

Chapitre 3

Protéines et acides aminés

Introduction

En 1839, le chimiste hollandais Gerrit MULDER publia des résultats sur l'analyse de la fibrine du sang, des albumines du sérum sanguin et de l'œuf. Ceux-ci indiquaient que c'étaient des composés quaternaires (C, H, O, N) avec des pourcentages quasiment identiques pour ces quatre atomes et qui contenaient des traces variables de soufre et de phosphore. En 1938, sur la suggestion du chimiste suédois BERZELIUS, MULDER désigna ces composés sous le nom de **protéines** (du grec : prééminence).

Après une période d'identification des composants, **les acides α -aminés**, FISCHER et HOFMEISTER présentèrent chacun, le même jour, lors d'un congrès en 1902 le mode de liaison des acides aminés dans les protéines: **la liaison peptidique**.

Les protéines sont des biomolécules de première importance :

- par leur présence universelle dans le monde vivant, seuls des viroïdes en sont dépourvus.
- par leur abondance cellulaire : c'est le premier constituant après l'eau (10 fois plus que des glucides)
- par leur extrême diversité : elles assurent des fonctions vitales tant structurales que dynamiques et de plus elles sont le support de la spécificité des "espèces".

1. Les acides aminés standards des protéines

Sur un ensemble de quelques 300 aminoacides, pour le moment inventoriés, seuls 20 de ceux-ci composent les protéines en tenant compte du fait que certains aminoacides non standard, trouvés dans les protéines, sont modifiés après la traduction (modification post-traductionnelle). Les noms de ces 20 aminoacides, dont le dernier à être caractérisé fut la thréonine en 1935, n'obéissent à aucune nomenclature et évoquent soit leurs sources, soit leurs propriétés physiques ou encore un quelconque caractère analytique.

On a l'habitude d'utiliser des abréviations à trois lettres ou à une lettre pour cette série de vingt aminoacides.

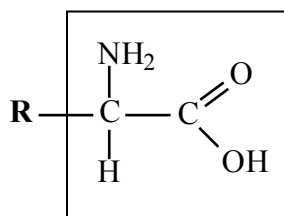
Les animaux supérieurs sont incapables de biosynthétiser la totalité de ces aminoacides. Chez l'homme, l'isoleucine, la leucine, la lysine, la méthionine, la phénylalanine, la thréonine, le tryptophane et la valine doivent être apportés par la ration alimentaire, ils sont qualifiés d'indispensables. A ceux-ci, on peut ajouter des aminoacides essentiels que l'organisme synthétise à une vitesse trop lente : l'arginine et l'histidine, qui sont indispensables pour le nouveau-né ou l'enfant.

Nom	Abréviations		Nom	Abréviations	
alanine	Ala	A	leucine	Leu	L
arginine	Arg	R	lysine	Lys	K
asparagine	Asn	N	méthionine	Met	M
acide aspartique	Asp	D	phénylalanine	Phe	F
cystéine	Cys	C	proline	Pro	P
acide glutamique	Glu	E	sérine	Ser	S
glutamine	Gln	Q	thréonine	Thr	T
glycine	Gly	G	tryptophane	Trp	W
histidine	His	H	tyrosine	Tyr	Y
isoleucine	Ile	I	valine	Val	V
Asp ou Asn	Asx	B	non identifiés par l'analyse		
Glu ou Gln	Glx	Z			
inconnu		X			

Formules générales

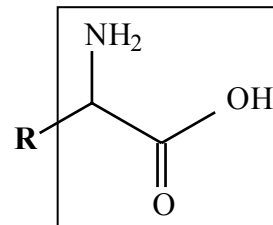
Les aminoacides ont en commun d'être des molécules bifonctionnelles portant un groupement **amine** (primaire) sur le carbone porteur du groupement **carboxyle**, dit carbone α . La fonction amine est une base et la fonction carboxyle est un acide (fonctions ionisables).

Ce sont des acides α -aminés (ou encore 2-amino-acides), exception pour la proline qui a une amine secondaire (acide α -iminé). Leur formule générique s'écrit :



acide α -aminé

R : chaîne latérale



Le résidu **R** est un résidu variable qu'on appelle la **chaîne latérale**. On distingue :

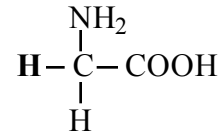
- les R aliphatiques à
 - chaîne carbonée de type carbure, linéaire ou branchée
 - chaîne carbonée portant des groupements fonctionnels (acide, amide, alcool, thiol, amine, guanidine)
- les R cycliques :
 - aromatiques
 - hétérocycles à azote

Sept groupes d'acides aminés peuvent être définis par rapport à leurs chaînes latérales :

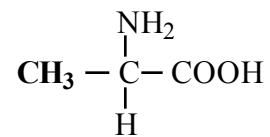
Groupe 1 : acides aminés aliphatiques

La chaîne latérale est une chaîne carbonée aliphatique linéaire ou ramifiée.

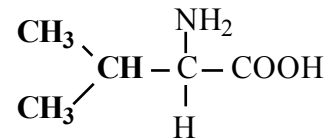
Glycine (Gly, G) :



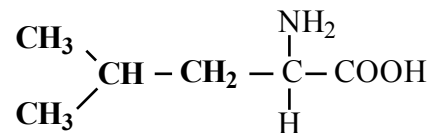
Alanine (Ala, A) : R est un groupement méthyle



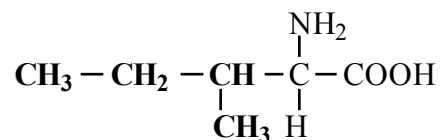
Valine (Val, V) : R est un groupement isopropyle



Leucine (Leu, L) : R est un groupement isobutyle



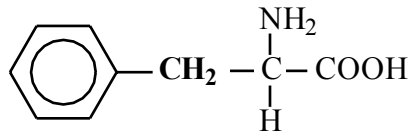
Isoleucine (Ile, I) : R est un groupement butyle secondaire



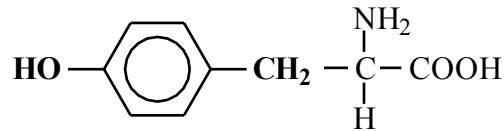
Groupe 2 : acides aminés aromatiques

La chaîne latérale contient un groupe aromatique, structure cyclique à 6 électrons délocalisés.

Phénylalanine (Phe, F) : R est un groupement phényle

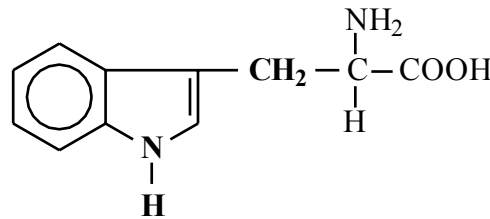


Tyrosine (Tyr, Y) : R est un groupement phénol



Les alcools aromatiques sont des acides très faibles dont la forme base conjuguée est un phénate.

Tryptophane (Trp, W) : R est un groupement indole

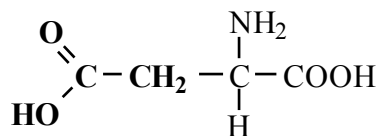


La délocalisation des électrons supprime les propriétés basiques de l'azote : le doublet électronique n'est plus un accepteur de protons.

Groupe 3 : acides aminés dicarboxyliques et leurs amides

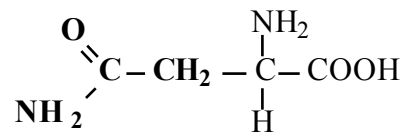
La chaîne latérale contient un groupement carbonyle libre ou sous forme d'amide.

Acide aspartique (Asp, D) : groupement β-carboxyle



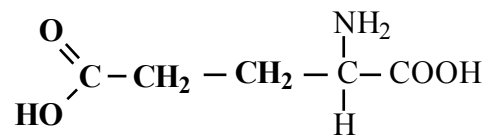
Le groupement β -carboxyle est ionisable. Il est chargé négativement à pH physiologique (forme base conjuguée).

Asparagine (Asp, N) : amide de l'acide aspartique



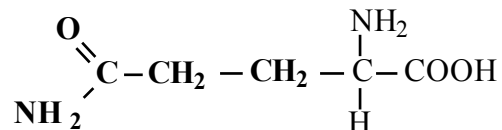
Le groupement amide n'est pas protonable : le doublet électronique de l'azote est délocalisé et engagé dans une orbitale hybride sp^2 avec les atomes C et O.

Acide glutamique (Glu, E) : groupement γ -carboxyle



Le groupement γ -carboxyle est ionisable. Il est chargé négativement à pH physiologique (forme base conjuguée).

Glutamine (Gln, Q) : amide de l'acide glutamique

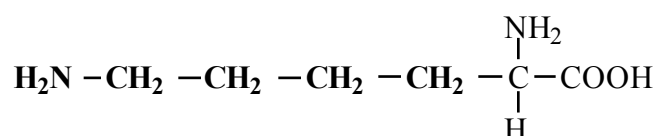


Le groupement amide n'est pas protonable (voir l'asparagine).

Groupe 4 : acides aminés dibasiques

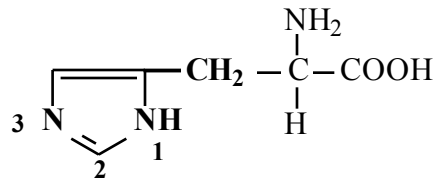
La chaîne latérale contient une fonction amine qui porte sous la forme acide conjuguée une charge positive.

Lysine (Lys, K) : groupement ϵ -amino



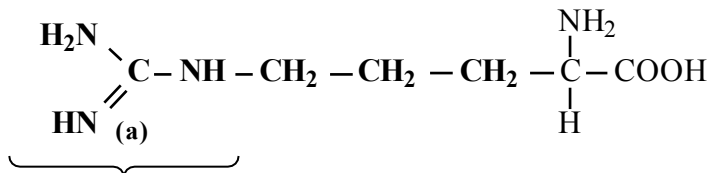
Le groupement ϵ -amino est un accepteur de proton (forme acide conjugué : ion ammonium).

Histidine (His, H) : groupement imidazole



Le doublet libre de l'azote en position 3 est un accepteur de proton. Le doublet de l'azote en position 1 participe à la conjugaison des doubles liaisons et n'est pas disponible pour accepter un proton. Bien évidemment, les rôles des deux azotes peuvent être échangés (formes mésomères).

Arginine (Arg, R) : groupement δ -guanidyle

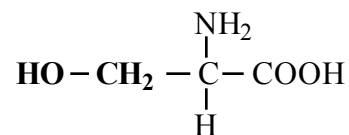


La double liaison de l'azote (a) et les doublets libres des deux autres azotes forment un hybride de résonance. Seul le doublet de l'azote (a) est libre et peut fixer un proton.

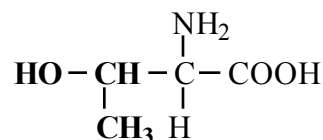
Groupe 5 : acides aminés alcools

La chaîne latérale contient une fonction alcool. Les groupes OH ne sont pas ionisables.

Sérine (Ser, S) : alcool primaire



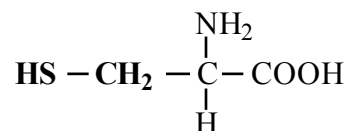
Thréonine (Thr, T) : alcool secondaire



Groupe 6 : acides aminés soufrés

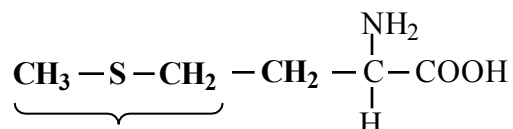
La chaîne latérale contient un atome de soufre.

Cystéine (Cys, C) : groupement thiol



Le groupement thiol (SH) ou sulfhydryle est un donneur de proton, c'est un acide très faible (forme base conjuguée : thiolate).

Méthionine (Met, M) : groupement thioéther

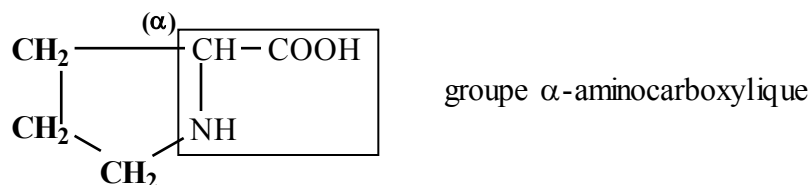


Groupe 7 : iminoacide

L'amine de l'acide aminé est une amine secondaire (imine).

Proline (Pro, P) :

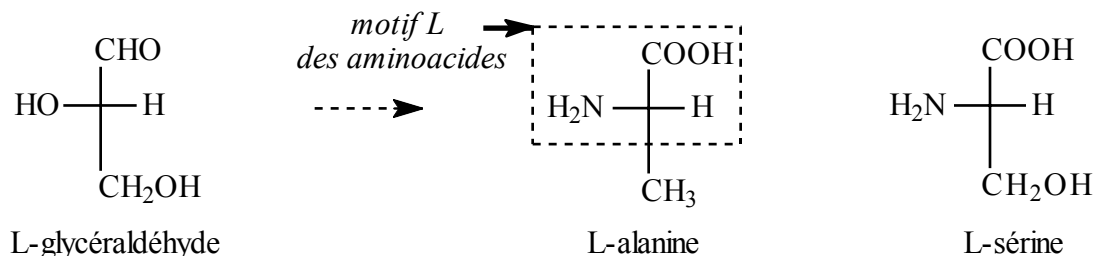
Le groupe α -amino est engagé dans une structure cyclique. L'amine est une amine secondaire (imine) dont l'azote présente un doublet libre, accepteur de proton : la fonction base d'un acide aminé est donc conservée.



Propriétés physiques

La chiralité

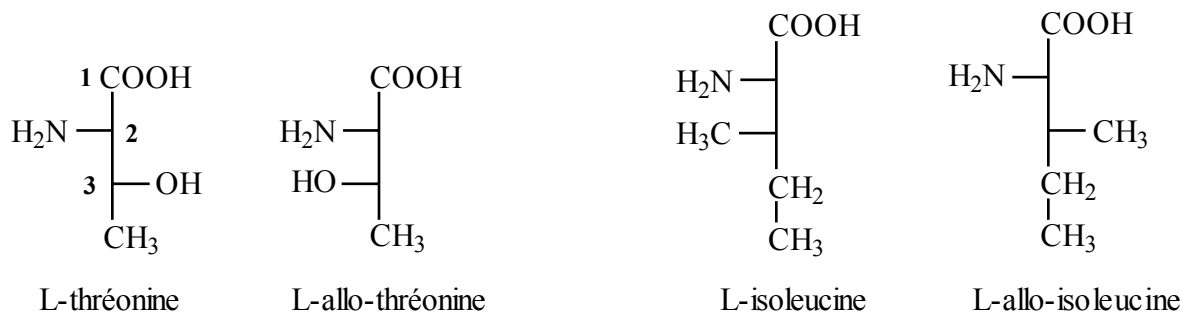
A l'exception de la glycine, le carbone α porte quatre substituants différents : c'est donc un centre chiral dont la conformation définira les stéréoisomères, isomères optiques à pouvoir rotatoire spécifique opposé. Les deux énantiomères sont définis de la même manière que pour les oses en prenant le glycéraldéhyde comme référence dans la représentation de Fischer :



Les acides aminés des protéines appartiennent tous à la série L. Comme pour les oses, aucune prédiction du pouvoir rotatoire ne peut être faite : un aminoacide de la série L peut être lévogyre ou dextrogyre.

Cas d'acides aminés ayant un deuxième centre chiral

Le carbone 3 (β) de la thréonine et de l'isoleucine est aussi un centre chiral : leur énantiomère (L) existera sous deux formes épimères. On affecte le préfixe "**allo**" à l'épimère que l'on ne trouve pas dans les protéines :



En utilisant la nomenclature stéréochimique **R/S**, la thréonine des protéines est de configuration (2S,3R), et l'isoleucine (2S,3S).

La racémisation et les acides aminés D

La racémisation est le passage d'un énantiomère à un autre. Certains microorganismes peuvent utiliser ou produire des aminoacides D, par exemple les antibiotiques peptidiques sécrétés par des bactéries :

- une D-Phe dans la gramicidine S et la tyrocidine A
- 6 aminoacides D (D-Leu et D-Val) sur les 15 de la gramicidine A.

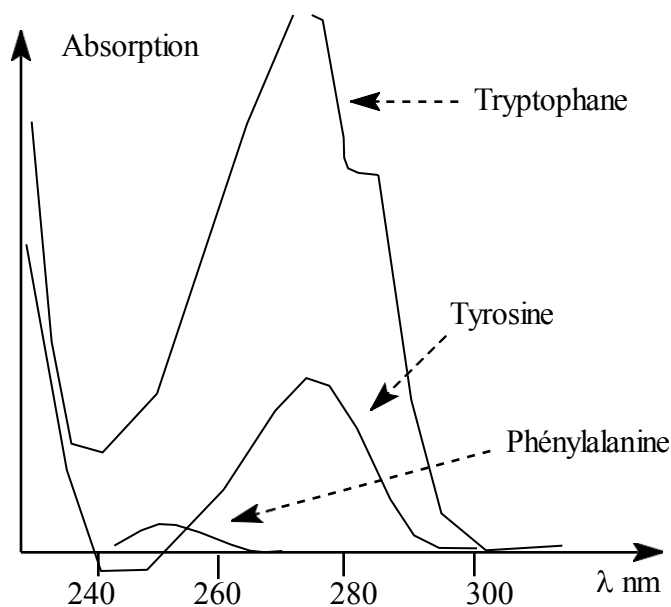
Cette particularité augmente la résistance de ces peptides à la dégradation par des enzymes protéolytiques dont une spécificité est d'agir que sur des aminoacides de série L.

