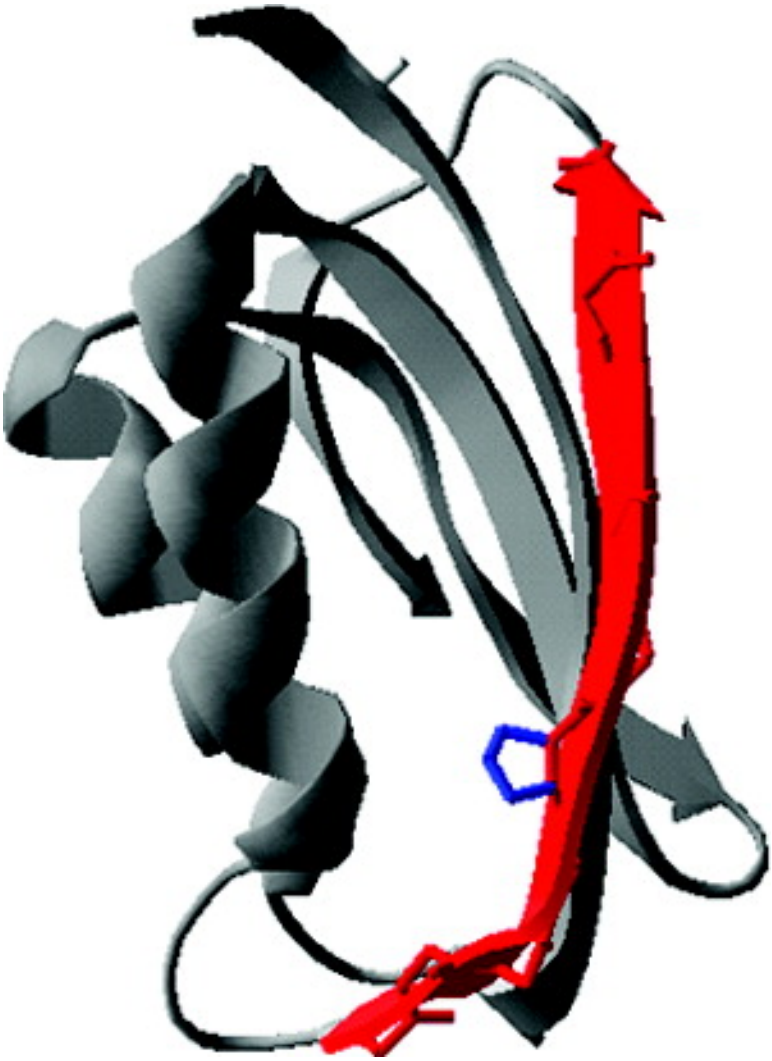
Article

The Degree of Structural Protection at the Edge β-Strands Determines the Pathway of Amyloid Formation in Globular Proteins

Gemma Soldi, Francesco Bemporad, and Fabrizio Chiti

J. Am. Chem. Soc., **2008**, 130 (13), 4295-4302 • DOI: 10.1021/ja076628s



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/12/2008

The Degree of Structural Protection at the Edge β-Strands Determines the Pathway of Amyloid Formation in Globular Proteins

Gemma Soldi, Francesco Bemporad, and Fabrizio Chiti*

Dipartimento di Scienze Biochimiche, Università di Firenze, Viale Morgagni 50, 50134 Firenze, Italy

Received September 3, 2007; E-mail: fabrizio.chiti@unifi.it

Abstract: The assembly of proteins into highly organized fibrillar aggregates is a key process in biology, biotechnology, and human disease. It has been shown that proteins retain a small, yet significant propensity to aggregate when they are folded into compact globular structures, and this may be physiologically relevant, particularly when considering that proteins spend most of their lifespan into such compact states. Proteins from the acylphosphatase-like structural family have been shown to aggregate via different mechanisms, with some members forming native-like aggregates as a first step of their aggregation process and others requiring unfolding as a first necessary step. Here we use the acylphosphatase from *Sulfolobus solfataricus* to show that assembly of folded protein molecules into native-like aggregates is prevented by single-point mutations that introduce structural protections within one of the most flexible region of the protein, the peripheral edge β-strand 4. The resulting mutants do not form native-like aggregates, but can still form thioflavin T-binding and β-structured oligomers, albeit more slowly than the wild-type protein. The kinetic data show that formation of the latter species proceeds via an alternative mechanism that is independent of the transient formation of native-like aggregates.

Introduction

The conversion of proteins and peptides from their soluble states into fibrillar aggregates, characterized by an extended cross-β structure and generally referred to as amyloid or amyloid-like fibrils, is a generic property and an essential behavior of polypeptide chains.[1] It is also an important process in human pathology,[2] biology,[3] and even biotechnology.[4,5] It is generally accepted that proteins adopting a well-defined three-dimensional fold need to unfold, at least partially, to aggregate into amyloid fibrils. Evidence for this hypothesis comes from a large body of experimental data.[2,6] Although such a "conformational change hypothesis" is undoubtedly correct, recent observations have suggested that in some cases the major conformational change associated with amyloid aggregation may not take place until after the initial aggregation step.

The formation of amyloid-like fibrils by insulin, for example, is preceded by an oligomerization phase in which a native-like content of α-helical structure is almost completely retained.[7] Aggregates with a morphology reminiscent of amyloid protofibrils and with a high content of β-structure appear later in the process. As a second example, the acylphosphatase from *Sulfolobus solfataricus* (Sso AcP) was found to aggregate under conditions in which the folded state is thermodynamically more stable than the dominant partially unfolded state.[8] Under such conditions, native Sso AcP has a flexibility higher than that under conditions in which aggregation does not occur, as indicated by hydrogen/deuterium exchange data.[9] The first event in the aggregation of Sso AcP under these conditions is the formation of oligomers that do not bind to thioflavin T (ThT) or Congo red (CR) and, remarkably, retain enzymatic activity.[10] This aggregation phase is more rapid than unfolding, indicating that it does not require a transition into a partially unfolded conformation across the major energy barrier for unfolding.[10] These native-like oligomers then undergo structural reorganization to form amyloid protofibrils that have extensive β-structure, bind ThT and CR, and are not enzymatically active.[10] Human β2-microglobulin and hen lysozyme were found to aggregate at neutral pH and physiological temperature (under mild agitation and in the presence of salts), into assemblies of protein molecules in their native-like states; these subsequently can convert, upon heating, into amyloid fibrils.[11] Related mechanisms of aggregation from native or native-like states have also been described for other proteins.[12−14] Remarkably, mechanisms

