

Bac to basics

2000 novembre

Les protéines humaines

Protéiformes et ubiquitaires, les protéines sont au cœur des processus chimiques essentiels à la vie de tous les organismes. Mais combien sont-elles ? Les estimations varient, comme pour le nombre de gènes humains, du simple au quadruple. Quant à l'incroyable diversité des fonctions qu'elles remplissent, tout indique qu'elle est loin d'être recensée...

Les protéines, ça se mange, non ?

Oui, elles représentent une partie importante de notre alimentation. Nos cellules fabriquent leurs propres protéines, mais elles ne savent pas en synthétiser toutes les briques, c'est-à-dire les vingt acides aminés nécessaires. Nous avons donc impérativement besoin d'un apport extérieur en protéines. Nous les trouvons essentiellement dans la viande et les poissons, mais certaines catégories de légumes, choisis de manière équilibrée, peuvent tout aussi bien en fournir en quantités suffisantes. Récemment, des chercheurs ont démontré que le goût typique de la viande est reconnu par des cellules gustatives très sensibles à la présence d'une seule des briques protéiques, l'acide glutamique : le besoin d'aliments riches en acides aminés s'est révélé tellement essentiel pour notre survie que l'évolution nous a donc conduits à favoriser le développement d'un sens spécialisé dans leur reconnaissance. Une fois ingérées, les protéines animales ou végétales sont décomposées dans l'estomac. Une partie des acides aminés est ensuite directement fournie aux cellules pour être incorporée dans de nouvelles protéines humaines.

Pourquoi en avons-nous besoin ?

Les protéines ne sont rien de moins que les molécules qui nous font vivre. Leur nom vient d'ailleurs du grec *protos*, qui signifie le premier, le plus important : il a été forgé par le chimiste suédois J.J. Berzelius en 1838 pour souligner leur rôle primordial dans notre organisme. La plupart des fonctions élémentaires des cellules de notre corps sont remplies par des protéines. Certaines agissent comme des moteurs mécaniques, d'autres sont des catalyseurs chimiques (les enzymes), d'autres encore servent d'éléments structuraux. Certaines hormones (l'insuline, par exemple) sont des protéines, de même que toutes les « antennes » qui transmettent les signaux hormonaux aux cellules. Il est donc peu surprenant d'apprendre qu'un organisme aussi complexe que vous et moi ait besoin de plusieurs dizaines de milliers de protéines

différentes ! En fait, nous ne savons même pas combien nous en possédons. La difficulté de les inventorier provient du fait que chaque type de cellules ne produit qu'un nombre restreint et spécifique de protéines. Certes, toutes les protéines sont représentées dans l'ADN du génome humain que l'on est en train de déchiffrer. Mais on ne dispose pas encore de critères bien établis pour identifier, au sein d'une séquence d'ADN, la présence d'un gène et, de plus, le gène d'une protéine peut être fragmenté et distribué en plusieurs endroits du génome. Le nombre total de nos gènes est encore très spéculatif : d'après les estimations actuelles, il y en aurait entre 30 000 et 120 000 ! Ce n'est qu'avec l'analyse complète du génome humain que nous pourrions savoir enfin à combien de protéines il correspond, et déterminer ainsi ce que l'on appelle le protéome.

Les protéines des autres animaux ressemblent-elles aux nôtres ?

