



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique



Université Abbes Laghrour-Khenchela
Faculté des Sciences de la Nature et de la Vie
Département de biologie moléculaire et cellulaire

Cours de

Biologie Moléculaire

Dr BOUHALIT Samira

Destiné aux étudiants de 3^{ème} année Licence Biochimie

2020-2021

Préface

La biologie moléculaire est la compréhension des mécanismes de fonctionnement de la cellule au niveau moléculaire. C'est l'étude des molécules porteuses du message héréditaire (ADN, ARN), de leur structure, leur synthèse, leur altération et du contrôle de l'expression des gènes. On a consacré un intérêt particulier pour les macromolécules de l'ADN et de l'ARN ainsi qu'aux processus de réplication, de transcription et de traduction. Les nouvelles technologies expérimentales mises en jeu pour la manipulation de ces molécules, appelées aussi techniques de génie génétique, sont essentielles à la biologie moléculaire moderne ; celles-ci présentent non seulement des informations fondamentales sur les molécules et la compréhension des gènes dans les processus biologique fondamentaux, mais possèdent également des applications pratiques considérables dans le développement des produits nouveaux sert pour le diagnostic, le traitement, et la prévention.

Le présent polycopié constitue un support pédagogique d'une grande importance, destiné aux étudiants ayant dans leur cursus le module de Biologie Moléculaire et les Techniques en Biologie Moléculaire.

Le polycopié abordera plusieurs thèmes qui permettra aux étudiants de comprendre la structure et le dynamique de l'ADN, l'organisation d'un génome, les mécanismes de mutations et de réparation de l'ADN, la mutagenèse et le diagnostic génotypique avec la transmission et la conservation de l'information génétique et la régulation de l'expression des gènes. Le dernier chapitre est consacré à l'extraction des acides nucléiques ainsi que les techniques de manipulation d'acides nucléiques et les techniques de bases (PCR, séquençage, électrophorèse, l'hybridation moléculaire, transfert sur membrane et marquage).

Liste des tableaux

Tableau I :	Principales caractéristiques géométriques des ADN A, B et Z.	7
Tableau II:	Correspondance entre les triplets de bases de l'ARNm et les acides aminés et codon stop.	53
Tableau III :	État de la chromatine en fonction de l'acétylation des histones et la méthylation de l'ADN	67

Liste des figures

Figure 01 :	Exemple de nucléotide et nucléoside	1
Figure 02:	Structure des bases azotées qui constituent l'ADN et l'ARN.	2
Figure 03:	Structure du ribose et du désoxyribose constitue respectivement l'ARN et l'ADN.	3
Figure 04 :	Formation d'un brin d'ADN par polymérisation des nucléotides	3
Figure 05 :	Représentation schématique de la double hélice d'ADN	4
Figure 06:	Profil de dénaturation de l'ADN sous l'effet de la chaleur.	5
Figure 07 :	Représentation schématique des formes A, B et Z de l'ADN.	6
Figure 08 :	Les différentes échelles de l'organisation hiérarchique du chromosome	8
Figure 09 :	Tautomérisation spontanée des bases : appariements illicites	12
Figure 10 :	Bases endommagées par des radicaux oxygène.	12
Figure 11:	Principales lésions de l'ADN	13
Figure 12 :	Position des mutations sur un gène et leurs conséquences	14
Figure 13:	Exemples de lésions photochimiques sur des bases pyrimidiniques	16
Figure 14:	Mésappariements dus à la désamination des bases par l'acide nitreux	17
Figure 15:	Mutagenèse ciblée /Nucléase à doigt de zinc (ZFN), TALENs, et CRISPR/Cas9	19
Figure 16:	Transgénèse, cisgénèse et intragénèse	20
Figure 17:	Stratégie de recherche de mutations par technique de CSGE puis séquençage.	26
Figure 18 :	Expérience de Meselson et Stahl	29
Figure 19:	Processus de la réplication au niveau d'une fourche de réplication chez <i>E.coli</i> .	31

Figure 20:	Modèle de la réplication de l'ADN chez l'eucaryote <i>Saccharomyces cerevisiae</i> .	32
Figure 21:	Mécanisme d'action de la télomérase	32
Figure 22:	Domages de l'ADN, mécanisme de réparation et conséquences.	34
Figure 23:	Le système de réparation MMR (methyl Mismatch Repair)	36
Figure 24:	Représentation des voies constituant la réparation par excision de base	36
Figure 25 :	Réparation par réversion directe du dimère cyclobutane de thymine	38
Figure 26:	Réparation par recombinaison homologue	38
Figure 27 :	Séquences palindromique cible des enzymes de restriction	41
Figure 28 :	Polymorphismes de longueur des fragments de restriction (RFLP)	42
Figure 29 :	Représentation du site promoteur des Procaryotes.	45
Figure 30 :	La transcription chez les bactéries	45
Figure 31:	Mise en place du complexe protéique d'initiation	48
Figure 32:	Schéma d'un gène eucaryote et d'un transcrit ARNm mature prêt à être traduit.	49
Figure 33:	Addition de la coiffe: 7m-Guanosine.	50
Figure 34:	Formation de splicéosome	51
Figure 35:	Structure générale et composition des ribosomes procaryotes et eucaryotes	55
Figure 36 :	Structures secondaire et tertiaire d'une molécule d'ARNt.	56
Figure 37:	Formation du complexe d'initiation et élongation de la traduction chez les eucaryotes.	59
Figure 38:	Fonctionnement de l'opéron lactose en présence de lactose	62
Figure 39:	L'opéron tryptophane chez <i>E. coli</i>	63
Figure 40:	Processus de l'atténuation de l'opéron tryptophane d' <i>E. coli</i>	64
Figure 41:	Mécanisme de contrôle positif de l'opéron <i>malEFG</i>	65
Figure 42:	Interaction entre facteurs et séquences régulatrices	68
Figure 43 :	Protéines en fonction différente	69
Figure 44 :	Exemple de la synthèse de la ferritine ; inhibition de la lecture de l'ARN par RE	70
Figure 45 :	Exemple de la synthèse de β -globine ; inhibition des facteurs de traduction.	70

Figure 46 :	(A) exemple de la biosynthèse de l'ubiquitine (protéolyse), (B) exemple de la myostatine (clivage protéolytique d'un dimère)	71
Figure 47:	Influence des différents facteurs sur les récepteurs nucléaires	72
Figure 48:	Structure des récepteurs nucléaires	72
Figure 49:	Types de récepteurs nucléaires	73
Figure 50:	Distance de migration en fonction du logarithme de la taille des fragments d'ADN.	82
Figure 51:	Hybridation des acides nucléiques.	84
Figure 52:	Schéma des différentes étapes d'un Southern blot	85
Figure 53:	Illustration des positions sur l'ADN et des fragments après migration	86
Figure 54 :	Amplification d'ADN par PCR.	88
Figure 55 :	Evolution de la température et des différents types de brins d'ADN au cours des 4 premiers cycles de la PCR	89
Figure 56:	Les différentes formes des nucléotides.	91
Figure 57:	Séquençage par la méthode enzymatique de Sanger	92
Figure 58:	Détection par fluorescence et séparation par électrophorèse ; Séquençage d'ADN méthode de Sanger	92
Figure 59:	Types d'interactions ADN-Protéine.	94
Figure 60:	L'interaction des protéines homodimériques (a) le motif leucine zipper ; (b) le motif hélice boucle hélice	95
Figure 61 :	(A) Technique de décalage de mobilité (Technique EMSA), (B) Principe du shift et supershift	95
Figure 62:	Etapes de l'immunoprécipitation de la chromatine pontée.	96
Figure 63:	Caractérisation de l'interaction ADN-Protéine par ADNase I	97
Figure 64:	Principe du simple hybride	98
Figure 65:	Principe du Biacore.	98

Sommaire

Chapitre I : Le support de l'information génétique, l'ADN	1
I. Structure et dynamique de l'ADN	1
1. Composition chimique de l'ADN	1
2. Structure de la double Hélice	3
3. Propriétés de la molécule d'ADN	4
4. Topologie de l'ADN	6
5. Structure des chromosomes	7
6. Implications biologiques du surenroulement de l'ADN	8
II. Structure et organisation du génome des procaryotes et eucaryotes	9
Chapitre II : Mutations, mutagenèse et détection	10
I. La mutation génique	10
1. Les différents types de mutation	10
1.1. Mutations ponctuelles	10
1.2. Mutation réverse	11
1.3. Mutation dominante négative	11
1.4. Mutation conditionnelle	11
2. Cause des mutations géniques	11
2.1. Instabilité chimique intrinsèque de l'ADN	11
2.2. Les sous-produits du métabolisme cellulaire	12
2.3. Les agents environnementaux (lésions exogènes)	13
3. Les conséquences et les effets des lésions	13
II. Mutagenèse	15
1. Mutagenèse aléatoire	15
1.1. Mutagenèse physique	15
1.2. Mutagenèse chimique	16
1.3. Mutagenèse biologiques	17
2. La mutagenèse dirigée	17
2.1. Techniques d'édition du génome	18
2.1.1. Les nucléases à doigt de zinc (ZFN)	18
2.1.2. Les TALENs	18
2.1.3. La technique CRISPR/Cas	19
2.1.4. Les méganucléases	19

2.1.5. Introduction transitoire d'ADN recombinant/ Mutagenèse dirigée par oligonucléotides (ODM)	19
2.2. Cisgénèse et intragénèse	20
2.3. Méthylation de l'ADN dépendante de l'ARN (RdDM)	20
2.4. Agro-infiltration	21
2.5. Sélection inverse	21
2.6. Greffage	21
III. Diagnostic génotypique	22
1. Le diagnostic de maladies pré-symptomatique	22
2. Approches de diagnostic génotypique	22
2.1. Diagnostic direct	23
2.1.1. Détection de mutations connues (hot spot)	23
2.1.2. Détection de mutations dans un gène sans hot spot	24
2.2. Diagnostic indirect	26
Chapitre III : Transmission et conservation de l'information génétique	28
I. La réplication de l'ADN et sa régulation	28
1. Généralité	28
2. Expériences de Meselson et Stahl	28
3. Mécanisme de la réplication chez les procaryotes	29
4. La réplication chez les eucaryotes	31
5. Régulation de la réplication	33
II. La réparation de l'ADN et détection du pouvoir mutagène	34
1. Nature et origine des dommages à l'ADN	34
2. Les mécanismes de réparation faisant suite à la réplication	35
2.1. La réparation des erreurs de réplication	35
2.2. Correction des mésappariements après la réplication	35
2.3. Mécanisme d'excision resynthèse (BER)	36
2.4. Réparation par excision/réparation des plusieurs nucléotides (NER)	37
3. Réparation des lésions (naturelles ou induites)	37
3.1. Réparation par réversion direct	37
3.2. La réparation par recombinaison	38
3.3. Une réparation mutagène, le système SOS	39
III. Systèmes de restriction-modification	39

1. Origines de méthylation d'ADN dans les bactéries	39
2. Types de systèmes de restriction-modification dans les bactéries	40
3. Nomenclature des enzymes de restriction	40
4. Séquences cibles et types de clivage	41
5. Applications des enzymes de restriction	41
5.1. Technologie de l'ADN recombinant	42
5.2. Cartographie de restriction	42
6. Polymorphisme de restriction	42
6.1. Polymorphismes de longueur des fragments de restriction (RFLP)	42
6.2. Analyse RFLP	43
6.3. Les applications de la technique de RFLP	43
Chapitre IV: L'expression de l'information génétique et son contrôle	44
I. La transcription et la maturation de l'ARN	44
1. Généralités	44
2. Caractéristiques générales de la transcription	44
2.1. Les éléments nécessaires à la transcription	44
2.2. Les étapes de la transcription chez les procaryotes.	44
2.3. Les ARNs polycistroniques	46
3. Transcription de l'ADN eucaryote	46
3.1. Les ARN-polymérases eucaryotes	46
3.2. La formation d'un ARN prémessager	47
3.3. Maturation des transcrits primaires	49
II. La traduction et la maturation des protéines	52
1. Le code génétique	52
2. Décryptage du code génétique	53
3. Les acteurs de la traduction	54
4. Les différentes étapes de la traduction	56
III. Régulation de l'expression des gènes	60
1. Régulation de l'expression des gènes dans la cellule procaryote	60
1.1. Au niveau transcriptionnel	60
1.1.1. Contrôle négatif de la transcription "Répression et induction"	61
A. Régulation de l'expression des gènes impliqués dans les voies cataboliques	61
B. Régulation de l'expression des gènes impliqués dans les voies anaboliques	63

1.1.2. Contrôle positif de la transcription	65
1.2. Régulation post-transcriptionnelle	65
1.3. Au niveau traductionnel	66
2. Régulation de l'expression des gènes dans la cellule eucaryote	66
2.1. Au niveau chromatinien	66
2.2. Au niveau transcriptionnel	67
2.3. Au niveau post-transcriptionnel	68
2.3.1. Epissage alternatif	68
2.3.2. Modification éditorial de l'ARNm	69
2.3.3. Régulation de la durée de vie des ARNm	69
2.4. Régulation traductionnelle	70
2.5. Régulation post-traductionnelle	71
IV. Voies de régulation des gènes par les signaux extracellulaires	71
1. La signalisation intracellulaire par facteurs environnementaux	71
2. Modèle de régulation de l'expression d'un gène par un récepteur nucléaire	72
2.1. Structure des récepteurs nucléaires	72
2.2. Types de récepteurs nucléaires	73
3.3. Mécanisme de régulation par les récepteurs nucléaires	73
Chapitre V : Méthodologie et biologie moléculaire	75
1. Techniques d'extraction et de purification des acides nucléiques	75
1.1. Méthodes d'extraction	75
1.1.1. Lyse des cellules	75
1.1.2. Elimination des protéines	76
1.1.3. Réactifs divers ajoutés au tampon d'extraction	77
1.1.4. Elimination des ARN lors de l'extraction de l'ADN	77
1.2. Préparation de l'ADN plasmidique	78
1.3. Méthodes de purification	78
1.3.1. Extraction/Précipitation	78
1.3.1.1. Purification au phénol-chloroforme	79
1.3.1.2. Précipitation/concentration par l'éthanol ou l'isopropanol	79
1.3.2. Chromatographie	80
1.3.3. Centrifugation	80
1.3.4. Séparation par affinité	80

2. Séparation et Visualisation des acides nucléiques (Electrophorèse)	81
3. Analyse de la pureté de l'ADN extrait	82
4. Quantification de l'ADN	83
4.1. Dosage fluorimétrique classique de l'ADN	83
4.2. Dosage de l'ADN par fluorescence en présence de BET	83
4.3. Dosage colorimétrique de l'ADN	83
4.4. Dosage par absorption U.V. de la concentration en ADN	83
4.5. Dosage par la PCR temps réel	83
5. Hybridation des acides nucléiques	84
5.1. Généralités	84
5.2. Types d'hybridation	85
6. Amplification in vitro des acides nucléiques	87
6.1. Réaction de polymérisation en chaîne PCR	87
6.2. RT-PCR	90
7. Le séquençage de l'ADN	90
8. Interaction protéine-ADN	93
8.1. Généralités	93
8.2. Mécanisme d'interaction ADN-Protéine	93
8.3. Motifs protéiques de fixation aux acides nucléiques	94
8.4. Techniques d'étude des interactions ADN-protéine	95
8.4.1. Le retard sur gel (EMSA ou <i>electrophoretic mobility shift assay</i>)	95
8.4.2. Immunoprécipitation de chromatine (ChIP)	96
8.4.3. L'empreinte à la DNase (<i>DNase footprinting assay</i>)	97
8.4.4. Simple Hybride	97
8.4.5. Système Biacore (SPR ou surface plasmon resonance)	98

I. Structure et dynamique de l'ADN :

Les acides nucléiques sont des macromolécules biologiques qui, comme leur nom indique, ont été tout d'abord isolés du noyau, mais ce terme a néanmoins été conservé.

Présentes dans toutes les cellules vivantes et également chez les virus. Ils sont impliqués dans la conservation, transmission et expression de l'information génétique. Les acides nucléiques sont de deux types, l'ADN et les ARN. L'ADN (**A**cide **D**ésoxyribo **N**ucléique) est le support de l'information génétique. Les ARN (**A**cide **R**ibo**N**ucléique) ont soit un rôle de support de l'information afin d'être traduit en protéines (ARN messager), ou bien un rôle structurale (ARN ribosomiques, ARN de transferts et autres petits ARN).

1. Composition chimique de l'ADN

Les acides nucléiques sont des polymères d'unités élémentaires appelées **nucléotides**. Chaque nucléotide est composé d'acide phosphorique, base azotée (A, T, C, G) et un pentose (ribose pour l'ARN ou désoxyribose pour l'ADN). L'association d'une base et d'un pentose par une liaison N-glycosidique est appelée un **nucléoside**. De cette manière les bases puriques rajoutent le suffixe « osine » et les bases pyrimidiques rajoutent le suffixe « idine » ; on parle ainsi d'adénosine, de guanosine, de cytidine, d'uridine et de thymidine. L'hypoxanthine est une exception et devient l'inosine.

Un nucléotide résulte de l'estérification de la fonction alcool d'un nucléoside par une molécule d'acide ortho-phosphorique. On parle alors d'adénylate (ou AMP pour adénosine monophosphate), de guanylate (ou GMP), de thymidylate (ou TMP) et de cytidylate (ou CMP). L'hypoxanthine est à nouveau soumise à exception, on parle de l'inosinate.

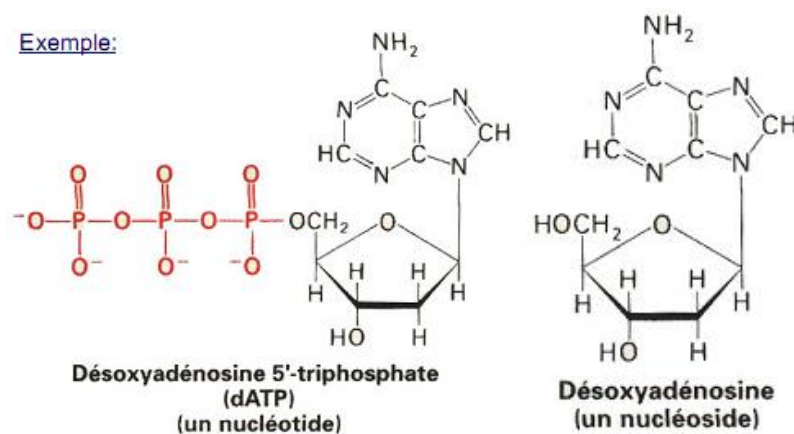


Figure 01 : Exemple de nucléotide et nucléoside

a. Les bases :

Les **bases** sont des structures coplanaires présentant une résonance grâce à leurs doubles liaisons conjuguées. Elles sont de deux types :

- Les bases pyrimidiques (1 cycle): la cytosine, la thymine (ADN) et l'uracile (ARN).
- Les bases puriques (2 cycles): noyau pyrimidine + noyau imidazole: la guanine, l'adénine et l'hypoxanthine (précurseur des bases puriques et présent au niveau des ARNt).

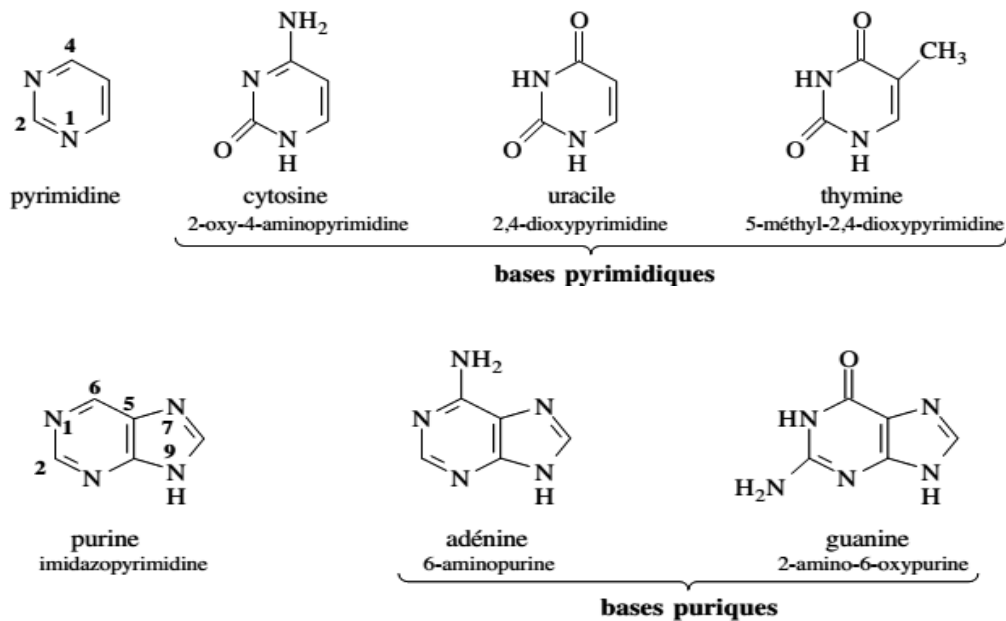


Figure 02: Structure des bases azotées qui constituent l'ADN et l'ARN.

La structure et la composition en bases azotées d'une molécule d'ADN double-brin sont définies comme suit :

- La quantité d'adénine est proportionnelle à la quantité de thymine et la quantité de guanine est proportionnelle à la quantité de cytosine ($A=T$ et $G=C$). Le rapport de Chargaff (somme des bases puriques/somme des bases pyrimidiques ou $A+G / T+C$) est donc égal ou à peu près égal à 1.
- Toutes les cellules d'un même organisme ont un ADN de même composition en bases. Le rapport $(A+T)/(C+G)$ est propre à chaque espèce et il permet de déterminer la composition en bases d'une espèce. Ce rapport est en général supérieur à un (> 1) chez les animaux et les végétaux. (E. Coli : 1 ; Homme : 1,50 ; Mouton : 1,36)

b. Pentoses Il s'agit : - du ribose (sous forme β -D ribofuranose) : constitutif de l'ARN

- du désoxyribose (sous forme β -D 2-désoxy ribofuranose) : constitutif de l'ADN

Les atomes de C ont la particularité d'être numérotés de 1' à 5'. Le désoxyribose est une forme beaucoup plus stable que le ribose. Cette forme désoxy entraîne la stabilisation de l'hélice.



Figure 03: Structure du ribose et du désoxyribose constitue respectivement l'ARN et l'ADN.

c. Acide phosphorique

C'est un triacide dont une fonction acide est dissociée permettant de donner une charge négative à l'ADN, et dont les deux autres peuvent former des liaisons phosphodiester.

Les nucléotides s'associent grâce à des liaisons phosphodiester formées entre le phosphate en 5' d'un nucléotide et le OH en 3' du nucléotide suivant. La séquence qui code l'information génétique est toujours dans le sens 5'→3' (sens dans lequel les enzymes polymérase copient l'ADN) (figure 04).

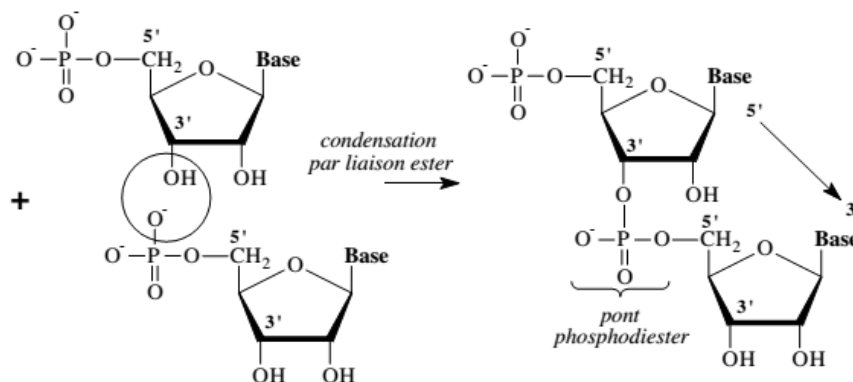


Figure 04 : Formation d'un brin d'ADN par polymérisation des nucléotides

2. Structure de la double Hélice

La découverte fondamentale de Watson et Crick est que L'ADN est formé de deux chaînes enroulées l'une autour de l'autre pour former une double hélice. La partie sucre-phosphate constituant le squelette est située à l'extérieur de l'hélice, les bases azotées se

trouvent au centre. Le sens de polymérisation d'un brin étant inverse de celui du brin qui lui est associé, Les deux hélices sont réunies par les liaisons hydrogènes qui forment entre les bases azotées une complémentarité bien définie : A-T ; G-C. L'espacement entre les deux hélices est tel qu'à chaque fois une base purine interagit avec une base pyrimidique.

La double hélice effectue un tour toutes les dix paires de bases, la distance entre bases sur l'axe est $3,4 \text{ \AA}$. Un tour fait 34 \AA , le diamètre de l'hélice est de 20 \AA (figure 05). Ceci est la structure classique de la molécule de DNA connue comme étant la forme B, c'est la forme que l'on rencontre dans les conditions physiologiques de température et de pH.

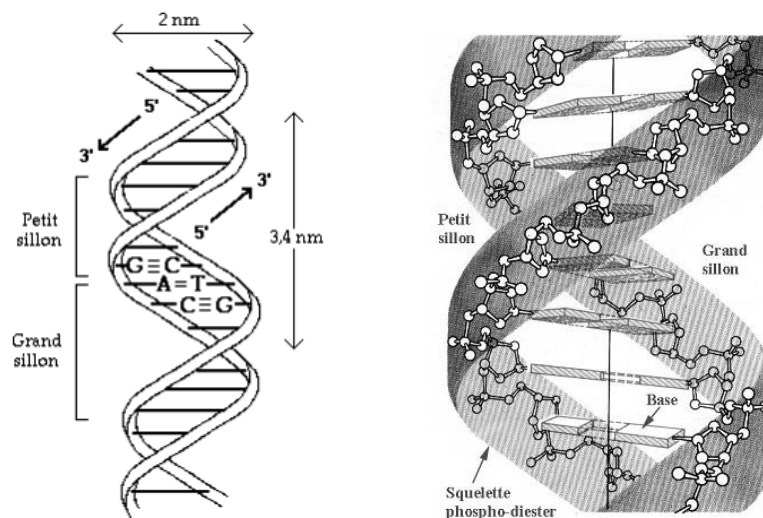


Figure 05 : Représentation schématique de la double hélice d'ADN

3. Propriétés de la molécule d'ADN :

Elles sont dues d'une part à sa composition (nucléotides) et à sa structure (longues chaînes, double hélice).

a) Solubilité : L'ADN devient un sel d'acide en milieu aqueux et est ainsi soluble. Il précipite en présence d'éthanol et d'une forte concentration saline (phénomène de « relarguage»). Cette propriété permet sa purification, Ils sont récupérés sous forme d'un long filament translucide (méduse).

b) Viscosité : Les solutions d'ADN ont une grande viscosité en raison de la grande longueur de la double hélice et de sa rigidité; les mesures de viscosité permettent de suivre la dénaturation de l'ADN.

c) Absorption des UV : Les acides nucléiques absorbent les UV à une longueur d'onde de 260 nm. Ceci en raison de l'interaction entre la lumière UV et les cycles des bases puriques

et pyrimidiques. L'ADN simple brin absorbe plus que l'ADN double brin: cela vient du fait que dans l'ADN double brin, les bases sont partiellement masquées.

Pour estimer la contamination par les protéines lors de la purification des acides nucléiques et vérifier le degré de pureté d'un échantillon d'ADN après son extraction en calculant le rapport A_{260}/A_{280} nm. Si ce rapport est supérieur ou égal à 1,8 alors l'échantillon d'ADN est considéré comme étant pur (>2 il sera considéré comme étant contaminé par de l'ARN). Par contre, si ce rapport est inférieur à 1,8 alors l'échantillon d'ADN est considéré comme étant contaminé par des protéines.

d) Dénaturation thermique

Les liaisons hydrogène, relativement fragiles, peuvent être détruites par chauffage, dans ces conditions les structures secondaires disparaissent. Pour l'ADN, le résultat est la séparation complète des deux brins qui le composent : il y a dénaturation de la molécule. Comme l'ADN dénaturé est simple brin ; la viscosité diminue et l'absorbance augmente c'est **l'effet hyperchrome**. On peut suivre l'absorbance en fonction de la température; on obtient une courbe sigmoïde de dénaturation dont le point d'inflexion définit la température de fusion T_m (melting température = 50 % de dénaturation) de la molécule (Figure 06). Lorsque le plateau de la courbe est atteint (densité optique maximale), la dénaturation est complète et la solution ne contient que de l'ADN simple-brin.

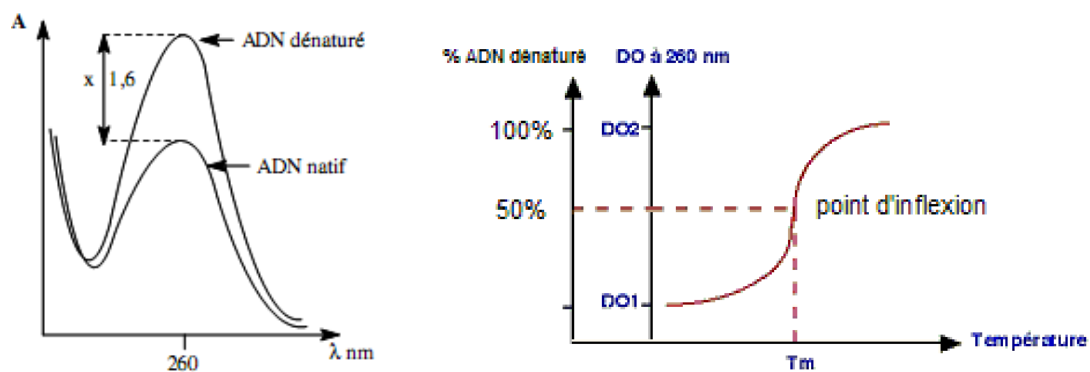


Figure 06: Profil de dénaturation de l'ADN sous l'effet de la chaleur.

Les simples brins peuvent, dans des conditions appropriées de température (refroidissement lent), se réassocier par complémentarité entre les bases pour reformer un ADN bi caténaire, dans ce cas on parle de renaturation. La dénaturation est irréversible si le refroidissement est trop brutal. Les possibilités de renaturation sont utilisées pour faire de l'hybridation moléculaire pour étudier l'ADN humain : l'ADN cible (ADN génomique) est hybridé avec un ADN sonde (ADN de référence). L'augmentation des forces ioniques du

milieu réactionnel (NaCl) permet une renaturation plus facile, les cations neutralisant les forces répulsives de l'ADN. Il faut noter que la température de fusion T_m est dépendante du pH et de la force ionique du milieu et qu'elle diminue lorsque cette dernière augmente, mais surtout, le T_m dépend de la longueur et la composition de l'ADN en GC. T_m est d'autant plus élevée que le pourcentage de bases G + C (Impliquant 3 liaisons H) est grand

4. Topologie de l'ADN :

L'axe des paires de bases ne passe pas par l'axe de la double hélice : on aura donc une partie plus large, le **grand sillon**, et une partie plus petite, le **petit sillon**. Le grand sillon favorise la fixation de protéines ou d'ARN régulateur, le petit sillon favorise la liaison aux histones, selon la forme des sillons alternés on définit différente configuration d'ADN dont les plus connus sont: ADN A, ADN B et ADN Z, les deux formes A et B sont composées d'hélice droite emmêlée en torsade par contre la forme Z est composée d'hélice gauche en zigzag, il semble que la forme ADN B soit le plus courante (figure 07).

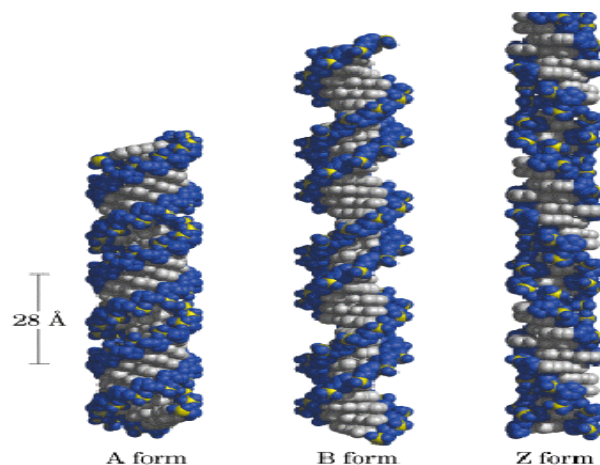


Figure 07 : Représentation schématique des formes A, B et Z de l'ADN.

L'ADN B est une double hélice régulière parfaite, sauf que les paires de bases opposées ne se trouvent pas exactement au même niveau. Ils sont tordus de manière hélicoïdale. De cette façon, l'ADN peut facilement être plié sans provoquer de changements essentiels dans les structures locales. L'ADN-A, lorsque la teneur d'eau d'une solution contenant une molécule d'ADN est diminuée (trouvé dans les spores bactériennes) la molécule change la conformation en conformation A. Dans l'ADN-Z, le squelette sucre-phosphate a un motif en zigzag; le sillon unique de l'ADN-Z a une plus grande densité de molécules chargées négativement. L'ADN-Z peut apparaître dans des segments limités in vivo. Un segment d'ADN B constitué de paires de GC peut être converti en ADN-Z lorsque les bases sont

tournées de 180°. Normalement, l'ADN-Z est thermodynamiquement relativement instable. Cependant, la transition vers l'ADN-Z est facilitée lorsque la cytosine est méthylée en position 5 (C5). La modification de l'ADN par la méthylation de la cytosine est fréquente dans certaines régions de l'ADN des eucaryotes. Il existe des protéines spécifiques qui se lient à l'ADN-Z, mais leur rôle pour la régulation de la transcription n'est pas clair.

Tableau I : Principales caractéristiques géométriques des ADN A, B et Z.

Caractéristique géométrique	ADN A	ADN B	ADN Z
Hélicité	Droite	Droite	Gauche
Unité de répétition	1pb	1pb	2pb
Rotation/pb	33.6°	35.9°	60°/2pb
Nombre de pb par période	10.7	10.0	12
Inclination pb par rapport à l'axe	+19°	-1.2°	-9°
Distance entre pble longe de l'axe (rise)	2.3A°	2.3A°	3.8A°
Longueur par période (pitch)	24.6A°	33.2A°	45.6A°
Diamètre	25.5A°	23.7A°	18.4A°

Deux molécules d'ADN circulaires ayant exactement la même séquence de bases, peuvent différer entre elles par ce que l'on appelle le nombre d'enlacements (tours). On appelle alors topoisomères. Il existe deux possibilités de surenroulement.

>Surenroulement positif: le nombre d'enlacement a augmenté, l'enroulement de la double hélice s'effectue dans le même sens → super hélice droite.

>Surenroulement négatif: le nombre d'enlacement a diminué, l'axe de l'hélice s'enroule dans le sens opposé à celui de la double hélice, ceci favorise le déroulement et la séparation locale des deux brins.

5. Structure des chromosomes :

Une série d'enroulement et de surenroulement provoquent la compaction de l'ADN pour aboutir à la forme finale d'un chromosome. La compaction de l'ADN est rendue possible grâce à 4 types de protéines histones qui sont : H2A / H2B / H3 / H4. L'ADN s'enroule autour de ces protéines pour former un nucléosome ; fibrille élémentaire 100 A° de diamètre, cette fibrille subit un nouvel enroulement pour former une fibre de 300A°, chacun de ces fibres

s'enroulent sous forme de spirale (comportant 50 fibres), ces spirales se regroupent sous forme de mini bande. Le chromosome est enfin composé de plusieurs mini bandes (≈un million). Le chromosome apparaît constitué de 2 sous unités = chromatides.

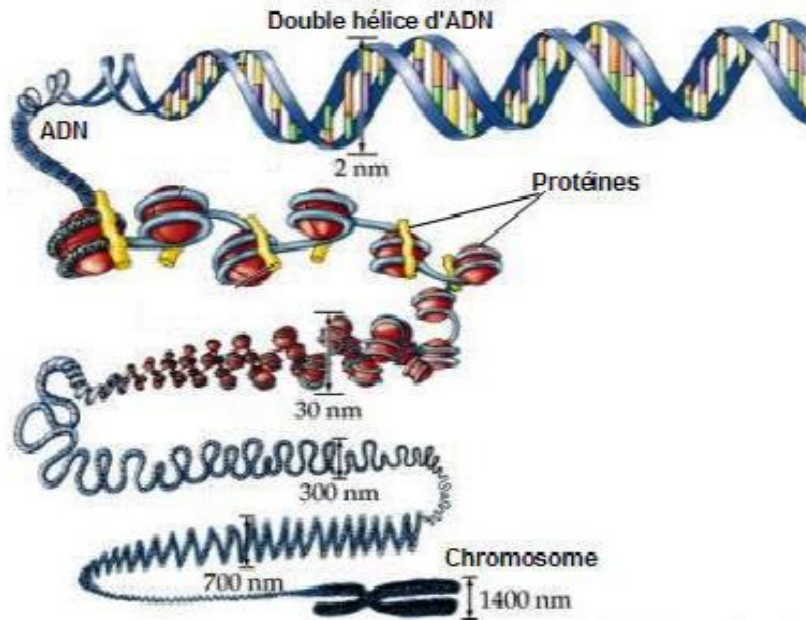


Figure 08 : Les différentes échelles de l'organisation hiérarchique du chromosome

6. Implications biologiques du surenroulement de l'ADN

Le surenroulement a des implications biologiques et joue un rôle important dans un grand nombre de processus biologiques, notamment la condensation de l'ADN. Il module en effet l'entortillement de l'ADN, la somme de la torsade et de l'entortillement correspondant à l'enlacement de la molécule, lequel a un effet sur le volume occupé par cette dernière.