Une solution d'acide aminé L évolue très lentement vers un l'équilibre racémique. Après la mort d'un organisme vivant qui ne contient que des acides aminés de série L, on aura une évolution lente vers l'équilibre racémique pour chacun d'entre eux : l'évaluation du rapport D/L de l'acide aspartique est utilisée comme méthode de datation de fossiles.

Absorption et fluorescence

Absorption

- les aminoacides n'absorbent pas la lumière visible, leurs solutions sont incolores.
- les bandes d'absorption dans l'infrarouge sont caractéristiques de leurs chaînes latérales
- les chaînes latérales aromatiques des aminoacides ont des spectres d'absorption caractéristiques dans l'ultraviolet moyen :



Spectres d'absorption des aminoacides aromatiques dans l'ultra-violet

La phénylalanine absorbe peu et le tryptophane est 4 fois plus absorbant que la tyrosine au maximum d'absorption, proche de 280 nm. Cette propriété est très souvent utilisée pour le dosage des peptides et des protéines.

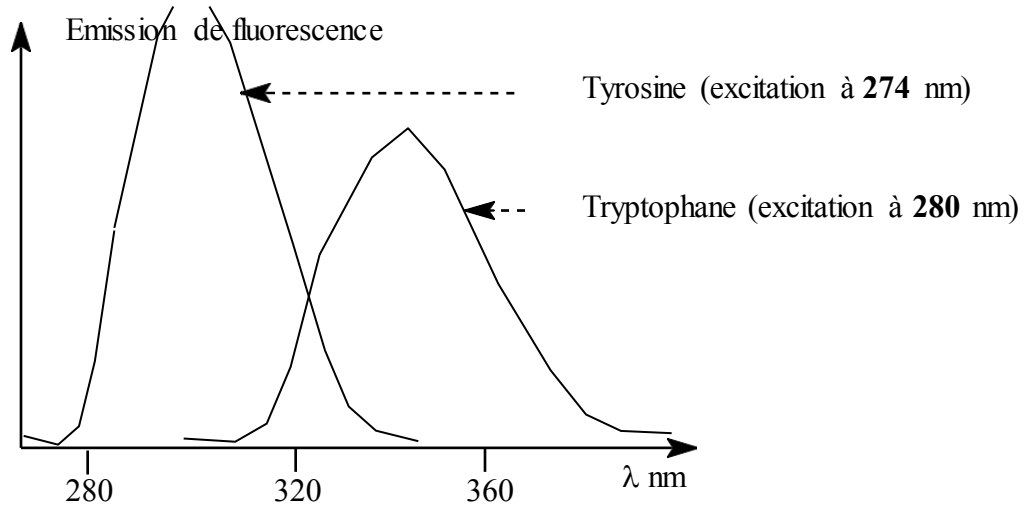
Remarquons que l'absorption de la tyrosine dans l'UV sera dépendante de l'état d'ionisation du phénol et par conséquent du pH.

Fluorescence

Certaines molécules, lorsqu'elles sont excitées par une lumière incidente à une longueur d'onde où elles absorbent ce rayonnement émettent une lumière de longueur d'onde plus grande : c'est le phénomène de fluorescence qui est maximum pour une longueur d'onde excitatrice égale à leur maximum d'absorption. Cette émission est très dépendante des

molécules voisines : cette dépendance permet des études fines de l'environnement des molécules fluorescentes.

C'est le cas du tryptophane et de la tyrosine dont la fluorescence permet l'étude de leur environnement proche dans les protéines (analyse de structure tridimensionnelle ou de mécanisme catalytique).



Solubilité

La solubilité des aminoacides dans l'eau (de un gramme à une centaine par litre) va dépendre essentiellement de deux facteurs :

- le double groupement fonctionnel commun qui peut s'ioniser et donc favoriser la dissolution
- la chaîne latérale qui peut avoir un caractère plus ou moins polaire ou apolaire.

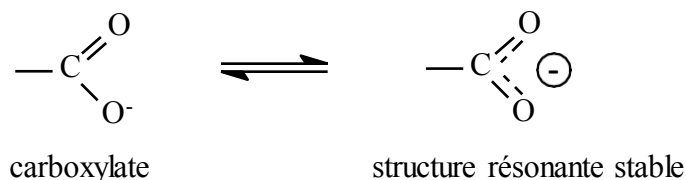
La solubilité dans les solvants organiques est faible de quelques mg/L et encore moins dans les solvants plus apolaires. En présence de deux phases liquides (éthanol/eau), les aminoacides se répartissent dans les deux phases avec des coefficients de partage spécifique : cette propriété est utilisée pour les classer (voir plus loin).

Propriétés chimiques

Lorsque l'acide aminé est libre, il a au moins deux groupements fonctionnels sur le même carbone α et pour certains un troisième sur la chaîne latérale.

Le groupement carbonyle (α -carboxylique)

Dans l'eau, à un pH physiologique, l'ion carboxylate est fortement stabilisé par résonance et donc peu réactif :



Estérification par les alcools

Cette réaction est utilisée pour séparer les aminoacides en phase gazeuse (et même en phase liquide) en produisant des dérivés esters butyliques.

Dans la synthèse peptidique, il faut lier le carboxyle d'un aminoacide avec l'amine du suivant, cela se fait dans les cellules par "activation" du carboxyle en anhydride d'acide avec l'ATP.

Décarboxylation

Cette réaction est présente dans les organismes vivants pour produire à partir des aminoacides des dérivés (amines) qui peuvent être des précurseurs d'autres molécules et ce par des décarboxylases.



Quelques produits de décarboxylation d'acides aminés :

- **sérine** : produit de décarboxylation : éthanolamine qui est le précurseur de la choline des phospholipides
- **histidine** : produit de décarboxylation : histamine qui est un vasodilatateur intervenant dans les réactions d'allergie ou d'inflammation
- **acide glutamique** : produit de décarboxylation : 4-aminobutanoïque ou "GABA" qui est un neurotransmetteur.

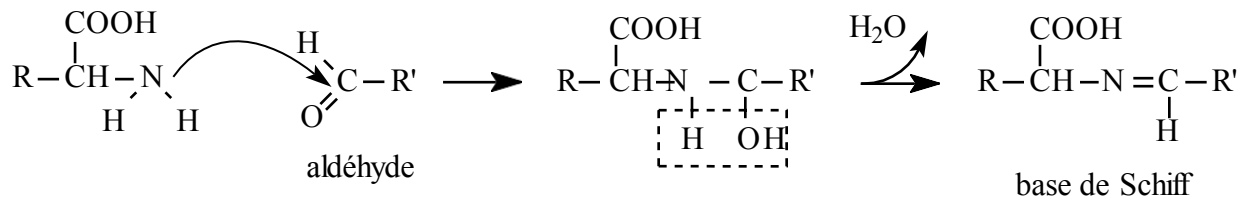
Le groupement amine (α -aminé)

Le groupement amine est très réactif : le doublet électronique de l'azote est un puissant nucléophile. Cette réactivité a été utilisée dans l'identification des aminoacides et l'élucidation des structures primaires des protéines.

Addition de carbonyle

Les fonctions α -aminés des aminoacides réagissent réversiblement avec les aldéhydes pour donner des bases de Schiff qui sont relativement labiles. Ces bases de Schiff apparaissent très souvent comme intermédiaires dans des réactions enzymatiques impliquant les aminoacides

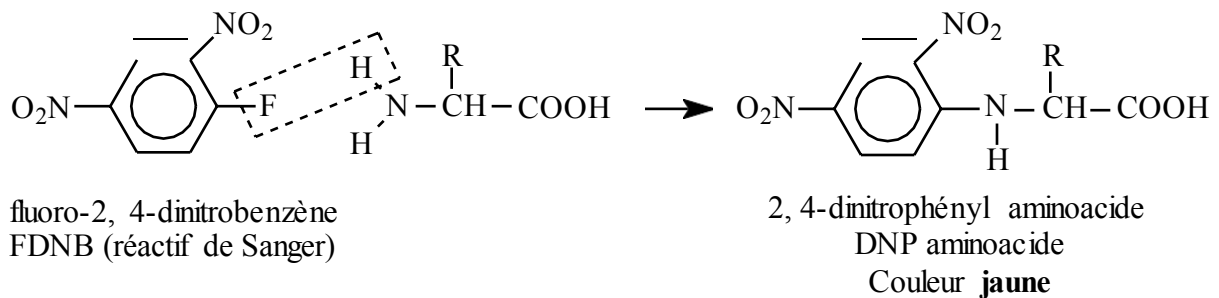
comme substrat. La proline qui contient une fonction amine secondaire ne réagit pas avec les aldéhydes.



Un des moyens très sensibles de détection des aminoacides utilise cette réaction : l'aldéhyde utilisé est le 1, 2-dialdéhyde benzénique. Le produit d'addition est très fluorescent.

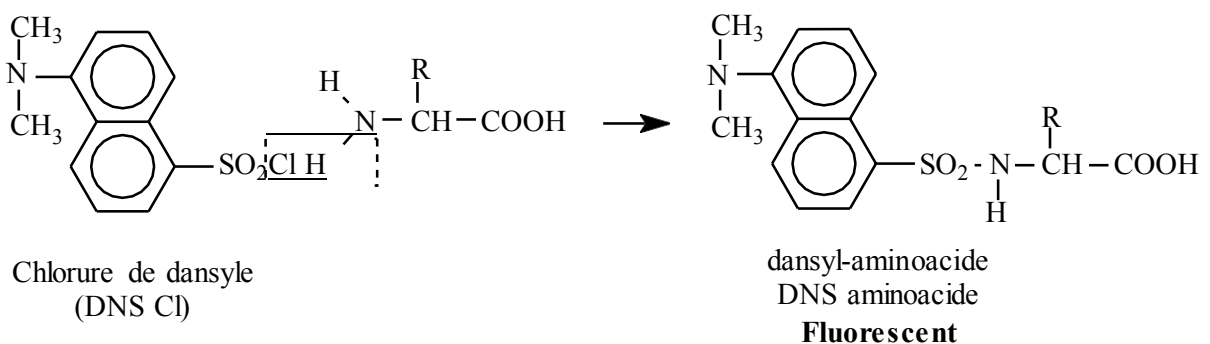
Arylation

Cette réaction à l'aide d'un dérivé aromatique activé a permis à Frederik SANGER (1953) d'établir la première structure primaire d'une protéine : l'insuline, hormone pancréatique qui contrôle la production et l'utilisation du glucose.



Acylation

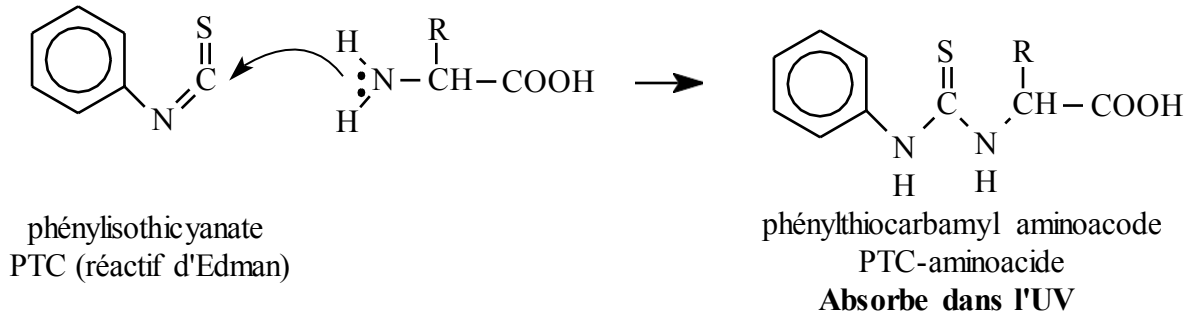
Le réactif de Sanger a été supplanté par un réactif donnant un produit plus stable et fluorescent permettant une plus grande sensibilité dans la détection : c'est le chlorure de dansyle (1-diméthyl-amino-naphtalène-5-sulfonyle).



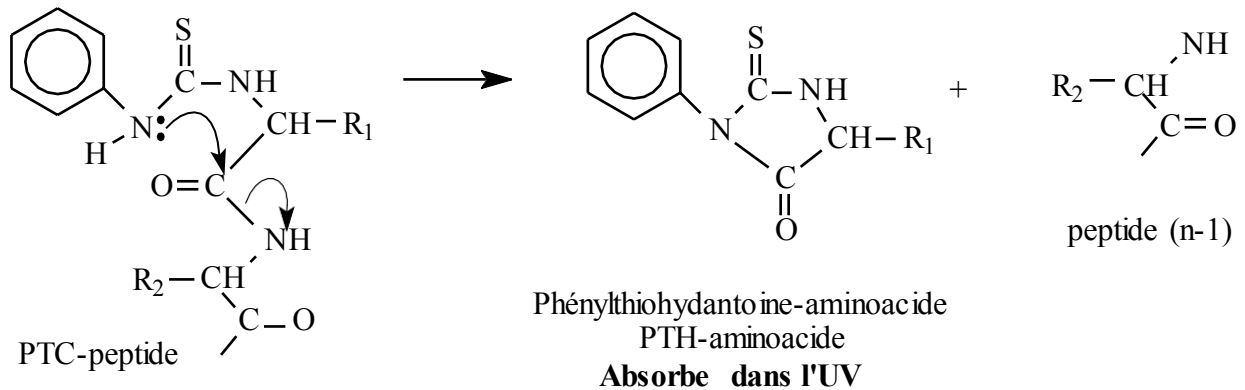
Dans les cellules, après leur biosynthèse, nombreuses sont les protéines subissent une acylation sur leur extrémité -NH₂ par l'acide acétique "activé".

Carbamylation

La carbamylation avec le phénylthiocyanate (PTC), à un pH basique de 9, donne des dérivés qui absorbent dans l'ultraviolet et facilement séparable par chromatographie. De plus la réaction avec l'acide aminé terminal d'une protéine libre le dérivé d'addition et une protéine amputée de son acide aminé N-terminal : en itérant le processus, la détermination de la structure primaire de la protéine sera possible (dégradation récurrente d'Edman).

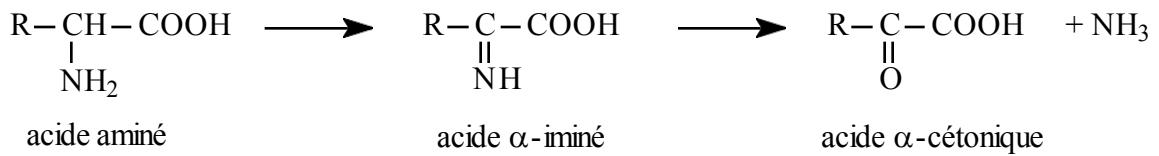


Dans le cas d'un peptide de **n** acides aminés, le PTC-peptide va subir une cyclisation et une coupure à un pH légèrement acide, libérant un phénylthiohydantoïne-aminoacide identifiable (**PTH-aminoacide**), qui absorbe dans l'UV, et un peptide de (**n-1**) acides aminés.

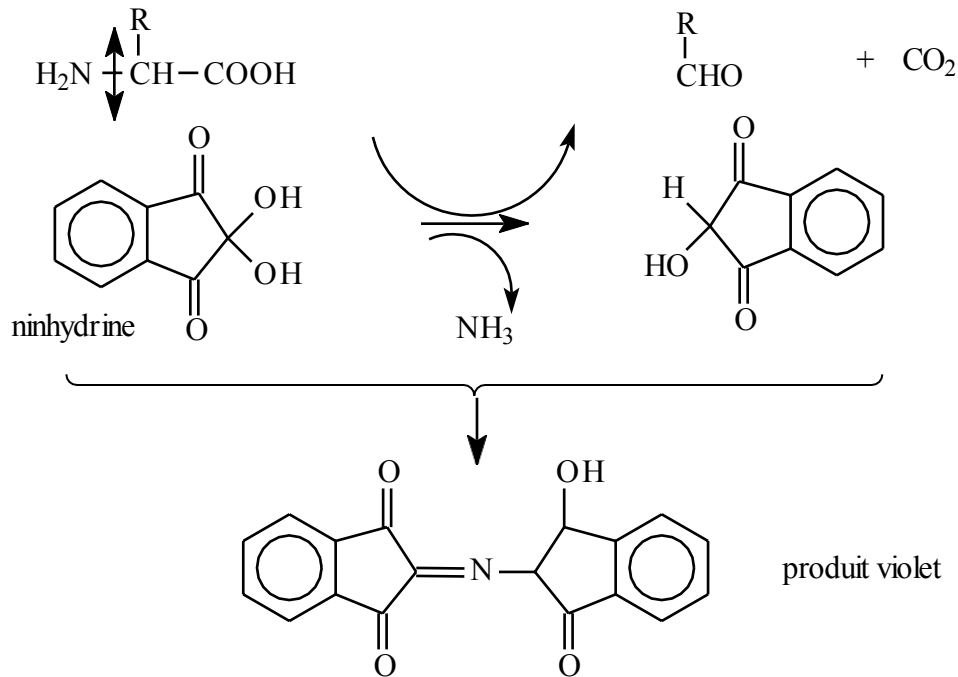


Désamination

Pour maintenir la réserve intracellulaire des 20 acides aminés servant à la synthèse protéique, le métabolisme passera par des désaminations avec oxydation qui produiront des acides α-cétoniques, source principale, sinon la seule, à partir de laquelle les acides aminés sont synthétisés.



