

Structure des protéines

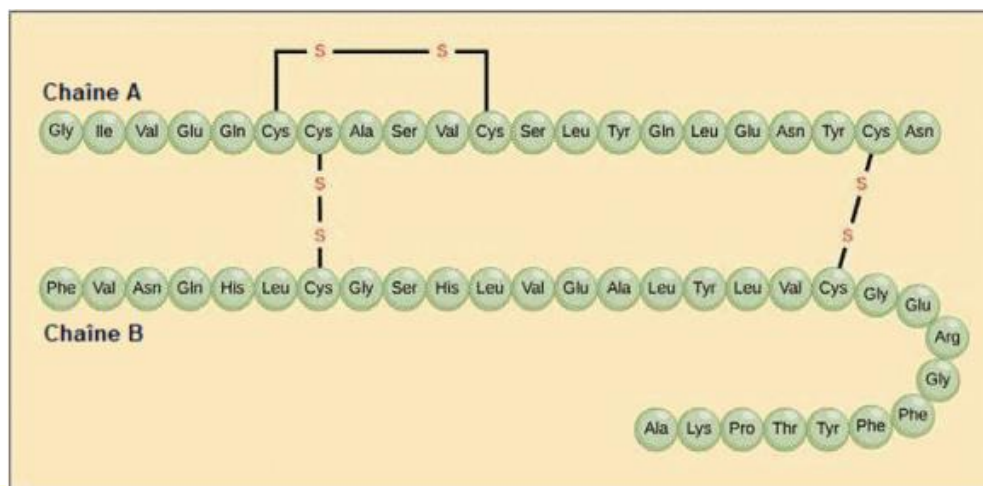
La forme d'une protéine est essentielle à sa fonction. Par exemple, une enzyme peut se lier à un substrat spécifique sur un site appelé site actif. Si ce site actif est altéré en raison de changements locaux ou de changements dans la structure globale des protéines, l'enzyme peut être incapable de se lier au substrat.

Pour comprendre comment la protéine prend sa forme ou sa conformation finale, nous devons comprendre les quatre niveaux de la structure protéique : primaire, secondaire, tertiaire et quaternaire.

Structure primaire.

La séquence unique d'acides aminés dans une chaîne polypeptidique correspond à sa *structure primaire*.

Par exemple, l'hormone pancréatique insuline a deux chaînes polypeptidiques, A et B, et elles sont liées entre elles par des liaisons disulfures (deux atomes de soufre liés par une liaison covalente simple). L'acide aminé N terminal de la chaîne A est la glycine, tandis que l'acide aminé C terminal est l'asparagine. Les séquences d'acides aminés dans les chaînes A et B sont uniques à l'insuline.

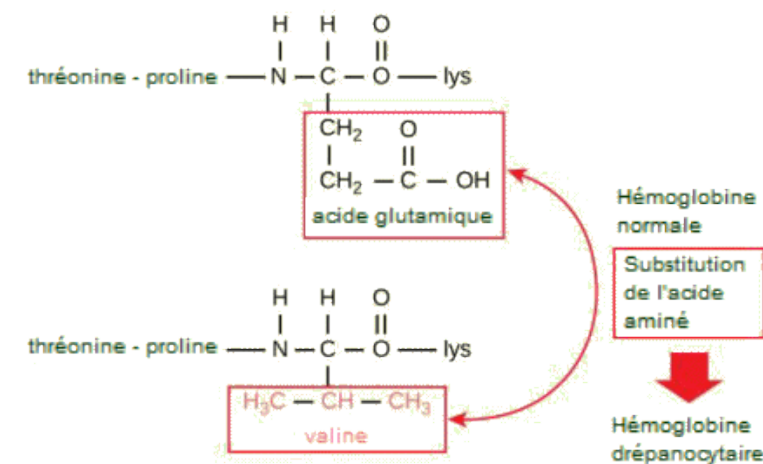


L'insuline sérique bovine est une hormone protéique constituée de deux chaînes peptidiques : A (21 acides aminés) et B (30 acides aminés). Dans chaque chaîne, la structure primaire est indiquée par des abréviations à trois lettres qui représentent les noms des acides aminés dans l'ordre où ils sont présents. L'acide aminé cystéine (cys) a un groupe sulfhydryle (SH) comme chaîne latérale. Deux groupes sulfhydryle peuvent

réagir en présence d'oxygène pour former une liaison disulfure (S-S). Deux liaisons disulfure relient les chaînes A et B ensemble, et une troisième aide la chaîne A à se replier dans la forme correcte. Notez que toutes les liaisons disulfure ont la même longueur, mais sont dessinées de tailles différentes pour plus de clarté.

La séquence unique pour chaque protéine est finalement déterminée par le gène codant pour la protéine. Un changement dans la séquence nucléotidique de la région codante du gène peut conduire à l'ajout d'un acide aminé différent à la chaîne polypeptidique en croissance, provoquant un changement dans la structure et la fonction des protéines.

Dans l'anémie falciforme ou drépanocytose, la chaîne β de l'hémoglobine (dont une petite partie est représentée sur la figure ci-dessous) a une seule substitution d'acides aminés, provoquant un changement dans la structure et la fonction des protéines. Plus précisément, l'acide aminé glutamique est substitué dans la chaîne β par la valine.

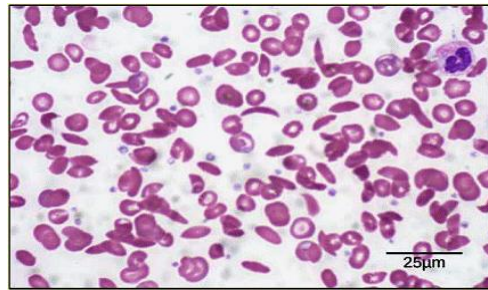


La chaîne bêta de l'hémoglobine a une longueur de 147 résidus, mais une seule substitution d'acide aminé conduit à l'anémie falciforme. Dans l'hémoglobine normale, l'acide aminé en position sept est le glutamate. Dans l'hémoglobine drépanocytaire, ce glutamate est remplacé par une valine.

Ce qui est le plus remarquable à considérer, c'est qu'une molécule d'hémoglobine est composée de deux chaînes alpha et de deux chaînes bêta qui se composent chacune d'environ 150 acides aminés. La molécule contient donc environ 600 acides aminés. La différence structurelle entre une molécule d'hémoglobine normale et une molécule dépanocytaire (= molécule induisant la drépanocytose, maladie qui diminue considérablement l'espérance de

vie) est un seul acide aminé des 600. Ce qui est encore plus remarquable, c'est que ces 600 acides aminés sont codés par trois nucléotides chacun, et la [mutation](#) est causée par un changement de base unique (mutation ponctuelle), 1 sur 1800 bases.

