P22 Arc Repressor: Transition State Properties Inferred from Mutational Effects on the Rates of Protein Unfolding and Refolding†

Marcos E. Milla,‡ Bronwen M. Brown,§ Carey D. Waldburger, and Robert T. Sauer*

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

Received June 28, 1995; Revised Manuscript Received September 1, 1995⊗

ABSTRACT: The kinetics of unfolding and refolding have been measured for a set of Arc repressor mutants bearing single amino acid substitutions at 44 of the 53 residue positions. Roughly half of the mutations cause significant changes in the unfolding and/or refolding rate constants. These substitutions alter the hydrophobic core, tertiary hydrogen bonds and salt bridges, and glycines with restricted backbone conformations. Overall, the mutations cause larger changes in the unfolding rates than the refolding rates, indicating that significantly less side-chain information is used between the denatured state and transition state than between the transition state and native state. The set of mutants displays reasonable Brønsted behavior, suggesting that many native interactions are partially formed in the transition state. Taken together, these observations suggest that the overall structure of most of the protein must be somewhat native-like in the transition state but without close, complementary packing of the hydrophobic core or good hydrogen bond geometry. Such a transition state is inconsistent with a model in which monomers fold to their correct conformations and then dock to form the dimer but supports a model in which folding and dimerization are concurrent processes.

Understanding the states of a protein that are populated at different times during a folding reaction, how interactions mediated by specific residues contribute to the free energies of these states, and the nature of the transition states which determine the heights of the kinetic barriers along the folding pathway are all parts of the protein-folding problem. Studies of the unfolding and refolding of wild-type and mutant proteins in a variety of model systems have begun to provide answers to these questions [for reviews, see Goldenberg (1988), Kim and Baldwin (1990), Creighton (1992), Matthews (1993), Fersht and Serrano (1993), Shortle (1993), Evans and Radford (1994), and Fersht (1995)]. However, most of the model proteins which have been studied are monomers. Far less is known about proteins which only fold as dimers or higher oligomers, despite the fact that such proteins are relatively common in biological systems.

The Arc repressor of phage P22 has several features that make it an attractive model system for studying the folding of a protein whose native state is dimeric. The polypeptide chain is small (53 residues; Sauer et al., 1983; Vershon et al., 1985), and the protein has a well-defined, roughly globular, tertiary structure (Breg et al., 1990; Raumann et al., 1994). As shown in Figure 1, the dimer consists of an intersubunit, antiparallel β-sheet (residues 8−14 and 8′−14′) and four α-helices (residues 15−30, 15′−30′, 32−48, and 32′−48′). These elements of secondary structure and the intervening turns enclose a conventional, tightly-packed hydrophobic core. Formation of the native Arc dimer from denatured monomers occurs reversibly and without detectable intermediates in both equilibrium and kinetic studies, suggesting that folding and dimerization can be modeled as concerted processes (Bowie & Sauer, 1989a; Milla & Sauer, 1994). Finally, Arc acts as a simple genetic repressor *in vivo* [for review, see Susskind and Youderian (1983)], and straightforward assays and selections for intracellular function have been devised and used to characterize collections of mutants (Vershon et al., 1986; Bowie & Sauer, 1989b).

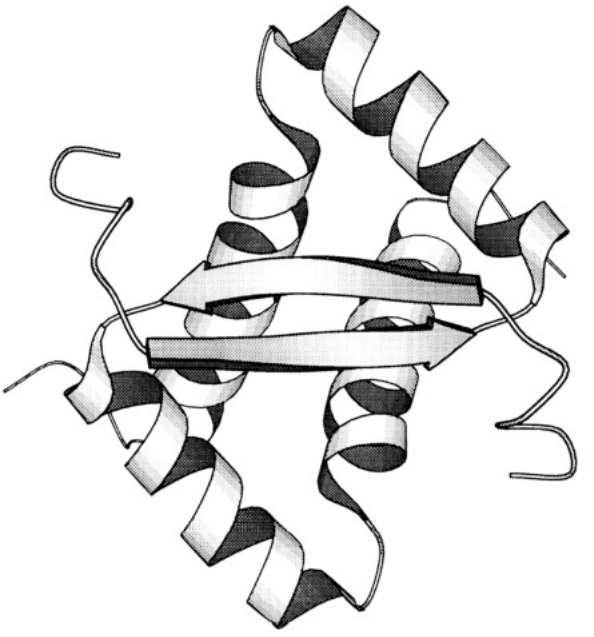

FIGURE 1: View of the Arc dimer as a polypeptide backbone ribbon trace. Prepared with MOLSCRIPT (Kraulis, 1991).

In previous work, we reported the construction of a complete library of alanine-substitution mutations and the effects of these alterations on the equilibrium stability of Arc

† Supported by NIH Grant AI-15706 and by postdoctoral fellowships from the Jane Coffin Childs Fund for Medical Research (M.E.M.) and the National Institutes of Health (C.D.W.).

