

SV-D-2.4 poly 02

Texte du cours

Partie D : Organisation fonctionnelle des molécules du vivant

Les grandes familles biochimiques

Chapitre 4 : Acides aminés et protéines

Les protéines, hétéropolymères séquencés d'acides aminés, sont responsables du phénotype moléculaire, en permettant de réaliser un grand nombre de fonction dans la cellule et l'organisme.

I. Les acides α -aminés

Les acides aminés sont les monomères constitutifs des polypeptides (dont les protéines). Ils sont caractérisés par la diversité des fonctions chimiques qu'ils présentent, ce qui leur permet d'avoir des propriétés variées.

Il existe 20 acides aminés naturels différents codés par le code génétique au sein des organismes vivants.

A. La nature d'un acide aminé

Les AA sont des acides carboxyliques α aminés. Ils possèdent un carbone asymétrique (sauf Glycine), le carbone α , qui porte :

- Une fonction amine NH_2
- Une fonction acide carboxylique COOH
- Un H
- Un radical (R) propre à chaque AA

La fonction carboxylique est acide, et la fonction amine une base. Ces fonctions pouvant être ionisées en fonction du pH, on peut distinguer trois états ioniques possibles pour les AA (isolés) à chaîne latérale non ionisable : chargé positivement, forme zwitterion (charge positive équilibrant une charge négative), chargé négativement.

Dans le cas d'un radical ionisable, l'état d'ionisation dépend à la fois des fonctions COOH , NH_2 et de l'état d'ionisation du radical.

Remarque : l'existence de forme zwitterion ne concerne que les acides aminés isolés ; les fonctions amine et acide carboxylique des acides aminés dans les protéines ne sont pas libres.

B. La diversité des acides aminés

Les 20 AA naturels peuvent être regroupés en plusieurs groupes, en fonction des propriétés de leur radical

A (G,A,V,L,I) + Proline P : aliphatiques simples $\rightarrow$ hydrophobes, GAV petits*

B (S,T) + Tyrosine Y : hydroxylés $\rightarrow$ phosphorylables, hydrophiles (pas T) ; ST petits*

C (C,M) : soufrés $\rightarrow$ plutôt hydrophobes. Cystéine : petit, et possibilité de liaison covalente $\text{Cys-SH} + \text{HS-Cys} \rightarrow \text{Cys-S-S-Cys}$

D, D', F (E,D,N,Q,R,Q) : diacides, amides, diamminés $\rightarrow$ polaires, parfois ionisables (ED charge - ; RQ charge +)

Histidine H : globalement hydrophobe, mais possibilité de charge +

E (F,Y,W) : aromatiques $\rightarrow$ hydrophobes (= pas de solubilité dans l'eau) mais polaires pour Y et W

Proline : petit, et l'azote étant inclus dans un cycle, l'angle entre les liaisons $\text{C}\alpha\text{-N}$ et $\text{C}\alpha\text{-C}$ est imposé ! (alors que cet angle est libre pour tous les autres AA).

* petits : G,A,S,C + T,M,N,E,V,P

C. Des propriétés de certains acides aminés permettant des modifications des protéines

Certains acides aminés possèdent des chaînes latérales glycosylables :

- Chaînes latérales avec fonction hydroxyle (Sérine en particulier)
- Chaînes latérales avec fonction amine (Asparagine en particulier)

Ils permettent l'ajout de résidus osidiques sur les protéines.

Les acides aminés possédant une fonction OH sont phosphorylables :

- Sérine et Thréonine
- Tyrosine

Ils permettent la phosphorylation des protéines.

