

## STRUCTURE ET FONCTIONS DES PROTEINES

### PROTEINES

**Introduction :** Les protéines sont des polymères d'acides aminés, formés généralement par l'enchaînement de quelques centaines d'acides aminés. Elles sont complexes et présentent une grande diversité. Dans notre organisme il y'aurait environ 100000 protéines différentes, par leurs structures et leurs fonctions, dont moins de 5% ont été étudiées et caractérisées.

Les protéines représentent le principal composant des cellules avec plus de 30% de leur poids sec. Elles sont essentielles voir même vitales pour tous les organismes vivants. Elles sont spécifiques à chaque espèce vivante et à chaque organe et même à chaque type cellulaire. Les protéines ont des poids moléculaires très variables allant de quelques kilos Daltons à quelques centaines de kilos Daltons.

Selon leur forme structurale, les protéines peuvent être divisées, en 2 principaux groupes, les **protéines globulaires** (solubles) qui sont les plus nombreuses dans la cellule, avec une structure compacte et les **protéines fibreuses** qui sont des molécules allongées. On parle aussi de protéines fonctionnelles (transport, signalisation, catalyse, défense ...) et de protéines de structure (protéines fibreuses).

#### \*\*\*Protéines fibreuses

Principalement, le collagène, la kératine, le fibrinogène et les protéines musculaires.

**Le collagène :** une protéine très abondante chez les vertébrés. Il se trouve principalement dans l'os, la peau, les cartilages...etc.

**La kératine :** une protéine très stable et bien résistante, chargée de la protection de l'organisme contre le milieu extérieur. Elle se trouve dans les couches supérieures de l'épiderme, les cheveux, les ongles, les écailles, les plumes ... etc.

**Les protéines musculaires :** la myosine, l'actine, l'élastine et l'actomyosine sont les meilleurs exemples de cette catégorie.

**Le fibrinogène :** la protéine plasmatique responsable de la coagulation du sang.

### \*\*\*Protéines globulaires

Contrairement aux protéines fibreuses, les protéines globulaires sont sphériques et hautement solubles. Elles jouent un rôle important dans les différents métabolismes. Les enzymes, les hormones et les albumines sont de bons exemples.

**Enzymes** : protéines capables de catalyser des réactions métaboliques avec des vitesses extrêmement rapides.

**Hormones protéiques** : assurent le contrôle de certains processus très importants, exemple insuline/glucagon.

**Anticorps** (immunoglobulines) : présents dans le sang, ils neutralisent et facilitent la dégradation des antigènes.

**Microtubules** : protéines globulaires qui s'assemblent en des tubes creux très fins. Elles servent comme squelette des cellules et comme moyen de transport intracellulaire.

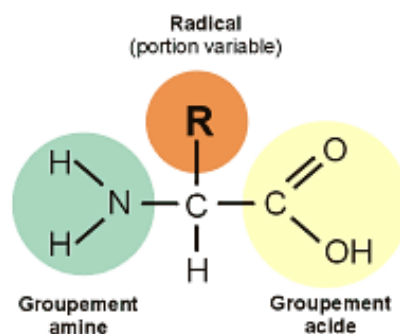
## 1. Acides Aminés

Les acides aminés sont formés d'un carbone alpha central ( $C\alpha$ ) auquel sont liés un groupement amine ( $NH_2$ ), un groupement carboxyle ( $COOH$ ), un hydrogène ( $H$ ) et un radical **R** très variable. C'est ce radical **R** qui crée la différence entre les acides aminés (**Fig.1**). Dans la nature, on trouve plus d'une centaine d'acides aminés, dont 20 seulement rentrent dans la structure des protéines (**Fig.2**)

