

JMB

The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates

M. C. R. Shastry and Jayant B. Udgaonkar*

National Centre for Biological Sciences, TIFR Centre
PO Box 1234, Indian
Institute of Science Campus
Bangalore 560012, India

The mechanism of folding of the small protein barstar in the pre-transition zone at pH 7, 25°C has been characterized using rapid-mixing techniques. Earlier studies had established the validity of the three-state $U_S \rightleftharpoons U_F \rightleftharpoons N$ mechanism for folding and unfolding in the presence of guanidine hydrochloride (GdnHCl) at concentrations greater than 2.0 M, where U_S and U_F are the slow-refolding and fast-refolding unfolded forms, respectively, and N is the fully folded form. It is now shown that early intermediates, I_{S1} and I_{S2} as well as a late native-like intermediate, I_N , are present on the folding pathways of U_S , and an early intermediate I_{F1} on the folding pathway of U_F , when barstar is refolded in concentrations of GdnHCl below 2.0 M. The rates of formation and disappearance of I_N , and the rates of formation of N at three different concentrations of GdnHCl in the pre-transition zone have been measured. The data indicate that in 1.5 M GdnHCl, I_N is not fully populated on the $U_S \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ pathway because the rate of its formation is so slow that the $U_S \rightleftharpoons U_F \rightleftharpoons N$ pathway can effectively compete with that pathway. In 1.0 M GdnHCl, the $U_S \rightarrow I_{S1} \rightarrow I_N$ transition is so fast that I_N is fully populated. In 0.6 M GdnHCl, I_N appears not to be fully populated because an alternative folding pathway, $U_S \rightarrow I_{S2} \rightarrow N$, becomes available for the folding of U_S , in addition to the $U_S \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ pathway. Measurement of the binding of the hydrophobic dye 1-anilino-8-naphthalenesulphonate (ANS) during folding indicates that ANS binds to two distinct intermediates, I_{M1} and I_{M2} , that form within 2 ms on the $U_S \rightarrow I_{M1} \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ and $U_S \rightarrow I_{M2} \rightarrow I_{S2} \rightarrow N$ pathways. There is no evidence for the accumulation of intermediates that can bind ANS on the folding pathway of U_F .

*Corresponding author

Keywords: barstar; folding kinetics; protein folding; molten globule; ANS

Introduction

The mechanism by which the primary structure directs the folding of the protein to the fully folded state is poorly understood. The folding of a protein usually follows one or a few kinetic pathways in which structural intermediates are populated (Kim & Baldwin, 1990; Matthews, 1993). Multiple pathways usually arise because of the heterogeneity of the unfolded state that results from *cis-trans* isomerization of *X-Pro* bonds upon unfolding (Brandts *et al.*, 1975), with different unfolded forms folding via parallel pathways (Garel & Baldwin, 1973; Schmid, 1983, 1986).

A major difficulty in experimental investigations of folding and unfolding of proteins is the reliable

identification and structural description of the intermediate states (Kim & Baldwin, 1990; Ptitsyn, 1994). Only a very restricted number of methods can be used for obtaining structural descriptions of kinetic intermediates that exist only transiently, usually in the millisecond time domain. Nevertheless, in recent years, several new methodologies have been applied to the study of the partly formed structures that serve as intermediates on the folding pathways of proteins (Evans & Radford, 1994). These include stopped-flow circular dichroism (Kuwajima *et al.*, 1987), electron spin resonance spectroscopy (Semisotnov *et al.*, 1987), and proton nuclear magnetic resonance spectroscopy used in conjunction with amide hydrogen exchange (Udgaonkar & Baldwin, 1988; Roder *et al.*, 1988). Stopped-flow fluorescence or absorbance studies provide a means to characterize the folding pathway of a protein and to identify the conditions in which structural intermediates accumulate, and facilitate the application of these newer structural methodologies.

Abbreviations used: BSCCAA, a Cys40Cys82 → Ala40Ala82 double mutant form of barstar; GdnHCl, guanidine hydrochloride; ANS, 1-anilino-8-naphthalenesulphonate.

More importantly, these more traditional studies are absolutely necessary for a correct interpretation of results obtained using the newer structural methods (Udgaonkar & Baldwin, 1990).

The small 89 amino acid residue protein barstar, which functions as an intracellular inhibitor of barnase in *Bacillus amyloliquefaciens* is an ideal model protein for folding studies. Barstar is available in a good expression system (Hartley, 1988). It undergoes completely reversible unfolding transitions whether denatured by GdnHCl, urea, temperature or pH (Khurana & Udgaonkar, 1994), its X-ray crystal structure is known (Guillet *et al.*, 1993), and its solution structure has been solved by NMR (Lubienski *et al.*, 1994). At low pH, barstar adopts a molten globule form (Khurana & Udgaonkar, 1994; Swaminathan *et al.*, 1994). The folding pathway of BSCCAA, a Cys40Cys82 → Ala40Ala82 double mutant form of barstar has also been studied at pH 8, using urea as the denaturant (Schreiber & Fersht, 1993).

At concentrations of GdnHCl greater than 2.0 M a three-state model, $U_S \rightleftharpoons U_F \rightleftharpoons N$, incorporating two unfolded forms, a fast-refolding U_F and a slow-refolding U_S , and the fully folded protein, N , completely describe the kinetics of folding and unfolding (Shastry *et al.*, 1994). U_S and U_F appear to differ in possessing *trans* and *cis* conformations, respectively, of the Tyr47-Pro48 bond (Schreiber & Fersht, 1993). Folding in marginally stable conditions (1.2 M GdnHCl) was shown to occur by two parallel pathways: $U_F \rightarrow N$ and $U_S \rightarrow I_N \rightarrow N$ (Shastry *et al.*, 1994). The native-like intermediate, I_N has been shown to be capable of inhibiting barnase activity (Schreiber & Fersht, 1993). In the case of BSCCAA, very early intermediates (I_1) were also implicated on both of the folding pathways in the pre-transition region, but there was no direct data on their roles (Schreiber & Fersht, 1993).

Here, studies on the folding of U_F confirm that a very early intermediate I_{F1} is indeed present on the folding pathway of U_F , when barstar is folded in low concentrations of GdnHCl. It is also shown that in such stabilizing conditions, there are two competing pathways for the folding of U_S , on only one of which the native-like intermediate I_N accumulates. On each of the two competing pathways, a molten globule-like intermediate, capable of binding the dye ANS, accumulates within 2 ms. It is also demonstrated that no intermediate that can bind ANS accumulates on the folding pathway of U_F . To account for the change in fluorescence that accompanies the folding of U_S , it has been necessary to include two intermediates I_{S1} and I_{S2} on the two folding pathways.

Results

Folding of transiently unfolded barstar (U_F)

Figure 1 illustrates the kinetics of folding of U_F (see Materials and Methods) in 1.5 M GdnHCl. It is observed that a fast folding reaction is followed by a

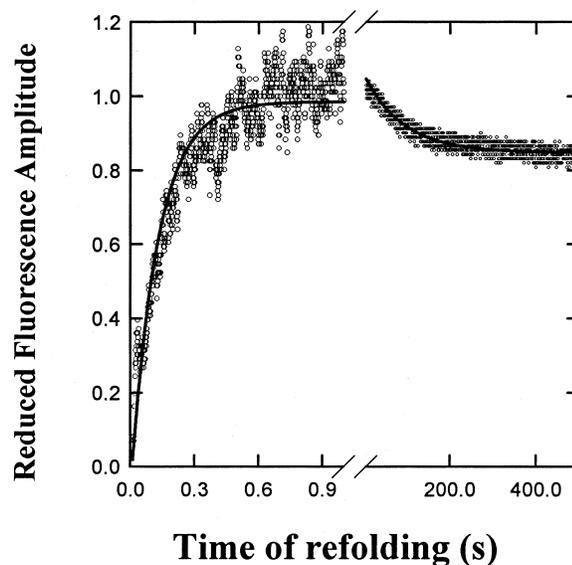


Figure 1. Folding of transiently unfolded barstar at pH 7, 25°C. Barstar in native buffer was unfolded in 3.7 M GdnHCl for 5 seconds, before being diluted into the final folding conditions of 1.5 M GdnHCl. The change in fluorescence is plotted against time of refolding. The continuous line drawn through the data is a non-linear least-squares fit of the data to equation (1), and yields values for λ_1 , λ_2 , A_1 and A_2 of 0.01 s⁻¹, 12 s⁻¹, -0.21 and 1.04, respectively.

slow unfolding reaction. At equilibrium, only 85% of the molecules are fully folded, as expected from an equilibrium denaturation curve (Shastry *et al.*, 1994).

Comparison of the folding kinetics of U_F to equilibrium-unfolded barstar

Figure 2a to c shows the dependence of the folding kinetics of U_F on the final concentration of GdnHCl. Also shown for comparison are the folding kinetics of equilibrium-unfolded barstar, which contains both U_F and U_S in slow equilibrium (Shastry *et al.*, 1994). In Figure 2a, it is seen that both of the observed rate constants for the folding of U_F are similar to the observed rate constants for the folding of equilibrium-unfolded barstar. A linear extrapolation of the log λ_2 values measured at low GdnHCl concentrations to zero GdnHCl according to equation (2), yields the rate of folding of U_F in water of 34 s⁻¹. In Figure 2b are shown the dependences of the individual reduced amplitudes of both the fast and the slow phases of folding of U_F . It is observed that the reduced amplitude of the slow phase is negative, indicating that it corresponds to an unfolding and not a folding reaction. In Figure 2c, the dependence of the total reduced amplitude (sum of the reduced amplitudes of the fast and slow phases) of the folding reaction of U_F on the concentration of GdnHCl is compared with the total reduced amplitude of the folding reaction of equilibrium-unfolded barstar. The good agreement between the

