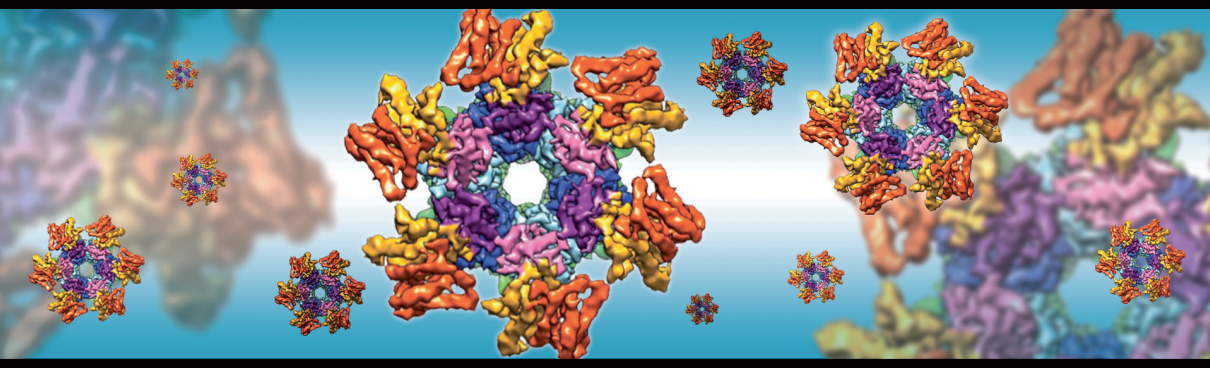


QuinteSciences

La structure des protéines

Jeannine Yon-Kahn, Guy Hervé et Jean-Luc Popot

edp sciences

La structure des protéines

Jeannine Yon-Kahn
Guy Hervé
Jean-Luc Popot

edp sciences

Remerciements

Les auteurs remercient bien vivement David Perahia, Florence Lederer et Liliane Mouawad pour avoir lu attentivement cet ouvrage et pour leurs suggestions constructives, ainsi qu'Annie Rousselet et Fabrice Giusti pour leur aide compétente pour la préparation de certaines figures et le Professeur Jean Bornarel pour ses encouragements à le publier. Les auteurs souhaitent également remercier Sophie Hosotte pour le soin et la compétence avec lesquels elle a assuré l'ordonnancement et la présentation de cet ouvrage.

Illustrations de couverture : Composition graphique à partir d'une structure obtenue en cryo-microscopie électronique du complexe cyanobactérien KaiCBA. Le complexe comporte six copies de la protéine KaiC (en bleu et vert), six de KaiB (en rose et violet) et douze de KaiA (en jaune et orange). En présence d'ATP, il oscille spontanément entre états assemblés et désassemblés avec une période d'environ 24 h, déterminant le cycle circadien de la bactérie. D'après Snijder et al., Science 355: 1181-1184, 2017; reproduit avec la permission de AAAS, copyright (2019).

Imprimé en France

ISBN : 978-2-7598-2242-3
ISBN : 978-2-7598-2243-0
Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, 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 intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1er de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.

© EDP Sciences 2019

Table des matières

Introduction 1

Chapitre 1 • Les acides aminés, motifs élémentaires 3

1.1 Identification 3

1.1.1 Asymétrie structurale 6

1.2 Propriétés optiques 7

1.2.1 Absorption de la lumière 7

1.2.2 Fluorescence 8

1.2.3 Propriétés rotatoires 9

1.3 Propriétés ioniques 10

1.4 Modifications chimiques du groupe aminé 11

1.4.1 Réactions d'addition sur une double liaison 15

1.4.2 Attaque par un réactif électrophile 16

1.5 Modifications chimiques du groupe carboxyle 16

1.6 Modifications chimiques des chaînes latérales 17

1.6.1 Groupes α-aminés et β-carboxyliques 17

1.6.2 Histidine 18

1.6.3 Sérine et thréonine 20

1.6.4 Tyrosine 20

1.6.5 Cytéine 22

1.6.6 Méthionine 26

1.6.7 Arginine 26
1.6.8 Tryptophane 26
1.7 Techniques de séparation des acides aminés 27
1.7.1 Électrophorèse 27
1.7.2 Chromatographie 27
Chapitre 2 • La structure primaire des protéines 31
2.1 Définition 31
2.2 Détermination de la structure primaire des protéines 32
2.2.1 Avancées historiques 32
2.2.2 Principes des méthodes de détermination des séquences protéiques 35
2.2.3 Spectrométrie de masse 35
2.3 Composition globale des protéines en acides aminés 36
2.4 Étude des alignements de séquences : taxonomie des protéines 37
2.5 Synthèse peptidique 37
Chapitre 3 • La structure secondaire 41
3.1 Définition 41
3.2 Conformations du squelette peptidique 43
3.2.1 Hélices 45
3.2.2 Déformation des hélices α 47
3.2.3 Structures β 47
3.2.4 Retournement de la chaîne : les virages 48
3.2.5 Autres structures régulières 49
3.2.6 Caractéristiques des structures non régulières 50
3.2.7 Méthodes d'étude des structures secondaires 50
3.3 Transformation des hélices α en structures β 52
Chapitre 4 • La structure supersecondaire 55
4.1 Définition 55
4.2 Association d'hélices α 55
4.3 Association de brins β 56
4.4 Interactions entre hélices α et segments β 58

