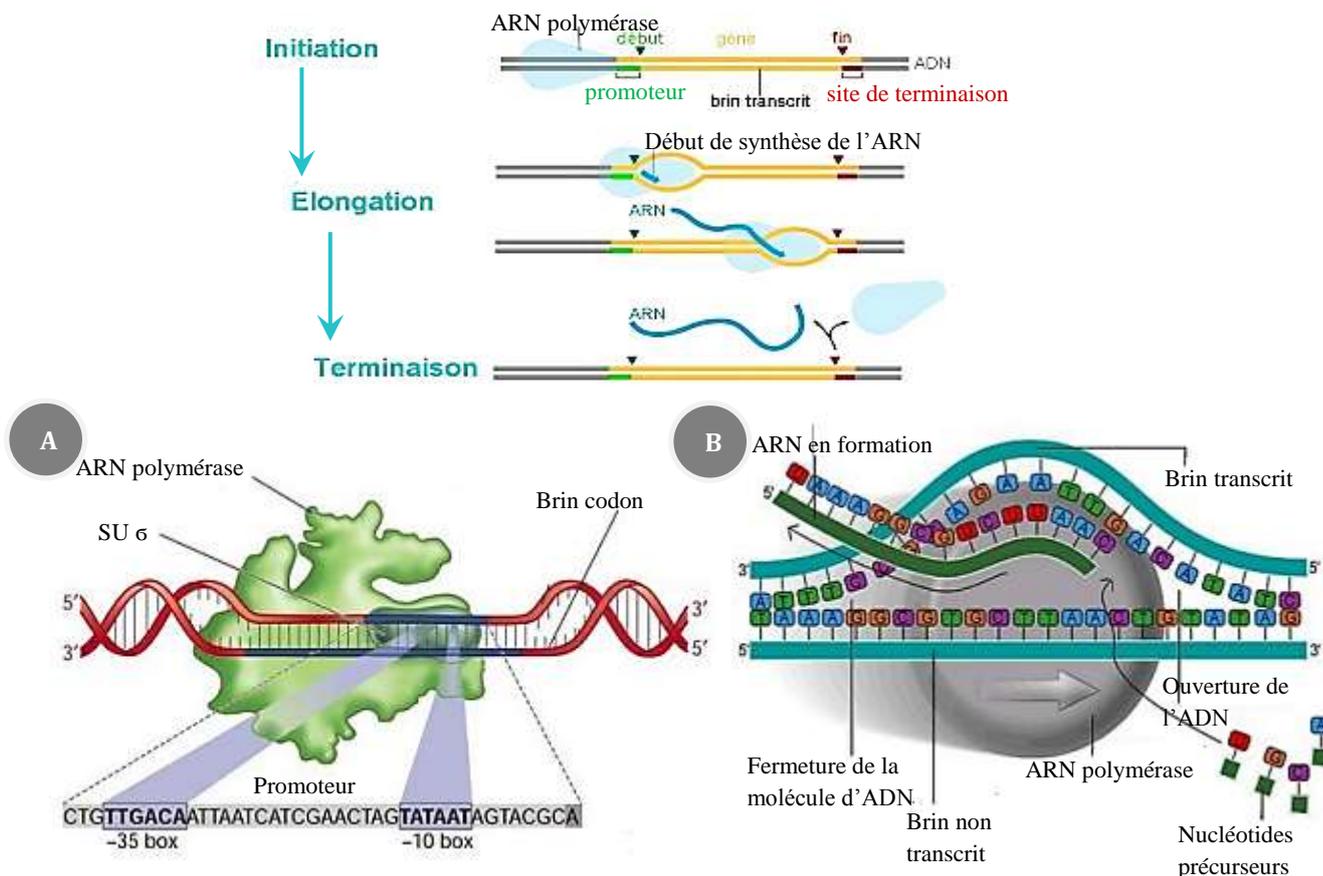


Fig.02: Mécanisme de la transcription chez les procaryotes. A : Initiation, B : Elongation

