

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA  
RECHERCHE SCIENTIFIQUE

UNIVERSITE DE SAIDA « Dr TAHAR MOULAY »  
FACULTE DES SCIENCES

DEPARTEMENT DE BIOLOGIE



*Polycopie*

## **Cours de Biologie Moléculaire Appliquée**

**Destiné aux étudiants de 1<sup>ère</sup> année Master BPC**

**« Biologie et Physiologie Cellulaire »**

**Réalisé par Boukabene Fouzia Kheira  
(2021)**

---

*Aux étudiants qui suivent mon cours de Biologie Moléculaire Appliquée (BMAp) et qui, année après année, m'inspirent par leur intelligence, leur curiosité et leur enthousiasme.*

---

## Préface

*Traditionnellement, le fonctionnement des cellules et des organismes a été abordé de trois façons différentes par les généticiens, les biochimistes et les physiologistes. Les généticiens identifiaient des mutations pour rechercher les gènes qui au fond de la boîte tiraient les ficelles. Résolument réductionnistes, les biochimistes purifiaient des molécules, analysaient leur structure et leur fonctionnement, pour recréer en tube à essai les réactions de la vie. Observant la réponse finale de systèmes plus ou moins perturbés, les physiologistes recherchaient des boucles de régulation et cherchaient à comprendre quelles structures (cellulaire, tissulaire...) en étaient le siège. C'est du dialogue entre ces trois disciplines que devait nécessairement naître l'image complète du vivant que tous recherchaient, mais ce dialogue était difficile : les objets étudiés (gènes immatériels, molécules chimiques, réseaux de régulation ou structures physiques) étaient différents, les langages avaient profondément divergé, et les méthodologies sophistiquées propres à chaque discipline créaient autant d'écrans entre spécialistes.*

*La biologie moléculaire n'est pas à proprement parler une discipline au sens où on entend généralement ce mot. Elle est née dans les années 40 -du fait en particulier de quelques physiciens curieux de savoir si "la vie obéissait aux mêmes lois que la physique"- comme une approche intégrant les trois disciplines fondamentales (génétique, biochimie, physiologie) pour obtenir sur un système modèle le plus simple possible (un phage, une bactérie) une vision complète du fonctionnement d'un être vivant. Après plus de 60 ans d'effort, nous ne comprenons toujours pas complètement lambda ou Escherichia coli, mais cet effort nous a fourni des concepts universels, et surtout une extraordinaire panoplie d'outils collectivement regroupés sous le nom de génétique moléculaire. Ces outils et les concepts qu'ils véhiculent ont contribué à unifier puissamment la biologie moléculaire aujourd'hui et sont à la base de ce cours.*

*Ils sont aussi à l'origine d'applications très diverses qui modifient profondément nos relations avec l'ensemble du monde vivant et posent des choix difficiles à nos sociétés: puissent ce cours "**Biologie Moléculaire Appliquée**" apporter à ce qui les liront les informations utiles à leur compréhension. Par rapport aux concepts déjà abordés en deuxième et troisième année, l'objectif de ce cours est de reprendre, de développer et d'étendre la description de l'ensemble des grands mécanismes impliquant l'ADN et l'ARN. Les grands principes de bases seront considérés comme acquis (structure de l'ADN, principes de la transcription, de la traduction, code génétique...). Quelques rappels seront effectués, pour la lisibilité du texte. En revanche, on abordera les aspects plus avancés qui seront sensiblement plus développés, en complément et au delà des notions élémentaires vues les années précédentes.*

*Comme dans la plupart des processus biologiques, la fonction est souvent dictée par la structure des macromolécules impliquées, c'est pourquoi, chaque fois que c'est possible, l'étude des relations entre la structure et la fonction des molécules sera abordée, ainsi que les mécanismes moléculaires impliqués.*

*Enfin, on essaiera de détailler les outils et approches expérimentales qui sont utilisés pour mettre en évidence et étudier ces processus. Il est important de maîtriser le principe de ces applications (techniques et biotechnologiques) pour être en mesure de pouvoir comprendre une démarche expérimentale en biologie moléculaire et pour pouvoir analyser et interpréter les publications scientifiques.*

# Sommaire

## Préface

I. Les acides nucléiques.....	1
1. Les nucléotides.....	1
- Les bases azotées.....	2
- Les nucléosides.....	7
- Les nucléotides.....	9
2. Les acides nucléiques.....	12
- La structure primaire des polymères.....	13
- Les différents groupes ionisables des acides nucléiques.....	14
- L'hydrolyse des acides nucléiques.....	17
- Structure spatiale des acides désoxyribonucléiques.....	19
- Structure spatiale des acides ribonucléiques.....	27
II. La réplication.....	30
1. La double hélice.....	30
- Les différentes formes de la double hélice.....	30
- La réplication chez les procaryotes.....	31
- La réplication chez les eucaryotes.....	32
- Les étapes de mécanisme génétique de réplication (in vivo).....	32
2. Particularité du mécanisme de la réplication chez les procaryotes.....	33
3. Particularité du mécanisme de la réplication chez les eucaryotes.....	34
4. Organisation de l'ADN dans la chromatine.....	35
III. Application à la PCR et au séquençage.....	37
1. Amplification in vitro des acides nucléique PCR et RT-PCR.....	37
- Les amorces.....	38
- Les températures.....	39
- L'ADN matriciel.....	40
- L'ADN polymérase.....	40
- Le tampon.....	40
- Aspect quantitatifs de la réaction PCR.....	40
2. RT-PCR.....	41
3. Utilisation des produits PCR.....	42
IV. Le séquençage de l'ADN.....	43
1. Les différents acteurs du séquençage.....	43
2. Méthode de Sanger.....	44
- Principe de séquençage par la méthode Sanger.....	44
- Automatisation du séquençage.....	46
V. La transcription.....	48
1. Structure générale d'un gène.....	48
2. Mécanisme de la transcription.....	48
1. Début de la transcription.....	49
2. Elongation.....	50
3. Fin de la transcription.....	51
4. Les produits de la transcription.....	51
5. Modifications post transcriptionnelles.....	53
VI. La traduction.....	55
1. Les éléments nécessaires à la traduction.....	56
- Les différentes étapes.....	56
2. Code génétique.....	60
- Le code à trois lettres (3bases, 3nucleotides).....	60
- Les caractéristiques du code génétique.....	60
- Les codons.....	61

VII. La mutagénèse.....	62
1. Mutagénèse dirigée.....	62
- Remplacement de bases.....	63
- Modification de la séquence d'ADN.....	63
- Endommagement.....	64
2. Mutagénèse aléatoire.....	64
VIII. Application au clonage moléculaire (génie génétique).....	65
1. Les enzymes de restriction.....	65
- Séquences d'ADN reconnues par les enzymes de restriction.....	66
- Origine des enzymes de restriction.....	67
- Utilisation des enzymes de restriction.....	71
2. Les éléments nécessaires au clonage.....	72
- Les vecteurs de clonage.....	72
1- Les plasmides.....	72
2- Les phages.....	78
3- Les cosmides.....	82
4- Autres types de vecteurs.....	83
- Les banques d'ADN.....	84
- Les vecteurs d'expression.....	84
3. Les applications du clonage moléculaires (génie génétique).....	86
Références Bibliographiques.....	88

## I. Les acides nucléiques :

Le médecin suisse MIESCHER isola en 1869, à partir de leucocytes de pus de plaies infectées, une substance organique remarquablement riche en phosphore qu'il baptisa nucléine. Sa nature chimique fut élucidée par les travaux initiateurs de KOSSEL à partir de 1882 jusqu'à ceux de LEVENE en 1927 : il s'agissait de l'**acide désoxyribonucléique** dont l'abréviation est **ADN** (DNA : anglo-saxon), polymère d'unités dénommées **nucléotides**.

Les polynucléotides biologiques sont :

- le support moléculaire de l'information génétique : l'ADN (et ARN pour certains virus) est le support de l'hérédité et du codage des composés biologiques (les ARN, les protéines)
- des effecteurs de l'expression de l'ADN en peptides et protéines : **acide ribonucléique** dont l'abréviation est **ARN** (RNA : anglo-saxon) regroupés en trois classes :
  - les ARN messagers (ARNm)
  - les ARN de transfert (ARNt)
  - les ARN ribosomiaux (ARNr)

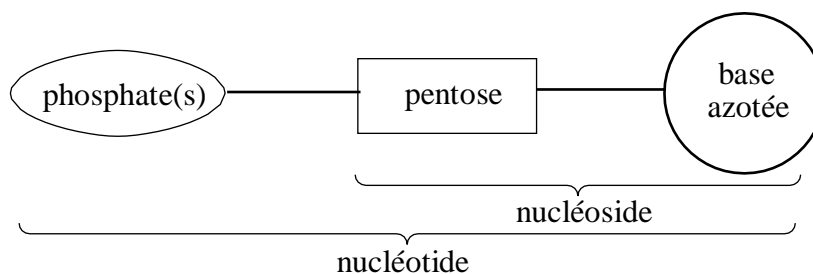
Les nucléotides ont des fonctions variées et importantes, ce sont :

- des composés à "haut potentiel énergétique" dont dépendent l'activation de molécules entrant dans des réactions endergoniques dans les réactions cellulaires,
- des composés structuraux de coenzymes,
- des seconds messagers intracellulaires de signaux et des médiateurs extracellulaires,
- des régulateurs d'activité de protéines lors de modifications covalentes.

### 1. Les nucléotides :

Un nucléotide résulte de :

- 1) la condensation d'un ose (pentose) avec une base nucléique (hétérocycle azoté) qu'on appelle **nucléoside**.
- 2) l'estérification de l'ose d'un nucléoside par l'acide phosphorique produit un **nucléotide**.

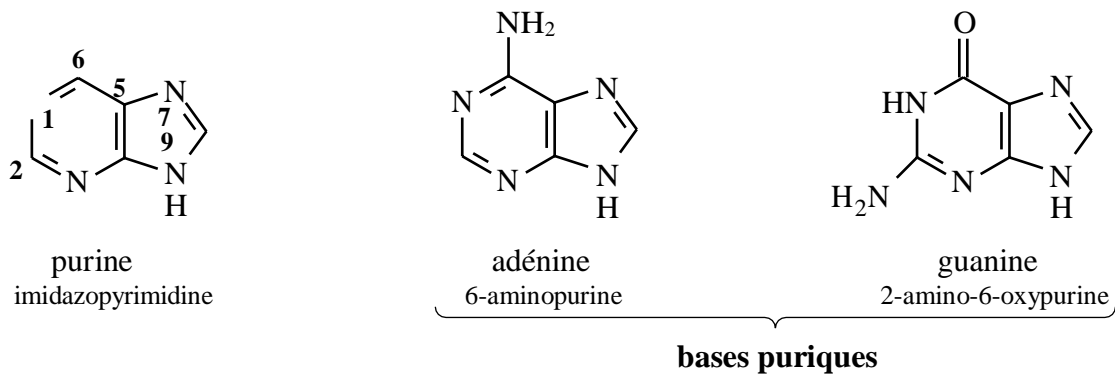
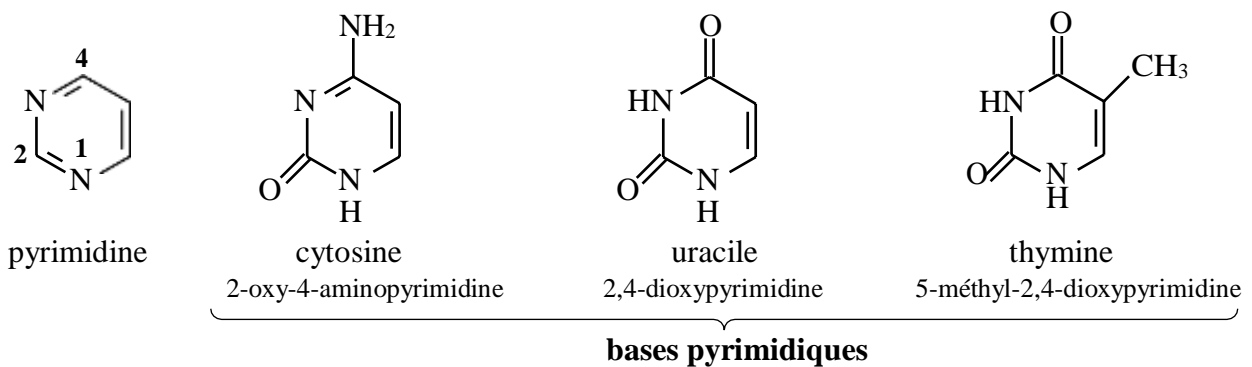


▪ **Les bases azotées :**

Cinq bases majeures, partagées en deux séries, entrent dans la composition des nucléotides et leurs polymères. Elles vont conférer aux composés biologiques dont elles font partie des propriétés capitales.

**Les bases pyrimidiques et les bases puriques**

Les dérivés **oxy** ou/et **amino** de la pyrimidine et de la purine forment les deux familles de base des nucléotides naturels.



Les noms courants des différentes bases n'ont aucun lien avec la nomenclature classique de la chimie organique, certains font référence à leurs conditions de découverte (thymine : thymus de veau).

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

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

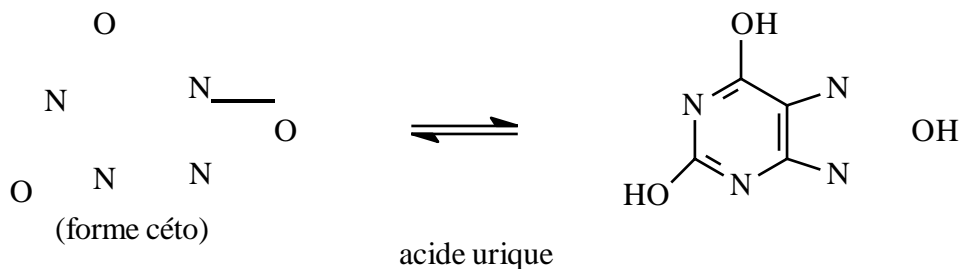
## Les bases modifiées dans l'ADN ou l'ARN

Les modifications peuvent avoir lieu sur des sites cycliques ou exocycliques :

- la **5-méthylcytosine** est trouvée dans l'ADN des plantes et des animaux sauf les insectes. Cette méthylation est un signal négatif de la régulation de l'expression des gènes, le groupe méthyle favorisant une conformation de l'ADN qui ne peut fixer un facteur de transcription.
- la **N6-méthyladénine** est présente dans les bactéries. Cette méthylation permet aux enzymes de restriction de la bactérie de reconnaître son propre ADN vis-à-vis d'ADN étrangers (virus). D'autres méthylations permettent le fonctionnement d'un système de correction des éventuelles erreurs de réplication de fonctionner.
- Les ARN et principalement les ARNt contiennent une variété étendue de dérivés : des dérivés hydrogénés (5, 6-dihydrouracile) ou soufrés (**thiouracile** ou 2-oxy-4-thiopyrimidine) des pyrimidines, ou encore des formes altérées de la guanine, la **xanthine** (2, 6-oxypurine) et l'**hypoxanthine** (6-oxypurine).

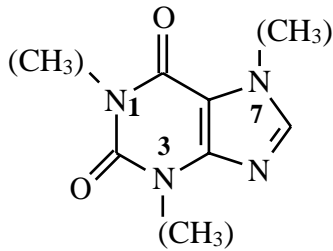
## Des dérivés : molécules d'intérêt biologique

- Lorsqu'elles ne sont pas recyclées, les bases puriques sont dégradées en **acide urique** par passage par des formes désaminées hypoxanthine et xanthine. Celui-ci, très peu soluble, est excrété par les primates.



- des produits de métabolisme des alcaloïdes végétaux sont des produits à usage pharmacologique : caféine (stimulant), théobromine et théophylline (stimulants cardiaques, relaxant des muscles lisses et vasodilatateurs)





1,3-diméthylxanthine ou théophylline

3,7-diméthylxanthine ou théobromine

1,3,7-triméthylxanthine ou caféine

produits de dégradation d'alcaloïdes végétaux

### Des analogues synthétiques

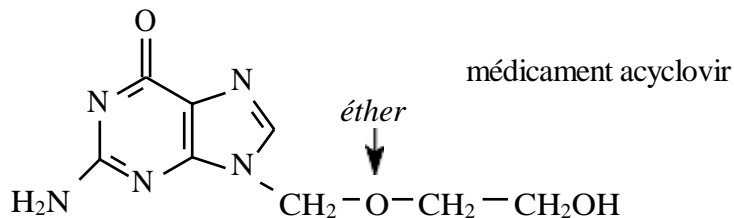
Des analogues des bases nucléiques sont utilisés :

- comme molécules de marquage en biologie moléculaire : 5-bromouracile sur lequel peuvent être greffés des molécules marqueurs

- comme agents thérapeutiques, agissant en compétition avec les bases naturelles, ils bloquent la multiplication des bactéries, la mitose :

- antitumoraux : ce sont les dérivés fluorés (5-fluorouracile), thiols (6-thiopurine) ou encore aza (N remplace un C : 8-azaguanine) qui sont utilisés

- un antiviral classique (acyclovir) est un dérivé de la guanine :  
(2-hydroxyéthoxy)9-méthylguanine



### Des propriétés importantes des bases azotées

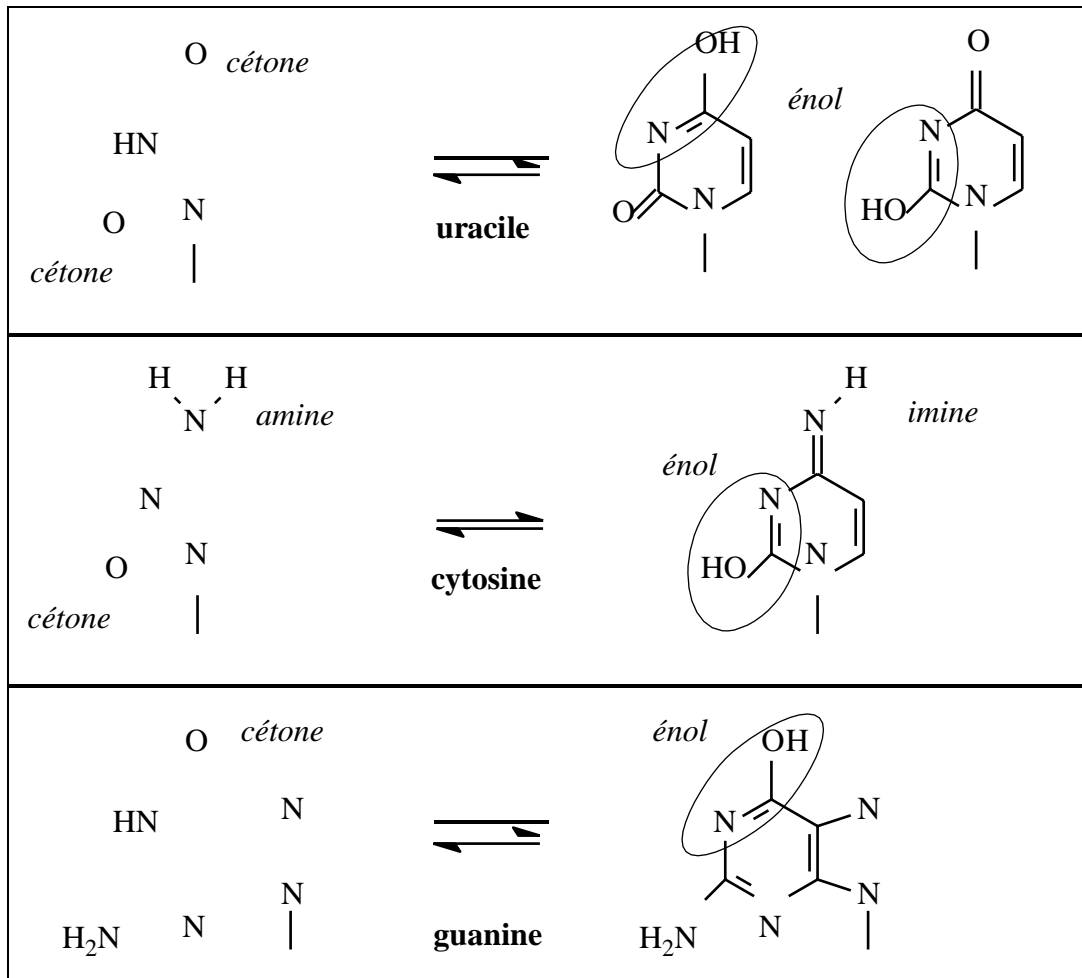
Leurs formules chimiques indiquent :

- les hétérocycles azotés sont susceptibles d'ionisation
- les doubles liaisons créent des systèmes conjugués pour lesquels certaines propriétés physiques sont remarquables (spectre, hydrophobicité, empilement (stacking))

### La conjugaison des doubles liaisons

La résonance entre de nombreux atomes délocalise les électrons  $\Pi$  des doubles liaisons avec les conséquences suivantes :

- la molécule est fortement stabilisée dans une configuration plane
- la molécule existe sous différentes formes tautomères



Equilibres tautomériques à pH 7 des différentes bases liées à un ose (nucléoside ou nucléotide) :

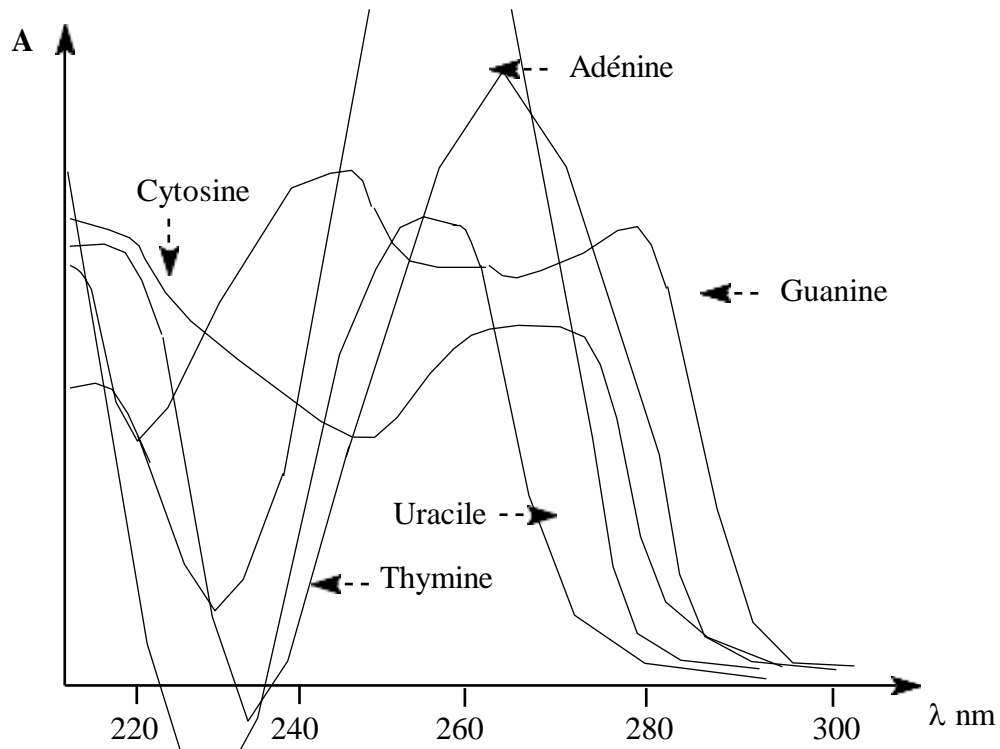
- forme **lactame** à gauche (cétone)
- forme **lactime** à droite (énol)

Les formes prépondérantes à pH7 sont les formes lactame (cétone) et amino.

### Les propriétés spectrales

Les hétérocycles des différentes bases ainsi que leurs dérivés, nucléosides ou nucléotides, présentent des spectres d'absorption caractéristiques dans l'ultraviolet, spectres dépendant du pH. L'aire de ces spectres dans cette région est plus élevée pour les purines (à deux cycles) : leurs absorptions sont donc plus importantes. Ces propriétés optiques sont communément utilisées pour la détection, le dosage et le contrôle de pureté d'acides nucléiques.

La fluorescence de ces bases est par contre inutilisable : l'émission se situe dans la région UV 300-320 nm et elle est très faible (400 fois plus faible que celle du tryptophane pour les purines et 2500 fois pour les pyrimidines).



### Les transformations chimiques des bases

1) la lente désamination est spontanée dans les cellules (100 fois moins importante pour les purines par rapport aux pyrimidines) :



2) les radiations altèrent les bases :

- l'irradiation dans l'ultraviolet ouvre les liaisons de deux bases superposées et les pontes par des liaisons covalentes
- les radiations ionisantes (rayons X ou gamma) ouvrent les cycles et les cassent.

3) de nombreux agents chimiques réagissent avec les bases :

- l'acide nitreux ( $\text{HNO}_2$ ) et l'hydrogénosulfite de sodium  $\text{HSO}_3^-$  ont une action désaminante. Ils font partie des conservateurs dans l'industrie alimentaire.
- les espèces réactives de l'oxygène (peroxydes, radicaux libres) font subir des dommages oxydatifs

▪ **Les nucléosides :**

Une liaison covalente (liaison N-osidique) fixe les bases à un pentose.

**Le pentose des nucléosides**

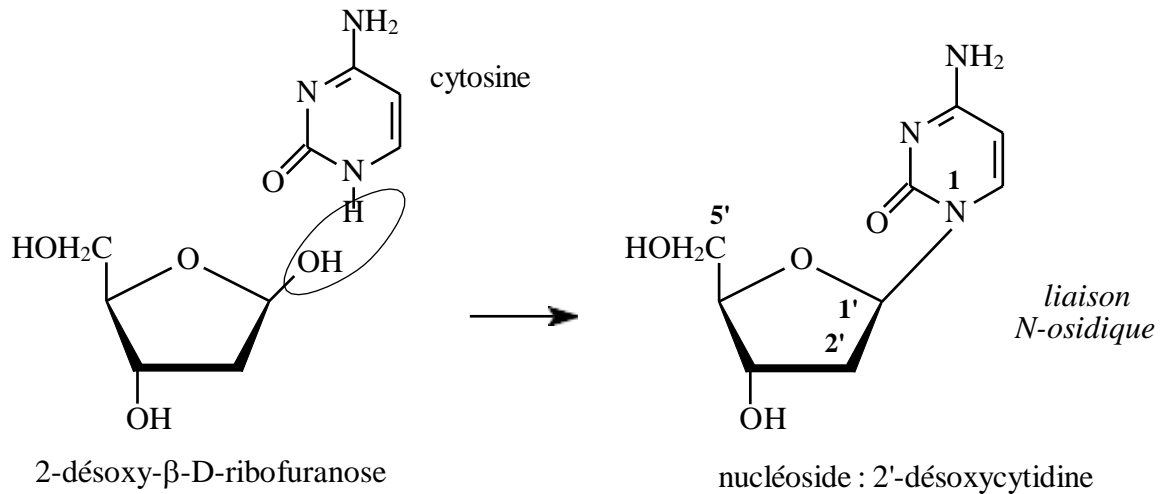
Il s'agit du ribose dans les acides ribonucléiques (ARN) et de son dérivé 2-désoxyribose dans les acides désoxyribonucléiques (ADN) avec les caractéristiques suivantes :

- l'énantiomère est de la série D
- l'ose est sous forme hémiacétalique (furanose)
- l'anomère est en conformation  $\beta$

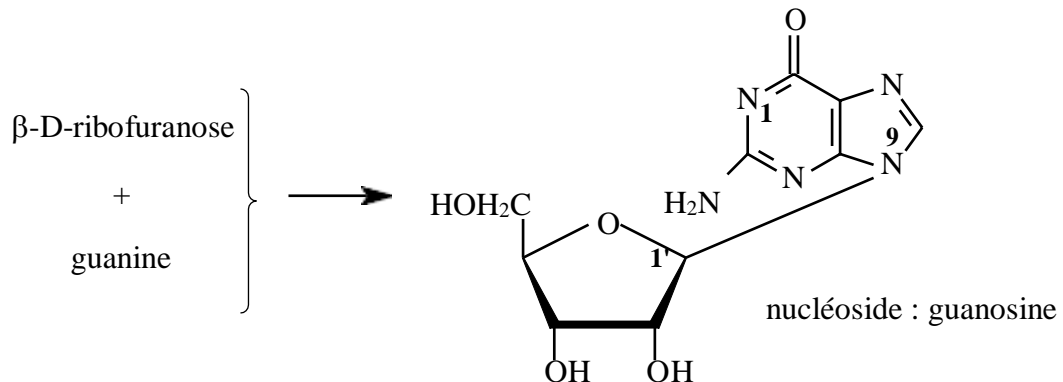
**La liaison osidique**

La liaison avec la base est de type **N-osidique** entre le carbone **1'** (carbone anomérique) du furanose en conformation  $\beta$  et l'azote **N1** des pyrimidines et **N9** des purines.

