Protein folding of the SAP domain, a naturally occurring two-helix bundle

Charlotte A. Dodson a,b,⁎, Eyal Arbely a,1

a MRC Centre for Protein Engineering, Hills Road, Cambridge CB2 0QH, UK
b Molecular Medicine, National Heart & Lung Institute, Imperial College London, SAF Building, London SW7 2AZ, UK

1 Current address: Department of Chemistry and The National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel.

⁎ Corresponding author at: Molecular Medicine, National Heart & Lung Institute, Imperial College London, SAF Building, London SW7 2AZ, UK.

E-mail address: c.dodson@imperial.ac.uk (C.A. Dodson).

Article history:
Received 8 May 2015
Revised 1 June 2015
Accepted 1 June 2015
Available online 11 June 2015
Edited by Stuart Ferguson

Keywords:
SAP domain
Protein folding
Φ-value
Transition state analysis
Tho1

Abstract

The SAP domain from the Saccharomyces cerevisiae Tho1 protein is comprised of just two helices and a hydrophobic core and is one of the smallest proteins whose folding has been characterised. Φ-value analysis revealed that Tho1 SAP folds through a transition state where helix 1 is the most extensively formed element of secondary structure and flickering native-like core contacts from Leu35 are also present. The contacts that contribute most to native state stability of Tho1 SAP are not formed in the transition state.

© 2015 The Authors. Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The principles which govern the folding and unfolding of proteins have fascinated the scientific community for decades [1]. One of the most successful approaches has been to apply chemical transition state theory and treat the folding reaction as a barrier-limited process between two conformational ensembles of proteins populated at equilibrium: the native and denatured state ensembles.

The transition state, the high energy conformation transiently adopted by the polypeptide chain as the protein crosses the barrier between native and denatured ensembles, provides information on the structural mechanism of protein folding. This information cannot be determined by traditional structural techniques, and instead is inferred from kinetics and mutational analysis using the technique of Φ-value analysis [2–7]. The Φ-value of a mutation is defined as the change in stability of the transition state upon making the mutation (ΔΔG N→D = ΔG N WT – ΔG D WT – ΔG N→D mut) expressed as a fraction of the change in stability of the native state (ΔΔG N,D = ΔG N WT – ΔG D WT) for the same mutation: i.e. Φ = ΔΔG N→D mut

http://dx.doi.org/10.1016/j.febslet.2015.06.002
0014-5793/© 2015 The Authors. Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
contain antiparallel or perpendicular helices in a helix-turn-helix arrangement [10–14] – and Tho1 SAP is one of the smallest proteins whose folding has been studied experimentally. It is therefore of interest to study the folding of Tho1 SAP in more detail.

In this paper we have conducted a U-value analysis of the Tho1 SAP domain. The U-values we obtained were fractional, indicating that Tho1 SAP folds through a transition state with transient formation of a core and flickering elements of helical structure. The best formed element of secondary structure was helix 1. Interestingly, the contacts which contributed most to native state stability were not formed in the transition state. In order to obtain a crude indication of the validity of our results across multiple temperatures, we measured the folding of L31W SAP across the range of 283–323 K. As judged by the change in solvent-accessible surface area upon folding (bT value and analogous ratio of heat capacities), there are no gross changes in the transition state of Tho1 SAP with temperature.

2. Materials and methods

2.1. Reagents

L31W SAP domain was expressed and purified as detailed previously [8]. Point mutations were generated using Stratagene Quikchange mutagenesis. Mutant proteins were expressed and purified as described for SAP L31W until completion of cleavage of the fusion protein, when tag and target were separated by flowing once more through Ni-charged IMAC resin (GE Healthcare BioSciences, Sweden) before concentration and gel filtration on an S75 column into 50 mM MES pH 6.0, total ionic strength of 500 mM made up to this value using NaCl. A single peak was obtained for all proteins and fractions within this peak pooled.

2.2. Equilibrium denaturation

Far-UV CD spectroscopy (thermal denaturation) and fluorescence emission (chemical denaturation) were carried out as described previously [8].

2.3. Kinetic measurements

We measured relaxation kinetics on the μ–ms timescale using T-jump fluorescence spectroscopy and temperature jumps of 3–5 K on a modified Hi-Tech PTJ-64 (Hi-Tech Ltd., Salisbury, UK) capacitor-discharge T-jump apparatus as previously described [8]. Arrhenius analysis of the plot of microscopic rate constant against temperature was carried out constraining the overall ΔH, ΔS and ΔCp to their equilibrium values at the thermal midpoint.

