

Unzipping the mysteries of amyloid fiber formation

Andrew D. Miranker*

Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06520-8114

Interest in the phenomenon of protein aggregation is as old as protein research. For example, changes in activity and covalent structure accompanying “coagulum” formation by insulin were important to the elucidation of the molecular origins of hormone action (1). In recent memory, the phenomenon of aggregation served only as an annoyance to the research biochemist and a serious economic consideration in the production and distribution of protein pharmaceuticals. Now, however, one class of aggregate has drawn considerable interest from a range of disciplines. These aggregates are readily identified ultrastructurally by the presence of nonamorphous, filamentous structures, termed amyloid fibers.

Function and Pathology

Two diverse groups share an interest in these systems. The first are those interested in the biological and biomedical relevance. Amyloid fibers are involved in a range of human conditions, e.g., Alzheimer’s disease and type II diabetes (2, 3). The conversion of a normally soluble protein into amyloid can give rise to a gain in toxic function or a loss of function, or can result in occlusion of normal cellular or organ function. Furthermore, there is growing evidence that amyloid can be used productively by organisms (4). For example, in *Escherichia coli*, the deposition of the protein curli as amyloid serves as a substrate for colony and biofilm formation (5). In humans, it has recently been suggested that similar conversions in conformation of cytoplasmic polyadenylation element-binding proteins are central to the maintenance of long-term memory (6). The second group are those interested in the development of novel materials (7, 8). Amyloid fibers template their own assembly, giving rise to reproducible structures on the nanometer scale. The precursors are readily synthesized or are biologically expressed, allowing for a range of derivitization. The reaction conditions for assembly can be environmentally friendly and yet yield fibers that are physically and chemically robust. For example, the fibrillogenic NM domain of the yeast prion Sup35 has recently been derivitized to create conductive gold wires with diameters of ≈ 100 nm (9).

Effective insight into biological problems and control in material sciences require a much deeper understanding of

the physical chemistry of amyloid formation. In this issue of PNAS, Kammerer *et al.* (10) rigorously characterize an important tool for making these insights. This tool is a 17-residue *de novo*-designed peptide, $cc\beta$, in which considerations relevant to amyloid formation were overlaid onto a coiled-coil protein design (11). The coiled-coil protein motif is both a well used motif in biological systems (12) [e.g., in regulation of influenza membrane fusion (13, 14)], and a basic unit of tertiary structure for *de novo* design. The basic structure can be visualized as a seven-residue or heptad repeat, $(\text{abcdefg})_n$, of amino acids with α -helical propensity. The intrinsic pitch of an α -helix is 3.6 residues per turn. By spacing small aliphatic side chains, notably leucine, at positions **a** and **d**, a hydrophobic stripe occurs on average every 3.5 residues. In solution, hydrophobic association gives rise to a supercoiling of the helices into dimer, trimer, or tetrameric assemblies with many of the properties of a folded protein. This association includes hydrogen–deuterium exchange protection, cooperative unfolding transitions, and well defined ^1H NMR spectra.

Structural Properties

To understand the interest and possibilities, it is important to understand the physical properties of amyloid systems. All amyloid fibers are composed of β -strands. This is true regardless of the structural origins of the precursor. For example, islet amyloid precursor polypeptide (IAPP) from type II diabetes is natively unfolded (15), β -2 microglobulin (β 2m) from dialysis related amyloidosis is a β -sandwich (16), and PrP from spongiform encephalopathy is predominantly α -helical (17). After conversion to fibers, these systems all adopt β -strand secondary structure. The even and odd numbered side chains of an extended β -strand form stripes on alternate sides of the polypeptide chain. Hence, it is possible to overlay a binary pattern of hydrophobic and hydrophilic side chains onto a coiled coil. Such patterning gives rise to amphipathicity and subsequent stabilization of β -sheets (18). Further structural insight has been gained from x-ray fiber diffraction studies of aligned fibers (19). This technique enabled the determination that β -strands run orthogonal to the long axis of the fiber. In addition, the hydrogen-bonding network of the

β -sheets is arranged in parallel to the long axis of the fiber. This observable result provokes a number of simple questions. How is it that closely related sequences will alternatively arrange themselves in parallel vs. antiparallel β -sheets? For example, $A\beta_{10-35}$ gives rise to parallel β -strands, whereas $A\beta_{16-22}$ forms antiparallel β -strands (20, 21). The facile incorporation of selective isotopic labels into $cc\beta$ enabled Kammerer *et al.* (10) to use solid-state NMR methods to determine that its strands are arranged in an antiparallel fashion. The origins of determinants of orientation should be readily explored by rational modification of the $cc\beta$ design.

