

# Structure et organisation de l'ADN (acide désoxyribonucléique)

L'ADN constitue le patrimoine génétique de pratiquement toutes les espèces vivantes (à l'exception de quelques virus qui sont à ARN), il est transmis de génération en génération. Le message porté par l'ADN spécifie le fonctionnement et la reproduction de chaque organisme vivant.

Chez les eucaryotes on retrouve la quasi-totalité de l'ADN dans le noyau, chez les procaryotes l'ADN circulaire baigne dans le cytoplasme.

L'ADN appartient à la famille chimique des acides nucléiques, c'est un grand polymère défini par une séquence linéaire d'unités simples répétées.

## **Structure de l'ADN**

### **Structure primaire**

L'ADN est un polynucléotide. L'unité de base est le désoxyribonucléotide, ce dernier est formé de :

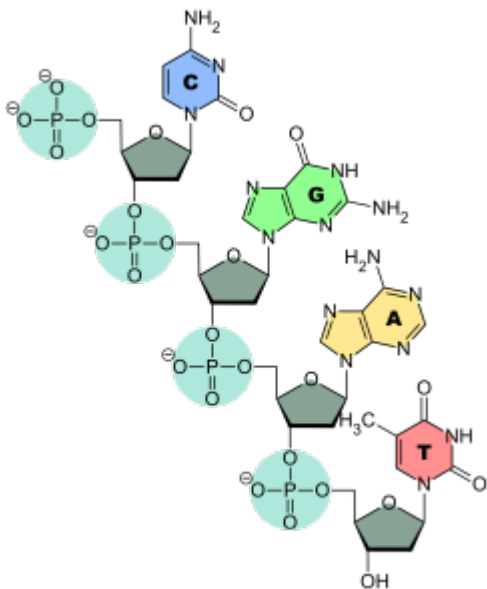


Figure. Structure primaire de l'ADN

a- Bases azotés : soit des bases pyrimidiques : la cytosine et la thymine ou des bases puriques adénine et guanine. Les bases pyrimidiques sont formées d'un noyau aromatique hexagonal (de 6 atomes) avec 4 carbones et 2 azotes. C : est un noyau pyrimidique dans lequel le C4 contient une fonction amine et le C2 une fonction cétone. T : le C2 et le C4 ont une fonction cétone et le C5 à une fonction méthyle (c'est de l'uracile méthylé).

Les bases pyrimidiques contiennent 2 noyaux cycliques accolés : un hexagonal et l'autre pentagonal les 2 noyaux ont 2 carbones en commun au milieu. A : est un noyau purique dont le C6 contient une fonction amine et c'est la seule base qui ne contient pas d'atomes d'oxygène. G : c'est un noyau purique dont C2 possède une fonction amine et le C6 une fonction cétone.

b- sucre : c'est un sucre à 5 carbones (pentose) : le désoxyribose. c- groupement phosphate : l'acide phosphorique  $H_3PO_4$  Le groupement phosphate se lie au sucre au niveau du C5 et la base azotée se lie au C1 du sucre par une liaison N-Osidique. Chaque nucléotide va se lier à un autre nucléotide de la manière suivante :

Le 1er nucléotide se lie au groupe  $P_{O_4}$  du 2e nucléotide au niveau du C3 de son désoxyribose, et ainsi de suite formant une succession de liaisons 5'-3' phosphodiesters ce qui va donner un enchaînement linéaire qui constitue la structure laire de l'ADN.

Par convention le sens d'un brin d'ADN commence par l'extrémité 5' libre & se termine par l'extrémité 3'OH du dernier nucléotide libre.

Un Nucléotide = nucléoside + acide phosphorique;

Un Nucléoside = base + sucre;

## Structure secondaire

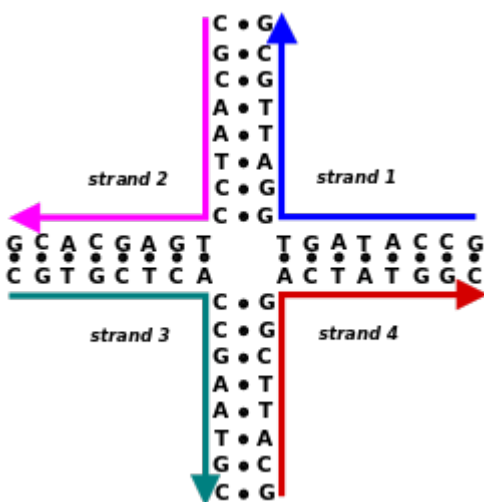


Figure. Structure secondaire d'ADN

In vivo, l'ADN se présente sous la forme d'un double brin dont chaque brin présente une orientation opposée, les brins sont maintenus par des liaisons hydrogène qui s'établissent entre les bases. A se lie toujours à T par 2 liaisons hydrogène et C toujours à G par 3 liaisons hydrogène. Cette liaison peut être facilement interrompue par la chaleur et les agents chimiques. Les brins d'ADN sont dits alors complémentaires et antiparallèles.

On compare l'ADN double brin à une échelle dont les montants

constituent le squelette sucre et les marches constituent les bases. Cette échelle va se tordre pour donner la fameuse configuration de double hélice qui représente la structure secondaire de l'ADN.

les différentes formes de l'ADN :

Il existe plusieurs formes d'ADN chez les eucaryotes, elles diffèrent par le nombre de nucléotides par tour d'hélice et par la distance entre 2 bases azotées adjacentes.

ADNb :

C'est la forme la plus fréquente (physiologique), c'est la forme décrite par CRICK et WATSON en 1953, c'est une double hélice droite, l'appariement de bases conduit à la formation d'un grand sillon et d'un petit sillon. La distance entre 2 bases successives est de 0,34 nm. La longueur d'un tour complet de spire est de 3,4 nm et donc chaque tour de spire contient 10 bases.

ADNa :

C'est une forme rare qui existe seulement à l'état déshydraté (cristallisé) , c'est une double hélice d'ADN à enroulement droit contenant des bases très inclinées constituant une hélice plus courte et plus large que celle de l'ADNb.

ADNz :

C'est une forme à enroulement gauche, la distance entre 2 bases est de 0,77 nm, le squelette sucre phosphate de la double hélice à une forme de zigzag. Elle est parfois retrouvée dans certaines conditions dans les cellules eucaryotes, mais son intérêt exact n'est pas encore connu.

Les autres formes d'ADN :

ADNc :

Très voisine de l'ADNb 3,3 nm par tour de spire et 9 paires de bases par tour de spire.

