

Les exemples évoqués jusqu'à présent sont des protéines qui se replient – soit sur elles-mêmes, soit autour d'autres protéines – lorsqu'elles accomplissent leur fonction (voir l'encadré page 60). Mais le désordre est souvent une partie intégrante du fonctionnement d'une protéine. Parfois, une région non structurée agit comme un minuteur contrôlant le moment auquel deux sites de liaison se rencontrent : si la région non structurée est longue, les deux sites de liaison passent plus de temps à se chercher que lorsque la région non structurée est courte, car la probabilité de « rencontre » des deux sites est inférieure. Dans certains cas, le fait d'être déstructurée permet à une protéine particulière de se faufiler à travers la membrane cellulaire. De telles protéines non structurées apparaissent dans les axones des neurones, où elles forment des structures protégeant l'axone en occupant le volume aux alentours et en empêchant les enzymes de destruction d'approcher.

On a même constaté que certaines protéines restent non structurées après leur liaison. C'est le cas de la protéine Sic1 de la levure (un organisme unicellulaire à noyau, ou eucaryote unicellulaire) ; elle est attachée à son partenaire grâce à plusieurs petits segments qui se fixent et se détachent en continu sur un site de liaison unique, tandis que le reste de Sic1 reste désordonné.

Un désordre généralisé

Le désordre existe aussi parmi les protéines des organismes plus simples et même des virus. Certains virus nommés phages (ou bactériophages), qui infectent les bactéries, se fixent sur la membrane d'une bactérie *via* des protéines reliées à la capsid du phage par l'intermédiaire de liens flexibles. La protéine d'attachement peut se mouvoir plus facilement dans l'espace que le phage entier et le réorienter lors du processus d'arrimage, facilitant ainsi sa pénétration dans la bactérie.

Jusqu'à présent, les chercheurs du monde entier ont identifié la fonction d'environ 600 protéines partiellement ou totalement non structurées. Mais nous pensons qu'il en existe beaucoup plus. Après tout, les scientifiques n'ont déterminé la structure que d'une petite fraction des millions de protéines qui existent dans l'organisme humain. De nouvelles études bio-informatiques menées par K. Dunker et ses collègues tendent à le confirmer.

L'approche bio-informatique est fondée sur des études théoriques antérieures de protéines, qui suggéraient qu'une chaîne d'acides aminés tout juste formée se repliait différemment selon sa composition. En particulier, les acides aminés qui sont volumineux et hydrophobes – sans affinité pour les molécules d'eau qui entourent naturellement les protéines – ont tendance à se positionner au cœur de la protéine. Au contraire, ceux qui se positionnent à la surface d'une protéine repliée sont généralement petits et hydrophiles – et ont tendance à s'attacher aux molécules d'eau environnantes.

LES BIOLOGISTES ONT DÉCOUVERT

que jusqu'à 35 pour cent des protéines humaines auraient de longues régions non structurées.

K. Dunker a voulu comparer les séquences d'acides aminés de protéines intrinsèquement désordonnées avec celles de protéines rigides. En utilisant divers algorithmes, son équipe a découvert en 1997 que les protéines intrinsèquement désordonnées tendent à être plus riches en acides aminés hydrophiles que les protéines rigides. Ainsi, la proportion d'acides aminés hydrophiles et hydrophobes permettrait de prévoir si une protéine se replie partiellement ou pas du tout.

Pour étudier les conséquences de ces découvertes, l'équipe de K. Dunker a comparé, en 2000, divers génomes du monde vivant. Les bio-informaticiens ont examiné les génomes de plusieurs organismes en recherchant des séquences d'ADN codant de longues chaînes d'acides aminés hydrophiles. Les protéines correspondantes seraient les meilleures candidates pour être au moins partiellement non structurées. Dans les organismes les plus simples, les bactéries et les archées (ou archéobactéries), on pense qu'il y a peu de protéines intrinsèquement désordonnées. Mais chez les eucaryotes (les organismes plus complexes, qui ont des cellules pourvues d'un noyau), les protéines déstructurées semblent être beaucoup plus répandues.

