Studying the Folding Process of the Acylphosphatase from *Sulfolobus solfataricus*. A Comparative Analysis with Other Proteins from the Same Superfamily†

Francesco Bemporad,‡ Cristina Capanni,‡ Martino Calamai,§ Maria Luisa Tutino,⊥ Massimo Stefani,‡ and Fabrizio Chiti*,‡

Dipartimento di Scienze Biochimiche, Università di Firenze, Viale Morgagni 50, 50134 Firenze, Italy, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K., and Dipartimento di Chimica Organica e Biochimica, Università di Napoli, Edificio MB, Via Cinthia, Monte S. Angelo, 80100 Napoli, Italy

Received November 17, 2003; Revised Manuscript Received March 22, 2004

ABSTRACT: The folding process of the acylphosphatase from *Sulfolobus solfataricus* (Sso AcP) has been followed, starting from the fully unfolded state, using a variety of spectroscopic probes, including intrinsic fluorescence, circular dichroism, and ANS binding. The results indicate that an ensemble of partially folded or misfolded species form rapidly on the submillisecond time scale after initiation of folding. This conformational ensemble produces a pronounced downward curvature in the Chevron plot, appears to possess a content of secondary structure similar to that of the native state, as revealed by far-UV circular dichroism, and appears to have surface-exposed hydrophobic clusters, as indicated by the ability of this ensemble to bind to 8-anilino-1-naphthalenesulfonic acid (ANS). Sso AcP folds from this conformational state with a rate constant of ca. 5 s^{-1} at pH 5.5 and 37 °C. A minor slow exponential phase detected during folding (rate constant of 0.2 s^{-1} under these conditions) is accelerated by cyclophilin A and is absent in a mutant of Sso AcP in which alanine replaces the proline residue at position 50. This indicates that for a lower fraction of Sso AcP molecules the folding process is rate-limited by the cis−trans isomerism of the peptide bond preceding Pro50. A comparative analysis with four other homologous proteins from the acylphosphatase superfamily shows that sequence hydrophobicity is an important determinant of the conformational stability of partially folded states that may accumulate during folding of a protein. A low net charge and a high propensity to form α-helical structure also emerge as possibly important determinants of the stability of partially folded states. A significant correlation is also observed between folding rate and hydrophobic content of the sequence within this superfamily, lending support to the idea that sequence hydrophobicity, in addition to relative contact order and conformational stability of the native state, is a key determinant of folding rate.

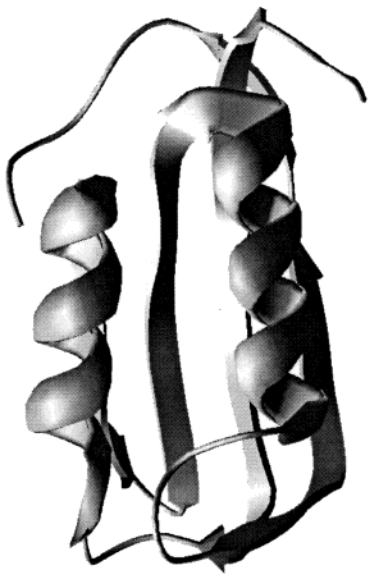
The recent sequencing of the genomes of a number of living organisms have enriched our databases with countless open reading frames potentially able to code for an equivalent number of amino acid sequences. To fully exploit the impressive mass of sequence information so far achieved we have to gain a full knowledge of the protein folding code, that is, that complex set of rules that determine the conversion of linear and unstructured amino acid sequences into functional three-dimensional structures. Considerable progress in our understanding of how proteins fold has come from comparative analyses of different protein systems that have allowed principles of general applicability to be obtained. The comparison between sequences of proteins known to exist in unstructured states under physiological conditions with those known to fold into a globular structure has first of all allowed the prediction to be made on whether an amino acid sequence adopts a globular structure or remains natively unfolded (*1*). Using published data on the conformational stabilities and folding/unfolding rates of a large number of proteins and mutated variants, Serrano and co-workers have achieved the goal of predicting, with a fair level of accuracy, the conformational stability of the native state of a protein of choice, the change of stability upon mutation, and even mechanisms of protein folding (*2*). Structural determination of the folding transition states for structurally related proteins has allowed topology to be identified as an important determinant in folding mechanisms (*3*−*6*). The relative contact order of a protein, that is, the average distance between residues forming an interaction in the native state, is an established important determinant of folding rate for single-domain proteins (*7*). Although in these and other cases the achieved possibility of predicting characteristics of protein folding has to be considered a starting point for further refinement analyses, rather than a definitive attainment, it is becoming clear that the comparison between closely related proteins and the statistical analyses of structurally diverse proteins is an important source of information for improving our current knowledge of protein folding.

† The Dipartimento di Scienze Biochimiche in Florence is supported by the Italian MIUR (Progetto FIRB "Folding di proteine: l'altra metà del codice genetico" and L. 449/97, Progetto "Strutture ed interazioni molecolari di prodotti genici", Settore "Genomica Funzionale") and the "Cassa di Risparmio di Firenze".

* Corresponding author. E-mail address: fchiti@scibio.unifi.it.
‡ Università di Firenze.
§ University of Cambridge.
⊥ Università di Napoli.

10.1021/bi030238a CCC: $27.50
Published on Web 06/19/2004

a)



b)

```
mAcP      -------STAQSLKSVDYEVFGRVQGVCFRMYTEDEARKIGVVGWVKNTSKGTVTGQVQGPEDKVNSMKSWLSKVGSPS-SRIDRTNFSNEKTISKLEYSNFSIRY
ctAcP     -------AEGNTLISVDYEIFGKVQGVFFRKHTQAEGKKLGLVGWVQNTDRGTVQGQLQGPISKVRHMQEWLETRGSPK-SHIDKANFNNEKVILKLDYSDFQIVK
Sso AcP   MKKWSDTEVFEMLKRMYARVYGLVQGVGFRKFVQIHAIRLGIKGYAKNLPDGSVEVVAEGYEEALSKLLERIKQ-GPPA-AEVEKVDYS--FSEYKGEFEDFETY
HypF-N    -------MAKNTSCGVQLRIRGKVQGVGFRPFVWQLAQQLNLHGDVCNDGDG-VEVRLRED---PETFLVQLY-QHCPPLARIDSVEREPFIW--SQLPTEFTIR
```

FIGURE 1: (a) Three-dimensional structure of Sso AcP obtained using the modeling procedure of the SwissPDBViewer software and the structure of ctAcP as a reference template and (b) sequences of mAcP, ctAcP, Sso AcP, and HypF-N aligned on the basis of sequence identities. Residues conserved in all four proteins are underlined and shown in bold.

In this work, we investigate the folding mechanism of a newly cloned protein, the acylphosphatase from the hyperthermophilic archaeon *Sulfolobus solfataricus* (Sso AcP).[1] The investigation of this acylphosphatase adds to our previous studies on human muscle acylphosphatase (mAcP) (*3*, *8*), human common-type acylphosphatase (ctAcP) (*9*), and the N-terminal acylphosphatase domain of the *Escherichia coli* HypF (HypF-N) (*10*). These proteins share sufficiently high sequence similarities (all pairs have sequence identities higher than 22%) to suggest a common origin from a unique ancestor and allow classification within a single acylphosphatase-like superfamily (Figure 1). The structural determination of mAcP, ctAcP, and HypF-N has basically revealed the same ferredoxin-like topology, consisting of a five-stranded antiparallel β-sheet facing on two antiparallel α-helices (*11*−*13*). Although the structure of Sso AcP has not yet been determined, the close sequence identity with the previously characterized acylphosphatases and procedures of structure modeling allow the same gross topology to be deduced for Sso AcP (Figure 1).

Despite their common origin and structural similarity, previous inspection of the folding processes of mAcP, ctAcP, and HypF-N has revealed different features. While mAcP and ctAcP have been found to fold in an apparent two-state fashion with no detectable misfolded or partially folded intermediates populated even at the lowest denaturant concentrations (*8*, *9*), HypF-N forms a partially structured state during folding (*10*). Under conditions close to physiological, this intermediate state is significantly more stable, has a stronger affinity for ANS, and has a higher intrinsic fluorescence than the fully unfolded protein (*10*). Another impressive difference between these systems involves the rate at which they fold. mAcP, ctAcP, and HypF-N fold with rate constants of 0.23 ± 0.04, 2.3 ± 0.5 and 80 ± 15 s^{-1}, respectively (*8*−*10*). Differences also exist as to the involvement of proline residues in slow and low-amplitude folding phases (*10*).

