Nonlinear Free Energy Relationships in Arc Repressor Unfolding Imply the Existence of Unstable, Native-like Folding Intermediates†

Thorlakur Jonsson, Carey D. Waldburger, and Robert T. Sauer*

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307

Received December 28, 1995; Revised Manuscript Received February 9, 1996⊗

ABSTRACT: Under strongly denaturing conditions, the logarithm of the rate constant for dissociation/unfolding of the wild-type Arc dimer varies in a nonlinear fashion with denaturant concentration. To assess the unfolding/dissociation behavior under conditions favoring the native structure, we mixed Arc variants labeled with fluorescence acceptor or donor groups and used energy transfer to monitor the increase in heterodimer with time. Under the conditions of this experiment, the rate at which the heterodimer concentration approaches its equilibrium value is determined by the rate of dissociation and unfolding of the protein. Using this method and traditional denaturant-jump experiments, rate constants for unfolding/dissociation were determined over a wide range of stabilizing and destabilizing conditions. In each case examined, plots of $\log(k_u)$ versus denaturant showed significant curvature under strongly denaturing conditions, even though other kinetic experiments indicate that the unfolding/dissociation reactions remain largely two-state. This curvature can be explained most readily by a series of unstable intermediates in the unfolding pathway, with denaturant-induced changes in the kinetic step that is rate-limiting. Alternatively, curvature might result from Hammond behavior in which the structure of the transition state becomes more native-like as the stability of native Arc decreases with increasing denaturant.

The Arc repressor of bacteriophage P22 is a small, homodimeric protein with a roughly globular structure formed by packing of an antiparallel β-sheet against four α-helices (Breg et al., 1990; Bonvin et al., 1994; Raumann et al., 1994). In equilibrium and kinetic experiments, Arc unfolding appears to be a two-state process in which native dimers and denatured monomers are the only significantly populated species (Bowie & Sauer, 1989; Milla & Sauer, 1994). Refolding and association of denatured Arc monomers occur with a second-order rate constant of $\approx 10^7$ M^{-1} s^{-1} (Milla & Sauer, 1994) while refolding of a variant (Arc-MYL) in which a buried salt-bridge is replaced with hydrophobic interactions is about 50-fold faster (Waldburger et al., 1995; 1996). Unfolding rate constants (k_u) for Arc have typically been measured by standard methods involving jumps to a set of higher denaturant concentrations (Milla & Sauer, 1994; Schildbach et al., 1995). Assuming that $\Delta G^{\ddagger}$ varies in a linear fashion with denaturant (Tanford, 1970; Pace, 1975), the rate constant in the absence of denaturant can then be estimated by linear extrapolation of a $\log(k_u)$ versus denaturant plot. For wild-type Arc, the linear extrapolation method gives an unfolding rate constant reasonably close to the value expected based on the refolding rate constant and the equilibrium constant. By contrast, this method gives an unfolding rate constant for the MYL mutant that is too large by greater than 10-fold. This led us to question the validity of the linear extrapolation method for Arc variants and to reexamine the unfolding behavior of wild-type Arc.

Using fluorescence energy transfer, we devised a method that allows measurement of rate constants for the dissociation/unfolding of Arc variants under a wide range of denaturant concentrations, including strongly native conditions. Basically, cysteine-substituted homodimers labeled with fluorescence donor or acceptor groups are mixed, and the rate of heterodimer formation is monitored by energy transfer between the protein-bound dyes. Under the conditions of the experiment, the rate of heterodimer formation is determined by the rate of protein dissociation and unfolding. Similar experiments have recently been described for unfolding/dissociation of leucine-zipper proteins (Wendt et al., 1995).

We find that rate constants measured for unfolding/dissociation of Arc or Arc variants show the expected linear relationship between $\log(k_u)$ and denaturant only under conditions favoring the native state. By contrast, under strongly denaturing conditions, plots of $\log(k_u)$ versus denaturant show significant curvature. We discuss the physical basis for the observed nonlinear free energy relationships in terms of denaturant-induced changes in the rate-limiting step for Arc unfolding and/or transition-state structure.

MATERIALS AND METHODS

Mutagenesis and Protein Purification. The RC23 mutation was constructed by cassette mutagenesis in the *arc-st*11 gene of pSA700 (Milla et al., 1993). This gene encodes a C-terminal sequence, $(His)_6$-Lys-Asn-Gln-His-Glu (*st*11), which allows affinity purification and improves expression of Arc mutants *in vivo*. The MYL-RC23 variant was constructed in the same manner, beginning with a derivative of pSA700 encoding Met31, Tyr36, and Leu40 (Waldburger et al., 1995). Variants of wild-type Arc and the MYL mutant with *st*11 tails were also used. All proteins were purified by chromatography on Ni^{2+} NTA resin (Qiagen) and SP-Sephadex (Pharmacia) as described (Milla et al., 1993).

† Supported by NIH Grant AI-15706 and by an NIH/NRSA postdoctoral grant to C.D.W.

* To whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1996.

0006-2960/96/0435-4795$12.00/0

Fluorescent Labeling of RC23 and MYL/RC23. To reduce any oxidized forms of the RC23 or MYL-RC23 variants, the purified proteins were incubated overnight in 300 mM DTT, 50 mM Tris (pH 7.5), 50 mM KCl, and 5 M GuHCl at 37 °C, followed by extensive dialysis against buffer containing 50 mM Tris (pH 8.5), 50 mM KCl, and 1 mM DTT. The proteins were concentrated to approximately 100 μM using a Centriprep 3 device (Amicon), and GuHCl was added to a final concentration of 5 M (final pH ≈8.0). The fluorophores IAF [5-(iodoacetamido)fluorescein] or IAEDANS (5-[[2-[(iodoacetyl)amino]ethyl]amino]naphthalene-1-sulfonic acid) (Molecular Probes, Inc.) were added to approximately 2 mM concentration, followed by a 2−3 h incubation at room temperature. The reaction mixture was passed over a G-15 column (Pharmacia), and the fractions containing protein were pooled and dialyzed against buffer containing 50 mM Tris (pH 7.5), 250 mM KCl, and 0.1 mM EDTA. Labeling of the cysteine residues appeared to be complete, as determined by A_{280}/A_{340} ratios (protein/AEDANS), A_{280}/A_{494} ratios (protein/AF), and amino acid analysis. The circular dichroism intensity of the modified proteins at 222 nm was that expected based on the protein concentration, after correction for the contribution of the fluorophore. The fluorescently-labeled proteins were stored at 4 °C in the dark.

