

Université Populaire de Marseille, printemps 2019
Cycle Hérité génétique et épigénétique
Jacques van Helden

Chapitre 2

Les mécanismes moléculaires de l'hérité

- Héritéité génétique
 - Comment une cellule ou un individu transmettent-ils leurs caractéristiques à leur descendance ?
- Mutations
 - Comment les caractéristiques génétiques se modifient-elles ?
- Mécanismes moléculaires
 - Comment les molécules du vivant mettent-elles en œuvre ces caractéristiques héritées ?
- Adaptation phénotypique
 - Comment une cellule s'adapte-t-elle à des modifications de son environnement ?
- Différentiation
 - Comment les cellules au génome identique adoptent-elles des phénotypes différents ?
- Morphogenèse
 - Comment les cellules s'organisent-elles pendant le développement embryonnaire ?

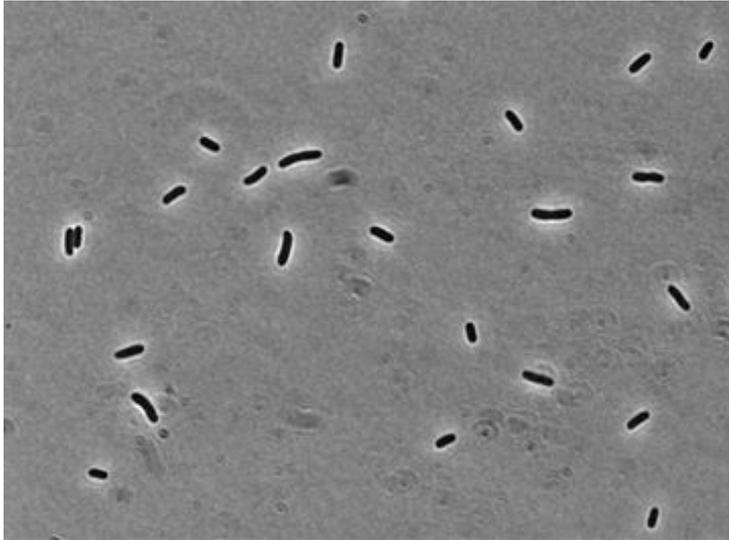
Les mêmes questions en termes d'information

- Hérité
 - Où réside l'information transmise à la descendance d'une cellule ou d'un individu ?
- Mutations
 - Comment cette information se modifie-t-elle au fil des générations ?
- Mécanismes moléculaires
 - Comment l'information se traduit-elle en molécules effectrices (protéines, protéines) à partir de son support héréditaire (ADN) ?
- Adaptation phénotypique
 - Comment l'information est-elle modulée par les changements de l'environnement (de la cellule, de l'organisme) ?
- Différentiation
 - Qu'est-ce qui régit l'exploitation différentielle de l'information héréditaire dans les différents types de cellules ?
- Morphogenèse
 - Comment l'information passe-t-elle du niveau d'une cellule à celui d'un organisme (où réside le « plan » de l'organisme) ?

Organisation cellulaire

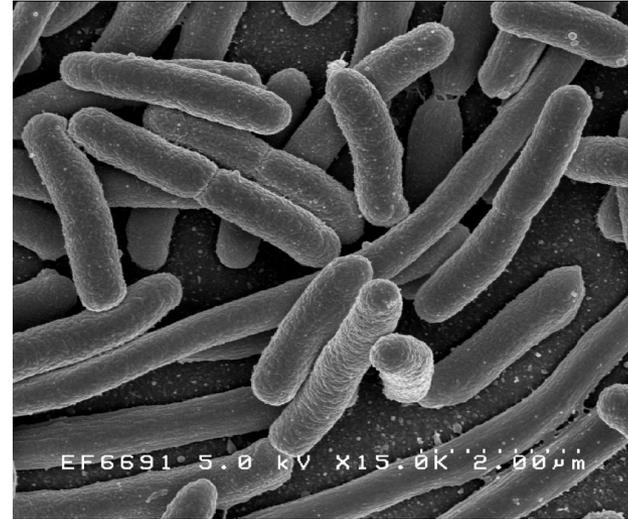
Cellule procaryote : l'entérobactérie *Escherichia coli*

Microscopie optique (grossissement 1000 fois)



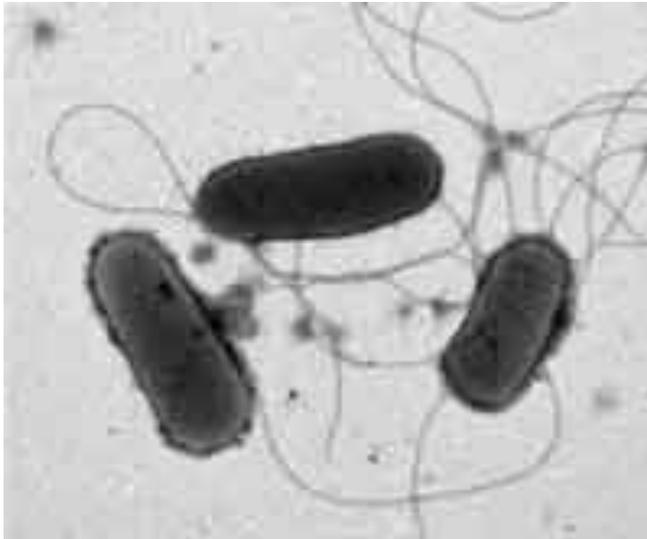
http://parts.igem.org/Part:BBa_K257021:Experience

Microscopie électronique à balayage



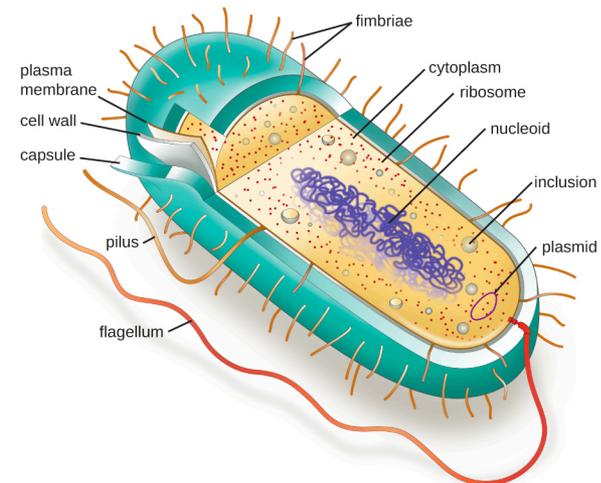
https://upload.wikimedia.org/wikipedia/commons/3/32/EscherichiaColi_NIAID.jpg

Microscopie électronique à transmission



<https://globalfoodsafetyresource.com/escherichia-coli/>

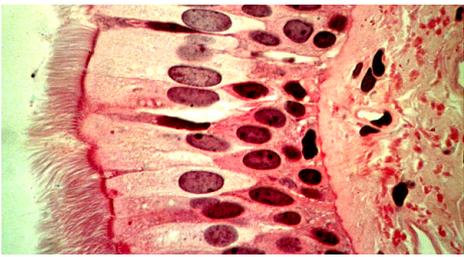
Représentation schématique d'une cellule procaryote



[https://bio.libretexts.org/Bookshelves/Microbiology/Book%3A_Microbiology_\(OpenStax\)/03%3A_The_Cell/3.3%3A_Unique_Characteristics_of_Prokaryotic_Cells](https://bio.libretexts.org/Bookshelves/Microbiology/Book%3A_Microbiology_(OpenStax)/03%3A_The_Cell/3.3%3A_Unique_Characteristics_of_Prokaryotic_Cells)

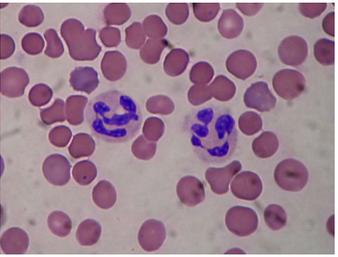
Cellule eucaryote

Cellules d'épithélium animal



<http://www.anatomybox.com/tag/light-microscopy/page/11/>

Cellules sanguines



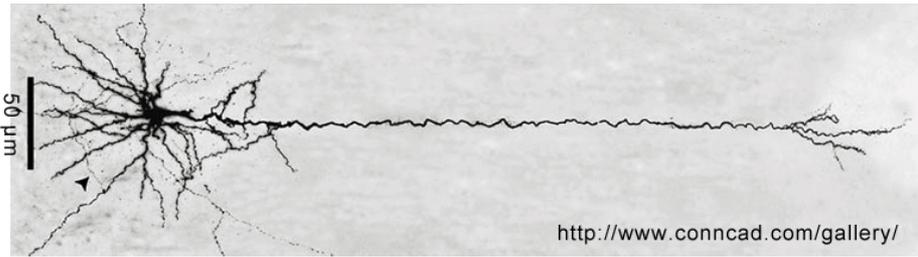
<https://www.futura-sciences.com/sante/definitions/medecine-globule-blanc-733/>

Cellules musculaires



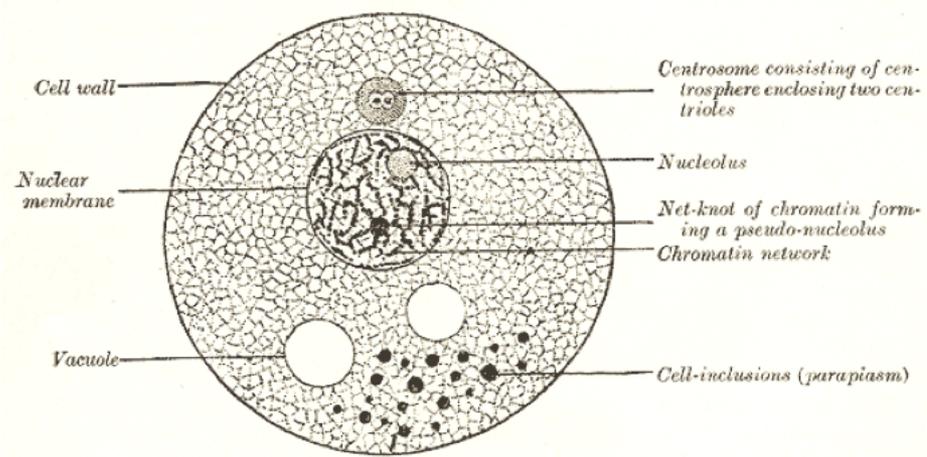
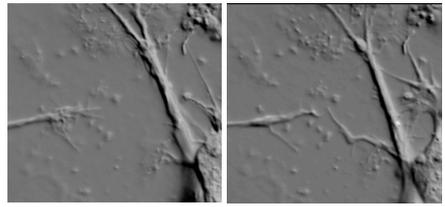
<https://planet-vie.ens.fr/article/1901/differences-muscle-squelettique-muscle-cardiaque>

Neurone



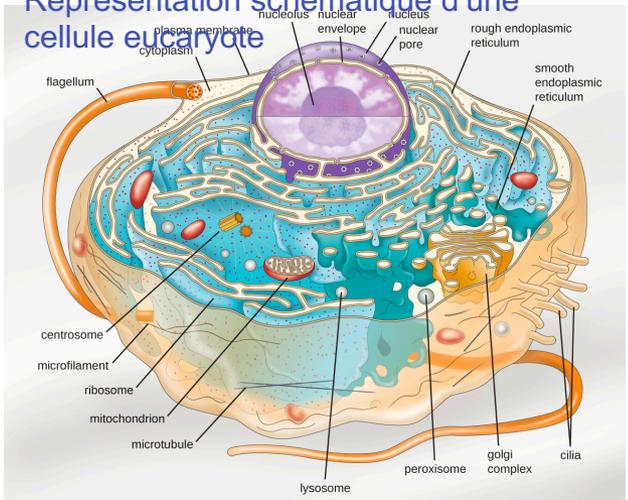
<http://www.conncad.com/gallery/>

Etablissement d'une jonction synaptique entre deux neurones



Henry Gray (1825–1861). Anatomy of the Human Body. 1918. Diagram of a cell. (Modified from Wilson.) <https://www.bartleby.com/107/illus1.html>

Représentation schématique d'une cellule eucaryote



[https://bio.libretexts.org/Bookshelves/Microbiology/Book%3A_Microbiology_\(OpenStax\)/03%3A_The_Cell/3.4%3A_Unique_Characteristics_of_Eukaryotic_Cells](https://bio.libretexts.org/Bookshelves/Microbiology/Book%3A_Microbiology_(OpenStax)/03%3A_The_Cell/3.4%3A_Unique_Characteristics_of_Eukaryotic_Cells)

Hérédité génétique

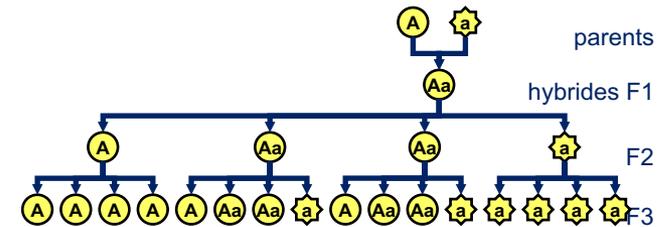
Les lois de l'hérédité de Mendel

- Gregor Mendel (1822-1884) réalise des expériences de croisement entre variétés de pois, et dénombre les différents phénotypes au fil de la descendance.
- Il en dérive trois « lois », qui synthétisent les régularités observées dans les résultats.

1. **Uniformité des hybrides** de première génération.
2. **Ségrégation des caractères** au fil des générations suivantes.
3. **Indépendance** des caractères.

- Mendel définit les concepts de **dominance** et **récessivité**.
- Ces lois s'appliquent également à d'autres organismes, mais seulement sous certaines **conditions**.
 - La ségrégation des caractères repose sur le mode de reproduction autogame après le premier croisement (hybrides).
 - Pour certaines paires de caractères, on observe des écarts à la loi d'indépendance. C'est sur base de ces écarts que Thomas. Hunt Morgan s'appuiera pour fonder la théorie chromosomique de l'hérédité (1915).
- Les lois de Mendel furent ignorées de son vivant, et « re-découvertes » en 1900 par trois chercheurs : Hugo De Vries (Pays-Bas), Carl Correns (Allemagne) et Erich Tschermak (Autriche).

	A	a	hybrides (=F1)	deuxième génération (=F2)		
				A	a	A/a
Forme de la graine à maturité	ronde	irrégulière	253	5,474	1,850	2.96
Couleur de l'albumen	jaune	vert	258	6,022	2,001	3.01
Couleur extérieure de la graine	gris-brun	blanche		705	224	3.15
Forme de la cosse à maturité	enflée	contractée		882	299	2.95
Couleur de la cosse avant maturité	verte	jaune		428	152	2.82
Position des fleurs	axiales	terminales		651	207	3.14
Longueur de la tige	6 pieds	1 pied		787	277	2.84



Génération après hybrides	Nom de la génération	A	Aa	a	A	Aa	a
0 (hybrides)	F1	0	1	0	0	1	0
1	F2	1	2	1	1	2	1
2	F3	6	4	6	3	2	3
3	F4	28	8	28	7	2	7
4	F5	120	16	120	15	2	15
5	F6	496	32	496	31	2	31
n					2ⁿ⁻¹	2	2ⁿ⁻¹

	♂ gamètes			
	R Y $\frac{1}{4}$	R y $\frac{1}{4}$	r Y $\frac{1}{4}$	
♀ gamètes	R Y $\frac{1}{4}$ RR YY $\frac{1}{16}$ RR Yy $\frac{1}{16}$ Rr Yy $\frac{1}{16}$ Rr yy $\frac{1}{16}$	R y $\frac{1}{4}$ RR Yy $\frac{1}{16}$ RR yy $\frac{1}{16}$ Rr Yy $\frac{1}{16}$ Rr yy $\frac{1}{16}$	r Y $\frac{1}{4}$ Rr Yy $\frac{1}{16}$ Rr yy $\frac{1}{16}$ rr Yy $\frac{1}{16}$ rr yy $\frac{1}{16}$	r y $\frac{1}{4}$ Rr Yy $\frac{1}{16}$ Rr yy $\frac{1}{16}$ rr Yy $\frac{1}{16}$ rr yy $\frac{1}{16}$
9 : 3 : 3 : 1				
Round, yellow Wrinkled, yellow Round, green Wrinkled, green				

Verts 1/4 Jaunes 3/4
Ridés 1/4 Lisses 3/4

Verts et ridés: $1/16 = 1/4 * 1/4$
 Verts et lisses: $3/16 = 1/4 * 3/4$
 Jaunes et ridés: $3/16 = 3/4 * 1/4$
 Jaunes et lisses: $9/16 = 3/4 * 3/4$

- En 1901, Hugo de Vries publie *La théorie des mutations*
 - Il y décrit l'apparition de nouveaux caractères dans une espèce de plante (l'onagre de Lamarck, *Oenothera lamarckiana*).
 - Il observe qu'après leur apparition, ces caractères se transmettent selon les lois de Mendel.
 - Selon Hugo de Vries, les mutations sont le principal mécanisme de l'évolution.
- Les découvertes combinées de Gregor Mendel (lois de l'hérédité) et Hugo de Vries (mutations) fournissent les maillons manquants de la théorie darwinienne de l'évolution : apparition (mutations) et transmission (hérédité) des «variations» au sein des individus d'une population.