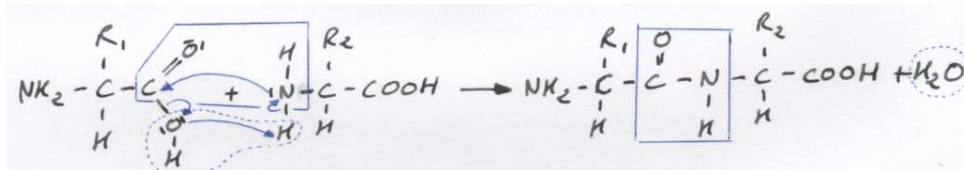
Les acides aminés possédant une chaîne latérale avec une fonction acide carboxylique portent une charge négative au pH cellulaire : aspartate, glutamate

Les acides aminés possédant une chaîne latérale avec une fonction amine portent une charge positive au pH cellulaire : arginine et lysine en particulier

II. La liaison peptidique

(voir SV-D-2.1 : la fonction amide)

Deux acides aminés peuvent s'associer par une liaison covalente : la **liaison peptidique**. Il s'agit d'une **liaison amide**, établie par réaction entre le COOH d'un AA et le NH₂ d'un autre AA.



La **liaison peptidique** est une **liaison très stable** (84 kJ/mol), **partiellement double** (à 40% ; longueur C-N de 0.130 nm contre 0.149 pour une liaison simple C-N et 0.127 pour une liaison double C=N) : **les quatre atomes C O N H forment un plan rigide**. Ces propriétés sont liées entre elles et sont dues à l'existence d'une **forme résonnante**.

Remarque : il existe de plus une isomérisation cis/trans par rapport à la position des radicaux vis-à-vis de la liaison ; dans la pratique, les contraintes d'encombrement stérique imposent une configuration.

L'association d'acides aminés conduit à la mise en place d'une chaîne orientée d'acides aminées : on a une séquence d'AA, orientée d'une extrémité N-term vers une extrémité C-term → il s'agit de la structure Primaire des protéines.

III. Les structures des protéines

Une protéine est formée par une chaîne linéaire d'au moins 50 acides aminés (en dessous, on parle de polypeptide, voire de peptide pour des courtes chaînes de moins de 15 AA). Une protéine est ainsi une molécule séquencée :

- Présence d'une séquence d'acides aminés, orientée d'une extrémité N-terminale vers une extrémité C-terminale ;
- Cette séquence d'AA est la structure primaire de la protéine ;
- Séquence colinéaire et dépendante de la séquence nucléotidique du gène correspondant.

A. Méthodes d'analyse des structures protéiques

La structure tridimensionnelle des protéines peut être étudiée par des études cristallographiques : diffraction aux rayons, RMN en particulier.

Il est de plus possible de dresser des profils hydrophatiques, qui permettent d'analyser la structure transmembranaire des protéines membranaires.

(voir aussi IV.B.1)

B. Les structures secondaires : importance des liaisons peptidiques

La structure primaire (enchaînement des AA) est la conséquence directe de la mise en place des liaisons peptidiques. À part pour la Proline, une libre rotation existe entre les carbones α des acides aminés et les liaisons peptidiques les « entourant » : cette situation permet à une chaîne polypeptidique de se replier selon une très importante diversité de formes spatiales.

Toutefois, ce repliement s'effectue dans la majorité des cas selon quelques motifs stéréotypés, qui sont donc retrouvés dans de très nombreuses protéines différentes. Ces motifs ne concernent qu'une partie de la protéine : ce sont les structures secondaires.

Il s'agit essentiellement de deux structures différentes, l'hélice α et le feuillet β, qui partagent comme point commun le fait que ce sont des liaisons faibles établies entre les atomes des liaisons peptidiques qui stabilisent ces structures. Les radicaux des AA impliqués n'interviennent ainsi que peu dans l'établissement des structures secondaires.

Remarque : toutefois, on peut trouver des séquences d'AA « classiques » (= des **motifs protéiques**) conduisant en général à la formation soit d'une hélice α soit d'un feuillet β.

1. L'hélice α

La chaîne d'AA s'enroule en hélice (enroulement droit). Chaque AA est relié à deux autres AA (3 positions en avant et 3 positions en arrière) par des liaisons hydrogènes entre les groupements C=O et N-H des liaisons peptidiques.

L'hélice fait 0,5 nm de diamètre, avec 3,6 résidus par tour (soit 0,54 nm).

Ex : de nombreux domaines transmembranaires sont formés par une succession d'une vingtaine d'acides aminés hydrophobes, organisés en hélice α .

2. Le feuillet β

La chaîne d'AA forme une structure ondulée, un pli faisant 0,7 nm.

