# Observation of Multistate Kinetics during the Slow Folding and Unfolding of Barstar<sup>†</sup>

Abani K. Bhuyan and Jayant B. Udgaonkar\*

National Centre for Biological Sciences, UAS-GKVK Campus, Bangalore 560065, India

Received February 5, 1999; Revised Manuscript Received April 26, 1999

ABSTRACT: The kinetics of the slow folding and unfolding reactions of barstar, a bacterial ribonuclease inhibitor protein, have been studied at  $23(\pm 1)$  °C, pH 8, by the use of tryptophan fluorescence, far-UV circular dichroism (CD), near-UV CD, and transient mixing <sup>1</sup>H nuclear magnetic resonance (NMR) spectroscopic measurements in the 0-4 M range of guanidine hydrochloride (GdnHCl) concentration. The denaturant dependences of the rates of folding and unfolding processes, and of the initial and final values of optical signals associated with these kinetic processes, have been determined for each of the four probes of measurement. Values determined for rates as well as amplitudes are shown to be very much probe dependent. Significant differences in the intensities and rates of appearance and disappearance of several resolved resonances in the real-time one-dimensional NMR spectra have been noted. The NMR spectra also show increasing dispersion of chemical shifts during the slow phase of refolding. The denaturant dependences of rates display characteristic folding chevrons with distinct rollovers under strongly native as well as strongly unfolding conditions. Analyses of the data and comparison of the results obtained with different probes of measurement appear to indicate the accumulation of a myriad of intermediates on parallel folding and unfolding pathways, and suggest the existence of an ensemble of transition states. The energetic stabilities of the intermediates estimated from kinetic data suggest that they are approximately half as stable as the fully folded protein. The slowness of the folding and unfolding processes ( $\tau = 10-$ 333 s) and values of 20.5 ( $\pm$ 1.4) and 18 ( $\pm$ 0.5) kcal mol<sup>-1</sup> for the activation energies of the slow refolding and unfolding reactions suggest that proline isomerization is involved in these reactions, and that the intermediates accumulate and are therefore detectable because the slow proline isomerization reaction serves as a kinetic trap during folding.

The diversity of the processes by which proteins can fold to their native states, for example, fast or slow folding in a two-state manner (1, 2) or in multiple kinetic phases with and without intermediates (3-5), casts doubt if all proteins fold via specific pathways in a well-determined manner as held classically (6, 7). During the past decade a new view of folding, which is based on the statistical characterization of the energy landscape of folding proteins, has emerged (8-12) that appears to capture the diversity of experimentally observed protein folding kinetics. Validation of this approach requires sufficient data over a range of conditions so that experimental folding work can be put into the context of the conceptual framework of the model.

The fundamental issue of protein folding, that of the problem of sampling out all possible polypeptide conformations, is poorly understood. It is generally believed that polypeptides avoid exhaustive sampling of conformations by undergoing an initial random collapse (13, 14). The collapsed structures then transform to the native state through a diffusive process. It is important to decipher the physical

forces and chain dynamics that are involved as well as the time needed for polypeptide collapse and for fast chain organization reactions within the collapsed species. Thus, much experimental work has focused on the millisecond and submillisecond processes of folding (15, 16). On the other hand, kinetics in the seconds to minutes time-scale have received meager attention, perhaps because they are often thought to arise from proline isomerization reactions or from protein—protein interactions.

In this article, we describe folding and unfolding kinetics in the seconds to minutes time domain to show that even slow phases of kinetics can provide a wealth of information related to chain organization and transition states of protein folding, in a manner consistent with the predictions of the landscape energy model, even though proline isomerization would appear to be inextricably associated with late structural events. We have used barstar, a bacterial ribonuclease inhibitor protein of 89 amino acids, for which an initial nonspecific hydrophobic collapse of the random polypeptide and some of the millisecond chain organization events within the collapsed structure have been described earlier (17). The folding and unfolding kinetics of barstar have been the subjects of several previous reports (refs 5, 17-22, for example).

In this study, tryptophan fluorescence, far-UV CD, near-UV CD, and <sup>1</sup>H NMR methods have been used to record

<sup>&</sup>lt;sup>†</sup>This work was funded by the Tata Institute of Fundamental Research and the Department of Biotechnology, Government of India. J.B.U. is the recipient of a Golden Jubilee Biotechnology Fellowship and a Swarnajayanti Fellowship from the Government of India.

<sup>\*</sup> Corresponding author. Fax: 91-80-8561662. E-mail: jayant@ncbs.res.in.

the kinetics of the slow phase of folding as well as unfolding by manually placing the protein in different concentrations of GdnHCl and at different temperatures. The dependences on GdnHCl concentration of the amplitude and the rate constant of the slow observable kinetic phase were measured. In addition, changes in the values of all the spectroscopic signals which occur in the unobservable burst phase (during the manual mixing dead time of 10 s) were measured as burst phase amplitudes as a function of GdnHCl concentration. Our results reveal the occurrence of a myriad of intermediates which appear to accumulate on multiple competing pathways, and, therefore, the occurrence of an ensemble of rate-limiting transition states.

### **MATERIALS AND METHODS**

Details of protein overexpression and the procedure for protein purification have been described (23). All D<sub>2</sub>O and H<sub>2</sub>O buffer solutions contained 20 mM sodium phosphate, pH 8, 300  $\mu$ M EDTA, and 250  $\mu$ M DTT. The reported pH value of the D<sub>2</sub>O buffer is the uncorrected pH meter reading. For NMR experiments, GdnHCl was deuterated by repeated lyophilization of its D<sub>2</sub>O solution. All optical experiments were performed at room temperature (23 ± 1 °C). NMR spectra were recorded at 24 °C.

Measurement and Analysis of Equilibrium Unfolding Data. Barstar solutions (3  $\mu$ M, 20  $\mu$ M, and ~80  $\mu$ M protein for fluorescence, far-UV CD, and near-UV CD measurements, respectively) containing different concentrations of GdnHCl were incubated at room temperature for  $\sim 12$  h. Tryptophan fluorescence excited at 287 nm was measured at 320 nm using a square quartz cell in a photon counting instrument (SPEX 320). For CD measurements at 222 and 270 nm, a Jasco J720 spectropolarimeter was used. After correcting for buffer background signals, each set of equilibrium data was scaled with respect to the signal of the corresponding native protein. The baselines in the pre- and post-transition regions, estimated by linear extrapolation, were assumed to be dependent linearly on GdnHCl concentration. The data were least-squares-fitted to a two-state N=U model as described previously (18):

$$S_{\rm obs} = \frac{S_{\rm f} + m_{\rm f}[D] + S_{\rm u} + m_{\rm u}[D] \exp\left(\frac{-\Delta G + m_{\rm G}[D]}{RT}\right)}{1 + \exp\left(\frac{-\Delta G + m_{\rm G}[D]}{RT}\right)}$$
(1)

where  $S_{obs}$  is the observed signal,  $S_f$  and  $S_u$ , and  $m_f$  and  $m_u$  represent intercepts and slopes of native and unfolded baselines, respectively,  $m_G$  is a parameter related to exposure of hydrophobic groups upon global unfolding of the protein, and [D] represents the concentration of GdnHCl. The baselines are assumed to depend linearly on GdnHCl concentration.