Fig. 1. Structure of L31W and energetic contribution of different residues to stability. (a) Connectivity of L31W. Trp31 is shown in stick representation. (b) Contribution of side chains to native state stability. −0.5 < ΔAGN,0 < 0.5 kcal mol⁻¹ (red), 0.5 < ΔAGN,0 < 1.0 kcal mol⁻¹ (orange), 1.0 < ΔAGN,0 < 1.5 kcal mol⁻¹ (yellow), 1.5 < ΔAGN,0 < 2.0 kcal mol⁻¹ (green), ΔAGN,0 > 2.0 kcal mol⁻¹ (blue). Trp31 is shown in grey. (c) Contribution of side chains to transition state stability. Φ-values coloured by Φ₁ < 0 (pink), 0 < Φ₁ < 0.3 (red), 0.3 < Φ₁ < 0.6 (yellow) and Φ₁ > 0.6 (blue). Trp31 and Leu22 shown in grey. (d) Putative hydrogen bonding network stabilising the native state of Tho1 SAP. Proposed hydrogen bonds shown with dashed yellow lines, other colours as for (b). (e) Φ-values for alanine-glycine scanning of helix 1 with A10G, A12G and A14G. Colours as for (c).
Urea denaturation (chevron) plots were fitted to an apparent two-state transition (Eq. 1) with each point weighted by the fitting error on the rate constant. No further constraints were placed upon the data fit:

\[
k_{obs} = k_f \left( \frac{m}{m_D - m} \right) + k_e \left( \frac{m_D}{m_D - m} \right)
\]

(1)

2.4. Estimating \( \Phi \)-values for severely destabilised mutants

We estimated the group (low, medium or high) into which the \( \Phi \)-values for L17A, R20A and L35A were likely to fall by using our measured data to place bounds on \( k_f \) for these severely destabilised mutants. The data for all three destabilised mutants describe the unfolding arm of the chevron well (\( m_{1-D} \) and \( k_f \)) and we used the average value of \( m_{1-D} \) from all other mutants (480 ± 21 cal mol\(^{-1}\) M\(^{-1}\)) as a fixed parameter in fitting these chevrons to enable convergence on a solution. We also fit the data with \( m_{1-D} \) fixed at 400 cal mol\(^{-1}\) M\(^{-1}\) and 600 cal mol\(^{-1}\) M\(^{-1}\) to check that constraining this value did not overly affect the value of \( k_f \) obtained (Supplementary Table SI). In addition, we fit the data for the R20A mutant using Eq. (1) with all parameters allowed to float freely. In all cases, the value of [\( D \)]\(_{obs} \) calculated from the kinetic data was used in calculating \( \Phi \) and was checked against the data from equilibrium urea denaturation to ensure that it was consistent with this.

2.4.1. R20A chevron

As an additional check on data fitting, we observed that the measured rate constant for R20A in 0 M urea is at, or is a little above, the midpoint of the transition (visual inspection of Fig. 2d). This places an upper bound of 1500 s\(^{-1}\) on \( k_f \) (half the measured observed rate constant), consistent with the values in