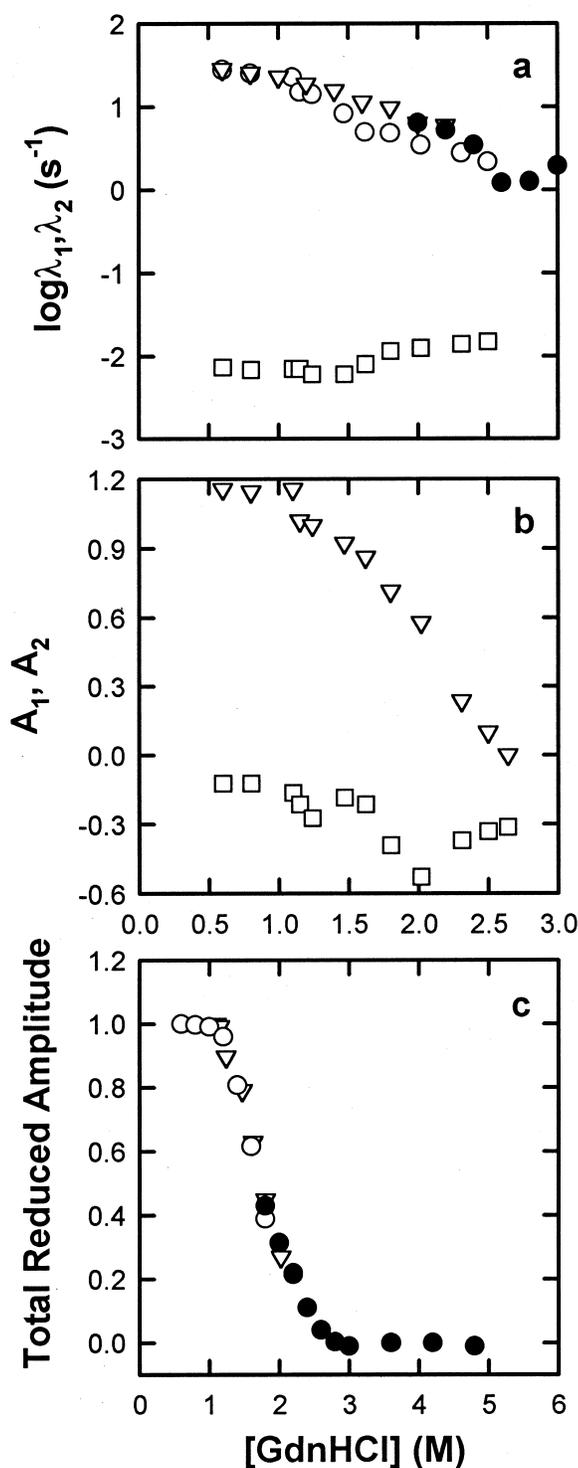


Figure 2. Kinetics of folding of U_F at pH 7, 25°C. Barstar in native buffer was unfolded in 3.7 M GdnHCl for 5 seconds, before being diluted into different final concentrations of GdnHCl. (∇) and (\square) Data obtained starting from U_F ; (\circ) and (\bullet) data obtained starting with an equilibrium mixture of U_F and U_S . a, Dependence of λ_2 (\circ , \bullet , ∇) and λ_1 (\square) on the concentration of GdnHCl present during folding. b, Dependence of the reduced amplitudes, A_1 and A_2 on the final concentration of GdnHCl present. c, Dependence of the total reduced amplitude ($A_1 + A_2$) on the concentration of GdnHCl present during refolding.

total reduced amplitudes at any concentration of GdnHCl indicates that the final composition of the protein molecules is the same irrespective of whether they originate from only U_F or from an equilibrium mixture of U_F and U_S .

In Figure 3a, the reduced amplitude of the fast folding reaction starting from U_F is compared with that of the fast folding reaction observed on starting from an equilibrium mixture of U_F and U_S . The reduced amplitude decreases to half its value when the GdnHCl concentration is raised to 1.3 M in the former case and to 2.0 M in the latter case.

An intermediate I_{F1} accumulates on the folding pathway of U_F

The dependence of λ_2 on GdnHCl concentration can be calculated for a two-state $U_F \rightleftharpoons N$ folding transition as follows (Matouschek *et al.*, 1990): The free energy of unfolding of barstar in water obtained from equilibrium unfolding measurements ($\Delta G_{app} = 4.9$ kcal/mol) yields a value of 1.27×10^{-4} for the apparent equilibrium constant K_{app} ($= K_{32}(1 + K_{21})$) for folding according to a $U_S \rightleftharpoons U_F \rightleftharpoons N$ model. K_{21} has a GdnHCl concentration-independent value of 2.2 (Shastry *et al.*, 1994) and K_{32} , the equilibrium constant characterizing the $U_F \rightleftharpoons N$ reaction, therefore has a value of 4.11×10^{-5} in water. The value of λ_2 for the fast unfolding reaction, λ_u , is equal to 0.096 s⁻¹ in water. Thus, the value of λ_2 for the fast folding reaction, λ_f ($= K_{32}\lambda_u$) is 2213 s⁻¹ in water. Since the linear dependences of $\log K_{app}$ and $\log \lambda_u$ on GdnHCl concentration have slopes of 1.8 M⁻¹ and 0.36 M⁻¹, respectively (Shastry *et al.*, 1994), that of $\log \lambda_f$ on GdnHCl concentration consequently has a slope, m_{λ_2} of 1.5 M⁻¹ ($1.86 - 0.36$). The values expected for λ_f and λ_u if the $U_F \rightleftharpoons N$ transition were two-state were therefore determined at all GdnHCl concentrations using equation (2). Figure 3b shows that the values calculated for λ_2^{obs} ($= \lambda_f + \lambda_u$) do not match the values for λ_2 determined experimentally for the folding of U_F in GdnHCl concentrations less than 1.0 M. Thus, the $U_F \rightleftharpoons N$ transition cannot be two-state, and an intermediate I_{F1} must be present on the folding pathway of U_F .

Accumulation of I_N and N measured in double-jump experiments

Folding in marginally native-like conditions (1.0 M GdnHCl)

Figure 4 shows that N and I_N accumulate to levels of 31% and 69%, respectively, within 0.1 second, the earliest experimental time-point. Subsequently, the population of N rises to 100% and that of I_N decreases to 0% after 300 seconds of refolding, at a rate of 17×10^{-3} s⁻¹, in each case. The kinetics of formation

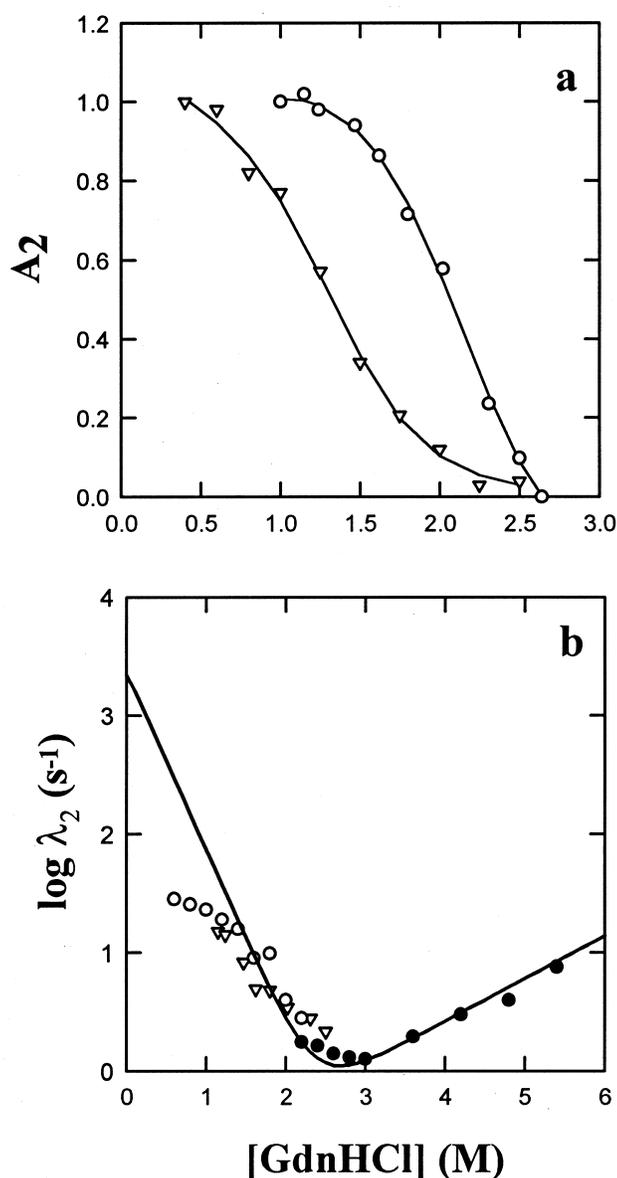


Figure 3. Dependence of the kinetics of the fast folding reaction on GdnHCl concentration at pH 7, 25°C. (∇) Data for U_F . (\circ) and (\bullet) Data for an equilibrium mixture of $U_F + U_S$. a, Dependence of the reduced amplitude A_2 . The continuous lines through the data have been drawn by inspection only. Visual interpolation shows that the concentration of GdnHCl at which the amplitude is reduced to half is 1.3 M (starting from U_F) and 2.0 M (starting from an equilibrium mixture of $U_F + U_S$). b, Dependence of the rate constant λ_2 . The folding rate constants have been taken from Figure 2a. The unfolding rate constants were obtained when native protein was unfolded to each of the concentrations of GdnHCl indicated and have been taken from Shastry *et al.* (1994). The continuous line is the GdnHCl concentration dependence of $\log \lambda_2^{\text{obs}} (= \log(\lambda_f + \lambda_u))$ calculated from equilibrium unfolding data and kinetic unfolding data for a 2-state model (see the text).

of I_N are too fast to be monitored. I_N and N together account for all protein molecules at all times of folding between 0.1 and 300 seconds.

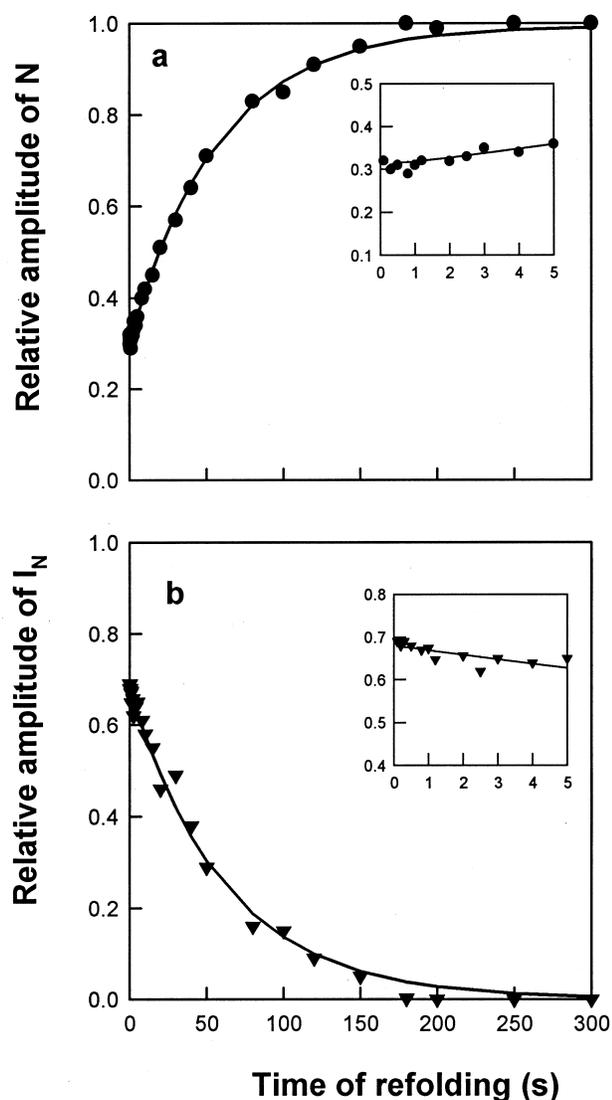


Figure 4. Kinetics of folding of equilibrium-unfolded barstar in 1 M GdnHCl, pH 7, 25°C. The amounts of N (a) and I_N (b) were determined at different times during the folding process by means of double-jump experiments in which various forms of the protein were identified by unfolding assays (see Materials and Methods). The amounts of N and I_N are plotted relative to the total amount of N that is formed when the folding reaction is allowed to go to completion (600 seconds) in 0.6 M GdnHCl. Each time-point was done in triplicate, and the standard errors in the determination of N and I_N , were less than $\pm 15\%$. The continuous lines through the data in a and b are non-linear fits of the data to Mechanism 2 in Scheme I, using equations (A1) to (A6). The value of k_2 was fixed to 18 s⁻¹, and the value obtained for k_1 is given in Table 1. The insets in a and b show the data and the fits on a shorter time-scale.

