Structure *14*, 993–1001, June 2006  DOI 10.1016/j.str.2006.03.014

# Exploring the Mechanism of Formation of Native-like and Precursor Amyloid Oligomers for the Native Acylphosphatase from *Sulfolobus solfataricus*

**Georgia Plakoutsi,[1] Francesco Bemporad,[1] Maria Monti,[2] Daniela Pagnozzi,[2] Piero Pucci,[2,3] and Fabrizio Chiti[1,*]**
[1] Dipartimento di Scienze Biochimiche
Università di Firenze
Viale Morgagni 50
50134 Firenze
Italy
[2] CEINGE Biotecnologie Avanzate
Napoli & SEMM
European School of Molecular Medicine, Naples Site
Via Comunale Margherita 482
80145 Napoli
Italy
[3] Dipartimento di Chimica Organica e Biochimica
Via Cinthia 4
80126 Napoli
Italy

## Summary

**Over 40 human diseases are associated with the formation of well-defined proteinaceous fibrillar aggregates. Since the oligomers precursors to the fibrils are increasingly recognized to be the causative agents of such diseases, it is important to elucidate the mechanism of formation of these early species. The acylphosphatase from *Sulfolobus solfataricus* is an ideal system as it was found to form, under conditions in which it is initially native, two types of prefibrillar aggregates: (1) initial enzymatically active aggregates and (2) oligomers with characteristics reminiscent of amyloid protofibrils, with the latter originating from the structural reorganization of the initial assemblies. By studying a number of protein variants with a variety of biophysical techniques, we have identified the regions of the sequence and the driving forces that promote the first aggregation phase and show that the second phase consists in a cooperative conversion involving the entire globular fold.**

## Introduction

Proteins and peptides have a generic tendency to convert from their soluble states into well-organized aggregates such as extracellular amyloid fibrils or intracellular inclusions with amyloid-like characteristics (Uversky and Fink, 2004; Chiti and Dobson, 2006). This intrinsic tendency can also be seen in the ability of proteins that have no link to human diseases to form amyloid-like fibrils under appropriate conditions in vitro (Stefani and Dobson, 2003; Uversky and Fink, 2004). Since these aggregates are toxic to the cells, especially in the form of early-forming oligomers and/or amyloid protofibrils, this essential behavior of polypeptide chains represents a constant challenge for every living organism (Buccianti et al., 2002). It was suggested that protein sequences and the cellular machinery have evolved, among other requirements, to inhibit uncontrolled protein aggregation (Dobson, 2003). Failure of the cell's defense mechanisms against protein aggregation often results in pathology. To date, over 40 human diseases, including Alzheimer's, Parkinson's, Creutzfeldt-Jacob's disease, and type II diabetes have been shown to be associated with the formation of amyloid fibrils or intracellular inclusions with amyloid-like properties (Chiti and Dobson, 2006). Interestingly, in addition to its involvement in cellular housekeeping and disease, amyloid fibril formation is also exploited by living organisms for the formation of fibrillar structures with specific biological functions, provided that this occurs under controlled conditions (Chapman et al., 2002; Chiti and Dobson, 2006). Perhaps, the most fascinating of these biological functions is the ability of such structures to serve as transmissible genetic traits distinct from DNA genes (Chien et al., 2004).

Concerning the detrimental features of the aggregated structures, there is an increasing consensus that prefibrillar aggregates, rather than mature fibrils, represent the pathogenic species, at least in some protein deposition diseases (Walsh and Selkoe, 2004). Consequently, the cell needs to constantly repress the formation of these pathological agents. Moreover, although some proteins that adopt a well-defined three-dimensional conformation under physiological conditions generally aggregate following partial unfolding, it has recently been found that some other folded proteins can aggregate following pathways in which unfolding is not the first step in this complex process (Bouchard et al., 2000; Pedersen et al., 2004; Plakoutsi et al., 2004). The observation that all-$\beta$ proteins have evolved structural adaptations to prevent their aggregation via interaction of their native $\beta$ sheets suggests that native-like aggregation is indeed a constant and plausible danger in the highly crowded environments of living organisms where proteins spend most of their lifetime in a folded conformation (Richardson and Richardson, 2001). Prion proteins such as HET-s from *P. anserina* and Ure2p from *S. cerevisiae* are suggested to be able to form amyloid-like fibrils in which the folded units retain in each case native-like conformations (Bousset et al., 2002; Balguerie et al., 2003; Baxa et al., 2003). Following these emerging themes, it becomes important to investigate the mechanism of formation of potentially toxic oligomers under conditions in which globular proteins adopt initially a native-like folded conformation prior to aggregation.

One particularly favorable system with which to study the formation of prefibrillar aggregates under such conditions is the acylphosphatase from *Sulfolobus solfataricus* (Sso AcP). This $\alpha/\beta$ 101 residue protein is characterized, by contrast with the other acylphosphatases so far investigated, by an 11 residue unstructured segment at the N terminus (Corazza et al., 2006) (Figure 1). Sso AcP forms stable amyloid-like protofibrils under conditions in which the protein is initially folded (Plakoutsi et al., 2004). This occurs via an exponential phase in the

*Correspondence: fabrizio.chiti@unifi.it

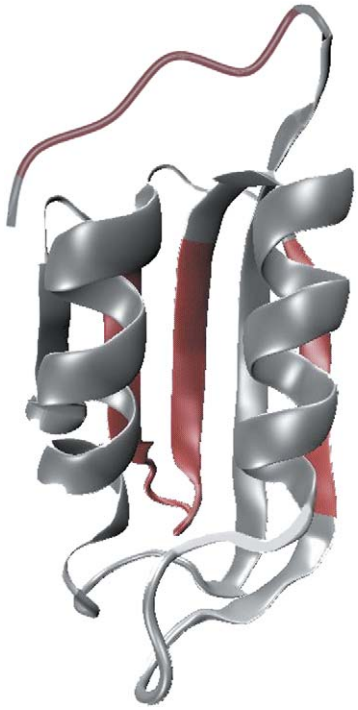

Figure 1. Three-Dimensional Structure of Sso AcP

The structure was determined by X-ray crystallography (Corazza et al., 2006). The PDB entry is 2BJD. The structure determined by NMR spectroscopy shows a similar topology (Corazza et al., 2006). The N-terminal tail, devoid of a precise conformation, is added arbitrarily and has only a descriptive purpose. The regions of native Sso AcP that are most exposed to the solvent and/or have a major flexibility, as identified by limited proteolysis, are indicated in red.

absence of any detectable lag phase. Although protofibril formation also occurs in the absence of added denaturants (Plakoutsi et al., 2004), solutions containing 15%–25% trifluoroethanol (TFE) are generally preferred as these conditions allow protofibril formation to occur more rapidly, on a time scale that is reasonable for lab experiments, while maintaining the protein in a folded conformation prior to aggregation (Plakoutsi et al., 2005).

