

Parlons « protéines »...

Connaissez-vous le point commun entre le collagène, qui confère la souplesse à votre peau ; la kératine qui constitue vos ongles et vos cheveux ; l'insuline qui régule le taux de sucre dans le sang ; l'hémoglobine, qui transporte le dioxygène ; les anticorps qui reconnaissent les microbes ; ou les enzymes qui facilitent les réactions chimiques ?

Savez-vous comment une cellule conserve sa forme ?

Pourquoi nos ongles et nos cheveux poussent ?

Comment l'organisme peut transporter les nutriments et les métaboliser ?

Comment les cellules parviennent à communiquer entre elles ?

Comment parviennent-elles à se défendre ?

Où comment les spermatozoïdes sont capables de se mouvoir ?

La réponse à ces questions se résume en un mot : les « protéines ».

Les protéines sont des matériaux essentiels à la construction et au fonctionnement de l'organisme de tous les êtres vivants, de la bactérie à l'homme en passant par les champignons et les plantes.

Ces protéines sont des polymères d'acides aminés, ou, en d'autres mots, des molécules formées d'une succession d'acides aminés. Ces protéines peuvent être plus ou moins longues, de deux acides aminés pour les plus petites protéines, comme l'aspartame qui sucre vos sodas, à 30000 acides aminés pour la titine, une protéine qui assure l'élasticité de vos muscles. Il s'agit cependant d'exemples extrêmes, la plupart des protéines ne comprenant que quelques centaines d'acides aminés.

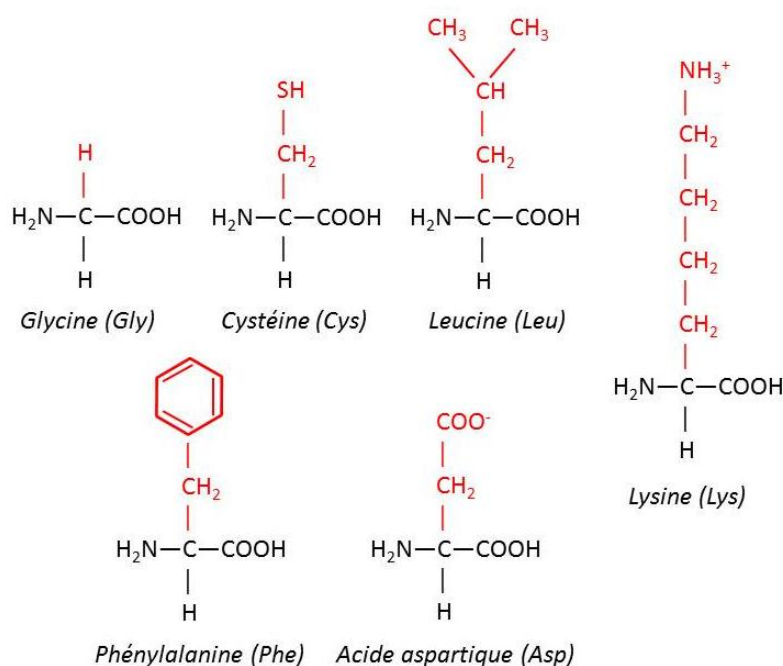
À ce jour, on ne sait pas précisément combien de protéines différentes existent. Chez l'homme, on avait évalué ce nombre à environ 20.000 il y a quelques années, puis 60.000, et il n'est à présent pas rare d'entendre des estimations dépasser les 100.000 protéines différentes. Mais une chose est certaine : les protéines représentent environ 60% d'une cellule, très loin devant les lipides, les glucides et les acides nucléiques comme l'ADN ou les ARN. En somme, un élément fondamental de la vie.

Vous avez dit, acide aminé ?

Acide à minet ? Miaouuu !

Plus sérieusement, les acides aminés sont les briques nécessaires à la construction de toutes les protéines. Il s'agit de petites molécules qui, mises bout à bout, forment les longues chaînes que sont les protéines. Comme leur nom l'indique, ces molécules possèdent à la fois un groupement acide ($-\text{COOH}$) et un groupement amine ($-\text{NH}_2$).

Il existe dans la nature une multitude d'acides aminés différents ; cependant, les êtres vivants n'en utilisent principalement que 20 pour construire les protéines.



Chaque acide aminé diffère des autres par une partie variable appelée « radical », représentée en rouge dans les exemples ci-dessus. Ce radical confère des propriétés chimiques spécifiques, influençant localement le repliement et les propriétés de la protéine.

On peut noter que, dans la nomenclature, on appelle rarement les acides aminés par leur nom complet. Ainsi, chaque acide aminé peut être représenté par trois lettres, voire par une seule. La glycine s'abrège « Gly » ou « G » ; la phénylalanine s'abrège « Phe » ou « F » ; etc...

Par exemple, voici comment peut s'écrire la séquence de la chaîne α de l'insuline humaine : GIVEQCCTSICSLYQLENYCN.

1 lettre	3 lettres	Nom complet	Occurrence dans les protéines humaines (%)
A	Ala	Alanine	7,8
C	Cys	Cystéine	1,9
D	Asp	Acide aspartique	5,3
E	Glu	Acide glutamique	6,3
F	Phe	Phénylalanine	3,9
G	Gly	Glycine	7,2
H	His	Histidine	2,3
I	Ile	Isoleucine	5,3
K	Lys	Lysine	5,9
L	Leu	Leucine	9,1
M	Met	Méthionine	2,3
N	Asn	Asparagine	4,3
P	Pro	Proline	5,2
Q	Gln	Glutamine	4,2
R	Arg	Arginine	5,1
S	Ser	Sérine	6,8
T	Thr	Thréonine	5,9
V	Val	Valine	6,6
W	Trp	Tryptophane	1,4
Y	Tyr	Tyrosine	3,2

Et ces acides aminés, d'où viennent-ils ?

Une grande partie des acides aminés est apportée tout naturellement par l'alimentation.

Lorsque vous mangez des protéines (viandes, poissons, œufs, légumineuses, ...), celles-ci sont dégradées au cours de votre digestion, par les sucs gastriques au niveau de l'estomac tout d'abord, puis surtout dans votre intestin grêle par des « *aminopeptidases* », des enzymes capables de casser les liaisons entre chaque acide aminé.

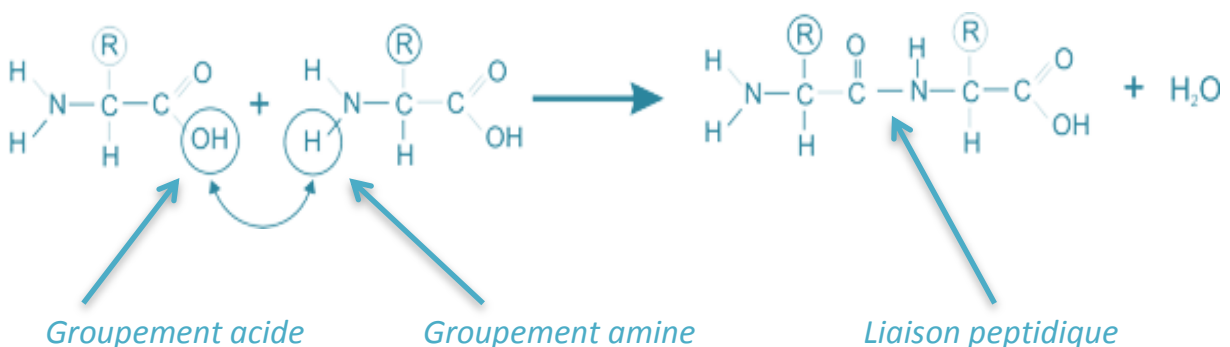
Les acides aminés libérés traversent ensuite la barrière intestinale et se retrouvent dans le sang, où ils pourront alimenter l'ensemble de l'organisme.

