

الجمهورية الجزائرية الديمقراطية الشعبية وزارة التعليم العالي والبحث العلمي

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

Année Universitaire 2024/2025

Ecole Supérieure en Sciences Biologique d'Oran (ESSB d'Oran)



Département des Classes Préparatoire ou Département du Second Cycle

Polycopie Pédagogique

Matière

Ingénierie des protéines

Niveau : 1ère année deuxième cycle

Option : Biologie Moléculaire

Spécialité : Biotechnologie

Filière : Sciences Biologiques

Domaine : Science de la Nature et de la Vie

Réalisé par :

Dr BOUKADOUM Ali

Matière enseignée pendant les Années Universitaires :

2022/2023

Matière : Ingénierie des protéines

Semestre : 2

UEF 2.2 : Ingénierie des protéines

Crédits : 4

Coefficient : 2

Unité d'Enseignement	VHS	V.H hebdomadaire				Coef	Crédits	Mode d'évaluation	
	14-16 sem	C	TD	TP	Autres*VHS			Continu	Examen
UE fondamentales						9	17	40 %	60 %
UEF1(O/P)									
UEF2 (O/P)									
Ingénierie des Protéines	45h00	01h30	01h30	-	50h00	2	4	X	X

a) Objectifs de l'enseignement

L'objectif de la formation est d'approfondir la compréhension des propriétés des protéines naturelles et d'introduire l'étudiant au domaine de l'ingénierie des protéines ainsi qu'aux notions théoriques et pratiques permettant d'appréhender les problématiques de la stabilité et du repliement des protéines.

b) Connaissances préalables recommandées

Biochimie structurale et fonctionnelle.

c) Contenu de la matière

Chapitre 1. Stabilité, repliement et dynamique des protéines

Chapitre 2. Thermodynamique de la stabilité et du repliement des protéines

Chapitre 3. Dénaturation des protéines

Chapitre 4. Repliement des protéines *in vivo*

Chapitre 5. Ingénierie de nouvelles interactions protéine-ligand

Chapitre 6. Ingénierie des enzymes

Chapitre 7. Modification chimique des protéines *in vitro* et *in vivo*

Chapitre 8. Ingénierie des protéines *in silico*

d) Mode d'évaluation : Contrôle continu : 40 % ; Examen : 60 %.

e) Références : Livres et photocopiés, sites internet, etc.

Table des matières	1
Avant-propos	4
Chapitre I : Composition et structures des protéines	5
I Acides aminés des protéines	5
I.1 Liaison peptidique	7
I.2 Acides aminés non standard	10
II Structure des protéines	11
II.1 Définition	11
II.2 Les liaisons qui stabilisent les protéines repliées	11
II.3 Détermination de la structure primaire	13
III Structure tridimensionnelles des protéines	18
III.1 Structures en hélice	19
III.2 Les structures β	21
III.3 Structure en coude et en boucle	22
IV Structure tertiaire	23
IV.1 Définition	23
IV.2 Structure tertiaire et conformation	23
V Domaines structuraux	26
V.1 Les structures α	29
V.2 Les structures β	30
V.3 Les structures α et β	32
VI Structure quaternaire	34
VII Symétrie dans les protéines	35
VII.1 Symétrie cyclique	35
VII.2 Symétrie dièdre	36
VII.3 Symétrie de rotation	37
Chapitre II : Repliement des protéines, dynamique et évolution structurale	38
I Repliement des protéines	38
I.1 Renaturation des protéines	38
I.2 Déterminants du repliement des protéines	41

a. Présence des hélices et des feuillets	41
b. Le repliement des protéines se fait essentiellement sous l'influence de résidus internes	42
c. La structure des protéines s'établit selon une hiérarchie	42
d. Les structures des protéines sont adaptables	43
I.3 Mécanismes de repliement	43
a. Il faut des mesures rapides pour suivre le repliement d'une protéine	43
b. Les étapes le plus précoces du repliement d'une protéine sont amorcées par un effondrement hydrophobe	45
c. La théorie du paysage pour le repliement des protéines	45
d. Le repliement des protéines suit une hiérarchie	47
e. L'inhibiteur trypsique pancréatique bovin (BPTI) prend sa conformation native selon un mécanisme séquentiel	48
Chapitre III : Repliement des protéines <i>in vitro</i> et <i>in vivo</i>, et rôle des protéines auxiliaires	49
I Repliement des protéines	49
I.1 Le modèle de nucléation-condensation	49
I.2 Le modèle de diffusion - collision, proposé par Karplus et Weaver (1976)	49
I.3 Le repliement Folding Units	50
I.4 Le repliement séquentiel et hiérarchique	50
I.5 Le modèle de la fermeture éclair hydrophobe (<i>Hydrophobic Zipper</i>), proposé par Dill et al. (1993)	50
I.6 Le modèle <i>Framework</i> (Ptitsyn, 1991)	50
I.7 Le modèle d'effondrement (collapse) hydrophobe	51
I.8 Le modèle du <i>Jigsaw Puzzle</i> (Harrison, Durbin, 1985)	51
II Rôle des protéines auxiliaires du repliement	53
II.1 Protéines disulfure isomérase	53
II.2 Réactions catalysées par la protéine disulfure isomérase	54
II.3 Peptidyl prolyl cis-trans isomérase	56
II.4 Chaperons moléculaires: le système GroEL/ES	56
II.5 Le système GroEL/ES fonctionne comme un moteur à deux temps	58
Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles	60
I Prédiction des structures secondaires	60
II Prédiction des structures tertiaires	63
III Dynamique des protéines	63

IV Maladies conformationnelles; amyloïdes et prions	64
IV.1 Les protéines amyloïdes	64
IV.2 Les maladies à prions	65
Chapitre V : Interaction protéine ligand et interaction allostérique	67
I Introduction	67
II Les ligands	67
III Mécanismes de liaison protéine-ligand	68
III.1 Modèle serrure et clé	69
III.2 Modèle ajustement induit	69
III.3 Modèle de sélection conformationnelle	71
IV Interaction allostérique	71
Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt	73
I Synthèse chimique des peptides et des polypeptides	73
II Méthodes de synthèse	74
II.1 La méthode de synthèse en phase solide SPPS	74
II.2 Exigences de la synthèse peptidique	77
II.3 Ancrage de la chaîne au support inerte	78
II.4 Activation des groupements fonctionnels	79
II.5 Détachement du polypeptide de la résine	80
III Synthèse chimique par fragment	81
IV Synthèse et bioproduction des protéines	81
Références bibliographiques	83

Avant-propos

À l'ère des technologies de pointe, les avancées en cristallographie, en spectroscopie et en bio-informatique ont révolutionné notre compréhension des protéines. Les protéines, ces macromolécules, présentes dans tous les organismes vivants, remplissent des fonctions vitales, et la capacité de manipuler leurs propriétés ouvre des perspectives infinies dans des secteurs tels que la biotechnologie, la médecine, l'industrie pharmaceutique et l'environnement.

L'ingénierie des protéines repose sur la compréhension détaillée de la relation structure-fonction des protéines. En exploitant les connaissances issues de la biochimie, de la biologie moléculaire et de la bio-informatique, les chercheurs peuvent altérer des séquences d'acides aminés pour conférer à une protéine des caractéristiques spécifiques, telles qu'une stabilité accrue, une activité enzymatique améliorée ou une affinité accrue pour un ligand particulier. Cette discipline permet également la création de nouvelles protéines qui n'existent pas dans la nature, capables de remplir des fonctions inédites ou d'améliorer des processus biologiques naturels.

Grâce aux technologies, l'ingénierie des protéines permet de résoudre des problèmes complexes en médecine, comme la conception de médicaments biologiques, le développement d'enzymes thérapeutiques ou la mise au point de vaccins de nouvelle génération. Les protéines recombinantes, par exemple, sont déjà utilisées dans des traitements contre le cancer, les maladies neurodégénératives et les troubles métaboliques.

Au-delà du domaine médical, cette discipline impacte des industries diverses. Dans l'agriculture, elle permet de créer des enzymes plus efficaces pour la production de bioénergies ou l'amélioration des cultures. En environnement, des protéines spécialement conçues aident à la bioremédiation, en décomposant des polluants ou en recyclant des matériaux. Ce domaine en constante évolution offre des opportunités inégalées pour façonner l'avenir des sciences biologiques et relever des défis globaux, avec le potentiel de transformer profondément notre compréhension des systèmes biologiques et d'apporter des solutions concrètes à des problèmes complexes.

Chapitre I : Composition et structures des protéines

I Acides aminés des protéines

Les protéines sont les macromolécules douées des propriétés les plus variées des systèmes vivants. Ce sont des polymères linéaires construits à partir d'unités monomériques appelées les acides aminés. Les protéines sont construites à partir d'un répertoire de 20 acides aminés. La séquence des aminoacides liés entre eux est appelée structure primaire. Les protéines se replioient spontanément en des structures tridimensionnelles. La structure tridimensionnelle formée par les liaisons hydrogène qui s'établissent entre les aminoacides proches les uns des autres s'appelle structure secondaire, tandis que la structure tertiaire est formée par des interactions à longue distance. La fonction des protéines est dépendante de structure tridimensionnelle. De nombreuses protéines ont une structure quaternaire ce qui signifie que la protéine fonctionnelle est composée de plusieurs chaînes polypeptidiques distinctes.

Les aminoacides sont les unités structurales de base des protéines. Un α -aminoacide est constitué d'un atome de carbone central, appelé carbone α lié à un groupe amine primaire à l'exception de la proline, un acide carboxylique, un atome d'hydrogène et un groupe R qui constitue la chaîne latérale (**Figure 01**). Les α -aminoacides sont des molécules chirales.

Seuls les aminoacides L sont des constituants des protéines.



Figure 01 : Structure générale d'un α -aminoacide.

Les aminoacides en solution à pH neutre se trouvent essentiellement sous forme d'ions dipolaire, le groupe amine est protoné et le groupe carboxyle déprotoné. En solution acide, le groupe amine est protoné. Les chaînes latérales, se différencient par leur dimension, leur forme, leur charge et leur capacité à contracter des liaisons hydrogène en (**Tableau 01**) :

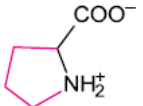
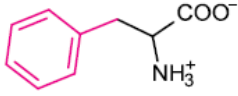
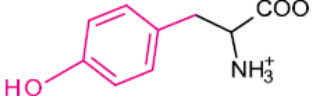
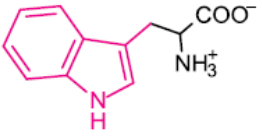
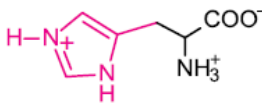
Aminoacides hydrophobes avec chaîne latérale non polaire.

Aminoacides avec chaîne latérale polaire non chargé ou neutre.

Aminoacides chargés positivement et aminoacides chargés négativement.

Chapitre I : Composition et structures des protéines

Tableau 01 : Nomenclature des α -aminoacides.

Nom	Codes		Formule (chaîne latérale en couleur)	Remarque
Glycine	Gly	G	$\text{H}-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	Non chiral
Chaîne latérale (R) apolaire aliphatique				
Alanine	Ala	A	$\text{CH}_3-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Valine	Val	V	$(\text{CH}_3)_2\text{CH}-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Leucine	Leu	L	$(\text{CH}_3)_2\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Isoleucine	Ile	I	$\text{CH}_3-\text{CH}_2-\text{CH}(\text{CH}_3)-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	Carbone β chiral
Proline	Pro	P		Acide aminé cyclique, fonction amine secondaire
Méthionine	Met	M	$\text{CH}_3-\text{S}-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	δ Méthyl thioéther
Chaîne latérale (R) aromatique				
Phénylalanine	Phe	F		β Phényl
Tyrosine	Tyr	Y		β <i>p</i> -Hydroxyphényl $\text{pK}_a = 10.6$
Tryptophane	Trp	W		β Indolyl
Chaîne latérale (R) polaire non chargée à pH 7				
Sérine	Ser	S	$\text{HO}-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Thréonine	Thr	T	$\text{CH}_3-\text{CH}(\text{OH})-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	Carbone β chiral
Cystéine	Cys	C	$\text{HS}-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	β Thiol $\text{pK}_a = 8,2$
Asparagine	Asn	N	$\text{NH}_2-\text{C}(=\text{O})-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Glutamine	Gln	Q	$\text{NH}_2-\text{C}(=\text{O})-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	
Chaîne latérale (R) chargée - à pH 7				
Aspartate	Asp	D	$^- \text{OOC}-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	β Carboxyle $\text{pK}_a = 3,8$
Glutamate	Glu	E	$^- \text{OOC}-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	γ Carboxyle $\text{pK}_a = 4,2$
Chaîne latérale (R) chargée + à pH 7				
Lysine	Lys	K	$^+ \text{H}_3\text{N}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	ϵ Amine $\text{pK}_a = 10,6$
Arginine	Arg	R	$\text{H}_2\text{N}-\text{C}(=\text{NH}_2^+)-\text{NH}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_3^+)-\text{COO}^-$	δ Guanidino $\text{pK}_a = 12,5$
Histidine	His	H		β Imidazole $\text{pK}_a = 6$

Chapitre I : Composition et structures des protéines

Les fonctions amine et carboxyle du carbone alpha disparaissent au cours de la synthèse des protéines, suite à la formation des liaisons peptidiques. Dès lors, comme on ne peut plus parler d'acides aminés, on dira qu'une chaîne polypeptidique est une séquence de résidus d'acides aminés. Les codes, à une ou trois lettres, facilitent l'archivage et l'analyse informatique de ces séquences.

I.1 Liaison peptidique

Les peptides et les protéines sont le résultat de l'enchaînement d'acides aminés reliés entre eux par une liaison covalente qui est en fait une liaison amide formée par déshydratation entre le groupement amine d'un acide aminé et le groupement carboxyle d'un autre acide aminé. Cette liaison est également appelée liaison peptidique, d'où le nom de polypeptides également donné aux protéines (**Figure 02**).

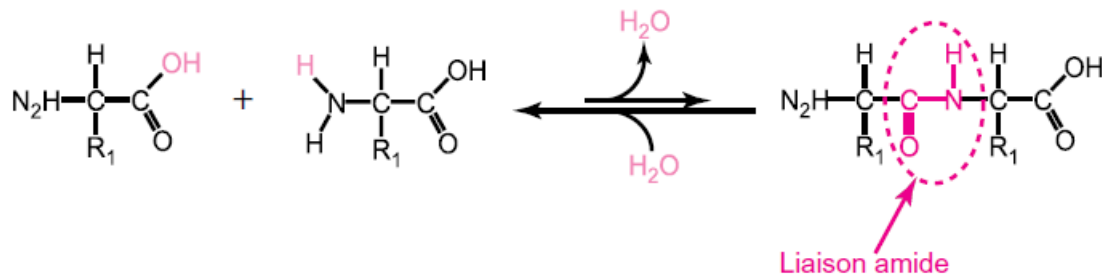


Figure 02 : Formation de la liaison peptidique.

La liaison peptidique est très stable à pH neutre malgré un équilibre de réaction en faveur de l'hydrolyse. La résistance particulière à l'hydrolyse est liée à la résonance des électrons entre les atomes d'oxygène et d'azote de la liaison amide. Ce phénomène de résonance est également à l'origine de deux autres propriétés importantes de la liaison peptidique : (i) il augmente la polarité de la liaison lui octroyant un moment dipolaire et (ii) il donne à la liaison peptidique un caractère de double liaison (**Figure 03**).

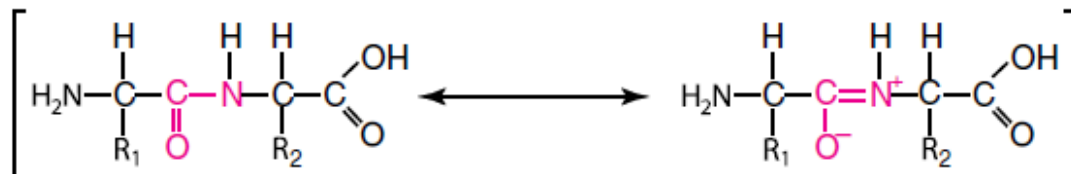


Figure 03 : Caractère double liaison de la liaison peptidique.

Chapitre I : Composition et structures des protéines

La libre rotation autour de la liaison amide est limitée à deux configurations, soit la configuration *trans* où les deux carbones α sont localisés de part et d'autre de la liaison amide et la configuration *cis*, où les deux carbones α sont localisés du même côté de la liaison (Figure 04).

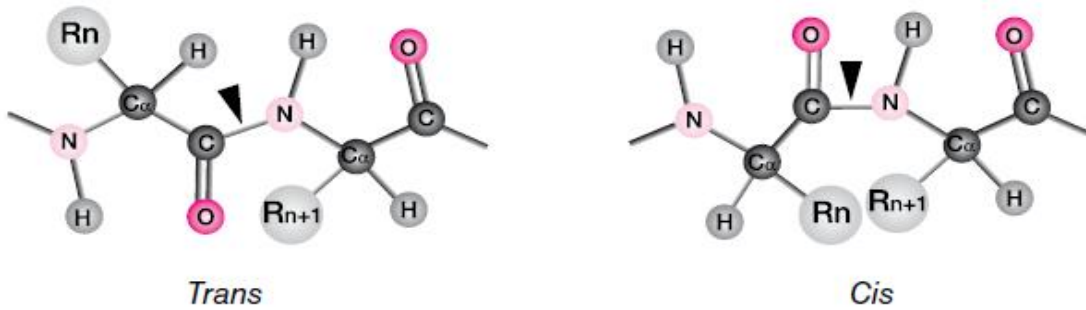


Figure 04 : Liaisons peptidiques cis et trans.

Dans les protéines, les liaisons peptidiques sont en configuration *trans* à cause des encombrements stériques des chaînes latérales.

Il y a toutefois une rotation possible entre le carbone α et l'azote de la liaison amide, ainsi qu'entre le carbone α et le carbone de la liaison amide. Ce sont les valeurs de ces angles qui définissent pour partie le type de structure secondaire adoptée par le polypeptide (Figure 05).

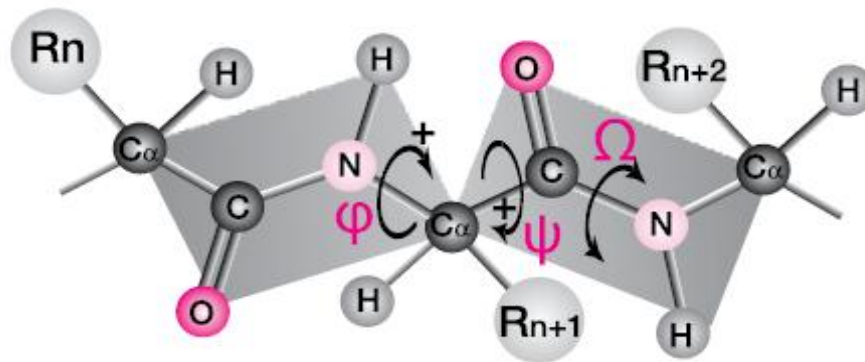


Figure 05 : Planéité de la liaison peptidique et liberté de rotation.

Les protéines sont des polymères linéaires par union de groupe α carboxyle d'un acide aminé avec un groupe α amine d'un autre acide aminé. Les liaisons peptidiques sont très stables parce que leur vitesse d'hydrolyse est extrêmement lente, leur durée de vie en solution aqueuse en l'absence de catalyseur est très longue. Une chaîne polypeptidique est constituée

Chapitre I : Composition et structures des protéines

d'une partie répétitive régulière, appelée chaîne principale ou squelette, et d'une partie variable comportant des chaînes latérales spécifiques. Le squelette polypeptidique est riche en liaisons hydrogènes potentielles. Ces groupes entrent en interaction l'un avec l'autre, ainsi qu'avec les groupes fonctionnels des chaînes latérales pour stabiliser les structures.

L'enchaînement successif des acides aminés reliés entre eux par une liaison peptidique constitue la structure primaire ou séquence de la protéine. Dans cette structure chaque acide aminé prend le nom de résidu. La chaîne polypeptidique obtenue a une polarité avec un groupe α -amine présent à une extrémité (extrémité N-terminale) et un groupe α -carboxyle présent à l'autre extrémité (extrémité C-terminale). Par convention une séquence en acides aminés est toujours écrite en partant du résidu N-terminal. Pour décrire la séquence des acides aminés, le suffixe « -yl » est ajouté à tous les résidus, sauf au C-terminal. La chaîne peptidique contient une partie répétitive nommée chaîne principale ou squelette et une partie variable correspondant aux chaînes latérales des résidus d'acides aminés (**Figure 06**).

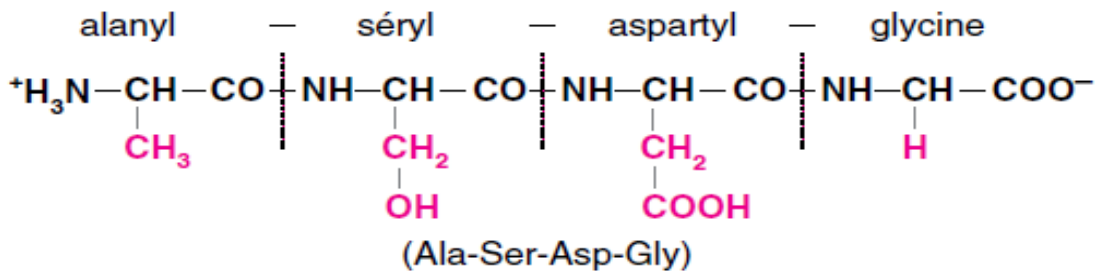


Figure 06 : Séquence successive d'une liaison peptidique.

Le moyen le plus courant et peut être le plus utile de classer les acides aminés des protéines est de considérer les polarités de leurs chaînes latérales (groupe R). En effet, les protéines se replient pour prendre leurs conformations natives essentiellement en vertu de la tendance à éloigner leurs chaînes latérales hydrophobes du contact de l'eau et à hydrater leurs chaînes latérales hydrophiles.

Les chaînes latérales non polaires ont des formes et des dimensions variées.

Les chaînes latérales polaires non chargées ont des groupes hydroxyle, amide ou thiol.

Les chaînes latérales polaires peuvent être chargées positivement ou négativement.

Chapitre I : Composition et structures des protéines

I.2 Acides aminés non standards

Les 20 acides aminés protéiques ne sont pas les seuls acides aminés présents dans les systèmes biologiques. On trouve des acides aminés non standards en proportions souvent importantes dans les protéines ainsi que dans les polypeptides biologiquement actifs (**Figure 07**). Dans tous les cas connus, ces acides aminés insolites résultent de la modification spécifique d'un résidu d'acide aminé intervenue après la synthèse de la chaîne polypeptidique.

Citons la 4-hydroxyproline et la 5-hydroxylysine. Ces deux résidus d'acide aminé sont des constituants structuraux importants d'une protéine fibreuse, le collagène. Les acides aminés des protéines qui forment des complexes avec les acides nucléiques sont souvent modifiés. Par exemple, les protéines ribosomiales et les protéines de chromosomes appelées histones peuvent être méthyles, acétylées ou phosphorylés.

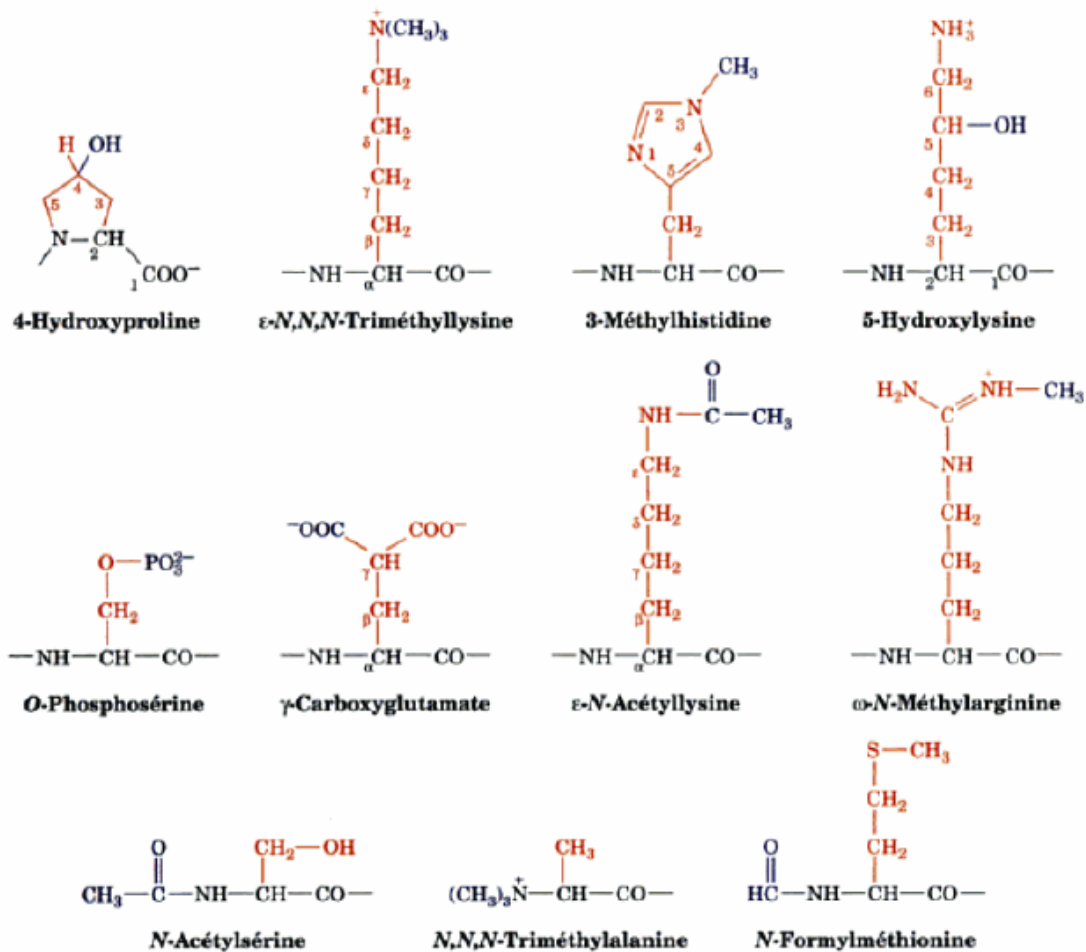


Figure 07 : Acides aminés non standards.

Chapitre I : Composition et structures des protéines

II Structure des protéines

II.1 Définition

Les protéines sont au cœur de l'action dans les processus biologiques. Tout d'abord, elles régulent l'ensemble des réactions chimiques. Elles agissent ici directement, en tant qu'enzymes, et indirectement en tant que messagers chimiques, hormones et comme récepteurs de ces hormones. Les protéines participent activement au contrôle de l'expression de l'information génétique, tout en étant des produits de cette expression.

La fonction d'une protéine ne peut être comprise que par sa structure, c'est-à-dire les relations tridimensionnelles entre les atomes qui les constituent. La description des protéines se fait selon quatre niveaux d'organisations.

