

## 4. La diversité des molécules organiques : du simple au très complexe

Lorsqu'on parle de la chimie de la vie, mais aussi des molécules que les astronomes découvrent dans le milieu interstellaire ou à la surface de certaines planètes, on entend souvent les mots « **molécules organiques** ». Souvent, on pense à tort que « organique » sous-entend l'existence de la vie, mais **il n'en est rien**. En réalité, même si la vie se fonde sur la fabrication, l'utilisation, et la transformation de molécules organiques, de tels composés ne présupposent pas du tout l'existence de la vie. En fait, l'adjectif « organique » renvoie bien à l'idée de « vivant », mais cette désignation remonte au XIX<sup>e</sup> siècle, lorsque les chimistes pensaient encore que les molécules qui composent les organismes vivants se différenciaient de celles trouvées « ailleurs » (principalement dans le monde minéral) par l'intervention d'une *force vitale*, supposée leur donner leurs propriétés. Aujourd'hui, on sait qu'il n'en est rien, et que les molécules « minérales » et « organiques » sont construites selon des principes identiques. Aujourd'hui, les molécules (ou composés) organiques désignent *tout assemblage d'atomes (molécules) construites principalement autour du carbone*. Réciproquement, *toute molécule qui contient du carbone est une molécule organique*. Vous savez à présent ce qu'est un atome, et, mieux encore, ce qu'est une molécule. Vous allez maintenant apprendre ce que sont les molécules organiques, et leurs principales caractéristiques. Ces nouvelles informations nous permettront bientôt de comprendre comment fonctionne le vivant.

### 1. Du carbone, du carbone, et encore du carbone.

Une molécule organique contient du carbone. Nous avons déjà rencontré cet atome (voir **Chap. 1**), et vous savez qu'il peut former jusqu'à quatre liaisons simples avec d'autres atomes (**Fig. 1**). De la même façon, il peut former une liaison double et deux liaisons simples, deux liaisons doubles, ou encore une liaison triple et une liaison simple. Cette plasticité rend le carbone particulièrement intéressant, car il conduit à l'obtention de molécules parfois très grandes, et surtout très complexes. « Jusqu'à quatre liaisons » signifie qu'on peut greffer au carbone à peu près n'importe quelle combinaison d'atomes (ou groupes d'atomes). Par exemple, quatre hydrogènes, pour former le méthane (CH<sub>4</sub>) (**Fig. 1**), ou encore trois atomes de chlore et un hydrogène (CHCl<sub>3</sub>) pour produire le chloroforme. Les possibilités sont infinies.

Lorsqu'une molécule est constituée uniquement de carbone et d'hydrogène, on dit qu'il s'agit d'un **hydrocarbure** (*hydrogène* et *carbone*). L'éthane, le propane, le butane, ainsi que le pentane et l'hexane sont tous des hydrocarbures. Comme les carbones sont « alignés » (**Fig. 1**), on dit que ce sont des molécules *linéaires*. Par ailleurs, en chimie, on nomme les hydrocarbures en fonction du nombre d'atomes de carbone qu'ils contiennent. Ainsi, on utilise les préfixes meth-, eth-, pro-, but-, pent-, hex-, hept-, oct-, non- et dec- pour signifier qu'une molécule contient 1, 2, 3, 4, 5, 6, 7, 8, 9 et 10 atomes de carbone respectivement. Lorsqu'il s'agit d'une molécule ne contenant que du carbone et de l'hydrogène, et rien d'autre, on dit qu'il s'agit d'un **alcane**, et on ajoute la terminaison -ane. On obtient ainsi le méthane, l'éthane, le propane, le butane, le pentane, l'hexane, l'heptane, l'octane, le nonane et le décane (**Fig. 1**).

Bien entendu, une molécule organique composée uniquement de carbone et d'hydrogène n'est pas toujours linéaire, mais peut être cyclique, c'est-à-dire refermée sur elle-même. Il ne s'agit pas, dans ce cas d'un alcane à

proprement parler. Pour obtenir une molécule cyclique, il faut que le nombre d'atomes de carbone soit suffisamment grand pour leur permettre de se refermer. On peut ainsi parfois rencontrer le cyclohexane (**Fig. 1**).

Valence	Liaisons simples	Liaisons doubles	Liaisons triples
	 méthane (CH <sub>4</sub> )	 éthène (C <sub>2</sub> H <sub>4</sub> )	 acétylène (C <sub>2</sub> H <sub>2</sub> )

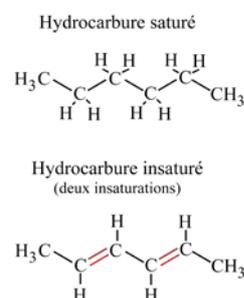
Hydrocarbures linéaires		Hydrocarbures cycliques
①	CH <sub>4</sub> méthane	
②	H <sub>3</sub> C—CH <sub>3</sub> éthane	
③	propane	
④	butane	
⑤	pentane	
⑥	hexane	

**Figure 1 : Molécules organiques.** Elles sont construites autour d'un ou plusieurs atomes de carbone (quatre valences), qui peuvent permettre plusieurs types de liaisons). La partie basse de la figure présente quelques exemples de molécules organiques simples (uniquement carbone et hydrogène).