Parmi les 20 acides aminés, douze peuvent aussi être synthétisés directement par l'organisme. Cependant, huit d'entre eux sont dits « essentiels » car ils ne peuvent être synthétisés par l'organisme, d'où l'importance d'une alimentation équilibrée permettant un apport suffisant en protéines.

Casser les liaisons ? Mais quelles liaisons ?

Les liaisons peptidiques, pardi !

Au sein d'une protéine, les acides aminés sont reliés entre eux par une liaison appelée liaison peptidique. Il s'agit d'une liaison covalente qui lie le groupement acide d'un acide aminé avec le groupement amine de l'acide aminé suivant.



Et comment nos cellules synthétisent-elles les protéines ?

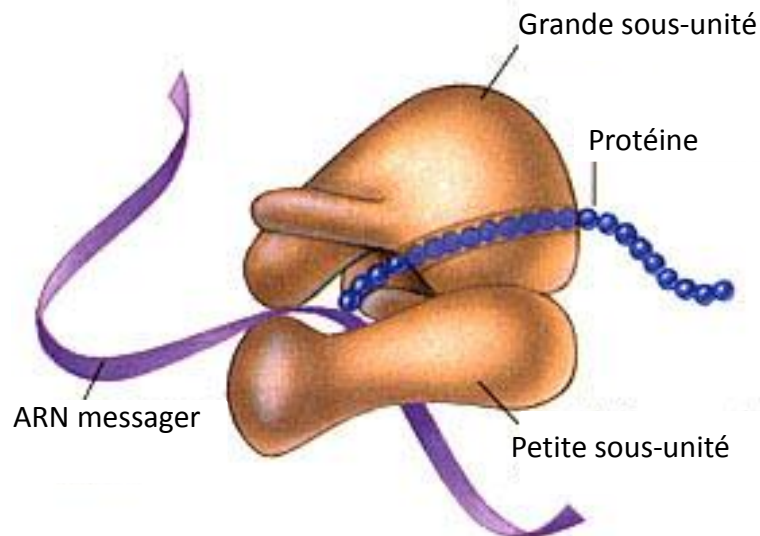
Chacune de nos cellules contient de l'ADN, qui est le « support » de l'information génétique. C'est à ce niveau que tout commence. Imaginons l'ADN comme un dictionnaire, contenant la définition d'une multitude de mots : les protéines. Malheureusement, le dictionnaire de l'ADN n'est pas écrit dans la même langue que les protéines... Mais rien n'arrête nos cellules !

Plus précisément, l'ADN est le support des gènes. Il existe différents types de gènes et beaucoup d'entre eux codent pour des protéines. La cellule doit donc décoder un gène donné (concrètement, un enchaînement de nucléotides A, T, C et G) pour produire la protéine voulue (un enchaînement d'acides aminés).

Dans un premier temps, le gène ciblé va être transcrit par des enzymes, les ARN polymérases, sous forme d'un ARN messager ou ARNm. Celui-ci va alors voyager depuis le noyau de la cellule vers le cytoplasme, à la rencontre de ribosomes, d'énormes molécules qui jouent le rôle de « traducteur ». Le ribosome est capable de lire l'ARNm, et de transformer sa séquence en une suite d'acides aminés, pour former la protéine.

L'ARN est chimiquement très semblable à l'ADN. Mais pourquoi alors ne pas traduire directement l'ADN ? Il y a plusieurs raisons à cela. Tout d'abord, l'ARN, contrairement à l'ADN, est très instable. En effet, un ARN ne survivra pas plus que quelques heures dans la cellule. Ainsi, lorsque la cellule arrête de produire un ARNm, il est détruit assez rapidement et la synthèse de protéine s'arrête. C'est donc un bon moyen de ne pas surproduire une protéine dont on a plus besoin. Ensuite, la production d'un intermédiaire entre le gène et la protéine est un bon moyen pour sauvegarder le « fichier original », l'ADN qui, s'il est dégradé ou muté, provoquera des conséquences plus ou moins graves sur la cellule.

Et comment fonctionne-t-il ce ribosome, précisément ?

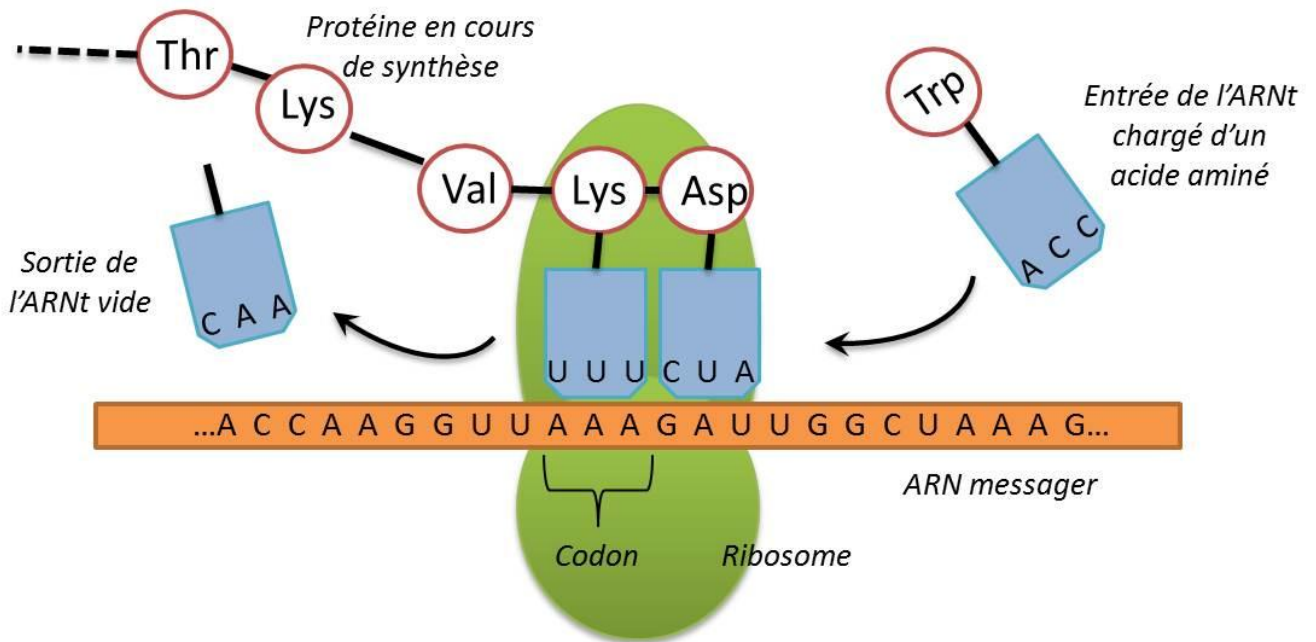


Le ribosome est un très gros assemblage de plusieurs protéines et ARN ribosomiaux (ou ARNr), composé de deux sous-unités : une grande et une petite. Ces deux sous-unités s'assemblent et forment, à leur interface, une « tête de lecture » capable de décoder l'ARN messager.

