The "Correctly Folded" State of Proteins: Is It a Metastable State?

Ehud Gazit*

Infectious, genetic, and sporadic protein misfolding diseases are currently of a major medical, social, and even political concern. Such concern calls for a fresh consideration of the central dogmas of the field of protein folding. Experimental data indicate that concentrated aqueous solutions of numerous disease-related but also generic proteins undergo spontaneous aggregation. Such aggregation is associated with a structural transition from the "correctly folded" state of the proteins into a collapsed β-sheet-rich structure. The process of protein aggregation is often thermodynamically irreversible and the large aggregates formed are very stable. These observations and others leads us to suggest that the aggregated form of proteins may represent a true global minimum of free-energy for ensembles of protein molecules in general and not for aggregation-prone proteins only. In line with this hypothesis, the situation of an ensemble of "correctly folded" proteins in solution may represent a kinetically-trapped metastable state. We hypothesize that at an infinite time any protein solution above a critical concentration will eventually undergo structural transition into the aggregated state.

Introduction

There has been an increasing interest in the study of protein misfolding and aggregation in recent years.[1] There are two main reasons for this: 1) The sharp increase in the occurrence of genetic and sporadic amyloidosis and other protein-misfolding diseases, many of which are correlated with the increased life expectancy of the western population. 2) The outbreak of a major infectious protein-misfolding disease, bovine spongiform encephalopathy (BSE), also known as the mad cow disease.

We currently know of more than twenty unrelated protein-misfolding diseases.[1] A partial list of such diseases includes Alzheimer's disease, type II diabetes, Creutzfeldt–Jakob disease (CJD), and various unrelated forms of amyloidosis. Although these diseases are of unrelated origin, they all share several biochemical and biophysical properties. In each case there is a progressive transition of normal, correctly folded cellular proteins into an aggregated state. In the aggregated state the proteins undergo a major structural change from their normal secondary structure into a β-sheet-rich conformation. Usually protein-misfolding diseases are spontaneous ones. For some unknown reason, at a certain stage, mostly in old age, normal proteins are transformed into the pathological aggregated state. In a minority of the cases there is a genetic background in which the proteins prone to aggregation have a different amino acid sequence than the corresponding normal proteins in the population. Such mutations lead to an early onset of the misfolding diseases. Examples for genetic factors are the Flemish mutation in the β-amyloid (βA) peptide, which results in an early manifestation of Alzheimer's disease;[2] the Japanese mutation in the islet amyloid polypeptide (IAPP), which causes an early onset of type II diabetes;[3] and the mutations in the PrP protein that lead to the inherited spongiform encephalopathy forms of CJD, Gerstmann–Straussler–Scheinker disease, or atypical dementia.[4]

Infectious protein diseases or prion diseases are directly linked to the spontaneous and genetic protein-misfolding diseases.[4] Actually, BSE is the infectious form of the spontaneous or inherited CJD. In BSE, the infecting agent is an aggregate of cellular PrP proteins that have undergone structural transition into the infectious Pr(Sc) form. This aggregate of Pr(Sc) interacts with normal, correctly folded PrP cellular proteins and causes an irreversible transition of the latter into the aggregated and misfolded state.

These two forms of protein-misfolding diseases, the spontaneous and the infectious, not only have an enormous medical and social impact; they also raise significant doubts regarding the central dogma of the protein folding field, as is discussed below. The epidemic of BSE in Europe made very clear that infectious protein-misfolding diseases are also a major economical and political matter. Better mechanistic understanding of the aggregation process is needed since, unlike for other major diseases, gene therapy will not be a very useful way to solve protein-misfolding maladies. The interesting fact in all cases of misfolding diseases, from the thermodynamic point of view, is the irreversibility of the aggregation process and the exceptional stability of the aggregates. Actually, the extraordinary stability of the prion infectious agent causes a significant medical problem. The protein aggregates are resistant to disaggregation by disinfecting agents, a situation that raises a major worry regarding the possibility of infection of patients by contaminated surgical and dental tools.[5]

[*] Dr. E. Gazit
Department of Molecular Microbiology and Biotechnology
Tel-Aviv University, Tel-Aviv, 69978 (Israel)
Fax: (+972) 3-640-9407
E-mail: ehudg@post.tau.ac.il

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 1433-7851/02/4102-0257 $ 17.50+.50/0