‡ Current address: Glaxo-Wellcome Research Institute, Research Triangle Park, NC 27709.

§ Current address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403.

⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1995.

0006-2960/95/0434-13914$09.00/0

(Milla et al., 1994). Here, we extend our characterization of this library of mutant proteins by determining rate constants for protein unfolding and refolding. Mutations causing significant changes in these kinetic parameters alter positions that are involved in formation of much of the hydrophobic core, residues involved in hydrogen bonds and salt bridges, and glycines with conformationally-restricted backbone conformations. In general, mutations at these positions affect both the folding and unfolding rates in a manner expected for interactions that are partially formed in the transition state. This suggests that the overall structure must be somewhat native-like in the transition state. Overall, however, the effects of the mutations on the rates of refolding are far smaller than on the rates of unfolding, suggesting that the transition state is closer to the denatured state than to the native state in terms of the energetic importance of side-chain information. This finding suggests that the hydrophobic core is not well-packed in the transition state and that hydrogen bonds do not have good geometry in this state.

MATERIALS AND METHODS

Mutants. The construction and purification of the mutants used in this paper have been described (Brown et al., 1994; Milla et al., 1994). To reduce intracellular proteolysis and to allow affinity purification (Milla et al., 1993), each mutant contains a short, unstructured C-terminal extension designated *st6* (His-His-His-His-His-His) or *st11* (His-His-His-His-His-His-Lys-Asn-Gln-His-Glu). Neither tail sequence causes significant changes in the folding or stability of otherwise wild-type Arc (Milla et al., 1993). In addition, the properties of mutants containing the *st6* tail can be compared with otherwise wild-type Arc-*st6* and those containing the *st11* tail can be compared with otherwise wild-type Arc-*st11*. Mutants are named using the one-letter code for the wild-type residue, the mutant residue, and the residue position. Thus, FA10 refers to a mutant in which Phe10, the wild-type residue, is replaced by Ala.

Unfolding and Folding Kinetics. Experiments were performed at 25 °C using an Applied Photophysics DX17.MV stopped-flow instrument to monitor changes in the fluorescence of Trp14 at 320 nm after excitation at 280 nm. To measure rates of unfolding, mutant proteins at a concentration of 10−50 μM in 50 mM Tris·HCl (pH 7.5), 0.2 mM EDTA, and 0.25 M KCl were diluted into the same buffer containing 2−8 M urea. Unfolding data fit well to the single exponential function $F = F_0 + F_1(1 - e^{-k_u t})$, where k_u is the unfolding rate constant (Milla & Sauer, 1994).

Rates of refolding were measured by pH-jump experiments in which 10 μM protein in low-pH denaturation buffer (10 mM phosphoric acid, 0.2 mM EDTA, and 0.25 M KCl) was mixed with an equal volume of buffer containing 100 mM Tris·HCl, 0.2 mM EDTA, and 0.25 M KCl (the pH following mixing was 7.5). The folding rate of the hyperstable PA8 mutant could not be measured by this method because it was not fully unfolded in the low-pH denaturation buffer. For this mutant, refolding rate constants (k_f) were measured following jumps from 9 M urea to lower urea concentrations (1.5−2.75 M) in buffer containing 50 mM Tris·HCl (pH 7.5), 0.2 mM EDTA, and 0.25 M KCl. A plot of log(k_f) vs [urea] was approximately linear ($r = 0.999$), allowing a value of k_f in the absence of urea to be calculated by linear regression. For each refolding experiment, a value of the refolding/dimerization rate constant (k_f) was obtained by nonlinear least-squares fitting of data using the program NONLIN (Johnson & Frasier, 1985; Brenstein, 1989) to equations that include both the bimolecular association and unimolecular dissociation reactions (Milla & Sauer, 1994; eq 6−12). The values of k_u calculated in the absence of denaturant from the unfolding experiments were used for these calculations. However, for all but the most unstable mutants (those with values of $k_u > 3$ s^{-1}), values of k_f calculated in the manner described were within 30% of values calculated by fitting to a hyperbolic equation for a simple homodimerization reaction:

$$F = F_0 + F_1\left(\frac{[P_t]k_f t}{1 + [P_t]k_f t}\right)$$

where [P_t] is the total protein concentration in monomer equivalents (Milla & Sauer, 1994). Additional details of the assay procedures and fitting protocols are described in Milla and Sauer (1994). Equilibrium values of ΔG_u (25 °C, pH 7.5, and 250 mM KCl, 1 M standard state) and m for the mutants were taken from Milla et al. (1994). Equilibrium values of $\Delta\Delta G_u$ were calculated as $\Delta G_u^{mut} - \Delta G_u^{wt}$. Values of $\Delta\Delta G_u$ from kinetics were calulated as $-RT \ln(k_u^{mut}k_f^{wt}/k_u^{wt}k_f^{mut})$. Although the ΔG_u values for a dimeric protein depend upon the choice of a standard state concentration, $\Delta\Delta G_u$ values are independent of concentration.