Folding in the transition zone (1.5 M GdnHCl)

In 1.5 M GdnHCl, only 85% of all barstar molecules are fully folded at equilibrium (Shastry *et al.*, 1994). Figure 5a shows that N accumulates to 30% of its final equilibrium value within 0.1 second of folding, and its population increases to its final equilibrium value after 300 seconds of folding.

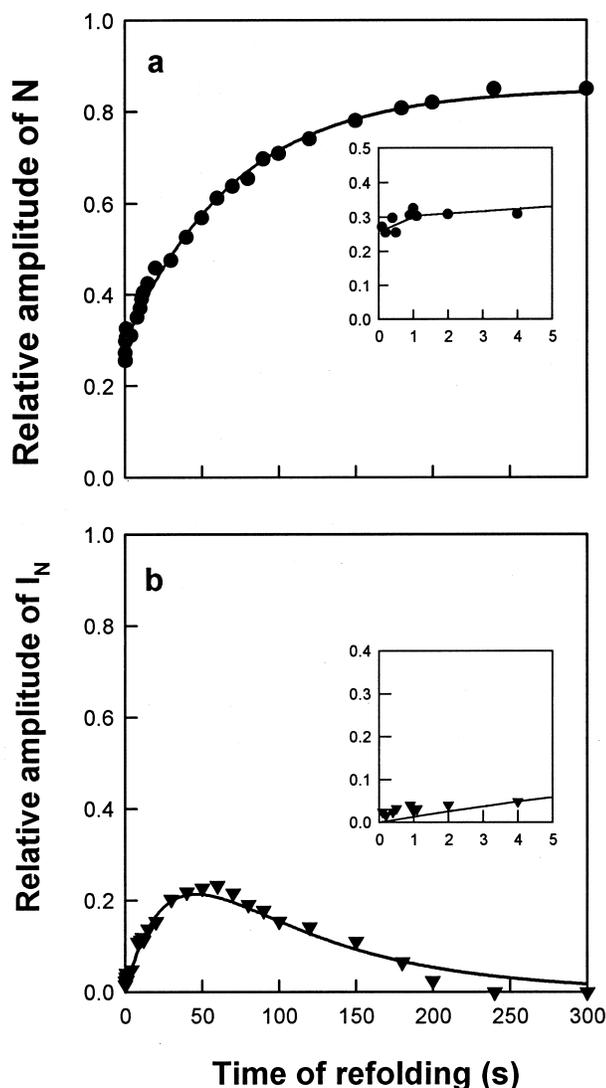


Figure 5. Kinetics of folding of equilibrium-unfolded barstar in 1.5 M GdnHCl, pH 7, 25°C. The amounts of N (a) and I_N (b) were determined at different times during the folding process by means of double-jump experiments, in which various forms of the protein were identified by unfolding assays (see Materials and Methods). The amounts of N and I_N are plotted relative to the total amount of N that is formed when the folding reaction is allowed to go to completion in 0.6 M GdnHCl. Each time-point was repeated in triplicate, and the standard errors in the determination of N and I_N were less than $\pm 15\%$. The continuous lines through the data are non-linear least squares fits of the data to Mechanism II in Scheme I, using equations (A1) to (A6). The value of k_2 was fixed to 12 s^{-1} , and the values obtained for k_3 and k_4 are given in Table 1. The insets in a and b show the data and the fits on a shorter time-scale.

Figure 5a shows also that there is apparently no lag in the formation of N, and that the maximal rate of formation of N is not when the population of I_N is maximal (at 60 seconds), but earlier. The rate of formation of I_N is 0.04 s^{-1} and the maximum population of I_N , 60 seconds after the commencement of folding (Figure 5b), is only 30% of the number of barstar molecules that are folded at equilibrium.

I_N and N do not account for all of the barstar molecules present until after 60 seconds. The residual amplitude (not shown) indicates that approximately 70% of the molecules that are folded at equilibrium are in conformations other than I_N and N at 100 ms, but this population decreases to 0% after 60 seconds of folding in a first-order process characterized by a rate constant of 0.04 s^{-1} . This suggests that these 70% molecules represent only one conformation that directly transforms to I_N , which is observed to form at the same rate (Figure 5b).

Folding in strongly native-like conditions (0.6 M GdnHCl)

In Figure 6, it is seen that N again accumulates to approximately 30% of its final equilibrium value within 0.1 second. There is a subsequent lag in further accumulation of N during the next five seconds (Figure 6a). In these five seconds, I_N accumulates at a rate of 1.0 s^{-1} (Figure 6b). The observed rate of formation of N is maximum at five seconds when the population of I_N is maximum. N is fully formed 300 seconds after initiation of refolding (Figure 6a). The maximum level of accumulation of I_N is only 35%, and it then disappears in a first-order process with a rate constant of 0.017 s^{-1} .

Again, I_N and N do not account for all barstar molecules until after 50 seconds of folding. The residual amplitude indicates that approximately 70% of the molecules are in other conformations at 100 ms. The residual amplitude (not shown) decays in a two-exponential process: about half of it disappears with a rate constant of 1 s^{-1} and the other half disappears with a rate constant of 0.07 s^{-1} . This suggests that there are at least two populations of molecules other than I_N and N at 100 ms. The 35% of the total number of molecules that disappear with a rate constant of 1.0 s^{-1} probably transform into I_N , which appears at the same rate and which accumulates also to 35%. The 35% of the total number of molecules that disappear with a rate constant of 0.07 s^{-1} are assumed to directly transform to N (see the legend to Figure 6).

ANS binding-monitored folding kinetics

The hydrophobic molecule ANS binds to hydrophobic clusters on proteins that are hydrated and therefore accessible (Stryer, 1965), and such binding is accompanied by a large change in the fluorescence of ANS. Since ANS does not bind to fully folded or full unfolded barstar (Khurana & Udgaonkar, 1994), the occurrence of ANS binding during the refolding of barstar provides a method of identifying kinetic intermediates (Figures 7 to 9). Kinetic intermediates that bind ANS are expected to be similar to molten globule forms of proteins (Ptitsyn, 1994), and the rates of formation and disappearance of such intermediates can be monitored by monitoring the fluorescence change that occurs when ANS binds to or dissociates from these intermediates.

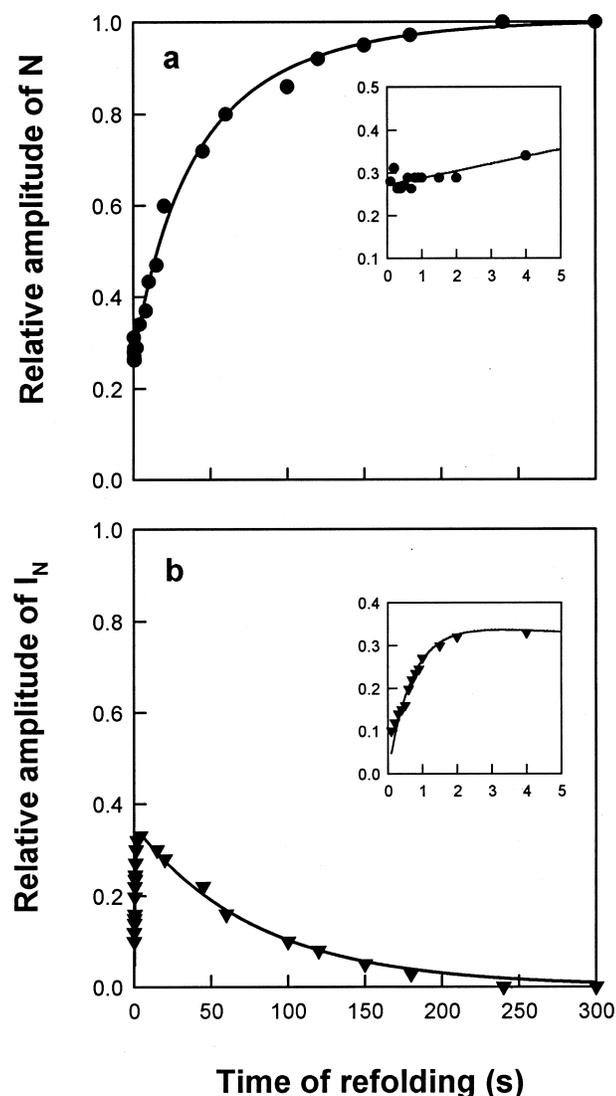


Figure 6. Kinetics of folding of equilibrium-unfolded barstar in 0.6 M GdnHCl, pH 7, 25°C. The amounts of N (a) and I_N (b) were determined at different times during the folding process by means of double-jump experiments, in which various forms of the protein were identified by unfolding assays (see Materials and Methods). The amounts of N and I_N are plotted relative to the total amount of N that is formed when the folding reaction is allowed to go to completion. Each time-point was repeated in triplicate, and the standard errors in the determination of N and I_N were less than $\pm 15\%$ of the values shown. The continuous lines through the data are non-linear least squares fits of the data to Mechanism III in Scheme I, using equations (A1) to (A6). The value of k_2 was fixed to 25 s^{-1} , and the values obtained for k_3 , k_4 and k_5 are given in Table 1. The insets show the data and fits on a shorter time-scale.

ANS has no effect on the folding kinetics of barstar

Figure 7a shows that the kinetics of folding of barstar (20 μM), as monitored by change in intrinsic tryptophan fluorescence at 320 nm, are not affected by the presence of 50 μM ANS during folding in 0.6 M GdnHCl. In Figure 7b and c, the effect of

varying the concentration of ANS during refolding on the tryptophan fluorescence-monitored kinetics is shown. Figure 7b shows that the rate constants of the fast and the slow reactions, λ_2 and λ_1 , respectively, are not affected by ANS concentrations up to 125 μM . Above this concentration of ANS, λ_2 increases marginally. Figure 7c shows that the reduced amplitudes of the fast and the slow phases are not affected by ANS concentrations up to 125 μM , after which the amplitudes decrease, possibly because of the inner filter effect. Thus, the maximum concentration of ANS that can be used without altering the folding kinetics of barstar was determined to be 125 μM .