est catalysée par l'ARN polymérase suite à l'utilisation de l'énergie fournie par les ribonucléotides triphosphates comme substrat. Au fur et à mesure que l'ARN polymérase se déplace, l'ARN naissant se détache de la matrice d'ADN. Les régions d'ADN situées derrière la polymérase regagnent leur forme en double brin par rétablissement des liaisons hydrogènes entre les bases complémentaires. Seule une petite portion de la double hélice est déroulée à la fois. Dont la zone déroulée comporte à la fois l'ARN néosynthétisé apparié avec l'ADN matrice qui s'étend sur 12 à 17 bases (hélice hybride ADN-ARN). Cette zone doit rester de faible longueur car dérouler la double hélice dans une région implique de la « super-enrouler » dans les régions adjacentes, ce qui impose des tensions au niveau de la molécule d'ADN.

**C. Terminaison :** Définie comme étant le processus conduisant à la dissociation du complexe ARN polymérase de l'extrémité 3'OH du gène et la libération de polymère d'ARN transcrit. L'arrêt de la transcription chez les procaryotes fait intervenir des séquences « **terminateur** » et des **facteurs de terminaison (facteur rho)**. Il existe deux mécanismes de terminaison de la transcription chez les procaryotes, l'un qui fait intervenir le facteur de terminaison rho, et l'autre non:

✓ **La terminaison Rho-indépendante :** Dans ce cas, le terminateur se présente sous la forme d'un palindrome : une séquence répétée inversée suivie d'une série de T (uraciles sur l'ARN transcrit). Ce palindrome entraîne une complémentarité de séquence au niveau de l'ARN qui permet la mise en place d'une structure en tige-boucle qui déstabilise l'ARN-polymérase jusqu'à dissociation. La structure tige-boucle riche en paires de bases G-C, est suivie d'une séquence poly-U d'environ 6 nucléotides permettant une dissociation plus facile de l'hybride ADN-ARN (Fig. 3).

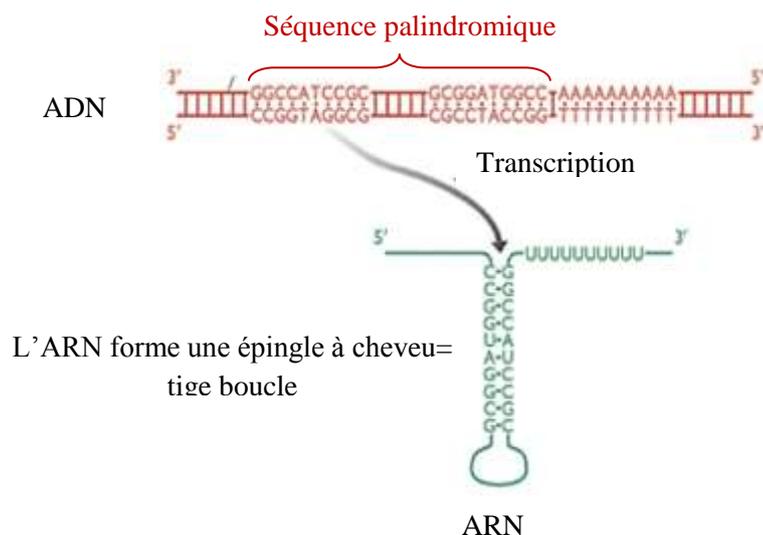


Fig.03: Terminaison Rho-indépendante (palindrome).

✓ **La terminaison Rho-dépendante** : constitués d'une séquence **consensus** reconnue par la protéine Rho. Le facteur rho a une affinité pour les ARN en court de synthèse, le parcourant de 5' vers 3' jusqu'à trouver l'ARN-polymérase. Le facteur rho est ATP dépendante, dont l'hydrolyse permettra la dissociation du complexe (l'ARN polymérase de l'ADN matrice)

### I.1.3. Mécanisme de la transcription chez les eucaryotes

La transcription a lieu chez les eucaryotes de la même manière que chez les procaryotes. Toutefois certaines différences existent :

**A.** la transcription est réalisée par trois enzymes : les ARN polymérase I, II et III, dont chacune transcrit un certain nombre de gène d'une façon légèrement différente. Ces enzymes différents elles-mêmes par leur localisation dans le noyau, par la nature des ARN formés et par leur sensibilité à des inhibiteurs tels que l' $\alpha$ -amanitine.