RESULTS

At 25 °C, pH 7.5, and 250 mM KCl, the wild-type Arc dimer unfolds/dissociates with a rate constant of ≈0.1 s^{-1} and refolds/dimerizes with a second-order rate constant of ≈10^7 M^{-1} s^{-1} (Milla & Sauer, 1994). To determine the effects of mutations throughout the protein sequence on the rates of denaturation and renaturation, we determined rate constants for unfolding (k_u) and refolding (k_f) for 43 alanine-substitution mutants and one leucine-substitution mutant of Arc. Of the 51 possible alanine-substitution mutants of Arc, seven of the purified proteins (WA14, VA22, EA36, IA37, RA40, VA41, and FA45) were so thermodynamically unstable that they did not form native dimers or formed dimers that were too unstable for reliable kinetic studies (Milla et al., 1994; Milla & Sauer, 1995; Waldberger et al., 1995). Two additional positions (26 and 53) are alanine in wild-type Arc.

Unfolding/Dissociation. Rate constants for unfolding/dissociation were determined for each mutant at a series of urea concentrations. Figure 2 shows representative data for Arc and five mutants. Plots of log(k_u) versus [urea] were reasonably linear for each mutant ($r > 0.99$), allowing values of k_u in the absence of urea to be calculated by extrapolation. Table 1 lists the values of the mutant rate constants relative to the wild-type parent ($k_{u\text{-}rel} = k_u^{mut}/k_u^{wt}$). These values for different variants span a range of approximately 4 orders of magnitude. For example, unfolding of the PA8 mutant is more than 100-fold slower than wild type, while unfolding of the YA38 variant is more than 100-fold faster. Almost half of the mutants, however, unfold with rate constants within 2.5-fold of wild type.

Figure 2 also shows that the slopes of the log(k_u) vs [urea] plots can differ significantly for different mutants. These

Table 1: Properties of Mutant Proteins[a]

	$\frac{k_f^{wt}}{k_f^{mut}}$	$\frac{k_u^{mut}}{k_u^{wt}}$	$\frac{\partial \log(k_u)}{\partial[\text{urea}]}$	$\frac{m_u}{m}$	structural role of wild-type side chain	H-bond partners of wild-type side chain
MA1–*st6*	1.09	2.20	0.24	0.20	flexible arm	
KL2–*st6*	0.96	0.92	0.31	0.26	flexible arm	
GA3–*st6*	0.90	1.16	0.28	0.21	flexible arm	
MA4–*st6*	0.91	1.08	0.28	0.23	flexible arm	
SA5–*st6*	0.82	1.31	0.26	0.22	flexible arm	
KA6–*st6*	0.78	0.75	0.30	0.26	flexible arm	
MA7–*st6*	1.05	2.61	0.24	0.21	partially buried	
PA8–*st6*	0.41	0.0048	0.32	0.31	partially buried	
QA9–*st6*	0.99	1.34	0.26	0.22	surface	
FA10–*st6*	2.24	21.69	0.15	0.16	hydrophobic core	
NA11–*st6*	0.85	1.33	0.17	0.16	surface	
LA12–*st11*	4.34	15.93	0.17	0.19	hydrophobic core	
RA13–*st6*	1.11	2.29	0.19	0.19	surface	
PA15–*st11*	1.32	58.33	0.09	0.08	core, N-cap	
RA16–*st6*	0.67	0.93	0.29	0.25	salt bridge	D20
EA17–*st6*	1.14	0.89	0.32	0.28	hydrogen bond	backbone NH E17
VA18–*st6*	1.10	0.32	0.37	0.30	hydrophobic core	
LA19–*st6*	1.79	2.20	0.31	0.28	hydrophobic core	
DA20–*st6*	1.07	1.71	0.29	0.25	salt bridges	R16, R23
LA21–*st11*	7.22	49.07	0.20	0.22	hydrophobic core	
RA23–*st11*	1.59	2.21	0.24	0.23	salt bridge	D20
KA24–*st11*	1.27	2.28	0.27	0.23	surface	
VA25–*st6*	0.86	0.97	0.21	0.23	hydrophobic core	
EA27–*st6*	1.09	0.85	0.27	0.22	surface	
EA28–*st11*	1.84	5.33	0.23	0.20	salt bridge	R50′
NA29–*st11*	2.14	31.67	0.23	0.18	hydrogen bond	E36
GA30–*st11*	4.05	5.00	0.30	0.28	positive ϕ,ψ	
RA31–*st11*	4.76	82.87	0.28	0.26	salt bridge	E36
SA32–*st11*	7.17	87.22	0.20	0.22	N-cap H-bond	backbone NH S35
VA33–*st11*	3.45	14.35	0.28	0.23	hydrophobic core	
NA34–*st11*	0.58	2.59	0.20	0.20	hydrogen bonds	backbone NH, O R13′
SA35–*st6*	0.49	0.96	0.27	0.28	surface	
YA38–*st11*	5.07	160.93	0.21	0.20	hydrophobic core	
QA39–*st11*	0.86	1.40	0.20	0.20	surface	
MA42–*st11*	8.32	36.94	0.15	0.14	hydrophobic core	
EA43–*st6*	1.00	2.25	0.28	0.26	salt bridge	K46, K47
SA44–*st11*	1.85	50.74	0.27	0.22	hydrogen bonds	R40′, backbone O R40
KA46–*st11*	1.02	4.24	0.20	0.17	salt bridge	E43
KA47–*st11*	1.00	15.83	0.18	0.16	salt bridge	E43
EA48–*st11*	1.78	61.67	0.14	0.14	salt bridge	R40
GA49–*st11*	1.40	25.00	0.17	0.16	positive ϕ,ψ	
RA50–*st11*	1.71	22.31	0.17	0.15	salt bridge	E28′
IA51–*st11*	1.82	11.48	0.21	0.21	hydrophobic core	
GA52–*st11*	1.26	1.11	0.28	0.24	surface	
Arc–*st6*	1.00	1.00	0.30	0.26		
Arc–*st11*	1.00	1.00	0.27	0.22		