(1) Dobson, C. M. *Nature* **2003**, *426*, 884−890.
(2) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333−66.
(3) Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W. *Trends Biochem. Sci.* **2007**, *32* (5), 217−24.
(4) Hamada, D.; Yanagihara, I.; Tsumoto, K. *Trends Biotechnol.* **2004**, *22*, 93−7.
(5) Rajagopal, K.; Schneider, J. P. *Curr. Opin. Struct. Biol.* **2004**, *14*, 480−6.
(6) Uversky, V. N.; Fink, A. L. *Biochim. Biophys. Acta* **2004**, *1698*, 131−153.
(7) Bouchard, M.; Zurdo, J.; Nettleton, E. J.; Dobson, C. M.; Robinson, C. V. *Protein Sci.* **2000**, *10*, 1960−7.
(8) Plakoutsi, G.; Taddei, N.; Stefani, M.; Chiti, F. *J. Biol. Chem.* **2004**, *279*, 14111−14119.
(9) Plakoutsi, G.; Bemporad, F.; Monti, M.; Pagnozzi, D.; Pucci, P.; Chiti, F. *Structure* **2006**, *14*, 993−1001.
(10) Plakoutsi, G.; Bemporad, F.; Calamai, M.; Taddei, N.; Dobson, C. M.; Chiti, F. *J. Mol. Biol.* **2005**, *351*, 910−922.
(11) Sasahara, K.; Yagi, H.; Naiki, H.; Goto, Y. *Biochemistry* **2007**, *46*, 3286−93.
(12) Soldi, G.; Bemporad, F.; Torrassa, S.; Relini, A.; Ramazzotti, M.; Taddei, N.; Chiti, F. *Biophys. J.* **2005**, *89*, 4234−44.

of fibril formation in which the native state is retained in the fibrillar end products have been described for Ure2p and lithostatin.[15,16]

Aggregation processes that initiate from the folded states of globular proteins may be relevant in biology and pathology, considering that the majority of protein molecules spend most of their lifetime in a folded state. Interestingly, the yeast Hsp40 chaperone Ydj1 was found to inhibit fibril formation by the yeast prion Ure2p by binding preferentially to the globular domain of the native protein, suggesting that folded proteins or domains are species to be targeted by molecular chaperones in vivo to prevent their effective aggregation.[17] Perhaps the most compelling evidence that "native" aggregation processes represent a real possibility for globular proteins and a constant challenge for living organisms is shown by the finding that natural all-β proteins have developed strategies during evolution to prevent their assembly through a direct interaction of folded units.[18]

We have previously shown that two proteins from the acylphosphatase-like structural family, namely, the acylphosphatase from *Sulfolobus solfataricus* (Sso AcP) and the second acylphosphatase from *Drosophila melanogaster* (AcPDro2) have the ability to aggregate via the transient formation of oligomers in which the protein molecules retain native-like conformations.[10,12] By contrast, two other proteins from the same structural family, human muscle acylphosphatase (mAcP) and the N-terminal domain of HypF from *Escherichia coli* (HypF-N), have been shown to require unfolding to aggregate.[19,20] In particular, aggregation of HypF-N was shown to occur, even under conditions in which the native structure is the most populated species (>98%), through the self-assembly of partially folded molecules that form after the major free-energy barrier of unfolding and that are in rapid equilibrium with the native state.[20] The observation of such differences in aggregation pathways within the same structural family provides a unique opportunity to explore the determinants of either mechanism. To this purpose we have compared the sequences and structures of these four proteins. On the basis of the differences resulting from this comparative inspection, we have designed mutants of Sso AcP that are able to prevent the aggregation pathway involving the formation of native-like aggregates.

Results and Discussion

A Structural Comparison between mAcP, HypF-N, Sso AcP, and AcPDro2. The native structures of the four proteins considered here have all been solved with X-ray crystallography, facilitating the search of possible structural differences that can explain the different aggregation pathways of these four proteins.[21−24] Table 1 shows that in the two proteins that need to unfold prior to aggregation (mAcP and HypF-N), the number of residues forming loops or adopting an unfolded conformation is lower than in the two proteins aggregating via the assembly of native-like structures (Sso AcP and AcPDro2). Deletion of the unstructured 11 residues at the N-terminus of Sso AcP yields a mutant (ΔN11 Sso AcP) that no longer forms native-like aggregates.[9] This indicates that the extent of unfolded regions in the native state is an important determinant of the aggregation pathway and may suggest that the self-assembly of folded protein molecules is facilitated by the interaction of unstructured portions, such as loops or unstructured C- and N-terminal tails.

Table 1. Structural Comparison between Proteins from the Acylphosphatase-like Structural Family

variant	no. of residues with unordered secondary structure[a]	no. of protective features at the edge β strands[b]	$\Delta G_{F-U}^{H_2O}$ (kJ mol^{-1})[c]	surface hydrophobicity (kcal mol^{-1})[d]	no. of UWHB[e]
AcPDro2	40 (39.2%)	1−2, 1	-15.5 ± 0.7	25.7	5
Sso AcP	39 (38.6%)	1, 1	-48.7 ± 0.7	21.5	2
HypF-N	34 (37.4%)	3, 2	-29 ± 3	19.0	0
mAcP	32 (32.7%)	3, 3	-21.3 ± 2.3	28.3	9

[a] Calculated using *MolMol*. The numbers indicate residues that do not form α-helical or β-sheet structures, that is, residues adopting an unfolded conformation or forming loops. [b] Number of structural features able to protect edge β-strands from edge-to-edge aggregation, defined as described.[18] The table indicates, for each protein, the number of such features for both β-strand 4 and β-strand 5 (before and after the comma, respectively). [c] Obtained in 50 mM acetate buffer, pH 5.5, at $T = 25$ °C for AcPDro2, and at $T = 28$ °C for Sso AcP, HypF-N, and mAcP. [d] Calculated as the sum of the hydrophobicity values of all the amino acid residues of the protein, each weighted by the percent of solvent exposure (calculated using *Swiss-PdbViewer* and the hydrophobicity scale of the 20 amino acids reported in ref 32. [e] Number of backbone hydrogen bonds having three or fewer hydrophobic residues in close proximity. An hydrophobic residue is considered in close proximity if its β-carbon is contained within the space created by two spheres with 5 Å radii, centered on the α-carbons of the two residues forming the hydrogen bond. The ordering of proteins does not change if considering four (or less) and five (or less) hydrophobic residues. This counting has been made as described.[26]

Nevertheless, further inspection of the four proteins indicates other important differences. It has been suggested that folded proteins can self-assemble via a direct interaction of their edge β-strands, that is, the peripheral strands of a native β-sheet.[18] Indeed, edge β-strands have only one adjacent strand in the same sheet and can therefore interact with the edge β-strands of other folded protein molecules. Such strands are generally protected by a combination of structural strategies.[18] The incorporation of one of such structural features (inward pointing charges) on the edge β-strands of newly designed β-proteins has actually been shown to be an effective strategy to prevent aggregation of these proteins.[25] All four proteins analyzed here have one β-sheet in their folded conformation, with strands 4 and 5 representing the edge β-strands in all cases. mAcP and HypF-N, the two proteins that need to unfold prior to aggregation, have 2−3 protections on both edge β-strands (Table 1). By contrast, AcPDro2 and Sso AcP, the two proteins aggregating

