Published in final edited form as:

Trends Biotechnol. 2007 June ; 25(6): 254–261. doi:10.1016/j.tibtech.2007.03.011.

Protein aggregation *in silico*

Troy Cellmer¹, Dusan Bratko^{2,3}, John M. Prausnitz^{3,4}, and Harvey W. Blanch³

¹National Institutes of Health, Bethesda, MD 20892, USA

²Department of Chemistry, Virginia Commonwealth University, Richmond, VA 23284, USA

³Department of Chemical Engineering, University of California, Berkeley, CA 94720, USA

⁴Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Abstract

Protein aggregation is a challenge to the successful manufacture of protein therapeutics; it can impose severe limitations on purification yields and compromise formulation stability. Advances in computer power, and the wealth of computational studies pertaining to protein folding, have facilitated the development of molecular simulation as a tool to investigate protein misfolding and aggregation. Here, we highlight the successes of protein aggregation studies carried out *in silico*, with a particular emphasis on studies related to biotechnology. To conclude, we discuss future prospects for the field, and identify several biotechnology-related problems that would benefit from molecular simulation.

Protein aggregation and its relationship to biotechnology

Protein aggregation can impose severe limitations on the yields of biologically active proteins at nearly all stages of the manufacturing process [1,2]. During storage, the presence of air and solid interfaces, as well as temperature fluctuations, cause proteins to populate non-native, aggregation-prone states. Aggregated protein reduces the efficacy of formulations and leads to enhanced immunogenicity. Aggregation also limits yields of biologically active products during purification. For example, when *E. coli* is the chosen host for recombinant protein expression, the overexpressed protein often forms intracellular agglomerations termed inclusion bodies. Although inclusion bodies are a relatively pure form of the protein, the protein is not biologically active and therefore must be refolded. During the refolding step, aggregation competes with folding and often imposes severe limitations on refolding yields [3]. Protein aggregation is also the subject of intense biomedical study; it has been associated with more than 20 human diseases, including Alzheimer's and Parkinson's [4]. Finally, there has been considerable interest in using ordered protein aggregates as novel nanomaterials [5,6].

In this article, we highlight the recent successes of molecular simulation studies of protein aggregation. Special attention is devoted to studies related to biotechnology that have experimental implications. Also discussed are the limitations of the technique as well as new avenues for future research.

Protein folding and molecular simulation

All-atom simulations are computationally intractable for all but the smallest of proteins. To access the pertinent time scales, most simulation studies of folding require some detail of the model to be sacrificed: the length of a simulation scales, roughly, as the number of particles to

the third power. The ‘simplified’ models usually consist of a necklace of beads on chain (Figure 1), where each bead represents an amino acid. The solvent is treated implicitly in the potential function that sets the interactions between the different species in the system. ‘Coarse-graining’ these details makes it possible to completely characterize the thermodynamics and kinetics of folding. The success of minimalist models in developing the conceptual understanding of protein folding has been the subject of excellent reviews [7–9].

Simulations of minimalist models have facilitated the construction of free-energy landscapes, which afford a visual description of folding mechanisms [9,10]. Figure 2 shows the free-energy landscape for a lattice-model 64-mer that we have studied extensively [11–13]. These plots describe the conformational preferences of a protein as a function of progress variables that typically reflect the particular similarity (or dissimilarity) of a conformation to its native state. Free-energy landscapes reveal the existence of different thermodynamic states, the free-energy differences between the states, and the barrier heights separating them. The funnel-like nature of the free-energy landscapes has important biological consequences [14].

We have extended the landscape approach to investigate misfolding and aggregation [13]. Figure 2 shows the free-energy surface for the lattice-model 64-mer, simulated with three other chains. Without external agents (e.g. molecular chaperones), the landscape is radically altered: the funnel-like nature disappears and the protein spends nearly equal time in the folded and misfolded states. Misfolding can be directly attributed to interactions with neighboring molecules.

