

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université A/Mira de Bejaia
Faculté des Sciences de la Nature et de la Vie
Département de Science Alimentaire



Cours sur les outils de la biologie moléculaire

Destiné aux
Master 1 Production et Transformation Laitière (PTL)

D^r TAFININE née MOUHOUBI Zina

Maitre de Conférences classe B

2018-2019

Table des matières

Première partie: Biologie moléculaire

Chapitre I : L'ADN, molécule de l'information génétique

I. Structure de l'ADN	2
II. Caractéristiques de l'ADN	5
III. Réplication de l'ADN	9

Chapitre II : Expression de l'information génétique

I. Transcription.....	12
II. Traduction.....	18

Chapitre III : Régulation de l'expression génétique

I. Chez les procaryotes	30
II. Chez les Eucaryotes.....	38

Deuxième partie : Génie génétique

Chapitre I : Les techniques du gène recombiné (clonage moléculaire)

I. Objectif	45
II. Principe du clonage	45
III. Les enzymes utilisés dans le clonage.....	46
IV. Les vecteurs de clonage	50
IV. Les cellules hôtes	55

Chapitre II : Les banques d'ADN

I. Définition	57
I.1. Les banques d'ADNc.....	57
I.2 Les banques d'ADN génomiques	62
II. Criblage des clones recombinants	64

Chapitre III : Méthodes d'analyse du gène purifié

I. Détection des fragments d'ADN et d'ARN par des techniques de transfert	70
II. Séquençage des acides nucléiques.....	75

Liste des figures

Figure 1 : Les bases pyrimidiques.....	2
Figure 2 : Les bases puriques	3
Figure 3 : Liaison <i>sucre-base</i>	4
Figure 4 : Etapes de l'élongation dans la transcription chez les Procaryots.....	14
Figure 5 : Excision et épissage du transcrit primaire	17
Figure 6 : Formation d'un ARNt.....	19
Figure 7 : Structure d'un ARNt.....	19
Figure 8 : Différence entre le ribosome procaryotique et eucaryotique.....	20
Figure 9 : Initiation chez les procaryotes et les eucaryotes	25
Figure 10 : Etapes de l'élongation.....	28
Figure 11 : Etapes de la terminaison	29
Figure 12 : Organisation des gènes de l'opéron lactose.....	32
Figure 13 : Régulation de l'expression de l'opéron lactose	32
Figure 14 : Régulation de l'opéron lactose par la protéine CAP	33
Figure 15 : Régulation de l'opéron tryptophane	34
Figure 16 : Régulation de l'opéron trp par atténuation	35
Figure 17 : mécanisme de l'autorégulation chez les procaryotes (exp : proteine r).....	37
Figure 18 : Schéma des principaux types de facteurs de transcription	41
Figure 19 : Régulation post transcriptionnelle chez les eucaryotes	43
Figure 20 : régulation de la traduction chez les eucaryotes (exp : la ferritine).....	43
Figure 21 : Les étapes du clonage moléculaire	45
Figure 22 : Action de la phosphatase alcaline	50
Figure 23 : Les plasmides pBR322 et Puc 19	52
Figure 24 : Cycles de multiplication du phage lambda.....	54
Figure 25 : Construction d'une banque d'ADNc	58
Figure 26 : Isolement des ARNm.....	59
Figure 27 : Synthèse de l'ADNc par la Technique originelle	60
Figure 28 : Technique de synthèse de l'ADNc par addition de queue uniforme (tailing)	61
Figure 29 : Technique de synthèse de l'ADNc par cassure à la RNase H.....	62
Figure 30 : Construction d'une banque d'ADN génomique	63
Figure 31 : Criblage par des sondes nucléaires	65

Figure 32: Marquage par Nick translation (translation de coupure)	66
Figure 33: Marquage par Random priming (amorçage au hasard)	67
Figure 34: Technique de criblage par des anticorps.....	69
Figure 35: Technique de southern blot.....	71
Figure 36 : La technique de Northern blot	72
Figure 37: Les puces à ADN	74
Figure 38 : Méthode enzymatique de Sanger.....	76
Figure 39 : Séquençage de l'ADN selon Maxam et Gilbert	77

Liste des tableaux

Tableau I: Le code génétique.....	21
Tableau II: Facteurs de l'élongation.....	28
Tableau III: Facteurs de terminaison.....	29

Première partie

Biologie moléculaire

Chapitre I

ADN, support de l'information génétique

Préambule

La variété des organismes vivants est extraordinaire, on estime qu'il ya de nos jours plus de 10 millions d'espèces. Chaque espèce est différente des autres. Cette différence est due au contenu en information génétique de chaque espèce. Les généticiens ont envisagé que des molécules spécifiques étaient porteuses de cette information génétique. Voici quelques expériences qui ont conduit à la découverte de l'ADN comme matériel génétique:

En 1869: Fredrich Micher utilisa une protéase (pepsine: découverte en 1836) pour séparer le noyau du cytoplasme. Il a observé une petite tache de poudre grise qui s'était détachée d'un liquide claire jaunâtre (cytoplasme). Il l'appela nucléine.

En 1889: Le biologiste Altman a précisé que la nucléine c'est un acide nucléique.

En 1928 : L'expérience de Griffith a montrée que la souche S tuée ne tue pas la souris, mais lorsqu'elle est mélangée avec la souche R vivante, la souris injectée meurt. Comme conclusion la souche R est transformée en souche S, et le facteur transformant est l'ADN.

En 1944: Avery O.T., Macleod C. M. et Mc Carty M. (Rockfeller institute - New York) ont repris les expériences effectués en 1928 par Fred Griffith en utilisant les extraits de différentes biomolécules de la cellule S traités ou non par des enzymes spécifiques. Les résultats obtenus montrent qu'il ya transformation d'un nombre de souches R en souche S lorsque l'ADN purifié est utilisé (selon la fréquence de transformation). Les mêmes résultats ont été obtenus en dégradant les protéines et les ARN par des enzymes spécifiques. En conclusion, l'ADN est le facteur transformant et le support de l'information génétique

En 1946: Lederberg et Tatum ont recombiné génétiquement des bactéries de souches auxotrophes incapables de pousser sur un milieu dépourvu de facteurs de croissance).

En 1950: Chargaff a fait l'équivalence en bases dans l'ADN des animaux, végétaux, bactéries et phages. En conclusion il a trouvé que le rapport $A + G / T + C \approx 1$ (purines / pyrimidines ≈ 1) chez tous les organismes testés et que le rapport AT / GC est variable selon les espèces.

En 1952-1953: - Franklin a obtenu une excellente photographie d'ADN par diffraction des rayons X. La même année Watson et Crick utilisant les données de Chargaff et l'image obtenue par Franklin ont établi le modèle de la double hélice (la forme B).

Chapitre I : ADN, support de l'information génétique

I. Structure de l'ADN

Les acides nucléiques sont de très longues molécules, formées par la répétition de sous unités que l'on appelle « nucléotides ». La molécule d'ADN est composée de deux chaînes complémentaires formées par l'enchaînement d'un nombre élevé de monomères (de nucléotides).

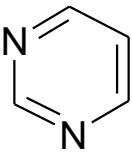
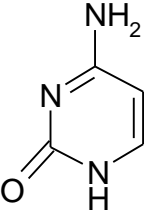
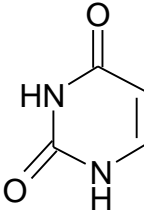
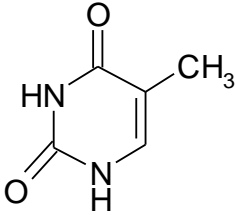
I. 1. Les nucléotides et leurs constituants

Un nucléotide est lui-même constitué de trois éléments :

Nucléotides = Acide phosphoriques + Sucre + Base

I.1.1. La base (les bases azotée)

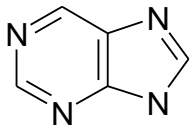
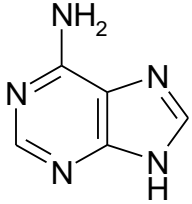
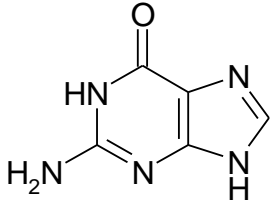
a) *Les bases dites pyrimidiques* : elles possèdent toutes un cycle pyrimidine et divers constituants qui viennent se greffer sur ce cycle. Le noyau pyrimidine est le plus simple : c'est un noyau aromatique à six atomes, quatre carbones et deux azotes, les deux azotes en position méta (n° 1 et 3) (figure 1).

			
Pyrimidine	Cytosine	Uracile	Thymine
	2-oxy-4-aminopyrimidine	2,4dioxypyrimidine	5-méthyl-2,4dioxypyrimidine

- Numérotation de 1 à 6 des atomes

Figure 1 : Les bases pyrimidiques

b) *Les bases dites puriques* : le noyau purine est constitué de deux noyaux hétérocycliques accolés, un de six atomes et l'autre de cinq atomes, ayant deux carbones en commun au milieu. Par rapport à ces carbones communs, les azotes occupent des positions symétriques (n°1 et 3 à gauche, n° 7 et 9 à droite) (figure 2).

		
Purine	Adénine	Guanine
Imidazopyrimidine	6-aminopurine	2-amino-6oxypurine

- Numérotation de 1 à 9 des atomes

Figure 2 : Les bases puriques

Les doubles liaisons et la nature cyclique confèrent une grande rigidité à ces molécules. La présence de doubles liaisons conjuguées confère également à ces molécules la propriété d'absorber fortement les photons (les bases absorbent fortement dans l'ultraviolet à 260nm).

Les nucléotides de l'ADN, comme ceux de l'ARN ne comportent que quatre de ces bases azotées :

- deux puriques communes aux deux types d'acides nucléiques
- une pyrimidique commune : la cytosine
- une pyrimidique spécifique : l'uracile pour l'ARN et son dérivé méthylé, la thymine pour l'ADN.

I.1.2. Le sucre

On trouve deux types de sucres dans les acides nucléiques :

a) *Le ribose (D-ribose)*

Le ribose est un ose en cinq carbones. Ce nom provient des initiales de l'institut où il a été découvert « **R**ockefeller **I**nstitutes of **B**iochemistry », à New York (« R.I.Bose »), on numérote les atomes de carbone du ribose avec des « primes » pour éviter les confusions avec les numéros des bases.

b) *Le désoxyribose (2'-désoxy-D-ribose)*

Le désoxyribose est un ribose dans lequel il manque un OH en 2' (cet OH ayant été remplacé par un H).

C'est un ose à 5 atomes de carbone, donc un pentose. La numérotation des carbones est affectée d'un exposant prime (') pour éviter toute confusion avec les numéros des carbones des autre constituants de l'ADN.

Il comporte trois fonctions alcool : une fonction alcool primaire portée par le carbone 5', deux fonctions alcool secondaire au niveau des carbones 1' et 3'. Comportant de nombreux hydroxyles.

I.1.3. L'acide phosphorique

L'acide phosphorique est un tri-acide, Il est ionisé au pH cellulaire. Deux des trois fonctions acides seront estérifiées dans les DNA et RNA.

I.1.4. Association des trois éléments constituant le nucléotide

a. Liaison : Sucre-Base

La liaison qui unit l'ose et la base est une liaison β -osidique (N glycosidique) (figure 3). Cette liaison se forme par élimination d'une molécule d'eau entre l'OH semi-aldéhyde de l'ose et un H de la base pyrimidique (H en 1) ou purique (H en 9). Cet assemblage *ose-base* s'appelle « **nucléoside** » (ne pas confondre avec nucléotide).

Par rotation autour de la liaison N-glycosidique, la base peut être soit du même côté, par rapport à cette liaison, que le pentose (*conformation syn*), soit de l'autre côté (*conformation anti*).

Dans le DNA (A- DNA OU B-DNA), toutes les bases puriques et pyrimidiques ont la même conformation (anti). Dans une autre forme de DNA (Z-DNA) la cytosine est sous forme anti, la guanine et sous forme syn (les conformations anti et syn alternent).

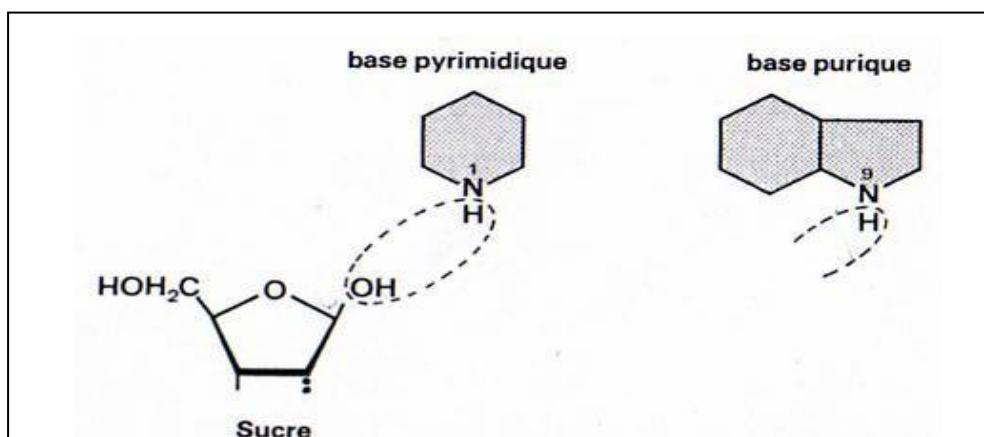


Figure 3 : Liaison sucre-base

b. Liaison : acide phosphorique-sucre

La liaison entre l'ose et l'acide phosphorique est une liaison ester. Cette liaison se forme par élimination d'une molécule d'eau entre :

- OH d'un acide : il s'agit ici d'un OH de H_3PO_4 .
- H d'un alcool : il s'agit ici de l'hydrogène de la fonction alcool en 5' de l'ose.

c. Liaisons reliant les nucléotides

Dans un acide nucléique, les nucléotides sont assemblés entre eux par des liaisons esters. Une molécule d'eau est éliminée entre :

- OH d'un acide : il s'agit ici d'un OH de H_3PO_4 .
- H d'un alcool : il s'agit ici d'hydrogène de la fonction alcool en 3'

Ainsi, l'acide phosphorique engage deux fonctions acides dans des liaisons dites « phosphodiester » (une fonction ester servant à former la nucléotide, la 2^{ème} à relier deux nucléotides entre eux). La 3^{ème} fonction acide de H_3PO_4 reste libre et confère donc des propriétés acides aux acides nucléiques DNA et RNA.

II. Caractéristiques de l'ADN

II.1. Caractéristiques propres aux ADN

Trois caractéristiques sont propres aux ADN, et vont le différencier des ARN.

- **L'ose** : Comme les initiales « DNA » l'indiquent, l'ose entrant dans la constitution du DNA est du désoxyribose (et non pas du ribose comme ce sera le cas dans les RNA).
- **Les bases** : Les bases constituant les nucléotides du DNA sont : A, G, C et T.
- **Les deux chaînes de nucléotides** : Une molécule de DNA est habituellement formée de deux chaînes (on dit aussi deux brin) de nucléotides, alors qu'une molécule de RNA n'en comprend qu'une. On note cependant des exceptions chez certains virus.

II.2. Caractéristiques des deux chaînes de DNA

Ces deux chaînes ont trois propriétés essentielles:

II.2.1. Antiparallèles : Signifie, que les deux brins de nucléotides sont parallèles mais dans des directions opposées.

II.2.2. Complémentaires : La règle de complémentarité est la suivante, en face de A on a T et en face de C on a G. ainsi, si on connaît les bases du brin I, celles du brin II seront, grâce à

la règle de complémentarité, obligatoirement connues. Cette complémentarité s'explique pour les raisons suivantes :

- **Pour des raisons stériques** (c'est-à-dire pour des raisons de place, d'encombrement). « deux purines (= 4 cycle au total) prendraient trop de place, deux pyrimidines (= 2 cycle en tout) seraient trop éloignées pour former des liaisons stables. Chaque paire de base a donc la même dimension et ceci rend possible la structure régulière de la double hélice. Ceci pourrait laisser supposer qu'en face de A (base purique), on peut trouver aussi bien T que C (base pyrimidique).

- **Pour une raison de liaison hydrogène** : les bases complémentaires situées face à face sont liées entre elles par des liaisons hydrogène. Cette dernière est une liaison de faible énergie entre deux atomes attirés l'un vers l'autre pour des raisons électrostatiques l'un étant riche en électrons donc nucléophile (azote ou oxygène) et l'autre n'ayant que des protons de son noyau donc électrophile (hydrogène).

Lorsque un acide est en solution les molécules forment des liaisons hydrogènes associant les nucléotides deux par deux, de sorte qu'un nucléotide à adénine se lie avec un nucléotide à thymine (ou à uracile dans un RNA) et un nucléotide à guanine avec un nucléotide à cytosine. On désigne cette liaison sous le terme d'hybridation.

L'hybridation adénine-thymine est moins stable (2liaison hydrogène) que celle entre guanine et cytosine (3liaison hydrogène).

II.2.3. Hélicoïdales : les deux chaînes de DNA présentent dans l'espace une configuration hélicoïdale. Elles s'enroulent autour d'un axe central imaginaire en formant une double hélice droite (Nous parlerons plus loin d'une autre forme de DNA (Z DNA) qui s'enroule à gauche) dont les caractéristiques sont les suivantes :

- 2nm de diamètre ;
- 3.4 nm de pas (le pas de l'hélice est de 3.4nm et il correspond à 10 paires de base, cela implique que la distance entre deux bases successives est de 0.34nm);
- Contenant environ 10 paires de base (pb) par pas (10.4exactement) :
- Les deux squelettes pentoses-phosphates sont à l'extérieur de la double hélice, les bases sont à l'intérieur.

II.3. Les différentes formes d'ADN

Plusieurs conformations correspondant à des sens d'enroulement différents ou des pas différents ont été trouvées, les principales étant :

II.3.1. Conformation B

C'est celle du modèle décrit par Watson et Crick et que l'on trouve comme la forme principale native dans les conditions physiologiques suivantes:

- enroulement droit
- pas : 3,4 nm
- 10 pb par tour

II.3.2. Conformation A

Lorsque la teneur en eau d'une solution contenant une molécule d'ADN est diminuée, la molécule change de conformation et adopte une conformation notée A. Ce changement est réversible. La conformation A a les spécificités suivantes :

- hélice droite ;
- hélice plus compacte ;
- pas de 2,8 nm (11 pb par tour) ;
- distance entre les plans de bases successives de 0,26 nm

Cette conformation est trouvée *in vivo* dans l'ADN de certaines spores bactériennes, formées en réponse à la dessiccation du milieu, ainsi que dans les hybrides ADN-ARN qui se forment transitoirement à l'amorce de la réplication, et pendant la transcription.

II.3.3. Conformation Z

Cette conformation est présente dans des régions courtes de l'ADN dans une conformation générale de type B (native). Ces régions spécifiques sont en général des segments de séquence alternée Pur/Pyr (G-C-G-C). La conformation Z a les spécificités suivantes :

- hélice gauche
- pas de 4,5 nm (12 pb par tour) : la molécule est plus étirée dans cette conformation
- distance entre les plans de bases successives de 0,37 nm

Cette conformation est trouvée *in vivo* pour des segments de la molécule d'ADN, avec des enchaînements alternés Pur/Pyr.

II.4. Propriétés physico-chimiques de l'ADN

L'ADN possède plusieurs propriétés physico-chimiques qui sont :

- La structure de la double hélice donne une nature fibreuse à la molécule d'ADN dont les propriétés sont exploitées dans de nombreuses expériences de biologie moléculaire :
- Les alcools, et en particulier l'éthanol, précipitant les molécules d'ADN sous forme d'agglomérat en longues fibres.
- La charge des molécules à pH physiologique est négative et directement proportionnelle à la longueur (nombre de nucléotides). Cette propriété est utilisée pour les séparer par électrophorèse.
- Quelle que soit l'origine de l'ADN, le nombre de purines est toujours égal au nombre de pyrimidines : $[Pur] = [Pyr]$ ou encore $[A] + [G] = [T] + [C]$, cette caractéristique est désignée sous le nom de règle de Chargaff 1940.
- Les proportions ($[A] + [T]$) et ($[G] + [C]$) ne sont pas égales et varient de 35 à 75% selon l'ADN.
- L'ADN présente une absorbance dans l'ultraviolet avec un maximum à 260nm. Quand il est dénaturé a une absorption à 260 nm plus élevée que l'ADN natif, d'un facteur 1,6. Cette propriété est appelée *l'effet hyperchrome* ou *hyperchromicité* (la fusion est accompagnée d'une augmentation de l'absorption de 40%) (figure).

L'absorption de l'ADN natif, à 260 nm, en fonction de la température (courbe de fusion), présente l'allure d'une sigmoïde : le point d'inflexion de cette courbe, qui correspond à la demi-variation d'absorbance, est la température de fusion de la molécule, notée T_m (melting temperature).

Lors d'un refroidissement lent, l'absorption suit la courbe de fusion en sens inverse. Lors d'un refroidissement rapide, l'absorption ne suit pas la courbe de fusion en sens inverse mais une autre courbe qui n'aboutit à la même valeur originale.

- La température de fusion T_m de l'ADN est dépendante de la force ionique du milieu et qu'elle diminue lorsque cette dernière augmente (dans des milieux où $[NaCl] > 1M$). La valeur de la température de fusion T_m est d'autant plus élevée que le pourcentage de bases G + C est grand

III. Réplication de l'ADN

Les informations contenues dans l'ADN nécessaires pour la synthèse des différentes protéines vont être transmises de génération en génération par le processus de la réplication. Lorsqu'une cellule se divise pour donner deux cellules filles, il faut que l'ADN de ces cellules filles soit l'exacte réplique de l'ADN de la cellule mère (les cellules filles doivent posséder le même patrimoine génétique que la cellule mère).

III.1. Mode de réplication de l'ADN

Trois modes de réplication peuvent être théoriquement envisagés :

- **Réplication conservative** : l'ADN parental est entièrement conservé ; les deux brins de la molécule fille sont néoformés.
- **Réplication semi - conservative** : Un brin de la molécule mère est conservé intégralement et sert de modèle pour la reconstitution du second brin : chaque molécule fille contient l'un des brins de la molécule d'origine.
- **Réplication dispersive** : les fragments de l'ADN se répartissent entre les brins des deux molécules filles.

III.2. Eléments nécessaires à la réplication

La polymérisation des désoxyribonucléiques au cours de la réplication de l'ADN est réalisée par des **ADN polymérases** spécifiques qui catalysent le transfert d'une unité 5' désoxyribonucléoside triphosphate (dNTP) sur le groupements hydroxyle en 3' d'une chaîne d'ADN en cours de synthèse. Ces enzymes exigent :

- Une matrice d'ADN constitué par un brin parental
- La présence de nucléotides propres à l'ADN sous forme de désoxynucléotides 5'triphosphate :dATP, dTTP, dGTP et Datp.
- La présence de nombreux enzymes et protéines, pour séparer les deux brins d'ADN, les maintenir séparés, résoudre les problèmes topologiques (ADN polymérase, hélicase, primase,...)
- Une amorce associée au brin parental
- La présence des ions Mg^{++}

III.3. Caractéristiques de la réplication

III.3.1. Synthèse des nouveaux nucléotides

Le nouveau brin est toujours synthétisé :

- dans le sens 5'---3', chaque nouveau nucléotide étant ajouté à l'extrémité 3' du brin en cours de synthèse.
- de façon complémentaire, selon les règles d'appariement des bases (A/T et G/C).
- et de façon antiparallèle : le brin matrice est lu dans le sens 3'.....5'.

III.3.2. La propagation de la réplication est bidirectionnelle

Cela veut dire que la réplication se propage simultanément à droite et à gauche du point d'initiation (on peut comparer la réplication à un œil qui s'agrandit jusqu'à ce que la réplication soit terminée).

III.3.3. La réplication est discontinue pour l'un des deux brins

Au niveau de chaque fourche de réplication, un brin matrice est lu (dans le sens 3'----5') et le brin fils synthétisé (dans le sens 5'---3') dans le même sens que celui de propagation de la fourche, tandis que l'autre brin est lu et le brin fils est synthétisé dans le sens opposé à celui de propagation de la fourche.

La réplication est continue pour un brin (le brin dit « avancé ou précoce») et discontinue pour l'autre brin (le brin dit « retardé ou tardif »).

III.4. Etapes de la réplication

III.4.1. Intervention d'une ARN polymérase (primase)

L'addition des nucléotides les uns aux autres, pour former un brin d'ADN, s'effectue grâce à une enzyme (ADN polymérase), cette enzyme n'a aucun « esprit initiative », elle ne sait pas commencer une chaîne. Elle ne sait qu'allonger une chaîne de nucléotides (c'est-à-dire qu'elle ne sait qu'ajouter un nucléotide à l'extrémité 3'OH d'un acide nucléique). C'est là que va intervenir l'ARN polymérase, cette enzyme est capable de commencer une chaîne d'acide nucléique. La synthèse d'un nouveau brin d'ADN commence par un petit fragment d'ARN.

III.4.2. Hydrolyse et remplacement des amorces d'ARN

Les amorces d'ARN seront ensuite détruites, hydrolysées et, remplacées par de l'ADN. C'est la même enzyme qui enlève les amorces d'ARN et resynthétise à la place de l'ADN. Il

s'agit de l'ADN polymérase (I) douée à la fois de propriétés exonucléasiques 5'---3' (qui servent à hydrolyser les amorces de l'ARN) et de propriétés polymérasiques (permettant d'ajouter les désoxyribonucléotides en 3' du fragment de l'ADN précédent).

Les fragments de l'ADN, libérés de leur amorce ARN seront soudés les uns aux autres par une ligase.

Chapitre II

Expression de l'information génétique

Chapitre II : Expression de l'information génétique

I. Transcription

La transcription est la synthèse d'ARN à partir d'une matrice d'ADN par l'ARN polymérase. Un seul brin de l'ADN est transcrit au cours d'une transcription mais ce n'est pas toujours le même brin qui est copié tout au long de la molécule d'ADN. La séquence d'ARN est complémentaire de celle du brin matrice et est identique à celle du brin codant sauf que l'ARN présente des bases U au lieu de T dans l'ADN.

La synthèse d'un ARNm s'effectue dans le sens 5'→3' de façon antiparallèle et complémentaire.

I.1. Eléments nécessaires pour la transcription

Pour synthétiser *in vivo* un ARNm, il faut :

- Des nucléotides sous forme de nucléosides triphosphates (ATP, GTP, CTP, UTP).
- L'ARN polymérase qui est l'enzyme responsable de souder les nucléotides les uns aux autres pour former ce polymère qui est l'ARNm.
- Un modèle d'ADN représentant la matrice.

I.2. Transcription chez les procaryotes

Les ARN polymérases les mieux connues sont celles des Eubactéries notamment celle d'*Escherichia coli*. Un seul type d'ARN polymérase est à l'origine de la synthèse de presque tous les ARNm et de tous les ARNt et ARNr. Cette enzyme présente une masse moléculaire de près de 465 kDa chez *E. coli*. Elle possède cinq sous-unités : $\alpha_2\beta\beta'\sigma$ qui forment l'holoenzyme.

□ La sous unité σ , encore appelée facteur sigma, permet à l'ARN polymérase de reconnaître les sites promoteurs (initiation de la transcription puis se dissocie après cette étape) et assure une fixation stable de l'ARN polymérase au niveau de ce site.

□ Le noyau $\alpha_2\beta\beta'$ contient le site catalytique directement responsable de l'élongation : la sous unité β intervient dans la fixation des nucléosides triphosphates et la sous unité β' intervient dans la fixation du template (ADN matrice).

I.2.1. Etapes de la transcription

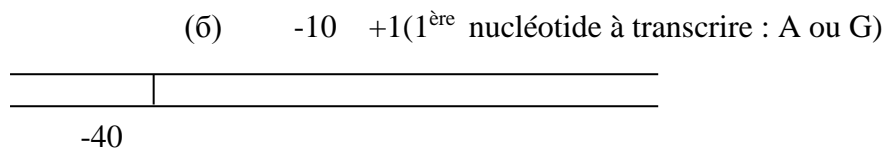
a) *Initiation*

La sous unité σ se lie au promoteur, région de l'ADN comprenant environ 40 paires de nucléotides située juste avant le début de la région où démarrera la transcription.

Chez *E. coli* deux éléments de séquences reconnus par l'ARN polymérase sont présents au sein des promoteurs : la séquence -10 dont la séquence consensus est TATAAT et la séquence -35 dont la séquence consensus est TTTGACA.

Séquence -10 (Pribnow box)

C'est une séquence conservée se trouvant dans le promoteur, elle est située à environ 10 paires de nucléotides en amont du point de départ de la transcription, elle est riche en liaisons faibles A et T. C'est une région d'avertissement pour l'ARN polymérase.



L'holoenzyme se fixe sur le promoteur au niveau de la région -35. L'enzyme recouvre près de 80 pb d'ADN. Ensuite, elle se lie à une courte région de l'ADN conduisant à son ouverture dans une région riche en A-T au niveau de la région -10.

b) *Elongation*

Le noyau de l'enzyme se détache du promoteur et se déplace le long de la matrice d'ADN. Les nucléosides triphosphates apportent à la fois le nucléoside monophosphate et l'énergie nécessaire pour relier chaque nucléotide au précédent. Seul le premier nucléotide de l'ARNm conserve son groupement triphosphate.

Au fur et à mesure que l'œil de transcription parcourt la molécule d'ADN, l'ARN synthétisé se dissocie de l'ADN matrice et l'hélice se referme (Figure 4).