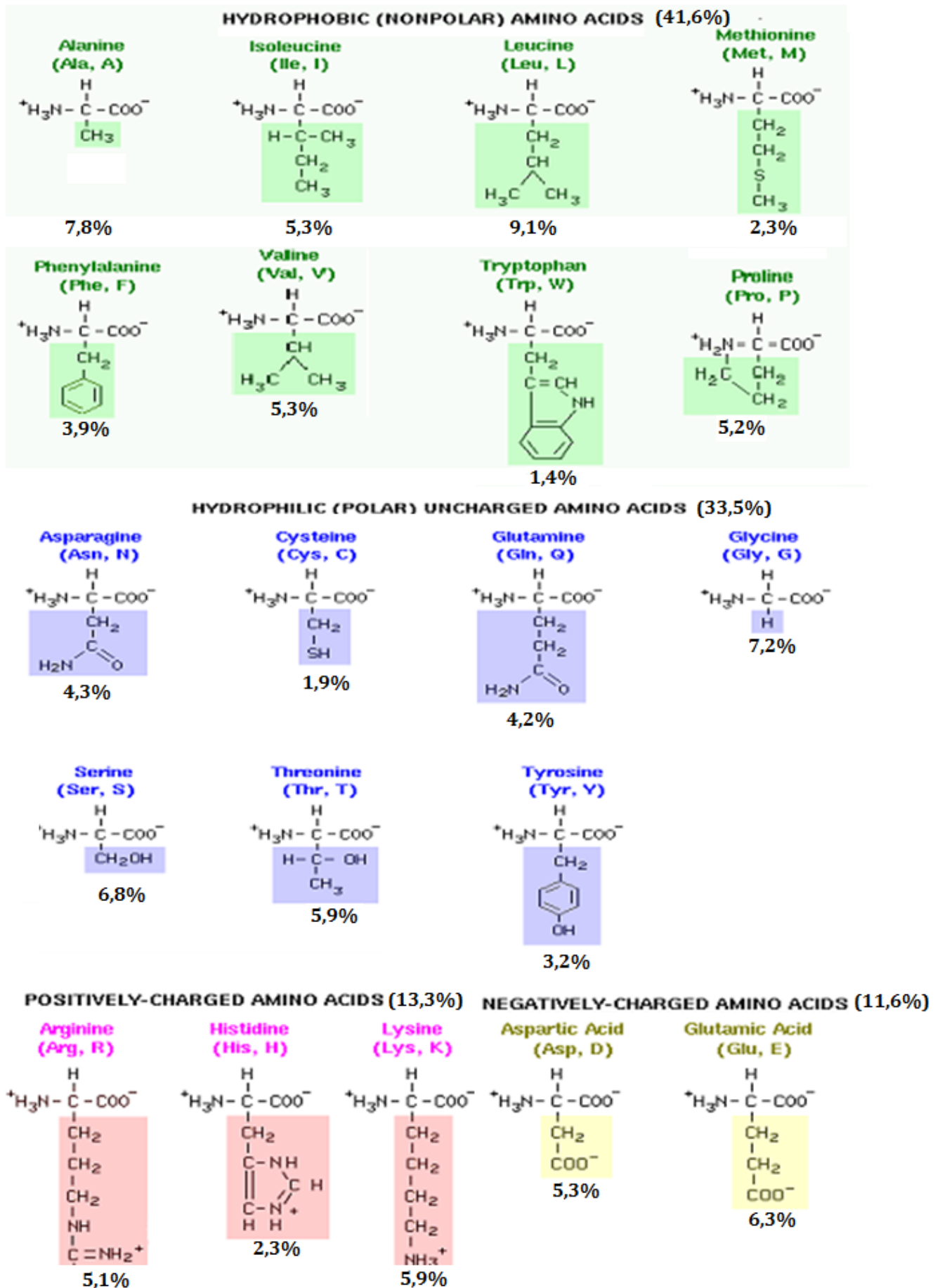
\*En solution, le groupement  $NH_2$  est sous la forme  $NH_3^+$  et le groupement  $COOH$  sous la forme  $COO^-$ .

Sur la base de la polarité de R, les acides aminés sont classés en 4 groupes (**Fig. 2**).

- Acides aminés non polaires (hydrophobes)
- Acides aminés polaires non chargés (hydrophiles)
- Acides aminés chargés positivement (basiques, hydrophiles)
- Acides aminés chargés négativement (acides, hydrophiles)



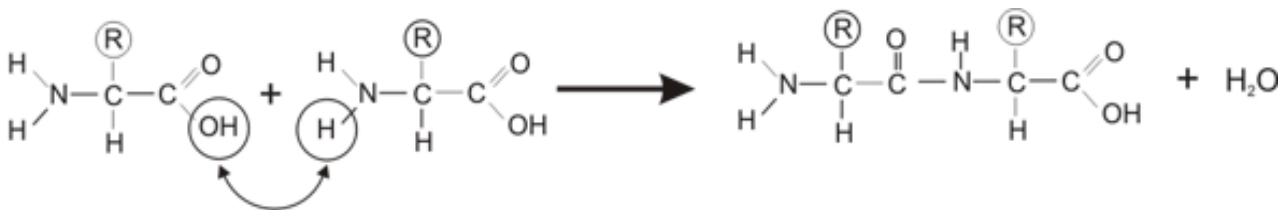
**Fig.1:** Structure d'un acide aminé



**Fig.2:** Les 4 groupes d'acides aminés classés selon la polarité du radical R, avec leur taux de présence de chaque acide aminé et chaque groupe dans les protéines humaines.