Kinetic Experiments. Stopped-flow kinetic experiments were performed using an Applied Photophysics Model DX-17MV instrument, with a 10 mm path length. Refolding experiments for wild-type Arc and fluorescently-labeled RC23 variants were performed by pH jumps from 10 mM phosphoric acid (pH 2.0), 50 mM KCl into buffer containing 60 mM Tris (pH 7.6), for a final pH of 7.5 (1:5 mixing). All other folding and unfolding experiments were carried out in buffer containing 50 mM Tris (pH 7.5), 50 mM KCl, and different concentrations of urea or GuHCl. The temperature in all experiments was 25 °C. Kinetic data were fit by nonlinear least squares methods using the instrument software or the program NONLIN (Johnson & Frasier, 1985; Brenstein, 1989). Energy-transfer experiments and unfolding experiments were fit as single exponentials. Refolding data were fit to a hyperbolic equation for homodimerization, as previously described (Milla & Sauer, 1994).

Kinetic Test of Two-State Folding/Unfolding. Stopped-flow unfolding and refolding experiments were performed at different initial urea concentrations but the same final urea concentrations and fit as described above. Final urea concentrations in these experiments were 0.5 M (wild-type refolding), 2 M (MYL refolding), 6 M (wild-type unfolding), and 8 M (MYL unfolding). If the reaction is two-state, then the observed rate constants should be independent of the initial urea concentration and the amplitudes should reflect the fraction of folded or unfolded protein (Tanford et al., 1973; Schindler et al., 1995).

Equilibrium Stability. Protein stability at equilibrium was determined by urea denaturation, monitoring the circular dichroism signal at 230 nm in a buffer containing 50 mM Tris (pH 7.5), 50 mM KCl, 25 °C, using an AVIV 60DS circular dichroism spectrometer. Fitting of the data was performed by a nonlinear least squares procedure using NONLIN and equations previously described (Milla et al., 1994).

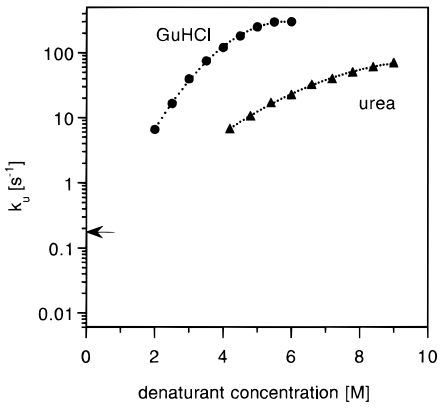

FIGURE 1: Dependence of the logarithm of the rate constant (k_u) for wild-type Arc unfolding/dissociation at 25 °C on denaturant concentration. Protein in 50 mM Tris (pH 7.5), 50 mM KCl was mixed with buffer containing urea or GuHCl to give the final denaturant concentration and a protein concentration of 1 μM. Rate constants were determined from exponential fits of the unfolding trajectories which were monitored by Trp fluorescence. Errors in the rates are smaller than or comparable to the symbol size. The arrow shows the zero-denaturant value of k_u calculated from K_u and k_f, the equilibrium denaturation constant and refolding rate constant, respectively. The dotted lines illustrate the curvature in the data but have no theoretical foundation.

RESULTS

Unfolding Rates Measured by Denaturant-Jumps. Arc unfolding at different urea or GuHCl concentrations was monitored by the change in fluorescence of the single tryptophan residue (Trp14) which is buried in the hydrophobic core of the native dimer (Milla & Sauer, 1994). The unfolding trajectories could be fit well as single exponentials with the expected amplitudes (Milla & Sauer, 1994). As shown in Figure 1, plots of log(k_u) versus denaturant concentration are clearly nonlinear. In fact, at the highest GuHCl concentrations, the unfolding reaction appears to become independent of denaturant. In each case, the lowest denaturant concentrations used were determined by the position of the denaturation transition zone, and the highest concentrations were determined by denaturant solubility. Thus, we could not test whether the unfolding rate also reaches a plateau at higher urea concentrations. Nevertheless, although GuHCl is a more potent denaturant than urea and causes greater curvature, the observation of significant nonlinearity in both the urea and GuHCl plots suggests that the effect is not denaturant-specific.

The existence of curvature in the log(k_u) versus denaturant plots makes extrapolation to lower denaturant values problematic. Previous studies of Arc unfolding using narrower ranges of urea (3−5 M; Milla & Sauer, 1994) and GuHCl (2.5−4 M; Schildbach et al., 1995) showed reasonable linear correlation coefficients ($r > 0.99$) for plots of log(k_u) vs denaturant. The lower half of each plot in Figure 1, which encompasses a similar range of denaturant, is also roughly linear ($r = 0.991-0.995$).

Unfolding Rates Measured by Energy Transfer. To examine unfolding rates in the absence of denaturant, Arc dimers were labeled with fluorescent acceptor or donor groups (see below) to allow measurement of the rate of heterodimer formation following mixing (Figure 2a). If the labeling reactions do not alter the kinetics of protein unfolding/dissociation or refolding/association, then the overall equilibrium between folded dimers and unfolded

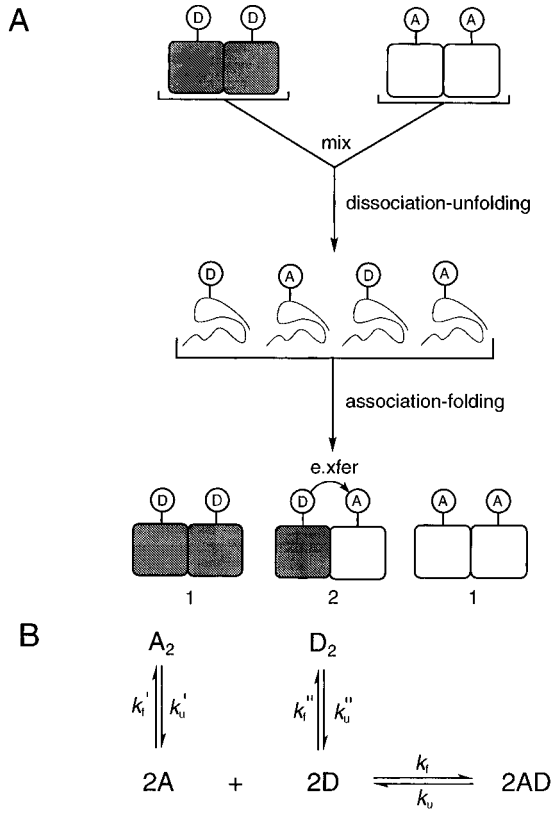

FIGURE 2: Scheme illustrating the procedure for measuring unfolding rates by energy transfer. (A) Arc dimers, labeled with fluorescent donors or acceptors, are mixed, and the formation of heterodimers is monitored by energy transfer between the donor and acceptor. (B) Rate constants determining the equilibration between homo- and heterodimers following mixing. If labeling with the fluorescent groups does not affect the kinetics of protein unfolding or refolding (i.e., $k = k' = k''$), then the overall equilibrium between native dimers and denatured monomers is not perturbed, and the mixture will approach equilibrium with a rate constant corresponding to k_u, the unfolding rate constant of the heterodimer.