The results presented in this manuscript for Sso AcP contribute to introduction of a higher variability in the folding process of this superfamily than previously determined. We will show that Sso AcP forms an ensemble of partially folded conformations even more stable than that observed for HypF-N, folds with a rate intermediate between the two extremes of mAcP and HypF-N, and displays a slow folding phase that originates from a proline residue that is absent in all other acylphosphatases. The considerable heterogeneity observed in the folding behavior within the acylphosphatase superfamily and the analysis of such variability presented here sheds further light on the determinants of protein folding.

MATERIALS AND METHODS

Materials. Bovine cyclophilin A (CypA), 8-anilino-1-naphthalenesulfonic acid (ANS) and guanidinium chloride

[1] Abbreviations: AcPDro2, acylphosphatase no. 2 from *D. melanogaster*; ANS, 8-anilino-1-naphthalenesulfonic acid; ctAcP, human common-type acylphosphatase; CypA, cyclophilin A; far-UV CD, circular dichroism in the far ultraviolet; GdnHCl, guanidinium chloride; HypF-N, N-terminal domain of *E.coli* HypF; mAcP, human muscle acylphosphatase; Sso AcP, acylphosphatase from *Sulfolobus solfataricus*.

(GdnHCl) were purchased from Sigma-Aldrich. Benzoyl phosphate was synthesized and purified as described (*14*). DNA oligonucleotides were purchased from M-Medical (Milano, Italy).

Purification of Sso AcP and Its Mutated Variants. The gene coding for Sso AcP was originally inserted in a pEMBL plasmid. The DNA fragment corresponding to the Sso AcP gene was amplified by PCR using two primers that contained the restriction sites for *Bam*HI and *Eco*RI. The resulting amplified fragments were first purified using the Qia Quick Purification Kit by Qiagen, digested with *Bam*HI and *Eco*RI for 2 h, purified again, and then ligated into the pGEX-2T plasmid, previously digested with the same restriction enzymes. After the resulting plasmid was checked with DNA sequencing, this was transformed into *E. coli* DH5α cells. Gene expression in the DH5α cells and purification of the resulting protein were carried out as described for mAcP (*15*). Protein purity was checked by SDS−polyacrylamide gel electrophoresis and electrospray mass spectrometry. The resulting sequence of the protein is GS[1]MKKWSDTEV-FEMLKRMYARVYGLVQGVGFRKFVQIHAIRLGIKG-YAKNLPDGSVEVVAEGYEEALSKLLERIKQGPPAAE-VEKVDYSFSEYKGEFEDFETY. The Gly-Ser dipeptide at the N-terminus results from the cloning in pGEX-2T. The following methionine residue is residue no. 1. An extinction coefficient (ϵ_{280}) of 1.24 cm^{-1} mg^{-1} mL was used to determine protein concentration by UV absorption measurements. Mutants of Sso AcP were obtained using the QuickChange site-directed mutagenesis kit from Stratagene. The presence of the desired mutations was assessed by DNA sequencing. All the mutants were expressed and purified similarly to the wild-type protein.

Equilibrium Studies with Intrinsic Fluorescence and Far-UV CD. Fluorescence spectra from 290 to 450 nm (excitation at 280 nm) were acquired for 21 equilibrated samples containing 0.04 mg mL^{-1} Sso AcP and different concentrations of GdnHCl, ranging from 0 to 6 M in 50 mM acetate, pH 5.5, 37 °C. Spectra were acquired using a PerkinElmer LS 55 (Wellesley, Massachusetts). CD spectra from 210 to 260 nm were also acquired for 26 equilibrated samples containing 0.2 mg mL^{-1} protein and different concentrations of GdnHCl, ranging from 0 to 6.3 M, under the same conditions. Spectra were acquired using a Jasco J-810 spectropolarimeter (Great Dunmow, Essex, United Kingdom). The ratio of fluorescence values at the two major peaks (312 and 355 nm) and the mean residue ellipticity at 222 nm were plotted versus GdnHCl concentration. The two resulting plots were fitted to a two-state transition according to the equation described by Santoro and Bolen (*16*), to obtain the free energy difference between the unfolded and the native states in the absence of denaturant ($\Delta G_{U-F}^{H_2O}$), the dependence of ΔG_{U-F} on GdnHCl concentration (*m* value), and the midpoint of denaturation (C_m).

Folding and Unfolding Kinetics Monitored with Intrinsic Fluorescence. Unfolding and refolding reactions were followed using a Bio-logic SFM-3 stopped-flow device coupled to a fluorescence detection system (Claix, France). An excitation wavelength of 280 nm and a band-pass filter to monitor emitted fluorescence above 320 nm were used. All the experiments were performed in 50 mM acetate, pH 5.5, 37 °C, at final protein concentrations of 0.02−0.04 mg mL^{-1}. The unfolding experiments were initiated by dilution of the native protein into solutions containing GdnHCl at concentrations ranging from 3.5 to 6.5 M. The refolding reactions were similarly initiated by a dilution of the GdnHCl-denatured Sso AcP into solutions containing concentrations of GdnHCl between 0.2 and 3.5 M. The dead time was generally 10.4 ms. The unfolding and refolding traces were fitted to simple and double exponential functions, respectively, to determine the rate constants of unfolding and of the two phases of folding, together with their relative amplitudes. The fitting procedure was made using the general equation

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

where $y(t)$ is the fluorescence signal recorded as a function of time, A_i and k_i are the amplitude and the rate constant of the *i*th phase, respectively, q is the fluorescence value at equilibrium, and n is the number of observed phases.

Folding Kinetics Monitored with Far-UV CD. Folding of Sso AcP was monitored with far-UV CD at 222 nm using an Applied Photophysics PiStar-180 stopped-flow device coupled to a CD detection system (Leatherhead, U.K.). The folding reaction was initiated by an 11-fold dilution of 0.715 mg mL^{-1} Sso AcP denatured in 6 M GdnHCl into refolding buffer. Final conditions were 0.065 mg mL^{-1} Sso AcP, 0.55 M GdnHCl, 50 mM acetate, pH 5.5, 37 °C. Immediately after injection, the ellipticity at 222 nm was followed for up to 50 s. The dead time of the experiment was 20 ms. The measurement was repeated 35 times, and the resulting traces were averaged to obtain a trace with a maximized signal-to-noise ratio. The signal of the buffered solution with no protein was subtracted from the averaged trace. The averaged trace from 0 to 2 s was fitted to a single-exponential function (eq 1). No detectable changes of ellipticity were found after 2 s.

Folding Kinetics Monitored with ANS Fluorescence. Folding of Sso AcP was monitored in the presence of ANS using the Bio-Logic stopped-flow fluorimeter. The folding reaction was carried out by a 10-fold dilution of 0.5 mg mL^{-1} Sso AcP denatured in 5.5 M GdnHCl into refolding buffer containing 110 μM ANS in 50 mM acetate, pH 5.5. Final conditions were 0.05 mg mL^{-1} Sso AcP, 0.55 M GdnHCl, 99 μM ANS, 50 mM acetate, pH 5.5, 37 °C. Immediately after injection, the emission of the dye above 475 nm was followed using an excitation wavelength of 370 nm. The dead time of the stopped-flow measurement was 10.4 ms. The resulting ANS fluorescence trace was fitted to a double-exponential function (eq 1).

Folding in the Presence of CypA. Folding of Sso AcP was studied in the presence of CypA on the Bio-Logic stopped-flow apparatus by monitoring intrinsic fluorescence emission above 320 nm with excitation wavelength of 280 nm. Folding was initiated at 28 °C by a 30-fold dilution of 0.6 mg mL^{-1} Sso AcP denatured in 6 M GdnHCl into a refolding buffer containing 0.1 M Tris/trifluoroacetic acid, pH 8.0 and CypA. Final protein and GdnHCl concentrations were 0.02 mg mL^{-1} and 0.2 M, respectively. Final CypA concentrations ranged from 0 to 0.6 μM.