L'ARN messager est lu par série de trois nucléotides, qu'on appelle codon ou triplet, de telle sorte qu'à un codon donné correspond un acide aminé. C'est ce qu'on appelle le « code génétique ».

Ainsi, lorsque le ribosome lit l'enchaînement GGC (Guanine-Guanine-Cytosine), il sait qu'il devra insérer une glycine dans la protéine. Celle-ci lui sera apportée par l'intermédiaire d'un ARN de transfert (noté ARNt), spécifique pour cet acide aminé et ce codon. A chaque acide aminé peuvent correspondre plusieurs codons et un ou plusieurs ARNt.

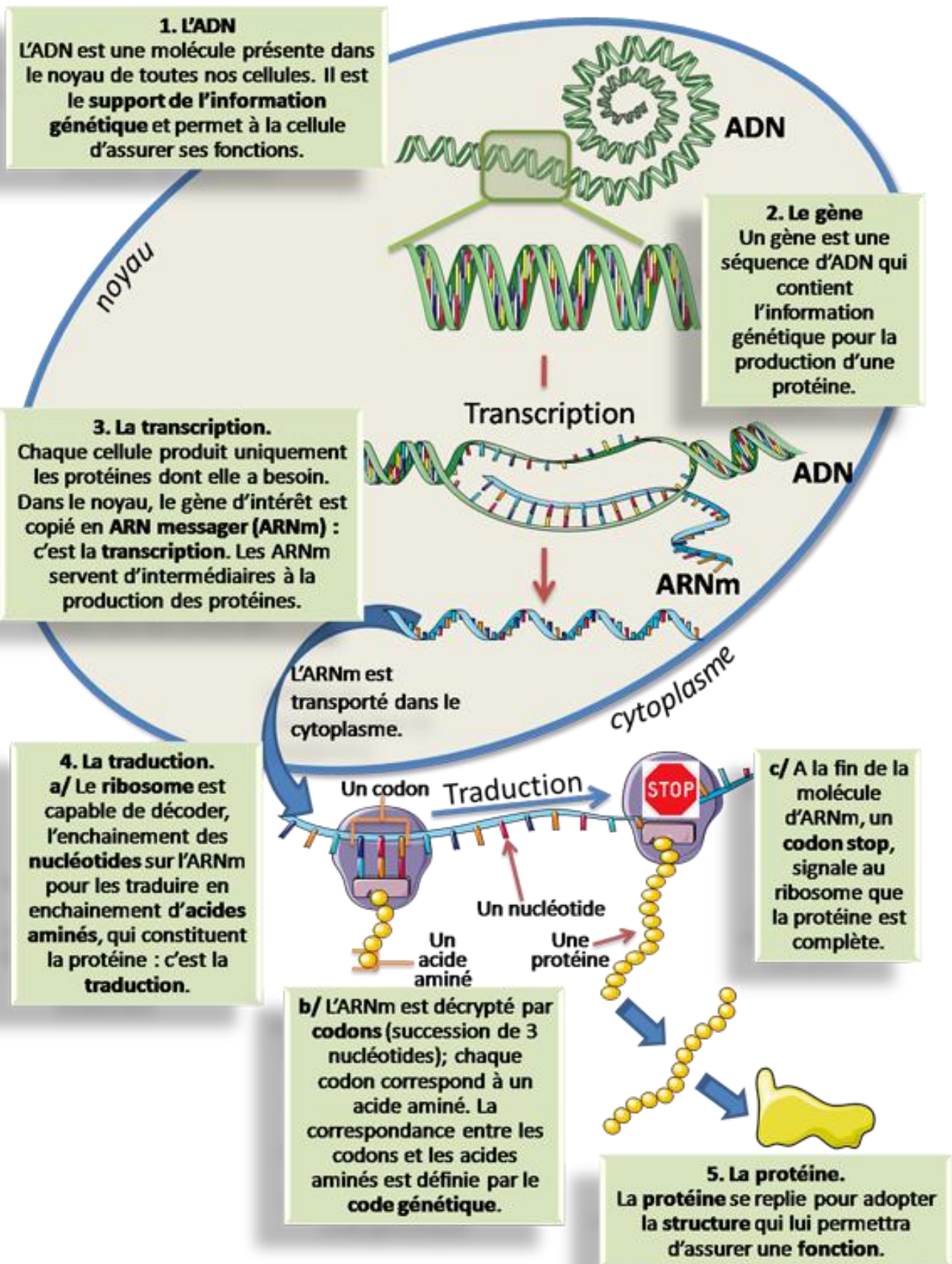
La fin de la traduction de l'ARN messager est dictée par la présence d'un codon spécial, appelé codon stop (classiquement il en existe 3, UAA, UAG et UGA), qui correspond simplement à un signal d'arrêt pour le ribosome car il n'existe aucun ARNt pouvant décoder ces codons. La protéine nouvellement formée est alors libérée dans la cellule.



Le code génétique

		Deuxième lettre									
		U		C		A		G			
Première lettre	U	UUU	Phénil-alanine	UCU	sérine	UAU	tyrosine	UGU	cystéine	U	Troisième lettre
		UUC	alanine	UCC		UAC	UGC	codon stop	C		
		UUA	leucine	UCA		UAA	codons	UGA	tryptophane	A	
		UUG		UCG		UAG	stop	UGG	G		
	C	CUU	leucine	CCU	proline	CAU	histidine	CGU	arginine	U	
		CUC		CCC		CAC	CGC	C			
		CUA		CCA		CAA	CGA	A			
		CUG		CCG		CAG	CGG	G			
	A	AUU	isoleucine	ACU	thréonine	AAU	asparagine	AGU	sérine	U	
		AUC		ACC		AAC	AGC	C			
		AUA	ACA	AAA		lysine	AGA	arginine	A		
		AUG	méthionine	ACG		AAG	AGG	G			
G	GUU	valine	GCU	alanine	GAU	acide	GGU	glycine	U		
	GUC		GCC		GAC	aspartique	GGC		C		
	GUA		GCA		GAA	acide	GGA		A		
	GUG		GCG		GAG	glutamique	GGG		G		

Le code génétique est dit « redondant », car un acide aminé peut être codé par plusieurs codons différents. Par exemple, les codons UUA, UUG, CUU, CUC, CUA et CUG correspondent tous les six à l'acide aminé leucine.



Un gène = Une protéine ?

Faux ! C'est ce qu'on a pendant longtemps cru, mais nous sommes conscients aujourd'hui qu'il existe beaucoup plus de protéines qu'il n'y a de gènes...

