pH-Jump-Induced Folding and Unfolding Studies of Barstar: Evidence for Multiple Folding and Unfolding Pathways †

ARTICLE *in* BIOCHEMISTRY · JANUARY 2002

Impact Factor: 3.02 · DOI: 10.1021/bi011701r · Source: PubMed

2 AUTHORS, INCLUDING:

Jayant B Udgaonkar
Tata Institute of Fundamental Research

147 PUBLICATIONS **4,286** CITATIONS

SEE PROFILE

pH-Jump-Induced Folding and Unfolding Studies of Barstar: Evidence for Multiple Folding and Unfolding Pathways†

Bhadresh R. Rami and Jayant B. Udgaonkar*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 560065, India

Received August 21, 2001; Revised Manuscript Received September 26, 2001

ABSTRACT: Equilibrium and kinetic characterization of the high pH-induced unfolding transition of the small protein barstar have been carried out in the pH range 7−12. A mutant form of barstar, containing a single tryptophan, Trp 53, completely buried in the core of the native protein, has been used. It is shown that the protein undergoes reversible unfolding above pH 10. The pH 12 form (the D form) appears to be as unfolded as the form unfolded by 6 M guanidine hydrochloride (GdnHCl) at pH 7 (the U form): both forms have similar fluorescence and far-UV circular dichroism (CD) signals and have similar sizes, as determined by dynamic light scattering and size-exclusion chromatography. No residual structure is detected in the D form: addition of GdnHCl does not alter its fluorescence and far-UV CD properties. The fluorescence signal of Trp 53 has been used to monitor folding and unfolding kinetics. The kinetics of folding of the D form in the pH range 7−11 are complex and are described by four exponential processes, as are the kinetics of unfolding of the native state (N state) in the pH range 10.5−12. Each kinetic phase of folding decreases in rate with increase in pH from 7 to 10.85, and each kinetic phase of unfolding decreases in rate with decrease in pH from 12 to 10.85. At pH 10.85, the folding and unfolding rates for any particular kinetic phase are identical and minimal. The two slowest phases of folding and unfolding have identical kinetics whether measured by Trp 53 fluorescence or by mean residue ellipticity at 222 nm. Direct determination of the increase in the N state with time of folding at pH 7 and of the D form with time of unfolding at pH 12, by means of double-jump assays, show that between 85 and 95% of protein molecules fold or unfold via fast pathways between the two forms. The remaining 5−15% of protein molecules appear to fold or unfold via slower pathways, on which at least two intermediates accumulate. The mechanism of folding from the high pH-denatured D form is remarkably similar to the mechanism of folding from the urea or GdnHCl-denatured U form.

Changes in pH are expected to alter the stability of a protein when the native (N) and unfolded (U) states differ in the number of protons bound, because the N ⇌ U unfolding reaction is coupled to the protonation reactions of the N and U states (*1*−*3*). It is not unusual for the N and U forms to have different numbers of protons bound, because ionizable amino acid side chains will very often have different pK_a values in the two forms. In the N state, pK_a values will be perturbed if side chains are buried, hydrogen bonded, or otherwise involved in electrostatic interactions (*4*, *5*), while in the U state, repulsion between like-charged side chains may perturb pK_a values (*5*, *6*). A determination of the pH dependence of stability or structure in a peptide or protein very often allows identification and characterization of specific electrostatic interactions that make significant contributions in determining the free-energy difference between the folded and unfolded states (*3*, *7*−*10*). Similarly, changes in pH are expected to affect the kinetics of folding and unfolding when the ground and transition states differ in the number of protons bound and determining the pH dependence of folding and unfolding kinetics should allow delineation of the contributions of specific electrostatic interactions to the energy barriers that separate folded and unfolded proteins (*11*−*14*). Nevertheless, studies characterizing the pH dependence of protein folding and unfolding kinetics are scarce.