*Optical Measurements and Analysis of Folding–Unfolding Kinetics.* Folding was initiated manually by diluting a constant volume of the unfolded protein solution (prepared in buffered 6 M GdnHCl) into a constant volume of the refolding buffer containing variable amounts of GdnHCl. Thus, the final protein concentration in the refolding medium

was held constant:  $\approx 4 \ \mu$ M,  $\approx 60 \ \mu$ M, and  $\approx 37.5 \ \mu$ M for fluorescence, far-UV CD, and near-UV CD measurements, respectively. The optical signals of the native and of the unfolded protein at appropriate concentrations were noted to facilitate normalization of the data. The measurement dead-time was  $\sim 10$  s. Refractive indices of the final solutions were read to calculate the concentration of GdnHCl in the refolding medium. The time of measurement (2–2100 s) and the number of traces averaged, typically 2–5, depended on the denaturant concentration in the folding solution. Unfolding measurements were carried out similarly.

The averaged traces were fitted to single-exponential functions to obtain the folding (unfolding) rates,  $\lambda_1$ , the amplitude of the total signal change during the reaction, and signal values at t = 0 ( $S_0^{\text{fol}}$  or  $S_0^{\text{unfol}}$ ) and at  $t = \infty$  ( $S_\infty^{\text{fol}}$  or  $S_\infty^{\text{unfol}}$ ). The  $S_0^{\text{fol,unfol}}$  and  $S_\infty^{\text{fol,unfol}}$  values were normalized with reference to the fluorescence or ellipticity signals of the native protein. The observed signal amplitude at a particular GdnHCl concentration was converted to a reduced amplitude by dividing the observed amplitude by the difference in the values for the fully folded and unfolded proteins at that GdnHCl concentration. Conversion to reduced amplitudes allows the GdnHCl concentration dependences of the amplitudes of the slow kinetic measurements obtained with different probes to be compared directly.

Real-Time NMR Measurement of Folding-Unfolding Kinetics. The method used for real-time NMR measurement of folding-unfolding kinetics is the same as that used previously for measurement of fast hydrogen exchange rates (24, 25). Briefly, one part of the protein solution and nine parts of the buffer, the two contained separately in gastight syringes, are injected simultaneously into the NMR tube positioned inside the magnet. The solutions are delivered via two Teflon flow lines.

The present set of experiments was performed in  $D_2O$ solutions. The reactions were initiated by mixing 60  $\mu$ l of  $\sim$ 8 mM native or unfolded protein solution (prepared in  $\sim$ 5.5 M deuterated guanidine hydrochloride, GdnDCl) with 540  $\mu$ l of the buffer containing the desired concentration of the denaturant, directly in the NMR tube. The dead-time of measurement was  $\sim$ 7 s. Following injection of the solutions, 512 FIDs were collected in the arrayed mode. The number of scans per FID, 2 or 4, depended on the concentration of GdnDCl in the final medium. The kinetic preacquisition delay (the parameter PAD in the VNMR software of Varian), which is the delay between successive FIDs, was set to zero. The spectra were of 4K data points over a spectral width of 9000 Hz. The delay before the 90° pulse was 0.772 s. Since the pulse width increases at higher concentrations of the denaturant in the medium, it was calibrated at each concentration of GdnHCl using protein samples at equilibrium prepared under matching solution conditions. Other acquisition parameters were held constant, although the  $T_1$  relaxation times ( $\sim 0.5-0.6$  s in the native protein) are expected to be longer at higher concentrations of the denaturant. All spectra were recorded at 24 °C in a Varian Unity plus spectrometer operating at a <sup>1</sup>H frequency of 600.051 MHz. Data were processed with a line broadening of 6 Hz using Felix 95 (BIOSYM) and VNMR. Intensities of resolved resonances were determined by the use of the Lorentzian curve-fitting routine of VNMR.



FIGURE 1: Representative manual-mixing kinetic traces of folding (a) and unfolding (b) of barstar at 23  $(\pm 1)$  °C in 20 mM phosphate buffer, pH 8, containing 300  $\mu$ M EDTA and 250  $\mu$ M DTT. The final concentrations of GdnHCl in the folding and unfolding medium are 0.8 and 2.1 M, respectively. (O) Far-UV CD (222 nm); ( $\bullet$ ) fluorescence (320 nm); and ( $\Box$ ) near-UV CD (270 nm). Kinetics were recorded long enough for the reaction to reach equilibrium. For illustration, the data after 600 s have been truncated.

## RESULTS

Optically Monitored Slow Folding and Unfolding. Figure 1 shows a few representative traces illustrating the kinetics of changes in signals of tryptophan fluorescence, far-UV CD, and near-UV CD during the folding and unfolding reactions of barstar. The kinetics of folding in 0.8 M GdnHCl when monitored by the three optical probes (Figure 1a) appear distinct in terms of rates, initial signal intensities, and observable amplitudes. For all three optical probes, the major fraction of the optical change occurs in the initial unobservable burst phase of 10 s. Only the final amplitude of the burst phase change in signal intensity,  $S_0^{\text{fol}}$ , can be determined, not its rate. Under strongly refolding conditions (<0.4 M GdnHCl), the kinetics are fastest and the observable amplitudes smallest when measured by far-UV CD, and the kinetics are slowest and the observable amplitudes largest when measured by near-UV CD. This result is in agreement with an earlier study (17) of the folding of barstar in 1 M GdnHCl at pH 7.

The kinetics of unfolding measured by the three optical probes are also heterogeneous, albeit to a lesser extent. In Figure 1b, it is seen that the rates as well as the observable amplitudes are distinct when unfolding in 2.1 M GdnHCl is monitored by the three probes. Similar burst phase unfolding amplitudes are exhibited by all three probes.

*NMR-Monitored Slow Folding and Unfolding*. The slow phases of folding and unfolding of barstar have provided an opportunity to monitor the time-dependent changes of chemical shifts and resonance intensities of the resolved



FIGURE 2: Aromatic region of 600 MHz real-time <sup>1</sup>H NMR spectra of barstar at 24 °C in 20 mM phosphate buffer, pH 8, containing ~300  $\mu$ M EDTA and ~250  $\mu$ M DTT. The spectra were recorded during (a) refolding by rapidly diluting the protein solution containing 6 M GdnDCl into a D<sub>2</sub>O buffer to yield a final denaturant concentration of 1 M; and (b) unfolding of the native protein, initially dissolved in the D<sub>2</sub>O buffer, by diluting into a buffer containing 2 M GdnDCl to yield a final denaturant concentration of 1.8 M.

protons. Rapid mixing <sup>1</sup>H and <sup>19</sup>F NMR experiments useful for protein folding and unfolding have been described previously (26-30).

