

République Algérienne démocratique et populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université A.MIRA-BEJAIA  
Faculté des Sciences de la Nature et de la Vie  
Département des Troncs Communs



# Biologie moléculaire

*(Notions de bases en biologie moléculaire)*

**Dr. BOUREBABA Yasmina**

Cours destiné aux étudiants en Licence et Master SNV

## *Préambule*

Le grand développement de la biologie moléculaire au XXe siècle a marqué le développement de la biologie moderne. Si les progrès de la génétique sont à la fois fascinants et effrayants, on oubliera que l'histoire de la discipline est avant tout une épopée humaine mythique.

Il est important de comprendre l'origine, les fondements ainsi que le développement de la biologie moléculaire afin de comprendre ce qu'elle représente à l'heure actuelle. De manière convaincante, avec le développement rapide de la science, l'essor de cette technologie a commencé il y'a plus d'un siècle. Grâce à de subtiles expériences sur les pois, elle a été exposée dans le jardin discret d'un monastère, puis décrite. Elle a pris énormément d'ampleur suite au séquençage du génome humain ainsi que celui de nombreuses espèces vivantes par de grands laboratoires au XXe siècle, ce qui a ouvert des horizons plein d'espoir quant au future de l'humanité.

Ce travail fournit aux lecteurs un support pédagogique destiné à les initier à la biologie moléculaire et notamment aux notions élémentaires. La pédagogie employée lors de la conception de ce cours repose sur l'utilisation d'un langage simple, accompagné d'exemples et de nombreuses illustrations afin de rendre les notions abordés plus accessible. Ce cours intitulé "*Notions élémentaires en biologie moléculaire*" est scindé en trois principaux chapitres abordant essentiellement une description de ce qu'est la biologie moléculaire, un récit historique concis des découvertes sur le terrain et de l'évolution de cette discipline, des caractéristiques des acides nucléiques, et des mécanismes moléculaires à savoir la réplication de la molécule d'ADN, sa transcription en ARN, la traduction des ARN messagers en protéines ainsi que la régulation de l'expression génique.

**« *Tout va trop vite* »**

*Simone Gilgenkrantz*

## Préambule

# *Table des Matières*

Liste des figures

Liste des tableaux

Objectifs généraux du cours

## *Chapitre I. Introduction à la biologie moléculaire*

I. Définition de la biologie moléculaire.....	01
II. Historique relatif à la biologie moléculaire et au génie génétique .....	01
III. Quelques définitions .....	07

## *Chapitre II. Caractéristiques des acides nucléiques*

I. Définition .....	09
I.1. Acides nucléiques .....	09
I.2. Nucléotides .....	09
II. Liaisons des acides nucléiques .....	11
III. Structure de la molécule d'ADN .....	13
III.1. Lois de Chargaff.....	13
III.2. Formes de l'ADN .....	14
III.3. Importance de l'ADN dans la cellule.....	16
III.4. Principaux types d'ADN .....	16
III.4.1. Matériel génétique des virus .....	16
III.4.2. ADN des procaryotes (cas des bactéries).....	16
III.4.3. ADN des eucaryotes.....	17
IV. Structure de la molécule d'ARN .....	17
IV.1. Types et rôles des ARN.....	18
IV.1.1. ARN messagers (ARNm).....	18
IV.1.2. ARN de transfert (ARNt).....	18
IV.1.3. ARN ribosomiques (ARNr) .....	19
IV.1.4. ARN small nuclear (ARNsn) et ARN small nucleolar (ARNsno).....	19
IV.1.5. ARN micro (ARNmi) et Small interfering RNA (ARNsi) .....	19
IV.1.6. ARN long noncoding (LncRNA).....	19
V. propriétés physico-chimiques des acides nucléiques .....	20

## *Chapitre III. Mécanismes moléculaires*

I. Réplication de la molécule d'ADN .....	22
I.1. Introduction .....	22
I.1.1. Définition .....	22
I.1.2. Cycle cellulaire .....	22
I.2. Lois fondamentales de la réplication .....	23
I.3. Mécanismes de la réplication .....	25
I.3.1. Réplication chez les procaryotes ( <i>E. coli</i> ).....	27
I.3.2. Réplication chez les eucaryotes .....	29
I.3.3. Réplication de l'ADN mitochondrial.....	32
I.3.4. Réplication du matériel génétique des rétrovirus.....	32
II. Mutation/Altération de la molécule d'ADN et mécanismes de réparation de l'ADN .....	33
II.1. Introduction .....	33
II.2. Mutations et dommages sur l'ADN.....	33
II.2.1. Comment apparaissent ces mutations? .....	33
II.2.1.1. Erreurs commises lors de la réplication .....	33
II.2.1.2. Altération de l'ADN survenant en dehors de la réplication .....	35
II.2.1.2.1. Lésions spontanées .....	35
II.2.1.2.2. Mutation induites par des agents mutagènes .....	37
II.3. Réparation des lésions de l'ADN .....	39
II.3.1. Mécanismes de réparation .....	40
II.3.1.1. Correction des erreurs d'appariement lors de la réplication .....	40
II.3.1.1.1. Correction immédiate : fonction d'édition de l'ADN polymérase .....	40
II.3.1.1.2. Réparation des mésappariements par le système methyl mismatch repair .....	40
II.3.1.2. Correction des autres mutations de l'ADN .....	42
II.3.1.2.1. Réparation direct par retour a l'état antérieur .....	42
II.3.1.2.2. Réparation par excision .....	43
II.3.1.2.3. Réparation par recombinaison homologue .....	46
II.3.1.2.4. Système SOS .....	47
III. Recombinaison et transposition de l'ADN.....	48
III.1. Recombinaison génétique .....	48
III.1.2. Rôle de la recombinaison .....	48

III.1.3. Mécanismes de recombinaison .....	49
III.1.3.1. Recombinaisons intra-chromosomiques .....	49
III.1.3.2. Recombinaisons inter-chromosomiques .....	49
III.1.3.3. Recombinaison homologue .....	50
III.1.3.4. Recombinaison site-spécifique.....	54
III.1.3.5. Recombinaison illégitime.....	56
III.2. Transposition de l'ADN .....	59
III.2.1. Transposons.....	59
III.2.2. Intérêts des transposons.....	59
III.2.3. Classification des éléments transposables.....	59
III.2.4. Types de transposition.....	61
III.2.4.1. Transposition répllicative .....	61
III.2.4.2. Transposition conservative.....	62
III.2.5. Transposase et son mode d'action.....	63
IV. Transcription de l'ADN .....	65
IV.1. Définition .....	65
IV.2. Eléments nécessaires à la transcription.....	65
IV.3. Mécanismes de transcription.....	67
IV.3.1. Transcription chez les procaryotes.....	67
IV.3.2. Transcription de l'ADN chez les eucaryotes .....	71
V. Traduction des ARN messagers .....	77
V.1. Définition.....	77
V.2. Rôle de la traduction.....	77
V.3. Acteurs de la traduction.....	77
V.4. Mécanisme de la traduction.....	81
V.4.1. Traduction chez les procaryotes .....	81
V.4.2. Traduction chez les eucaryotes.....	83
V.5. Code génétique .....	86
VI. Régulation de l'expression des gènes .....	89
VI.1. Définition .....	89
VI.2. Mécanismes de régulation.....	89
VI.2.1. Chez les procaryotes .....	89

VI.2.1.1. Les opérons .....	90
VI.2.1.2. Opéron lactose.....	90
VI.2.1.2.1. Organisation de l'opéron lactose.....	91
VI.2.1.2.2. Fonctionnement de l'opéron lactose .....	91
VI.2.2. Chez les eucaryotes.....	93
VI.2.2.1. Au niveau chromatinien .....	93
VI.2.2.2. Au niveau transcriptionnel.....	94
VI.2.2.3. Au niveau post transcriptionnel .....	94
VI.2.2.4. Au niveau traductionnel .....	95
VI.2.2.5. Au niveau post traductionnel .....	95
<b>Références</b> .....	<b>96</b>

# Liste des figures

<i>N°</i>	<i>Titre</i>	<i>Pages</i>
<b>1</b>	Structures chimiques des bases puriques et des bases pyrimidiques des acides nucléiques.....	09
<b>2</b>	Pentoses respectifs de l'ARN et de l'ADN.....	10
<b>3</b>	Molécule d'acide phosphorique.....	11
<b>4</b>	Structure détaillée d'un désoxyribonucléoside triphosphate.....	11
<b>5</b>	Position <i>anti</i> et <i>syn</i> des bases azotées reliées au sucre.....	12
<b>6</b>	Principe de formation d'un brin d'ADN par la polymérisation des nucléotides.....	12
<b>7</b>	Structure plane et en 3D des deux chaînes antiparallèles de l'ADN.....	13
<b>8</b>	Liaisons hydrogènes des différentes bases azotés.....	13
<b>9</b>	Différentes formes de l'ADN.....	15
<b>10</b>	Brin d'ARN.....	18
<b>11</b>	Représentation de l'ARNt.....	18
<b>12</b>	Représentation de la dénaturation de l'ADN sous l'effet de la température.....	21
<b>13</b>	Renaturation des brins d'ADN après dénaturation.....	21
<b>14</b>	Interaction du cis-platine avec la molécule d'ADN.....	21
<b>15</b>	Etapes du cycle cellulaire.....	23
<b>16</b>	Différents modèles de répllication de l'ADN.....	24

<b>17</b> : Représentation de l'expérience de Meselson et Stahl.....	25
<b>18</b> : Fourche et sens de réplication.....	26
<b>19</b> : Processus d'action de la protéine DnaA lors de l'initiation de la réplication.....	27
<b>20</b> : Processus de terminaison de la réplication procaryote .....	29
<b>21</b> : Mode d'action de la télomérase .....	31
<b>22</b> : Réplication de l'ADN mitochondrial.....	32
<b>23</b> : Nomenclature des substitutions de base .....	34
<b>24</b> : Mutation de l'ADN par délétion de base .....	34
<b>25</b> : Mutation par insertion de bases .....	35
<b>26</b> : Phénomène de dépurination de l'ADN et création d'un site apurique (AP).....	35
<b>27</b> : Désamination de différentes bases azotées.....	36
<b>28</b> : Oxydation de la guanine en oxoguanine.....	36
<b>29</b> : Analogues de bases azotées .....	37
<b>30</b> : Modification de la guanine par le 6-ethylguanine .....	37
<b>31</b> : Agents intercalants.....	38
<b>32</b> : Dimère de thymine.....	38
<b>33</b> : Lésions d'ADN par radiations ionisantes .....	39
<b>34</b> : Action de l'aflatoxine B1 sur la molécule d'ADN .....	39
<b>35</b> : Formation de pontage d'un brin d'ADN.....	39
<b>36</b> : Fonction d'édition et de correction immédiate de l'ADN polymérase.....	40
<b>37</b> : Mismatch repair chez <i>E. coli</i> .....	41

<b>38</b> : Réparation des mésappariements .....	42
<b>39</b> : Réparation par photolyase d'un dimère de thymine .....	43
<b>40</b> : Réparation par excision de base .....	44
<b>41</b> : Réparation par excision de nucléotides.....	45
<b>42</b> : Mécanisme NER chez l'homme .....	46
<b>43</b> : Recombinaison de l'ADN chez les procaryotes .....	47
<b>44</b> : Recombinaison in vivo entre le génome et un fragment d'ADN par appariement de régions homologues .....	48
<b>45</b> : Recombinaison Intra-chromosomique .....	49
<b>46</b> : Recombinaison inter-chromosomique .....	50
<b>47</b> : Représentation du modèle de Holliday .....	51
<b>48</b> : Modèles moléculaires de recombinaison permettant d'expliquer la conversion génique chez les eucaryotes. Présentation succincte des modèles de DSBR, et SDSA .....	53
<b>49</b> : Double "break-induced replication" .....	54
<b>50</b> : Réaction de recombinaison site spécifique entre le phage $\lambda$ et le chromosome bactérien.....	56
<b>51</b> : Exemples de jonctions de recombinaison homologue et non homologue .....	57
<b>52</b> : Différents rétro-transposons.....	60
<b>53</b> : Quelques transposons de type II .....	61
<b>54</b> : Processus de transposition répllicative .....	62
<b>55</b> : Transposition conservative .....	63
<b>56</b> : Représentation schématique des différents processus biologique de la cellule.....	65

<b>57</b> : Promoteurs des gènes procaryotes .....	68
<b>58</b> : Mécanisme direct (Rho indépendant) .....	70
<b>59</b> : Mécanisme indirect avec une protéine Rho à activité ATPasique.....	70
<b>60</b> : Transcription de l'ADN chez les eucaryotes .....	71
<b>61</b> : Phase d'élongation de la transcription .....	73
<b>62</b> : Ajout de la coiffe en 5' .....	74
<b>63</b> : Ajout de la queue poly(A) en 3' .....	75
<b>64</b> : Excision des introns (épissage).....	76
<b>65</b> : Représentation d'un ARN messager eucaryote .....	77
<b>66</b> : Structure d'un ribosome.....	78
<b>67</b> : ARNt et sa liaison avec l'acide aminé ainsi que l'ARNm.....	79
<b>68</b> : Structure chimique des acides aminés .....	80
<b>69</b> : Processus de traduction de l'ARNm en protéine .....	81
<b>70</b> : Etapes de la traduction chez les eucaryotes .....	84
<b>71</b> : Initiation de la traduction des eucaryotes.....	86
<b>72</b> : Représentation schématique d'un opéron .....	90
<b>73</b> : Représentation structurale de l'opéron lactose .....	91
<b>74</b> : Mécanisme de l'opéron lactose.....	93
<b>75</b> : Représentation d'un micro-ARN .....	95

---

# Liste des tableaux

<i>N</i>	<i>Titre</i>	<i>Pages</i>
<b>I</b>	Représentation symboliques des différentes bases azotées .....	10
<b>II</b>	Paramètres structurels indicatifs des trois principales formes d'ADN .....	15
<b>III</b>	Propriétés en taille, masse et longueur de l'ADN des différents types cellulaires.....	20
<b>IV</b>	Différentes ARN polymérase et leur fonction chez les eucaryotes .....	67
<b>V</b>	Représentation des wobble.....	79
<b>VI</b>	Correspondance entre codons et acides aminés (base du code génétique).....	87

## *Objectifs généraux du cours*

Par le biais de ce cours, il est donc voulu donner aux apprenants des bases solides sur l'historique, et la méthodologie moléculaire. Ce cours a été conçu dans l'optique que des étudiants venant d'horizons différents, ayant ou n'ayant pas eu au préalable des notions de biologie moléculaire, puissent le suivre facilement. Néanmoins, ce cours ne pourra être bien compris que si certaines notions scientifiques de bases ont été acquises.

### *Pré-requis*

- Notions de Biologie Générale ;
- Notions de Chimie/Biochimie élémentaire ;
- Notions en génétique ;
- Notions en biologie cellulaire.

*A l'issue du cours, les apprenants devraient en théorie pouvoir répondre à la liste des objectifs suivants*

- Décrire la composition et la structure de l'ADN et des différents ARN ;
- Donner la structure générale des nucléotides, des bases puriques et pyrimidiques ;
- Connaître la réplication (réplicon, origine de réplication, ADN polymérase) ;
- Connaître les dommages que peut subir la molécule d'ADN ;
- Connaître et comprendre les mécanismes de réparation ainsi que les lésions prise en charge ;
- Savoir ce qu'est la recombinaison génétique ainsi que la transposition ;
- Connaître la transcription (promoteur, transcrit, ARN polymérase) ;
- Connaître la maturation des transcrits (intron, exon) ;
- Connaître la traduction (ARNt, ARNr, ARNm, ribosome) ;
- Connaître le code génétique (codon) ;
- Connaître ce qu'est la régulation de l'expression génique et quand intervient-elle.

# *Chapitre I*

---

*Introduction à la biologie moléculaire*

## Introduction à la biologie moléculaire

### Objectifs spécifiques

Au terme de ce cours qui est une introduction à la biologie moléculaire, vous devez être capable de :

- Connaitre et pouvoir définir ce qu'est la biologie moléculaire ;
- Connaitre ce qu'est le génie génétique ;
- Distinguer les différentes et les relations entre les deux disciplines ;
- Comprendre la terminologie relié à ces deux disciplines ;
- Avoir une culture générale sur l'histoire de l'évolution de ces disciplines.

### I. Introduction

On entend tous ce terme 'biologie moléculaire' qui est très en vogue en ce moment, mais que veut-il dire ? C'est quoi la biologie moléculaire ?

C'est une discipline consacrée à l'étude des molécules porteuses du message héréditaire en particulier les acides nucléiques (ADN, ARN) et la compréhension des mécanismes de fonctionnement au niveau moléculaire (de leur structure, synthèse, altérations (mutations), expression des gènes,...).

C'est une discipline récente au croisement de la génétique, de la biochimie et de la physique qui c'est considérablement développée au cours de la deuxième partie du 20<sup>ème</sup> siècle et qui continue sans cesse d'évoluer.

#### Remarque

Il faut savoir que la biologie moléculaire utilise tout un ensemble de techniques et de manipulations des acides nucléiques que l'on appelle « techniques de génie génétique ».

#### - Génie génétique

C'est l'ensemble des outils et des techniques de la biologie moléculaire permettant, de manière contrôlée, l'étude de la modification des gènes : leur isolement, leur clonage, leur séquençage, leur découpage, ... dans un but de recherche fondamentale ou appliquée.

La biologie moléculaire et le génie génétique sont à la base des biotechnologies, en particulier, celles basées sur l'ADN recombiné et les transformations génétiques.

### II. Historique relatif à la biologie moléculaire et au génie génétique

La biologie moléculaire, la génétique et le génie génétique sont des disciplines en plein développement, presque chaque année apparaissent de nouvelles techniques, certaines représentent des avancées si considérables qu'elles rendent caduques les techniques encore routinières l'année d'avant.

Le terme « biologie moléculaire », a été utilisé pour la première fois en 1938 par **Warren Weaver**, mais suite à quoi ?

Comment sommes-nous arrivé à ces disciplines qui se complètent l'une l'autre ? On va retracer certaines grandes lignes de ces disciplines ensemble :



**Warren Weaver**

Il faut tout d'abord savoir que c'est au cours de la deuxième moitié du 19<sup>ème</sup> siècle que ces disciplines ont vue une évolution significative.

**1869 :**

- **Johann Friedrich Miescher** (1844-1895), médecin et biochimiste suisse isola à partir des noyaux de cellules du pus une substance non protéique et non lipidique riche en phosphore jusque-là inconnue qu'il nomma 'nucléine'.



**Johann Friedrich Miescher**

**1872 :**

- **Johann Friedrich Miescher** (1844-1895), démontre la présence de la nucléine dans les spermatozoïdes de plusieurs espèces.
- Il émet l'hypothèse que cette substance joue un rôle dans la transmission de l'hérédité.

**1879 :**

- **Walther Flemming** (1843-1905), biologiste allemand et l'un des fondateurs de la biologie cellulaire.



**Walther Flemming**

- Il introduit la notion de « chromatine » (chroma = couleur) pour désigner la substance nucléaire qui se colore avec un colorant à base d'aniline (comme le carmin acétique).
- Il décrit pour la première fois la mitose du grec *mitos* = filament.

**1888 :**

- **Heinrich Wilhelm Gottfried Waldeyer Hartz** (1836-1921), anatomiste allemand.
- Il créa le terme de « chromosome » pour désigner les filaments colorés lors de la division cellulaire.



**Heinrich Wilhelm Gottfried Waldeyer Hartz**

**1891 :**



**Theodor Boveri**

- **Theodor Boveri** (1862-1915), biologiste allemand, démontre et affirme que les chromosomes sont indispensables à la vie.

**1903 :**

- **Walter Stanborough Sutton** (1877-1916), médecin et généticien américain.
- Ces travaux portaient sur l'étude de la méiose chez la sauterelle. Il observa pour la première fois la méiose en 1902.
- Il propose que les chromosomes soient le support de l'information génétique.



**Walter Stanborough Sutton**

**1909 :**

- **Wilhelm Ludvig Johannsen** (1857-1927), botaniste, physiologiste et généticien danois, crée le terme « gène ».
- En **1911**, il fait la différence entre l'aspect d'un être (phénotype) et son gène (génotype).
- Il définit aussi le terme « population » en biologie.



**Wilhelm Ludvig  
Johannsen**

**1910 :**

- **Thomas Hunt Morgan** (1866-1945), embryologiste et généticien Américain, cherchait à comprendre le développement de l'embryon des vertébrés, jugea nécessaire de se mettre à l'étude de l'hérédité et l'évolution.
  - Il étudia les variations phénotypiques chez la mouche du vinaigre '*Drosophila melanogaster*'.
  - Il démontre l'existence de mutations en conduisant des expériences sur la drosophile.
  - Il montre que les chromosomes sont les supports des gènes, grâce à la découverte des liaisons génétiques (Deux allèles de deux gènes différents ont tendance à être transmis ensemble d'un individu à sa descendance) et des recombinaisons génétiques (brassage génétique).



**Thomas Hunt Morgan**

**1913 :**

- **Morgan**, avec l'aide d'**Alfred Sturtevant**, publie la première carte génétique du chromosome X de la drosophile, montrant l'ordre et la succession des gènes le long du chromosome.



**Alfred Sturtevant**

*Et c'est là que Tout commence réellement*

**1924 :**

- **Robert Feulgen** et **Heinrich Rössenbeck** avaient montré que les chromosomes contiennent de l'ADN.

**1932 :**

- **Fred Griffith**, microbiologiste anglais, en recherchant un vaccin contre la pneumonie, démontra que des pneumocoques tués par la chaleur pouvaient transmettre certains de leurs caractères, notamment leur virulence, à des souches de pneumocoques vivantes non virulentes. Résultats qu'il ne publia que 4 ans plus tard.
- Seul le transfert d'une substance chimique entre des bactéries mortes et des bactéries vivantes pouvait expliquer cette transformation.



**Fred Griffith**

**1941 :**

- **George Beadle** et **Edward Tatum** émettent l'hypothèse qu'un gène code une (et uniquement une) enzyme en étudiant *Neurospora crassa*.



**George Beadle**



**Edward Tatum**

1943 :

- La diffraction au rayon X de l'ADN par **William Astbury** permet d'émettre la première hypothèse concernant la structure de la molécule : une structure régulière et périodique qu'il décrit comme une pile de pièces de monnaie.



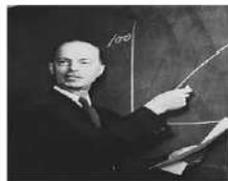
*William Astbury*

1944 :

- L'Américain **Oswald Avery**, **Colin McLeod** et **Maclyn McCarty** découvrent que l'ADN est responsable de la transformation génétique des bactéries et que ce serait bien le support de l'hérédité. Alors que tous les biologistes étaient convaincus que la transmission des caractères héréditaires d'une génération à une autre dépendait des protéines.



*Oswald Avery*



*Colin McLeod*



*Maclyn McCarty*

- **Barbara McClintock** montre que les gènes peuvent se déplacer et que le génome est beaucoup moins statique que prévu. Elle reçoit le prix Nobel de Médecine en 1983.



*Barbara McClintock*

1950 :

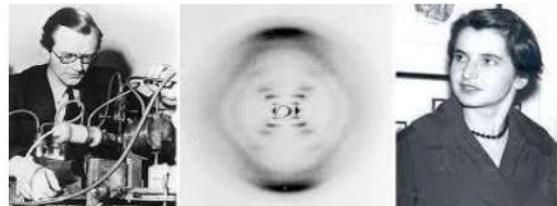


*Alexander Robert Todd Phoebus Levene*

- **Phoebus Levene** et **Alexander Robert Todd** définissent la structure chimique de l'ADN.

1953 :

- **Maurice Wilkins** et **Rosalind Franklin** réalisèrent un cliché d'une molécule d'ADN.



- **James Watson** et **Francis Crick** présentent le modèle en double hélice de l'ADN, expliquant ainsi que l'information génétique puisse être portée par cette molécule. Watson, Crick et Wilkins recevront en 1962 le prix Nobel de médecine pour cette découverte.

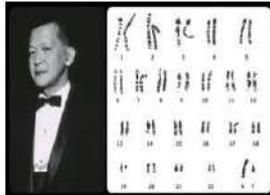
1954 :

- **Georges Gamow** découvre que l'ADN obéit à un code génétique.



**Georges Gamow**

1955 :



- **Joe Hin Tjio** fait le premier compte exact des chromosomes humains : 46.

- **Arthur Kornberg** découvre l'ADN polymérase, une enzyme permettant la réplication de l'ADN.



**Arthur Kornberg**

1957 :

- Le mécanisme de réplication de l'ADN est mis en évidence.

1962 :



**Carl Woese**

- **Carl Woese** démontre que l'ARN des ribosomes de certaines bactéries est très différent des autres organismes.

- **Crick, Watson et Wilkins** reçoivent le prix Nobel de médecine pour avoir établi que les triplets de bases étaient des codes. Le comité Nobel évoquera « la plus grande réussite scientifique de notre siècle ».



**Francis Harry Compton Crick**  
(1916-2004)



**James Dewey Watson**  
(1928 - )



**Maurice Hugh Frederick Wilkins**  
(1916-2004)

1966 :

- **J.L. Hubby et Richard C. Lewontin** ouvrent la voie au domaine de la recherche sur l'évolution moléculaire en introduisant les techniques de la biologie moléculaire comme l'électrophorèse sur gel dans la recherche sur la génétique des populations.

1968 :

- **Marshall W. Nirenberg et Har Gobind Khorana** se voient attribuer un prix Nobel pour avoir déchiffré le code génétique des 20 acides aminés.
- Par la suite, les chercheurs ont pu conclure que le code



**Nirenberg**



**Khorana**

génétique est universel chez toutes les créatures vivantes.

**1972 :**

- **Paul Berg** réussit la recombinaison des brins d'ADN découpés par des enzymes de restriction ce qui ouvre la voie au début de manipulations génétiques.



**Paul Berg**

**1973 :**

- Une première transgénèse est réalisée, lorsqu'un gène d'un amphibien africain est inséré dans l'ADN d'une bactérie.

**1978 :**

- Un gène humain codant pour l'insuline est introduit dans la bactérie *Escherichia coli*, afin que cette dernière produise de l'insuline humaine.

**1983 :**

- Le Canada autorise la production commerciale d'insuline à partir d'*E. coli* GM. Aujourd'hui, cette insuline est utilisée dans le traitement du diabète.

**1986 :**

- Des chercheurs ont réalisé le premier essai en champ de plante transgénique (Un tabac résistant à un antibiotique).

**1989 :**

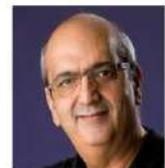
- Il a été décidé de décoder les 03 milliards de paires de bases du génome humain pour identifier les gènes afin de comprendre, dépister et prévenir les maladies génétiques et tenter de les soigner.
- Une première équipe se lance dans la course : le Human Genome Project, coordonné par le NIH (National Institutes of Health) et composé de 18 pays dont la France avec le Génoscope d'Evry qui sera chargée de séquencer le chromosome 14.

**1993-1996 :**

- Les premières cartes génétiques du génome humain sont publiées par **Jean Weissenbach** et **Daniel Cohen** dans un laboratoire du Généthon.



Jean Weissenbach



Daniel Cohen

**1998 :**

- Créée par **Craig Venter** et **Perkin Elmer** (leader dans le domaine des séquenceurs automatiques), la société privée Celera Genomics commence elle aussi le séquençage du génome humain en utilisant une autre technique que celle utilisée par le NIH.



**1999 :**

- Un premier chromosome humain, le 22, est séquencé par une équipe coordonnée par le centre Sanger, au Royaume-Uni.

**2000 :**

- Nombreux sont les organismes vivants dont l'ADN a été décrypté au cours des années 2000.
  - La première plante, la petite arabette des dames (*Arabidopsis thaliana*) dévoile son génome en 2000.
  - Suivie par la souris de laboratoire (*Mus musculus*) en 2002.

**2003 :**

- La course effrénée pour le séquençage du génome humain, sur plus de dix ans, prend fin en 2003, soit cinquante ans après la découverte de la structure de l'hélice d'ADN par **James Watson** et **Francis Crick**.

Le séquençage de l'ADN humain fut porteur d'espoirs pour la santé avec la médecine personnalisée, le dépistage de maladies et la thérapie génique.

**III. Quelques définitions**

- **Cellule**

Aussi diverse soit-elle, l'unité fondamentale et la plus élémentaire du vivant, c'est la cellule, qui est délimitée par une membrane et contient tous les composants permettant à la cellule d'être vivante. Les cellules sont toujours filles d'autres cellules et ne proviennent jamais de l'assemblage de leurs constituants.

Tout être vivant (donc tout organisme) est soit une cellule isolée, soit une association de plusieurs cellules. Il y a deux types de cellules vivantes : les procaryotes et les eucaryotes.

- **Procaryotes**

Les procaryotes (du grec *pro*, avant et *karyon*, noyau) sont des êtres unicellulaires, dépourvus de noyau et bordés d'une membrane. Elles ne possèdent donc pas de noyau, de ce fait le matériel génétique se retrouve au contact du cytoplasme sans oublier l'absence d'autres organites telles que les mitochondries, l'appareil de golgi,...

Il y a deux types de procaryotes, les eubactéries qui sont les bactéries les plus connues et les plus ordinaires, et les archéobactéries qui ont une structure membranaire très différente de celle des eubactéries leur permettant de pousser dans des milieux extrêmes (méthane, hautes températures,...).

- **Eucaryotes**

A la différence des procaryotes, les eucaryotes possèdent un vrai noyau délimité par une membrane et qui contient la majeure partie du matériel génétique.

Les eucaryotes peuvent être uni- ou multicellulaires. La levure est un eucaryote monocellulaire qui possède un noyau et 16 chromosomes, c'est un modèle très courant en biologie moléculaire. Il existe un certain nombre de modèles pluricellulaires utilisés dans la plupart des laboratoires pour chacun des règnes animaux ou végétaux, la *Drosophila* pour les

insectes, *Danio rerio* (le poisson zèbre) pour les poissons, *Arabidopsis thaliana* (le cresson) pour les plantes, la souris pour les mammifères.

- **Virus**

Les virus sont à la limite du monde des vivants, ce sont les plus importants, ce ne sont pas des cellules vivantes, car ils n'ont pas de membrane externe et sont délimités par une capsidie ou une enveloppe protéique. Ils ont un matériel génétique qui peut être de l'ADN ou de l'ARN, mais n'ont ni ribosomes, ni métabolisme propre et sont obligatoirement parasites et demandent à être hébergés par un hôte cellulaire dont ils utilisent les ressources pour se reproduire.

Les plus connus des virus sont les rétrovirus qui sont des virus à ARN qui utilisent l'ADN de l'hôte pour se répliquer, le plus célèbre des rétrovirus est le virus du SIDA.

L'hôte du virus peut également être une bactérie, le virus s'appelle alors bactériophage ou phage.

- **Viroïdes**

Ce sont des virus à ARN sans revêtement protéique qui n'infectent que des plantes.

- **Prions**

Ce sont des agents infectieux de nature protéique dépourvus d'acides nucléiques dont la structure spatiale est anormale, ils peuvent infecter le tissu nerveux central et y entraîner une dégénérescence du système nerveux central.

- **Plasmides**

Ce sont des molécules d'ADN capables de se répliquer et qui n'ont pas de phase extracellulaire, ils sont en permanence portés par des cellules hôtes.

- **Transposons**

Ce sont des séquences d'ADN capable de se déplacer de manière autonome dans un génome, par un mécanisme appelé transposition. Cette transposition est rendue possible sous l'effet d'une enzyme, la transposase.

- **Gène**

C'est une unité de l'hérédité contrôlant un caractère particulier. Cet élément génétique correspondant à un segment d'ADN ou d'ARN (virus), situé à un endroit bien précis (locus) sur un chromosome. Chaque région de l'ADN qui produit une molécule d'ARN fonctionnelle est un gène.

**Exemple**

Le noyau de la cellule est « la bibliothèque » qui renferme tout le patrimoine héréditaire de l'individu. Le chromosome est « un livre » de cette bibliothèque et le gène « une page » de ce livre.

# *Chapitre II*

---

*Caractéristiques des acides nucléiques*

## Caractéristiques des acides nucléiques (ADN et ARN)

### Objectifs spécifiques

Au terme de ce cours qui est caractéristiques des acides nucléiques (ADN et ARN), vous devez être capable de :

- Connaître et pouvoir identifier ce que sont les acides nucléiques ;
- Distinguer les différences entre un ADN et un ARN ;
- Comprendre les caractéristiques et les spécificités entre l'ADN et l'ARN ;
- Pouvoir décrire les paramètres physico-chimiques de l'ADN et l'ARN.

## I. Définitions

### I.1. Acides nucléiques

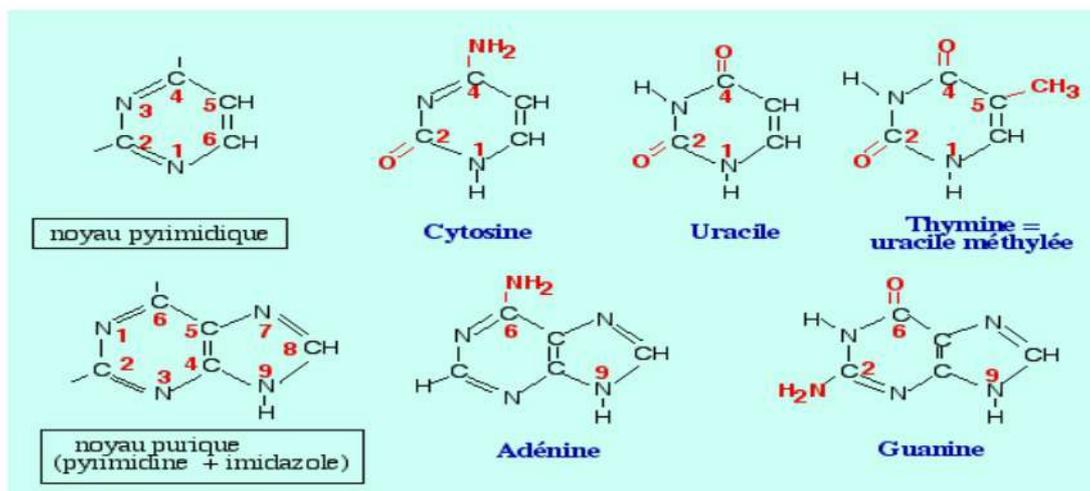
Les acides nucléiques sont des macromolécules présentes dans toutes les cellules et composées d'un enchaînement d'unités structurales appelées '*nucléotides*', se sont donc des polynucléotides. On peut en distinguer deux grands types :

- Les acides désoxyribonucléiques (ADN) : constituant le support de l'information génétique ;
- Les acides ribonucléiques (ARN) : participant à l'expression de l'information génétique.

### I.2. Nucléotides

Les nucléotides sont des molécules composées de :

- **Bases azotés** : dérivés de 2 noyaux hétérocyclés azotés
  - 1) Noyau Purine : Adénine (A), Guanine (G)  $\Rightarrow$  Bases Puriques
  - 2) Noyau Pyrimidine : Cytosine (C), Uracile (U), Thymine (T)  $\Rightarrow$  Bases Pyrimidiques



**Figure 01** : Structures chimiques des bases puriques et des bases pyrimidiques.

### Remarque

La Thymine (T) est uniquement présente au niveau de l'ADN et l'Uracile (U) est uniquement présent sur l'ARN.

On peut assister à une désamination oxydative de la cytosine en uracile. C'est une réaction fréquente.

Dans l'ADN, l'uracile sera réparé par des enzymes de réparation car la présence de cette base entraînerait trop de mutation. Au contraire, au niveau de l'ARN, ces changements ne sont pas graves car la durée de vie de l'ARN et des protéines est courte.

L'absence de 'T' permet la reconnaissance de l'ARN pour les enzymes de dégradation et un gain d'énergie (la greffe de la fonction méthyle demande de l'énergie).

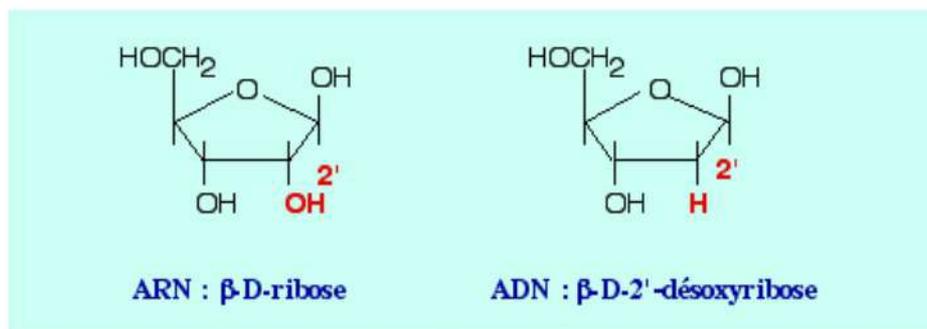
Afin de faciliter la tâche d'écriture pour les biologistes moléculaire, un consensus au niveau international concernant la symbolique de représentation des différentes bases azotés dans les séquences étudiées a été élaboré ce qu'on appelle « Code IUB » pour International Union of Biochemistry.

**Tableau I.** Représentation symboliques des différentes bases azotées.

Symbole	Bases désignées
<b>G</b>	Guanine
<b>A</b>	Adénine
<b>T</b>	Thymine
<b>C</b>	Cytosine
<b>U</b>	Uracile
<b>R</b>	puRine (G ou A)
<b>Y</b>	pYrimidine (A ou C)
<b>M</b>	aMino (A ou C)
<b>K</b>	Keto (G ou T)
<b>S</b>	Strong interaction (3 liaisons H : G ou C)
<b>W</b>	Weak interaction (2 liaisons H : A ou T)
<b>N</b>	aNy (G, A, T ou C)

Si l'on souhaite se référer à une base purine G ou A, au lieu d'écrire G ou A, on va tout simplement y faire référence avec un R.

- **Pentose** : est soit le Ribose (pour l'ARN) ou le Désoxyribose (pour l'ADN), ose à cinq atomes de carbone. Les atomes C du pentose portent les chiffres 1' à 5'. Le signe prime ajouté à ces derniers les distingue de ceux des bases azotés.

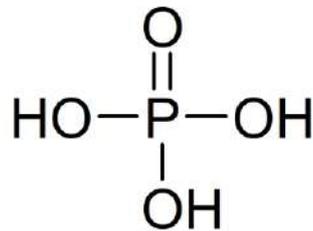


**Figure 02** : Pentoses respectifs de l'ARN et de l'ADN.

**Remarque**

Le désoxyribose est une forme beaucoup plus stable que le ribose. Cette forme désoxy entraîne la stabilisation de l'hélice.

- **L'acide phosphorique ( $H_3PO_4$ )** : possède trois fonctions acides. Deux sont estérifiées dans les ADN et les ARN, la 3<sup>ème</sup> est libre.