4.5 Interactions entre hélices α ou brins β et segments dépourvus de structure régulière 62
4.6 Principaux motifs structuraux présents dans les protéines 62

Chapitre 5 • La structure tertiaire 67

5.1 Définition 67
5.2 Caractéristiques de la structure protéique 68
5.2.1 Absence de nœud dans la chaîne polypeptidique 68
5.2.2 Compacité des protéines globulaires 68
5.2.3 Les résidus hydrophobes ont tendance à se rassembler à l'intérieur de la protéine 68
5.2.4 Domaines structuraux 69
5.2.5 Protéines intrinsèquement non structurées 73
5.3 Classification des protéines basée sur la structure 73
5.4 Métalloprotéines 74
5.5 Méthodes de détermination des structures protéiques 75
5.5.1 Cristallographie aux rayons X 76
5.5.2 Résonance magnétique nucléaire 77
5.5.3 Microscopie électronique 77

Chapitre 6 • La structure quaternaire 79

6.1 Définition 79
6.2 Mode d'association des sous-unités 79
6.3 Protéines oligomériques possédant des sous-unités différentes 84
6.4 Assemblage de protéines 85
6.5 Méthodes d'étude des structures quaternaires 87

Chapitre 7 • Structure des protéines membranaires 89

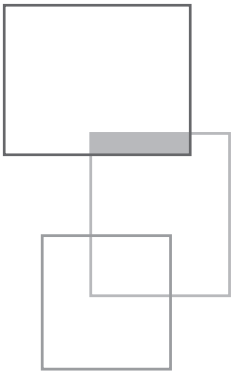
7.1 Localisation et fonction des protéines membranaires 89
7.2 Environnement membranaire 93
7.2.1 Composition et organisation 93
7.3 Expression, extraction, purification et manipulation des protéines membranaires en solution aqueuse 98
7.3.1 Expression 98
7.3.2 Extraction 100

7.3.3 Instabilité des protéines membranaires en solution détergente et alternative aux détergents 102
7.4 Étude structurale des protéines membranaires 107
7.4.1 Microscopie électronique 107
7.4.2 Radiocristallographie 109
7.4.3 Résonance magnétique nucléaire 110
7.4.4 Autres approches fournissant des informations structurales 111
7.5 Structure des protéines membranaires 112
7.5.1 Mode d'association avec la membrane. Structure quaternaire. Supercomplexes 113
7.5.2 Structure secondaire et tertiaire des régions transmembranaires 120
7.5.2 Interactions stabilisantes. Cofacteurs. Lipides 122

Chapitre 8 • Énergétique conformationnelle 129
8.1 Interactions intramoléculaires résultant de facteurs intrinsèques à la protéine 130
8.1.1 Interactions non covalentes 130
8.1.2 Interactions covalentes : les ponts disulfure 134
8.2 Interactions intramoléculaires influencées par le solvant 135
8.2.1 Liaisons hydrogène 135
8.2.2 Interactions électrostatiques 139
8.3 Interactions intramoléculaires déterminées par le solvant 140
8.3.1 Structure de l'eau liquide 140
8.3.2 Interactions hydrophobes 141
8.3.3 Rôle des interactions hydrophobes dans la stabilisation de la structure protéique 142
8.3.4 Surface accessible 143
8.4 Interactions entre le solvant et les molécules protéiques 144
8.5 Énergie conformationnelle totale d'une protéine 147

Chapitre 9 • Formation de la structure tridimensionnelle : le repliement des protéines 149
9.1 Le postulat d'Anfinsen et le paradoxe de Levinthal 151
9.2 Chemins de repliement 151
9.2.1 Modèles de repliement 151
9.2.2 Détection et caractérisation des intermédiaires de repliement 153