ANS binds to two distinct very early intermediates

Figure 8a shows the kinetic trace obtained when folding in 0.6 M GdnHCl was followed by monitoring the change in ANS fluorescence during the folding process. When ANS is present during refolding, it binds to the folding protein within 2 ms, the dead-time of mixing in the stopped-flow machine, and consequently, the rise in ANS fluorescence that accompanies binding cannot be monitored. The decrease in ANS fluorescence that occurs as folding proceeds is, however, easily monitored (Figure 8a) and occurs in two phases in 0.6 M GdnHCl (Figure 8a), a fast phase with a rate constant of $23(\pm 3) \text{ s}^{-1}$ and a slow phase characterized by a rate constant of $7(\pm 3) \times 10^{-3} \text{ s}^{-1}$. This indicates that there are two distinct transient populations of folding barstar molecules that bind to ANS within the mixing dead-time. One population has a mean life-time of less than 50 ms, while the other has a mean life-time of approximately 140 seconds.

ANS concentration-dependence of ANS fluorescence-monitored kinetics.

Figure 8b shows that the rate constants of both ANS fluorescence-monitored kinetic phases are independent of the concentration of ANS used. The mean value for the fast rate constant is $23(\pm 3) \text{ s}^{-1}$ and that of the slow rate constant is $7(\pm 3) \times 10^{-3} \text{ s}^{-1}$, over the entire range of ANS concentrations used. In Figure 8c, it is shown that the amplitudes of both of the ANS fluorescence-monitored phases increase with an increase in ANS concentration, when the protein is refolded in 0.6 M GdnHCl. The ANS concentration-dependent data were fit to equation (3) to obtain the dissociation constants for binding. Binding to the molecules that undergo a fast folding reaction is characterized by a dissociation constant of 100 μM , while binding to those that undergo a slow folding reaction is characterized by a dissociation constant of 125 μM . The quantum yield of ANS fluorescence when ANS is bound to the fast-transforming intermediate is higher than when it is bound to the slow-transforming intermediate (Figure 8c).

GdnHCl dependence of ANS fluorescence-monitored kinetics.

Figure 9 shows the dependences of the relative amplitudes and the rate constants of the two phases observed on the concentration of GdnHCl present during refolding. Figure 9a shows that the amplitude of the fast phase of change in ANS fluorescence decreases continuously from 0.6 M GdnHCl and reduces to zero at around 1.0 M GdnHCl. At concentrations of GdnHCl greater than 1.0 M, only the slow phase of change in ANS fluorescence is observed, and the decay in ANS fluorescence was fitted to a single exponential. The amplitude of the slow phase increases as the concentration of

GdnHCl is increased from 0.6 M, and is at a maximum in 1.0 to 1.2 M GdnHCl, after which it decreases continuously and becomes zero in 2.0 M GdnHCl.

Also shown in Figure 9a are the maximum levels to which I_N is observed to be populated as a function of the concentration of GdnHCl present during refolding. The population of I_N is maximum at a level of 70% in 1.0 M (Figure 4) and 1.2 M GdnHCl (Shastry *et al.*, 1994), and is less in both lower and higher GdnHCl concentrations. The dependence of I_N on GdnHCl concentration when measured using the double-jump experiments of Figures 4 to 6 is the same as that of the population of molecules from which ANS dissociates slowly.

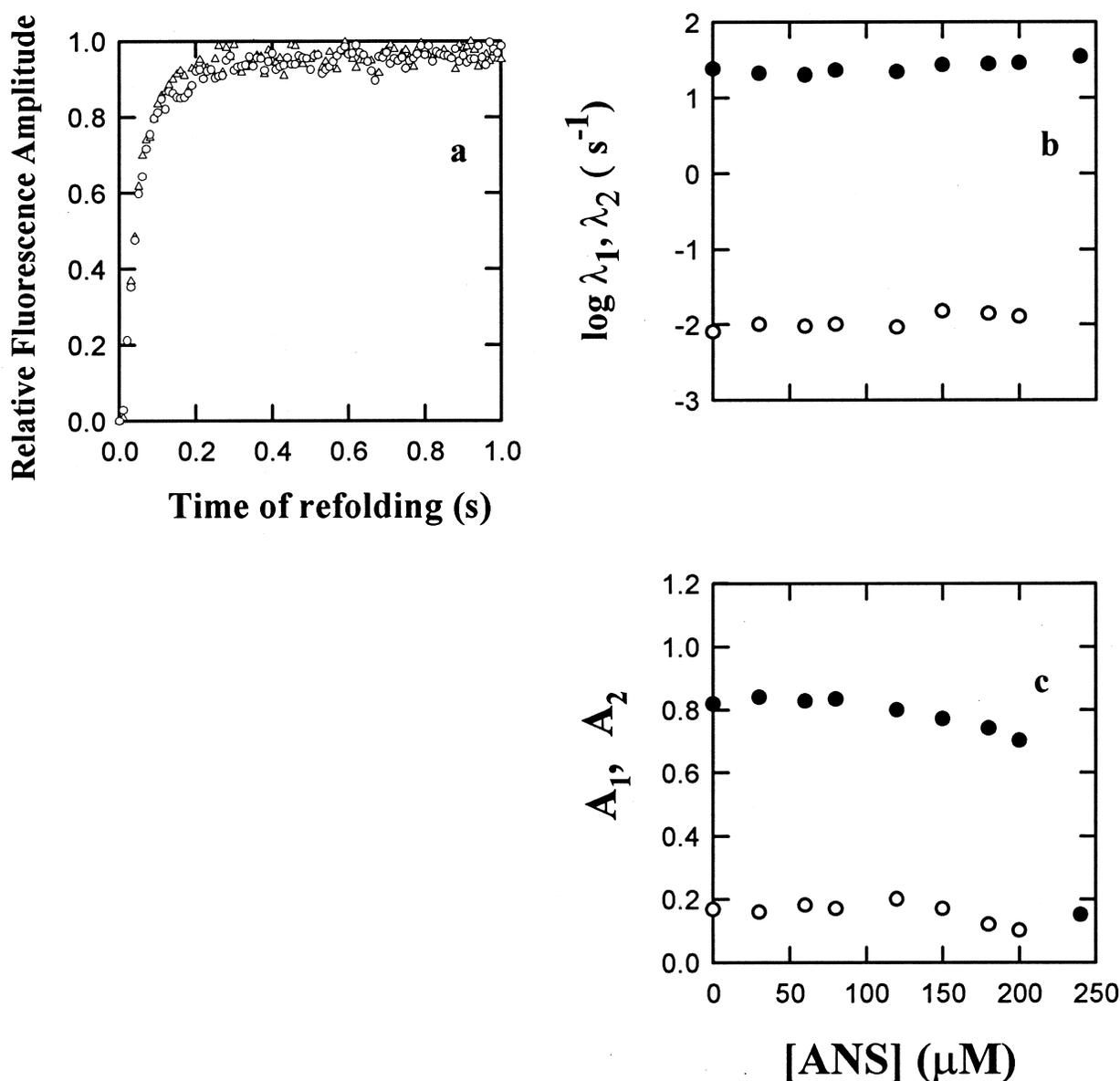


Figure 7. Effect of ANS on the tryptophan fluorescence-monitored kinetics of folding of barstar in 0.6 M GdnHCl at pH 7, 25°C. a, The fluorescence at 320 nm (in arbitrary units) on excitation at 278 nm is plotted as a function of the time of refolding in 0.6 M GdnHCl in the absence (○) or presence (△) of 50 μM ANS. b, Dependence of the fast (λ_2 , ●) and the slow (λ_1 , ○) rate constants, on ANS concentration. c, Dependence of the reduced amplitudes of the fast (A_2 , ●) and slow (A_1 , ○) phases on ANS concentration.

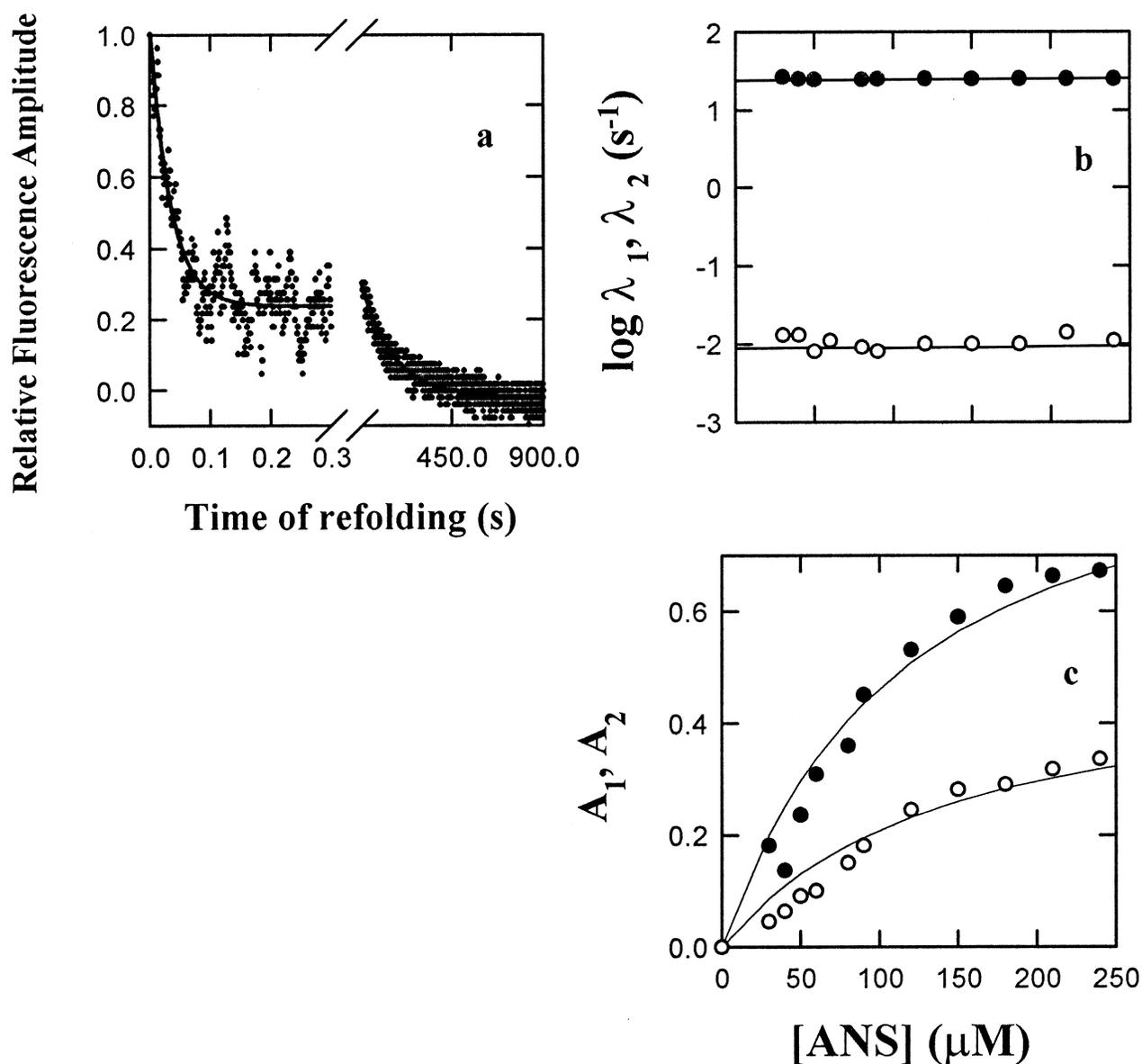


Figure 8. a, ANS fluorescence-monitored kinetics at pH 7, 25°C. Barstar that had been unfolded to equilibrium in 6.0 M GdnHCl was refolded in 0.6 M GdnHCl in the presence of 50 μM ANS. The kinetics of folding were followed by monitoring total ANS fluorescence above 450 nm. The kinetic trace was fitted using non-linear least-squares analysis to equation (1), and the values obtained for λ_1 , λ_2 , A_1 and A_2 are $6 \times 10^{-3} \text{ s}^{-1}$, 26 s^{-1} , 0.1 and 0.24, respectively. b and c, Dependence of ANS fluorescence-monitored folding kinetics on ANS concentration at pH 7, 25°C. Equilibrium-unfolded barstar (in 6.0 M GdnHCl) was refolded in 0.6 M GdnHCl in the presence of the concentrations of ANS indicated, and the folding kinetics measured by monitoring total ANS fluorescence above 450 nm. b, Dependence of λ_1 (\circ) and λ_2 (\bullet) on ANS concentration. Continuous horizontal lines through the data in b represent the ANS concentration-independent mean values for λ_1 and λ_2 , of $7(\pm 3) \times 10^{-3} \text{ s}^{-1}$ and $23(\pm 3) \text{ s}^{-1}$, respectively. c, Dependence of A_1 (\circ) and A_2 (\bullet) on the ANS concentration. The continuous lines through the data for A_1 and A_2 in c are non-linear least-squares fits of the data to equation (3), and yielded values for K_D of 100 μM for the A_2 data and 120 μM for the A_1 data.