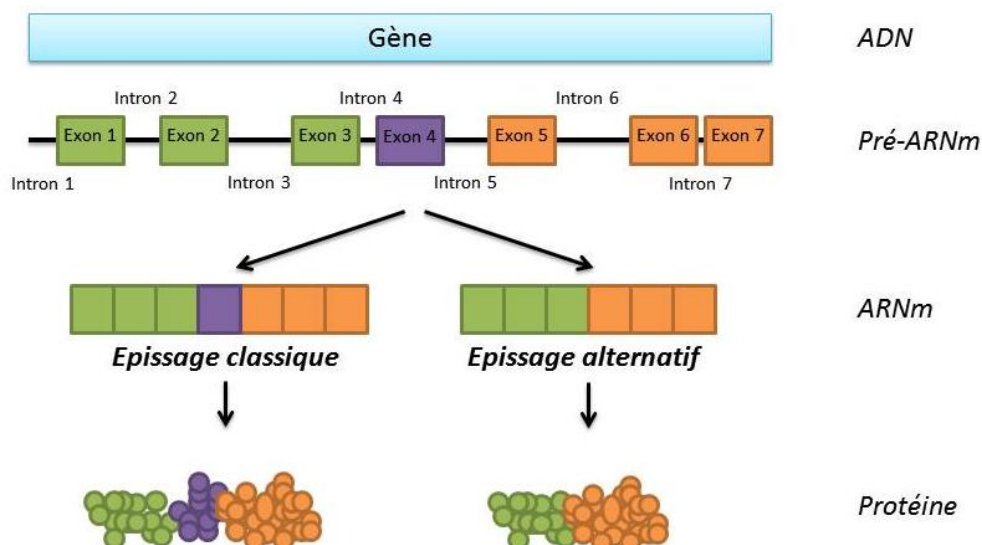
Mais comment cela est-il possible ? En réalité, l'équation à retenir est plutôt du type : un gène = plusieurs ARNm = plusieurs protéines. Et cette équation est rendue possible par un mécanisme : l'épissage alternatif.

L'épissage, c'est quoi ? Et l'épissage « alternatif »... comment ça marche ?

Un gène, sur toute sa longueur, est en réalité composé de parties dites codantes (les exons) et de parties non codantes (les introns). Pour former une protéine, il ne faut conserver que les exons et supprimer les introns. C'est ce qu'on appelle l'épissage. Imaginez une phrase. L'essentiel de l'information réside dans les mots, et non dans les espaces ou la ponctuation. Ces derniers ne sont présents que pour nous aider à lire l'information plus simplement. Mais on pourrait tout à fait s'en passer. C'est ce que la cellule fait avec l'épissage !

Mais parfois, au lieu de ne supprimer que les espaces, la machinerie cellulaire supprime également un mot au milieu de la phrase. L'information finale est donc différente, même si certains éléments demeurent les mêmes.

Revenons maintenant à la notion « d'alternatif ». C'est comme le courant, une fois il passe, une fois non. Pour un ARNm, une fois il y a épissage d'un intron, une fois non. C'est ainsi que deux protéines différentes peuvent être formées à partir d'un plan de montage initial identique.



Et à quoi ça ressemble, une protéine ?

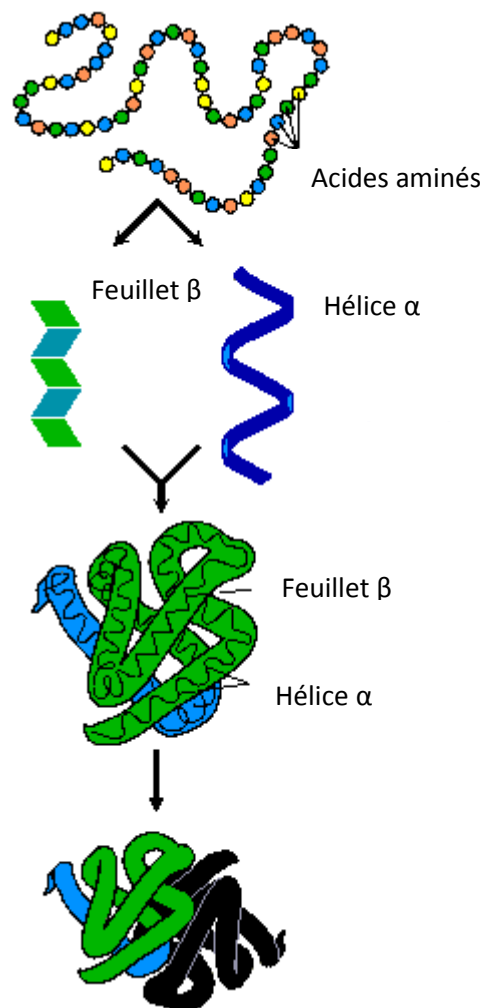
Comme nous l'avons dit, une protéine est une succession d'acides aminés. Cependant, une protéine ne se résume pas à simple un collier de perle, et toutes les protéines ne se ressemblent pas... Elles possèdent toutes une structure caractéristique, c'est-à-dire une forme dans l'espace.

On définit quatre niveaux de structure dans une protéine. La structure primaire est le niveau le plus simple. Il s'agit simplement de la séquence en acides aminés. Le collier de perles.

Puis la structure secondaire correspond à des repliements locaux au sein du collier. Certaines portions peuvent prendre la forme d'une hélice α , comme le câble d'un téléphone, ou d'un feuillet β , où les acides aminés forment une surface plane.

Au-delà, ces différentes structures peuvent interagir les unes avec les autres, par le biais de liaisons. C'est la structure tertiaire. Ainsi, l'hélice N°2 pourra par exemple former une liaison hydrogène avec l'hélice N°7, forçant l'ensemble de la protéine à se plier de manière à mettre en contact ces deux partenaires. D'autres types de liaisons peuvent intervenir : les liaisons ioniques, les interactions hydrophobes, ou les ponts disulfures.

Pour finir, la structure quaternaire est observée dans certaines protéines comme l'hémoglobine, où plusieurs chaînes peptidiques, plusieurs « sous-protéines » doivent s'assembler pour former un grand complexe protéique.



Une structure = une fonction.

La fonction d'une protéine découle directement de sa structure tertiaire et/ou quaternaire. Ce n'est que dans cet état replié que la protéine pourra assurer son activité biologique. On parle d'état natif. Connaître la structure des protéines, à l'atome près, et pouvoir suivre leur changement de conformation en fonction de la présence d'un partenaire et des changements environnementaux est donc d'importance, puisqu'on pourra ainsi comprendre plus aisément leur mécanisme d'action.

