

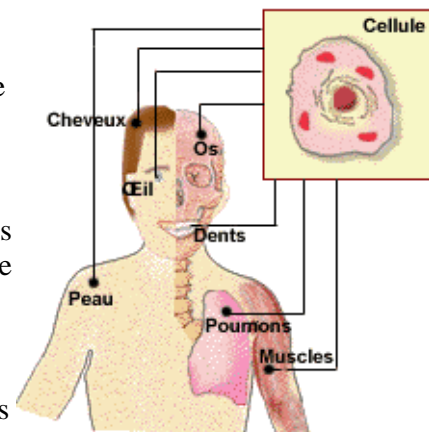
## A) Introduction générale

### 1) Qu'est-ce qu'une cellule?

*La cellule est l'unité du monde vivant* et les millions de types différents d'organismes qui peuplent la Terre ont tous un dénominateur commun : ils sont constitués de cellules. Elles sont très petites. Une bactérie, qui est formée d'une seule cellule, a généralement un diamètre d'un micron ( $\mu\text{m}$ ). Pour avoir une idée de la longueur d'un micron, imaginez que vous divisez la plus petite division d'une règle graduée (1 millimètre) en 1 000 segments égaux. Chacun de ces segments mesure un micron. Les êtres humains sont constitués de milliards de cellules, qui sont dix fois plus grosses qu'une cellule bactérienne ( $10 \mu\text{m}$ ), mais qui ne peuvent être observées à l'oeil nu. Il faut un microscope.

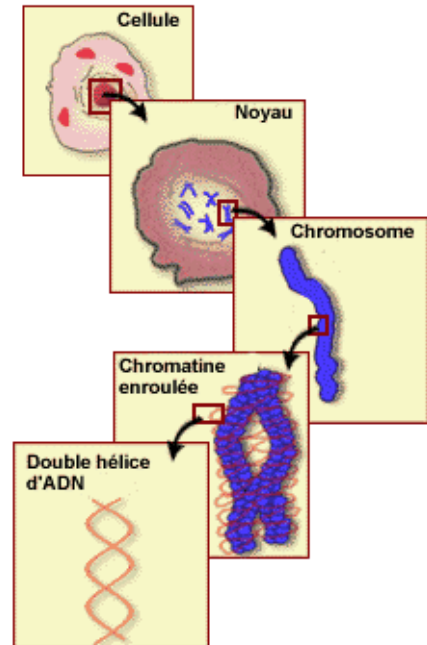
Les cellules appartenant à de grands organismes (comme les êtres humains) ne sont pas toutes pareilles. Ces cellules différentes ont une fonction et un aspect fort différents car elles se sont spécialisées dans certaines tâches très diverses. Par exemple, les cellules T, qui font partie de notre système immunitaire, sont très différentes de nos cellules nerveuses.

Les cellules T sont structurées de sorte à nous aider à combattre l'infection, tandis que les cellules nerveuses doivent pouvoir transmettre et recevoir les impulsions électriques qui caractérisent la fonction cérébrale. Mais toutes les cellules humaines ont quelque chose en commun, depuis les cellules du cerveau jusqu'à celles des yeux en passant par les cellules des muscles et de la peau : elles contiennent de l'ADN, le « patrimoine génétique ». Et chacune d'entre elles, peu importe sa forme ou sa fonction, contient une série complète d'ADN. C'est ce même ADN que l'on retrouve dans toute cellule du corps d'un individu.



## 2) Qu'est-ce que l'ADN?

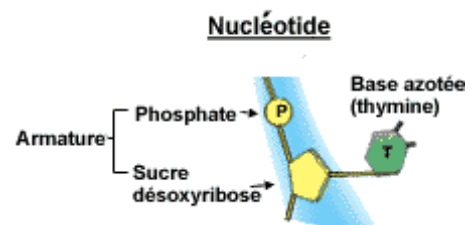
L'ADN, abréviation d'**acide désoxyribonucléique**, se trouve **dans le noyau** de la plupart des types de cellules. Il contient les instructions propres à la cellule et détermine comment les traits d'une personne seront transmis d'une génération à l'autre. Dans le noyau d'une cellule humaine, on compte 23 paires de chromosomes, soit 46 chromosomes en tout. **Chaque chromosome est formé de chromatine enroulée qui est composée d'ADN enrobant des protéines appelées histones.** Les 23 paires de chromosomes dans le noyau font office de « **manuel d'instructions** » pour le développement d'un individu. L'ADN détermine si la personne aura les yeux bleus ou bruns ou les cheveux foncés ou blonds. Comment les molécules (et les brins d'ADN sont des molécules) peuvent-elles donner des « instructions »? Pour le comprendre, examinons comment nous communiquons et comprenons les livres à partir desquels nous obtenons des informations.

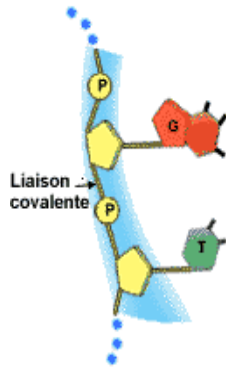


Au niveau le plus fondamental, nous ne comprenons les livres que s'ils utilisent un *code* que nous pouvons comprendre. Le code s'appelle par exemple *français ou anglais*. Vous comprendrez le texte que vous êtes en train de lire si vous comprenez chacun des mots de français ou d'anglais qu'il contient. Par ailleurs, les mots véhiculent rarement une information complète ou compréhensible. L'information est mieux communiquée si l'on regroupe des mots, et, comme vous le savez déjà, la série de mots qui véhicule une pensée complète s'appelle une phrase. Le langage de l'ADN, comme l'anglais et le français, comprend des mots et des phrases également! Chaque « mot » est une unité de la molécule d'ADN appelée **nucléotide**. Chaque « phrase » est une longue chaîne de nucléotides appelée **gène**. Mais parlons d'abord des mots de l'ADN.

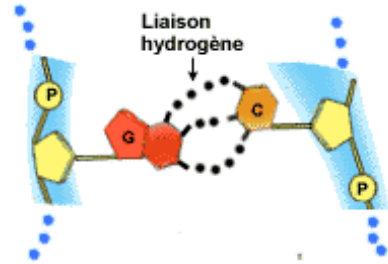
### Formation d'une molécule d'ADN : Nucléotides

Les « mots » de l'ADN sont de petites molécules appelées nucléotides. Le **génome humain, constitué de 23 paires de chromosomes**, contient au total quelque **trois milliards de nucléotides**. Chaque nucléotide comprend une armature et une **base azotée**. L'armature sert à attacher les nucléotides ensemble.



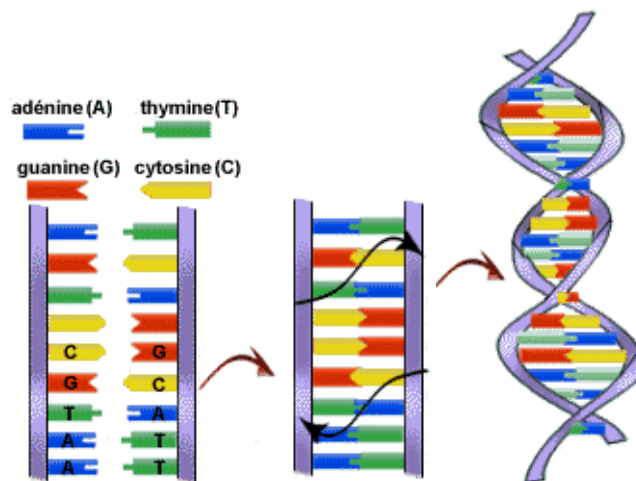


Il convient de parler d'un autre aspect important des nucléotides : l'adénine (A) se lie uniquement à la thymine (T) et la cytosine (C) uniquement à la guanine (G). De ce fait, on dit que A est **associée** à T et que C est **associée** à G. Il est bien plus facile de rompre les liaisons A-T et C-G (appelés **liaisons hydrogènes**) que de rompre des liaisons reliant ensemble l'armature des nucléotides dans la chaîne d'ADN (appelés **liaisons covalentes**). Cette propriété prendra de l'importance par la suite, lorsque nous parlerons de la **synthèse des protéines**.