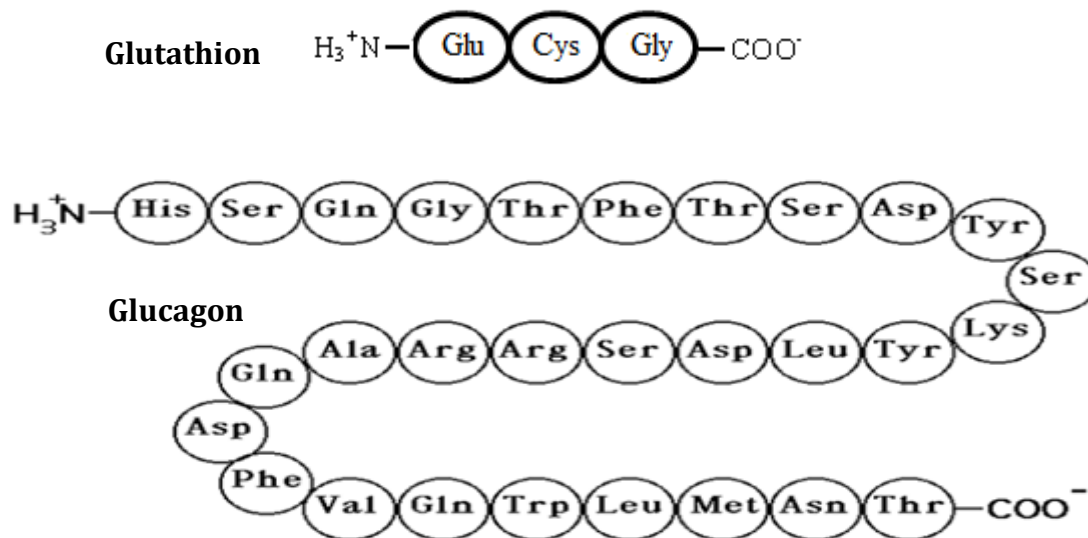
## 2. Peptides

La liaison entre 2 acides aminés est une liaison covalente, forte dite peptidique. Elle se forme entre le groupement COOH d'un acide aminé et le groupement NH<sub>2</sub> d'un autre acide aminé, avec libération d'une molécule d'eau (**Fig. 3**). L'enchainement de plusieurs acides aminés forme un **peptide** (petite chaîne de moins de 50 acides aminés) ou un **polypeptide** (longue chaîne). La protéine est le polypeptide ayant une fonction.



**Fig.3:** Formation d'une liaison peptidique

La taille des polypeptides naturels est très variable, elle passe de 3 acides aminés chez le glutathion jusqu'à quelques dizaines chez le glucagon (**Fig.4**) et même à quelques centaines d'acides aminés chez les macromolécules protéiques.



**Fig.4:** Structure primaire du glutathion (3 aa) et du glucagon (29 aa).

Il faut noter que plus les espèces se rapprochent, plus leurs protéines se ressemblent et inversement. Par exemple, le lysozyme humain (130 aa) est identique à celui du chimpanzé, mais il diffère de celui du poulet (voir TD: exo1-3).

**Voir 1<sup>ère</sup> série de TD**

### 3. Structure des protéines

Les protéines ont une structure primaire, qui n'est que la séquence linéaire des acides aminés. Mais la présence d'un grand nombre de forces d'attraction et de répulsion font que la protéine ne peut jamais garder cette forme linéaire. En effet, **les liaisons hydrogène, les ponts disulfures, les charges positives et négatives, les radicaux hydrophobes et hydrophiles, les forces de Van der Waals...**, imposent à la molécule généralement une structure secondaire soit en hélice  $\alpha$  ou en feuillet  $\beta$  ou les 2 ensembles ou une forme en pelote. Les molécules deviennent encore plus compactes si elles adoptent une structure tertiaire.

Lorsqu'une protéine est constituée de sous unités, on parle dans ce cas d'une structure quaternaire.

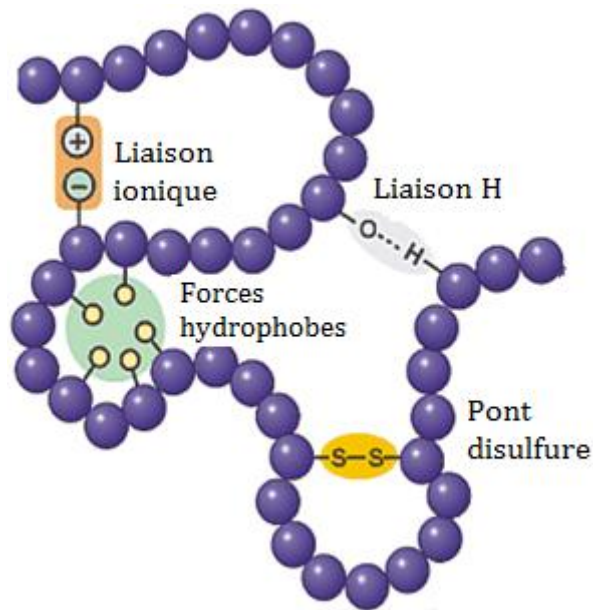
Il est clair que les protéines ont 4 niveaux d'organisation structurale: structure primaire, secondaire, tertiaire et quaternaire.

#### 3.1. Structure Primaire

La structure primaire c'est le nombre et l'ordre (séquence) du premier au dernier acide aminé constituant le polypeptide. Cette séquence s'écrit de gauche à droite de l'extrémité NH<sub>2</sub> (N-terminale) vers l'extrémité COOH (C-terminale). Cette structure primaire, est d'une grande importance, car c'est elle qui dicte d'une manière ou d'une autre la structure tridimensionnelle de la molécule et par conséquent sa fonction.