(13) Pedersen, J. S.; Christensen, G.; Otzen, D. E. *J. Mol. Biol.* **2004**, *341*, 575−588.
(14) Jahn, T. R.; Parker, M. J.; Homans, S. W.; Radford, S. E. *Nat. Struct. Mol. Biol.* **2006**, *13*, 195−201.
(15) Bousset, L.; Thomson, N. H.; Radford, S. E.; Melki, R. *EMBO J.* **2002**, *17*, 2903−2911.
(16) Laurine, E.; Gregoire, C.; Fandrich, M.; Engemann, S.; Marchal, S.; Thion, L.; Mohr, M.; Monsarrat, B.; Michel, B.; Dobson, C. M.; Wanker, E.; Erard, M.; Verdier, J. M. *J. Biol. Chem.* **2003**, *278*, 51770−8.
(17) Lian, H. Y.; Zhang, H.; Zhang, Z. R.; Loovers, H. M.; Jones, G. W.; Rowling, P. J.; Itzhaki, L. S.; Zhou, J. M.; Perrett, S. *J. Biol. Chem.* **2007**, *282*, 11931−40.
(18) Richardson, J. S.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2754−2759.
(19) Chiti, F.; Taddei, N.; Bucciantini, M.; White, P.; Ramponi, G.; Dobson, C. M. *EMBO J.* **2000**, *19*, 1441−9.
(20) Marcon, G.; Plakoutsi, G.; Canale, C.; Relini, A.; Taddei, N.; Dobson, C. M.; Ramponi, G.; Chiti, F. *J. Mol. Biol.* **2005**, *347*, 323−35.
(21) Thunnissen, M. M.; Taddei, N.; Liguri, G.; Ramponi, G.; Nordlund, P. *Structure* **1997**, *5*, 69−79.
(22) Rosano, C.; Zuccotti, S.; Bucciantini, M.; Stefani, M.; Ramponi, G.; Bolognesi, M. *J. Mol. Biol.* **2002**, *321*, 785−96.
(23) Zuccotti, S.; Rosano, C.; Ramazzotti, M.; Degl'Innocenti, D.; Stefani, M.; Manao, G.; Bolognesi, M. *Acta Crystallogr., D: Biol. Crystallogr.* **2004**, *60*, 1177−9.
(24) Corazza, A.; Rosano, C.; Pagano, K.; Alverdi, V.; Esposito, G.; Capanni, C.; Bemporad, F.; Plakoutsi, G.; Stefani, M.; Chiti, F.; Zuccotti, S.; Bolognesi, M.; Viglino, P. *Proteins* **2006**, *62*, 64−79.
(25) Wang, W.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2760−5.

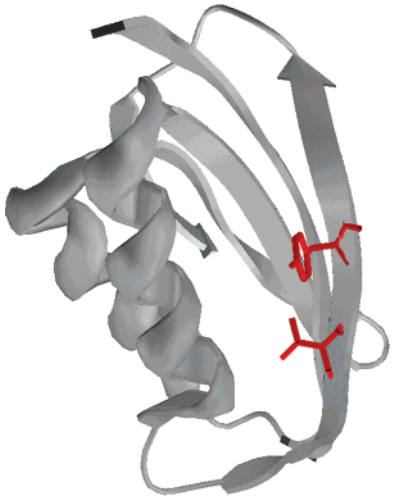


Figure 1. Three-dimensional structure of native Sso AcP as determined by X-ray crystallography (PDB code: 2BJD). Side chains of the residues mutated here (Val84 and Tyr86) are highlighted in red. The structure has been obtained using the *Swiss-PDBViewer* software.

via the assembly of native structures, have only 1−2 protective features (Table 1). A detailed description of the protective features existing for each of the strands in all four proteins is presented in the Supporting Information.

No other obvious differences were found between the four proteins that might explain their different aggregation behaviors. AcPDro2 and Sso AcP, the two proteins aggregating via the formation of native-like aggregates, have the lowest and highest conformational stabilities, respectively (Table 1). This rules out the importance of the native state stability in determining differences in the aggregation processes of the four proteins. Nor do the four proteins exhibit a clear trend for their content of solvent-exposed hydrophobic residues, with AcPDro2 and Sso AcP having intermediate values of superficial hydrophobicity, compared with the other two cases (Table 1). The four proteins were also compared for their number of underwrapped hydrogen bonds (UWHBs)—that is, backbone hydrogen bonds surrounded by three or fewer hydrophobic residues and thus vulnerable to water attack. A high number and a high local density of UWHBs have been proposed to dramatically increase the aggregation propensity of a folded protein and even to be a distinctive feature of folded proteins associated with pathological amyloid aggregation.[26] However, AcPDro2 and Sso AcP do not clearly exhibit more UWHBs than the other two proteins (Table 1).

From this structural comparison it appears that the lower number of protections on the edge β-strands of native Sso AcP and AcPDro2 could be a reason why such proteins can form native-like aggregates as a first step of the process of amyloid formation. Sso AcP has been studied in deeper detail than AcPDro2. Moreover, the formation of the native-like aggregates is faster than the subsequent reorganization into amyloid protofibrils, allowing accumulation and direct monitoring of native-like assemblies. β-strand 4, as opposed to the other edge β-strand 5, has been shown to play a key role in the formation of native-like aggregates of this protein.[9] For these reasons, we have chosen to focus our attention on the edge β-strand 4 of Sso AcP. Three mutants were produced and purified. Figure 1 shows the X-ray structure of Sso AcP along with the positions of the mutated residues, whereas Table 2 lists the mutants produced. The V84D and Y86E mutations generate an additional negative charge between β-strand 4 and α-helix 2 (inward pointing charge), whereas the V84P variant creates a distortion in the β-strand conformation. Both features have been proposed to inhibit the self-assembly of folded proteins via edge-to-edge interactions.[18] The former has also been shown to inhibit the aggregation of newly designed β-sheet proteins that would otherwise fail to remain soluble.[25]

Analysis of the Structure, Stability, and (Un)folding rate of the Sso AcP Mutants. The three variants have far-UV circular dichroism (far-UV CD) spectra indistinguishable from that of the wild-type protein, indicating that the mutants maintain a topology and secondary structure content similar to those of the wild-type protein (Figure 2a). The apparent hydrodynamic diameters of the mutants, as determined with dynamic light scattering (DLS), are within the experimental error range of the wild-type protein diameter, suggesting that the mutations do not induce changes in the size and compaction of the native structure (Figure 2b, Table 2). All variants have enzymatic activity, providing further indication on the persistence of a native topology upon mutation (Figure 2c). Interestingly, all three variants have significantly lower specific activity compared with that of the wild-type protein, suggesting that the mutations introduce subtle distortions at the site of mutations that are propagated up to the active site of the enzyme. This phenomenon has also been observed for mutations of mAcP that were not in close proximity to the active site.[27]

