

# TOUTE LA BIOCHIMIE

***Serge Weinman***

Professeur émérite  
à la Faculté de médecine Saint-Antoine, Paris.  
Ancien chef de service de biochimie  
et de biologie moléculaire de l'hôpital Rothschild, Paris

***Pierre Méhul***

Illustrateur

**DUNOD**

## Illustration de couverture : *Pierre Méhul*

Ce pictogramme mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du **photocopillage**.

Le Code de la propriété intellectuelle du 1<sup>er</sup> juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les

établissements d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la

possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation du Centre français d'exploitation du droit de copie (**CFC**, 20 rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2004

© Dunod, Paris, 2023 pour la nouvelle présentation  
11 rue Paul Bert, 92240 Malakoff

[www.dunod.com](http://www.dunod.com)

ISBN : 978-2-10-085533-9

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

# Table des matières

Introduction

1

## PARTIE I

---

### ARCHITECTURE ET LOGIQUE DES ÊTRES VIVANTS STRUCTURE DES BIOMOLÉCULES

CHAPITRE 1 • Aminoacides, peptides et protéines	7
Les $\alpha$ -aminoacides	8
La chaîne polypeptidique et ses conformations	14
Repliement des chaînes polypeptidiques des protéines globulaires hydrosolubles	18
Des motifs de repliement, ou structures supersecondaires, sont formés à partir d'hélices $\alpha$ et/ou de brins $\beta$	20
Les chaînes polypeptidiques sont organisées en domaines	22
Les structures des protéines peuvent être réparties en trois classes	22
Les protéines transmembranaires	24
La prédiction de la structure tridimensionnelle des protéines est-elle possible ?	24
Flexibilité conformationnelle des protéines	26
Liaisons et interactions des protéines	26
Liaison d'un ligand sur un seul site	27

---

Liaison coopérative d'un même ligand sur des sites multiples semblables : interactions homotropiques	30
Liaison de plusieurs ligands sur des sites différents : interactions hétérotropiques	30
Concept d'allostérie	31
Modèle concerté	31
Modèle séquentiel	32
Les protéines sont-elles des démons de Maxwell ?	32
Classification fonctionnelle des protéines	34
Appendice : structures X et RMN des protéines	34
<b>CHAPIRE 2 • Nucléotides et acides nucléiques</b>	37
Les nucléosides	38
Les nucléotides	38
Liaison phosphodiester et acides nucléiques	40
Acide désoxyribonucléique ou DNA	42
Formes tridimensionnelles alternatives du DNA	44
Acides ribonucléiques ou RNA	46
<b>CHAPITRE 3 • Glucides et glycoconjugués</b>	51
Structure linéaire des oses	52
Centres d'asymétrie des oses : le glycéraldéhyde	52
Filiation des oses	54
Structure cyclique des oses	58
Dérivés des oses	64
La liaison glycosidique	66
Amidon et glycogène	68
Cellulose et chitine	70
Glycosaminoglycanes	72
Protéoglycanes	72
Les glycoprotéines	74
Les glycolipides	78
Les lipopolysaccharides	78

CHAPITRE 4 • <b>Lipides et membranes cellulaires</b>	81
Les acides gras	82
Lipides de réserve : les triacylglycérols	82
Lipides membranaires : les glycérophospholipides et les sphingolipides	84
Stérols et stéroïdes	86
Double couche lipidique et membranes biologiques	88

## PARTIE II

### CATALYSE BIOLOGIQUE : STRUCTURE ET MÉCANISME D'ACTION DES ENZYMES

CHAPITRE 5 • <b>Cinétique enzymatique</b>	93
Relations entre la vitesse de réaction et la concentration du substrat ou de l'enzyme	94
Théorie du complexe stéréospécifique et propriétés cinétiques des enzymes	94
Équation de Michaelis-Menten	96
Détermination expérimentale des valeurs de $V_{\max}$ et de $K_M$	97
Signification physique de $K_M$	98
Signification physique de la constante catalytique $K_{\text{cat}}$	98
Le critère $K_{\text{cat}}/K_M$ mesure la spécificité et l'efficacité des enzymes	98
Réactions mettant en présence plusieurs substrats	100
Effet de la température sur les réactions enzymatiques	102
Les enzymes diminuent l'énergie d'activation	104
Nature du pouvoir catalytique et de la spécificité des enzymes	104
Inactivation thermique des enzymes	105
CHAPITRE 6 • <b>Stratégies moléculaires de la catalyse enzymatique</b>	107
Le modèle serrure-clé et l'adaptation induite	108
Des résidus aminoacide polaires forment le centre catalytique des enzymes	110
La vitesse des réactions enzymatiques est affectée par le pH	110

---

La catalyse acide-base	112
La catalyse covalente	112
La liaison stéréospécifique des réactifs joue un rôle important dans la catalyse enzymatique	112
Effet de proximité	114
Effet Circé	114
Effet de stabilisation de l'état de transition	116
Mécanisme de l'hydrolyse d'une liaison osidique par le lysozyme	118
Mécanisme de l'activation de la tyrosine par une aminoacyl-tRNA synthétase	120
Mécanisme d'action et classification des coenzymes	122
Les coenzymes nucléotidiques	124
Le motif de liaison des nucléotides et les déshydrogénases NAD-dépendantes	126
Le domaine de liaison du FAD est semblable au domaine de liaison du NAD	128
Le domaine de liaison de l'ATP est, lui aussi, semblable au domaine de liaison du NAD	128
Le coenzyme A	130
Le thiamine pyrophosphate	130
Le pyridoxal phosphate	130
Rôle des cofacteurs métalliques	132
<b>CHAPITRE 7 • Mécanismes de régulation de l'activité enzymatique</b>	137
Inhibition compétitive des enzymes	138
Inhibition non compétitive des enzymes	140
Inhibition incompétitive des enzymes	142
Inhibition par excès de substrat	142
Inactivation des enzymes	144
Des inhibiteurs de l'activité enzymatique sont des réactifs biochimiques ou des médicaments puissants et précieux	144
Enzymes allostériques	148
L'aspartate transcarbamylase : exemple d'un enzyme allostérique	150
Modification covalente des enzymes	150
Les protéine kinases	152
Activation des proenzymes	152

## PARTIE III

**ÉNERGIE CELLULAIRE  
CONVERSION ET MISE EN RÉSERVE**

<b>CHAPITRE 8 • Métabolisme du glucose : glycolyse, voie des pentoses phosphate, glycogénogenèse et glycolyse, gluconéogenèse</b>	159
Origine et transport du glucose	160
Métabolisme du glucose : vue d'ensemble	162
Glycolyse	164
Bilan de la glycolyse	172
Cancer et glycolyse	173
Métabolisme du galactose et du fructose	174
Voie des pentoses phosphate et création de NADPH et de pentoses	176
Glycogénogenèse	178
Glycogénolyse	180
Gluconéogenèse	184
Devenir du pyruvate	187
<b>CHAPITRE 9 • Métabolisme des acides gras : <math>\beta</math>-oxydation et biosynthèse</b>	189
Origine et transport des acides gras	190
$\beta$ -oxydation mitochondriale des acides gras	192
Formation et utilisation des corps cétoniques	194
Biosynthèse cytosolique des acides gras	194
Contrôle du métabolisme des acides gras	198
<b>CHAPITRE 10 • Métabolisme des aminoacides : désamination oxydative</b>	201
Origine des aminoacides	202
Élimination et devenir du groupe amine des aminoacides	204
Destinée métabolique des squelettes carbonés des aminoacides	210
Le pyruvate comme point d'entrée dans le flux métabolique	212
l'oxaloacétate et l' $\alpha$ -cétoglutarate comme point d'entrée dans le flux métabolique	212

---

Le fumarate comme point d'entrée dans le flux métabolique	212
Le succinyl CoA comme point d'entrée dans le flux métabolique	214
La dégradation des aminoacides aromatiques fait intervenir des oxygénases	216
Les aminoacides à chaîne ramifiée, leucine, valine et isoleucine, sont dégradés par des voies qui ont des étapes communes	218
Erreurs innées du métabolisme des aminoacides	218
<b>CHAPITRE 11 • Cycle de l'acide citrique</b>	221
Formation de l'acétyl CoA à partir du pyruvate : le complexe pyruvate déshydrogénase	222
Oxydation de l'acétyl CoA par le cycle de l'acide citrique : vue d'ensemble	224
Réactions du cycle de l'acide citrique	226
Stœchiométrie du cycle de l'acide citrique	232
Des intermédiaires du cycle de l'acide citrique sont d'importants précurseurs biosynthétiques	232
Le cycle de l'acide citrique est contrôlé en plusieurs points	233
Cycle du glyoxylate	234
<b>CHAPITRE 12 • Phosphorylation oxydative</b>	237
Transfert d'électrons, pompage de protons et phosphorylation oxydative	238
Les électrons de haut potentiel de transfert du NADH et du FADH <sub>2</sub> activent la chaîne respiratoire mitochondriale	240
Origine du NADH et du FADH <sub>2</sub>	242
NADH-Q oxydoréductase	244
Les électrons du FADH <sub>2</sub> des flavoprotéines ont l'ubiquinone pour point d'entrée dans la chaîne respiratoire mitochondriale	246
Q-cytochrome c oxydoréductase	248
Cytochrome c	252
Cytochrome c oxydase	252
L'ATP synthase couple la synthèse d'ATP à la rentrée des protons dans la matrice	254
Bilan de la phosphorylation oxydative	256
<b>CHAPITRE 13 • Photosynthèse</b>	259
Photorécepteurs moléculaires	260
Membranes thylacoïdes et photosystèmes	262



