P22 Arc Repressor: Folding Kinetics of a Single-Domain, Dimeric Protein†

Marcos E. Milla and Robert T. Sauer*

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

Received August 25, 1993; Revised Manuscript Received November 29, 1993®

ABSTRACT: The rate constants for refolding and unfolding of the P22 Arc repressor dimer have been determined by stop-flow fluorescence experiments. Under most conditions, refolding is described well as a two-state reaction with a bimolecular rate-limiting step ($k_f \approx 10^7$ M^{-1} s^{-1}). A unimolecular step appears to become co-rate limiting at high protein concentrations. The urea dependence of the refolding reaction suggests that about 75% of the total burial of hydrophobic surface occurs between the unfolded state and the transition state for folding. Hydrophobic interactions are also evidenced by the temperature dependence of the refolding reaction; the rate increases with temperature and Arrhenius plots are curved, as expected for a reaction that proceeds with a significant heat capacity change. The refolding of Arc also proceeds more rapidly as the salt concentration is raised, presumably because repulsive interactions between monomers are screened. At a protein concentration of 10 μM, the apparent rate constant for refolding of the Arc dimer is $\approx$100 s^{-1}, as fast as the refolding of many monomeric proteins. The rate constant for unfolding is $\approx$0.1 s^{-1}, corresponding to a half-life of less than 10 s for the folded Arc dimer. This rate of unfolding is very fast in comparison to that of other characterized proteins and implies that a free Arc molecule must unfold and refold hundreds of times per generation in the cell.

During the last few years, dramatic advances in our understanding of the folding of monomeric proteins have occurred, largely as a consequence of experiments employing stop-flow spectroscopy, proton and disulfide quenching, peptide models, and mutagenic analyses [for reviews, see Kim and Baldwin (1990) and Matthews (1993)]. These studies generate a picture of the time frame for early folding events, such as hydrophobic collapse and secondary structure formation, and also provide information concerning the roles of specific interactions in stabilizing folding intermediates.

In terms of molecular detail, much less is known about the folding of oligomeric proteins [for review, see Jaenicke (1987)]. For many such proteins, including members of the dehydrogenase family (Krebs et al., 1979; Herman et al., 1981; Vaucheret et al., 1987), aspartate transcarbamylase (Bothwell & Schachman, 1980a,b; Burns & Schachman, 1982a,b), tryptophan synthase (Blond & Goldberg, 1986), cytochrome *c* (Elöve et al., 1992), and enzymes of the glycolytic pathway (Herman et al., 1983; Le Bras et al., 1989), the kinetics of folding are complex. In general, the folding of monomers proceeds to a significant extent before subunit association of these proteins. Moreover, association appears to be a slow process requiring minutes to hours, which competes with nonproductive pathways in which monomers aggregate irreversibly. There is a second class of oligomeric proteins typified by the trimeric tailspike protein of phage P22 (Goldenberg & King, 1982; Villafane & King, 1988) and the *Escherichia coli* Trp repressor dimer (Gittelman & Matthews, 1990; Mann & Matthews, 1993). In these cases, protein folding and association are more concerted processes, presumably because monomers are intertwined in the final structure, resulting in a blurring of the distinction between tertiary and quaternary structure. However, folding appears to be slow, even for these proteins. These examples seem to suggest that the folding of oligomeric proteins is a fundamentally slow process. In fact, for this reason it has been suggested that folding of oligomeric proteins in the cell requires the assistance of molecular chaperones (Rothman, 1989). It is possible, however, that the slow folding of many of these proteins is not an inherent property of all oligomeric proteins but rather results from other factors such as large size or the presence of multiple structural domains.

In this paper, we report studies examining the kinetics of refolding and unfolding of the P22 Arc repressor dimer. Arc is a small polypeptide (53 amino acids; Sauer et al., 1983; Vershon et al., 1985). In the structure of the globular Arc dimer, each monomer is interwound with the other, forming a single folded domain with a conventional hydrophobic core (Breg et al., 1990). The kinetic studies described here add to an extensive list of structural, genetic, and biochemical information about Arc. In the cell, Arc functions as a transcriptional regulator by binding to a single phage operator site and repressing transcription from the P_{ant} promoter (Susskind & Youderian, 1983). The solution structure of the Arc dimer has been solved by NMR (Breg et al., 1990), the crystal structure of the protein is known (C. Kissenger, U. Obeysekare, B. Raumann, L. Keefe, R. Sauer, and C. Pabo, in preparation), and the cocrystal structure of the repressor–operator complex has recently been determined (Raumann et al., 1994). In equilibrium experiments, the denaturation and renaturation of the Arc dimer is a reversible, two-state process in which only folded dimers and unfolded monomers are significantly populated (Bowie & Sauer, 1989a). A wealth of genetic information is also available, which identifies amino acid positions important for the folding, stability, and function of Arc (Vershon et al., 1986; Bowie & Sauer, 1989b; Brown et al., 1994).

Using rapid-mixing, stop-flow techniques, we have followed changes in the fluorescence emission of Arc's single tryptophan

† This work was supported by NIH Grant AI-15706 and by a postdoctoral fellowship to M.E.M. from the Jane Coffin Childs Fund for Medical Research.

® Abstract published in *Advance ACS Abstracts*, January 15, 1994.

0006-2960/94/0433-1125$04.50/0

residue (located at position 14 in the hydrophobic core) after changes in pH or in urea concentration. The refolding and unfolding reactions are quite fast, generally going to completion in the subsecond time scale. We have also studied the salt, temperature, and urea dependences of these reactions. The latter studies indicate that hydrophobic interactions play an important role in the kinetics of the Arc folding reaction.

MATERIALS AND METHODS

Purification. Arc was purified from *E. coli* strain X90 containing the overproducing plasmid pSA300 (Vershon et al., 1986; Bowie & Sauer, 1989b) by chromatography on CM-Accell and Sephadex G-50 as previously described (Bowie & Sauer, 1989c). Typical yields of Arc from 12-L preparations were 70–90 mg, and purity was greater than 95% as determined by Coomassie blue staining of samples after SDS–PAGE (Laemmli, 1970). The concentration of Arc in moles of monomer equivalents per liter was determined by using an extinction coefficient of 6756 M^{-1} cm^{-1} at 280 nm (Brown et al., 1990).

Renaturation Kinetics. Arc was denatured by dialyzing against a low-pH buffer containing 10 mM phosphoric acid, 0.2 mM EDTA, and either 0, 0.1, or 0.25 M KCl. The final pH was approximately 2. Denaturation was confirmed by monitoring the fluorescence emission spectrum and the circular dichroism spectrum as previously described (Bowie & Sauer, 1989a). In addition, equilibrium sedimentation experiments using a Beckman XL-A analytical ultracentrifuge were performed to confirm that denatured Arc was monomeric under the low-pH conditions. Duplicate samples were centrifuged for 18 h, and radial absorbance scans were taken in 10-μm steps. Twenty-five scans were averaged for each experiment. Experiments were performed at both 27 000 and 39 000 rpm. Molecular weight values were calculated by fitting the data to a one-species function assuming a partial specific volume of 0.7366 cm^3/g (van Holde, 1971; Laue et al., 1992). For the spectral and sedimentation experiments, native Arc protein in 50 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, and either 0, 0.1, or 0.25 M KCl was used as a control.

