

Structure et activité des protéines

Les sciences du vivant tiennent aussi leurs nanomachines : on les appelle les protéines. Chacune exécute, au sein ou au profit d'une cellule biologique, une ou plusieurs tâches spécifiques, rendant ainsi possibles les différents processus nécessaires au développement et à la survie d'un organisme. **Mieux comprendre les protéines est donc un défi incontournable**, ne serait-ce qu'en vue d'élargir la pharmacopée.

La fonction d'une **protéine** se trouve encodée dans sa **séquence** : celle-ci détermine non seulement sa structure, mais aussi sa dynamique. La dynamique est à la structure des protéines ce que fut la magie au pantin Pinocchio.

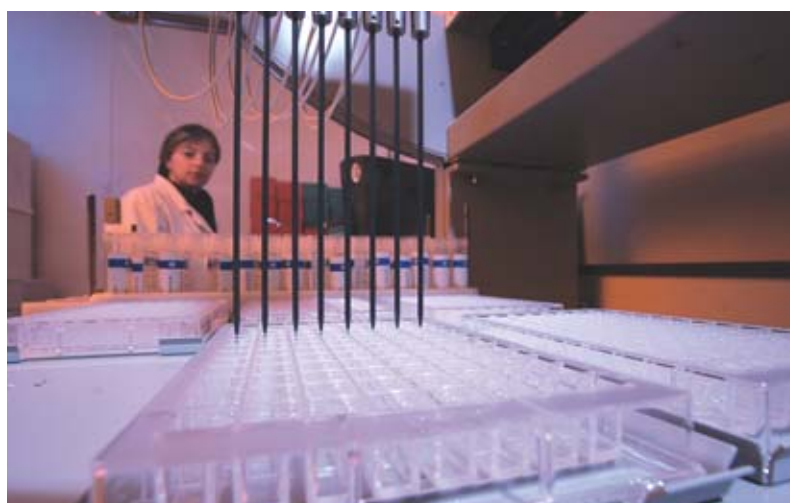
Les nanomachines du vivant

Fabriquées selon les instructions génétiques cryptées dans l'**ADN**, les protéines figurent comme les véritables « ouvrières » du vivant et cela, aussi bien au-dehors qu'au-dedans des cellules biologiques. Dans les processus complexes à l'œuvre au sein des organismes vivants, chacune de ces protéines interprète son, ou ses rôles. Dans le sang, par exemple, l'**hémoglobine** permet d'assurer le transport de l'oxygène, et l'**insuline**, de gérer la régulation du taux de sucre. Ces rôles très spécifiques conduisent à une classification des protéines selon leur type de fonction biologique : **enzymes, hormones, récepteurs, anticorps**, protéines de signalisation, de transport, de stockage...

Une de ces protéines intéresse particulièrement les chercheurs de l'équipe Dynamique structurale des protéines, à l'Institut de biologie structurale (IBS), un institut co-financé par le CEA, le **CNRS** et l'**université Joseph Fourier**. Il s'agit de l'acétylcholinestérase (AChE), « l'interrupteur *off* » des **synapses** cholinergiques, c'est-à-dire des synapses utilisant l'acétylcholine (ACh) comme **neurotransmetteur**. De telles synapses se rencontrent au niveau des jonctions neuromusculaires ainsi que dans les zones du cerveau en charge des fonctions **cognitives** (orientation, mémoire, langage...) : il s'agit donc d'une enzyme essentielle. La modulation de son activité sera souhaitable lors des traitements symptomatiques du **glaucome**, de la **myasthénie** (*myasthena gravis*) ou de la **maladie d'Alzheimer**. Par exemple, rien que pour cette démence, quatre des cinq médicaments approuvés s'avèrent être des inhibiteurs réversibles de l'AChE. Néanmoins, son inhibition complète et irréversible sera létale – par exemple, lors de l'exposition à une dose massive d'insecticides (**méthamidophos, monocrotophos, parathion...**) ou à certains agents **neurotoxiques** (VX, soman, **sarin...**).

