Normal and Aberrant Biological Self-Assembly: Insights from Studies of Human Lysozyme and Its Amyloidogenic Variants

MIREILLE DUMOULIN, JANET R. KUMITA, AND CHRISTOPHER M. DOBSON*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Received April 3, 2006

ABSTRACT

Studies of lysozyme have played a major role over several decades in defining the general principles underlying protein structure, folding, and stability. Following the discovery some 10 years ago that two mutational variants of lysozyme are associated with systemic amyloidosis, these studies have been extended to investigate the mechanism of amyloid fibril formation. This Account describes our present knowledge of lysozyme folding and misfolding, and how the latter can give rise to amyloid disease. It also discusses the significance of these studies for our general understanding of normal and aberrant protein folding in the context of human health and disease.

1. Introduction

Since its discovery by Alexander Fleming in 1922, human lysozyme has been a rich source of information on many fundamental aspects of protein science. Moreover, in the early 1990s, it was found that two lysozyme variants (I56T and D67H) are associated with a form of hereditary nonneuropathic systemic amyloidosis in which amyloid fibrils containing the full-length variant proteins are deposited in various organs[1] [Figure 1]. More recently, other naturally occurring amyloidogenic variants (F57I, F57I/T70N, W64R, and T70N/W112R)[2−4] and one non-amyloidogenic variant (T70N)[5] have been reported. The depth of knowledge that has accumulated on wild-type lysozyme (for reviews see refs 6 and 7) has enabled detailed studies of the effects of these mutations on the properties of the protein including its folding, stability, and aggregation behavior. In this Account, we summarize the results of these studies and describe how they provide molecular insights into both the mechanism by which lysozyme converts in vitro into its fibrillar state and the origin of the amyloid disorder with which this process is associated in vivo. In addition, we discuss the significance of the study of lysozyme for understanding the more general issues of protein misfolding and the nature of protein deposition diseases. Finally, we show how the results of these studies give rise to new insights into the structures of proteins and the evolutionary pressures that have resulted in their remarkable and varied characteristics.

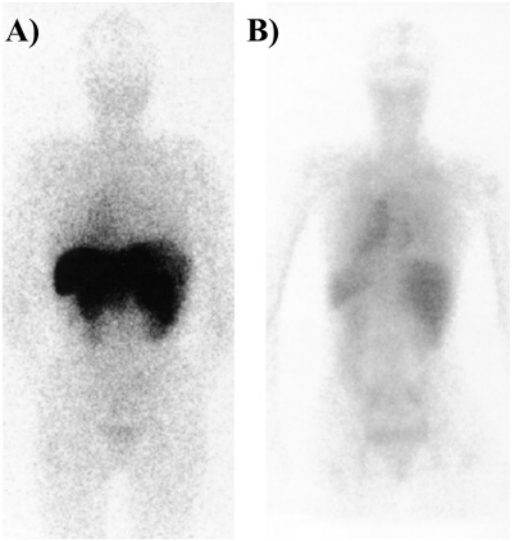

FIGURE 1. Posterior whole-body scintigraphic images following intravenous injection of ^{123}I-human serum amyloid protein that binds selectively to amyloid deposits: (A) from a patient with lysozyme amyloidosis showing heavy amyloid deposition in the liver, spleen, and kidneys; (B) from a normal subject showing the absence of protein deposits and distribution of the tracer throughout the blood pool. Adapted from ref 35.

2. Folding and Unfolding of Human Lysozyme

Wild-type human lysozyme is a 14 kDa enzyme, whose native structure consists of two closely interacting structural regions called the α-domain and the β-domain, and

Mireille Dumoulin was born in Romans sur Isère (France) in 1971. She received her undergraduate degree in Food Sciences in 1994 from the Ecole Nationale Supérieure des Industries Alimentaires (ENSIA, Paris). She then moved to Japan, where she obtained her Ph.D. at the University of Kyoto in 1999, after working on the behavior of proteins under high hydrostatic pressure under the supervision of Prof. Rikimaru Hayashi. The same year, she moved to the Centre for Protein Engineering at the University of Liège for a postdoctoral stay with Dr. André Matagne during which she studied the thermodynamic and folding properties of single-domain antibody fragments from camelids. Since 2001, she has been a postdoctoral fellow in the group of Prof. Christopher Dobson at the University of Cambridge. Her current interests focus on the study of the mechanism of amyloid fibril formation by human lysozyme, using camelid single-domain antibody fragments as structural probes.

Janet R. Kumita was born in Hamilton, Canada, in 1974. She attended the University of Waterloo (Canada) where she received a B.Sc. in Biochemistry. She completed her Ph.D. in Chemistry at the University of Toronto in the laboratory of Prof. G. Andrew Woolley, investigating reversible photocontrol of peptide structure and activity. She began working at the University of Cambridge with Prof. Christopher Dobson in 2003 as an NSERC Postdoctoral Research Fellow and is currently a Postdoctoral Research Associate in the Dobson lab. Her current interests involve studies of the mechanism of amyloid fibril formation by human lysozyme via site-directed mutagenesis, as well as the inhibition of fibril formation by chaperones.

Christopher M. Dobson received his doctorate from the University of Oxford in 1976, having worked on the application of NMR spectroscopy to define the structures and dynamics of proteins in solution. After a short period as a Research Fellow, he moved to Harvard University as an Assistant Professor of Chemistry. In 1980, he returned to the University of Oxford first as a Lecturer and later as Professor and Director of the Oxford Centre for Molecular Sciences. His research interests increasingly focused on defining the mechanism of protein folding and more recently on understanding the consequences of misfolding particularly in terms of its relationship to human disease. In 2001, he moved to the University of Cambridge as John Humphrey Plummer Professor of Chemical and Structural Biology.

* Corresponding author. E-mail: cmd44@cam.ac.uk. Tel: +44 1223 763 070. Fax: +44 1223 763 418.

