Salt-induced detour through compact regions of the protein folding landscape

Daniel E. Otzen and Mikael Oliveberg*

Department of Biochemistry, Lund University, P.O. Box 124, 221 00 Lund, Sweden

Edited by Peter G. Wolynes, University of Illinois, Urbana, IL, and approved August 5, 1999 (received for review May 27, 1999)

In several cases, inorganic salts have been used to induce partly structured states in protein folding. But what is the nature of these states: Do they represent key stepping stones in the folding process, or are they circumstantial pitfalls in the energy landscape? Here we report that, in the case of the two-state protein S6, the salt-induced collapsed state is off the usual folding routes in the sense that it is prematurely collapsed and slows down folding by several orders of magnitude. Although this species is over-compact, it is not a dead-end trap but may fold by alternative channels to the native state.

2-state | intermediate | collapse | protein engineering

Certain inorganic salts such as phosphates and sulfates favor compact protein conformations because of preferential exclusion from the protein surface (1, 2). This makes them potent stabilizers of the native state and also provides a useful tool for characterizing partially structured states in the refolding process. These may otherwise go undetected even at low [GdmCl] but can be induced to accumulate transiently at sufficiently high salt concentrations (3, 4). A central issue has been whether such structures represent on-pathway or off-pathway intermediates, i.e., whether they are obligate stepping-stones in the conformational search. Kinetic data are often compatible with both scenarios (5). In the case of ubiquitin, fitting the folding rates and amplitudes to an off-pathway model leads to unrealistically fast folding rates whereas a conventional on-pathway model gives more plausible values (3). Provided the intermediate forms and decays slowly, pulse–chase competition experiments can distinguish between on- and off-pathway schemes, based on the assumption that an off-pathway species must unfold to reach the native state. Baldwin and coworkers have used this technique to infer that RNase A's intermediate is indeed on-pathway (6). A triangular folding pathway was recently proposed for hen egg-white lysozyme, in which the denatured state can fold to the native state either directly or somewhat slower via an intermediate. Two relaxation phases are resolved in the refolding reaction, and a comprehensive global analysis fits better to an on-pathway than an off-pathway model (7, 8). In triangular schemes, the distinction between on- and off-pathway is not clear-cut. However, in other cases, it is clear that the protein has to back-track from partly structured states, at least to some extent, to find the native conformation. Examples include cytochrome *c*, where the heme cofactor can misligate (9), bovine pancreatic trypsin inhibitor, where covalent side chain cross links can give rise to incorrect disulfide pairing (10, 11), and U1A, where intermolecular association takes place during refolding at protein concentrations >1 μM (12).

Here, we investigate a salt-induced collapsed state of the 101-residue ribosomal protein S6 from *Thermus thermophilus.* This collapsed species is off-pathway in the sense that it is over-compact. However, it need not always reverse back to the starting point but also may fold via an alternative and overall more compact route, albeit much more slowly.

Materials and Methods

S6 was produced by overexpression in *Escherichia coli* and was purified as described (13). S6 was unfolded in 4.4 M guanidinium chloride (GdmCl). Stopped-flow fluorescence experiments were performed on an Applied Photophysics (Surrey, U.K.) SX-18.MV apparatus in 50 mM Mes (pH 6.3) at 25°C with a final protein concentration of 0.5 μM, unless otherwise stated. Stopped-flow CD was performed by using the Aviv Associates (Lakewood, NJ) SF CD spectrometer 202SF with a bandwidth of 1.5 nm, a time constant of 10 ms, and a final protein concentration of 20 μM; experiments were performed at 7°C to slow the observed process and improve the signal.

Data Analysis. *Dead-end scheme (C⇔D⇔N) (Scheme 1).* If, as is common practice (5), we assume rapid equilibrium between the collapsed state C and the denatured state D, the following relationship holds:

$$k_{\mathrm{obs}} = f_{\mathrm{D}} k_{\mathrm{f}} = \frac{\mathrm{D}}{\mathrm{C} + \mathrm{D}} k_{\mathrm{f}} = \frac{k_{\mathrm{f}}}{1 + K_{\mathrm{C}}}, \qquad [1]$$

where k_{obs} is the measured rate constant, $K_{\mathrm{C}} = [\mathrm{C}]/[\mathrm{D}]$, and k_{f} is the refolding rate constant for the D→N step. This gives the following equation:

$$\log k_{\mathrm{obs}} = \log\left[\frac{10^{\log k_{\mathrm{f}}^{\mathrm{water}} + m_{\mathrm{f}} X}}{1 + 10^{\log K_{\mathrm{C}}^{\mathrm{water}} + m_{\mathrm{C}} X}}\right], \qquad [2]$$

where X is [GdmCl] and m_{C} and m_{f} are the linear dependencies of $\log K_{\mathrm{C}}$ and $\log k_{\mathrm{f}}$ on [GdmCl] (and in some stated cases on $[Na_2SO_4]$) (14). m_{f} and $\log k_{\mathrm{f}}$ are not linked to m_{C} and $\log K_{\mathrm{C}}$; the two pairs of parameters are determined independently by virtue of the long stretches of linearity, particularly in 1.0 M Na_2SO_4 (cf. Fig. 1*B*) above and below the "rollover" region.