monomers will not be perturbed (Figure 2B). Homodimers (D_2 or A_2) will initially be in dynamic equilibrium with a small population of monomer (D or A); following mixing, the equilibrium concentration of monomers should remain constant, while that of heterodimer (AD) will increase (as A_2 and D_2 decrease) with a rate given by

$$\frac{\partial[\mathrm{AD}]}{\partial t} = k_f[\mathrm{A}][\mathrm{D}] - k_u[\mathrm{AD}] \quad (1)$$

If the monomer concentrations remain constant during the experiment, however, the k_f[A][D] term will also be constant. In this case

$$[\mathrm{AD}] = [\mathrm{AD}]_{eq}(1 - e^{-k_u \cdot t}) \quad (2)$$

where $[AD]_{eq}$ is the concentration of heterodimer at equilibrium. Hence, the concentration of heterodimer should approach its equilibrium value in an exponential fashion with a rate determined by the rate of unfolding/dissociation.

Arg23 was chosen as the site for introduction of a cysteine for labeling with fluorescent dyes because substitution of alanine for this residue is not destabilizing and the β-carbons of positions 23 and 23′ are approximately 25 Å apart on opposite sides of the protein surface in the crystal structure of the wild-type Arc dimer (Milla et al., 1994; Raumann et al., 1994). The RC23 mutant of Arc was constructed and purified, and different batches were modified to generate the AF-labeled and AEDANS-labeled molecules needed for energy-transfer experiments. The free energies of unfolding/dissociation (ΔG_u) and the refolding rate constants (k_f) of these AF-modified and AEDANS-modified proteins were found to be within error of the wild-type protein (Table 1), indicating that the cysteine substitutions and modifications do not significantly alter equilibrium stability or the kinetics of refolding. By extension, the kinetics of unfolding should also be unaltered.

Table 1: Equilibrium Stability and Folding Kinetics for Wild-Type Arc and the Fluorescent-Labeled RC23 Variant[a]

	wild type	RC23/AF	RC23/AEDANS
ΔG_u (kcal/mol)	10.4	10.8	10.5
m_{eq} [(kcal/(mol·M)]	1.5	1.4	1.4
k_f (M^{-1} s^{-1})	0.8×10^7	1.0×10^7	1.2×10^7

[a] Experiments were performed in 50 mM Tris (pH 7.5), 50 mM KCl, 25 °C. Kinetic refolding experiments were performed by pH-jumps, as described under Materials and Methods, monitoring the change in Trp fluorescence. ΔG_u values were calculated for a standard state of 1 M. The errors in the values for ΔG_u are about 0.5 kcal/mol, while those for m_{eq} and k_f are about 10% and 20%, respectively.

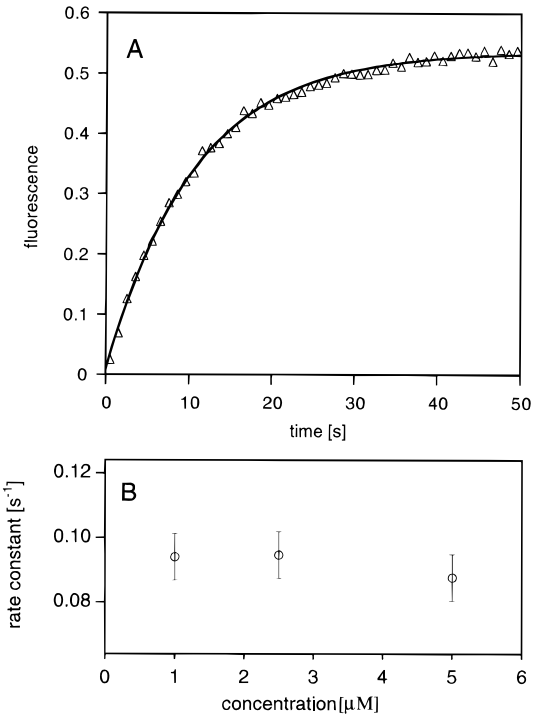

FIGURE 3: (A) Kinetics of formation of AF/AEDANS heterodimers of Arc-RC23 at 25 °C monitored by energy transfer. AF and AEDANS homodimers (1 μM) in 50 mM Tris (pH 7.5), 50 mM KCl were mixed at time 0. The concentration of heterodimer was monitored by the fluorescence emission (>530 nm), which arises from energy transfer, following excitation at 340 nm. The solid line is a single-exponential fit. (B) Rate constants determined for AF/AEDANS-R23C heterodimer formation from fits of experiments like that shown in panel A are independent of protein concentration, as expected if the rate is determined by the rate of unfolding/dissociation rather than the rate of refolding/association (see Figure 2B and text).

Figure 3A shows a stopped-flow experiment, monitored by energy transfer, following mixing of the AF-labeled and AEDANS-labeled RC23 proteins. The concentration of heterodimer approaches its equilibrium value in an exponential fashion with a rate constant of 0.095 ± 0.006 s^{-1}. Figure 3B shows that the rate of heterodimer formation is