Grande diversité des structures tridimensionnelles

Notre « mémoire génétique », l'ADN, s'écrit par succession de « triplets » d'un alphabet à 4 lettres – les bases de l'ADN ; il y a donc $4^3=64$ « caractères » disponibles. Néanmoins, au sein des protéines traduites à partir de l'ADN, on ne trouve que 20 **acides aminés** « naturels ». En effet, les différents types d'acides aminés sont codés par



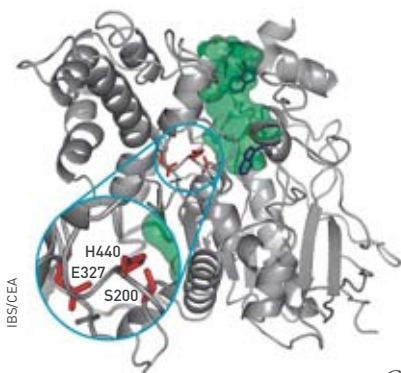
P. Avavian/CEA

un nombre variable de « caractères » de l'ADN : on parle de dégénérescence du **code génétique**.

Pour décrire simplement une protéine, on peut l'envisager comme un « collier » composé par l'alternance de vingt « perles » différentes. Plus ou moins longs, plus ou moins enrichis en chacun des types de « perles », les différents « colliers » de protéines se distinguent avant tout par leur séquence en « perles » : cette séquence va déterminer la structure adoptée par la protéine. Ainsi, malgré l'apparente faible diversité de la « bibliothèque » d'acides aminés naturels, les protéines se présentent sous une multitude de structures tridimensionnelles. La séquence – ou « structure primaire » – d'une protéine dicte également sa dynamique, qui s'avère tout aussi importante du point de vue biologique. En effet, c'est la subtile animation de sa structure, par une combinaison de mouvements plus ou moins longs et plus ou moins amples, qui confère, *in fine*, sa fonction à une protéine. Pour filer la métaphore, figurons-nous un papillon : son envol relève autant du battement de ses ailes que de leur forme.

Les protéines peuvent être classées en deux catégories : les protéines solubles, baignant en solution aqueuse, et les protéines membranaires, autrement dit insérées dans une membrane biologique. Dans ces deux catégories, la vaste majorité des protéines possède une structure tridimensionnelle unique et spécifique, puisque adaptée à leur(s) rôle(s) biologique(s). Néanmoins, des protéines dites « intrinsèquement dépliées » ont été découvertes il y a une décennie environ. Elles se spécifient par l'absence de structure tridimensionnelle unique et sont mieux décrites à partir d'un ensemble de conformations transitant rapidement les unes entre les autres.

Robot de cristallisation des protéines.



IBS/CEA

Figure 1. Vue d'ensemble de la structure de l'acétylcholinestérase. Les hélices- α et feuillets- β sont représentés par des rubans hélicoïdaux et des flèches, respectivement ; ce sont les éléments de structure secondaire. La surface verte correspond à la gorge menant au site actif de l'enzyme ; les bâtonnets rouges figurent les chaînes latérales des acides aminés de la « triade catalytique », c'est-à-dire une sérine, une histidine et un acide glutamique en position 200 [S200], 440 [H440] et 327 [E327], respectivement.

L'AChE entre dans la catégorie des protéines solubles et structurées. Sa structure tridimensionnelle illustre les deux éléments majeurs de « structure secondaire » au sein des protéines (figure 1), à savoir, les hélices- α (représentés par des rubans hélicoïdaux) et les feuillets- β (représentés par des flèches) ; ces derniers se trouvent reliés par des boucles souvent peu structurées.