**Figure 03** : Molécule d'acide phosphorique.

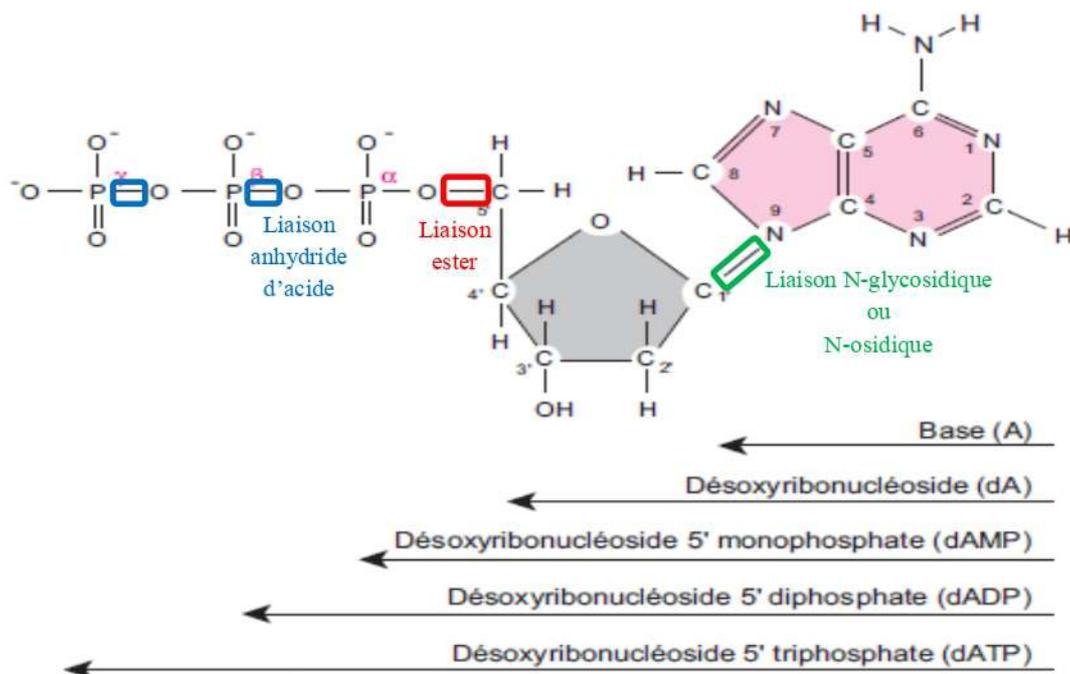
### Remarques

- Les phosphates permettent la solubilisation de l'ADN dans l'eau grâce à leurs charges négatives.
- Les phosphates permettent la polymérisation<sup>1</sup> des acides nucléiques (des nucléotides).

## II. Liaisons des acides nucléiques

L'ose, par son carbone 1', établit une liaison N-glycosidique avec l'azote 1 des pyrimidines ou l'azote 9 des purines. Par son carbone 5', il forme une liaison ester avec un premier groupe phosphate (liaison phosphoester) (désigné pour cette raison phosphate  $\alpha$ ). Ce dernier peut former des liaisons anhydride d'acide (Fig. 4) avec des groupes phosphate  $\beta$  et  $\gamma$  ou des liaisons ester avec un groupe OH porté par le carbone 3' du pentose précédent.

L'union d'une base et d'un pentose s'appelle un nucléoside. Dès qu'un groupe phosphate est présent, l'ensemble est désigné par le terme nucléotide.



**Figure 04** : Structure détaillée d'un désoxyribonucléotide triphosphate.

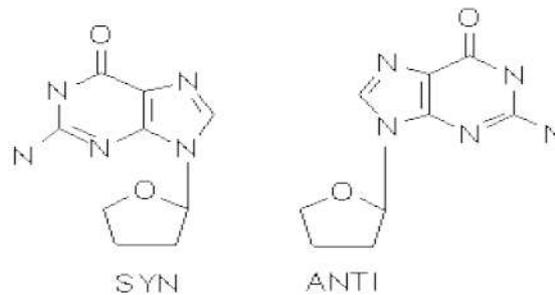
En remplaçant la lettre A (Adénine) par les lettres G, C, T/U et le désoxyribose par le ribose, la nomenclature est généralisable à tous les nucléotides de l'ADN et de l'ARN.

**Remarques**

La liaison Pentose–base est une liaison N-osidique. Cette association est appelée nucléoside et en fonction des bases on aura :

- Thymidine = Thymine + Désoxyribose
- Uridine = Uracile + Ribose
- ◌ Union de plusieurs molécules d'un composé pour former une grosse molécule.
  - Adénosine = Adénine + Ribose / Désoxyribose
  - Cytidine = Cytosine + Ribose / Désoxyribose

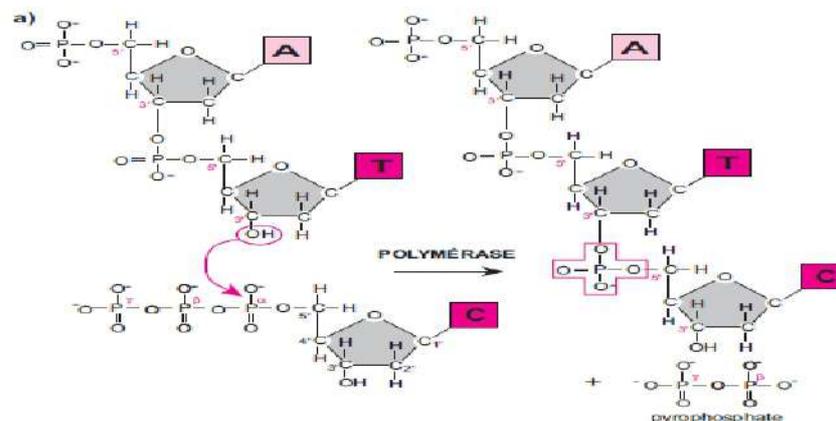
Une base reliée à un sucre peut être en position *anti* ou *syn*, évidemment la position la plus courante dans l'ADN est *anti*.



**Figure 05 :** Position *anti* et *syn* des bases azotées reliées au sucre.

La structure d'un polynucléotide (polymère de nucléotides) repose sur la formation d'un nombre parfois élevé (plusieurs millions) de liaisons phosphodiester entre les nucléotides constitutifs. Chaque liaison phosphodiester s'établit entre le groupe -OH du carbone 3' d'un pentose et le groupe  $-H_2PO_4$  du carbone 5' du pentose constituant le nucléotide suivant. C'est la liaison phosphodiester (3'-5'). Ainsi se crée une molécule avec un enchaînement ordonné de nucléotides.

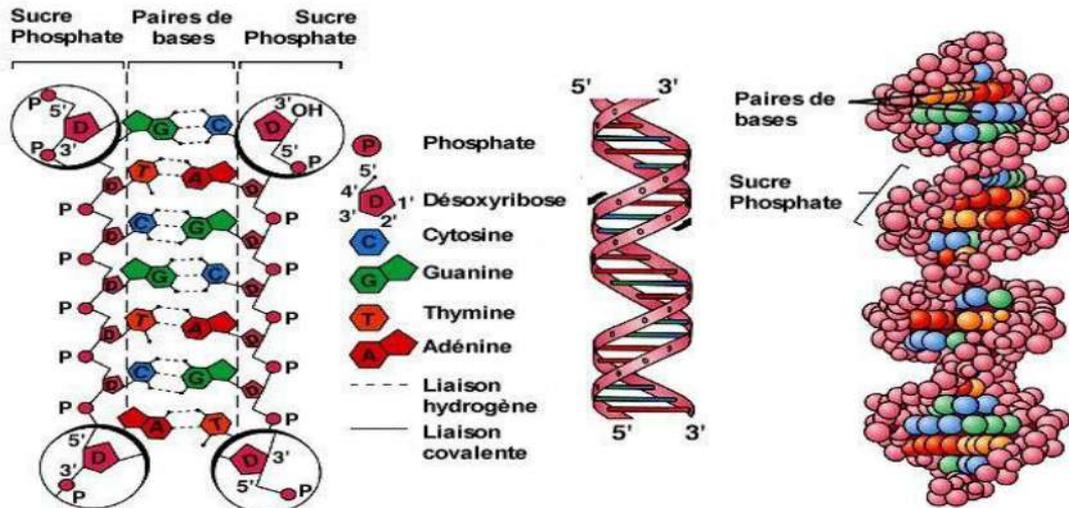
Il en résulte que l'orientation du polynucléotide se trouve définie par la position du groupe phosphate libre du premier nucléotide, généralement le phosphate porté par le carbone 5', et par le groupe -OH libre porté par le carbone 3' du dernier nucléotide. On dit ainsi que le polynucléotide est orienté dans le sens 5' vers 3' (Fig.06). Les nucléotides d'une même chaîne sont associés par des liaisons fortes.



**Figure 06 :** Principe de formation d'un brin d'ADN par la polymérisation des nucléotides.

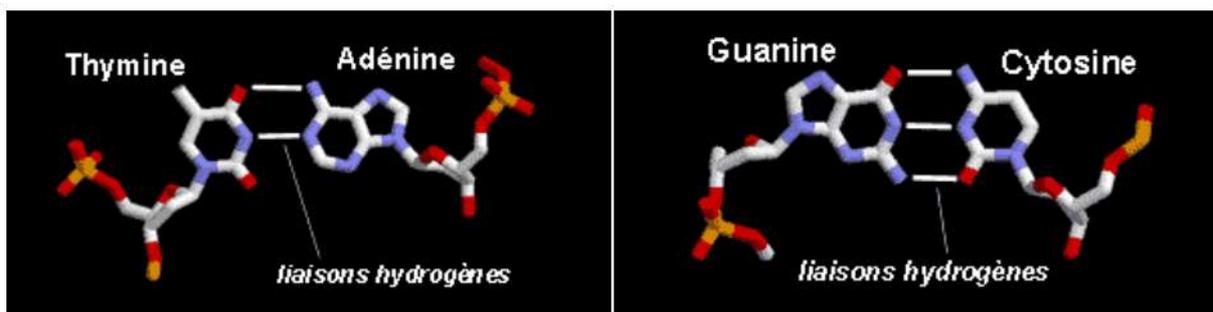
### III. Structure de la molécule d'ADN

L'ADN (Acide **D**éoxyribo Nucléique) est une **macromolécule** de longue chaîne polynucléotidique **bicaténaire**, c'est-à-dire constitué de deux brins, **hélicoïdales** et **antiparallèle**, c'est-à-dire disposés dans des directions opposées ( $5' \rightarrow 3'$  et  $3' \rightarrow 5'$ ), associés par des liaisons hydrogène entre les bases formant de ce fait dans l'espace une double hélice et assurant le maintien de cette structure.



**Figure 07 :** Structure plane et en 3D des deux chaînes antiparallèles de l'ADN.

Les bases azotées liées par ces liaisons hydrogènes sont tournées vers l'intérieur, tandis que les riboses et les acides phosphoriques, hydrophiles sont tournés vers l'extérieur. Celles-ci s'établissent toujours entre une purine de l'un des brins et une pyrimidine de l'autre brin et sont dites '**liaisons faibles**'. Dans cet appariement complémentaire, l'adénine (A) est toujours associée à la thymine (T) par deux liaisons hydrogène et la guanine (G) interagit avec la cytosine (C) grâce à trois liaisons hydrogène formant de ce fait des paires (pb) (Fig. 8).



**Figure 08 :** Liaisons hydrogènes des différentes bases azotées.

#### III.1. Lois de Chargaff

Les travaux d'Erwin Chargaff fait en 1950, sur la quantité de bases contenues dans l'ADN de tout organisme ont mené aux déductions suivantes appelées « lois de Chargaff »:

- La quantité totale de pyrimidines (T+C) est toujours égale à la quantité totale de purines (A+G) donc  $[A+G = C+T = 1]$ .

- La quantité de 'T' est toujours égale à la quantité de 'A' ( $T = A$ ), et la quantité de 'C' est toujours égale à la quantité de 'G' ( $G = C$ ).
- Cependant, la quantité de 'A+T' n'est pas toujours égale à 'G+C'. Ce rapport est différent selon les organismes  $A+T \neq G+C$ .

### III.2. Formes de l'ADN

Dans l'espace les deux chaînes présentent une configuration hélicoïdale. Elles s'enroulent autour d'un axe imaginaire pour constituer une double hélice à rotation droite (Forme A et B) ou bien à rotation gauche (Forme Z).

Les paires de bases forment des plateaux, de ce fait, l'hélice est un ensemble de plateaux qui forment une espèce d'escalier. La rotation entre 2 plateaux consécutifs est de l'ordre de  $36^\circ$ . Il y a donc 10 paires de bases par tour d'hélice. Ce même tour d'hélice fait 34 Å de long pour un diamètre de 20 Å. On a donc 34 Å entre 2 paires de bases consécutives.

L'hélice d'ADN peut avoir plusieurs configurations (formes) suivant la séquence et la concentration ionique des milieux. Celles-ci peuvent se distinguer par :

- Nombre de paires de bases dans chaque tour d'hélice ;
- Pas de l'hélice (longueur) ;
- Diamètre hélicoïdal de la molécule ;
- Sens de la double hélice (droite, gauche).

#### 1) Forme B

C'est généralement la forme biologique la plus importante retrouvée dans la cellule. Elle correspond à celle décrite par Watson et Crick en 1953. Cette forme est caractérisée par un pas (sens de rotation vers la droite) de l'hélice de 10.5 pb (34 Å) et un diamètre de 2,0 nm qui se traduit par la formation de sillons, sillons de taille différentes avec une alternance régulière petit sillon / grand sillon, la présence de ces sillons est un élément essentiel reconnu par un nombre de protéine régulatrice de l'expression des gènes.

#### 2) Forme A

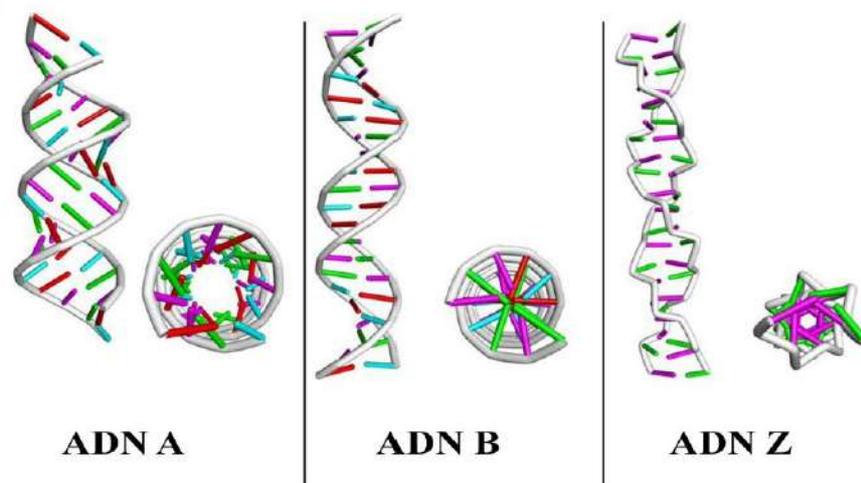
L'ADN A est une forme de la double hélice d'ADN proche de l'ADN B dans laquelle les bases nucléiques sont plus inclinées par rapport à l'axe de la double hélice, ce qui augmente le diamètre de la structure ainsi que le nombre de paires de bases par unité de longueur de la double hélice. Le diamètre d'un ADN A est typiquement de 2,3 nm tandis que la double hélice s'allonge en moyenne de 0,24 nm par paire de bases, contre 0,32 nm pour l'ADN B). Ceci a pour effet d'élargir le grand sillon et de rendre le petit sillon plus étroit.

Cette forme d'ADN dériverait de la forme B lorsque les conditions physiques du milieu à savoir l'hygrométrie (taux d'humidité) est inférieur à 70-75% ce qui crée un épaissement de la double hélice. Cet ADN A est retrouvé en particulier dans les spores végétales.

#### 3) Forme Z

C'est une double hélice à rotation gauche c'est-à-dire inverse de celui de l'ADN B. Le nombre de paires de bases par tour d'hélice est de 12, le pas d'hélice est de 4.56 nm, et un diamètre de 1.8 nm due au changement dans le sens de rotation de l'hélice. Cette forme présente les caractéristiques suivantes:

- Le squelette présente une conformation en zigzag favorisée par un taux de GC élevé (moins lisse que l'ADN B). La forme Z n'est en fait qu'une séquence alternée de C et de G : CGCGCGCG ;
- Il y'a un seul sillon, qui ressemble au petit sillon de l'ADN B ;
- Les phosphates sont plus proches les uns des autres que dans l'ADN B ;
- Cette structure empêche l'interaction avec certaines protéines de régulations ;
- L'ADN Z ne peut pas former de nucléosomes, important dans le conditionnement de l'ADN cellulaire ;
- La transformation de l'ADN B en ADN Z se fait lors de la transcription des gènes (au site d'initiation de la transcription).



**Figure 09 :** Différentes formes de l'ADN.

Toutes les bases ont la même conformation par rapport à l'ose et sont à l'intérieur de l'hélice. Les groupements phosphates et les désoxyriboses sont à l'extérieur.

**Tableau II.** Paramètres structurels indicatifs des trois principales formes d'ADN.

Paramètres	ADN A	ADN B	ADN Z
Sens de l'hélice	droite	droite	Gauche
Motif répété	1 bp	1 bp	2 bp
Rotation par paire de bases	32,7°	34,3°	60°/2
Paire de bases par tour d'hélice	11	10,5	12
Pas de l'hélice par tour	2,82 nm	3,32 nm	4,56 nm
Allongement de l'axe par paire de bases	0,24 nm	0,32 nm	0,38 nm
Diamètre	2,3 nm	2,0 nm	1,8 nm
Inclinaison des paires de bases sur l'axe	+19°	-1,2°	-9°
Torsion moyenne ( <i>propeller twist</i> )	+18°	+16°	0°
Orientation des substituants de bases	anti	anti	<u>Pyrimidine</u> : anti, <u>Purine</u> : syn
Torsion endocyclique du furanose	C3'- <i>endo</i>	C2'- <i>endo</i>	<u>Cytosine</u> : C2'- <i>endo</i> , <u>Guanine</u> : C2'- <i>exo</i>

### III.3. Importance de l'ADN dans la cellule

En février 1953, on raconte que Francis Crick aurait surgi dans le pub 'The Eagle' à Cambridge en annonçant qu'il avait, avec James Watson, découvert rien de moins que le secret de la vie. C'est la découverte de la structure de l'ADN qui allait permettre, dans les années suivantes, de comprendre les bases moléculaires de l'hérédité.

L'ADN constitue le matériel génétique de la cellule, et le matériel génétique d'un être vivant doit :

#### 1) Détenir l'information génétique propre à son espèce

L'ordre d'enchaînement des nucléotides constitue un message, une information permettant à la cellule d'assembler dans le bon ordre les acides aminés de ses protéines. Et c'est le type de protéine que fabrique un être vivant qui détermine ses caractéristiques physiques.

#### 2) Pouvoir se reproduire avec le moins d'erreurs possibles

À chaque fois qu'une cellule se divise en deux cellules, l'information doit être reproduite intégralement de façon à ce que chacune des deux cellules obtenues détienne l'ensemble de l'information qui était dans la cellule de départ.

#### 3) Pouvoir se modifier (mais pas trop)

L'ADN n'est pas une molécule d'une stabilité à toute épreuve. Des modifications accidentelles peuvent modifier le message. Un nucléotide, par exemple, peut être remplacé par un autre ce qui modifie le message. Ce sont ces modifications qui permettent l'évolution. Sans elles, l'évolution ne serait pas possible.

#### 4) Pouvoir être traduit en caractéristiques physiques

Le message contenu dans l'ADN doit pouvoir être "lu" par la cellule. Ce n'est pas tout de détenir de l'information, mais plus encore, cette information doit pouvoir être utilisée, elle doit pouvoir être traduite en quelque chose de concret, en protéines dans le cas de l'ADN.

Chacune de ces propriétés fondamentales de la vie découle de la structure de l'ADN. La découverte de la structure de l'ADN a permis de comprendre sous quelle forme est l'information dans la cellule, comment cette information peut se reproduire, comment elle peut se modifier permettant ainsi l'évolution et comment elle peut être traduite en caractéristiques physiques transmissibles à travers les générations.

### III.4. Principaux types d'ADN

#### III.4.1. Matériel génétique des virus

La molécule d'ADN ou d'ARN constituant le matériel génétique des virus peut être double ou simple brin, elle est généralement de petite taille. Ce matériel génétique est souvent associé à des protéines qui forment une capsid.

#### III.4.2. ADN des procaryotes (cas des bactéries)

Chez les procaryotes, tels que les bactéries, l'ADN est en général présent sous la forme d'un seul chromosome double brin circulaire super-enroulé et associé à des protéines pour

former le nucléoïde. Il est éventuellement possible de retrouver un plasmide (ADN double brin circulaire de petite taille et présent en un nombre variable de copies).

### III.4.3. ADN des eucaryotes

Chez les eucaryotes, l'ADN est présent dans le noyau cellulaire principalement, mais aussi dans les mitochondries et les chloroplastes, organites cellulaires.

#### a. ADN chromosomique

Dans le noyau, l'ADN 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 ainsi que des protéines non-histones formant de ce fait la chromatine.

#### b. ADN mitochondrial (ADNmt)

L'ADN mitochondrial est présent en 4 à 6 molécules bicaténaire et circulaire, dont la composition nucléotidique diffère de celle du noyau avec un taux en GC% faible. Il est essentiel au bon fonctionnement de ces organites.

Chez les mammifères, l'ADNmt représente moins de 1 % de l'ADN total, mais dans les plantes, la quantité est variable. Il code les ARNr, les ARNt et quelques protéines mitochondriales, souvent essentielles à la respiration cellulaire (jusqu'à 30 chez les animaux).

Dans la plupart des cas, cet ADN est transmis par la mère.

#### c. ADN chloroplastique (ADNcp)

C'est un ADN présent au niveau des chloroplastes. Chaque chloroplaste peut contenir jusqu'à 100 copies, ils sont généralement linéaires ou ramifiées, de petite taille (37 à 220 kb) formant une centaine de gènes. Même si, le chloroplaste possède un petit génome, le grand nombre de chloroplastes par cellule implique une proportion significative d'ADN chloroplastique par rapport à l'ADN total dans une cellule végétale.

Les ribosomes sont constitués d'ARNr, synthétisés dans les chloroplastes, et de protéines codées par les génomes nucléaires et chloroplastiques.

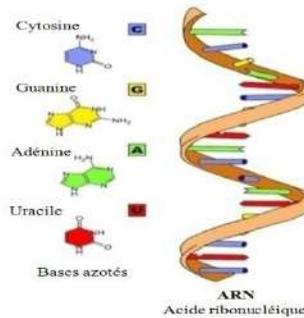
## IV. Structure de la molécule l'ARN

L'ARN (Acide RiboNucléique) est une macromolécule composée d'un seul brin nucléotidique linéaire (molécules monocaténares). Cependant, dans une même chaîne d'ARN des portions peuvent se replier facilement sur elle-même et être sous forme bicaténaire en s'appareillant avec une complémentarité suivant la règle (02 liaisons hydrogène entre A et U et 03 liaisons hydrogène entre C et G).

Il est en général synthétisé dans les cellules à partir d'une matrice d'ADN dont il est une copie. Les cellules utilisent en particulier l'ARN comme un support intermédiaire des gènes pour synthétiser les protéines dont elles ont besoin.

Les nucléotides d'ARN, contiennent du *ribose* ce qui rend l'ARN chimiquement moins stable que l'ADN (ce sont des *ribonucléotides*). Les bases sont A, U, C et G. Donc l'Uracile (U) remplace toujours la Thymine (T).

Les molécules d'ARN présentes dans les cellules sont plus courtes, leur taille variant de quelques dizaines à quelques milliers de nucléotides, contre quelques millions à quelques milliards de nucléotides pour l'ADN.



**Figure 10 :** Brin d'ARN.

#### IV.1. Types et rôles des ARN

Les ARN sont très divers en forme, en taille et par leurs rôles. Il existe 03 types d'ARN qui jouent un rôle essentiel dans la transcription (ARNm) et la traduction (ARNt et ARNr) de l'information génétique qui aboutit à la biosynthèse des protéines :

##### IV.1.1. ARN messagers (ARNm)

Les ARNm sont monocaténaire et correspondent à un faible pourcentage des ARN cellulaires (2 à 4 %). Ils ont une durée de vie très courte (1 à 2 minutes chez les bactéries, plus longue chez les eucaryotes).

Ce sont des copies en ARN de l'ADN contenu dans les gènes : leur séquence est complémentaire de certaines séquences d'ADN. Ces ARNm sont transcrits dans le noyau, puis exportés dans le cytoplasme, où elles sont ensuite traduites en protéines grâce aux ribosomes. Sa taille est proportionnelle à celle des protéines pour laquelle il code.

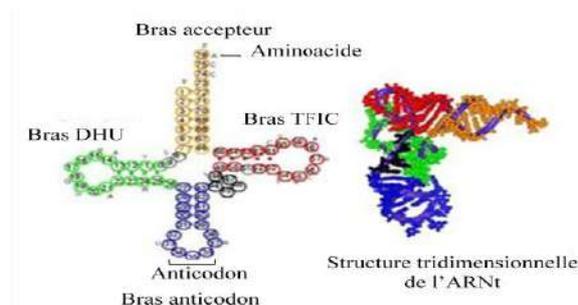
##### IV.1.2. ARN de transfert (ARNt)

L'ARNt possède une structure tertiaire par repliement de la chaîne nucléotidique dans l'espace. Il possède une double spécificité qui joue un rôle d'adaptateur moléculaire :

- Une pour l'acide aminé (extrémité de fixation de l'AA) ;
- L'autre pour l'ARNm (anticodon permettant de connaître le codon de l'ARNm).

C'est un vecteur de petite taille (73 à 93 nucléotides), qui va reconnaître les acides aminés dans le cytoplasme pour les amener jusqu'au brin d'ARN messenger où s'effectue la synthèse protéique ; ce sont donc des sorte de transporteurs d'acides aminés.

Il existe environ 70 ARNt différents alors qu'il existe 20 acides aminés différents (Compte tenu de dégénérescence du code génétique). Ils représentent environ 15% de l'ARN cellulaire total. La structure schématique d'un ARNt est représenté en figure 13.



**Figure 11 :** Représentation de l'ARNt.

#### IV.1.3. ARN ribosomiques (ARNr)

Les ARN ribosomiques sont les ARN les plus représentés dans les cellules, ils représentent environ 80% des ARN cellulaires totaux. Ils participent à la constitution et le maintien de l'intégrité des ribosomes, organites cellulaires, qui constituent la tête de lecture de l'information génétique transcrite par l'ARNm et donc impliqué dans la traduction protéique.

On distingue différents types d'ARN ribosomiques en fonction de leur coefficient de sédimentation.

Il est à noter que les ARNr sont synthétisés dans le noyau à partir de l'ADN et s'assemblent au niveau des nucléoles avec les protéines, puis passent dans le cytoplasme.

On retrouve aussi de manière minoritaire des petits ARN, principalement au niveau du noyau et donc présent chez les eucaryotes :

#### IV.1.4. ARN small nuclear (ARNsn) et ARN small nucleolar (ARNsno)

Ils sont présents uniquement chez les eucaryotes et représentent < 1% des ARN totaux. Ils ont des rôles très différents :

- Les ARNsn vont intervenir dans le mécanisme d'épissage donc maturation des ARNm ;
- Les ARNsno vont intervenir dans les modifications chimiques des ARN ribosomique (maturation de ces ARNr).

#### IV.1.5. ARN micro (ARNmi) et Small interfering RNA (ARNsi)

Ils représentent < 1% des ARN totaux. Ils ont un rôle dans la stabilité des ARNm. Ils ont la propriété d'être complémentaires au niveau de la séquence de certains ARNm. Leur hybridation avec leur cible induit un mécanisme qui aboutit à la dégradation des ARNm concernés.

C'est un mécanisme de régulation de l'expression des gènes, bien que cela se produise après la réplication. Un micro-ARN donné peut cibler différents ARNm. Il semblerait aussi que certains ARN réguleraient la traduction de certains ARNm en protéines, donc ils ont un probable rôle dans la traductibilité.

Ces micro ARN peuvent aussi agir comme des oncogènes ou suppresseurs de tumeurs en fonction de l'ARNm avec lequel ils s'hybrident. Donc des dérégulations de certains miARN entraînent des passages à l'état cancéreux. Ils peuvent aussi agir comme des cofacteurs pour réguler l'expression des gènes en contrôlant les facteurs d'initiation. Ils semblent également jouer un rôle dans l'épissage et dans la traduction.

#### IV.1.6. ARN long noncoding (LncRNA)

Ils représentent < 1% des ARN totaux. Ce sont des ARN de grande taille, synthétisés par transcription (ils peuvent être constitués de plusieurs kilobases). Ils jouent un rôle très important en cancérologie (ex : dans le cancer du sein). Ils interviennent dans la régulation de la transcription d'ARNm en se liant aux facteurs de transcription et en régulant l'expression de ces facteurs de transcription. Ils ont aussi un rôle dans l'épissage et la traduction.

**Small interfering RNA** = Petits ARN d'interférence

**Small nuclear RNA** = Petit ARN nucléaire

**Small nucleolar RNA** = Petit ARN nucléolaire

**Long noncoding RNA** = Long ARN non codant

***Tous ces ARN proviennent de la transcription. Mais seuls les ARNm sont traduits en protéines. Les autres ARN exercent une activité propre en tant qu'ARN.***

## V. Propriétés physico-chimiques des acides nucléiques

- 1) La taille de l'ADN ainsi que sa masse varie en fonction du type de cellule. Pour l'ARN, la taille varie de 15 à 20 nucléotides à quelques milliers de nucléotides.

**Tableau III.** Propriétés en taille, masse et longueur de l'ADN des différents types cellulaires.

Cellule	ADN d'une cellule		
	Taille	Masse	Longueur
Adénovirus	$\approx 35 \times 10^3$ pb	$\approx 10^7$ Da	$\approx 1.1 \times 10^{-6}$ m
<i>E. coli</i> (Bactérie)	$\approx 3.5 \times 10^6$ pb	$\approx 10^9$ Da	$\approx 1.3 \times 10^{-3}$ m
Homme	$\approx 2 \times 3 \times 10^9$ pb	$\approx 10^{12}$ Da	$\approx 2$ m

1 kpb = 1 000 pb, 1 Mpb = 1 000 000 pb

- 2) La densité des acides nucléiques est en fonction de la nature des nucléotides.
- 3) Les acides nucléiques (ADN et ARN) sont chargés négativement en raison des groupements phosphates. Placés dans un champ électrique (électrophorèse), ces molécules migrent vers le pôle positif qui est l'anode.
- 4) Les bases azotées (A, T, G, C, U) ont une absorption maximale à  $\approx 260$  nm.
- 5) Comme les liaisons hydrogènes sont faibles, elles peuvent donc être facilement rompues. Une chaleur élevée, une faible salinité, à pH élevé/solution basique et à pH faible/solution acide, peut dissocier les deux chaînes de la double hélice d'ADN. Cette séparation est appelée 'fusion ou dénaturation de l'ADN'. Cette fusion est réversible : les deux chaînes peuvent s'hybrider à nouveau.

La température à laquelle 50 % de l'ADN bicaténaire est dissocié en deux molécules d'ADN monocaténaire est dite température de fusion ou température de semi-dénaturation de l'ADN, notée  $T_m$ . On peut la mesurer en suivant l'absorption optique à 260 nm de la solution contenant l'ADN : cette absorption augmente au cours du désappariement, ce qu'on appelle hyperchromicité.

La stabilité d'une double hélice d'ADN dépend essentiellement du nombre de liaisons hydrogène à briser pour en séparer les deux brins. Par conséquent, plus la double hélice est longue, plus elle est stable.

Cependant, les paires GC étant unies par trois liaisons hydrogène au lieu de deux pour les paires AT, la stabilité des molécules d'ADN bicaténaires de même longueur croît avec le nombre de paires GC qu'elles contiennent, mesuré par leur taux en GC.

La température de fusion de l'ADN dépend par conséquent de la longueur des molécules, de leur taux en GC, de leur séquence, de leur concentration dans le solvant et de la force ionique dans celui-ci.

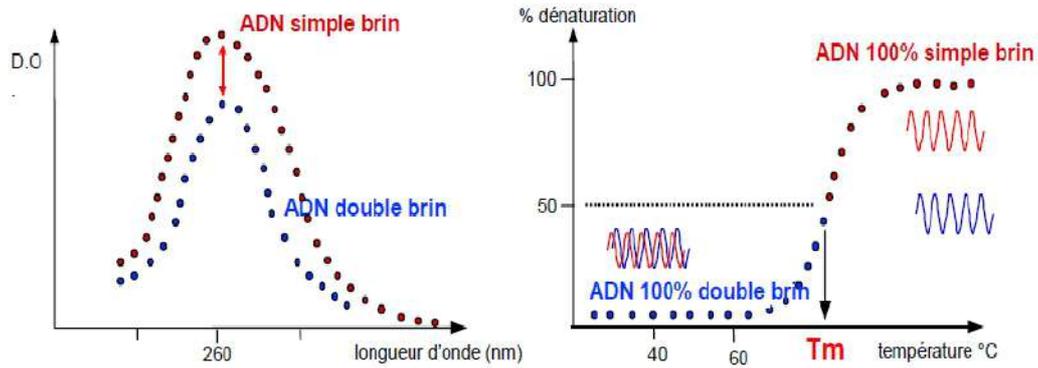


Figure 12 : Représentation de la dénaturation de l'ADN sous l'effet de la température.

6) La molécule d'ADN peut se renaturaliser suite à une dénaturation et cela de façon spontanée par réappariement des deux brins complémentaires.

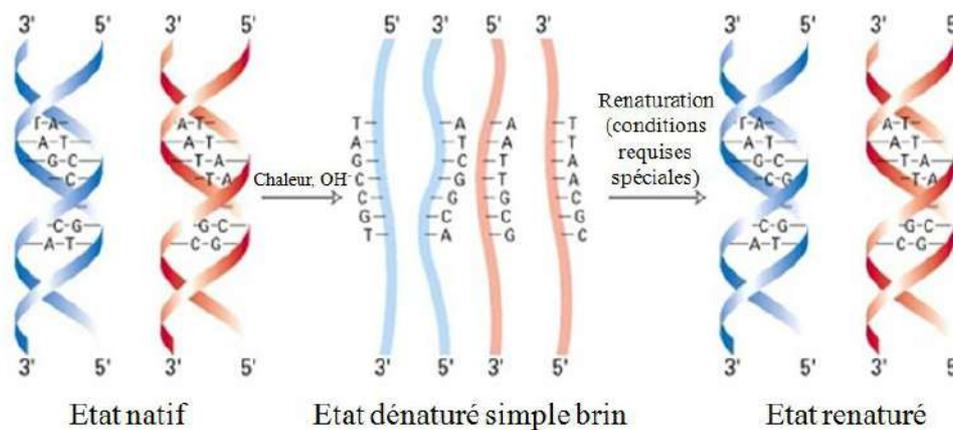


Figure 13 : Renaturation des brins d'ADN après dénaturation.

7) L'interaction de certaines molécules peut modifier la structure de l'ADN et altérer sa fonction. Ainsi le cis-platine, molécule utilisée dans le traitement de certains cancers inhibe la réplication de l'ADN des cellules tumorales en modifiant la structure de la double hélice.

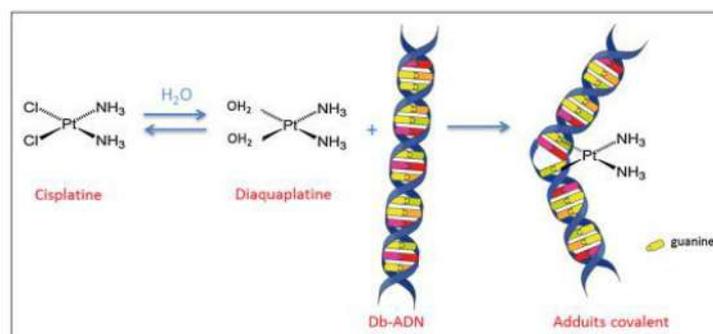


Figure 14 : Interaction du cis-platine avec la molécule d'ADN.

# *Chapitre III*

---

*Mécanismes moléculaire*

## Mécanismes moléculaires

### (Réplication de la molécule d'ADN)

#### Objectifs spécifiques

Au terme de ce cours qui est la réplication de l'ADN, vous devez être capable de :

- Connaître ce qu'est la réplication de l'ADN ;
- Identifier la phase du cycle cellulaire durant laquelle a lieu la réplication de l'ADN ;
- Connaître les différentes étapes de la réplication de l'ADN ;
- Pouvoir décrire et connaître les facteurs intervenant dans la réplication de l'ADN.

### I.1. Introduction

L'ADN a l'importante propriété de pouvoir être reproduit à l'identique, ce qui permet à l'information génétique de se transmettre d'une cellule mère aux cellules filles lors de la division cellulaire.

#### I.1.1. Définition

La **réplication de l'ADN**, appelée aussi **duplication de l'ADN**, est le processus au cours duquel l'ADN est synthétisé grâce à un complexe enzymatique « l'ADN polymérase ».

Ce mécanisme permet d'obtenir, à partir d'une molécule d'ADN initiale, deux molécules identiques à la précédente, en vue de leur distribution aux deux cellules filles pendant la mitose. C'est donc au cours du cycle cellulaire, plus précisément pendant la phase S, avant la mitose, que l'ADN se dédouble.

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 de ce fait la vitesse de polymérisation.

#### I.1.2. Cycle cellulaire

La vie de la cellule se déroule entre deux mitoses. Chez les Mammifères, cette période dure en moyenne 30 heures bien qu'il y ait des cellules dont la vie soit très courte ou au contraire très longue. Durant ces trente heures la cellule traverse quatre phases :

- **Phase G1** où le génome étant diploïde, chaque gène est représenté en deux exemplaires. La chromatine est accessible aux ARN-polymérases qui transcrivent les gènes en messagers, qui seront à leur tour traduits.
- **Phase S**, vers la moitié du cycle commence la réplication de l'ADN : les ADN-polymérases vont mettre environ 8 heures pour recopier en double l'ADN de chaque chromosome. Durant cette phase la transcription est inhibée.
- Puis la cellule entre en **phase G2** où chaque gène est représenté en quatre exemplaires. La chromatine est à nouveau accessible aux ARN-polymérases qui recommencent à transcrire.
- Enfin survient la **mitose**, qui donne naissance à deux cellules filles. Chacune recevra une des copies identiques d'ADN de chaque chromosome et chaque gène y sera représenté en deux exemplaires.

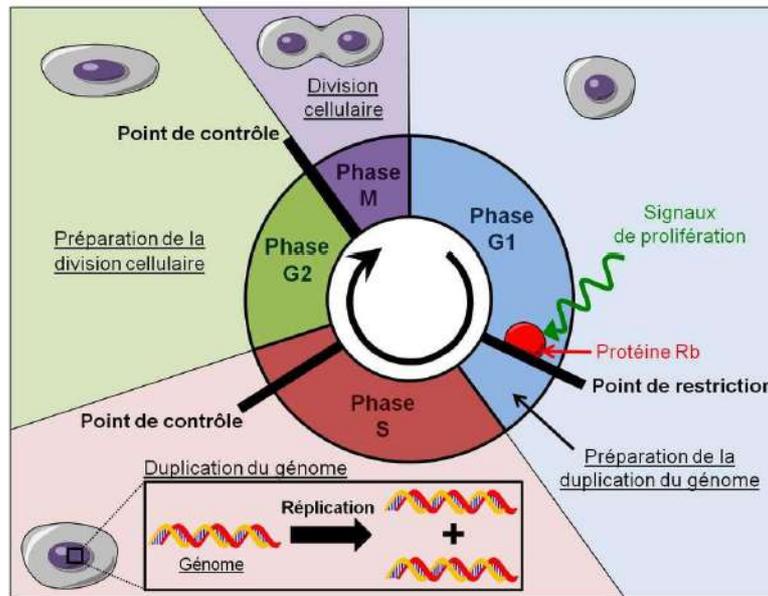


Figure 15 : Etapes du cycle cellulaire.

