

Biological function in a non-native partially folded state of a protein

Francesco Bemporad¹, Joerg Gsponer²,
Harri I Hopearuoho², Georgia Plakoutsi¹,
Gianmarco Stati¹, Massimo Stefani¹,
Niccolò Taddei¹, Michele Vendruscolo²
and Fabrizio Chiti^{1,*}

¹Dipartimento di Scienze Biochimiche, Università degli Studi di Firenze, Firenze, Italy and ²Department of Chemistry, University of Cambridge, Cambridge, UK

As structural flexibility is known to be required for enzyme catalysis and pattern recognition and a significant fraction of eukaryotic proteins appear to be unfolded or contain unstructured regions, biological activity of conformational states distinct from fully folded structures could be more common than previously thought. By applying a procedure that allows the recovery of enzymatic activity to be monitored in real time, we show that a non-native state populated transiently during folding of the acylphosphatase from *Sulfolobus solfataricus* is enzymatically active. The structural characterization of this partially folded state reveals that enzymatic activity is possible even if the catalytic site is structurally heterogeneous, whereas the remainder of the structure acts as a scaffold. These results extend the spectrum of biological functions carried out in the absence of a folded state to include enzyme catalysis.

The EMBO Journal (2008) 27, 1525–1535. doi:10.1038/emboj.2008.82; Published online 1 May 2008

Subject Categories: proteins; structural biology

Keywords: enzyme dynamics; folding nucleus; intrinsically disordered proteins; Phi-value; protein evolution

Introduction

Proteins are among the most abundant macromolecules in living systems and carry out a vast number of functions, including the catalysis of virtually every chemical transformation occurring in cell biology and the transduction of signals inside and between cells. Although it is well known that the attainment of well-defined, folded three-dimensional structures is crucial in determining their function, increasing evidence is accumulating about the existence of proteins or protein domains that adopt unstructured but functional states under physiological conditions (Dunker *et al.*, 2001; Fink, 2005). This raises the important point that biological function can be attained by proteins under conformational states

distinct from fully folded structures. The mechanisms, however, by which biological function can be carried out in the absence of a well-defined fold is not completely understood.

In this study, we have focused our attention on the acylphosphatase from the archaeon *Sulfolobus solfataricus* (Sso AcP). We will show that this protein retains an ability to function as an enzyme when adopting a non-native state in which the catalytic site is largely unstructured and flexible. Sso AcP is a 101-residue protein belonging to the acylphosphatase-like structural family. The structure of the native state of Sso AcP was recently determined by nuclear magnetic resonance spectroscopy and X-ray crystallography (Corazza *et al.*, 2006). This protein shares the same $\beta\alpha\beta\beta\alpha\beta$ topology, typical of the ferredoxin-like fold, with the other acylphosphatases so far characterized (Thunnissen *et al.*, 1997; Zuccotti *et al.*, 2004; Miyazono *et al.*, 2005; Pagano *et al.*, 2006). By contrast to related proteins, however, Sso AcP contains an unstructured, 12-residue N-terminal tail (Corazza *et al.*, 2006). Sso AcP is able to hydrolyse benzoylphosphate (BP), with k_{CAT} and K_{M} values of $198 \pm 20 \text{ s}^{-1}$ and $0.36 \pm 0.04 \text{ mM}$, respectively, at pH 5.3 and 25°C, and to be competitively inhibited by inorganic phosphate (Corazza *et al.*, 2006). The k_{CAT} value of the enzyme is low at 25°C, but increases with temperature and reaches at 81°C—the living temperature for the Archaeon *S. solfataricus*—a value close to those previously reported for the mesophilic enzymes at 25°C (Corazza *et al.*, 2006). The native state of Sso AcP is thermodynamically very stable with a free energy change of unfolding ($\Delta G_{\text{U-F}}^{\text{H}_2\text{O}}$) of $47 \pm 1 \text{ kJ mol}^{-1}$ at 37°C (Corazza *et al.*, 2006). The midpoint of thermal unfolding of the protein (T_{m}) is $100.8 \pm 4.1^\circ\text{C}$ and at 81°C the $\Delta G_{\text{U-F}}^{\text{H}_2\text{O}}$ is as high as $20.6 \pm 0.3 \text{ kJ mol}^{-1}$, similar to that of human muscle acylphosphatase at 28°C (Corazza *et al.*, 2006).

The folding mechanism of Sso AcP was described previously at pH 5.5 and 37°C (Bemporad *et al.*, 2004). After removal of the denaturant, the unfolded state of this protein collapses on the microsecond timescale into an ensemble of partially folded conformations. This ensemble is capable of binding the fluorescent dye 8-anilino-1-naphthalenesulfonic acid suggesting the presence of hydrophobic clusters exposed to the solvent and presents a far-UV mean residue ellipticity comparable to that of the fully native state, indicating that a native-like secondary structure is already formed in this state (Bemporad *et al.*, 2004). The partially folded ensemble converts into the fully folded state with a rate constant of $5.4 \pm 1.0 \text{ s}^{-1}$; a small fraction of molecules ($\sim 10\%$) folds slower with a rate constant of $\sim 0.2 \text{ s}^{-1}$ as their folding process is rate-determined by the *cis* to *trans* conversion of the Leu49–Pro50 peptide bond (Bemporad *et al.*, 2004). The presence of a relatively stable partially folded state accumulating during folding of Sso AcP offers a very favourable opportunity to study the function of a protein in a conformational state different from the native and folded one, but still populated under non-denaturing conditions.

*Corresponding author. Dipartimento di Scienze Biochimiche, Università di Firenze, viale Morgagni 50, Firenze I-50134, Italy. Tel.: +39 055 459 8319; Fax: +39 055 459 8905; E-mail: fabrizio.chiti@unifi.it

Received: 7 February 2008; accepted: 31 March 2008; published online: 1 May 2008

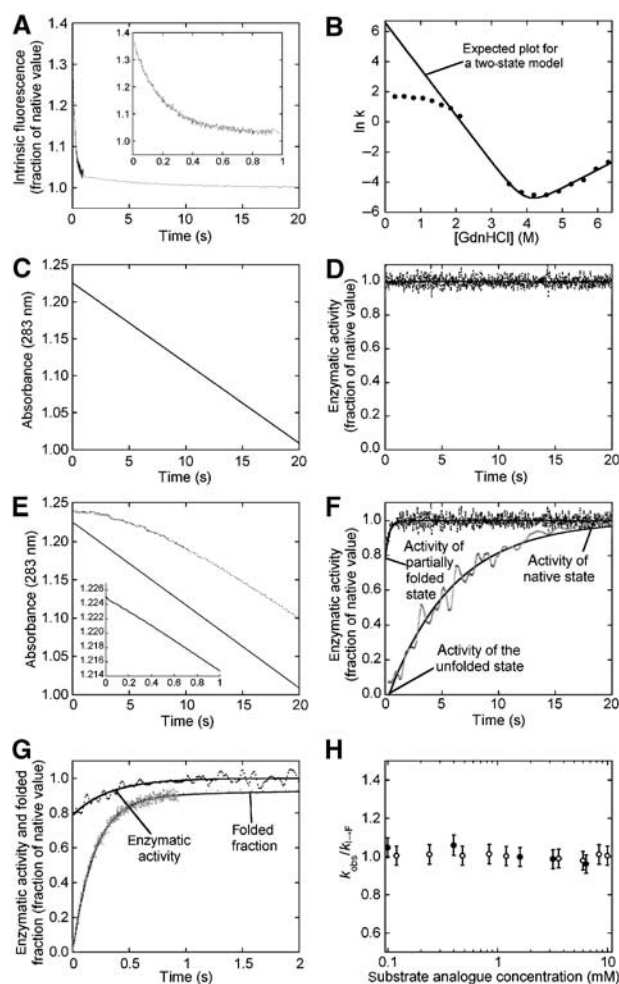


Figure 1 The partially folded ensemble of Sso AcP displays enzymatic activity. **(A)** Folding trace of Sso AcP recorded using intrinsic fluorescence as a probe in 0.275 M GdnHCl, 50 mM acetate buffer pH 5.5, 37°C. The inset shows the first second of recording. **(B)** Observed folding/unfolding rate constant versus denaturant concentration (readapted from Bemporad *et al* 2004). The continuous line represents the expected plot for a two-state model. The presence of a downward curvature at low GdnHCl concentrations, along with other experimental evidence, indicates the accumulation of a partially folded ensemble during folding. **(C)** Time course of BP absorbance recorded at 283 nm in the presence of native Sso AcP. **(D)** Time course of enzymatic activity of the native state, calculated from the trace reported in (C) (see Materials and methods). The continuous line represents the best fit to Equation (1). **(E)** BP absorbance recorded at 283 nm after dilution of GdnHCl-unfolded Sso AcP into a refolding buffer; final conditions are 0.275 M GdnHCl (continuous line) and 0.275 M GdnHCl, 7 M urea (dotted line). In both conditions, the native protein is thermodynamically more stable than any other states, making it possible to monitor the kinetics of folding. The inset shows the first second of recording. **(F)** Development of enzymatic activity, calculated from the traces reported in (E) (see Materials and Methods), during the refolding of wild-type Sso AcP in the absence (upper trace) and presence (lower trace) of 7 M urea. The continuous lines represent the best fits to Equation (1). The activities of fully folded, partially folded and unfolded states are shown. **(G)** A comparison between the time courses (first 2 s) of recovery of native conformation and enzymatic activity. After 3.6 ms, when the partially folded ensemble is populated more than 99%, Sso AcP exhibits 79% of the native enzymatic activity. **(H)** Ratio between the main folding rate constant recorded in the presence of phosphate (○) or phenyl phosphate (●) (k_{obs}) and that recorded without substrate-analogue in the sample ($k_{1 \rightarrow F}$) plotted versus substrate-analogue concentration.