Comme nous l'avons mentionné plus haut, la fonction principale de l'AChE est de « catalyser » le clivage du neurotransmetteur ACh. Il s'agit donc d'une enzyme, c'est-à-dire une protéine chargée d'accélérer une réaction chimique. Les enzymes sont invariablement structurées, leur structure globale correspondant à l'ossature la mieux adaptée pour le fonctionnement du « site actif » assurant la catalyse. Ainsi celui de l'AChE n'implique que trois protagonistes parmi les quelque 600 acides aminés qui la composent : d'où l'expression de « triade catalytique » (agrandie et en rouge dans la figure 1). Cette triade se situe au fond d'une « gorge » étroite (surface en vert dans la figure 1). Sa position a rendu difficilement compréhensible l'organisation du trafic du substrat (ACh) et des produits au sein de la gorge. Ce trafic est apparu d'autant plus mystérieux que l'AChE figure parmi les enzymes les plus rapides de la nature (jusqu'à 20 000 molécules d'ACh hydrolysées par seconde). Une telle efficacité catalytique ne peut être rationalisée que par la prise en compte de sa dynamique. L'existence d'une porte annexe de sortie, invisible dans la structure de l'enzyme au repos, a notamment été postulée ; son ouverture transitoire à la base du site actif permettrait l'évacuation rapide des produits de catalyse, libérant ainsi le site actif pour une nouvelle molécule de substrat. Nous verrons plus bas comment cette hypothèse a pu être confirmée, grâce à des expériences qui ont littéralement « filmé » l'AChE en action.

Comment déterminer la structure d'une protéine ?

Pour résoudre la structure d'une protéine à résolution atomique, les chercheurs disposent de deux méthodes.

- La **cristallographie aux rayons X**, exigeant que la protéine ait été préalablement cristallisée. La cristallisation d'une protéine est parfois longue et son issue toujours incertaine. La radio-cristallographie correspond néanmoins à une « mesure directe » permettant de localiser chacun des atomes constituant la protéine.

Spectromètre de résonance magnétique nucléaire (RMN) à haut champ (800 MHz) de l'Institut de biologie structurale (IBS).



P. Avrami/CEA



IBS/CEA

Expérience de cristallographie aux rayons X à l'ESRF, sur la ligne de lumière FIP (French beamline for Investigation of Proteins). Cette plateforme du Partnership for Structural Biology, à Grenoble, est financée en partie par le CEA. Cette ligne permet notamment l'évaluation de la qualité de diffraction de très petits cristaux de protéines, sans les extraire de leur boîte de cristallisation (visible en bas à droite).

- La **résonance magnétique nucléaire (RMN)** n'a pas besoin de recourir aux cristaux mais ne concerne que les protéines de relativement petite taille (≈ 300 acides aminés). La reconstruction d'une structure par RMN sera toujours indirecte car elle se base sur des mesures de distances interatomiques dans la protéine. Les modèles produits sont ceux qui satisfont au mieux ces contraintes de distances. Une troisième méthode, la cryo-microscopie électronique, permet, dans certains cas favorables, d'atteindre des résolutions quasi atomiques, voire atomiques.

Par conséquent, 90 % des 70 000 structures de protéines déposées à la Protein Data Bank (PDB) ont été résolues par cristallographie aux rayons X, contre 10 % par la RMN. Le fleurissement, à l'échelle mondiale, de **synchrotrons** dits de 3^e génération a largement contribué à cet état de fait – par exemple, l'European Synchrotron Radiation Facility (ESRF/Installation européenne de rayonnement synchrotron) situé à Grenoble, le synchrotron SOLEIL de Saclay ou l'Advanced Photon Source (APS) de Chicago. Brillants et intenses, les faisceaux de **rayons X** délivrés par ces synchrotrons permettent de réaliser, en une minute, des expériences qui nécessiteraient plusieurs jours avec une source de laboratoire. Le **rayonnement synchrotron** a ainsi permis la résolution de structures, à partir de cristaux de très petite taille et dont le pouvoir diffractant était très faible. C'est notamment le cas du **ribosome**, un complexe **macromoléculaire** géant capable de traduire l'information génétique en protéines, et dont les cristaux présentent les deux carences mentionnées. Sans les synchrotrons de 3^e génération, Venkatraman Ramakrishnan, Thomas Steitz et Ada Yonath n'auraient pu obtenir le prix Nobel de chimie, en 2009, pour leurs travaux sur la structure et la fonction des ribosomes.