9.2.3 La nouvelle vision du repliement : le paysage énergétique et l'entonnoir de repliement 154
9.3 Repliement des protéines dans l'environnement cellulaire 156
9.4 Repliements incorrects, agrégation et conséquences pathologiques 158
9.5 Repliement des protéines membranaires, *in vivo* et *in vitro* 163
9.5.1 Synthèse *in vivo* 164
9.5.2 Repliement *in vitro* des protéines membranaires 168
Chapitre 10 • Prédictions conformationnelles 173
10.1 Introduction 173
10.2 Problèmes rencontrés dans les calculs conformationnels 174
10.2.1 Le problème des multiminima 174
10.2.2 Le grand nombre d'interactions 174
10.2.3 L'effet de solvant 175
10.3 Prédictions de structures secondaires 175
10.4 Prédictions de la structure tertiaire 178
Chapitre 11 • Dynamique structurale 183
11.1 Différents types de mouvements dans les protéines 184
11.2 Preuves expérimentales des mouvements dans les protéines 184
11.3 Approches théoriques : la dynamique moléculaire 186
11.4 Autres méthodes 189
Index 195
Crédits photographiques 199



Introduction

Les protéines, molécules fonctionnelles, participent à toutes les activités biologiques. Elles assurent une grande variété de fonctions. La plupart d'entre elles possèdent une propriété élémentaire qui est la reconnaissance spécifique et la fixation de ligands. Parmi les protéines, les enzymes qui exercent une activité catalytique sont impliqués dans toutes les nombreuses étapes du métabolisme. Certaines protéines sont impliquées dans des mécanismes de transport ou la transduction de signaux. Ainsi l'hémoglobine a pour fonction la fixation et le transport de l'oxygène. Des protéines qui possèdent une fonction mécanique sont responsables de la contraction musculaire et de la motilité cellulaire. Des mécanismes aussi divers que ceux qui participent aux phénomènes de l'immunité ou de la vision mettent aussi en jeu des protéines.

La diversité des fonctions assurées par les protéines n'a d'égale que l'étonnante complexité de leur construction. Les acides aminés constituent le motif élémentaire de toutes les protéines. Associés par des liaisons peptidiques, leur agencement dans la chaîne polypeptidique détermine la structure et la spécificité fonctionnelle de l'édifice.

L'émergence de la connaissance des protéines remonte au XIX^{e} siècle avec les premières identifications des acides aminés. Au cours du temps, les progrès dans l'identification et la purification des protéines, puis la détermination de la séquence grâce aux développements des méthodes chimiques, enfin la résolution de leur structure tridimensionnelle que permit l'introduction de méthodes physiques puissantes comme la cristallographie aux rayons X, la cryo-microscopie électronique et la résonance magnétique nucléaire ont conduit à une connaissance précise de ces molécules essentielles des organismes vivants. Les connaissances structurales résultent de la convergence d'un ensemble de méthodes relevant de la biologie, de la chimie et de la physique et plus récemment de l'informatique (Kessler & Ben-Tal, 2018). La Protein Data Bank comporte plus de 120 000 structures de protéines à haute résolution.

L'architecture des molécules protéiques se présente selon un ordre hiérarchique et montre plusieurs niveaux d'organisation. La structure primaire représente l'enchaînement séquentiel des acides aminés qui forment la chaîne polypeptidique. La structure secondaire désigne les éléments de structure régulière ou périodique. La structure supersecondaire résulte de l'interaction entre segments de structure secondaire. La structure tertiaire correspond à la conformation spatiale relativement compacte de la molécule et les protéines de grande taille peuvent comporter des domaines structuraux. La structure quaternaire est formée par l'association non covalente de plusieurs sous-unités identiques ou non. Ces structures ne sont pas figées. Elles sont animées de mouvements d'amplitude et de constantes de temps variées. Ces différents niveaux d'organisation seront abordés successivement.

En plus de ces connaissances fondamentales, le présent ouvrage introduit certains aspects nouveaux. C'est d'abord le cas du chapitre 1 dans lequel les réactions chimiques impliquant les chaînes latérales des acides aminés, si utiles aux chimistes des protéines et aux enzymologistes, font l'objet d'une présentation détaillée, car jusqu'ici cette information était dispersée dans la littérature, et de ce fait, difficilement accessible. Par ailleurs, les progrès considérables de ces dernières années en ce qui concerne la spécificité structurale et fonctionnelle des protéines membranaires justifient qu'un long chapitre leur soit consacré. La résolution de leur structure a nécessité le développement d'un mode de pensée et de technologies spécifiques. Du point de vue méthodologique, le développement de la RMN du solide et le développement de nouveaux logiciels pour l'exploration de la dynamique moléculaire sont également pris en compte. Enfin, lorsqu'il y a lieu, le développement historique de l'établissement de certains concepts fondamentaux est présenté.

Dans cet ouvrage, la structure des protéines est présentée dans sa dimension spatio-temporelle. La seule dimension spatiale risquait de faire apparaître les molécules protéiques comme des objets rigides. La dimension temporelle intervient dans les événements qui conduisent à la formation de la structure protéique et à la dynamique interne des molécules qui permet l'expression de leurs propriétés fonctionnelles.