Formation d'une molécule d'ADN : Comment les nucléotides se lient :

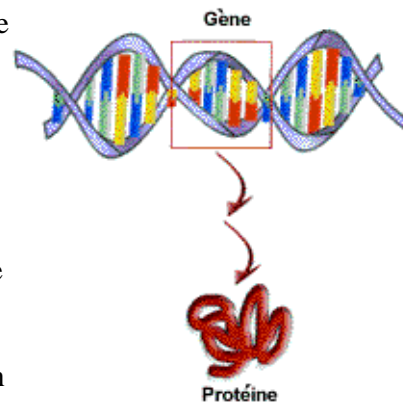
Les molécules d'ADN sont formées en réalité de deux chaînes parallèles de nucléotides. On dit que chaque chaîne est complémentaire de l'autre, car chaque nucléotide d'une chaîne se lie à son partenaire complémentaire de l'autre côté. Il serait utile de représenter la molécule d'ADN comme une échelle, dont les deux montants sont composés des armatures de nucléotides reliées entre elles, et dont les échelons sont les paires de base complémentaires A-T et G-C. Ainsi, si un côté de l'échelle a la séquence AATGC, le côté complémentaire aura la séquence TTACG. En réalité, l'échelle d'ADN est entortillée et forme une double hélice. Comme les liaisons hydrogènes (reliant G à C ou T à A) sont plus faibles que les liaisons covalentes reliant les nucléotides entre eux, les deux chaînes complémentaires formant l'échelle entortillée peuvent facilement être déroulées et séparées.



### 3) Gènes: les phrases de l'ADN

Un gène est une série de nucléotides (mots d'ADN) qui constitue une unité d'information héréditaire. Chacun des chromosomes dans le noyau d'une cellule humaine contient des milliers de gènes. **Tout l'ADN contenu sur les 23 paires de chromosomes dans une cellule humaine renferme les 80 000 gènes (et en conséquence, toutes les instructions génétiques) qui constituent le génome humain.**

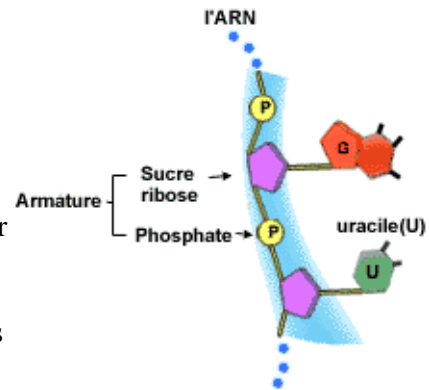
Mais qu'entend-on par « unité d'information héréditaire »? On peut donner une définition plus précise du gène : **région d'ADN qui est transcrite**. La transcription, processus biologique assuré par les enzymes constitue la première étape du processus de synthèse de protéines. Comme son nom l'indique, le processus de synthèse des protéines donne lieu à la production d'une protéine. Les protéines sont les molécules biologiques qui donnent aux cellules vivantes leurs formes et fonctions diverses. Ainsi, un gène est une séquence d'A, de T, de C et de G - dans un ordre particulier - qui code pour une fonction biochimique précise, en général par la production d'une protéine particulière. Ce sont les protéines produites à l'aide de gènes comme *matrice*, qui sont responsables des caractéristiques d'une cellule ou d'un organisme particulier.



Les gènes ont des tâches spécifiques, à des moments particuliers. **Tous les gènes ne sont pas « actifs » en même temps**. Par exemple, de nombreux gènes ne sont « actifs » (entraînant la production de protéines) que pendant le développement du fœtus humain. Une fois l'enfant né, les protéines associées à ces gènes ne sont plus requises, et ces derniers sont « désactivés », peut-être à jamais, sauf lorsqu'ils sont transmis aux générations suivantes. D'autres gènes ne sont « activés » que lorsque le corps en a besoin. Par exemple, le gène responsable de la production d'insuline est régi par la quantité de glucose (sucre) présente dans le sang d'une personne. Après la consommation d'un repas, le sang renferme une plus forte concentration de glucose, ce qui déclenche la production d'une quantité accrue d'insuline par les cellules du pancréas. L'insuline stimule l'absorption de glucose par des tissus comme les muscles squelettiques et les tissus adipeux (graisse), ce qui ramène la concentration de glucose dans le sang à un niveau normal. Une fois que le taux de glucose dans le sang est revenu à la normale, la quantité d'insuline sécrétée par les cellules du pancréas est également réduite.

## 4) Qu'est-ce que l'ARN?

L'ARN est l'abréviation d'**acide ribonucléique**. Tout comme l'ADN, les molécules d'ARN sont fabriquées dans le noyau de la cellule. Toutefois, contrairement à l'ADN, l'ARN ne se limite pas au noyau. Il peut migrer dans d'autres parties de la cellule. De l'ARN, appelé **ARN messager** communique le message génétique que l'on retrouve dans l'ADN au reste de la cellule afin de favoriser la synthèse de protéine. La façon dont une séquence de gènes dans l'ADN se traduit par la suite en une protéine correspondante fait l'objet de la section sur la synthèse des protéines.



Tout comme l'ADN, les molécules d'ARN sont composées de nucléotides. Toutefois, alors que les molécules d'ADN sont formées de deux brins parallèles de nucléotides, une molécule d'ARN n'en compte qu'un.

Par ailleurs, la structure chimique de l'« armature » **du nucléotide de l'ARN** est légèrement différente de la structure de celle de l'ADN. L'**armature de l'ADN** contient des molécules de sucre **désoxyribose** (d'où le D dans ADN) et **celui de l'ARN** des molécules de sucre **ribose** (d'où le R dans ARN).

Trois des quatre bases azotées pouvant se lier à l'armature de l'ARN sont les mêmes que celles pour l'ADN. Tout comme l'ADN, les nucléotides de l'ARN peuvent avoir des bases de guanine (G) et de cytosine (C) et tout comme dans les nucléotides d'ADN, le guanine s'associe (se lie) à la cytosine (G-C). Une troisième base que l'on retrouve dans l'ARN - adénine (A) - est également la même que dans l'ADN. Mais **au lieu de la thymine (T), l'ARN comporte de l'uracile (U)** qui se lie à l'adénine (U-A).

## 5) Les protéines

### À quoi servent les protéines?

De nombreux gènes codent pour des **chaînes polypeptidiques** particulières. Et les **protéines** comprennent une ou plusieurs chaînes polypeptidiques.

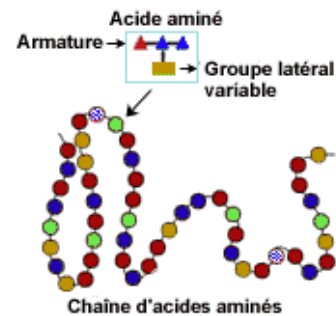
Ce sont les protéines qui sont responsables des caractéristiques d'un organisme ou d'une cellule. La façon dont les protéines sont construites, selon une matrice génétique, est décrite dans la section sur la synthèse des protéines.

Les protéines ont diverses vocations et donnent aux cellules vivantes leurs diverses formes et fonctions. Certaines protéines ont une **fonction structurale**; ces protéines fabriquent le cartilage, les cheveux et les ongles, par exemple. Une catégorie spéciale de protéines, les

enzymes, catalysent d'importantes réactions chimiques dans la cellule qui ne pourraient normalement se produire en leur absence. Certaines protéines servent de **canaux membranaires** qui facilitent le passage des particules moléculaires vers ou hors de la cellule. **Certaines hormones**, comme l'insuline, sont des protéines qui régulent les fonctions du corps (l'insuline contrôle le taux de sucre dans le sang). Les protéines sont également requises pour la **contraction des muscles** et aident les **cellules du corps à se défendre** contre les envahisseurs étrangers. Les fonctions susmentionnées ne constituent qu'une partie des diverses fonctions des protéines.

### Structure de la protéine

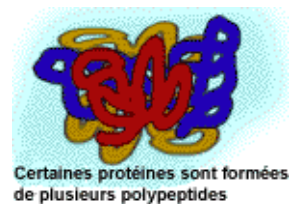
Toutes les protéines sont constituées d'une ou de plusieurs longues molécules appelées **polypeptides**. Chaque polypeptide est composé de petites molécules reliées bout à bout et appelées **acides aminés**. Les 20 types d'acides aminés utilisés par les cellules vivantes ont tous une structure d'armature identique, qui sert à lier ensemble les acides aminés en une longue chaîne. Chaque type d'acide aminé possède également ce qu'on appelle un **groupe latéral**, distinct sur le plan chimique, selon le type d'acide aminé. Bien qu'il ne soit pas nécessaire d'exposer en détail la façon dont varient les structures des groupes latéraux, mentionnons qu'ils peuvent être regroupés en plusieurs catégories.



Par exemple, certains groupes latéraux sont **non polaires**, tandis que d'autres sont **polaires**. Les molécules polaires et non polaires restent généralement éloignées les unes des autres. (*Avez-vous jamais remarqué que l'huile de cuisson et l'eau ne se mélangent pas, mais demeurent ensemble dans de grosses bulles? C'est justement parce que les molécules d'huile sont non polaires, et les molécules d'eau, polaires*) Les molécules d'eau sont polaires, et comme les molécules non polaires n'aiment pas s'associer à des molécules polaires, nous appelons souvent les molécules non polaires **hydrophobes** (du grec, « craignant l'eau »). Par ailleurs, les molécules polaires sont **hydrophiles** (du grec, « aimant l'eau »), car elles aiment interagir avec l'eau.



