

M/S : médecine sciences



Le repliement des protéines : seconde traduction du message génétique

Protein folding: the second translation of the genetic message

Michel E. Goldberg

Volume 21, Number 6-7, juin–juillet 2005

Repliement des protéines

URI: <https://id.erudit.org/iderudit/011179ar>

[See table of contents](#)

Publisher(s)

SRMS: Société de la revue médecine/sciences
Éditions EDK

ISSN

0767-0974 (print)
1958-5381 (digital)

[Explore this journal](#)

Cite this article

Goldberg, M. E. (2005). Le repliement des protéines : seconde traduction du message génétique. *M/S : médecine sciences*, 21(6-7), 563–566.

Article abstract

Translation of the message encoded in a gene proceeds in two steps: one translates the gene's base sequence into the protein's aminoacid sequence; the second translates the latter into a complex, very precisely defined folded structure, the single one endowed with the specific biological properties. The second step, named « protein folding », is currently a central issue in biological research. The present editorial introduces a series of articles dedicated to this problem. It gives a brief historical overview of how our understanding of the folding mechanisms has evolved, which will be described in more details by J. Yon-Kahn. It also places the problem of protein folding in the context of basic research and its applications in several domains: (1) medicine, where folding diseases (reviewed by G. Grateau) and in particular neurodegenerative diseases caused by prions (see article by R. Melki), are more and more recognized as a public health issue ; (2) biotechnology, where the production of recombinant proteins for research, diagnostic or therapy purposes is often hampered by misfolding when a protein is expressed at high concentration in a foreign host (modern approaches to overcome this difficulty will be discussed by J.M. Betton and A. Chaffotte, and the way in which « chaperones » prevent misfolding in the cell under natural conditions will be described by A.P. Arrigo) ; (3) postgenomics, where the interpretation of the immense amount of base-sequence information collected through genome sequencing needs to be translated as rapidly as possible into structural and functional information (the computer approaches used to predict the 3D structure of a protein from its aminoacid sequence will be outlined by T. Simonson). This series comes timely, at a moment when the efforts of experimentalists, theoreticians, and « users » of protein folding start converging, and boost the power of molecular biology.

La traduction du message contenu dans un gène se fait en deux étapes : l'une traduit la séquence des bases du gène en séquence d'acides aminés de la protéine ; la seconde traduit cette séquence en une structure tridimensionnelle complexe, très précisément définie, seule capable de conférer à la protéine son activité biologique. La seconde étape, ou « repliement de la protéine », est à l'heure actuelle l'un des problèmes centraux de la recherche biologique. Cet éditorial introduit une série d'articles consacrée à ce problème. Il fait un bref survol historique de l'évolution des connaissances sur les mécanismes de repliement et décrit leur impact sur les recherches et les applications en médecine (maladies du repliement, notamment maladies neurodégénératives), en biotechnologie (production de protéines recombinantes) et en postgénomique (génomique structurale et fonctionnelle).

Le problème du repliement

Les macromolécules constituant les cellules vivantes peuvent, schématiquement, être classées en deux catégories : les acides nucléiques, molécules porteuses de l'information génétique, et les protéines, molécules « fonctionnelles » constituant la charpente des cellules vivantes, pour leur donner leur forme et leurs mouvements, ou portant les nombreuses activités biologiques nécessaires au fonctionnement de la cellule et de l'organisme (enzymes, anticorps, hormones protéiques, récepteurs...).

En dehors des très rares cas où il est lui-même porteur d'une activité catalytique, un acide nucléique a pour fonction biologique essentielle de porter le message génétique nécessaire à la synthèse des protéines spécifiques de chaque cellule, de chaque organisme. Ce message se présente sous forme d'une information « linéaire » écrite dans un alphabet à quatre « lettres », les bases nucléiques A, T, G et C. C'est la séquence exacte dans laquelle se succèdent les quelques centaines ou milliers de bases d'un gène qui dicte l'exacte structure chimique de la protéine codée par ce gène. Chimiquement, les protéines se présentent aussi sous la forme d'une information linéaire, écrite cette fois dans un alphabet à vingt lettres, les acides aminés. Et c'est la séquence exacte dans laquelle se succèdent les quelques centaines ou milliers d'acides aminés d'une protéine qui lui confère sa spécificité chimique, dite « structure primaire », d'où résulte sa spécificité fonctionnelle.