---

Le photosystème de type II des bactéries pourpres	264
Photosystème I des cyanobactéries	267
Photosystème II des chloroplastes	270
Couplage du photosystème II au photosystème I	272
Photosystème I des chloroplastes	272
Ferrédoxine et NADPH	274
Gradient de protons, ATP synthase et photophosphorylation	274
Conversion du CO <sub>2</sub> en polysaccharides	276
<b>CHAPITRE 14 • Intégration et régulation du métabolisme énergétique</b>	279
Interconversion des métabolites	280
Contrôle des voies métaboliques énergétiques essentielles	280
Contrôle hormonal du métabolisme énergétique	282

## PARTIE IV

---

### SYNTHÈSE DES BIOMOLÉCULES

<b>CHAPITRE 15 • Biosynthèse des aminoacides</b>	291
Incorporation de l'azote dans les biomolécules	292
Biosynthèse des aminoacides	294
Régulation de la biosynthèse des aminoacides	302
Biomolécules dérivées des aminoacides	304
<b>CHAPITRE 16 • Biosynthèse des nucléotides</b>	307
Synthèse <i>de novo</i> des ribonucléotides pyrimidiques	308
Synthèse <i>de novo</i> des ribonucléotides puriques	312
Les désoxyribonucléotides sont synthétisés par réduction des ribonucléotides	314
Le thymidylate est synthétisé par méthylation du désoxyuridylate	314
Recyclage des bases nucléotidiques	316

---

Synthèse du NAD <sup>+</sup> , du NADP <sup>+</sup> , du FAD et du coenzyme A	316
Dégradation des purines	318
Syndrome de Lesch-Nyhan	318
<b>CHAPITRE 17 • Biosynthèse des lipides membranaires et des stéroïdes</b>	321
Biosynthèse des triacylglycérols et des glycérophospholipides	322
Biosynthèse des sphingolipides	324
Biosynthèse du cholestérol	326
Régulation de la biosynthèse du cholestérol	330
Les triglycérides et le cholestérol sont transportés par des lipoprotéines	330
Dérivés du cholestérol	332
Hormones eicosanoïdes	336
Vitamines A	338
Vitamines E, K et quinones	338
<b>CHAPITRE 18 • Réplication du DNA</b>	341
La réplication du DNA, mécanisme universel	342
Le DNA est synthétisé par des DNA polymérase	344
Réplication du DNA	346
Télomères	348
<b>CHAPITRE 19 • Biosynthèse du RNA</b>	351
La transcription est effectuée par des RNA polymérase DNA-dépendantes	352
La transcription chez les Procaryotes	352
La transcription chez les Eucaryotes	356
La maturation du RNA	358
Maturation des mRNA des Eucaryotes	358
Maturation des rRNA et des tRNA	362
Certains RNA catalysent l'épissage des introns	362
Régulation de l'initiation de la transcription	364
Domaines d'interaction protéines régulatrices-DNA	364
Domaines d'interaction protéine-protéine	368
Régulation de la transcription chez les Procaryotes	372

Régulation de la transcription chez les Eucaryotes	374
Superfamille des récepteurs nucléaires	376
<b>CHAPITRE 20 • Biosynthèse des protéines</b>	381
Le code génétique	382
Les ribosomes	382
Activation des aminoacides	384
Initiation	386
Élongation	388
Terminaison	390
Repliement de la chaîne polypeptidique et modifications post-traductionnelles	390
Régulation de la synthèse des protéines	390

## PARTIE V

---

### BIOMOLÉCULES ET ENVIRONNEMENT

<b>CHAPITRE 21 • Oxygène : transport et mise en réserve</b>	395
Mécanisme moléculaire du transport de l'oxygène	396
Rôle du fer dans le transport de l'oxygène	396
Le repliement globinique, motif structural commun	398
Structure des myoglobines et oxymyoglobines	398
Structure des hémoglobines et des oxyhémoglobines	400
La famille des gènes de globine	402
Évolution des gènes de globine	402
Fonction des myoglobines et des hémoglobines	404
L'hémoglobine F a plus d'affinité pour l'oxygène que l'hémoglobine A	406
Les ions H <sup>+</sup> et le CO <sub>2</sub> favorisent la libération de l'oxygène	406
Le 2,3-bisphosphoglycérate diminue l'affinité de l'hémoglobine pour l'oxygène	406
Drépanocytose et hémoglobine S	408
La désoxyhémoglobine S est très peu soluble	408
L'hémoglobine S confère une résistance au paludisme	410
Pathologie des gènes de l'hémoglobine	410

CHAPITRE 22 • <b>Système immunitaire</b>	413
Vue d'ensemble du système immunitaire	414
Protéines du complexe majeur d'histocompatibilité et présentation des antigènes cellulaires	420
Récepteur des cellules T et reconnaissance des complexes antigène-CMH	424
Marqueurs CD et profil spécifique des lymphocytes	426
Destruction des cellules cible par les lymphocytes T cytotoxiques	426
Immunoglobulines et fixation des antigènes	428
<b>Index</b>	439

# Introduction

Lors du processus de création et d'expansion de l'Univers qui évolue depuis environ quinze milliards d'années, deux événements nous concernant plus directement sont survenus : l'organisation du système solaire avec la formation de notre planète, la Terre, il y a quatre milliards et demi d'années et l'apparition de ce qu'il est convenu d'appeler la vie, il y a quatre milliards d'années.

Mais, qu'est la vie ? Bien qu'il n'y ait pas de réponse à cette question, on peut cependant considérer qu'un être vivant, s'il se présente comme un ensemble constitué d'atomes régi par les lois universelles de la physique, est en fait, au sein de son environnement terrestre, un système susceptible de se suffire à lui-même et de subir l'évolution darwinienne. Il peut, en effet, de façon autonome, se procurer les éléments dont il a besoin pour s'auto-construire et obtenir l'énergie nécessaire à son fonctionnement. Il est aussi capable de se reproduire par formation de copies identiques à lui-même, sauf dans quelques rares cas où apparaissent des variations qu'une sélection naturelle rejettera ou adoptera. Ces deux considérations, très synthétiques, résultent d'études biochimiques conduites, pour l'essentiel, au XX<sup>e</sup> siècle.

La biochimie a analysé les constituants des cellules et étudié les réactions qui s'effectuent au sein de ces dernières. Il est ainsi apparu que toutes les cellules, délimitées par des membranes phospholipidiques, étaient constituées essentiellement de macromolécules de deux types, les protéines, éventuellement associées à des glucides, et les acides nucléiques, elles-mêmes formées par polymérisation de modules relativement simples, les aminoacides et les nucléotides, respectivement, tous susceptibles d'être intégralement synthétisés en laboratoire.

Les protéines, qui constituent le protéome, sont en fait responsables de toutes les « performances téléonomiques des êtres vivants » (Jacques Monod, *Le Hasard et la Nécessité*, 1970), y compris la catalyse par des enzymes de tout le réseau des réactions qui assument la transformation des biomolécules et la transduction de l'énergie.

L'acide désoxyribonucléique (DNA), qui constitue le génome, contient l'ensemble de l'information génétique. Cette dernière est tout d'abord transcrite en acides ribonucléiques messagers (mRNA) avant d'être traduite en protéines grâce à l'action conjuguée des acides ribonucléiques de transfert (tRNA) et des acides ribonucléiques ribosomiques (rRNA) ; les mRNA, les tRNA et les rRNA constituent ensemble le transcriptome. Par ailleurs, le DNA est susceptible de se répliquer et d'être transmis, inchangé, ou parfois muté, ou parfois recombinaison, d'une cellule mère à ses cellules filles.

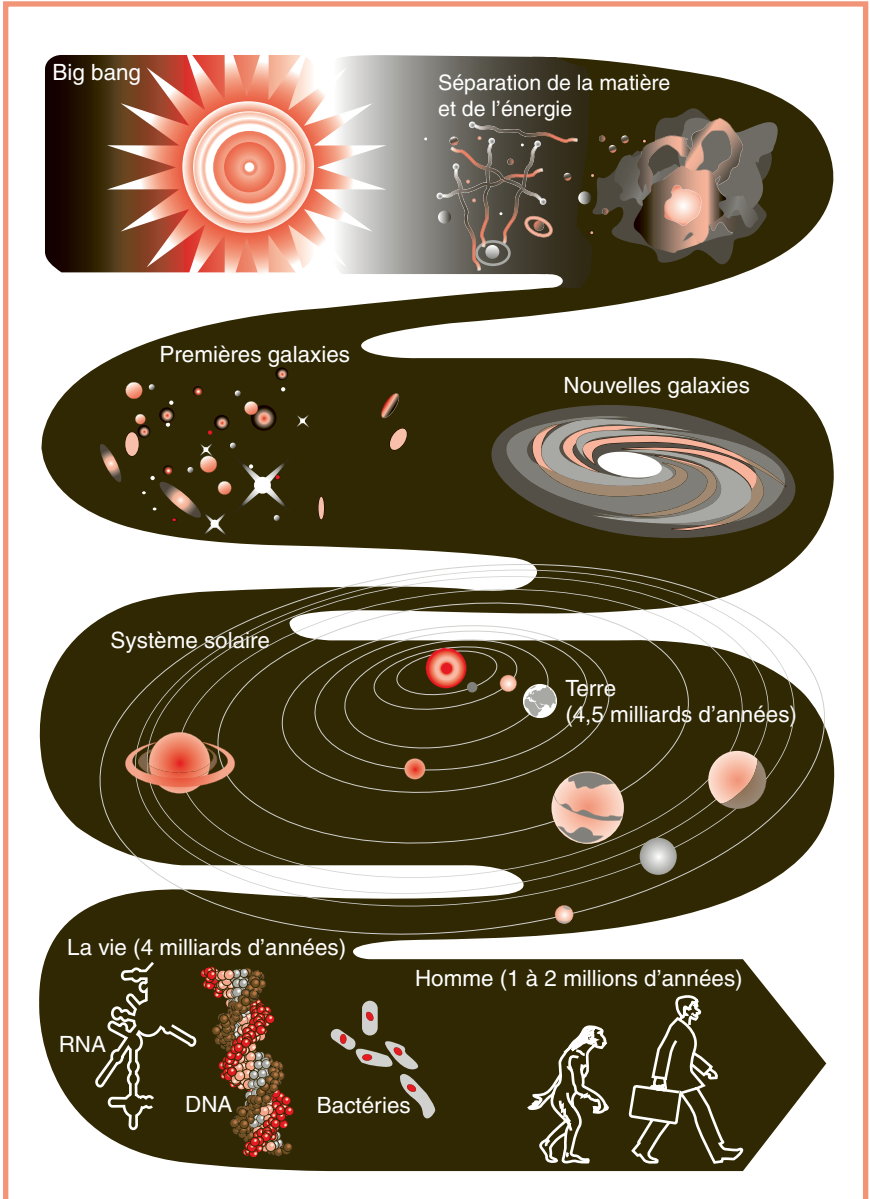
Le schéma de l'évolution des réactions qui s'effectuent chez un être vivant s'établit donc ainsi :



