

Figure. Carte d'interaction protéine-protéine impliquant 911 interactions caractérisées chez l'homme.

INTRODUCTION

Les protéines sont des macromolécules biologiques présentes dans les cellules et dont les activités peuvent être très variées. Elles sont découvertes en 1835 par Gerrit Mulder à Rotterdam en analysant le blanc d'œuf. Le terme de protéine est apparu en 1838, proposé par Jöns Jacob Berzelius dans une lettre à Gerrit Mulder. Il est issu du grec "protos" qui signifie "premier", associé à la découverte des protéines comme étant des éléments de base, essentiels au vivant. Une autre explication est apparue après la découverte de la plasticité des protéines, proposant que le terme "protéine" fasse référence au dieu grec Protée qui pouvait changer de forme, d'où l'adjectif "protéiforme".

Sous le nom générique de protides, on inclut les protéines, des substances de nature protéique telles que les enzymes, les acides aminés, et tous les composés donnant par hydrolyse un ou plusieurs acides aminés. Les peptides englobent les produits (naturels ou synthétiques) formés par l'union de plusieurs acides aminés et dans lesquels la liaison, dite peptidique ($-\text{CO}-\text{NH}-$), a lieu par perte d'une molécule d'eau entre un groupement aminé d'une molécule et un groupement carboxylique d'une autre

Structure

Une fois synthétisée, la chaîne polypeptidique se replie et prend une conformation spatiale particulière. Une protéine est alors caractérisée par sa structure, c'est à dire par la position relative des atomes qui la compose. **Il existe 4 niveaux de structure** (Fig) :

- 1. La structure primaire** correspond à la séquence des acides aminés (ou résidus) constituant la protéine, sans référence à une configuration spatiale. Ces résidus sont liés par des liaisons peptidiques, et forment une chaîne polypeptidique polarisée avec une extrémité N-terminale (amine libre) et une extrémité C-terminale (carboxylate libre).
- 2. La structure secondaire** décrit les repliements locaux de la chaîne de résidus. Ces structures locales correspondent à des repliements énergétiquement stables, dont le nombre est limité. De plus, certaines de ces conformations sont renforcées par des liaisons hydrogènes. Les trois principales catégories de conformations sont les hélices, les feuillets et les coudes. Ainsi, une protéine peut aussi être décrite par un enchaînement d'éléments de structures secondaires.

3. Les différentes structures secondaires sont agencées les unes par rapport aux autres pour former **la structure tertiaire**, souvent renforcée par des ponts disulfures. Les forces qui gouvernent ce repliement sont les forces physiques classiques. La structure tertiaire d'une protéine correspond au repliement de la chaîne polypeptidique dans l'espace. On parle plus couramment de structure tridimensionnelle. La structure tridimensionnelle d'une protéine est intimement liée à sa fonction : lorsque cette structure est détruite par l'emploi d'agent dénaturant par exemple, la protéine perd sa fonction.

4. Dans le cas des protéines formées par l'agencement de plusieurs chaînes, **la structure quaternaire** décrit la position relative des sous-unités les unes par rapport aux autres. La structure quaternaire des protéines regroupe l'association d'au moins deux chaînes polypeptidiques (identiques ou différentes) par des liaisons non-covalentes, dites faibles (liaison H, liaison ionique, interactions hydrophobes et force de Van der Waals), rarement par des ponts disulfures, qui ont pour rôle de créer les liaisons intra-chaînes. L'effet hydrophobe est un facteur prépondérant dans l'assemblage des éléments structuraux, y compris dans l'association des sous-unités. Chacune de ces chaînes est appelée "monomère" (ou sous-unité) et l'ensemble "oligomère" ou "protéine multimérique". Par exemple, la structure quaternaire de l'hémoglobine est constituée par 4 sous-unités : 2 sous-unités (de 141 acides aminés chacune) et 2 sous-unités (de 146 acides aminés chacune).