Exemple pour une base pyrimidique :



Exemple pour une base purique :



## Nomenclature

Les noms des nucléosides ont comme suffixe :

- "osine" pour les nucléosides puriques
- "idine" pour les nucléosides pyrimidiques

Bases azotées		Ribonucléosides		Désoxyribonucléosides	
Nom	Symbole	Nom	Symbole	Nom	Symbole
cytosine	Cyt	cytidine	Cyd	2'-désoxycytidine	dCyd
uracile	Ura	uridine	Urd	2'-désoxyuridine	dUrd
thymine	Thy	thymidine*	Thd	(2'-désoxy)thymidine*	
adénine	Ade	adénosine	Ado	2'-désoxyadénosine	dAdo
guanine	Gua	guanosine	Guo	2'-désoxyguanosine	dGuo
xanthine	Xan	xanthosine	Xao	2'-désoxyxanthosine	dXao
hypoxanthine	Hyp	inosine	Ino	2'-désoxyinosine	dIno

\* le nom de thymidine a été consacré pour la 2'-désoxythymidine : pour différencier ce ribonucléoside, on le désigne sous le nom de ribosylthymidine.

## Les nucléosides naturels

Les nucléosides apparaissent en tant qu'intermédiaires dans le métabolisme des nucléotides, mais aussi comme cofacteurs ou coenzymes :

- la **S-adénosylméthionine** est un cofacteur de transfert de groupement méthyle. Le carbone 5' de l'adénosine est lié de manière covalente avec le groupement sulfhydryl de l'acide aminé méthionine.
- le coenzyme B12 (ou vitamine B12), dont la fonction permet de déplacer des groupes de manière intramoléculaire, est la 5'-désoxyadénosine liée au cobalt du noyau corrine.

## Des nucléosides naturels et singuliers

- On trouve dans les ARN de transfert (ARNt) de l'uracile qui est lié au ribose non par l'azote mais par le carbone **C5** : c'est une liaison C-osidique. On trouve aussi des dérivés 2'-O-méthylés du ribose : la présence du méthyle donne une résistance à l'hydrolyse enzymatique ou même alcaline à l'ARNt.
- Certains microorganismes produisent des nucléosides libres dans lesquels l'ose est l'arabinose (quelquefois appelés spongonucléosides pour avoir été isolés dans des spongiaires).

L'arabinosylcytosine est utilisé comme médicament dans le traitement de l'herpès, l'arabinosyladénosine dans les leucémies.

- La puromycine est un antibiotique sécrété par *Streptomyces*. La base est l'adénosine N-méthylée, le pentose la ribosamine dans lequel le carbone **C3'** est lié par une liaison amide à une O-méthyltyrosine.

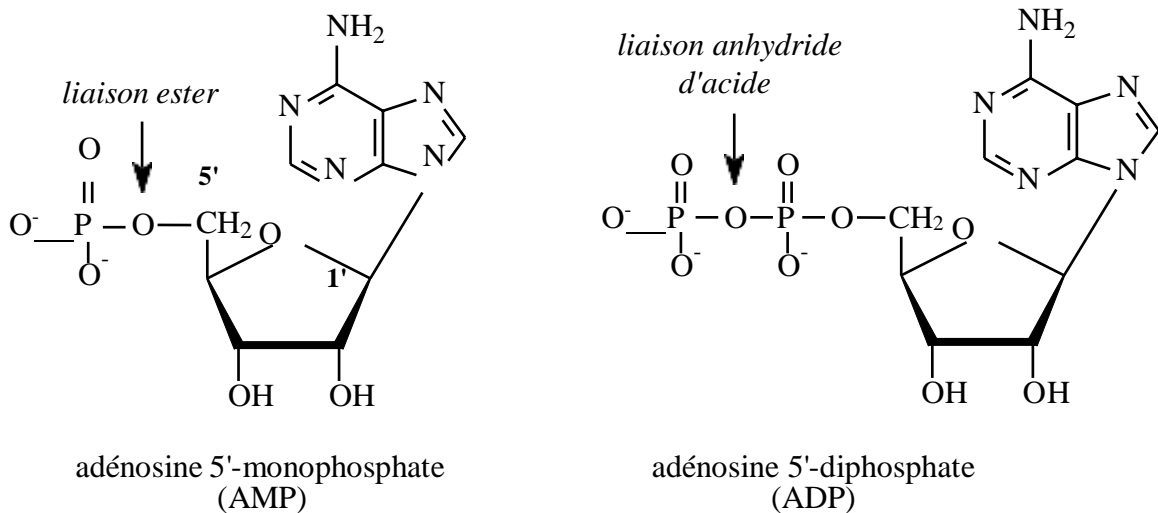
### Des analogues synthétiques

- les formes désoxy, halogénées ou non, des nucléosides puriques ou pyrimidiques sont des bactériostatiques, par exemple la 3'-désoxyadénosine.
- la 6-azauridine (carbone 6 remplacé par un azote) sont des médicaments utilisés dans les maladies prolifératives de l'épiderme.

### ▪ Les nucléotides :

Ce sont des **esters-phosphates** de nucléosides (condensation alcool-acide).

Exemple de nucléotides :



- la phosphorylation peut concerner un ou plusieurs hydroxyles de l'ose.
- le groupe ester-phosphate peut être engagé soit :
  - avec d'autres molécules d'acide phosphorique (ADP, ATP)
  - avec un autre acide (adénosine 3'-phosphate 5'-phosphosulfate)
  - avec un autre hydroxyle par une deuxième liaison ester dans les nucléotides cycliques

## Nomenclature

Les noms des nucléotides obéissent à la règle suivante :

nucléoside **n'**[,n"] – **nb** phosphate : (n' numéro du carbone portant le phosphate)

nucléoside **n'**[:n"] – **nb** phosphate, cyclique pour les nucléotides cycliques

Les symboles ou abréviations classiques sont pour les nucléotides à une seule liaison ester :

- symbole **nucléoside n' P**[P]

- symbole **nucléoside** (une lettre) **N P** (N nombre de phosphate : M 1, D 2, T 3) pour le symbole à 3 lettres qui suppose que la liaison ester est porté par le carbone 5'.

Lorsque l'ose est le désoxyribose, le symbole est précédé de d (exemple dAMP)

Pour leur caractère acide, dû au groupement phosphate, les nucléotides monophosphates ont aussi la dénomination suivante :

acide **n'** [désoxy] (radical de la base) idylique : (n' numéro du carbone portant le phosphate)

Par exemple :

- **adénosine 3', 5'-biphosphate** : l'adénosine porte 2 groupements phosphate, l'un sur la carbone **C3** et l'autre sur le **C5**.

- **adénosine 3'-(mono)phosphate 5'-phosphosulfate** (PAPS) : l'adénosine porte un groupement phosphate en **C3'** et un groupement phosphate en **C5'** qui est lui-même lié à un sulfate par une liaison anhydride d'acide.

- **adénosine 5'-triphosphate** (Ado5' PPP ou ATP) : le plus connu des nucléotides où l'adénosine porte un groupement phosphate en **C5'**, lié à un autre phosphate par une liaison anhydride d'acide, lui-même lié à un autre phosphate par une liaison anhydride d'acide.

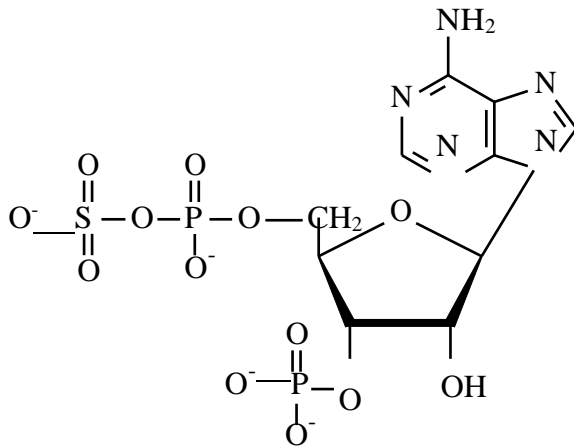
- **adénosine 3':5'-(mono)phosphate, cyclique** (Ado3':5' P ou AMPc) : 2 liaisons esters sur les carbones **C3'** et **C5'** avec un seul groupement phosphate

- **désoxyadénosine 5'-(mono)phosphate** (dAdo5' P ou dAMP) : l'adénosine porte un groupement phosphate en **C5'**, (ou encore acide 5' désoxy adénylique)

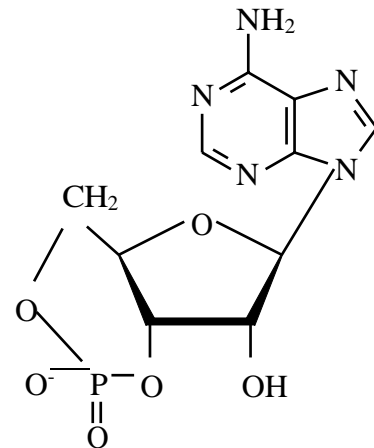
- **adénosine 5'-(mono)phosphate** (Ado5' P ou AMP) : l'adénosine porte un groupement phosphate en **C5'**, (ou encore acide 5' adénylique)

- **cytidine 5'-(mono)phosphate** (Cyd5' P ou CMP) : la cytidine porte un groupement phosphate en **C5'**, (ou encore acide 5' cytidylique)

- **uridine 5'-(mono)phosphate** (Urd5' P ou UMP) : la cytidine porte un groupement phosphate en **C5'**, (ou encore acide 5' uridylique)



adénosine 3'-phosphate 5'-phosphosulfate



adénosine 3': 5'-(mono)phosphate, cyclique

## Les nucléotides d'intérêt biologique

### Les nucléosides monophosphates

La phosphorylation sur le carbone C5' est la plus courante. L'AMP est l'un des composés que l'on trouve dans de nombreuses réactions du métabolisme.

Les nucléosides 5'-phosphates sont les éléments de bases de polymères tels que l'ADN et l'ARN que l'on trouve dans tous les organismes vivants.

### Les nucléosides diphosphates

Le PAPS (adénosine 3'-(mono)phosphate 5'-phosphosulfate) est impliqué dans des réactions de sulfatation des glycanes et des lipides.

L'ADP (et les autres nucléosidesPP classiques) est un composé qui a de nombreuses différentes fonctions, citons parmi elles :

- molécule à haut potentiel énergétique (hydrolyse de la liaison anhydride d'acide)
- molécule intermédiaire dans la production d'ATP
- activateur d'enzyme allostérique comme la L-glutamate-déshydrogénase

### Les nucléosides triphosphates

- L'une des molécules les plus "universelles" est l'ATP : ses deux liaisons anhydride d'acide (enthalpie libre d'hydrolyse d'une liaison de l'ordre de  $30\text{kJ}\cdot\text{mol}^{-1}$ ) sont une source d'énergie utilisable :

- par des réactions de catalyse enzymatique impossibles sans couplage
- pour le transport transmembranaire actif
- pour la contraction musculaire
- etc...

D'autres nucléosides 5' triphosphates jouent un rôle identique à celui de l'ATP mais à un degré beaucoup plus minoritaire.



Les nucléosides 5' triphosphates sont aussi des donneurs de phosphates.

- Les nucléosides 5' triphosphates sont les précurseurs de base dans la biosynthèse des acides nucléiques.

### **Des constituants de coenzymes**

Citons :

- l'adénosine 3'-phosphate 5'-diphosphate est une partie de la molécule d'un coenzyme d'acétylation, le **coenzyme A**.

- l'adénosine 5'-diphosphate fait partie de la structure de deux coenzymes d'oxydo-réduction : le nicotinamide adénine dinucléotide (NAD) et le flavine adénine dinucléotide (FAD)

le PAPS est impliqué dans des réactions de sulfatation des glycanes et des lipides

### **Les seconds messagers**

Un second messenger est une molécule qui assure le relais intracellulaire d'un signal extracellulaire. L'AMPc fut identifié comme le second messenger universel et dans quelques cas, ce fût le GMPc. Il est produit sur la face intracellulaire de la membrane plasmique après fixation externe d'une molécule signal.

## **2. Les acides nucléiques :**

Les acides nucléiques sont des enchaînements de nucléosides 5'-phosphates dont l'assemblage est réalisé par une **liaison phosphodiester**. Les deux types d'acides nucléiques sont :

**ADN** (acide désoxyribonucléique) composé de :

un ose qui est le 2'-désoxyribose

la base est soit : **adénine ou guanine** (purine), soit **cytosine ou thymine** (pyrimidine)

**ARN** (acide ribonucléique) composé de :

un ose qui est le ribose

la base est soit : **adénine ou guanine** (purine), soit **cytosine ou uracile** (pyrimidine)

Remarquons que l'ADN peut contenir des bases méthylées et que l'ARN contient de nombreuses différentes bases modifiées (voir le paragraphe des bases).

La différence entre les 2 oses a des conséquences très importantes entre ces deux polymères : la présence de l'hydroxyle en 2' du ribose interdit tout appariement pour former des duplex de chaînes.

## ▪ La structure primaire des polymères :

### La taille des polymères nucléiques

La taille des acides nucléiques est exprimée en trois unités selon l'usage :

- la longueur
- la masse moléculaire en Dalton (Da)
- le nombre de nucléotides (ou bases), noté *b*, pour les molécules simple brin et le nombre de paires de base, noté *pb*, pour les molécules double brin.

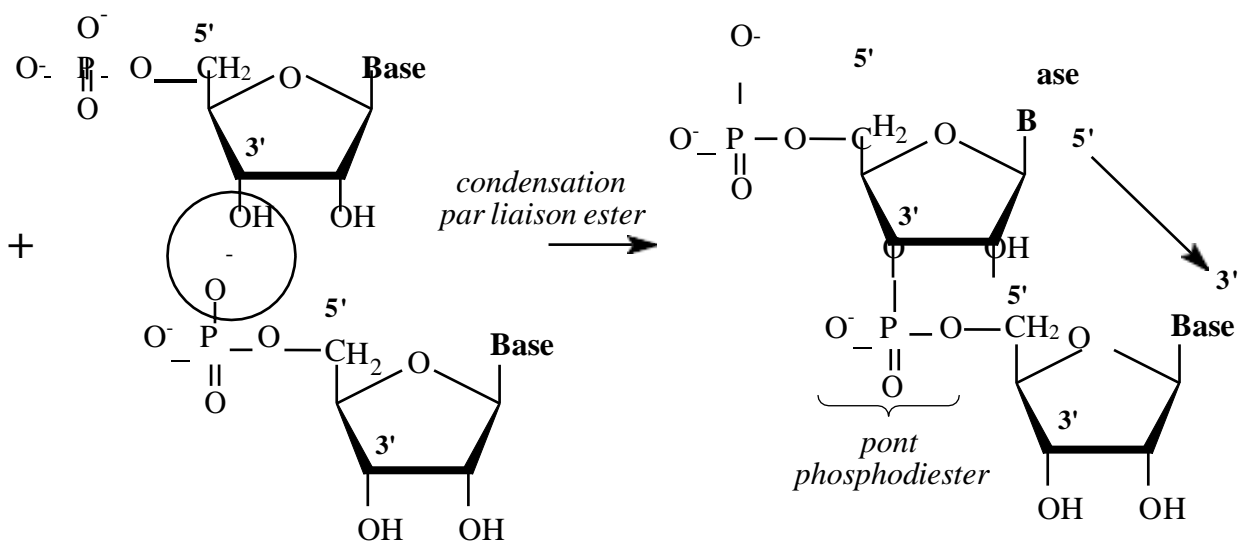
Pour les ARN, le nombre de nucléotides varie de plusieurs dizaines à plusieurs milliers :

- ARN ribosomiaux : de 100 à 5000 *b*
- ARN de transfert : de 75 à 90 *b*
- ARN messagers : fonction du gène transcrit

Pour les ADN, le nombre de nucléotides varie de 5000 à plus de 100 millions de *pb*.

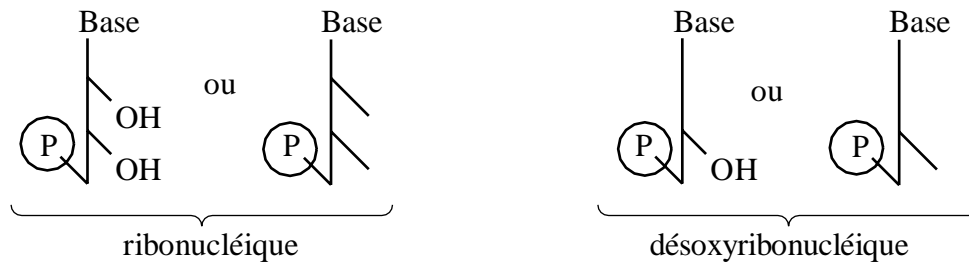
### La liaison phosphodiester

Les acides nucléiques sont des enchaînements de nucléosides 5'-phosphate dont l'assemblage est réalisé par une liaison phosphodiester :

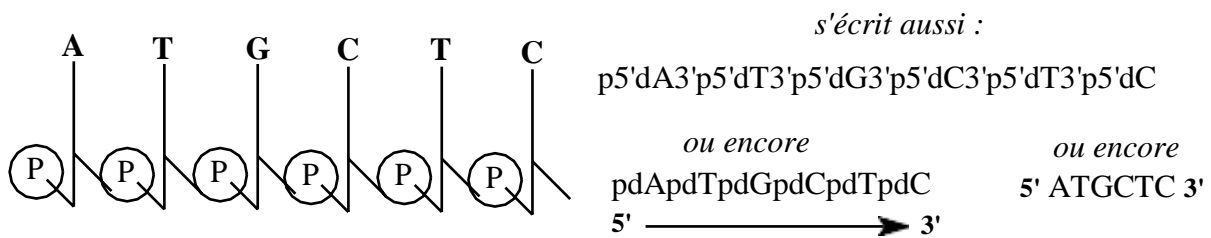


La chaîne est **vectorisée** : elle est écrite de gauche à droite et dans le sens, extrémité phosphate 5' → 3'. C'est le sens dans lequel les séquences d'acides nucléiques sont utilisées comme molécules informationnelles (transcription, traduction).

L'usage a consacré des représentations simplifiées d'un polymère à l'aide d'abréviations ou sigles suivants :



Exemple d'une polymère de type ADN :

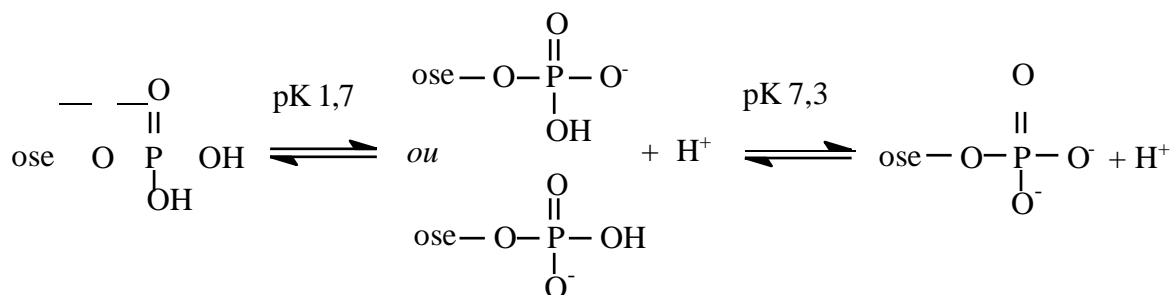


▪ **Les différents groupes ionisables des acides nucléiques :**

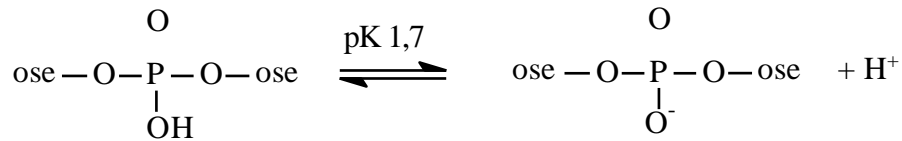
Les différents groupes ionisables des acides nucléiques sont de trois types et d'un type particulier correspondant aux formes tautomères de l'équilibre forme céto et forme énol (lactame et lactime). La nature de l'ose n'a aucun effet pour ces phénomènes d'ionisation.

**Les groupes phosphates**

- le groupe phosphomonoester possède deux fonctions acides dont les valeurs des pK sont respectivement 1,5 à 2 et 7,2 à 7,5.



- le groupe phosphodiester, participant aux ponts phosphodiester, possède une seule fonction acide de pK égal à 1,5 à 2



Chacun de ces groupes possède donc au pH physiologique (6,5) un seul groupement ionisé : l'acide nucléique portera un nombre de charges négatives (contribution des phosphates) égal au nombre de nucléotides.

### Les azotes des cycles des bases pyrimidiques et puriques

Lorsque les bases puriques ou pyrimidiques sont liées à un ose (liaison N-osidique), la protonation d'un azote du cycle correspond au cas plus simple des bases de type pyrrole ou pyridine. Il faut noter que la conjugaison des liaisons du doublet libre de l'azote avec les doublets des électrons  $\pi$  aboutit à la règle simple suivante :

- la protonation d'un azote d'un cycle empêche la protonation des autres azotes.
- l'uracile et la thymine n'ont aucun azote protoné.

### Les groupes amino liés aux cycles des bases

Les groupements amino liés au cycle peuvent être protonés et ce plus facilement qu'un azote du cycle. Le doublet de l'azote de ce groupement amino peut se conjuguer partiellement avec les électrons du cycle, empêchant la protonation des autres azotes.

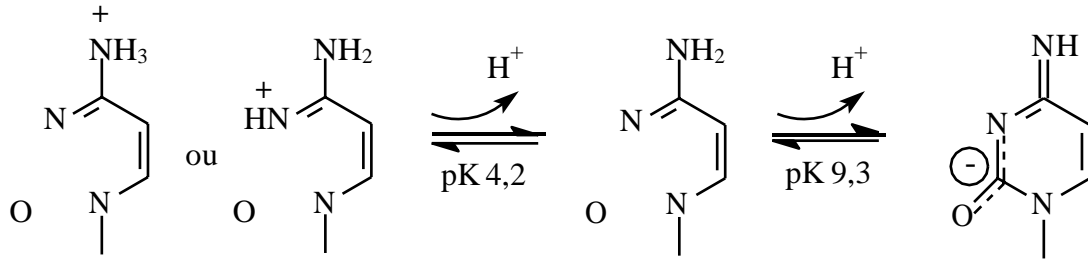
### Les formes tautomères (lactame et lactime)

Nous avons vu que la présence d'un groupement céto lié à un cycle entraînait l'existence de formes tautomères par le passage de la forme céto en forme énol. Cette dernière fonction peut libérer un proton à pH basique (acide faible de pK d'environ 9,2 à 9,5), entraînant l'apparition d'une charge négative.

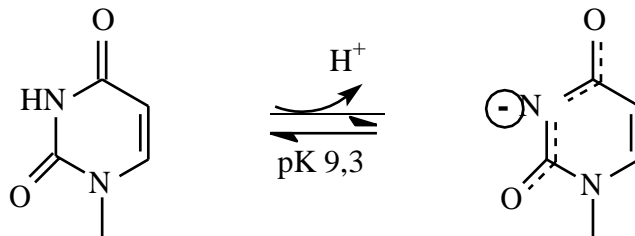
### En guise de conclusion

Les bases puriques et pyrimidiques, liées à un ose (liaison N-osidique), portent des groupes ionisables qui peuvent libérer des protons :

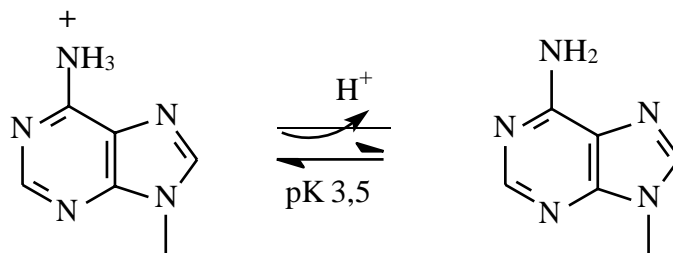
### Cytosine



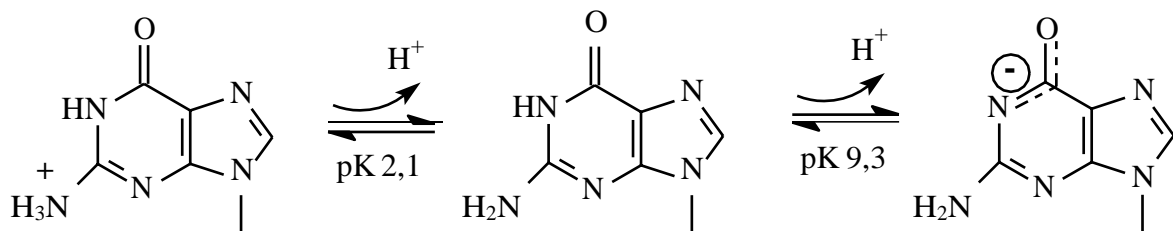
### Uracile ou thymine



### Adénine



### Guanine



A un pH voisin du pH physiologique (6,5), les bases puriques ou pyrimidiques ne portent aucune charge.

Les acides nucléiques dans une solution dont le pH est voisin du pH physiologique sont des **polymères chargés négativement**, charges dont la contribution est uniquement due aux groupements phosphates.

#### ▪ **L'hydrolyse des acides nucléiques :**

La dégradation d'un polynucléotide peut être chimique ou enzymatique, elle concerne :

- l'enchaînement phosphodiester
- les unités nucléotidiques : composants et liaison osidique

### **L'hydrolyse chimique**

#### **Le traitement acide affecte de la même façon les ADN et les ARN**

- la dégradation du squelette phosphodiester est obtenue dans des conditions drastiques (acide concentré et chauffage) auxquelles ne résistent pas les autres liaisons, cette dégradation conduit à la libération d'un mélange de phosphates, oses et bases.

- dans des conditions douces (pH 4), seules les liaisons N-osidique avec les purines sont hydrolysées.

#### **Les ARN et les ADN réagissent différemment à l'hydrolyse alcaline**

- les ADN résistent aux pH basiques : par exemple, à pH 13 et à 37°C on a une dizaine de coupures de ponts phosphodiester par million de ponts

- les ARN sont totalement hydrolysés en leurs ribonucléotides en quelques minutes à 37°C et à pH 11. C'est la présence de l'hydroxyle libre en 2' qui permet cette hydrolyse qui donne un intermédiaire cyclique 2':3' P, aboutissant à des nucléotides 2' P ou 3' P.

### **L'hydrolyse enzymatique**

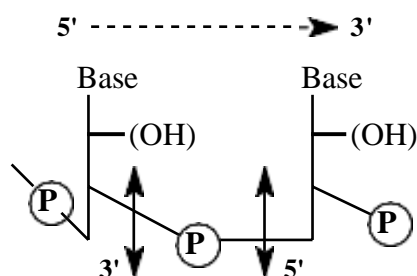
Les enzymes qui catalysent l'hydrolyse de la liaison phosphodiester des acides nucléiques, présents dans la plupart de toutes les cellules, sont des phosphodiesterases spécifiques appelées **nucléases**. Des endonucléases de très haute spécificité sont présentes dans les bactéries, ce sont des désoxyribonucléases désignées sous le nom **d'enzyme de restriction**.