Enzymatic Activity Measurements. The enzymatic activity of Sso AcP was measured by a continuous spectrophoto-

metric method, using benzoyl phosphate as a substrate as previously described (*17*). All enzymatic activity measurements were subtracted by the spontaneous hydrolysis of the substrate measured in the absence of Sso AcP. Sso AcP (0.3 mg mL^{-1}) denatured in 5.5 M GdnHCl was diluted 10 times into a refolding buffer to yield final conditions of 0.03 mg mL^{-1} protein, 0.55 M GdnHCl, 50 mM acetate buffer, pH 5.5, 37 °C. At regular time intervals, starting from 35 s after initiation of the refolding reaction, 50 μL of the resulting sample was added to 950 μL of 5 mM benzoyl phosphate freshly dissolved in 50 mM acetate buffer, pH 5.5. As a control experiment, 50 μL of Sso AcP dissolved at a concentration of 0.03 mg mL^{-1} in 0.55 M GdnHCl, 50 mM acetate buffer, pH 5.5, 37 °C, were mixed with 950 μL of 5 mM benzoyl phosphate dissolved in 50 mM acetate buffer, pH 5.5. The enzymatic activity after refolding is expressed as percent of that measured in the control experiment.

RESULTS

Equilibrium Denaturation. The denaturation of Sso AcP was studied at equilibrium at pH 5.5, 37 °C, using GdnHCl as a chemical denaturant and intrinsic fluorescence and far-UV circular dichroism as spectroscopic probes to detect conformational changes of the protein following denaturation. Sso AcP contains seven tyrosine residues and a single tryptophan residue, the latter at position 4, a region that appears to be rather unstructured in acylphosphatases (*11*−*13*). For this reason, the major peak at 355 nm in the fluorescence spectrum does not undergo any evident transition from 0 to 6 M GdnHCl, while a significant, although weak, decrease was observed from ca. 3.5 to 5 M GdnHCl for the small peak at 312 nm that typically arises from tyrosine residues. Remarkable changes within this range of denaturant concentration were also detected in the far-UV CD spectrum.

The two plots reported in Figure 2 show the changes of the ratio of intrinsic fluorescence between 312 and 355 nm (excitation 280 nm) and mean residue ellipticity at 222 nm following progressive addition of GdnHCl. From these curves, the equilibrium unfolding of Sso AcP appears to be a reversible and cooperative event. The main thermodynamic parameters of the unfolding reaction were inferred from the equilibrium curves assuming a two-state model and using the equation described by Santoro and Bolen (*16*). Analysis of the curve obtained with far-UV CD led to values of 48 ± 4 kJ mol^{-1}, 11.3 ± 1.1 kJ mol^{-1} M^{-1}, and 4.2 ± 0.1 M for the free energy change upon unfolding in the absence of denaturant ($\Delta G_{U-F}^{H_2O}$), the dependence of ΔG_{U-F} on denaturant concentration (*m* value), and the midpoint of denaturation (C_m), respectively. Despite the poor signal-to-noise ratio of the curve obtained with intrinsic fluorescence, values of 45 ± 7 kJ mol^{-1}, 10.5 ± 2.5 kJ mol^{-1} M^{-1}, and 4.3 ± 0.2 M were obtained for the $\Delta G_{U-F}^{H_2O}$, *m*, and C_m values, respectively, in close agreement with those obtained with far-UV CD.

Folding and Unfolding Kinetics Monitored with Intrinsic Fluorescence. The kinetics of the folding and unfolding processes of Sso AcP were investigated at pH 5.5, 37 °C, using a stopped-flow fluorimeter. The folding reaction was initiated by diluting GdnHCl-denatured Sso AcP in refolding buffer and studied in the presence of final GdnHCl concentrations ranging from 0.2 to 3.5 M. At all denaturant concentrations investigated here, the resulting traces were fitted satisfactorily by double-exponential functions (eq 1), indicative of two major phases in the refolding process. A representative refolding trace, obtained in the presence of 0.55 M GdnHCl, is shown in Figure 3a. The fast and slow phases have apparent rate constants of 5.4 ± 1.0 and 0.26 ± 0.05 s^{-1} under these conditions (Table 1). The fast phase, with a relative amplitude of 92% ± 2%, accounts for the majority of the fluorescence change upon folding (Table 1). No further changes of fluorescence were observed after the slow phase. In addition, the enzymatic activity of Sso AcP after the slow phase is identical to that of the native protein that has not undergone unfolding (data not shown). This indicates that the folding process has gone to completion after the slow phase.

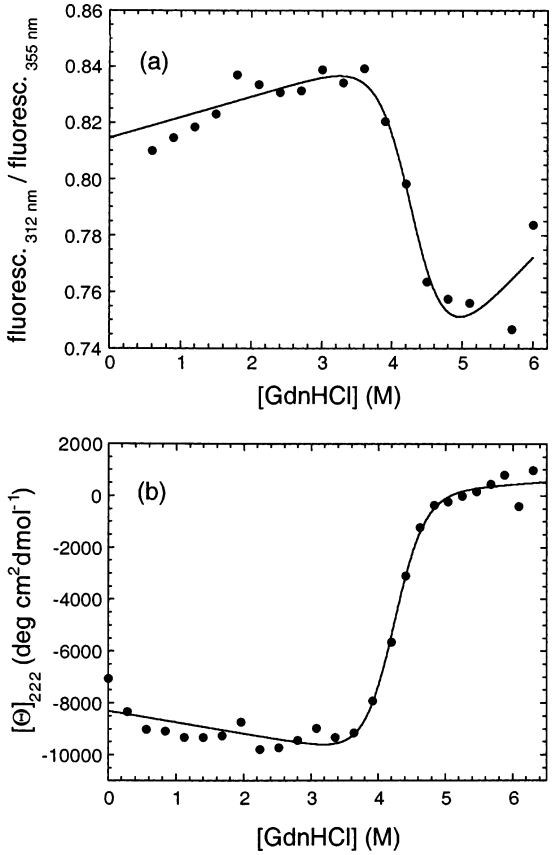


FIGURE 2: Equilibrium GdnHCl-denaturation curves of Sso AcP monitored in 50 mM acetate buffer, pH 5.5, 37 °C using (a) intrinsic fluorescence and (b) far-UV CD. Fluorescence data are reported as the ratio of the intrinsic fluorescence values at 312 and 355 nm (excitation of 280 nm). Mean residue ellipticity data are reported at 222 nm. The continuous line in each plot represents the best fit of the data points to the two-state equation described by Santoro and Bolen (*16*). The parameters derived from the fitting procedures are reported in the text.

To assess the occurrence of early conformational changes within the dead time of the stopped-flow experiment, the fluorescence signal of fully denatured Sso AcP at 0.55 M GdnHCl, determined by linear extrapolation from fluorescence measurements performed at high GdnHCl concentrations, was compared with the value extrapolated to time zero of the recorded fluorescence trace. The significant discrepancy between these two fluorescence values indicates the existence of an additional burst-phase that occurs on the