Bibliographie

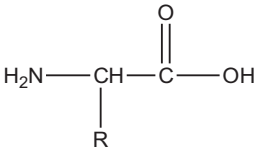
Kessler, A. & Ben-Tal, N. (2018) *Introduction to Proteins: Structure, Function and Motion.* Chapman and Hall/CRC.

1

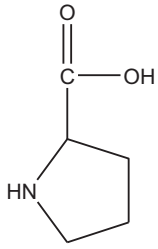
Les acides aminés, motifs élémentaires

1.1 Identification

Les protéines sont constituées d'un enchaînement d'acides aminés qui diffèrent par leur chaîne latérale. Les acides aminés ont pour formule générale :



Ils comportent un groupe aminé NH_2 et un groupe carboxylique COOH ainsi qu'une chaîne latérale R qui spécifie les différents acides aminés. L'atome de carbone central, le C^{α}, est lié aux groupes aminé et carboxyle. La proline fait exception, c'est un acide aminé dans lequel le groupement aminé est engagé dans une liaison covalente avec la chaîne latérale. Sa formule est la suivante :



Il existe vingt-deux acides aminés naturels qui sont codés par les triplets des nucléotides des gènes. Dans les premiers temps de la biologie moléculaire, seuls les vingt premiers acides aminés étaient connus. Deux autres acides aminés plus rares, la sélénocystéine et la pyrrolysine, ont été identifiés par la suite. Ainsi que le montre la figure 1.1, un alphabet à quatre lettres, les bases qui constituent les acides nucléiques, spécifient les vingt-deux acides aminés naturels.

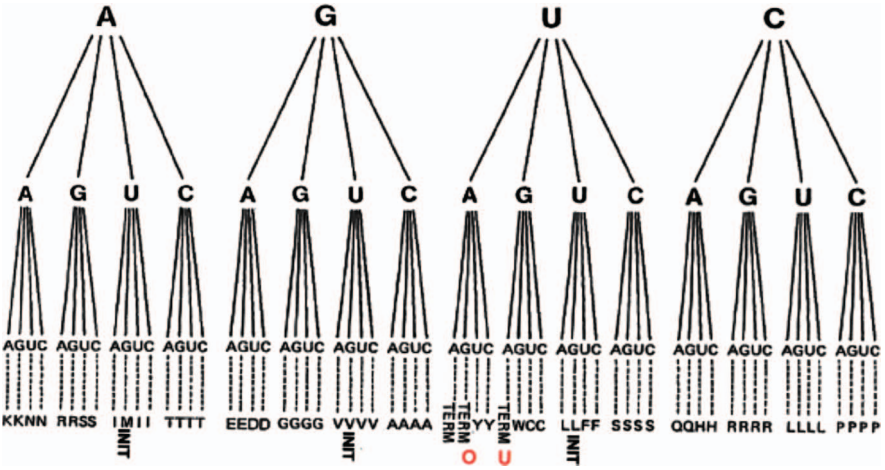


Figure 1.1 *Le code génétique. Dans l'ARN messager, les quatre bases (A adénine, G guanine, U uracile et C cytosine) spécifient les codons triplets. INIT et TERM indiquent les codons d'initiation et de terminaison, respectivement ;* **O** *et* **U** *sont la pyrrolysine et la sélénocystéine respectivement.*

Les différents acides aminés furent découverts entre 1805 et 1920. L'asparagine fut isolée en 1805 par Robiquet et Vauquelin à partir de jus d'asperges ; la leucine par Proust en 1818, et la glycine par Braconnet en 1820 à partir d'un hydrolysat de déchets animaux. En 1846, Liebig obtint la tyrosine. La glutamine fut isolée en 1866 à partir de gluten de blé, l'arginine et l'histidine en 1886. La lysine fut obtenue par Drechsel en 1899 à partir de la caséine. La proline et l'hydroxyproline furent identifiées par E. Fischer en 1901, l'isoleucine par Erlich en 1904. La méthionine fut isolée en 1921. La phénylalanine et le tryptophane furent ensuite identifiés. En 1986, la sélénocystéine codée par le codon de terminaison UGA a été découverte. Seules quelques protéines ont une sélénocystéine dans leur séquence. La pyrrolysine, 22^{e} acide aminé, fut identifiée en 2002 au site actif d'une méthyl transférase d'une archée ; elle est codée par un autre codon de terminaison UAG.

Les structures des acides aminés sont présentées sur la figure 1.2.

Le tableau 1.1 ci-après donne la liste des acides aminés naturels avec leurs codes à trois lettres et à une lettre.