La création de supertours demande de l'énergie : les supertours constituent une réserve d'énergie ($-20 \text{ kJ mol}^{-1} / 100 \text{ pb}$) qui peut faciliter les transitions structurales (la séparation des brins lors de la réplication, transcription, recombinaison requiert de $50 \text{ à } 200 \text{ kJ mol}^{-1} / 10 \text{ pb}$). Cette "réserve" est suffisante pour "aider" les enzymes à entrer.

Par ailleurs, certaines enzymes telles que les ADN topoisomérases sont capables de modifier la topologie de l'ADN pour faciliter certains processus tels que la réplication de l'ADN ou sa transcription en ARN. Il en existe deux classes :

>Les topoisomérases I: coupent un brin de l'ADN bicaténaire au niveau d'une liaison phosphodiester, l'extrémité 3'OH reste libre tandis que le phosphate en 5' estérifie momentanément une tyrosine de l'enzyme. Le brin coupé peut alors tourner librement autour du brin intact, puis la liaison phosphodiester est rétablie.

>Les topoisomérases II: ou gyrases clivent les deux brins de l'ADN et les font tourner l'un autour de l'autre pour introduire des supertours négatifs en présence d'ATP. Chez *E. coli*, la gyrase introduit plus de 100 supertours /minute, un toutes les 10 tours d'hélice. Les composés qui l'inhibent comme la novobiocine et l'acide nalidixique sont des antibiotiques puissants.

II. Structure et organisation du génome des procaryotes et eucaryotes

Les cellules procaryotes, telles que les bactéries, ne possèdent pas de noyaux et possèdent un ADN sous la forme d'un seul chromosome circulaire superenroulé, situé dans le cytoplasme et haploïdes à l'état végétatif. Cependant en plus de l'ADN génomique ou chromosomique, on trouve aussi des petits morceaux indépendants de l'ADN principal, ce sont les **plasmides** qui se répliquent indépendamment du chromosome.

On appelle virus des entités composées d'un seul type d'acide nucléique (virus à ADN et virus à ARN) inclus dans une structure protéique rigide, constituant la nucléocapside. Certains virus ont en plus une enveloppe lipidique. Ils possèdent les acides nucléiques les plus courts et parfois simple brin. Pour se reproduire, ils utilisent le matériel de la cellule qu'ils infectent. Hors de celle-ci, ils sont inertes. Leur taille est bien inférieure à celle des bactéries. Les gènes des bactéries sont généralement plus petits que ceux des eucaryotes, car ils sont "continus", alors que les gènes des eucaryotes sont discontinus (ils contiennent des régions codantes et non codantes, respectivement des exons et des introns).

Chez les eucaryotes, l'ADN est présent dans le noyau cellulaire principalement, mais aussi dans les mitochondries et les chloroplastes. Dans le noyau, il est linéaire et est scindé en plusieurs ADN formant des chromosomes. Il est plus ou moins compacté et associé à des protéines comme les histones et les **protamines**. Le chromosome soit en un seul exemplaire (pour les espèces haploïdes) ou en double exemplaire (pour les espèces diploïdes) on parle alors de paire de chromosomes homologues.

Dans les mitochondries et les chloroplastes, l'ADN peut prendre de nombreuses formes différentes, circulaires, linéaires ou encore ramifiés. Le code génétique mitochondriale et légèrement différent du code nucléaire (code des ARNr, t, et m). L'hérédité de l'ADN mitochondrial est cytoplasmique: seul l'ADN mitochondriale de la mère est transmis.

I. La mutation génique

Deux grands processus sont responsables de la variation génétique: la recombinaison et la mutation. La mutation est le processus par lequel des changements se produisent dans la séquence d'ADN d'un gène. Une mutation est rarement réversible : le plus souvent le gène muté est réparé ou détruit.

On doit distinguer entre les mutations germinales qui affectent les gamètes et les mutations somatiques qui affectent les autres cellules : les premières sont transmissibles, les autres sont une cause importante de cancers. Seules les mutations survenant dans les séquences codantes sont susceptibles d'avoir des répercussions sur le phénotype et peuvent avoir des conséquences délétères sur l'organisme.

1. Les différents types de mutation :

Une mutation peut être de forte amplitude (quelquefois visible au niveau du chromosome) ou très ponctuelle.

1.1. Mutations ponctuelles : Il en existe deux principales catégories:

a) Transition-transversion ; Purine remplacée par une purine différente ou pyrimidine remplacée par une pyrimidine différente ou l'inverse pour la transversion (A → C, A → T).

Ces mutations n'ont pas les mêmes conséquences:

- Mutation synonyme (**silencieuse**): Le triplet code le même acide aminé. le code est redondant. Exemple: AGG CGG (les deux codent Arg).
- Mutation **faux-sens conservative**: Le codon spécifie un acide aminé fonctionnellement équivalent. Exemple: AAA AGA (change la Lys en Arg basique elle aussi)
- Mutation **faux-sens non conservative**: Le codon spécifie un acide aminé fonctionnellement différent.
- Mutation **non-sens**: Le codon signale la terminaison de la chaîne. Exemple: CAG UAG (change la Gln en un codon STOP)

b) Insertion-délétion (indel) Mutation par décalage du cadre de lecture: toute addition ou délétion de paires de bases modifie le cadre de lecture et conduit à de nouveaux acides aminés à partir de ce site mutationnel et parfois à une terminaison anticipée (ou plus lointaine) de la chaîne polypeptidique.

L'altération à grande échelle la plus fréquente est la délétion. Elle peut concerner 10 bases comme plusieurs centaines de milliers de bases. Pour qu'il y ait délétion, il faut que l'ADN double brins soit rompu à deux endroits et qu'une ligation se produise entre les deux.

Il arrive également que lors de la réplication, s'il y a des séquences répétées, elles s'apparient entre elles, en formant alors une boucle d'ADN qui sera excisée.

1.2. Mutation réverse La mutation inverse peut résulter d'une réversion vraie ou d'une suppression intragénique ou extragénique. Les mêmes mécanismes qui produisent des mutations peuvent produire des réversions ou des suppressions.

• **La réversion** : C'est la mutation elle-même qui est annulée par une seconde mutation, elle provoque le retour à l'état ancestral du gène (réversion vraie) ou de la fonction (réversion équivalente) en restaurant le même acide aminé mais codé par un autre codon.

• **La suppression** : Ce sont les conséquences de la mutation qui sont annulées par une deuxième mutation. Dans la suppression de type intragénique, la mutation a lieu dans le même gène en dehors du codon muté et son effet annule les effets de la première mutation. Par contre la suppression extragénique qui a lieu en dehors du gène.

1.3. Mutation dominante négative : Cette classe de mutation agit activement en bloquant l'activité du gène provoquant ainsi un phénotype avec perte de fonction même en présence d'une copie normale du gène. Ce phénomène se produit quand le produit du gène mutant interfère avec la fonction de produit normal du gène.

1.4. Mutation conditionnelle : Mutation ne s'exprimant que sous certaines conditions (conditions restrictives). Ces mutations provoquent des effets phénotypiques. Les mutations létales ne peuvent exister à l'état homozygote dans une cellule que sous forme conditionnelle, par exemple si elles ne s'expriment qu'à certaines températures.

2. Cause des mutations géniques :

Les lésions de l'ADN sont causées principalement par trois types d'agents:

2.1. Instabilité chimique intrinsèque de l'ADN (lésions spontanées)

Des erreurs dans la réplication de l'ADN, des lésions spontanées et même des éléments génétiques transposables peuvent être à l'origine des mutations spontanées. Ces lésions ne sont pas soumises à des agents exogènes et sont ponctuelles. On observe :

➤ **Dépurations et formation de sites abasiques**: Hydrolyse spontanée des nucléotides par la chaleur, les acides et certaines enzymes (uracile N-glycolase, hypoxanthine-N-glycolase...) peuvent entraîner une perte de base par hydrolyse de la liaison N-

glycosidique → sites abasiques. Au cours de la réplication, des sites apuriques produits ne peuvent plus spécifier la base complémentaire de la purine d'origine.

- **Désaminations** : correspondent à des pertes de groupement amine sur les bases C, A et G. L'adénine est transformée en hypoxanthine, la guanine en xanthine et la 5-méthylcytosine en thymine, la cytosine donne de l'uracile.
- **Tautomérisation des bases** : Les bases peuvent se présenter sous deux formes tautomères ; Cétone (la plus fréquente), Imino (pour A et C) ou Enol (pour T et G, les plus rares) qui diffèrent par un proton et une double liaison. L'insertion d'un mauvais tautomère d'une base standard peut mener à un mésappariement, susceptible de créer une mutation au cours de la réplication.

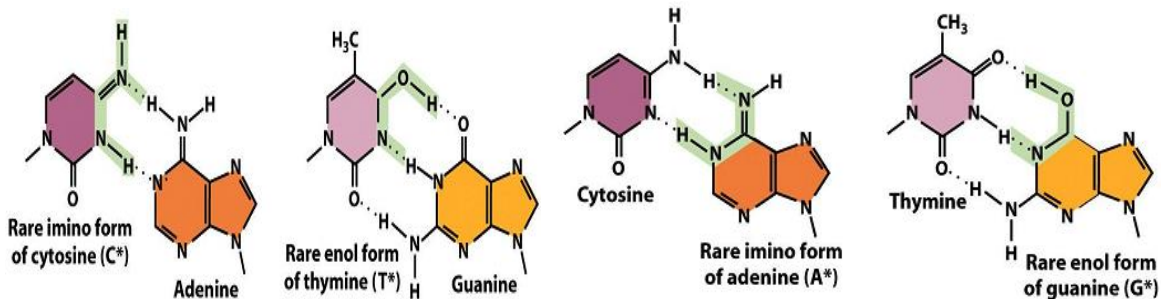


Figure 09 : Tautomérisation spontanée des bases : appariements illicites

- **Des erreurs de méthylations**, Les erreurs de méthylations donnent des alkylations sur le carbone C6 au lieu du carbone C5 entraînant des absences de formation de liaisons H entre bases au niveau des îlots CpG.

2.2. Les sous-produits du métabolisme cellulaire (lésions endogènes)

Dégâts causés par les radicaux libres de l'oxygène (O2*, OH*, H2O2) proviennent du métabolisme oxydatif. Le résultat peut être :

- Bases oxydées et hydroxylées (8-OH G; Foraminopyrimidines; Glycol de thymine)
- Cassures de brin d'ADN (simples et doubles)

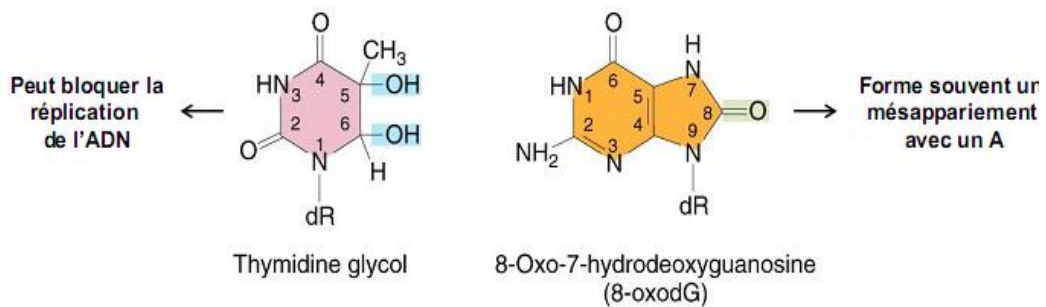


Figure 10 : Bases endommagées par des radicaux oxygène. dR=désoxyribose.

2.3. Les agents environnementaux (lésions exogènes)

Ce sont des mutations causées par des effets directionnels des facteurs physiques ou chimiques, nommés agents mutagènes

2.3.1. Les mutagènes physiques :

La température, les éléments radioactifs et les ondes électromagnétiques; l'énergie est croissante dans cet ordre : infrarouge, visible, rayons ultraviolets, rayons X et gamma. Ils peuvent provoquer des délétions et remaniements assez importants au niveau de l'ADN ; des altérations ou perte de bases, des dimères entre bases pyrimidiques et des ruptures dans l'un ou les deux brins (pouvant conduire à des réarrangements, délétions, perte de fragments de chromosome) (figure 11).

2.3.2. Les mutagènes chimiques

Des substances chimiques interagissant avec l'ADN (ou peut-être avec l'ARN) tels que pesticides, dérivés de benzène, solvants, colchicine, etc. D'autres mutagènes chimiques comprennent: les agents oxydants forts ou agents réducteurs (ex, nitrates, nitrites, ERO), certains additifs alimentaires (ex, hydrocarbures aromatiques, cyclamates), produits du raffinage du pétrole, les solvants organiques, les médicaments (ex, les cytostatiques, les immunosuppresseurs), les plus répandus sont les pesticides (ex, herbicides, fongicides) et les déchets industriels.

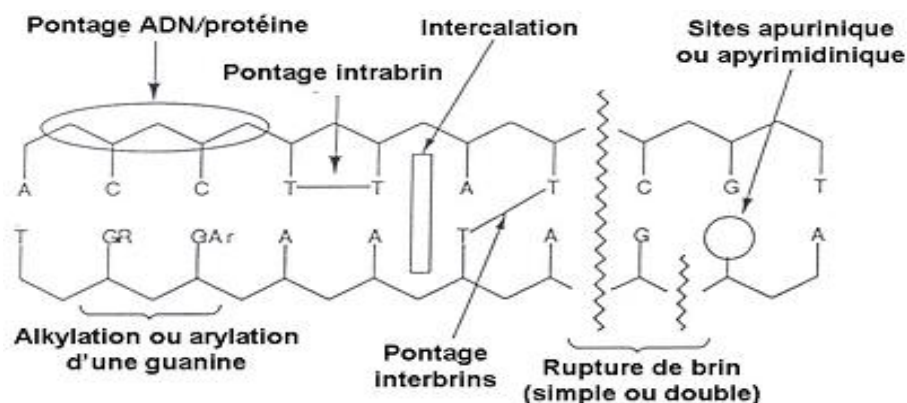


Figure 11: Principales lésions de l'ADN

3. Les conséquences et les effets des lésions

La conséquence de toute mutation dépend de son effet fonctionnel, qui peut être neutre, conduire à l'amélioration d'une fonction (diversité, évolution) ou à l'altération d'une fonction

(effet pathogène). Une anomalie de la séquence codante d'un gène, conduisant à une anomalie de la séquence en acides aminés de la protéine codée par ce gène, responsable d'une altération qualitative et/ou quantitative de la protéine. Les conséquences délétères des microlésions du génome dépendent essentiellement de leur localisation et de leur type, elles sont classées en deux grandes catégories : la perte et le gain de fonction.

- Perte de fonction: par un effet délétère dû à la diminution ou absence de production de la protéine active et/ou par effet dû à la fonctionnalité de la protéine (présence d'un produit moins actif voir inactif) constitue la cause majoritaire des maladies récessives.

La perte de fonction peut résulter d'un effet délétère au niveau de l'ARN messager : par effet sur la régulation de la transcription, par altération de la maturation de l'ARNm (notamment l'épissage), par altération de la stabilité de l'ARNm entraînant sa destruction. Ceci va conduire à une diminution ou absence de production de la protéine active.

- Le gain de fonction : résulte habituellement d'un effet délétère dû à une surproduction de protéines (sur-expression conduisant à un excès de produit), ou par la présence d'un produit plus actif ou alors doué d'une nouvelle propriété, qui remplace l'activité antérieure ou s'ajoute à elle. Le produit se révélant toxique ou responsable d'un effet pathologique, il s'agit de la cause majoritaire des maladies dominantes.

L'effet délétère peut aussi se situer au niveau de l'ARN messager, entraînant par exemple la séquestration d'ARNm mutés, ou la séquestration de protéines par des ARNm mutés.

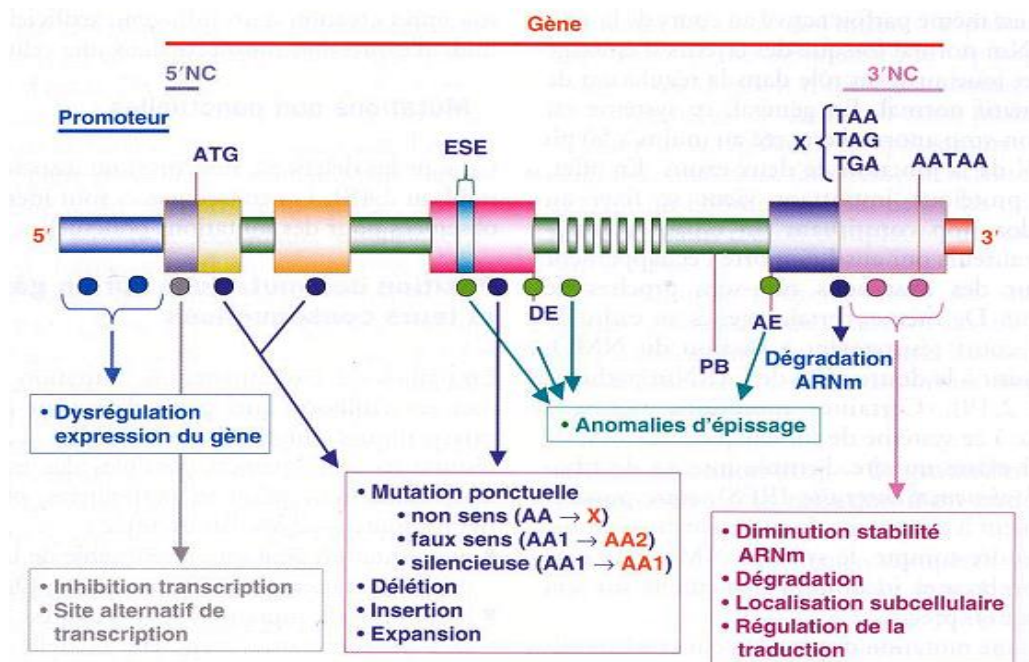


Figure 12 : Position des mutations sur un gène et leurs conséquences

II. MUTAGENESE

Procédé, naturel ou artificiel, consiste en l'induction d'une ou plusieurs mutations dans un génome, par utilisation d'agents physiques ou chimiques (mutagenèse aléatoire), par un virus intégratif ou à un élément transposable, ou par utilisation d'une technique de biologie moléculaire, ou par utilisation de petits segments d'ADN (mutagenèse dirigée).

Le but de la mutagenèse est de modifier l'activité biologique qui en découle, l'étude de gènes spécifiques (informations sur le rôle et la fonction) et produire de la diversité génétique et des nouveaux génotypes d'intérêt (amélioration des plantes, sélection animale, bactéries). La mutagenèse joue un rôle central dans le domaine de thérapie génique. Le terme de mutagenèse peut être associé à trois modifications distinctes d'un fragment d'ADN :

- la délétion, qui consiste à éliminer des nucléotides du fragment d'ADN d'intérêt ;
- l'insertion, qui consiste à en rajouter ;
- la substitution, qui consiste à remplacer une ou plusieurs bases par un même nombre de bases de nature différente.

Deux méthodes sont possibles, la mutagenèse aléatoire et la mutagenèse dirigée/ciblée.

1. Mutagenèse aléatoire : Consiste à utiliser un agent mutagène (physique, chimique ou biologique), qui induira de manière aléatoire des mutations dans le génome. L'endroit et la nature des mutations ne sont pas prévisibles et ne peuvent pas être contrôlés.

1.1. Mutagenèse physique : Comprennent:

- Température excessivement élevée ou basse: son effet sur l'ADN se traduit par l'élimination de la liaison N-glycosidique entre les purines et leur sucre, il en résulte un site un site apurinique au niveau de l'ADN.
- Les éléments radioactifs (ex. le radium, le radon, les isotopes du potassium, du carbone, etc.)
- Les rayonnements ionisants et ultraviolets (Les rayons X, les neutrons..).

1.1.1. Radiations ionisantes

Les rayonnements du spectre électromagnétique. Les rayons X et gamma γ sont assez énergétiques pour produire des ions réactifs quand ils interagissent avec les molécules

biologiques. Les radiations agissent par action directe ou en provoquant la production de radicaux libres qui peuvent rompre la double hélice. Au niveau de l'ADN, il s'ensuit:

- Des ruptures dans l'un ou les deux brins (qui peuvent conduire à des réarrangements, délétions, perte de fragments de chromosome, ou la mort de la cellule en l'absence de réparation).
- Une altération ou perte de bases (mutations).
- Un enchevêtrement de l'ADN sur lui-même ou avec des protéines.

1.1.2. Radiations non-ionisantes : ultra-violet (UV, environ 260nm)

Leurs longueurs d'onde sont absorbées préférentiellement par les bases de l'ADN (plus utilisés = UV-B) pour former des dimères cyclobutanes ou des photoproduits (6-4) au niveau de deux bases pyrimidiques adjacentes bloquant la transcription et la réplication.

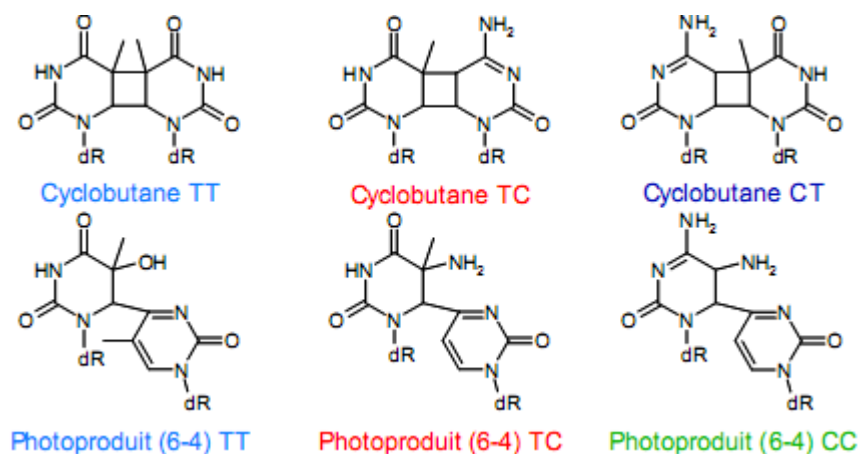


Figure 13: Exemples de lésions photochimiques sur des bases pyrimidiques

1.2. Mutagenèse chimique

Provoque généralement des mutations ponctuelles ou des petites délétions et permet d'obtenir une forte densité de mutations («saturation» des génomes). Les différents modes d'action :

- Altération de la structure et l'appariement des bases :

- désamination (ex: acide nitreux, hydroxylamine)
- alkylation : rajout d'un groupe méthyl ou éthyl par alkylsulfonates (éthyl-méthane-sulfonate EMS) ou nitroso-urées (éthyl-nitroso-urée ENU)

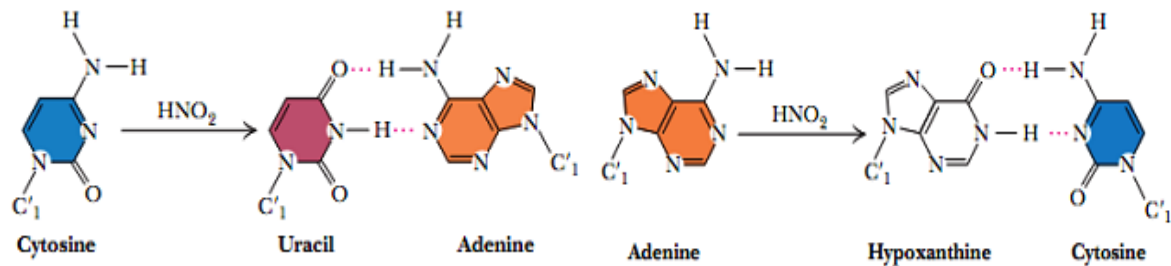


Figure 14: Mésappariements dus à la désamination des bases par l'acide nitreux

- Analogues des bases Exp: 5-bromouracile incorporée à la place d'une T. Il a une forte tendance à se tautomériser en « enol ». Elle s'apparie alors à G. La 2-amino-purine est un analogue de A. Elle provoque des transitions A-T en G-C.

- Agents intercalants tels que l'acridine, la proflavine et le bromure d'éthidium, sont des molécules qui s'insèrent entre les plateaux de paires de bases, la polymérase insère alors une base surnuméraire en face de la molécule étrangère.

- Altération de la structure de l'ADN : telles que la liaison de grosses molécules à des bases qui deviennent ainsi « non codantes », l'introduction de liaisons intra et inter brins (ex : le psoralène, G-G cis-platine, mitomycine; chlorambucil) ou des ruptures dans l'ADN (ex : peroxydes) n'induisent pas directement des mutations mais sont à l'origine de processus de réparation qui sont mutagéniques.

1.3. Mutagénèse biologiques Comprennent:

- Certains virus (ex, la rougeole, la rubéole, la grippe)
- Les produits métaboliques ex, les produits de peroxydation lipidique notamment les aldéhydes comme le 4-hydroxy-nonén-2-al, ou les dérivés chlorés issus de l'inflammation.
- Les antigènes de certains microbes et parasites.

Les agents mutagènes utilisés pour la mutagenèse aléatoire par insertion sont: les transposons (rétro-transposons (ex: Ty, copia, Tnt1, SINE, LINE) transposons à ADN (ex: Tn, Tc1, élément P, Ac/Ds) rétrovirus (virus à ARN codant pour une transcriptase inverse) ADN-T d'Agrobacterium (plantes, fungi).

2. La mutagenèse dirigée : consiste à introduire des mutations spécifiques et ciblées au niveau d'un ADN plasmidique double brin. Elle présente les applications suivantes :

- Étude des changements d'activité protéique résultant de la modification de l'ADN ;

- Sélection ou criblage de mutations (au niveau de l'ADN, de l'ARN ou de la protéine) présentant une propriété recherchée ;
- Ajout ou suppression de sites d'endonucléase de restriction ou de tags.

Les premières techniques de mutagenèse dirigée étaient basées sur l'amplification d'un fragment d'ADN linéaire, qui devait ensuite être introduit dans un plasmide par clonage en utilisant des enzymes de restriction. De multiples approches sont ensuite possibles pour obtenir des modifications génétiques localisées du génome dans lesquelles un ou plusieurs morceaux d'ADN sont insérés, remplacés ou retirés d'un génome permet une étude fine de leur fonctionnement. Plusieurs techniques de transformation génétique existent :

2.1. Techniques d'édition du génome,

Technique de génie génétique reposant sur l'utilisation d'endonucléases de restriction, y compris les nucléases à doigts de zinc (ZFN), TALENs, CRISPR /Cas, les méganucléases et la mutagenèse dirigée par des oligonucleotides (ODM), pour la re-écriture des parties du génome en supprimant, en remplaçant ou en ajoutant des séquences d'ADN dans des endroits prédéfinis. Le double brin d'ADN modifié est ressoudé par recombinaison homologue ou par jonction d'extrémités non homologues. L'insertion de séquences génétiques par recombinaison homologue après coupure double brin de l'ADN permet de corriger, modifier ou inhiber la séquence génétique cible.

2.1.1. Les nucléases à doigt de zinc (ZFN, zinc finger nucleases)

Ces protéines artificielles permettent de couper une séquence particulière d'ADN. Le « doigt de zinc » en tant que composant du ZFN peut reconnaître une section courte d'ADN (9-12 bases) et l'autre composant, la nucléase (N) coupe à travers l'ADN à ce point. Cette coupure déclenche un mécanisme de réparation de l'ADN par la cellule.

2.1.2. Les nucléases effectrices de type activateur de transcription (TALEN, *Transcription activator-like effector nuclease*)

Ce sont des protéines sécrétées par la bactérie phytopathogène *Xanthomonas* de la famille des *Pseudomonadaceae*. Également utilisés par paires, ciblant deux séquences d'ADN proches. Ils comprennent un assemblage unique d'une nucléase qui assure la coupure double brin avec une protéine TAL, qui reconnaît une séquence particulière d'ADN et agit comme guide. Les TALENs sont plus faciles à produire et présentent une très bonne efficacité.

2.1.3. La technique CRISPR/Cas

Elle est basée sur la capacité des ciseaux moléculaires (la nucléase Cas) à déclencher une coupure à double brin dans l'ADN à un endroit dont la précision est définie par un ARN guide (ARNg), appelé CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats), qui s'amarre à la séquence d'ADN recherchée qui lui est identique. Cet outil est utilisé pour couper l'ADN qui sera suivie par une réparation de l'ADN au travers de laquelle il sera possible d'insérer ou de supprimer des séquences génétiques plus ou moins longues et ainsi modifier les propriétés d'un gène. Cet outil possède plusieurs atouts par rapport aux meilleures nucléases ZFN et TALEN: la simplicité, la rapidité et beaucoup moins coûteux.

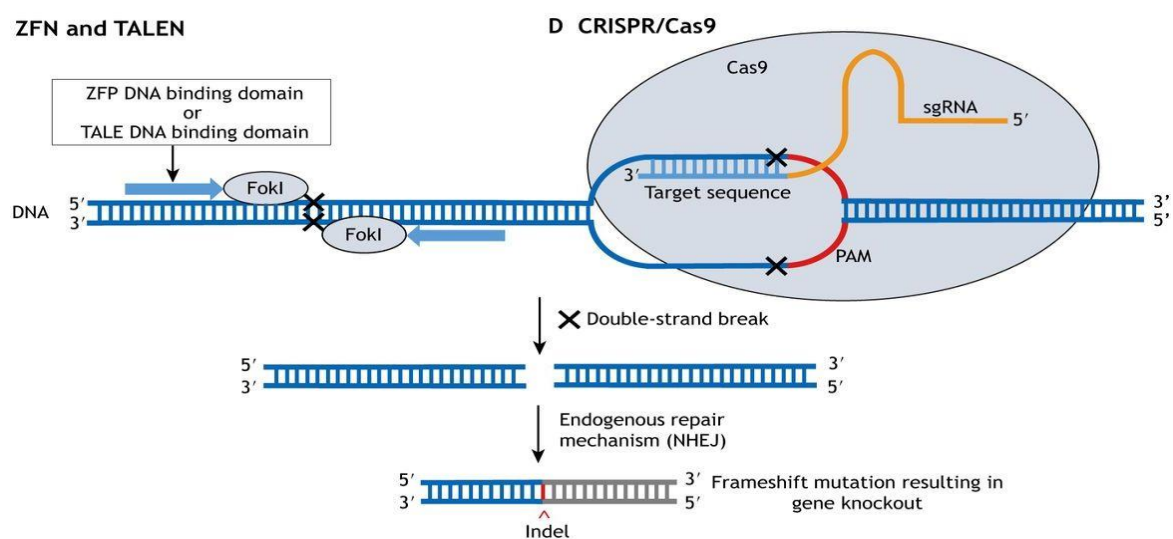


Figure 15: Mutagénèse ciblée /Nucléase à doigt de zinc (ZFN), TALENs, et CRISPR/Cas9

2.1.4. Les méganucléases:

Contrairement aux ZFN et TALEN, ces protéines sont des enzymes de restriction extrêmement spécifiques, capables de reconnaître et de cliver une séquence d'ADN très longues en s'assemblant par paire de sous-unités identiques (homodimères). Mais ces séquences sont définies de façon naturelle et ne se prêtent pas à une modification par ingénierie génétique, leur utilisation est très limitée.

2.1.5. Introduction transitoire d'ADN recombinant/Mutagénèse dirigée par oligonucléotides (ODM)

Dans ce processus, il n'y a pas de coupure d'ADN; des segments synthétiques d'ADN ou des hybrides ADN/ARN (oligonucléotides), presque identiques à la séquence d'ADN du gène

cible sont introduits dans la cellule. La présence de ces oligonucléotides modifiés active également les mécanismes de réparation cellulaire qui déclenchent des mutations aux sites du génome qui présentent une similarité avec cet oligonucléotide synthétique.

2.2. Cisgénèse et intragénèse

Insertion, dans l'ADN d'une espèce, d'un fragment d'ADN – intact – provenant de la même espèce ou d'une espèce sexuellement compatible. Dans le cas de la cisgénèse, les gènes insérés, les introns associés et les éléments régulateurs sont contigus et inchangés. Dans le cas de l'intragénèse, l'ADN inséré – remanié par l'homme – peut être une nouvelle combinaison de fragments d'ADN de l'espèce elle-même ou d'une espèce compatible (figure 16). Les deux approches visent à conférer une nouvelle propriété à la plante modifiée.

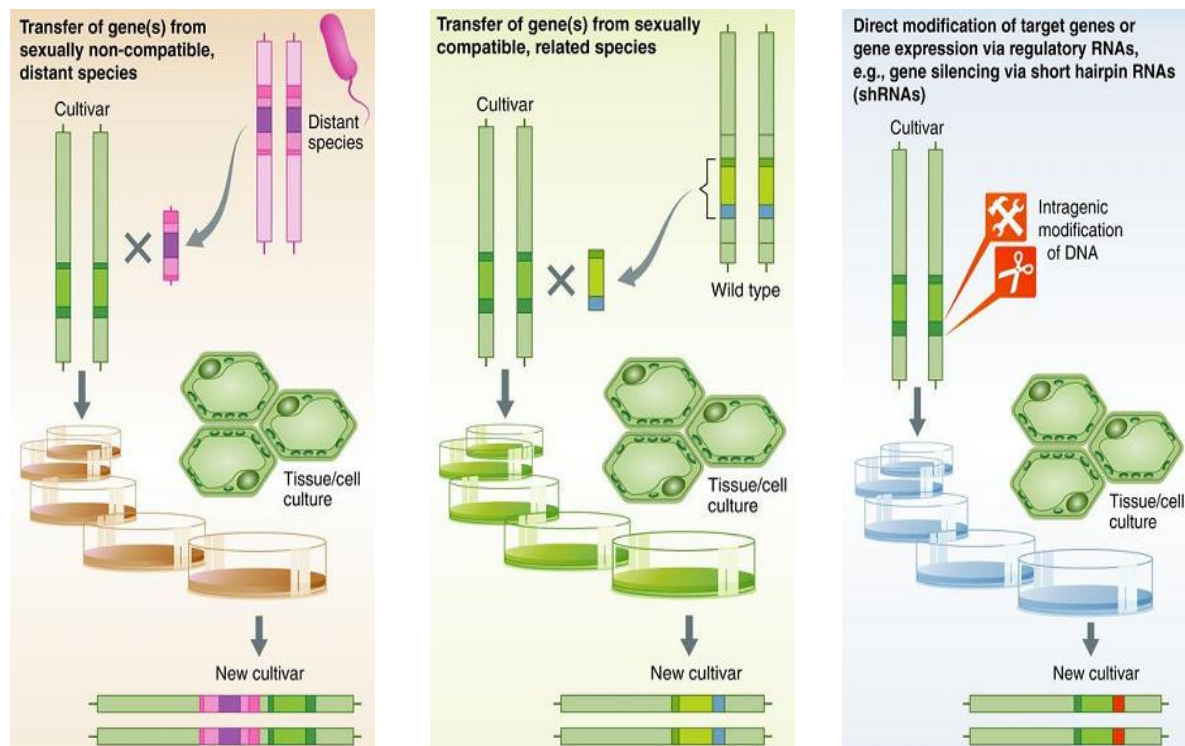


Figure 16: Transgénèse, cisgénèse et intragénèse

2.3. Méthylation de l'ADN dépendante de l'ARN (RdDM)

Silence des gènes spécifiques d'une manière qui disparaît généralement après plusieurs générations, aucune modification ou mutation n'est effectuée dans la séquence nucléotidique l'expression génique est modifiée par épigénétique. La RdDM induit le silence de gène transcriptionnel (TGS) des gènes ciblés via la méthylation de séquences promotrices par

exemple le retardement du murissement de la tomate, la production d'insecticide dans les pommes.

2.4. Agro-infiltration Livrer du matériel génétique à une plante transitoirement, pour un maximum d'une génération par exemple un vaccin, la production d'anticorps; la recherche dans des plantes modèles. Utilisation d'une bactérie *Agrobacterium tumefaciens* comme une seringue pour injecter de l'ADN étranger dans une cellule. Cet ADN étranger ne s'intègre pas dans l'ADN hôte. L'ADN étranger, après production de certaines protéines, sera éliminé par la cellule.

2.5. Sélection inverse

Gènes sous silence en charge de la recombinaison génétique dans le processus de reproduction sexuelle (comme une étape dans le processus global) création d'hybrides par exemple dans le maïs, dans les arbres fruitiers. Elle facilite la production de lignées parentales homozygotes qui, une fois hybridées, reconstituent la composition génétique d'une plante hétérozygote d'élite, sans qu'il soit nécessaire d'effectuer un rétrocroisement et une sélection.

2.6. Greffage

Combinez un scion non-OGM avec un OGM rhizome (ou vice versa), le composant végétal supérieur d'une plante (scion) est attaché à un composant inférieur enraciné (porte-greffe) d'une autre plante pour produire un organisme chimère avec des caractéristiques de culture améliorées. Par exemple, arbres à fruit résistant aux maladies

III. DIAGNOSTIC GENOTYPIQUE

Les altérations diverses du patrimoine génétique, créant un immense éventail de mutations. Ces mutations, affectant l'intégralité et la fonctionnalité de la protéine, peuvent être responsables de maladies génétiques, ou modifier le phénotype du porteur. Le test diagnostique permet d'évaluer l'ampleur de ces anomalies dans chaque génome.