## I.2. Lois fondamentales de la réplication

Ces lois sont identiques chez les procaryotes et les eucaryotes.

- **Sens de la réplication**

La réplication se fait dans le sens  $5' \rightarrow 3'$ , de façon complémentaire, selon les règles d'appariement des bases :  $A = T / G \equiv C$ . Pourquoi ? Et bien il se trouve que l'ADN polymérase a besoin d'une extrémité  $3'OH$  libre pour rajouter des nucléotides dans le sens  $5' \rightarrow 3'$ . Ainsi, il est impossible d'allonger un brin dans le sens  $3' \rightarrow 5'$  (et on verra plus tard que c'est précisément ça qui pose problème au niveau du brin retardé).

Elle se fait selon un mode antiparallèle (le brin matrice complémentaire est copié dans le sens  $3' \rightarrow 5'$ ).

- **La réplication est bidirectionnelle**

La réplication démarre à l'intérieur de la molécule d'ADN en un point appelé **origine de réplication** puis se poursuit dans les deux directions en copiant les deux brins à chaque fourche. Chaque origine de réplication engendre deux fourches de réplication.

Dans un chromosome circulaire (chez les bactéries et certains virus), une origine est souvent suffisante et les deux fourches qui en partent se confondent du côté opposé du cercle pour finir la réplication.

Les longs chromosomes eucaryotes comportent un grand nombre d'origines : les deux fourches partant d'une origine donnée progressent jusqu'à rencontrer les fourches partant d'origine voisines.

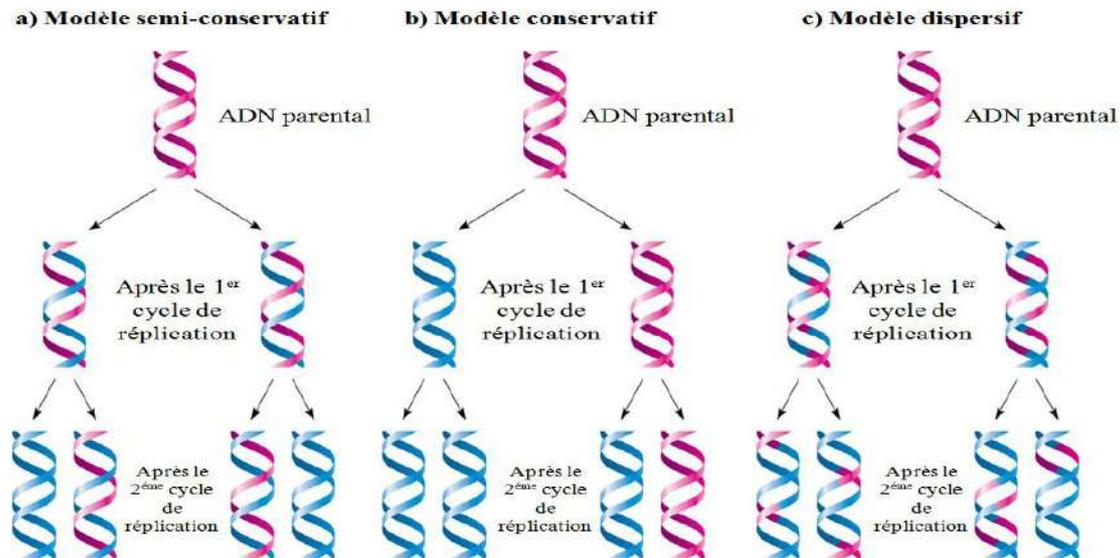
- **Modèles de la réplication**

- a) *Model conservatif*

A partir d'une molécule d'ADN bicaténaire "mère", on forme une nouvelle molécule d'ADN bicaténaire totalement néo-formée.

### b) Model dispersif

Aucun brin ne reste intact, les nouveaux brins sont constitués à la fois d'ADN de la molécule mère et d'ADN néo-formé.



**Figure 16 :** Différents modèles de réplication de l'ADN.

### c) Model semi-conservatif

Celle-ci est tenue pour universel aussi bien chez les procaryotes que chez les eucaryotes. Une réplication semi-conservative veut dire que sur les deux brins toute molécule d'ADN, est toujours formé d' :

- un brin d'ADN parental (brin ancien)
- un brin d'ADN fils (brin nouvellement formé).

### Expériences de Meselson et Stahl (1957)

L'expérience de Meselson et Stahl est une expérience réalisée pour la première fois en 1957 par **Matthew Meselson** et **Franklin Stahl**, qui démontra la réplication semi-conservative de l'Acide désoxyribonucléique.

Ces chercheurs ont utilisé *Escherichia coli* car sa culture dans un milieu minimal est facile et rapide et son temps de génération est court, ce qui permet de suivre l'évolution de l'ADN à travers les générations successives.

*E. coli* a été cultivé pendant plusieurs générations dans un milieu contenant comme seule source d'azote le  $\text{NH}_4\text{Cl}$  composé du  $\text{N}^{15}$ . Le résultat est que tous les composés cellulaires de la bactérie contenant de l'azote entre autres les purines et les pyrimidines de l'ADN étaient marqués par le  $\text{N}^{15}$ , ce qui fait que son ADN était plus dense que l'ADN normal contenant le  $\text{N}^{14}$ . La bactérie a été ensuite transférée rapidement dans un milieu contenant le  $\text{N}^{14}$ .

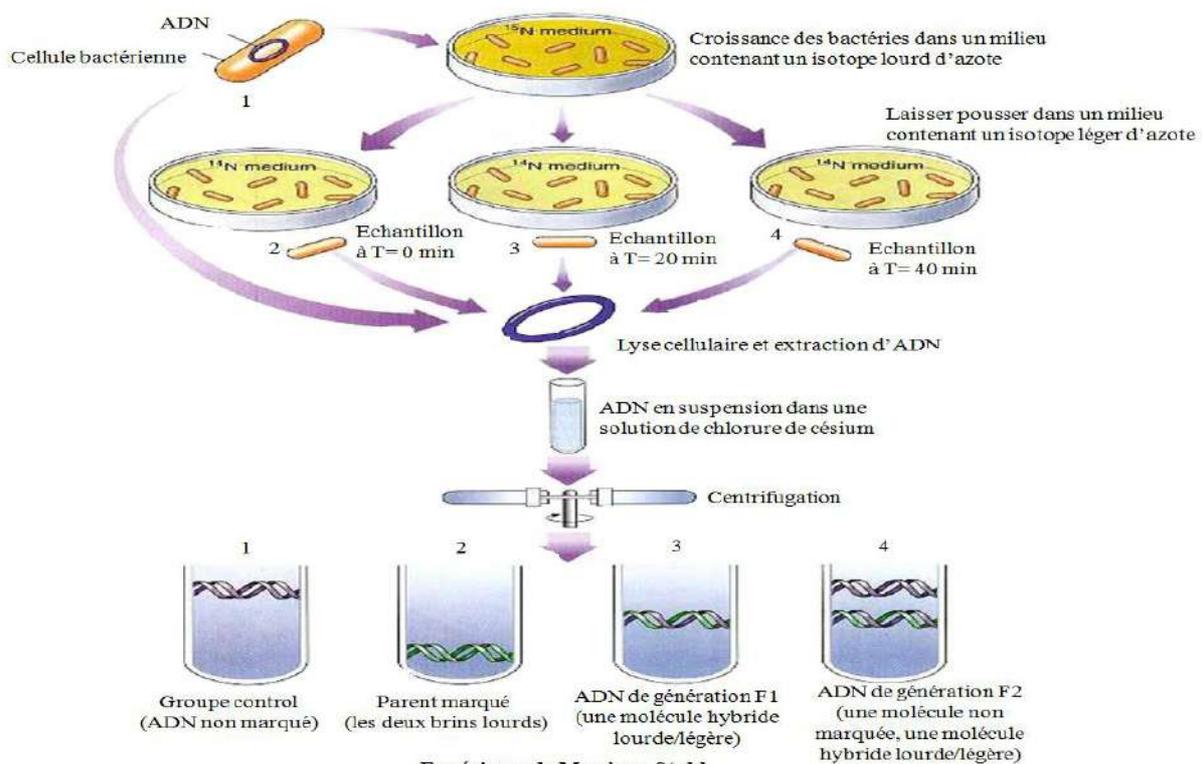
La distribution du  $\text{N}^{14}$  et  $\text{N}^{15}$  a été révélée par la technique de centrifugation sur gradient de densité. Elle consiste à dissoudre une petite quantité d'ADN dans une solution concentrée de Chlorure de Césium qui a une densité proche de celle de l'ADN ( $1,7\text{g}/\text{cm}^3$ ). Cette solution

est centrifugée jusqu'à atteindre l'équilibre. Un gradient stable de densité est alors établi et qui varie de 1,66 à 1,76 g/cm<sup>3</sup>. Les molécules d'ADN formeront une bande à l'endroit du gradient qui correspond à leur densité.

Un mélange d'ADN N<sup>14</sup> et d'ADN N<sup>15</sup> a donné des bandes séparées parce que leur densité est différente de 1%. Si on extrait l'ADN d'une génération après l'avoir transférée dans un milieu contenant le N<sup>14</sup>H<sub>4</sub>Cl, une seule bande apparaît qui est intermédiaire entre la bande contenant seulement le N<sup>15</sup> et celle contenant seulement le N<sup>14</sup>.

Ceci indique que cette bande est formée d'un hybride de N<sup>15</sup> et de N<sup>14</sup>, ce qui veut dire que la moitié de son ADN est parental et l'autre moitié est nouvellement synthétisée.

Si on extrait l'ADN deux générations après le transfert des bactéries dans un milieu contenant le N<sup>14</sup>, on trouve deux bandes : l'une est la bande intermédiaire que nous avons déjà trouvée après la première génération et l'autre correspond à la bande qui contient entièrement du N<sup>14</sup>.



**Expérience de Meselson-Stahl**  
Les cellules bactériennes ont été cultivées pendant plusieurs générations dans un milieu contenant un isotope lourd d'azote (N<sup>15</sup>) puis ont été transférées dans un nouveau milieu contenant l'isotope plus léger normal (N<sup>14</sup>). Par la suite, à divers moments des échantillons de bactéries ont été collectés et leur ADN a été dissous dans une solution de chlorure de césium, qui a été centrifugée rapidement dans une centrifugeuse.

**Figure 17** : Représentation de l'expérience de Meselson et Stahl.

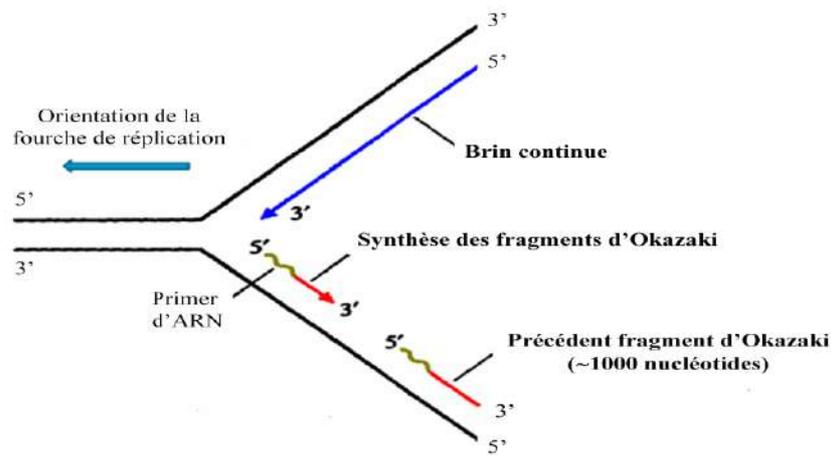
### I.3. Mécanismes de la réplication

La réplication est une réaction asymétrique qui s'effectue uniquement dans le sens 5'→3', de façon **complémentaire**, selon les règles d'appariement : A-T / G-C et un mode antiparallèle.

Ce processus fait intervenir une enzyme de polymérisation nucléotidique qui est la polymérase d'ADN ou **ADN polymérase**. Celle-ci a besoin des composés suivants afin de pouvoir synthétiser une chaîne d'ADN :

- Une matrice d'ADN (ADN parental) ;
- Les quatre désoxyribonucléosides 5'-triphosphate (dATP, dGTP, dTTP, dCTP), substrats de cette enzyme qui sont les désoxyribonucléosides triphosphates en apportant l'énergie nécessaire à la formation de la liaison phosphodiester et intégrés dans le brin sous forme monophosphate ;
- Des ions  $Mg^{2+}$  qui stabilisent les dNTP en les protégeant d'une hydrolyse enzymatique mais également importants pour l'ADN polymérase (Assemblage / dissociation du site actif des complexes de l'ADN polymérase avant et après la réaction chimique d'incorporation de nucléotide) ;
- Une amorce d'environ 10 nucléotides dont le dernier résidu doit être apparié et avoir un groupement hydroxyle libre (3'-OH) ;
- Les enzymes de réplication qui forment un complexe enzymatique dont :
  - Les **hélicases** ou déroulases (une sur chaque brin) séparant progressivement les deux brins d'ADN par rupture des liaisons hydrogènes. Elles sont ATP dépendante.
  - Les **topoisomérases**, qui contrôlent la structure topologique de l'ADN en générant des coupures transitoires dans celui-ci et en catalysant le passage des segments d'ADN à travers ces coupures avant de les refermer. Elles permettent notamment d'ajouter et d'enlever des super-tours dans les molécules d'ADN.
  - La **primase** qui est une ARN polymérase ADN dépendante synthétise des amorces d'une 10<sup>ème</sup> de nucléotides.
  - Les **ADN polymérases** ajoutent des désoxyribonucléotides à l'extrémité 3'-OH de l'amorce.
  - Les **ADN ligases**, enzymes qui catalysent la formation de liaisons phosphodiesters entre 2 chaines polynucléotidiques assurant la liaison des fragments d'Okazaki.

La réplication de l'ADN se fait en trois étapes qui sont : l'activation ou l'initiation de la réplication, l'élongation et la terminaison de la réplication.



**Figure 18** : Fourche et sens de réplication.

### I.3.1. Réplication chez les procaryotes (*E. coli*)

#### 1) Initiation de la réplication

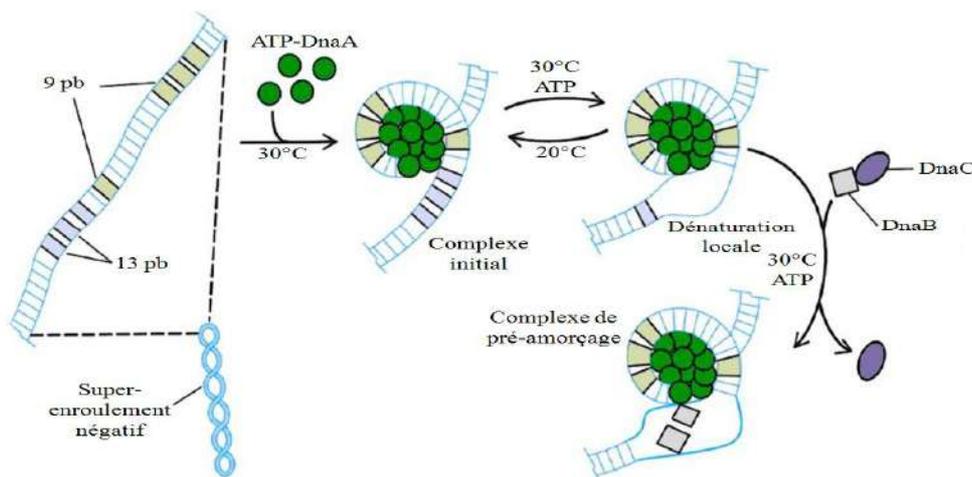
La réplication chez les procaryotes est initiée ou débute au niveau d'une région bien précise qui est l'**origine de réplication** « ORI C » pour *Origin Chromosome*, composée entre 100 et 245pb, dont il en existe qu'une seule en règle générale (chez les *Archaea*, le nombre d'origines de réplication varie entre une et trois selon les espèces).

Les séquences clés sont formées de deux séries de courtes répétitions :

- Trois répétitions d'une séquence de 13 pb « GATCTNTTNTTTT » ;
- Quatre répétitions d'une séquence de 9 pb « TTATCCACA ».

Sachant que N représente n'importe quel autre nucléotide.

La première étape débute par la fixation de la **protéine DnaA** (complexe multimérique capable de reconnaître ces origines) sur l'origine de réplication. Il s'ensuit un enroulement de l'ADN autour de la protéine qui provoque une dénaturation locale des deux brins, au niveau des 3 répétitions consécutives de 13 pb. Une **hélicase** s'y engouffre et sépare les brins dans les deux sens en rompant les liaisons hydrogène (ou "ponts hydrogènes") entre les deux brins de la double hélice d'ADN.



**Figure 19 :** Processus d'action de la protéine DnaA lors de l'initiation de la réplication.

Deux fourches de réplication sont ainsi créées, délimitant l'**œil de réplication**. D'autres protéines se lient à l'ADN simple brin (monocaténaire) ainsi formé et évitent la reformation de la double hélice. Ce sont les **protéines SSB** (*single strand binding*) chez les bactéries et les **protéines RPA** (*replication protein A*) chez les eucaryotes et les archées. De part et d'autre des deux fourches de réplifications, une **gyrase** (topoisomérase de classe II) enlève les super-tours positifs engendrés par la formation de l'œil de réplication.

Chez les procaryotes, le complexe multienzymatique composé de l'hélicase et de la primase est appelé le **primosome**.

#### 2) Elongation

L'**hélicase DnaB** ouvre la double hélice en aval de la polymérase puis synthétise des amorces d'ARN par la **primase**. La fourche de réplication se déplace le long du brin matrice qui est progressivement dénaturé par les hélicases et les brins fils sont synthétisés par la **réplicase** (ADN polymérase III) qui incorpore les nucléotides. La fourche de réplication se déplace dans le

sens  $5' \rightarrow 3'$  sur un brin et  $3' \rightarrow 5'$  sur l'autre. Les ADN polymérase n'incorporent des nucléotides qu'à une extrémité  $3'OH$  libre (synthèse dans le sens  $5' \rightarrow 3'$ ).

On distingue un brin dit avancé et un brin dit retardé :

- Pour le brin avancé, la synthèse de l'ADN se fait dans le sens d'ouverture de la molécule d'ADN, donc dans le sens de propagation de la fourche de réplication. C'est le brin continu.
- Pour le brin retardé ou discontinu, la synthèse d'ADN se fait dans le sens contraire de l'ouverture de la molécule d'ADN : l'ADN polymérase n'exerce son activité polymérasique que dans le sens  $5' \rightarrow 3'$  de la chaîne en cours de synthèse, or sur ce brin d'ADN, le sens de propagation de la fourche de réplication est  $3' \rightarrow 5'$ . La synthèse de ce brin est donc discontinue. Elle va s'effectuer de manière **rétrograde**, à partir de multiples fragments d'ARN. On obtient donc de multiples fragments d'ADN (fragments de 1000 à 2000 nt), appelés **fragments d'Okazaki**.

Il y'a donc mise en jeu de deux molécules d'ADN polymérase III qui travaille en même temps au niveau de la fourche de réplication : une qui synthétise de manière continue le brin avancé et l'autre qui synthétise de manière discontinue le brin retardé. Le brin d'ADN parental, qui sert de matrice au brin d'ADN retardé, forme une boucle pour repositionner l'amorce de façon à ce que l'allongement du fragment d'Okazaki par l'ADN polymérase se fasse dans le sens  $5' \rightarrow 3'$  mais aussi dans le sens de propagation de la fourche de réplication.

#### Remarque

La processivité est la capacité de polymériser sur une longue distance. L'ADN polymérase III est naturellement peu processive : elle a tendance à synthétiser spontanément de l'ADN sur une courte distance et à se dissocier aussi spontanément de l'ADN parental. Cela est en adéquation avec la synthèse du brin discontinu : synthèse d'un fragment, puis détachement et synthèse d'un autre fragment. Mais sur le brin continu, l'ADN polymérase III doit synthétiser sur une longue distance. L'ADN pol III est donc maintenu au brin d'ADN parental par un collier coulissant (clamp).

Pour *E. Coli*, c'est la sous-unité  $\beta$  du complexe holoenzyme qui joue le rôle de clamp. Cela permet à l'ADN pol III de rester associé au brin parental.

### 2.1. Types d'ADN polymérase intervenant chez les procaryotes

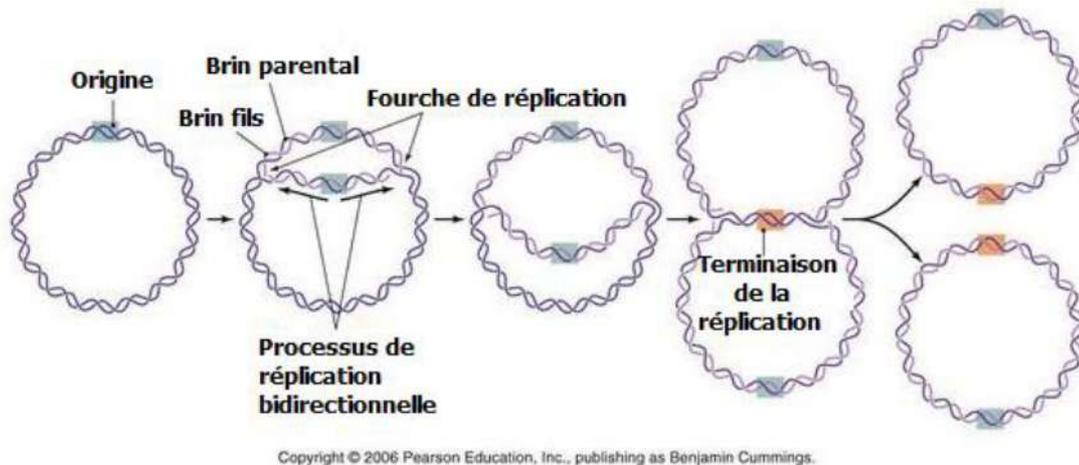
- **Pol I** : Impliquée dans la réparation et la réplication de l'ADN. Elle possède une activité polymérase  $5' \rightarrow 3'$ , une activité exonucléase  $3' \rightarrow 5'$  pour la relecture, et enfin une activité exonucléase  $5' \rightarrow 3'$  qui lui permet d'éliminer les amorces d'ARN des fragments d'Okazaki pendant la réplication.  
Son activité polymérase lui permet de synthétiser de l'ADN destiné à être immédiatement suturé par une ADN ligase.  
La Pol I est peu processive et se dissocie de la matrice après l'ajout d'une vingtaine de nucléotides.
- **Pol II** : Impliquée dans la réplication de l'ADN endommagé, elle possède une activité polymérase  $5' \rightarrow 3'$  et une activité exonucléase  $3' \rightarrow 5'$ . La Pol II est peu processive.
- **Pol III** : C'est la principale polymérase, qui intervient dans l'élongation de la chaîne d'ADN lors de la réplication au niveau du brin avancé et de la synthèse des fragments d'Okazaki.

### 3) Terminaison

Cette phase correspond à l'arrêt de la réplication lorsque deux fourches de réplication se rencontrent à 180° d'ORIC au niveau des **sites de terminaison** (Ter) de la réplication. Il y a : terA, terD, terB et terC qui freinent les fourches de réplication.

Ces sites sont situés à l'opposé de l'origine de réplication reconnue par la **protéine Tus** (*Terminator Utilisation Substance*). Le complexe **Ter-Tus** bloque les fourches de réplication, mettant fin à la réplication. Lorsque la réplication d'un chromosome circulaire est terminée, les deux 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).



Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings.

Figure 20 : Processus de terminaison de la réplication procaryote.

#### I.3.2. Réplication chez les eucaryotes

La réplication chez les eucaryotes se déroule également en 3 étapes et elle a lieu au cours de la phase S du cycle cellulaire. Elle est en grande majorité semblable à celle de la bactérie *E. coli* cependant avec quelques spécificités.

##### 1) Initiation

La réplication débute au niveau des **origines de réplication**. Ce sont des séquences d'initiation de réplication que l'on appelle « **ARS** » pour *Autonom Replication System*. Chez les eucaryotes à l'inverse des procaryotes, il y'a plusieurs origines de réplication entre 20 000 à 100 000 sites par cellule. Elle va donc commencer simultanément en plusieurs milliers de points d'un chromosome appelé « réplicon », ce qui est nécessaire vue la taille du génome.

Ces séquences ARS sont très riches en appariements **AT** qui sont beaucoup plus facile à rompre que les appariements **GC**, puisqu'ils ne sont formés que de deux liaisons hydrogènes. Ces origines de réplifications doivent se situer dans des zones d'euchromatine, c'est-à-dire des zones où l'ADN est peu condensé, là où la double hélice est accessible.

L'**antigène T** reconnaît ce site d'initiation, s'y fixe et entraîne l'ouverture de la double hélice d'ADN par son activité **hélicase** ; la **protéine RPA** (*Replication Protein A*) se fixe aux régions d'ADN simple-brin pour prévenir leur réassociation.

L'**ADN polymérase  $\alpha$**  (Pol  $\alpha$ ) synthétise l'amorce ARN-ADN de chaque brin continu et de chaque brin discontinu, grâce à son activité **primase**, suivi d'un fragment d'ADN de 20 à 30 nucléotides grâce à son activité ADN polymérase.

## 2) Elongation

La protéine **RF-C** (*Replication factor C*) reconnaît le complexe matrice-amorce et recrute la protéine **PCNA** (*Proliferating Cell Nuclear Antigen*) qui provoque le remplacement de l'**ADN polymérase  $\alpha$**  par l'**ADN polymérase  $\delta$**  ou l'**ADN polymérase  $\epsilon$** , cette dernière reprend la synthèse pendant que l'antigène T continu à fonctionner comme hélicase et que la topoisomérase I relâche la tension devant la fourche de réplication.

- Le brin continu est synthétisé sans interruption par l'ADN polymérase  $\delta$  ou l'ADN polymérase  $\epsilon$  (ADN pol  $\delta$  participe également à la réparation de l'ADN).
- Lorsque sur le **brin discontinu** l'ADN polymérase  $\delta$  (ou  $\epsilon$ ) s'approche de l'amorce d'ARN-ADN précédemment synthétisée, cette dernière est éliminée par une **RNase H1** (une endonucléase qui détache la séquence des ribonucléotides sauf celui lié au premier désoxynucléotide qui est éliminé par l'exonucléase **FEN-1** "*Flap endonucléase*").

L'**ADN polymérase  $\delta$**  engagé dans la synthèse du fragment d'Okasaki comble le vide laissé par le départ de l'amorce. L'**ADN ligase I** ligature enfin l'extrémité « 5'-P » de la séquence de l'amorce du fragment d'Okasaki en amont à l'extrémité 3'-OH du fragment d'Okasaki en aval et ainsi de suite.

Chez les eucaryotes, les fourches de réplication se déplacent 10 fois plus lentement que chez les procaryotes.

### 2.1. Quelques caractéristiques des ADN polymérases des eucaryotes

On connaît au moins six ADN polymérases chez les eucaryotes qui ont des rôles ou des fonctions bien distinctes et qui sont :

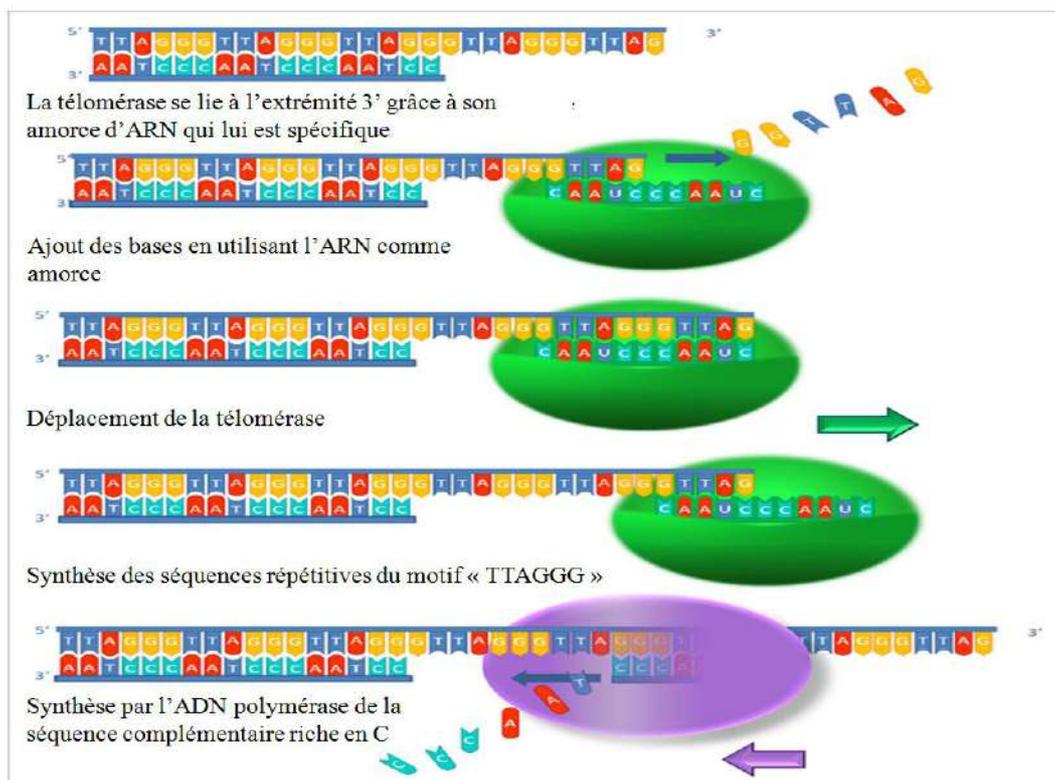
- **La polymérase alpha ( $\alpha$ ) / primase** : Cette polymérase est impliquée dans l'initiation de la réplication ("priming"). Elle possède 04 sous-unités, deux ont une activité ADN polymérase et les deux autres ont une activité primase, et dépourvu d'une activité exonucléasique 3'  $\rightarrow$  5' ;
- **La polymérase bêta ( $\beta$ )** : Elle semble impliquée dans la réparation de courts fragments d'ADN. Elle ne possède pas de fonction exonucléasique. Elle correspond à la Pol II bactérienne ;
- **La polymérase delta ( $\delta$ )** : Enzyme principale de la réplication du brin précoce et des fragments d'Okasaki, prend le relais de la polymérase  $\alpha$ . Elle possède aussi une activité exonucléasique 3'  $\rightarrow$  5' intervenant dans la correction des erreurs et dans des processus de réparation ;
- **La polymérase gamma ( $\gamma$ )** : Cette polymérase est à localisation mitochondriale, bien que codée par un gène du noyau cellulaire. Sa fonction est donc la réplication de l'ADN mitochondrial et possède une activité exonucléasique 3'  $\rightarrow$  5' ;

- **La polymérase epsilon ( $\epsilon$ )** : Possède la même fonction que l'ADN polymérase  $\delta$  mais plus processive que ces dernières. Elle possède une activité polymérase  $5' \rightarrow 3'$  et une activité exonucléase  $3' \rightarrow 5'$ . Elle intervient dans la réplication et la réparation de l'ADN. Elle lit dans le sens  $3' \rightarrow 5'$  et synthétise dans le sens  $5' \rightarrow 3'$  la zone du télomère qui ne peut être synthétisée par la Pol  $\delta$  ;
- **La polymérase zeta ( $\zeta$ )** : Elle semble impliquée dans la réparation de l'ADN.

### 3) Terminaison

Chez les eucaryotes, le problème majeur lors de la réplication de l'ADN est l'élimination potentielle de l'amorce d'ARN la plus externe ce qui pourrait entraîner un raccourcissement de l'ADN à chaque cycle de réplication. Ce processus à l'extrémité des chromosomes eucaryotes (place où les amorces d'ARN ont été éliminées) fait intervenir une enzyme particulière appelée « **téломérase** ». Elle ajoute des séquences spécifiques en 3' d'un brin d'ADN afin d'éviter ce problème de raccourcissement.

La télomérase va se positionner à l'extrémité 3' du brin d'ADN grâce à une matrice d'ARN qui lui est propre puis synthétise d'un motif d'ADN « TTAGGG » face à la matrice d'ARN répétée quelques centaines de fois avant le 3'OH final. Celle-ci va donc constituer le **téломère**. Pour synthétiser le brin télomère complémentaire riche en C, une **primase** interviendrait avec synthèse d'une amorce d'ARN par copie de l'extrémité 3'. Puis l'ADN polymérase allongerait la chaîne à partir de l'amorce d'ARN. Finalement, une **ligase** assurerait la soudure finale. Après hydrolyse de l'amorce d'ARN, les télomères riches en C sont légèrement plus courts que les télomères riches en G. Quand la synthèse des 2 copies est achevée, la **topoisomérase II** décatène les 2 chromosomes.



**Figure 21** : Mode d'action de la télomérase.

### I.3.3. Réplication de l'ADN mitochondrial

La réplication de l'ADN mitochondrial circulaire n'est pas limitée à la phase S du cycle cellulaire. Elle utilise deux origines de réplication et fait intervenir une structure intermédiaire à 3 brins.

L'ADN polymérase **Gamma** est responsable de la réplication de l'ADN mitochondrial dont la réplication est indépendante de l'ADN nucléaire. Cette réplication de l'ADNm est donc contrôlée par des gènes nucléaires et fait intervenir de nombreux facteurs protéiques.

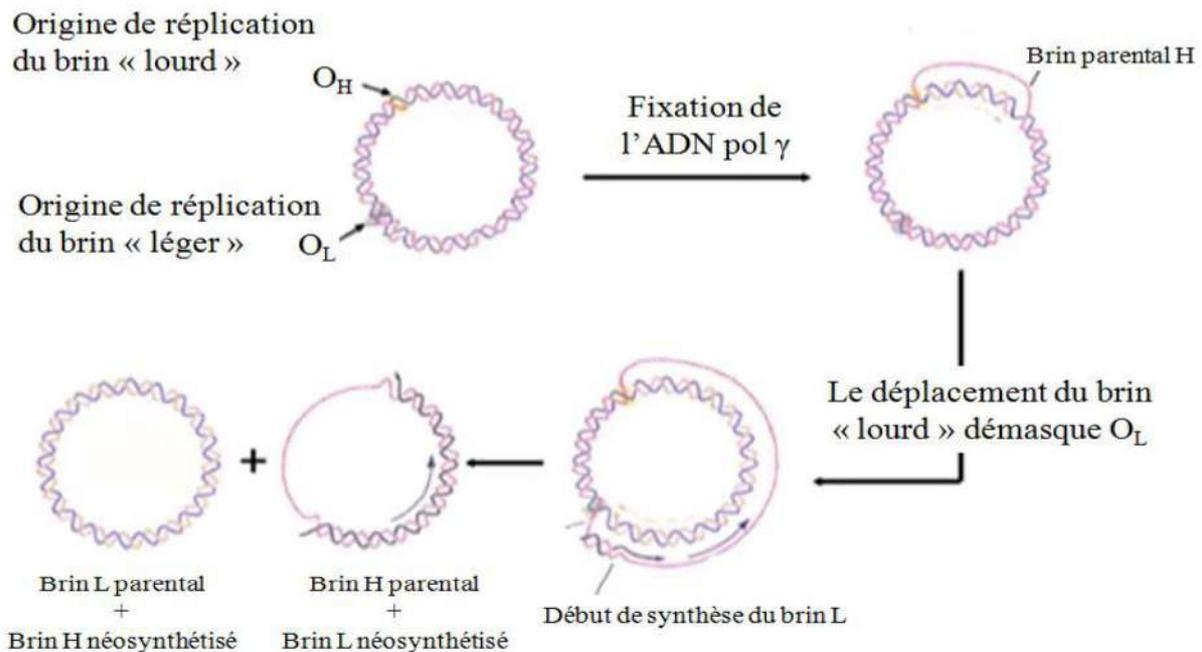


Figure 22 : Réplication de l'ADN mitochondrial.

### I.3.4. Réplication du matériel génétique des rétrovirus

Les rétrovirus sont des virus à ARN monocaténaire dont le génome passe au cours du cycle viral, par une intégration sous la forme d'ADN, dans le génome de la cellule hôte. Le plus connu de ces rétrovirus est le virus de l'immunodéficience humaine (VIH) ou **virus du SIDA**.

La réplication du matériel génétique des rétrovirus permet ainsi le passage d'un ARN simple brin à un ADN double brin et ceci grâce à trois principales enzymes :

- Une **ADN polymérase ARN dépendante**, qui n'est autre que la **transcriptase inverse** responsable de la transcription reverse de l'ARN virale. Elle a la caractéristique de synthétiser dans la direction 5' vers 3', et nécessite une amorce, une matrice, ainsi que les désoxyribonucléosides triphosphate ; elle ne présente par contre pas d'activité exo-nucléasique 3' vers 5'.
- Une **RNase** responsable de la lyse de l'ARN virale.
- Une **ADN polymérase ADN dépendante** responsable de la formation de l'ADN double brin qui sera intégré dans le génome de la cellule hôte.

## Mécanismes moléculaires

### (Mutation/Altération de la molécule d'ADN et mécanismes de réparation de l'ADN)

#### *Objectifs spécifiques*

Au terme de ce cours qui est la réparation de l'ADN, vous devez être capable de :

- Connaître ce que signifie la réparation de l'ADN ;
- Identifier les différents types de mutations que subie la molécule d'ADN ;
- Connaître les différents mécanismes de réparation de l'ADN ;
- Pouvoir décrire chaque mutation, ces causes et son mode de réparation.

#### II.1. Introduction

Le système de dédoublement de l'information génétique n'est pas parfait, il peut donc y avoir des anomalies et d'autre part il peut être altéré par des agressions chimiques ou physiques externes. Qui dit réplication sous-entend donc également mécanismes de réparation. On ne peut parler de réplication d'ADN sans parler de réparation ni sans parler des lésions ou dommages que subit l'ADN.

#### II.2. Mutations et dommages sur l'ADN

Dans les cellules, l'acide désoxyribonucléique (ADN) est continuellement soumis à des activités métaboliques normales et à des facteurs environnementaux portant atteinte à son intégrité. Les mutations sont des modifications transmissibles (héritables) de la séquence du génome d'un organisme. Ces événements modifient donc la séquence d'ADN et ainsi le phénotype<sup>1</sup> de l'individu.

Les mutations peuvent être transmises à la descendance si elles se réalisent dans des génomes de cellules germinales ou précurseurs de cellules germinales. Les mutations de cellules somatiques entraînent des modifications au sein même de l'individu. Ces mutations permettent l'évolution de l'espèce, mais malheureusement elles sont responsables d'un grand nombre de maladies. On estime entre mille et plus d'un million le nombre de lésions par cellule et par jour. Beaucoup de ces lésions provoquent de tels dommages que la cellule elle-même ne pourrait se reproduire ou donnerait naissance à des cellules-filles non viables si n'intervenaient les différents processus de réparation.

##### II.2.1. Comment apparaissent ces mutations?

On retrouve généralement deux cas pendant lesquels peuvent apparaître des mutations ou lésions qui sont :

###### II.2.1.1. Erreurs commises lors de la réplication

Ce sont des altérations **non corrigées** par le système de correction d'édition de l'ADN polymérase. Ces mutations peuvent conduire soit à une **délétion**, soit à une **substitution**, soit à une **insertion**.

- **Mutation par substitution**

Cela correspond à une mauvaise incorporation de nucléotide sur le brin fils. Une base est alors remplacée par une autre.

On distingue deux types de mutation par substitution :

<sup>1</sup> Ensemble des traits observables d'un organisme.