| Mutation (L31W+) | \( T_m \) (K) | \( \Delta H_{Urea} \) (at \( T_m \)) (kcal mol\(^{-1}\)) | \( \Delta AG_{Urea} \) at 322 K (kcal mol\(^{-1}\)) | \( m_{0-N} \) (cal mol\(^{-1}\) M\(^{-1}\)) | [\( D \)]\(_{obs} \) (M) | \( \Delta AG_{O,D} \) at 4.8 M urea (kcal mol\(^{-1}\)) |
|------------------|--------------|---------------------------------|---------------------------------|-----------------|-----------------|---------------------------------|
| Wild-type        | 322.4 ± 0.2  | 33.4 ± 0.4                      | N/A                             | 690 ± 50        | 4.8 ± 0.1        | N/A                             |
| Y5A             | N/D          | N/D                             | N/A                             | N/D             | N/D             | N/A                             |
| Y5F             | 299          | 16.8                            | 2.4                             | 620             | 1.7             | 1.8 ± 0.2                       |
| S6A             | 323          | 34.4                            | -0.1                            | 530             | 4.6             | 0.2 ± 0.1                       |
| S6G             | 321          | 33.4                            | 0.2                             | 930             | 4.5             | 0.2 ± 0.1                       |
| L18             | 303          | 31.4                            | 2.0                             | 660             | 1.2             | 2.0 ± 0.2                       |
| T9S             | 323          | 35.5                            | -0.1                            | 520             | 4.2             | 0.4 ± 0.1                       |
| T9A             | N/D          | N/D                             | N/D                             | 670             | 2.2             | 1.7 ± 0.2                       |
| V10A            | 321          | 33.3                            | 0.1                             | 730             | 4.4             | 0.3 ± 0.1                       |
| V10G            | 316          | 33.6                            | 0.7                             | 680             | 3.5             | 0.9 ± 0.1                       |
| V10A/V11A        | 326          | 35.4                            | -0.3                            | 730             | 5.0             | -0.4 ± 0.1                      |
| V11A            | 326          | 33.9                            | -0.4                            | 320             | 6.9             | -0.4 ± 0.1                      |
| V11G            | 323          | 34.4                            | -0.1                            | 470             | 5.1             | 0.0 ± 0.1                       |
| Q12A            | 317          | 32.6                            | 0.6                             | 720             | 3.4             | 0.9 ± 0.1                       |
| Q12G            | 303          | 21.3                            | 2.0                             | 730             | 2.2             | 1.9 ± 0.2                       |
| L13A            | 317          | 31.0                            | 0.5                             | 610             | 3.9             | 0.6 ± 0.1                       |
| K14A            | 318          | 31.4                            | 0.4                             | 570             | 3.3             | 0.8 ± 0.1                       |
| K14G            | 304          | 27.5                            | 1.9                             | 690             | 2.0             | 1.9 ± 0.2                       |
| L16A            | 310          | 29.8                            | 1.3                             | 760             | 2.9             | 1.3 ± 0.1                       |
| L17A            | N/D          | N/D                             | N/D                             | N/D             | N/D             | N/D                             |
| K19A            | 319          | 36.5                            | 0.3                             | 380             | 4.6             | 0.4 ± 0.1                       |
| R20A            | N/D          | N/D                             | N/D                             | N/D             | N/D             | N/D                             |
| L22A            | N/D          | N/D                             | N/D                             | N/D             | N/D             | N/D                             |
| S23A            | 317          | 31.9                            | 0.6                             | 590             | 3.3             | 0.3 ± 0.1                       |
| V24A            | 317          | 29.9                            | 0.5                             | N/D             | N/D             | N/D                             |
| K28A            | 316          | 29.3                            | 0.6                             | 550             | 4.8             | -0.0 ± 0.1                      |
| Q33A            | 323          | 35.3                            | -0.1                            | 600             | 5.0             | -0.1 ± 0.1                      |
| R34A            | 315          | 24.0                            | 0.8                             | 570             | 3.4             | 0.8 ± 0.1                       |
| L35A            | N/D          | N/D                             | N/D                             | N/D             | N/D             | N/D                             |
| I36A            | 314          | 29.7                            | 0.9                             | 590             | 3.0             | 1.1 ± 0.1                       |
| I36V            | 320          | 33.5                            | 0.3                             | 660             | 4.0             | 0.5 ± 0.1                       |
| D38A            | 323          | 30.6                            | -0.1                            | 610             | 4.2             | 0.4 ± 0.1                       |
| D39A            | 301          | 13.7                            | 2.2                             | N/D             | N/D             | N/D                             |
| D39N            | 298          | 13.3                            | 2.6                             | 610             | 1.3             | 2.1 ± 0.2                       |
| E40Q            | 322          | 33.3                            | 0.1                             | 660             | 4.6             | 0.1 ± 0.1                       |
| E41G            | 319          | 31.5                            | 0.4                             | 660             | 4.1             | 0.5 ± 0.1                       |

N/A not applicable, N/D not determined.

a S.E.M. reported for wild-type (n = 3). Standard deviation on repeat measurements on wild-type and all fitting errors on mutants <1 K for \( T_m \) and <0.6 kcal mol\(^{-1}\) for \( \Delta H \). 

b \( \Delta AG_{O,D} \) calculated using the Scheffelin equation [30]: \( \Delta AG = (T_m^{WT} - T_m^{mutant}) \Delta H^{WT}/T_m^{mutant} \). All reported values have a propagated error of 0.1 kcal mol\(^{-1}\) (propagation using error of 1 K for \( T_m \) of mutant, and S.E.M. for wild-type parameters).

c S.E.M. reported for wild-type (n = 6). Fitting error on all values reported for mutants <10 cal mol\(^{-1}\) M\(^{-1}\).