Figure 2 presents an expanded view of the evolution of the aromatic region of the NMR spectrum of barstar during refolding and unfolding at 24 °C in the presence of 1 and 1.8 M GdnDCl, respectively. These representative spectra demonstrate that the spectra recorded were of sufficiently good quality to provide sufficient resolution of kinetic parameters without large errrors. In Figure 2a, the initial spectrum is characterized by relatively less resolved resonances, lower peak intensities, and poorly dispersed chemical shifts. The resonances grow in intensity, and the dispersion of chemical shifts evolves as nativelike structure is formed during folding. Conversely, unfolding is characterized by decaying intensities of resonances of the folded conformation accompanied by development of degeneracy of chemical shifts (Figure 2b). The evolution of the F56 H<sup> $\xi$ </sup> resonance during folding and unfolding reactions (Figure 2a,b) exemplifies these spectral changes. In Figure 2b, the peak intensities do not decay to zero, because the protein has been unfolded using a concentration of GdnDCl corresponding to one in the unfolding transition region. In these experiments, no significant line broadening or shifts of resolved



FIGURE 3: Illustrative time dependences of change of NMR intensities during folding and unfolding of barstar under the conditions stated in the caption to Figure 2. The intensity of any resonance during folding is normalized with respect to the final intensity of that resonance at the end of 1300 s of folding, under the folding conditions used (in 1 M GdnHCl in this case). The intensities during unfolding were normalized by dividing the observed values by the intensity of the resonance in the native state. The solid lines through the data represent single-exponential fits.

resonances (F56 H<sup> $\epsilon$ </sup> and F56 H<sup> $\delta$ </sup>, for example) were detected, indicating that the rates of interconversion between different folded and unfolded conformations of barstar under these conditions and within the observable time window are slow relative to frequency differences of the resonances in these conformational states.

Figure 3 illustrates the time courses of intensity changes for a few of the resolved resonances during folding and unfolding of barstar. The rates of appearance of different resonances are different. There is also a range of folding burst phase changes in resonance intensities: the recorded intensity at the end of the unobservable 10 s burst phase  $(S_0^{\text{fol}})$  is different for various resonances. In NMR-monitored kinetics, the burst phase change in resonance intensity may arise from different chemical shifts of a resonance before the earliest time of interrogation, or from chemical exchange effects due to exchange of the resonance between two or more environments. The chemical shift heterogeneity and exchange processes may also affect the observable rate of transition of an intermediate to the folded or unfolded states. When exchange effects are absent during the observed course of kinetics, the rate calculated from the resonance area and the resonance intensity should be comparable. In the present study, the rates obtained from line heights and line areas for the resolved resonances were roughly comparable. An experimental demonstration of the discrepancy of refolding rates calculated from resonance heights and resonance integrals for certain tryptophan side chain resonances of *E. coli* dihydrofolate reductase is found in the work of Frieden and co-workers (27).

The burst phase change in resonance intensity varies from  $\sim 11\%$  for F56 H<sup> $\delta$ </sup> to  $\sim 50\%$  for E52 H<sup> $\alpha$ </sup> and I5 H<sup> $\gamma$ 2</sup>. This result is consistent with the observation made in optically monitored folding experiments in 0.8 M GdnHCl (Figure 1a), where near-UV CD reports only  $\sim 65\%$  of the change in total signal as missing in the burst phase, whereas far-UV CD associates  $\sim 90\%$  of the total signal with burst phase kinetics.

Burst phase changes in the intensities of various resonances are notable during unfolding as well. As seen in Figure 3, different resonances undergo burst phase changes during unfolding to different extents. The resonances also display a dispersion in the rate constants for changes in intensities during the observable slow phase.

Kinetic vs Equilibrium Amplitudes of Slow Folding and Unfolding. Figure 4 presents the [GdnHCl] dependences of the values of  $S_0^{\text{fol}}$  and  $S_{\infty}^{\text{fol}}$ , the signal intensities at zero and infinite times, respectively, obtained from kinetic data along with equilibrium unfolding profiles of barstar monitored by all three optical probes. For each set of kinetic measurements, the  $S_{\infty}^{\text{fol}}$  values reproduce the respective equilibrium unfolding signals closely, indicating that the folding kinetics were recorded long enough for the protein to relax to equilibrium. From the GdnHCl concentration dependence of the  $S_0^{\text{fol}}$ values, it is seen that the burst phase increase in fluorescence, far-UV CD as well as near-UV CD, which occurs initially upon folding, all change cooperatively with a decrease in GdnHCl concentration in the folding medium. While the equilibrium folding transitions measured by the three probes are clearly coincident, the burst phase folding transitions are clearly not. The values of the thermodynamic parameters obtained by fitting the burst phase folding transitions to twostate folding transitions (eq 1) depend on which optical probe is used to measure the transitions (see legend to Figure 4). This noncoincidence emphasizes that the unobservable burst phase kinetics cannot be two-state but are clearly multistate.

Figure 4 also presents the [GdnHCl] dependences of the initial and final signal values ( $S_0^{\text{unfol}}$  and  $S_\infty^{\text{unfol}}$ , respectively) obtained from kinetic experiments in which unfolding was monitored by all three optical probes. Again, in each case, the  $S_\infty^{\text{unfol}}$  values reproduce the respective equilibrium unfolding signals closely, indicating that the unfolding kinetics were recorded long enough for the protein to relax to equilibrium. The GdnHCl concentration dependences of the  $S_0^{\text{unfol}}$  values show that the decrease in burst phase signals recorded by fluorescence, far-UV CD as well as near-UV CD, which occurs initially during unfolding, all change cooperatively with an increase in GdnHCl concentration. Relative to the burst phase folding transition, the probe dependence of the burst phase unfolding transition is less pronounced: two-state fits to the burst phase unfolding