[a] Kinetic values are calculated relative to the Arc-*st6* ($k_u = 0.12\ s^{-1}$; $k_f = 7.1 \times 10^6\ M^{-1}\ s^{-1}$) or Arc-*st11* ($k_u = 0.14\ s^{-1}$; $k_f = 1.09 \times 10^7\ M^{-1}\ s^{-1}$) proteins depending on the C-terminal tail of the mutant. The standard errors in the values of $k_{f\text{-rel}}$, $k_{u\text{-rel}}$, $\partial \log(k_u)/\partial[\text{urea}]$, and m_u/m are approximately ±25%, ±10%, ±5%, and ±20% of the values, respectively. Errors were calculated from linear regression or nonlinear least-squares fitting procedures and were propagated by standard methods. Structural information is taken from the 2.2 Å resolution crystal structure of the wild-type Arc dimer (C. Kissinger, U. Obeysekare, L. J. Keefe, B. Raumann, R. T. Sauer, and C. O. Pabo, unpublished).

slopes are listed for each mutant in Table 1, as are values of m_u/m (which is equal to $\partial \log(k_u)/\partial[\text{urea}]$ divided by $\partial \log(K_u)/\partial[\text{urea}]$). m_u/m is a fractional value that represents the increase in the solvent exposure of the transition state relative to the native state, normalized by the increase in solvent exposure between the native and denatured states (Tanford, 1970). A small value of m_u/m indicates that the native and transition states have similar solvent accessibilities and *vice versa*. For wild-type Arc, for example, m_u/m is about 0.25, indicating that approximately 25% of the total buried hydrophobic surface is exposed in going from the native state to the transition state (Milla & Sauer, 1994). The m_u/m values for the PA15, MA42, and EA48 mutants are only 0.08–0.14, suggesting that significant changes occur in the solvent exposure of the transition states of these mutants, bringing them closer to the native states.

Refolding/Dimerization. Rate constants for the refolding/dimerization of each mutant were determined by pH-jump or urea-jump experiments, and values relative to the wild-type parent ($k_{f\text{-rel}} = k_f^{wt}/k_f^{mut}$) are listed in Table 1. Three mutants appear to refold faster than wild type by a factor of approximately 2, the great majority of mutants refold at approximately the wild-type rate, and only 10 mutants refold more than 2-fold more slowly than wild type. Overall, there is a much smaller range in the values of $k_{f\text{-rel}}$ for the mutants than in the values of $k_{u\text{-rel}}$.