The protein-refolding problem

Protein misfolding due to aggregation is encountered when a therapeutic protein must be refolded from inclusion bodies [3]. During the refolding step, aggregation competes with folding, limiting the yield of biologically active protein. The protein-refolding problem has been investigated extensively using molecular simulation. Gupta and co-workers used multi-chain simulations of a two-dimensional lattice model [15] to study the competition between folding and aggregation. Their studies show that there is an optimum denaturant concentration in the refolding solution that maximizes refolding yield: a slightly elevated denaturant concentration destabilizes aggregation-prone intermediates relative to the native state, and thus promotes folding. Similar behavior is observed, experimentally, for hen egg-white lysozyme [16]. A subsequent study probed the effect of various renaturation protocols on refolding yield [17]. These protocols were designed to mimic experiments such as dialysis refolding and one-step renaturation. The effectiveness of each protocol is strongly dependent on the protein concentration. This finding suggests that these process parameters need to be optimized simultaneously.

Our group examined the protein-refolding problem [18] using an off-lattice 46-mer (Figure 1) originally developed by Honeycutt and Thirumalai [19]. In each simulation, three proteins were equilibrated without the presence of attractive interactions. At time zero, the attractive interactions were initiated, and the proteins folded or aggregated. Figure 3 shows snapshots of an aggregate. Increasing the simulation volume, that is, reducing the protein concentration, increases the refolding yield, as expected; however, the improvement in yield slows as the concentration is further decreased. Because the optimal refolding conditions reflect a balance between yield and overall economics (smaller protein concentrations mean, for example, larger tanks and buffer volumes), our results suggest this optimum might reside at intermediate values of protein concentration.

Aggregation is abated in many refolding experiments by the addition of folding enhancers [3]. Typical folding enhancers include polymers, detergents, certain amino acids and molecular chaperones. However, the proper additive is often protein dependent, and the mechanisms of

additive-promoted refolding are poorly understood. A recent computational study from our laboratory found that a chaperone mimic could improve refolding [20]. As expected from the experiments, the chaperone shielded partially folded species from forming non-native aggregates. A lattice-model study by Liu and co-authors showed that polymer additives improve refolding yields by reducing the population of trapped intermediates [21].

Sequence and structure effects on protein aggregation

In addition to manipulating solution conditions, aggregation behavior can be affected by sequence alterations. One approach is to choose mutations that enhance protein stability. We illustrate this point in Figure 4, which shows melting curves for three variants of a lattice-model protein with identical native conformations, as well as plots of the number of inter-protein contacts as a function of temperature [11]. The degree of inter-protein association is strongly correlated to the decrease in native structure: as the protein unfolds, it exposes its hydrophobic core, which then becomes available for inter-protein interactions. Several experimental and computational studies have examined mutation effects on protein stability. As our focus is to review aggregation, we refer the interested reader to several excellent articles [22–24].

Aggregation can also be prevented through mutations, by disrupting inter-protein interactions; to do so, the amino acids involved in inter-protein association must be identified. Molecular simulations are ideal for this task, provided that the system is computationally tractable. We found that 12 amino acids in the lattice-model 64-mer were particularly prone to form inter-protein interactions [13]. Half of these amino acids were buried in the core of the native state. This was expected because inter-protein interactions increase as the protein unfolds (Figure 4b). Many of the most probable amino acid contacts between chains were also native contacts. This is similar to domain swapping, which has been observed in other simulation studies [25, 26] and experimentally [27,28]. Domain swapping can either facilitate aggregation [29] or inhibit it by forming 'closed' dimers that do not have an interface for further monomer addition [25].

In the earliest molecular simulation work involving multi-protein systems, Istrail and co-authors found that contiguous stretches of 'sticky' residues were conducive to aggregation [30]. A survey of naturally occurring amino acid sequences subsequently showed that contiguous blocks of hydrophobic residues were disfavored in nature [31], consistent with the simulation result. Similar ideas are inferred from experiments. Protein engineering studies show that aggregation is facilitated by segments of strongly interacting amino acids [32]. Further study yielded an algorithm for predicting 'aggregation-prone' amino-acid segments, as well as 'aggregation-susceptible' segments, or segments that will become aggregation-prone once mutated [33]. Kinetic data from a variety of unstructured peptides shows the effects of mutations on aggregation rates can be rationalized using a few physiochemical parameters [34]. Principles derived from these studies were used to rationally design an aggregation-resistant variant of the pharmaceutical peptide calcitonin [35]. Serrano and co-workers have developed the user-friendly web program TANGO to identify the nucleating regions of unstructured peptides [36].