10.1021/ar050070g CCC: $33.50

Published on Web 07/20/2006

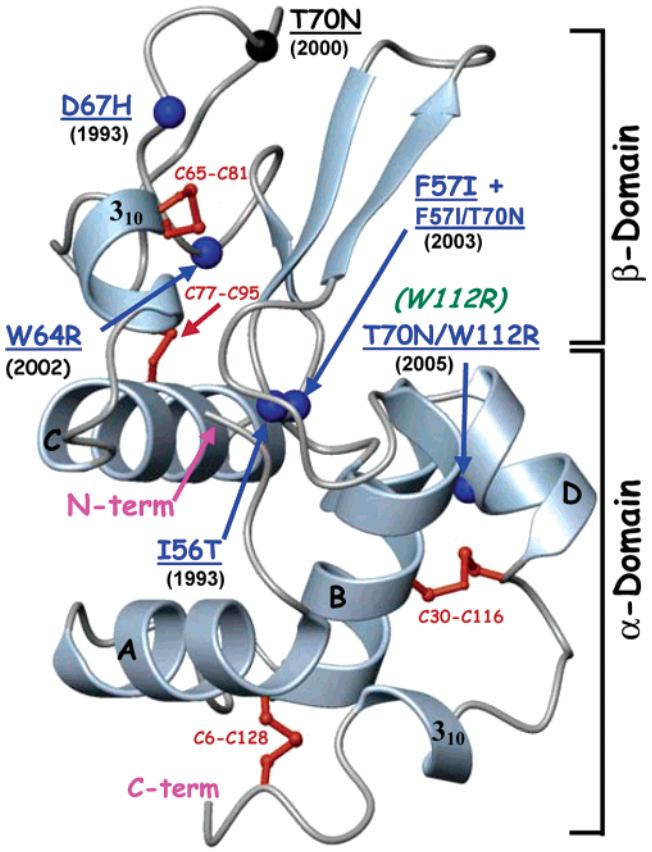

FIGURE 2. Ribbon diagram of the structure of human wild-type lysozyme showing the locations of the known natural mutations, along with the year of their discovery. The six mutations give rise to seven amyloidogenic variants shown in blue and one nonamyloidogenic variant shown in black. The single point mutant W112R has not yet been detected. The four α-helices in the α-domain are labeled A through D. The four disulfide bonds are shown in red.

is cross-linked by four disulfide bonds, one of which (77−95) links the two domains [Figure 2]. All natural mutations discovered so far are located in the β-domain with the exception of the W112R mutation, which is in the D helix of the α-domain [Figure 2]; this latter mutation has, however, only been found in conjunction with the β-domain mutation T70N. X-ray crystallographic data have shown that the I56T, D67H, and T70N variants have the same overall native fold as the wild-type protein with all disulfide bonds correctly formed.[8,9] These variants are functional, with specific activities very similar to that of the wild-type protein.[8,10]

Folding from the Reduced State. Because lysozyme is an extracellular protein, its folding in vivo takes place in the oxidizing environment of the lumen of the endoplasmic reticulum prior to secretion via the Golgi apparatus. The refolding of the wild-type, I56T, and D67H lysozymes from their reduced states in guanidinium chloride (where the protein is highly denatured and with all eight cysteine residues in the free thiol form) has therefore been studied in vitro under oxidative conditions designed to mimic, at least in general terms, protein folding in the endoplasmic reticulum.[11]

Under these conditions, the reduced proteins all correspond to ensembles of unfolded conformers, and in the initial stages of folding these species collapse rapidly via a large number of parallel pathways to form a multitude of relatively unstructured intermediates having one or two disulfide bonds. The majority of these species then fold to form a native-like three-disulfide intermediate lacking the (77−95) bond. The final and slowest step involves a conformational rearrangement requiring at least local unfolding of these latter species in order to allow the remaining free thiol groups to form the fourth disulfide linkage and hence to generate the fully oxidized native protein. Although, the I56T and D67H variants refold in a manner qualitatively similar to that of wild-type lysozyme, an interesting and initially unexpected finding is that they fold faster by a factor of 2 and 3, respectively.[11] The origin of this observation, however, can be attributed primarily to the fact that the lower stabilities of the native-like intermediates of the variants compared with those of the wild-type protein facilitate the conformational rearrangements associated with the final folding step.

Folding in the Presence of Disulfide Bonds. The efficient in vitro oxidative folding of the variant proteins is consistent with the finding that ex vivo lysozyme fibrils are composed only of full-length variant proteins containing all native disulfide bonds.[1,8] Moreover, because these pathological fibrillar deposits are predominantly extracellular, it is likely that the amyloidogenic variants are not only synthesized but also secreted correctly in significant quantities; the folding and unfolding processes underlying their later conversion into amyloid fibrils will therefore in all probability involve only species in which all four disulfide bonds are correctly formed. We and others have studied these processes in vitro in detail for wild-type lysozyme and some of its natural variants. These studies have been described extensively in previous review articles,[6,7] and we shall therefore simply summarize here the most important findings.

The unfolding kinetics of the I56T, D67H, and T70N variants and wild-type lysozyme in high concentrations of guanidinium chloride have all been found to fit well to single-exponential functions. However, the T70N, I56T, and D67H variants unfold, respectively, about 3, 30, and 160 times faster than wild-type lysozyme.[10,12] The refolding of wild-type lysozyme from its guanidinium chloride denatured state occurs via multiple parallel tracks and through a series of well-defined metastable intermediates that become progressively structured in discrete domains. The majority of the protein molecules follow a slow track, the first step of which involves the stabilization of the A and B helices and the C-terminal 3_{10} helix in a locally cooperative manner.[12,13] This step is followed by the cooperative folding of helices C and D leading to an intermediate in which the α-domain has a native-like structure and then by the folding of the β-domain. As found for the homologous hen egg-white lysozyme, it is likely that a final step involving the docking of the two domains is required to generate the native close-packed structure with a functional active site.[6,7] About 10% of the