Triangular scheme between D, C, and N (Scheme 2). Again assuming rapid equilibrium between D and C, we obtain

$$k_{\mathrm{obs}} = f_{\mathrm{D}} k_{\mathrm{f}} + f_{\mathrm{C}} k_2 = \frac{k_{\mathrm{f}}}{1 + K_{\mathrm{C}}} + \frac{k_2}{1 + \frac{1}{K_{\mathrm{C}}}}, \qquad [3]$$

where k_2 is the refolding rate constant for the C→N step, giving

$$\log k_{\mathrm{obs}} = \log\left[\frac{10^{\log k_{\mathrm{f}}^{\mathrm{water}} + m_{\mathrm{f}} X}}{1 + 10^{\log K_{\mathrm{C}}^{\mathrm{water}} + m_{\mathrm{C}} X}} + \frac{10^{\log k_2^{\mathrm{water}} + m_2 X}}{1 + 10^{-\log K_{\mathrm{C}}^{\mathrm{water}} - m_{\mathrm{C}} X}}\right], \qquad [4]$$

where m_2 is the linear dependency of $\log k_2$ on [GdmCl].

Φ-values for the collapsed state (Φ_{C}) are calculated according to (15):

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: C, collapsed state; D, denatured state; GdmCl, guanidinium chloride; N, native state; TS, transition state.

*To whom reprint requests should be addressed. E-mail: mikael.oliveberg@biokem.lu.se.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

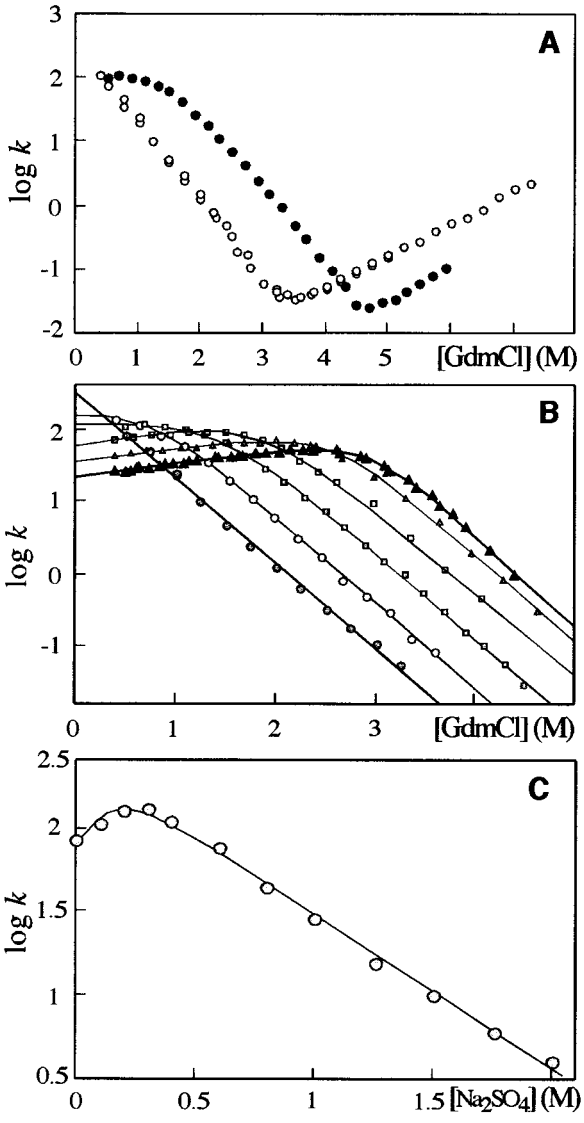


Fig. 1. (*A*) Chevron plots for S6 wild-type in the absence of Na_2SO_4 (○) and in 0.4 M Na_2SO_4 (●). (*B*) Refolding of 0.5 μM wild-type S6 from 4.4 M GdmCl into lower [GdmCl] at different [Na_2SO_4]: 0 M (◉), 0.2 M (○), 0.4 M (▣), 0.6 M (□), 0.8 M (△), and 1.0 M (▲). Note that the positive slope at low [GdmCl] becomes more pronounced as [Na_2SO_4] increases. The lines represent the best fits to Eq. **2**, in which [GdmCl], rather than [Na_2SO_4], is varied. (*C*) Refolding of wild-type S6 into 0.4 M GdmCl at different [Na_2SO_4]. The solid line represent the best fit to Eq. **2**. All rates are in units of s^{-1}.

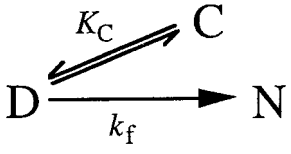
$$\Phi_C = \frac{\Delta\Delta G_C}{\Delta\Delta G_{D-N}} = \frac{-1.36RT(\log[K_C^{water,mutant}] - \log[K_C^{water,wild\text{-}type}])}{\Delta\Delta G_{D-N}}. \quad [5]$$

For each mutant, $\Delta\Delta G_{D-C}$ is obtained by using Eq. **3** for refolding data in 1 M Na_2SO_4 and 0.4–4.4 M GdmCl. $\Delta\Delta G_{D-N}$, the destabilization of N, is calculated from equilibrium denaturation data (16) measured in the absence of salt because of limitations in GdmCl solubility. Data on the stability of the transition state in the absence and presence of salt suggest that this is unlikely to affect the value of $\Delta\Delta G_{D-N}$ (see *Results and Discussion*).