#### **Les nucléases**

Elles présentent des niveaux de spécificité et sont classées par :

- leur mode d'attaque de la chaîne : extrémité (exo) ou intérieur (endo)
- leurs spécificité vis-à-vis du substrat : ADN, ARN ou les deux et de la structure, simple ou double brin

- leur spécificité de reconnaissance des sites : bases ou leur enchaînement
- le type de coupure de la liaison phosphodiester



Différents types de coupure d'un pont phosphodiester

Exemple de quelques nucléases avec leurs spécificités :

nucléases	substrats	type coupure	spécificité coupure
<b>exonucléases</b>			
phosphodiesterase de venin	ARN, ADN (s)	3'	extrémité 3'
phosphodiesterase de la rate	ARN, ADN (s)	5'	extrémité 5'
exonucléase I d' <i>E.Coli</i>	ADN (s)	3'	extrémité 3'
exonucléase III d' <i>E.Coli</i>	ADN (d)	3'	extrémité 3'
<b>endonucléases</b>			
endonucléase S1 d' <i>Aspergillus</i>	ARN, ADN (s, d)	3'	aléatoire
ARNase T1 d' <i>Aspergillus</i>	ARN (s)	5'	-G # N-
ARNase de pancréas	ARN (s)	5'	-Pyr # N-
ADNase II de thymus	ADN (s)	5'	-dPyr # dPur-

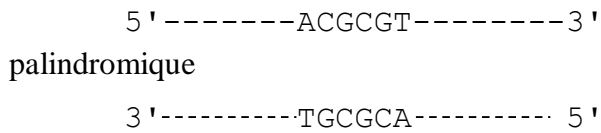
s : simple brin, d : double brin

N : représente n'importe lequel des nucléotides

### Les enzymes de restriction

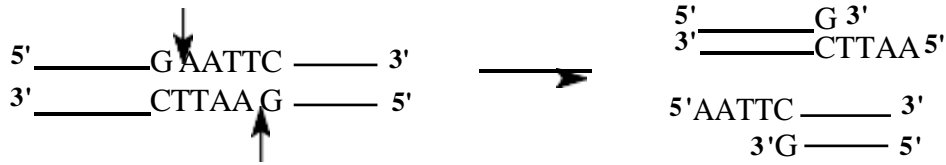
Chaque espèce de bactéries produit une collection d'endo-désoxyribonucléases, dont le site est **spécifique d'une séquence de 4 à 10 nucléotides**, qui leur permet de s'opposer à l'infection de certains virus en hydrolysant leur ADN, sans hydrolyser leur propre ADN protégé par le biais d'une méthylation des séquences partielles spécifiques. On parle de **restriction virale** et on a appelé ces nucléases des enzymes de restriction.

Ces désoxyribonucléases hydrolysent un pont phosphodiester de chacune des 2 chaînes d'acide désoxyribonucléique dont la structure est un double brin. L'ADN double brin est formé de deux chaînes complémentaires de sens opposé (voir plus loin). La plupart du temps la séquence nucléotidique reconnue (lue dans le même sens) est identique sur les deux brins, on dit que le site est **palindromique** :

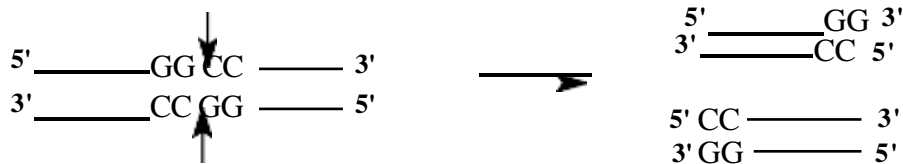


ADN double brin : site

Voici deux exemples d'enzymes de restriction :



Hydrolyse par l'enzyme de restriction : EcoRI  
 site : **GAATTC** (coupure débordante)



Hydrolyse par l'enzyme de restriction : HaeIII  
 site : **GGCC** (coupure franche)

Ces enzymes sont un outil de choix pour toutes les techniques de biologie moléculaire d'autant que le nombre d'enzymes de restriction purifiés est relativement grand : environ 500.

▪ **Structure spatiale des acides désoxyribonucléiques :**

**Quelques résultats expérimentaux sur les acides désoxyribonucléiques**

**La composition en bases**

Les molécules d'ADN présentent la caractéristique suivante :

- quelle que soit l'origine de l'ADN, le nombre de purines est toujours égal au nombre de pyrimidines : [Pur] = [Pyr] ou encore [A] + [G] = [T] + [C]

- de plus, les fractions molaires des bases sont telles que :

$$A = T \text{ et } G = C$$

Cette caractéristique est désignée sous le nom de règle de Chargaff qu'il observa en 1940. Les bases A et T sont dites **complémentaires**, il en est de même pour G et C. Bien sûr les proportions ([A] + [T]) et ([G] + [C]) ne sont pas égales et varient de 35 à 75% selon l'ADN étudié.

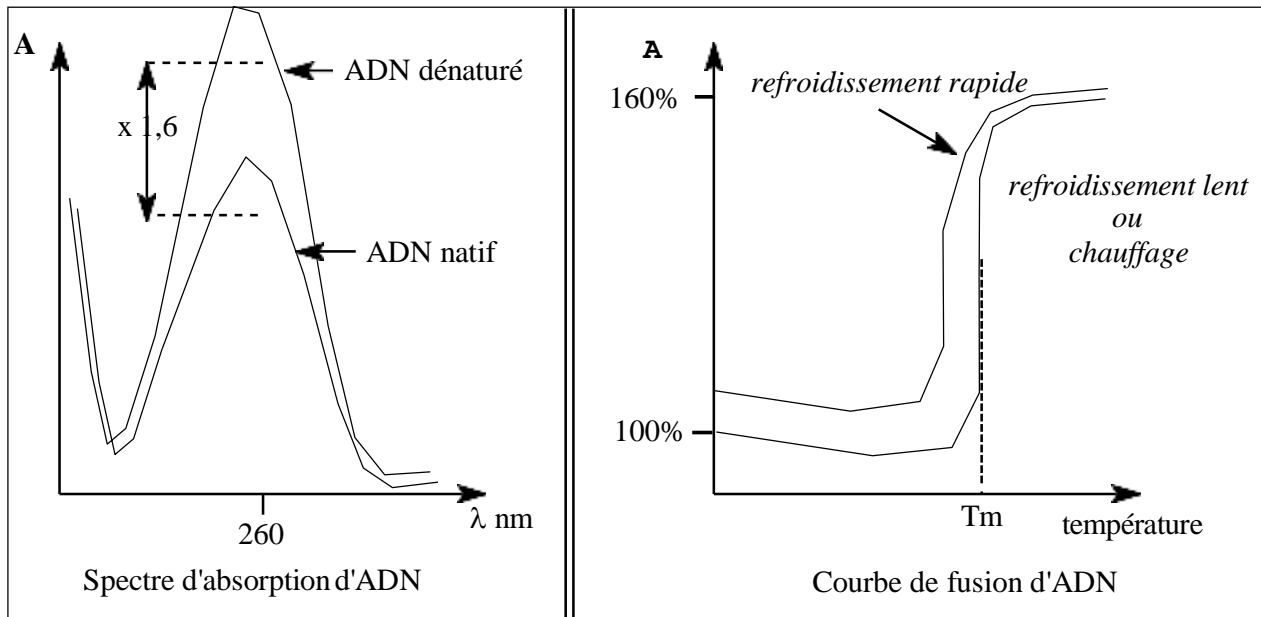


## Le titrage des groupes basiques

Si on titre les bases de l'ADN (charge moyenne de la molécule en fonction du pH) en présence d'un agent dénaturant, comme l'urée, le pK moyen obtenu est de l'ordre de 4,5.

Si on titre les bases de l'ADN dans des conditions non dénaturantes, on trouve un pK moyen augmenté d'une unité, environ 5,5

## L'effet hyperchrome



- Le spectre d'absorption de l'ADN natif n'est pas identique à celui du même ADN dénaturé par la chaleur (chauffage à 100°C) ou par l'urée ou encore à pH très alcalin. L'ADN dénaturé a une absorption à 260 nm plus élevée que l'ADN natif, d'un facteur 1,6. Cette propriété est appelée **l'effet hyperchrome** ou **hyperchromicité**.

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

- Lors d'un refroidissement lent, l'absorption suit la courbe de fusion en sens inverse. Lors d'un refroidissement rapide, l'absorption ne suit pas la courbe de fusion en sens inverse mais une autre courbe qui n'aboutit à la même valeur originale de l'absorption, mais à une valeur plus élevée : c'est le phénomène d'**hystérésis**.

- Il faut noter que la température de fusion  $T_m$  est dépendante de la force ionique du milieu et qu'elle diminue lorsque cette dernière augmente (dans des milieux où  $[NaCl] > 1M$ ).

- La valeur de la température de fusion  $T_m$  est d'autant plus élevée que le pourcentage de bases G + C est grand :

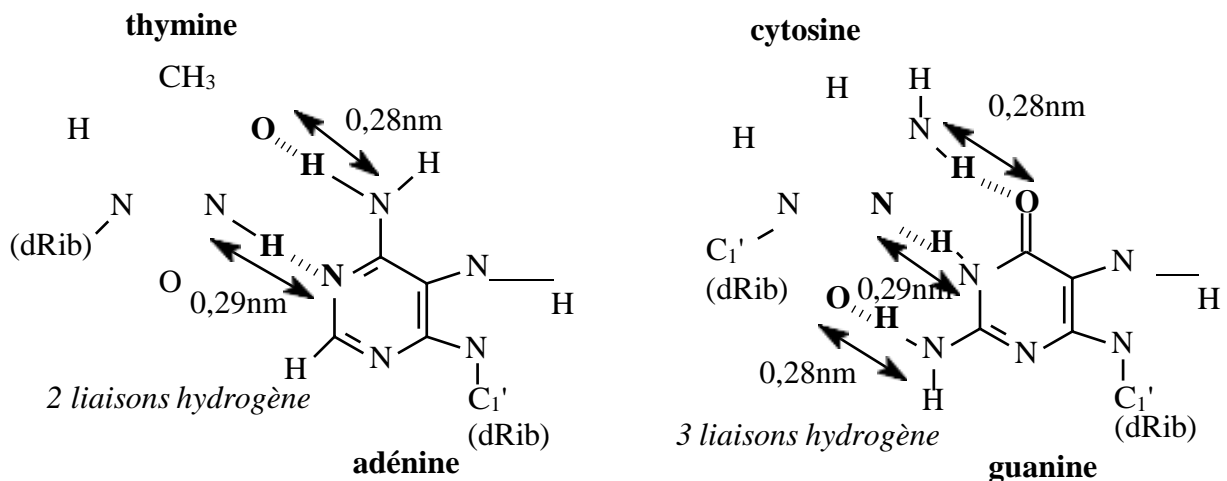
- 65 °C pour l'ADN de *E. Coli* où G+C est égal à 50%,

- 76 °C pour l'ADN de *P. aeruginosa* où G+C est égal à 68%

## La structure révélée par la diffraction aux rayons X

L'analyse aux rayons X de cristaux de molécules d'ADN, réalisée en 1953 par Watson et Crick, indique une structure ayant la configuration suivante :

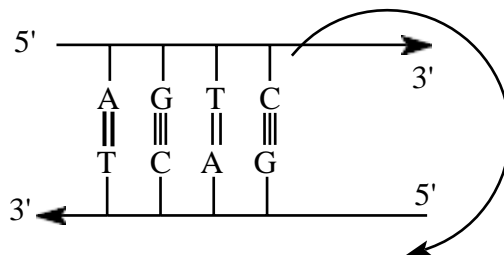
- une double hélice formée de deux chaînes d'ADN dont les paires de bases sont complémentaires
- les deux chaînes sont antiparallèles
- les enchaînements oses-phosphates forment les squelettes hélicoïdaux parallèles extérieurs, les plans des oses étant presque perpendiculaires au plan des bases
- les plans des bases sont perpendiculaires à l'axe de l'hélice et chacune des bases d'une chaîne est appariée à celle de l'autre chaîne par des liaisons hydrogène, les deux bases étant situées dans un même plan
- les bases appariées sont complémentaires (A / T) et (G / C) et sont à l'intérieur du cylindre central de l'hélice.



Appariements pyrimidines-purines selon Watson et Crick

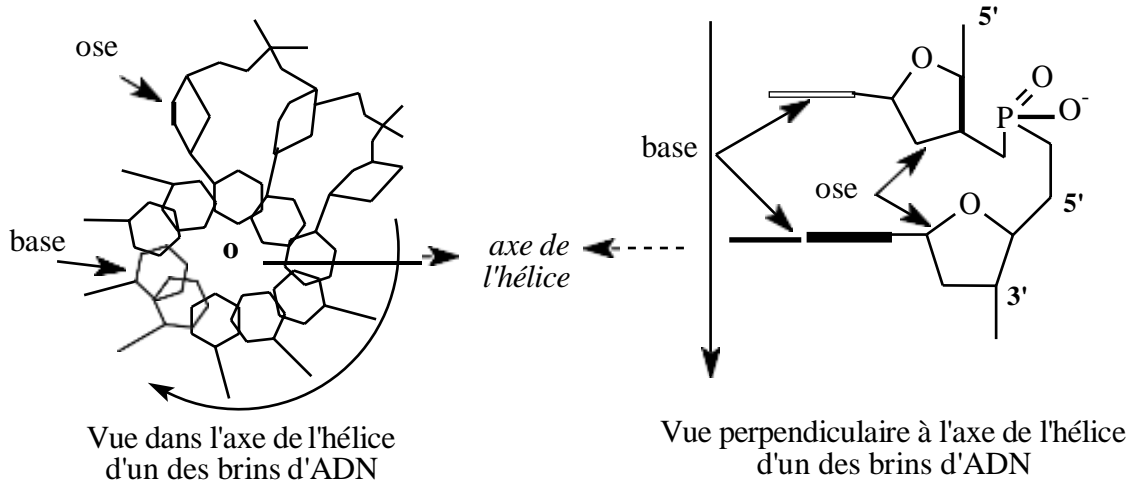
## Les doubles hélices

Une double hélice est formée de deux chaînes antiparallèles d'ADN qui s'enroulent soit à droite (sens du tire bouchon ou sens des aiguilles d'une montre) soit à gauche.

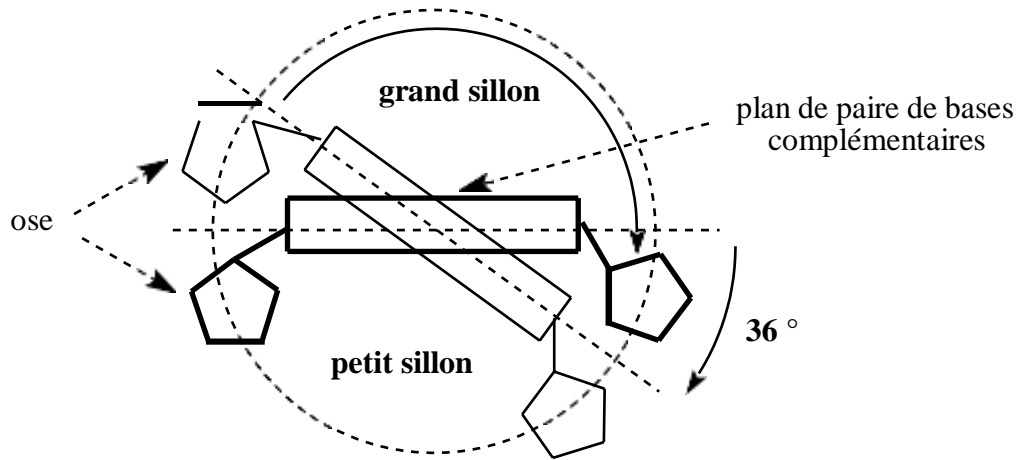


Cette double hélice est caractérisée par :

- son axe principal,
- le sens d'enroulement,
- le pas.



Du fait de la position de l'axe au centre de chaque paire de base complémentaire et de leur attachement dissymétrique sur le squelette ose-phosphate, deux sillons d'ouverture et de profondeurs différentes vont alterner sur les flancs de la double hélice :



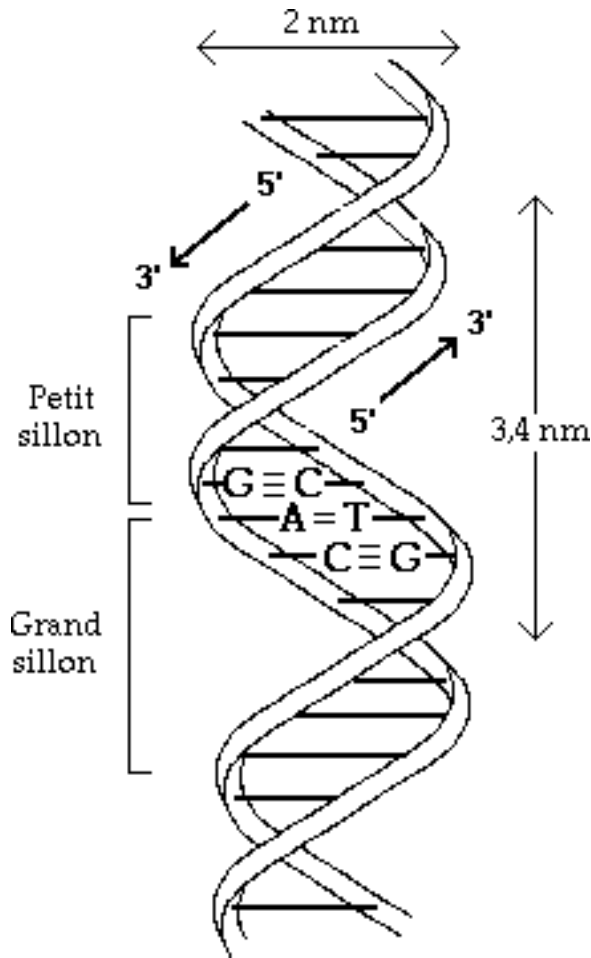
Vue dans l'axe de la double hélice : représentation d'un enchaînement

La stabilité de la structure secondaire de ces différentes double hélices d'ADN est essentiellement due aux :

- liaisons hydrogène entre les bases complémentaires de chacun des brins
- interactions hydrophobes et électrostatiques des bases successives empilées dans la structure de l'hélice, dont la distance des plans varie de 0,26 à 0,37 nm.

Cette stabilité n'entraîne pas une rigidité de la molécule d'ADN et celle-ci peut adopter des conformations différentes selon les régions.

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



**Modèle de Watson et Crick**

**Conformation B :**

c'est le modèle de Watson et Crick, le plus stable dans les conditions physiologiques.

- enroulement droit
- pas : 3,4 nm
- 10 pb par tour
- rotation du plan des bases :  $36^\circ$

**Conformation A :**

- enroulement droit
- pas : 2,8 nm
- 11 pb par tour
- rotation du plan des bases :  $33^\circ$

**Conformation Z :**

- enroulement gauche
- pas : 4,5 nm
- 12 pb par tour
- rotation du plan des bases :  $-30^\circ$

La stabilité de la structure secondaire de ces différentes double hélices d'ADN est essentiellement due aux :

- liaisons hydrogène entre les bases complémentaires de chacun des brins
- interactions hydrophobes et électrostatiques des bases successives empilées dans la structure de l'hélice, dont la distance des plans varie de 0,26 à 0,37 nm.

Cette stabilité n'entraîne pas une rigidité de la molécule d'ADN et celle-ci peut adopter des conformations différentes pour différentes régions.

**Conformation B**

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

C'est une hélice droite de pas égal à 3,4 nm (10 pb par tour).

- L'inclinaison des bases par rapport à leur axe de rotation perpendiculaire à l'axe principal de l'hélice est de 1°.
- La distance entre deux plans de bases successifs est de 0,34 nm.

### **Conformation A**

Lorsque la teneur en eau d'une solution contenant une molécule d'ADN est diminuée, par exemple lors de la cristallisation, la molécule change de conformation et adopte une conformation notée A. Ce changement est réversible.

La conformation A a les spécificités suivantes :

- hélice droite (idem conformation B)
- hélice plus compacte
  - pas de 2,8 nm (11 pb par tour)
  - distance entre les plans de bases successives de 0,26 nm
- les plans des bases sont tournés de 180° par rapport à la conformation B
- l'inclinaison des bases par rapport à leur axe de rotation perpendiculaire à l'axe principal de l'hélice est de 20°.

Cette conformation est trouvée *in vivo* dans :

- l'ADN de certaines spores bactériennes, formées en réponse à la dessiccation du milieu
- les hybrides ADN-ARN qui se forment transitoirement à l'amorce de la réplication, et pendant la transcription.

### **Conformation Z**

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

La conformation Z a les spécificités suivantes :

- **hélice gauche**
- pas de 4,5 nm (12 pb par tour) : la molécule est plus étirée dans cette conformation
- distance entre les plans de bases successives de 0,37 nm
- l'inclinaison des bases par rapport à leur axe de rotation perpendiculaire à l'axe principal de l'hélice est de 9°.

Cette conformation est trouvée *in vivo* pour des segments de la molécule d'ADN, avec des enchaînements alternés Pur/Pyr dont les bases sont souvent méthylées. Celle-ci aurait un rôle dans l'expression des gènes.

Ces différentes conformations décrivent la structure secondaire des molécules d'ADN double brin. Ces conformations sont très stables, mais possèdent une flexibilité assez grande qui peut définir des structures tertiaires (structure dans l'espace d'une double hélice) différentes :

celles-ci vous seront décrites dans des cours ultérieurs et vous verrez, par exemple, les phénomènes de surenroulement.

### **Retour sur les quelques résultats expérimentaux sur l'ADN natif**

La structure en double hélice, les liaisons hydrogène entre les bases complémentaires et les interactions hydrophobes entre bases successives peuvent éclairer les observations du paragraphe 3.4.1 (Quelques résultats expérimentaux sur les acides désoxyribonucléiques) :

- composition en bases : évident
- titrage : l'engagement de l'hydrogène d'un azote dans une liaison hydrogène avec la base complémentaire fait apparaître un doublet électronique sur cet hydrogène qui par délocalisation renforcera la liaison avec un hydrogène supplémentaire. L'ion  $H^+$  sera plus difficilement libéré et le pKa de ce groupe augmentera.
- effet hyperchrome : l'arrangement spatial des bases successives dans la double hélice d'un brin d'ADN, empilement des plans avec une distance pouvant varier de 0,24 à 0,37nm, permet des interactions hydrophobes entre bases successives (contact de Van der Waals) qui ont des répercussions sur le nuage électronique des cycles et donc sur leurs propriétés spectrales. Ce phénomène diminue l'absorption des bases dans l'ultraviolet : effet hyperchrome. Cet effet peut être enlevé en dénaturant la molécule : destruction de la double hélice qui supprime les interactions entre bases qui retrouvent leurs propriétés spectrales originales. La dénaturation, qui libère chacun des deux brins d'ADN, est obtenue par destruction des liaisons hydrogènes soit :
  - par addition d'urée en concentration  $> 6M$
  - par augmentation de la température
- phénomène d'hystérésis : après dénaturation par chauffage, les deux brins d'ADN sont libres. Un refroidissement lent permet un réappariement des deux brins d'ADN qui aboutit à la double hélice originale. Un refroidissement rapide ne permet pas un réappariement total des deux brins, seules certaines régions complémentaires des deux brins reforment des doubles hélices partielles. L'absorption de la molécule aura une valeur intermédiaire entre celle de la molécule native et de la molécule dénaturée : phénomène d'hystérésis.
- température de fusion : la température de fusion de molécule d'ADN en double hélice est dépendante de la composition du pourcentage de bases G + C. Elle est d'autant plus élevée que celui-ci est élevé : ceci peut s'expliquer par le fait que les appariements de ces bases fait intervenir trois liaisons hydrogène alors que les appariements A-T en font intervenir seulement deux. Une présence importante d'ions dans la solution (force ionique élevée : NaCl 1M) perturbera les liaisons hydrogène et provoquera une diminution de la température de fusion.

## Quelques exemples de molécules d'ADN

Les molécules d'ADN des organismes vivants sont toutes des molécules double brin à de rares exceptions près, comme par exemple l'ADN du phage  $\phi$ X174.

Taille de génome en nombre de bases

Organisme	Nb bases (kb)			Nb chromosomes
phage $\phi$ X174	5,4	simple brin	circulaire	
E. Coli	4000	double brin	circulaire	
Drosophile	~ 137 000	double brin	linéaire	4
Drosophile / Chr 1	27 000	double brin	linéaire	
Homme	~ 3 000 000	double brin	linéaire	23
Homme / Chr9	145 000	double brin	linéaire	

Les molécules d'ADN peuvent être circulaires par formation de liaison ester 3'-5' de ses extrémités : certains virus ou bactéries, etc.. L'ADN des organites (chloroplastes ou mitochondries) sont circulaires à l'exception de génomes mitochondriaux de quelques algues et protozoaires.

## Des propriétés physico-chimiques souvent utilisées

La structure de la double hélice donne une nature fibreuse à la molécule d'ADN dont les propriétés sont exploitées dans de nombreuses expériences de biologie moléculaire :

- les alcools, et en particulier l'éthanol, précipitent les molécules d'ADN sous forme d'agglomérat en longues fibres
- la densité des molécules d'ADN est telle qu'on peut les séparer par ultracentrifugation dans des gradients de densité (chlorure de césium)
- la charge de ces molécules à pH physiologique est négative et directement proportionnelle à leur longueur (nb de nucléotides). Cette propriété est utilisée pour les séparer par électrophorèse.
- clonage et séquençage (structure primaire) de l'ADN qui exploitent la complémentarité des bases

## Pour la suite

Dans ce module, nous nous sommes intéressés uniquement aux propriétés essentiellement physico-chimiques de l'ADN. Dans des cours ultérieurs, seront développées les propriétés biologiques de ces molécules :

- molécule support d'information : code génétique, régions spécifiques de régulation, recombinaison génétique
- interactions ADN-protéines et régulation de l'expression génique

▪ **Structure spatiale des acides ribonucléiques :**

L'ARN à deux brins appariés est l'exception de quelques rares virus. Les différents types d'ARN sont des molécules formées d'un seul brin et quelquefois des molécules faisant partie d'hybrides ARN-ADN.

Les principaux types de molécules d'ARN présentes dans les cellules d'organismes vivants sont :

- les **ARN génomiques** : virus (poliovirus, virus de la grippe, etc..), rétrovirus (oncogènes, sida, herpès, etc..)
- les **ARN ribosomiques** (ARNr) [80% ARN totaux] : les ribosomes sont des complexes nucléo-protéique contenant 3 types d'ARNr, qui sont le siège de la biosynthèse des protéines (traduction)
- les **ARN de transfert** (ARNt) [15% ARN totaux] : interviennent dans l'élongation de la chaîne polypeptidique, lors de la traduction
- les **ARN messagers** (ARNm) [5% ARN totaux] : support de l'information transcrite à partir de l'ADN, qui interviennent dans la traduction
- les **ARN hétérogènes nucléaires** (ARNhn) [< 2% ARN totaux] : sont les précurseurs des ARNm
- et les **petits ARN stables** [< 2% ARN totaux], soit cytosoliques (ARNsc), soit nucléaires (ARNsn) qui existent sous forme de ribonucléoprotéines appelées snRNP (*small nuclear ribonucleoproteins*).

L'usage a consacré de caractériser les ARNr, ARNt par leurs propriétés de sédimentation qui sont liées à leur masse volumique, y compris l'eau d'hydratation, et on parle par exemple d'ARN 23S où est le Svedberg ( $1 S = 10^{-13}$  seconde).