Despite the existence of roughly 20 aggregation-related diseases, it is clear that there is a strong evolutionary pressure to avoid aggregation [37]. Although a stable native structure is one mode of protection, many proteins are natively unfolded, suggesting that nature has employed additional strategies for aggregation abatement. This is indeed the case; many natively unfolded proteins contain a high net charge and low intrinsic hydrophobicity [38]. Furthermore, it has been suggested that strategically placed charged residues act as gatekeepers, and disrupt interactions between strongly interacting amino acid segments by electrostatic repulsion [39]. The effectiveness of this approach has been demonstrated in several systems [40,41] and, when

coupled with an algorithm for determining aggregation-prone amino acid segments, it offers great promise for the *a priori* design of aggregation-abating mutations.

We applied the gatekeeper approach to the refolding of the off-lattice model 46-mer (Figure 1). Disruption of strands one and three (the ‘sticky’ strands) with hydrophobic to hydrophilic mutations leads to a reduction in aggregation (Figure 5). However, despite having the same levels of stickiness, strand one forms inter-protein interactions more often. Consequently, the variants with hydrophobic to hydrophilic mutations on strand one have the highest refolding yields. This result was rationalized in terms of the propensity of each strand to be structured at the early stages of folding when most aggregation events occur. Strand three is flanked in the primary sequence by strands two and four; therefore, it is more likely to form intra-protein interactions during the early steps of folding. Interactions with these strands, which have substantial repulsive character, shield strand three from inter-protein interactions. Fawzi and co-authors observed similar behavior when simulating the self-association of the proteins L and G [42]. They found that protein G aggregates more slowly because it forms an aggregation-protected intermediate early in the folding process. Experiments have also demonstrated that aggregation can be abated by structure formation in non-native states. A protein engineering study of the DNA-repair protein ADA2H (human procarboxypeptidase A2) showed that stabilizing a helix in the unfolded form prevented the protein from forming amyloid fibrils [43]. It is worth mentioning that, traditionally, partial folding has been believed to be conducive to aggregation [44]; however, as the aforementioned studies show, this is not always the case.

Amyloid fibril formation

Toxic protein aggregates have been associated with more than 20 human diseases. Although the form of the toxic species remains controversial, the aggregates eventually adopt a β -sheet-rich, fibrillar structure called an amyloid. Fibrillation of proteins and peptides occurs under moderately destabilizing conditions, and thus poses a hazard during protein storage or production [43]. Protein aggregates can be infectious [45], as in the case of the human form of mad cow disease (also known as new variant Creutzfeldt-Jacob’s disease); therefore, it is imperative that their presence is avoided.

Computational studies pertaining to amyloid fibril formation have been prevalent in recent years. Many have used models at the atomic resolution, to study the conformational preferences of the monomer [46–48] or to evaluate different proposals for aggregate structure [49,50]. Taking the former approach, Armen and co-workers [51] have proposed that the aggregation of polyglutamine, implicated in Huntington’s disease, occurs through an α extended-chain conformation. This result, along with other simulation work from the same laboratory, has led to the supposition that the formation of an α -sheet is a common structural transition in the formation of amyloid fibrils [52,53]. Knowledge of such structures can facilitate the rational design of aggregation inhibitors [51].

Because of the time scales involved, simplified models must be used to observe spontaneous fibrillation. Intermediate-resolution models are well suited for this task because they contain explicit representations of the protein backbone and side chains. This is crucial for studies of amyloid fibril formation because backbone hydrogen bonds are a major stabilizing force of the aggregates. Nguyen and Hall used an intermediate-resolution model to observe amyloid fibril formation by polyalanine peptides [54–57] (Figure 6). A discontinuous molecular dynamics algorithm enabled simulations of systems containing up to 96 peptides. The simulations show that fibrillation is facilitated by the formation of disordered aggregates. These aggregates provide a high local concentration of peptide and therefore facilitate nucleation. Similar behavior has been observed, experimentally, for the aggregation of the prion protein [58]. Following nucleation, fibril growth takes place by the association of β sheets, and by the

addition of monomers to the fibril ends. The fibrils exhibit structural characteristics, such as C_{α} - C_{α} distances, similar to those observed experimentally.