Sur ce solide fond de connaissances apportées par la biochimie, la génétique a pu établir que des segments de DNA, les gènes, sont au cœur de tous les événements cellulaires, puis la biologie moléculaire a su montrer que toutes les propriétés, héréditaires ou téléonomiques, des êtres vivants s'expliquent par la structure et les interactions des biomolécules qui constituent ces derniers.

La biochimie a aussi montré l'unité de structure et de fonction sous-jacente à la diversité biologique du monde vivant. L'utilisation du DNA par tous les êtres vivants et l'universalité du code génétique sont à la base de la constatation selon laquelle tous les organismes actuels apparaissent remarquablement uniformes au niveau moléculaire et ne se distinguent que par des variations somme toute relativement mineures. Cette notion implique aussi qu'ils viennent tous d'un ancêtre commun et que la diversité de la vie dans le monde moderne résulte de processus évolutifs opérant sur des processus fondamentaux. Bien plus que les mutations éventuelles, c'est une combinatoire d'un nombre limité de segments d'ADN qui produit l'énorme variété des structures formant les principaux constituants cellulaires. L'évolution biochimique est donc due avant tout à la duplication de gènes et à leur réassortiment ; la constitution en mosaïque des gènes et la structure en domaines des protéines qui s'ensuit donne à ces dernières des possibilités d'interactions multiples, accrues encore par la possibilité de formation de complexes multiprotéiques parfois très volumineux.

Ces étonnantes conquêtes de la biochimie sont actuellement présentées dans nombre de traités très complets, mais un guide permettant d'en revoir rapidement l'essentiel peut être aussi un instrument de travail utile pour l'étudiant. Ce livre a été conçu pour répondre à ce besoin ; abondamment illustré, il apporte une vue d'ensemble des principaux concepts de la biochimie et, bien que rédigé sous une forme concise, chaque chapitre tente d'en présenter les fondements en prenant en compte les découvertes les plus récentes.



Le livre comprend cinq parties.

- La Partie I : «Architecture et logique moléculaires des êtres vivants» est une présentation de la structure et des propriétés des classes de biomolécules les plus importantes : protéines, modules fondamentaux des êtres vivants, acides nucléiques, porteurs de l'information génétique, glucides et lipides, réserves d'énergie mais aussi éléments structuraux.
- La Partie II : «Catalyse biologique : structure et mécanisme d'action des enzymes» est consacrée aux processus qui permettent la mise en œuvre du vaste réseau de réactions assurant l'édification des structures des cellules et la production d'énergie utilisable par ces dernières.
- La Partie III : «Énergie cellulaire, conversion et mise en réserve» traite de la création de l'ATP par phosphorylation oxydative ou photosynthèse. L'interconversion des différentes formes de l'énergie métabolique et les réactions de voies telles que la glycolyse et la gluconéogenèse, la  $\beta$ -oxydation et la biosynthèse des acides gras, le métabolisme des aminoacides, le cycle de l'acide citrique, la voie des pentoses phosphate sont discutées en fonction de mécanismes enzymatiques établis. Un dernier chapitre étudie certains aspects de la régulation et de l'intégration du métabolisme énergétique.
- La Partie IV : «Synthèse des biomolécules» couvre la biosynthèse des aminoacides, des nucléotides, des lipides membranaires, des stérols et des stéroïdes, la réplication du DNA, la synthèse du RNA et son épissage, la synthèse des protéines.
- La Partie V : «Biomolécules et environnement» montre comment les cellules répondent ou s'adaptent aux changements environnementaux. En particulier, les molécules qui assument les réponses du système immunitaire illustrent la façon dont ce processus de défense est intégré dans les organismes multicellulaires.



## PARTIE I

---

# ARCHITECTURE ET LOGIQUE DES ÊTRES VIVANTS STRUCTURE DES BIOMOLÉCULES

La vie est la manifestation de l'une des nombreuses organisations de la matière et de l'énergie qui constituent l'Univers et les êtres vivants apparaissent comme formés d'atomes et de molécules qui se conforment à toutes les lois physiques déterminant les propriétés de cette matière et de cette énergie. De plus, les molécules des êtres vivants n'ont pas de propriétés particulières que n'auraient pas celles des corps inertes. Cependant, les organismes vivants se distinguent des autres collections de molécules par une architecture et une logique moléculaires qui leur confèrent des fonctions propres.

Les systèmes vivants présentent une étonnante diversité. Néanmoins, ils sont construits à partir de constituants moléculaires très semblables qui ne se distinguent que par des variations somme toute mineures. De plus, leurs processus biochimiques essentiels sont communs. Ces constatations impliquent que tous les organismes qui ont vécu ou vivent actuellement sur la Terre dérivent d'un ancêtre commun et que leur diversité résulte de phénomènes évolutifs qui ont progressivement adapté aux conditions environnementales les processus fondamentaux déjà présents chez cet ancêtre commun.

### Architecture moléculaire

Tous les organismes vivants sont constitués de cellules. La cellule est l'unité du vivant ; elle en a toutes les propriétés. L'analyse d'une cellule révèle qu'elle est constituée de deux types de molécules : des macromolécules et des petites molécules.

Les macromolécules sont des polymères construits par répétition d'une même réaction ; à chaque étape, un même type de module est ajouté de façon à former un ensemble constitué de centaines, voire de milliers de résidus. Les macromolécules cellulaires sont de quatre types : (1) les protéines, édifiées à partir de vingt aminoacides différents, qui assument toutes les activités et performances des êtres vivants ; (2) l'acide désoxyribonucléique (DNA), élaboré à partir de seulement quatre types de désoxyribonucléotides, qui assure la conservation de l'information génétique d'une cellule et sa transmission, inchangée, d'une génération à une autre ; (3) les acides ribonucléiques (RNA), assemblés à partir de quatre types principaux de ribonucléotides, qui permettent le transfert vers les protéines de l'information contenue dans le DNA et enfin (4) les polysaccharides, constitués de résidus glucose, qui constituent une partie de la réserve d'énergie chimique nécessaire à la cellule.

Les petites molécules se forment généralement d'une toute autre manière au cours d'une suite de réactions où, à chaque étape, un atome ou un groupe limité d'atomes est ajouté ou retranché sous l'action d'un enzyme particulier qui se comporte en catalyseur spécifique. Les petites molécules ont des fonctions très diverses. Elles peuvent être, tels les aminoacides, les nucléotides ou les oses simples, les modules élémentaires à partir desquels se construisent les édifices macromoléculaires ; elles peuvent aussi, tels les phospholipides et le cholestérol, s'associer pour former les membranes biologiques qui assurent la compartimentation de la cellule ; elles peuvent enfin, tels les oses simples, les acides gras et les aminoacides, constituer un ensemble de molécules immédiatement disponibles à partir desquelles la cellule crée de l'énergie directement utilisable lorsqu'elle en a besoin pour ses activités.

### **Logique moléculaire**

Le vivant se distingue de l'inanimé par deux caractères. Tout d'abord, les êtres vivants ont la capacité de reproduction invariante et de morphogénèse. De plus, ils extraient et transforment continûment l'énergie contenue dans les molécules ou dans la lumière solaire de leur environnement, puis ils l'utilisent pour maintenir leurs structures complexes et ordonnées ou effectuer un travail osmotique, chimique ou mécanique. Ainsi, ils ne sont jamais en équilibre avec leur environnement mais se présentent dans un état dynamique de création continue. La matière inanimée, elle, tend vers un état désordonné qui la met en équilibre avec ce qui l'entoure.

L'objet de la Partie I de ce livre est d'exposer la structure et les propriétés des principaux types de biomolécules qui constituent la cellule.