Results and Discussion

Using Salt To Collapse the Denatured Ensemble. S6 has no co-factors, disulfide bridges or cis-prolines (17). Refolding is described by two phases: a fast major phase that corresponds to gain of native structure and a 10-fold slower minor phase that could be associated with folding-assisted cis-trans isomerization in the unfolded state (13) or refolding from the aggregated state, as observed for CI2 (18). The analysis in this study is based on the true refolding phase only, i.e., the major phase that describes two-state folding with a V-shaped chevron plot (Fig. 1*A*). It accounts for at least 90% of the total amplitude for all refolding experiments in the presence of Na_2SO_4. Although S6 has no detectable intermediates, it has been proposed that intermediates may still occur in the folding process but are too unstable to accumulate in the absence of stabilizing salts. To investigate this, we measured refolding rates as a function of both denaturant (GdmCl) and stabilizing salt (Na_2SO_4). S6 responds in an unusual manner to this treatment. At low [Na_2SO_4] (0.2 M), the chevron plot levels off at low [GdmCl] in a manner compatible with the accumulation of a folding intermediate I (D⇔I⇔N) (5). As I accumulates in the mixing time at low [GdmCl], the rate-limiting step becomes the conversion of I to N rather than the conversion of D to N via I. Above 0.4 M Na_2SO_4, however, the folding rate is augmented by low [GdmCl] (Fig. 1*B*). A similar effect has been observed previously for carbonic anhydrase in the absence of salt (19) and has been predicted independently from theory (30). From mass-action, a positive dependence of a given rate on [GdmCl] shows that more denaturant binds to the transition state than to the kinetic ground state C. This means that the transition state exposes more surface area than C, which accordingly is prematurely collapsed. Thus, C appears to be off the ideal route. The point is made even clearer when the refolding of S6 is measured at increasing [Na_2SO_4] and constant [GdmCl] (Fig. 1*C*): Although its refolding rate initially increases by 50% as [Na_2SO_4] increases to 0.3 M, it subsequently plummets to a value that amounts to 5% of the rate in the absence of salt. In a progressive folding scheme (D⇔I⇔N), the major transition state between I and N (TS) will be more stabilized by Na_2SO_4 than the less compact intermediate, leading to a monotonic increase in the refolding rate with increasing [Na_2SO_4]. This is in clear contrast to the experimental data (Fig. 1*C*). Note that it is necessary to go above 0.2 M Na_2SO_4 to reveal this over-compact character.

To a first approximation, the data are consistent with a minimalist dead-end scheme in which the over-compact species C equilibrates with the denatured state in the dead-time of the experiment (Eq. **2**)



Scheme 1.

This rapid equilibration is consistent with general views on the formation of compact species (5). The dead-end scheme also may be solved without the assumption of a rapid equilibrium between D and C, using simulated microscopic rate constants, k_1 and k_{-1}, for the interconversion of D and C. The characteristic equation gives two relaxation rates. The faster phase corresponds to the establishment of equilibrium within the dead-time of the experiment. The slower phase matches the rate constant derived from Eq. **2**, provided $k_1 > 10k_f$. The equation for Scheme 1 (Eq. **2**) is also based on the general view that both log K_C and

Table 1. Effect of Na_2SO_4 and GdmCl on the stability of C (K_C) and the refolding rate (k_f)

Parameters in 1 M Na_2SO_4 extrapolated to 0 M GdmCl*			
log K_C	m_C^{GdmCl}, M^{-1}	log k_f	m_f^{GdmCl}, M^{-1}
4.38 ± 0.06	−1.47 ± 0.02	5.71 ± 0.07	−1.30 ± 0.02
Parameters extrapolated to 0 M GdmCl and 0 M Na_2SO_4†			
log K_C	$m_C^{Na_2SO_4}$, M^{-1}	log kf	$m_f^{Na_2SO_4}$, M^{-1}
−0.02 ± 0.11	4.34 ± 0.14	2.56 ± 0.12	3.10 ± 0.17

*Data were obtained by fitting refolding data over 0.4–4.4 M GdmCl at fixed [Na_2SO_4] (1 M) to Eq. **1**. k_f was measured in s^{-1}.

†Values of log K_C and log k_f at six different concentrations of Na_2SO_4 were obtained from the data in Fig. 1*B* (Eq. **2**) and were fitted to a linear plot verseus [Na_2SO_4]. k_f was measured in s^{-1}.

log k_f depend linearly on [GdmCl] (with slopes m_C and m_f) as well as on [Na_2SO_4] (Table 1) (14). This is in line with the view that GdmCl and Na_2SO_4 both affect protein stability through weak interactions in which the cosolutes are either preferentially bound to or excluded from the protein surface (1). The values of m_C and m_f are significantly lower for GdmCl than for Na_2SO_4, probably because the guanidinium ion is larger than the sulfate ion and therefore has less access to the protein surface. Note that, in 0 M Na_2SO_4, log $K_C = -0.41 \pm 0.07$ in 0.4 M GdmCl (the lowest [GdmCl] under which refolding is measured). Thus, C becomes 29% populated during the dead-time of mixing even in the absence of salt. This is within the experimental error of two-state behavior (13).