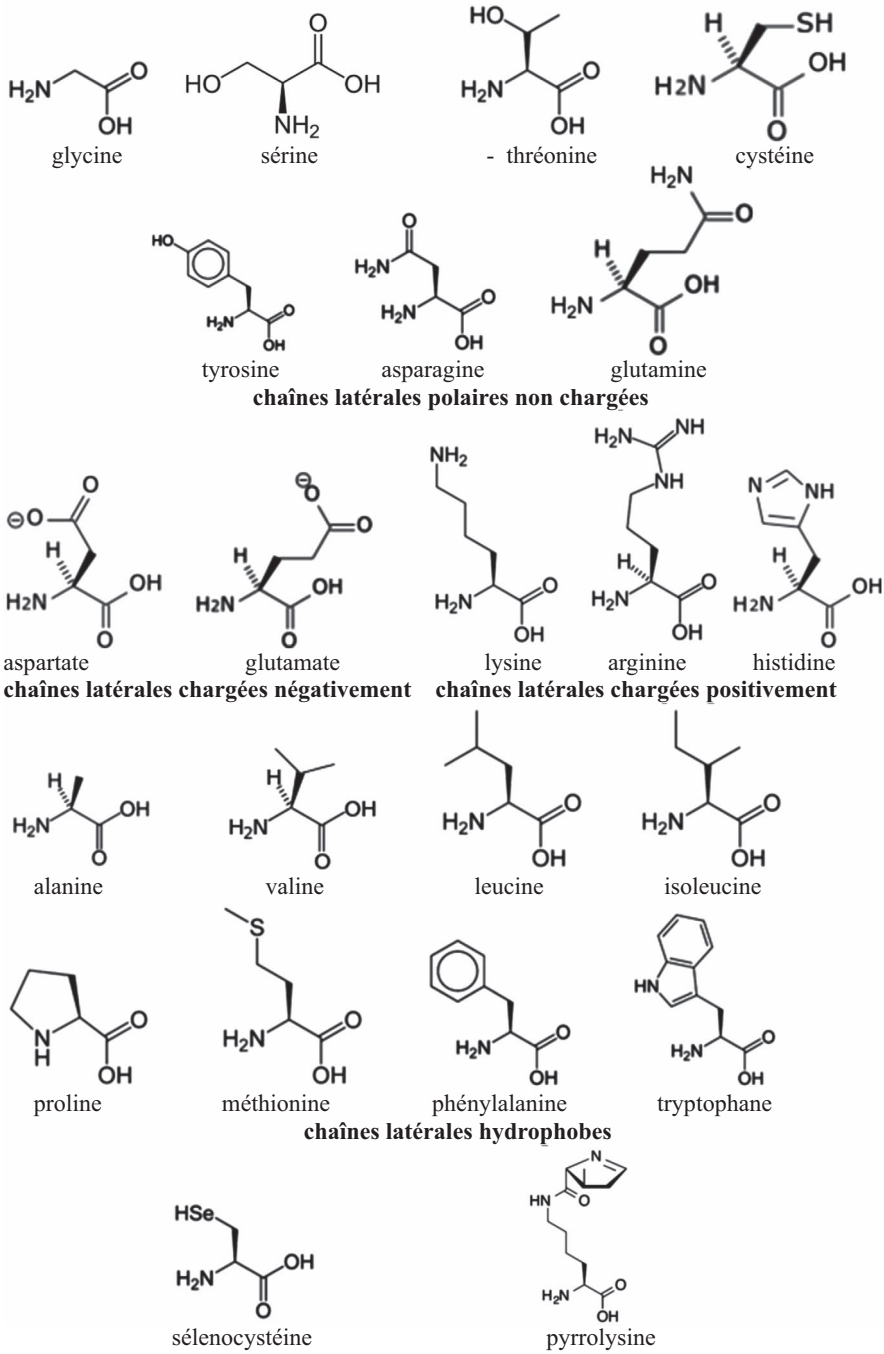


Figure 1.2 *Les acides aminés naturels.*

Tableau 1.1 *Liste des acides aminés avec leurs codes.*

Acide aminé	Code à trois lettres	Code à une lettre	Acide aminé	Code à trois lettres	Code à une lettre
Alanine	Ala	A	Lysine	Lys	K
Arginine	Arg	R	Méthionine	Met	M
Asparagine	Asn	N	Phénylalanine	Phe	F
Aspartate	Asp	D	Proline	Pro	P
Cystéine	Cys	C	Pyrrolysine	Pyl	O
Glutamate	Glu	E	Sérine	Ser	S
Glutamine	Gln	Q	Sélénocystéine	Sec	U
Glycine	Gly	G	Thréonine	Thr	T
Histidine	His	H	Tryptophane	Trp	W
Isoleucine	Ile	I	Tyrosine	Tyr	Y
Leucine	Leu	L	Valine	Val	V

1.1.1 *Asymétrie structurale*

Comme la plupart des molécules biologiques, les acides aminés sont asymétriques. À l'exception de la glycine qui porte deux atomes d'hydrogène sur le carbone α (R=H), les acides aminés naturels existent sous la configuration L.