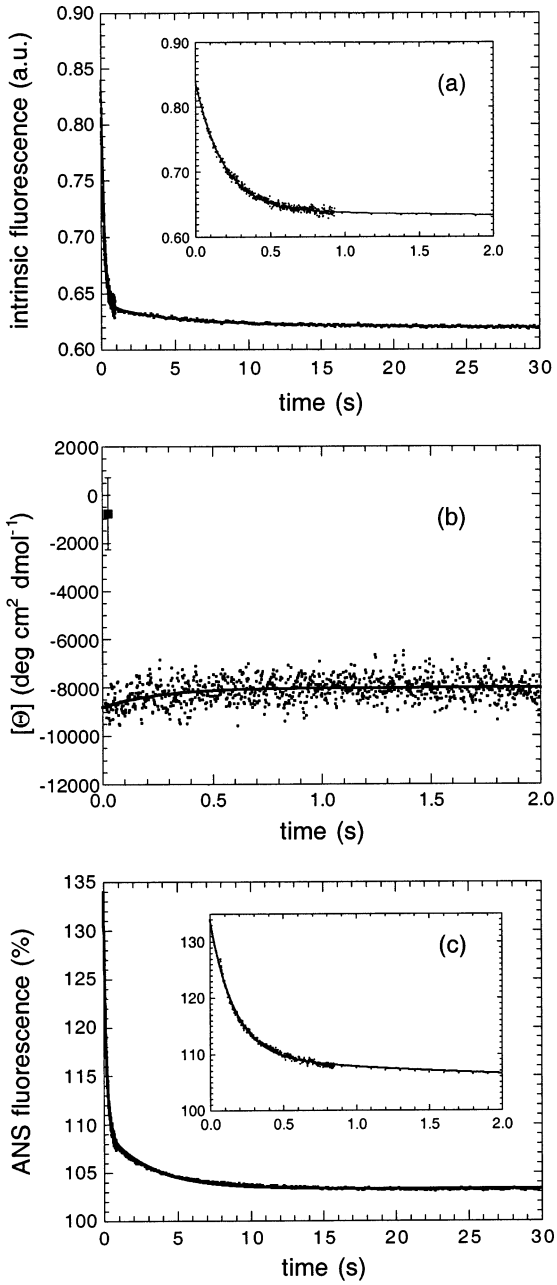


FIGURE 3: Representative folding traces of Sso AcP obtained in 0.55 GdnHCl, 50 mM acetate buffer, pH 5.5, 37 °C. The folding process was followed with (a) intrinsic fluorescence and protein concentration of 0.02 mg mL^{-1}, (b) far-UV CD at 222 nm and protein concentration of 0.065 mg mL^{-1}, and (c) ANS fluorescence and protein concentration of 0.05 mg mL^{-1}. Insets in panels a and c report the same folding traces of the corresponding panels expanded to show the first 2 s. The filled square in panel b represents the mean residue ellipticity of the fully unfolded protein extrapolated at the reported conditions from values measured at high GdnHCl concentrations. No detectable change of CD signal was observed after 2 s, most probably due to the poor signal-to-noise ratio that masks possible exponential phases with small amplitudes. In panel c, the data are reported as percent of the fluorescence emission of unbound ANS measured under identical conditions. In all three panels, the continuous lines through the data points represent the best fit of the data to double-exponential functions (eq 1). The rate constants and relative amplitudes for the two observed exponential phases obtained from the equations of best fit are reported in Table 1 for all three experiments.

submillisecond time scale and corresponds to a marked increase of the fluorescence emission intensity.

Table 1: Rate Constants (k) and Relative Amplitudes (A) for the Two Observed Phases of Sso AcP Folding Monitored by Different Methods[a]

	fast phase		slow phase	
spectroscopic probe	k (s^{-1})	A (%)	k (s^{-1})	A (%)
intrinsic fluorescence[b]	5.4 ± 1.0	92 ± 2	0.26 ± 0.05	8 ± 2
far-UV CD[c]	4.0 ± 2.0	nd	nd	nd
ANS fluorescence[d]	5.9 ± 1.0	80 ± 4	0.32 ± 0.05	20 ± 4

[a] Experimental conditions were 50 mM acetic acid, 0.55 M GdnHCl, pH 5.5, 37 °C. All traces were acquired after dilution of GdnHCl-denatured Sso AcP in refolding buffer. Recorded traces were analyzed by fitting to a double-exponential function (eq 1). [b] Folding reaction monitored with intrinsic fluorescence emission above 320 nm with an excitation wavelength of 280 nm. Final protein concentration was 0.02 mg ml^{-1}. [c] Folding reaction monitored with far-UV CD at 222 nm. Final protein concentration was 0.065 mg ml^{-1}. The slow phase was not detectable (nd) due to its low relative amplitude and the relatively poor signal-to-noise ratio in the recorded trace. [d] Folding reaction followed in the presence of 99 μM ANS by monitoring fluorescence emission above 475 nm with an excitation wavelength of 370 nm. Final protein concentration was 0.05 mg ml^{-1}.

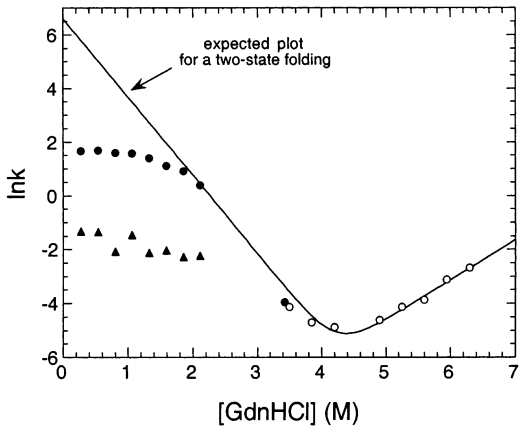


FIGURE 4: Dependence of the rate constants for the fast phase of folding (●), slow phase of folding (▲), and unfolding (○) on GdnHCl concentration in 50 mM acetate buffer, pH 5.5, 37 °C. The rate constants for the folding process were not determined between 2.2 and 3.3 M GdnHCl due to the partial or total superimposition of the two phases under these conditions (see text for further details). The continuous line represents the expected profile of the rate constants for the major fast phase of folding and unfolding if a two-state model accounted for the folding of Sso AcP. The expected profile was traced from the $\Delta G^{H_2O}_{U-F}$, m, and C_m values measured at equilibrium with far-UV CD and the unfolding rate constants measured with fluorescence. Deviation of the observed rate constants for folding from theoretical values at the low GdnHCl concentrations indicates deviation from a two-state mechanism under these conditions. Folding and unfolding rate constants extrapolated in the absence of denaturant are 5.4 ± 1.0 and $(6.1 \pm 1.0) \times 10^{-6}$ s^{-1}.

The unfolding process of Sso AcP was examined by GdnHCl concentration jump experiments at final denaturant concentrations ranging from 3.5 to 6.5 M. The obtained traces were all fitted satisfactorily to monoexponential functions, indicating that the unfolding reaction appears to be a monophasic process at all GdnHCl concentrations studied here. The dependence on denaturant concentration of the two rate constants for folding and of the unique rate constant for unfolding was determined (Figure 4). From 2.2 to 3.3 M GdnHCl, the major fast phase of folding does not occur on a time scale distinct from that of the minor slow phase. Under these conditions, accurate determination of the two rate

constants of folding is problematic (*18*) and contributes to creation of a blind area in the resulting plot (Figure 4). At GdnHCl concentrations lower than 2.2 M and higher than 3.3 M, the fast phase is distinct from the slow phase, making it possible to determine its rate constant (Figure 4). It is clear that the major fast phase for folding has a decreasing dependence on GdnHCl concentration as the denaturant decreases in concentration (Figure 4). This phenomenon, consisting of a downward curvature in the plot of ln *k* versus denaturant concentration, is often referred to as roll-over (*4*, *10*, *19*−*23*).

The natural logarithm of the rate constant of unfolding appears to have a positive and linear dependence on GdnHCl concentration, at least at concentrations higher than 5 M beyond the unfolding transition region (Figure 4). The solid line reported in Figure 4 indicates the profile that the major rate constant for folding should have if Sso AcP folded in a two-state fashion (traced on the basis of the parameters determined with far-UV CD at equilibrium and the observed unfolding rate constants). The considerable discrepancy between observed and expected rate constant for folding in the absence of denaturant clearly indicates deviation from a two-state model for folding of Sso AcP. Folding appears to be ca. 140-fold slower than that expected on the basis of a two-state mechanism.

Such a deviation has been observed for a number of proteins and generally attributed to the formation, on the submillisecond time scale, of partially folded intermediates under low-denaturant solvent conditions (*4*, *10*, *19*−*23*). Transient aggregation during folding or modifications of the folding transition state with denaturant concentration have also been indicated as possible causes of roll-overs of this type (*24*, *25*). However, the rate constants and relative amplitudes of the two exponential phases observed during folding do not change with protein concentration, at least in the range of 0.004−0.3 mg mL^{-1} (data not shown). This observation rules out the involvement of transient protein aggregation to account for the observed roll-over. Moreover, the jump of intrinsic fluorescence during the dead time of the stopped-flow experiment suggests that the roll-over arises from the accumulation of an intermediate, rather than modifications of the transition state with denaturant.

Folding Kinetics Monitored with Far-UV CD and ANS Fluorescence. To investigate further whether an ensemble of partially folded species accumulates within the dead time of the stopped-flow experiments, the refolding reaction of Sso AcP was followed with far-UV CD at 222 nm. Similarly to the experiment in which fluorescence was used as an optical probe, the refolding reaction was started by diluting, with a stopped-flow device, the GdnHCl-deantured protein to a final denaturant concentration of 0.55 GdnHCl. The resulting CD trace shows that a considerable increase of CD signal occurs within the dead time of the instrument (Figure 3b). Immediately after dilution, the mean residue ellipticity jumps from a value of -750 ± 1500 deg cm^2 $dmol^{-1}$ (corresponding to that of the unfolded protein and extrapolated from values measured at high denaturant concentration where the protein is fully unfolded) to a value of -8800 ± 700 deg cm^2 $dmol^{-1}$. As the protein refolds, the ellipticity undergoes a slight increase up to the value of the native protein (Figure 3b). Such a change occurs with a rate constant of 4.0 ± 2.0 s^{-1}, which is, within experimental error, that measured with intrinsic fluorescence (Table 1). Because of the poor signal-to-noise ratio, it was not possible to detect any subsequent slow and low amplitude change of the CD signal, corresponding to that of the second phase observed with fluorescence.