It has been shown that in 20% (v/v) TFE (pH 5.5) 25ºC, Sso AcP assembles initially, within 1–2 min, into aggregates that (1) do not bind thioflavin T (ThT), Congo red (CR), or 1-anilino-8-napthalenesulfonic acid (ANS); (2) do not contain an extensive β-sheet structure; and (3) retain considerable enzymatic activity (Plakoutsi et al., 2005). In spite of the seemingly native topology, far-UV circular dichroism (CD) and Fourier transform infrared (FT-IR) spectroscopies indicate that subtle conformational changes are associated with the formation of these species (Plakoutsi et al., 2005). These oligomers convert, with no apparent need of disassembly and renucleation and with a rate that is independent of protein concentration, into spherical aggregates that (1) possess a diameter of 3–5 nm and assemble further to form chain-like structures; (2) bind ThT, CR, and ANS; (3) possess extensive β-sheet structure; and (4) do not retain enzymatic activity (Plakoutsi et al., 2005). We will refer to the first and second species as the "initial aggregates" and "amyloid protofibrils," respectively. Interestingly, the two processes occur with rates that are, correspondingly, three and two orders of magnitude more rapid than Sso AcP unfolding under the same conditions, indicating that unfolding is not required to initiate this self-assembly process. This has led to a mechanism of aggregation that can be summarized with a simplified scheme of the type $N \leftrightarrow (N)_n \leftrightarrow PF$, where N is the monomeric folded state, and $(N)_n$ and PF represent the initial aggregates with native-like topology and the amyloid protofibrils, respectively (Plakoutsi et al., 2005). Although PF can form directly from N, the kinetic analysis shows that most of the molecules transit through the initial aggregates (Plakoutsi et al., 2005).

The driving forces and the mechanism, at a molecular level, of this peculiar aggregation process remain to be elucidated. By combining a number of experimental approaches, we will identify the regions of the sequence and the driving forces that promote the first aggregation phase and show that the second phase consists in a cooperative conversion involving the entire globular fold as a whole, rather than specific regions of the sequence. We will also emphasize that such an aggregation mechanism is conceptually distinct from those proposed for other globular proteins and for systems with prion-like properties.

## Results

### The Unstructured N-Terminal Stretch Is Essential for Aggregation of Sso AcP

The presence of an 11 residue unstructured tail at the N terminus of native Sso AcP (Figure 1) prompted us to investigate the role of this sequence segment in the aggregation of the full-length protein. Three deletion mutants were produced, lacking 5, 8, and 11 residues at the N terminus (ΔN5, ΔN8, and ΔN11 Sso AcP, respectively). The aggregation of the truncated variants was studied at a protein concentration of 0.4 mg $ml^{-1}$ in 20% (v/v) TFE, 50 mM acetate buffer (pH 5.5), 25ºC, conditions under which Sso AcP still adopts a folded structure prior to aggregation (Plakoutsi et al., 2004). While the ΔN5 and ΔN8 variants of Sso AcP are still able to form the ThT binding protofibrils, albeit with rates that appear progressively slower upon truncation, the ΔN11 variant does not convert into such aggregates even after prolonged exposure to these conditions (Figure 2A).

The CD spectrum of the ΔN11 variant does not undergo detectable changes following incubation of the mutant under these conditions and remains similar to those of native wild-type and native ΔN11 Sso AcP acquired at pH 8.0 in the absence of TFE (Figure 2B). This spectrum is readily distinguishable from those of the initial aggregates and protofibrils that are recorded for the wild-type and the other truncated variants during aggregation (Figure 2B). The self-assembly of the native-like protein into initial aggregates can be monitored by following the mean residue ellipticity at 208 nm ($[\theta]_{208}$), as the difference of the signal between the two conformational states is maximal at this wavelength (Figure 2B). The increase of $[\theta]_{208}$ becomes progressively slower with the shortening of the N-terminal tail, with the ΔN11 mutant maintaining a stable native-like signal up to 4 hr (Figure 2C). The formation of the initial aggregates can also be monitored effectively with light scattering, as this phase corresponds to a rapid high-amplitude increase of the light scattering

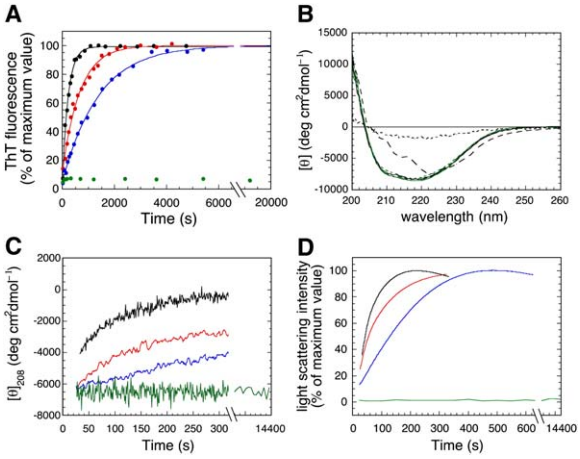


Figure 2. Aggregation Kinetics of Sso AcP

The traces are reported as changes in the fluorescence intensity of ThT at 485 nm (A), in far-UV CD spectra (B), in mean residue ellipticity at 208 nm (C), and in light-scattering intensity during aggregation (D). The kinetic traces in (A), (C), and (D) refer to the wild-type (black), ΔN5 (red), ΔN8 (blue), and ΔN11 (green) Sso AcP. The ThT and light scattering traces (A and D) were normalized to attribute 100% to the maximum value. In all experiments, conditions were 0.4 mg $ml^{-1}$ protein, 20% (v/v) TFE, 50 mM acetate (pH 5.5) at 25°C. The far-UV CD spectra reported in (B) for native wild-type and native ΔN11 Sso AcP in 10 mM Tris (pH 8.0), 25°C, are roughly superimposable to those of ΔN11 Sso AcP incubated under aggregating conditions in 20% TFE for 1 min, 10 min, and 4 hr. By contrast, the spectra of the initial aggregates (2 min) and protofibrils (30 min) for wild-type Sso AcP are distinct (dashed and dotted lines, respectively).

intensity, followed by a slower low-amplitude decrease as the initial aggregates convert into protofibrils (Plakoutsi et al., 2005). Neither increase nor decrease was observed with ΔN11 Sso AcP (Figure 2D). From these data we conclude that the unstructured segment at the N terminus is essential to promote formation of the initial aggregates, and consequently of the amyloid protofibrils, in the full-length Sso AcP.

## The Isolated Unstructured Tail of Sso AcP Does Not Aggregate

To assess whether the first step of Sso AcP aggregation consists in the stacking of N-terminal tails, with the globular units remaining folded to decorate the core of such assembled tails, an 11 residue peptide with a sequence identical to the N-terminal tail of Sso AcP was synthesized (sequence MKKWSDTEVFE). The lyophilized peptide was suspended at concentrations ranging from 30 to 350 μM in the absence or presence of 20% (v/v) TFE for 30 min (50 mM acetate [pH 5.5], 25°C). The peptide concentrations for the various samples, determined spectrophotometrically after filtration with 0.02 μm filters, are in good agreement with the concentrations of the total peptide before filtration (Figure 3A). This indicates that the peptide remained largely soluble in all samples, even at molar concentrations ten times higher than those used here to promote aggregation of the full-length protein. The high solubility of the peptide probably results from the high number of polar charged residues present in its sequence.