Les remarquables avancées de la biologie moléculaire au cours de la seconde moitié du xx^e siècle ont permis non seulement de décoder la correspondance entre la séquence des bases d'un gène et la séquence d'acides aminés de la protéine correspondante (travaux de Marshall W. Nirenberg), mais aussi de connaître en détails, au niveau atomique, la structure de l'énorme complexe macromoléculaire (le ribosome) au sein duquel se produit la traduction du message nucléaire en chaîne d'acides aminés, ou chaîne polypeptidique [1].

Cependant, l'obtention d'une chaîne polypeptidique de séquence appropriée ne suffit pas à l'expression de sa fonction biologique. Pour qu'elle soit « active », il faut en effet que la chaîne polypeptidique adopte une conformation spatiale, un « repliement » sur elle-même, une structure tridimensionnelle très précisément définie où chacun des milliers d'atomes de la protéine est positionné par rapport aux autres à une petite fraction d'angström près. Cette structure, unique parmi un nombre immense de conformations possibles, est appelée conformation « native ». C'est cette conformation tridimensionnelle unique, rigoureusement définie, qui confère à la protéine sa fonction biologique. Ainsi, entre le gène et la forme native existent deux niveaux de traduction de l'information génétique : le premier est la traduction d'une information linéaire (la séquence des bases) en une autre information linéaire (la séquence des acides aminés) ; le second est la traduction d'une information linéaire (la séquence des acides aminés) en une information tridimensionnelle (la structure native) (Figure 1). C'est cette seconde traduction du message génétique que l'on appelle le « repliement » des protéines, problème auquel *médecine/sciences* a choisi de consacrer cette série d'articles.

L'importance du processus de repliement

Comme nous l'avons déjà mentionné, le repliement de la chaîne polypeptidique en sa structure native est indispensable à l'émergence de sa fonction. C'est en effet ce repliement complexe qui positionne à proximité l'un de l'autre, dans un arrangement relatif d'une extrême précision, les atomes de quelques acides aminés souvent éloignés les uns des autres dans la séquence, afin de former le « site spécifique » de la protéine : site catalytique, site de liaison de l'antigène, site de reconnaissance d'un récepteur, site d'interaction entre deux protéines... Mais cette conformation complexe de la chaîne

peptidique a aussi d'autres effets, comme d'assurer la stabilité et la solubilité de la protéine, de mettre la chaîne peptidique à l'abri de l'action des protéases ou, encore, d'éviter des interactions illégitimes de la protéine avec d'autres molécules. Le processus de repliement doit donc nécessairement aboutir à l'enroulement de la chaîne peptidique dans cette conformation native que l'évolution a si subtilement sélectionnée.

Mais d'autres contraintes pèsent sur le processus du repliement pour le rendre encore plus complexe et fascinant. Pour être significatif au niveau de la vie d'une cellule, le repliement doit aussi s'effectuer assez vite pour que l'apparition de la fonction biologique serve à la cellule qui a synthétisé la protéine. Le repliement doit par ailleurs être « efficace » (c'est-à-dire que la majorité des chaînes synthétisées doivent se replier vers la forme native) pour éviter que l'énergie et les métabolites nécessaires à la synthèse de la protéine ne soient gaspillés dans la production de protéines inactives. Enfin, le repliement doit se faire au bon endroit dans la cellule, dans le bon compartiment cellulaire : par exemple, le transport d'une protéine à travers une membrane ne peut se faire efficacement que si la protéine est dépliée ; il faut donc éviter qu'une protéine extracellulaire se replie dans le cytoplasme où elle est synthétisée. Notons que, dans ce phé-

nomène de translocation à travers la membrane, le repliement n'est sans doute pas qu'un phénomène subi passivement par la chaîne peptidique, puisque certaines théories en vogue postulent que l'énergie libérée par la protéine lors de son repliement « de l'autre côté » de la membrane serait la force motrice responsable de sa translocation.