The folding of Sso AcP was also followed at a final GdnHCl concentration of 0.55 M in the presence of the fluorescent dye 8-anilino-1-naphthalenesulfonic acid (ANS), which is known to bind to exposed hydrophobic regions of proteins (*26*). After the dead time of the refolding experiment, ANS fluorescence was found to be ca. 35% higher than that extrapolated for the fully unfolded protein under the same conditions, the latter being similar to that of ANS in the absence of protein (Figure 3c). This suggests that folding of Sso AcP proceeds through the formation of metastable, partially folded species with significant unburied hydrophobic clusters. The trace recorded after the dead time shows a biphasic decrease of ANS fluorescence with rate constants of 5.9 ± 1.0 and 0.32 ± 0.05 s^{-1}, in agreement with those obtained using intrinsic fluorescence and far-UV CD as spectroscopic probes (Table 1). When the second slower process reaches equilibrium the fluorescence value is only 3% higher than that of unbound ANS indicating only weak binding of ANS, if any, for the native state (Figure 3c).

Proline Isomerization. The lack of a marked dependence of the slow exponential phase of Sso AcP folding on denaturant concentration suggests that this phase arises from proline cis−trans isomerism (*27*). The folding process of Sso AcP was followed in the presence of catalytic amounts of bovine cyclophilin A (CypA), a peptidyl-prolyl isomerase that catalyses the cis−trans interconversion of X-Pro peptide bonds (*27*). Folding was followed at pH 8.0, 28 °C, and in the presence of low concentrations of residual GdnHCl (0.2 M), conditions found to be optimal for CypA catalysis (*28*). We found that while the rate of the fast phase was not perturbed by CypA, the slow phase was significantly affected with an acceleration of over 2-fold at the highest concentration of CypA utilized (Figure 5a).

A site-directed mutagenesis approach was also employed to identify which proline residue is responsible for the observed slow phase of folding. Sso AcP has three proline residues at positions 50, 76, and 77. Three single mutants of the protein, each with a proline residue replaced with alanine, were therefore produced. The folding reactions of the three variants and the wild-type protein were studied at a final GdnHCl concentration of 1.1 M, pH 5.5 and 37 °C, using intrinsic fluorescence as a spectroscopic probe. The P76A and P77A mutants of Sso AcP were found to fold with two exponential phases displaying rates and relative amplitudes comparable to those of the wild-type protein (Table 2). By contrast, the P50A mutant folds with a single exponential phase (Table 2 and Figure 5b). This unique kinetic phase, which accounts for 100% of the recovery of native fluorescence in this mutant, has a rate lower than that of the fast phase observed for the wild-type and other protein variants (Table 2). This is probably due to a decelerating effect of the amino acid substitution on the folding process of Sso AcP. Importantly, however, no slow phase is observed in the folding process of this variant. The change of fluorescence observed during the unique exponential phase corresponds to the sum of the two phases observed for the wild-type protein (Figure 5b). These results indicate that the

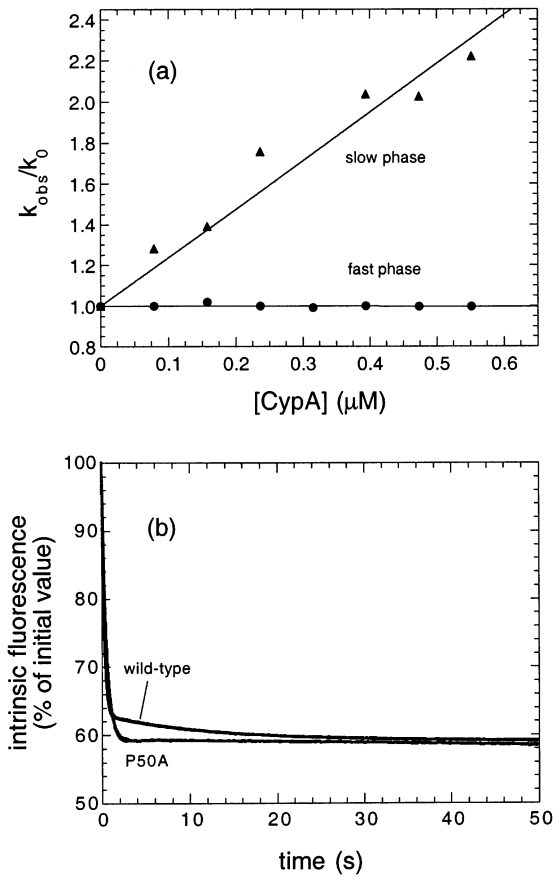


FIGURE 5: (a) Effect of CypA on the fast and slow phases of Sso AcP folding. The folding reaction was monitored at a protein concentration of 0.02 mg mL^{-1} in 0.2 M GdnHCl, 0.1 M Tris, pH 8.0, 28 °C, in the presence of increasing quantities of CypA. Data are reported as ratios of the rate constants (k_{obs}) obtained for the fast (●) and slow (▲) phases in the presence of CypA and the corresponding values in the absence of enzyme (k_0). The continuous lines represent best fits to linear functions. Panel b shows Folding traces recorded with intrinsic fluorescence for wild-type and P50A Sso AcP. Final conditions were 1.1 M GdnHCl, 50 mM acetate, pH 5.5, 37 °C. The kinetic parameters for the observed exponential phases are reported in Table 2.

Table 2: Rate Constants (k) and Relative Amplitudes (A) for Folding of Wild-Type and Proline to Alanine Mutants of Sso AcP[a]

	fast phase		slow phase	
mutant	k (s^{-1})	A (%)	k (s^{-1})	A (%)
wild-type	3.9 ± 0.7	90 ± 2	0.080 ± 0.030	10 ± 2
P50A	2.0 ± 0.4	100 ± 1		
P76A	5.0 ± 0.8	91 ± 3	0.075 ± 0.030	9 ± 3
P77A	4.6 ± 0.7	92 ± 3	0.080 ± 0.030	8 ± 3

[a] In all cases, final conditions were 0.04 mg ml^{-1} protein, 1.1 M GdnHCl, 50 mM acetate, pH 5.5, 37 °C.

slow phase of folding of Sso AcP is entirely attributable to the cis−trans isomerism of the Leu49-Pro50 peptide bond.

DISCUSSION

The Enhanced Thermodynamic Stability of Sso AcP Results from a Modification of Folding and Unfolding Rates. Sso AcP displays a conformational stability higher than any acylphosphatase from mesophilic organisms so far studied (*8*−*10*, *29*, *30*). The high stability of proteins from hyperthermophiles is due to a number of adaptive mechanisms including a high content of hydrophobic and charged amino acids, as opposed to polar uncharged amino acids, and a high stability of secondary structure elements (*31*). Recently, particular importance has been attributed to the high content of charged residues and to ion pair networking at the protein surface (*32*). Sso AcP contains a relatively high number of both charged (32% for Sso AcP versus 25.5% for ctAcP and mAcP) and hydrophobic residues (56% for Sso AcP versus 43% for mAcP and 48% for ctAcP), thus supporting the central importance of this modification in the amino acid content between mesophilic and hyperthermophilic proteins. Our data show that the increased stability of Sso AcP is the consequence of a decreased rate constant for unfolding, rather than an increased rate for folding (*8*−*10*). Indeed, Sso AcP was found to fold with a rate similar to that of ctAcP and even lower than that of HypF-N. This is in agreement with the comparative analysis carried out on cold-shock proteins from different organisms, showing that proteins with higher stability do not necessarily fold faster (*33*).