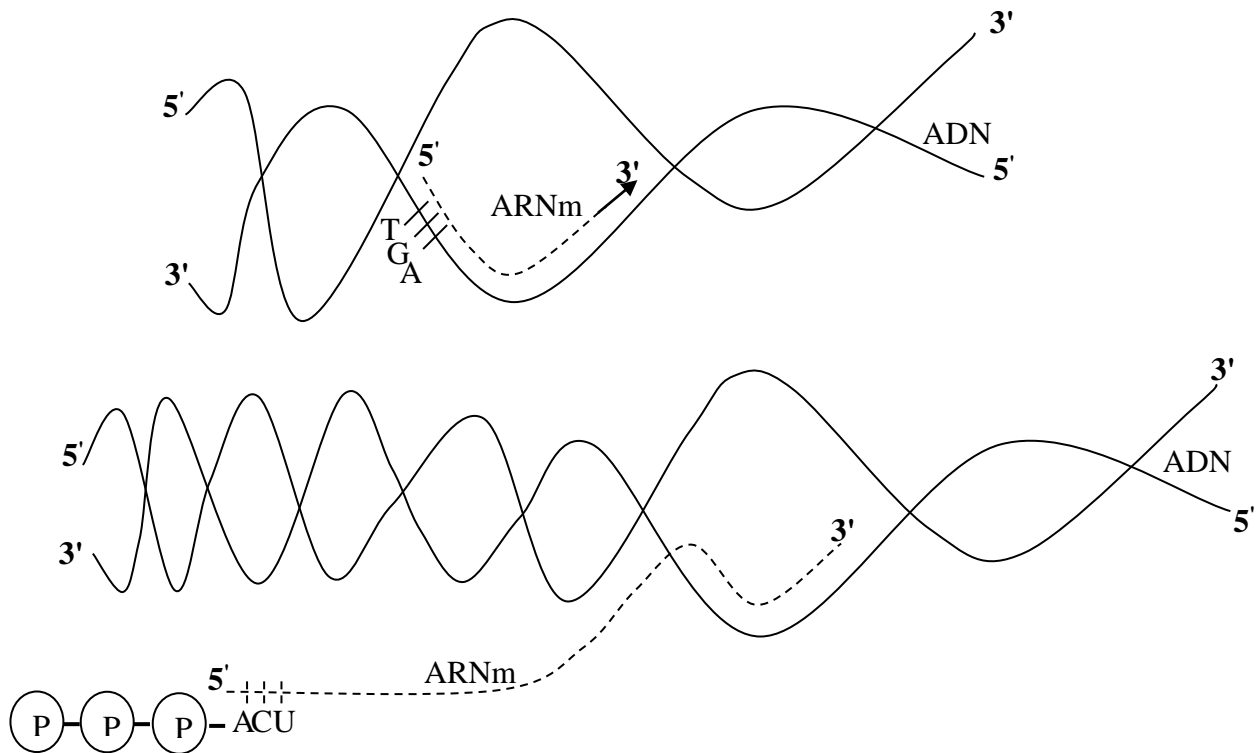


Figure 4 : Etapes de l'élongation dans la transcription chez les Procaryotes.

c) *Terminaison*

La transcription se termine par la reconnaissance de sites spécifiques ou terminateurs. Dans la plupart des cas, la terminaison dépend de la formation dans la structure secondaire de l'ARN d'une structure dite en épingle à cheveux. Cette dernière provoque un ralentissement voir un arrêt de l'ARN polymérase.

La structure en épingle à cheveux est suivie d'une série de nucléotides à U (séquence poly-U). Cette séquence est indispensable au détachement de l'ARN polymérase du brin matrice lors de son arrêt sur l'épingle à cheveux.

Un autre facteur intervient dans la terminaison de la transcription, c'est la présence d'un facteur protéique rho (ζ) qui est un hexamère de sous unités identiques de 419 acides aminés. La protéine rho détruit les appariements entre les bases de l'ARN et de l'ADN matrice lorsque l'ARN polymérase s'arrête après le motif en épingle à cheveux.

I.3. Transcription chez les eucaryotes

Il existe trois ARN polymérases :

- ⊠ ARN polymérase I ----synthétise----> ARN ribosomiaux (18s, 5.8s et 28s).
- ⊠ ARN polymérase II -----> Précurseurs des ARN messagers.
- ⊠ ARN polymérase III -----> ARN 5s, ARNt et ARNsn.

I.3. 1. Etapes de la transcription

a) *Initiation*

Les promoteurs chez les eucaryotes présentent des éléments conservés :

- ⊠ Une séquence nommée boîte TATA (TATA box) située à -25 avec une séquence consensus TATA (A/T)A(A/T).
- ⊠ Une séquence nommée CAAT box située entre la position -75 et -80, cette séquence augmente l'efficacité du promoteur.
- ⊠ Une séquence GC box ou GGGCGG située à la position -90 qui agit également sur la régulation de la transcription.

L'initiation de la transcription chez les eucaryotes nécessite la formation d'un complexe d'initiation. L'ARN polymérase II nécessite la présence de plusieurs facteurs de transcription pour permettre sa liaison à l'ADN (exemple de facteur : TFIID). Les facteurs de transcription se fixent dans un ordre précis : TFIID puis TFIIA et TFIIB. L'ARN polymérase II se fixe ensuite, suivie par TFIIF, E, H et J pour produire finalement un complexe fonctionnel capable d'initier la transcription.

b) *Elongation*

ARN (n nucléotides) + XTP -----> ARN (n+1) nucléotides +PP

XTP = CTP, ATP, UTP, GTP

c) *Terminaison*

Les séquences signalant l'arrêt de la transcription ne sont pas identifiées, mais certaines ressemblent à ceux des procaryotes car ils comportent une structure en épingle à cheveux suivie d'une série de résidus uridyliques (u,u,u,...)

I.3.2. Maturation de l'ARNm

La maturation des ARN est l'étape qui permet d'obtenir après la transcription un ARN opérationnel. A partir d'un transcrit primaire on obtient un ARN dit mature selon les étapes suivantes :

- *Mise en place de la coiffe (Capping)*

Les ARNm des eucaryotes subissent au niveau de leur extrémité 5' une modification appelée **capping** qui correspond à l'addition d'un nucléotide modifié, la 7-méthylguanosine = cap = chapeau = coiffe. Ce chapeau est ajouté par l'enzyme guanyltransférase qui lie le GTP au premier nucléotide de l'ARNm. Cette réaction se déroule avant que l'ARNm possède 20 nucléotides.

Des enzymes nommées méthyl-transférases additionnent ensuite un groupement CH₃ à l'azote 7 du cycle de la guanine et généralement, au groupement 2'-hydroxyl des riboses des deux nucléotides suivants.

Le capping évite que l'ARNm soit dégradé depuis son extrémité 5' par certaines enzymes (phosphatase, nucléases). C'est aussi un signal autorisant le ribosome à reconnaître le début de la molécule d'ARNm (Voire figure 3).

- *Polyadénylation de l'extrémité 3'*

La majorité des pré-ARNm des eucaryotes sont modifiés en leurs extrémités 3' par l'addition d'une séquence pouvant compter jusqu'à 250 adénines consécutives appelée queue poly A. L'enzyme impliquée est l'enzyme poly A polymérase.

La queue poly A protège l'ARNm de la dégradation par les exonucléases, mais aussi l'aide à passer du noyau vers le cytoplasme.

NB : Les ARNm codant pour les histones ne possèdent pas la queue poly A.

-*Excision des introns et épissage des exons*

Les éléments qui jouent un rôle important lors de l'excision-épissage du transcrit primaire sont (Figure 5):

-La jonction exon-intron : Tous les introns d'un transcrit primaire commencent par GU « ce site 5' est appelé site donneur d'épissage » et se terminent par AG « site 3' appelé site accepteur d'épissage ».

- **Le site de branchement du lasso** : site à environ -30 nucléotides de l'extrémité 3', se trouve un AMP appelé « A du branchement ».

Le mécanisme de l'excision-épissage peut être décomposé en deux étapes :

1^{ère} étape : formation d'un premier clivage en 5' de l'intron, entre la fin de l'exon amont et le début de l'intron. L'extrémité 5' de l'intron GMP vient alors se souder par une liaison covalente à l'A du branchement, formant ainsi un lasso.

2^{ème} étape : Formation du deuxième clivage, en 3' de l'intron, entre la fin de l'intron et le début de l'exon suivant situé en aval.

Il se produit simultanément : - soudure des deux exons

- libération du lasso qui sera dégradé par des nucléases cellulaires.

NB : L'épissage (splicing en anglais) est assuré par un complexe nommé particule d'épissage (ou spliceosome en anglais), contenant des petits ARN nucléaires (ARNsn) associés à des protéines sous la forme de petites ribonucléoprotéines (ou snRNP pour small nuclear ribonucléoproteins) que l'on nomme « snurp ».

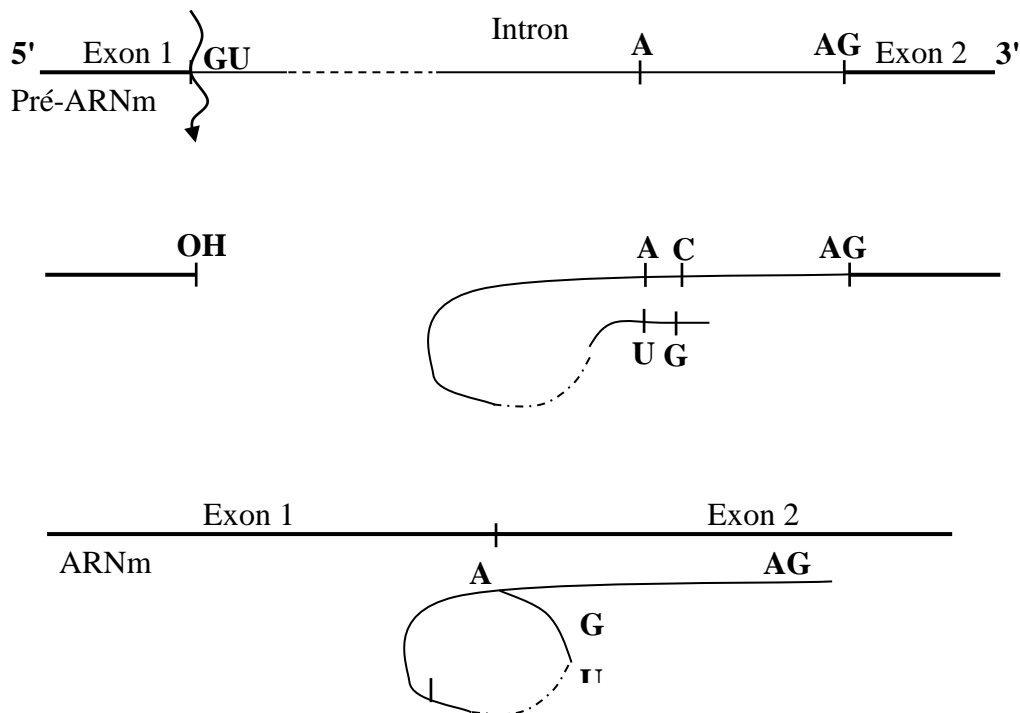


Figure 5 : Excision et épissage du transcrit primaire

II. Traduction

La traduction est le processus par lequel une cellule assemble une chaîne protéique en combinant des acides aminés isolés présents dans son cytoplasme, guidé par l'information contenue dans l'ARN_m. Elle représente la deuxième étape de la synthèse protéique après la transcription.

II.1. Siège de la traduction

Chez les eucaryotes, la traduction se déroule dans le réticulum endoplasmique au niveau des ribosomes, alors que la transcription se fait dans le noyau. En revanche, chez les procaryotes, les deux étapes ont lieu dans le cytoplasme et peuvent être simultanées, la traduction débutant alors que la transcription n'est pas encore achevée.

II.2. Acteurs de la traduction

La traduction fait intervenir :

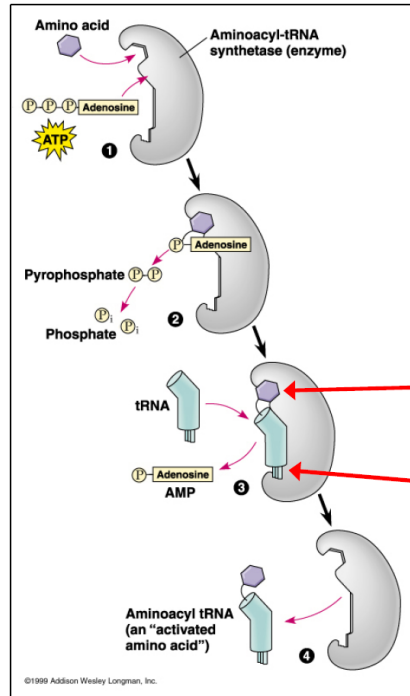
II.2.1. Les ARN messenger (ARN_m)

L'ARN messenger porte le message de l'ADN. Il indique par un codage nucléotidique l'ordre (ou séquence) des acides aminés dans la chaîne polypeptidique.

II.2.2. Les ARN de transfert (ARN_t)

Les ARN_t sont les molécules adaptatrices qui alignent chaque acide aminé avec son codon correspondant sur l'ARN messenger. A leurs extrémités 3', tous les ARN_t portent la séquence 5' CCA 3'. C'est à cette extrémité de l'ARN_t qu'est attaché l'acide aminé par une aminoacyl-tARN synthétase. Chaque aminoacyl-ARN_t synthétase reconnaît un unique acide aminé et l'ARN_t correspondant. Chaque ARN_t possède aussi une boucle contenant l'anticodon qui va s'apparier avec le codon correspondant sur l'ARN messenger (figure 6 et 7).

L'acide aminé est attaché au bon ARNt par l'enzyme *aminoacyl-ARNt synthétase*. Il existe plusieurs sortes d'*aminoacyl-ARNt synthétase*. Chacune peut attacher un acide aminé particulier à un ARNt particulier.



Le site actif de l'enzyme reconnaît:

un acide aminé particulier
et
un anticodon particulier.

L'enzyme unit l'acide aminé à l'ARNt

Figure 6 : Formation d'un ARNt

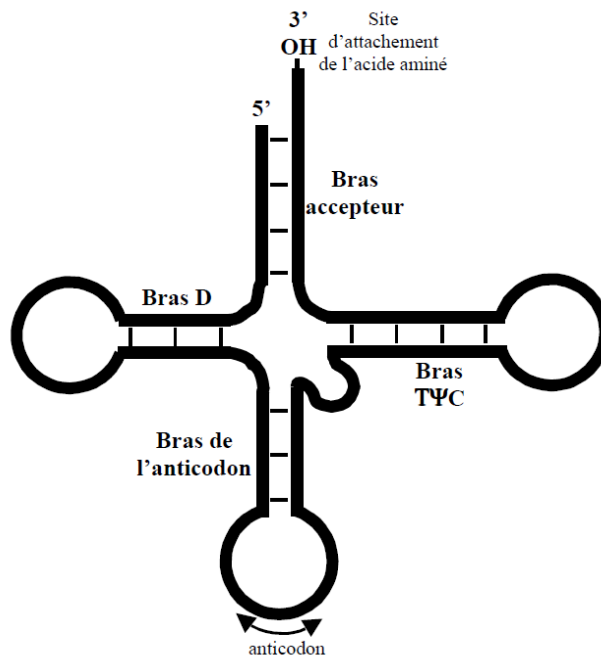


Figure 7 : Structure d'un ARNt

.2.3. Les ribosomes

Les ribosomes sont des complexes de grande taille composés d'ARN et de protéines et organisés en deux sous unités. La petite sous unité se lie à l'ARNm et à l'aminocyl-ARNt, permettant la reconnaissance des codons. La grande sous-unité catalyse la formation de la liaison peptidique. Il y a un site de liaison à l'ARNm, un site de liaison (P) du peptidyl-ARNt (c'est-à-dire de l'ARNt portant la chaîne peptidique en croissance) et un site de liaison à l'aminocyl ARNt (A). La figure 8 montre les différences existantes entre le ribosome procaryotique et eucaryotiques.

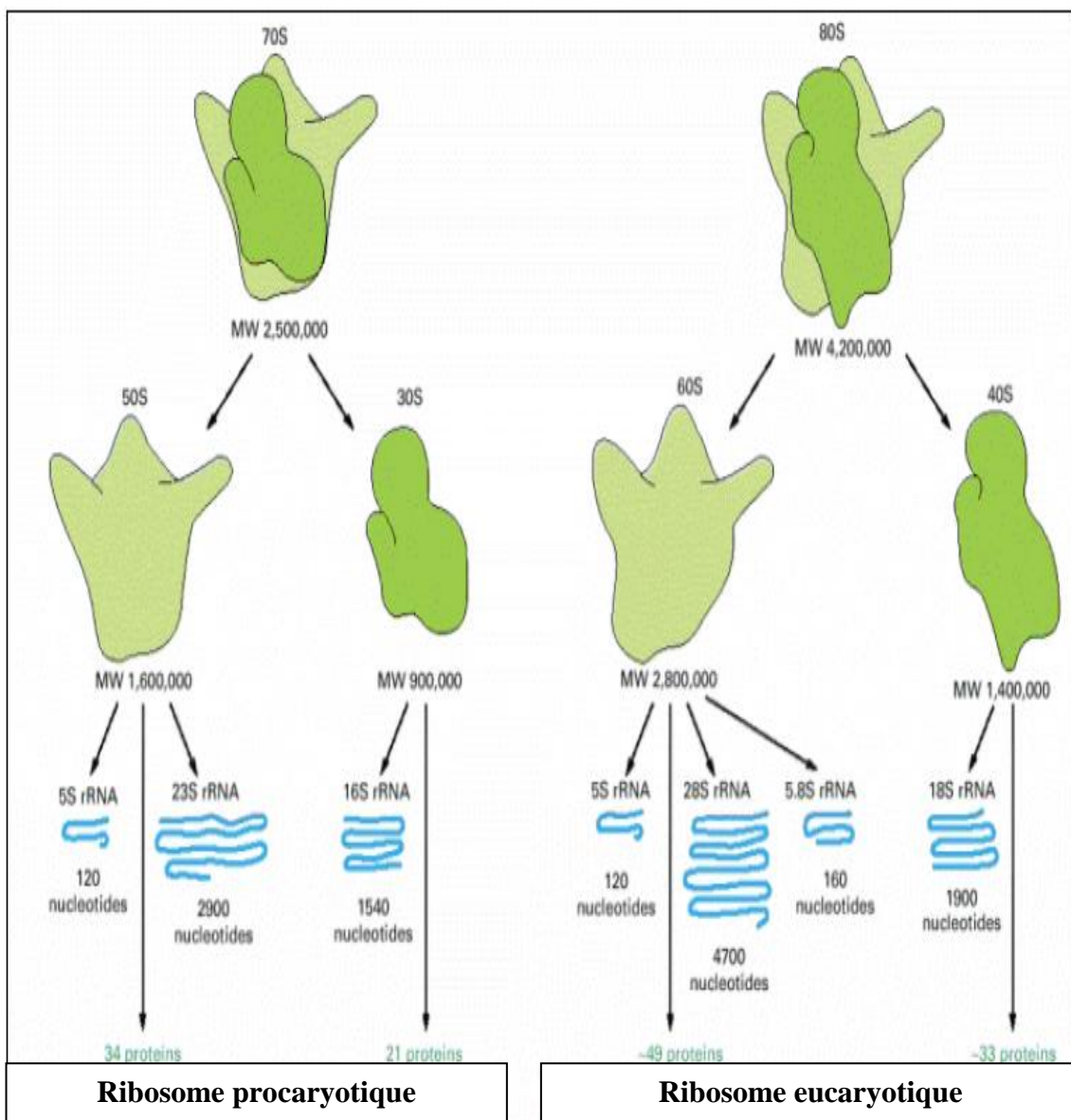


Figure 8 : Différence entre le ribosome procaryotique et eucaryotique

II.3. Le code génétique

II.3.1. Définition

Le code génétique (tableau I) représente le système de correspondance entre la séquence des nucléotides dans les acides nucléiques et celle des acides aminés dans les protéines.

La correspondance n'est pas un nucléotide pour un acide aminé puisqu'il n'existe que 4 nucléotides différents pour 20 acides aminés différents. Le code génétique repose sur une combinaison de nucléotides. Une combinaison de 2 parmi 4 possibles ne peut suffire, une combinaison de 3 (64 possibilités) est vraisemblable et le système de triplets (3 nucléotides pour spécifier un acide aminé) est effectivement celui qui est universellement utilisé.

Tableau I: Le code génétique

1 ^{ère} lettre (extrémité 5')	2 ^{ème} lettre				3 ^{ème} lettre (extrémité 3')
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

II.3.1. Les différents codons

- **Les codons d'initiation**

En général le codon (AUG) désignant l'acide aminé Méthionine est le signal d'initiation de la synthèse protéique chez les eucaryotes. Les bactéries peuvent parfois utiliser les codons GUG (val) ou UUG (leu) comme signal d'initiation

- **Les codons stop, ou codons de terminaison ou encore codons non-sens**

Les codons UAA, UAG et UGA ne correspondent à aucun acide aminé ce qui provoque l'arrêt de la synthèse protéique.

- **Les codons d'élongation**

Ce sont tous les autres codons qui désignent un acide aminé qui sera incorporé lors de la synthèse protéique.

II.3.2. Caractéristique du code génétique

On peut résumer les caractéristiques du code génétique en trois points :

- **Le code est universel**

Que ce soit chez une plante, un animal ou encore un virus, le code génétique est le même. Cependant il existe quelques exceptions ; ainsi pour le procaryote *Mycoplasma capricolum*, UGA ne code pas pour un codon stop mais pour le tryptophane. Chez la levure du genre *Candida*, CUG code pour la serine au lieu de l'isoleucine et le codon UAG est utilisé comme codon stop. En ce qui concerne le génome mitochondrial, celui-ci comporte un certain nombre de modifications. Le changement commun à toutes les mitochondries (sauf celles des plantes) concerne le codon UGA qui code pour le tryptophane au lieu d'un codon stop.

- **Le code est dégénéré ou redondant**

Il existe 20 acides aminés différents et il y a 61 codons possibles. Il y a donc plus de codons que d'acides aminés. De ce fait, plusieurs codons désignent le même acide aminé. On parle alors de dégénérescence du code et de codons synonymes. Fréquemment la troisième position du codon n'intervient pas dans la désignation de l'acide aminé. Cette position est dite **flottante** lorsque le codon désigne un même acide aminé quelque soit la nature du nucléotide

à cette troisième position. Cependant une exception apparaît avec la méthionine et le tryptophane qui ne sont désignés que par un seul codon AUG et UGG respectivement.

- **Le code est non ambigu**

Alors qu'un même AA peut être codé par plusieurs codons différents (jusqu'à six dans le cas extrême), un même codon ne peut coder que pour un seul AA c'est pour cette raison que l'on dit qu'il y a absence d'ambiguïté.

II.3. Les étapes de la traduction

II.3.1. Initiation

L'initiation de la traduction (figure 9) débute par la fixation de la petite sous unité ribosomique sur une séquence d'ARN_m dite séquence de liaison des ribosomes. Cette séquence est située avant le codon d'initiation de l'ARN_m. Chez les procaryotes la petite sous unité reconnaît une séquence particulière dite de *Shine Dalgarno* de séquence AGGAGGU, située une dizaine de nucléotides avant le codon d'initiation.

Chez les eucaryotes, la petite sous unité reconnaît la coiffe en 5' de l'ARN_m. La grande sous unité rejoint ensuite le complexe d'initiation formé de l'ARN_m et de la petite sous-unité ribosomique. Le début de la synthèse protéique commence au signal d'initiation qui est le codon Méthionine (AUG). Parfois les bactéries utilisent les codons GUG ou UUG. La méthionine débutant la protéine doit être formyle (COH)

Lorsque le ribosome complet est formé, la synthèse protéique peut alors commencer. Dans le complexe d'initiation, l'ARN_t initiateur portant la Méthionine se place dans le site P porté par la petite sous unité.

Des facteurs d'initiation, différents chez les procaryotes et les eucaryotes, interviennent lors de cette étape (figure).

a) *Initiation chez les procaryotes*

Chez les bactéries, les facteurs d'initiation sont indispensables pour assurer la liaison entre l'ARN_m et l'ARN_t, la sous unité 30S étant incapable d'assurer à elle seule cette fonction. Ces facteurs se fixent à l'unité 30S lorsque celle-ci s'associe à l'unité 50S pour former un ribosome complet. Trois facteurs d'initiation notée IF sont nécessaires à l'initiation :

- Le facteur IF3 est impliqué dans la reconnaissance du site de fixation de l'ARN_m par la sous unité 30S. Ce facteur se lie aux sous unités 30S libres et empêche leur

association avec l'unité 50S. Il doit être libéré du complexe d'initiation pour que les deux unités du ribosome s'associent.

- Le facteur IF1 se fixe au complexe d'initiation et le stabilise.
- Le facteur IF₂ lie GTP et l'ARN_t initiateur. Ce facteur contrôle l'entrée de l'ARN_t initiateur dans le ribosome en le dirigeant vers le site P.
- Une protéine ribosomique clive le GTP lié au facteur IF2 et favorise l'assemblage des deux sous unités. Les différents facteurs sont alors libérés.

b) Initiation chez les eucaryotes

Chez les eucaryotes, le processus est similaire à celui des procaryotes mais nécessite d'avantage de facteurs d'initiation. Par exemple, dans les réticulocytes (globules rouges immatures). On trouve jusqu'à 9 facteurs d'initiation. On les note eIF (e pour eucaryotes).

L'initiation de la traduction, chez les eucaryotes, peut se résumer comme suit :

- Le facteur eIF2 est impliqué au début de l'initiation dans la formation d'un complexe ternaire contenant également l'ARN_t initiateur et du GTP ; le complexe ternaire s'associe alors à la petite sous unité 40S ;
- Le facteur eIF3 agit comme le facteur IF3 des procaryotes en se liant à la petite sous unité 40S et intervient dans son maintien sous sa forme libre
- Le facteur eIF4 reconnaît la coiffe et est responsable de l'élimination des structures secondaire qui peuvent se former dans l'ARN_m, en utilisant l'ATP ;
- Le facteur eIF6 se fixe à la grande sous unité 60S et prévient son association avec la petite sous unité ;

L'association des deux sous unités est assurée par le facteur eIF5 qui est une GTPase et nécessite la libération des facteurs eIF6, eIF2 et eIF3. Le ribosome complet 80S est alors formé.

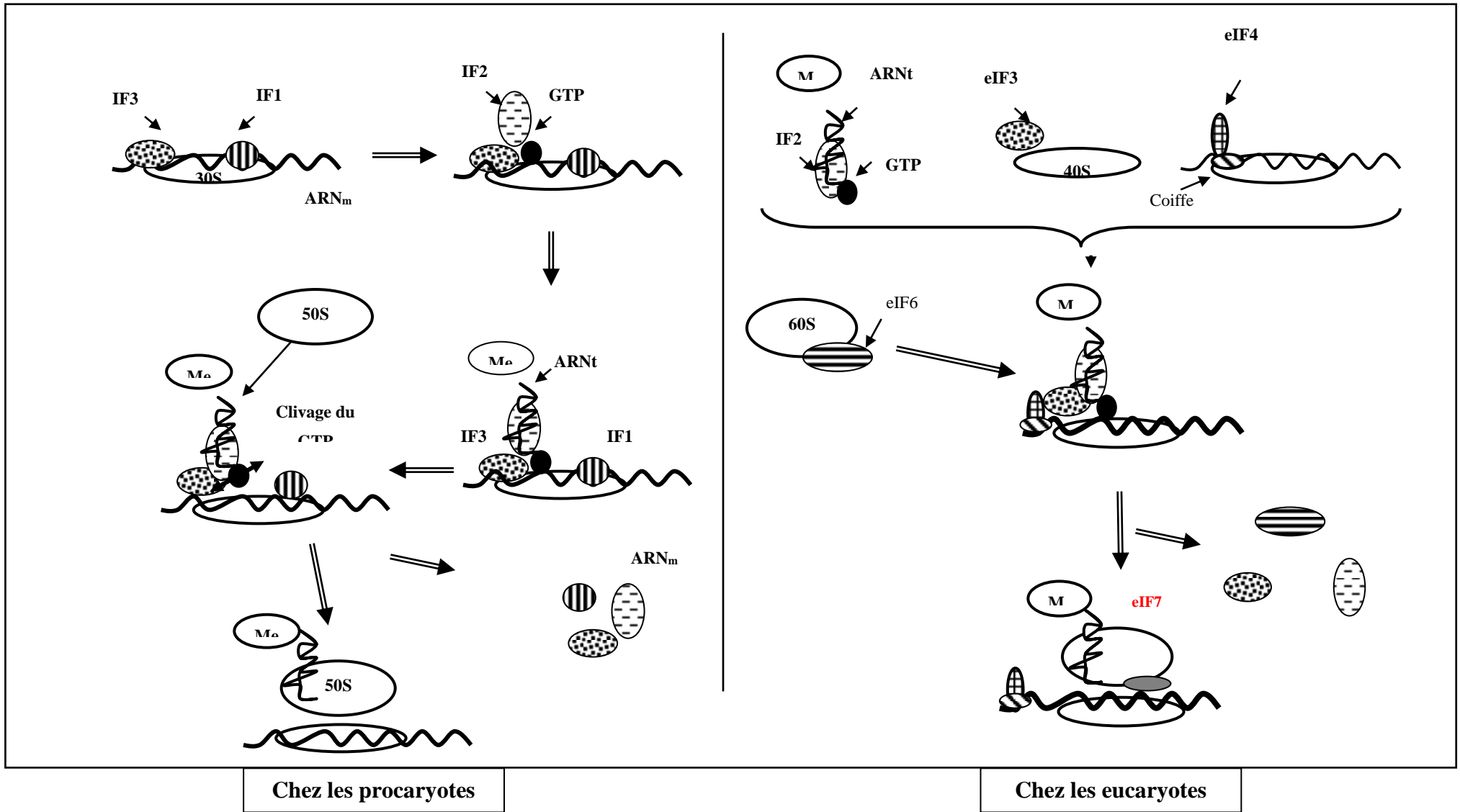


Figure 9: Initiation chez les procaryotes et les eucaryotes

II.3.2. Elongation

Lorsque la grande sous unité rejoint la petite sous unité, le site A du ribosome est alors disponible pour l'entrée du deuxième aminoacyl-ARN_t dont l'anticodon est complémentaire du deuxième codon de l'ARN_m. La première liaison peptidique se forme alors entre l'Art initiateur et l'aminoacyl-ARN_t.