Les tests diagnostiques sont utilisés pour identifier les porteurs d'une maladie héréditaire récessive (évaluer le risque de transmission d'une maladie génétique à leur descendance), d'obtenir le diagnostic d'une condition génétique soupçonnée ou d'évaluer la prédisposition d'un individu à développer une maladie génétique et d'identifier les facteurs de risque génétiques prédisposant à la survenue de maladies multifactorielles dans un but de prévention

1. Le diagnostic de maladies pré-symptomatique :

- **Pour les maladies monogéniques**

Les tests génétiques prédictifs sont effectués chez des personnes asymptomatiques, afin de prédire le risque de développer ultérieurement une maladie. La mucoviscidose des nouveau-nés; révélation précoce, la Chorée de Huntington ; révélation tardive. Ces tests peuvent donner des informations importantes sur la prise en charge de l'individu ou éclairer certaines décisions.

- **Pour les maladies multifactorielles**

Il existe des tests prédictifs en cancérologie, apportent une indication concernant le risque de développer la maladie lorsqu'une mutation a déjà été identifiée dans la famille: les facteurs environnementaux et personnels contribuent largement à la survenue d'un cancer, et les mutations génétiques recherchées lors de ces tests ne sont insuffisantes à l'apparition d'un cancer (BRCA1 et BRCA2 pour le cancer du sein). L'identification de cette susceptibilité permet une surveillance plus précoce et rigoureuse des sujets à risque (l'ablation des seins et des ovaires recommandée en fonction de l'âge et du projet parental). De même, un test de susceptibilité aux cancers colorectaux héréditaires sans polypose (HNPCC ou syndrome de Lynch) est également disponible.

2. Approches de diagnostic génotypique

On distingue classiquement deux approches de diagnostic génotypique :

2.1. Diagnostic direct

Le diagnostic direct est le plus spécifique, consiste à rechercher les anomalies qui sont au niveau du gène directement responsables de la maladie; criblages (screening). La stratégie utilisée dépendra des connaissances des types de défauts moléculaires connus pour le gène en cause. On a deux cas : soit existence de mutations fréquentes ou hot spots, soit mutations familiales privées (pas de hot spot). L'ensemble de ces techniques est appliqué à des fragments de gène amplifiés par PCR ce qui permet de copier en grand nombre une séquence ciblée d'ADN.

2.1.1. Détection de mutations connues (hot spot)

Les mutations y seront préférentiellement recherchées avec les techniques les mieux adaptées (Southern ou amplification). Les méthodes de criblage ont pour objectif de pouvoir cribler rapidement tout le gène pour y repérer la région où est localisée la mutation. Seule cette région sera séquencée.

- **PCR avec migration sur gel (Polymerase Chain Reaction)**

Technique moléculaire permettant de mettre en évidence une perte. Cette technique permet l'amplification en une seule réaction de plusieurs séquences d'ADN distinctes cibles (amplicons), de longueurs différentes. Les amplicons sont séparés par électrophorèse en fonction de leur taille et comparés aux références. La PCR multiplexes fluorescentes permet de mettre en évidence un panel de mutations différentes (exp dans la mucoviscidose). On fait intervenir soit un agent intercalant (bromure d'éthylum qui est révélé à la lumière UV), soit des composés fluorescents.

- **Southern Blot**

Permet de visualiser une portion du génome par hybridation d'une sonde marquée sur des fragments de restriction d'ADN séparés par électrophorèse. Recherche de grand délétions ou amplifications (exemple la maladie de l'X fragile).

- **Séquençage direct selon la technique de Sanger**

Permet de visualiser une portion du génome par hybridation d'une sonde marquée sur des fragments de restriction d'ADN séparés par électrophorèse elle permet la recherche de mutations ponctuelles, petites insertions ou délétions. Repose sur l'incorporation de

didésoxynucléotides (ddNTPs) qui bloquent la synthèse de l'ADN par les ADN polymérase. Les nucléotides ddNTPs marqués par des fluorochromes permet d'obtenir un ensemble de fragments d'ADN de différentes tailles, correspondant aux emplacements d'un nucléotide donné. Les échantillons sont soumis à une électrophorèse sur gel et lus à travers rayon x (des bandes noires apparaissent). ~100% de sensibilité, le cout est beaucoup diminué.

2.1.2. Détection de mutations dans un gène sans hot spot

Pas de mutation principale mais un grand nombre de mutations rares, il faut donc analyser tout le gène, sans préjuger de leur localisation ou de leur nature ; c'est le cas de la majorité des maladies, il existe une très grande hétérogénéité allélique.

Les techniques de balayage ont pour objectif d'identifier une différence de comportement électrophorétique entre l'allèle normal et l'allèle muté. Pour cela, soit l'on procédera à une modification chimique ou enzymatique avant la migration électrophorétique (technique à la RNase, modification par carbodiimide, formation d'hétéroduplex) soit l'on jouera sur les conditions électrophorétiques (gradient de dénaturation, gradient de température, gel non dénaturant pour la technique SSCP...), suivies d'une caractérisation par séquençage des produits identifiés comme anormaux

- **Technique DGGE (Denaturing Gradient Gel Electrophoresis)**

La perturbation de séquence induit une modification de la température de fusion d'un produit PCR (ADN double brin). La modification de T_m par une mutation ponctuelle est mise en évidence par électrophorèse sur gel de polyacrylamide en présence d'un gradient d'agent dénaturant (urée et formamide). Les premiers complexes à se séparer si on les chauffe sont les hétéroduplex (hétérozygotes), on repère donc les fragments qui se dénaturent plus rapidement et on les séquences. Sensibilité de 95% pour des fragments de taille inférieure à 500 pb.

- **Technique dHPLC (Denaturing High Performance Liquid Chromatography)**

Méthode chromatographique en phase liquide, établissement de profils spécifiques de migration/élution selon le type de mutation mais ne détecte pas les homozygotes mutés. Rapide, permet la détection de mutations de fragments de 500 pb à partir d'ADN génomique avec une sensibilité proche de 100%. Cette technique.

- **Technique HRM (High Resolution Melting) :**

Même principe que DDGE mais on utilise des composés fluorescents qui s'intercalent dans l'ADN double brin. La région de l'ADN potentiellement porteuse de la mutation est d'abord amplifiée par PCR. L'amplicon est alors chauffé les deux brins d'ADN se séparent. En l'absence de double brin, ils ne peuvent pas se fixer et leur niveau de fluorescence est faible. La courbe de dénaturation diffère de celle d'une séquence témoin. Séquençage direct par la méthode de Sanger uniquement pour les amplicons aberrants.

- **Le Protein Truncation Test (PTT)**

Le PTT, est un test spécifique permettant de détecter des mutations non-sens (responsables de l'apparition de codon stop prématuré (CSP)) ou qui décalent le cadre de lecture de l'ARNm (mutations frameshift) et provoquent une terminaison de traduction prématurée donc de produits protéiques tronqués. Il est fondé sur l'analyse de produits polypeptidiques RT-PCR qui sont transcrits et traduits in vitro. Le principe de base de la technique est la différence de taille (et par conséquent de migration électrophorétique) entre une protéine normale et une variante mutante, tronquée. La taille des exons étant généralement de l'ordre de 100 à 300 pb.

- **L'électrophorèse d'ADN hétéroduplex (HMA et CSGE)**

HMA (Heteroduplex Mobility Assay), l'ADN hétéroduplex correspond à l'ADN double-brin obtenu après dénaturation et renaturation d'un mélange de fragments homologues d'ADN normal et mutant. A l'endroit d'une mutation, cet ADN hétéroduplex contiendra des mésappariements. Sensibilité de 90% ont été rapportés pour des fragments de 300 pb.

La méthode CSGE (Conformation Sensitive Gel Electrophoresis) permet de détecter les mésappariements dans des fragments de plus grande taille (figure 17).

- **La technique SSCP (Single-Strand Chain Polymorphism)**

Le polymorphisme de conformation de l'ADN en simple brin (SSCP) est basé sur l'analyse électrophorétique des produits PCR d'ADN normaux et mutés qui sont dénaturés sous forme de fragments simple brin. Une mutation ponctuelle au sein d'une séquence modifie suffisamment la structure secondaire de l'ADN monobrin pour qu'il en résulte des changements de migration électrophorétique sur gel de polyacrylamide non dénaturant. La visualisation par marquage radioactif mais une détection par coloration argentique ou au bromure d'éthidium a également été utilisée avec succès. Sensibilité: 60-80% de mutations détectées, limitée à l'analyse de courts fragments d'ADN 200pb.

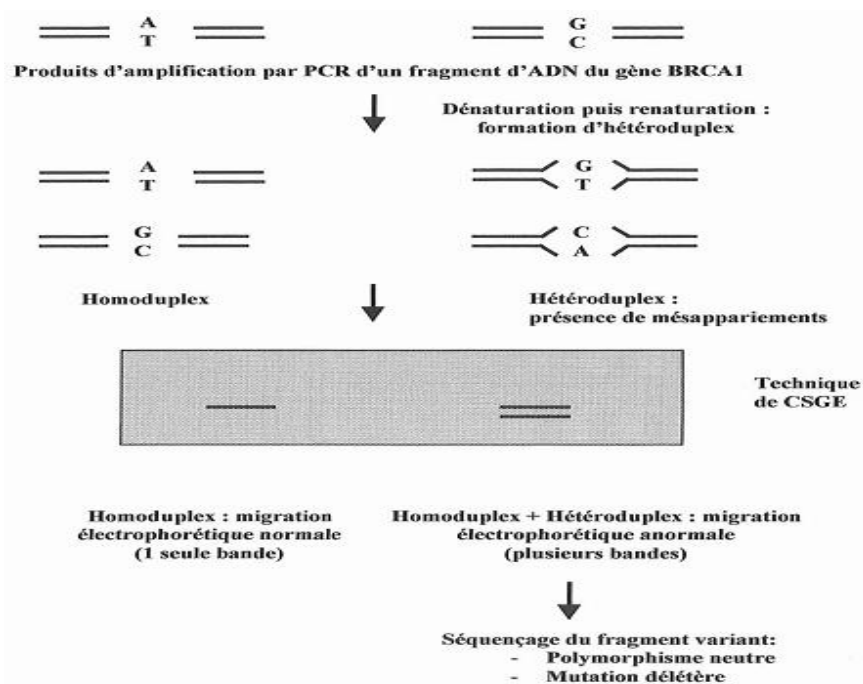


Figure 17: Stratégie de recherche de mutations par technique de CSGE puis séquençage.

2.2. Diagnostic indirect

L'approche indirecte s'appuie sur l'identification de l'allèle du gène porteur de la mutation familiale. Différentes stratégies seront utilisées selon la taille du gène, l'existence ou non de phénomènes de recombinaison intragénique, ou si le gène est seulement localisé. Utilisation de marqueurs pour analyser la coségrégation d'un phénotype avec un allèle particulier dans une famille, permet de suivre indirectement la transmission d'une mutation. Plusieurs catégories de marqueurs :

- SNP (Single Nucleotide Polymorphisms) : se trouvant dans les régions codantes (SNPc) et dans les régions régulatrices des gènes seront particulièrement intéressants pour réaliser la cartographie des maladies multifactorielles (étude d'association de gènes candidats impliqués dans ces maladies).
- RFLP (Restriction Fragments Length Polymorphism) : variation entre individus ou souches dans le profil d'ADN obtenu après coupure par diverses enzymes de restriction. Le polymorphisme de taille des fragments de restriction (RFLP) reflète directement des variations de séquence de segments précis d'ADN et est utilisé comme marqueur sur les cartes génétiques et physiques.
- Variable Number Tandem Repeats (VNTR)

- STS (Sequenced Tag Site) : C'est une courte séquence représentée de façon unique dans le génome. Il est facilement amplifiable par PCR et est archivé sous forme des amorces oligonucléotidiques qui le définissent. Certains d'entre eux correspondent à des EST (séquences codantes) ou à des marqueurs microsatellites, ce qui, dans ce dernier cas, permet d'intégrer la carte physique à la carte génétique.
- Microsatellites ou Short Tandem Repeats (STR) : répétitions de 5 nucléotides maximum (di ou tri-nucléotides) en tandem situés tout le long du génome (environ 1 tout les 30 à 40 kb), très polymorphes (nombre de répétitions différentes dans la population générale), facilement analysables par PCR. Marqueurs de choix pour le diagnostic indirect des maladies génétiques et autre utilisation pour les empreintes génétiques, médecine légale et le typage moléculaire de certains cancers et pour construire une carte génétique.

Les principaux problèmes posés:

- nécessite de connaître le gène impliqué (ou le locus) pour choisir les marqueurs à utiliser (problème posé si hétérogénéité génétique)
- nécessite une certitude du diagnostic clinique
- nécessite une famille « informative »: parfois non concluant (la famille recrutée soit suffisamment grande)
- risque d'erreur par recombinaison méiotique : les crossing-over ou recombinaison pouvant existé entre le marqueur et le gène qu'il cible peuvent induire des erreurs diagnostiques.

I. La réplication de l'ADN et sa régulation

1. Généralité

La réplication de l'ADN est un mécanisme complexe au cours duquel la quantité du matériel génétique cellulaire double. Elle se déroule au cours de la phase S (synthèse) de l'interphase, l'ADN est alors en double exemplaire dans la cellule mère pour que chaque cellule fille reçoive une copie complète de l'ADN.

- La réplication est un processus semi-conservatif (Démontrer par les expériences de Taylor (1957- incorporation de thymidine ^3H) et de Meselson et Stahl (1958-incorporation de ^{15}N): chaque brin d'ADN sert de matrice pour la formation d'un nouveau brin complémentaire (brin synthétisé ou néoformé). C'est la complémentarité des bases des nucléotides qui assurera la précision du processus de réplication de l'ADN.
- La réplication est une réaction rapide : 1000 nucl/sec/fourche de réplication chez les bactéries et 50 nucl/sec chez les mammifères. Chez ces derniers, la réplication s'accompagne de la synthèse protéique des histones qui s'assemblent pour former la structure chromatinienne. Cette étape supplémentaire ralentit la vitesse de polymérisation.
- La réplication est une réaction asymétrique qui s'effectue uniquement dans le sens $5' \rightarrow 3'$. Cette asymétrie est liée à l'activité des ADNs polymérase qui ont besoin de certains composés :
 - les nucléotides triphosphates (dATP, dGTP, dTTP, dCTP) en présence des ions Mg^{+2} ;
 - nécessité d'avoir une amorce ;
 - le dernier résidu de l'amorce doit être avoir un groupement hydroxyle libre ($3'\text{OH}$) ;
 - nécessité d'avoir une matrice.

Au cours de la réplication, des modifications du génome peuvent apparaître comme les mutations ou des phénomènes de recombinaison génétiques. La cellule possède des mécanismes de contrôle, pour limiter ces erreurs. Ces mécanismes peuvent être complétés par des mécanismes de réparation de l'ADN pour corriger ces erreurs

2. Expériences de Meselson et Stahl (1958)

Meselson et Stahl ont mis au point un protocole permettant de distinguer l'ADN nouvellement répliqué par rapport à l'ADN « ancien » en utilisant des isotopes marqueurs. L'expérience a permis de démontrer que la réplication s'effectue par un mécanisme semi-conservatif (les deux autres hypothèses impliquaient des modèles conservatif ou dispersif).

Pour savoir quel modèle est le bon, Des bactéries cultivées depuis longtemps en présence de molécules azotées lourde (^{15}N) sont repiquées sur un milieu contenant des

molécules azotées léger (^{14}N) et permettant la synchronisation des divisions. Des fractions sont prélevées après différents temps correspondant à 1, 2, 3... divisions. L'ADN est extrait, placé dans la solution de chlorure de césium (CsCl) et centrifugé. La position des ADN est repérée par une mesure de la densité optique. Cette manipulation permet de séparer les molécules d'ADN selon leur poids, l'ADN forme une bande qui est localisée d'autant plus près du fond du tube que la molécule contient de l'azote lourd (figure 18).

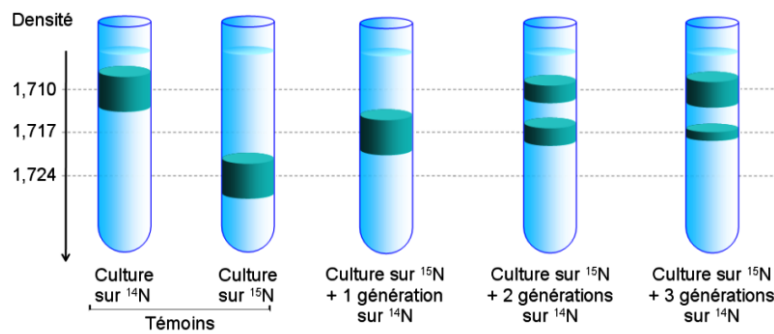


Figure 18 : Expérience de Meselson et Stahl

Après 1ère génération (donc première réplication de l'ADN), tout l'ADN est **hybride** (contenant ^{14}N et ^{15}N). Ensuite, après la deuxième réplication, l'ADN hybride disparaît progressivement au profit d'ADN "léger" (^{14}N). Cette configuration ne peut correspondre que à l'hypothèse du modèle semi-conservatif.

3. Mécanisme de la réplication chez les procaryotes

En 1960, par une expérience de Pulse-Chase en thymidine ^3H , Cairns a montré par de techniques isotopiques que la réplication s'effectuait de manière bidirectionnelle, se fait à partir des deux brins anciens, en un point du chromosome bactérien à partir d'une (procaryotes) ou de plusieurs origines de réplication (eucaryotes). Chaque origine de réplication est donc constituée de deux fourches de réplication partent d'une origine donnée et progressent jusqu'à ce qu'elles rencontrent les fourches parties d'une origine adjacente. Chaque région copiée par une origine de réplication s'appelle un replicon.

On distingue ainsi au niveau de chaque fourche de réplication, un brin précoce qui subit une synthèse continue et un brin tardif dont la synthèse s'effectue de manière discontinue dans un sens opposé à celui de la progression de la fourche de réplication.

On distingue trois étapes au cours de la réplication chez les procaryotes:

3.1. L'initiation

Chez E. Coli, il n'y a qu'un seul chromosome, et environ $4,6 \cdot 10^6$ paires de bases. Il existe une seule origine de réplication (appelée ori C: environ 245 pb), contenant des séquences répétées riches en A et T. L'ouverture de la double hélice sur 40 pdb est permise

par reconnaissance de l'origine de réplication par les DNA A. Les hélicases (DNA B) se mettent alors en place pour permettre le déroulement des deux brins; en rompant les liaisons hydrogène (ponts hydrogènes) entre les deux brins de la double hélice d'ADN permettant ainsi la progression de la réplication. Localement on va avoir le passage à la forme simple brin, c'est l'oeil de réplication. La séparation des deux brins d'ADN nécessite aussi une autre protéine appelée DNA C qui sera relâchée ou libérée juste après la fixation de la protéine DNA B au niveau de l'origine « ori C ».

D'autre part les protéines SSB (SSB: Single Strand Binding protéine) vont stabiliser les deux brins d'ADN séparés pour les empêcher de se réenrouler. De part et d'autre des deux fourches de réplifications, un gyrase (topoisomérase de classe II) enlève les supertours positifs engendrés par la formation de l'œil de réplication.

Enfin, une ARN primase va se fixer au niveau de l'ADN hélicase et constituer ainsi le primosome. Cette enzyme va synthétiser de courtes séquences d'ARN (10 nds.) qui vont par la suite servir d'amorce (primers) à l'ADN polymérase III. Ces enzymes, en s'accrochant aux amorces d'ARN, commenceront à allonger ou synthétiser la chaîne polynucléotidique (nouveau brin d'ADN) complémentaire au brin d'ADN matrice.

3.2. L'élongation

Le sens de propagation de la réplication doit s'effectuer dans le sens d'ouverture de la boucle d'ADN ($5' \rightarrow 3'$). La réplication est donc discontinue sur un des deux brins. L'ADN polymérase va pouvoir synthétiser en continu à partir d'une chaîne parentale un nouveau brin d'ADN, le brin précoce. Par contre sur l'autre chaîne parentale, la synthèse s'effectue de manière discontinue dans le sens contraire de l'ouverture la molécule d'ADN. La primase va régulièrement synthétiser des amorces d'ARN à partir desquelles l'ADN polymérase va synthétiser les fragments d'Okasaki. Ces fragments sont reliés les uns aux autres par l'ADN ligase, après que l'ADN polymérase I ait dégradé l'amorce ARN (grâce à son activité exonucléase $5' \rightarrow 3'$) puis re-synthétisé la partie manquante (figure 19).

3.3. La terminaison

La fin, ou la terminaison, de la réplication de l'ADN aura lieu lorsque les deux fourches de la réplication se rencontrent ou lorsqu'une fourche rencontre un signal de terminaison de la réplication; une protéine, appelée « Tus », se fixe sur le site de terminaison et arrête la progression de l'enzyme hélicase qui se détache de l'ADN. Lorsque la réplication d'un chromosome circulaire est terminée, les 2 molécules obtenues sont reliées ensemble, comme les maillons d'une chaîne (concaténées). La séparation et la ligation se font par une topoisomérase IV (topoisomérase de type 2).

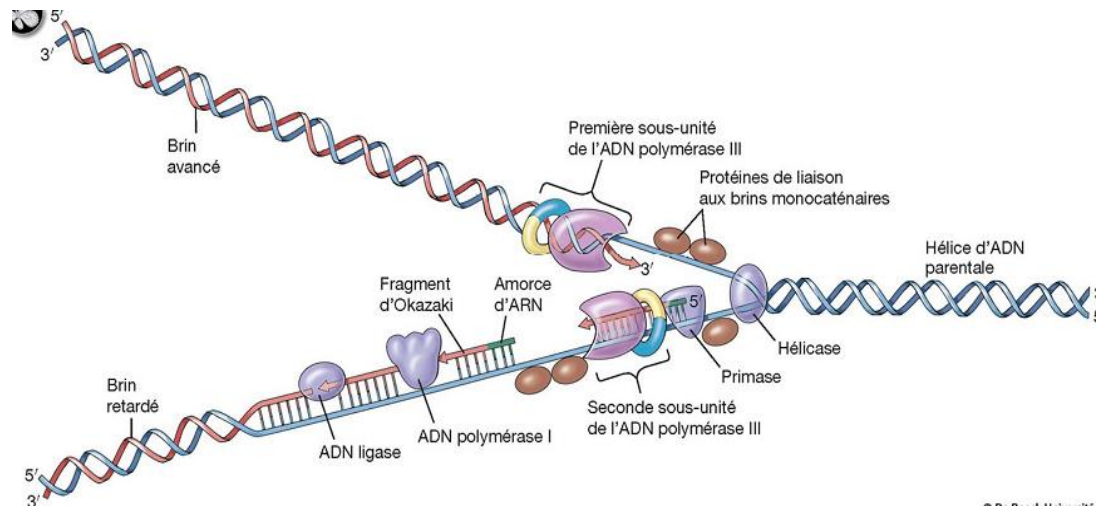


Figure 19: Processus de la réplication au niveau d'une fourche de réplication chez *E.coli*.

4. La réplication chez les eucaryotes.

Même principe que procaryotes avec le brin avancé et le brin retardé. Spécificités liées aux génomes eucaryotes:

- Génome beaucoup plus grand (12 Mbp dans *S. cerevisiae* à 3 Gbp chez l'homme), qui se réplique à une vitesse plus lente (50 nucléotides/sec) et qui est associé à des structures chromatiennes.
- L'ADN est intégré dans la chromatine associée aux histones, au moment de la réplication il y a production d'histones, il faut que les nucléosomes se dissocient pour permettre le déroulement de l'ADN et sa réplication. Les molécules d'ADN nouvellement formées s'organisent très rapidement en nucléosomes en s'associant aux histones
- Ceci compensé par de multiples origines de réplication activées de manière synchronisée et progressant à la même vitesse ; (mieux caractérisées chez *Saccharomyces cerevisiae*). Les chromosomes de *S. cerevisiae* contiennent des séquences conservées appelées ARS (*autonomous replicating sequence*) qui sont responsables de l'initiation de la réplication.
- Autres protéines intervenant au niveau de la fourche de réplication :
 - L'hélicase est le complexe MCG qui permet la séparation des deux brins de l'hélice.
 - RFC: Facteur de réplication C, c'est le clamp loader. Le chargeur de pince permet l'ouverture et la mise en place de PCNA (*prolifération cell nuclear antigen*): (Pol30 chez *S. cerevisiae*) qui encercle l'ADN.
 - DéroutasesRPA : réplication de protéine A, se fixent sur l'ADN simple brin et maintiennent l'ADN sous forme monocaténaire et protègent l'ADN simple brin de la dégradation.
- Plusieurs polymérases 5'→3', presque 15 ADN pol, agissant simultanément tout en ayant des fonctions différentes: Pol α : synthèse de l'amorce, élongation et réparation de l'ADN,

Pol β : réparation de l'ADN (équivalent de la I chez les bactéries), Pol γ : réplification de l'ADN mitochondriale, Pol δ : Elongation des brins plus réparation de l'ADN, Pol ϵ : Réparation et remplacement de l'ARN au niveau du brin retardé (Similaire à Pol).

- Seules les ADN Pol δ Pol ϵ Pol γ ont une activité autocorrection (3'→5' exonucléases). Aucune des ADN polymérase eucaryote n'a la capacité d'enlever les amorces d'ARN (pas d'activité 5'→3' exonucléase), cette étape est assurée par d'entre enzyme (FEN1, RNase H). La réparation des césures entre les fragments d'Okazaki par l'ADN pol β .

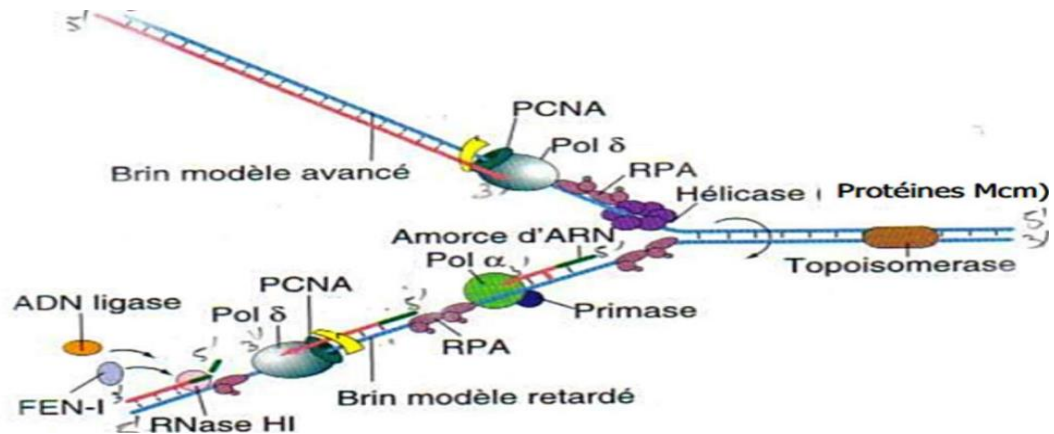


Figure 20: Modèle de la réplification de l'ADN chez l'eucaryote *Saccharomyces cerevisiae*.

- La télomérase empêche le brin retardé de se raccourcir progressivement lors de la réplification. Le mécanisme réplcatif ne permet pas de répliquer la totalité du télomère car à chaque division cellulaire on a une perte en 3', donc perte au niveau du télomère, donc le télomère va finir par disparaître. On a alors un mécanisme qui fait intervenir la télomérase : intervient avant la répllication : le rôle est d'allonger les extrémités 3' des télomères, pas au hasard, il faut ajouter AGGGTT plusieurs fois à l'extrémité 3', en recopiant une séquence qu'elle a dans son propre site actif, une séquence d'ARN : la télomérase est donc une transcriptase inverse (elle va utiliser de l'ARN pour former de l'ADN) (figure 21).

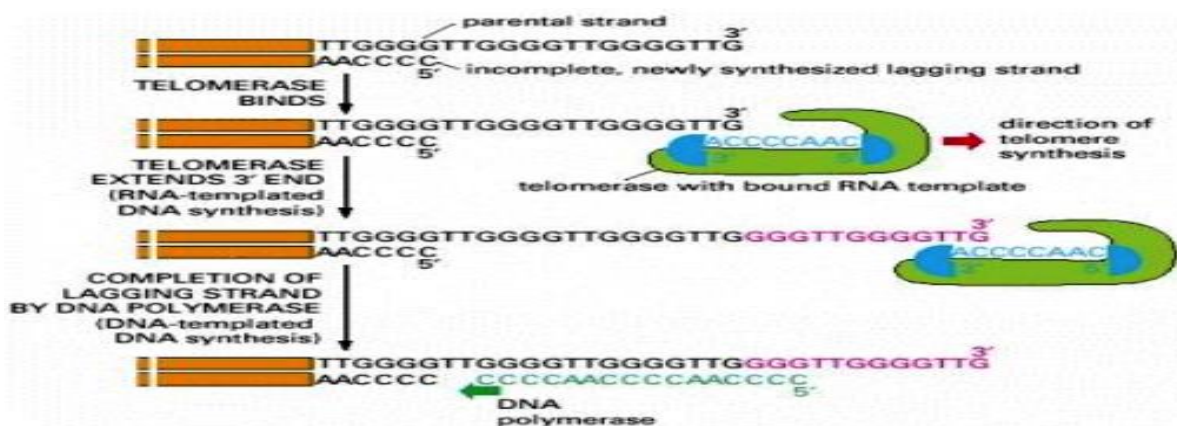


Figure 21: Mécanisme d'action de la télomérase

5. Régulation de la réplication

- **La méthylation des séquences GATC au niveau de l'origine de réplication :** l'initiation de la réplication nécessite la méthylation des séquences GATC sur les deux brins par la protéine Dam (pour DNA adénine méthylase). L'hémi-méthylation bloque la réinitiation.
- **Certains gènes sont régulés par un mécanisme épigénétique de méthylation :** L'hypométhylation entraîne le déverrouillage du gène, alors que l'hyperméthylation entraîne le verrouillage du gène. Exemple : Le gène de l'insuline est présent dans toutes les cellules, l'information n'est exprimée que dans les cellules pancréatiques, où il y a donc hypométhylation. Dans toutes les autres cellules, il y a hyperméthylation.
- **Rôle de la Dna A :** L'accumulation de Dna A induit l'initiation de la réplication, dont la synthèse est contrôlée et qui est, de plus, en état d'équilibre entre deux formes, l'une active, l'autre inactive.
- Chez les levures de type sauvage, l'hydroxyurée provoque l'arrêt de la réplication. Certains mutants, au contraire, n'arrêtent pas leur cycle en présence de ce produit et peuvent entrer en mitose même si leur génome n'est pas complètement dupliqué.
- la vitesse de la réplication est aussi contrôlée, au moins chez *S. cerevisiae*, en réponse à des lésions sur l'ADN. Un des gènes impliqués dans ce contrôle semble être l'homologue fonctionnel d'un gène humain qui est muté chez les malades atteints d'ataxie-télangiectasie.
- Un complexe multiprotéique, appelé ORC (*origin recognition complex*) est présent tout au long du cycle cellulaire sur les ARS de *S. cerevisiae* et change de structure après la mitose. Seule la forme postmitotique apparaît capable de permettre le démarrage de la réplication à partir d'une ARS, après interaction avec une protéine signal activée spécifiquement en fin de phase G1. Après le démarrage, le complexe ARS/ORC change de structure et apparaît inactivé sur les deux copies chromosomiques jusqu'à la mitose, empêchant ainsi une nouvelle réplication à partir de la même origine avant la division cellulaire.

II. La réparation de l'ADN et détection du pouvoir mutagène

Les lésions générées sont potentiellement dangereuses car elles peuvent modifier le fonctionnement des cellules qui constituent notre organisme et conduire à leur transformation en cellules cancéreuses. Heureusement, il existe des mécanismes cellulaires chargés de repérer ces lésions et de les éliminer.

1. Nature et origine des dommages à l'ADN

Le métabolisme cellulaire normal produit entre autre des espèces réactives de l'oxygène, sources de nombreux dommages. D'autre part, l'environnement soumet en permanence un organisme à différentes agressions, tel le rayonnement (UV), les radiations ionisantes ou des agents génotoxiques. Parfois, des molécules toxiques pour l'ADN sont utilisées comme outil thérapeutique, notamment en chimiothérapie. Les lésions ainsi générées sont de natures très diverses : bases altérées ou perdues, liens intra- ou inter-brins, dimères de thymines, cassures simple ou double brin. Il est intéressant de noter que le métabolisme de l'ADN lui-même est source de dommages. Ainsi, les polymérase, qui réalisent la copie à l'identique de l'ADN lors de la réplication, commettent parfois des erreurs qui rompent la complémentarité entre les deux brins.

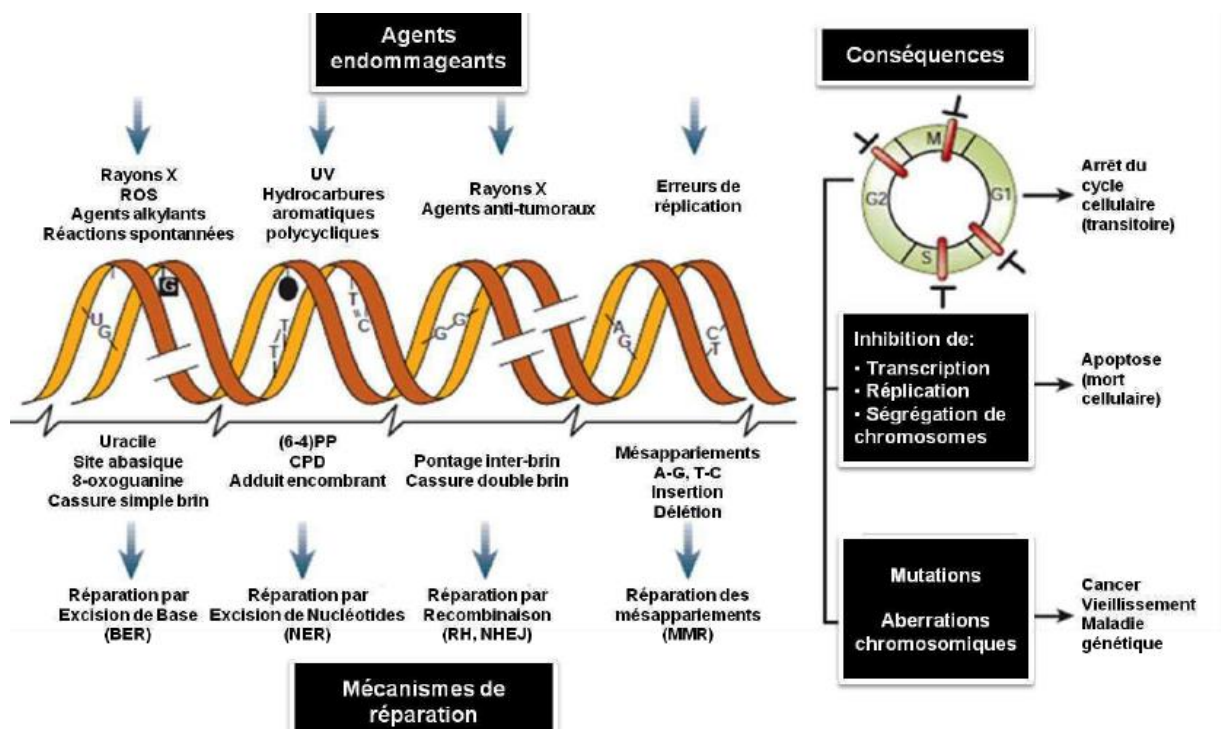


Figure 22: Dommages de l'ADN, mécanisme de réparation et conséquences.

2. Les mécanismes de réparation faisant suite à la réplication

2.1. La réparation des erreurs de réplication

De nombreuses ADN-polymérase possèdent une activité exo-nucléase capable de limiter les mésappariements, intervient directement lors de la polymérisation. Les enzymes ont donc deux ou trois activités : polymérase 5' vers 3', exonucléase 3' vers 5' (proof reading), exonucléases 5' vers 3'. Grâce à cette activité, l'ADN pol I permet une réplication vingt fois plus conservative. L'ADN pol III réduit par mille le même type d'erreurs.

2.2. Correction des mésappariements après la réplication :

C'est un système de réparation permettant la correction d'un mésappariement oublié par la fonction d'édition. Grâce à ce système, on aboutit à un taux d'erreur de 1 pour 10⁹ bases. Ce système existe aussi bien chez les cellules eucaryotes et que chez les cellules procaryotes.

- Chez les procaryotes, le système de réparation repère la méthylation des adénines des séquences GATC et fait intervenir les enzymes MUT (figure 26).
- Chez les eucaryotes, le système repère la méthylation des cytosines des séquences CG et fait intervenir les enzymes hMSH, hMLH, hPMS.