If the refolding and unfolding experiments probe the same transition state, as expected if the use of different conditions (no urea for refolding; high urea for unfolding) does not change the rate-determining step, then changes in protein stability from equilibrium experiments ($\Delta\Delta G_u$; Milla et al., 1994) and kinetic experiments should be similar. Figure 3

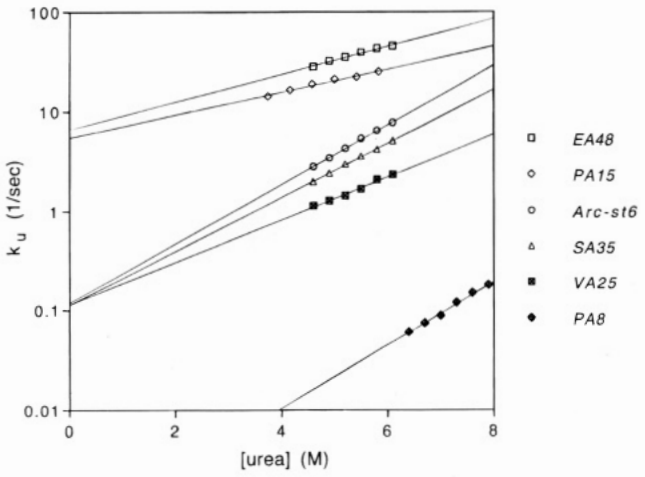

FIGURE 2: Logarithmic dependence of the rate constants for unfolding/dissociation (k_u) on urea concentration for wild-type Arc-*st6* and five mutants. The solid lines are theoretical fits and have correlation coefficients (r) of 0.99 or higher.

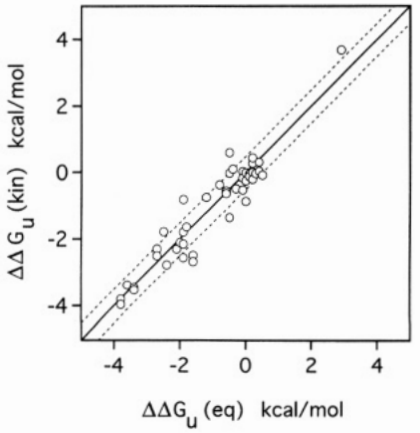

FIGURE 3: Relationship of $\Delta\Delta G_u$ values calculated from equilibrium and kinetic experiments. The solid line is expected for a perfect correlation. The dotted lines represent an error of ±0.5 kcal/mol, the 67% confidence levels for the equilibrium $\Delta\Delta G_u$ values (Milla et al., 1994).

shows this comparison for the mutants studied here. For the majority of mutants, the $\Delta\Delta G_u$ values agree to within ±0.5 kcal/mol, the 67% confidence level for the equilibrium $\Delta\Delta G_u$ measurements. In a few cases, however, the $\Delta\Delta G_u$ values calculated from the kinetic experiments are approximately 1 kcal/mol larger or smaller than those calculated from equilibrium experiments. In these cases, there may be changes in the rate-determining step or the errors in the measured values of the equilibrium or kinetic constants may be larger than estimated from the fitting procedures.

DISCUSSION

The variation in the values of m_u/m for some of the alanine-substituted proteins suggests that their transition states occur at different positions along a reaction coordinate defined by solvent accessibility (Matouschek & Fersht, 1993). For the mutations analyzed here, the fractional increase in the solvent exposure of the transition states for different mutants varies from 0.08 to 0.31, with an average of 0.22 (±0.04) which is close to the wild-type value. If unfolding can occur by distinct, parallel pathways, a mutation might change the preferred path or even create a new path. By this model, mutants with significantly different values of m_u/m would

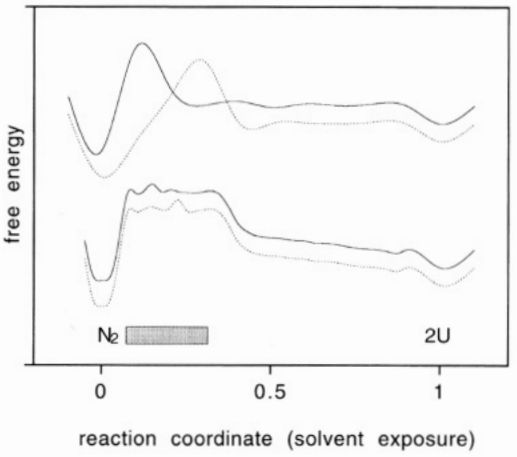

FIGURE 4: Models for changes in the position of the transition state for different mutants. The reaction coordinate is defined by solvent exposure of surface area as measured by m_u/m values. The top traces represent hypothetical alternative, parallel pathways. The bottom traces indicate how changes in the rate-determining step could occur along a single hypothetical pathway. The solid bar shows the distribution of m_u/m values for the mutants listed in Table 1. Because Arc-refolding kinetics are second-order in protein concentration, the highest point on the free energy profile could be the transition state for the dimerization reaction. If, however, there is a rate-determining preequilbrium involving dimerization, then the highest point could be the transition state for a unimolecular reaction.

unfold by largely different pathways in which the relationship of solvent exposure to energy could be quite different (Figure 4, top traces). A more conservative interpretation is that the wild type and the mutants go through the same basic series of unfolding steps along a single path but some mutations change the step that is rate determining. For example, if all protein states with values of m_u/m between 0.05 and 0.35 lie on a high, bumpy energy plateau, then mutations could simply change the positions of the highest bumps (Figure 4, bottom traces). In either case, information inferred from the kinetic properties of the set of mutants will refer to an average or composite transition state.