The "Thermodynamic Theory" of Protein Folding

The central dogma of the protein folding field was put forward by Chris B. Anfinsen more than forty years ago.[6] The basic premise of this theory, labeled the thermodynamic theory, stresses that the energetically favored state of proteins is their "correctly folded" one. In a series of very elegant experiments, Anfinsen and his colleagues proved that all the information needed for a protein to be folded correctly is contained within its primary structure. Moreover, an unfolded protein will spontaneously fold to its correctly folded form, which is the state of the lowest Gibbs energy for a given protein. For this theory Anfinsen was awarded the Nobel Prize in Chemistry in 1972. While it is presumably true that the correctly folded state is the state of lowest Gibbs energy for a given isolated protein molecule—that is, in an infinitely diluted solution as explicitly assumed by Anfinsen's model (see below)—experimental evidence raises doubt as to whether this is valid for concentrated ensembles of proteins in aqueous solution.

Experimental Studies of Protein Aggregation

It is well known that concentrated aqueous solutions of numerous unrelated proteins undergo aggregation, and this process was studied extensively.[1] The aggregation process is normally composed of two parts (Figure 1). The first step is a

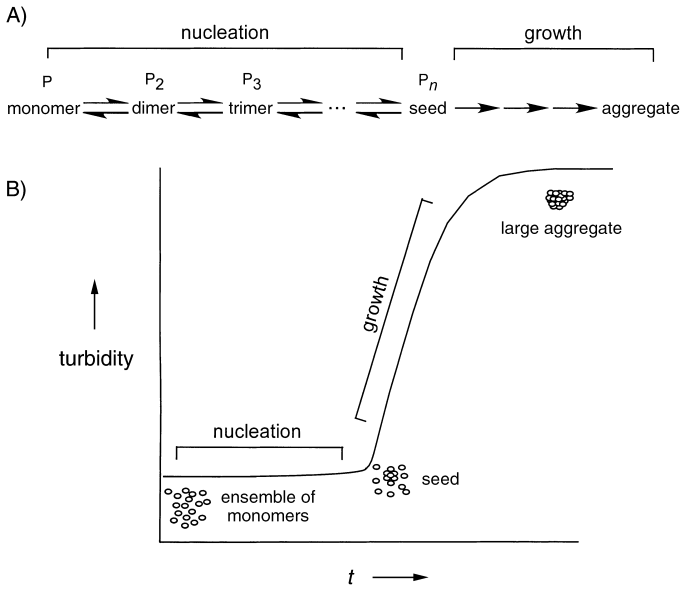

Figure 1. A) The mechanism of protein aggregation involves two stages: A nucleation process is followed by growth of the nuclei to a critical mass. B) A typical aggregation curve. Aggregation is usually determined by monitoring the turbidity of the solution as a function of time. The required initial nucleation results in a delayed appearance of solution turbidity.

reversible nucleation. At this stage proteins are sequentially added to a growing nucleus in a thermodynamically reversible manner. The second part of the process occurs after the growing nucleus reaches a critical mass. Then there are further irreversible additions of protein molecules to the nucleus. This eventually leads to the formation of a large aggregate. The aggregates are very stable, and there is no evidence for an equilibrium between the aggregated state and the correctly folded state.

It is very interesting that not only disease-related proteins undergo aggregation in solution. It was also observed that disease-unrelated proteins—such as the SH3 domain, myoglobin, and a bacterial cold shock protein—form aggregated structures in solution.[7] One of the generic types of aggregates formed by unrelated proteins is amyloid fibrils. These fibrils are presumably built by intermolecular backbone hydrogen bonding interactions.[7] These observations led to the suggestion that the aggregated form, and most notably the amyloid form, may represent a generic form of protein.[1] Moreover, non-amyloidal aggregated structures are also extremely interesting since this aggregated state may be the actual agent that causes cell death.[8] It is also common for any person studying protein chemistry to witness that concentrated protein preparations eventually aggregate and precipitate after long storage.

The Effect of Chemical Chaperons

Chemical chaperons are small molecules that have a very robust effect on the thermodynamic stability of various unrelated proteins.[9] The list of chemical chaperons includes naturally occurring organic osmolytes as well as solvents. Chemical chaperons are assumed to change the thermodynamic landscape of proteins by controlling protein folding through a preferential hydration of the exposed polypeptide backbone and side chains of partially unfolded structures.

