BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

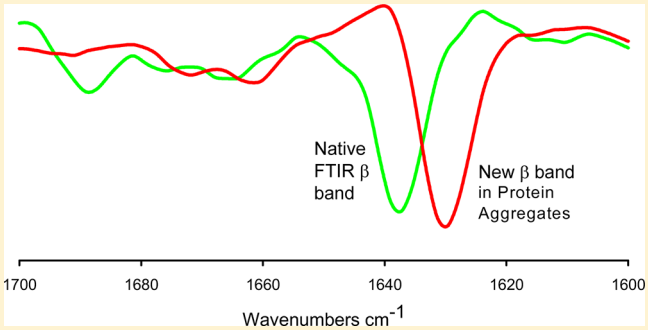
pubs.acs.org/biochemistry

Distinct β-Sheet Structure in Protein Aggregates Determined by ATR−FTIR Spectroscopy

Bhavana Shivu,*[,†] Sangita Seshadri,[‡] Jie Li,[§] Keith A. Oberg,[∥] Vladimir N. Uversky,[⊥] and Anthony L. Fink

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, United States

ABSTRACT: Attenuated total reflectance Fourier transform infrared spectroscopy (ATR−FTIR) was used to study the conformation of aggregated proteins in vivo and in vitro. Several different protein aggregates, including amyloid fibrils from several peptides and polypeptides, inclusion bodies, folding aggregates, soluble oligomers, and protein extracts from stressed cells, were examined in this study. All protein aggregates demonstrate a characteristic new β structure with lower-frequency band positions. All protein aggregates acquire this new β band following the aggregation process involving intermolecular interactions. The β sheets in some proteins arise from regions of the polypeptide that are helical or non β in the native conformation. For a given protein, all types of the aggregates (e.g., inclusion bodies, folding aggregates, and thermal aggregates) showed similar spectra, indicating that they arose from a common partially folded species. All of the aggregates have some nativelike secondary structure and nonperiodic structure as well as the specific new β structure. The new β could be most likely attributed to stronger hydrogen bonds in the intermolecular β-sheet structure present in the protein aggregates.



INTRODUCTION

The critical importance of protein aggregation in a wide variety of situations, including protein-deposition diseases, protein-folding studies, the stability of protein drugs, and the industrial preparation of proteins, is now widely recognized. However, little is understood about the details of the structural changes underlying the process. One of the relatively few techniques that can be used to determine the conformation of aggregated protein is attenuated total reflectance Fourier transform infrared spectroscopy (ATR−FTIR) in which a sample is dried on the surface of a high-refractive-index material, such as a germanium or zinc selenide crystal (an internal reflectance element, IRE).[1−4] Thin-film ATR−FTIR is very sensitive compared to normal transmission FTIR because the major water peak is not present when the hydrated thin-film sample is analyzed, which gives a highly sensitive method to analyze protein secondary structure in water. In transmission mode, the water peak is usually suppressed by exchanging the sample with D_2O. More recently, a group has used thin films in the transmission mode to study the insoluble protein extracts of ataxin[5] and hence overcame the use of D_2O in the transmission mode. The advantage of thin-film ATR−FTIR/FTIR is that because hydrogen bonds strongly contribute to IR bands, it is important to have hydrogen instead of D_2O for the collection and analysis of IR data. By analysis of the amide I band in the 1600−1700 cm^{-1} region (predominantly corresponding to the amide C═O stretch), it is possible to determine the secondary-structure composition of the sample.[6] Fourier transform infrared spectroscopy has been used to study the secondary structure of insoluble protein aggregates such as inclusion bodies,[4,7] heat-gelled proteins,[8] and amyloid fibrils.[9]

Amyloid fibrils are the most commonly known in vivo protein aggregates in several human diseases. Amyloid deposits are observed to be extracellular in Alzheimer's diseases in the brain and in several systemic amyloid diseases, intracellular in Lewy bodies in Parkinson's disease, and intranuclear in Huntington's disease. Amyloid fibrils from various ex vivo and in vitro sources have been demonstrated to have a crossed-β structure by X-ray fiber diffraction.[10] A crossed-β structure is described as a β sheet where the β strands are perpendicular to the fiber axis. Other in vivo protein aggregates are inclusion bodies that are formed in bacterial cells resulting from the overexpression of proteins. Protein structural changes were also monitored in cultured mammalian cells before and after stress.

Under suitable in vitro conditions, most proteins acquire native structure, but altered conditions lead to alternate conformations that are either monomeric partially folded, soluble oligomers, or insoluble aggregates. London et al. first suggested 25 years ago that partially folded intermediates might be responsible for aggregation during protein folding,[11] and increasing evidence to support this hypothesis is accumulating.[12−20] In fact, for the same protein, two different partially folded intermediates were shown to give rise to either amorphous aggregates or amyloid fibrils.[21] Knowledge of the structural basis of protein interactions within aggregates may provide critical insight to the underlying aggregation mechanism and therefore suggest ways of preventing protein aggregation.

Received: May 17, 2013
Revised: July 9, 2013
Published: July 9, 2013

   dx.doi.org/10.1021/bi400625v | *Biochemistry* 2013, 52, 5176−5183

MATERIALS AND METHODS

Protein Aggregates. Amyloid fibrils from Aβ peptides were a gift from Prof. D. Teplow. Insulin and Ig light-chain fibrils were prepared from the purified proteins by incubating 0.5 mg/mL of protein in 20−100 mM HCl with a 100 mM NaCl solution in glass vials with stirring using Teflon-coated magnets at 37 °C for 6−24 h. α-Synuclein fibrils were prepared by stirring a 1 mg/mL solution at pH 7.5 in Tris buffer with 100 mM NaCl.

Thermal Aggregates. Thermal aggregates of apomyoglobin were generated by dissolving 4.8 mg of apomyoglobin in 480 μL of 20 mM phosphate buffer pH 7.4 and heating at 75 °C in a water bath for 15 min. The white precipitate was pelleted and applied to the IRE. Thermal gels of BSA were formed by heating solutions at pH 7 for several hours at 60−70 °C.

Stress-Induced Cellular Aggregates. Stress-induced aggregates in mammalian cells were prepared from heat-shocked HeLa cells at a density of $(2-4 \times 10^5)$/ml by heating in a water bath at 43 °C for 45 min. This was followed by the addition of fresh medium at 37 °C into the culture, and the cells were left at 37 °C for varying periods of time prior to ATR−FTIR analysis.

