

## Introduction à la DSP - Les protéines et leur structure

### Vos objectifs:

A la fin de cette leçon, vous devrez être capable de décrire les protéines et leur structure.

Les protéines sont composées de chaînes d'acides aminés. Une protéine typique contient entre 200 et 300 acides aminés. Chaque acide aminé a une formule chimique et une structure spécifique.

Les cellules fabriquent des protéines distinctes en disposant les différents acides aminés en séquences uniques. Pour chaque type de protéine, le code génétique de la cellule dicte quels acides aminés sont ajoutés et dans quel ordre.

Ces acides aminés soigneusement disposés forment une structure en chaîne maintenue par des liaisons chimiques. Comme ces liaisons chimiques sont faibles, elles peuvent être facilement affectées par des facteurs tels que la température et le pH.

Les protéines ne peuvent remplir leur fonction que si elles conservent leur forme. Une modification de la forme d'une protéine l'empêche de fonctionner correctement. Si la protéine cible (produit) est exposée à une température ou un pH inappropriés, la chaîne protéique perd sa forme et sa capacité à fonctionner complètement. La perte de structure de la protéine est appelée dénaturation. Une protéine dénaturée ne peut pas retrouver sa forme initiale.

**La structure des protéines** est la disposition tridimensionnelle des atomes dans une molécule de chaîne d'acides aminés. Les protéines sont des polymères - plus précisément des **polypeptides** - formés de séquences d'acides aminés, les monomères du polymère. Un acide aminé monomère peut également être appelé un résidu, indiquant une unité répétitive d'un polymère. Les protéines sont formées par des acides aminés qui subissent des réactions de condensation, par lesquelles les acides aminés perdent une molécule d'eau par réaction afin de s'attacher les uns aux autres, ce qui constitue une liaison peptidique. Par convention, une chaîne de moins de 30 acides aminés est souvent identifiée comme un peptide plutôt que comme une protéine. Pour pouvoir remplir leur fonction biologique, les protéines se replient en une ou plusieurs conformations spatiales spécifiques, sous l'effet d'un certain nombre d'interactions non covalentes, telles que la liaison hydrogène, les **interactions** ioniques, les forces de Van der Waals et l'effet hydrophobe. Pour comprendre les fonctions des protéines au niveau moléculaire, il est nécessaire de déterminer leur structure tridimensionnelle. Le domaine scientifique de la biologie structurale utilise des techniques telles que la cristallographie aux rayons X, la spectroscopie RMN, la microscopie électronique cryogénique (cryo) et l'interférométrie à double polarisation pour déterminer la structure des protéines.

Les structures des protéines varient en taille, de quelques dizaines à plusieurs milliers d'acides aminés. Par leur taille physique, les protéines sont classées en nanoparticules de 1 à 100 nm.

De très grands agrégats peuvent être formés à partir de sous-unités de protéines. Par exemple, plusieurs milliers de molécules d'actine s'assemblent en un microfilament.

Une protéine subit généralement des modifications structurales réversibles lorsqu'elle remplit sa fonction biologique. Les structures alternatives d'une même protéine sont appelées isomères de conformation différents, ou simplement conformations, et les transitions entre elles sont appelées changements de conformation.