La réaction est catalysée par un complexe enzymatique, nommé peptidyl transferase de la grande sous unité ribosomique. Après la formation de la liaison peptidique, les liaisons entre la méthionine et son ARN_t sont rompues par une enzyme l'ARN_t déacylase. Le ribosome effectue ensuite une translocation, il se déplace jusqu'au codon suivant. L'ARN_t qui n'est plus chargé est expulsé un site P. le peptide porté par l'ARN_t du second acide aminé s'y engage. Le site A est alors libéré et un troisième aminoacyl-ARN_t ayant un anticodon de séquence complémentaire du troisième codon vient s'y placer (figure 10).

Chez les bactéries, il existe un site E où l'ARN_t déchargé transite avant de quitter le ribosome.

Chez les eucaryotes, l'ARN_t est libéré directement dans le cytosol. L'élongation se poursuit selon le même cycle et la chaîne polypeptidique se constitue. Le ribosome se déplace au cours de l'élongation le long de l'ARN_m ce qui libère le site d'initiation pour une nouvelle traduction par un autre ribosome. La structure formée par un ARN_m et plusieurs ribosomes synthétisant la même chaîne polypeptidique mais de longueur différentes, se nome polysome.

L'élongation requiert différents facteurs (tableau II).

Chez les procaryotes, trois facteurs EF-Tu, EF-Ts et EF-G interviennent au cours de deux étapes :

- le facteur EF-Tu permet aux aminoacyl-ARN_t de s'engager dans le site A. Ce facteur se fixe au GTP formant un complexe binaire EF-Tu-GTP. Ce complexe se lie à l'aminoacyl-ARN_t formant un complexe ternaire qui se fixe au site A seulement si le site P est déjà occupé par un peptidyl-ARN_t. Une fois l'aminoacyl-ARN_t positionné, la reconnaissance codon-anticodon provoque un changement de la conformation de EF-Tu. L'hydrolyse du GTP libère le complexe binaire EF-Tu-GDP. Sous cette forme EF-Tu est inactif. Le facteur

EF-Ts a pour rôle de rendre de nouveau actif EF-Ts. Cette réaction fait intervenir le GTP.

- Le facteur EF-G intervient dans la translocation. ce facteur se fixe au ribosome dans un complexe contenant du GTP. Le facteur EF-G est libéré par l'hydrolyse du GTP qui fournit l'énergie pour la translocation.

Chez les eucaryotes, le facteur eEF-1 comporte EF-1 α qui joue le même rôle que le facteur EF-Tu des procaryotes et EF-1beta gamma qui a la même rôle que le facteur EF-Tu des procaryotes. De même le facteur eEF-2 joue le même rôle que le facteur EF-G.

II.3.3. Terminaison de la traduction

La terminaison de la traduction (figure 11, tableau III) a lieu lorsqu'un codon stop (UAA, UAG ou UGA) se place au niveau du site A. Aucun ARN_t ne correspond à un codon stop. Ce codon est reconnu par des facteurs de terminaison encore nommés facteurs de libération notés RF (pour release factors)

Chez les procaryotes, le facteur RF1 reconnaît les codons UAA et UAG, le facteur RF2 reconnaît les codons UAA et UGA. Ces facteurs agissent au niveau du site P. un troisième facteur EF3 fixant le GTP intervient également dans la terminaison.

L'hydrolyse du GTP conduit à la libération du polypeptide du site P par addition d'une molécule d'eau à la place d'un autre aminoacyl-ARN_t. Le ribosome libère ensuite l'ARN_m et se dissocie.

Chez les eucaryotes il n'y a qu'un seul facteur de terminaison eRF qui nécessite du GTP pour se fixer au ribosome. L'hydrolyse du GTP conduit également à la libération du polypeptide.

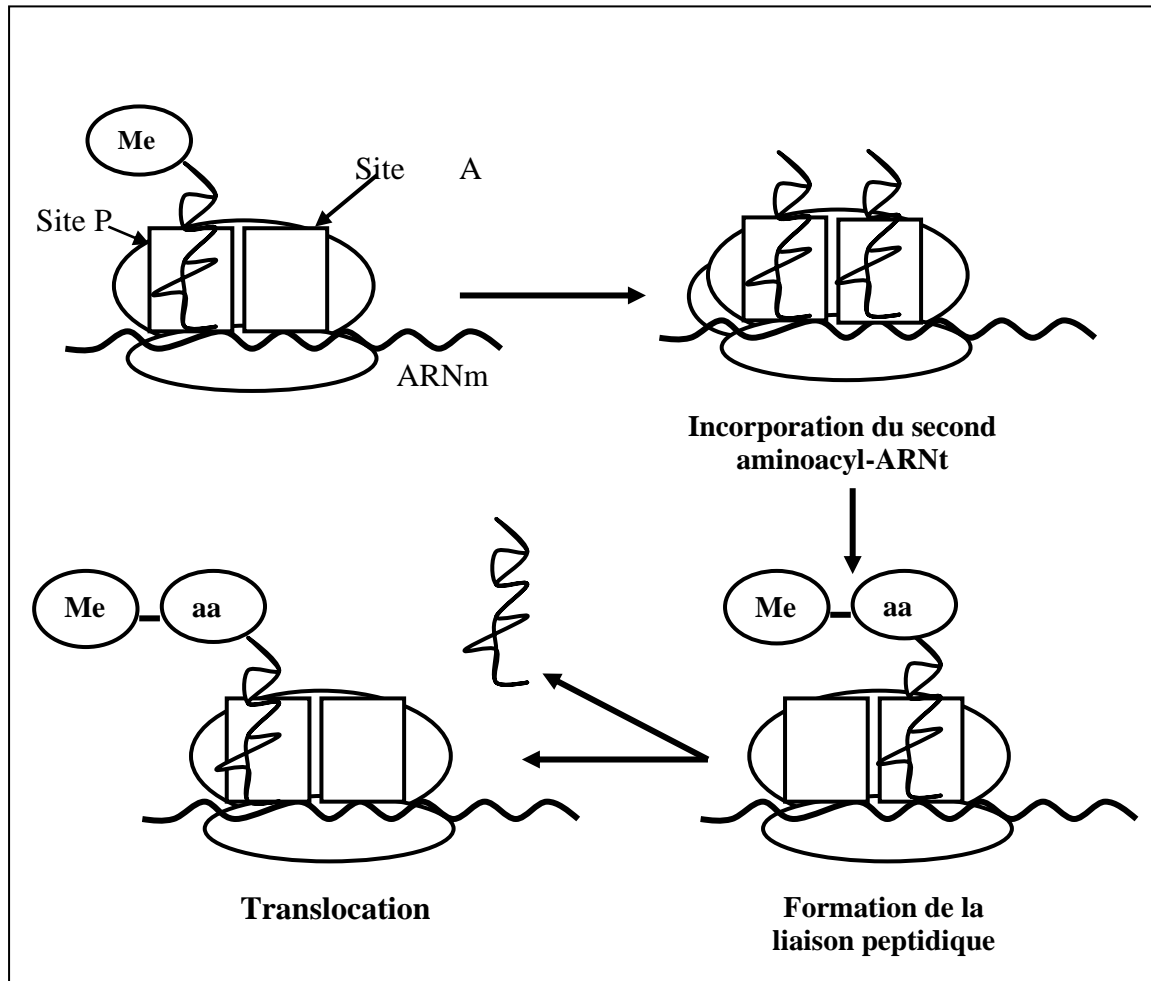


Figure 10: Etapes de l'élongation

Tableau II: Facteurs de l'élongation

Procaryotes	Eucaryotes	Rôle
EF-Tu	eEF-1 α	lient l'ARNt au site A
EF-Ts	eEF-1 $\beta\gamma$	Réactivent le facteur eEF-1 α / EF-Ts
EF-G	eEF-2	Translocation

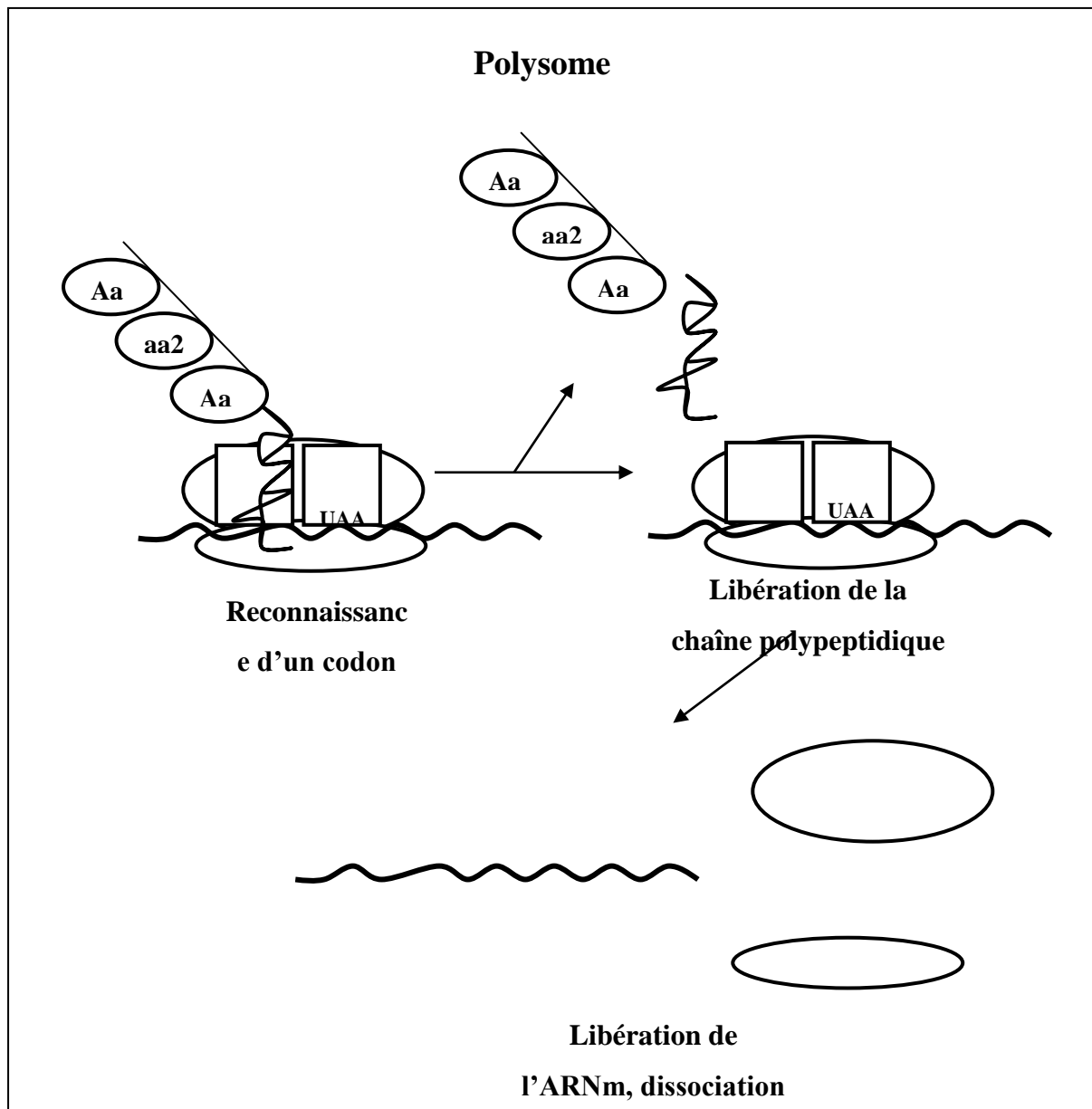


Figure 11: Etapes de la terminaison

Tableau III: Facteurs de terminaison

Procaryotes	Eucaryotes	Rôle
RF-1 (UAA,UAG)	eRF-1 (UAA,UAG, UGA)	Reconnaissance des codons de terminaison
RF-2 (UGA,UAA)		
RF-3	eRF-3	Coopèrent avec les RF-1 et RF-2

Chapitre III

Régulation de l'expression génétique

Chapitre III : Régulation de l'expression génique

I. Chez les procaryotes

Pour qu'une bactérie puisse vivre, il n'est pas nécessaire que tous ses gènes soient transcrits en même temps. Afin de conserver de l'énergie et de ressources, les bactéries régulent l'activité de leurs gènes de façon à ce que seuls les produits génétiques nécessaires à leur fonctionnement soient fabriqués. Ceci leur permet de s'adapter aux changements environnementaux. Elles peuvent faire varier la quantité des produits des gènes en faisant varier les taux de transcription et de traduction

I.1. Régulation de la transcription

I.1.1. Organisation des gènes bactériens en opérons

Une caractéristique importante, qui détermine comment est régulée la transcription des gènes chez les bactéries est leur organisation en opérons. C'est-à-dire en unités transcriptionnelles contenant plusieurs gènes qui codent généralement des protéines ayant des fonctions apparentées et qui sont régulés ensemble. On distingue deux types d'opérons, les opérons inductibles et répressibles :

a) les opérons inductibles

Ils contiennent des gènes qui codent des enzymes impliqués dans des voies métaboliques. L'expression des gènes est alors contrôlée par le substrat de la voie métabolique. Un exemple d'opéron inductible ; l'opéron lactose, qui code des enzymes nécessaires au métabolisme du lactose.

b) Les opérons répressibles

Ces opérons contiennent des gènes codant des enzymes impliquées dans des voies de biosynthèse et dont l'expression est contrôlée par le produit terminal de la voie, qui peut réprimer l'expression de l'opéron. L'exemple d'opérons répressible est l'opéron tryptophane, qui code des enzymes impliqués dans la biosynthèse du trp

I.1.1.1. L'opéron lactose (opéron lac) (figure 12)

a) Régulation par un répresseur

L'opéron lactose est un opéron inductible et la régulation est négative, il contient trois gènes Z, Y et A codant respectivement pour la B galactosidase (qui hydrolyse le lactose à

l'intérieur de la cellule), la lactose perméase qui transporte le lactose à l'intérieur de la cellule et la transacétylase qui est également impliquée dans l'hydrolyse du lactose. Dans cet exemple d'opéron, la régulation se fait par l'intermédiaire du substrat (le lactose).

En présence de lactose dans le milieu, la concentration de ces trois enzymes augmente près de mille fois dans la cellule. En l'absence de lactose l'opéron est réprimé.

En amont de l'opéron se trouve une séquence régulatrice ou gène lac I codant pour un répresseur. En absence de lactose, ce répresseur se fixe sur le site appelé site O ou site opérateur qui est positionné entre le promoteur et le gène lac Z. Le répresseur empêche la transcription des trois enzymes en bloquant la progression de la polymérase fixée sur le promoteur.

En présence de lactose, le répresseur forme avec celui-ci un complexe ne lui permettant plus de se fixer sur l'opéron. Le substrat est également appelé inducteur car il permet à la transcription de s'effectuer en piégeant le répresseur. Le répresseur est inactivé par un changement allostérique.

La présence de substrat induit la transcription (opéron inductible). La présence d'un répresseur (répresseur lac) régule la transcription en l'inhibant (régulation négative).

a) Répression catabolique de l'opéron lactose (régulation positive) (figure 12 et 13)

Lorsque les bactéries, par chance, disposent en même temps de glucose et de lactose, la situation se complique !

Le répresseur de l'opéron lactose est inactivé par l'allolactose, le site opérateur de l'opéron est donc libre et les gènes pourraient être transcrits. Or, tant que du glucose est présent, la bactérie va le métaboliser préférentiellement: elle n'a donc pas besoin des enzymes nécessaires au métabolisme du lactose. Ceci implique l'existence d'un autre mécanisme de régulation que l'on appelle la répression catabolique. Ce n'est que lorsque la concentration en glucose diminue que le métabolisme du lactose devient nécessaire. Un signal de carence alimentaire est alors déclenché sous forme d'une augmentation du taux d'AMPc. Cet AMPc forme un complexe avec la protéine CAP (pour Catabolite gene Activator Protein, ou CRP pour cAMP Receptor Protein). Ce complexe se lie à l'ADN en amont du site de fixation de l'ARN polymérase. L'interaction du complexe **CAP-AMPc** va agir comme un **inducteur** et augmenter l'affinité de l'ARN polymérase pour le promoteur de l'opéron. Cette régulation positive peut permettre d'augmenter d'un facteur 50 la transcription de l'opéron lactose. On

comprend alors qu'en présence de glucose, il n'y a pas de complexes CAP-AMPC disponibles : le niveau de transcription de l'opéron lactose est donc très faible.

	<i>P</i>	<i>lacI</i>	<i>P_O</i>	<i>lacZ</i>	<i>lacY</i>	<i>lacA</i>	
ADN							
ARNm		1040	82	3510	780	825	
Polypeptide		360		1021	260	275	Acides aminés
		38 000		116 000	30 000	30 000	Daltons
Protéine		Tétramère		Tétramère	Protéine membranaire	Dimère	Structure
		152 000		500 000	30 000	60 000	Daltons
Fonction		Répresseur		β -galactosidase	Perméase	Transacétylase	

Figure 12: Organisation des gènes de l'opéron lactose

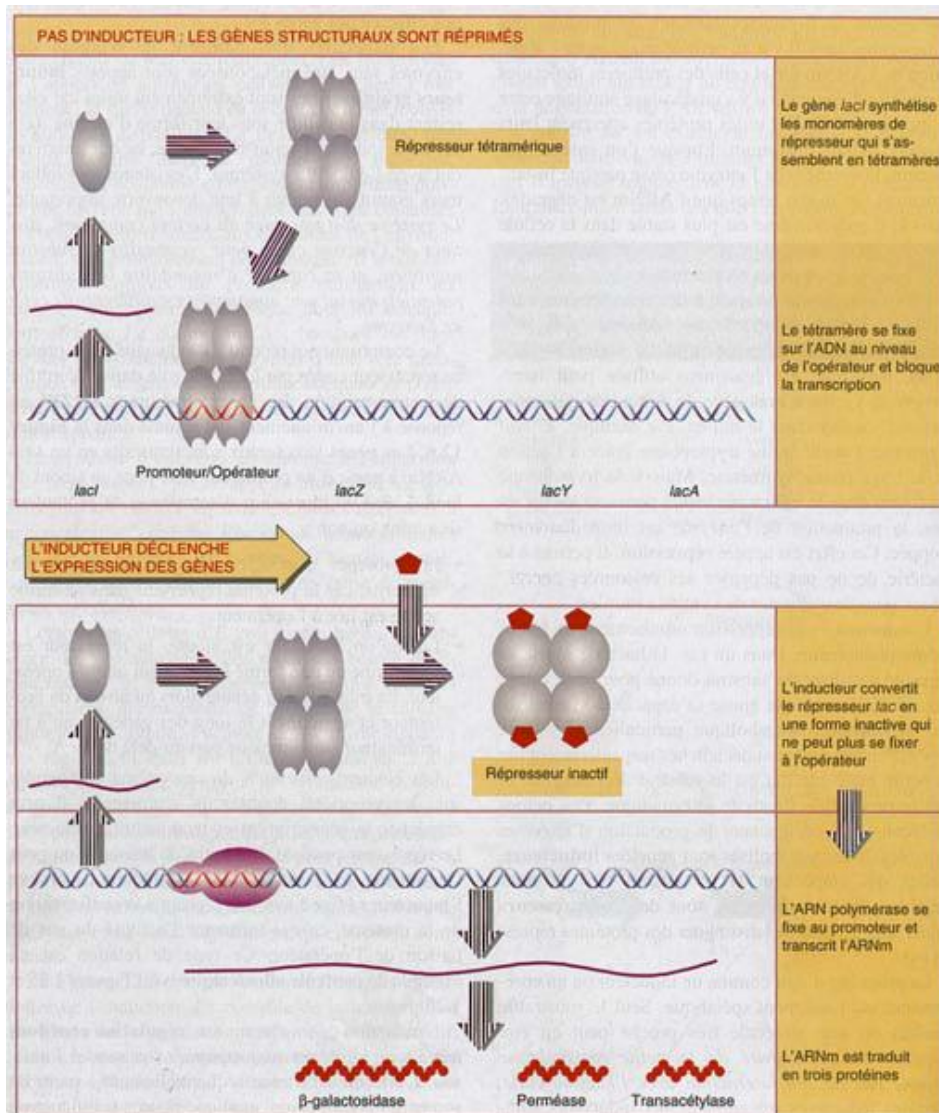


Figure 13 : Régulation de l'expression de l'opéron lactose

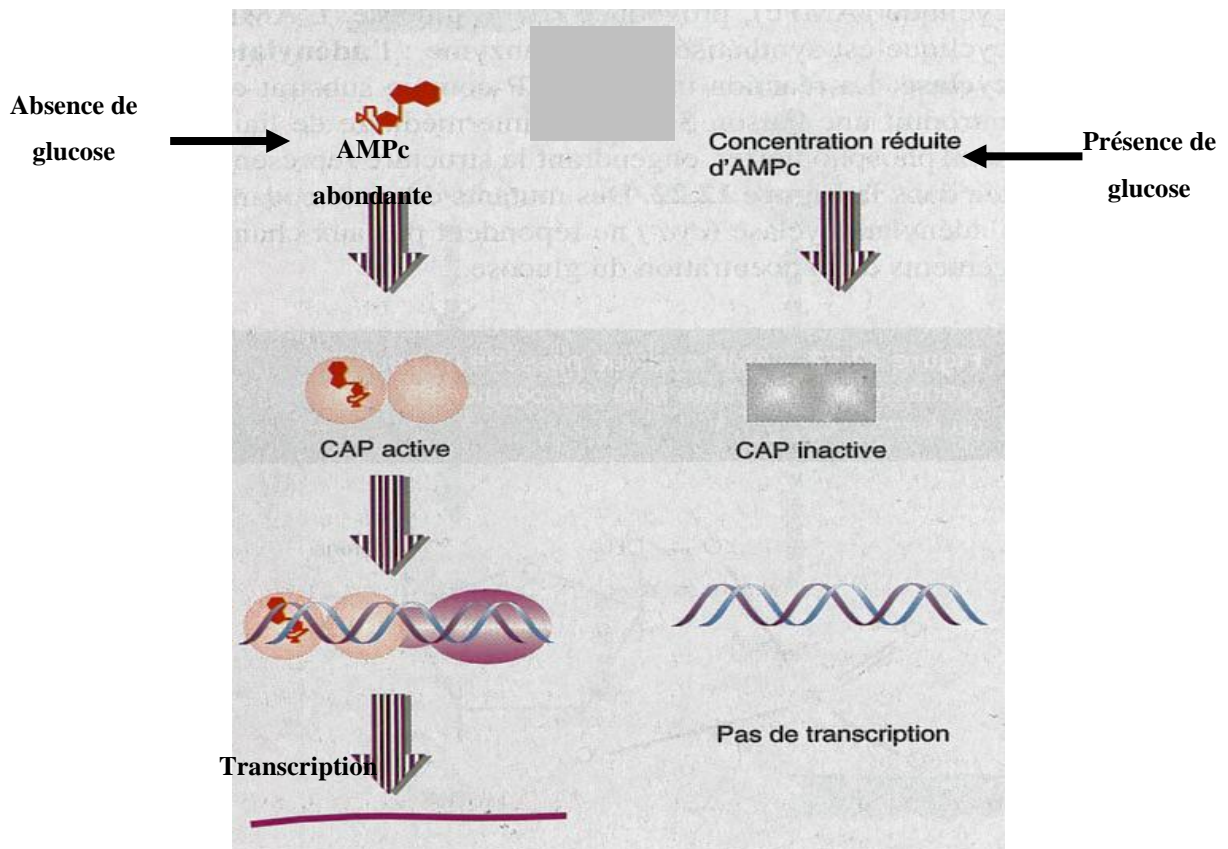


Figure 14 : Régulation de l'opéron lactose par la protéine CAP

I.1.1. 2. L'opéron tryptophane

a) Régulation par un répresseur

C'est un opéron répressible dont la régulation est négative, il présente cinq gènes codant pour des enzymes impliquées dans la biosynthèse du tryptophane. La transcription est régulée par le taux du produit terminal c-a-d le tryptophane présent dans la cellule. En amont de l'opéron, se trouve une séquence régulatrice codant pour un répresseur.

Si le tryptophane est présent dans la cellule, il se fixe au répresseur. Le répresseur est alors activé et peut se fixer sur l'opérateur trp situé juste en aval du promoteur de l'opéron trp. Ceci empêche la progression de l'ARN polymérase et donc la transcription.

Le tryptophane agit comme un corépresseur. Si le tryptophane est absent, le répresseur ne peut pas se fixer à l'opérateur et la transcription a lieu (figure 15).

La présence de l'acide aminé réprime la transcription (opéron répressible). Lorsqu'il y a régulation (répresseur trp + co-répresseur), la transcription est inhibée (régulation négative).

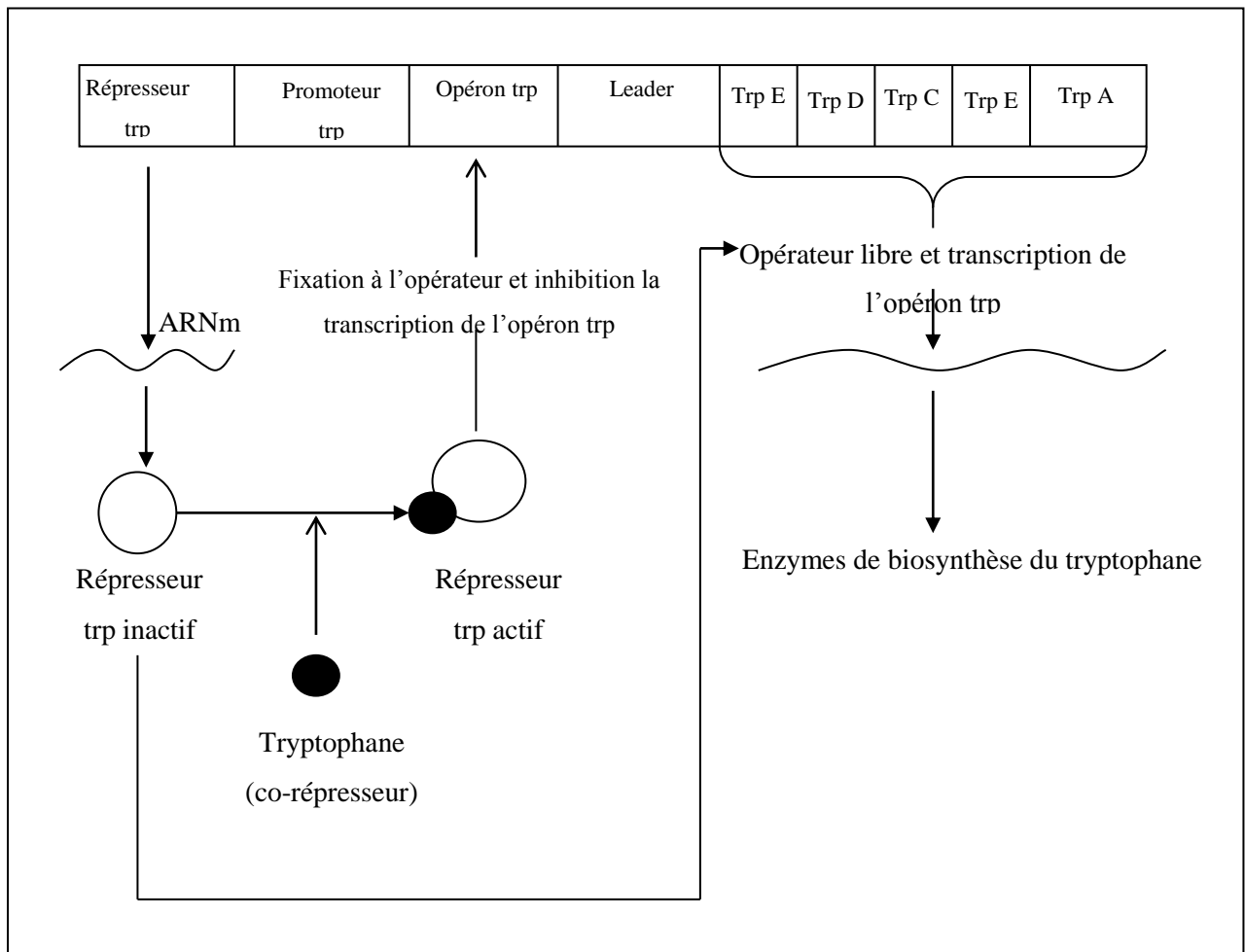


Figure 15: Régulation de l'opéron tryptophane

b) Régulation par atténuation (arrêt précoce de la transcription)

L'atténuation (figure 16) est une stratégie alternative que l'opéron tryptophane utilise pour contrôler la transcription quand le trp est present mais pas en quantité suffisante pour fixer le répresseur.