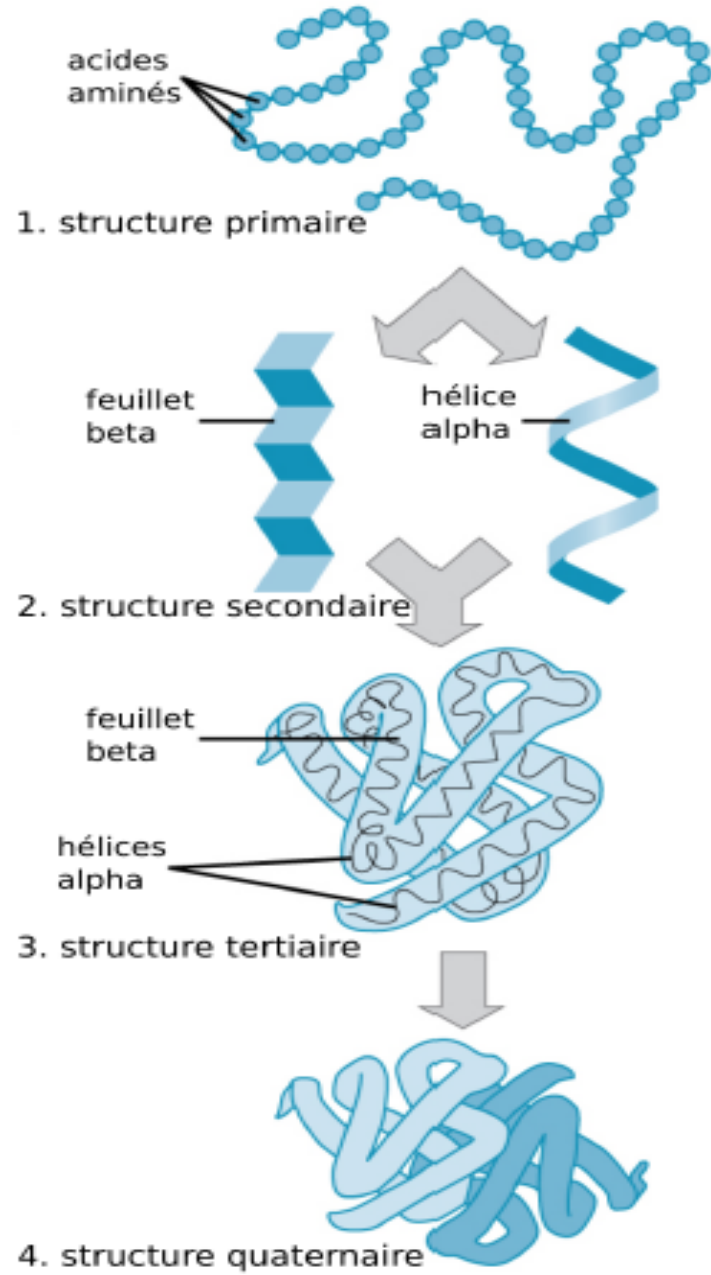


FIGURE – Les quatre niveaux de structure d’une protéine.

Synthèse des protéines

Les protéines sont des assemblages d'acides aminés, dont la nature est déterminée par l'information présente dans les gènes. La transformation de l'information génétique en protéine se fait en deux étapes :

- la transcription, lorsque l'information de l'ADN est transcrite en ARN messager,
- la traduction, lorsque l'information de l'ARN messager est traduite en protéine, en fonction du code génétique.

Chez les eucaryotes, la transcription a lieu dans le noyau, où peuvent aussi avoir lieu des modifications post-transcriptionnelles. Ces modifications peuvent être de différentes natures, telles que l'excision épissage, permettant de former différents ARN messagers à partir d'un même gène, et/ou l'ajout d'une coiffe et d'une queue poly(A), permettant de stabiliser la molécule d'ARN.

La traduction, quant à elle, se fait dans le cytoplasme ou le réticulum endoplasmique par un ribosome. **La protéine nouvellement produite** peut aussi subir des modifications post-traductionnelles, telles que des glycosylations dans l'appareil de Golgi. Il est important de noter que grâce aux étapes de modifications post-transcriptionnelles et post-traductionnelles, un gène peut coder pour différentes isoformes d'une même protéine, ce qui rend complexe le lien entre gènes et protéines, car ces isoformes peuvent avoir des fonctions très différentes