Les hydrocarbures sont très abondants sur Titan, la plus grande lune de Saturne. On a ainsi détecté du méthane et de l'éthane, ainsi que d'autres composés organiques plus complexes. De même, les comètes contiennent un certain nombre de molécules organiques.

## 2. Des décorations infinies.

Le cas des alcanes (molécules d'hydrocarbures linéaires) est simple. Il suffit d'enchaîner des atomes de carbones entre eux par des liaisons **simples**, et leur ajouter, à chacun, le nombre adéquat d'hydrogènes (3 aux extrémités, et deux à l'intérieur des chaînes) (**Fig. 1**). Si c'est le cas, on dit qu'il s'agit d'un hydrocarbure **saturé**. Or, dans la nature, on peut trouver des molécules constituées d'enchaînements de carbones, mais qui, parfois, comportent des liaisons doubles. Ces molécules sont dites **insaturées** (**Fig. 2**).



**Figure 2 : Les hydrocarbures** peuvent être saturés, c'est-à-dire constitués d'un enchaînement de carbones (ici 6) reliés entre eux par des liaisons simples uniquement, ou insaturés, c'est-à-dire comportant des carbones reliés entre eux par des liaisons doubles (en rouge).

Dans d'autres cas encore, on rencontre des molécules hydrocarbonées, mais dont les atomes de carbone sont reliés à des atomes – ou groupes d'atomes – qui ne sont pas des hydrogènes. Un certain nombre d'exemples sont représentés sur la **Fig. 3**. On peut ainsi ajouter un groupe  $\text{-NH}_2$  (un azote, deux hydrogènes), encore appelé « groupe amine », ou un groupe  $\text{-OH}$  (un oxygène et un hydrogène), appelé groupe « hydroxyle » ou « alcool », ou un groupe  $\text{-COOH}$  (un carbone, deux oxygènes, un hydrogène), dit « groupe acide, ou carboxyle », et tout un tas d'autres groupements. On les appelle **groupements fonctionnels**. Il en existe des centaines (**Fig. 3**).

On peut trouver des groupements fonctionnels seuls ou dans différentes combinaisons au sein de molécules organiques. Par exemple, une chaîne de seize carbones comportant un groupe « carboxyle » ( $\text{-COOH}$ ) à l'une de ses extrémités représente l'acide palmitique, un acide gras trouvé dans les cellules. De manière générale, toute chaîne d'hydrocarbure comportant un groupe carboxyle à son extrémité est un **acide gras**. Si les carbones sont saturés (uniquement des liaisons simples), c'est un acide gras saturé, mais s'il existe un ou plusieurs liaisons doubles, on dit que c'est un **acide gras insaturé** ou **poly-insaturé** (**Fig. 4**). Si une molécule d'éthane (**Fig. 1**) est « décorée » d'une fonction hydroxyle ( $\text{-OH}$ ) (**Fig. 4**), elle devient l'éthanol (suffixe « -ol »). Toute molécule comportant un hydroxyle  $\text{-OH}$  fait partie de la famille des alcools. L'éthanol est la molécule présente dans les boissons alcoolisées.

La diversité des molécules organiques est gigantesque. On ne peut pas rentrer dans les détails, c'est pourquoi j'ai résumé quelques exemples de groupements fonctionnels et de molécules qui les contiennent dans les **Fig. 1 à 4**. Ce sont les groupements que nous rencontrerons le plus souvent dans les molécules organiques utilisées et produites par les organismes vivants.

Groupe	Nom	Nom «courant»
$\text{R}-\text{OH}$	Hydroxyle	Alcool
$\text{R}-\text{C}(=\text{O})\text{H}$	Aldéhyde	Aldéhyde
$\text{R}-\text{C}(=\text{O})\text{OH}$	Carboxyle	Acide
$\text{R}-\text{C}(=\text{O})\text{NH}_2$	Amide	Amide
$\text{R}_1-\text{C}(=\text{O})-\text{R}_2$	Carbonyle	Cétone
$\text{R}-\text{NH}_2$	Amine	Amine
$\text{R}-\text{C}_6\text{H}_5$	Phényle	Phényle/Benzène
$\text{R}-\text{C}_6\text{H}_4\text{OH}$	Phénolyle	Phénol
$\text{R}=\text{NH}$	Imine	Imine
$\text{R}-\text{P}(=\text{O})(\text{O}^-)_2$	Phosphoryle	Phosphate
$\text{R}-\text{S}(=\text{O})(\text{O}^-)_2$	Sulfuryle	Sulfate
$\text{R}-\text{C}(=\text{O})\text{CH}_3$	Acétyle	Acétate

**Figure 3 : Groupement fonctionnel (ou « décorations »)** les plus courantes rencontrées dans les molécules organiques. Dans ce tableau, la lettre R représente n'importe quelle molécule organique sur laquelle le groupement représenté peut être greffé. Dans chaque cas, le nom du groupement est indiqué, ainsi que son nom « courant ». Dans le cas des groupements « phényle » et « phénolyle », il s'agit d'un cycle à 6 carbones *aromatique* (voir **Chap. 2, § 2.11**). Tous ces groupements fonctionnels sont extrêmement courants dans les molécules biologiques. Remarquez qu'ils peuvent contenir de l'oxygène (O), de l'azote (N), du phosphore (P) et même du soufre (S).