Chaîne d'acides aminés pliée - structure de protéines en trois dimensions (un seul polypeptide)



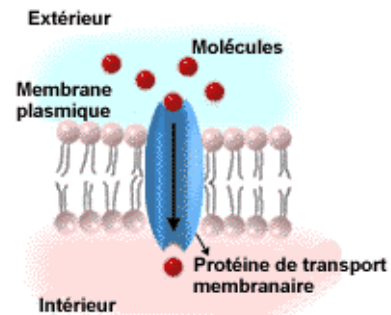
Certaines protéines sont formées de plusieurs polypeptides

**Les protéines, qui flottent dans la cellule ou n'importe où dans votre corps, sont entourées d'un milieu principalement aqueux.** Qu'arrive-t-il à la longue chaîne d'acides aminés, dont certains sont hydrophobes et d'autres hydrophiles?

**La protéine se plie en une structure en trois dimensions** où la plupart des acides aminés hydrophobes sont tournés vers l'intérieur de la structure (s'écartant de l'eau) et où la plupart des acides aminés hydrophiles se trouvent en surface, tournés vers l'eau. En conséquence, les

types d'acides aminés et l'ordre dans lequel ils se situent dans la chaîne détermineront comment la protéine finira par se plier dans l'eau, et, dès lors, sa structure en trois dimensions dans votre corps.

Cette structure tridimensionnelle est essentielle au bon fonctionnement de la protéine. Une **protéine de transport membranaire**, par exemple, est intégrée dans la membrane de la cellule et a la forme d'un tunnel ou d'un corridor reliant chaque côté de la membrane à l'autre. Elle a pour tâche de permettre à certaines molécules qui le peuvent d'entrer ou de sortir de la cellule. De toute évidence, la forme de la protéine de transport est très importante pour qu'elle puisse remplir correctement sa fonction! Une protéine de transport mal formée peut avoir un corridor « bloqué », ce qui signifie que les grosses molécules ne peuvent entrer ou sortir de la cellule. Les enzymes constituent un autre exemple de catégorie de protéines dont la forme est essentielle à leur bon fonctionnement.



## 6) Les enzymes

### Qu'est-ce qu'un enzyme?

Un enzyme est un *catalyseur biologique*. Mais qu'est-ce qu'un catalyseur?

Un *catalyseur* est une substance qui **accélère la vitesse d'une réaction biochimique** sans être altérée dans le processus. Des centaines de réactions chimiques différentes se produisent sans arrêt dans nos cellules et dans notre corps. Dans notre estomac et notre intestin grêle, des réactions chimiques décomposent les aliments que nous mangeons en particules plus petites qui peuvent être absorbées par nos cellules. Par exemple, une molécule de sucre complexe présente dans les produits laitiers, appelée lactose, doit d'abord être décomposée en deux molécules -- glucose et galactose -- avant de pouvoir être absorbée par les cellules somatiques. Cette réaction se produit normalement avec l'aide d'un enzyme appelé *lactase* présent dans l'intestin grêle.

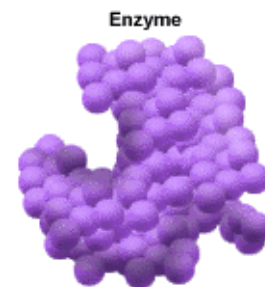
De nombreuses réactions chimiques, y compris la décomposition du lactose, ne se produisent pas spontanément. Le lactose ne se décompose pas s'il ne réside pas suffisamment longtemps dans du lait ou du fromage, car les molécules en cause doivent posséder une certaine quantité d'énergie pour que la réaction puisse se produire.



La quantité d'énergie requise pour qu'une réaction donnée ait lieu s'appelle *énergie d'activation*. Bien que la décomposition du lactose soit *possible*, elle ne se produira que si les molécules du lactose possèdent l'énergie requise. Dans le lait, le fromage ou la crème glacée, la molécule moyenne de lactose ne dispose tout simplement pas de l'énergie nécessaire pour subir la réaction de décomposition, ce qui fait qu'il faudrait attendre bien longtemps avant que cette réaction ne se produise par elle-même.

Comment les catalyseurs accélèrent-ils donc les réactions chimiques? Ils permettent à la réaction de se produire alors que *l'énergie d'activation est insuffisante*. En d'autres termes, en présence d'un catalyseur adéquat, les molécules réactantes auront besoin de moins d'énergie pour se transformer en produits. **Le catalyseur ne réagit pas lui-même et n'est pas altéré par la réaction. Les catalyseurs facilitent la réaction** en permettant à une série de molécules réactantes de se transformer en produits, pour ensuite aider d'autres molécules à subir la même réaction. Certains catalyseurs biologiques (enzymes) sont si efficaces qu'un seul suffit pour qu'à chaque seconde, plus de 600 000 molécules réactantes se convertissent en molécules de produit!

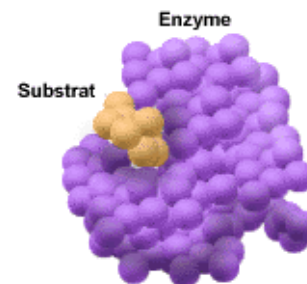
Il convient de noter que les enzymes sont très spécialisés. La *lactase* qui aide les molécules de lactose à se décomposer en molécules de galactose et de glucose, est structurée de sorte à ne pouvoir catalyser qu'un seul type de réaction. Les enzymes sont si sélectifs qu'ils ignorent les milliers de molécules dans les cellules somatiques et les fluides organiques pour lesquels ils ne sont pas conçus.



On appelle **substrat** la molécule qu'un enzyme aide à réagir. Ainsi, le lactose est le substrat de la lactase. Pour comprendre comment les enzymes peuvent être spécifiques, nous examinons la *structure de l'enzyme* et le *modèle clé-serrure* de la fonction enzymatique.

### Structure de l'enzyme : le « modèle clé-serrure »

À l'exception de quelques enzymes composés d'ARN, les enzymes sont des protéines. Souvenez-vous qu'une protéine est composée d'une ou de plusieurs chaînes d'**acides aminés** reliées, et que chaque chaîne d'acides aminés prend une forme tridimensionnelle selon la séquence d'acide aminé et la façon dont les acides aminés de la chaîne interagissent entre eux et avec la solution environnante.



Les enzymes se plient de telle façon qu'on observe une échancrure ou une poche à leur surface. On appelle cette poche *site actif*. Le **modèle clé-serrure** repose sur le principe selon lequel les formes des molécules réagissantes (les substrats) et le site actif de l'enzyme s'emboîtent comme une clé dans la serrure pour laquelle elle est conçue. Ainsi, la molécule de lactose s'adapte parfaitement au site actif de la lactase, ce qui signifie que cet enzyme peut uniquement catalyser la décomposition du lactose.

### Les enzymes sont importants!

Les enzymes produisent des centaines de réactions chimiques essentielles à notre survie. Ils



provoquent les réactions nécessaires à la digestion des aliments, à la formation et à la décomposition de l'ADN et de l'ARN, ainsi qu'à de nombreux autres processus vitaux. Une déficience en lactase, par exemple (relativement courante chez les êtres humains) rend impossible la décomposition du lactose et entraîne une intolérance au lactose, état qu'il est possible de surmonter en prenant des pilules contenant des enzymes de lactase avant de consommer des produits laitiers.

### **LISTE DES OUVRAGES DE RÉFÉRENCE**

1. S.S. Zumdahl, *Chemical Principles*, D.C. Health, Lexington, 1995
2. L.A. Moran, K. G. Scrimgeour, H. R. Horton, R. S. Ochs et J. D. Rawn.  
*Biochemistry*, Neil Patterson, New Jersey, 1994

## 7) Synthèse des protéines

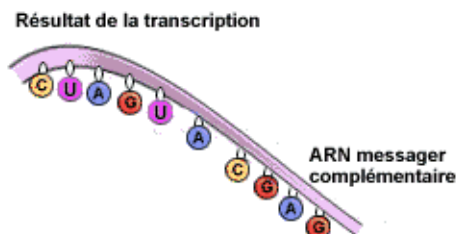
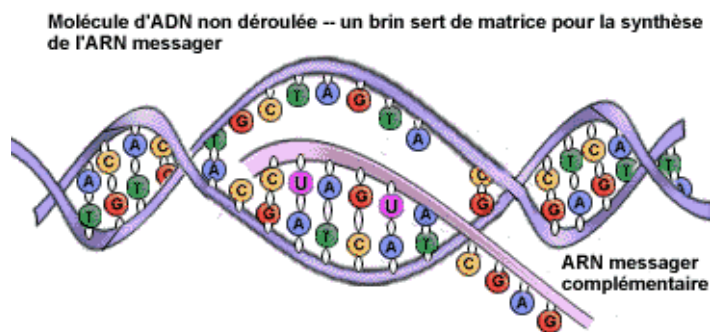
Voici un aperçu de la façon dont un gène (une section de la molécule d'ADN) sert de matrice pour la synthèse d'une protéine. Le processus peut être divisé en deux phases : la **Transcription**, suivie de la **Traduction**.