Le système MMR (methyl Mismatch Repair) d'*E.coli* mobilise un ensemble de gènes dont les produits ont une action spécifique :

- MutS patrouille sur l'ADN à la recherche de mésappariement. Il détecte le brin porteur de l'erreur (nouvellement synthétisé) en distinguant le degré de méthylation des deux brins. En effet, la Dam méthylase, lors de la réplication, reconnaît et méthyle le brin parental.
- Fixation de Mut L qui stabilise. L'hélicase, MutU déroule la double hélice en hydrolysant les liaisons Hydrogènes
- L'endonucléase MutH incise et coupe le brin muté à une certaine distance de part et d'autre de la base erronée.
- L'ADN pol III resynthétise le fragment manquant.
- La ligase permet de réassembler les deux morceaux.

Ce système doit agir dans le laps de temps où la méthylation du brin fils n'est pas encore effectuée.

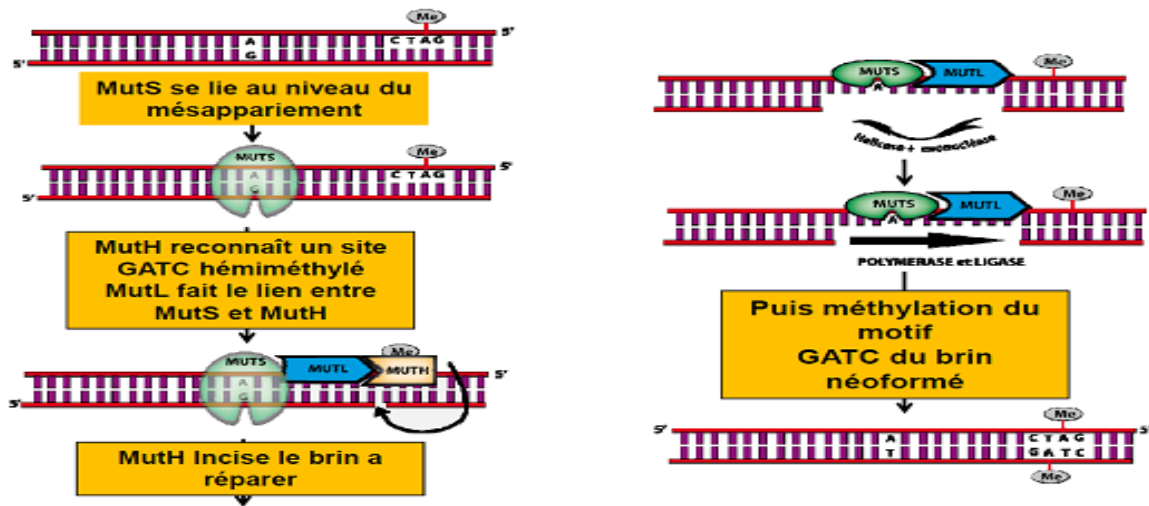


Figure 23: Le système de réparation MMR (methyl Mismatch Repair)

2.3. Mécanisme d'excision resynthèse (BER)

Le mécanisme BER (Base Excission Repair) est un des systèmes de réparation des ADN procaryotes et eucaryotes, impliqué dans les réparations de mutations endogènes (bases modifiées, méthylation, désamination, hydroxylation) jusqu'à 4 nucléotides par excision de bases. Les ADN glycosylases reconnaissent et enlèvent la base modifiée (hydrolyse de la liaison N-glycosidique). Cela produit un site AP (APurique ou APyrimidique) sur l'ADN. Une endonucléase 3'-5' coupe la liaison phosphodiester adjacente au site AP, l'ADN polymérase I enlève le site AP et synthèse le morceau d'ADN manquant, puis ligase met en place la liaison phosphodiester manquante.

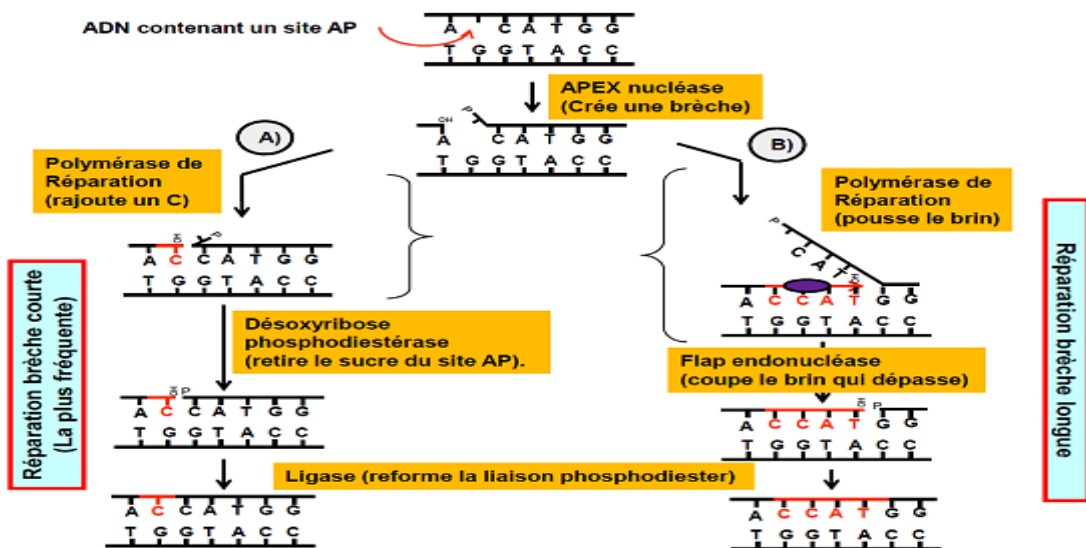


Figure 24: Représentation des voies constituant la réparation par excision de base

2.4. Réparation par excision/réparation des plusieurs nucléotides (NER)

Ce mécanisme est particulièrement efficace pour des lésions plus volumineuses (dimères de pyrimidines, adduits avec le cis-platine ou les psoralènes...), induisant des changements conformationnels substantiels de l'hélice d'ADN, où il faut exciser plusieurs nucléotides.

❖ **Chez *E. coli*** : La réparation est assurée par un complexe multienzymatique : l'exinucléase ABC. Les constituants de ce complexe sont des protéines codées par *uvrA*, *uvrB* et *uvrC*. Ce complexe se positionne sur la lésion et reconnaît la modification. Après s'être fixé, il induit une distorsion de la double hélice. L'ADN hélicase permet de séparer les deux brins et de les dissocier, ce qui permet à l'exinucléase de cliver un fragment de 13 nt, contenant les lésions. L'ADN POL I se fixe sur l'extrémité 3'OH libre pour synthétiser l'ADN complémentaire et anti parallèle. La dernière liaison phosphodiester est réalisée par l'ADN ligase.

❖ **Chez l'Homme** : Ce sont les protéines ERCC1 et les protéines XP (A, B, C, ...) (xeroderma pigmentosum) qui réalisent la NER dans les cellules. La protéine XPC reconnaît la lésion. Certaines XP possèdent l'activité hélicase (XPB et XPD). L'excision est réalisée par XPF, ERCC1 et XPG. La synthèse d'ADN se fait par l'ADN POL δ et l'ADN POL ϵ et la dernière liaison phosphodiester est réalisée par l'ADN ligase.

3. Réparation des lésions (naturelles ou induites)

3.1. Réparation par réversion direct

Les mécanismes de réparation ne nécessitent pas l'action de nucléases qui coupent l'ADN : ils réparent ou agissent directement sur la mutation.

3.1.1. La déalkylation C'est le transfert de groupement méthyle ou éthyle sur des bases (ou des phosphates) de l'ADN, la guanine pouvant être mutée en O₆-méthyl-guanine. Celle-ci peut alors s'apparier par erreur avec une thymine, induisant une mutation G-C vers A-T. Cette réparation s'effectue de façon direct (sans incision de l'ADN) par une enzyme, la O₆-méthyle-guanine-méthyle-transférase (ou MGMT). Elle transfère le groupement méthyle de la base vers un de ses acides aminés, une cystéine. L'enzyme de réparation est alors alkylé de façon irréversible. En cas d'alkylation très élevée, le système peut alors être saturé, et la réparation ne pourra pas se faire correctement.

3.1.2. La photoréactivation

La réparation des dimères pyrimidiques induit par une exposition aux UV dépend d'un système utilisant la lumière visible pour dissocier ses dimères. Ce clivage est effectué par une enzyme de photoréaction, la photolyase activé par la lumière du soleil.

L'absorption d'un photon par le folate permet l'interaction de ce dernier avec le cofacteur redox de la protéine, $FADH_2$, qui aboutit à la réduction du dimère et à la rupture des liaisons du cycle cyclobutane entre les deux bases adjacentes. Ce système est largement présent chez les eucaryotes et les procaryotes, mais pas chez l'Homme.

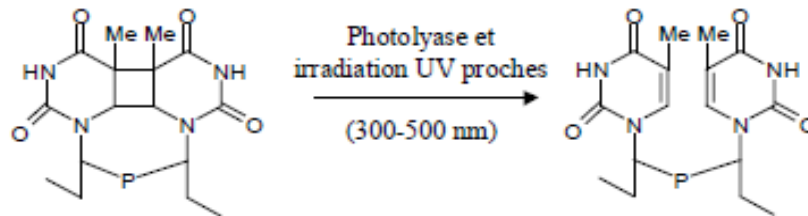


Figure 25 : Réparation par réversion directe du dimère cyclobutane de thymine

3.2. La réparation par recombinaison

C'est un système de réparation constitutif, basé sur la recombinaison homologue, qui intervient quand les systèmes précédents n'ont pas pu réparer les lésions majeures de l'ADN, soit car ils étaient défectueux, soit parce qu'ils étaient débordés.

❖ **Chez les procaryotes** (figure 30): S'il n'y a pas de réparation avant réplication, une lacune (ou brèche post-répllicative) est laissée par l'ADN polymérase (1), lorsqu'elle rencontre une lésion. Cette lacune va être réparée par recombinaison homologue, elle est remplie par la séquence du brin parental opposé identique à ce brin fils (2), qui est prélevée par la protéine REC A. Cela va fournir la séquence correcte et créer une deuxième lacune sur le brin parental opposé (3). Puis le dimère est excisé par des enzymes d'excision (4). Les deux lacunes (à la place du dimère de thymines et sur le brin parental opposé) sont ensuite corrigées grâce à l'ADN polymérase qui synthétise la séquence complémentaire et antiparallèle de chaque brin parental (5) (figure 27).

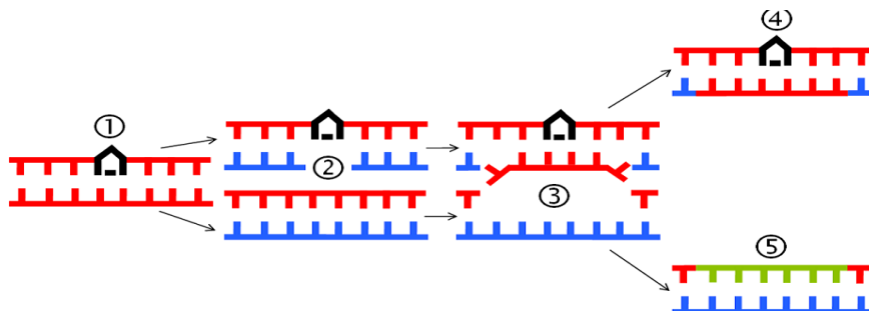


Figure 26: Réparation par recombinaison homologue

❖ Chez l'Homme :

Le système est équivalent mais il fait intervenir chez l'Homme la protéine RAD 51 (homologue eucaryote de REC A), associée aux protéines BRCA1 et BRCA2.

3.3. Une réparation mutagène, le système SOS

Lorsque les lésions sont graves et nombreuses, la synthèse de l'ADN est stoppée, ce qui provoque l'apparition de zones monocaténares. Ces zones sont reconnues par l'enzyme de réparation et recombinaison RecA qui agit d'une part en assurant les événements de recombinaison et d'autre part en hydrolysant par une activité protéolytique de la protéine répresseur LexA, qui s'attache au complexe RecA-ADN simple brin, et est clivé. Le clivage entraîne son inactivation et crée un signal SOS, permettant l'expression d'une vingtaine de gènes impliquée dans la réparation de l'ADN. On observe entre autre :

- L'activation du gène UvrA, impliqué dans les mécanismes de réparation par excision.
- La synthèse en grande quantité de RecA (induit sa propre expression).
- La synthèse des gènes UmuD et UmuC qui jouent un rôle important dans la mutagènesesOS.

III. Systèmes de restriction-modification**1. Origines de méthylation d'ADN dans les bactéries**

La méthylation de l'ADN fournit un contrôle réversible épigénétique du programme génétique. Cette modification dans les bactéries peut moduler des procédés variés, tels que la régulation de l'expression des gènes et la réplication de l'ADN. La méthylation d'ADN est devenue comme méthode pour protéger l'ADN bactérien contre l'ADN extérieur du bactériophage. Ce mécanisme de défense se nomme le système de restriction-modification (R-M), et il se compose de deux composantes : Méthyltransférase d'ADN et endonucléase de restriction qui reconnaît et clive les séquences d'ADN étrangères sur des sites spécifiques au niveau des liaisons phosphodiester, générant des surplombs ou des extrémités franches de 5' ou 3', tandis que la méthyltransférase assure la discrimination entre l'ADN de soi et de non-soi, en transférant des groupes méthyle à la même séquence d'ADN spécifique dans le génome de l'hôte. Yuan et Meselson (1968) ont élucidé la base biochimique du phénomène RM de l'hôte, par la purification de l'endonucléase de restriction d'*E.coli* K12. Cela a soulevé la perspective d'une manipulation contrôlée de l'ADN.

2. Types de systèmes de restriction-modification dans les bactéries

Les systèmes de restriction permettent aux bactéries de surveiller l'origine de l'ADN entrant et de le détruire. Quatre systèmes de restriction-modification ont été décrits, selon leurs modes et sites d'action.

- **Les systèmes de type I** sont les plus complexes, consistant en trois polypeptides: R (restriction), M (modification) et S (spécificité). Le complexe résultant peut à la fois cliver et méthyler l'ADN. Les deux réactions nécessitent de l'ATP et le clivage se produit souvent à une distance considérable du site de reconnaissance. Exemples : EcoKI et EcoR124I.
- **Les systèmes de type II** a des enzymes d'endonucléase et de méthyltransférase de restriction codées comme deux protéines séparées et agissent indépendamment sur le même site de reconnaissance et sont donc en compétition pour l'activité (ne nécessite pas d'ATP, et coupe au niveau de courtes séquences palindromiques de 4 à 8 pb). Les enzymes de type II sont les plus étudiées et les plus répandus.
- **Les systèmes de type III** se compose d'une méthyltransférase qui contient le domaine et les composés obligatoires de spécificité d'ADN avec de l'endonucléase de restriction pour le fendage. Ces enzymes reconnaissent de courtes séquences asymétriques de 5 à 6 pb, se transloquent le long de l'ADN et clivent le côté 3' du site cible à une distance de ~25 pb (nécessite de l'ATP).
- **Les systèmes de type IV** ne sont pas de vrais systèmes RM car ils ne contiennent qu'une enzyme de restriction et non une méthylase. Contrairement aux autres types, les enzymes de restriction de type IV reconnaissent et ne coupent que l'ADN modifié.

3. Nomenclature des enzymes de restriction

La nomenclature des enzymes de restriction est précise. Leur nom comporte 3 ou 4 lettres correspondant entre autres à la bactérie à partir de laquelle on a extrait cette enzyme :

- La première lettre est en majuscule cela correspond à la première lettre du genre de la bactérie d'où a été extraite l'enzyme.
- La seconde et troisième lettre sont en minuscule et correspondent aux deux premières lettres de l'espèce bactérienne
- La quatrième lettre correspond à la souche bactérienne, elle est écrite en majuscule mais n'est pas présente dans toutes les enzymes de restriction
- Enfin un chiffre romain indique le numéro d'ordre de caractérisation de l'enzyme

Pour faire la distinction entre les activités de restriction et de méthylation, on utilise respectivement les préfixes R et M, par exemple ; R.*SmaI* et M.*SmaI*.

Exemple:

Enzymes **BamHI** : Bacillus (genre), amyloliquefaciens (espèce), souche H, 1^{ère} enzyme isolée

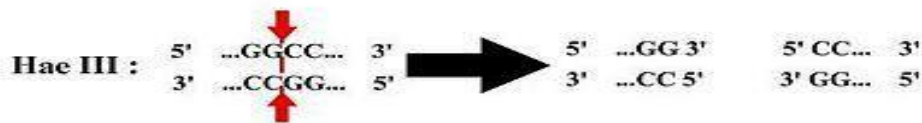
Les enzymes **EcoRI**: Escherichia (genre), coli (espèce), souche RY13, 1^{ère} enzyme isolée.

4. Séquences cibles et types de clivage.

Les enzymes de restriction reconnaissent et coupent des séquences de quatre à huit nucléotides très spécifiques appelées sites de restriction. Ces sites sont pour la plupart palindromiques, c'est-à-dire qu'ils sont composés de séquences nucléotidiques identiques sur les deux brins mais en orientations inversées ou antiparallèles.

Les extrémités générées peuvent être cohésives (des bouts collants) ou franches ; la coupure de la liaison phosphodiester laisse une extrémité 3'-OH et une extrémité 5'-phosphate (figure 28). Les parties simples brins peuvent s'apparier, d'où la possibilité de lier des fragments d'ADN d'origine différentes : c'est le phénomène de la recombinaison génétique.

Extrémités non cohésives :



Extrémités cohésives :

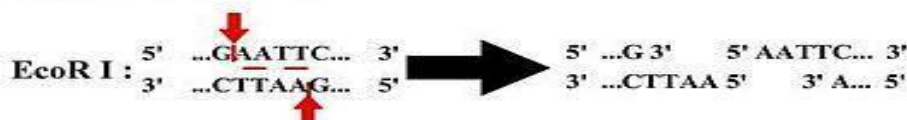


Figure 27 : Séquences palindromique cible des enzymes de restriction

- Les enzymes de restriction qui possèdent le même site de reconnaissance et clivent ce site de façon identique sont des **Isoschizomères**. Par contre, celles qui possèdent le même site de reconnaissance mais clivent différemment l'ADN sont des **Neoschizomères**. Exp: *SmaI* (CCC/GGG) et *XmaI* (C/CCGGG).
- Deux enzymes de restriction sont dites compatibles quand elles génèrent après digestion des fractions aux extrémités cohésives complémentaires. Ces fragments peuvent être facilement ligatures. Exemple: Les enzymes Bam HI (G / GATCC) et Mbo I (/ GATC):

5. Applications des enzymes de restriction

Les enzymes de restriction sont des ciseaux moléculaires et qui sont largement utilisés dans de nombreux domaines :

5.1. Technologie de l'ADN recombinant

Les enzymes de restriction sont utilisés en biotechnologie pour l'insertion de gènes ou de fragment d'ADN dans des molécules appartenant à d'autres organismes. Ces enzymes génèrent des extrémités qui peuvent s'apparier entre elles. Ainsi un gène de résistance aux champignons est cloné dans un plasmide.

5.2. Cartographie de restriction

L'identification des sites de coupure des enzymes de restriction qui reconnaissent des points précis du génome, permet d'établir une carte de restriction extrêmement détaillée. Ce qui fournit des informations utiles pour la caractérisation des molécules d'ADN.

Cette cartographie a trouvée de nombreuses applications, comme la détection d'une maladie génétique telles que la mucoviscidose, la localisation des gènes pour les troubles génétiques, la détermination du risque de maladie pour recenser un transporteur d'une mutation de pathogène dans une famille et les tests de paternité . En effet, la RFLP (Polymorphisme de longueur des fragments de restriction) couplé à une PCR permet l'établissement de l'empreinte génétique.

6. Polymorphisme de restriction.

6.1. Polymorphismes de longueur des fragments de restriction (RFLP) :

L'analyse de la longueur des fragments de restriction ou RFLP, à la recherche de variations individuelles est une des techniques d'analyse de la séquence primaire de l'ADN à la recherche de substitutions, d'insertions ou de délétions qui affectent la spécificité de la spécificité de clivage REase et modifient le nombre de sites de restriction, donc la longueur des fragments de restriction. Les RFLP sont utilisés comme marqueurs sur les cartes génétiques, En règle générale, l'électrophorèse sur gel est utilisée pour visualiser les RFLP

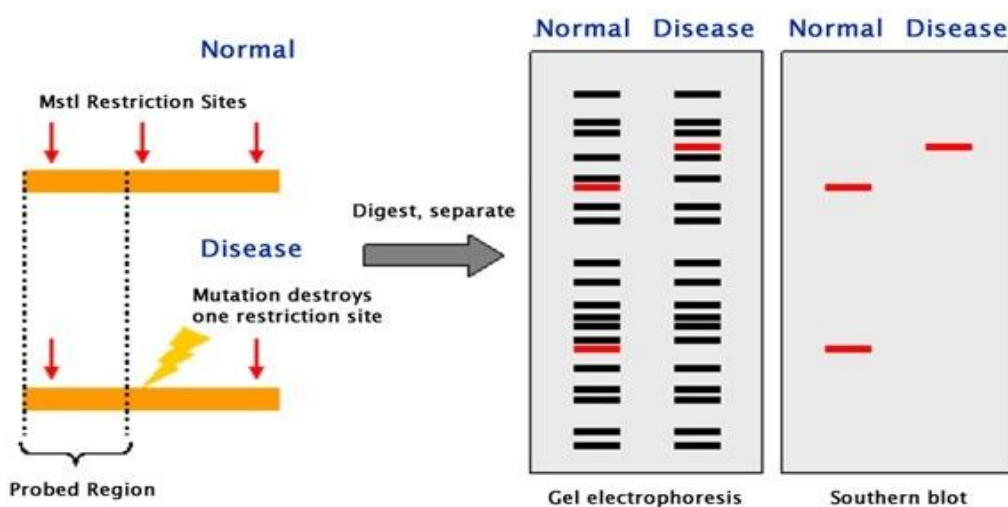


Figure 28 : Polymorphismes de longueur des fragments de restriction (RFLP)

6.2. Analyse RFLP

Après isolement et purification de l'ADN, celui-ci est incubé avec différentes enzymes de restriction ; c'est l'extrême spécificité de ces enzymes qui est exploitée pour la mise en évidence du polymorphisme. Les fragments de digestion sont séparés selon leur taille par une électrophorèse en gel d'agarose et transférés sous forme dénaturée sur une membrane de nylon. L'hybridation de la membrane à une sonde ADN marquée détermine alors la longueur des fragments complémentaires de la sonde. L'ADN cible digéré est visualisé par autoradiographie, si la sonde est marquée au P32 (sonde chaude) ou par méthodes biochimiques si la sonde est non radioactive (sonde froide).

Le polymorphisme détecté est dû à des mutations au niveau des sites de restriction de l'enzyme (polymorphisme de site de restriction) et/ou à des délétions/insertions d'un fragment d'ADN au voisinage de la zone génomique reconnue par la sonde. C'est le couple enzyme/sonde qui constitue le marqueur. Bien que cette technique soit co-dominante et permette une analyse génétique complète, elle est lente et laborieuse. Les étapes de transfert et d'hybridation empêchent une automatisation du travail.

6.3. Les applications de la technique de RFLP

Le PLFR a été employé pour plusieurs applications d'analyse génétique depuis son invention. Enumérer les applications de la technique de RFLP revient à énumérer les utilisations des marqueurs génétiques, que l'on peut diviser en deux grandes catégories :

- les études de diversité, avec les analyses de structuration intra ou interpopulation, le criblage de variantes dans les populations, l'identification variétale, etc.
- la construction de cartes de liaisons génétiques, avec leurs multiples intérêts : localisation de gènes majeurs, analyse de la recombinaison, cartographie de gènes à effets quantitatifs, etc. la construction de cartes est ainsi très répandue chez les végétaux.

I. La transcription et la maturation de l'ARN

1. Généralités

La transcription constitue l'ensemble des mécanismes par lequel l'ARN messager (ARNm) est synthétisé. L'ARNm est une copie d'une portion de l'ADN. Seules certaines portions de l'ADN sont transcrites, ces séquences d'ADN sont appelées gènes. Enfin, seul l'un des deux brins d'ADN est copié en ARNm. La transcription constitue l'étape préliminaire essentielle pour la biosynthèse protéique (ou traduction). La transcription ne produit pas seulement des ARNm, mais aussi ARNt, ARNr et ARNsn.

Chez les procaryotes une seule ARN polymérase-ADN dépendante effectue la transcription dans le sens 5' → 3', pour tous les types d'ARN, tandis que chez les eucaryotes, trois ARN polymérases (ARNpol) interviennent : l'ARNpol I pour les ARN ribosomiques transcrits dans le nucléole (28S, 18S et 5,8S), l'ARNpol II pour les ARNm, et l'ARNpol III pour les petits ARN (ARNt, ARNr 5S, ARNsn). Pour certains virus à ARN enfin, l'ARN est transcrit par une ARNpol-ARN dépendante appelée aussi réplacase

2. Caractéristiques générales de la transcription

2.1. Les éléments nécessaires à la transcription

La synthèse d'un mARN à partir d'ADN s'effectue dans le sens 5' → 3', de manière antiparallèle et complémentaire par rapport à la portion d'ADN copiée. La synthèse d'un ARNm nécessite:

- La présence de nucléotides propres à l'ARNm, contenant du ribose, des bases A, U, G et C et sous forme de nucléosides triphosphates: ATP, UTP, CTP et GTP.
- La présence d'une enzyme: l'ARN-polymérase. Les sels de Mg sont indispensables.
- La présence d'une matrice d'ADN servant de modèle.

2.2. Les étapes de la transcription chez les procaryotes.

La transcription ne concerne qu'une portion de l'ADN. Il est bien évident qu'il faut définir le début et la fin de la transcription et préciser les mécanismes par lesquels l'ARN-polymérase reconnaît la portion d'ADN à transcrire. La synthèse de l'ARN messager comprend trois phases successives:

2.2.1. L'initiation

L'initiation de la transcription se fait au niveau d'une séquence appelée promoteur. Ce promoteur est constitué de deux séquences très conservées de six bases, situées respectivement 35 et 10 pb en amont du site d'initiation de la transcription (+1). Les ARN polymérases ne nécessitent pas d'amorce et ne possèdent pas d'activité exonucleasique et donc de correction d'erreur.

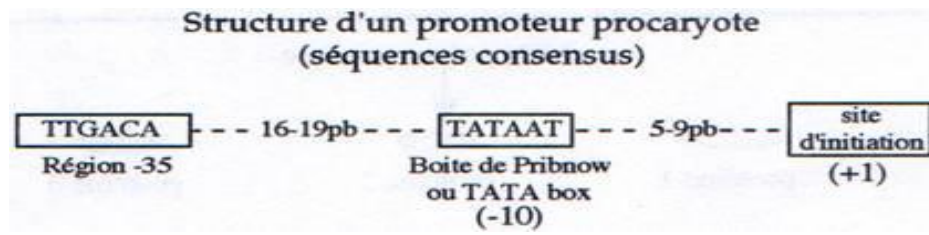


Figure 29 : Représentation du site promoteur des Procaryotes.

Le core-enzyme (forme inactive de l'enzyme contient les sous-unités α β ω) a une affinité faible pour les séquences d'ADN double brin. Par contre la présence du facteur δ (élément *trans*) dans l'holoenzyme confère à l'ARN polymérase une très forte affinité pour le promoteur. L'holoenzyme explore donc l'ADN par liaison non spécifique, il recouvre une région située entre les positions -60 et +20 et se lie fortement aux régions promotrices qu'elle dénature localement, se fait de la position -10 à la position +1, pour commencer la transcription. L'initiation va durer le temps que la polymérase associe 7 ou 8 ribonucléotides sous forme d'un polymère hybridé au brin matrice. Comme le core-enzyme présente une très forte affinité pour les hétéroduplexes ARN/ADN, le facteur δ se décroche et c'est le core-enzyme qui va seul continuer la transcription.

2.2.2. L'élongation de la chaîne polynucléotidique.

La progression de la transcription se fait dans le sens 5'→3', nécessite de former des liaisons ester entre les nucléotides. Le core-enzyme continue la polymérisation tout en déroulant l'ADN en aval et en le réenroulant une fois la séquence copiée. L'ARN initialement apparié au brin d'ADN transcrit se détache. Les liaisons hydrogène se reforment après passage de l'ARN polymérase et les deux brins d'ADN reprennent leur forme hélicoïdale. Parmi les sous-unités de l'ARN polymérase, c'est la sous-unité β qui joue le rôle catalytique, alors que β' servirait essentiellement à maintenir la liaison entre l'ADN et l'ARN polymérase.

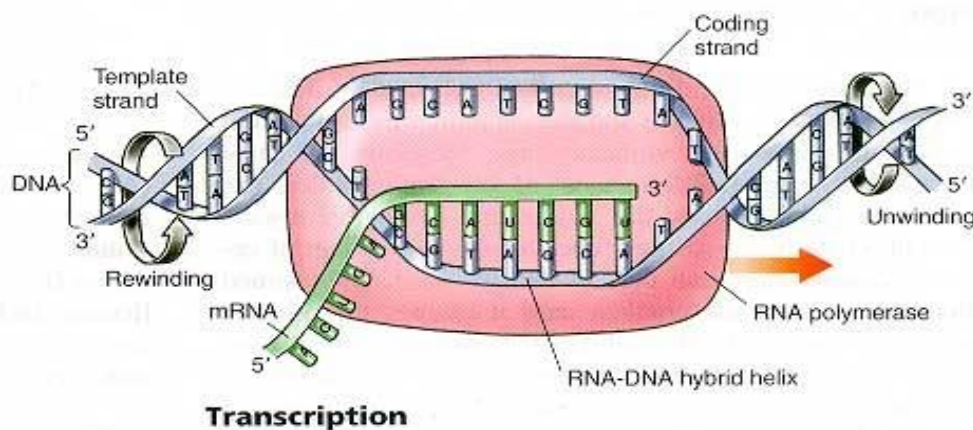


Figure 30 : La transcription chez les bactéries

2.2.3. Terminaison

Le terminateur se présente sous la forme d'un palindrome qui peut être parfait ou imparfait. Ce palindrome entraîne une complémentarité de séquence au niveau de l'ARNm qui permet la mise en place d'une structure en épingle à cheveux (appariement intra-chaîne) qui déstabilise l'ARN-polymérase jusqu'à dissociation. Elle peut être facilitée par un facteur rho ρ suivant la séquence du terminateur, on met ainsi en évidence des terminateurs rho indépendant (environ les 2/3) et des terminateurs rho dépendant (environ 1/3) :

- Pour les terminateurs rho indépendant on trouve une structure en épingle à cheveux riche en paires de bases G-C, suivie d'une séquence poly-U d'environ 6 nucléotides permettant une dissociation plus facile de l'hybride ADN-ARN.
- Pour les terminateurs rho dépendant on trouve une structure en épingle à cheveux plus courte et qui n'est pas riche en paires de bases G-C et qui est non-suivie d'une séquence poly-U. Il y a donc nécessité du facteur rho qui a une affinité pour les ARN en court de synthèse, le parcourant de 5' vers 3' jusqu'à trouver l'ARN-polymérase. Le facteur rho est ATP dépendante, dont l'hydrolyse permettra la dissociation du complexe.

2.3. Les ARNs polycistroniques

Certains gènes procaryotes sont rassemblés en opérons structures qui renferment plusieurs gènes, impliqués dans une même voie métabolique, sous le contrôle d'un même promoteur. On retrouve donc au sein d'un même ARN, des séquences codant pour plusieurs protéines.

3. Transcription de l'ADN eucaryote

Chez les eucaryotes le mécanisme de base de la transcription est identique à ce qui a été décrit pour les procaryotes. Cependant, la structure des promoteurs est différente et les transcrits primaires obtenus sont toujours monocistroniques. Enfin, une des différences majeures concerne les modifications post-transcriptionnelles des ARN eucaryotes.

3.1. Les ARN-polymérases eucaryotes

Trois ARN-polymérases eucaryote ont été mis en évidence. Elles diffèrent par leur localisation dans le noyau, par la nature des ARN formés et de par leur sensibilité à des inhibiteurs tels que l' α -amanitine.

- ARN-polymérase I dans le nucléole pour les ARNr 5,8 ; 18 et 28 S, et est insensible à l' α -amanitine qui inhibe l'élongation de la transcription
- ARN-polymérase II dans le nucléoplasme pour les ARNm, ARNsn et est sensible à l' α -amanitine
- ARN-polymérase III dans le nucléoplasme pour les ARNt, ARNr 5 S et pour les petits ARN, elle est également sensible à l' α -amanitine mais à hautes doses.

3.2. La formation d'un ARN prémessager

3.2.1. L'initiation de la transcription

Tout l'ADN n'est pas transcrit par la polymérase II, seules les régions correspondant à des gènes le sont. Et encore, cette expression peut être régulée selon le stade de développement, le type cellulaire, l'environnement, etc... Dès lors, un acteur doit intervenir pour déterminer à quel endroit une région d'ADN doit commencer à être transcrite : c'est le rôle du promoteur. Le promoteur correspond à une région non transcrite de l'ADN, généralement juste en amont du début de la région transcrite, dont la séquence permet le recrutement de l'ARNpol II. Certaines séquences du promoteur (surnommées "boîtes") sont reconnues spécifiquement par différentes protéines appartenant au complexe d'initiation (figure 31) :

- la "boîte TATA" riche en thymine et adénine, entre -25 à -30 nt du site de démarrage (+1), elle semble jouer un rôle dans la précision du premier nucléotide transcrit;
- des éléments proximaux :
 - la "boîte CAAT" (facultative), contenant de la cytosine, entre -120 à -80 nt.
 - la "boîte GC" (facultative), peut être présente entre la boîte CAAT et TATA.

Signalons que si ces séquences sont souvent bien conservées, une variabilité non négligeable peut également être observée. Ainsi, il existe des promoteurs sans "boîte TATA".

Ces boîtes constituent les sites de fixation des facteurs de transcription et permettent ainsi la modulation de l'activité du promoteur minimum. Elles représentent donc un signal pour le complexe d'initiation de la polymérase II mais leur assemblage en différentes combinaisons possibles permet déjà une diversification qui va intervenir dans la régulation de la transcription. D'autres modules, parfois très éloignés du promoteur "de base", jouent également un rôle dans la régulation.

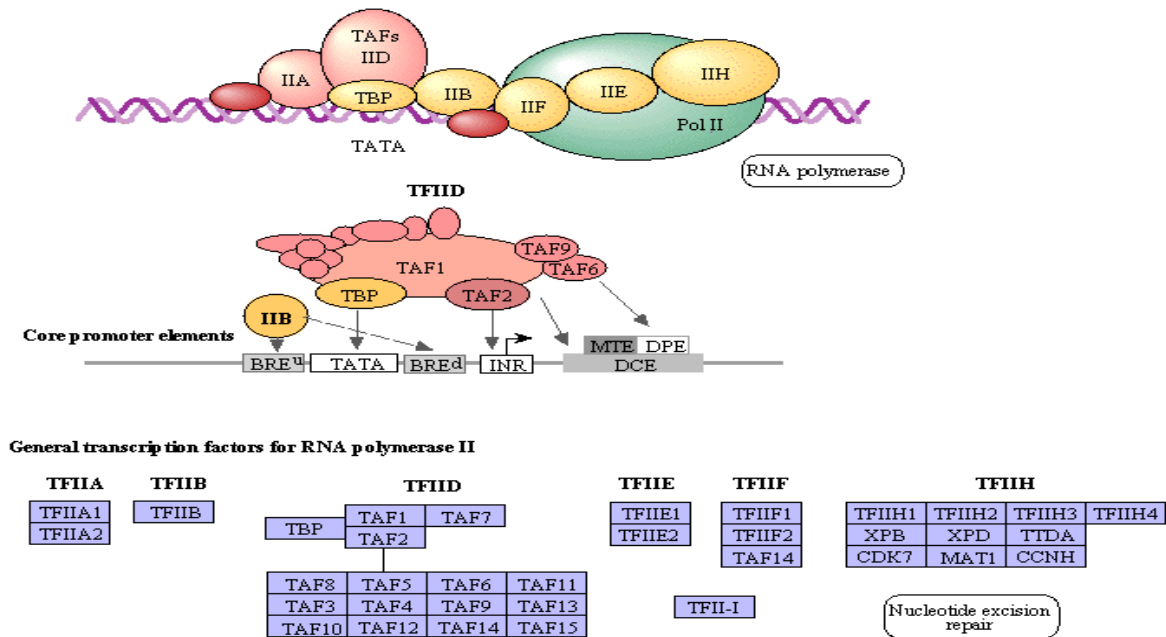


Figure 31: Mise en place du complexe protéique d'initiation

La protéine TBP (TATA box-Binding Protein) reconnaît la boîte TATA. TFII-D reconnaît le promoteur et permet la fixation de TFII-A puis des interactions protéiques entre cet assemblage, TFII-B et la polymérase B permettent la fixation de celle-ci, son maintien est assuré par TFII-E. TFII-H avec activité hélicase participe à la modification de la topologie de l'ADN, et une activité kinase responsable de la phosphorylation de la queue C-terminale de l'ARNpolymérase II qui entraînant la dissociation du complexe d'initiation et le début de la transcription.

3.2.2. L'élongation

L'ARN pol II est équipé de facteurs protéiques d'élongation qui facilitent sa progression au travers d'une chromatine dont ils relâchent la structure. Un ARN pré-messager complémentaire du brin matrice de l'ADN (brin antisens), donc identique au brin codant de l'ADN aux riboses et uraciles près, commence à être synthétisé selon la direction 5'→3'.