- ✓ **ARN-polymérase I** : catalyse la transcription des gènes codant pour trois des quatre ARN ribosomiaux (5.8S, 18S et 28S), localisée au niveau du nucléole et elle est insensible à l' $\alpha$ -amanitine.
- ✓ **ARN-polymérase II** : cette enzyme transcrit les gènes qui codent les protéines (ARNm), elle est localisée au niveau du nucléoplasme et elle est sensible à l' $\alpha$ -amanitine.
- ✓ **ARN-polymérase III** : cette enzyme transcrit de courts gènes codant les ARNt et l'ARNr 5S, elle est localisée au niveau du nucléoplasme et elle est également sensible à l' $\alpha$ -amanitine mais à hautes doses.

**B.** l'initiation est un processus plus complexe : La reconnaissance du promoteur implique de nombreux facteurs protéiques spécifiques des différentes ARN polymérase : ARNpol II (TFIIA-H), ARN pol I (TAFs), ARN pol III (TFIIIA-C).

**C.** Le mécanisme de terminaison de la transcription chez les procaryotes est différent de celui des eucaryotes, chez ces derniers la terminaison ne fait pas intervenir de structure tige-boucle.

## I.2. Maturation de l'ARN transcrit

La maturation se déroule entre le moment où ils sont synthétisés les ARN et celui où ils jouent un rôle dans la synthèse des protéines. Les ARN subissent un certain nombre de modifications post-transcriptionnelles qui diffèrent selon le type d'ARN transcrit :

- **ARNt** : le précurseur d'ARNt subira des clivages, d'addition d'un motif « CCA » en 3' par une nucléotidyltransférase et d'autres modifications comme des méthylations (formant de la thymine à partir de l'uracile) et des désaminations (désamination de l'Adénosine et hypoxanthine). Le transcrit primaire d'ARNt chez les eucaryotes diffère de son homologue chez les procaryotes par la présence d'un intron de petite taille qui est enlevé au cours de sa maturation.

- **ARNr** : Ces modifications correspondent à des clivages successifs conduisant à la perte de certains segments du transcrit initial (mais pas l'épissage).

- **ARNm** : La maturation de l'ARNm est un phénomène complexe est propre à la cellule eucaryote. Chez les procaryotes ce phénomène n'existe pas, le début de la traduction de l'ARNm se faisant avant la fin de la transcription. Cependant chez les eucaryotes, l'ARNm doit subir d'importantes modifications avant qu'il puisse être traduite, et avant de subir ces modifications, l'ARNm est appelé transcrit primaire ou ARN pré-messager.

**A. Addition de la coiffe « cap » au niveau de l'extrémité 5' :** La coiffe correspond à l'ajout d'un nucléotide modifié : la 7-méthylguanosine dit « m7G » au premier nucléotide de l'ARNm par une liaison 5'-5' triphosphate inhabituelle. Ce groupement m7G correspond à l'addition de trois groupements phosphates et d'une molécule de GTP au niveau de l'extrémité 5' du transcrit primaire grâce à l'énergie libérée par l'hydrolyse de la molécule de GTP. Le nucléotide G va ensuite être méthylé sur le septième atome (C7) pour donner la 7-méthyl-guanosine (Fig.04). La coiffe est ajoutée grâce à un complexe protéique appelé « Cap-Binding-Complex » possédant une activité triphosphatase, une activité guanylyl-transférase et une activité méthyl-transférase. La coiffe protège l'extrémité 5' des ARN contre la dégradation par des exonucléases 5' -> 3', augmentant ainsi leur temps de demi-vie et elle participe au transport des ARN vers le cytoplasme.