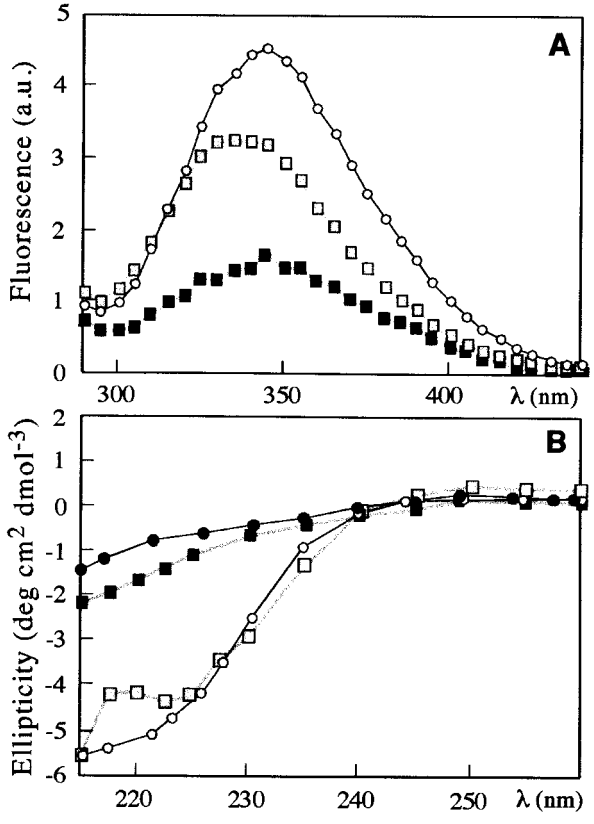


Fig. 2. (*A*) Fluorescence dead-time spectra for refolding of wild-type S6 into 0.4 M GdmCl at 0 M (■) and 1.0 M (▥) Na_2SO_4. The spectrum of the native state (○) is included for comparison. With increasing [Na_2SO_4], the fluorescence intensity of the dead-time species increases and becomes more blue-shifted, indicating a structure more compact than the unfolded state in the absence of Na_2SO_4. (*B*) Far-UV CD dead-time spectra for refolding into 0.9 M GdmCl at 0 M Na_2SO_4 (●) and 1.0 M Na_2SO_4 (▥). The equilibrium spectra of native S6 (○, 0.9 M GdmCl) and unfolded S6 (■), 4.4 M GdmCl) are included for comparison.

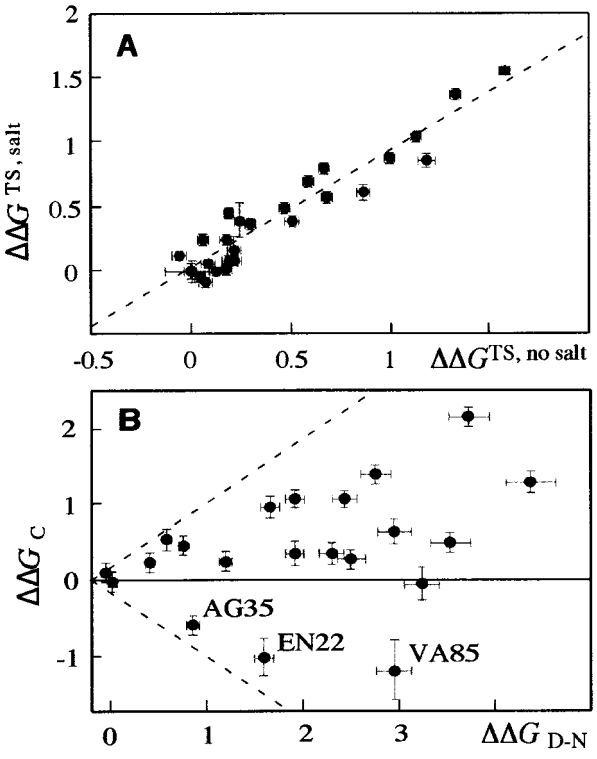


Fig. 3. (*A*) Effect of salt on the structure of the transition state of folding. $\Delta\Delta G^{TS}$, the mutational destabilization of the transition state for folding under given conditions, is calculated as $-1.36RT$ ($\log k_f^{mutant} - \log k_f^{wild\ type}$). Values in 3 M GdmCl and 1 M Na_2SO_4 are plotted versus values in 1 M GdmCl in the absence of salt. There is a clear linear correlation, with intercept 0.05 ± 0.04 and slope 0.91 ± 0.06, showing that the direct refolding step is the same in the absence and presence of Na_2SO_4. (*B*) Change in the free energy of formation of the collapsed state of S6 in 1 M GdmCl relative to the wild type ($\Delta\Delta G_C$) versus the change in the free energy of unfolding ($\Delta\Delta G_{D-N}$, from equilibrium denaturation experiments in the absence of salt). The dashed lines through Origo have slopes of 1 and −1, respectively, and are drawn purely for reference. There is no correlation between the two parameters, and there are three mutations that stabilize the collapsed state while destabilizing the native state. This suggests that the collapsed state contains specific but non-native interactions. All values are in units of kcal·mol^{-1}.

Although Scheme 1 turns out to be an over-simplification (see below), it can be used to determine the compactness of C and TS as follows. The *m*-value for each step is a rough indication of the change in solvent-exposure associated with the step (Table 1) (14). As $m_C = -1.47\ M^{-1}$ and $m_f = -1.30\ M^{-1}$, C is ≈10% more compact than TS. Because $m_{D-N} = 1.75\ M^{-1}$ (13), we conclude that C and TS have reached 85% and 75% of the native state's amount of surface area burial.

An alternative explanation for the decline in refolding rate at low [GdmCl] is aggregation, which is known to occur transiently in other systems (12). Although this cannot be entirely excluded for S6, it plays a minor role because a 40-fold increase in protein concentration (to 20 μM) has no significant effect on the folding kinetics (data not shown).