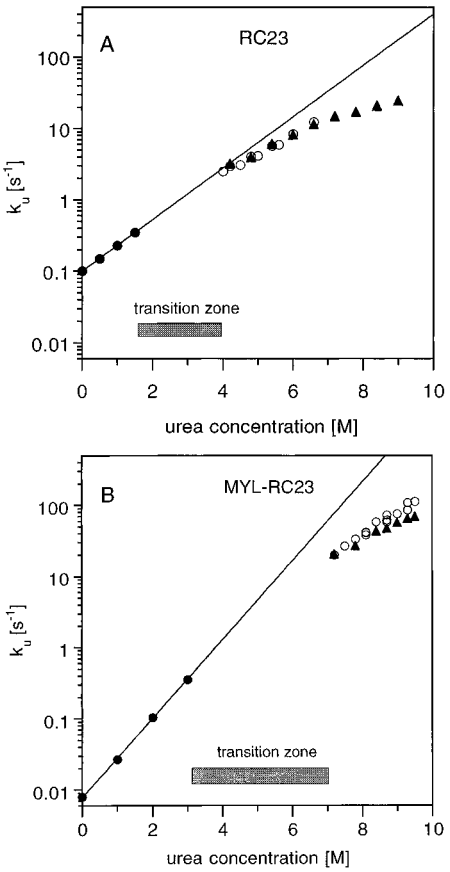

FIGURE 4: Urea dependence of unfolding rates for AF/AEDANS heterodimers of the RC23 (A) and MYL-RC23 (B) Arc variants at 50 mM Tris (pH 7.5), 50 mM KCl, 25 °C. Rates at urea concentrations below the transition zone (5−95% denaturation at 1 μM protein) were determined by energy transfer (closed circles; see Figures 2 and 3). Rates above the transition zone were determined by loss of energy transfer between AF/AEDANS heterodimers (open circles) or Trp fluorescence (closed triangles). Errors in rates are smaller than or comparable to the symbol size. The solid lines illustrate the linearity of the data at low urea concentrations ($r > 0.999$). Curvature at high urea concentrations causes significant deviations from these lines.

independent of protein concentration. This is an important control because it shows that refolding/association processes do not contribute to the observed rate of heterodimer formation as would be expected if eq 2 is valid.

Urea-Dependence of RC23 Unfolding. Figure 4A shows rate constants measured for unfolding/dissociation of the RC23 heterodimer at different concentrations of urea. In the low-urea regime, the rates were measured by the increase in heterodimer following mixing at the appropriate urea concentration. To measure unfolding rates at high [urea], a mixture of AF/AEDANS heterodimers and homodimers was allowed to form in the absence of urea, and unfolding of the heterodimer population was monitored by the loss of energy transfer between the protein-bound dyes following urea jumps. As a control, unfolding/dissociation rates for the entire mixture of heterodimer (≈50% of the sample) and homodimers (≈25% each) were determined by changes in the Trp fluorescence (closed triangles) and found to be within error of those determined by loss of energy transfer (open circles; Figure 4A). Unfolding rates determined for the individual AF-modified and AEDANS-modified homodimers were also within error of those determined for the mixture of heterodimer and homodimers (data not shown).

Table 2: Equilibrium Stability and Folding Kinetics for the MYL and Fluorescent-Labeled MYL-RC23 Variants of Arc[a]

	MYL	MYL-RC23/ AF	MYL-RC23/ AEDANS
ΔG_u (kcal/mol)	12.9	13.0	12.6
m_{eq} [kcal/(mol·M)]	1.2	1.1	1.0
k_f (M^{-1} s^{-1})	2.8×10^8	1.6×10^8	2.2×10^8

[a] Experiments were performed in 50 mM Tris (pH 7.5), 50 mM KCl, 25 °C. ΔG_u values were calculated for a standard state of 1 M. The errors in reported values for ΔG_u are about 0.5 kcal/mol, while those for m_{eq} and k_f are about 10% and 25%, respectively.

As shown in Figure 4A, the log(k_u) versus urea plot for the RC23 variant is quite linear ($r > 0.999$) below the transition zone but shows significant curvature (similar to that observed for wild-type Arc) above the transition zone. Because of this curvature, linear extrapolations using all of the high urea data would overestimate k_u in the absence of denaturant by a factor of 4-fold, while extrapolations using data between 4 and 6 M urea would overestimate this value by roughly 2-fold. We assume that estimates of the unfolding rate constant for wild-type Arc based on linear extrapolations would be in error by similar margins. An error of two-fold in the value of k_u for wild-type Arc would result in a 0.4 kcal/mol error in the value of ΔG_u calculated from the refolding and unfolding rate constants. Because experimentally determined values of ΔG_u for Arc have errors of approximately ±0.5 kcal/mol, a 2-fold error in k_u should not lead to a serious discrepancy between free energy changes calculated from kinetic and equilibrium experiments. Indeed, these values have been reported to be within error for wild-type Arc and most alanine-substitution mutants (Milla & Sauer, 1994; Milla et al., 1995). As discussed below, however, for certain Arc variants the linear extrapolation method can lead to much larger errors in both k_u and the value of ΔG_u calculated from kinetic experiments.

Unfolding/Dissociation of Arc Variants Bearing the MYL Mutation. In the MYL mutant of Arc, a partially buried salt bridge formed by the wild-type side chains of Arg31, Glu36, and Arg40 is replaced by hydrophobic interactions between Met31, Tyr36, and Leu40 (Waldburger et al., 1995). These substitutions increase the equilibrium stability of the MYL variant substantially, and thus higher concentrations of urea are required for unfolding jumps. In denaturant-jump experiments monitoring Trp fluorescence, we found that plots of log(k_u) vs denaturant for MYL were curved (7−9.5 M urea) and that the zero-denaturant value of k_u calculated by linear extrapolation using these data was greater than 10-fold larger than that calculated from the equilibrium constant and refolding rate constant (data not shown).

To allow measurement of unfolding over a much wider range of denaturant concentrations, we constructed and purified the MYL-RC23 protein, and prepared AF-modified and AEDANS-modified derivatives. As in the wild-type case, modification of MYL by a cysteine substitution at position 23 and subsequent labeling with AF and AEDANS did not affect stability or folding kinetics to any appreciable extent (Table 2). The urea dependence of MYL-RC23 unfolding is shown in Figure 4B. At concentrations of urea between 0 and 3 M, the log(k_u) vs urea plot is quite linear ($r = 0.9999$), but it begins to curve sharply above this range. As inspection of Figure 4B makes clear a linear extrapolation

based on the high-urea data would severely overestimate the real value of k_u in the absence of urea.

The unfolding rate constant (0.0093 ± 0.0009 s^{-1}) measured for the AF/AEDANS heterodimer of MYL-RC23 in the absence of urea is about 10-fold smaller than that observed for the RC23 heterodimer. As a result, at least part of the enhanced stability conferred by the MYL mutations results from a decrease in the unfolding rate. The remaining increase in stability is caused by an increase in the refolding rate from 1×10^7 to 2×10^8 M^{-1} s^{-1} (Tables 1 & 2). This increase in the refolding rate for MYL-RC23 compared to RC23 is similar to that reported for the MYL mutant compared to wild-type Arc (Waldburger et al., 1996).