## Chapitre 1

---

# Aminoacides, peptides et protéines

**P**résentes chez tous les êtres vivants, des Bactéries à l'Homme, les protéines sont de très gros biopolymères formés à partir de seulement 20 monomères différents. Ces monomères sont des  $\alpha$ -aminoacides qui s'unissent par des liaisons peptidiques pour former des chaînes polypeptidiques linéaires. Ces  $\alpha$ -aminoacides constituent un alphabet universel, apparu il y a plus de deux milliards d'années, grâce auquel sont écrites des milliers de séquences, toutes différentes, propres chacune à une protéine.

Dans une chaîne polypeptidique, la séquence des  $\alpha$ -aminoacides, c'est-à-dire l'ordre dans lequel ces derniers sont unis, détermine la structure primaire. Les protéines possèdent la propriété de morphogenèse autonome en ce sens qu'un déterminisme interne assure la formation spontanée de leur structure tridimensionnelle, quelle qu'en soit la complexité, à partir de leur structure primaire. Les protéines ont donc une structure d'ensemble imposée par la nature et la séquence de leurs  $\alpha$ -aminoacides qui toutes deux déterminent des actions morphogénétiques internes décisives, même si des interactions avec le milieu peuvent moduler certaines conformations. Des régions du squelette peptidique adoptent une conformation régulière, en hélice  $\alpha$  ou en feuillet plissé  $\beta$ , dite structure secondaire. Les hélices et les feuillets plissés reliés par des boucles de conformation non régulière, se reploient en des unités compactes et globulaires appelées domaines dont l'organisation tridimensionnelle constitue la structure tertiaire de la chaîne polypeptidique. Certaines protéines ont une seule chaîne polypeptidique, d'autres en contiennent plusieurs disposées en une structure quaternaire spécifique.

On distingue, selon leur forme, deux groupes principaux de protéines, les protéines globulaires, solubles ou transmembranaires, de beaucoup les plus nombreuses, où le repliement des chaînes polypeptidiques conduit à une structure compacte, et les protéines fibreuses qui sont des molécules très allongées.

## LES $\alpha$ -AMINOACIDES

Les  $\alpha$ -aminoacides ont pour formule générale  $R - C_{\alpha} (H, NH_3^+) - COO^-$  où un carbone tétraédrique chiral  $C_{\alpha}$  est uni à un carboxyle  $-COO^-$ , une amine primaire  $-NH_3^+$ , un hydrogène  $-H$  et une chaîne latérale  $-R$  propre à chaque  $\alpha$ -aminoacide (Figure 1.1A). Les  $\alpha$ -aminoacides ont trois caractères structuraux fondamentaux qui déterminent leurs propriétés :

### 1. La configuration du $C_{\alpha}$

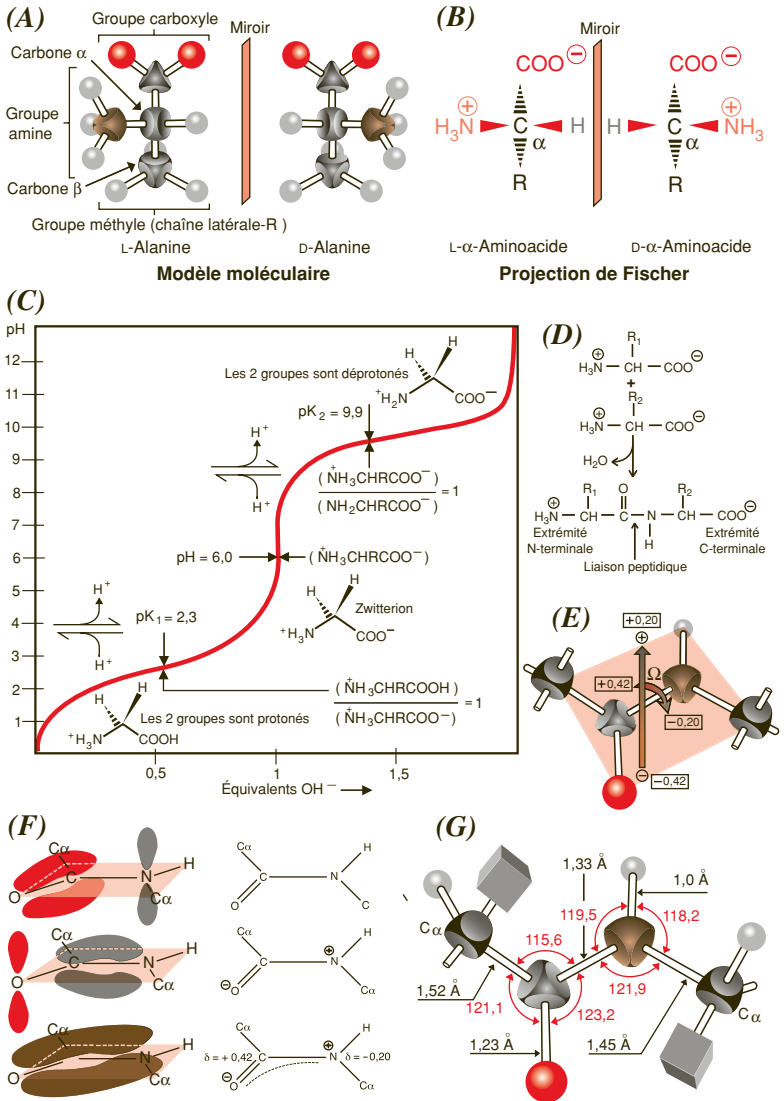
Lorsque les modèles des  $\alpha$ -aminoacides sont disposés selon la convention de Fischer, la chaîne carbonée est verticale et vue par sa convexité ; le  $-COO^-$ , dont le niveau d'oxydation est le plus élevé, est placé vers le haut et le  $-NH_3^+$  peut alors se situer soit à gauche, soit à droite du  $C_{\alpha}$  chiral ; les  $\alpha$ -aminoacides appartiennent alors à la série L (du latin *laevus*, côté gauche) ou à la série D (du latin *dexter*, côté droit), respectivement (voir les modèles de l'alanine, Figures 1.1A et B). Il existe donc, pour chaque  $\alpha$ -aminoacide, sauf pour la glycine où  $R = H$ , deux stéréoisomères (énantiomères), images l'un de l'autre dans un miroir ; les  $\alpha$ -aminoacides des protéines de tous les êtres vivants connus appartiennent à la série L.

### 2. La fonction $\alpha$ -aminoacide

En milieu aqueux, les  $\alpha$ -aminoacides sont ionisés et l'état d'ionisation est fonction du pH (voir la courbe de titration de l'alanine, Figure 1.1C). En milieu très acide, le carboxyle et l'amine sont protonés et l' $\alpha$ -aminoacide porte une charge positive ; à pH 2,3 ( $pK_a$  du carboxyle), un premier point d'inflexion apparaît ; à pH 6 (pH isoélectrique ou point isoélectrique pI), la charge positive est égale à la charge négative et l' $\alpha$ -aminoacide se présente comme un ion dipolaire (zwitterion) ; à pH 9,9 ( $pK_a$  de l'amine), un deuxième point d'inflexion apparaît ; en milieu très alcalin, le carboxyle et l'amine sont déprotonés et l' $\alpha$ -aminoacide porte une charge négative.

Deux  $\alpha$ -aminoacides peuvent s'unir par condensation du carboxyle de l'un avec l'amine de l'autre, ce qui conduit à un dipeptide présentant une liaison peptidique  $-CO-NH-$  dont le  $-CO$  est situé du côté de l'extrémité N-terminale et le  $-NH$  du côté de l'extrémité C-terminale (Figure 1.1D). La liaison peptidique est plane, rigide et polaire (Figure 1.1E) ; C, O, N, H et les deux  $C_{\alpha}$  voisins sont dans un plan où ils constituent une unité peptidique ; l'angle de torsion  $\Omega$  autour de la liaison C-N ne peut prendre que la valeur  $0^\circ$ , configuration *cis* où O et H sont d'un même côté de l'axe C-N, ou  $+180^\circ$ , configuration *trans* où O et H sont de part et d'autre de l'axe C-N ; pour toutes les liaisons peptidiques des protéines, la configuration *trans* est énergétiquement la plus favorable, sauf dans celles où intervient la proline. Ces caractères résultent du fait que la liaison peptidique est un hybride de résonance où les électrons sont localisés dans une orbitale moléculaire  $\pi$  qui recouvre les atomes O, C et N (Figure 1.1F) et empêche la libre rotation autour de l'axe C-N. Par ailleurs, l'oxygène étant plus électronégatif que l'azote, les électrons sont déplacés vers O. La liaison peptidique possède donc une géométrie fixe et les longueurs et les angles de liaison peuvent être mesurés (Figure 1.1G).

Figure 1.1



### 3. La structure et les propriétés de la chaîne latérale



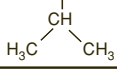
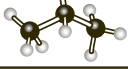
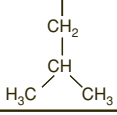
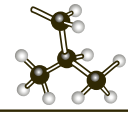
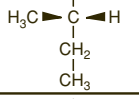
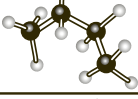

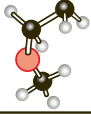
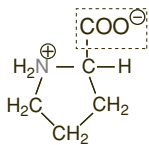
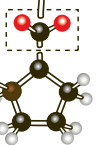
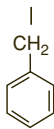
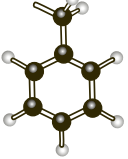
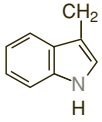
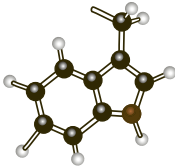
Les 20  $\alpha$ -aminoacides à partir desquels sont construites toutes les protéines connues sont caractérisés par leur chaîne latérale ; chaque  $\alpha$ -aminoacide porte un nom, d'abord abrégé selon un code à trois lettres puis à une lettre, qui permet d'écrire et de comparer les séquences des protéines. Selon la nature de la chaîne latérale, on distingue trois groupes d' $\alpha$ -aminoacides : apolaires, polaires neutres et polaires ionisables. Le caractère apolaire ou polaire des aminoacides est fondamental car il conditionne leur hydrophilie ou leur hydrophobicité et détermine le repliement et la conformation des chaînes polypeptidiques au sein des cellules. De plus, la structure des chaînes latérales détermine la réactivité de chaque  $\alpha$ -aminoacide. L'ensemble de ces réactivités confère aux protéines la possibilité d'effectuer les interactions discriminatives qui constituent leur fonction.