En raison de ce changement d'un acide aminé dans la chaîne, les molécules d'hémoglobine forment de longues fibres qui déforment les globules rouges biconcaves ou en forme de disque et leur font prendre une forme de croissant ou de faucille qui obstrue les vaisseaux sanguins. Cela peut entraîner une myriade de problèmes de santé graves tels que l'essoufflement, des étourdissements, des maux de tête et des douleurs abdominales pour les personnes touchées par cette maladie.



Dans ce **frottis sanguin**, visualisé au grossissement 535x en utilisant la microscopie à fond clair, l'anémie falciforme se manifeste par la présence de cellules (globules rouges) en forme de croissant ou de faucille, tandis que les cellules normales sont en forme de disque. (crédit: Ed Uthman).

Structure secondaire.

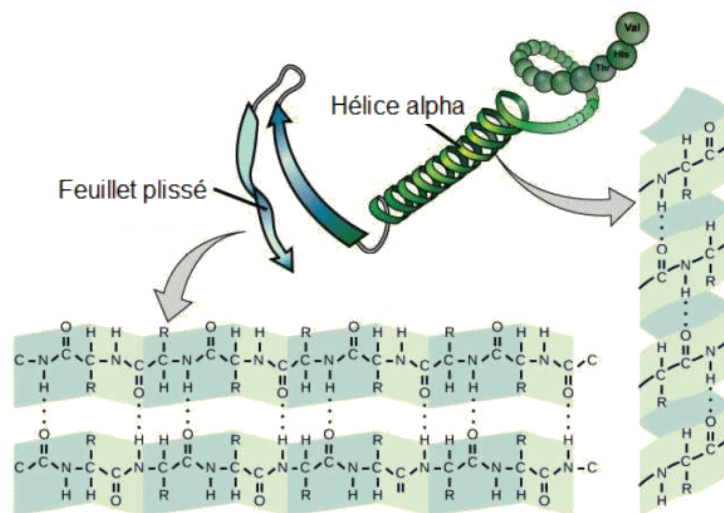
Le repliement local du polypeptide dans certaines régions donne naissance à la structure secondaire de la protéine. Les structures les plus courantes sont l'hélice et le feuillet plissé. Les deux structures sont maintenues ensemble par des liaisons hydrogène. Dans la structure en hélice, les liaisons hydrogène se forment entre l'atome d'oxygène du groupe carbonyle d'un acide aminé et un autre acide aminé situé quatre acides aminés plus loin le long de la chaîne.

Chaque tour hélicoïdal dans une hélice correspond à 3,6 résidus d'acides aminés. Les groupes R du polypeptide dépassent de la chaîne de l'hélice α .

Dans le feuillet plissé β , les "plis" sont formés par une liaison hydrogène entre les atomes sur le squelette de la chaîne polypeptidique. Les groupes R sont attachés aux carbones et

s'étendent au-dessus et en dessous des plis du pli. Les segments plissés s'alignent parallèlement ou antiparallèlement les uns aux autres, et des liaisons hydrogène se forment entre l'atome d'azote partiellement positif du groupe amino et l'atome d'oxygène partiellement négatif du groupe carbonyle du squelette peptidique.

Les structures en hélice α et en feuillet plissés β se retrouvent dans la plupart des protéines globulaires et fibreuses et jouent un rôle structurel important.



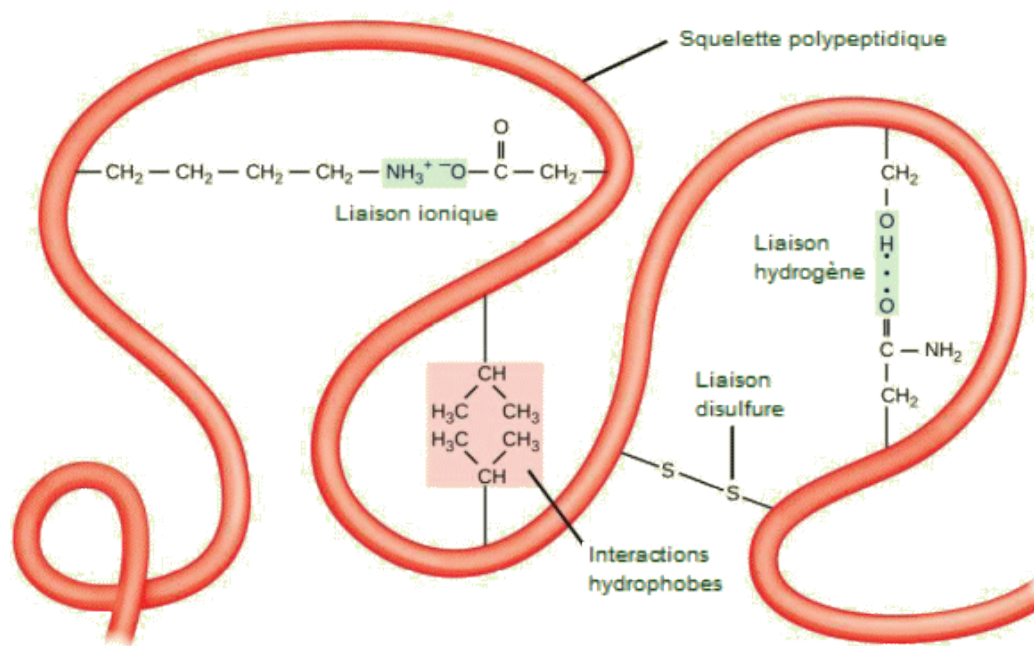
Structure secondaire d'une protéine. - L'hélice alpha et le feuillet plissé bêta sont des structures secondaires de protéines qui se forment en raison de la liaison hydrogène entre les groupes carbonyle et amino dans le squelette peptidique. Certains acides aminés ont une propension à former une hélice alpha, tandis que d'autres ont une propension à former un feuillet plissé bêta.

Structure tertiaire.

La seule structure tridimensionnelle d'un polypeptide est sa structure tertiaire. Cette structure est en partie due aux interactions chimiques à l'oeuvre sur la chaîne polypeptidique.

Ce sont principalement, les interactions entre les groupes R qui créent la structure tertiaire tridimensionnelle complexe d'une protéine. La nature des groupes R que l'on trouve dans les acides aminés impliqués peut contrecarrer la formation des liaisons hydrogène décrites pour les structures secondaires standard. Par exemple, les groupes R avec des charges similaires sont repoussés les uns par les autres et ceux avec des charges différentes sont attirés les uns vers les autres (liaisons ioniques).