FIGURE 4: Quantitative analysis of equilibrium unfolding signals along with the initial and final (i.e., at t = 0 and  $t = \infty$ ) signals during kinetics of slow folding and unfolding reactions of barstar measured by fluorescence (a), far-UV CD (b), and near-UV CD (c), at 22  $(\pm 2)$  °C in 20 mM phosphate buffer, pH 8, containing 300  $\mu$ M EDTA and 250  $\mu$ M DTT. (**A**) Initial signal during folding  $(S_0^{\text{fol}}); (\Delta)$  initial signal during unfolding  $(S_0^{\text{unfol}}), (\Phi)$  final signal during folding  $(S_{\infty}^{\text{fol}})$ ; and (O) final signal during unfolding  $(S_{\infty}^{\text{unfol}})$ . These signal changes are shown relative to the changes of the corresponding optical signals observed during equilibrium unfolding of barstar,  $S_{\infty}^{eq}(\blacksquare)$ , under identical experimental conditions. At a particular denaturant concentration, the difference between  $S_0^{\text{fol}}$  and the unfolded baseline extrapolated linearly to nativelike conditions (lower dotted line) and between  $S_0^{\text{unfol}}$  and the native baseline extrapolated linearly to higher concentrations of GdnHCl (upper dotted line) describe, respectively, the amplitudes associated with fast reactions unobservable in manual mixing experiments (deadtime  $\sim$ 5–10 s). The thin solid lines in all three panels represent nonlinear least-squares fits of  $S_{\infty}^{\text{fol,unf,eq}}$  to a two-state N=U reaction with iterated values of  $\Delta G_{\rm NU}$  (kcal mol<sup>-1</sup>) and  $m_{\rm NU}$  (kcal mol<sup>-1</sup>  $M^{-1}$ ) as 5.1 and 2.5 for fluorescence (a), 4.5 and 2.3 for far-UV CD (b), and 4.8 and 2.5 for near-UV CD (c). The boldface solid lines drawn through  $S_0^{\text{fol}}$  ( $\blacktriangle$ ) and  $S_0^{\text{unf}}$  ( $\triangle$ ) values represent, respectively, the melting transitions for U $\rightleftharpoons$ I and N $\rightleftharpoons$ U<sub>F</sub> reactions. The fit parameters  $\Delta G$  and *m* (values in units of kcal mol<sup>-1</sup> and kcal mol<sup>-1</sup>  $M^{-1}$ , respectively) are as follows: the U=I reactions, 2.1 and 1.45 for fluorescence (a), 2.4 and 1.55 for far-UV CD (b), and 2.5 and 2.1 for near-UV CD (c); the N=U<sub>F</sub> reaction, 5.6 and 2.2 for fluorescence (a), 5.4 and 2.2 for far-UV CD (b), and 4.8 and 2.1 for near-UV CD (c). Standard deviations of signal amplitudes (error bars) were calculated from three or more sets of data. Standard errors for  $\Delta G$  and *m* are, respectively, 0.1 and 0.05.

transitions measured by the three probes yield similar m values as well as  $C_m$  values (see legend to Figure 4).



FIGURE 5: Optically monitored kinetics of slow folding and unfolding. (a) GdnHCl concentration dependence of the reduced amplitudes of folding. (b) GdnHCl concentration dependence of the reduced amplitudes of unfolding. (c) GdnHCl concentration dependence of the apparent rates of folding. (d) GdnHCl concentration dependence of the apparent rates of unfolding. ( $\bigcirc$ ) Fluorescence; ( $\bigcirc$ ) far-UV CD; ( $\square$ ) near-UV CD. The error bars represent standard deviations obtained from 3 or 4 repetitions of each experiment. Lines have been drawn by inspection only.

GdnHCl Concentration Dependences of Optically Monitored Folding and Unfolding Kinetics. The dependence on GdnHCl concentration of the reduced amplitude of the slow phase of folding was measured by all three optical probes (Figure 5a). The reduced amplitude is the amplitude of the observed change for folding to a particular final concentration of GdnHCl relative to the difference in the signals of fully folded and unfolded proteins at the same concentration of GdnHCl. The slow folding reaction is measurable in the presence of GdnHCl concentrations less than  $\sim 2.5$  M, and the reduced amplitude shows a relatively broad peak at  $\sim 1.6$ M GdnHCl where its value is maximum. This concentration of GdnHCl is fairly close to the  $C_{\rm m}$  values for cooperative transitions of  $S_0^{\text{fol}}$  values monitored by the three optical probes (see Figure 4). In 1.6 M GdnHCl, 50%, 41%, and 82% of the expected changes in fluorescence, far-UV CD, and near-UV CD, respectively, occur during the observable slow folding reaction. The values of the reduced amplitudes at concentrations of GdnHCl greater than the  $C_{\rm m}$  value are identical for all three probes. As the denaturant concentration in the folding medium falls below  $\sim 1.6$  M, the reduced amplitude decreases in a probe-specific manner. As more nativelike conditions are approached, the far-UV CD signal decreases sharply, and virtually no amplitude is measured in the presence of less than  $\sim 0.8$  M GdnHCl (Figure 5a), indicating complete formation of secondary structure within the dead-time of manual mixing experiments (see also Figure 4). Fluorescence and near-UV CD, however, present measurable signals even under the most strongly nativelike conditions (Figure 5a), yielding constant values for the reduced amplitudes of 0.1 and 0.2, respectively.

The reduced amplitude of the slow phase of the unfolding reaction of barstar, measured by all three optical probes as a function of GdnHCl concentration, is presented in Figure 5b. In GdnHCl concentrations less than  $\sim 1.3$  M, no slow

unfolding was detectable. For all three probes, the distribution of the reduced amplitude determined from unfolding kinetics peaks at approximately 2.2 M GdnHCl. This concentration of GdnHCl is close to the  $C_m$  for the transition of the  $S_0^{\text{unfol}}$ values (see Figure 4a–c). As the GdnHCl concentration in the unfolding medium is increased, the value of the reduced amplitude decreases gradually for all three optical probes. For unfolding in 2.3 M GdnHCl, 63% of the expected changes in fluorescence and near-UV CD, and 42% of the expected change in far-UV CD, occur during the slow phase of unfolding. An intriguing feature of these data common to all three probes of measurement is that the reduced amplitudes obtained from unfolding measurements are not coincident with those obtained from folding measurements (Figure 5a,b).

The dependences on GdnHCl concentration of the apparent rate constant of folding,  $\lambda_1$ , measured by the three optical probes are presented in Figure 5c, and display several important differences: (1) Values of  $\lambda_1$  measured by the use of far-UV CD are larger than those yielded by fluorescence or near-UV CD. (2) In the range of GdnHCl concentrations,  $\sim 1.4-0.6$  M, the values of log  $\lambda_1$  increase linearly when measured by fluorescence and near-UV CD. The rates measured by near-UV CD are, however, slightly larger than those measured by fluorescence. (3) When nativelike conditions are approached (<0.5 M GdnHCl), the denaturant dependences of the apparent folding rates measured by fluorescence and near-UV CD become nonlinear, more strikingly so for the latter.

The denaturant dependences of the apparent unfolding rate constants,  $\lambda_1$ , measured by the three optical probes are shown in Figure 5d. There are small differences in rates measured by the optical probes. The near-UV CD-measured rates are slightly faster than the fluorescence-monitored rates, which in turn are a bit faster than those measured by far-UV CD. Unlike the apparent folding rates, the unfolding rates show no rollovers in the rate-denaturant profile throughout the range of GdnHCl covered. The kinetics of slow unfolding in the presence of higher concentrations of the denaturant (~2.7 M GdnHCl) could not be monitored reliably, because the rates are too fast to measure by manual mixing of solutions.