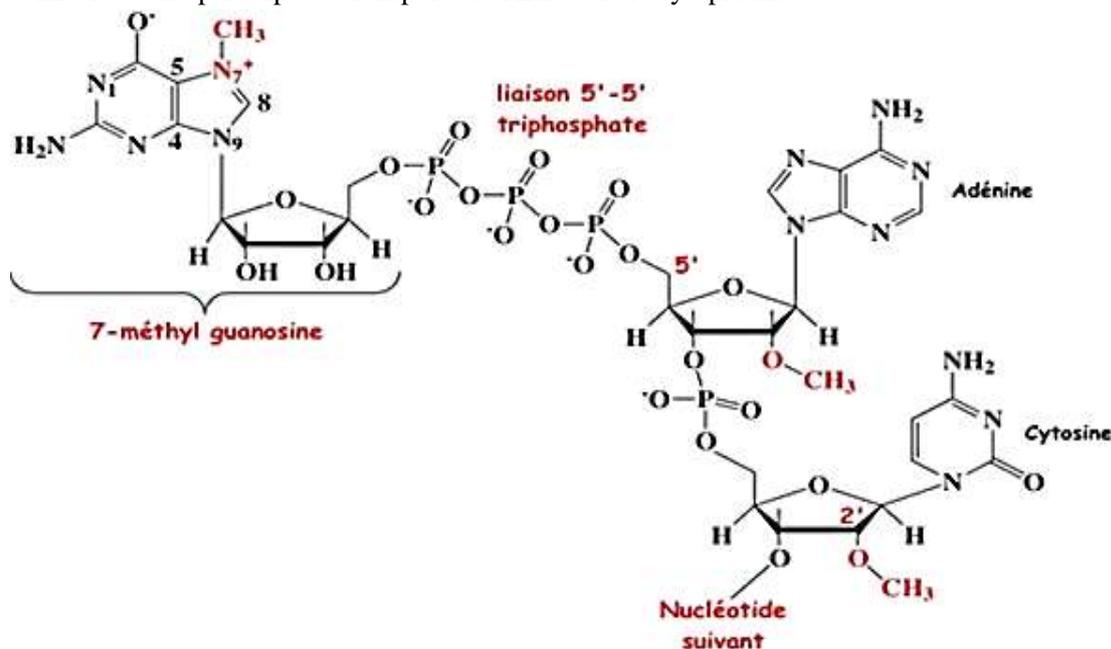


Fig.04: Addition de la coiffe au niveau de l'extrémité 5' de l'ARN

**B. Addition du poly adénine (Poly-adénilation) en extrémité 3'** La poly-adénilation correspond à l'ajout de 100 à 200 adénines à l'extrémité 3' du transcrit primaire et ceci sans matrice. La queue poly (A) est ajoutée après transcription par une enzyme appelée la poly-A-polymérase utilisant l'ATP comme substrat. On pense que la queue poly (A) aiderait au passage de l'ARNm du noyau vers le cytoplasme et qu'elle protégerait l'ARNm au cours de la traduction.

**C. Excision des introns et épissage des exons (*splicing*):** Après l'addition de la coiffe et la poly-adénilation, le transcrit primaire est encore soumis à l'excision des introns (opération correspond à la coupure et l'élimination des introns) et l'épissage des exons (réunion des segments d'exons qui vont donc être soudés bout à bout). Ceci est possible par la présence de site donneur d'épissage (dinucléotide GU) à l'extrémité 5' des introns et de site accepteur d'épissage à l'extrémité 3' des introns (Fig.05)