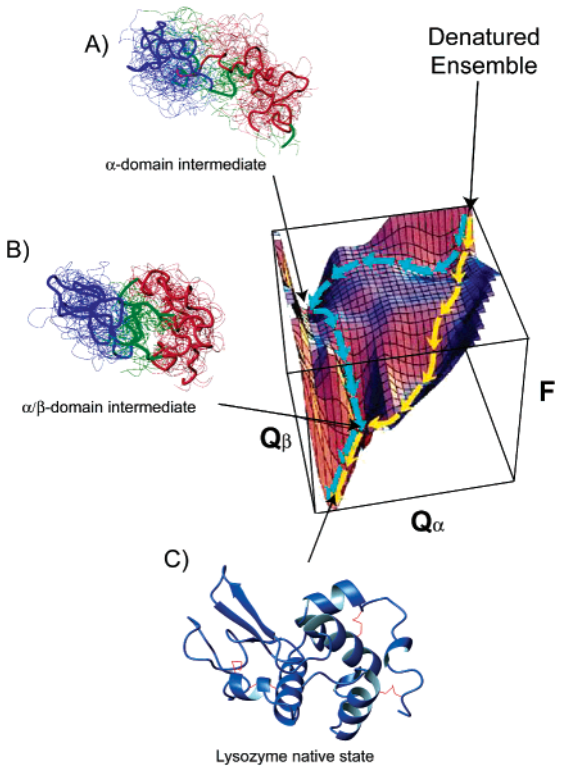

FIGURE 3. A schematic free-energy (F) surface illustrating the folding of hen lysozyme. Q_α and Q_β represent the number of native contacts present in the α-domain and β-domain, respectively, at different stages of folding. Protein molecules are initially in a denatured ensemble and two trajectories are illustrated by which they can achieve the native fold. One represents a "fast track" (yellow trajectory) and shows a lysozyme molecule in which native-like structure in the two domains forms concurrently and the molecule populates an intermediate state (B) briefly prior to achieving the native state (C). Another lysozyme molecule is illustrated following a "slow track" (cyan trajectory) in which the polypeptide chain becomes temporarily trapped in an intermediate state (A), which contains persistent structure in the α-domain but not the β-domain; further folding of this intermediate involves either a transition over a higher barrier or partial unfolding to enable the remainder of the folding process to occur along the fast track. The ensembles of structures shown for the intermediate states (A and B) are based on restrained molecular dynamics simulations using experimental data from studies of human α-lactalbumin, a protein homologous to hen lysozyme and for which stable intermediates can be generated and studied in detail.[36]

molecules, however, fold along a fast track in which the native state is formed more efficiently than it is for the majority of molecules, because the β-domain becomes structured concomitantly with the α-domain. These observations are similar to those obtained for the folding of hen lysozyme and the homologous α-lactalbumins [Figure 3]. All of these studies have contributed to the development of a "new view" of protein folding in terms of a stochastic search process on a highly evolved energy landscape determined by the protein sequence[14,15] [Figure 3].

The refolding of both the I56T and D67H variants also occurs via multiple parallel pathways and through a series of well-defined intermediates. The refolding of the D67H variant is virtually identical to that of the wild-type protein, whereas for the I56T variant, the coalescence of the β-domain onto the folded α-domain is about 10 times slower.[12] These findings can be attributed to the fact that residue 67 is located in a loop region within the β-domain, while residue 56 is located at the crucial interface between the α and β domains.

3. Misfolding and Aggregation of Human Lysozyme

Reduced Stability and Cooperativity of the Amyloidogenic Variants. The thermostability of all four single-point amyloidogenic variants (I56T, F57I, W64R, and D67H) is remarkably decreased by almost exactly the same amount (12 ± 2 °C at pH 5.0) relative to the wild-type protein;[8,16,17] by contrast, the nonamyloidogenic T70N variant is only marginally destabilized (by ~4 °C).[9,10] Moreover, not only do the amyloidogenic mutations drastically reduce the stability of the protein, but they also significantly reduce the cooperativity of the thermal unfolding process.[8,16−18] Most importantly, the global cooperativity of the amyloidogenic variants is substantially reduced at equilibrium under physiologically relevant conditions;[16,19] hydrogen/deuterium exchange experiments monitored by mass spectrometry and NMR spectroscopy have clearly demonstrated that at pH 5−8 and 35−37 °C, both the I56T and D67H variants populate transiently a partially unfolded species [Figure 5a]. In this intermediate species, which is remarkably similar for both variants despite the different locations of the mutations, the regions of the protein that form the β-domain and the adjacent C-helix in the native state are simultaneously unfolded, whereas the regions forming the remainder of the α-domain maintain their native-like structure [Figure 5f].[16,19]

This partial and transient unfolding event is not detectable in vitro for the nonamyloidogenic T70N variant or for the wild-type protein under physiologically relevant conditions; it is, however, detectable for these two proteins under more destabilizing conditions.[9] Taken together, these results suggest that T70N lysozyme is not sufficiently destabilized to allow population of the amyloidogenic intermediate under physiological conditions at a level that results in pathogenic aggregation. This conclusion rationalizes the fact that the T70N mutation has not been found to form amyloid fibrils in people carrying this mutation and suggests that there may be only a relatively narrow window of stability that results in pathogenic behavior. In accord with this conclusion, although patients suffering from lysozyme amyloidosis are all heterozygotes, the deposits appear to be composed exclusively of the amyloidogenic variants.[1,8]

To begin to explore further the link between stability and amyloid disease, we have investigated the levels of secretion of a range of mutational variants of lysozyme in *Pichia pastoris* into which the lysozyme gene has been inserted.[17] In this eukaryotic organism, we have observed a strong correlation between the thermostability of each variant and its secretion level, presumably because the latter reflects a lower probability of satisfying the quality

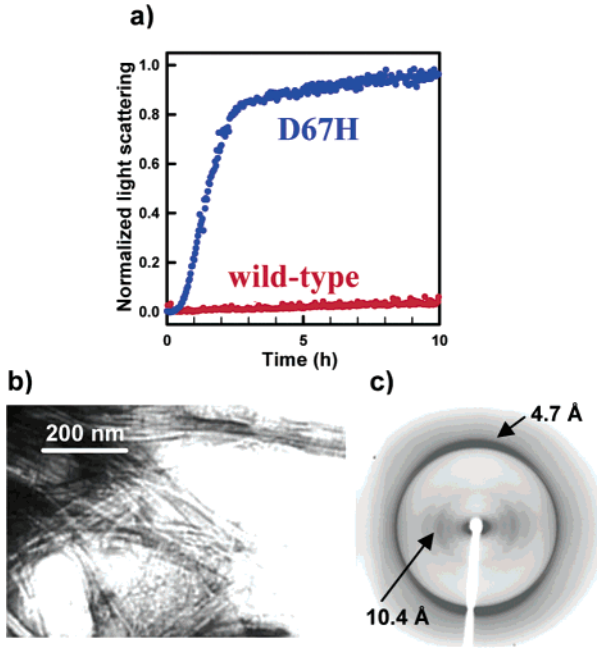

FIGURE 4. In vitro formation of amyloid fibrils by human lysozyme: (a) time course of the aggregation of the D67H and wild-type proteins monitored by light-scattering; (b) representative transmission electron microscopy image of fibrils formed from the D67H variant; (c) X-ray fiber diffraction pattern of the same type of fibrils showing a prominent meridional reflection at 4.7 Å and an equatorial reflection at 10.4 Å, features typical of the cross-β structure of amyloid fibrils. Adapted from ref 23.