Les deux énanthiomères L et D sont représentés sur la figure 1.4.

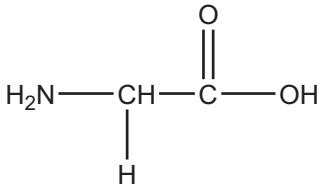


Figure 1.3 *Structure de la glycine.*

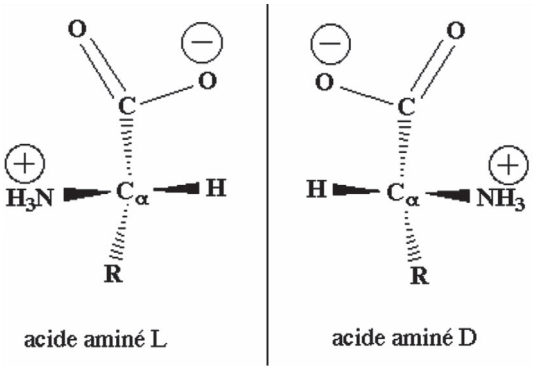


Figure 1.4 *Configuration des acides aminés.*

Il existe une règle simple pour reconnaître la configuration des acides aminés. Dans la configuration L, les trois groupes R, N et C apparaissent dans le sens des aiguilles d'une montre, et dans le sens opposé dans la configuration D. Ces deux conformations peuvent être distinguées par leurs propriétés rotatoires.

1.2 Propriétés optiques

1.2.1 *Absorption de la lumière*

Les acides aminés absorbent les rayons infrarouges par les groupes carboxyles et aminés et par les chaînes latérales. L'absorption du groupe carboxyle est à 5,9 μ ; elle est déplacée à 6,3 μ pour l'ion carboxylate. L'eau absorbant dans la zone des longueurs d'ondes utilisées, les mesures s'effectuent à l'état solide qui est la forme zwitterionique des acides aminés. Les fréquences caractéristiques se situent pour le groupe NH_3^+ à 3 200 cm^{-1} ; dans la région entre 1 700 et 1 300 cm^{-1}, on trouve les fréquences correspondant à l'élongation symétrique et asymétrique du groupe COO^-, à la déformation du groupe NH_3^+ et du $C^{\alpha}H$, et dans la région située entre 1 300 et 650 cm^{-1}, l'élongation de la liaison C-N et de la liaison C-C, ainsi que la déformation du COO^-.

Les acides aminés absorbent la lumière dans l'ultraviolet lointain ($\lambda < 190$ nm). Ceux qui possèdent une chaîne latérale aromatique absorbent dans le proche ultraviolet entre 200 et 300 nm. C'est le cas du tryptophane, de la tyrosine et de la phénylalanine. La tyrosine à pH 7 présente un maximum d'absorption à 275 nm, le tryptophane à 280 nm et la phénylalanine à 257 nm. La figure 1.5 montre le spectre d'absorption de ces acides aminés aromatiques à molarité égale.

Le tableau 1.2 indique leur coefficient molaire d'absorbance à la longueur d'onde maximum.

L'absorption de la lumière est une méthode simple et rapide pour déterminer la concentration d'un composé chromophore. L'absorption d'une lumière monochromatique dans un milieu homogène isotrope est proportionnelle à la longueur du trajet lumineux dans la solution et à la concentration du composé absorbant :

$$I = I_o\, e^{-klc}$$

Soit encore :

$$2{,}3 \log_{10} I_o/I = kcl$$

I_o étant l'intensité de la lumière incidente, I celle de la lumière sortante, c la concentration du composé absorbant, l la longueur du trajet optique. La longueur unitaire l est de 1 cm et $\varepsilon = k/2{,}3$ définit le coefficient d'extinction molaire de la substance. L'absorbance de la solution est donc :