Chimie pour le vivant : conception rationnelle de médicaments basée sur la structure de protéines

La plupart des médicaments agissent en ciblant une protéine spécifique, impliquée dans tel ou tel désordre **métabolique**. Selon le cas, sera recherchée l'activation

ou l'inhibition de cette protéine. Concernant l'AChE, cible de nombreuses molécules actives, quatre des cinq médicaments utilisés dans le traitement symptomatique de la maladie d'Alzheimer agissent en modulant son activité. En effet, une diminution du taux d'ACh étant invariablement observée dans le cerveau des victimes de la maladie, l'inhibition réversible de l'AChE va permettre de le restaurer et donc de ralentir la progression de la maladie. Parmi les médicaments inhibiteurs figure le **donépézil**, une molécule qui se lie non seulement à la base du site actif, mais aussi à d'autres acides aminés de la gorge (figure 2). Ceci illustre comment la connaissance de la structure tridimensionnelle d'une protéine peut assister la conception rationnelle d'un médicament.

Quoique la chimie reste maîtresse du jeu en ce qui concerne la synthèse, la connaissance de la structure d'une protéine permettra de simplifier grandement la quête pour de nouvelles molécules actives, guidant la conception (*structure-based drug-design* et *docking*) et rationalisant les observations (résolution de la structure des complexes). Ainsi, un très grand nombre de petites molécules chimiques, constituant une chimiothèque, peut être criblé d'abord *in silico*, puis *in vitro* et *in crystallo*, répondant ainsi aux questions suivantes :

- la molécule mérite-t-elle considération ?
- si oui, la molécule est-elle active ?

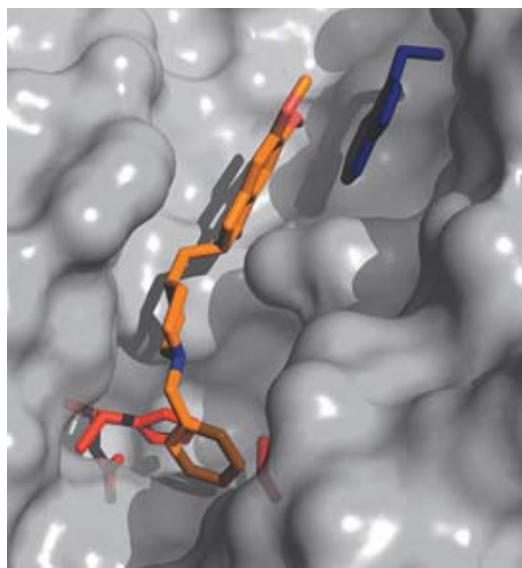


Figure 2. Une molécule de donépézil (orange) interagit dans la gorge (en vert dans la figure 1) avec l'acétylcholinestérase, l'inhibant ainsi réversiblement⁽¹⁾. En rouge sont montrés les acides aminés de la triade catalytique et en bleu un **tryptophane** à l'entrée de la gorge qui sert de site périphérique de fixation de substrats. La donépézil [commercialisée sous le nom Aricept^{MD}] est utilisée pour le traitement palliatif de la maladie d'Alzheimer.

(1) Kryger (G), Silman (I), Sussman (J-L), « Structure of Acetylcholinesterase complexed with E2020 (Aricept) : Implications for the Design of New Anti-Alzheimer Drugs », *Structure* 7, 297-307 (1999).