The kinetics of Arc refolding were monitored by changes in fluorescence after 1:1 dilution of low pH-denatured protein into renaturing buffer [100 mM Tris-HCl (pH 8.0), 0.2 mM EDTA, and either 0, 0.1, or 0.25 M KCl]. Following mixing, the pH of the buffer was 7.5. For experiments in which the urea dependence of Arc refolding was analyzed, the renaturing buffer also contained different concentrations of urea. All kinetic experiments were carried out using an Applied Photophysics DX17.MV stop-flow instrument, equipped with 2.5-mL syringes (Hamilton Accudil). Temperature was controlled via a thermostated water bath. Tryptophan fluorescence was excited at 280 nm, and a Corion bandpass lens (320 ± 6 nm) was used to filter fluorescence emission. For each experiment, 400 time points were collected. Generally, for each set of conditions, 13 experiments were averaged prior to analysis.

Denaturation Kinetics. Native Arc in buffer containing 50 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, and either 0, 0.1, or 0.25 M KCl was diluted into the same buffer containing different amounts of urea. Stop-flow fluorimetry determinations were performed as described in the preceding section.

Thermodynamic Stability at Equilibrium. Equilibrium denaturation of Arc was assayed by monitoring changes in tryptophan fluorescence or circular dichroism as a function of urea concentration (Bowie & Sauer, 1989a). Arc was diluted to a final concentration of 20 μM and equilibrated at 20 °C for 1 h in buffer containing 50 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, either 0.1 or 0.25 M KCl, and different concentrations of urea. Denaturation was monitored by fluorescence emission intensity at 320 nm (excitation 280 nm) using a Perkin-Elmer LS-50 luminescence spectrometer and by circular dichroism ellipticity at 234 nm using an Aviv 60DS spectropolarimeter. In both cases, readings were taken in a thermostated cell at 20 °C and the denaturation reactions were fully reversible.

Data Analysis. Kinetic and equilibrium data were fit by nonlinear least-squares procedures using a Macintosh version of the program NonLin (Johnson & Frasier, 1985; Brenstein, 1989). For kinetic analyses, data were fit to eq 2 for refolding conditions (0–1.2 M urea) and eq 5 for unfolding conditions (>3.3 M urea). In both cases, these analyses ignore the competing unfolding or folding reactions and thus introduce some error. However, simulations of the complete reactions by numerical integration of the linearized rate equations (using a range of rate constants) indicate that the likely error is within the reproducibility of the experimental rate constant measurements (10–20%). Moreover, treating the refolding reactions as relaxations to equilibrium (see below) using the experimentally determined unfolding rate constants did not change the calculated values of the folding rate constants by more than 10%. Between 1.2 and 3.3 M urea, both the refolding and unfolding rates of Arc are significant relative to each other, and the reaction was treated as a relaxation to equilibrium using eq 12 for refolding experiments and eq 13 for unfolding experiments.

For equilibrium denaturation experiments, data were fit to equations for a reaction in which unfolded monomers are in equilibrium with folded dimers, ΔG_u (the free energy of dissociation and unfolding) is a linear function of urea concentration, and the slopes and intercepts of the native and denatured base lines are variables (Pace, 1986; Santoro & Bolen, 1988; Bowie & Sauer, 1989a).

RESULTS

To allow studies of refolding kinetics, we first sought pH conditions where the Arc protein is denatured. At pH 2, 20 °C, and a protein concentration of 64 μM, Arc is unfolded by the following criteria: (i) the circular dichroism spectrum shows very little secondary structure and increased random coil content compared with native Arc (Figure 1A); (ii) the fluorescence spectrum of the single tryptophan at position 14 is red shifted and shows reduced intensity compared to native Arc (Figure 1B); and (iii) the protein behaves as a monomer in sedimentation equilibrium experiments (av M_w 5560), whereas native Arc is a dimer (av M_w 12 346) (Figure 1C). In similar experiments, we also determined that Arc is denatured at KCl concentrations of 0.1 and 0.25 M at pH 2.

Rate-Limiting Steps in Arc Refolding. The kinetics of Arc refolding were examined in stop-flow experiments in which the pH was increased from 2 to 7.5 by rapid mixing. As shown in Figure 2A, the refolding reaction monitored by fluorescence appears to consist of a single phase, and greater than 95% of the native amplitude is recovered in less than 200 ms at pH 7.5, 20 °C, and 16 μM Arc. Because native Arc is dimeric and denatured Arc is monomeric, the simplest kinetic model that could describe the refolding reaction under strongly refolding conditions is:

$$2U \xrightarrow{k_f} N_2 \qquad \frac{d[N_2]}{dt} = k_f[U]^2 \qquad (1)$$

In this model, the refolding kinetics are second-order in protein

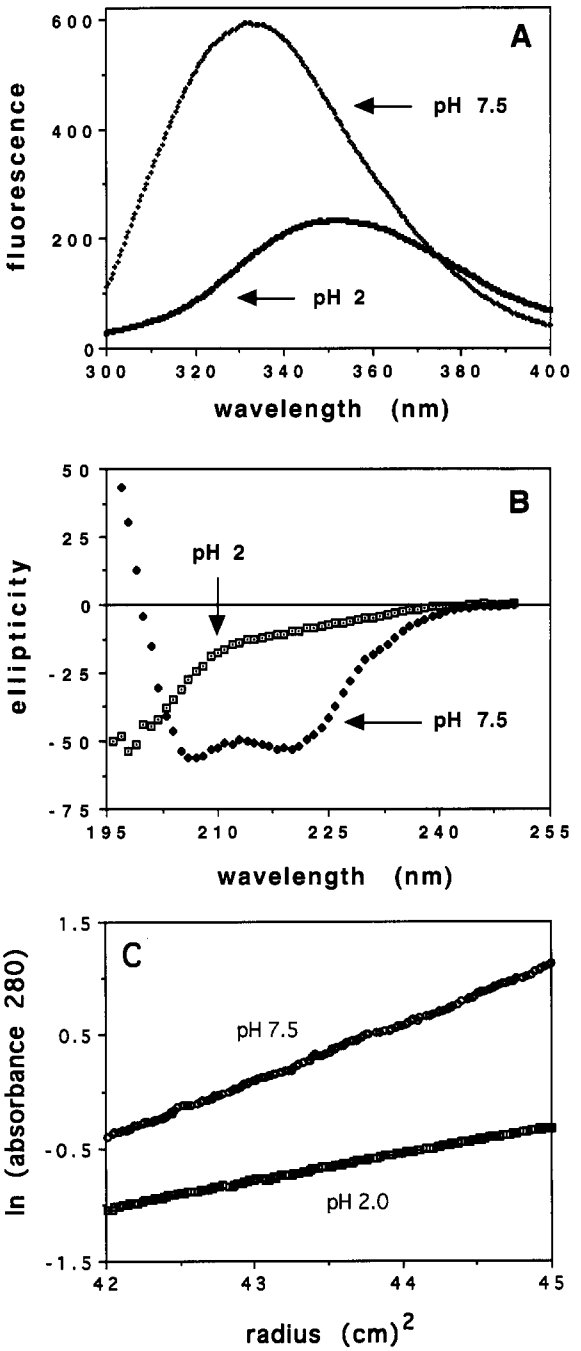

FIGURE 1: Characterization of native Arc (pH 7.5) and denatured Arc (pH 2) by (A) fluorescence emission (excitation 280 nm); (B) circular dichroism; and (C) sedimentation equilibrium centrifugation. These experiments were performed at 20 °C using an Arc concentration of 64 μM. In (C), the natural logarithm of the absorbance at 280 nm is plotted against the square of the distance from the center of the rotor following 16 h of centrifugation at a speed of 27 000 rpm. The linearity of these plots is indicative of single oligomeric species (monomers at pH 2; dimers at pH 7.5). The ordinate values for the pH 7.5 data have been increased by 1 log unit for clarity.

concentration and thus are described by the hyperbolic equation:

$$F = F_0 + F_1\left(1 - \frac{[U]}{[P_t]}\right) = F_0 + F_1\left(\frac{k_{app}t}{1 + k_{app}t}\right) \quad (2)$$

where

$$k_{app} = [P_t](k_f) \quad (3)$$

In these equations, [U] and [N_2] are the concentrations of

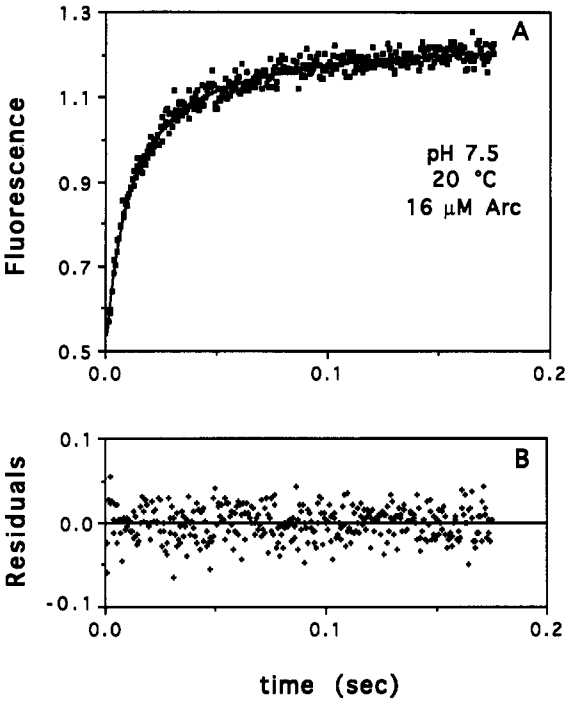

FIGURE 2: Kinetics of Arc refolding as monitored by fluorescence following a jump from pH 2 to 7.5. The solid line in panel A is the best fit for a bimolecular folding reaction (k_{app} = 90.99 s^{-1}) in which unfolding is ignored (see eqs 1–3). Panel B shows the residuals of the fit.

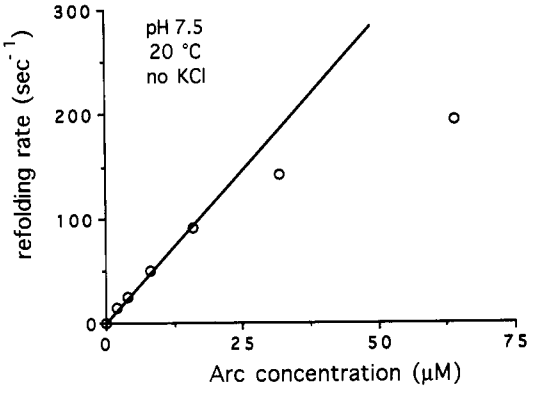

FIGURE 3: Dependence of the observed rate of refolding (k_{app}) on Arc protein concentration. The straight line is the best fit using eq 3 (with k_f = 6.2 × 10^6 M^{-1} s^{-1}) calculated from the 0–16 μM Arc data and should pertain as long as the bimolecular step is the rate-limiting step in Arc refolding. At 32 and 64 μM Arc, the refolding rate is less than expected, suggesting that, at these Arc concentrations, a unimolecular step becomes co-rate limiting for folding.

denatured and native Arc, respectively; F corresponds to the fluorescence amplitude at time t; F_0 is the fluorescence amplitude at time zero; F_1 is the amplitude change during the refolding reaction (under strongly refolding conditions; F_1 corresponds to the fluorescence change between the denatured and native states); k_{app} is the apparent rate constant for the reaction; k_f is the actual second-order rate constant; and [P_t] is the total Arc concentration. As shown by the even distribution of the residuals in Figure 2B, the Arc refolding data are fit well by eq 2.

A second prediction of the bimolecular model is that the rate of the refolding reaction should change with protein concentration. To test this, refolding experiments were carried out at 20 °C, pH 7.5, using Arc concentrations ranging from 2 to 64 μM. Figure 3 shows a plot of the refolding rate (measured by k_{app}, the reciprocal of the reaction half-time)

as a function of Arc concentration. As expected, the refolding reaction proceeds more rapidly at higher protein concentrations and more slowly at lower protein concentrations. Figure 3 also shows that the refolding rate is a linear function of Arc concentration ($r^2 = 0.99$) from 0 to 16 μM protein. This relationship is expected from eq 3. However, significant deviation from linearity occurs at higher protein concentrations, where the rate is slower than predicted. It is possible that, as the bimolecular reaction becomes faster, a unimolecular step (either before or after the bimolecular step) becomes co-rate limiting for the folding reaction. Alternatively, the kinetic parameters for high protein concentrations could be inaccurate because the dead time of mixing for the stop-flow instrument becomes significant in comparison to the time course of refolding. If the latter explanation were correct, however, then refitting of the data after removal of the first several milliseconds of data should yield a larger apparent rate constant, but this does not occur. As a result, we favor the possibility that a unimolecular step becomes co-rate limiting as the Arc concentration is raised above 16 μM.

Regardless of the reasons for the nonlinearity of the folding rate at high protein concentrations, as long as we restrict our attention to Arc concentrations at or below 16 μM, the quality of the fit of the kinetic data to the hyperbolic function (eq 2) and the linear variation of the refolding rate with protein concentration (eq 3) indicate that the bimolecular step is rate limiting for Arc folding. Moreover, the rate constant for this second-order reaction can be measured with reasonable accuracy. For example, from the data in the linear portion of Figure 3, we calculate that k_f has an average value (±sd) of 6.2 (±0.6) × 10^6 M^{-1} s^{-1} at 20 °C and pH 7.5.