Les chaînes latérales (radicaux R) des acides aminés ont des structures et des propriétés chimiques différentes (hydrophobie, hydrophilie, charge négative ou positive...). Par ces propriétés, ces radicaux R forment différents types de liaisons qui vont aboutir à la formation de la structure secondaire.

Dans une protéine, on trouve des liaisons covalentes, principalement la liaison peptidique et occasionnellement la liaison disulfure (S-S) et les liaisons non-covalentes (liaison hydrogène, liaison hydrophobe, liaison ionique et forces de Van der Waals... ) comme les montre la **figure 5**.



**Fig.5:** Forces non covalentes participant à l'édification de la structure des protéines

**A compléter par** (travail personnel)

*\*Liaisons hydrogène*

*\*Forces hydrophobes*

*\*Liaisons ioniques*

*\*Forces de Vander Waals*

### **Détermination de la structure primaire des protéines (Séquençage)**

Le séquençage des protéines permet de déterminer l'enchaînement des acides aminés, cet enchaînement est essentiel pour le maintien de la structure tertiaire de la molécule et par conséquent, le maintien de la fonction de la protéine. Le séquençage se fait par la méthode d'Edman, qui consiste en la dégradation séquentielle des acides aminés à partir de l'extrémité N terminale, ce qui permet de les identifier l'un après l'autre. Schématiquement, cette méthode se déroule en 4 étapes, comme le montre la **figure 6**.

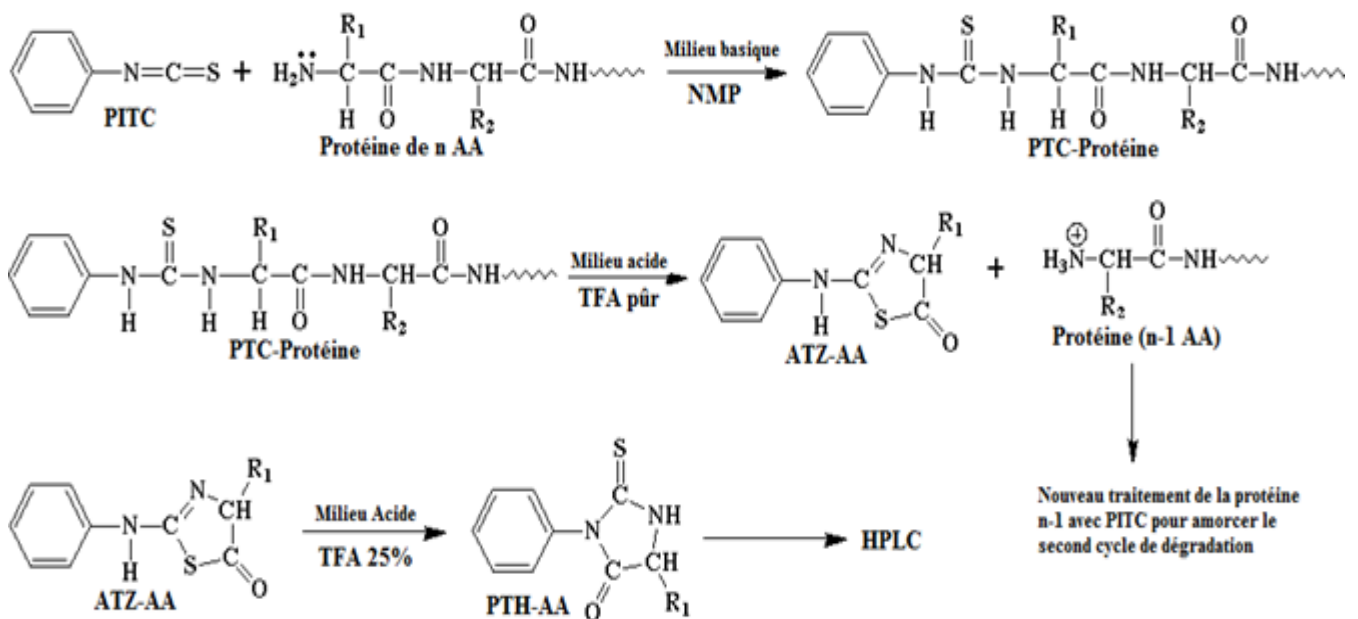
**1<sup>ère</sup> étape : Couplage (Dérivatisation):** Le Phényl-Iso-Thio-Cyanate (PITC) se couple à la fonction amine terminale de la protéine en présence de la base N-méthyl pipéridine (NMP) pour donner le PhenylThioCarbamyl-Protéine (PTC-Protéine).

**2<sup>ème</sup> étape : Coupure:** Ce couplage fragilise la liaison peptidique terminale, ce qui va faciliter sa coupure par l'acide trifluoroacétique (TFA pur) et d'une manière spécifique, car les autres liaisons peptidiques restent intacts. Cette action génère

l'AnilinoThiaZolinone (ATZ) de l'acide aminé N°1 instable, ainsi, la protéine perd son 1<sup>er</sup> AA, il reste donc une protéine à n-1 acides aminés. Le ATZ-AA instable sera transformé en un produit stable durant la 3<sup>ème</sup> étape

**3<sup>ème</sup> étape : Conversion:** Le ATZ-AA est converti en milieu acide (TFA à 25 %) en PhénylThioHydantoïne-AA (PTH-AA) stable.