Lors du repliement des protéines, les groupes R hydrophobes d'acides aminés non polaires se trouvent à l'intérieur de la protéine, tandis que les groupes R hydrophiles se trouvent à l'extérieur. Les premiers types d'interactions sont également connus sous le nom d'*interactions hydrophobes*. L'interaction entre les chaînes latérales de la cystéine forme des liaisons disulfure en présence d'oxygène : ce sont les seules liaisons covalentes se formant pendant le repliement des protéines.



La structure tertiaire des protéines est déterminée par une variété d'interactions chimiques. Il s'agit notamment des interactions hydrophobes, des liaisons ioniques, des liaisons hydrogène et des liaisons disulfure.

Toutes ces interactions, faibles et fortes, déterminent la forme tridimensionnelle finale de la protéine. Lorsqu'une protéine perd sa forme tridimensionnelle, elle peut ne plus être fonctionnelle.

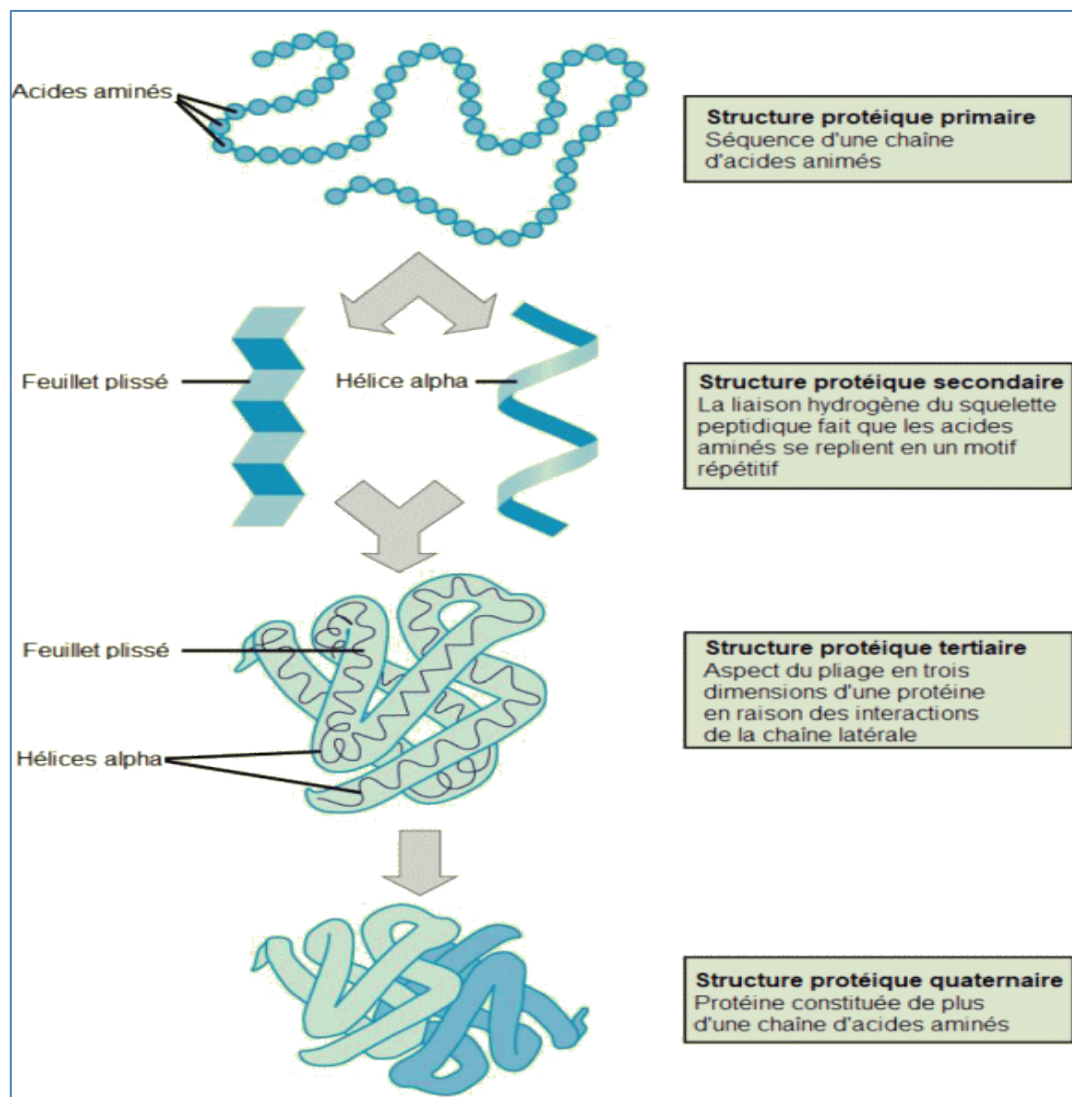
Structure quaternaire.

Dans la nature, certaines protéines sont formées de plusieurs polypeptides, également appelés sous-unités, et l'interaction de ces sous-unités forme la structure quaternaire. De faibles interactions entre les sous-unités contribuent à stabiliser la structure globale.

Par exemple, l'insuline (une protéine globulaire) possède une combinaison de liaisons hydrogène et de liaisons disulfure qui la rendent principalement agglomérée en forme de

boule. L'insuline commence comme un seul polypeptide et perd certaines séquences internes en présence d'une modification post-traductionnelle après la formation des liaisons disulfure qui maintiennent ensemble les chaînes restantes.

La soie (une protéine fibreuse), cependant, a une structure en feuillet plissé β c'est le résultat d'une liaison hydrogène entre différentes chaînes.



Résumé des quatre niveaux de structure protéique(crédit: National Human Genome Research Institute)

Dénaturation et repliement des protéines

Chaque protéine a sa propre séquence et forme unique qui sont maintenues ensemble par des interactions chimiques. Si la protéine est sujette à des changements de température, à des changements de pH ou à une exposition à des produits chimiques, la structure des protéines peut changer, perdant sa forme sans perdre sa séquence principale dans ce que l'on appelle la *dénaturation*.

La dénaturation est souvent réversible car la structure primaire du polypeptide est conservée dans le processus si l'agent dénaturant est éliminé, permettant à la protéine de reprendre sa fonction. Parfois, la dénaturation est irréversible, entraînant une perte de fonction.

On a un exemple de dénaturation irréversible des protéines lorsqu'un oeuf est frit : la protéine d'albumine dans le blanc d'oeuf liquide est dénaturée lorsqu'elle est placée dans une poêle chaude.

Toutes les protéines ne sont pas dénaturées à des températures élevées; par exemple, les bactéries qui survivent dans les sources hydrothermales ont des protéines qui fonctionnent encore à des températures proches de l'ébullition. L'estomac est également très acide (pH bas) et dénature les protéines dans le cadre du processus de digestion; cependant, les enzymes digestives de l'estomac conservent leur activité dans ces conditions.