### Transcription :

La transcription est le processus par lequel un morceau particulier d'ARN - appelé **ARN messager (ARNm)** - est construit à l'aide d'une séquence génétique particulière (ADN) comme matrice.

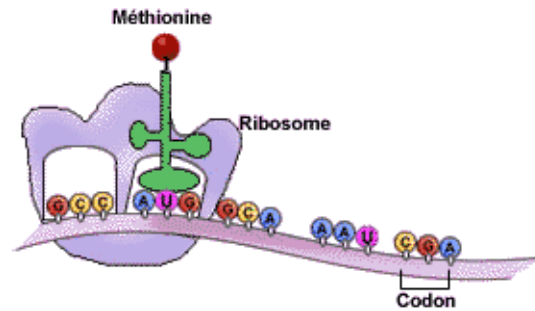
Tout d'abord, les enzymes déroulent une partie de l'hélice d'ADN bicaténaire et rompent les liaisons entre les paires de base complémentaires dans la section non déroulée. Ensuite, un brin complémentaire d'ARN messager est synthétisé, utilisant comme matrice l'un des brins de l'ADN non déroulés. Par exemple, si une partie d'un brin de l'ADN non défait lit GATCAT, la séquence d'ARN messager complémentaire lira CUAGUA. (Souvenez-vous que les nucléotides d'uracile [U] prennent la place de la thymine dans l'ARN). Enfin, une fois que l'ARN messager complémentaire est formé, le segment d'ADN reprend sa forme originale, soit celle d'une double hélice.

La transcription donne lieu à la création d'une molécule d'ARN messager complémentaire à une section donnée d'ADN (qui constitue un gène). **Contrairement aux molécules d'ADN, les molécules d'ARN messager sont libres de sortir du noyau par les pores de la membrane nucléaire pour voyager dans le reste de la cellule** (appelé cytosol). C'est dans le cytosol que prend place la **traduction**.

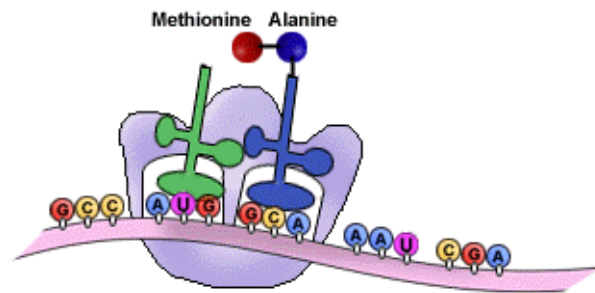


### Traduction :

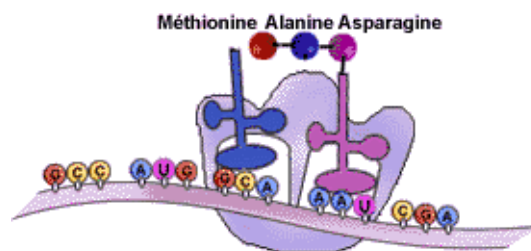
C'est le processus de fabrication d'une molécule, en fonction de l'information contenue dans une molécule d'ARN messenger. Mentionnons d'abord que comme l'ADN et l'ARN, les protéines sont des chaînes de petits éléments reliés entre eux. Dans le cas de l'ADN, ces petits éléments s'appellent nucléotides, et dans le cas des protéines, on les appelle acides aminés. La séquence de nucléotides dans l'ARN messenger est simplement transformée en une séquence d'acides aminés, selon un code uniforme.



Chaque séquence de trois bases d'ARN messenger code pour un acide aminé particulier. Par exemple, la séquence d'ARN messenger AUG code pour un acide aminé appelé méthionine. Les ribosomes, les « machines » qui assurent la synthèse des protéines, s'attachent au brin d'ARN messenger et descendent, « lisant » ainsi la séquence de nucléotides et reconstituant la protéine adéquate à mesure qu'ils se déplacent. La première série de trois nucléotides que lit le ribosome est toujours AUG, et ce parce que la séquence AUG sert de balise, indiquant au ribosome où il « doit commencer à lire ».



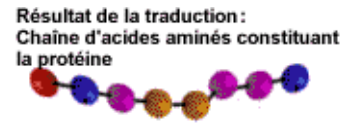
À mesure que le ribosome descend le long de l'ARN messenger, il ajoute l'acide aminé adéquat à la chaîne grossissante correspondant à chaque série de trois nucléotides. Chaque triplet de nucléotides qui code pour un acide aminé particulier s'appelle **codon**. Les vingt acides aminés employés pour fabriquer des protéines biologiques ont au moins un codon correspondant.



Par exemple, le codon CGA code pour un acide aminé appelé alanine. Et le codon AAU code pour un acide aminé appelé asparagine. En conséquence, une partie d'une séquence d'ARN

messager qui se lit AUG GCA AAU donnera lieu à la chaîne suivante d'acides aminés : méthionine-alanine-asparagine.

En lisant toute la séquence d'ARN messager, le ribosome construit une longue chaîne d'acides aminés, qui constituent la protéine.



## OUVRAGE DE RÉFÉRENCE

1. P. H. Raven et G. B. Johnson. *Biology*, troisième édition, Mosby Year Book, St. Louis, 1992, p. 308.

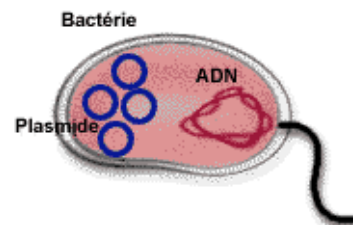
## 8) Les bactéries

Les bactéries sont des organismes unicellulaires. Bien que certaines bactéries soient un agent infectieux dans de nombreuses maladies humaines, il existe de nombreuses souches inoffensives voire même essentielles aux êtres humains. De nombreuses souches sont très importantes dans les laboratoires de biotechnologie! **Les cultures bactériennes** servent, entre autres, à la production de protéines utiles. Par exemple, une espèce de bactérie appelée *E. coli* peut être génétiquement modifiée de façon à produire d'importantes quantités d'insuline humaine, qui peut être administrée aux diabétiques. Voici un aperçu de certaines **caractéristiques des bactéries** qui font qu'elles conviennent parfaitement à de nombreuses applications biotechnologiques.



Photographie : Conseil national de recherches du Canada

Les bactéries sont des **procaryotes**, ce qui signifie qu'elles ne contiennent pas de noyau. Bien qu'ils prolifèrent souvent en groupes où les bactéries adhèrent l'une à l'autre, les procaryotes sont composés d'une seule cellule. On appelle ces groupes de bactéries des colonies. Le **génome bactérien** comprend un grande molécule circulaire d'ADN bicaténaire située dans le cytoplasme cellulaire. Cette **grande molécule d'ADN**, le chromosome bactérien, contient la plupart des gènes bactériens. Outre cette grande molécule d'ADN, les bactéries renferment souvent de **petites molécules d'ADN circulaires appelées plasmides**. Ces plasmides contiennent également des gènes, mais contrairement au grand chromosome circulaire, ils sont extrêmement mobiles. Ils peuvent passer facilement



d'une bactérie à l'autre, et, de cette façon, les gènes sont transmis entre bactéries. Les molécules de plasmide, une fois dans la cellule bactérienne hôte, peuvent s'intégrer en permanence au grand chromosome bactérien.

La **capacité des plasmides à pénétrer dans les cellules bactériennes** et à s'intégrer au chromosome de ces cellules en fait des **outils très utiles pour insérer un gène dans une cellule bactérienne**. Nous expliquerons comment les plasmides sont employés de cette façon dans la section portant sur le génie génétique.

## 9) Les virus

### Qu'est-ce qu'un virus?

Les virus sont constitués de matériel génétique (soit d'ADN soit d'ARN) entouré d'une couche protectrice de protéines. Certains virus d'animaux sont également entourés d'une membrane de lipides (gras). Un virus n'est pas un organisme vivant de manière autonome. Les virus n'existent que pour se multiplier, et à moins qu'un virus ne se trouve dans une cellule vivante, il est inactif et ne peut se reproduire. Lorsqu'un virus ou une partie de virus parvient à pénétrer dans une cellule, on parle d'**infection**.