A chemical chaperon with a remarkable ability to affect protein folding is trimethylamine *N*-oxide (TMAO). The claim is that agents like TMAO can "force thermodynamically unfolded proteins to fold".[9] Unlike helix-promoting solvents such as trifluoroethanol, TMAO does not seem to favor a particular form of secondary structure. It was recently shown that TMAO can manipulate in vitro the aggregation state of the β-amyloid peptide involved in Alzheimer's disease.[10] Upon addition of TMAO there was a dose-dependent increase in the aggregation of the peptide.[10] We observed that this same effect of TMAO is true for a variety of unrelated proteins (unpublished results). Therefore, the same agent that forces proteins to fold into the correctly folded state also causes protein aggregation. Our molecular explanation for this phenomenon is that in the same manner that TMAO forces proteins into the folded state, it further forces them into the aggregation state by overcoming energy barriers.

The "Metastability Theory" of Protein Folding

The combination of the issues raised above led us to suggest a new way to consider the thermodynamics of protein misfolding. Our approach diverts from the traditional view of each isolated protein molecule as an independent thermodynamic system, but considers the ensemble of proteins in aqueous solution as one large thermodynamic system (Figure 2). According to this view, a solution of proteins that are prone to aggregation but correctly folded represents a system in a transient state that will eventually reach the global Gibbs energy minimum of the aggregated state. The fact that many

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 1433-7851/02/4102-0258 $ 17.50+.50/0 *Angew. Chem. Int. Ed.* **2002**, *41*, No. 2

generic proteins undergo similar aggregation processes calls for an expansion of the model. It seems to be that many proteins will sooner or later undergo aggregation. For some proteins this may take minutes or hours and for others it can take days or months. We hypothesize that at infinite time *any* protein solution above a critical concentration C_{ag} will eventually undergo structural transition into an aggregated state. The exact value of C_{ag} varies from one type of protein to the other and is also dependent on parameters such as temperature, ionic strength, and pH. To test this model and to understand protein aggregation as a generic process, future research should be conducted not only for hours and days (as is done, for example, for the βA peptide), but also for weeks or months.

The important point is that the thermodynamically favored state of an ensemble of proteins is likely to be the aggregated

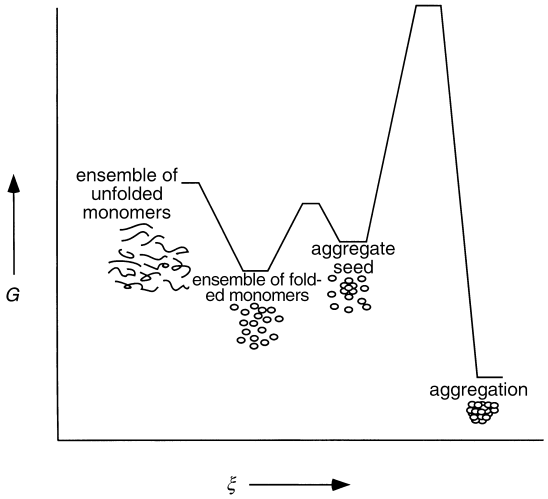

Figure 2. A model for the aggregation of proteins in which the ensemble of proteins in aqueous solution is viewed as one thermodynamic system. Shown is the relative Gibbs energy of unfolded, folded, pre-aggregated, and aggregated proteins. The aggregated state represents the global minimum of Gibbs energy. However, the large activation energy kinetically traps the proteins of the ensemble in the correctly folded state. G = Gibbs energy, ξ = reaction coordinate.

or misfolded one. In line with this model, a concentrated solution of correctly folded proteins represents a metastable state. A classical definition of a metastable state is a state of a local minimum that is isolated from the global minimum by an energy barrier and is hence maybe kinetically trapped. Therefore, in spite of the fact that the metastable state is only a local minimum and not the global minimum of the system, for practical reasons in some cases it is the state observed for a considerable amount of time.

Indeed, some experimental data hints that the transition from the correctly folded state to the global minimum of the aggregated state is achieved by overcoming an energy barrier. Often heating a protein solution causes rapid aggregation. The common explanation for this is that heating causes partial unfolding of the proteins, and the partial unfolding accelerates aggregation. However, this is in contrast to the observation that chemical chaperons, which inhibit partial unfolding, actually accelerate the aggregation process. We speculate that in fact heating helps the protein solution to overcome the large energy barrier separating the aggregated state from the kinetically trapped correctly folded state. In the same manner, chemical chaperons, which force proteins into the folded state, further force them into the aggregation state by overcoming such large energy barriers.