The conformational stabilities of the three variants have been investigated by means of equilibrium unfolding curves induced by guanidine hydrochloride (GdnHCl) (Figure 3a). All variants appear to be destabilized, relative to the wild-type (Figure 3a). Despite the observed destabilization, all variants appear to be folded in the absence of GdnHCl (Figure 3a). Analysis of the equilibrium curves with a two-state model[28] allows the determination of the values of free-energy change of folding in the absence of denaturant ($\Delta G_{F-U}^{H_2O}$) and of the GdnHCl concentra-

Table 2. Parameters of Sso AcP Variants in the Absence of TFE

variant	apparent hydrodynamic diameter (nm)[a]	specific activity (IU mg^{-1})[b]	C_M (M)[c]	$\Delta G_{F-U}^{H_2O}$ (kJ mol^{-1})[c]	$k_{F\rightarrow U}^{H_2O}$ (s^{-1})[d]	$k_{I\rightarrow F}^{H_2O}$ (s^{-1})[d]	$\Delta G_{F-I}^{H_2O}$ (kJ mol^{-1})[e]
wild-type	4.58 ± 0.09	893 ± 89	4.23 ± 0.2	−46.9 ± 2.2	2.3(±0.2) × 10^{-5}	4.6 ± 0.5	−31.5 ± 4.4
V84P	4.49 ± 0.03	208 ± 21	2.52 ± 0.2	−28.0 ± 2.2	2.3(±0.2) × 10^{-2}	4.6 ± 0.5	−13.6 ± 1.9
V84D	4.49 ± 0.03	495 ± 50	2.53 ± 0.2	−28.1 ± 2.2	2.2(±0.2) × 10^{-2}	4.3 ± 0.5	−13.6 ± 1.9
Y86E	4.46 ± 0.10	306 ± 31	3.43 ± 0.2	−38.1 ± 2.2	2.8(±0.3) × 10^{-4}	4.8 ± 0.5	−25.1 ± 3.5

[a] Calculated in 50 mM acetate buffer, pH 5.5, 25 °C, using DLS. [b] Calculated in 50 mM acetate buffer, pH 5.5, 25 °C, using 4 mM benzoylphosphate as a substrate. [c] Data obtained by GdnHCl unfolding curves at equilibrium, in 50 mM acetate buffer, pH 5.5, 37 °C. [d] Data obtained by a stopped-flow device in 50 mM acetate buffer, pH 5.5, 37 °C, in 0 M GdnHCl. [e] Calculated from $k_{F\rightarrow U}^{H_2O}$ and $k_{I\rightarrow F}^{H_2O}$ values using eq 2.

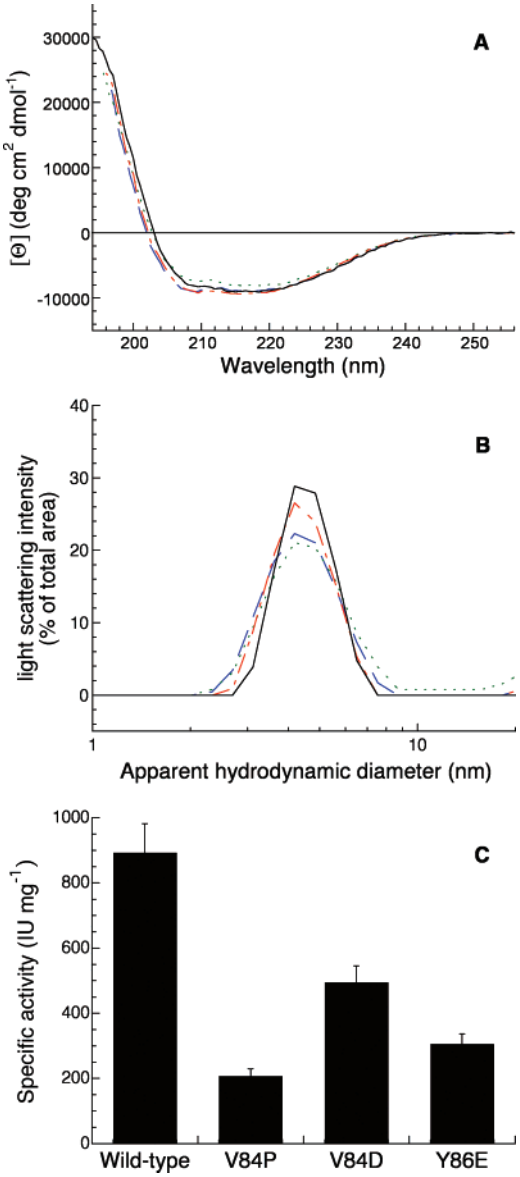


Figure 2. Analysis of the native states of Sso AcP variants. (a) Far-UV CD spectra for wild-type (black solid line), V84P (blue dashed line), V84D (green dotted line), and Y86E (red dashed/dotted line) Sso AcP. Spectra were recorded in 10 mM Tris/HCl, pH 8.0, 25 °C. (b) Size distribution of all variants (color code as in figure 2a), measured by DLS in 50 mM acetate buffer, pH 5.5, 25 °C. (c) Enzymatic activity of Sso AcP mutants, measured in 50 mM acetate buffer, pH 5.5, 25 °C, using 4 mM benzoylphosphate as a substrate.