Plusieurs chaînes côte à côte conduisent à une structure en feuillet (entre deux chaînes, par exemple : coude formé de 4 AA, avec souvent des proline). Les chaînes côte à côte peuvent être disposés de manière parallèle (dans le même sens) ou antiparallèle (sens opposés). Des liaisons hydrogènes s'établissent entre les squelettes des chaînes adjacentes.

3. Bilan

La structure secondaire est le premier niveau de repliement d'une protéine. Il repose sur l'établissement de liaisons hydrogènes au niveau des liaisons peptidiques.

La formation d'hélice ou de feuillet rejette les chaînes latérales vers l'extérieur : ces chaînes peuvent alors interagir avec leur environnement.

Il est à noter qu'il peut exister d'autres structures secondaires. Il peut s'agir de structures non répétitives (bien que stables !), il peut aussi s'agir d'autres structures répétitives : c'est le cas par exemple de **l'hélice du collagène** (protéine des MEC animales – motif répété Gly-X-Pro ou Gly-X-hydroxyPro).

C. Structure tertiaire et radicaux des acides aminés

La structure tertiaire est la forme (= conformation) dans l'espace d'une unique chaîne polypeptidique.

La **myoglobine**, par exemple, est une protéine de 153 AA chez l'Homme, qui présente une structure globulaire avec le regroupement de 8 hélices alpha encadrant un hème.

Hème : groupement prosthétique formé d'un noyau tétrapyrrolique fermé avec un atome de Fer ferreux Fe^{2+} . Il est présent dans une poche hydrophobe, sans liaison covalente avec l'apoprotéine.

La myoglobine comporte donc une partie protéique (l'apoprotéine) et une partie non protéique (le groupement prosthétique).

C'est la **nature chimique des chaînes latérales** des AA qui permet aux AA d'établir des liaisons entre eux et/ou avec le milieu environnant. Ce sont ces liaisons qui permettent d'obtenir une structure 3D stable.

- Liaisons **covalentes** : ponts disulfures entre 2 cystéines.
- Liaisons **faibles** : liaisons hydrogènes, électrostatiques, de Van Der Waals, interactions hydrophobes.
- Interaction avec **l'environnement** : les AA hydrophiles ont tendance à se placer en périphérie de la protéine (en contact avec le solvant : établissement de liaisons H avec l'eau), alors que les AA hydrophobes se regroupent au cœur de la protéine (par interactions hydrophobes + VdW) (sauf cas des protéines transmembranaires). Les AA hydrophobes peuvent aussi être « tournés vers l'extérieur », dans le cas de protéines membranaires (*cf. profil d'hydropathie !*).

D. L'existence de protéines à structure quaternaire

Certaines protéines sont oligomériques : elles possèdent plusieurs chaînes d'AA, nommées sous-unités. La structure formée par l'ensemble de ces sous-unités est d'ordre supérieur à la structure tertiaire : il s'agit de la structure quaternaire.

On peut prendre comme exemple l'hémoglobine (Hb). Cette protéine est présente en abondance dans le cytosol des hématies (globules rouges), où elle joue un rôle de transporteur de dioxygène : elle peut en effet fixer le O_2 de manière réversible.

L'hémoglobine est formée de quatre sous-unités. Chaque sous-unité est un polypeptide de la famille des globines : il s'agit de protéines globulaires possédant une partie non protéique (= groupement prosthétique ; ici un hème ferreux).

Une hémoglobine peut ainsi fixer 4 molécules de O_2 au niveau des ions Fe^{2+} de chaque globine :



IV. Une forte relation structure-fonction

A. Différentes propriétés structurales

1. Des interactions protéine – ligand pour des rôles fonctionnels

a. Une diversité d'interactions protéine - ligand

Les propriétés de certaines protéines sont dues à leur capacité à se lier à une autre molécule, nommée **ligand**. C'est le cas en particulier :

- Des **enzymes** : le ligand est le substrat (ou les substrats), dont la transformation en produit(s) est accélérée par l'enzyme.
- De protéines **transmembranaires** permettant le passage de substances de part et d'autre de la membrane : le ligand est la substance traversant la membrane grâce à la protéine. Par exemple, le glucose est le ligand des transporteurs GluT.
- De **transporteurs** : le ligand est la substance transportée par la protéine. Par exemple, le dioxygène O₂ est un ligand de l'hémoglobine.

