doi:10.1016/j.jmb.2005.01.032 J. Mol. Biol. (2005) 347, 355–366

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Comprehensive Analysis of Protein Folding Activation Thermodynamics Reveals a Universal Behavior Violated by Kinetically Stable Proteases

Sheila S. Jaswal[1], Stephanie M. E. Truhlar[2], Ken A. Dill[3] and David A. Agard[1]*

[1]*Department of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of California, San Francisco San Francisco, CA 94143 USA*

[2]*Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco San Francisco, CA 94143 USA*

[3]*Department of Pharmaceutical Chemistry, University of California, San Francisco San Francisco, CA 94143 USA*

**Corresponding author*

α-Lytic protease (αLP) and *Streptomyces griseus* protease B (SGPB) are two extracellular serine proteases whose folding is absolutely dependent on the existence of their companion pro regions. Moreover, the native states of these proteins are, at best, marginally stable, with the apparent stability resulting from being kinetically trapped in the native state by large barriers to unfolding. Here, in an effort to understand the physical properties that distinguish kinetically and thermodynamically stable proteins, we study the temperature-dependences of the folding and unfolding kinetics of αLP and SGPB without their pro regions, and compare their behavior to a comprehensive set of other proteins. For the folding activation thermodynamics, we find some remarkable universal behaviors in the thermodynamically stable proteins that are violated dramatically by αLP. Despite significant variations in $\Delta C^{\ddagger}_{P,F}$, the maximal folding speed occurs within the narrow biological temperature range for all proteins, except for αLP, with its maximal folding speed shifted lower by 200 K. This implies evolutionary pressures on folding speed for typical proteins, but not for αLP. In addition, the folding free energy barrier in the biological temperature range for most proteins is predominantly enthalpic, but purely entropic for αLP. The unfolding of αLP and SGPB is distinguished by three properties: a remarkably large $\Delta C^{\ddagger}_{P,U}$, a very high $\Delta G^{\ddagger}_{U}$, and a maximum $\Delta G^{\ddagger}_{u}$ at the optimal growth temperature for the organism. While other proteins display each of these traits to some approximation, the simultaneous optimization of all three occurs only in the kinetically stable proteins, and appears to be required to maximize their unfolding cooperativity, by suppressing local unfolding events, and slowing the rate of global unfolding. Together, these properties extend the lifetime of these enzymes in the highly proteolytic extracellular environment. Attaining such functional properties seems

Present addresses: S. S. Jaswal, Department of Biological Sciences, Stanford University, Stanford, CA 94305-5430, USA; S. M. E. Truhlar, Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0378, USA.

Abbreviations used: αLP, α-lytic protease from *Lysobacter enzymogenes*; SGPB, *Streptomyces griseus* protease B; *C*p, heat capacity; Δ*H*, enthalpy change; Δ*S*, entropy change; Δ*G*, free energy change; TS, transition state ensemble; N, the native state; I, the intermediate ensemble; U, unfolded ensemble; k_U, microscopic unfolding rate constant; GdnHCl, guanidine hydrochloride; $\Delta X^{\ddagger}_{U} = X_{TS} - X_N$, activation function of unfolding where *X* is *G*, *S*, *H*, or C_p; $\Delta X^{\ddagger}_{F} = X_{TS} - X_I$, activation function of refolding; $\Delta G^{\ddagger}_{U,max}$, maximum value of $\Delta G^{\ddagger}_{U}$ with temperature; $T^{\ddagger}_{S,U}$, temperature at which $-T\Delta S^{\ddagger}_{U}$ crosses 0; $\Delta G^{\ddagger}_{F,min}$, minimum value of $\Delta G^{\ddagger}_{F}$ with temperature; $T^{\ddagger}_{S,F}$, temperature at which $-T\Delta S^{\ddagger}_{F}$ crosses 0; ACP, human muscle acylphosphatase; α-specSH3, PWT variant of α-spectrin Src homology region 3 (SH3) domain with second and third residues substituted for a Gly residue; CD2.d1, domain 1 (residues 1–98) of T-cell adhesion protein CD2; CI2, chymotrypsin inhibitor 2; Bc CspB, *Bacillus caldolyticus* cold-shock protein; Bs CspB, *Bacillus subtilis* cold-shock protein; FKBP, human FK506 binding protein 12; Hpr, *Escherichia coli* histidine-containing phosphocarrier protein; N-PGK, *Bacillus stearothermophilus* N-terminal domain (residues 1–175) of phosphoglycerate kinase; NTL9, *Bacillus stearothermophilus* N-terminal domain (residues 1–56) of L9; protein L, Y43W point mutant of protein L; tendamistat, *Streptomyces tendae* α-amylase inhibitor tendamistat; y. isocyto-*c*, *Saccharomyces cerevisiae* iso-2 cytochrome *c*.

E-mail address of the corresponding author: agard@msg.ucsf.edu

possible only through the gross perturbation of the folding thermodynamics, which in turn has required the co-evolution of pro regions as folding catalysts.

Keywords: α-lytic protease; *Streptomyces griseus* protease B; kinetic stability; protein folding kinetics; transition state

Introduction

It is axiomatic that the amino acid sequence of a protein determines uniquely the three-dimensional structure appropriate for its particular function (the native state) and the free energy favoring it over unfolded states.[1] It was once thought that a common property of proteins that fold to stable native states would become evident by comparing the sequence and structural features with the thermodynamic properties of many proteins. There are some universal features of the free energies of protein folding, such as the hydrophobic temperature-dependence,[2] but the enthalpy and entropy convergence temperatures that were once observed in a small set of proteins[3] do not appear to be generally applicable to the broader set of proteins that has been studied since then.[4,5]

By contrast, it now appears that there may be more universal behavior evident in the barriers that separate the native and unfolded states. The free energy barriers between conformations are determined by measuring folding and unfolding rates. With folding kinetics measured for an increasing number of proteins, several empirical relationships have been found linking the folding rate to various properties of protein structure.[6–8] Plaxco *et al.* have shown that the folding rate of a large set of proteins correlates extremely well with the topology of the protein, which is quantified using a parameter known as contact order.[9–11] Proteins with mostly local interactions fold rapidly, while those with more long-range interactions fold more slowly.

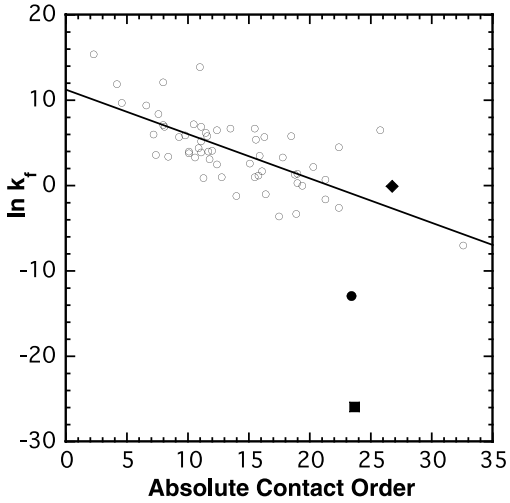

Figure 1. Correlation between topology and folding rate. The folding rates for a set of 29 two-state folders and 28 multi-state folders (black open circles) correlate well (correlation coefficient 0.7) with their absolute contact order (not normalized to length).[11] While trypsin (filled diamond), SGPB (filled circle), and αLP (filled square) all have nearly the same topology (contact order), only trypsin obeys the correlation between contact order and folding rate.

However, as shown in Figure 1, there are quite dramatic exceptions to this observation. The measured folding rates for the extracellular bacterial protease, α-lytic protease (αLP), and its homologue *Streptomyces griseus* protease B (SGPB) are orders of magnitude slower than that predicted by their contact order.[12,13] On the other hand, their mammalian homologue trypsin, with the same overall topology, folds much faster, placing it close to the previously observed correlation between $\ln k_F$ and absolute contact order.[13]

In contrast to trypsin, which folds independently,[14] αLP and SGPB are members of a class of enzymes that are synthesized with pro regions required for their folding.[15] The pro regions catalyze the folding of their respective protease and are subsequently degraded, releasing the mature enzymes.[16,17] The co-evolution of a transient pro region folding catalyst has effectively separated the evolution of the folding landscape from the functional native landscape for these proteases.