- Substitution par **transition** : base pyrimidique substituée par une base pyrimidique ou base purique substituée par une base purique. Ainsi une paire de bases A-T est remplacée par une paire de bases G-C.
- Substitution par **transversion** : base purique substituée par une base pyrimidique ou base pyrimidique substituée par une base purique.

Les mutations entraînant le changement d'acides aminés sont des **mutations faux-sens**. Lorsque la mutation n'entraîne pas de modification d'acides aminés, et ceci dû à la redondance du code génétique en position 3, on parle de **mutation silencieuse**. Lorsque la substitution entraîne l'apparition d'un codon stop (UAA, UGA et UAG), la mutation est une **mutation non-sens**.

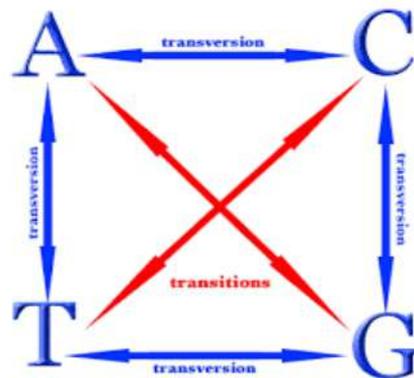


Figure 23 : Nomenclature des substitutions de base.

- **Mutations par délétion**

Il s'agit d'un **oubli** d'incorporation d'un nucléotide par l'ADN polymérase. Au bout de deux cycles successifs, une séquence est totalement différente par rapport aux séquences parentales (25% ADN mutant).

Lorsque le brin fils va servir d'ADN matrice, cela entraîne un **décalage** du cadre de lecture.

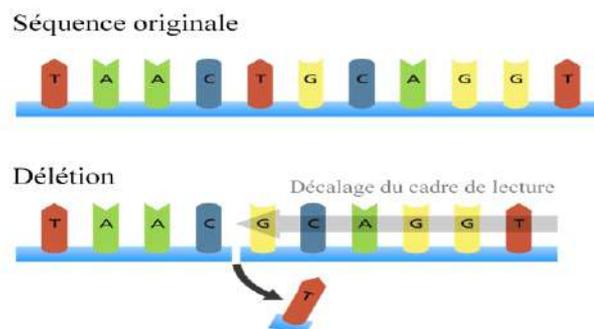


Figure 24 : Mutation de l'ADN par délétion de base.

- **Mutations par insertion**

Il s'agit d'une **introduction** d'un nucléotide en trop par l'ADN polymérase. Si l'addition de nucléotides n'est pas un multiple de 3, il y aura décalage du cadre de lecture (*frame-shift*).

En effet si l'addition ou la délétion est un multiple de 3, il y aura addition ou délétion d'acides aminés au niveau de la protéine finale. Le décalage du cadre de lecture peut entraîner des séquences d'acides aminés totalement différentes et des apparitions de codons stop.

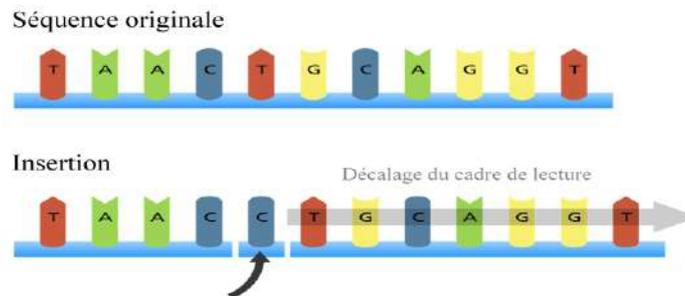


Figure 25 : Mutation par insertion de bases.

### II.2.1.2. Altération de l'ADN survenant en dehors de la réplication

En dehors de la réplication, on peut avoir des lésions **spontanées** ou des altérations de l'ADN, en conséquence à des **agents mutagènes** (physiques ou chimiques).

Les agents mutagènes physiques correspondent aux rayonnements X ou  $\gamma$ , aux rayonnements UV et à la chaleur.

Les agents mutagènes chimiques sont essentiellement sous la forme de radicaux super-oxydes ( $O_2^-$ ) qui peuvent oxyder de manière non catalytique les molécules biologiques et les engager dans des réactions chimiques incontrôlées.

#### II.2.1.2.1. Lésions spontanées

Les mutations **spontanées** sont des événements à **grande fréquence** qui aboutissent à des lésions majeures de l'ADN s'il n'y a pas de réparation.

- **Dépuration spontanée**

C'est la perte de bases par interruption (hydrolyse) d'une liaison  $\beta$ -N-glycosidique entre la base purique (A ou G) et le désoxyribose, donc perte d'un résidu guanine ou adénine de l'ADN. Cela conduit donc à la « Création d'un **site apurique** (ou apurinique) ».

Ces pertes sont spontanées à pH acide par rupture de la liaison N-glycosidique. Suite à ces pertes d'informations la polymérase ne sait pas quelle base incorporer, il y a ainsi formation d'un site AP. La dépuration est un **phénomène très fréquent** de l'ordre de 5 à 10 000 dépurations par jour et par cellule chez les mammifères. Il est donc nécessaire d'avoir un système de réparation efficace pour réparer ce type de mutation.

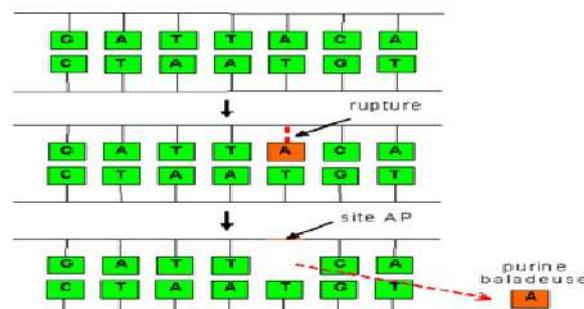


Figure 26 : Phénomène de dépuration de l'ADN et création d'un site apurique (AP).

**Remarque**

Les sites apyrimidiques (clivage entre base pyrimidique et un désoxyribose) sont beaucoup **moins fréquents** et beaucoup plus lents. Tous ces sites apuriques ou apyrimidiques sont dits abasiques ou site AP.

- **Désamination spontanée des bases naturelles**

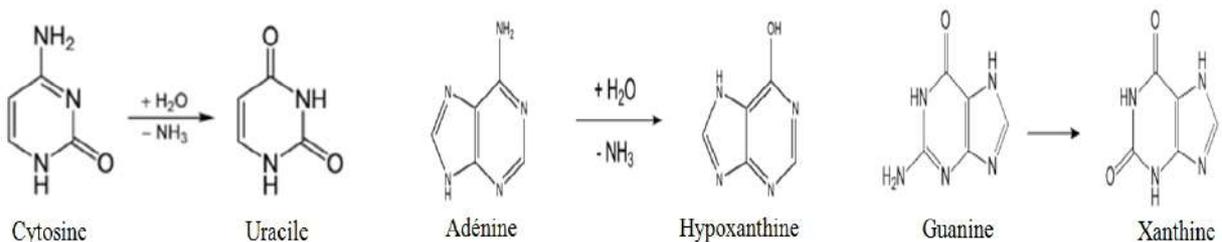
Aussi appelées désaminations oxydatives, elles donnent naissance à des bases modifiées (bases désaminées due à la perte du groupement amine), qui établissent des liaisons particulières avec les autres bases et entraînent donc des mésappariements. Les désaminations sont dues à des excès de chaleur. Une mutation apparaît de manière stable après 2 réplifications.

La désamination la plus fréquente est celle de la **cytosine** en **uracile**, environ 100 bases par jour et par cellule subissent cette désamination. L'**uracile** est ensuite capable de **s'apparier** à une **adénine**, donc sur le brin fils, il y aura incorporation d'une guanine. Après deux cycles cellulaires, il y aura 25% de mutation.

La désamination spontanée de l'**adénine** donne de l'**hypoxanthine**. L'hypoxanthine, s'apparie de façon sélective avec la cytosine au lieu de la thymine. Cela mène à une mutation transitoire post-répllicative, où la paire de bases originelle A-T se transforme en paire de bases G-C.

La désamination de la **guanine** donne de la **Xanthine** qui reste apparié à une cytosine.

La **thymine** est la seule base qui ne subit **pas de désamination**.

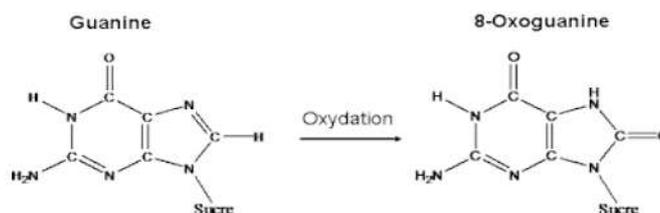


**Figure 27** : Désamination de différentes bases azotées.

- **Bases altérées par oxydation**

La production de molécules chimiques très réactives entraîne l'oxydation des bases de l'ADN. Ce sont les **radicaux superoxydés** ( $O_2^-$ ), les **radicaux hydroxylés** ( $OH^\cdot$ ) et les **peroxyde d'hydrogène** ( $H_2O_2$ ) qui sont créés normalement par le système aérobie (il y en a beaucoup dans la mitochondrie où l'on retrouve la chaîne respiratoire).

Cette oxydation est souvent à l'origine de maladies génétiques. L'oxydation de la guanine en **oxoguanine** permet un appariement avec l'adénine (GC devient TA après deux réplifications).



**Figure 28** : Oxydation de la guanine en oxoguanine.

### II.2.1.2.2. Mutation induites par des agents mutagènes

Plusieurs facteurs sont dits agents mutagènes car ils affectent la molécule d'ADN.

- **Incorporation d'analogues de bases dans la molécule d'ADN**

Les analogues de bases sont des composés chimiques qui ressemblent suffisamment aux bases azotées normales pour parfois être incorporées à la place de celle-ci lors de la réplication au niveau du brin fils.

Ces analogues ont des propriétés d'appariement différentes, ce qui entraîne des mutations après réplication.

#### Exemples

- Le 5-Bromo-uracile (5BU) : analogue de la thymine
- Le 2-Amino-purine (2AP) : analogue de l'adénine

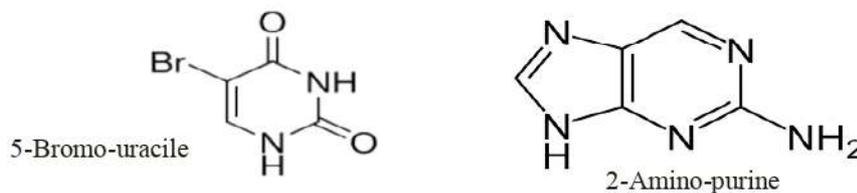


Figure 29 : Analogues de bases azotées.

- **Modification de bases par les agents alkylants**

Ces agents altèrent les bases en ajoutant des **groupes éthyle** ou **méthyle** sur l'atome d'oxygène de la **guanine** ou de la **thymine**. Les bases modifiées ont des propriétés d'appariement différentes : par exemple, la 6-O-méthylguanosine s'apparie avec T (GC → AT).

#### Mutagenèse par Ethyl méthane Sulfonate (EMS)

#### Exemples d'agents alkylants

- Ethylméthane sulfate
- Nitrosoguanidine
- Diméthylsulfate

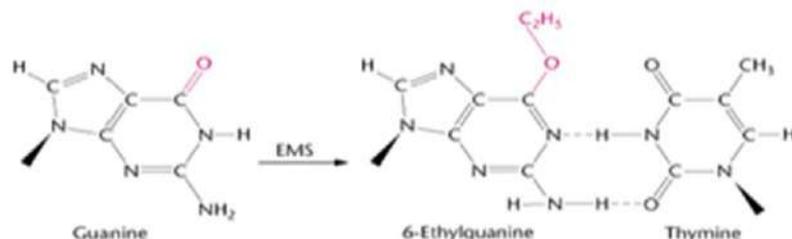


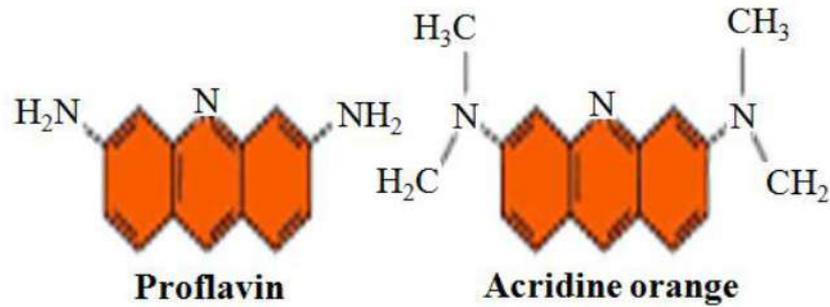
Figure 30 : Modification de la guanine par le 6-ethylguanine.

- **Insertions ou délétions de bases par des agents intercalants**

Les agents intercalants ont des structures planes à 3 cycles s'apparentant à des paires de bases et qui vont ainsi s'insérer entre les paires de bases de la double hélice. Elles induisent une distorsion causant à la réplication suivante une insertion de bases (fréquente) ou une délétion.

#### Exemples d'agents intercalants

- L'acridine orange
- La proflavine



**Figure 31** : Agents intercalants.

- **Lésions de bases par des agents mutagènes**

Ce sont des lésions très importantes, qui entraînent en l'absence de réparation un blocage de la réplication, donc de la division cellulaire.

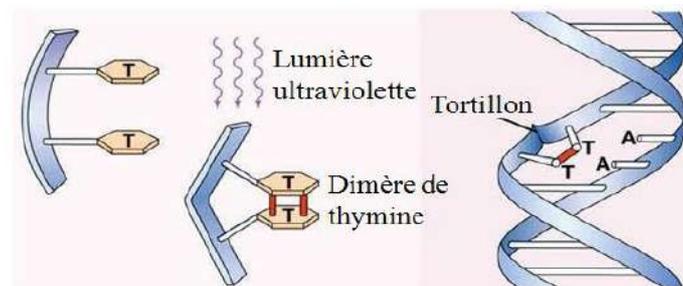
→ **Lésions par la lumière UV**

Cela aboutit à la création de liaisons covalentes sur deux pyrimidines adjacentes appartenant au même brin d'ADN : formation de **dimère de pyrimidine**.

Ces liaisons covalentes s'établissent **souvent entre deux thymines** adjacentes ce qui aboutit à la création d'un **dimère de thymines**. Cela entraîne une distorsion de la molécule d'ADN ce qui bloque la réplication, s'il n'y a pas de réparation.

Les UV solaires sont capables d'induire ce types de lésions, au niveau des cellules de la peau, et s'il n'y a pas de réparation, comme chez certaines personnes prédisposées, il y aura le développement du cancer de la peau : les mélanomes.

**Dimère de pyrimidine**



**Figure 32** : Dimère de thymine.

→ **Lésions par radiations ionisantes**

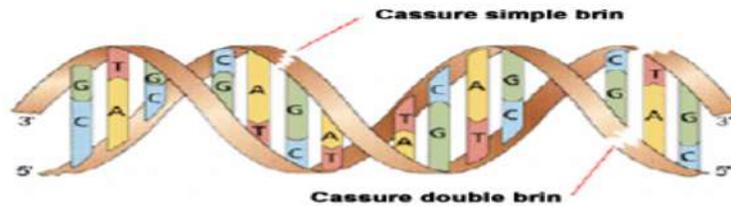
Certains agents mutagènes (comme les radiations ionisantes) peuvent aboutir à des **cassures de brins**, qui touchent soit un brin (cassure simple brin), soit les deux brins (cassure double brin).

**Les radiations ionisantes** capables d'induire ce type de lésions sont les rayons cosmiques, la radioactivité et les rayons X (utilisés à des fins thérapeutiques). Les mécanismes d'action de lésions par ces radiations sont multiples :

- *Action directe* : cassure de brins puis **création de sites AP** (Sites Apurique/Apyrimidique) ;
- *Action indirecte* : met en jeu la création de **réactifs oxydatifs** qui vont créer les lésions oxydantes.

Les systèmes de réparations sont normalement capables de réparer ces lésions.

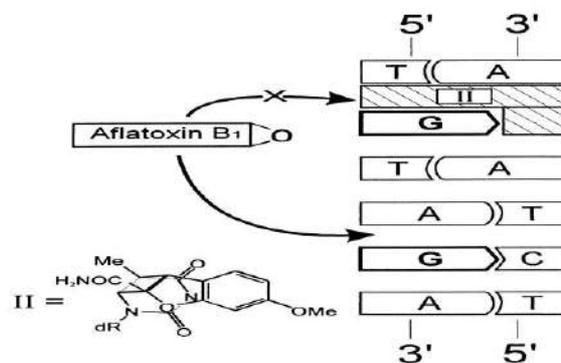
La radiothérapie des cancers utilise ce principe : les radiations sont localisées et entraînent des cassures non réparables, ce qui aboutit à la mort des cellules irradiées : elles rentrent en apoptose.



**Figure 33 :** Lésions d'ADN par radiations ionisantes.

#### → *Aflatoxine B1*

Il s'agit d'une mycotoxine, élaborée par les champignons, qui est capable de créer des **sites AP** (Sites Apurique/Apyrimidique) au niveau de la molécule d'ADN exposée.

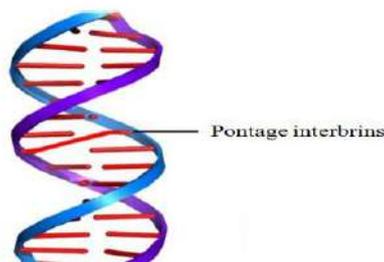


**Figure 34 :** Action de l'aflatoxine B1 sur la molécule d'ADN.

- **Pontage de brin**

Il s'agit de la distorsion de la molécule d'ADN, à cause de la formation d'une liaison covalente entre les deux bases appartenant à deux brins différents.

Plusieurs espèces chimiques sont capables d'introduire ce type de lésions. Par exemple, la **mitomycine** qui est un antibiotique responsable du pontage de brins.



**Figure 35 :** Formation de pontage d'un brin d'ADN.

### II.3. Réparation des lésions de l'ADN

La **réparation de l'ADN** est un ensemble de mécanismes de sauvegarde essentiel au cours de l'évolution, par lesquels une cellule identifie et corrige les dommages causés aux molécules d'ADN qui codent son génome.

### II.3.1. Mécanismes de réparation

La réparation de l'ADN est mise en œuvre *via* une grande variété de mécanismes adaptés à chaque type de lésion :

- Réparation directe ;
- Réparation des mésappariements causés par le processus de réplication ;
- Réparation par excision et échange de base ;
- Réparation par excision et échange de nucléotides ;
- Réparation des cassures double brin par recombinaison homologue ou non homologue.

Chacun de ces mécanismes met en jeu de nombreuses protéines au sein de complexes moléculaires fonctionnels.

L'impossibilité de réparer des lésions de l'ADN conduit normalement à la mort cellulaire, mais des altérations des mécanismes de réparation peuvent favoriser l'instabilité génétique et participer ainsi à l'oncogenèse.

#### II.3.1.1. Correction des erreurs d'appariement lors de la réplication

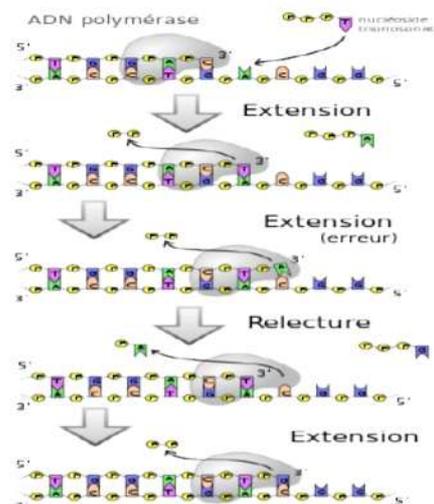
Les premières fonctions de correction sont celles qui s'exercent pendant la réplication :

- La **fonction exonucléasique** de l'ADN polymérase, en cas de reconnaissance d'un mésappariement.
- Le **système MMR**.

##### II.3.1.1.1. Correction immédiate : fonction d'édition de l'ADN polymerases

Ce mécanisme est aussi bien présent chez les procaryotes que chez et les eucaryotes et se déroule en même temps que la réplication de l'ADN.

En plus de sa fonction de synthèse (polymérisation), l'**ADN polymérase** possède aussi une fonction **exonucléasique** qui lui permet de corriger au fur et à mesure de sa polymérisation les mésappariements et diminuer le taux d'erreurs de  $2 \log (10^5 \text{ à } 10^7)$ .



**Figure 36** : Fonction d'édition et de correction immédiate de l'ADN polymérase.

##### II.3.1.1.2. Réparation des mésappariements par le système methyl mismatch repair

Ce mécanisme désigne le système de reconnaissance et de réparation des mésappariements de l'ADN. Ce système constitue la deuxième ligne de défense contre les

erreurs de réplication après celle de relecture du brin nouvellement synthétisé par l'ADN polymérase elle-même. Ce mécanisme est partagé par les eucaryotes et le procaryote dû à sa conservation au cours de l'évolution.

Ce mécanisme vise à corriger les erreurs commises lors de la synthèse de l'ADN par les polymérases  $\delta$  et  $\epsilon$ , erreurs non corrigées par son activité d'édition. Grâce à ce système, on aboutit à un taux d'erreur de 1 pour  $10^9$  bases.

Chez les **procaryotes**, le système de réparation repère la méthylation des **adénines** des séquences **GATC** et fait intervenir les enzymes **Mut**.

Chez les **eucaryotes**, le système repère la méthylation des **cytosines** des séquences **CG** et fait intervenir les enzymes **hMSH**, **hMLH** et **hPMS**.

La reconnaissance du brin matrice par rapport au brin nouvellement synthétisé est basée sur la présence de méthylations spécifiques dans le premier, ce qui permet à la cellule de reconnaître la copie de l'original et d'éviter de modifier ce dernier.

- **Chez *E. coli***

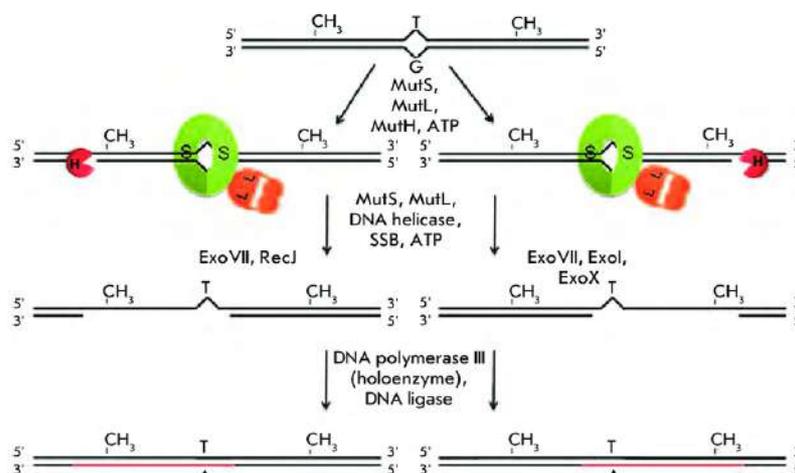
Le système enzymatique de réparation est **multi-protéique**. Il est capable de reconnaître un mésappariement et une séquence GATC située dans l'environnement du mésappariement (jusqu'à 1000 pb), afin de vérifier le brin méthylé.

Ce complexe multi-enzymatique est **codé** par les **gènes *mut* très conservés** et constitué des **enzymes Mut**. Ce nom vient du fait que leur dysfonctionnement engendre l'apparition de nombreuses mutations.

Il peut se positionner sur le mésappariement et détecter le brin méthylé. Il a une **activité endonucléasique** : il peut cliver les liaisons phosphodiester à l'intérieur d'une chaîne.

Ce complexe excise le fragment d'ADN simple brin qui contient le mésappariement (donc sur le brin non méthylé). Cela entraîne la formation d'une lacune qui doit être comblée. L'**ADN pol III** se positionne alors en 3'-OH en synthétisant l'ADN complémentaire et antiparallèle. La dernière liaison phosphodiester est effectuée par l'**ADN ligase**.

Ce système doit agir dans le laps de temps où la méthylation du brin fils n'est pas encore effectuée. Une fois ce laps de temps écoulé, les ADN méthylase rentrent en jeu et méthylent en miroir le brin fils. Le complexe ne peut alors plus reconnaître le brin fils.

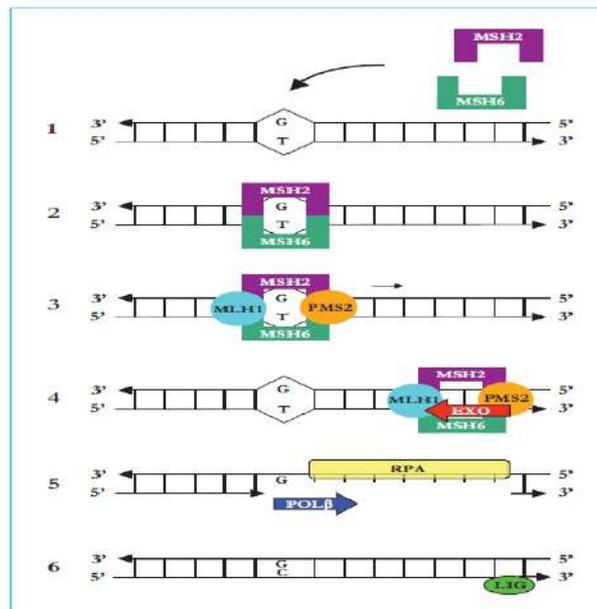


**Figure 37 :** Mismatch repair chez *E. coli*.

- *Chez l'homme*

Le **complexe multi-enzymatique** a les mêmes fonctions, mais il est basé sur les méthylations des cytosines des séquences répétées GC. Les enzymes qui interviennent ne sont pas les enzymes Mut mais les **enzymes hMSH** (*MutS homolog 2*), **hMLH** (*MutL homolog 1*) et **hPMS** (*postmeiotic segregation increased*), (h pour humain).

Chez l'homme, l'altération des gènes correspondants est responsable de l'apparition de cancers précoces avec une forte prévalence.



**Figure 38** : Réparation des mésappariements.

### II.3.1.2. Correction des autres mutations de l'ADN

Il existe d'autres altérations de l'ADN qui peuvent conduire à l'arrêt du cycle cellulaire, afin de laisser le temps à la cellule de réparer les dommages, par soucis de transmettre aux cellules filles un patrimoine le plus intègre possible.

#### II.3.1.2.1. Réparation direct par retour à l'état antérieur

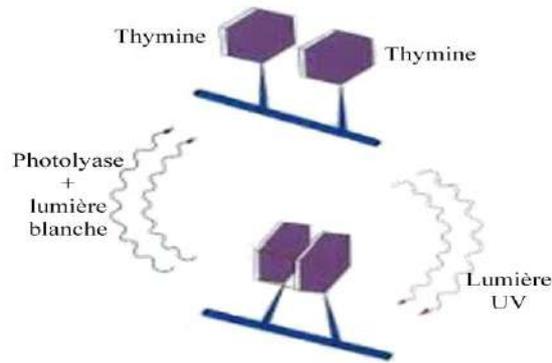
Ce mécanisme est associé à des altérations spécifiques où l'objectif général est d' « inverser » le mécanisme qui a conduit à l'altération de l'ADN. Voici les deux exemples les plus courants :

- *Mécanisme faisant intervenir la photolyase*

La **photolyase** est une enzyme qui permet de réparer des **lésions** induite par la **lumière UV** en particulier les dimères de thymines ou photodimères. Elle agit sous l'action d'une **lumière visible**. Quand elle est activée, elle excite un cofacteur qui est la flavine réduite (FAD) ce qui permet à l'enzyme de se lier aux dimères de thymine et de transférer un électron afin d'entraîner la scission (scinder) du dimère et faire disparaître la liaison et revenir à deux thymines adjacentes.

La photolyase n'existe que chez certains organismes : elle est présente chez les bactéries, ou chez certains eucaryotes inférieurs (levures, drosophile, végétaux,...).

*!\\ Elle n'est pas présente chez l'Homme!\\.*



**Figure 39 :** Réparation par photolyase d'un dimère de thymine.

- **Mécanisme faisant intervenir les alkyltransférases**

Les alkyltransférases sont des enzymes intervenant dans l'inversion directe des lésions causées à l'ADN. Ils vont avoir pour rôle de réparer les **liaisons** induites par les **agents alkylants**. Ces enzymes éliminent certains groupes alkyles qui ont été ajoutés aux positions O-6 de la guanine par des substances mutagènes telles que la nitrosoguanidine et l'éthyl-méthanesulfonate.

La formation de la O6-méthylguanine est une mutation qui peut se transmettre. La réparation se fait grâce à l'enzyme **O6-méthylguanine méthyltransférase (MGMT)**. Elle se lie et transfère le groupement méthyle au niveau d'une cystéine interne situé sur cette enzyme, où est localisé le site actif de l'enzyme. On a donc un retour à la guanine, et une dégradation du groupement méthyle.

Cependant, ce transfert inactive l'enzyme, de sorte que ce système de réparation peut être rapidement saturé si le niveau d'alkylation est suffisamment élevé.

### II.3.1.2.2. Réparation par excision

C'est un mécanisme de réparation **simple brin**. On peut distinguer deux types de mécanismes [*nucleotide excision repair* (NER) ou *base excision repair* (BER)] en fonction de l'excision de la partie altérée. Ce mécanisme agit sur les lésions présentes sur un seul brin. C'est un mécanisme multi-étapes :

- a- **Reconnaissance** de la lésion ;
- b- **Excision** de la partie altérée : soit sur une base (BER) soit sur plusieurs nucléotides (NER) ;
- c- **Réparation** par réplication pour combler la lacune (ADN polymérase et ligase).

- **Réparation par excision de base (BER)**

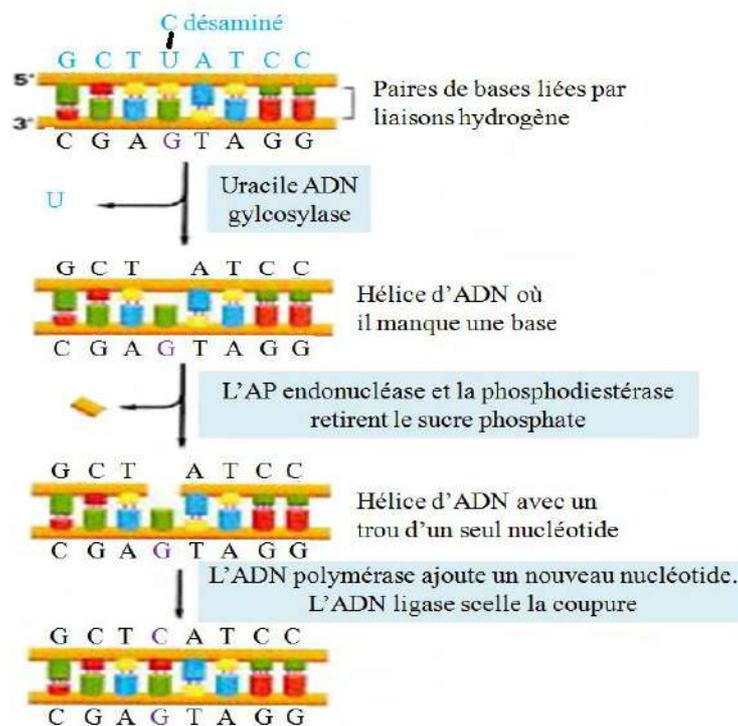
La réparation par excision de bases (BER), est aussi appelé mécanisme de correction courte. Ce mécanisme vise le plus souvent à remplacer les bases azotées altérées par un processus oxydatif endogène. Il est donc capable de réparer par exemple les **désaminations**, les **dépurinations** ou les **dépyrimidations spontanées**. Cette réparation aboutit à la **réparation d'un site AP** (dans le cas des désaminations spontanées, il y aura au préalable création du site AP).

Ce mécanisme met en jeu une **ADN-glycosylase**, dont la nature varie selon la base endommagée. Cette ADN-glycosylase va reconnaître et exciser spécifiquement une base modifiée, par clivage de la liaison N-Glycosidique (entre la base et le désoxyribose).

Lorsque ces ADN-glycosylases agissent, elles aboutissent à la **formation d'un site AP** (site abasique). L'ADN-glycosylase qui rentre en jeu étant spécifique à la base qui est endommagée (ex : *Uracile ADN-glycosylase*).

Il faut ensuite réparer le site AP. Cette réparation est faite avec deux autres enzymes. Tout d'abord l'**AP-endonucléase** (*apurinic-apyriminidic endonuclease*) qui a pour rôle de couper le squelette désoxyribose phosphate contenant le site AP, et donc d'enlever le sucre. Il y a donc clivage (hydrolyse) de la liaison phosphodiester, et retrait du site AP ce qui génère une coupure simple brin.

Le trou sur la chaîne d'ADN est alors complété par l'**ADN polymérase** qui se positionne en 3'-OH pour ajouter par polymérisation le nucléotide complémentaire. La dernière liaison phosphodiester est faite par l'**ADN ligase** afin de restaurer la continuité de l'ADN.



**Figure 40 :** Réparation par excision de base.

- **Réparation par excision de nucléotides (NER)**

Ce deuxième mécanisme est basé sur le même principe que le précédent, mais il correspond à des lésions plus volumineuses, où il faut exciser plusieurs nucléotides. Ce mécanisme s'adresse donc à des lésions ou des modifications structurales importantes (ex : dimères de thymines ou pyrimidines).

Ce mécanisme de réparation passe par l'excision sur le brin endommagé d'un segment de plusieurs nucléotides encadrant la lésion. Il nécessite l'intervention d'un ensemble de facteurs protéiques : des facteurs de reconnaissance qui se lient initialement à la région

endommagée, des hélicases pour ouvrir le duplexe d'ADN de part et d'autre, des endonucléases simple-brin pour couper en 5' et en 3' de la lésion.

Après excision, une ADN polymérase synthétise à nouveau le brin enlevé en utilisant le brin intact comme matrice. La réaction est terminée par l'action d'une ADN ligase qui suture le squelette phosphodiester de l'ADN.

La réparation par excision de nucléotides est un mécanisme présent dans toutes les cellules vivantes. Chez les bactéries, elle est prise en charge par le **complexe UvrABC**. Chez l'homme, ce sont les protéines de la **famille XP**, associées au facteur de transcription **TFIIH** qui assurent cette fonction. Cette dénomination, XP, est l'acronyme de *Xeroderma pigmentosum*, une affection génétique sévère dont souffrent les patients déficients dans cette voie de réparation de l'ADN.

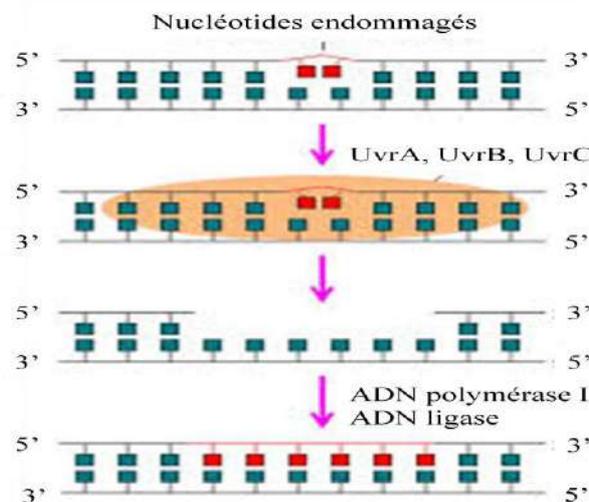
### Chez *E. coli*

La réparation est assurée par un complexe multi-enzymatique : 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**.



**Figure 41** : Réparation par excision de nucléotides.

### Chez l'Homme

Ce sont les **protéines ERCC1** et les **protéines XP** (A, B, C, ...) 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  $\delta$**  et l'**ADN pol  $\epsilon$**  et la dernière liaison phosphodiester est réalisée par l'**ADN ligase**.

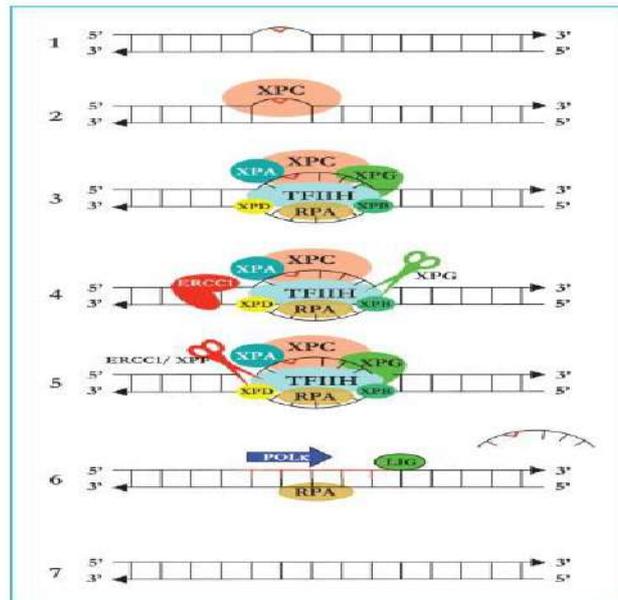


Figure 42 : Mécanisme NER chez l'homme.

#### Remarque

La maladie *Xeroderma pigmentosum* (enfant de la lune) est une maladie où les sujets sensibles à la lumière UV sont souvent porteurs de modifications des gènes XP, c'est-à-dire qu'ils ont un système de réparation altéré, ce qui aboutit à une incapacité de réparer les paires de thymines.

#### II.3.1.2.3. Réparation par recombinaison homologue

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, soit car ils étaient défectueux, soit parce qu'ils étaient débordés.

Il agit contre les **lésions majeures de l'ADN**, lorsque le système NER est débordé ou défectueux. S'il n'y a pas de réparation avant réplication, une lacune (ou brèche post-répllicative) est laissée par l'ADN polymérase, lorsqu'elle rencontre une lésion. Cette lacune va être réparée par recombinaison homologue. Ce mécanisme agit surtout contre les dommages de coupures doubles brin induits par les radiations ionisantes ou les agents oxydants.

La réparation par recombinaison correspond à la **synthèse translésionnelle (TLS)** qui consiste à poursuivre la réplication de l'ADN au niveau d'une lésion du brin matriciel de l'ADN ne permettant aucun appariement. Elle se réalise en même temps que la réplication.

L'ADN-polymérase II a un rôle important dans la reprise de la synthèse d'ADN après la lésion, l'ADN-polymérase III est alors transitoirement expulsée pour que la réplication puisse continuer.

L'ADN-polymérase III arrive au niveau d'une erreur et est éjectée par les enzymes de la TLS qui remplacent la base erronée ou alors qui permet la poursuite de la réplication un peu plus loin (ceci suivant l'ADN-polymérase utilisé (II, IV ou V)).

- **Chez les procaryotes**

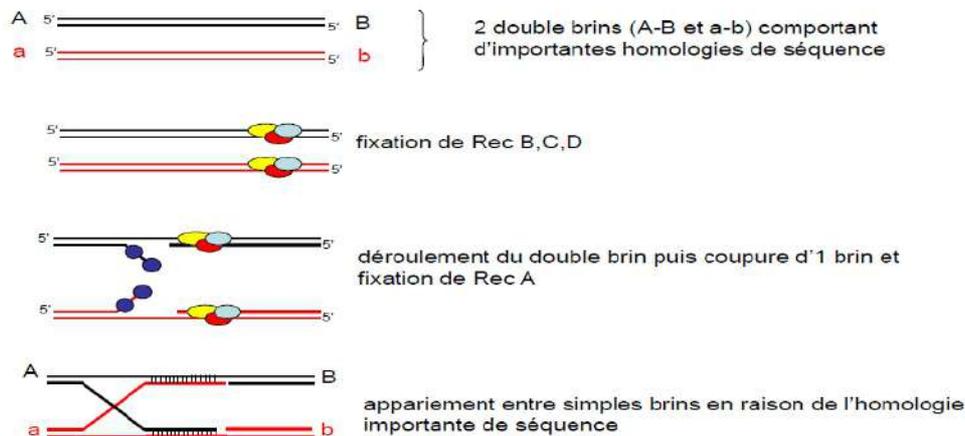
Lors de la réplication, une lacune est laissée sur le brin fils en face du dimère. Cette lacune est remplie par la séquence du brin parental opposé identique à ce brin fils, qui est

prélevée par la **protéine RecA**. Cela va fournir la séquence correcte et créer une deuxième lacune sur le brin parental opposé. Puis le dimère est excisé par des enzymes d'excision.