Celles des singes beaucoup, celles des escargots un peu moins... De nombreuses protéines de fonction fondamentale se retrouvent en fait dans tous les organismes que nous connaissons : même si leur composition en acides aminés diffère, elles se ressemblent beaucoup dans leur forme extérieure. Le degré de parenté entre les organismes se reflète dans le degré de parenté entre leurs protéines : pour des protéines apparentées, l'identité d'acides aminés en une position spécifique varie entre 30 % et 100 %. Tous les organismes cellulaires produisent et utilisent au minimum quelques centaines de protéines communes, ce qui signifie que le rôle et la répartition des tâches entre l'ADN, l'ARN et les protéines étaient déjà mis en place à l'époque de l'espèce ancestrale la plus lointaine commune à toutes les espèces actuelles (voir le dossier sur les origines de la vie dans ce numéro). Depuis cette époque, de nombreuses diversifications ont eu lieu, des embranchements se sont créés, éloignés ou même croisés. La duplication accidentelle d'un gène donné peut, par exemple, conduire à un processus de diversification des deux copies dont les protéines correspondantes vont finalement aboutir à des fonctions différentes (évolution divergente). L'effet inverse joue également un rôle important dans l'histoire de l'évolution : des protéines différentes obéissant à des pressions évolutives identiques peuvent se transformer et devenir très semblables (évolution convergente). A cause des ambiguïtés produites par ce genre de phénomènes, il reste toujours très difficile d'étudier l'évolution des organismes en comparant uniquement leurs protéines : on préfère ainsi effectuer des comparaisons sur un

groupe de plusieurs gènes ou bien carrément sur le génome entier d'un organisme par rapport à un autre.

Quelle est la relation exacte entre le génome et les protéines ?

Chaque protéine est définie par la séquence des acides aminés qui la constituent. Au début des années 1930, c'était encore loin d'être une évidence pour tout le monde. On se disputait même pour savoir si une protéine était une molécule unique ou plutôt un assemblage de nombreuses petites molécules. La résolution de la controverse vint de la cristallographie. En effet, en dépit de leur nature flexible et de formes souvent peu symétriques, de nombreuses protéines peuvent former des cristaux réguliers comme ceux des sels. C'est Dorothy Hodgkin et John Bernal qui, en observant que les cristaux de pepsine (une enzyme du suc gastrique) donnaient des réflexions bien distinctes par diffraction des rayons X, démontrèrent pour la première fois en 1934 qu'une protéine possédait une structure bien définie dans l'espace : si les molécules dans le cristal avaient eu des structures différentes ou mal ordonnées, les rayons X auraient été dispersés aléatoirement, et non pas diffusés de manière spécifique dans l'espace comme c'était le cas. Cette conception moderne qu'une protéine était une macromolécule linéaire fut confirmée définitivement par Frederick Sanger, qui détermina en 1953 la séquence des acides aminés de l'insuline, puis par John Kendrew qui acheva de déchiffrer la structure tridimensionnelle de la myoglobine en 1959. La touche finale fut apportée par Christian Anfinsen quand il démontra que sa petite protéine modèle, l'enzyme ribonucléase, pouvait se replier toute seule en partant de la chaîne polypeptidique* complètement dépliée.

On savait désormais que la séquence linéaire des acides aminés était la seule information nécessaire pour qu'une protéine atteigne sa structure tridimensionnelle. Normalement, ce n'est que dans cet état correctement replié, dit état natif, qu'une protéine peut exercer son activité biologique. Par la suite, dans les années 1960, on réussit à déchiffrer le code génétique qui, à chaque acide aminé, fait correspondre un ou plusieurs triplets de bases d'ADN (les codons). On apprit ainsi que la séquence des chaînes polypeptidiques est directement dictée par celle des gènes. Les gènes sont écrits en ADN et cette information précieuse est souvent protégée à l'intérieur d'un noyau. Les gènes actifs dans une cellule donnée sont donc copiés (« transcrits ») sur un ARN messager par une enzyme composée de plusieurs molécules de protéines (l'ARN-synthétase). Cet ARN messager transmet ensuite aux ribosomes

l'information génétique, qu'ils traduisent en suivant le code génétique, aboutissant à la succession des acides aminés demandée. Enfin, ils catalysent la formation des liaisons peptidiques entre les acides aminés successifs. Le ribosome lui-même est constitué de plus de cinquante protéines, mais les deux tiers de son poids proviennent de molécules d'ARN.

Comment une protéine devient-elle active ?