At a general level, the alanine substitutions cause significantly larger changes in the unfolding/dissociation rates than in the refolding/dimerization rates. This can clearly be seen in Figure 5 which shows the relative changes in rate constants for each of the 44 mutants studied here. To permit quantitative comparisons of the net effects of these mutations, we calculated the sum of the free energy changes ($\Sigma\Delta\Delta G_u$ = 45.5 kcal/mol; Milla et al., 1994), and the sums of energetic contributions from the changes in unfolding rates ($\Sigma RT \ln(k_{u\text{-rel}})$ = 37 kcal/mol) and refolding rates ($\Sigma RT \ln(k_{f\text{-rel}})$ = 10 kcal/mol). By this very rough measure, about 78% of the information contributed by these 44 side chains is utilized between the native state and transition state, and only about 22% is utilized between the denatured state and transition state. Hence, in terms of the importance of side-chain information, the transition state appears to be closer to the denatured state than to the native state.

What do the kinetic parameters for the Arc mutants tell us about the importance of specific side chains in the processes of acquisition and maintenance of structure for the Arc dimer? Of the mutations studied, 19 increase the rate of unfolding by 3-fold or more, with the maximum effect being 160-fold. These 19 substitutions affect side chains which play a variety of structural roles, including packing

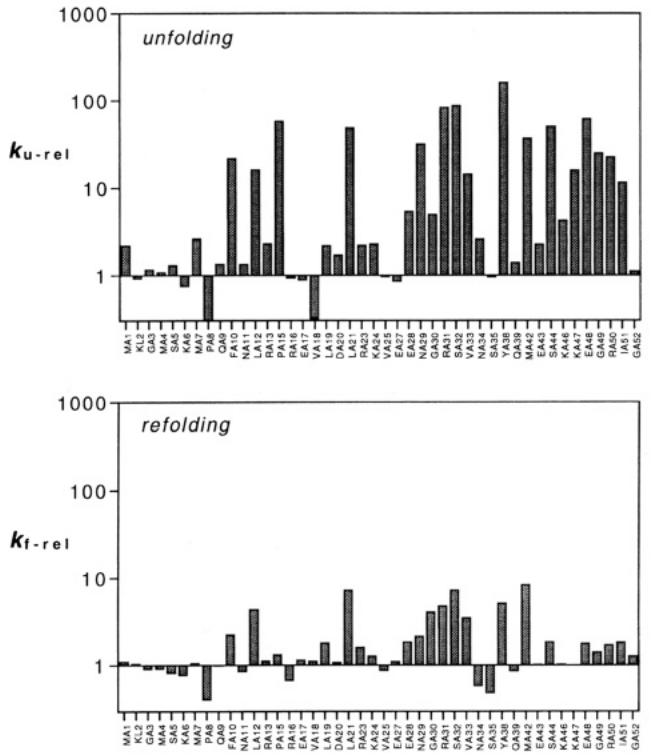

FIGURE 5: Summary of the effects of mutations on the relative rates of unfolding/dissociation (top) and refolding/dimerization (bottom). The *y*-axis is a log scale.

of the hydrophobic core (FA10, LA12, PA15, LA21, VA33, YA38, MA42, and IA51), formation of hydrogen bonds and salt bridges (EA28, NA29, RA31, SA32, SA44, KA46, KA47, EA48, and RA50), and formation of restricted polypeptide backbone conformations (GA30 and GA49). The eight mutations affecting hydrophobic core residues and nine mutations affecting hydrogen bonds and salt bridges have roughly equivalent net effects ($\Sigma RT \ln(k_{u\text{-}rel}) = 16.3$ and 17.3 kcal/mol, respectively), suggesting that both classes of interactions contribute about equally to kinetic stabilization of the native structure. Although most of the destabilizing mutations also decrease the rate of refolding, only 10 of the mutations decrease this rate by more than 2-fold and the maximum effect is only 8-fold.