(2) Colletier (J-P), Fournier (D), Greenblatt (H-M), Stojan (J), Sussman (J-L), Zaccari (G), Silman (I) et Weik (M), « Structural insights into substrate traffic and inhibition in acetylcholinesterase », *Embo Journal.*, 25, 2746-2756 (2006).

(3) Weik (M) et Colletier (J-P), « Temperature-dependent macromolecular X-ray crystallography », *Acta Crystallographica Section D*, 66, 437-446 (2010).

(4) Chapman (H-N), Fromme (P) et al. « Femtosecond X-ray protein nanocrystallography », *Nature*, 470, 73-77 (2011).

- si oui encore, comment rationaliser voire améliorer cette activité ?

Jusqu'à encore quelques années, le *Structure-based drug-design* se limitait au modèle « clef/serrure », selon lequel un médicament épouse forcément la structure de repos du site actif de la protéine-cible. Néanmoins, d'autres conformations dudit site s'avèrent parfois plus propices à la fixation d'un médicament putatif. Le caractère dynamique des structures protéiques se trouve donc pris en compte, de plus en plus régulièrement, dans la conception de médicaments. Ceci est au prix évident d'une complexité accrue, mais reflète mieux la réalité biologique et peut parfois faire une différence majeure.

La cristallographie cinétique pour montrer les protéines en action

Nous connaissons les structures de plus de 70000 protéines mais il ne s'agit encore que d'une collection d'images statiques d'objets pourtant dynamiques. Une meilleure compréhension de leur fonctionnement appelle la caractérisation de leur dynamique structurale, autrement dit de l'ensemble des conformations – proches mais tout de même différentes – qui leur sont accessibles. Pour une enzyme, cette caractérisation s'avère particulièrement importante puisque ce sont les transitions entre sous-états conformationnels qui lui confèrent son activité *in fine*.

La **cristallographie cinétique** est à la cristallographie classique ce que le cinéma est à la photographie : elle a pour but de visualiser les protéines « en action ». Concrètement, il s'agit de déclencher, au sein même du cristal, la réaction biochimique dont une protéine est responsable, puis de caractériser les différents sous-états conformationnels entre lesquels elle transite pour remplir son cycle catalytique. Ici encore, l'utilisation du rayonnement intense d'un synchrotron sera un avantage certain.

Dans le cas de l'AChE, le déroulement du trafic du substrat et des produits enzymatiques depuis l'entrée de la gorge – site périphérique – jusqu'au site actif a pu être « filmé »⁽²⁾ (figure 3). La cristallographie cinétique a également prouvé l'existence d'une porte annexe de sortie pour l'évacuation rapide des produits, dont l'ouverture synchrone a été permise en prenant avantage soit d'un flash **laser**, soit du rayonnement synchrotron lui-même⁽³⁾.

Dans un futur proche, les sources de rayons X de 4^e génération, capables de produire un rayonnement mille milliards de fois plus intense que celui des synchrotrons actuels, fourniront des instantanés de protéines uniques en action (**lasers à électrons libres** ou X-ray Free Electron Lasers/XFEL). La preuve en a été fournie par le XFEL de Stanford (États-Unis)⁽⁴⁾. Deux autres XFELs sont déjà en construction : l'*European XFEL* à Hambourg et la *SPring-8 SASE Compact Source* au Japon. Certes, il serait prématuré d'affirmer que ces machines géantes deviendront des « cinémas moléculaires ». Une chose paraît d'ores et déjà acquise : les rayons X ont encore beaucoup à nous apprendre sur les molécules et processus du vivant.

➤ **Martin Weik et Jacques-Philippe Colletier**
Institut de biologie structurale (IBS)
Direction des sciences du vivant
CEA Centre de Grenoble



Figure 3. Une enzyme en action. Différentes étapes clés du trafic de substrats et de produits au sein de la gorge (verte dans la figure 1) de l'acétylcholinestérase, identifiées par cristallographie cinétique.