1. La **structure primaire** d'une protéine est la séquence en acides aminés de sa chaîne polypeptidique.
2. La **structure secondaire** est l'arrangement spatial local des atomes du squelette d'un polypeptide.
3. La **structure tertiaire** désigne la structure tridimensionnelle du polypeptide.
4. La **structure quaternaire** désigne l'arrangement spatial des sous-unités.

II.2 Les liaisons qui stabilisent les protéines repliées

Dans la plupart des protéines, les liaisons amides du squelette sont les seules liaisons covalentes qui fixent les résidus les uns aux autres, il peut y avoir des liaisons covalentes supplémentaires présentes sous la forme de ponts disulfure entre les chaînes latérales. Cependant la stabilisation d'une protéine repliée provient de liaison ou interaction polaire faible non covalente. Dans une protéine repliée, les principales interactions sont : les liaisons hydrogènes et les liaisons de van der Waals.

Les interactions de van der Waals se produisent à chaque fois que des nuages électroniques fluctuants sur un atome ou un groupe d'atomes.

Les liaisons hydrogène sont établies lorsqu'un atome hydrogène possédant une charge positive partielle due au fait qu'il est lié à un atome plus électronégatif, est attiré par un atome voisin possédant une charge négative partielle (**Tableau 02**).

Chapitre I : Composition et structures des protéines

Tableau 02 : Les propriétés des liaisons hydrogène de l'eau ont des effets importants sur la stabilité des protéines.

Interaction	Exemples	$\Delta_r G'^{\circ}$ (kJ·mol ⁻¹) *	Intervention
Liaison H		-10 à -40	Acides nucléiques (appariement des bases) Protéines (structures secondaires)
Force de van der Waals (London)		- 4**	Membranes biologiques Repliement des protéines ADN (empilement des paires de bases)
Liaison ionique		- 40	Protéines (ponts salins)
Interaction hydrophobe		- 4 à -13***	Protéines (stabilisation du repliement) Membranes biologiques

Chapitre I : Composition et structures des protéines

II.3 Détermination de la structure primaire

La plupart des chaînes polypeptidiques naturelles contiennent entre 50 et 2000 résidus aminoacide et sont habituellement appelées protéines. Les plus grandes protéines connues sont les protéines musculaires qui sont constituées de 27000 aminoacides. Le poids moléculaire moyen d'un résidu aminoacide est de 110g/mol environ. Le poids moléculaire de la plupart des protéines est donc compris entre 5500 et 220000 g/mol. La masse d'une protéine peut être exprimée en dalton, un dalton est égal à une unité de masse atomique. Une protéine d'un poids moléculaire de 50000 g/mol a une masse de 50000 daltons, ou 50 kd. Dans certaines protéines, la chaîne polypeptidique linéaire présente des liaisons croisées. Les plus habituelles sont les liaisons disulfure, formées par l'oxydation d'un couple de résidus cystéine. Le bloc formé par la liaison de deux cystéines est appelé cystine (**Figure 08**).

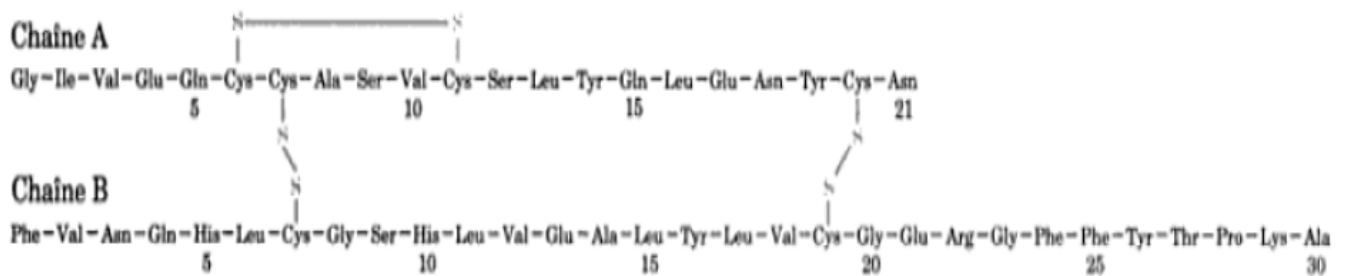


Figure 08 : Structure primaire de l'insuline bovine.

Les chaînes polypeptidiques sont flexibles et essentiellement plane. Ainsi pour un couple d'acides aminés, six atomes se situent dans un même plan ; atome de carbone α et le groupe CO du premier acide, et l'atome de carbone α et le groupe NH du second acide aminé. Cette liaison résonne entre la simple et la double liaison. La résonance (la délocalisation des électrons entre plusieurs atomes) a deux conséquences importantes ;

Elle augmente la polarité de la liaison peptidique. La polarité de la liaison peptidique peut jouer un rôle de premier plan dans le comportement des protéines repliées.

De plus, la liaison peptidique possède le caractère d'une double liaison partielle, ce qui signifie que la libre rotation des atomes est limitée (**Figure 09**).

Chapitre I : Composition et structures des protéines

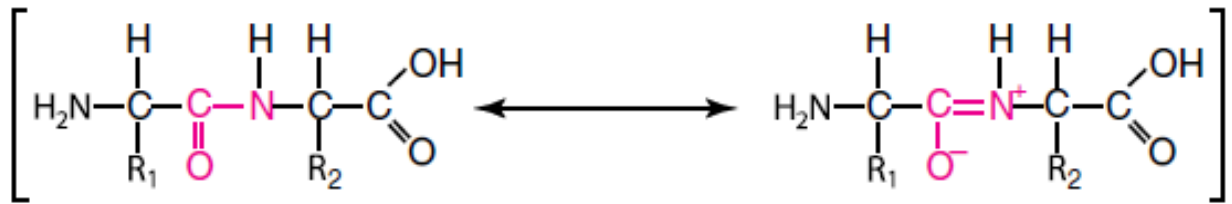


Figure 09 : caractère double liaison de la liaison peptidique.

L'angle de la liaison peptidique N-C α à la liaison peptidique adjacente s'appelle l'angle de rotation phi ϕ

L'angle de la liaison peptidique C-C α à la liaison peptidique adjacente s'appelle l'angle de rotation psi ψ

Une chaîne polypeptidique a deux types de liberté de rotation qui lui permettent d'adopter les conformations qui caractérisent sa structure secondaire (**Figure 10**) :

- la liberté de rotation d'un angle ϕ autour de la liaison unissant le carbone α à l'azote amidique (liaison C α -N) ;
- la liberté de rotation d'un angle ψ autour de la liaison unissant ce même carbone α au carbone du carbonyle (liaison C α -C).

Les deux unités peptidiques rigides adjacentes peuvent tourner autour de ces liaisons et prendre diverses orientations. Cette liberté de rotation de deux liaisons de chaque résidu aminoacide permet aux protéine de se replier de nombreuses manières différentes.

Est-ce que toutes les combinaisons de rotation ϕ et ψ sont possibles ?

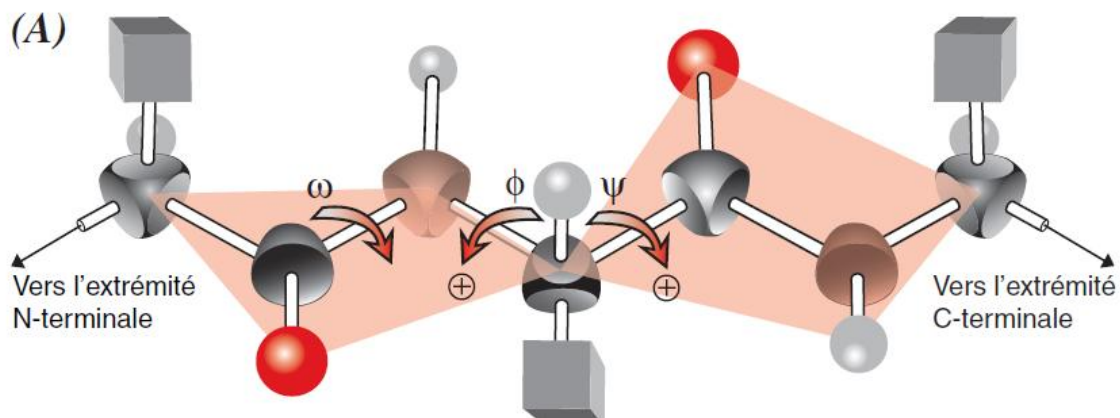


Figure 10 : Liberté de rotation autour de la liaison lipidique.

Chapitre I : Composition et structures des protéines

Ramachandran a pu montrer que beaucoup de combinaisons sont interdites à cause des collisions stériques entre atomes. Les valeurs permises peuvent être visualisées sur un graphe bidimensionnel appelé diagramme de Ramachandran. Les trois quarts de combinaisons possible sont simplement exclues à cause des chocs stériques locaux. Les empêchements stériques, et le fait que deux atomes ne puissent pas être à la même place et au même moment, peuvent représenter un principe puissant d'organisation. Les régions les plus favorables sont représentées en vert foncé, les régions limite sont représentées en vert clair (**Figure 11**).

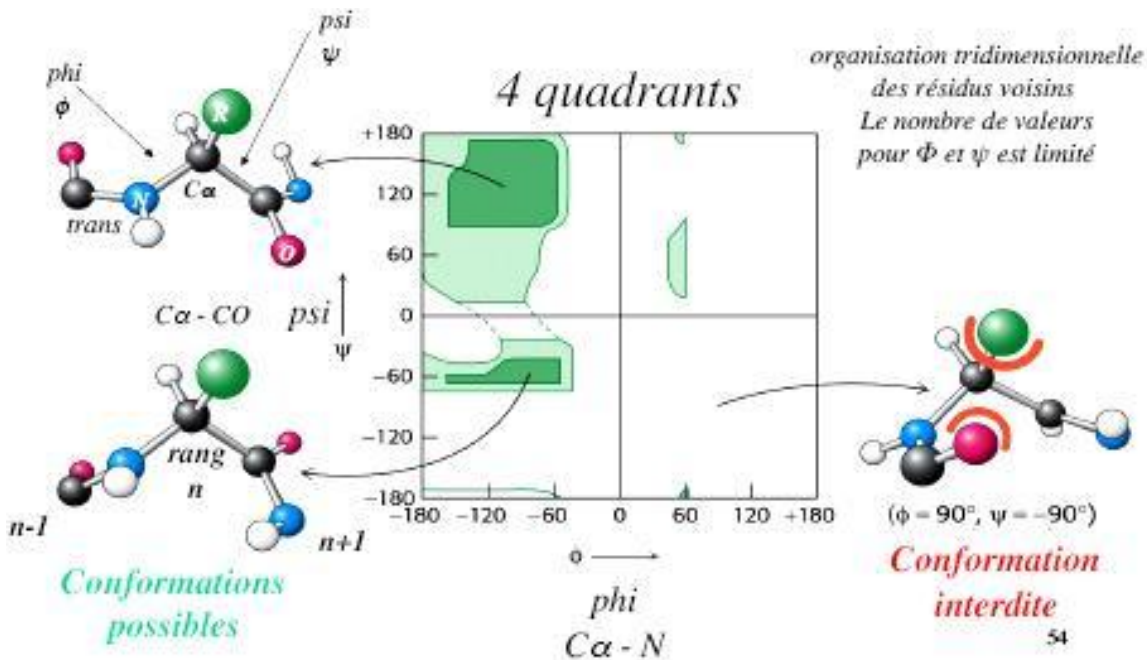


Figure 11 : Diagramme de Ramachandran.

Selon une convention internationale, les rotations mesurées par ϕ et ψ ont le sens de 180° . Ces deux angles ont la valeur de $+180^\circ$ (ou, ce qui revient au même, de -180°), lorsque la chaîne est totalement étirée.

En examinant les conformations définies par ϕ et ψ , il est possible de reconnaître celles qui ne sont pas possibles en raison d'un encombrement stérique et d'établir les tableaux de Ramachandran.

Chapitre I : Composition et structures des protéines

Les valeurs de ϕ comprises entre -50° et -60° ou entre -120° et -140° apparaissent comme privilégiées. Certaines valeurs de ψ déterminent des conformations régulières de la chaîne principale lorsque plusieurs résidus consécutifs ont les mêmes couples de valeurs de ϕ et de ψ . Ce sont des conformations hélicoïdales droites, dites α , ($\psi = -47^\circ$) ou formées à partir de brins β , dites en feuillets plissés β antiparallèles ou parallèles ($\psi = +135^\circ$ et $+113^\circ$, respectivement) (**Figure 12**).

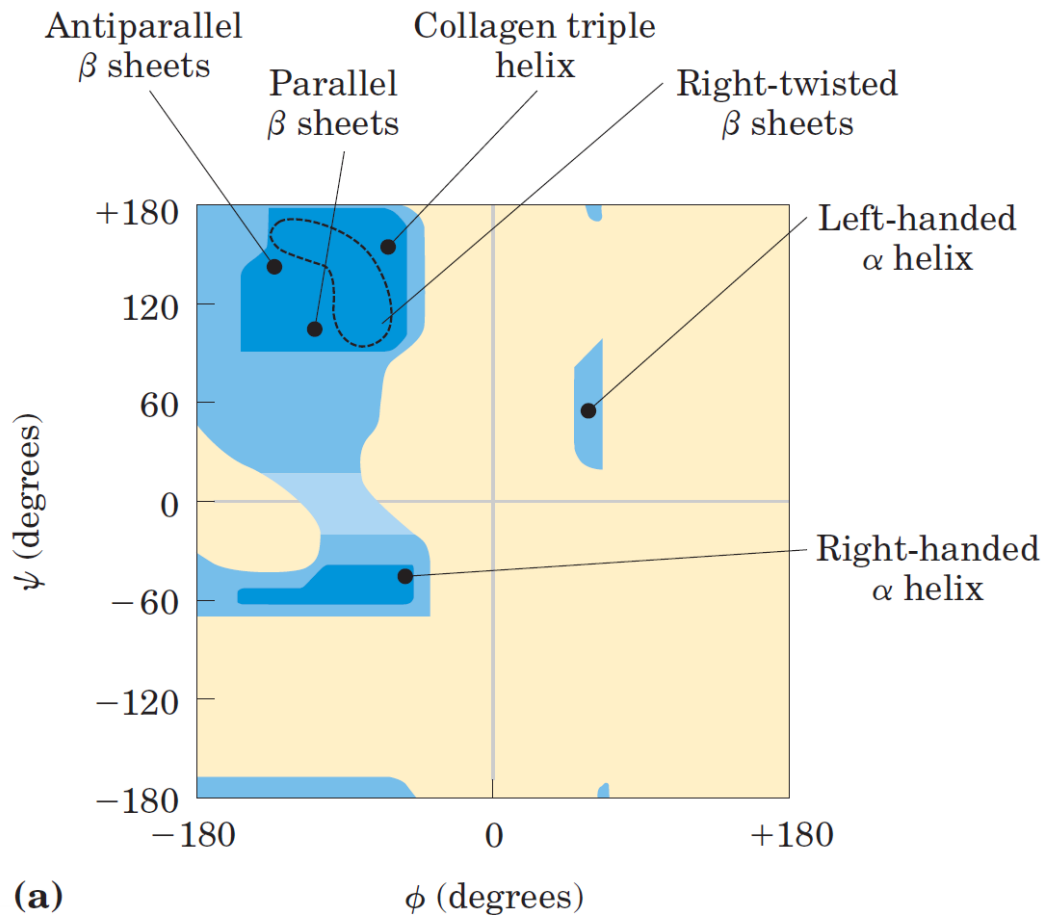


Figure 12 : Projection structurale dans le diagramme de Ramachandran.

Chapitre I : Composition et structures des protéines

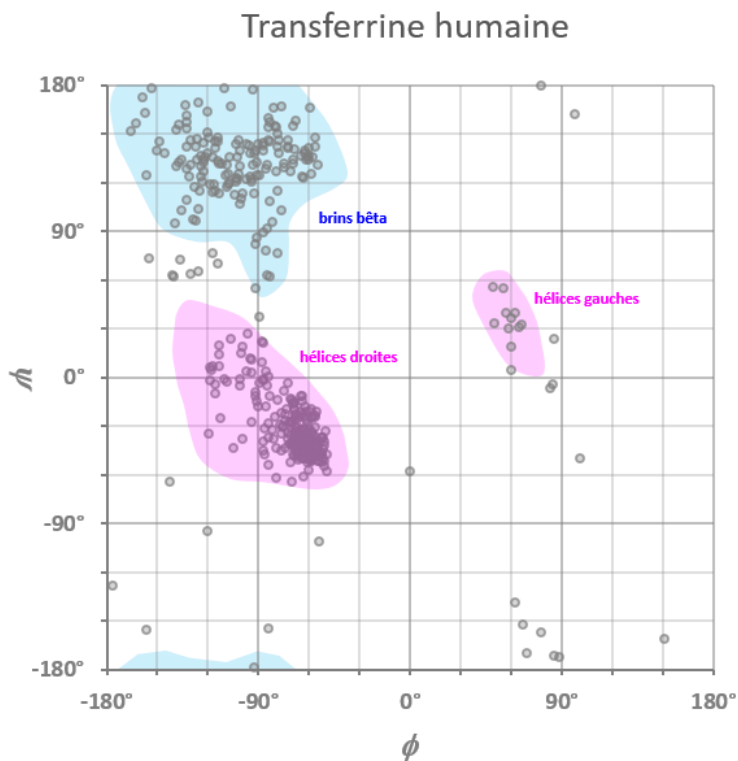
Pour la plupart des résidus, les couples de valeurs stériquement permis sont groupés essentiellement dans deux régions (**Tableau 03 et 04**).

Tableau 03 : Valeurs de ϕ et de ψ dans les conformations structurales.

Conformation	Angles (degrés)			Résidus/tour	Translation/résidu (Å)
	Φ	Ψ	ω		
Hélice α droite	-57	-47	180	3,6	1,50
Hélice α gauche	+57	+47		3,6	1,50
Hélice 3_{10} droite	-49	-26	180	3,0	2,00
Hélice π droite	-57	-70	180	4,4	1,15
Hélice gauche du collagène	-51	+153		3,0	3,13
Feuillet plissé β					
antiparallèle	-139	+135	-178	2,0	3,40
parallèle	-119	+113	180	2,0	3,20
Chaîne étirée	180	180			

Chapitre I : Composition et structures des protéines

Tableau 04 : Exemple d'une détermination de la structure d'une protéine, la transferrine.



III Structure tridimensionnelles des protéines

Les propriétés d'une protéine dépendent essentiellement de sa structure tridimensionnelle. La structure secondaire d'un polymère correspond à la conformation locale de son squelette, qui conduit à définir les motifs issus du repliement régulier de son squelette polypeptidique ; hélices, feuillets plissés et coudes.

Les conformations du squelette peptidique peuvent être décrites par leurs angles de torsion. La conformation du squelette polypeptidique est déterminée par les valeurs des angles de torsion ou angles de rotation (**Figure 13**).

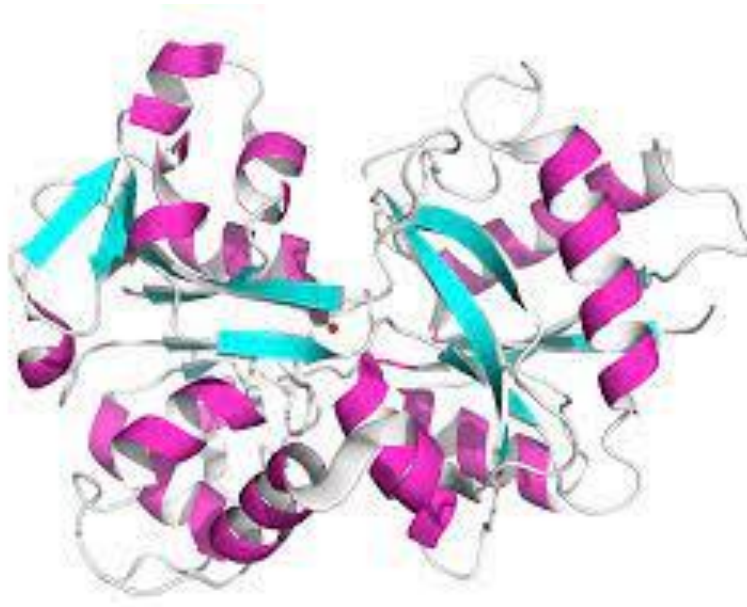


Figure 13 : Structure tridimensionnelles des protéines.

III.1 Structures en hélice

Les hélices sont les éléments les plus frappants de la structure secondaire des protéines.

L'hélice est stabilisée par des liaisons hydrogène entre les groupes NH et CO de la chaîne principale. En particulier, le groupe CO de chaque acide aminé contracte une liaison hydrogène avec le groupe NH de l'acide aminé situé quatre résidus plus en avant dans la séquence. 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 α , en particulier au sein des régions hydrophobes des protéines (**Figure 14**).

Chapitre I : Composition et structures des protéines

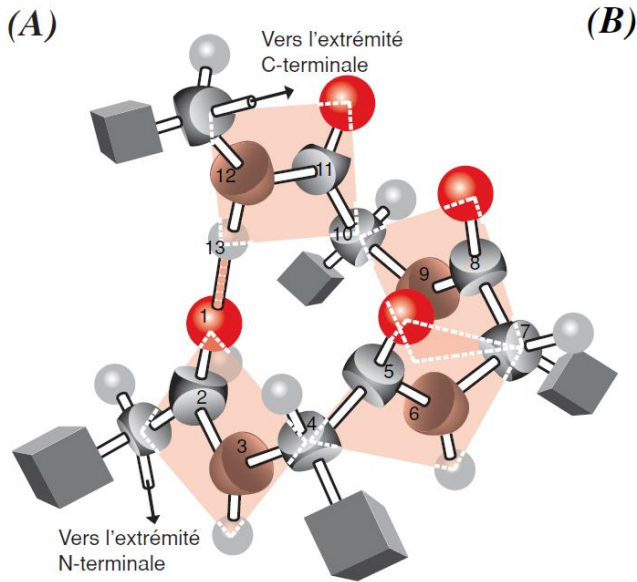


Figure 14 : Les liaisons dans les structures en hélices.

L'orientation des groupes peptidiques étant la même tout au long de l'hélice α , leurs moments dipolaires se cumulent et une hélice α de n résidus a un moment dipolaire, 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. Beaucoup d'hélices α 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 α peut varier considérablement, de quatre ou cinq résidus aminoacide à plus d'une quarantaine, avec une moyenne de 10. Dans la représentation schématique d'une protéine, on donne aux hélices la forme d'un ruban enroulé en spirale (**Figure 15**).

Chapitre I : Composition et structures des protéines

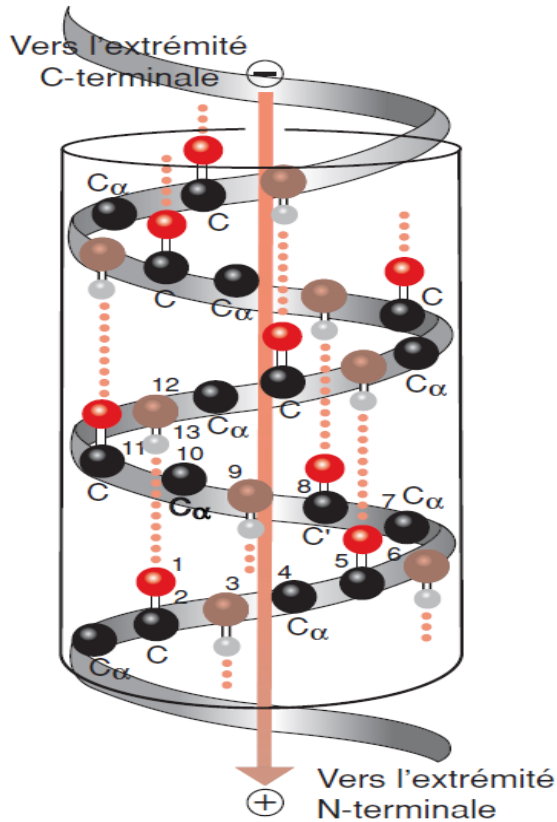


Figure 15 : Représentation schématique d'une hélice.

Les aminoacides distants de trois ou quatre résidus dans la séquence linéaire sont spatialement très proches l'un de l'autre dans une hélice, en revanche, les aminoacides distants de deux résidus dans la séquence linéaire sont situés sur des côtés opposés de l'hélice et ainsi ne peuvent pas entrer en contact. Le pas de l'hélice est la longueur parcourue le long de l'axe de l'hélice après un tour complet. Le sens d'enroulement d'une hélice peut être droit (dans le sens des aiguilles d'une montre) ou gauche (sens inverse). Presque toutes les hélices rencontrées dans les protéines sont droites. La proportion des hélices dans les protéines varie largement, de 0 à 100 %, par exemple 75% des résidus de la ferritine sont engagés dans des hélices.

III.2 Les structures β

Au contraire de l'hélice α , le Feuillet beta plissé comporte des liaisons hydrogène entre des groupements du squelette provenant de résidus distants les uns des autres dans la séquence linéaire. Dans les feuillets beta, deux chaînes ou plus qui peuvent être très éloignées dans la séquence protéique, sont placées côte à côte avec des liaisons hydrogènes entre les brins en sens

Chapitre I : Composition et structures des protéines

inverse (Feuillet antiparallèle) ou orientées dans le même sens (Feuillet parallèle), Il existe également des structures mixtes comportant à la fois des chaînes parallèles et antiparallèles (Figure 16).

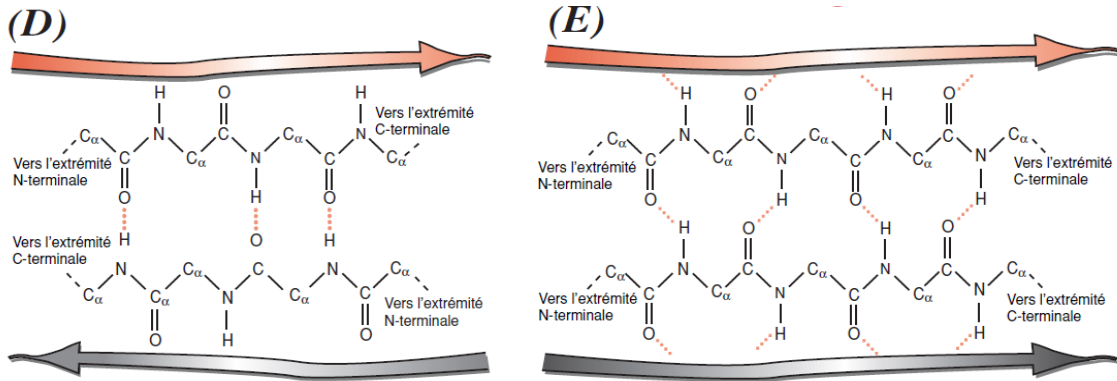


Figure 16 : Les liaisons dans les structures en feuillet.

Dans l'arrangement antiparallèle les groupes NH et CO de chaque aminoacide sont respectivement unis par des liaisons hydrogène aux groupes CO et NH d'un partenaire de la chaîne adjacente. Dans l'arrangement parallèle, la disposition des liaisons hydrogène est légèrement plus compliquée. Pour chaque aminoacide, le groupe NH est lié au groupe CO d'un aminoacide du brin adjacente, tandis que le groupe CO est uni par liaison au groupe NH de l'acide situé deux résidus plus loin dans la chaîne.