In this work, the functional properties of the partially folded state of Sso AcP accumulating during folding are investigated using a procedure that allows the recovery of enzymatic activity during folding to be determined in real time (Chiti *et al*, 1999). The protein engineering method and Φ -value analysis (Matouschek *et al*, 1989) are then used to obtain information on the degree of structure formation, at the level of the mutated residues, in both the partially folded and transition states of the protein. We shall show that the partially folded state of Sso AcP accumulating during folding is enzymatically active despite structural heterogeneity at the level of the active site. In addition, molecular dynamics (MD) simulations using Φ -values as restraints illustrate how this state is made up by an ensemble of conformations displaying native-like topology. This forces catalytic residues to be in close proximity and allows this conformational state to retain enzymatic activity.

Results

The partially folded state of Sso AcP populated during folding exhibits enzymatic activity

We first monitored the folding process of Sso AcP using intrinsic fluorescence as a spectroscopic probe. This protein possesses one tryptophan in the N-terminal segment and seven tyrosines at various positions along the sequence, mainly positioned in the β -sheet (Supplementary Figure S1). Figure 1A shows the change of intrinsic fluorescence when one volume of Sso AcP unfolded in 5.5 M guanidinium hydrochloride (GdnHCl) is mixed with 19 volumes of refolding buffer. In this trace, which is in very good agreement with that previously reported (Bemporad *et al*, 2004), three phases are observed. A first rapid increase of fluorescence, occurring on a timescale shorter than 10 ms and escaping experimental detection with the stopped-flow device utilized here, was shown to correspond to the conversion of the fully unfolded state into the partially folded state (Bemporad *et al*, 2004). The following rapid decrease of fluorescence, with a rate constant ($k_{1 \rightarrow F}$) of $5.3 \pm 1.0 \text{ s}^{-1}$ and a relative amplitude of $90 \pm 2\%$, was shown to correspond to the conversion of this state into the fully native conformation, whereas the second slower decrease, with a rate constant of $0.18 \pm 0.04 \text{ s}^{-1}$ and a relative amplitude of $10 \pm 2\%$, arises from the *cis-trans* isomerization of a small fraction of protein molecules with the Leu49-Pro50 peptide bond initially in a non-native *cis* configuration. The downward curvature observed at low GdnHCl concentrations in the plot reporting $\ln(k_{1 \rightarrow F})$ versus denaturant concentration (Figure 1B) was shown to arise from the formation of the partially folded ensemble in the dead time of the stopped-flow experiment (Bemporad *et al*, 2004).

The time course of recovery of enzymatic activity has then been studied in real time during Sso AcP folding. The substrate BP, unlike its hydrolysis products benzoate and phosphate, has a significant optical absorption at 283 nm. By measuring the decay of absorbance at 283 nm in small and sequential time intervals Δt , in the presence of Sso AcP undergoing the folding process, the time course of enzymatic activity during folding can be reconstructed (see Materials and methods for details). In a first control experiment, a solution containing native Sso AcP has been mixed with the refolding buffer containing BP and a non-denaturing concentration of GdnHCl. Final conditions were 0.02 mg ml^{-1} pro-

tein, 5 mM BP, 0.275 M GdnHCl, 50 mM acetate buffer, pH 5.5, 37°C. The change of absorbance at 283 nm has been monitored in real time (Figure 1C) and analysed to yield the time course of enzymatic activity (Figure 1D). In this control experiment, the protein is expected to remain native all the way through. Indeed, the absorbance at 283 nm arising from BP decays linearly with time (Figure 1C) and the enzymatic activity remains high and constant during the 20 s of recording (Figure 1D), as a result of ongoing catalysis.

To monitor the time course of the enzymatic activity during Sso AcP refolding, a sample of GdnHCl-unfolded Sso AcP has been mixed with the refolding buffer containing BP. Final conditions were the same as in the first experiment. Figure 1E and F show the time-dependent changes of absorbance at 283 nm and of the corresponding enzymatic activity, respectively. Immediately after mixing, when the partially folded state is maximally populated, enzymatic activity is already present, corresponding to $79 \pm 10\%$ of that of the native protein under the same conditions (Figure 1F). The activity then shows a small increase with a rate constant of $3.5 \pm 1.5 \text{ s}^{-1}$, in good agreement with the $k_{I \rightarrow F}$ value determined with intrinsic fluorescence for the conversion of the partially folded state into the folded state (Figure 1F). The time-dependent changes of intrinsic fluorescence and enzymatic activity in the first two seconds of recording are summarized in Figure 1G. This figure emphasizes that at the beginning of the folding trace, when the fraction of native protein is equal to 0 and the protein populates the partially folded ensemble, the activity is equal to 79% of the native protein.

The early enzymatic activity of Sso AcP is not contributed by the unfolded or native states

The GdnHCl-unfolded protein has been diluted into a solution containing urea to a final concentration of 7 M. Owing to the high conformational stability of Sso AcP, these conditions are not yet denaturing and refolding is possible. However, the plot reporting $\ln(k)$ versus urea concentration shows a downward curvature in the range of 0–5 M, indicating that in 7 M urea, the partially folded state is destabilized and the protein folds according to a two-state model (Supplementary Figure S2). The results show that enzymatic activity is absent immediately after mixing, when only the unfolded state is populated (Figure 1F). Upon refolding, the activity then increases with a rate constant of $0.08 \pm 0.02 \text{ s}^{-1}$ (Figure 1F) corresponding to the folding rate constant under these conditions (Supplementary Figure S2).

To rule out that a substantial fraction of the native protein is present after the dead time of the stopped-flow experiment, when enzymatic activity is observed already, a double jump experiment has been carried out (Supplementary Figure S3). In this experiment, the GdnHCl-unfolded protein has been diluted into the refolding buffer (first jump) and then, after 10 ms, transferred again to solutions containing GdnHCl at final concentrations ranging from 4.2 to 7.0 M (second jump). Although such final conditions promote the unfolding of native Sso AcP and produce a single exponential change of intrinsic fluorescence (Bemporad *et al*, 2004), no significant fluorescence changes have been observed in any of these experiments, suggesting that the native protein is not present 10 ms after the folding process has been initiated (Supplementary Figure S3). This result indicates that the

enzymatic activity observed at 10 ms does not arise from Sso AcP adopting a native conformation.

Sso AcP refolding has also been followed using intrinsic fluorescence in the presence of phosphate and phenyl phosphate, two competitive inhibitors of Sso AcP that are stable analogues of BP. Experimental conditions were the same as for the enzymatic-activity experiments. Neither the first nor the second folding rate constants are affected by either compound (Figure 1H). These findings rule out the possibility that folding of Sso AcP is accelerated by BP and that the enzymatic activity observed at the beginning of the folding process is due to an early substrate-induced folding of the protein. Taken together, these data suggest the idea that the partially folded ensemble accumulating during folding of Sso AcP possesses significant enzymatic activity.

The enzymatic activity observed in the Sso AcP partially folded state is highly sensitive to mutations

The recovery of enzymatic activity during folding has also been recorded for a set of Sso AcP mutants. Several mutants, such as the K92A variant, show a behaviour similar to that of the wild-type protein with enzymatic activity detected for both the partially folded and native states (Figure 2A and B, Table I). The traces recorded for the N48A and R30A variants show full inactivation of their partially folded and native states (Figure 2C and D, Table I). This result confirms the key role of these two residues in the catalysis of the native state of acylphosphatases and suggests their major role in the catalytic mechanism of the partially folded state as well (Stefani *et al*, 1997). The partially structured ensembles of the V24A and V27A mutants do not show residual enzymatic activity (Figure 2E and F, Table I). In the case of V24A, the enzymatic activity increases with a rate constant of $3.2 \pm 1.0 \text{ s}^{-1}$, a value that is, within experimental error, similar to that measured by following folding of the mutant using intrinsic fluorescence (Table II). This observation indicates that the observed recovery of enzymatic activity in the V24A variant is due to the conversion of the partially folded ensemble into the native state during folding. Finally, two variants with mutations within the 49–52 loop have been investigated. Similarly to V24A, both P50A and G52A variants display fully inactive, partially folded states, but native structures with significant activity that are recovered with rate constants highly consistent with those measured for folding using intrinsic fluorescence (Figure 2G and H, Tables I and II). These data show that the ability to hydrolyse the substrate of the partially folded ensemble is more sensitive to mutations than the native state and provide insight into regions of the structure that are more important for enabling enzymatic activity.