### **Structure primaire**

La structure primaire d'une protéine fait référence à la séquence des acides aminés dans la chaîne polypeptidique. La structure primaire est maintenue par des liaisons peptidiques formées au cours du processus de biosynthèse des protéines. Les deux extrémités de la chaîne polypeptidique sont appelées carboxyle terminus (C-terminus) et amino terminus (N-terminus), en fonction de la nature du groupe libre à chaque extrémité. Un comptage des résidus commence toujours par l'extrémité N-terminale (groupe NH<sub>2</sub>), qui est l'extrémité où le groupe amino n'est pas impliqué dans une liaison peptidique. La structure primaire d'une protéine est déterminée par le gène correspondant à cette protéine. Une séquence spécifique de nucléotides dans l'ADN est transcrite en ARNm, qui est lu par le ribosome dans un processus appelé traduction. La séquence des acides aminés de l'insuline a été découverte par Frederick Sanger, qui a établi que les protéines ont des séquences d'acides aminés déterminées. La séquence d'une protéine est unique à cette protéine et définit la structure et la fonction de la protéine. La séquence d'une protéine peut être déterminée par des méthodes telles que la dégradation d'Edman ou la spectrométrie de masse en tandem. Souvent, cependant, elle est lue directement à partir de la séquence du gène, en utilisant le code génétique. Il est strictement recommandé d'utiliser les mots "résidus d'acides aminés" lorsqu'on parle de protéines car, lorsqu'une liaison peptidique est formée, une molécule d'eau est perdue, et les protéines sont donc constituées de résidus d'acides aminés. Les modifications post-traductionnelles telles que les phosphorylations et les glycosylations sont généralement considérées comme faisant partie de la structure primaire et ne peuvent pas être lues à partir du gène. Par exemple, l'insuline est composée de 51 acides aminés répartis en deux chaînes. Une chaîne comporte 31 acides aminés et l'autre 20 acides aminés.

### **Structure secondaire**

La structure secondaire désigne des sous-structures locales très régulières sur la chaîne principale du polypeptide. Deux principaux types de structure secondaire, l'hélice  $\alpha$  et le brin  $\beta$  ou feuillet  $\beta$ , ont été suggérés en 1951 par Linus Pauling et al. Ces structures secondaires sont définies par des modèles de liaisons hydrogène entre les groupes peptidiques de la chaîne principale. Elles ont une géométrie régulière, étant contraintes à des valeurs spécifiques des angles dièdres  $\psi$  et  $\phi$  sur le tracé de Ramachandran. L'hélice  $\alpha$  et le feuillet  $\beta$  représentent tous deux un moyen de saturer tous les donneurs et accepteurs de liaisons hydrogène dans la chaîne peptidique. Certaines parties de la protéine sont ordonnées mais ne forment pas de structures régulières. Il ne faut pas les confondre avec la bobine aléatoire, une

chaîne polypeptidique dépliée dépourvue de toute structure tridimensionnelle fixe. Plusieurs structures secondaires séquentielles peuvent former une "unité supersecondaire".

### Structure tertiaire

La structure tertiaire fait référence à la structure tridimensionnelle des molécules de protéines monomères et multimères. Les hélices  $\alpha$  et les feuillets plissés  $\beta$  sont repliés en une structure globulaire compacte. Le repliement est entraîné par les interactions hydrophobes non spécifiques, l'enfouissement des résidus hydrophobes par l'eau, sauf que la structure n'est stable que lorsque les parties d'un domaine protéique se verrouillent en place à l'aide d'interactions tertiaires spécifiques, telles que les ponts salins, les liaisons hydrogène, et l'emballage serré des chaînes latérales et des liaisons disulfure. Les liaisons disulfure sont extrêmement rares dans les protéines cytosoliques, car le cytosol (fluide intracellulaire) est généralement un environnement réducteur.

### Structure quaternaire

La structure quaternaire est la structure tridimensionnelle constituée par l'agrégation de deux ou plusieurs chaînes polypeptidiques individuelles (sous-unités) qui fonctionnent comme une seule unité fonctionnelle (multimère). Le multimère résultant est stabilisé par les mêmes interactions non covalentes et les mêmes liaisons disulfure qu'une structure tertiaire. Il existe de nombreuses organisations possibles de structures quaternaires. Les complexes de deux polypeptides ou plus (c'est-à-dire plusieurs sous-unités) sont appelés multimères. Plus précisément, on parle de dimère lorsqu'il contient deux sous-unités, de trimère avec trois sous-unités, de tétramère pour quatre sous-unités et de pentamère avec cinq. Les sous-unités sont souvent reliées entre elles par des opérations de symétrie, comme un axe double dans un dimère. Les multimères constitués de sous-unités identiques sont désignés par le préfixe "homo", tandis que ceux constitués de sous-unités différentes portent le préfixe "hétéro", comme dans le cas d'un hétérotétramère : deux chaînes alpha et deux chaînes bêta d'hémoglobine.