La séquence d'ARN_m transcrite à partir de la région leader comporte 4 domaines (1, 2, 3, 4) qui est capable de former deux structures en épingle à cheveux mais qui ne peuvent se former en même temps. La plus grosse, n'influence pas la transcription et est présente en amont de la seconde épingle à cheveux, plus petite, qui met fin à la transcription.

L'atténuation est possible parce que la traduction peut avoir lieu en même temps que la transcription. Un ribosome progresse en même temps que l'ARN polymérase. La fixation du ribosome à l'ARN_m détermine celle des deux structures en épingle à cheveux qui sera formée.

La région 1 contient 2 codons *trp*. Lorsque la concentration en *trp* est élevée, cette région est traduite rapidement par le ribosome qui suit de près la polymérase, empêchant ainsi la formation de la grande épingle et permettant la formation de la petite épingle qui mettra fin à la transcription.

En absence de *trp*, le ribosome est bloqué au niveau des codons *trp*, l'ARN polymérase avance et la grande épingle peut se former empêchant la formation de la petite épingle qui va laisser la transcription avoir lieu

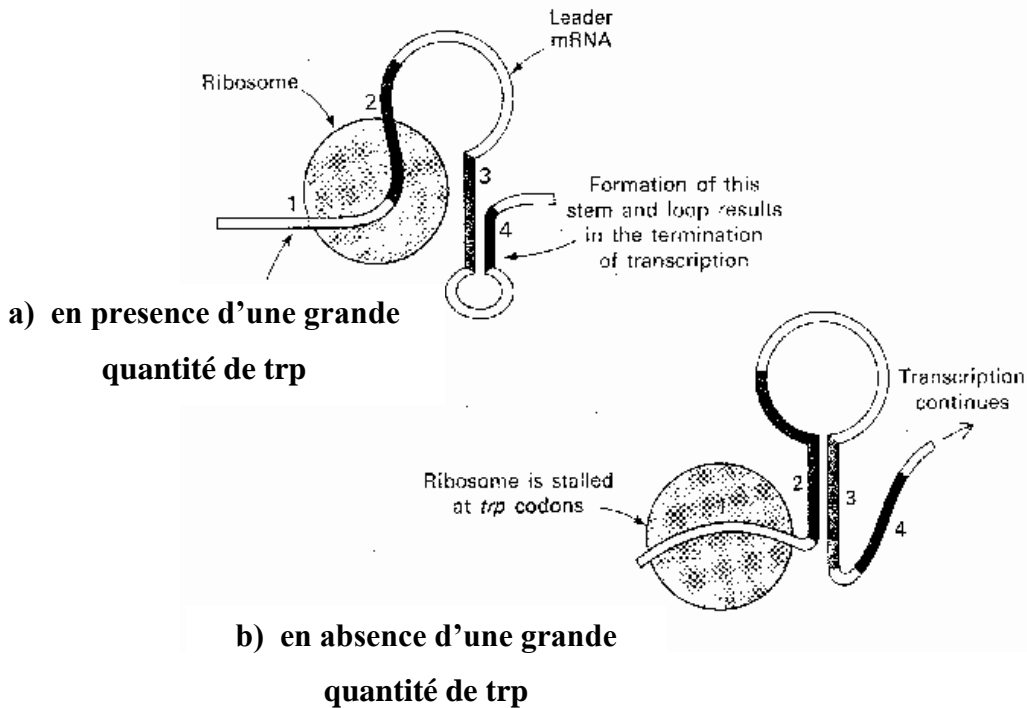


Figure 16: Régulation de l'opéron *trp* par atténuation

I.1.2. Le contrôle négatif et contrôle positif

Les gènes soumis à un contrôle négatif ne sont pas transcrits si un répresseur est lié à l'opérateur. C'est le cas des opérons lactose ou tryptophane. Les gènes soumis à un contrôle positif ne sont transcrits efficacement que si une protéine régulatrice activatrice favorise l'initiation. L'opéron lactose est également soumis à un contrôle positif par la protéine CAP.

I.1.3. Autre mécanismes de régulation de la transcription

I.1.3.1. Régulation par des facteurs sigma alternatifs

L'ARN polymérase bactérienne est composée de 5 sous unités polypeptidiques individuelles. Une de ces sous unités, le facteur δ est responsable de la reconnaissance des promoteurs. Les bactéries dont *E coli*, fabriquent différents facteurs δ alternatifs et font que

l'ARN polymérase transcrit différents ensembles de gènes. Ceci est utilisé comme moyen pour réguler l'expression des gènes en réponses à des changements environnementaux.

Exemple :

- **Choc thermique**

Lorsqu'elle est exposée à une température anormalement élevée, *E.coli* transcrit un ensemble de 17 protéines qui aident la cellule à s'adapter aux nouvelles conditions environnementales. Un facteur sigma alternatif, σ^{32} , est produit et il reconnaît les promoteurs des gènes du choc thermique

- **Sporulation chez *Bacillus subtilis***

La sporulation de cette bactérie requiert l'arrêt de la plupart des synthèses protéiques et la production des protéines nécessaires à la spore. La bactérie réalise ses modifications en utilisant des facteurs alternatifs.

- **Facteurs de bactériophage**

Certains phages utilisent l'ARN polymérase de la bactérie hôte et lui fournissent des facteurs qui lui « apprennent » à transcrire préférentiellement les gènes phagiques.

I.2. Régulation de la traduction

I.2.1. L'autorégulation

Ce mode de régulation est assuré par une protéine ayant la fonction d'un répresseur qui se fixe à une région cible de l'ARN_m. La région d'initiation de la traduction n'est plus reconnue par le ribosome.

Ce type d'autorégulation (figure 17) concerne notamment les gènes codant pour des protéines ribosomiques ou protéines r. lorsque de l'ARN_r est présent, les protéines r s'y associent. En revanche, lorsqu'il n'y a pas d'ARN_r, les protéines r, s'accumulent ce qui entraîne l'inhibition de leur synthèse par fixation de ces protéines à leur ARN_m. Ainsi, dès qu'il y a un excès de protéines r par rapport aux ARN_r, il y a une autorégulation par répression, au niveau des ARN_m codant pour ces protéines. Ce mode de régulation est de type négatif.

L'autorégulation est un type de contrôle courant parmi les protéines incorporées dans des assemblages macromoléculaires. La particule assemblée complète peut ne pas convenir comme régulateur en raison de sa taille trop grande (exemple : tubuline associées aux microtubules). La tubuline emprisonnée dans les microtubules ne joue aucun rôle dans la

régulation mais leur concentration à l'état libre détermine si d'autres monomères seront synthétisés ou non.

Les opérons inductibles et répressibles sont régulés par un contrôle extrinsèque (avec inducteur ou corépresseur) mais l'autorégulation est un système intrinsèque : le paramètre critique est la concentration de la protéine elle-même.

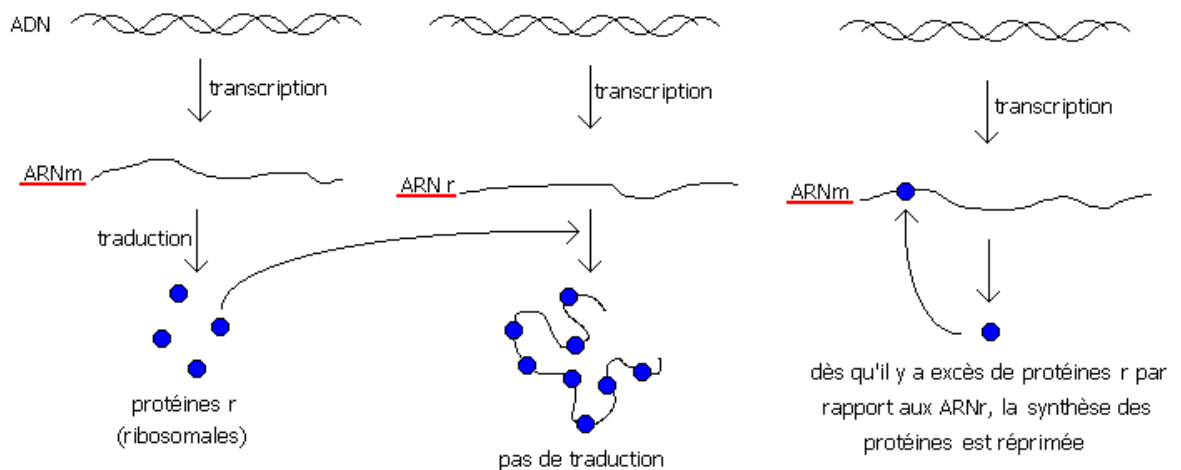


Figure 17 : mécanisme de l'autorégulation chez les procaryotes (exp : proteine r)

I.2.2. Régulation par des petites molécules d'ARN

Les régulateurs ne sont pas que des protéines. De petites molécules d'ARN se fixent sur une séquence spécifique simple brin ce qui entraîne la formation d'une région double brin. Celle-ci peut empêcher la traduction de façon directe en emprisonnant le site d'initiation de la traduction, ou indirecte en modifiant la conformation d'une autre région.

Ces petits ARN régulateurs sont appelés anti-sens.

Différents mécanismes de régulation sont décrits chez les bactéries.

- L'ARN anti-sens peut se fixer au niveau de l'ARN_m et masquer, pour le ribosome, le site d'initiation de la traduction,
- L'ARN antisens en se fixant sur l'ARN_m peut engendrer une région double brin reconnue par une endonucléase qui clive et dégrade cette région,
- L'ARN antisens peut se fixer sur l'ARN_m et faire office de terminateur.

Les ARN antisens régulent également l'expression chez les eucaryotes.

II. Chez les eucaryotes

Chez les eucaryotes, la régulation de l'expression des gènes est beaucoup plus compliquée que chez les procaryotes. En fait, une cellule peut contrôler les protéines qu'elles synthétisent en contrôlant :

- 1- la transcription
- 2- la maturation de l'ARN
- 3- Le transport de l'ARN_m
- 4- La traduction
- 5- La dégradation de l'ARN_m
- 6- L'activité des protéines

Parmi les différents niveaux de contrôle possibles, le contrôle transcriptionnel est le seul que permette d'empêcher la synthèse d'intermédiaires superflus.

II.1. Contrôle de l'expression des gènes par la structure chromatinienne

La chromatine est constituée d'un mélange d'ADN et de protéines. Des nucléosomes forment un élément essentiel de la chromatine. Ils sont constitués par environ 200 paires de bases d'ADN associées à un octamère d'histones. La structure chromatinienne constitue le premier niveau de régulation des gènes, une région est active lorsqu'elle est fréquemment transcrite. Ces régions actives sont moins condensées qu'ailleurs et les nucléosomes y seraient dépourvus d'histones de condensation H1. Les gènes codant les ARN ribosomiaux, qui sont donc les plus fréquemment transcrits, semblent se situer dans des régions dépourvues de nucléosomes. Les gènes " domestiques ", qui codent en permanence de nombreuses protéines de la vie cellulaire, sont eux aussi situés en dehors des nucléosomes

II.2. Régulation de la transcription

II.2.1. Les facteurs de la transcription basaux

Dans la cellule eucaryote, les gènes codant les protéines sont tous transcrits par l'ARN polymérase II. La transcription est initiée par la formation du complexe d'initiation de la transcription (TIC) qui met en jeu la fixation de l'ARN polymérase II et d'un certain nombre de protéines associées, les facteurs basaux, au promoteur au niveau d'une séquence caractéristique appelée la TATA BOX

II.2.2. Les facteurs de transcription trans régulateurs

L'efficacité de l'initiation de la transcription est influencée par des facteurs de transcription supplémentaires qui se fixent à d'autres séquences d'ADN et capables d'interagir avec des protéines du TIC, affectant la stabilité de ce dernier. De tels facteurs de transcription peuvent augmenter ou diminuer le taux de transcription.

II.2.3. Les séquence cis régulatrices

a) Les séquences régulatrices proximales

Généralement situées entre 40 et 110 paires de bases en amont du site d'initiation, les séquences proximales peuvent avoir un effet activateur ou répresseur en fonction de l'activité de la protéine qui les reconnaît. Les deux séquences les mieux étudiées sont les boîtes CAAT et les motifs riches en GC. Ces séquences d'ADN sont en général reconnues par des protéines qui régulent la transcription par contacts directs avec la machinerie de transcription de base.

b) Les séquences régulatrices distales

Les éléments distaux du promoteur peuvent être situés jusqu'à quelques milliers de paires de bases en amont ou en aval du site d'initiation. Ces éléments peuvent activer la transcription, on les appellera alors 'enhancers', ou la réprimer, on parlera de 'silencers'. Les séquences activatrices peuvent agir indépendamment de leur orientation en cis. Le repliement de l'ADN permettrait le rapprochement de ces séquences régulatrices du site d'initiation et, de ce fait, les interactions avec la machinerie de transcription peuvent se créer.

II.2.4. Les domaines des facteurs de transcription

Il existe deux domaines principaux, un domaine de fixation à l'ADN (DBD, pour *DNA binding domain*) et un domaine d'activation de la transcription (AD, pour *activation domain*). Le rôle du domaine de fixation à l'ADN permet de placer le domaine d'activation à proximité du complexe d'initiation. Ces domaines présentent des motifs particuliers (figure 18):

a) Les motifs en doigts de zinc.

Ce type de protéine comporte des éléments répétitifs avec une forme de doigt de gant. On peut retrouver de nombreux doigts. Ces protéines tirent leur nom de leur structure, dans

laquelle un groupe d'acides aminés conservés lient un ion de zinc. L'ion Zn^{2+} sert à stabiliser le motif sous forme de doigt.

b) Les motifs hélice-coude-hélice

Ce motif est composé de 2 hélices séparées par un coude. Une des hélices, appelée l'hélice de reconnaissance de l'ADN en entrant en contact avec le sillon majeur de la double hélice.

c) Les glissières à leucine.

Les protéines avec glissière à leucine se fixent à l'ADN sous forme de dimères. Elles possèdent deux régions importantes. Une région à base du dimère, constituée par deux hélices face à face riches en leucine et interagissant par des liaisons de type hydrophobe. Une autre région riche en charges positives se fixe sur les groupes phosphates de l'ADN.

d) Les motifs hélice/boucle/hélice.

Comme son nom l'indique, ce motif est caractérisé par deux segments en hélice α séparé par une boucle intercalaire.

Certains facteurs trans-régulateurs possèdent un troisième domaine qui permet de fixer un ligant permettant de réaliser l'action d'un message extérieur à la cellule, comme un message hormonal. On a pu ainsi définir la superfamille des récepteurs nucléaires qui sont des facteurs de transcription inductibles par un ligand (comme une hormone stéroïde, mais aussi par exemple certains lipides intracellulaires.

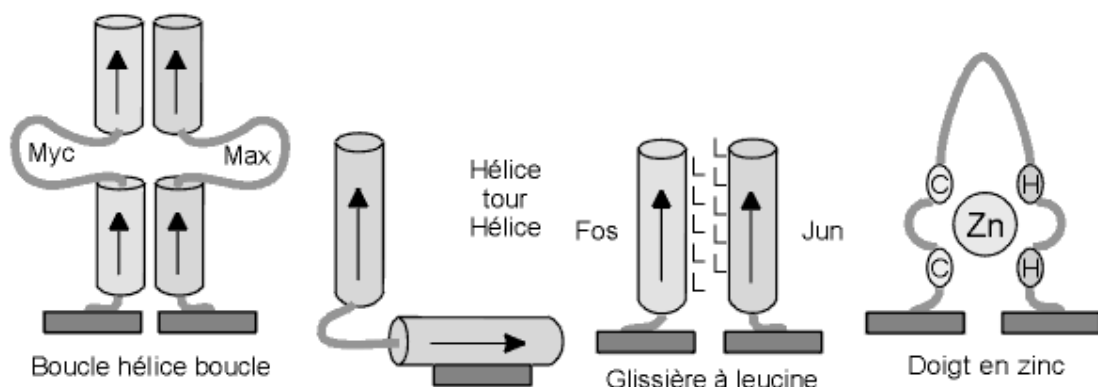


Figure 18 : Schéma des principaux types de facteurs de transcription.

II.3. Régulation post-transcriptionnelle

II.3.1. L'épissage alternatif ou différentiel (alternative RNA splicing).

L'une des principales modifications du RNA transcrit est l'épissage (splicing) (figure 19) qui consiste en l'élimination des introns et à la conservation des exons. Généralement, une cellule peut épisser le transcrit primaire de différentes façons, et donner ainsi plusieurs protéines différentes à partir d'un seul gène. En fait, dans certains cas, le complexe d'épissage n'est pas capable de distinguer nettement entre les différents sites d'excisions. Ce processus est appelé épissage alternatif ou épissage différentiel.

L'épissage de l'ARN peut être contrôlé soit négativement, par une molécule régulatrice qui empêche au complexe d'épissage l'accès à un site d'épissage particulier sur l'ARN, soit positivement, par une protéine régulatrice qui dirige la machinerie d'épissage vers un site d'épissage qui sinon n'aurait pas été reconnu

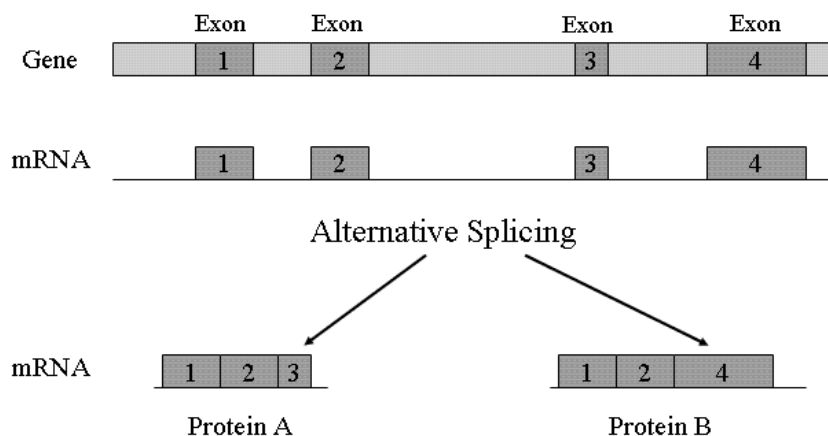


Figure 19 : Régulation post transcriptionnelle chez les eucaryotes

II.3.2. L'édition.

On connaît depuis peu, un processus qui altère la correspondance entre ADN et ARNm : ce processus est appelé édition (*editing*) de l'ARN. La transcription produit un préARNm au niveau duquel il peut y avoir insertion ou délétion d'un ou de plusieurs nucléotides, ce qui peut changer la séquence de la protéine. De cette définition est exclue bien entendu, l'excision des introns.

a) L'édition par insertion d'une base et décalage du cadre de lecture.

Ce cas a été décrit pour une protéine du virus de la rougeole. Le gène P peut coder pour un ARNm lui correspondant exactement nucléotide à nucléotide. Mais on peut avoir également

une insertion d'un nucléotide supplémentaire (G) ce qui provoque un changement du cadre de lecture dans la succession des triplets (codons) nucléotidiques. Dès lors, on peut dire d'un gène unique peut exprimer deux protéines différentes.

b) L'édition avec changement d'un seul codon.

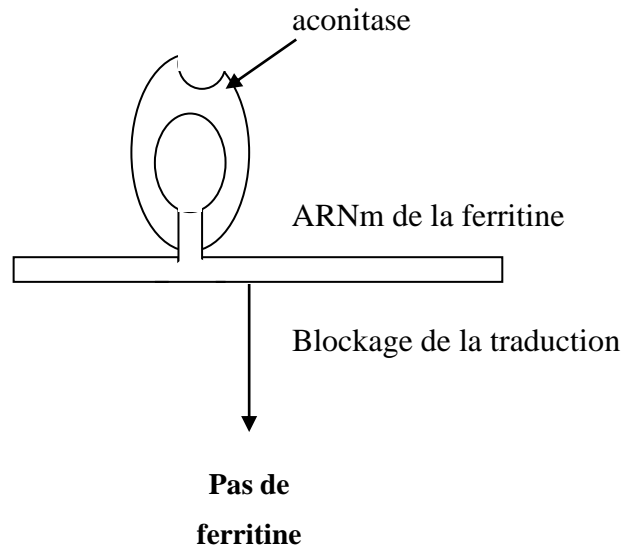
L'exemple typique est celui d'une apolipoprotéine B chez l'Homme. L'apparition d'un codon stop dans le messager de cette apolipoprotéine conduit à une protéine raccourcie du fait de l'apparition prématurée de ce codon-stop. Ainsi, un même gène peut ainsi donner naissance à deux ARNm différents, traduits en deux protéines différentes, l'une dans le foie (apolipoprotéine B100), l'autre dans l'intestin (apolipoprotéine B48).

II.4. Régulation de la traduction

Chez les eucaryotes, une forme de contrôle négatif de la traduction permet le réglage rapide de la synthèse de la protéine de stockage intracellulaire du fer, la ferritine (figure 20).

La régulation ferrique dépend d'une séquence d'environ 30 nucléotides dans la région 5' du messager de la ferritine. Cet élément de réponse au fer se replie en une structure tige boucle qui fixe un répresseur de la traduction appelé l'aconitase, qui réprime la traduction de toutes séquence d'ARN en aval. L'aconitase est une protéine de liaison au fer, et l'addition de fer à la cellule dissocie cette protéine de l'ARNm de la ferritine, et permettant ainsi la traduction de l'ARNm de la ferritine.

Carence en fer



Excé de fer

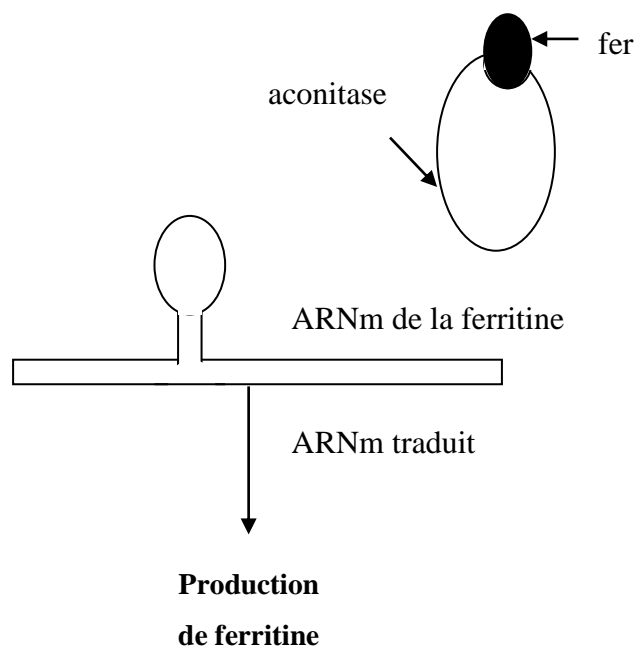


Figure 20 : régulation de la traduction chez les eucaryotes (exp : la ferritine)

Deuxième partie

Génie génétique

Préambule

Le génie génétique est principalement basé sur le clonage moléculaire qui permet la recombinaison d'un fragment double brin d'ADN avec un vecteur et son introduction dans un hôte approprié.

Cette technologie a d'abord concerné les organismes les plus simples qui ont souvent des génomes de petite taille comme les bactéries (Ex. *E. coli*, *B. subtilis*, *Agrobacterium tumefaciens*...etc.), et les levures (*S. cerevisiae*). Ce n'est que plus tard que la transgénèse stable des cellules végétales et animales.

Actuellement les OGM (végétaux, animaux transgéniques) sont nombreux et certains commencent à parvenir dans nos assiettes.

Le premier apport de cette technologie à la médecine fut de rendre des protéines thérapeutiques telle: insuline, HGH (human growth hormone), disponibles en quantités suffisantes. La médecine a retiré bien d'autres bénéfices de la technologie de l'ADN recombinant, vaccins (anti hépatite, rage, ...etc.), sondes (diagnostique, identification ...etc.), agents thérapeutiques (acides nucléiques anti sens, thérapie génique, ...etc.).

La technologie de l'ADN recombinant a touché d'autres produits d'intérêt agro-alimentaire et industriel, enzymes (présure, protéases, ...etc.), acides aminés (glutamate, succinate, ...etc.), vitamines (B2, B12, C ...etc.).

Dans les deux dernières décennies plusieurs milliers d'entreprises de biotechnologie ont été créés dans le monde surtout en USA et le Japon.

Les différentes étapes successives à entreprendre en génie génétique sont les suivants:

- 1-** Isolement d'une séquence d'acide nucléique, idéalement un gène.
- 2-** L'insertion de ce gène dans un vecteur (Ex. plasmide ou phage).
- 3-** L'incorporation de ce vecteur dans un hôte (Ex. *E. coli*).
- 4-** L'hôte transformé peut assurer la propagation du vecteur à chaque cycle de réplication.

Chapitre I

Les techniques de l'ADN recombiné (clonage moléculaire)

Chapitre I : Les techniques de l'ADN recombiné (Clonage moléculaire)

I. Objectif

Les techniques de l'ADN recombiné, appelées également clonage moléculaire ont pour objectif l'introduction d'un gène d'un organisme dans une cellule hôte, dans laquelle il peut être cloné au fil des générations et étudié.

II. Principe du clonage

L'ADN est extrait d'un organisme donneur et coupé par des enzymes de restriction. Chaque fragment issu du traitement enzymatique (insert d'ADN) est attaché (lié, ligaturé) à un vecteur de clonage qui possède la propriété de se répliquer dans une cellule hôte particulière pour former une molécule d'ADN recombinée complète. Chaque construction est introduite dans une cellule hôte, ou elle se perpétue (figure 21).

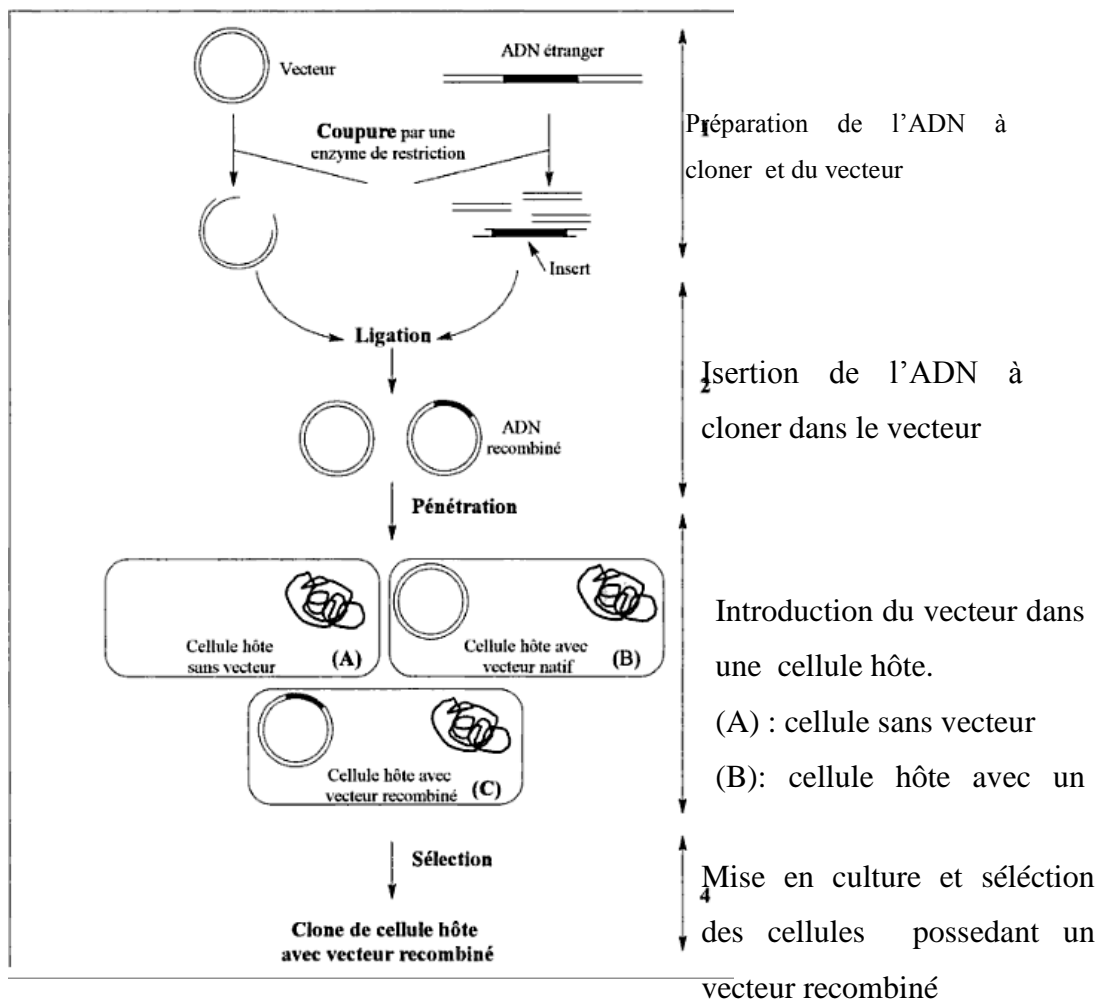


Figure 21 : Les étapes du clonage moléculaire

III. Les enzymes utilisées dans le clonage

III.1. Les enzymes de restriction

III.1.1. Définition

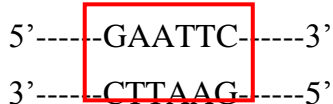
Les enzymes de restriction ou endonucléases de restriction sont des enzymes qui coupent l'ADN. Elles sont obtenues à partir des bactéries qui les produisent à fin de se protéger contre tout ADN exogène. Une enzyme donnée reconnaît une séquence palindromique spécifique de 4 à 8 (le plus souvent 6) nucléotides (sites de restriction) à l'intérieur de laquelle elle clive l'ADN. La taille des fragments d'ADN produits dépend de la répartition des sites de restriction. Plus de 400 types d'enzymes de restriction différents ont été isolés.