III.3 Structure en coude et en boucle

Les hélices et des feuillets plissés sont connectés par des régions de conformation non répétitive et souvent irrégulière, de longueur variable, les coudes ou les boucles. Dans de nombreux coudes, le groupe CO d'un résidu i est uni par liaison hydrogène au groupe NH du résidu $i+3$. Cette interaction stabilise les changements de direction de la chaîne polypeptidique. Les boucles sont des changements de direction plus élaborées.

Chapitre I : Composition et structures des protéines

IV Structure tertiaire

IV.1 Définition

La structure tertiaire d'une protéine correspond à la conformation qu'adoptent les éléments de structure secondaire dans l'espace. Également nommée structure tridimensionnelle ou 3D, elle dépend non seulement de la séquence en acides aminés, mais également des conditions environnementales (température, solvant, pH).

La structure tertiaire est une caractéristique plutôt associée aux protéines globulaires au niveau desquelles, des segments de chaîne polypeptidique se replient et créent des liaisons faibles (forces intermoléculaires) et des ponts disulfure entre des résidus cystéine. Dans cette structure, les chaînes latérales polaires interagissent entre elles ou avec le solvant alors que les chaînes latérales hydrophobes ont tendance à se retrouver à l'intérieur de la structure (**Figure 17**).

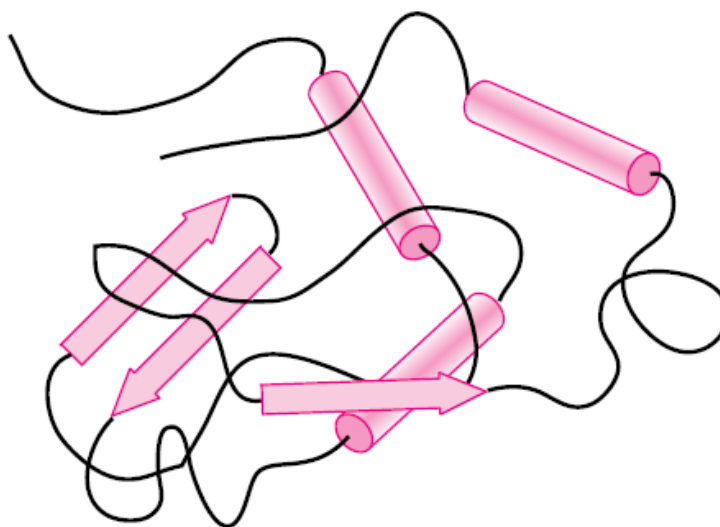


Figure 17 : Structure tertiaire d'une protéine.

IV.2 Structure tertiaire et conformation

Dans cette structure, les chaînes latérales polaires interagissent entre elles ou avec le solvant alors que les chaînes latérales hydrophobes ont tendance à se retrouver à l'intérieur de la structure.

Chapitre I : Composition et structures des protéines

En milieu aqueux les résidus d'acides aminés hydrophobes sont pour la plupart d'entre eux localisés vers l'intérieur de la molécule tandis que les acides aminés polaires sont plus représentés en surface. Des distributions de charges ou de groupements polaires ou hydrophobes peuvent contribuer à conférer à la protéine une fonction donnée.

Un grand nombre de protéines se complexifient jusqu'à la *structure tertiaire*, une structure très spécifique formée à partir de la structure secondaire. Dans une structure tertiaire, des régions hélicoïdales ou plissées de la chaîne polypeptidique se replient les unes sur les autres et forment une molécule en forme de boule, ou molécule globulaire. La structure tertiaire est maintenue par des liaisons (covalentes, hydrogène ...) entre des acides aminés souvent très éloignés sur la chaîne primaire.

La structure tertiaire d'une protéine est son agencement tridimensionnel, c'est à dire le repliement de ses éléments à structure secondaire, ainsi que les dispositions spatiales de ses chaînes latérales. Les principaux types de structures secondaires, les hélices et les feuillets plissés, se trouvent fréquemment dans les protéines globulaires mais en proportions et combinaisons variables. 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.

Chapitre I : Composition et structures des protéines

Les chaînes latérales des protéines globulaires sont réparties dans l'espace selon leurs polarités :

1. Les résidus non polaires (Val, Leu, Ile, Met, Phe) se trouvent essentiellement à l'intérieur de la protéine, à l'abri du solvant aqueux. Les interactions hydrophobes à l'origine de cette distribution, sont en grande partie responsables des structures tridimensionnelles.
2. Les résidus polaires chargés (Arg, His, Lys, Asp, Glu) sont essentiellement situés à la surface de la protéine au contact du solvant aqueux. Ces groupements ont souvent une fonction chimique spécifique ; assurer la catalyse ou participer à la liaison d'un ion métallique.
3. Les groupements polaires non chargés (Ser, Thr, Asn, Gln, Tyr, Trp) sont aussi habituellement à la surface de la protéine mais peuvent aussi se trouver à l'intérieur de la molécule.

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. *La chaîne principale est constituée par les enchaînements $-C\alpha-CO-NH-C\alpha-$ 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 $-N-H-O-$

Chapitre I : Composition et structures des protéines

V Domaines structuraux

De nombreuses protéines sont constituées de domaines qui représentent des unités structurales et fonctionnelles fondamentales. *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 et 200 aa. 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 (**Figure 18**).

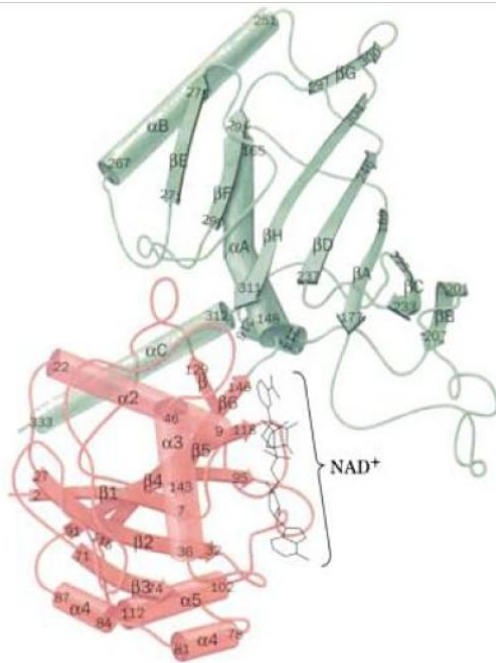


Figure 18 : Une sous-unité de glycéraldéhyde -3- phosphate déshydrogénase.

Lorsqu'une protéine globulaire possède plus de 200 résidus, elle se replie en domaines (deux ou plus) qui vont avoir une fonction spécifique pour l'activité de la protéine (domaine de liaison à un substrat, domaine catalytique, etc.). Des protéines qui partagent des domaines fonctionnels équivalents font partie d'une même famille protéique. Ainsi, des domaines de structure semblable sont souvent retrouvés dans des protéines différentes, avec des fonctions différentes et des séquences d'acides aminés différentes (**Figure 19**).

Chapitre I : Composition et structures des protéines

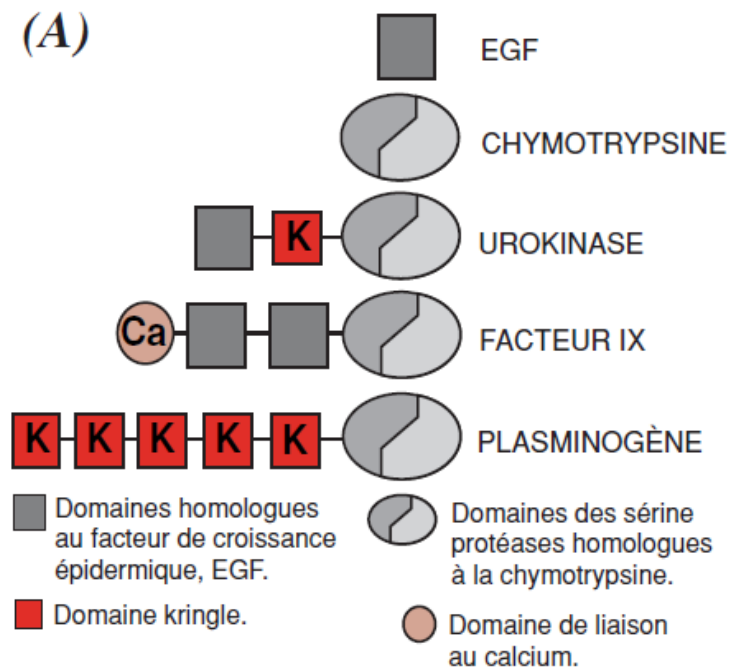


Figure 19 : Domaines structuraux d'une protéine.

Les motifs structuraux retrouvés dans les domaines (**Figure 20**);

Les motifs $\beta\alpha\beta$ les plus courant (feuillet parallèle)(a)

Le motif en épingle à cheveux β (feuillet β antiparallèle)(b)

Le motif $\alpha\alpha$ (antiparallèle) (c)

Le motif en clé grecque (feuillet β antiparallèle)(d)

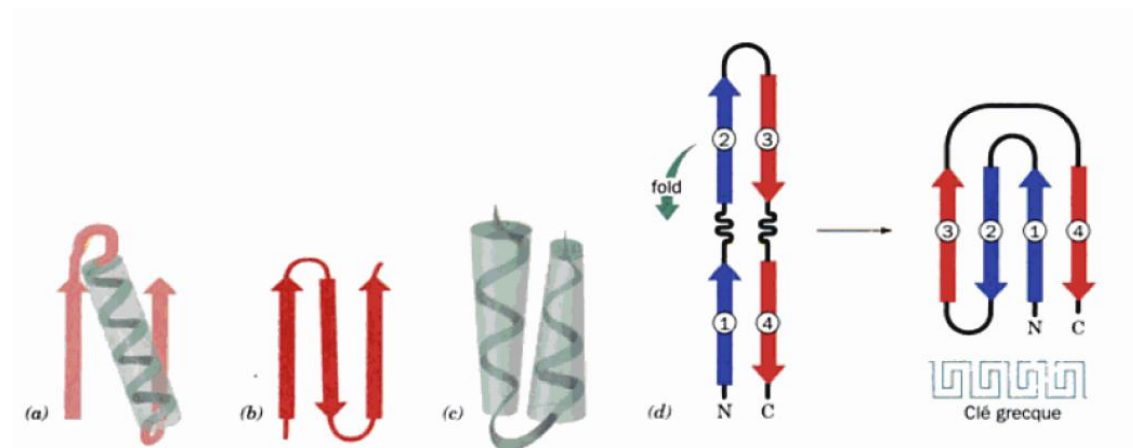


Figure 20 : Motifs structuraux des protéines.

Chapitre I : Composition et structures des protéines

Sur la base de leur structure, on classe les domaines en 4 groupes (**Tableau 05**);

Domaine	Description	Protéine
Domaines α	Structures constituées exclusivement d'hélices α organisées en faisceau.	Hormone de croissance
Domaines β	Structures constituées de brins β antiparallèles organisés en feuillets qui adoptent souvent un enroulement en tonneau.	Immunoglobuline
Domaines α/β	Dans cette structure, un feuillet β central est flanqué d'hélices α . Ce type de domaine est également appelé tonneau TIM (triose phosphate isomérase).	Triose phosphate isomérase
Domaines à feuillets β ouverts	Structures constituées d'un feuillet β central avec de part et d'autre des hélices α qui relient les brins β successifs	Adénylate kinase de porc

Tableau 05 : Différents domaines structuraux d'une protéine.

LES STRUCTURES DES PROTÉINES PEUVENT ÊTRE RÉPARTIES EN TROIS CLASSES ;

les structures α

les structures β

les structures α et β

Dans les structures α , les domaines sont constitués uniquement d'hélices α , tandis que dans les structures β , ils sont formés de feuillets β antiparallèles.

Les structures α et β résultent de combinaisons de motifs $\beta\alpha\beta$ qui forment un feuillet β parallèle entouré d'hélices α (**Figure 21**).

Chapitre I : Composition et structures des protéines

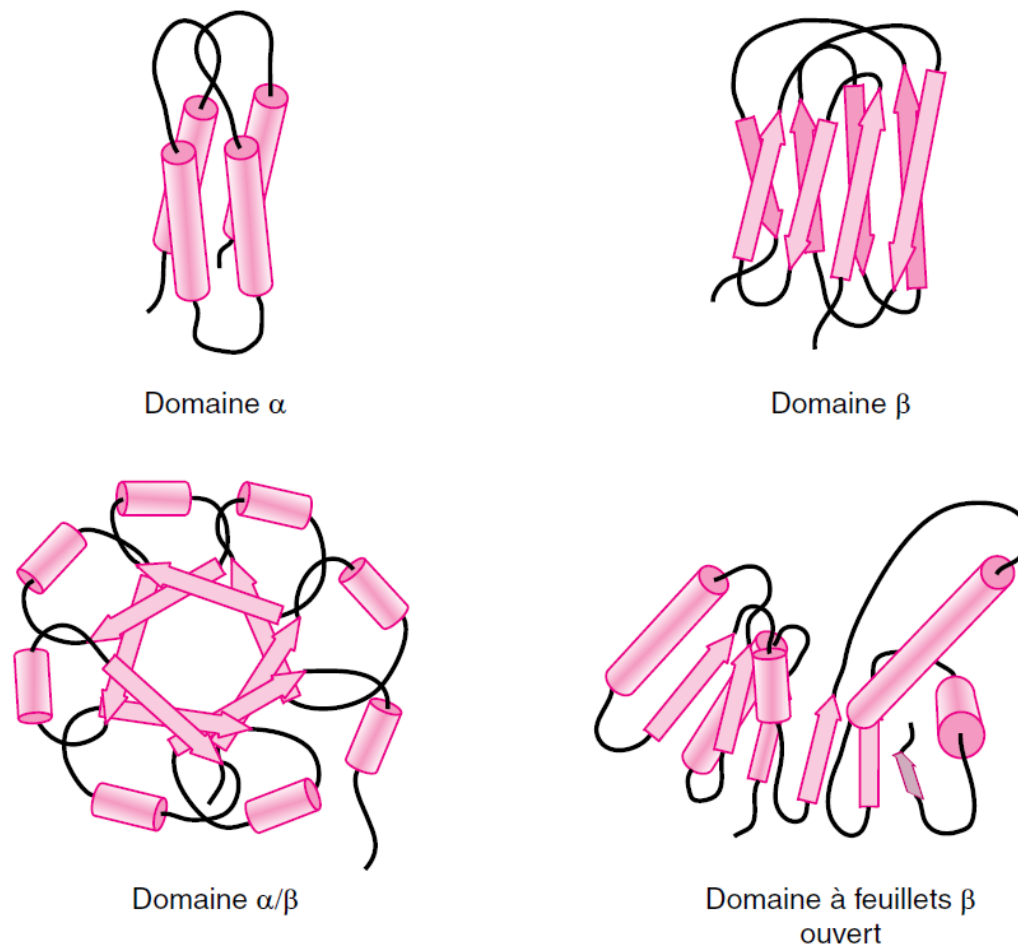


Figure 21 : Différents domaines structuraux d'une protéine.

V.1 Les structures α

Parmi les structures α , le faisceau de quatre hélices, au sein duquel des résidus hydrophobes délimitent un site actif central (les hélices sont inclinées de sorte que leurs chaînes latérales perd le contact avec le milieu aqueux, et s'organisent en topologie haut-haut-bas-bas ou haut-bas-haut-bas), est rencontré dans des protéines ex; les cytochromes c et b562 et la ferritine (**Figure 22**).

Chapitre I : Composition et structures des protéines

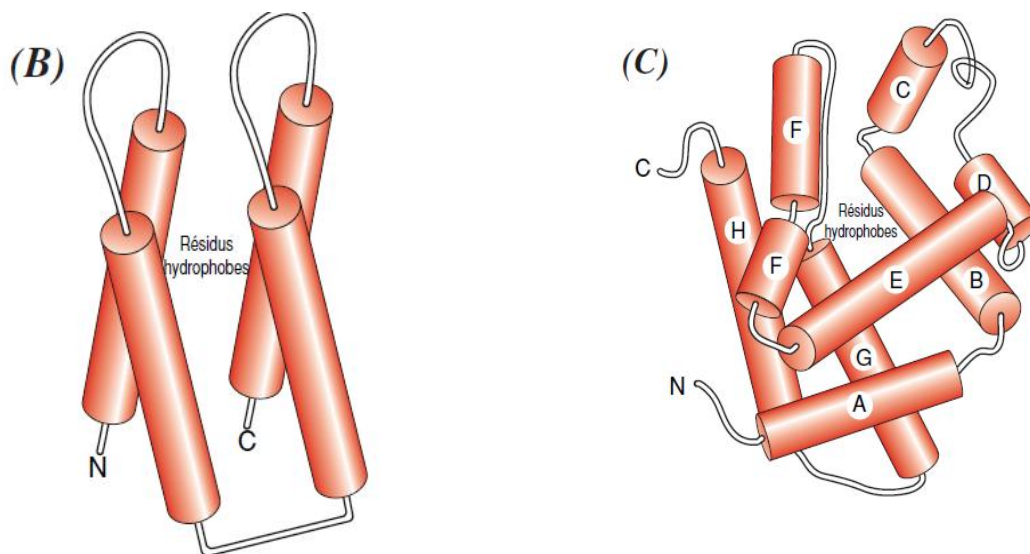


Figure 22 : Structure α .

De même, le repliement globinique, où là aussi des résidus hydrophobes délimitent un site actif central, est présent dans toutes les protéines d'un important groupe incluant les myoglobines et les hémoglobines. 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.

V.2 Les structures β

Les structures β sont constituées de brins β dont le nombre peut varier de quatre à plus de dix, souvent répartis en deux feuillets antiparallèles plaqués l'un contre l'autre pour former un tonneau β affecté d'une certaine torsion (**Figure 23**).

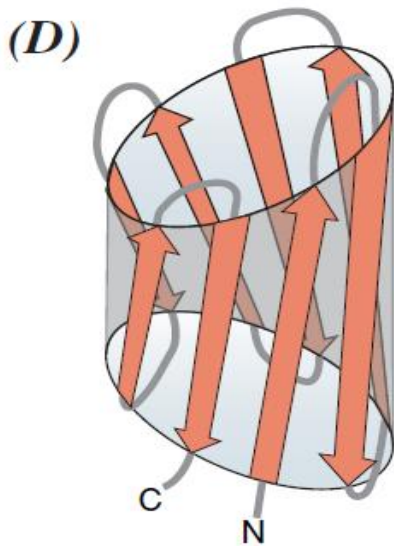


Figure 23 : Structure β .

Le pli immunoglobuline, qui forme la structure du domaine de base de la plupart des protéines du système immunitaire, est constitué d'un feuillet β antiparallèle à 3 brins en contact, face à face avec un feuillet β antiparallèle à 3 brins. Le tonneau B haut et bas fait de 8 brins antiparallèles successifs comme dans la protéine de liaison du rétinol (Figure 24).

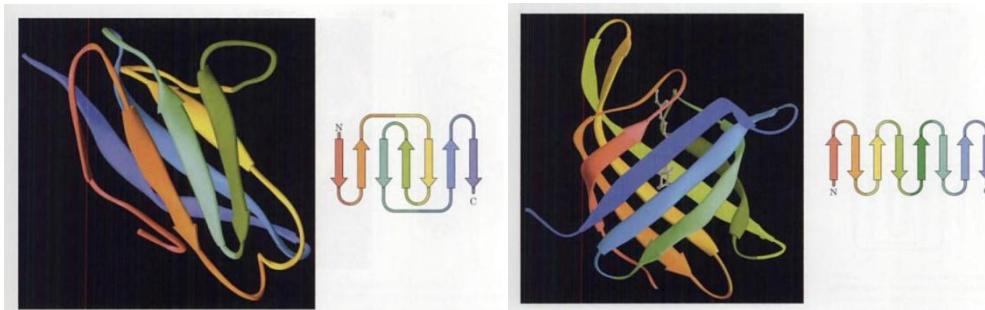


Figure 24 : Structure β (pli immunoglobuline).

Les structures β ont un core hydrophobe constitué par les résidus aminoacide hydrophobes des brins β qui tapissent la partie interne du tonneau. Leur surface est constituée par les résidus aminoacide hydrophiles des brins β ainsi que par ceux des boucles qui unissent ces derniers. Les structures β sont présentes dans des protéines fonctionnellement très différentes, telles que des enzymes, des protéines de transport et des anticorps (Figure 25).

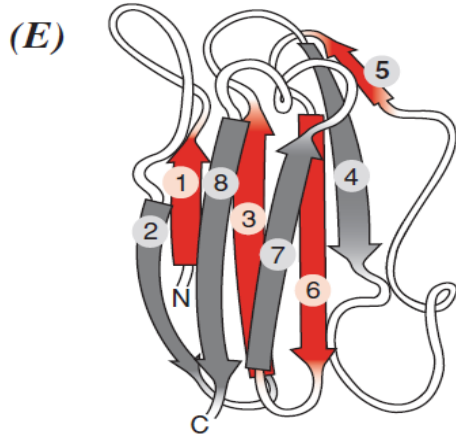


Figure 25 : Structure β (core hydrophobe).

V.3 Les structures α et β

Les structures $\alpha\beta$, très fréquemment rencontrées, sont constituées d'un feuillet β central, parallèle ou mixte, entouré d'hélices α . Elles se répartissent en deux classes. Les structures $\alpha\beta$ de la première classe présentent un core de huit brins β parallèles disposés comme les douves d'un tonneau mais affectés d'une certaine torsion.

8 motifs $\alpha\beta$ superposés en hélice de pas à droite pour former un tonneau β à 8 brins à l'intérieur d'un tonneau de 8 hélices α . tous les brins β sont antiparallèles (Figure 26).

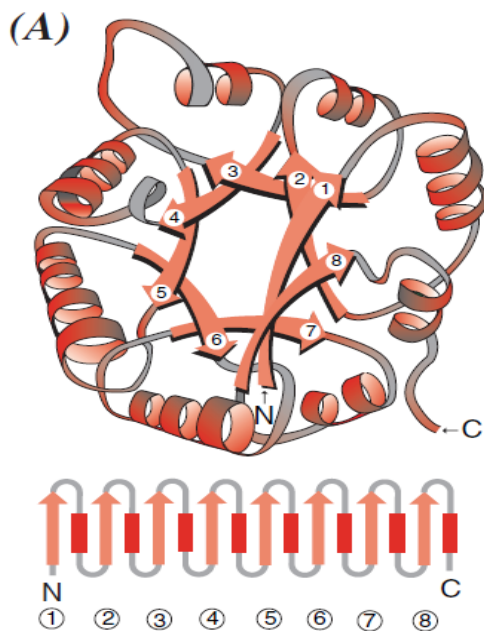


Figure 26 : Structures α et β .

Chapitre I : Composition et structures des protéines

Les résidus aminoacide hydrophobes des brins β se disposent à l'intérieur du tonneau. Les hélices α qui connectent les brins β parallèles se disposent à la périphérie du tonneau. Les structures $\alpha\beta$ de la deuxième classe possèdent un feuillet β ouvert entouré de chaque côté par les hélices α . ces plis sont constitués de feuillet β central, parallèle, relie de part et d'autre par une hélice α (Figure 27).

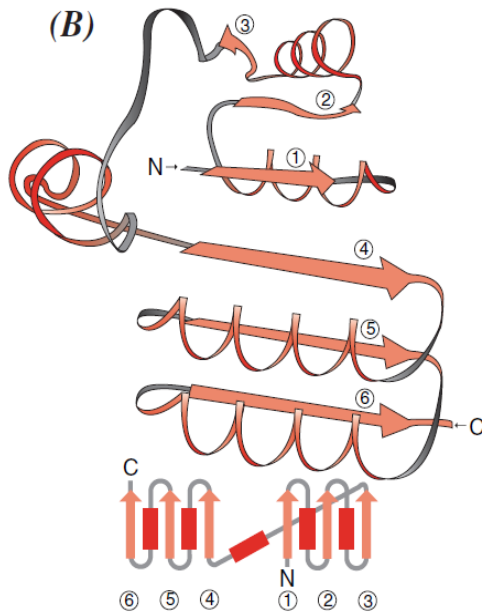


Figure 27 : Structures α et β .

Dans les structures $\alpha\beta$, les boucles forment des cavités qui permettent la fixation de ligands et peuvent éventuellement posséder une activité catalytique. En raison de ces propriétés, nombre de protéines de transport et d'enzymes possèdent une structure $\alpha\beta$; le domaine de fixation des coenzymes nucléotidiques, des enzymes de la glycolyse, tels que la triose phosphate isomérase et la pyruvate kinase, de même que beaucoup d'autres enzymes, en sont des exemples (Figure 28).

Chapitre I : Composition et structures des protéines

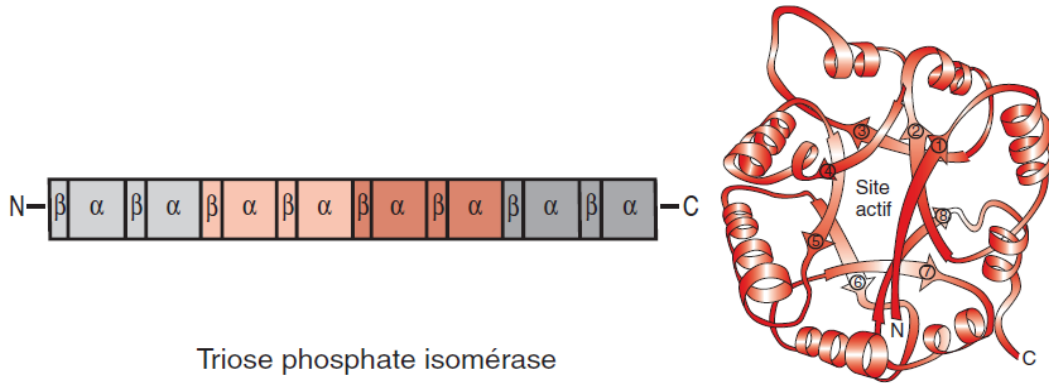


Figure 28 : Structures α et β de la triose phosphate isomérase.

VI Structure quaternaire

La structure quaternaire correspond à l'association non covalente de plusieurs sous-unités (protomères) protéiques identiques ou différents. La structure protéique résultant de l'association est appelée oligomère. Les oligomères constitués de deux, trois, quatre ou plus de sous-unités sont appelés dimères, trimères, tétramères etc. Lorsque les sous-unités d'un oligomère sont identiques, le préfixe homo- est utilisé (ex. homodimère). Le préfixe hétéro- sera utilisé lorsque les sous-unités d'un oligomère sont différentes (ex. hétérotétramère).

L'association entre les sous unités dépend à la fois de la complémentarité de forme et surtout des interactions faibles entre les sous-unités : liaisons électrostatiques, liaisons hydrogène, liaisons de Van der Waals, interactions hydrophobes (**Figure 29**).