Théorie chromosomique de l'hérédité mendélienne

Drosophile de « type sauvage »



http://www.steve.gb.com/images/science/drosophila_melanogaster.jpg

Mutant white



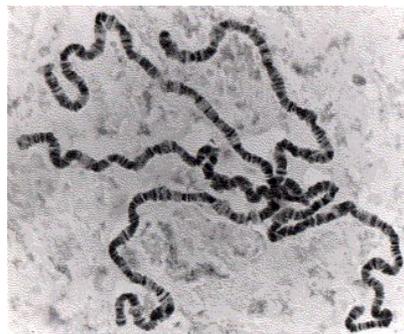
http://www.scq.ubc.ca/quarterly012/white_drosophila.gif

Mutant curly



<https://www.biologie.uni-halle.de/entwicklungsgenetik/lehre/studenten/drosophila/mutanten/>

Double mutant curly + white

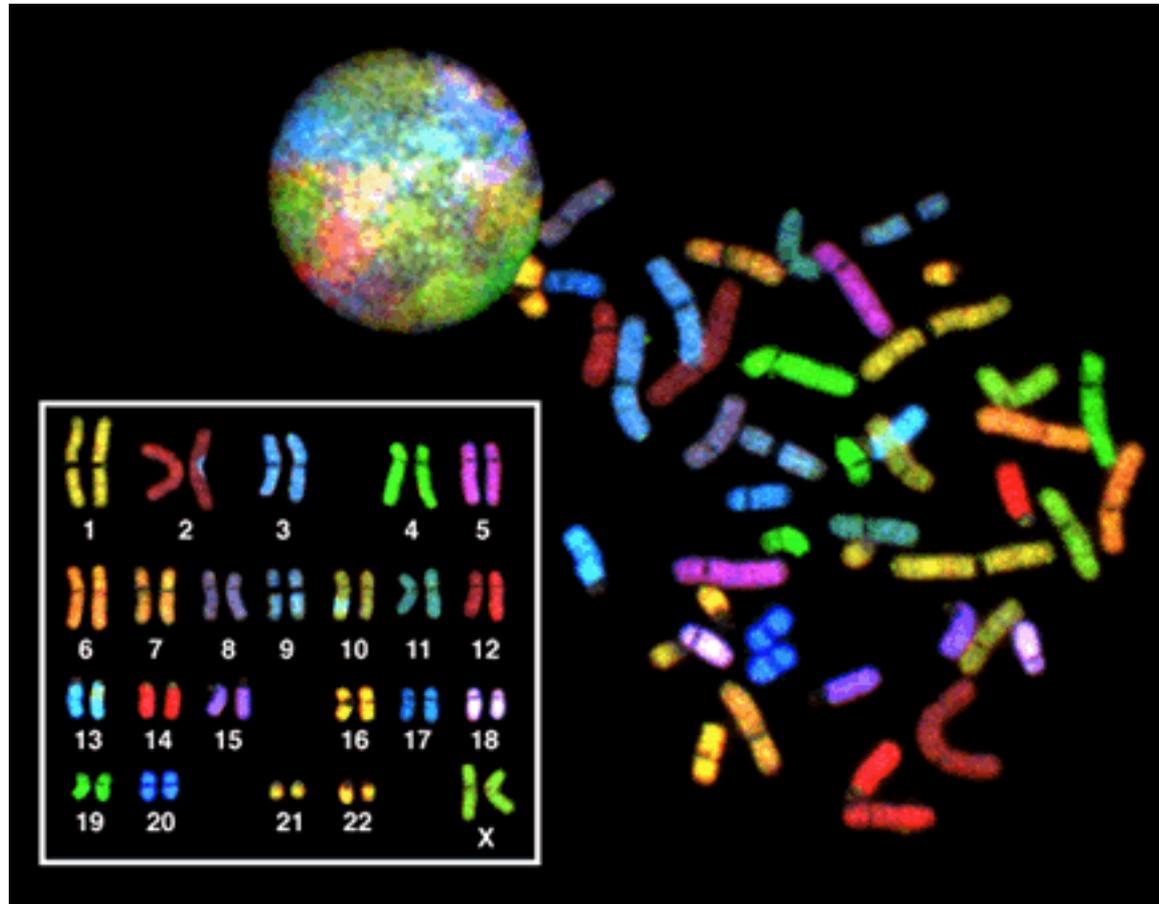


<http://users.rcn.com/kimball.ma.ultranet/BiologyPages/P/Polytene.jpg>

- Thomas H. Morgan (1866-1945) choisit comme organisme modèle la mouche à vinaigre *Drosophila melanogaster*.
- Isole des **mutants**, notamment certaines mouches aux yeux blancs (le type sauvage a les yeux rouges).
- Etudie la transmission de caractères à la descendance.
- Certains caractères sont transmis de façon dépendante du sexe.
 - En 1910, Morgan émet l'hypothèse que ces caractères sont portés par le chromosome X.
- Certaines paires de caractères sont transmises de façon corrélée.
 - Il introduit la notion de *liaison génétique*. Il met en évidence 4 groupes de liaison.
- il formule la *théorie chromosomique de l'hérédité* (*Mécanismes de l'hérédité Mendélienne*, 1915).
 - Les 4 groupes de liaison génétique de la drosophile correspondent aux 4 chromosomes.
 - Les chromosomes sont porteurs des caractères transmis de façon héréditaire.
 - Sur chaque chromosome, les gènes sont ordonnés de façon linéaire.

Caryotype en couleurs

- Caryotype :
 - formule décrivant le nombre de copies de chaque chromosome dans une cellule;
 - image de microscopie permettant de caractériser ces nombres.
- Principe: les cellules sont bloquées en métaphase (qui contient la forme compacte des chromosomes) et écrasées sur une lame de microscope puis colorées.
- A partir de la photo initiale (photo de droite) on ordonne les chromosomes (encart de gauche).
- Le **caryotype spectral** est une préparation microscopique où chaque chromosome est marqué par une couleur spécifique.



https://en.wikipedia.org/wiki/Karyotype#/media/File:Sky_spectral_karyotype.png

Un cristal apériodique

- En 1944, Erwin Schrödinger publie un petit livre intitulé « Qu'est-ce que la vie ? », où il formule les grandes questions que pose la biologie de son époque, en les interprétant du point de vue de physicien.
 - *Pour donner vie et couleur à [ma]déclaration, laissez-moi anticiper ce qui sera expliqué plus tard de façon beaucoup plus détaillées, en l'occurrence que la partie la plus essentielle d'une cellule vivante – la fibre chromosomique – pourrait à juste titre être qualifiée de cristal apériodique.*
- *Il souligne la différence de complexité qui distingue les cristaux périodiques et apériodiques*
 - *En physique, nous avons jusqu'à présent traité uniquement de cristaux périodiques. Pour l'esprit d'un humble physicien, il s'agit d'objets très intéressants et compliqués; ils constituent l'une des structures matérielles les plus fascinantes et complexes par lesquelles la nature déconcerte son intelligence. Pourtant, en comparaison avec les cristaux apériodiques, ils sont plutôt simples et ternes. La différence de structure est du même ordre que celle entre un papier-peint ordinaire, dans lequel le même motif se répète sans cesse selon une périodicité régulière, et un chef d'œuvre de broderie, par exemple une tapisserie de Raphaël, qui ne révèle aucune répétition terne, mais un trait élaboré, cohérent, riche de sens, tracé par un grand maître.*

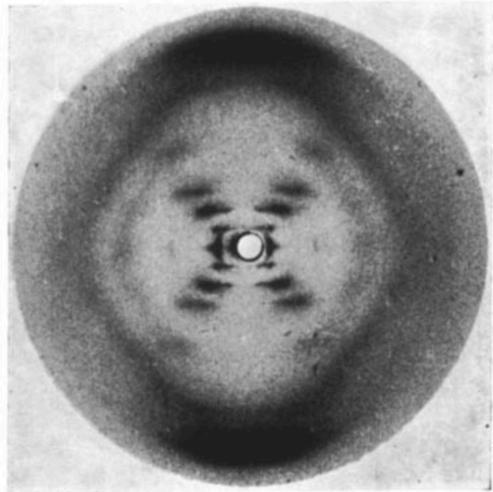
Papier peint



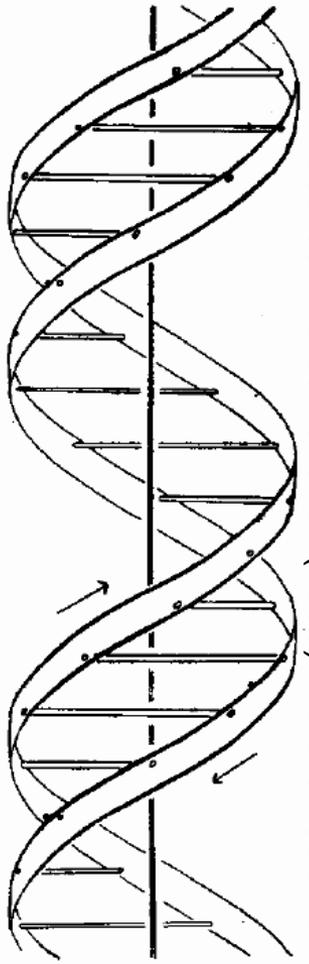
Raphaël – La pêche miraculeuse.

Structure de l'ADN - la double hélice

- En 1953, Watson et Crick proposent un modèle pour la structure B de l'ADN, inspiré par la structure cristallographique caractérisée par Rosalind Franklin.
- L'ADN est une double hélice, dont chacun des deux "montants" est formé d'une chaîne de désoxyribose (un sucre) unis par des groupes phosphate.
- Chaque "barreau" est formé par une paire de nucléotides liés par des ponts hydrogènes.
 - Une Guanine se trouve toujours face à une Cytosine (3 ponts hydrogène).
 - Une Adénine se trouve toujours face à une Thymine (2 ponts hydrogène).

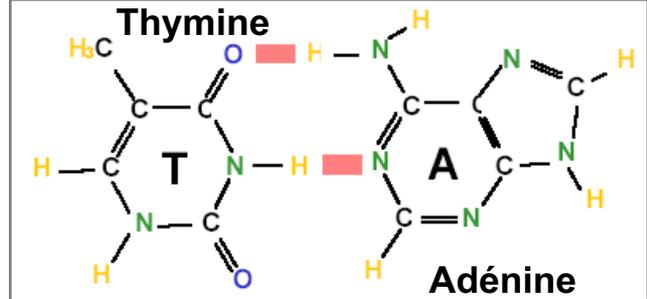
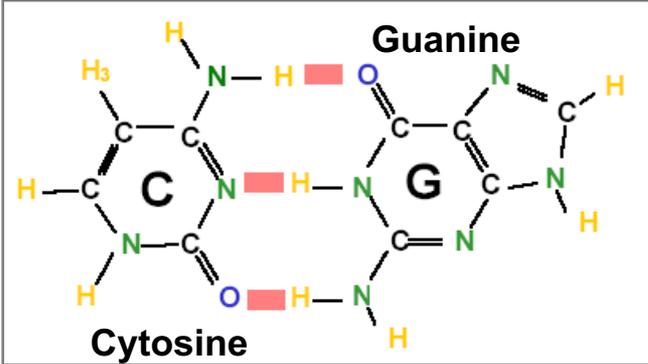


Sodium deoxyribose nucleate from calf thymus. Structure B
X-Ray fiber diagram of deoxyribonucleic acid (R.E. Franklin and R. Gosling, 1953)



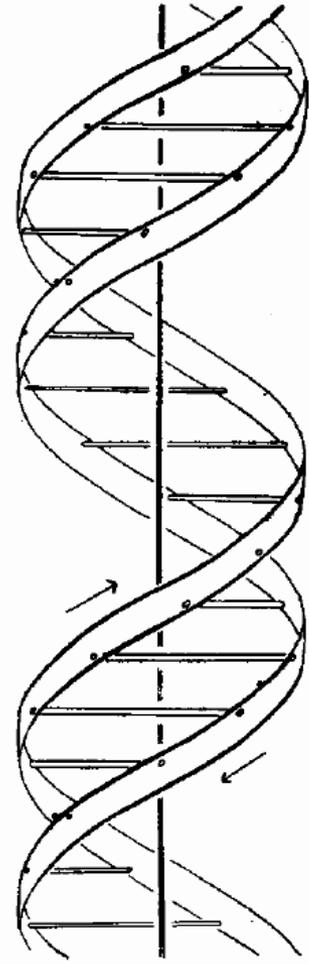
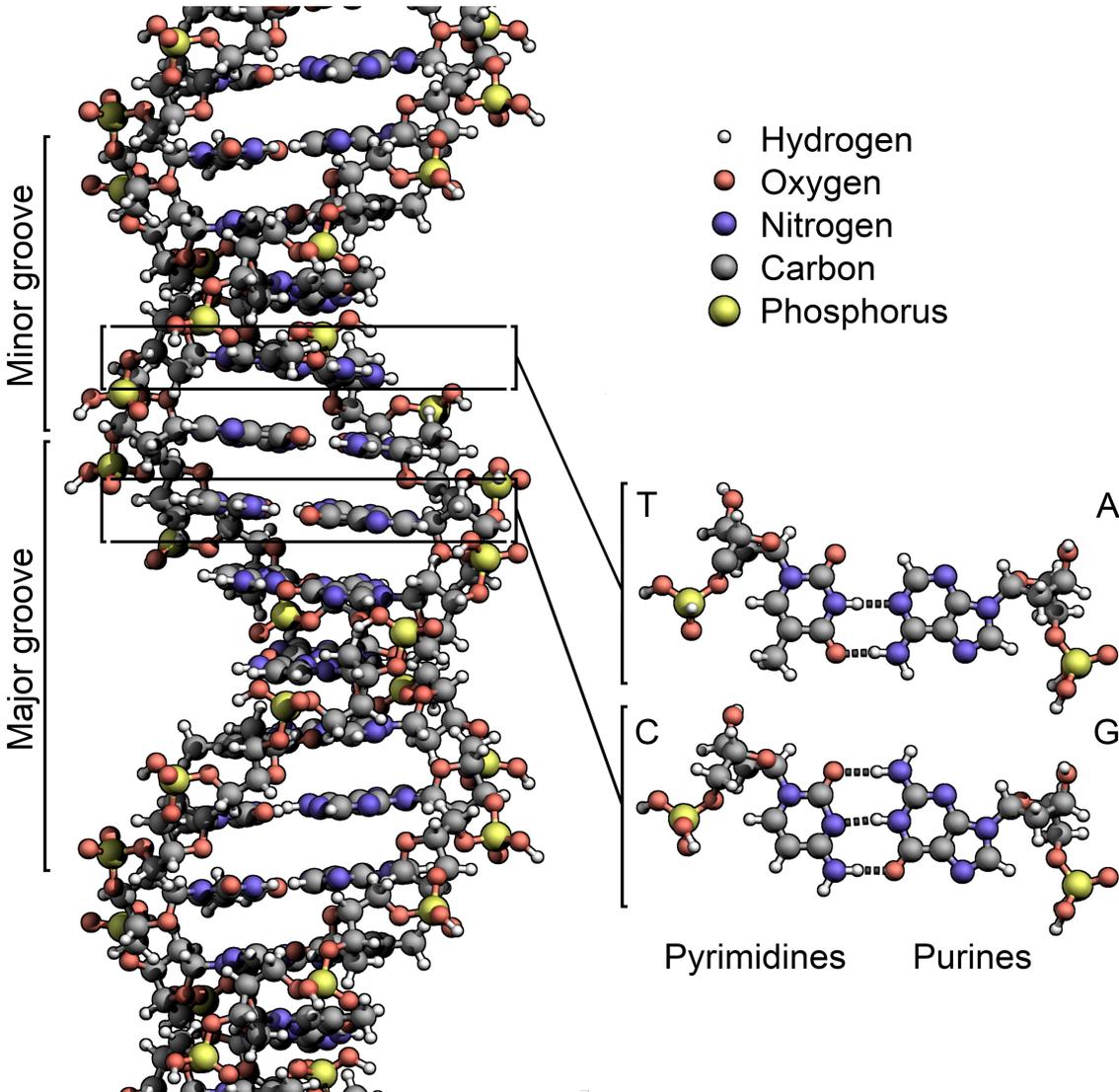
Modèle de la structure de l'ADN (Watson and Crick, 1953b)

Source: Watson & Crick (1953). Nature 4356:737-738



- WATSON, J.D. and CRICK, F.H. (1953a) The structure of DNA. *Cold Spring Harb Symp Quant Biol*, **18**, 123–131.
- Watson, J. and Crick, F. (1953b) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
- WATSON, J.D. and CRICK, F.H. (1953c) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, **171**, 964–967.
- Franklin, R.E. and Gosling, R.G. (2003) Molecular configuration in sodium thymonucleate. 1953.