Figure 9b shows the variation of the corresponding rate constants of the two phases. Both the fast and slow rate constants are observed to be independent of GdnHCl. The mean value of the fast rate constant in the range of GdnHCl concentrations between 0.6 M and 1.0 M is $23(\pm 3) \text{ s}^{-1}$, while the mean value of the slow rate constant in the range of GdnHCl concentrations between 0.6 M and 2.0 M is $7(\pm 3) \times 10^{-3} \text{ s}^{-1}$.

ANS binding does not occur on the folding pathway of U_F

After five seconds of unfolding in 3.7 M GdnHCl, 95% of barstar molecules exist as U_F (see Materials and Methods). U_F obtained in this way was refolded in 0.6 M GdnHCl in the presence of 50 μM ANS. No change in ANS fluorescence was observed, indicating that no intermediate capable of binding ANS accumulates on the folding pathway of U_F .

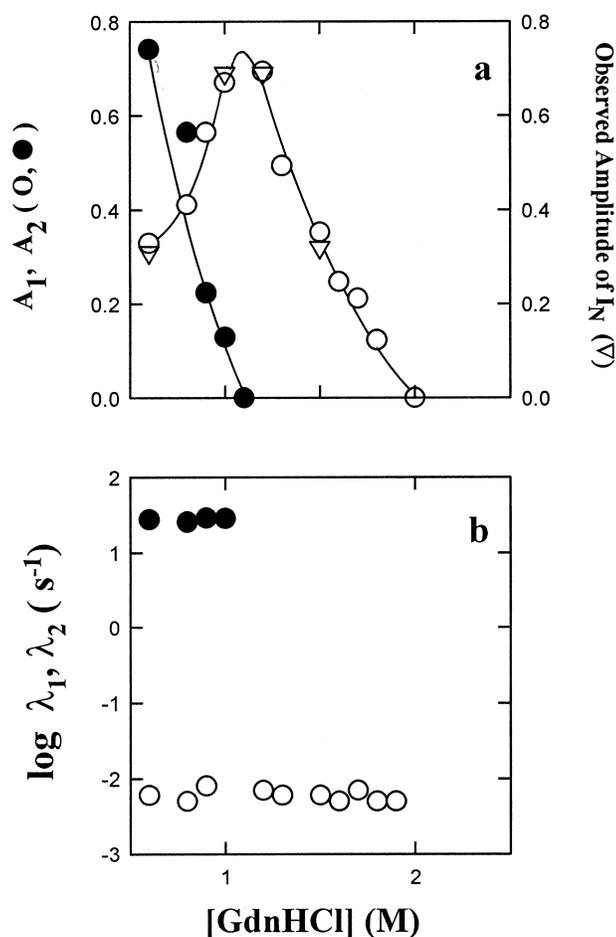


Figure 9. Dependence of the ANS fluorescence-monitored folding kinetics on the concentration of GdnHCl at pH 7, 25°C. Barstar that had been unfolded to equilibrium in 6 M GdnHCl was refolded in the presence of the concentrations of GdnHCl indicated, and in the presence of 50 μ M ANS. The folding kinetics were monitored by measuring the total ANS fluorescence above 450 nm. (○) Kinetics of the slow phase; (●) kinetics of the fast phase. a, Dependences of the observed amplitudes of the 2 kinetic phases. The fluorescence amplitudes are shown relative to the total change in ANS fluorescence observed (i.e. the sum of the fluorescence amplitudes of both the fast and slow phases) when the folding reaction was initiated by jumping the GdnHCl concentration from 6.0 M to 0.6 M, and are therefore analogous to the reduced amplitudes of tryptophan fluorescence-monitored kinetics (Figure 2). Shown in a also is the dependence of the observed maximum population of I_N (∇) on the concentration of GdnHCl at pH 7, 25°C. The observed maximum of I_N in 0.6 M, 1.0 M and 1.5 M GdnHCl are from Figures 4 to 6, and in 1.2 M from Shastry *et al.* (1994). The lines through the points have been drawn by inspection only. b, Dependences of the 2 observed rate constants.

Discussion

Folding of U_F versus folding of an equilibrium mixture of $U_F + U_S$

There are two notable differences between the kinetics of folding U_F and folding an equilibrium mixture of U_F and U_S . Firstly, the mid-point of the

GdnHCl melting curve defined by the amplitudes of the fast folding reaction increases from a value of 1.3 M for the folding of U_F to a value of 2.0 M for the folding of $U_F + U_S$ (Figure 3a). Secondly, no slow folding reaction is seen; instead, a slow unfolding reaction is observed when folding in the transition zone commences from U_F (Figure 1). The unfolding reaction, which is driven by equilibration between U_F and U_S (Kiefhaber *et al.*, 1992) results in the negative amplitude of the slow folding reaction in the transition zone (Figure 2). U_F first folds rapidly to N, and the final equilibrium mixture of N and U seen at any concentration of GdnHCl in the transition zone can be reached only when some of the accumulated N unfolds. Thus, the total observed amplitude (the sum of fast and slow amplitudes, the latter being negative for folding commencing from U_F and positive for folding commencing from an equilibrium mixture of U_F and U_S) is the same whether folding commences from U_F or from an equilibrium mixture of U_F and U_S (Figure 2c).

I_{F1} forms rapidly on the folding pathway of U_F

The data in Figure 3 indicate that an intermediate I_{F1} forms on the folding pathway of U_F when folding occurs in concentrations of GdnHCl less than 1.0 M. The presence of I_{F1} has been inferred only indirectly, and there is no direct information on the rate of its formation, but it must form very fast from U_F , much faster than the subsequent transformation to N.

Fast folding reactions occur on the folding pathway(s) of U_S

Equilibrium-unfolded barstar is composed of 69% U_S and 31% U_F (Shastry *et al.*, 1994; Figure 2a). When folding kinetics are monitored by the change in intrinsic tryptophan fluorescence, the value of the rate constant of the fast refolding step, λ_2 , is the same whether folding begins from U_F alone (34 s^{-1}), or from an equilibrium mixture of U_F and U_S (37 s^{-1}). The relative amplitude of the fast folding reaction characterized by λ_2 is greater than 31% when folding is carried out in the presence of any concentration of GdnHCl less than 2.0 M (Shastry *et al.*, 1994), and in 0.6 M GdnHCl it is 95%. Since only 31% of the molecules fold on the $U_F \rightarrow N$ pathway, and U_F and U_S have the same fluorescence properties, this results suggests that the fast change in tryptophan fluorescence occurs on the folding pathway of U_F and of U_S . U_S folds slowly compared with U_F only because *trans-cis* isomerization of a X-Pro bond must occur when U_S but not U_F refolds. This isomerization reaction appears to be silent to fluorescence change (Schreiber & Fersht, 1993; Shastry *et al.*, 1994).

Folding and unfolding in >2.0 M GdnHCl

No intermediate is seen in folding or unfolding reactions in concentrations of GdnHCl greater than 2.0 M, and the $U_S \rightleftharpoons U_F \rightleftharpoons N$ mechanism adequately

accounts for all the kinetic data (Shastry *et al.*, 1994). This mechanism is not valid for GdnHCl concentrations lower than 2.0 M, which is not surprising because intermediates are expected to accumulate in more stable folding conditions. Thus, as the GdnHCl concentration is reduced from 2.0 M, alternative pathways defined by partly folded kinetic intermediates become available for folding, and the contribution of the $U_S \rightleftharpoons U_F \rightleftharpoons N$ mechanism to folding diminishes.

Folding in 1.0 M-2.0 M GdnHCl

It has been shown that the native-like intermediate I_N accumulates on the folding pathway of U_S (Shastry *et al.*, 1994) under such marginally stable conditions. It is shown here that the folding reactions that occur in these conditions, $U_S \rightleftharpoons U_F \rightleftharpoons N$ and $U_S \rightarrow I_N \rightarrow N$ (see Scheme 1) account for the kinetic data when the GdnHCl concentration is decreased from 2.0 M to 1.0 M, from the midpoint to the beginning of the equilibrium folding transition. With a decrease in GdnHCl concentration, the $U_S \rightarrow I_N \rightarrow N$ pathway becomes more important at the cost of the $U_S \rightleftharpoons U_F \rightleftharpoons N$ pathway, because I_N becomes more stable at lower GdnHCl concentrations resulting in an increase in the rate of the $U_S \rightarrow I_N$ reaction, while the rate of the competing $U_S \rightarrow U_F$ transition does not change. Thus, in 1.5 M GdnHCl there is no apparent lag in the formation of N because the $U_S \rightleftharpoons U_F \rightleftharpoons N$ pathway productively competes with the $U_S \rightarrow I_N \rightarrow N$ pathway.

The productive competition of the $U_S \rightleftharpoons U_F \rightleftharpoons N$ pathway is also the reason why I_N is maximally populated to an extent of only 24% in 1.5 M GdnHCl. Another contributing factor is that the rate of formation of I_N is only twice the rate of its conversion to N (Figure 5). Thus, I_N starts disappearing before it can reach its maximal population.

As the concentration of GdnHCl nears 1.0 M, the $U_S \rightarrow I_N$ reaction becomes so fast that the $U_S \rightleftharpoons U_F \rightleftharpoons N$ pathway plays an insignificant role, and I_N is maximally populated. In 1.0 M (Figure 4) or 1.2 M GdnHCl (Shastry *et al.*, 1994), all molecules of U_S (70% of all molecules) fold to I_N .