d [\( D \)]\(_{obs} \) obtained using average \( m_{0-N} \) of 645 ± 14 cal mol\(^{-1}\) M\(^{-1}\) for fits of mutant data. S.E.M. reported for wild-type (n = 6), fitting error <0.1 M for all mutants.

e \( \Delta AG \) calculated from \( \Delta [D]_{obs} \) and average \( m_{0-N} \). Error propagation used an error of 0.2 M in \( [D]_{obs} \) and 120 cal mol\(^{-1}\) M\(^{-1}\) in \( m_{0-N} \) for mutant proteins. S.E.M. was used for wild-type parameters.

f Poor protein expression.

g Mutation too destabilising to fit and report parameters for thermal or chemical denaturation curves.
Supplementary Table SI. As a final check on our estimate of \( \Phi \), we recalculated this parameter using a \( \Delta \Delta G_{N,D} \) of 2.1 kcal mol\(^{-1}\). This is the most destabilised transition we have been able to fit (D39N) and represents the minimum value of \( \Delta \Delta G_{N,D} \) (maximum value of \( \Phi \)) for this mutant. We are thus confident in an estimate of \( \Phi \) for R20A as <0.3.

2.4.2. L17A chevron and L35A chevrons

All data fits resulted in a \( \Phi \)-value of 0.1 for L17A and 0.3 for L35A. We were a little concerned that the data fits indicated that the L35A mutant was less stable than L17A when the chevrons suggest the reverse (urea concentration of maximum measurable signal higher for L17A). In order to determine the impact of this on the \( \Phi \)-value calculated, we used the fitted values of \( k_f \) to calculate \( \Phi \)-values with destabilisations corresponding to \([D]_{\text{50\%}}\) values of 1.3 M (\( \Delta \Delta G_{N,D} = 2.1 \text{ kcal mol}^{-1} \)), 0 M and −2.0 M urea (L35A) and 1.3 M, −2.0 M and −4.0 M urea (L17A). For L17A, using the minimum value of 2.1 kcal mol\(^{-1}\) for \( \Delta \Delta G_{N,D} \) raised \( \Phi \) to 0.2 for all but the lowest value of \( m_{1,D} \). Other values of \([D]_{\text{50\%}}\) left \( \Phi \) = 0.1 (unchanged) indicating a likely value of 0.1. For L35A, the maximum value of \( \Phi \) was 0.5/0.6 (\( \Delta \Delta G_{N,D} = 2.1 \text{ kcal mol}^{-1} \)) with a value of 0.4 when \([D]_{\text{50\%}} = 0 \text{ M} \) and unchanged (\( \Phi \) = 0.3) when \([D]_{\text{50\%}} = -2.0 \text{ M} \). Since we estimate the midpoint of this mutant to be between 0 M and −2.0 M, we assign the \( \Phi \)-value for L35A to the ‘medium’ category.

3. Results

3.1. Contribution of specific side-chains to native state stability

As a preparation for \( \Phi \)-value analysis, and in order to probe the contribution of different residues to the native state stability of

![Fig. 2.](image-url)
L31W, we made a series of thirty-five non-disruptive side-chain deletion mutations (Table 1, Fig. 1b and Supplementary Fig. S1). The greatest contribution to L31W stability (>2 kcal mol⁻¹) was made by a cluster of residues on the buried faces of the C-terminal end of both helices and at the beginning of the connecting loop. Of the five residues making up this cluster, Leu17 and Leu35 are buried in the hydrophobic core of the SAP domain, and Leu22 and the alky chain of Arg20 pack against them. Asp39 and the guanidinium group of Arg20 are solvent-exposed.

To probe the role of the surface-exposed residues further, we mutated Asp39 to both alanine and asparagine. Both mutations destabilised L31W by similar amounts (2.2 and 2.6 kcal mol⁻¹) indicating that the carboxylic acid group in the aspartic acid is forming a stabilising interaction. Mutating Arg20 to alanine destabilised L31W by a greater amount (>2.6 kcal mol⁻¹) than mutating Asp39. 2.0–2.5 kcal mol⁻¹ of this destabilisation is likely to arise from loss of a hydrogen bond or salt bridge between the guanidinium group of Arg20 and the carboxylic acid of Asp39 (Fig. 1d), and the remaining destabilisation is likely to be due to the loss of the Arg20 alkyl chain packing against the hydrophobic core. We speculate that a further hydrogen bond may be formed between Asp39 and Tyr5. Like Arg20, Tyr5 is close to Asp39 and removal of the tyrosine hydroxyl destabilised L31W considerably (Y5F mutation; 2.4 kcal mol⁻¹).