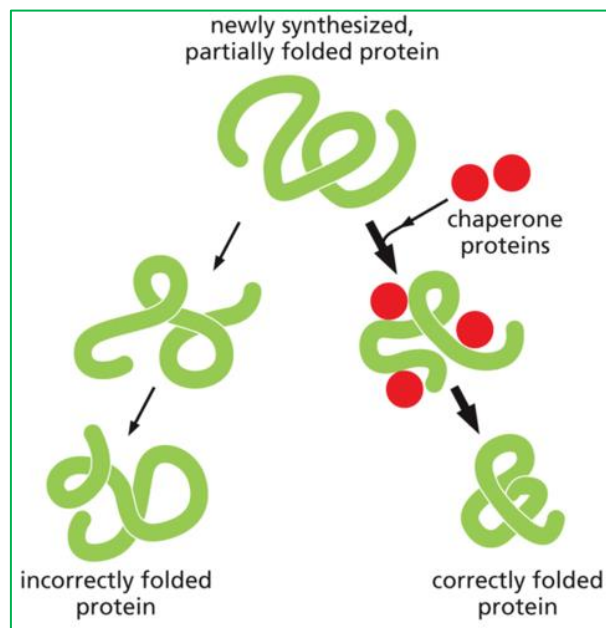
Rôle des protéines chaperons.

Le repliement des protéines est essentiel à leur fonction. On pensait à l'origine que les protéines elles-mêmes étaient responsables du processus de pliage. Ce n'est que récemment qu'il a été constaté que, souvent, elles reçoivent une assistance dans le processus de pliage de la part d'aides protéiques appelées *protéines chaperons* (certaines sont appelées *chaperonines*) qui s'associent à la protéine cible pendant le processus de pliage. Elles agissent en empêchant l'agrégation des polypeptides qui composent la structure protéique complète, et elles se dissocient de la protéine une fois que la protéine cible est repliée.

Mode d'action

Les protéines chaperon aident au repliement des protéines en modifiant leur conformation. Elles reconnaissent les protéines par leur surface hydrophobe (située en temps normal au

centre des protéines mais à l'extérieur dans ce cas), les masquent et, grâce à leur flexibilité, modifient leur structure en mettant les zones hydrophiles au contact de l'eau.



Une **protéine chaperon** est une protéine dont la fonction est d'**assister d'autres protéines** dans leur maturation en évitant la formation d'agrégats via les domaines hydrophobes présents sur leur surface lors de leur repliement tridimensionnel. Beaucoup de protéines chaperons sont des protéines de choc thermique (Heat shock proteins - Hsp), c'est-à-dire des protéines exprimées en réponse à des variations de température ou d'autres types de stress cellulaire, tel que le stress oxydant.

La dénaturation des protéines

La dénaturation des protéines est un processus au cours duquel l'application d'une contrainte externe fait perdre aux protéines toutes leurs structures présentes dans leur état natif. Cependant, les réactions de dénaturation ne sont pas assez fortes pour rompre les liaisons peptidiques qui ne perdent donc pas leur structure primaire (figure 1). Ces contraintes favorisent la rupture des interactions protéine-protéine et peuvent prendre la forme d'une contrainte chimique, d'une contrainte physique, ou les deux.

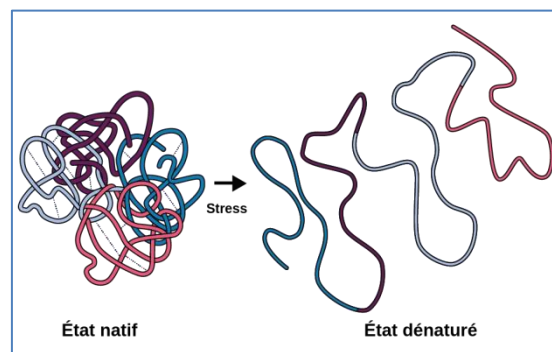


Figure 1 : Une représentation schématique du processus de dénaturation des protéines

Lorsque les protéines d'une cellule vivante sont dénaturées, l'activité cellulaire est perturbée, ce qui peut entraîner la mort cellulaire. Ce processus est très commun en science alimentaire où la dénaturation des protéines sert d'outil pour créer de nouveaux ingrédients et améliorer certains procédés de cuisine.

- **Facteurs de dénaturation :**

- *pH* : Des valeurs extrêmes peuvent déstabiliser la protéine.
- *Température* : Au-delà de 40°C, les protéines des organismes classiques se dénaturent.
- *Substances chimiques* : Détergents, urée et autres perturbent les liaisons.
- *Agents réducteurs* : Ex. mercaptoéthanol ou dithiothréitol affectent les ponts disulfure.

- **Conséquences :**

- La dénaturation entraîne une perte de fonction.
- Certaines protéines peuvent être *renaturées* si les conditions sont inversées (exemple : ribonucléase A).
- Certaines dénaturations sont **irréversibles**, comme la cuisson des œufs.

Voici les principaux **mécanismes chimiques** impliqués :

- **Modification des liaisons hydrogène** : Les variations de température ou de pH peuvent perturber les liaisons hydrogène qui stabilisent la structure secondaire et tertiaire.
- **Altération des interactions hydrophobes** : Certains agents chimiques, comme les détergents, peuvent perturber les interactions hydrophobes qui maintiennent la conformation des protéines.
- **Rupture des ponts disulfure** : Des agents réducteurs comme le **dithiothréitol (DTT)** ou le **β -mercaptoéthanol** brisent les liaisons disulfure entre les résidus cystéine, modifiant la structure de la protéine.
- **Effet des agents chaotropiques** : L'urée et le chlorure de guanidinium perturbent les interactions non covalentes, entraînant un dépliement de la protéine.
- **Modification du pH** : Un environnement acide ou basique peut provoquer la protonation ou la déprotonation des résidus d'acides aminés, modifiant les interactions électrostatiques et entraînant la dénaturation.

Ces mécanismes sont souvent exploités en laboratoire pour étudier la structure des protéines ou dans l'industrie agroalimentaire pour modifier leurs propriétés fonctionnelles.

Mécanismes de Repliement des Protéines et Maladies Associées

Introduction

Le repliement des protéines est un processus crucial dans la biologie cellulaire. Il permet à une chaîne d'acides aminés, formée lors de la traduction de l'ADN, de prendre une structure tridimensionnelle fonctionnelle. Une fois correctement repliée, une protéine est capable d'accomplir ses fonctions biologiques spécifiques. Cependant, un mauvais repliement ou un défaut dans ce processus peut entraîner diverses maladies, notamment des maladies neurodégénératives et des cancers. Dans cet article, nous explorerons les mécanismes de repliement des protéines et les pathologies associées.