Intermediate I_{S1} accumulates on the $U_S \rightarrow I_N \rightarrow N$ pathway

The rate of the fast change in tryptophan fluorescence that occurs on the folding pathway of U_S (see above) decreases from 25 s^{-1} in 0.6 M GdnHCl to 12 s^{-1} in 1.5 M GdnHCl, 25-fold and 250-fold faster, respectively, than the rate of formation of I_N (see Figures 4 to 6). This suggests that U_S does not fold directly to I_N but first folds to an intermediate, I_{S1} . The folding pathway of U_S in which I_N is populated must therefore be $U_S \rightarrow I_{S1} \rightarrow I_N \rightarrow N$, with the fast change in tryptophan fluorescence occurring during the $U_S \rightarrow I_{S1}$ step. The double-jump experiments monitoring the formation of I_N (Figures 4 to 6) measure the kinetics of the $I_{S1} \rightarrow I_N$ step.

Folding in 0.6 M GdnHCl

In 0.6 M GdnHCl, I_N is again not populated to the maximal extent of 70%, but only to a level of 35%. The rate of formation of I_N is also less than that in 1.0 M GdnHCl. There are several possible explanations. (1) It is possible that I_N is less stable in 0.6 M than in 1.0 M GdnHCl (the $I_{S1} \rightleftharpoons I_N$ equilibrium shifts to favour I_{S1}), which would be consistent with the rate of its accumulation being slower. The observed rate of accumulation of I_N is, however, more than 50-fold faster than its conversion to N (Figure 6), and hence, I_N would still be expected to be populated maximally. Thus, the rate of formation of I_N being slower in 0.6 M than in 1.0 M GdnHCl cannot be the reason why I_N is not populated maximally. Moreover, the stability of N is lowered more in 1.0 M GdnHCl than in 0.6 M GdnHCl (unpublished results) and it is unlikely that the stability of I_N would show an opposite trend. (2) It is possible that the double-jump experiments do not correctly assay for the amount of I_N in different GdnHCl concentrations because of a strong GdnHCl-dependence of fluorescence intensity of I_N . No such strong dependence is seen for N, however, making such an explanation unlikely. (3) The third and most likely explanation is that there is an alternative, competing pathway for the folding of U_S for folding in <1.0 M GdnHCl.

Evidence for two folding pathways of U_S from double-jump experiments

When folding in 0.6 M GdnHCl is monitored by double-jump assays (Figure 6), the amplitudes corresponding to I_N and N do not account for the total folding amplitude. The residual amplitude decreases to 0 only after 50 seconds, after which I_N and N are the only two forms present. The data in Figure 6 indicate that the residual amplitude is due to the presence of two different, less folded forms because its decrease follows a two-exponential process. One form clearly folds directly to I_N (it disappears at the rate at which I_N appears) with a rate constant of 1.0 s^{-1} , and the other appears to fold directly to N with a rate constant of 0.07 s^{-1} (see Results). The folding of U_S in concentrations of GdnHCl below 1.0 M can therefore be represented by two different pathways: $U_S \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ and $U_S \rightarrow I_{S2} \rightarrow N$. The fast change in tryptophan fluorescence that occurs during the latter pathway occurs during the $U_S \rightarrow I_{S2}$ step, as it does for the $U_S \rightarrow I_{S1}$ step for the former pathway (see above). The $I_{S2} \rightarrow N$ reaction occurs at the slow rate of disappearance of the residual amplitude (0.07 s^{-1}).

Evidence for two folding pathways of U_S from ANS binding experiments

The data in Figures 8 and 9 indicate that there are two distinct kinetic intermediates that bind ANS. Although, the formation of both of these intermediates is too fast to be monitored by stopped-flow kinetic measurements, their disappearance is easily

followed (Figure 7). One intermediate (I_{M2}) accumulates only in the presence of GdnHCl concentrations less than 1.0 M, while the other (I_{M1}) is not populated when the GdnHCl concentration during folding is >2.0 M.

I_{M1} and I_{M2} must be on two separate folding pathways

If I_{M1} and I_{M2} were on the same folding pathway, then the populations of both are expected to decrease with increasing GdnHCl concentration. Instead, the population of I_{M1} increases as the GdnHCl concentration is increased from 0.6 M to 1.0 M, at the expense of the population of I_{M2} , which decreases to 0 at 1.0 M GdnHCl (Figure 9).

I_{M1} is on the $U_S \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ pathway

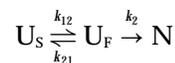
The rate constant for the disappearance of ANS fluorescence bound to I_{M1} is independent of GdnHCl and has a value of $7(\pm 3) \times 10^{-3} \text{ s}^{-1}$. The value is similar to that of the apparent rate of the $I_N \rightarrow N$ reaction ($16(\pm 4) \times 10^{-3} \text{ s}^{-1}$), which is also independent of GdnHCl concentration. I_{M1} is detected in the same range of GdnHCl concentrations (0.6 to 2.0 M) as is I_N (Figure 9). This suggests that I_{M1} is on the $U_S \rightarrow I_{M1} \rightarrow I_N \rightarrow N$ pathway. The fast change in tryptophan fluorescence occurs during the $I_{S1} \rightarrow I_N$ step, and the ANS that initially binds to I_{M1} dissociates from the folding protein molecules only during the last detected step, the $I_N \rightarrow N$ transition. This suggestion is supported by the observation that the maximum amplitude of the change in ANS fluorescence associated with ANS binding to I_{M1} occurs at the concentration of GdnHCl (1.0 M) at which the maximum accumulation of I_N is seen in double-jump experiments (Figure 9a).

The native-like intermediate, I_N , is capable, like N , of inhibiting the activity of barnase (Schreiber & Fersht, 1993). The observation that it possesses solvent-accessible hydrophobic clusters to which ANS can bind is surprising. It suggests that while regions in barstar that interact with barnase have correctly folded in I_N , the rest of the structure is still open to the solvent. Solvent is excluded from the hydrophobic core only after *trans-cis* isomerization of the Tyr47-Pro48 bond occurs during the $I_N \rightarrow N$ transition.

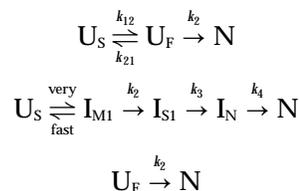
I_{M2} is not on the $U_F \rightarrow N$ pathway

Figure 3 indicates that an early intermediate I_{F1} accumulates on the $U_F \rightarrow N$ pathway when refolding is carried out in <1.0 M GdnHCl. Figure 9 indicates that I_{M2} too is present only when the concentration of GdnHCl is less than 1.0 M. Moreover, the rate constant for the disappearance of I_{M2} has a value of $23(\pm 3) \text{ s}^{-1}$, which corresponds to the rate of the $I_{F1} \rightarrow N$ reaction ($\lambda_2 = 25(\pm 3) \text{ s}^{-1}$) when folding in 0.6 M GdnHCl is monitored by the change in intrinsic tryptophan fluorescence. These observations suggested that I_{M2} might be identical with I_{F1} . No fast (or slow) phase of ANS release is, however,

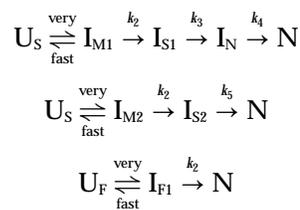
Mechanism I. Folding in >2.0 M GdnHCl:



Mechanism II. Folding in 1.0 to 2.0 M GdnHCl:



Mechanism III. Folding in <1.0 M GdnHCl:



Scheme I. Folding mechanisms of barstar. All folding reactions, except for the $U_S \rightleftharpoons U_F$ conversion, have been assumed to be irreversible, and have been analysed using the equations in the Appendix. The values of the rate constants k_{12} and k_{21} , $5.3 \times 10^{-3} \text{ s}^{-1}$ and $11.7 \times 10^{-3} \text{ s}^{-1}$, respectively, are independent of the concentrations of GdnHCl (Shastry *et al.*, 1994).

detected (see Results) when folding is commenced from only U_F molecules. Thus, I_{M2} cannot be on the $U_F \rightarrow N$ pathway, and is distinct from I_{F1} .

I_{M2} is on the $U_S \rightarrow I_{S2} \rightarrow N$ pathway

Both I_{M2} and I_{S2} accumulate only when the concentration of GdnHCl is less than 1.0 M. The simplest way to account for this observation, after first discounting the presence of I_{M2} on both the $U_S \rightarrow I_{M1} \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ and $U_F \rightarrow I_{F1} \rightarrow N$ pathways (see above), is that both intermediates are on the same pathway. Thus, in the $U_S \rightarrow I_{M2} \rightarrow I_{S2} \rightarrow N$ pathway, I_{M2} is formed within 2 ms and the fast change in tryptophan fluorescence as well as the release of ANS, both characterized by a rate constant of $25(\pm 3) \text{ s}^{-1}$, occur during the $I_{M2} \rightarrow I_{S2}$ step. Complete folding to N occurs in the $I_{S2} \rightarrow N$ step that, like the $I_N \rightarrow N$ step on the $U_S \rightarrow I_{M1} \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ pathway, is silent to fluorescence change.

The double-jump experiments in conjunction with the ANS binding experiments therefore strongly suggest that an alternative competing pathway is present when refolding is carried out in <1.0 M GdnHCl. This pathway accounts for I_{M2} , which together with I_{S2} , defines it. Other explanations for why I_N does not accumulate to its expected level (see above) cannot account for the presence of I_{M2} .

Folding mechanism of barstar

The minimal folding mechanism of barstar is depicted in Scheme I. U_F folds to N by a single

pathway, while at least three alternative pathways are available for the folding of U_s . The pathway(s) by which U_s molecules fold depends on the concentration of GdnHCl present. The $U_s \rightarrow I_{M2} \rightarrow I_{S2} \rightarrow N$ pathway is populated only when folding is carried out in less than 1.0 M GdnHCl. This pathway does not contribute significantly to the folding of U_s , and I_{F1} does not accumulate significantly on the folding pathway of U_F , when folding is carried out in concentrations of GdnHCl greater than 1.0 M. The folding of U_s in concentrations of GdnHCl between 1.0 and 2.0 M proceeds via the $U_s \rightarrow I_{M1} \rightarrow I_{S1} \rightarrow I_N \rightarrow N$ and $U_s \rightleftharpoons U_F \rightleftharpoons N$ pathways with the contribution of the former becoming less significant with an increase in GdnHCl concentration, until finally, at concentrations of GdnHCl greater than 2.0 M, the latter pathway is the only pathway available for U_s to fold.

Are I_{M2} and I_{S2} productive intermediates?

I_N is clearly a productive intermediate leading to the formation of N because the rate at which it disappears is similar to the rate of formation of N (Figures 4, 5 and 6; see also Shastry *et al.*, 1994). It has, however, not been possible to experimentally demonstrate that I_{S2} is a productive intermediate that leads to the formation of N. Although the simulation according to Scheme I correctly predicts all experimentally observed parameters (see Figure 6), the role of I_{S2} , and especially how it is different from I_N , is puzzling. Like I_N , I_{S2} must possess the non-native *trans* conformation of the Tyr47-Pro48 peptide bond.