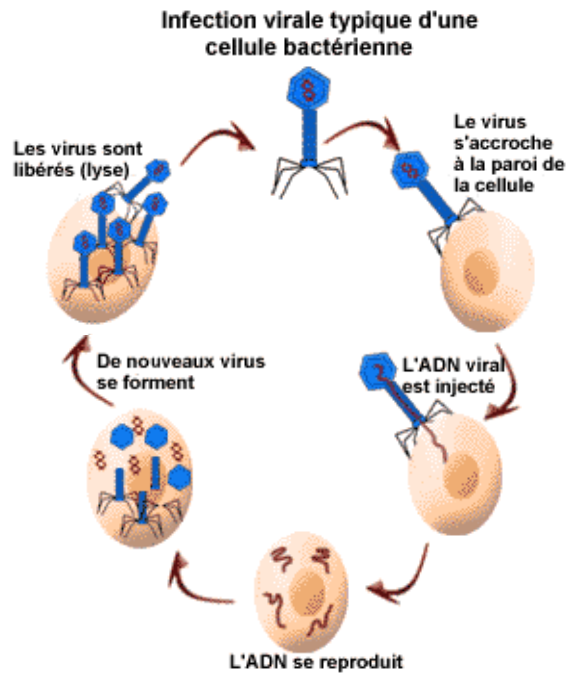
Selon le virus, c'est le virus tout entier qui pénètre dans la cellule ou seulement le matériel génétique qui est « injecté » dans la cellule tandis que la couche externe demeure à l'extérieur. Dans le cas du bactériophage T14 -- un type de virus qui infecte certaines bactéries --, l'ADN interne est injecté dans la cellule à infecter. En revanche, tout le virus du sida (appelé VIH) pénètre dans les cellules T de l'être humain pour les infecter.

Dans les deux cas, **par suite de l'infection virale, le matériel génétique du virus pénètre dans le cytoplasme de la cellule**, qui renferme tous les enzymes nécessaires et d'autres matériels indispensables à la reproduction du matériel génétique du virus et à la synthèse de ses protéines.

### Pourquoi une infection virale peut-elle nuire à une cellule?

Un virus nuit à la cellule qu'il infecte, car il « prend les commandes » du gène de la cellule et de la machine à fabriquer les protéines, ce qui donne lieu à la production de morceaux de virus uniquement. Une fois ceux-ci fabriqués, ils forment une myriade de nouveaux virus, qui remplissent la cellule.

Ces nouveaux virus quittent la cellule, quelques-uns à la fois (bourgeoisement) ou par un processus appelé lyse, où l'on assiste à une rupture de la membrane cellulaire, qui libère toutes les particules du virus en même temps, ce qui a pour effet de tuer la cellule hôte, tandis que les particules du virus libérées s'en vont infecter d'autres cellules.

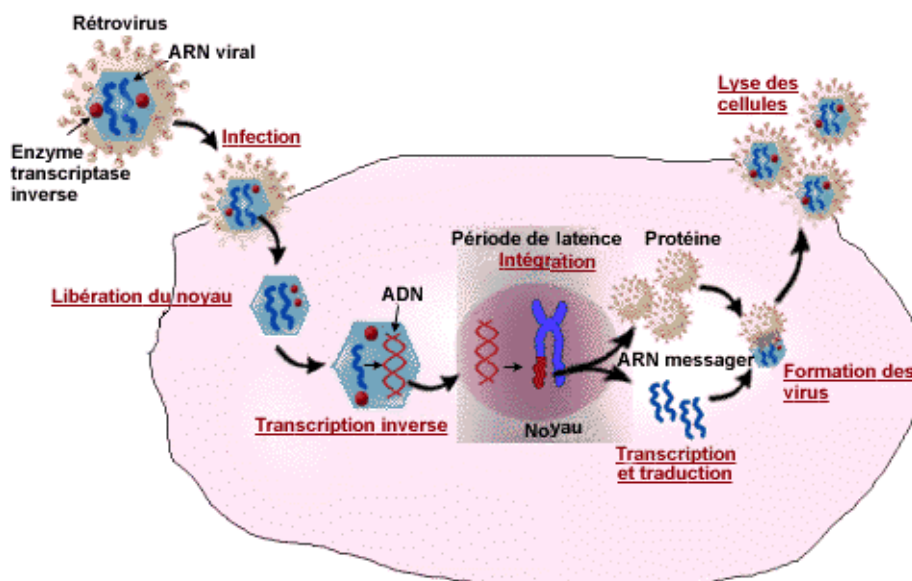


### Rétrovirus -- un type d'infection différent

Parfois, un virus ne prend pas les commandes de la machine à fabriquer des cellules dès qu'il l'infecte. Les rétrovirus, qui possèdent de l'ARN comme principal matériel génétique, portent également un enzyme spécial qui utilise l'ARN pour fabriquer une molécule d'ADN bicaténaire complémentaire. L'enzyme (connu sous le nom de **transcriptase inverse**) synthétise l'ADN à partir de l'ARN du virus, et cet ADN peut s'intégrer au génome de la cellule hôte situé dans le noyau.

Pendant une **période de latence**, les gènes viraux sont dormants dans les chromosomes de la cellule hôte. On pense que de nombreux segments du génome humain sont formés de **rétrovirus endogènes**, qui sont de l'ancien ADN défectueux d'un rétrovirus, qui s'est intégré il y a des milliers d'années et qui est présent, sans causer d'effets nocifs depuis. On pourrait dire que la « période de latence » de ces virus est pratiquement infinie ! Par ailleurs, le virus du sida (un rétrovirus) a une période de latence beaucoup plus courte (une moyenne d'environ 8 ans).

Après la période de latence, les gènes viraux seront activés et, selon le processus ordinaire de synthèse de protéines, ils prendront les commandes de la machinerie cellulaire, rendant viraux l'ARN et les protéines et entraînant la production de particules de virus.



Comme les rétrovirus parviennent bien à incorporer leur propre matériel génétique au génome de leur cellule hôte, ils **sont souvent utilisés comme vecteurs d'ADN recombinant**. En d'autres termes, si nous voulons intégrer un gène particulier qui a été isolé, mis au point ou modifié à l'aide du **génie génétique** (c'est ce qu'on entend par **recombinant**), dans le génome d'une cellule, nous ajoutons le gène dans l'ADN du rétrovirus, enlevons les parties nocives de l'ADN du rétrovirus qui provoquent la « prise des commandes » de la cellule, et **utilisons le rétrovirus pour transporter le gène voulu dans la cellule**. Lorsque nous permettons au virus modifié d'infecter la cellule hôte, l'ADN viral ainsi que le nouveau gène s'intègrent au génome de la cellule hôte. On trouvera une description plus complète de ce processus dans les sections sur le génie génétique et sur la thérapie génique.

## 10) Génie génétique

### QU'EST-CE QUE LE GÉNIE GÉNÉTIQUE?

Le génie génétique est le processus par lequel on identifie et isole l'ADN d'une cellule vivante ou morte pour l'introduire dans une autre cellule vivante. Avant d'introduire le matériel génétique, on peut le modifier en laboratoire. Lorsque la manipulation génétique réussit, le nouvel ADN est intégré à jamais dans les chromosomes de la nouvelle cellule, et apparaîtra également dans l'ADN des cellules des descendants. Comment les scientifiques peuvent-ils faire des manipulations génétiques? Ils utilisent la **technologie de recombinaison de l'ADN**.

### TECHNOLOGIE DE RECOMBINAISON DE L'ADN

On appelle technologie de recombinaison de l'ADN les méthodes mises au point pour isoler, manipuler, amplifier, couper et épisser des séquences identifiables d'ADN. Dans les trois prochaines sections, nous présentons plusieurs techniques de recombinaison de l'ADN employées pour localiser, isoler et amplifier l'ADN. Dans la dernière section, nous montrons comment utiliser d'autres techniques de recombinaison pour introduire du nouvel ADN dans

les cellules.

## 11) Localisation d'un gène

De nombreuses techniques en biologie moléculaire sont employées pour déterminer où se trouve un gène spécifique dans le génome humain. Cette tâche est loin d'être facile, puisque le génome humain contient des milliers de gènes, dont bon nombre n'ont pas encore été découverts ou séquencés. Les **sondes d'ADN** sont bien utiles à cette fin.

### Qu'est-ce qu'une sonde d'ADN?

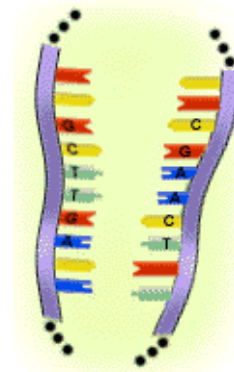
Il est possible de trouver un gène particulier à l'aide d'une **sonde d'ADN**, une molécule d'ADN monocaténaire relativement courte qui est complémentaire de la séquence sur le gène recherché. En d'autres termes, si un segment du gène recherché est connu comme étant AGTTCG, le segment complémentaire de la sonde d'ADN sera TCAAGC, car A se lie à T et C se lie à G. (En fait, en raison de certains détails structurels sur les molécules d'ADN, la séquence complémentaire s'écrirait en fait comme suit : CGAACT, ce qui donne TCAAGC à l'envers. Toutefois, nous utilisons le format techniquement incorrect ici, pour éviter toute confusion.)