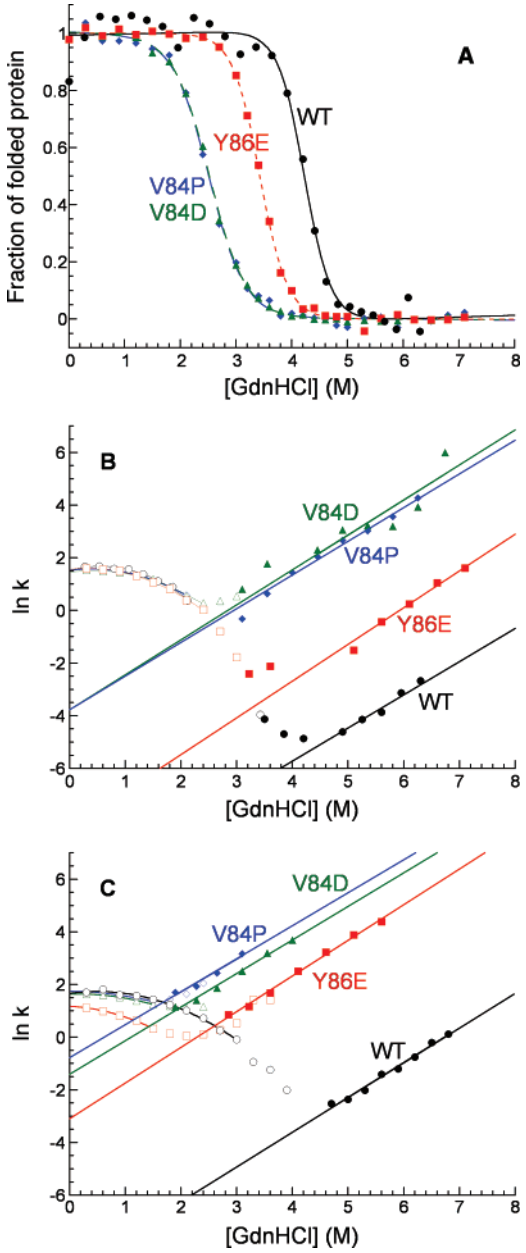


Figure 3. Analysis of the conformational stability and folding and unfolding rates of Sso AcP variants. (a) Equilibrium GdnHCl unfolding curves in 50 mM acetate buffer, pH 5.5, 37 °C, for wild-type (black circles), V84P (blue diamonds), V84D (green triangles), and Y86E (red squares) Sso AcP. Lines represent the best fits to the two-state equation.[28] (b and c) Rate constants for folding from the partially folded state (open symbols) and unfolding (filled symbols) as a function of GdnHCl concentration, determined in 50 mM acetate buffer, pH 5.5, 37 °C, for wild-type (black circles), V84P (blue diamonds), V84D (green triangles) and Y86E (red squares) Sso AcP. The curves were obtained in the absence (b) or in the presence of 3% TFE (c) with the exception of Y86E for which 10% TFE was used (c). In all cases, folding and unfolding data have been fitted to second-order polynomial and linear functions, respectively (solid lines), to obtain folding and unfolding rates in the absence of GdnHCl.

tion at which 50% of the protein is unfolded (C_M) for all variants (Table 2).

As previously reported, the folding process of wild-type Sso AcP is complex, with three kinetic phases detected during folding at low GdnHCl concentration and one phase observed for unfolding at high GdnHCl concentration.[29] Figure 3b reports the dependence of the rate constant for unfolding ($k_{F\rightarrow U}$, filled symbols) and of the rate constant for the conversion of the partially folded state into the fully native state ($k_{I\rightarrow F}$, empty symbols) on GdnHCl concentration for all variants. The Y86E variant appears to have ln($k_{F\rightarrow U}$) values 1 order of magnitude higher than those of the wild-type protein, when these are compared at the same denaturant concentration; the V84D and V84P variants unfold even more rapidly, 3 orders of magnitude faster than the wild-type (Figure 3b). The dependence of ln($k_{F\rightarrow U}$) on GdnHCl concentration appears to be linear and similar

(26) Fernandez, A.; Kardos, J.; Scott, L. R.; Goto, Y.; Berry, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6446−51.
(27) Taddei, N.; Modesti, A.; Bucciantini, M.; Stefani, M.; Magherini, F.; Vecchi, M.; Raugei, G.; Ramponi, G. *FEBS Lett.* **1995**, *362*, 175−9.
(28) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063−8068.
(29) Bemporad, F.; Capanni, C.; Calamai, M.; Tutino, M. L.; Stefani, M.; Chiti, F. *Biochemistry* **2004**, *43*, 9116−26.

Table 3. Parameters of Sso AcP Variants in the Presence of the Indicated TFE Concentrations

variant	$k_{F\rightarrow U}^{H_2O}$ (s^{-1})[a]	$k_{I\rightarrow F}^{H_2O}$ (s^{-1})[a]	$\Delta G_{F-I}^{H_2O}$ (kJ mol^{-1})[b]	k_2^{ThT} (s^{-1})
wild-type (20% TFE)	$4.1(\pm0.4) \times 10^{-5}$	$7.4(\pm 0.7) \times 10^{-2}$	-19.3 ± 2.7	$3.7(\pm0.4) \times 10^{-3}$
wild-type (3% TFE)	$1.4(\pm0.1) \times 10^{-4}$	5.3 ± 0.6	-27.1 ± 3.8	n.d.
V84P (3% TFE)	$4.6(\pm0.5) \times 10^{-1}$	5.7 ± 0.5	-6.5 ± 0.9	$6.1(\pm0.6) \times 10^{-4}$
V84D (3% TFE)	$2.5(\pm0.3) \times 10^{-1}$	5.1 ± 0.5	-7.8 ± 1.1	$8.5(\pm 0.8) \times 10^{-4}$
Y86E (10% TFE)	$4.5(\pm0.5) \times 10^{-2}$	3.3 ± 0.3	-11.0 ± 1.6	$6.0(\pm0.6) \times 10^{-4}$

[a] Data obtained in 50 mM acetate buffer, pH 5.5, 37 °C, in the absence of GdnHCl and in the presence of the TFE concentrations reported by each variant. The data for the wild-type in 20% TFE were reported previously.[8] [b] Calculated from $k_{F\rightarrow U}^{H_2O}$ and $k_{I\rightarrow F}^{H_2O}$ values using eq 2.

in the four proteins (Figure 3b). For each variant, the linear extrapolation of the $\ln(k_{F\rightarrow U})$ data to 0 M GdnHCl yields the $\ln(k_{F\rightarrow U})$ value in the absence of GdnHCl (Table 2). The $\ln(k_{I\rightarrow F})$ values of all three mutants are similar to those of the wild-type protein (Figure 3b, empty symbols). The dependence of $\ln(k_{I\rightarrow F})$ on GdnHCl concentration is not linear in this case, and a polynomial function was used to extrapolate the $\ln(k_{I\rightarrow F})$ values in 0 M GdnHCl (Table 2).

Following the measurement of the $k_{I\rightarrow F}$ and $k_{F\rightarrow U}$ values in the absence of GdnHCl, the use of well-established thermodynamic laws (eq 2) enables the determination, for each variant, of the free-energy change resulting from the conversion of the partially folded state into the native structure, $\Delta G_{F-I}^{H_2O}$ (Table 2).

As described above, the three mutated forms of Sso AcP are destabilized relative to the wild-type counterpart (Figure 3a, Table 2). Under the destabilizing conditions that promote the aggregation of wild-type Sso AcP—typically 20% (v/v) trifluoroethanol (TFE)—the three mutants are either unfolded or heavily destabilized (data not shown). For this reason, the aggregation processes of the mutants were compared in the presence of different TFE concentrations: 20% (v/v) TFE for wild-type, 10% (v/v) TFE for Y86E, and 3% (v/v) TFE for V84D and V84D. The $k_{I\rightarrow F}$, $k_{F\rightarrow U}$, and $\Delta G_{F-I}^{H_2O}$ values were re-determined for all the three mutants under the new conditions using the stopped-flow apparatus (Figure 3c). All the re-determined values are reported in Table 3, which also reports the same values for the wild-type protein determined in both 3 and 20% (v/v) TFE. Under their respective conditions of TFE concentration, all variants are folded prior to aggregation ($k_{I\rightarrow F} \gg k_{F\rightarrow U}$ and $\Delta G_{F-I}^{H_2O} \ll 0$ in 0 M GdnHCl). This provides an opportunity to assess whether the three mutations protect Sso AcP against the formation of native-like assemblies.