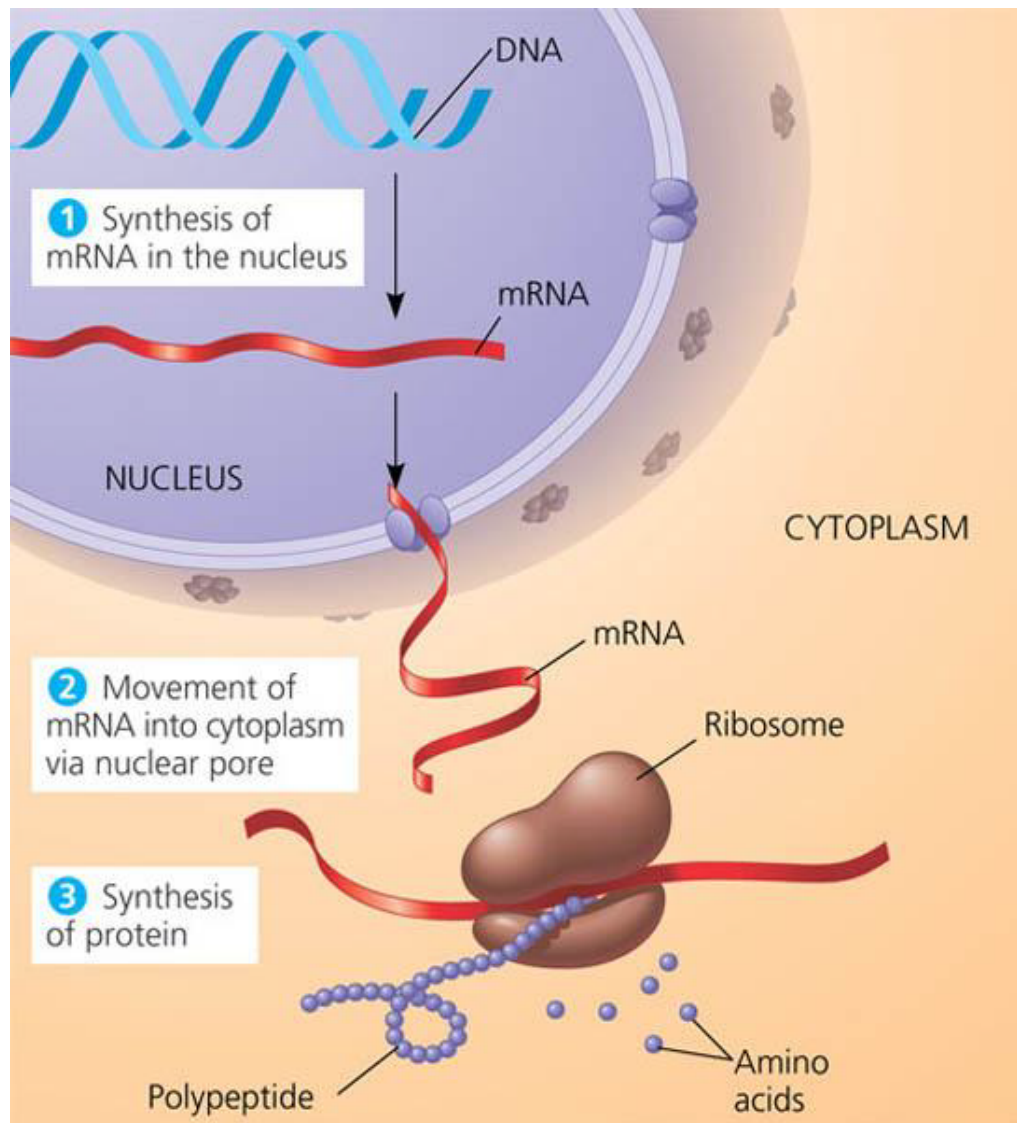


FIGURE – Les étapes de la synthèses des protéines.

La complexité de la fonction

Le terme fonction, bien que très largement employé dans la littérature scientifique, n'est pas toujours bien défini. La fonction biochimique est directement liée à la séquence d'ADN, qui code pour une séquence d'acides aminés, et qui se replie selon une conformation spécifique pour assurer une activité biochimique. Ces liens séquence-fonction et structure-fonction sont généralement portés par des éléments appelés domaines. De façon simple, un domaine protéique est une partie d'une protéine dont la structure et la fonction peuvent être stables indépendamment du reste de la chaîne protéique. La fonction d'un domaine est liée à sa séquence et/ou à sa structure. Si la séquence est modifiée, cela peut i) altérer la composition en acides aminés de l'interface active (le site actif pour les enzymes), et donc modifier directement la fonction moléculaire, et ii) modifier le repliement de la structure tridimensionnelle, pouvant avoir pour conséquence une modification de la fonction. C'est la raison pour laquelle les analyses de séquences d'ADN, d'acides aminés, et de structures des protéines sont très utilisées pour la compréhension de la fonction biochimique.