Conclusions and future prospects

The improving market for protein pharmaceuticals and the prevalence of aggregation-related diseases ensures that protein aggregation will continue to be an area of intense research. Molecular simulation serves as an important complement to experimental studies aimed at understanding and abating unwanted aggregation. The importance of the phenomenon, and advances in computational algorithms, will continue to stimulate novel simulation studies.

Biotechnology stands to benefit considerably from this work because many problems of biotechnological interest are amenable to molecular simulation (Box 1). These extend beyond the protein refolding problem discussed within this review. Particularly promising are studies pertaining to the role of excipients in stabilizing formulations. Recent work has dealt explicitly with crowding effects on protein folding [59], and could readily be extended to probe the role of additives on aggregation behavior. Surface effects are also of considerable interest, and much could be learned from work such as that reported by Anderson and co-authors [60]. Another interesting avenue for research would be to probe the effects of shear on protein structure and aggregation. Proteins are constantly exposed to shear forces throughout purification, and shear considerably affects aggregation rates. Molecular simulation stands to offer a unique insight into these issues.

Because all-atom simulations of spontaneous aggregate formation are computationally intractable, coarse-grained models will continue to be used to investigate multi-protein systems. Despite their relative simplicity, they capture many of the aspects of folding and aggregation observed in experiment. Furthermore, by reducing the detail of the protein representation, minimalist-model studies facilitate generalizations. Such generalizations can be used to guide experiments, even if specific answers are not provided. For example, several studies reviewed here showed that structure in the unfolded state blocks aggregation. Thus, empirical approaches currently used to predict mutation effects on peptide aggregation rates could benefit from the addition of a term that accounts for structure-forming propensity.

The simplifications required to simulate multi-protein systems come at the steep expense of certainty in the findings. Although many of the results discussed in this review do have some level of experimental corroboration, it is often indirect. One particularly clever approach to overcome this problem is the use of multi-scale modeling. This approach uses information from coarse-grained models as constraints for studies with all-atom models, which are more easily compared with experimental data [29,61–63], and has been used to probe the aggregation associated with Huntington's, Alzheimer's and dialysis-related amyloidosis.

The best glimpse into the future of molecular simulation comes from recent work using the discontinuous molecular-dynamics algorithm. The efficiency of this approach has made it possible to simulate systems containing ~100 peptides [54–57,64,65]. This is a crucial attribute, because aggregation is inherently a multi-body problem. It is possible that studies on smaller systems might miss important events, such as the formation of critical nuclei. As shown by Nguyen and Hall, simulations containing a large number of peptides make it possible to obtain mechanistic insight into the different stages of aggregation, from nucleus formation to elongation. These simulations could be used to explain kinetic results for specific systems and further refine the aforementioned empirical approaches for predicting mutation effects on aggregation rates. Finally, as with protein folding, free-energy landscapes of fibril-forming systems would facilitate identification of difficult-to-observe intermediates, and consequently reveal new targets for aggregation abatement.

Box 1. Biotechnology-related questions that might be addressed by molecular simulations

- a. What effects do shear forces have on protein structure? What mixing conditions are optimal for aggregation prevention during protein folding (Figure I)?
- b. What are the strategies for preventing protein unfolding and aggregation at interfaces (Figure I)?
- c. How can an excipient be designed to stabilize protein formulations (Figure I)?

Acknowledgements

For financial support, the authors are grateful to the National Science Foundation under award BES-0432625, the Virginia Center of Aging ARDRAF Fund, and to the Office for Basic Sciences of the US Department of Energy.