Another noticeable effect of high salt folding is that k_f turns out very fast ($0.15 \times 10^6\ s^{-1}$) when extrapolated to 0.4 M GdmCl, the lowest concentration at which refolding was measured. Scheme 1 implies that the formation of the species C must be even faster than this to allow the intermediate to accumulate; i.e., the D⇔C relaxation must be in the nanosecond–

Table 2. Data for refolding of S6 mutants in M Na_2SO_4

Mutant	log K_C^{water}*	m_C^{GdmCl}, M^{-1}	log $k_f^{3M\ GdmCl}$, s^{-1}*	m_f^{GdmCl}, M^{-1}	$\Delta\Delta G_{D-N}$, kcal·mol^{-1}†	Φ_C‡
Wild-type	4.37 ± 0.06	−1.47 ± 0.01	1.82 ± 0.02	−1.29 ± 0.02	≡ 0	
VA6	4.01 ± 0.08	−1.54 ± 0.02	0.80 ± 0.02	−1.44 ± 0.02	3.56 ± 0.16	0.14 ± 0.1
VA8	2.78 ± 0.07	−1.17 ± 0.02	0.69 ± 0.02	−1.18 ± 0.02	3.74 ± 0.17	0.59 ± 0.02
VA9	3.64 ± 0.10	−1.45 ± 0.02	1.64 ± 0.02	−1.27 ± 0.03	1.65 ± 0.08	0.61 ± 0.02
LA21	4.30 ± 0.13	−1.50 ± 0.03	1.76 ± 0.03	−1.29 ± 0.03	0.57 ± 0.03	0.19 ± 0.05
EQ22	5.11 ± 0.18§	−1.67 ± 0.07	1.82 ± 0.07‡	−1.34 ± 0.05	1.60 ± 0.08	−0.64 ± 0.03
IA26	3.33 ± 0.06	−1.40 ± 0.02	1.05 ± 0.02	−1.22 ± 0.02	2.76 ± 0.13	0.53 ± 0.01
LA30	4.38 ± 0.15	−1.58 ± 0.04	1.18 ± 0.04	−1.42 ± 0.04	3.26 ± 0.15	−0.00 ± 0.01
YA33	4.19 ± 0.06	−1.45 ± 0.01	1.77 ± 0.02	−1.26 ± 0.02	0.38 ± 0.02	0.67 ± 0.03
AG35	4.78 ± 0.07	−1.55 ± 0.02	1.83 ± 0.02	−1.39 ± 0.02	0.84 ± 0.04	−0.67 ± 0.02
VA37	4.14 ± 0.09	−1.52 ± 0.02	1.53 ± 0.02	−1.35 ± 0.02	2.50 ± 0.19	0.13 ± 0.01
VA40	3.97 ± 0.10	−1.45 ± 0.02	1.64 ± 0.03	−1.29 ± 0.03	0.59 ± 0.04	0.96 ± 0.04
IA48	4.04 ± 0.06	−1.43 ± 0.01	1.79 ± 0.02	−1.27 ± 0.02	0.77 ± 0.04	0.61 ± 0.02
FA60	4.19 ± 0.07	−1.51 ± 0.02	1.71 ± 0.02	−1.29 ± 0.02	1.21 ± 0.06	0.21 ± 0.01
VA65	3.89 ± 0.10	−1.52 ± 0.03	1.17 ± 0.02	−1.37 ± 0.03	2.95 ± 0.13	0.23 ± 0.01
MA67	4.11 ± 0.09	−1.49 ± 0.02	1.40 ± 0.02	−1.33 ± 0.02	2.30 ± 0.11	0.16 ± 0.01
LA75	4.11 ± 0.11	−1.55 ± 0.03	1.31 ± 0.03	−1.27 ± 0.03	1.92 ± 0.10	0.19 ± 0.01
LA79	3.41 ± 0.09	−1.32 ± 0.02	1.23 ± 0.02	−1.15 ± 0.03	4.39 ± 0.20	0.30 ± 0.01
VA85	5.24 ± 0.23§	−1.84 ± 0.09	1.49 ± 0.01‡	−1.13 ± 0.04	2.97 ± 0.15	−0.40 ± 0.01
VA88	3.57 ± 0.04	−1.33 ± 0.01	1.53 ± 0.10	−1.18 ± 0.01	1.93 ± 0.09	0.58 ± 0.02
VA90	3.57 ± 0.06	−1.37 ± 0.01	1.55 ± 0.02	−1.23 ± 0.02	2.44 ± 0.12	0.46 ± 0.01

*Data fitted to Eq. **2** except where indicated.
†Calculated from equilibrium denaturation data (16).
‡Calculated according to Eq. **5**.
§Data fitted to Eq. **4** (triangular model).

microsecond regime. Such collapse rates for the denatured ensemble are reported from laser-induced temperature jumps (20) and also are observed for reconfiguration of the denatured state on ligation of heme (21, 22). However, it is plausible that the extrapolated value of k_f is an overestimate because previous studies show that S6 is prone to transition state movements (13): The S6 transition state in 1.0 M Na_2SO_4 could well be shifted to an earlier part of the activation barrier at low [GdmCl] (23). In effect, this will cause a downward kink or curvature of the refolding limb of the chevron plot, lowering the folding rate in water (24).