La réaction avec la **ninhydrine** est l'une des plus connue et utilisée, elle aboutit à un produit violet pour les amines primaires et à un autre dérivé de couleur jaune pour les amines secondaires. L'acide aminé est complètement dégradé par une réaction de désamination et de décarboxylation.



Les groupes latéraux

Les groupes latéraux réactifs des aminoacides sont :

- les carboxyles et amines : ils sont de réactivité identique à celles des groupes α
- les chaînes latérales contenant du soufre, des groupes imidazole ou aromatique autre que le phényle inerte.

Ces réactions sont présentes dans les cellules pour modification post-traductionnelle, comme sites actifs dans les protéines. Elles sont aussi utilisées par les biochimistes dans l'étude des protéines.

Les chaînes latérales des lysines

La réactivité de la fonction amine se retrouve dans :

- la condensation des aldoses et cétooses sur les protéines : glycation des protéines du plasma
- la fixation des coenzymes prosthétiques sur leur apoprotéine

- l'établissement de liaisons covalentes entre molécules de protéines fibreuses (collagène) qui passe par l'oxydation enzymatique d'une chaîne latérale de lysine en aldéhyde et ensuite la condensation de l'aldéhyde avec une amine d'une lysine d'une autre chaîne (base de Schiff).

Les chaînes latérales sulfurées

Le groupe thiol de la cystéine est très réactif :

- son oxydation permet la formation des ponts disulfures que l'on trouve dans les protéines (pont intra-chaîne ou entre chaînes polypeptidiques).



- des thiols sont indispensables à l'expression de la fonction de nombreuses protéines
- les dérivés mercuriels se fixent spécifiquement sur les groupements thiols
- les protéines peuvent fixer des métaux par l'intermédiaire du groupe thiol de la cystéine (l'histidine peut aussi fixer des métaux sur le groupement imidazole)
- le sélénium (oligo-élément) peut remplacer le soufre du groupe thiol : **sélocystéine**.
- dans les études de protéines, l'acide iodoacétique est utilisé pour bloquer les thiols des protéines : l'acide aminé modifié obtenu est une carboxyméthylcystéine.

Le groupe thioéther de la méthionine est très peu réactif :

- dans les cellules, l'alkylation enzymatique du soufre par un nucléoside produit le dérivé S-Adénosylméthionine qui est un coenzyme d'activation et de transfert du groupement méthyle.
- dans les études de protéines, le bromure de cyanogène (BrCN) est utilisé. Il réagit avec le soufre d'une méthionine et coupe la chaîne polypeptidique à cet endroit.

Les chaînes latérales alcools et amides

Ces groupes sont peu réactifs :

- la sérine intervient dans des mécanismes de catalyse enzymatique
- la sérine, la thréonine et la tyrosine sont des sites potentiels de phosphorylation
- la sérine, la thréonine et l'asparagine sont des sites potentiels de O et N-glycosylations
- l'hydroxyle de la tyrosine donne au noyau phényle une réactivité assez forte qui favorise les substitutions électrophiles. L'iodation enzymatique de la tyrosine est une étape dans la biosynthèse des prohormone (T4) et hormone (T3) thyroïdiennes.

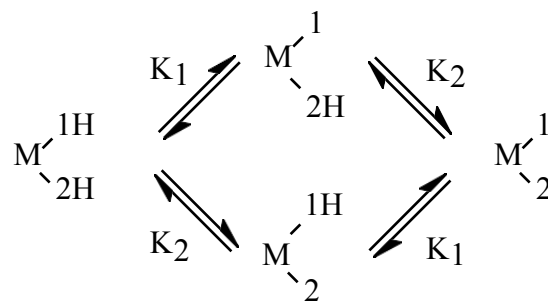
Propriétés ioniques

Lorsque l'acide aminé est libre, il a au moins deux groupements fonctionnels sur le même carbone α et pour certains un troisième sur la chaîne latérale.

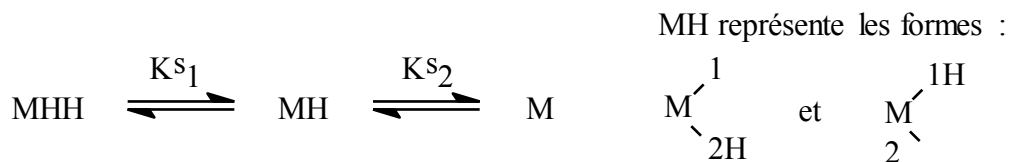
L'un des deux groupements est un acide (carboxylique) et l'autre est une base (amine). C'est une molécule amphotère. Avant d'étudier les équilibres des diverses formes qu'on peut trouver en solution aqueuse pour un aminoacide et les courbes de titrage, nous allons faire une remarque générale.

Remarque générale

Considérons une molécule (M) portant deux fonctions ionisables et étudions les diverses formes en solution en équilibre : une fonction notée 1 et une fonction notée 2 avec leurs constantes **individuelles** acides de dissociation notées respectivement K_1 et K_2 . Nous pouvons représenter le système ainsi dans le cas où les **groupes sont indépendants** (l'état de fixation du site 1 n'a aucune influence sur l'équilibre du site 2 et vice-versa).



Nous pouvons aussi l'écrire en considérant uniquement le nombre d'H porté par la molécule indépendamment du groupe concerné en définissant de nouvelles constantes acides de dissociation, constante **apparente** ou **successive** : K_1^s et K_2^s (avec $K_1^s > K_2^s$)



Pour ce cas particulier de deux fonctions, les relations entre les constantes individuelles et successives s'écrivent en utilisant la relation $[\text{MH}] = [\text{M12H}] + [\text{M1H2}]$ et bien sûr les relations d'équilibre pour les deux façons d'écrire le système :

$$(KS) \quad K_1^s = K_1 + K_2 \quad \text{et} \quad K_2^s = \frac{K_1 K_2}{K_1 + K_2}$$

Remarquons que si les deux groupes sont identiques et bien sûr indépendants, nous avons :

$$K_1^s = 2K \quad \text{et} \quad K_2^s = \frac{K}{2}$$

d'une manière générale, on peut démontrer que si nous avons **n**

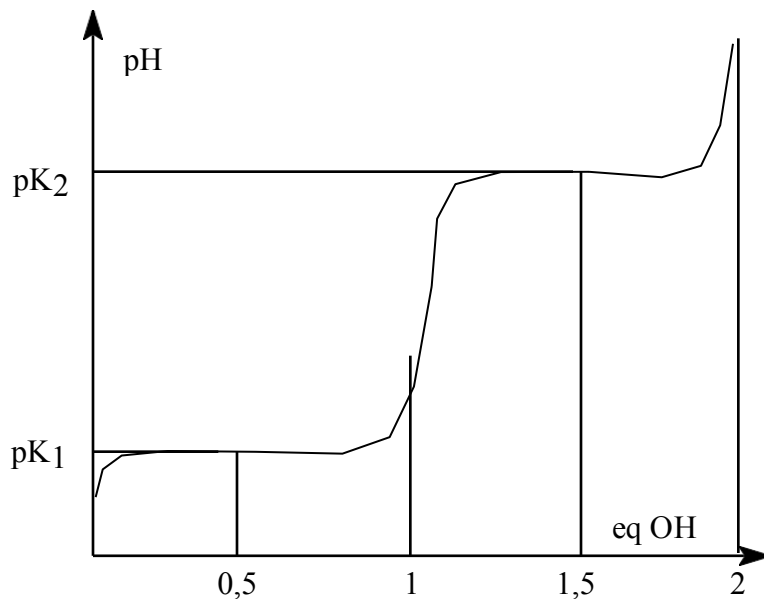
groupes identiques et indépendants, les constantes successives sont : $K_j^s = \frac{n-j+1}{j} K$ où

K est la constante individuelle.

Pour les aminoacides, les constantes tabulées des fonctions ionisables sont **les constantes individuelles acides de dissociation** et leurs pK correspondant.

Examinons les courbes de titrage que nous pouvons obtenir :

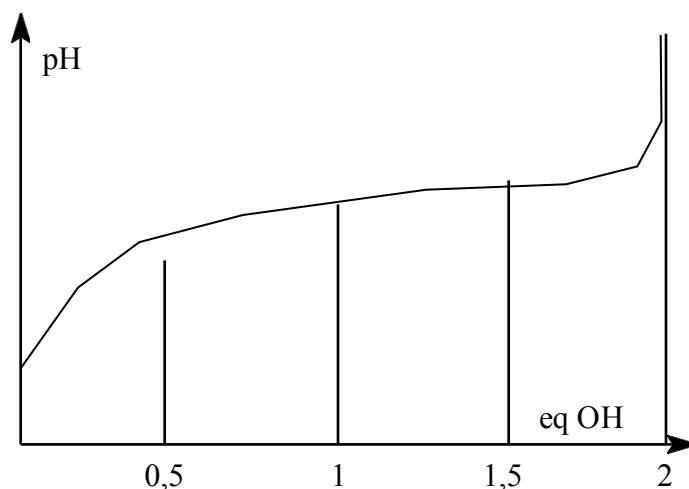
1) Si K_1 et K_2 ont des **valeurs assez différentes**, la courbe de titrage ressemblera à la juxtaposition de chacune des courbes relatives à chacun des deux groupes ionisables : nous avons déjà vu que pour des valeurs de pH extérieures au segment $[pK-1 \dots pK+1]$, l'équilibre est complètement déplacé.



Courbe de titrage

Cas où les constantes individuelles de dissociation sont différentes d'un facteur **supérieur** à 100 (2 unités pH entre les pK)

2) Si K_1 et K_2 ont des **valeurs assez proches**, la courbe de titrage sera une combinaison de chacune des courbes de titrage



Courbe de titrage

Cas où les constantes individuelles de dissociation sont différentes d'un facteur **inférieur** à 10 (1 unité pH entre les pK)

Caractéristiques des courbes de titrage :

1) K_1 et K_2 ont des **valeurs assez différentes** :

- pour des valeurs de pH voisines de pK_1 , le seul équilibre concerné est celui du groupe 1, la valeur du pH au premier point de demi-équivalence correspond au pK de ce groupe (1).
- pour des valeurs de pH voisines de pK_2 , le seul équilibre concerné est celui du groupe 2, la valeur du pH au deuxième point de demi-équivalence correspond au pK de ce groupe (2).

Le système d'équations (KS) peut se simplifier et on retrouve bien dans l'écriture avec les équilibres successifs ce que nous avons dit dans les lignes précédentes, $[MH]$ représente uniquement la forme $[M_1H_2]$ et cela se passe comme si on pouvait étudier chacun des équilibres séparément :

$K_1 \gg K_2 \Rightarrow K_1^s \approx K_1$ et $K_2^s \approx K_2$, les constantes successives sont égales aux constantes individuelles (à 10% pour une différence de 1 unité entre pK_1 et pK_2 , de 3% pour 1,5 unité de différence) .

2) K_1 et K_2 ont des **valeurs assez proches**

Les points de la courbe de titrage, qui correspondent aux demi-équivalences successives dans le cas précédent, n'ont plus la valeur de pH correspondant aux pK des groupes : la variation de pH implique un déplacement concomitant des équilibres de chacun des deux groupes.

3) K_1 et K_2 ont des **valeurs identiques** : K

Dans ce cas particulier, le point de la courbe de titrage qui correspond à la demi-équivalence globale a une valeur de pH correspondant au pK du groupe ionisable. Toutefois, dans l'écriture des équilibres successifs, il y a deux équilibres avec des constantes successives différentes.

Définition du pH isoélectrique

Considérons une molécule portant des fonctions ionisables telles que différentes formes chargées de la molécule puissent exister en solution aqueuse. Le **pH isoionique (pI)** est la **valeur du pH de la solution dans laquelle la charge totale nette moyenne de la molécule est nulle** : (on dit aussi pH isoélectrique)

$$\sum_i z_i [f_i^p] - \sum_j z_j [f_j^n] = 0 \quad \text{où } [f_i^p] \text{ est la concentration de la forme } i \text{ portant } z_i \text{ charges nettes positives et } [f_j^n] \text{ est la concentration de la forme } j \text{ portant } z_j \text{ charges nettes négatives.}$$

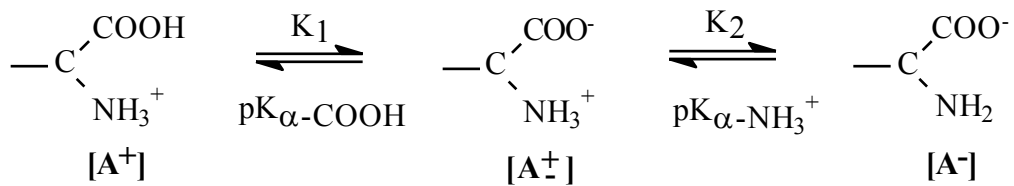
Les aminoacides qui portent au moins une fonction carboxylique et une fonction amine se présenteront en solution aqueuse sous diverses formes en équilibre, chargées positivement ou négativement.

Nous pouvons faire la remarque importante suivante :

- pour un **pH inférieur à la valeur du pI**, la charge nette moyenne de l'acide est **positive**
- pour un **pH supérieur à la valeur du pI**, la charge nette moyenne de l'acide est **négative**

Aminoacides à chaîne latérale ne comportant pas de groupe ionisable

Les pK des groupes α -COOH et α -aminé sont très différents, celui correspondant à l'acide étant beaucoup plus faible. Les équilibres successifs de dissociation s'écrivent avec les constantes individuelles de dissociation (égales aux constantes successives dans ce cas) :



La forme qui porte une charge nette nulle est un ion mixte ou bipolaire ou encore zwitterion.

pH isoélectrique : défini par $[A^+] - [A^-] = 0$

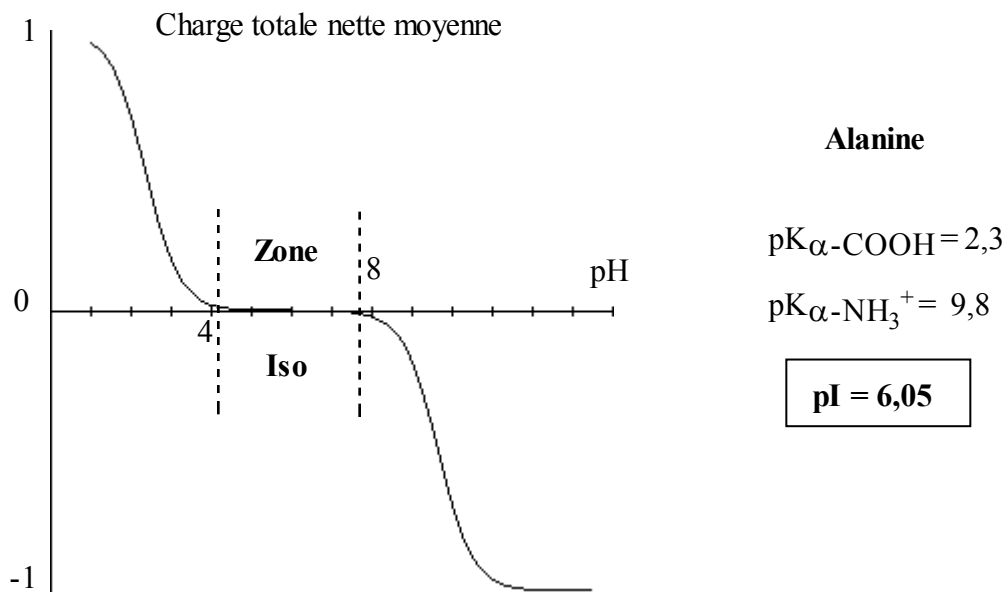
Nous avons à résoudre le système d'équations suivant :

$$\left\{ \begin{array}{l} K_1 = \frac{[A^\pm][H^+]}{[A^+]} \\ K_2 = \frac{[A^-][H^+]}{[A^\pm]} \end{array} \right. \quad \text{d'où } [H^+]^2 = \frac{K_1 K_2}{[A^+]} \quad \text{ou encore } \boxed{pI = \frac{pK_1 + pK_2}{2}}$$

$[A^-] = [A^+]$

Aminoacide	pK α -COOH	pK α -amine	pI
Gly, Ala, Val, Leu, Ile	2,3	9,8	6,05
Ser	2,2	9,2	5,7
Thr	2,6	10,4	6,5
Met	2,3	9,2	5,75
Asn	2,0	8,8	5,4
Gln	2,2	9,1	5,65
Phe	2,6	9,2	5,9
Trp	2,4	9,4	5,9
Pro	2,0	10,6	6,3

Pour ces aminoacides avec une chaîne latérale ne comportant pas de fonction ionisable, la courbe de la charge totale nette moyenne en fonction du pH a l'allure suivante :



La zone de pH où l'acide aminé a une charge totale nette moyenne nulle est **appelée zone isoionique ou isoélectrique** : celle-ci sera d'autant plus importante que les pK de la fonction carboxylique et de la fonction amine seront éloignés et on peut l'estimer par excès à $[(pK_{\alpha-COOH} + 1) \dots (pK_{\alpha-NH_3^+} - 1)]$.