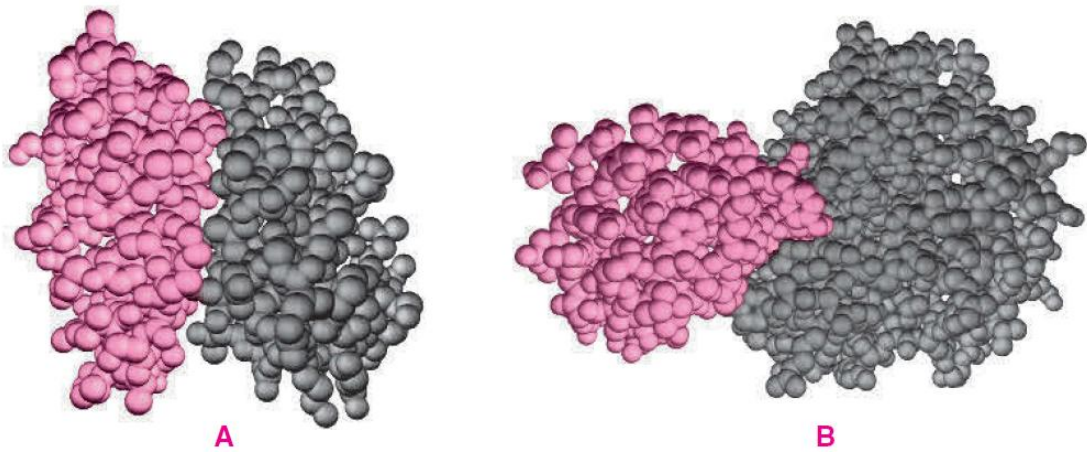


Figure 29 : Illustration de structures quaternaires (A : Homodimère , B : Hétérodimère).

Chapitre I : Composition et structures des protéines

Il s'agit d'une organisation spatiale d'assemblages de molécules protéiques dans une structure tertiaire donnée par des liaisons ionique, hydrogène, hydrophobe ou par pont disulfure. Cet assemblage symétrique ou non est en général constitué par un nombre limité de sous-unités (monomères) identiques ou non. L'entité moléculaire, dont la masse molaire peut dépasser plusieurs millions, possède des propriétés souvent remarquables; chaque association a une structure souvent unique et contribue à la fonction biologique. Les chaînes qui constituent ce complexe sont des protomères ou sous unités. Chacune ayant une structure tertiaire définie.

L'oligomère présente un axe de symétrie. L'activité biologique dépend de cette structure oligomérique et disparaît lorsque la structure quaternaire est dissociée.

L'assemblage des protomères en oligomère fait intervenir des liaisons hydrogènes, hydrophobes, ioniques et de Van der Waals. Une protéine a plusieurs chaînes polypeptidiques identiques ou différentes. Nous appellerons des protéines ayant des sous-unités identiques des oligomères ou protéines oligomériques et ces sous-unités identiques des protomères. Oligomère, peut être constituée d'une chaîne polypeptidique ou plusieurs chaînes polypeptidiques différentes. Ex l'hémoglobine est un dimère ; un oligomères de deux protomères $\alpha_2\beta_2$.

VII Symétrie dans les protéines

Lorsque les sous-unités d'un complexe protéique sont identiques, celles-ci vont occuper des positions géométriques équivalentes au sein de l'oligomère, et donc présenter une symétrie par rapport à un axe (symétrie de rotation). Une symétrie peut également être produite par rapport à des dimères (ou plus) de structure différente. C'est le cas de l'hémoglobine, protéine tétramérique constituée de deux dimères symétriques eux-mêmes issus de l'association de deux sous-unités différentes notées α et β .

Dans pratiquement toutes les protéines oligomériques, les protomères sont disposés symétriquement. Les protomères se réunissent autour d'un seul point pour former une entité fermée.

VII.1 Symétrie cyclique

Dans le type de symétrie de rotation le plus simple, la symétrie cyclique, les sous-unités sont reliées (amenées à coïncider) par un seul axe de rotation.

Chapitre I : Composition et structures des protéines

Dans la symétrie cyclique (C_n), les sous-unités sont symétriques par rapport à un seul axe de rotation. La symétrie est notée C_2 , C_3 , ..., C_n (symétries d'ordre 2, 3, ..., n), le nombre en indice représentant le nombre de sous-unités constitutives de l'oligomère (**Figure 30**).

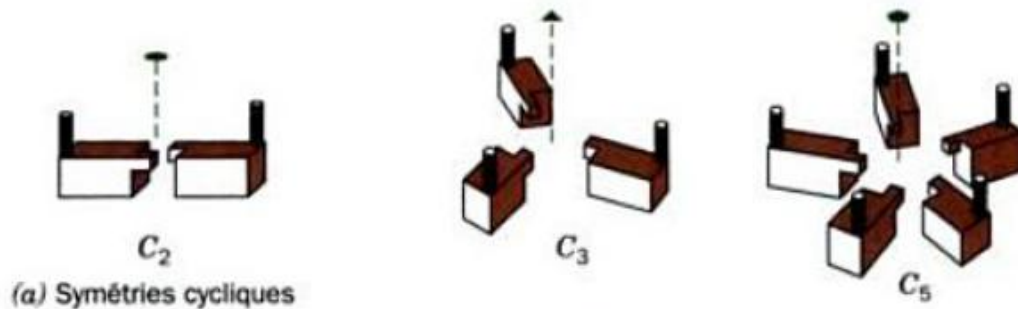


Figure 30 : Symétrie cyclique.

La symétrie C_2 est la symétrie la plus courante dans les protéines; des symétries cycliques supérieures sont relativement rares.

VII.2 Symétrie dièdre

Dans la symétrie diédrique (D_n), un type de symétrie de rotation plus compliquée, l'axe de rotation d'ordre 2 est coupé par un axe de rotation d'ordre n . Les oligomères ont une symétrie notée D_2 , D_3 , ..., D_n , le nombre en indice représentant la moitié du nombre de protomères constitutifs de la structure (**Figure 31**).

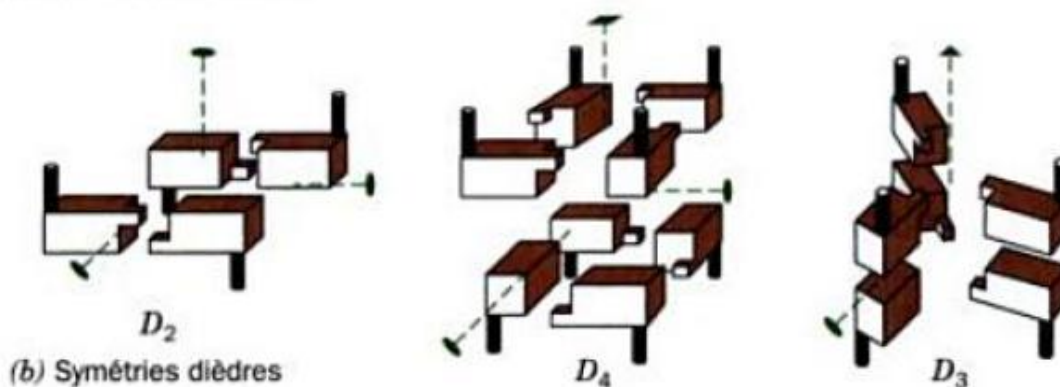


Figure 31 : Symétrie dièdre.

Chapitre I : Composition et structures des protéines

La symétrie D_2 est la symétrie dièdre la plus courante dans les protéines (Figure 32).

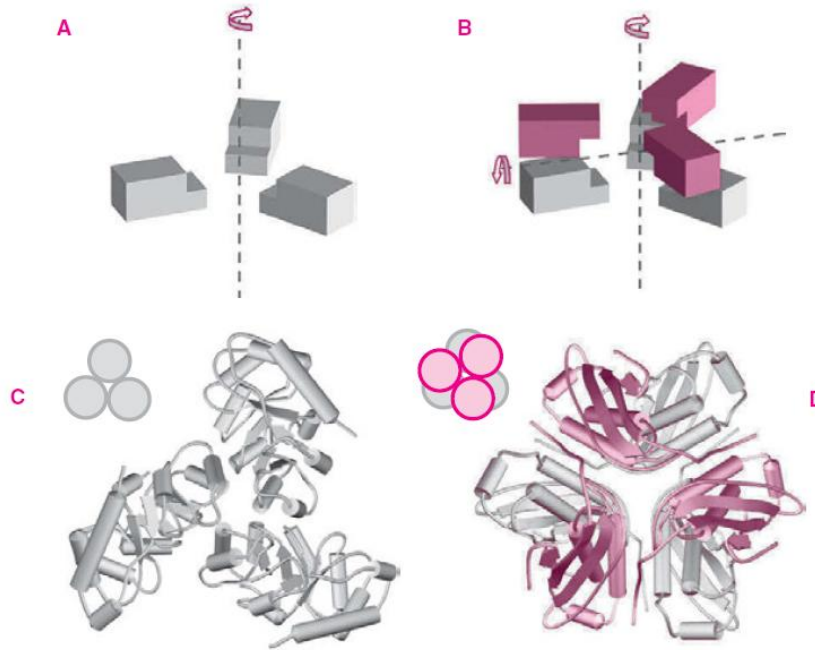


Illustration de deux symétries de rotation.

A : Symétrie cyclique C_3 ; B : Symétrie diédrique D_3 ; C : KDGP Aldolase (2-Céto-3-désoxygluconate-6-phosphate aldolase), protéine à symétrie cyclique C_3 ; D : Protéine C de biosynthèse du cofacteur molybdène, protéine à symétrie diédrique D_3 . Les axes de rotation sont représentés en pointillés.

Figure 32 : Symétrie cyclique et dièdre.

VII.3 Symétrie de rotation

Il existe également des oligomères présentant des symétries particulières faisant référence à des polyèdres réguliers (symétrie de rotation d'un tétraèdre (4 sous-unités), d'un octaèdre (8 sous-unités), d'un icosaèdre (20 sous-unités)) (Figure 33).

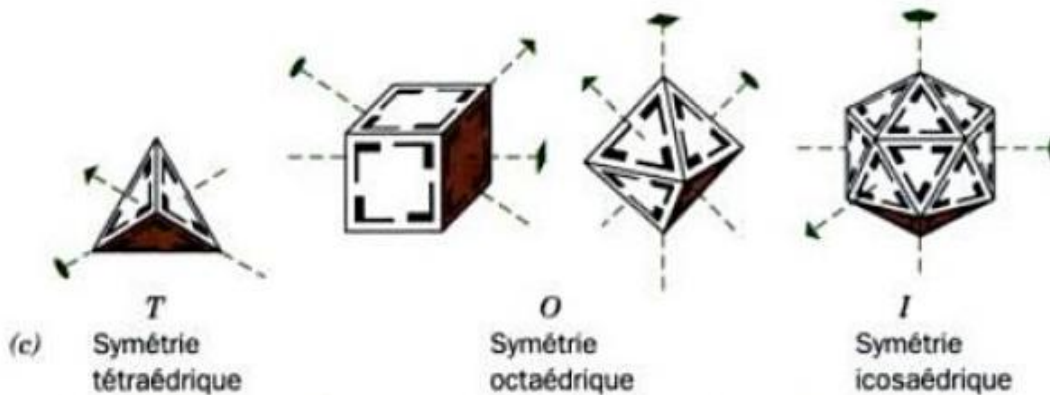


Figure 33 : Symétrie de rotation.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

I Repliement des protéines

Les protéines se forment à partir de leurs différents constituants. Nous devons savoir quelles sortes de mouvements internes la protéine peut et doit subir afin d'assurer sa fonction biologique et aussi comment elle parvient à sa structure native. Bien que les mécanismes catalytiques de nombreuses enzymes de structure connue aient été étudiés en détail, on ne peut dire qu'on comprenne ces mécanismes avec précision. Nous n'avons qu'une connaissance incomplète des modes d'interaction des groupements qui constituent les protéines, et en sommes donc souvent réduits à des hypothèses.

Les protéines se replient spontanément pour prendre leur conformation native dans des conditions physiologiques. Ceci implique que la structure primaire d'une protéine dicte sa structure tridimensionnelle. Les protéines sont douées d'auto-assemblage.

I.1 Renaturation des protéines

En 1957, les expériences de *Christian Anfinsen* sur l'ARNase A de pancréas bovin ont abordé l'étude de la renaturation des protéines sur une base quantitative. L'ARNase A, protéine à une seule chaîne de 124 résidus, se déroule entièrement et ses quatre ponts disulfure sont rompus par réduction dans une solution d'urée 8M contenant du 2-mercaptoéthanol. Cependant, après élimination de l'urée par dialyse et mettant une solution obtenue en présence d'O₂ à pH 8, on obtient une protéine qui est pratiquement à 100% enzymatiquement active à l'ARNase A native. La protéine s'est donc renaturée spontanément. La renaturation de l'ARNase A nécessite que ses quatre ponts disulfure se reforment dans des conditions de renaturation (**Figure 34**).

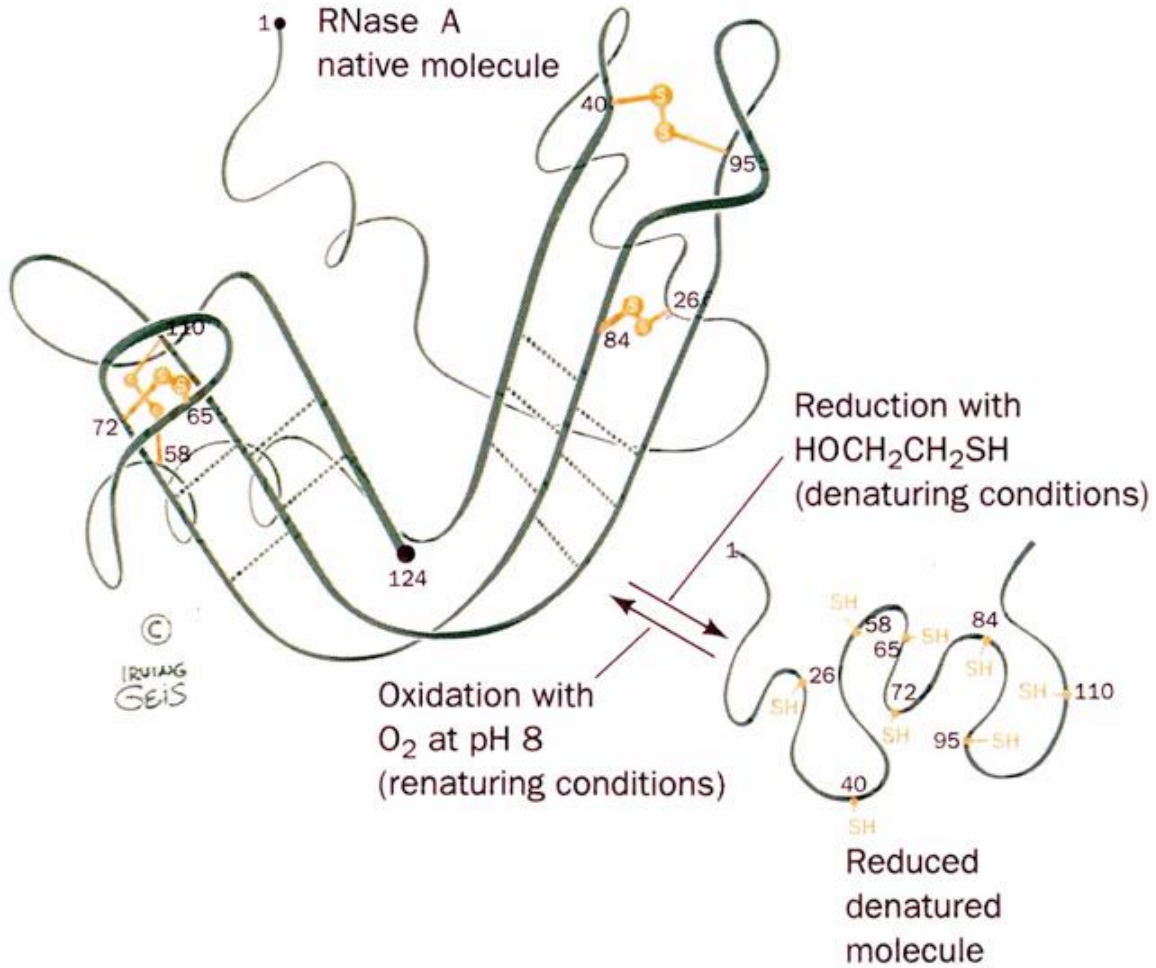


Figure 34 : Structure de l'ARNase.

Si l'on réoxyde l'ARNase A en présence d'urée 8M de sorte que les ponts disulfure se reforment alors que la chaîne polypeptidique est déroulée, on trouve après élimination de l'urée que l'ARNase A n'a qu'environ 1% de son activité enzymatique. Cette ARNase A brouillée peut devenir totalement active si on ajoute des traces de 2-mercaptoéthanol qui, durant une période de 10 heures, catalysent des réactions de réarrangement des ponts disulfure jusqu'à ce que la structure native soit retrouvée.

L'état natif de l'ARNase A dans des conditions physiologiques correspond donc vraisemblablement à sa conformation thermodynamiquement la plus stable. Le temps nécessaire à la renaturation de l'ARNase A brouillée est réduit à environ 2 min par la protéine disulfure isomérase (PDI), enzyme qui catalyse les réactions de réarrangement des ponts

Chapitre II : Repliement des protéines, dynamique et évolution structurale

disulfure. Le site actif de la PDI contient deux résidus Cys, qui doivent être sous la forme-SH pour que l'isomérase soit active (**Figure 35**).

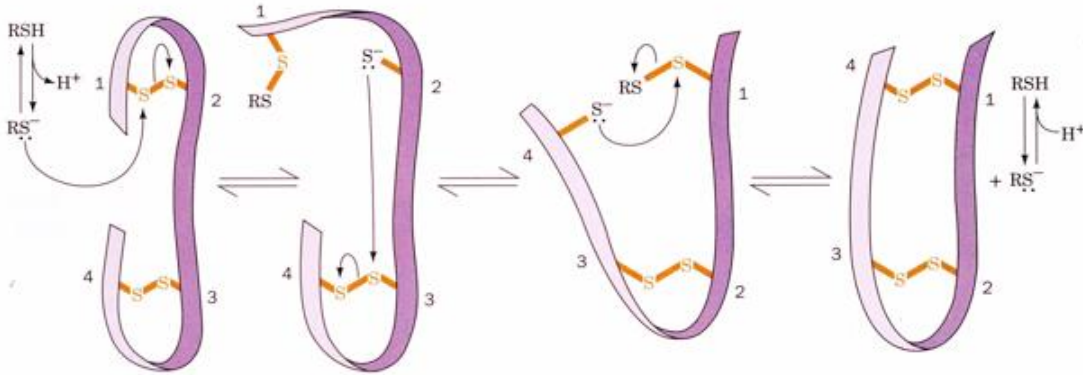


Figure 35 : Site actif de la PDI.

Beaucoup de protéines "brouillées" peuvent être renaturées par la PDI car elle n'affecte pas leur état natif. Cependant, pour des protéines modifiées après leur traduction, les ponts disulfures maintiennent la protéine sous une conformation native qui autrement serait instable.

L'insuline, hormone polypeptidique de 51 résidus formée de deux chaînes reliées par deux ponts disulfure, est inactivée par la PDI. Cette observation a conduit à la découverte que l'insuline est issue de la proinsuline, un polypeptide de 84 résidus. Ce n'est qu'après la formation de ses ponts disulfure que la proinsuline est transformée en hormone active suite à l'excision protéolytique d'un segment interne de 33 résidus appelé chaîne C ou peptide C.

Sa chaîne C est excisée par protéolyse entre ses chaînes A et B pour donner l'hormone mature. Le peptide C ne dirige pas le repliement des chaînes A et B, mais plutôt il les maintient ensemble pendant qu'elles forment leurs pont disulfure natif (**Figure 36**).

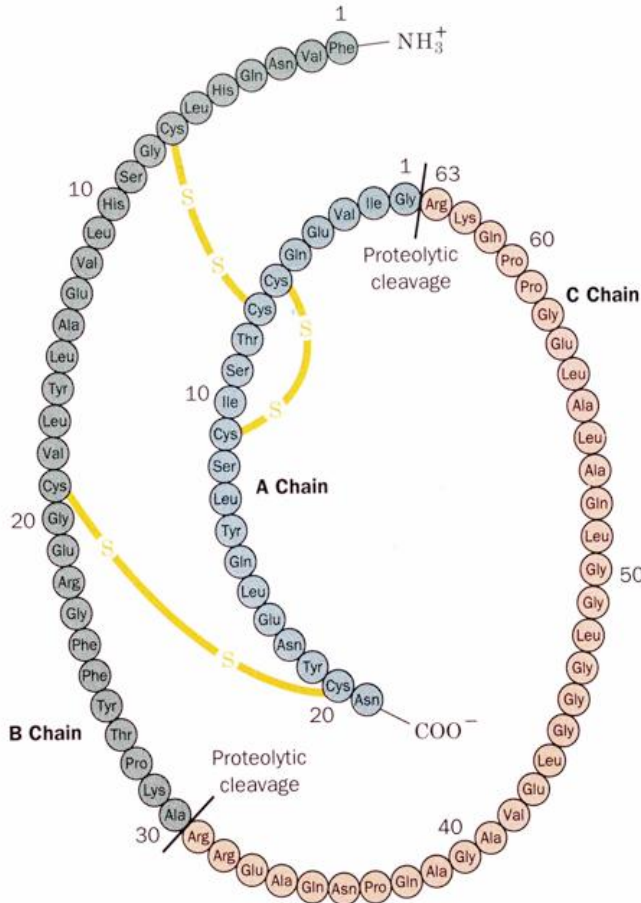


Figure 36 : Structure de l'insuline.

I.2 Déterminants du repliement des protéines

a. Présence des hélices et des feuillets

Les interactions hydrophobes sont responsables de la nature compacte et non polaire du cœur des protéines. La formation des hélices et des feuillets est la conséquence de contraintes stériques au sein des protéines compactes. Ces éléments de

Structure secondaire s'organisent via des liaisons hydrogène, des interactions ioniques et des forces de van der Waals (la plupart des liaisons hydrogène dans les protéines s'établissent entre des résidus proches dans la séquence).

Si les hélices et les feuillets prédominent dans les protéines, c'est sans doute

Simplement parce qu'ils occupent bien l'espace.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

b. Le repliement des protéines se fait essentiellement sous l'influence de résidus internes

Dans une étude les groupements amines primaires libres de l'ARNase A (les résidus Lys et le groupement N-terminal) ont été couplés, sans entraîner une modification de la conformation native de la protéine, ni sa faculté à se replier. Ces groupements amines libres situés à la surface de l'ARNase. Le repliement des protéines est sous la dépendance de forces hydrophobes et ce sont les résidus internes d'une protéine qui déterminent son repliement. Le repliement des protéines se fait essentiellement sous l'influence de résidus internes.

c. La structure des protéines s'établit selon une hiérarchie

Les grosses sous-unités des protéines sont constituées de domaines, qui sont des segments compacts, mais physiquement séparables de la chaîne polypeptidique. Les domaines sont à leur tour composés de sous-domaines et ainsi de suite.

L'organisation des protéines est hiérarchique, les chaînes polypeptidiques forment localement des structures compactes qui s'associent avec des structures voisines dans la séquence, pour former des structures compactes plus vastes (**Figure 37**).

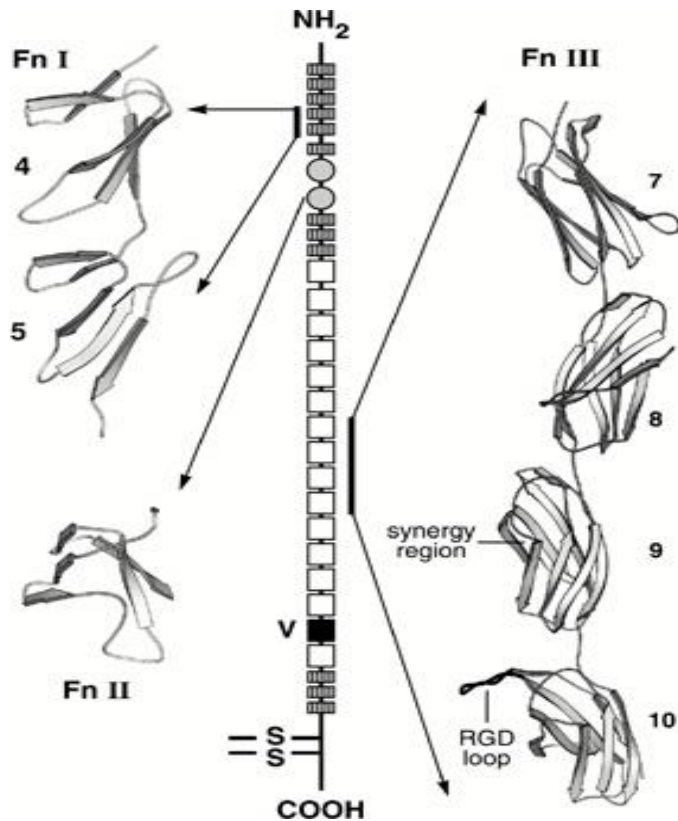


Figure 37 : Organisation des domaines d'une protéine.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

d. Les structures des protéines sont adaptables

Les protéines globulaires ont des densités de compactage car les chaînes latérales à l'intérieur de la protéine s'assemblent avec une complémentarité. Le compactage dépend du repliement natif, on retrouve des protéines qui ont des repliements identiques et qui n'ont pas de similitudes de séquences.

I.3 Mécanismes de repliement

On pourrait imaginer qu'une protéine explore au hasard toutes les conformations possibles jusqu'à ce qu'elle "tombe" sur sa conformation native.

Cyrus Levinthal a calculé qu'elle peut trouver environ 10¹³ conformations par seconde et pour une petite protéine de 100 résidus, chacun ayant 3 conformations stables, le temps nécessaire pour que la protéine explore toutes les conformations possibles est de l'ordre de 10⁸⁷ ans. En réalité, il faut moins que quelques secondes.

Levinthal a proposé que le repliement d'une protéine doit donc faire intervenir une série de mécanismes séquentiels au cours desquels l'évolution vers l'état natif s'accompagne d'une augmentation très importante de la stabilité conformationnelle (**diminution d'énergie Libre**).

Paradoxe de Levinthal (1968):

- chaque résidu n'a que deux possibilités de conformation.
- une protéine de 100 résidus aurait $2^{100} \approx 10^{30}$ conformations possibles.

a. Il faut des mesures rapides pour suivre le repliement d'une protéine

L'étude du repliement de plusieurs petites protéines à domaine unique comme l'ARNase, le cytochrome c, montre que ces protéines se replient dans la milliseconde, qui suit leur transfert de conditions dénaturantes en conditions natives.

On peut suivre les premières phases de repliement avec un mélangeur rapide appelé appareil à **flux interrompu** ("stopped flow«). Ici une solution dans laquelle la protéine est dénaturée par le pH, le chlorure de guanidinium ou l'urée, est rapidement modifiée pour déclencher le repliement de la protéine (**Figure 38**).