Investigation of the partially folded and transition states of Sso AcP using Φ -value analysis

To characterize the partially folded and transition states of Sso AcP, we have carried out a Φ -value analysis using 34 single mutants (Matouschek *et al*, 1989, 1992). This method introduces the Φ -value for a certain mutation in the partially folded state Z (Φ^Z) according to the equation: $\Phi^Z = \Delta\Delta G_{Z-U} / \Delta\Delta G_{F-U}$, where $\Delta\Delta G_{Z-U}$ and $\Delta\Delta G_{F-U}$ represent the free energy changes upon mutation of the partially folded and native states, respectively. The rationale behind this approach is that the elimination of targeted atoms from the side chain of a residue already structured in the partially folded state Z

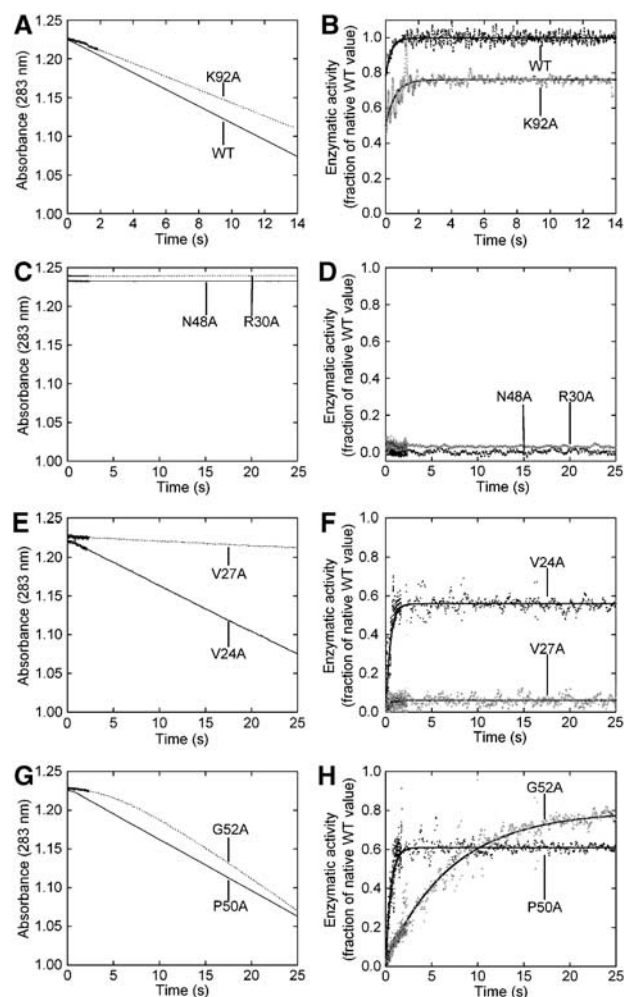


Figure 2 The enzymatic activity of the partially folded state of Sso AcP is sensitive to mutations. (A, C, E, G) BP absorbance at 283 nm after dilution of the indicated GdnHCl-unfolded mutants into a refolding buffer. (B, D, F, H) Development of enzymatic activity, calculated from the traces reported in the corresponding panels on the left (see Materials and methods), during the refolding of the indicated Sso AcP mutants. The continuous lines represent the best fits to Equation (1). All traces are labelled to indicate the mutants to which they refer.

will affect the conformational stabilities of both native and partially folded states to the same extent, resulting in a Φ^Z value equal to 1. Conversely, any side-chain change of a residue unstructured in the partially folded ensemble does not affect its conformational stability, resulting in a Φ^Z value equal to 0. Mutations have been chosen to probe (1) the hydrophobic core, (2) the active site (four of the investigated mutations, namely R30A, N48A, V24A and V27A, involve the catalytic site and have been found here to abolish or decrease dramatically the enzymatic activity of Sso AcP), (3) the salt bridges that are present on the protein surface and contribute to the enhanced structural stability of the protein (Corazza *et al*, 2006).

An equilibrium GdnHCl-induced unfolding curve has been acquired for each mutant to yield the change in conformational stability upon mutation ($\Delta\Delta G_{U-F}^{H_2O}$). Conditions were 50 mM acetate buffer, pH 5.5, 37°C. Figure 3A shows representative equilibrium unfolding curves for the wild-type

protein, the only mutant found to be stabilized relative to the wild type (R71A), and three destabilized mutants (V20A, A46G, L65A). All plots have been analysed with Equation (2). The results show that several mutants are destabilized (Table II).

The destabilized (or stabilized) mutants with $\Delta\Delta G_{U-F}^{H_2O}$ values higher than 3.2 kJ mol^{-1} or lower than -3.2 kJ mol^{-1} have been analysed to obtain the Φ -values of the corresponding mutations for the partially folded and transition states. For each of these mutants, kinetic traces for folding and unfolding have been acquired at various denaturant concentrations, using intrinsic fluorescence and far-UV ellipticity as probes for folding and unfolding, respectively. Figure 3B and C show representative traces for folding and unfolding, respectively. All mutants showed, at low denaturant concentrations, a downward curvature in the folding limb of the plot reporting the folding/unfolding rate constant versus denaturant concentration. This deviation from the two-state model is similar to that observed for the wild-type protein (Figure 1B) and suggests that the partially folded ensemble forms in all mutants. For each mutant, the kinetic traces have been analysed as described in the Materials and methods, to yield the folding ($k_{1 \rightarrow F}^{H_2O}$) and unfolding ($k_{F \rightarrow U}^{H_2O}$) rate constants in the absence of denaturant ($k_{1 \rightarrow F}^{H_2O}$ refers to the rate of formation of the native state regardless of the on- or off-pathway nature of the partially folded state). A list of the $k_{1 \rightarrow F}^{H_2O}$ and $k_{F \rightarrow U}^{H_2O}$ values obtained for all mutants is reported in Table II. The thermodynamic and kinetic data have been combined to obtain the Φ -values for the partially folded and transition states. The Φ -values for the partially folded ensemble (Φ^1) are generally lower than the corresponding ones for the transition state (Φ^\ddagger), showing a gain of structure along the folding coordinate (Table II).

In the partially folded state, the catalytic site does not appear to be fully structured. The R30A and N48A variants have not been analysed, due to their low $\Delta\Delta G_{U-F}^{H_2O}$. Nevertheless, the Φ^1 values obtained for the V24A and V27A variants are close to 0, suggesting that the catalytic 22–28 loop does not display a native-like structure in this ensemble. These data suggest that the catalytic properties of the partially folded ensemble do not result from a native-like structure of the active site in this state. It cannot be excluded that the substrate directly promotes the appropriate positioning of the catalytic residues in the partially folded state and hence determines a local folding of the active site. Nevertheless, the control experiments described above using substrate analogues indicate that substrate-induced folding cannot involve the whole Sso AcP structure.

The analysis of all mutants used in this study enables the folding pathway of Sso AcP to be characterized. The region that appears most structured in the ensemble of partially folded conformations is the interface between β -strand 1 and α -helix 2 (Figure 4, Table II). The rest of the molecule shows Φ -values close to 0. The transition state displays a more compact structure, with generally increased Φ -values (Figure 4; Table II). In this state, the establishment of native contacts is propagated to the β -hairpin formed by β -strands 2 and 3. Four residues (Val20, Gly52, Ala58 and Arg71) appear to drive structure formation in the transition state ensemble (Figure 4). Interestingly, Arg71 does not appear important for the overall stabilization of the native structure, suggesting a specific role for this residue in transition state formation.

Table I Catalytic parameters for a set of Sso AcP variants measured in 50 mM acetate buffer at pH 5.5, 37°C using BP as a substrate

Variant	Native state k_{CAT} (s ⁻¹)	Partially folded state k_{CAT} (s ⁻¹) ^a	Activity in the partially folded state (% of that from the native state) ^b
WT	222 ± 20	175 ± 30	79 ± 10
V24A	124 ± 10	7 ± 5	5.6 ± 10
V27A	13 ± 1	0 ± 10	0.0 ± 10
R30A	7 ± 1	ND	ND
N48A	2 ± 2	ND	ND
P50A	137 ± 14	0 ± 10	0.0 ± 10
G52A	178 ± 18	6 ± 5	3.3 ± 10
K92A	169 ± 20	111 ± 30	66 ± 10

^aValues determined from the best fits of the traces reported in Figure 2 to Equation (1).