Les chaînes latérales des  $\alpha$ -aminoacides apolaires (*Tableau 1.1*) sont de nature aliphatique (G, A, V, L, I, M et P) ou aromatique (F et W) et donc très hydrophobes. G, avec sa chaîne latérale réduite à un hydrogène, joue un rôle particulier dans la conformation des chaînes polypeptidiques (p. 14). Les chaînes latérales de A, V, L et I, de plus en plus volumineuses, ont un caractère hydrophobe croissant ; I, avec son C $\beta$  chiral, a quatre stéréoisomères dont un seul est rencontré dans les protéines. M, avec son méthyl thioéther apolaire, a été sélectionnée de préférence à la norleucine  $[-(\text{CH}_2)_3-\text{CH}_3]$  en raison de son soufre nucléophile. P, avec son hétérocycle pyrrolidine, est un  $\alpha$ -aminoacide qui joue un rôle particulier dans la conformation des protéines (p. 14). F et W peuvent s'associer par recouvrement de leur nuage électronique  $\pi$  à nombre d'autres constituants biochimiques.

Les chaînes latérales des  $\alpha$ -aminoacides polaires non ionisables (*Tableau 1.2*) ont des groupes fonctionnels hydroxyle, sulfhydryle ou amide. Les hydroxyles de S et de T, polaires mais très peu ionisables, confèrent un certain caractère hydrophile à la chaîne latérale aliphatique et l'hydroxyméthyle de S peut être très réactif au sein des protéines ; T a deux centres de chiralité, mais un seul stéréoisomère est rencontré dans les protéines. L'hydroxyle de Y est plus ionisable ( $\text{p}K_a = 10,5$ ) que celui de S ou de T, ce qui rend Y moins hydrophobe que F. Bien que la chaîne latérale de C soit assez hydrophobe, son sulfhydryle légèrement ionisable ( $\text{p}K_a = 8,4$ ) lui confère une grande réactivité ; C peut former des liaisons H avec des atomes O ou N, perdre son proton et devenir un ion thiolate chargé négativement, et enfin contracter avec une autre C, au sein d'une chaîne polypeptidique ou entre deux chaînes polypeptidiques, une liaison covalente, dite liaison disulfure, susceptible de contribuer fortement à la stabilité de la conformation d'une protéine. L'amide de N et de G, qui ne s'ionise pas, est peu réactif mais il est polaire et peut jouer le rôle d'accepteur ou de donneur dans des liaisons H.


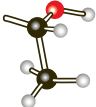
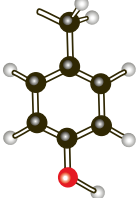

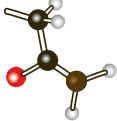
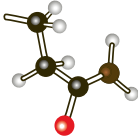
Les chaînes latérales des  $\alpha$ -aminoacides polaires ionisables (*Tableau 1.3*) possèdent des groupes fonctionnels très hydrophiles dont la charge est fonction du pH. Au pH physiologique, les chaînes latérales de D et de E se terminent par un  $\beta^-$  ou un  $\gamma$ -carboxyle ionisé qui contribue à la charge négative des protéines lorsqu'il se situe à leur périphérie. Les chaînes latérales de K, R et H se terminent par des groupes azotés basiques,  $\epsilon$ -amine primaire, groupe guanidinium ou hétérocycle imidazole, respectivement ; au pH physiologique, celles de K et de R sont protonées et apportent des charges positives aux protéines ; celle de H est un catalyseur extrêmement efficace car dans sa forme non ionisée, l'azote portant l'hydrogène est un électrophile et un donneur de proton tandis que l'autre azote est un nucléophile et un accepteur de proton.

Tableau 1.1

AMINOACIDE		CHAÎNE LATÉRALE	
Nom	Abréviations	Groupe fonctionnel	Modèle moléculaire
Glycine	Gly, G	H	
Alanine	Ala, A	CH <sub>3</sub>	
Valine	Val, V		
Leucine	Leu, L		
Isoleucine*	Ile, I		
Méthionine	Met, M		
Proline	Pro, P		
Phénylalanine	Phe, F		
Tryptophane	Trp, W		

\* Seul isomère rencontré dans les protéines.

Tableau 1.2

AMINOACIDE		CHAÎNE LATÉRALE	
Nom	Abréviations	Groupe fonctionnel	Modèle moléculaire
Sérine	Ser, S	$\begin{array}{c}   \\ \text{CH}_2 \\   \\ \text{OH} \end{array}$	
Thréonine*	Thr, T	$\begin{array}{c}   \\ \text{H} \text{---} \text{C} \text{---} \text{OH} \\   \\ \text{CH}_3 \end{array}$	
Tyrosine	Tyr, Y	$\begin{array}{c}   \\ \text{CH}_2 \\   \\ \text{C}_6\text{H}_4 \\   \\ \text{OH} \end{array}$	
Cystéine	Cys, C	$\begin{array}{c}   \\ \text{CH}_2 \\   \\ \text{S H} \end{array}$	
Asparagine	Asn, N	$\begin{array}{c}   \\ \text{CH}_2 \\   \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{NH}_2 \end{array}$	
Glutamine	Gln, Q	$\begin{array}{c}   \\ \text{CH}_2 \\   \\ \text{CH}_2 \\   \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{NH}_2 \end{array}$	

\* Seul isomère rencontré dans les protéines.

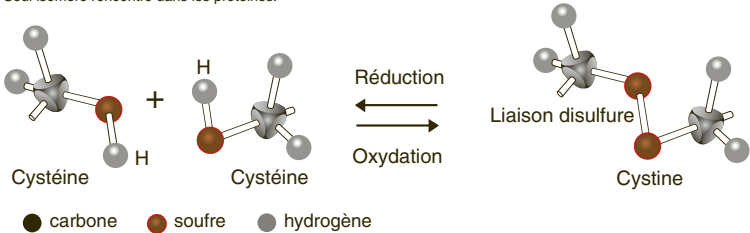




Tableau 1.3

AMINOACIDE		CHAÎNE LATÉRALE			
Nom	Abréviations	Groupe fonctionnel		Modèle moléculaire	
Acide aspartique	Asp, D	groupe carboxyle protoné	ou	déprotoné	
		$\beta$ CH <sub>2</sub>   O=C-OH	$\xrightarrow{H^+}$ $pK_a = 3,9$ $\xleftarrow{H^+}$	$\beta$ CH <sub>2</sub>   C   O <sup>-</sup> O <sup>-</sup> ion aspartate	
Acide glutamique	Glu, E	groupe carboxyle protoné	ou	déprotoné	
		$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   O=C-OH	$\xrightarrow{H^+}$ $pK_a = 4,1$ $\xleftarrow{H^+}$	$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   C   O <sup>-</sup> O <sup>-</sup> ion glutamate	
Lysine	Lys, K	groupe amine protoné	ou	déprotoné	
		$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   $\delta$ CH <sub>2</sub>   $\epsilon$ CH <sub>2</sub>   ion alkylammonium ⊕ NH <sub>3</sub>	$\xrightarrow{H^+}$ $pK_a = 10,5$ $\xleftarrow{H^+}$	$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   $\delta$ CH <sub>2</sub>   $\epsilon$ CH <sub>2</sub>   :NH <sub>2</sub>	
Arginine	Arg, R	groupe guanidine protoné	ou	déprotoné	
		$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   $\delta$ CH <sub>2</sub>   NH   ion guanidinium C   H <sub>2</sub> N <sup>⊕</sup> NH <sub>2</sub>	$\xrightarrow{H^+}$ $pK_a = 12,5$ $\xleftarrow{H^+}$	$\beta$ CH <sub>2</sub>   $\gamma$ CH <sub>2</sub>   $\delta$ CH <sub>2</sub>   NH   HN=C:NH <sub>2</sub>	
Histidine	His, H	hétérocycle imidazole protoné	ou	déprotoné	
		$\beta$ CH <sub>2</sub>   ion imidazolium HN <sup>⊕</sup> NH	$\xrightarrow{H^+}$ $pK_a = 6,0$ $\xleftarrow{H^+}$	$\beta$ CH <sub>2</sub>   HN-NH	