Inclusion Bodies and Folding Aggregates. Inclusion bodies were isolated from overexpression systems in *E. coli* using low-speed centrifugation followed by washing (initially with detergent or urea and then with water). SDS-PAGE indicated that the inclusion bodies were relatively homogeneous, usually >90%. Folding aggregates were prepared by 10-fold dilution of the unfolded protein at a high concentration in 6 M GdnHCl into native buffer to form a precipitate of the folding aggregate. Soluble aggregates of barstar were formed at pH <3 in 20 mM glycine buffer. Partially folded intermediates of apomyoglobin were prepared as described previously.[1,22,23]

ATR−FTIR Spectroscopy and Data Analysis. Hydrated (H_2O) thin-film spectra were collected using a Nicolet 800 FTIR spectrometer equipped with an MCT detector that was purged with dry air. All samples were scanned in an out-of-compartment horizontal ATR accessory (SPECAC) with a high-throughput 73 × 10 × 6 mm^3, 45° trapezoidal germanium crystal (IRE). To collect spectra, the samples were applied (50 μL of a 1 to 2 mg/mL solution) and dried while being spread constantly with a spatula in a stream of dry nitrogen. Water-vapor spectra were collected using a clean IRE with a reduced purge rate such that the absorbance of the highest water-vapor peak in the amide I and II regions was between 0.03 and 0.05. The water-vapor spectrum was subtracted from the sample spectrum until the region between 1700 to 1800 wavenumbers was featureless. Data processing was carried out using GRAMS32 (Galactic Industries) with additional BioPsi software (developed by K. A. Oberg) as previously described.[3,24,25] Raw spectra were further processed to enhance the component bands by second-derivative analysis using the Savitsky−Golay routine with no more than nine-point smoothing, as shown in Figures 3A−E and 4. Raw spectra after water-vapor subtraction are shown in Figure 1. Secondary-structure analysis was performed using curve-fit spectra. Curve fitting was performed by detailed spectral analysis first using second-derivative and Fourier self-deconvolution to determine the peak positions of the components of the raw spectrum followed by curve fitting and determination of the area under each component peak, which was then used to calculate the percentage of each band, as described previously.[1,4,6,24,26] The fraction of nativelike structure was determined from the curve fit spectra. From a comparison of the spectra of the native and aggregated species for each common component, the fraction of the total structure was converted to a ratio and summed.

Secondary-Structure Assignments. From several studies over the past several years, a consensus has emerged for the assignment of many IR components in the amide I region.[4,27−29] In general, bands (in H_2O) in the 1643−1615 and 1692−1697 cm^{-1} regions are ascribed to a β-sheet/extended conformation, 1647−1654 cm^{-1}, to disordered, 1651−1663 cm^{-1}, to loops, 1653−1660 cm^{-1}, to α helix, and 1663−1695 cm^{-1}, to turns. Curve-fitting analysis of thin-film ATR amide I data has been shown to provide accurate secondary-structure analysis for proteins,[1,2] indicating that the procedure does not have an adverse effect on the secondary structure of most proteins.

RESULTS AND DISCUSSION

We have included various in vivo and in vitro protein aggregates in our study in an attempt to understand their structural features. In vivo aggregates include amyloid fibrils, inclusion bodies, other intra- and extracellular inclusions and deposits (e.g., similar to those found in Lewy body diseases and light-chain deposition disease), and insoluble aggregated protein accumulated in mammalian cells upon stress. In vitro aggregates included in our study are refolding aggregates (insoluble associated states of proteins formed during in vitro protein folding or under weakly native conditions, especially at higher protein concentrations), thermal aggregates (formed at temperatures just below the beginning of the thermally unfolded transition, particularly at higher protein concentrations), gels (formed after heating at high temperatures for some proteins), and the soluble and insoluble aggregates of partially folded intermediates such as molten globules. We have specifically excluded the consideration of alcohol-induced aggregates from this investigation because alcohols are well known to induce helical structure in polypeptides.

Amyloid Fibrils. X-ray fiber-diffraction experiments have demonstrated that amyloid fibrils are rich in β-sheet structure, with a core of cross-β structure.[10,30,31] FTIR spectra of a number of proteins in fibrillar aggregates, such as the Alzheimer's Aβ 1−42 peptide,[32] insulin, and the immunoglobulin light-chain variable domain SMA,[33] confirm that they are all rich in β structure (Figure 1A, Table 1). All of the protein fibril spectra studied have β-band positions around 1630 cm^{-1} (1632 for insulin, 1630 for SMA, and 1628 for Aβ and α-synuclein) (Figure 1A). Zandomeneghi and co-workers have demonstrated using transmission FTIR that transthyretin amyloid fibrils have a β band that has lower wavenumbers compared to its native β band, but the peak positions are different because of the use of D_2O in their study.[34] In addition to the polypeptide amyloid fibrils, we have also included two 13-residue peptide fibrils (Table 1 and Figure 1B): synthetic peptide KLKLKLELELELG (KLEG)[35] and peptide EDVAVYYCHQYYS (ED), which is a fragment of the immunoglobulin light-chain SMA. These fibrils are also rich in β structure as demonstrated by the observed major β bands, but the position of the corresponding amide I β-sheet components varies significantly from a low of around 1623 cm^{-1} for the KLEG peptide to 1634 cm^{-1} for the ED peptide. Furthermore, with the exception of the peptide KLEG, all of the spectra for the other fibrils indicated the presence of significant non-β