L'état natif d'une protéine fait opposition à son état dénaturé, dans lequel la structure tertiaire est irréversiblement perdue (suite à l'action de la chaleur ou d'agents chimiques par exemple). Cette notion de dénaturation est importante : chauffez de l'ADN, il va se dénaturer mais aussitôt que la température baissera, il va pouvoir se « renaturer ». Ce n'est pas le cas des protéines. Prenez un œuf. Lorsque vous le cassez, le blanc est transparent, visqueux. Mettez-le dans une poêle avec une noisette de beurre, un peu de sel et de poivre. Le blanc va devenir opaque et dur. Vous avez fait un œuf au plat ! Et accessoirement, vous avez dénaturé la protéine de l'œuf : l'albumine. Bon appétit ! ... mais impossible de revenir en arrière, si vous essayez de faire un flan en utilisant votre œuf dur, laissez tomber et allez plutôt chez le pâtissier du coin !

Généralement, les acides aminés exposés à la surface de la protéine, en contact avec le milieu environnant, sont des acides aminés importants pour la fonction de la protéine. Ce sont eux qui vont être capable d'interagir avec des partenaires (d'autres protéines, de l'ADN, des ARN, des lipides, etc...) pour assurer leur fonction.

Schématiquement, leur organisation va permettre de former des bosses et des creux, complémentaires de leurs partenaires. Comme une serrure sera complémentaire de la clef qui lui correspond. Une protéine a donc une fonction propre, mais aussi très souvent, une spécificité de substrat (la molécule avec laquelle elle va et doit interagir).

Concrètement, des protéines, pour quoi faire ??

Les protéines assurent de très nombreuses fonctions au sein d'un organisme.

Elles peuvent ainsi avoir une fonction structurale, permettant à la cellule de maintenir son organisation spatiale et sa forme (le collagène par exemple).

Elles peuvent également avoir un rôle de transporteur, comme l'hémoglobine qui transporte le dioxygène, la transferrine qui transporte le fer, etc... Il peut également s'agir de canaux, des protéines qui se trouvent dans la membrane des cellules et qui y forment un tunnel, permettant à certaines molécules d'entrer ou de sortir de la cellule, par exemple les porines. Comme leur nom l'indique, ces protéines forment des pores, des trous, et permettent l'accès entre deux compartiments cellulaires, comme le tunnel du Mont-Blanc entre la France et l'Italie.

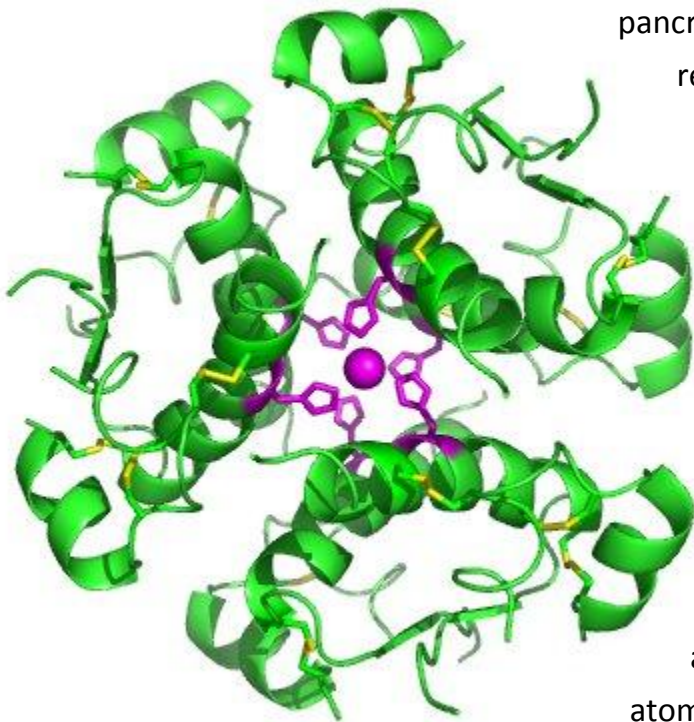
Il peut s'agir de protéines motrices, permettant aux muscles de se contracter, aux spermatozoïdes de se déplacer, ou encore aux chromosomes de se séparer lors de la mitose. Des protéines telles que l'actine ou la myosine seront responsables de cette fonction.

Les protéines de signalisation assurent la communication au sein de l'organisme, entre les cellules et à l'intérieur d'une cellule. On peut citer l'exemple des récepteurs hormonaux, des protéines qui, au contact d'une hormone, changent de structure pour induire une cascade de réponses.

Certaines protéines portent une fonction immunitaire, c'est-à-dire qu'elles sont capables de reconnaître la présence d'organismes étrangers et d'alerter l'organisme en cas d'infection. Il s'agit des immunoglobulines, plus connues sous le nom d'anticorps.

Enfin, les protéines peuvent avoir une fonction de catalyseur chimique. On parle alors d'enzyme. Elles sont capables de transformer des substrats, de les dégrader, ou encore d'en assembler plusieurs pour créer un nouveau produit.