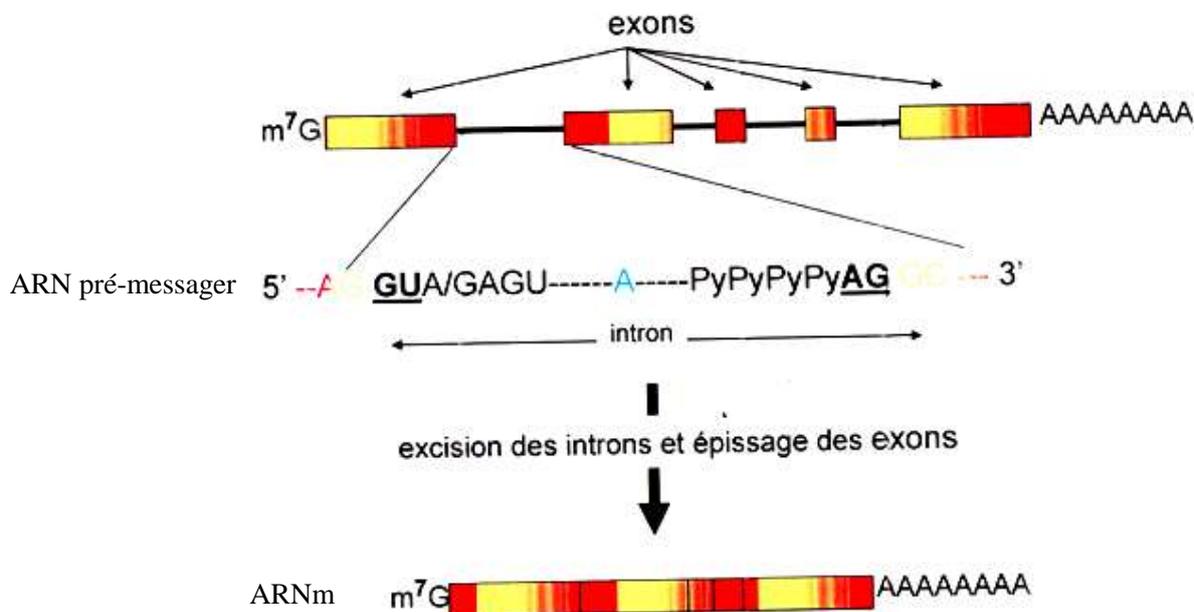


Fig.05 : Excision des introns et épissage des exons (*splicing*)

### I.3. Traduction

C'est le processus par lequel l'information contenue dans la séquence des bases d'ARNm est traduite en séquence spécifique d'acide aminé par les ribosomes. L'ARNm porte dans sa structure un message constitué d'une série de « **codons** » alignés sur cette molécule.

#### I.3.1. Code génétique

Le code génétique est l'ensemble des règles permettant de passer du langage nucléique élaboré par les 4 bases de l'ARNm au langage protéique élaboré par 20 signes représentés par les 20 acides aminés susceptibles d'entrer dans la constitution d'un polypeptide. L'unité élémentaire de ce code est le **codon**, formé par trois bases successives sur l'ARNm. Puisque il y a 4 nucléotides différents et que chaque codon en comporte 3, il existe au total  $4^3 = 64$  codons possibles. Comme il y a 20 acides aminés, il y a donc 44 codons supplémentaires. Trois codons correspondent à des codons **stops** ou non-sens, les autres sont des synonymes qui codent pour différents acides aminés (Tab.01). Le code génétique est dit dégénéré puisque chaque acide aminé est codé par plusieurs codons (à l'exception du tryptophane et de la méthionine). L'initiation des chaînes protéiques est assurée par le codon AUG (code méthionine), les triplets UAA, UAG et UGA sont des codons stop assurant la terminaison des

chaines peptidiques. Les codons sont lus dans le sens 5' vers 3' le long de la chaîne de l'ARNm. A un codon correspond un **anticodon** (séquence de 3 nucléotides successifs sur un ARNt), spécifiant un acide aminé.

Tab.01 : Le code génétique au niveau de l'ARNm

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

En effet, il a été observé une utilisation préférentielle de certains codons pour une espèce donnée, ainsi l'usage des codons diffère d'une espèce à l'autre. Enfin l'universalité du code génétique présente quelques exceptions. Où par exemple, le codon AGA correspond à une arginine dans une cellule eucaryote et à un codon stop dans une mitochondrie.

### I.3.2. Mécanisme de la traduction

La traduction est très proche chez les procaryotes et chez les eucaryotes et à lieu en trois étapes essentielles à savoir : l'initiation après l'activation des acides aminés, l'élongation et la terminaison.