Figure 2 illustrates these decoupled landscapes for αLP and SGPB: the folding landscape in the presence of, and the native landscape in the absence of, their pro regions (P). Instead of folding to the native state when diluted from denaturant, both enzymes form stable molten-globule intermediates (I). In the presence of their pro regions, the intermediates for both proteases are folded rapidly to their thermodynamically stable states (native state–pro region (N·P) complexes) due to the stabilization of both their folding transition states and their native states (N) by their pro regions. Following degradation of the pro region to release the folded mature protease, the landscapes change dramatically.

In their native landscapes, the folding free energy barriers are so large that there is virtually no spontaneous folding ($t_{1/2}$ for αLP folding is ~1700 years, $t_{1/2}$ for SGPB folding is about three days).[13] The most striking feature of these landscapes is that the αLP native state is thermodynamically unstable and the SGPB native state is marginally stable. Their native states persist solely due to kinetic stability resulting from the large free energy barriers that prevent unfolding ($t_{1/2}$ for αLP

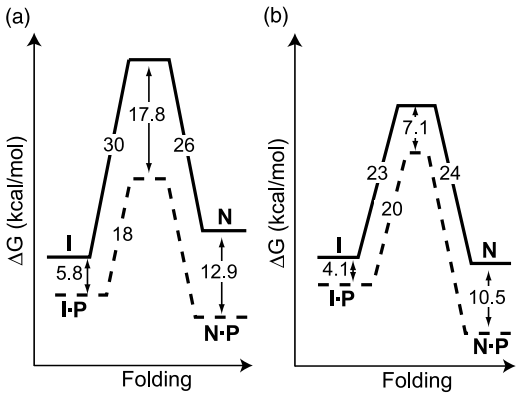

Figure 2. Free energy landscapes for αLP and SGPB folding at 0 °C. Free energies of activation were calculated from folding and unfolding rates using transition state theory.[49] Continuous lines indicate folding in the absence of, and broken lines indicate folding in the presence of, the pro region. Previous experiments demonstrated that the αLP and SGPB intermediate states are equivalent in energy in the absence of the pro region.[13] The diagrams were aligned by positioning the intermediate states at equivalent energies. Binding of the pro region to the native states was measured at 25 °C. (a) The αLP folding landscape.[12] In the absence of its pro region, unfolded αLP folds spontaneously to a molten-globule-like intermediate (I), which proceeds at an extremely slow rate to the native state (N) through a high-energy folding transition state ensemble (TS). The N state is thermodynamically less stable than either the I state or the fully unfolded molecule. The addition of the pro region (P) provides a catalyzed folding pathway with a smaller folding barrier that folds to a thermodynamically stable pro region–native state complex (N·P). (b) The SGPB folding landscape.[13] Unlike αLP, the native state of SGPB (N) is marginally stable compared to its molten-globule-like intermediate (I).[12,13]

unfolding is about one year, $t_{1/2}$ for SGPB unfolding is ~11 days).[13]

This appears to be a result of the functional optimization of their native states in order to survive in the highly proteolytic conditions where they function. Since even subtle subglobal unfolding events can render a protein susceptible to exogenous proteolysis,[18–23] αLP and SGPB have evolved highly rigid native states that lack the typical breathing and partial unfolding motions[13,24] observed in native proteins. Only after their very slow global unfolding do they sample proteolytically vulnerable conformations. This is demonstrated by the significantly extended lifetimes of the kinetically stable enzymes when mixed in proteolysis survival assays with their thermodynamically stable homologs trypsin and chymotrypsin, which undergo more typical native state dynamics that limit their functional longevity.[13,24]

Evolving this functionally desirable protease resistance through the high level of cooperativity of the unfolding process and the large barrier to unfolding has a steep energetic cost on the folding landscape of the mature protease. Comparative studies have revealed that for every 2.4–8-fold increase in proteolytic resistance, there is a 7 kcal/mol increase in the free energy barrier for folding and a 5–10 kcal/mol loss in the thermodynamic stability of the native state. Overcoming these penalties requires assistance from increasingly effective pro regions during folding.[13]

As a first step in discovering the structural principles that lead to the very high kinetic barriers that are the hallmark of these pro-dependent proteases, we have mapped the thermodynamic activation parameters of their folding and unfolding transition states. In addition, we present the first comprehensive compendium of the thermodynamic activation parameters underlying unfolding and folding kinetics for a large set of thermodynamically stable proteins. The comparison of αLP and SGPB activation parameters with those of thermodynamically stable proteins reveals how kinetic stability differs specifically, and the extent that thermodynamic restrictions limit a single landscape responsible for both folding and native state function.

Results and Analysis

The αLP and SGPB unfolding barriers

We determined the thermodynamic activation parameters heat capacity ($\Delta C^{\ddagger}_{P,U}$), entropy ($\Delta S^{\ddagger}_{U}$), and enthalpy ($\Delta H^{\ddagger}_{U}$), of the αLP and SGPB unfolding free energy barriers by monitoring the temperature-dependence of their unfolding kinetics. The unfolding rates for αLP at 4 °C and SGPB at 0 °C in the absence of denaturant have been determined from linear extrapolations of denaturant-dependent unfolding data.[12,13] The unfolding rates for αLP and SGPB from 15–70 °C are shown in Figure 3(a) and (b), respectively. At high temperatures, the unfolding rates were measured directly in the absence of denaturant. At other temperatures, the unfolding rates were measured in the presence of various concentrations of denaturant, and were extrapolated to zero denaturant. The denaturant-binding model was chosen because it empirically provided the best fit to the observed curvature. Evidence described in Materials and Methods has validated its use in extrapolation and ruled out transition state movement as the cause of the curvature. The thermodynamic activation parameters $\Delta H^{\ddagger}_{U}$, $\Delta S^{\ddagger}_{U}$, and $\Delta C^{\ddagger}_{P,U}$ were determined from a fit of the temperature-dependence of the αLP and SGPB unfolding rate constants in the absence of denaturant to equation (9) (Figure 3(c), see Materials and Methods).

The Eyring analysis in Figure 3(c) shows a striking difference in the temperature-dependence of αLP and SGPB unfolding kinetics. Both αLP and SGPB have a very large $\Delta C^{\ddagger}_{P,U}$, but $\Delta C^{\ddagger}_{P,U}$ for SGPB is much larger than that for αLP (Table 1), leading to the narrower shape of the curve. The