Type d'ARN	Nb nucléotides	coeff sédimentation
ARNm	variable	
ARNr	120 - 5000	(5S, 16S, 23S)* (5,8S, 18S, 28S)**
ARNt	75 - 90	4S
ARNhn	variable	
ARNsc	90 - 330	
ARNsn	58 - 220	

\* : procaryotes - \*\* : eucaryotes

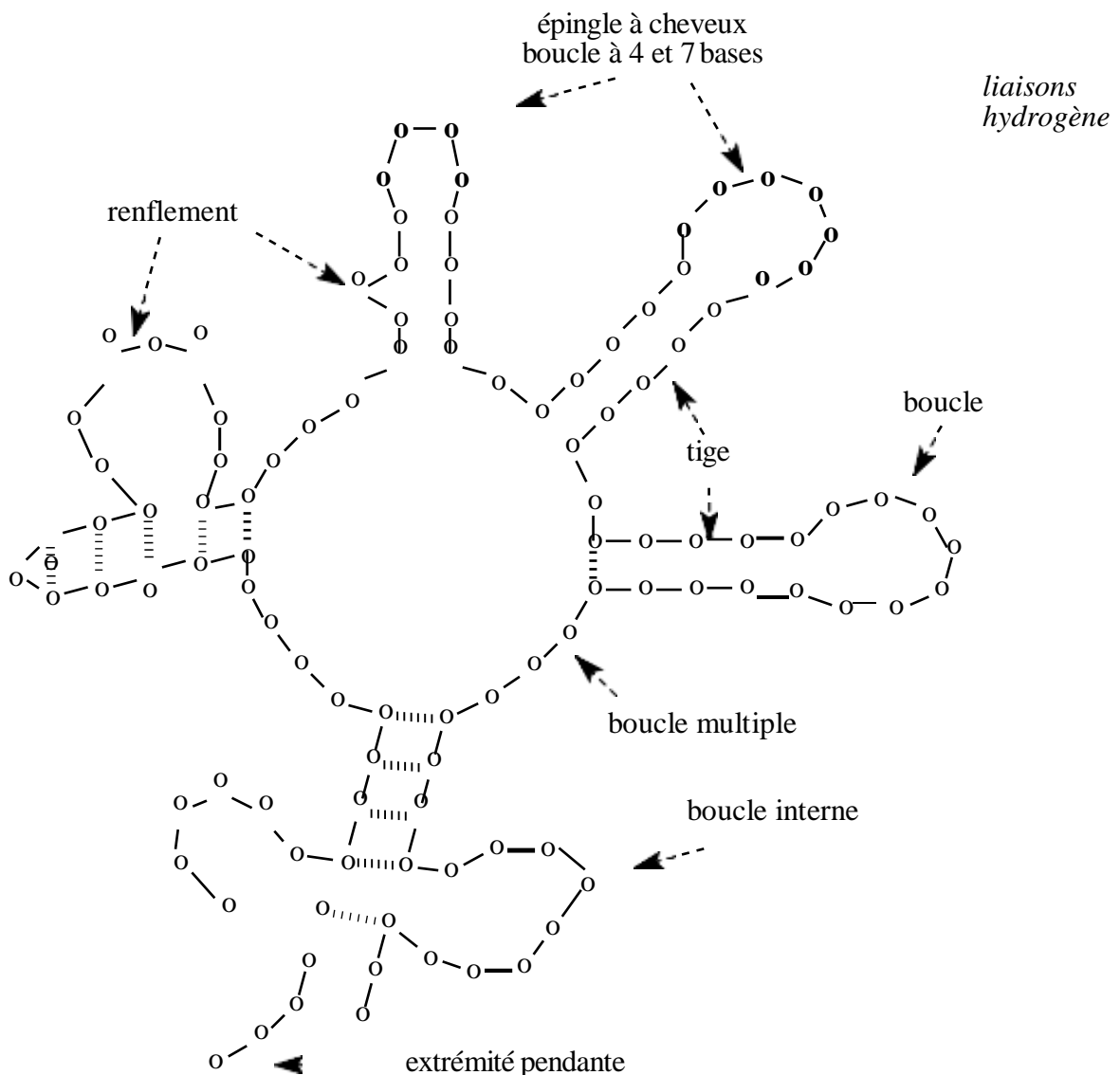
**Les hélices d'ARN**

Sous l'action des forces d'empilement (interaction de Van der Waals) entre bases successives, le squelette d'une molécule d'ARN tend à former spontanément une hélice simple droite irrégulière.



Les conformations stabilisantes sont des régions en double hélice que l'on trouve, en dehors de deux brins d'ARN ou hydrides ARN-ADN, dans des régions où **deux segments distants du même brin d'ARN présentent une complémentarité suffisante** pour un nombre d'appariements supérieur ou égal à 3 (appariement A/U par deux liaisons hydrogène et G/C par trois liaisons hydrogène). Ce type de conformation implique des structures secondaires de type épingle à cheveux.

### Les motifs élémentaires de structure secondaire d'ARN



Ces différents motifs de structure secondaire ont été trouvés dans les ARN ribosomiques et les ARN de transfert :

- les tiges sont des hélices dont la conformation est proche de celle de la conformation A de l'ADN : les hydroxyles en 2' des riboses s'opposent à un enroulement de type B.

-les boucles : celles participant à un motif d'épingle à cheveux stabilisent cette structure, en particulier les tétraboucles de séquence UNCG (N représentant l'une des quatre bases).

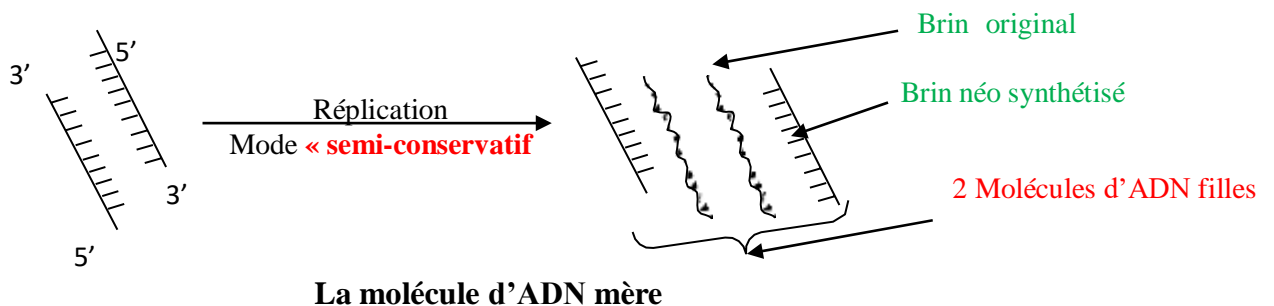
### **Pour la suite**

Dans des cours ultérieurs, seront développées les propriétés biologiques et les structures des molécules d'ARN :

- molécule support d'information : transcription de l'ADN en ARN et traduction des ARN messagers
- interactions ARN-protéines: biosynthèse des protéines avec en particulier la structure des ARN ribosomiques et des ARN de transfert (feuille de trèfle).

## II. La réplication :

- C'est la synthèse d'ADN reproduit exactement par le même génome d'une cellule avant sa division
- La réplication est sous la mode « semi-conservatif », chaque molécule d'ADN copie qui contient un brin dérivé de la molécule mère et un brin néo synthétisé (nouveau).



### 1. La double hélice :

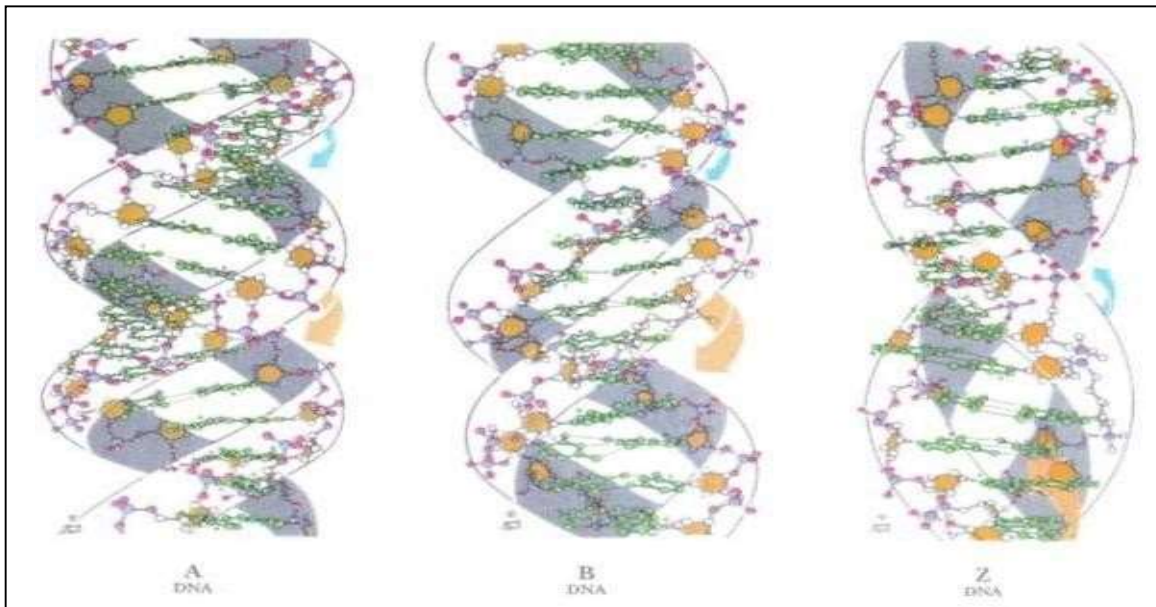
Découvert par Watson et Crick en 1953, sur la base de photo de diffraction aux rayons X prise par certains Franklin et Wilkins

- L'ADN formé par deux (02) chaînes reliées entre elles pour former la double hélice.
- Il effectue une tournée pour 10 paires de base (Pb), le pas de l'hélice est environ **34 Å**
- La distance moyenne entre 2 bases est environ **3.4 Å**, avec un diamètre de **20 Å**
- Les deux hélices ne sont régulières on distingue un sillon majeur et un sillon mineur, c'est très important dans l'interaction avec les protéines, dans sa réplication, et aussi dans l'expression de l'information génétique
- La double hélice elle est dextre, le squelette (sucre et phosphate) est sur la droite.

#### a. Les différentes formes de la double hélice

L'ADN double brin peut avoir trois structures secondaires : A, B et Z, La forme B est généralement la forme la plus fréquente dans la cellule. La forme d'ADN-A est la plus compacte

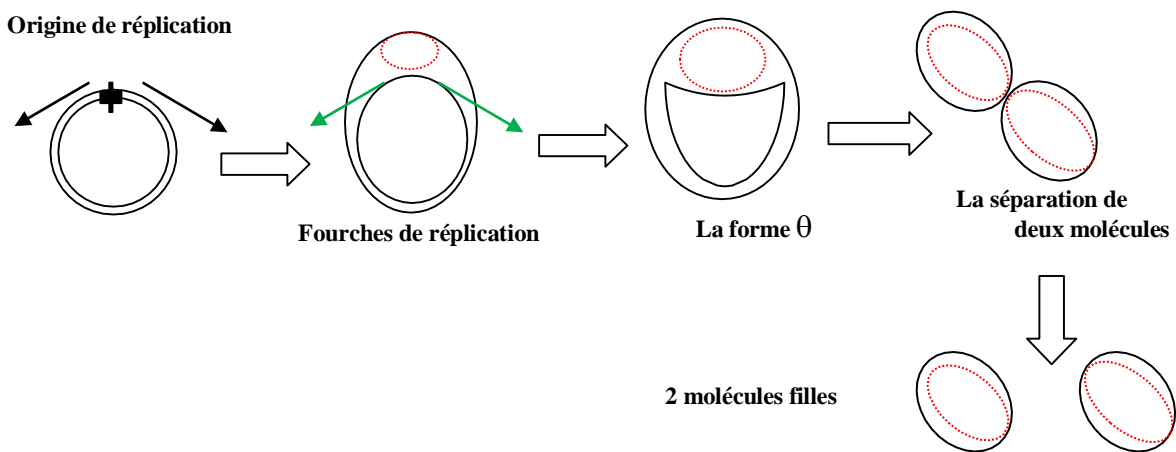
La forme Z est identifiée sur des régions chromosomiques qui présentent une double hélice tournée vers la gauche.



**Règle : la synthèse d'ADN est toujours dans le sens 5' --- -- 3'**

**b. La réplication chez les procaryotes**

L'ADN chez les procaryotes est circulaire = **Chromonème (chromosome procaryote)**



**Les étapes de réplication d'ADN chez les procaryotes**

La réplication commence à partir d'un endroit appelé "origine de réplication" il est à l'origine de formation de deux brins ce qui a appelé « **une fourche de réplication** » et se progressé en les 02 sens opposés.

On obtient une forme intermédiaire appelée  $\theta$  les fourches finissent par une région de fusion (La réplication s'achève).

L'ADN des procaryotes se réplique à partir d'une seule origine de réplication appelée : **Réplicon.**

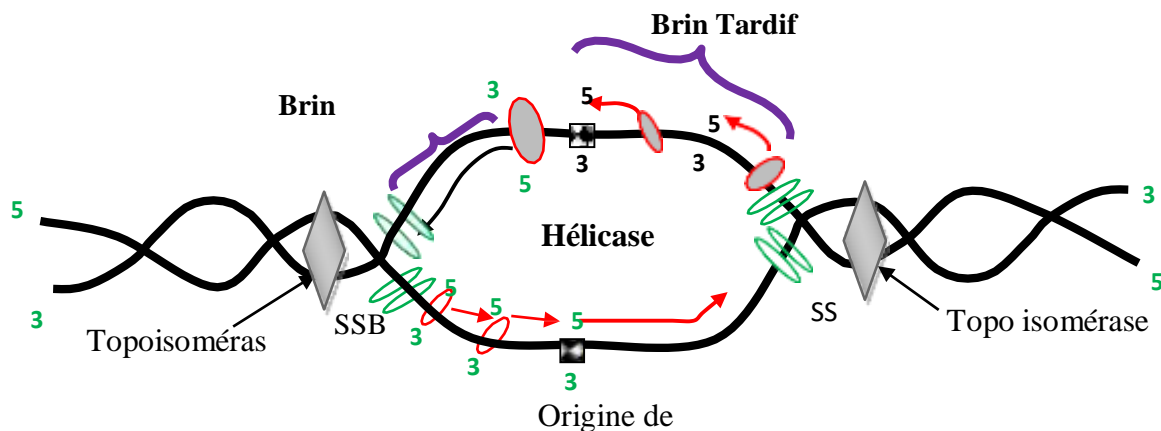
### c. La réplication chez les eucaryotes

- Les chromosomes des eucaryotes sont très longs, pour cela il existe plusieurs origines de réplication permettent de se répliquer l'ADN dans un temps raisonnable (8heures).
- Les 2 fourches de réplication progressent dans les sens différents puis sont entrées en **collision (fusion de l'ensemble).**
- L'ADN des eucaryotes possède plusieurs réplicons.
- Une cellule mammifère possède 50milles à 100 milles réplicons chacun représente une copie d'ADN de 40à 200 Kilos base

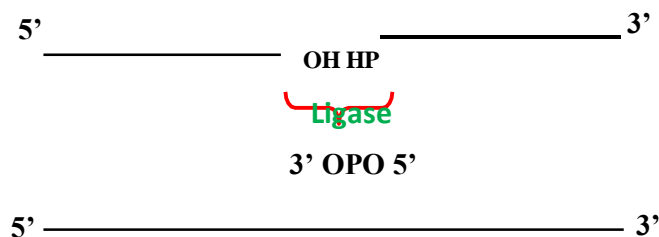
### d. Les étapes de mécanisme génétique de réplication (*in vivo*)

**1- Formation des fourches de réplication :** la double hélice d'ADN commence à dérouler progressivement au niveau de l'origine de réplication, ce déroulement se fait dans les deux sens grâce à un enzyme qu'appelé **Topoisomérase.**

**2- Séparation de deux brins :** un enzyme appelle **Hélicase** sépare les 2 brins, puis les protéines **SSB (Single Stand Binding)** ce lies aux brins pour les garder séparé durant le temps de réplication.



- 3- **Synthèse du brin précoce et brin tardif** : la synthèse de brin précoce se fait de façon continue par contre la synthèse de brin tardif se fait de façon discontinue appelée **Fragment OKAZAKI**
- 4- **Amorçage** : les ADN polymérase ne peuvent pas ajouter des nucléotides qu'à des chaînes nucléotides pré existantes, ainsi le début de réplication soit pour le brin précoce et tardif nécessite la synthèse de fragment d'ARN (de 4 à 12 n) pour servir de point de départ d'amorçage à la polymérisation  
 Pour le brin précoce un seule amorce initiale et nécessaire par contre le brin retardé (tardif) chaque fragment Okazaki a besoin sa propre amorce.  
 Des amorces sont synthétisées par un ensemble de protéines appellent **primosomes** par le principal enzyme c'est le **primase**.
- 5- **La ligation** : qu'après l'enlèvement des amorces d'ARN les vides laissés sont remplacés par des fragments d'ADN synthétisé par ADN polymérase, après avoir finir un autre enzyme (**Ligase**) intervient pour relier les fragments OKAZAKI par les liaisons **Phosphodiester**

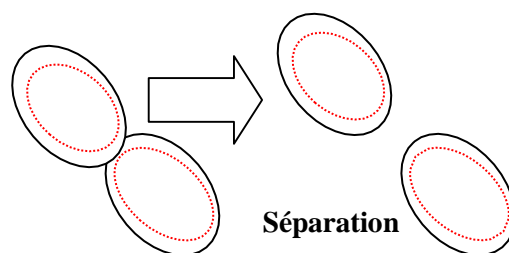


## 2. Particularité du mécanisme de la réplication chez les procaryotes :

On compte 3 ADN polymérase chez les eucaryotes

1. **ADN polymérase type I** : c'est l'enzyme de synthèse de l'ADN dans le sens 5' --- 3', possède deux activités une **Endonucléase** (synthèse) 5' --- 3' et une autre **Exonucléase** 5' --- 3' (Cette activité permet de corriger les erreurs de synthèse de l'ADN).
2. **ADN polymérase type II et III** : permettent de continuer la polymérisation, élongation, catalyse et adhésion des nucléotides à l'extrémité 3'

Les 2 molécules d'ADN sont séparées (après la réplication) par l'enzyme **Topoisomérase** qui il peut de produire une rupture transitoire, sont l'ADN polymérase type II et III qui intervient pour rejoindre les 2 extrémités de brin cassé.

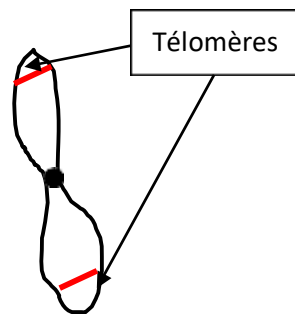


### 3. Particularité du mécanisme de la réplication chez les eucaryotes :

**Les télomères :** Ce sont des séquences répétitives d'ADN, ils sont situés aux extrémités de chromosome

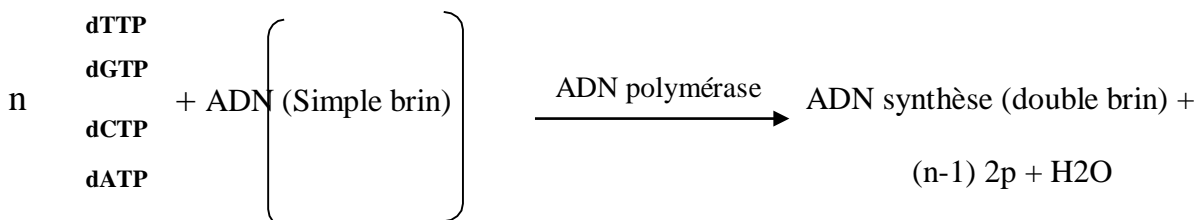
**Exemple 1 :** une séquence télomère chez les protozoaires ciliés (Tétrahymétra) qu'est riche en Guanine 5'TTGGGC3'

**Exemple 2 :** Chez l'homme une séquence télomère riche aussi en Guanine 5'TTAGGG3' cette séquence est synthétisé par un enzyme appelé Télomérase, dont le rôle principal est protégé l'ADN codant



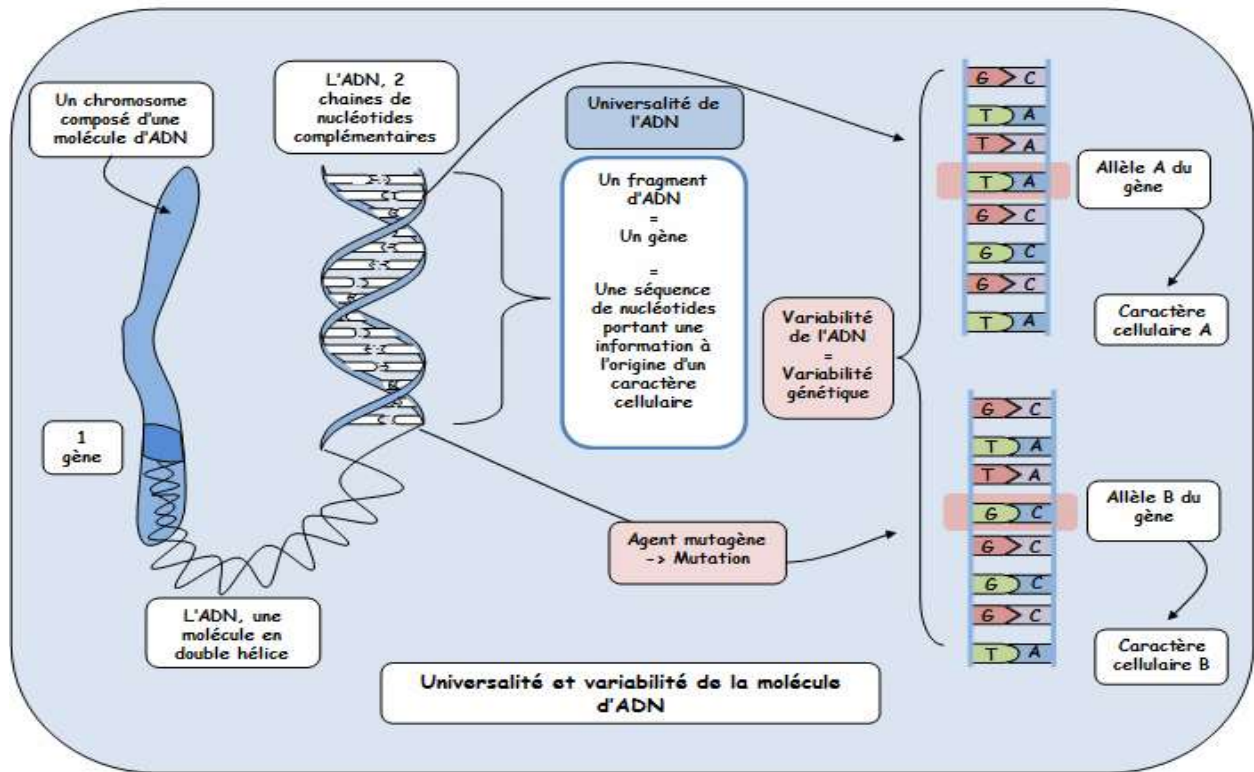
➤ **Synthèse d'ADN in vitro** (Dans un milieu synthétique) Pour synthétiser un ADN in vitro on met en présence :

- 1- ADN matrice : simple brin (séparation par température)
- 2- Les nucléotides triphosphates : (dTTP, dGTP, dCTP, dATP)
- 3- Un enzyme de l'ADN polymérase et des ions  $Mg^{+2}$



Une technique de polymérisation d'ADN *in vitro* utilisé dans la recherche est appelée : Polymérase Chain Réaction (PCR).

#### 4. Organisation de l'ADN dans la chromatine :

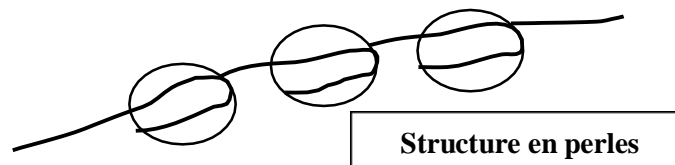


La chromatine c'est une structure associant l'ADN et les protéines dans noyau pendant l'interphase.

La chromatine est constituée principalement d'ADN avec 30% et des protéines de type Histones (H1, H2A, H2B, H3 H4) et une faible quantité d'ARN.

Structure de **nucléofilament** : la structure de base de chromatine est nucléofilament, tout simplement c'est une molécule d'ADN associée à des protéines, de diamètre 10nm. Le point d'association entre l'ADN et la protéine est appelé : **Nucléosome**.

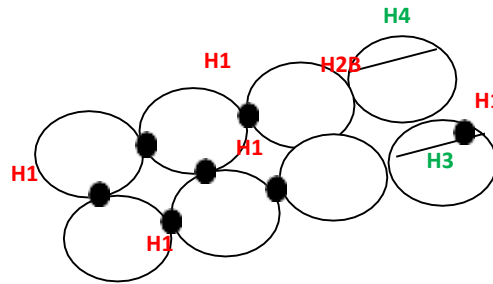
#### Nucléosome



Le segment d'ADN qui entoure la protéine est de longueur 200 pb. Au niveau du nucléosome l'ADN effectue deux (02) tours séparés et successifs.



**La fibre chromosomique** : l'association des nucléosomes permet un degré de repliement simple grâce au 5ème type des Histones (H1) en obtenant la fibre chromosomique.



**Organisation en boucle de la chromatine** : La fibre chromosomique s'organise en suite en domaine sous forme des boucles. Ex : chromosome humain à une taille moyenne comporte 243 milles domaines.

### III. Application à la PCR et au Séquençage :

#### 1. Amplification *in vitro* des acides nucléiques PCR et RT-PCR :

Cette méthode de biologie moléculaire a été mise au point en 1985 par Kary Mullis, qui obtint pour ces travaux le prix Nobel de Chimie en 1993. Aujourd'hui, ce procédé révolutionnaire couplé à l'utilisation d'une ADN polymérase thermorésistante permet d'obtenir, sans clonage, une amplification considérable d'un fragment donné d'ADN.

##### Principe :

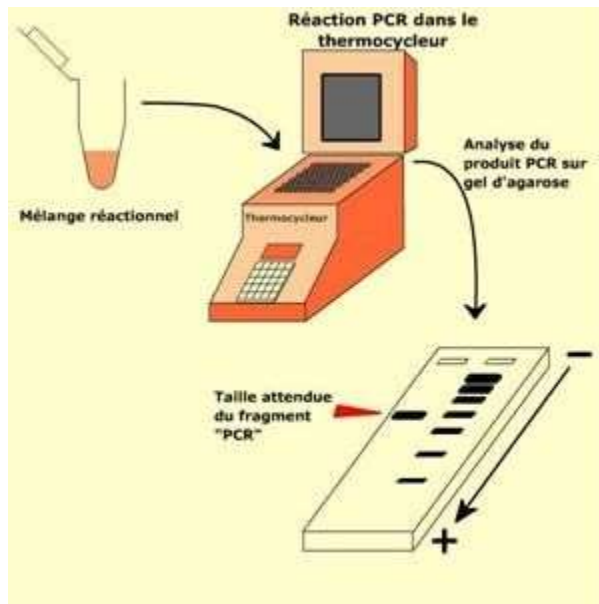
La réaction PCR (Polymerase Chain Reaction) permet d'amplifier *in vitro* une région spécifique d'un acide nucléique donné afin d'en obtenir une quantité suffisante pour le détecter et l'étudier.

Pour se faire, une série de réactions permettant la réplication d'une matrice d'ADN double brin est répétée en boucle. Ainsi, au cours de la réaction PCR, les produits obtenus à la fin de chaque cycle servent de matrice pour le cycle suivant, l'amplification est donc exponentielle.

Pour avoir réplication d'un ADN double brin, il faut agir en trois étapes.

- 1- Il faut dénaturer l'ADN pour obtenir des matrices simple brin : **dénaturation**.
- 2- Borner et amorcer la réplication de la séquence à amplifier à l'aide d'oligonucléotides amorces spécifiques : **hybridation**.
- 3- Réaliser la réaction de polymérisation du brin complémentaire. A la fin de chaque cycle, les produits sont sous forme d'ADN double brin : **polymérisation**.

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique. Pour effectuer ces transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil programmable : un thermocycleur. Cet appareil permet d'exposer les tubes à des températures choisies et pour des durées déterminées par l'expérimentateur. La réaction PCR est extrêmement rapide et ne dure que quelques heures (2 à 3 heures pour une PCR de 30 cycles).



### Réaction d'amplification d'ADN dans le thermocycleur et visualisation du produit de PCR par électrophorèse.

#### ▪ Les amorces

Dans la mise au point de la réaction PCR, le choix des amorces est crucial. Elles vont avoir un double rôle : en s'hybridant à l'ADN matrice, elles délimitent la région d'ADN à amplifier (étape 2 du cycle) et avec leur extrémité 3' OH libre servir d'amorce pour l'ADN polymérase (étape 3 du cycle).

Les oligonucléotides amorces s'hybrident aux extrémités de la séquence qui va être amplifiée, il faut donc connaître les séquences nucléotidiques des extrémités de la région ADN amplifiée. C'est en effet au niveau de ces extrémités que les oligonucléotides amorces vont s'hybrider. Pour faciliter le choix des séquences des deux amorces, des logiciels d'analyse de séquences permettent de vérifier les points suivants :

- des **T<sub>m</sub>** comparables. Les oligonucléotides amorces doivent s'hybrider à l'ADN matrice dans les mêmes conditions de température
- des séquences qui ne soient pas complémentaires entre elles (en particulier dans la région 3')

- des séquences qui ne contiennent pas de séquences répétées inversées (pas de repliement intramoléculaire).