Chapitre II : Repliement des protéines, dynamique et évolution structurale

Il faut pouvoir suivre le repliement de la protéine par une technique qui détecte des changements rapides dans la structure de la protéine comme la spectroscopie par **dichroïsme circulaire (CD)**.

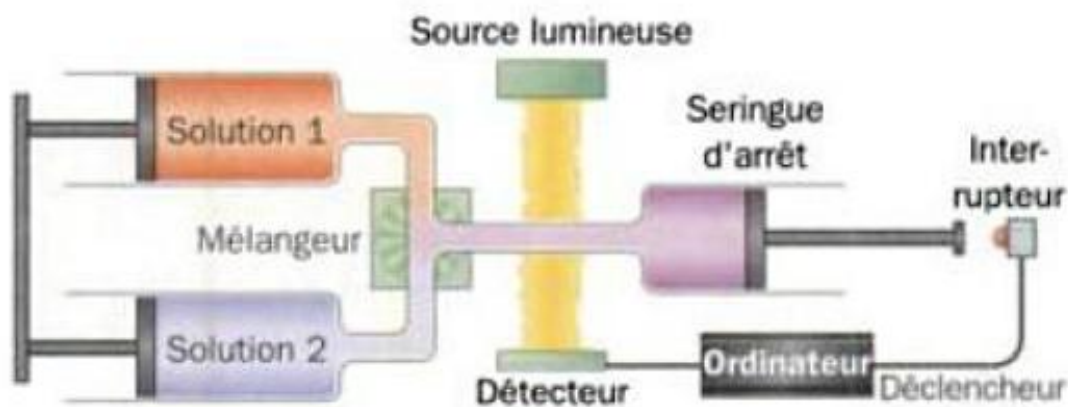


Figure 38 : Flux interrompu ("stopped flow«).

Une solution contenant un soluté qui absorbe la lumière.

Selon la loi de Beer-Lambert:

$$A = \epsilon \cdot c \cdot l$$

où A est l'absorbance, l est la longueur en cm du trajet optique,

c est la concentration et ϵ est le coefficient d'extinction molaire.

Pour des molécules chirales comme les protéines, les valeurs de ϵ sont différentes pour la lumière polarisée en cercle vers la gauche (ϵ_L) ou vers la droite (ϵ_R).

Les polypeptides absorbent fortement la lumière dans la région ultraviolette, essentiellement parce que leurs chaînes latérales aromatiques ont un coefficient d'extinction molaire élevés.

Pour les protéines les valeurs de la lumière polarisée donne le spectre CD. Dans les protéines, les hélices, les feuillets et les enroulements ont des spectres CD caractéristiques.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

b. Les étapes le plus précoces du repliement d'une protéine sont amorcées par un effondrement hydrophobe

Les mesures de CD-stopped flow montrent que pour des petites protéines à domaine unique, la majorité des structures secondaires dans la protéine native se forment dans les quelques millisecondes qui suivent le déclenchement du repliement. Puisque les protéines globulaires ont un coeur hydrophobe compact, il est probable que ce qui déclenche leur repliement soit un "effondrement hydrophobe", au cours duquel les groupements hydrophobes se rassemblent pour expulser la majorité des molécules d'eau. L'état initial "affaissé" ou "comprimé" d'une protéine qui se replie est appelé globule fondu ou « molten globule ». Cependant les chaînes latérales du globule fondu sont très désordonnées et une stabilité thermodynamique minime. Malgré cela la chaîne polypeptidique poursuit son repliement et atteint l'état natif, sans subir de profonds réarrangements. Après la phase rapide d'effondrement hydrophobe et formation du globule fondu, il y a des étapes intermédiaires dans le repliement durant 5 à 1000 ms. C'est à ce stade que la structure secondaire se stabilise et que sa structure tertiaire commence à se former.

c. La théorie du paysage pour le repliement des protéines

- le repliement des protéines passe par une série d'intermédiaires bien définis.
 - le repliement d'une protéine enroulée au hasard commence par la formation aléatoire de courts segments de structure secondaire.
 - des hélices α et des tournants β joueraient le rôle de **centres de nucléation**.
 - ces centres de nucléation se développeraient en structures stables et ces régions ordonnées croissent spontanément de façon coopérative pour former un domaine de la protéine native.
- D'après la **théorie du paysage ("energy landscape")** le repliement se déroule sur une surface énergétique ou paysage, qui représente les états d'énergie conformationnelle auxquels la protéine a accès.

Entonnoirs de repliement (a) - la chaîne polypeptidique établit un nombre croissant de contacts intracaténaux, son énergie libre interne diminue ainsi que sa liberté conformationnelle. Des polypeptides de conformations différentes suivent différentes voies (A, B) pour atteindre l'état natif (N) (**Figure 39**).

Chapitre II : Repliement des protéines, dynamique et évolution structurale

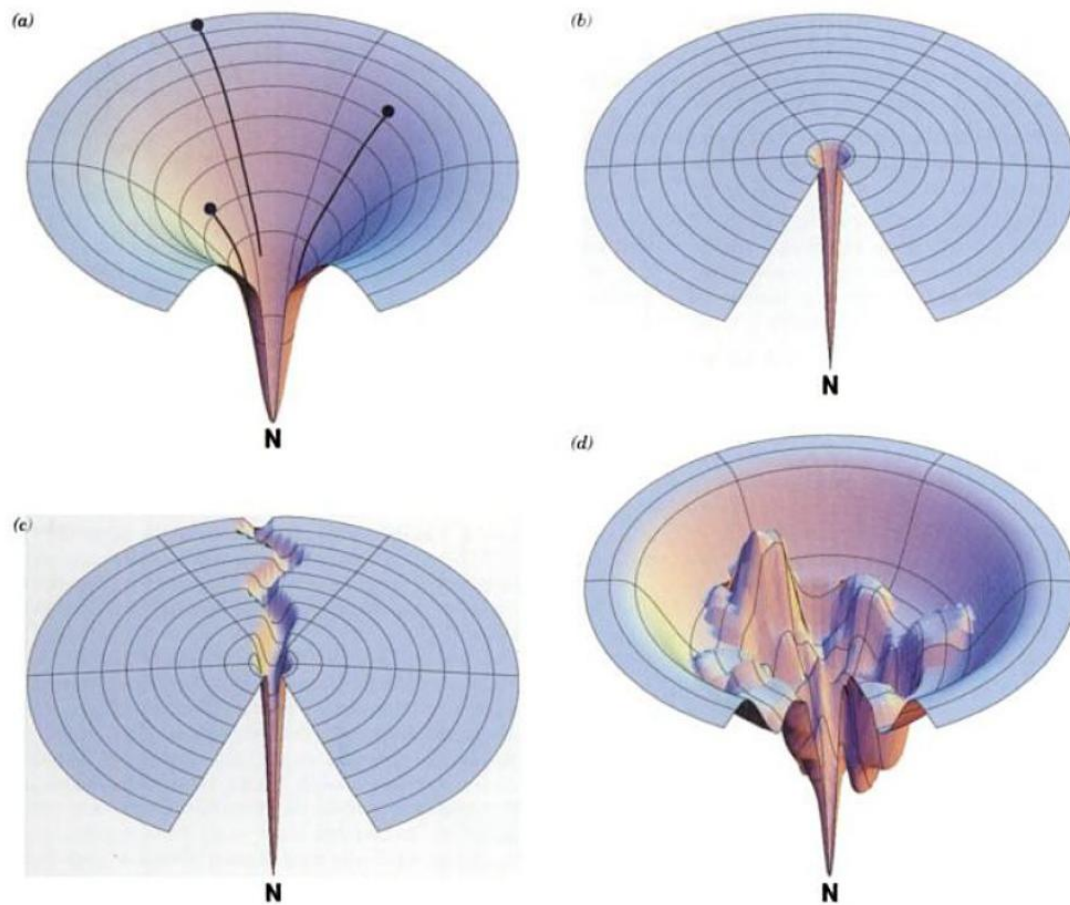


Figure 39 : Entonnoirs de repliement.

- Paysage en entonnoir idéalisé. A mesure que la chaîne polypeptidique établit un nombre croissant de contact, son énergie libre interne (sa hauteur au dessus de l'état natif, N) décroît ainsi que sa liberté conformationnelle. Des polypeptides de conformations différentes suivent différentes voies (lignes noires) pour atteindre l'état natif.
- Paysage de Levinthal en terrain de Golf dans lequel la chaîne polypeptidique doit trouver son état natif (le trou) par recherche aléatoire, sur une surface énergétique plane.
- Paysage de repliement classique où la chaîne se déplace au hasard sur une surface énergétique plane jusqu'à ce qu'elle rencontre une gorge qui conduit à l'état natif.
- Surface énergétique accidentée présentant des minima locaux dans lesquels un polypeptide en cours de repliement peut être piégé momentanément.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

Les protéines se replient via une suite d'ajustements conformationnels qui diminuent leur énergie libre et leur entropie jusqu'à ce qu'elles atteignent l'état natif (**Figure 40**).

Beginning of helix formation and collapse

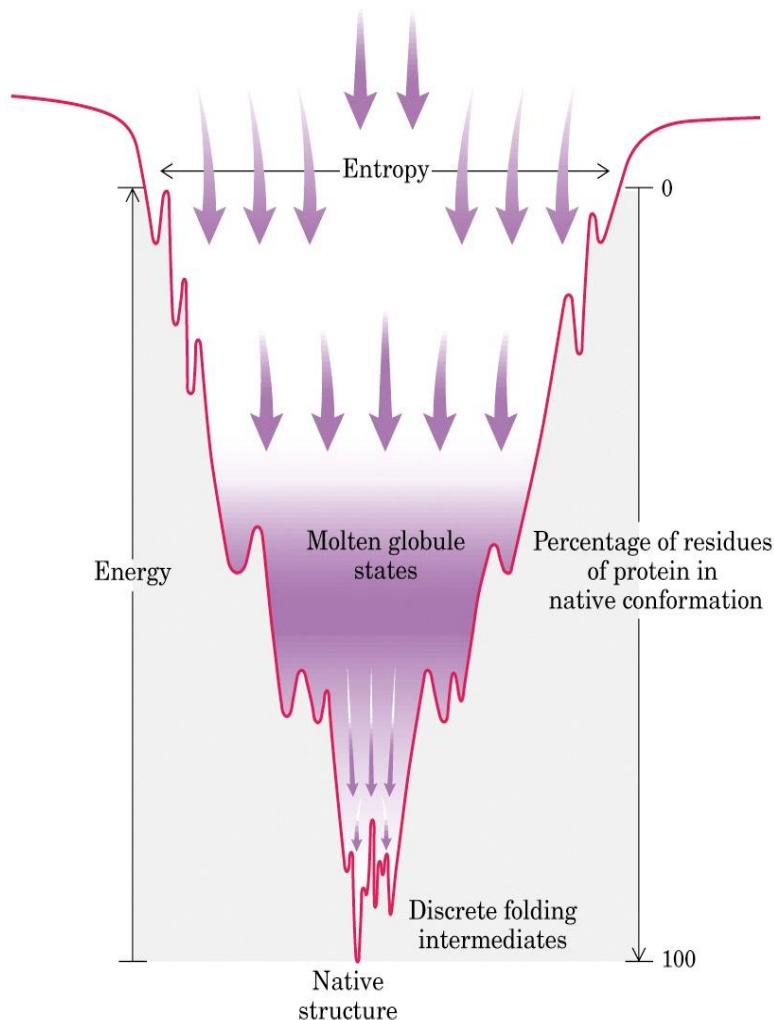


Figure 40 : Formation de l'état Molten globule.

d. Le repliement des protéines suit une hiérarchie

La théorie du paysage est en accord avec le repliement hiérarchique. In vivo, le repliement d'un polypeptide commence avec sa synthèse dès la sortie du ribosome. Dans ces conditions, la protéine atteindra plus facilement l'état natif si son repliement est Hiérarchique.

Chapitre II : Repliement des protéines, dynamique et évolution structurale

e. L'inhibiteur trypsique pancréatique bovin (BPTI) prend sa conformation native selon un mécanisme séquentiel

Le BPTI complètement réduit est entièrement déroulé ("random coil"). Ceci a permis à *Thomas Creighton* de déterminer l'ordre de formation des ponts disulfure dans le BPTI en amorçant son repliement par addition d'un réactif sulfhydryl sous forme oxydée (**Figure 41**).

Au fur et à mesure que la protéine se renature, les intermédiaires formés sont piégés en diminuant le pH, ce qui inhibe la formation de l'anion thiolate requis pour les réactions d'échange de ponts disulfure. Les intermédiaires sont séparés par chromatographie.

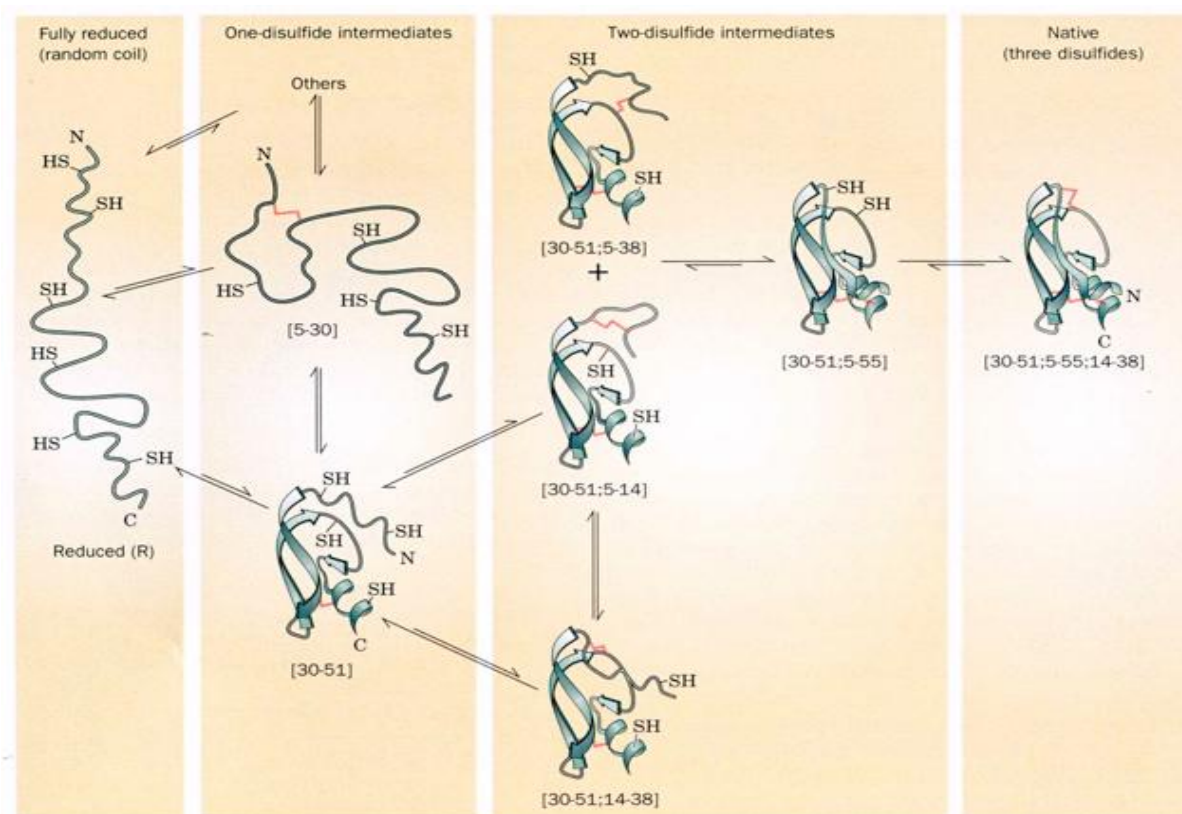


Figure 41 : Squelette polypeptidique et ponts disulfure du BPTI natif.

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

I Repliement des protéines

La plupart des protéines déroulées se renaturent *in vitro* sur des périodes qui vont de quelques minutes à plusieurs jours. Les chaînes polypeptidiques prennent des conformations non natives quasi-stables et/ou forment des agrégats non spécifiques.

Le repliement *in vitro* représente un bon modèle pour comprendre les mécanismes par lesquels une chaîne polypeptidique naissante acquiert sa structure tridimensionnelle.

Cependant, l'environnement intracellulaire diffère de celui utilisé *in vitro*, l'intérieur des cellules est très concentré en macromolécules. *In vivo*, les polypeptides prennent leur conformation native de manière efficace au fur et à mesure de leur synthèse, processus qui demande normalement que quelques minutes ou moins.

Ceci parce que toutes les cellules contiennent trois types de protéines auxiliaires dont la fonction est d'aider les polypeptides à se replier pour prendre leur conformation native et d'assurer leur assemblage pour atteindre leur structure quaternaire ; ce sont les protéines disulfure isomérase (PDI), les peptidyl prolyl cis-trans isomérase, et les protéines chaperons moléculaires.

QUELS SONT LES DIFFÉRENTS MODÈLES DE CHEMINS DE REPLIEMENT ?

On distingue des modèles construits à partir de considérations théoriques. Différents modèles de repliement furent proposés soit à partir d'observations expérimentales, soit à partir de considérations théoriques ou de simulations. Un certain nombre de modèles seront présentés brièvement afin de faire le point sur les différents mécanismes qui sont proposés :

I.1 Le modèle de nucléation-condensation

Le modèle de nucléation-condensation est selon lequel se formeraient des noyaux de structures locales stabilisés par des interactions à longue distance.

Il inclut une étape de nucléation suivie par une propagation rapide de la structure. L'étape limitante est le processus de nucléation. Ce modèle, sensé représenter les premiers résultats cinétiques, prend en compte la coopérativité du repliement.

I.2 Le modèle de diffusion - collision, proposé par Karplus et Weaver (1976)

La nucléation intervient simultanément en plusieurs points de la chaîne polypeptidique et génère des structures qui diffusent et s'associent puis deviennent coalescentes (soudure/adhérence de deux surfaces) jusqu'à former des sous-structures ayant une conformation

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

native. Ces microstructures ont une durée de vie contrôlée par la diffusion, et le repliement d'une chaîne de 100 ou 200 acides aminés peut avoir lieu en moins d'une seconde. Selon ce modèle, le repliement intervient en passant par plusieurs étapes de diffusion collision.

I.3 Le repliement Folding Units

Un modèle modulable de repliement a été proposé en considérant la structure tertiaire des protéines. Wetlaufer (1981) considère les domaines comme des unités de repliement (*Folding Units*) ; il suggère que les sous-domaines pourraient se replier de façon indépendante en formant, comme intermédiaires de repliement, des modules structuraux qui s'assemblent pour donner la structure native.

I.4 Le repliement séquentiel et hiérarchique

Plusieurs segments de structures sont formés et assemblés à différents niveaux suivant un chemin de repliement unique. Schulz (1977) a proposé une hiérarchie dans le repliement des protéines correspondant aux structures primaire – secondaire – tertiaire – et éventuellement quaternaire. Dans ce modèle, le premier événement, la nucléation, est suivi par la formation des structures secondaires qui s'associent pour former des superstructures secondaires, puis des domaines et éventuellement un monomère actif suivant un chemin unique. L'association entre les sous-unités des protéines oligomériques se ferait à l'étape finale du processus de repliement.

I.5 Le modèle de la fermeture éclair hydrophobe (*Hydrophobic Zipper*), proposé par Dill et al. (1993)

Ce modèle suggère que la formation de segments de structures secondaires est simultanée à la formation du collapse hydrophobe.

I.6 Le modèle *Framework* (Ptitsyn, 1991)

Ce modèle très semblable au **repliement séquentiel et hiérarchique**, pose comme hypothèse que la structure secondaire est formée à une étape précoce du repliement avant la mise en place de la structure tertiaire. Un modèle modulaire fondé sur l'examen des structures tridimensionnelles de protéines considère les domaines, et même les sous-domaines, comme des unités autonomes de repliement. Ce modèle accepte pour chaque module un chemin de repliement unique et hiérarchique. Ces modèles privilégient le rôle des interactions à courte distance dans les premiers événements du repliement.

Chapitre III : Repliement des protéines *in vitro et in vivo*, et rôle des protéines auxiliaires

I.7 Le modèle d'effondrement (collapse) hydrophobe

Ce modèle implique que le premier évènement du repliement des protéines est un collapse hydrophobe ayant lieu avant la formation des structures secondaires. Ce collapse conduirait au repliement des protéines et à la stabilisation de la structure tertiaire. L'enfouissement des résidus hydrophobes peut avoir lieu dans les étapes précoces du repliement.

I.8 Le modèle du *Jigsaw Puzzle* (Harrison, Durbin, 1985)

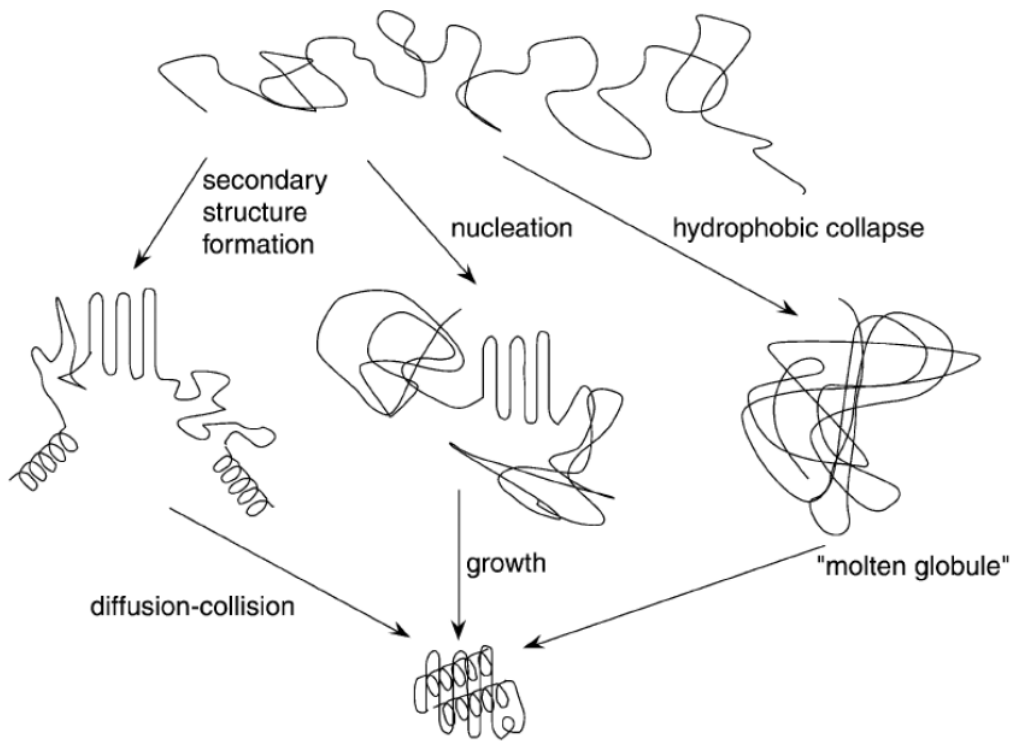
Dans ce modèle, le repliement est considéré comme un assemblage en puzzle avec existence de plusieurs chemins menant à une solution unique. Les intermédiaires du lysozyme se réarrangeraient pour générer un globule fondu à partir duquel la protéine native se formerait en une ou plusieurs étapes séquentielles, en suivant le *Jigsaw Puzzle*. Ce modèle

présente des similarités avec celui de Karplus et Weaver. L'identification d'intermédiaires représente une description cinétique plutôt qu'une description structurale, chaque intermédiaire étant constitué d'espèces hétérogènes. Le modèle du *Jigsaw*

Puzzle apporte un éclairage nouveau car on a longtemps postulé que le repliement des protéines suivait un chemin unique en passant par la création séquentielle d'intermédiaires instables le long de ce chemin.

Ces différents modèles ont été construits à partir de résultats expérimentaux. Comme on peut le constater, il n'y a pas de consensus sur les modèles de repliement proposés dans la littérature. On peut simplement dire qu'à la lumière de telle ou telle expérience, un modèle paraît plus crédible qu'un autre pour illustrer le ou les chemins de repliement.

Modèles classiques pour le repliement



Modèles classiques pour le repliement

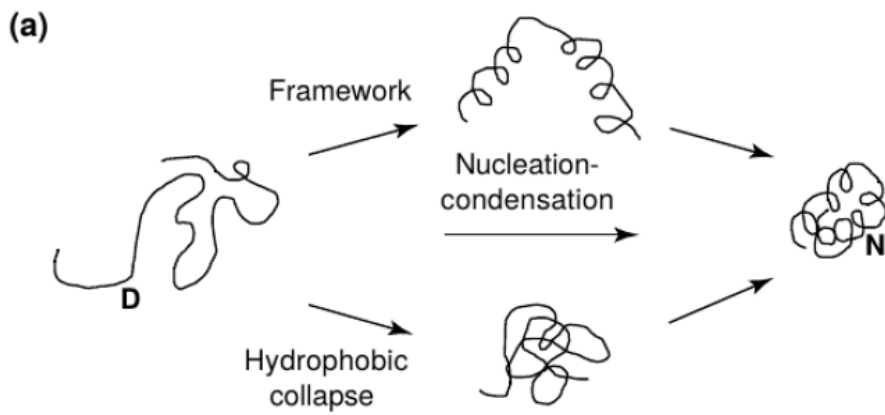


Figure 42 : Représentation schématique des modèles de repliement.

II Rôle des protéines auxiliaires du repliement

Toutes les cellules contiennent trois types de protéines auxiliaires dont la fonction est d'aider les polypeptides à se replier pour prendre leur conformation native et d'assurer leur assemblage pour atteindre leur structure quaternaire.

II.1 Protéines disulfure isomérase

La **protéine disulfure isomérase (PDI)** est une enzyme homodimérique d'eucaryote comportant 486 résidus par sous unité. Sous sa forme réduite, la PDI catalyse les réactions d'échange entre liaisons disulfure, facilitant ainsi le brassage des ponts disulfure des protéines jusqu'à ce qu'ils trouvent leur appariement définitif caractéristique de la protéine native. De plus, la PDI doit intervenir pour faciliter le repliement correct des protéines qui se dénaturent en absence de leurs ponts disulfure natifs. La sous unité PDI est constituée de 4 domaines de ± 100 résidus, qui sont homologues à la **thiorédoxine**, protéine rédox ubiquitaire contenant un pont disulfure (**Figure 43**).

Les groupements thiol de ces deux cystéines sont actifs à cause de leur pKa (7,3) plus faible que le pKa habituel d'un résidu Cys dans une protéine. Le site actif de la PDI contient le motif **Cys-Gly-His-Cys**, dans lequel le premier résidu Cys, sous sa forme $-SH$, participe à la réaction d'échange de ponts disulfure. La fonction du deuxième résidu Cys, sous sa forme $-SH$, est de libérer la PDI des ponts disulfure stables que le premier résidu Cys peut établir avec les substrats protéiques. On obtient ainsi d'une part des substrats réduits et d'autre part une PDI où les deux résidus Cys du site actif sont reliés par un pont disulfure.