Dans une zone de pH $< pK_{\alpha-COOH} - 1$, la forme prépondérante est : A^+ (cation)

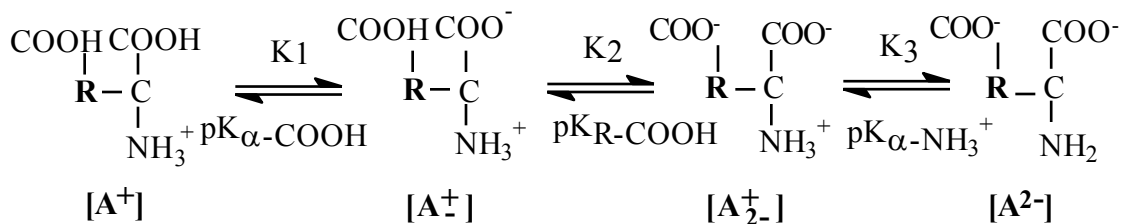
Dans la zone de pH "iso", la forme prépondérante de l'acide aminé est : A^{\pm}

Dans une zone de pH $> pK_{\alpha-NH_3^+} + 1$, la forme prépondérante est : A^- (anion)

- pour un pH $< pI$, l'acide aminé a une charge totale nette moyenne positive
- pour un pH $> pI$, l'acide aminé a une charge totale nette moyenne négative

Aminoacides à chaîne latérale comportant une fonction acide

Dans le cas d'une fonction carboxylique, le phénomène d'ionisation est représenté par les équilibres successifs suivants (le pK du COOH de la chaîne latérale est plus élevé que le pK du α -COOH et plus faible que le pK de l' α -amine) :



pH isoélectrique : défini par $[A^+] - [A^+]_{2-} - 2[A^{2-}] = 0$

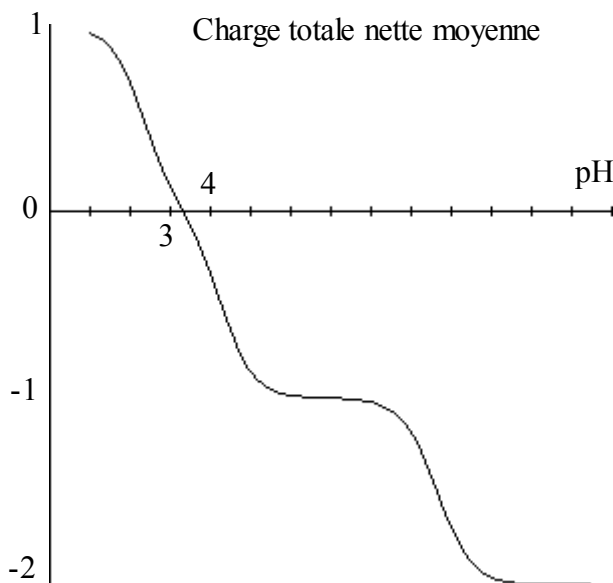
Nous avons à résoudre le système d'équations suivant :

$$\left\{ \begin{array}{l} K_1 = \frac{[A^+][H^+]}{[A]} \quad \alpha\text{-carboxylique} \\ K_2 = \frac{[A_{2-}^+][H^+]}{[A^+]} \quad R\text{-carboxylique} \\ K_3 = \frac{[A^{2-}][H^+]}{[A_{2-}^+]} \quad \alpha\text{-amine} \end{array} \right. \quad \text{d'où } \boxed{[H^+]^3 - K_1K_2[H^+] - 2K_1K_2K_3 = 0}$$

$$\left\{ \begin{array}{l} [A^+] - [A_{2-}^+] - 2[A^{2-}] = 0 \end{array} \right.$$

Les équilibres successifs de dissociation en fonction du pH nous indiquent que si la forme A^{2-} est prépondérante, alors la forme A^+ sera en concentration très faible et l'équation de contrainte du pH isoélectrique ne pourra être satisfaite : nécessairement la valeur du pH sera inférieure à la valeur de pK_3 ($[H^+] \gg K_3$). L'équation du troisième degré en H peut donc être approximée :

$$[H^+]^2 = K_1K_2 \quad \text{d'où } \boxed{pI = \frac{pK_1 + pK_2}{2}}$$



Acide glutamique

$pK_{\alpha\text{-COOH}} = 2,2$

$pK_{R\text{-COOH}} = 4,3$

$pK_{\alpha\text{-NH}_3^+} = 9,7$

$pI = 3,25$

Contrairement aux aminoacides à chaîne latérale ne comportant pas de fonctions ionisables, il n'y a pas de zone isoélectrique pour l'acide glutamique : la forme portant une charge nette nulle est au centre des deux équilibres correspondant aux deux fonctions carboxyliques et leurs pK ont des valeurs trop proches. Il en est de même pour l'acide aspartique.

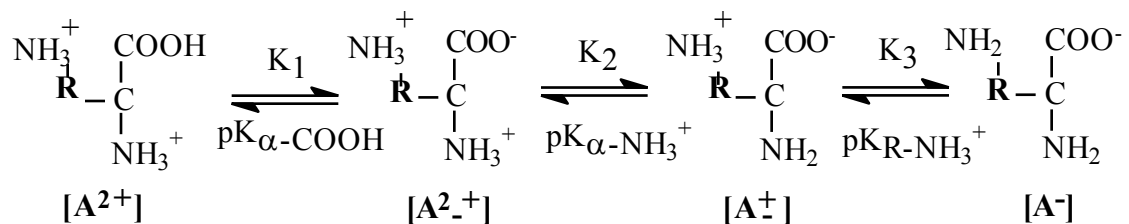
Aminoacide	pK α -COOH	pK α -amine	pK R-acide	pI
Asp	2,1	9,8	3,9	3,0
Glu	2,2	9,7	4,3	3,25
Cys	1,7	10,8	8,3	5,0
Tyr	2,2	9,1	10,1	5,65

Pour la cystéine, la différence des pK est élevée (1,7 et 8,3) et on retrouvera le cas du paragraphe précédent avec une zone isoélectrique [2,7 ... 7,3] (par excès).

Pour la tyrosine, le deuxième équilibre **ne concerne pas la deuxième fonction acide** mais la fonction α -aminée, le pI sera la moyenne des pK α -COOH et α -amine. La différence des pK est élevée (2,2 et 9,1) et on aura une zone isoélectrique [3,2 ... 8,1] (par excès)

Aminoacides à chaîne latérale comportant une fonction base

Ces aminoacides portent une amine sur leur chaîne latérale. , le phénomène d'ionisation est représenté par les équilibres successifs suivants (le pK de la fonction α -amine est plus faible que celui de la fonction R-amine) :



pH isoélectrique : défini par $2[\text{A}^{2+}] + [\text{A}^{2,+}] - [\text{A}^-] = 0$

Nous avons à résoudre le système d'équations suivant :

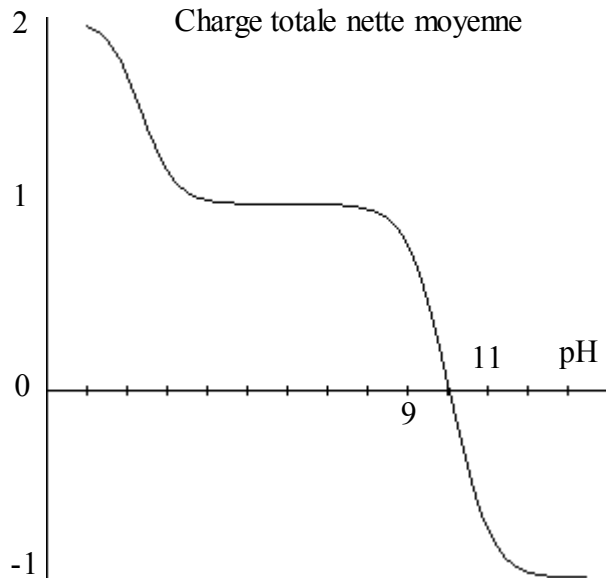
$$\left\{ \begin{array}{l}
 \text{K}_1 = \frac{[\text{A}^{2+}][\text{H}^+]}{[\text{A}^{2,+}]} \quad \alpha\text{-carboxylique} \\
 \text{K}_2 = \frac{[\text{A}^+][\text{H}^+]}{[\text{A}^{2,+}]} \quad \alpha\text{-amine} \\
 \text{K}_3 = \frac{[\text{A}^-][\text{H}^+]}{[\text{A}^+]} \quad \text{R-amine}
 \end{array} \right. \quad \text{d'où } \boxed{2[\text{H}^+]^3 + \text{K}_1[\text{H}^+]^2 - \text{K}_1\text{K}_2\text{K}_3 = 0}$$

$$\left\{ \begin{array}{l}
 2[\text{A}^{2+}] + [\text{A}^{2,+}] - [\text{A}^-] = 0
 \end{array} \right.$$

Les équilibres successifs de dissociation en fonction du pH nous indiquent que si la forme A^{2+} est prépondérante, alors la forme A^- sera en concentration très faible et l'équation de contrainte du pH isoélectrique ne pourra être satisfaite : nécessairement la valeur du pH sera

supérieure à la valeur de pK_1 ($[H^+] \ll K_1$). L'équation du troisième degré en H peut donc être approximée :

$$K_1[H^+]^2 = K_1K_2K_3 \text{ d'où } \boxed{pI = \frac{pK_2 + pK_3}{2}}$$



Lysine

$$pK_{\alpha\text{-COOH}} = 2,2$$

$$pK_{\alpha\text{-NH}_3^+} = 8,95$$

$$pK_{R\text{-NH}_3^+} = 10,5$$

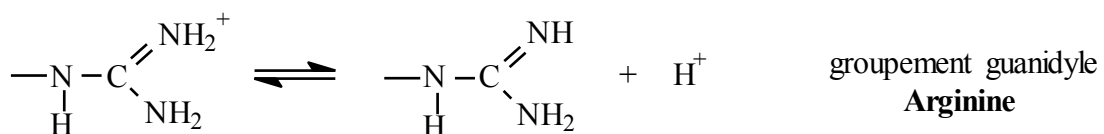
$$\boxed{pI = 9,725}$$

Contrairement aux aminoacides à chaîne latérale ne comportant pas de fonctions ionisables, il n'y a pas de zone isoélectrique pour la lysine : la forme portant une charge nette nulle est au centre des deux équilibres correspondant aux deux fonctions amines et leurs pK ont des valeurs trop proches. Il en sera de même pour l'histidine et l'arginine avec toutefois une "petite" zone isoélectrique puisque les deux équilibres concernés ont des valeurs de pK différentes de 3 unités.

Aminoacide	pK α -COOH	pK α -amine	pK R-amine	pI
His	1,8	9,2	6	7,6
Lys	2,2	8,95	10,5	9,725
Arg	2,2	9,05	12,5	10,775

Dans l'écriture des équilibres successifs, il faut remarquer que, pour l'histidine, la fonction R-amine perdra son proton avant celle α -amine.

Rappel :





Récapitulatif des pK des fonctions ionisables

	Asp	Glu	His	Cys		Tyr	Lys	Arg
	β -COOH	γ -COOH	NH ⁺	SH	α -NH ₃ ⁺	Φ -OH	ϵ -NH ₃ ⁺	NH ₂ ⁺
	1,8 - 2,6	3,9	6	8,3	8,9 - 10,8	10,1	10,5	12,5

Autres classifications des aminoacides

Dans le paragraphe 2.1 (Formules des aminoacides), les aminoacides standard ont été classés en sept groupes par rapport aux propriétés de leurs chaînes latérales.

D'autres classifications ont été définies dont les deux les plus utilisées sont :

- propriété d'hydrophilicité
- valeur du pI

Hydrophilicité

Par rapport à la structure et la fonction des peptides et des protéines, la classification des aminoacides a été faite en tenant compte essentiellement des interactions de l'acide avec son environnement, c'est-à-dire les liaisons faibles autres que les interactions de Van der Waals : liaisons électrostatiques, hydrogène, interactions hydrophobes. On distingue trois classes sur le critère de la **polarité de la chaîne latérale (R)** :

- R apolaire (hydrophobe)
Gly, Ala, Val, Pro, Met, Leu, Ile, Phe, Tyr, Trp
- R polaire mais non chargé à pH physiologique (hydrophile)
Ser, Cys, Asn, Thr, Gln
- R polaire et chargé (plus ou moins hydrophile selon le pH)
Asp, Glu, Lys, His, Arg

La valeur de la polarité de l'acide est soit calculée théoriquement à partir des valeurs des groupes simples, ou obtenue expérimentalement en mesurant le coefficient de partage entre deux phases liquides éthanol/eau.

pH isoélectrique (pI)

Trois catégories d'acides aminés sont définies d'après la valeur de leur pI :

- acide aminé "acide" : Asp, Glu
- acide aminé "basique" : Lys, Arg, His
- acide aminé "neutre" : les autres

2. Les acides aminés non standard

Les acides aminés non standard sont soit :

- acide aminé d'une protéine modifié après la traduction (modification post-traductionnelle)
- des intermédiaires de la biosynthèse d'autres acides aminés, des éléments de construction d'autres molécules (lipides, coenzymes) ou encore des molécules actives

Aminoacides subissant une modification post-traductionnelle

Ces modifications peuvent concerner dans les protéines, les groupes des extrémités terminales ou ceux qui appartiennent à la chaîne latérale.

Modification	Acide aminé	Protéines ou fonctions
hydroxylation	Pro -> 4-hydroxyPro Lys -> 5-hydroxyLys	établissement de liaisons H sites d'O-glycosylation (collagène)
méthylation	azote de l'histidine	- protéines contractiles - protéines se complexant aux acides nucléiques (histones)
carboxylation	Glu -> 4-carboxyGlu	protéine fixant le Ca ⁺⁺ (facteurs de coagulation et de l'ostéogénèse)
acétylation	α - NH ₂ terminaux ϵ - NH ₂ des Lys	résistance à la dégradation interactions avec l'ADN (histones)
O et N-glycosylations	O : Ser, Thr N : Asn	glycoprotéines
fixation de lipides	α - NH ₂ terminaux ou -SH	inclusion des protéines dans les membranes
phosphorylation	OH des Ser et Thr OH des Ser, Thr et Tyr	phosphoprotéines (caséine) modulation de l'activité

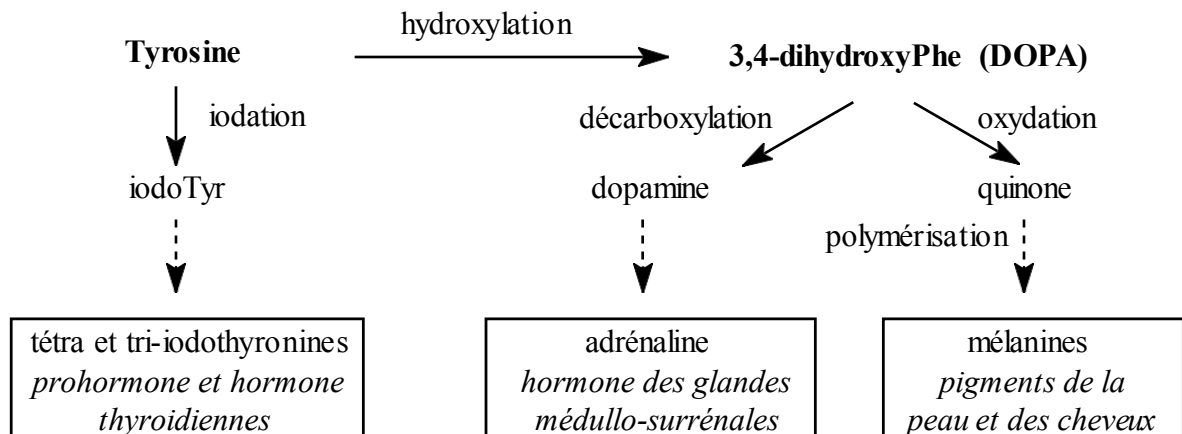
Autres aminoacides

Citons quelques dérivés importants :

Les dérivés obtenus par décarboxylation :

- histidine décarboxylée en histamine qui entre en jeu dans les réactions d'inflammation
- acide glutamique décarboxylé en 4-aminobutanoïque (GABA) qui est un neurotransmetteur

Les dérivés de la tyrosine :



Chez les mammifères, citons :

- l'ornithine (homologue à 5 carbones de la lysine) et la citrulline (dérivé amide du précédent) sont des intermédiaires du cycle qui convertit dans le foie l'ion ammonium en urée.
- la taurine est un acide 2-aminosulfonique (oxydation de la cystéine en acide cystéique puis décarboxylation de ce dernier) qui entre dans la composition des sels biliaires nécessaires à l'absorption des lipides.

Chez les plantes, champignons et bactéries, citons :

- le poison β -cyanoalanine, sécrété par des plantes et des champignons
- la bactérie *Streptomyces* sécrète des antibiotiques comme la D-cyclosérine et l'azasérine, cette dernière est employée dans des thérapeutiques pour ses propriétés antifongique et antitumorale.