Structure de l'ADN



Source: Watson & Crick (1953). Nature 4356:737-738.

Modèle de la structure de l'ADN (Watson and Crick, 1953b)

https://en.wikipedia.org/wiki/DNA#/media/File:DNA_Structure%2BKey%2BLabelled.pn_NoBB.png

Réplication
L'ADN, support de l'hérédité

Implications de la structure de l'ADN

- Il n'a pas échappé à notre attention que l'appariement spécifique que nous avons postulé suggère immédiatement un mécanisme possible pour copier le matériel génétique.

Watson & Crick (1953b)

- Notre modèle d'acide désoxyribonucléique constitue en fait une paire de modèles, chacun étant complémentaire de l'autre. Nous imaginons qu'avant la duplication* les liens d'hydrogènes sont rompus, et que les deux chaînes se déroulent et se séparent. Chaque chaîne agit alors comme modèle pour la formation, sur elle-même, d'une nouvelle chaîne compagne, de telle sorte qu'à la fin nous obtenons deux paires de chaînes là où il n'y en avait qu'une auparavant. De plus, la séquence des paires de bases aura été dupliquée exactement.

Watson & Crick (1953c)

Note: aujourd'hui on appelle ce processus "réplication", le terme "duplication" étant réservé pour désigner un type particulier de mutation.

GENETICAL IMPLICATIONS OF THE STRUCTURE OF DEOXYRIBONUCLEIC ACID

By J. D. WATSON and F. H. C. CRICK

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge

THE importance of deoxyribonucleic acid (DNA) within living cells is undisputed. It is found in all dividing cells, largely if not entirely in the nucleus, where it is an essential constituent of the chromosomes. Many lines of evidence indicate that it is the carrier of a part of (if not all) the genetic specificity of the chromosomes and thus of the gene itself.

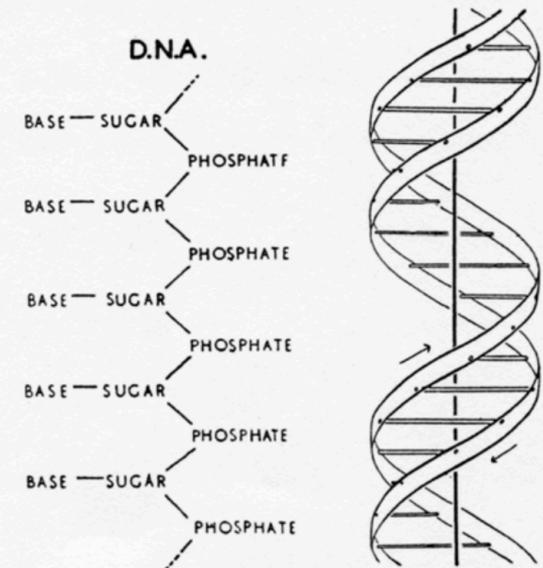


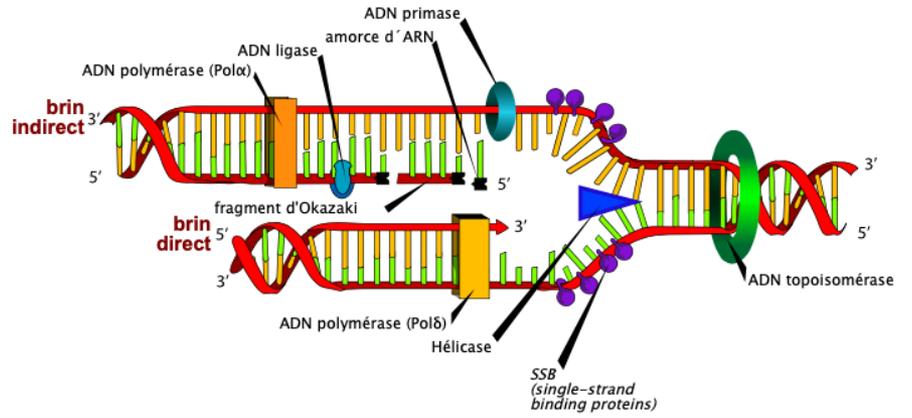
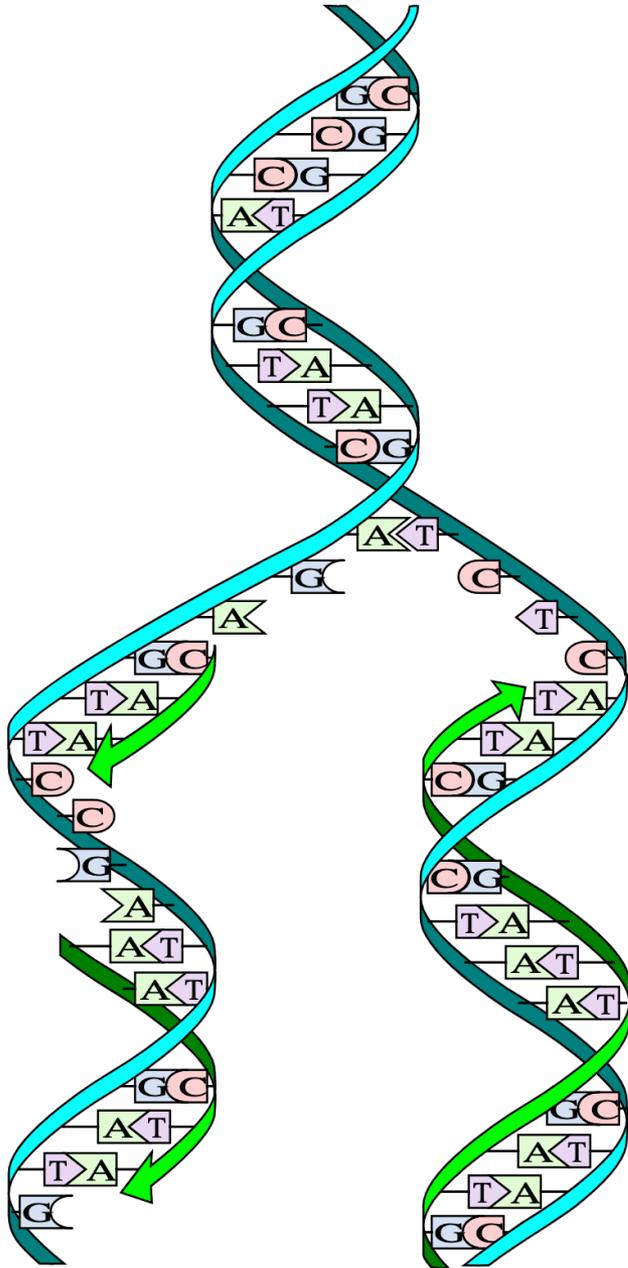
Fig. 1. Chemical formula of a single chain of deoxyribonucleic acid

Fig. 2. This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

Reprinted by permission of Macmillan & Company, Ltd., from NATURE, 171, 964-969 (1953).

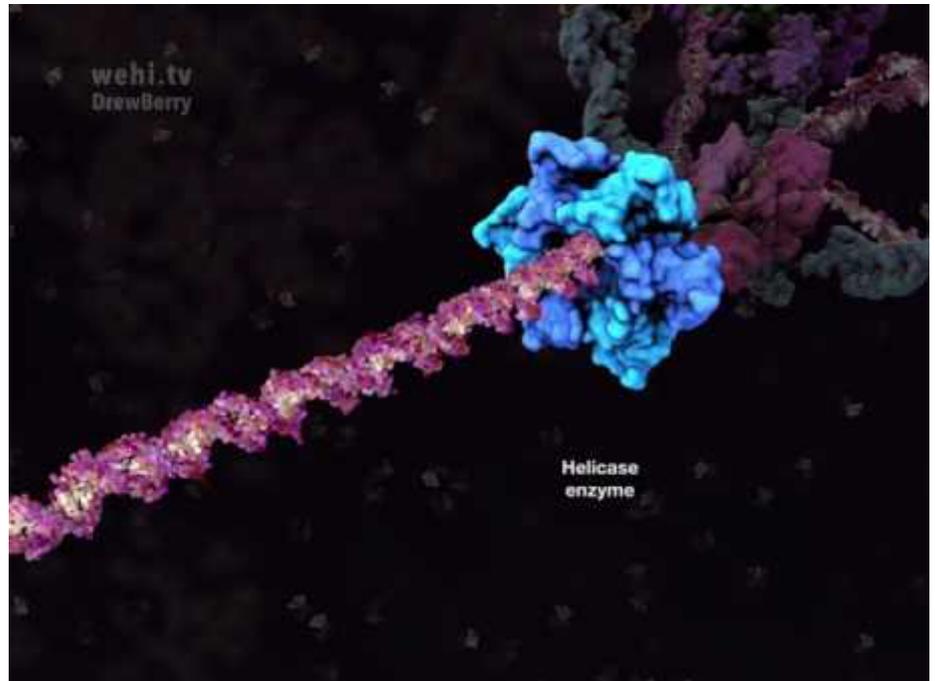
- Watson, J. and Crick, F. (1953b) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, 737-738.
- WATSON, J.D. and CRICK, F.H. (1953c) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171, 964-967.

Réplication



https://upload.wikimedia.org/wikipedia/commons/1/1c/DNA_replication_fr.svg

[Animation « DNA replication » par Drew Berry](https://www.youtube.com/watch?v=6j8CV3droDw)
<https://www.youtube.com/watch?v=6j8CV3droDw>

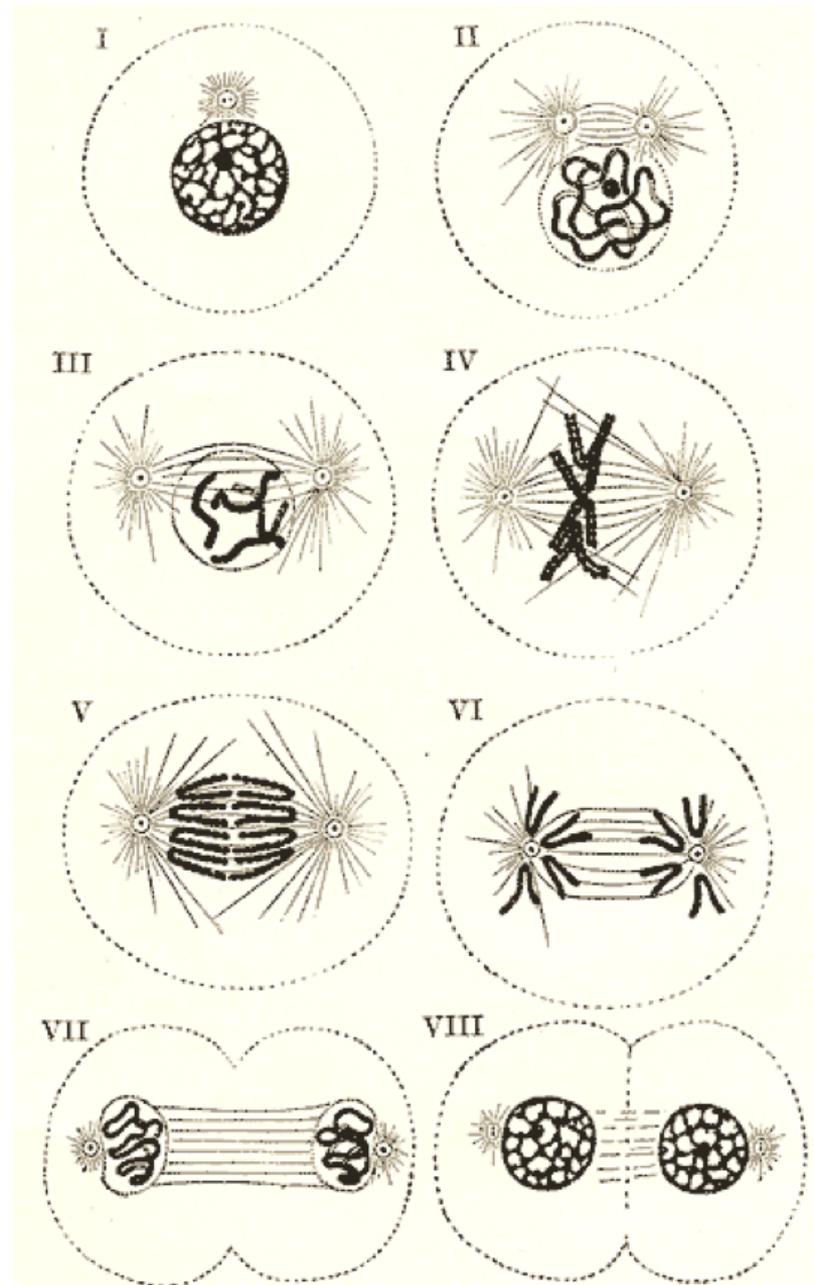


Mitose

Diagram showing the changes which occur in the centrosomes and nucleus of a cell in the process of mitotic division. (Schäfer.) I to III, prophase; IV, metaphase; V and VI, anaphase; VII and VIII, telophase.

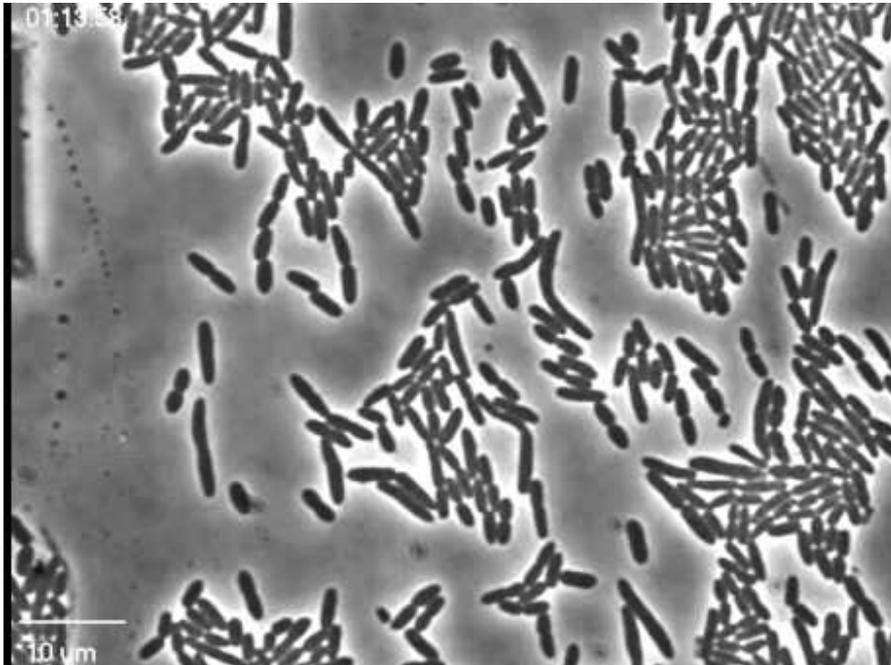
Henry Gray (1825–1861). *Anatomy of the Human Body*. 1918.