On conçoit donc tout l'intérêt qu'il y a, du point de vue fondamental, à analyser et comprendre les mécanismes moléculaires du repliement, mécanismes abordés d'un point de vue théorique et expérimental dans l'article de Jeannine Yon-Kahn (→). Cependant, depuis une ou deux décennies, l'intérêt des scientifiques pour le repliement déborde largement le cadre strictement conceptuel : en effet, les connaissances acquises par les expérimentateurs et les théoriciens sur le repliement des protéines trouvent des applications dans trois domaines aujourd'hui à la pointe des recherches en biologie et médecine :

- les maladies dites « du repliement », sujet de l'article de Gilles Gateau *et al.* (→), qui résultent de la présence de protéines mal repliées. L'origine peut en être une mutation qui affecte soit la stabilité, soit le mécanisme de repliement d'une protéine. Mais elle peut aussi être un agent infectieux non conventionnel comme le prion, particule infectieuse dépourvue de matériel génétique et constituée d'une protéine mal repliée qui, lorsqu'elle entre en contact avec une protéine homologue normale, l'oblige à changer de conformation pour adopter à son tour la conformation infectieuse. C'est ce type d'agent infectieux qui est responsable de la tristement célèbre maladie de Creutzfeldt-Jakob chez l'homme, de la vache folle chez le bovin, de la tremblante chez le mouton. L'article de Ronald Melki et Luc Bousset (→) fera le point de nos connaissances actuelles sur le prion ;

- la production de protéines recombinantes, qui trouve des applications aussi bien dans le domaine médical (production à bas prix et faible risque de protéines à usage thérapeutique comme l'insuline, l'hormone de croissance humaine...) qu'en recherche fondamentale ou appliquée, et qui se heurte encore souvent à l'expression de protéines dont la structure chimique est parfaite, mais qui ne peuvent adopter spontanément la conformation native dans l'environnement cellulaire où elles sont exprimées et se retrouvent dans la cellule sous forme de « corps d'inclusion », particules faites de protéines mal repliées, insolubles, agrégées. L'article de Jean-Michel Betton et Alain Chaffotte (→) présentera les approches expérimentales et les stratégies nouvelles développées pour surmonter ces difficultés et permettre l'expression de protéines sous leur forme native fonctionnelle ;

- la génomique structurale et fonctionnelle, dont l'une des stratégies prometteuses à long terme est la prédiction de la struc-