Il est alors possible de faire synthétiser les deux oligonucléotides (d'une vingtaine de bases). La synthèse chimique d'ADN permet d'obtenir des oligonucléotides dont l'extrémité 5' n'est pas phosphorylée (5'OH) contrairement aux ADN naturels. Ainsi, tous les fragments d'ADN amplifiés par PCR sont déphosphorylés à leurs extrémités 5'.

### ▪ Les températures

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique :

- l'étape de **dénaturation**, est réalisée à environ 95°C, pour une dissociation complète des deux brins d'ADN.
- l'étape d'**hybridation** se fait à une température qui sera définie selon la nature des amorces (cette température varie de 50 à 60°C). Cette température va déterminer la stabilité des hybrides une fois que l'appariement amorces / matrice est réalisé.
- l'étape de **polymérisation** est à environ 72°C, température de « travail » de l'ADN polymérase thermorésistante utilisée. Au cours de cette étape, les brins complémentaires d'ADN sont synthétisés à partir des extrémités 3'OH libre des amorces hybridées.

Les températures de dénaturation et de polymérisation sont fixes, seule la température d'hybridation devra être calculée pour chaque nouvelle PCR. Cette température d'hybridation dépend de la composition en bases des oligonucléotides amorces. Elle est légèrement inférieure (environ de 5°C) au T<sub>m</sub> qui est la température de demi-dénaturation.

Pour calculer le T<sub>m</sub> d'un oligonucléotide inférieur à 30 nucléotides, on utilise la relation suivante :

$$T_m = 2(A + T) + 4(G + C)$$

(Où A, T, G et C sont respectivement le nombre de chacune de ces bases dans l'oligonucléotide).

## ▪ L'ADN matriciel

En théorie une copie ADN de la séquence recherchée est suffisante pour avoir une amplification, il faut cependant tenir compte de la probabilité de « rencontre » des molécules d'ADN matrice avec les amorces. Dans la pratique plusieurs copies sont nécessaires pour avoir un résultat correct. Mais attention, la mauvaise qualité et/ou une quantité trop importante d'ADN matrice peut conduire à une amplification aspécifique, voire à une inhibition enzymatique. Il est possible d'expliquer cette inhibition par la présence de contaminants provenant de l'échantillon ou des réactifs utilisés dans les protocoles d'extraction d'ADN.

## ▪ L'ADN polymérase

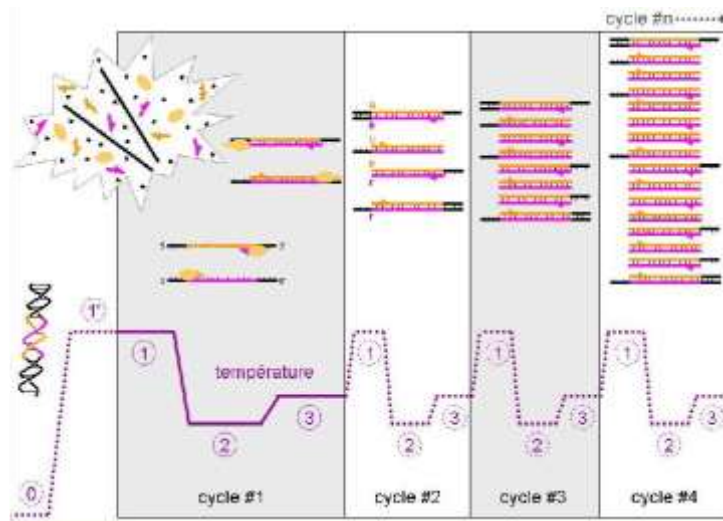
Les premières ADN polymérases utilisées provenaient d'une bactérie thermophile (résistante à des températures très élevées), comme par exemple *Thermus aquaticus* (Taq polymérase). De nos jours, les enzymes utilisées sont dites recombinantes, ce qui simplifie considérablement leur obtention, et leurs propriétés ont été largement modifiées pour les rendre plus efficaces et plus fidèles.

## ▪ Le tampon

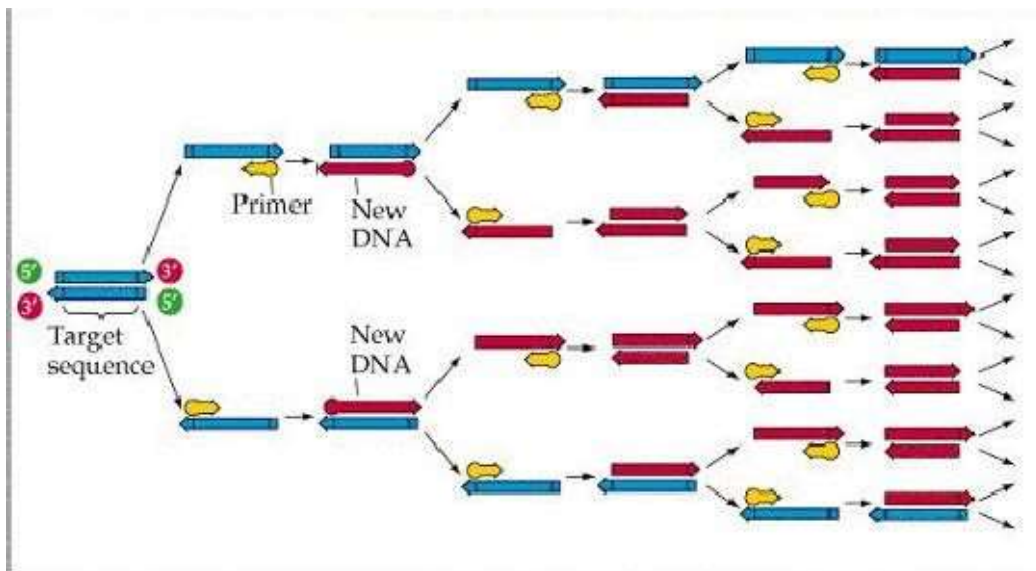
Le tampon utilisé pour la réaction PCR sert à maintenir stable le pH du milieu réactionnel au niveau optimal pour la Taq polymérase (Tris HCl à pH basique 8,5 à 9). Il contient des cations bivalents  $Mg^{2+}$ , cofacteurs indispensables pour la réaction de polymérisation avec la Taq polymérase. La présence dans le milieu réactionnel des cations bivalents  $Mg^{2+}$  et de cations monovalents ( $K^+$  ou  $NH_4^+$ ) vont neutraliser les charges négatives des groupements phosphates au niveau de l'ADN et ainsi stabiliser les hybrides ADN/ADN. En pratique, la concentration en sel doit cependant rester compatible avec l'activité de l'ADN polymérase.

## ▪ Aspects quantitatifs de la réaction PCR

Le nombre de cycle est défini par l'expérimentateur selon le degré d'amplification souhaité (n), dans la plupart des cas, ce nombre est d'environ 30 (seulement en théorie!) il y a un facteur d'amplification de l'ordre de  $2^{30}$  ( $2^n$ ) soit un rendement d'amplification d'un milliard de copies d'ADN. En pratique les conditions expérimentales évoluent au fur et à mesure de l'avancée des cycles, le rendement de la réaction de polymérisation évolue de même, on peut espérer obtenir avec 30 à 40 cycles, un facteur d'amplification de l'ordre de  $10^5$  à  $10^6$ .



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



**Amplification d'ADN par PCR. Le produit d'un cycle sert de matrice pour le cycle qui suit.**

## **2. RT-PCR :**

La PCR en temps réel (RT-PCR= Real Time PCR), se déroule en deux phases. Une première phase correspond à la copie d'ARN messenger en ADN complémentaire (ADNc) et une seconde phase correspond à une réaction PCR classique sur l'ADNc synthétisé.

Dans la première phase, l'ARN messenger à étudier est repéré en utilisant une sonde oligonucléotidique spécifique (amorce 1 qui s'hybride à l'extrémité 3' du seul mARN auquel on s'intéresse), puis la transcriptase inverse (ou rétrotranscriptase) permet la synthèse du brin complémentaire (sous une forme de ADNc simple brin), une seconde amorce oligonucléotidique spécifique (amorce 2) permettra la synthèse du second brin par extension.

L'ADN complémentaire synthétisé servira ensuite de matrice pour une réaction PCR classique.

La technique RT-PCR a permis de montrer que la transcription de tous les gènes s'effectuait dans tous les tissus et ceci même pour les gènes qui présentent une très grande spécificité tissulaire. On parle dans ces conditions de transcription illégitime. Il est évident qu'avant les techniques d'amplification génique, la sensibilité des méthodes classiques n'avait pas permis de mettre en évidence un tel phénomène.

### **3. Utilisations des produits PCR :**

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

-Mise en évidence de mutations ponctuelles par hybridation des produits PCR avec des sondes oligonucléotidiques :

Les produits PCR peuvent après transformation en monobrins être hybridés avec des sondes oligonucléotidiques fixées sur un support solide. Ces sondes correspondent à des séquences normales et pathologiques (présence de mutations ponctuelles par exemple) pour un gène donné.

-Analyse de restriction :

Le produit PCR est soumis à une digestion enzymatique par une enzyme de restriction. Si une mutation ponctuelle modifie le site de restriction initialement présent, la taille des fragments d'ADN obtenus après digestion sera modifiée et décelable après électrophorèse des fragments d'ADN (sur gel d'agarose ou gel de polyacrylamide).

-Introduction du produit PCR dans un vecteur: clonage du produit PCR.

-Séquençage direct du produit PCR.

## IV. Le séquençage de l'ADN :

La séquence d'ADN contient l'information nécessaire aux êtres vivants pour survivre et se reproduire. Déterminer cette séquence est donc utile aussi bien pour les recherches visant à savoir comment vivent les organismes que pour des sujets appliqués. En médecine, elle peut être utilisée pour identifier, diagnostiquer et potentiellement trouver des traitements à des maladies génétiques et à la virologie. En biologie, l'étude des séquences d'ADN est devenue un outil important pour la classification des espèces.

Les premières techniques de séquençage ont été développées en parallèle au milieu des années 1970. Les méthodes de **Sanger** (Grande-Bretagne) et **Gilbert** (Etats-Unis) ont toutes deux été récompensées d'un prix Nobel de chimie en 1980.

Le premier organisme a été séquencé en 1977. Il s'agissait du virus bactériophage  $\phi$ X174, possédant un ADN simple brin ne nécessitant donc pas l'étape de dénaturation utilisé dans les méthodes de Sanger et Maxam et Gilbert.

Depuis une trentaine d'année, l'amélioration de la technique de production des amorces, de l'amplification des brins et la généralisation des traceurs fluorescents ont permis d'améliorer considérablement les techniques de séquençage. L'apparition des **séquenceurs automatiques** a notamment permis l'automatisation de ces technologies et ont contribué à la finalisation du génotypage humain en 2003.

Le **séquençage de l'ADN** est une technique d'analyse de l'ADN, consiste à déterminer l'ordre d'enchaînement des nucléotides pour un fragment d'ADN donné. Actuellement la technique de séquençage repose majoritairement sur la méthode enzymatique de Sanger.

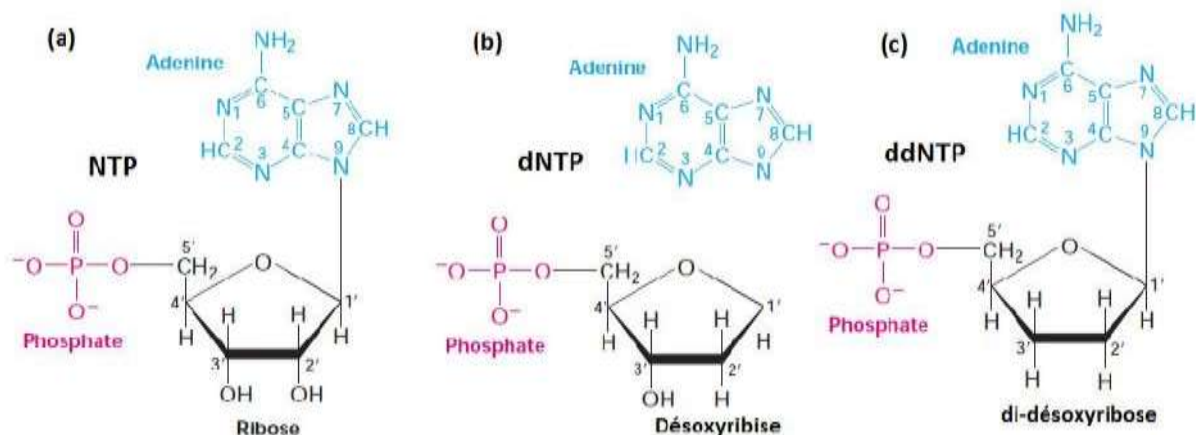
### 1. Les différents acteurs du séquençage :

- l'ADN provient des organismes dont on souhaite séquencer le génome. L'ADN, le plus souvent sous la forme double-brin, est dénaturé afin de séparer les deux brins. Celui qui sera séquencé s'appelle le brin « matrice »



- les **nucléotides**, ou plutôt désoxynucléotides, sont les bases de l'ADN (A, C, G ou T). Ils sont attachés les uns aux autres grâce à des liaisons chimiques nécessitant la présence d'un groupement OH particulier (sur le carbone 3' du ribose)
- les **didésoxynucléotides** sont des nucléotides privés du groupement OH en position 3'. Leur incorporation dans une chaîne d'ADN interrompt définitivement la synthèse de l'ADN (voir figure 1).
- une **amorce** est un brin d'ADN très court (une vingtaine de nucléotide), qui peut s'hybrider à une séquence complémentaire spécifique
- l'**ADN polymérase** est une enzyme qui a pour rôle de copier l'ADN, en synthétisant un brin complémentaire au brin matrice. L'ADN polymérase ne fonctionne qu'en ajoutant des nucléotides à **une amorce** déjà présente, en fonction de la succession de nucléotides du brin matrice (un A est toujours positionné en face d'un T et un C toujours en face d'un G).

## 2. Méthode de Sanger :



Les différentes formes des nucléotides.

Dans le ddATP, le **groupement 3'-OH** est remplacé par un **hydrogène**. Cette modification empêche la poursuite de la synthèse de l'ADN qui continue normalement sur le 3-OH

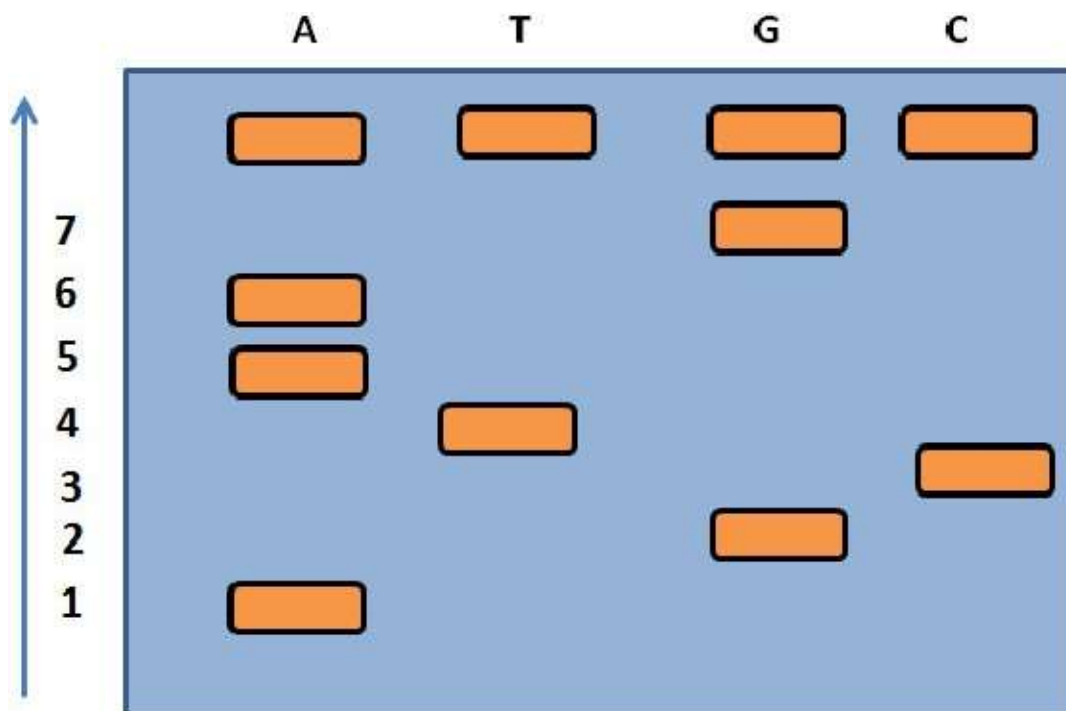
### ▪ Principe du séquençage par la méthode de Sanger

Le principe de cette méthode consiste à initier la polymérisation de l'ADN à l'aide d'un petit oligonucléotide (amorce) complémentaire à une partie du fragment d'ADN à séquencer. L'élongation de l'amorce est réalisée par une ADN polymérase I dépourvue d'activité exonucléase 5'→3' et maintenue par des ADN polymérases thermostables, celles qui sont

utilisées pour la PCR. On utilise quatre tubes séparés contenant les quatre désoxyribonucléotides (dATP, dCTP, dGTP, dTTP), ainsi qu'une faible concentration de l'un des quatre didésoxyribonucléotides (ddATP, ddCTP, ddGTP ou ddTTP).

Ces didésoxyribonucléotides agissent comme des « poisons » terminateurs de chaîne : une fois incorporés dans le nouveau brin synthétisé, ils empêchent la poursuite de l'élongation. Cette terminaison se fait spécifiquement au niveau des nucléotides correspondant au didésoxyribonucléotide incorporé dans la réaction. Pour le séquençage complet d'un même fragment d'ADN, on répète cette réaction quatre fois en parallèle, avec les quatre didésoxyribonucléotides différents.

Par exemple, dans la réaction où on a ajouté du ddGTP, la synthèse s'arrête au niveau des G. Le mélange réactionnel contenant, à la fois du dGTP et un peu de ddGTP, la terminaison se fait de manière statistique suivant que l'ADN polymérase utilise l'un ou l'autre de ces nucléotides. Il en résulte un mélange de fragments d'ADN de tailles croissantes, qui se terminent tous au niveau d'un des G dans la séquence. Ces fragments sont ensuite séparés par électrophorèse sur un gel de polyacrylamide, ce qui permet ainsi de repérer la position des G dans la séquence.



**Séquençage de l'ADN par la méthode de Sanger.**

La détermination de l'enchaînement des nucléotides dans un fragment d'ADN simple brin par la méthode de Sanger comporte deux étapes:

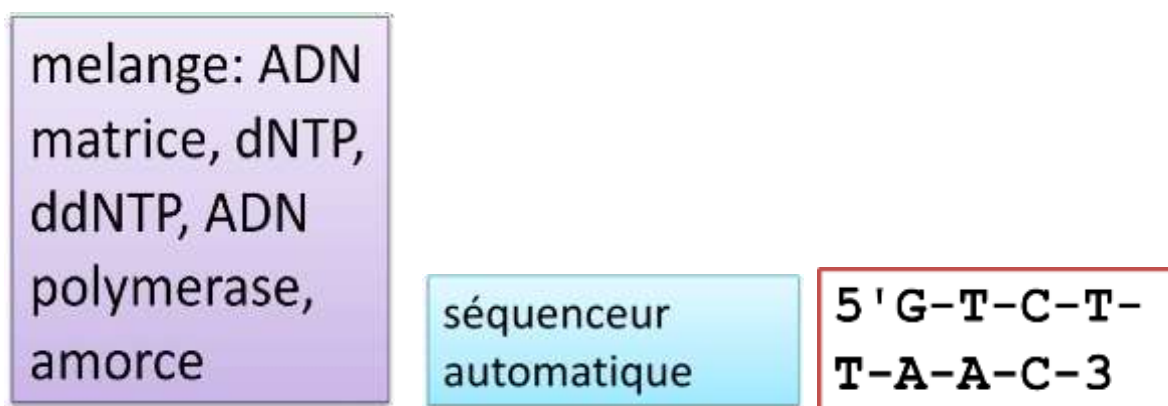
1- Synthèse du brin complémentaire d'un brin matrice à l'aide d'une polymérase à partir d'une amorce, en présence des 4 dNTPs et d'un ddNTP qui joue le rôle de terminateur de chaîne (absence de 3'-OH). Quatre réactions en parallèle sont nécessaires, chacune avec l'un des terminateurs présents en faible quantité.

2- Séparation par électrophorèse des fragments synthétisés, le nombre de ces fragments dépend du nombre de résidus nucléotidiques. Donc la position des terminateurs dans la séquence.

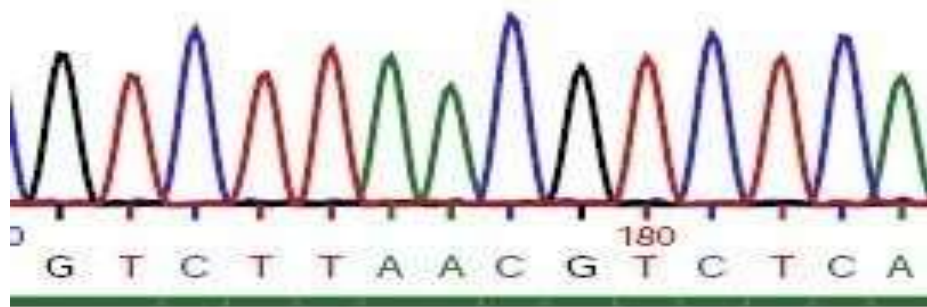
La détection des fragments ainsi synthétisés se fait en incorporant un traceur dans l'ADN synthétisé. Initialement ce traceur était radioactif ; aujourd'hui, on utilise des traceurs fluorescents, attachés soit à l'oligonucléotide, soit au didésoxyribonucléotide.

#### ▪ Automatisation du séquençage

Des séquenceurs automatiques sont capables de réaliser les réactions de séquence puis de les lire.



- Les didésoxynucléotides sont marqués avec des fluorescences différentes, cela permet de faire une seule réaction et une seule lecture. Le séquenceur détecte la fluorescence et repère la taille des fragments d'ADN.



- On obtient des courbes de fluorescences qui sont interprétées en termes de nucléotides.



## 1- Début de la transcription

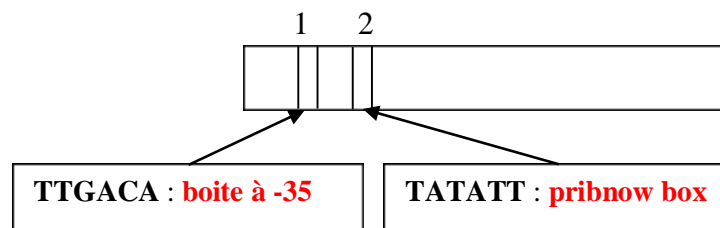
Par convention on appelle **+1** le premier nucléotide à partir lequel la transcription démarrera et **-1** le nucléotide qui précède

Le signal de début de la transcription est le **promoteur** : c'est une région de l'ADN à peu près (4 nucléotides) elle est située juste avant le début de la région dite **transcrite** (où démarrera la transcrite)

**Chez les procaryotes :**

*Exemple* : E-coli cette zone promotrice comporte des séquences conservées, se sont de courtes séquences jusqu'à (6N), il existe 2 : **-35** et **-10**

La première séquence est située à environ -35 paires de nucléotides et en amont de point de départ de la transcription



La 2<sup>ème</sup> séquence est située à environ -10 paires de nucléotides en amont de point de départ de la transcription, c'est une séquence riche en thymine

Chez l'E-coli l'ARN polymérase est composé des sous unités, l'ensemble de ces sous unités ( $\alpha \beta \beta' \sigma$ ) forment un **Holoenzyme**

Le sous unité  $\sigma$  permet à l'ARN polymérase de reconnaître les sites promoteurs (**-35, -10**), puis  $\sigma$  se dissocie après l'initiation de la transcription

Le noyau de sous unités  $\alpha \beta \beta'$  contiens le site catalytique responsable de l'élongation (synthèse de l'ARN)

**Chez les eucaryotes :**

Elle est beaucoup plus complexe dans la région promotrice se trouve essentiellement les séquences :

**a.** TATA box : chez presque tous les gènes et se localise dans (**-27 ; -22**), elle est riche en Adénine et Thymines et elle est équivalente de **Pribnowbox**

La TATA box participe au complexe d'initiation

ARN polymérase Boite box Facteur de Transcription (aide de l'ARN Polymérase)
--

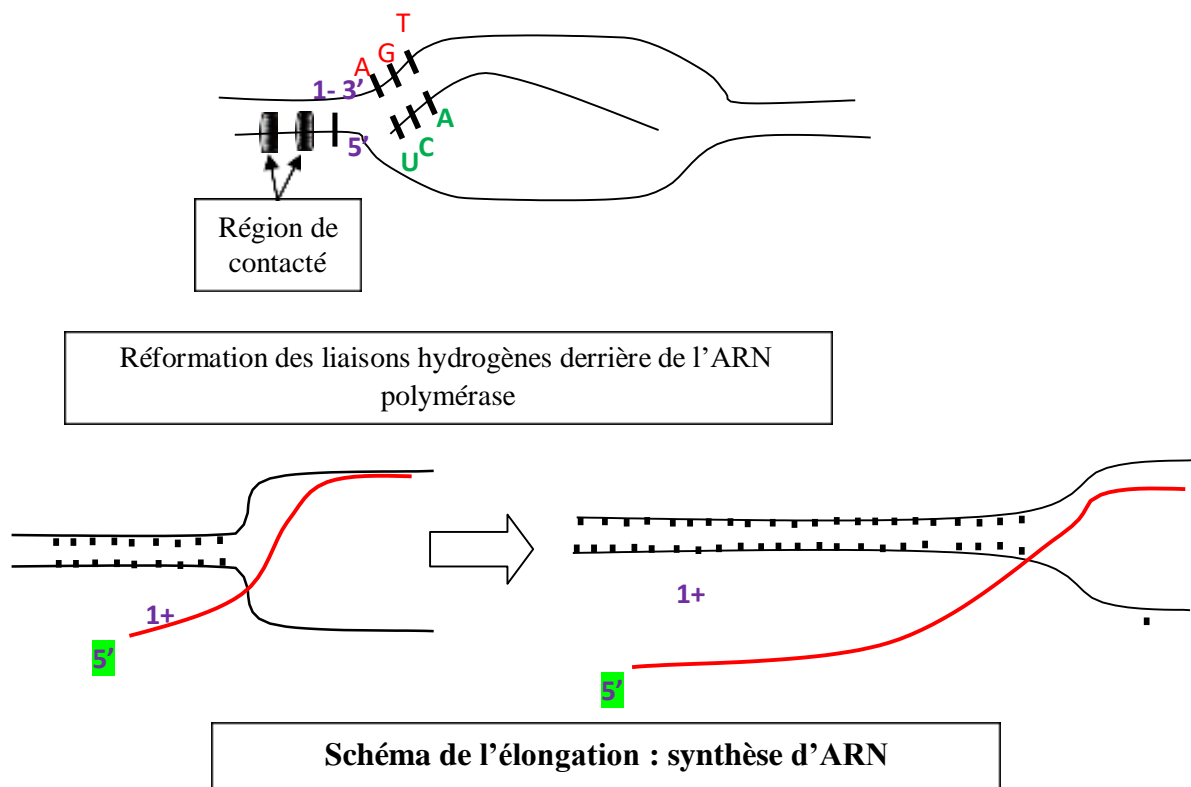
b. GC box (boite) : riche en Guanine et cytosine, elle se trouve dans la zone située entre (-140 et -40)

c. CCAAT box : elle représente dans de nombreux gènes, elle se situe dans la zone promotrice entre (-120 et -80)

## 2- Élongation

C'est la synthèse ou la transcription proprement dite se fait toujours dans le sens **5'.....3'**

Les 2 brins de l'ADN s'écartent par effet de rupture des liaisons hydrogènes puis l'ARN polymérase se lie et se déplace sur les brins d'ADN de façon transitoire et puis se détache, les liaisons hydrogènes entre les 2 brins d'ADN se reforment derrière de l'ARN polymérase.