^bValues obtained by combining data of columns 1 and 2.

Table II Thermodynamic and kinetic parameters of (un)folding for a set of Sso AcP mutants

Variant	C_m (M)	m (kJ mol ⁻¹ M ⁻¹)	$\Delta\Delta G_{U-F}^{H_2O}$ (kJ mol ⁻¹) ^a	$k_{U-F}^{H_2O}$ (s ⁻¹)	$k_{F-U}^{H_2O}$ (s ⁻¹)	Φ^I	Φ^\ddagger
WT	4.23 ± 0.07	11.3 ± 1.1	—	5.436 ± 0.272	(6.10 ± 0.30) · 10 ⁻⁶	—	—
<i>Hydrophobic core residues</i>							
M16A	3.10 ± 0.07	11.0 ± 1.1	12.59 ± 1.13	7.179 ± 0.359	(2.04 ± 0.10) · 10 ⁻⁴	0.34 ± 0.06	0.28 ± 0.07
A18G	3.21 ± 0.07	10.3 ± 1.0	11.36 ± 1.12	4.249 ± 0.212	(3.21 ± 0.16) · 10 ⁻⁵	0.57 ± 0.05	0.62 ± 0.04
V20A	3.78 ± 0.07	11.7 ± 1.2	5.01 ± 1.11	1.477 ± 0.074	(3.3 ± 0.17) · 10 ⁻⁶	0.64 ± 0.09	1.31 ± 0.08
F29L	3.58 ± 0.07	10.9 ± 1.1	7.24 ± 1.11	6.421 ± 0.321	(1.47 ± 0.07) · 10 ⁻⁴	-0.07 ± 0.17	-0.13 ± 0.18
I42V	4.11 ± 0.07	11.5 ± 1.1	1.34 ± 1.10	ND	ND	ND	ND
A46G	3.07 ± 0.07	9.9 ± 1.0	12.92 ± 1.13	0.855 ± 0.043	(1.66 ± 0.083) · 10 ⁻⁴	-0.03 ± 0.09	0.34 ± 0.06
V54A	3.11 ± 0.07	10.6 ± 1.1	12.47 ± 1.13	0.754 ± 0.038	(1.81 ± 0.09) · 10 ⁻⁴	-0.11 ± 0.10	0.30 ± 0.06
A58G	3.89 ± 0.07	10.6 ± 1.1	3.79 ± 1.10	1.567 ± 0.078	(6.14 ± 0.31) · 10 ⁻⁶	0.15 ± 0.26	1.00 ± 0.05
Y61A	4.18 ± 0.07	14.5 ± 1.4	0.56 ± 1.10	ND	ND	ND	ND
Y61L	4.18 ± 0.07	14.0 ± 1.4	0.56 ± 1.10	ND	ND	ND	ND
L65A	2.59 ± 0.07	10.3 ± 1.0	18.27 ± 1.15	2.354 ± 0.118	(2.78 ± 0.14) · 10 ⁻⁴	0.34 ± 0.04	0.46 ± 0.04
L68A	2.54 ± 0.07	10.0 ± 1.0	18.82 ± 1.15	1.064 ± 0.053	(3.35 ± 0.17) · 10 ⁻⁴	0.23 ± 0.05	0.45 ± 0.04
I72V	3.78 ± 0.07	11.8 ± 1.2	5.01 ± 1.11	4.364 ± 0.218	(2.22 ± 0.11) · 10 ⁻⁵	0.22 ± 0.18	0.34 ± 0.15
P76A	3.73 ± 0.07	12.1 ± 1.2	5.57 ± 1.11	5.003 ± 0.250	(3.83 ± 0.19) · 10 ⁻⁴	-0.96 ± 0.39	-0.92 ± 0.38
V81A	3.13 ± 0.07	10.0 ± 1.0	12.25 ± 1.12	4.761 ± 0.238	(7.72 ± 0.38) · 10 ⁻⁴	-0.05 ± 0.10	-0.02 ± 0.09
V84A	3.26 ± 0.07	10.6 ± 1.1	10.80 ± 1.12	4.729 ± 0.236	(2.98 ± 1.49) · 10 ⁻⁴	0.04 ± 0.10	0.07 ± 0.10
F88A	2.92 ± 0.07	12.7 ± 1.3	14.59 ± 1.13	5.176 ± 0.259	(2.01 ± 0.10) · 10 ⁻³	-0.03 ± 0.08	-0.02 ± 0.08
G93A	4.05 ± 0.07	10.8 ± 1.1	2.00 ± 1.10	ND	ND	ND	ND
<i>Residues important in catalysis</i>							
V24A ^b	3.28 ± 0.07	9.7 ± 1.0	10.58 ± 1.12	4.290 ± 0.214	(1.94 ± 0.10) · 10 ⁻⁴	0.10 ± 0.10	0.16 ± 0.09
V27A ^b	3.78 ± 0.07	11.1 ± 1.1	5.01 ± 1.11	6.473 ± 0.324	(4.43 ± 0.22) · 10 ⁻⁵	0.07 ± 0.21	-0.02 ± 0.23
R30A ^c	4.41 ± 0.07	12.2 ± 1.2	-2.00 ± 1.10	ND	ND	ND	ND
N48A	4.34 ± 0.07	11.2 ± 1.1	-1.23 ± 1.10	ND	ND	ND	ND
<i>Salt bridge residues</i>							
R15A	3.88 ± 0.07	11.9 ± 1.2	3.90 ± 1.10	3.012 ± 0.151	(2.07 ± 0.10) · 10 ⁻⁵	-0.20 ± 0.35	0.19 ± 0.23
R19A	3.56 ± 0.07	10.4 ± 1.0	7.46 ± 1.11	0.652 ± 0.033	(2.25 ± 0.11) · 10 ⁻⁵	-0.18 ± 0.18	0.55 ± 0.07
E59A	3.52 ± 0.07	12.0 ± 1.2	7.91 ± 1.11	2.343 ± 0.117	(5.49 ± 0.27) · 10 ⁻⁵	0.01 ± 0.14	0.28 ± 0.10
R71A	4.52 ± 0.07	12.6 ± 1.3	-3.23 ± 1.10	10.090 ± 0.504	(7.01 ± 0.35) · 10 ⁻⁶	0.62 ± 0.15	1.11 ± 0.07
<i>Other residues</i>							
A37G	3.36 ± 0.07	10.1 ± 1.0	9.69 ± 1.12	4.996 ± 0.250	(1.43 ± 0.07) · 10 ⁻⁴	0.14 ± 0.10	0.16 ± 0.10
L49A	3.73 ± 0.07	11.2 ± 1.1	5.57 ± 1.11	1.422 ± 0.071	(1.49 ± 0.07) · 10 ⁻⁵	-0.03 ± 0.21	0.59 ± 0.09
P50A	4.20 ± 0.07	11.8 ± 1.2	0.33 ± 1.10	ND	ND	ND	ND
G52A	3.34 ± 0.07	11.3 ± 1.1	9.91 ± 1.12	0.181 ± 0.009	(1.21 ± 0.06) · 10 ⁻⁵	-0.06 ± 0.12	0.82 ± 0.03
P77A	4.19 ± 0.07	10.1 ± 1.0	0.45 ± 1.10	ND	ND	ND	ND
S89A	4.17 ± 0.07	11.2 ± 1.1	0.67 ± 1.10	ND	ND	ND	ND
K92A	4.07 ± 0.07	10.1 ± 1.0	1.78 ± 1.10	ND	ND	ND	ND
F98L	3.01 ± 0.07	8.2 ± 0.8	13.59 ± 1.13	4.211 ± 0.211	(7.61 ± 3.80) · 10 ⁻⁴	0.04 ± 0.08	0.08 ± 0.08

^aThe $\Delta\Delta C_{U-F}^{H_2O}$ have been calculated according to Equation (2) using the average of the m values for all mutants.

^bThese residues also belong to the group of 'hydrophobic core residues'.

^cThese residues also belong to the group of 'Salt bridge residues'.