La protéine naissante fabriquée par le ribosome en sort sous la forme d'une chaîne linéaire, mais elle ne possède pas encore son activité biologique, puisque, pour la plupart des protéines, celle-ci dépend de leur structure spatiale. Pour atteindre sa structure spécifique, la chaîne polypeptidique doit encore se replier dans l'espace. Depuis les années 1970, le repliement des protéines a été le sujet de très nombreuses études : les chercheurs espéraient déchiffrer cet autre « code » qui aurait permis de prédire la structure tridimensionnelle d'une protéine en lisant simplement sa séquence. Malgré tous les efforts accomplis, le problème du repliement des protéines n'est toujours pas résolu. Ce que l'on sait, c'est qu'au cours de son repliement, la chaîne polypeptidique doit approcher une conformation favorable par rapport à l'énergie libre du système, en réalisant un grand nombre de liaisons non covalentes, dites faibles. Dans certains cas, le processus va se dérouler « automatiquement » en moins d'une seconde, mais dans d'autres, la protéine néo-synthétisée va devoir se faire aider et faire appel à d'autres protéines, les chaperons moléculaires, ou protéines chaperons.

Et les chaperons, que font-ils au juste ?

Strictement parlant, les chaperons ne sont pas des enzymes, parce qu'ils n'accélèrent pas la réaction dont ils s'occupent. Ils peuvent même ralentir le repliement, car leur véritable rôle est d'augmenter la proportion de molécules qui arrivent au but à atteindre, c'est-à-dire dans leur état correctement replié. Ils garantissent un repliement réussi en protégeant les chaînes dépliées de « liaisons dangereuses » inappropriées qui mèneraient vers des états mal repliés. Ce domaine de recherche étant encore tout récent, les connaissances portent encore beaucoup plus sur des organismes simples, les bactéries, que sur les hommes. En ce qui concerne au moins les bactéries, on pense que les protéines qui ont besoin de cette assistance passent par toute une chaîne de chaperons différents qui gardent la molécule les unes après les autres. Le dernier maillon de la chaîne est une protéine en forme de tonneau qui reçoit les protéines dépliées dans sa cavité. Ce chaperon, nommé GroEL, va fixer et libérer la

protéine au cours de plusieurs cycles, jusqu'à ce qu'elle atteigne son repliement correct.

Quelles sont les activités typiques des protéines ?

Evidemment, si nous possédons des dizaines de milliers de protéines différentes, c'est qu'il y a autant de tâches différentes à accomplir ! Il est probable que nous en ignorons 80 %. Sachant que le groupe des molécules chaperons n'a été découvert qu'il y a dix ans, il est fort possible que des fonctions tout aussi importantes se cachent derrière des protéines que nous ne connaissons même pas. On comprend mieux alors l'enjeu de la recherche sur le génome humain, et surtout le passage à l'étape suivante, l'étude du protéome.

Parmi les protéines que nous connaissons déjà, on peut répertorier quelques groupes importants :

- les enzymes, des catalyseurs qui accélèrent des réactions chimiques ;
- les facteurs de transcription, sortes d'interrupteurs qui contrôlent l'activité des gènes et donc, indirectement, toute l'activité de la cellule ;
- les messagers, porteurs de signaux, qui assurent la communication entre cellules (par exemple l'insuline) ;
- les transporteurs de petites molécules (par exemple, l'hémoglobine qui transporte l'oxygène dans le sang) ;
- les récepteurs, qui reçoivent les messages des hormones et sont responsables des sensations fondamentales comme le goût, la douleur, les odeurs, la lumière, etc. ;
- les protéines moteurs, qui utilisent l'énergie chimique pour produire des mouvements (comme la myosine des muscles) ;
- les anticorps, que l'on appelle immunoglobulines et qui protègent l'organisme en luttant contre les agressions d'organismes étrangers ;
- les protéines structurales (par exemple, la kératine des cheveux ou le collagène).

Et elles poursuivent ces activités jusqu'à la mort de la cellule ?