dx.doi.org/10.1021/bi400625v | *Biochemistry* 2013, 52, 5176−5183

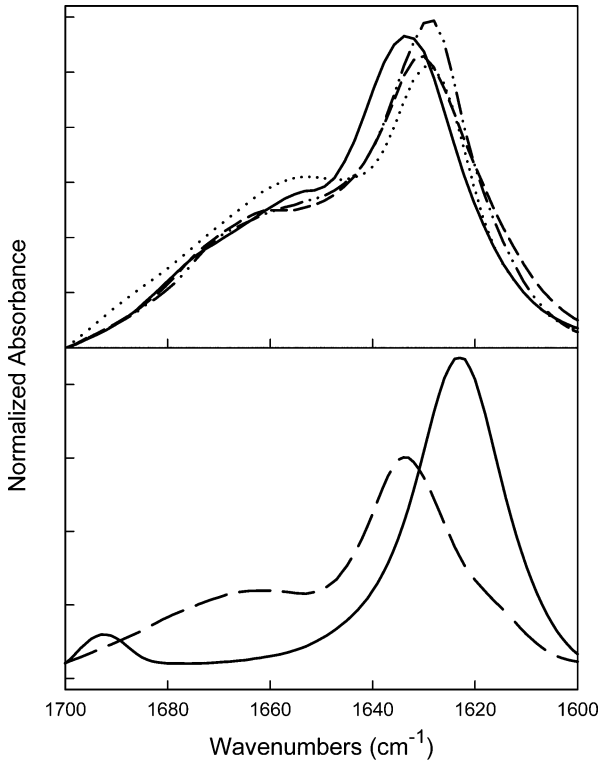


Figure 1. ATR−FTIR spectra of the amide I region of amyloid fibrils formed in vitro. The top panel shows SMA (dashed line, maximum at 1630 cm^{-1}), insulin (solid line, maximum at 1632 cm^{-1}), Aβ (1−40) (Alzheimer's disease β peptide) (dotted line, maximum at 1628 cm^{-1}), α-synuclein (dashed-dotted line, maximum at 1628 cm^{-1}). The bottom panel shows peptide fibrils from a 13 aa synthetic peptide designed to have β-helical structure and a 13 aa peptide fragment (ED) from SMA. KLEG(29) (solid line) shows the lowest-wavenumber β band among all the fibrils (1623 cm^{-1}), and ED (dashed line) shows the highest-wavenumber β band (1634 cm^{-1}).

Table 1. Amide I Band Positions for Amyloid Fibril Spectra Obtained from Proteins and Peptides[a]

	fibrils	
proteins and peptides	non-β bands	β bands
SMA	1660 (15)	1630 (49)
	1647 (14)	
α-synuclein	1656 (46)	1628 (52)
insulin	1656 (11)	1633 (70)
Aβ (1−42)	1654 (32)	1628 (37)
ED (EDVAVYYCHQYYS)	1665 (10)	1634 (58)
	1655 (11)	
KLEG (KLKLKLELELELG)		1624 (94)
		1692 (6)

[a]The numbers in parentheses correspond to the percent area of the peak.

structure shown by bands at 1670−1685 cm^{-1} (especially turns and loops, Table 1), indicating that in addition to the core of the β sheet, other secondary structure is also present in the fibrils (Figure 1).

The generally wider peak widths for the amorphous aggregates compared to the corresponding fibrils indicate that there is a broader distribution of conformations in the amorphous aggregates (Figure 2). Curve-fit analysis of the raw spectra for native, fibril, and amorphous forms of the light-chain SMA showed significant differences (Figure 2). Compared to the native state, the fibrils showed a small increase in β structure (about 8%) along with a shift in the position of the β component from 1637 to 1630 cm^{-1}, which is consistent with stronger hydrogen bonds. The amorphous aggregate, however, does not show an increase in overall β content but shows a new low-frequency β band at 1628 cm^{-1}. Interestingly, we have shown that the amorphous aggregate and amyloid fibrils arise from a nativelike and a relatively unfolded-like partially folded conformation, respectively.[21] Again, this confirms that partially folded conformations are precursors of the aggregates.

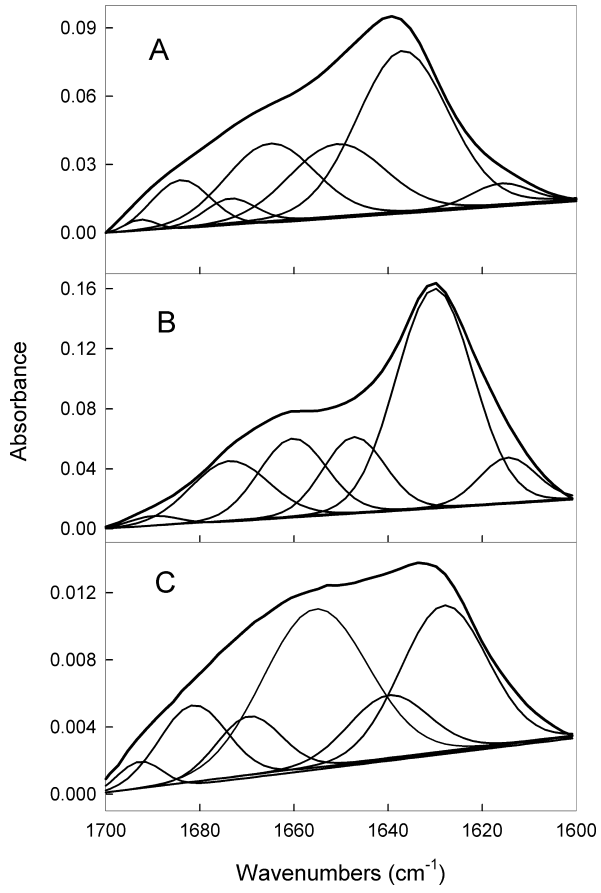


Figure 2. Raw ATR−FTIR spectra and their component peaks for the native conformation of SMA (A), amyloid aggregate of SMA (B), and amorphous aggregate of SMA (C).

Precipitates of Native Protein. To confirm that precipitation per se does not affect the conformation of a protein, we examined the spectra of native IL-2, leptin, and apomyoglobin that had been salted-out of solution by ammonium sulfate. The amide I spectra of these precipitates were indistinguishable, within experimental error, from those of the native proteins (in solution) and showed no increase in absorbance in the β-structure region (Figure 3A).