$$A_\lambda = -\log_{10} I/I_o = \varepsilon c$$

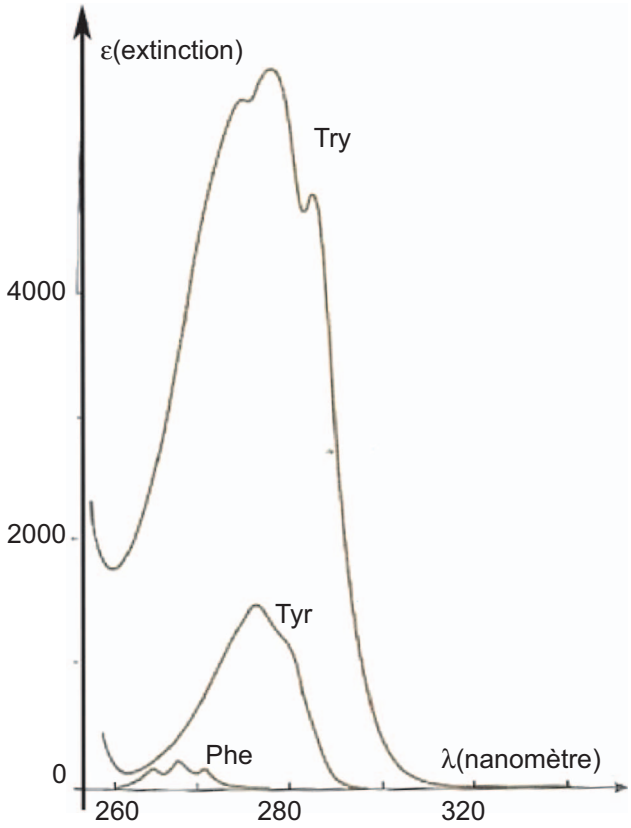


Figure 1.5 *Spectre d'absorption des acides aminés aromatiques.*

Tableau 1.2 *Coefficient molaire d'absorbance des acides aminés aromatiques à* λ_{max}.

Acide aminé pH 7, 0 25 °C	λ_{max}	$\varepsilon/M^{-1}\ cm^{-1}$
Tryptophane	280	5 600
Tyrosine	274	1 400
Phénylalanine	257	200

1.2.2 *Fluorescence*

La fluorescence est une émission lumineuse provoquée par l'excitation d'une molécule par absorption d'un photon. L'énergie absorbée porte les électrons dans l'état excité ; cette énergie est restituée sous forme lumineuse lorsque les électrons retournent à leur état fondamental. Dans la plupart des cas, une partie seulement de l'énergie est restituée sous forme lumineuse. La fréquence ν = E/h (E étant l'énergie et h la

constante de Planck) sera donc plus faible que celle de la lumière d'excitation et la longueur d'onde d'émission supérieure. Parmi les acides aminés, seul le tryptophane et la tyrosine présentent une fluorescence appréciable.

1.2.3 *Propriétés rotatoires*

Les acides aminés ont un pouvoir rotatoire. Les composés asymétriques dévient la lumière polarisée ; ils provoquent une rotation de son plan de polarisation. Le composé est lévogyre si la rotation est vers la gauche, dextrogyre si elle est vers la droite. L'angle de rotation du plan de polarisation représente le pouvoir rotatoire. Celui-ci dépend de la longueur d'onde de la lumière, de la température et de la concentration de la substance en solution. La loi de Biot définit le pouvoir rotatoire :

$$\alpha = [\alpha]_{\lambda T} \; l \; c$$

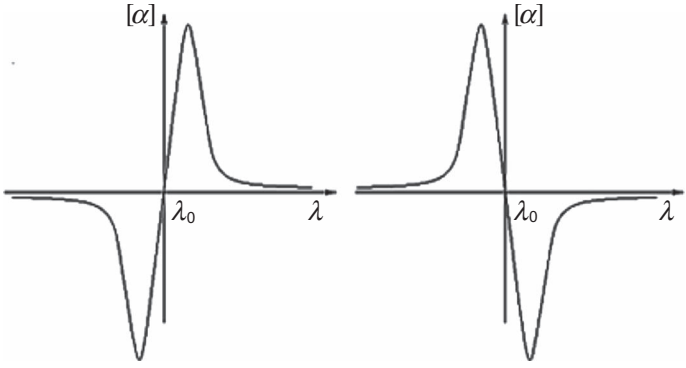
α est le pouvoir rotatoire ou angle de rotation du plan de polarisation exprimé en degrés.

l est l'épaisseur de la solution traversée exprimée en dm.

c est la concentration de la substance dissoute exprimée en $g \cdot cm^{-3}$.

$[\alpha]_{\lambda T}$ est le pouvoir rotatoire spécifique exprimé en $dm^{-1}\ g^{-1}\ cm^3$. Il dépend de la longueur d'onde qui est généralement celle de la raie D du sodium (λ = 589,3 nm) ; T, la température, les valeurs de $[\alpha]_{\lambda T}$ étant généralement données à 20 °C.

La variation de α avec la longueur d'onde définit la dispersion optique rotatoire. Cette dispersion présente des anomalies au niveau des chromophores lorsque ceux-ci sont engagés dans une structure asymétrique. Ces anomalies se traduisent par une inversion du pouvoir rotatoire appelée effet Cotton. La figure 1.6 montre l'allure des effets Cotton, positif et négatif.