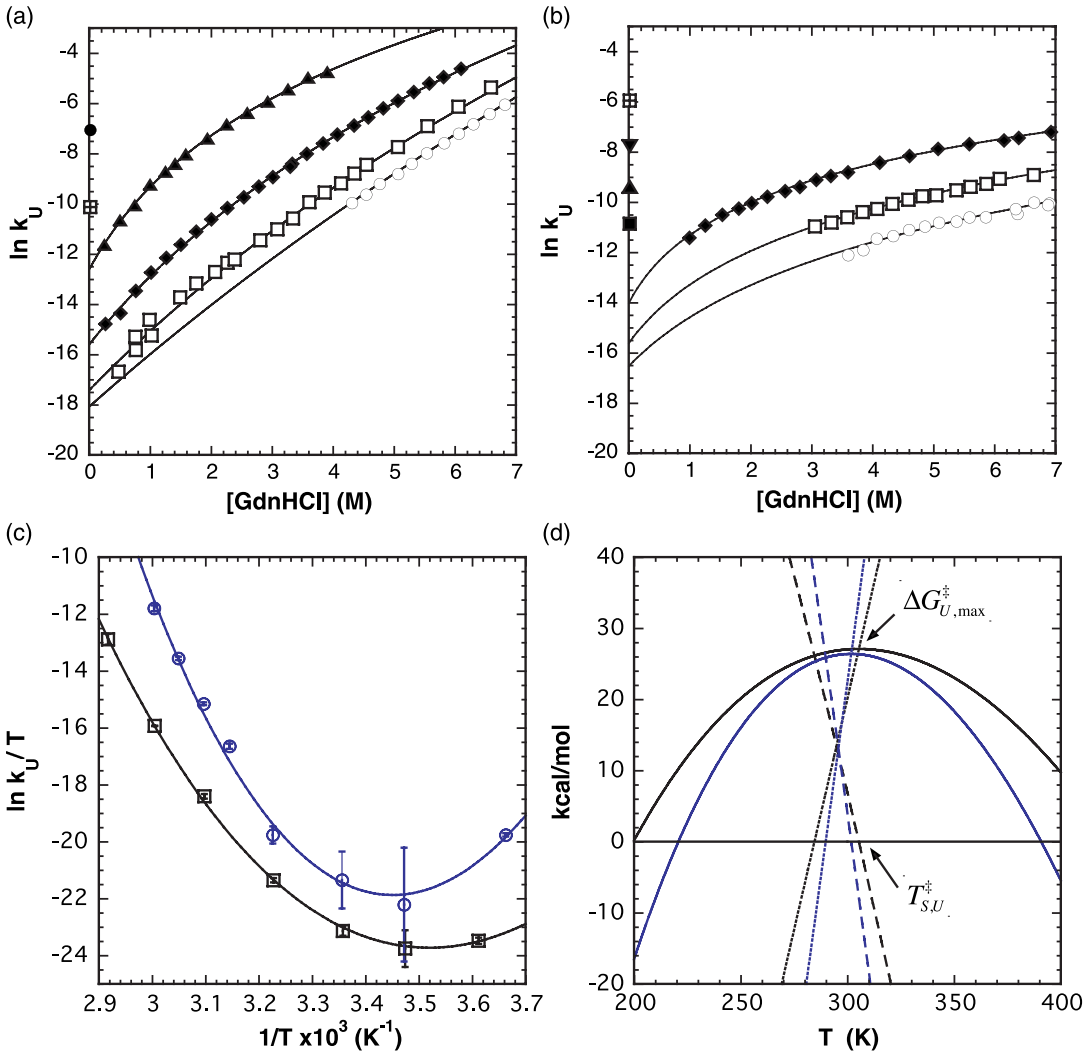

Figure 3. Temperature-dependence of αLP and SGPB unfolding kinetics. (a) Unfolding rate constants for αLP as a function of denaturant at 15 °C (open circles), 25 °C (open squares), 37 °C (diamonds), and 50 °C (triangles). Unfolding rates in the absence of denaturant were extrapolated using the denaturant binding model (continuous lines, see Materials and Methods). At 60 °C (hatched square) and 70 °C (filled circle), unfolding was measured directly in 0 M GdnHCl. (b) Unfolding rate constants for SGPB from 15–37 °C were extrapolated from denaturant-dependent unfolding data using the denaturant-binding model. From 45–55 °C, SGPB unfolding rates were measured directly in 0 M GdnHCl. Symbols and lines as in (a) and 45 °C (filled square), 55 °C (inverted triangle). (c) Temperature-dependence of αLP and SGPB unfolding kinetics. ln (k_U/T) *versus* $1/T$ for αLP (black squares) and SGPB (blue circles) are well fit by an Eyring analysis ($R = 0.99973$ and 0.99627, respectively). The αLP unfolding rate at 4 °C was taken from Sohl *et al.*[12] and the SGPB unfolding rate at 0 °C was taken from Truhlar *et al.*[13] Error bars indicate the propagated standard error in the extrapolated or measured k_U. (d) αLP and SGPB unfolding barrier stability curves. Temperature-dependence of $\Delta G^{\ddagger}_U$ (continuous line), $\Delta H^{\ddagger}_U$ (dotted line), and $-T\Delta S^{\ddagger}_U$ (broken line), (unfavorable >0, favorable <0) for αLP (black) and SGPB (blue) unfolding. The maximum barrier height ($\Delta G^{\ddagger}_{U,max}$) occurs at $T^{\ddagger}_{S,U}$ (305 K for αLP, 302 K for SGPB) where $-T\Delta S^{\ddagger}_U$ crosses 0, switching from activation entropy opposing unfolding to favoring it.

change in heat capacity upon unfolding is highly correlated with the exposure of surface.[25–27] The larger $\Delta C^{\ddagger}_{P,U}$ for SGPB indicates that its transition state is more unfolded (exposed) than the αLP transition state.

Since $\Delta C^{\ddagger}_{P,U}$ determines the slopes of $\Delta H^{\ddagger}_U$ and $-T\Delta S^{\ddagger}_U$ with temperature (Materials and Methods, equations (7) and (8)), the large $\Delta C^{\ddagger}_{P,U}$ for αLP and SGPB unfolding indicates that these change rapidly with temperature. Therefore, instead of analyzing the thermodynamic activation parameters of the unfolding free energy barriers at just one temperature, we investigated the overall thermodynamic profiles of the activation parameters, including temperature-dependence. To this end, we adapted the protein stability curve analysis developed by Becktel & Schellman, which illustrates the temperature-dependence of ΔG, ΔH, and $-T\Delta S$ for equilibrium unfolding,[28] and was first used in a comprehensive manner by Agashe & Udgaonkar.[29] This approach has been used to examine thermodynamic differences between mesophilic and thermophilic homologs and other protein variants.[30–32]

Table 1. Thermodynamics of SGPB unfolding and αLP unfolding and folding at 10 °C

	SGPB	αLP		
Parameter	N unfolding to TS (X_{TS}–X_N)	N unfolding to TS (X_{TS}–X_N)[a]	I refolding to TS (X_{TS}–X_I)[b]	N unfolding to I (X_I–X_N)[c,d]
ΔG (kcal/mol)	+25±3	+26±1	+30±1[c]	−4±0.1
ΔH (kcal/mol)	−15±3	−2±1	−20±2	+18±2
$-T\Delta S$ (kcal/mol)	+40±3	+28±1	+50±2	−22±2
ΔC_p (kcal/(mol K))	+2.2±0.2	+1.3±0.05	−0.17±0.005	+1.47±0.05[b]

[a] From this work.
[b] Calculated from **a** and **d**.
[c] Measured at 4 °C.
[d] From Sohl *et al.*[12]

We generated unfolding barrier stability curves by plotting the temperature-dependence of $\Delta G^{\ddagger}_{U}$, $\Delta H^{\ddagger}_{U}$, and $-T\Delta S^{\ddagger}_{U}$. The result is a thermodynamic fingerprint (Figure 3(d)), where the unfolding barrier for each protein is characterized by the degree to which its thermodynamic activation parameters change with temperature as dictated by $\Delta C^{\ddagger}_{P,U}$. This is illustrated in the degree of curvature of $\Delta G^{\ddagger}_{U}$ with temperature and the steepness of the $\Delta H^{\ddagger}_{U}$ and $-T\Delta S^{\ddagger}_{U}$ curves. In addition, the $\Delta G^{\ddagger}_{U}$ curve is characterized by both its maximum value and the temperature where that maximum is reached, which occurs where $-T\Delta S^{\ddagger}_{U}$ crosses zero ($T^{\ddagger}_{S,U}$).

As with most proteins, in the experimental temperature range the free energy of activation for unfolding is unfavorable ($\Delta G^{\ddagger}_{U} > 0$) for αLP and SGPB. For both proteins, the unfolding free energy barriers at low temperature are dominated completely by unfavorable activation entropy ($-T\Delta S^{\ddagger}_{U} > 0$). It is not until rather high temperatures that both proteins reach their $T^{\ddagger}_{S,U}$ (∼300 K) and the favorable configurational entropy is able to compensate for the unfavorable entropy from solvating the exposed hydrophobic surfaces. In contrast to the activation entropy, at low temperatures the activation enthalpy actually favors unfolding to the transition state for both proteins ($\Delta H^{\ddagger}_{U} < 0$), presumably from additional solvent-based interactions in the transition state, but as the temperature increases, the activation enthalpy becomes unfavorable. Above ∼300 K, activation enthalpy dominates the unfolding free energy barriers due to the loss of favorable protein–protein interactions present in their native states.