Les interactions protéine-protéine sont un aspect essentiel des processus biologiques. Elles sont fortement impliquées dans la formation de structures macromoléculaires, dans la signalisation, dans la régulation et dans les différentes voies métaboliques. Leur étude est donc cruciale pour la compréhension des réseaux d'interaction protéiniques, but majeur dans l'étude des systèmes biologiques. Les interactions protéine-protéine ont un rôle conséquent dans l'induction de beaucoup d'états pathologiques et dans les processus importants pour la pathogenèse des infections bactériennes et virales

Les interactions entre protéines

La plupart des protéines assurent leurs fonctions biologiques en interagissant avec une ou plusieurs autres protéines. Elles peuvent former de larges complexes protéiques tel que le protéasome, qui est un assemblage d'environ 50 sous-unités protéiques agissant ensemble pour dégrader d'autres protéines, et jouant un rôle primordial dans l'homéostasie. Ces interactions sont très diverses, selon leurs composition, leurs affinités ou leur nature permanente ou transitoire.

Homo-oligomère ou hétéro-oligomère.

Les interactions peuvent exister entre protéines identiques ou différentes (i.e. homo- ou hétérooligomères). Pour les homo-oligomères, l'interaction peut avoir lieu sur une même surface pour les deux monomères (isologue), ou sur deux surfaces différentes (hétérologue), auquel cas la formation d'agrégats est possible.

Obligatoire ou non-obligatoire.

Les complexes formés peuvent être obligatoires ou non. Une interaction est obligatoire si les monomères impliqués n'ont pas de structure stable in vivo en l'absence de cette interaction. Dans ce cas, il est fréquent que la fonction des protéines impliquées soit dépendante de cette interaction. La plupart des complexes hétéro-oligomériques impliquent des interactions non-obligatoires, ce qui signifie que les protéines sont stables en l'absence d'interactions.

Permanente ou transitoire.

On peut aussi distinguer les interactions selon leur dynamique. Les interactions permanentes sont très stables, et les protéines impliquées ne sont présentes que sous leur forme complexée. Les interactions transitoires sont beaucoup plus dynamiques, les partenaires s'associent et se dissocient rapidement *in vivo*. Les interactions transitoires peuvent être faibles, c'est à dire dépendantes de la concentration de chacun des partenaires dans le milieu. L'interaction est contrôlée par un équilibre oligomérique dynamique en solution, dans laquelle les interactions se font et se défont continuellement. Les interactions transitoires peuvent aussi être fortes lorsqu'elles sont contrôlées par un mécanisme moléculaire. L'équilibre oligomérique dépend alors d'une phosphorylation ou d'une déphosphorylation, ou de la présence d'un autre partenaire tel que la guanosine triphosphate. Bien souvent, les interactions obligatoires sont aussi permanentes.

Il est important de noter que la plupart des interactions ne tombent pas exactement dans chacune de ces catégories, mais qu'un continuum existe entre interactions obligatoires et non-obligatoires, et que la stabilité des complexes dépend beaucoup des conditions physiologiques et de l'environnement. Une interaction peut être principalement transitoire *in vivo*, mais devenir permanente sous certaines conditions cellulaires.

Les interactions sont régulées par différents mécanismes. On identifie 3 types de contrôles :

La localisation.

L'association de deux protéines dépend d'une rencontre des surfaces d'interaction, et requiert une co-localisation dans l'espace et le temps, c'est à dire une co-expression ou co-localisation dans un compartiment. Dans le cas où les protéines sont dans des compartiments différents, des transports dirigés entre les différentes localisations sont nécessaires.

La concentration locale.

Ce paramètre est contrôlé par divers mécanismes, tels que : l'expression des gènes, les niveaux de sécrétion, la dégradation des protéines, le stockage temporaire, l'environnement moléculaire, et la diffusion ou viscosité du milieu.

L'environnement physico-chimique local.