3. Les peptides

Les chaînes peptidiques sont le produit de la polymérisation covalente des aminoacides par une **liaison peptidique**. Elles diffèrent par le nombre, la nature et l'ordre des aminoacides. On définit arbitrairement :

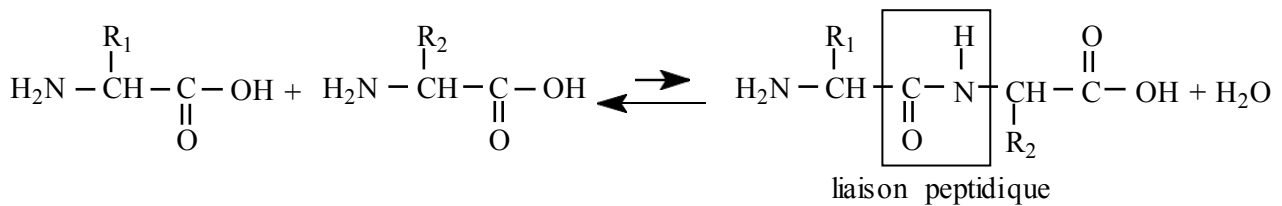
- peptide : enchaînement d'un nombre d'acide aminé inférieur à 50. Parmi ceux-ci, on parle d'oligopeptide pour un nombre d'acides aminés inférieur à 10 et de polypeptide pour un nombre supérieur à 10.

- protéine : enchaînement d'un nombre d'acides aminés au-delà de 50.

La liaison peptidique

Type de liaison

La liaison est de type amide substituée : élimination d'eau entre les groupes α -COOH et α - NH₂ de deux acides aminés. Cette liaison amide lie les deux carbones C α des deux acides aminés.



Cette réaction est thermodynamiquement en faveur de l'hydrolyse :

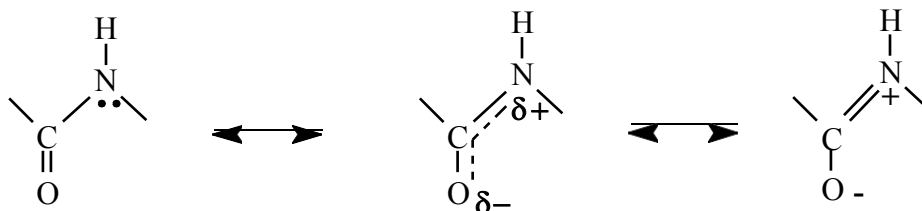
- la synthèse chimique d'un peptide n'est possible qu'en milieu anhydre et une forme activée de l'acide

- la biosynthèse cellulaire, en milieu aqueux, est dépendante d'une "activation énergétique" qui a lieu par un couplage avec l'hydrolyse d'un "donneur d'énergie" (ATP).

Cette liaison, une fois formée, est très **stable** et son hydrolyse spontanée est quasiment nulle.

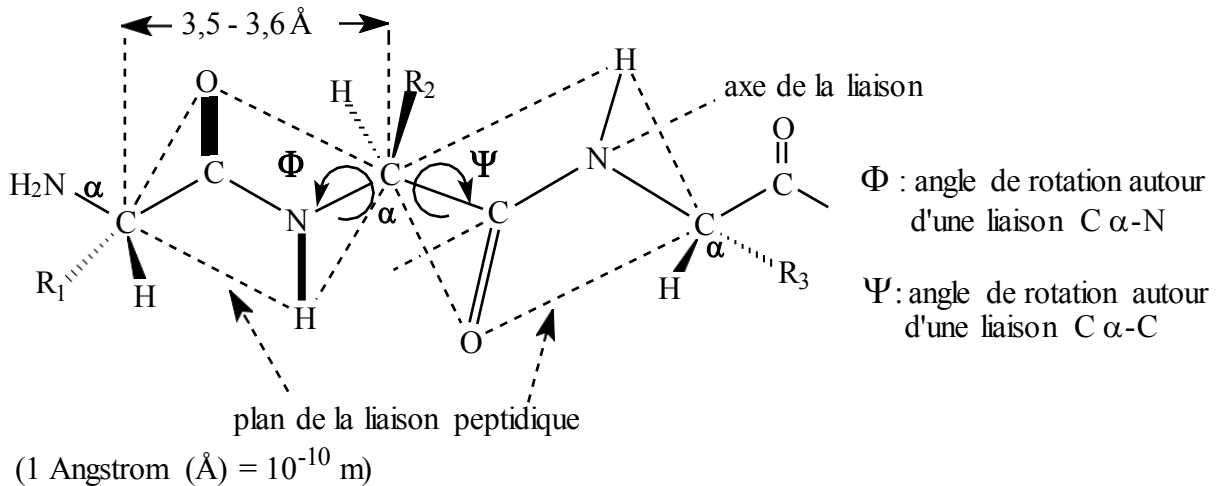
Géométrie de la liaison peptidique

Les électrons π du groupe carbonyle et le doublet électronique libre de l'azote sont très proches. La résonance de ces électrons donne au groupe peptidique des structures intermédiaires entre deux formes mésomères.



Cette liaison est intermédiaire entre une simple et une double liaison qui implique les propriétés suivantes :

- la structure du groupe peptidique est **rigide** : les 6 atomes sont **coplanaires**. Les angles des liaisons pour le carbone et pour l'azote avec leurs substituants sont de 120° .
- les 2 carbones $C\alpha$ se placent de part et d'autre du pont C-N dans la configuration **trans** la plus favorable thermodynamiquement
- de part et d'autre de cette structure rigide, les rotations des groupes des liaisons $C\alpha-N$ et $C\alpha-C$ sont libres et seulement limitées par l'encombrement stérique.



Les atomes d'oxygène et d'hydrogène d'une liaison peptidique sont d'excellents accepteurs et donneurs de liaisons hydrogène.

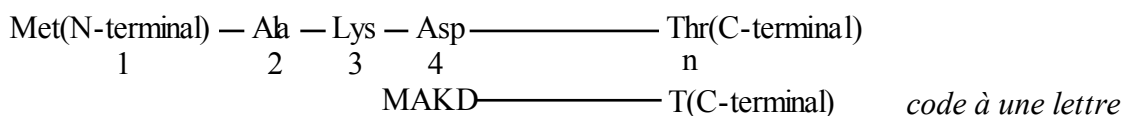
Les chaînes peptidiques et leur nomenclature

Les chaînes peptidiques sont vectorisées : les liaisons peptidiques attachent les acides aminés dans un ordre spécifique. Les conventions sont les suivantes :

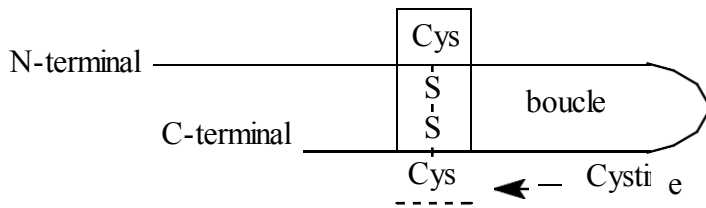
- les aminoacides engagés dans une chaîne peptidique sont appelés **résidus**. Leur nom est celui de l'acide aminé auquel on ajoute le suffixe **yl**.
- les deux aminoacides aux extrémités de la chaîne sont appelés : **N-terminal** pour celui qui a sa fonction α -aminée libre et **C-terminal** pour celui qui a sa fonction α -COOH libre
- on numérote les aminoacides en écrivant l'enchaînement de gauche à droite à partir de l'extrémité N-terminal.

On distingue trois types de peptides :

- peptide formé d'une seule chaîne (monocaténaire) et linéaire



- peptide formé d'une seule chaîne (monocaténaire) et cyclique

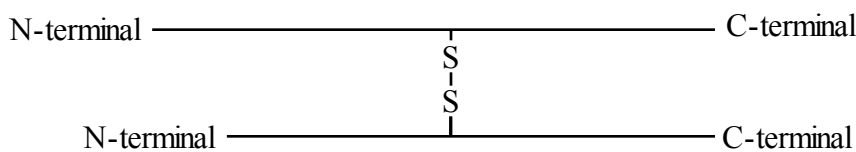


Pont disulfure
Oxydation de thiol

Les 2 cystéines impliquées
sont appelées demi-cystine

Une liaison covalente (pont S-S) intra-chaîne est présente et réalisée par l'oxydation de deux fonctions thiol de deux cystéines

- peptide formé de plusieurs chaînes (polycaténaire)



Pont disulfure
inter-chaînes

Une liaison covalente (pont S-S) inter-chaînes est présente et réalisée par l'oxydation de deux fonctions thiol de deux cystéines appartenant à deux chaînes peptidiques différentes

Ionisation des peptides

Les groupes ionisables d'un peptide sont :

- le groupe α - NH^+ de l'acide aminé N-terminal

- le groupe α -COOH de l'acide aminé C-terminal

- les groupes ionisables des chaînes latérales des différents acides aminés du peptide, à l'exception toutefois des thiols (cystéine) ayant subi une oxydation et qui sont engagés dans une liaison S-S.

Nous allons voir quelques exemples de calcul du pI d'un peptide de détermination des formes prépondérantes en équilibre dans une solution à un pH donné. Pour traiter ces problèmes, nous comptabiliserons toutes les fonctions ionisables, nous les classerons par leurs valeurs croissantes de pK et nous écrirons les équilibres successifs. Nous **supposerons l'indépendance complète des différents groupes ionisables entre eux.**

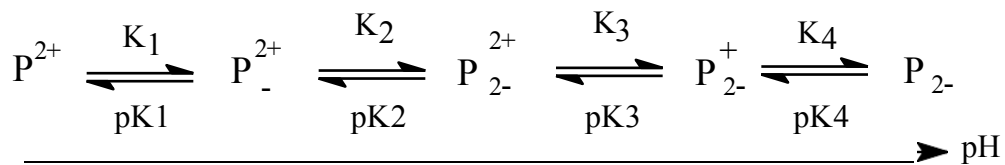
Exemple 1

Peptide Ala-Lys-Leu-Met-Asp-Ile (AKLMDI)

Fonctions ionisables :

- Ile _α -COOH	pK1 = 2,6	K ₁ = 2,5 10 ⁻³ M
- Asp _R -COOH	pK2 = 3,9	K ₂ = 1,25 10 ⁻⁴ M
- Ala _α -NH ₃ ⁺	pK3 = 9,8	K ₃ = 1,6 10 ⁻¹⁰ M
- Lys _R -NH ₂ ⁺	pK4 = 10,5	K ₄ = 3,16 10 ⁻¹¹ M

Considérons que les valeurs des constantes individuelles sont assez différentes pour écrire les équilibres successifs avec celles-ci. Le cas litigieux est celui des deux derniers équilibres (différence de 0,7 entre les deux pK). Les équilibres successifs s'écrivent :



pH isoélectrique

L'équation de contrainte s'écrit :

$$2[P^{2+}] + [P_{-}^{2+}] - [P_{2-}^{+}] - 2[P_{2-}] = 0 \quad (C)$$

Au lieu d'écrire le système complet décrivant ce phénomène d'équilibres, nous allons voir si nous ne pouvons pas le simplifier en sachant que la valeur du pI sera nécessairement dans une zone de pH particulière.

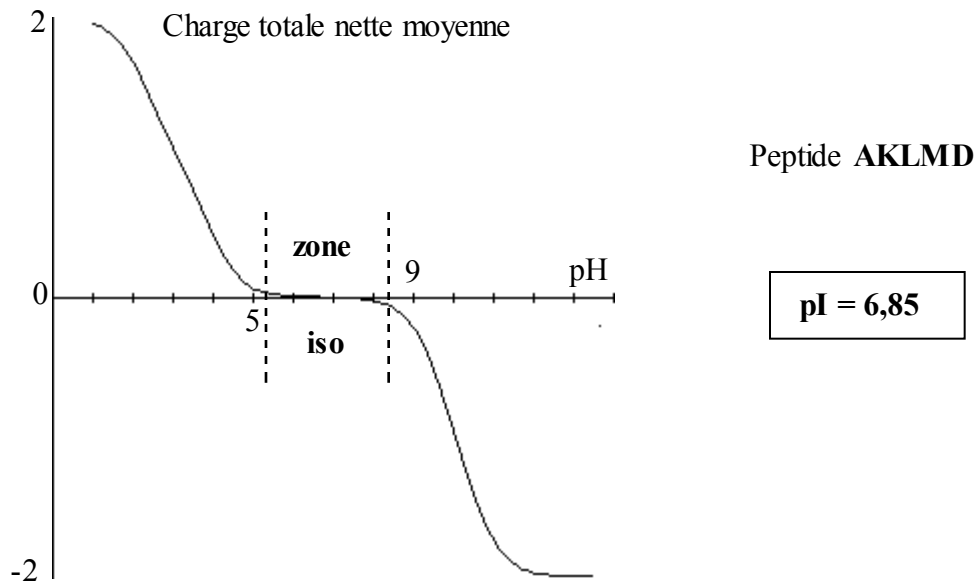
L'équation de contrainte du pH isoélectrique (C) nous indique :

- si le pH a une valeur supérieure à pK₃, les formes prépondérantes seront chargées négativement et l'équation (C) ne pourra être satisfaite
- de manière symétrique, si le pH a une valeur inférieure à pK₂, les formes prépondérantes seront chargées positivement et l'équation (C) ne pourra être satisfaite

La valeur du pI sera comprise entre pK₂ et pK₃, vu d'une part, les différences entre pK₂ et pK₁ et d'autre part, les différences entre pK₃ et pK₄, les équilibres concernés seront uniquement ceux qui entourent la forme qui a une charge nette moyenne nulle P₂₋²⁺. Ecrivons le système approximé d'équations :

$$\left\{ \begin{array}{l} K_2 = \frac{[P_{2-}^{2+}][H^+]}{[P_{-}^{2+}]} \\ K_3 = \frac{[P_{2-}^{+}][H^+]}{[P_{2-}^{2+}]} \\ [P_{-}^{2+}] - [P_{2-}^{+}] = 0 \end{array} \right. \quad \text{d'où} \quad pI = \frac{pK_2 + pK_3}{2} = \frac{3,9 + 9,8}{2} = 6,85$$

Courbe de titrage : les valeurs de pK2 et de pK3 sont très différentes, on note donc une zone de pH isoionique.



Pourcentage des formes prépondérantes à pH = pI (6,85)

$$(S1) \left\{ \begin{array}{l} K_1 = \frac{[P_{2-}^{2+}][H^+]}{[P_{2-}^{2+}]} \\ K_2 = \frac{[P_{2-}^{2+}][H^+]}{[P_{2-}^-]} \\ K_3 = \frac{[P_{2-}^+][H^+]}{[P_{2-}^{2+}]} \\ K_4 = \frac{[P_{2-}^+][H^+]}{[P_{2-}^+]} \\ [P_{2-}^{2+}] + [P_{2-}^-] + [P_{2-}^{2+}] + [P_{2-}^-] + [P_{2-}^+] = [P_T] \text{ concentration totale} \end{array} \right.$$

En exprimant chacune des formes en fonction de P_{2-}^{2+} à l'aide des quatre premières équations d'équilibre et en les remplaçant dans l'équation de conservation, on obtient :

$$[P_{2-}^{2+}] \left(1 + \frac{K_3}{[H^+]} + \frac{[H^+]}{K_2} + \frac{[H^+]^2}{K_2 K_1} + \frac{K_4 K_3}{[H^+]^2} \right) = [P_T] \quad (P1)$$

Les deux derniers termes représentent les contributions des formes P_{2-}^{2+} et P_{2-}^- dont les valeurs respectives sont : $6,2 \cdot 10^{-6}$ et $2,57 \cdot 10^{-7}$ pour une valeur de pH de 6,85

$[H^+] = 1,4 \cdot 10^{-7} \text{ M}$), termes absolument négligeables. L'équation (P1) peut être approximée par l'équation :

$$[P_{2-}^{2+}] \left(1 + \frac{K_3}{[H^+]} + \frac{[H^+]}{K_2} \right) = [P_T]$$

Le calcul nous donne un pourcentage de la forme P_{2-}^{2+} de 99,7 % et pour les formes P_{2-}^{2+} et P_{2-}^+ un pourcentage de 0,15 %

Remarque :

Nous pouvons simplifier le calcul en écrivant un système d'équations déjà approximé.

Compte tenu des remarques précédentes pour le calcul du pI, les formes significatives à ce pH seront P_{2-}^{2+} , P_{2-}^{2+} et P_{2-}^+ , les équilibres concernés sont les deuxième et troisième : le système approximé d'équations s'écrit :

$$\begin{cases} K_2 = \frac{[P_{2-}^{2+}][H^+]}{[P_{2-}^+]} \\ K_3 = \frac{[P_{2-}^+][H^+]}{[P_{2-}^{2+}]} \\ [P_{2-}^{2+}] + [P_{2-}^{2+}] + [P_{2-}^+] = [P_T] \text{ concentration totale de la protéine} \end{cases}$$

On obtient les mêmes valeurs numériques : pourcentage de la forme P_{2-}^{2+} de 99,7 % et pour les formes P_{2-}^{2+} et P_{2-}^+ un pourcentage de 0,15 %.

Pourcentage des formes prépondérantes à pH 10,15

Reprenons le système d'équations (S1), et calculons chacune des formes en fonction de P_{2-}^+ à l'aide des quatre premières équations d'équilibre. En les remplaçant dans l'équation de conservation, on obtient :

$$[P_{2-}^+] \left(1 + \frac{K_4}{[H^+]} + \frac{[H^+]}{K_3} + \frac{[H^+]^2}{K_3K_2} + \frac{[H^+]^3}{K_3K_2K_1} \right) = [P_T] \quad (P2.1)$$

Les deux termes qui représentent les contributions des formes P_{2-}^{2+} et P_{2-}^{2+} sont respectivement égaux à $2,45 \cdot 10^{-7}$ et $6,8 \cdot 10^{-15}$, contribution absolument négligeable.