The narrower temperature range for the SGPB $\Delta G^{\ddagger}_{U}$ curve and the steeper temperature-dependence of the $\Delta H^{\ddagger}_{U}$ and $-T\Delta S^{\ddagger}_{U}$ curves, compared to those for αLP, again highlight SGPB's larger $\Delta C^{\ddagger}_{P,U}$. As a result of this larger $\Delta C^{\ddagger}_{P,U}$, the unfolding free energy barrier of SGPB decreases more rapidly from its maximum than does the αLP $\Delta G^{\ddagger}_{U}$, allowing αLP to maintain a larger unfolding free energy barrier at all temperatures compared to SGPB. In a similar study that measured the thermodynamics underlying the different equilibrium stabilities of thermophilic and mesophilic RNase H* (RNase H* is a cysteine-free variant of RNase H1), Marqusee and co-workers demonstrated that the thermostable variant has a smaller $\Delta C_{P,NU}$ (1.8 *versus* 2.7 kcal/(mol K)),

Table 2. Thermodynamic activation parameters of αLP, SGPB and 13 thermodynamically stable proteins

Protein	N_r	$\Delta C^{\ddagger}_{P,U}$ (kcal/(mol K))	$\Delta G^{\ddagger}_{U,max}$ (kcal/mol)	$T^{\ddagger}_{S,U}$ (K)	$\Delta C^{\ddagger}_{P,F}$ (kcal/(mol K))	$\Delta G^{\ddagger}_{F,min}$ (kcal/mol)	$T^{\ddagger}_{S,F}$ (K)
αLP[a]	198	1.3±0.05	27.1	305	−0.17±0.005	11.1	100
SGPB	185	2.2±0.2	26.4	302	n.d.	n.d.	n.d.
ACP[b,39]	98	0.91±0.05	24.1	271.5	−0.61±0.04	19.4	285.5
α-specSH3[40]	62	0.17	20.1	298	−0.55±0.02	16.4	284.5
CD2.d1[a,c,41]	98	0.57±0.1	21.9	283.5	−0.41±0.05	16.5	296.5
CI2[50]	64	0.48±0.02	23.4	270	−0.49±0.01	14.9	294
Bc CspB[31]	67	0.26±0.02	18.1	269	−0.53±0.02	13.2	288
Bs CspB[31]	67	0.07±0.1	16.6	209.5	−0.64±0.07	13.2	289
FKBP[43]	107	n.d.	n.d.	n.d.	−0.67±0.01	16.6	290.5
Hpr[d,32]	85	0.74±0.1	19.1	296	−0.77±0.04	16.6	298
N-PGK[a,c,44]	175	0.45±0.06	20.0	301.5	−0.28±0.08	16.0	306
NTL9[c,45]	56	0.28±0.11	17.9	285	−0.60±0.08	13.3	295.5
Protein L[c,46]	62	0.19±0.03	21.9	269.5	−0.32±0.02	15.9	275
Tendamistat[47]	74	0.60±0.03	23.5	289	−0.49±0.01	14.8	281
y. isocyto-c[48]	103	0.6±0.3	19.9	272	−1.0±0.3	16.6	289.5

[a] Folding from an intermediate.
[b] Measured in 1.1 M urea.
[c] Original $k'T$ used was 1.00×10^{10}.
[d] Measured in 0.55 M GdnHCl.

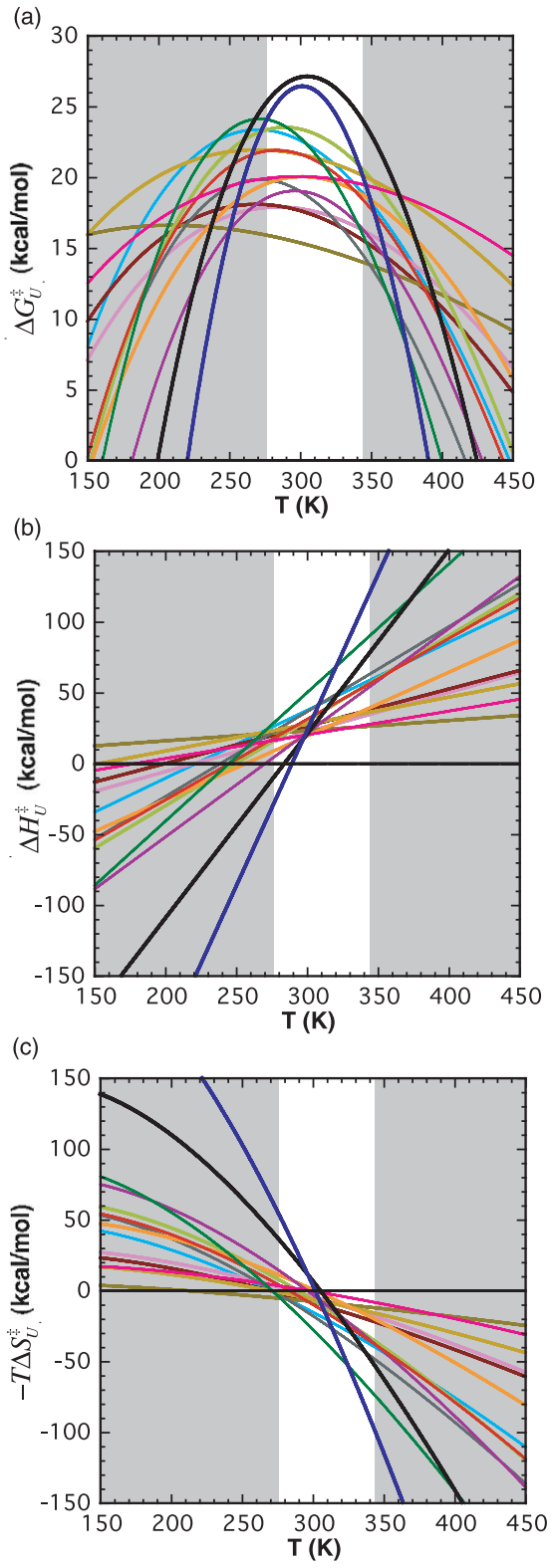

Figure 4. Temperature-dependence of unfolding activation thermodynamics. Temperature-dependence of activation thermodynamic parameters for αLP (black), SGPB (blue), ACP (green), α-spec SH3 (pink), CD2.d1 (red), CI2 (cyan), Bc CspB (brick), Bs CspB (brown), Hpr (purple), N-PGK (orange), NTL9 (lilac), protein L (gold), tendamistat (light green), and y- isocyto-*c* (gray) unfolding. For ease of comparison, curves have been extrapolated to a wide temperature range; however, the regions outside 0–70 °C are shaded to indicate their exclusion from the experimental temperature range. (a) $\Delta G^{\ddagger}_{U}$, (b) $\Delta H^{\ddagger}_{U}$, and (c) $-T\Delta S^{\ddagger}_{U}$.

allowing it to maintain greater stability at the higher temperatures at which it normally functions.[30]

While αLP maintains a larger unfolding free energy barrier than SGPB at all temperatures, the difference between them at physiological temperatures is much smaller than at 0 °C, where they had been compared previously.[13] In fact, SGPB's $\Delta G^{\ddagger}_{U,max}$ is only 0.7 kcal/mol lower than that for αLP (Table 2). Furthermore, $\Delta G^{\ddagger}_{U,max}$ for both αLP and SGPB occur at ~300 K ($T^{\ddagger}_{S,U}$, Table 2), coinciding with the optimal growth temperatures for their bacterial hosts (303 K and 298 K, respectively), indicating that they are maximally stable under physiological conditions.

Thermodynamic activation parameters of unfolding barriers

The similar profiles with temperature for the unfolding activation parameters of αLP and SGPB suggest a common thermodynamic mechanism by which proteins develop highly cooperative unfolding transitions and high degrees of kinetic stability. Do these hallmarks (large $\Delta C^{\ddagger}_{P,U}$ and large $\Delta G^{\ddagger}_{U}$ with maxima at physiological temperatures) distinguish them from thermodynamically stable proteins? How widely do unfolding barrier activation parameters vary among proteins that are constrained to concurrently evolve folding and function in a single energy landscape? To answer these questions, we collected data on activation thermodynamics from 13 proteins in the literature for a comprehensive barrier analysis (Table 2, see Materials and Methods).

Figure 4 shows the (a) $\Delta G^{\ddagger}_{U}$, (b) $-T\Delta S^{\ddagger}_{U}$, and (c) $\Delta H^{\ddagger}_{U}$ curves for αLP (black), SGPB (blue), and the thermodynamically stable proteins. The biggest difference observed is that αLP and SGPB have the two largest $\Delta C^{\ddagger}_{P,U}$ values (larger by 0.4 and 1.3 kcal/(mol K), respectively), and therefore the narrowest free energy barrier curves and the steepest entropy and enthalpy curves. Despite the narrowness of their $\Delta G^{\ddagger}_{U}$ curves, their maximum free energy barriers are so large that SGPB maintains a larger $\Delta G^{\ddagger}_{U}$ than the others between 4 °C and 69 °C, and the free energy barrier of αLP is the largest between 4 °C and 97 °C. Their $\Delta G^{\ddagger}_{U,max}$ values are the highest (by 3.0 kcal/mol and 2.3 kcal/mol, respectively), and occur at the highest $T^{\ddagger}_{s,u}$ (305 K and 302 K, respectively).