GdnHCl Concentration Dependence of NMR-Monitored Folding Kinetics. Figure 6a shows the denaturant dependences of burst phase changes in NMR intensities  $(S_0^{\text{fol}})$  for a few resolved resonances. Since the intensity of a resolved resonance in the unfolded state of the protein is zero, the ratio of  $S_0^{\text{fol}}$  to  $S_{\infty}^{\text{fol}}$  yields the fraction of the final intensity that recovers during the dead-time (burst phase) of the folding measurement (~10 s). For all resonances, it is seen that the fraction of intensity that recovers in the burst phase increases as the GdnDCl concentration in the folding medium is lowered, and tends to level off in more nativelike conditions. More importantly, the data display considerable dispersion in burst phase changes in resonance intensities, for all concentrations of GdnDCl studied.

The dependences on GdnDCl concentration of the apparent rates of folding, as measured by the rates of development of the intensities of some of the resolved resonances, are shown in Figure 6b. The NMR-measured rates for all resonances are faster than the rates measured by any of the optical probes (Figure 5c). Figure 6b demonstrates the heterogeneity of



FIGURE 6: NMR-monitored slow kinetics of folding and unfolding. (a) The relative change in the intensity of NMR resonances during the burst phase of folding (given by  $S_0^{\text{fol}}/S_{\infty}^{\text{fol}}$ ) is plotted as a function of GdnDCl concentration. (b) GdnDCl concentration dependence of the apparent rate of folding. (c) GdnHCl concentration dependence of the apparent rate of unfolding. (•) L20 H<sup> $\delta$ 2</sup>; (•) L41 H<sup> $\delta$ 1</sup>; (•) L51 H<sup> $\delta$ 2</sup>; (□) I5 H<sup> $\gamma$ 2</sup>; (▲) L26 H<sup> $\delta$ 1</sup>; (△) F56 H<sup> $\epsilon$ </sup>; (▼) F56 H<sup> $\delta$ </sup>; (∇) W53 C<sub>α</sub>H; and (◇) an unassigned resonance (7.47 ppm) of the unfolded polypeptide. All lines were drawn by inspection only. The error bars represent standard deviations in values obtained from 2 repetitions.

GdnDCl concentration dependences of apparent folding rates for some sidechain resonances. Differences in denaturant dependences of folding rates between the aliphatic and aromatic side-chain resonances exist. The folding rates for the two aromatic resonances, F56 H<sup> $\epsilon$ </sup> and F56 H<sup> $\delta$ </sup>, show less denaturant dependence producing a distinct rollover in the plot of log  $\lambda_1$  vs GdnDCl concentration under strongly nativelike conditions, consistent with the results obtained by fluorescence and near-UV CD measurements (see Figure 5c).

GdnHCl Concentration Dependence of NMR-Monitored Unfolding Kinetics. The [GdnDCl] dependences of the apparent rates of unfolding, as measured by the rates of decrease in the intensities of some of the resolved resonances, are shown in Figure 6c. Again, the NMR-measured rates are significantly faster than the rates measured by any of the optical probes (see Figures 5d), but the dispersion in the NMR-measured unfolding rates is less than that seen for the NMR-measured folding rates. Under unfolding conditions,



FIGURE 7: Arrhenius plots for the slow folding and unfolding reactions of barstar at pH 8, in 20 mM sodium phosphate buffer containing 250  $\mu$ M EDTA and 300  $\mu$ M DTT. (a) Data measured by fluorescence: folding ( $\bigcirc$ ) and unfolding ( $\bigcirc$ ) in the presence of 0.1 and 3.0 M GdnHCl, respectively. Activation energies calculated by the use of eq 2 are 22.2 and 18.3 kcal mol<sup>-1</sup> at 0.1 and 3 M GdnHCl, respectively. (b) Data measured by far-UV CD: folding ( $\bigcirc$ ) and unfolding ( $\bigcirc$ ) in the presence of 1.0 and 2.3 M GdnHCl, with activation energies of 20.3 and 18.3 kcal mol<sup>-1</sup>, respectively. Activation energies for folding to different concentrations of GdnHCl are closely similar. Shown in the inset is the temperature dependence of folding in the presence of 0.6 ( $\blacksquare$ ), 1.1 ( $\triangle$ ), and 1.4 M ( $\bigtriangledown$ ) GdnHCl, for which the values of  $E_a$  are, respectively, 21.9 (0.6 M), 20.9 (1.1 M), and 19 (1.4 M) kcal mol<sup>-1</sup>. The error bars show standard deviations of rates determined from 3 or more measurements.

the rates measured for W38 H<sup> $\alpha$ </sup> and the unassigned 7.47 ppm peak associated with the unfolded state are at least 3-fold higher than the rest. Rollovers in the unfolding arms of plots of log  $\lambda_1$  vs GdnDCl concentration are apparent for all the resonances.

Since many of the conclusions (see Discussion) are based on differences in the kinetics determined by different probes, it was important to determine whether these are real or whether they reflect large errors in the measurement of kinetics. This was especially important for folding experiments where the data are more noisy (Figure 3) than in unfolding experiments because a much lower concentration of protein was used. To obtain an estimate of the error, measurement of folding in 0.42 (±0.02) M GdnHCl was repeated. The spread in the values around the mean value of either burst phase amplitude or rate for any proton was significantly smaller than the dispersion in mean values for different protons (Figure 6a,b). In similar folding experiments at pH 7 (data not shown), the errors determined were also similarly small in comparison to the differences between proton probes.

Temperature Dependence of Optically Monitored Folding and Unfolding Kinetics. The folding and unfolding reactions were also carried out as a function of temperature in the range 283–313 K for different final concentrations of GdnHCl in the reaction medium (Figure 7). Both far-UV CD (Figure 7a) and fluorescence (Figure 7b) were used to monitor the reactions. To obtain the activation energies, plots of  $\ln(\lambda_1)$ vs 1000/*T* were fit to the equation:

$$\ln(\lambda)_1 = \ln(A) - E_a/kT \tag{2}$$

where *A* is the frequency factor (s<sup>-1</sup>) and  $E_a$  is the activation energy (kcal mol<sup>-1</sup>). The values determined for  $E_a$  are listed in the legend to Figure 7. For both probes of measurement, the  $E_a$  for unfolding is less than the  $E_a$  for refolding by  $\sim 2$ kcal mol<sup>-1</sup>.

#### DISCUSSION

In this paper the results of a detailed study of the slow folding and unfolding reactions of barstar have been presented. In the following we discuss that the observed slow kinetics are the consequences of a coupling between the fast and the slow folding phases, and that proline isomerization that is believed to produce the slow phases acts as a kinetic trap for the accumulation of a multitude of protein intermediate states formed during the fast phase of folding and unfolding.

The Slow Phase of Refolding of Barstar and Proline Isomerization. The folding of barstar has been described by

$$UF \rightarrow N$$
 (3)

$$\mathbf{U}_{\mathrm{S}} \rightarrow \mathbf{I}_{\mathrm{N}} \rightarrow \mathbf{N} \tag{4}$$

where  $U_F$  and  $U_S$  represent, respectively, 31% fast folding and 69% slow folding molecules, and  $I_N$  is a long-lived, late folding intermediate. Early characterization of the folding of barstar showed that  $U_S \rightarrow I_N \rightarrow N$  constitutes the major folding pathway. As expected for the three-state scheme in eq 4, the folding kinetics of barstar display two observable kinetic phases, when monitored directly by fluorescence or far-UV CD or near-UV CD (17). The two phases are well separated in time: the faster phase characterized by an observed rate constant  $\lambda_2$  occurs in the subsecond time domain, while the slower phase characterized by a rate constant  $\lambda_1$  occurs in the 10–1000 s time domain (19–20).