An Ensemble of Partially Structured Conformations Form during Folding of Sso AcP. Sso AcP folds via a mechanism that involves at least two steps. After removal of denaturant, the unfolded state collapses on the microsecond time scale into an ensemble of partially folded conformations. This converts subsequently into the fully folded state with a rate constant of ca. 5 s^{-1} under the conditions of temperature and pH studied here. A small fraction of Sso AcP molecules folds more slowly with a rate constant of ca. 0.2 s^{-1}. In this case, the folding process is rate-limited by the cis−trans conversion of the Leu49-Pro50 peptide bond. Formation of a burst-phase intermediate state that precedes formation of the native structure is revealed by three distinct lines of evidence: (i) a mean residue ellipticity similar to that of the native state is reached within an instrumental dead time of a few milliseconds, well before the observed fast and slow phases of folding occur (Figure 3b); (ii) the maximum affinity for ANS, with consequent increase of its fluorescence, is found just after the dead time of our stopped-flow experiment; ANS fluorescence decreases as the protein folds through the two observed exponential phases (Figure 3c); (iii) a roll-over is observed in the plot reporting the natural logarithm of the rate constant for the fast phase versus denaturant concentration; the folding rate in the absence of denaturant is ca. 140 times slower than that expected for a two-state mechanism (Figure 4).

Factors Influencing the Population and Stability of Intermediate States in Protein Folding. On the assumption that the intermediate and unfolded states fold through the same transition state, the conformational stability of this intermediate state can be determined using

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

where $\Delta G_{I-U}^{H_2O}$ is the free energy change for the conversion of the unfolded into the intermediate state in the absence of denaturant, $k_{F-I}^{H_2O}$ is the observed folding rate starting from the intermediate state in the absence of denaturant, and $k_{F-U}^{H_2O}$ is the folding rate constant from the unfolded state in the absence of denaturant extrapolated from the equilibrium and unfolding rate data as shown in Figure 4. Our results allow a value of −12.5 ± 2.0 kJ mol^{-1} to be obtained for

$\Delta G_{I-U}^{H_2O}$. A similar analysis applied to the intermediate state forming during folding of HypF-N and using previously published data (*10*) yields a value of -2.6 ± 0.3 kJ mol^{-1} for $\Delta G_{I-U}^{H_2O}$. Values higher than 0 kJ mol^{-1} hold for the $\Delta G_{I-U}^{H_2O}$ parameter that may potentially pertain to any folding intermediate of mAcP and ctAcP, because no intermediate states appear to be significantly populated during the folding processes of these two proteins (*8*, *9*). These results indicate that the intermediate state forming during the folding process of Sso AcP, consisting of an ensemble of either partially folded or misfolded conformations, has the highest conformational stability relative to the unfolded state among the acylphosphatases so far characterized. This agrees with the idea that the stability of a protein refers to the native state as well as any intermediates of its folding process (*31* and references therein). Interestingly, the scale of the four proteins considered here, established on the basis of the stabilities of the intermediate states, mirrors that established on the basis of the stabilities of the native states.

While the heterogeneity found in the stabilities of protein folding intermediates supports the view that the distinction between two-state and multistate folding simply depends on whether the intermediate states under refolding conditions are sufficiently stable for experimental detection (*4*, *34*), it is interesting to investigate what factors determine the stabilities of such intermediate states. The ensemble of partially structured conformations detected during folding of Sso AcP is reminiscent of the molten globule state given the high content of secondary structure and the solvent exposure of hydrophobic clusters (*35*). Since partially folded states of this type are stabilized by hydrophobic interactions between fully or partially formed α-helices (*35*), we have analyzed the hydrophobicity and propensity to form α-helical structure of the sequences of mAcP, ctAcP, HypF-N, and Sso AcP (Figure 6). The two proteins forming intermediate states during folding, that is, Sso AcP and HypF−N, possess sequences with the highest peaks, less negative minima of hydrophobicity, or both (Figure 6a). The highest hydrophobicity of these two proteins relative to mAcp and ctAcP also results from the average hydrophobicity values calculated from the entire sequences (see legend to Figure 6). In addition, Sso AcP and HypF-N have net charges of −1 and +1, respectively, while ctAcP and mAcP have net charge values of +6 and +5. The relatively high hydrophobic content and a charge state close to neutrality can favor the rapid hydrophobic collapse of the unfolded protein for Sso AcP and HypF-N before the polypeptide chain reorganizes to form the native conformation. Nevertheless, differences in hydrophobicity and net charge among the four proteins analyzed here cannot explain why the intermediate state of Sso AcP is more stable than that of HypF-N because the latter appears to be more hydrophobic.

The α-helical propensity profiles reveal that Sso AcP has a relatively large propensity to form α-helical structure in at least two distinct regions (Figure 6b). The Sso AcP peaks are significantly more intense than the largest peaks of the HypF-N sequence and remarkably more marked than those of the ctAcP sequence, the latter possessing very little propensity to form α-helices (Figure 6b). mAcP also has a relatively large propensity to form α-helical structure at least in one region of the sequence. Nevertheless, mAcP has the lowest hydrophobicity among the proteins considered here therefore explaining why potential folding intermediates remain undetectable (see Figure 6a and legend to Figure 6).

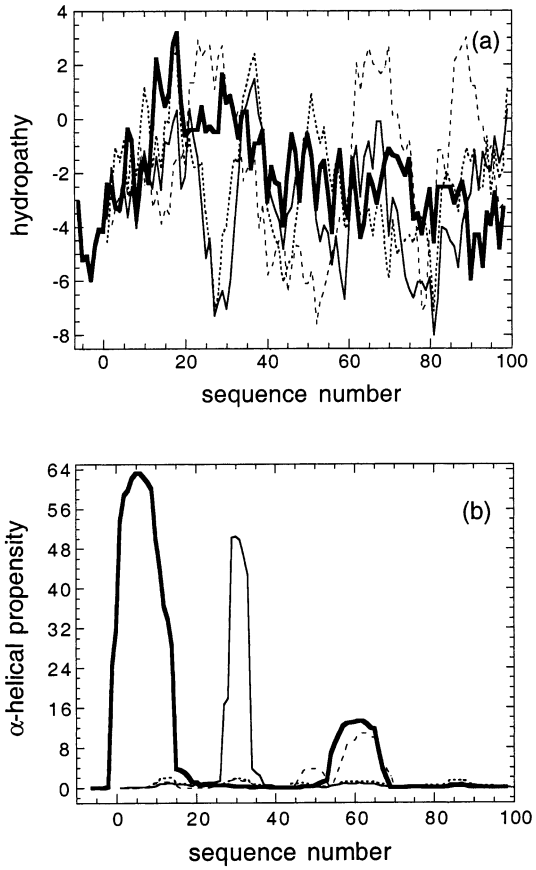


FIGURE 6: Hydropathy (a) and α-helical propensity (b) profiles of the sequences of Sso AcP (thick continuous line), mAcP (thin continuous line), ctAcP (dotted line), and HypF-N (dashed line). The hydropathy profile was obtained using the scale of hydrophobicity values of the 20 amino acid residues edited by Roseman (*46*). The reported hydrophobicity value for each amino acid residue represents the average hydrophobicity of a window of nine residues centered on that residue, except for residues at the C- and N-termini where narrower windows are used. The α-helical propensity profile for each sequence was obtained using the whole sequence as an input for the AGADIR algorithm (*47*) and plotting the calculated propensities for each amino acid residue. High values indicate high hydrophobicity and α-helical propensities. In panel a, a value of zero indicates a hydrophobicity equal to a stretch of nine glycine residues. In panel b, a value of zero indicates no propensity to form α-helical structure. Sequences were aligned to obtain the highest superimposition of identical residues. The numbering of the sequence reported in the figure refers to that of mAcP and ctAcP. Average hydrophobicities calculated for the entire sequences are −0.478, −0.717, −0.552, and −0.416 for Sso AcP, mAcP, ctAcP, and HypF-N, respectively (less negative numbers indicate higher hydrophobicities). Average α-helical propensities are 9.79, 3.27, 0.64, and 1.42% for Sso AcP, mAcP, ctAcP, and HypF-N, respectively.

The importance of α-helical propensity in determining the appearance of partially folded species during folding also results from the comparison of structurally related proteins from other superfamilies. Im7 and Im9, two *E. coli* proteins sharing considerable sequence identity and belonging to the Colicin E immunity family, fold with different behaviors because only Im7 populates a partially structured species during folding (*23*). While the two proteins have equivalent hydropathy profiles, Im7 has a higher α-helical propensity in two of the four segments adopting helical structure in the

native state (analysis of hydrophobicity and α-helical propensity is performed similarly to that for acylphosphatases, see legend to Figure 6). The higher α-helical propensity of Im7 also results from the average values that are 2% and 1% for Im7 and Im9, respectively. The engrailed homeodomain from *Drosophila melanogaster* (En-HD) and the c-Myb DNA-binding repeat III from mouse (c-Myb) belong to the homeodomain-like superfamily. Although En-HD has a slightly lower hydrophobicity than c-Myb, its higher propensity to form α-helical structure is compelling (average values are 8.44% and 2.73% for En-HD and c-Myb, respectively). Interestingly, only En-HD collapses into a partially folded species before attaining the native state (*36*, S. Gianni, personal communication).