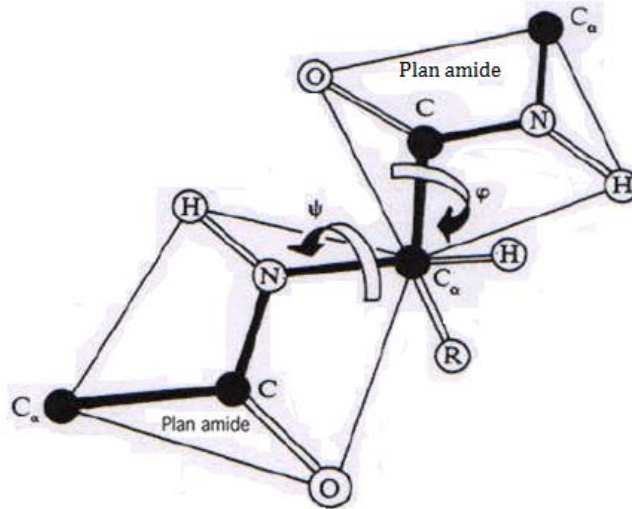
**4<sup>ème</sup> étape: chromatographie:** Le PTH-AA va être analysé par HPLC et la détection se fera à 280nm. Une gamme d'étalonnage PTH-AAs (tous les aa) permet de déterminer la nature de chaque PTH-AA obtenu. Le cycle de réaction est répété autant de fois jusqu'à ce que la protéine soit totalement épuisée et c'est ainsi que la séquence sera déterminée.



**Fig. 6:** Séquençage des protéines par la méthode d'Edman

### 3.2. Structure secondaire

La structure secondaire est l'ensemble des conformations régulières et répétitives adaptées par un polypeptidique. La conformation du squelette peptidique est décrite par les positions respectives des plans peptidiques qui la constituent. La position des plans peptidiques de part et d'autre du carbone  $\alpha$  est caractérisée par la valeur des angles de rotation d'un plan amide par rapport au carbone  $\alpha$ . Ces plans peptidiques peuvent pivoter de  $360^\circ$  autour des liaisons N- $C_\alpha$  et C- $C_\alpha$ . L'angle  $\varphi$  (phi) correspond à l'angle de rotation N- $C_\alpha$  et l'angle  $\psi$  (psi) correspond à l'angle C- $C_\alpha$  ces 2 angles peuvent varier de  $+180^\circ$  à  $-180^\circ$  (**Fig. 7**). Le point zéro correspond à la position où les plans peptidiques de part et d'autre du  $C_\alpha$  sont dans le même plan.



**Fig.7:** Plan amide avec l'angle  $\varphi$  et l'angle  $\psi$ . L'angle de rotation d'un plan amide par rapport à  $C_{\alpha}$  du côté CO est noté  $\psi$  et du côté N avec le même  $C_{\alpha}$  est noté  $\varphi$ . Les angles  $\varphi$  et  $\psi$  sont 2 angles entre 2 plans on les appelle des angles dièdres.

Les angles possibles et les structures qu'ils engendrent le plus souvent sont représentés sur la table et le diagramme de **Ramachandran (Fig. 8)**. Les couples  $\psi\varphi$  qui conduisent à des configurations stables vont se répéter pour donner une structure régulière de type hélice alpha ou feuillet beta qui forment la **structure secondaire**.