3.2.3. La terminaison

L'ARN pol II est également équipée de facteurs de terminaison, et reconnaît ainsi un ou plusieurs signaux de terminaison portés par le brin progressivement parcouru et qui annoncent la fin de la transcription sur le brin d'ADN matrice (TTATTT par exemple, parfois aussi plus

en aval ATACAAC...). Elle arrête bientôt son travail de transcription et libère l'ARN_{pm} qu'elle vient d'assembler.

3.3. Maturation des transcrits primaires

Le transcrit primaire n'est pas utilisé tel quel pour la synthèse protéique. Il doit subir des modifications qui répondent à plusieurs impératifs (augmentation de la demi-vie, modification de la séquence). Toutes ces modifications sont réalisées au fur et à mesure de la progression de la synthèse du préARN_m dans le nucléoplasme. Il existe trois grands types de modifications, catalysées chacune par des enzymes protéique ou ribonucléique (figure 32).

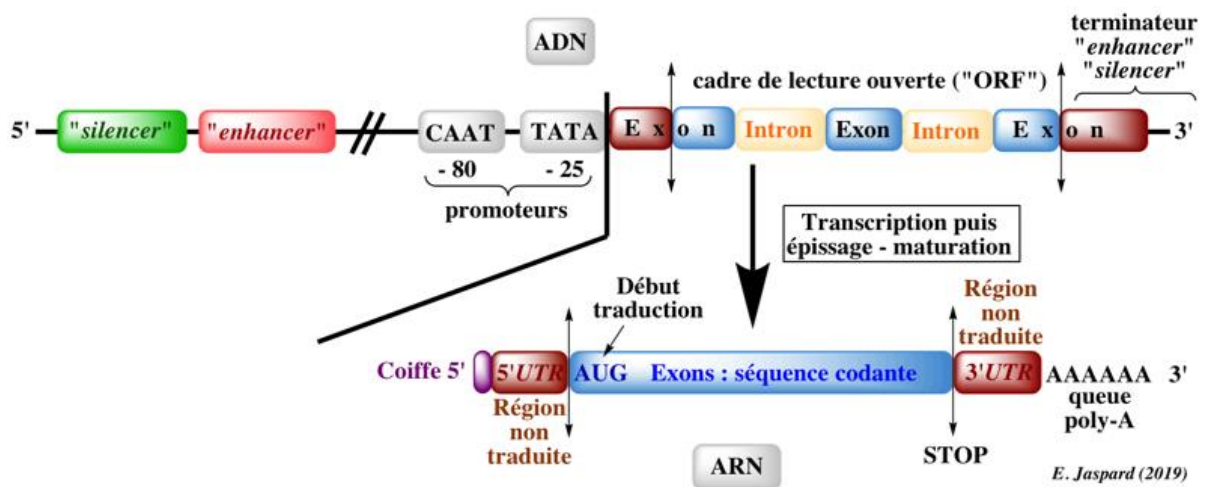


Figure 32: Schéma d'un gène eucaryote et d'un transcrit ARNm mature prêt à être traduit.

Coiffe ajoutée, la queue polyA ajoutée, introns éliminés et jonction des exons.

3.3.1. Addition de la coiffe en 5' (ou capping)

Elle a lieu dès le début de la transcription avant que la chaîne ne compte plus de 30 nucléotides. Elle consiste en l'ajout d'un nucléotide à guanine sur l'extrémité 5' de l'ARN suivi de sa méthylation sur l'azote 7 de la base, ainsi que de la méthylation en 2' du ribose des un ou deux premiers nucléotides du transcrit primaire (figure 33). Il en résulte que l'extrémité 5' de l'ARN_m n'est pas porteuse des trois acides phosphoriques libres habituels, mais d'un GMP, ce qui limite la réactivité de cette extrémité et sa reconnaissance par les exonucléase (protection contre la dégradation). Elle est également nécessaire à l'exportation de l'ARN_m vers le cytoplasme (reconnue par un complexe de fixation de la coiffe "*cap-binding complex*") et à la liaison de ce dernier avec la petite sous-unité du ribosome lors de l'étape d'initiation de la traduction.

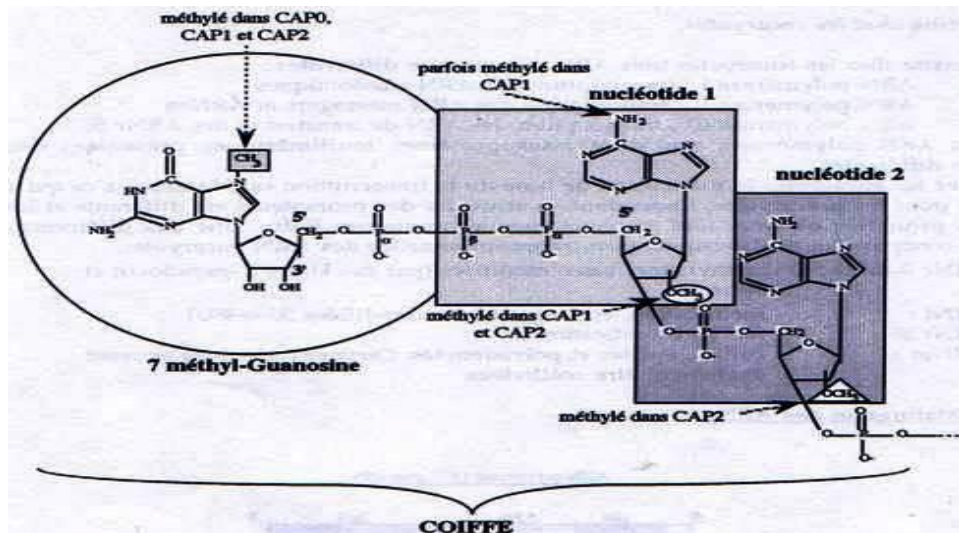


Figure 33: Addition de la coiffe: 7m-Guanosine.

3.3.2. Poly-adénilation en 3' par la poly-A polymérase

La poly-adénilation correspond à l'ajout de jusqu'à 200 adénines à l'extrémité 3' du transcrit primaire sans matrice par la poly-A-polymérase. La poly-A-polymérase reconnaît le signal de poly-adénilation qui n'est autre que la séquence CPSF (AAUAAA). 20 nucléotides en aval de cette séquence la poly-A-polymérase utilise son activité endo-nucléasique et son activité A-polymérasique. Elle semble assurer la protection de l'ARNm contre les dégradations enzymatiques et jouerait un rôle dans l'adressage du messenger vers le cytoplasme et stimule la terminaison de la transcription.

3.3.3. Excision des introns et épissage des exons (splicing)

Le transcrit primaire est encore soumis à l'excision des introns et l'épissage des exons ; les introns sont ainsi éliminés suite à la présence de site donneur d'épissage (dinucléotide GU) à l'extrémité 5' des introns et de site accepteur d'épissage (dinucléotide CAG) à l'extrémité 3' des introns.

Les jonctions d'épissage sont reconnues par les **snRNPs** (ou **snurps** pour *Small-Nuclear-Ribonucleo-protein-Particules*). Les snRNP correspondent à l'association de snRNA (snRNA U1, U2, U3, U4, U5, U6) et de protéines et l'ensemble des snRNPs s'appelle le **splicéosome**. Le splicéosome enlève les introns et effectue la jointure des exons pour former un ARNm mature. Les réactions catalysées consomment plusieurs molécules d'ATP, et elles sont réalisées par les ARNs plutôt que les protéines, et même la localisation des introns se fait par les

snRNA (Ex: U1, U2, U2AF). La taille de ces snRNA varie de 100 à 300 nucléotides chez la plupart des eucaryotes. Le snRNP U1 reconnaît le site donneur et le snRNP U2 reconnaît le site de branchement et le site accepteur. L'excision des introns et l'épissage des exons se fait en plusieurs étapes :

- Le snRNP U1 permet la reconnaissance du site donneur d'épissage et entraîne la rupture de la liaison phosphodiester entre le premier exon et l'intron.
- Cette rupture de la liaison phosphodiester entraîne la formation d'un lasso, qui n'est autre que l'extrémité 5' de l'intron. Ce lasso forme une liaison avec le site de branchement, lui-même situé sur le même intron qui se repli ainsi sur lui-même. Le site de branchement est reconnu par le snRNP U2 et permet la liaison par l'intermédiaire d'une adénosine.
- Le snRNP U2 permet également la reconnaissance du site accepteur d'épissage. Suite à cette reconnaissance il y a rupture de la liaison phosphodiester au niveau de l'extrémité 3' de l'intron.
- Le groupement 3'OH du premier exon peut ainsi réagir avec l'extrémité 5' phosphate du deuxième exon pour former une liaison phosphodiester et permettre la libération de l'intron qui sera dégradé.

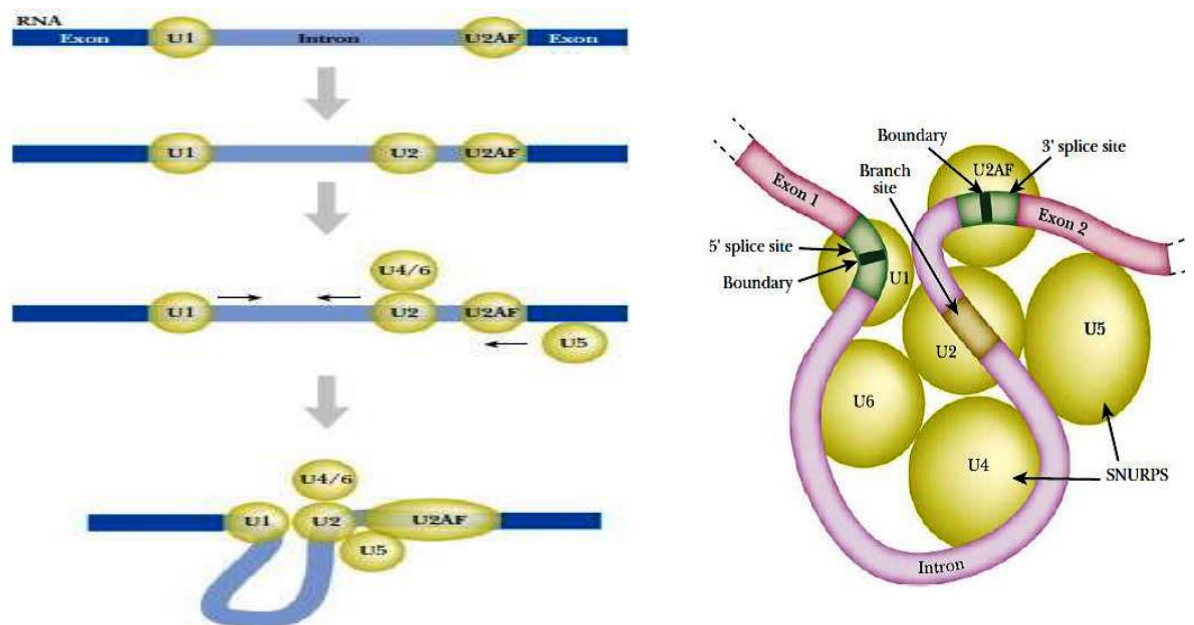


Figure 34: Formation de splicéosome, il ya six SNURPs (U1, U2, U4, U5 et U6, le U3 est un snRNA se trouve dans le nucléole), U1 s'apparie sur le site d'épissage en 5', U2 s'apparie avec le site de branchement, et U2AF avec le site d'épissage en 3'.

II. La traduction et la maturation des protéines

La traduction correspond au fait que l'ARNm est traduit en protéine : passage de séquences de nucléotides à des séquences d'acides aminés par respect du code génétique. La traduction s'effectue dans le cytoplasme de la cellule cytoplasme au niveau des ribosomes et nécessite la présence d'ARN de transfert (ARNt), de l'énergie sous forme de GTP et des dizaines de protéines et d'autres molécules différentes.

1. Le code génétique

Le code génétique est un code qui permet la conversion d'une séquence de nucléotides (ADN puis ARN) en séquence d'acides aminés (protéines). Le code implique les bases A, C, U et G ainsi que les 20 acides aminés. Le code génétique possède différentes caractéristiques :

- Les codons sont des triplets de nucléotides et ils codent pour un acide aminé.
- La séquence du gène et la séquence de la protéine codée sont colinéaires ; c'est-à-dire que la longueur du gène et la longueur de la structure primaire de la protéine finale sont proportionnelles.
- Le code génétique est quasiment universel, dans quelques rares espèces le code a évolué et est très légèrement modifié. Notamment dans la synthèse des protéines mitochondriales, UGA ne code plus Stop mais Trp, AGG et AGA ne sont pas des codons Arg mais des codons Stop, AUA ne code plus Ile mais Met.
- Le code génétique est redondant (ou dégénéré). Plusieurs codons codent pour un même acide-aminé : on trouve 64 codons et 20 acides aminés.
- Le code possède un système de ponctuation ; le codon d'initiation est le codon AUG et les codons de terminaison sont les codons UAA, UAG et UGA.
- Le code génétique est non-chevauchant ; les nucléotides d'un codon ne participe qu'au code d'un seul acide aminé.
- Le cadre de lecture est défini par le codon d'initiation, ainsi le véritable codon de terminaison sera le premier codon qui sera dans le cadre de lecture imposé par ce codon d'initiation.

Seuls deux acides aminés ne montrent pas de redondance de leur code: la méthionine et le tryptophane. Le nombre de codons synonymes n'est pas le même pour tous les acides aminés, on distingue des duos (histidine, lysine, asparagine, glutamine, acide glutamique, asparagine,

tyrosine, cystéine et phénylalanine), des trios (isoleucine et codons stop), des quatuors (thréonine, proline, alanine, glycine et valine) et des sextuors (arginine, leucine et sérine) (tableau II).

Tableau II: Correspondance entre les triplets de bases de l'ARNm et les acides aminés et codon stop.

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

2. Décryptage du code génétique

Plusieurs expériences ont été nécessaires pour déterminer que la suite des acides aminés correspondait à l'ordre des nucléotides, que les codons étaient composés de triplets et quelle était la correspondance entre un triplé et un acide aminé. En 1961 le code génétique est totalement décrypté grâce aux expériences de Nirenberg et Matthaei. En ajoutant un ARN poly U à un extrait cellulaire (sans ARNm), ils obtiennent un poly Phénylalanine: donc UUU code pour la phénylalanine. De même, l'ARN poly A et poly C donne une polylysine et polyproline. Un Poly AC porte deux codons possible ACA ou CAC quel que soit le cadre de lecture, sa traduction donne un mélange équimolaire de la thréonine et l'histidine. Les autres codons ont été élucidés par la traduction des polyribonucléotides di/tri ribonucléiques.

L'identification des codons de ces aa est donnée par une autre expérience, il s'agit de synthétiser un polymère de 5A et 1C. Les codons AAA, A₂C (AAC.ACA.CAA) seront plus

fréquents que CCC, C₂A. L'analyse de la composition en aa des produits de la traduction indique que seule l'His peut avoir AC₂. Ils ont résulté que CAC code l'His et donc ACA code la thréonine. Cette correspondance codon-acide aminé permet de résumer le code génétique sous forme d'une table associant chacun des 64 codons ou triplets possibles (4³), dont 3 sont des codons de terminaison, avec l'un des 20 acides aminés présents dans les protéines.

3. Les acteurs de la traduction

Les acteurs de la traduction sont l'ARN messager (ARNm), les ARN de transfert (ARNt), les ribosomes, les acides aminés, les amino-acyl RNA synthétases, le Mg²⁺, le GTP et l'ATP.

3.1. Les ribosomes

Particules cytoplasmiques constituées d'une grande sous unité et une petite sous unité, toutes deux constituées d'ARN ribosomiques (ARNr). Les deux sous unités sont souvent libres et ne s'associent qu'au moment de la traduction de l'ARNm. La fonction principale de ribosome est d'orienter correctement les aminoacyl ARNt par rapport à l'ARNm et de créer les liaisons peptidiques.

- Les ribosomes procaryotes (70S) sont constitués d'une petite sous-unité 30S et d'une grande sous-unité 50S.
 - La sous-unité 30S est constituée d'un ARNr 16S (1541 nucléotides) et de 21 protéines.
 - La sous-unité 50S est constituée des ARNr 23S (2904 nucléotides) et 5S (120 nucléotides) ainsi que de 32 protéines.
- Les ribosomes eucaryotes (80S) sont constitués d'une petite sous-unité 40S et d'une grande sous-unité 60S.
 - La sous-unité 40S est constituée d'un ARNr 18S et de 33 protéines.
 - La sous-unité 60S est constituée des ARNr 28S, 5,8S et 5S ainsi que de 49 protéines.

Le ribosome bactérien comporte des sites spécifiques :

- **Site A** : (= site Acide-aminé ou Accepteur) fixation des acides aminés.
- **Site P** : (= site Peptidique ou Donneur) fixation de f-Met (Met chez les eucaryotes).
- **Site E** : (= site Exit) sortie de l'ARN de transfert.
- **Site EF-G** : présent au niveau de la grande sous-unité.
- **Site EF-Tu** : présent au niveau de la petite sous-unité.

L'enchaînement des ribosomes sur l'ARNm forme le **polysome**, il permet d'augmenter l'efficacité de la traduction.

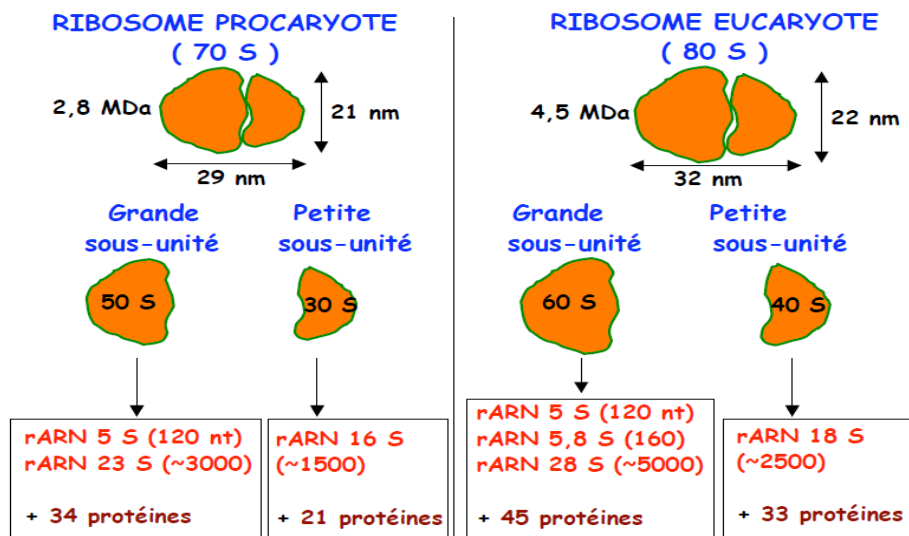


Figure 35: Structure générale et composition des ribosomes procaryotes et eucaryotes

3.2. Les ARNt

Les acides aminés n'ont pas d'affinité spécifique pour l'ARNm, ils s'attachent aux matrices d'ARNm par l'intermédiaires d'adaptateurs les ARNt.

3.2.1. Structure des ARNt et ARNt iso-accepteur :

Les ARNt ont une structure secondaire en forme de trèfle à 3 feuilles et une structure tertiaire en forme de L à l'envers (75 bases). Les ARNt portent l'un des 20 acides aminés attaché par une liaison ester à leur extrémité 3'-OH (CCA) et transportent celui-ci au ribosome. L'extrémité 5'-OH est toujours occupée par G. Dans une boucle située à l'autre extrémité de la molécule, se trouve une séquence de trois nucléotides, spécifique de l'acide aminé, appelée l'anticodon. L'anticodon s'apparie au codon sur l'ARNm assurant ainsi la correspondance entre codon et acide aminé, conformément au code génétique. Il existe donc plusieurs ARNt différents pour un acide aminé, on les appelle ARNt iso-accepteurs, alors que d'autres, les codons stop, ne sont reconnus par aucun.

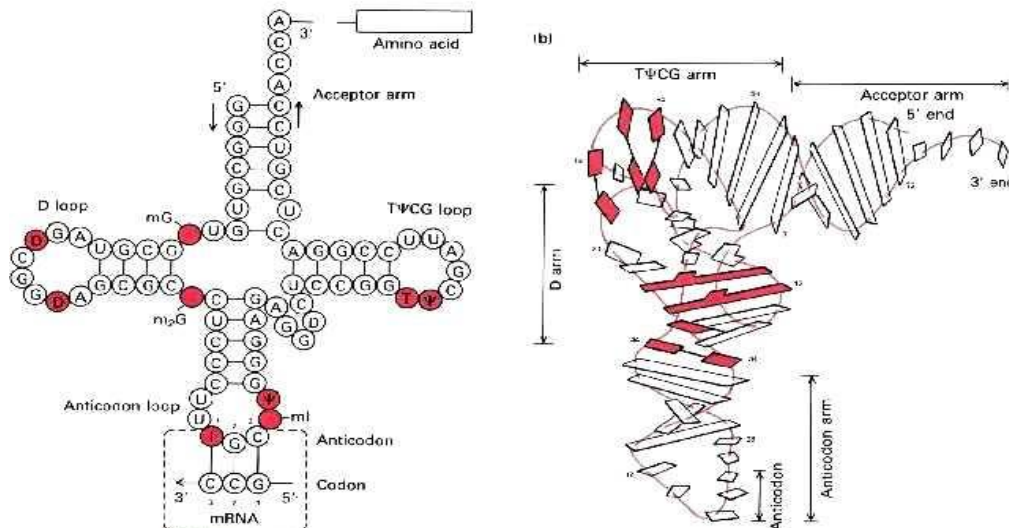


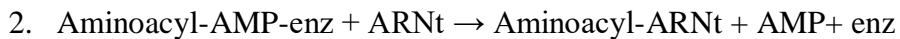
Figure 36 : Structures secondaire et tertiaire d'une molécule d'ARNt.

3.2.2. Chargement de l'acide aminé sur l'ARNt :

La formation du complexe amino-acyl-tARN (aa-tARN) nécessite une Amino-acyl-tRNA-synthétase spécifique de l'acide aminé, qui doit ainsi reconnaître toutes les formes de codon de cet acide aminé. Le chargement correct de l'ARNt est un élément important dans la fidélité de la traduction. L'acide aminé est tout d'abord activé et cette activation nécessite de l'énergie sous forme d'ATP pour permettre la formation d'aa-AMP (liaison anhydride mixte).



Dans une deuxième étape l'aa est transféré sur l'extrémité 3'OH de l'ARNt et forme une aminoacyl-ARNt (ARNt chargé)



La liaison formée entre l'ARNt et l'acide aminé est une liaison covalente de type carboxy-ester. Les Amino-acyl-tRNA-synthétase sont au nombre de 20 dans la cellule, autant qu'il y a d'acides aminés qui rentrent en compte dans la traduction.

4. Les différentes étapes de la traduction

4.1. L'initiation Un ribosome reconnaît le début de la séquence codante, il utilise des signaux d'adressage en amont entre -8 et -13 du codon initiateur (AUG) qui correspond à la

séquence de Shine-Dalgarno ou RBS (AGGAGG). Il y a appariement antiparallèle de bases entre l'ARNm et la petite sous-unité (30S) du ribosome, dû à une complémentarité de séquences entre l'ARNm et l'ARNr 16S. La grande sous-unité 50S du ribosome se joint à ce complexe pour former le ribosome 70S fonctionnel.

Les bactéries nécessitent un acide aminé particulier pour l'initiation ; cet acide aminé est la méthionine et elle nécessite une **formylation** sur l'extrémité NH₂ (ajout d'un formyl) pour former la f-Met dont le codon est AUG, c'est un phénomène pré-translationnel. Cette formylation est réalisée par vitamine B9 (ou tétrahydrofolate) qui reconnaît l'ARNt caractéristique et responsable du transport de la f-Met. La particularité de conformation de cet ARNt lui permet d'être placé directement dans le site P et non pas dans le site A.

L'initiation est permise grâce à la présence de facteurs d'initiation (IF) :

- **IF 1** est le facteur de dissociation du ribosome 70S.
- **IF 2** est un facteur assurant la fidélité de reconnaissance entre l'ARNt et l'acide aminé. Il possède également une activité GTPasique (d'hydrolyse du GTP).
- **IF 3** facteur nécessaire à la fixation spécifique de 30S sur l'ARNm et de contrôle de l'équilibre entre la forme associé et dissocié du ribosome (facteur anti-réassociation).

Chez les eucaryotes, les ribosomes reconnaissent la coiffe méthylée grâce à des facteurs Protéines fixatrices de la coiffe. Ils glissent le long de l'ARNm à partir de l'extrémité 5' jusqu'à la rencontre du 1^{er} codon AUG initiateur. Des facteurs d'initiation additionnels favorisent à cette étape l'effacement des structures secondaires (épingles) susceptibles d'arrêter le processus.

4.2. L'élongation

L'élongation correspond à une synthèse protéique par ajout d'acides aminés à l'extrémité C-terminale de la chaîne peptidique naissante, réaction catalysée par l'activité peptidyl-transférase de la grande SU des ribosomes. La lecture de l'ARNm par le ribosome se fait de 5' vers 3'. Il y a formation d'une liaison peptidique, les deux fonctions formant la liaison étant portées par le carbone- α de deux acides aminés différents.

La fixation du ribosome à l'ARNm s'opère de telle manière que fMet-ARN^{fMet} soit positionné dans le site P afin d'apparier au codon AUG de l'ARNm, une molécule

d'aminoacyl-ARNt se lie au site A adjacent au site P occupé. Une liaison peptidique est établie entre COOH du Met et NH₂ du second acide aminé grâce à l'enzyme peptidyl-transférase de la grande SU des ribosomes. L'ARN^{Met} déchargé quitte le site P. Le nouveau peptidylARNt est ensuite transloqué du site A au site P par un mouvement du ribosome de 5' vers 3' le long du complexe ARNm-ARNt-peptide.

L'allongement de la chaîne polypeptidique requiert 2GTP qui sont hydrolysées pour chaque acide aminé ajouté. L'élongation également est permise par la présence de facteurs d'élongation (EF) :

- **EF-Tu** ; réagit avec GTP et l'ARNt chargé pour former le complexe aminoacylARNt-GTP-EF-Tu et placer, grâce à l'hydrolyse du GTP en GDP, l'aminoacyl-ARNt au site A du ribosome.
- **EF-Ts** participe à la régénération du complexe **EF-Tu-GDP**.
- **EF-G** translocation du peptidyl-ARNt du site A au site P, qui agit en s'associant au GTP et au ribosome, il permet la translocation et la libération de l'ARNt libre au site P. La réutilisation de ce facteur nécessite là encore l'hydrolyse d'un GTP en GDP et Pi.

Chez les eucaryotes, les facteurs d'élongation eEF-1 et eEF-2 assurent les mêmes rôles que EF-Tu, EF-Ts et EF-G.

4.3. Terminaison

La terminaison de la traduction se fait au niveau des codons stop UAA, UAG et UGA qui ne codent pour aucun acide aminé. Ces codons stop sont reconnus par les facteurs de terminaison RF 1, RF 2 et RF 3 (RF pour *Releasing Factor*) :

- **RF 1** reconnaît UAA et UAG.
- **RF 2** reconnaît UAA et UGA.
- **RF 3** stimule l'activité des 2 autres facteurs.

Ils favorisent l'interaction de la chaîne en élongation avec l'eau et la libération de la chaîne.

La liaison ester unissant l'ARNt au dernier acide aminé de la chaîne peptidique est hydrolysée par la peptidyl-transférase. Le ribosome se dissocie en deux sous-unités qui pourront recommencer de nouvelles lectures d'ARNm.

La traduction eucaryote, les facteurs de terminaison sont du type eRF (pour *eukaryotic Releasing Factor*). Il y a pas de ré initiation de la synthèse protéique après la rencontre d'un codon Stop, chez les procaryotes, l'ARNm est couvert de ribosomes espacés d'environ 80 nts, c'est la polyribose ou polysome qui permet de produire plusieurs copies de la protéine à partir de la molécule d'ARNm .

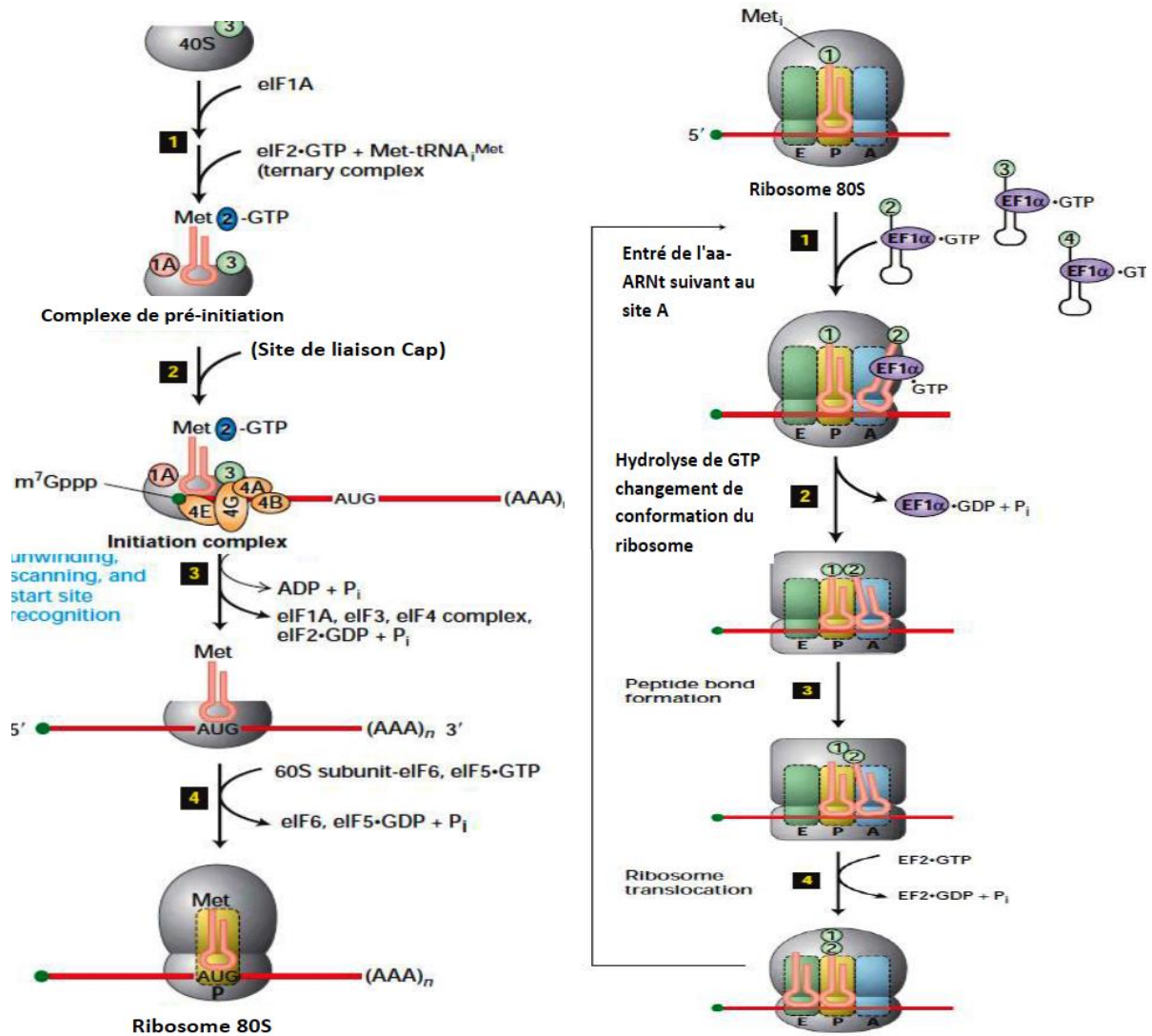


Figure 37: Formation du complexe d'initiation et élongation de la traduction chez les eucaryotes.

III. Régulation de l'expression des gènes

Les cellules peuvent s'adapter pour utiliser les ressources du milieu de manière optimale, ou elles peuvent se différencier. Cela suggère que des cellules, de matériel génétique identique, expriment leurs gènes de manière différente. Le contrôle de l'expression des gènes permet essentiellement à la cellule d'ajuster des synthèses en fonction des besoins nutritionnels, face à un environnement changeant, de façon à assurer la croissance et division cellulaire. La régulation des gènes se fait aux différentes étapes de l'expression et permet son activation ou sa répression. La régulation peut être positive grâce à des activateurs ou négative grâce à des répresseurs. Elle est assurée de plusieurs façons :

→ Régulation **quantitative** :

- Modulation du niveau de transcription d'un gène
- Régulation de l'accumulation des ARNm

→ Régulation **qualitative** :

- Structure des ARNm
- Modifications de la séquence codante

1. Régulation de l'expression des gènes dans la cellule procaryote

1.1. Au niveau transcriptionnel

L'expression d'un gène peut être régulée à plusieurs étapes, la plus commune étant celle de l'initiation de la transcription. Il y a deux raisons qui peuvent justifier cela.

- Le coût en énergie et en ressources pour la transcription
- La régulation lors de l'initiation est plus facilement réalisable.

Cette régulation est souvent contrôlée par des signaux extracellulaires, dans le cas des bactéries, ces signaux sont donnés par les molécules présentes dans le milieu. Ils sont ensuite transmis aux gènes par des protéines régulatrices. Ces protéines régulatrices contrôlent l'accès de l'ARN polymérase aux promoteurs de l'ADN bactériens (contrôle du taux de la transcription). Donc ces protéines aident les cellules d'évaluer leur environnement. Les gènes codant ces protéines sont appelés les gènes régulateurs.

Il ya deux types de protéines régulatrices :

- Les activateurs (régulateurs positifs: augmente la transcription d'un gène)
- Les répresseurs (régulateurs négatifs: diminue ou bloque sa transcription)

1.1.1. Contrôle négatif de la transcription "Répression et induction"

Chez les bactéries, un opéron c'est un ensemble de gènes transcrit en un seul ARNm polycistronique sous le contrôle d'un système régulateur unique et sous la direction d'un seul promoteur. Chaque opéron comporte un nombre variable de gènes de structure appelés cistrons. En général, toutes les protéines codées par un même ARNm polycistronique sont produites simultanément et de manière coordonnée, et interviennent dans une même séquence biochimique. Les gènes d'un opéron donné sont soumis à un contrôle coordonné.

Il existe deux grands types d'opérons :

Les opérons inductibles: codent pour des enzymes de la voie catabolique (voie de dégradation). Exemple : opéron lactose.

Les opérons répressibles : codent pour les enzymes de la voie anabolique (biosynthèse). Exemple: opéron tryptophane.

A. Régulation de l'expression des gènes impliqués dans les voies cataboliques

La cellule bactérienne a besoin d'une source de carbone, qu'elle trouve dans le catabolisme des sucres. Le lactose n'étant pas utilisable comme source directe, on a besoin d'un clivage du lactose (grâce à une β -galactosidase) pour obtenir du glucose et secondairement du galactose. L'opéron lactose est composé d'un gène de structure avec 3 cadres de lectures :

-**Lac Z** : code pour la **β -galactosidase** (clive le lactose en glucose et galactose)

-**Lac Y** : code pour la **β -galactosyl-perméase** (assure la perméabilité de la bactérie au lactose)

-**Lac A** : code pour la **β -galactoside trans-acétylase** (rôle inconnu de l'acétylation)

Les gènes de ces protéines faisant partie d'une même unité de transcription, on parle alors d'unité polycistronique, sous la direction d'une région régulatrice comprend le promoteur et l'opérateur et un gène régulateur (*lacI*) qui code le répresseur et agit en inhibant l'expression des gènes de l'opéron lactose par transactivation en se liant au niveau de l'opérateur son affinité pour l'opérateur sera modifiée en présence de lactose.

Le glucose pouvant être métabolisé directement par la bactérie et sera utilisé préférentiellement, il exerce une répression catabolique sur l'opéron lactose. Les gènes de l'opéron *lac* ont un fort niveau d'expression seulement en présence du lactose, et en absence

du glucose. Un signal de carence alimentaire est alors déclenché sous forme d'une augmentation du taux d'AMPc (synthétisée à partir de l'ATP par l'adénylate cyclase). Cet AMPc forme un complexe avec la protéine CAP (Catabolite Activator Protein, connue aussi par CRP pour cAMP Receptor Protein). Ce complexe se lie à l'ADN sur le site CAP en amont du site de fixation de l'ARN polymérase. L'ADN subit alors une inclinaison de 90°, qui stimule la fixation des éléments *trans* de l'ARN polymérase aux éléments *cis* du promoteur pour augmenter le taux de la transcription de 50 fois. L'interaction du complexe CAP-AMPc va agir comme un inducteur (contrôle positif) et augmenter l'affinité de l'ARN polymérase pour le promoteur de l'opéron *lac*.

En absence de lactose, le gène *Lac I* est exprimé et entraîne la formation du tétramère qui se fixe sur l'opérateur. Cette fixation entraîne une incapacité de l'ADN-polymérase à transcrire le gène dont le promoteur se situe avant l'opérateur.