L'équation (P2.1) se réduit à :

$$[P_{2-}^+] \left(1 + \frac{K_4}{[H^+]} + \frac{[H^+]}{K_3} \right) = [P_T] \quad (P2.2)$$

où le terme $\frac{K_4}{[H^+]}$ représente la contribution de la forme P_{2-} , et le terme $\frac{[H^+]}{K_3}$ celle de la forme P_{2-}^{2+} .

Calcul numérique : $P_{2-}^{2+} = 23,5\%$ $P_{2-}^+ = 53\%$ et $P_{2-} = 23,5\%$ (R1)

Comme précédemment, on aurait pu approximé le système d'équations avant de le résoudre : pour une telle valeur de pH (10,15) celui-ci s'écrit :

$$\left\{ \begin{array}{l} K_3 = \frac{[P_{2-}^+][H^+]}{[P_{2-}^{2+}]} \\ K_4 = \frac{[P_{2-}]}{[P_{2-}^+][H^+]} \\ [P_{2-}^{2+}] + [P_{2-}^+] + [P_{2-}] = [P] \end{array} \right. \quad T$$

et on retrouve bien l'équation (P2.2)

La charge totale nette moyenne du peptide est égale à -1 à ce pH de 10,15 ($0,53 + (2 \times 0,235)$).

Remarque importante :

Nous avons vu au paragraphe 2.4.1 (Remarque générale) que lorsque les constantes individuelles ont des valeurs trop proches (différence pour les pK de l'ordre de 1), les constantes successives ne peuvent pas être remplacées par les constantes individuelles (erreur de 10% pour une différence pour les pK de l'ordre de 1). Ici nous avons les valeurs suivantes des pK : 9,8 et 10,5. Reconnaissons le calcul précédent avec les constantes successives.

$$K_3 = K_3 + K_4 \quad \text{et} \quad K_4 = \frac{K_3 K_4}{K_3 + K_4} \quad \text{d'où :} \quad \left\{ \begin{array}{l} K_3^s = 1,9 \cdot 10^{-10} \text{ M} \quad pK_3^s = 9,72 \\ K_4^s = 2,6 \cdot 10^{-11} \text{ M} \quad pK_4^s = 10,58 \end{array} \right.$$

En remplaçant ces valeurs dans l'équation (P1), le calcul numérique donne : $P_{2-}^{2+} = 21,5\%$ $P_{2-}^+ = 57\%$ et $P_{2-} = 21,5\%$, à comparer avec les résultats (R1).

Exemple 2

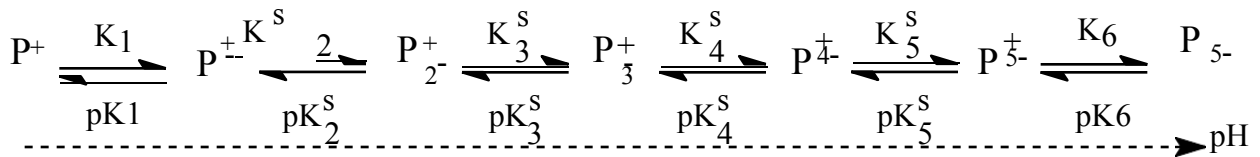
Peptide Ala-Glu-Glu-Met-Glu-Glu-Ile (AEEMEEI)

Fonctions ionisables :

- Ile $_{\alpha}$ -COOH pK1 = 2,6 K₁ = 2,5 · 10⁻³ M
- 4 Glu $_R$ -COOH pK = 4,3 K = 5 · 10⁻⁵ M
- Ala $_{\alpha}$ -NH⁺₃ pK6 = 9,8 K₆ = 1,6 · 10⁻¹⁰ M

Pour écrire les équilibres successifs, il faut expliciter les constantes successives de dissociation des 4 groupes COOH des chaînes latérales des résidus acide glutamique. Pour des fonctions ionisables **indépendantes et identiques**, $K_j = \frac{1}{j} K$ (ici n=4), d'où :

$$\begin{aligned} K_2^s &= 4K = 2 \cdot 10^{-4} \text{ M} & \text{pK}_2^s &= 3,7 \\ K_3^s &= \frac{3}{2} K = 7,5 \cdot 10^{-5} \text{ M} & \text{pK}_3^s &= 4,125 \\ K_4^s &= \frac{2}{3} K = 3,33 \cdot 10^{-5} \text{ M} & \text{pK}_4^s &= 4,48 \\ K_5^s &= \frac{1}{4} K = 1,25 \cdot 10^{-5} \text{ M} & \text{pK}_5^s &= 4,91 \end{aligned}$$



pH isoélectrique

L'équation de contrainte s'écrit :

$$[\text{P}^+] - [\text{P}_2^-] - 2[\text{P}_3^-] - 3[\text{P}_4^-] - 4[\text{P}_5^-] - 5[\text{P}_{5-}] = 0 \quad (\text{C2})$$

Au lieu d'écrire le système complet décrivant ce phénomène d'équilibres, comme précédemment, simplifions le système d'équations en sachant que la valeur du pI sera nécessairement dans une zone de pH particulière.

L'équation de contrainte du pH isoélectrique (C2) nous indique :

- la valeur du pI sera comprise entre pK1 et pK₂^s; les équilibres concernés sont ceux qui entourent la forme de charge nette nulle.
- les différences de valeur du pI avec les valeurs des pK₄^s, pK₅^s et pK6 nous indiquent que

seuls les trois premiers équilibres seront concernés.

Le système approximé d'équations s'écrit :

$$\left\{ \begin{array}{l} K_1 = \frac{[\text{P}^+][\text{H}^+]}{[\text{P}^{\pm}]} \\ K_2 = \frac{[\text{P}_2^+]}{[\text{P}_3^-][\text{H}^+]} \\ K_3 = \frac{[\text{P}_3^+]}{[\text{P}_4^-][\text{H}^+]} \end{array} \right. \quad \text{d'où} \quad [\text{H}^+]^3 - K_1 K_2 [\text{H}^+] - 2K_1 K_2 K_3 = 0 \quad (\text{C3})$$

$$[\text{P}^+] - [\text{P}_2^-] - 2[\text{P}_3^-] = 0$$

Nous pouvons simplifier l'équation (C3) en constatant que $K_1 K_2 [H^+] \gg 2K_1 K_2 K_3$ d'au

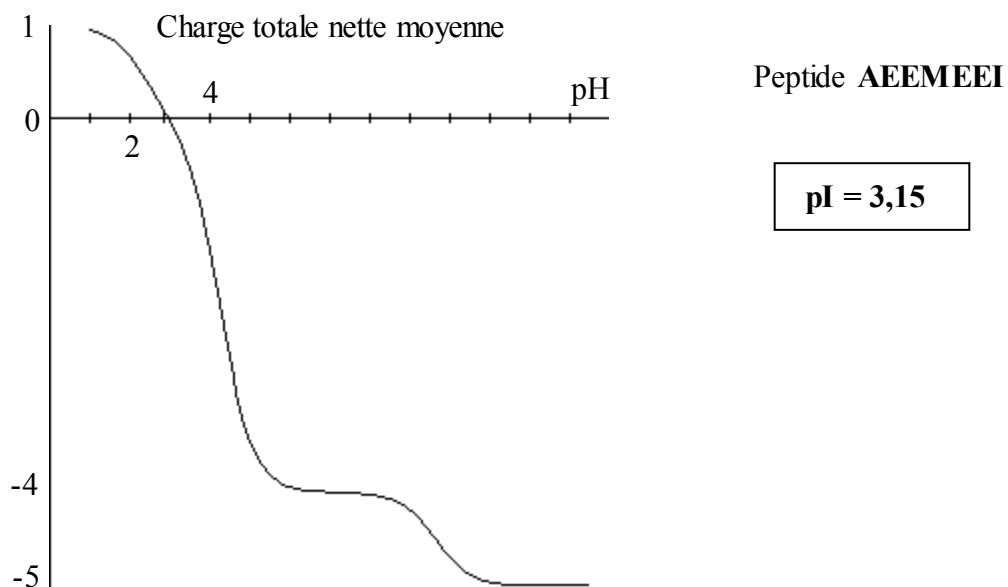
moins un facteur 10 et obtenir la valeur du pI avec une précision supérieure à 5% :

L'équation (C3) se simplifie en :

$$[H^+]^3 - K_1 K_2 [H^+] = 0 \quad \text{d'où} \quad pI = \frac{pK_1 + pK_2}{2} = \frac{2,6 + 3,7}{2} = 3,15$$

Si on avait conservé la valeur de la constante individuelle pour les fonctions ionisables des acides glutamiques, on aurait trouvé : $pI = 3,45$

Courbe de titrage : les valeurs de pK_1 et de pK_2 sont assez proches, il n'y a pas de zone isoionique. On note une zone de pH où la charge est constante et égale à -4, elle est située entre pK_5 et pK_6 qui ont des valeurs très différentes.



Pourcentage des formes prépondérantes à pH = pI (3,15)

$$(S2) \left\{ \begin{array}{l} K_1 = \frac{[P^+][H^+]}{[P^+]} \\ K_2 = \frac{[P^+][H^+]}{[P^+]} \\ K_3^s = \frac{[P_{3-}^+][H^+]}{[P_{2-}^+]} \\ K_4^s = \frac{[P_{4-}^+][H^+]}{[P_{3-}^+]} \\ K_5^s = \frac{[P_{5-}^+][H^+]}{[P_{4-}^+]} \\ K = \frac{[P_{5-}][H^+]}{[P_{4-}^+]} \\ [P^+] + [P_{3-}^+] + [P_{2-}^+] + [P_{3-}^+] + [P_{4-}^+] + [P_{5-}^+] + [P_{5-}] = [P]_T \end{array} \right.$$

A l'aide des 6 premières équations, exprimons les différentes formes en fonction de la forme P_{2-}^+ et remplaçons celles-ci dans l'équation de conservation :

$$(P3) [P_{2-}^+] \left(1 + \frac{[H^+]}{K_1} + \frac{2}{[H^+]} + \frac{2^2}{[H^+]^2} + \frac{2^3}{[H^+]^3} + \frac{2^4}{[H^+]^4} + \frac{2^5}{[H^+]^5} \right) = [P]_T$$

Pour la valeur de pH de 3,15 ($[H^+] = 7 \cdot 10^{-4}$ M), les termes des contributions des différentes formes P_{3-}^+ , P_{4-}^+ , P_{5-}^+ et P_{5-} sont négligeables (pour le plus élevé : 3%) et l'équation (P3) se

réduit à :

$$[P_{2-}^+] \left(1 + \frac{[H^+]}{K_1} + \frac{2}{[H^+]} \right) = [P]_T$$

Le calcul nous donne un pourcentage de la forme P_{2-}^+ de 64 % et pour les formes P_{3-}^+ et P_{4-}^+ un pourcentage de 18 %.

Pourcentage des formes prépondérantes à pH = 4,48

Reprenons le système d'équations (S2), et calculons chacune des formes en fonction de P_{4-}^+ à l'aide des six premières équations d'équilibre. En les remplaçant dans l'équation de conservation, on obtient :

$$[P_{4-}] \left(1 + \frac{[H^+]}{K_4} + \frac{[H^+]^2}{K_3 K_4} + \frac{[H^+]^3}{K_2 K_3 K_4} + \frac{[H^+]^4}{K_1 K_2 K_3 K_4} \right) = [P_T] \quad (P4)$$

Les termes représentant les contributions des formes P_{5-} , P^+ et P^+ sont négligeables (à 5% près pour le terme le plus élevé) pour une valeur de pH de 4,48 ($[H^+] = 3,11 \cdot 10^{-5} \text{ M}$).

L'équation (P4) se simplifie en :

$$[P_{4-}] \left(1 + \frac{[H^+]}{K_4} + \frac{[H^+]^2}{K_3 K_4} \right) = [P_T]$$

Le calcul nous donne les valeurs suivantes :

$P_{2-} = 15,4 \%$	$P_{3-} = 35,6 \%$	$P_{4-} = 36 \%$	$P^+ = 13 \%$
--------------------	--------------------	------------------	---------------

La charge totale nette moyenne de ce peptide à pH 4,48 est de -2,46.

Détermination de la structure d'un peptide

La détermination de la structure primaire d'un peptide est conduite en deux étapes :

- 1) détermination de la composition en aminoacides
- 2) détermination de l'ordre des enchaînements des résidus

Ces deux étapes ont comme point commun l'hydrolyse de la liaison peptidique.

Hydrolyse de la liaison peptidique

La liaison peptidique est très stable, son hydrolyse spontanée est quasiment nulle.

Hydrolyse chimique complète

L'action de l'acide chlorhydrique (HCl) 6M sur un peptide, à ébullition pendant au moins 24 heures, aboutit à un **hydrolysate** contenant les aminoacides avec toutefois les restrictions suivantes :

- l'acide tryptophane est entièrement détruit
- les amides (Asn, Gln) sont hydrolysés en ammoniac et acides correspondants (Asp, Glu)
- certains aminoacides (Tyr, Ser, Thr) peuvent être partiellement détruits (un temps d'hydrolyse plus faible permet de résoudre le problème).

Hydrolyse chimique spécifique

Certains réactifs hydrolysent une liaison peptidique avec une spécificité sur un des aminoacides participant à la liaison :

- le bromure de cyanogène (BrCN) hydrolyse la liaison peptidique du côté carboxyle de la méthionine : cette dernière devient alors un résidu C-terminal transformé en résidu homosérine lactone

- le 2-nitro-5-thiocyanobenzoate (NTCB) hydrolyse la liaison peptidique du côté amine de la cystéine.

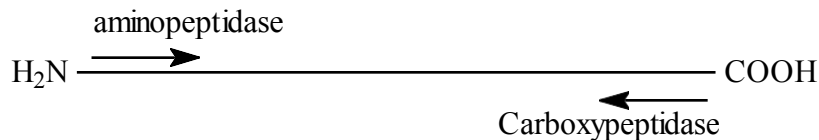
Hydrolyse enzymatique

L'hydrolyse des liaisons peptidiques peut être réalisée par des enzymes protéolytiques (ou **protéases** ou encore peptidases) qui sont des hydrolases. La spécificité principale de ce groupe d'enzymes est l'hydrolyse des liaisons peptidiques.

Leur spécificité secondaire permet de les classer en deux groupes :

- exopeptidase

L'enzyme n'hydrolyse que la première liaison peptidique (**aminopeptidase**) ou la dernière liaison peptidique (**carboxypeptidase**) en libérant l'acide aminé terminal. Bien évidemment, le processus recommence sur le peptide amputé d'un acide aminé (un temps d'hydrolyse court permet de libérer un seul acide aminé).

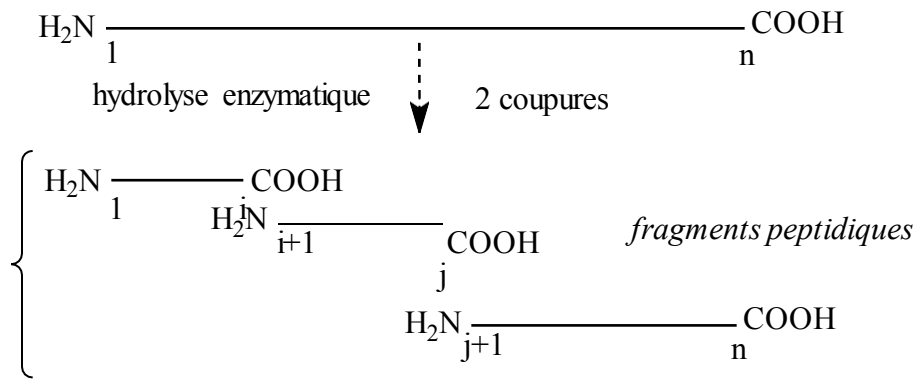


Certaines exopeptidases ont des spécificités secondaires particulières. Exemples :

Type	Nom	Source	Particularité
<i>aminopeptidase</i>	leucine aminopeptidase	rein de porc	- sauf Pro
	aminopeptidase M	rein de porc	
	aminopeptidase K	moisissure	- autres que basiques - arrêtée par Pro
<i>carboxypeptidase</i>	carboxypeptidase A	pancréas de boeuf	- autres que basiques - arrêtée par Pro
	carboxypeptidase B	pancréas de boeuf	
	carboxypeptidase C	feuille de citronnier	
	carboxypeptidase P	moisissure	- sauf Ser et Gly

- endopeptidase

L'enzyme hydrolyse des liaisons peptidiques internes entre deux acides aminés i , $(i+1)$. Il peut être spécifique du résidu en position i ou $(i+1)$. L'hydrolyse d'un peptide par une endopeptidase donnera plusieurs fragments peptidiques : si on a m coupures (m liaisons peptidiques hydrolysées), le peptide sera dégradé en $(m+1)$ fragments peptidiques.



Exemples d'endopeptidases avec leurs spécificités :

Enzyme	Source	résidu i	résidu (i+1)	particularité
trypsine	pancréas de boeuf	Arg, Lys		sauf (i+1) = Pro
chymotrypsine	pancréas de boeuf	Phe, Tyr, Trp		
Sa protéase	<i>Staphylococcus aureus</i>	Asp, Glu		
Fm protéase	<i>Flavobacterium meningosepticum</i>	Pro		
thermolysine	<i>Bacillus thermoprotéolyticus</i>		Ala, Val, Leu Ile, Met	

Remarque : on peut dire que la thermolysine est une endopeptidase spécifique des aminoacides Ala, Val, Leu, Ile, Met avec une coupure du côté amino de la liaison peptidique. Les autres enzymes sont spécifiques des aminoacides indiqués avec une coupure du côté carboxyle de la liaison peptidique.