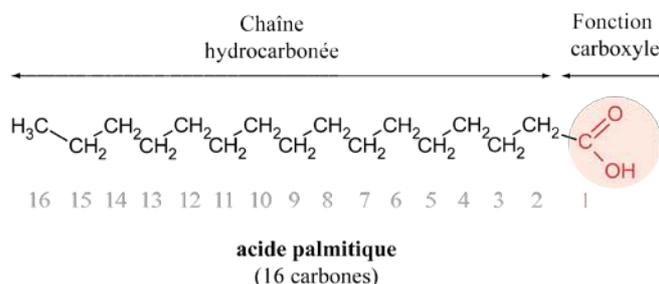
(A)

Nombre de carbones	Racine du nom	①	②	③	④	⑤
		Alcane simples	Alcool simple	Aldéhyde simple	Acide	Amine simple
1	meth-	méthane	méthanol	méthanal	méthanoïque	aminométhane
2	eth-	éthane	éthanol	éthanal	éthanoïque	aminoéthane
3	prop-	propane	propanol	propanal	propanoïque	aminopropane
4	but-	butane	butanol	butanal	butanoïque	aminobutane
5	pent-	pentane	pentanol	pentanal	pentanoïque	aminopentane
6	hex-	hexane	hexanol	hexanal	hexanoïque	aminohexane
7	hept-	heptane	heptanol	heptanal	heptanoïque	aminoheptane
8	oct-	octane	octanol	octanal	octanoïque	aminooctane
9	non-	nonane	nonanol	nonanal	nonanoïque	aminononane
10	dec-	decane	decanol	decanal	decanoïque	aminodécane

(B)

Nombre de carbones	Alcane simples	Alcool simple	Aldéhyde simple	Acide simple	Amine simple
	①	②	③	④	⑤
1	CH <sub>4</sub>	CH <sub>3</sub> —OH			CH <sub>3</sub> —NH <sub>2</sub>
2	H <sub>3</sub> C—CH <sub>3</sub>	H <sub>3</sub> C—CH <sub>2</sub> —OH			H <sub>3</sub> C—CH <sub>2</sub> —NH <sub>2</sub>

(C) Exemple d'acide gras



**Figure 4 : Notions de nomenclature.** (A) Pour nommer les molécules hydrocarbonées, on utilise les mêmes notions que pour nommer les alcanes (Fig. 1). Des molécules à 1, 3, 3, 4, ... 10 atomes de carbone sont nommées à partir des racines meth-, eth-, prop-, ... dec-, etc... Si l'y a que des liaisons simples, et uniquement des carbones et des hydrogènes, ce sont des alcanes, et on parle de **methane**, **ethane**, **propane** (1), etc... Si l'on ajoute une fonction « hydroxyle », on doit ajouter le suffixe -ol, et on obtient le **méthanol**, l'**éthanol**, le **propanol** (2), etc... En ajoutant une fonction « aldéhyde », on obtient **methanal**, **ethanal**, **propanal** (3), et autres dérivés (suffixe -al). Dans le cas de la fonction « acide » (ou encore appelée « carboxyle »), on obtient l'**acide méthanoïque**, l'**acide éthanoïque**, l'**acide propanoïque** (4), etc... Pour finir, en ajoutant une fonction « amine », on obtient l'**aminométhane**, l'**aminoéthane**, ou l'**aminopropane** (5). (B) La partie B de la figure montre quelques exemples avec le cas du méthane (méthanol, méthanal, acide méthanoïque et aminométhane) et de l'éthane (éthanol, éthanal, acide éthanoïque et amminoéthane). (C) Exemple d'un acide gras, composé d'une chaîne hydrocarbonée de 16 carbones (saturée) terminée par une fonction carboxyle. La numérotation part du groupement fonctionnel. Il s'agit de l'**acide palmitique** (acide hexadécanoïque).

### 3. Des acides aminés aux protéines.

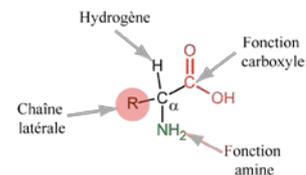
Les acides aminés sont des molécules assez connues. On les rencontre dans *toutes* les formes vivantes, et ceci parce qu'ils sont utilisés pour la fabrication (on dit la **biosynthèse**) des **protéines**. Ils sont également célèbres du fait qu'on les rencontre également dans l'espace, en particulier dans les météorites. En effet, un sous-type particulier de météorites, appelé **chondrites carbonées** contient de petites inclusions riches en carbone (en matière organique !), lesquelles sont composées, entre autres, d'acides aminés.