En absence du glucose et la présence du lactose, le gène *Lac I* est également exprimé mais cette fois-ci chaque monomère du tétramère fixe l'allolactose (β -D-galactopyranosyl-(1-6)- β -glucopyranose, une molécule inductrice synthétisée à partir du lactose (β , 1-4) par une réaction de transglycosylation). Cette fixation entraîne la modification de la structure du répresseur qui ne peut plus se fixer sur l'opérateur permettant la transcription des gènes de l'opéron. De cette manière le lactose est l'inducteur physiologique.

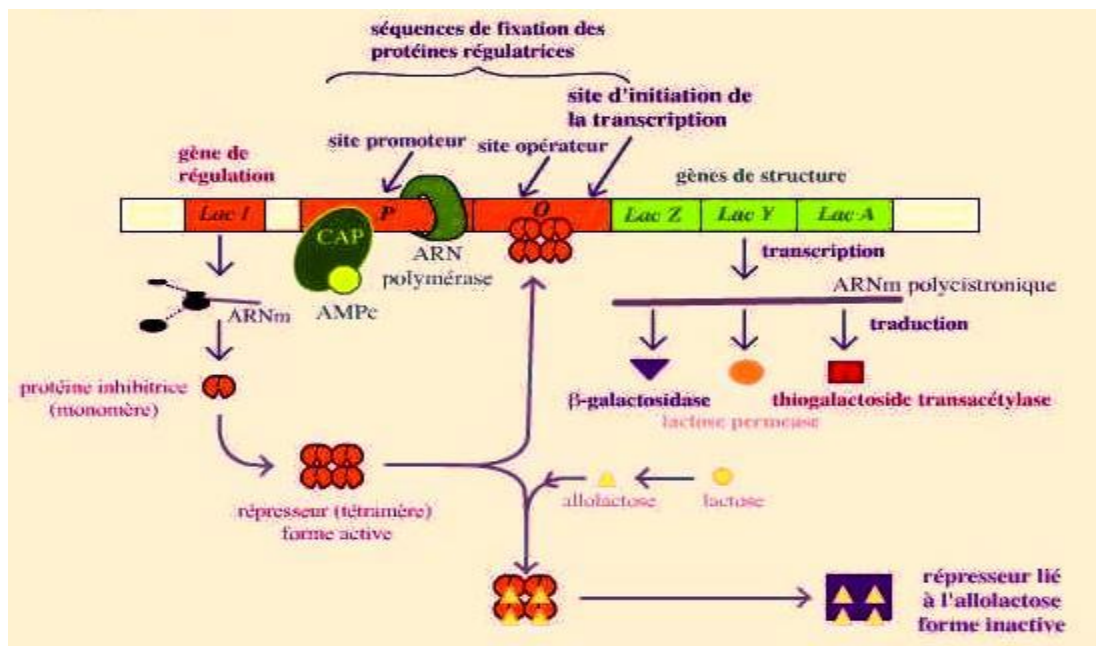


Figure 38: fonctionnement de l'opéron lactose en présence de lactose

B. Régulation de l'expression des gènes impliqués dans les voies anaboliques

Le contrôle de l'expression génétique par atténuation dépend du fait que la transcription et la traduction sont étroitement couplées. Chez *E. coli* les opérons de biosynthèse des acides aminés sont contrôlés par ce mécanisme. L'opéron Trp est un exemple de gènes répressibles, il est composé de cinq gènes de structure adjacents codant les enzymes qui synthétisent le tryptophane à partir du chorismate. Ces gènes sont contrôlés par un répresseur et ne s'expriment qu'à l'épuisement de tryptophane.

En absence de tryptophane, le répresseur ne se lie pas à l'opérateur ce qui entraîne la transcription des gènes de structure de l'opéron. En présence de tryptophane, le répresseur se lie au tryptophane et devient ainsi « actif », pouvant se lier à l'opérateur, ce qui entraîne l'inhibition de la transcription des gènes de structure de l'opéron. Le tryptophane est ici un co-répresseur.

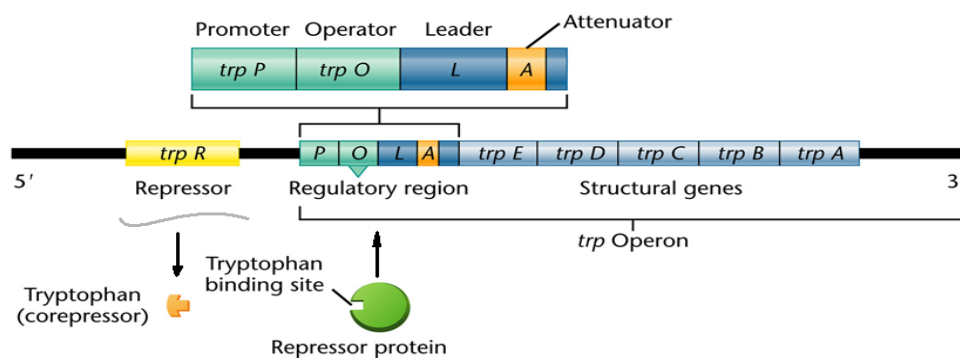


Figure 39: L'opéron tryptophane chez *E. coli*

Cependant, cette répression n'est pas totale et un second mécanisme entre en jeu lorsque le trp est abondant; c'est l'atténuation. Le phénomène d'atténuation contrôle la décision de transcrire un ARNm entier. Si les taux de tryptophane sont élevés, la plupart des transcrits terminent leur synthèse de façon prématurée (incomplète). Au contraire, si les taux sont insuffisants, la polymérase transcrit les gènes en entier.

L'atténuation est provoquée par un terminateur facultatif qui est situé au début de l'unité de transcription de l'opéron et que l'on appelle atténuateur. Une séquence leader de 161 nucléotides, comprise entre les nucléotides 27 à 68, localisés en amont du premier codon du gène *trpE*, elle contient 4 régions (1, 2, 3, et 4) (figure). Cette séquence permettant la synthèse d'un peptide de 14 acides aminés de longueur : dont deux codons tryptophane (en positions 10

et 11), alors s'il ya manque de tryptophane la synthèse du peptide leader sera incomplet à cause du manque des ARNt chargés par le Trp.

L'atténuation dépend de la capacité de l'ARN à former des structures secondaires entre des régions complémentaires (2=3 ou 3=4) (figure 40).

En fonction de la présence ou non de Trp, la vitesse du ribosome est différente pour traduire la séquence leader car la séquence leader code pour Z TRPs. Le ribosome pourra avoir deux positions sur l'ARNm et selon ces positions, il pourra y avoir formation de deux types de structures en épingle à cheveux différentes :

- Formation d'une boucle 3/4 ; boucle de terminaison, si le tryptophane est abondant: le ribosome ne s'arrête pas au niveau des codons trp; il continue à traduire la séquence leader, le reste de l'ARNm leader forme une structure secondaire entre les régions 3 et 4 constitue la séquence « atténuateur » et fonctionne comme signal de terminaison et suivi par une séquence riche en uracile qui provoque finalement la terminaison. La transcription s'achève avant que l'ARN polymérase atteigne les gènes permettant la synthèse du tryptophane.
- Formation d'une boucle 2/3 ; boucle d'anti-terminaison, le tryptophane est absent ou en quantité insuffisante, le ribosome s'arrête aux codons trp ce qui couvre la région 1, la région 2 s'apparie à la région 3, cette structure n'est pas un signal de terminaison de la transcription et empêche la formation du signal de terminaison. L'ARN polymérase continue à transcrire l'ensemble de la séquence codante ce qui permet la synthèse complète de l'ARNm.

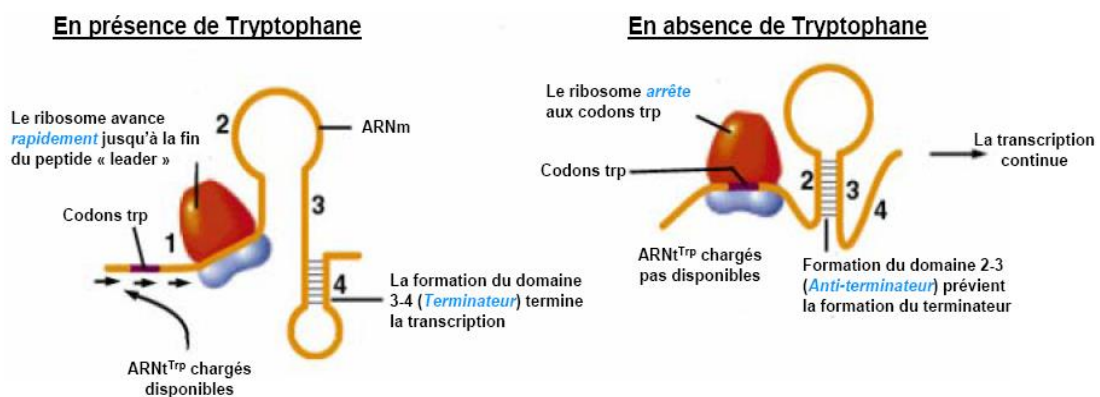


Figure 40: Processus de l'atténuation de l'opéron tryptophane d'*E. coli*

La régulation par atténuation est observée aussi pour au moins six opérons qui codent pour la biosynthèse des acides aminés par exemple: Thr, His, et le Phe. Les peptides leader

des opérons de synthèses de ces acides aminés contiennent dans leur séquences plusieurs monomères de l'acide aminé concerné (8 Thr, 7 His, 7 Phe). L'ensemble de ces opérons ne présentent pas la même combinaison d'éléments de contrôle de la régulation que celle de l'opéron *trp*. (Exemple: pour l'opéron *his*, l'atténuation est le seul mécanisme de contrôle).

1.1.2. Contrôle positif de la transcription

Dans ce type de contrôle, une protéine régulatrice active la fixation de l'ARN polymérase d'où le terme positif. Un excellent exemple de ce type de contrôle est l'opéron *mal*/EFG (l'opéron du catabolisme du maltose) chez *E. coli*.

En absence d'inducteur, ni la protéine activatrice, ni l'ARN polymérase ne peuvent se lier à l'ADN. La liaison d'une molécule inductrice dans ce cas-là c'est le maltose à la protéine activatrice provoque un changement de sa conformation (protéine allostérique), à son tour, se lie au site de fixation de l'activateur. Ceci permet à l'ARN polymérase de se fixer sur le promoteur et commencer la transcription (figure 41).

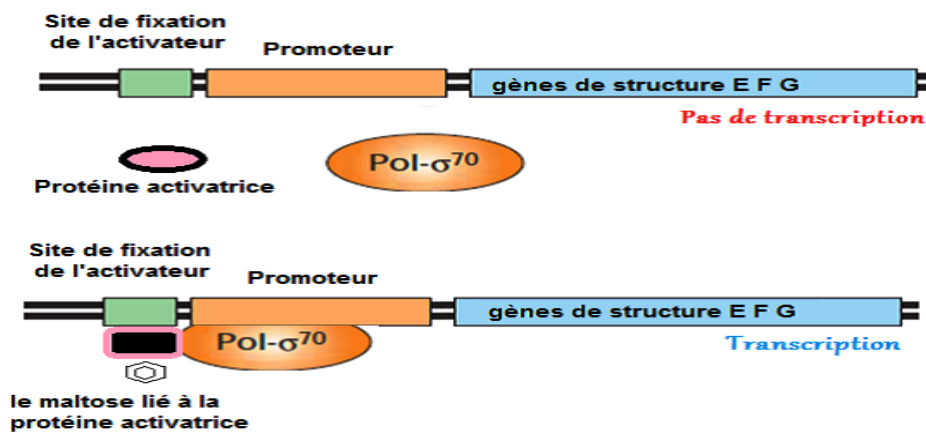


Figure 41: Mécanisme de contrôle positif de l'opéron *mal*/EFG chez

1.2. Régulation post-transcriptionnelle

Les ARNm ont un temps de demie vie très courte (*E.coli*). Ce temps de demi-vie dépend de l'importance de protéine, plus la protéine est importante plus son temps de demi-vie est élevé. Les enzymes d'un même opéron peuvent être produites en quantités inégales même si les gènes sont sous la dépendance d'un même promoteur. Ce phénomène est lié au fait que la dégradation commence souvent du côté 3'. Et à la quantité d'ARNm transcrit au départ et à la vitesse de propagation. Certains ribosomes n'ont pas le temps de synthétiser a b c et d et on peut avoir que du a ou que du a b etc. et donc on aura un niveau de protéine pas le même en fonction de la position sur l'ARNm. La régulation de l'expression des gènes dépend de l'efficacité des liaisons entre les Ribosomes et l'ARNm

1.3. Au niveau traductionnel

La régulation au niveau traductionnel est peu utilisée chez les procaryotes. On pourra donner l'exemple de la régulation de la synthèse de protéines ribosomiques dont les gènes sont organisés en opérons. Un excès de celles-ci entraîne l'inhibition de leur propre traduction.

2. Régulation de l'expression des gènes dans la cellule eucaryote

Dans les organismes multicellulaires l'expression de gènes différents est à l'origine d'une spécialisation cellulaire. Chez l'Homme on compte 250 types cellulaires différents de leur morphologie, leur biochimie et leur rôle dans l'organisme. L'expression des gènes eucaryotes est régulée au niveau chromatinien, au niveau transcriptionnel, au niveau post-transcriptionnel, au niveau traductionnel et au niveau post-traductionnel. L'ADN peut exister sous forme ;

- Compactée, l'hétérochromatine = état condensé les gènes ne sont pas accessibles aux polymérase, transcriptionnellement inactive, constitutive ou facultative
- lâche appelée euchromatine= organisé en boucle de 40-100 Kpb fixées à la matrice nucléaire (MAR = matrice associated region).

2.1. Au niveau chromatinien

Chez les eucaryotes, la chromatine peut se présenter sous forme d'hétérochromatine (ADN condensé autour d'un octamère d'histones : formation de nucléosomes) et d'euchromatine (ADN décondensé). Seules les séquences de l'euchromatine sont transcriptionnellement active et sensible à la DNase. L'hétérochromatine rend impossible la transcription. La régulation de la structure de la chromatine se fait par méthylation des cytosines ou acétylation des histones ainsi que par les complexes de remodelage de la chromatine (CRC) qui permettent la rupture des liaisons entre histones et ADN.

Les histones composant les nucléosomes peuvent présenter des modifications post-traductionnelles transitoires (réversibles) qui affectent la transcription (fonctionnent comme des répresseurs généralisés). Parmi ces modifications on trouve : l'acétylation, la méthylation, la phosphorylation et l'ubiquitination, ainsi que d'autres modifications plus rares (Tab. III). La modification des histones a un impact sur la structure de l'ADN et donc l'expression des

gènes : si l'ADN est plus condensé, il est moins accessible et aura moins de chances de s'exprimer. Les interactions entre ces histones et leur état sont complexes, et les modifications post-traductionnelles des histones ont été proposées comme formant un « code » qui dicterait l'état chromatinien et la régulation d'un gène ou d'un locus.

L'acétylation des histones induit une décondensation de la chromatine : en effet l'addition d'un groupement acétyles sur les histones crée des charges qui ont un rôle répulsif entre les nucléosomes. Ce phénomène est réversible et est contrôlée par des histones acétylases transférases (HAT). L'élimination d'un groupement acétyl se fait par des désacétylases.

La méthylation diminue la transcription, plus de 70% des séquences CpG sont méthylées dans l'ADN humain. 50% des promoteurs ont des îlots CpG et la méthylation d'un promoteur bloque la transcription, ainsi les promoteurs méthylés correspondent à des ADN satellites (hétéro-chromatine non transcrite).

Tableau III : État de la chromatine en fonction de l'acétylation des histones et la méthylation de l'ADN :

	Transcription active	Transcription inactive
État de la chromatine	Ouverte, sensibilité à la DNase I (euchromatine)	Condensée (hétérochromatine)
Acétylation des histones	Histones acétylées	Histones désacétylées
Méthylation de l'ADN (îlots CpG)	Faible méthylation sur promoteur	Forte méthylation sur promoteur

2.2. Au niveau transcriptionnel

La régulation concerne en général la phase d'initiation faisant intervenir les différents facteurs de transcription d'initiation qui se fixent à l'ADN et provoquent des effets négatifs ou positifs sur la transcription. La transcription peut aussi être régulée par des signaux extracellulaires (hormones stéroïdes, thyroïdiennes...qui agissent au niveau des récepteurs nucléaires).

Les gènes des eucaryotes possèdent des séquences régulatrices, souvent présentes en amont du promoteur de ces gènes. On appelle le motif d'ADN régulé un cis régulateur du type enhancers, silencers et insulators, et le facteur de transcription se fixant spécifiquement au cis régulateur de manière à l'activer ou à l'inhiber un trans-régulateur.

- Les cis régulateurs sont des séquences de 6 à 15 nucléotides, pouvant être placées en amont, entre, ou dans les introns de la séquence codante, régulent l'activité des promoteurs. Une séquence enhancer active les promoteurs de certains gènes en utilisant des facteurs de transcription. Les séquences silencers inhibent la transcription. Les promoteurs des gènes sont des séquences cis-régulatrices permettant la fixation de l'ARN polymérase II.
- Les trans-régulateurs reconnaissent ces séquences cis et activent ou inhibent leur expression. Ces facteurs protéiques possèdent un domaine de fixation sur l'ADN, un domaine d'action (répression ou activation) et souvent un domaine d'interaction avec d'autres ligands, dont les motifs structuraux sont :
 - **Les motifs de type hélice – tour – hélice** : structure courbe qui peut s'enchâsser dans le grand sillon de l'ADN, composée de 2 hélices alpha unies par une petite séquence. Ils agissent souvent sous forme monomérique.
 - **Les doigts de zinc** : structure stabilisée par un atome de zinc (Zn^{2+}) qui établit des liaisons covalentes avec 4 Cys ou avec 4 His ou avec 2Cys et 2 His. Ils fonctionnent en dimère.
 - **Les glissières à leucines** : structure en hélice alpha avec répétition de Leu tous les 7 AA ce qui crée un alignement de leucine (plateau hydrophobe) permettant l'appariement des protéines par liaisons hydrophobes. Il s'agit d'une structure hélicoïdale très stable. Elles fonctionnent en dimère.
 - **Les homéodomains** à charge positive avec des structures hélice–boucle–hélice : fixation au grand sillon de l'ADN, sous forme de dimère. C'est une structure composée de deux hélices alpha unies par une boucle.

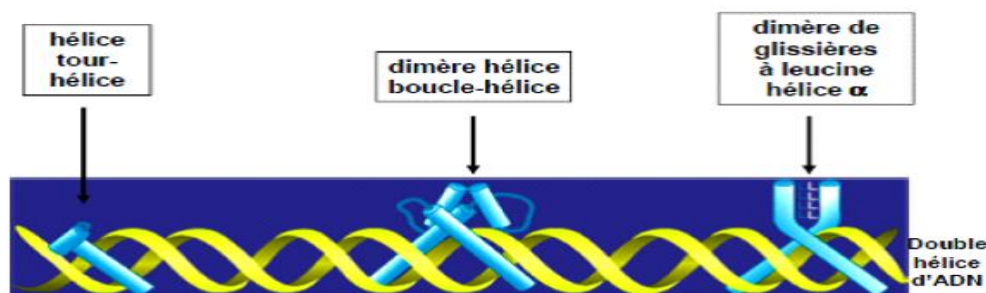


Figure 42: Interaction entre facteurs et séquences régulatrices

2.3. Au niveau post-transcriptionnel

2.3.1. Epissage alternatif

Production, à partir d'un gène commun de différentes protéines en fonction de l'environnement ou des tissus, il y aura une homologie de structure, ceci permet de créer des

familles de protéines aux fonctions voisines. Par exemple : le transcrite primaire de la calcitonine (pré ARNm) contient 6 exons et peut donner 2 ARNm matures différents ; Un qui contient les 4 premiers exons, il est produit dans la cellule thyroïdienne et donne la calcitonine. L'autre ARNm mature ne possède pas le quatrième exon (donc constitué des exons : 1, 2, 3, 5, 6), il code une protéine proche de la calcitonine dans les cellules de l'hypothalamus (CGRP : calcitonin rene related product).

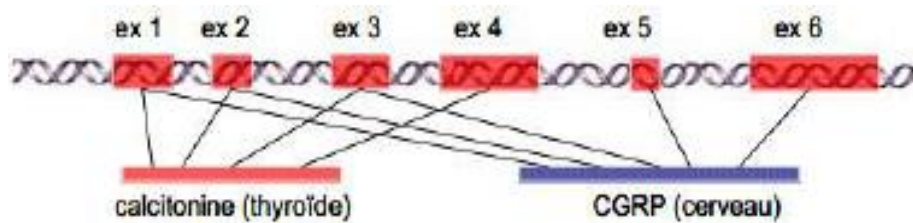


Figure 43 : Protéines en fonction différente

2.3.2. Modification éditorial de l'ARNm

Addition, suppression ou substitution d'un ou plusieurs nucléotides aboutissant à un ARNm dont la séquence nucléotidique diffère de celle du brin sens du gène. Exemple édition apolipoprotéine B (maturation en fonction du tissu) l'apolipoprotéine B codée par le gène Apo (protéine de transport des lipides dans le sang) ;

Dans le foie, l'ARNm donne une protéine de 4536 acides aminés (ApoB 100).

Dans les cellules intestinales, un C est désaminé en U → un codon CAA (Gln) est modifié en un codon UAA (codon stop). L'ARNm transcrit dans l'intestin donne une protéine de 2152 aa. (Apo B48). L'enzyme qui intervient est un éditosome : son activité est modulée par des facteurs spécifiques du tissu

2.3.3. Régulation de la durée de vie des ARNm

- Polyadénylation différentielle qui rend plus ou moins stable.
- Dégradation de l'ARNm par le produit de traduction (ex: les β -tubulines)
- Stabilisation de l'ARN en 3'
 - Éléments Riches en AU (ARE) et protéines de fixation (ARE-BP)
 - Protéines Puf et séquences riches en UG
 - Séquences de stabilisation riche en pyrimidines (α et β -globine, α -collagène)
- Stabilisation de l'ARNm par RE (responcive element) en 3' (ex: récepteur de la transferrine)
- stockage des ARNm dans les ovocytes.

2.4. Régulation traductionnelle

La traduction d'un ARNm commence à partir du premier codon AUG rencontré, cependant un ribosome peut sauter un codon initiateur. La détection de ce codon dépend des codons voisins et de la présence des facteurs de traduction (eIF).

- inhibition de la lecture de l'ARN par RE (responsive element) en 5' exemple de la synthèse de la ferritine

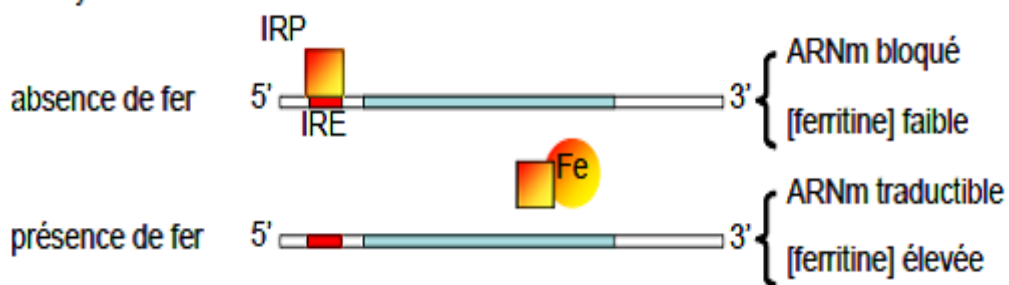


Figure 44 : Exemple de la synthèse de la ferritine ; inhibition de la lecture de l'ARN par RE

- inhibition des facteurs de traduction (initiation, élongation, terminaison), exemple de la synthèse de β -globine.

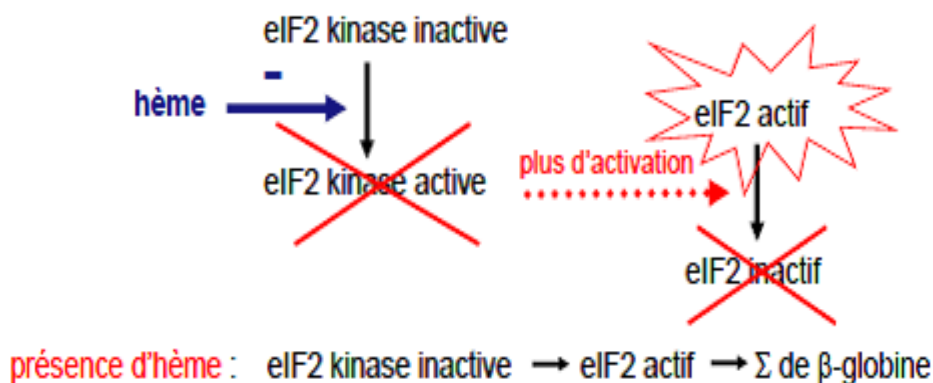


Figure 45 : Exemple de la synthèse de β -globine ; inhibition des facteurs de traduction.

- Régulation par les **micro-ARN** qui permettent un blocage de la traduction. Cette inhibition post-transcriptionnelle de l'expression des gènes se fait par interférence grâce à l'utilisation d'ARN anti-sens (si-ARN ou micro-ARN) qui s'apparient avec l'ARN sens par des séquences complémentaires. Cette inhibition peut être réalisée artificiellement par l'utilisation de trans-gènes qui correspondent à des gènes apportés artificiellement

2.5. Régulation post-traductionnelle

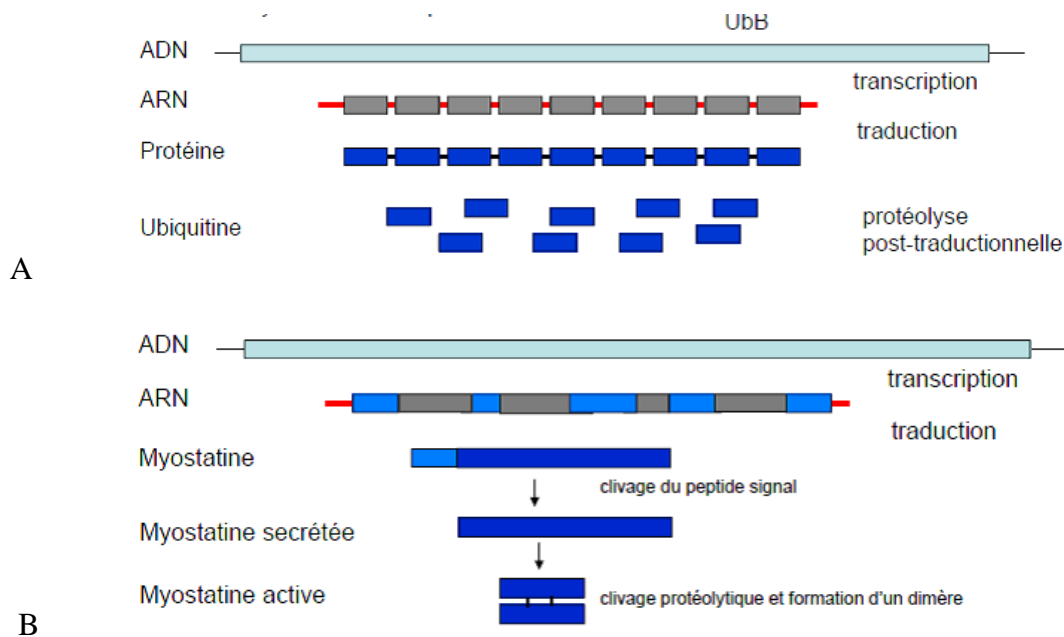


Figure 46 : (A) exemple de la biosynthèse de l'ubiquitine (protéolyse), (B) exemple de la myostatine (clivage protéolytique d'un dimère)

IV. Voies de régulation des gènes par les signaux extracellulaires

1. La signalisation intracellulaire par facteurs environnementaux

Les cellules sont soumises à l'action de facteurs environnementaux qui induisent une signalisation intracellulaire :

- **Les facteurs hydrophobes** pénètrent directement dans les cellules et agissent via des récepteurs nucléaires (hormones stéroïdes, certaines acides gras, hormones thyroïdiennes...).
- **Les facteurs protéiques** agissent via des récepteurs au niveau de la membrane cytoplasmique qui induisent une signalisation intracellulaire très souvent au moyen de cascades de phosphorylation de protéines.
- **Certains facteurs physiques** (stress mécanique ou thermique) ou **chimiques** (stress toxique) peuvent induire des signaux intracellulaires qui génèrent également des cascades de modifications biochimiques des protéines.

Tous ces signaux générés vont induire des modifications post-traductionnelles et fonctionnelles des facteurs régulateurs de transcription, qui agissent souvent en interaction entre eux. Les récepteurs sont donc au centre du réseau de signalisation.

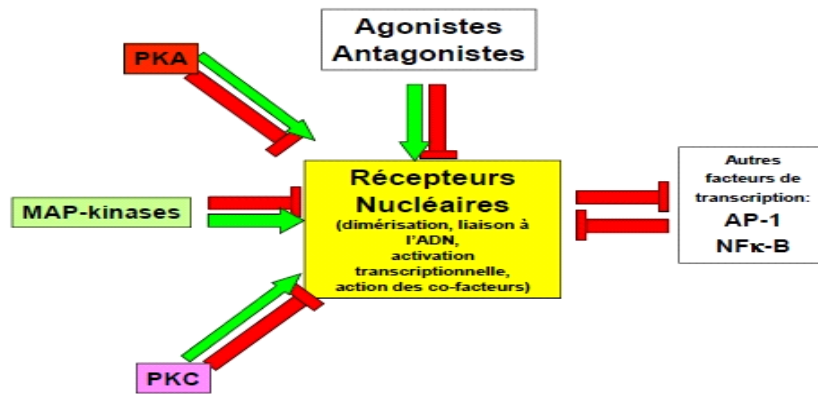


Figure 47: Influence des différents facteurs sur les récepteurs nucléaires

2. Modèle de régulation de l'expression d'un gène par un récepteur nucléaire

Un récepteur nucléaire est un facteur de transcription ligands-dépendants, comportant un domaine de liaison à l'ADN de structure en 2 doigts de zinc (chaque doigt à 4 Cys) et capable de fixer des hormones ((hormones stéroïdiennes et hormones thyroïdiennes) ou d'autres molécules lipophiles.

2.1. Structure des récepteurs nucléaires

Les récepteurs nucléaires comportent plusieurs domaines :

- Un domaine de liaison au ligand ("Ligand Binding Domain" - LBD).
- Un domaine de liaison à l'ADN ("DNA Binding Domain"- DBD) sur des séquences d'ADN particulières qui se trouvent à proximité des gènes qu'elles régulent. Ces séquences sont appelées éléments de réponse à l'hormone (HRE).
- La région N-terminale (domaine A / B) est la plus variable en ce qui concerne la taille et la séquence en acides aminés.
- Il existe une forte homologie de séquence au sein des domaines DBD et LBD, respectivement.
- Domaines activateurs de la transcription AF1 et AF2: vont interagir avec d'autres facteurs de transcription.

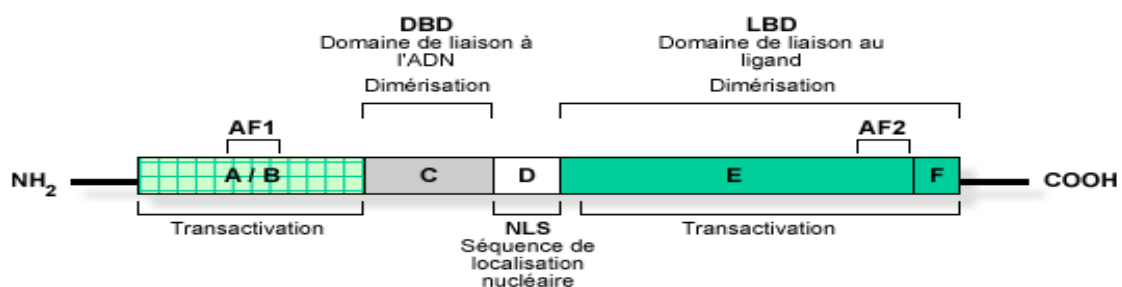


Figure 48: Structure des récepteurs nucléaires

2.2. Types de récepteurs nucléaires

Environ 48 récepteurs nucléaires ont été identifiés dans le génome humain. Ces derniers peuvent être regroupés en 3 classes, selon le type de leur ligand :

• Groupe 1 : homodimères

Récepteurs pour les hormones stéroïdes (glucocorticoïdes, progestérone, oestrogène, minéralocorticoïdes, androgènes...). Ils fonctionnent sous forme de dimères formés de récepteurs identiques après fixation de l'hormone, se fixant sur des régions spécifiques de l'ADN (HRE) : séquences répétées inversées.

• Groupe 2 : hétérodimères

Récepteurs pour quelques hormones et des médiateurs. Ils fonctionnent sous forme de dimères formés d'un récepteur généralement couplé à un partenaire constant, se fixant sur des séquences répétées dans le même sens.

• Groupe 3 : monomères (ou orphelins)

Récepteurs fonctionnant sous forme de monomères et se fixant sur des séquences uniques. La substance fixée est soit inconnue, soit inexistante (n'a pas besoin de fixer une substance).

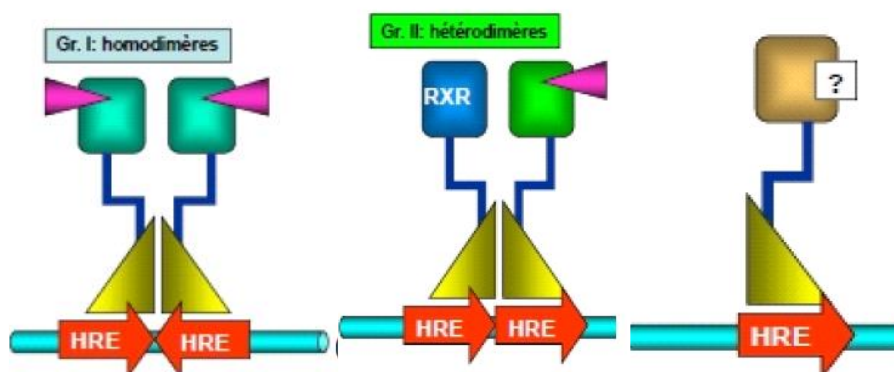


Figure 49: Types de récepteurs nucléaires

3.3. Mécanisme de régulation par les récepteurs nucléaires

Selon leur mécanisme d'action et leur distribution sub-cellulaire en absence de leur ligand, les récepteurs nucléaires peuvent être classés en 2 principales catégories :

- **Les récepteurs nucléaires de type I :** Situés dans le cytosol puis délocalisés dans le noyau. La fixation du ligand sur ce type de récepteur initialement situé dans cytosol induit :
 - la dissociation d'une protéine de choc thermique
 - l'homo-dimérisation du récepteur nucléaire

-la translocation du récepteur nucléaire (via un transport actif) du cytoplasme dans le noyau et enfin la fixation du récepteur nucléaire sur une séquence spécifique de l'ADN appelée élément de réponse à l'hormone (HRE).

Exemples: récepteurs des androgènes, récepteurs d'oestrogènes, récepteur des glucocorticoïdes et récepteur de la progestérone.

- **Les récepteurs nucléaires de type II :**

Quel que soit leur état (ligand fixé ou non fixé), les récepteurs de type II sont toujours maintenus dans le noyau et ils sont fixés sous forme d'hétéro-dimères (par exemple avec un autre récepteur de type "retinoid X receptor") à l'ADN.

1. en l'absence de ligand (hormone), les récepteurs nucléaires de type II sont souvent complexés à des protéines co-répresseurs
2. la fixation du ligand sur les récepteurs nucléaires de type II, ce qui entraîne un changement de conformation du récepteur
3. la dissociation des co-répresseurs et le recrutement de protéines co-activatrices de la transcription Ce qui entraîne le relâchement de la chromatine et permet la transcription.
4. des protéines supplémentaires (dont l'ARN polymérase) sont ensuite recrutées par le complexe [récepteur nucléaire de type II/ADN] pour la traduction de l'ADN en ARN messager.

Exemples: récepteur de l'acide rétinoïque, le récepteur X des rétinoïdes et le récepteur de l'hormone de la thyroïde.

Il existe 2 autres catégories "mineures" de récepteurs nucléaires :

- les récepteurs nucléaires de type III (principalement la sous-famille 2) ressemblent aux récepteurs nucléaires de type I puisqu'ils se fixent à l'ADN sous forme d'homodimères. Cependant, les récepteurs nucléaires de type III se fixent à une séquence HRE répétée directe au lieu d'une séquence HRE répétée inversée (exemple de séquence répétée inversée : 5'TTACGnnnnCGTAA3').
- les récepteurs nucléaires de type IV se fixent à l'ADN sous forme de monomères ou de dimères, mais seul un domaine de liaison à l'ADN du récepteur se fixe sur une seule moitié de la séquence HRE.

1. Techniques d'extraction et de purification des acides nucléiques

L'extraction et la purification des acides nucléiques sont essentielles pour un grand nombre d'études en biologie moléculaire et dans toutes les techniques de génies génétique. Le rendement et la pureté des acides nucléiques sont deux éléments importants pour assurer l'efficacité et la fiabilité des analyses. Pour certains échantillons, il est plus difficile d'obtenir un rendement optimal et/ou une pureté satisfaisante.