ADNp :

Retrouvée chez le virus PFL en moyenne 3,6 bases par tour de spire.

## Structure tertiaire

L'ADN est étroitement lié à certaines protéines pour qu'il soit condensé et compacté au maximum, dans le cas contraire il ne tiendrait pas dans le noyau. Si on déroule l'ADN humain il mesurera 1,8 à 2 m. La liaison entre l'ADN et les protéines vont nous donner la structure tertiaire, elle représente en fait la fibre chromatinienne.

Les principales protéines associées à l'ADN sont des protéines basiques caractérisées par leur richesse en lysine et en arginine ce sont les Histones. On en distingue généralement 5 types (un 6ème type a été retrouvé chez les oiseaux et les amphibiens c'est le H5)

Les types d'histones sont : H1, H2a, H2b, H3, H4.

Lorsque l'on isole la fibre chromatinienne elle se présente comme un collier de perles dont le fil serait l'ADN reliant des structures nommées Nucléosomes (la perle); L'ensemble est appelé fibre A chromatinienne.

Un nucléosome comporte un octamère d'histones constitué de 2 molécules de chacune des protéines histones suivantes : H2a H2b H3 et H4 et un segment d'ADN de 200 paires de bases.

On peut subdiviser l'ADN en :

- un noyau nucléosomique : l'octaèdre + un segment d'ADN de 140 paires de bases qui fait un tour et 3 quarts autour de l'octaèdre.
- 2 liens inter-nucléosomique : ce sont les segments d'ADN

qui relie les noyaux nucléosomique ( $2 \times 30$  paires de bases).

- Le nucléosome constitue le premier degré de condensation de l'ADN (la fameuse fibre A).

Le deuxième degré de condensation consiste en l'enroulement d'une succession de nucléosomes en un segment hélicoïdal pour former une fibre Solénoïde (c'est la fibre B). La condensation fait intervenir les protéines H1 qui agissent au niveau des liens inter-nucléosomiques. Le degré final de condensation est dû à l'enroulement de la fibre solénoïde en super boule. Le degré extrême de condensation est : Le chromosome métaphasique.

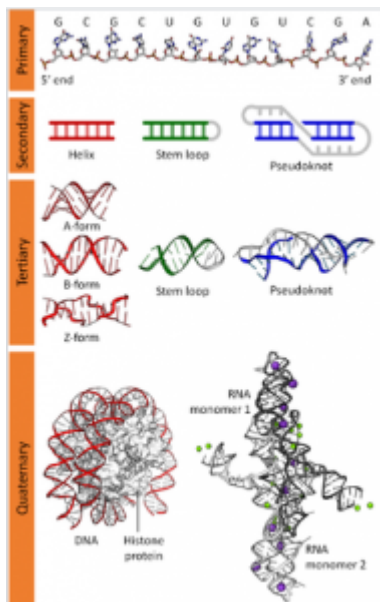


Figure. Structures de l'ADN

## Chromatine Inter-phasique

Morphologie : la chromatine est constituée d'ADN organisé en fibres nucléotidiques (complexe nucléoprotéique). Selon le type cellulaire elle peut apparaître en MP après coloration soit en grosse motte régulière, ou en granulations fines. La finesse de la chromatine exprime le degré d'activité de la cellule.

Pendant la division cellulaire la chromatine va se condenser

pour former le chromosome. On peut subdiviser la chromatine en hétérochromatine ou en Euchromatine.

Hétérochromatine :

Elle représente 80% à 90% de l'ADN nucléaire, c'est une chromatine dense qui n'est pas transcrite. C'est de l'ADN inactif, elle a une structure solénoïde, elle est très riche en Histones H1 et existe sous deux formes :

- constitutive : c'est de l'ADN non génique, elle est présente autour des centromères et au niveau du bras long du chromosome Y
- facultative : contient des gènes réprimés, la séquence des gènes réprimés diffère d'un type cellulaire à un autre, c'est ce mécanisme qui représente la différenciation cellulaire par ex : le gène de la cytokératine qui est exprimé dans les kératinocytes et réprimé dans les fibroblastes

La quasi-totalité d'un des chromosomes X chez la personne de sexe féminin est réprimé et forme le corpuscule de BARR « c'est une structure biconvexe plaquée contre la face interne de la membrane nucléaire

Euchromatine:

Elle est non visible au MP, ce sont les fibres chromatiniennes despiralisées (fibre A) ce sont des fibres actives génétiquement, elle correspond à des molécules d'ADN en cours de transcription ou de réplication.

## Biochimie

Le contenu du noyau est constitué de 50% de protéines non histones, 23% d'Histones, 23% d'ADN, 4% d'ARN. Les protéines histones peuvent subir des modifications à savoir: phosphorylation, méthylation, acétylation (joue un rôle dans la régulation d'expression des gènes). L'histone H1 est en

quantité considérable au niveau de l'hétérochromatine et peut subir une phosphorylation pour donner une compaction accrue de la chromatine.

Les protéines non histones sont des protéines acides, on peut trouver:

- les protéines de structuration du noyau ex : la Lamine;
- les protéines enzymatiques ex : ARN polymérase;
- les protéines de la régulation de la transcription : les protéines en doigts de zinc, les protéines leu-zipper (riches en leucine, et possèdent une forme de fermeture éclair);
- les protéines qui pénètrent temporairement dans le noyau ex : Ubiquitine qui dégrade les protéines altérées.

## **Organisation de l'information génétique**

### **Génome d'une espèce**

L'ensemble de l'ADN contenu dans une cellule d'une espèce constitue son génome ou son patrimoine génétique.

- chez les eucaryotes la quasi-totalité du génome se trouve dans le noyau, une petite quantité d'ADN est retrouvée dans la mitochondrie (génome mitochondrial) et dans le chloroplaste chez les végétaux
- chez les procaryotes en plus de leur génome qui est représenté par l'ADN circulaire qui baigne dans le cytoplasme, on retrouve le génome plasmidique.



## Taille du génome « valeur C »

La taille du génome est définie comme la quantité d'ADN présente dans les gamètes (haploïdes) d'une espèce donnée. Les mesures sont effectuées en générale sur les spermatozoïdes afin d'éviter les variations de teneur en ADN qui ont lieu lors du cycle cellulaire de la cellule somatique.

La quantité d'ADN étant constant pour une espèce donnée on la nomme la valeur C.