Effects of Ionic Strength. Some insight into the nature of the interactions involved in the bimolecular step of Arc refolding can be gained by studying the variation of the rate constant for this step as the solvent conditions are varied. To determine the effect of salt on Arc refolding, experiments like those shown in Figure 2A were also carried out at higher KCl concentrations. These experiments were performed at pH 7.5, at 20 °C, and at Arc concentrations of 2, 4, 8, and 16 μM. The average rate constants calculated from these experiments were 9.7 (±2.6) × 10^6 M^{-1} s^{-1} for 0.1 M KCl and 12.6 (±2.6) × 10^6 M^{-1} s^{-1} for 0.25 M KCl. As described above, in the absence of KCl, the rate constant was 6.2 (±0.6) × 10^6 M^{-1} s^{-1}. Hence the rate constant for the bimolecular step of Arc refolding increases modestly with increasing ionic strength.

One interpretation of the finding that the folding reaction is faster at higher ionic strength is that salt helps to screen repulsive electrostatic interactions between Arc monomers, allowing them to collide more frequently or more efficiently. In Figure 4, $\log(k_f)$ is plotted against the square root of the ionic strength for the salt dependence data. According to Debye–Huckle theory, the slope of this plot should correspond to $1.02z^2$, where z is the effective charge involved in the reaction (Laidler, 1965). For the data shown in Figure 4, the slope is 0.61 and z is 0.78, indicating that, on average, less than one formal charge per monomer is involved in the bimolecular step of the folding reaction. Similar salt dependence experiments performed at different temperatures (see below and Table 1) give an average z value of 0.75 (±0.05).

Temperature Dependence. Refolding experiments using Arc at a concentration of 4 μM were performed at temperatures from 9 to 31 °C, and at several different KCl concentrations. As shown in Table 1, the rate constant for Arc refolding increases as the temperature is raised, for each salt concentration tested. The observed increase in the refolding rate

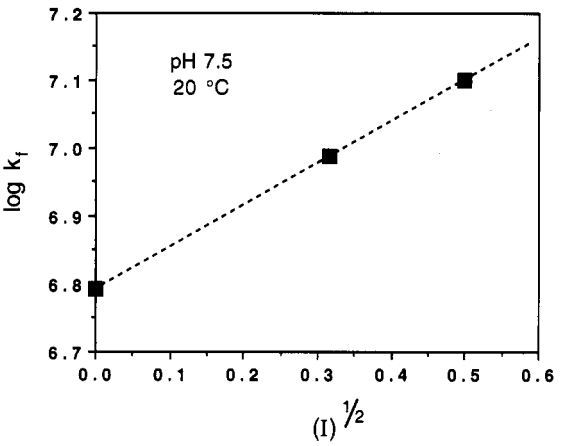

FIGURE 4: Dependence of the rate constant for Arc refolding on ionic strength. The linear relationship is expected from Debye–Huckle theory.

Table 1: Salt and Temperature Dependence of Refolding Rate Constants (k_f) (in Units of M^{-1} s^{-1}) Determined at pH 7.5, Using 4 μM Arc[a]

	9 °C	14 °C	19 °C	26 °C	31 °C
0 M KCl	1.5 × 10^6	3.1 × 10^6	4.2 × 10^6	8.0 × 10^6	11.2 × 10^6
0.1 M KCl	2.5 × 10^6	4.2 × 10^6	6.9 × 10^6	11.6 × 10^6	12.8 × 10^6
0.25 M KCl	3.0 × 10^6	5.8 × 10^6	8.9 × 10^6	14.3 × 10^6	16.0 × 10^6

[a] Values were calculated from nonlinear least-squares fitting (using eqs 2 and 3) of experiments like that shown in Figure 2.

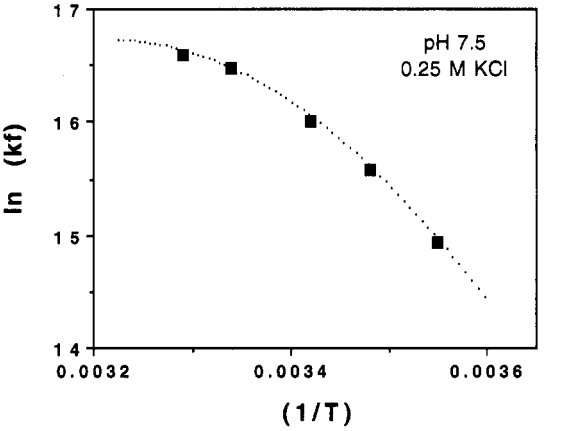

FIGURE 5: Arrhenius plot of the temperature dependence of the second-order rate constant for Arc refolding. The downward curvature of the plot is expected for a reaction that proceeds with a negative heat capacity change. The dotted line is a nonlinear least-squares fit to the equation $y = A + (B)(293/T) + (C)\ln(293/T)$ (Chen et al., 1989), with $A = 474.9$, $B = 458.8$, and $C = 437.2$.

with temperature is too great to be explained by an increase in the Arc diffusion rate. As a consequence, thermal energy appears to be required during the folding process, presumably for breaking noncovalent bonds or for bond rotation. An Arrhenius plot of the data for 0.25 M KCl is shown in Figure 5 (plots of the 0 and 0.1 M KCl data are similar). We interpret the smoothness of this plot to indicate that the bimolecular step remains rate limiting in Arc refolding in this temperature range. The curvature of the plot suggests that the activation energy (roughly 15 kcal/mol at 20 °C) changes with temperature, i.e., that there is a substantial heat capacity change during the refolding process (Chen et al., 1989). Changes in heat capacity usually accompany burial of hydrophobic surface, and some of the activation energy could be required for disruption of water organized around such surfaces prior to burial.

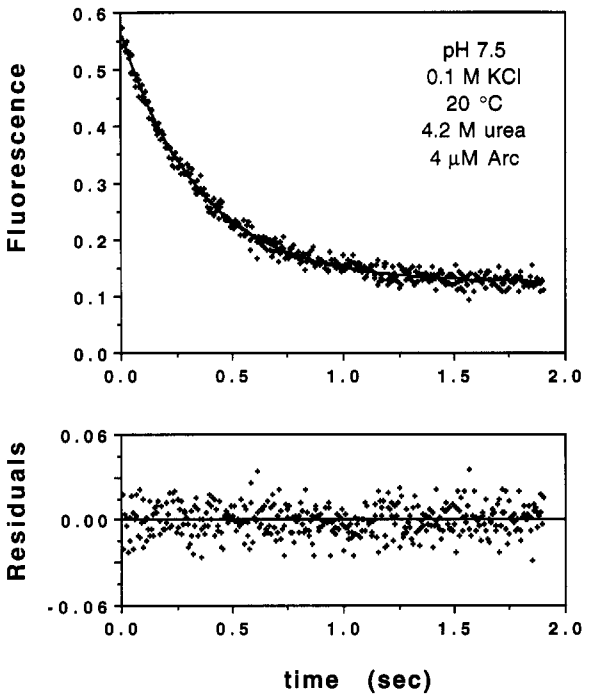

FIGURE 6: Kinetics of Arc unfolding following a jump from 0 to 4.2 M urea. The solid line in panel A is the best fit for a unimolecular unfolding reaction (k_u = 2.83 s^{-1}) in which refolding is ignored (see eqs 4 and 5). Panel B shows the residuals of the fit.