Conclusion

Protein misfolding is a major medical, economical, and even political issue. This leads to the need for a fresh view on protein folding with an emphasis on the misfolded state. While it is true that for isolated protein molecules the correctly folded state is the state of the lowest Gibbs energy, this is apparently not the case for an ensemble of many protein molecules in solution. All the experimental evidence leads to the conclusion that for many proteins the favorable state or the state of the lowest Gibbs energy is the aggregated one. We speculate that this is true for all proteins. Therefore, the correctly folded state of any generic protein ensemble may be a local minimum of Gibbs energy. If this is true, at an infinite time all proteins will be transformed to an aggregated state.

[1] Reviews: a) G. Taubes, *Science* **1996**, *271*, 1493 – 1495; b) J. D. Harper, P. T. Lansbury, Jr., *Annu. Rev. Biochem.* **1997**, *66*, 385 – 407; c) C. M. Dobson, *Trends Biochem. Sci.* **1999**, *24*, 329 – 332; d) C. M. Dobson, R. J. Ellis, *EMBO J.* **1998**, *17*, 5251 – 5254; e) J. D. Sipe, A. S. Cohen, *J. Struct. Biol.* **2000**, *130*, 88 – 98; f) J. W. M. Höppener, B. Ahrén, C. J. M. Lips, *New Engl. J. Med.* **2000**, *343*, 411 – 419.

[2] L. Hendriks, C. De Jonghe, P. Cras, J.-J. Martin, C. Van Broeckhoven, *Ciba Found. Symp.* **1996**, *199*, 170 – 180.

[3] a) S. Sakagashira, T. Sanke, T. Hanabusa, H. Shimomura, S. Ohagi, K. Y. Kumagaye, K. Nakajima, K. Nanjo, *Diabetes* **1996**, *45*, 1279 – 1281; b) S. Sakagashira, H. J. Hiddinga, K. Tateishi, T. Sanke, T. Hanabusa, K. Nanjo, N. L. Eberhardt, *Am. J. Pathol.* **2000**, *157*, 2101 – 2109.

[4] a) S. B. Prusiner, M. R. Scott, S. J. DeArmond, F. E. Cohen, *Cell* **1998**, *93*, 337 – 348; b) S. A. Priola, B. Chesebro, *J. Biol. Chem.* **1998**, *273*, 11980 – 11985.

[5] a) L. Manuelidis, *J. Neurovirol.* **1997**, *3*, 62 – 65; b) E. Zobeley, E. Flechsig, A. Cozzio, M. Enari, C. Weissmann, *Mol. Med.* **1999**, *5*, 240 – 243.

[6] Reviews: a) C. B. Anfinsen, *Science* **1973**, *181*, 223 – 230; b) C. B. Anfinsen, H. A. Scheraga, *Adv. Protein Chem.* **1975**, *29*, 205 – 300.

[7] a) J. T. Jarrett, P. T. Lansbury, Jr., *Biochemistry* **1992**, *31*, 12345 – 12352; b) J. I. Guijarro, M. Sunde, J. A. Jones, I. D. Campbell, C. M. Dobson, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4224 – 4228; c) M. Gross, D. K. Wilkins, M. C. Pitkeathly, E. W. Chung, C. Higham, A. Clark, C. M. Dobson, *Protein Sci.* **1999**, *8*, 1350 – 1357; d) T. Konno, K. Murata, K. Nagayama, *FEBS Lett.* **1999**, *454*, 122 – 126; e) M. Fandrich, M. A. Fletcher, C. M. Dobson, *Nature* **2001**, *410*, 165 – 166.

[8] M. S. Goldberg, P. T. Lansbury, Jr., *Nat. Cell Biol.* **2000**, *2*, E115 – E119.

[9] A. Wang, D. W. Bolen, *Biochemistry* **1997**, *36*, 9101 – 9108; I. Baskakov, D. W. Bolen, *J. Biol. Chem.* **1998**, *273*, 4831 – 4834; P. H. Yancey, J. F. Siebenaller, *J. Exp. Biol.* **1999**, *24*, 3597 – 2603; E. Gazit, R. T. Sauer, *J. Biol. Chem.* **1999**, *274*, 2652 – 2657; J. A. Burrows, L. K. Willis, D. H. Perlmutter, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1796 – 1801.

[10] D. S. Yang, C. M. Yip, T. H. Huang, A. Chakrabartty, P. E. Fraser, *J. Biol. Chem*. **1999**, *274*, 32970 – 32974.

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 1433-7851/02/4102-0259 $ 17.50+.50/0