Une véritable sonde d'ADN serait probablement constituée d'au moins quelques douzaines de nucléotides associés à un segment de la même longueur sur le gène recherché. La sonde est conçue de façon à être radioactive, de sorte qu'elle puisse être détectée facilement.

Comme la sonde d'ADN se lie à l'ADN monocaténaire, on a recours à une technique appelée **transfert de Southern** qui permet de séparer l'échantillon d'ADN bicaténaire en un seul brin et de le transférer à une membrane en nylon. Lorsque les sondes sont incubées avec la membrane dans une solution, elles se lient à des régions complémentaires dans l'ADN et « adhèrent » à la membrane. Ensuite, la membrane est mise en contact avec une pellicule photographique sensible

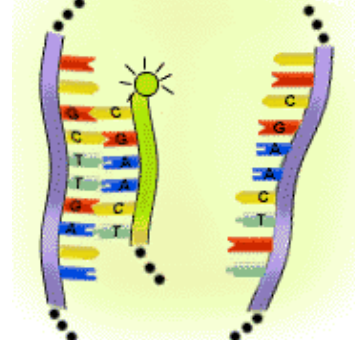


La technique du transfert de Southern sépare les brins d'ADN et les transfère sur le nylon



Ajout d'une sonde

Le sonde d'ADN radioactif se lie uniquement au gène complémentaire recherché





aux émissions radioactives. Les sections de l'échantillon où se trouve le gène ressortent en foncé sur le papier, car ce sont les seules sections liées à une sonde radioactive.

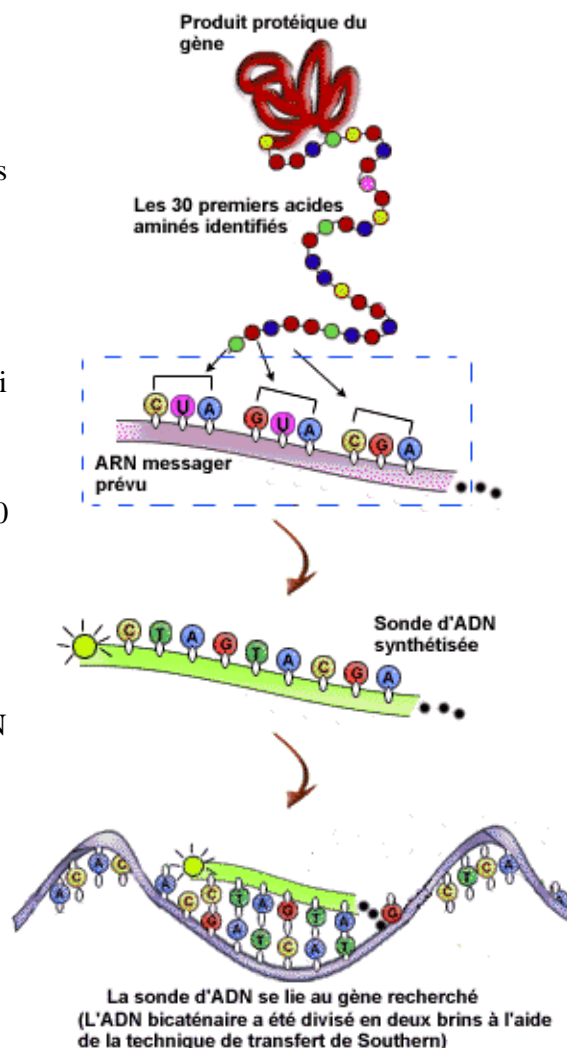
Les sondes d'ADN ont plusieurs applications en biologie moléculaire, y compris les méthodes génétiques de prédiction et de diagnostic des maladies, où l'on emploie des sondes qui identifient les séquences de gènes que l'on sait responsables de maladies. Les sondes d'ADN peuvent également être employées pour les empreintes génétiques, technique à laquelle les experts en criminalistique ont souvent recours pour déterminer si l'ADN retrouvé sur les lieux d'un crime correspond à celui du suspect.

### Comment construit-on les sondes d'ADN?

Il est possible de construire une sonde d'ADN bien avant de connaître la séquence des gènes elle-même! Pour ce faire, on travaille à partir du produit protéique du gène. Rappelez-vous que dans notre description de la fabrication des protéines, nous avons dit qu'un gène était transcrit en ARN messager (ARNm), selon les simples règles de la **complémentarité des bases**. L'ARN messager est transporté hors du noyau et utilisé comme matrice pour la formation d'une chaîne d'acides aminés, qui se transforme en une protéine.

Nous pouvons isoler la protéine produite par le gène qui nous intéresse, et trouver quels sont les 30 premiers acides aminés de la protéine. Selon cette information, nous pouvons déterminer les 90 premiers nucléotides de cette matrice d'ARN messager de la protéine (n'oubliez pas que chaque triplet de nucléotides code pour un acide aminé, d'où le rapport 90:30). Et comme la matrice d'ARN messager est complémentaire du gène recherché, nous savons que notre sonde d'ADN devrait avoir une séquence complémentaire des 90 premiers nucléotides du gène recherché.

Pour construire une sonde d'ADN, nous utilisons une « machine à gènes » capable de synthétiser en quelques heures à peine une molécule courte d'ADN monocaténaire contenant la séquence voulue de nucléotides.

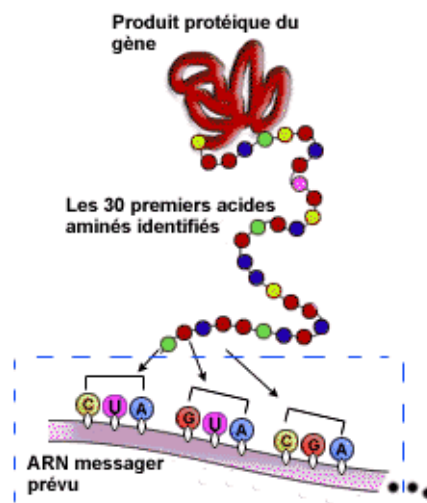


## 12) Isolement d'un gène

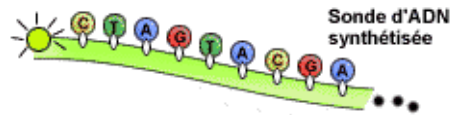
Si nous voulions introduire un gène humain dans une autre cellule, il ne nous suffirait pas de savoir où se trouve ce gène dans le génome humain. Nous devrions également isoler une copie du gène, de façon à pouvoir l'insérer dans la nouvelle cellule.

Par exemple, le gène de l'insuline humaine doit être isolé des cellules humaines de sorte à pouvoir être introduit dans les cellules de la bactérie *E. coli*. Par suite de l'incorporation du gène, les cellules bactériennes produisent la protéine de l'insuline humaine que l'on peut administrer aux diabétiques.

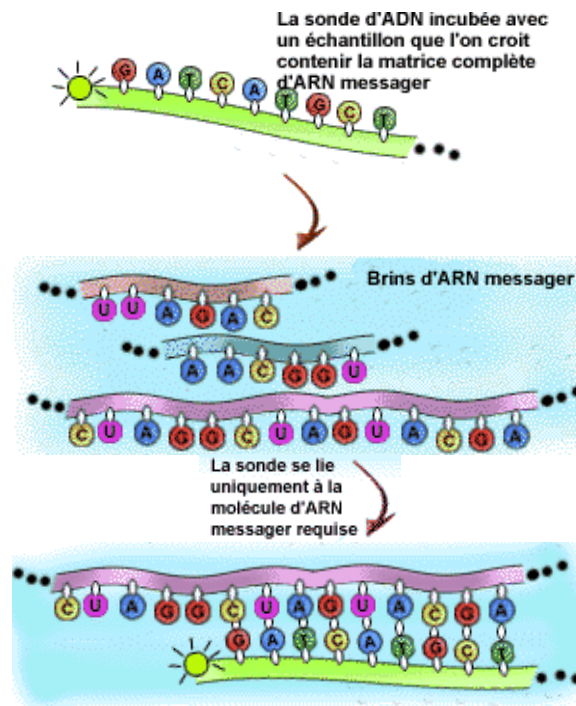
Pour isoler un gène, on peut travailler à rebours à partir de son produit protéique. Tout d'abord, au moins une partie de la protéine est *séquencée*, ce qui signifie que l'ordre des acides aminés qui constituent la chaîne protéique est déterminé. En général, il suffit de connaître les 30 premiers acides aminés de la protéine. Ensuite, selon la séquence connue d'acides aminés et si l'on comprend le processus de synthèse des protéines, on peut prédire la séquence de nucléotides d'une partie de la matrice d'ARN messenger de la protéine.