control mechanisms within the cell for protein of reduced stability. As far as amyloid disease is concerned, a lower level of protein secretion will, at least in the absence of other factors, reduce the probability of aggregation in the extracellular environment, but the reduced stability of the native state will increase the probability of aggregation. A complex interplay between reduced stability, lower secretion levels, and intrinsic aggregation propensity is therefore likely to determine the types of mutation that give rise to familial forms of amyloid disease.[17]

Links between Normal Folding and Amyloid Formation. The I56T, D67H, and T70N variants, and indeed wild-type human lysozyme, are all able to form amyloid fibrils in vitro when they are incubated under conditions where a significant fraction of the protein molecules are at least partially unfolded, such as low pH, high temperature, or moderate concentrations of denaturant.[6,7] However, as a result of their lower stability and cooperativity, and therefore their higher ability to populate partially unfolded states, the amyloidogenic variants form fibrils in vitro much more easily than the T70N variant and especially than the wild-type protein.[9,18] Fibrils formed in vitro from the various lysozyme species studied in this way do, however, resemble those formed in vivo, and their amyloid character has been shown by a wide variety of techniques,[6,7] some of which are illustrated in Figure 4. The kinetics of in vitro aggregation are sigmoidal, showing a lag phase followed by exponential growth [Figure 4a]. Moreover, fibril formation by I56T, D67H, and wild-type lysozymes are greatly accelerated by seeding with preformed fibrils formed from either a variant or the wild-type protein.[18] These results are all consistent with a nucleation-dependent growth process that increasingly appears to be a common feature of fibril formation.[6,7]

One of the most striking features of the transiently populated intermediates characterized for the amyloidogenic I56T and D67H variants under physiologically relevant conditions is that they have marked similarities not just to each other but also to the dominant α-domain intermediate detected during their nonoxidative folding. This observation can be rationalized on the basis of structural data which suggest that, for both variants, mutations destabilize the interface between the α- and β-domains. Indeed, perturbation of this region is likely to be the key factor in the reduction of the cooperativity between the two domains of the native proteins and of the stability of the β-domain and the C-helix, which are linked via the 77−95 disulfide bond.[8] As a substantial part of the region destabilized by the mutations forms β-sheet structure in the native protein, the initial steps in the aggregation process could simply involve the formation of intermolecular interactions such as hydrogen bonds in the region of the protein that is transiently exposed in the intermediate, rather than the intramolecular interactions characteristic of the native structure [Figure 5c]. These observations suggest that the processes of protein folding and aggregation are closely coupled and underline the need to study both processes in detail to understand properly how they occur and how they are connected to each other.

When the structurally homologous hen lysozyme is incubated at low pH and elevated temperatures, peptide fragments corresponding to parts of the β-domain and to all of the C-helix are readily cleaved from the protein.[20] Remarkably, these fragments correspond almost exactly to the regions of the I56T and D67H human variants that unfold to form the intermediates discussed above. In addition, these fragments have been found to form amyloid fibrils rapidly after purification; by contrast, the fragments corresponding to the rest of the protein remain predominantly soluble under the same conditions. This result indicates that the region of lysozyme corresponding to the β-domain and the C-helix not only can unfold with local cooperativity but also has a higher intrinsic propensity to aggregate than other regions of the protein. Interestingly, a similar result is predicted for human lysozyme by analyzing its sequence with an algorithm developed to rationalize the relative aggregation propensities of different polypeptides.[21] The highly amyloidogenic character of this region is probably the reason all but one of the known amyloidogenic mutations are localized in the β-domain. The fact that the single α-domain mutation, W112R, has only been found in conjunction with the β-domain mutation T70N[4] suggests that this substitution results in additional destabilization of the T70N variant enabling it to form pathogenic aggregates in vivo.

Inhibition of Amyloid Formation. The detail in which the processes of lysozyme folding and amyloid formation are understood provides an important opportunity to design strategies to inhibit the conversion of the variants