## Classes d'ADN

Pseudo gène : structure qui ressemble à la structure d'un gène donne, mais qui est non fonctionnelle ex : pseudo gène de la famille des globines.

ADN satellite : c'est une séquence de 100 à 6500 paires de bases répétées en tandem.

ADN mini- satellite : 10 à 20 paires de bases répétées en tandem.

ADN microsatellite : 2 à 5 paires (c'est le plus fréquent)

Line : (long interspersed repeat séquences) : ce sont des séquences répétées dans le génome en longueur de 6500 paires de bases.

Sine: (short interspersed repeat sequences)

LTR: (long terminal repeat)

R! : Chez les procaryotes la majorité de l'ADN est codante et la séquence d'un gène ne comporte pas des parties non codantes

## Les gènes

Définition : c'est l'unité de l'hérédité, c'est une séquence

d'ADN qui est nécessaire à la production d'un produit fonctionnel. Ex : un gène qui code pour une enzyme.

## Structure

Ce type de gène s'appelle : gène morcelé ou en mosaïque car il contient des introns (fragments non exprimés).

Selon leur fonction les gènes sont divisés en :

- Gènes de structure: codent pour des protéines.
- Gènes régulateurs: régulent l'expression ou l'inhibition d'autres gènes.

On peut classer les gènes selon le nombre de leurs copies en :

- gène unique ou quasi-unique: c'est la grande majorité des gènes;
- les familles de gènes: ce sont des gènes qui codent pour des protéines grossièrement analogues ex : les gènes qui codent pour l'actine.

« Selon la théorie de l'évolution ça serai à l'origine un seul genet qui se serai dupliqué, mais qui aurai subit à chaque fois des mutations »

- les super familles de gènes: regroupent des gènes dont l'homologie est partielle (se ressemblent moins que ceux des familles de gènes), mais qui conservent toujours des séquences fonctionnelles superposables par ex : super famille des immunoglobulines.
-

# Introduction à l'étude de la génétique

Il est courant de discuter des ressemblances entre les personnes de la même famille. Les caractères qui les unissent se transmettent de génération en génération des parents aux enfants à travers les gamètes.

L'étude de la transmission de ces caractères est à la base de la création d'une discipline scientifique qui est la science de l'hérédité qu'on appelle génétique (genos = origine).

Cette discipline est relativement jeune, elle a pratiquement un siècle d'existence mais elle a fait de pas de géants dans son développement.

## **Historique**

Le début de la génétique a été marqué par les expériences de Mendel sur les petits pois qui a été publié en 1866, ces publications passèrent inaperçues, ce n'est que 40 ans plus tard qu'on a reconnu leur importance. En 1905, Bateson définit le terme de gène et de génétique (Mendel parlait de caractère et non pas de gène).

En 1941, Beadle et Tatum établissent la correspondance entre gène et protéine, c'est à dire la traduction de gènes en enzymes.

En 2001, publication de séquençage d'environ 94% du génome humain et estimation de 30000 à 40000 gènes grâce à la bioinformatique.

## **Les différents domaines de la**

# **génétique**

## **Génétique moléculaire**

La génétique moléculaire s'intéresse à :

Etude des bases moléculaires de l'hérédité : Les Acides nucléiques : ADN et les ARN.

Organisation de l'information génétique : structure du gène et du chromosome.

Les variations génétiques : les polymorphismes, les mutations et la mutagenèse : la différence entre les polymorphisme et les mutations est que les polymorphismes sont les variations qui n'entraînent pas de maladies alors que les mutations sont des variations qui peuvent entraîner des maladies ex : les cancers.

La mutagenèse : l'étude des facteurs qui peuvent entraîner des mutations.

La régulation de l'expression d'un gène : étudie les phénomènes qui régissent l'expression d'un gène plutôt qu'un autre dans des conditions bien précises car toutes les cellules d'un organisme ont le même ADN mais au niveau de chaque type cellulaire il y 'a des gènes spécifiques qui s'expriment et il y'a d'autres silencieux.

La génétique bactérienne : (bactérie-virus) : Pour son intérêt dans la génie génétique et dans la thérapie génétique

Les techniques d'extraction de l'ADN (PCR) : amplification d'un gène séquençage : pour connaître la structure de l'ADN.

## **Génétique formelle (Mendélienne) et**

## **humaine : ou classique**

La génétique classique s'intéresse à:

- a- transmission classique : mono hybridisme.
- b- non classique (mitochondriale) : emprunte parentale.
- c- établissement d'un arbre généalogique (pedigree).
- d- le conseil génétique.

## **Cytogénétique : génétique des chromosomes (cytogénétique du cancer)**

La cytogénétique étudie le caryotype humain normal, les anomalies du caryotype constitutionnel, du nombre et de la structure ou des anomalies acquises ex : cancers, leucémie myéloïde chronique LMC

## **Génie génétique**

La manipulation des gènes à des fins industrielles et pharmaceutiques ex : OGM (organisme génétiquement modifié), insuline, interférons.

## **Thérapie génétique**

Utilisation des virus pour remplacer les gènes défectueux par des gènes sains.

## **Génétique des populations**

Etude des races-ethnicités : étudie l'évolution des caractères dans les populations.

## **Génétique médicale**

S'occupe des maladies d'origine génétique qui sont divisées en :

a- Maladies mono-géniques : dues à l'altération d'un seul gène  
ex : Hémophilie (problème dans la coagulation sanguine), le Daltonisme.

b- Maladies chromosomiques : dues à une anomalie du nombre ou de structure. Ex : trisomie 21.

c- Maladies polygéniques ou multifactorielles : ce sont des maladies qui résultent de l'altération de plusieurs gènes  
ex : diabète sucré, HTA.

d- Maladies mitochondriales.

e-onco-génétique : mécanismes génétiques retrouvés dans les cancers ex : cancer du sein (2 gènes BRCA1 et BRCA 2) et le chromosome philadelphie dans la LMC.