We also note that the slope of the linear, low-urea portion of the log(k_u) vs urea plot for MYL-RC23 (1.254 ± 0.017) is significantly steeper than the slope of the corresponding region of the RC23 plot (0.834 ± 0.004). Division of these slopes by those of the corresponding log(K_u) vs urea slopes for each protein provides an estimate of the position of the transition state along a reaction coordinate defined by solvent accessibility (Tanford, 1970). By these calculations, the transition state for MYL-RC23 has about 30% as much buried surface as native MYL-RC23, while the transition state for RC23 has about 70% as much buried surface as native RC23. Thus, the transition state for MYL-RC23 is apparently less compact and seems to occur later in the unfolding pathway than the transition state for the RC23 protein. Comparison of the refolding kinetics of wild-type Arc and the MYL mutant has also led to the conclusion that the MYL transition state is less solvent-protected than the wild-type transition state (Waldburger et al., 1996). This, in turn, has been attributed to the existence of an unstable, dimeric intermediate in both the wild-type and MYL refolding reactions. In MYL, the transition state is postulated to involve formation of this intermediate dimer from denatured monomers. For wild type, the transition state is postulated to involve conversion of the intermediate dimer to the native dimer (Waldburger et al., 1996).

Kinetic Test of Apparent Two-State Unfolding/Refolding. The existence of intermediates in Arc unfolding/refolding has been proposed in several studies (Peng et al., 1993; Milla & Sauer, 1994; Burgering et al., 1995; Waldburger et al., 1996) and could explain the nonlinear free energy relationships observed here (see Discussion). However, the equilibrium and kinetic unfolding/refolding reactions of Arc behave as two-state transitions between denatured monomers and native dimers, both in kinetic and in equilibrium experiments (Bowie & Sauer, 1989; Milla & Sauer, 1994). The existence of intermediates and the apparent two-state nature of the unfolding/refolding reactions can be reconciled if the intermediates are unstable and thus are not significantly populated with respect to denatured monomers and/or native dimers under the experimental conditions used.

Using the method proposed originally by Tanford et al. (1973) and advanced more recently by Schindler et al. (1995), we measured unfolding and folding rates for wild-type Arc and MYL using the same final urea concentrations but different initial urea concentrations. For both proteins, the refolding and unfolding trajectories were monophasic, and the rates were independent of the initial urea concentration, as expected if any intermediates present are not significantly populated. Figure 5A shows these data for wild-type Arc. Moreover, plotting the amplitudes of these unfolding and refolding transitions produces denaturation curves (Figure 5B) which can be fitted by ΔG_u values of of 10.8 kcal/mol (wild type) and 13.1 kcal/mol (MYL). These values are within error of the values determined for these proteins by equilibrium measurements (10.4 ± 0.5 kcal/mol and 12.9 ± 0.5 kcal/mol, respectively). This coincidence of kinetic and equilibrium values is expected for reactions in which denatured monomers and native dimers are the major populated species at all concentrations of denaturant since the amplitudes will then be a measure of the amount of folded or unfolded protein (Schindler et al., 1995).

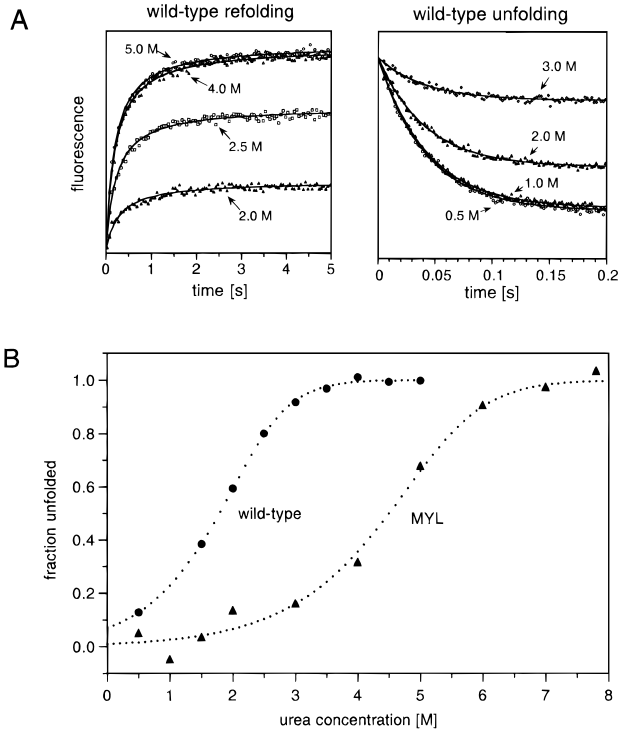

FIGURE 5: Urea denaturation of wild-type Arc and the MYL variant determined by kinetic refolding and unfolding experiments [50 mM Tris (pH 7.5), 50 mM KCl, 1 μM final protein, 25 °C]. (A) Refolding (left panel) and unfolding (right panel) experiments for wild-type Arc were performed starting at different initial urea concentrations (indicated for each trace) and ending at the same final urea concentration (0.5 M for folding; 6.0 M for unfolding). Data from different experiments were adjusted by adding or subtracting an offset value so that each trajectory begins at the same fluorescence value. The solid lines represent hyperbolic fits of refolding trajectories ($k_{app} = k_f P_t = 4.7 \pm 0.2$ s^{-1} for each experiment) and exponential fits of unfolding trajectories ($k_u = 23.9 \pm 1.2$ s^{-1} for each experiment). (B) Denaturation curves for wild-type Arc and the MYL variant are plotted using the amplitudes from experiments like those shown in panel A as a measure of the fraction of unfolded protein. Refolding experiments were performed at initial urea concentrations of 1.5−5.0 M (wild type) or 4.0−8.0 M (MYL), while unfolding experiments were performed at initial urea concentrations of 0.5−3.5 M (wild type) and 0.5−7.0 M (MYL). At urea concentrations where multiple determinations were made, the average values of fraction unfolded protein are shown. The dotted lines are calculated denaturation curves for wild type [$\Delta G_u = 10.8$ kcal/mol; $m_{eq} = 1.5$ kcal/(mol·M)] and MYL [$\Delta G_u = 13.1$ kcal/mol; $m_{eq} = 1.1$ kcal/(mol·M)].