Quelques exemples de structures

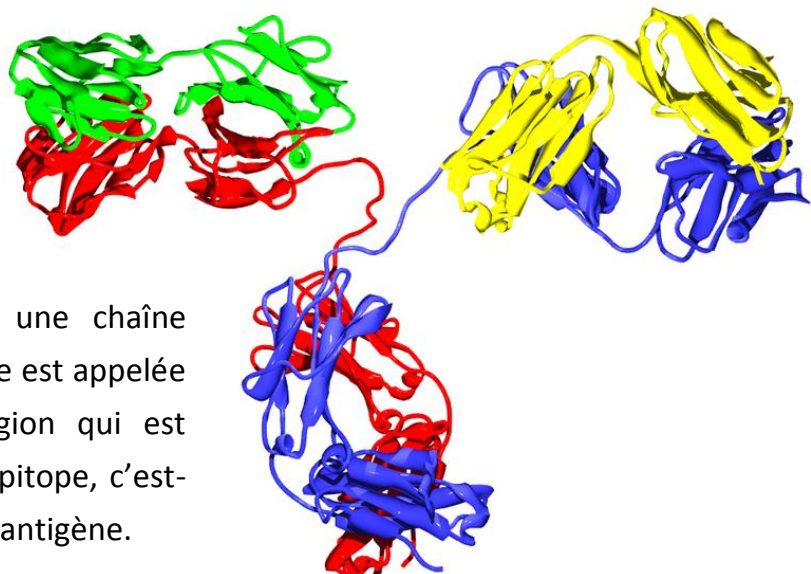


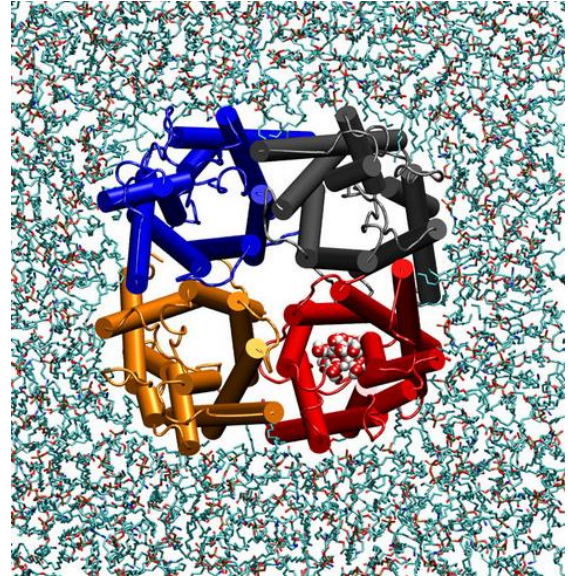
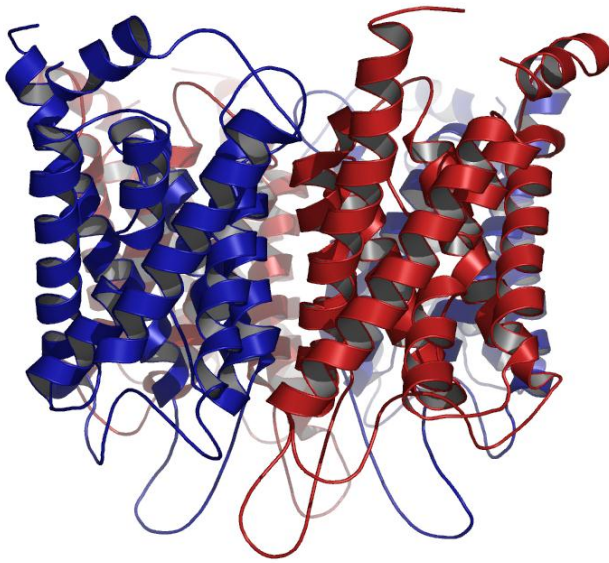
L'insuline, une hormone produite par le pancréas, est responsable de la régulation du taux de glucides dans le sang. Lorsqu'elle est stockée, l'insuline s'assemble par six pour former un hexamère. Lorsqu'elle est active en revanche, elle agit sous forme isolée ou monomère.

On voit ici un hexamère d'insuline, les six protéines étant liées ensemble par l'intermédiaire d'un acide aminé histidine coordonnant un atome de zinc (en violet).

Les immunoglobulines G (IgG) constituent un isotype d'immunoglobulines ou « anticorps ». Il s'agit du type d'anticorps le plus représenté dans le sang. Ils sont produits en réponse à la détection de molécules étrangères dans notre organisme.

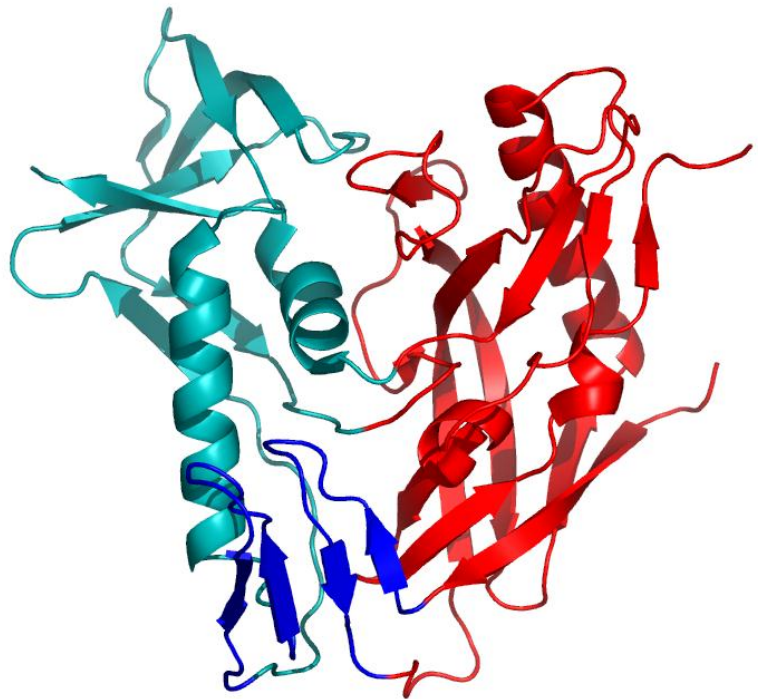
Ici, l'IgG est composée de quatre chaînes : deux dites lourdes (en rouge et bleu) et deux dites légères (en vert et jaune). L'interface entre une chaîne lourde et une chaîne légère est appelée paratope. C'est cette région qui est capable de reconnaître l'épitope, c'est-à-dire le déterminant d'un antigène.

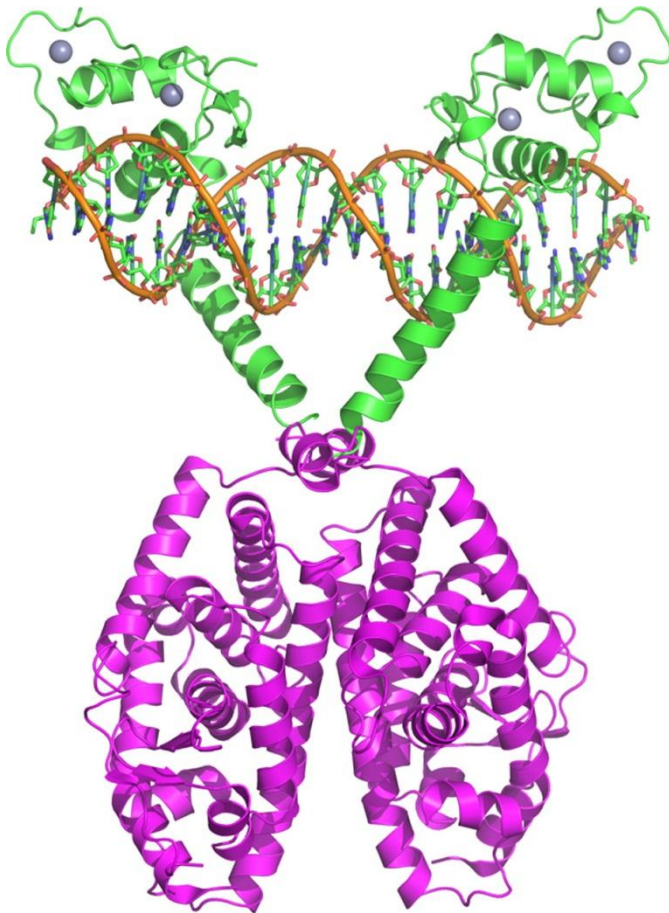




Une aquaporine est une protéine membranaire, c'est-à-dire une protéine incluse à l'intérieur de certaines membranes d'une cellule. Sa particularité est de former un canal ouvert, perméable aux molécules d'eau, comme le montre la structure vue de dessus à droite (aquaporine entourée de lipides formant la membrane).

La protéine gp120 est un des constituants du VIH, le virus responsable du SIDA. Cette protéine se trouve à la surface du virus et permet la liaison du virus aux cellules cible (via les récepteurs CD4 présents à la surface de nos cellules), provoquant ainsi l'infection.