Structure of the partially folded and transition states of Sso AcP

We have used MD simulations using the experimentally determined Φ -values as restraints in the simulations (Vendruscolo *et al*, 2001; Gsponer *et al*, 2006) to generate two ensembles of structures representing the partially folded

and transition states, respectively (See Materials and methods and Supplementary data). This method has been validated in the case of barnase where the resulting structures were used to predict with high accuracy the results of double-mutant cycle experiments that were not used as restraints in those calculations (Salvatella *et al*, 2005). The resulting structural

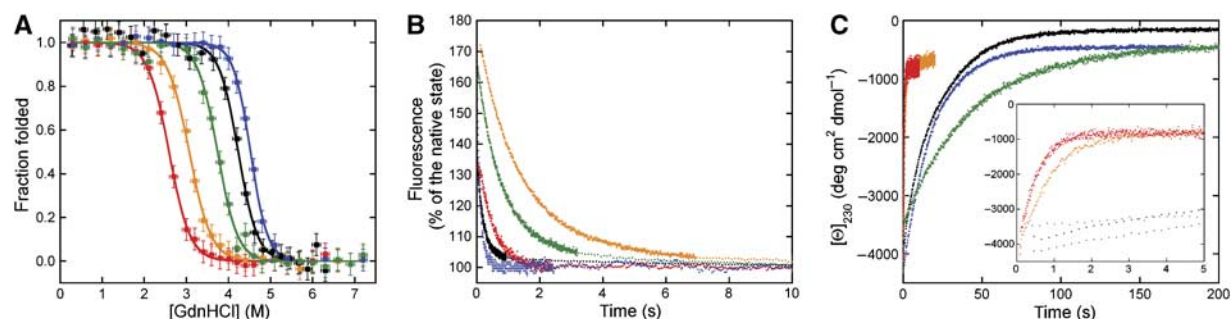


Figure 3 Thermodynamics and kinetics of (un)folding for Sso AcP variants. **(A)** Equilibrium unfolding curves for a set of Sso AcP variants in 50 mM acetate buffer, pH 5.5, 37°C. The continuous lines represent the best fits to the equation reported by Santoro and Bolen (Santoro and Bolen, 1988). The obtained parameters of conformational stability are reported in Table II. **(B)** Folding traces recorded in 0.275 M GdnHCl, 50 mM acetate buffer at pH 5.5, 37°C. **(C)** Unfolding traces recorded in 6 M GdnHCl, 50 mM acetate buffer at pH 5.5, 37°C. The inset shows the first 5 s of recording. In all plots, the traces refer to wild-type (black), R71A (blue), A46G (orange), L65A (red) and V20A (green) Sso AcP.

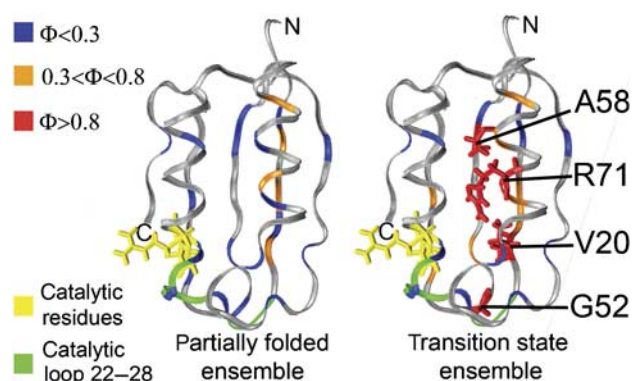


Figure 4 Native Sso AcP color-coded to show the obtained Φ^1 (left) and Φ^2 (right) values. Residues are shown in blue if their Φ -value is lower than 0.3, orange if the value is between 0.3 and 0.8 and red if the value is greater than 0.8. In the latter case, the residue is labelled and shown also in a stick representation. The two catalytic residues (Arg30 and Asn48) are indicated in yellow. The catalytic loop 22–28 is indicated in green.

ensembles do not depend on whether the starting conformational state used in the simulations is the native or unfolded state (Settanni *et al*, 2004).

From this analysis, the partially folded ensemble of Sso AcP appears to be characterized by a significant structural heterogeneity and by the presence of several non-native interactions (Figure 5A). The native topology is, however, rather well preserved, as shown by the comparison of the structures of the partially folded and native states (Figure 5B) and by the energy maps, which provide an illustration of the most strongly interacting regions (Figure 5C). The catalytic site (formed by residues in α -helix 1, β -strand 2 and the loop between β -strand 1 and α -helix 1) is very flexible in the modelled ensemble, although the Arg30 and Asn48 residues, which are the most important for the catalytic activity, remain closer than 6 Å to each other in more than 20% of the structures in the ensemble. The structural flexibility determined for the catalytic site does not result from the absence of restraints in this region, but rather from Φ -values close to 0 determined for this region and used here as restraints in the simulations.

Overall, the native architecture is particularly well conserved in the region of α -helix 2 and β -strands 1 and 4. By contrast, the regions corresponding to α -helix 1, β -strand 2,5

and 3 are much less well structured. To quantify these observations, we have divided the protein structure in two parts, with the first corresponding to α -helix 1, β -strands 2, 5 and 3 (catalytic region) and the second one corresponding to α -helix 2, β -strands 1 and 4. We have then calculated the probability distributions of the C α -carbon root mean square distance from the native state (C α -RMSD) (Figure 6A and B). The C α -RMSD value of a given residue provides information on the extent to which that residue in the partially folded state is spatially distant from its position in the native state. Low and high values of C α -RMSD value indicate structural similarity and conformational heterogeneity, respectively.

The results reported in Figure 6A show that the high heterogeneity of the partially folded state is due to the catalytic region (blue), which shows C α -RMSD values comparable to those of the entire protein (black). By contrast, the remainder of the molecule (red) virtually shows a native-like fold, with an average distance from the native state equal to about 2 Å. We refer to the latter region as the 'scaffold region'. The presence of this scaffold implies that, although the catalytic residues and the catalytic loop are highly dynamic in the partially folded ensemble, the overall topology of the protein is already formed in this state. This decreases the number of its accessible conformations, that is its entropy. Thus, the structure determination that we present here provides a result that is not apparent from an immediate inspection of the experimental Φ -values that are all small in this region. This analysis provides a structural basis for rationalizing the maintenance of the catalytic activity in the partially folded state.

For comparison, the transition state ensemble is much more native like and comprises a particularly well-structured region that includes parts of β -strands 1, 2, 3 and 4, α -helices 1 and 2, bearing a more folded catalytic site (Figures 5B,C and 6B). The region around the catalytic residue Asn48 is structurally very different from the native state in partially folded ensemble and becomes more native like in the transition state ensemble (Figure 6C).

Discussion

Enzymatic activity in the presence of a highly dynamic catalytic site

We have shown that the partially structured ensemble of Sso AcP accumulating during folding before the major energy