Il existe au moins 20 enzymes de type aminoacyl-ARNt synthétase qui permettent la charge des acides aminés sur les ARNt. Ce processus peut être simplifié en deux étapes :



**A. Initiation :** La traduction commence par la fixation d'une petite sous-unité ribosomale à l'ARNm au niveau de sa séquence de Shine-Dalgarno chez les procaryotes et au niveau de la coiffe chez les eucaryotes. Une fois fixée, La sous unité migre ensuite le long de l'ARNm jusqu'à rencontrer le codon AUG de méthionine. Un ARNt chargé avec de la Met se fixe à l'AUG localisé par la petite sous unité ribosomale. La combinaison de l'ARNm, de la petite sous unité ribosomale et de l'ARNt est appelée complexe d'initiation. Chez les bactéries, la méthionine Met-ARNt est formylée et les facteurs d'initiation (protéines) sont trois : IF1, IF2 et IF3. Les deux facteurs IF1 et IF3 empêchent que la grosse sous unité ribosomale se fixe avant que l'initiation ne soit achevée. Tandis que le IF2 transporte

le Met-ARNt jusqu'au complexe d'initiation. Chez les eucaryotes, au moins 9 facteurs d'initiation sont mis en jeu, eIF1 et eIF2 ont des fonctions similaires à celles d'IF1 et d'IF2 des procaryotes. Après l'ajout de la grande sous unité, le ribosome est maintenant constitué et fonctionnel. Les ribosomes comportent trois sites de liaisons :

- **site A** (de l'Aminoacyl-ARNt) : Où viendra se placer l'ARNt porteur de l'acide aminé.
- **site P** (du Peptidyl-ARNt) : Où viendra se placer l'ARNt porteur de la chaîne peptidique en cours d'élongation.
- **site E** (du sortie ou Exit) : où viendra se placer l'ARNt avant d'être libéré du ribosome

**B. Elongation** : Dès que le facteur d'initiation est formé, une grosse sous-unité ribosomale se lie au complexe d'initiation. L'élongation commence lorsqu'un ARNt entre dans le site A et s'apparie avec le second codon. Quand les deux sites sont occupés par des ARNt chargés, les acides aminés attachés sont placés en contact étroit et une liaison peptidique peut se former entre le groupement carboxyle de la méthionine et le groupement amine du second acide aminé. La réaction est catalysée par un complexe enzymatique appelée peptidyl transférase. La formation de la liaison peptidique donnant le dipeptide et le processus de détachement du premier ARNt (chargé avec Met) se font simultanément. A ce moment il y a formation d'un dipeptide logé dans le site A et porté par le 2<sup>ème</sup> l'ARNt. Après la formation de la liaison peptidique, une translocation a lieu et le ribosome se déplace jusqu'au codon suivant, autrement le ribosome va avancer d'un cran (3 nucléotides) sur l'ARNm dans la direction 5' 3'. Le dipeptide formé et qui est attaché au 2<sup>ème</sup> ARNt se déplace jusqu'au site P en éjectant l'ARNt non chargé qui se trouve dans le site E et le site A devient vacant. Un 3ème ARNt entre dans le site A et le cycle d'élongation est répété.

**C. Terminaison** : La traduction se termine lorsque le ribosome en avançant d'un cran sur l'ARNm trouve un codon stop (UAA, UAG ou UGA). Il n'existe pas d'ARNt capables de se fixer aux codons stop, mais des protéines appelées facteurs de terminaison (ou de libération, *release*) qui entrent dans le site A et provoquent la libération du polypeptide achevé. La liaison entre le dernier ARNt et la chaîne peptidique est hydrolysée par la peptidyl transférase libérant la chaîne peptidique. Le ribosome se dissocie en 2 sous unités qui pourront recommencer de nouvelles lectures de l'ARNm (Fig.06).