It seemed intuitively possible that I_{S2} could be a non-productive off-pathway intermediate. A kinetic simulation of a folding scheme was therefore carried out, in which I_{S2} is a non-productive, dead-end intermediate formed reversibly from U_s via the $U_s \rightleftharpoons I_{M2} \rightleftharpoons I_{S2}$ pathway. While this simulation could correctly predict the accumulation of I_N and N, as well as the tryptophan fluorescence-monitored kinetics, it could not correctly predict the results of the ANS binding experiments. Thus, I_{S2} like I_N must be a productive intermediate leading to the formation of N.

Table 1

GdnHCl concentration-dependence of folding kinetics at pH 7, 25°C

Rate constant (s ⁻¹)	Concentration of GdnHCl (M)		
	1.5	1.0	0.6
k_2	12	18	25
k_3	0.04	fast	1.0
k_4	0.017	0.017	0.017
k_5	—	—	0.07

k_2 , k_3 , k_4 and k_5 are defined in Scheme I. The values for k_2 are the values determined for λ_2 , the rate constant of the fast change in tryptophan fluorescence that accompanies folding. The values of k_3 , k_4 and k_5 were obtained from non-linear least-squares fits of the data in Figures 4 to 6 to equations (A1) to (A6)

Structures of the intermediates I_{M1} and I_{M2}

A large body of experimental data indicates that formation of the native structure of a protein occurs via compact intermediates with pronounced secondary structure (Dolgikh *et al.*, 1985; Kim & Baldwin, 1990; Matthews, 1993; Ptitsyn, 1994). Usually, a large part of the protein secondary structure forms very fast (in <10 ms) concurrently with a collapse of the polypeptide chain into essentially native-like compactness and the formation of the rigid tertiary structure of protein proceeds more slowly (Kuwajima *et al.*, 1987). Folding intermediates that possess a pronounced secondary structure and are not compact have not been detected. Nevertheless, such intermediates might accumulate during protein refolding under strongly native conditions. At present, there is no information available on the compactness or secondary structure content of the early intermediates I_{M1} and I_{M2} (or I_{F1} , I_{S1} and I_{S2}), but stopped-flow circular dichroism experiments in progress in this laboratory will provide useful information.

Barstar forms an equilibrium molten globule-like form, the A form, at low pH (Khurana & Udgaonkar, 1994; Swaminathan *et al.*, 1994). ANS binding to the A form is characterized by a dissociation constant of 10 μ M (Khurana & Udgaonkar, 1994). The affinities of the kinetic intermediates I_{M1} and I_{M2} for ANS are approximately tenfold lower (Figure 8c). The lower affinity at neutral pH as compared with acid pH is probably because barstar is negatively charged at pH 7 and positively charged at pH 3, because ANS is negatively charged. So far, it has not been possible to determine whether I_{M1} or I_{M2} is structurally related to the A form. In the case of α -lactalbumin, the A state formed at low pH appears to be populated on the kinetic pathway of folding (Kuwajima *et al.*, 1985).

Rates of ANS release monitor conformational changes

The fluorescence change that accompanies ANS dissociating from a folding intermediate has been used to monitor the disappearance of the intermediate on the folding pathway. It is crucial that the binding of ANS to, as well as its dissociation from, the intermediate must be fast with respect to the conformational changes occurring in the intermediate, otherwise the kinetics of the change in ANS fluorescence will not reflect the kinetics of the conformational changes. The bimolecular rate constants for the binding of small molecules to proteins are typically in the range of 10^6 to 10^8 M⁻¹ s⁻¹ (Hammes, 1982), less than the expected diffusion-controlled rate constant, which is typically 10^9 to 10^{10} M⁻¹ s⁻¹. Thus, ANS at a concentration of 50 μ M is expected to bind with a rate of >50 s⁻¹. The equilibrium constant for the dissociation of ANS is approximately 100 μ M (Figure 8c), which requires a dissociation rate of >100 s⁻¹. This dissociation rate is more than fourfold faster than the fast phase observed in ANS fluorescence-monitored kinetics.

Thus, the change in ANS fluorescence truly monitors conformational changes occurring on the folding pathway.

Significance of the alternative pathway

Studies that utilize amide-hydrogen exchange in conjunction with proton NMR to study folding pathways have identified the presence of transient kinetic barriers (Udgaonkar & Baldwin, 1990; Baldwin, 1993; Sosnick *et al.*, 1994) to folding in the case of both ribonuclease A (Udgaonkar & Baldwin, 1990) and ribonuclease T₁ (Mullins *et al.*, 1993). In the former case, the barrier leads to an early intermediate not accumulating to its expected level, and it was suggested that the barrier might be the presence of one or more *cis* peptide bonds other than at proline residues (Udgaonkar & Baldwin, 1990). In the latter case, it was suggested that a rapid hydrophobic collapse starting from the same population of unfolded molecules leads to the formation of two molten globule forms, one structured and the other unstructured, with the latter folding by a separate pathway (Mullins *et al.*, 1993). In this work, it has been demonstrated that in strongly native-like conditions, the U_S state of barstar folds very rapidly (within 2 ms) to two distinct molten globule-like forms, I_{M1} and I_{M2}, which then fold to native protein by separate pathways. The presence of two pathways for the folding of U_S appears to be the reason why I_N does not populate to its expected level. NMR characterization of the folding intermediates of barstar, utilizing amide hydrogen exchange, has been initiated in this laboratory, and is expected to provide greater insight into their structures and roles on the folding pathway of barstar.

Materials and Methods

Protein purification

Barstar was expressed in *Escherichia coli* strain MM294 using the plasmid pMT316, a kind gift from R. W. Hartley (1988). The method used to purify barstar has been described in detail (Khurana & Udgaonkar, 1994). Protein concentrations were measured using an extinction coefficient of 23,000 M⁻¹ cm⁻¹ (Khurana & Udgaonkar, 1994).

Buffers and solutions

Native buffer is 5 mM sodium phosphate (pH 7), 250 μM EDTA, 0.25 mM DTT. Unfolding buffer is native buffer with 6 M GdnHCl (ultra-pure GdnHCl from Sigma). The concentrations of stock solutions of GdnHCl were determined by measuring the refractive index using an Abbe 3L refractometer. For ANS binding experiments, a 10 mM stock solution of ANS (obtained from Sigma) was prepared, and the concentration was determined using an extinction coefficient of 5000 M⁻¹ cm⁻¹ at 350 nm (Stryer, 1965). All experiments were done in reducing conditions, i.e. in the presence of DTT, which prevents a small amount (<5%) of dimer formation through an inter-molecular disulphide bond that is seen in non-reducing conditions (Shastry *et al.*, 1994).

Kinetic experiments

All kinetic experiments, both single-jump and double-jump, were performed on a Biologic SFM-3 instrument. Folding and unfolding were monitored using two different probes; (1) the intrinsic tryptophan fluorescence, and (2) the change in fluorescence upon binding of ANS. Tryptophan fluorescence was excited at 287 nm using a bandwidth of 20 nm, and monitored at 320 nm using an Oriel band-pass filter with a bandwidth of 10 nm. For experiments in which the binding of ANS was monitored, an excitation wavelength of 350 nm was used with a bandwidth of 20 nm, and the total fluorescence emission above 450 nm was monitored using an Oriel 450 nm low-cut filter. The protein concentration during refolding for tryptophan fluorescence-monitored experiments was typically 2 μM.

Intrinsic tryptophan fluorescence-monitored folding experiments

Folding of equilibrium-unfolded barstar. Barstar was unfolded in 6 M GdnHCl (unfolding buffer) for at least three hours. Single-jump refolding experiments were then performed in which the concentration of GdnHCl was reduced to between 0.6 M and 2.3 M by appropriate dilution in the stopped-flow machine (Shastry *et al.*, 1994).

Folding of U_F (transiently unfolded barstar). Barstar in native buffer was unfolded in 3.7 M GdnHCl for five seconds by a ten-fold dilution into unfolding buffer. In 3.7 M GdnHCl barstar unfolds completely to U_F with a rate constant of 4(±1) s⁻¹, and the rate of the U_F → U_S reaction is only 11.7 × 10⁻³ s⁻¹ (Shastry *et al.*, 1994). Thus, five seconds after unfolding in 3.7 M GdnHCl, 95% of the barstar molecules are present as U_F. The concentration of GdnHCl was then reduced to between 0.6 M and 2.5 M by a second dilution of the transiently unfolded protein into native buffer, and the refolding was monitored in the stopped-flow machine.

Double-jump experiments. The relative amounts of I_N and N at any time in the folding process between 0.1 and 600 seconds, were determined by double-jump experiments (Schmid, 1983) as described (Shastry *et al.*, 1994). Equilibrium-unfolded barstar in 6 M GdnHCl (see above) was diluted to a final GdnHCl concentration of either 0.6 M, 1.0 M or 1.5 M. After various times (0.1 to 600 seconds) of refolding, the protein solution was diluted once more to raise the GdnHCl concentration to 3.7 M. The rates and amplitudes of the two kinetic phases observed in the re-unfolding reaction were monitored in the stopped-flow machine. The rates observed in this unfolding assay for the amounts of I_N and N were used to identify the form of barstar unfolding (N and I_N unfold at rates of 4(±1) s⁻¹ and 28(±4) s⁻¹, respectively, in 3.7 M GdnHCl (Shastry *et al.*, 1994)), and the relative amplitudes indicated the relative amounts of N and I_N present at the time of initiation of re-unfolding in 3.7 M GdnHCl. The relative amplitudes were compared with the amplitude of the unfolding reaction observed when the unfolding assay was performed after refolding in 0.6 M GdnHCl was allowed to proceed for 900 seconds, when N is the only form of the protein present.

ANS fluorescence-monitored folding experiments.

Folding of equilibrium-unfolded barstar. The concentration of ANS was varied between 10 and 250 μM by first mixing

native buffer containing 300 μM ANS with native buffer without ANS in different proportions. Then 30 μM of equilibrium-unfolded barstar (in 6 M GdnHCl) was diluted tenfold into 270 μl of the ANS-containing native buffer in a second mixing step. This procedure ensured that the protein concentration (20 μM) and GdnHCl concentration (0.6 M) during the refolding reaction in the presence of different concentrations of ANS did not change. For GdnHCl concentration-dependent studies also, the stopped-flow machine was used in the sequential mixing mode. The concentration of GdnHCl was varied between 0.6 M and 2.5 M by first mixing unfolding buffer (6.0 M GdnHCl) containing 55 μM ANS with native buffer also containing 55 μM ANS in different proportions. Then 30 μl of unfolded barstar (in 6 M GdnHCl) was diluted tenfold into 270 μl of solution containing 55 μM ANS and a variable concentration of GdnHCl in the second dilution step. This procedure ensured that the concentration of protein (20 μM) as well as of ANS (50 μM) did not change during the refolding reactions in different concentrations of GdnHCl.