The thermodynamically stable proteins themselves vary more widely in their behavior than might be expected from the modest variations in $\Delta C^{\ddagger}_{P,U}$ (0.07–0.91 kcal/(mol K)), with broad ranges observed for both their maximum $\Delta G^{\ddagger}_{U}$ (16.6–24.1 kcal/mol) and their $T^{\ddagger}_{S,U}$(209.5–301.5 K). Although αLP and SGPB are the largest proteins

in the set, there is no correlation between the number of residues and any of the parameters that uniquely characterize the unfolding barrier stability curves ($\Delta G^{\ddagger}_{U,max}$, $T^{\ddagger}_{S,U}$, or $\Delta C^{\ddagger}_{P,U}$). Interestingly, there is also no clear correlation between $\Delta C^{\ddagger}_{P,U}$ and the $\Delta G^{\ddagger}_{U,max}$ or $T^{\ddagger}_{S,U}$. Acylphosphatase (ACP), the protein with the largest $\Delta C^{\ddagger}_{P,U}$ after αLP and SGPB, does have the highest maximum $\Delta G^{\ddagger}_{U}$ of the thermodynamically stable proteins, but its $T^{\ddagger}_{S,U}$ is 271.5 K. On the other hand, the histidine-containing phosphocarrier protein Hpr, which has the next largest $\Delta C^{\ddagger}_{P,U}$ displays nearly the highest $T^{\ddagger}_{S,U}$ (296 K) of the set, but its $\Delta G^{\ddagger}_{U,max}$ is among the lowest.

The αLP folding barrier

Because of the thermal instability of the I state, it is not possible to perform an Eyring analysis on αLP and SGPB refolding to extract the thermodynamic activation components of their folding free energy barriers directly. However, the previously determined thermodynamics of the equilibrium between the αLP N and I[12] allowed us to use the unfolding activation parameters determined here to calculate the αLP folding barrier activation thermodynamic parameters as well (Table 1). This reveals that, as with the unfolding free energy barrier, unfavorable activation entropy dominates the folding free energy barrier of αLP at 10 °C. Previous titration calorimetry results showed that the I state is stabilized over the N state from an excess of favorable entropy, likely resulting from the additional configurational entropy due to αLP's unusually high glycine content.[12] In folding to the TS state, the loss of this excess configurational entropy contributes to the dominating unfavorable activation entropy.

We then calculated the folding barrier stability curve for αLP (Figure 5(a)). Its appearance is inverted compared to the unfolding barrier stability

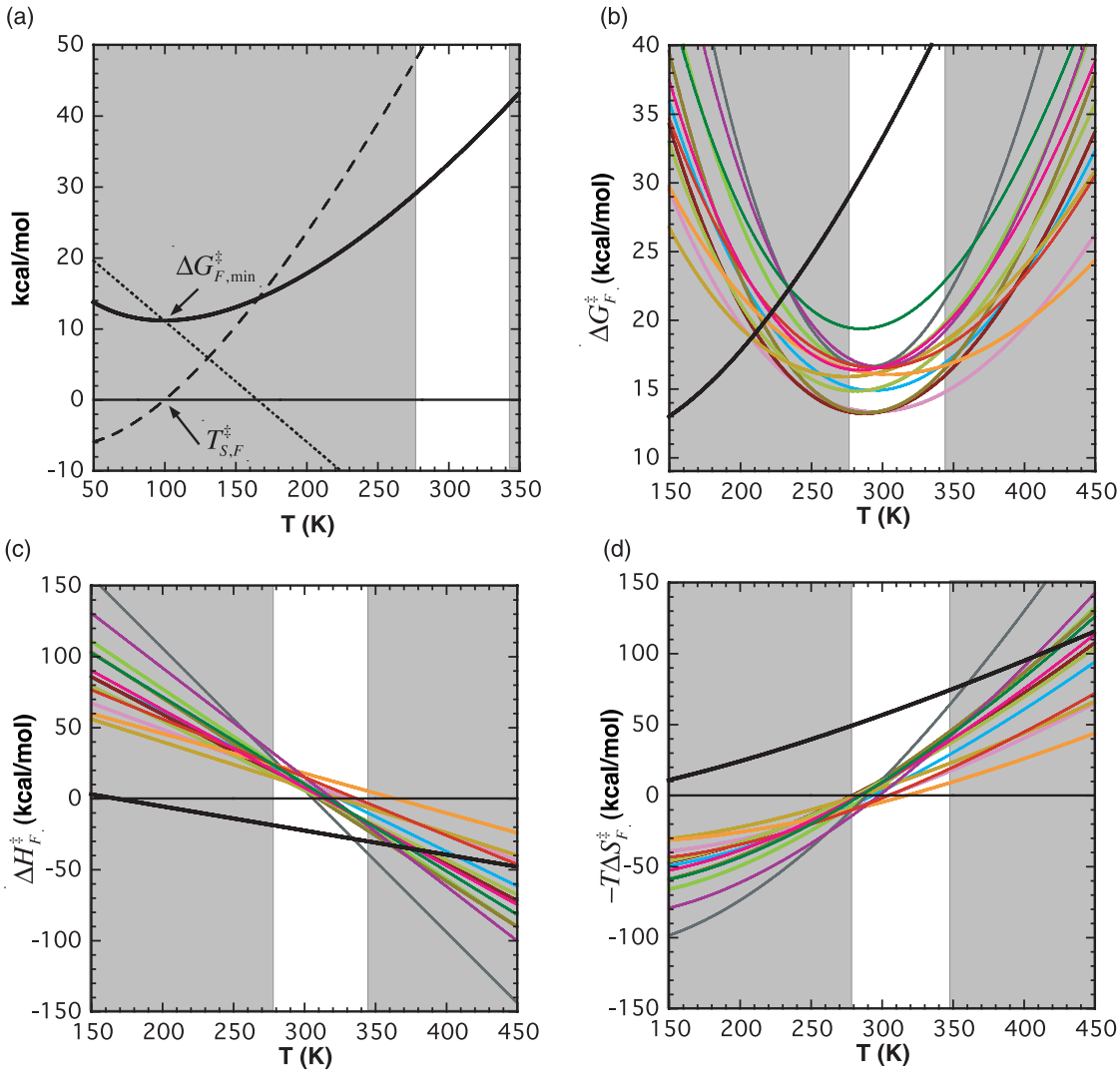

Figure 5. Temperature-dependence of folding activation thermodynamics. (a) αLP folding barrier stability curve. For the αLP I to TS folding transition, $\Delta G^{\ddagger}_{F}$ reaches a minimum at the temperature where $-T\Delta S^{\ddagger}_{F}$ crosses 0 K ($T^{\ddagger}_{S,F}$), switching from favorable to unfavorable. (b) $\Delta G^{\ddagger}_{F}$, (c) $\Delta H^{\ddagger}_{F}$, and (d) $-T\Delta S^{\ddagger}_{F}$ with temperature for the entire set of proteins. Colors are as in Figure 4, and neon green for FKBP.

curve, because surface area is buried during folding, leading to a negative change in activation heat capacity. αLP's minimum $\Delta G^{\ddagger}_{F}$ occurs at a $T^{\ddagger}_{S,F}$ of 100 K. As a result, the folding free energy barrier of αLP increases with temperature over the entire accessible temperature range, due to increasing unfavorable activation entropy of folding. In contrast to the large $\Delta C^{\ddagger}_{P,U}$ upon unfolding, the $\Delta C^{\ddagger}_{P,F}$ of αLP is relatively small, leading to the observed shallow dependencies of $\Delta G^{\ddagger}_{F}$, $\Delta H^{\ddagger}_{F}$, and $-T\Delta S^{\ddagger}_{F}$ on temperature.

Thermodynamic activation parameters of folding barriers

While the unfolding activation parameters of αLP and SGPB are at the extreme of the range observed for other proteins, dramatically different behavior is revealed for αLP when its folding activation parameters are compared with the 13 thermodynamically stable proteins from the literature. It is immediately apparent in Figure 5(b) that the entire folding barrier stability profile for αLP is shifted to far lower temperatures than those of the other proteins. Its minimum occurs at a $T^{\ddagger}_{S,F}$ nearly 200 K below the next lowest $T^{\ddagger}_{S,F}$. In addition, αLP's activation entropy for folding becomes increasingly unfavorable throughout the accessible temperature range, and is far larger than all the others up to 330 K. At 300 K, $-T\Delta S^{\ddagger}_{F}$ for αLP is more unfavorable by 40 kcal/mol than the next largest entropic penalty. Clearly, the folding free energy barrier of αLP possesses significant excess activation entropy relative to the other proteins.