Chez les eucaryotes il ya 3 types de l'ARN messages qui assure la transcription :

- ARN polymérase I : elle s'occupe la transcription des ARN r **18S, 28S et 5.8s** (s : **unité de sédimentation, S= Sueberg**)
- ARN polymérase II : pour la transcription des **ARN m et ARNsm**
- ARN polymérase III : pour la transcription les ARNt et ARN r (5s)

Plus ARN est long plus son degré de sédimentation sera élevé

Les ARN polymérase II et III se trouvent le nucléosome et l'ARN polymérase I se trouve dans le nucléole.

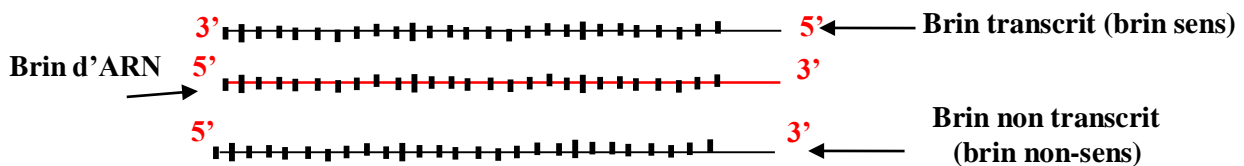
### 3- Fin de la transcription

**Chez les procaryotes :**

La zone de terminaison est riche en Adénines et thymines, c'est une région plus lâche qui laisse facilement au l'ARN polymérase de se détacher et de sortir, ARN transcrit va se terminer par une boucle de **l'épingle à cheveux**

**Chez les eucaryotes :**

Le signal de la fin de gène est une séquence de type AATATTA etc, qui lie sur le brin non transcrit = **appelé brin sens**

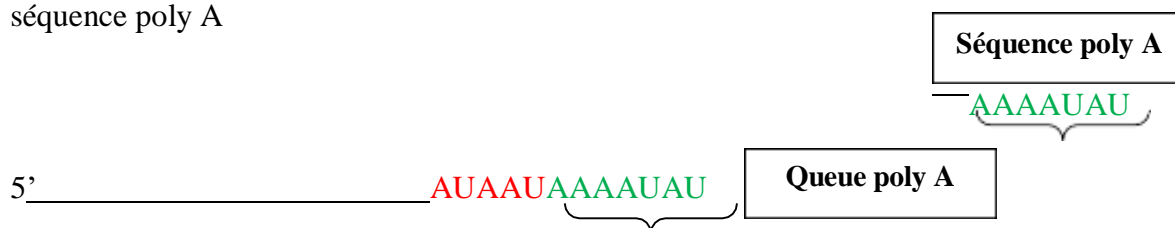


L'ARN et le brin d'ADN sens ont la même composition en base et même sens sauf au niveau de base de thymine (T) est remplacé par la base uracile (U) au niveau de l'ARN

AATATTA : cette séquence est appelée : séquence de **polyadénylation** ou encore séquence **poly A**

L'ARN qui a transcrit se terminer par AUAAAU= fin de transcription.

Un clivage (cassure) va se produire grâce à une enzyme nucléase puis l'adhésion d'une séquence poly A



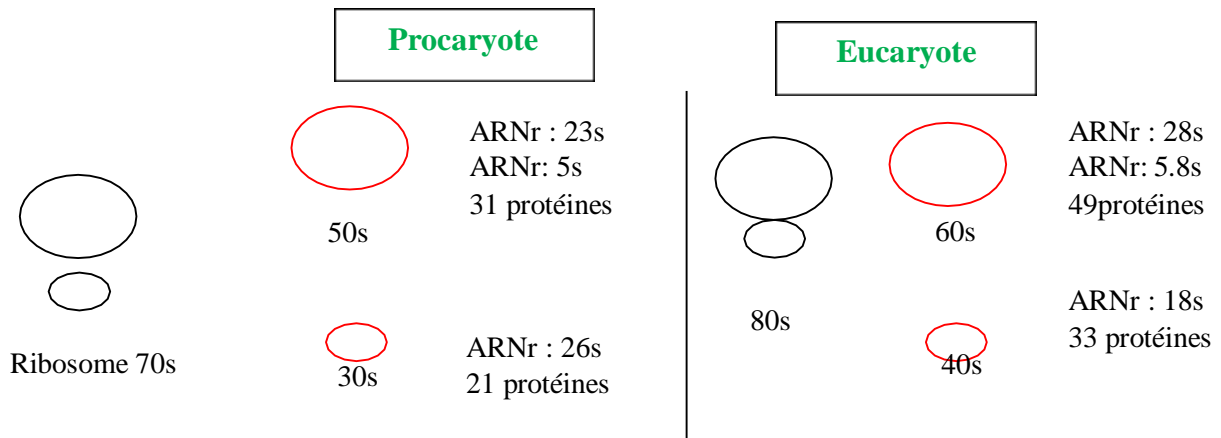
La queue poly A est synthétisé par une enzyme **polyA polymérase AAUAAA**

### 4- Les produits de la transcription

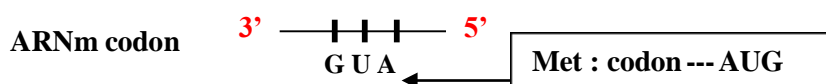
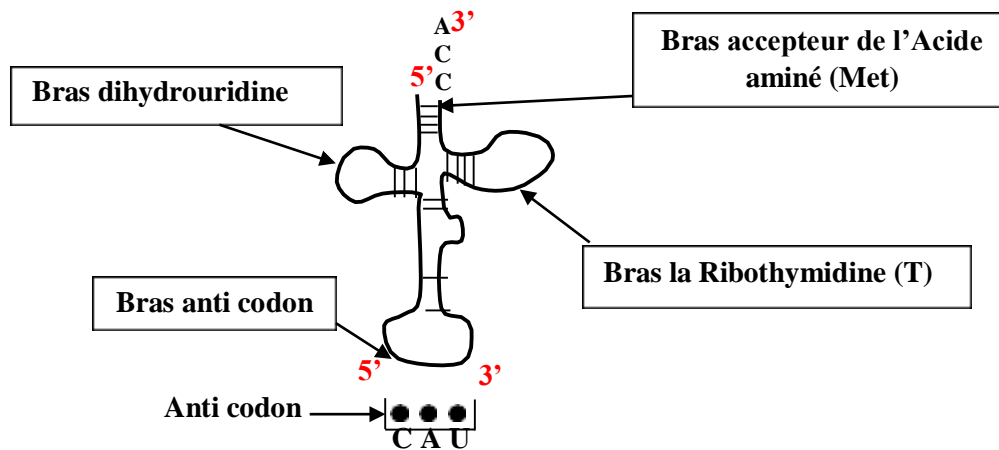
Sont donc les ARNm, ARNr et ARNt (aussi ARNsn et ARNsc).

a) **ARNr** : il rentre dans la structure biochimique de ribosome, ce dernier est constitué par 2 sous unités



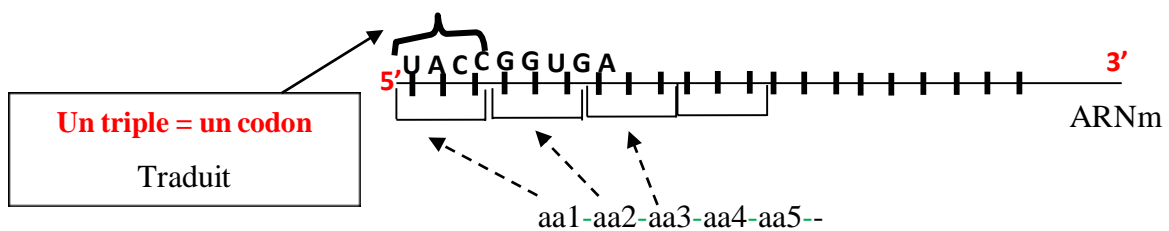


b) **ARNt** : sont de petite taille, cette molécule se réplique sur elle-même donnant une structure d'un **feuillelet Trèfle**.



Structure d'un ARNt en feuillelet de Trèfle

c) **ARNm** : leur taille est hétérogène (8s---30s), lors de la biosynthèse de protéines les ARNm



sont transformés en séquence d'acides aminés.

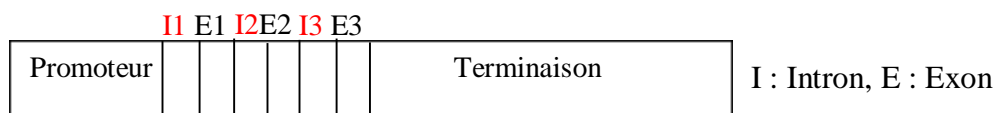
## 5- Modifications post transcriptionnelles

### Chez les procaryotes :

Pratiquement il n'y a pas des modifications, la traduction commence avant même la transcription se termine à l'extrémité 3' d'ARNm.

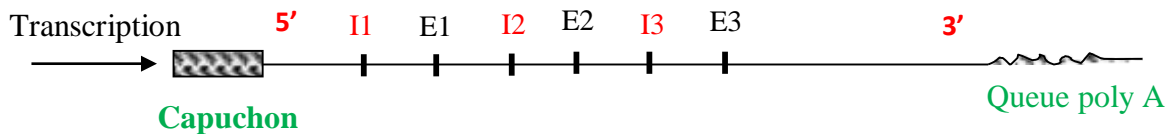
### Chez les Eucaryotes :

Déjà la structure des gènes est spécifiques, elle comprend des Exons (contient l'information génétique à traduire) et des Introns (qui s'intercale au milieu de 2 exons) ils sont transcrits, mais ne pas traduit.



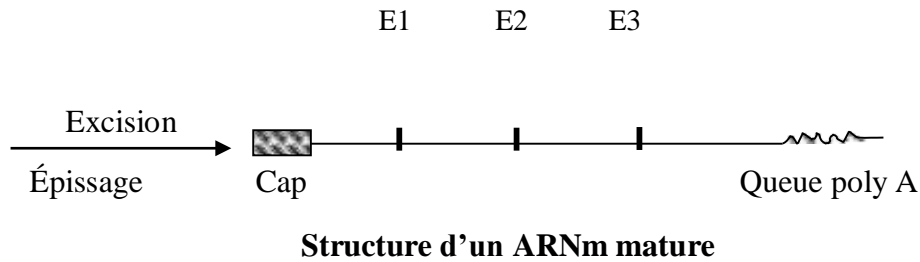
Structure d'un gène des eucaryotes

Des modifications font se produire pendant la transcription au niveau de l'ARNm seulement :

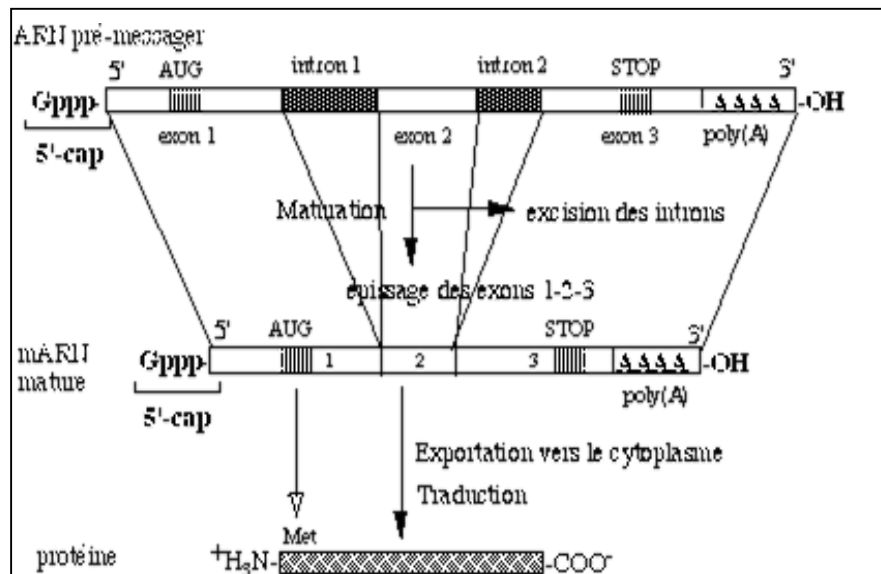


Structure d'un précurseur d'ARNm (pré-ARNm)

- L'addition d'une queue **poly A**.
- L'addition d'un capuchon =**Cap** est un GMP méthyle.
- Maturation de l'ARNm précurseur « pré-ARN » va subir des transformations.  
« processing » elles consistent à des excisions (coupé)/épissage (lié) c'est-à-dire couper les introns et lient les exons.



Les ARN messagers des eucaryotes nécessitent une maturation avant leur traduction.

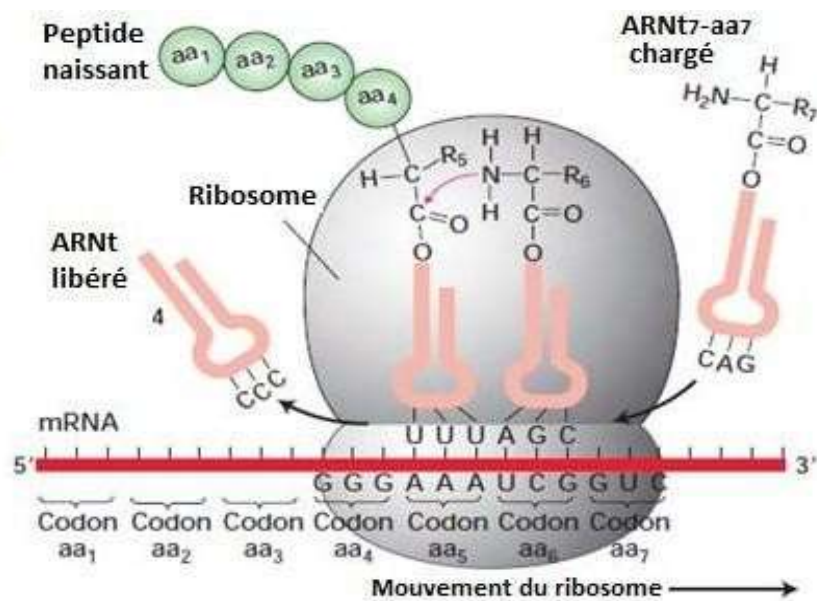


## VI. La traduction :

C'est le mécanisme par lequel l'ARNm est décodé, c'est la 2<sup>ème</sup> partie essentielle de la biosynthèse des protéines, elle s'effectue dans le cytoplasme au niveau du ribosome.

Le ribosome sera donc le siège de synthèse des protéiques cellulaires et les éléments nécessaires pour cette synthèse sont : les tARNs et le mARN.

Le ribosome contient les éléments enzymatiques nécessaires pour la constitution de la chaîne polypeptidique.



**Décodage de l'ARNm au cours de la traduction.**

## 1. Les éléments nécessaires à la traduction :

- Les ribosomes.
- Les acides aminés.
- La constitution de la liaison peptidique.
- L'ARN messenger.
- Les ARN de transfert.

### ▪ Les différentes étapes

La traduction se déroule en trois étapes successives (comparablement aux deux mécanismes précédents, réplication et transcription) : *l'initiation, l'élongation et la terminaison.*

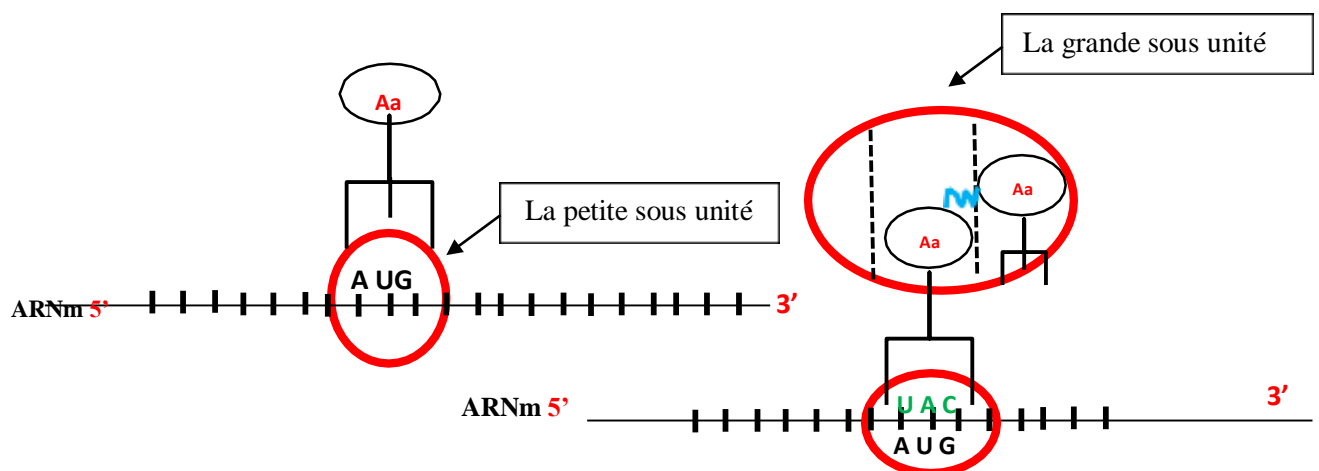
Il est important de connaître l'ordre des événements dans la traduction :

- La synthèse protéique se déroule de l'extrémité N-terminale à l'extrémité C-terminale de la protéine.
- Les ribosomes lisent l'ARNm dans le sens **5'--3'**.

### 1- Initiation :

De côté et de près de l'extrémité 5'phosphate de l'ARNm se trouve un codon signal qui indique le début de la traduction, c'est un codon initiateur presque toujours le même « 5'AUG 3' » se code pour l'acide aminé **Méthionine**

Les petites sous unités de ribosome se fixent en un point juste en amont de codon AUG, elles se déplacent en aval jusqu'à rencontre le 1<sup>er</sup> codon AUG, l'initiation chez eucaryote est similaire à celle des procaryotes, cependant la Méthionine (Met) chez les eucaryotes n'est pas **formylée** par contre elle est **formylée** chez les procaryotes est appelé **formyl-méthionine** (f Met) (Il s'agit d'une méthionine dont le groupe aminé est **formylé**).



À la phase d'initiation, la petite sous unité forme un complexe avec d'une part l'ARNm au niveau de codon AUG et d'autre part avec l'ARNt porteur acide aminé initial la méthionine

Lorsque la grande sous unité s'ajoute au complexe alors le ribosome est maintenant fonctionnel

Le ribosome possède 2 sites actifs pour lier les ARNt

**Site A :** (A=pour Acide aminé) c'est un site où s'installe l'ARNt porteur l'acide aminé

**Site P :** (P= pour peptide) c'est un site pour l'ARNt porteur de la chaîne polypeptidique en cours de sa l'élongation

**ARNt :** qui transporte la Méthionine initiale a une conformation qui lui permet d'être logé automatiquement et exceptionnellement dans le site peptidique.

## **2- Élongation :**

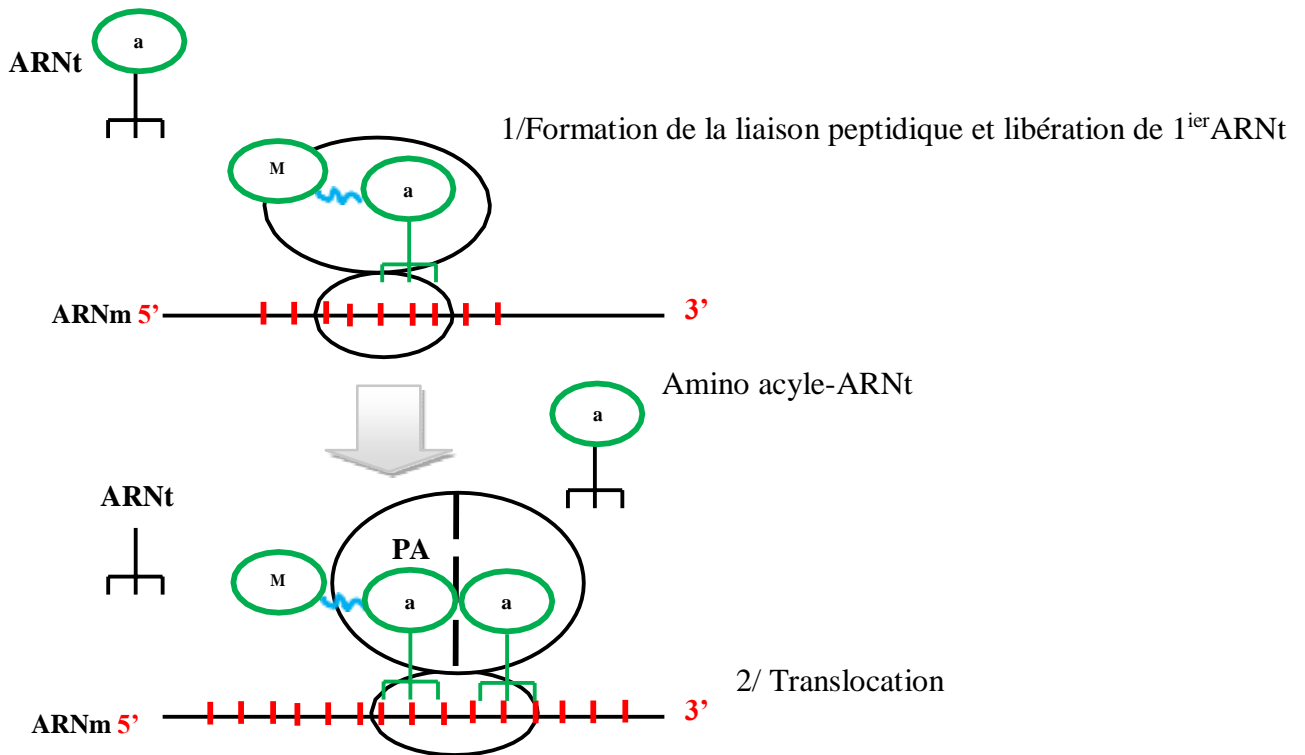
A la formation de la première liaison peptidique l'étape qui suit est l'élongation pour chaque acide aminé qui va s'accrocher et donc pour chaque liaison peptidique à fabriquer un même cycle est chaque fois d'écrire

### **Pour un cycle :**

1<sup>ère</sup> étape : **l'accrochage** : d'un amino acyle-ARNt dans le ribosome : c'est-à-dire il va s'accrocher dans le site A et puis le choix du 2<sup>ème</sup> acide aminé est déterminé par 2<sup>ème</sup> codon qui suit.

2<sup>ème</sup> étape : **la formation de liaison peptidique.**

3<sup>ème</sup> étape : **la translocation** : le ribosome il va avancer d'un codon (triplet 3 nucléotides) sur l'ARNm en toujours dans la direction (sens) 5' 3'.



### 3- Terminaison :

La fin de la traduction se produit lorsque le ribosome trouve (découvert) un **codon stop**, sont de nombre 3 : **UAA, UAG, UGA**, ils ne déterminent aucun acide aminé, il se produira une coupure entre le dernier ARNt et la chaîne peptidique.

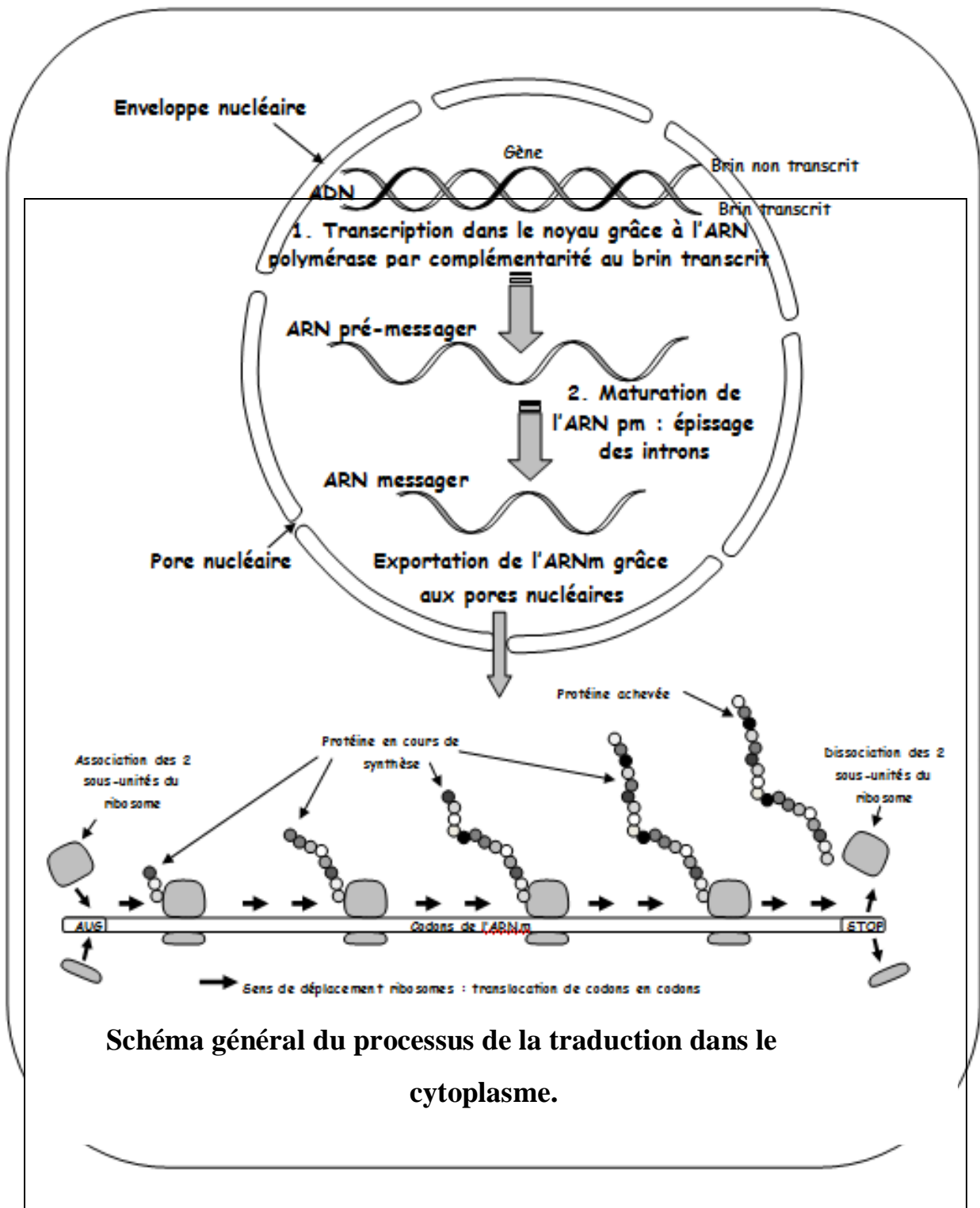


Schéma récapitulatif de l'expression de gène (transcription et traduction) dans une cellule eucaryote.





Les 2 premières bases d'un codon sont rigoureusement complémentaires d'anticodon alors que la 3<sup>ème</sup> base du codon peut être avoir une complémentarité moins rigoureuse.

Base présente en 5' de l'anticodon	Base présente en 3' de codon
G	C. U
C	G
A	U
U	U. G
I (Inosine télomère de Guanine)	U. C. G

### ▪ Les codons

Le tableau suivant donne la signification standard de chaque codon de trois bases nucléiques d'ARN messager. Les principaux codages alternatifs sont indiqués après une barre oblique :

		2 <sup>e</sup> base								
		U		C		A		G		
1 <sup>re</sup> base	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	<b>STO</b> <b>P</b>	UGA	<b>STO</b> <b>P</b>	A
		UUG	L Leu	UCG	S Ser	UAG	<b>STO</b> <b>P</b>	UGG	W Tr p	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	CCG	P Pro	CAG	Q Gln	CGG	R Arg	G
	A	AUU	I Ile	ACU	T Thr	AAU	N As n	AGU	S Ser	U
		AUC	I Ile	ACC	T Thr	AAC	N As n	AGC	S Ser	C
		AUA	I Ile	ACA	T Thr	AAA	K Lys	AGA	R Arg	A
		AUG	M Met et <b>STA</b> <b>RT</b>	ACG	T Thr	AAG	K Lys	AGG	R Arg	G
	G	GUU	V Val	GCU	A Ala	GAU	D As p	GGU	G Gly	U
		GUC	V Val	GCC	A Ala	GAC	D As p	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

## VII. La mutagénèse :

Une mutation est une modification de l'information génétique contenue dans l'ADN. Elle affecte la séquence par le remplacement d'un ou plusieurs nucléotides, l'insertion ou la délétion de quelques nucléotides. Elle peut être due à l'instabilité du génome, à des erreurs de copie ou de réparation lors de la multiplication des cellules ou de la reproduction sexuée, à des coupures de l'ADN, à des conditions environnementales ou à l'action ciblée du sélectionneur.