Folding of U_F . Barstar in native buffer was unfolded in 3.7 M GdnHCl for five seconds by means of tenfold dilution into unfolding buffer. The concentration of GdnHCl was then reduced to 0.6 M by a second appropriate dilution of the transiently unfolded protein into native buffer containing 50 μM (final) ANS, and the refolding was monitored in the stopped-flow machine.

Data analysis

All data were fit using the Biologic Biokine software or the SigmaPlot version 4.02 software.

Raw fluorescence amplitude data obtained from single-jump kinetic experiments were first converted to reduced amplitude data by dividing the observed amplitude of each kinetic phase by the total fluorescence amplitude observed when the folding reaction was performed from the fully unfolded state (6 M GdnHCl) to the fully folded state (0.6 M GdnHCl). The folding kinetics in the pre-transition regions are described by a two-exponential process:

$$A(t) = A(\infty) - A_1 e^{-\lambda_1 t} - A_2 e^{-\lambda_2 t} \quad (1)$$

where $A(t)$ and $A(\infty)$ are the observed reduced amplitudes at times t and at infinity, λ_1 and λ_2 are the apparent rate constants of the slow and the fast phases, and A_1 and A_2 are the respective reduced amplitudes.

The dependence of λ_2 on GdnHCl concentration, in either the pre-transition region or the post-transition region, is given by the following equation (Tanford, 1970):

$$\log \lambda_2 = \log \lambda_2(\text{H}_2\text{O}) + m_{\lambda_2}[\text{D}] \quad (2)$$

where $\lambda_2(\text{H}_2\text{O})$ is the rate constant of the fast phase when the protein is either unfolded or refolded in water, and m_{λ_2} is the slope of the linear dependence of $\log \lambda_2$ on the concentration of GdnHCl.

The dependence of the ANS fluorescence amplitudes on the concentration of ANS in kinetic experiments was fit to equation (3) (de Prat Gay & Fersht, 1994):

$$\Delta F(L) = \frac{[\text{P}] + [\text{L}] + K_D + \sqrt{([\text{P}] + [\text{L}] + K_D)^2 - 4[\text{P}][\text{L}]}}{2} \quad (3)$$

$\Delta F(L)$ is the change in the fluorescence on addition of ANS at concentration $[\text{L}]$ to protein at concentration $[\text{P}]$. Equation (3) describes the binding of ANS to a single site on the protein, and is applicable when the concentration of

ligand is similar to the value of the dissociation constant K_D , when the formation of the complex significantly alters the concentrations of unbound ligand and unbound protein.

Acknowledgements

We thank R. L. Baldwin and the two anonymous referees for their constructive comments on the manuscript. 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 Biotechnology Career Fellowship from the Rockefeller Foundation.

References

- Baldwin, R. L. (1993). Pulsed H/D-exchange studies of folding intermediates. *Curr. Opin. Struct. Biol.* **3**, 84–91.
- Brandts, J. F., Halvarson, H. R. & Brennan, M. (1975). Consideration of the possibility that the slow step in the protein denaturation reaction is due to *cis-trans* isomerization of proline residues. *Biochemistry*, **14**, 4953–4963.
- de Prat Gay, G. & Fersht, A. R. (1994). Generation of family of protein fragments for structure folding studies. 1. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. *Biochemistry*, **33**, 7957–7963.
- Dolgikh, D. A., Abaturov, L. K., Bolotina, I. A., Brazhnikov, E. V., Bushev, V. N., Bychkova, V. E., Gilmanshin, R. I., Lebedev, Y. O., Semisotnov, G. V., Yiktopulo, E. I. & Ptitsyn, O. B. (1985). Compact state of a protein molecule with pronounced small scale mobility: bovine α -lactalbumin. *Eur. J. Biophys.* **13**, 109–121.
- Evans, P. A. & Radford, S. E. (1994). Probing the structure of folding intermediates. *Curr. Opin. Struct. Biol.* **4**, 100–106.
- Garel, J. R. & Baldwin, R. L. (1973). Both the fast- and slow-refolding reactions of ribonuclease A yield native enzyme. *Proc. Nat. Acad. Sci., U.S.A.* **70**, 3347–3350.
- Guillet, V., Laphorn, A., Hartley, R. W. & Mauguen, Y. (1993). Recognition between a bacterial ribonuclease barnase, and its natural inhibitor, barstar. *Structure*, **1**, 165–176.
- Hammes, G. G. (1982). *Enzyme Catalysis and Regulation*, Academic Press, New York.
- Hartley, R. W. (1988). Barnase and barstar. Expression of its cloned inhibitor prevents expression of a cloned ribonuclease. *J. Mol. Biol.* **202**, 913–915.
- Khurana, R. & Udgaonkar, J. B. (1994). Equilibrium unfolding studies of barstar: evidence for an alternative conformation which resembles a molten globule. *Biochemistry*, **33**, 106–115.
- Kiefhaber, T., Kohler, H. A. & Schmid, F. X. (1992). Kinetic coupling between protein folding and prolyl polymerization. I. Theoretical models. *J. Mol. Biol.* **224**, 217–229.
- Kim, P. S. & Baldwin, R. L. (1990). Intermediates in the folding reaction of small proteins. *Annu. Rev. Biochem.* **59**, 631–660.
- Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. (1985). Comparison of the transient folding intermediates in lysozyme and α -lactalbumin. *Biochemistry*, **24**, 874–881.
- Kuwajima, K., Yamaya, H., Miwas, S., Sugai, S. & Nagamura, T. (1987). Rapid formation of secondary structure frame work in protein folding studied by

- stopped-flow circular dichroism. *FEBS Letters*, **221**, 115–118.
- Lubienski, M. J., Bycroft, M., Freund, S. M. V. & Fersht, A. R. (1994). Three-dimensional solution structure and ^{13}C assignments of barstar using nuclear magnetic resonance spectroscopy. *Biochemistry*, **33**, 8866–8877.
- Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature (London)*, **346**, 440–445.
- Matthews, C. R. (1993). Pathways of protein folding. *Annu. Rev. Biochem.* **62**, 653–683.
- Mullins, L., Pace, C. N. & Raushel, F. (1993). Folding intermediates by hydrogen-deuterium amide exchange two-dimensional NMR spectroscopy. *Biochemistry*, **32**, 6152–6156.
- Ptitsyn, O. B. (1994). Kinetic and equilibrium intermediates in protein folding. *Protein Eng.* **7**, 593–596.
- Roder, H., Elöve, G. A. & Englander, S. W. (1988). Structural characterization of the folding intermediate in cytochrome C by H-exchange labelling and proton NMR. *Nature (London)*, **335**, 700–704.
- Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow folding is a sequential reaction via structural intermediates. *Biochemistry*, **22**, 4690–4696.
- Schmid, F. X. (1986). Fast-folding and slow-folding forms of unfolded proteins. *Methods Enzymol.* **131**, 70–82.
- Schreiber, G. & Fersht, A. R. (1993). The refolding of *cis* and *trans*-peptidyl prolyl isomers of barstar. *Biochemistry*, **32**, 11195–11203.
- Semisotnov, G. V., Radionova, N. A., Kutysenko, V. P., Ebert, B., Blanck, J. & Ptitsyn, O. B. (1987). Sequential mechanism of refolding of carbonic anhydrase B. *FEBS Letters*, **224**, 9–13.
- Shastry, M. C. R., Agashe, V. R. & Udgaonkar, J. B. (1994). Quantitative analysis of the kinetics of renaturation and denaturation of barstar in the folding transition zone. *Protein Sci.* **3**, 1409–1417.
- Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994). The barriers in protein folding. *Nature Struct. Biol.* **1**(1), 149–156.
- Stryer L. (1965). The interaction of naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol.* **13**, 482–495.
- Swaminathan, R., Periasamy, N., Udgaonkar, J. B. & Krishnamoorthy, G. (1994). Molten globule-like conformation of barstar: a study by fluorescence dynamics. *J. Phys. Chem.* **98**, 9270–9278.
- Tanford, C. (1970). Protein denaturation. Part C. Theoretical models for the mechanism of denaturation. *Advan. Protein Chem.* **21**, 1–95.
- Udgaonkar, J. B. & Baldwin, R. L. (1988). NMR evidence for an early intermediate in the folding pathway of ribonuclease A. *Nature (London)*, **335**, 694–699.
- Udgaonkar, J. B. & Baldwin, R. L. (1990). Early folding intermediate of ribonuclease A. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 8197–8201.

Appendix

The mechanisms shown in Scheme I involve up to three reaction pathways occurring in parallel, with each reaction pathway involving up to two kinetically distinct intermediates.

A reaction pathway comprising three consecutive, first-order reactions can be represented as:



The equations governing this scheme, which give the time dependence of A , B , C and D , are well known (Szabo, 1969):

$$A = A_0 e^{-k_1 t} \quad (\text{A2})$$

$$B = A_0 \left[1 + \frac{1}{(k_1 - k_2)} (k_2 e^{-k_1 t} - k_1 e^{-k_2 t}) \right] \quad (\text{A3})$$

$$C = \frac{A_0 k_1 k_2}{(k_2 - k_1)(k_3 - k_1)(k_3 - k_2)} [(k_3 - k_2)e^{-k_1 t} - (k_3 - k_1)e^{-k_2 t} - (k_1 - k_2)e^{-k_3 t}] \quad (\text{A4})$$

$$D = A_0 \left[1 - e^{-k_1 t} \frac{k_1}{(k_2 - k_1)} (e^{-k_1 t} - e^{-k_2 t}) - \frac{k_1 k_2}{(k_2 - k_1)(k_3 - k_1)(k_3 - k_2)} \{ (k_3 - k_2)e^{-k_1 t} - (k_3 - k_1)e^{-k_2 t} - (k_1 - k_2)e^{-k_3 t} \} \right] \quad (\text{A5})$$

A_0 is concentration of A at time $t = 0$, and k_1 , k_2 and k_3 are the rate constants describing the reaction pathway comprising three consecutive reactions. The corresponding expressions for A , B and C when there are only two consecutive reactions are easily obtained from equations (A2), (A3) and (A4) by setting k_3 equal to zero.

In the series of consecutive reactions describing a reaction pathway, A may represent different forms, A_1 , A_2 , A_3 , etc that are in rapid pre-equilibrium with each other. The expression for the time dependence of A ($= A_1 + A_2 + A_3 + \dots$) with time (equation (A2)) is still the same.

When there are two or more parallel reaction pathways each yielding the same final product, the contribution of any one reaction to the final product is given by the fraction of molecules undergoing that reaction. The expression for the time dependence of the final product, $D(t)$, is therefore given by

$$D(t) = \sum f_i D_i(t) \quad (\text{A6})$$

where f_i is the fraction of molecules that form D by reaction pathway i , and $D_i(t)$ is the expression for the time dependence of D on reaction pathway i .

Reference

- Szabo, Z. G. (1969). Kinetic characterization of complex reaction systems. In *Comprehensive Chemical Kinetics* (Bamford, C. H. & Tipper, C. F. H., eds), Vol. 2, pp. 1–80, Elsevier, New York.

Edited by R. Huber

(Received 7 November 1994; accepted 27 January 1995)