Ces résultats ont été généralisés en 2004 par l'équipe de David Jones, de l'*University College* de Londres, qui a effectué des comparaisons similaires en incluant des données humaines. Les biologistes ont découvert que jusqu'à 35 pour cent de toutes les protéines humaines portent de

longues régions non structurées. Ainsi, environ un tiers des protéines humaines présenteraient de grandes régions pour lesquelles le concept clé-serrure est inadapté.

On ignore la raison d'un tel résultat, mais il est possible que les protéines ayant des caractéristiques structurales de type clé-serrure aient surtout des fonctions enzymatiques et de régulation. Les organismes procaryotes (les bactéries) contiennent tous leurs constituants dans un seul compartiment. Au contraire, les organismes eucaryotes sont dotés de multiples compartiments intracellulaires, tels le noyau, l'appareil de Golgi, les mitochondries, etc. Or, il faut que ces différents orga-

nismes communiquent et ils nécessitent une régulation importante. Les organismes pluricellulaires doivent également disposer de signaux pour coordonner les actions des différentes cellules et des tissus. La protéine p27 peut, grâce à sa flexibilité, transporter des messages chimiques le long des voies de signalisation d'une cellule : ces messages sont codés dans sa conformation, dans ses modifications chimiques telle la phosphorylation, et dans les molécules partenaires auxquelles elle se lie (qu'elle inhibe ou régule).

Le secret le mieux gardé de l'évolution

La faible concentration de protéines intrinsèquement désordonnées chez les bactéries signifie peut-être que ces protéines sont apparues tardivement au cours de l'évolution. Toutefois, cela ne semble pas être le cas. D'une part, de nombreux systèmes de signalisation bactériens utilisent plutôt des protéines déstructurées que des protéines structurées. D'autre part, dans les machines moléculaires anciennes (du point de vue évolutif), qui sont constituées d'assemblages d'ARN et de protéines, presque toutes ces protéines sont partiellement ou entièrement déstructurées quand elles ne sont pas liées à leurs ARN « partenaires ». Ces anciens complexes hybrides (complexes ribonucléoprotéiques) comprennent le spliceosome et le ribosome. Les recherches sur l'origine de la vie

LES AUTEURS



Keith DUNKER est biophysicien à l'École de médecine de l'Université de l'Indiana, aux États-Unis, où il dirige le Centre de biologie computationnelle et de bio-informatique. Richard KRIWACKI est biologiste à l'Hôpital pour enfants de St Jude à Memphis.

✓ BIBLIOGRAPHIE

Jensen *et al.*, Intrinsic disorder in measles virus nucleocapsids, *Proc. Natl. Acad. Sci.*, vol. 108, pp. 9839-9844, 2011.

Peter Tompa *et al.*, Structural disorder throws new light on moonlighting. *Trends in Biochemical Sciences*, vol. 30, n° 9, pp. 484-489, 2005.

H. Jane Dyson et Peter E. Wright, Intrinsically unstructured proteins and their functions, *Nature Reviews Molecular Cell Biology*, vol. 6, pp. 197-208, 2005.

A. Keith Dunker *et al.*, Identification and functions of usefully disordered proteins, *Advances in Protein Chemistry*, vol. 62, pp. 25-49, 2002.

Richard W. Kriwacki *et al.*, Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-Bound state : conformational disorder mediates binding diversity, *Proceedings of the National Academy of Sciences USA*, vol. 93, n° 21, pp. 11504-11509, 1996.

Base de données de protéines non structurées connues : www.disprot.org
Base de données de protéines structurées : www.rcsb.org/pdb

plaident aussi en faveur de l'ancienneté des protéines déstructurées. Selon l'une des principales hypothèses, les premiers organismes vivants étaient fondés sur l'ARN et non sur l'ADN. Cet ARN agissait à la fois comme molécule catalytique et comme réservoir de l'information génétique – les rôles qui, dans les cellules actuelles, sont respectivement joués par les protéines et l'ADN.