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.

La bactérie possède près de 1 milliard de protéines RecA, qui sont normalement présentes en quantité suffisante pour effectuer les recombinaisons.

Le pool des protéines RecA présentes dans la cellule est contrôlé par un **inhibiteur LexA** qui est le répresseur de RecA. Cet inhibiteur bloque la synthèse de protéines RecA et régule donc le nombre de protéines présentes dans la cellule.



**Figure 43** : Recombinaison de l'ADN chez les procaryotes.

- **Chez l'homme**

Le système est équivalent mais il fait intervenir chez l'homme la **protéine Rad51** (homologue eucaryote de RecA), associée aux protéines **BrcA1 et BrcA2**.

#### II.3.1.2.4. Système SOS

Le système SOS regroupe un ensemble de gènes (env. 30) qui est impliqué dans la réplication de l'ADN, dans la réparation de l'ADN et dans la division cellulaire et dont l'expression est contrôlée par une altération de l'ADN.

Si les dommages dans l'ADN sont trop importants : le **système SOS** se met en place. C'est le dernier système pour tenter de réparer les dommages de l'ADN. C'est un système de réparation par recombinaison homologue, sauf qu'il est **inductible**.

Lorsque les systèmes de recombinaisons sont débordés, la réplication est stoppée. Le système SOS se met en place. Il active la deuxième fonction de la protéine RecA : son **activité protéolytique** (=dégradation de protéines). Cette activité aboutit à la dégradation de son propre répresseur LexA. Il y a alors synthèse de protéines RecA et d'une vingtaine d'autres protéines issues des gènes SOS. Cela permet de stimuler le système de recombinaison homologue.

#### Remarque

Contrairement aux autres systèmes de réparation, le système SOS est **inductif** : il s'adapte, alors que les réparations par excision sont **constitutives**.

## Mécanismes moléculaires (Recombinaison et transposition de l'ADN)

### Objectifs spécifiques

Au terme de ce cours qui est la recombinaison et la transposition de l'ADN, vous devez être capable de :

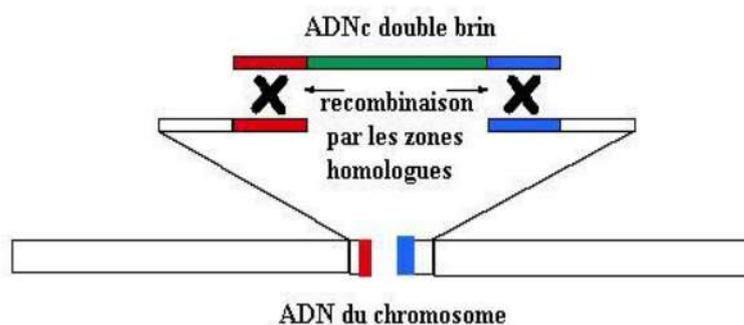
- Connaître ce que signifie la recombinaison ainsi que la transposition de l'ADN ;
- Identifier les différents éléments transposable ;
- Identifier et comprendre les différents mécanismes de la recombinaison et de la transposition d'ADN.

### III.1. Recombinaison génétique

Dans son sens le plus large, la recombinaison c'est la réorganisation du génome soit par réassortiment au hasard des chromosomes à la méiose, soit plus spécifiquement par cassure et recollage des brins d'ADN.

La **recombinaison génétique** désigne les processus qui conduisent à changer l'association physique entre deux segments d'ADN/ARN c'est-à-dire un **échange d'information génétique** entre deux **génomes différents**. En biologie moléculaire le terme recombinaison génétique est souvent utilisé comme synonyme de la recombinaison de l'ADN. Il s'agit en général d'un échange entre fragments d'ADN. Mais chez certains virus comme celui de la grippe, il s'agit d'échange d'ARN.

Chez les *eucaryotes*, elle se produit lors de la **reproduction** sexuée, grâce à la méiose, quand se forment les gamètes, et grâce à la fécondation. Chez les *procaryotes* (bactéries), elle se produit grâce à la **conjugaison** bactérienne, la **transformation** bactérienne, ou la **transduction** bactérienne. Chez les *virus*, la recombinaison peut avoir lieu au sein des cellules infectées par deux virus différents. L'apparition du nouveau virus H5N1 est par exemple, le résultat de recombinaisons génétiques.



**Figure 44 :** Recombinaison in vivo entre le génome et un fragment d'ADN par appariement de régions homologues.

#### III.1.2. Rôle de la recombinaison

La recombinaison de l'ADN joue un rôle fondamental à la fois dans la diversification des génomes et dans le maintien de leur homogénéité. Cela permet de créer de nouvelles combinaisons génétiques donc des génomes nouveaux.

La recombinaison est un phénomène naturel et universel dans le monde vivant, et c'est un des facteurs essentiels permettant de maintenir la diversité génétique dans une population. La formation de nouvelles combinaisons génétiques assure le brassage génétique et le maintien de la diversité génétique dans une population, ce qui augmente la possibilité pour une espèce de s'adapter à une modification de l'environnement. La recombinaison est donc l'un des processus essentiels de l'évolution des espèces.

### III.1.3. Mécanismes de recombinaison

Les approches moléculaires les plus récentes soulignent la grande diversité des mécanismes de recombinaison. Ceux-ci peuvent être classés en quatre grands groupes : recombinaison homologue, transposition, recombinaison site-spécifique et recombinaison illégitime.

#### III.1.3.1. Recombinaisons intra-chromosomiques

Les deux chromatides d'une même paire portent des allèles différents à un certain nombre de locus. Au cours de la prophase de la 1<sup>ère</sup> division méiotique, les chromosomes homologues s'apparient et s'enchevêtrent au niveau des chiasmas. Il se produit des échanges de segments entre ces chromosomes.

Ce phénomène est l'enjambement (ou *crossing-over*) : un allèle<sup>2</sup> d'un chromosome donné peut être échangé avec l'allèle porté par le chromosome homologue. Tous les gènes situés sur une paire de chromosomes peuvent être « brassés » grâce à l'enjambement ce qui modifie l'association d'allèles portés par chacun des chromosomes.

Ce brassage entre allèles d'une paire homologue est qualifiée d'intra-chromosomique. Un chromosome comporte en moyenne 1 000 gènes, alors il y aurait théoriquement  $2^{1000}$  possibilités. Mais ce calcul ne prend pas en compte les probabilités relatives au nombre de chevauchements possibles sur un même chromosome : il n'y en a en moyenne que de 1 à 5 par chromosome.

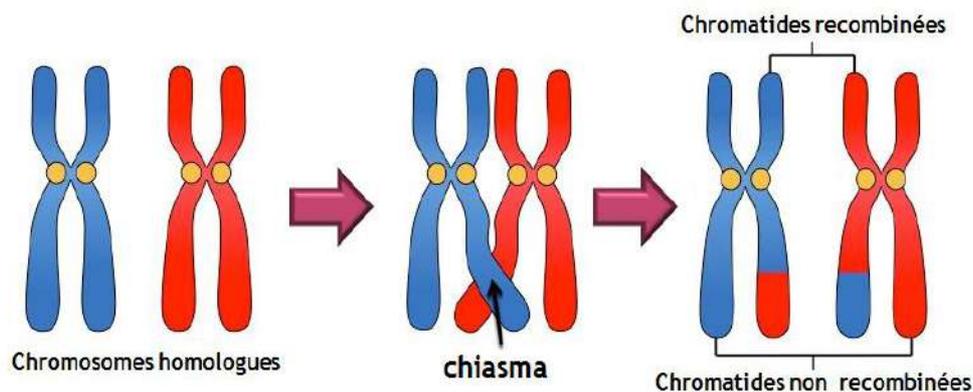
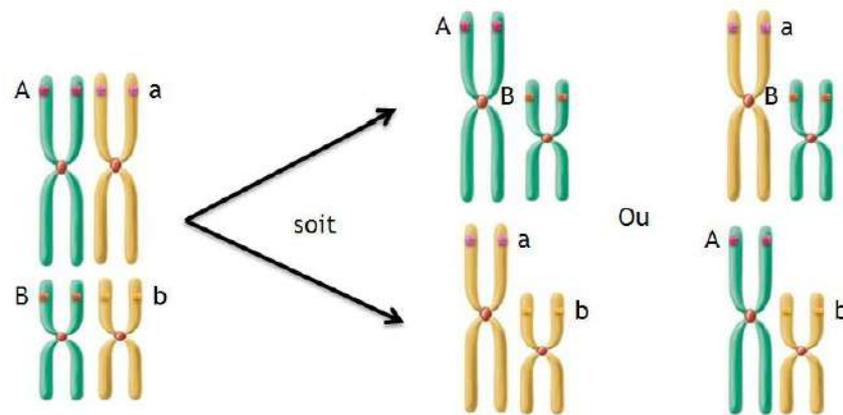


Figure 45 : Recombinaison Intra-chromosomique.

#### III.1.3.2. Recombinaisons inter-chromosomiques

Elles ont lieu pendant la méiose, en anaphase ou lorsque la cellule se sépare en deux, il y a une séparation aléatoire des chromosomes: si par exemple, il y a deux paires de chromosomes noté "Aa" et "Bb", il existe 4 possibilités soit "AB" et "ab" ou bien "aB" et "Ab".

<sup>2</sup> Un allèle est une version variable d'un même gène, c'est-à-dire une forme variée qui peut être distinguée par des variations de sa séquence nucléotidique.



**Figure 46 :** Recombinaison inter-chromosomique.

Chez l'homme et la femme il y a 23 paires de chromosomes, ce qui fait  $2^{23}$  possibilités soit plus de 8 millions. La dernière étape du brassage génétique est la fécondation: la rencontre au hasard des gamètes. Elle multiplie encore plus les possibilités, soit dans notre exemple  $8 \text{ millions} \times 8 \text{ millions} = 64 \text{ 000 milliards}$ .

Donc le brassage inter-chromosomique et la fécondation apportent plus de 64 000 milliards ( $64 \times 10^{12}$ ) de combinaisons génétiques originales et donc de descendants différents si on exclut les cas de forte consanguinité.

En rajoutant les recombinaisons intra-chromosomiques, le nombre de combinaisons originales et donc de descendants différents augmente encore et rend la probabilité d'avoir deux individus avec le même génome (hors vrais-jumeaux) quasi nulle.

### III.1.3.3. Recombinaison homologue

Le terme recombinaison homologue aussi appelée recombinaison générale, correspond à un événement de recombinaison génétique (Echange de fragments d'ADN) entre deux séquences identiques situées sur 2 molécules d'ADN différentes, ou distantes l'une de l'autre sur la même molécule.

La recombinaison homologue est impliquée dans la réparation des dommages subis par l'ADN ; chez la levure c'est la principale stratégie de réparation des cassures double-brin.

#### 1) Caractéristiques

La recombinaison homologue nécessite des régions génomiques identiques afin qu'elle puisse avoir lieu. Lorsque l'étendue de cette homologie entre les régions recombinantes descend au-dessous de quelques dizaines de paires de bases chez les procaryotes et de 200 paires de bases chez les mammifères (dans le cas de recombinaisons intra- ou inter-plasmidiques), la fréquence de recombinaison diminue de manière brutale.

Une autre caractéristique de la recombinaison homologue est qu'elle peut se produire n'importe où dans le génome. Pourtant, elle n'est pas forcément équiprobable, certaines régions du génome pouvant être plus fréquemment impliquées que d'autres dans des événements de recombinaison.

Dans les cellules de mammifère, par reconstitution d'un gène à partir de deux copies tronquées de ce gène contenant de moins en moins d'identité entre elles, le nombre minimal de nucléotides consécutifs identiques nécessaire à la recombinaison homologue est estimé à quatorze pour la recombinaison intramoléculaire et à vingt-cinq pour la recombinaison intermoléculaire.

## 2) Phases du processus

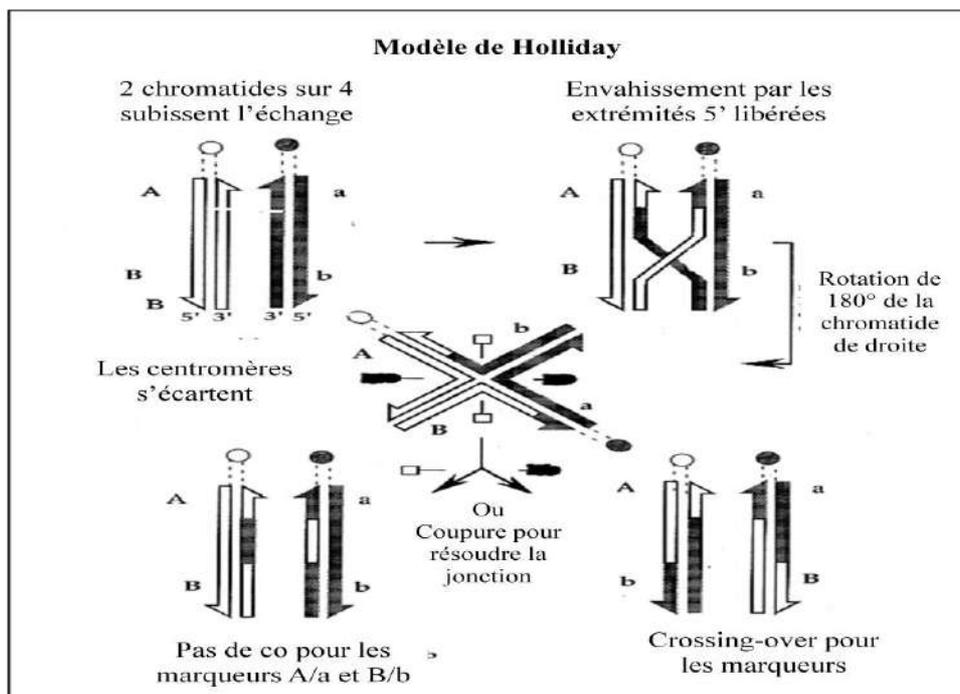
Plusieurs modèles ont été proposés pour décrire la recombinaison homologue. Tous ont en commun l'invasion de la terminaison 3'-OH d'une séquence d'ADN simple brin par un deuxième ADN double brins. L'appariement initial peut se produire à n'importe quelle position de la région homologue.

La réaction d'échange entre brins commence quand les deux molécules sont alignées et que l'extrémité de l'ADN est libre. S'établit alors une des étapes clef de la recombinaison homologue : la jonction des deux molécules (jonction de Holliday) formant une région hétéro-duplex.

S'il y a coupure puis religature des deux brins qui ont subi l'échange il n'y aura pas de crossing-over visible, seul un très petit fragment sera échangé, sur un seul brin, ce phénomène est appelé conversion génique. Cette recombinaison s'effectue sans modification des sites donneurs et le plus souvent sans échanges des marqueurs flanquants, donc dans ce cas sans crossing-over.

- **Modèle de Holliday (1964)**

Il permet d'expliquer le crossing-over et la conversion génique. Diverses améliorations de ce modèle ont été proposées pour rendre compte de phénomènes de recombinaison très rares.



**Figure 47 :** Représentation du modèle de Holliday.

Le mécanisme de la recombinaison homologue, conservé dans le monde vivant, suit trois grandes étapes initiées par une cassure double-brins de l'ADN et qui sont :

→ Une **phase pré-synaptique** durant laquelle le site de cassure est reconnu, partiellement dégradé, et où les protéines de recombinaison sont assemblées.

Suite à la formation d'une cassure double-brin de l'ADN, la recombinaison homologue est initiée par la dégradation, de part et d'autre de la cassure, du brin d'ADN orienté dans le sens 5' → 3'. Cette étape, dite de *résection*, expose un ADN simple brin orienté dont l'extrémité est 3'.

L'ADN simple brin généré par la résection est immédiatement lié par RPA, une protéine abondante à haute affinité pour l'ADN. Cette protéine doit être déplacée pour permettre à la recombinase Rad51/RAD51 (RecA chez les procaryotes) de s'assembler en un filament sur cet ADN simple-brin. Chez l'homme, la protéine BRCA2 catalyse ce remplacement. D'autres protéines s'associent au filament pour assurer sa stabilité.

→ Une **phase synaptique**, qui englobe la recherche et l'échange de brin d'ADN avec une séquence similaire (homologue) intacte pour opérer la réparation de la cassure.

La recombinase RAD51 utilise l'information des séquences de l'ADN simple brin alentours pour rechercher une séquence identique. Le mécanisme de cette recherche d'homologie est encore assez mal connu. Une fois la séquence homologue identifiée au sein du filament de RAD51, la protéine RAD54 glisse le long du filament et, en même temps qu'elle enroule le brin d'ADN "chercheur" et son complémentaire pour former un ADN hétéro-duplex, elle détache la protéine RAD51. La structure résultant de cet échange de brins, qui comporte un ADN double-brin hétéro-duplex et un ADN simple-brin déplacé, est connu sous le nom de D-loop (pour Displacement loop).

→ Une **phase post-synaptique** durant laquelle les intermédiaires d'échanges de brins d'ADN sont résolus à la suite de la synthèse d'ADN requis pour restaurer la séquence perdue au site de cassure.

L'extrémité 3' de la molécule cassée se retrouve hybridé à son homologue intact au sein de la D-loop. Une ADN polymérase étend cette extrémité en utilisant l'ADN complémentaire intact comme matrice. C'est cette étape de synthèse qui permet la ré-acquisition de l'information perdue au site de cassure.

Une fois l'étape de synthèse accomplie, la D-loop peut être démantelée ou résolue. Plusieurs sous-mécanismes branchent à partir de cet intermédiaire.

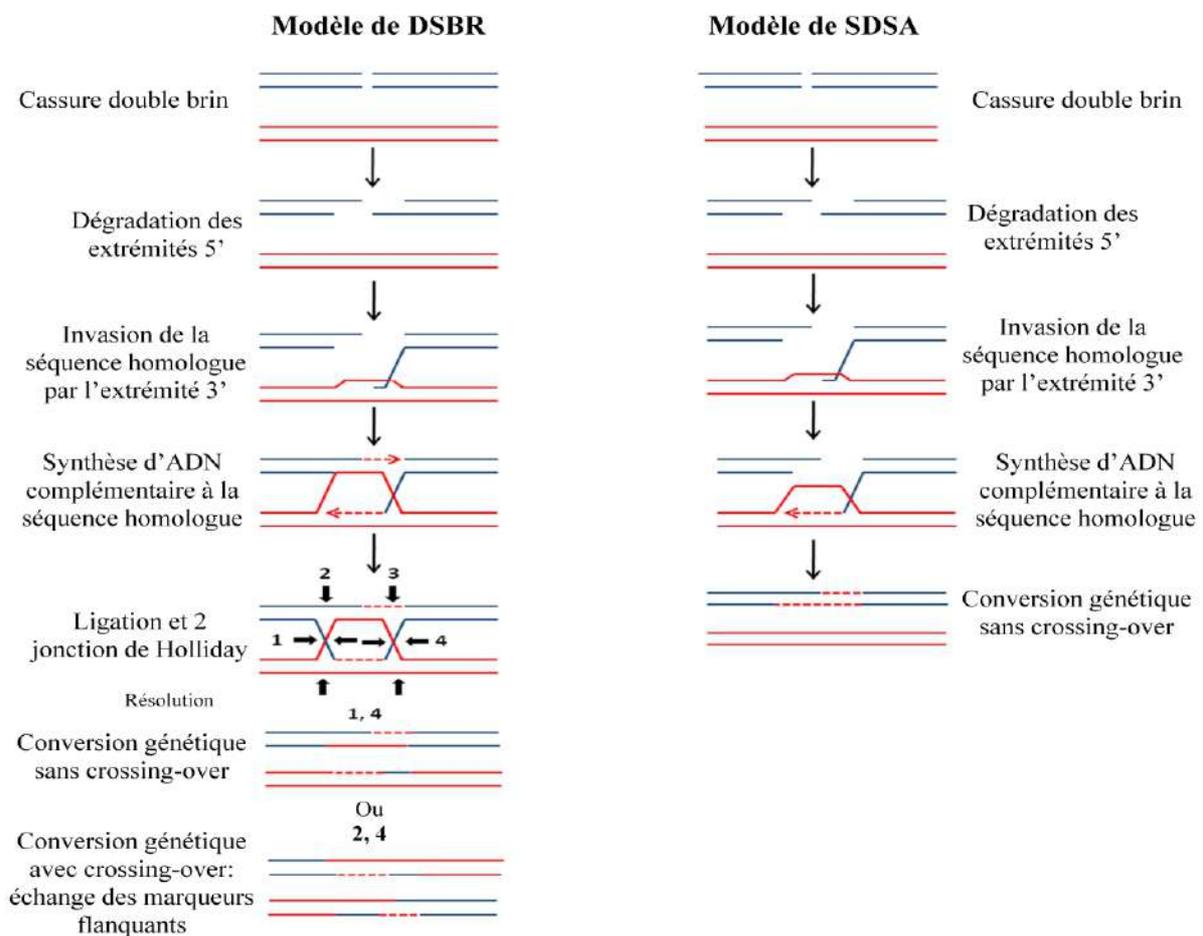
- **SDSA (Synthesis-Dependent Strand Annealing "synthèse dépendante du brin apparié")**

Dans le modèle du SDSA, la D-loop est démantelée, et l'extrémité d'ADN étendue peut se ré-associer à l'autre extrémité de la cassure grâce à la séquence nouvellement acquise. Suit alors une étape de synthèse d'ADN qui comble l'ADN simple-brin généré par la résection sur les deux molécules.

- **DSBR (Double-Strand Break Repair "cassure double brins et réparation")**

Le modèle du DSBR postule la capture du brin déplacé par la synthèse au sein de la D-loop par le brin situé de l'autre côté de la cassure. Cette capture et la synthèse d'ADN qui s'ensuit conduit à la formation d'un intermédiaire contenant deux hétéro-duplexes, avec à leurs frontières deux jonctions de Holliday. Des enzymes appelées "structure-selective endonucleases" reconnaissent et clivent ces jonctions.

Suivant les brins clivés, la résolution conduira ou non à un crossing-over. Un crossing-over implique la translocation entre la molécule initialement cassée et la molécule qui servait de matrice pour la réparation. Cette issue est généralement évitée dans les cellules somatiques mais est requise dans les cellules germinales pour la réalisation de la première division méiotique.



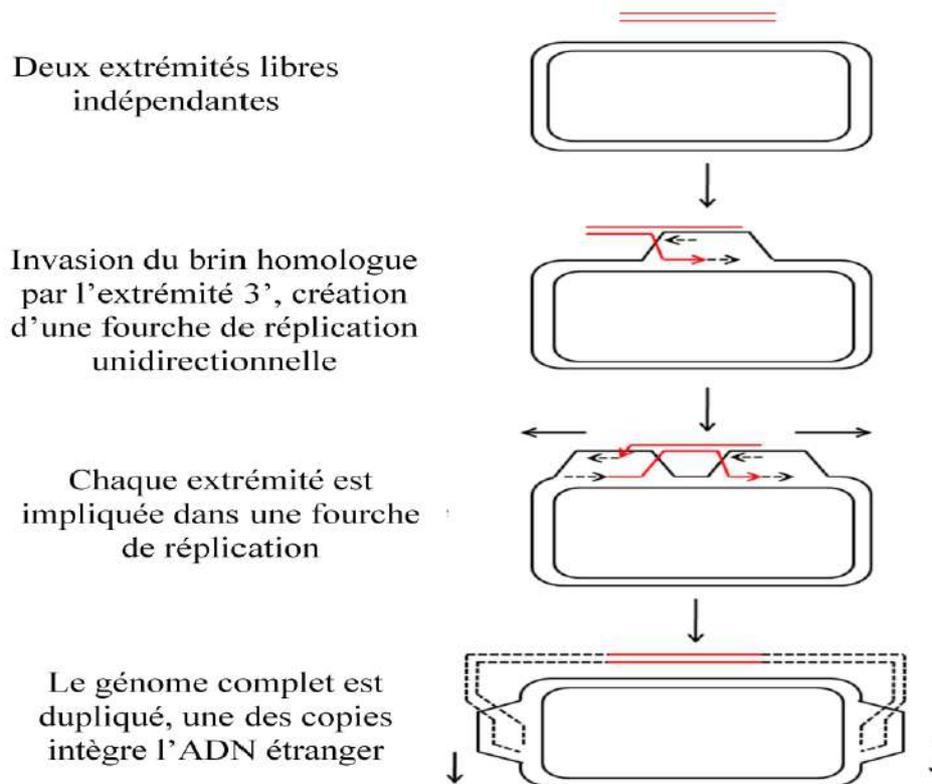
**Figure 48 :** Modèles moléculaires de recombinaison permettant d'expliquer la conversion génique chez les eucaryotes. Présentation succincte des modèles de DSBR, et SDSA.

Les séquences homologues sont présentées de deux couleurs différentes. La synthèse d'ADN est symbolisée par une flèche, et l'ADN nouvellement synthétisé par des pointillés.

- **BIR (Break Induced Replication "Réplication Induite par une cassure")**

Dans le modèle du BIR, la synthèse d'ADN initiée au niveau de la D-loop se poursuit de façon unidirectionnelle jusqu'à l'extrémité du chromosome. Ce mode de réparation est

principalement mis en œuvre au niveau d'une cassure d'ADN ne possédant pas de deuxième extrémité, comme celles générées par la cassure d'une fourche de réplication.



**Figure 49 :** Double “break-induced replication”.  
Modèle de recombinaison couplé à la réplication chez *E. coli*.

**Remarque**

La recombinaison homologue est un processus très complexe qui implique plus d'une trentaine de gènes chez *E. coli*. Parmi ceux-ci on peut citer *recA* qui contrôle l'appariement des homologues, les autres gènes *rec*, les gènes *ruv* sont impliqués dans l'échange des brins.

### III.1.3.4. Recombinaison site-spécifique

La recombinaison site spécifique décrit un processus impliquant des événements de recombinaison programmés très spécialisé qui implique un échange réciproque entre des sites bien définis et précis des génomes.

Les sites les plus simples ont une longueur comprise entre 20 et 30 nucléotides et sont composés de deux courtes répétitions inversées auxquelles se lie la recombinase entourant une zone centrale où se produit le crossing-over.

#### 1) Etapes de ce processus

Les principales étapes de ce mode de recombinaison sont :

- La reconnaissance des sites par la recombinase ;
- La formation du complexe synaptique ;
- Le clivage de l'ADN ;
- L'échange des brins, puis la religature de la molécule et ;
- Enfin la rupture du complexe synaptique qui permet la libération des recombinants.

L'énergie de la liaison phosphodiester est conservée après le clivage et est réutilisée pour la ligature. C'est pourquoi, il n'y a pas nécessité de requérir à un cofacteur fournisseur d'énergie tel que l'ATP dans cette réaction.

## 2) Résultats de cette recombinaison

Selon la disposition initiale des sites de recombinaison, la recombinaison sites spécifiques peut engendrer trois résultats : l'intégration, l'excision ou l'inversion d'une molécule d'ADN.

- L'**intégration** résulte de la recombinaison entre deux molécules séparées (dont au moins une qui est circulaire).
- Par contre, lorsque les sites sont localisés sur le même chromosome, le résultat de la recombinaison sera une **excision** entre des sites ayant la même orientation, et une **inversion** lorsque l'échange se produit entre des sites dont l'orientation est inversée.

Il existe deux familles de recombinases, les tyrosines recombinases, appelées aussi intégrases et les sérines recombinases ou résolvases/invertases. Leur nom provient de l'acide aminé nucléophile qui sera lié à l'ADN pendant l'échange des brins d'ADN.

Les sérines recombinases coupent les deux brins d'ADN aux deux sites d'échange, tous les brins sont clivés avant que l'échange soit initié. Les tyrosines recombinases, quant à elles, clivent un brin à la fois dans chaque duplex. Il doit être joint à son nouveau partenaire avant que le second brin soit coupé, ce qui produit une jonction de Holliday.

Ces dernières sont celles qui sont le plus souvent rencontrées chez les procaryotes, l'exemple le plus connu est l'intégration du phage  $\lambda$  dans le génome d'*E. coli*.

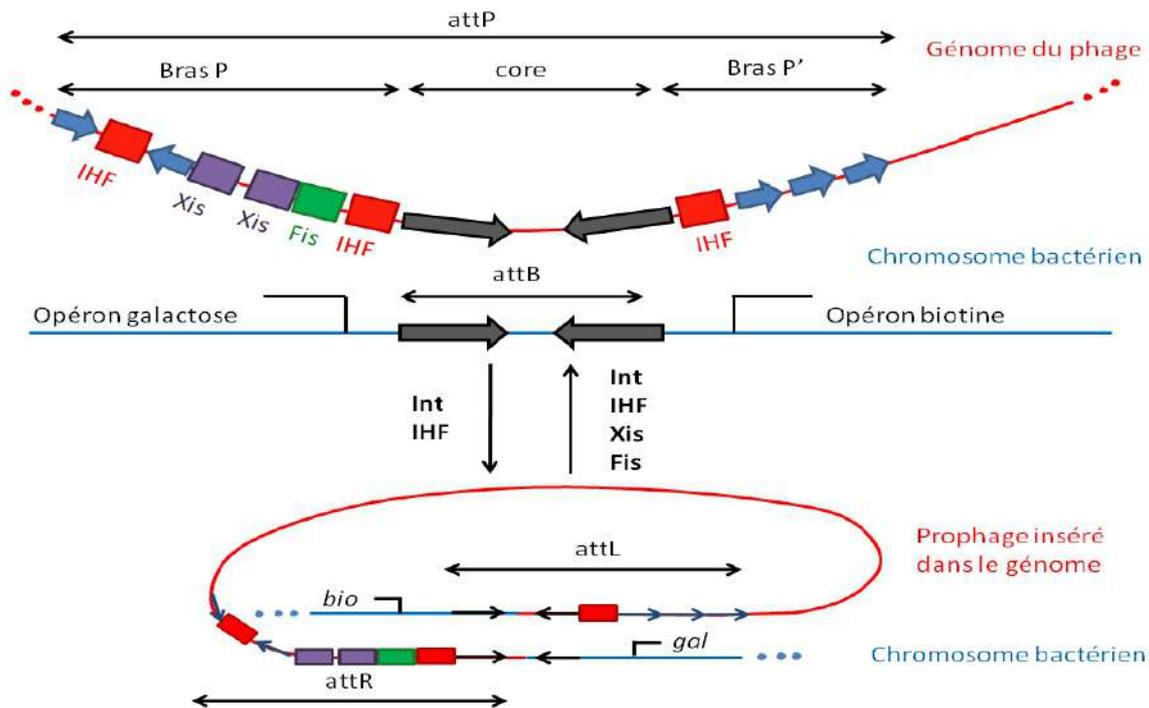
### Exemple

Le phage  $\lambda$  dans le cas où il suit un cycle lysogénique est en dormance sous la forme d'un prophage intégré dans le génome bactérien. La recombinaison site spécifique permettant cette intégration est une des plus connues.

Le seul et unique site dans le chromosome d'*E. coli* où peut s'intégrer le phage  $\lambda$  est situé entre les opérons *gal* (galactose) et *bio* (biotine). Il est appelé « attB » pour site d'attachement bactérien. Ce site est constitué de seulement 30 nucléotides, la séquence centrale en comporte 15. Le site de recombinaison présent sur le phage (attP) est plus complexe. Il contient la même séquence de 15 nucléotides bornée par des sites de liaison protéiques appelés bras P et P'.

L'intégrase (Int) phagique et le facteur d'intégration dans l'hôte (IHF) bactérien se lient aux sites P et P' pour former un complexe dans lequel les sites attB et attP seront alignés. Les brins sont alors coupés, échangés et joints à nouveau.

Le résultat de la recombinaison est que le prophage intégré est flanqué par deux sites d'attachement modifiés (attR et attL). Xis (excisionase) et Fis sont deux protéines qui se lient à des sites spécifiques sur le bras P et qui, avec Int et IHF vont former un complexe autour de attR et attL alignés et catalyser la réaction inverse permettant ainsi l'excision du prophage (Fig. 50).



**Figure 50 :** Réaction de recombinaison site spécifique entre le phage  $\lambda$  et le chromosome bactérien.

*Int* et *IHF* sont nécessaires à l'intégration de l'ADN phagique dans le génome bactérien. *Int*, *IHF*, *Xis* et *Fis* sont requises lors de son excision.

### III.1.3.5. Recombinaison illégitime

La recombinaison illégitime ou non homologue regroupe les types de recombinaison ne répondant pas aux définitions précédentes. C'est un processus cellulaire pouvant se faire à de nombreux sites et conduisant à la réunion de deux fragments d'ADN ne présentant pas ou très peu d'homologie de séquence.

Elle est à l'origine de diverses anomalies du génome (délétion, insertion, translocation, amplification génique, intégration virale) impliquées dans l'apparition de maladies génétiques ou le développement tumoral.

Elle permet l'élimination des cassures double-brin accidentelles de la chaîne nucléotidique, consécutives à des erreurs du métabolisme normal de l'ADN (réplication, transcription, réparation) ou à la présence de lésions sur l'ADN produites par les agents génotoxiques.

La recombinaison illégitime regroupe en fait différents mécanismes moléculaires selon le contexte. Trois types de jonctions de recombinaison illégitime sont possibles suivant qu'il existe, sur le site de recombinaison :

- Quelques nucléotides communs aux deux séquences recombinaisonnées (Fig. 51.B) ;
- Aucun nucléotide commun aux deux séquences réarrangées (Fig. 51.C) ou ;
- Des nucléotides supplémentaires n'étant auparavant présents sur aucune des deux séquences impliquées dans la jonction de recombinaison (Fig. 51.D).



**Figure 51** : Exemples de jonctions de recombinaison homologue et non homologue.

Les points indiquent les nucléotides communs aux deux séquences impliquées dans la recombinaison. Les bases communes à la jonction de recombinaison sont soulignées et les nucléotides supplémentaires à cette jonction sont en gras. Les nucléotides en rouge correspondent aux nucléotides communs aux deux séquences parentales.

### 1) Conséquences

Une des hypothèses quant aux conséquences biologiques de la recombinaison illégitime est qu'elle permettrait une meilleure survie des cellules endommagées grâce à l'élimination des cassures double-brin accidentelles de la chaîne nucléotidique, consécutives à des erreurs du métabolisme normal de l'ADN (réplication, transcription, réparation).

Elle peut être considérée comme un mécanisme de sauvegarde cellulaire permettant de réparer les coupures double-brin accidentelles au prix de la production de chromosomes endommagés. Elle semble en général assez peu fréquente chez *E. coli*, pourtant, sa fréquence augmente de manière importante lorsque la cellule est soumise à des agents préjudiciables à l'ADN comme les rayons ultraviolets.

### 2) Phénomènes

La recombinaison illégitime est impliquée dans différents types de réarrangements. Cela implique que, sous le terme « recombinaison illégitime », sont regroupés des phénomènes ne faisant pas nécessairement appel à des mécanismes moléculaires identiques. Un modèle de recombinaison développé pour un type de remaniement ne peut pas être automatiquement

étendu à l'ensemble des phénomènes de recombinaison non homologue. Il n'existe donc pas de modèle général de la recombinaison illégitime.

- **Recombinaison illégitime entre de courtes séquences répétées**

Il peut s'agir de recombinaison entre de courtes séquences homologues. Chez *E. coli*, elle serait mise en œuvre entre des séquences homologues comportant moins de 20 pb. Au dessus, ce serait la recombinaison homologue (RecA dépendante) qui interviendrait, cette limite étant dépendante des organismes.

Les courtes séquences homologues sont dans ce cas toujours des répétitions intrachromosomiques relativement proches l'une de l'autre. Il a également été observé que la présence de palindromes (formant des structures secondaires en tige-boucle) augmente la fréquence de la recombinaison illégitime.

Les protéines impliquées ne sont pas encore toutes connues, mais il semblerait que l'exonucléase RecJ intervienne. L'hélicase RecQ paraît quant à elle réprimer cette forme de recombinaison illégitime. DnaB, une autre hélicase semble elle aussi jouer un rôle.

- **Recombinaison illégitime associée à des éléments sites spécifiques**

Il semblerait qu'en cas de stress important, l'activité des transposons puisse s'intensifier. Ce processus pourrait faire intervenir la recombinaison illégitime. La transposition intramoléculaire induirait des réarrangements tels que des délétions, des insertions, des inversions.

Il a été également observé que les recombinases impliquées dans la recombinaison site spécifique présentée ci-dessus peuvent se tromper de site d'action (on parle de pseudo-site ou de quasi-site). Par exemple, une erreur lors de la transposition de certain transposon (Mu, Tn10) induit une délétion d'un morceau de séquence adjacent au transposon. Dans ces cas, on parle également de recombinaison illégitime.

- **Recombinaison illégitime entre des séquences non homologues**

Certains réarrangements peuvent se produire sans qu'il n'y ait aucune homologie de séquence. Par exemple, les **gyrases** (topoisomérases II) coupent l'ADN double brins pour modifier son enroulement puis religuent les fragments.

Ces enzymes peuvent, lorsqu'elles sont liées à deux molécules d'ADN différentes, ou à deux endroits différents de la même molécule d'ADN, échanger leurs sous-unités entre elles ce qui entraîne un réarrangement après liaison des deux fragments.

Les topoisomérases de type I, permettent le relâchement du super-enroulement de l'ADN par coupure transitoire et ligation d'un seul brin. Il a été proposé qu'une topoisomérase I puisse lier deux extrémités de deux fourches de réplication différentes bloquées aux sites de terminaison ce qui entraînerait la délétion d'un fragment d'ADN sous forme circulaire.

### III.2. Transposition de l'ADN

Au cours de l'évolution des espèces, le génome s'est continuellement agrandi suite aux duplications de gènes et aux transpositions. La transposition est un mécanisme d'insertion ou de déplacement d'une séquence d'ADN que l'on appelle « transposon », provenant d'un même ADN ou d'un autre ADN (viral ou d'un ARN cytoplasmique rétro-transcrit).

Cette transposition est rendue possible grâce à l'intervention d'enzymes spécifiques que sont les *intégrases* ou les *transposases* et c'est habituellement le transposon lui-même qui code pour ces protéines. Cette transposase coupe la chaîne d'ADN, qui est ensuite réparée. Le déplacement qui en résulte peut-être simple (sans réplication du transposon) ou répllicative.

L'ADN qui contient l'élément transposable est dit « ADN hôte » et celui qui reçoit l'élément mobile après transposition est dit « ADN cible ».

#### III.2.1. Transposons

Un transposon ou **élément transposable** ou encore appelé parfois **élément mobile**, est une séquence d'ADN capable de se déplacer de manière autonome dans un génome d'un endroit à un autre sur un même brin d'ADN ou sur un autre brin, par un mécanisme appelé **transposition**.

#### III.2.2. Intérêts des transposons

Les transposons jouent un rôle très important dans la dynamique des génomes et donc dans l'évolution de la biodiversité notamment dans l'adaptation des organismes à leur environnement.