### 3.2.1. Hélice $\alpha$

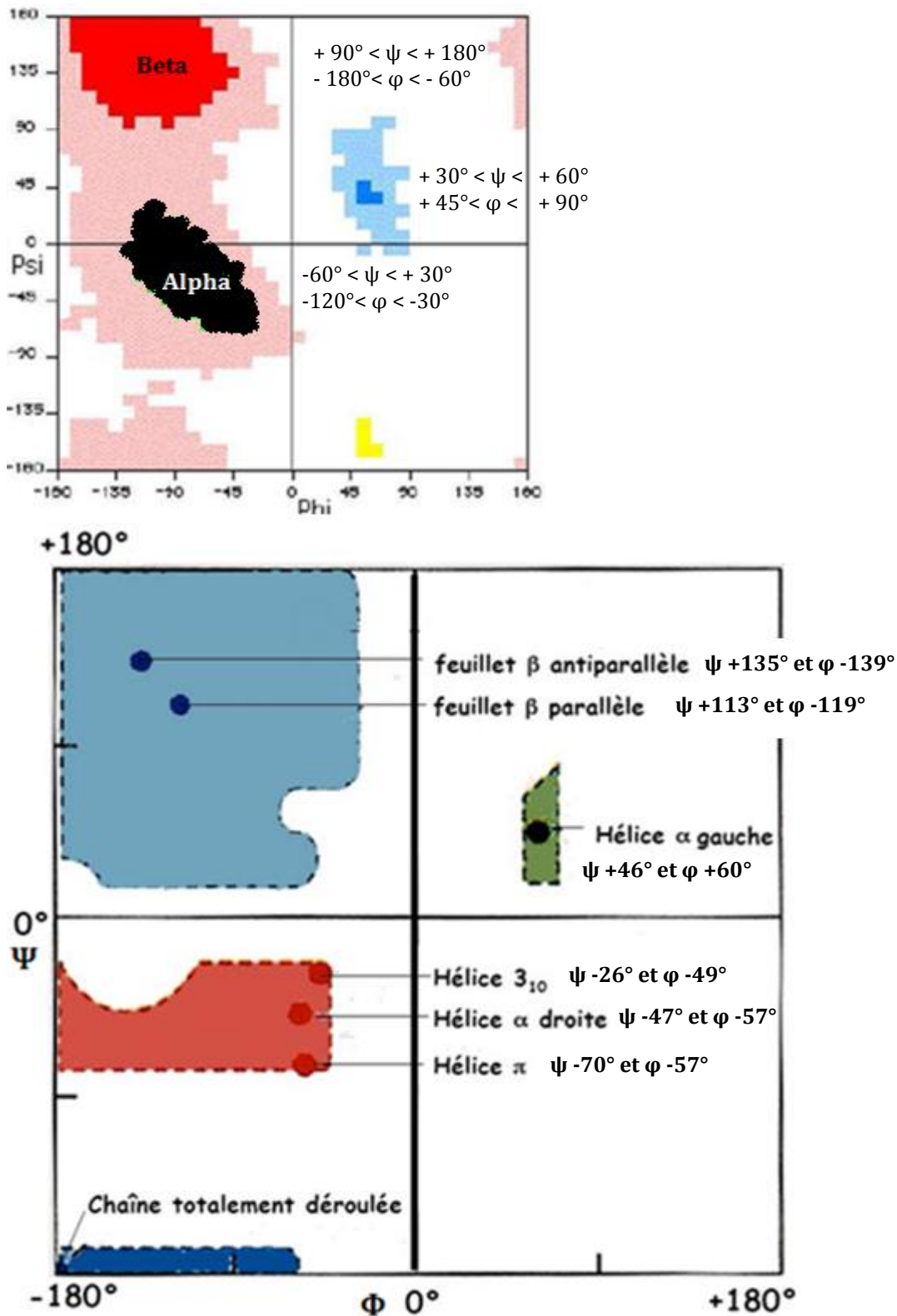
La structure secondaire la plus fréquente est l'hélice  $\alpha$ , dans laquelle la chaîne polypeptidique s'enroule autour d'elle-même en hélice droite de telle sorte que les chaînes latérales (**R**) seront orientées vers l'extérieur avec un pas (tour) de 3,6 aa (**Fig.9**) Cette structure est stabilisée par une liaison hydrogène s'établissant entre l'oxygène carbonyle et l'hydrogène de l'amide qui se trouve à 4 résidus plus loin. Donc c'est entre le CO de l'aa N°1 et l'amine de la liaison peptidique qui précède l'aa N°5 et ainsi de suite (on retient que la liaison H se fait donc entre aa N° i et aa N° i+4).

Cette organisation en hélice  $\alpha$  correspond à la structure la plus stable des polypeptides. Elle se forme quand  $\varphi = -57^{\circ}$  et  $\psi = -47^{\circ}$

La formation de l'hélice  $\alpha$  dépend étroitement de la structure primaire, car certains acides aminés favorisent sa formation et d'autres la défavorisent. Par exemple, l'alanine ( $R = CH_3$ ) est très fréquente et très favorable à la formation de l'hélice  $\alpha$ , par contre, la proline et les acides aminés basiques et les acides aminés acides sont défavorables à la formation de l'hélice  $\alpha$ .