Urea Dependence of Refolding and Unfolding Rates. Rates for Arc refolding at low concentrations of urea (0–1.2 M) were studied by jumping from pH 2 into pH 7.5 buffer with urea, and analyzing the data using eq 2. To study Arc unfolding, stop-flow experiments were performed by starting in pH 7.5 buffer with no urea and jumping to higher urea concentrations. For urea concentrations above 3.3 M, the unfolding reactions are described by the simple model,

$$N_2 \xrightarrow{k_u} 2U \qquad \frac{d[U]}{dt} = 2k_u[N_2] \tag{4}$$

and the exponential function

$$F = F_0 + F_1(1 - e^{-k_u t}) \tag{5}$$

where F, F_0, and F_1 have the same definitions as in eq 2 and k_u is the unfolding rate constant. Figure 6 shows an unfolding experiment at a urea concentration of 4.2 M. Under the conditions of this experiment, the data are well fit by a single exponential and k_u is approximately 3 s^{-1}. In control experiments, we determined that the half-life of the unfolding reaction did not change as the protein concentration is changed, as expected for a unimolecular reaction.

For refolding jumps or unfolding jumps in which the final concentration of urea was 1.5–3.3 M, the rates of both the unfolding and refolding reactions are significant relative to each other and neither eq 2 (refolding) nor eq 5 (unfolding) provided a good fit of the data. The data from these experiments were analyzed using a model in which the system relaxes to equilibrium:

$$2U \underset{k_u}{\overset{k_f}{\rightleftarrows}} N_2 \tag{6}$$

In this model,

$$\frac{d[U]}{dt} = -k_f[U]^2 + 2k_u[N_2] = -\frac{k_f[P_t]}{2}\left(\frac{2}{[P_t]}[U]^2 + \frac{2k_u}{k_f[P_t]}[U] - \frac{2k_u}{k_f}\right) \tag{7}$$

Rearrangement, integration, and evaluation of the integration constant at time zero yields

$$\frac{1}{sq}\ln\left(\frac{4[U]/[P_t] + b - sq}{4[U]/[P_t] + b + sq}\right) = -\frac{k_f}{2}[P_t]t + \frac{1}{sq}\ln\left(\frac{4[U_0]/[P_t] + b - sq}{4[U_0]/[P_t] + b + sq}\right) \tag{8}$$

where

$$b = \frac{2k_u}{k_f[P_t]} \quad \text{and} \quad sq = (b^2 + 8b)^{1/2} \tag{9}$$

Equation 8 can be rearranged to give

$$[U] = \frac{[P_t]}{4}\left[\frac{-b + sq + bz + (sq)z}{1 - z}\right] \tag{10}$$

where

$$z = \left(\frac{4[U_0]/[P_t] + b - sq}{4[U_0]/[P_t] + b + sq}\right)\exp\left[\frac{-(sq)k_f t[P_t]}{2}\right] \tag{11}$$

For refolding experiments, $[U_0] \approx [P_t]$ and data were fit to

$$F = F_0 + F_1\left(1 - \frac{[U]}{[P_t]}\right) \tag{12}$$

For unfolding experiments, $[U_0] \approx 0$ and data experiments were fit to

$$F = F_0 + F_1\left(\frac{[U]}{[P_t]}\right) \tag{13}$$

In eqs 12 and 13, F_0 is the initial fluorescence of the folded or unfolded state and F_1 is the fluorescence change between the folded and unfolded states. Since [U] in eqs 13 and 14 is a function of $[P_t]$, k_u, and k_f, nonlinear least-squares fitting should, in principle, be able to determine values for the four variables F_0, F_1, k_u, and k_f. We found, however, that reproducible values for k_u and k_f could only be determined if the value of F_1 was kept constant during the fitting procedure.[1] As shown in Figure 7, values of F_1 in the urea transition zone can be determined by linear interpolation between F_1 values determined at low urea concentrations in refolding jumps and F_1 values determined at high urea concentrations in unfolding jumps.

The rate constants for Arc refolding and unfolding are plotted as a semilogarithmic function of urea concentration in Figure 8, panels A and B, respectively, for data collected at pH 7.5, 20 °C, and 0.1 M KCl. The refolding data are fit well by linear regression (r^2 = 0.99) and give a zero urea intercept of 9.0 × 10^6 M^{-1} s^{-1}, essentially the same value as measured directly in pH jump experiments. The rate constant for refolding is quite urea sensitive, decreasing about 7-fold for every 1 M increase. The unfolding data are also fit well by linear regression (r^2 = 0.98) and give a k_u value of 0.2 s^{-1} at zero urea. This corresponds to a half-life for unfolding and

[1] Reproducible values of k_u and k_f could also be obtained without fixing F_1 if the value of the equilibrium constant K_u (which fixes the ratio of k_f and k_u) was specified instead.

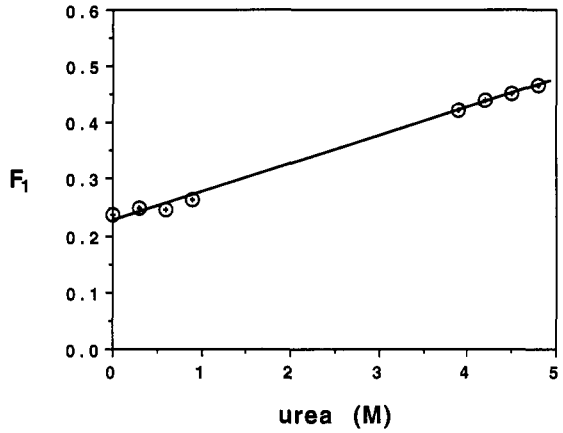

FIGURE 7: Urea dependence of F_1, the fluorescence difference between the native and denatured protein, for 4 μM Arc at 20 °C, pH 7.5, and 100 mM KCl. The linear regression line (r^2 = 0.995) has a y-intercept of 0.29 and a slope of 0.049. Values of F_1 were obtained either from refolding jumps from pH 2 into pH 7.5 buffer containing 100 mM KCl and 0–0.9 M urea (fitted to eq 2) or from unfolding jumps from pH 7.5 buffer containing 100 mM KCl into the same buffer containing 3.9–4.8 M urea (fitted to eq 5).