The size distributions acquired by dynamic light scattering (DLS) show that the peptide has an apparent hydrodynamic diameter of 1.5 ± 0.3 and 1.95 ± 0.5 in 0% and 20% (v/v) TFE, respectively (Figure 3B). These values are within error to that expected for an 11 residue peptide in its unstructured monomeric form, corresponding to 1.85, 1.73, and 1.61 nm with the formulas reported by three independent groups (Damaschun et al., 1998; Wilkins et al., 1999; Kohn et al., 2005). Since the intensity of the scattered light scales with the square of the molecular mass of the light-scattering particle, the species with diameters higher than 4 nm in the reported size distributions are quantitatively irrelevant.

The far-UV CD spectra of the peptide samples in 0% and 20% (v/v) TFE have both a negative peak at 198–200 nm, indicating a largely unstructured conformation in both solution conditions (Figure 3C). The results obtained with the solubility assay, DLS, and far-UV CD were substantially similar when the experiments were repeated with an extended 14 residue peptide (sequence MKKWSDTEVFEMLK), corresponding to the first 14 residues of Sso AcP (data not shown). In summary, while full-length Sso AcP aggregates under the investigated conditions, neither the N-terminal tail nor the globular unit can do so when separated from each other.

## Native Sso AcP Has Portions of the Sequence with a Preferential Flexibility

To map out the specific residues or regions of the sequence of Sso AcP that appear most exposed to the solvent and/or flexible, and therefore available for intermolecular interactions, a limited proteolysis approach was used. Under the conditions used for these experiments (pH 7.5 and 20 μM protein concentration), aggregation was sufficiently slow in both the absence and presence of TFE as to allow the folded monomeric full-length form of the protein to be characterized. Size-exclusion high-performance liquid chromatography (SE-HPLC) confirmed that Sso AcP maintains a monomeric state in all tested conditions over the time required for the completion of the experiment (data not shown). In the absence of TFE, six proteases were used with the aim of creating conditions in which the selectivity of the cleavage was not limited by the specificity of individual enzymes. The proteases were trypsin, chymotrypsin, subtilisin, elastase, endoprotease Glu-C, and thermolysin. The latter two were found to be inactive in the presence of TFE, and the experiments were carried out with the four remaining proteases. In each experiment, aliquots of 1 μl of the digestion mixture were withdrawn at different time intervals and analyzed by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) in linear mode, as described in the Experimental Procedures section. Direct peptide analysis by MALDI-MS allowed the extent of proteolytic digestions to be monitored on a much shorter time scale with much fewer protein samples. Moreover, small peptide components could also be detected due to the higher sensitivity of the mass spectral analysis. A proteolytic site between residues *a* and *b* was evaluated to occur on the full-length protein provided that the two complementary fragments (from the N terminus to *a* and from *b* to the C terminus) could be identified by MALDI-MS.

The results obtained in 0%, 15%, and 25% (v/v) TFE are summarized in Figure 4. In all conditions, preferential proteolytic sites clustered within three regions of the protein,

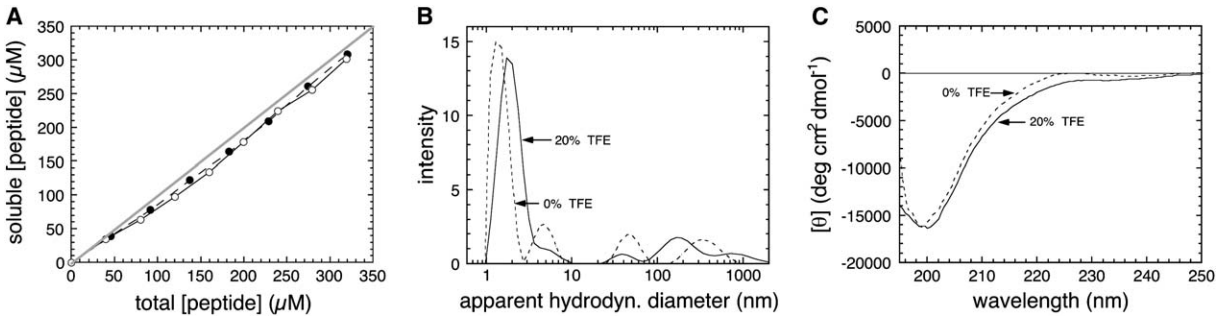


Figure 3. Characteristics of the Synthetic Peptide Corresponding to the 11 Residue N-Terminal Tail of Sso AcP

(A) Concentration of the peptide, determined spectrophotometrically after filtration with 0.02 μm filters, versus the peptide concentration before filtration.

(B) Distribution of apparent hydrodynamic diameter for a solution containing 400 μM peptide, estimated by DLS.

(C) Far-UV CD spectra of the peptide at a concentration of 85 μM.

In all cases (A–C), the measurements were taken without (dashed line) or with (solid line) 20% (v/v) TFE, in 50 mM acetate (pH 5.5) at 25ºC.

encompassing approximately residues 1–12, 44–61, and 83–91. These three regions of the sequence correspond, in the native protein, to the unstructured N-terminal tail, to the β hairpin formed by strands 2 and 3, and to β strand 4 plus a short segment of the following loop, respectively (Figure 4). In native Sso AcP, β strands 2 and 3 occupy internal position in the β sheet, while strand 4 is at one edge of the sheet (Figure 1). The observation that the pattern of proteolytic sites does not change significantly with the addition of TFE is an indication that a preferential solvent-exposure and/or chain flexibility is showed by essentially the same stretches of the polypeptide chain under the three sets of experimental conditions.

### The Conversion of the Initial Aggregates into Amyloid Protofibrils

In the light of the limited proteolysis results, several mutants of Sso AcP were produced (Table 1). One mutant involves a double substitution within the region 1–12 that appears to be susceptible to proteolysis (V9A/F10A). This mutant adds to the ΔN5, ΔN8, and ΔN11 deletion variants described above. Six other variants involve single substitutions of residues from the flexible region 44–61 (Y45A, K47A, and N48A) and 83–91 (V84A, Y86A, and F88A). Two additional single-point mutants from nonflexible regions were produced as controls (F29L and I72V). All variants were incubated at a concentration of 0.4 mg $ml^{-1}$ in 20% (v/v) TFE, 50 mM acetate buffer (pH 5.5), 25ºC. The conversion of the monomeric folded structure into native-like enzymatically active aggregates was followed by monitoring the change of $[\theta]_{208}$ and light scattering. The second step of aggregation (the conversion of the initial aggregates into amyloid-like protofibrils) was monitored by ThT fluorescence. For simplicity, we shall describe the second step first.