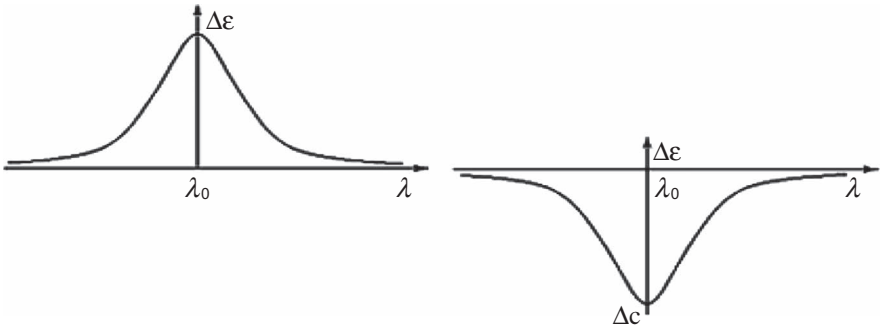


(a) Effet Cotton positif (b) Effet Cotton négatif

Figure 1.6 *Effets Cotton positif et négatif.*

On peut rendre compte du **dichroïsme circulaire** par l'existence de deux coefficients d'extinction différents pour les ondes polarisées circulairement droite et gauche.

$$\Delta\varepsilon = \varepsilon_G - \varepsilon_D$$



(a) Effet Cotton positif (b) Effet Cotton négatif

Figure 1.7 *Différence des coefficients d'extinction droite et gauche dans le cas des effets Cotton positif et négatif.*

Le dichroïsme circulaire est mesuré à des longueurs d'onde proches des bandes d'absorption des molécules. En pratique, on mesure l'ellipticité :

$$\Psi = K(\varepsilon_G - \varepsilon_D)$$

La mesure de l'ellipticité dans l'ultraviolet lointain dans la zone d'absorption de la liaison peptidique est une méthode expérimentale importante car elle donne des informations sur la structure secondaire des protéines.

1.3 Propriétés ioniques

Les acides aminés ont une structure bi-ionique faisant intervenir les deux groupes COOH et NH_2 susceptibles de gagner ou de perdre un proton.

$$\text{-COOH} \rightleftharpoons \text{-COO}^- + \text{H}^+$$
$$\text{-N}^+\text{H}^+_3 \rightleftharpoons \text{-NH}_2 + \text{H}^+$$

À l'état solide ainsi qu'en solution neutre, les acides aminés présentent deux groupes chargés ; ils sont dipolaires (zwitterion) :

$$^+\text{H}_3\text{N}-\underset{\displaystyle \text{R}}{\underset{|}{\text{CH}}}-\text{COO}^-$$

Leur état d'ionisation varie avec le pH. Le pK des groupes carboxyliques est environ de 3-4, celui du groupe aminé est de l'ordre de 10. Le point isoélectrique qui correspond

à la neutralité globale de la molécule est égal à la demi-somme des pK. La présence d'une chaîne latérale (radical R) possédant un groupe ionisable participe à la valeur du point isoélectrique. Le tableau 1.3 ci-après donne les valeurs du pK et du point isoélectrique des acides aminés. Dans une protéine, l'environnement peut modifier considérablement le pK et la réactivité de ces groupes.

Par suite de leur caractère amphotère, les acides aminés sont solubles dans les solvants polaires de constante diélectrique élevée tels que l'eau et ne sont que très peu ou pas solubles dans les solvants apolaires de faible constante diélectrique. La dissolution d'un acide aminé entraîne une augmentation de la constante diélectrique du solvant.

Tableau 1.3 *Valeur du pK et du point isoélectrique des acides aminés.*

Acide aminé	pK du groupe COOH	pK du groupe NH^{3+}	pK des chaînes latérales	Point isoélectrique
Glycine	2,4	9,8		6,06
Alanine	2,4	9,9		6,1
Valine	2,3	9,7		6,0
Leucine	2,3	9,7		6,03
Isoleucine	2,3	9,7		6,04
Sérine	2,2	9,4		5,7
Thréonine	2,1	9,1		5,6
Méthionine	2,2	9,3		5,7
Cystéine	1,9	10,3	8,3	5,1
Proline	2,0	10,6		6,3
Phénylalanine	2,6	9,2		5,9
Tyrosine	2,2	9,1	10,9	5,6
Tryptophane	2,4	9,4		5,9
Acide aspartique	2,0	10	3,9	2,8
Acide glutamique	2,1	10	4,3	3,2
Lysine	2,2	9,2	10,8	9,6
Arginine	1,8	9,0	12,5	11,2
Histidine	1,8	9,2	6,0	7,6
Asparagine	2,0	8,8		
Glutamine	2,2	9,1		

1.4 Modifications chimiques du groupe aminé

Le groupe aminé possède un pouvoir nucléophile en raison du doublet disponible : $-NH_2$ qui lui confère sa réactivité. Il participe ainsi à divers types de réactions à condition qu'il ne soit pas protoné.

L'**acylation** est l'attaque d'un carbone trigonal sp^2 par un groupe nucléophile, le doublet de l'azote du groupe aminé conduisant à la formation d'une liaison du type R′-CO-NH-R. Pour cette réaction, on utilise généralement des anhydrides d'acides :