La vie terrestre utilise actuellement **vingt** acides aminés de base qui servent à la biosynthèse des protéines, mais également un nombre important d'autres dérivés qui servent à d'autres processus. Généralement, ce sont des acides dits **alpha-aminés** (nous verrons pourquoi dans quelques instants). Dans les météorites, on trouve un certain nombre d'acides aminés semblables à ceux des organismes vivants, mais surtout un très grand nombre – et une très grande proportion – de molécules moins conventionnelles.

On parle d'acide aminé parce que ces molécules organiques sont construites autour d'un carbone central (le carbone alpha) autour duquel sont greffés quatre groupements fonctionnels différents : tout d'abord, on trouve toujours une fonction **amine** ( $-\text{NH}_2$ ) et une fonction **carboxyle** ( $-\text{COOH}$ ), dite également « fonction **acide** » (Fig. 4) Le troisième groupe est un hydrogène. Le quatrième groupe est variable dans chacun des acides aminés biologiques (Fig. 5). Comme il existe vingt acides aminés, il y a vingt groupements fonctionnels. On dit qu'il s'agit de **chaînes latérales**. Comme la fonction amine et la fonction acide se trouvent sur le même carbone alpha, on parle d'acides **alpha-aminés**. Il existe également des acides beta-aminés, pour lesquels la fonction amine et la fonction acide se trouvent sur deux carbones consécutifs différents, mais aussi des acides gamma-aminés, etc... Dans la nature, les acides beta- ou gamma-aminés sont rares, les formes « alpha » étant les plus courantes. Les vingt acides aminés trouvés dans les protéines sont résumés dans la Fig. 5. Ceux qu'on retrouve aussi dans les météorites sont indiqués.

Les protéines sont des **polymères** d'acides aminés (nous reparlerons des protéines plus en détail dans un chapitre ultérieur). Cela signifie, pour simplifier, qu'il s'agit d'enchaînements d'acides aminés reliés entre eux par des liaisons covalentes. Les cellules utilisent un mécanisme assez complexe pour effectuer cette biosynthèse, et je ne la décrirai pas ici, mais il faut simplement savoir que l'assemblage de deux acides aminés nécessite une réaction de **condensation** (voir Chap. 3) entre la fonction carboxyle d'un premier acide aminé, et la fonction amine du suivant, et ainsi de suite. Lorsqu'ils sont incorporés de cette manière dans une molécule plus grande, on ne parle plus d'acide aminés, mais de **résidus** d'acides aminés. De cette manière, on peut obtenir des enchaînements d'acides aminés qui varient de quelques résidus (des **polypeptides**), à plusieurs centaines, voire plusieurs milliers de résidus (des **protéines**) (Fig. 6). Ainsi, l'alcool déshydrogénase (aussi appelée ADH), une protéine impliquée dans la dégradation de l'éthanol, et qui est présente chez les levures, est-elle composée de 348 résidus d'acides aminés. Les protéines sont également, à proprement parler, des molécules, mais comme elles sont très grandes (Fig. 6), on dit que ce sont des **macromolécules**. Les organismes vivants produisent énormément de macromolécules, dont les protéines, mais aussi l'ADN et autres espèces moléculaires aux noms compliqués.

**Figure 5 : Les acides aminés.** Les explications figurent dans le texte.



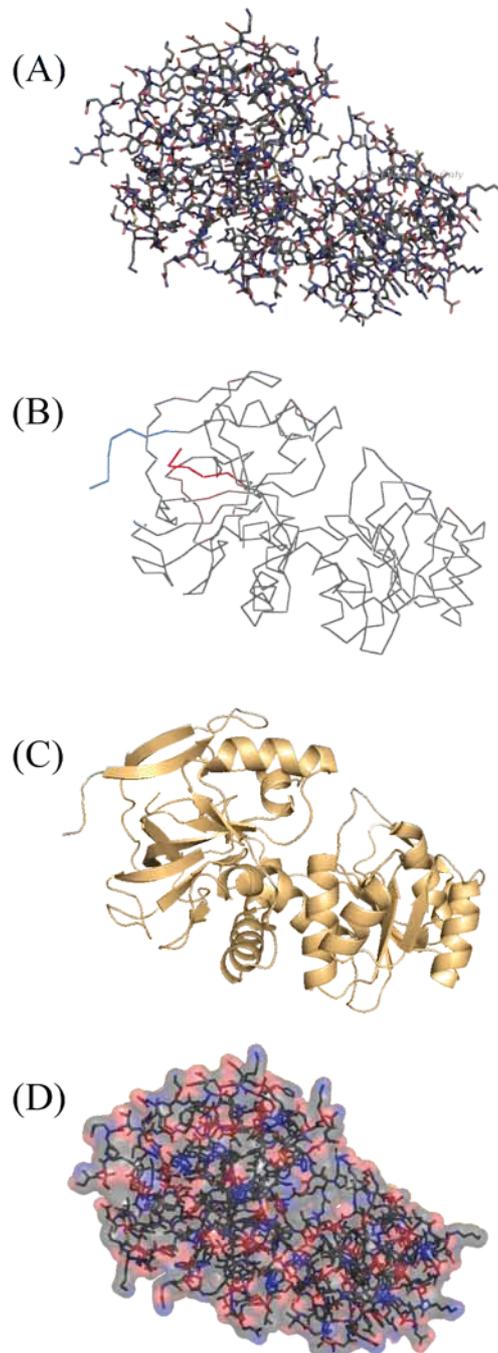
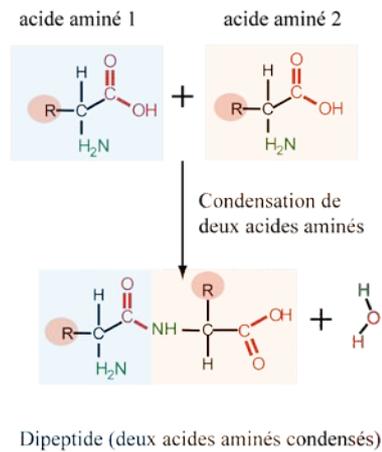
Nom	Abréviation (trois lettres)	Abréviation (une lettre)	Structure moléculaire
Glycine	Gly	G	
Alanine	Ala	A	
Valine	Val	V	
Leucine	Leu	L	
Isoleucine	Ile	I	
Phénylalanine	Phe	F	
Tyrosine	Tyr	Y	
Sérine	Ser	S	
Cystéine	Cys	C	
Thréonine	Thr	T	
Nom	Abréviation (trois lettres)	Abréviation (une lettre)	Structure moléculaire
Proline	Pro	P	
Histidine	His	H	
Tryptophane	Trp	W	
Acide aspartique	Asp	D	
Asparagine	Asn	N	
Acide glutamique	Glu	E	
Glutamine	Gln	Q	
Lysine	Lys	K	
Arginine	Arg	R	
Méthionine	Met	M	