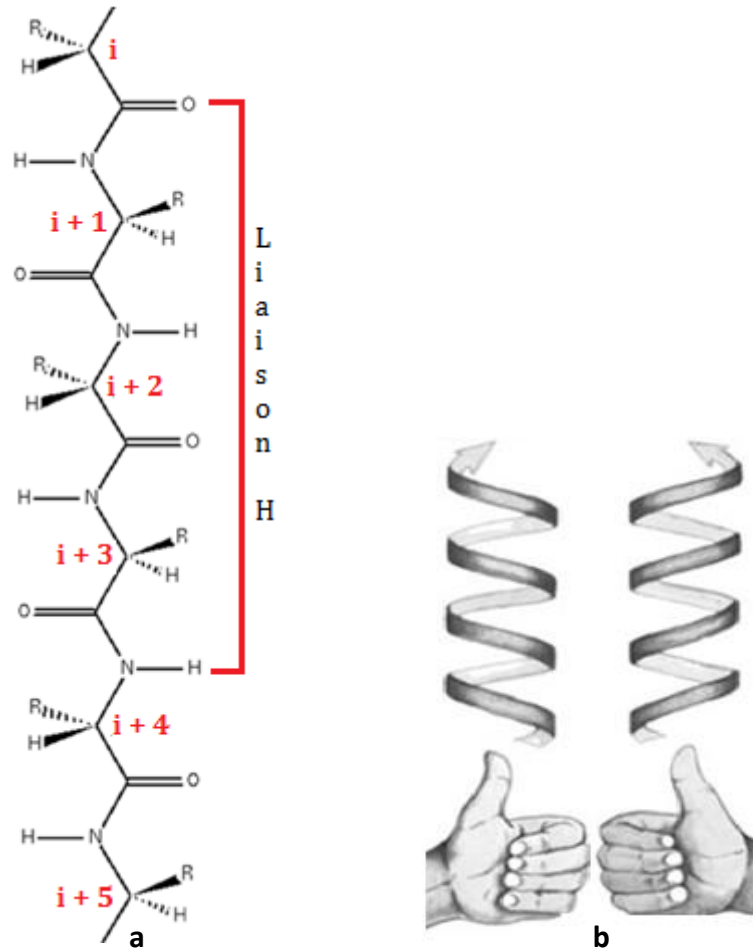


A compléter par (travail personnel) Hélice  $3_{10}$  Hélice  $3,6_{13}$  Hélice  $\pi$  Hélice gauche



**Fig.8:** Diagramme de Ramachandran: c'est une représentation graphique des conformations possibles de la chaîne polypeptidique. Pour chaque aa d'une protéine, on porte la valeur de l'angle  $\varphi$  (Phi) en abscisse et celle de l'angle  $\psi$  (Psi) en ordonnée, pour des valeurs de  $-180$  à  $+180$ . L'analyse de la structure des protéines, montre que la majorité des acides aminés ont des coordonnées  $(\varphi, \psi)$  qui les situent dans 3 zones dites favorables, à savoir les 2 principales

régions correspondant aux structures secondaires régulières des protéines (hélices  $\alpha$  et feuillets  $\beta$ ) et une troisième région, plus petite correspondant à l'hélice  $\alpha$  gauche.