(→) m/s
2005, n° 6-7,
p. 601

(→) m/s
2005, n° 6-7,
p. 627

(→) m/s
2005, n° 6-7,
p. 634

(→) m/s
2005, n° 6-7,
p. 613

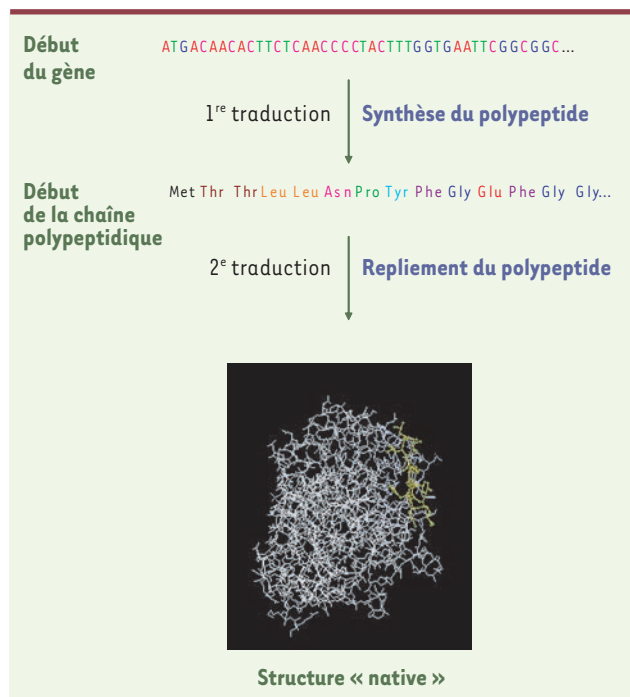


Figure 1. Les deux traductions du gène *trpB* d'*Escherichia coli*. Trois « lettres » (bases nucléiques) du message génétique forment un « mot » pour coder un acide aminé. Ainsi, les trois premières lettres (ATG) codent pour une méthionine (Met), les trois suivantes (ACA) pour une thréonine (Thr), et ainsi de suite. Les 14 premiers acides aminés de la protéine, dont la séquence est indiquée dans la figure, sont indiqués en jaune dans la structure tridimensionnelle de la protéine native.



ture tridimensionnelle et de la fonction de protéines inconnues dont la structure primaire a été révélée par le séquençage des génomes. L'article de Thomas Simonson (→) montrera comment, à l'heure actuelle, on aborde par l'informatique le problème de la « seconde traduction des génomes ».

(→) m/s
2005, n° 6-7,
p. 609

Évolution des idées sur le processus de repliement

L'étude, au niveau moléculaire, du repliement des protéines a longtemps été négligée par les biologistes, au point que l'expression « biologie moléculaire » s'est appliquée aux acides nucléiques et à la 1^{re} traduction, mais pas à la seconde. Cela tient certainement au remarquable succès, dans les années 60, des travaux pionniers du groupe d'Anfinsen [2], travaux qui ont masqué la complexité du processus et donné aux biologistes l'illusion que le problème était résolu. Comme le rappellera J. Yon-Kahn dans son article, les travaux initiaux d'Anfinsen, puis les études du groupe de Tanford sur l'équilibre *in vitro* entre la forme native et la forme dénaturée de quelques petites protéines hautement purifiées semblaient compatibles avec un mécanisme simple d'équilibre entre ces deux formes, sans qu'il y ait entre elles d'intermédiaire partiellement replié. L'idée dominante était donc que l'état natif est identique à l'état thermodynamiquement le plus stable du système protéine/solvant, et que l'acquisition de la forme native est simplement la recherche de la conformation de plus faible énergie. Cependant, la très grande rapidité du processus de repliement observé *in vitro* (de quelques millisecondes à quelques secondes seulement dans bien des cas) excluait la recherche « au hasard » de la conformation de plus faible énergie, et suggérait l'existence d'un chemin de repliement privilégié, jalonné d'intermédiaires partiellement structurés. Ce modèle, s'il expliquait la rapidité du repliement, n'était cependant pas compatible avec le modèle à deux états ci-dessus. De plus, le passage par des intermédiaires « obligés » pouvait parfaitement conduire à un état métastable, distinct de l'état de plus faible énergie, dans lequel la protéine pourrait être « piégée » par une barrière énergétique élevée.

Des études fines à l'équilibre, mais surtout l'utilisation de techniques de cinétique rapide adaptées à l'étude du repliement des protéines ont confirmé l'existence d'intermédiaires de repliement, et donné naissance à des modèles séquentiels d'acquisition progressive de la structure native à travers une succession d'intermédiaires de plus en plus organisés. Cependant, des considérations d'ordre génétique sur la fréquence des mutations affectant le repliement [3], puis des considérations d'ordre physicochimique venues des théoriciens ont bousculé l'idée d'un chemin de repliement unique, et récemment imposé celle de l'existence d'une multiplicité de chemins de repliement permettant chacun à la protéine d'évoluer vers son état natif, guidée en cela par la minimisation de son énergie. C'est la *New view*, la « nouvelle vision », du processus de repliement, selon laquelle la chaîne peptidique testerait une multitude de con-

formations par approximations successives dans un paysage énergétique plus ou moins bosselé [4] : un peu comme une bille qui chercherait à rouler du sommet du Mont-Blanc vers la vallée de Chamonix, ralentie parfois dans tel ou tel creux (état intermédiaire), aboutissant parfois dans une crevasse pour s'y perdre (échec du repliement), mais réussissant souvent à atteindre son but (l'état natif). Comme le montrera l'article de Jeannine Yon-Kahn, nous vivons actuellement une époque enthousiasmante, où les travaux des expérimentateurs et ceux des théoriciens, jusqu'ici menés indépendamment, ont fini par converger pour donner une représentation cohérente, et sans doute utilisable dans l'avenir, du processus de repliement.