<https://www.bartleby.com/107/illus2.html>



Divisions cellulaires en microscopie optique

Divisions cellulaires de la bactérie *Escherichia coli*



<https://www.youtube.com/watch?v=5bGPa-QXV4>

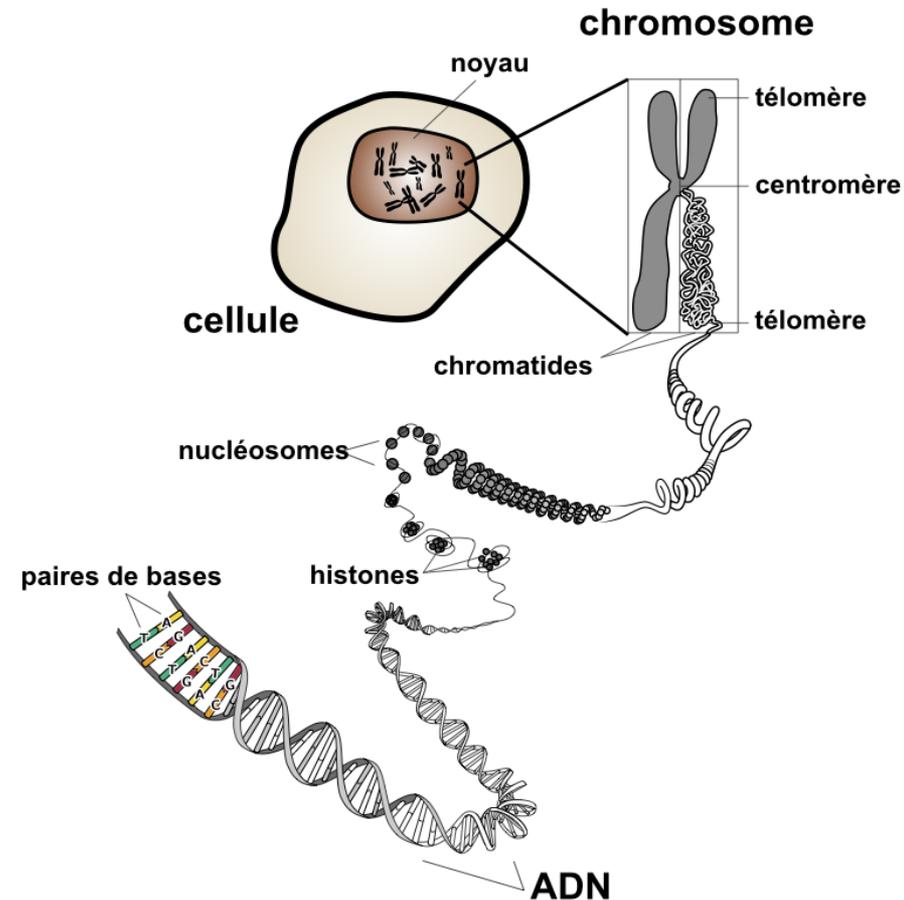
Division d'une cellule eucaryote



https://www.youtube.com/watch?v=L61Gp_d7evo

Déroulons un chromosome

- Cette représentation artistique montre l'organisation structurale d'un chromosome avec différents niveaux de détails
 - ❑ Appariement des paires de bases
 - ❑ Structure en double hélice de l'ADN
 - ❑ Enroulement de l'ADN autour des histones
 - ❑ Super-enroulement de la chromatine
 - ❑ Appariement des chromatides homologues au cours des division cellulaire
 - ❑ Localisation des chromosomes au sein du noyau cellulaire



Source de l'image:

https://upload.wikimedia.org/wikipedia/commons/2/2a/Chromosome_fr.svg

L'ADN, source de variabilité génétique
Mutations

Exemples de phénotypes mutants chez la drosophile

Drosophile de « type sauvage »



Mutant white



http://www.steve.gb.com/images/science/drosophila_melanogaster.jpg

http://www.scg.ubc.ca/quarterly012/white_drosophila.gif

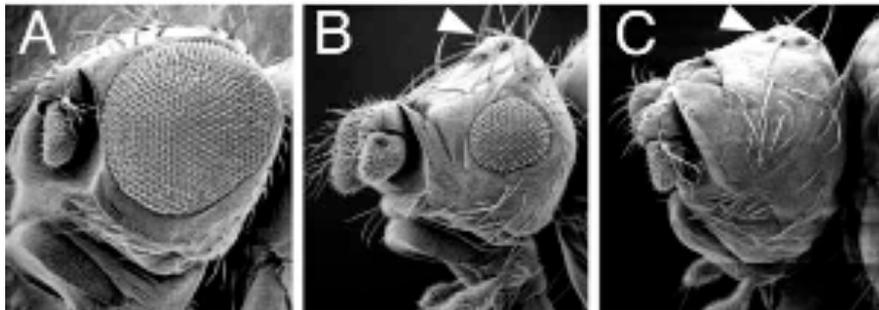
Mutant curly



Double mutant curly + white



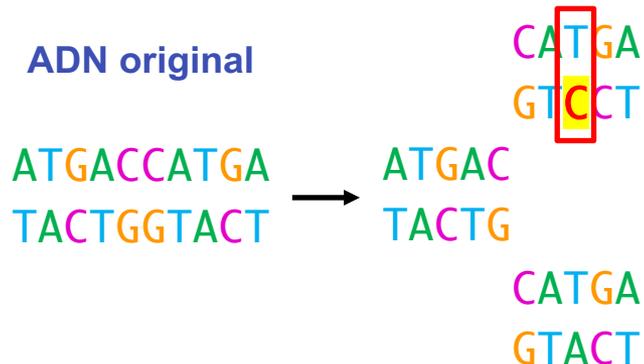
Têtes de mouches sauvage (A) et mutantes eyeless (B, C)



Substitution

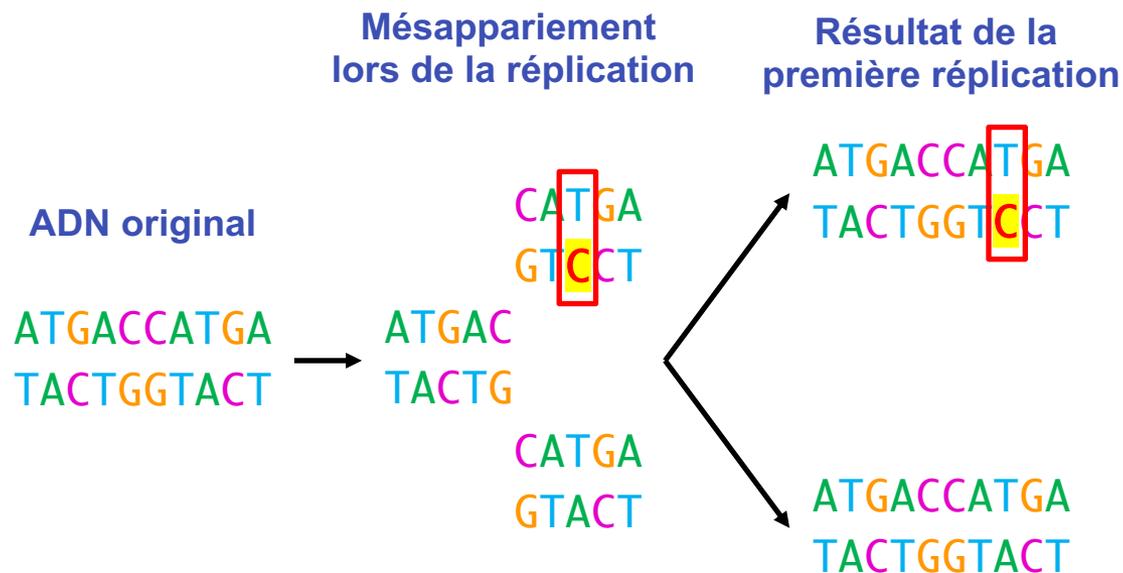
- Substitution
 - Remplacement d'un résidu (une lettre) par un autre
- Origine:
 - Lors d'une réplication, la polymérase de l'ADN incorpore un nucléotide incorrect sur un des brins (mésappariement).

Mésappariement lors de la réplication



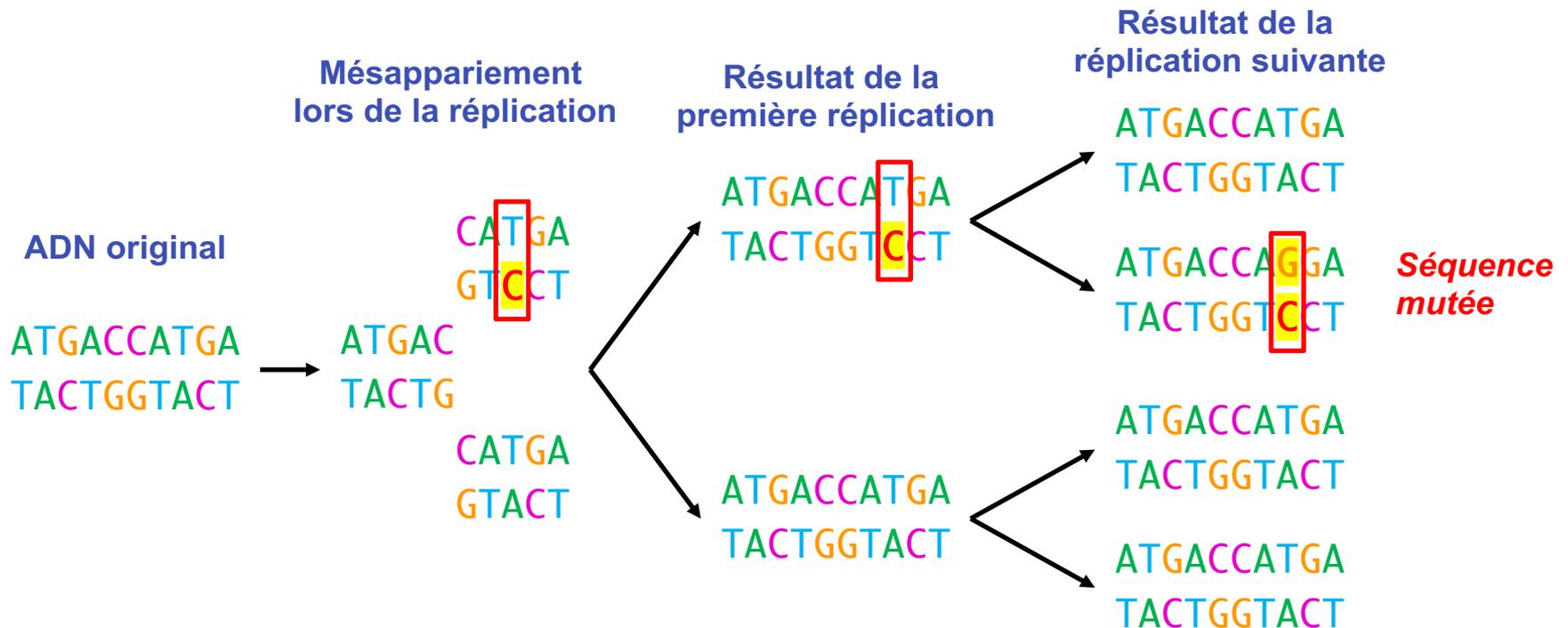
Substitution

- Substitution
 - Remplacement d'un résidu (une lettre) par un autre
- Origine:
 - Lors d'une réplication, la polymérase de l'ADN incorpore un nucléotide incorrect sur un des brins (mésappariement).



Substitution

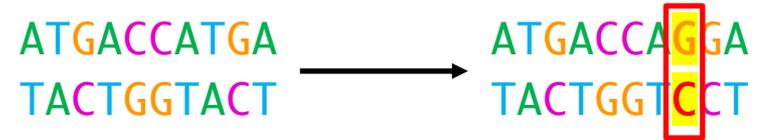
- Substitution
 - Remplacement d'un résidu (une lettre) par un autre
- Origine:
 - Lors d'une réplication, la polymérase de l'ADN incorpore un nucléotide incorrect sur un des brins (mésappariement).



Typologie des mutations

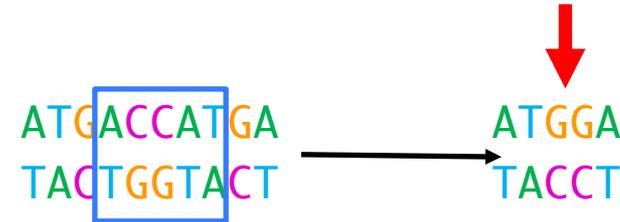
■ Substitution

- Remplacement d'un résidu (une lettre) par un autre



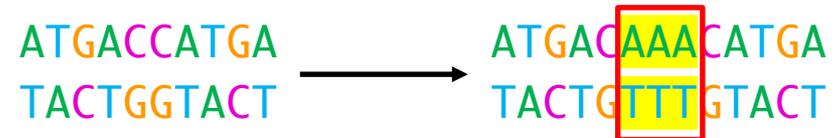
■ Délétion

- Suppression d'un fragment de la molécule



■ Insertion

- Ajout d'un fragment de molécule

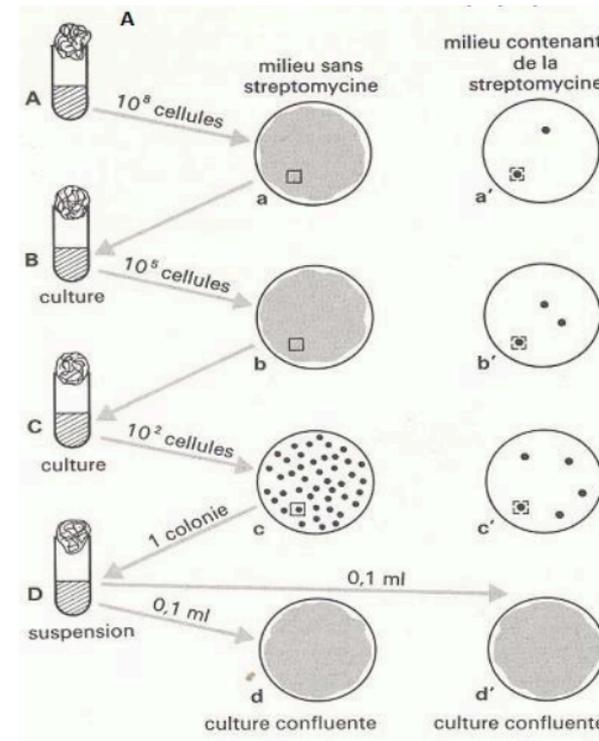


Les mutations sont indépendantes de la présence d'un agent sélectif

- Méthode des répliques par tampons de velours (Lederberg & Lederberg (1952))
- Sélection indirecte
 - Avec des tampons de velours, on obtient deux répliques d'un tapis bactérien
 - On applique un agent toxique pour la bactérie (phage, antibiotique) sur un des répliques, et on localise les mutants.
 - On sélectionne la même région sur la réplique non-soumise à l'agent toxique, on la met en culture et on génère deux nouvelles répliques.
 - On constate un enrichissement en colonies résistantes sur la réplique soumise à l'agent toxique.
 - Sur la réplique non soumise à l'agent toxique, on sélectionne la région correspondant à un clone résistant.
 - Au bout de quelques générations, on obtient des colonies purs de résistants.
- Lederberg & Lederberg :
 - « *Nous concluons que la résistance à la streptomycine, ainsi que celle au phage, est une mutation spontanée qui se produit indépendamment de la présence de l'agent sélectif.* »



<https://www.youtube.com/watch?v=uYb-hYSVYw>



https://fsnv.univ-setif.dz/telecharger/EDT2017/GENETIQUE_MICROBIENNE.pdf

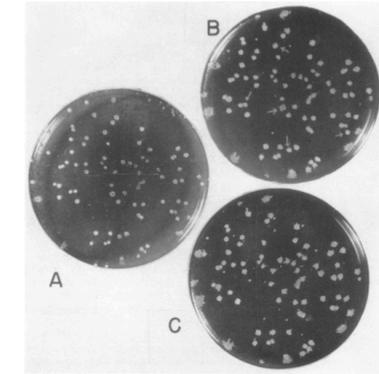


Figure 1. Replica plating for the isolation of auxotrophic colonies. A, Initial plate; B, Replica; both on complete agar medium. C, Second replica to minimal agar. The arrows designate the auxotrophic colonies which fail to grow on minimal medium. The resolution of these replicas is of fair to average quality.

Lederberg, J., and Lederberg, E.M. (1952).

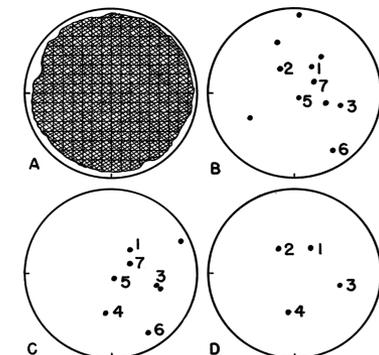


Figure 2. Clonal occurrence of mutants resistant to phage T-1. A, Initial or replica plate on plain agar with diffuse, confluent growth, (semidiagrammatic). B, C, and D, Successive replicas from A to agar coated with phage, from tracing of a typical experiment. Superimposable colonies of resistant cells are numbered. These are concluded to be derived from small clones of resistant mutants already present at corresponding sites on the plain agar plate, A.

Lederberg, J., and Lederberg, E.M. (1952).

- Lederberg, J., and Lederberg, E.M. (1952). Replica plating and indirect selection of bacterial mutants. J. Bacteriol. 63, 399–406.

« DNA makes RNA makes protein »

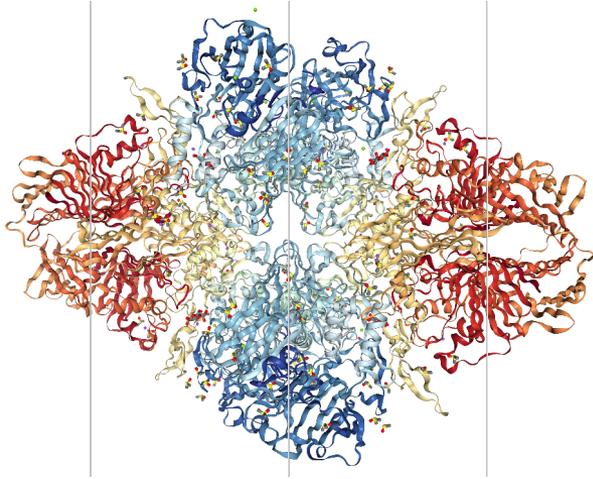
L'ADN, support de l'information moléculaire

Structure tridimensionnelle d'une protéine

Beta-galactosidase (structure 1JYN)

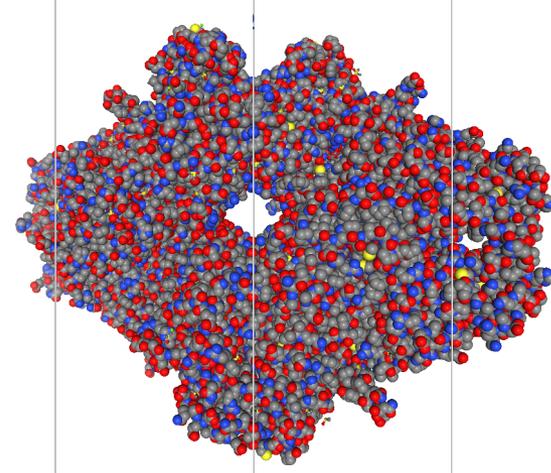
Représentation « cartoon », coloration par chaîne peptidique

<https://www.rcsb.org/3d-view/1JYN/1>



Beta-galactosidase (structure 1JYN)

Représentation « spacefill », coloration par type d'atome (CPK)



The image part with relationship ID r1d7 was not found in the file.