Dans les météorites, on trouve beaucoup d'acides aminés, mais surtout la glycine – l'acide aminé le plus simple et le plus abondant –, l'acide aspartique, l'acide glutamique, l'alanine, la valine et la sérine (voir référence [Sephton et al, 2002](#) ; [Martins & Sephton, 2009](#)). D'autres sont plus rares, mais peuvent être détectés. Un certain nombre sont très souvent absents, sans qu'on sache réellement pour quelle raison. Toutefois, comme quelques acides aminés possèdent des chaînes latérales complexes, ou qui contiennent du soufre, ils sont probablement plus difficiles à obtenir par synthèse chimique abiotique (non reliée à la vie) dans le premier cas, et facilement oxydés (altérés) dans le second cas. De plus, nous verrons plus tard que, dans les météorites, on peut trouver deux formes d'acides aminés, dites L et D, tandis que la vie n'utilise que des formes L pour la construction des protéines. Les implications de ce fait seront discutées dans un chapitre ultérieur qui traitera des molécules organiques dans l'espace (météorites, comètes, milieu interstellaires, etc.).

#### 4. Des nucléotides à l'ADN.

Une autre classe importante de molécules organiques rencontrées dans les cellules sont les nucléotides et leurs dérivés. S'ils sont moins connus que les acides aminés, ils n'en sont pas moins importants, car ils entrent dans la composition et la biosynthèse des acides nucléiques, autrement dit, l'ADN (entre autres). Les bases azotées, qu'on retrouve dans les nucléotides sont également abondantes dans les météorites ([Callahan et al, 2011](#)).

**Figure 6 : Les acides aminés, les peptides et les protéines.** Deux acides aminés différents (acide aminé 1 sur fond bleu et acide aminé 2 sur fond orange) peuvent réagir entre eux dans une réaction dite de *condensation*. La fonction carboxyle (en rouge) du premier et la fonction amine (en vert) du second réagissent entre elles (avec libération d'une molécule d'eau) pour former une liaison covalente (appelée liaison amide, ou liaison *peptidique*). Les deux acides aminés sont alors liés dans une seule molécule, appelée un **dipeptide**. Au-delà d'une dizaine d'acides aminés, on parle de **protéine**. Certaines peuvent être très grandes. Dans l'exemple suivant, il s'agit de l'alcool déshydrogénase (ADH), qui représente un enchaînement de 348 acides aminés (résidus). **(A)** Il s'agit d'une représentation de la structure cette protéine (ADH), où toutes les liaisons covalentes des acides aminés sont représentées par des bâtonnets. A leurs extrémités, les atomes ont été figurés selon leur nature : en noir, le carbone (très abondant, évidemment), en rouge, l'oxygène, en bleu, l'azote, en orange, le soufre. Les hydrogènes (normalement en blanc) n'ont pas été représentés, pour plus de clarté. **(B)** La représentation « simplifiée » ici présentée montre *uniquement la direction générale et l'orientation relative des acides aminés* au sein de l'enchaînement avec le début de la protéine (les dix premiers acides aminés) en bleu, et les dix derniers en rouge. **(C)** Une autre représentation « simplifiée », très utilisée, qui peut remplacer la précédente. La direction générale et l'orientation des acides aminés sont représentés par des bandes. On voit apparaître des motifs : des hélices pour les acides aminés qui forment des spirales, des brins épais pour ceux qui forment des structures plus linéaires, et des boucles fines pour ceux qui ne possèdent pas de direction privilégiée les uns par rapport aux autres. Cette représentation est très utile pour comparer l'architecture générale des protéines entre elles. **(D)** La dernière représentation montre, comme dans la première, les liaisons en fonction des atomes présents, tandis que le « fond » en transparence montre l'espace qu'occuperaient le volume des atomes composant tous ces acides aminés, si on les faisait apparaître. Dans ce dernier cas, la protéine n'apparaîtrait que comme un amas diffus de sphérules. On comprend aisément l'utilité des représentations « simplifiées ».