The Y47–P48 prolyl peptide bond is *cis* in fully folded barstar, and is *trans* in U<sub>S</sub>. Several observations have suggested that *trans*—*cis* isomerization of the Y47–P48 bond occurs during the slow  $I_N \rightarrow N$  step shown in eq 4, and that  $I_N$  has at least some nativelike structure: (1) The slow phase observed directly in fluorescence-monitored kinetics is catalyzed by peptidyl-prolyl isomerase, albeit in stoichiometric quantities (19). (2) Like N,  $I_N$  is capable of inhibiting barnase (19), suggesting that the slow phase may involve only local structural rearrangements, consistent with the scenario of the  $I_N \rightarrow N$  step being a local proline isomerization event. (3) Consistent with the activation energy of proline isomerization, the values for  $E_a$  of the reaction are 20.5 (±1.4) kcal mol<sup>-1</sup> for folding and 18 (±0.5) kcal mol<sup>-1</sup> for unfolding (see Figure 7). (4) The rate of the  $I_N \rightarrow N$ reaction has been measured directly by double-jump experiments and is similar in value to the rates observed for other proteins where *trans-cis* isomerization has been implicated more directly. (5) The directly measured rate of the  $I_N \rightarrow N$ reaction is independent of GdnHCl concentration.

Since the structural transition in the  $U_S \rightarrow I_N$  step and the proline isomerization reaction in the  $I_N \rightarrow N$  step are coupled, both  $\lambda_2$  and  $\lambda_1$  are determined by all four microscopic rate constants that define the forward and backward transitions in eq 4. The significant contributions of the forward and backward microscopic rate constants defining the  $U_S \rightarrow I_N$ step to the value of  $\lambda_1$  are attested to by the following experimental observations: (1)  $\lambda_1$  is significantly different from the rate of the  $I_N \rightarrow N$  reaction measured directly from double-jump experiments. (2) Unlike the rate of the  $I_N \rightarrow N$ reaction,  $\lambda_1$  is sensitive to denaturant concentration, and the values of  $\lambda_1$  display chevrons with distinct rollovers (Figures 5 and 6). (3) Conservative mutations that should have no effect on the proline isomerization reaction have significant effects (31) on  $\lambda_1$ . (4)  $\lambda_1$  displays a small but significant dependence on pH (in the range  $\sim 6.7-9$ ), as determined from NMR and fluorescence-monitored kinetics (unpublished results).

In the present study, folding has been carried out in concentrations of GdnHCl corresponding to those either in the unfolding transition zone or smaller, where the coupling of the fast structural folding transitions to the slow  $I_N \rightleftharpoons N$ reaction produces an observable slow folding phase. Similarly, unfolding has been carried out in concentrations of GdnHCl corresponding to those in the unfolding transition zone, where the coupling of the fast structural unfolding transitions to the slow  $U_F \rightleftharpoons U_S$  reaction yields an observable slow unfolding phase. Thus, characterization of the observable slow folding and unfolding kinetics leads indirectly to characterization of the structural folding and unfolding transitions. Characterization of the fast structural folding or unfolding transitions by the real-time NMR method used in this study would not have been possible in the absence of coupling to the slower  $I_N \rightleftharpoons N$  and  $U_F \rightleftharpoons U_S$  reactions, respectively. The proline isomerization reaction also serves as a kinetic trap for real structural intermediates (32), and many different probes for folding, optical as well as NMR, have been used to study the nature of the products of subsecond folding as well as unfolding reactions, which accumulate because of the kinetic traps.

Multiple Folding Intermediates Are Present at 10 s of Folding. The three optical probes used to monitor the slow phase of folding yield different folding kinetics. The amplitude of the unobservable burst phase change in an optical probe as well as its dependence on GdnHCl concentration is probe dependent (Figure 4). The noncoincidence of the burst phase changes in the optical signals is also reflected in the different dependences of the reduced amplitude of the observable slow phase on GdnHCl concentration for the three optical probes (Figure 5a). Similarly

the apparent rate constants of the slow observable phase and their dependences on GdnHCl concentration are different for each probe. The optical results alone suggest that not one but at least three folding intermediates, I<sub>1</sub>, I<sub>2</sub>, and I<sub>3</sub>, are the products of the 10 s burst phase of folding in the present experiments. These intermediates differ in their fluorescence properties as well as in their far-UV and near-UV CD properties.

Not just the optical probes, but also the different resonances in the NMR spectra used to monitor folding gave different kinetics. Burst phase changes in resonance intensities and their denaturant dependences are resonance specific (Figure 6a). These observations along with the considerable dispersion in the NMR-measured rates for folding, determined for various resonances (Figure 6b), the differences in the GdnHCl concentration dependences of these rates, and the differences between the NMR-measured and optically measured rates provide substantial evidence for the accumulation of not just three but of a large number of folding intermediates,  $I_L^n$  (n = 1, 2, 3, ...) at 10 s of folding. The subscript L in  $I_L^n$  has been used to indicate that the intermediates accumulate at a later stage of the fast phase of folding kinetics.

Accumulation of the Fast Refolding Unfolded Form,  $U_F$ , *during the Burst Phase of Unfolding*. Unlike the burst phase folding transitions, the burst phase unfolding transitions measured by the three optical probes are fairly coincident, suggesting that only one species may form at the end of the 10 s burst phase (Figure 4). Previous work has shown that the unfolding reaction of barstar in the transition zone and the post-transition zone of denaturant concentration can be described by a three-state model (5, 19, 20):  $N \rightleftharpoons U_F \rightleftharpoons U_S$ , where  $U_F$  and  $U_S$ , respectively, are fast and slow refolding forms of the unfolded polypeptide. The  $U_S \rightleftharpoons U_F$  equilibrium, characterized by an equilibrium constant  $K_{21} = U_F/U_S = 0.47$ (20), produces 69% of  $U_S$  and 31% of  $U_F$  molecules. Three important properties of the GdnHCl dependence of  $S_0^{\text{unfol}}$ values (Figure 4) indicate that the burst phase species detected in slow unfolding reactions is U<sub>F</sub>. First, the rate-GdnHCl profile for the fast unfolding reaction exhibits a broad minimum at  $\sim$ 2.4 M GdnHCl (5) which, within the limit of experimental error, corresponds to the value of  $C_{\rm m}$ for denaturant titration of the  $S_0^{\text{unfol}}$  values (Figure 4). Second, the denaturant dependence of the reduced amplitude of the slow unfolding phase exhibits a maximum around the value of  $C_{\rm m}$  (Figure 5b), as expected (33). Third, the difference in free energy of U<sub>F</sub> and of the equilibrium mixture of U<sub>F</sub> and  $U_{\rm S}$  is given by (24)

$$\Delta\Delta G = -RT \ln \frac{K_{21}}{1 + K_{21}} \tag{5}$$

With  $K_{21} = 0.47$ , the value of  $\Delta\Delta G \sim 0.7$  kcal mol<sup>-1</sup>. The value of  $\Delta G$  for the transition of the unfolding burst amplitudes measured by fluorescence and far-UV CD is higher than the equilibrium unfolding free energy by  $\sim 0.5$  kcal mol<sup>-1</sup> (Figure 4), which implies that the final product of the unfolding burst phase is U<sub>F</sub> itself.