The increase of ThT fluorescence followed apparent single-exponential kinetics for most of the mutants, as shown previously for the ΔN5 and ΔN8 variants (Figure 2A). The resulting rate constant, $k_2^{ThT}$, was found for nearly all mutants to be significantly different from that of the wild-type protein, including the two control variants F29L and I72V (Table 1). A statistically significant correlation was found between the logarithm of $k_2^{ThT}$ and the conformational stability in this group of protein variants (the latter is expressed as the change of free energy change of unfolding, $\Delta\Delta G_{U\text{-}F}$, as determined from GdnHCl unfolding curves at equilibrium) (Figure 5A). The r and p values (0.95 and <0.001, respectively) indicate that the correlation is significant with all data points deviating by less than 2 SD from the line of best fit (Figure 5A).

This relationship indicates that the conversion of individual protein molecules from the native-like conformation adopted within the initial aggregates into that of the amyloid protofibrils is a cooperative process. The acceleration of this process correlates with the extent of the destabilization of the folded state induced by mutation and the resulting ability of the protein to populate fully or partially unfolded conformations that are competitive for amyloid protofibril formation. In this mechanism, the whole protein molecule undergoes a substantial structural reorganization. The mutation-induced acceleration of this conversion is independent of the position of the

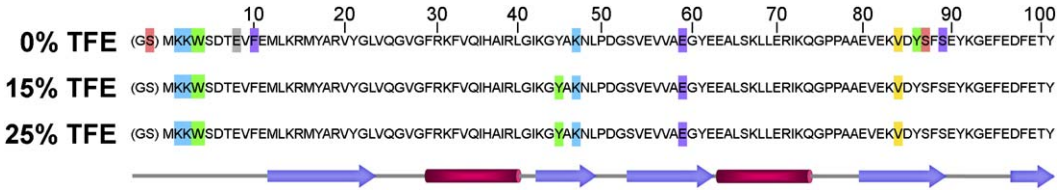


Figure 4. Location of Proteolytic Sites Obtained for the Folded, Monomeric, Wild-Type Sso AcP

The three diagrams show the results obtained in 0%, 15%, and 25% (v/v) TFE with trypsin (blue), chymotrypsin (green), subtilisin (purple), thermolysin (pink), elastase (yellow), and Glu-C (grey). In all cases, the proteolytic cleavage took place between the colored residue and the following one in the sequence. Glu-C and thermolysin were not used in 15% and 25% (v/v) TFE because they were not active under these conditions. The figure also shows the location of β strands and α helices along the sequence according to the X-ray structure (Corazza et al., 2006) and the previously proposed algorithm (Kabsch and Sander, 1983).

Table 1. Parameters of Conformational Stability and Aggregation Rate of Sso AcP Variants

| Mutant | $C_m$ (M)[a] | $\Delta\Delta G_{U-F}$ (kJ mol$^{-1}$)[b] | $k_1^{CD}$ (s$^{-1}$)[c] | $k_1^{LS}$ (s$^{-1}$)[c] | $k_2^{ThT}$ (s$^{-1}$)[c] |
|---|---|---|---|---|---|
| Wild-type | 4.23 ± 0.07 | — | 0.014 | 0.0174 | 0.0045 |
| ΔN5 | n.d. | n.d. | n.d. | n.d. | n.d. |
| ΔN8 | 4.59 ± 0.07 | 4.0 ± 1.7 | 0.0054 | 0.0055 | 0.00072 |
| ΔN11 | 4.59 ± 0.07 | 4.0 ± 1.7 | 0.00005[d] | 0.00005[d] | — |
| V9A/F10A | 4.45 ± 0.07 | 2.45 ± 1.6 | 0.0925 | 0.0084 | 0.0217 |
| F29L | 3.58 ± 0.07 | −7.25 ± 1.55 | 0.25 | 0.328 | 0.0572 |
| Y45A | 3.25 ± 0.07 | −10.9 ± 1.5 | 0.18 | 0.434 | 0.116 |
| K47A | 4.24 ± 0.07 | 0.1 ± 1.6 | 0.044 | 0.0819 | 0.00799 |
| N48A | 4.34 ± 0.07 | 1.2 ± 1.6 | 0.0179 | 0.0324 | 0.0116 |
| I72V | 3.78 ± 0.07 | −5.0 ± 1.6 | 0.079 | 0.0928 | 0.0111 |
| V84A | 3.20 ± 0.07 | −11.5 ± 1.5 | 0.034 | 0.065 | 0.111 |
| Y86A | 3.94 ± 0.07 | −3.25 ± 1.6 | 0.147 | 0.44 | 0.0117 |
| F88A | 2.92 ± 0.07 | −14.6 ± 1.5 | 0.11 | 0.11 | 0.183 |

All the experiments were performed in 20% (v/v) TFE, 50 mM acetate (pH 5.5) at 25ºC.
[a] The GdnHCl concentration required to unfold 50% of the protein molecules.
[b] Determined using the equation $\Delta\Delta G_{U-F} = \langle m \rangle (C'm - Cm)$, where <m> is the average m value of all mutants (<m> = 11.14 + 0.20 kJ mol$^{-1}$ M$^{-1}$) and Cm and C′m are the midpoint of denaturation for wild-type and mutant, respectively.
[c] Aggregation rate constants estimated by CD, LS and ThT analyses. Experimental uncertainties were 12%, 10% and 4% for $k_1^{CD}$, $k_1^{LS}$, and $k_2^{ThT}$, respectively.
[d] These values are not real estimates but experimental upper bonds.

mutated site along the sequence and appears to depend only on the magnitude of the destabilization of the overall fold caused by the amino acid substitution.

No correlation was found between the observed change of $k_2^{ThT}$ upon mutation and the change predicted for the conversion of an unfolded state into amyloid-like aggregates, determined on the basis of numerous proposed algorithms (DuBay et al., 2004; Fernandez-Escamilla et al., 2004; Tartaglia et al., 2004; Chiti et al., 2003). This indicates further that destabilization of the native-like fold within the initial aggregate is a major factor driving its conversion, compared to the establishment of a specific β core within the subsequent amyloid-like aggregates.

### The Conversion of the Native-like Monomeric Protein into the Initial Aggregates

The assembly into native-like aggregates from the monomeric folded state also occurs with different rates for the various mutants investigated here (Table 1). However, unlike the second step of aggregation, a correlation between the logarithm of the aggregation rate constant determined with far-UV CD ($k_1^{CD}$) or light scattering ($k_1^{LS}$) and conformational stability ($\Delta\Delta G_{U-F}$) was only found with some of the mutants from the flexible region 1–12 (ΔN8 and V9A/F10A), from the flexible region 44–61 (N48A, K47A, and Y45A), and with the two control mutants (F29L and I72V).