Selon les cas, le ligand peut donc être modifié (substrat d'une enzyme) ou pas (autres cas) suite à sa fixation avec la protéine.

De même, une protéine peut fixer un seul ligand, ou plusieurs ligands en même temps :

- Fixation de plusieurs substrats au niveau du site actif d'une enzyme (lorsque la réaction catalysée met en jeu plusieurs réactifs) ;
- Fixation d'un ligand « principal » sur un site, et d'un ligand à rôle régulateur sur un autre site.

b. Le site de fixation du ligand

L'interaction entre la protéine et le ligand est réalisée par la mise en place de liaisons faibles entre le ligand et les chaînes latérales de la protéine. Ce sont donc les propriétés chimiques des chaînes latérales qui expliquent la possibilité de la fixation du ligand.

c. Spécificité et affinité de l'interaction

La fixation du ligand est **spécifique** : Les liaisons faibles réalisées ne peuvent s'établir qu'entre la protéine et un ligand bien précis, présentant des propriétés chimiques compatibles avec la mise en place des liaisons faibles nécessaires.

Cette spécificité peut être plus ou moins importante :

- L'hexokinase est une enzyme permettant la phosphorylation de tous les hexoses : l'ensemble des hexoses sont donc des ligands de cette enzyme ;
- La glucokinase (présente dans le foie) ne permet la phosphorylation que du glucose : elle a donc une spécificité plus restreinte, et seul le glucose est un ligand de cette enzyme.

De même, lorsque l'on compare deux ligands d'une même protéine, ou un ligand capable de se lier à deux protéines différentes, l'interaction protéine – ligand peut être mise en place avec plus ou moins de facilité : Ceci correspond à l'**affinité** de la protéine pour son ligand.

L'affinité de la protéine pour un ligand peut être évaluée à partir de la courbe présentant l'activité de la protéine (réaction catalysée, flux transmembranaire, etc.) en fonction de la concentration du ligand :

- L'activité protéique plafonne à une valeur maximale, traduisant la saturation des sites de fixations ;
- La concentration en ligand correspond à la moitié de cette activité maximale correspond au **K_M** (pour une enzyme ; constante de Michaelis) ou au **K_d** (pour une autre protéine ; constante de dissociation) : le K_M ou le K_d est inversement proportionnel à l'affinité :

$$K_M \text{ (ou } K_d) \equiv \frac{1}{\text{affinité}}$$

2. Des propriétés structurales

Les propriétés de certaines protéines sont structurales : Elles sont la conséquence des propriétés mécaniques de la protéine.

Le collagène est une protéine structurale de la matrice extracellulaire animale, dans laquelle cette protéine joue un rôle de charpente. Ce rôle est rendu possible par une structure tridimensionnelle particulière, sous forme de triple hélice gauche.

En effet, le collagène présente une séquence primaire répétitive (motif Gly-X-(Hydroxy)Pro) qui conduit à un repliement spécifique en hélice gauche, permettant la formation d'une protéine linéaire (adapté à sa fonction dans la MEC).

B. Des structures dynamiques

1. Des structures basées sur des liaisons faibles

Les travaux d'Anfinsen, réalisés dans les années 1960, sur la **ribonucléase** pancréatique (enzyme hydrolysant l'ARN). Il s'agit d'une protéine formée d'une unique chaîne de 124 AA. Quatre paires d'AA cystéine sont reliées de manière covalente : on parle de **pont disulfure**.

On traite la ribonucléase par du **β -mercaptoéthanol (qui rompt les ponts disulfure)** et de l'**urée (qui déstabilise les liaisons faibles)**. On observe alors une perte d'activité de l'enzyme, corrélativement à sa dénaturation (perte de structure 3D). Après élimination par dialyse des agents dénaturant, on retrouve la fonction et la structure de la ribonucléase.

Cette expérience illustre le **rapport structure – fonction** au sein des protéines, et montre que les **liaisons établies entre les AA de la protéine** (forte –ponts disulfures- et faibles) **sont indispensables à cette structure**. Ces liaisons –et en particulier les liaisons faibles– permettent de stabiliser une structure protéique qui serait, sinon, très variable du simple fait de l'agitation moléculaire.