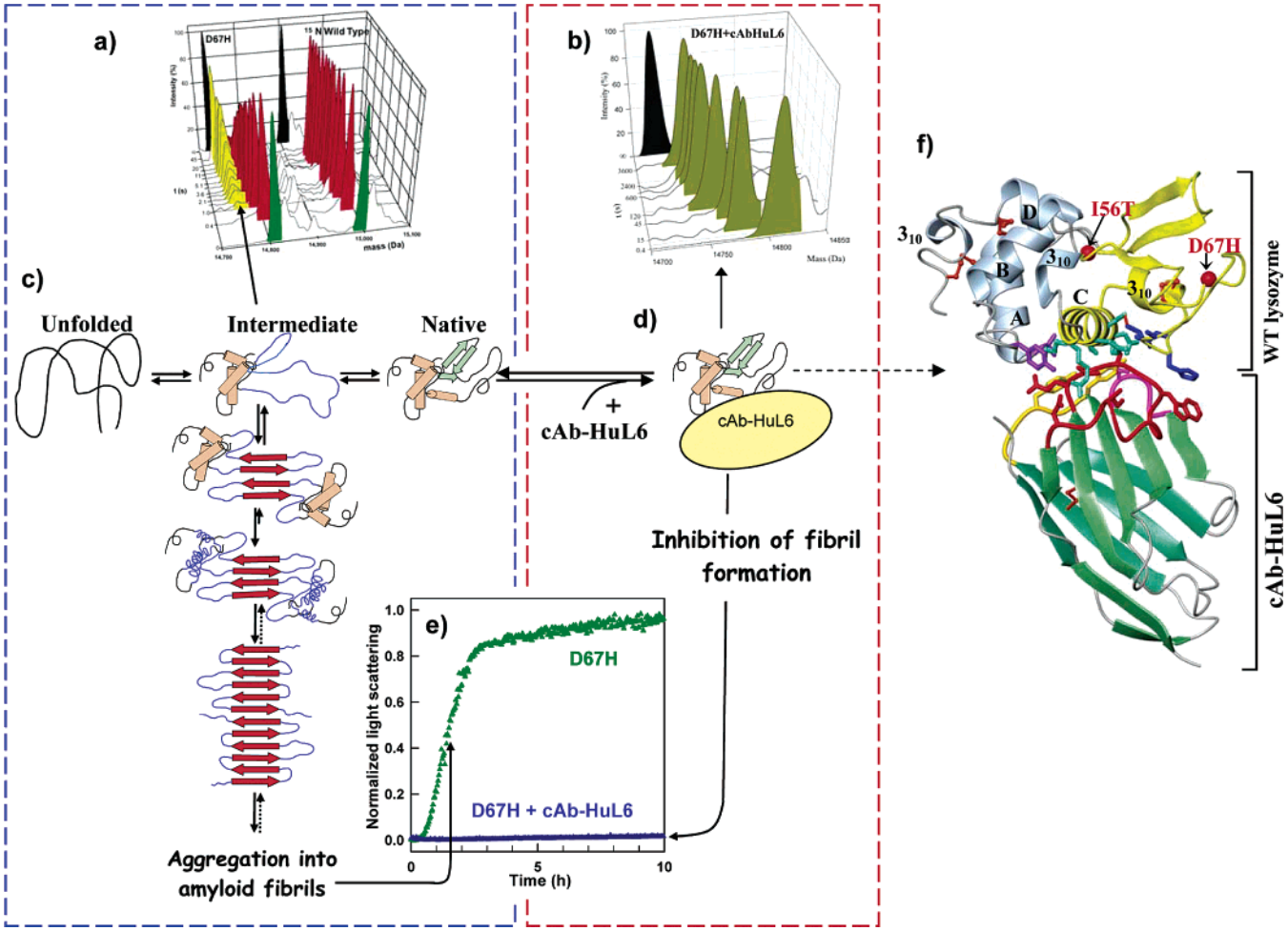

FIGURE 5. Schematic mechanism of human lysozyme aggregation and of its inhibition by the cAb-HuL6 antibody fragment. Panel a shows pulse-labeling hydrogen/deuterium exchange experiments analyzed by mass spectrometry. Two peaks are present in the spectra D67H variant, whereas for wild-type lysozyme under similar conditions a single peak is observed. The additional peak observed for the amyloidogenic variant (yellow) is attributed to the transient formation of an intermediate in which the β-domain and the C-helix are cooperatively unfolded but the remainder of the α-domain is native-like (these regions are colored yellow in panel f). In these experiments, the I56T variant has been found to behave like the D67H variant, whereas the T70N variant behaves like the wild-type protein. Panel c shows the pathway of aggregation of amyloidogenic lysozyme variants. The formation of intermolecular interactions between molecules in the intermediate state is likely to be the origin of the aggregation events that ultimately lead to the formation of amyloid fibrils. Note that the native disulfide bridges, though not represented in this scheme, are present in the fibrils. Panels b and d show stabilization of the lysozyme variants by the antibody fragment. In the presence of cAb-HuL6 (panel b), the D67H variant appears as a single peak indicating that the binding of cAb-HuL6 inhibits its locally cooperative unfolding (as shown on Figure 5a). A similar stabilizing effect was observed with the I56T variant. In panel d, the binding of cAb-HuL6 prevents the ready conversion of the D67H lysozyme into its aggregated state. Again a similar inhibition of aggregation has been observed with the I56T variant. Panel f shows the X-ray structure of the complex between wild-type lysozyme and cAb-HuL6. Adapted from ref 37.

into amyloid fibrils. In light of the conclusion that the lower stability and cooperativity of the amyloidogenic variants is likely to be a primary factor in their enhanced amyloidogenicity, we have explored the possibility of stabilizing these proteins through binding to another molecule. To test this strategy, we have investigated the effects of binding a heavy-chain camelid antibody fragment raised against wild-type human lysozyme and referred to as cAb-HuL6.[22] These experiments reveal that the specific binding of cAb-HuL6 reverses the loss of global structural cooperativity that results from the mutations; that is, in the presence of an equimolar quantity of the antibody fragment virtually none of the molecules of the I56T and D67H proteins undergoes even a single unfolding event of the type observed for the lysozyme variants when unbound. The result of this binding is to prevent the ready in vitro conversion of the lysozyme variants into their aggregated states [Figure 5b,d,e].[16,23]

Structural studies of the complex between wild-type human lysozyme and cAb-HuL6 have revealed that the epitope on the lysozyme molecule includes neither the sites of the mutations nor most of the residues in the region of the structure that is destabilized by the mutations[23] [Figure 5f]. Thus, the effects of binding are not simply to mask this latter region and prevent its unfolding from the remainder of the structure. Rather, it appears that the binding of cAb-HuL6 acts by restoring the global cooperativity of the native structure via, at least in part,

the transmission of small but significant long-range conformational effects through the protein to the interface between the two structural domains. These results demonstrate the close link between the reduction of global cooperativity in the amyloidogenic variants, and hence their increased tendency to populate a partially unfolded intermediate state and their propensity to convert into amyloid fibrils. These conclusions therefore reinforce the view that the cooperativity of native protein structures is an essential evolutionary development to enable otherwise marginally stable structures to resist aggregation under conditions in which they exert their biological function.[24]