Repliement spontané, avorté ou assisté

Après les travaux novateurs du groupe d'Anfinsen sur la renaturation (repliement d'une protéine dénaturée vers son état natif) *in vitro* de la ribonucléase, la littérature scientifique s'est rapidement remplie de publications décrivant la renaturation d'un grand nombre de protéines purifiées, donnant à croire que le repliement spontané était toujours possible. Cela a très longtemps masqué les nombreux essais infructueux, les échecs persistants rencontrés par les chercheurs s'attaquant à certaines protéines. Modestie, ou humiliant sentiment d'incompétence de la part de ces chercheurs qui n'ont pas voulu publier leurs échecs ? Étroitesse de vue du système d'évaluation des articles soumis pour publication dans des revues qui répugnent à publier des expériences infructueuses ? Toujours est-il qu'il a fallu des décennies pour réaliser que le repliement d'une protéine, *in vitro* aussi bien que lors de sa biosynthèse dans une cellule, pouvait échouer dans bien des cas. Ce repliement abortif aboutit très généralement à une forme insoluble, agrégée, inactive de la protéine. Dès les années 70, il est apparu que l'agrégation n'était pas le résultat, mais bien la cause elle-même du mauvais repliement, que l'efficacité du repliement dépendait d'une compétition cinétique entre le « bon » repliement et l'agrégation néfaste [5], et que l'agrégation résultait d'interactions stéréospécifiques entre chaînes polypeptidiques homologues partiellement repliées sous forme d'intermédiaires de repliement [6]. Ce phénomène a été identifié plus tard *in vivo* [7]. La compréhension des interactions responsables de la formation de tels agrégats est à l'origine de l'utilisation d'additifs (dénaturants dilués, détergents, sulfobétaïnes non détergentes, arginine...) ou de procédés (basses températures, faibles concentrations de protéines...) améliorant le résultat du repliement *in vitro*. Et la spécificité des interactions responsables de l'agrégation conduit à l'homogénéité chimique des agrégats, ce qui facilite la purification de la protéine agrégée (voir l'article de J.M. Betton et A. Chaffotte). Cependant, malgré ces « trucs » inventés par les spécialistes du repliement, la renaturation *in vitro* d'une protéine demeure un problème délicat, qui demande de difficiles et souvent longues mises au point spécifiques à chaque protéine.

Qu'en est-il alors du repliement « naturel », dans la cellule, des protéines récalcitrantes à la renaturation *in vitro* ? En fait, la nature s'est heurtée au même difficile problème du repliement, et l'a résolu par l'émergence, au cours de l'évolution, d'une série de protéines qui aident les polypeptides nouvellement synthétisés à se replier correctement. Il s'agit de quelques enzymes (les proline-isoméras, qui accélèrent l'isomérisation lente entre les deux stéréo-isomères de la proline au sein d'une chaîne polypeptidique, et les enzymes impliqués dans l'oxydation et l'échange de ponts disulfures) et des protéines chaperons, une variété de protéines capables de piéger les chaînes mal ou incomplètement repliées pour les empêcher de s'agréger et leur donner une nouvelle chance de progresser sans encombre vers l'état natif. Notons que ces protéines chaperons, très ubiquitaires puisque retrouvées dans tous les organismes vivants, ne sont généralement pas des enzymes : elles ralentissent souvent le repliement (au contraire d'une enzyme dont le rôle est d'accélérer une réaction), mais favorisent le « bon » repliement par rapport aux agrégats dans la compétition cinétique entre repliement natif et agrégation. L'article d'André-Patrick Arrigo (→) sur les protéines chaperons décrira ces macromolécules, essentielles à la vie cellulaire non seulement parce qu'elles assistent de nombreuses protéines dans leur repliement, mais aussi parce qu'elles agissent en tant que régulateurs du trafic intracellulaire, du transport transmembranaire, de la dégradation par les protéases et, très probablement, de diverses modifications chimiques post-traductionnelles.