Recently, trifluoroethanol and hexafluoroisopropanol also have been used to convert apparent two-state folding into multistate behavior (25). These alcohols are generally thought to stabilize secondary structure, and their effect on folding is attributed to enhanced levels of secondary structure in the

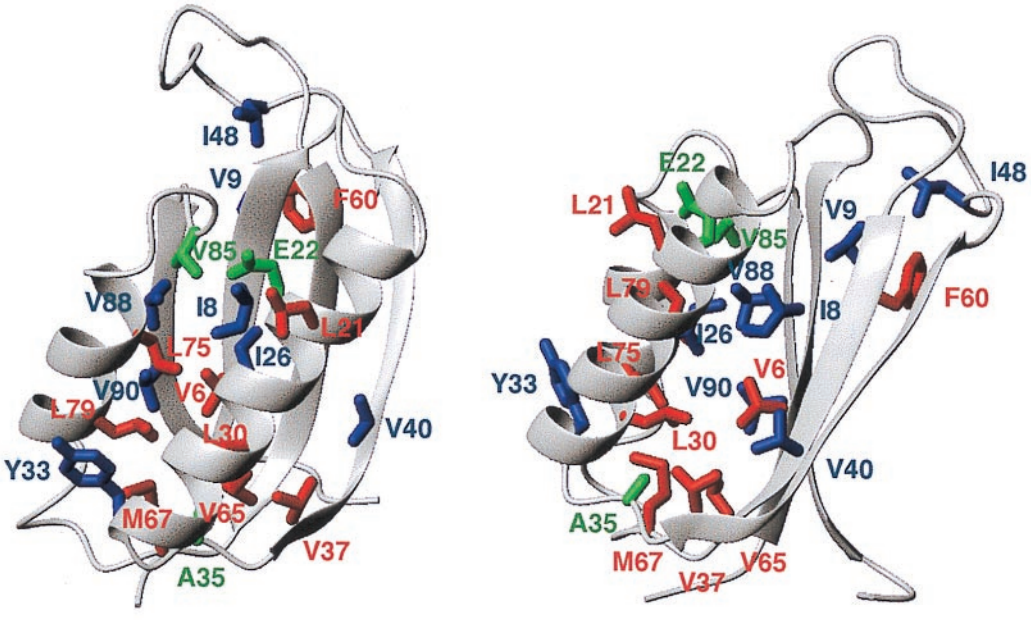


Fig. 4. Two views of S6 (17), showing the Φ_C-values of different mutated residues, calculated according to Eq. **3**. Residues 8,9, 26, 33, 40, 48, 88, and 90 (blue) have Φ-values around 0.6 whereas residues 6, 21, 30, 37, 60, 65, 67, 75, and 79 (red) have Φ-values below 0.2. Residues 22, 35, and 85 (green) have negative Φ-values (see text).

denatured ensemble. The kinetic effects of the alcohols is very similar to that of salt (compare Fig. 1*C* with Fig. 2*A* in ref. 25), although the maximum in k_f is suggested to coincide with native-like levels of secondary structure in the denatured ensemble. In contrast to Na_2SO_4, however, high alcohol concentrations unfold the protein by weakening the intramolecular hydrophobic interactions, so we do not *a priori* expect them to induce the same conformational states as salt. This is also supported by the lack of the characteristic rollover at low [GdmCl] when S6 is refolded in high concentrations of trifluoroethanol (data not shown). It may therefore be important to distinguish between effects on secondary structure alone and in conjunction with overall compaction.

Snapshots of a Prematurely Collapsed State. S6 contains a single Trp that we used to probe the level of structure in C. The spectrum of the dead-time species during refolding in 0.4 M GdmCl and 1.0 M Na_2SO_4 (which is 99.98% C) is intermediate between D and N in terms of intensity and is blue-shifted beyond N (Fig. 2*A*). This suggests a compact state with non-native contacts. The blue-shift, in particular, implies that Trp62 is less solvent-accessible in C than in N, where it is only partly buried. Reconstruction of the far-UV CD-spectrum under these conditions indicates a substantial accumulation of secondary structure that does not occur in the absence of salt (Fig. 2*B*). The GdmCl-dependent decrease in folding amplitude obtained by CD and fluorescence coincide (data not shown), which indicates that the collapse is cooperative. But, although high levels of secondary and tertiary structure are the spectroscopic hallmarks of a classical folding intermediate (26), the accumulation of C is kinetically disadvantageous in the case of S6.

Protein Engineering Analysis of the Structure of C. We have also characterized the structure of C by protein engineering using a large number of mutants. Judging from the refolding kinetics, the transition state structure in the direct route between D and N is not significantly affected by Na_2SO_4 because it is destabilized by mutations to the same extent in the absence and presence of 1 M Na_2SO_4 (Fig. 3*A*). Based on the Φ-value formalism (15), this transition state contains a rather diffuse nucleus with decreasing levels of structure at the edges; the highest Φ-value is ≈0.4[†] (data not shown). As concerns the Φ-value analysis of C, there is no clear correlation between the mutational destabilizations of C and N (Fig. 3*B*). With all caveats attributable to the intrinsic complexities in carrying out Φ-value analysis on non-native-like species, the mutants fall into three distinct classes. The first class has Φ_C-values around 0.6 whereas the second class has Φ_C-values below 0.2 (Table 2 and Fig. 4). Mutants in the third class (EQ22, AG35, and VA85) stabilize C despite destabilizing N and are discussed in more detail below.