Analysis of the Aggregation Properties of the Sso AcP Mutants. The first phase of Sso AcP aggregation, the aggregation of the native state into native-like, enzymatically active assemblies, is very rapid and is detectable as a rapid change of the far-UV CD spectrum.[10] Figure 4a compares the CD spectrum of native wild-type Sso AcP, recorded in the absence of TFE, and that recorded immediately after incubation in 20% (v/v) TFE. Such a change in the CD spectrum is not observed for any of the mutants, following the addition of TFE (Figure 4b–d). The maximum difference of mean residue ellipticity between the native protein and the native-like assemblies appears to occur at 208 nm (Figure 4a). The time course of mean residue ellipticity at this wavelength produces a rapid, well-defined kinetic profile for wild-type Sso AcP in 20% TFE, whereas the CD signal does not appear to change for the mutants on the same time-scale (Figure 5a). The first phase of Sso AcP aggregation can also be monitored with light scattering.[10] Similarly to the results obtained with far-UV CD, a rapid increase of light scattering intensity is observed only for the wild-type protein (Figure 5b).

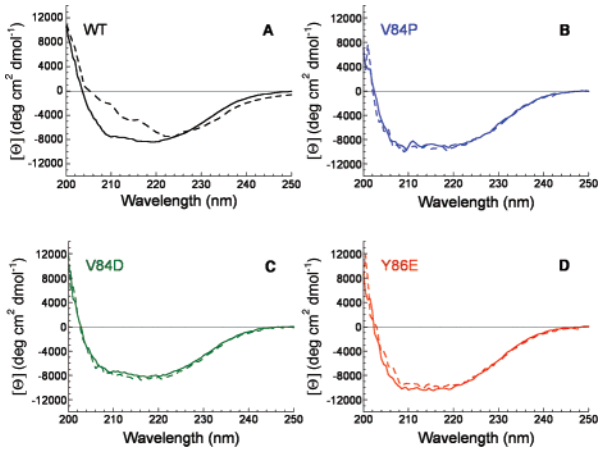


Figure 4. Secondary structure of Sso AcP variants under conditions promoting aggregation. (a–d) Far-UV CD spectra obtained for all variants under conditions in which the native state is stable, i.e., in 10 mM Tris/HCl, pH 8.0, 25 °C (continuous lines), and immediately after incubation in conditions promoting aggregation, that is 0.4 mg mL^{-1} protein, 50 mM acetate buffer, pH 5.5, 3–20% (v/v) TFE, 25 °C (dashed lines). TFE concentration was 20% for wild-type (a), 3% for V84P (b) and V84D (c), and 10% for Y86E (d).

The second phase of Sso AcP aggregation, the conversion of the native-like oligomers into amyloid-like protofibrils, can be monitored effectively with ThT fluorescence, as the protofibrils, unlike the native-like oligomers, bind to ThT.[9,10] The incubation of wild-type and three mutant Sso AcP in the presence of their respective TFE concentrations yields an increase of ThT fluorescence in all cases, with the wild-type protein exhibiting the most rapid increase (Figure 5c). To assess further that all mutants form aggregates and to exclude that the increase of ThT fluorescence arises from a potential binding of ThT with partially unfolded states of the mutants, we have analyzed all samples with DLS, after the ThT fluorescence had reached an apparent plateau. The size distributions of all samples clearly show the presence of aggregates (Figure 5d). Since the intensity of the scattered light is proportional to the square of the light scattering particles in solution, the size distributions obtained with DLS are enriched with the large clusters of protofibrils. Indeed, it has been previously shown with both DLS and transmission electron microscopy that Sso AcP protofibrils tend to associate and form large clusters.[8]

The observed kinetic profiles of far-UV CD, light scattering, and ThT fluorescence have several implications. First, the lower TFE concentrations used here for the three mutants relative to the wild-type protein do not inhibit aggregation per se, as an increase of ThT fluorescence is observed in all cases. Second, the mutants aggregate into ThT-binding protofibrils, but do not form transiently native-like assemblies. In other words, the