Les mutations sont la source de la variabilité dans une même espèce.

### Les mutations spontanées

Les mutations spontanées sont rares ( $10^{-6}$ ,  $10^{-7}$  par gène et par génération) mais néanmoins significatives. La plupart sont neutres (n'influencent pas la valeur sélective), les moins favorables sont éliminées par la sélection naturelle. Les agriculteurs et les sélectionneurs qui ont choisi depuis des siècles de façon plus ou moins empirique les plantes les plus intéressantes ont souvent choisi en fait des plantes mutantes.

### Les mutations induites

Elles peuvent être provoquées par des traitements physiques (UV, rayons X...) ou chimiques (divers agents alkylants...) qui augmentent significativement le taux de mutation (plus de 500 fois, par exemple, chez *Arabidopsis* par l'éthyl méthane sulfonate – EMS). La probabilité de produire ou modifier un caractère intéressant pour le sélectionneur est alors ainsi fortement augmentée.

La mutagénèse est le processus d'apparition d'une mutation, il peut être naturel ou artificiel (par exposition de l'ADN à un « agent mutagène »)

La **mutagénèse dirigée ou aléatoire** est aussi une approche utilisée par le génie génétique et la biologie pour comprendre la fonction des gènes ; elle consiste en l'introduction volontaire de mutations par l'action d'agents mutagènes chimiques ou physiques dans une séquence ADN afin de déduire des informations sur le rôle des gènes, à partir de l'analyse des effets de ces mutations.

### 1- Mutagénèse dirigée :

L'observation des mutants est une des méthodologies de bases de la génétique. Cependant, la fréquence d'apparition de mutants dans une population d'organismes est relativement basse, la probabilité d'obtenir de manière naturelle un mutant pour une fonction ou un processus biologique particulier est donc

particulièrement faible. Pour augmenter cette probabilité, les organismes sont traités par un agent mutagène, puis les mutants obtenus sont sélectionnés en fonction d'un crible.

Deux méthodes sont possibles, la mutagenèse aléatoire et la mutagenèse dirigée. La première consiste à utiliser un agent mutagène, qui induira de manière aléatoire des mutations dans le génome de l'organisme étudié. La seconde utilisera des méthodes de biologie moléculaire pour induire une mutation précise dans le gène ciblé.

Les mutagènes induisent l'apparition de mutations par trois mécanismes différents au moins. Ils peuvent remplacer, modifier ou endommager des parties de la séquence d'ADN.

### ▪ Remplacement de bases

- incorporation d'analogues de bases:

Certains composés chimiques ressemblent suffisamment aux bases azotées pour être incorporés à leurs places dans l'ADN. Ce sont des analogues de bases qui pourront se lier à n'importe quelle base du brin complémentaire contrairement à la règle d'appariement des bases dans l'ADN. Ils peuvent donc provoquer l'insertion face à eux de nucléotides incorrects. Par exemple, le 5-bromouracile (ou 5-BU) est un analogue de la thymine et le 2-aminopurine (ou 2-AP) est un analogue de l'adénine.

### ▪ Modification de la séquence d'ADN

- mésappariement spécifique:

Certains mutagènes modifient une base pour qu'elle ne suive plus le principe d'appariement (A-T, G-C). Parmi ces mutagènes, on trouve des agents alkylants qui ajoutent des groupements alkyles sur les bases. D'autres agents chimiques peuvent provoquer la désamination (enlève un  $-NH_2$ ) des bases.

- **agents intercalants** : Ce sont des molécules capables de se glisser entre les bases azotées. Elles peuvent provoquer une inhibition du procédé de réplication de l'ADN. On trouve par exemple la proflavine et l'acridine orange.
- **PCR** (Réaction en chaîne par polymérase) : On utilise une amorce mutée qui permet de provoquer, à une position précise, la mutation voulue. Le gène ainsi modifié sera ensuite cloné un grand nombre de fois grâce à la réplication de l'ADN.

- **par « cassettes »** : On synthétise chimiquement des oligonucléotides complémentaires de façon à obtenir une petite séquence ADN double-brin appelée cassette, qui contient la mutation souhaitée. Cette cassette est ensuite insérée dans le gène cible<sup>6</sup>.

- **Endommagement**

Il existe un grand nombre de mutagènes qui endommagent une ou plusieurs bases, il n'y a alors plus d'appariement spécifique et cela entraîne un blocage de la réplication.

- **lumière ultraviolette** : Elle implique différentes lésions dans l'ADN. Certaines relient deux pyrimidines adjacentes sur le même brin.
- **radiations ionisantes** : Elles forment des molécules excitées et ionisées qui peuvent endommager l'ADN, soit indirectement par l'intermédiaire d'espèces d'oxygène actif (OH, O<sub>2</sub><sup>-</sup>, H<sub>2</sub>O<sub>2</sub>), soit directement en provoquant la rupture de la liaison entre le sucre et la base (liaison N-glycosidique).
- **agent chimique mutagène** : Par exemple, l'EMS (Méthanesulfonate d'éthyle) est souvent utilisé pour modifier la séquence génétique de la plante à muter.

## 2- Mutagenèse aléatoire :

La **mutagenèse aléatoire** consiste à induire des mutations n'importe où dans l'ADN, afin d'obtenir un grand nombre de mutants, pouvant être par la suite sélectionnés. Ces mutations peuvent être obtenues en modifiant les propriétés du milieu (concentration en ions, enzymes...), par l'action de composés mutagènes chimiques ou encore par irradiation.

La mutagenèse aléatoire sert principalement à localiser des gènes codant pour un phénotype particulier. En effet, après avoir obtenu des mutations dans diverses régions du génome, les mutants sont triés à l'aide d'un crible et seuls seront gardés ceux qui présenteront le caractère que l'on veut observer. Une fois le(s) gène(s) identifié(s), on pourra utiliser la mutagenèse dirigée pour créer des mutations plus ciblées et étudier leurs impacts sur le phénotype.

Lorsqu'on sait peu de choses sur la fonction d'une protéine, la mutagenèse aléatoire est souvent utilisée comme première étape d'investigation. L'analyse des mutants obtenus donne plus d'informations sur la fonction et les gènes associés à cette protéine. L'intérêt de cette stratégie est qu'elle permet de réduire le champ de recherches à un fragment beaucoup plus petit.

## **VIII. Application au clonage moléculaire (génie génétique) :**

Le génie génétique est un ensemble de techniques de biologie moléculaire permettant d'isoler des gènes spécifiques, de les reconstruire puis de les réinsérer dans des cellules ou des organismes. Ces techniques ont fourni à la médecine et à l'industrie un moyen efficace de produire en grandes quantités des protéines spécifiques, qui, auparavant, n'étaient disponibles (si elles l'étaient) qu'en quantité extrêmement faibles. Ces techniques ont permis également d'étudier la régulation de leur expression et ainsi de mieux comprendre le développement de maladies génétiques.

Les différentes étapes passent par : la construction d'une banque d'ADN, le criblage de la banque et l'expression du gène.

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

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

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

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

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

### **1. Les enzymes de restriction :**

Découvertes à partir de 1973. Les enzymes de restriction sont capables de reconnaître spécifiquement une courte séquence, de 4 à 10pb, et de cliver l'ADN au site reconnu. Ils permettent de fragmenter l'ADN en segments de taille réduite, ou de le couper à tel ou tel site désiré. Certains enzymes coupent le site en son milieu et produisent deux fragments dont les extrémités sont franches. Cependant, la

plupart réalisent une coupure dissymétrique : on parle dans ce cas d'extrémités cohésives (chaque fragment possède une chaîne qui dépasse l'autre de quelques bases). Plusieurs centaines de ces enzymes ont été caractérisés : ils reconnaissent une grande variété de sites de coupure.

Les enzymes de restriction sont utilisés pour établir une carte de restriction de toute molécule d'ADN que l'on souhaite caractériser. Cela consiste à déterminer l'ordre des sites de restriction le long de cette molécule, qui vont produire, après "digestion enzymatique" de cette molécule, des fragments de tailles différentes dont la taille pourra être définie par électrophorèse.

Les enzymes de restriction appartiennent à la classe des endonucléases, c'est-à-dire des enzymes capables de cliver les liaisons phosphodiester entre deux nucléotides à l'intérieur d'un acide nucléique. Les endonucléases se différencient des exonucléases qui dégradent la molécule d'ADN à partir de l'une de ses extrémités (3' ou 5').

#### ▪ Séquences d'ADN reconnues par les enzymes de restriction

Les séquences de nucléotides reconnues par les enzymes de restriction sont habituellement des séquences dites palindromiques. Les séquences palindromiques sont des séquences où la succession des nucléotides lue dans le sens 5' à 3' (gauche-droite) pour le premier brin est identique à la séquence lue dans le sens droite-gauche pour le second brin (sens 5' à 3'). Ces séquences palindromiques sont le plus souvent constituées de 4, 5 ou 6 paires de bases. Il faut remarquer que dans les ADN, on rencontre statistiquement des séquences reconnues par des enzymes de restriction. Ainsi, la séquence **GATC** reconnue par l'enzyme **Mbo I** est présente avec une fréquence statistique de 1 / 256 paires de bases (1/4<sup>4</sup>).

En effet, la fréquence de coupure d'un ADN par une enzyme de restriction donnée peut être approchée statistiquement en considérant le nombre de nucléotides de la séquence spécifique reconnue par l'enzyme. Ainsi, par exemple, dans la séquence de six nucléotides: **GGATCC** reconnue par l'enzyme **Bam HI**, on aura donc une fréquence de coupure statistique de 1 / 4<sup>6</sup>, soit 1 coupure tous les 4096 nucléotides.

## ▪ Origine des enzymes de restriction

Les enzymes de restriction sont extraites de micro-organismes, le plus souvent des bactéries. Les bactéries peuvent être parasitées par des virus à ADN. Les bactéries fabriquent des enzymes de restriction qui sont capables de cliver les ADN étrangers. Pour éviter une auto- destruction de leur propre ADN, elles se protègent contre leurs propres enzymes de restriction par une modification des sites de restriction correspondants.

### 1. Nomenclature des enzymes de restriction

Les enzymes de restriction présentent une nomenclature bien précise. Leur nom comporte plusieurs lettres (3 ou 4). La première lettre de dénomination de l'enzyme est écrite en majuscule, elle correspond au **genre de la bactérie** d'où a été extraite l'enzyme. La seconde lettre et la troisième lettre (en minuscules) correspondent à **l'espèce de la bactérie** d'où l'enzyme est extraite. On peut avoir une quatrième lettre écrite en majuscule correspondant à **la souche bactérienne**. Enfin pour terminer, un chiffre romain indique **l'ordre** de caractérisation de ces enzymes.

Exemples:

**Eco RI** Extraite d'*Escherichia coli* RYB site reconnu: **G / AATTC** **Sma I** Extraite de *Serratia marcescens* site reconnu: **CCC / GGG** **Pst I** Extraite de *Providencia stuartii* site reconnu:

**CTGCA / G**

### Notion d'isoschizomères

Des enzymes de restriction différentes peuvent reconnaître des mêmes sites spécifiques, on les appelle isoschizomères. Les isoschizomères fournissent souvent après clivage enzymatique des fragments dont les extrémités sont différentes.

**Exemple :**

Soit la séquence suivante: **GGTACC**, cette séquence est coupée par l'enzyme **Kpn I** et l'enzyme **Acc65 I**:



### **Kpn I:**

5'-G-G-T-A-C/C-3'

3'-C/C-A-T-G-G-5'

### **Acc65 I:**

5'-G/G-T-A-C-C-3'

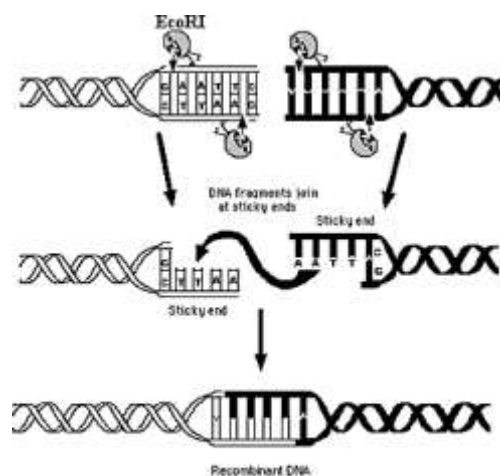
3'-C-C-A-T-G/G-5'

Ces enzymes sont des isoschizomères, elles reconnaissent en effet la même séquence nucléotidique GGTACC. On voit tout de suite que les extrémités des fragments obtenus diffèrent.

## **2. Types de coupures réalisées par les enzymes de restriction**

Les enzymes de restriction peuvent donner deux types de coupures: la coupure à bouts francs et la coupure à bouts collants.

- **La coupure à bouts francs** aboutit à une coupure au milieu de la séquence palindromique.
- **La coupure à bouts collants** (ou à extrémités adhésives) correspond à une coupure qui se fait de part et d'autre du centre de symétrie.



**Restriction Enzyme  
Action of EcoRI**

**Coupure à bouts francs et coupure à bouts collants réalisées par les enzymes de restriction.**

Les sites de restriction sont repérés dans l'ADN par l'enzyme de restriction qui coupe l'ADN en principe autant que fois qu'il y a de sites de restriction. Ceci est valable pour une enzyme de restriction donnée, pour une autre enzyme, la coupure se fera en une position différente sur l'ADN. On voit tout de suite les possibilités considérables de ce type d'outils enzymatiques. L'ADN est découpé en fragments variables et ceci aussi bien l'ADN circulaire des bactéries ou des plasmides que l'ADN linéaire.

### 3. Les types d'enzymes de restriction

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

**Les enzymes de type I** reconnaissent des séquences sur l'ADN sans aucune symétrie. Ces enzymes coupent non pas au niveau de la séquence de reconnaissance mais 1000 à 5000 paires de bases plus loin.

**Les enzymes de type III** présentent également un site de reconnaissance mais coupent une vingtaine de nucléotides plus loin.

#### Exemples d'enzymes de restriction de type II

Enzyme	Source	Séquence reconnue	Coupure	Extrémités (après coupure)
<u>Eco RI</u>	<u>Escherichia coli</u>	5'GAATTC 3'CTTAAG	5'---G AATTC---3' 3'---CTTAA G---5'	Cohésives
BamHI	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'	Cohésives
HindIII	<u>Haemophilus influenzae</u>	5'AAGCTT 3'TTCGAA	5'---A AGCTT---3' 3'---TTCGA A---5'	Cohésives
MstII	<i>Microcoleus species</i>	5'CCTNAGG 3'GGAMTCC	5'---CC TNAGG---3' 3'---GGAMT CC---5'	Cohésives
TaqI	<u>Thermus aquaticus</u>	5'TCGA 3'AGCT	5'---T CGA---3' 3'---AGC T---5'	Cohésives

NotI	<i>Nocardia otitidis</i>	5'GCGGCCGC 3'CGCCGGCG	5'—GC GGCCGC--3' 3'--CGCCGG CG--5'	Cohésives
HinI	<i>Haemophilus influenzae</i>	5'GANTC 3'CTNAG		
AluI	<i>Arthrobacter luteus</i>	5'AGCT 3'TCGA	5'---AG CT---3' 3'---TC GA---5'	Franches
BgIII	<i>Bacillus globigii</i>	5'AGATCT 3'TCTAGA	5'---A GATCT---3' 3'---TCTAG A---5'	Cohésives
HaeIII	<i>Haemophilus aegyptius</i>	5'GGCC 3'CCGG	5'---GG CC---3' 3'---CC GG---5'	Franches
HhaI	<i>Haemophilus haemolyticus</i>	5'GCGC 3'CGCG	5'---GCG C---3' 3'---C GCG---5'	Cohésives
PstI	<i>Providencia stuartii</i>	5'CTGCAG 3'GACGTC	5'---CTGCA G---3' 3'---G ACGTC---5'	Cohésives
SmaI	<i>Serratia marcescens</i>	5'CCCGGG 3'GGGCC	5'---CCC GGG---3' 3'---GGG CCC---5'	Franches

Le N et le M dans la séquence de MstII peuvent être remplacés par une des 4 bases et son complément. La séquence de restriction du MstII est un exemple de palindrome imparfait puisqu'il comporte un nombre impair de paires de bases.

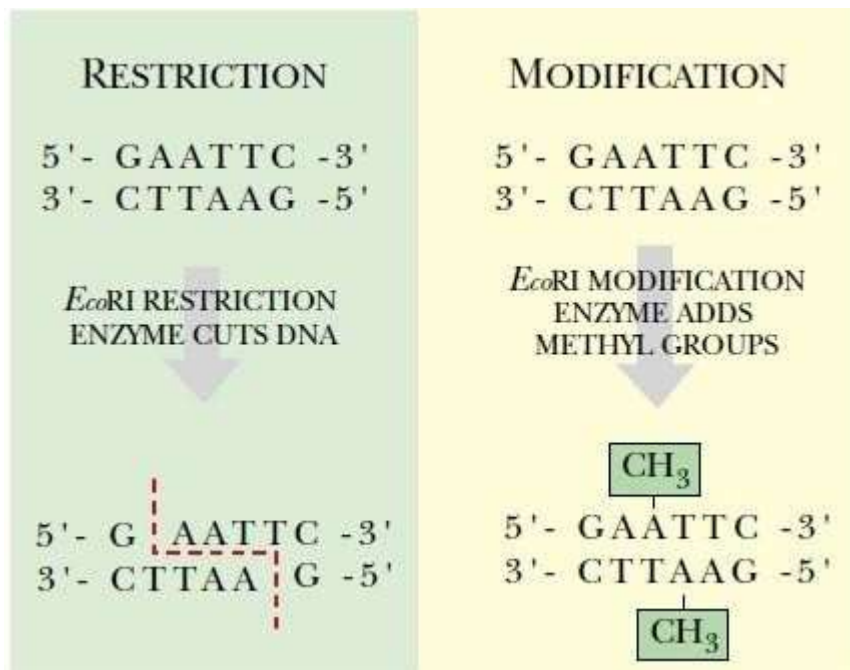
#### 4. Méthylation des sites de restriction et inactivation des enzymes de restriction

Dès 1953, on avait observé que lorsque des molécules d'ADN provenant d'une souche d'E.coli sont introduites dans une autre souche, elles ne sont que très rarement exprimées. En fait, l'ADN étranger est la plupart du temps rapidement coupé en petits fragments par une enzyme dite de restriction. Pour qu'il puisse se répliquer dans une nouvelle souche bactérienne, l'ADN introduit doit être protégé par une enzyme de modification qui ajoute des groupements méthyles à l'ADN.

En accord avec ces expériences sur E. coli, Hamilton Smith à l'Université Johns Hopkins s'aperçoit fortuitement que la bactérie *Haemophilus influenzae* dégrade rapidement un ADN étranger alors que son propre ADN reste intact. Cette découverte l'a amené à purifier HindII,

la première enzyme de restriction coupant à un site spécifique et responsable de la dégradation de l'ADN étranger.

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



### Systèmes de restriction et de méthylation.

#### ▪ Utilisations des enzymes de restriction

Les utilisations des enzymes de restriction sont très nombreuses en biologie moléculaire.

- Par exemple, elles permettent de fractionner l'ADN en multiples fragments susceptibles d'être séparés par les techniques d'électrophorèse.
- Les enzymes de restriction peuvent être utilisées pour préparer un fragment d'ADN d'un gène donné (insert) à être inséré dans un vecteur comme un plasmide.

- Les enzymes de restriction sont utilisées couramment pour rechercher des mutations dans le génome.
- Les enzymes de restriction sont utilisées pour rechercher dans l'ADN des cellules eucaryotes les méthylations de bases. Ces méthylations ont une signification complètement différente des méthylations de bases observées chez les procaryotes. Elles sont en relation directe avec des modifications de l'expression des gènes des eucaryotes. La méthylation provoque le verrouillage de l'expression de tel ou tel gène dans un tissu. Les méthylations dans le génome des eucaryotes concerne les cytosines impliquées dans les doublets dinucléotidiques CG. La recherche de ces doublets et de la présence ou non d'une méthylation sur les cytosines est réalisée à l'aide de deux enzymes de restriction, une enzyme insensible à la méthylation des cytosines et une autre enzyme sensible à la méthylation des cytosines.

## **2. Les éléments nécessaires au clonage :**

### **▪ Les vecteurs de clonage**

On appelle vecteur l'ADN dans lequel on insère le fragment d'ADN à étudier. L'ADN inséré est appelé insert ou ADN étranger ou ADN exogène. Cette séquence nucléotidique est capable de s'auto-répliquer. Les vecteurs sont donc des petits ADN dans lesquels on insère un fragment d'ADN que l'on veut étudier. Ces petits ADN sont généralement des plasmides ou des bactériophages. Il est nécessaire d'introduire ces bactériophages ou plasmides dans les bactéries pour réaliser une multiplication de ceux-ci.

#### **1. Les plasmides**

Les plasmides sont des petits fragments d'ADN circulaire présents dans la cellule bactérienne et indépendants du génome bactérien.

##### **- Propriétés générales des plasmides**

Les plasmides présentent les caractéristiques suivantes:

Leur ADN est bicaténaire, circulaire avec un nombre de nucléotides inférieur à 10 kb (1 kilobase = kb = 1000 nucléotides).

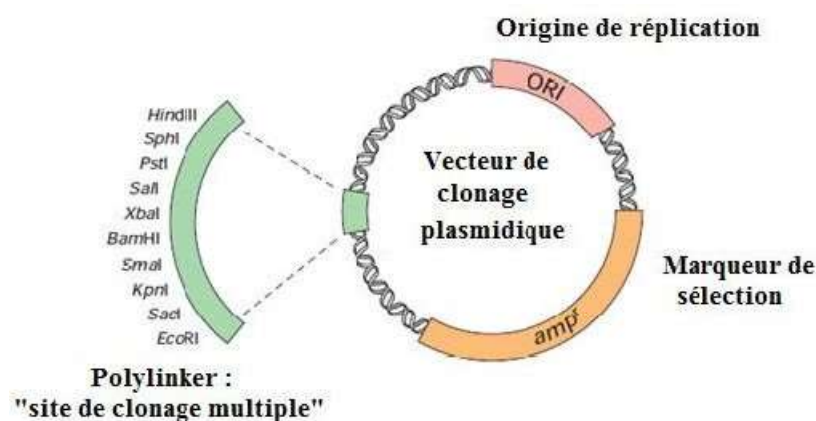
Le nombre de plasmides dans une cellule bactérienne peut être considérable (plusieurs centaines). Les plasmides portent normalement des gènes qui leur confèrent un avantage sélectif, par exemple une résistance à un antibiotique, résistance aux métaux lourds, dégradation de composés aromatiques. Leur réplication est indépendante de celle du génome bactérien.

Les plasmides naturels possèdent dans leur génomes un locus nommé "**par**" qui fonctionne à la façon de centromère des chromosomes eucaryotes, ce locus contrôle la répartition précise des plasmides parmi les cellules filles lors de la division cellulaire. Pour réduire la taille des plasmides qui servent de vecteurs de clonage, le locus "par" est très souvent éliminé. C'est notamment le cas du plasmide pBR322 et leurs dérivés.

Les plasmides sont de bons vecteurs de clonage chez les bactéries parce qu'ils se multiplient en nombre de copies important et qu'ils sont aisément purifiables. Les marqueurs de sélection (Ex. gènes de résistance aux antibiotiques) permettent l'identification des bactéries recombinantes (transformées) qu'ils les portent.

#### - Préparation des plasmides

Les bactéries contenant les plasmides sont mises en culture dans un grand volume (jusqu'à 2 litres de milieu de culture). Quand la concentration bactérienne atteint une valeur suffisante, on traite par un antibiotique (comme le chloramphénicol) qui bloque la réplication du chromosome bactérien mais pas celle du plasmide. On récupère les bactéries par centrifugation. Puis, on perméabilise la paroi bactérienne par un traitement doux permettant d'obtenir un passage en solution des plasmides qui sont plus légers que l'ADN bactérien. La purification des plasmides peut être ensuite réalisée par chromatographie liquide à haute performance ou ultra-centrifugation en gradient de densité.

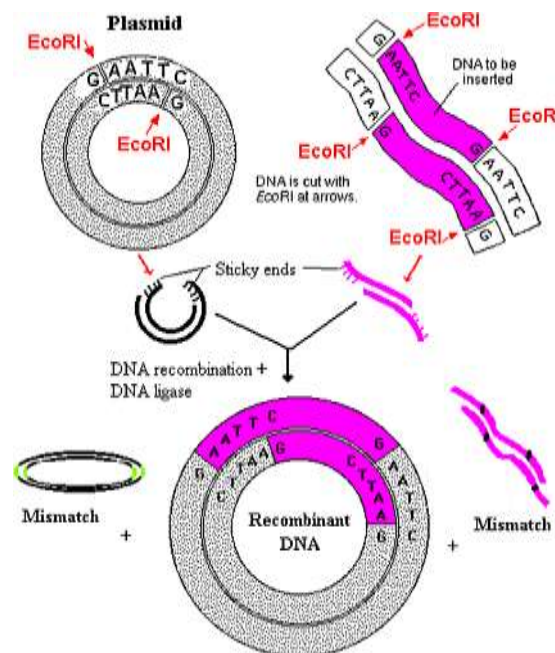


**Composants basiques d'un vecteur plasmidique chez *E. coli***

## - Préparation des plasmides pour la recombinaison, constitution de l'hybride et transformation bactérienne

L'insertion d'une séquence d'ADN bicaténaire dans un plasmide nécessite un traitement préalable par une enzyme de restriction. On choisit une enzyme qui a un site unique dans la séquence plasmidique. Après action de l'enzyme de restriction, le plasmide circulaire est linéarisé. Pour maintenir la linéarisation, il est indispensable de traiter le plasmide linéaire par une phosphatase alcaline qui enlève les groupements phosphates en 5'. Le plasmide linéarisé et traité par une phosphatase alcaline est ajouté au fragment d'ADN à insérer en présence d'une ligase. L'ensemble de ces étapes produit **un plasmide recombinant**.

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



### Insertion d'ADN dans le plasmide

## -Sélection des bactéries transformées par des plasmides

Le taux de transformation des bactéries est généralement très faible. Il est fondamental de disposer de méthodes permettant d'obtenir des bactéries transformées par les plasmides et seulement celles-ci. Pour cela, l'artifice consiste à utiliser *une résistance à un antibiotique* (par exemple: l'ampicilline).

La transformation d'une bactérie sensible à un antibiotique par des plasmides portant un gène de résistance au même antibiotique aboutit à l'apparition d'une résistance uniquement pour les bactéries ayant incorporé les plasmides. Cette technique permet de séparer les bactéries comportant des plasmides et les bactéries sans plasmides car ces dernières sont tuées. Cependant, *elle ne permet pas de distinguer les bactéries avec plasmides intacts des bactéries avec plasmides recombinants*. On utilise pour cela une résistance à un second antibiotique, par exemple: une tétracycline. L'insertion du fragment d'ADN dans le plasmide doit inactiver le gène de résistance au deuxième antibiotique. Les bactéries transformées par les plasmides recombinants sont donc résistantes au premier antibiotique (ampicilline) et sensibles au deuxième antibiotique (tétracycline). Les bactéries non transformées sont sensibles aux deux antibiotiques.

### **-Un exemple de plasmide: le puc19**

De nouvelle génération de plasmides plus puissants sans cesse croissantes ont été développés depuis pBR322 et ces dérivés. C'est le cas de la famille pUC appelés (les vecteurs de clonage de seconde génération).