The image part with relationship ID r1d7 was not found in the file.

Séquence d'ADN

Séquence nucléique (ADN) du gène de la beta-galactosidase d'*Escherichia coli*.

ATGACCATGATTACGGATTCCTACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTTCC
GGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCATCTTCTGAGGCCGATACTGTCGTCGTCCTCAAACCTGGCAGATGCACGGTTACGAT
GCGCCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATG
TTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGAACGGGCGCTGGGTGCGTTA
CGGCCAGGACAGTCGTTTGCCGTCGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGCTGGAGTGAC
GGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGCATTTCCTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCC
ATGTTGCCACTCGCTTTAATGATGATTTACGCCGCGCTGACTGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGT
TTCTTTATGGCAGGGTCAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCGTCACA
CTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCAGCTGA
TTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAA
CCGTCACGAGCATCATCCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACCTTTAACGCCGTG
CGCTGTTCCGATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCA
TGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACCGTAACCGGAATGGTGCAGCGCGATCGTAATCACCCGAG
TGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCCGGTG
CAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCGATGTACGCGCGCTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGA
AATGGTCCATCAAAAATGGTTTTGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGGTTT
CGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGCTGGGACTGGGTGGATCAGTCGCTGATTAATATGATGAAAAC
GGCAACCCGTGGTTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCG
ATCCAGCGCTGACGGAAGCAAACACCAGCAGCAGTTTTTCCAGTTCGTTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCA
TAGCGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAACAG
TTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGAAG
CCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCTCCCACGCCATCCCGCATCTGACCACAG
CGAAATGGATTTTTGTCATCGAGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTTACAGATGTGGATTGGCGATAAAAAACAA
CTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGACCCTAACGCCTGGGTG
AACGCTGGAAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGTGATGCGGTGCTGATTACGACCGCTCA
CGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTG
GCGAGCGATACCCGCATCCGGCGCGGATTGGCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAA
ACTATCCCGACCGCCTTACTGCCGCTGTTTTGACCGCTGGGATCTGCCATTGTGACACATGTATAACCCCGTACGTCTTCCCAGCGAAAACGGTCT
GCGCTGCGGGACGCGCGAATTGAATTATGGCCACACCAGTGGCGCGGCGACTTCCAGTTC AACATCAGCCGCTACAGTCAACAGCAACTGATGGAA
ACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGTTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGT
CAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGGTGTCAAAAATAA

Séquence d'ADN

Séquence nucléique (ADN) du gène de l'opsine verte chez l'humain

>NG_011606.1 Homo sapiens opsin 1, medium wave sensitive (OPN1MW), RefSeqGene on chromosome X
TGGCGGGAAGCAAATTTGTA CTACAGCCACCTTAGAAAACCACTTGACAGTATCACAAAGTTGAAGATAT
ACCTACCAAATGATCCAGCAACTGCATTCATAGGCGATATTCCCAGGAGGCAAGTACACTGATGTTCTCA
ATATCCAACA ACTAGCAACAAACCACCTGCCAGAAAGTGTAAAGTAACTGTAATATATGCAACAACATA
CTACTGTACAGCAATGAGAATGAACTACAGCCTTGAAAAACATGAATGCATTTCAAGCATAATGTTGA
ACAAAAGAAACCAGACAGAAAATCCATACTATATATGGTTCATTTACATAAATTCCAAGAGCAGGCAAA
ACTAAACAATATTGTGTAGGGATGACTACACAGGCAGGAAAGCTAAAGAGCAGCATGGATGTTGGCAACA
GAAACCCAGGATGGTGGTCCCCTAGAGCGGGAGAAGGGGATTATGACTGGGGTACTGTGGAGGCTG
GGGTGCCTGGCCTGAGTAGTGGTTAAATGTGTGTTACCTTATCGTTACTTCTTAAACATACATCAATG
TTTTCTGCACCTTTATATACATATGCCAGATTTACAGTTTTCAA ACTATTTAAATGTTCAATATAGGCC
AGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGAGGTCAGGAGGTC
GACCAGTCTGGCCAACATGGTGAACCCCGTCTATACAAAATACAAAAA AAAAAAAAAAATGCCTGGCA
CAGTGGTTCACACCTGACTTCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACCTGAGGTCGGGAGTT
CAAGACCAGCCTGACCAACACAGAGAAACCCCATCTCTACTAGAAATACAAAATTAGCCAGGCGTGGTGG
CACACGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAACCGCTTGAACCCGGGGTGCAGAAGT
TGCAGTGAGCTGAGATCGTGCCACTGCACTCCAGCCTGGGTGGTAGAGTGAGACTCCGTCTCAAAAAATA
AATAAATAATAAATAAATAATATGTTCAATTGATGGATTTAATGGCAGATTTGATATGGCTGGAGAGACTG
TTGGTAAATTGTACAGAGGAGCTAAAGAAAGTATCCAGGCCGGGCGCGGTGGCTCACACGCCTGTAATCC
CAGCATTTTTGGGAGGCCAAGGCGGGTGGATCACCTGACGTCAGGAGTTTGAGGCTAACCTGACCAACATG
GTGAAACCCTGTCTCTACTAAAAAATACAAAAAATTA ACTGGGCATGGTGGTGGGCACCTGTAATCCCAG
CTACTTGGGAGGCTGAGGCAGGAGAATGGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAGATCAC
GCCACTGCACTCCAACCTGGGCAACAAGAGCAAACTCCATCCCAAGAAAGAAAGAGAGAGAGAGAGAGA
GAGAGAGAGAGAGAAAGAGAAAAGAAAGAAAGAAAAA AAGTATCCAGAATGCAGCACAGAGGCACAAAA
AAATAGAAAAC TTGAAAAATAGAAGACATGGAAGGTAGCATGAGTAGGTCTAACAAAAATCCAATTGTCA
CACCAGAGAGAGAAGACAGACAGAATGAAGCAGCAGCAATATATACATATTTTTCTTCTTTTTTTTTTAA
ATTGAAATGGAGTCACACTCTGTTGCATAGGCGGAGTGCCGTGGCACAATCTCGGCTCACGGCCGACCTC
TGCTGCCC GGTTCAAGCACTTCTCGTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCGTGCCATC
ACAGCTAGCTAATTC TTTTGTATTTTTAGTAGAGATGGGGTTTCATCATGTTGGCAAGACTGTTCTCGAA

...

Séquence: https://www.ncbi.nlm.nih.gov/nucleotide/NG_011606.1?report=fasta

Info sur le gène: <https://www.ncbi.nlm.nih.gov/gene/2652>

Séquence peptidique de l'opsine verte humaine

```
>sp|P04001|OPSG_HUMAN Medium-wave-sensitive opsin 1 OS=Homo sapiens OX=9606 GN=OPN1MW PE=1
SV=1 MAQQWSLQRLAGRHPQDSYEDSTQSSIFTYTNNSNSTRGPFEGPNYHIAPRWVYHLTSVWM
IFVVIASVFTNGLVLAATMKFKKLRHPLNWILVNLAVADLAETVIASTISVVNQVYGYFV
LGHPMCVLEGYTVSLCGITGLWSLAIISWERWMVCKPFGNVRFDKLAIVGIAFSWIWA
AVWTAPPIFGWSRYWPHGLKTSCGPDVFSGSSYPGVQSYMIVLMVTCCITPLSIIIVLCYL
QVWLAIRAVAKQQKESESTQKAEKEVTRMVVVMVLAFCFCWGPYAFFACFAAANPGYPFH
PLMAALPAFFAKSATIYNPVIYVFMNRQFRNCILQLFGKKVDDGSELSSASKTEVSSVSS
VSPA
```

Les acides aminés, composants des protéines

Protéines (polypeptides)

- Polymères d'acides aminés
- 20 acides aminés distincts:
 - Arginine (A), Histidine (H), Lysine (K), Acide aspartique (D), Acide glutamique (E), Sérine (S), Thréonine (T), Asparagine (N), Glutamine (Q), Cystéine (C), Glycine (G), Proline (P), Alanine (A), Valine (V), Isoleucine (I), Leucine (L), Méthionine (M), Phénylalanine (P), Tyrosine (T), Tryptophane (W)

Séquence peptidique (protéique) de la beta-galactosidase d'Escherichia coli.

MTMITDLSLAVVLQRRDWNPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNG
 EWRFAWFPAPEAVPESWLCDLPEADTVVVPNSWQMHGYDAPYITNVYTPITVNPFF
 VPTENPTGCYSLTFNVDES WLQEGQTRIIIFDGVNSAFHLWCNGRWVGYGQDSRLPSE
 FDLSAFLRAGENRLAVMVLRWSDGSYLEDDQMWRMSGIFRDVSL LHKPTTQISDFHV
 ATRFNDDFSRAVLEAEVQMC GELRDYL RVTVSLWQGETQVASGTAPFGGEIIDERGG
 YADRVTLRNLNENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGL
 LLLNGKPLLIRGVNRHEHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLW
 YTLCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVII
 WSLGNESGHGANHDALYRWIKSVDPSPVQYEGGADTTATDIICPMYARVDEDDQPF
 PAVPKWSIKKWLSPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWD
 WVDQSLIKYDENGPNWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQQQF
 FQFRLSGQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGGKQLIELP
 ELQPESAGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHL
 TTSEMDFCIELGNKRWQFNRSGLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSE
 ATRIDPNAWVERWKAAGHYQAEALLQCTADTLADAVLITTAHAWQHGGKTLFISRK
 TYRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQVAERNWLGLGPQENYDRLT
 AACFDRWDLPLSDMYTPYVFPSENGLRCTRELNYGPHQWRGDFQFNISRYSQQQLM
 ETSRHLHLHAEEGTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK

Twenty-One Amino Acids

⊕ Positive ⊖ Negative
 * Side chain charge at physiological pH 7.4

A. Amino Acids with Electrically Charged Side Chains

Positive

- Arginine (Arg) **R**: NC(=O)C(CCN=[NH2+])N (pKa 12.10)
- Histidine (His) **H**: NC(=O)C(Cc1c[nH]c1)N (pKa 6.04)
- Lysine (Lys) **K**: NC(=O)C(CCN)N (pKa 10.67)

Negative

- Aspartic Acid (Asp) **D**: NC(=O)C(C(=O)[O-])N (pKa 3.71)
- Glutamic Acid (Glu) **E**: NC(=O)C(CCC(=O)[O-])N (pKa 4.15)

B. Amino Acids with Polar Uncharged Side Chains

- Serine (Ser) **S**: NC(=O)C(CO)N (pKa 9.05)
- Threonine (Thr) **T**: NC(=O)C(C(C)O)N (pKa 9.06)
- Asparagine (Asn) **N**: NC(=O)C(C(=O)N)N (pKa 8.70)
- Glutamine (Gln) **Q**: NC(=O)C(CCC(=O)N)N (pKa 9.00)

C. Special Cases

- Cysteine (Cys) **C**: NC(=O)C(CS)N (pKa 10.28)
- Selenocysteine (Sec) **U**: NC(=O)C(CSe)N (pKa 10)
- Glycine (Gly) **G**: NC(=O)CN (pKa 9.58)
- Proline (Pro) **P**: C1CCN1C(=O)N (pKa 10.47)

D. Amino Acids with Hydrophobic Side Chain

- Alanine (Ala) **A**: NC(=O)C(C)N (pKa 9.71)
- Valine (Val) **V**: NC(=O)C(C(C)C)N (pKa 9.52)
- Isoleucine (Ile) **I**: NC(=O)C(C(C)CC)N (pKa 9.60)
- Leucine (Leu) **L**: NC(=O)C(C(C)C)C(C)C (pKa 9.58)
- Methionine (Met) **M**: NC(=O)C(CSC)N (pKa 9.05)
- Phenylalanine (Phe) **F**: NC(=O)C(Cc1ccccc1)N (pKa 9.09)
- Tyrosine (Tyr) **Y**: NC(=O)C(Cc1ccc(O)cc1)N (pKa 9.04)
- Tryptophan (Trp) **W**: NC(=O)C(Cc1c[nH]c2ccccc12)N (pKa 9.34)

Découvrir la pierre de rosette

Acide déoxyribonucléique (ADN)

- Polymères de nucléotides
- 4 nucléotides possibles:
 - Adénosine (A)
 - Cytidine (C)
 - Guanosine (G)
 - Thymidine (T)

Séquence nucléique (ADN) du gène de la beta-galactosidase d'*Escherichia coli*.

```
ATGACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
CCTGGCGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGT
AATAGCGAAGAGGCCCGACCGATCGCCCTTCCAACAGTTGCGCAGCCTGAATGGC
GAATGGCGCTTTGCCTGGTTTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAG
TGCGATCTTCTGAGGCCGATACTGTCGTCGTCCCTCAAACCTGGCAGATGCACGGT
TACGATGCGCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTT
GTTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGC
TGGCTACAGGAAGGCCAGACGCGAATTAATTTTTGATGGCGTTAACTCGGCGTTTCAT
CTGTGGTGCAACGGGCGCTGGGTGCGTTACGGCCAGGACAGTCGTTTTGCCGCTGAA
TTTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTG
CGCTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATT
TTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTT
GCCACTCGCTTAATGATGATTTACGCCGCGCTGACTGGAGGCTGAAGTTCAGATG
TGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAAACG
CAGGTGCGCAGCGGCACCGCGCCTTTCCGGCGGTGAAATTAATCGATGAGCGTGGTGGT
TATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCC
GAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTG
ATTGAAGCAGAAGCCTGCGATGTCGGTTTTCCGCGAGGTGCGGATTGAAATGGTCTG
CTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAAACCGTCACGAGCATCAT
CCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATG
AAGCAGAACAACCTTAACGCCGTGCGCTGTTTCGATTATCCGAACCATCCGCTGTGG
TACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACC
CACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGCGATG
AGCGAACCGTAACCGCAATGGTGCAGCGCGATCGTAATACCCGAGTGTGATCATC
TGGTCGCTGGGGAATGAATCAGGCCACGGCGTAATCACGACGCGCTGATCGCTGG
```

Protéines (polypeptides)

- Polymères d'acides aminés
- 20 acides aminés distincts:
 - Arginine (A), Histidine (H), Lysine (K), Acide aspartique (D), Acide glutamique (E), Sérine (S), Thréonine (T), Asparagine (N), Glutamine (Q), Cystéine (C), Glycine (G), Proline (P), Alanine (A), Valine (V), Isoleucine (I), Leucine (L) Méthionine (M), Phénylalanine (P), Tyrosine (T), Tryptophane (W)

Séquence peptidique (protéique) de la beta-galactosidase d'*Escherichia coli*.

```
MTMITDSLAVVLQRRDWNENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNG
EWRFAWFPAPEAVPESWLECDLPEADTVVPSNWQMHGYDAPYITNVYPTITVNPFF
VPTENPTGCYSLTFNVDESWLQEQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPSE
FDLSAFLRAGENRLAVMVLRWSDGSYLEDDQDMWRMSGIFRDVSLLLHKPTTQISDFHV
ATRFNDDFSRAVLEAEVQMCGERLDYL RVTVSLWQGETQVASGTAPFGGEIIDERGG
YADRVTLRLNVENPKLWSAEIPNL YRAVVELHTADGTLIEAEACDVGFREVRIENGL
LLLLNGKPLLIRGVNRHEHPLHGQVMDEQTMVQDILLMKQNNFNAVRC SHYPNHPLW
YTLCDRYGLYVVDEANIE THGMVPMNRLTDDPRWL PAMSERVTRMVQRDRNHPSV I I
WSLGNESGHGANHDALYRWIKSVDP SRPVQYEGGGADTTATDII CPMYARVDEDQPF
PAVPKWSIKKWL SLPGETRPLIL CEYAHAMGNSL GGF AKYWQAFRQYPR LQGGFVWD
WVDQSLIKYDENGPNWSAYGGDFG DTPNDRQFCMNGLVFADRTPHPAL TEAKHQQQF
FQFRLSGQTEIETSEYLFRHSDNELLHMMVALDGKPLASGEVPLDVAPOGKQLIELP
ELPQPE SAGQLWLTVRVVQPNATAWSEAGHISAWQQWR LAENL SVTLPAASHAIPHL
TTSEMDFCIELGNKRWQFNRSQSGFLSQMWWIGDKKQLL TPLRDQFTRAPLDNDIGVSE
ATRIDPNAWVERWKAAGHYQAEALLQCTADTLADAVLITTAHAWQHOGKTLFISRK
TYRIDGSGQMAITVDVEVASDTPH PARIGLNCQLAQVAERNWLGLGPQENYPRDLT
AACFDRWDLPLSDMYTPYVFPSENGLR CGTRELNYGPHQWRGDFQFNISRYSQQLM
ETSHRHLLHAEEGTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK
```

Découvrir la pierre de rosette

Séquence nucléique (ADN) du gène de la beta-galactosidase d'*Escherichia coli*.