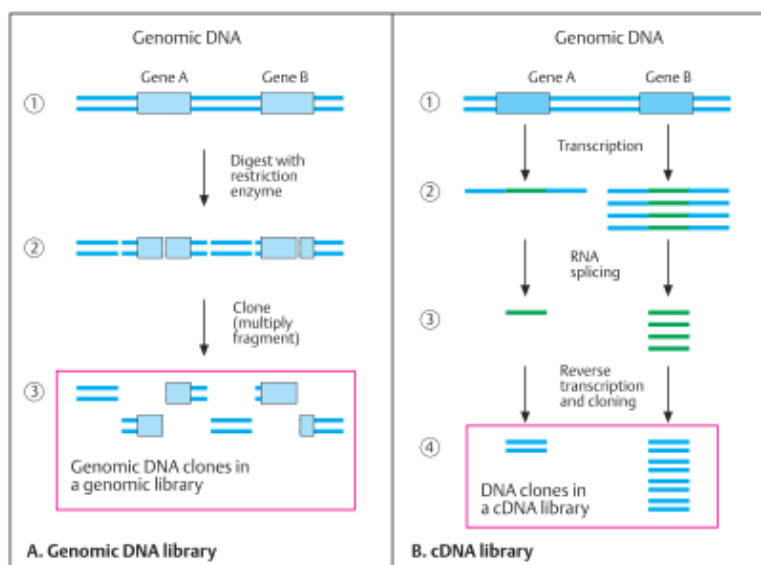
---

## Bibliothèque d'ADN

Une bibliothèque d'ADN est une collection de fragments d'ADN qui représentent dans son ensemble le génome, c'est-à-dire un gène d'ADN particulier. C'est le point de départ pour le clonage d'un gène de localisation chromosomique inconnue. Pour produire une banque d'ADN, l'ADN total est digéré avec une enzyme de restriction et les fragments résultants sont incorporés dans des vecteurs et répliqués dans des bactéries. Un nombre suffisant de clones doit être présent au moins une fois. Ceci est une question de la taille du génome étudié et de la taille du fragment. Les plasmides et les phages sont utilisés comme vecteurs. Pour des fragments d'ADN plus grands, des cellules de levure peuvent être utilisées. Il existe deux types de bibliothèques: l'ADN génomique et l'ADNc.

# Bibliothèque d'ADN génomique

Les clones d'ADN génomique sont des copies de fragments d'ADN de tous les chromosomes (1). Ils contiennent des séquences codantes et non codantes. Les enzymes de restriction sont utilisées pour cliver l'ADN génomique en plusieurs fragments. Ici, quatre fragments sont représentés schématiquement, contenant deux gènes, A et B (2). Ceux-ci sont incorporés dans des vecteurs, par exemple dans de l'ADN de phage, et sont répliqués dans des bactéries. La collection complète de molécules d'ADN recombinant, contenant toutes les séquences d'ADN d'une espèce ou d'un individu, est appelée une bibliothèque génomique. Pour trouver un gène particulier, une procédure de dépistage est nécessaire (voir B).



ADN génomique et bibliothèque d'ADNc

# Bibliothèque d'ADNc

Contrairement à une bibliothèque génomique, qui est complète et contient de l'ADN codant et non codant, une bibliothèque d'ADNc se compose uniquement de séquences d'ADN codantes. Cette spécificité offre des avantages considérables par rapport à l'ADN génomique. Cependant, il nécessite que l'ARNm soit disponible et ne donne pas d'informations sur la

structure du gène. L'ARNm peut être obtenu uniquement à partir de cellules dans lesquelles le gène respectif est transcrit, c'est-à-dire dans lequel l'ARNm est produit (1). Chez les eucaryotes, l'ARN formé pendant la transcription (transcription primaire) subit un épissage pour former l'ARNm. L'ADN complémentaire (ADNc) est formé à partir de l'ARNm par l'enzyme transcriptase inverse (3). L'ADNc peut servir de matrice pour la synthèse d'un brin d'ADN complémentaire, de sorte qu'un ADN double brin complet peut être formé (clone d'ADNc). Sa séquence correspond aux séquences codantes des exons du gène. Ainsi, il est bien adapté pour une utilisation en tant que sonde (sonde d'ADNc). Les étapes suivantes, l'incorporation dans un vecteur et la répllication dans des bactéries, correspondent à celles de la procédure pour produire une banque génomique. Les clones d'ADNc ne peuvent être obtenus qu'à partir de régions codantes d'un gène actif (producteur d'ARNm); ainsi, les clones d'ADNc de différents tissus diffèrent selon l'activité génétique. Puisque les clones d'ADNc correspondent aux séquences codantes d'un gène (exons) et ne contiennent pas de sections non codantes (introns), l'ADNc clone est le matériau de départ préféré lorsque des informations supplémentaires sur un produit génique sont recherchées en analysant le gène. La séquence d'acides aminés dans une protéine peut être déterminée à partir d'ADNc cloné et séquencé. De même, de grandes quantités d'une protéine peuvent être produites en ayant le gène clone exprimé dans des bactéries ou des cellules de levure.

## **Dépistage d'une bibliothèque d'ADN**

Les bactéries qui ont absorbé les vecteurs peuvent se développer sur une boîte de Pétri enduite d'agar, où elles forment des colonies (1). Une réplique empreinte de la culture est prise sur une membrane (2), et l'ADN qui adhère à la membrane est dénaturé avec une solution alcaline (3). L'ADN du segment de gène recherché peut ensuite être identifié par hybridation avec une sonde marquée radioactivement (ou



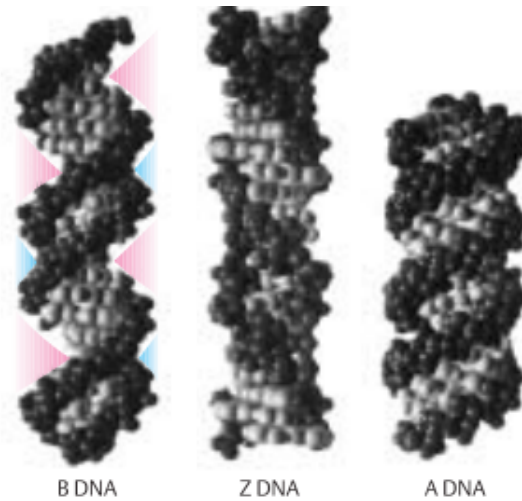
autrement) (4). Après l'hybridation, un signal apparaît sur la membrane au site du segment de gène (5). L'ADN complémentaire de la sonde marquée est localisé ici; sa position exacte dans la culture correspond à celle du signal sur la membrane (6). Une sonde est prélevée dans la zone correspondante de la culture (5). Il contiendra le segment d'ADN désiré, qui peut maintenant être encore répliqué (cloné) dans les bactéries. Par ce moyen, le segment désiré peut être enrichi et est disponible pour des études ultérieures.