References

1. Frokjaer S, Otzen DE. Protein drug stability: a formulation challenge. *Nat. Rev. Drug Discov* 2005;4:298–306. [PubMed: 15803194]
2. Chi EY, et al. Physical stability of proteins in aqueous solution: mechanism and driving forces in non-native protein aggregation. *Pharm. Res* 2003;20:1325–1336. [PubMed: 14567625]
3. Clark ED. Protein refolding for industrial processes. *Curr. Opin. Biotechnol* 2001;12:202–207. [PubMed: 11287238]
4. Stefani M, Dobson CM. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med* 2003;81:678–699. [PubMed: 12942175]
5. Zhang S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol* 2003;21:1171–1178. [PubMed: 14520402]
6. Koga T, et al. Controlled self-assembly of amphiphilic oligopeptides into shape-specific nanoarchitectures. *Chemistry* 2006;12:1360–1367. [PubMed: 16163755]
7. Head-Gordon T, Brown S. Minimalist models for protein folding and design. *Curr. Opin. Struct. Biol* 2003;13:160–167. [PubMed: 12727508]
8. Dokholyan NV. Studies of folding and misfolding using simplified models. *Curr. Opin. Struct. Biol* 2006;16:79–85. [PubMed: 16413773]
9. Dinner AR, et al. Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci* 2000;25:331–339. [PubMed: 10871884]
10. Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat. Struct. Biol* 1997;4:10–19. [PubMed: 8989315]
11. Bratko D, et al. Effect of single-point sequence alterations on the aggregation propensity of a model protein. *J. Am. Chem. Soc* 2006;128:1683–1691. [PubMed: 16448142]
12. Cellmer T, et al. Thermodynamics of folding and association of lattice-model proteins. *J. Chem. Phys* 2005;122:174908. [PubMed: 15910070]
13. Cellmer T, et al. Protein-folding landscapes in multichain systems. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:11692–11697. [PubMed: 16081531]
14. Wolynes PG. Energy landscapes and solved protein-folding problems. *Philos Transact A Math Phys. Eng. Sci* 2005;363:453–464. [PubMed: 15664893]
15. Gupta P, et al. Effect of denaturant and protein concentrations upon protein refolding and aggregation: a simple lattice model. *Protein Sci* 1998;7:2642–2652. [PubMed: 9865959]
16. De Bernardes Clark E, et al. Oxidative renaturation of hen eggwhite lysozyme. Folding vs aggregation. *Biotechnol. Prog* 1998;14:47–54. [PubMed: 9496669]
17. Nguyen HD, Hall CK. Effect of rate of chemical or thermal renaturation on refolding and aggregation of a simple lattice protein. *Biotechnol. Bioeng* 2002;80:823–834. [PubMed: 12402328]

18. Cellmer T, et al. The competition between protein folding and aggregation: off-lattice minimalist model studies. *Biotechnol. Bioeng* 2005;89:78–87. [PubMed: 15540197]
19. Honeycutt JD, Thirumalai D. The nature of folded states of globular proteins. *Biopolymers* 1992;32:695–709. [PubMed: 1643270]
20. Bratko D, Blanch HW. Competition between protein folding and aggregation: a three-dimensional lattice-model simulation. *J. Chem. Phys* 2001;114:561–569.
21. Lu D, Liu Z. Molecular simulation of polymer-assisted protein refolding. *J. Chem. Phys* 2005;123:134903. [PubMed: 16223327]
22. Voigt CA, et al. Computationally focusing the directed evolution of proteins. *J Cell Biochem* 2001;37:58–63.
23. Johannes TW, Zhao H. Directed evolution of enzymes and biosynthetic pathways. *Curr. Opin. Microbiol* 2006;9:261–267. [PubMed: 16621678]
24. Eijssink VG, et al. Rational engineering of enzyme stability. *J. Biotechnol* 2004;113:105–120. [PubMed: 15380651]
25. Ding F, et al. Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. *J. Mol. Biol* 2002;324:851–857. [PubMed: 12460582]
26. Clark L. Protein aggregation determinants from a simplified model – cooperative folders resist aggregation. *Protein Sci* 2005;14:653–662. [PubMed: 15689507]
27. Janowski R, et al. 3D domain-swapped human cystatin C with amyloid-like intermolecular beta-sheets. *Proteins* 2005;61:570–578. [PubMed: 16170782]
28. Bennett MJ, et al. Deposition diseases and 3D domain swapping. *Structure* 2006;14:811–824. [PubMed: 16698543]
29. Chen Y, Dokholyan NV. A single disulfide bond differentiates aggregation pathways of β 2-microglobulin. *J. Mol. Biol* 2005;354:473–482. [PubMed: 16242719]
30. Istrail S, et al. Lattice simulations of aggregation funnels for protein folding. *J. Comput. Biol* 1999;6:143–162. [PubMed: 10421520]
31. Schwartz R, et al. Frequencies of amino acid strings in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues. *Protein Sci* 2001;10:1023–1031. [PubMed: 11316883]
32. Chiti F, et al. Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Biol* 2002;9:137–143. [PubMed: 11799398]
33. Pawar AP, et al. Prediction of aggregation-prone and ‘aggregation-susceptible’ regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol* 2005;350:379–392. [PubMed: 15925383]
34. Chiti F, et al. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 2003;424:805–808. [PubMed: 12917692]
35. Fowler SB, et al. Rational design of aggregation-resistant bioactive peptides: reengineering human calcitonin. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:10105–10110. [PubMed: 16006528]
36. Fernandez-Escamilla AM, et al. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol* 2004;22:1302–1306. [PubMed: 15361882]
37. Thirumalai D, et al. Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol* 2003;13:146–159. [PubMed: 12727507]
38. Uversky VN, et al. Why are ‘natively unfolded’ proteins unstructured under physiologic conditions? *Proteins* 2000;41:415–427. [PubMed: 11025552]
39. Otzen DE, et al. Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. *Proc. Natl. Acad. Sci. U. S. A* 2000;97:9907–9912. [PubMed: 10944185]
40. Richardson JS, Richardson DC. Natural β -sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. U. S. A* 2002;99:2754–2759. [PubMed: 11880627]
41. Wang WX, Hecht MH. Rationally designed mutations convert de novo amyloid-like fibrils into monomeric β -sheet proteins. *Proc. Natl. Acad. Sci. U. S. A* 2002;99:2760–2765. [PubMed: 11880628]
42. Fawzi NL, et al. Influence of denatured and intermediate states of folding on protein aggregation. *Protein Sci* 2005;14:993–1003. [PubMed: 15772307]