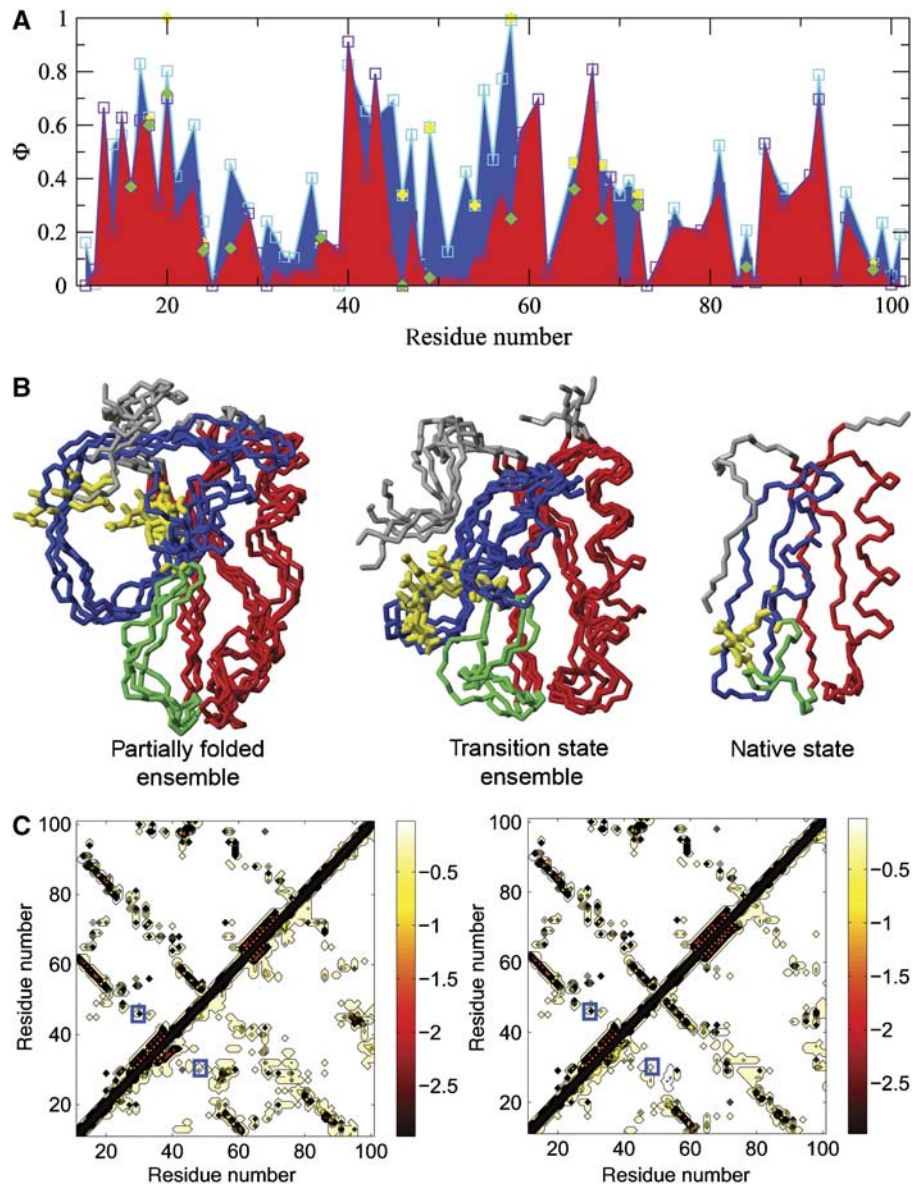


Figure 5 Structural properties of the partially folded and transition state ensembles of Sso AcP. **(A)** Profiles of the experimentally determined Φ -values (Φ^{exp}) and those calculated from the simulations (Φ^{calc}) for partially folded and transition state ensembles. The ensemble average Φ^{calc} values of the partially folded and transition state ensembles are shown as purple squares (with red area) and cyan squares (with blue area), respectively. Φ^{exp} values of the partially folded and transition state ensembles are indicated as green and yellow diamonds, respectively. **(B)** Comparison of the structures of the four biggest clusters in the partially folded ensemble (left) and transition state ensemble (middle) with the X-ray structure of the native state (right); the ‘scaffold region’ (residues 13–23 and 60–90) is shown in red, the region around the catalytic site (residues 24–59) in blue; residues Arg30 and Asn48 are highlighted in yellow; the catalytic loop 22–28 is indicated in green. **(C)** Energy maps of partially folded (left) and transition state (right) ensembles. The pairwise interaction energies of the native state are shown above the diagonal, those of the partially folded and transition state ensembles below the diagonal. The interaction between Arg30 and Asn48 is highlighted by blue squares.

barrier to the native state is enzymatically active. Control experiments have shown that the observed enzymatic activity does not arise from a fraction of native protein present in the sample, nor does it derive from a substrate-induced folding of the partially folded state. The catalytic site of Sso AcP appears to be highly heterogeneous in the partially folded ensemble. This conclusion follows from the observation that the two Φ -values of residues in the catalytic 22–28 loop are close to 0. The Φ -value of residue 29, which is near to both the catalytic 22–28 loop and the substrate-binding residue Arg30, is also close to 0. Moreover, the overall structural analysis carried out with all of the experimentally determined

Φ -values confirms that the catalytic site of the protein is not yet fully structured. As the region of the catalytic site is exhibiting significant fluctuations, the memory of the initial structure used in the MD simulations is rapidly lost and the interactions that we identified are not resulting from the starting conformation. We cannot exclude that the catalytic site becomes structured upon binding of the ligand. Importantly, however, folding of the entire protein is not accelerated in the presence of substrate analogues, suggesting that no global substrate-induced folding occurs.

These findings indicate that an enzyme can be an efficient catalyst even in the absence of its stable native conformation

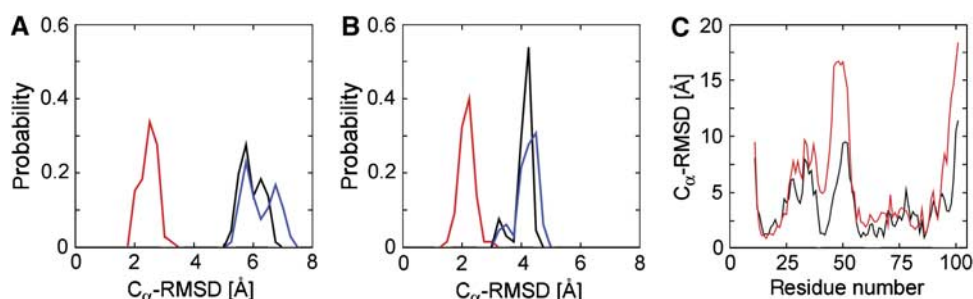


Figure 6 The structural heterogeneity of the partially folded state is due to the catalytic region. (A, B) Distribution of the C α -RMSD values for the entire sequence (black), the scaffold region (red) and the catalytic region (blue) from the X-ray structure in the partially folded (A) and transition state (B) ensembles. (C) The C α -RMSD values of the partially folded (red) and transition state (black) ensembles from the native state are plotted versus residue number.

and provide clues to the characterization of the structural and dynamic features of a protein that allow the catalysis to take place in the absence of a fully structured catalytic site. Indeed, we have shown that the presence of a scaffold region, whose structure is already formed in the partially folded state, determines the topology of the entire molecule; thus, the catalytic residues, although highly dynamic, remain in close proximity and can hydrolyze the substrate. The activity of the partially folded state is highly sensitive to mutations, suggesting that, in the partially structured ensemble, a small number of native contacts stabilize the overall structure and enable the protein to hydrolyse the substrate; substitutions may therefore cause a higher degree of flexibility and allow a full inactivation of the enzyme. The importance of conformational changes and flexibility in enzyme catalysis has been widely recognized (Osborne *et al*, 2001; Eisenmesser *et al*, 2002; Benkovic and Hammes-Schiffer, 2003; Poulsen *et al*, 2003; Garcia-Viloca *et al*, 2004). The detection of enzymatic activity in the absence of a structured catalytic site and the ability to determine the distribution of structures in these non-native states represent a key step forward in the elucidation of the protein dynamics that are required for enzyme catalysis.

In addition, the identification of a highly dynamic functional site in the presence of a scaffold region that restricts the conformational space of the flexible region suggests a potentially important concept in molecular biology. It has been shown, for example, that it is possible to change the catalytic activity of an existing protein scaffold by substituting several loops and then introducing point mutations to tune the enzyme activity (Park *et al*, 2006). Moreover, it has been proposed that the order of formation of native contacts during folding recapitulates the emergence of topology in molecular evolution (Nagao *et al*, 2005). The fact that the catalytic site of Sso AcP is highly heterogeneous in the partially folded ensemble and folds afterwards—that is, concomitantly to the formation of the native state (Figure 6)—indicates the presence of an efficient evolutionary mechanism in which the scaffold of a protein is maintained and the regions or residues directly involved in function (for example, in catalysis or binding) are allowed to mutate without compromising the overall stability of the structure. This mechanism may help to increase the rate of development of new activities and could explain how proteins with the same topology can possess very different functions.

Catalysis in the absence of a well-defined structured fold is more common than previously thought

One question that naturally arises at this point is whether enzyme catalysis in the absence of well-defined folded structure is a unique, strange feature of Sso AcP and possibly other few cases, or rather a common yet unappreciated property of enzymes. A few other systems have been shown to bear enzymatic activity in the absence of folded structure. An intermediate in the folding of ribonuclease T1 is characterized by extensive secondary and tertiary structure, a hydrophobic core with low solvent accessibility and partial enzymatic activity (Kiefhaber *et al*, 1992). A monomeric chorismate mutase obtained by topological redesign of a dimeric helical bundle enzyme from *Methanococcus jannaschii* shows properties typical of a molten globule state. Such a conformational state undergoes ordering upon addition of a substrate analogue indicating that a substrate-induced folding can transform the protein into an enzymatically active state (Vamvaca *et al*, 2004). In a more recent report, it was shown that the complex between the molten globule state and the ligand retains a considerable flexibility, indicating that efficient catalysis can be coupled with global conformational plasticity (Pervushin *et al*, 2007). In another study, two variants of dihydrofolate reductase were shown to adopt molten globule features; in the presence of methotrexate and NADPH, they adopt a native-like structure and become enzymatically active (Uversky *et al*, 1996). These reports suggest that catalysis is not necessarily coupled to a stable and persistent fold and that an ability to carry out catalysis outside the native state is a characteristic more common than previously thought. The study presented here offers structural insight to explain how catalysis can occur in the absence of the native state.

Enzyme catalysis in the absence of a native state: possible biological implications

It has been proposed that a considerable fraction of eukaryotic proteins are either fully unstructured or contain significant portions of their sequence—that is, regions longer than 50 residues—in an unstructured state (Dunker *et al*, 2001; Fink, 2005). Intriguingly, these intrinsically disordered polypeptide chains are mostly involved in fundamental biological processes, such as transcription, translation and regulation of the cell cycle (Nakayama *et al*, 2001). It has been suggested that the high flexibility of intrinsically disordered

proteins brings many advantages, including binding to many substrates and targets (Dunker *et al*, 2005), coupling high specificity with low affinity (Dunker *et al*, 2001), enlarging interaction surfaces (Dunker *et al*, 2005), increasing rapid turnover and thus rapid response to cell signaling in fundamental points of the protein network (Fink, 2005).