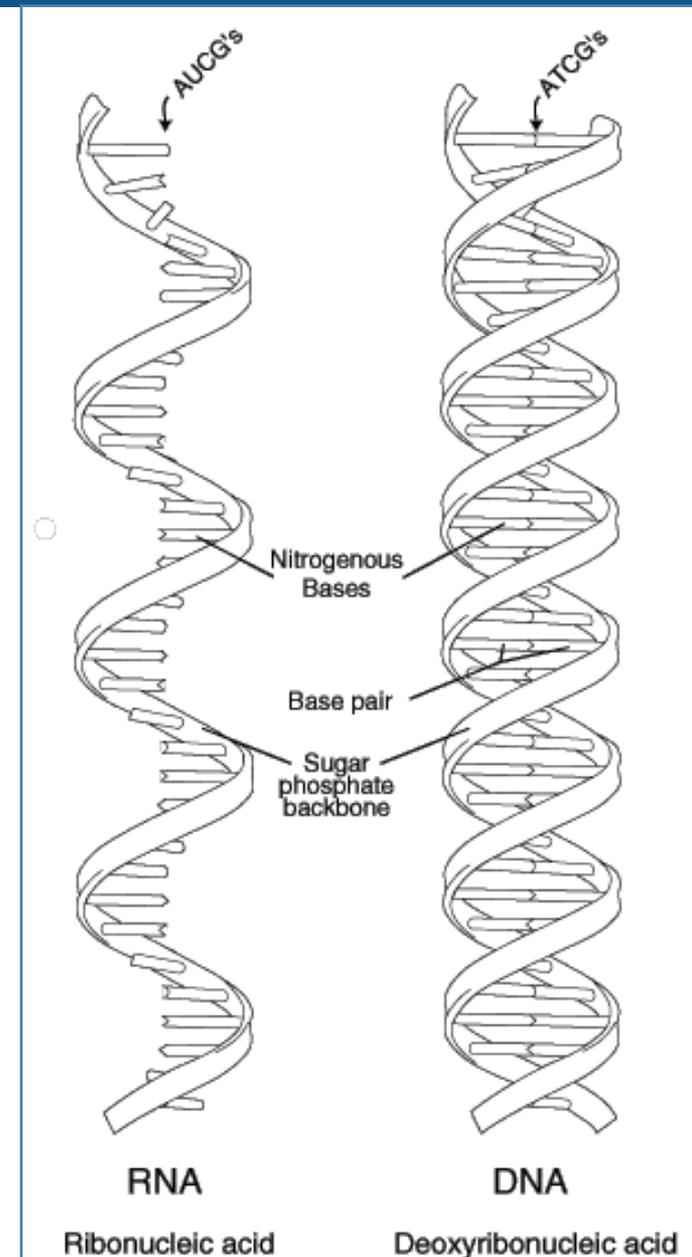
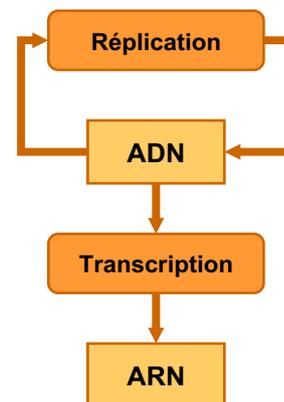
ATGACCATGATTACGGATTACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTG
GCGTTACCCAACCTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGA
AGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTT
GCCTGGTTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCATCTTCTCTGAGG
CCGATACTGTCGTGCTCCCTCAAACGGCAGATGCACGGTTACGATGCGCCCATCTACAC
CAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTCCCACGGAGAATCCGACGGGT
TGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGCCAGACGCGAATTA
TTTTTGATGGCGTTAACTCGCGTTTTATCTGTGGTCAACGGGCGCTGGGTGCGTTACGG
CCAGGACAGTCGTTTCCGCTGTAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAAC
CGCCTCGCGGTGATGGTGTGCGCTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGT
GGCGGATGAGCGGCATTTCCGCTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAG
CGATTTCCATGTTGCCACTCGCTTAAATGATGATTTACGCCGCGCTGACTGGAGGCTGAA
GTTCAGATGTGCGGCGAGTTGCGTGAATCTACCTACGGGTAACAGTTTCTTTATGGCAGGGTG
AAACGCAGGTCGCCAGCGGCACCGCGCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGG
TTATGCCGATCGCGTACACATACGCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAA
ATCCCGAATCTCTATGTCGCGGTGGTGAACGTCACACCGCCGACGGCAGCGTATTGAAG
CAGAAGCTCGCATGTCGTTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAA
CGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAG
GTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACCTTAAACG
CCGTGCGCTGTTGCGATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGG
CCTGTATGGTGGATGAAGCCAATATTGAAACCCACGGCATGGTGCATGAATCGTCTG
ACCGATGATCCGCGCTGGCTACCGCGATGAGCGAACGCGTAACGCGAATGGTGCAGCGCG
ATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAA
TCACGACGCGCTGTATCGCTGGATCAAATCTGTGATCCTTCCCGCCCGGTGCAGTATGAA
GGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGCGCTGGATG
AAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTGCTACCTGG
AGAGACGCGCCCGAATGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGGT
TTCGCTAAATACTGGCAGGCGTTTTGCTCAGTATCCCCGTTTACAGGGCGGCTTCTGCTGGG
ACTGGGTGGATCAGTCGCTGATTAATATGATGAAAACGGCAACCCGTTGGTGGCTTACGG
CGGTGATTTTGGCGATACGCCAAGCATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCC
GACCGCACGCCGATCCAGCGCTGACGGAAGCAAACACCAGCAGCAGTTTTTCCAGTTCC
GTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGA
GTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTG
GATGTCGCTCCACAAGGTAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCG
CCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAAGAGC
CGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCC
GCCGCGTCCACGCCATCCCGCATCTGACCACAGCGAAATGGATTTTTGCATCGAGCTGG
GTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTACAGATGTGGATTGGCGA
TAAAAACAACCTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGAC
ATTGGCGTAAGTGAAGCGACCCGATTGACCCTAACGCTGGGTGAAACGCTGGAAGGCGG
CGGGCCATTACAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGGTGTGATGC
GGTGTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGG

Séquence peptidique (protéique) de la beta-galactosidase d'*Escherichia coli*.

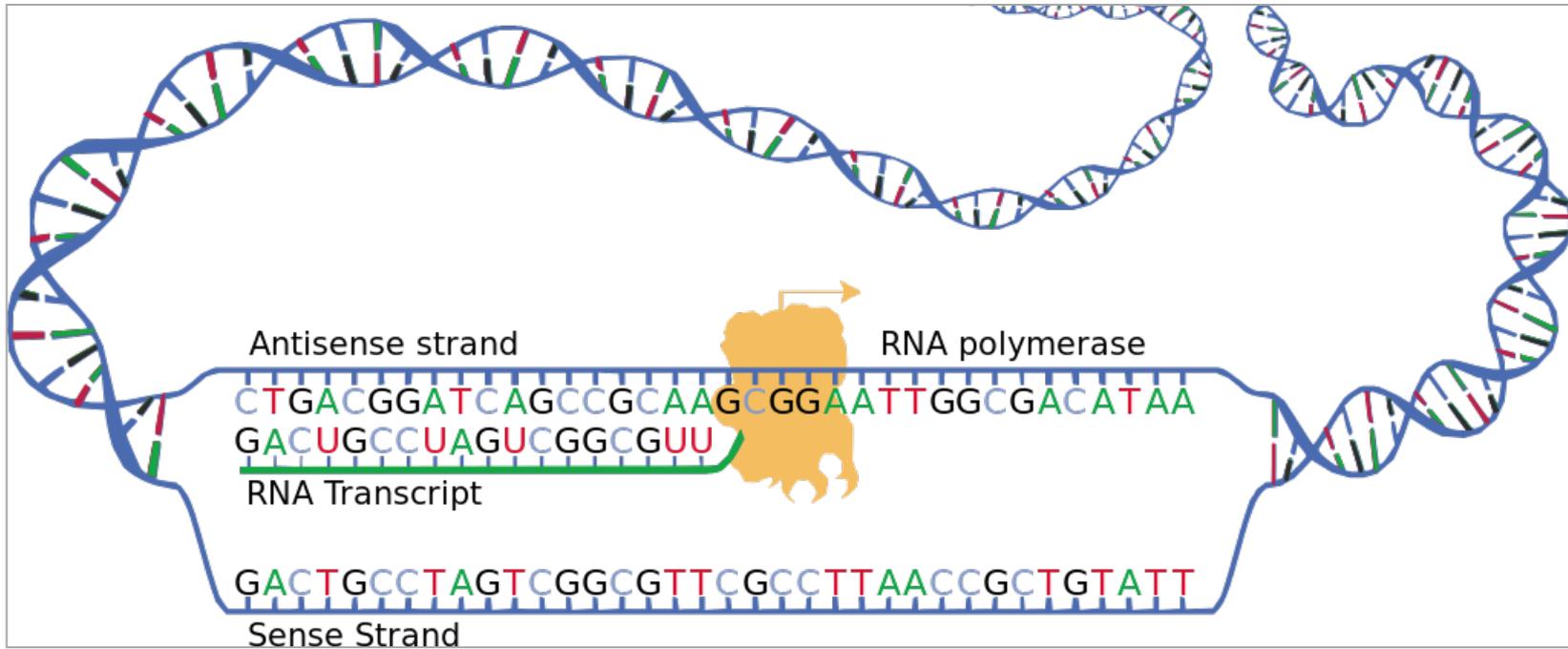
MTMITDSLAVVLQRRDWNPGVTQLNRLAAHPPFASWRNSEEARTRPSOQLRSLNGEWR
FAWFPAPEAVPESWLECDLPEADTVVVPNSWQMHGYPAPIYTNVYPIITVNPFPVPTENP
TGCYSLTFNVDESWLQEQTRIIIFDGVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAFLRA
GENRLAVMVLWSDGSYLEQDMWRMSGIFRDVSLLLHKPTTQISDFHVATR FNDDFSRAV
LEAEVQMCGELRDYLRTVSLWQGETQVASGTAPFGGEIIDERGGYADRVTLLNLVNPK
LWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLIRGVNRHEH
HPLHGQVMDEQTMVQDILLMKQNNFNVRCSHYPNHPLWYTLCDRYGLYVVDEANIETHG
MVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHDALYRWIKSVD
PSRPVQYEGGGADTTATDIIICPMYARVDEQPFPAVPKWSIKKWLSPGETRPLILCEYA
HAMGNSLGGFAKYWQAFRQYPRLQGGFVWDVQSLIKYDENGPNWSAYGGDFGDTPNDR
QFCMNLVFADRTPHPALTEAKHQQQFFQFRLSGQTIETSEYLFRHSDNELLHWMVALD
GKPLASGEVPLDVAPQKQLIELPELPQESAGQLWLTVRVVQPNATAWSEAGHISAWQQ
WRLAENLSVTLPAASHAPHLTTSEMFDCIELGNKRWFNRQSGFLSQMWIGDKKQLLTP
LRDQFTRAPLDNDIGVSEATRIPNAWVERWKAAGHYQAEALLQCTADTLADAVLITTA
HAWQHKGKTLFISRKTYRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQVAERNWLGL
LGPQENYDRLTAACFDRWDLPLSDMYTPYVFPSENGLRGCTRELNYGPHQWRGDFQFNI
SRYSQQLMETSHRHLLHAEETWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRHYHQLV
WCQK

L'ARN - acide ribonucléique

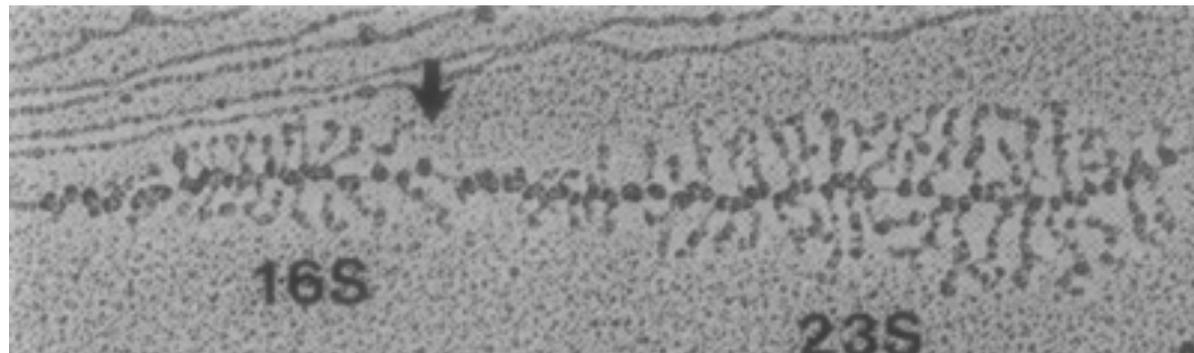
- L'acide ribonucléique (ARN) est généralement constitué d'une seule chaîne de nucléotides.
- La synthèse d'ARN à partir d'un modèle d'ADN est appelée transcription.
- Il y a une correspondance de un à un entre les nucléotides du modèle ADN et ceux de la copie ARN
 - ▣ Adénine → Uracile
 - ▣ Thymine → Adénine
 - ▣ Cytosine → Guanine
 - ▣ Guanine → Cytosine



Transcription de l'ADN en ARN



http://www.wikiwand.com/fr/Acide_d%C3%A9soxyribonucl%C3%A9ique

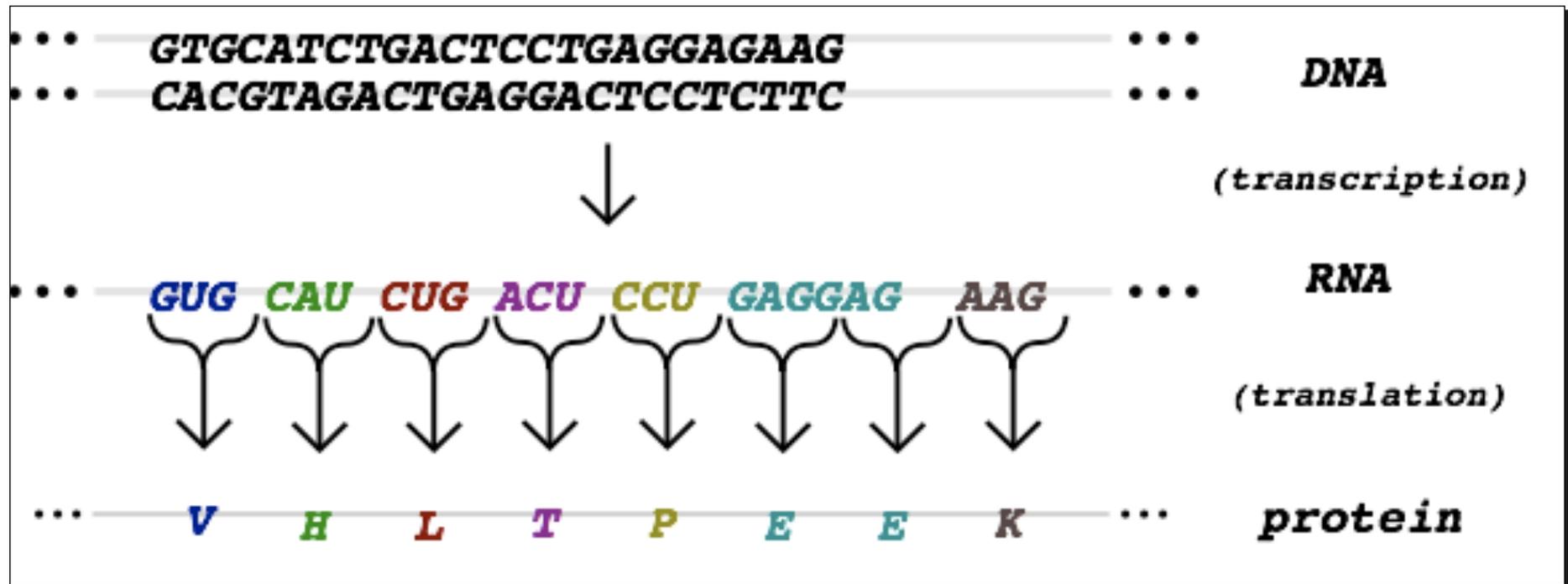


The 16S rRNA and 23S rRNA genes are co-transcribed. The arrow points to a region between these two genes. Intervening gene transcripts had been selectively removed by RNase III.