Détermination de la séquence

La détermination de la structure primaire d'un peptide obéit en général à la stratégie suivante :

1) Identification des aminoacides terminaux

- l'extrémité N : on utilise en général le chlorure de dansyl et après hydrolyse chimique complète du peptide, on identifie le dansyl-aminoacide (voir paragraphe 2.3.2). La présence de lysine dans le peptide va perturber cette méthode puisque la chaîne latérale de ce résidu porte un groupe $-NH_2$ qui réagira avec le DNS-Cl.

- l'extrémité C : on utilise en général une dégradation limitée à l'aide des carboxypeptidases

2) Résolution des problèmes

- l'absence d'acide aminé terminal indiquera que ceux-ci ont été modifiés et sont sûrement cycliques ou semi-cycliques

- la présence de ponts disulfure intra-chaîne posera des problèmes quant à la composition en acide aminé. La réduction de ceux-ci par un thiol, suivie d'une alkylation supprimera les problèmes

- la présence de plusieurs groupements N et C-terminaux indiquera que le peptide est formé de plusieurs chaînes liées des ponts disulfure. La réduction de ceux-ci par un thiol, suivie d'une alkylation, séparera les différentes chaînes.

3) Dégradation des peptides en fragments

En utilisant les différentes méthodes de coupure de liaisons peptidiques qui fragmentent le peptide et en répétant l'étape 1 (identification des acides aminés terminaux), on peut positionner certains acides aminés dans la séquence. En répétant le processus 3 et 1 de manière itérative sur chacun des fragments, on peut arriver à la détermination complète de la structure primaire du peptide original.

Détermination de la séquence : la dégradation récurrente d'Edman

On a vu au paragraphe 2.3.2 la réaction d'Edman avec l'acide aminé terminal : formation d'un PTC-peptide à pH 9. Par un changement de pH (légèrement acide), il y a cyclisation et libération d'un dérivé PTH-acide aminé et d'un peptide amputé de son acide aminé N-terminal. En répétant ces cycles :

- action du PTC à pH 9 pour donner un PTC-peptide
- puis libération du PTH-acide aminé par passage à un pH acide, et du peptide (n-1) amputé du côté N-terminal,

on a, à chaque cycle, la libération de l'acide aminé suivant dans la séquence du peptide original.

Des appareils (séquenceur) sont capables d'effectuer ces cycles automatiquement. Toutefois l'analyse est techniquement limitée à des séquences dont le nombre d'acides aminés est inférieur à 200.

Avec cette technique, il faudra résoudre le problème d'acides aminés qui sont cyclisés sur l'acide aminé N-terminal.

Détermination de la séquence : technique récente

On utilise de plus en plus la **spectrométrie de masse** (déviations d'une particule chargée dans un champ électrique). Un spectromètre de masse est un appareil capable de mesurer avec grande précision le rapport masse/(charge électrique) de molécules. Cette technique comporte deux étapes :

- le bombardement d'un peptide par des atomes "rapides" d'un gaz rare (argon ou xénon) produit l'espèce peptidique ionique $(P, H)^+$ isolée du milieu par un premier spectromètre,
- l'ion peptidique est ensuite bombardé par des atomes neutres d'hélium qui provoquent des ruptures à partir des deux extrémités. Une série de fragments est obtenue, et ces derniers sont soumis à une analyse par un second spectromètre de masse.

Les résultats sont analysés et un traitement informatique reconstitue la séquence primaire.

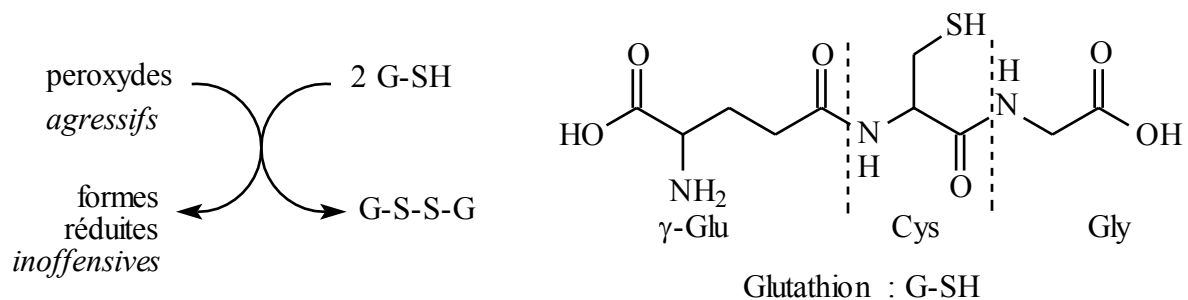
Cette technique ne peut analyser que des peptides d'environ 30 aminoacides, il faudra dans le cas de peptides de nombre d'acides aminés supérieur les fragmenter.

Les peptides d'intérêt biologique

Les peptides participent à la structure de membranes comme éléments de construction, jouent des rôles importants dans des activités biologiques. En voici quelques exemples :

Peptides à rôle physico-chimique

Le glutathion (tripeptide γ Glu-Cys-Gly : liaison de Glu par le carbonyle γ) joue un rôle central dans la défense cellulaire contre le dioxygène et ses dérivés actifs. C'est un couple Red-Ox très efficace contre les peroxydes :



Citons un peptide synthétique très employé comme édulcorant : l'aspartame. C'est un dipeptide dont l'extrémité COOH est estérifiée : aspartyl-phénylalaninylméthylester

Peptides à activité de médiateur

Cette famille de peptides se retrouve dans les grandes de fonctions de communication dans le système endocrine, la transmission et la modulation nerveuses, la motricité des vaisseaux sanguins, les réactions inflammatoires. Citons quelques exemples :

Les peptides hormonaux

- Hypothalamus contient des neurones, à fonction endocrine, qui sécrètent 2 nonapeptides à activité hormonale :

- l'**ocytocine** qui déclenche les contractions des muscles lisses (utérus pour l'accouchement, glande mammaire pour la réjection du lait)

- la **vasopressine** : effet antidiurétique au niveau du rein (action hypertensive comme médicament)

- d'autres neurones produisent des peptides activateurs (**libérines**) ou inhibiteurs (**inhibines**) de la sécrétion d'hormones par le lobe antérieur de l'hypophyse.

- le lobe antérieur de l'hypophyse biosynthétise et sécrète de nombreuses hormones peptidiques : citons l'ACTH (**hormone adrénocorticotrope**), polypeptide de 39 aminoacides.

- le pancréas endocrine sécrète :

- l'**insuline**, polypeptide formé de deux chaînes (une de 21 et une de 30 aminoacides), qui régule le métabolisme du glucose (hypoglycémie)

- le **glucagon**, peptide de 29 aminoacides, qui provoque une augmentation de la glycémie (hyperglycémie)

- la muqueuse duodénale produit la **sécrétine**, peptide de 27 aminoacides, qui stimule le pancréas exocrine au cours de la digestion

Les neuropeptides

Ce sont deux familles de molécules, ligands de récepteurs cérébraux : la famille des **enképhalines** et la famille des **endorphines**. Ces molécules participent au contrôle de la douleur. Elles contiennent le même tétra peptide N-terminal (Tyr-Gly-Gly-Phe).

Les peptides vasomoteurs

Citons des peptides vasoconstricteurs :

- l'**angiotensine II** est un octapeptide

- les **endothélines**, dont la production est localisée dans le tissu pulmonaire sont des peptides de 21 aminoacides,

des peptides vasodilatateurs :

- l'oreillette du cœur sécrète une famille de **peptides natriurétiques atriaux** (ANP) qui agissent au niveau rénal en augmentant la sécrétion d'ions Na^+ et dans la diurèse (élimination d'eau) : chez l'homme, l' α -ANP est un polypeptide de 28 aminoacides.

- la **bradykinine** est un hypotenseur libéré lors d'une réaction inflammatoire.

Les molécules de l'inflammation

Une agression tissulaire déclenche une réaction de défense d'abord locale : c'est la réaction inflammatoire dont les manifestations sont douleur, chaleur, rougeur et œdème provoqués par des molécules vasodilatatrices :

- les amines **histamine** et **sérotonine**, produits de décarboxylation de l'histidine pour l'histamine et de l'hydroxytryptophane pour la sérotonine
- une famille de peptides : les **kinines** dont le prototype est la bradykinine qui est un nonapeptide. Le venin de guêpe contient un décapeptide : la glycyl-bradykinine.

Peptides antibiotiques

Certains micro-organismes synthétisent des peptides, inhibiteurs de la synthèse protéique, qui sont des "armes" de défense et de colonisation. Ce sont des peptides cycliques et qui intègrent des aminoacides D ou non-standard, propriétés qui les protègent des dégradations protéolytiques.

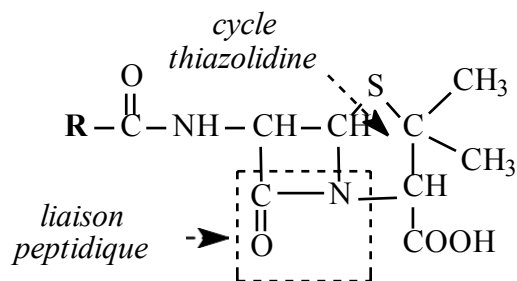
Les peptides bactériens

On peut isoler de différentes souches de *Bacillus* :

- la famille des **tyrocidines**, décapeptides cycliques comprenant une D-Phe et l'ornithine (homologue à 5 carbones de la lysine)
- la **bacitracine A** est un peptide à 12 aminoacides dont trois appartiennent à la série D (Glu, Asn, Phe) et dont un est un dérivé, l'ornithine.

Les peptides fongiques

La moisissure *Penicillium* produit la **pénicilline** qui est un dipeptide de deux dérivés d'acides aminés :



Pénicilline
R dépend de la souche de *Penicillium*

Peptides immunomodulateurs

Le champignon micromycète *Tolyocladium* sécrète des métabolites secondaires dont une famille de peptides antifongiques, les **cyclosporines**, qui sont des immunosuppresseurs utilisés dans les transplantations d'organes. Ils inhibent l'activation des lymphocytes T sans affecter les autres systèmes de défense ni la moelle osseuse.

La **cyclosporine A** est un undécapeptide cyclique (11 aminoacides) comportant de nombreux aminoacides atypiques :

- une alanine D
- 7 aminoacides N-méthylés : ce qui se traduit par 7 liaisons peptidiques modifiées

- un acide aminé non-standard (en C9), spécifique des cyclosporines : 2-amino-3-hydroxyl-4-méthyl-6octénoïque

La synthèse peptidique

Spécialité très technique, nous n'en dirons que quelques mots.

La synthèse chimique doit intégrer les difficultés suivantes :

- activation des carboxyles et blocage des groupes réactifs indésirables
- éviter les problèmes de racémisation
- élimination des différents produits à chaque étape.

Des synthétiseurs automatiques sont capables de produire des peptides jusqu'à une centaine d'acides aminés et cela en 4 à 5 jours. Une bactérie effectue une telle synthèse en 4 à 5 secondes.

Les micro-organismes sont utilisés pour obtenir des peptides particuliers comme les antibiotiques, les immunomodulateurs qui contiennent un acide aminé non-standard et des cyclisations.

4. Les protéines

L'étude des protéines et de leurs fonctions est la partie la plus importante de l'enseignement de biochimie dans le second cycle. Aussi, nous nous limiterons à quelques notions structurales qui vous seront développées ultérieurement.

Les protéines, peptides de plus de 50-100 acides aminés, sont classées en deux groupes principaux :

- les **holoprotéines** qui ne sont constituées que d'acides aminés
- les **hétéroprotéines** qui contiennent :
 - un **groupe prosthétique** minéral, métallique ou organique
 - une partie glucidique liée de manière covalente : **glycoprotéines**
 - une partie lipidique liée de manière covalente : **lipoprotéines**

Les caractéristiques spatiales des protéines sont la clé de leurs fonctions. La structure des protéines est définie à plusieurs niveaux :

- structure **primaire** : composition et séquence des acides aminés
- structure **secondaire** : c'est une structure locale qui rend compte de l'organisation spécifique d'un fragment d'acides aminés consécutifs dans l'espace
- structure **tertiaire** : c'est l'organisation spatiale complète des structures locales de la molécule
- structure **quaternaire** : c'est l'organisation de protéines oligomériques qui sont des assemblages non covalents de sous-unités (protomères)

Structure primaire

Elle est définie par la composition et l'enchaînement des aminoacides dont la détermination peut se faire par :

- des méthodes directes
- des méthodes indirectes à partir de la séquence d'un ARN messager dont la traduction est basée sur le code génétique.

Méthodes directes

Les méthodes utilisées sont celles qui sont décrites dans le paragraphe 4.4 (Détermination de la structure d'un peptide). Les méthodes les plus utilisées sont soit la dégradation récurrente d'Edman, soit la spectrométrie de masse. Toutefois celles-ci obligent à utiliser des fragments peptidiques d'une longueur inférieure à 200 aminoacides pour la première et 30 aminoacides pour la seconde : il faut donc utiliser les techniques classiques pour obtenir des fragments analysables comme l'hydrolyse de la protéine par des enzymes protéolytiques ou des hydrolyses chimiques spécifiques avec BrCN ou NTCB.

Les problèmes rencontrés seront identiques que ceux déjà décrits pour la détermination de la structure d'un peptide :

- présence de ponts disulfure intra-chaîne ou inter-chaîne
- cyclisation de l'acide aminé N-terminal pour la méthode d'Edman

Les apports de la connaissance des structures primaires

Les bases de données de protéines et les programmes informatiques disponibles permettent à partir de la structure primaire d'une protéine d'étudier :

- des familles moléculaires de protéines
- des problèmes d'évolution moléculaire
- de déterminer des "signatures" spécifiques de protéines ayant la même fonction, ce sont des motifs où pour chaque position, on peut évaluer la probabilité de présence d'un ou plusieurs aminoacides spécifiques.

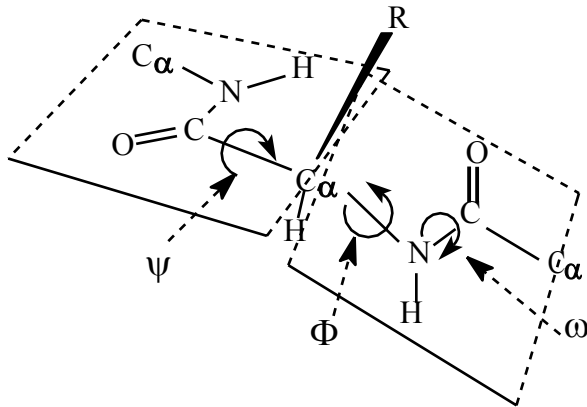
Structure secondaire

C'est l'arrangement, dans l'espace, de fragments courts de séquences en 4 grands types de forme. Les liaisons hydrogène potentielles que peuvent avoir les atomes d'oxygène (C=O) et d'hydrogène (N-H) d'une liaison peptidique et l'impact de la géométrie de la liaison peptidique sur la structure locale sont très importants.

Les contraintes de la liaison peptidique

Cet impact sera double :

- la rigidité de la liaison peptidique conduira à un squelette défini et stable
- les libertés de rotation des atomes aux extrémités de cette liaison permettront à la structure locale d'adopter différentes conformations dans l'espace.