Most residues within helix 1 played little role in stabilising the native state. However, comparison of T9S and T9A mutants enabled us to determine the energetic contribution of the N-capping Thr9 hydroxyl group to the overall stability of L31W (i.e. to determine the contribution of the virtual mutation S9A) to be 1.3 kcal mol⁻¹. Leu8 acts as a hydrophobic staple (L8A gave a ΔΔG_D,N of 2.0 kcal mol⁻¹).

In order to build a picture of the transition state of the SAP domain, we undertook a Φ-value analysis of L31W. Twenty of our thirty-five point mutants were suitable for Φ-value analysis [15]. All gave rise to apparent two-state chevron plots with linear arms (Fig. 2) which, in most cases, could be fitted with high confidence to yield Φ-values (Table 2). Three mutants, L17A, R20A and L35A, were so destabilised that it was not possible to fit the folding arm of the chevron reliably. However by placing bounds on kₘ and on the equilibrium denaturation midpoint, and by careful analysis of the resulting data fits, we could estimate into which category the Φ-values for these mutants fell (see Section 2.4 for full details).

Most Φ-values measured were fractional, which we interpret to indicate that the transition state ensemble for SAP L31W does not contain extensive regions of fully consolidated secondary structure (Fig. 1c & e). The highest Φ-values (Φ = 0.5–0.6) were measured for Leu8, Thr9 and Leu16. Leu8 and Thr9 form the base of helix 1, while Leu16 (on helix 1) participates in core hydrophobic packing. Alanine to glycine mutations at the N-terminus of helix 1 (which probe the degree of helix formation [2,16]) gave Φ-values ranging from 0.3 to 0.7, although the error on the largest value is considerable. Together, these results indicate that helix 1 is partially formed in the transition state for folding.

Our Φ-value mutants provide several probes of the degree of formation of the hydrophobic core in the L31W transition state. Leu13 and Arg20 both have low Φ-values indicating that their core contacts are not formed in the transition state. Leu35 (helix 2) and Leu16 (helix 1) have medium Φ-values, with that for Leu35 being lower than that for Leu16. It was not possible to measure a Φ-value for Leu22 since the L22A mutant was too destabilised for kinetic

### Table 2

| Mutation (L31W) | kᵢ (s⁻¹) | -mᵢ,D (cal mol⁻¹ M⁻¹) | kₑ (s⁻¹) | mₑ,N (cal mol⁻¹ M⁻¹) | Calc [D]ₜₜₜ (M) | Φ (s) |
|----------------|----------|------------------------|----------|------------------------|------------------|-------|
| Wild-type      | 3160 ± 30 | 420 ± 10               | 210 ± 10 | 5.4                    | 630 ± 10         | N/A   |
| Y5F            | 2480 ± 20 | 400 ± 10               | 110 ± 10 | 1.7                    | 680 ± 10         | 0.5 ± 0.1 |
| T9A            | 500 ± 10  | 520 ± 10               | 160 ± 10 | 2.8                    | 720 ± 10         | 0.6 ± 0.1 |
| V10G           | 1950 ± 20 | 420 ± 10               | 270 ± 90 | 4.5                    | 690 ± 10         | 0.3 ± 0.1 |
| A10G           | N/A       | N/A                    | N/A      | N/A                    | N/A              | 0.3 ± 0.1 |
| Q12A           | 3090 ± 40 | 710 ± 10               | 20 ± 10  | 2.6                    | 730 ± 10         | 0.0 ± 0.1 |
| L13A           | 3280 ± 40 | 430 ± 10               | 20 ± 10  | 2.2                    | 690 ± 10         | 0.3 ± 0.1 |
| K14G           | 3580 ± 20 | 430 ± 10               | 220 ± 50 | 4.0                    | 650 ± 10         | -0.1 ± 0.1 |
| K14G           | 1550 ± 10 | 540 ± 10               | 120 ± 10 | 1.9                    | 660 ± 10         | 0.2 ± 0.1 |
| A14G           | N/A       | N/A                    | N/A      | N/A                    | N/A              | 0.4 ± 0.2 |
| L16A           | 1040 ± 10 | 650 ± 10               | 80 ± 10  | 2.0                    | 720 ± 10         | 0.5 ± 0.1 |
| L17A           | N/D       | N/D                    | N/D      | N/D                    | N/D              | -0.1³ |
| R20A           | N/D       | N/D                    | N/D      | N/D                    | N/D              | -0.3–0.6⁴ |
| S23A           | 3850 ± 60 | 490 ± 10               | 150 ± 10 | 3.4                    | 640 ± 10         | -0.1 ± 0.1 |
| V24A           | 3000 ± 40 | 430 ± 10               | 150 ± 10 | 4.3                    | 580 ± 10         | 0.1 ± 0.1 |
| R34A           | 6170 ± 270| 590 ± 40               | 3 ± 20   | 2.1                    | 590 ± 50         | -0.5 ± 0.1 |
| L35A           | N/D       | N/D                    | N/D      | N/D                    | N/D              | -0.3–0.6⁴ |
| I36A           | 2140 ± 20 | 360 ± 10               | 290 ± 10 | 3.8                    | 650 ± 10         | 0.2 ± 0.1 |
| L36V           | 2390 ± 40 | 470 ± 10               | 100 ± 10 | 4.0                    | 570 ± 10         | 0.3 ± 0.1 |
| D39A           | 2490 ± 10 | 340 ± 30               | 120 ± 30 | 2.1                    | 470 ± 40         | 0.1 ± 0.1 |
| D39N           | 2600 ± 80 | 440 ± 30               | 110 ± 20 | 1.4                    | 550 ± 40         | 0.1 ± 0.1 |