Ensuite, on construit une **sonde d'ADN** monocaténaire *complémentaire* de la séquence prévue d'ARN messenger. Par exemple, si une partie de la séquence prévue d'ARN messenger est CUA GUA CGA, la section correspondante de la sonde d'ADN serait GAT CAT GCT, car G s'associe à C et A à T. (En fait, en raison de certains détails structurels relatifs aux molécules d'ADN, la séquence complémentaire devrait réellement s'écrire comme suit : TCG TAC TAG, qui correspond à GAT CAT GCT à l'envers. Toutefois, nous utilisons un format techniquement incorrect ici, pour éviter la confusion.) La sonde d'ADN est conçue pour être radioactive, de sorte qu'elle soit décelable lorsqu'elle se lie à son « image symétrique » d'ADN.

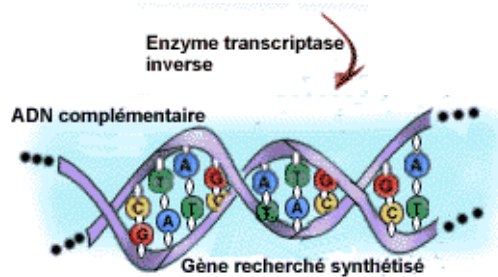


La sonde d'ADN monocaténaire est ensuite incubée avec un échantillon censé contenir la matrice complète de la protéine d'ARN messenger. Lorsque nous isolerons l'ARN messenger auquel se lie la sonde d'ADN, nous aurons probablement trouvé l'ARN messenger que nous cherchions.



Une fois que l'on a trouvé le brin d'ARN messenger, il suffit de travailler à rebours -- nous devons synthétiser le brin d'ADN qui aurait servi de matrice pour l'ARN messenger. Il nous faut synthétiser de l'ADN à partir d'ARN. Il existe un enzyme appelé **transcriptase inverse** qui nous permet de faire cela. On retrouve cet enzyme dans certains virus appelés **rétrovirus**. Ces virus emploient l'ARN comme matériel génétique et utilisent la transcriptase inverse pour générer de l'ADN une fois qu'ils ont infecté une cellule hôte. Les biotechnologistes peuvent mélanger une transcriptase inverse avec des ARN messagers *in vitro* (à l'extérieur de cellules vivantes, en laboratoire, généralement dans un petit tube en plastique). Par conséquent, la séquence d'ADN pour le gène recherché est synthétisée par l'enzyme, selon le brin d'ARN messenger présenté. Comme l'ADN produit a été fabriqué de manière artificielle en vue d'être complémentaire de l'ARN messenger, on l'appelle **ADN complémentaire**.





### 13) Amplification de l'ADN : amplification en chaîne par la polymérase

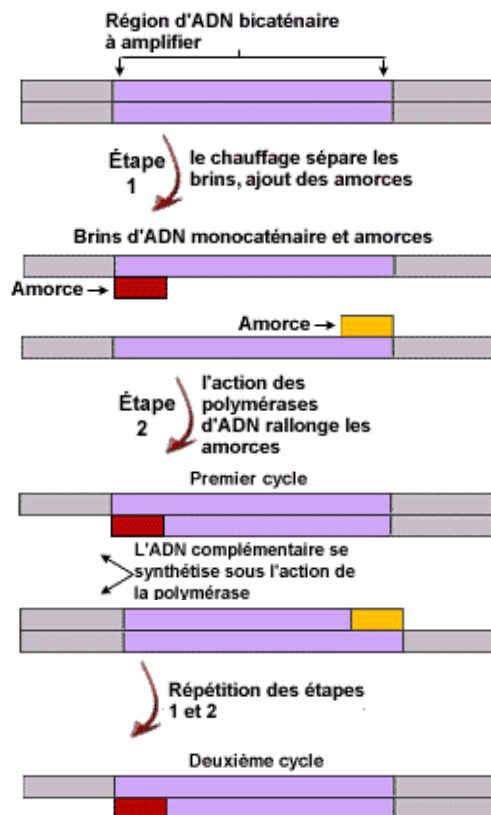
Souvent, la quantité des échantillons d'ADN est trop petite pour qu'on puisse les utiliser. Heureusement, on peut avoir recours à une technique inventée dans les années 1980, l'**amplification en chaîne par la polymérase (ACP)** pour « amplifier » les quantités d'ADN de ces échantillons.

La machine d'ACP n'est en fait rien de plus qu'un dispositif très précis de chauffage et de refroidissement. La machine comporte de petites fentes où sont insérés de petits tubes contenant l'échantillon d'ADN et d'autres ingrédients nécessaires à la réaction. Ces ingrédients supplémentaires englobent une bonne quantité de nucléotides (A, T, C et G), de courtes molécules d'ADN monocaténaire appelées **amorces** et un enzyme appelé **polymérase *Thermus aquaticus* (polymérase Taq, en abrégé)**. La polymérase Taq est dérivée de bactéries qui vivent dans des sources chaudes et comptent parmi les rares enzymes capables de fonctionner à de très hautes températures.

Le cycle commence lorsque la machine chauffe le tube à une température d'environ 90-95 °C, ce qui entraîne la séparation de chaque molécule d'ADN bicaténaire de l'échantillon d'origine en deux brins.

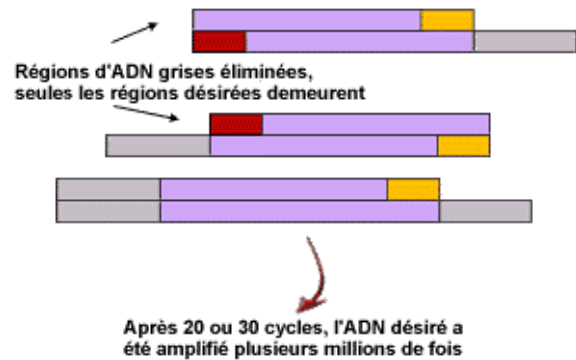
(Souvenez-vous que les *liaisons hydrogènes* qui relient deux brins complémentaires d'une double hélice d'ADN sont bien plus faibles que les liaisons covalentes qui relient les nucléotides formant chaque chaîne. Le chauffage provoque la rupture des liaisons hydrogènes, qui déroulent et séparent les deux brins, tandis que les liaisons covalentes ne sont pas touchées.)

Ensuite, on abaisse la température légèrement, ce qui permet aux amorces d'ADN de se lier aux brins séparés. Les amorces se lient, car elles sont *complémentaires* de certaines séquences de chaque brin d'ADN qui « flanquent » l'ADN à répliquer au milieu. Une fois que les amorces se sont attachées aux brins, la polymérase Taq se synthétise, en utilisant les nucléotides flottant dans le tube, un brin complémentaire pour chaque



brin monocaténaire d'origine. Cette réaction boucle le **premier cycle**, et double la quantité d'ADN présente dans le tube.

Au cours du cycle suivant, la machine d'ACP chauffe et refroidit comme auparavant, ce qui entraîne la séparation des nouvelles molécules d'ADN bicaténaire, et la synthèse des nouveaux brins complémentaires par la polymérase Taq.



**Chaque fois qu'un cycle est complété, la quantité de copies de la séquence d'ADN désirée (située entre les deux amorces) est, en théorie, multipliée par deux.** Après une trentaine de cycles (qui durent généralement environ trois heures), on disposera de suffisamment de copies de cette séquence d'ADN pour l'application d'autres techniques biotechnologiques.

## 14) Introduction d'un gène dans une cellule

Un gène, probablement isolé sous forme d'ADN complémentaire, peut être introduit dans une cellule à l'aide d'un vecteur. Un vecteur est un véhicule par lequel de l'ADN étranger est transmis d'une cellule à une autre.

Par exemple, les virus modifiés et les plasmides sont des vecteurs.

### Les virus en tant que vecteurs

Les virus constituent d'excellents vecteurs, car au terme de longues périodes d'évolution, ils ont acquis la capacité d'éviter d'être détruits par le système immunitaire humain et d'insérer leur propre matériel génétique dans certaines cellules. Comme nous l'avons vu dans la section portant sur les virus, une infection virale se produit lorsque du matériel génétique étranger (viral) pénètre dans la cellule et emploie la machine à fabriquer des protéines et de l'acide nucléique de la cellule pour produire ses propres protéines ainsi que son propre ADN et ARN. Pour utiliser un virus comme vecteur, on remplace les segments nocifs de son ADN par l'ADN complémentaire que l'on souhaite introduire dans la cellule. Ensuite, on laisse le virus « infecter » la cellule hôte et, si tout va bien, l'ADN complémentaire pénètre dans la cellule et servira à fabriquer la protéine voulue.