Plans de deux liaisons peptidiques

Les propriétés de la liaison peptidique sont :

- 1) les atomes N, H, O, C α sont coplanaires
- 2) l'angle ω ne prend que deux valeurs :
 - 180° pour les C α en trans
 - 0° pour les C α en cis

Distance en Amstrong :

C-N	1,33	
C=O	1,23	N-C α 1,45
C-C α	1,52	N-H 1,0

L'angle dièdre est compté positivement quand on déplace dans le sens des aiguilles d'une montre l'atome situé en avant pour qu'il se superpose à l'atome situé en arrière (dans la représentation de Newmann). Les angles ϕ et ψ sont caractéristiques de la géométrie locale des deux plans de liaisons peptidiques consécutifs

- ϕ : rotation autour de la liaison C α -N
- ψ : rotation autour de la liaison C α -C

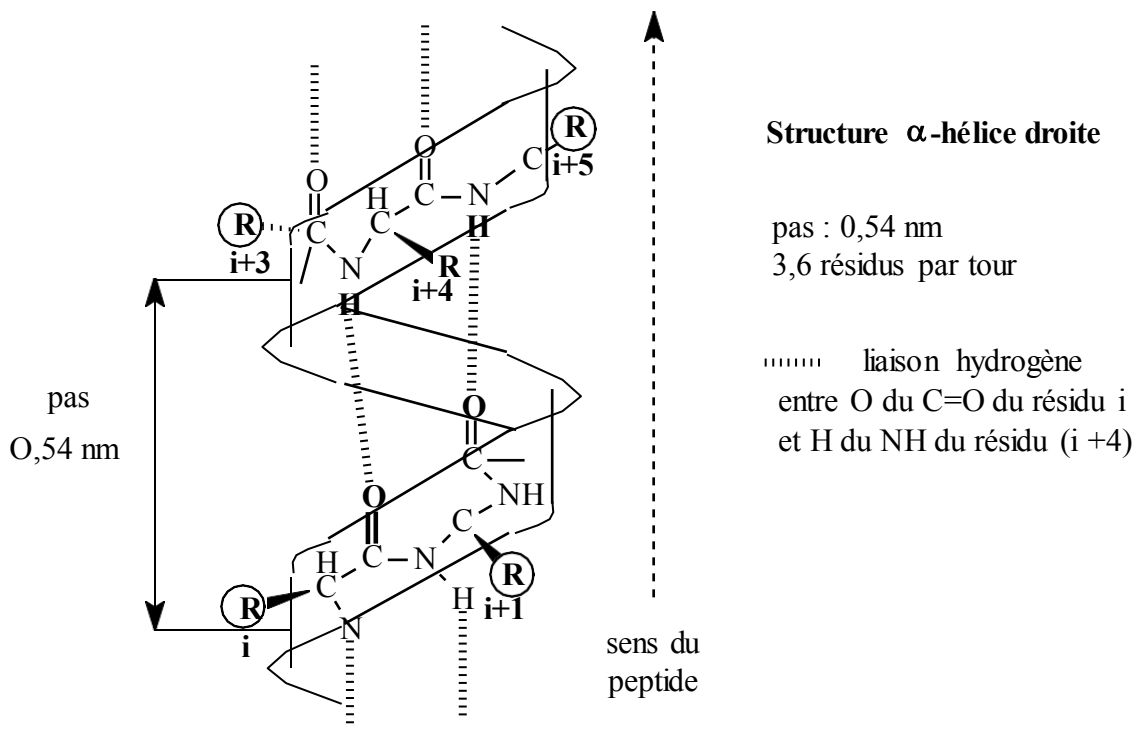
α -hélice

C'est une structure locale répétitive et relativement compacte dont les caractéristiques principales sont :

- la structure est stabilisée par les liaisons hydrogène de l'atome d'oxygène (C=O d'une liaison peptidique i avec l'atome d'hydrogène (N-H) de la liaison peptidique $(i+4)$)
- les radicaux des résidus sont à l'extérieur de l'hélice, ce qui minimise les encombrements stériques
- l'hélice ne peut s'établir qu'avec des résidus de la même série (L ou D)
- la configuration L des aminoacides privilégie un enroulement à droite (dans un enroulement à gauche, les chaînes latérales recouvrent trop le squelette de l'hélice)
- les plans des liaisons peptidiques sont parallèles à l'axe de l'hélice et forment le squelette de l'hélice.

Caractéristiques de l'hélice :

- pas : 0,54 nm par tour
- nb de résidus par tour : 3,6
- translation par résidu : 0,15 nm
- diamètre de l'hélice : 0,50 nm
- angles dièdre : $\phi = -57^\circ$ et $\psi = -47^\circ$
- la liaison hydrogène C=O-----N-H, d'une longueur de 0,286 nm, est presque parallèle à l'axe de l'hélice



Cette structure est favorisée par les résidus dont les chaînes latérales ne portent pas de charges et dont l'encombrement stérique est faible.

Certains résidus déstabilisent l'hélice par la présence de charge dans leurs chaînes latérales (Asp, Glu, Arg et Lys).

La proline est un point de rupture d'une hélice pour deux raisons :

- il n'existe pas de NH pour une liaison hydrogène
- le cycle rigide pyrrolidone engagé dans la liaison peptidique bloque la rotation, détruisant la continuité de l'hélice.

Feuillet β

Les liaisons hydrogène des atomes d'oxygène (C=O) et d'hydrogène (N-H) d'une liaison peptidique se répètent non plus sur une portion continue de la chaîne peptidique mais entre des segments différents qui peuvent appartenir à la même chaîne ou à des chaînes différentes.

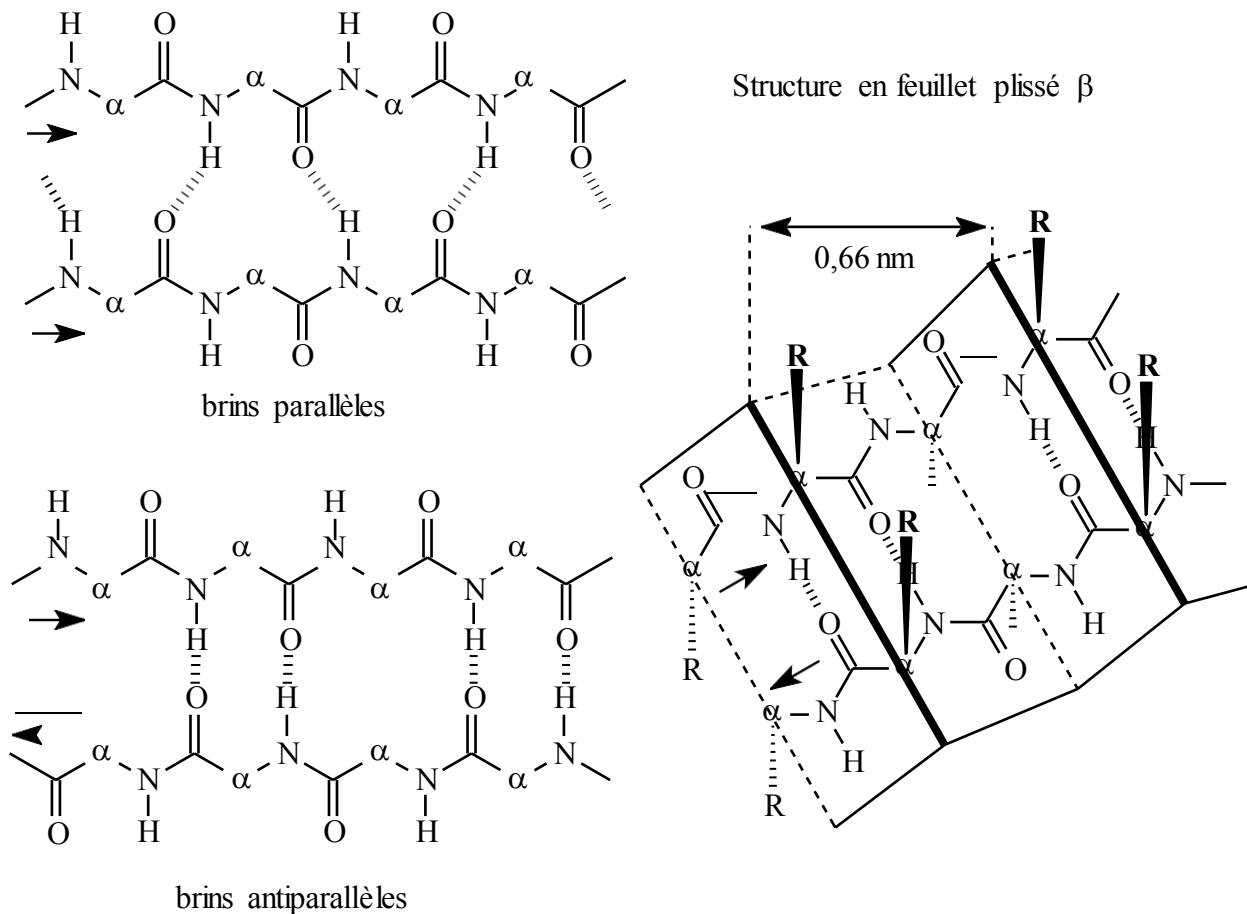
Cette structure est une structure à plat, étirée et étalée : **feuillelet β plissé**.

Le sens peptidique des deux brins adjacents peut être identique ou contraire, les feuillettes sont respectivement dits **parallèles** ou **anti-parallèles**. Ces derniers sont plus stables, les liaisons H subissant de plus faibles distortions.

Les chaînes latérales des résidus se distribuent alternativement de part et d'autre du plan moyen.

Les feuillettes β parallèles et anti-parallèles ont des caractéristiques très similaires :

- β parallèle
 - translation : 0,32 nm
 - angles dièdre : $\phi = -119^\circ$ et $\psi = 113^\circ$
- β anti-parallèle
 - translation : 0,34 nm
 - angles dièdre : $\phi = -139^\circ$ et $\psi = 135^\circ$



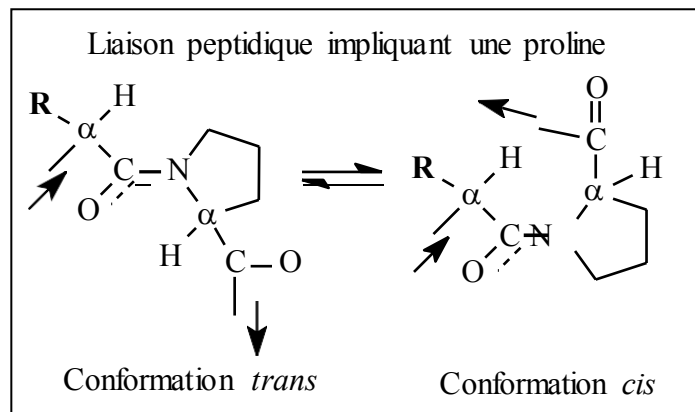
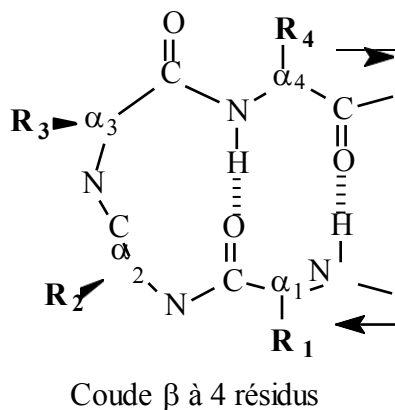
Une liaison peptidique sur deux ne participe pas à une liaison hydrogène : cela implique que nous pourrions avoir une structure en feuillet β plissé, composée de plus de deux brins.

Le coude

Certaines régions protéiques ne sont pas structurées dans des conformations périodiques. Toutefois leurs structures sont semblables par le fait qu'elles imposent un changement brusque de direction de 180°: on les appelle **coude ou tour β** (β turn). La lettre β rappelle qu'ils sont indispensables pour des feuillets de chaînes anti-parallèles.

Cette structure peut être réalisée de différentes manières, toutefois on peut dire :

- c'est un court segment peptidique de 2 à 4 résidus
- une ou deux liaisons hydrogène se forment entre le premier et le dernier résidu du coude
- la configuration de la proline est telle qu'elle provoque un changement de direction et peut donc être presque à elle seule un coude.



Dans le coude à 4 résidus, les résidus R2 et R3 avec des chaînes latérales portant des charges favoriseront une telle structure (chaînes latérales à l'extérieur et interagissant avec l'eau). Les résidus R1 et R4 avec des chaînes latérales de faible encombrement stérique favoriseront cette structure.

La pelote statistique

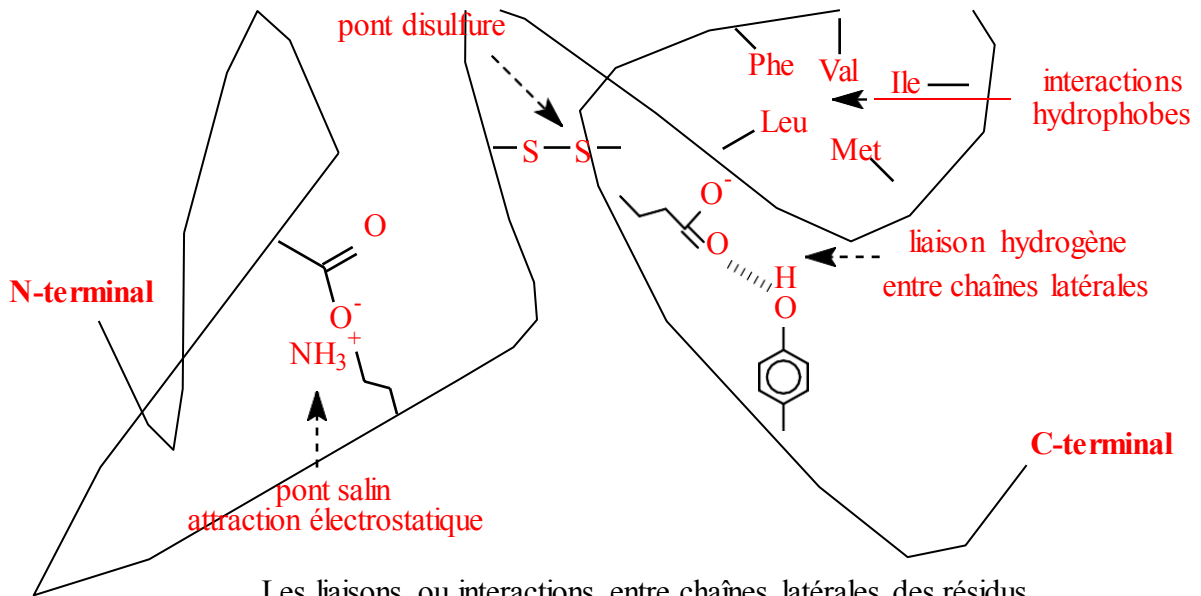
Certaines régions protéiques ne sont pas structurées dans des conformations périodiques comme l' α -hélice ou le feuillet plissé β : leur forme irrégulière est qualifiée de **pelote statistique** (random coil). Cette structure n'est pas pour autant inorganisée, elle obéit aux contraintes locales de voisinage.

Structure tertiaire

L'arrangement spatial des structures secondaires locales aboutit à une forme globale spécifique de la protéine maintenue par des interactions qui peuvent être de nature différente :

- liaison covalente : pont disulfure
- liaisons ioniques entre groupements chargés de signe opposé (pont salin)

- interactions électrostatiques entre dipôles permanents et groupements ionisés ou encore entre deux dipôles : les plus répandues sont les liaisons hydrogène
- attractions hydrophobes (force de Van der Waals entre groupes apolaires subissant des forces de répulsion par l'eau, ce qui favorise leur rapprochement)
- interactions des chaînes latérales des résidus avec le solvant : dans l'eau, les chaînes latérales polaires pourront être exposées au solvant, alors que les chaînes latérales apolaires auront tendance à "s'enfouir" dans des poches hydrophobes de la protéine.



Les liaisons ou interactions entre chaînes latérales des résidus, impliquées dans la structure tertiaire des protéines

La structure tertiaire d'une protéine est le paramètre fondamental dont dépend l'expression de ses fonctions biologiques qu'elles soient structurales ou dynamiques (protéines des membranes ou du cytosquelette ou etc, récepteurs, enzymes, etc). Cette structure est très dépendante des interactions que nous venons de décrire. Ces interactions subissent l'influence du milieu dans lequel elles se trouvent (solvant, température, pH, force ionique, agents détruisant ces interactions, etc).

La conformation **native** de la protéine est la structure tertiaire qui correspond à celle qui exprime sa fonction biologique dans les conditions physiologiques. Un traitement détruisant cette structure qui a pour conséquence de supprimer sa fonction est appelé **dénaturation** : elle aboutit à un état plus désorganisé de la molécule.

L'étude des protéines est la partie la plus importante de l'enseignement de deuxième cycle, nous n'aborderons pas les relations structure-fonction, ni les propriétés physicochimiques. Avant de nous arrêter, citons les deux grands types de structure des protéines :

- protéines **fibreuse** : elles forment des agrégats ordonnés de protéines qui sont en général sous une unique conformation principale, soit α -hélice (la kératine a est sous formes de

torsades de dimères en α -hélice) ou feuillets plissés β (les fibroïnes sont des empilements de protéines en feuillets plissés β).

- les autres protéines sont dites **globulaires** : leurs structures tertiaires sont très dépendantes des interactions. Elles peuvent comprendre des motifs de structure α -hélice ou feuillets plissés β . On les subdivise en groupes par rapport à la dominance des motifs de structure

secondaire :

- protéines à domaine α
- protéines à domaine β
- protéines α/β : elles comprennent des domaines α et des domaines β (les plus nombreuses).

5.3.1. Méthodes de détermination de structure tertiaire des protéines

Les méthodes de détermination de la structure tertiaire d'une protéine les plus utilisées et performantes sont :

- analyse par diffraction aux rayons X de cristaux de protéines : cette technique a été utilisée à partir de 1937 et permet de déterminer des structures partielles (premières études sur l'hémoglobine par Max PERUTZ). La première résolution complète de la structure d'une protéine (myoglobine de cachalot) à 0,28 nm fut réalisée par John KENDREW en 1968. Depuis, des centaines de structures tertiaires de protéines ont été résolues.

- par résonance magnétique nucléaire (RMN), limitée à des fragments d'au maximum 80 à 100 résidus.

D'autres méthodes prédictives, essentiellement informatiques, sont utilisées. Elles sont basées sur la comparaison de structure primaire d'une protéine avec d'autres protéines dont on connaît la structure tertiaire (banques de structure 3D).

Structure quaternaire

C'est l'organisation des protomères qui définit la structure quaternaire d'une protéine formée de multimères. De manière très succincte, indiquons que :

- l'association ou la dissociation des protomères permettent à ces protéines d'avoir des fonctions de signalisation ou de contrôle (l'enzyme protéine-kinase dépendante de l'AMP cyclique)

- l'interaction entre les protomères (sous-unités) est un moyen très précis de régulation : phénomène coopératif de fixation de ligands, enzymes allostériques, récepteurs membranaires

- les isoformes des protéines : elles ont les mêmes fonctions, toutefois leur structure

est dépendante du tissu soit au niveau du nombre de protomères dans l'association soit une différence au niveau du protomère.