All the three variants from the flexible region 83–91 (V84A, Y86A, and F88A) were found to have log($k_1^{CD}$) and log($k_1^{LS}$) values that deviate by more than 2 SD ($p < 0.05$) from the corresponding lines of best fit, i.e., from the values expected on the basis of only their conformational stability (Figures 5B and 5C, empty squares). The ΔN11 mutant also deviates from the expected correlation as the upper bound values of log($k_1^{CD}$) and log($k_1^{LS}$) for this mutant are both remarkably distant from the corresponding lines of best fit (Figures 5B and 5C, empty squares). This indicates that the flexible region 83–91, similarly to the unstructured N-terminal tail, plays a key role in the first phase of aggregation of Sso AcP that is detected here. For clarity, all the rate constants defined here ($k_1^{CD}$, $k_1^{LS}$, and $k_2^{ThT}$) do not have any relationship with the lag phase as this is not detected

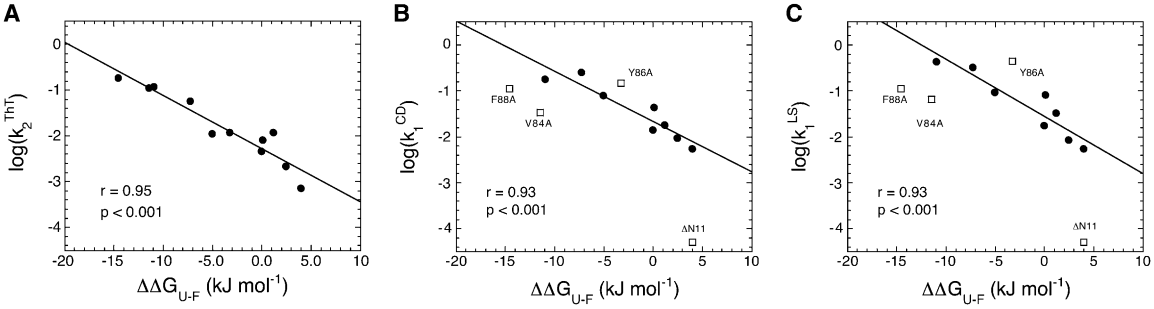


Figure 5. Aggregation Rates of Sso AcP Variants

Dependencies of the apparent rate constants for the formation of protofibrils as followed by using ThT fluorescence, $k_2^{ThT}$ (A), for the formation of the initial aggregates as followed by using CD, $k_1^{CD}$ (B), and for the formation of the initial aggregates as followed by using light scattering, $k_1^{LS}$ (C), on the change of conformational stability of Sso AcP upon mutation, $\Delta\Delta G_{U-F}$. In all plots, filled circles and empty squares indicate values that deviate by less and more than 2 SD from the corresponding lines of best fit, respectively. The log($k_1^{CD}$) and log($k_1^{LS}$) values of the ΔN11 mutant are not real values but upper bonds of such values.

for the aggregation of Sso AcP under the investigated conditions. In addition to having unusual values of $k_1^{CD}$ and $k_1^{LS}$, all three variants having mutations within the region 83–91 give rise to initial aggregates with CD spectra that are significantly different from those of the wild-type protein and all other variants (data not shown). This lends further support to the idea that this region of the sequence plays a crucial role in this first aggregation step of Sso AcP.

The presence of a correlation between $k_1^{CD}$ (or $k_1^{LS}$) and $\Delta\Delta G_{U\text{-}F}$ for all other variants and the finding that the line of best fit has a slope similar to that observed in the relationship between $k_2^{ThT}$ and $\Delta\Delta G_{U\text{-}F}$ (cf. Figures 5A, 5B, and 5C) suggests that the assembly of native-like molecules into the initial aggregates requires a flexibility of the folded state that correlates with its conformational stability. This step does not require, however, full unfolding because (1) this aggregation step is three orders of magnitude more rapid than the global unfolding of the nonaggregated protein under the same conditions, and (2) the resulting aggregates preserve a residual native-like topology as witnessed by the persistence of enzymatic activity (Plakoutsi et al., 2005). In the next paragraph, we will show that in the folded state of Sso AcP significant structural fluctuations occur more rapidly than unfolding by three orders of magnitude.

### The Folded Unit of Sso AcP Maintains a Native-like Conformation Prior to Aggregation

The ability of ΔN11 Sso AcP to resist aggregation, under conditions of TFE and pH in which the full-length protein by contrast self-assembles, provides an opportunity to investigate the structure, dynamics, and stability of the globular unit of Sso AcP under conditions that promote aggregation. The near-UV CD spectrum of ΔN11 Sso AcP in 20% (v/v) TFE is very similar to that of the wild-type protein in its absence, with a major positive peak at ca. 270–280 nm in both cases (Figure 6A). The enzymatic activity of ΔN11 Sso AcP, determined as the rate of catalysed hydrolysis of benzoylphosphate, decreases linearly with TFE concentration, with the enzyme exhibiting ca. 77% of the initial activity in 20% (v/v) TFE (Figure 6B). The linear decrease, in the range of 0%–50% (v/v) TFE, suggests that such reduction originates from solvent effects rather than caused by a significant structural transition. Incubation of the full-length protein under the same conditions in 20% TFE leads to the formation of initial aggregates and late protofibrils with enzymatic activities of ca. 63% and 0% of the corresponding enzyme with no TFE (Figure 6B).

The conformational stability of ΔN11 Sso AcP in 20% (v/v) TFE, studied by means of GdnHCl-induced equilibrium unfolding, is substantially lower than those of wild-type and ΔN11 Sso AcP in 0% (v/v) TFE (Figure 6C). Although unfolding of ΔN11 Sso AcP in 20% (v/v) TFE occurs at a denaturant concentration ($C_m$) higher than in its absence, the cooperativity of the former transition is considerably lower, as indicated by a lower m value (3.7 ± 1.2 kJ mol$^{-1}$ M$^{-1}$ for ΔN11 Sso AcP in 20% [v/v] TFE versus 11.14 ± 0.20 kJ mol$^{-1}$ M$^{-1}$ for wild-type and ΔN11 Sso AcP in 0% [v/v] TFE). The values of free energy change of unfolding ($\Delta G_{U\text{-}F}$), that result from the products of $C_m$ and m values, are 22.6 ± 7.3, 48.0 ± 1.2, and 52.1 ± 1.2 kJ mol$^{-1}$ for ΔN11 in 20% (v/v) TFE, wild-type in 0% (v/v) TFE, and ΔN11 Sso AcP in 0% (v/v)TFE, respectively.

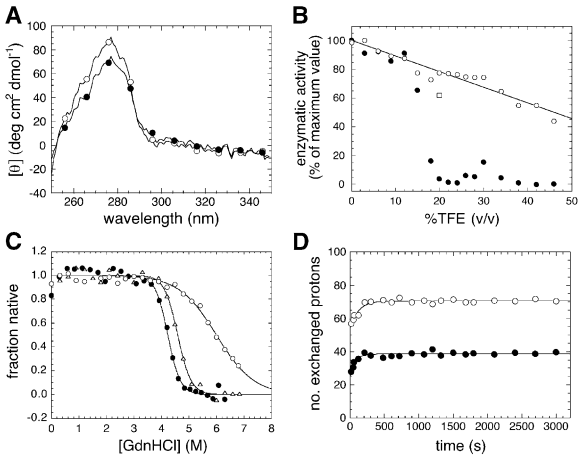


Figure 6. Properties of Native Monomeric Sso AcP under Conditions Promoting Aggregation

(A) Near-UV CD spectra of wild-type Sso AcP in the absence of TFE (filled circles) and of ΔN11 Sso AcP in 20% (v/v) TFE (empty circles). Conditions were 50 mM acetate buffer (pH 5.5), 25ºC.