The RC23 and MYL-RC23 unfolding and refolding reactions also appear two-state by the criteria of monophasic unfolding and refolding trajectories, and good correspondence between ΔG_u values calculated from rate constants (using k_u from energy-transfer experiments) and those determined in equilibrium experiments (Table 3). The agreement between ΔG_u values determined from kinetic and

Table 3: Comparison of Thermodynamic Parameters for the RC23 and MYL-RC23 Variants of Arc Measured in Equilibrium Experiments and Calculated from Kinetics[a]

	RC23	MYL-RC23
ΔG_{eq} (kcal/mol)	10.6 ± 0.5	12.8 ± 0.5
m_{eq} [kcal/(mol·M)]	1.4 ± 0.1	1.1 ± 0.1
ΔG_{kin} (kcal/mol)	10.5 ± 0.2	13.8 ± 0.5
m_{kin} [kcal/(mol·M)]	1.5 ± 0.1	1.1 ± 0.1

[a] Reported values are for unfolding in 50 mM Tris (pH 7.5), 50 mM KCl, 25 °C, as determined by equilibrium denaturation and calculated from the urea dependence of folding and unfolding kinetic parameters. Values for ΔG_{eq} and m_{eq} are the average of the corresponding values for AF- and AEDANS-modified proteins (Tables 1 and 2), while values for ΔG_{kin} and m_{kin} are based on k_f and k_u values determined in buffer containing no urea and on the urea dependence of these rate constants under conditions where ln(k) vs urea plots are linear (0−2 M urea for wild-type, 0−3 M for MYL).

equilibrium experiments is expected for a two-state reaction or for a multi-state reaction in which the apparent rate constants (k_u and k_f) contain the rate constants for all of the elementary reactions in the overall unfolding/refolding reaction. This latter situation will occur if there are rate-determining preequilibria in either the unfolding and/or the refolding reactions.

DISCUSSION

Measuring Unfolding Rates under Physiological Conditions. Using fluorescence energy transfer, the rates of subunit dissociation and unfolding (k_u) of variants of the Arc repressor dimer have been measured in the absence of denaturants. The method requires different batches of protein to be labeled with a fluorescence donor and a fluorescence acceptor group, which is most easily accomplished if the protein contains a unique and accessible cysteine residue. Since this was not the case for wild-type Arc, the RC23 mutant background was used for the studies reported here. As we have shown, however, neither the RC23 mutation nor modification of this cysteine with the AF or AEDANS groups causes significant changes in stability or folding kinetics in the wild-type Arc background or in MYL, a hyperstable Arc variant in which a buried salt-bridge is replaced by hydrophobic interactions. The half-lives for unfolding were found to be 7 s for the RC23 protein and 74 s for the MYL-RC23 protein (50 mM KCl, pH 7.5, 25 °C). Both proteins maintain stably folded structures, despite these relatively rapid unfolding rates, because refolding is very fast. At concentrations of 10 μM, the MYL-RC23 protein refolds with a half-time of 0.3 ms, and the RC23 protein refolds with a half-time of 10 ms [50 mM Tris (pH 7.5), 50 mM KCl, 25 °C]. Thus, the increased equilibrium stability conferred by the MYL mutations comes both from slower unfolding and from faster refolding. Detailed studies of the effects of the MYL mutations on the refolding of otherwise wild-type Arc have been reported elsewhere (Waldburger et al., 1996).

Curvature in log(k_u) vs Denaturant. Plots of log(k_u) vs denaturant for the RC23 and MYL-RC23 variants of Arc are quite linear at low urea concentrations but become significantly curved at high concentrations (Figure 4). Traditional methods for the determination of the zero-denaturant value of k_u employ linear extrapolation based on log(k_u)values determined at denaturant concentrations above the transition zone. Based on data obtained immediately above the transition zone, this procedure leads to 2−3-fold errors in the zero-denaturant value of k_u for the RC23 protein but greater than 10-fold errors for the MYL-RC23 protein. In both cases, the error increases when rate constants determined at very high denaturant are included in the linear extrapolation. Curvature in log(k_u) vs denaturant plots is not unique to the Arc system. For example, similar but less dramatic nonlinear behavior has also been reported for mutants of barnase (Matouschek & Fersht, 1993; Matouschek et al., 1994).

Models for Nonlinear Free Energy Relationships. The reaction-coordinate diagrams shown in Figure 6A,B depict two models that can explain curvature in log(k_u) vs. denaturant plots for Arc and its variants. The first is a two-state denaturation model, in which the position of the transition state changes with denaturant along a reaction coordinate defined by solvent accessibility (Figure 6A). The second is a multistate model in which the relative heights of kinetic barriers between intermediate states change as a function of urea and thus different barriers become rate-limiting at different denaturant concentrations. Why do the models shown in Figure 6A,B predict curvature in plots of log(k_u) vs denaturant? The free energy difference between two states (ΔG(D)) is expected to change with denaturant (D) in a fashion that depends on the solvent accessibilties of state 1 (A_1) and state 2 (A_2) (Schellman, 1977):

$$\Delta G(\mathrm{D}) = \Delta G^\circ + c(A_1 - A_2)[\mathrm{D}] = \Delta G^\circ + m_{eq}[\mathrm{D}] \quad (3)$$

In this equation, c is a constant that includes the free energy change per angstrom squared of solvent-accessible surface. Assuming that the transition-state description of rate constants is valid for protein folding reactions, a relationship for the dependence of the unfolding rate constant on denaturant can be formulated:

$$\Delta G_u^{\ddagger} = -2.3RT \log(k_u) + 2.3RT \log(k_B T/h) \quad (4)$$

$$\frac{\partial \log(k_u)}{\partial[\mathrm{D}]} = -c(A_n - A_{ts})[\mathrm{D}]/2.3RT = -m_u^{\ddagger}[\mathrm{D}]/2.3RT \quad (5)$$

where k_B and h represent Boltzmann's and Planck's constants, respectively. Thus, the free energy difference between the native state and transition state will vary linearly with denaturant if, and only if, ($A_n - A_{ts}$) remains constant at different concentrations of denaturant. If ($A_n - A_{ts}$) changes with denaturant, however, then the slope of log(k_u) vs denaturant plots will change, giving rise to curved plots. Hence, the nonlinear behavior observed for Arc and its variants can be explained by denaturant-induced changes in the structure and thus the solvent accessibility of the transition state (the accessibility of the native state could also change, but this seems less likely).