Pour construire un **nucléotide**, il faut tout d'abord disposer d'une **base azotée** (Fig. 7). Il en existe cinq différentes : l'adénine (abrévée A), la guanine (G), la cytosine (C), la thymine (T) et l'uracile (U)<sup>1</sup>. Les quatre premières sont retrouvées dans l'ADN, tandis que l'uracile n'est retrouvée que dans l'ARN, un autre type de molécule composé de nucléotides, très important pour les cellules. Il faut ensuite ajouter l'une de ces bases azotées sur une molécule de **ribose** (ou de désoxyribose) (Fig. 7). A ce stade, on obtient ce qu'on appelle un **nucléoside** – une base liée au ribose<sup>2</sup>. L'adénosine, la guanosine, la cytidine et la thymidine sont les nucléosides comportant respectivement l'adénine, la guanine, la cytosine et la thymine (Fig. 7). Si vous utilisez un ribose, vous construirez un **ribonucléoside**, comme ceux qu'on retrouve dans l'ARN (acide **ribonucléique**), et s'il s'agit d'un **désoxyribose**, vous obtenez un **désoxyribonucléoside**, c'est-à-dire ceux qu'on retrouve dans l'ADN (acide **désoxyribonucléique**). Toutefois, pour être utilisables et incorporés dans l'ARN ou l'ADN, ces nucléosides doivent être modifiés en **nucléotides** (avec un *t*). Il faut alors leur ajouter un groupement phosphate, auquel cas on obtient un **nucléoside monophosphate**. L'ajout de deux groupes phosphate consécutifs donne un **nucléoside diphosphate**, et trois, un **nucléoside triphosphate**. Tous les nucléosides phosphatés sont des **nucléotides** (avec un *t*) (Fig. 7). L'adénosine triphosphate (ou encore abrégé ATP) est, par exemple, composée de l'adénine (base azotée), du ribose (sucre), et de trois groupements phosphates consécutifs (triphosphate) (Fig. 7).

En faisant réagir entre eux deux nucléotides, ils seront reliés entre eux par une liaison covalente (en termes techniques, on parle de *liaison phosphodiester*), et vous obtiendrez un dinucléotide. En **polymérisant** ainsi des dizaines, des centaines, des milliers, voire des millions de *désoxyribonucléotides*, vous construirez un brin d'ADN (Fig. 7). Physiologiquement, l'ADN n'est quasiment jamais retrouvé sous forme de brin isolé comme celui représenté dans la Fig. 7, mais plutôt sous forme de double-brin. Nous verrons cela plus en détail en temps voulu, car à chaque jour suffit sa peine !

Ce chapitre avait pour but de vous présenter schématiquement le monde dense et mystérieux des molécules organiques. Vous avez découvert que leurs propriétés et leur capacité à former des assemblages très divers peuvent conduire, chez les organismes vivants, à l'obtention de **macromolécules** gigantesques comme les protéines ou les acides nucléiques (ADN, ARN). Armés de ces connaissances, vous pourrez bientôt découvrir les secrets des molécules biologiques. Nous verrons très bientôt que l'univers est riche en matière organique, et que, même si la vie est probablement née sur Terre, ses constituants, eux, sont retrouvés même dans le vide spatial.

<sup>1</sup> Attention ! L'adénine, la guanine, la cytosine, la thymine et l'uracile, ainsi que leurs dérivés sont représentés, pour simplifier, par les abréviations bien connues A, G, C, T et U. Cependant, il ne faut pas les confondre avec les A, G, C et T de l'alanine, glycine, cystéine et thréonine des acides aminés. La même abréviation est utilisée, mais il ne s'agit pas du tout du même type de molécule. Si vous parlez de protéines, il s'agirait des acides aminés (alanine, glycine, etc...), et si vous parlez d'acides nucléiques (ADN, ARN, nucléotides), vous parlez d'adénine, guanine, etc... La science est ainsi faite ! Dans la plupart des films, les lettres A, T, G et C représentent des nucléotides, c'est-à-dire les lettres de l'ADN.

<sup>2</sup> Les sucres sont également appelés **oses**. Pour cette raison, tous les sucres et leurs dérivés comportent la terminaison **ose** ou **oside** respectivement.