Aggregation of Partially Folded Intermediates. To substantiate more directly the hypothesis that the key precursors of protein aggregation are partially folded conformations, which may arise either from the native or unfolded state, we determined whether the aggregation of monomeric partially folded intermediates correlates with an increase in β sheet. Apomyoglobin is an all-α protein[22] whose acidic partially folded intermediates (A-states) are also rich in α

dx.doi.org/10.1021/bi400625v | Biochemistry 2013, 52, 5176−5183

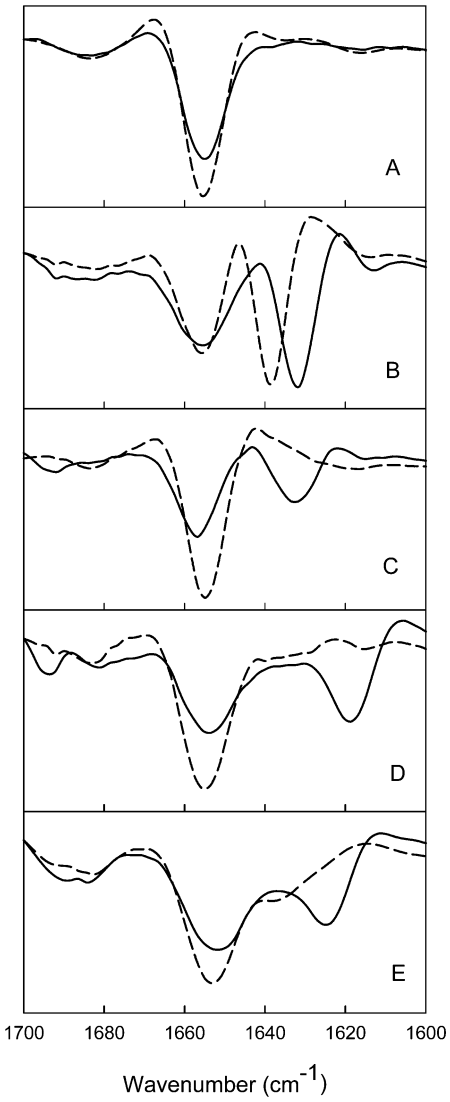


Figure 3. Second-derivative ATR−FTIR spectra of the amide I region of various aggregated proteins and a native control. The spectra of the native states are shown as dashed lines. (A) Spectrum of ammonium sulfate-precipitated IL-2 (solid line) is essentially identical to that of the native protein in solution, indicating that precipitation per se does not result in the presence of β structure. (B) Soluble, associated altered conformation of barstar shows the presence of new β-sheet structure (solid line). (C) Association of an apomyoglobin (ApoMb) partially folded intermediate (A state) from its monomeric state (dashed line) leads to increased β structure (solid line). (D) Heat-gelled bovine serum albumin (solid line) formed by prolonged thermal denaturation. (E) Spectra of isolated, washed Triton X-100-insoluble material from heat-shocked HeLa cells shows increased β structure (solid line) compared to that of the non-heat-shocked control (dashed line).

helices.[23,24] At low protein concentrations, these intermediates are soluble and monomeric; however, at higher protein concentrations they associate, initially forming soluble aggregates that eventually become insoluble precipitates. Figure 3C shows a spectrum of a native and monomeric partially folded intermediate (dashed line) and soluble/insoluble oligomers of an apomyoglobin intermediate (solid line), suggesting that the monomeric partially folded intermediates of apomyoglobin contain no β structure, yet on association a substantial amount of β structure appears. This indicates that the new β structure upon aggregation arises from the initial association of partially folded intermediate conformations, leading to intermolecular β-strand interactions that are maintained in the subsequent insoluble aggregates.

Similarly, an acidic non-native conformation of barstar (an α/β protein), which exists as a soluble oligomer (16 mer),[36,37] also demonstrates a shift in its β band to lower frequencies compared to the native β band, whereas the remaining bands are surprisingly native like (Figure 3B, dashed line is native and the solid line is for the acidic soluble oligomer). Interestingly, these observations suggest that for barstar aggregates the lower-frequency β band (new β) arises from the parts of the protein that are β in the native conformation.

Heat-gelled proteins have frequently been used as models of intermolecularly associated proteins and typically show a strong IR band in the 1620 cm^{-1} region arising from intermolecular β structure.[26] A heat-gelled BSA spectrum (Figure 3D, solid line) is compared to its native spectrum (Figure 3D, dashed line). The position of the β band in the heat gel is 1620 cm^{-1}, which is the lowest value observed among all of the aggregates studied. This might help us to understand the basis of the shift toward lower frequencies of the β band position upon aggregation.

Stress-Induced Insoluble Protein in Cells. When cells are subjected to stress, such as heat shock or ATP depletion, one of the many consequences is the formation of insoluble protein (including the cytosolic Hsp70s), especially in the nucleus and nucleolus.[27−29] ATR−FTIR spectra of the insoluble protein from HeLa cells subjected to heat shock reveal a substantial and reproducible appearance of the new β band compared to that of the unstressed controls (Figure 3E).

Inclusion Bodies and Folding Aggregates. Inclusion bodies are amorphous protein aggregates commonly found in overexpressing cells and are generally thought to arise from the aggregation of unfolded protein. They are typically rather homogeneous and usually can be dissolved by treatment with high concentrations of denaturant. We examined the ATR−FTIR spectra of the inclusion bodies of several proteins. The second-derivative spectra of inclusion bodies from representative all-α (apomyoglobin and interferon-γ, Figure 4A,B), all-β (interleukin-1β, Figure 4C), and α-β proteins (haptoglobin, Figure 4D) are compared to the corresponding spectra of the native and folding aggregates in Figure 4. The spectra of the folding aggregates are the same, within experimental error, as those of the corresponding inclusion bodies (Figure 4). This suggests that both inclusion bodies and folding aggregates are formed from the same partially folded species. In addition, folding aggregates from proteins that did not form inclusion bodies were examined and also exhibited low-wavenumber β bands (e.g., citrate synthase, data not shown). The most notable changes observed by comparing the spectra of the native conformations in the inclusion bodies and folding aggregates are the appearance of new lower-wavenumber β bands (new β) and the large amount of nativelike secondary structure present (determined by detailed spectral analysis by first using second-derivative and Fourier self-deconvolution to determine the components of the raw spectrum followed by curve fitting to determine the percentage of each peak). Secondary-structure analysis of all of the inclusion bodies revealed that there is typically no more than 25% nonperiodic (disordered) structure in the inclusion bodies (which is comparable to that present in the native states), 20−25% new β structure appears, and 50−70% of the nativelike