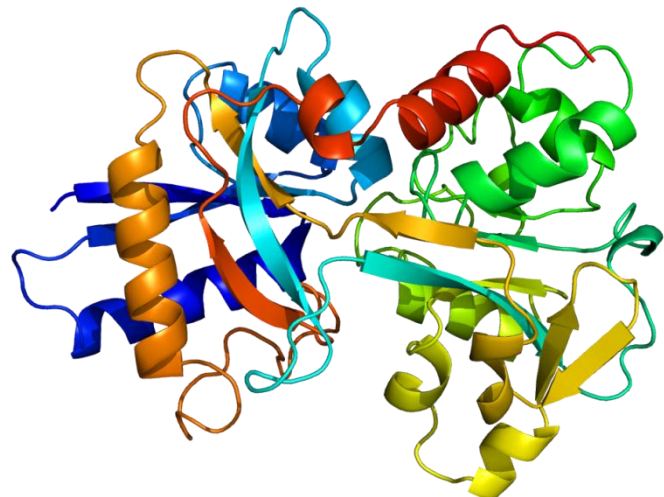




Les récepteurs des hormones thyroïdes sont des protéines capables de lier les hormones thyroïdiennes comme la T3 ou la T4. Ces récepteurs jouent un rôle dans le métabolisme et le contrôle du rythme cardiaque. Ils se lient à l'ADN (représenté par l'hélice en orange) pour réguler l'expression de certains gènes. En absence d'hormone, le récepteur inhibe l'expression de ces gènes, empêchant l'ARN polymérase de se fixer sur le promoteur. En revanche, en présence d'hormone thyroïdienne, le récepteur change de forme, et laisse la place à l'ARN polymérase. Le gène peut alors être

transcrit et la protéine synthétisée. Ce type de mécanisme est très courant, et permet de réguler l'expression des protéines à un moment donné.

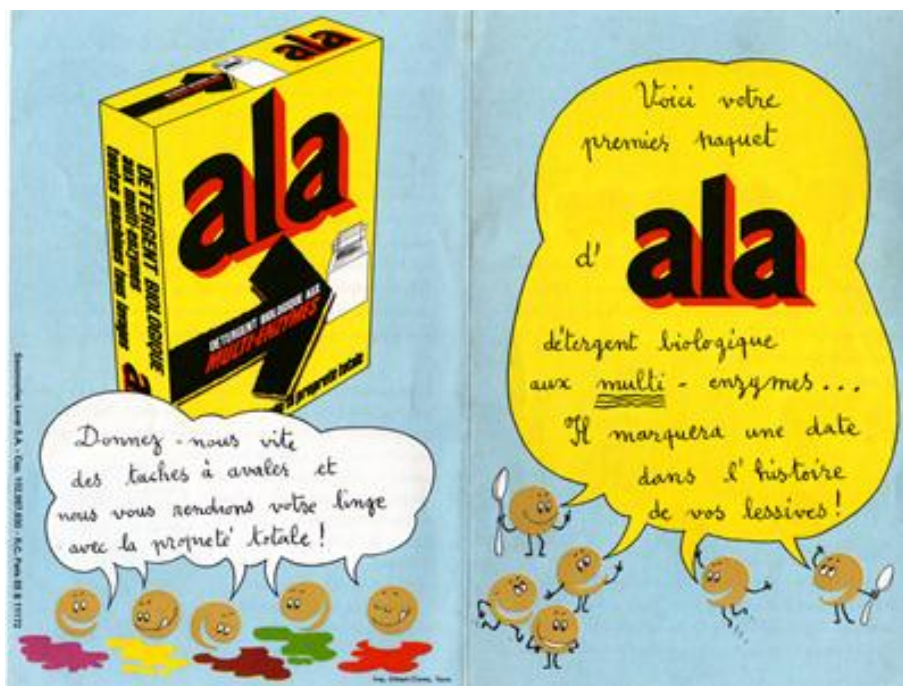
La transferrine est une protéine qui est capable de lier des atomes de fer et de les transporter dans l'organisme. Une autre protéine, le récepteur de la transferrine, reconnaît les transferrines chargées en fer et les dirige vers les cellules hépatiques, qui servent de lieu de stockage.



Un mot sur les enzymes

Tout le monde a déjà entendu ce nom quelque part, mais savez-vous ce qu'est exactement une enzyme ?

En 1969, la compagnie Unilever commercialise une lessive qui se veut révolutionnaire ! Elle contient des enzymes qui dévorent les tâches : les enzymes glutons... Mais Unilever fait une grosse erreur de marketing en représentant les enzymes comme des petites bêtes qui mangent tout sur leur passage... Les clients n'adhèrent pas, de peur que ces bêtêtes dévorent leur linge !



Une erreur marketing, et une erreur biologique !! Car les enzymes ne sont pas des êtres vivants... Oubliez les bêtêtes, les microbes... Il s'agit « simplement » de protéines ! Ou peut-être pas si « simplement » en fait. Car ces protéines sont très spécialisées, et leur fonction hautement spécifique.

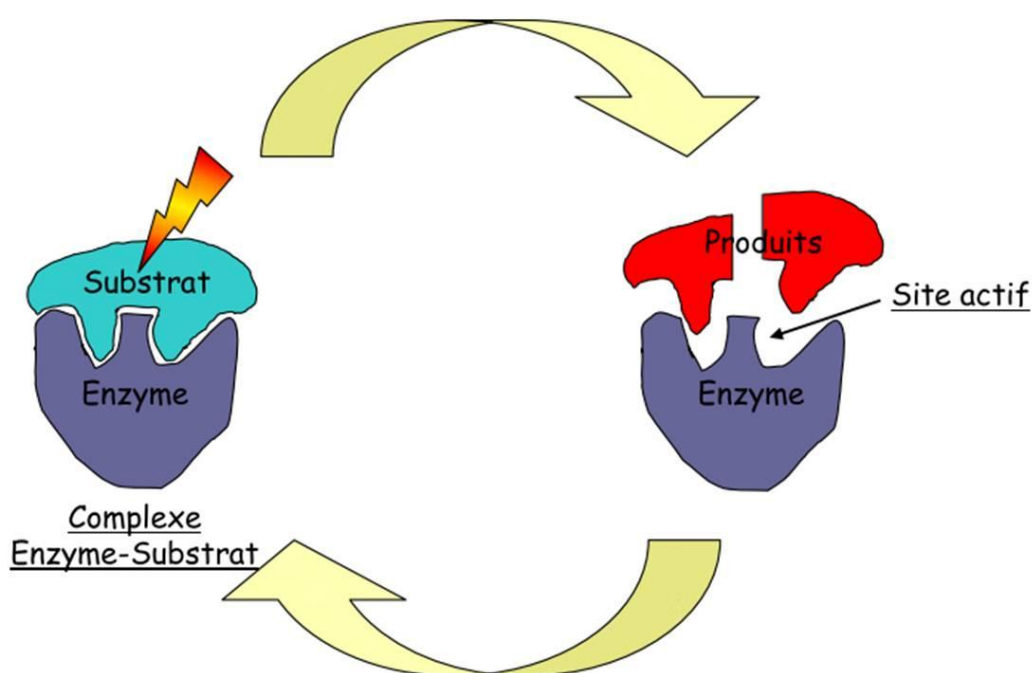
Les enzymes sont ce qu'on appelle des biocatalyseurs : elles sont capables en quelques secondes, voire moins, de catalyser des réactions chimiques qui sans elles, auraient pris des heures, des jours voire carrément des années ! Et tout ça, à 37°C, dans vos cellules...

Comment fonctionne une enzyme ?