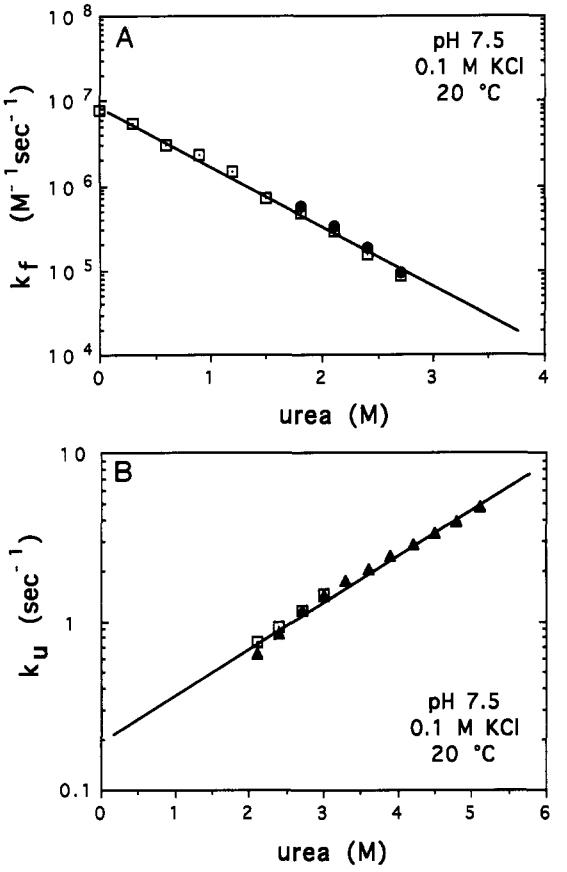

FIGURE 8: Urea dependence of the rate constants for Arc refolding (A) and unfolding (B). Open symbols represent data from refolding experiments. Closed symbols represent data from unfolding experiments. The solid lines are best fits from linear regression with r^2s of 0.992 (A) and 0.988 (B). The zero urea intercepts are 9.05×10^6 M^{-1} s^{-1} (A) and 0.20 s^{-1} (B). Refolding and unfolding experiments were performed at 4 μM Arc, 20 °C, pH 7.5, and 100 mM KCl. Refolding experiments were fit to eq 2 for data from 0 to 1.2 M urea and to eq 12 for data from 1.5 to 2.7 M urea. Unfolding experiments were fit to eq 5 for data from 3.6 to 5.2 M urea and to eq 13 for data from 2.1 to 3.3 M urea.

dissociation of the Arc dimer of approximately 3.5 s under roughly physiological conditions. As we discuss later, this is remarkably fast when compared with many other protein

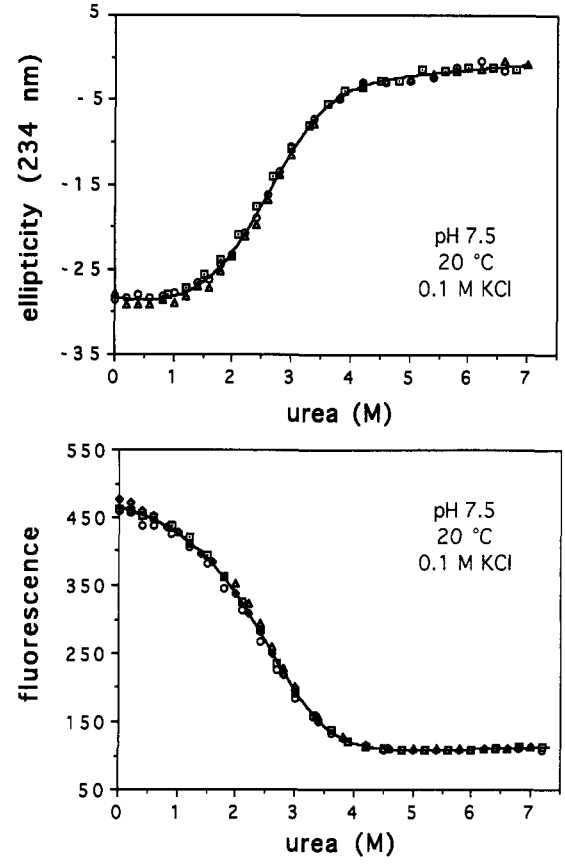

FIGURE 9: Equilibrium denaturation of Arc monitored by circular dichroism (top) and fluorescence (bottom). Each plot includes data points from 3–4 independent experiments performed using 20 μM Arc, pH 7.5, 20 °C, and 100 mM KCl. The solid lines are nonlinear least-squares fits with ΔG_u's of 9.52 kcal/mol (top) and 9.58 kcal/mol (bottom), and m values of −1.39 kcal/(mol·M) (top) and −1.37 kcal/(mol·M) (bottom). The standard deviation of the fits are 0.51 (top) and 5.57 (bottom).

unfolding reactions. The unfolding rate constant for Arc is considerably less urea sensitive than the refolding rate constant, increasing by a factor of roughly 2.7 for each 1 M increase in urea. The linearity of the Figure 8 plots suggests that there is no change in the rate-determining step for refolding or unfolding as the urea concentration is raised. Similar results were obtained when the urea dependences of refolding and unfolding were measured at pH 7.5, 20 °C, and 0.25 M KCl (data not shown). Under these conditions, the zero urea intercept was 1.2×10^7 M^{-1} s^{-1} for k_f and 0.1 s^{-1} for k_u (half-life 7 s).

Several observations indicate that the two-state model of eq 6 provides a good kinetic description of both refolding and unfolding in the Arc system. First, at some points in the transition zone of Figure 8, k_u and k_f can be accurately determined from both refolding and unfolding experiments and are in close agreement as expected from the two-state model. Second, if the two-state model is valid, then the kinetic constants will be related to the equilibrium constant for Arc denaturation, since

$$K_u = \frac{2k_u}{k_f} \tag{14}$$

Equilibrium denaturation of Arc was monitored by fluorescence and by circular dichroism as shown in Figure 9. The experimentally derived values of K_u are listed in Table 2, as are those expected from the kinetic results. The predicted and experimentally determined values of K_u are identical within

Table 2: K_u, ΔG_u, and m Values from Equilibrium and Kinetic Experiments[a]

method	K_u (M)	ΔG_u (kcal/mol)	m [kcal/(mol·M)]
	20 °C, pH 7.5, 100 mM KCl		
equilibrium CD	7.4×10^{-8}	9.52 (±0.21)	−1.39 (±0.12)
equilibrium F	6.7×10^{-8}	9.58 (±0.28)	−1.37 (±0.10)
kinetics	4.4×10^{-8}	9.82	−1.31
	20 °C, pH 7.5, 250 mM KCl		
equilibrium CD	1.1×10^{-8}	10.61 (±0.35)	−1.47 (±0.13)
equilibrium F	1.2×10^{-8}	10.58 (±0.43)	−1.43 (±0.10)
kinetics	1.7×10^{-8}	10.36	−1.31

[a] Error limits for the equilibrium experiments are from the nonlinear least-squares fitting procedure. Values of m_f and m_u were −0.95 and 0.36, respectively, for the 100 mM KCl data and −0.97 and 0.34, respectively, for the 250 mM KCl data.

error limits. Finally, the two-state model relates the m value for equilibrium unfolding (the slope of ΔG_u vs urea) to the m values for the kinetic experiments:

$$m = \frac{-\mathrm{d}[RT\ln(K_u)]}{\mathrm{d[urea]}} = \frac{-\mathrm{d}[RT\ln(k_u)]}{\mathrm{d[urea]}} + \frac{\mathrm{d}[RT\ln(k_f)]}{\mathrm{d[urea]}} = m_u - m_f \quad (15)$$

The m values from the equilibrium denaturation experiments are between 1.37 and 1.47, while those predicted from the kinetic results are 1.31 for both salt concentrations. The error limits on all of these m values are about 10%.