III.1.2. Nomenclature

Le nom d'une enzyme de restriction est composé de 3 ou 4 lettres. L'une des premières endonucléases de restriction identifiées et employées pour cloner des gènes a été purifiée à partir de la bactérie d'*Escherichia coli*. Cette enzyme est appelée EcoRI. Eco désigne la bactérie dont elle est extraite avec E pour le genre et co pour l'espèce ; la lettre R indique la souche (c'est une endonucléase) et si c'est M c'est une méthylase. Le chiffre romain I, ordre de caractérisation des enzymes, veut dire que cette enzyme est la première endonucléase de restriction extraite d'*E. coli*. Cette nomenclature est employée pour toutes les endonucléases de restriction, quoique l'on ne mentionne pas toujours la souche.

Exemple :

- ***EcoRI*** : cet enzyme reconnaît le palindrome :



- ***HpaI*** : isolé de *Haemophilus parainfluenzae*, cet enzyme reconnaît le palindrome :



III.1.3. Les différents types

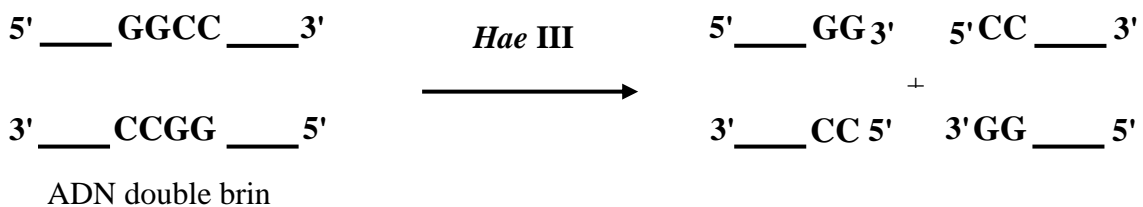
La plupart des enzymes de restriction utilisées au laboratoire présentent un site de reconnaissance et un site de coupure identique ou proche du site de reconnaissance, ce sont des enzymes de type II. Certaines bactéries possèdent d'autres types d'enzymes de restriction.

Les enzymes de type I reconnaissent des séquences sur l'ADN sans aucune symétrie. Ces enzymes coupent non pas au niveau de la séquence de reconnaissance mais 1000 à 5000 paires de bases plus loin. Les enzymes de type III présentent également un site de reconnaissance mais coupent une vingtaine de nucléotides plus loin.

III.1.4. Clivage de l'ADN par des endonucléases de restriction

Les enzymes de restriction utilisées coupent dans la séquence selon deux modes : coupure franche (bouts francs) et coupure décalée (coupure cohésives) sur les deux brins.

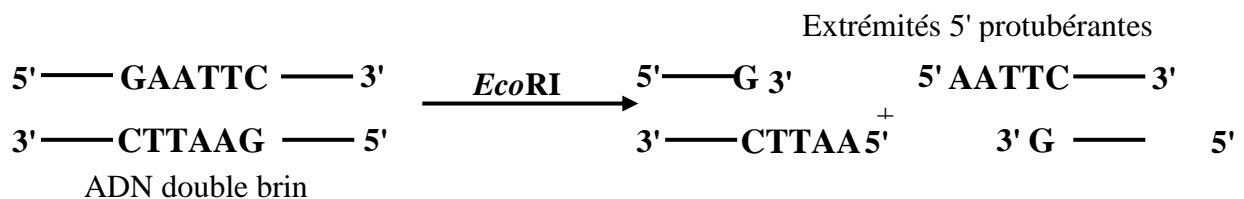
a) Coupure franche : Les extrémités franches sont obtenues par coupure au même endroit sur les deux brins. Exemple : l'enzyme *Hae* III isolée de *Haemophilus aegyptius* coupe l'ADN double brin au niveau de la séquence GG/CC :



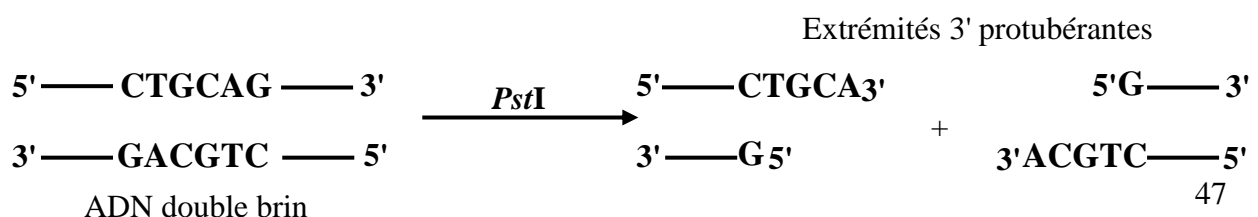
b) Coupure décalée sur les deux brins : Elles sont générées des extrémités monocaténares (simple brin) ayant des séquences complémentaires appelées extrémités cohésives ou bouts collants.

Exemples :

1. L'enzyme de restriction *Eco*RI coupe l'ADN double brin dans la séquence palindromique G/AATTC.



2. L'enzyme de restriction *Pst*I isolée de la bactérie *Providencia stuarti* coupe l'ADN double brin dans la séquence palindromique CTGCA/G.



III.1.5. Isoschizomères

Les isoschizomères sont des enzymes de restriction qui proviennent de souches bactériennes différentes, qui portent donc des noms différents, mais ils coupent l'ADN au niveau d'un même site de restriction. Les isoschizomères fournissent souvent après clivage enzymatique des fragments dont les extrémités sont différentes.

3.1.6. Enzymes compatibles

Les enzymes compatibles, ne reconnaissent pas le même site de restriction, mais ils génèrent des fragments aux extrémités « cohésives » complémentaires, qui peuvent donc être ligaturés.

Exp: Les enzymes Bam HI et Mbo I sont compatibles

Premier fragment avant clivage par Bam H I 5'GGA TCC3' 3'CCTAGG5'	Après clivage par Bam H I 5'G3' (1) + 5'GA TCC3' (2) 3'CCTAG5'	
Premier fragment avant clivage par Mbo I AGATGAGC CTAGGGC	Après clivage par Mbo I 5'A3' (3) + 5'GATCAGC3' (4) 5'TCTAG5' 3'TCG5'	
Ligatures possibles	1+4 : 5'GGA TCA GC3' 3'CCTAGTCG5'	2+3 : 5'AGA TCC3' 3'TCTAGG5'

III.1.7. Cartographie de fragments d'ADN au moyen d'enzymes de restriction

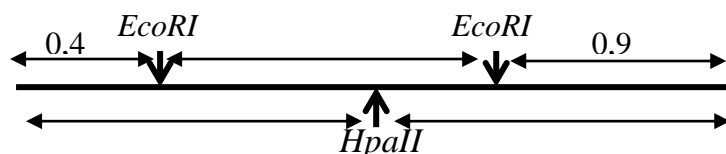
Les endonucléases de restriction fournissent le moyen de fragmenter des génomes complexes en séries de fragments plus petits. L'analyse des fragments produits par plusieurs endonucléases de restriction utilisées séparément ou de façon combinées permet de déduire l'ordre des fragments dans la molécule d'ADN originale. Le résultat est une carte physique de l'ADN.

Exemple de carte de restriction:

Etablir la carte de restriction d'un gène de 3 Kb.

- Digestion totale *EcoRI* : 1,7Kb, 0,9Kb et 0,4 Kb
- Digestion totale *HpaII* : 1,6Kb et 1,4Kb
- Double digestion *EcoRI* et *Hpa II* : 0,4 kb, 0,5Kb, 0,9Kb et 1,2Kb.

Corrigé :



III.2. Autre enzymes

III.2.1. Enzymes de modification (méthylase)

L'ADN bactérien présente des sites de restriction susceptibles d'être repérés par les enzymes de restriction que possède la bactérie. Pour éviter une auto-destruction, les enzymes de modification de l'ADN bactérien interviennent. Ces enzymes de modification sont des méthylases bactériennes (ou enzymes de méthylation). La méthylation de la cytosine (sur le carbone 5) ou de l'adénine (sur l'azote 6) appartenant à des sites de restriction aboutit à la non reconnaissance de ce dernier par l'enzyme de restriction correspondante. Cette méthylation peut se réaliser sur une base ou sur plusieurs bases appartenant au site de restriction. Les méthylases bactériennes sont très spécifiques puisque la méthylation porte uniquement sur les cytosines et adénines appartenant à des sites de restriction.

III.2.2. La DNase

La DNase utilisée au laboratoire est extraite du pancréas de bovin. Il s'agit d'une endonucléase qui coupe l'ADN double brin (mais aussi l'ADN simple brin). Elle conduit à des coupures ou « nicks » tout à fait au hasard, sans reconnaissance d'un site spécifique (ce qui la distingue des enzymes de restriction). On obtient des fragments de tailles variées (ou oligonucléotides) qui possèdent en leur extrémité 5' un groupement phosphate.

III.2.3. La nucléase S1

Cette enzyme extraite d'un champignon (*Aspergillus oryzae*), n'attaque que l'ADN simple brin. Elle n'attaque pas en principe les ADN doubles brins et les hybrides ADN-ARN.

III.2.4. La ligase

Pour lier (souder) deux fragments d'ADN, on utilise une ligase, en présence d'ATP. Elle favorise la formation d'une liaison ester entre un fragment 5' phosphate et un autre fragment 3'OH. La ligase provenant d'*E. coli* ne peut que lier deux ADN à extrémité cohésives, mais la ligase virale T4 peut aussi lier des bouts francs.

III.2.4. La phosphatase (déphosphorylase)

Les phosphatases dites alcaline (active à pH alcalin), catalysent l'élimination d'un groupement phosphate en 5' d'une chaîne de DNA. Elles sont extraites de bactéries ou d'origine animale (intestins). Elles sont utilisées pour déphosphoryler un vecteur ouvert par une

enzyme de restriction, afin d'éviter une refermeture de ce vecteur (auto-ligation) qui empêcherait alors d'insérer le fragment de DNA à étudier (figure 22).

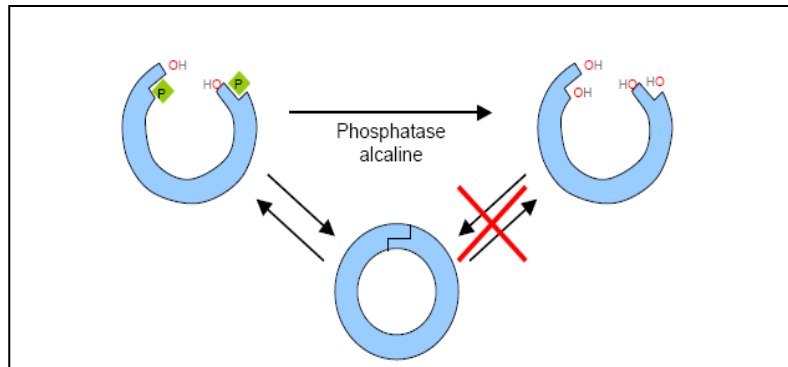


Figure 22: Action de la phosphatase alcaline

III.2.5. Les Kinase (Phosphorylase)

Elles sont extraites de bactéries et elles permettent le transfert d'un groupement phosphate à partir d'une molécule d'ATP vers l'extrémité d'un ADN déphosphorylé.

III.2.6. La rétrotranscriptase ou transcriptase reverse

C'est une ADN polymérase ARN –dépendant. Elles permettent la synthèse in vitro d'ADN à partir d'une matrice d'ARN. La réaction catalysée par les transcriptases inverses se déroule selon un mécanisme analogue à celui de l'ADN polymérase classique.

IV. Les vecteurs de clonage

Les vecteurs sont des petits ADN dans lesquels on insère le fragment d'ADN à étudier. Ces petits ADN sont généralement des *bactériophages* ou des *plasmides*. Ils possèdent dans leur génome les signaux nécessaires pour leur répllication, mais ils ne savent pas se multiplier seuls. Ils doivent être introduits dans des cellules hôte (bactéries par exemple).

IV.1. Les plasmides

IV.1.1. Définition

Les plasmides utilisables pour le clonage possèdent trois caractéristiques communes : un site de répllication, un ou plusieurs marqueurs de sélection et un ou plusieurs sites de clonage. Le site de répllication est l'origine de la répllication. Le marqueur de sélection est classiquement un gène conférant une résistance à un antibiotique et le site de clonage est séquence possédant plusieurs sites de restriction.

IV.1.2. Exemple de plasmides

- **pBR322** (proposé par Bolivar et Rodriquez) : C'est un plasmide qui a été très utilisé. Il possède deux gènes de résistance à des antibiotiques l'un **à l'ampicilline**, l'autre à la **tétracycline**. Ces deux gènes contiennent des sites uniques de restriction, utiles pour le clonage. L'insertion d'un ADN étranger peut se faire dans l'un ou l'autre de ces deux sites. Seule persisterait la résistance à l'antibiotique due à celui des deux gènes n'ayant pas reçu d'insert, ce qui permettrait ainsi une sélection des plasmides recombinants (insertion d'un ADN étranger dans le gène tetR inactive ce dernier et cela supprime la résistance à la tétracycline et la cellule devient sensible.

- **pUC 18**: C'est un plasmide (Plasmid of University of California) de 2.69Kb, Il contient un gène de résistance à l'ampécilline et le gène lacZ responsable de la fabrication de la B-Galactosidase. ce dernier contient un poly linker, c'est-à-dire une séquence poly nucléotidique synthétique correspondant à une série de site unique de restriction. L'insertion d'un ADN étranger au sein de ce gène, inactive ce dernier et rend la bactérie incapable de synthétiser la B- Galactosidase.

IV.1.2. Préparation des plasmides pour la recombinaison

L'insertion d'une séquence d'ADN bi caténaire dans un plasmide nécessite un traitement préalable par une enzyme de restriction. On choisit une enzyme qui a un site unique dans la séquence plasmidique. Après action de l'enzyme de restriction, le plasmide circulaire est liniarisé. Pour maintenir la linéarisation, il est indispensable de traiter le plasmide par une phosphatase alcaline qui enlève les groupements phosphates en 5'. Le plasmide est ensuite ajouté au fragment d'ADN à insérer en présence d'une ligase. L'ensemble de ces étapes produit un plasmide recombinant. L'incorporation des plasmides dans les bactéries est appelée transformation bactérienne.

Les bactéries peuvent être placées dans certaines conditions physico-chimiques pour permettre aux plasmides d'être incorporés à l'intérieur de celles-ci. Les bactéries traitées et prêtes à recevoir l'ADN étranger sont appelées bactéries compétentes. L'incorporation des plasmides dans les bactéries est appelée transformation bactérienne.

IV.1.3. Sélection des bactéries ayant incorporé le plasmide recombiné

a) Système du second gène de résistance aux antibiotiques (figure 23a)

Le DNA étranger est inséré au sein de ce second gène de résistance pour un antibiotique, cette insertion va se traduire par une inactivation de ce gène, donc par la perte de la résistance à cet antibiotique.

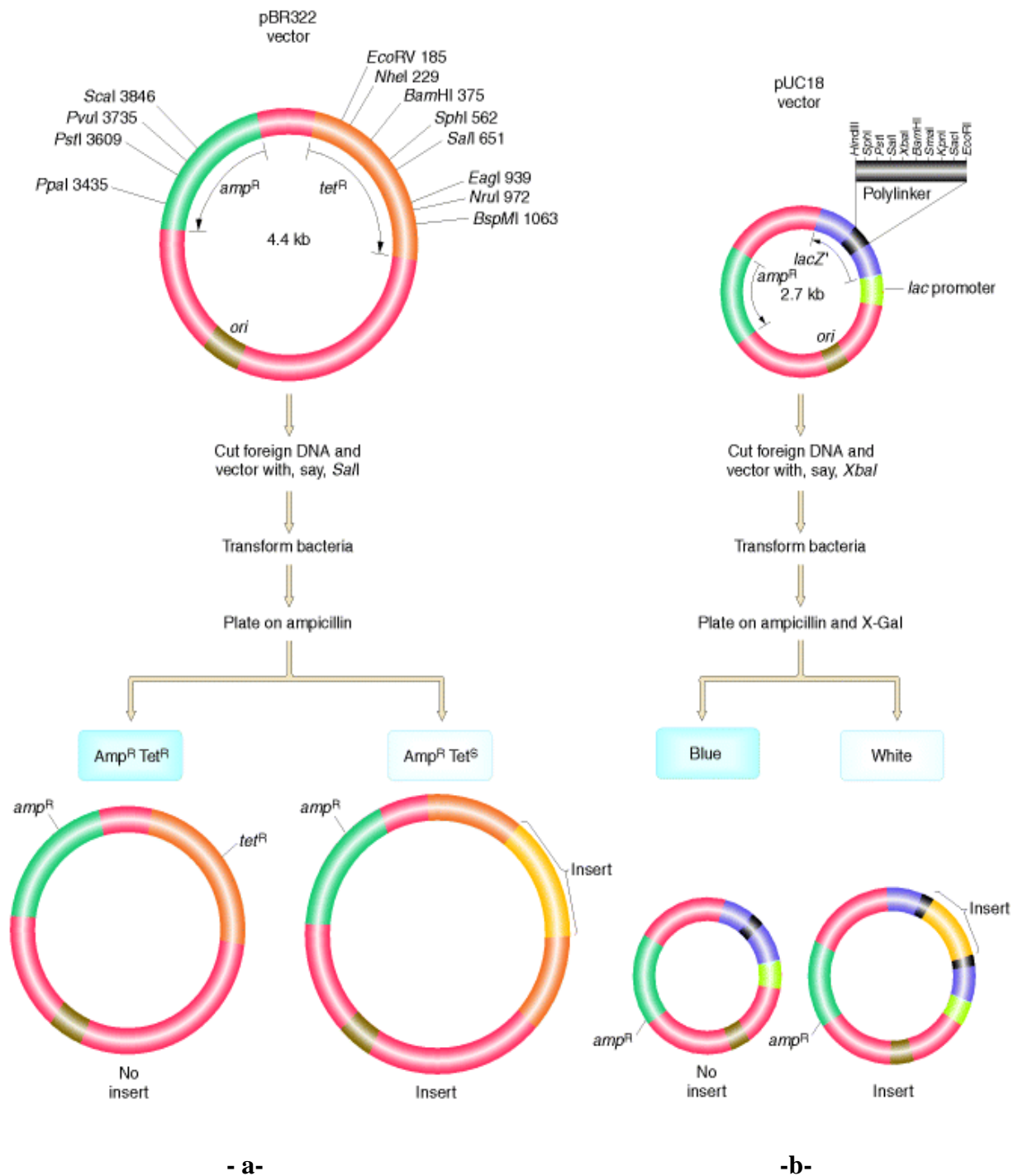


Figure 23: Les plasmides pBR322 et Puc 19

Les bactéries qui ont incorporé un plasmide recombinant contenant un DNA inséré seront résistantes au premier antibiotique, mais sensibles au second. Par contre les bactéries qui auront incorporé un plasmide sans DNA inséré seront résistantes aux deux antibiotiques.

b) Système de l'opéron lactose

Pour sélectionner les recombinants, il est très pratique d'utiliser un système visuel : colonies bleu pour les non recombinants qui synthétisent la B-galactosidase et blanche pour les recombinants. La transformation des bactéries lac⁻ par un plasmide contenant un gène lacZ (Galactosidase) confère à cette bactérie le phénotype lac⁺ qui peut être caractérisé directement au niveau des colonies en utilisant un substrat chromogène. Pour cela les bactéries sont cultivées en présence de l'X-Gal (5-bromo-4-chloro-3-indolyl-B-D-galactosidase), un galactoside dont la couleur passe de l'incolore au bleu quant il est clivé par la galactosidase. Donc sur un milieu de culture contenant l'X-Gal, les bactéries n'ayant pas été recombinée (Lac⁺), seront colorées en bleu car elles sont capable de dégrader l'X-gal, par contre les bactéries, ayant été recombinés (lac⁻), auront une coloration blanche. En effet, l'insertion d'un DNA étranger au sein du gène lac Z inactive ce dernier et les colonies correspondantes resteront blanches (figure 23b).

IV.2. Les bactériophages

Les bactériophages sont des virus qui infectent les bactéries. Deux phages sont très utilisés comme vecteurs : les phages lambda et M13.

IV.2.1. Phage lambda

-Les différentes parties du phage lambda : On distingue deux parties dans ce phage ; la tête qui renferme l'ADN viral et la queue qui permettra au virus de se fixer sur la cellule-hôte bactérienne. Le phage lambda est un phage double brin, linéaire, d'une longueur de 48.5Kb. Il comprend plusieurs dizaines de gènes. Les extrémités de cet ADN sont simple brin sur une longueur de 12 nucléotides, et complémentaire l'une de l'autre. Elles sont appelées « extrémités cohésives » ou séquences COS.

Sur le plan importance, le génome du phage lambda peut être divisé en partie essentielle et non essentielles. La première partie comprend les gènes codant pour les protéines situées dans la tête et la queue du virus, les sites COS, et le site de répllication de l'ADN viral. La seconde partie est représentée par les gènes impliqués dans la lysogénie.

- **Les cycles du phage lambda** : Le phage lambda se multiplie selon deux modes possibles : multiplication lytique ou lysogénique (figure 24).

- **La multiplication lytique**

Le virus s'adsorbe spécifiquement à la surface des bactéries. Il injecte son ADN à la cellule-hôte, l'ADN viral se circularise dans la bactérie. Puis, une phase complexe de réplication débute et aboutit à la constitution de nombreuses particules virales à l'intérieur du cytoplasme bactérien. La lyse bactérienne provoque la libération des particules virales.

- **La multiplication lysogénique**

Comme dans la multiplication lytique, le virus s'adsorbe et pénètre dans la bactérie. L'ADN viral s'intègre à l'ADN bactérien et sera répliqué en même temps que lui.

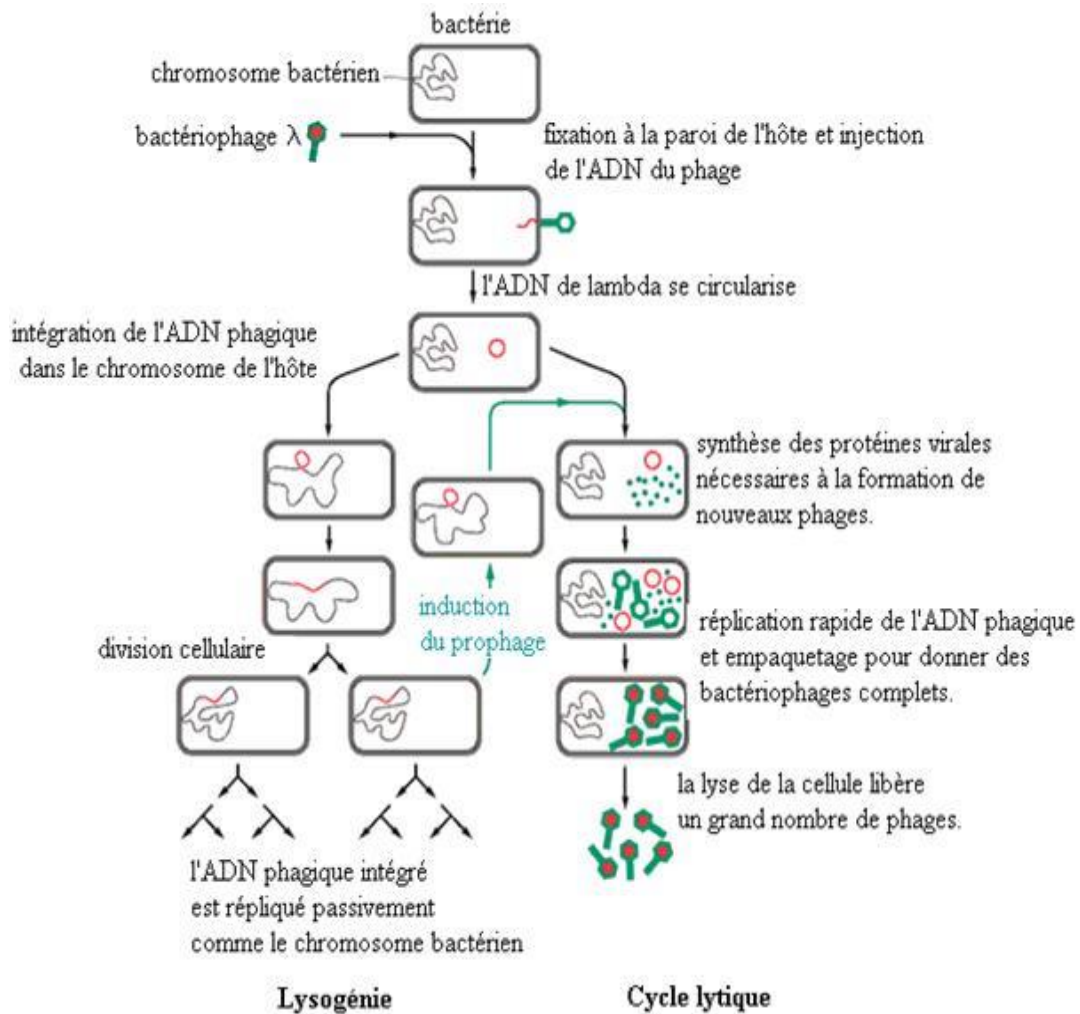


Figure 24 : Cycles de multiplication du phage lambda

IV.2.2. Préparation d'un phage pour la construction d'un recombinant

Deux stratégies sont utilisables : l'insertion simple et la délétion-remplacement, le choix dépend de la taille du DNA à introduire.

- ***La technique d'insertion simple***

Le principe est identique à celui utilisé avec les plasmides. Le phage est coupé en un site unique où le DNA à cloner est inséré. Par cette technique il est possible d'insérer jusqu'à 12Kb.

- ***La technique de délétion-remplacement***

La partie centrale du phage, non indispensable à son cycle de vie, est déletée et remplacée par le DNA à insérer, avec cette stratégie il est possible d'insérer de 8 à 22KB de DAN étranger.

V. Les cellules hôtes

En biotechnologie, le génie génétique est utilisé à des fins commerciales comme pour la production de nouveaux vaccins, de grandes quantités de protéines valorisables, ou l'introduction de gènes spécifiques dans un organisme animal ou végétale. Dans chaque cas, le choix de l'hôte est essentiel puisqu'il nous orientera vers un type de vecteur adapté.

V.1. L'hôte idéal

Pour obtenir de grande quantité d'ADN cloné, l'hôte idéal doit :

- Se développer rapidement dans un milieu de culture peu onéreux.
- Etre non pathogène.
- Être capable d'incorporer l'ADN.
- Etre stable en culture
- Posséder des enzymes appropriées pour la réplication du vecteur.

Les hôtes répondant à ces critères sont des microorganismes eucaryotes ou procaryotes dont les génomes sont bien connus car entièrement séquencés et génétiquement manipulables. Il faut noter que *E. coli* est l'organisme le plus utilisé en clonage moléculaire. Malgré que cette bactérie soit comptée parmi la flore normale de l'intestin de l'homme et des animaux, il est aussi un pathogène potentiel (surtout les souches sauvages). Aussi, *E. coli* retiens des protéines extracellulaires dans son espace périplasmique, ce qui peut rendre l'isolement et la purification des protéines recombinantes difficiles et coûteuse.

Avec *Bacillus subtilis*, l'inconvénient majeur reste la difficulté de maintenir la réplication plasmidique dans les sous cultures, ce qui engendre souvent la perte de l'ADN cloné.

Des vecteurs plasmidique et des YAC (Yeast Artificial Chromosome) ont été développés pour le clonage dans la levure *Saccharomyces cerevisiae*. L'avantage que présente cet hôte est qu'il possède les ARN et les systèmes post traductionnels complexes nécessaire à la synthèse de produits de gènes d'organismes supérieurs. Les processus post traductionnels peuvent être à l'origine de problème de clonage.

La culture de cellules de mammifères présente un coût élevé et des difficultés de production à grande échelle. En plus le niveau d'expression des gènes clonés est souvent faible (aussi pour les insectes, plantes, etc.).

On parle de transformation dans les cas des bactéries pour le processus d'intégration de l'ADN étranger, mais pour les eucaryotes on parle de transfection. Car, la transformation des cellules de mammifère désigne habituellement la conversion en cellules malignes (tumoraux, cancéreuses).

V.2. Méthodes d'introduction de l'ADN à cloner dans la cellule hôte

Il y'a plusieurs méthodes sont largement utilisées pour introduire l'ADN dans les cellules hôtes.

V.2.1. Électroporation

Cette technique implique l'exposition de l'hôte à des décharges électriques afin d'ouvrir les pores (temporairement) dans la membrane par lesquels, l'ADN cloné, ajouté dans le milieu, peut pénétrer sans lyse des cellules.

V.2.2. Un canon à Particules

La transfection des cellules cibles se fait par des billes métalliques (généralement de tungstène) recouvertes d'acides nucléiques, en perçant parois et membranes plasmiques sans provoquer de lyse cellulaire. Cette technique a été utilisée sur des levures, des algues, des cellules de plantes et même des mitochondries et chloroplastes. De plus, contrairement à l'électroporation, cette technique peut être utilisée pour introduire de l'ADN dans des tissus intacts comme des graines de plantes.

V.2.3. Microinjection

Dans les cellules animales, l'ADN peut être injecté dans le noyau par microinjection.