4. Relationship to Other Protein Aggregation Diseases

Generic Aspects of Amyloid Formation and Disease. A chance observation, made just as we had started to explore the mechanism of lysozyme fibril formation, suggested that perfectly ordinary proteins, unrelated to any known disease, can form amyloid fibrils with all the characteristics of those found in lysozyme amyloidosis, or indeed in the 20 or so other related misfolding conditions including Alzheimer's disease and type II diabetes.[25] This finding, which was closely followed by an independent report of a similar observation,[26] stimulated us to begin a series of investigations of the abilities of a range of peptides and proteins to form amyloid fibrils, of the mechanisms by which they do so, and of the properties of the aggregates. These studies led us to the conclusion that many, perhaps in principle nearly all, proteins have the intrinsic ability to convert into amyloid or amyloid-like fibrils in vitro if appropriate conditions can be found.[24,27] Further studies also led to the suggestion that there are likely to be many common structural and mechanistic features underlying protein misfolding and aggregation.[27] This conclusion is consistent with the fact that the fibrillar aggregates associated with all the known amyloid-related diseases appear similar in their overall character and morphology despite little, if any, similarities in the sequences or native structures of the peptides or proteins (whether intact or fragmented) that represent the major components of the pathological deposits.[27]

On the basis of this and other such evidence, we believe that the formation of amyloid fibrils is a consequence of the inherent physicochemical properties of polypeptide chains and that amyloid disorders result fundamentally from the reversion of a peptide or protein in its native state to this alternative and generic form of polypeptide structure.[24,27,28] In general, as in the case of lysozyme, the structure of at least some, and perhaps the majority, of regions of a protein in its fibrillar form will differ substantially from that of the protein in its functional state, whether the latter is globular or natively unfolded. In the case of polypeptides that normally fold to globular structures, specific packing between the side chains and exclusion of many of the hydrophobic groups from water stabilize states in which the main chain is folded into unique and often highly intricate native structures in which the surface regions are designed to avoid inappropriate intermolecular interactions.[29] In these structures, intramolecular hydrogen bonds are formed between the main-chain amide and carbonyl groups in the familiar α-helices, β-sheets, loops, and turns.[29] The overall architecture of a given fold and its net stability relative to that of other globular folds, the denatured state, or indeed the amyloid structure, depend on the specific nature of the side-chain interactions within that fold and so is encoded in the sequence of the protein. Native proteins, even under physiological conditions, are at best only marginally stable relative to these alternatives states.[29]

While globular proteins in their native states are usually highly resistant to aggregation because of their close-packed structures, when such a protein is even partially or transiently unfolded, its propensity to aggregate can be dramatically enhanced. In vitro aggregation, including formation of amyloid fibrils, is therefore often associated with the disruption of the native structure of a protein provided the latter occurs under conditions in which relatively stable noncovalent interactions, particularly hydrogen bonding, can still occur.[24] In vivo, despite the plethora of housekeeping mechanisms such as molecular chaperones and the unfolded protein response, there is a finite chance that any protein molecule produced in a cell may fail to fold correctly or to remain correctly folded and hence is prone to aggregate.[28] In the case of lysozyme, our knowledge of the effects of natural amyloidogenic mutations on the properties of the protein points to the conclusion that the formation of fibrils in vivo results from a series of factors that combine to generate pathogenic behavior. First, the amyloidogenic variants are able to fold sufficiently well to escape degradation by the cellular quality control systems and hence to be secreted in significant amounts into extracellular space where they normally function. Second, the mutations decrease the stability and global cooperativity of the protein sufficiently to enable the variants to unfold partially and populate intermediate states at a significant level. And third, the region of the polypeptide chain that is exposed as the result of this unfolding event is highly aggregation prone and hence can initiate the events that ultimately lead to the formation of amyloid fibrils. One can speculate from these findings that the relative rarity of amyloidoses is likely to reflect not just the efficiency of the biological mechanisms that defend against protein aggregation but also the importance of multiple risk factors of this type for a given protein.

New Paradigms of Protein Structure. Within the amyloid core structure, an array of inter- and intramolecular hydrogen bonds is formed between the main-chain amide and carbonyl groups that are common to all peptides and proteins (except where proline residues are present). Both solid-state NMR and X-ray studies suggest that the side chains of the component amino acid residues can possess a very high degree of order, akin to the situation when peptides and other molecules form three-dimensional crystals.[30,31] The interactions between side chains (and sometimes simply their burial from solvent