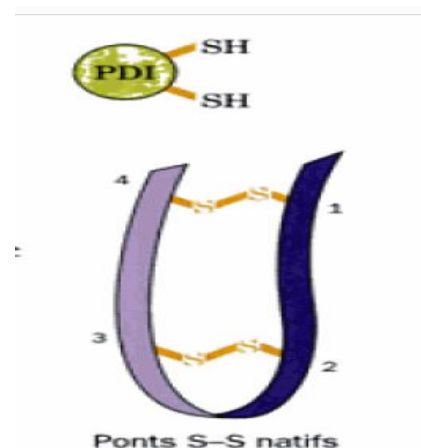


Figure 43 : Protéine disulfure isomérase (PDI).

II.2 Réactions catalysées par la protéine disulfure isomérase

La PDI réduite catalyse le réarrangement des ponts disulfure non natifs dans un substrat protéique via des échanges disulfure pour donner les ponts disulfure natifs.

Un pont disulfure s'établit entre la PDI et le substrat protéique, à cet échange la PDI est oxydée et le substrat protéique est réduit (**Figure 44**).

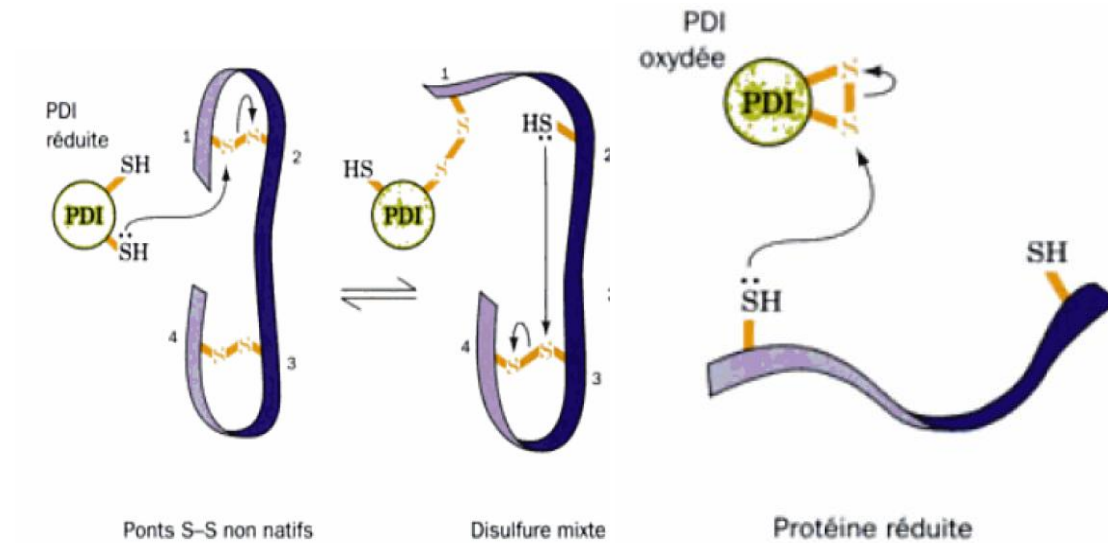


Figure 44 : Réactions catalysées par la protéine disulfure isomérase.

La réaction entraîne la formation d'un disulfure mixte entre la PDI et la protéine (**Figure 45**).

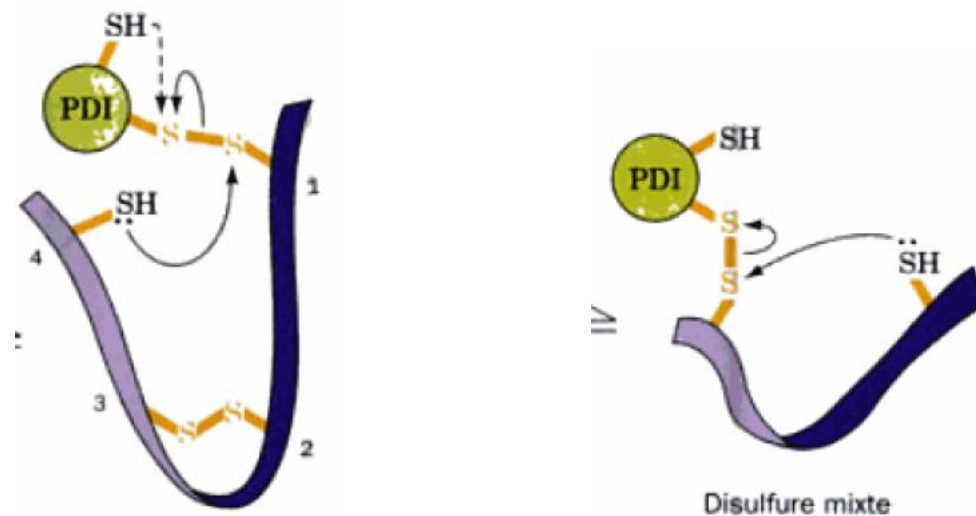


Figure 45 : Réactions de formation d'un disulfure mixte par la protéine disulfure isomérase.

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

Le produit réduit de la réaction réagit avec des agents oxydants de la cellule pour régénérer la PDI oxydée (Figure 46).

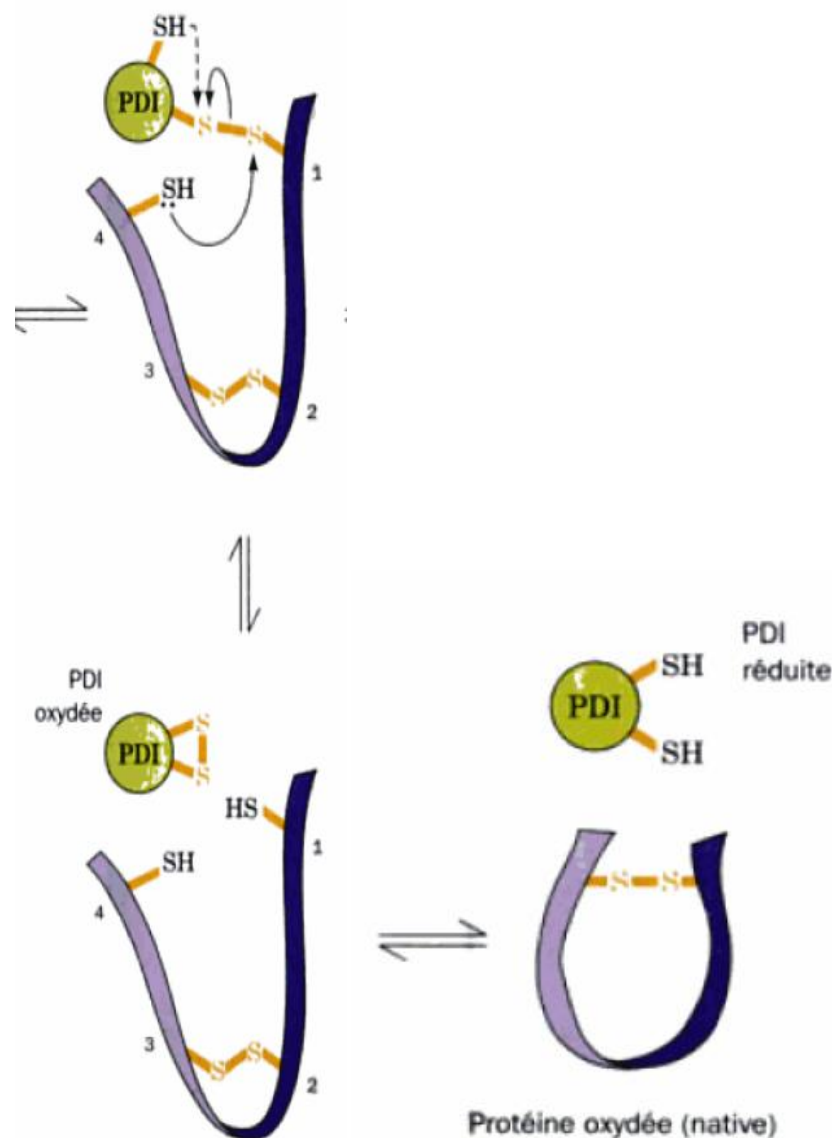


Figure 46 : Régénération de la PDI.

Le site actif de la PDI occupe une région qui semble capable de lier des segments polypeptidiques déroulés. La PDI oxydée est moins stable que la réduite et possède un pont disulfure hautement réactionnel, très oxydant. Ceci permet à la PDI oxydée d'introduire directement par réaction d'échange des ponts disulfure dans les polypeptides. Pour assurer cette

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

fonction en continu, la PDI réduite doit être réoxydée, son pont disulfure rétabli par les agents oxydants cellulaires.

II.3 Peptidyl prolyl cis-trans isomérase

Les peptidyl prolyl cis-trans isomérase (PPI; appelées aussi rotamases) catalysent l'interconversion des liaisons peptidiques Xaa-Pro entre leurs configurations cis et trans, accélérant ainsi le repliement des polypeptides contenant des résidus Pro. On a caractérisé deux familles de Ppi appelées immunophilines ; les cyclophilines (ainsi appelées car elles sont inhibées par le médicament immunosuppresseur cyclosporine A, un peptide cyclique de 11 résidus produit par un champignon) et la famille dont la FK506 binding protein (protéine de 12 kD appelée FKBP12).

Les deux familles de PPI (les cyclophilines et la FK506 binding protéine de 12 kDa ou FKBP12) sont inhibés par les médicaments immunosuppresseur cyclosporine A et FK506 (lactones macrocyclique produites par des champignons) utilisés pour le traitement des maladies auto-immunes et pour empêcher le rejet d'organes transplantés. La cyclosporine A et la FK506 empêchent l'activation des lymphocytes T du système immunitaire. Le mystère c'est qu'il n'y a pas de relation entre les propriétés immunosuppressives de ces substances et les activités des rotamases.

II.4 Chaperons moléculaires: le système GroEL/ES

Les chaperons moléculaires sont des protéines qui empêchent la formation des agrégats intra- et intermoléculaires pendant le repliement. Pour cela, elles se lient aux surfaces hydrophobes exposées au solvant d'un polypeptide déroulé ou agrégé puis elles s'en détachent, ces deux opérations pouvant être répétées de façon à faciliter leur repliement et ou l'obtention de la structure quaternaire.

Beaucoup de chaperons moléculaires sont des ATPases (hydrolyse d'ATP pour se détacher du polypeptide).

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

Les chaperons moléculaires comprennent plusieurs classes parmi lesquels on trouve ; les protéines de choc thermique, comme Hsp70, les chaperonines, Hsp90, et les nucléoplasmines.

Les protéines de choc thermique 70 (Hsp70) (**Figure 47**) sont des protéines monomériques aussi bien chez les procaryotes que chez les eucaryotes (ces derniers en possèdent différents types dans le cytosol, le réticulum endoplasmique, les mitochondries et les chloroplastes). Elles sont ainsi appelées car leur vitesse de synthèse augmente fortement à température élevée. Les Hsp70 utilisent l'énergie de l'ATP pour inverser la dénaturation et l'agrégation des protéines, et faciliter le repliement correct des polypeptides dès leur sortie du ribosome. En association avec le cochaperon protéique Hsp40, Hsp 70 lie et lide petites régions hydrophobes dans les protéines mal repliées.

Les chaperonines forment des complexes de grande taille. Elles se lient aux surfaces hydrophobes accessibles des protéines globulaires mal repliées, dans un processus ATP dépendant, empêchent leur agrégation.

Les protéines Hsp90, sont impliquées surtout dans le repliement de protéines qui jouent un rôle dans la transduction du signal, comme les récepteurs des hormones stéroïdes.

Les nucléoplasmines, protéines nucléaires acides, dont la présence est nécessaire pour l'assemblage correct *in vivo* des nucléosomes (particules autour desquelles l'AND des eucaryotes est enroulé) à partir de leurs constituants, l'AND et les histones.

Les protéines Hsp60 (GroEL chez *E. coli*) sont constituées de 14 sous-unités identiques de 60 kDa disposées en deux anneaux accolés de 7 sous-unités chacun, entourant une cavité centrale hydrophobe.

Les 7 sous-unités de GroES forment une structure en dôme qui ferme une des sorties du cylindre GroEL. Un anneau GroEL, qui lie 7 ATP et un substrat protéique mal replié, se lie à GroES.

Chapitre III : Repliement des protéines *in vitro* et *in vivo*, et rôle des protéines auxiliaires

Un changement de conformation libère le substrat protéique dans la cavité où il commence à se replier. La cavité tapissée de groupements hydrophiles (**cage d'Anfinsen**) empêche l'agrégation du substrat protéique.

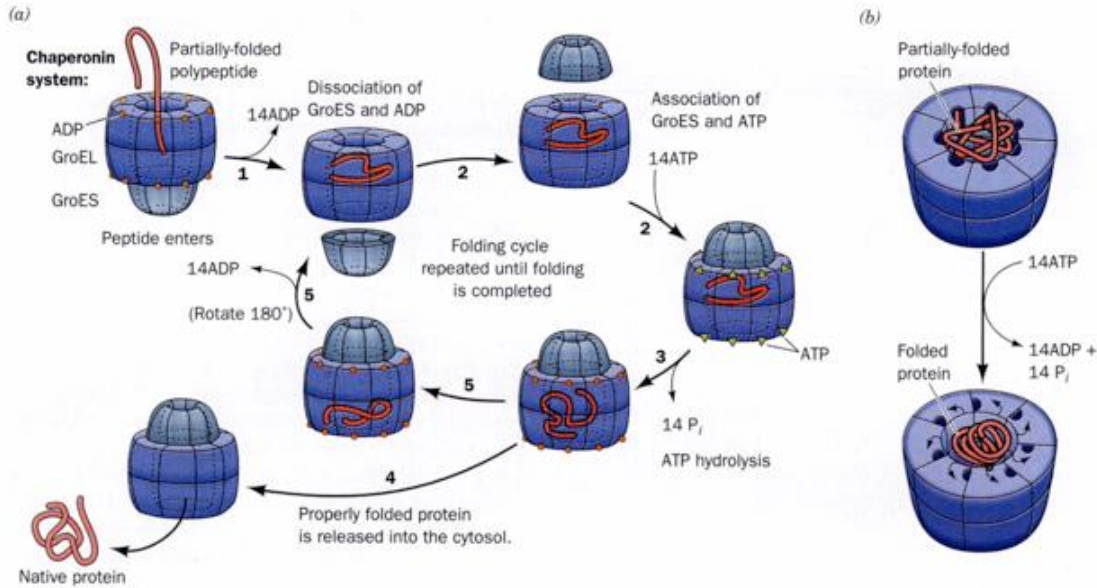


Figure 47 : Chaperons moléculaires: le système GroEL/ES.

II.5 Le système GroEL/ES fonctionne comme un moteur à deux temps: (Figure 48)

1. Un anneau GroEL cis qui lie 7 ATP et un substrat protéique mal replié via les régions hydrophobes, se lie à GroES. Ceci induit un changement conformationnel dans l'anneau GroEL ce qui libère le substrat protéique dans la cavité agrandie et fermée, où ce substrat commence à se replier.

2. Dans les 13 s environ l'anneau catalyse l'hydrolyse de ses 7 ATP en ADP+ P_i , et ceci affaiblit les interactions qui unissent GroES et GroEL.

3. Une deuxième molécule de substrat protéique, puis 7 ATP se lient à l'anneau trans.

4. La liaison du substrat protéique et de l'ATP à l'anneau trans provoque la libération de GroES liée à l'anneau cis, ainsi que des 7 ADP et du substrat protéique mieux replié. Seuls l'ATP et le substrat protéique mal replié restent liés à l'ancien anneau trans de GroEL qui devient anneau cis en liant à nouveau GroES.

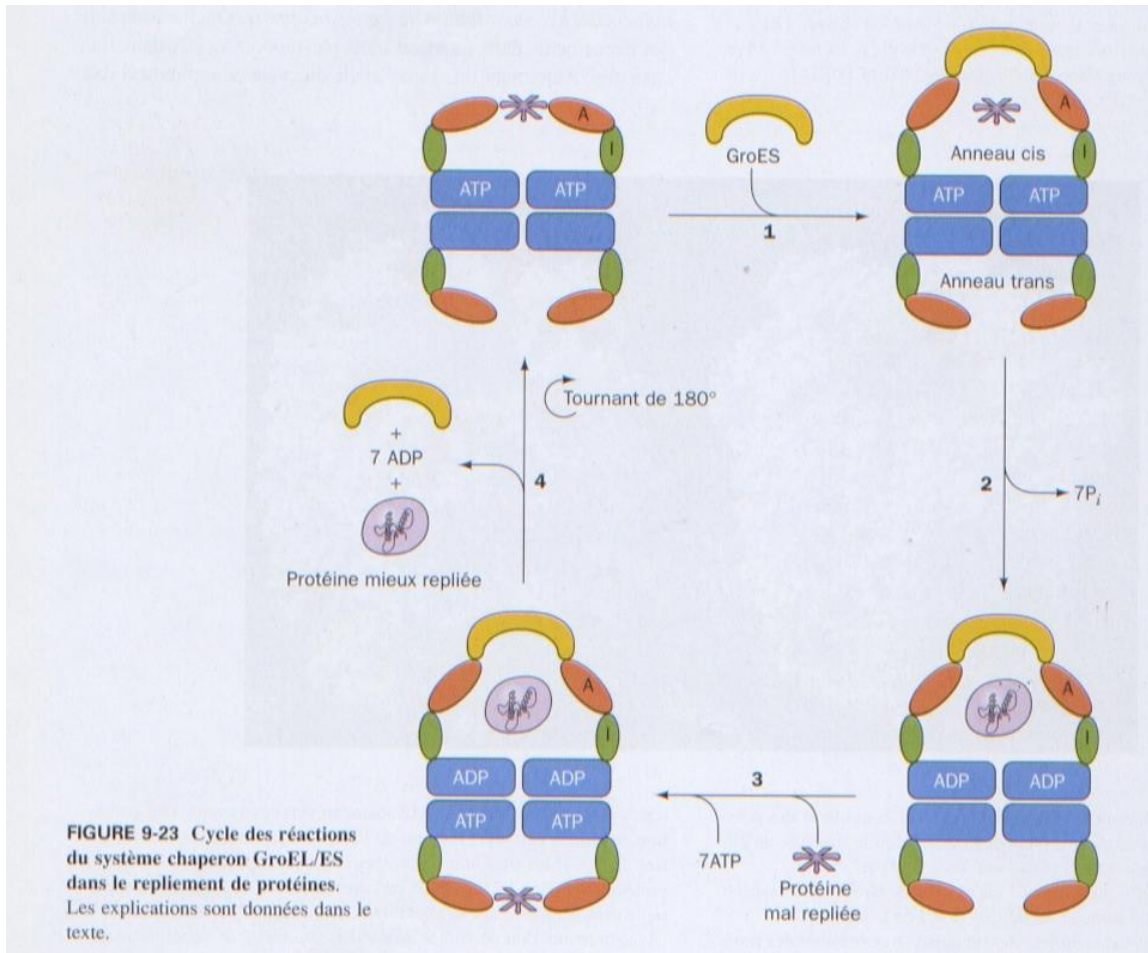


Figure 48 : Cycle des réactions du système GroEL/ES.

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

I Prédiction des structures secondaires

Dans la mesure où la structure primaire d'une protéine spécifie sa structure tridimensionnelle, on devrait pouvoir en principe prévoir la structure native d'une protéine en ne connaissant que sa structure primaire.

On utilise une approche théorique fondée sur des principes physicochimiques, soit par des méthodes qui utiliseraient des schémas prévisionnels obtenus après analyse de structures connues de protéines. La manière la plus fiable de déterminer la structure secondaire prise par un polypeptide est de comparer sa séquence en acides aminés avec celle d'un homologue de structure connue. En l'absence d'une telle structure, il faudra recourir aux méthodes prédictives. Les méthodes théoriques de prédiction des structures tertiaires, prédisent nécessairement aussi les structures secondaires.

a. La méthode Chou-Fasman

Un résidu Pro ne peut s'insérer à l'intérieur d'hélices ou de feuillets réguliers car son cycle remplirait l'espace dans la chaîne et il ne présente pas le groupement NH du squelette permettant d'établir une liaison hydrogène.

La méthode **Chou-Fasman** est basée sur une classification des résidus d'acide aminé selon leur proportion pour les conformations en hélice α ou en feuillet β (fondée sur l'analyse de 29 structures par rayons X). Pro (par déformation) et Gly (trop souple) déstabilise l'hélice α

Le schéma de prévision de structure mis au point par Peter Chou et Gerald Fasman est très populaire car facile à appliquer et assez fiable.

$$f_{\alpha} = n_{\alpha} / n$$

f_{α} est la fréquence de l'apparition d'un résidu donné dans une hélice α

n_{α} est le nombre de résidus de l'acide aminé donné que l'on trouve dans des hélices α

n est le nombre total de résidus de ce même acide dans la série

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

La propension P d'un résidu particulier d'acide aminé à se trouver dans une hélice est définie par ;

$$P_{\alpha} = f_{\alpha} / \langle f_{\alpha} \rangle$$

$\langle f_{\alpha} \rangle$ est la valeur moyenne de f_{α} pour les 20 résidus différents.

Une valeur de P supérieur à 1 signifie qu'un résidu se trouve dans une hélice avec une fréquence supérieure à la moyenne.

La propension d'un résidu à se trouver dans un feuillet se définit de la même manière.

La méthode Chou-Fasman est basée sur une classification des résidus d'acide aminé selon leur propension pour les conformations en hélice α ou en feuillet β (fondée sur l'analyse de 29 structures par rayons X). Selon la valeur de sa propension, un résidu est classé comme ;

(H) très favorable, (h) favorable, (I) peu favorable, (i) indifférent, (b) défavorable, (B) très défavorable

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

Résidue	P_{α}	Classification de l'hélice	P_{β}	Classification pour le feuillet
Ala	1,42	H_{α}	0,83	i_{β}
Arg	0,98	i_{α}	0,93	i_{β}
Asn	0,67	b_{α}	0,89	i_{β}
Asp	1,01	l_{α}	0,54	B_{β}
Cys	0,70	i_{α}	1,19	h_{β}
Gln	1,11	h_{α}	1,10	h_{β}
Glu	1,51	H_{α}	0,37	B_{β}
Gly	0,57	B_{α}	0,75	b_{β}
His	1,00	l_{α}	0,87	h_{β}
Ile	1,08	h_{α}	1,60	H_{β}
Leu	1,21	H_{α}	1,30	h_{β}
Lys	1,16	h_{α}	0,74	b_{β}
Met	1,45	H_{α}	1,05	h_{β}
Phe	1,13	h_{α}	1,38	h_{β}
Pro	0,57	B_{α}	0,55	B_{β}
Ser	0,77	i_{α}	0,75	b_{β}
Thr	0,83	i_{α}	1,19	h_{β}
Trp	1,08	h_{α}	1,37	h_{β}
Tyr	0,69	b_{α}	1,47	H_{β}
Val	1,06	h_{α}	1,70	H_{β}

Source : Chou, P.Y. and Fasman, G.D., *Annu. Rev. Biochem.* **47**, 258 (1978).

Tableau 06 : Classification des résidus d'acide aminé selon leur propension pour les conformations en hélice et ou en feuillet.

Ces règles assez simples permettent de prévoir les positions de segments en hélice ou en feuillet dans une protéine avec une fiabilité moyenne de l'Ordre de 50 à 80%.

La méthode Chou-Fasman permet aussi de prévoir la position des coudes.

Cependant il existe un grand nombre d'algorithmes informatiques pour la prédiction des structures secondaires comme le **Jpred**, d'accès libre.

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

II Prédiction des structures tertiaires

Les bases de données de séquences contiennent celles de > 300 000 polypeptides. Cependant, on ne connaît les structures que d'environ 20 000 protéines. On peut prédire les structures tertiaires de protéines par la méthode de modélisation comparative ou par homologie.

On aligne la séquence d'intérêt avec celle d'une protéine homologue de structure connue, et l'on corrige pour les substitutions en acides aminés, les insertions et les délétions, par modélisation et calculs de minimisation d'énergie.

La technique de calcul dite de la reconnaissance du repliement ou du glissement, tente de déterminer le repliement inconnu d'une protéine en recherchant si sa séquence est compatible avec un membre d'une bibliothèque de structures protéiques connues. Ces méthodes utilisent des algorithmes d'alignements de séquences.

III Dynamique des protéines

Les protéines sont des molécules souples qui oscillent en permanence et dont les mouvements structuraux ont une signification fonctionnelle.

Les mouvements intramoléculaires des protéines sont classés en trois grandes catégories ;

1. Fluctuations d'ordre atomique, telles que les vibrations de liaisons individuelles.

2. Mouvements collectifs, ou les groupes d'atomes liés par covalence, dont la taille est comprise entre celles de chaînes latérales d'acides aminés et celles de domaines entiers, se déplacent comme des entités.

3. Changements conformationnels provoqués, ou des groupes d'atomes, de tailles qui vont de la chaîne latérale à la sous-unité entière, se déplacent en réponse à des stimuli spécifiques tels que la liaison d'une petite molécule par exemple liaison de l'ATP à GroEL.

a. Les protéines ont des structures fluctuantes, par exemple, la myoglobine a un cœur rigide qui entoure son noyau hème et que les régions vers la périphérie de la molécule ont un caractère plus mobile. De même, le domaine apical de GroEL et la boucle mobile de GroES sont tous deux très

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

flexibles dans les protéines séparées, mais deviennent nettement plus rigides lorsqu'ils interagissent au sein du complexe GroEL-GroES-ADP.

En effet de nombreuses protéines se rigidifient suite à l'interaction avec leurs molécules cibles, cette flexibilité initiale facilite la formation d'un complexe spécifique entre une protéine et son partenaire de liaison.

b. Les mouvements des atomes dans les protéines sont dues aux écartements par rapport aux longueurs des liaisons covalentes, des angles et des angles de torsion, ainsi qu'aux interactions non covalentes, cette approche est appelée **dynamique moléculaire**.

c. La structure native d'une protéine consiste en une série de nombreux sous états conformationnels qui ont des stabilités identiques. Ces sous états ont des arrangements atomiques légèrement différents, qui s'interagissent au hasard à des vitesses qui augmentent avec la température. En théorie en température en dessous de -53°C , les mouvements collectifs dans les protéines sont arrêtés, laissant que les fluctuations atomique comme mouvement dominant.

IV Maladies conformationnelles; amyloïdes et prions

IV.1 Les protéines amyloïdes

Dans l'organisme la plupart des protéines gardent leur conformation native ou, en cas de dénaturation partielle, se renaturent avec l'aide des chaperons moléculaires, ou sont dégradées par protéolyse. Cependant il existe des maladies humaines qui sont associées à des dépôts extracellulaires de protéines normalement solubles. Ces protéines s'accumulent dans certains tissus sous forme d'agrégats insolubles connus sous le nom d'amyloïde. Les dépôts d'amyloïde interfèrent avec les fonctions normales des cellules, ce qui peut causer leur mort et la défaillance de l'organe atteint. Ces maladies comprennent la **maladie d'Alzheimer**, neurodégénérative qui touche les personnes âgées, les **encéphalopathies spongiformes transmissibles**, une famille de maladies infectieuses neurodégénératives, et les **amyloïdoses**, qui regroupent des maladies causées par le dépôt d'une protéine souvent mutée dans des organes.