N/A not applicable, N/D not determined.

a Errors quoted are fitting errors. A minimum error of 10 is applied to all fitting errors with the exception of values of kₑ that less than 10, where a minimum error of 5 is applied.

b Calculated at 0 M urea. Φ = RT ln (kₑ/kᵢ)/Δ[D]ₜₜₜ. Minimum error of 0.1 reported on Φ.

c mₑ,N/μmolN/m, mᵢ,D/N-D: mᵢ,D-N = mᵢ,D-D

d Kinetic parameters for detailed fits of these very destabilised mutants given in Supplementary Table SI.

e Φ for D39A calculated using average equilibrium m-value and estimated [D]ₜₜₜ.
3.3. Solvent accessibility of the Tho1 SAP transition state

Our measurements of SAP domain folding (here and in ref [8]) have all been made at 283 K because the baselines of our equilibrium chemical denaturation are best defined at this temperature and our measurements most accurate. However, many protein folding experiments are carried out at higher temperatures and we wanted some assurance that our conclusions were valid across a range of temperatures.

The denaturant m-value of a protein unfolding transition reports on changes in solvent-accessible surface area (ΔSASA) upon protein denaturation [17]. In order to determine whether the ΔSASA between unfolded and transition states (ΔSASA_{u,D}) varied with temperature, we measured the kinetics of SAP domain folding at 10 degree intervals between 283 K and 323 K (Fig. 3a and Table 3).

The Tanford b-value (b_{T} = S_{T}/S_{D}) reports on ΔSASA_{u,D} as the fractional ΔSASA upon formation of the transition state. Changes in b_{T} with temperature or [chemical denaturant] can indicate Hammond behaviour [18–20] or one of many mechanistic changes [21–25]. The average value of b_{T} for Tho1 SAP was 0.70 ± 0.03 (mean ± S.E.M.) and did not vary with temperature.

We also probed the folding of SAP L31W in buffer alone by measuring the observed rate constant, k_{obs}, as a function of temperature (Fig. 3b). We included the rate constants in the absence of denaturant from Table 3 as they overlaid the other kinetic data excellently and helped define the curvature in k_{f} at low temperatures. The fractional position of the L31W transition state at 322 K (calculated from the ratio ΔC_{p(u,D)}/ΔC_{p(N,D)} [26]) is also a measure of fractional ΔSASA upon formation of the transition state analogous to b_{T} [17]. The ratio of heat capacities for SAP L31W was 0.9.

Both b_{T} and the ratio of heat capacities for Tho1 SAP are high, indicating that the greatest change in solvent accessible surface area occurs between denatured and transition states. Thus the transition state of Tho1 SAP is a compact, partially dehydrated, structured species. The b_{T} value for Tho1 SAP did not vary with temperature, giving us confidence that the model of the transition state we have determined is likely to be relevant across a range of temperatures.