1. Le Repliement des Protéines : Processus Fondamental

Le repliement des protéines commence dès qu'une chaîne polypeptidique est synthétisée par le ribosome dans le cytoplasme. Le polypeptide naissant interagit avec des chaperonnes moléculaires qui l'aident à adopter une conformation fonctionnelle. Ce processus implique plusieurs étapes :

- **Synthèse du Polypeptide** : Le ribosome synthétise la chaîne d'acides aminés, qui se déploie alors dans l'espace intra-cellulaire.
- **Chaperonnes Moléculaires** : Les protéines chaperonnes, telles que la **hsp70** et la **hsp60** (heat shock proteins), facilitent le repliement en guidant les chaînes polypeptidiques vers leur forme correcte, en évitant les agrégats non fonctionnels.
- **Repliement Correct** : Après l'intervention des chaperonnes, la protéine acquiert sa structure tertiaire ou quaternaire qui est nécessaire à son activité biologique.
- **Contrôle de Qualité** : Les cellules possèdent des mécanismes de surveillance pour vérifier que les protéines sont correctement repliées avant de les envoyer dans leur compartiment fonctionnel.

2. Mécanismes de Défaillance du Repliement et Agrégation Protéique

Un repliement incorrect peut survenir en raison de mutations génétiques, de stress cellulaire, de vieillissement, ou d'autres facteurs environnementaux. Lorsque cela se produit, la protéine ne fonctionne pas correctement, voire peut devenir toxique pour la cellule.

- **Mutations Génétiques** : Certaines mutations peuvent modifier la séquence d'acides aminés et perturber le repliement de la protéine. Un exemple typique est la **maladie de Huntington**, où une expansion d'une répétition CAG dans le gène codant pour la huntingtine conduit à une agrégation protéique toxique.
- **Stress Cellulaire** : Des facteurs tels que le stress oxydatif ou la privation de nutriments peuvent altérer la fonction des chaperonnes et des autres protéines impliquées dans le repliement, entraînant des erreurs.
- **Vieillesse** : Avec l'âge, la capacité des cellules à maintenir un repliement protéique adéquat diminue, augmentant ainsi le risque de maladies liées à l'agrégation des protéines.

3. Maladies Associées à un Mauvais Repliement des Protéines

Les erreurs de repliement des protéines sont impliquées dans un certain nombre de maladies graves. Voici quelques exemples :

a. Maladies Neurodégénératives

Les maladies neurodégénératives telles que **Alzheimer**, **Parkinson** et la **Sclérose Latérale Amyotrophique (SLA)** sont souvent causées par l'accumulation de protéines mal repliées dans le cerveau.

- **Alzheimer** : L'agrégation de la protéine **tau** et des **plaques amyloïdes** (dérivées de l'amyloïde- β) dans le cerveau est l'un des principaux marqueurs de cette maladie. Ces agrégats perturbent les cellules nerveuses et entraînent leur déclin progressif.
- **Parkinson** : L'agrégation de la protéine **α -synucléine** forme des corps de Lewy, qui perturbent le fonctionnement neuronal et conduisent à la dégénérescence des neurones dopaminergiques.
- **SLA** : Des protéines mal repliées telles que la **superoxyde dismutase 1 (SOD1)** sont impliquées dans la dégénérescence des motoneurones.

• b. Amyloïdoses

- Les **amyloïdoses** sont des maladies où des protéines mal repliées s'agrègent et forment des dépôts insolubles appelés **amyloïdes**. Ces dépôts peuvent se former dans différents organes, y compris les reins, le cœur et le foie, entraînant des dommages tissulaires. Un exemple bien connu est l'**amyloïdose AL**, associée à des dépôts de chaînes légères immunoglobulines.

• c. Cystic Fibrosis (Mucoviscidose)

- La **mucoviscidose** est une maladie génétique causée par des mutations du gène **CFTR**, qui code pour une protéine impliquée dans le transport des ions. La mutation la plus courante, **F508del**, perturbe le repliement de la protéine CFTR, la rendant incapable de se transporter correctement à la surface cellulaire. Cela conduit à des problèmes respiratoires graves et à d'autres complications.
- **d. Cancer et Repliement des Protéines**

Le mauvais repliement des protéines est également associé au cancer. Certaines mutations dans des protéines clés, telles que celles impliquées dans la **réparation de l'ADN** ou dans le contrôle du cycle cellulaire, peuvent être liées à un mauvais repliement, ce qui favorise la transformation maligne des cellules. Par exemple, des mutations dans la **p53**, une protéine qui joue un rôle crucial dans la régulation du cycle cellulaire et la réparation de l'ADN, peuvent altérer son repliement et son fonctionnement, contribuant ainsi à la tumorigénèse.

4. Conclusion

Le repliement des protéines est essentiel pour la fonction cellulaire et la santé globale de l'organisme. Lorsque ce processus est perturbé, il peut entraîner une variété de maladies graves, allant des maladies neurodégénératives aux cancers. La compréhension des mécanismes de repliement et de leurs défaillances offre des perspectives intéressantes pour le développement de nouvelles thérapies, comme l'utilisation de molécules chaperonnes pour améliorer le repliement des protéines défectueuses ou l'inhibition des agrégats protéiques. Les recherches sur le repliement des protéines continuent d'évoluer, et des progrès dans ce domaine pourraient bien ouvrir la voie à de nouvelles stratégies thérapeutiques pour traiter ces maladies dévastatrices.

Méthodes classiques pour l'étude de la structure tridimensionnelle des protéines

L'étude de la structure des protéines repose sur plusieurs méthodes complémentaires, chacune offrant des informations spécifiques :

- **Cristallographie aux rayons X** : Technique permettant de déterminer la structure tridimensionnelle des protéines en analysant la diffraction des rayons X à travers un cristal de protéine.
- **Résonance magnétique nucléaire (RMN)** : Utilisée pour étudier la structure des protéines en solution, cette méthode fournit des informations sur la dynamique et les interactions moléculaires.
- **Cryo-microscopie électronique** : Technique avancée qui permet d'observer des protéines à haute résolution sans nécessiter leur cristallisation.
- **Spectroscopie infrarouge et UV-visible** : Utilisée pour analyser la conformation des protéines et leurs interactions avec d'autres molécules.
- **Électrophorèse et western blot** : Techniques permettant de séparer et identifier les protéines en fonction de leur taille et de leur charge.
- **Modélisation informatique** : Approches bioinformatiques qui prédisent la structure des protéines à partir de leur séquence d'acides aminés.