---

## Les structures alternatives de l'ADN

L'expression et la transcription des gènes peut être influencée par des modifications de la topologie de l'ADN. Cependant, ce type de contrôle de l'expression génique est relativement universel et non spécifique. Ainsi, il est plus approprié pour la suppression permanente de la transcription, par exemple, dans les gènes qui sont exprimés seulement dans certains tissus ou sont actifs seulement pendant la période embryonnaire et devenant plus tard définitivement inactifs.

### **Trois formes d'ADN**



Les trois formes d'ADN

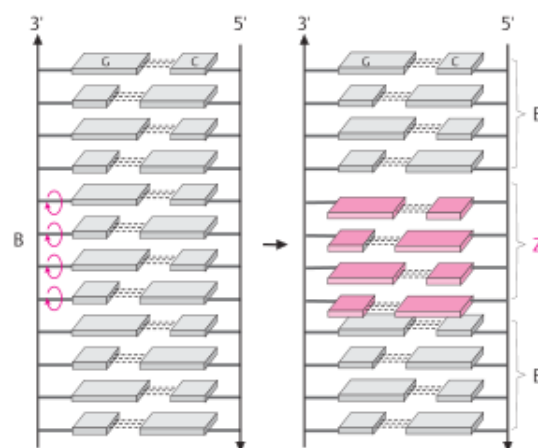
La double hélice d'ADN ne se présente pas comme une structure unique, mais représente plutôt une famille structurale de différents types. La forme classique originale, déterminée par Watson et Crick en 1953, est l'ADN-B. La caractéristique structurale essentielle de l'ADN-B est la formation de deux sillons, l'un grand (sillon principal) et l'autre petit (sillon mineur). Il existe au moins deux autres formes alternatives de la double hélice d'ADN, de l'ADN-Z et de la forme rare de l'ADN-A. Alors que l'ADN-B forme une hélice droite, l'ADN-Z montre une conformation gauchère. Cela conduit à une plus grande distance (0,77 nm) entre les paires de bases que dans l'ADN B et une forme de zigzag (d'où la désignation Z-ADN). L'ADN A est rare. Il n'existe que dans l'état déshydraté et diffère de la forme B par une rotation de 20 degrés de l'axe perpendiculaire de l'hélice. L'ADN-A a un sillon majeur profond et un sillon mineur plat (Figures de Watson et al, 1987).

## Sillons majeurs et mineurs dans l'ADN-B

L'appariement des bases dans l'ADN (adénine-thymine et guanine-cytosine) conduit à la formation d'un grand et d'un petit sillon car les liaisons glycosidiques au désoxyribose

(dRib) ne sont pas diamétralement opposées. Dans l'ADN B, les anneaux de purine et de pyrimidine sont espacés de 0,34 nm. L'ADN a dix paires de bases par tour de la double hélice. La distance d'un tour complet à l'autre est de 3,4 nm. De cette manière, des courbes localisées apparaissent dans la double hélice. Le résultat est un sillon un peu plus grand et un autre un peu plus petit. <fn> Stryer, L.: Biochimie, 4 e éd. W.H. Freeman & Co., New York, 1995. </ Fn>

## Transition entre l'ADN B et l'ADN-Z



Transition de l'ADN-B  
vers l'ADN-Z

L'ADN B est une double hélice régulière parfaite, sauf que les paires de bases opposées ne se trouvent pas exactement au même niveau. Ils sont tordus de manière hélicoïdale. De cette façon, l'ADN peut facilement être plié sans provoquer de changements essentiels dans les structures locales. Dans l'ADN-Z, le squelette sucre-phosphate a un motif en zigzag; le sillon unique de l'ADN-Z a une plus grande densité de molécules chargées négativement. L'ADN-Z peut apparaître dans des segments limités in vivo. Un segment d'ADN B constitué de paires de GC peut être converti en ADN-Z lorsque les bases sont tournées de 180 degrés. Normalement, l'ADN-Z est thermodynamiquement relativement instable. Cependant, la transition vers l'ADN-Z est facilitée lorsque la cytosine est

méthylée en position 5 (C5). La modification de l'ADN par la méthylation de la cytosine est fréquente dans certaines régions de l'ADN des eucaryotes. Il existe des protéines spécifiques qui se lient à l'ADN-Z, mais leur rôle pour la régulation de la transcription n'est pas claire.

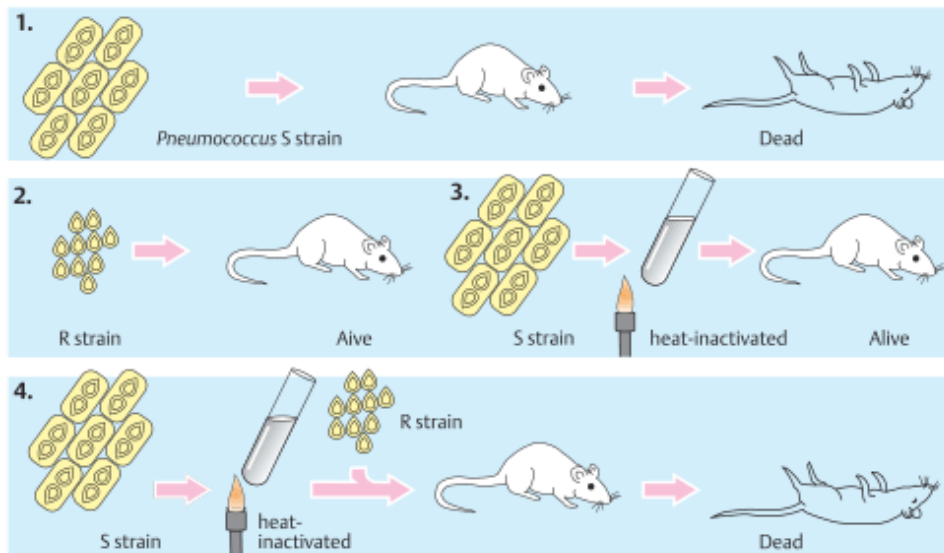
Source: Watson, J.D. et al .: Molecular Biology of the Gene. 3ème éd. Benjamin / Cummings Publishing Co., Menlo Park, Californie, 1987.

---

## L'ADN comme un support de l'information génétique

Bien que l'ADN ait été découvert en 1869 par Friedrich Miescher comme une nouvelle substance acide contenant du phosphore composée de très grosses molécules qu'il a baptisées «nucléine», son rôle biologique n'a pas été reconnu. En 1889, Richard Altmann a introduit le terme «acide nucléique». En 1900, les bases puriques et pyrimidiques étaient déjà connues. Vingt ans plus tard, les deux types d'acides nucléiques, l'ARN et l'ADN, ont été distingués. Une observation accidentelle mais précise (1928) et des investigations pertinentes (1944) ont indiqué que l'ADN pourrait être le vecteur de l'information génétique.