(B) Enzymatic activity of wild-type (filled circles) and ΔN11 (empty circles) Sso AcP as a function of TFE concentration. Wild-type Sso AcP was preincubated for 30 min under the indicated TFE concentrations before the measurements. The activity of the initial aggregates formed from wild-type Sso AcP in 20% (v/v) TFE is also indicated (empty square). The final protein concentration was in all cases 0.0015 mg ml$^{-1}$ in 50 mM acetate buffer (pH 5.5), 25ºC.

(C) Equilibrium GdnHCl unfolding curves of wild-type Sso AcP with no TFE (filled circles), ΔN11 Sso AcP with no TFE (empty triangles) and ΔN11 Sso AcP in 20% TFE (empty circles). The fitting was based on the two-state equation as described (Santoro and Bolen, 1988).

(D) Kinetic profiles for global H/D exchange of wild-type Sso AcP in 0% (filled circles) and 25% (empty circles) TFE at pH 7.5.

A shift of pH from 5.5 to 7.5 slows down considerably the aggregation of the full-length protein in 15%–25% TFE, allowing the conformational state adopted by Sso AcP in the presence of TFE to be explored, albeit at a different pH. Full-length Sso AcP was incubated in the presence of 0% and 25% (v/v) deuterated TFE in 10 mM ammonium acetate containing $D_2O$ rather than $H_2O$, as a solvent. In both conditions, the H/D exchange, quantified at regular time points by liquid chromatography mass spectrometry (LCMS), is approximately a biphasic process, with some of the atoms exchanging on a subminute time scale and others exchanging on a slower time scale of ca. 300 s (Figure 6D). However, it is clear that the number of atoms exchanged during both phases is higher in the presence of TFE than in its absence (Figure 6D). Importantly, the slower H/D exchange phase in 25% (v/v) TFE is more rapid than unfolding of the protein under the same conditions by two orders of magnitude, indicating that the H/D exchange occurs due to rapid structural fluctuations rather than full unfolding of the protein.

Overall, these data, along with the far-UV CD analysis shown in Figure 2B, indicate that the globular unit of Sso AcP maintains, under aggregating conditions but before aggregation occurs, a secondary structure, packing around aromatic residues and positioning of residues in the catalytic site (Arg30, Asn48, and the loop $^{22}$GLVQGVG$^{28}$) that are indistinguishable from those of

the protein under conditions in which it is stable in its native state. In spite of this native-like structure, the folded state of Sso AcP appears to be less stable and more flexible in aggregating conditions than in fully native conditions.

## Discussion

In this work, we have investigated the mechanism by which a globular protein in its fully folded state self-assembles to form those amyloid protofibrils that are increasingly recognized to represent the pathogenic species in protein deposition diseases. The globular unit of Sso AcP adopts initially, prior to aggregation, a native-like structure. Aggregation is promoted by region 1–12 (the unstructured N-terminal tail) and region 83–91 (approximately β strand 4) to form initial aggregates in which the individual protein molecules retain characteristics that are still reminiscent of the native state and far from those of amyloid structures. This process does not require full unfolding of the protein, as it occurs with a rate that is three orders of magnitude higher than unfolding under the same conditions (Plakoutsi et al., 2004). However, it occurs more efficiently in the presence of a solution medium in which the native state appears to be less cooperatively folded and sufficiently flexible to transiently expose most of its backbone amides to hydrogen exchange with the solvent. In the presence of mutations that destabilize the folded state, this native-like assembly occurs more rapidly, lending further support to the idea that local fluctuations are needed to promote the process.

The rationale behind the involvement of the two regions of the sequence 1–12 and 83–91 lies in the preferential solvent exposure and/or flexibility of these two segments. The N-terminal tail, in particular, does not possess any degree of secondary or long-range interactions and is therefore the most susceptible region to nucleate aggregation. The region 83–91 corresponds to one of the two edge strands of the native β sheet of Sso AcP. Edge β strands are so effective in promoting aggregation of native structures that evolution has devised a number of adaptations to prevent this phenomenon (Richardson and Richardson, 2001). Region 44–61 also appears to be solvent exposed and flexible. Nevertheless, since it forms a β hairpin that is internal in the native β sheet of Sso AcP (strands 2 and 3), it does not in fact promote the assembly of native-like molecules.

Unlike the first step, the subsequent conversion of the initial aggregates into amyloid protofibrils does not appear to be promoted by specific residues or regions of the sequence. The effect of mutations on the rate of this process does not depend on the position of the mutated residues in the sequence or in the three-dimensional fold but only on the extent by which the native state appears to be destabilized. It therefore appears to be a cooperative process that involves a global structural reorganization of each individual molecule from a native-like fold into a nonnative amyloid-like conformation able to bind ThT, CR, and with extensive β sheet structure.

The aggregation mechanism outlined here for Sso AcP is distinct from that described previously for other globular proteins, which involve a significant unfolding of the native globular state before structured aggregates can form (Uversky and Fink, 2004). The first phase of Sso AcP aggregation appears to require structural fluctuations of the native state, but the resulting conformations remain on the native side of the major free-energy barrier of unfolding. The aggregation mechanism described here for Sso AcP is also distinct from that proposed for proteins with prion-like properties such as Sup35p and Ure2p from *S. cerevisiae* and HET-s from *P. anserina*. These proteins resemble Sso AcP as they are all characterized by an unstructured tail, either at the N or C terminus, and a folded domain. Models of fibril formation in which the tails assemble autonomously to form the core of the fibrils have been proposed for all prion systems (Glover et al., 1997; Balguerie et al., 2003; Baxa et al., 2003). In analogy with this idea, the unstructured domains can form amyloid-like fibrils in vitro and maintain a prion status in vivo when dissected from their respective folded domains (Masison and Wickner, 1995; Glover et al., 1997; Kochneva-Pervukhova et al., 1998; Taylor et al., 1999; Balguerie et al., 2003). In contrast to these models, the N-terminal tail of Sso AcP remains fully soluble and monomeric when isolated from the folded domain and needs to interact with another portion of the protein to promote aggregation of the full-length protein. In another study, however, Ure2p was found to form fibrils in which the unstructured tails interact with the folded domains, with the latter maintaining a native-like conformation (Bousset et al., 2002). This mechanism is similar to that proposed here for the initial aggregates of Sso AcP.

The proposed mechanism, along with the increased awareness that folded proteins retain a subtle but significant tendency to aggregate, has implications for our understanding of disease-related aggregation processes in vivo, therapeutic treatment, and, more generally, potentially harmful aggregation processes that may occur physiologically and need to be actively combated by living organisms.

### Experimental Procedures

#### Materials

ThT, TFE, GdnHCl, urea, and $D_2O$ were purchased from Sigma-Aldrich. Benzoylphosphate was synthesized and purified as described (Camici et al., 1976). DNA oligonucleotides were purchased from M-Medical (Milano, Italy). The mutated genes of Sso AcP were created with the QuickChange Site-Directed Mutagenesis Kit from Stratagene. The desired mutations were verified by DNA sequencing. Sso AcP was expressed, purified, tested for purity, and stored as described previously (Plakoutsi et al., 2004). The synthetic peptides were purchased from Genscript Corporation and had amidated C termini.