In two cases, alanine substitution mutations were found to cause significant decreases in the rate at which the mutant protein unfolds relative to wild type. The VA18 mutation alters a hydrophobic core residue in α-helix A. Because this mutation decreases the unfolding rate without affecting the refolding rate, it appears that mutation of the valine side chain to alanine lowers the energy of the native state relative to the transition state. This might occur because a bad steric contact is relieved or because the mutation results in structural rearrangements which permit new and favorable interactions at nearby positions. We note that a variety of other hydrophobic side chains (Phe, Ile, Leu, Met) are allowed at position 18 in randomization experiments (Bowie & Sauer, 1989b), suggesting that the structure can relax to accommodate side chains of different sizes at this core position. In the other case of a stabilizing substitution, the PA8 mutation alters a residue at the beginning of the antiparallel β-sheet. This mutation causes an approximate 200-fold decrease in the rate of unfolding and a 2-fold increase in the rate of refolding, resulting in a significant increase in thermodynamic stability. The decreased unfolding rate and increased stability of the PA8 mutant are similar to those of a previously-characterized mutant, PL8, bearing a substitution at the same position (Schildbach et al., 1995). The crystal structure of the PL8 protein shows that the substitution of an amino acid for the wild-type imino acid allows a new hydrogen bond to form at each end of the β-sheet. It is likely that this same effect is responsible for much of the decreased unfolding rate and increased thermodynamic stability of the PA8 protein. One difference between PL8 and PA8 is that the former mutation does not affect the folding rate whereas the latter mutation causes a slight increase. This suggests that bulkier side chains at position 8 may slow folding to a small degree. Both the PA8 and PL8 mutants bind operator DNA poorly, suggesting that Pro8 is needed for high-affinity DNA binding even though it destabilizes the wild-type protein (Brown et al., 1994; Schildbach et al., 1995).

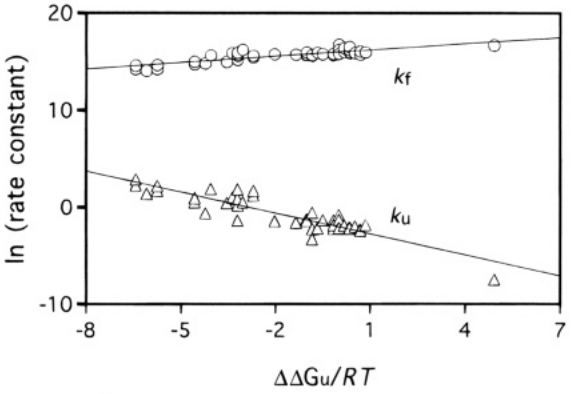

FIGURE 6: Brønsted plots for the unfolding and refolding constants of the 44 mutants listed in Table 1. $\Delta\Delta G_u/RT = \ln(k_{u\text{-}rel}) + \ln(k_{f\text{-}rel})$. The solid lines are linear regression fits.

Only a few mutations affect the unfolding rate significantly without affecting the folding rate, and no mutations alter the refolding rate significantly without also changing the unfolding rate. This suggests that very few side chains in Arc mediate interactions that are completely broken before the transition state for unfolding or are completely formed before the transition state for refolding. Rather, almost all of the core and polar residues implicated in the unfolding reaction by significant mutational effects are also implicated in the refolding reaction (Figure 5). Such behavior is consistent either with a reaction in which most of these interactions are partially formed in the transition state of a single kinetic pathway, or with a reaction composed of a set of parallel pathways, each with a distinct transition state in which specific interactions are completely formed or broken. Fersht et al. (1994) have noted that reactions of the former type should display Brønsted behavior, in which plots of $\Delta\Delta G_u/RT$ vs $\ln(k_u)$ and $\ln(k_f)$ for a set of mutants are approximately linear with slopes of $-\beta$ and $1 - \beta$, respectively, as long as changes in rate constants and bond energies are related in a simple fashion. As shown in Figure 6, the entire set of Arc mutants shows reasonable Brønsted behavior, for the unfolding reaction (slope $= -0.72 \pm 0.05$; $r = 0.90$) and refolding reaction (slope $= 0.22 \pm 0.02$; $r = 0.84$). The assumption that changes in rate constants reflect changes in noncovalent bond energies is expected to be better for mutations that replace hydrophobic groups with alanine than for those that replace polar residues with alanine (Fersht et al., 1992). Indeed, for the 14 Arc mutations that involve hydrophobic

to alanine substitutions, the fit to the Brønsted relationshipis slightly better for both unfolding (slope = -0.80 ± 0.08; $r = 0.93$) and refolding (slope = 0.22 ± 0.03; $r = 0.90$). The ratio of the Brønsted slopes ($-\beta/(1-\beta) \approx 3.5$) also gives an estimate of the partitioning of side-chain information between the unfolding and refolding reactions that is similar to that calculated from the ratio of the total side-chain contributions to both reactions ($\Sigma RT \ln(k_{u\text{-}rel})/\Sigma RT \ln(k_{f\text{-}rel}) \approx 3.7$). Taken together, the sensitivity of the position of the transition state to mutations and the approximate Brønsted behavior, suggest that the set of mutants share a series of high-energy, compact states in which many native interactions are partially formed.