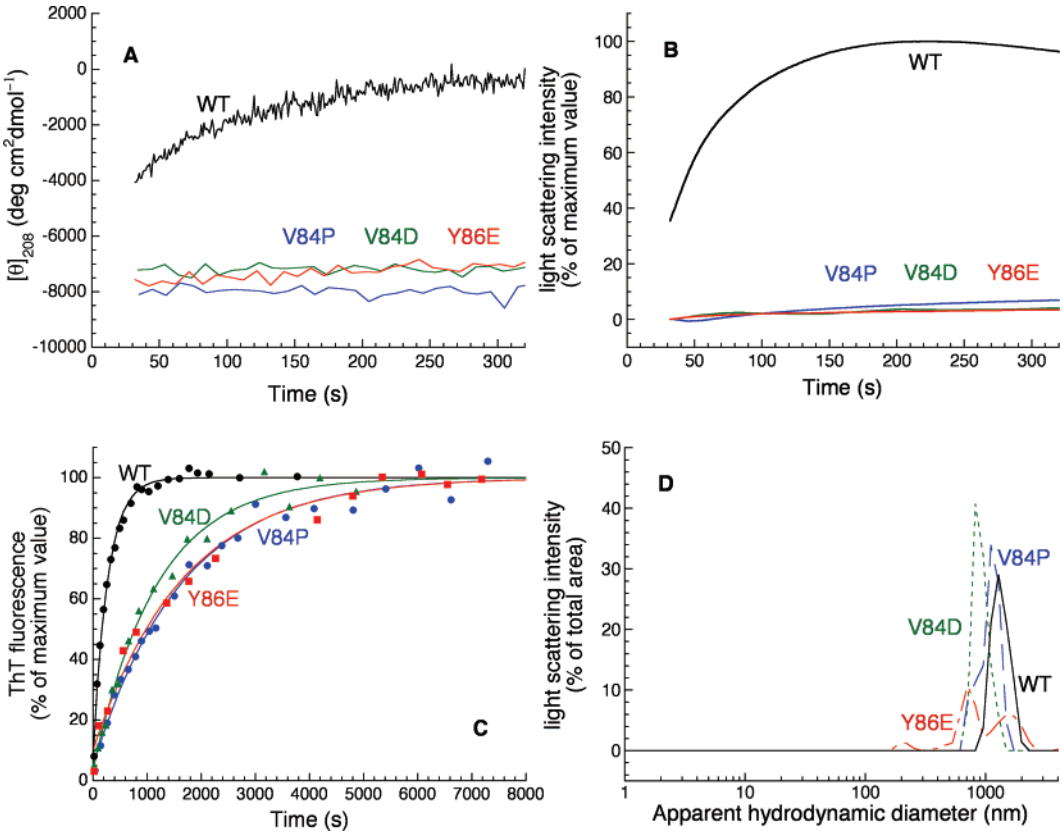


Figure 5. Aggregation of Sso AcP variants. (a and b) First aggregation phase (assembly of the native state into native-like aggregates) for wild-type (black), V84P (blue), V84D (green), and Y86E (red) Sso AcP, monitored by mean residue ellipticity at 208 nm (a) and light scattering intensity (b). (c) Second aggregation phase (conversion of native-like aggregates into amyloid-like protofibrils) for Sso AcP variants (color code as in figure 5a) monitored by ThT fluorescence. (d) Size distribution of all variants (color code as in figure 5a), measured by DLS after ThT fluorescence plateau has been reached. In all four panels, conditions were 0.4 mg mL^{-1} protein, 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of 3% TFE (V84P and V84D), 10% TFE (Y86E), and 20% TFE (wild-type).

conditions used here promote the aggregation of the mutants, but this occurs via a different pathway, as protofibril formation is not preceded by the accumulation of native-like assemblies. While the first and second phases of aggregation for wild-type Sso AcP in 20% TFE are both faster than unfolding under the same conditions, the ThT fluorescence kinetics of the mutants in 3 or 10% TFE are much slower than unfolding under the same conditions (compare $k_{F\rightarrow U}^{H_2O}$ and k_2^{ThT} values in Table 3). This suggests that aggregation of the mutants may occur via unfolding. Third, despite the fact that all mutants in their respective TFE concentrations are remarkably less stable than the wild-type protein in 20% TFE (Table 3), the formation of ThT-binding protofibrils is slower for the mutants. The rates of the first and second phases of aggregation were previously found to correlate inversely with the conformational stability of the folded state in a group of Sso AcP variants.[9] The trend observed here is opposite, with the destabilized variants forming the protofibrils more slowly and not forming the native-like aggregates at all. This rules out the possibility that the changes of aggregation behavior observed for the mutants arise from the different conditions employed.

Conclusions

Amyloid fibril formation is a complex process in which the initial steps consist of the formation of aggregates that can reorganize or assemble further into other species, until critical nuclei are formed to trigger fibril elongation.[2] The detailed investigation of the early steps of aggregation by folded proteins reveals that such proteins can aggregate either following their unfolding into fully or partially unfolded states or via the assembly of protein molecules in their native or native-like state that subsequently can reorganize into β-sheet-containing species. Both pathways have been observed within the acylphosphatase-like structural family, providing a very favorable opportunity to explore the determinants of either process through a comparative analysis of the structural characteristics of the proteins involved.

The two members following the "native-like aggregation pathway" (Sso AcP and AcPDro2) differ from the two proteins aggregating via the unfolding route (mAcP and HypF-N) in their higher content of residues in an unfolded structure and in the paucity of protections present at the edge strands of the native β-sheet. A protein engineering approach has shown that the unstructured N-terminus and the edge β-strand 4 play a key role in the formation of the native-like aggregates by Sso AcP.[9] It was previously shown that it is possible to inhibit the formation of the native-like aggregates of Sso AcP by decreasing the unfolded content of the native state, i.e., by removing the unstructured N-terminus.[9] This work has shown that the same effect can be achieved by introducing structural protections at edge β-strand 4. This results in the suppression or retardation of the formation of the later protofibrils, as these have to form through an alternative pathway that probably proceeds following denaturation. It will be important to extend these approaches to other structural families, in an attempt to more broadly

delineate the factors governing the early steps of a process as fundamental as amyloid formation.

Methods

Materials, Protein Purification, and Site-Directed Mutagenesis. Thioflavin T (ThT), trifluoroethanol (TFE), and guanidine hydrochloride (GdnHCl) were purchased from Sigma-Aldrich. Benzoylphosphate was synthesized and purified as described.[30] The Sso AcP gene sequence was cloned in the pGEX-2T plasmid (Amersham Biosciences, Piscataway, NJ) and expressed in DH5α *E. coli* cells as a protein fused to glutathione-S-transferase (GST). The resulting protein was purified and separated from GST according to the plasmid manufacturer. The sequence of the purified protein is GS[1]MKKWSDTEVFEMLKRMYARVYGLVQGVGFRKFVQIHAIRLGIKGYAKNLPDGSVEVVAEGYEEALSKLLERIKQGPPAAEVEKVDYSFSEYKGEFEDFETY. The Gly-Ser dipeptide at the N-terminus results from the cloning in pGEX-2T. The following Met residue is residue no. 1. Protein purity was checked by sodium dodecyl sulfate polyacrylamide gel electrophoresis and electrospray ionization mass spectrometry. The mutated genes of Sso AcP were obtained using the QuickChange site-directed mutagenesis kit from Stratagene (La Jolla, CA). The presence of the desired mutation was assessed by DNA sequencing of the entire gene. All mutational variants were expressed in *E. coli* XL-1 Blue cells and purified similarly to the wild-type protein. All Sso AcP variants were maintained in 10 mM Tris-HCl, pH 8.0, and stored at −20 °C. The protein concentration was determined spectrophotometrically using an ϵ_{280} value of 1.37 mL mg^{-1} cm^{-1} for all variants except Y86E, for which an ϵ_{280} value of 1.24 was used. Before starting each experiment, each protein solution was centrifuged at 16000*g* for 3−5 min and filtered using 0.22 μm Millex-GV filters (Millipore Corp., Billerica, MA).

Enzymatic Activity. Enzymatic activity was tested using a final protein concentration of 0.5 μg mL^{-1} and 4 mM benzoylphosphate as a substrate, as described.[31] Experimental conditions were 50 mM acetate buffer, pH 5.5, 25 °C. A Perkin-Elmer λ 4 B UV−vis spectrophotometer was used (Wellesley, MA). The noncatalyzed spontaneous hydrolysis of benzoylphosphate was subtracted from all measurements.

Far-UV Circular Dichroism (CD). Far-UV CD spectra were first acquired, for all variants, under conditions in which Sso AcP has a native structure, i.e., in 10 mM Tris/HCl buffer, pH 8.0, 25 °C, at a protein concentration of 0.2 mg mL^{-1}, using a 1 mm path length quartz cuvette and a Jasco J-810 spectropolarimeter (Tokyo, Japan) equipped with a thermostatted cell holder attached to a Thermo Haake C25P water bath (Karlsruhe, Germany). The spectra result from 10 accumulations. Far-UV CD spectra were also acquired under conditions promoting aggregation, in 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of the TFE concentration reported in the text, at a protein concentration of 0.4 mg mL^{-1}. In this case, the spectra were acquired immediately after the addition of TFE.