43. Villegas V, et al. Protein engineering as a strategy to avoid formation of amyloid fibrils. *Protein Sci* 2000;1700:1700–1708. [PubMed: 11045616]
44. Fink AL. Protein aggregation: folding aggregates, inclusion bodies, and amyloid. *Fold. Des* 1998;3:R9–R23. [PubMed: 9502314]
45. Prusiner SB. Prions. *Proc. Natl. Acad. Sci. U. S. A* 1998;95:13363–13383. [PubMed: 9811807]
46. Bernstein SL, et al. Amyloid β -protein: monomer structure and early aggregation states of A β 42 and its Pro19 alloform. *J. Am. Chem. Soc* 2005;127:2075–2084. [PubMed: 15713083]
47. Baumketner A, et al. Structure of the 21–30 fragment of amyloid β -protein. *Protein Sci* 2006;15:1239–1247. [PubMed: 16731963]
48. Baumketner A, et al. Amyloid β -protein monomer structure: a computational and experimental study. *Protein Sci* 2006;15:420–428. [PubMed: 16501222]
49. Ma B, Nussinov R. Stabilities and conformations of Alzheimer's β -amyloid peptide oligomers (A β 16–22, A β 16–35, and A β 10–35): sequence effects. *Proc. Natl. Acad. Sci. U. S. A* 2002;99:14126–14131. [PubMed: 12391326]
50. Ma B, Nussinov R. Molecular dynamics simulations of alanine-rich β -sheet oligomers: insight into amyloid formation. *Protein Sci* 2002;11:2335–2350. [PubMed: 12237456]
51. Armen RS, et al. Characterization of a possible amyloidogenic precursor in glutamine-repeat neurodegenerative diseases. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:13433–13438. [PubMed: 16157882]
52. Armen RS, et al. Pauling and Corey's alpha pleated-sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:11622–11627. [PubMed: 15280548]
53. Armen RS, et al. Anatomy of an amyloidogenic intermediate: conversion of β -sheet to α -sheet structure in transthyretin at acidic pH. *Structure* 2004;12:1847–1863. [PubMed: 15458633]
54. Nguyen HD, Hall CK. Spontaneous fibril formation by polyanalines; discontinuous molecular dynamics simulations. *J. Am. Chem. Soc* 2006;128:1890–1901. [PubMed: 16464090]
55. Nguyen HD, Hall CK. Phase diagrams describing fibrillization by polyaniline peptides. *Biophys. J* 2004;87:4122–4134. [PubMed: 15465859]
56. Nguyen HD, Hall CK. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:16180–16185. [PubMed: 15534217]
57. Nguyen HD, Hall CK. Kinetics of fibril formation by polyaniline peptides. *J. Biol. Chem* 2005;280:9074–9082. [PubMed: 15591317]
58. Serio TR, et al. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* 2000;289:1317–1321. [PubMed: 10958771]
59. Cheung MS, et al. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:4753–4758. [PubMed: 15781864]
60. Anderson RE, et al. Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. *J. Chem. Phys* 2000;112:9167–9185.
61. Urbanc B, et al. Molecular dynamics simulation of amyloid beta dimer formation. *Biophys. J* 2004;87:2310–2321. [PubMed: 15454432]
62. Khare SD, et al. Molecular origin of polyglutamine aggregation in neurodegenerative diseases. *PLoS Comput Biol* 2005;1:230–235. [PubMed: 16158094]
63. Ding F, et al. Direct observation of protein folding, aggregation, and a prion-like conformational conversion. *J. Biol. Chem* 2005;280:40235–40240. [PubMed: 16204250]
64. Urbanc B, et al. *In silico* study of amyloid β -protein folding and oligomerization. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:17345–17350. [PubMed: 15583128]
65. Khare SD, et al. The rate and equilibrium constants for a multistep reaction sequence for the aggregation of superoxide dismutase in amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:15094–15099. [PubMed: 15475574]