Chapitre II

Les banques d'ADN

Chapitre II : Les banques d'ADN

I. Définition

Une banque d'ADN est un ensemble de fragments d'ADN recombinants différents qui vont servir à former une collection de clones différents dont lesquels on recherchera le gène d'intérêt. Une Banque est caractérisée par :

- le type de vecteur
- la nature des inserts (ADN génomique, ADNc) et leur origine
- le pourcentage de vecteurs possédant un insert
- la taille moyenne des inserts
- la taille de la banque c.a.d la quantité de clones

Il existe deux grandes catégories de banques, les banques d'ADN complémentaires et les banques génomiques

I.1. Les banques ADNc

I.1.1. Définition

Un ADNc (ADN complémentaire) est une copie sous forme d'ADN d'un ARNm, et ne possède donc pas d'introns et l'ensemble de clones préparés à partir d'ARNm constitue une banque d'ADNc. Ces banques (figure 25) sont donc spécifiques du type cellulaire. Ainsi une banque de cerveau ne possédera pas d'ADNc de l'insuline. Pour être représentative une banque d'ADNc devra contenir au moins une copie de chaque ARNm présent dans la cellule d'où proviennent ces messagers.

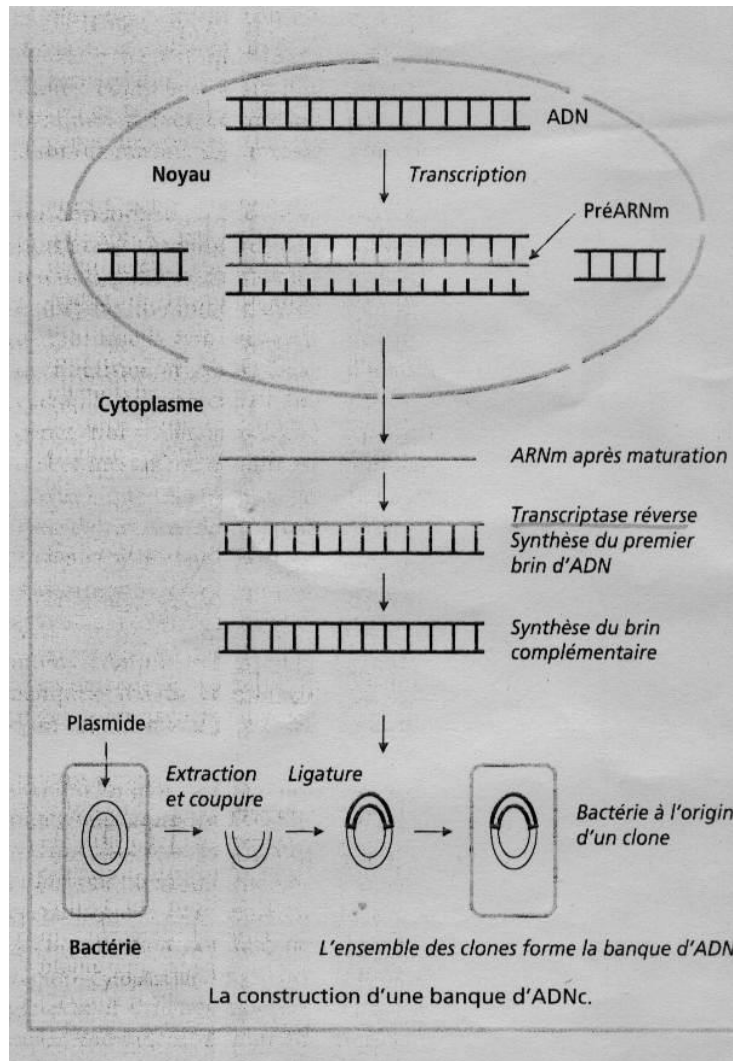


Figure 25 : construction d'une banque d'ADNc

I.1.2. Constitution d'une banque d'ADNc

1- Isolement des ARNm (figure 26)

Pour préparer l'ADNc, il faut d'abord isoler l'ARNm correspondant. Extrémité 3'OH des eucaryotes est polyadénylée ; cette queue poly A servira à séparer l'ARNm des autres ARN par chromatographie sur oligo (dT)-cellulose selon les étapes suivantes :

- En présence de NaCl 0,5 M, les queues des polyA s'hybrident de façon covalente sur un support insoluble (cellulose - oligo (dT)). Les autres ARN passent directement à travers la colonne de chromatographie
- La colonne est rincée avec un excès de solution saline pour éliminer les contaminants
- L'ARNm est récupéré par élution avec de l'eau car l'appariement entre les bases de la queue polyA et des chaînes de l'oligo (dT) n'est pas stable à faible force ionique.

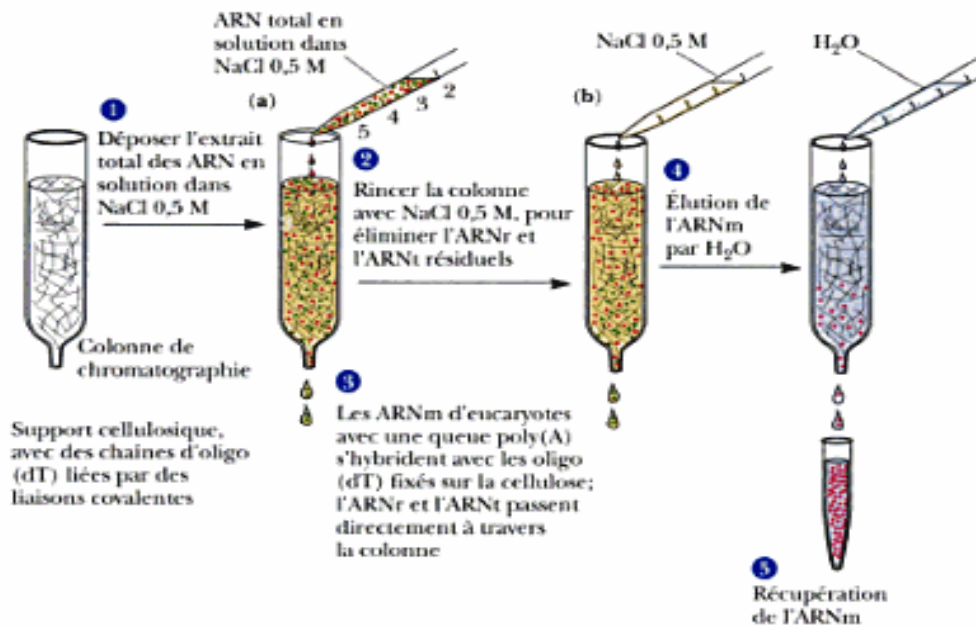


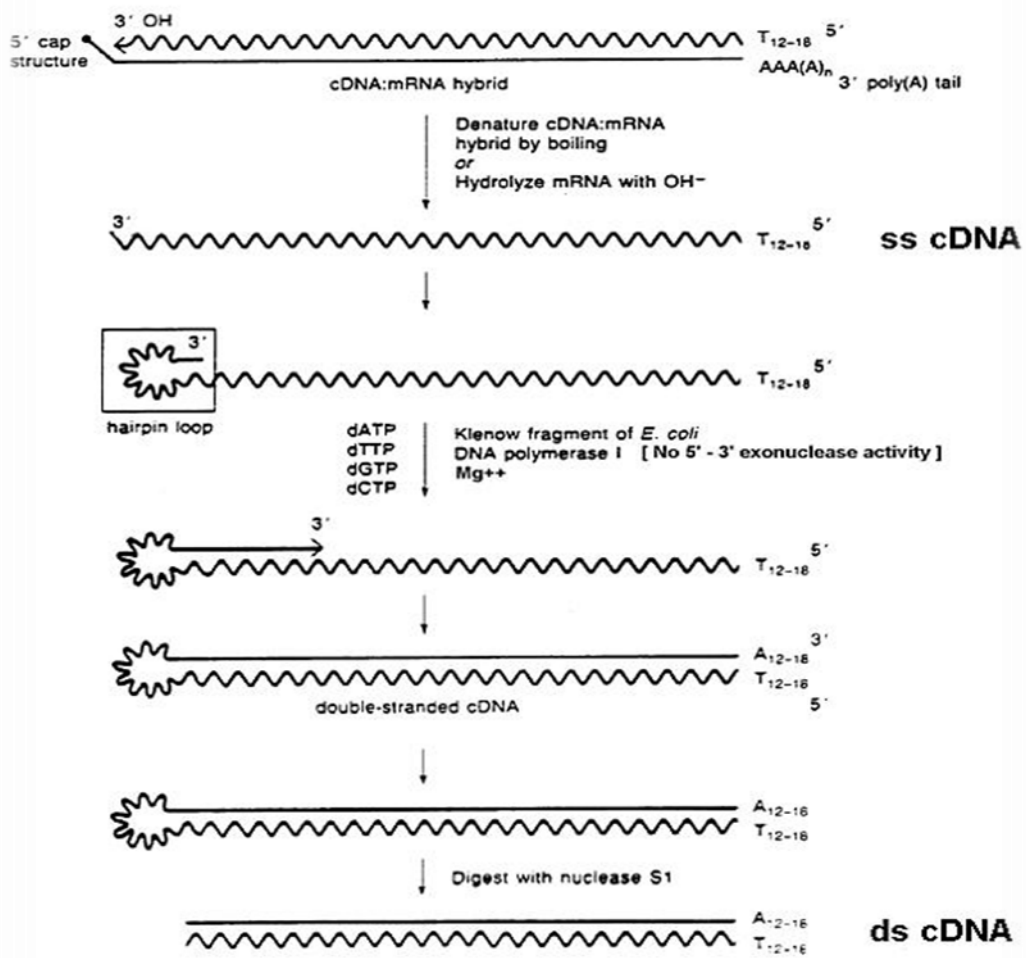
Figure 26: Isolement des ARNm

2- Synthèse de l'ADNc par la transcriptase inverse

-Technique originelle (figure 27)

Les ARNm sont incubés en présence de poly dT court, de transcriptase réverse et d'un cocktail des 4 dNTP. Le poly dT s'hybride au poly A de l'ARNm et sert d'amorce pour la transcriptase, qui synthétise alors le brin d'ADNc. L'ARNm est alors détruit par de la soude et une ARNase, ce qui permet la synthèse du brin d'ADNc par une DNA polymérase. Cette enzyme nécessite une amorce qui lui sera fournie par l'extrémité 3' OH du brin synthétisé par la transcriptase en s'hybridant avec la première séquence complémentaire qu'il trouvera sur le brin. Il en résulte la formation d'une boucle qui devra être détruite par une nucléase S1 lorsque la synthèse du second brin sera achevée afin de permettre l'intégration du DNAc dans un vecteur.

Traditional cDNA synthesis



Synthesis of double-stranded cDNA by the self-priming method.

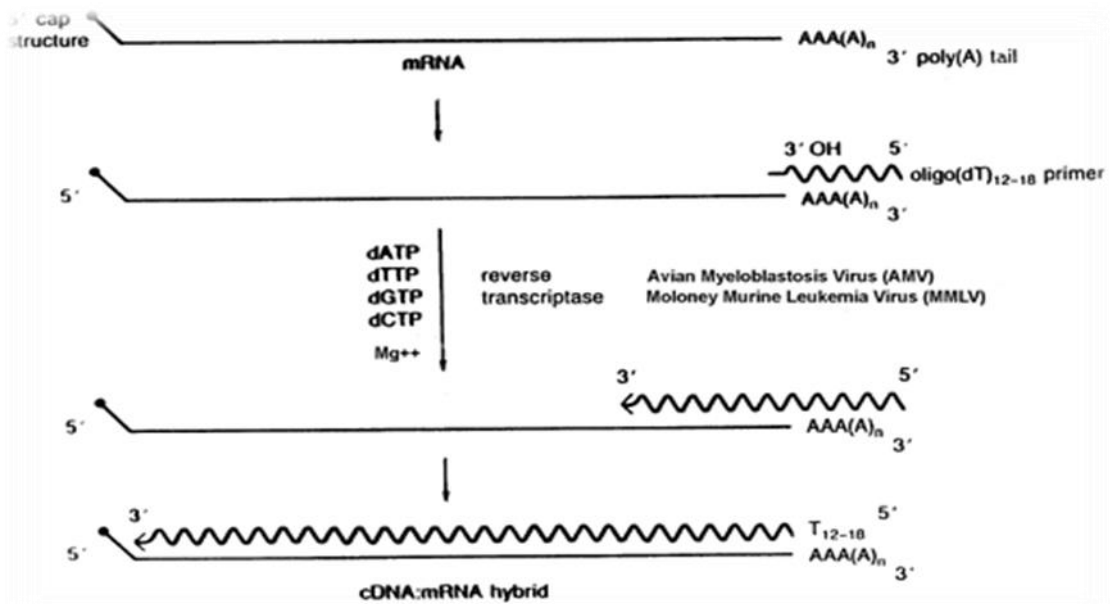


Figure 27: Synthèse de l'ADNc par la Technique originelle

-Technique par addition de queue uniforme (tailing)

Après synthèse du premier brin d'ADNc par la transcriptase inverse, une queue poly dC est créée à l'extrémité 3' de ce cDNA grâce à l'action de la terminale- transférase. L'ARN qui a servi de matrice est détruit. Un poly dG synthétisé est apporté, il s'hybride sur la queue poly dC et sert d'amorce pour la synthèse du second brin (figure 28).

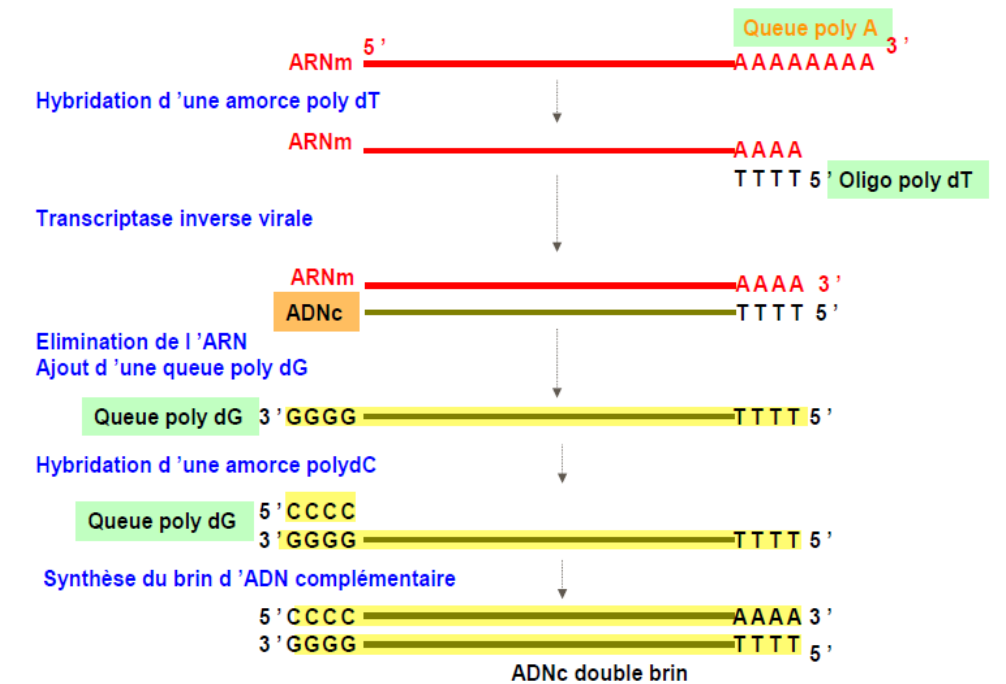


Figure 28: Technique de synthèse de l'ADNc par addition de queue uniforme (tailing)

-Technique par cassure à la RNase H

Après synthèse du premier brin d'ADNc, l'ARN n'est pas détruit en totalité mais seulement à moitié par une RNase H. Les court fragments de l'ARN restants vont servir de matrice pour la ADN polymérase I qui synthétise le brin de l'ADNc par son activité polymérase 5'–3', et détruit les restes de RNA au fur et à mesure par son activité exonucléasique 3'–5'. Les fragments qui dans le second brin d'ADN sont séparés, sont ligaturés par une ADN ligase. On obtient finalement un ADNc bicaténaire qui sera inséré dans un vecteur (figure 29).

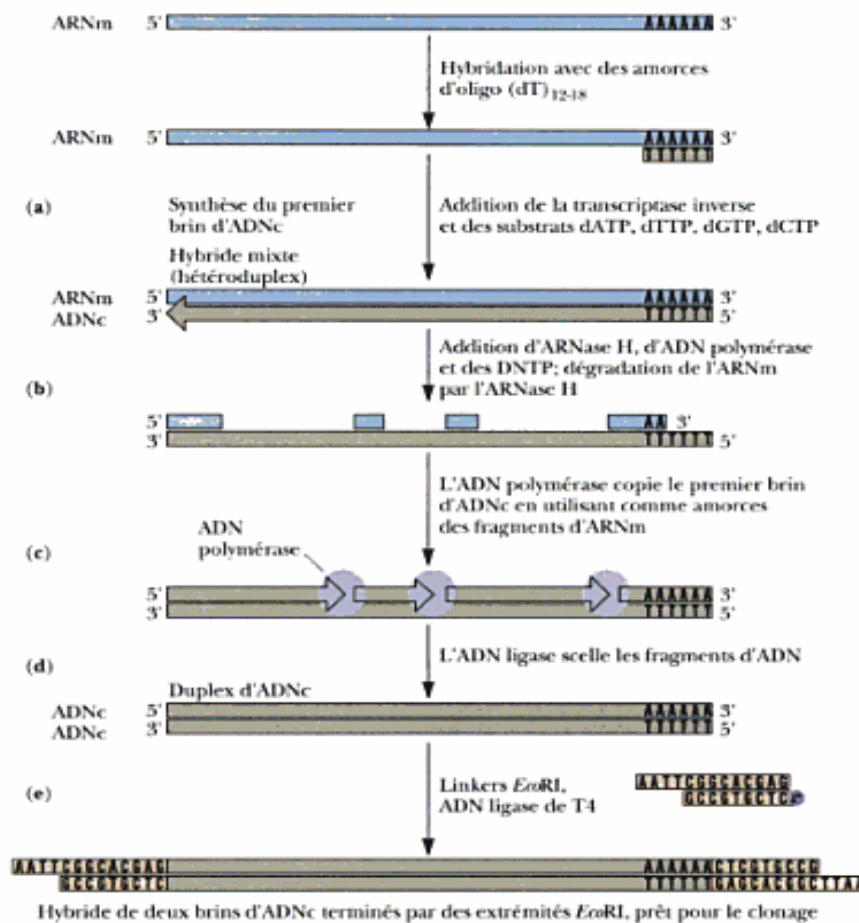


Figure 29: Technique de synthèse de l'ADNc par cassure à la RNase H

I.2. Les banques génomiques

I.2.1. Définition

Le clivage de la totalité du génome d'une cellule aboutit à un très grand nombre de fragments d'ADN qui peuvent contenir des séquences codantes ou non codantes. Ces fragments vont produire des millions de colonies différentes de cellules bactériennes transfectées. Chacune de ces colonies sera composée d'un clone dérivé d'une cellule ancestrale unique et hébergera par conséquent un plasmide recombinant avec la même séquence intégrée d'ADN génomique. La collection complète de ces plasmides constitue une banque d'ADN génomique.

I.2. 2. Constitution d'une banque génomique

Après extraction de l'ADN génomique, on fait agir des enzymes de restriction sur le mélange de cet ADN et du vecteur. Ce traitement va permettre l'ouverture des deux ADN et la formation d'extrémités collantes complémentaires. Les rencontres des fragments d'ADN plasmidique et génomique se feront au hasard aboutissant à la formation de nouveaux plasmides un peu plus grands et portant des fragments d'ADN génomique. Ces plasmides recombinants sont ensuite réintroduits dans des bactéries afin de les cloner en grand nombre (figure 30).

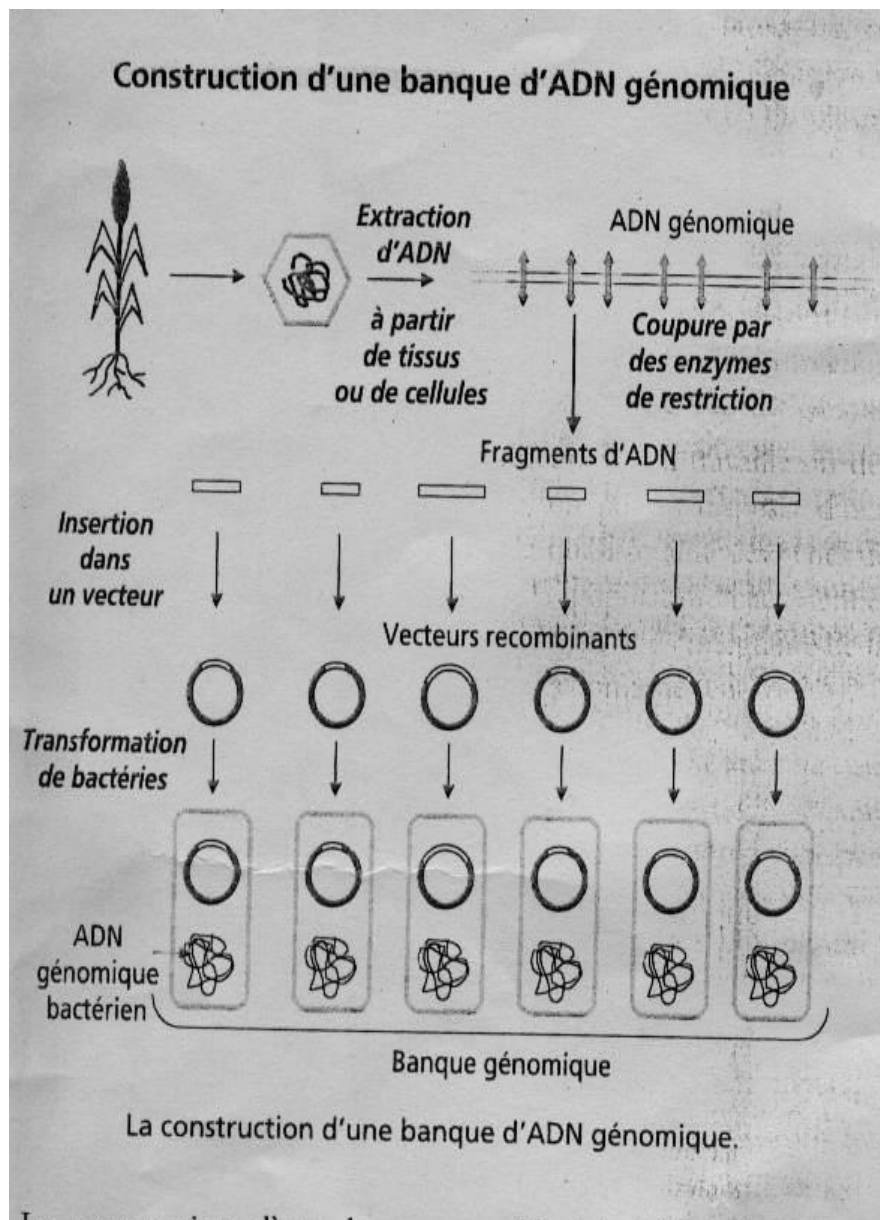


Figure 30: Construction d'une banque d'ADN génomique

I.3. Le choix entre une banque d'ADN génomique et une banque d'ADNc

Le choix dépend de la situation ; si l'on recherche un gène spécifique, actif dans un type précis de tissus chez une plante ou un animal alors il est astucieux d'utiliser ce tissu pour en extraire ses ARNm et les convertir en ADNc pour fabriquer une banque d'ADNc. En revanche, il faut fabriquer une banque d'ADN génomique si l'on veut étudier des introns, des séquences régulatrices ou encore des gènes dont la localisation tissulaire est inconnue.

II. Criblage des clones recombinants

Le criblage consiste à distinguer les cellules ayant intégré des vecteurs comportant la séquence d'ADN recherchée du reste des cellules.

II.1. Criblage grâce à une sonde d'acide nucléique

C'est une des techniques communément employées est l'hybridation de l'ADN des cellules avec des sondes nucléotidiques marquées (figure 31).

Une réplique, est réalisée par transfert de quelques cellules sur filtre de nitrocellulose. Les filtres sont traités avec de la soude pour casser les cellules et séparer les brins de leurs molécules d'ADN. La sonde nucléotidique marquée radioactivement est placée au contact du filtre permettant ainsi l'hybridation avec la séquence complémentaire de l'ADN recherché. La sonde non hybridée est éliminée par des lavages répétés et les colonies portant la séquence d'intérêt sont déterminées par autoradiographie.

II.1.1. Préparation d'une sonde nucléaire

Une sonde peut être un fragment d'ADN génomique, un fragment d'ADNc, une molécule d'ARN (ribosonde) ou un oligonucléotide de synthèse. Pour permettre la détection de séquences nucléaires par appariement, il faut que la sonde soit simple brins.

Pour qu'un oligonucléotide puisse servir de sonde, il doit être suffisamment long (environ 20 nucléotides) pour que sa séquence existe exclusivement dans le clone recherché. Si on connaît une fraction de la séquence d'acides aminés de la protéine, alors on peut synthétiser une sonde d'ADN correspondant à une petite région du gène d'après le code génétique. Cependant comme le code génétique est dégénéré, une sonde basée sur la séquence d'acides aminés doit comprendre tous les oligonucléotides possibles qui pourraient théoriquement coder cette séquence peptidique. L'un des oligonucléotides s'hybridera parfaitement avec le clone recherché. Le marquage des sondes peut se faire par plusieurs techniques :

a) Marquage par Nick translation (translation de coupure) (figure 32)

Cette technique permet d'incorporer plusieurs atomes de P^{32} à partir de ^{32}P désoxyribonucléotide triphosphate (dNTP). L'ADN double brin est d'abord traité à l'ADNase I qui effectue des coupures de manière aléatoires. Dans un deuxième temps, l'ADN polymérase I est utilisée pour dégrader l'ADN grâce à son activité exonucléasique 3'-5', et reforme le brin d'ADN grâce à son activité polymérasique (5'-3') en incorporant des dNTP radioactifs.