La conformation 3D d'une protéine est donc basée en grande partie sur des liaisons faibles, qui sont par définition **labiles** (c'est-à-dire faciles à rompre et à établir) : une protéine est ainsi une structure déformable.

2. L'ajustement induit

Initialement (Fisher, 1894), il a été proposé que les enzymes reconnaissent leur substrat grâce à une stricte complémentarité de forme en le substrat et le site actif de l'enzyme : c'est le modèle « **clé-serrure** ».

Cette idée générale reste valable, mais il est toutefois apparu, lorsque des études des structures tridimensionnelles ont été possibles, que la fixation du substrat pouvait induire un changement de la forme 3D de l'enzyme : c'est le modèle de l'**ajustement induit**, dont l'exemple classique est l'hexokinase, dont le site actif final est formé par un repliement de la protéine, induit par la fixation du substrat (le glucose).

Il semblerait maintenant que l'ajustement induit soit la règle, les cas extrêmes du type de l'hexokinase restant toutefois minoritaires : même dans une situation « clé-serrure », la fixation du substrat induit une légère modification de la conformation de l'enzyme.

3. Allostérie et coopérativité

Dans le cas de protéines à structure quaternaire, la fixation d'un ligand (molécule pouvant se fixer de manière spécifique au niveau d'un site de fixation) sur une sous-unité de la protéine peut entraîner un changement de conformation de cette sous-unité, qui induit un changement de conformation de l'ensemble de la protéine : on parle de coopérativité.

Ce changement de conformation augmente l'affinité de la protéine pour son ligand : Il s'agit donc aussi d'allostérie, la transition allostérique permettant le passage d'une forme T (tendue, *tense*, à faible affinité) à une forme R (relâchée, *relax*, à forte affinité).

On parle ainsi :

- De **coopérativité** : le changement de conformation d'une sous-unité induit le changement de conformation des autres sous-unités de la protéine.
- **D'allostérie** : le changement de conformation s'accompagne d'un changement d'affinité de la protéine. Cette allostérie se traduit par une cinétique sigmoïde.

4. Ajustement induit et allostérie : l'exemple de l'hémoglobine

Au niveau d'une sous-unité (globine) de l'hémoglobine, la **fixation de O₂ entraîne une modification de la conformation de cette sous-unité** : l'établissement de la sixième coordinance conduit à une contraction du rayon atomique du Fer, qui peut alors se placer dans le plan de l'hème. Ce mouvement du Fer entraîne avec lui l'histidine proximale (F8) à laquelle il est relié par une liaison de coordinance. C'est toute la structure 3D de la sous-unité qui est légèrement modifiée par ce mouvement de l'hélice α F.

Le changement de conformation de la globine induit la **rupture de liaisons** ioniques entre les sous-unités de l'hémoglobine, ce qui se traduit par un changement de conformation des autres globines de la protéine, qui présentent alors une **meilleure affinité** pour le dioxygène.

La désoxyhémoglobine correspond ainsi à la forme T, à faible affinité, de l'hémoglobine, alors que l'oxyhémoglobine correspond à la forme R, à forte affinité.

La transition allostérique se réalise au niveau de la gamme de concentrations en O₂ au niveau des tissus, ce qui permet un relargage du O₂ variable en fonction des besoins cellulaires.

C. Une possibilité de modifications post-traductionnelles

1. Glycosylations dans la voie réticulum-Golgi

Les protéines peuvent être glycosylées : un **oligosaccharide** (ensemble de quelques oses, souvent ramifié) peut être branché sur certains acides aminés (hydroxylés ou aminés) par une liaison osidique.

Cette glycosylation est réalisée tout d'abord dans la lumière du **réticulum endoplasmique**, puis modifiée lors du passage des protéines par l'**appareil de Golgi** : Elle ne concerne donc que les protéines dont la synthèse utilise cette voie, et les glycosylations ne sont pas cytosoliques (→ glycosylation des protéines extracellulaires, parties extracellulaires des protéines membranaires et protéines lysosomiales).