4. Discussion

4.1. Brønsted/Leffler analysis of Tho1 SAP

As an alternative method of determining the extent of structure formation in the Tho1 SAP transition state, we plotted ΔΔG_{u,D} against ΔΔG_{D,N} for each of our b-value mutants (Fig. 4). By comparison with Brønsted analysis [27] and Leffler α-values [28] of organic chemistry, linearity among residues present in the same region of the protein (e.g. within an element of secondary structure) indicates that this region is equally formed in the transition state. The gradient of the line gives the extent of formation of this region on a reaction coordinate of free energy [29]. Outliers in this analysis can indicate long-range contacts. Depending on the position of the outlier (i.e. above or below the trend line), a point may indicate long-range contacts formed in the transition state, or may indicate contacts formed in the native state but not present in the transition state.

For the SAP domain, there is a weak correlation between ΔΔG_{u,D} and ΔΔG_{D,N} for those residues in helix 1 which make only local contacts in the native state (solid red and black points in Fig. 4; slope of best fit line = 0.4 ± 0.1; R² = 0.8). Two points in helix 1 and one point in helix 2 are outliers in this analysis (open circles in Fig. 4). These points are from residues which make long-range contacts in the SAP native state (L17A, R20A and L35A) and
for Leu35 is much higher than for and the remained constant across the range 283–323 K indicating of Tho1 SAP (Fig. 4. Fitting equation was formulated as $DD = k_f (s^{-1}) + k_{-f} (s^{-1})$ for this fit. Reported values for $k_f$ and $k_{-f}$ are calculated from $k_{0f}$ (the microscopic rate constant at the denaturation midpoint).

$\Delta D_{G,\text{N}}$ is determined from equilibrium chemical denaturation. Fitting error $<10$ cal mol$^{-1}$, and standard deviation of repeat measurements at 283 K is 100 cal mol$^{-1}$ M$^{-1} (n = 6)$.

$[D]_{0.5(\text{eqm})}$ determined by equilibrium chemical denaturation (283 K, 293 K, 303 K) or estimated by calculating $\Delta G_{D,\text{N}}$ from thermal denaturation and dividing by $m_D$ (313 K, 323 K).

The contacts which contributed most to Tho1 SAP native state stability were not formed in the transition state. We thus conclude that the late stages of Tho1 folding involve ‘locking’ the diffuse transition state ensemble into the native state using strong, long-range interactions.

5. Summary & conclusions

We carried out a $\Phi$-value analysis of the folding reaction of the Tho1 SAP domain in order to build a structural model of the transition state. All $\Phi$-values were fractional, indicating that no elements of secondary structure were fully formed in the transition state. Overall, helix 1 was the most highly structured region in the transition state with occasional native-like core contacts from Leu16 and Leu35.

We also probed the folding of the SAP domain by determining the degree of $\Delta$SASA at different temperatures. Both $\beta_f$ and the analogous ratio of heat capacities indicated a compact transition state. $\beta_f$ remained constant across the range 283–323 K indicating no gross changes in the folding mechanism.

The contacts which contributed most to Tho1 SAP native state stability were not formed in the transition state. We thus conclude that the early stages of SAP domain folding involve chain collapse and exclusion of solvent water, while the late stages involve formation of strong long-range interactions.

Acknowledgements

We are grateful to Neil Ferguson for his advice, support and for critical reading of the manuscript, to Chris Johnson for his technical help and to Alan Fersht for provision of laboratory space and equipment. We would also like to thank our anonymous peer reviewers whose constructive comments have improved this manuscript. This work was funded by an EMBO Long-Term Fellowship (ALTF-650-2006 to EA), a Medical Research Council (MRC), UK Research Studentship (ref G78/7940) and an Imperial College Junior Research Fellowship (to CD).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2015.06.002.
References