Dans cette théorie d'un « monde à ARN », une difficulté majeure tient au fait que l'ARN se replie difficilement dans sa forme active et reste souvent figé dans des conformations inactives. Dans les cellules actuelles, des protéines dites chaperonnes assurent un repliement correct de l'ARN et d'autres le stabilisent dans sa conformation active. Ces deux types de protéines n'ont pas de structure stable avant de se lier à l'ARN. Leur apparition aurait ainsi permis de résoudre l'épineux problème du repliement de l'ARN.

Une épée à double tranchant

L'analyse de l'origine du code génétique conforte encore l'hypothèse de l'apparition précoce des protéines non structurées. Le code génétique est l'ensemble des règles utilisées par les cellules pour transcrire l'information codée dans les acides nucléiques (ARN et ADN) en une séquence d'acides aminés. Certains acides aminés auraient été codés très tôt au cours de l'évolution, alors que d'autres seraient apparus plus tard.

Les acides aminés hydrophobes, volumineux, qui poussent une protéine à se replier, seraient apparus tardivement, et donc les protéines constituées des premiers acides aminés devaient très probablement rester dépliées quand elles étaient seules. Si ces hypothèses sur l'évolution du code génétique sont correctes, alors les premières protéines apparues sur Terre se repliaient mal ou pas du tout. Les acides aminés apparus ultérieurement ont manifestement permis aux protéines de se structurer. Elles ont alors acquis des sites enzymatiques actifs de type clé-serrure et, dès lors, ont assuré les fonctions remplies par les ARN dans les cellules pendant des millions d'années.

Étant donné le rôle central que jouent les protéines en biologie, il n'est pas surprenant que beaucoup d'entre elles soient impliquées dans les maladies. Cette

nouvelle conception du désordre intrinsèque des protéines modifiera certainement la façon dont nous comprenons et traitons les maladies. Tout d'abord, dans certains cas, le manque de structure d'une protéine peut être nuisible : quand une cellule produit en excès des protéines non structurées, ces dernières forment des agrégats. Dans le cerveau, ces plaques sont incriminées dans des maladies neurodégénératives telles la maladie d'Alzheimer, la maladie de Parkinson et la chorée de Huntington.

Plus généralement, il semble que la production des protéines non structurées doive être rigoureusement contrôlée pour assurer le fonctionnement normal de l'organisme. Une étude à grande échelle sur la levure, la souris et l'homme, dirigée par Madan Babu, du Laboratoire de biologie moléculaire du Centre de recherche médicale de Cambridge, en Grande-Bretagne, a montré, en 2008, que les cellules régulent les protéines désordonnées plus étroitement que les protéines rigides.

Sachant que les protéines intrinsèquement désordonnées peuvent être impliquées dans certaines maladies, on imagine de nouvelles pistes thérapeutiques. Les protéines qui interagissent avec des protéines non structurées offrent souvent à leurs partenaires des points d'ancrage où les chercheurs pourraient fixer des molécules qui, par exemple, bloqueraient des interactions délétères ou, au contraire, avantageuses. Ainsi, en empêchant un gène suppresseur de tumeur d'interagir avec l'un de ses régulateurs, des biologistes ont obtenu certains succès dans la lutte contre le cancer chez des animaux de laboratoire. Des essais cliniques vont commencer chez l'homme pour tester ces molécules. R. Kriwacki et ses collègues mettent au point une stratégie similaire pour traiter le rétinoblastome, un cancer de l'œil qui touche en particulier les enfants. Les premiers tests chez l'animal ont donné des résultats encourageants.

Le modèle clé-serrure qui a si longtemps expliqué la fonction des protéines a perdu sa toute-puissance. Aujourd'hui, on sait que certaines fonctions biologiques sont remplies par des protéines rigides, et d'autres par des protéines dont la structure fluctue. Mais plus de 100 ans après l'énonciation du dogme voulant que les protéines aient une structure rigide, l'histoire des liens entre la structure des protéines et leurs fonctions reste à écrire. ■