Figure 5 emphasizes the relatively small $\Delta C^{\ddagger}_{P,F}$ of αLP (only -0.17 kcal/(mol K)), which impacts $-T\Delta S^{\ddagger}_{F}$ as well, indicating only a small favorable solvent entropy gain from surface burial upon folding from the I state to the TS state. Two other proteins that fold from intermediates, CD2.d1 and NPGK, also have two of the smallest $\Delta C^{\ddagger}_{P,F}$ values (-0.41 kcal/(mol K) and -0.28 kcal/(mol K)). However, their minima are not shifted to lower temperatures like that of αLP, but instead occur at two of the highest $T^{\ddagger}_{S,F}$ observed (296.5 K and 306 K). Therefore, while folding from an intermediate and/or having a small $\Delta C^{\ddagger}_{P,F}$, may contribute partially to αLP's unusually entropic folding free energy barrier, the primary cause must be the loss of the excess configurational entropy of the I state.

In addition to highlighting αLP's unusual folding activation parameters, Figure 5 illuminates an apparent universal behavior among the thermodynamically stable proteins. Despite the wide range of $\Delta C^{\ddagger}_{P,F}$ (-0.28 kcal/(mol K) to -1.6 kcal/(mol K)), they all reach their $\Delta G^{\ddagger}_{F,min}$ in a very narrow temperature range. This is visually apparent in the centered behavior of the $\Delta G^{\ddagger}_{F}$ curves and the tightly clustered intersections in the $-T\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}_{F}$ plots. In contrast to the $T^{\ddagger}_{S,U}$ values, which range from 209.5 K to 301.5 K, the $T^{\ddagger}_{S,F}$ values range from 275 K to 306 K.

Discussion

The large $\Delta C^{\ddagger}_{P,U}$ observed for αLP and SGPB is the most significant feature of their unfolding thermodynamic activation parameters. Exposure of such a large amount of surface area in both transition states implies the existence of a significant network of interactions in their native states that is released only in a highly concerted unfolding transition. Therefore, their large $\Delta C^{\ddagger}_{P,U}$ values are likely to be linked directly to the observed high cooperativity in their unfolding reactions (i.e. the suppression of local unfolding transitions).[13,24] There may be other subtleties in the balance of surface exposure and the degree of structure in the transition state ensemble, and the presence of strain in the native state that contribute to optimizing the thermodynamic activation parameters.

When compared with the thermodynamically stable proteins, in addition to having the largest $\Delta C^{\ddagger}_{P,U}$ values, the $T^{\ddagger}_{S,U}$ and $\Delta G^{\ddagger}_{U,max}$ for ΔLP and SGPB are the highest in the set as well. This indicates that their unfolding barriers are maximized at temperatures closer to physiological, which is more critical for αLP and SGPB because their stability is determined solely from their unfolding free energy barriers. In addition, because $T^{\ddagger}_{S,U}$ denotes the temperature at which activation entropy switches from unfavorable to favorable, activation entropy dominates the unfolding barriers of αLP and SGPB at higher temperatures than for the other proteins.

The unfolding activation parameters observed for the thermodynamically stable proteins vary widely and in an uncorrelated fashion. It appears that there are various mechanisms for increasing $\Delta G^{\ddagger}_{U,max}$ and $\Delta C^{\ddagger}_{P,U}$ or reaching $T^{\ddagger}_{S,U}$ at physiological temperatures, which operate independently in most proteins. While the values of these parameters for αLP and SGPB merely extend the already broad range observed previously, these proteins appear to be unique, in having evolved to maximize all three of these parameters simultaneously.

In contrast to unfolding kinetics, our comprehensive analysis of folding kinetics has revealed that thermodynamically stable proteins themselves exhibit a remarkable coincidence of their folding activation parameters. The $\Delta H^{\ddagger}_{F}$ and $-T\Delta S^{\ddagger}_{F}$ curves intersect in, and $\Delta G^{\ddagger}_{F}$ curves are centered around, a tight temperature range ($T^{\ddagger}_{S,F}$ 290($\pm$8) K). This is observed for proteins ranging in size from 56–175 amino acid residues, with $\Delta C^{\ddagger}_{P,F}$ values from -0.28 kcal/(mol K) to -1.0 kcal/(mol K), and includes those folding through intermediates, as well as mesophilic *versus* thermophilic homologues (CspB) differing by twofold in equilibrium stability. This striking result suggests that the evolutionary pressure on a landscape required to optimize folding and native properties simultaneously imposes quite a narrow restriction on the activation thermodynamics of folding free energy barriers.

Convergence of behavior in folding transition states was observed by Munoz and co-workers in a

separate analysis of the temperature-dependence of unfolding and folding kinetics for six proteins.[33] Using a different treatment of transition state theory and additional assumptions to separate solvation from conformational contributions to activation entropy, they found that all of the transition states appear to possess a similar fraction of conformational entropy per residue. Despite the different approaches, both their analysis and ours point to a distinct restriction of folding thermodynamics for normal proteins.

The violation of this universal behavior by αLP demonstrates that, with separate folding and native landscapes it is possible to evolve a native state with dramatically different activation parameters comprising the folding barrier separating it from unfolded states. Presumably, αLP folding in the context of the pro region would conform to the universal behavior. Although the full thermodynamic profile for the pro region-catalyzed folding reaction is not known, the free energy barrier to folding in the presence of the pro region at 273 K is 18 kcal/mol,[34] which places it squarely in the range of the $\Delta G_F^{\ddagger}$ observed for the thermodynamically stable proteins.

Interestingly, Baker and co-workers have shown that, when the pressure for native state function is relieved, variants of existing proteins can be redesigned completely to fold and/or unfold faster than the functional parent.[35] In addition, they computationally designed and optimized a 93 amino acid residue sequence to adopt a completely novel structure.[36] This sequence, which was evolved *in silico*, solely for native state stability but not folding or function, does in fact fold to a very stable native state, but with more complex folding kinetics than observed typically for small proteins. These examples further suggest that balancing the multiple evolutionary pressures of folding, stability, and function within a single conformational landscape restricts the activation thermodynamics for both folding and unfolding dramatically. Relieving some of those evolutionary pressures allows for the exploration of otherwise inaccessible thermodynamic space.

Nature appears to have already set a more generous range for unfolding than folding thermodynamic activation parameters, as a large range is observed in unfolding parameters with no evidence for a universal behavior. This is presumably required to accommodate the vast array of different needs for function, longevity, and regulation. However, above some threshold, the small improvements to simultaneously tune these parameters to the extreme observed for αLP and SGPB incur huge costs on folding. The pro region-containing proteins overcome these penalties by decoupling their folding and native landscapes. Proteins that fold with chaperone assistance, those that function in environments different from where they fold, and those functioning as components of larger complexes may offer examples of other strategies that decouple folding and native landscapes to allow extreme optimization of native states.

Materials and Methods

Materials

Guanidine hydrochloride (GdnHCl) was from ICN (Cleveland, OH). Stocks were filtered through a 0.22 μm pore-size filter and the concentration was determined using refractive index measurements.[37] All other chemicals were from Fisher.

Measurement of αLP unfolding kinetics

The proteolytically inactive SA195 variant of αLP, which was expressed and purified as described,[12] was used for unfolding experiments. SA195 has a crystal structure identical with that of wild-type αLP. Unfolding reactions were performed in 10 mM potassium acetate (pH 5.0), and they were initiated by manual mixing to a final protein concentration of 0.1–1.75 μM. At 15 °C, 25 °C, 37 °C and 50 °C, unfolding was carried out at a series of concentrations of GdnHCl. Refractive index measurements were used to measure the concentration of GdnHCl in each sample accurately.[37] At 60 °C and at 70 °C unfolding was measured directly in the absence of denaturant. Unfolding was followed by measuring change in intrinsic tryptophan fluorescence ($\lambda_{excitation}$ = 283 nm, $\lambda_{emission}$ = 322 nm), using an 8100SLM-Aminco fluorometer connected to an external, thermostatically controlled bath. Previous experiments demonstrating the coincidence of unfolding rates monitored by fluorescence and circular dichroism support the assumption of two-state unfolding for αLP.[12]