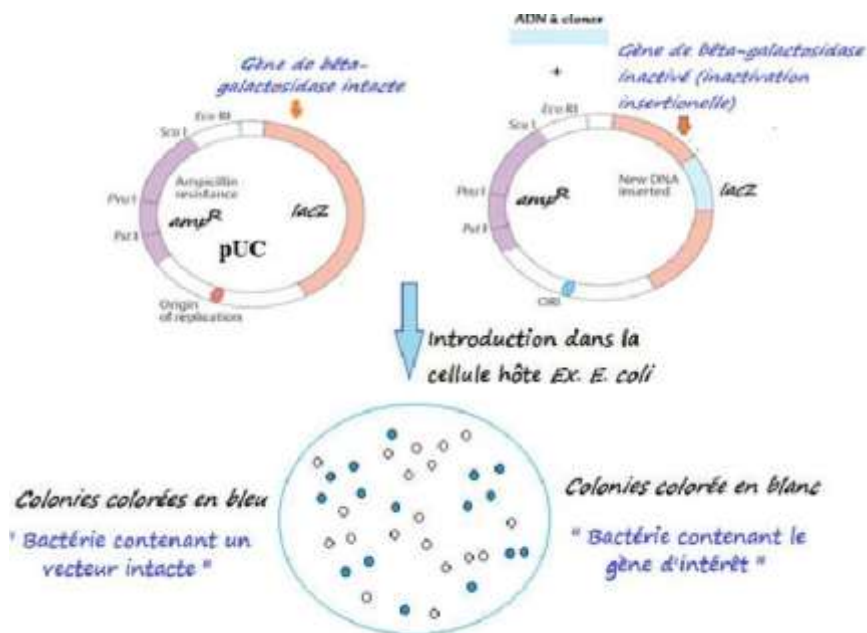
Un plasmide pUC est un plasmide pBR dans lequel on a remplacé le gène de résistance à la tétracycline (qui sert à repérer les plasmides recombinés) par un gène bactérien lacZ qui code la (galactosidase (revoir l'opéron lactose)). Cette enzyme a pour propriété d'hydrolyser le lactose en glucose et galactose. Un produit non toxique pour la cellule bactérienne est également substrat de la galactosidase : le Xgal (5-bromo-4chloro-3-indonyl- $\beta$ -D- galactoside), l'hydrolyse de ce dernier en dibromo-5,5-dichloro-4, 4-indigo (de couleur bleu), indique la production de  $\beta$ -galactosidase.

Un plasmide classique est le pUC19. La séquence complète est connue (2686 nucléotides). Il comporte les particularités suivantes:

- *Un gène de résistance à l'ampicilline (antibiotique)*. La présence de ce gène permettra à la bactérie porteuse de ce plasmide de ne pas être sensible à l'effet de cet antibiotique.
  
- *Le gène lac Z qui code pour la b-galactosidase dans l'opéron lactose*.
  
- *Enfin, une région avec des sites multiples et uniques pour des enzymes de restriction connues*. Cette région est appelée " *polylinker* ". Son rôle est de permettre l'insertion du fragment d'ADN étranger.







**Sélection des bactéries transformées portant un plasmide pUC recombiné.**

**-Avantages et désavantages des plasmides**

L'inconvénient qui limite l'utilisation des plasmides est l'impossibilité d'insérer des larges fragments d'ADN (La capacité maximum de ce plasmide est de 10kb d'ADN étranger). D'autres choix de vecteurs sont utilisés pour dépasser ce problème de la taille d'insert.

## 2. Les phages

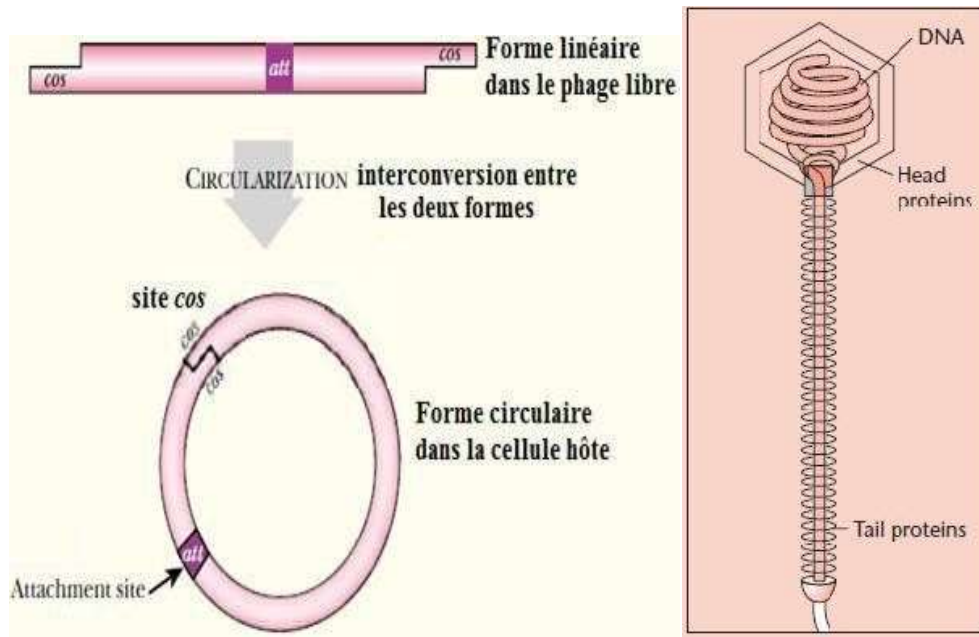
Les phages sont des virus qui infectent les bactéries. Deux phages sont très utilisés comme vecteurs: le phage lambda et le phage M13 (mais maintenant les phages dérivés de ces deux phages). Les séquences de ces phages utilisés comme vecteurs sont connues (40 à 50 kb d'ADN double brin). Les extrémités de cet ADN sont simples brins sur une longueur réduite, et complémentaires l'une de l'autre et surtout formant *des extrémités cohésives*. Ces séquences sont appelées *séquences cos* (*pour cohésives*).

### -Bactériophage $\lambda$

Le bactériophage  $\lambda$  a été découvert par E.M. Lederberg en 1950. C'est un virus *d'E. coli*, l'ADN de ce phage est une molécule linéaire d'ADN double brin de 48 kb. A chaque extrémité 5' se trouve une région monocaténaire de 12 nucléotides, l'une complémentaire de l'autre et leur association donne une structure circulaire à l'ADN dans la cellule hôte. L'association de ces extrémités cohésive naturelles forme le site cos [cos: Des éléments important pour la réplication et l'encapsidation de bactériophage  $\lambda$ ].

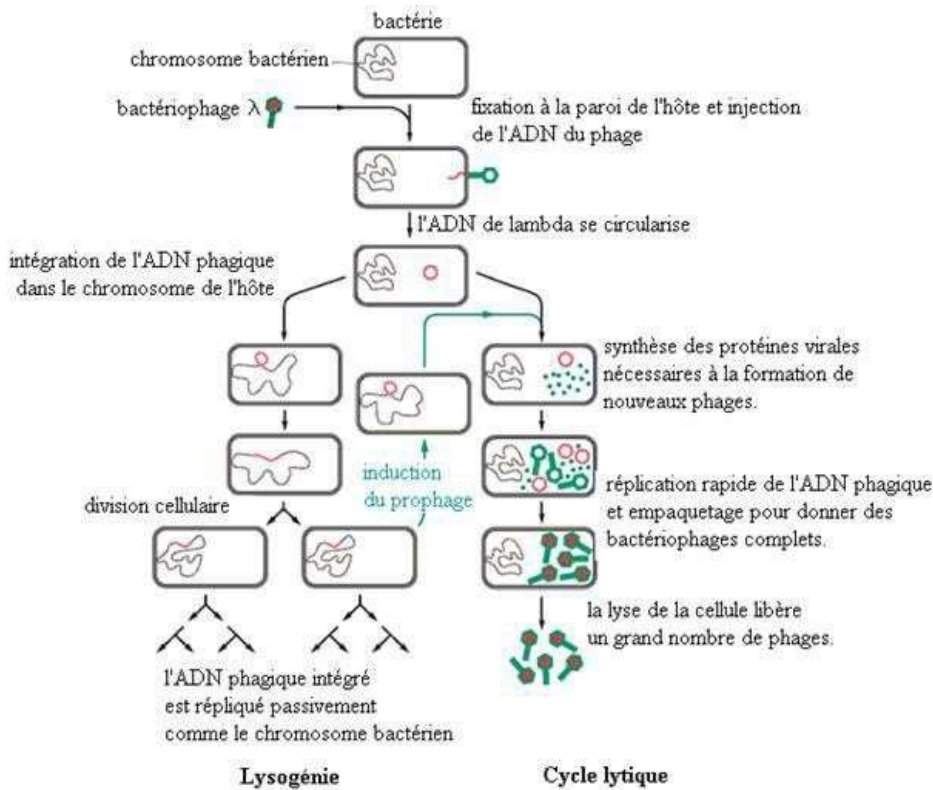
Deux parties sont individualisées dans le phage lambda:

- La tête du virus: elle renferme l'ADN viral.
- La queue du virus: elle renferme des protéines. Elle permet la fixation du virus sur la cellule hôte bactérienne.



### Structure du bactériophage $\lambda$ et les différentes formes de son génome.

Le virus lambda se multiplie selon deux modes possibles: soit par *multiplication lytique*, soit par *multiplication lysogénique*. Comme dans la multiplication lytique, le virus s'adsorbe et pénètre dans la bactérie. L'ADN viral s'intègre à l'ADN bactérien et sera répliqué en même que lui.



### Les cycles de vie du bactériophage $\lambda$ .

Le génome du phage lambda peut être divisé en parties essentielles (les gènes codant pour les protéines de la tête et les protéines de la queue du virus, les sites *cos*, le site de réplication de l'ADN viral) et en parties non essentielles (les gènes impliqués dans la lysogénie). En effet, il est possible d'éliminer les parties non essentielles.

Ces propriétés du bactériophage ont permis de concevoir un nouveau vecteur capable de transporter plus d'ADN étranger que les plasmides. On a alors un vecteur de remplacement de taille voisine de celle du génome naturel du bactériophage. Ce phage accepte des fragments d'ADN de 9 kb à 23 kb.

### **-Principales étapes de la construction d'un vecteur phagique**

Produire une grande quantité de phage, la purifier, puis en extraire son ADN génomique, qui sera digéré par une enzyme de restriction.

Hybrider les deux bras du phage avec le fragment d'ADN à cloner (ce dernier doit avoir une taille adéquate) puis souder par l'ADN ligase.

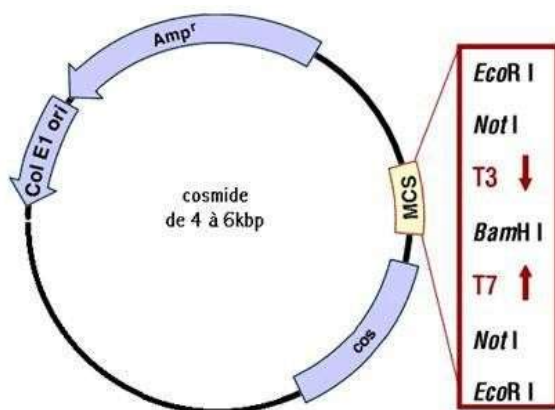
Procéder à l'encapsulation in vitro de l'ADN recombinant en ajoutant les protéines phagiques de tête et de la queue. Ces derniers ils vont s'auto assembler pour former les nouveaux virions recombinants infectieux.

Infester des bactéries (cellules hôtes) et les étaler sur boîte de Pétri, chaque plaque de lyse correspond à un phage recombinant qui peut être récupéré.

Vérifier la présence d'un insert dans l'ADN recombinant par toute procédure appropriée (hybridation ADN-ADN, séquençage).

### 3. Les cosmides

Les cosmides sont des vecteurs artificiels hybrides: phage lambda-plasmides. En fait, ils se comportent comme des plasmides avec des sites de restriction permettant l'insertion d'ADN étranger. Ils renferment également un gène de résistance aux antibiotiques (ampicilline). De plus, un site cos d'un virus lambda a été inclus dans leur ADN circulaire ce qui permettra au cosmide d'être empaqueté dans la tête d'un virus lambda.



**Schéma d'un cosmide.** (Les éléments viennent des vecteurs précédents)

Les cosmides, après assemblage in vitro sont capables de se fixer sur la paroi de la cellule hôte et d'injecter leur ADN. Celui-ci grâce aux extrémités cos se circularise et se réplique comme un plasmide mais aucune des fonctions phagiques ne peut être exprimée donc les cellules survivent et forment des clones. La taille des fragments insérables dans les cosmides peut atteindre 50 kb.

#### 4. Autres types de vecteurs

D'autres vecteurs existent avec possibilité de cloner des fragments d'ADN de taille élevée (au moins jusqu'à 500 kb).

##### -Les BAC : (Bacterial Artificial Chromosome)

Sont construits à partir du plasmide F (99.2 kb) à cause de ces propriétés intéressantes (Ex. Phénotype Hfr: mobilisation du ou souvent d'une partie du chromosome bactérien d'une cellule donatrice à une cellule réceptrice, cela après l'intégration au chromosome). Il contient qu'une petite fraction du plasmide F, et comme les vecteurs plasmidiques, il possède un polylinker, gène de résistance à un antibiotique comme marqueur de sélection (*cat*: résistance au chloramphénicol). Une origine de réplication et *rep E* qui sont nécessaires à la réplication. Des régulateurs de la réplication (*sop B*, *sop A*) pour maintenir un faible nombre de copies, sa taille est de 6.5 kb.

L'hôte typique d'un BAC est une souche mutante de *E. coli*, cette souche ne dispose pas des systèmes de modification et de restriction de la souche sauvage, pour empêcher la destruction du BAC. Ainsi, elle a perdu les capacités de recombinaison normales, cela interdit la recombinaison et les réarrangements de l'ADN cloné du BAC avec le chromosome de l'hôte. Le BAC a la capacité d'insérer jusqu'au **300 kb**.

##### -Les PAC : (P1-derived Artificial Chromosome)

Peuvent cloner des fragments d'ADN de taille variant de 100 à 200kb.

##### -Les YAC (Yeast Artificial Chromosome)

Chromosome artificiel de levure : peuvent intégrer des fragments de grande taille (jusqu'à 400kb). Indispensables pour analyser les grands génomes (notamment, le génome humain). A noter : pour obtenir de grands fragments d'ADN, on utilise des enzymes de restriction qui ne coupent que très rarement.



## ▪ Les Banques d'ADN

Le clivage de la totalité du génome d'une cellule avec une nucléase de restriction spécifique pour le clonage d'un gène est parfois qualifié de pêche à la ligne. En effet, on obtient des millions de fragments d'ADN qui produisent des millions de colonies différentes de cellules transformées. Ce qui constitue donc une banque d'ADN génomique (ADNg).

Une autre stratégie possible est de commencer le processus de clonage en sélectionnant les seules séquences d'ADN qui sont transcrites en ARN et qui sont donc supposées correspondre à des gènes : les ARN messagers.

Cette méthode consiste à extraire l'ARNm à partir des cellules et à faire ensuite une copie d'ADN complémentaire (ADNc) de chaque molécule d'ARNm présente grâce à une transcriptase inverse.

La transcriptase inverse (ou reverse) isolée des rétrovirus est capable de catalyser la synthèse d'une chaîne d'ADN à partir d'une matrice d'ARN. Les molécules d'ADN monocaténaire synthétisées par cette enzyme sont ensuite converties en molécules d'ADN bicaténaire par l'ADN polymérase.

Ces molécules d'ADNc sont ensuite insérées dans des plasmides et clonées comme précédemment. On obtient alors une banque d'ADNc.

Quand on dispose d'une banque d'ADN d'intérêt (banque génomique ou banque d'ADNc) dont on a vérifié que le gène cloné est celui que l'on veut faire exprimer, il faut alors transformer à nouveau des cellules avec le vecteur de clonage purifié (avec les mêmes techniques de transformation présentées précédemment). Cependant, pour que le gène s'exprime dans la cellule, celui-ci doit être entouré d'un ensemble de signaux que la bactérie (ou toute autre cellule transformée) devra reconnaître.

## ▪ Les vecteurs d'expression

Les vecteurs d'expression possèdent un ou plusieurs promoteurs qui doivent être puissants et aisément régulés. En effet, chaque cellule possède des promoteurs faibles de transcription de gènes codant pour des molécules requises en faible quantité par la cellule et des promoteurs forts pour des substances devant être produites en quantité importantes (Ex : production d'insuline humaine).

De plus, l'expression d'un gène peut être induite ou réprimée par la présence d'un composé spécifique.

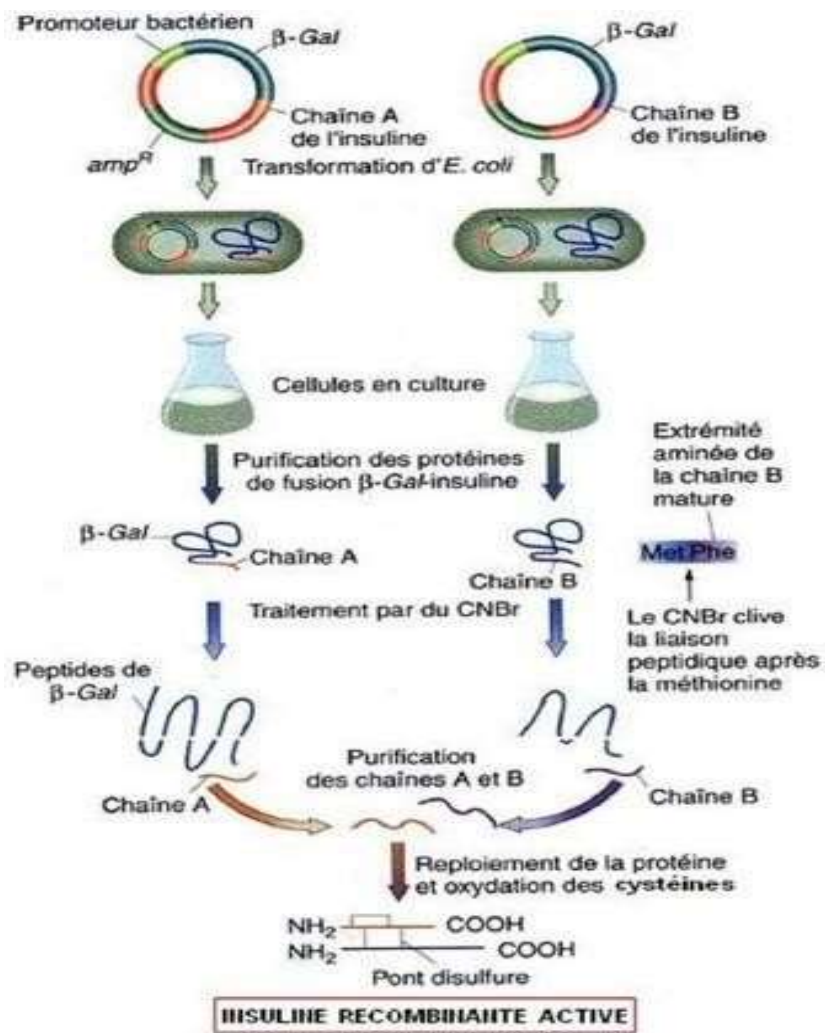
Un des promoteurs le plus couramment utilisé est le promoteur Lac qui est la séquence contrôlant la transcription du gène Lac Z codant pour la  $\beta$ -galactosidase bactérienne. Ce promoteur est induit par l'IPTG. Il est aussi possible d'utiliser d'autres promoteurs de transcription comme celui du gène de la tryptophane synthétase (Tryp E).

Il est possible d'augmenter la puissance du promoteur par des séries de mutations ou de délétions dans la séquence promotrice. On arrive alors dans certains cas à obtenir une efficacité de transcription multipliée par un facteur 10.

Le gène de la protéine à exprimer doit donc être cloné de telle sorte que sa transcription soit sous la

dépendance de ce promoteur et introduit dans la bonne phase de lecture.

Ensuite, l'expression peut être améliorée sous les conditions environnementales qui normalement activent le promoteur : c'est le cas par addition d'IPTG dans le cas du promoteur Lac.



### Vecteur d'expression

### 3. Les applications du clonage moléculaire (génie génétique) :

Les méthodes et les techniques de biologie moléculaire développées durant ces dernières décennies, pour séquencer, « recombiner », transférer et analyser les produits de l'expression du matériel génétique des diverses espèces vivantes ont donné lieu à une application sous le nom de « **génie génétique** ».

Le transfert transitoire ou stable d'un DNA étranger dans une cellule procaryote ou eucaryote s'opère grâce à des vecteurs naturels et à de nouveaux vecteurs issus des constructions: plasmides, bactériophages, virus, cosmides etc. Ces techniques doivent permettre une meilleure compréhension génétique et fonctionnelle des organismes vivants.

Il s'agit donc d'un outil, aux applications très variées qui permet d'intervenir avec une grande précision sur le patrimoine génétique de tous les organismes vivants. Le résultat obtenu au cours de cette manipulation est la protéine originelle du premier organisme. Cependant, le génie génétique permet également d'apporter des modifications à des gènes par une mutagenèse dirigée qui par conséquent produiront des protéines modifiées ou recombinantes.

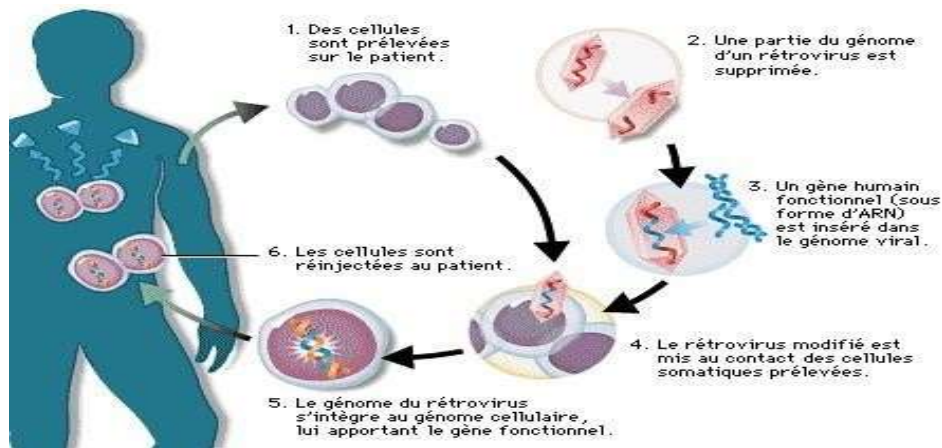
Les applications du clonage moléculaire sont multiples et peuvent être organisées selon quatre domaines d'application : en recherche (mutagenèse dirigée et création d'animaux transgéniques), en médecine (diagnostic des maladies héréditaires et leur soin à travers la thérapie génique), en industrie pharmaceutique (production de protéines recombinantes humaines) et en industrie agroalimentaire (amélioration génétique des espèces animales et végétales «OGM »).

**Exemple d'application : « La thérapie génique »** ; Autrement dit, les applications de la biologie moléculaire en génétique humaine.

La thérapie génique est une méthode visant à traiter une maladie en modifiant son information génétique. Le principe de la thérapie génique est de remplacer, à l'intérieur des cellules malades, un gène muté défectueux responsable de la maladie par un gène fonctionnant normalement dans le but de produire des protéines thérapeutiques nécessaires pour corriger la maladie visée. Le ou les gènes responsables ont été identifiés préalablement. La thérapie génique somatique qui est différente de la thérapie génique germinale, consiste à transférer certains gènes dans les cellules du patient pour prévenir l'apparition d'une maladie ou en ralentir l'évolution. Ce type de génothérapie suscite de très grands espoirs. L'introduction d'un gène médicament à l'intérieur de la cellule cible du patient requiert un vecteur dont le rôle est d'intégrer le gène d'intérêt et de le transférer dans un nombre suffisant de cellules cibles, pour arriver au niveau du noyau de la cellule, et permettre l'expression du gène et la production de protéines sur une période suffisante pour obtenir un effet thérapeutique. Les vecteurs les plus utilisés sont des vecteurs viraux, AAV, lentivirus etc....

Bien sûr le virus utilisé est modifié et atténué afin de transférer le matériel génétique souhaité dans les

cellules du patient sans entraîner chez celui-ci des réactions immunitaires non désirées. La difficulté est que le virus risque de s'attaquer à une grande variété de cellules cibles. Le transfert de gènes dans les cellules du patient peut être opéré *in vivo* (à l'intérieur de l'organisme), *Ex vivo* (en dehors de l'organisme) ou *In situ* (au sein du tissu cible).



### **Thérapie génique somatique *ex vivo* utilisant des virus modifiés comme vecteur.**

La thérapie génique est aujourd'hui d'actualité car elle constitue une approche thérapeutique complètement nouvelle. Des développements considérables seront néanmoins nécessaires avant que l'on sache si une proportion significative des millions de malades atteints de maladies génétiques pourra bénéficier de tels traitements.

Le génie génétique possède, sans doute, aujourd'hui un champ d'application très large et nous ouvre des perspectives nouvelles. Il s'agit incontestablement d'une technologie clé pour ce troisième millénaire. Cependant, comme toute technique, le génie génétique a ses limites. Ainsi, dans un organisme vivant, la production d'une protéine de nature et de fonction données peut nécessiter des informations présentes à plusieurs endroits de l'ADN. Ces informations plus complexes et plus complètes peuvent amener un gène à produire de façon différente une même protéine selon l'état physiologique de la cellule, de son stade de développement ou de différenciation. Plusieurs gènes doivent donc s'associer dans ce processus, afin de fournir à l'organisme les quantités et les qualités précises de cette protéine en fonction de son rôle biologique.

Si la technologie génétique sait parfaitement identifier, isoler et modifier un gène particulier, elle a beaucoup plus de difficultés à l'heure actuelle pour déterminer les liens existant entre les différents gènes. Leur transfert et leur expression, parfois ectopique, posent encore des problèmes qui restreignent, pour l'instant, le champ d'application du génie génétique.

## Références bibliographiques :

**Ameur Ameur Abdelkader, (2015).** Génétique générale. Ed El-Djazair. 101p.

**Aouf Abdelhakim, (2016).** Cours de Biologie moléculaire et génie génétique. Université Ferhat-Abbas Setif 1, 133p.

**Beaumont Simon, (2007).** Biologie moléculaire PCEM1, Dunod, Paris 180-193p.

**Carl Herrmann, (2006).** La régulation de la transcription. Institut de Biologie du développement Marseille Luminy. Université de la Méditerranée 56p.

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

**Cyril Esnault, (2007).** Étude des interactions entre le Médiateur et les facteurs généraux de la transcription par l'ARN polymérase II. Thèse pour obtenir le grade de Docteur de l'université Paris XI Orsay 8-35p.

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

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

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

**Karouche Saida, (2018).** Cours de Biologie moléculaire et génie génétique. Université Larbi Ben Mhidi. Oum el bouaghi, 161p.

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

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

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

**Medraoui Leila, (2015).** Cours Biologie Moléculaire ; Initiation aux techniques usuelles de biologie moléculaire. Faculté des sciences. Université Mohamad V Rabat 132p.

**Passarge E, (2007).** Color Atlas of Genetics. 3<sup>rd</sup> Ed. Flexibook. Thieme Stuttgart., NY. USA. 486p.

**Perrin Pascale, (2010).** Le contrôle de l'expression génétique FLBI399. Université Montpellier 2-53p.

**Perrin-Schmitt Fabienne, (2011).** Biotechnologies et applications (génie génétique), PACES. UE2. Ellipses 70-74p.

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

**Roberts R. J., Belfort M., Bestor T., Bhagwat A.S, (2003).** « SURVEY AND SUMMARY: A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes », Nucleic Acids Res. 31 : (7), 1805–12.

**Simpore J. K, (1999).** Cours de génie génétique. Université de

Ouagadougou. 107p.

**Smith H. O., Nathans D, (1973)** « Letter: A suggested nomenclature for bacterial host modification and restriction systems and their enzymes », *J. Mol. Biol.*, 81 : (3), 419–23.

**Strachan Tom., Read Andrew, (2004).** Génétique Moléculaire Humaine. Hybridation des acides nucléiques : principes et applications. 4<sup>e</sup> édition ; Médecine Sciences publications. Lavoisier 16p.

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

**Tagu D., Moussard C, (2003).** Principes des techniques de biologie moléculaire. 2<sup>e</sup> éd., INRA-Quæ (ISBN 2-7380-1067-9).

**Vincent Ecochard., Didier Fournier., Laurence Nieto., Laurent Paquereau, (2011).** Techniques et Stratégies en Biologie Moléculaire. Première partie. Université Paul Sabatier, Toulouse, 201p.

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