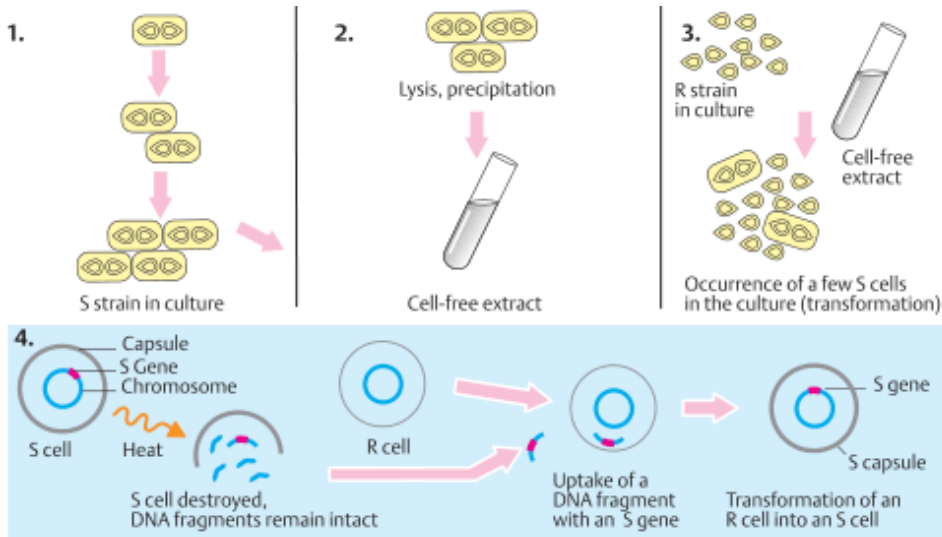
## **L'observation de Griffith**



## L'observation de Griffith

En 1928, le microbiologiste anglais Fred Griffith a fait une observation remarquable. En étudiant différentes souches de *Pneumococcus*, il a déterminé que les souris ayant reçu une injection de la souche S (lisse) sont mortes (1). D'autre part, les animaux injectés avec la souche R (rugueuse) vivaient (2). Quand il a inactivé la souche S létale par la chaleur, il n'y a pas eu de séquelles et l'animal a survécu (3). De façon surprenante, un mélange de la souche R non létale et de la souche S inactivée par la chaleur avait un effet léthal comme la souche S (4). Et il a trouvé des pneumocoques vivants normaux de la souche S dans le sang de l'animal. Apparemment, les cellules de la souche R ont été transformées en cellules de la souche S. Pendant un certain temps, ce résultat surprenant n'a pas pu être expliqué. Sa pertinence pour la génétique n'était pas apparente.

**Le principe de transformation est l'ADN**



L'origine de la

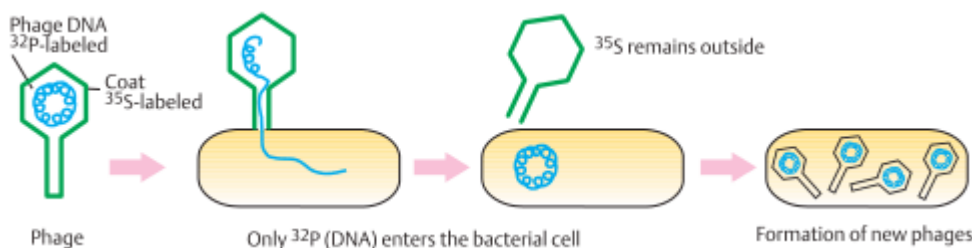
transformation est l'ADN

Les conclusions de Griffith ont servi de base aux enquêtes d'Avery, MacLeod et McCarty (1944). Avery et ses collègues du Rockefeller Institute de New York ont élucidé la base chimique du principe de transformation. A partir de cultures d'une souche S (1), ils ont produit un extrait de cellules lysées (extrait acellulaire) (2). Après que toutes ses protéines, ses lipides et ses polysaccharides aient été éliminés, l'extrait conservait la capacité de transformer les pneumocoques de la souche R en pneumocoques de la souche S (principe de transformation) (3). Avec d'autres études, Avery et ses collègues ont déterminé que cela était attribué à l'ADN seul. Ainsi, l'ADN doit contenir l'information génétique correspondante. Ceci explique l'observation de Griffith. L'inactivation par la chaleur avait laissé intact l'ADN des chromosomes bactériens. La section du chromosome contenant le gène responsable de la formation de la capsule (gène S) pourrait être libérée des cellules S détruites et être absorbée par certaines cellules R dans les cultures subséquentes. Après que le gène S ait été incorporé dans son ADN, une cellule R a été transformée en une cellule S (4).

**L'information génétique est**

# transmise par l'ADN seul

La preuve finale que l'ADN, seul, et aucune autre molécule, transmet l'information génétique a été fournie par Hershey et Chase en 1952. Ils ont marqué la protéine capsulaire des bactériophages avec du soufre radioactif ( $^{35}\text{S}$ ) et l'ADN avec du phosphore radioactif ( $^{32}\text{P}$ ). Lorsque les bactéries ont été infectées par le bactériophage marqué, seuls les  $^{32}\text{P}$  (ADN) sont entrés dans les cellules, et non les  $^{35}\text{S}$  (protéine capsulaire). La formation ultérieure de nouvelles particules de phages complètes dans la cellule a prouvé que l'ADN était le vecteur exclusif de l'information génétique nécessaire pour former de nouvelles particules de phages, y compris leur protéine capsulaire. Ensuite, la structure et la fonction de l'ADN devaient être clarifiées. Les gènes de toutes les cellules et de certains virus sont constitués d'ADN, une molécule filiforme à longue chaîne.



---

## L'ADN recombinant et les enzymes de restriction

La technologie de l'ADN recombinant est un moyen extrêmement important dans la biotechnologie, cette technologie permet de construire n'importe quelle protéine en l'occurrence des anticorps, tout simplement par changement de la carte génétique. Deux enzymes naturelles clés sont utilisées sont

ADN ligase et les enzymes de restriction qui permettent de moduler l'information génétique dans un seul brin d'ADN. Avant cette découverte, on avait recouru à des produits chimiques ou de rayonnements ionisants qui réagissent de façon aléatoire, actuellement, on peut agir sur le code génétique précisément et à l'échelle atomique.