Measurement of SGPB unfolding kinetics

SGPB was expressed and purified as described.[13] Unfolding reactions containing 1.75 μM SGPB in 10 mM potassium acetate (pH 5.0) were initiated by manual mixing and followed by measuring change in intrinsic tryptophan fluorescence ($\lambda_{excitation}$ = 290 nm, $\lambda_{emission}$ = 350 nm) using a Fluoromax-3 (J. Y. Horiba), connected to an external water-bath. Unfolding rates from 15 °C to 37 °C were measured in a series of concentrations of GdnHCl, where refractive index measurements were used to measure the concentration of GdnHCl accurately. From 45 °C to 60 °C, unfolding was measured directly in the absence of denaturant. Autolysis was minimized by performing all unfolding reactions at pH 5.0. Even at high temperatures, inhibition of the proteolytic activity resulted in the identical unfolding rate. Previous experiments demonstrating the coincidence of unfolding rates monitored by fluorescence and loss of enzymatic activity support the assumption of two-state unfolding for SGPB.[13]

Analysis of unfolding kinetics

Non-linear regression analysis with Kaleidagraph (version 3.6; Synergy Software) was used to obtain the mono-exponential rate constants for αLP unfolding using:

$$y = a - b\exp^{-k_U t} \quad (1)$$

SGPB unfolding data were fit to either equation (1) or equation (2):

$$y = a - b\exp^{-k_U t} - ct \quad (2)$$

when necessary, to determine the rate constants for unfolding. Due to the curvature in the dependence of $\ln k_U$ on [GdnHCl] at 15 °C and above for both αLP and SGPB, the data were best fit by the denaturant-binding model.[37] For equilibrium unfolding, this model assumes a discrete number of binding sites for denaturant on the unfolded molecule:

$$\ln k_{U,GdnHCl} = \ln k_{U,H_2O} + \Delta n \ln(1 + K[\text{GdnHCl}]) \quad (3)$$

where Δn is the difference in the number of binding sites between the native and the unfolded molecule, and K is the equilibrium constant for denaturant binding. However, although this model has been applied to the denaturant-dependence of unfolding kinetics,[38] it is not clear that a physical interpretation of the fit parameters in terms of number of denaturant binding sites on the transition state and binding constant of denaturant to the transition state is warranted, due to the highly transitory nature of the transition state. Therefore, we have treated equation (3) as an empirical equation that best fits the data. In addition, we have validated its use for αLP by showing that the resulting rate constant extrapolated to water from measurements in GdnHCl matches exactly the directly measured unfolding rate constant at 60 °C, where direct measurement of unfolding is possible (S.S.J. & D.A.A., unpublished results).

Eyring analysis of unfolding kinetics

Applying transition state theory by the method of Chen *et al.*,[38] we determined the activation free energies ($\Delta G^{\ddagger}_U$) at each concentration of denaturant from the unfolding rate constants:

$$\Delta G^{\ddagger}_U = -RT \ln(k_U/k') \quad (4)$$

where k' is the pre-exponential factor, which accounts for the rate of transition state decay, originally taken from chemical collision theory to be simply the rate of a molecular vibration:

$$k' = k_B T/h \quad (5)$$

where k_B is Boltzman's constant (1.3805×10^{-23} J/K), and h is Planck's constant (6.625×10^{-34} J s).

$\Delta G^{\ddagger}$ is comprised of a balance of enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$):

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} \quad (6)$$

$\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ vary with temperature, due to the difference in the heat capacities ($\Delta C^{\ddagger}_P$) of the N state and the TS state:

$$\Delta H^{\ddagger} = \Delta H^{\ddagger}_{T0} + \Delta C^{\ddagger}_P(T - T_0) \quad (7)$$

$$\Delta S^{\ddagger} = \Delta S^{\ddagger}_{T0} + \Delta C^{\ddagger}_P \ln(T/T_0) \quad (8)$$

where T_0 is a reference temperature at which $\Delta H^{\ddagger}_0$ and $\Delta S^{\ddagger}_0$ are known. When they are not known at any temperature, they can be determined by monitoring the change in k_U with temperature, which is how we determined them here. The data were fit to the following equation, which combines equations (4)–(8):[38]

$$\ln\left(\frac{k_U}{T}\right) = \frac{\Delta S^{\ddagger}_{T_0,U} - \Delta C^{\ddagger}_{P,U}}{R} + \ln\left(\frac{h}{k_B}\right) + \frac{\Delta C^{\ddagger}_{P,U} - \Delta H^{\ddagger}_{T_0,U}/T_0}{R}\left(\frac{T_0}{T}\right) - \frac{\Delta C^{\ddagger}_{P,U}}{R}\ln\left(\frac{T_0}{T}\right) \quad (9)$$

Calculation of αLP folding parameters at 10 °C

The thermodynamic activation parameters for αLP folding (from I to the TS) were calculated using the unfolding $\Delta H^{\ddagger}_U$ and $-T\Delta S^{\ddagger}_U$ at 10 °C determined here, and the previously determined ΔH_{NI} and $-T\Delta S_{NI}$ at 10 °C.[12] We then used those values with the previously determined k_f at 4 °C and at 25 °C to calculate $\Delta C^{\ddagger}_{P,F}$ using a minimum least-squares calculation (Table 1).

Comparison with activation parameters of thermodynamically stable proteins

Thermodynamic activation parameters for the unfolding and folding transition states of thermodynamically stable proteins were collected from the literature (Table 2). The proteins included are human muscle acylphosphatase (ACP),[39] PWT variant of α-spectrin Src homology region 3 (SH3) domain with second and third residues substituted for a Gly residue (α-specSH3),[40] domain 1 (residues 1–98) of T-cell adhesion protein CD2 (CD2.d1),[41] chymotrypsin inhibitor 2 (CI2),[42] *Bacillus caldolyticus* cold-shock protein (Bc CspB),[31] *Bacillus subtilis* cold-shock protein (Bs CspB),[31] human FK506 binding protein 12 (FKBP),[43] *Escherichia coli* histidine-containing phosphocarrier protein (Hpr),[32] *Bacillus stearothermophilus* N-terminal domain of phosphoglycerate kinase (residues 1–175) (N-PGK),[44] *B. stearothermophilus* N-terminal domain (residues 1–56) of L9 (NTL9),[45] Y43W point mutant of protein L (protein L),[46] *Streptomyces tendae* α-amylase inhibitor tendamistat (tendamistat),[47] and *Saccharomyces cerevisiae* iso-2 cytochrome *c* (y. isocyto-*c*).[48]

In order to compare the thermodynamic activation parameters, it was necessary to ensure that all raw unfolding and refolding rates were transformed using the same pre-exponential factor k' in equation (4), to calculate $\Delta G^{\ddagger}$ and $-T\Delta S^{\ddagger}$ (the absolute values of $\Delta H^{\ddagger}$ and $\Delta C^{\ddagger}_P$ are independent of k'). Although there is ongoing discussion about which value constitutes the most appropriate pre-exponential factor for protein folding transition states,[44] for the purposes of relative comparisons, as in our study, this choice is not critical. Since the majority of studies used k_BT/h for k', we used that as the standard pre-exponential factor for comparing the thermodynamics of various transition states. For those studies using an alternative k', the raw k_U or k_F was calculated from the reported $\Delta G^{\ddagger}_{exp}$, alternative k', and experimental temperature (T_{exp}) using equation (4). The new $\Delta G^{\ddagger}_{calc}$ for our comparison was then calculated using the extracted rate constant and k_BT/h, again using equation (4). By subtracting the reported $\Delta H^{\ddagger}$ from the new $\Delta G^{\ddagger}_{calc}$, $-T\Delta S^{\ddagger}_{calc}$ was calculated. For all proteins, with the $\Delta H^{\ddagger}$ and standardized $-T\Delta S^{\ddagger}$ known at T_{exp}, the reported $\Delta C^{\ddagger}_P$ was used in equations (7) and (8) to calculate $\Delta H^{\ddagger}$ and $-T\Delta S^{\ddagger}$ for unfolding and folding at temperatures ranging from 150–450 K. This wide range was used simply to help illustrate the full profiles and

assist in their comparison. Assumption of a constant $\Delta C_P^{\ddagger}$ is likely to be valid only in the experimentally relevant temperature range 0–70 °C.[5]

Acknowledgements

We thank B. Kelch for many helpful discussions. S.S.J. was supported by a Howard Hughes Medical Institute Predoctoral Fellowship while this work was carried out. S.M.E.T. was supported by a National Science Foundation Predoctoral Fellowship. Funding was provided by the Howard Hughes Medical Institute.