Il y a quatre étapes principales pour extraire les acides nucléiques

- La lyse des cellules
- La dénaturation des protéines et des complexes nucléoprotéiques
- L'inactivation des nucléases cellulaires par les agents chaotropiques (urée, chlorure de guanidine, thiocyanate de guanidine ...)
- La purification des acides nucléiques (ADN ou ARN) de débris cellulaires

Il existe aujourd'hui des kits commerciaux permettant de réaliser rapidement l'extraction et la purification à l'aide de réactifs prêts à l'emploi.

1.1. Méthodes d'extraction

On commence en général par une lyse des cellules ou des tissus, consistant éventuellement en un broyage, suivi d'une extraction par des détergents, qui vont disperser les bicouches lipidiques des membranes et dénaturer les protéines, et en particulier celles qui sont associées à l'ADN dans la chromatine. La solution obtenue est en général très visqueuse, car l'ADN ainsi libère forme de très longs filaments qui s'opposent aux écoulements hydrodynamiques.

1.1.1. Lyse des cellules

Les cellules peuvent être lysées, suivant le type de cellules concerné, par des méthodes chimiques, enzymatiques, ou mécaniques. Et doit être suffisamment rigoureuse pour briser le matériau de départ complexe, mais suffisamment douce pour préserver l'acide nucléique cible.

- **Lyse mécanique (Broyage ou lyse hypotonique)**

Cette méthode est particulièrement préconisée lors des extractions à partir des cellules sans paroi les bactéries, peuvent se glisser entre les faisceaux de forces de cisaillement et sont donc parfois très difficiles à lyser mécaniquement. Il est également possible d'utiliser cette méthode sur des cellules procaryotes ou des levures en ajoutant des microbilles ou du sable

(de fontainebleau, qui contient 95% de silice) pour faciliter la destruction des parois. Néanmoins, ce procédé est de plus en plus délaissé

- **Lyse chimique et enzymatique**

Implique l'utilisation de détergents, d'agents chaotropiques, la protéinase K.... Il se fait en plusieurs étapes :

-Fragilisation enzymatique des parois de cellules bactériennes, végétales et fongiques par des hydrolases spécifiques comme le lysozyme.

-On réalisera obligatoirement une lyse chimique pour extraire sélectivement les plasmides des bactéries. La lyse chimique permettent de récupérer dans le lysat uniquement les plasmides sans l'ADN génomique, par contre la lyse mécanique, on parle alors de lysat clair.

-Désorganisation des membranes par des détergents et solubilisation des lipides membranaires : Sodium Dodécyl Sulfate (SDS), triton X100, sarcosyl sont des détergents qui vont solubiliser les lipides membranaires sous forme de micelles. Cela permet de créer des pores membranaires suffisamment larges pour libérer le contenu du cytoplasme hors des cellules. Suivant leur force, les détergents vont aussi dénaturer les protéines membranaires et contribue également à la lyse de la cellule.

1.1.2. Elimination des protéines

Cette étape sera très importante pour l'élimination des histones eucaryotes. On peut procéder de deux façons différentes :

- **Déprotéinisation par hydrolyse enzymatique :**

Elle est réalisée en faisant agir une endoprotéase non spécifique comme la protéinase K, active jusqu'à 65°C c'est donc une enzyme très stable. Cette digestion est souvent conduite en présence d'un détergent dénaturant comme le SDS. Le SDS, facilite l'action de la protéinase K car il déploie la chaîne protéique.

- **Précipitation des protéines en utilisant un agent chaotrope :**

Un agent chaotrope est un ion qui modifie la solubilité des molécules (protéines ou acides nucléiques) et qui peut provoquer leur précipitation. Certains sont dénaturants (le perchlorate de sodium (NaClO_4), le thiocyanate de guanidine (TCG), l'iodure de sodium (NaI) et le chlorure de lithium) d'autres pas (le chlorure de sodium (NaCl) à forte concentration).

→ L'agent chaotrope peut agir en neutralisant certaines charges ioniques requises en surface pour le maintien de la solubilité.

- Il peut également agir en interférant dans les interactions que les protéines établissent avec l'eau ce qui modifie la solubilité des protéines. L'agent chaotrope entre en compétition avec la protéine pour établir des interactions avec les molécules d'eau disponibles dans la solution. L'eau se lie préférentiellement à l'agent chaotrope ce qui prive les protéines des molécules d'eau et provoque leur précipitation.
- pour certains, il peut aussi dénaturer les protéines, on le qualifie alors de dénaturant précipitant. Par exemple pour les protéines, il provoque la rupture des liaisons hydrogènes qui maintiennent leur structure tertiaire entraînant ainsi le démasquage de leurs régions hydrophobes. Les régions hydrophobes ont tendance à s'agréger et les protéines précipitent (défécation). Après centrifugation, il reste peu de protéines dans le surnageant qui contient, lui, tous les acides nucléiques.

1.1.3. Réactifs divers ajoutés au tampon d'extraction

Lorsqu'on utilise les agents chaotropiques pour éliminer les protéines par précipitation, on ajoute quelque fois dans le tampon d'extraction des thiols pour empêcher la reformation de ponts disulfures des protéines qui restent ainsi à l'état dénaturé. Une forte concentration saline (NaCl 0,15 M) du milieu d'extraction ne contenant pas d'agents chaotropiques empêche la séparation des deux brins de l'ADN en formant un écran protecteur de contre-ions autour de la double hélice : les ions Na^+ s'associent avec les groupements phosphate PO_4^- , les neutralisent et diminuent les forces de répulsion qui s'exercent entre les deux brins de même charge électrique, les ions Na^+ stabilisent ainsi la structure en double hélice. Le citrate de sodium et l'acétate de sodium aussi utilisés dans les tampons d'extraction jouent le même rôle = stabiliser l'ADN natif c'est-à-dire sous sa forme bicaténaire.

Le tampon contient souvent une substance chélatrice telle que l'**EDTA** qui séquestre par complexation les ions divalents ou le **citrate** qui possède 3 groupements carboxyle qui s'ils sont chargés COO^- piègent les ions positifs. Ces deux réactifs sont utilisés pour piéger notamment le magnésium, cofacteur des DNases et RNases afin de mieux préserver les acides nucléiques par inhibition des nucléases.

1.1.4. Elimination des ARN lors de l'extraction de l'ADN

On peut hydrolyser les ARN de deux façon, par :

- **Hydrolyse chimique** avec NaOH. A pH alcalin l'ARN est hydrolysé et pas l'ADN qui est protégé de la lyse alcaline par l'absence de groupement hydroxyle en 2' sur le désoxyribose
- **Digestion enzymatique** on utilise une RNase « *DNase free* », c'est-à-dire dépourvue d'activité DNase ; pour cela les DNases contaminant éventuellement les préparations de

RNase du commerce sont dénaturées par un chauffage (par exemple 5 min à 100°C) auquel résiste la RNase, enzyme particulièrement thermostable. La RNase est souvent apportée dès le début de l'extraction-purification grâce à sa grande stabilité.

1.2. Préparation de l'ADN plasmidique

Le principe de l'extraction est connu sous le nom de lyse alcaline. Cette méthode permet de préparer sélectivement l'ADN du plasmide contenu dans les bactéries, tout en éliminant l'ADN du chromosome bactérien.

Le principe de cette méthode consiste effectuer la lyse des cellules au moyen d'un détergent = SDS (dodécyl sulfate de sodium) en présence de soude, à pH 13. A ce pH très alcalin, l'ADN est dénaturé, c'est-à-dire que les deux brins de la double-hélice sont séparés. En effet, les groupements phosphate sont tous chargés négativement à pH basique et les forces de répulsion entre les deux brins sont à leur maximum. On neutralise ensuite rapidement la solution, ce qui provoque la renaturation brutale (réappariement des brins du duplex d'ADN). L'ADN chromosomique, très long (quelques Méga (10⁶) paires de base), ne parvient pas à se réappairier complètement et forme des enchevêtrements insolubles.

L'ADN plasmidique, court (~10³ paires de base), parvient à se reformer et reste en solution. On sépare alors les entités par centrifugation. Les protéines précipitées, sont également éliminées avec le détergent et l'ADN chromosomique. Cette solution ne contenant que très peu d'ADN chromosomique et de protéines est appelée lysat clair. L'ADN plasmidique est ensuite concentré par précipitation à l'alcool.

1.3. Méthodes de purification

Les méthodes de purification des acides nucléiques à partir d'extraits cellulaires sont généralement des combinaisons de deux ou plusieurs des techniques suivantes :

1.3.1. Extraction/Précipitation

L'extraction par solvants est souvent utilisée pour éliminer les contaminants d'acides nucléiques. Le procédé classique est l'extraction par le couple phénol-chloroforme, sert fréquemment à supprimer les protéines. La précipitation par l'isopropanol ou l'éthanol est généralement utilisée pour concentrer les acides nucléiques. Si la quantité d'acides nucléiques cibles est faible, un véhicule inerte (tel que le glycogène) peut être ajouté au mélange afin d'accroître l'efficacité de la précipitation. D'autres méthodes de précipitation des acides nucléiques incluent la précipitation sélective à l'aide de fortes concentrations de sel (« relargage ») ou la précipitation de protéines en utilisant les changements au niveau du pH.

1.3.1.1. Purification au phénol-chloroforme

Cette méthode utilise la solubilité différentielle des molécules (acides nucléiques/contaminants comme les protéines et les lipides) entre deux phases non miscibles. On mélange vigoureusement l'extrait d'acides nucléiques avec une phase hydrophobe. Après centrifugation, on récupère la phase aqueuse supérieure contenant les acides nucléiques.

- **L'extraction phénolique** est utilisée pour débarrasser les acides nucléiques des protéines car le phénol est un déprotéinisant puissant (très corrosif). Les protéines précipitent, elles restent à l'interface de la phase aqueuse et la phase phénolique qui est une phase hydrophobe. Les débris membranaires lipidiques vont aller dans la phase phénol hydrophobe. Le phénol doit être très pur et saturé en tampon (pH 8 pour extraire l'ADN).
- **L'extraction au chloroforme** : elle complète toujours l'extraction précédente pour éliminer toutes traces de phénol aqueux qui peut inhiber l'action ultérieure d'enzymes utilisées dans l'amplification, la restriction...etc.). Le chloroforme est additionné d'alcool isoamylique (AIA = 3-méthyl-1-butanol = $(\text{CH}_3)_2\text{CHCH}_2\text{CH}_2\text{OH}$) qui est un agent antimousse (agent déstabilisant de l'émulsion).

1.3.1.2. Précipitation/concentration par l'éthanol ou l'isopropanol.

Avant l'ajout de l'éthanol, il faut ajouter à la solution d'acides nucléiques, une quantité importante de cations (apport de Na^+ sous forme NaCl 3 M). Les cations Na^+ neutralisent les groupements PO_4^- de l'ADN ou de l'ARN et forment un écran protecteur de contre-ions autour de la double hélice, l'acide nucléique devient moins soluble car il établit moins d'interactions avec l'eau.

La précipitation est réalisée par addition de l'éthanol, elle doit être effectuée à haute force ionique (2.5 volumes d'éthanol 95° contre un volume d'échantillon). Pour les faibles concentrations, il faut prolonger le temps de précipitation (> 10h), comme il est possible d'accélérer la précipitation par le froid (-20°C à -70°C). Les très petits fragments de DNA même à hautes concentrations ne sont pas précipités, ce qui permet de les éliminer. Le lavage par l'éthanol 70° permet d'éliminer les sels. L'ADN est récupéré par centrifugation, et le séchage permet l'élimination de l'éthanol (s'évapore). Resuspendre dans un tampon adéquat 10mM tris EDTA, l'eau purifiée de nucléases...etc. Le sel n'est pas nécessaire dans la précipitation à l'isopropanol et les petits fragments d'ADN sont éliminés car non précipités. Dans ce cas, on procède à un mélange volume à volume.

1.3.2. Chromatographie

Les méthodes chromatographiques peuvent utiliser différentes techniques de séparation, telles que la filtration sur gel, l'échange d'ions, l'adsorption sélective ou la liaison par affinité. La filtration sur gel exploite les propriétés du tamisage moléculaire de particules de gel poreuses. Une matrice avec des pores d'une taille définie permet aux petites molécules de traverser les pores par diffusion, tandis que les plus grosses molécules sont exclues et éluées. Les molécules sont donc éluées afin de diminuer la taille moléculaire. La chromatographie par échange d'ions est une autre technique qui a recours à l'interaction électrostatique entre une molécule cible et un groupe fonctionnel sur la matrice à colonne. Les acides nucléiques (polyanions linéaires à forte charge négative) peuvent être élués des colonnes d'échange d'ions grâce à de simples tampons de sel. Dans la chromatographie par adsorption, les acides nucléiques sont fixés sélectivement par adsorption sur des silices ou du verre en présence de certains sels (ex. : des sels chaotropiques), alors que d'autres molécules biologiques ne se fixent pas. Un tampon ou une eau faible en sels peut ensuite éluer les acides nucléiques et produire ainsi un échantillon à utiliser directement dans des applications en aval

1.3.3. Centrifugation

La centrifugation sélective est une méthode de purification puissante. À titre d'exemple, l'ultracentrifugation isopycnique en gradients de CsCl à des forces g élevées a été longtemps utilisée pour la purification de plasmides. La centrifugation est souvent combinée à une autre méthode. Un exemple d'une telle utilisation est la chromatographie à colonne rotative qui combine la filtration sur gel et la centrifugation afin de débarrasser l'ADN ou l'ARN des contaminants de plus petit format (sels, nucléotides, etc.), pour l'échange de tampon ou pour la sélection de taille. Certaines procédures combinent l'adsorption sélective sur matrice chromatographique (cf. paragraphe ci-dessus « Chromatographie ») à l'élution centrifuge pour purifier sélectivement un type d'acide nucléique.

1.3.4. Séparation par affinité

La méthode est basée sur l'interaction entre les molécules de surface chargées positivement (ex. DEAE : diéthylaminoéthyl cellulose) et les phosphates présents dans les squelettes des acides nucléiques chargés négativement. Les molécules de surface sont fixées à un support solide souvent une résine de billes de silice. Les acides nucléiques se fixent sur les molécules de surface lorsque que la concentration en sel est faible. Les impuretés comme les protéines, les métabolites ... sont éliminés par des lavages avec des tampons contenant des concentrations modérées de sel.

Les ARN et les ADN peuvent également être éliminés sélectivement en optimisant la concentration en sels et le pH. Les acides nucléiques sont élués dans un tampon à concentration élevée en sel. Une précipitation à l'alcool en présence de sel est généralement recommandée pour concentrer les acides nucléiques après élution.

2. Séparation et Visualisation des acides nucléiques (Electrophorèse)

Technique qui permet de séparer des molécules en fonction de leur taille et de leur charge en utilisant un courant électrique. On peut ainsi analyser et purifier dans un milieu gélifié (gel d'agarose, gel de polyacrylamide...) l'ADN, l'ARN, les protéines.

Les fragments d'ADN après digestion par les enzymes de restriction peuvent être séparés par électrophorèse sur un gel d'agarose. Les migrations des fragments d'ADN dépendent de la taille du fragment plus que de la charge de celui-ci. Les fragments de petite taille auront une distance de migration la plus élevée que les fragments de taille élevée. Des marqueurs de taille sont déposés et migrent en parallèle avec les échantillons à analyser.

La détection de l'ADN sur ce type de gel est réalisée par exposition aux rayons UV après réaction avec un réactif spécifique, **bromure d'éthidium** par exemple, agent s'intercalant entre les bases. La mobilité relative est donc le rapport entre la distance de migration d'une bande et la distance de migration du front de migration. La droite $\log(\text{nombre de kb}) = f(\text{mobilité relative})$, permet de déterminer la taille d'une molécule d'ADN inconnue (figure).

L'électrophorèse des fragments d'ADN en gel d'agarose permet des séparations jusqu'à 20 kpb. Des fragments d'ADN de taille restreinte (inférieure à 1000 paires de bases) peuvent être séparés par électrophorèse sur gel de polyacrylamide. Le tampon utilisé pour la migration électrophorétique a un pH basique (par exemple: 8,3 dans le cas du tampon appelé TBE = Tris-Borate-EDTA).

D'autres techniques électrophorétiques existent comme l'électrophorèse en champ pulsé qui permet de séparer des grands fragments d'ADN (taille supérieure à 50 kb). Le support d'électrophorèse est constitué par de l'agarose et on applique au gel un champ électrique de nature variable. Ainsi, le champ électrique peut être appliqué dans une direction donnée pendant 1 minute puis dans une direction perpendiculaire au champ précédent pendant une minute et ainsi de suite. Le processus de réorientation des macromolécules dans le champ électrique dépend de leur taille.

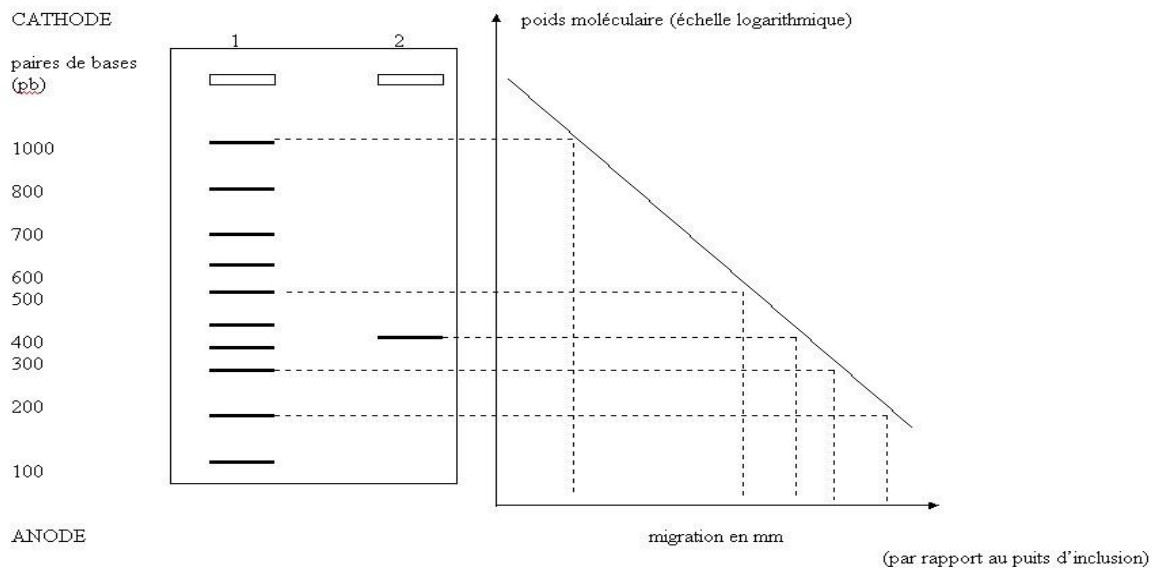


Figure 50: Distance de migration en fonction du logarithme de la taille des fragments d'ADN. 1: Marqueurs de taille (en pb), 2: Echantillon d'ADN dont la taille est à déterminer.

3. Analyse de la pureté de l'ADN extrait

Plusieurs paramètres basés sur les absorptions spectrophotométriques sont utilisés pour analyser la pureté d'une solution d'acides nucléiques. Le maximum d'absorption des acides nucléiques se situe à 260 nm, et pour les protéines à 280 nm.

3.1. Le ratio A260 nm/280 nm

L'intérêt de l'utilisation du ratio A260 nm/280 nm pour détecter la contamination d'acides nucléiques présente dans une solution par les protéines ou par les ARN. D'une manière générale, l'ADN double brin pur devrait avoir un ratio d'environ 1,8, tandis que l'ARN pur devrait avoir une valeur d'environ 2,0. Une contamination par les ARN se traduit par une augmentation du rapport $R > 1,8$. Alors une contamination par les protéines se traduit par un rapport $R < 1,8$. Le coefficient moyen d'absorption d'un nucléotide est supérieur à celui du même nucléotide dans la double hélice à cause de l'hypochromisme.

3.2. Le ratio A260 nm/A230 nm

L'absorption à 230 nm reflète la contamination de l'échantillon par des substances telles que les hydrates de carbone, les peptides, les phénols ou les composés aromatiques. Dans le cas d'échantillons purs, le ratio A260/A230 devrait être d'environ 2,2. Une méthode alternative, en l'occurrence la méthode de la plaque d'agarose au bromure d'éthidium, est utile lorsqu'on ne dispose que de faibles quantités d'acide nucléique. La quantité d'acide nucléique peut être estimée par comparaison à une gamme de concentrations en utilisant l'intensité de la fluorescence émise par le bromure d'éthidium lorsque celui-ci est irradié par la lumière UV.

4. Quantification de l'ADN

L'estimation des quantités d'ADN est indispensable après extraction d'ADN à partir d'un matériel biologique.

4.1. Dosage fluorimétrique classique de l'ADN

Le principe est un fluorochrome qui se fixe sur le DNA et dont le taux de fixation est directement lié à la quantité de DNA émettra une quantité de fluorescence qui sera proportionnelle à la quantité de DNA présente. On compare ensuite les valeurs de ces intensités avec celle d'une solution d'ADN standard (courbe standard). Les différents fluorochromes:

- Se fixant spécifiquement sur des paires de bases: Bases A-T: Hoechst 332 et DAPI
Bases G-C: mithramycine
- Intercalants: Iodure de propidium, Bromure d'ethidium et Acridine orange

4.2. Dosage de l'ADN par fluorescence en présence de BET

Le Bromure d'Ethidium (BET) interagit avec l'ADN en s'y intercalant et une fois intercalé, le rendement de fluorescence du colorant devient cent fois plus important. Le principe de la quantification consiste à comparer l'intensité de la fluorescence émise par l'ADN sur un gel d'électrophorèse. La quantification exacte consiste alors à établir une courbe d'étalonnage donnant l'intensité de fluorescence d'une solution de BET à laquelle on ajoute des quantités croissantes de DNA de concentration connue (Gamme d'étalonnage). Par cette méthode on peut déceler des quantités de DNA de l'ordre de 50 ng/ml.

4.3. Dosage colorimétrique de l'ADN

Il est basé sur la réaction spécifique des 2-désoxypentoses avec la diphénylamine. En milieu acide à chaud, le 2-désoxyribose des nucléotides purs peut être libérés et former un composé bleu dont le maximum d'absorption se situe à 595 nm. C'est la méthode classique pour mesurer des quantités importantes de DNA (de l'ordre de quelques milligrammes).

4.4. Dosage par absorption U.V. de la concentration en ADN

A 260 avec un trajet optique de 1 cm, une unité d'absorbance correspond à une concentration d'ADN double brin de 50 µg/ml ou à l'absorption d'une solution d'ADN simple brin (ou d'ARN) à la concentration de 25 µg/ml. Cette méthode est peu sensible pour manipuler des concentrations d'ADN inférieures à 250 ng/ml.

4.5. Dosage par la PCR temps réel

La PCR en temps réel permet de mesurer l'accumulation du produit de PCR à chaque cycle au cours de la réaction d'amplification. Le principe est d'utiliser un marquage

fluorescent du produit de PCR. L'appareil de PCR en temps réel n'est autre qu'un thermocycleur classique sur lequel vient s'ajouter un détecteur de fluorescence, il mesure l'intensité de la fluorescence en fonction du nombre de cycles (cette technique sera détaillée ultérieurement).

5. Hybridation des acides nucléiques

5.1. Généralités

L'hybridation est une propriété fondamentale des acides nucléiques qui repose sur les règles de complémentarité. Il est possible d'apparier des brins d'ADN ou ARN avec des oligonucléotides qui reconnaissent spécifiquement des séquences sur les brins d'ADN de manière antiparallèle et complémentaire. Ces oligonucléotides sont appelés sondes nucléiques (figure 51). L'un des deux brins est marqué par couplage chimique avec une molécule pouvant générer un signal. L'hybridation moléculaire est utilisée surtout dans la détection de l'homologie entre les molécules d'ADN de sources différentes. La complémentarité dépendra de la spécificité et de la sensibilité.

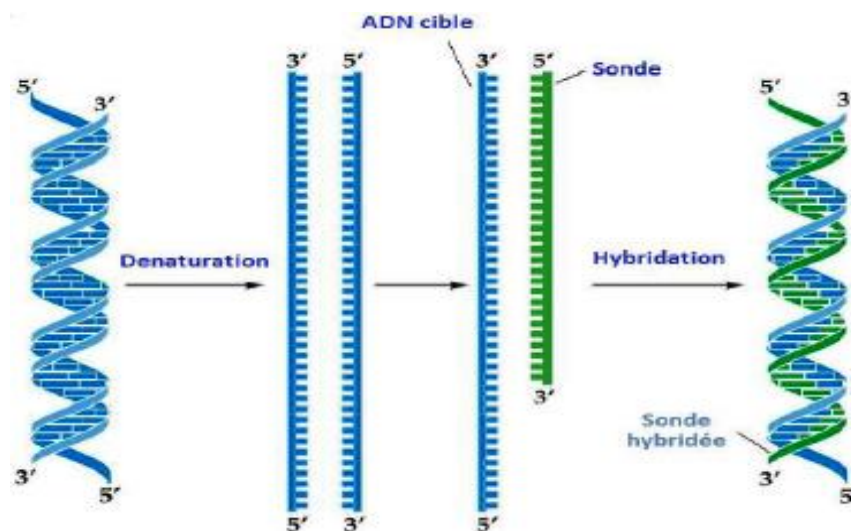


Figure 51: Hybridation des acides nucléiques.

La formation et la stabilité des duplex dépendent de nombreux facteurs en plus de la composition en bases : longueur des duplex, complexité des séquences, la température d'hybridation et la force ionique.

Pour le cas de deux brins complémentaires de la même molécule d'ADN (brins homologues) on parle de renaturation (hybridation à 100%). Chaque 1% de non homologie diminue la température d'hybridation (T_h) de 1°C . La T_h est estimée en utilisant plusieurs

formules, cela dépend de la longueur du fragment (nombres de nucléotides), composition du milieu, mis-appariement et le contenu en C+G

5.2. Types d'hybridation

L'objectif de l'hybridation est la détection de la présence d'un acide nucléique d'une séquence donnée par l'utilisation d'un fragment d'ADN complémentaire = sonde. Cela peut avoir lieu en solution ou sur support solide (immobilisation de la cible sur une membrane [nitrocellulose, nylon], sur verre, colonies bactériennes, chromosomes, plage de lyse ...etc.).

5.2.1. Hybridation d'ADN sur support solide : Southern blot

Cette méthode consiste à détecter spécifiquement des fragments d'ADN transférés sur filtre par leur hybridation à des séquences complémentaires marquées par un radioisotope. Les étapes successives de cette technique sont les suivantes : (figure 52)

- 1- Extraction de l'ADN génomique.
- 2- Digestion par les enzymes de restriction
- 3- Electrophorèse du produit de la digestion sur gel d'agarose.
- 4- Dénaturation des fragments par un traitement alcalin du gel d'électrophorèse. Ce traitement transforme les fragments d'ADN bi caténaire en monocaténaire.
- 5- Transfert sur membrane de nitrocellulose ou de nylon par exemple
- 6- Hybridation avec une sonde spécifique marquée.
- 7- Lavages pour éliminer les sondes non appariées
- 8- Révélation par autoradiographie

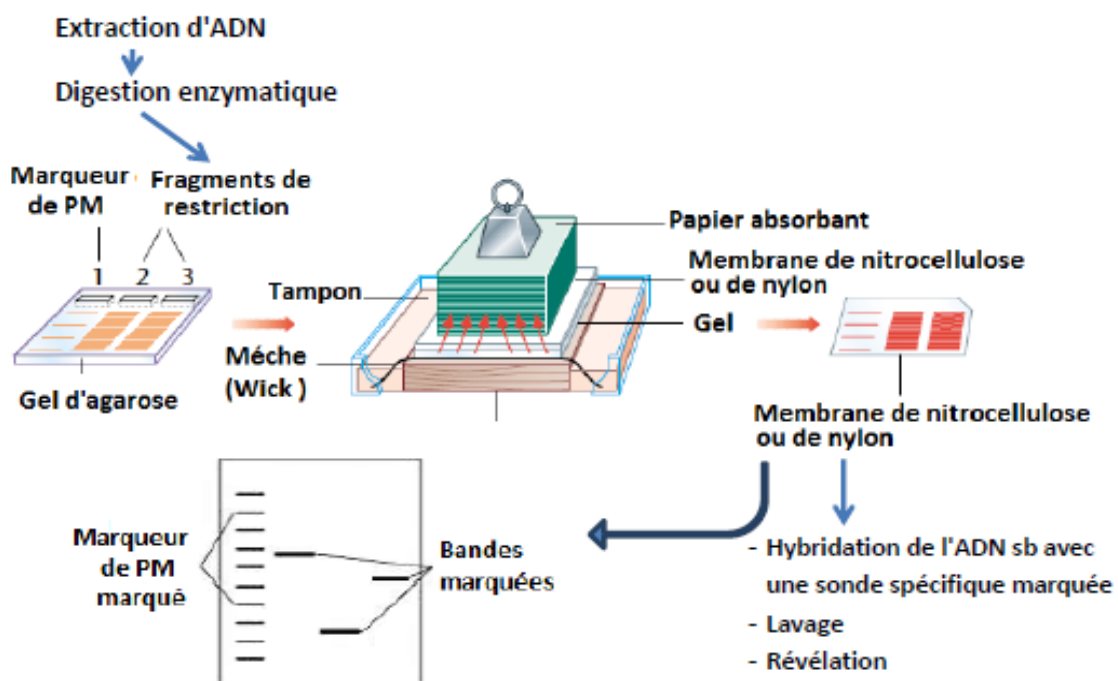


Figure 52: Schéma des différentes étapes d'un Southern blot

On peut aussi révéler la présence des séquences spécifiques dans les ARN (Northern Blot) et des protéines (Western Blot) utilisant le même principe.

Nous citerons parmi les applications de la technique de SOUTHERN:

- **Carte de restriction de l'ADN**

Les enzymes de restriction coupent l'ADN au niveau de séquences parfaitement définies. Il est donc possible d'établir une véritable carte d'un gène donne par la technique de Southern. Cette carte porte le nom de carte de restriction.

En pratique, l'ADN à étudier est reparti en différentes fractions. Chaque fraction est traitée par une enzyme de restriction ou un couple d'enzymes de restriction. Ainsi dans l'illustration ci-dessous. Dans le puits 1, on dépose l'ADN digéré par Eco RI seule; dans le puits 2, on dépose l'ADN digéré par Hind III seule et dans le puits 3, l'ADN digéré par Eco RI et Hind III. La détection des différents fragments est réalisée à l'aide d'une sonde marquée du gène étudié.

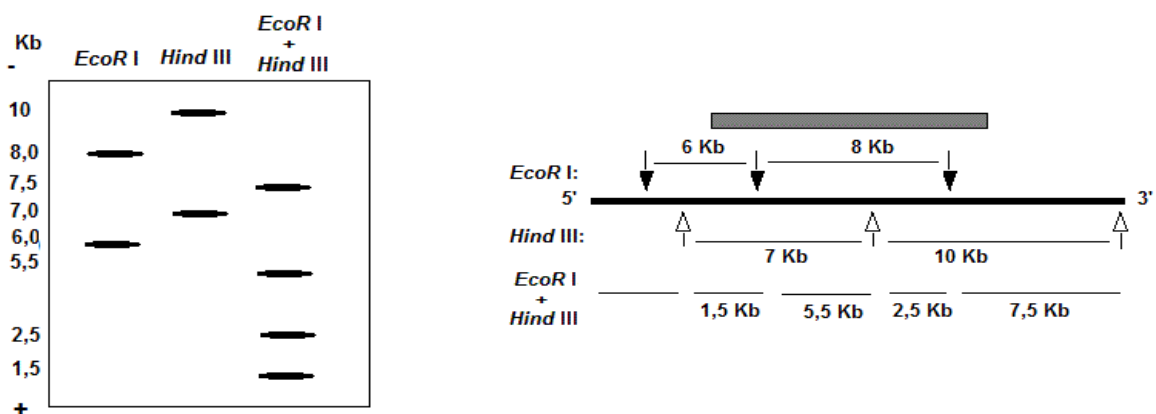


Figure 53: Illustration des positions sur l'ADN et des fragments après migration

- Si on dispose de sondes spécifiques du gène à explorer. Mise en évidence des délétions dans un gène ou une zone juxta-génique. L'étendue de la délétion peut être estimée par la détermination de la taille des fragments d'ADN.
- Comparaison de l'ADN de différents individus avec mise en évidence des polymorphismes de restriction. En présence de mutation(s) ponctuelle(s) dans des sites de restriction, des variations dans la taille de certains fragments seront observées et ceci à l'intérieur d'une même population (polymorphisme de taille des fragments de restriction (RFLP)).
- Mise en évidence des recombinaisons entre des gènes.

5.2.2. Hybridation *in situ* (HIS)

Cette technique permet de déterminer sur quelle séquence et dans quelle région se trouve le gène d'intérêt. Elle est très proche, des Southern et des Northern blots et repose, comme eux,

sur l'hybridation d'une sonde marquée avec une séquence complémentaire d'acides nucléiques que l'on cherche à identifier et à localiser. Les chromosomes fixés sur lame sont hybridés avec une sonde marquée (H3). L'examen microscopique révèle les signaux positifs sous forme de grains noirs disposés sur le chromosome. Comme exemple d'application est la détection des gènes viraux intégrés dans les chromosomes.

5.2.3. Hybridation sur colonie de bactéries, sur plage de lyse

Le principe est le même qu'en Southern ou en northern mais les fragments d'ADN sont séparés par clonage dans des bactéries. Ce type d'hybridation permet la détection parmi un grand nombre de bactéries ou de phages recombinants (plage de lyse) qui contient le fragment d'ADN cible. La réalisation de cette technique passe par les étapes suivantes:

- Culture pure sur boîte de Petri (colonies ou plage de lyse)
- Transfert sur membrane de nylon ou de nitrocellulose
- Lyse alcaline des bactéries et dénaturation de l'ADN
- Fixation par la chaleur ou les UV
- Hybridation avec une sonde marquée
- Lavage
- Autoradiographie
- Localisation des clones d'intérêt.

6. Amplification in vitro des acides nucléiques

6.1. Réaction de polymérisation en chaîne PCR

La technique de Réaction de polymérisation en chaîne PCR décrite en 1985 par Kary MULLIS, permet d'amplifier des séquences d'ADN de manière spécifique et d'augmenter de manière considérable la quantité d'ADN dont on dispose initialement. Elle nécessite de connaître la séquence des régions qui délimitent l'ADN à amplifier. Ces séquences serviront à synthétiser des amorces oligonucléotidiques complémentaires (de longueur de 20 à 30 nucléotides en général). Ces oligonucléotides serviront à délimiter la portion d'ADN à amplifier. L'ADN polymérase les utilisera comme amorces.

6.1.1. Les étapes de PCR :

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique. Pour effectuer ces transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil

programmable : un thermocycleur. Le nombre de cycles est généralement compris entre 30 et 40, ils génèrent 10^6 à 10^9 fois la séquence cible.

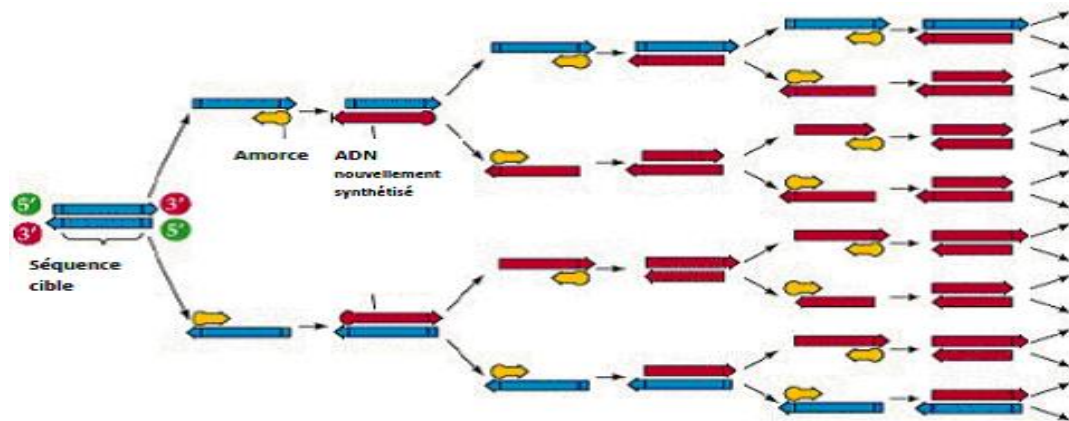


Figure 54 : Amplification d'ADN par PCR.

La dénaturation de la double hélice s'effectue à haute température, en utilisant une enzyme résistante aux hautes températures c'est l'ADN Taq polymérase thermiquement stable extraite d'une bactérie thermophile: *Thermus aquaticus*.