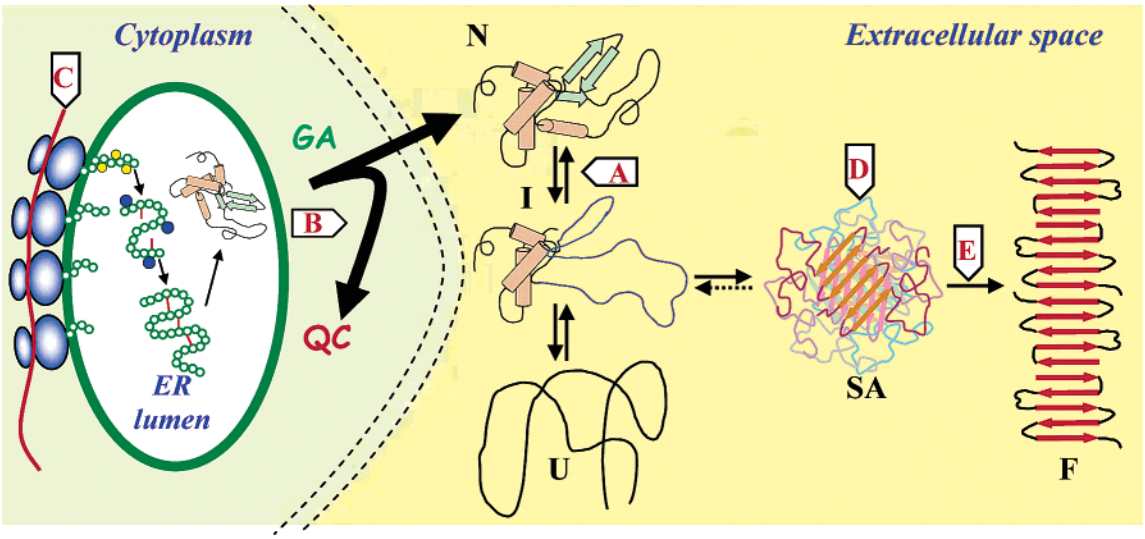

FIGURE 6. Schematic representation of in vivo amyloid fibril formation by lysozyme. After synthesis on the ribosome, the protein folds in the endoplasmic reticulum, aided by molecular chaperones. The correctly folded protein is secreted from the cell and functions normally in its extracellular environment, although it is likely that some misfolded protein is retained within the cell. Under certain circumstances, however, it unfolds, at least partially, and becomes prone to aggregation. This unfolding event can result in the formation of fibrils and other aggregates that accumulate in tissue. Some possible targets for therapeutic intervention are indicated by arrows; they include the following: (A) stabilizing the native state; (B) stimulating clearance of misfolded proteins, for example, by boosting their proteolytic degradation; (C) altering protein synthesis; (D) neutralizing or preventing accumulation of fibril precursors; (E) perturbing fibril assembly. GA, golgi apparatus; QC, quality control; N, native state; I, intermediate state; U, unfolded state; SA, small aggregates; F, fibrils. Adapted from ref 24.

water) will undoubtedly contribute to the stability of the amyloid structure and explains, at least in part, the fact that the propensity to convert to this state varies dramatically with the composition and sequence of any given polypeptide chain.[21] In addition, the repetitive nature and degree of order of the long-range interactions within the highly organized fibril core are important in stabilizing the amyloid state.

Fibrils formed from wild-type human lysozyme can seed the formation of fibrils from solutions of hen lysozyme, a protein with approximately 60% sequence identity.[32] The efficiency of seeding, however, is lower than that carried out with preformed fibrils from proteins having a higher sequence similarity to hen lysozyme including turkey lysozyme or a single-point mutational variant of hen lysozyme.[32] These results show clearly that the side chains of the residues involved in the core of the fibrils, and perhaps also those of the residues that make up the regions peripheral to the fibril core, influence the propensity to form the amyloid structure relative to the many alternative states accessible in principle to a polypeptide chain. Side-chain residues have also been shown to affect the specific details of the fibril structures, for example, the separation between β-sheets and the manner in which protofilaments assemble to form fibrils.[33] However, by contrast to the situation for native states, where the specific interactions of particular side chains is essential to stabilize the overall fold, we believe that in the amyloid form of a peptide or protein, the side chains, whatever their nature, generally accommodate themselves as favorably as possible in a structure whose essential character is determined by common features of the polypeptide chain.[24]

5. Future Perspectives

The diseases associated with protein misfolding are rapidly emerging as among the most feared and debilitating in the modern world. We believe that the rapidly increasing prevalence of disorders such as Alzheimer's disease and type II diabetes results ultimately from the fact that our proteins and the environments in which they function have been selected by evolution to be resistant to aggregation only to the extent necessary for the competitive success of our species.[24,27] Under other conditions, the most vulnerable of our proteins will progressively tend to revert to the generic amyloid structure. In today's society, we strive to go beyond our evolutionary fitness through, for example, extended life spans and radically altered lifestyles.[34] Nevertheless, we believe that studies of the type described here have begun to generate a sufficiently profound understanding of the distinctive differences between normal and aberrant folding behavior and, hence, identified a wide variety of novel strategies through which rational therapeutic intervention could lead to effective prevention or treatment of perhaps the entire family of protein misfolding disorders[28] [Figure 6].

We are grateful to the many students, research fellows, and colleagues who have worked with us on the topics discussed here. Much of this Account is based on research that has been carried out with the generous support of the Leverhulme Trust and the Wellcome Trust.

References

(1) Pepys, M. B.; Hawkins, P. N.; Booth, D. R.; Vigushin, D. M.; Tennent, G. A.; Soutar, A. K.; Totty, N.; Nguyen, O.; Blake, C. C. F.; Terry, C. J.; Feest, T. G.; Zalin, A. M.; Hsuan, J. J. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* **1993**, *362*, 553−557.