dx.doi.org/10.1021/bi400625v | *Biochemistry* 2013, 52, 5176−5183

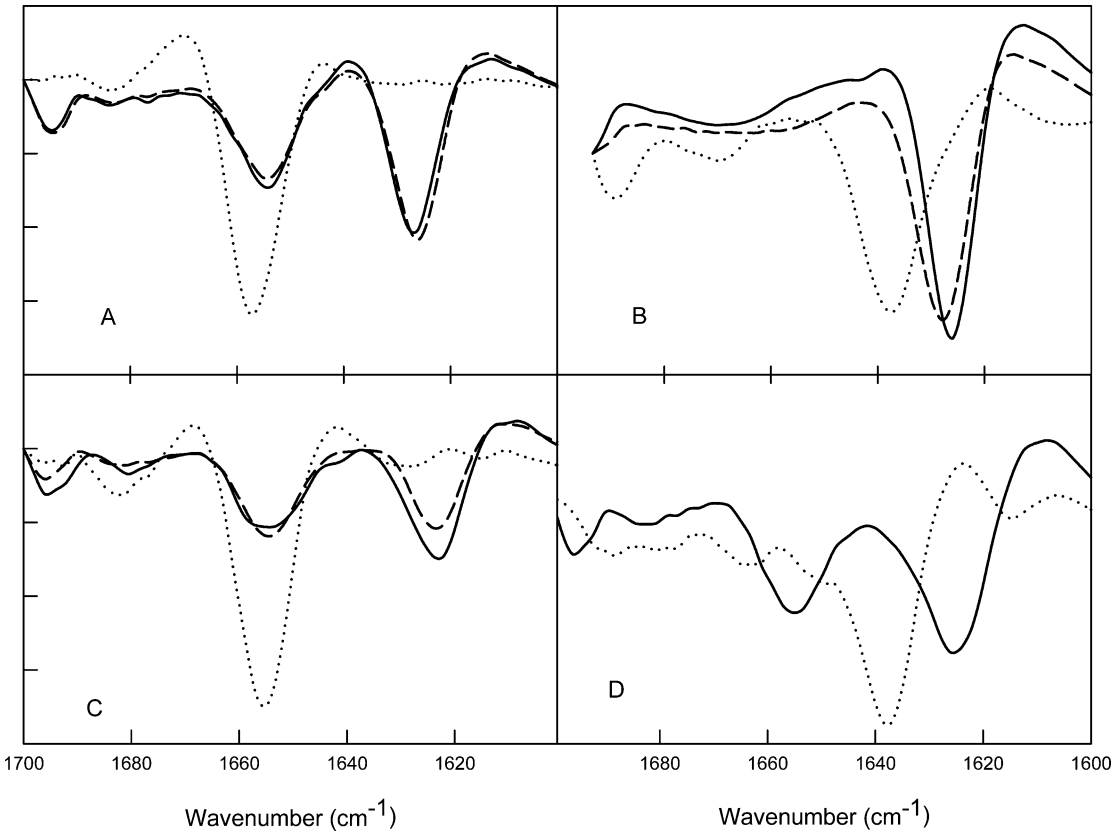


Figure 4. Second-derivative ATR−FTIR spectra of the amide I region of samples of native (dotted line), inclusion bodies (solid line), and folding aggregates (dashed line) showing the increased β structure (1623−1633 cm^{-1}). Apomyoglobin and interferon-γ (INF-γ) are all-helical proteins, interleukin-1β (IL-1β) is an all-β protein, and haptoglobin is a mixed α/β protein.

secondary structure is present. This indicates that the inclusion bodies are formed from a partially folded intermediate with substantial nativelike structure already present that is also responsible for the formation of folding aggregates. The fact that new β structure only appears upon aggregate or inclusion-body formation suggests that the intermolecular interactions involve β structure.

Aggregated protein may also be obtained by taking solutions of native protein at a relatively high concentration and subjecting them to marginally destabilizing conditions, such as raising the temperature close to the beginning of the thermal denaturation transition or by using low-to-moderate concentrations of chemical denaturant. The examination of aggregates formed in this manner yielded FTIR spectra (data not shown) that matched those of the corresponding refolding aggregates and inclusion bodies. This suggests that such aggregates are formed from similar intermediates to those responsible for the folding aggregates and inclusion bodies.

Because the secondary structure of the partially folded intermediates of apomyoglobin is predominantly α helix with some turn and disordered structure, the β strands must arise from regions that are in a non-β conformation in both the aggregation-competent precursor and the native state. This is also apparent from the data in Figure 4 and for proteins such as IL-2 and interferon-γ that are also all-helical proteins in the native state.

Implications for Protein Misfolding. Because the spectra of inclusion bodies, folding aggregates, and aggregates formed under destabilizing conditions for a given protein are so similar, we infer that the underlying structures in these different aggregates are the same. This indicates that a given protein will have at least one partially folded conformation that is especially stable or long-lived and that is particularly prone to aggregation. Consequently, most (or all) aggregates of that protein will arise from that conformation and thus have similar structures. The data for the folding aggregates and inclusion bodies indicate that this intermediate is present during normal folding both in vivo and in vitro.

This aggregation-competent intermediate conformation is likely to have a core of nativelike secondary structure with the remainder being relatively disordered: aggregation interactions may involve both the core and/or the disordered regions and will preserve the core. Furthermore, in all cases, even for all-β proteins,[38] significant new-β structure compared to that in the native conformation was observed. This indicates that the intermolecular interactions leading to the aggregation involve β-sheet-like interactions, which are most probably from multiple β strands (e.g., there could be three additional β strands in three different locations that form three β ladders, leading to propagation in three dimensions). A few previous reports on individual proteins have also revealed increased β structure on protein aggregation.[39−41] A recent report[42] has studied five different aggregates of the same protein, including fibrils, inclusion bodies, heat-aggregated, concentration-dependent, and TCA-precipitated forms using FTIR. This group found a new β band at lower wavenumbers for all aggregates compared to the soluble protein, and the fibrils had the lowest band at 1624 cm^{-1}.

The structures of the amorphous aggregates studied here differ from heat-gelled proteins (Figure 3D), which typically

dx.doi.org/10.1021/bi400625v | *Biochemistry* 2013, 52, 5176−5183

show a strong IR band in the 1620 cm^{-1} region;[26,40,43] this likely arises from long contiguous stretches of intermolecular β-sheet interactions resulting from the association of a significantly unfolded polypeptide. The low-frequency bands in heat-gelled proteins probably reflect very strong H bonding corresponding to a more perfect β-sheet structure, arising from fewer tertiary-structure constraints on the β strands in the intermolecular interactions compared to those in most native globular proteins where the topological and packing constraints of the rest of the native structure frequently lead to distortions, such as β bulges and bifurcated H bonds in the strands.