- Création de nouvelles mutations, de recombinaisons, de réarrangements des génomes. Ils contribuent au polymorphisme<sup>3</sup> ;
- Création de nouvelles combinaisons d'introns-exons ;
- Apparition de nouvelles fonctions cellulaires : constitution des centromères, système de réponse aux interférons, réparation des coupures double brin. Chez la drosophile les télomères<sup>4</sup> sont des transposons.
- On observe des cas de coévolution du génome de l'hôte et du transposon.
- En génétique moléculaire appliquée, nombreux usages: Mutagenèse et sélection de mutants, clonage de gènes, vecteurs pour transfert de gènes.

#### III.2.3. Classification des éléments transposables

Il existe deux classes d'éléments transposables selon leurs mécanismes de transposition :

##### 1) Classe I ou rétro-transposon

Ce sont des éléments qui utilisent un intermédiaire, l'ARN dont la mobilité est sous la dépendance de trois enzymes. L'élément est copié sous forme d'ARN simple brin par une **ARN polymérase** de l'hôte, puis rétro-transcrit en ADN par une **reverse transcriptase** et intégré dans un nouveau site par une **intégrase**. Le préfixe rétro vient du fait que leur ARN est rétro-transcrit en ADN.

Pour les rétro-transposons on parle de transposition de type copier-coller ou transposition répllicative puisqu'il y a une augmentation du nombre de copie de l'élément

<sup>3</sup> Caractérise la capacité à se présenter sous différentes formes.

<sup>4</sup> Un télomère est une région hautement répétitive d'ADN à l'extrémité d'un chromosome.

donnant naissance à des séquences répétées dispersées. Ils codent pour la fabrication de différentes protéines dont une **transcriptase inverse** (participant à la répllication de l'ADN), une **intégrase** (permettant l'intégration du transposon répliqué à l'ADN cible) et une **protéase** (impliqué dans la maturation des protéines de l'élément).

Il existe de nombreux types de rétro-transposons et leur nomenclature est encore sujette à débat. Ils se différencient en fonction de :

- Leurs différences structurales ;
- La présence ou l'absence de longues régions terminales répétitives LTR (Pour Long Terminal Repeats) non inversées ;
- Leur capacité à former des particules virales.

De ce fait nous distinguons entre deux classes de rétro-transposons : les rétro-transposons à LTR ou sans LTR. Ces deux classes d'éléments diffèrent également par leur mécanisme d'intégration.

- **Ceux possédant des LTR** séquences de 1000 à 10000 nucléotides chez l'humain, vont synthétiser un ADN complémentaire (ADNc) à partir de leur ARNm (reverse transcription) dans des particules pseudo-virales (supposées strictement cytoplasmiques), puis intégrer cet ADNc à un nouveau locus chromosomique.
- **Ceux n'ayant pas de LTR** vont synthétiser cet ADNc directement au site cible d'intégration, provoquant parfois eux-mêmes la coupure du double brin d'ADN nécessaire à cette intégration.

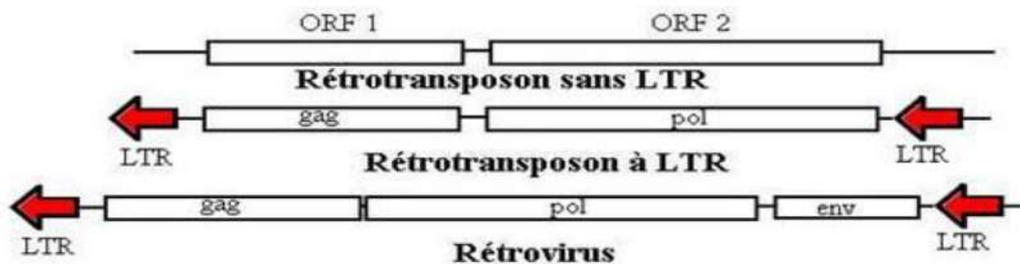


Figure 52 : Différents rétro-transposons.

## 2) Classe II ou transposon

Ces éléments sont coupés de l'ADN hôte (donneur) puis réintégré dans un autre site. Une seule enzyme catalyse la coupure et la réintégration c'est la **transposase**.

Pour les transposons à ADN, il s'agit d'une transposition de type couper-coller ou transposition conservatrice puisque le nombre de copies reste constant.

Au cours des 30 dernières années, le nombre des transposons chez les procaryotes et les eucaryotes identifiés n'a cessé à croître, avec une grande diversité dans leur structure et une remarquable conservation des mécanismes qui assurent leur mobilité.

Ils sont essentiellement caractérisés par la présence des ITR (*Inverted Terminal Repeat* ou *répétition terminal inversé*) en orientation inverse qui flanquent les transposons à chaque extrémité.

Certains de ces transposons sont dits autonomes car ils codent une enzyme la transposase (permettant son propre transfert) et d'autres non autonomes, devant utiliser la machinerie des éléments autonomes (MITEs, *Miniature Inverted repeat Transposable Elements*).

#### Remarque

Récemment, il a été mis en évidence des éléments transposables dont on ignore le mode de transposition : *miniature inverted repeat transposable elements* ou MITEs dans les génomes d'*Oryza sativa* "Riz asiatique" ( $10^5$  éléments qui représentent 6% du génome) et *Caenorhabditis elegans* "petit ver transparent".

Ils ont des extrémités inversées répétées de 15 bp séparées par 400bp, insuffisante pour coder une protéine. Ces éléments sont également présents chez l'homme, le xénope (Crapaud à griffe), le pommier, etc.

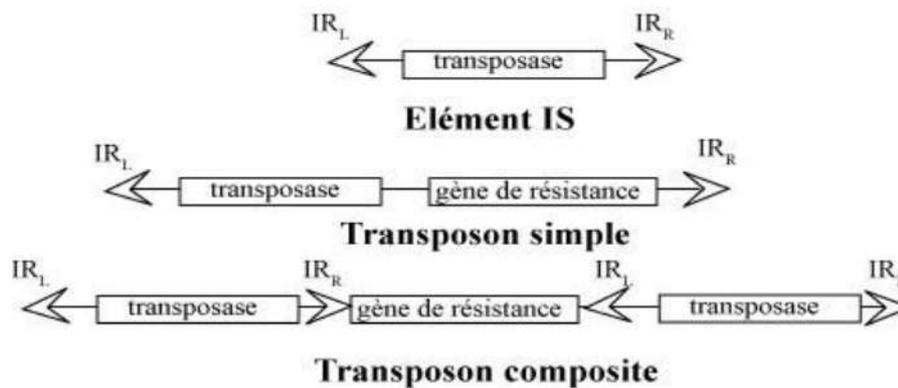


Figure 53 : Quelques transposons de type II.

### III.2.4. Types de transposition

La transposition revêt deux principaux mécanismes répllicative ou conservative. Selon les éléments transposables, un mode ou l'autre, ou les deux seront employés. Une variante de la transposition appelée rétro-transposition est toujours répllicative car l'élément transposable n'est pas transféré en tant que tel, mais simplement copié.

A l'endroit d'insertion (cible, site accepteur), une courte séquence de 4 à 12 pb est dupliquée et encadre l'élément transposable. Ces deux séquences forment une répétition directe  $5' \rightarrow 3'$  et constituent une marque de reconnaissance de la transposition.

Certains éléments transposables s'insèrent préférentiellement à un locus du génome, d'autres sont plus généralistes et s'insèrent sur un chromosome particulier ou encore sur n'importe quel chromosome. Cela dépend de l'élément transposable lui-même.

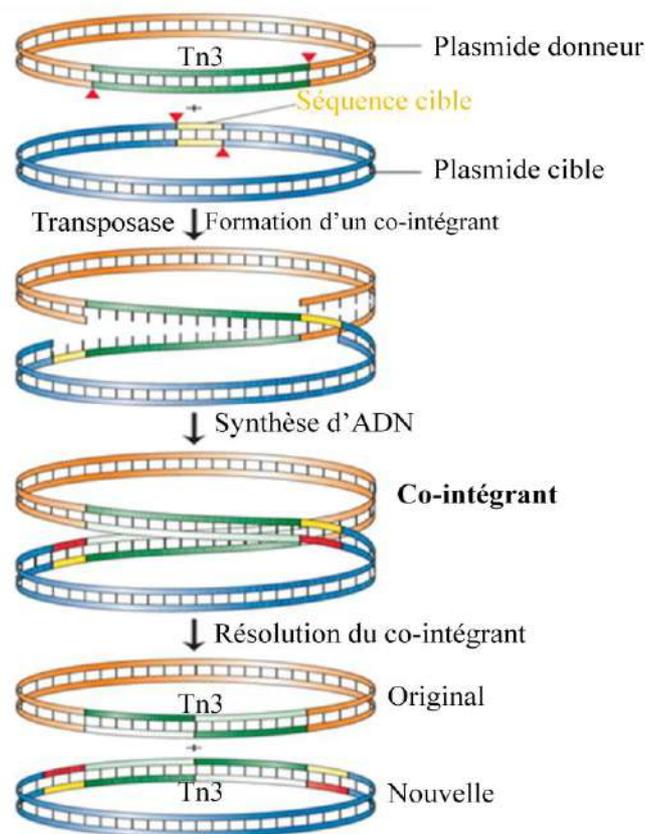
#### III.2.4.1. Transposition répllicative

Appelé aussi "co-intégration" ou "copier-coller", ce mécanisme a été proposé par **James A. Shapiro** en 1979, dans lequel l'élément transposable est dupliqué pendant la réaction, de sorte que l'entité de transposition est une copie de l'élément d'origine.

Dans ce mécanisme, les séquences d'ADN donneur et récepteur forment une configuration "thêta" intermédiaire caractéristique, parfois appelée "intermédiaire Shapiro". La transposition répllicative est caractéristique des rétro-transposons et se produit de temps en temps dans les transposons de classe II.

Ce processus implique la formation d'une structure moléculaire de type "*crossing-over*" appelée co-intégrat. Cette structure résulte de la fusion entre un réplicon (molécule circulaire capable de se répliquer) donneur et un réplicon récepteur avec une copie du transposon à chacune des deux jonctions entre ces deux molécules. Généralement, cette structure ne persiste pas et une recombinaison entre les deux copies du transposons résout le co-intégrat.

La transposase produit une coupure à un site bien précis de chaque côté du transposon ainsi qu'au site d'insertion, ensuite ces extrémités sont liées et une réplication se produit en utilisant les fonctions répllicative de la cellule hôte pour former le co-intégrat qui est détaché par une résolvasse codée par le transposon et qui agit au niveau du site *res* également porté par le transposon. Dans certains cas, une intégrase achève la résolution du co-intégrat en agissant sur un site appelé *att*.



**Figure 54** : Processus de transposition répllicative.

#### III.2.4.2. Transposition conservative

Egalement appelé transposition "non répllicative" ou "couper-coller", "cut-paste", ce modèle correspond à une simple insertion.

La transposase se fixe aux ITRs pour former une structure compacte appelée "complexe synaptique" ou "transpososome", qui rapproche les ITRs, ensuite une coupure double brin est effectuée par la transposase conduisant à la formation de l'intermédiaire de transposition (une molécule de transposase fortement associée au transposon).

En présence d'ions calcium ou magnésium, la transposase coupe au site cible d'une façon décalée et une structure branchée de type *crossing-over* est formée entre les brins d'ADN du transposons donneur et la séquence cible.

La réparation des extrémités décalées conduit à la duplication de la séquence cible. Cette transposition se distingue de la précédente par le fait que l'intégration du transposon dans le site cible se fait sans le passage par une étape de réplication.

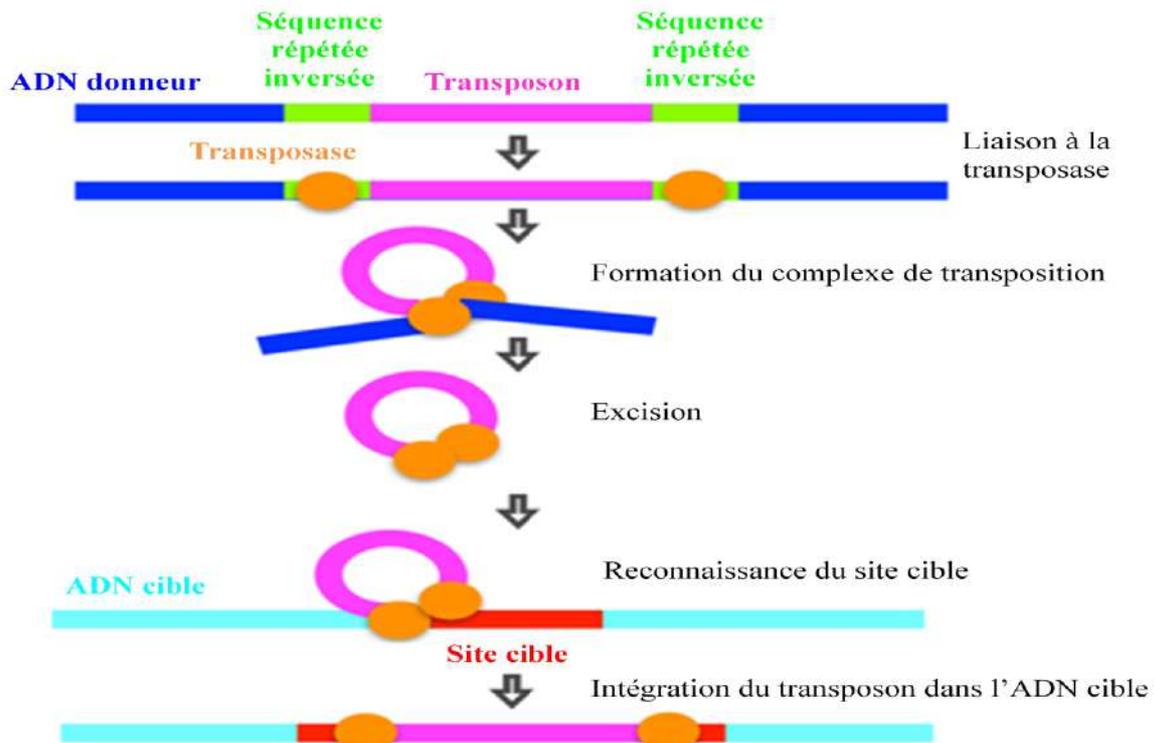


Figure 55 : Transposition conservative.

### III.2.5. Transposase et son mode d'action

Une transposase est une **nucléotidyltransférase** qui catalyse le déplacement d'un élément transposable (transposon) d'un endroit à un autre du génome. Cette enzyme se lie à une extrémité du transposon, clive la liaison entre celui-ci et le reste du génome, et réalise la suture de cet élément transposable à un autre endroit du génome.

Le terme *transposase* a été introduit en 1979 par la première équipe à avoir cloné l'enzyme nécessaire à la transposition du transposon Tn3. Les gènes codant ces endonucléases sont très répandus chez la plupart des êtres vivants et sont les gènes les plus abondants connus.

Les transposases sont peu conservées dans leur structure primaire, mais contiennent très souvent un motif conservé D, DxE/D où le résidu glutamate (E) ou aspartate (D) est généralement séparé des deux premiers résidus aspartate par un nombre connu d'acide aminé fréquemment conservé pour les membres d'une même famille, ce motif est impliqué dans la liaison de deux ions métalliques bivalents (magnésium ou calcium) et fait partie du site catalytique de l'enzyme, le mode d'action de la transposase est généralement comme suit :

- Les ions bivalents piégés par le motif D, D-E/D, interagissent avec les extrémités de l'élément et fragilisent un pont phosphodiester qui devient plus sensible à une attaque nucléophile par une molécule d'eau qui libère des extrémités 3'-OH.
- Une seconde attaque se lance contre les ponts phosphodiester de l'ADN cible y libérant des extrémités 5'-P sortantes.
- Selon que la transposase a clivé ou non l'extrémité 5' du transposon, la transposition sera conservative ou répliquative.

Dans le premier cas, les deux brins du transposon se désolidarisent simultanément du site donneur et la synthèse d'ADN s'arrête au niveau du transposon qui n'est pas répliqué.

En revanche lorsque la transposition est répliquative, le second brin du transposon n'est pas clivé, la synthèse à l'extrémité 3'-OH de la cible se poursuit à travers le transposon et le duplique, en engendrant un co-intégrant.

## Mécanismes moléculaires (Transcription de l'ADN)

### Objectifs spécifiques

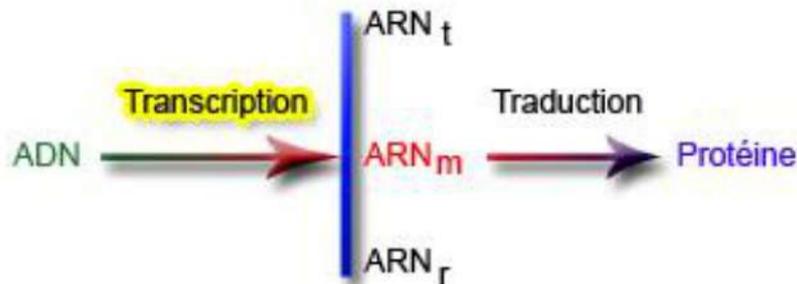
Au terme de ce cours qui est la transcription de l'ADN, vous devez être capable de :

- Connaître ce que signifie la transcription de l'ADN ;
- Identifier les différents étapes de la transcription d'ADN ;
- Connaître les différents complexes enzymatiques impliqués lors de la transcription de l'ADN chez les procaryotes et les eucaryotes.

### IV.1. Définition

La **transcription** est un processus biologique, au niveau de la cellule, qui fait intervenir un système enzymatique afin de **convertir** des régions de la molécule d'**ADN** en une molécule d'**ARN**. Cette molécule est appelée **transcrit**. En effet, si la molécule d'ADN est le support universel de l'information génétique, ce sont les molécules d'ARN qui sont reconnues par la machinerie de traduction en séquences protéiques.

Il faut savoir que ce n'est pas toute l'information génomique qui peut être transcrite. Les portions qui sont transcrites sont appelées les gènes. Les gènes des cellules **eucaryotes** (non bactériennes), ont leurs séquences codantes (appelées **exons**) interrompues par des séquences non codantes (appelées **introns**).



**Figure 56** : Représentation schématique des différents processus biologique de la cellule.

Les mécanismes généraux qui assurent la transcription des ARN (messager, ribosomique, de transfert, etc.) sont globalement identiques. Par contre, l'organisation des gènes codant ces différents ARN et le contrôle de leur expression sont très différents d'un type de gène à l'autre.

### IV.2. Éléments nécessaires à la transcription

La synthèse de l'ARN ne nécessite pas d'amorce et se fait dans le sens 5' → 3', de manière antiparallèle à l'ADN matrice et complémentaire de l'un des deux brins de l'ADN. Plusieurs composantes sont donc nécessaires afin que la transcription de l'ADN en ARN puisse se produire de façon correcte qui sont :

#### 1) Matrice de synthèse

L'ADN double brin sert de matrice à l'ARN polymérase qui synthétise un ARN simple brin dans le sens 5' → 3' grâce à l'énergie apportée par le relargage du phosphate.

Le choix du brin matrice se fait en fonction du sens de déplacement de l'ARN polymérase :

- Le brin *sens* est le brin *non transcrit* ou non matrice ;
- Le brin *anti sens* sera le *brin transcrit* et donc le **brin matrice**. Le brin anti-sens est le brin utilisé comme brin matrice par l'ARN polymérase. Il sera donc lu et transcrit.

#### Exemple

Si l'ARN polymérase lit le brin 3'→5', la synthèse d'ARN se fait de manière anti parallèle et complémentaire du brin anti-sens. Et donc l'ARN sera identique au brin sens (ici 5'→3'), sauf que la thymine est remplacée par de l'uracile.

Pour un même chromosome, les deux brins peuvent servir de matrice pour la transcription. Selon le gène considéré, le brin utilisé comme matrice peut différer, mais pour un gène donné c'est toujours le même brin qui est utilisé comme matrice. C'est la direction du mouvement de l'ARN polymérase le long de l'ADN qui détermine le brin sens et le brin anti-sens. La direction du mouvement est conditionnée par la position du site de l'initiation et de la région promotrice du gène.

## 2) Les ribonucléotides

Pour la synthèse de l'ARN, les ribonucléotides rentrent sous forme triphosphate (ATP, CTP, GTP et UTP), ce qui amène l'énergie nécessaire à la polymérisation par libération d'un pyrophosphate. On trouve également dans certains ARN des bases mineures :

- Puriques : hypoxanthine et 7-méthyl-guanine.
- Pyrimidiques : pseudo-uracile et 4-thiouracile.

## 3) L'ARN polymérase ADN-dépendante

L'enzyme qui catalyse cette réaction de transcription est appelée **ARN polymérase**. Ces ARN polymérases sont très complexes dont la structure est très variable selon les espèces. On peut distinguer 4 grands types structuraux d'ARN polymérases un type "procaryote" et 3 types "eucaryotes".

### • Chez les procaryotes

Il n'existe qu'une seule ARN polymérase, constituée de plusieurs sous unités :  $\alpha$ ,  $\beta$ ,  $\beta'$  et  $\sigma$  (c'est le complexe multimérique). Cette enzyme a deux dénominations : **enzyme cœur** ou **holoenzyme**, selon sa constitution. Comment ?

Le complexe holoenzyme correspond à l'association des sous unités  $\alpha$ ,  $\beta$ ,  $\beta'$  et  $\sigma$  qui initie la transcription. Puis le facteur  $\sigma$  est relargué pour former l'enzyme cœur, constituée des sous unités  $\alpha$ ,  $\beta$  et  $\beta'$ , qui synthétise tous les ARNs.

L'ARN polymérase est donc constitué de :

- ✓ 2  $\alpha$  : Assemblage de l'enzyme, assure la liaison au promoteur.
- ✓ 1  $\beta$  : Assure la liaison des nucléotides.
- ✓ 1  $\beta'$  : Assure la liaison à la matrice d'ADN.
- ✓ 1  $\sigma$  : Assure la reconnaissance du promoteur, initiation de la transcription.

- **Chez les eucaryotes**

Il existe trois ARN polymérase différentes, localisées dans le noyau cellulaire. Chacun de ces ARN est spécifique à la synthèse de certains transcrits. Ces différentes ARN polymérases étudiées *in vitro* travaillent globalement de la même façon.

**Tableau IV.** Différentes ARN polymérase et leur fonction chez les eucaryotes.

Enzymes	Localisation nucléaire	Transcrits	Activité cellulaire
ARN Polymérase I	Nucléole	ARNr (5.8S, 18S, 28 S)	60 à 70 % ARN
Polymérase II	Nucléoplasme	ARNm, ARNsno, LncRNA, miARN et certains ARNsn	10 à 30 %
ARN Polymérase III	Nucléoplasme	ARNt, ARNr 5S et certains ARNsn	~ 10 %

### IV.3. Mécanismes de transcription

Une unité de transcription s'étend toujours d'un site d'initiation de la transcription jusqu'à la région de terminaison de la transcription. Le mécanisme général de la transcription pourra donc être divisé en 03 étapes distinctes :

- **L'initiation de la transcription** : reconnaissance du début de l'unité de transcription ;
- **L'élongation de la chaîne ribonucléotidique** : polymérisation de la chaîne d'ARN ;
- **La terminaison de la transcription** : nécessitant la reconnaissance de la région de terminaison.

#### IV.3.1. Transcription chez les procaryotes

Comme vous l'avez bien compris, ici, l'ensemble des gènes (codant ARNm, ARNt, ARNr) sont transcrit par la même ARN polymérase. Cette enzyme est variable en fonction des espèces de bactéries considérées, mais globalement son organisation et son fonctionnement reste à peu près comparable dans l'ensemble des procaryotes.

##### 1) Initiation

Le promoteur est une séquence d'une centaine de nucléotides située dans la région régulatrice et désignant le début de la transcription. Il est situé en amont du site d'initiation et porte des éléments de séquence reconnus par l'ARN polymérase et déterminant le sens de la transcription. Le promoteur est constitué de courtes séquences conservées d'une unité de transcription à l'autre et appelées séquences consensus (conservées au cours de l'évolution). Ces séquences sont des séquences hexamériques (6nt) riches en T et en A, surtout celle positionnée en -10.

- En région -10 du site d'initiation on trouve la **TATA box** ou **boîte de Pribnow** : «3'-TATAAT-5'», très importante, composée principalement de thymine et d'adénine.
- En région -35 du site d'initiation on trouve : «3'-TTGACA-5'», lieu de fixation du complexe enzymatique.



**Figure 57 :** Promoteurs des gènes procaryotes.

L'initiation correspond à la synthèse de la première liaison phosphodiester réalisé par la sous-unité  $\beta$  qui correspond à la sous-unité catalytique de l'ARN-polymérase.

L'ARN polymérase doit être sous sa forme holoenzyme, donc liée au facteur  $\sigma$ , car la forme holoenzyme est la forme active et compétente pour initier la transcription. Celle-ci reconnaît et se fixe sur la séquence -35 de l'ADN génomique, on obtient un complexe binaire dit fermé (enzyme + ADN). Ensuite, l'ARN polymérase migre et se déplace jusqu'au promoteur -10 (boîte de Pribnow).

L'ARN polymérase est capable d'ouvrir la double hélice de l'ADN génomique sur une courte région (déroulement sur 14 nucléotides). On a donc un passage à l'état simple brin, pour permettre la lecture de l'ADN par l'ARN polymérase, c'est la **boucle de transcription**. L'ADN se re-naturalise spontanément après le passage de l'ARN polymérase et revient à sa forme double brin. Le premier nucléotide mis en place par l'ARN polymérase est très souvent un A ou un G en 5' et un allongement de 4 à 5 nucléotides est nécessaire avant que le facteur sigma ne puisse se détacher.

**Remarque**

L'interaction de cette sous-unité est inhibée par la **rifampicine** qui inhibe ainsi de manière irréversible la transcription de l'ADN, c'est le cas de la tuberculose.

Une mutation dans la sous unité  $\beta$  induit l'apparition de souches bactériennes résistantes à la rifampicine.

## 2) Elongation

L'élongation correspond au déplacement de la boucle de transcription le long de la molécule d'ADN. La région désappariée est alors de 70 paires de bases (Elimination des liaisons hydrogènes). Pendant la transcription, l'ARN forme un court appariement avec le brin matriciel de l'ADN formant une hélice hybride ADN-ARN sur une dizaine de paires de bases.

Il y a incorporation du premier ribonucléotide sous forme triphosphate, puis création de la liaison phosphodiester au niveau du 3'-OH avec un autre ribonucléotide qui va libérer son pyrophosphate, ce qui libère l'énergie nécessaire à la réaction. Donc la synthèse de l'ARN se fait bien dans le sens 5'  $\rightarrow$  3'.

Au fur et à mesure que la boucle progresse au sein de la double hélice d'ADN, des torsions de l'ADN sont induites. Il y a alors nécessité de l'intervention d'une topoisomérase pour supprimer ces tensions et ces torsions : c'est la **gyrase** bactérienne.

Il est à noter que l'ARN polymérase doit rester associée à son brin d'ADN matrice tout au long de son action. Elle est donc associée à des protéines, appelées **facteurs d'élongation**, qui diminuent la probabilité de dissociation de l'ARN polymérase vis-à-vis du brin d'ADN matrice qu'elle est en train de transcrire.

L'ARN polymérase possède deux fonctions de correction pour corriger les erreurs possibles de nucléotides. Elles sont différentes des fonctions correctrices de l'ADN polymérase.

- **Correction pyrophospholytique**

C'est une réaction miroir par rapport à la fonction de polymérisation. C'est la réaction qui catalyse l'enlèvement du ribonucléotide incorrect par incorporation forcée d'un pyrophosphate. Le ribonucléotide qui avait été incorporé était sous forme monophosphate, il se retrouve alors sous forme triphosphate et est retiré de la chaîne en cours de synthèse.

- **Correction hydrolytique**

L'enzyme « recule » de quelques ribonucléotides, clive l'ARN pour éliminer sur quelques ribonucléotides la séquence qui contient l'erreur, et de par son activité polymérasique reprend la synthèse d'ARN. (Ce n'est pas une fonction exonucléasique). Ces fonctions de corrections sont différentes de la fonction d'édition (exonucléasique) de l'ADN polymérase.

**Remarque**

L'élongation peut être inhibée par des aminosides ou amino-glucosides.

### 3) **Terminaison/Maturation des transcrits primaires**

L'ARN polymérase reconnaît sur l'ADN génomique en cours de lecture des séquences particulières : les signaux de terminaison de la transcription, qui aboutissent à l'arrêt de la transcription, à la libération de la molécule d'ARN et à la dissociation de l'enzyme vis-à-vis de l'ADN matrice en cours de lecture.

Chez les bactéries, il y a deux mécanismes de terminaison de la transcription, l'un qui fait intervenir une protéine spécifique appelé *facteur de terminaison Rho*, l'autre qui ne nécessite pas l'intervention de facteur extérieur, ce qui conduit à distinguer deux types de terminateurs :

- **Mécanisme direct : Rho indépendant**

Sur le brin d'ADN génomique en cours de lecture, il y a une séquence particulière qui correspond à une séquence palindromique imparfaite. L'ARN polymérase lit cette séquence et la transcrit en ARN. Ce palindrome imparfait adopte une structure secundo-tertiaire dite en épingle à cheveux, avec établissement local de liaisons hydrogène intra-chaînes, ce qui stabilise la structure.

Cette séquence est suivie en 3' par une succession de nombreux uraciles. La structure en épingle à cheveux entraîne l'instabilité de l'hybride ARN-ADN et provoque donc une dissociation de l'ARN polymérase. On a ainsi un arrêt de la progression de l'ARN polymérase, et de ce fait la libération de l'ARN synthétisé et la dissociation de l'enzyme de l'ADN génomique.

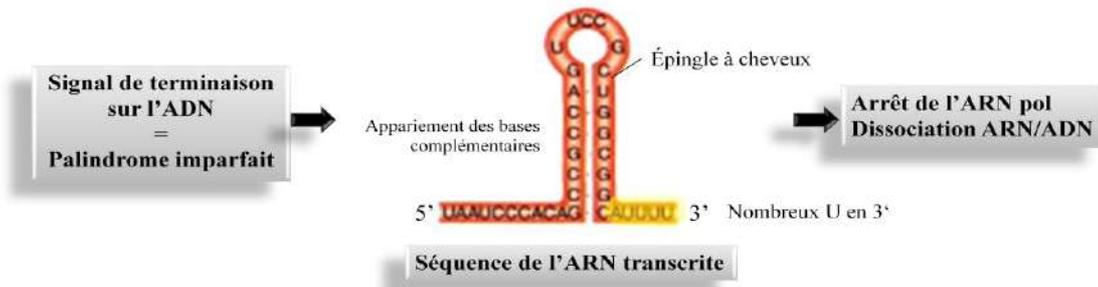


Figure 58 : Mécanisme direct (Rho indépendant).

• **Mécanisme indirect avec une protéine Rho à activité ATPasique**

Ce mécanisme se met en place lorsque la transcription du signal de terminaison n'est pas suffisamment instable pour arrêter la transcription. Sur certains ARN bactériens, au niveau de la région 5' de l'ARN en cours de synthèse, on a une séquence qui est reconnue par des protéines appelées *protéines Rho*, à activité ATPasique.

Le site d'enroulement de l'ARN autour de Rho est une région d'environ 70 nucléotides (jusqu'à 100 nucléotides), riche en cytosine et pauvre en guanine, appelée « *ruh* » pour *rho utilization site* située en amont de la séquence appelée "terminateur".

Elles s'agglutinent sur l'ARN en cours de synthèse après reconnaissance de cette séquence. Par hydrolyse de molécules d'ATP, ces protéines Rho ont à leur disposition l'énergie suffisante pour se déplacer elle-même le long de la molécule d'ARN vers l'extrémité 3' pour rattraper la boucle de transcription et l'ARN polymérase.

A un moment donné, l'ARN polymérase lit et transcrit le signal de terminaison ce qui induit la formation d'une structure en *épingle à cheveux*. Pour ces gènes Rho-dépendants, l'hybride ADN/ARN n'est pas suffisamment instable pour induire l'arrêt de la transcription et la dissociation. La structure en épingle à cheveux ralentit ou provoque une pause de la boucle de transcription. Cet événement laisse le temps au facteur Rho de rejoindre le complexe et d'entraîner la dissociation et l'arrêt de la transcription.

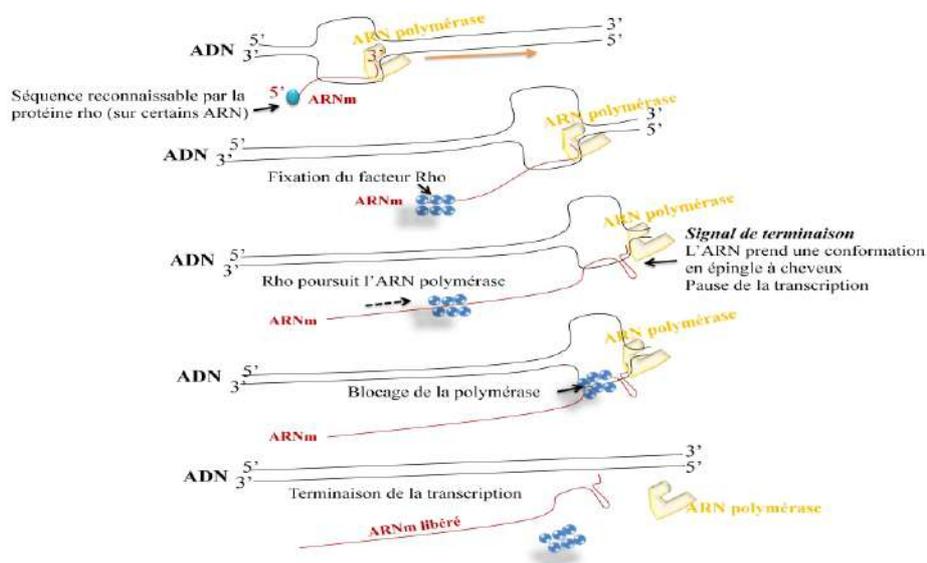


Figure 59 : Mécanisme indirect avec une protéine Rho à activité ATPasique.

Le **transcrit primaire** correspond à l'ARN non mature qui nécessite une maturation sous forme de clivages ou de modifications de bases. Cette maturation n'est pas obligatoire. Après maturation on obtient de l'ARN mature. Généralement, il y a peu de modifications pour les ARNm ; beaucoup d'ARNm sont traduits en protéines alors qu'ils sont encore transcrits.

Le transcrit primaire code soit pour un produit, on parle d'**ARN monocistronique**, soit plusieurs produits, on parlera alors d'**ARN polycistronique**.

#### IV.3.2. Transcription de l'ADN chez les eucaryotes

De nombreuses différences existent par rapport à la transcription chez les procaryotes. La transcription se déroule dans le noyau cellulaire. La chromatine doit au préalable avoir été décompactée (euchromatine) pour permettre à la machinerie protéique d'accéder à l'ADN. Contrairement aux procaryotes, l'ARN produit par la transcription n'est pas directement utilisable par les ribosomes pour la traduction et devra subir plusieurs étapes de maturation post-transcriptionnelle.

Enfin, l'ARN polymérase ne peut se fixer seule au promoteur du brin matrice d'ADN : chez les eucaryotes, elle nécessite des facteurs de transcription, protéines qui servent d'intermédiaires à la liaison de l'ARN polymérase sur le promoteur. Cette association entre le promoteur, les facteurs de transcription qui y sont liés et l'ARN polymérase forme le complexe d'initiation de la transcription, nécessaire au commencement de la transcription.

Le processus de transcription comporte les mêmes phases que chez les procaryotes : initiation, élongation et terminaison.

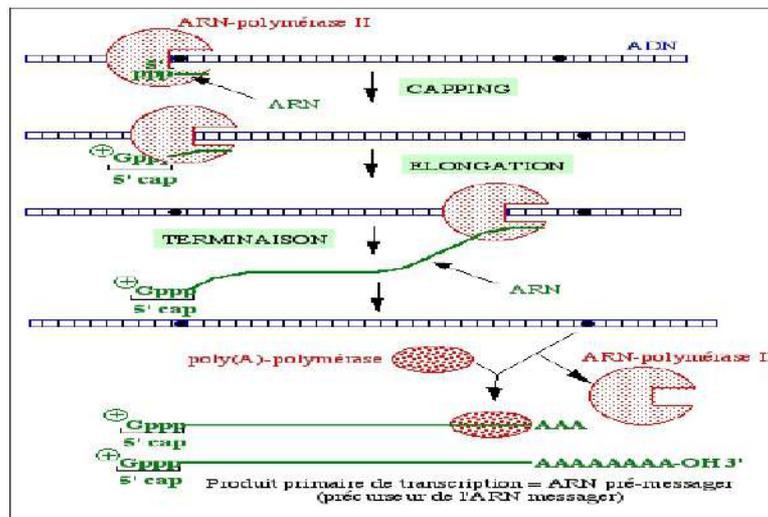


Figure 60 : Transcription de l'ADN chez les eucaryotes.

##### 1) Initiation

Comme pour les gènes procaryotes, on retrouve une région promotrice qui fait intervenir plusieurs régions consensus, conservée au cours de l'évolution (même nomenclature).

- La plus importante est la « boîte **TATA** » (promoteur), qui est l'équivalente de la boîte Pribnow bactérienne. Elle se positionne vers -30 à -25 nucléotides (séquence hexamérique) du site de démarrage de la transcription (noté +1) et est riche en A et en

T. Cette boîte TATA représente une séquence promotrice où vient se fixer un complexe d'initiation de la transcription qui fait intervenir l'ARN polymérase II ainsi que d'autres protéines : les facteurs généraux de la transcription.

- La « boîte **CAAT** » (facultative), contenant de la cytosine, est située vers -120 à -80 nucléotides du site de démarrage de la transcription.
- La « boîte **GC** » (facultative également), riche en guanine et en cytosine, peut être présente entre la boîte CAAT et la boîte 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 ».

Contrairement à ce qui se passe chez les procaryotes, l'ARN polymérase II des eucaryotes ne reconnaît pas seule le promoteur proximal. Elle effectue ce travail en compagnie de nombreux co-facteurs protéiques formant avec elle un complexe d'initiation. Ces facteurs sont notés TFII pour "*Transcription Factor for RNA polymerase II*".

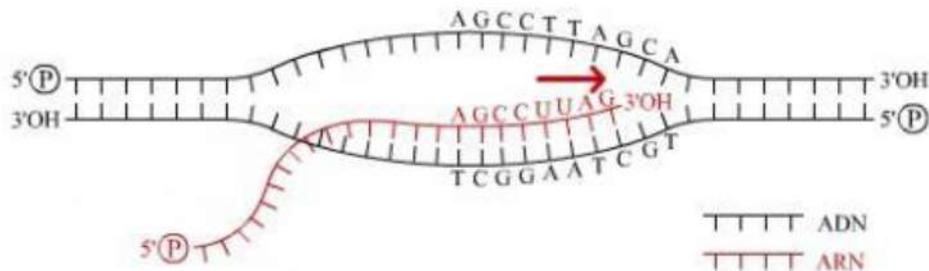
Les TF sont les facteurs généraux de la transcription, qui s'associent dans un ordre précis pour initier la transcription du gène, contrôlée par la boîte TATA au niveau de la région promotrice. Dans cette région promotrice, on trouve aussi les séquences d'amont, qui varient selon les promoteurs des gènes considérés. Les plus fréquentes sont la CCAAT box et la GC box. Elles régulent la fréquence d'initiation de la transcription.