In the models of Figure 6A, the transition-state position varies with urea along the reaction coordinate. In the MYL-RC23 case, for example, the transition state is positioned close to the denatured state in the absence of urea (≈0.2) but close to the native state in 8 M urea (≈0.8). In the RC23 case, the transition state moves from a position at ≈0.6 in the absence of urea to one at ≈0.9 in 8 M urea. Thus, for both proteins, the solvent accessibility of the transition state becomes more native-like as the urea concentration increases and the native state becomes less stable. This behavior is

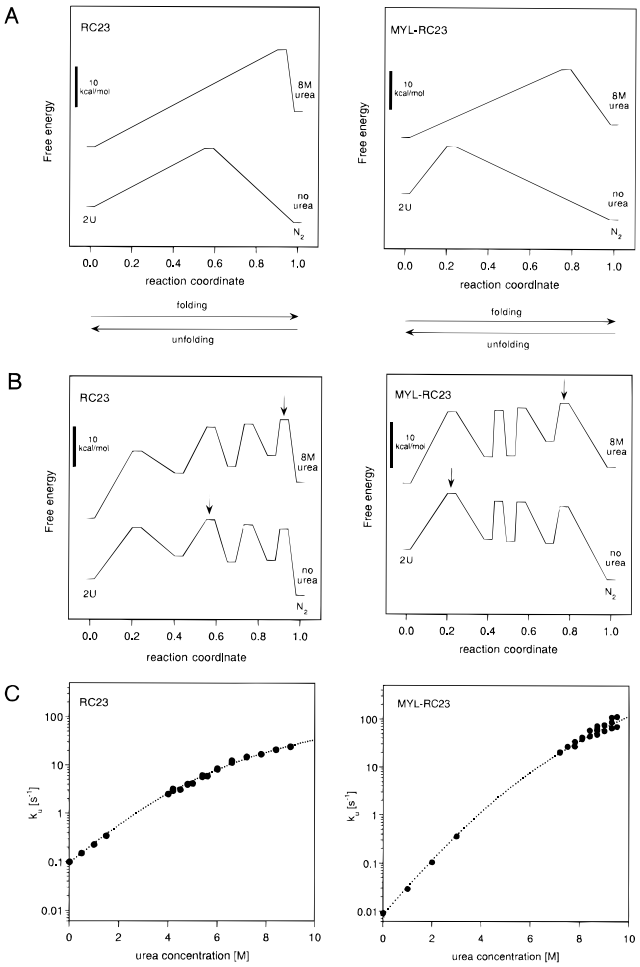

FIGURE 6: Panels A and B show denaturation models that may explain curvature in log(k_u) vs urea plots for the RC23 and MYL-RC23 variants of Arc. In each case, the position of the unfolded state on the free energy axis is arbitrary. The position of each state on the reaction coordinate axis is defined as $(A_x - A_u)/(A_n - A_u)$ where A_x is the solvent-accessible surface area of any state, and A_n and A_u are the solvent-accessible surface areas of the native and denatured states, respectively. Free energy profiles for no urea and 8 M urea are shown in each case. (A) Two-state models for RC23 and MYL-RC23 denaturation in which the position of the transition state along the reaction coordinate changes with urea. Hammond behavior results in a transition state that is more native-like as the stability of the native state decreases with increasing urea. (B) Multistate models for RC23 and MYL-RC23 denaturation in which the heights of different kinetic barriers change as a function of urea and thus the position of the rate-determining step along the reaction coordinate (marked by arrow) changes with urea. The number of equilibrium and associated transition states, the free energies of these states, and the positions of these states along the reaction coordinate were adjusted by trial-and-error to obtain a satisfactory fit of the experimental data, as shown in panel C. Experimental values of m_u/m_{eq} were used to set the range of possible transition state positions at each urea concentration. (C) The dotted lines show the urea dependence of log(k_u) calculated from the denaturation models shown in panel B. Values for each microscopic rate constant were calculated based on the height of each free energy barrier and used to calculate the net rate constant for the unfolding transition. This procedure was then repeated for different urea concentrations, after adjusting the barrier and free energy heights based on *m*-values. The closed circles are the experimental data from Figure 4A,B.

consistent with the Hammond postulate (1955). In our view, however, there are several unattractive features of this model. First, the Hammond postulate suggests that the properties of the transition state should be most similar to the least stable equilibrium state, which is not true for RC23 in the absence of urea. Second, ΔG_u changes continuously with urea, and thus the accessibility of the transition state might also be expected to change continuously. This, however, would give rise to plots of log(k_u) vs urea that are smoothly curved, which is not observed. Third, a transition state whose physical properties can change as dramatically as would be necessary in the MYL-RC23 case is difficult to conceptualize in structural terms.

In the models of Figure 6B, the positions of the kinetic barriers along the reaction coordinate do not change with urea. Rather, the free energy difference between kinetic barriers changes with urea, so that different barriers are rate limiting at different urea concentrations. The urea-dependent change in free energy between two barriers can be calculated as

$$\frac{A_{ts2} - A_{ts1}}{A_n - A_d} m_{eq}[D] \qquad (6)$$

where m_{eq} is approximately 1.5 kcal/(mol·M) for wild-type Arc and RC23. Thus, if two barriers were separated by 0.5 on the reaction coordinate, then the height of the more native-like barrier would increase by 6 kcal/mol relative to the other barrier between 0 and 8 M urea. If the less native-like barrier were 4 kcal/mol higher in the absence of urea, this would lead to inversion in 8 M urea. By contrast, if the earlier barrier were 10 kcal/mol higher in the absence of urea, it would still be 4 kcal/mol higher in 8 M urea. By adjusting the relative heights and positions of different transition-state barriers, log(k_u) vs denaturant plots with both linear and curved regions can easily be generated.

The heights and positions of the equilibrium states and kinetic barriers shown in Figure 6B were determined by trial-and-error to fit the unfolding data for RC23 or MYL-RC23 (see Figure 6 legend). The quality of these fits is very good, as shown in Figure 6C. For both MYL-RC23 and RC23, there are three equilibrium intermediates in the denaturation models. Only two intermediates (near 0.7 and 0.9) were strictly needed to fit the RC23 unfolding data, but three were included for consistency between the two proteins, and because the presence of the intermediate near 0.4 is also consistent with refolding data for wild-type Arc (Waldburger et al., 1996). The free energies of the intermediate states shown in Figure 6B are arbitrary and can be changed without altering the fit of the kinetic data. As long as these states are kinetically unstable with respect to the native or the denatured state, then the observed unfolding and refolding kinetics will appear two-state. Thus, although the models shown in Figure 6B are probably not unique, they are physically reasonable, and explain both the observed curvature in plots of log(k_u) vs denaturant and the apparent two-state nature of the folding and unfolding reactions. Models of this type are also consistent with the observation that some alanine-substitution mutants of Arc show significant changes in the urea dependence of unfolding (Milla et al., 1995), since a mutation could easily alter the height or accessibility of a specific kinetic barrier and thus alter which steps become rate limiting at different urea concentrations.