The first two classes of mutations are not randomly distributed throughout the protein (Fig. 4). Although they do not show a conventional nucleation pattern, they still form contiguous patches; the highly structured region centers around strands 1 and 4 and the "hook region" between strands 2 and 3 whereas the unstructured region involves strands 2 and 3 and α-helix 1. If the residues in the structured region indeed are close in space, then C must be significantly different from the native state. According to the CD spectra (Fig. 2*B*), this could involve native-like secondary structure in a non-native tertiary context (cf. ref. 25). The high level of structure in C, which yields an average Φ_C-value of 0.33,[‡] contrasts with TS, which has an average Φ-value of 0.15 and a very delocalized nucleus (data not shown). Thus, C is more structured than TS, both in terms of side-chain interactions and overall compactness. The latter parameter is much higher than the average Φ-values, in line with the view that burial of hydrophobic surface area precedes orderly side chain arrangement (27).

Prefactor Effects. A relevant question concerns whether the salt-induced retardation of the folding rate could be explained simply by changes in the kinetic prefactor, i.e., an increase in the diffusional reconfiguration time as the polypeptide becomes more compact and experiences a higher degree of internal friction and steric conflicts (28). Although the build-up of native-like interactions in this case need not differ from that in the absence of salt, such a collapse would still be considered a compact detour because it restricts the protein to more compact regions of conformational space. However, the mutations' specific effects on the refolding kinetics suggest that the detour also involves ordered contact patterns as well as backbone configurations that are normally disfavored. Nevertheless, the reconfiguration time is expected to be closely coupled to the compactness of the polypeptide, and experiments to deconvolute these contributions are underway.

Alternative Folding Routes. The C-stabilizing mutants EQ22 and VA85 deviate from conventional dead-end behavior. At low [GdmCl], the positive slope flattens out, reaching an apparently constant value for the observed rate constant (Fig. 5*A*). The most plausible explanation for this is the opening of a direct route between C and D (Scheme 2 and Eq. **4**). The rate constant for this new passage (k_2) is 10.5 ± 0.6 for EQ22 and 7.4 ± 0.2 s^{-1} for VA85. Again, analysis including microscopic rate constants for the D⇔C relaxation matches Eq. **4**, provided $k_1 > 10\ k_f$. Interestingly, the effect also may be provoked in wild-type S6

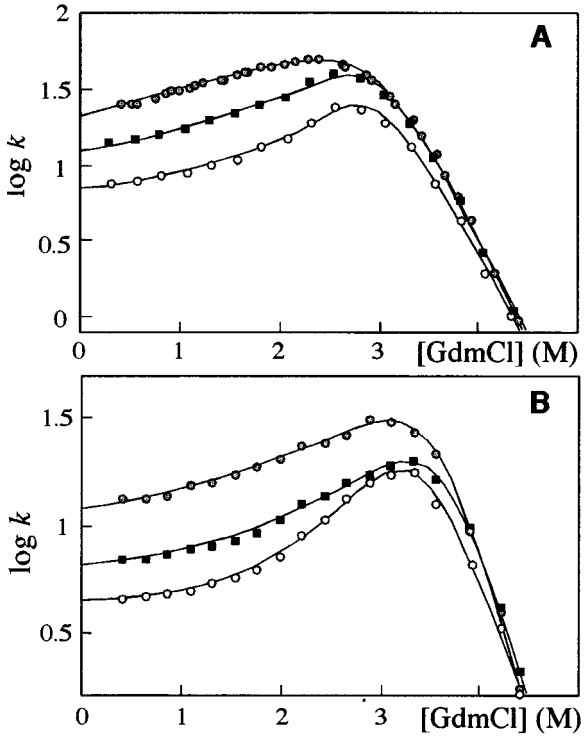
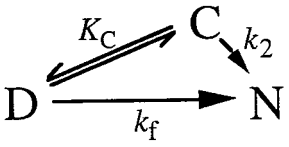


Fig. 5. Refolding of S6 wild-type (◉), VA85 (○), and EQ22 (■) in 1 M Na_2SO_4 (*A*) and 1.25 M Na_2SO_4 (*B*). At 1 M Na_2SO_4, wild-type data were fitted to the dead-end model (Scheme 1 and Eq. **2**). All other data were fitted to the triangular folding model (Scheme 2 and Eq. **4**). Rates are in units of s^{-1}.

[†]Φ = 0 indicates that the mutated side chain is just as unstructured as in the denatured state whereas Φ = 1 indicates native-like structure (15).

[‡]Excepting the three non-native mutations EQ22, AG35, and VA85 (see below).

by increasing [Na_2SO_4] above 1 M; in 1.25 M Na_2SO_4, wild-type and certainly the two mutants show a distinct flattening at low [GdmCl] (Fig. 5*B*).



Scheme 2.

The kinetic *m*-value for the C→N conversion (m_2) is ≈0, suggesting that the associated change in solvent exposure is very small (m_2 is -0.02 ± 0.08 for VA85, the mutant with the highest degree of stabilization of C). In view of the compact nature of C, this is not surprising. Note that this C → N conversion, which appears only to involve internal reconfiguration, could even represent a case of "glassy" downhill folding, cf. Sabelko *et al.* (29). AG35, the third mutation that stabilizes C, does not show flattening at low [GdmCl]. An explanation could be that AG35 does not stabilize C to the same degree as EQ22 and VA85 (Table 2). Of interest, E22, V85, and A35 have Φ-values close to zero in the transition state for the normal folding reaction (TS) and thus have no significant effect on the stability of this state (data not shown). Nor are they involved in stabilizing interactions in C because side-chain truncations increase the stability of C. Assuming that D is unstructured, it is more likely that these residues act as gatekeepers that steer the wild-type protein away from the "wrong" route by destabilizing C. The mutations EQ22 and VA85 may open up the "wrong" route by removing electrostatic or steric clashes whereas AG35 may favor C by increasing the flexibility of the backbone. Consistent with this idea, all three residues separate sequences of low and high Φ_C-values, suggesting that they obstruct the docking of substructures rearranging in a non-native fashion (Table 4). Because the premature collapse of the denatured ensemble profoundly decreases the folding rate (28–34), it seems reasonable that evolution selects gatekeeping residues that actively disfavor these rugged delays. Such a selection mechanism also explains the seemingly smooth folding funnels (28) of small proteins, where the folding behavior is well described by native interactions alone (35, 36).