[1] Dill, K.A. and MacCallum, J.L. (2012) The protein-folding problem, 50 years on. Science (New York, NY) 338, 1042–1046.
[2] Fersht, A.R., Matouschek, A. and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. J. Mol. Biol. 224, 771–782.
[3] Matouschek, A., Religa, T.L., Serrano, L. and Fersht, A.R. (1989) Mapping the transition state and pathway of protein folding by protein engineering. Nature 340, 122–126.
[4] Fersht, A.R. (1999) Structure and Mechanism in Protein Science, W. H. Freeman and Co., New York.
[5] Ferguson, N., Day, R., Johnson, C.M., Allen, M.D., Daggett, V. and Fersht, A.R. (2005) Simulation and experiment at high temperatures: ultrafast folding of a thermophilic protein by nucleation–condensation. J. Mol. Biol. 347, 855–870.
[6] Sato, S., Religa, T.L. and Fersht, A.R. (2004) Phi-value analysis and the nature of protein-folding transition states. Proc. Natl. Acad. Sci. U.S.A. 101, 7976–7981.
[7] Giri, R., Morrone, A., Travaglini-Allocatelli, C., Jemth, P., Brunori, M. and Gianni, S. (2012) Folding pathways of proteins with increasing degree of sequence identities but different structure and function. Proc. Natl. Acad. Sci. U.S.A. 109, 17772–17776.
[8] Dodson, C.A., Ferguson, N., Rutherford, T.J., Johnson, C.M. and Fersht, A.R. (2010) Engineering a two-helix bundle protein for folding studies. Protein Eng. Des. Sel. 23, 357–364.
[9] Aravind, L. and Koonin, E.V. (2000) SAP – a putative DNA-binding motif involved in chromosomal organization. Trends Biochem. Sci. 25, 112–114.
[10] Jemth, P., Day, R., Gianni, S., Khan, F., Allen, M., Daggett, V. and Fersht, A.R. (2005) The structure of the major transition state for folding of an FF domain from experiment and simulation. J. Mol. Biol. 350, 363–378.
[11] Mayor, U., Guydosh, N.R., Johnson, C.M., Grossmann, J.C., 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.
[12] Sato, S., Religa, T.L., Daggett, V. and Fersht, A.R. (2004) Testing protein-folding simulations by experiment: B domain of protein A. Proc. Natl. Acad. Sci. U.S.A. 101, 6952–6956.
[13] Ferguson, N., Sharpe, T.D., Schartau, P.J., Sato, S., Allen, M.D., Johnson, C.M., Rutherford, T.J. and Fersht, A.R. (2005) Ultra-fast barrier-limited folding in the peripheral subunit-binding domain family. J. Mol. Biol. 353, 427–446.
[14] Ferguson, N., Capaldi, A.P., James, R., Kleanthous, C. and Radford, S.E. (1999) Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Imf. J. Mol. Biol. 286, 1597–1608.
[15] Fersht, A.R. and Sato, S. (2004) Phi-value analysis and the nature of protein-folding transition states. Proc. Natl. Acad. Sci. U.S.A. 101, 7976–7981.
[16] Fersht, A.R. (1994) Characterising transition states in protein folding: an essential step in the puzzle. Curr. Opin. Struct. Biol. 5, 79–84.
[17] Myers, J.K., Pace, C.N. and Schulz, J.M. (1995) Denaturant in values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci. 4, 2138–2148.
[18] Matouschek, A. and Fersht, A.R. (1993) Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. Proc. Natl. Acad. Sci. U.S.A. 90, 7814–7818.
[19] Matthews, J.M. and Fersht, A.R. (1995) Exploring the energy surface of protein folding by structure-reactivity relationships and engineered proteins: observation of Hammond behavior for the gross structure of the transition state and anti-Hammond behavior for structural elements for unfolding/folding of barnase. Biochemistry 34, 6805–6814.
[20] Hammond, G.S. (1955) A correlation of reaction rates. J. Am. Chem. Soc. 77, 334–338.
[21] Sanchez, I.E. and Kiefhaber, T. (2003) Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states. J. Mol. Biol. 327, 867–884.
[22] Oliverberg, M., Tan, Y.J., Silow, M. and Fersht, A.R. (1998) The changing nature of the protein folding transition state: implications for the shape of the free-energy profile for folding. J. Mol. Biol. 277, 933–943.
[23] 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.
[24] Wright, C.F., Steward, A. and Clarke, J. (2004) Thermodynamic characterisation of two transition states along parallel protein folding pathways. J. Mol. Biol. 338, 445–451.
[25] Oliverberg, M. (1998) Alternative explanations for ‘multistate’ kinetics in protein folding: transient aggregation and changing transition-state ensembles. Acc. Chem. Res. 31, 765–772.
[26] Chen, B.L., Baase, W.A. and Schellman, J.A. (1989) Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. Biochemistry 28, 691–699.
[27] Oliveberg, M. (1998) Alternative explanations for ‘multistate’ kinetics in protein folding: transient aggregation and changing transition-state ensembles. Acc. Chem. Res. 31, 765–772.
[28] Chen, B.L., Baase, W.A. and Schellman, J.A. (1989) Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. Biochemistry 28, 691–699.
[29] Oliveberg, M. (1998) Alternative explanations for ‘multistate’ kinetics in protein folding: transient aggregation and changing transition-state ensembles. Acc. Chem. Res. 31, 765–772.
[30] Becktel, W.J. and Schellman, J.A. (1987) Protein stability curves. Biopolymers 26, 1859–1877.