Certains virus peuvent produire leur propre ADN et l'insérer dans le génome de la cellule hôte. Ces **rétrovirus** à base d'ARN sont les vecteurs viraux les plus courants utilisés en thérapie génique où les gènes ayant une valeur thérapeutique sont introduits dans les rétrovirus qui, au moment de l'infection, les insèrent dans le génome de la cellule hôte.

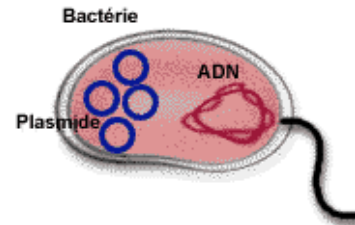
Il convient de noter que les virus employés comme vecteurs sont conçus de sorte à ne pouvoir se reproduire. En d'autres termes, les



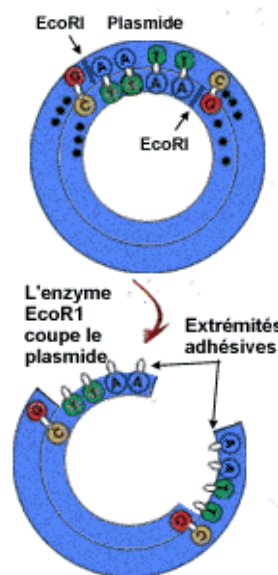
segments nocifs du génome viral qui servent à produire un nombre accru de particules virales ont été enlevés et remplacés par une séquence qui code pour la protéine recherchée.

### Les plasmides en tant que vecteurs

L'ADN complémentaire de l'insuline humaine est introduit dans des cellules bactériennes à l'aide d'un plasmide. Un plasmide est simplement une molécule d'ADN circulaire contenant des gènes qui peuvent facilement se répandre dans des cellules bactériennes ou en sortir. Bien que les plasmides soient présents naturellement dans certaines bactéries, ceux employés en vue d'introduire un gène étranger dans une cellule et de l'en extraire ont été modifiés de telle sorte que les séquences qu'ils renferment sont fort différentes des plasmides présents à l'état naturel desquels ils sont tirés.



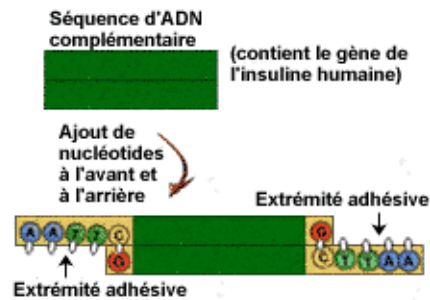
En premier lieu, le plasmide contient plusieurs séquences courtes spécialisées appelées *sites de restriction*. Les enzymes appelés **endonucléases de restriction** reconnaissent ces sites et coupent l'ADN du plasmide. Par exemple, un enzyme de restriction appelé *EcoRI* reconnaît la séquence GAATTC et procède à la coupure entre le G et le premier A. Il convient de noter que la séquence complémentaire est CTTAAG, soit GAATTC à l'envers! Ainsi, l'enzyme coupe les deux brins de plasmide comme suit :



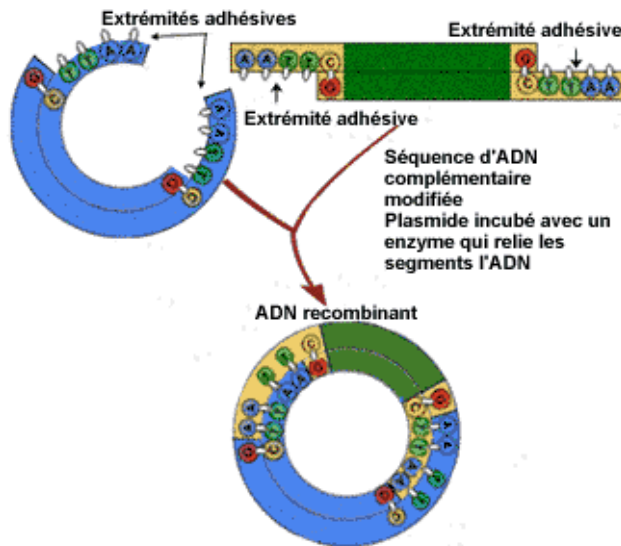
Mentionnons que la découpe à l'aide de l'*EcoRI* donne lieu à deux « extrémités adhésives » qui consistent en un seul brin de nucléotides qui se lieront à une série complémentaire d'« extrémités adhésives » monocaténaïres. Qui plus est, le plasmide est conçu de sorte que cette séquence particulière de restriction n'est présente qu'*une seule fois*, coupera le plasmide bicaténaire en un seul site.

L'ADN complémentaire (qui renferme le gène de l'insuline humaine) à insérer dans le plasmide est modifié selon l'enzyme de restriction employé pour couper le plasmide. Pour revenir à notre exemple avec l'*EcoRI*, il faudrait les extrémités adhésives suivantes à chaque extrémité de

l'ADN complémentaire :



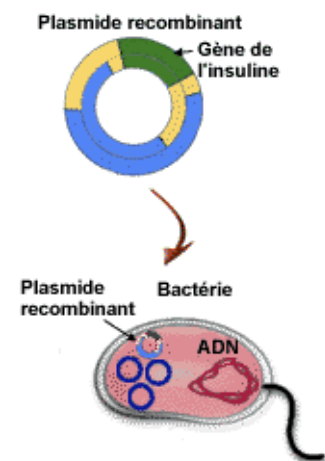
Maintenant, nous incubons la séquence d'ADN complémentaire modifiée dans une solution renfermant un plasmide coupé à l'aide de l'enzyme EcoR1 et un enzyme qui relie les segments d'ADN (appelé **ADN ligase**). On obtient alors un plasmide circulaire fermé qui renferme l'ADN complémentaire et, par là même, le gène de l'insuline humaine.



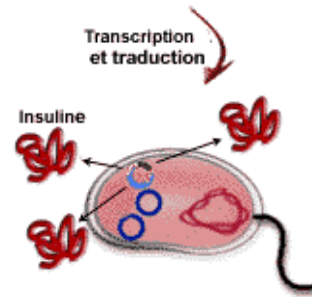
Le plasmide est ensuite incubé avec des cellules bactériennes (dans le cas du processus de fabrication de l'insuline, l'espèce de bactérie employée est *E. coli*) dans des conditions particulières qui favorisent l'absorption du plasmide par les cellules bactériennes.

En théorie, le plasmide contenant le gène de l'insuline humaine pénétrera dans toutes les cellules bactériennes et toutes ces cellules transcriront la protéine et produiront de l'insuline humaine, qui pourra ensuite être récupérée et utilisée pour traiter les patients diabétiques.

Malheureusement, les cellules bactériennes n'absorberont pas toutes le plasmide. En fait, dans la plupart des cas, elles seront peu nombreuses à le faire. Comment les biotechnologues peuvent-ils uniquement sélectionner les cellules bactériennes qui ont absorbé le plasmide?



Ils le peuvent en fonction des conditions de la culture bactérienne et d'une autre modification spéciale intégrée dans les plasmides issus du génie génétique. Les bactéries, après avoir été incubées en présence du plasmide (et certaines l'ont absorbé), sont cultivées dans un milieu qui renferme un antibiotique, comme l'ampicilline. L'ampicilline tuera la bactérie *E. Coli*, à moins qu'elle ne soit protégée d'une façon ou d'une autre. Le plasmide que les bactéries ont absorbé contient également un gène qui confère une résistance à l'ampicilline. En conséquence, seules les bactéries ayant absorbé le plasmide résisteront à l'antibiotique et survivront. Comme le plasmide renferme également le gène de l'insuline humaine, nous ne permettons qu'aux bactéries capables de produire de l'insuline de survivre et de se reproduire.



### OUVRAGES DE RÉFÉRENCE

1. P. H. Raven et G.B. Johnson. *Biology*, troisième édition, Mosby Year Book, St. Louis, 1992
2. W. Bains. *Biotechnology: From A to Z*, Oxford, New York, 1993
3. L.A. Moran et K.G. Scrimgeour, *Biochemistry Resource Book*, Neil Patterson, New Jersey, 1994
4. L. A. Moran et K. G. Scrimgeour, H.R. Horton, R.S. Ochs et J.D. Rawn. *Biochemistry*, Neil Patterson, New Jersey, 1994
5. J. L. Southin, *Forty-Eight Lectures in Modern Biology*, Kendall/Hunt, Dubuque, 1993
6. E. S. Grace, *Biotechnology Unzipped: Promises & Realities*, Trifolium, Toronto, 1997