Les affinités mutuelles des composants d'un complexe peuvent être modifiées par la présence d'une molécule effectrice (e.g. ATP, Ca²⁺) ou par un changement des conditions physiologiques.

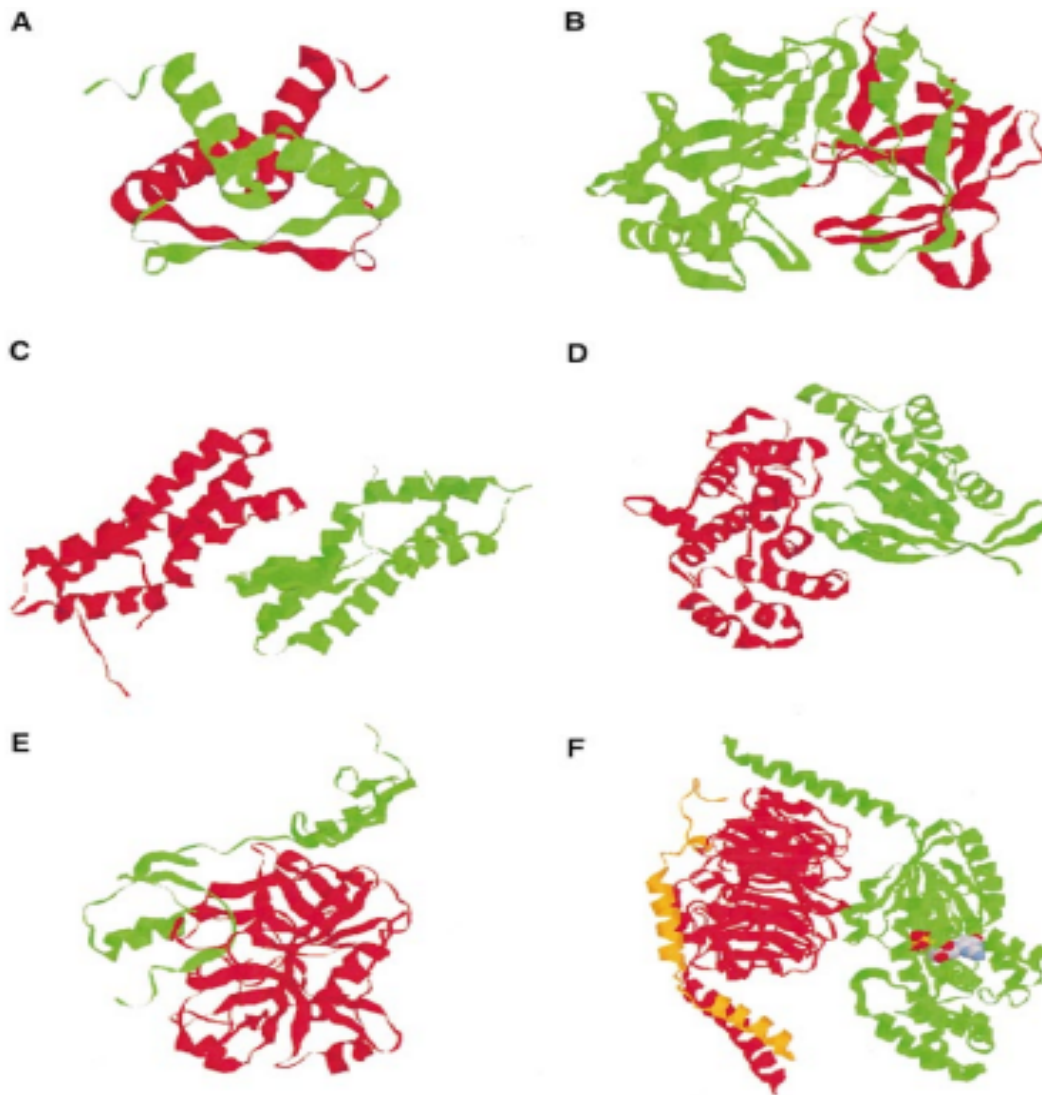


FIGURE – Exemples de différents types d'interactions décrits dans le texte : (A) homodimère obligatoire ; (B) hétérodimère obligatoire ; (C) homodimère non obligatoire ; (D) hétérodimère non obligatoire ; (E) hétérodimère permanent non obligatoire ; (F) hétérodimère transitoire non obligatoire.