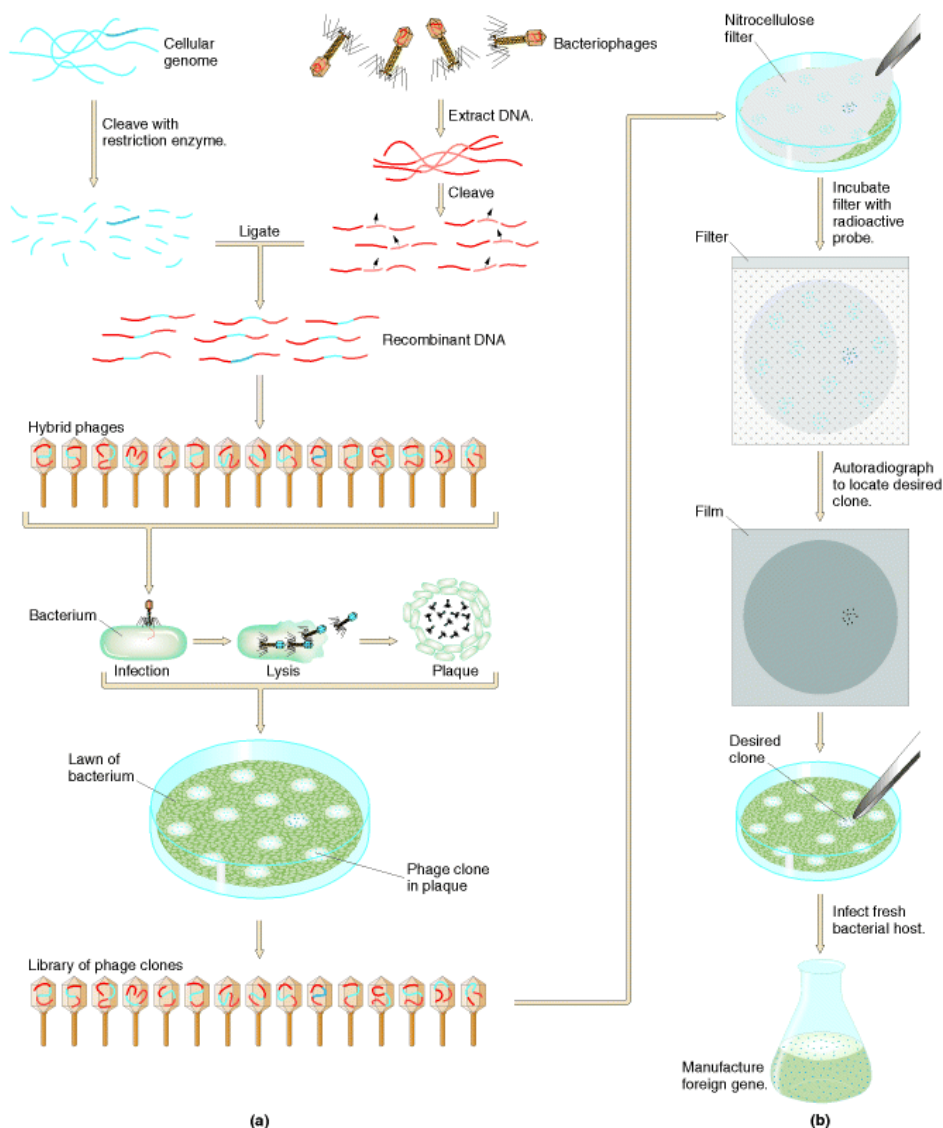


Figure 31: Criblage par des sondes nucléaires

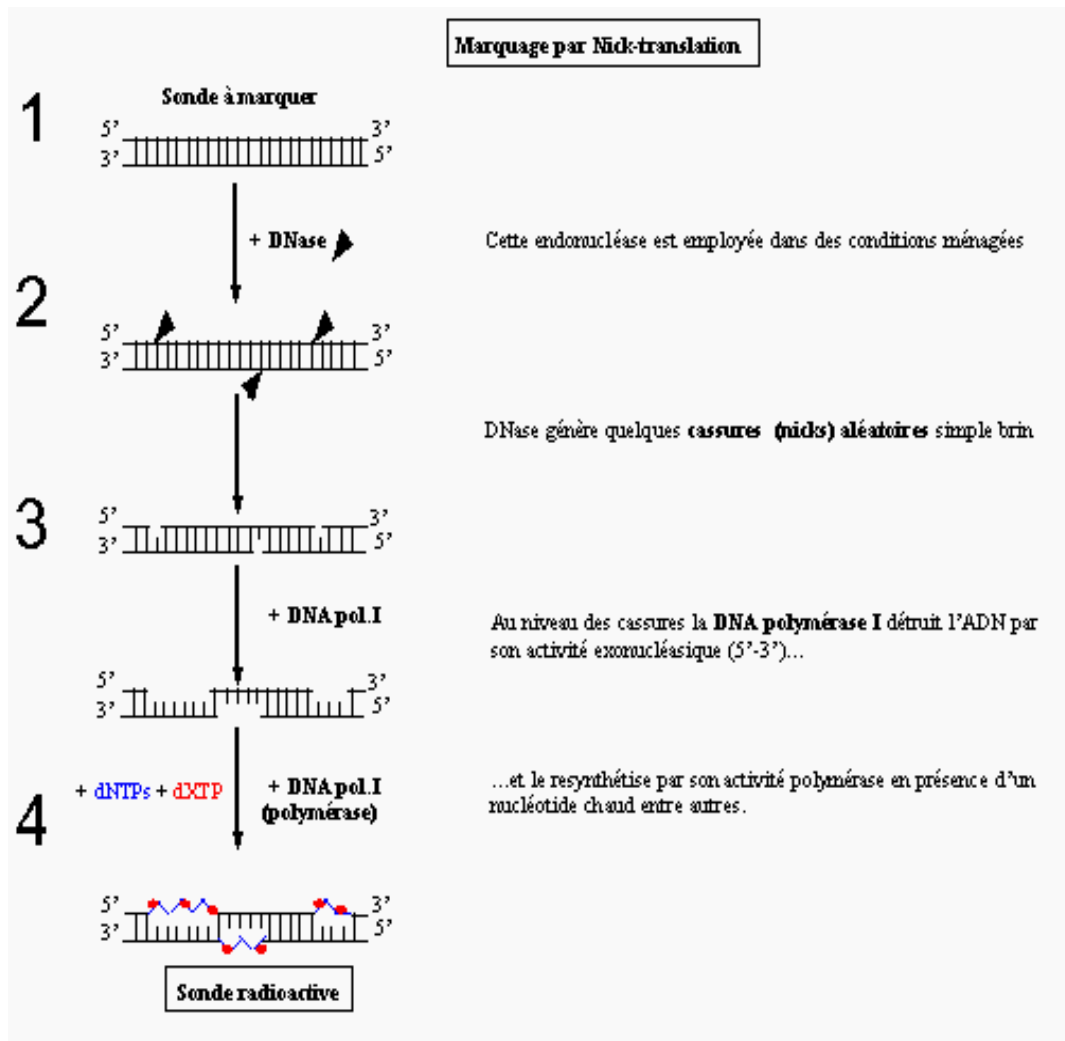


Figure 32: Marquage par Nick translation (translation de coupure)

b) Marquage par Random priming (amorçage au hasard) (figure 33)

Partant d'un ADN simple brin ou bien d'un ADN double brin auquel on a fait subir une dénaturation par chauffage, on ajoute un mélange d'héxanucléotides synthétiques dont certains vont s'hybrider au hasard mais spécifiquement avec la sonde, pour servir d'amorce. La synthèse du brin amorcé est effectuée par une polymérase en présence de dNTP radioactif.

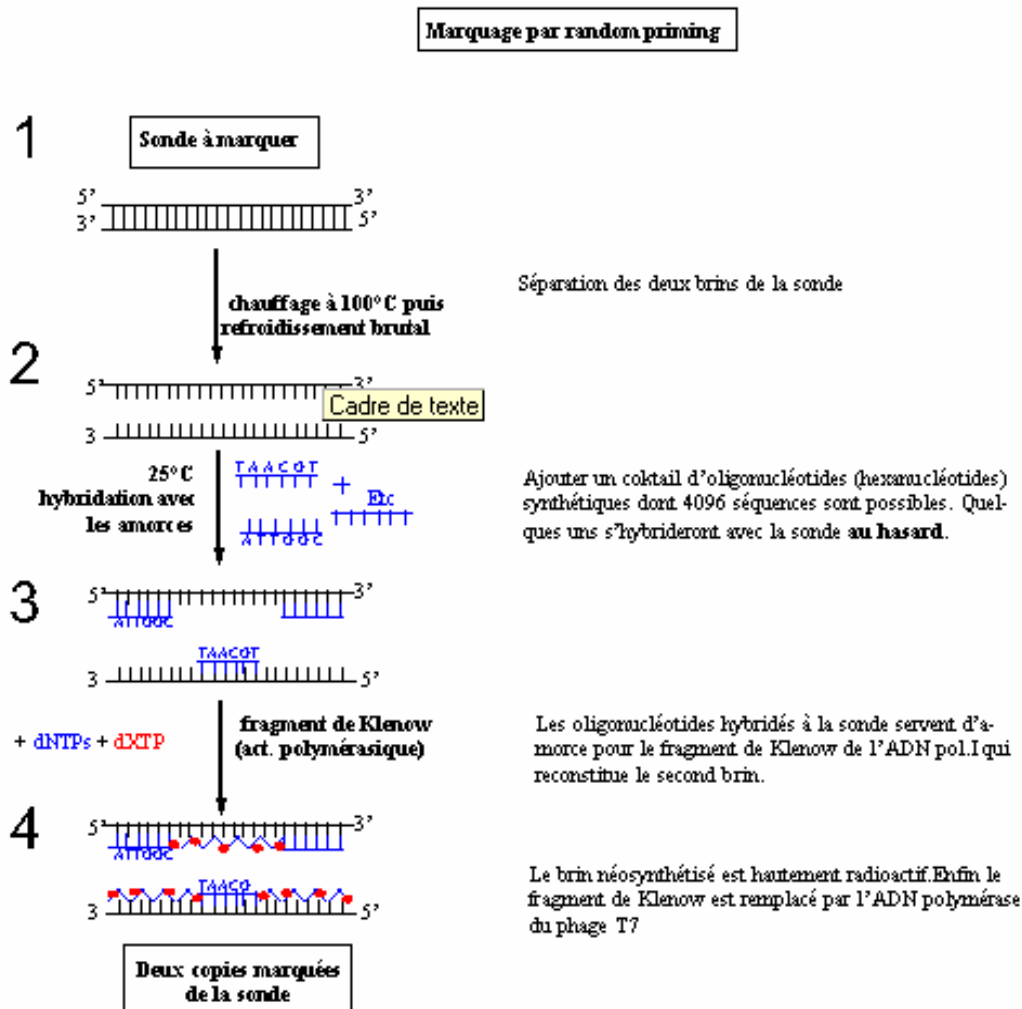


Figure 33: Marquage par Random priming (amorçage au hasard)

c) Marquage des oligonucléotides de synthèse en 5'

Les oligonucléotides sont synthétisés chimiquement puis marqué radioactivement en utilisant une kinase pour transférer un groupement phosphate marqué au P^{32} de l'ATP vers l'extrémité 5' de chaque oligonucleotide. Ces nucléotides peuvent être marqués avec des molécules non radioactives telles que les fluorochromes qui émettront une lumière mesurable de longueur d'onde donnée.

II.2. Criblage par des anticorps

Parfois aucune connaissance sur la séquence de la protéine n'est disponible. Des anticorps spécifique à la protéine étudiée peuvent être utilisés pour rechercher les bons recombinants portant le gène responsable de cette protéine. Ceci suppose que la banque d'ADN a été construite dans un vecteur d'expression : le site d'insertion des ADN dans le vecteur se trouve en aval d'un promoteur bactérien qui est le plus souvent le promoteur de l'opéron lactose. Ainsi toutes les bactéries contenant ces vecteurs recombinants sont capables de synthétiser la protéine correspondant à la séquence d'ADN d'intérêt.

II.2.1. Etapes du criblage par anticorps (figure 34)

Comme dans les méthodes de criblage par des sondes nucléaire, les colonies de la banque sont répliquées sur un filtre de nitrocellulose. Cette membrane est a son tour mise à incuber, dans des conditions qui induisent la synthèse de la protéine correspondant à l'insert d'intérêt, puis les cellules sont traitées de façon à libérer la protéine synthétisée. La protéine se fixe fermement au filtre qui est ensuite mis en contact avec des anticorps primaires dirigés contre la protéine. Après élimination des anticorps non fixés, le complexe stable protéine-anticorps est ensuite révélé par un autre anticorps secondaire radioactif qui se lie à l'anticorps primaire. Cette radioactivité est mise en évidence par un film d'autoradiographie

- ***Remarque concernant les techniques de criblage***

Le bon clone ne peut pas être récupéré du premier coup car lors du premier criblage les colonies sont très rapprochées ; la précision du repérage est donc très faible. Il est absolument indispensable de réétaler les clones sélectionnés à plus faible concentration et de renouveler le criblage. Pour les plasmides et les cosmides au moins deux tours sont nécessaires ; pour les phages, il en faut au moins trois.

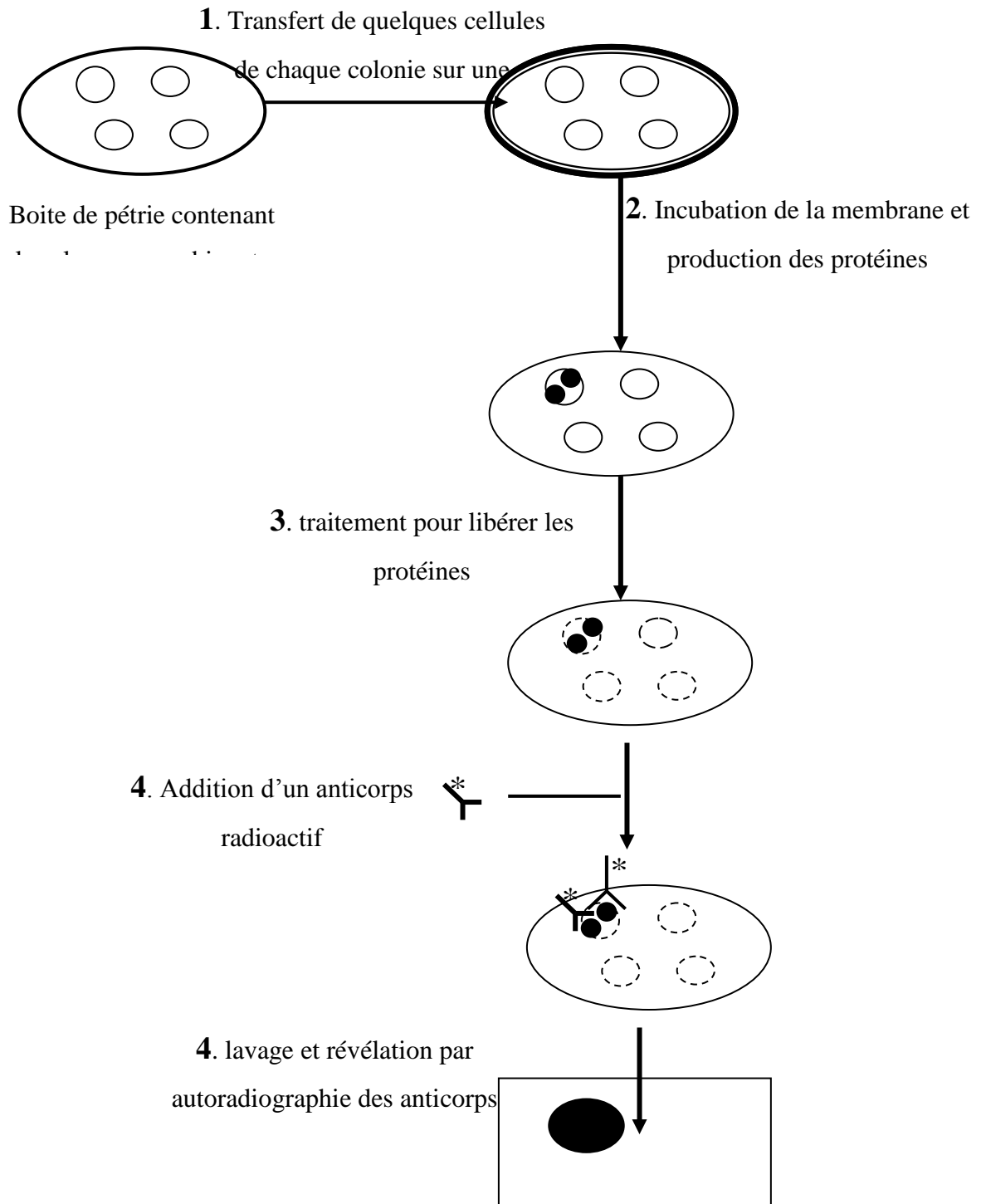


Figure 34: Technique de criblage par des anticorps

Chapitre III

Méthodes d'analyses du gène purifié

Chapitre III : Méthodes d'analyse du gène purifié

I. Détection des fragments d'ADN et d'ARN par des techniques de transfert

Deux techniques très efficaces pour détecter une séquence particulière d'ADN ou d'ARN dans un mélange complexe combinent la séparation par électrophorèse sur gel et l'hybridation à l'aide d'une sonde complémentaire marquée radioactivement.

I.1. Le transfert de type Southern blot

L'une des techniques majeures dans l'analyse d'un gène est la technique d'hybridation selon Southern (southern blot) (figure 35).

Il suffit de posséder une sonde s'hybridant spécifiquement avec la séquence que l'on veut étudier. L'ADN génomique est digéré par des enzymes de restriction et les fragments sont séparés selon leur taille par électrophorèse en gel d'agarose. Le gel est ensuite plongé dans une solution alcaline pour rendre l'ADN sous forme monobrin. Afin de permettre l'hybridation, les fragments d'ADN sont transférés sur une membrane de nitrocellulose (blotting). La membrane est ensuite immergée dans une solution contenant la sonde marquée radioactivement. Après lavage de la membrane, les fragments hybridés avec la sonde sont révélés par autoradiographie.

Grace à cette technique, il est possible de réaliser une carte de restriction d'un fragment d'ADN, de détecter les délétions et remaniements.

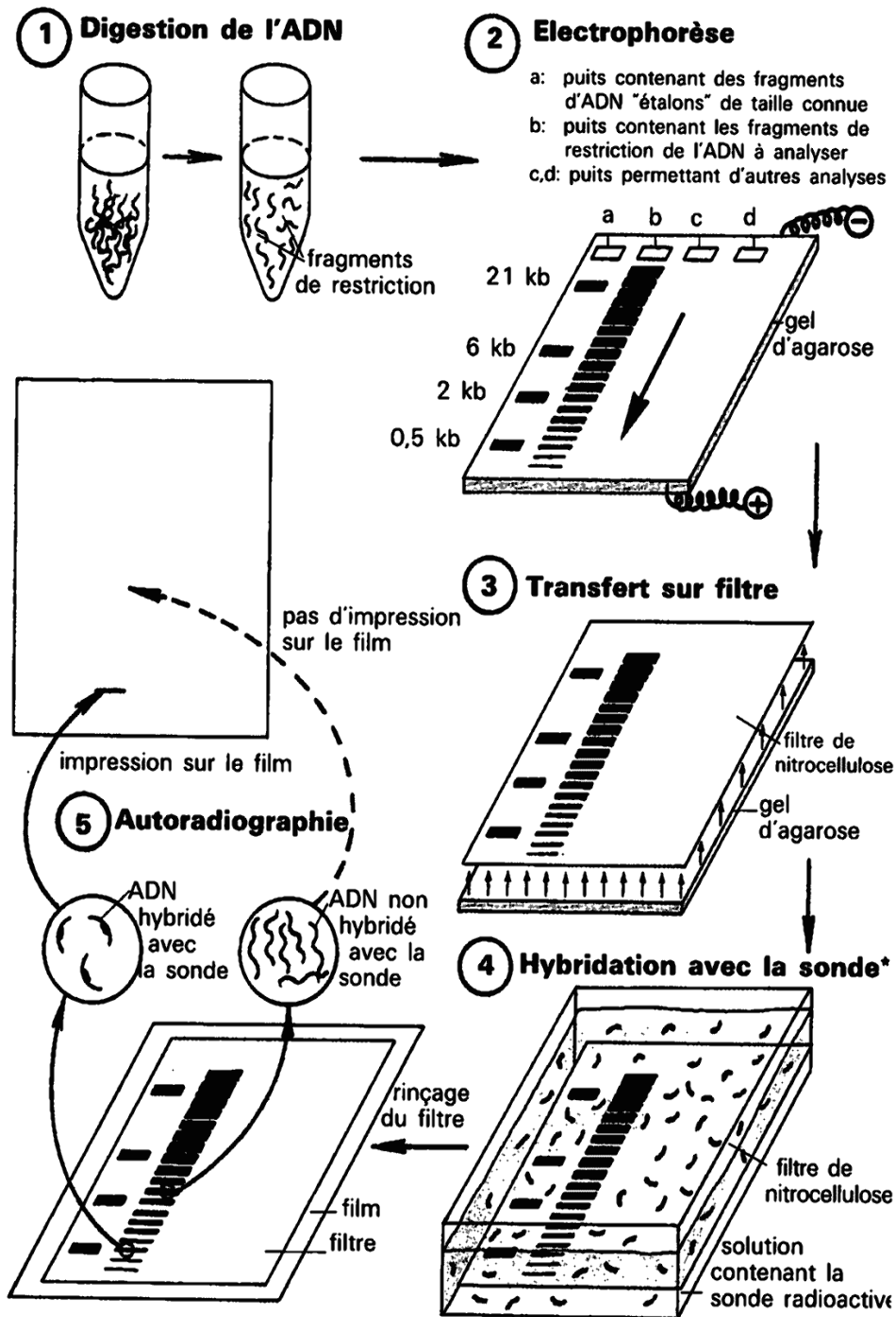
I.2. Le transfert de type Northern Blot

Il s'agit d'une technique analogue à celle décrite par southern pour l'ADN, appelée de manière humoriste (Northern) (figure 36).

Les ARNm sont séparés selon leur taille par migration dans un gel d'agarose contenant un agent dénaturant (formaldéhyde) de manière à détruire les éventuelles structures secondaires de l'ARN (telles que les liaisons hydrogène entre les paires de bases). Ainsi les molécules d'ARN se retrouveront sous leur forme linéaire, non repliée, pour la migration. Les étapes de transfert et d'hybridation sont tout à fait similaires à celles utilisées pour l'ADN.

Cette technique permet de :

- apprécier sa distribution dans les tissus, étudier son abondance relative
- déterminer sa taille
- détecter les intermédiaires de maturation



*sonde moléculaire : courte séquence de nucléotides obtenue de diverses manières et capable de s'hybrider de façon stable avec les séquences d'ADN qui lui sont complémentaires

Figure 35: Technique de southern blot

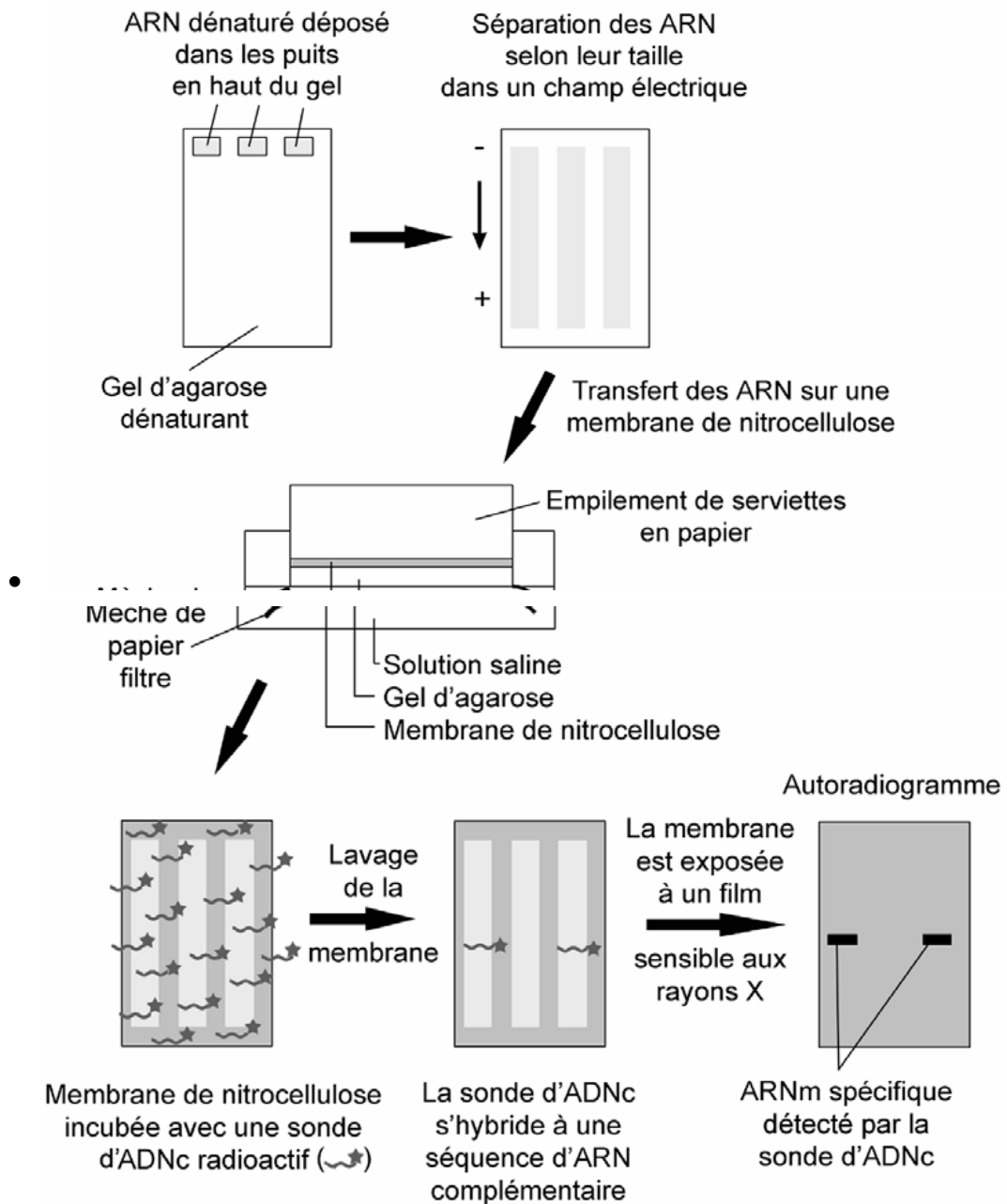


Figure 36 : La technique de Northern blot.

I.3. Mise en évidence et quantification d'une protéine

La mise en évidence d'une protéine peut se faire par la technique de l'immunoblot ou western blot. L'extrait protéique contenant la protéine à caractériser est déposé sur un gel de polyacrylamide contenant du SDS. Après migration, les protéines sont transférées sur une membrane et mise en présence de l'anticorps spécifique marqué radioactivement. L'autoradiographie révélera une bande caractéristique de la protéine.

Le dosage quantitatif est souvent effectué par la méthode d'ELISA (enzyme-linked immunosorbant assay) (dosage d'immunosorption liée à enzyme). Un premier anticorps spécifique de l'antigène est absorbé sur un support solide. La solution contenant la protéine à doser (antigène) est mise en contact avec l'AC. Après lavage, un 2eme AC couplé à une enzyme est mis en contact avec le premier anticorps fixé sur l'antigène, ensuite le substrat de l'enzyme est apporté qui une fois fixé sur l'enzyme prendra une couleur spécifique. La quantité de produit coloré obtenue est proportionnelle à la quantité d'antigène.

I.4. Le transfert de type dot blot

Le terme *dot blot* regroupe un ensemble de techniques de transfert de molécules permettant de vérifier la présence de molécules spécifiques dans un milieu, elle représente une simplification des technique de Southern blot, northern blot et western blot. En effet dans cette technique, on ne procède pas a une séparation par électrophorèse, la détection se fait directement sur la membrane grâce à des sondes nucléaire pour le southern et le northern et avec des anticorps marqué dans le cas de western blot. La technique de Dot blot est rapide mais présente l'inconvénient de ne pas donner d'information concernant la taille des biomolécules, elle permet uniquement de mettre en évidence la présence ou l'absence de biomolécules.

I.5. Hybridation in situ

Cette technique diffère des autres techniques classiques d'hybridation car elle est utilisée pour détecter les séquences d'acides nucléiques qui sont présents dans des cellules intactes. Par exemple, on peut l'utiliser pour montrer que l'ARNm de l'insuline est produite dans les seules cellules du pancréas.

De fines tranches de tissu (coupes histologique) sont placées sur une lame de microscope. On ajoute la sonde aux cellules de la lame et on la fait pénétrer dans le cytoplasme ou elle se fixe aux ARNm cibles pour former des hybrides. Ceux ci peuvent être détectés par le marquage de la sonde radioactivement et révélés par autoradiographie.

I.6. Les puces à ADN (puces à gènes ou biopuces)

I.6.1.Principe de fonctionnement

Le principe de la puce à ADN est basé sur la technique d' hybridation . Immobilisés sur un support solide (matrice), des oligonucléotides (simples brins) spécifiques de différents gènes

connus constituent les sondes dont le rôle est de détecter des cibles marquées complémentaires, présentes dans le mélange complexe à analyser (ARNm extraits de cellules, tissus ou organismes entiers et convertis en ADNc).

Les signaux d'hybridation sont détectés selon le type de marquage, radioactivité ou fluorescence, par mesure radiographique ou par fluorescence (figure 37).

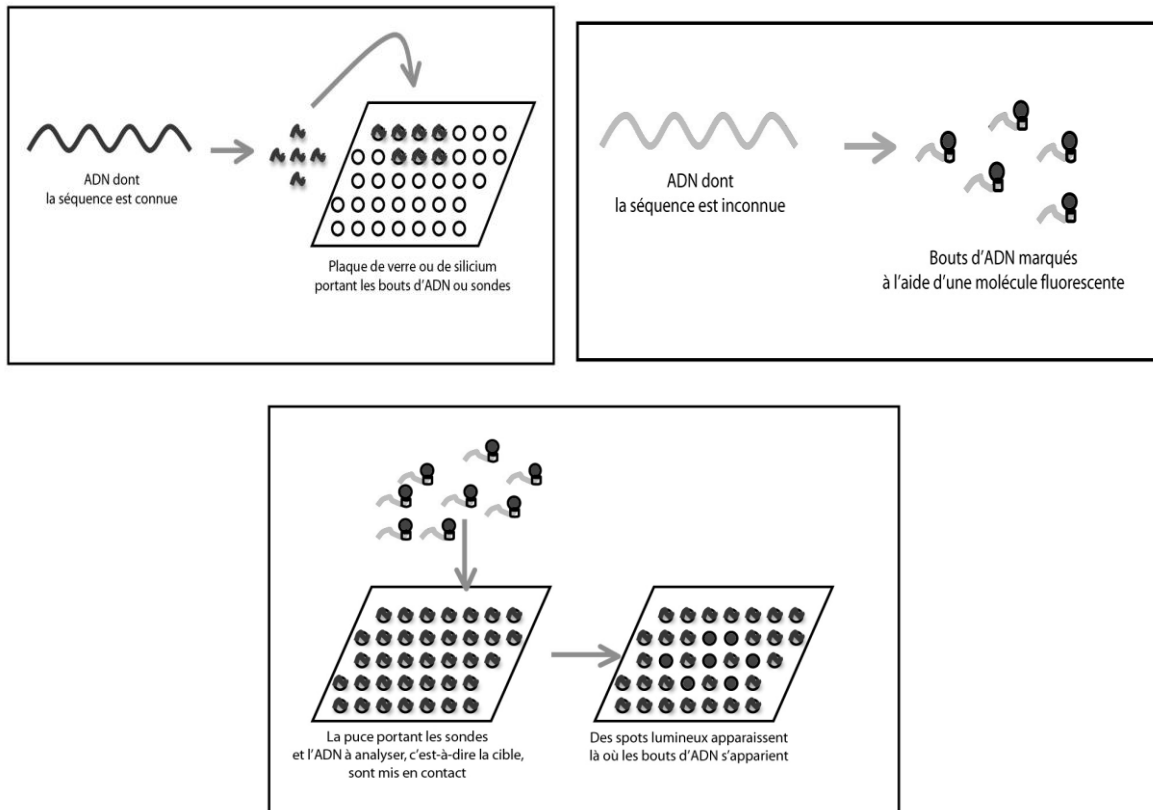


Figure 37: Les puces à ADN

I.6.2. Le support

Le support sur lesquels sont fixées les sondes se présente sous la forme d'une surface (matrice) inférieure à 1cm², plane ou poreuse (percées de puits) et composés de matériaux tels que le verre, le silicium ou l'or et le platine.

I.6.3. L'hybridation sonde-cible

Une séquence d'ADN cible est marquée par fluorescence et mise en contact avec la biopuce. Le rôle de chaque sonde est de reconnaître, dans le mélange appliqué à la surface de la biopuce, cette séquence d'ADN cible. La phase d'hybridation est réalisée et suivie d'un lavage destiné à débarrasser la puce des cibles nucléiques non hybridées.