## Les enzymes de restriction

Une **enzyme de restriction** est une protéine qui est capable d'inciser un fragment de DNA au niveau d'une séquence nucléotidique spécifique appelée site de restriction. Chaque **enzyme de restriction** reconnaît ainsi un site spécifique

Les enzymes de restriction sont immensément utiles, certains poissons hébergent un grand nombre de bactéries pour lutter contre les infections virales. Les bactéries fabriquent des enzymes de restriction qui coupent l'ADN viral à un ordre spécifique, modifient bases en même temps, pour les protéger de l'action de la même enzyme.

Beaucoup d'enzymes de restriction coupent les deux brins d'ADN à la fois, au lieu de couper un seul. La biotechnologie intervient dans ce cas, les terminaisons du gène sont terminées par des extrémités adhésives à d'autres extrémités similaires, donc les enzymes de restriction peuvent être utilisées pour couper et créer de nouvelles régions adhésives au niveau des séquences cibles, qui sont ensuite collées ensemble selon les besoins.

Aujourd'hui, la technologie de l'ADN recombinant a développé, de nouvelles méthodes, nombreuses, ont été découvertes pour moduler les fonctions des bionanomachines cellulaires de façon continue. On peut extraire les gènes cibles à partir d'une cellule, en reproduire une quantité importante de copies, provoquer des mutations, recombinaisons, épissage, afin de créer entièrement de nouveaux gènes. En fin on peut remplacer des gènes cellulaires par ces gènes pour modifier son



information génétique.

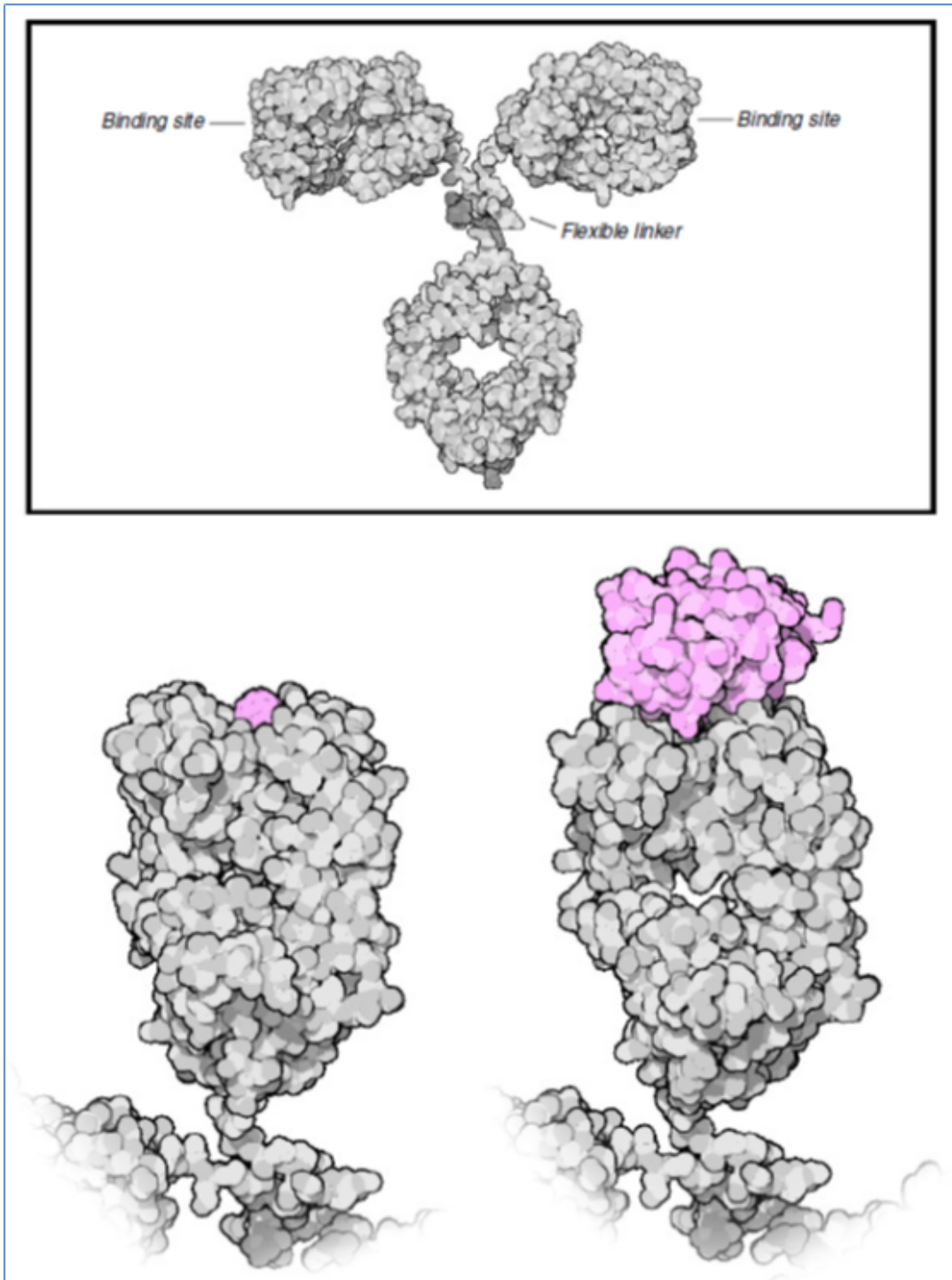


Figure. La technologie d'ADN recombinant est basée sur deux enzymes: E de restriction (EcoRI à gauche) et les ligases.

**Disponibilité des enzymes de**

# restriction

plus de 100 enzymes de restriction sont disponibles, chacune coupe l'ADN à un ordre spécifique de base, une ligase, une DNA polymérase.

Il existe des méthodes automatisées qui permettent la synthèse des brins d'ADN de 100 nucléotides de longueur. Une fois qu'une copie soit réalisée, elle passerait à la phase de reproduction: par clonage et réaction en chaîne de la polymérase, on emploie souvent ainsi les cellules bactériennes en raison de leur multiplication rapide, cette méthode consiste à incorporer la séquence voulue dans un bactériophage qui infecte la bactérie, la séquence génétique est insérée et on attend le résultat après de nombreuses divisions. Les plasmides à côté du matériel génétique bactérien, sont les plus employés. La PCR peut être employée aussi, elle tire profit de la polymérase pour produire plusieurs copies de la séquence génétique.

une fois que la séquence est réalisée, on a recourt à des vecteurs d'expression, les plasmides qui contiennent le gène codant pour une protéine possèdent une séquence dite promoteur ayant une grande spécificité. Le promoteur qui est souvent pris d'un virus, provoque la création de grandes quantités d'ARN messager, la cellule synthétiserait en l'occurrence, plusieurs protéines correspondantes. Les bactéries sont plus idéales à ces opérations, car peuvent créer un nombre plus important de ces protéines, et comprendre une quantité importante entre 1-10% de la totalité des protéines cellulaires. Il existe même des techniques de fermentation, peu coûteuses, permettant une croissance massive des bactéries avec des besoins nutritionnels minimes.