TRENDS in Biotechnology

Figure 1.

Snapshots of coarse-grained proteins. The top half of the figure (**a–c**) shows an example of (**a**) denatured-extended, (**b**) denatured-compact, and (**c**) native states of an off-lattice 46-mer studied by Cellmer et al. [18]. There are three amino-acid or bead types: hydrophobic (blue); hydrophilic (red); and neutral (orange). Neutral beads have relaxed dihedral potentials that permit the formation of turns. Note that the core of the protein comprises the hydrophobic beads, and the hydrophilic beads are on the surface. The bottom portion of the figure (**d–f**) shows snapshots of (**d**) denatured-extended, (**e**) denatured-compact, and (**f**) native states of the lattice-model 64-mer studied in [11–13]. Twenty amino acids were used in this model. For the representation shown here, the amino acids are placed into one of four groups: hydrophobic (blue); hydrophilic (red); positively charged (yellow); and negatively charged (green).

TRENDS in Biotechnology

Figure 2.

(Left) Free-energy landscape for a lattice-model 64-mer in isolation. F is the free energy, V_{native} is the potential energy from native interactions, and $V_{\text{non-native}}$ is the potential energy from non-native, intramolecular interactions. (Right) Free-energy landscape for a lattice-model 64-mer in the presence of three other identical molecules. Plots were originally published in [13], copyright the National Academy of Sciences 2005.

TRENDS in Biotechnology

Figure 3. Snapshots of a trimer formed in refolding simulations of the off-lattice 46-mer **(a)** side view, **(b)** top view. The figure clearly shows the hydrophobic core (blue beads) of the aggregate. The two figures are different perspectives of the same snapshot. The coloring scheme is the same as in Figure 1.

TRENDS in Biotechnology

Figure 4.

(a,b) Average number of native contacts (per molecule) in a two-protein system as a function of reduced temperature. These data are similar to those found in denaturing experiments. The blue line is the wild-type (lattice-model 64-mer) protein, the red line is the M54T variant, and the green line is the L36T variant. Average number of inter-protein contacts (per molecule) in a two-protein system as a function of reduced temperature. Inter-protein contacts are used to monitor aggregation. The color coding is the same as in (c). Note that as the amount of native structure decreases, the degree of aggregation increases. This occurs at lower temperatures for the destabilized mutants. (c,d) Snapshots of a dimer consisting of two wild-type proteins. The coloring scheme for (c) is the same as in Figure 1. In (d), each chain has been colored differently, so the proteins can be distinguished.

Figure 5. Mutation effects on refolding yield of the off-lattice 46-mer. The diagram is a two-dimensional representation of its native structure. For simplicity, the hydrophobic strands (one and three) are blue, the remainder is grey.

Figure 6. Snapshots of amyloid fibril formation. Figure courtesy of Hung Nguyen and Carol Hall. The figure was originally published in reference [56] (copyright National Academy of Sciences 2004).

TRENDS in Biotechnology

Figure 1. Representation of the various structural states proteins might adopt in solution, indicating some external factors that affect the various equilibria.