The concentration dependence of the wild-type refolding rate indicates that the transition state is dimeric, and the urea dependence of the refolding and unfolding rates indicates that about three-quarters as much solvent accessible surface is buried in the transition state as in the native structure (Milla & Sauer, 1994). These properties suggest that, with respect to burial of hydrophobic groups and oligomeric form, the transition state is more similar to the native state than to the denatured state. However, as discussed above, from the perspective of side-chain information, the transition state seems more similar to the denatured state. These findings are not necessarily in conflict. The chemical details of side-chain structure are likely to be most important in the complementary, close packing of side chains in the hydrophobic core and in the formation of hydrogen bonds with the proper geometry between donor and acceptor atoms. Since only 20%−25% of the side-chain information seems to be utilized in the transition state, it appears that hydrophobic groups are likely to be loosely packed and that polar interactions, although present to some extent, have poor geometry. By this model, most tight core packing and improvement of hydrogen bond geometry would occur after the transition state in the refolding reaction. If this is true, however, then these processes must occur very rapidly since they do not appear to be rate limiting for the overall Arc-refolding reaction, which can be completed on the millisecond time scale (Milla & Sauer, 1994).

For some dimeric proteins, monomer folding precedes the docking reaction. In a folding-before-docking model, intramonomer interactions should form relatively early with monomer−monomer interactions forming later. The kinetic results reported here do not support such a model, however, as most native interactions seem to be partially formed in the composite transition state. This implies that folding and dimerization of Arc repressor occur more or less concurrently. In the native Arc dimer, the monomers wrap around and interact with each other over an extensive surface contributed by many different parts of the protein (Figure 1). This architecture also makes it unlikely that monomers fold to their correct conformations before docking. The initial interaction of Arc monomers may be similar to the hydrophobic collapse step postulated for monomeric proteins. By the time the protein reaches the transition state, however, numerous hydrogen bonds and salt bridges, as well as hydrophobic interactions, have begun to contribute to stabilization. This suggests that many of the tertiary interactions must already be approximately in place in the transition state, requiring further exclusion of solvent, tight packing of the core, and proper alignment of polar interactions to attain the native structure.

ACKNOWLEDGMENT

We thank Thor Jonsson, Matt Cordes, and Frank Solomon for discussion and comments on the manuscript.

REFERENCES

Bowie, J. U., & Sauer, R. T. (1989a) *Biochemistry 28*, 7139−7143.
Bowie, J. U., & Sauer, R. T. (1989b) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2152−2156.
Breg, J. N., Opheusden, J. H. J., Burgering, M. J. M., Boelens, R., & Kaptein, R. (1990) *Nature 346*, 586−589.
Brenstein, R. J. (1989) *Robelko Software, version 0.9 8b5*, Carbondale, IL.
Brown, B. M., Milla, M. E., Smith, T. L., & Sauer, R. T. (1994) *Nature Struct. Biol. 1*, 164−168.
Creighton, T. E. (1992) *BioEssays 15*, 195−199.
Evans, P. A., & Radford, S. E. (1994) *Curr. Opin. Struct. Biol. 4*, 100−106.
Fersht, A. R. (1995) *Curr. Opin. Struct. Biol. 5*, 79−84.
Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol. 3*, 75−83.
Fersht, A. R., Matoushek, A., & Serrano, L. (1992) *J. Mol. Biol. 224*, 771−782.
Fersht, A. R., Itzhaki, L. S., ElMasry, N. F., Matthews, J. M., & Otzen, D. E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10426−10429.
Goldenberg, D. P. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 481−507.
Johnson, M., & Frasier, S. (1985) *Methods Enzymol. 117*, 301−342.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Matouschek, A., & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7814−7818.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
Milla, M. E., & Sauer, R. T. (1995) *Biochemistry 34*, 3344−3351.
Milla, M. E., Brown, B. M, & Sauer, R. T. (1993) *Protein Sci. 2*, 2198−2205.
Milla, M. E., Brown, B. M., & Sauer, R. T. (1994) *Nature Struct. Biol. 1*, 518−523.
Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994) *Nature 367*, 754−757.
Sauer, R. T., Krovatin, W., DeAnda, J., Youderian, P., & Susskind, M. (1983) *J. Mol. Biol. 168*, 699−713.
Schildbach, J. F., Milla, M. E., Jeffrey, P. D., Raumann, B. E., & Sauer, R. T. (1995) *Biochemistry 34*, 1405−1412.
Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
Susskind, M. M., & Youderian, P. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., & Weisberg, R. A., Eds.) pp 347−364, Cold Spring Harbor Laboratory, NY.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
Vershon, A. K., Youderian, P., Susskind, M. M., & Sauer, R. T. (1985) *J. Biol. Chem. 260*, 12124−12129.
Vershon, A. K., Bowie, J. U., Karplus, T. M., & Sauer, R. T. (1986) *Proteins 1*, 302−311.
Waldburger, C. D., Schildbach, J. F., & Sauer, R. T. (1995) *Nature Struct. Biol. 2*, 122−128.

BI9514651