References

1. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.
2. Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, **29**, 7133–7155.
3. Privalov, P. L. & Gill, S. J. (1988). Stability of protein structure and hydrophobic interaction. *Advan. Protein Chem.* **39**, 191–234.
4. Makhatadze, G. I. & Privalov, P. L. (1995). Energetics of protein structure. *Advan. Protein Chem.* **47**, 307–425.
5. Robertson, A. D. & Murphy, K. P. (1997). Protein structure and the energetics of protein stability. *Chem. Rev.* **97**, 1251–1268.
6. Calloni, G., Taddei, N., Plaxco, K. W., Ramponi, G., Stefani, M. & Chiti, F. (2003). Comparison of the folding processes of distantly related proteins. Importance of hydrophobic content in folding. *J. Mol. Biol.* **330**, 577–591.
7. Makarov, D. E., Keller, C. A., Plaxco, K. W. & Metiu, H. (2002). How the folding rate constant of simple, single-domain proteins depends on the number of native contacts. *Proc. Natl Acad. Sci. USA*, **99**, 3535–3539.
8. Miller, E. J., Fischer, K. F. & Marqusee, S. (2002). Experimental evaluation of topological parameters determining protein-folding rates. *Proc. Natl Acad. Sci. USA*, **99**, 10359–10363.
9. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994.
10. Grantcharova, V., Alm, E. J., Baker, D. & Horwich, A. L. (2001). Mechanisms of protein folding. *Curr. Opin. Struct. Biol.* **11**, 70–82.
11. Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D. & Finkelstein, A. V. (2003). Contact order revisited: influence of protein size on the folding rate. *Protein Sci.* **12**, 2057–2062.
12. Sohl, J. L., Jaswal, S. S. & Agard, D. A. (1998). Unfolded conformations of alpha-lytic protease are more stable than its native state. *Nature*, **395**, 817–819.
13. Truhlar, S. M. E., Cunningham, E. L. & Agard, D. A. (2004). The folding landscape of *S. griseus* protease B reveals the energetic costs and benefits associated with evolving kinetic stability. *Protein Sci.* **13**, 381–390.
14. Zajicek, J. L., Carter, R. M. & Ghiron, C. A. (1981). A spectroscopic analysis of the thermally induced folding-unfolding transition of beta-trypsin. *Biophys. J.* **35**, 23–29.
15. Baker, D., Shiau, A. K. & Agard, D. A. (1993). The role of pro regions in protein folding. *Curr. Opin. Cell Biol.* **5**, 966–970.
16. Cunningham, E. L. & Agard, D. A. (2004). Disabling the folding catalyst is the last critical step in alpha-lytic protease folding. *Protein Sci.* **13**, 325–331.
17. Baardsnes, J., Sidhu, S., MacLeod, A., Elliott, J., Morden, D., Watson, J. & Borgford, T. (1998). *Streptomyces griseus* protease B: secretion correlates with the length of the propeptide. *J. Bacteriol.* **180**, 3241–3244.
18. Linderstrom-Lang, K. (1938). Peptide bonds in globular proteins. *Nature*, **142**, 996.
19. Linderstrom-Lang, K. (1950). Structure and enzymatic break-down of proteins. *Cold Spring Harbor Symp. Quant. Biol.* **14**, 117–126.
20. Ottesen, M. (1967). Induction of biological activity by limited proteolysis. *Annu. Rev. Biochem.* **36**, 55–76.
21. Anfinsen, C. B. & Scheraga, H. A. (1975). Experimental and theoretical aspects of protein folding. *Advan. Protein Chem.* **29**, 205–300.
22. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E. & Zambonin, M. (1997). Probing the partly folded states of proteins by limited proteolysis. *Fold. Des.* **2**, R17–R26.
23. Imoto, T., Yamada, H. & Ueda, T. (1986). Unfolding rates of globular proteins determined by kinetics of proteolysis. *J. Mol. Biol.* **190**, 647–649.
24. Jaswal, S. S., Sohl, J. L., Davis, J. H. & Agard, D. A. (2002). Energetic landscape of alpha-lytic protease optimizes longevity through kinetic stability. *Nature*, **415**, 343–346.
25. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
26. Murphy, K. P. & Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavior in proteins. *Advan. Protein Chem.* **43**, 313–361.
27. Spolar, R. S., Livingstone, J. R. & Record, M. T., Jr (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry*, **31**, 3947–3955.
28. Becktel, W. J. & Schellman, J. A. (1987). Protein stability curves. *Biopolymers*, **26**, 1859–1877.
29. Agashe, V. R., Shastry, M. C. & Udgaonkar, J. B. (1995). Initial hydrophobic collapse in the folding of barstar. *Nature*, **377**, 754–757.
30. Hollien, J. & Marqusee, S. (1999). A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochemistry*, **38**, 3831–3836.
31. Perl, D., Jacob, M., Bano, M., Stupak, M., Antalik, M. & Schmid, F. X. (2002). Thermodynamics of a diffusional protein folding reaction. *Biophys. Chem.* **96**, 173–190.
32. Van Nuland, N. A., Meijberg, W., Warner, J., Forge, V., Scheek, R. M., Robillard, G. T. & Dobson, C. M. (1998). Slow cooperative folding of a small globular protein HPr. *Biochemistry*, **37**, 622–637.
33. Akmal, A. & Munoz, V. (2004). The nature of the free energy barriers to two-state folding. *Proteins: Struct. Funct. Genet.* **57**, 142–152.
34. Derman, A. I. & Agard, D. A. (2000). Two energetically disparate folding pathways of alpha-lytic protease share a single transition state. *Nature Struct. Biol.* **7**, 394–397.

35. Kuhlman, B. & Baker, D. (2004). Exploring folding free energy landscapes using computational protein design. *Curr. Opin. Struct. Biol.* **14**, 89–95.
36. Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L. & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science*, **302**, 1364–1368.
37. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
38. Chen, X. & Matthews, C. R. (1994). Thermodynamic properties of the transition state for the rate-limiting step in the folding of the alpha subunit of tryptophan synthase. *Biochemistry*, **33**, 6356–6362.
39. Chiti, F., Taddei, N., van Nuland, N. A., Magherini, F., Stefani, M., Ramponi, G. & Dobson, C. M. (1998). Structural characterization of the transition state for folding of muscle acylphosphatase. *J. Mol. Biol.* **283**, 893–903.
40. Martinez, J. C., Viguera, A. R., Berisio, R., Wilmanns, M., Mateo, P. L., Filimonov, V. V. & Serrano, L. (1999). Thermodynamic analysis of alpha-spectrin SH3 and two of its circular permutants with different loop lengths: discerning the reasons for rapid folding in proteins. *Biochemistry*, **38**, 549–559.
41. Lorch, M., Mason, J. M., Sessions, R. B. & Clarke, A. R. (2000). Effects of mutations on the thermodynamics of a protein folding reaction: implications for the mechanism of formation of the intermediate and transition states. *Biochemistry*, **39**, 3480–3485.
42. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition state of folding. *Biochemistry*, **30**, 10436–10443.
43. Main, E. R., Fulton, K. F. & Jackson, S. E. (1999). Folding pathway of FKBP12 and characterisation of the transition state. *J. Mol. Biol.* **291**, 429–444.
44. Parker, M. J., Lorch, M., Sessions, R. B. & Clarke, A. R. (1998). Thermodynamic properties of transient intermediates and transition states in the folding of two contrasting protein structures. *Biochemistry*, **37**, 2538–2545.
45. Kuhlman, B., Luisi, D. L., Evans, P. A. & Raleigh, D. P. (1998). Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9. *J. Mol. Biol.* **284**, 1661–1670.
46. Scalley, M. L. & Baker, D. (1997). Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability. *Proc. Natl Acad. Sci. USA*, **94**, 10636–10640.
47. Schonbrunner, N., Pappenberger, G., Scharf, M., Engels, J. & Kiefhaber, T. (1997). Effect of preformed correct tertiary interactions on rapid two-state tendamistat folding: evidence for hairpins as initiation sites for beta-sheet formation. *Biochemistry*, **36**, 9057–9065.
48. Panda, M., Benavides-Garcia, M. G., Pierce, M. M. & Nall, B. T. (2000). Cytochrome c folds through a smooth funnel. *Protein Sci.* **9**, 536–543.
49. Glasstone, S., Laidler, K. & Eyring, H. (1941). *The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity, Diffusion, and Electrochemical Phenomena* (1st edit.), McGraw-Hill Book Company, New York.
50. Tan, Y. J., Oliveberg, M. & Fersht, A. R. (1996). Titration properties and thermodynamics of the transition state for folding: comparison of two-state and multi-state folding pathways. *J. Mol. Biol.* **264**, 377–389.

Edited by C. R. Matthews

(Received 14 September 2004; received in revised form 22 December 2004; accepted 13 January 2005)