Ces méthodes sont essentielles pour comprendre la fonction des protéines et leur rôle dans les processus biologiques.

Détermination de la structure tridimensionnelle :

Pour être actives, elles adoptent un repliement au sein duquel les atomes s'organisent de façon spécifique dans l'espace. On parle de *structure tridimensionnelle* des protéines. Pour déterminer ces structures, il existe plusieurs méthodes qui diffèrent de par leur **résolution**. Deux techniques permettent d'atteindre une résolution atomique : la spectroscopie par résonance magnétique nucléaire **RMN** et la radiocristallographie.

1. Spectroscopie par Résonance Magnétique Nucléaire (RMN) :

La spectroscopie par résonance magnétique nucléaire (RMN) repose sur l'analyse des noyaux atomiques de la molécule.

L'environnement des noyaux influencera la façon dont ils se comportent quand on les irradie à la longueur d'onde appropriée ; la RMN permet donc de déterminer la nature des liens entourant un atome.

On peut faire transiter un noyau de l'état d'énergie faible à l'état d'énergie élevé en l'irradiant avec une fréquence donnée. Le retour du noyau à son état d'énergie antérieur est enregistré par un détecteur. Ce phénomène de transition d'un niveau d'énergie à un autre est la résonance qui donne le R de RMN.

L'apport de la RMN est essentiel non seulement pour déterminer la structure tridimensionnelle des protéines avec une résolution atomique, mais aussi pour élucider les mécanismes qui régissent l'assemblage dynamique de ces molécules.

On utilise la RMN pour étudier la structure de petites protéines (jusqu'à 30 KDa environ) en solution sans avoir besoin de les cristalliser. La RMN utilise le spin des particules subatomique pour nous renseigner sur la structure des protéines.

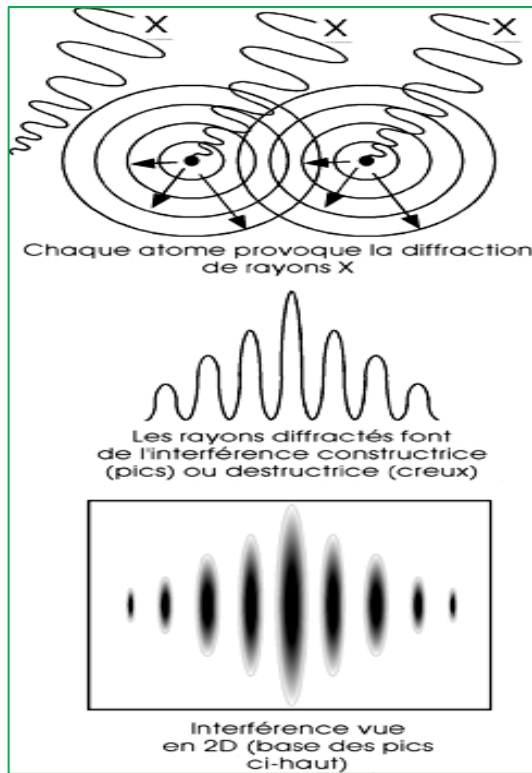
Le principe de la résonance magnétique nucléaire consiste à plonger une protéine en solution dans un **champ magnétique** intense (entre 10 et 20 **teslas**). Naturellement, seuls les atomes présentant un spin nucléaire non nul sont observables. Les protéines étant presque exclusivement composées d'atomes de carbone, d'oxygène, d'azote et d'**hydrogène**, les **protons** ont longtemps été les observables privilégiés de la RMN des protéines.

L'environnement électronique des protons module leurs fréquences de résonance qui varient de quelques parties par million (ppm) par rapport à une fréquence de référence. La valeur de la fréquence de résonance rapportée à une fréquence de référence et exprimée en ppm est appelée *déplacement chimique*. Celui-ci est fortement sensible au type d'atome lié au proton (carbone, oxygène, azote) (1), à son implication dans des liaisons non **covalentes** comme les liaisons hydrogène (2), à la proximité de noyaux aromatiques (3), ou encore d'un centre **paramagnétique**, **ion** métallique ou radical (4) contenant un **électron** non apparié. Ainsi, chaque proton résonne à un déplacement chimique qui le caractérise.

2. Diffraction aux rayons X (Cristallographie):

Le principe de la détermination d'une protéine par cristallographie au rayon X est de déterminer la position de ses atomes afin d'en déduire un modèle structural. La position précise de chaque atome d'une molécule peut être déterminée seulement si la molécule est cristallisée (l'organisation des atomes forme des cristaux). Lorsque les rayons X frappent une molécule cristallisée, les électrons entourant chaque atome courbent ou diffractent le faisceau rayon X; ce phénomène permet de déduire un modèle de diffraction par rayon X par analyse de la densité des électrons. Un ordinateur interprète alors mathématiquement ce modèle et reconstruit la position des atomes pour finalement donner un modèle quasi exact de la structure de la molécule.

Le nombre d'électrons dans les atomes de la protéine détermine le degré de diffraction des rayons, et c'est pourquoi les atomes lourds seront plus efficaces pour faire dévier les rayons X.



Chaque atome peut induire la diffraction de rayons X incidents. Les rayons déviés se répartissent dans l'espace. Si plus d'un atome est présent, les rayons déviés par l'un et l'autre se feront de l'interférence; cette interférence peut être constructive quand deux pics, ou encore deux creux, se superposent. Elle peut aussi être destructive quand ce sont un pic et un creux qui se superposent.

La succession d'interférence constructive et destructive donne le graphique du milieu; vue en 2D elle donne la figure du bas, ci-contre.

En combinant le patron de diffraction d'un cristal obtenu à partir de différents angles, on obtient assez d'information pour nous aider à déterminer l'agencement des atomes de la protéine dans l'espace.

La cristallisation demande une grande quantité de protéine pure.

Comparaison entre RMN et cristallographie :

La RMN appliquée aux macromolécules biologiques se distingue de la cristallographie par plusieurs points forts :

- les mesures de structures ne sont pas dommageables pour les échantillons protéiques et elles sont réalisées en solution, alors que la radiocristallographie des protéines nécessite au préalable de disposer d'un cristal de protéine de très bonne qualité.
- La RMN permet également de mettre en évidence les fluctuations conformationnelles des protéines, qui jouent souvent un rôle majeur dans la reconnaissance entre molécules ou dans le fonctionnement intime et dynamique des enzymes.