The β structure observed in the amorphous aggregates was at lower frequencies compared to the β bands in the native conformation. In native globular proteins, β-component bands below 1630 cm^{-1} are quite rare. Because the major factor determining the position of the amide I carbonyl-stretch frequency in the IR is the strength of the H bond, the lower frequencies of β bands are frequently indicative of stronger hydrogen bonds (the other factors are the Φ and Ψ angles and transition-dipole coupling). This suggests that the intermolecular interactions leading to aggregation involve β strands with stronger H bonds than those typically found in globular proteins. Because the typical amide I β-component band of the amorphous aggregates is between 1624 and 1630 cm^{-1} and because in native proteins such bands are usually at higher frequencies, we conclude that the nativelike structure in the aggregation-competent intermediates imposes some constraints on the β ladders that may form, but these are much less than those in the native state. Interestingly, the amyloid fibrils from the proteins had a corresponding value between 1628 and 1632 cm^{-1}, presumably reflecting tertiary constraints on the β strands within the regular quaternary structure of the fibrils (such as the twist in the β sheet and the tight side-chain packing in the core of the fibril) that are absent in amorphous aggregates.

Model for Protein Aggregation and Misfolding. The following general model for protein aggregation is proposed based on the hypothesis that specific intermolecular interactions between the hydrophobic patches of structural subunits in partially folded intermediates are responsible for protein aggregation, including inclusion bodies and amyloid. A key feature is that in aggregation/misfolding the normal process of folding is interrupted. Given the key position of the critical partially folded intermediate on the pathway of aggregation, conditions that lead to an increased concentration of this aggregation-competent intermediate will be crucial in pathological protein aggregation.

Salient features of the model include the following. (a) Protein folding involves the formation of the native state by the sequential interaction of substructural units, the building blocks, which may be metastable on their own but are significantly stabilized by interactions with other such building blocks.[24,44] (b) Although the major driving force for the interaction of these structural subunits are hydrophobic interactions, electrostatic interactions (including hydrogen bonding) may be important for providing additional specificity as well as additional energetic stabilization. (c) The formation of the native state involves the intramolecular interaction of the hydrophobic faces of the structural subunits. Specificity will arise from a variety of features, of which the geometric shape, the extent of the hydrophobic patches, the topological constraints of the polypeptide-chain, and the presence of other structural subunits are probably the most important. (d) Aggregation occurs when these hydrophobic surfaces interact in an intermolecular manner and lead to the conversion of nearby disordered or extended polypeptide into an intermolecular β sheet. Three-dimensional propagation of this process leads to large aggregates. Initially, the aggregates will be soluble, but eventually their size will exceed the solubility limit. The fact that they may still have significant solvent-exposed hydrophobic surfaces would also minimize their solubility.

Any viable model for aggregation must be able to account for both amorphous and fibrillar aggregates. We have observed that for protein systems that form both types of aggregates amorphous aggregates are usually formed more rapidly than fibrils. We believe that this is a consequence of amorphous aggregation being under kinetic control, whereas fibril formation is under thermodynamic control, reflecting lower free-energy barriers in the former and deeper free-energy minima in the latter (along with higher barriers).

We propose the following model to account for amorphous and fibrillar aggregation (Figure 5). In the amorphous pathway, the interactions between hydrophobic surface patches on the partially folded intermediates would be very fast and could bring disordered regions of the two interacting molecules into close contact where they could form intermolecular β ladders. Thus, these initial self-associated states would have no long-range order and would be held together by a combination of both hydrophobic and hydrogen-bonding interactions, leading to amorphous aggregates. We have observed a large variation in the size of amorphous aggregates, again indicating the lack of specific interactions. These interactions are visualized as being reversible because of the less than optimal (from an free energy point of view) intermolecular interactions resulting from topological constraints. In certain cases, domain swapping in nativelike intermediates may be an important factor in the formation of amorphous aggregates.

Over longer time periods, the nonassociated intermediate conformations would have time to seek alternate forms of intermolecular interaction, leading to fibrillar aggregates. The partitioning to fibrils involves much more precise alignment of the associated intermediates, with concomitant greater entropy costs. Because fibrils are considerably more difficult to dissolve compared to amorphous aggregates, we postulate that they will have a lower free energy than the corresponding amorphous aggregates. It is this lower free energy that ultimately leads to the formation of the fibrils. Fibrils are believed to consist of at least two infinitely long β sheets, with their great stability arising from a combination of tight hydrophobic interactions between the sheets and the continuous β-sheet structure.[31] A major contributing factor to the increased stability of the fibrils will be the tight side-chain packing between the adjacent β sheets in the fibril. The necessity of the correct side-chain alignment/interactions in the fibril nucleus in part accounts for the slower formation of fibrils compared to the formation of amorphous aggregates. Only protein sequences that can provide these specific, tight-interacting side-chain arrangements will lead to fibrils.

One simple explanation to account for whether aggregation leads to amorphous or fibril aggregates is the nature of the initial hydrophobic interactions between the partially folded conformations: strong initial interactions of the core regions leads to the formation of the amorphous aggregates as discussed above, whereas weaker initial interactions of the aggregation-competent partially folded intermediate leads to fibrils. In this case, the disordered regions will be able to sample different modes of association until a highly favorable