I.6.4. La détection des hybridations

Il s'agit ici de déterminer les sondes qui ont été effectivement hybridées, c'est à dire de repérer à quelles adresses les sondes sont complémentaires des cibles de l'échantillon test. Une lecture optique permet de révéler les sondes d'ADN devenues fluorescentes (hybridées avec les cibles marquées).

I.6.5. Utilisation des biopuces

Cette technique permet de déterminer la composition d'un ADN et de savoir si un patient est porteur d'une maladie ou non en testant l'hybridation des brins d'ARN extraits du patient (et rétrotranscrits en ADN) avec les brins d'ADN synthétiques représentatifs de la maladie se trouvant au niveau de la puce à ADN. Avant d'être déposé sur la puce, l'ADN doit être marqué.

II. Séquençage des acides nucléiques

II.1. Méthode enzymatique de Sanger

II.1.1.Principe

Le principe de cette technique consiste à effectuer la synthèse de multiples brins d'ADN complémentaires du brin dont on veut déterminer la séquence. L'astuce consiste à utiliser des didésoxynucléotides (ddNTP), analogues structuraux des dNTP mais ne possède pas de gpt OH en 3'. L'extension de la chaîne est stoppée après incorporation d'un ddNTP du fait de l'impossibilité de former la liaison phosphodiéster.

II.1.2.Procédé

Quatre réactions sont menées en parallèle contenant toutes l'ADN à séquencer sous forme monobrin, les 4 dNTP, une amorce marquée et une ADN polymérase. Chacun des tubes contient en outre un ddNTP particulier à faible concentration. Pour chaque tube, on obtient un ensemble de molécules dont la longueur dépend de la position du nucléotide complémentaire du ddNTP introduit. Le contenu de chacun des tubes est déposé sur un gel de polyacrylamide après dénaturation. Ce dernier est soumis à une autoradiographie. La séquence du brin néoformé d'ADN se lie horizontalement de l'anode (+) (correspondant au côté 5'P) vers la cathode (-) (correspondant au côté 3'OH). Concernant la séquence du brin complémentaire, il est facile d'en déduire la séquence du brin matrice (figure 38).

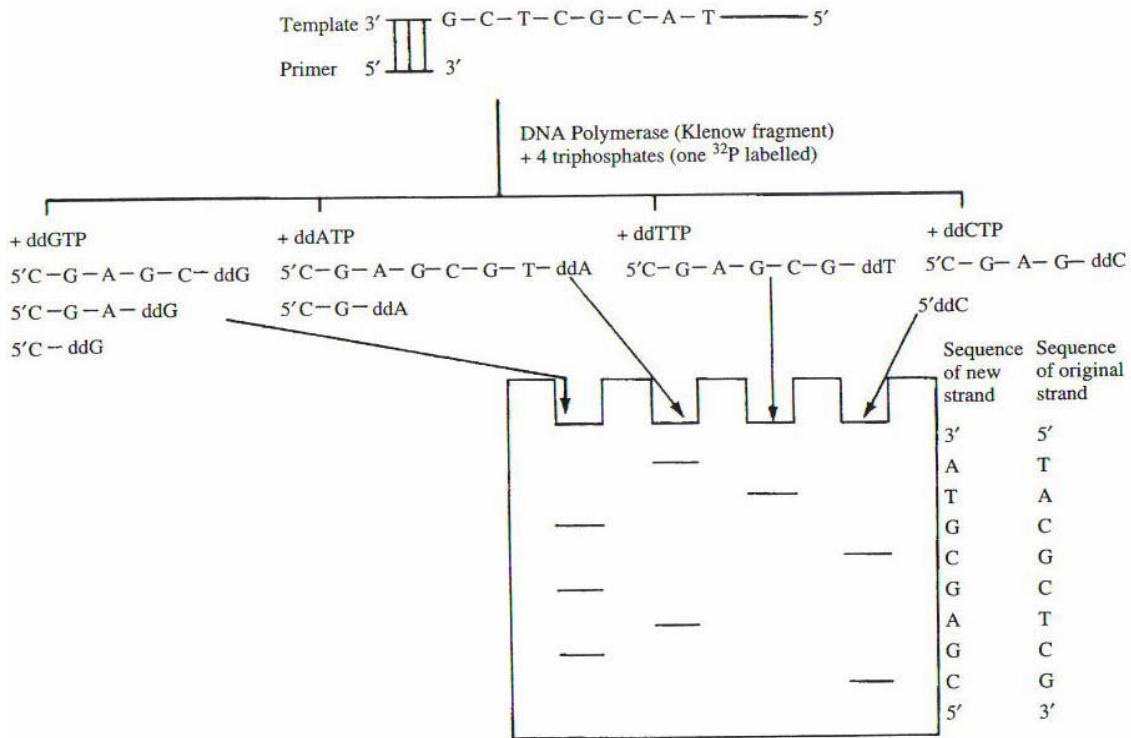


Figure 38 : Méthode enzymatique de Sanger

II.1.2. Automatisation du Séquençage de Sanger

Le séquençage est de plus en plus automatisé. Le principe reste le même mais le marquage et le mode de détection est différent. Ce sont les ddNTP qui sont marqués avec des fluorochromes différents, par exemple un marqueur fluorescent pour A, un bleu pour T, un vert pour G et un jaune pour C. A la fin de la réaction, les contenus des quatre tubes sont réunis et déposés en un même point haut du gel d'électrophorèse. A mesure que les oligonucléotides atteignent le bas du gel, ils sont illuminés par un rayon laser. La fluorescence de chaque colorant est détectée par des capteurs reliés à un ordinateur. Les résultats apparaissent alors sous forme de graphe de couleur correspondant à chaque nucléotide. Un logiciel traduit alors le pictogramme en séquences.

II.2. Méthode chimique de Maxam et Gilber

Cette méthode utilise de l'ADN monocaténaire marqué sur l'une de ces extrémités. Ce dernier est ensuite réparti dans 4 tubes et chacun d'eux est traité avec un agent chimique qui modifie une ou deux des quatre bases. Il y a une réaction spécifique pour G et une réaction spécifique pour les purines qui éliminé à la fois A et G. De même pour les bases pyrimidiques, il y a une réaction spécifique pour C et T qui devient spécifique pour C quant elle est effectuée en présence de NaCl concentré.

La modification des bases est telle que leur liaison à l'ose est clivée, une réaction à la pipéridine provoque ensuite l'excision du désoxyribose de la chaîne du squelette et donc la libération d'un fragment marqué et un autre non marqué. Vu le nombre important d'ADN dans chaque tube, les produits finaux de chaque réaction sont une collection de fragments de différentes tailles qui seront soumises en parallèle à une électrophorèse. Le gel est ensuite autoradiographié et le profil des bandes sur le film est lu pour déterminer la séquence du fragment d'ADN (figure 39).

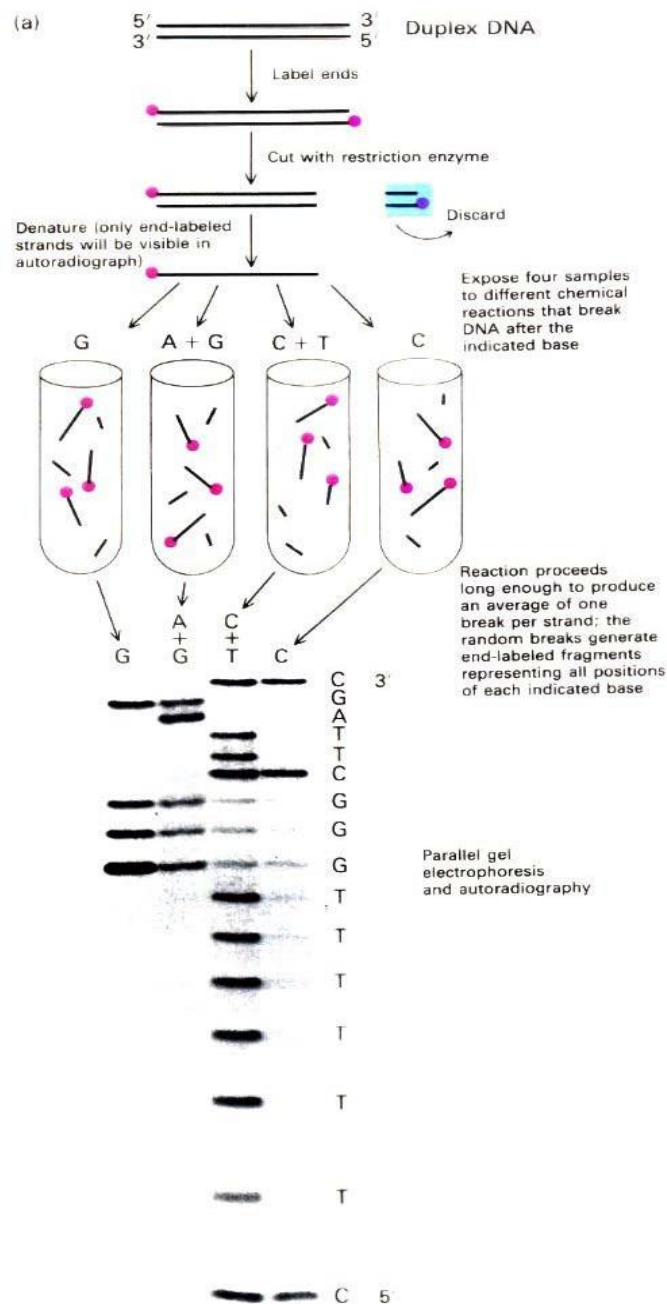


Figure 39 : Séquençage de l'ADN selon Maxam et Gilbert

II.3. Séquençage par hybridation ou séquencing by hybridization (SBH)

II.3.1. Principe

Le principe de séquençage par hybridation est basé sur l'hybridation du fragment à séquencer avec des sondes fixées sur un support (puce),

II.3.2. Procédé

Une série d'oligonucléotides de 8 bp, appelée sonde, représentant toutes les combinaisons possibles, est synthétisée (4^8 oligonucléotides différentes sont utilisées). Ces oligonucléotides sont fixés de manière ordonnée sur des supports pour constituer une puce

Le fragment à séquencer est amplifié par PCR puis marqué à la fluorescence. Il est ensuite mis en contact avec la puce portant plusieurs sondes de séquence connues, après l'étape d'hybridation et de lavage, la lecture de la puce est réalisée par des systèmes de détection de fluorescence. Le traitement informatique des données est ensuite réalisé par différents logiciels.

La séquence X, dont on cherche à déterminer la composition par la technique de séquençage par hybridation, sera hybridée au niveau de différentes sondes d'une manière chevauchante et on pourra alors reconstituer la séquence du brin analysé.

II.3.3. Limites de la technique

- Présence de bruits de fond qui sont une conséquence de l'hybridation non spécifique. Pour y remédier, des sondes plus courtes sont utilisées.
- Formation de structures secondaires en boucles et en épingle à cheveux à l'intérieur de l'ADN simple brin à séquencé.

II.4. Méthode du pyroséquençage

Le pyroséquençage est une technique qui permet d'effectuer un séquençage rapide et à moindre coût qu'un séquençage par la méthode de Sanger. Elle permet une lecture directe de la séquence obtenue après le séquençage.

Ici, les nucléotides ne sont pas ajoutés tous ensemble comme dans une réaction de séquençage normale mais l'un après l'autre. Si le Nucléotide ajouté dans le milieu réactionnel correspond à celui attendu par la Polymérase, il est incorporé dans le brin et libère un Pyrophosphate (PPi). Une ATP sulfurylase vient alors transformer ce dernier en ATP qui est alors utilisé, couplé à une Luciférine, par une Luciférase. On a alors production

d'Oxyluciférine et d'un signal Lumineux (une Apyrase dégrade les nucléotides en surplus). C'est le séquenceur qui va capter ce signal lumineux et le reproduire sous forme d'un pic sur le Pyrogramme. La hauteur de ce pic est fonction de l'intensité du signal lumineux, elle-même proportionnelle au nombre de nucléotides incorporés en même temps. On peut donc déduire la séquence à partir de la taille des pics obtenus.

Travaux dirigés

Série TD 1

Exercice I : Représenter les extrémités des fragments de restriction obtenus après digestion par les enzymes de restrictions suivantes en les classant selon la nature de ces extrémités.

- *Bam*H1 G/GATCC
- *Hpa*II C/CGG
- *Sma*I CCC/GGG

Exercice II : La digestion d'un ADN linéaire de 900pb avec trois enzymes de restriction et avec une combinaison de ces enzymes a donné les résultats suivants :

<i>Enzyme</i>	<i>Taille du fragment obtenu</i>
<i>Eco</i> R1	200 pb, 700pb
<i>Hind</i> III	300pb, 600pb
<i>Bam</i> H1	50pb, 350pb, 500pb
<i>Eco</i> R1 + <i>Hind</i> III	100pb, 200pb, 600pb
<i>Eco</i> R1+ <i>Bam</i> H1	50pb, 150pb, 200pb, 500pb
<i>Hind</i> III + <i>Bam</i> H1	50pb, 100pb, 250pb, 500pb

- Donner la carte de restriction de cet ADN

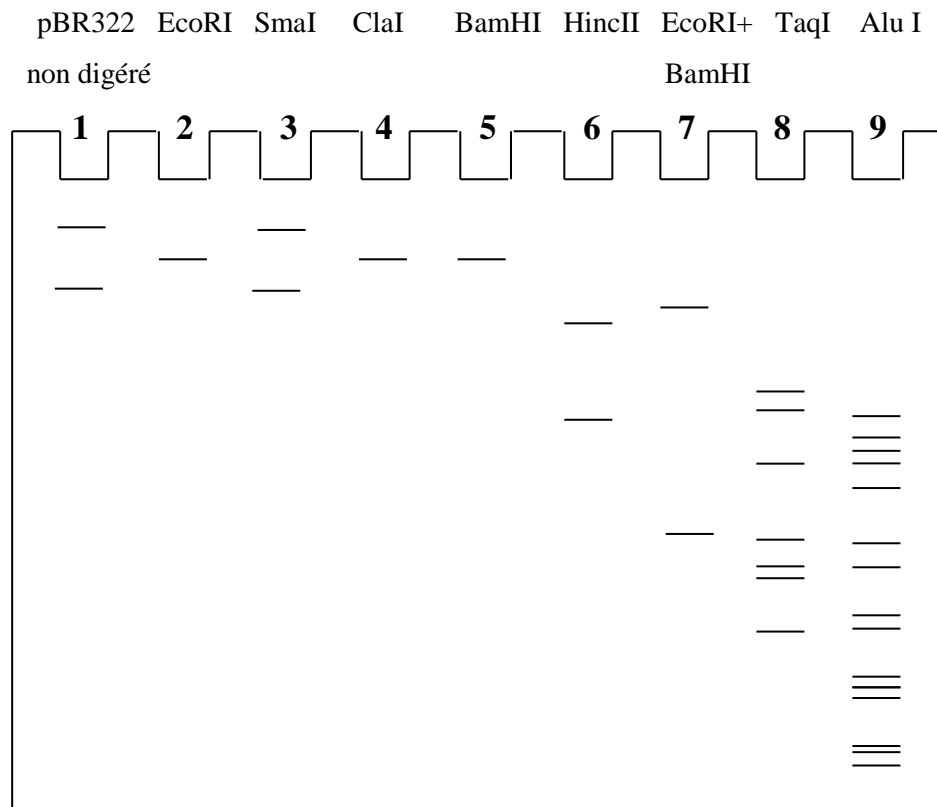
Exercice IV : Un plasmide recombinants est digéré par les enzymes de restriction *Bam*HI et *Eco*RI. Après migration et séparation des fragments d'ADN sur gel d'agarose, on obtient les profils de restriction suivants :

<i>Eco</i> RI	<i>Bam</i> HI	<i>Eco</i> RI + <i>Bam</i> HI
— 8Kb		
— 6Kb	— 5,5 Kb	
	— 4,5 Kb	
	— 4 Kb	— 4 Kb
		— 3,5 Kb
		— 3 Kb
		— 2,5 Kb
		— 1Kb

Série TD 2

1) L'ADN du plasmide pBR 322 est circulaire et comprend 4363pb. Il est soumis à l'action de diverses endonucléases de restriction dans des conditions où l'hydrolyse est totale. Les produits des digestions enzymatiques sont soumis à une électrophorèse dans un gel d'agarose. Le résultat de la migration est présenté sur le graphique I.

Graphique I



- Pourquoi trouve t'on deux bande dans le puit 1 ?
 - Combien pBR322 contient-il de séquences reconnues par les endonucléases de restriction *EcoRI*, *Sma I* et *Hind CII* ? Pourquoi ?
 - Le profil de digestion par ces enzymes aurait-il été le même si la molécule de pBR322 avait été linéaire au lieu d'être circulaire? Justifier cas par cas.
- 2) Les séquences reconnus (site de restriction) par les endonucléases utilisées sont les suivantes (un seul brin est représenté ; Py : pyrimidine ; pu : purine).

Alu I 5'AGTCT.....3'
Bam HI 5'GGATCC.....3'
Cla I 5'ATCGAT.....3'
EcoRI 5'GAATTC.....3'
Hind cII 5'GTPyPuAC.....3'
Sma I 5'CCCGGG.....3'
Taq I 5'TCGA.....3'

Les tailles des fragments d'ADN générés par l'action de *Taq I* sur pBR322 sont : 1444pb, 1307pb, 475pb, 368pb, 316pb, 312pb et 141pb.

- L'observation et la comparaison de ces sites de restriction permet de mettre en évidence une caractéristique de la séquence reconnu par ces endonucléases : laquelle ?
- Quelle est, en pb, la plus petite distance qui, d'après le graphique I, sépare les sites *EcoRI* et *BamHI* ? Pourquoi ?
- Donner en g et en Da les PM des 2 fragments générés par l'action simultanée de *Bam HI* et *Eco RI*.

3) L'endonucléases de restriction *EcoRI* génère des extrémités 5P débordantes selon le schéma suivant :

Site *EcoRI* 5'GAATTC.....3'
 3'CTTAAG.....5'
Extrémités 5'G AATTC.....3'
générés 3'CTTAA G.....5'

L'expérimentateur soumet pBR322 au protocole expérimentale suivant :

- Il faut agir *EcoRI* sur pBR322 puis il inactive l'endonucléases ;
 - Il complète les extrémités 5p débordantes par l'ADN pol ;
 - Il soude entre les extrémités complétées grâce à une ligase ;
 - Il transforme une souche bactérienne avec produit de la ligation pour amplifier l'ADN ;
 - Il purifie à partir d'une culture de bactéries transformées le nouveau plasmide pBR422.
- Quelle est, en pb, la taille de pBR422 ?

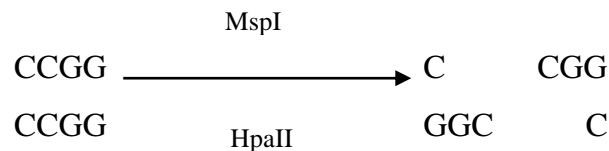
Série TD 3

Exercice I :

La digestion d'une molécule d'ADN avec l'enzyme de restriction *HindIII* conduit à l'obtention de deux fragments de 2.2Kb et de 2.8Kb. *EcoRI* coupe cette même molécule d'ADN en générant deux fragments de 1.8 Kb et 3.2Kb. Lorsque cette molécule d'ADN est digérée par un mélange des deux enzymes de restriction, on obtient quatre fragments de 0.8Kb, 1Kb, 1.2Kb et 2Kb. Dessinez la carte de restriction de cette molécule d'ADN.

Exercice II :

MspI et *HpaII* sont deux endonucléases qui clivent le même motif palindromique CCGG/GGCC. Cette coupure entraîne l'apparition des bouts cohésifs 3'cohésifs.

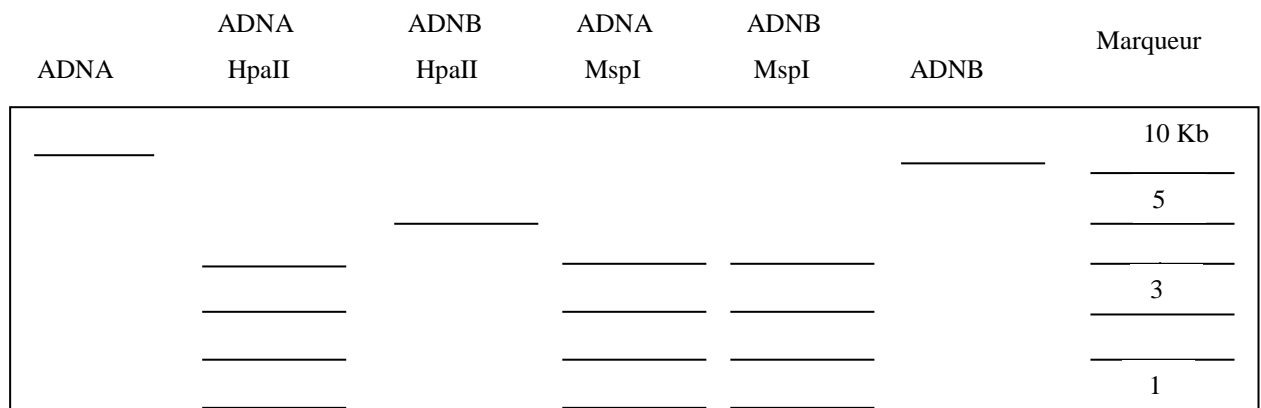


Il est par ailleurs démontré que la méthylation, des cytosine qui précèdent les guanosines protège les sites CCGG/GGCC de l'attaque de *HpaII* et non de celle de *MspI*.

1- Comment appelle-t-on les enzymes *MspI* et *HpaII* ?

L'étude de deux fragments d'ADN(A) et (B) isolés à partir de deux tissus différents d'un même individu, est réalisé à l'aide de ces deux enzymes.

L'électrogramme ci-après est obtenu après coloration au BET des sous-fragments de digestion des deux ADN, préalablement séparés par électrophorèse.

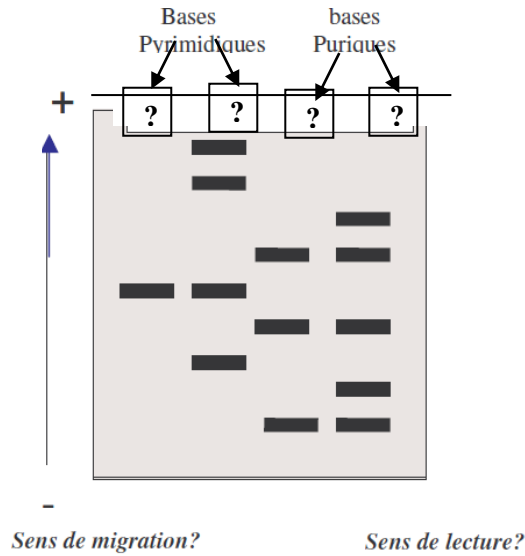


2- Que peut-on dire des ADN A et B ? Justifier votre réponse.

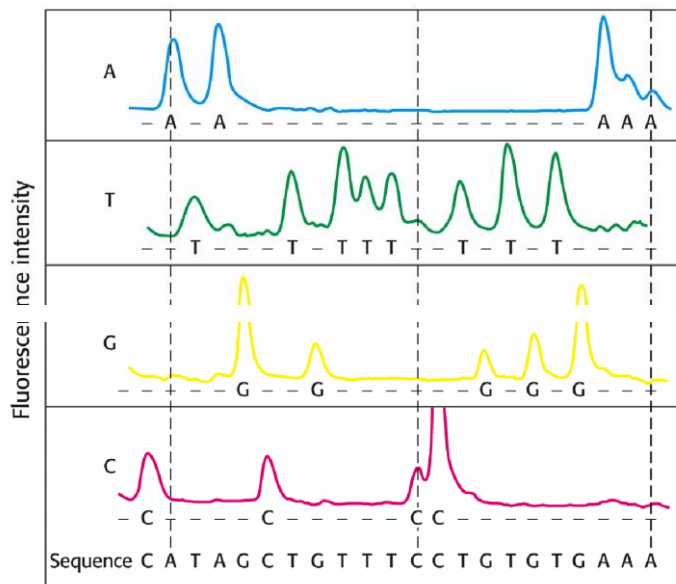
3- Donnez la cartographie des ADN A et B.

Série TD 4

Exercice 1: Compléter les résultats ci-dessous et donner la séquence du monobrin de DNA



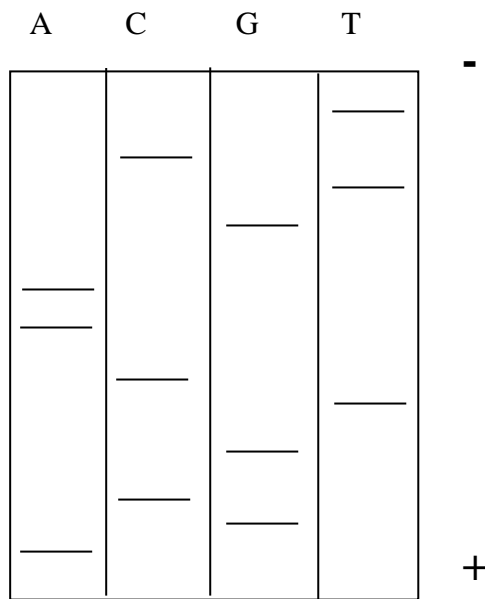
Exercice 2 : Voici un pictogramme obtenu après séquençage d'un fragment d'ADN par un séquenceur automatique, déduisez la séquence du brin complémentaire de cet ADN



Exercice 3 :

Le résultat de l'électrophorèse d'un séquençage d'oligonucléotide par la méthode de sanger est donné sur le diagramme ci-dessous. Quel est la séquence du nucléotide analysé en supposant que le dernier nucléotide est un A?

Un second échantillon de cet oligonucléotide est marqué à son extrémité 3'OH et séquencé par la méthode de Maxam et Gilbert. Tracer le diagramme des bandes observées sur un autoradiogramme du gel de séquençage.



Références bibliographiques

References Bibliographiques

Brown T. A. (2010). Gene Cloning and DNA Analysis: An Introduction. 6th Ed, Wiley Blackwell. London, UK. 782p.

Clark D. (2005). Molecular Biology: Understanding the Genetic Revolution. Southern Illinois University. USA Elsevier Academic Press. 784p.

Gubta P. K. (2005). Microbiology, Cell Physiology and Biotechnology. 2nd Ed. National Offset Printers, Meerut, India. 311p.

Hogg S. (2005). Essential Microbiology. British Library Cataloguing in Publication Data. John Wiley & Sons. UK. 468p.

Housset C. et A. Raisonnier. (2006). Biologie Moléculaire. Biochimie PCEM1 Université Paris-VI. 204p.

Klug S. W., M. R. Cummings, C. A. Spencer & M. A. Palladino. (2013). Essentials of genetics. 8th Ed. Pearson Education. New Jersey. USA. 608p.

Lodish H., J. Darnell, D. Baltimore. (1989). La cellule: Biologie Moléculaire. Edition VIGOT. Bibliothèque nationale du Québec-Canada. 1189P.

Lodish, Berk, Matsudaira, Kaiser, Krieger, Scott, Zipursky, Darnell. (2003). Molecular Cell Biology. 6th Ed. 937p.

Madigan M., J. Martinko. (2007). Biologie des microorganismes. 11e Ed. Pearson Education-Paris-France. 1047p.

Misener S. and S. A. Krawetz. (2000). Bioinformatics Methods and protocols. Humana Press Inc, New Jersey. USA. 500p.

Nicholl S. T. D. (2008). *An introduction to Genetic Engineering*. Third Edition. Cambridge University Press. UK.292p.

Okafor N. (2007).*Modern Industrial Microbiology and Biotechnology*. Science Publishers, Enfield, NH, USA. 530p.

Passarge E. (2007).*Color Atlas of Genetics*. 3rdEd. Flexibook. Thieme Stuttgart, NY. USA. 486p.

Prescott M. L., J. P. Harley, et D. A. Klein. (2002). *Microbiology*. 5thEd. The McGraw–Hill Companies.1139p.

Prescott M. L., J. P. Harley, et D. A. Klein. (2007). *Microbiologie*. 2eEd. Edition De Boek Université. Paris. 1137p.

Primose S. B., R. M. Twyman and R.W. Old. (2002) *Principles of Gene Manipulation*. 6th Ed. Blackwell Scientific Publishing, Oxford. UK. 319p.

Sambrook, J., and D. Russell. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory. New York. USA. 2344p.

Singh B. D. (2007). *Molecular Biology and genetic engeneering*. Kalyani Publishers. India. 311p.

Smith E. J. (2005). *Biotechnology*. 3th Ed. Cambridge University Press. UK. 234p.
123

Streips N. U. & R. Yasbin. (2002). *Modern Microbial Genetics*. 2nd Ed. Wiley-Liss, Inc. New York. USA; 603p.

Turner P. C., A. G. Mc Lennan, A. D. Bates, M. R. H. White. (2002). *Instants Notes in Molecular Biology*. Bios scientific Publishers Limited. Oxford. UK. 386p.

Ussery D. W., T. M. Wassenaar, and S. Borini. (2009). *Computing for Comparative Microbial Genomics: Bioinformatics for Microbiologists*. Computational Biology Series. Springer-Verlag London Limited. UK.270p.

Watson J., T. Baker, S. Bell, A. Gann, M. Levine, R. Losick. (2009). *Biologie Moléculaire du gène*. 6eEd. Pearson Education- Paris-France. 688p.