Les bactéries posent en fait quelques problèmes, elles ne disposent pas de matériel de maturation, cependant, de nombreuses cellules végétales possèdent à leurs surfaces des

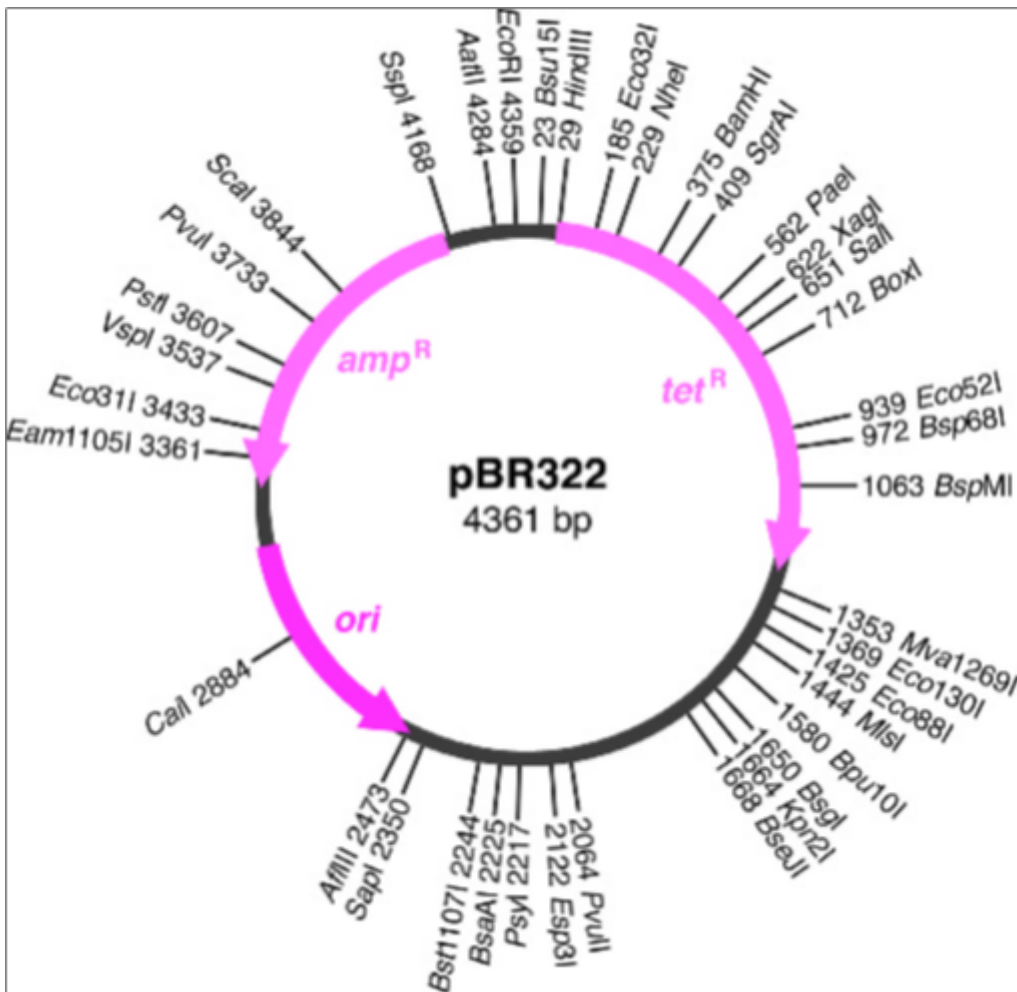
groupes d'hydrate de carbone, qui sont incorporées aux protéines à la fin de leur synthèse, les bactéries n'en possèdent pas (limitation sévère dans l'industrie pharmaceutique). Le système immunitaire peut dans certains cas réagir dangereusement avec les hydrates de carbone, attachées aux protéines cellulaires (système ABO).

En outre, les protéines ont tendance, dans certains moments, de former des agrégats spécifiques avec un arrangement soigneusement calculé, facilement mis en évidence par microscopie. Le manque de certains groupements peut être à l'origine d'un réarrangement aléatoire, et la formation des corps durs, la solubilisation de ces corps se trouvera difficile.

Il est tout à fait possible de renaturer les protéines des inclusions dures, ils sont les plus lourdes, alors une centrifugation des composants cellulaires peut aider à les extraire.

Les protéines peuvent être également construites sans avoir recouru à des cellules vivantes, cela est fait dans des éprouvettes spéciales. L'ARN qui se trouve dans les cytoplasmes cellulaires, et un excellent appui technique, en effet, l'extraction des composés cytoplasmiques contenant des bionanomachines de synthèse protéique, mais avec peu de chances de réussir à accomplir parfaitement la tâche, car le stock énergétique limité et la présence des protéases et de nucléases mettent rapidement fin à cette aventure. Des systèmes continus sans cellule, ont surmonté ce problème. Des tentatives de recréer des protéines avec des préparations purifiées ont réussi, mais avec un rendement modeste dû à la complexité de ces systèmes. En gros le développement des systèmes de production de protéines non dépendant de l'activité cellulaire sont l'objet des recherches.

# Carte de restriction : exemple du plasmide pBR322



Le plasmide pBR322 est l'un des vecteurs les plus employés sur l'E. coli. Il contient 4361 paires de bases d'ADN, et comprend des régions qui codent à diverses fonctions, *amp<sup>R</sup>*, *tet<sup>R</sup>*. En provoquant des coupures dans des régions spécifiques par les enzymes de restriction, on pourrait coller de nouveaux gènes, par exemple si le nouveau fragment d'ADN est ajouté à la région *PstI*, à la position 3607, il provoquerait des perturbations au niveau du gène de *-amp<sup>R</sup>*, la bactérie peut être facilement identifiée et séparée (sensible à la tétracycline).