#### ThT Kinetics

Sso AcP and its variants were incubated in 20% (v/v) TFE, 50 mM acetate buffer (pH 5.5), 25°C, at a protein concentration of 0.4 mg $ml^{-1}$. The experiments were performed as described (Plakoutsi et al., 2004). The value of the ThT fluorescence intensity in the absence of protein was subtracted from all the fluorescence measurements in its presence, and the resulting values were normalized so that the final fluorescence intensity at the endpoint of the kinetic trace was 100%.

#### Far-UV CD

Far-UV CD spectra were acquired at 25°C with a Jasco J-810 spectropolarimeter (Tokyo, Japan) equipped with a thermostated cell holder and a quartz cuvette of 1 mm path length. Wild-type and mutant Sso AcP were incubated at a concentration of 0.4 mg $ml^{-1}$ in either 20% (v/v) TFE, 50 mM acetate buffer (pH 5.5), or 0% TFE, 10 mM Tris (pH 8.0),

25ºC. In another set of experiments, time-course measurements were performed at 208 nm for 0.4 mg $ml^{-1}$ wild-type and mutant Sso AcP in 50 mM acetate buffer (pH 5.5), 25ºC, immediately after addition of 20% (v/v) TFE. The spectra of the 11 and 14 residue peptides were acquired immediately after suspension in 0% (v/v) TFE and 30 min after the addition of 20% (v/v) TFE at a concentration of ca. 85 μM and 152 μM, respectively (conditions were 50 mM acetate buffer [pH 5.5], 25ºC). In all cases, the spectra were blank subtracted and normalized to mean residue ellipticity.

### Dynamic Light Scattering

The 11 and 14 residue peptides corresponding to the N-terminal tail of Sso AcP were suspended at a concentration of 400 μM in 50 mM acetate (pH 5.5), 25ºC, in the presence and absence of 20% (v/v) TFE. Size distributions by light scattering intensity were acquired with a Zetasizer Nano S DLS device from Malvern Instruments (Malvern, Worcestershire, UK). Low-volume 12.5 × 45 mm disposable cuvettes were used. A Peltier thermostatting system maintained the temperature at 25ºC. The viscosity and refractive index parameters were set for each solution. The buffer and stock protein solutions were centrifuged (18,000 rpm, 5 min) and filtered with 0.02 μm Anotop 10 filters (Whatman, Maidstone, UK) before the measurements.

### Solubility Assay of the Sso AcP Peptides

The lyophilized 11 and 14 residue peptides were dissolved in 0% or 20% (v/v) TFE, 50 mM acetate (pH 5.5), at concentrations ranging from 50 to 400 μM, and were incubated for 30 min at 25ºC. The samples were then centrifuged (18,000 rpm, 5 min) and filtered (0.02 μm). Peptide concentration was determined spectrophotometrically with an Ultrospec 2000 spectrophotometer (Pharmacia Biotech, Cambridge, UK) and a 1 mm path length quartz cuvette. The total peptide concentration present before centrifugation and filtration was determined spectrophotometrically for each peptide with a solution containing 150 mM peptide in 5 M urea, 10 mM Tris (pH 8.0), 25ºC.

### Limited Proteolysis

Limited proteolysis experiments were carried out by incubating 20 μM Sso AcP with either trypsin, chymotrypsin, subtilisin, elastase, endoprotease Glu-C, or thermolysin. Enzymatic digestions were all performed in 20 mM Tris-HCl (pH 7.5), 25ºC, with enzyme-to-substrate ratios ranging from 1/3000 to 1/20 (w/w). Enzymatic digestions at different concentrations of TFE were carried out under the same conditions in the presence of 15% and 25% (v/v) TFE with trypsin, chymotrypsin, subtilisin, and elastase with enzyme-to-substrate ratios ranging from 1/960 to 1/28 (w/w). The extent of the proteolytic digestion was monitored on a time-course basis by sampling the incubation mixture at different time intervals and directly analyzing the released peptides by MALDI-MS. For each measurement, an aliquot of 1 μl of peptide mixture was applied to a sample slide and mixed with 5 μl of a 5 mg/ml sinapinic acid solution in acetonitrile/0.2% TFA (70:30, v/v) before air drying. The samples were then analyzed on a linear Voyager DE MALDI-TOF mass spectrometer (Applied Biosystems, Foster City, CA). The mass range was internally calibrated with the $[M+H]^+$ and $[M+H_2]^{2+}$ ions from the entire Sso AcP, always present in the proteolytic mixture.

### Equilibrium Unfolding

Samples of wild-type and mutated Sso AcP (8.5 μM) were incubated for 1 hr in 50 mM acetate (pH 5.5), 37ºC, and different concentrations of GdnHCl ranging from 0 to 6.8 M. The values of mean residue ellipticity at 222 nm ($[\theta]_{222}$) of the samples were measured with the same CD apparatus and cuvette described above. The plot of $[\theta]_{222}$ versus GdnHCl concentration was fitted to a two-state transition equation, as described (Santoro and Bolen, 1988), to determine the free energy change upon unfolding in the absence of denaturant ($\Delta G_{U\text{-}F}^{H2O}$), the dependence of $\Delta G_{U\text{-}F}$ on GdnHCl concentration (m value), and the midpoint of unfolding ($C_m$).

### Near-UV CD

Near-UV CD spectra were acquired at 25ºC from 250 to 350 nm with the Jasco J-810 spectropolarimeter and a 5 mm path length quartz cuvette. The final protein concentration of the samples was 17 μM (0.2 mg $ml^{-1}$). The spectrum of wild-type Sso AcP in 0% (v/v) TFE, 10 mM Tris (pH 8), 25ºC, and the spectrum of the ΔN11 Sso AcP in 20% (v/v) TFE, 50 mM acetate buffer (pH 5.5), 25ºC, were recorded as the average of 40 and 30 separate scans, respectively. The spectra were blank subtracted and normalized to mean residue ellipticity.

### Enzymatic Activity Measurements

ΔN11 Sso AcP was initially unfolded in 6 M GdnHCl, 50 mM acetate (pH 5.5), 25ºC. The enzymatic activity was measured, as described (Plakoutsi et al., 2004), immediately after diluting GdnHCl out and adding different concentrations of TFE ranging from 0 to 46% (v/v). The final protein concentration was 0.0015 mg $ml^{-1}$. The enzymatic activity of the wt Sso AcP was evaluated in the presence of various TFE concentrations with a 30 min preincubation under the same conditions.