## 5. Des familles macromoléculaires à ne pas confondre.

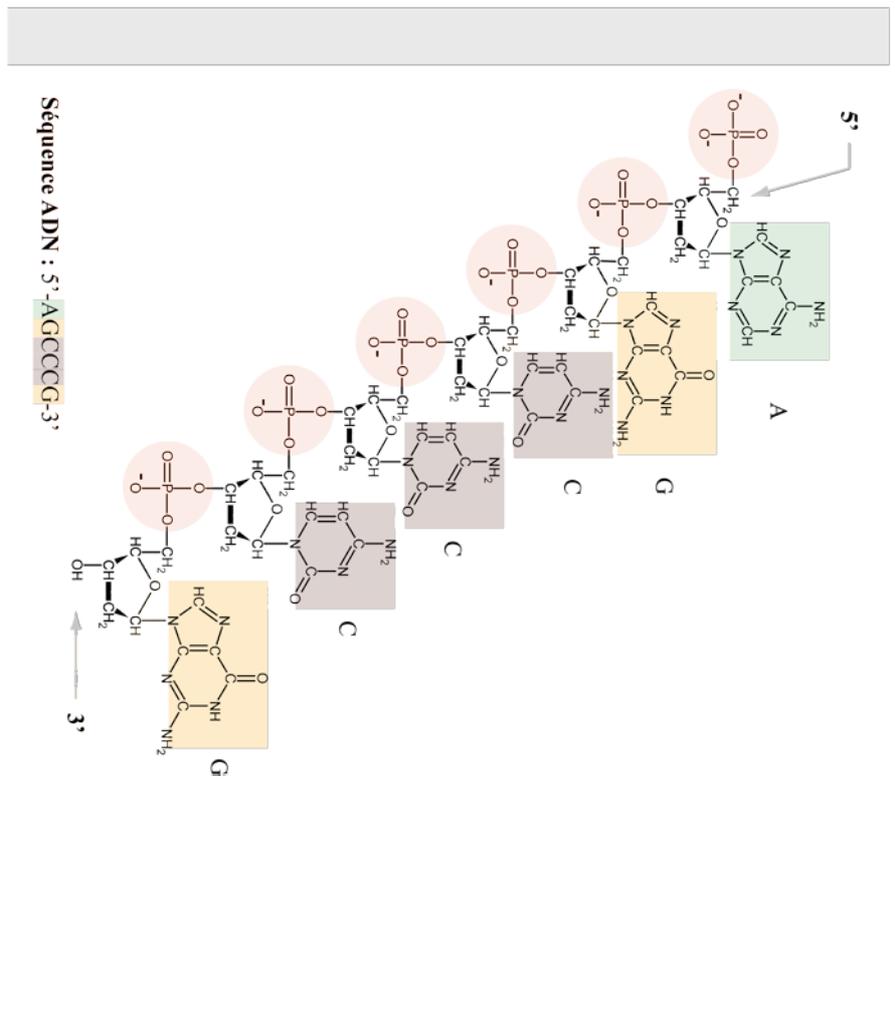
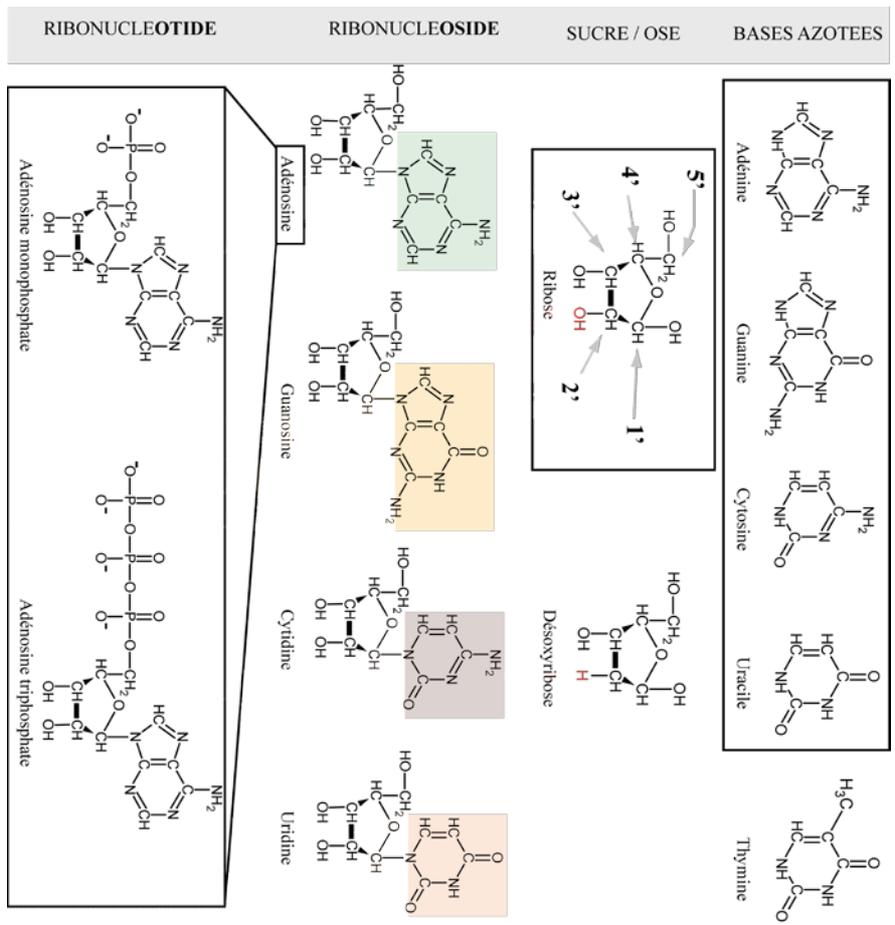
Comme je l'ai déjà abordé dans l'introduction de ces exposés, le problème majeur avec la biologie, la chimie ou la biochimie, ce sont les *termes techniques*. Ainsi, il faut encore clarifier quelques points avant de conclure ce chapitre.