Chez les eucaryotes, l'ARN polymérase doit faire face aux structures chromatiniennes, puisque l'ADN est associé aux nucléosomes. Elle a donc besoin de faire partie d'un complexe où elle est associée à ces facteurs généraux de la transcription. Certains de ces facteurs ont un rôle particulier :

- ✓ Le facteur **TFII D**, sa sous-unité TBP reconnaît et se fixe à la boîte TATA, ce qui induit un début de distorsion de la molécule d'ADN. Ce mécanisme induit le recrutement d'autres facteurs TFII et de l'ARN polymérase II, dans une certaine conformation.
- ✓ **TFII A** interagit avec l'ADN en amont de la TATA box.
- ✓ **TFII B** interagit avec l'ADN en aval de la TATA box au niveau du site d'initiation.
- ✓ **TFII F** agit lors de l'élongation.
- ✓ Le facteur **TFII H** : il a deux fonctions :
  - Rôle d'hélicase : favorise la dénaturation locale de l'ADN, en éliminant les liaisons hydrogène. Elle nécessite de ce fait de l'ATP. Ce complexe peut commencer à initier la transcription, donc synthèse d'ARN sur une courte distance.
  - Rôle de protéine-kinase : capable de phosphoryler l'ARN polymérase II, cela va induire un changement de conformation de l'enzyme et la dissociation des facteurs généraux (plus besoin car initiation déjà faite).

## 2) Elongation

L'élongation fait intervenir des facteurs protéiques d'élongation, qui permettent et facilitent la progression ainsi que le maintien le plus longtemps possible de l'ARN polymérase II sur l'ADN matrice. Des topoisomérases interviennent également afin de permettre l'élimination des surenroulements positifs créés par l'avancée de l'ARN polymérase II. Des facteurs protéiques de correction vont aussi agir afin de stimuler la fonction de correction hydrolytique de l'ARN pol II. Chez les eucaryotes, la vitesse de transcription est d'environ 40 nucléotides par seconde.



**Figure 61** : Phase d'élongation de la transcription.

L'ARN polymérase lit le brin patron (ou brin anti sens), qui est complémentaire du brin codant (ou brin sens), depuis son extrémité 3' vers son extrémité 5'. Le brin d'ARN néosynthétisé est donc identique au brin codant d'ADN, aux uraciles et riboses près.

## 3) Terminaison

L'ARN polymérase II est également équipée de facteurs protéiques de terminaison. Elle 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**, etc.). Elle continue sa transcription un peu après ces motifs avant d'arrêter son travail de transcription et libérer l'ARNpm qu'elle vient d'assembler. La transcription proprement dite est terminée mais l'ARN obtenu n'est pas fonctionnel pour autant et doit subir 03 étapes de maturation.

## 4) Maturation des transcrits spécificité eucaryote

L'élongation des ARNm par l'ARN polymérase II s'accompagne d'une maturation du transcrit. Certains événements de cette maturation s'effectuent en même temps que la transcription (dans le noyau cellulaire). La maturation comprend à la fois la modification covalente des extrémités 3' et 5' de l'ARNm et le retrait des introns grâce au mécanisme d'épissage.

La séquence transcrite d'un gène est composée de séquences introniques et exoniques. De manière générale, la plupart des séquences exoniques seront exprimées, donc traduites en séquence protéiques, sauf les régions 5'UTR et 3'UTR qui ne seront pas traduites. Tandis que les introns seront transcrits mais non traduits. Lors de la transcription, toute la séquence est transcrite en ARN pré-messager (ou précurseur).

C'est seulement après maturation qu'on aura un ARN messenger mature, où il y aura eu des modifications covalentes des extrémités en 3' par l'ajout d'une queue polyA

(modification résultant d'une polyadénylation) et en 5' par l'ajout d'une coiffe (ou cap) ainsi que le retrait des séquences introniques.

- **Ajout de la coiffe en 5' (= cap)**

La coiffe est présente sur tous les ARNm eucaryotes (qui proviennent uniquement de la transcription par l'ARN polymérase II). Il s'agit d'une modification covalente qui se fait dans le noyau au cours de l'élongation par l'ARN polymérase II.

Cette coiffe se met en place dès le début de la transcription, après 20 à 30 ribonucléotides incorporés dans l'ARN en cours de synthèse, en 5' du transcrit primaire. Il y a d'abord élimination d'un phosphate en 5' du premier nucléotide triphosphate par une phosphatase puis soudure d'une 7-méthyl-guanine (GMP méthylée sur l'azote en 7) par une guanyle-transférase et une méthyle-transférase. Il y a ensuite méthylation en 2' sur les riboses des 2 premiers nucléotides dans la plupart des ARNm.

Sur cette coiffe viennent se fixer des protéines spécifiques qui jouent un rôle fonctionnel important. Les protéines qui se lient à la coiffe constituent le "cap binding complexe" (CBC). La coiffe a pour rôle de permettre :

- ✓ La distinction (au niveau de la cellule) entre les ARNm et les autres ARN ;
- ✓ L'export dans le cytoplasme de l'ARNm pour traduction ;
- ✓ Le recrutement de la petite sous-unité du ribosome à l'extrémité 5' permettant l'initiation de la traduction.

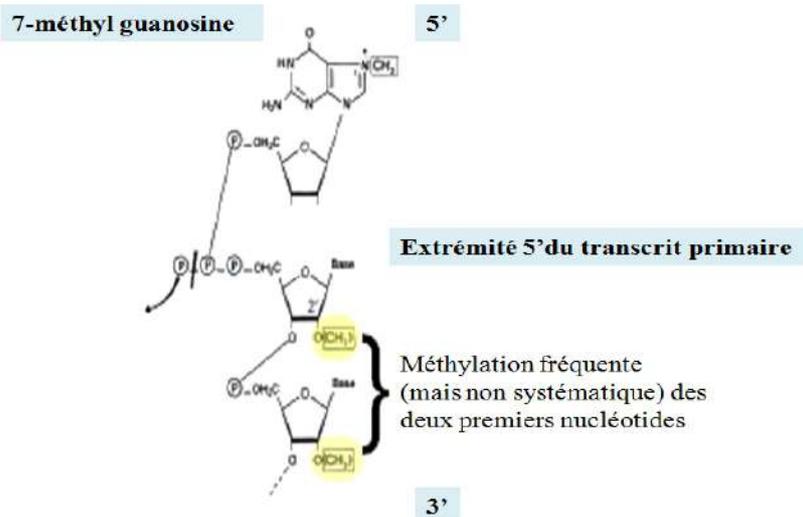


Figure 62 : Ajout de la coiffe en 5'.

- **La queue poly(A) coupure et polyadénylation des ARNm**

Pour arrêter la transcription chez les eucaryotes, il existe une séquence signal de terminaison de la transcription aussi appelée « séquence signal de polyadénylation ». C'est une séquence conservée au cours de l'évolution, qui se trouve à la fin de la séquence génique.

L'ARN polymérase II lit cette séquence et la transcrit en ARN. On aura donc une séquence dans le sens 5' qui sera "AAUAAA". Donc cette séquence provient de l'ADN génomique.

Cette séquence sur l'ARNm induit le recrutement de protéines spécifiques qui s'associent à la séquence transcrite du signal de terminaison de transcription. Ceci induit ensuite l'activité d'une ribonucléase spécifique qui clive la chaîne du pré-ARNm en cours de synthèse environ après 10 à 15nt en 3' de la séquence "AAUAAA".

Très rapidement, la polyadénylate polymérase ajoute par polymérisation une succession de A (200 à 250) en 3'-OH de cet ARNm : c'est la queue poly(A). Elle ne provient pas de la séquence présente sur l'ADN génomique.

Sur ces queues poly(A) peuvent aussi se fixer des protéines : Les poly(A)-Binding proteins. Elles jouent plusieurs rôles :

- ✓ Dans la reconnaissance des ARNm ;
- ✓ Dans l'exportation hors du noyau ;
- ✓ Probablement un rôle de protection de l'ARNm lors de la traduction.

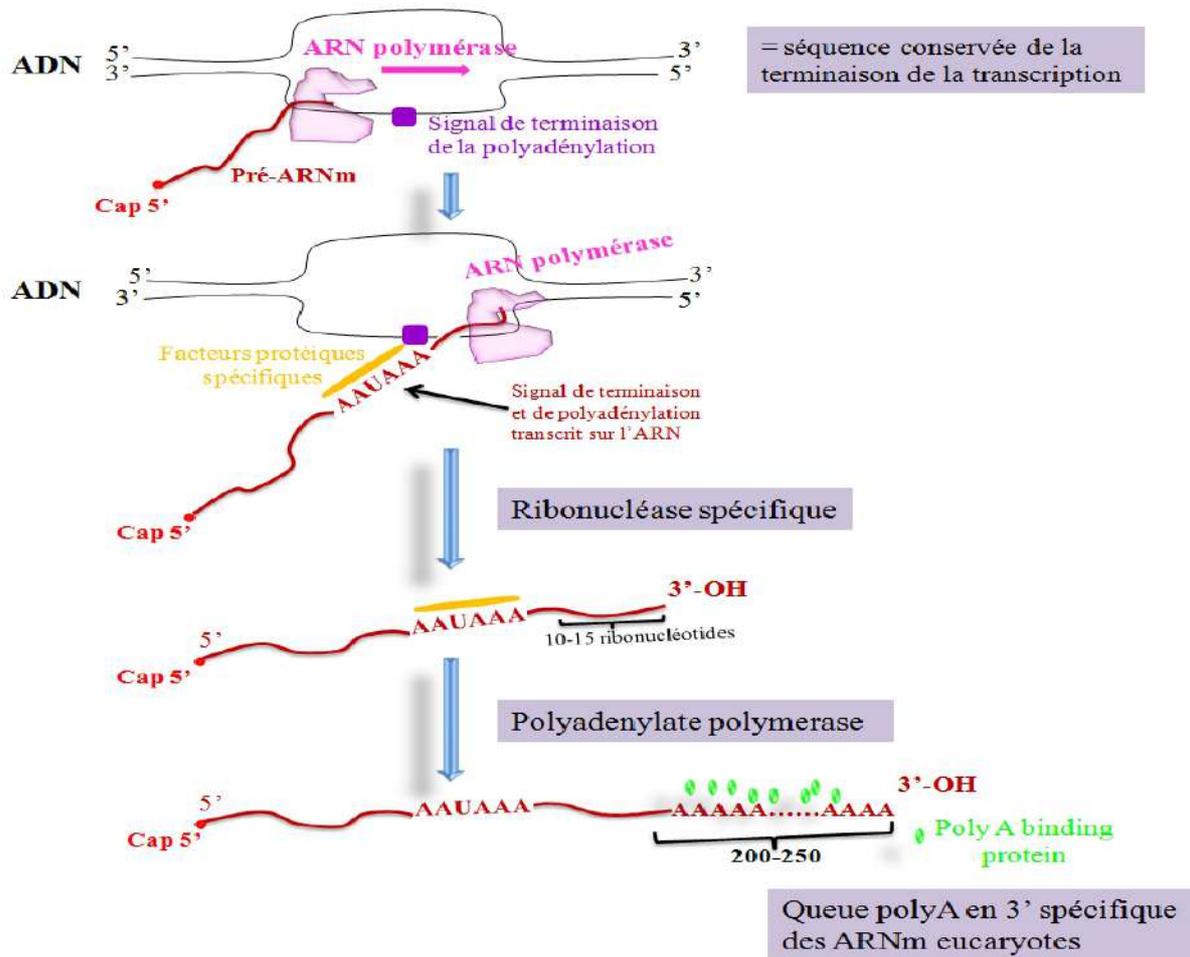


Figure 63 : Ajout de la queue poly(A) en 3'.

#### • Epissage

Il s'agit d'un mécanisme qui aboutit à l'excision des introns et à la ligature des exons, pour aboutir à un ARN fonctionnel et mature. Il existe plusieurs mécanismes dans la cellule

qui peuvent toucher les différentes classes : on ne parlera ici que de celui spécifique aux ARNm. La séquence intronique des ARNm contient :

- ✓ En 5' une séquence GU : site donneur d'épissage ;
- ✓ En 3' une séquence AG : site accepteur d'épissage ;
- ✓ Un A de branchement à environ 30 nt du site accepteur (3'). Le A de branchement est trouvé de manière fréquente, mais non obligatoire. Pour le Spliceosome, le A de branchement est nécessaire.

Sur ces séquences vient se fixer un complexe multimérique (qui fait intervenir plusieurs partenaires), appelé le spliceosome. Il est constitué de 5 snRNP (small nuclear ribonucléoprotéines particules). Les snRNP sont constitués à la fois d'ARNsn et de protéines.

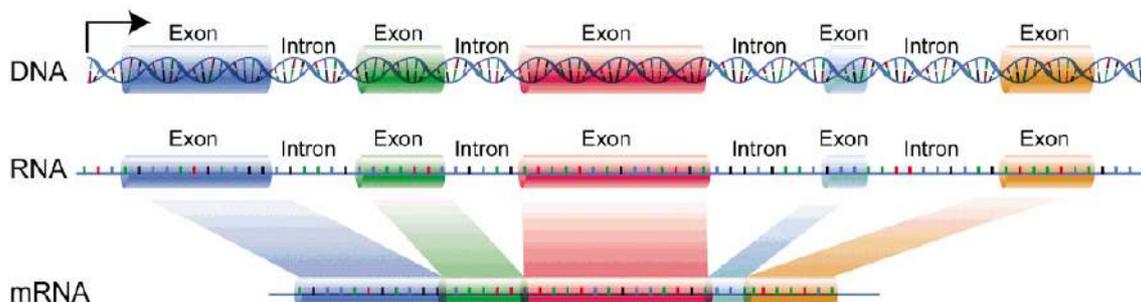
On trouve 5 snRNP : U1, U2, U4, U5, U6, (appelées U car riche en uracile). Certaines ont un rôle bien précis :

- ✓ U1 se place sur le site donneur (5') ;
- ✓ U5 se place sur le site accepteur (3') ;
- ✓ U2 se place sur le A de branchement ;

L'élimination d'un intron est réalisée après deux réactions de trans-estérification. C'est un mécanisme nécessitant de l'énergie sous forme d'ATP :

- ✓ Le A de branchement est reconnu en premier par 2 protéines BBP et U2AF ;
- ✓ Puis U2 déplace ces deux protéines et forme un appariement de base au niveau du A de branchement ;
- ✓ U1 se positionne sur le site 5' et U5 sur 3' ;
- ✓ Les autres snRNP arrivent et entraînent la formation d'un lasso ;
- ✓ Clivage des sites d'épissage 5' et 3' ;
- ✓ Ligature des exons et libération des introns sous forme de boucles.

Les snRNP sont dissociées de l'intron et sont recyclées dans un nouveau spliceosome. La séquence d'ARN (intron) sous forme de lasso est dégradée par une nucléase cellulaire. Le spliceosome est très consommateur d'énergie donc à chaque étape, une molécule d'ATP est hydrolysée.



**Figure 64 :** Excision des introns (épissage).

## Mécanismes moléculaires

### (Traduction des ARN messagers)

#### Objectifs spécifiques

Au terme de ce cours qui est la traduction des ARNm, vous devez être capable de :

- Connaître et pouvoir déterminer l'intérêt de la régulation de l'expression des gènes ;
- Distinguer les différents niveaux de l'expression des gènes ;
- Comprendre les mécanismes de la régulation génique ;
- Pouvoir décrire les mécanismes de régulation de l'expression de gènes chez les procaryotes et chez les eucaryotes.

#### V.1. Définition

En biologie moléculaire, la **traduction** est un processus permettant la synthèse d'une chaîne polypeptidique (protéines) par les ribosomes, à partir de l'information génétique contenue dans les brins d'ARN messagers (ARNm).

Elle s'accomplit (lieu) dans le cytoplasme des cellules au niveau des ribosomes, cœur de la machinerie de synthèse des protéines cellulaires où il y'a interprétation de l'information contenue dans l'ARNm sous forme de codons (triplets de nucléotides), qu'il traduit en acides aminés assemblés dans la protéine, selon l'ordre donné par les codons portés par l'ARNm.

Cette synthèse protéique s'effectue de l'extrémité N-terminale de la protéine vers l'extrémité C-terminale.

#### V.2. Rôle de la traduction

La traduction assure l'expression des gènes portés par l'ADN, et constitue la deuxième grande étape de ce processus après la transcription.

#### V.3. Acteurs de la traduction

La traduction nécessite l'intervention d'un grand nombre d'acteurs dont les plus importants sont :

##### 1) ARNm

S'il existe 20 acides aminés différents dans les **protéines** et quatre bases nucléotidiques différentes dans l'ARNm, on en déduit qu'il faut un enchaînement de trois bases pour définir l'ensemble des 20 acides aminés. Le codon est donc la succession de 3 nucléotides.



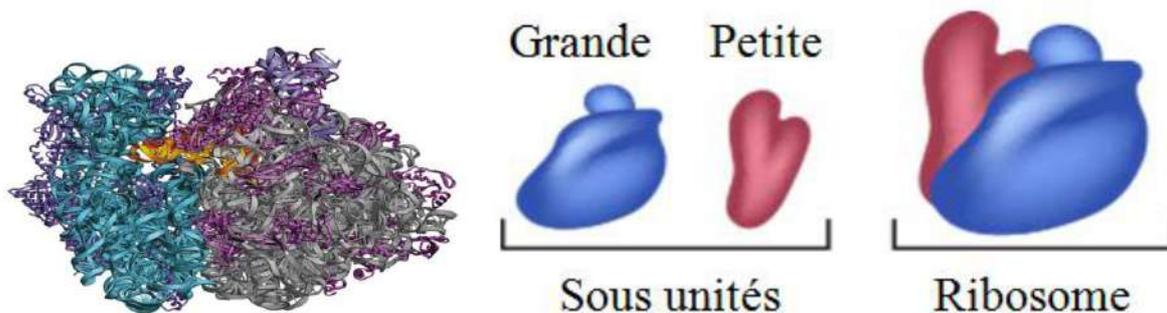
Figure 65 : Représentation d'un ARN messager eucaryote.

## 2) Ribosomes

Plus petites structures cellulaires : visibles au microscope électronique seulement. Ce sont des organites intra-cellulaires situés dans le cytoplasme. Ils constituent l'usine de fabrication des protéines. Les ribosomes sont constitués d'ARN ribosomiques (ARNr) et de protéines et sont structurés sous forme de deux sous-unités une grande et une petite associées entre elles, que ce soit chez les procaryotes ou chez les eucaryotes. Leur taille est définie en unité Svedberg.

- 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.

L'enchaînement des ribosomes sur l'ARNm forme le **polysome**, il permet d'augmenter l'efficacité de la traduction. La distance minimale qui sépare deux ribosomes est de 100 nucléotides. Il faut savoir que la grande sous unité du ribosome, portent deux sites accepteurs pour les ARNt chargés de leur acide aminé (site P et site A) en plus du site E pour l'expulsion des ARNt.



**Figure 66** : Structure d'un ribosome.

## 3) ARN de transfert (ARNt)

L'ARNt est la macromolécule dont la structure secondaire est en forme de trèfle à 3 feuilles et une structure tertiaire en forme de 'L' inversé. Lors du mécanisme de traduction, il y a un appariement antiparallèle entre l'ARNm et l'ARNt : reconnaissance codon-anticodon au niveau de la boucle de l'anticodon.

Chaque ARNt est caractérisé par son anticodon, de ce fait un ARNt ne transporte pas n'importe quel acide aminé: ça dépend de l'anticodon par exemple : l'ARNt AAA transporte toujours l'acide aminé Phe et l'ARNt GAU transporte toujours l'acide aminé Leu.

Cependant, certains ARNt peuvent également reconnaître plusieurs codons. Car la complémentarité entre les nucléotides occupant la troisième place dans le triplet, n'est pas stricte "wobble position". Il existe donc une flexibilité dans l'appariement des bases en position 3 du codon et en position 1 de l'anticodon, cette flexibilité s'appelle le **wobble**.

En faisant le total de tous les Wobble, on arrive à un total de 32 ARNt et non 61, ce qui représente une économie pour la cellule. De plus, il semblerait que ces appariements atypiques sont plus lâches et favorisent une accélération de la traduction.

**Tableau V** : Représentation des wobble.

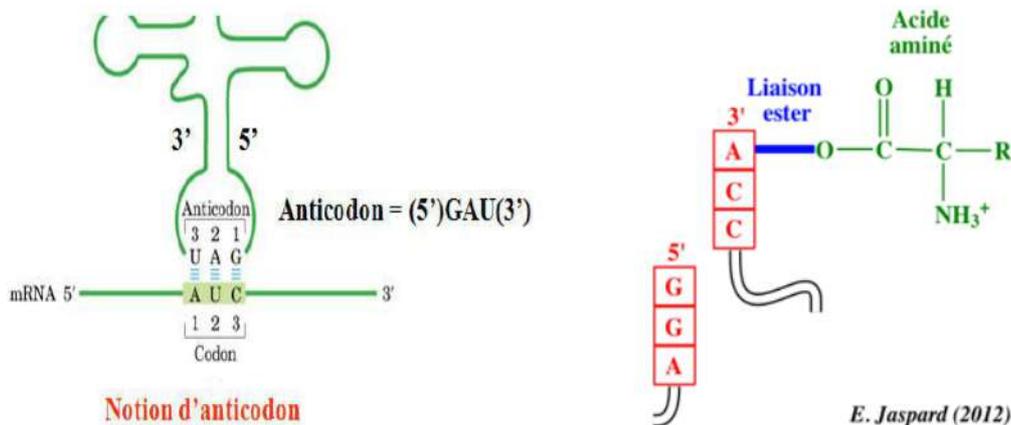
Codon en position 3	Anticodon en position 1
U	A, G, I
C	G, I
A	U, I
G	C, U

**Remarque**

**I** est l'inosine, base modifiée, constitué de ribose et d'hyoxanthine. Ce sont les propriétés chimiques de l'inosine qui expliquent que cette base particulière peut s'apparier indifféremment avec U, C ou A mais pas G.

L'inosine se retrouve fréquemment dans l'ARNt, en particulier au niveau de la première position de l'anticodon, où elle permet à un même ARNt de s'apparier à plusieurs codons synonymes. Dans l'ARN, l'inosine est formée par désamination de l'adénosine. Cette réaction est catalysée par l'adénosine désaminase.

Les ARNt possèdent également un bras pour l'acide aminé qui le fixe en 3' (CCA) sur le ribose, il s'agit d'une liaison covalente : liaison ester riche en énergie. Les acides aminés ne vont ainsi par arriver libre sur le ribosome mais associés à leurs ARNt respectifs. On trouve 40 à 60 ARNt différents par cellule, il existe donc plusieurs ARNt différents pour un acide aminé, on les appelle **ARNt iso-accepteurs**.



**Figure 67** : ARNt et sa liaison avec l'acide aminé ainsi que l'ARNm.

#### 4) Aminoacyl-ARNt synthétase

C'est une famille de ligases. Elles catalysent l'estérification des acides aminés sur l'extrémité 3' des ARNt. Ce sont des protéines essentielles, conservées chez tous les êtres vivants.

Chez la plupart des organismes, il existe une aminoacyl-ARNt synthétase pour chacun des acides aminés. Hormis pour la sélénocystéine, liée à son ARN de transfert par un mécanisme indirect passant par la sérine. Chacune de ces enzymes reconnaît un acide aminé et un ou plusieurs ARNt iso-accepteurs. Ce sont celles-ci qui garantissent que l'acide aminé qui est ainsi estérifié à l'extrémité de l'ARNt correspond bien au bon anticodon.

#### 5) Acides aminés

Comme pour la réplication de l'ADN ou sa transcription où il y'a nécessité de la présence de nucléotides à incorporer, là aussi pour la traduction de l'ARNm, il est nécessaire qu'il y ait cette fois ci non pas des nucléotides libres mais des acides aminés libres afin qu'ils soient incorporé progressivement à la chaîne polypeptidique en formation transporté par les ARNt selon un ordre bien précis dicté par le code génétique.

L'acide aminé est attaché au bon ARNt par l'enzyme dite « aminoacyl-ARNt 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. Ces enzymes sont au nombre de 20 dans la cellule, autant qu'il y a d'acides aminés qui rentrent en compte dans la traduction.

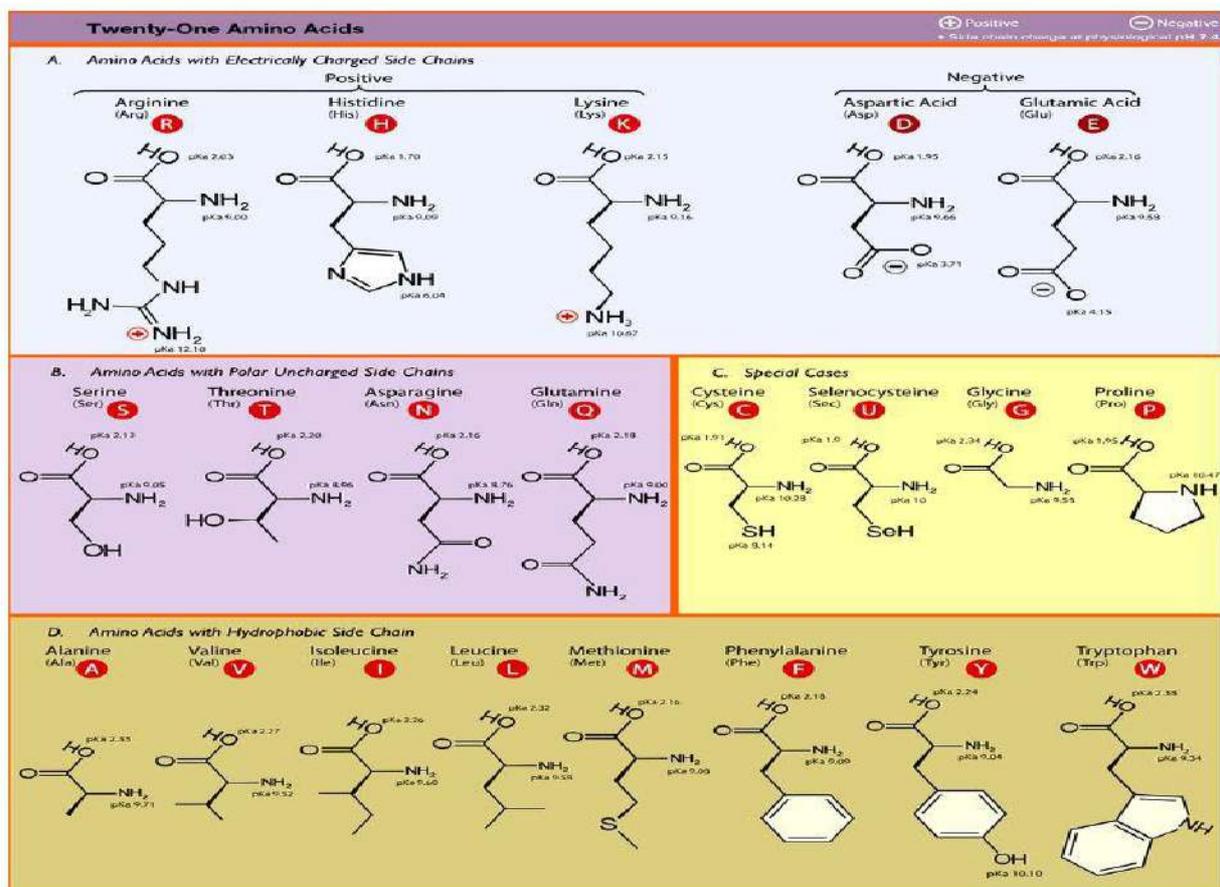


Figure 68 : Structure chimique des acides aminés.

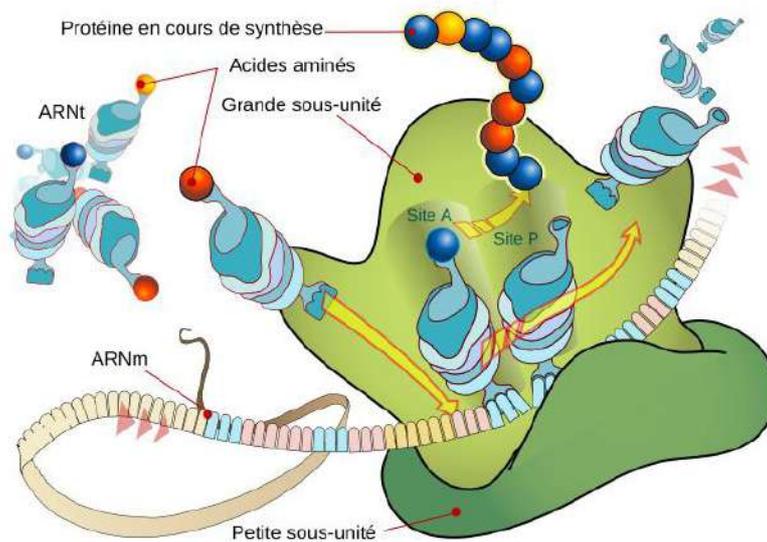
#### V.4. Mécanisme de la traduction

La traduction a lieu au niveau des ribosomes, ce processus est assez **complexe**. Il est **continu** mais il est possible d'y distinguer trois étapes: l'initiation, l'élongation et la terminaison. Au delà des aminoacyl-ARNt synthétases, des ARNt, les ribosomes, et les ARNm cette synthèse nécessite plusieurs protéines appelées : facteurs d'initiation, d'élongation et de terminaison, tandis que l'énergie est fournie par la molécule de GTP.

Les structures différentes dans les ARNm procaryotes et eucaryotes impliquent des voies différentes pour ces événements.

##### V.4.1. Traduction chez les procaryotes

Chez les procaryotes, transcription et traduction sont simultanées alors que chez les eucaryotes, la traduction s'effectue dans le cytoplasme à partir de l'ARNm mature.



**Figure 69** : Processus de traduction de l'ARNm en protéine.

##### 1) Initiation

Afin qu'il puisse y avoir initiation de la traduction, la petite sous unité du ribosome (30S) doit reconnaître le début de la séquence codante sur l'ARNm, il utilise des signaux en amont entre -8 et -13 du codon initiateur (AUG) qui correspond à la **séquence de Shine-Dalgarno** ou **RBS** pour **Ribosome Binding Site** (dont son consensus est AGGAGG) situé sur l'extrémité 5' de l'ARNm. 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.

Cette étape implique donc une seule des deux sous-unités du ribosome qui doivent être au préalable dissociées grâce au facteur d'initiation IF3 associé à la petite sous unité 30S.

Les bactéries nécessitent un acide aminé particulier pour l'initiation ; cet acide aminé est la **méthionine formylée** sur sa fonction amine (la méthionine nécessite une **formylation** sur l'extrémité  $\text{NH}_2$  = ajout d'un formyl) pour former la N-formylméthionine (f-Met), c'est un phénomène pré-traductionnel. Cette formylation est

réalisée par un cofacteur, la vitamine B9 (ou tétrahydrofolate) qui reconnaît l'ARNt caractéristique et responsable du transport de la f-Met.

L'initiation de la traduction est permise grâce à la présence de **facteurs d'initiation** (IF pour *Initiation Factor*) qui sont :

- **IF 1** est le facteur de dissociation du ribosome 70S (il va donc se fixer dans le site A de la sous-unité 30S du ribosome et permettre le recrutement direct de l'ARNt de démarrage au site P) ;
- **IF 2** est un facteur assurant la fidélité de reconnaissance entre l'ARNt et l'acide aminé. Il possède également une activité GTPasique (c'est-à-dire d'hydrolyse du GTP (**guanosine triphosphate**), coenzyme de transfert de groupements phosphate.) ;
- **IF 3** est un facteur qui assure la dissociation préalable des sous-unités 30S et 50S et vérifie que l'interaction entre le codon de démarrage AUG et l'anticodon de l'ARNt est correcte (facteur anti-réassociation).

Une fois la petite sous unité 30S et l'IF3 fixé sur l'ARNm, il y'a recrutement de l'ARNt chargé de la méthionine formylé (fMet). L'ARNm et le ribosome sont disposés de telle sorte que l'ARNt<sup>f-Met</sup> initiateur et le codon d'initiation se trouvent positionné au site P du ribosome.

En parallèle au recrutement de l'ARNt<sup>f-Met</sup>, il y'a la fixation de IF1 sur le site A de la sous unité 30S du ribosome afin de favoriser la fixation de cet ARNt<sup>f-Met</sup> au site P. IF2 vient s'associer aux autres facteurs d'initiation afin de veiller à la fidélité de reconnaissance entre l'ARNt et l'acide aminé. Il possède une activité GTPasique (c'est-à-dire d'hydrolyse du GTP).

Lorsque le complexe ternaire ARNm/ARNt<sup>f-Met</sup>/sous unité 30S c'est formé et stabilisé. La sous unité 30S change de conformation libérant ainsi le IF3 ce qui permet l'association de la grande sous unité 50S. Hydrolyse du GTP lié à IF2 et libération des autres facteurs d'initiation du complexe.

## 2) Elongation

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 mais aussi grâce à la présence de **facteurs d'élongation** (EF pour *Elongation Factor*) : **EF-Tu** ; **EF-Ts** et **EF-G**.

Après la lecture du codon d'initiation, le ribosome va se déplacer séquentiellement le long de l'ARNm dans le sens 5' → 3'. Le ribosome se décale exactement d'un triplet de nucléotides et libère un site pour un nouvel ARNt chargé et ainsi de suite.

Le 2<sup>ème</sup> ARNt vient avec l'acide aminé n°2 dans le site A (site acide aminé) de la grande sous unité complexé au facteur d'élongation EF-Tu lié au GTP. Lors de l'accommodation correcte, ce facteur d'élongation hydrolyse le GTP en GDP. C'est le codon n°2 placé sur l'ARNm après le codon AUG, qui détermine donc le choix du 2<sup>ème</sup> anticodon c'est-à-dire du 2<sup>ème</sup> ARNt, donc du 2<sup>ème</sup> acide aminé. Cette étape correspond à l'accrochage du nouvel aminoacyl-ARNt (aa-ARNt). EF-Tu-GDP quitte par la suite le ribosome puis sera régénéré en EF-Tu-GTP grâce au facteur d'échange EF-Ts.

Les acides aminés sont ajoutés les uns à la suite des autres par des liaisons peptidiques entre  $COOH$  de l'acide aminé n°1 (f-Met) et  $NH_2$  de l'acide aminé n°2 porté par l'ARNt n°2, grâce à une activité peptidyl-transférase de l'ARNr 23S portée par le ribosome qui est appelé ribozyme. Puis les acides aminés sont détachés des ARNt. Les ARNt libres (sans acide aminé) sont expulsés du ribosome au niveau du site E (Exit site).

Toutefois, le  $COOH$  du 1<sup>er</sup> acide aminé n'était pas libre puisqu'il était déjà engagé dans une liaison avec l'ARNt. La formation de la liaison peptidique donnant le dipeptide (accroché maintenant sur le 2<sup>ème</sup> ARNt) et le détachement du 1<sup>er</sup> ARNt se font simultanément en entraînant l'élimination d'une molécule d'eau. A ce stade, il y'a donc eu formation d'un dipeptide qui est logé dans le site acide aminé. Il est porté par l'ARNt n°2.

Par la suite, le ribosome va avancer d'un cran sur l'ARNm dans la direction 5' vers 3' (un cran signifie 3 nucléotides ou encore codon). Un nouveau codon (codon n°3) se trouve donc maintenant en face du site acide aminé. Simultanément, l'ARNt n°2 qui portait le dipeptide (formé de la f-Met et de l'acide aminé n°2) est passé du site A (site acide aminé) au site P (site peptidique). Il a changé de loge, d'où le nom de translocation mais il se trouve toujours en face de son codon (n°2). Cette étape nécessite l'intervention d'un troisième facteur d'élongation, EF-G et l'hydrolyse d'une seconde molécule de GTP.

Ces différentes parties de l'élongation vont se succéder au fur et à mesure. Au cours de celles-ci, la peptidyl-transférase va, comme son nom l'indique, couper, transférer, souder. A chaque fois le peptide s'allonge ainsi, d'un acide aminé.

### 3) Terminaison

La fin de la traduction se produit lorsque le ribosome en avançant sur l'ARNm trouve un codon stop : UAA, UAG ou UGA. Ces codons ne codent pour aucun acide aminé.

Il est à savoir qu'il n'existe aucun ARNt ayant un anticodon complémentaire à l'un de ces 3 codons 'non-sens'. Il se produira alors une coupure entre le dernier ARNt et la chaîne peptidique. La liaison ester unissant ce dernier ARNt au dernier acide aminé est hydrolysée, libérant ainsi la chaîne peptidique.

Les facteurs protéiques de relargage RF (*Release Factor*) de classe I (RF1 et RF2) reconnaissent directement les codons stop et vient occuper le site A, transforme la peptidyl transférase en une hydrolase afin de libérer la chaîne polypeptidique ainsi que le dernier ARNt alors que le facteur de classe II (RF3) a une activité GTPase qui a pour but de stimuler le relargage des facteurs de classe I. Le facteur de recyclage du ribosome (RRF pour *Ribosome Recycling Factor*) mime la structure d'un ARNt et libère le ribosome de l'ARNm en dissociant les deux sous-unités.

#### Remarque

Chez les procaryotes, les facteurs de terminaison sont au nombre de 3 (RF1, RF2, RF3). RF1 reconnaît les codons UAG et UAA, et RF2 UGA, UAA.

### V.4.2. Traduction chez les eucaryotes

Bien que les mécanismes de traduction des gènes soient similaires chez les procaryotes et chez les eucaryotes, des différences importantes les séparent.

En premier lieu, les ARN messagers procaryotes contiennent souvent plusieurs gènes (ou cistrons), alors que les ARN messagers eucaryotes sont presque toujours monocistroniques ; de plus, chez les eucaryotes, les ARN messagers portent à l'extrémité 5' une coiffe, formée d'un ribonucléotide triphosphate méthylé et, à l'extrémité 3', une queue, polyadénylée. L'autre différence encore se situe au niveau du ribosome qui est de taille différente et composé d'ARN ribosomiques différents bien que la structure générale et l'activité soit comparable aux procaryotes.

Comme chez les procaryotes on distingue trois étapes dans la synthèse protéique chez les eucaryotes (initiation, élongation et terminaison).