Multiple Intermediates Are Populated Transiently during Unfolding to  $U_F$ . It is clear that only  $U_F$  is stable enough to accumulate in the 10 s burst phase of the slow unfolding reaction studied here. Optically monitored unfolding studies

by themselves do not suggest that any intermediate is populated transiently during the unfolding of N to U<sub>F</sub>. While there is some dispersion in the values of  $\lambda_1$  and the reduced amplitudes for unfolding measured by the three optical probes, the dispersion is small in comparison to that observed for the folding reaction (Figure 5a-d). The NMR results, however, provide considerable evidence that unfolding proceeds through many intermediates. Nonidentical values of  $\lambda_1$  observed for various resonances (Figure 6c), the noncoincidence of the NMR-measured and optically measured values, and the differences in the denaturant dependences of the values of  $\lambda_1$  determined from various optical and NMR probes all imply that a large number of intermediates,  $I_U^1$ ,  $I_U^2$ ... $I_U^n$ , precede U<sub>F</sub> on the unfolding pathway.

*Multistate Kinetics and Multiple Trajectories for Folding.* Two minimal representations can be considered to describe the folding of U to N involving the late structural intermediates,  $I_L^1$ ,  $I_L^2$ , ... $I_L^n$ . In the multiple pathway representation:



the unfolded molecules can fold rapidly to the native state by convergent folding routes. Folding is slowed by accumulation of *n* number of late intermediates:  $I_L^1$ ,  $I_L^2$ , ... $I_L^n$ .

In the single pathway ensemble mechanism:



 $I_L^1$  forms rapidly from U, and preequilibrates rapidly with an ensemble of intermediates. The U =  $I_L^1$  equilibrium and the equilibria between the intermediates comprising the ensemble are all dependent on GdnHCl concentration.

Both schemes are not inconsistent with the experimental observation that folding occurs in two observable kinetic phases, one fast and one slow. While each of the  $I_L \rightarrow N$  reactions in the multiple pathway mechanism might be expected to yield a separate slow exponential, only one slow phase of folding is seen because the rates of these reactions are very similar as they are determined predominantly by the proline isomerization reaction. Different probes yield slightly different values for  $\lambda_1$  because the rates of formation

of the different I<sub>L</sub> forms, to which the proline isomerization reaction is coupled, differ. According to the ensemble mechanism, again only one slow phase is expected corresponding to the I<sub>L</sub>  $\rightarrow$  N reaction, and different probes yield different values for  $\lambda_1$  because the spectroscopic properties of the intermediates constituting the ensemble are different.

Available evidence favors the multiple pathway representation. (1) In this scenario, the rates of formation of the intermediates  $I_L^1$ ,  $I_L^2$ , ... $I_L^n$ , which accumulate at the end of the 10 s burst phase of the measurements reported here, should be different. Since the fluorescence, far-UV CD, and near-UV CD properties of these intermediates appear to be different, the differences in the rates should be apparent when the rates of folding in the subsecond time domain are measured in the millisecond time domain using the three different optical probes. Significant differences in both the rate constants and amplitudes of folding were indeed observed when the sub-second folding of barstar in 1 M GdnHCl was studied with the three optical probes (17) used in the present study. (2) A detailed characterization of the folding of barstar in the millisecond time domain has shown that under nativelike conditions (<1.0 M GdnHCl), 3 folding routes are available for the protein to refold (5, 19, 20):

$$U_{S} \rightleftharpoons I_{M1} \rightarrow I_{S1} \rightarrow I_{N} \xrightarrow{\text{slow}} N$$
$$U_{S} \rightleftharpoons I_{M2} \rightarrow I_{S2} \xrightarrow{\text{slow}} N$$
$$U_{F} \rightleftharpoons I_{F1} \rightarrow N$$

It is likely that three of the many parallel folding pathways identified here from characterization of slow folding kinetics correspond to these three parallel folding pathways. Of the six intermediates identified in these three folding pathways, four are very short-lived, and also only marginally stable as expected for early intermediates: they would therefore not be expected to be populated to any significant extent at 10 s of folding. Only two intermediates, I<sub>N</sub> and I<sub>S<sup>2</sup></sub>, are longlived (50-100 s) and sufficiently stable, and might be expected to correspond to two of the late intermediates:  $I_{L}^{1}$ ,  $I_L^2 \dots I_L^n$ . It should be noted that although  $I_M^1$ ,  $I_M^2$ ,  $I_F^1$ , and  $I_{S}^{1}$  are too unstable to have accumulated to any significant extent at the end of the 10 s burst phase of folding, their presence on the folding pathway is at least partially responsible for the noncoincidence in the values of  $\lambda_1$ observed using different optical and NMR probes, and for the differences in the dependences of  $\lambda_1$  on GdnHCl concentration observed. (3) The consequence of the occurrence of intermediates in parallel pathways versus in a sequential ensemble can be considered in terms of the mvalues. The m value for a given kinetic step in a folding reaction must be less than or equal to the *m* value associated with the equilibrium folding reaction. In the parallel pathway representation of folding, the *m* value for each  $U \rightarrow I_L^n$ reaction is less than the *m* value for the overall  $U \rightleftharpoons N$ equilibrium transition. On the other hand, for the sequential ensemble mechanism, the sum of *m* values of the multiple  $S_0^{\text{fol}}$  transitions exceeds overwhelmingly the *m* value associated with the equilibrium transition measured by any probe (see legend to Figure 4), which would mean that the intermediate whose rate of formation is slowest of all has

more nonpolar surface buried relative to that in the native state. Clearly this is unlikely, so the folding intermediates must exist on parallel routes.

*Multistate Kinetics and Multiple Pathways of Unfolding.* Previous millisecond measurements of the unfolding of barstar (22) had revealed the presence of at least two competing unfolding pathways:



where  $I_U^1$  and  $I_U^2$  are two very early intermediates, the former devoid of much of the secondary structure and the other with a perturbed hydrophobic core, which unfold to U<sub>F</sub> in the 50-500 ms time domain. Evidence for unfolding intermediates also comes from hydrogen exchange measurements of barstar in the presence of partially denaturing concentrations of GdnHCl (24). For several amide hydrogens, the free energies of opening of structure to exchange ( $\Delta G_{op}$ ) are larger than the free energy of global unfolding ( $\Delta G_{op}$ ), and plots of  $\Delta G_{op}$ vs GdnHCl concentration display significant dispersion even at concentrations of GdnHCl approaching  $C_{\rm m}$ . These results had also therefore suggested that multiple unfolding intermediates  $(I_U^1, I_U^2...I_U^n)$  precede U<sub>F</sub> on the unfolding pathway. The present studies cannot by themselves implicate these intermediates on parallel competing pathways. Nevertheless, the results of the previous millisecond measurements of the unfolding of barstar (22) suggest strongly that  $I_U^1$ ,  $I_U^2$ ... $I_U^n$ accumulate on competing pathways.