## LA CHAÎNE POLYPEPTIDIQUE ET SES CONFORMATIONS

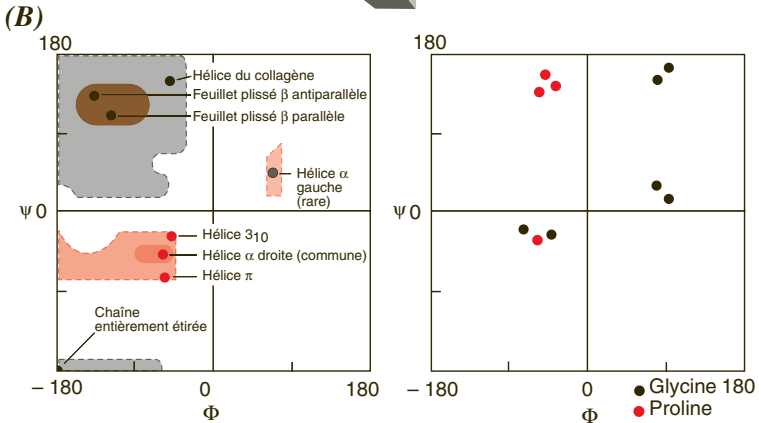
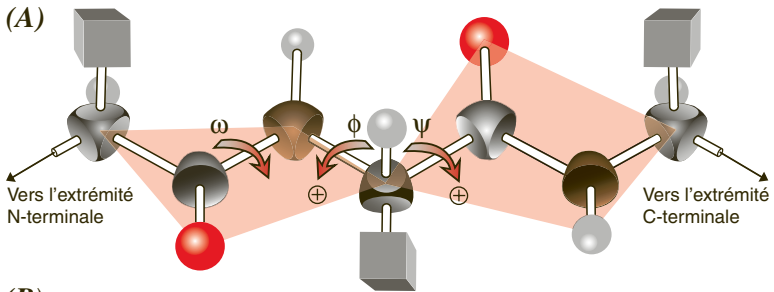
Une chaîne polypeptidique (*Figure 1.2A*) est formée par des aminoacides unis par des liaisons peptidiques. Elle se présente donc comme une succession de plans contenant des unités peptidiques  $-C\alpha-CO-NH-C\alpha-$  qui crée une chaîne principale d'où se projettent les chaînes latérales, ou résidus aminoacide. Elle est définie par la séquence des aminoacides qui détermine la structure primaire. Les propriétés propres à la liaison peptidique, associées aux propriétés particulières des résidus aminoacide unis selon une séquence donnée, déterminent les interactions morphogénétiques internes qui imposent spontanément une structure et une conformation propres à chaque chaîne polypeptidique.

Une chaîne polypeptidique a deux types de liberté de rotation (*Figure 1.2A*) qui lui permettent d'adopter les conformations qui caractérisent sa structure secondaire :

- la liberté de rotation d'un angle  $\Phi$  autour de la liaison unissant le carbone  $\alpha$  à l'azote amidique (liaison  $C\alpha-N$ ) ;
- la liberté de rotation d'un angle  $\Psi$  autour de la liaison unissant ce même carbone  $\alpha$  au carbone du carbonyle (liaison  $C\alpha-C$ ).

Selon une convention internationale, les rotations mesurées par  $\Phi$  et  $\Psi$  ont le sens indiqué dans la *Figure 1.2A* et leur champ est de  $180^\circ$ . Ces deux angles ont la valeur de  $+180^\circ$  (ou, ce qui revient au même, de  $-180^\circ$ ), lorsque la chaîne est totalement étirée. En examinant les conformations définies par  $\Phi$  et  $\Psi$ , il est possible de reconnaître celles qui ne sont pas possibles en raison d'un encombrement stérique et d'établir les tableaux de Ramachandran (*Figure 1.2B*). Pour la plupart des résidus, les couples de valeurs stériquement permis sont groupés essentiellement dans deux régions. Les valeurs de  $\Phi$  comprises entre  $-50^\circ$  et  $-60^\circ$  ou entre  $-120^\circ$  et  $-140^\circ$  apparaissent comme privilégiées. Certaines valeurs de  $\Psi$  déterminent des conformations régulières de la chaîne principale lorsque plusieurs résidus consécutifs ont les mêmes couples de valeurs de  $\Phi$  et de  $\Psi$ . Ce sont des conformations hélicoïdales droites, dites  $\alpha$ , ( $\Psi = -47^\circ$ ) ou formées à partir de brins  $\beta$ , dites en feuillets plissés  $\beta$  antiparallèles ou parallèles ( $\Psi = +135^\circ$  et  $+113^\circ$ , respectivement). Elles sont favorisées et stabilisées par les liaisons hydrogène qui s'établissent entre les CO et les NH polaires de la chaîne principale qui devient alors beaucoup moins hydrophile. La glycine est une importante exception ; avec son hydrogène comme chaîne latérale, cet  $\alpha$ -aminoacide peut prendre un beaucoup plus grand nombre de conformations que les autres résidus et permettre à la chaîne principale d'adopter certaines conformations. La proline, en raison de la présence de son cycle pyrrolidine, voit son angle  $\Phi$  bloqué à  $-65^\circ$  environ. Les hélices et les feuillets plissés sont universellement présents dans les chaînes polypeptidiques des protéines dont ils constituent la structure secondaire.

Figure 1.2



Conformation	Angles (degrés)			Résidus/tour	Translation/résidu (Å)
	$\phi$	$\psi$	$\omega$		
Hélice $\alpha$ droite	- 57	- 47	180	3,6	1,50
Hélice $\alpha$ gauche	+ 57	+ 47		3,6	1,50
Hélice $3_{10}$ droite	- 49	- 26	180	3,0	2,00
Hélice $\pi$ droite	- 57	- 70	180	4,4	1,15
Hélice gauche du collagène	- 51	+ 153		3,0	3,13
Feuillet plissé $\beta$					
antiparallèle	- 139	+ 135	- 178	2,0	3,40
parallèle	- 119	+ 113	180	2,0	3,20
Chaîne étirée	180	180			

● carbone ● oxygène ● azote ● hydrogène

## 1. Hélice $\alpha$

Dans une hélice  $\alpha$  (Figures 1.3A et 3B), les chaînes latérales des résidus aminoacides se projettent vers l'extérieur et n'interfèrent pas avec l'hélice elle-même, sauf celle du résidu prolyl dont le dernier carbone est lié à l'azote amidique. Les oxygènes des CO pointent vers l'extrémité C-terminale, les hydrogènes des NH vers l'extrémité N-terminale et des liaisons hydrogène s'établissent entre les CO et les NH ; individuellement, ces liaisons n'ont pas une grande énergie de stabilisation mais leur accumulation stabilise fortement l'hélice  $\alpha$ , en particulier au sein des régions hydrophobes des protéines. L'orientation des groupes peptidiques étant la même tout au long de l'hélice, leurs moments dipolaires se cumulent et une hélice  $\alpha$  de  $n$  résidus a un moment dipolaire résultant de  $n \times 3,5$  unités Debye, avec son extrémité N-terminale chargée positivement et son extrémité C-terminale chargée négativement. Ces charges peuvent attirer des ligands possédant des charges opposées. Ainsi, des ligands chargés négativement, en particulier ceux qui contiennent des groupes phosphoryl, se lient fréquemment à l'extrémité N-terminale des hélices  $\alpha$  qui, de plus, peut les positionner favorablement par des liaisons hydrogène spécifiques grâce à ses groupes NH libres. Les hélices qui ne font pas partie d'un site de liaison ont souvent un résidu aminoacide chargé négativement à leur extrémité N-terminale ou positivement à leur extrémité C-terminale. Ce résidu est dit compensateur de dipôle car il interagit avec le dipôle de l'hélice et stabilise la conformation hélicoïdale. L'extrémité C-terminale des hélices  $\alpha$ , où les derniers CO sont libres, est polaire et presque toujours localisée à la surface des protéines. Beaucoup d'hélices  $\alpha$  sont amphotères, en ce sens qu'elles présentent essentiellement des résidus apolaires sur un côté et des résidus polaires sur l'autre. Dans les protéines globulaires, la longueur des hélices  $\alpha$  peut varier considérablement, de quatre ou cinq résidus aminoacide à plus d'une quarantaine, avec une moyenne de 10, soit trois tours sur une longueur de 15 Å environ. Dans la représentation schématique d'une protéine, on donne aux hélices  $\alpha$  la forme d'un ruban enroulé en spirale.

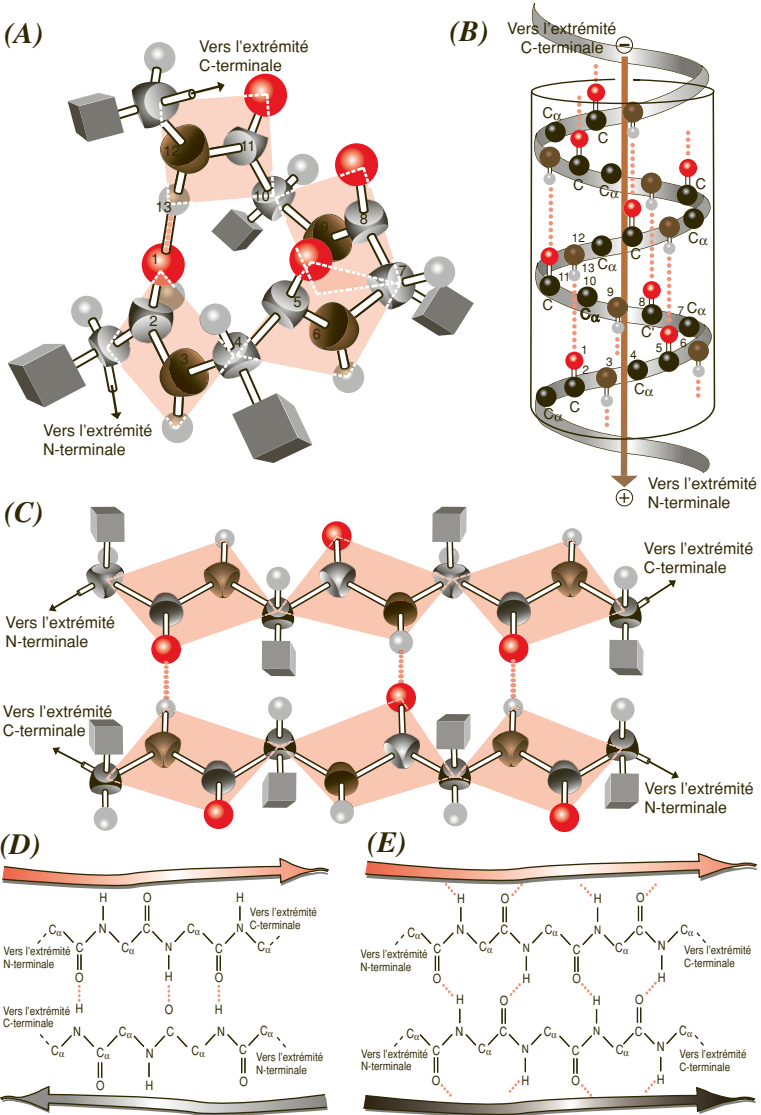
## 2. Brins $\beta$ et feuilletts plissés $\beta$