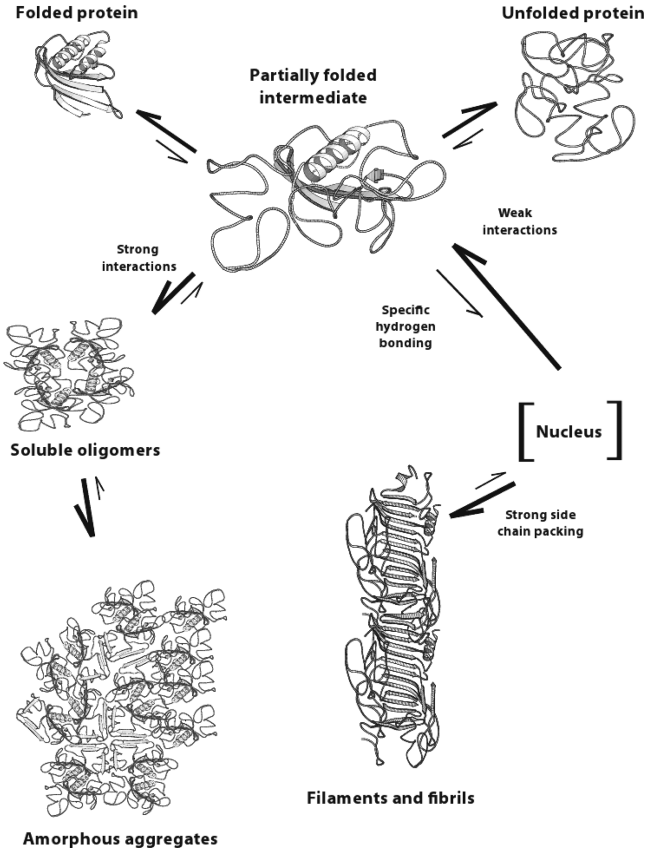


Figure 5. Model representing the formation of amyloid fibrils and amorphous aggregates from partially folded intermediates. The partially folded intermediate can result from either native or unfolded (including natively unfolded) conformations. The formation of the amorphous aggregate involves the association of the solvent-exposed hydrophobic surfaces of the nativelike cores of the intermediates accompanied by the formation of intermolecular β sheets between disordered loops, which first yields small soluble oligomers that increase in size to generate larger aggregates. Amyloid fibril formation, however, involves initial weaker hydrophobic interactions, resulting in a soluble oligomeric prenucleus that probably has more β-sheet formation between the disordered segments of the intermediate. More specific backbone-hydrogen-bonding interactions then give rise to aggregates in the nucleus. The initial fibrillar species, the filaments, results on the formation of tight side-chain packing along with the formation of the crossed-β-sheet core. The filaments then assemble into protofibrils and fibrils, as described previously.

interaction association is found. The core region of the intermediate may remain outside the core of the fibril (as with τ)[45] or become incorporated in the β structure of the fibril. This is most likely to occur with proteins that are high in β-sheet content to begin with. In such a case, rearrangements of the structure in the core to form the crossed-β-sheet structure of the fibril is likely to be a slower process than the simple association of the intermediate core regions to give an amorphous precipitate. An additional critical factor in determining the partitioning between amorphous and fibrillar deposits concerns the structure of the aggregation-competent intermediate. Intermediates that are more nativelike would be expected to favor the amorphous pathway, whereas those that are more unfolded will favor the fibril pathway. Thus, it is ultimately the relative balance of forces involved in the intermolecular interactions that determine the preference for amorphous or fibrillar aggregation.

AUTHOR INFORMATION

Corresponding Author

*E-mail: bhavanashivu@gmail.com. Tel: 507-216-6347.

Present Addresses

[†]Department of Biochemistry and Molecular Biology, Mayo Clinic, 200 1st Street South West, Rochester, Minnesota 55905

[‡]Pharmaceutical Consultant, P.O. Box 3162, Saratoga, California 95070

[§]Department of Neurophysiology, Genentech, 1 DNA Way, South San Francisco, California 94080

[∥]Alfred Mann Foundation, 25134 Rye Canyon Loop, Valencia, California 91355

[⊥]Department of Molecular Medicine, USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida 33612, United States; Institute for Biological Instrumentation, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

Author Contributions

S.S. performed all of the experiments on the ammonium sulfate precipitates, folding aggregates, and inclusion bodies. J.L. carried out the experiments on the protein extracts from cells with and without stress. K.A.O. performed the experiments on the heat-gelled thermal aggregates and developed the software for data analysis. V.N.U. designed the model proposed in the Article. B.S. conducted the experiments on the amyloid fibrils and the barstar native and low-pH form and helped to develop the model and write the manuscript. A.L.F. conceived the idea for the manuscript and the proposed model and wrote the manuscript. B.S. was previously known as Ritu Khurana. All of her earlier publications are in the name of Ritu Khurana.

Funding

This work was supported by grants from the National Science Foundation and National Institutes of Health to A.L.F.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge T. Arakawa and M. Zhang for their contributions. We kindly thank Juhi Juneja and Jayant B. Udgaonkar for the kind gift of purified barstar, Jonathan King for the P22 tailspike protein, Don Downing for KLEG, Youcef Fezoui and Dave Teplow for the Aβ peptide, and Liza Nielsen for the insulin.

DEDICATION

This Article is dedicated to Prof. Anthony L. Fink.

REFERENCES

(1) Harrick, N. J. (1967) *Internal Reflection Spectroscopy*, Wiley Interscience, New York.

(2) Goormaghtigh, E., Raussens, V., and Ruysschaert, J. M. (1999) Attenuated total reflection infrared spectroscopy of proteins and lipids in biological membranes. *Biochim. Biophys. Acta 1422*, 105−185.

(3) Oberg, K. A., and Fink, A. L. (1998) A new attenuated total reflectance Fourier transform infrared spectroscopy method for the study of proteins in solution. *Anal. Biochem. 256*, 92−106.

(4) Seshadri, S., Khurana, R., and Fink, A. L. (1999) Fourier transform infrared spectroscopy in analysis of protein deposits. *Methods Enzymol. 309*, 559−576.

(5) Invernizzi, G., Aprile, F. A., Natalello, A., Ghisleni, A., Penco, A., Relini, A., Doglia, S. M., Tortora, P., and Regonesi, M. E. (2012) The relationship between aggregation and toxicity of polyglutamine-