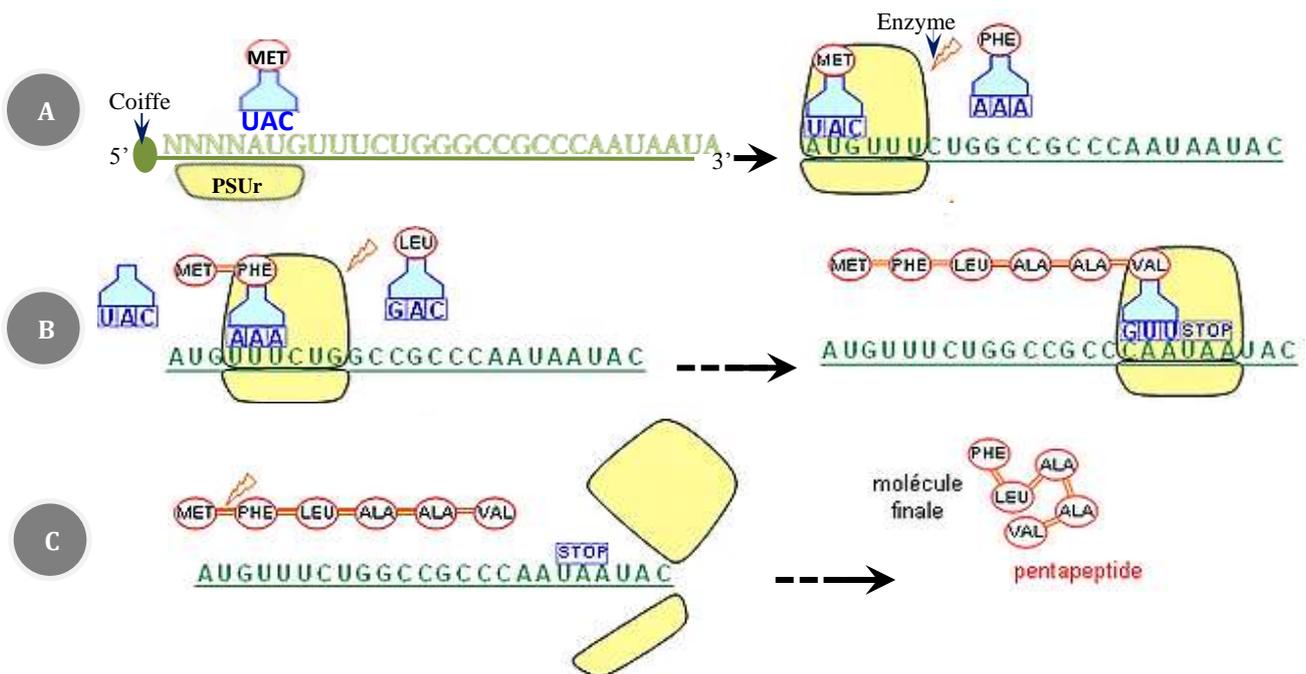


Fig.06: Différentes étapes de l'expression de l'information génétique (synthèse protéique). **A** : Initiation, **B** : Elongation, **C** : Terminaison.

## II. Régulation de l'expression des gènes

Dans cette partie, nous allons utiliser la cellule procaryote comme exemple de contrôle de l'expression de gènes pour dégager des notions fondamentales de régulation. L'opéron lactose est le premier et l'un des exemples les plus connus de la régulation des gènes procaryotes. Dans l'opéron lactose, on trouve les trois gènes indispensables à la dégradation du lactose, il code pour les protéines suivantes:

- La  $\beta$ -galactosidase (gène **lacZ**) : catalyse l'hydrolyse de la liaison  $\beta$ 1-4 des  $\beta$ -galactosides.
- La lactose perméase (gène **lacY**) : cette protéine membranaire permet l'entrée du lactose.
- La thiogalactoside transacétylase (gène **lacA**). Son rôle n'est pas bien connu. Elle acétyle les  $\beta$ -galactosides non métabolisables qui peuvent alors être éliminés hors de la cellule par diffusion à travers la membrane plasmique.

Ces trois gènes de structure sont précédés par une région responsable de la régulation de leur expression. Cette région régulatrice comprend le promoteur et l'opérateur. En amont de l'opéron lactose on trouve le gène **Lac I** qui est un gène régulateur. Il a son propre promoteur (non inductible) et exprime "en continu" un répresseur qui bloque l'expression de l'opéron lactose.