Les différents types de protéines amyloïdogènes ne sont pas apparentés et leurs structures natives adoptent des repliements très différents.

Dans la **maladie d'Alzheimer**, une affection neurodégénérative qui touche les personnes âgées de plus de 65 ans, se caractérise par la présence dans le tissu cérébral d'un grand nombre de **dépôts amyloïdes** ou **plaques séniles**. Les plaques séniles sont des dépôts de fibrilles amyloïdes

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

constituées d'une protéine de 40 à 42 résidus, appelée le **peptide amyloïde β ($A\beta$)**, ce dernier est un fragment d'une protéine transmembranaire appelée **précurseur du peptide amyloïde APP**.

la protéolyse du précurseur du peptide amyloïde (APP) dans le cerveau donne naissance au peptide amyloïde β ($A\beta$) lequel forme les fibrilles amyloïdes qui tuent les neurones. La relation entre maladie d'Alzheimer et âge, suggère que le dépôt amyloïde est un processus évolutif. Les mutations du gène APP dont les mutations touchent la région **$A\beta$ et dont les porteurs sont atteints de MA**. Ces mutations affectent le découpage protéolytique de l'APP, de telle sorte que la production du peptide amyloïde est accrue.

IV.2 Les maladies à prions

- la "**tremblante**", caractérisée par un dérèglement neurologique chez les ovins, appelée "**scrapie**" en anglais.

- l'**encéphalopathie spongiforme bovin (ESB ou maladie de la vache folle)** qui affecte de manière semblable les bovins.

- la **maladie de Creutzfeld-Jacob (CJD)**, **maladie** dégénérative du cerveau chez l'homme avec les mêmes symptômes. Elles sont toutes caractérisées par l'apparition dans les neurones de grandes vacuoles qui en microscopie donnent au tissu cérébral l'aspect d'une éponge.

La protéine de la tremblante est appelée **PrP ("Prion Protein"**; Prion = "proteinaceous infectious particle").

Il existe deux formes de PrP: la forme cellulaire normale, PrPC, une protéine conservée ancrée dans la membrane à la surface cellulaire des neurones; et PrPSc, qui possède une conformation différente.

La protéine de la tremblante est appelée **PrP ("Prion Protein«**; Prion = "proteinaceous infectious particle«). Il existe deux formes de PrP: la forme cellulaire normale,

PrPC, une protéine conservée ancrée dans la membrane à la surface cellulaire des neurones; et PrPSc, qui possède une conformation différente. On suppose que le PrPC est convertit en PrPSc par un processus autocatalytique, l'on présume être directement responsable de la formation de la plaque amyloïde et de la dégénérescence neuronale (**Figure 49**).

Chapitre IV : Prédiction de la structure, dynamique des protéines et maladies conformationnelles

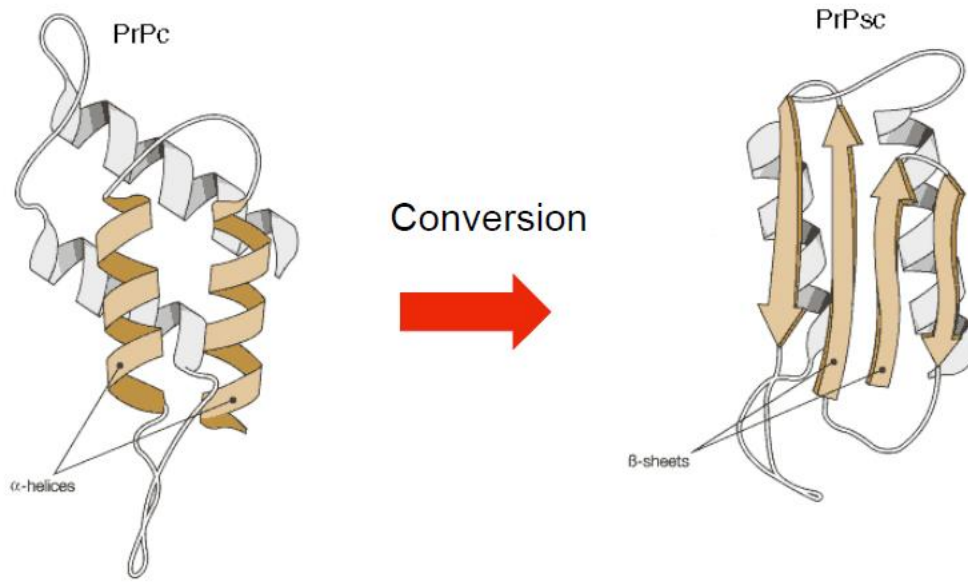


Figure 49 : Structure tridimensionnelle de PrPC et de PrPSc.

Chapitre V : Interaction protéine ligand et interaction allostérique

I Introduction

Un principe fondamental de tous les processus biologiques est l'organisation moléculaire et la reconnaissance, par exemple, les interactions protéine-ligand et la liaison.

Essentiellement, les protéines remplissent leur fonction par interaction avec des molécules telles que des protéines et des peptides, des acides nucléiques, des ligands et des substrats, et d'autres petites molécules telles que l'oxygène ou les ions métalliques.

Ainsi, une compréhension détaillée des processus biologiques nécessite d'étudier non seulement le mécanisme de repliement des protéines mais aussi le mécanisme de liaison protéine-ligand.

II Les ligands

Un ligand (du latin ligandum, liant) Ce sont des molécules de signalisation qui en se fixant sur des récepteurs spécifiques (de surface ou intracellulaire) déterminent une réponse cellulaire. Le ligand est de type endogène ou naturelle comme : les hormones, neurotransmetteurs, cytokines, ou ligand exogène : comme les médicaments qui miment ou bloquent les effets d'un ligand endogène.

Un ligand, par définition, est toute substance qui peut former un complexe non-covalent avec une biomolécule. La reconnaissance moléculaire est un processus par lequel les macromolécules biologiques interagissent les unes avec les autres ou avec de petites molécules pour former un complexe spécifique. Il a deux caractéristiques importantes: la spécificité et l'affinité. La spécificité est que les macromolécules biologiques sont capables de distinguer le ligand hautement spécifique d'autres moins spécifiques. Afin de rendre une interaction spécifique, les liens entre les partenaires corrects devraient être forts, alors que pour les autres partenaires ne présentant que des différences mineures de structure, ils devraient être faibles ou même répulsifs. Par conséquent, une fois qu'un ligand correct se lie à sa molécule d'accepteur, l'affinité entre eux devrait être suffisamment forte pour empêcher la dissociation.

L'aspect le plus important de la reconnaissance moléculaire est que la reconnaissance n'est généralement pas un processus en soi mais un élément d'un mécanisme plus complexe et fonctionnellement important. Comme exemple la régulation allostérique de l'activité enzymatique, la transduction du signal, le repliement des protéines ou la formation de

Chapitre V : Interaction protéine ligand et interaction allostérique

structures multi-sous-unités et supramoléculaires sont des mécanismes de la reconnaissance moléculaire. Cela nécessite des changements importants et parfois spectaculaires dans les propriétés des partenaires en interaction.

III Mécanismes de liaison protéine-ligand

Les mécanismes de liaison protéine-ligand ont évolué depuis les premiers «**serrure et clé**» jusqu'à «**l'ajustement induit**» et aux modèles de «**sélection conformationnelle**» (Figure 50).

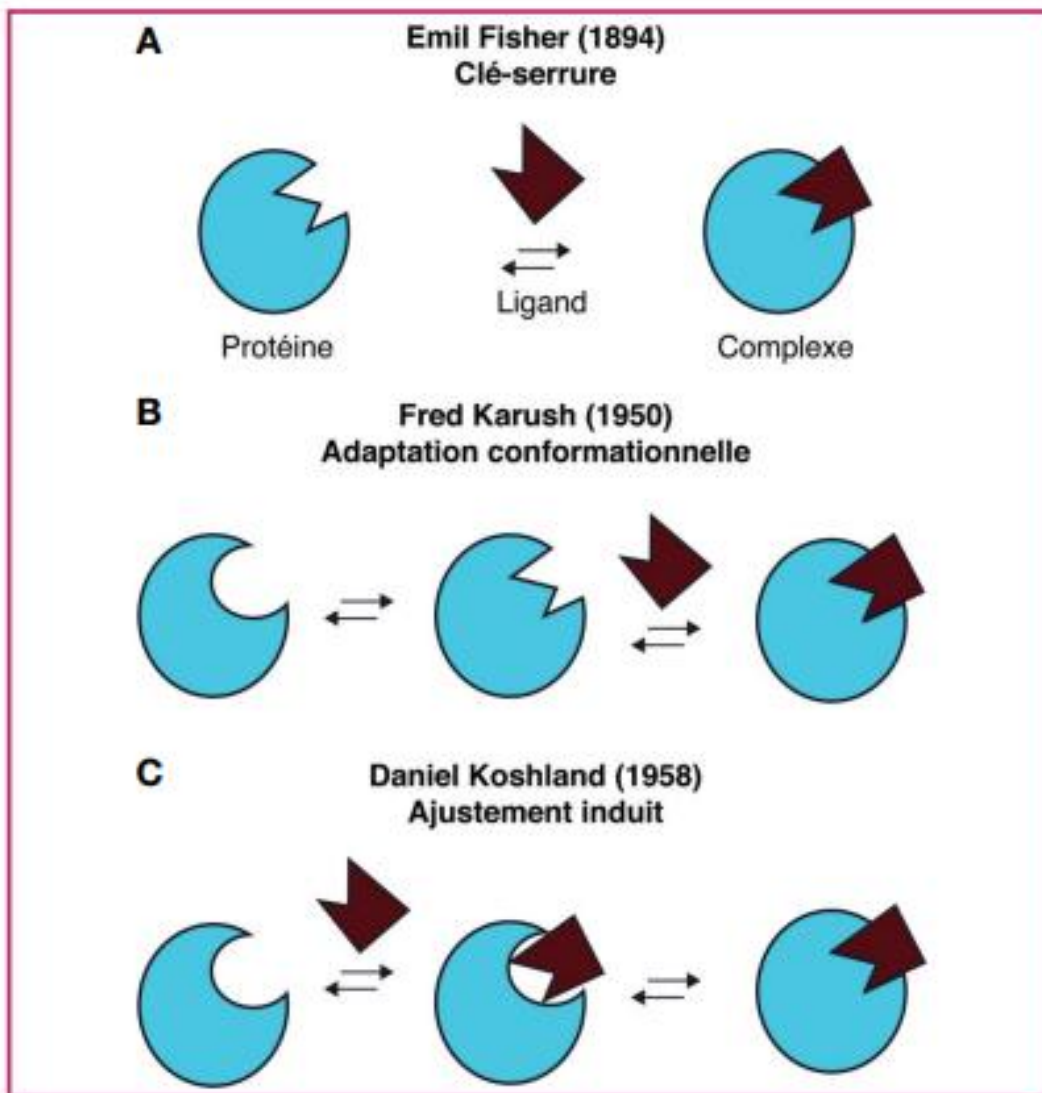


Figure 50 : Illustration des différents modèles d'interaction entre protéine et ligand.

Chapitre V : Interaction protéine ligand et interaction allostérique

III.1 Modèle serrure et clé

En 1894, Emil Fischer, qui étudie les réactions enzymatiques, propose que la liaison du substrat dans le site actif d'une protéine nécessite deux structures tridimensionnelles parfaitement complémentaires. Il introduit alors le concept de « **clé-serrure** » pour expliquer ses observations».

En 1950, Fred Karush est le premier à découvrir que le modèle « clé-serrure » n'est pas généralisable à toutes les protéines et propose l'hypothèse d'« **adaptation conformationnelle** ». Il découvre qu'une protéine, l'albumine, peut se lier avec une grande affinité à de nombreux ligands différents. Il postule alors que la structure tridimensionnelle d'une protéine n'est pas rigide et propose qu'un échange entre plusieurs conformations du site actif de l'albumine ait lieu.

Le modèle de **serrure et clé** a été proposé par (Fischer, 1894) pour expliquer la liaison d'un seul substrat à l'enzyme. Dans cette analogie, la serrure est l'enzyme et la clé est le substrat. La condition préalable de ce modèle est que l'enzyme et le substrat soient à la fois rigides et que leurs surfaces soient complémentaires, et donc seule la clé (substrat) correctement dimensionnée s'adapte dans le trou clé (site actif) de la serrure (enzyme); les clés de taille incorrecte ou mal positionnées (molécules de substrat de forme ou de taille incorrecte) ne rentrent pas dans la serrure (enzyme) car la clé et la serrure ne peuvent pas changer de forme et de taille. Cependant, toutes les preuves expérimentales ne peuvent pas être expliquées de manière adéquate par le modèle «**serrure et clé**». Par conséquent, une modification appelée le modèle d'ajustement induit a été proposée pour compenser la protéine et le ligand rigides, structurellement invariables.

III.2 Modèle ajustement induit

En 1958 Daniel Koshland propose qu'un changement de conformation du site actif de la protéine induit par le ligand ait lieu lors de l'interaction : c'est le modèle dit d'« **ajustement induit** ». Les résultats de Karush et Koshland constituent les premières preuves de l'importance de la dynamique conformationnelle dans la fonction des protéines.

Chapitre V : Interaction protéine ligand et interaction allostérique

Le modèle **d'ajustement induit** suppose que le ligand joue un rôle dans la détermination de la forme finale de la protéine, c'est-à-dire que la liaison du ligand induit un changement conformationnel de la protéine. Cependant, il semble que le mécanisme d'ajustement induit ne convient que pour les protéines présentant un changement conformationnel mineur ou modéré après les liaisons du ligand.

L'ajustement induit n'est possible que si la correspondance entre les sites en interaction est suffisamment forte pour fournir au complexe initial suffisamment de force et de longévité pour que l'ajustement induit prenne lieu dans un délai raisonnable.

Par conséquent, le mécanisme d'ajustement induit seul ne peut pas bien expliquer l'association entre le ligand et les protéines subissant de grands changements conformationnels après la liaison (**Figure 51**).

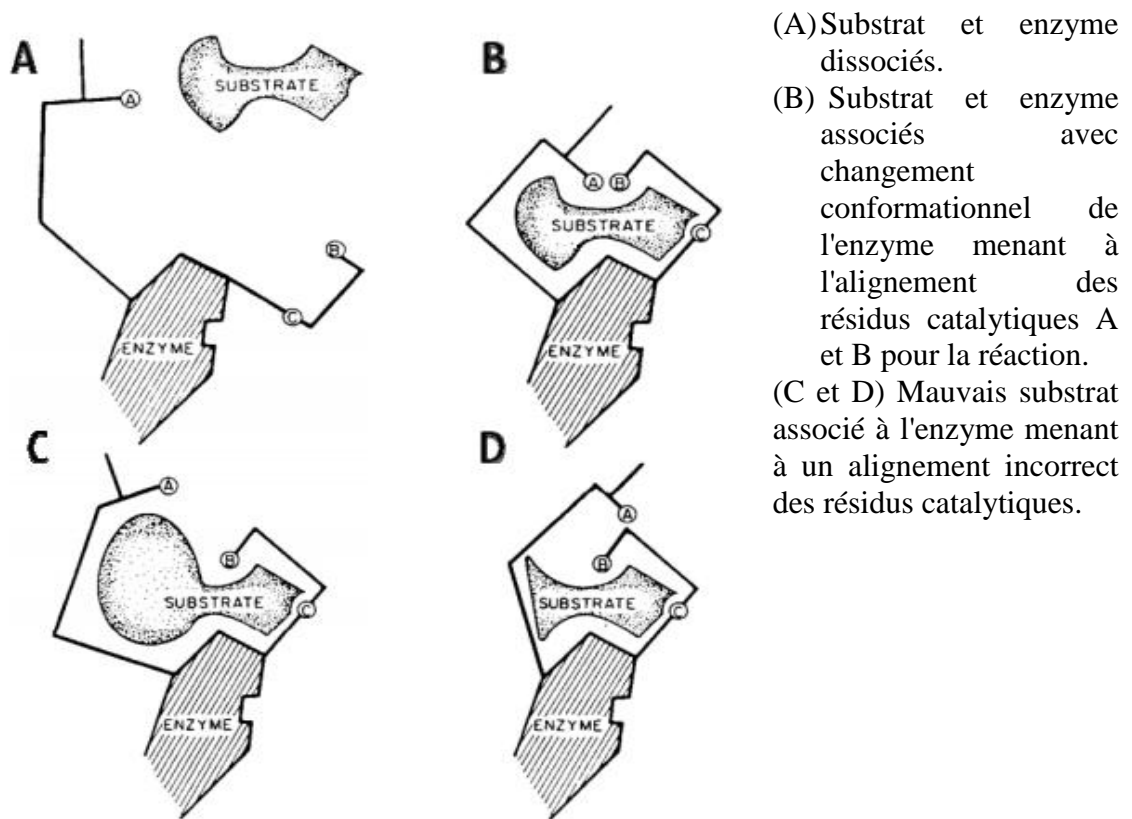


Figure 51 : Représentation du mécanisme de liaison par ajustement induit. Les lignes noires représentent les chaînes de la protéine qui portent les groupements catalytiques **A** et **B** ainsi que le site de liaison **C**.

Chapitre V : Interaction protéine ligand et interaction allostérique

III.3 Modèle de sélection conformationnelle

Le modèle de **sélection conformationnelle** (également appelé équilibre préexistant, sélection de la population, ajustement de la fluctuation ou ajustement choisi), est apparu comme une alternative à l'ajustement induit et devient de plus en plus répandu.

L'état natif d'une protéine n'est pas défini par une seule conformation rigide mais plutôt par un ensemble de sous-produits conformationnels étroitement liés coexistant en équilibre, et le ligand peut sélectionner de se lier au conformère (isomères de conformation) le plus approprié parmi ces sous-états conformationnels, déplaçant l'équilibre vers la formation complexe.

Généralement, les changements de conformation dans ce modèle sont au-delà des réarrangements de chaînes latérales locales près du site de liaison, mais affichent plutôt de grands mouvements concertés ou des mouvements de domaine entiers.



Figure 52 : Représentation du mécanisme de liaison par sélection conformationnelle.

La protéine P est en équilibre entre une conformation inactive (P₁) et une conformation active (P₂) en absence de ligand. La liaison du ligand (L) avec l'état actif (P₂) le stabilise et déplace l'équilibre vers cet état.

IV Interaction allostérique

La liaison de ligand aux protéines par l'intermédiaire des interactions allostériques qui sont des interactions coopératives se produisent lorsque la liaison des ligands à un site spécifique est influencée par la liaison d'un autre ligand appelé **effecteur** ou **modulateur** à un autre site de la protéine.

Si les ligands sont identiques on dit qu'il y a interaction **homotrope** et si ils sont différents on dit qu'il y a interaction **hétérotrope**. Ces effets sont dit **positifs** ou **négatifs** selon que l'effecteur augmente ou diminue l'affinité de la protéine (**Figure 53**).

Chapitre V : Interaction protéine ligand et interaction allostérique

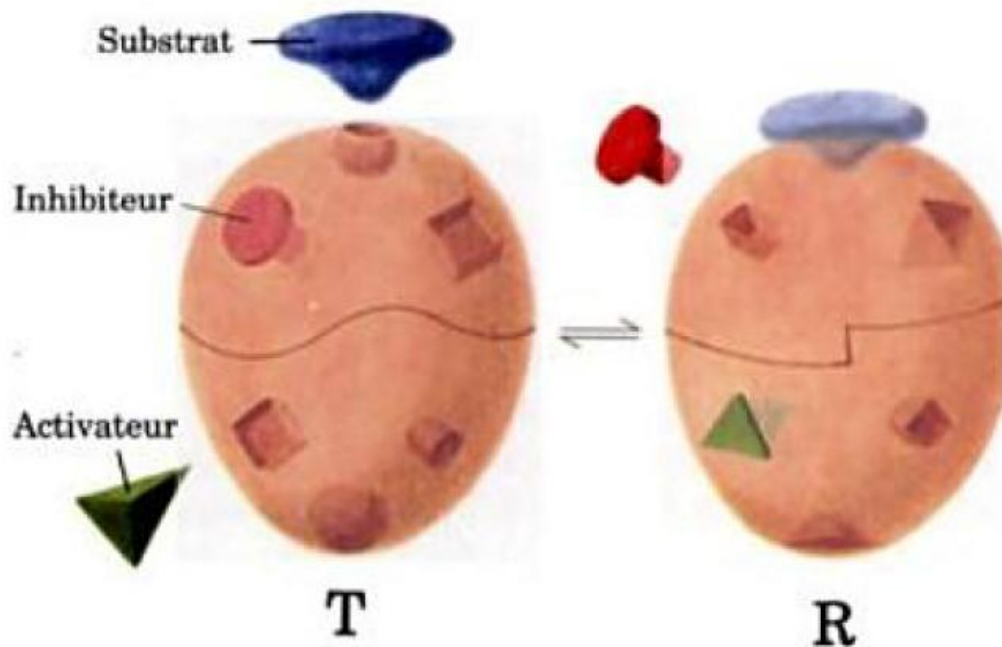


Figure 53 : Les Interactions hétérotropes dans le modèle de l'allostérie. Les effets hétérotropes se manifestent quand substrat et activateur se lient ; L'État R à droite, et substrat et inhibiteur ; l'état T à gauche. La liaison du substrat et de l'activateur à l'oligomère facilite les liaisons ultérieure du substrat et de l'activateur.

Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt

I Synthèse chimique des peptides et des polypeptides

Cette synthèse présente un grand intérêt au plan biologique mais aussi aux plans technologiques. Il est en effet nécessaire d'activer au moins l'un des groupements d'atomes qui participent à la condensation.

La difficulté majeure de cette synthèse réside alors dans le fait que les groupements activés nécessaires à la formation de la liaison peptidique réagissent avec certains des groupements fonctionnels non impliqués dans cette liaison. Il faut donc protéger certaines fonctions des chaînes latérales pour procéder à la synthèse. Il existe de nombreuses méthodes de synthèse comme celle au carbodi-imide. La technique de Merrifield en phase solide repose sur la condensation d'acides aminés dont la fonction amine (et latérale si nécessaire) est protégée et la fonction carboxylique activée.

La possibilité de fabriquer des polypeptides pharmacologiquement actifs offre un intérêt biomédical considérable. L'une des applications les plus prometteuses de la synthèse polypeptidique est la production de vaccins synthétiques.

Les vaccins, qui sont classiquement constitués de virus <tués> (inactivés) ou atténués (<vivants> mais mutés afin de ne pas donner de maladie à l'homme), stimulent le système immunitaire qui synthétise des anticorps dirigés spécifiquement contre ces virus, conférant ainsi l'immunité. L'utilisation des vaccins n'est cependant pas sans risque, de tels problèmes seront éliminés si l'on prépare des vaccins à partir de polypeptides synthétiques qui contiennent les séquences en acides aminés des déterminants antigéniques viraux (groupements moléculaires qui stimulent le système immunitaire pour la synthèse d'anticorps).

Les premiers polypeptides synthétisés chimiquement ne contenaient qu'un seul type d'acide aminé et pour cette raison on les appelle des **homopolypeptides** ex, **polyglycine**, la **polysérine** et la **polylysine**.

La première synthèse d'un polypeptide biologiquement actif a été celle de l'hormone **ocytocine**, un nonapeptide de 9 résidus d'acide aminé en 1953 par Vincent du Vigneaud.

Des améliorations méthodologiques ont permis la synthèse de nombreux polypeptides et plusieurs protéines.

Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt

II Méthodes de synthèse

Les polypeptides sont synthétisés chimiquement en liant par covalence l'un après l'autre, les acides aminés à l'extrémité de la chaîne polypeptidique. Si on synthétise un polypeptide depuis son extrémité C-terminale vers son extrémité N-terminale, cela signifie que la chaîne en formation se termine par un groupement aminé libre.

Donc, le groupement aminé de chaque acide que l'on veut incorporer à la chaîne doit être chimiquement protégé (bloqué), sinon il pourrait réagir avec des molécules. Une fois que le nouvel acide aminé a été couplé, son groupement amine N-terminale doit être déprotégé (débloqué) afin que la liaison peptidique suivante puisse être formée.

Chaque addition d'acide aminé nécessite donc une étape de couplage et une étape de déblocage. Les chaînes latérales réactives doivent être bloquées pour éviter qu'elles ne participent aux réactions de couplage, puis débloquées à la fin de la synthèse.

II.1 La méthode de synthèse en phase solide SPPS

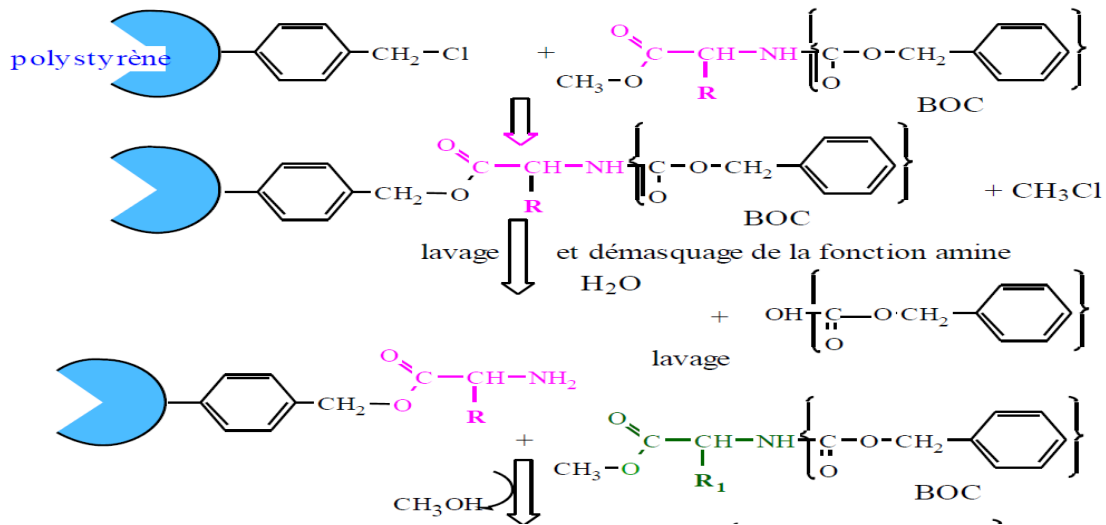
Les premières synthèses de polypeptides telles que celle de l'ocytocine ont été réalisées entièrement en solution, par la mise au point de la **synthèse en phase solide SPPS**. Selon cette méthode la chaîne polypeptidique en formation est ancrée par covalence généralement par son extrémité C-terminale, à un support solide insoluble tel que les résines en polystyrène ensuite les acides aminés convenablement bloqués et les réactifs sont ajoutés dans un ordre approprié. Cela permet la récupération quantitative et la purification des produits intermédiaires par simple filtration et lavage.