containing ataxin-3 in the intracellular environment of *Escherichia coli*. *PLoS One 7*, e51890.
(6) Byler, D. M., and Susi, H. (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers 25*, 469−487.
(7) Lee, S. H., Carpenter, J. F., Chang, B. S., Randolph, T. W., and Kim, Y. S. (2006) Effects of solutes on solubilization and refolding of proteins from inclusion bodies with high hydrostatic pressure. *Protein Sci. 15*, 304−313.
(8) Dong, A., Prestrelski, S. J., Allison, S. D., and Carpenter, J. F. (1995) Infrared spectroscopic studies of lyophilization- and temperature-induced protein aggregation. *J. Pharm. Sci. 84*, 415−424.
(9) Caughey, B. W., Dong, A., Bhat, K. S., Ernst, D., Hayes, S. F., and Caughey, W. S. (1991) Secondary structure analysis of the scrapie-associated protein PrP 27−30 in water by infrared spectroscopy. *Biochemistry 30*, 7672−7680.
(10) Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol. 273*, 729−739.
(11) London, J., Skrzynia, C., and Goldberg, M. E. (1974) Renaturation of Escherichia coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. *Eur. J. Biochem. 47*, 409−415.
(12) Speed, M. A., Wang, D. I., and King, J. (1995) Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains. *Protein Sci. 4*, 900−908.
(13) Speed, M. A., Wang, D. I. C., and King, J. (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat. Biotechnol. 14*, 1283−1287.
(14) Wetzel, R. (1992) in *Stability of Protein Pharmaceuticals, Part B; In Vivo Pathways of Degradation and Strategies for Protein Stabilization* (Ahern, T. J., Manning, M. C., Eds.), pp 43−88, Plenum Press, New York, NY.
(15) Wetzel, R. (1994) Mutations and off-pathway aggregation of proteins. *Trends Biotechnol. 12*, 193−198.
(16) Wetzel, R. (1996) For protein misassembly, it's the "I" decade. *Cell 86*, 699−702.
(17) Kelly, J. W. (1996) Alternative conformations of amyloidogenic proteins govern their behavior. *Curr. Opin. Struct. Biol. 6*, 11−17.
(18) Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry 31*, 8654−8660.
(19) 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., and Pepys, M. B. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature 385*, 787−793.
(20) Fink, A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding Des. 3*, R9−23.
(21) Khurana, R., Gillespie, J. R., Talapatra, A., Minert, L. J., Ionescu-Zanetti, C., Millett, I., and Fink, A. L. (2001) Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry 40*, 3525−3535.
(22) Eliezer, D., and Wright, P. E. (1996) Is apomyoglobin a molten globule? Structural characterization by NMR. *J. Mol. Biol. 263*, 531−538.
(23) Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate. *Science 249*, 1544−1548.
(24) Fink, A. L., Oberg, K. A., and Seshadri, S. (1998) Discrete intermediates versus molten globule models for protein folding: characterization of partially folded intermediates of apomyoglobin. *Folding Des. 3*, 19−25.
(25) Fink, A. L., Seshadri, S., Khurana, R., and Oberg, K. A. (2000) in *Infrared Analysis of Peptides and Proteins: Principles and Applications* (Singh, B. R., Ed.), pp 132−144, American Chemical Society, Washington, DC.
(26) Holzbaur, I. E., English, A. M., and Ismail, A. A. (1996) FTIR study of the thermal denaturation of horseradish and cytochrome c peroxidases in D_2O. *Biochemistry 35*, 5488−5494.
(27) Kampinga, H. H., Muller, E., Brunsting, J. F., Heine, L., Konings, A. W., and Issels, R. D. (1993) Association of HSP72 with the nuclear (TX-100-insoluble) fraction upon heating tolerant and non-tolerant HeLa S3 cells. *Int. J. Hyperthermia 9*, 89−98.
(28) Kabakov, A. E., and Gabai, V. L. (1993) Protein aggregation as primary and characteristic cell reaction to various stresses. *Experientia 49*, 706−713.
(29) Laszlo, A., Wright, W., and Roti Roti, J. L. (1992) Initial characterization of heat-induced excess nuclear proteins in HeLa cells. *J Cell. Physiol. 151*, 519−532.
(30) Inouye, H., Fraser, P. E., and Kirschner, D. A. (1993) Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction. *Biophys. J. 64*, 502−519.
(31) Blake, C., and Serpell, L. (1996) Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous beta-sheet helix. *Structure 4*, 989−998.
(32) Fezoui, Y., Hartley, D. M., Harper, J. D., Khurana, R., Walsh, D. M., Condron, M. M., Selkoe, D. J., Lansbury, P. T., Jr, Fink, A. L., and Teplow, D. B. (2000) An improved method of preparing the amyloid beta-protein for fibrillogenesis and neurotoxicity experiment. *Amyloid 7*, 166−178.
(33) Stevens, P. W., Raffen, R., Hanson, D. K., Deng, Y. L., Berrios-Hammond, M., Westholm, F. A., Schiffer, M., Stevens, F. J., Murphy, C., Solomon, A., Eulitz, M., and Wetzel, R. (1995) Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterization of light-chain amyloid proteins. *Protein Sci. 4*, 421−432.
(34) Zandomeneghi, G., Krebs, M. R., McCammon, M. G., and Fändrich, M. (2004) FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. *Protein Sci. 13*, 3314−3321.
(35) Lazo, N. D., and Downing, D. T. (1997) Beta-helical fibrils from a model peptide. *Biochem. Biophys. Res. Commun. 235*, 675−679.
(36) Khurana, R. (1995) pH Dependence of the Conformation of the Small Protein Barstar, Ph.D. Thesis, Tata Institute of Fundamental Research, Mumbai, India.
(37) Juneja, J., Bhavesh, N. S., Udgaonkar, J. B., and Hosur, R. V. (2002) NMR identification and characterization of the flexible regions in the 160 kDa molten globule-like aggregate of barstar at low pH. *Biochemistry 41*, 9885−9899.
(38) Oberg, K., Chrunyk, B. A., Wetzel, R., and Fink, A. L. (1994) Nativelike secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. *Biochemistry 33*, 2628−2634.
(39) Przybycien, T. M., Dunn, J. P., Valax, P., and Georgiou, G. (1994) Secondary structure characterization of beta-lactamase inclusion bodies. *Protein Eng. 7*, 131−136.
(40) Ismail, A. A., Mantsch, H. H., and Wong, P. T. (1992) Aggregation of chymotrypsinogen: portrait by infrared spectroscopy. *Biochim. Biophys. Acta 1121*, 183−188.
(41) Kendrick, B. S., Cleland, J. L., Lam, X., Nguyen, T., Randolph, T. W., Manning, M. C., and Carpenter, J. F. (1998) Aggregation of recombinant human interferon gamma: kinetics and structural transition. *J. Pharm. Sci. 87*, 1069−1076.
(42) Wang, L., Schubert, D., Sawaya, M. R., Eisenberg, D., and Riek, R. (2010) Multidimensional structure-activity relationship of a protein in its aggregated states. *Angew. Chem., Int. Ed. 49*, 3904−3908.
(43) Clark, A. H., Saunderson, D. H., and Suggett, A. (1981) Infrared and laser-Raman spectroscopic studies of thermally-induced globular protein gels. *Int. J. Pept. Protein Res. 17*, 353−364.
(44) Fink, A. L. (1995) Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct. 24*, 495−522.
(45) Goedert, M. (1999) Filamentous nerve cell inclusions in neurodegenerative diseases: tauopathies and alpha-synucleinopathies. *Philos. Trans. R. Soc., B 354*, 1101−1118.