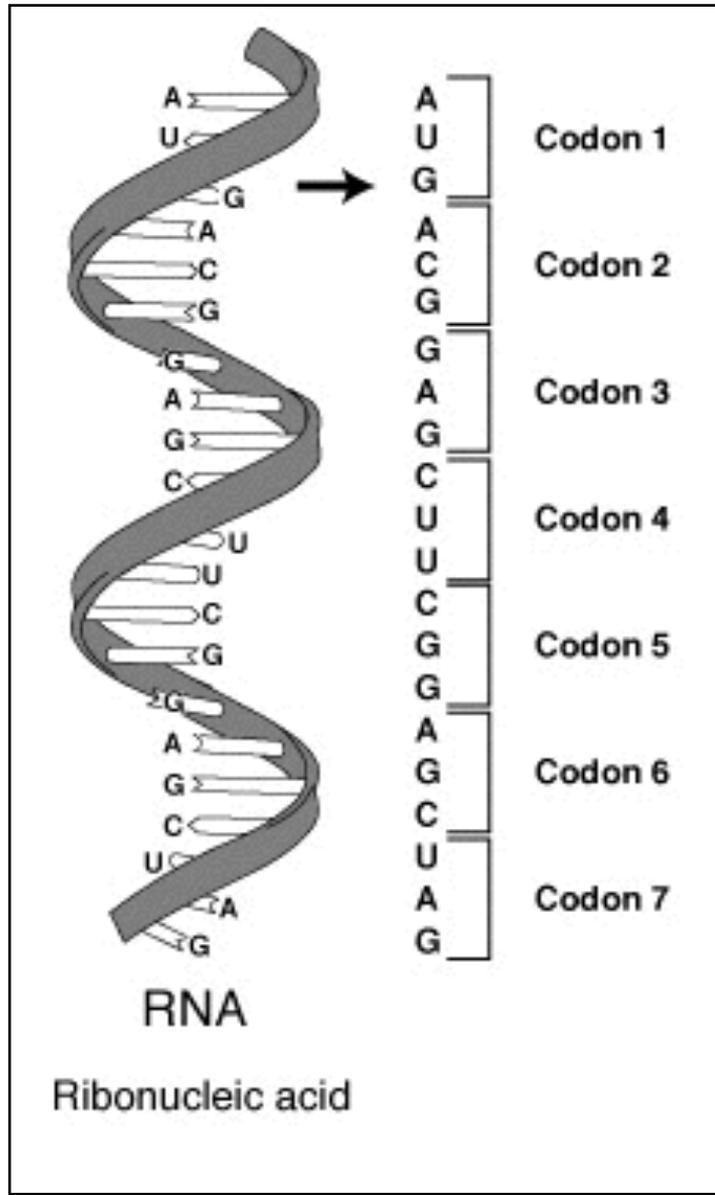
Direction of transcription is from left to right

(<http://schaechter.asmblog.org/schaechter/2013/03/pictures-considered.html>).

Le chemin de l'information : ADN → ARN → protéine



Le code génétique



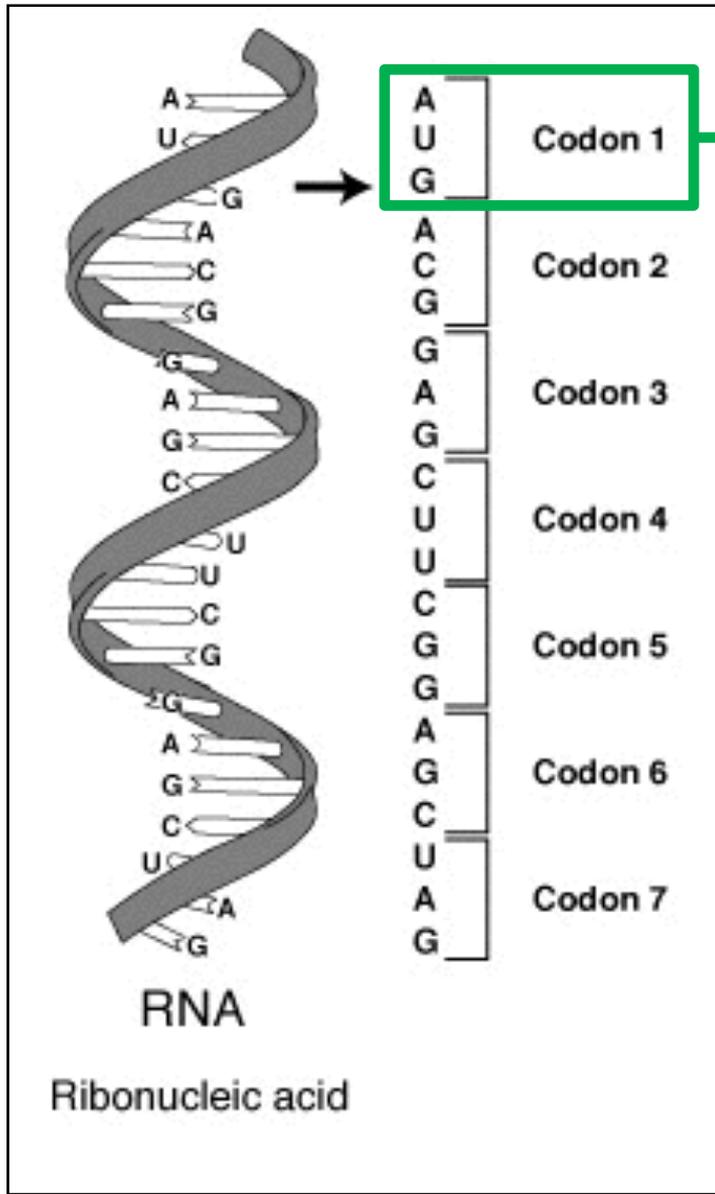
The genetic code is triplet

		Second base			
		U	C	A	G
First base	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA } STOP UAG }	UGU } Cys UGC } UGA } STOP UGG } Trp
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }
	A	AUU } AUC } Ile AUA } AUG } Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }

FIGURE 9.1 All the triplet codons have meaning: Sixty-one represent amino acids, and three cause termination (STOP).

Source: Genes IX

Le code génétique



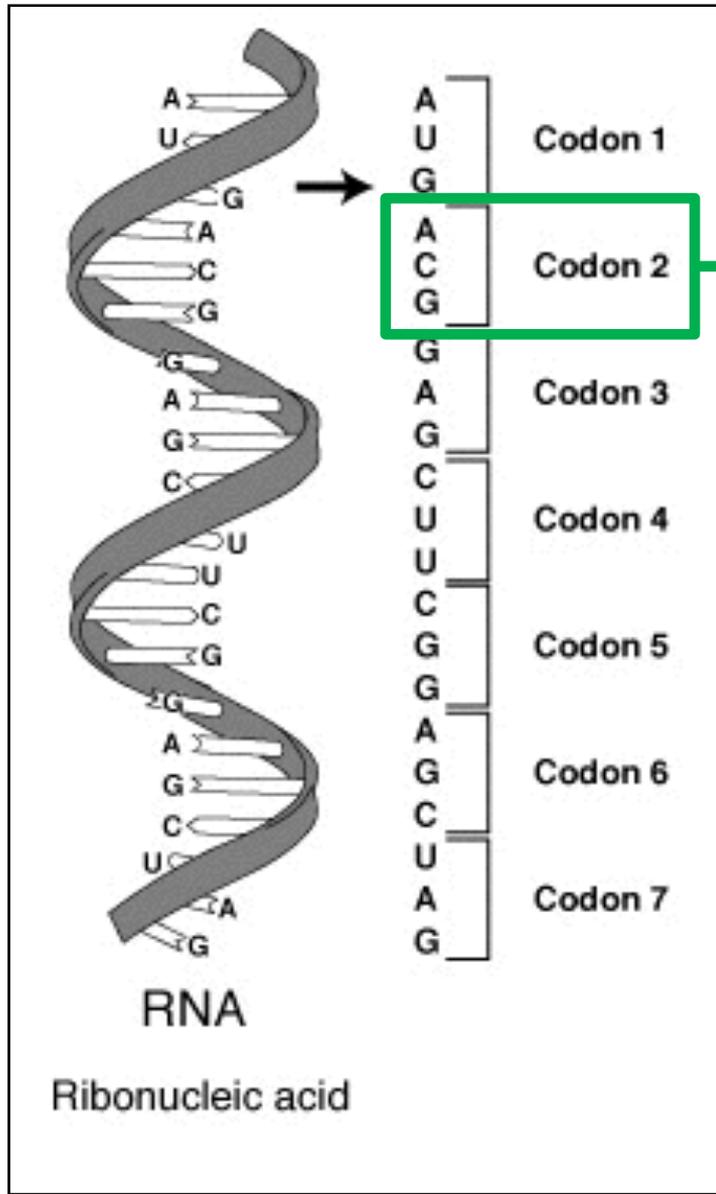
The genetic code is triplet

		Second base						
		U	C	A	G			
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Leu	UCC		UAC	STOP	UGC	STOP
	UUA		UCA	UAA	UGA			
	UUG	UCG	UAG	UGG	Trp			
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC		CCC		CAC	CGC		
	CUA		CCA	CAA	Gln	CGA		
	CUG		CCG	CAG	CGG			
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC		ACC		AAC	AGC		
	AUA	ACA	AAA	Lys	AGA	Arg		
	AUG	ACG	AAG	AGG				
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC		GCC		GAC	GGC		
	GUA		GCA	GAA	Glu	GGA		
	GUG		GCG	GAG	GGG			

FIGURE 9.1 All the triplet codons have meaning: Sixty-one represent amino acids, and three cause termination (STOP).

Source: Genes IX

Le code génétique



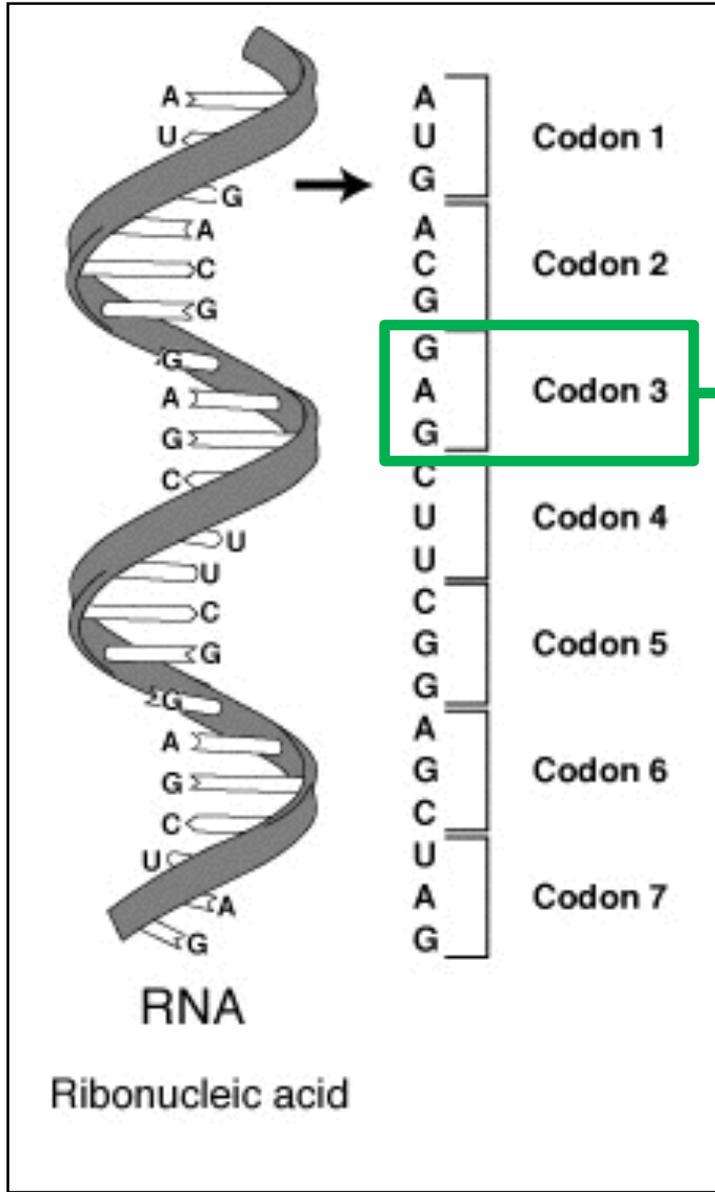
The genetic code is triplet

		Second base			
		U	C	A	G
U	U	UUU } Phe	UCU } Ser	UAU } Tyr	UGU } Cys
	U	UUC } Phe	UCC } Ser	UAC } Tyr	UGC } Cys
	U	UUA } Leu	UCA } Ser	UAA } STOP	UGA } STOP
	U	UUG } Leu	UCG } Ser	UAG } STOP	UGG } Trp
C	U	CUU } Leu	CCU } Pro	CAU } His	CGU } Arg
	C	CUC } Leu	CCC } Pro	CAC } His	CGC } Arg
	C	CUA } Leu	CCA } Pro	CAA } Gln	CGA } Arg
	C	CUG } Leu	CCG } Pro	CAG } Gln	CGG } Arg
A	U	AUU } Ile	ACU } Thr	AAU } Asn	AGU } Ser
	A	AUC } Ile	ACC } Thr	AAC } Asn	AGC } Ser
	A	AUA } Ile	ACA } Thr	AAA } Lys	AGA } Arg
	A	AUG } Met	ACG } Thr	AAG } Lys	AGG } Arg
G	U	GUU } Val	GCU } Ala	GAU } Asp	GGU } Gly
	G	GUC } Val	GCC } Ala	GAC } Asp	GGC } Gly
	G	GUA } Val	GCA } Ala	GAA } Glu	GGA } Gly
	G	GUG } Val	GCG } Ala	GAG } Glu	GGG } Gly

FIGURE 9.1 All the triplet codons have meaning: Sixty-one represent amino acids, and three cause termination (STOP).

Source: Genes IX

Le code génétique



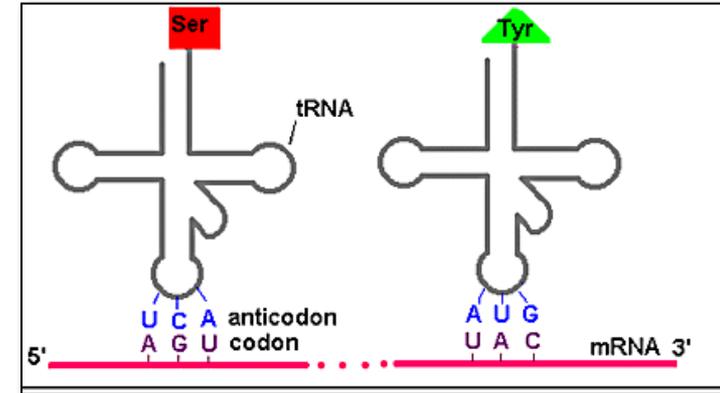
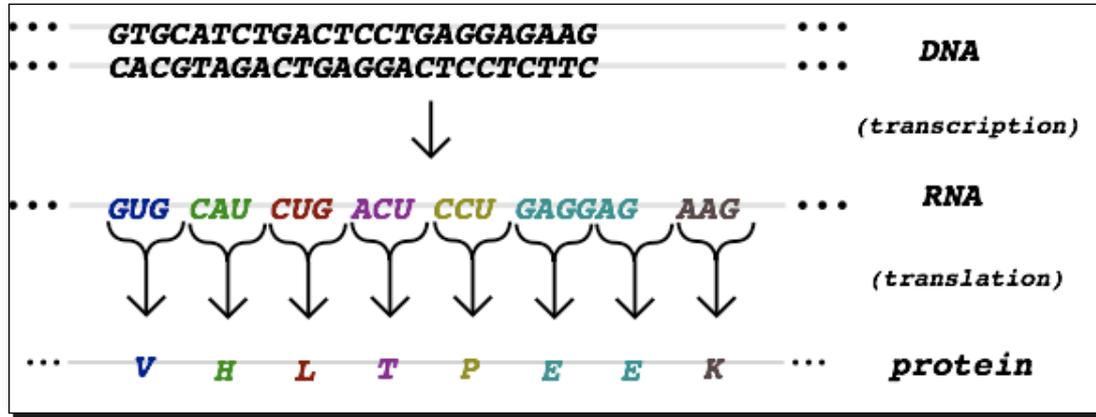
The genetic code is triplet

		Second base			
		U	C	A	G
First base U	U	UUU } Phe	UCU } Ser	UAU } Tyr	UGU } Cys
	U	UUC } Phe	UCC } Ser	UAC } Tyr	UGC } Cys
	U	UUA } Leu	UCA } Ser	UAA } STOP	UGA } STOP
	U	UUG } Leu	UCG } Ser	UAG } STOP	UGG } Trp
First base C	C	CUU } Leu	CCU } Pro	CAU } His	CGU } Arg
	C	CUC } Leu	CCC } Pro	CAC } His	CGC } Arg
	C	CUA } Leu	CCA } Pro	CAA } Gln	CGA } Arg
	C	CUG } Leu	CCG } Pro	CAG } Gln	CGG } Arg
First base A	A	AUU } Ile	ACU } Thr	AAU } Asn	AGU } Ser
	A	AUC } Ile	ACC } Thr	AAC } Asn	AGC } Ser
	A	AUA } Met	ACA } Thr	AAA } Lys	AGA } Arg
	A	AUG } Met	ACG } Thr	AAG } Lys	AGG } Arg
First base G	G	GUU } Val	GCU } Ala	GAU } Asp	GGU } Gly
	G	GUC } Val	GCC } Ala	GAC } Asp	GGC } Gly
	G	GUA } Val	GCA } Ala	GAA } Glu	GGA } Gly
	G	GUG } Val	GCG } Ala	GAG } Glu	GGG } Gly

FIGURE 9.1 All the triplet codons have meaning: Sixty-one represent amino acids, and three cause termination (STOP).