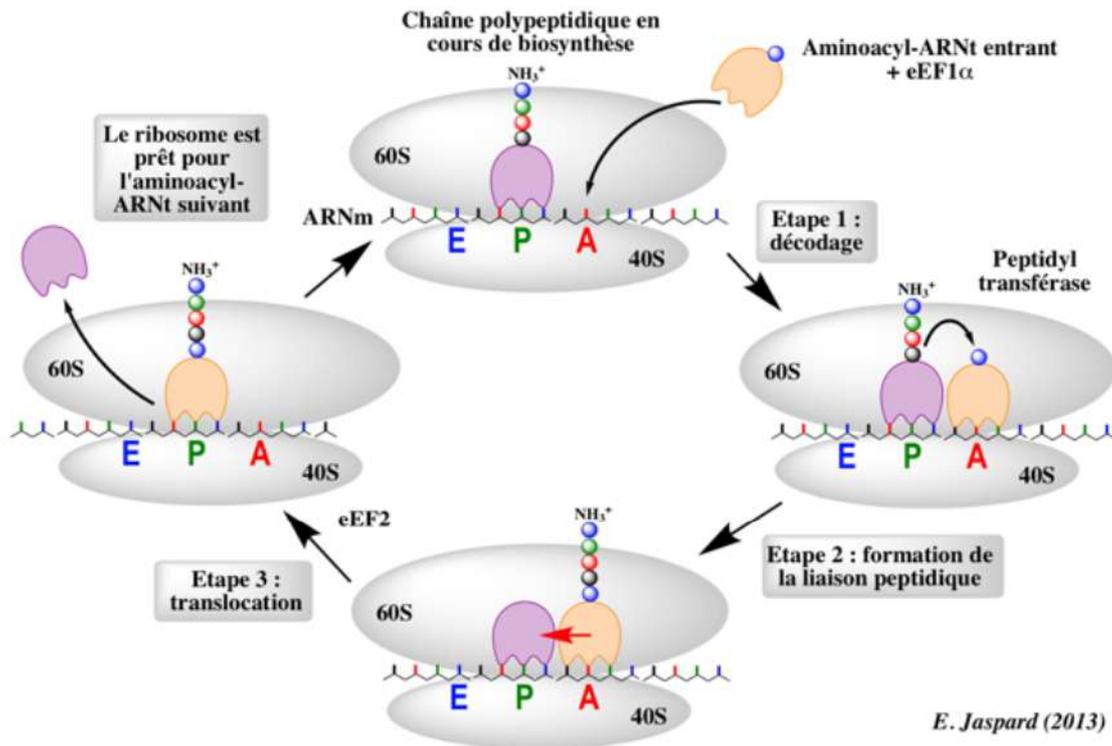


Figure 70 : Etapes de la traduction chez les eucaryotes.

### 1) Initiation

L'initiation chez les eucaryotes est similaire à celle des procaryotes. Cependant, elle débute par la reconnaissance et l'attachement d'un complexe comprenant la petite sous-unité ribosomale (40S) sur la coiffe, à l'extrémité 5' de l'ARNm et d'au minimum une dizaine de facteurs protéiques, formés pour la plupart de plusieurs sous-unités, connus sous le nom de facteurs d'initiation ou eIF (pour *eukaryotic Initiation Factors*).

Ce complexe progresse ensuite le long de la **région 5' non traduite (5' UTR 3')** jusqu'à atteindre le codon d'initiation (étape de « balayage ») où le ribosome complet (80S) sera assemblé. Au cours de ce processus, chacun des facteurs d'initiation a un rôle précis qui permet au final de positionner un ribosome compétent pour la traduction sur le « bon » codon d'initiation. Ce mécanisme est connu sous le nom d'initiation dépendante de la coiffe.

Avant le début de la synthèse, quatre facteurs d'initiation, eIF1, eIF1A, eIF3 et eIF5 se lient à la petite sous-unité du ribosome. Ces facteurs empêchent l'association de la petite sous-unité avec la grande et la liaison des ARNt au site A (complexe 1 ou complexe 43S).

Le Met-ARNt<sub>i</sub><sup>Met</sup> est recruté par le facteur eIF2 lié au GTP, ce qui forme un complexe ternaire. Ce complexe se fixe ensuite au site P de la petite sous-unité pour constituer le complexe de pré-initiation. La liaison de l'ARNt initiateur à la petite sous-unité du ribosome précède toujours l'association avec l'ARNm. L'ARNt charge une méthionine non modifiée (non formylé). On l'appelle Met-ARNt<sub>i</sub><sup>Met</sup>.

Une fois le complexe assemblé, il est recruté à l'extrémité 5' de l'ARNm par l'intermédiaire d'un complexe de quatre protéines ayant chacune une fonction :

- eIF4E reconnaît la coiffe située à l'extrémité 5' de l'ARNm et s'y fixe.
- eIF4A possède une activité hélicase et ATPase permettant de dérouler les éventuelles structures secondaires proches de la coiffe de l'ARNm.
- eIF4B stimule l'activité hélicase de eIF4A.
- eIF4G est un adaptateur entre les différents facteurs d'initiation.

Ce dernier se lie à la coiffe, à eIF4A, à eIF4E et à l'ARNm via la protéine PABP qui est une protéine associée à la queue poly(A) (complexe 2). C'est ce qui permet la circularisation de l'ARNm pendant la traduction. Chez les eucaryotes l'interaction entre le facteur eIF4G et la PABP qui recouvre la queue poly(A) maintient l'ARNm dans une configuration circulaire. Cette configuration est très stable puisqu'elle se maintient pendant de nombreux cycles de traduction. Ainsi, lorsque le ribosome a terminé la synthèse de la protéine, il se retrouve très proche de la coiffe ce qui favorise une ré-initiation rapide de la traduction du même ARNm.

Par la suite, il y'a donc recrutement du premier complexe 43S au niveau du deuxième complexe lié à l'extrémité 5' de l'ARNm ce qui forme le complexe 48S. Le ribosome va pouvoir balayer l'ARNm dans le sens 5' → 3' à la recherche du premier codon d'initiation, encadré par la séquence de Kozak.

La séquence de Kozak est une séquence consensus (5'-RNNAUGG-3') où R représente une purine conservée en amont du AUG et G immédiatement en aval.

Le balayage de l'ARNm consomme de l'énergie sous forme d'ATP en raison de l'activité hélicase de eIF4A.

L'appariement correct entre l'anticodon de l'ARNt initiateur chargé (positionné au site P) et le codon initiateur entraîne l'hydrolyse du GTP lié à eIF2. Puis la libération des facteurs eIF2-GDP, eIF1, eIF3 et eIF5 de la petite sous-unité. La grande sous-unité est alors recrutée pour former le ribosome 80S.

Cette étape nécessite l'intervention d'un autre facteur, eIF5B-GTP, qui se lie à l'ARNt initiateur chargé. Cette association nécessite l'hydrolyse du GTP associée à eIF5B-GTP, et la libération des facteurs d'initiation restants.

Le complexe d'initiation 80S est ainsi formé et le ribosome est prêt à accueillir un ARNt chargé au site A et former la première liaison peptidique.



- Le code génétique est **universel**. En effet chaque acide aminé dispose d'un ou de plusieurs codons et ceci au niveau d'une multitude d'organismes vivants procaryote et eucaryote.
- Le code génétique est **redondant** (ou dégénéré). Plusieurs codons codent pour un même acide-aminé : on trouve 64 codons et 20 acides aminés. Souvent se sont les deux premiers nucléotides du codon qui définissent l'acide aminé, la redondance est donc due au troisième nucléotide du codon.
- Le code génétique est **non-chevauchant**. Les nucléotides d'un codon ne participe qu'au code d'un seul acide aminé, ainsi le prochain acide-aminé sera codé par le prochain codon présent sur l'ARNm. On parle du **cadre de lecture** (ou *reading frame*).
- Le code possède un **système de ponctuation**. Le codon d'initiation est le codon AUG (GUG pour la mitochondrie) et les codons de terminaison sont les codons UAA (ocre), UAG (ambre) et UGA (opale). Le codon UGA (opale) n'est pas présent au niveau de la mitochondrie.

**Tableau VI.** Correspondance entre codons et acides aminés (base du code génétique).

Table des codons ARN									
1 <sup>re</sup> base	2 <sup>e</sup> base								3 <sup>e</sup> base
	U		C		A		G		
U	UUU	F Phe	UCU	S Ser	UAU	Y Tyr	UGU	C Cys	U
	UUC	F Phe	UCC	S Ser	UAC	Y Tyr	UGC	C Cys	C
	UUA	L Leu	UCA	S Ser	UAA	Stop <i>ocre</i>	UGA	Stop <i>opale</i> / U Sec / W Trp	A
	UUG	L Leu / initiation	UCG	S Ser	UAG	Stop <i>ambre</i> / O Pyl	UGG	W Trp	G
C	CUU	L Leu	CCU	P Pro	CAU	H His	CGU	R Arg	U
	CUC	L Leu	CCC	P Pro	CAC	H His	CGC	R Arg	C
	CUA	L Leu	CCA	P Pro	CAA	Q Gln	CGA	R Arg	A
	CUG	L Leu / initiation	CCG	P Pro	CAG	Q Gln	CGG	R Arg	G
A	AUU	I Ile	ACU	T Thr	AAU	N Asn	AGU	S Ser	U
	AUC	I Ile	ACC	T Thr	AAC	N Asn	AGC	S Ser	C
	AUA	I Ile	ACA	T Thr	AAA	K Lys	AGA	R Arg	A
	AUG	M Met & initiation	ACG	T Thr	AAG	K Lys	AGG	R Arg	G
G	GUU	V Val	GCU	A Ala	GAU	D Asp	GGU	G Gly	U
	GUC	V Val	GCC	A Ala	GAC	D Asp	GGC	G Gly	C
	GUA	V Val	GCA	A Ala	GAA	E Glu	GGA	G Gly	A
	GUG	V Val	GCG	A Ala	GAG	E Glu	GGG	G Gly	G

■	Acide aminé apolaire
■	Acide aminé polaire
■	Acide aminé acide
■	Acide aminé basique
■	Terminaison

**Phe** : Phénylalanine  
**Leu** : Leucine  
**Ile** : Isoleucine  
**Met** : Méthionine  
**Val** : Valine  
**Ser** : Sérine  
**Sec** : Sélénocystéine

**Pro** : Proline  
**Thr** : Thréonine  
**Ala** : Alanine  
**Tyr** : Tyrosine  
**Pyl** : Pyrrolysine  
**His** : Histidine  
**Trp** : Tryptophane  
**Gly** : Glycine

**Gln** : Glutamine  
**Asn** : Asparagine  
**Lys** : Lysine  
**Asp** : Acide aspartique  
**Glu** : Acide glutamique  
**Cys** : Cystéine  
**Arg** : Arginine

*Ocre, ambre et opale* est une terminologie liée à l'historique de leur découverte.

### Exemple

Le brin d'ARNm est :

A U G G C G U U C A G A A C U G A U A C G U A A

Les différents codons sont donc :

**AUG** · **GCG** · **UUC** · **AGA** · **ACU** · **GAU** · **ACG** · **UAA**

Les ARN de transfert se fixent

UAC CGC AAG UCU UGA CUA UGC codon-stop

par complémentarité et apportent

| | | | | | | reconnu par les

les acides aminés appropriés :

**Met** **Ala** **Phe** **Arg** **Thr** **Asp** **Thr** facteurs de terminaison

### Remarques

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. Parmi les 64 codons, 3 sont des codons de terminaison ou codon stop les 61 restants sont des codons codant.

## Mécanismes moléculaires (Régulation de l'expression des gènes)

### Objectifs spécifiques

Au terme de ce cours qui est la régulation de l'expression génique, vous devez être capable de :

- Connaître et pouvoir déterminer l'intérêt de la régulation de l'expression des gènes ;
- Distinguer les différents niveaux de l'expression des gènes ;
- Comprendre les mécanismes de la régulation génique ;
- Pouvoir décrire les mécanismes de régulation de l'expression de gènes chez les procaryotes et chez les eucaryotes.

### VI.1. Définition

Chez un organisme vivant aussi bien procaryote qu'eucaryote, les gènes ne sont pas tous exprimés en même temps. Il existe, un phénomène qu'on appelle « la régulation de l'expression des gènes ». La régulation de l'expression des gènes est le moyen de contrôler l'activité d'un gène donné. Elle définit si et dans quelle mesure, un gène est exprimé, à savoir, quand un gène doit être lu et copié en ARN. Elle a pour effet de moduler et d'accroître ou de décroître la quantité des produits de l'expression des gènes (ARN, protéines). Ce système permet donc l'activation ou la répression de la synthèse d'un produit issu d'un gène donné. La régulation peut être positive grâce à des **activateurs** ou négative grâce à des **répresseurs**.

Ces séquences régulent les gènes qui leur sont adjacents sur le même chromosome, on dit que ces séquences sont actives en « cis » et on parle d'**élément cis-régulateurs**. Sur ces séquences vont se fixer des facteurs de régulations, dit **trans-régulateurs**, ne provenant pas de séquences adjacentes mais de l'expression d'un autre gène situé ailleurs sur le génome comme l'exemple de l'*opéron lactose*.

Toutes les étapes allant de la séquence d'ADN au produit final peuvent être régulées, que ce soit la transcription, la maturation des ARNm, la traduction des ARNm ou la stabilité des ARNm et protéines. La régulation de l'expression génique est indispensable pour que le métabolisme de la cellule soit en adéquation avec son environnement. Elle est également fondamentale pour la différenciation cellulaire, structurelle et fonctionnelle. Elle est responsable du fait que, bien que possédant toutes le même génome, les cellules d'un individu ne l'expriment pas de la même façon.

### VI.2. Mécanismes de régulation

Chez les procaryotes comme chez les eucaryotes, les gènes sont régulés principalement à l'étape de l'initiation de la transcription, le plus souvent par des protéines capables d'activer ou d'inhiber le fonctionnement du gène. Les activateurs fonctionnent toujours par recrutement de l'ARN polymérase mais aussi, chez les eucaryotes, de complexes protéiques responsables de l'initiation de la transcription.

#### VI.2.1. Chez les procaryotes

Chez les procaryotes, la régulation de l'expression génétique sert à répondre aux conditions changeantes de l'environnement immédiat (sorte de mécanisme d'adaptation). Il

est à noter que les gènes participant à la réalisation d'une même fonction sont organisés en unités fonctionnelle qu'on appelle « opéron ».

### VI.2.1.1. Les opérons

Un opéron est une unité génétique d'expression et de régulation d'un ensemble de gènes adjacents qui seront transcrits à l'aide d'un même promoteur sous forme d'un seul ARNm (appelé polycistronique) et traduit en plusieurs protéines différentes. Les opérons sont trouvés uniquement chez les procaryotes et comprennent les gènes de structure, un ou plusieurs gènes régulateurs codants des protéines régulatrices (répresseurs ou activateurs), ainsi que des éléments de contrôle présents dans la séquence d'ADN (promoteur, operateur). On distingue généralement deux types d'opérons :

#### - *Les opérons inductibles*

Ce sont ceux qui codent pour des enzymes impliqués dans la voie catabolique comme l'opéron lactose l'exemple le plus connu.

#### - *Les opérons répressibles*

Ce sont ceux qui codent pour des enzymes impliqués dans la voie de la biosynthèse exemple de l'opéron tryptophane.

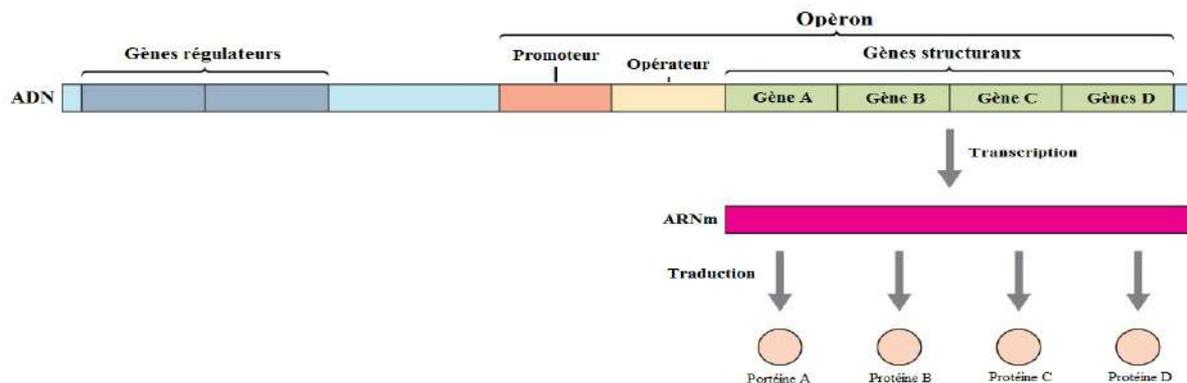


Figure 72 : Représentation schématique d'un opéron.

### Rappel

- **Promoteur**

C'est une région de l'ADN en amont du site d'initiation de la transcription sur laquelle l'ARN polymérase peut se lier.

- **Opérateur**

C'est une région de l'ADN sur laquelle se fixe un répresseur pour contrôler l'expression d'un gène ou d'un groupe de gènes.

- **Terminateur**

C'est une région de l'ADN qui signale la fin de la transcription.

### VI.2.1.2. Opéron lactose

L'opéron lactose ou plus communément appelé « **opéron lac** » est un opéron nécessaire au transport et au métabolisme du lactose chez *Escherichia coli*, ainsi que d'autres bactéries

de la flore intestinale. Cet opéron permet l'exploitation d'un autre sucre le lactose lorsqu'il y'a une pénurie en glucose. L'utilisation du lactose par la cellule fait appel à deux enzymes :

- La **perméase**, qui a pour rôle de faire rentrer le lactose à l'intérieur de la cellule ;
- La  **$\beta$ -galactosidase**, qui permet de le dégrader en glucose.

#### VI.2.1.2.1. Organisation de l'opéron lactose

Au niveau de l'opéron lactose on peut mettre en évidence quatre gènes dont trois gènes structuraux (*lacZ*, *lacY* et *lacA*) codant pour des protéines différentes et un gène régulateur (*lacI*) :

- Gène *lacZ* qui code pour la  **$\beta$ -galactosidase**, qui hydrolyse le lactose en glucose et galactose ;
- Gène *lacY* code pour la **lactose perméase** ou la  **$\beta$ -galactoside-perméase**. C'est une protéine membranaire qui permet l'entrée du lactose dans la cellule bactérienne ;
- Gène *lacA* qui code pour la **thiogalactoside-transacétylase** ou  **$\beta$ -galactoside-transacétylase**, dont le rôle est inconnu dans l'acétylation ;
- Gène *LacI* en 5' (gène régulateur I) possède son propre promoteur (Pi) et code pour une protéine appelée répresseur I.

Une région régulatrice de ces trois gènes structuraux comprend le **promoteur (P)** et l'**opérateur (O)**. En amont de ce promoteur on retrouve un site CAP (C'est une séquence d'ADN de 16 paires de bases en amont du promoteur) fixant une protéine CAP, (*Catabolite Activator Protein* ou protéine activatrice des catabolites), activée par l'adénosine monophosphate cyclique (AMPc) qui est une molécule signal se liant à la CAP. La fixation induit l'activation du promoteur et donc la transcription des gènes de structure. En aval : un opérateur, qui est une séquence d'ADN sur laquelle se fixe des protéines régulatrices permettant la régulation transcriptionnel de l'opéron.

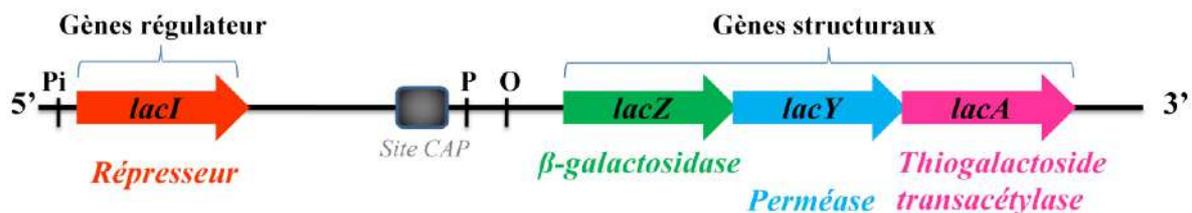


Figure 73 : Représentation structurale de l'opéron lactose.

#### VI.2.1.2.2. Fonctionnement de l'opéron lactose

L'opéron *lac* permet la digestion efficace du lactose en produisant la  $\beta$ -galactosidase afin de digérer le lactose en glucose et galactose. Toutefois, la production de l'enzyme est inutile quand le lactose n'est pas disponible, ou s'il y a une source d'énergie plus facilement exploitable comme le glucose. L'opéron *lac* utilise un mécanisme de contrôle en deux parties qui font que la cellule ne dépense pas d'énergie pour produire la  $\beta$ -galactosidase, la perméase  $\beta$ -galactoside et la thiogalactoside-transacétylase que lorsque c'est nécessaire.

### 1) En absence de lactose

En absence de lactose dans le milieu immédiat de la cellule, le gène *lacI* est normalement transcrit puis traduit entraînant la synthèse d'un répresseur monomérique de la protéine qui s'assemble pour former un complexe tétraédrique (04 sous unités identiques). Ce complexe va se fixer sur l'opérateur avec une grande affinité empêchant l'avancé de l'ARN polymérase (dont le promoteur est situé avant l'opérateur) et donc la transcription du gène.

### 2) En présence de lactose

Lorsque la cellule bactérienne est en présence de lactose, la perméase qui est présente en petite quantité dans la cellule va favoriser la pénétration du lactose à l'intérieur de la cellule bactérienne. Ce dernier va être transformé en allolactose possédant une affinité pour le répresseur. Le gène *lacI* est exprimé mais cette fois-ci chaque monomère du tétramère fixe l'allolactose. Cette fixation entraîne une modification de la conformation (structure) du répresseur qui ne peut plus se fixer sur l'opérateur (dissociation du complexe répresseur-opérateur). Ceci permet la progression de l'ARN polymérase et donc la transcription des gènes de structure de l'opéron (*lacZ*, *lacY* et *lacA*) et par la suite la production des enzymes chargées du métabolisme du lactose. Le **lactose** agit comme un inducteur physiologique alors que l'inducteur synthétique est l'isopropylthiogalactoside (**IPTG**) qui n'est lui pas métabolisable par la  $\beta$ -galactosidase. L'absence du glucose induit l'accumulation de l'AMP cyclique qui va entraîner la fixation du complexe CAP-AMPc au niveau du promoteur.

### 3) En présence de glucose et de lactose

Dans le cas où *E. coli* est en présence de glucose et de lactose en quantité défini, le métabolisme peut être départagé en 2 phases correspondant à la croissance d'*E. coli* dans le milieu de culture. *E. coli* va utiliser d'abord le glucose puis le lactose.

La première phase correspondant au catabolisme du glucose, avec inhibition de la transcription des gènes de l'opéron *lac* ; le glucose pouvant être métabolisé directement par la bactérie, il sera donc utilisé préférentiellement. L'augmentation de taux de glucose va entraîner la diminution de la concentration de l'AMPc (cyclique) donc pas de fixation de l'AMPc à la protéine CAP et par la suite pas de fixation de l'ARN polymérase donc pas d'expression de l'opéron lactose.

La deuxième phase correspondant au catabolisme du lactose, avec activation de la transcription des gènes de l'opéron *lac*. Sachant que la latence entre la phase I et II s'explique par le temps pris par la bactérie pour synthétiser les gènes de l'opéron. Après l'épuisement du glucose, la bactérie se trouve privée d'une source d'énergie, elle accumule l'AMPc qui va se lier à la protéine CAP formant le complexe CAP-AMPc qui lui se fixera au promoteur permettant ainsi la transcription par l'ARN polymérase et la production des enzymes nécessaires.

#### Remarque

Le complexe CAP-AMPc est un régulateur positif ;

Le répresseur est un régulateur négatif.

L'allolactose (galactose-( $\beta$ 1 $\rightarrow$ 6)-glucose) est un isomère du lactose (galactose-( $\beta$ 1 $\rightarrow$ 4)-glucose) et l'inducteur de l'opéron *lac*. Le lactose est converti en allolactose par la  $\beta$ -galactosidase dans une réaction alternative à une hydrolyse.

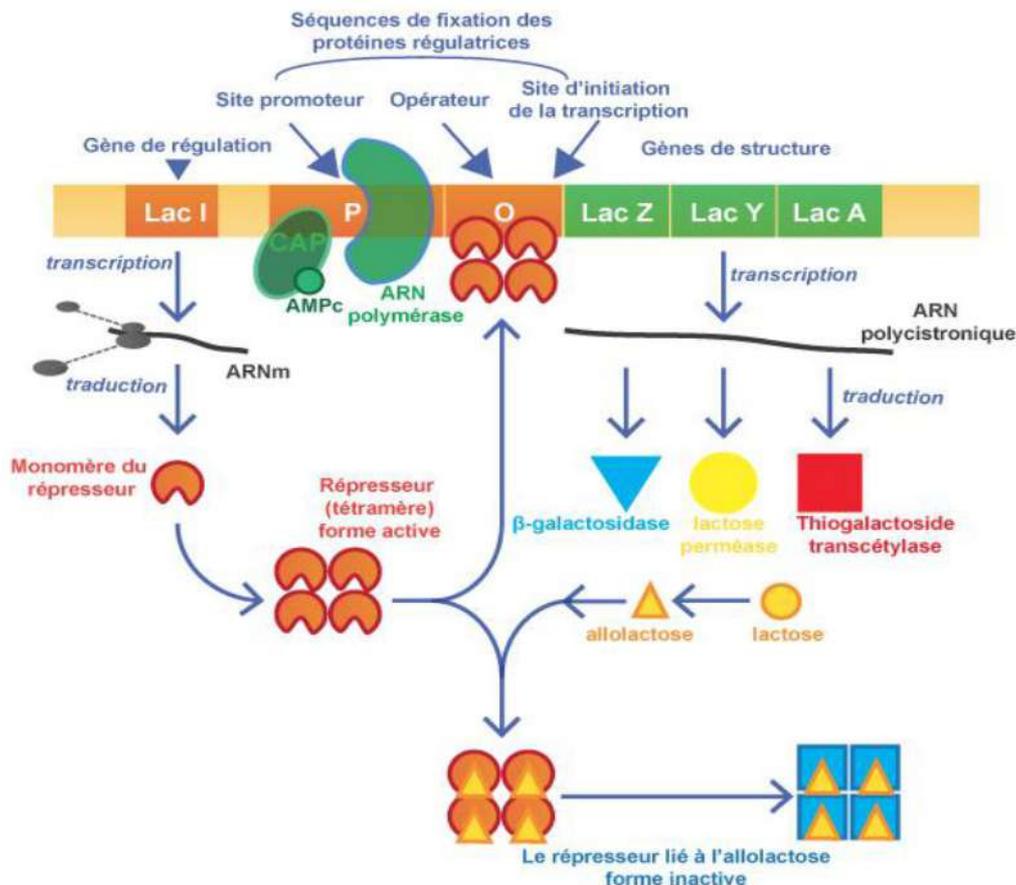


Figure 74 : Mécanisme de l'opéron lactose.

## VI.2.2. Chez les eucaryotes

Du fait de la complexité des organismes eucaryotes (multicellulaires), l'expression de gènes variés est à l'origine d'une spécialisation cellulaire. Chez l'Homme par exemple, on compte 250 types cellulaires différents de part leur morphologie, leur biochimie et leur rôle dans l'organisme. La régulation de l'expression génique peut se faire à différents niveaux qui sont :

### VI.2.2.1. Au niveau chromatinien

Comme vous le savez bien, chez les eucaryotes la chromatine peut se présenter sous deux formes à savoir **Hétérochromatine** qui est transcriptionnellement inactive (ADN condensé autour d'un octamère d'histones formant un nucléosome) et l'**Euchromatine** qui, elle est transcriptionnellement active (ADN décondensé). Les séquences télomériques et centromériques sont sous forme d'hétérochromatine, ces régions contiennent peu ou pas de gènes et afin qu'un gène soit transcrit, il doit être activé. Un gène activé est situé dans des régions non compactées de la chromatine (euchromatine) et non méthylé. On les appelle : régions sensibles ou hypersensibles. Cette régulation va avoir lieu par :

#### 1) Modification des histones

Certaines enzymes (acétylase, désacétylase, méthylase, déméthylase, phosphorylase, déphosphorylase, etc.) vont modifier l'état des histones en ajoutant ou en retirant des groupes chimiques. Par exemple, l'histone acétylase va ajouter des groupements acétyles, ce qui active

les gènes et donc la transcription. Les histones désacétylases sont capables de désacétyler les histones réprimant ainsi la transcription. En phosphorylant les histones (C'est-à-dire en ajoutant des groupements phosphates), cela inactive également les gènes.

### 2) Modification de la structure de l'ADN

La majorité de l'ADN est sous forme d'ADN B, mais il existe une forme particulière de l'ADN qui est la forme Z (transcriptionnellement inactif) modifiant la compaction de l'ADN.

### 3) Méthylation de l'ADN

La méthylation de l'acide désoxyribonucléique (ADN) est un processus épigénétique dans lequel certaines bases nucléotidiques peuvent être modifiées par l'addition d'un groupement méthyle. Cette modification de l'ADN est effectuée par des enzymes particulières appelées DNMTs pour "DNA methyl-transférase". Il est reconnu que l'ADN est méthylé (en carbone 5). La méthylation de l'ADN a généralement lieu au niveau des séquences cytosine-phosphate-guanine (CpG). Cette méthylation a un rôle de diminution de la transcription. En effet, pour qu'un gène soit transcrit, il doit être activé donc situé au niveau de l'euchromatine et qu'il soit non méthylé.

#### VI.2.2.2. Au niveau transcriptionnel

La régulation de la transcription fait appel à des facteurs régulateurs permettant de moduler positivement ou négativement la transcription en se fixant sur la molécule d'ADN. Cette régulation agit en réponse à des stimuli externes spécifiques. Les régulateurs sont soit synthétisés, soit activés à des instants précis et/ou dans des tissus spécifiques, et permettent donc le contrôle de la transcription dans le temps et dans l'espace en fonction des conditions environnementales. Cette régulation fait intervenir :

#### 1) Séquences cis-régulatrices

Ces séquences dites en cis se trouvent au niveau de l'ADN et comprennent le promoteur, les séquences activatrices ou modératrices, les séquences insulateurs (isolateurs) qui permet d'isoler le gène du reste de la chromatine, les séquences de réponse et la combinaison de plusieurs séquences.

#### 2) Facteurs de régulation dit trans

Certains de ces facteurs interviennent d'une manière générale, certains sont spécifiques d'autres inductibles et parfois ce sont des signaux extracellulaire comme les hormones thyroïdiennes, stéroïdes agissant au niveau des récepteurs nucléaires.

#### VI.2.2.3. Au niveau post transcriptionnel

Après la transcription de l'ARN plusieurs systèmes de régulation peuvent intervenir comme l'épissage alternatif. Il est reconnu qu'un gène peut coder des protéines différentes comme par exemple dans la thyroïde l'excision et épissage d'un gène aboutit à une protéine qui est la calcitonine (Métabolisme du calcium et du phosphore) alors qu'au niveau du système nerveux le même gène sera épissé d'une manière différente aboutissant à la formation d'une protéine différente CGRP (calcitonin gene-related peptide) qui a un effet vasodilatateur et un effet cardiaque inotrope et chronotrope donc le même gène peut donner naissance à deux protéines dont la fonction est différente suivant le tissu dans lequel le gène est exprimé.

#### VI.2.2.4. Au niveau traductionnel

Plusieurs mécanismes sont mis en jeu pour réguler la traduction protéique à savoir :

##### 1) Inhibition de la lecture de l'ARN

Ex : synthèse de la ferritine, qui est une molécule de stockage du fer en excès dans la cellule  
En absence de fer au niveau des cellules, il n'y a pas de nécessité de ferritine, de ce fait, l'ARNm sera bloqué afin d'éviter l'expression de ferritine. En revanche, en présence d'un excès de fer, l'ARNm sera traduit et il y aura production de ferritine.

##### 2) Inhibition des facteurs de traduction

Dans ce cas la on parle de ceux impliqués dans l'initiation, l'élongation et la terminaison.

##### 3) Micro ARN

Les micro ARN (miARN), non codants de 21 à 25 nucléotides sont codés par les gènes *miR* (dont certains sont présents dans les introns d'autres gènes) avec une structure particulière, dont le gène est composé de **2 séquences palindromiques** séparées d'un espaceur. Ils peuvent contrôler l'expression génique et modifier le comportement des ARNm (dégradation, inhibition, etc.) en s'appariant avec des ARN cibles portant une séquence homologue permettant la répression de la traduction de la protéine correspondante ou clivant l'ARNm cible au niveau du site de fixation du micro ARN.

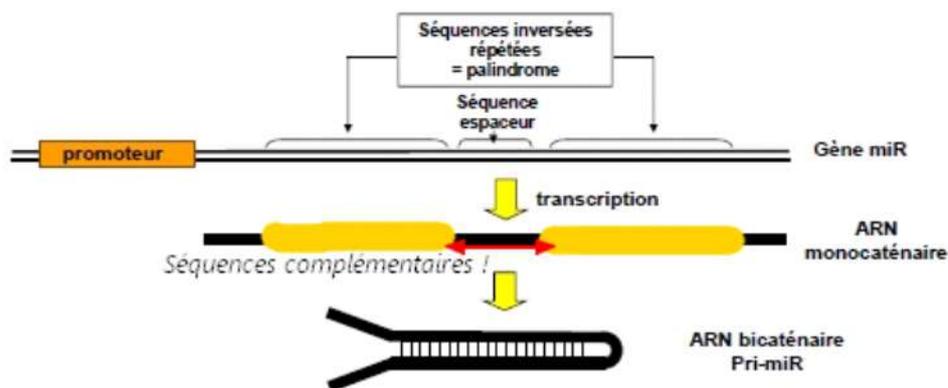


Figure 75 : Représentation d'un micro-ARN.

#### VI.2.2.5. Au niveau post traductionnel

Renferme les mécanismes mis en jeu dans la maturation et l'activation des protéines produites (Forme active ou inactive de la protéine selon les besoins de la cellule).

# *Références*

---

- Ameur Ameur, A. (2015) Génétique générale : Utile en biologie « génétique, biochimie, immunologie, écologie, microbiologie, zoologie, botanique ». Cours 2ème année. Edition Al-Gazzair.
- Ameziane, N., Bogard, M., Lamoril, J. (2006) Principes de biologie moléculaire en biologie clinique. Elsevier SAS.
- Berk, L., Kaiser, M., Scott, K., Darnell, Z. (2005) Biologie moléculaire de la cellule, 3ème édition. DE BOECK.
- Brunori, M., Gilson, É., Puisieux, A. (2001) Le système de réparation des mésappariements contrôle la stabilité des télomères. *médecine/sciences*. 17, 1184-6.
- Claverys, J.P. (1992) Correction des mésappariements et stabilité de l'information génétique. Société Française de Génétique m/s. 8(6), 1-11.
- Cohen, P. (2017) La réparation de l'ADN. Tutorat Santé Lyon Sud. 1-14.
- Cohen, P. (2017) La réplication de l'ADN. Tutorat Santé Lyon Sud. 1-17.
- Cohen, P. (2017) La transcription de l'ADN. Tutorat Santé Lyon Sud. 1-23.
- Dantzer, F., de Murcia, G. (1998) Quelles sont les ADN polymérase requises pour la réplication et la réparation de l'ADN chez les eucaryotes ? *médecine/sciences*. 14, 704-712.
- Dupont, A. (2008) La recombinaison homologue sur molécule unique d'ADN: mesures de torsion et de couple. Biophysique [physics.bio-ph]. Université Paris-Diderot - Paris VII, 2008. Français. <tel-00346315>.
- Etienne, J. (1999) Biochimie génétique & Biologie moléculaire, 5ème édition. Masson.
- Grindley, N.D., Whiteson, K.L., Rice, P.A. (2006) Mechanisms of site-specific recombination. *Annu Rev Biochem*. 75, 567-605.
- Haber, J.E. (1998) Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu Rev Genet*. 32, 561-99.
- Hoede, C. (2017) Impact des processus de mutation et de recombinaison sur la diversité génomique au sein de l'espèce *Escherichia coli*. Génétique. Université Paris 7 - Denis Diderot, 2010. Français. fftel-01621655v2.
- Holliday, R. (1964) A mechanism for gene conversion. *Gen Res*. 5, 282-304.
- Housset, C., Raisonnier, A. (2010) Biologie Moléculaire. Université Pierre et Marie Curie. 1-207.

- Jean-Jean, O., Cassan, M., Rousset, J.P. (1992) L'initiation de la traduction chez les eucaryotes, source de diversification et de modulation de l'expression des gènes. Société Française de Génétique. 1-11.
- Johnson, A., Raff, L., Walter, R. (2004) Biologie moléculaire de la CELLULE, 4<sup>ème</sup> édition, Médecine-Sciences chez Flammarion. 1464 p.
- Kempf, J. (1999) Fonction de l'ADN : Réplication et expression du génome. *Découvertes Ficus Nénuphar*.
- Lépine, F. (2006) Transcription et traduction de l'ADN. *SBI4U / Génétique*. 1-4.
- Lloyd, R.G., Low, K.B. (1996) Homologous Recombination. *Escherichia coli* and *Salmonella*: cellular and molecular biology, vol. 2. F. C. Neidhardt, Curtiss R. III, Ingraham L. et al. Washington DC, ASM Press: 2236-2255.
- Lunardi, J. (2012) Chapitre 7 : Régulation de l'expression du message génétique. Université Joseph Fourier de Grenoble.
- Maftah, A., Petit, J.M., Raymond, J. (2018) Mini manuel de biologie moléculaire. 4<sup>e</sup> édition, *Dunod*, Paris, 2007, 2011, 2015, 2018, ISBN 978-2-10-077368-8.
- Merlin, C., Toussaint, A. (1999) Les éléments transposables bactériens. *Soc Fran Géné.* 15(8-9), 1-13.
- Michel, B. (1999) Illegitimate recombination in bacteria. In Organization of the prokaryotic genome. R. I. Charlebois. ASM Press, Washington DC, 129-150.
- Motamedi, M.R., Szigety, S.K., Rosenberg, S.M. (1999) Double-strand-break repair recombination in *Escherichia coli*: physical evidence for a DNA replication mechanism in vivo. *Genes Dev.* 13(21), 2889-2903.
- Müller, S. (2008) Nucleic Acids from A to Z: *A Concise Encyclopedia*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-31211-5.
- Petit, J.M., Arico, S., Raymond, J. (2011) Mini manuel de génétique. 2<sup>e</sup> édition, *Dunod*, Paris, 2007, 2011. ISBN 978-2-10-055930-5.
- Pourquier P, Robert J. (2011) Présentation générale des mécanismes de réparation de l'ADN. *Bull Cancer.* 98(3), 229-237. doi : 10.1684/bdc.2011.1323.
- Renault, S., Bonnin-Rouleux, F., Periquet, G. (1997) Elements transposables et Evolution des Génomes. *Virologie.* 1, 133-144.
- Rossignol, J.L. (1990) La recombinaison homologue : mécanismes et conséquences. Société Française de Génétique. m/s, 7(6), 1-6.
- Samarut, J. (2017) Régulation de l'expression des gènes. Tutorat Santé Lyon Sud.

- Serre, J.L. (2006) GÉNÉTIQUE : Rappels de cours, exercices et problèmes corrigés. 3e édition. Dunod, Paris, 2001, 2004, 2006. ISBN 2 10 050524 6.
- Shiraishi, K., Imai, Y., Yoshiaki, S, Ikeda, H. (2005) Rep helicase suppresses short-homology-dependent illegitimate recombination in *Escherichia coli*. *Genes Cells*. 10(11), 1015-1023.
- Simon, M. (2018) Cours Pharmacie : Régulation de l'ADN. <https://www.cours-pharmacie.com/biologie-moleculaire/regulation-de-ladn.html>.
- Smith, G.R. (1988) Homologous recombination in procaryotes. *Microbiol Rev*. 52(1), 1-28.
- Strydom, A. (1994) La recombinaison illégitime dans les cellules de mammifère. *médecine/sciences*. 10, 986-994.
- Swynghedauw, B., Silvestre, J.S. (2008) Aide-mémoire Biologie et génétique moléculaires. 3e édition. Dunod, Paris, 2008, 2000. Nathan, 1994 pour la 1<sup>re</sup> édition. ISBN 978-2-10-053798-3.
- Watson, J.D., Crick, F.H.C. (1953) Molecular structure of nucleic acids. *Nature*. 4356, 737.
- Weisberg, R.L.A. (1983) Site specific recombination in phage lambda. *Lambda II*. C. S. H. Laboratory. Cold Spring Harbor, New York, 211-250.