The widespread presence of intrinsically disordered proteins in living organisms suggests that proteins do not have to adopt necessarily compact globular structures to be functional, at least for the molecular recognition of their targets. Although Sso AcP is not an intrinsically disordered protein and its partially folded state accumulating during folding is not as disordered as a natively unfolded protein, the evidence reported here indicates that a conformational state structurally distant from the folded structure—particularly in those loops or residues that form the substrate binding and catalytic site—is able to bind substrates, carry out catalysis and release products. It shows that even enzyme catalysis does not require a fully folded state and opens new perspectives in the exploration of the biological significance of enzymatically active non-native states. More generally, these findings suggest a re-consideration of the paradigm that biological function can only be associated with unique three-dimensional folds.

Materials and methods

Purification of wild-type and mutant Sso AcP

The Sso AcP gene was initially inserted in a pGEX-2T plasmid (Amersham, Little Chalfont, England) and the resulting construct was used to transform DH5 α *E. coli* cells (Invitrogen, Carlsbad, CA, USA). Mutated genes were obtained using the QuickChange site-directed mutagenesis kit from Stratagene (La Jolla, CA, USA). The presence of the desired mutations was assessed by DNA sequencing. Expression and purification of the wild-type and mutant proteins were carried out according to the protocol of the pGEX-2T manufacturer. Protein purity was checked by SDS-polyacrylamide gel electrophoresis. The extinction coefficient at 280 nm (ϵ_{280}) for each mutant was calculated as described (Gill and von Hippel, 1989).

Enzymatic activity assay

Enzymatic activity of native Sso AcP was measured in a continuous optical test at 283 nm using BP as a substrate (Ramponi *et al*, 1966) with a Lambda 4V Perkin Elmer spectrophotometer (Wellesley, MA, USA). Experimental conditions were 2.0 $\mu\text{g ml}^{-1}$ Sso AcP, 5.0 mM BP, 50 mM acetate buffer at pH 5.5, 37°C. BP was synthesized as described previously (Camici *et al*, 1976) and freshly dissolved before enzymatic activity measurements.

Development of enzymatic activity during folding

A Bio-logic SFM-3 stopped-flow device (Claix, France) coupled with an absorbance detection system and thermostated with an RTE-200 water circulating bath from Neslab (Newington, NH) was used to measure the recovery of enzymatic activity during folding. Sso AcP was initially unfolded at a concentration equal to 0.4 mg ml $^{-1}$ in 5.5 M GdnHCl (Sigma-Aldrich), 50 mM acetate buffer, pH 5.5, at 37°C. 20 μl aliquots of this sample were mixed with 380 μl of a solution containing 5.27 mM BP, 50 mM acetate buffer, pH 5.5. Final conditions in the assay test were 0.02 mg ml $^{-1}$ Sso AcP, 0.275 M GdnHCl, 10 mM BP, 50 mM acetate buffer, pH 5.5, 37°C. The experimental dead time ranged from 4 to 20 ms. The cuvette length was 1 cm. The signal at 283 nm (A_{283}) was acquired during protein refolding. As the decrease of this signal is proportional to the enzyme activity, for each obtained trace, the decay rate of absorbance at 283 nm was calculated, after a smoothing on a 0.02 s sliding window, as the opposite of the first-order derivative ($-dA_{283}/dt$). The obtained values were normalized to the activity of

the native wild-type protein, plotted versus time and fitted to

$$A(t) = B + C \exp(-k_{1 \rightarrow F} \cdot t) \quad (1)$$

where t is the time, $A(t)$ is the enzymatic activity measured during folding, B is the activity of the native state, C is the difference between the activities of the partially folded and the native states and $k_{1 \rightarrow F}$ is the main folding rate constant.

Equilibrium GdnHCl-induced unfolding curves

For each mutational variant of Sso AcP, 28 samples containing 0.2 mg ml $^{-1}$ of the tested protein, 50 mM acetate buffer, pH 5.5, and a GdnHCl concentration ranging from 0 to 7.2 M, were incubated for 2 h at 37°C to reach equilibrium. After this time, the circular dichroism (CD) signal at 222 nm was acquired for all samples using a 0.1 cm path-length cuvette in a Jasco J-810 CD spectropolarimeter (Great Dunmow, Essex, United Kingdom) thermostated with a C25P Thermo Haake water-circulating bath (Karlsruhe, Germany). The mean residue ellipticity was plotted versus GdnHCl concentration. The obtained plots were fitted to a two-state transition according to the equation described by Santoro and Bolen (Santoro and Bolen, 1988) to obtain the free energy difference between the unfolded and the native states in the absence of denaturant ($\Delta G_{U \rightarrow F}^{\text{H}_2\text{O}}$), the dependence of the $\Delta G_{U \rightarrow F}$ on GdnHCl concentration (m value) and the midpoint of denaturation (C_m). The fraction folded at each GdnHCl concentration was calculated as described (Chiti *et al*, 1998). For each mutant, the change in conformational stability upon mutation, $\Delta \Delta G_{U \rightarrow F}^{\text{H}_2\text{O}}$, was calculated according to

$$\Delta \Delta G_{U \rightarrow F}^{\text{H}_2\text{O}} = (C'_m - C_m) \bar{m} \quad (2)$$

where C'_m and C_m are the mid-denaturation concentrations for the considered mutant and the wild type, respectively, and \bar{m} is the average m value over the wild-type and 34 mutants that were studied here. The average m value, as opposed to the individual m values, was used after the argument previously raised (Matouschek and Fersht, 1991). In brief, the m value obtained from a single curve is affected by some experimental uncertainty. The m value averaged over a number of mutants is in most cases a best estimate and leads to more accurate values of $\Delta G_{U \rightarrow F}^{\text{H}_2\text{O}}$ for the mutants (Matouschek and Fersht, 1991). C_m , $\Delta G_{U \rightarrow F}^{\text{H}_2\text{O}}$ and individually calculated m values are reported in Table II.

Folding kinetics

Folding experiments were carried out using the Bio-logic SFM-3 stopped-flow device (Claix, France) equipped with a fluorescence detection system and thermostated with an RTE-200 water circulating bath from Neslab (Newington, NH). An excitation wavelength of 280 nm and a band-pass filter to monitor emitted fluorescence above 320 nm were used. The cuvette path-length was 0.15 cm. A total of 20 μl aliquots of 0.4 mg ml $^{-1}$ protein unfolded in 5.5 M GdnHCl were mixed with 380 μl of refolding buffer. Final conditions were 0.02 mg ml $^{-1}$ protein, 50 mM acetate buffer, pH 5.5, a GdnHCl concentration ranging from 0.2 to 2.5 M, 37°C. The experimental dead time was 10.4 ms. The obtained traces were fitted to a double exponential equation as described previously (Bemporad *et al*, 2004). The main folding rate constant was then plotted versus the denaturant concentration to extrapolate the folding rate constant in the absence of denaturant ($k_{1 \rightarrow F}^{\text{H}_2\text{O}}$) according to a second-order polynomial equation. However, $k_{1 \rightarrow F}^{\text{H}_2\text{O}}$ refers to the rate of formation of the native state regardless of the on- or off-pathway nature of the partially folded state. In another set of experiments, 20 μl aliquots of 0.4 mg ml $^{-1}$ wild-type Sso AcP unfolded in 5.5 M GdnHCl were mixed with 380 μl of refolding buffer containing different concentrations of inorganic phosphate and phenyl phosphate. Final conditions were 0.02 mg ml $^{-1}$ Sso AcP, 0.275 M GdnHCl, 50 mM acetate buffer, pH 5.5, a phosphate and phenyl phosphate concentration ranging from 0.1 to 10 mM, 37°C. The resulting traces were fitted to a double exponential equation, as described previously (Bemporad *et al*, 2004).

Unfolding kinetics

Unfolding experiments were carried out using a Bio-logic SFM-20 stopped-flow device (Claix, France) equipped with a Jasco J-810 CD detection system (Great Dunmow, Essex, United Kingdom) and thermostated with a C25P Thermo Haake water circulating bath (Karlsruhe, Germany). The signal was recorded at 230 nm with a slit window of 4.0 nm. The cuvette path-length was 0.2 cm. A total of 85 μl aliquots of 1.4 mg ml $^{-1}$ protein in 1.0 M GdnHCl were mixed with 215 μl of a solution containing 8.0 M GdnHCl, 50 mM acetate,

pH 5.5. Final conditions were 0.4 mg ml⁻¹ protein, 50 mM acetate buffer, pH 5.5, 6.0 M GdnHCl, 37°C. The experimental dead time was 74 ms. A total of 6–10 traces were averaged and fitted to a single exponential equation to obtain the unfolding rate constant ($k_{F \rightarrow U}$). The unfolding rate constant in the absence of denaturant ($k_{F \rightarrow U}^{H_2O}$) was obtained with a linear extrapolation method using a previously obtained slope (Bemporad *et al*, 2004).