In summary, there appear to be several folding routes available to S6. The stabilization of C by mutations or by excessive amounts of salt makes the swelling required for folding through the normal two-state route too costly. It is then better to continue folding through more compact conformations. Thus, all ways lead to Rome, also in the energy landscape; some are just very slow and arduous.

We thank Professor A. Fersht and Dr. C. Johnson for providing access to stopped-flow CD apparatus. This work was supported by the S. and E.-C. Hagbergs Foundation (The Royal Swedish Academy of Sciences), the Swedish Natural Science Research Council, and Kungliga Fysiografiska Sällskapet i Lund. D.E.O. gratefully acknowledges support from the Wenner-Grenska Foundation and the European Molecular Biology Organization.

1. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22,** 67–97.
2. Baldwin, R. L. (1996) *Biophys. J.* **71,** 2056–2063.
3. Khorasanizadeh, S., Peters, I. D. & Roder, H. (1996) *Nat. Struct. Biol.* **3,** 193–205.
4. Parker, M. J., Dempsey, C. E., Lorch, M. & Clarke, A. R. (1997) *Biochemistry* **36,** 13396–13405.
5. Baldwin, R. (1996) *Fold. Des.* **1,** R1–R8.
6. Laurents, D. V., Bruix, M., M., J. & Baldwin, R. L. (1998) *J. Mol. Biol.* **283,** 669–678.
7. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 9029–9033.
8. Wildegger, G. & Kiefhaber, T. (1997) *J. Mol. Biol.* **270,** 294–304.
9. Yeh, S.-R. & Rousseau, D. L. (1998) *Nat. Struct. Biol.* **5,** 222–228.
10. Creighton, T. E. (1975) *J. Mol. Biol.* **95,** 167–199.
11. Weissman, J. S. & Kim, P. S. (1991) *Science* **253,** 1386–1393.
12. Silow, M. & Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6084–6086.
13. Otzen, D. E., Kristensen, O., Proctor, M. & Oliveberg, M. (1999) *Biochemistry* **38,** 6499–6511.
14. Tanford, C. (1970) *Adv. Protein Chem.* **24,** 1–95.
15. Fersht, A. R., Matouschek, A. & Serrano, L. (1992) *J. Mol. Biol.* **224,** 771–782.
16. Clarke, J. & Fersht, A. R. (1993) *Biochemistry* **32,** 4322–4329.
17. Lindahl, M., Svensson, L. A., Liljas, A., Sedelnikova, S. E., Eliseikina, I. A., Fomenkova, N. P., Nevskaya, N., Nikonov, S. V., Garber, M. B., Muranova, T. A., *et al.* (1994) *EMBO J.* **13,** 1249–1254.
18. Silow, M., Tan, Y.-J., Fersht, A. R. & Oliveberg, M. (1999) *Biochemistry*, in press.
19. McCoy, L. F., Rowe, E. S. & Wong, K.-P. (1980) *Biochemistry* **19,** 4738–4743.
20. Gruebele, M., Sabelko, J., Ballew, R. & Ervin, J. (1998) *Acc. Chem. Res.* **31,** 699–707.
21. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11615–11617.
22. Parker, M. J., Lorch, M., Sessions, R. B. & Clarke, A. R. (1998) *Biochemistry* **37,** 2538–2545.
23. Silow, M. & Oliveberg, M. (1997) *Biochemistry* **36,** 7633–7637.
24. Oliveberg, M. (1998) *Acc. Chem. Res.* **31,** 765–772.
25. Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G. & Dobson, C. M. (1999) *Nat. Struct. Biol.* **6,** 380–387.
26. Ptitsyn, O. B. (1994) *Protein Eng.* **7,** 593–596.
27. Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995) *J. Mol. Biol.* **254,** 260–288.
28. Bryngelson, J., Onuchic, J. N., Socci, N. D. & Wolynes, P. (1995) *Proteins Struct. Funct. Genet.* **21,** 167–195.
29. Sabelko, T., Ervin, T. & Gruebele, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 6031–6036.
30. Camacho, C. J. & Thirumalai, D. (1996) *Protein Sci.* **5,** 1826–1832.
31. Sali, A., Shakhnovich, E. I. & Karplus, M. (1994) *Nature (London)* **369,** 248–251.
32. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995) *Biochemistry* **34,** 3066–3076.
33. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **4,** 561–602.
34. Bryngelson, J. D. & Wolynes, P. G. (1989) *J. Chem. Phys.* **93,** 6902–6915.
35. Shoemaker, B. A., Wang, J. & Wolynes, P. G. (1999) *J. Mol. Biol.* **287,** 657–674.
36. Shoemaker, B. A., Wang, J. & Wolynes, P. G. (1999) *J. Mol. Biol.* **287,** 675–694.