- ✓ En absence de lactose : le répresseur est sous sa forme active (tétramérique), il se fixe au niveau du site opérateur O, empêchant la fixation de l'ARN polymérase et donc l'expression de l'opéron lactose (Répression).
- ✓ En présence de lactose : le répresseur se complexe avec l'allolactose (isomère du lactose). Le répresseur lié à l'allolactose change de conformation, perd son affinité pour l'opérateur et se dissocie de l'opérateur lac. L'opéron lactose peut alors être exprimé (Induction).

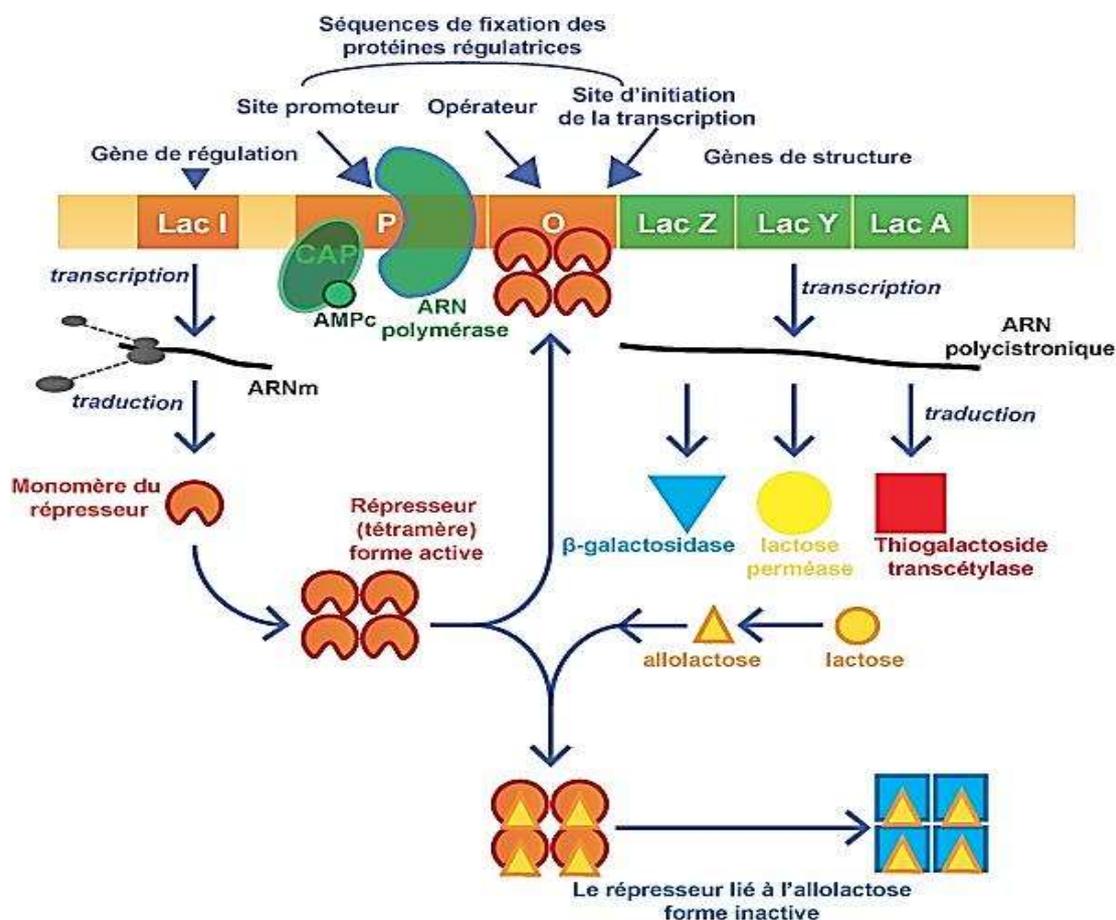


Fig.07 : Opéron lactose (*lac operon*).