(2) Valleix, S.; Drunat, S.; Philit, J. B.; Adoue, D.; Piette, J. C.; Droz, D.; MacGregor, B.; Canet, D.; Delpech, M.; Grateau, G. Hereditary renal amyloidosis caused by a new variant lysozyme W64R in a French family. *Kidney Int.* **2002**, *61*, 907−912.
(3) Yazaki, M.; Farrell, S. A.; Benson, M. D. A novel lysozyme mutation Phe57Ile associated with hereditary renal amyloidosis. *Kidney Int.* **2003**, *63*, 1652−1657.
(4) Rocken, C.; Becker, K.; Fandrich, M.; Schroeckh, V.; Stix, B.; Rath, T.; Kahne, T.; Dierkes, J.; Roessner, A.; Albert, F. W. ALys amyloidosis caused by compound heterozygosity in Exon 2 (Thr70Asn) and Exon 4 (Trp112Arg) of the lysozyme gene. *Hum. Mutat.* **2005**, *27*, 119−120.
(5) Booth, D. R.; Pepys, M. B.; Hawkins, P. N. A novel variant of human lysozyme (T70N) is common in the normal population. *Hum. Mutat.* **2000**, *16*, 180.
(6) Dumoulin, M.; Bellotti, V.; Dobson, C. M. Hereditary systemic amyloidosis associated with mutational variants of human lysozyme. In *The beta pleated sheet conformation and disease*; Sipe, J. D., Ed.; VCH Verlag GmbH & Co KgaA publisher: Weinheim, Germany, 2005; pp 635−656.
(7) Dumoulin, M.; Johnson, R. J. K.; Belotti, V.; Dobson, C. M. Human lysozyme. In *Protein misfolding, aggregation and conformational diseases. II. Molecular basis of conformational diseases*; Fink, A. L., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2006; Vol. II; in press.
(8) Booth, D. R.; Sunde, M.; Bellotti, V.; Robinson, C. V.; Hutchinson, W. L.; Fraser, P. E.; Hawkins, P. N.; Dobson, C. M.; Radford, S. E.; Blake, C. C. F.; Pepys, M. B. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **1997**, *385*, 787−793.
(9) Johnson, J. R. K.; Christodoulou, J.; Dumoulin, M.; Caddy, G.; Alcocer, M. J.; Murtagh, G. J.; Kumita, J. R.; Larsson, G.; Robinson, C. V.; Archer, D. B.; Luisi, B.; Dobson, C. M. Rationalising lysozyme amyloidosis: Insight from the structure and solution dynamics of T70N lysozyme. *J. Mol. Biol.* **2005**, *352*, 823−836.
(10) Esposito, G.; Garcia, J.; Mangione, P.; Giorgetti, S.; Corazza, A.; Viglino, P.; Chiti, F.; Andreola, A.; Dumy, P.; Booth, D.; Hawkins, P. N.; Bellotti, V. Structural and folding dynamics properties of T70N variant of human lysozyme. *J. Biol. Chem.* **2003**, *278*, 25910−25918.
(11) Wain, R.; Smith, L. J.; Dobson, C. M. Oxidative refolding of amyloidogenic variants of human lysozyme. *J. Mol. Biol.* **2005**, *351*, 662−671.
(12) Canet, D.; Sunde, M.; Last, A. M.; Miranker, A.; Spencer, A.; Robinson, C. V.; Dobson, C. M. Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in its disease-related variants. *Biochemistry* **1999**, *38*, 6419−6427.
(13) Hooke, S. D.; Radford, S. E.; Dobson, C. M. The refolding of human lysozyme: a comparison with the structurally homologous hen lysozyme. *Biochemistry* **1994**, *33*, 5867−5876.
(14) Dobson, C. M.; Evans, P. A.; Radford, S. E. Understanding how proteins fold: the lysozyme story so far. *Trends Biochem. Sci.* **1994**, *19*, 31−37.
(15) Dinner, A. R.; Sali, A.; Smith, L. J.; Dobson, C. M.; Karplus, M. Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* **2000**, *25*, 331−339.
(16) Dumoulin, M.; Canet, D.; Last, A. M.; Pardon, E.; Archer, D. B.; Muyldermans, S.; Wyns, L.; Matagne, A.; Robinson, C. V.; Redfield, C.; Dobson, C. M. Reduced global cooperativity is a common feature underlying the amyloidogenicity of pathogenic lysozyme mutations. *J. Mol. Biol.* **2005**, *346*, 773−788.
(17) Kumita, J. R.; Johnson, J. R. K.; Alcocer, M. J.; Dumoulin, M.; Holinquist, F.; McCammon, M. G.; Robinson, C. V.; Archer, D. B.; Dobson, C. M. Expression of the amyloidogenic human lysozyme variants F57I and W64R in Pichia pastoris: Impact of native state stability on protein secretion levels. *FEBS J.* **2006**, *273*, 711−720.
(18) Morozova-Roche, L. A.; Zurdo, J.; Spencer, A.; Noppe, W.; Receveur, V.; Archer, D. B.; Joniau, M.; Dobson, C. M. Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J. Struct. Biol.* **2000**, *130*, 339−351.
(19) Canet, D.; Last, A. M.; Tito, P.; Sunde, M.; Spencer, A.; Archer, D. B.; Redfield, C.; Robinson, C. V.; Dobson, C. M. Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. *Nat. Struct. Biol.* **2002**, *9*, 308−315.
(20) Frare, E.; Polverino de Laureto, P.; Zurdo, J.; Dobson, C. M.; Fontana, A. A highly amyloidogenic region of hen lysozyme. *J. Mol. Biol.* **2004**, *340*, 1153−1165.
(21) Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **2005**, *350*, 379−392.
(22) Dumoulin, M.; Conrath, K.; Van Meirhaeghe, A.; Meersman, F.; Heremans, K.; Frenken, L. G.; Muyldermans, S.; Wyns, L.; Matagne, A. Single-domain antibody fragments with high conformational stability. *Protein Sci.* **2002**, *11*, 500−515.
(23) Dumoulin, M.; Last, A. M.; Desmyter, A.; Decanniere, K.; Canet, D.; Larsson, G.; Spencer, A.; Archer, D. B.; Sasse, J.; Muyldermans, S.; Wyns, L.; Redfield, C.; Matagne, A.; Robinson, C. V.; Dobson, C. M. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature* **2003**, *424*, 783−788.
(24) Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **1999**, *24*, 329−332.
(25) Guijarro, J. I.; Sunde, M.; Jones, J. A.; Campbell, I. D.; Dobson, C. M. Amyloid fibril formation by an SH3 domain. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4224−4228.
(26) Litvinovich, S. V.; Brew, S. A.; Aota, S.; Akiyama, S. K.; Haudenschild, C.; Ingham, K. C. Formation of amyloid-like fibrils by self-association of a partially unfolded fibronectin type III module. *J. Mol. Biol.* **1998**, *280*, 245−258.
(27) Stefani, M.; Dobson, C. M. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **2003**, *81*, 678−699.
(28) Dobson, C. M. Principles of protein folding, misfolding and aggregation. *Semin. Cell Dev. Biol.* **2004**, *15*, 3−16.
(29) Branden, C.; Tooze, J. *Introduction to protein structure*, 2nd ed.; Garland Publishing: New York, 1999.
(30) Jaroniec, C. P.; MacPhee, C. E.; Bajaj, V. S.; McMahon, M. T.; Dobson, C. M.; Griffin, R. G. High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 711−716.
(31) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005**, *435*, 773−778.
(32) Krebs, M. R.; Morozova-Roche, L. A.; Daniel, K.; Robinson, C. V.; Dobson, C. M. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci.* **2004**, *13*, 1933−1938.
(33) Fandrich, M.; Dobson, C. M. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* **2002**, *21*, 5682−5690.
(34) Dobson, C. M. Getting out of shape. *Nature* **2002**, *418*, 729−730.
(35) Gillmore, J. D.; Booth, D. R.; Madhoo, S.; Pepys, M. B.; Hawkins, P. N. Hereditary renal amyloidosis associated with variant lysozyme in a large English family. *Nephrol., Dial., Transplant.* **1999**, *14*, 2639−2644.
(36) Vendruscolo, M.; Paci, E.; Karplus, M.; Dobson, C. M. Structures and relative free energies of partially folded states of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14817−14821.
(37) Dumoulin, M.; Dobson, C. M. Probing the origins, diagnosis and treatment of amyloid diseases using antibodies. *Biochimie* **2004**, *86*, 589−600.

AR050070G