La banque de structures protéiques PDB

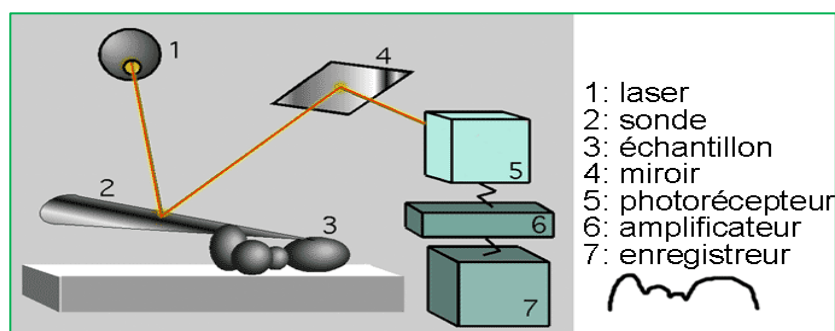
Protein Data Bank est la banque de référence des structures protéiques obtenues expérimentalement par cristallographie rayon X ou spectroscopie RMN. Jusqu'en mai 2003, 21000 structures protéiques ont été entrées dans la banque ; de 50 à 100 nouvelles structures sont déposées chaque semaine.

Les coordonnées des atomes formant la structure d'une protéine, le détail de la séquence, les conditions de cristallisation sont les principales informations disponibles pour chaque structure de la banque PDB. C'est à partir de cette banque que sont détectés les homologues structuraux. La majorité des séquences

ont une homologie structurale inférieure à 20% ; on évalue à environ 1000 le nombre total de structures protéiques originales qui suffirait à modéliser la quasi-totalité des protéines connues.

3. Microscopie de force atomique (AFM) :

Cette technique non-destructrice permet de relever des détails topographiques à très petite échelle. Elle avec un doigt. Dans ce cas-ci, le doigt est remplacé par un levier extrêmement ténu (la sonde). La position de cette sonde est suivie par la déviation du tracé d'un rayon laser qui rebondit en permanence sur sa surface réfléchissante et se rend jusqu'à un photodétecteur. Tout mouvement de la sonde quand elle grimpe par dessus un obstacle ou descend dans une cavité se traduit par une déviation du rayon, déviation qui est notée par le photodétecteur.



La microscopie de force atomique a permis de visualiser la structure de la chromatine, la position de polymérase sur l'ADN, et de nombreuses autres structures ne relevant pas toujours de la biologie.

Ce lien mène à de très belles images obtenues par AFM. On peut y suivre la condensation croissante d'un fragment de chromatine en fonction de la concentration en sel dans le milieu ambiant.

Les avantages de l'AFM sont nombreux. Elle ne requiert pas la destruction de l'échantillon, ne demande pas de coloration et donne une excellente idée de la forme de l'échantillon. Elle détecte les formes des molécules un peu comme vous le faites en tâtant une surface.

Prédiction sur les structures et fonction des protéines

Définition de la prédiction des structures de protéines

La prévision de la structure des protéines est l'inférence de la structure tridimensionnelle d'une protéine à partir de sa séquence d'acides aminés, c'est-à-dire la prédiction de son pliage et de sa structure secondaire et tertiaire de sa structure primaire. La prédiction de la structure est fondamentalement différente du problème inverse de la conception des protéines. La prédiction de la structure protéique est l'un des objectifs les plus importants poursuivis par la bioinformatique et la chimie théorique ; Elle est très importante en médecine (par exemple, dans la conception de médicaments) et en biotechnologie (par exemple, dans la conception de nouvelles enzymes). Tous les deux ans, la performance des méthodes actuelles est évaluée dans l'expérience CASP (Évaluation critique des techniques de prédiction des protéines). La structure d'une protéine donne beaucoup d'informations sur sa fonction et sa régulation. Mais ni la cristallographie aux rayons X ni la RMN n'est la panacée. En effet la première technique nécessite que la protéine étudiée cristallise, ce qui peut prendre plusieurs mois ou même ne jamais se réaliser. La seconde technique, la RMN, est limitée à des fragments d'une centaine d'acides aminés, et l'échantillon doit être soluble à forte concentration. Ainsi il est plus difficile, voire impossible pour l'instant, d'étudier des structures plus grosses, comportant des sous-unités ou flexibles que des plus petites, globulaires et rigides. De plus ces techniques demandent de longues et coûteuses manipulations, ce qui n'est pas conciliable avec le nombre de protéines dont la connaissance de structure présente un intérêt certain. C'est pourquoi, la détermination de structure de protéines par des outils bioinformatiques s'est rapidement révélé une nécessité, et des techniques comme la modélisation moléculaire par homologie ou la prédiction de structure tridimensionnelle de novo se développent.

Méthodes de prédiction de structures 3D

Aujourd'hui, les méthodes de prédiction de la structure tridimensionnelle des protéines peuvent être classées en trois catégories :

- **La modélisation par homologie (ou comparative)**
- **Les méthodes de reconnaissance de repliement (ou d'enfilage, threading en anglais).**
- **Les méthodes ab initio et de novo.**

Le choix de la méthode dépend des informations disponibles pour la réalisation du modèle. Le facteur le plus déterminant dépend de l'existence ou non dans la PDB d'une structure protéique similaire à celle de la protéine à modéliser et du taux d'identité de séquence entre ces protéines.

La panoplie des méthodes de prédiction de structure in-silico est quant à elle beaucoup plus large. Si plusieurs d'entre elles affichent des performances raisonnables, aucune n'est cependant parvenue à s'imposer. Toutefois, le concours CASP11 (Critical Assessment of Techniques for Protein Structure Prediction) permet de faire régulièrement le point sur l'efficacité des méthodes publiées. La fiabilité des prédictions est bien entendu fortement reliée au niveau de définition de la prédiction et du raffinement du modèle utilisé.

Modélisation moléculaire

La modélisation moléculaire est un ensemble de techniques pour modéliser ou simuler le comportement de molécules. Elle est utilisée pour reconstruire la structure tridimensionnelle de molécules, en particulier en biologie structurale, à partir de données expérimentales comme la cristallographie aux rayons X. Elle permet aussi de simuler le comportement dynamique des molécules et leurs mouvements internes. On l'utilise enfin pour concevoir de nouveaux médicaments. La modélisation moléculaire s'intéresse enfin au rendu visuel des molécules et des simulations en 3D, c'est le domaine du graphisme moléculaire.