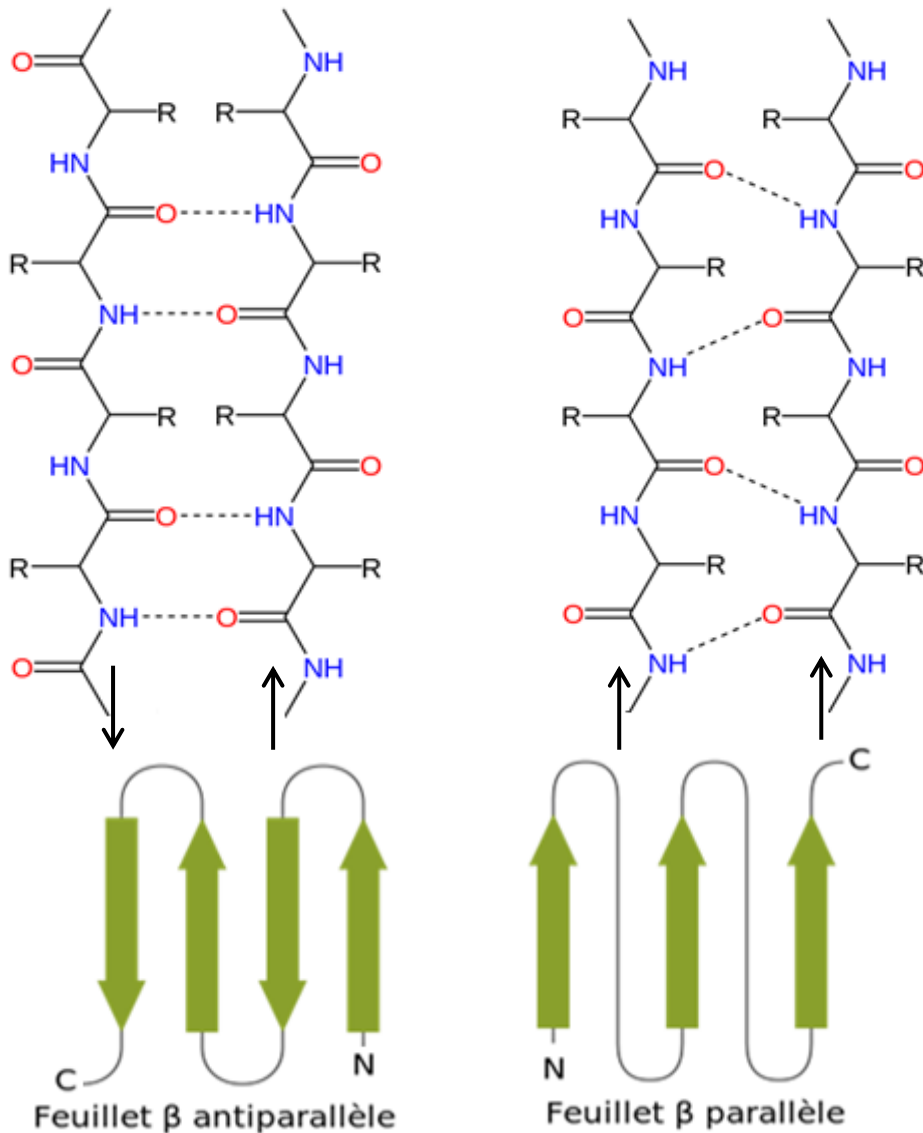


**Fig.9:** a= hélice  $\alpha$ , avec 3,6 aa/tour et 9 atomes carbone/tour ( $i \dots i + 4$ ). b= hélice gauche/droite

### 3.2.2. Feuille $\beta$ plissé

Dans ce cas du feuillet  $\beta$  plissé, le squelette peptidique ne s'enroule pas autour d'un axe mais se déplie en accordéon, pour former cette structure plate et étirée. Pour stabiliser cette structure, il se forme des liaisons hydrogène entre les groupements CO et NH de certains segments disposés parallèlement les uns par rapport aux autres (**Fig. 10**). Le feuillet  $\beta$  peut être **parallèle ou antiparallèle** selon que le sens des 2 chaînes est identique ou opposé (les parallèles sont moins stables, car les liaisons H subissent une certaine torsion (**Fig. 10**)).

La formation du feuillet  $\beta$  dépend étroitement de la séquence primaire du polypeptide. Certes, il n'y'a pas de règles précises de prédiction de la formation de cette structure, mais on peut noter que dans ces structures se forment plus fréquemment à partir de séquences riches en certains acides aminés, tel que Gly, Ala, Val ...etc.



**Fig.10:** Formation de feuillets  $\beta$  parallèles et antiparallèles

### Notes

\*Certaines protéines sont structurées en hélice  $\alpha$  seulement, d'autres en feuillet  $\beta$  seulement, mais la majorité des protéines ont les 2 structures en même temps.

\*Dans toutes les protéines structurées, ils existent des segments qui n'adoptent aucune structure secondaire. Ces segments permettent de relier les segments structurés et forment des zones flexibles nécessaires au repliement de la protéine. Le taux de séquences non structurées dans une protéine structurée est très variable d'une protéine à une autre, à titre d'exemple les segments non structurés dans la cytochrome C dépasse 50%.

### 3.2.3. Structure en pelote statistique (cas particulier)

Lorsque la conformation locale d'un segment protéique ne correspond à aucune de ces structures secondaires, on dit qu'il adopte une conformation en **pelote statistique** non périodique (*random coil*), par opposition aux hélices  $\alpha$  et aux feuillets  $\beta$  qui sont des structures périodiques. L'élasticité et la souplesse de l'élastine sont attribuées à sa structure en pelote statistique.

La structure en pelote statistique ne signifie pas l'absence de structuration. Il est évident que les protéines qui ne possèdent ni hélice  $\alpha$  ni feuillet  $\beta$  ont une structure parfaitement stable. Cette structure n'est pas pour autant inorganisée, mais elle obéit aux contraintes locales de voisinage. Donc, les pelotes ne sont pas moins organisées que les hélices  $\alpha$  et les feuillets  $\beta$ , mais plutôt elles sont simplement plus difficiles à décrire du fait de l'absence de structures répétitives. Si cette structure est aléatoire et n'est soumise à aucune règle, on n'obtiendrait jamais la même structure pour la même protéine à chaque fois.

Ce type de structure est souvent associé aux boucles. Certaines hormones et certaines toxines protéiques sont structurées en pelote.

### 3.2.4. Hélice gauche (cas particulier du collagène)

La structure en **hélice gauche** du collagène est un cas particulier dans la structure des protéines. En effet, les collagènes sont formés à partir de tropocollagène, qui adopte une conformation en **hélice gauche (tourne vers le haut dans les sens des aiguilles d'une montre)**. Contrairement à l'hélice  $\alpha$ , les hélices gauches ne sont stabilisées que par des liaisons hydrogène **intercaténares**, c'est-à-dire qu'on ne retrouve pas de liaisons hydrogènes entre les résidus d'une même hélice de collagène. L'hélice gauche compte moins de 3 résidus par tour alors que l'hélice  $\alpha$  contient 3,6 résidus par tour.

La conformation des molécules de tropocollagène est due à leur composition particulière en acides aminés, en fait, dans chaque chaîne polypeptidique, les résidus glycine représentent environ 1/3 et les résidus proline 1/4. Beaucoup de ces 2 résidus sont hydroxylés. La répétition du motif GlyXY où X est souvent une proline et Y une hydroxyproline est très fréquente. Le **scorbut** est dû à une anomalie de structure.

#### Notes

\*Le nombre de conformations d'un polypeptide de n acides aminés, susceptibles d'adopter S structures secondaires est égal à  $S^n$ .

\*Le temps d'exécution d'une conformation est estimé à  $10^{-12}$ s (1 picosec)

**A compléter par** (travail personnel) **Scorbut**

**Voir 2<sup>ème</sup> série de TD**