La durée de vie des protéines est fort variable. Certaines seront dégradées après quelques minutes. En particulier, si leur présence est responsable d'un quelconque signal (hormones, facteurs de transcription, etc.), il faut les éliminer dès que le signal a été reçu. Sinon, le plus souvent, elles survivront plus longtemps, avec une espérance de vie de vingt heures ou davantage. Les protéines du cristallin de nos yeux, quant à elles, représentent une exception remarquable : elles doivent tenir toute une

vie. Si elles commencent à se dénaturer, le cristallin devient opaque, ce qui conduit ensuite à la cataracte.

Comment la cellule se débarrasse-t-elle des protéines dont elle n'a plus besoin ?

Il existe plusieurs chemins, très strictement régulés et contrôlés. Les protéines coûtent relativement cher à la cellule, donc elle doit faire attention à ne pas en détruire plus que nécessaire et inutilement. La voie qui conduit vers la destruction commence avec un marquage moléculaire. Un système bien connu maintenant utilise une petite protéine, l'ubiquitine, pour marquer les protéines à dégrader. Les protéines marquées sont ensuite introduites dans un tonneau macromoléculaire, le protéasome : il s'agit d'une enzyme très agressive qui ôte les « boulons » assemblant les acides aminés. On comprend bien pourquoi cette activité doit être cloisonnée dans un tonneau, dans lequel seules les protéines portant le marquage fatal seront admises.

En quoi les protéines sont-elles utiles pour les biotechnologies ?

Il existe de plus en plus d'applications de protéines (humaines et autres) en médecine, biotechnologie, et dans l'industrie alimentaire. Parmi les exemples établis depuis les temps historiques, se trouve l'utilisation de la présure dans la fabrication du fromage. Le principe actif de la présure est en fait une enzyme, la chymosine de l'estomac de veau, qui a la propriété de faire cailler le lait (l'enzyme recombinante est fabriquée maintenant à l'échelle industrielle). Cette protéine appartient à la même famille d'enzymes que la pepsine dans notre suc gastrique.

En médecine, on sait depuis longtemps qu'une maladie peut être déclenchée par l'absence ou le dysfonctionnement d'une seule protéine. En 1921, des chercheurs canadiens Frederick Banting et Charles Best ont fait une célèbre découverte : ils ont montré que certains extraits pancréatiques, même s'ils proviennent d'une autre espèce, par exemple le cochon, pouvaient éviter les symptômes jusque-là fatals du diabète. L'injection d'insuline pour soigner le diabète fut la première application d'une protéine isolée comme médicament. L'insuline fut aussi la première protéine humaine produite par génie génétique à l'échelle industrielle. Un certain nombre de protéines sont aujourd'hui utilisées en thérapeutique humaine : l'hormone de croissance pour les enfants et les adolescents qui ont un retard de croissance, l'interféron pour lutter contre le cancer, l'érythropoïétine (EPO) pour favoriser la croissance des globules rouges, divers facteurs de croissance de globules blancs pour soigner certaines leucémies, des

facteurs de coagulation pour les hémophiles, etc. Pour certaines de ces protéines, les applications thérapeutiques sont parfois détournées par des sportifs en mal de performance...

Pour le diagnostic médical, on utilise des enzymes depuis longtemps : par exemple, l'uréase permet de déterminer la concentration de l'acide urique dans l'urine pour diagnostiquer la goutte. Plus récemment, des anticorps sont utilisés pour détecter la présence de certaines molécules dans le sang ou les urines : elles peuvent se retrouver jusque dans des produits à utiliser chez soi, par exemple les tests de grossesse. Pour tous ces domaines, les progrès dans le déchiffrement du génome et dans l'analyse du protéome sont porteurs de potentialités thérapeutiques et technologiques considérables. Et les nouvelles méthodes facilitant la production en masse des protéines, les applications promettent d'être très larges.

Michael Groß, Véronique Receveur