For amyloid investigations, this system is ideal because α -helical coiled coils possess a well defined interior and exterior and are spectroscopically distinct from fibers. In the case of $cc\beta$, crystallographic analysis enabled the trimeric precursor to be clearly characterized at the atomic level; therefore, modifications can be clearly assigned to alternative regions of the protein structure. This finding is essential for probing questions regarding the alternative roles of folded, partially folded, and unfolded states in amyloid assembly. The most basic assertion in amyloid formation is that it is partially folded states that serve as precursors for nucleation (3). In this context, however, studies on larger systems are challenged in their ability to distinguish the importance of stabilization of native states, stabilization of unfolded states, and the disruption of native oligomeric interfaces. For example, monomeric β 2m forms amyloid after exposure to divalent Cu(II) . However, both the folded and unfolded states of the protein are stabilized as a result of separate binding sites (22). In transthyretin amyloidosis, the soluble form is tetrameric. Mutations in transthyretin, which strongly promote familial amyloidosis, affect both tetramer dissociation and monomer stability (23). For $cc\beta$, two variants were produced. The first, $cc\beta$ -Met, substitutes two methionine residues for leucine and alanine at position **f** of the coiled coil. This point is furthest removed from the zipper interface and is a conservative mutation with respect to changes in hydrophobic-

See companion article on page 4435.

*E-mail: andrew.miranker@yale.edu.

© 2004 by The National Academy of Sciences of the USA

ity and intrinsic helicity. Not surprisingly, this variant's thermodynamic properties are comparable to its parent. Remarkably, oxidation of the methionines to methoxide, cc β -MetO, produced a construct wholly incapable of generating fibers. One principle of mutational analysis is to introduce the minimum perturbation required to generate a measurable effect. In the case of cc β -Met/cc β -MetO, a change of polarity with minimal change in size and weight illustrates the importance of specific side-chain-side-chain interactions in establishing and stabilizing the amyloid fiber.

Kinetics of Assembly

A protein placed under amyloidogenic conditions will initially remain soluble. This quiescent period or lag phase is followed by cooperative assembly into the aggregated state. Like crystallization, a hallmark of this process is the capacity of fiber formation to be bypassed by providing exogenous seed from a previously conducted reaction. From a physical chemist's point of view, it is seeding that blurs the distinction between infectious and degenerative forms of amyloid disease. For example, it may be the case that spontaneous conversion of the cellular conformation of prion protein, PrP^C, to a pathological conformation, PrP^{Sc}, takes place on time scales longer than a human lifetime. Only by encounter with preformed PrP^{Sc}, perhaps by ingestion of infected bovine tissue, will endogenous PrP^C convert. This example highlights a central question in amyloid formation. What is the molecular basis of a protein-only species barrier, or rather what is the structural basis of seeding? An extreme example of seeding specificity can be seen in studies of A β fibrillogen-

esis by using L- and D-amino acids (24). A β seed formed from D-A β can seed D-precursor, but not L-precursor (and vice versa). As described above, the β -strand orientations of a fiber project backbone donors and acceptors at the exposed end of the seed. Thus, most stereochemical elements of the peptide are buried from participation in the seed interface. Clearly, a more detailed understanding of fiber ends is required.

The crystalline nature of elongation assembly suggests the potential for stoichiometric inhibition of amyloid by poisoning of the ends. This inhibition can occur biologically, e.g., in the case of inhibition of IAPP fibrillogenesis by substoichiometric amounts of insulin (25). Alternatively, subpeptides whose sequence is based on the precursor of an amyloid can be used. For example, modified forms of A β ₁₇₋₂₁ potentially inhibit polymerization of A β ₁₋₄₀ (26). *In vivo*, the heterogeneous presence of alternative forms of precursor can have a dramatic effect on the efficiency of amyloid formation. For example, in tetrameric transthyretin, the mutation V30M gives rise to familial amyloidosis. However, heterozygous coexpression of V30M and T119M suppresses the disease (23). A systematic approach to understanding these phenomena can be addressed by using the cc β system. For example, the presence of the nonamyloidogenic derivative cc β -MetO resulted in suppression of amyloid formation by cc β -Met.

Peptide Model Systems

The benefit of studying fibrillogenesis in peptides is the ease of implementing design, incorporation of nonnative amino acids, and incorporation of site-specific labels for spectroscopy. Polypeptides have an intrinsic ability to form amyloids, the