Cette enzyme nommée ADN polymérase est stable à 95°C , l'ADN étant copié à 72°C et non plus 37°C , son utilisation augmente la spécificité de la PCR. Car aux hautes températures l'hybridation non spécifique d'amorces à l'ADN non cible est rare. Les produits de l'ADN Taq polymérase sont donc plus homogènes que ceux obtenus avec l'ADN polymérase III d'*E. coli*. La technique comporte des cycles successifs. Chaque cycle comprend une succession de trois phases:

- 1- **Dénaturation** ($\sim 95^{\circ}\text{C}$) : Sert à dénaturer l'ADN pour obtenir des matrices simple brin.
- 2- **Hybridation** ($\sim 55-60^{\circ}\text{C}$): avec les deux amorces spécifiques entre $55-60^{\circ}\text{C}$. La première amorce se fixe sur un brin d'ADN, l'autre sur le brin complémentaire.
- 3- **Extension** ($\text{à } 70-72^{\circ}\text{C}$) : L'ADN polymérase prolonge les amorces (côté 3'-OH libre) en utilisant les brins cibles comme matrice.
- 4- Après une période appropriée d'incubation, le mélange est chauffé de nouveau pour séparer les brins. Il est alors refroidi pour que les amorces s'hybrident aux régions complémentaires de l'ADN nouvellement synthétisé. Le processus complet est alors répété « cycle ».

Les résultats doivent être optimisés en fonction d'un certain nombre de paramètres: concentration en MgCl_2 , concentration en amorces, spécificité des amorces etc... Le choix des amorces est particulièrement crucial pour obtenir des résultats satisfaisants (spécificité, taille,

paramètres physico-chimiques.....). L'introduction de logiciels spécialisés et des bases de données nucléotidiques a permis de réaliser des choix plus rationnels.

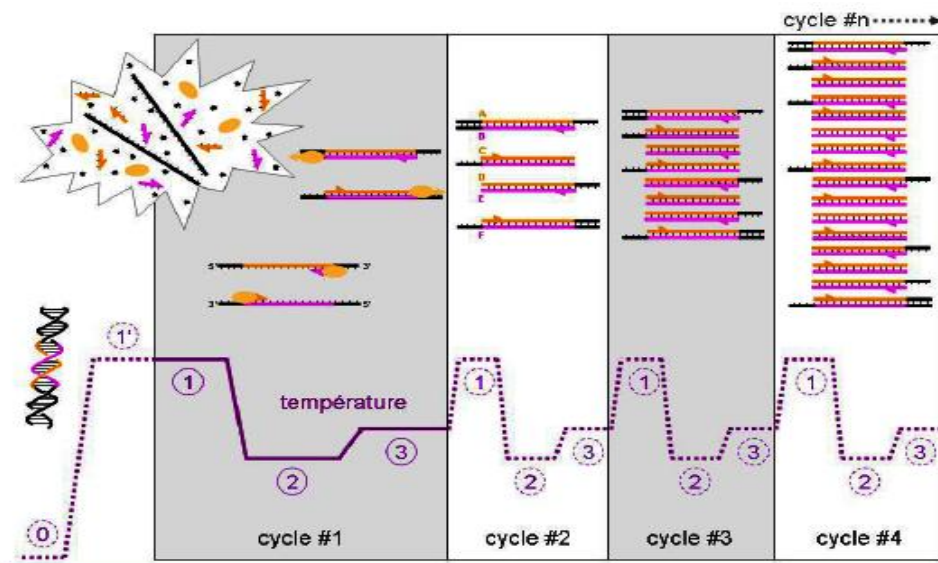


Figure 55 : Evolution de la température et des différents types de brins d'ADN au cours des 4 premiers cycles de la PCR

6.1.2. Types de PCR

Cette technique a évolué considérablement. De nouveaux types de PCR ont été introduits. Nous citerons brièvement:

- La PCR dite « Multiplex » pour amplifier des gènes avec de nombreux exons (le gène CFTR impliqué dans la mucoviscidose possède 27 exons), il est possible d'introduire dans le milieu d'amplification des couples d'amorces spécifiques différentes.
- La PCR dite « Nested PCR ». Elle correspond à une seconde PCR réalisée en utilisant des nouvelles amorces situées à l'intérieur du domaine défini par les amorces de la première PCR.
- La PCR quantitative. Dans ce type de PCR, on cherche à estimer le nombre de copies présent dans la séquence cible d'ADN ou d'ARN. La proportionnalité entre le nombre d'amplifications et le nombre de copies n'est valable que pour un nombre de cycles PCR faible.

6.1.3. Utilisations des produits PCR.

Les utilisations des produits PCR sont très variées, nous citerons quelques exemples:

- Le diagnostic microbiologique ou moléculaire des maladies génétiques et infectieuses (surtout les cancers et les microorganismes difficilement cultivés Ex. *Bartonella*, virus ...etc.).
- étude des mutations ponctuelles ;
- Analyse de restriction. Le produit PCR est soumis à une digestion enzymatique par une enzyme de restriction. Si une mutation ponctuelle modifie le site de restriction initialement présent, la taille des fragments d'ADN obtenus après digestion sera modifiée et décelable après électrophorèse des fragments d'ADN.
- Les études paléontologiques (fossiles : momies, plantes, animaux, ...etc.)
- Introduction du produit PCR dans un vecteur: clonage du produit PCR.
- Séquençage direct du produit amplifié.

6.2. RT-PCR :

La RT-PCR se déroule en deux phases. Une première phase correspond à la copie d'ARN messager en ADN complémentaire (ADNc) et une seconde phase correspond à une réaction PCR classique sur le ADNc synthétisé.

Dans la première phase, l'ARN messager à étudier est repéré en utilisant une sonde oligonucléotidique spécifique (amorce 1 qui s'hybride à l'extrémité 3' du seul mARN auquel on s'intéresse), puis la transcriptase inverse (ou rétrotranscriptase) permet la synthèse du brin complémentaire (sous une forme de ADNc simple brin), une seconde amorce oligonucléotidique spécifique (amorce 2) permettra la synthèse du second brin par extension. L'ADN complémentaire synthétisé servira ensuite de matrice pour une réaction PCR classique. La technique RT-PCR a permis de montrer que la transcription de tous les gènes s'effectuait dans tous les tissus et ceci même pour les gènes qui présentent une très grande spécificité tissulaire. On parle dans ces conditions de transcription illégitime. Il est évident qu'avant les techniques d'amplification génique, la sensibilité des méthodes classiques n'avait pas permis de mettre en évidence un tel phénomène.

7. Le séquençage de l'ADN

Le séquençage est une technique d'analyse de l'ADN, il permet de déterminer l'ordre d'enchaînement des nucléotides. Actuellement la technique de séquençage repose majoritairement sur la méthode enzymatique de Sanger.

❖ Méthode de Sanger

Le principe de cette méthode consiste à initier la polymérisation de l'ADN à l'aide d'un petit oligonucléotide (amorce) complémentaire à une partie du fragment d'ADN à séquencer. L'élongation de l'amorce est réalisée par une ADN polymérase I dépourvue d'activité exonucléase 5'→3' et maintenue par des ADN polymérases thermostables, celles qui sont utilisées pour la PCR. L'extension de la chaîne est stoppée après l'incorporation des didésoxynucléotides (ddNTP), analogues structuraux des désoxynucléosides triphosphates, mais ne possédant pas de groupement OH en 3' (impossibilité de former la liaison phosphodiester).

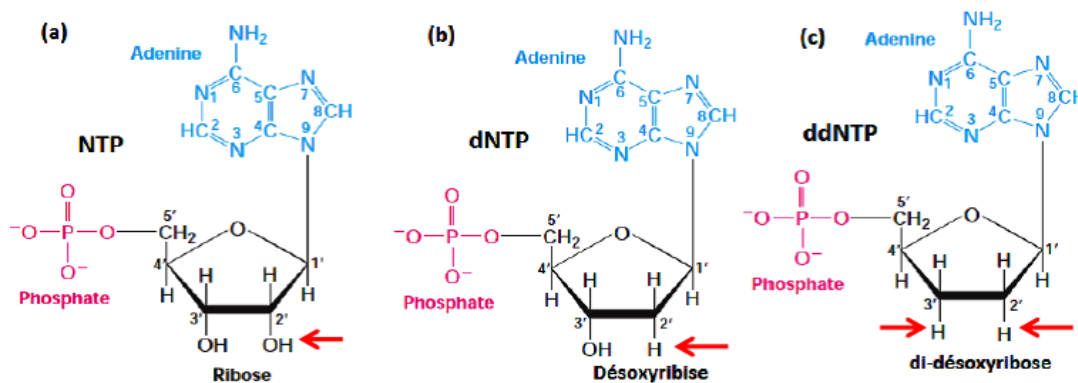


Figure 56: Les différentes formes des nucléotides.

Pour mieux comprendre le principe de séquençage, on utilise des ddNTP non marqués. Alors on utilise quatre tubes séparés contenant l'un des ddNTPs (ddATP, ddCTP, ddGTP, ddTTP). Chaque tube contient: l'ADN à séquencer, ADN polymérase, amorce, dNTPs (dATP, dTTP, dCTP, dGTP); eau purifiée de nucléases (WFN) et MgCl₂.

Pour chaque tube, on obtient un ensemble de molécules dont la longueur dépend de la position du nucléotide complémentaire du didésoxynucléotide introduit. Le contenu de chacun des tubes est déposé sur un gel de polyacrylamide (quatre pistes contiguës) (figure 64).

La résolution du gel est telle que l'on peut distinguer deux brins d'ADN dont la longueur ne diffère que d'un seul nucléotide. Le gel est séché puis soumis à une autoradiographie. On peut donc lire de haut en bas la séquence du brin d'ADN complémentaire et en déduire facilement la séquence du brin matrice.

Dans le séquençage automatique, des marqueurs fluorescents sont utilisés, un marqueur différent pour chaque ddNTP (bleu, rouge, vert, orange...etc.). La polymérisation se fait dans un seul tube contenant le mélange réactionnel et les quatre ddNTPs.

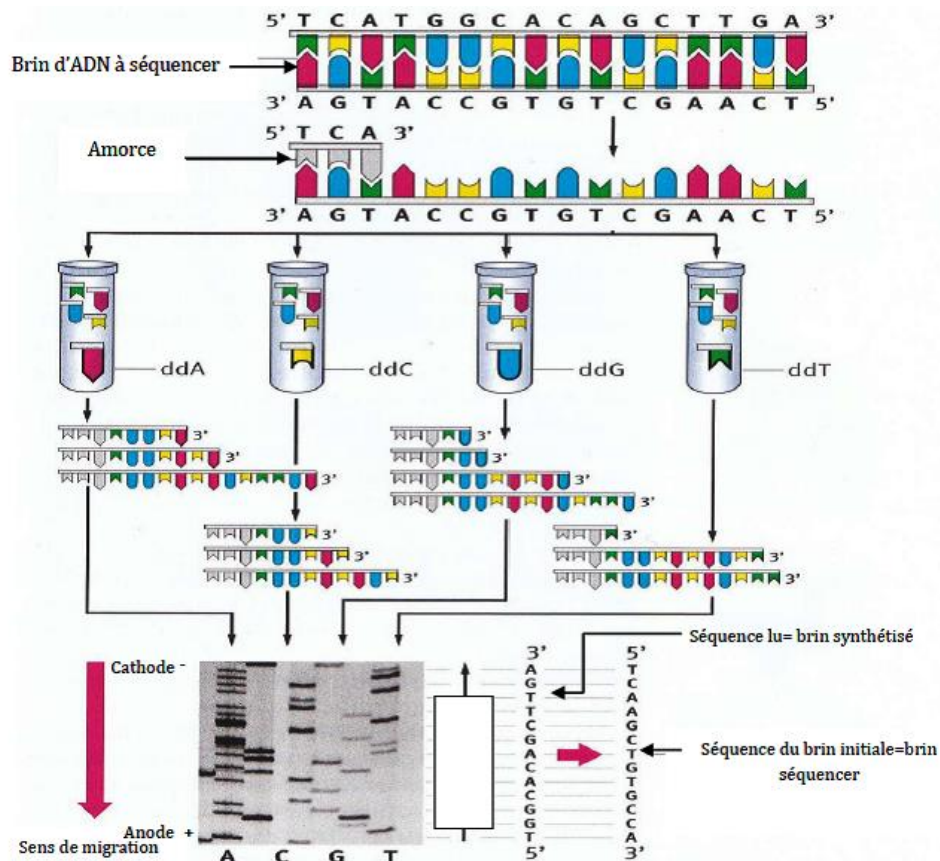


Figure 57: Séquencage par la méthode enzymatique de Sanger

• Automatisation du séquencage

La taille des fragments obtenus est déterminée par une chromatographie. Le séquenceur détecte la fluorescence sortant des colonnes de chromatographie, repérant ainsi les fragments d'ADN et leur taille précise).

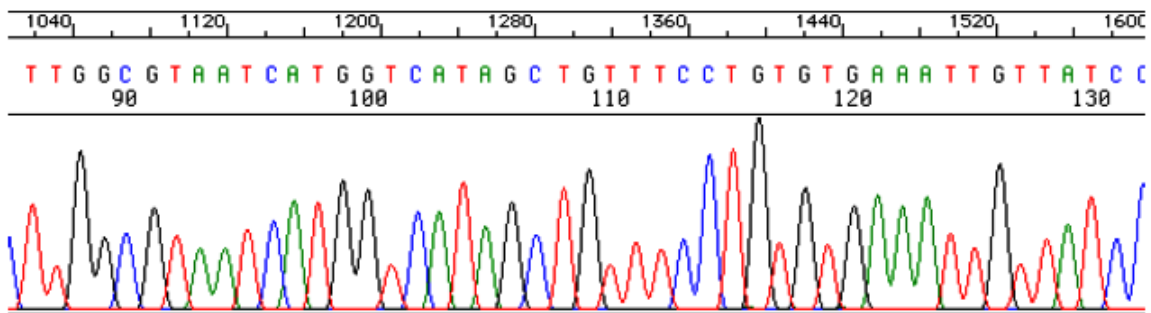


Figure 58: Détection par fluorescence et séparation par électrophorèse ; Séquencage d'ADN méthode de Sanger

8. Interaction protéine-ADN

8.1. Généralités

Les interactions protéine-ADN sont impliquées dans de nombreux processus cellulaires essentiels tels que la réparation de l'ADN, la réplication, la transcription, et la traduction. Les interactions des protéines avec des séquences d'ADN spécifiques dans la chromatine sont nécessaires pour moduler l'expression génétique. Ainsi, les facteurs de transcription sont des protéines qui se lient à l'ADN pour activer ou réprimer l'expression de certains gènes en interagissant avec des séquences nucléotidiques spécifiques. Les protéines histones, au sein des chromosomes, permettent le conditionnement de l'ADN en chromatine, par acétylation ou désacétylation ce qui modifie l'état de condensation de l'ADN plus ou moins propice à la transcription. Certaines enzymes, par exemple l'uracil-ADN glucosylase impliquée dans la réparation de l'ADN, interagissent avec l'ADN de façon très spécifique. Les protéines interagissent au niveau du grand sillon de la double hélice d'ADN (existe des exceptions). Concernant l'interaction proprement dite, ce sont des liaisons hydrogène et des liaisons de van der Waals qui se forment entre l'acide nucléique et la protéine.

Deux approches sont possibles : - Soit l'étude des conséquences de ces interactions.
- Soit l'analyse de l'interaction elle-même et l'identification de ces protéines de liaison à l'ADN

8.2. Mécanisme d'interaction ADN-Protéine

Quatre types d'interactions potentielles peuvent exister entre protéines et acides nucléiques. Du plus fort au plus faible, on peut avoir:

1. Des ponts salins formés entre les phosphates et les chaînes latérales d'acides aminés chargés positivement (N de la Lys, guanidinium de l'Arg, His protonée).
2. Des liaisons hydrogène pouvant s'établir entre les phosphates, sucres et bases des acides nucléiques et les liaisons peptidiques ou les chaînes latérales des acides aminés hydrophiles.
3. Des interactions d'empilement (stacking) impliquant les chaînes latérales des acides aminés aromatiques (Trp, Tyr, Phe, His voire le cycle indol des prolines) et les bases.
4. Des interactions hydrophobes entre les bases des acides nucléiques et les chaînes latérales des acides aminés non polaires.

Si l'on considère les acides aminés et les nucléotides comme des objets isolés, on s'aperçoit que les ponts salins et les liaisons H sont mieux adaptés et probablement dominants. Si on considère maintenant l'ADN et l'ARN sous leur forme double hélice, de nombreux sites de formation de liaison H vont être masqués par l'appariement des bases en Watson-Crick.

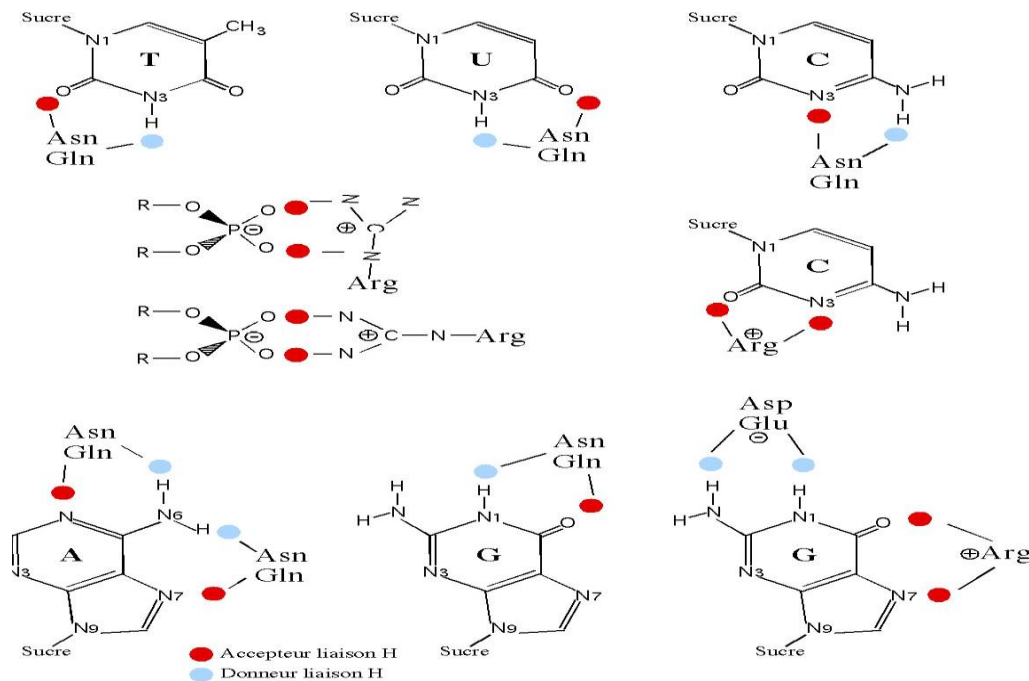


Figure 59: Types d'interactions ADN-Protéine.

8.3. Motifs protéiques de fixation aux acides nucléiques

Les domaines protéiques interagissant avec l'ADN peuvent être classés en fonction de leurs organisations structurales tridimensionnelles comme suite:

- Des motifs à doigts de zinc** : conçus pour se lier à des séquences spécifiques d'ADN, ce qui a servi au développement de nucléases à doigt de zinc.
- Des motifs de fermetures éclair à leucine (*leucine zipper*)** : C'est un motif qui consiste à une répétition périodique du résidu leucine après chaque six acide aminé en formant de 3 à 6 heptades. Ce motif interagit aussi avec le grand sillon de l'ADN par le côté N-terminale.
- Des motifs hélice-boucle-hélice** ; Structure semblable au motif leucine zipper, à l'exception du fait qu'une boucle non hélicoidale de la chaîne polypeptidique sépare deux régions en hélice α dans chaque monomère.
- Des motifs hélice-tour-hélice** : Cette classe est très reconnue pour les facteurs de transcriptions et les enzymes chez les procaryotes et les eucaryotes. La structure de cette

classe des protéines contient deux hélices α qui sont séparées part quarts résidus d'acides aminés qui forment un tour β .

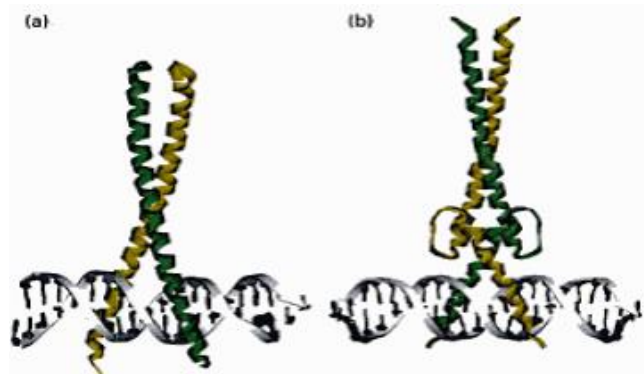


Figure 60: L'interaction des protéines homodimériques (a) le motif leucine zipper ; (b) le motif hélice boucle hélice

8.4. Techniques d'étude des interactions ADN-protéine

De nombreuses techniques ont été développées pour détecter et caractériser les complexes acide nucléique-protéine, on peut utiliser:

8.4.1. Le retard sur gel (EMSA ou *electrophoretic mobility shift assay*)

Sensible largement utilisée pour identifier de telles interactions, EMSA s'appuie sur le décalage électrophorétique d'oligonucléotides sur gel dans les bandes, le complexe ADN-protéine migrera moins vite dans un gel non-dénaturant qu'un ADN nu, qui sont systématiquement visualisé en utilisant la chimioluminescence pour détecter les marqueurs. On peut vérifier la spécificité de l'interaction ADN-protéine par l'incubation préalable du complexe avec un anticorps qui se fixe à la protéine. Le complexe ternaire ADN-protéine-anticorps a une migration encore plus retardée, qu'on qualifie de *supershift*.

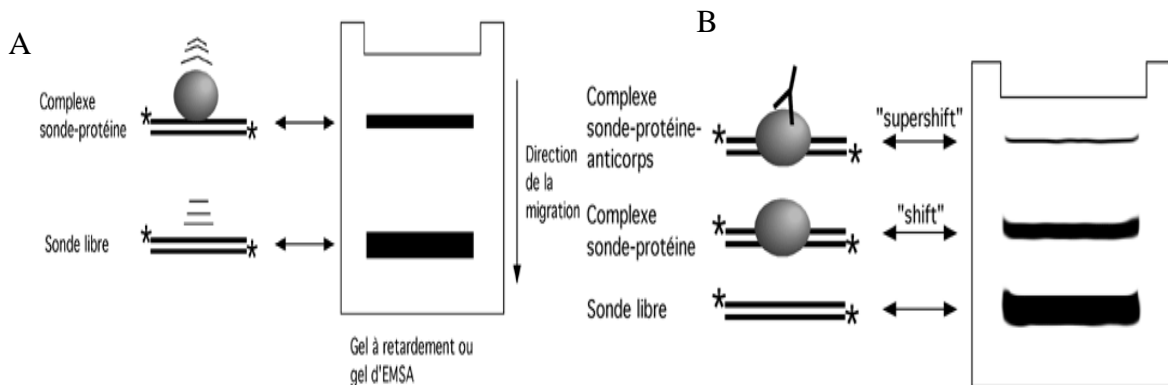
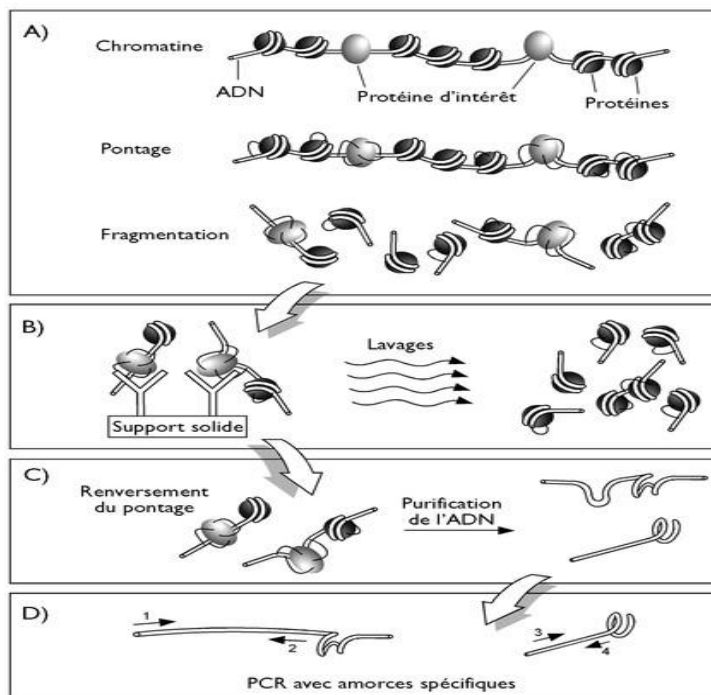


Figure 61 : (A) Technique de décalage de mobilité (Technique EMSA), (B) Principe du shift et supershift

Certaines protéines font se courber l'ADN. Dépendant de l'endroit où elle se lie sur une sonde d'ADN, une protéine qui fait courber l'ADN affectera de façon différente la migration sur le gel d'EMSA. Plus la courbure est au centre de la sonde et plus la migration est retardée. On peut utiliser une même sonde en y déplaçant le site de reconnaissance de la protéine pour juger de l'habileté d'une protéine à courber l'ADN.

8.4.2. Immunoprécipitation de chromatine (ChIP *Chromatine Immuno Precipitation*)

Sert à identifier la séquence de fragments d'ADN qui se lie à un facteur de transcription connu ; lorsqu'elle est combinée avec le séquençage haut débit (HTS ou *high-throughput sequencing*), cette technique est appelée Chip-Seq; lorsqu'elle combinée avec des puces à ADN, on l'appelle ChIP-chip. L'avantage majeur du ChIP est qu'il s'effectue *in vivo*. Ainsi, toutes les protéines qui touchent l'ADN y sont immobilisées par des liens covalents à l'endroit précis de leur interaction grâce à un traitement au formaldéhyde. Par la suite, les cellules sont lysées, puis les complexes ADN-protéines sont fragmentés en courts segments avant d'être immunoprécipités. L'immunoprécipitation, effectuée à l'aide d'un anticorps dirigé contre la protéine d'intérêt (l'activateur de transcription X, par exemple), permet de récupérer les protéines voulues ainsi que toutes les régions d'ADN auxquelles elles étaient liées lors du pontage initial au formaldéhyde;



A) Les protéines sont pontées de façon covalente à l'ADN, et ces complexes sont lysés et l'ADN est scindé en morceaux. B) Une immuno-précipitation avec un anticorps spécifique dirigé contre la protéine d'intérêt (pâle) et des lavages permettent d'isoler les fragments d'ADN qui lui sont liés.

C) Les liens entre l'ADN et les protéines isolées sont rompus, et l'ADN est ensuite purifié.

D) L'ADN immunoprécipité peut être détecté par des amplifications par PCR

Figure 62: Etapes de l'immunoprécipitation de la chromatine pontée.

8.4.3. L'empreinte à la DNase (*DNase footprinting assay*)

Utilisée pour identifier le site de liaison d'une protéine à l'ADN. Cette technique, utilise le fait qu'une protéine déposée sur l'ADN peut le protéger de l'activité de dégradation d'une nucléase. La DNase I ne coupe qu'un seul brin à la fois, aussi on marque un seul des brins au ^{32}P pour suivre son comportement sur gel de séquence. La dégradation partielle du fragment d'ADN protégé par la protéine générera une échelle de fragments de tailles différentes. En comparant l'échelle donnée par un ADN nu à celle donnée par un ADN protégé, on remarque que cette dernière présente un trou, ou une empreinte, correspondant aux sites rendus inaccessibles à l'enzyme par la protéine.

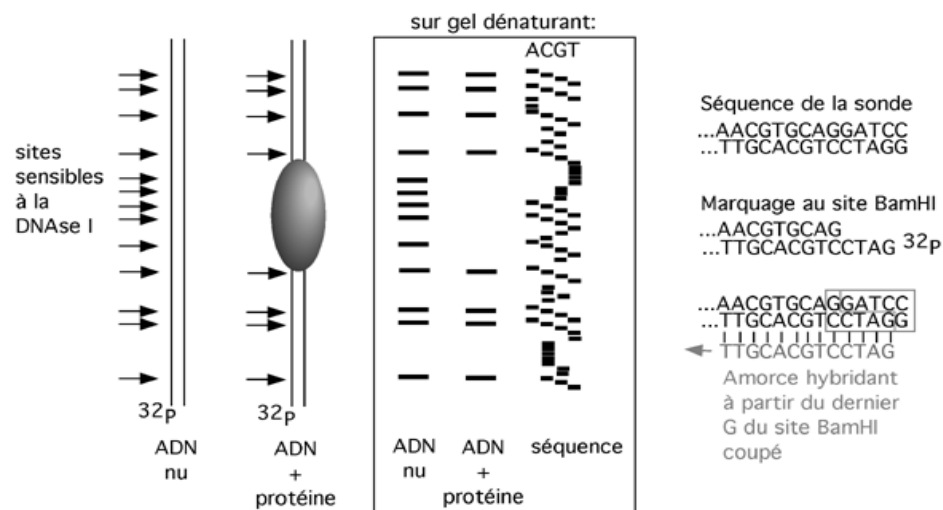


Figure 63: Caractérisation de l'interaction ADN-Protéine par ADNase I

8.4.4. Simple Hybride

Cette technique permet d'identifier de nouvelles protéines interagissant avec un fragment d'ADN connu de vérifier une interaction entre une séquence d'ADN et une protéine d'étudier les nucléotides et/ou acides aminés impliqués dans une interaction connue (figure 64).

Si la protéine X se lie sur la séquence cis régulatrice E, il y aura activation de la transcription du gène rapporteur. Si on fait pousser les levures ayant reçu ce vecteur sur un milieu en présence de X-gal, les colonies obtenues seront bleues. Afin d'améliorer le rendement de l'activation transcriptionnelle, la séquence cible E est répétée plusieurs fois.

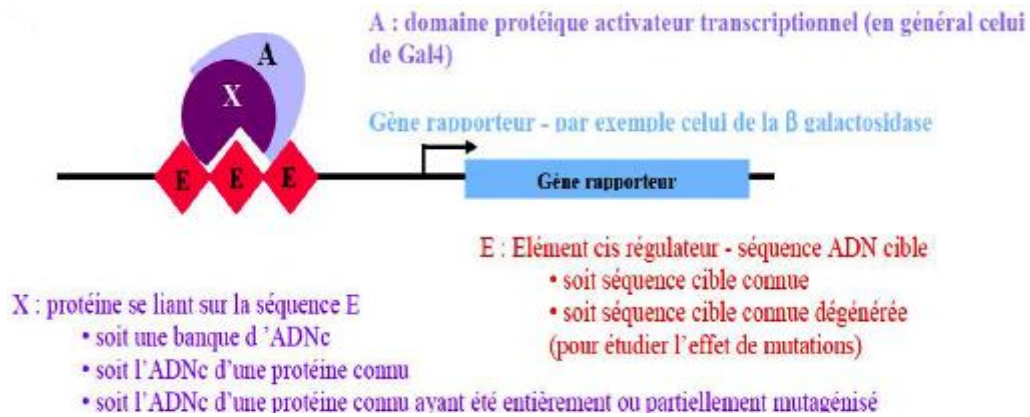


Figure 64: Principe du simple hybride

8.4.5. Système Biacore (SPR ou surface plasmon resonance)

On mesure ici l'interaction entre deux partenaires moléculaires grâce à une machine (Biacore) mesurant l'intensité d'un rayon de lumière. La SPR se produit quand un rayon lumineux est réfléchi sous certaines conditions par un film conducteur à l'interface entre deux milieux d'indices de réfraction différents (figure 65).

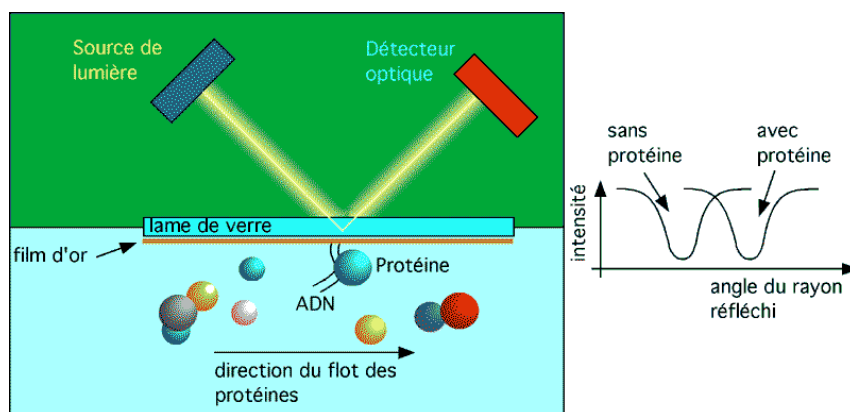


Figure 65: Principe du Biacore.

Le principe de détection par SPR quantifie des changements de l'indice de réfraction près de la surface, reliés à la variation de masse à la surface du biocapteur, due à la formation et à la dissociation des complexes moléculaires.

Dans le système Biacore^(tm) ces milieux sont (1) la solution dans laquelle se trouvent les molécules à analyser et (2) le verre d'une lame qui sert de senseur. Le film conducteur à l'interface des deux est un très fin film d'or à la surface de la lame de verre. La résonance de plasmon cause une réduction de l'intensité de la lumière réfléchie à un angle

spécifique de réflexion. Cet angle varie avec l'indice de réfraction près de la surface du côté *opposé* à la lumière réfléchie (du côté de l'échantillon). Le système détecte ce changement et enregistre une interaction. La mesure du niveau d'interaction en fonction du temps permet d'évaluer le taux d'association. Une unité de résonance (RU) correspond à un changement de 0.0001° dans l'angle donnant une intensité minimale de lumière réfléchie, ce qui pour la plupart des protéines correspond à un changement de concentration de l'ordre de 1 pg/mm^2 de surface sur la lame de verre. Le Biacore est donc extrêmement sensible.

1. Ameziane N, Bogard M, Lamoril J. 2005. Principes de biologie moléculaire en biologie clinique. Paris, Elsevier.
2. Aouf AA. 2015. Cours de Biologie Moléculaire et Génie Génétique, Université Ferhat Abbas-Sétif 1.
3. Brice A, Schorderet DF. 1999. Neurogénétique. Collection Traité de Neurologie. Doin.
4. Clemence A. 2006. Sondes fluorescentes pour l'ADN : marquages covalent et noncovalent. Université Pierre et Marie Curie - Paris VI, 260p.
5. Christian M. 2011. Biochimie et biologie moléculaire .de Boeck 2ème tirage, ISBN : 978-2-8041-6229-0.
6. Clark D. 2005. Molecular Biology: Understanding the Genetic Revolution. Southern Illinois University. USA Elsevier Academic Press. 784p.
7. Gelehrter TD. 1992. Collins FS. Principes de génétique moléculaire et médicale. Editions Pradel, Paris.
8. Goossens M, Mandel JL, Moustacchi E, Papado Poulos D. 1998. Mutations génétiques. In: Principes de génétique humaine Edited by J Feingold, M. Fellous, M. Solignac. Paris: Hermann, Editeurs des sciences et des arts.
9. Gupta PK. 2005. Microbiology, Cell Physiology and Biotechnology. 2nd Ed. National Offset Printers, Meerut, India. 311p.
10. Hogg S. 2005. Essential Microbiology. British Library Cataloguing in Publication Data. John Wiley & Sons. UK.468p.
11. Kaplan Jc, Delpech M. 1996. Biologie Moléculaire Et Médecine. Paris, Flammarion Médecine-Sciences.
12. Lewis R. 1999. Human genetics. Concepts and applications. WCB McGraw-Hill Ed, Boston.
13. Madigan M, Martinko J. 2007. Biologie des microorganismes. 11e Ed. Pearson Education-Paris-France. 1047p.
14. Maftah A, Petit Jm, Julien R. 2018. Mini Manuel de Biologie moléculaire 4ème édition. Dunod.
15. Moussard C. 2005. Biologie moléculaire. Biochimie des communications cellulaires. De Boeck Supérieur.
16. Passarge E. 2007. Color Atlas of Genetics. 3rdEd. Flexibook. Thieme Stuttgart, NY. USA. 486p.

17. Perrin-Schmitt F. 2011. Biotechnologies et applications (génie génétique), PACES. UE2. Ellipses 58-65p.
18. Petit Jm, Arico S. 2011. Julien R. Mini Manuel de Génétique 2ème édition. Dunod.
19. Pinto M. 1997. Outils et stratégie de biologie moléculaire dans le diagnostic génotypique prénatal. Tec & Doc Lavoisier.
20. Strachan T, Read A. 2004. Génétique Moléculaire Humaine. Hybridation des acides nucléiques: principes et applications. 4^e édition ; Médecine Sciences publications. Lavoisier 16p.
21. Swynghedauw B, Silvestre JS. 2008. Biologie et génétique moléculaires, Aide-mémoire. 3ème édition Dunod.
22. Tagu D, Moussard C. 2003. Principes des techniques de biologie moléculaire, 2e ed, INRA-Qua (ISBN 2-7380-1067-9).
23. Terwilliger JD, Ott, J. 1994. Handbook of Human Genetic Linkage. Baltimore, John Hopkins University Press.
24. Thompson M, McInnes R, Willard H. 1995. Génétique Médicale. Médecine-Sciences Flammarion Ed, Paris.
25. Watson J, Baker T, Bell S, Gann A, Levine M, Losick R. 2009. Biologie Moléculaire du gène. 6eEd. Pearson Education- Paris-France. 688p.
26. Yu L, He W, Xie J, Guo R, Ni J, Zhang X, Xu Q, Wang C, Yue Q, Li F, Luo M., Sun B, Ye L, Zheng K. 2019. In Vitro Biochemical Assays using Biotin Labels to Study Protein-Nucleic Acid Interactions. J Vis Exp. (149), e59830.