In summary, combination of hydrophobicity, net charge, and propensity to form α-helices can help rationalize the tendency of proteins to form folding intermediate states. Since the variability in sequence among homologous proteins is sufficiently high to generate differences in these determinants, it is not surprising that intermediates states are sufficiently stable for experimental detection only in a few members of a protein superfamily.

Hydrophobic Content of a Polypeptide Chain Is a Determinant of Its Folding Rate. Within the acylphosphatase-like superfamily, a significant correlation is found between folding rate and hydrophobic content of the sequence, the latter simply defined as the average hydrophobicity of the amino acid residues composing the sequence (Figure 7a). Sso AcP is the protein that most deviates from the line of best fit, a finding probably attributable to the relatively high stability of the intermediate state that causes its folding rate to decrease more than those for the other proteins (Figure 4). Despite this deviation, the *r* and *p* values of the analysis show that the correlation is significant (Figure 7a). Interestingly, a significantly correlation between folding rate and sequence hydrophobicity is also observed when the folding rate constant is that extrapolated from a two-state model, that is, $k_{F-U}^{H_2O}$ as opposed to $k_{F-I}^{H_2O}$ (Figure 7b). This indicates that the correlation described in Figure 7a is not due to artifacts arising from the formation of burst-phase intermediates in some proteins. The folding rate of proteins of the acylphosphatase-like superfamily is not found to correlate with any other sequence-related factors. Plots of folding rate versus parameters such as $\Delta G_{U-F}^{H_2O}$, helix propensity, or net charge of the sequence all produce *p* values higher than 50% (data not shown).

It is well-known that folding rate correlates well with the relative contact order of a protein, that is, the average distance of residues forming an interaction in the native state divided by the number of total residues, and other related metrics such as Q_D, that is, the number of pairs of sequence-distant residues that are in contact in the native state (*7*, *37*, *38*). Since the three-dimensional structure has been solved for only three proteins from this superfamily, it is difficult to assess whether relative contact order or other related metrics are important determinants of folding rate from an analysis performed with acylphosphatases only. The folding rate of those acylphosphatases for which the structure is known correlates well, however, with the line of best fit obtained using a large set of structurally unrelated proteins (*7*, *37*). It can therefore be concluded that relative contact order and

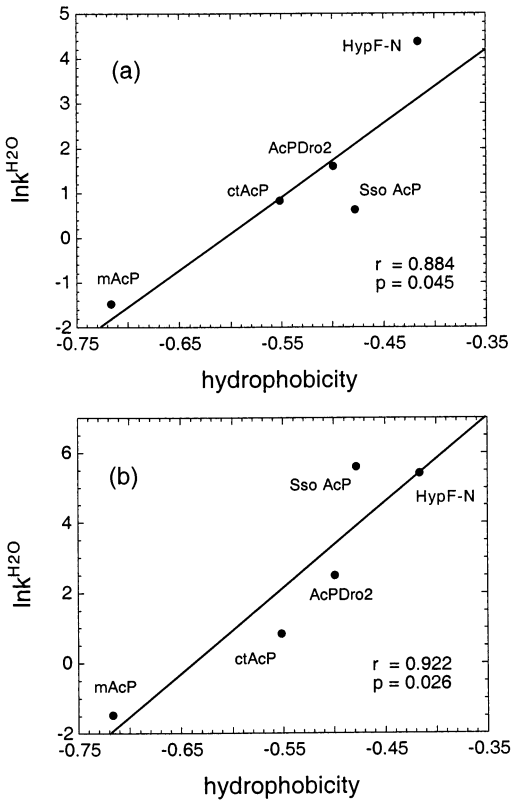


FIGURE 7: Correlation between rate constant for folding and hydrophobic content of the sequence for five acylphosphatases. The reported values of folding rate constant are those truly observed (a) or extrapolated from a two-state model, that is, assuming that folding starts from the unfolded rather than intermediate state (b). All rate constants refer to identical experimental conditions (50 mM acetate buffer, pH 5.5, no denaturant, 28 °C). The rate constants for Sso AcP considered in this plot were obtained at 28 °C after obtaining a plot similar to that reported in Figure 4 at 28 °C, rather than 37 °C (data not shown). To add statistical significance to our analysis, we have also included in the plot the folding rates of a new acylphosphatase that is currently under investigation from *D. melanogaster* (AcPDro2) (M. Ramazzotti et al., unpublished). Folding rate constants for mAcP, ctAcP, and HypF-N are reported elsewhere (*8*−*10*). The *p* values lower than 0.05 indicate that the correlations are significant.

hydrophobicity are the main determinants of folding rate for proteins of the acylphosphatase superfamily.

A statistically significant correlation between folding rate and sequence hydrophobicity was also found for proteins displaying the ferredoxin-like topology, provided the folding rate is considered at the midpoint of chemical denaturation to level out differences in conformational stabilities between the proteins (*10*). Furthermore, amino acid substitutions aimed at increasing the hydrophobicity at positions that are not involved in the folding nucleus can accelerate the folding process (*39*−*41*). The relatively fast folding in the presence of a high hydrophobic content can probably be attributed to the ability of hydrophobic moieties to facilitate encounters between distinct portions of the polypeptide chain and, consequently, stabilize the transition-state ensemble. It has also been attributed to the ability of nonnative hydrophobic contacts to provide additional energetically accessible pathways to the native state, therefore stabilizing the transition state ensemble by addition of an entropic bonus (*40*). Following these results from independent lines of investigation, it can be proposed that the hydrophobic content is an

important factor, in addition to relative contact order, in determining the folding rate of a protein.

Proline Residues Responsible for Slow Folding Events Are Not Conserved. Within the acylphosphatase superfamily, the involvement of proline residues in folding has been studied in detail for mAcP (*8*, *42*) and HypF-N (*10*), in addition to Sso AcP. In these three proteins, slow phases of folding, attributed to cis−trans proline isomerization and assigned to specific proline residues using a site-directed mutagenesis approach, have been observed with at least two spectroscopic techniques. The residues responsible for such slow phases are Pro54, Pro78, and Pro50 for mAcP, HypF-N and Sso AcP, respectively. It is clear form the alignment of the sequences presented in Figure 1 that these three prolines do not occupy homologous positions. Despite the obstacles raised by proline residues for fast and efficient folding, it appears that there is not an evolutionary pressure to maintain proline residues at specific positions in which they do not act as rate-determining residues. Nor does an evolutionary pressure seem to exist to maintain nonproline residues at positions where prolines could slow folding substantially. The variability of folding events associated with proline residues is also evident from the decelerated folding of a fraction of HypF-N molecules (*10*). This behavior, attributable to the cis configuration of a X-Pro peptide bond that is not prohibitive for folding, though causing it to be slower, has not been observed in the folding process of any other acylphosphatase (*10*). Proline isomerism was also reported to be highly variable in two homologues of ribonuclease H (*43*).

It is likely that peptidyl-prolyl isomerases, normally present in all compartments within cells and in all organisms regardless of their position on the evolutionary scale, can easily buffer potential problems arising from proline-induced slow folding and maintain folding efficiency above levels with which cells can easily manage. Molecular chaperones can cooperate with these enzymes by binding to unfolded or partially folded states that are persistent due to incorrect configurations of X-Pro peptide bonds (*44*). Protective mechanisms of this kind have probably allowed sequences to evolve independently of potential restraints associated with proline residues. The elasticity observed in the positioning of proline residues during evolution adds to the lack of evolutionary conservation of residues that play key roles in promoting folding (*45*).

CONCLUSIONS

The investigation of the folding process of Sso AcP presented here and the subsequent comparative analysis with other homologous proteins from the same superfamily have shown that folding rate, stability of partially folded states accumulating during folding and proline-associated folding phenomena are not evolutionarily conserved. Our analysis addresses sequence hydrophobicity as an important determinant of folding rate, conformational stability of the native state, and stability of the partially folded state of a protein. A low net charge and a high propensity to form α-helical structure also emerge as possibly important determinants of the stability of partially folded states. The results obtained here lend support to approaches of comparing proteins from one structural superfamily with the aim of learning the rules that the govern the hyperthermostability of proteins and the mechanism by which proteins fold.