sequence preferences of which are readily explored by this approach. For example, systematic analysis of the amyloidogenicity of 6-mer sequences has enabled the description of motifs suitable for genome-wide investigations (27). Whereas intrinsic propensities of polypeptides are readily investigated by this approach, caution must be exercised in studies of the subpeptides of parent amyloidogenic proteins. A consequence of the assertion that any protein can form an amyloid (28) is that any subpeptide of any protein can form an amyloid, provided suitable conditions are found. For example, the putative core of IAPP, IAPP₂₀₋₂₉, was suggested and then identified as amyloidogenic by using conditions of 10 mg/ml in 10% acetic acid (29). By contrast, full-length IAPP readily forms fibers at 10 μ g/ml. The origins of this disparity are structural with long-range aromatic/aromatic contacts central to the establishment of prefibrillogenic structure (30). Inhibition of parent precursor fiber formation by a subpeptide is therefore more suggestive of the subpeptides structural role in amyloid formation. For cc β , long-range interactions are accessible to measurement by using isotopic labeling and solid-state NMR. This permitted careful distinction between intra- and intermolecular contacts between disparate residues, specifically Ala-7 and Leu-14. Clearly, the coiled-coil design permits a breadth of analysis available only to synthetic methods. Furthermore, despite its small size, the cc β sequence yields a system in which the relationship between refolding transitions, oligomerization events, and fibrillogenesis can be systematically compared with those found in biomedical systems.

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health.

- Jensen, H. (1932) *Science* **75**, 614–618.
- Selkoe, D. J. (2003) *Nature* **426**, 900–904.
- Rochet, J. C. & Lansbury, P. T., Jr. (2000) *Curr. Opin. Struct. Biol.* **10**, 60–68.
- Huff, M. E., Balch, W. E. & Kelly, J. W. (2003) *Curr. Opin. Struct. Biol.* **13**, 674–682.
- Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S. & Hultgren, S. J. (2002) *Science* **295**, 851–855.
- Si, K., Lindquist, S. & Kandel, E. R. (2003) *Cell* **115**, 879–891.
- Ryadnov, M. G. & Woolfson, D. N. (2003) *Nat. Mater.* **2**, 329–332.
- Zhang, S. (2003) *Nat. Biotechnol.* **21**, 1171–1178.
- Scheibel, T., Parthasarathy, R., Sawicki, G., Lin, X. M., Jaeger, H. & Lindquist, S. L. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 4527–4532.
- Kammerer, R. A., Kostrewa, D., Zurdo, J., Detken, A., García-Echeverría, C., Green, J. D., Müller, S. A., Meier, B. H., Winkler, F. K., Dobson, C. M. & Steinmetz, M. O. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 4435–4440.
- Ciani, B., Hutchinson, E. G., Sessions, R. B. & Woolfson, D. N. (2002) *J. Biol. Chem.* **277**, 10150–10155.
- Burkhard, P., Stetefeld, J. & Strelkov, S. V. (2001) *Trends Cell Biol.* **11**, 82–88.
- Carr, C. M. & Kim, P. S. (1993) *Cell* **73**, 823–832.
- Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994) *Nature* **371**, 37–43.
- Higham, C. E., Jaikaran, E. T., Fraser, P. E., Gross, M. & Clark, A. (2000) *FEBS Lett.* **470**, 55–60.
- Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987) *Nature* **329**, 506–512.
- Zahn, R., Liu, A., Luhrs, T., Riek, R., von Schroetter, C., Lopez Garcia, F., Billeter, M., Calzolari, L., Wider, G. & Wuthrich, K. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 145–150.
- Rosenbaum, D. M., Roy, S. & Hecht, M. H. (1999) *J. Am. Chem. Soc.* **121**, 9509–9513.
- Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. F. (1997) *J. Mol. Biol.* **273**, 729–739.
- Hou, L. & Zagorski, M. G. (2004) *Biophys. J.* **86**, 1–2.
- Gordon, D. J., Balbach, J. J., Tycko, R. & Meredith, S. C. (2004) *Biophys. J.* **86**, 428–434.
- Eakin, C. M., Knight, J. D., Morgan, C. J., Gelfand, M. A. & Miranker, A. D. (2002) *Biochemistry* **41**, 10646–10656.
- Hammarstrom, P., Wiseman, R. L., Powers, E. T. & Kelly, J. W. (2003) *Science* **299**, 713–716.
- Esler, W. P., Stimson, E. R., Fishman, J. B., Ghilardi, J. R., Vinters, H. V., Mantyh, P. W. & Maggio, J. E. (1999) *Biopolymers* **49**, 505–514.
- Larson, J. L. & Miranker, A. D. (2004) *J. Mol. Biol.* **335**, 221–231.
- Findeis, M. A., Musso, G. M., Arico-Muendel, C. C., Benjamin, H. W., Hundal, A. M., Lee, J. J., Chin, J., Kelley, M., Wakefield, J., Hayward, N. J. & Molineaux, S. M. (1999) *Biochemistry* **38**, 6791–6800.
- de la Paz, M. L. & Serrano, L. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 87–92.
- Dobson, C. M. (1999) *Trends Biochem. Sci.* **24**, 329–332.
- Westermark, P., Engstrom, U., Johnson, K. H., Westermark, G. T. & Betsholtz, C. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5036–5040.
- Padrick, S. B. & Miranker, A. D. (2001) *J. Mol. Biol.* **308**, 783–794.