The New and Classical Views of Folding. Previous folding studies had suggested that only two distinct long-lived intermediates,  $I_N$  and  $I_S^2$ , accumulate on the folding pathways of barstar (5, 18, 19). The results of this study suggest that  $I_N$  and  $I_S^2$  are not unique structural forms, but represent a large number of partly folded, late intermediate forms ( $I_L^1$ ,  $I_L^2$  ... $I_L^n$ ) that accumulate on many folding pathways. Similarly, previous unfolding studies had suggested that two unique structural forms,  $I_U^1$  and  $I_U^2$ , accumulate on the unfolding pathways of barstar (21, 22). Again, the results of this study suggest that these two intermediates represent many intermediate forms.

The results reported here can be put into the context of convergent kinetic routes explained from the perspective of energy landscape theory (9-11, 13), according to which the 'ruggedness' of the energy surface leads to multistate kinetics or transient trapping of structures that are either partially folded or misfolded. At present, no information is available

on the structures of the intermediates detected here, and, hence, it cannot be ascertained whether any or some of these structures are misfolded. In the landscape view, parallel multiple pathways represent the multiple trajectories of reconfiguring polypeptide motions leading to the folded or the unfolded state.

How close are the  $I_L^n$  species to N, and  $U_F$  and  $U_S$  in terms of conformational free energy and entropy? There are two points to note: (1) according to the interpretation presented here, a large number of folding intermediates accumulate; and (2) the values for  $\Delta G$  associated with  $I_L^n \rightarrow N$  reactions are approximately half of the equilibrium unfolding free energy of barstar. Both of these observations would suggest that in a convergent folding funnel the intermediates would possess substantial configurational entropy. Nevertheless, the stabilities of  $I_L^n$  are significantly larger than the stabilities of partially folded molten globule states that are formed initially during the refolding of many proteins.

Multiple folding pathways also imply that no unique transition state governs the rate of folding and unfolding, contrary to the assumption made in the classical model of protein folding kinetics that a specific structure of highest free energy rate-limits the folding process. What these data reveal is the existence of an ensemble of transition states which limit the overall rate of folding. The results clearly delimit the original idea, generally referred to as the classical view (6), that in order to meet the Levinthal time requirement polypeptides proceed to fold quickly via specific pathways consisting of well-defined structural intermediates. It is hoped that these results, besides bringing theory and experiment closer, will lead to a greater appreciation of the importance and the complexity of slow folding kinetics.

# ACKNOWLEDGMENT

We thank Prof. R. L. Baldwin for his comments on the manuscript. The NMR spectra were recorded in the National Facility for High Resolution NMR, TIFR, Mumbai.

# REFERENCES

- Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol.* 2, 663–673.
- Van Nuland, N. A. J., et al. (1998) Biochemistry 37, 622– 637.
- Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10428–10435.
- 4. Matthews, C. R. (1993) Annu. Rev. Biochem. 62, 653-683.
- Shastry, M. C. R., and Udgaonkar, J. B. (1994) J. Mol. Biol. 247, 1013–1027.
- 6. Baldwin, R. L. (1995) J. Biomol. NMR 5, 103-109.
- 7. Baldwin, R. L. (1997) Nat. Struct. Biol. 4, 965-966.
- Bryngelson, J. D., and Wolynes, P. G. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 7524–7528.
- Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Proteins: Struct., Funct., Genet. 21, 167–195.
- 10. Dill, K. A., and Chan, H. S. (1997) Nat. Struct. Biol. 4, 10– 19.
- Chan, H. S., and Dill, K. A. (1998) Proteins: Struct., Funct., Genet. 30, 2–33.
- 12. Pande, V., Grosberg, A. Y., Tanaka, T., and Rokhsar, D. S. (1998) *Curr. Opin. Struct. Biol.* 8, 68–79.
- Leopold, P. E., Montal, M., and Onuchic, J. N. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 8721–8725.
- Sali, A., Shakhnovich, E., and Karplus, M. (1994) *Nature 369*, 248–251.

- Jones, C. M., et al. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 11860–11864.
- Eaton, W. A., Munoz, V., Thompson, P. A., and Hofrichter, J. (1997) *Curr. Opin. Struct. Biol.* 7, 10–14.
- 17. Agashe, V. R., Shastry, M. C. R., and Udgaonkar, J. B. (1995) *Nature 377*, 754–777.
- Agashe, V. R., and Udgaonkar, J. B. (1995) *Biochemistry 34*, 3286–3299.
- 19. Schreiber, G., and Fersht, A. R. (1993) *Biochemistry 32*, 11195–11203.
- Shastry, M. C. R., Agashe, V. R., and Udgaonkar, J. B. (1994) Protein Sci. 3, 1409–1417.
- 21. Nath, U., Agashe, V. R., and Udgaonkar, J. B. (1996) Nat. Struct. Biol. 3, 920–923.
- Zaidi, F. N., Nath, U., and Udgaonkar, J. B. (1997) Nat. Struct. Biol. 4, 1016–1024.
- 23. Khurana, R., and Udgaonkar, J. B. (1994) *Biochemistry 33*, 106–115.
- 24. Bhuyan, A. K., and Udgaonkar, J. B. (1998) Proteins: Struct., Funct., Genet. 30, 295–308.

- Bhuyan, A. K., and Udgaonkar, J. B. (1998) Proteins: Struct., Funct., Genet. 30, 241–247.
- Adler, M., and Scheraga, H. A. (1988) *Biochemistry* 27, 2471– 2480.
- 27. Hoeltzli, S. D., and Frieden, C. (1996) *Biochemistry 35*, 16843–16851.
- Kiefhaber, T., Labhardt, A. M., and Baldwin, R. L. (1995) *Nature* 375, 513–515.
- 29. Kiefhaber, T., and Baldwin, R. L. (1995) Proc. Natl. Acad. Sci. U.S.A. 92, 2657–2661.
- 30. Balbach, J., et al. (1995) Nat. Struct. Biol. 2, 865-870.
- Nath, U., and Udgaonkar, J. B. (1997) *Biochemistry 36*, 8602– 8610.
- Kiefhaber, T., Kohler, H. A., and Schmid, F. X. (1992) J. Mol. Biol. 224, 217–229.
- 33. Utiyama, H., and Baldwin, R. L. (1986) *Methods Enzymol.* 131, 51–70.

BI990285W