REFERENCES

1. Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins 41*, 415−427.
2. Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations, *J. Mol. Biol. 320*, 369−387.
3. Chiti, F., Taddei, N., White, P. W., Bucciantini, M., Magherini, F., Stefani, M., and Dobson, C. M. (1999) Mutational analysis of AcP suggests the importance of topology and contact order in protein folding, *Nat. Struct. Biol. 6*, 1005−1010.
4. Fowler, S. B., and Clarke, J. (2001) Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state, *Structure (Cambridge) 9*, 355−366.
5. Grantcharova, V., Alm, E. J., Baker, D., and Horwich, A. L. (2001) Mechanisms of protein folding, *Curr. Opin. Struct. Biol. 11*, 70−82.
6. Friel, C. T., Capaldi, A. P., and Radford, S. E. (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins, *J. Mol. Biol. 326*, 293−305.
7. Plaxco, K. W., Simons, K. T., Ruczinski, I., and Baker, D. (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics, *Biochemistry 39*, 11177−11183.
8. van Nuland, N. A. J., Chiti, F., Taddei, N., Raugei, G., Ramponi, G., and Dobson, C. M. (1998) Slow folding of muscle acylphosphatase in the absence of intermediates, *J. Mol. Biol. 283*, 883−891.
9. Taddei, N., Chiti, F., Paoli, P., Fiaschi, T., Bucciantini, M., Stefani, M., Dobson, C. M., and Ramponi, G. (1999) Thermodynamics and kinetics of folding of common-type acylphosphatase: comparison to the highly homologous muscle isoenzyme, *Biochemistry 38*, 2135−2142.
10. Calloni, G., Taddei, N., Plaxco, K. W., Ramponi, G., Stefani, M., and Chiti, F. (2003) Comparison of the folding processes of distantly related proteins. Importance of hydrophobic content in folding, *J. Mol. Biol. 330*, 577−591.
11. Pastore, A., Saudek, V., Ramponi, G., and Williams, R. J. (1992) Three-dimensional structure of acylphosphatase. Refinement and structure analysis, *J. Mol. Biol. 224*, 427−440.
12. Thunnissen, M. M., Taddei, N., Liguri, G., Ramponi, G., and Nordlund, P. (1997) Crystal structure of common type acylphosphatase from bovine testis, *Structure 5*, 69−79.
13. Rosano, C., Zuccotti, S., Bucciantini, M., Stefani, M., Ramponi, G., and Bolognesi, M. (2002) Crystal structure and anion binding in the prokaryotic hydrogenase maturation factor HypF acylphosphatase-like domain, *J. Mol. Biol. 321*, 785−796.
14. Camici, G., Manao, G., Cappugi, G., and Ramponi, G. (1976). A new synthesis of benzoyl phosphate: a substrate for acyl phosphatase assay, *Experientia 32*, 535−536.
15. Taddei, N., Stefani, M., Magherini, F., Chiti, F., Modesti, A., Raugei, G., and Ramponi, G. (1996) Looking for residues involved in the muscle acylphosphatase catalytic mechanism and structural stabilization: role of Asn41, Thr42, and Thr46, *Biochemistry 35*, 7077−7083.
16. 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.
17. Ramponi, G., Treves, C., and Guerritore, A. (1966) Aromatic acyl phosphates as substrates of acyl phosphatase, *Arch. Biochem. Biophys. 115*, 129−135.
18. Fersht, A. R. (1999) *Structure and mechanism in protein science. A guide to Enzyme catalysis and protein folding*, pp 132−168, W. H. Freeman and Company, New York.
19. Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) Transient folding intermediates characterised by protein engineering, *Nature 346*, 440−445.
20. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues, *Nat. Struct. Biol. 3*, 193−205.

21. Lopez-Hernandez, E., and Serrano, L. (1996) Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2, *Folding Des. 1*, 43−55.
22. Parker, M. J., and Clarke, A. R. (1997) Amide backbone and water-related H/D isotope effects on the dynamics of a protein folding reaction, *Biochemistry 36*, 5786−5794.
23. Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. (1999) Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9, *J. Mol. Biol. 286*, 1597−1608.
24. Silow, M., and Oliveberg, M. (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6046.
25. Sanchez, I. E., and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding, *J. Mol. Biol. 325*, 367−376.
26. Engelhard, M., and Evans, P. A. (1995) Kinetics of interaction of partially folded proteins with a hydrophobic dye: evidence that molten globule character is maximal in early folding intermediates, *Protein Sci. 4*, 561−602.
27. Schmid, F. X. (2001) Prolyl isomerases, *Adv. Protein Chem. 59*, 243−282.
28. Lang, K., Schmid, F. X., and Fischer, G. (1987) Catalysis of protein folding by prolyl isomerase, *Nature 329*, 268−270.
29. Pieri, A., Magherini, F., Liguri, G., Raugei, G., Taddei, N., Bozzetti, M. P., Cecchi, C., and Ramponi, G. (1998) *Drosophila melanogaster* acylphosphatase: a common ancestor for acylphosphatase isoenzymes of vertebrate species, *FEBS Lett. 433*, 205−210.
30. Degl'Innocenti, D., Ramazzotti, M., Marzocchini, R., Chiti, F., Raugei, G., and Ramponi, G. (2003) Characterization of a novel *Drosophila melanogaster* acylphosphatase, *FEBS Lett. 535*, 171−174.
31. Jaenicke, R., and Bohm, G. (1998) The stability of proteins in extreme environments, *Curr. Opin. Struct. Biol. 8*, 738−748.
32. Karshikoff, A., and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads, *Trends Biochem. Sci. 26*, 550−556.
33. Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R., and Schmid, F. X. (1998) Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins, *Nat. Struct. Biol. 5*, 229−235.
34. Dobson, C. M., Sali, A., and Karplus, M. (1998) Protein folding: A perspective from theory & experiment, *Angew. Chem., Int. Ed. 37*, 868−893.
35. Kuwajima, K., and Arai, M. (2000) The molten globule state: the physical picture and biological significance, in *Mechanism of protein folding* (Pain, R., Ed.) pp 138−174, Oxford University Press, Oxford, U.K.
36. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S., Freund, S. M., Alonso, D. O., Daggett, V., and Fersht, A. R. (2003) The complete folding pathway of a protein from nanoseconds to microseconds, *Nature 421*, 863−867.
37. Makarov, D. E., and Plaxco, K. W. (2003) The topomer search model: A simple, quantitative theory of two-state protein folding kinetics, *Protein Sci. 12*, 17−26.
38. Gunasekaran, K., Eyles, S. J., Hagler, A. T., and Gierasch, L. M. (2001) Keeping in the family: folding studies of related proteins, *Curr. Opin. Struct. Biol. 11*, 83−93.
39. Northey, J. G., Di Nardo, A. A., and Davidson, A. R. (2002) Hydrophobic core packing in the SH3 domain folding transition state, *Nat. Struct. Biol. 9*, 126−130.
40. Viguera, A. R., Vega, C., and Serrano, L. (2002) Unspecific hydrophobic stabilization of folding transition states, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5349−5354.
41. Cobos, E. S., Filimonov, V. V., Vega, M. C., Mateo, P. L., Serrano, L., and Martinez, J. C. (2003). A thermodynamic and kinetic analysis of the folding pathway of an SH3 domain entropically stabilised by a redesigned hydrophobic core, *J. Mol. Biol. 328*, 221−233.
42. Chiti, F., Taddei, N., Giannoni, E., van Nuland, N. A. J., Ramponi, G., and Dobson, C. M. (1999) Development of enzymatic activity during protein folding, *J. Biol. Chem. 29*, 20151−20158.
43. Hollien, J., and Marqusee, S. (2002) Comparison of the folding processes of *T. thermophilus* and *E. coli* ribonucleases H, *J. Mol. Biol. 316*, 327−340.
44. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein, *Science 295*, 1852−1858.
45. Larson, S. M., Ruczinski, I., Davidson, A. R., Baker, D., and Plaxco, K. W. (2002). Residues participating in the protein folding nucleus do not exhibit preferential evolutionary conservation, *J. Mol. Biol. 316*, 225−233.
46. Creighton, T. E. (1993) *Proteins. Structure and molecular properties*, 2nd ed., p 154, W. H. Freeman & Company, New York.
47. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) Elucidating the folding problem of a-helices: Local motifs, long-range electrostatics, ionic strength dependence and prediction of NMR parameters, *J. Mol. Biol. 284*, 173−191.

BI030238A