In another experimental set, the far-UV CD signal at 208 nm was followed to monitor the first phase of aggregation. Samples were prepared by diluting protein stock solution in 50 mM acetate, pH 5.5, 25 °C, 3−20% (v/v) TFE, at a final protein concentration of 0.4 mg mL^{-1}. The same cuvette and CD apparatus as those described above were used.

Dynamic Light Scattering (DLS). DLS measurements were performed at a protein concentration of 0.3 mg mL^{-1}, in 50 mM acetate buffer, pH 5.5, 25 °C. In another experimental set, the measurements were performed at 0.4 mg mL^{-1} protein concentration, after 2−3 h incubation in 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of 3−20% (v/v) TFE. Disposable plastic cuvettes with a 10 mm light path (Plastibrand, Brand GMBH, Wertheim, Germany) were used. Data were obtained with a Zetasizer Nano S DLS instrument (Malvern Instruments, Ltd., Worcestershire, U.K.) by setting the appropriate viscosity and refractive index parameters for each solution and keeping the temperature at 25 °C during the measurements by the means of a Peltier thermostating system.

Equilibrium Unfolding. The conformational stability of each Sso AcP variant was studied at equilibrium by preparing 25−30 samples containing 0.02 mg mL^{-1} protein, 50 mM acetate buffer, pH 5.5, and various GdnHCl concentrations ranging from 0 to 8 M. The samples were left to equilibrate for 10 min at 37 °C. The ellipticity at 222 nm was acquired using the CD apparatus described above. Ellipticity values were plotted versus GdnHCl concentration. The resulting plot was fitted to the equation edited by Santoro and Bolen[28] to obtain C_M and m values. The $\Delta G_{F-U}^{H_2O}$ value was calculated according to the equation $\Delta G_{F-U}^{H_2O} = C_M \cdot \underline{m}$, where $\underline{m}$ is the average m value over a set of ca. 30 Sso AcP variants (data not shown).

Stopped-Flow Kinetics. Folding and unfolding kinetics of Sso AcP and its mutants were studied with a Bio-logic SFM-3 stopped-flow device coupled to a fluorescence detection system (Claix, France) and thermostatted at 37 °C with a Neslab RTE-200 water-circulating bath (Newington, NH). The excitation wavelength was 280 nm, and the fluorescence emitted above 320 nm was monitored using a cutoff filter. Unfolding kinetics were monitored by preparing a protein sample at 0.4 mg mL^{-1} in 10 mM Tris/HCl, pH 8.0, and by a 20-fold dilution into solutions containing 50 mM acetate buffer, pH 5.5, and various GdnHCl concentrations. In refolding experiments all protein variants, except the wild-type, were initially unfolded in 5 M GdnHCl, 10 mM Tris/HCl pH 8, whereas the wild-type was unfolded in 6 M GdnHCl, 10 mM Tris/HCl pH 8. Reactions were started by a 20-fold dilution into solutions containing 50 mM acetate buffer, pH 5.5, and small GdnHCl concentrations. Unfolding and refolding kinetics were monitored at a final protein concentration of 0.02 mg mL^{-1}, with final GdnHCl concentrations ranging from 1.90 to 7.60 M for unfolding and from 0.30 to 3.90 M for refolding. These experiments were performed both in the absence and in the presence of 3% (v/v) TFE for all variants, except Y86E, for which a 10% (v/v) TFE concentration was also used. The dead time was 10.44 ms. The recorded traces were fitted to exponential functions of the following form

$$y(t) = q + \sum_{i=1}^{n} A_i \exp(-k_i t) \quad (1)$$

where $y(t)$ is the fluorescence at time t, q is the fluorescence value at equilibrium, A_i and k_i are the amplitude and rate constant of the ith phase, respectively, and n is equal to 1 and 2 for unfolding and folding, respectively. The natural logarithm of k_i was plotted versus GdnHCl concentration for each variant and for each TFE concentration used. In each case, the unfolding and refolding data from these plots were analyzed separately with a best-fitting procedure using linear and second-order polynomial functions, respectively. This allowed the values of folding rate constant from a partially folded state ($k_{I\rightarrow F}^{H_2O}$) and unfolding rate constant ($k_{F\rightarrow U}^{H_2O}$) to be extrapolated in the absence of GdnHCl from the lines of best fit. The free-energy change for the conversion of the partially folded state into the native state ($\Delta G_{F-I}^{H_2O}$) was determined using

$$\Delta G_{F-I}^{H_2O} = -RT \ln(k_{I\rightarrow F}^{H_2O}/k_{F\rightarrow U}^{H_2O}) \quad (2)$$

ThT Assay. All variants of Sso AcP were incubated at concentrations of 0.4 mg mL^{-1} in 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of 3, 10, or 20% (v/v) TFE. At regular times, 60 μL aliquots of each sample were added to 440 μL of a solution containing 25 μM ThT, 25 mM phosphate buffer, pH 6.0. The steady-state fluorescence values of the resulting samples were measured at 25 °C using a 2 × 10 mm path

(30) Camici, G.; Manao, G.; Cappugi, G.; Ramponi, G. *Experientia* **1976**, *32*, 535−6.
(31) Ramponi, G.; Treves, C.; Guerritore, A. *Arch. Biochem. Biophys.* **1966**, *115*, 129−35.
(32) Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. *Nature* **2003**, *424*, 805−808.

length quartz cuvette and a Perkin-Elmer LS 55 spectrofluorimeter (Wellesley, MA) equipped with a thermostatted cell holder attached to a Haake F8 water bath (Karlsruhe, Germany). The excitation and emission wavelengths were 440 and 485 nm, respectively. All measured fluorescence values are given after subtracting the fluorescence measured in the absence of protein and normalized so that the final fluorescence at the endpoint of the kinetic trace was 100%. ThT fluorescence was plotted versus time and fitted to

$$F(t) = F_{eq} + A \exp(-k_2^{ThT} t) \quad (3)$$

where $F(t)$ is the ThT fluorescence at time t, F_{eq} is the maximum ThT fluorescence obtained at the end of the observed exponential phase, A is the amplitude of the fluorescence change, and k_{obs} is the apparent rate constant.

Acknowledgment. This work was supported by grants from the EC (Project LSHM-CT-2006-037525), the Italian MIUR (FIRB Projects Nos. RBNE03PX83 and RBIN04PWNC, PRIN Projects Nos. 2005027330 and 2006058958), and the EMBO Young Investigator Programme. F.B. was supported by a studentship from the Accademia dei Lincei (Dott. Giuseppe Guelfi 2007).

Supporting Information Available: Detailed description of the protective features existing for each of the peripheral β-strands in all four proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

JA076628S