Le brin  $\beta$  est une conformation constituée d'un fragment de chaîne polypeptidique presque totalement étirée. Isolée, cette conformation n'est pas stable car il n'y a pas d'interactions entre atomes non directement liés ; en revanche, elle est stabilisée lorsqu'elle est incorporée dans des feuilletts plissés  $\beta$  où des liaisons hydrogène s'établissent entre les CO et les NH appartenant à des brins adjacents (Figure 1.3C). Les brins peuvent y être soit antiparallèles (Figure 1.3D), soit parallèles (Figure 1.3E). Dans la représentation schématique d'une protéine, on donne aux brins  $\beta$  la forme d'une flèche.

## 3. Boucles et coudes

Les hélices et des feuilletts plissés sont connectés par des régions de conformation non répétitive et souvent irrégulière, de longueur variable, les coudes ou les boucles.

Figure 1.3



## REPLOIEMENT DES CHAÎNES POLYPEPTIDIQUES DES PROTÉINES GLOBULAIRES HYDROSOLUBLES

L'étude de la structure à haute résolution de nombre de protéines a fait apparaître les principes fondamentaux qui gouvernent la structure des protéines globulaires (*Figure 1.4A*). Ces dernières contiennent des hélices et des feuillets plissés qui sont des éléments de structure secondaire régulière où les résidus consécutifs adoptent la même conformation. Les hélices et les feuillets sont connectés par des boucles de conformation non répétitive et de longueur variable, qui permettent les changements de direction de la chaîne polypeptidique nécessaires à la protéine pour qu'elle prenne sa forme globulaire. Si la combinaison d'hélices et/ou de feuillets plissés constitue le core hydrophobe interne stable de la protéine, les boucles, essentiellement constituées de résidus hydrophiles, se situent à la surface de cette dernière et peuvent donc entrer en interaction avec l'environnement aqueux. De plus, les groupes CO et NH de la chaîne principale des boucles n'établissent pas de liaisons hydrogène entre eux mais en contractent avec les molécules d'eau du milieu ambiant.

Les caractères les plus remarquables des molécules de protéines sont leur complexité et leur manque de symétrie. Une telle irrégularité de structure est en fait nécessaire aux protéines globulaires pour qu'elles puissent remplir leurs fonctions. Elles doivent en effet reconnaître des milliers de molécules très diverses au sein des cellules au moyen d'interactions tridimensionnelles très précises, ce qui nécessite des structures irrégulières et variées presque à l'infini.

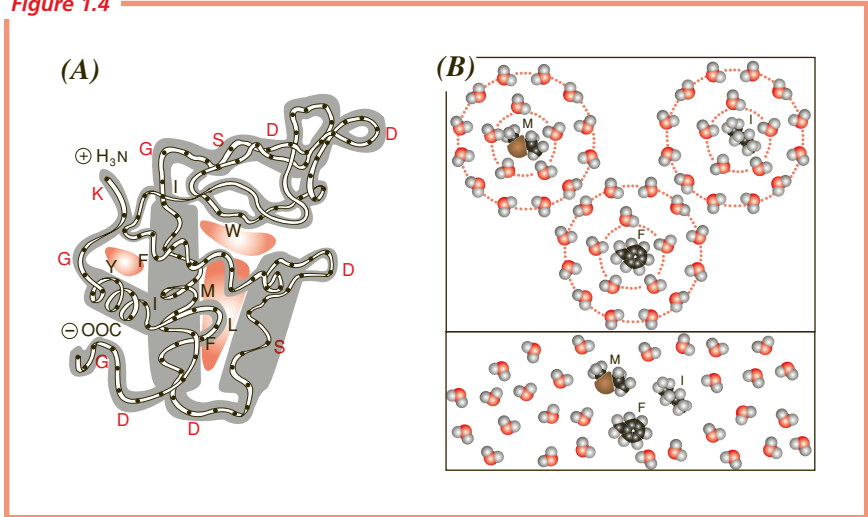
Les résidus aminoacide localisés à l'intérieur de la protéine sont presque exclusivement hydrophobes, tandis que ceux qui sont à sa périphérie sont hydrophiles. Il existe donc une force qui impose aux chaînes polypeptidiques des protéines globulaires hydrosolubles de se replier spontanément de façon à créer un core hydrophobe et une surface hydrophile. Quelle est la nature de cette force ?

### Le repliement des protéines globulaires hydrosolubles dépend de plusieurs types d'interactions

Le repliement des protéines globulaires et la stabilisation de leur conformation biologiquement fonctionnelle dépendent de plusieurs types d'interactions non covalentes : interactions hydrophobes, liaisons hydrogène, interactions de van der Waals et liaisons ioniques. Bien que prises individuellement ces interactions soient faibles, la somme de toutes celles qui entrent en jeu dans une protéine stabilise efficacement la conformation native de cette dernière. Le processus de repliement des protéines s'effectue grâce à la stabilisation progressive d'intermédiaires. Une chaîne polypeptidique se condense d'abord en un globule fondu qui possède l'essentiel de la structure secondaire, mais pas la structure tertiaire ; ce dernier évolue ensuite vers la conformation native de la protéine.

La formation des globules fondus résulte d'interactions hydrophobes (*Figure 1.4B*) qui constituent la composante la plus importante de la force qui dirige le repliement des protéines. Les protéines sont plus stables en milieu aqueux lorsque leurs résidus aminoacide hydrophobes sont dissimulés à l'intérieur de la protéine que lorsqu'ils sont solvatés par l'eau. Cette tendance des groupes hydrophobes à s'associer les uns avec les autres au sein de la molécule protéique leur permet d'éviter le contact avec les molécules d'eau de leur

Figure 1.4



environnement. L'agrégation de la plupart des résidus aminoacide hydrophobes à l'intérieur de la protéine impose alors un repliement de la chaîne polypeptidique caractérisé par la formation d'un core hydrophobe d'où l'eau est exclue et le déploiement des résidus polaires hydrophiles à la surface de la protéine où ils sont au contact direct de l'eau avec laquelle ils entrent en interaction (Figure 1.4A). La chaîne principale est constituée par les enchaînements  $-\text{C}\alpha-\text{CO}-\text{NH}-\text{C}\alpha-$  où les groupes CO et NH sont polaires et hydrophiles. Ces derniers perdent leur caractère polaire et leur hydrophilie lorsqu'ils sont unis par les liaisons hydrogène au sein des structures secondaires en hélice ou en feuillet plissé qui jouent donc, à ce niveau, un rôle essentiel dans la formation du core hydrophobe.

La stabilisation de la structure native est assurée par des liaisons hydrogène qui se forment entre les régions périphériques de la chaîne principale et l'eau, entre la chaîne principale et des résidus polaires, entre deux résidus polaires ou entre un résidu polaire et l'eau. La plupart des liaisons hydrogène des protéines sont du type  $-\text{N}-\text{H}-\text{O}-$ , avec une longueur variant de 2,6 à 3,4 Å.

Certains contacts de van der Waals entre groupes non polaires optimisent aussi le repliement de la chaîne polypeptidique et contribuent à la stabilité des protéines globulaires hydrosolubles.

Les interactions ioniques entre résidus aminoacide de charges opposées, ainsi que la solvataion des résidus chargés à la surface de la protéine, aident à la stabilisation.

Enfin, bien qu'absentes dans les protéines intracellulaires, les liaisons disulfure intrachânes, de nature covalente, stabilisent très efficacement la conformation native des protéines globulaires sécrétées et les rendent moins sensibles à la dénaturation et à la dégradation.