Φ -value analysis

A Φ -value analysis was carried out on the partially folded and transition state ensembles populated during the Sso AcP folding process. The Φ -values for the transition state ensemble, Φ^\ddagger (Matouschek and Fersht, 1991), were calculated according to

$$\Phi^\ddagger = 1 - \frac{RT \ln(k'_{F \rightarrow U} / k_{F \rightarrow U}^{H_2O})}{(C'_m - C_m) \bar{n}} \quad (3)$$

where C'_m and C_m are the mid-denaturation concentrations for the considered mutant and the wild type, respectively, R is the ideal gas constant (8.314 J mol⁻¹ K⁻¹), T is the temperature in K (310.15 K), $k'_{F \rightarrow U}$ and $k_{F \rightarrow U}^{H_2O}$ are the unfolding rate constants in the absence of denaturant for the mutant and the wild-type protein, respectively. The Φ -values for the partly folded state, Φ^1 (Matouschek *et al*, 1992), were calculated according to

$$\Phi^1 = 1 - \frac{RT \ln(k'_{F \rightarrow U} k_{1 \rightarrow F}^{H_2O} / k_{F \rightarrow U} k'_{1 \rightarrow F})}{(C'_m - C_m) \bar{n}} \quad (4)$$

where $k'_{1 \rightarrow F}$ and $k_{1 \rightarrow F}^{H_2O}$ are the folding rate constants in the absence of denaturant for the mutant and the wild-type protein, respectively.

References

- Bemporad F, Capanni C, Calamai P, Tutino ML, Stefani M, Chiti F (2004) Studying the folding process of the acylphosphatase from *Sulfolobus solfataricus*. A comparative analysis with other proteins from the same superfamily. *Biochemistry* **43**: 9116–9126
- Benkovic SJ, Hammes-Schiffer S (2003) A perspective on enzyme catalysis. *Science* **301**: 1196–1202
- Camicci G, Manao G, Cappugi G, Ramponi G (1976) A new synthesis of benzoyl phosphate: a substrate for acyl phosphatase assay. *Experientia* **32**: 535–536
- Chiti F, Taddei N, Giannoni E, van Nuland NAJ, Ramponi G, Dobson CM (1999) Development of enzymatic activity during protein folding. Detection of a spectroscopically silent native-like intermediate of muscle acylphosphatase. *J Biol Chem* **274**: 20151–20158
- Chiti F, van Nuland NAJ, Taddei N, Magherini F, Stefani M, Ramponi G, Dobson CM (1998) Conformational stability of muscle acylphosphatase: the role of temperature, denaturant concentration, and pH. *Biochemistry* **37**: 1447–1455
- 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 (2006) Structure, conformational stability, and enzymatic properties of acylphosphatase from the hyperthermophile *Sulfolobus solfataricus*. *Proteins* **62**: 64–79
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J* **272**: 5129–5148
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* **19**: 26–59
- Eisenmesser EZ, Bosco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. *Science* **295**: 1520–1523
- Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* **15**: 35–41
- Garcia-Viloca M, Gao J, Karplus M, Truhlar DG (2004) How enzymes work: analysis by modern rate theory and computer simulations. *Science* **303**: 186–195
- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochem* **189**: 319–326
- Gsponer J, Hopearuoho H, Whittaker SB, Spence GR, Moore GR, Paci E, Radford SE, Vendruscolo M (2006) Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7. *Proc Natl Acad Sci USA* **103**: 99–104
- Kiefhaber T, Schmid FX, Willaert K, Engelborghs Y, Chaffotte A (1992) Structure of a rapidly intermediate in ribonuclease T1 folding. *Protein Sci* **1**: 1162–1172
- MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE *et al* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**: 3586–3616
- Matouschek A, Fersht AR (1991) Protein engineering in analysis of protein folding pathways and stability. *Methods Enzymol* **202**: 82–112
- Matouschek A, Kellis Jr JT, Serrano L, Fersht AR (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**: 122–126
- Matouschek A, Serrano L, Fersht AR (1992) The folding of an enzyme. IV. Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure. *J Mol Biol* **224**: 819–835
- Miyazono K, Sawano Y, Tanokura M (2005) Crystal structure and structural stability of acylphosphatase from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Proteins* **61**: 196–205
- Nagao C, Terada TP, Yomo T, Sasai M (2005) Correlation between evolutionary structural development and protein folding. *Proc Natl Acad Sci USA* **102**: 18950–18955
- Nakayama KI, Hatakeyama S, Nakayama K (2001) Regulation of the cell cycle at the G1-S transition by proteolysis of cyclin E and p27Kip1. *Biochem Biophys Res Commun* **282**: 853–860
- Osborne MJ, Schnell J, Benkovic SJ, Dyson HJ, Wright PE (2001) Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism. *Biochemistry* **40**: 9846–9859
- Pagano K, Ramazzotti M, Viglino P, Esposito G, Degl'Innocenti D, Taddei N, Corazza A (2006) NMR solution structure of the acylphosphatase from *Escherichia coli*. *J Biomol NMR* **36**: 199–204
- Park HS, Nam SH, Lee JK, Yoon CN, Mannervik B, Benkovic SJ, Kim HS (2006) Design and evolution of new catalytic activity with an existing protein scaffold. *Science* **311**: 535–538

Structural analysis

Molecular dynamics simulations in explicit solvent with Φ -value restraints were carried out from the native state by using the CHARMM22 force-field (MacKerell *et al*, 1998), an all-atom protein representation, the TIP3P water model and periodic boundary conditions. All calculations used an atom-based truncation scheme with a list cut-off of 14 Å, a non-bond cut-off of 12 Å and the Lennard–Jones smoothing function initiated at 10 Å. Electrostatic and Lennard–Jones interactions were force-switched. MD simulations used a 2 fs integration time step and SHAKE of covalent bonds involving hydrogen atoms. For more details, see Supplementary data.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank Giampietro Ramponi for his critical reading and Elemir Simko for his technical assistance. This work was supported by grants from the Italian MIUR (FIRB projects no. RBNE01S29H and RBNE03PX83 and PRIN project no. 2005027330_003), the EMBO Young Investigator Program 2005, the Royal Society, the European Community (project no. HPRN-CT-2002-00241) and the Leverhulme Trust. JG was supported by the Swiss National Science Foundation. MV was supported by the Royal Society.

- Pervushin K, Vamvaca K, Vogeli B, Hilvert D (2007) Structure and dynamics of a molten globular enzyme. *Nat Struct Mol Biol* **14**: 1202–1206
- Poulsen TD, Garcia-Viloca M, Gao J, Truhlar DG (2003) Free energy surface, reaction paths, and kinetic isotope effect of short-chain Acyl-CoA dehydrogenase. *J Phys Chem B* **107**: 9567–9578
- Ramponi G, Treves C, Guerritore AA (1966) Aromatic acyl phosphates as substrates of acyl phosphatase. *Arch Biochem Biophys* **115**: 129–135
- Salvatella X, Dobson CM, Fersht AR, Vendruscolo M (2005) Determination of the folding transition states of barnase by using PhiI-value-restrained simulations validated by double mutant PhiIJ-values. *Proc Natl Acad Sci USA* **102**: 12389–12394
- Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**: 8063–8068
- Settanni G, Gsponer J, Caflisch A (2004) Formation of the folding nucleus of an SH3 domain investigated by loosely coupled molecular dynamics simulations. *Biophys J* **86**: 1691–1701
- Stefani M, Taddei N, Ramponi G (1997) Insights into acylphosphatase structure and catalytic mechanism. *Cell Mol Life Sci* **53**: 141–151
- Thunnissen MM, Taddei N, Liguri G, Ramponi G, Nordlund P (1997) Crystal structure of common type acylphosphatase from bovine testis. *Structure* **5**: 69–79
- Uversky VN, Kutysenko VP, Protasova NY, Rogov VV, Vassilenko KS, Gudkov AT (1996) Circularly permuted dihydrofolate reductase possesses all the properties of the molten globule state, but can resume functional tertiary structure by interaction with its ligands. *Protein Sci* **5**: 1844–1851
- Vamvaca K, Vogeli B, Kast P, Pervushin P, Hilvert D (2004) An enzymatic molten globule: efficient coupling of folding and catalysis. *Proc Natl Acad Sci USA* **101**: 12860–12864
- Vendruscolo M, Paci E, Dobson CM, Karplus M (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* **409**: 641–645
- Zuccotti S, Rosano C, Ramazzotti M, Degl'Innocenti D, Stefani M, Manao G, Bolognesi M (2004) Three-dimensional structural characterization of a novel *Drosophila melanogaster* acylphosphatase. *Acta Crystallogr D Biol Crystallogr* **60**: 1177–1179