One way in which the models of Figure 6A,B differ is whether the free energy profile is smooth or crenelated by the existence of unstable intermediates. We prefer the Figure 6B model because it makes physical sense that the polypeptide should pass through intermediate states during folding and unfolding. As noted above, the Figure 6B model

also appears better able to account both for linear and for curved regions of log(k_u) vs denaturant plots. We also considered but rejected models in which the free energy varies in an inherently nonlinear fashion with denaturant (Tanford, 1964, 1970; Aune & Tanford, 1969). Two models of this type, the denaturant binding model and the peptide free energy transfer model, both predict that the greatest curvature in log(k_u) vs. urea plots should occur at low denaturant concentrations and were unable to fit the RC23 and MYL-RC23 unfolding data in a satisfactory fashion (not shown).

Unfolding/Refolding Intermediates. We have recently proposed the existence of at least one intermediate in Arc refolding, based on the urea and viscosity dependence of refolding kinetics (Waldburger et al., 1996). The results presented here suggest the existence of several additional intermediates, at least one of which appears to be quite native-like in terms of solvent accessibility. In fact, the observation that the unfolding rate for wild-type Arc becomes nearly independent of denaturant at high GuHCl concentrations (Figure 1) suggests that a transition state with a solvent accessibility very close to that of the native state has become rate-limiting. The microscopic step associated with this barrier may involve processes such as the breaking of hydrogen bonds which do not involve significant changes in solvent accessibility. The view of relatively compact, native-like intermediates present late on the folding pathway of Arc is consistent with recent model studies of protein folding (Abkevich et al., 1994; Sali et al., 1994; Onuchic et al., 1995; Betancourt & Onuchic, 1995; Wolynes et al., 1995). These models, which are based on folding simulations on lattices, suggest that the transition state for folding consists of a large ensemble of partially folded structures. Kinetic intermediates appear later in folding, with most having a large fraction of the correct native contacts (Wolynes et al., 1995).

It is interesting to note that the model of Figure 6B predicts that the refolding reaction for MYL-RC23 follows a downhill free energy landscape in the absence of urea, with each successive barrier being of lower energy than the previous barrier. This may explain the efficient, submillisecond refolding kinetics of the MYL-RC23 protein. In several recent papers, the observation that certain small proteins show submillisecond refolding kinetics in an apparent two-state fashion has been interpreted as indicating that these refolding reactions occur without intermediates (Schindler et al., 1995; Fersht, 1995). This would be analogous to the two-state models shown in Figure 6A. Our results, however, are more consistent with the view that even in cases of very rapid, apparent two-state refolding, the polypeptide does pass through a series of intermediate states which are unstable and thus poorly populated. Unstable kinetic intermediates could play an important role in protein folding, acting to restrict the search of conformational space needed to locate subsequent states on the folding pathway.

ACKNOWLEDGMENT

We thank Phillip Bryan, Marcos Milla, Dennis Rentzeperis, Joel Schildbach, and Frank Solomon for advice and helpful discussions.

REFERENCES

Abkevich, V. I., Gutin, A. M., & Shaknovich, E. I. (1994) *Biochemistry 33,* 10026−10036.
Aune, K. C., & Tanford, C. (1969) *Biochemistry 8,* 4586−4590.
Betancourt, M. R., & Onuchic, J. N. (1995) *J. Chem. Phys. 103,* 773−787.
Bonvin, A. M., Vis, H., Breg, J. N., Burgering, M. J., Boelens, R., & Kaptein, R. (1994) *J. Mol. Biol. 236,* 328−341.
Bowie, J. U., & Sauer, R. T. (1989) *Biochemistry 28,* 7139−7143.
Breg, J. N., van Opheusden, J. H. J., Burgering, M. J., Boelens, R., & Kaptein, R. (1994) *Nature 346,* 586−589.
Brenstein, R. J. (1989) Robelko Software, version 0.9 8b5, Carbondale, IL.
Burgering, M. J. M., Hald, M., Boelens, R., Breg, J. N., & Kaptein, R. (1995) *Biopolymers 35,* 217−226.
Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92,* 10869−10873.
Hammond, G. S. (1955) *J. Am. Chem. Soc. 77,* 334−338.
Johnson, M., & Frasier, S. (1985) *Methods Enzymol. 117,* 301−342.
Matouschek, A., & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 7814−7818.
Matouschek, A., Matthews, J. M., Johnson, C. M., & Fersht, A. R. (1994) *Protein Eng. 7,* 1089−1095.
Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33,* 1125−1133.
Milla, M. E., Brown, B. M., & Sauer, R. T. (1993) *Protein Sci. 2,* 2198−2205.
Milla, M. E., Brown, B. M., Waldburger, C. D., & Sauer, R. T. (1995) *Biochemistry 34,* 13914−13919.
Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., & Socci, N. D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92,* 3626−3630.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3,* 1−43.
Peng, X., Jonas, J., & Silva, J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 1776−1780.
Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994) *Nature 367,* 754−757.
Sali, A., Shaknovich, E., & Karplus, M. (1994) *Nature 369,* 248−251.
Schellman, J. A. (1978) *Biopolymers 17,* 1305−1322.
Schildbach, J. F., Milla, M. E., Jeffrey, P. D., Raumann, B. E., & Sauer, R. T. (1995) *Biochemistry 34,* 1405−1412.
Schindler, T., Herrler, M., Marahiel, M. A., & Schmid, F. X. (1995) *Nat. Struct. Biol. 2,* 663−673.
Tanford, C. (1964) *J. Am. Chem. Soc. 86,* 2050−2059.
Tanford, C. (1970) *Adv. Protein Chem. 24,* 1−95.
Tanford, C., Aune, K. C., & Ikai, A. (1973) *J. Mol. Biol. 73,* 185−197.
Waldburger, C. D., Schildbach, J. F., & Sauer, R. T. (1995) *Nat. Struct. Biol. 2,* 122−128.
Waldburger, C. D., Jonsson, T., & Sauer, R. T. (1996) *Proc. Natl. Acad. Sci. U.S.A.,* in press.
Wendt, H., Berger, C., Baici, A., Thomas, R. M., & Bosshard, H. R. (1995) *Biochemistry 34,* 4097−4107.
Wolynes, P. G., Onuchic, J. N., & Thirumalai, D. (1995) *Science 267,* 1619−1620.

BI953056S