(→) m/s
2005, n° 6-7,
p. 619

Conclusions

Le repliement des protéines est un domaine de recherche en pleine évolution. Après une longue période de dormance, les applications potentielles des connaissances sur le repliement dans le domaine médical, en biotechnologie et en génomique ont braqué le projecteur sur ce processus complexe, délicatement ciselé par l'évolution pour permettre à une chaîne polypeptidique non seulement d'acquiescer sa structure biologiquement active (sur laquelle s'exerce la pression évolutive), mais aussi de l'acquiescer en un temps étonnamment court, avec une grande efficacité, là où il le faut et quand il le faut. La série d'articles qui suit apportera au lecteur, nous l'espérons, une vision générale actualisée des connaissances acquises sur les différents aspects de ce phénomène fascinant.

SUMMARY

Protein folding: the second translation of the genetic message

Translation of the message encoded in a gene proceeds in two steps: one translates the gene's base sequence into the protein's amino-acid sequence; the second translates the latter into a complex, very

precisely defined folded structure, the single one endowed with the specific biological properties. The second step, named « protein folding », is currently a central issue in biological research. The present editorial introduces a series of articles dedicated to this problem. It gives a brief historical overview of how our understanding of the folding mechanisms has evolved, which will be described in more details by J. Yon-Kahn. It also places the problem of protein folding in the context of basic research and its applications in several domains: (1) medicine, where folding diseases (reviewed by G. Grateau) and in particular neurodegenerative diseases caused by prions (see article by R. Melki), are more and more recognized as a public health issue ; (2) biotechnology, where the production of recombinant proteins for research, diagnostic or therapy purposes is often hampered by misfolding when a protein is expressed at high concentration in a foreign host (modern approaches to overcome this difficulty will be discussed by J.M. Betton and A. Chaffotte, and the way in which « chaperones » prevent misfolding in the cell under natural conditions will be described by A.P. Arrigo) ; (3) postgenomics, where the interpretation of the immense amount of base-sequence information collected through genome sequencing needs to be translated as rapidly as possible into structural and functional information (the computer approaches used to predict the 3D structure of a protein from its aminoacid sequence will be outlined by T. Simonson). This series comes timely, at a moment when the efforts of experimentalists, theoreticians, and « users » of protein folding start converging, and boost the power of molecular biology.

M.E. Goldberg
Unité de repliement et modélisation des protéines,
Institut Pasteur, 28, rue du Docteur Roux,
75724 Paris Cedex 15, France.
goldberg@pasteur.fr

RÉFÉRENCES

1. Zarivach R, Bashan A, Schluenzen F, *et al.* Initiation and inhibition of protein biosynthesis studies at high resolution. *Curr Protein Pept Sci* 2002 ; 3 : 55-65.
2. Anfinsen CB, Haber E, Sela M, White FH. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc Natl Acad Sci USA* 1961 ; 47 : 1309-14.
3. Harrison SC, Durrin R. Is there a single pathway for the folding of a polypeptide chain? *Proc Natl Acad Sci USA* 1985 ; 82 : 4028-30.
4. Wolynes P G, Onuchic J N, Thirumalai D. Navigating the folding routes. *Science* 1995 ; 267 : 1619-20.
5. Orsini G, Goldberg ME. The renaturation of reduced chymotrypsinogen A in guanidine HCl. Refolding versus aggregation. *J Biol Chem* 1978 ; 253 : 3453-8.
6. London J, Skrzynia C, Goldberg ME. Renaturation of *Escherichia coli* tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. *Eur J Biochem* 1974 ; 47 : 409-15.
7. Betts S, King J. There's a right way and a wrong way: *in vivo* and *in vitro* folding, misfolding and subunit assembly of the P22 tailspike. *Structure Fold Des* 1999 ; 7 : 131-9.

TIRÉS À PART

M.E. Goldberg