## Le repliement des protéines *in vivo* est assisté par des isomérases et des chaperons moléculaires

Toutes les liaisons peptidiques adoptent la configuration *trans*, à l'exception de certaines liaisons X-Pro, où X représente un résidu quelconque, qui peuvent prendre la configuration *cis* (p. 14). La *cis*-proline a un rôle structural important car elle favorise la formation de coudes. L'isomérisation spontanée de la liaison X-Pro est lente en raison du caractère de double liaison partielle de cette dernière et elle serait une étape limitante du repliement des chaînes polypeptidiques si, *in vivo*, elle n'était accélérée par des peptidyl prolyl isomérases qui déforment la liaison peptidique de telle façon que les atomes C, O, et N ne soient plus coplanaires et que la barrière d'activation soit abaissée. De plus, la formation de liaisons disulfure correctes dans les protéines naissantes est accélérée, *in vivo*, par la disulfure isomérase, enzyme de large spécificité pour des substrats contenant des résidus cystéine. Ce dernier, en brassant les liaisons disulfure permet aux protéines naissantes de trouver rapidement les appariements thermodynamiquement les plus favorables. Enfin, au sein de la cellule, milieu très encombré, deux classes de chaperons moléculaires, les Hsp 70 et les chaperonines, orientent au niveau moléculaire le repliement des chaînes polypeptidiques en empêchant des conformations incorrectes ou des liaisons inappropriées avec d'autres constituants du milieu. Isomérases et chaperons ne font qu'assister le repliement des chaînes polypeptidiques ; elles ne le dirigent pas et leur présence ne contredit en rien le dogme de la morphogénèse autonome qui seule détermine la conformation des protéines, comme celle des autres structures complexes des êtres vivants.

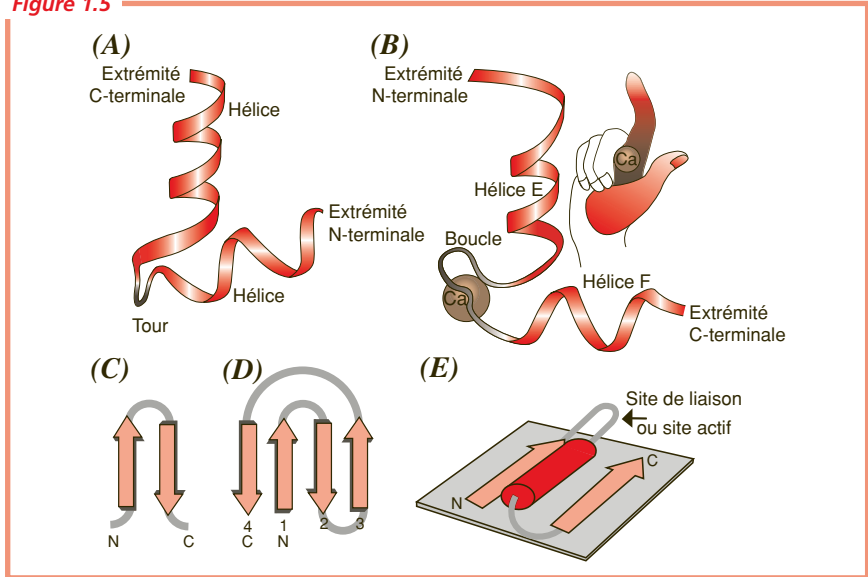
## DES MOTIFS DE REPLOIEMENT, OU STRUCTURES SUPERSECONDAIRES, SONT FORMÉS À PARTIR D'HÉLICES $\alpha$ ET/OU DE BRINS $\beta$

Au sein des protéines globulaires reploquées, des éléments de structure secondaire, hélices  $\alpha$  et/ou brins  $\beta$ , peuvent s'associer en des motifs appelés parfois structures supersecondaires et retrouvés fréquemment dans nombre de protéines. Ces motifs peuvent avoir une fonction particulière ou simplement être partie constituante d'unités structurales et fonctionnelles plus grandes et plus élaborées, les domaines.

Le motif doué de fonction le plus simple est constitué de deux hélices  $\alpha$  unies par un coude ou une boucle. Le motif hélice-coude-hélice (*Figure 1.5A*) s'associe spécifiquement aux régions régulatrices du DNA des Procaryotes. Le motif hélice-boucle-hélice, appelé aussi E-F hand (*Figure 1.5B*), dont la boucle est un site de liaison pour les ions  $\text{Ca}^{2+}$ , est régulièrement présent dans les protéines qui fixent le calcium et régulent de nombreux processus biologiques, telles les parvalbumines, la calmoduline, la troponine C ou certaines chaînes légères de la myosine.



Figure 1.5



Le motif en épingle à cheveux  $\beta$  (Figure 1.5C) est constitué de deux brins  $\beta$  adjacents antiparallèles connectés par un coude  $\beta$  en épingle à cheveux. Présent dans la plupart des structures  $\beta$ , soit isolément, soit comme constituant de feuillets plus complexes, il n'a pas en lui-même de fonction spécifique. La grecque est un motif constitué de quatre brins  $\beta$  où l'une des connexions n'est pas une épingle à cheveux (Figure 1.5D). Ce motif, souvent rencontré dans les protéines globulaires, n'a pas non plus en lui-même de fonction spécifique.

Le motif  $\beta$ - $\alpha$ - $\beta$  (Figure 1.5D) est constitué de deux brins  $\beta$  parallèles dont l'extrémité C-terminale de l'un est connectée par une hélice  $\alpha$  à l'extrémité N-terminale de l'autre. L'hélice  $\alpha$ , approximativement parallèle aux brins  $\beta$ , est étroitement associée à ces derniers dont elle dissimule les résidus aminoacide hydrophobes au milieu aqueux. Entre les brins  $\beta$  et l'hélice  $\alpha$ , il y a deux boucles dont les longueurs peuvent varier de deux à trois résidus aminoacide à plus d'une centaine. Un motif  $\beta$ - $\alpha$ - $\beta$  peut être considéré comme un enroulement hélicoïdal lâche d'un tour ; il est droit dans toutes les protéines connues, à l'exception de la subtilisine où il joue un rôle particulier. Que cet enroulement soit droit a d'importantes conséquences structurales et fonctionnelles lorsque les motifs  $\beta$ - $\alpha$ - $\beta$  sont inclus dans des domaines. Les deux boucles ont des fonctions différentes. Celle qui relie l'extrémité C-terminale du premier brin  $\beta$  à l'extrémité N-terminale de l'hélice  $\alpha$  est souvent impliquée dans la formation d'un site de liaison ou d'un site actif et possède des séquences d'acides aminés conservés dans les protéines homologues. L'autre boucle, en revanche, n'a pas de fonction connue.

## LES CHAÎNES POLYPEPTIDIQUES SONT ORGANISÉES EN DOMAINES

De nombreuses protéines sont constituées de domaines qui représentent des unités structurales et fonctionnelles fondamentales (*Figure 1.6A*). Les domaines sont des chaînes polypeptidiques ou des parties de longues chaînes polypeptidiques dont la taille peut varier considérablement, de 20 à 30 résidus aminoacide à plus de 300, avec une moyenne d'environ 100. Habituellement, chaque domaine a une fonction bien définie.

Les domaines sont constitués de combinaisons d'éléments de structure secondaire et de motifs connectés par des boucles. Le plus souvent, les hélices  $\alpha$  et les brins  $\beta$  des motifs formés à partir de régions voisines dans la structure primaire d'une chaîne polypeptidique sont proches dans la structure tridimensionnelle de cette dernière. Le nombre de combinaisons rencontrées dans les protéines est limité. De plus, certaines combinaisons semblent être favorisées. Ainsi, des domaines de structure semblable sont souvent retrouvés dans des protéines différentes, avec des fonctions différentes et des séquences d'acides aminés différentes.

## LES STRUCTURES DES PROTÉINES PEUVENT ÊTRE RÉPARTIES EN TROIS CLASSES

Trois groupes principaux peuvent être considérés : les structures  $\alpha$ , les structures  $\beta$  et les structures  $\alpha\beta$ . Dans les structures  $\alpha$ , les domaines sont constitués uniquement d'hélices  $\alpha$ , tandis que dans les structures  $\beta$ , ils sont formés de feuilletts  $\beta$  antiparallèles. Les structures  $\alpha\beta$  résultent de combinaisons de motifs  $\beta$ - $\alpha$ - $\beta$  qui forment un feuillet  $\beta$  parallèle entouré d'hélices  $\alpha$ . Par ailleurs, sont classées dans un groupe à part un certain nombre de protéines, le plus souvent de faible poids moléculaire, riches en liaisons disulfure ou en métaux et dont la structure présente une certaine distorsion.

Parmi les structures  $\alpha$ , le faisceau de quatre hélices (*Figure 1.6B*), au sein duquel des résidus hydrophobes délimitent un site actif central, est rencontré dans des protéines aussi différentes que la myohémérythrine, les cytochromes c et  $b_{562}$  et la ferritine. De même, le repliement globinique, où là aussi des résidus hydrophobes délimitent un site actif central (*Figure 1.6C*), est présent dans toutes les protéines d'un important groupe incluant les myoglobines et les hémoglobines, mais aussi les phycocyanines. Les caractéristiques structurales et les propriétés fonctionnelles du domaine globinique ont été très conservées tout au long de l'évolution, des Insectes et des nodules des racines des Végétaux, jusqu'aux Mammifères.

Les structures  $\beta$  sont constituées de brins  $\beta$  dont le nombre peut varier de quatre à plus de dix, souvent répartis en deux feuilletts antiparallèles plaqués l'un contre l'autre pour former un tonneau  $\beta$  affecté d'une certaine torsion (*Figure 1.6D*). Dans nombre d'entre elles, certains motifs sont des grecques. Les structures  $\beta$  ont un core hydrophobe constitué par les résidus aminoacide hydrophobes des brins  $\beta$  qui tapissent la partie interne du tonneau. Leur surface est constituée par les résidus aminoacide hydrophiles des brins  $\beta$  ainsi que par ceux des boucles qui unissent ces derniers. Les structures  $\beta$  sont présentes dans des protéines fonctionnellement très différentes, telles que des enzymes, des protéines de transport et des anticorps (*Figure 1.6E*).