DISCUSSION

The diffusion-controlled upper limit for collision of two Arc monomers can be calculated to be 9.6×10^8 M^{-1} s^{-1} using von Smoluchowski's equation (Eisenberg & Crothers, 1979). This calculation assumes a reaction radius of 5 Å and a diffusion constant of 1.2×10^{-6} cm^2 s^{-1} at 20 °C. The calculated diffusion limit is slightly smaller, 7×10^8 M^{-1} s^{-1}, when the Debye correction for electrostatic interactions is included (assuming $z^2 = 0.57$). We find that the rate constant for Arc folding and dimerization ranges from 5×10^6 M^{-1} s^{-1} to 2×10^7 M^{-1} s^{-1} under roughly physiological conditions. This indicates that productive interactions between Arc monomers occur at a rate roughly 100-fold slower than the diffusion limit. Nevertheless, the concerted folding and dimerization of Arc are still fast on a relative time scale as bimolecular rate constants for protein–protein interactions rarely exceed 10^7 M^{-1} s^{-1}, even between proteins that are already stably folded (Koren & Hammes, 1976; Raman et al., 1992).

How do the rates of Arc folding and dimerization compare with those of other protein folding reactions? The rate of Arc refolding varies with protein concentration, but at 10 μM, the half-time of refolding is less than 20 ms. This is as fast as the folding of monomeric proteins of similar size such as chymotrypsin inhibitor 2, hen lysozyme, and cytochrome c, and faster than the folding of barnase and T4 lysozyme (Matouschek et al., 1990; Jackson & Fersht, 1991; Elöve et al., 1992; Lu & Dahlquist, 1992; Radford et al., 1992). Thus, two Arc monomers can find each other and fold at rates comparable to the unimolecular folding of many monomeric proteins. It is worth noting, however, that fast refolding is not a general property of dimeric proteins. For example, the Trp repressor refolds 100-fold to 1000-fold more slowly than Arc (Gittelman & Matthews, 1990).

The half-life for unfolding of the Arc dimer is approximately 5 s under the conditions tested. This unfolding rate is very fast in comparison to unfolding rates measured for many other proteins. For example, DHFR and barnase have half-lives of several hundred seconds (Perry et al., 1987; Matouschek et al., 1989); the Trp synthase α-subunit and chymotrypsin inhibitor 2 have half-lives greater than 1 h (Beasty et al., 1986; Jackson & Fersht, 1991); and Trp repressor and human lysozyme have half-lives of several days or more (Gittelman & Matthews, 1990; Taniyama et al., 1992).

On the basis of the kinetic constants measured here, we would expect that free Arc would constantly be unfolding and refolding in the cell. This challenges a commonly held view that proteins only need to fold during or shortly after synthesis and perhaps following heat shock or some other type of cellular stress. Does Arc folding in the cell require molecular chaperones? Rothman (1989) has argued that chaperones may be especially important for the folding of dimers because some dimeric proteins appear incapable (on the basis of their folding rates *in vitro*) of folding sufficiently fast in the cell. It would be difficult, however, to argue that chaperones could accelerate Arc refolding to a point faster than the rates measured *in vitro*. The Arc refolding reaction already occurs at a rate comparable to the fastest protein–protein association rates, and since any chaperone-mediated Arc refolding reaction would add an additional association step, it would almost certainly be slower than the unassisted refolding reaction. Although definitive proof is lacking, there are several indications that Arc folding in the cell does not involve molecular chaperones: (i) overexpression of wild-type Arc or an unstable variant (SA32) fails to induce the expression of heat shock proteins such as GroEL and DnaK; and (ii) wild-type Arc and unstable variants are not found as complexes with cellular chaperones or other cellular proteins in crude cell lysates (unpublished data).

The comments made above should not be taken to indicate that Arc folds and unfolds frequently in the cell with little or no consequences. In fact, many Arc mutants are highly susceptible to intracellular degradation, and even wild-type Arc is degraded with a half-life of approximately 30 min (Vershon et al., 1986; Bowie & Sauer, 1989c). Because the dynamics of Arc unfolding and refolding are fast compared with this half-life, however, it seems likely that intracellular degradation of Arc is not under kinetic control of folding or unfolding but is under thermodynamic control instead. On average, a few percent of Arc molecules are probably unfolded in the cell at any given time and, thus are at risk for proteolysis.

Both the urea and temperature dependences of Arc refolding indicate that hydrophobic interactions play an important role in the rate-limiting bimolecular step of folding. The refolding rate is very urea sensitive and accounts, in fact, for about 75% of the urea dependence of the equilibrium constant. The urea dependence of unfolding accounts for only 25% of the equilibrium effect. This indicates that considerably more hydrophobic surface is buried between the unfolded state and the transition state than between the native state and the transition state. Similar findings have been obtained in folding studies of monomeric proteins such as T4 lysozyme and chymotrypsin inhibitor 2 (Chen et al., 1989; Jackson & Fersht, 1991). The temperature dependence of Arc refolding indicates a requirement for bond breaking and suggests that a substantial heat capacity change accompanies folding. Both findings are consistent with the burial of a considerable amount of hydrophobic surface between the unfolded ground state and the folding transition state.

The Arc refolding reaction shows a modest salt dependence. If we assume that this occurs because of screening of repulsive

interactions between Arc monomers, then each monomer would have an effective charge of roughly 0.75. On the basis of the number of acidic and basic residues, the Arc monomer would be expected to have a net charge of +4 at pH 7.5. We imagine that the effective repulsion between monomers is less than would be expected on the basis of net charge alone, because the charges are distributed throughout the native structure (Breg et al., 1990). The equilibrium stability of Arc is also salt dependent, and it has been proposed that the binding of two anions stabilizes the Arc dimer against denaturation (Bowie & Sauer, 1989a). The salt dependence of the refolding rate appears to be too small in magnitude to support a model in which ions participate directly in the bimolecular step of Arc folding, but binding of one or more anions could occur after the rate-limiting step.