Lorsque les chaînes polypeptidiques sont synthétisées par addition d'acides aminés vers leur extrémité N-terminale qui est le sens opposé à la biosynthèse des protéines, le groupement aminé de chaque acide ajouté successivement doit être protégé chimiquement pendant la réaction de couplage. A cet effet on utilise le plus souvent le groupement **Boc** : *tert-butylloxycarbonyl* ou encore le groupement Fmoc : **9-Fluorénylméthoxycarbonyl** (Figure 54).

Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt

Condensation

Elle est schématisée sur le schéma réactionnel indiqué selon Merrifield



Boc : *tert*-butyloxycarbonyl

Protection de la fonction acide, **résine ou support solide**

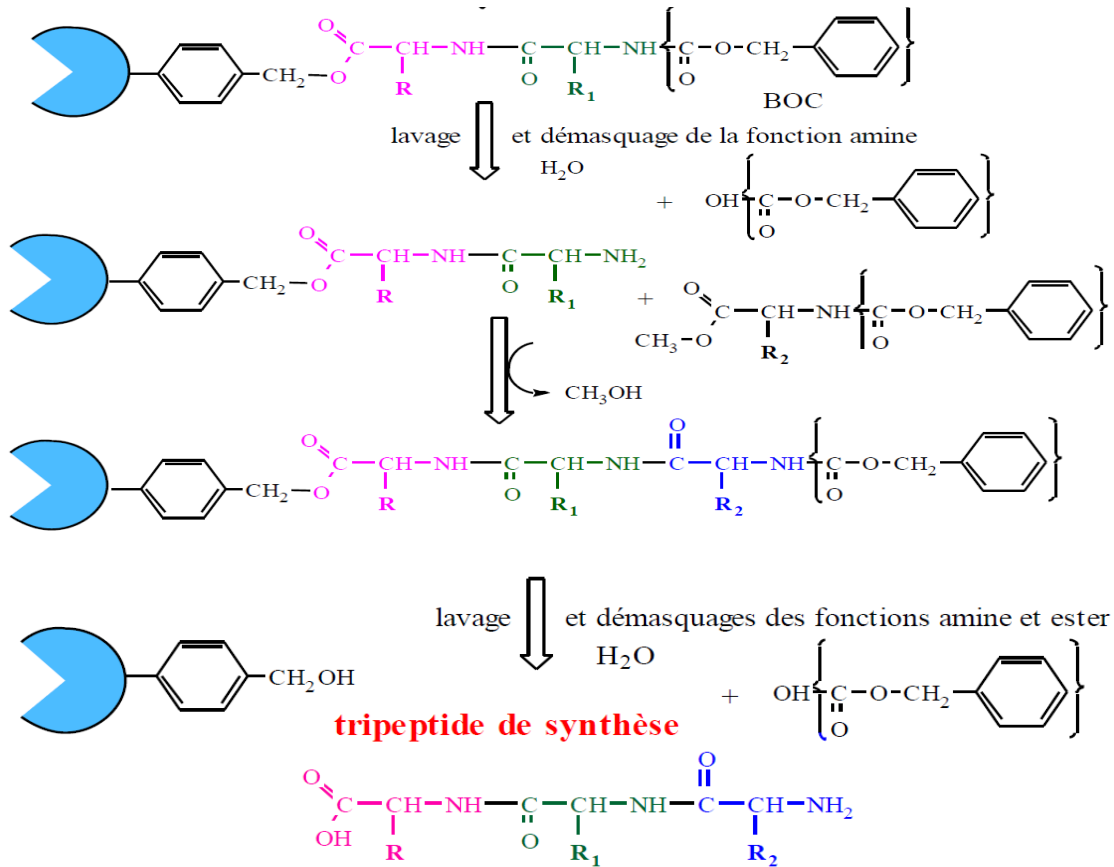


Figure 54 : synthèse peptidique en phase solide selon Merrifield.

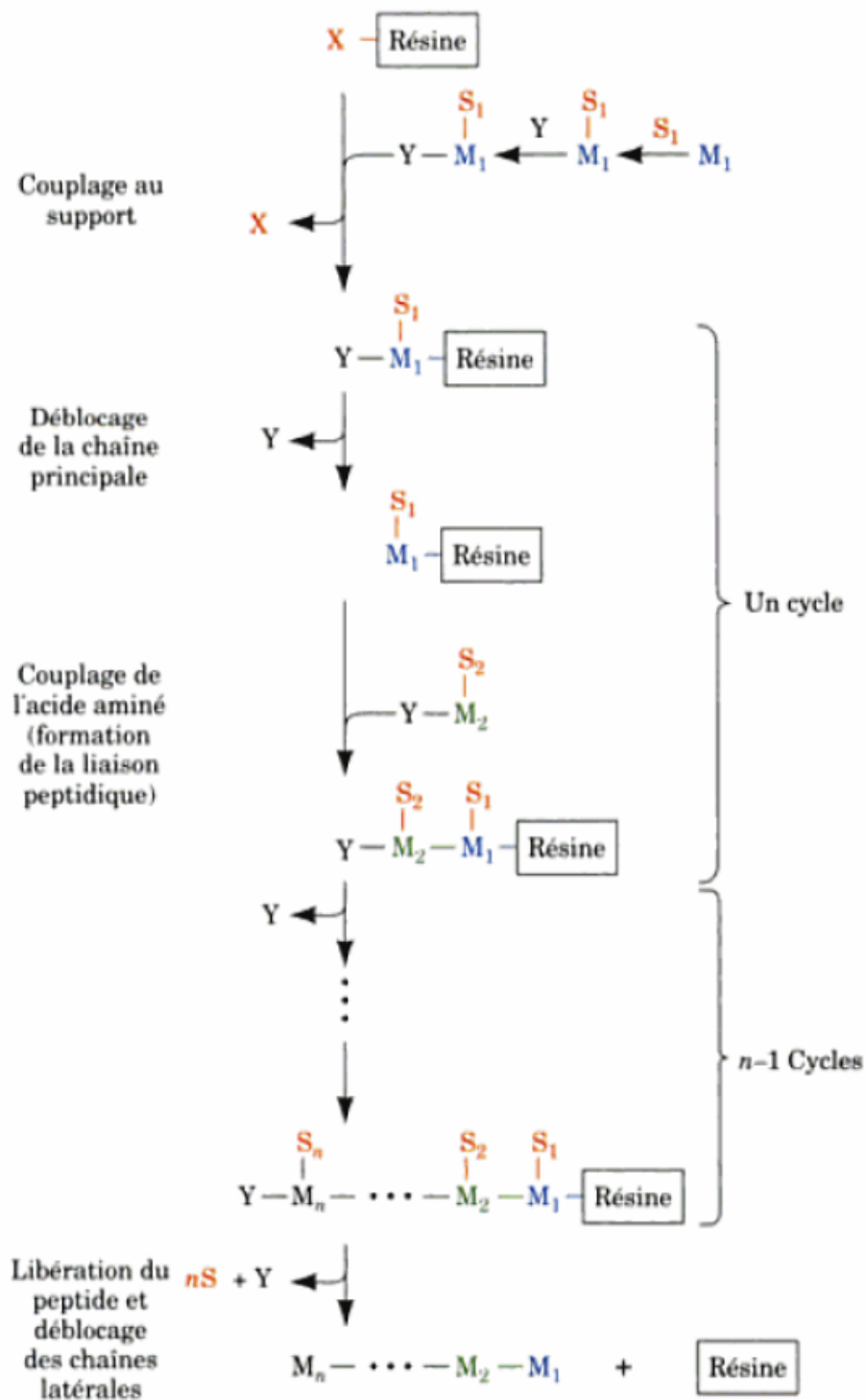
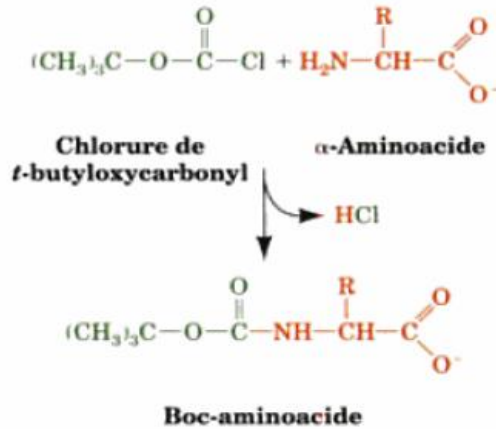


Figure 55 : Réaction pour la synthèse d'un polypeptide en phase solide.
M_i Correspond au résidu d'acide aminé qui doit être ajouté au polypeptide.
S_i Et le groupement protecteur de la chaîne latérale.
Y symbolise le groupement protecteur de la chaîne principale.

II.2 Exigences de la synthèse peptidique

- Protection des fonctions non engagées dans la liaison peptidique
- Activation de la fonction acide pour former la liaison peptidique
- Contrôle de l'épimérisation



Protection des fonctions non engagées dans la liaison peptidique et ancrage de la chaîne au support inerte

Couplage de l'acide aminé avec le groupement Boc

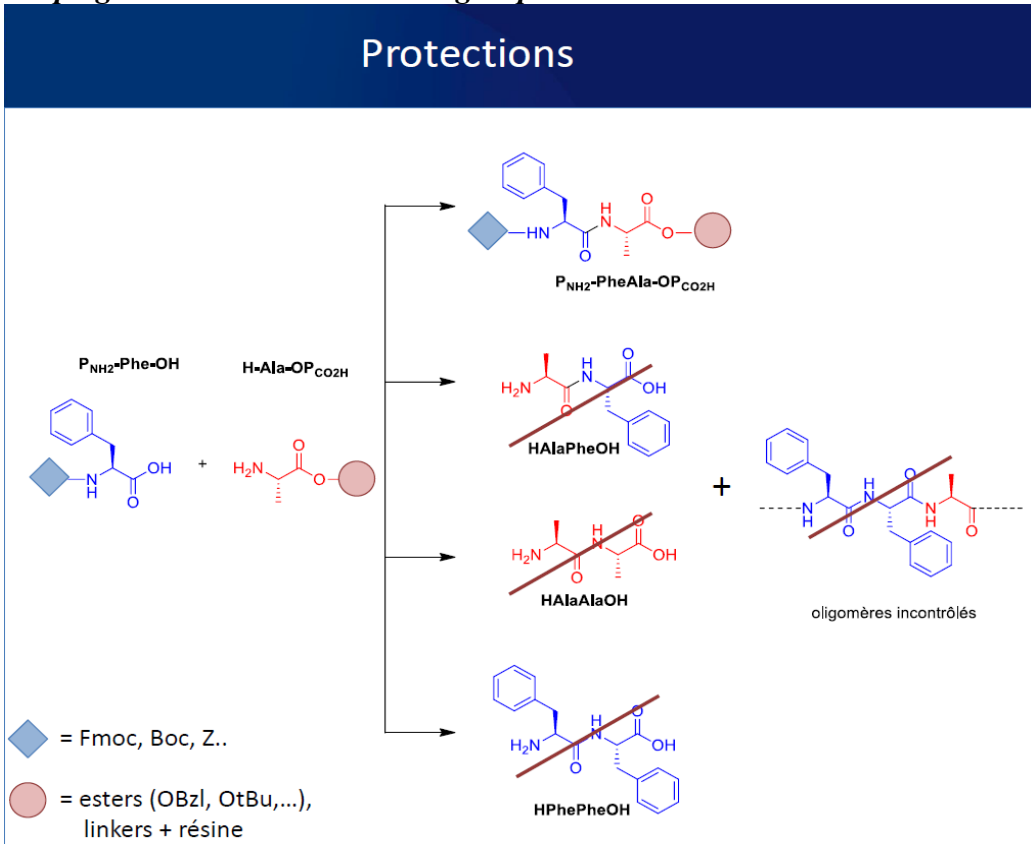
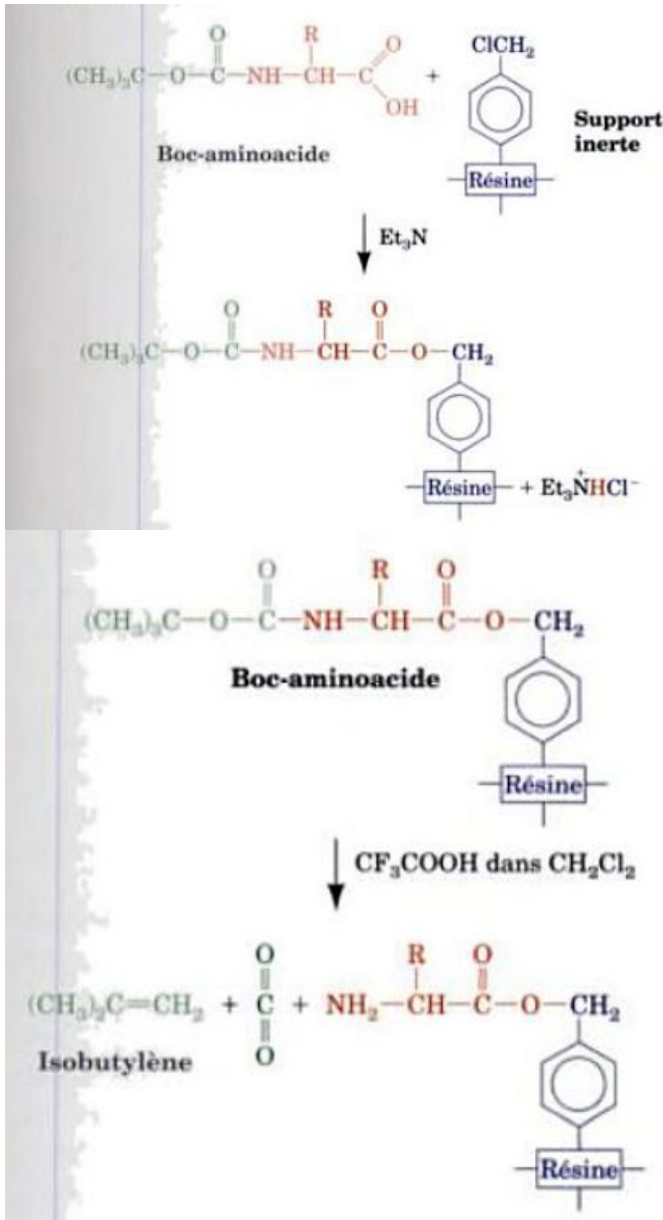


Figure 56 : Protection, activation et contrôle de l'épimérisation dans la méthode SPPS. Le groupement NH2 est protégé par acétylation, la fonction thiol est estérifiée, la fonction guanidique est nitrée etc.

II.3 Ancrage de la chaîne au support inerte

La première étape de la SPPS consiste à coupler l'acide aminé C-terminal au support solide. Le support le plus souvent utilisé est une résine en polystyrène réticulé sur laquelle sont greffés des groupements chlorométhyle. Le couplage à la résine se fait par la réaction suivante, la résine couplée à l'acide aminé est filtrée et lavée.



Le groupement aminé est alors débloquent par traitement avec un acide anhydre comme l'acide trifluoroacétique, qui n'affecte pas la liaison ester à la résine support.

Figure 57 : Réactions d'ancrage dans la méthode SPPS.

La réaction qui permet de coupler deux acides aminés par une liaison peptidique et **endergonique** et doit être activé pour avoir un rendement significatif. L'agent activateur le plus souvent utilisé est une carbodiimide comme le dicyclohexylcarbodiimide (DCCD).

II.4 Activation des groupements fonctionnels

L'activation se fait en milieu basique (amine déprotonée et formation d'un carboxylate préalablement au couplage) (**Figure 58**).

En alternant successivement les réactions de déblocage et de couplage on peut ainsi synthétiser un polypeptide de séquence voulue. La nature répétitive de ces opérations a permis d'automatiser facilement la méthode SPPS. Pendant la synthèse peptidique beaucoup de chaînes latérales doivent être protégées pour éviter qu'elles ne réagissent avec le réactif de couplage. Le groupement Benzyle et le plus souvent utilisé comme groupement de blocage.

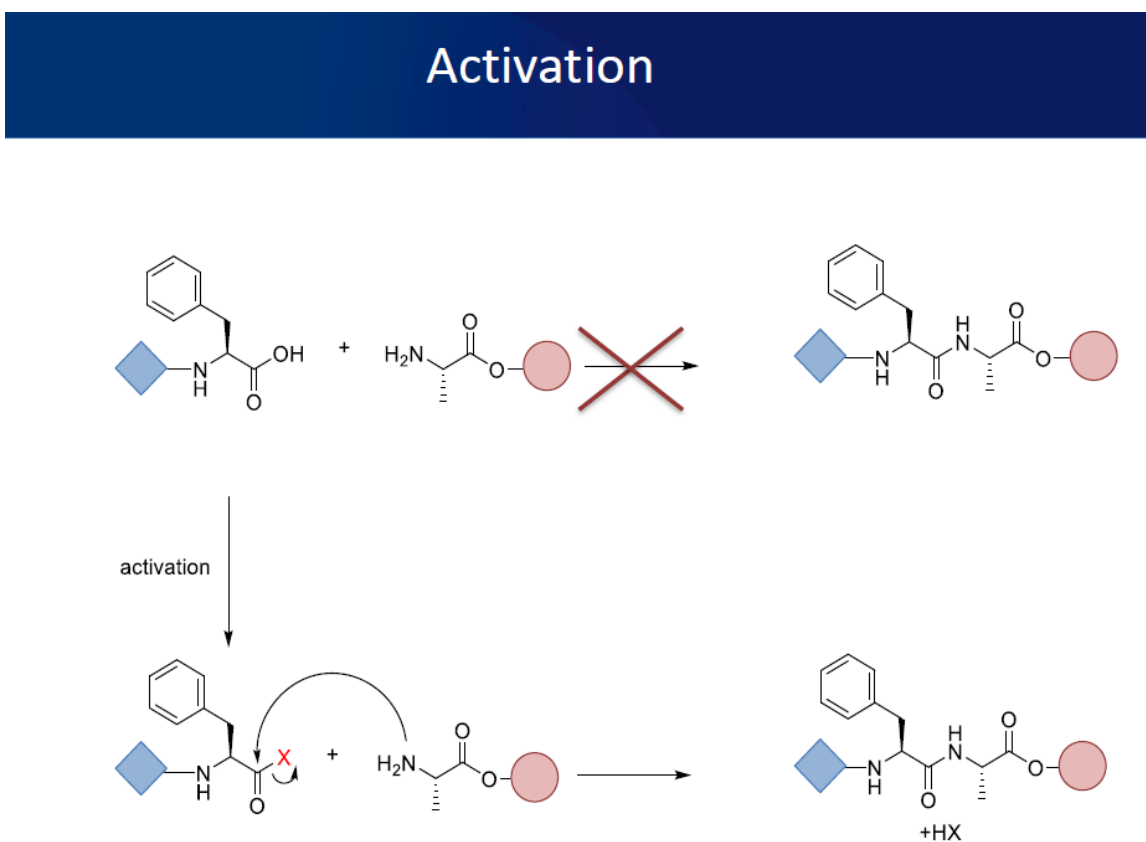


Figure 58 : Activation des groupements dans la méthode SPPS.

X; électroattracteur

Les groupements carboxyliques sont activés par formation d'esters méthyliques et les groupements aminés protégés par formation de dérivés benzyloxycarbonyl .

II.5 Détachement du polypeptide de la résine

La dernière étape de la SPPS est la libération du polypeptide de son support solide. la liaison benzyle Ester qui unit l'extrémité C terminale du polypeptide à la résine peut-être rompu par action de l'acide fluorhydrique liquide (**Figure 59**).

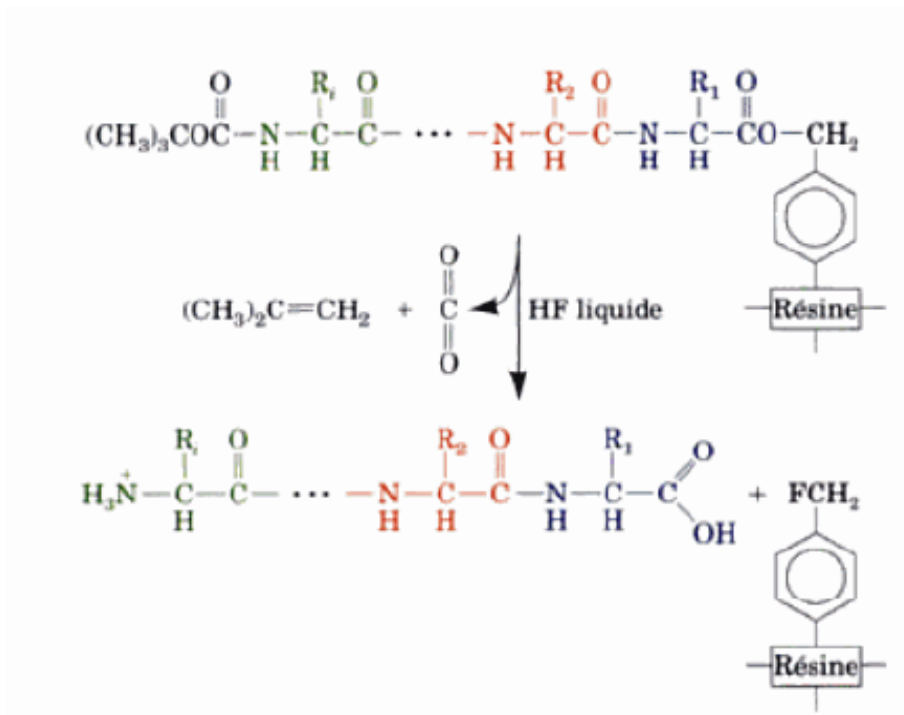


Figure 59 : Détachement du polypeptide de la résine dans la méthode SPPS.

Le groupement Boc lié à l'extrémité N terminale du polypeptide, ainsi que les groupements benzyle qui protègent les chaînes latérales sont aussi enlevés par ce traitement.

Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt

III Synthèse chimique par fragment

Avantages

- plus rapide
- possibilité de purification des intermédiaires
- différence de taille entre segments et produit de condensation : séparation plus facile
- en cas de problème de synthèse, le segment est plus facile à re-préparer.

Inconvénients

- difficulté de solubiliser les longs fragments
- faible concentration (problème de coût) des fragments : réactions lentes rendements bas
- problème d'épimérisation lors de l'activation de l'acide aminé C-terminal.

IV Synthèse et bioproduction des protéines (Tableau 07)

Les systèmes biologiques étant par nature beaucoup plus difficiles à contrôler que les procédés de production chimiques, les rendements variables, ainsi que la pureté et la sécurité, resteront des défis importants associés à ces technologies.

Les protéines recombinantes sont produites à partir d'un transgène introduit dans un organisme hétérologue par génie génétique. Ce procédé biotechnologique se déroule suivant cinq grandes étapes :

- 1. la construction d'un vecteur d'expression (en général, plasmide ou virus) jouant le rôle de transporteur du gène codant pour la protéine d'intérêt ;**
- 2. la modification d'une cellule hôte ou d'un organisme complexe par le vecteur d'expression ;**
- 3. la production de ce nouveau système d'expression, qui exécute les instructions fournies par le transgène, en l'occurrence synthétise la protéine d'intérêt ;**
- 4. la séparation de la protéine du milieu de culture, quand elle est sécrétée, ou son extraction du milieu intracellulaire le cas échéant, puis sa purification ;**
- 5. la caractérisation de la molécule et la vérification de son degré de pureté.**

Chapitre VI : Ingénierie des protéines et des polypeptides d'intérêt

Tableau 07 : Système de production des protéines.

	Bactéries	Levures	Cellule d'insectes + baculovirus	Cellules animales (cellules CHO)	Plantes transgéniques	Animaux transgéniques
Niveau de production théorique	+++++	+++++	+++	+	+++++	+++++
Niveau de production pratique	++(+)	++(+)	+	+	++	++++
Coût d'investissement	+++++	+++++	++	+	++++	+++
Coût de production	+++++	+++++	++	++	+++++	++++
Flexibilité	+++++	+++++	++	+	+++++	++++
Conservation	+++++	+++++	+++	+++	+++++	+++++
Stabilité	+++++	+++++	++++	+++	+++++	+++++
Délai pour la 1 ^{re} production	+++++	+++++	+++	+++++	++++	+++(+)
Augmentation	+++++	+++++	++	+	+++++	++++
Recueil	+++++	+++++	+++++	+++++	+++++	++++
Effets sur l'organisme	+++(+)	+++(+)	+++(+)	+++(+)	+++(+)	+++
Modifications post-translationnelles	+	++	+++	++++	+++	++++
Glycosylation	+	++	+++	++++	++	++++
Stabilité du produit	+++++	+++++	+++	+++	+++++	++++
Purification	+++	+++	+++	++++	+++	+++
Pathogènes contaminants	+++++	+++++	+++++	++++	+++++	++++
Propriété intellectuelle	++++	+++	+++	++	+++	+++
Produits sur le marché	++++	+++	+++	+++++	+	+++

Références bibliographiques

Biochimie de Harper. David A Bender, Kathleen M Botham, Peter J Kennelly. 6e Édition, Septembre 2017 deBoeck Supérieur. ISBN-13 9782807307247.

Biochimie 7E ED LAVOISIER. John L. Tymoczko, Jeremy M. Berg, Lubert Stryer. 7e Édition 29/01/2013 Médecine Sciences Publications. ISBN- 9782257204271.

Biochimie. Donald Voet, Judith G Voet. 3e Édition, Septembre 2016 deBoeck Supérieur. ISBN-13 9782804171018.

Biochimie métabolique. P Souetre. Éditions Pradel, 2003, 1e Édition. ISBN- 2-913996-24-8. ISSN 1628-4666.

Biochimie structurale. Claude Audigié, François Zonszain. Doin éditeurs, 2007. ISBN 2-7040-0655-5.

Biochimie maxi fiches. Françoise QUENTIN, Paul-françois GALLET, Michel GUILLOTON, Bernadette QUINTARD. DUNOD, 2e Édition, 2015. ISBN 978-2-10-073922-6.

Nora Benhabilès, Annick Thomas , Robert Brasseur. Les mécanismes de repliement des protéines solubles. *Biotechnol. Agron. Soc. Environ.* 2000 **4** (2), 71–81.

Goldberg, M. E. (2005). Le repliement des protéines : seconde traduction du message génétique. *M/S : médecine sciences*, 21(6-7), 563–566.

Yon-Kahn, J. (2005). Repliement des protéines : études *in vitro*. *M/S : médecine sciences*, 21(6-7), 601–607.

Gregory A Petsko, Dagmar Ringe. Structure et fonction des protéines, 1re édition, novembre 2008. deBoeck Supérieur. ISBN-13 9782804158880.

Références bibliographiques

N.A. Bushmarina, C. Blanchet, G. Vernier et V. Forge. Repliement des protéines : exemple de l' α -lactalbumine. *J. Phys. IV France* **130** (2005) 209–228.

Jean-Michel Betton, Alain Chaffotte. Repliement et production de protéines recombinantes. *MEDECINE/SCIENCES* 2005 ; 21 : 613-7.

Joëlle Dumas, Bénédicte Robert. Bioproduction de protéines thérapeutiques
Revue et perspectives. *MEDECINE/SCIENCES* 2009 ; 25 (hors série n° 2) : 18-26.