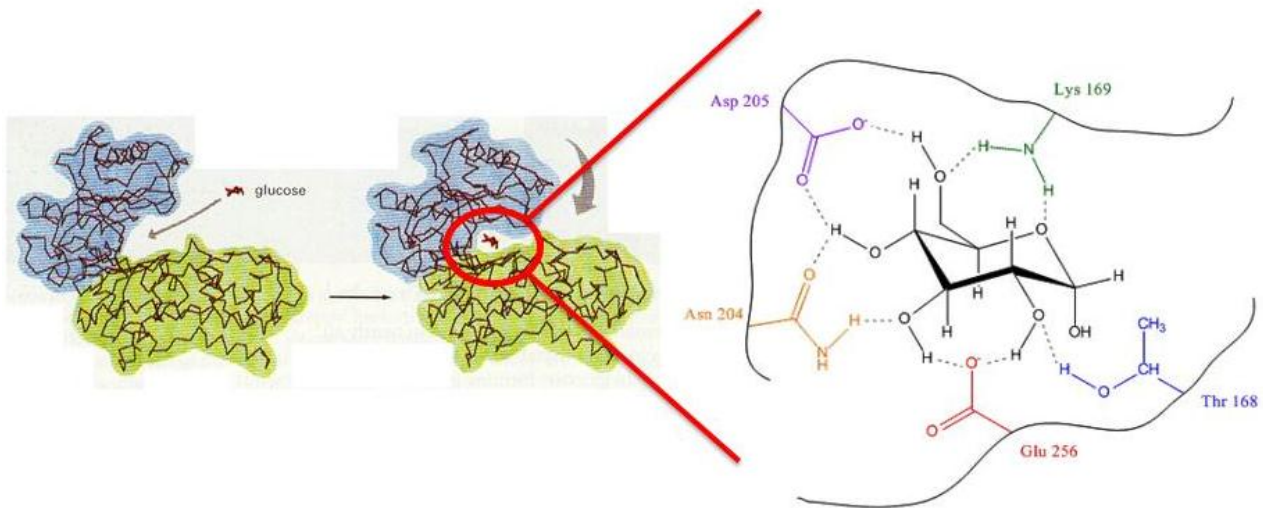
Comme expliqué précédemment, la structure tertiaire d'une protéine assure sa fonction. Ceci est d'autant plus vrai pour une enzyme, puisque son repliement en trois dimensions permet la formation d'un « site actif ». Classiquement, il s'agit d'un petit nombre d'acides aminés qui s'organisent de manière à former une cavité dans la protéine, dans laquelle le substrat va venir se loger.

Les acides aminés exposés dans le site actif confèrent à l'enzyme sa spécificité de substrat. Par liaisons hydrogènes, ioniques ou interactions hydrophobes, l'enzyme va être capable de reconnaître « uniquement » son substrat et de faire la différence entre deux molécules, même proches. Le même type de complémentarité que l'on observe pour une serrure et sa clef...

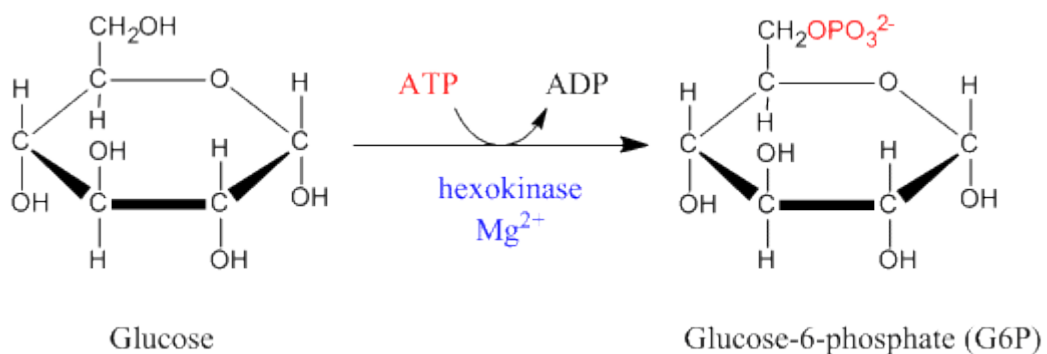
Cette complémentarité est essentielle pour permettre la catalyse de la réaction. Ainsi, le site actif de l'enzyme fournit tous les éléments nécessaires pour accélérer la réaction : il permet de rapprocher et d'orienter des substrats qui n'auraient naturellement pas été « attiré » l'un vers l'autre ; il fournit des accepteurs ou donneurs de protons, comme lors d'une réaction acido-basique ; ou encore, « prête » transitoirement des électrons pour accélérer une réaction.



Plus concrètement, voici un exemple d'enzyme, l'hexokinase, qui transforme le glucose en glucose-6-phosphate (la première étape de la réaction de glycolyse).



Observez comme la molécule de glucose (en noir) est immobilisée, comme prise au piège, à l'intérieur du site actif, par plusieurs liaisons avec les acides aminés de l'enzyme (en couleur). On peut aussi noter que les acides aminés exposés dans le site actif ne se suivent pas dans la séquence : on trouve la thréonine N° 168, la lysine N° 169, l'asparagine N° 204, l'acide aspartique N° 205 et l'acide glutamique N° 256 (en abrégé : T168-K169-N204-C205-E256). Ceci est la conséquence du repliement de la protéine, de sa structure tertiaire : des résidus éloignés dans la séquence peuvent se retrouver l'un à côté de l'autre en trois dimensions.



Quelques exemples d'enzyme

L'amylase est une enzyme produite par le pancréas et sécrétée dans l'intestin grêle. Son rôle est de dégrader les glucides complexes tels que l'amidon, en sucres simples comme le glucose ou le maltose, assimilables par l'organisme.

La lactase est une enzyme très utilisée dans l'industrie alimentaire. Elle permet de transformer le lactose en glucose et galactose. Ceci permet donc de supprimer le lactose des produits laitiers, certaines personnes y étant intolérantes.

Les hélicases sont des enzymes présentes dans nos cellules. Elles permettent de dénaturer le repliement en double hélice des acides nucléiques comme l'ADN, c'est-à-dire de séparer les deux brins. Cette réaction est très importante pour permettre la réplication de l'ADN par exemple. Sans ces enzymes, il faudrait chauffer notre ADN à plus de 90°C pour pouvoir le dénaturer...

La transcriptase inverse est une enzyme que nous trouvons chez les virus à ARN, tels que le VIH (Virus de l'Immunodéficience Humaine). Contrairement à nous, ces virus ne possèdent pas d'ADN, mais un ARN. Grâce à la transcriptase inverse, cet ARN va être « rétrotranscrit » en ADN pour infecter les cellules hôte.

L'alcool déshydrogénase est une enzyme produite par le foie, qui permet de dégrader les alcools comme l'éthanol. Cette enzyme est donc d'importance puisqu'elle permet de détoxifier l'organisme.

Les enzymes de restriction sont présentes dans les bactéries et clivent l'ADN de manière spécifique selon la séquence en nucléotides. Elles sont très utilisées au laboratoire en biologie moléculaire.