Under most of the conditions tested here, the refolding of the Arc dimer from two unfolded monomers appears to be two-state, suggesting that folding intermediates are poorly populated. Similar two-state behavior has been observed for refolding of the gene V protein dimer (Liang & Terwilliger, 1991). On the basis of structural considerations, however, it seems unlikely that Arc goes from two unfolded chains to a folded dimer in a single kinetic step. The structure of the Arc dimer (Breg et al., 1990) shows monomers that intertwine and interact over much of the sequence by hydrophobic packing between side chains and by hydrogen bonding between main-chain amides in a region of antiparallel β-sheet. As a consequence, monomers having the same conformation seen in the Arc dimer would not be expected to be stable in solution, as they would lack most of the hydrophobic and hydrogen-bonding interactions that appear to stabilize the structure. Hence, folding is not likely to be a simple process in which monomers fold to their proper conformations and then dock to form the dimer. Rather, monomers probably collide in an partially folded state (perhaps a molten globule) to form a dimeric intermediate, which then proceeds to form the native dimer via one or more unimolecular folding steps. A minimal model for such a folding pathway would be

$$2U \xrightarrow{k_1} 2I_m \xrightarrow{k_2} I_2 \xrightarrow{k_3} N_2 \qquad (16)$$

where I_m represents a partially folded monomeric intermediate and I_2 represents a partially folded dimeric intermediate. We emphasize that this model is based upon structural considerations and not upon kinetic evidence. There are, however, no inconsistencies between the observed kinetic results and this model. Moreover, although we have no kinetic evidence for I_m, there is evidence that pressure-denatured Arc forms a compact, molten globule state (Silva et al., 1992; Peng et al., 1993).

We have shown that the rate-limiting step in Arc refolding is bimolecular under most conditions. This could correspond to the k_2 step in the model above. Since unimolecular steps are not rate limiting (at least at Arc concentrations below 16 μM), we can estimate that any such step present in the folding pathway must have a rate constant faster than 100 s^{-1}. At higher Arc concentrations, the bimolecular step becomes fast enough that a unimolecular step appears to begin to limit the refolding rate. This unimolecular step could correspond to either the k_1 or k_3 step. At present, we have only followed the kinetics of folding by monitoring the fluorescence of the single tryptophan in Arc, which is buried in the dimer interface of the native protein. Other structural probes such as circular dichroism or fluorimetry in the presence of bisANS may be useful in revealing intermediates in the folding reaction. Kinetic studies of mutant proteins should also be useful in dissecting the mechanism of folding and/or in identifying interactions that are important in stabilizing the initial dimeric intermediate.

ACKNOWLEDGMENT

We thank Alan Davidson, Evan Kantrowitz, Peter Kim, Paul Schimmel, Kevin Shoemaker, JoAnne Stubbe, and Wei Wu for advice, help, and the use of instruments.

REFERENCES

Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T., Onuffer, J. J., & Matthews, C. R. (1986) *Biochemistry 25*, 2965–2974.

Blond, S., & Goldberg, M. E. (1986) *Proteins: Struct., Funct., Genet. 1*, 247–255.

Bothwell, M. A., & Schachman, H. K. (1980a) *J. Biol. Chem. 255*, 1962–1970.

Bothwell, M. A., & Schachman, H. K. (1980b) *J. Biol. Chem. 255*, 1971–1977.

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.

Bowie, J. U., & Sauer, R. T. (1989c) *J. Biol. Chem. 264*, 7596–7602.

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., Bowie, J. U., & Sauer, R. T. (1990) *Biochemistry 29*, 11189–11195.

Brown, B. M., Milla, M. E., Smith, T. L., & Sauer, R. T. (1994) *Nature* (submitted for publication).

Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G., & Gallagher, D. T. (1992) *Biochemistry 31*, 4937–4945.

Burns, D. L., & Schachman, H. K. (1982a) *J. Biol. Chem. 257*, 8638–8647.

Burns, D. L., & Schachman, H. K. (1982b) *J. Biol. Chem. 257*, 8648–8654.

Chen, B., Baase, W. A., & Schellman, J. A. (1989) *Biochemistry 28*, 691–699.

Eisenberg, D., & Crothers, D. (1979) *Physical Chemistry with Applications to the Life Sciences*, Benjamin Cummings, Inc., Menlo Park, CA.

Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. E. (1992) *Biochemistry 31*, 6876–6883.

Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011–7020.

Goldenberg, D., & King, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3403–3407.

Herman, R., Jaenicke, R., & Rudolph, R. (1981) *Biochemistry 20*, 5195–5201.

Herman, R., Rudolph, R., Jaenicke, R., Price, N. C., & Scobbie, A. (1983) *J. Biol. Chem. 258*, 11014–11019.

Jackson, S. E., & Fersht, A. R. (1991) *Biochemistry 30*, 10436–10443.

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.

Johnson, M., & Frasier, S. (1985) *Methods Enzymol. 117*, 301–342.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Koren, R., & Hammes, G. G. (1976) *Biochemistry 15*, 1165–1171.

Krebs, H., Rudolph, R., & Jaenicke, R. (1979) *Eur. J. Biochem. 100*, 359–364.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Laidler, K. J. (1965) *Chemical Kinetics*, McGraw-Hill, Inc., New York.

Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and*

Polymer Science (Harding, S., Rowe, A., and Horton, J., Eds.) pp 90–125, Royal Society of Chemistry, Cambridge.
Le Bras, G., Teschner, W., Deville-Bonne, D., & Garel, J. R. (1989) *Biochemistry 28*, 6836–6841.
Liang, H., & Terwilliger, T. C. (1991) *Biochemistry 30*, 2772–2782.
Lu, J., & Dahlquist, F. W. (1992) *Biochemistry 31*, 4749–4756.
Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32*, 5282–5290.
Matouschek, A., Kellis, J. J., Serrano, L., & Fersht, A. R. (1989) *Nature 340*, 122–126.
Matouschek, A., Kellis, J. J., Serrano, L., Bycroft, M., & Fersht, A. R. (1990) *Nature 346*, 440–445.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653–683.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.
Peng, X., Jonas, J., & Silva, J. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1776–1780.
Perry, K. M., Onuffer, J. J., Touchette, N. A., Herndon, C. S., Gittelman, M. S., Matthews, C. R., Chen, J.-T., Mayer, R. J., Taira, K., Benkovic, S. J., Howell, E. E., & Kraut, J. (1987) *Biochemistry 26*, 2674–2682.
Radford, S. E., Dobson, C. M., & Evans, P. A. (1992) *Nature 358*, 302–307.
Raman, C. S., Jemmerson, R., Nall, B. T., & Allen, M. J. (1992) *Biochemistry 31*, 10370–10379.
Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994) *Nature* (submitted for publication).
Rothman, J. E. (1989) *Cell 59*, 591–601.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063–8068.
Sauer, R. T., Krovatin, W., DeAnda, J., Youderian, P., & Susskind, M. M. (1983) *J. Mol. Biol. 168*, 699–713.
Silva, J. L., Silveira, C. F., Correia, A., Jr., & Pontes, L. (1992) *J. Mol. Biol. 223*, 545–555.
Taniyama, Y., Ogasahara, K., Yutani, K., & Kikuchi, M. (1992) *J. Biol. Chem. 267*, 4619–4624.
van Holde, K. E. (1971) *Physical Biochemistry*, pp 110–113, Prentice-Hall, Englewood Cliffs, NJ.
Vaucheret, H., Signon, L., Le Bras, G., & Garel, J. R. (1987) *Biochemistry 26*, 2785–2790.
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.
Villafane, R., & King, J. (1988) *J. Mol. Biol. 204*, 607–619.