La biologie structurale est la discipline qui a pour objet de reconstruire des modèles moléculaires, par l'analyse de données indirectes ou composites. L'objectif est d'obtenir une reconstruction tridimensionnelle présentant la meilleure adéquation avec les résultats expérimentaux. Ces données sont issues principalement d'analyses cristallographiques (étude des figures de diffraction des rayons X par un cristal), de résonance magnétique nucléaire, de cryomicroscopie électronique ou de techniques de diffusion aux petits angles (diffusion des rayons X ou diffusion des neutrons). Les données issues de ces expériences constituent des données (ou contraintes) expérimentales qui sont utilisées pour calculer un modèle de la structure 3D. Le modèle moléculaire obtenu peut être un ensemble de coordonnées cartésiennes des atomes composant la molécule, on parle alors de modèle atomique, ou une "enveloppe", c'est-à-dire une surface 3D décrivant la forme de la molécule, à plus basse résolution. L'informatique intervient dans toutes les étapes conduisant de l'expérimentation au modèle, puis dans l'analyse du modèle par la visualisation moléculaire.

Le docking moléculaire

Le docking moléculaire est une méthode informatique utilisée pour prédire la structure tridimensionnelle des complexes moléculaires, par la simulation de l'interaction entre une molécule cible, souvent une protéine, et un ligand, comme un médicament potentiel. En identifiant la position la plus favorable du ligand dans le site actif de la protéine cible, c'est-à-dire la plus stable, le docking moléculaire aide à prédire les interactions moléculaires et à améliorer leur efficacité. En somme, il révèle la disposition spatiale optimale des molécules lorsqu'elles se lient pour former un complexe, comme illustré Figure suivante.

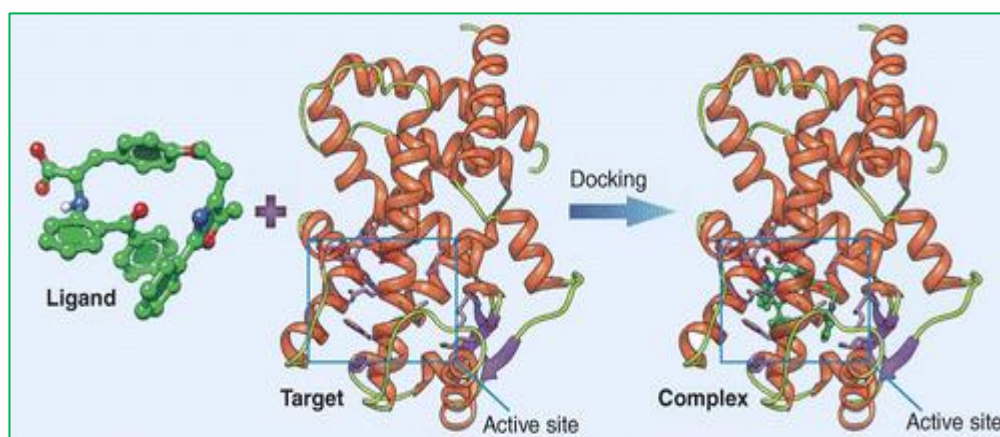


Figure 1 : principe du docking moléculaire, illustration de l'interaction entre une molécule (ligand) et une protéine (Target) au niveau de son site actif (Rectangle)

Depuis son développement dans les années 1990 pour la découverte de médicaments anti-VIH¹, le docking moléculaire est devenu une méthode standard et essentielle en recherche pharmaceutique² pour prédire les interactions protéine-ligand et aider la conception de médicaments. Ses applications s'étendent dans de nombreux autres domaines, quand l'optimisation de la disposition spatiale des macromolécules permet d'améliorer leur efficacité : en médecine personnalisée pour prédire la réponse des patients à certains médicaments en fonction de leur profil génétique et moléculaire, en industrie agroalimentaire pour optimiser des arômes et des additifs alimentaires, en agrochimie pour concevoir de nouveaux pesticides et herbicides plus spécifiques

Les avancées récentes en modélisation informatique des protéines sont impressionnantes, notamment grâce à l'intelligence artificielle. Voici quelques points clés :

- **AlphaFold3** : DeepMind a lancé une nouvelle version de son célèbre programme AlphaFold, qui améliore considérablement la précision de la prédiction des structures

protéiques. Ce modèle basé sur l'apprentissage automatique permet de mieux comprendre le repliement des protéines et leurs interactions moléculaires.

- **BETA et validation des prédictions** : Une nouvelle approche appelée BETA aide les chercheurs à éviter les erreurs liées aux données biaisées et améliore la fiabilité des prédictions.
- **Base de données massive** : DeepMind a également publié une base de données contenant plus de **214 millions** de modèles de protéines, facilitant la recherche en biologie et en médecine.
- **Applications médicales** : Ces avancées permettent d'accélérer la découverte de nouveaux médicaments et vaccins, notamment contre des maladies comme la malaria.

Ces progrès révolutionnent la biologie structurale et ouvrent la voie à des découvertes majeures.

Implication de l'intelligence artificielle dans les études de la structure des protéines

L'intelligence artificielle joue un rôle clé dans l'étude de la structure des protéines, notamment grâce à des avancées comme **AlphaFold**, un système développé par DeepMind qui prédit avec une grande précision la structure tridimensionnelle des protéines à partir de leur séquence d'acides aminés. Cette technologie révolutionne la biologie structurale en permettant aux chercheurs d'obtenir des modèles fiables sans recourir à des méthodes expérimentales coûteuses et longues.

Voici quelques implications majeures de l'IA dans ce domaine :

- **Prédiction rapide et précise** : L'IA permet de prédire la structure des protéines en quelques heures, contre plusieurs mois avec des techniques traditionnelles comme la cristallographie aux rayons X.
- **Accélération de la découverte de médicaments** : En comprenant mieux la structure des protéines, les chercheurs peuvent concevoir des médicaments ciblés plus efficacement.
- **Modélisation des interactions moléculaires** : L'IA aide à simuler comment les protéines interagissent entre elles et avec d'autres molécules, ce qui est essentiel pour la biologie intégrative.
- **Réduction des coûts de recherche** : Les prédictions basées sur l'IA permettent de limiter le recours à des expériences coûteuses et complexes.

alphaFold et structure protéique



AlphaFold est une avancée révolutionnaire dans la prédiction de la structure des protéines grâce à l'intelligence artificielle. Développé par **DeepMind**, ce modèle utilise l'apprentissage profond pour prédire la conformation tridimensionnelle des protéines à partir de leur séquence d'acides aminés.

Principales caractéristiques d'AlphaFold :

- **Précision remarquable** : AlphaFold atteint une précision comparable aux méthodes expérimentales comme la cristallographie aux rayons X.
- **Base de données massive** : Plus de **200 millions** de prédictions de structures protéiques sont accessibles librement.
- **Applications médicales et biotechnologiques** : Il accélère la découverte de médicaments et aide à comprendre les maladies liées aux protéines mal repliées.
- **Modélisation avancée** : AlphaFold utilise des algorithmes d'attention et des réseaux neuronaux entraînés sur des milliers de structures connues