Les glycosylations peuvent avoir des rôles variés, dont par exemple :

- étiquetage lors de l'adressage (mannos-6P → lysosomes)
- identité immunologique cellulaire (ex : groupes ABO)
- hydratation de la face externe des membranes plasmiques
- etc.

2. La phosphorylation des protéines

Les acides aminés hydroxylés (sérine, thréonine et tyrosine) peuvent être phosphorylés. Cette phosphorylation de la protéine est réalisée par les protéines kinases, et la déphosphorylation est réalisée par les phosphatases.

L'état de phosphorylation éventuel d'une protéine influe sur l'activité de cette dernière : par exemples, certaines enzymes sont actives une fois phosphorylée et inactives sous leur forme déphosphorylée ; la situation est inverse pour d'autres enzymes.

V. Bilan sur les stratégies d'étude des protéines

A. L'étude de la structure des protéines

La structure des protéines peut être étudiée à partir de plusieurs méthodes complémentaires :

- Séquençage de la protéine, soit direct, soit en passant par le séquençage du gène ;
- Cristallographie aux rayons X, RMN ; cryomicroscopie électronique (CryoEM) ;
- Profils d'hydropathie etc.

B. La purification de protéines

Lorsque l'on connaît un ligand d'une protéine, il est possible d'utiliser cette connaissance pour purifier la protéine.

On peut en effet réaliser, par exemple, des microbilles recouvertes du ligand en question (ou d'une forme stable du ligand : dans le cas d'enzymes, on utilise un dérivé du substrat qui peut se lier à l'enzyme mais n'est pas catalysé, afin d'obtenir une association durable). Après incubation des microbilles avec la protéine, celle-ci est fixée sur les microbilles, d'où il est ensuite possible de la détacher en changeant les conditions du milieu.

Ce type de stratégie peut être mise en œuvre lors d'une chromatographie d'affinité (microbilles placées dans une colonne), ou en récupérant les microbilles par centrifugation.

De même, il est possible de purifier des protéines en réalisant une immunoprécipitation.

C. Des stratégies de génétique moléculaire pour étudier la fonction des protéines

Les techniques de génétique moléculaire permettent l'étude des protéines.

En particulier, la **transgénèse** permet l'introduction d'un gène dans le génome d'une cellule : ceci permet de « rajouter » un gène codant une protéine donnée à une cellule, afin d'étudier les effets éventuels de la présence de cette protéine, ou de la surexpression de cette protéine.

Dans le cas d'une transgénèse, il est en effet possible d'utiliser une construction génétique dans laquelle les régions promotrices utilisées sont différentes de celles présentes dans le génome :

- Promoteur « fort », constitutif : expression forte et dans toutes les cellules de l'organisme, afin d'étudier l'effet d'une surexpression de la protéine.
- Promoteur modifié : expression dans des tissus différents, afin d'étudier l'effet de l'ajout de la protéine dans les tissus en question.

La **mutagenèse** peut aussi être réalisée en utilisant une séquence d'ADN modifiée : on parle alors de mutagenèse. Il est ainsi possible de modifier un ou quelques acides aminés de la protéine, afin d'étudier l'effet de ces modifications sur l'action de la protéine. Cette démarche permet de poser des hypothèses sur le rôle de certains acides aminés de la protéine.

D. La localisation des protéines

La localisation d'une protéine est un élément important de son rôle cellulaire. On peut en particulier l'étudier par des techniques d'immunocytochimie, en utilisant des anticorps spécifiques de la protéine.

Il est aussi possible de réaliser une transgénèse conduisant à l'expression d'une protéine chimère, constituée de la protéine étudiée et de la protéine GFP : la protéine GFP (*green fluorescent protein*) émet une fluorescence verte qui permet localiser la protéine chimère.

Le gène de la protéine GFP peut aussi (comme d'autres gènes) être utilisé comme gène rapporteur, dans le cas d'une étude des cellules exprimant la protéine dans un organisme pluricellulaire : on réalise une transgénèse avec une construction formée du gène de la GFP associé aux séquences régulatrices de la protéine étudiée → la protéine GFP est ainsi exprimée selon le même profil d'expression que la protéine étudiée.