### H/D Exchange

Two nanomoles of Sso AcP were dissolved in 50 μl of 10 mM ammonium acetate (pH 7.5), 25ºC, for 5 min and then diluted into 450 μl of the same buffer prepared with $D_2O$ (Sigma, Milan, Italy). The experiment was repeated with deuterated 25% (v/v) in 10 mM ammonium acetate buffer prepared in $D_2O$. Sample aliquots for each time point were loaded on a reverse-phase cartridge (Phenomenex widepore C4, 4 × 2.0 mm) and eluted directly into the mass spectrometer ion source at a flow rate of 200 μl/min with a fast desalting gradient. All data were acquired in positive-ion mode with an LCMS system including a triple quadrupole Quarttro Micro mass spectrometer (Waters, Micromass, Manchester, United Kingdom) equipped with an electrospray source interfaced with an Agilent 1100 HPLC system. Source conditions (capillary, cone voltages, desolvation nitrogen flow) were optimized for Sso AcP and used in all the experiments. Mass spectra were acquired at time intervals ranging between 15 s and 60 min, over the mass range 700–1400 m/z at a scan rate of 3 s/scan. The instrument was calibrated with myoglobin from horse heart. Data processing parameters were the same for all spectra (background subtraction, smoothing, and peak centroid). The total number of exchangeable sites in any experiment was calculated by the recorded molecular mass of Sso AcP on the basis of the theoretical numbers of 99 exchangeable amide protons.

### Acknowledgments

This work was supported by grants from the European Union (Project HPRN-CT-2002-00241) and the Italian Ministero dell'Istruzione, Università e Ricerca (FIRB Projects no. RBAU015B47, RBNE01S29H, and RBNE03PX83; PRIN Projects no. 2003025755 and 2004050405).

Received: February 21, 2006
Revised: March 20, 2006
Accepted: March 27, 2006
Published: June 13, 2006

### References

Balguerie, A., Dos Reis, S., Ritter, C., Chaignepain, S., Coulary-Salin, B., Forge, V., Bathany, K., Lascu, I., Schmitter, J.M., Riek, R., and Saupe, S.J. (2003). Domain organization and structure-function relationship of the HET-s prion protein of *Podospora anserina*. EMBO J. *22*, 2071–2081.

Baxa, U., Taylor, K.L., Wall, J.S., Simon, M.N., Cheng, N., Wickner, R.B., and Steven, A.C. (2003). Architecture of Ure2p prion filaments: the N-terminal domains form a central core fiber. J. Biol. Chem. *278*, 43717–43727.

Bouchard, M., Zurdo, J., Nettleton, E.J., Dobson, C.M., and Robinson, C.V. (2000). Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. Protein Sci. *9*, 960–967.

Bousset, L., Thomson, N.H., Radford, S.E., and Melki, R. (2002). The yeast prion Ure2p retains its native alpha-helical conformation upon assembly into protein fibrils in vitro. EMBO J. *17*, 2903–2911.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507–511.

Camici, G., Manao, G., Cappugi, G., and Ramponi, G. (1976). A new synthesis of benzoyl phosphate: a substrate for acyl phosphatase assay. Experentia *32*, 535–536.

Chapman, M.R., Robinson, L.S., Pinkner, J.S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S.J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851–855.

Chien, P., Weissman, J.S., and DePace, A.H. (2004). Emerging principles of conformation-based prion inheritance. Annu. Rev. Biochem. *73*, 617–656.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding and disease. Annu. Rev. Biochem., in press.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C.M. (2003). Rationalisation of the effects of mutations on peptide and protein aggregation rates. Nature *424*, 805–808.

Corazza, A., Rosano, C., Pagano, K., Alverdi, V., Esposito, G., Capanni, C., Bemporad, F., Plakoutsi, G., Stefani, M., Chiti, F., et al. (2006). Structure, conformational stability and enzymatic properties of acylphosphatase from the hyperthermophile *Sulfolobus solfataricus*. Proteins *62*, 64–79.

Damaschun, G., Damaschun, H., Gast, K., and Zirwer, D. (1998). Denatured states of yeast phosphoglycerate kinase. Biochemistry (Mosc.) *63*, 259–275.

Dobson, C.M. (2003). Protein folding and misfolding. Nature *426*, 884–890.

DuBay, K.F., Pawar, A.P., Chiti, F., Zurdo, J., Dobson, C.M., and Vendruscolo, M. (2004). Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. J. Mol. Biol. *341*, 1317–1326.

Fernandez-Escamilla, A.M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat. Biotechnol. *22*, 1302–1306.

Glover, J.R., Kowal, A.S., Schirmer, E.C., Patino, M.M., Liu, J.J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*. Cell *89*, 811–819.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.

Kochneva-Pervukhova, N.V., Poznyakovski, A.I., Smirnov, V.N., and Ter-Avanesyan, M.D. (1998). C-terminal truncation of the Sup35 protein increases the frequency of de novo generation of a prion-based [PSI+] determinant in *Saccharomyces cerevisiae*. Curr. Genet. *34*, 146–151.

Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., et al. (2005). Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. USA *101*, 12491–12496.

Masison, D.C., and Wickner, R.B. (1995). Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. Science *270*, 93–95.

Pedersen, J.S., Christensen, G., and Otzen, D.E. (2004). Modulation of S6 fibrillation by unfolding rates and gatekeeper residues. J. Mol. Biol. *341*, 575–588.

Plakoutsi, G., Taddei, N., Stefani, M., and Chiti, F. (2004). Aggregation of the *Acylphosphatase* from *Sulfolobus solfataricus*: the folded and partially unfolded states can both be precursors for amyloid formation. J. Biol. Chem. *279*, 14111–14119.

Plakoutsi, G., Bemporad, F., Calamai, M., Taddei, N., Dobson, C.M., and Chiti, F. (2005). Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. J. Mol. Biol. *351*, 910–922.

Richardson, J.S., and Richardson, D.C. (2001). Natural β-sheet proteins use negative design to avoid edge to edge aggregation. Proc. Natl. Acad. Sci. USA *99*, 2754–2759.

Santoro, M.M., and Bolen, D.W. (1988). Unfolding free-energy changes determined by the linear extrapolation method. 1. Unfolding of phenyl-methanesulfonyl alpha-chymotrypsin using different denaturants. Biochemistry *27*, 8063–8068.

Stefani, M., and Dobson, C.M. (2003). Protein aggregation and aggregate toxicity: new insight into protein folding, misfolding diseases and biological evolution. J. Mol. Med. *81*, 678–699.

Tartaglia, G.G., Cavalli, A., Pellarin, R., and Caflisch, A. (2004). The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates. Protein Sci. *13*, 1939–1941.

Taylor, K.L., Cheng, N., Williams, R.W., Steven, A.C., and Wickner, R.B. (1999). Prion domain initiation of amyloid formation in vitro from native Ure2p. Science *283*, 1339–1343.

Uversky, V.N., and Fink, A.L. (2004). Conformational constraints for amyloid fibrillation: the importance of being unfolded. Biochim. Biophys. Acta *1698*, 131–153.

Walsh, D.M., and Selkoe, D.J. (2004). Oligomers on the brain: the emerging role of soluble protein aggregates in neurodegeneration. Protein Pept. Lett. *11*, 213–228.

Wilkins, D.K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Jones, J.A., and Smith, L.J. (1999). Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. Biochemistry *38*, 16424–16431.