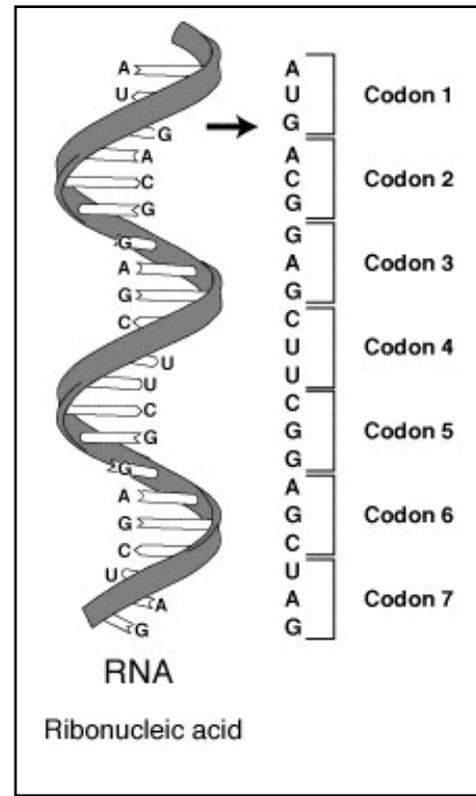
Source: Genes IX

Le code génétique



<http://www.accessexcellence.org/AB/GG/>

- Le code génétique a été élucidé en 1961.
- Les protéines sont synthétisées à partir de l'ARN.
- Contrairement à la transcription, il n'y a pas de correspondance de un à un entre nucléotides et acides aminés. En effet, l'ARN ne comporte que 4 nucléotides distincts (adénine, uracile, guanine et cytosine), tandis que les protéines sont formées de 20 acides aminés distincts.
- Chaque acide aminé est spécifié par une succession de 3 nucléotides (un *codon*).
- Il y a 64 triplets de nucléotides possibles. Plusieurs codons spécifient le même acide aminé (*dégénérescence du code*).
- On appelle *traduction* la synthèse d'un polypeptide à partir d'un modèle d'ARN.



The genetic code is triplet

		Second base			
		U	C	A	G
U	UUU	UCU	UAU	UGU	
	UUC	UCC	UAC	UGC	
	UUA	UCA	UAA	UGA	
	UUG	UCG	UAG	UGG	
C	CUU	CCU	CAU	CGU	
	CUC	CCC	CAC	CGC	
	CUA	CCA	CAA	CGA	
	CUG	CCG	CAG	CGG	
A	AUU	ACU	AAU	AGU	
	AUC	ACC	AAC	AGC	
	AUA	ACA	AAA	AGA	
	AUG	ACG	AAG	AGG	
G	GUU	GCU	GAU	GGU	
	GUC	GCC	GAC	GGC	
	GUA	GCA	GAA	GGA	
	GUG	GCG	GAG	GGG	

FIGURE 9.1 All the triplet codons have meaning: Sixty-one represent amino acids, and three cause termination (STOP).

Source: Genes IX

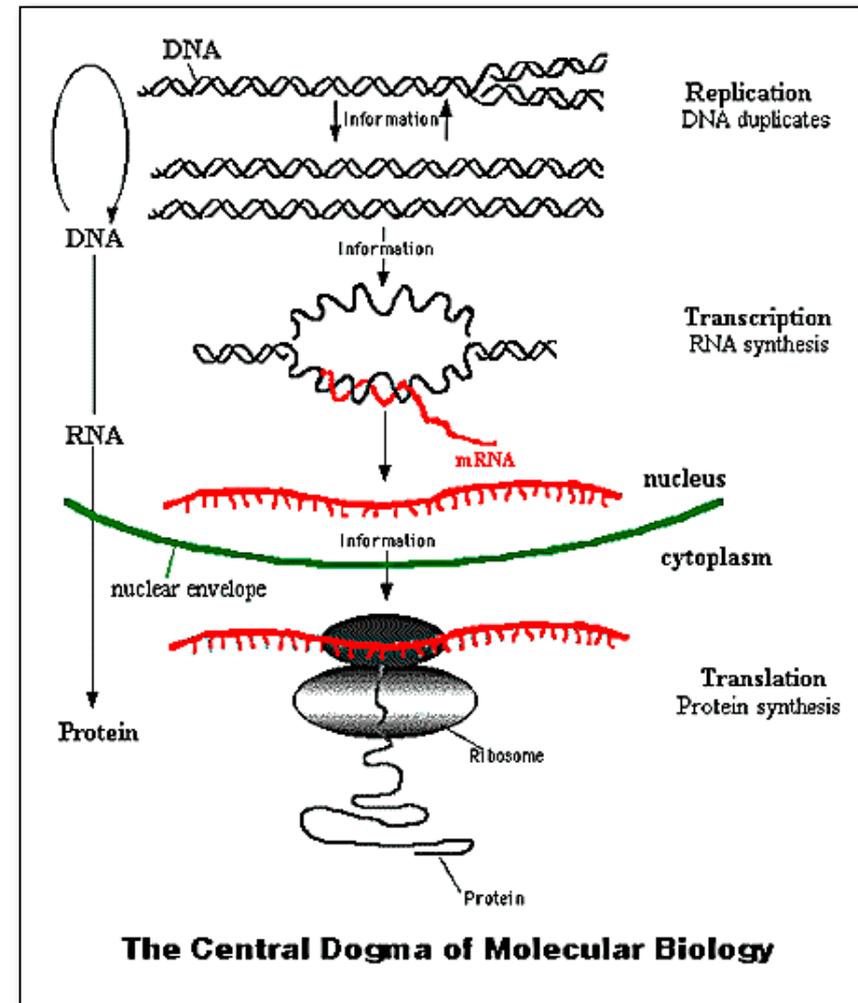
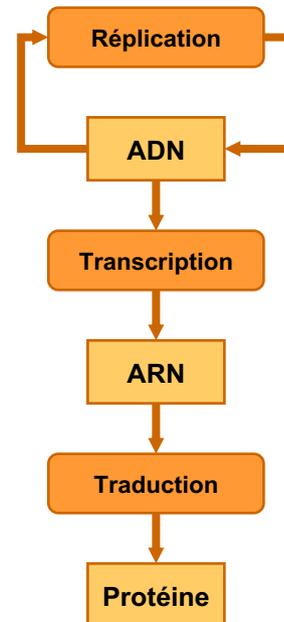
Le "dogme central"

- Le « **dogme central** » a été formulé en 1958 par Francis Crick
 - Crick, F. H. (1958). On protein synthesis. Symp Soc Exp Biol 12, 138-63.
 - Je recommande également de lire cette discussion ultérieure : Crick, F. (1970). Central dogma of molecular biology. Nature 227, 561-3.
- On le résume souvent de la façon suivante
 - *DNA makes RNA makes protein*
 - *(l'ADN fait l'ARN qui fait les protéines)*
- Cependant le dogme ne dit pas exactement cela. Il énonce les transferts d'information qui sont possibles ou impossibles entre les séquences d'acides nucléiques et celles des protéines.

Le dogme central stipule que, une fois que l' « information » est passée dans la protéine elle ne peut pas en ressortir. Plus précisément, le transfert d'information serait possible d'acide nucléique à acide nucléique, ou d'acide nucléique à protéine, mais le transfert de protéine à protéine, ou de protéine à acide nucléique est impossible.

Information signifie ici la détermination précise de la séquence, soit des bases dans l'acide nucléique, soit des résidus aminoacides dans la protéine.

Crick, F. H. (1958). On protein synthesis. Symp Soc Exp Biol 12, 138-63.



<http://www.accessexcellence.org/AB/GG/central.html>

Le dogme central a-t-il été réfuté ?

- On a à plusieurs reprises affirmé que le dogme central avait été réfuté.
 - Découverte de la transcription réverse.
 - Découverte du prion.
- Dès 1970 Crick publie une clarification, pour rappeler ce que dit le dogme central, et explique pourquoi la réverse transcription ne le réfute pas.
- Il distingue 3 classes de transfert d'information.
 - Transferts pour lesquels on dispose d'indications directes ou indirectes (flèches pleines) : $DNA \rightarrow DNA$, $DNA \rightarrow RNA$, $RNA \rightarrow Protein$, $RNA \rightarrow RNA$.
 - Possibles mais sans aucune indication d'existence (pointillés) : $RNA \rightarrow DNA$, $DNA \rightarrow Protein$.
 - Très invraisemblables : $Protein \rightarrow Protein$, $Protein \rightarrow DNA$, $protein \rightarrow RNA$.

All possible transfers

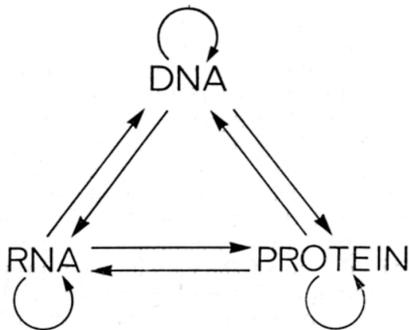


Fig. 1. The arrows show all the possible simple transfers between the three families of polymers. They represent the directional flow of detailed sequence information.

As it seemed in 1958

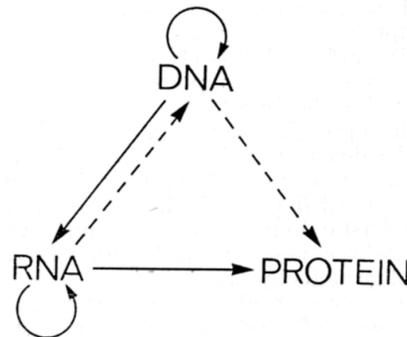


Fig. 2. The arrows show the situation as it seemed in 1958. Solid arrows represent probable transfers, dotted arrows possible transfers. The absent arrows (compare Fig. 1) represent the impossible transfers postulated by the central dogma. They are the three possible arrows starting from protein.

Revised with reverse transcription

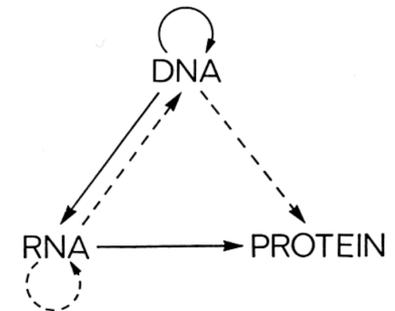
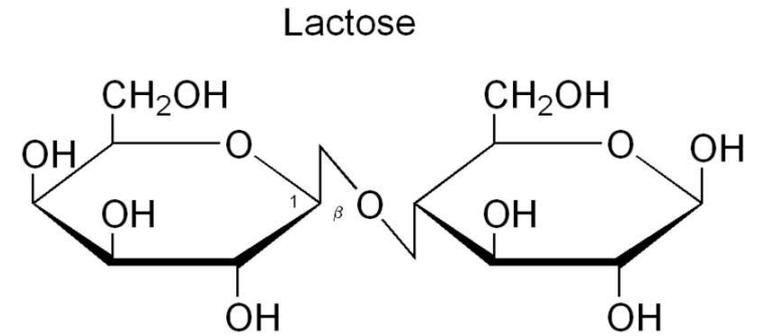


Fig. 3. A tentative classification for the present day. Solid arrows show general transfers; dotted arrows show special transfers. Again, the absent arrows are the undetected transfers specified by the central dogma.

OK, mais alors que font les protéines ?

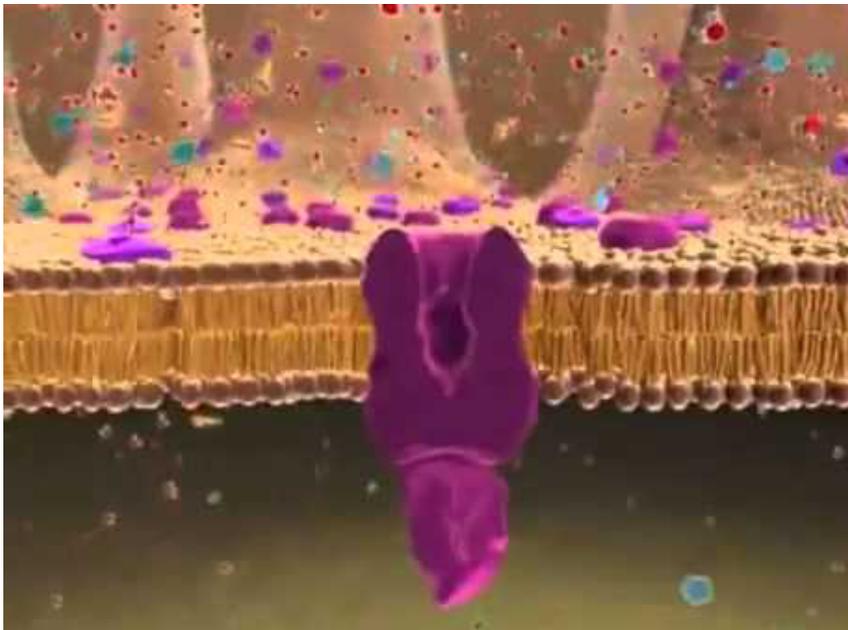
La perméase du lactose



<https://www.futura-sciences.com/sante/definitions/biologie-lactose-8703/>

Perméase du lactose

Animation illustrant le fonctionnement d'une perméase



<https://www.rcsb.org/3d-view/2Y5Y/1>

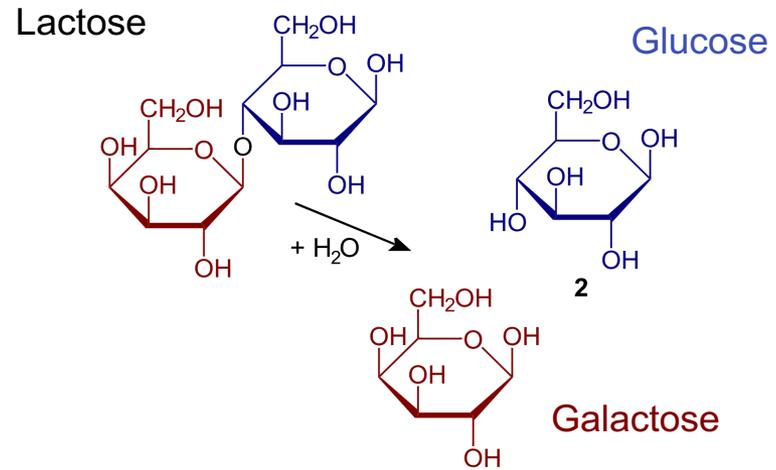
<https://www.youtube.com/watch?v=l1MZG6508IM>

Digestion du lactose chez l'entérobactérie

Beta-galactosidase



Réaction métabolique

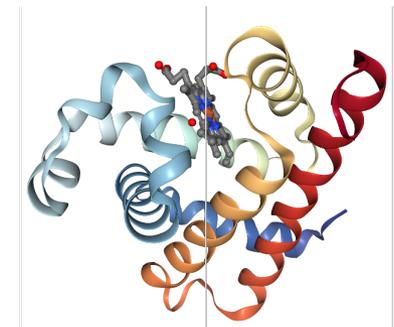


La séquence des protéines détermine leur structure tridimensionnelle

- La succession des acides aminés détermine la façon dont la protéine se replie.
- Certaines successions d'acides aminés forment des **hélices** (exemple: myoglobine) d'autres des **feuilletés** (exemple: porine).
- Les résidus **hydrophiles** se trouvent à l'**extérieur**, ils interagissent avec le cytoplasme
- Les résidus **hydrophobes** se retrouvent soit à l'**intérieur** de la protéine (ils interagissent entre eux plutôt qu'avec le cytoplasme), soit à l'interface entre la protéine et des membranes.

Myoglobine (PDB 1MBN)

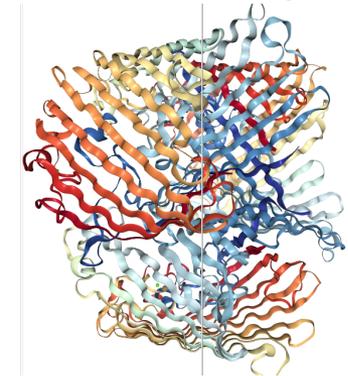
Noter les hélices alpha



<https://www.rcsb.org/3d-view/1MBN/1>

Porine (PDB 1A02)

Noter les hélices alpha



<https://www.rcsb.org/3d-view/1A0S/1>