Vous avez vu que les acides aminés sont de petites molécules organiques qui, si elles sont assemblées par polymérisation, peuvent donner des di-, tri-, polypeptides, ou, si le nombre de résidus est important, des **protéines**. Comme ces assemblages sont grands, et impliquent de nombreuses molécules organiques de base (qu'on appelle des précurseurs, puisqu'ils servent à construire autre chose), on dit que ce sont des **macromolécules**. De même, les nucléotides peuvent être polymérisés pour donner d'autres types de macromolécules, l'ADN ou l'ARN. L'ADN et l'ARN appartiennent à une famille de macromolécules qu'on appelle les *acides nucléiques*. Vous connaissez donc à présent deux grandes familles de macromolécules biologiques (il en existe quatre) :

- **Les protéines** : construites par polymérisation d'acides aminés,
- **Les acides nucléiques** (ADN, ARN), construits à partir de nucléotides.

Ce sont deux familles à ne pas confondre, car leurs rôles biologiques sont complètement différents. Il ne faut pas non plus confondre les *acides nucléiques*<sup>3</sup>, qui sont des **polymères** de nucléotides, et les *acides aminés*, qui sont les précurseurs servant à la construction des **protéines**.

<sup>3</sup> Le terme **acide nucléique** provient du fait qu'originellement, l'ADN a été découvert dans le **noyau** des cellules (adjectif correspondant : *nucléaire*, racine : « *nucléi-*, *nucléo-* »). Nous verrons plus tard que toutes les cellules ne possèdent pas de noyau. Toutefois, le terme a perduré, et on désigne aussi la grande famille des **polymères de nucléotides** comme les *acides nucléiques*.



**Figure 7 :** Il existe cinq bases azotées courantes : l'adénine (A), la guanine (G), la cytosine (C), l'uracile (U) et la thymine (T). A, G, C et T (mais pas U) sont retrouvés dans l'ADN, tandis que l'ARN est composé de A, G, C et U (mais pas T, sauf exception). Ces bases azotées entrent dans la composition des **nucleotides**. Pour cela, il faut les lier au **ribose** (dans l'ARN) ou au **désoxyribose** (dans l'ADN). Le **ribose** est un sucre (un ose, voir note 2, p. 6) cyclique composé de 5 carbones (4 internes au cycle) et un oxygène. On numérote généralement les carbones 1', 2', 3', 4' et 5' par convention. La base azotée (A, T, G, C ou U) est ajoutée sur le carbone 1'. Remarquez que le **ribose** et le **désoxyribose** se différencient par la présence d'une fonction hydroxyle sur le carbone 2' pour le ribose, et d'un hydrogène pour le désoxyribose. La combinaison d'une base et d'un sucre donne un **nucleoside**. La figure présente l'exemple de l'ajout de A, G, C, T et U sur le ribose. Pour le désoxyribose, le principe est le même, et il suffit de dessiner un H à la place du OH en position 2' (en rouge). Pour construire un **nucleotide** (avec un *t*), il faut rajouter un, deux ou trois phosphates (fonction phosphoryle, voir Fig. 3) en position 5' (*grossa modo*, de « l'autre côté » du nucleoside. L'exemple de l'adénosine est présenté. En ajoutant un seul phosphate, on obtient l'adénosine monophosphate (ou adénosine 5'-phosphate, abrégé AMP), et pour trois phosphates, on obtient l'adénosine triphosphate (ou adénosine 5'-triphosphate, ou ATP). Ces deux composés sont des nucleotides. Le principe est le même pour la guanine, la cytosine ou les autres nucleosides. **Les nucleotides peuvent être polymérisés** par l'intermédiaire de leurs phosphates en 5', et leur hydroxyle en 3'. Il en résulte une liaison dite **phosphodiester** (cercle rouge). L'assemblage présenté à droite correspond à la polymérisation de 6 **désoxyribonucleotides**. Il s'agit d'un **brin d'ADN**. Par convention, un scientifique lira toujours la séquence de ce brin en partant du nucleotide qui possède un phosphate **libre** en 5' (non engagé dans une liaison avec un autre nucleotide) pour aller vers celui qui possède une fonction 3'-OH libre (le dernier). Ainsi, vous verrez souvent qu'une séquence d'ADN (ici AGCCCCG) est orientée de 5' vers 3' : 5'-AGCCCCG-3'. Cette précision est très utile car le sens de lecture de l'ADN est important pour l'analyse du génome. Cette écriture simplifiée évite d'écrire la structure complète !