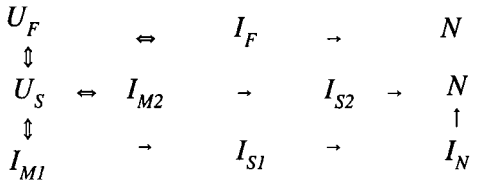
Many proteins unfold completely at extremes of pH (*15*−*17*). It is therefore possible (a) to determine folding and unfolding kinetics in the absence of denaturant; (b) to determine easily the pH dependence of these kinetics; and, most significantly, (c) to compare the kinetics determined in the absence of chemical denaturant at any fixed pH to what is expected from extrapolation to zero denaturant of the kinetics determined in the presence of chemical denaturant (*18*). Such comparison is important because low concentrations of chemical denaturant may alter the energy landscape of folding drastically by introducing or stabilizing local energy minima that act either as kinetic traps according to the new view of folding (*19*−*23*) or as folding intermediates according to the classical view of folding (*24*−*26*). Thus, it is surprising that studies of folding and unfolding reactions of proteins initiated by pH-jumps are so rare, while studies of such reactions initiated by jumps in the concentrations of chemical denaturants are commonly carried out.

† This work was supported by the Tata Institute of Fundamental Research and the Department of Biotechnology, Government of India. J.B.U. is the recipient of a Swarnajayanti Fellowship from the Government of India.

* To whom correspondence should be addressed. E-mail: jayant@ncbs.res.in. Fax: 91-80-3636662.

10.1021/bi011701r CCC: $20.00
Published on Web 11/20/2001

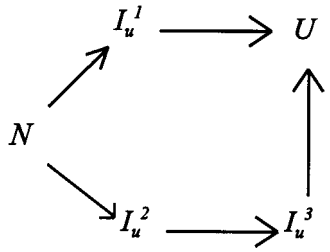
For the 89 residue protein barstar, the mechanism of folding from GdnHCl or urea-unfolded forms has been studied extensively using a large variety of physical methods with temporal resolution ranging from tens of microseconds to hundreds of seconds (*27*−*32*). In 6 M GdnHCl or 8 M urea, barstar exists in two unfolded forms; 30% of the molecules exist as the fast refolding U_F, and 70% as the slow refolding U_S. In U_F, the *Tyr 47*−*Pro 48* peptide bond is thought to be cis, like it is in the fully folded protein. U_S differs from U_F in accommodating this bond in the nonnative trans conformation. Multiple intermediates and competing pathways characterize the folding of U_F and U_S:



Mechanism 1

U_S appears to fold by two competing pathways in strongly stabilizing (low denaturant concentration) conditions, and on each pathway, folding commences by initial formation of a compact premolten globule form (I_{M1} or I_{M2}). Trans to cis isomerization of the *Tyr 47*−*Pro 48* peptide bond occurs during transformation of the late intermediates I_N and I_{S2} to N. U_F folds by only one pathway in which the intermediate I_F appears to form initially (*27*).

Unfolding of barstar by high concentrations of urea or GdnHCl also occurs via two competing pathways and multiple unfolding intermediates.



Mechanism 2

The kinetics of unfolding were measured by monitoring changes in far-UV CD and fluorescence (*33*), as well as a chemical labeling method which measured the kinetics of solvent exposure of a single Cys thiol that is fully buried in the native protein (*34*). The presence of competing unfolding pathways was suggested by three observations: (1) the three different probes yielded different rate constants of unfolding; (2) the dependence of the rate constant on denaturant concentration was different for each of the three probes; and (3) the existence of 5−8 ms burst phase changes in one or two of the probes.

Barstar is known to unfold at extremes of pH. At low pH, it unfolds partially to a molten globule-like A form (*35*). The A form possesses solvent-exposed hydrophobic patches capable of binding to the hydrophobic dye ANS, a property it shares with all the kinetic intermediates that accumulate during the folding of urea-unfolded or GdnHCl-unfolded barstar (mechanism 1), with the exception of I_F. At high pH, barstar undergoes reversible and complete unfolding (*36*). The pH 12 form (the D form) is not molten globule-like, but is unfolded to the same extent as protein unfolded in 6 M GdnHCl, pH 7 (the U form), as judged by far-UV CD and fluorescence.

In the present work, the kinetics of folding from the pH 12 unfolded form (D form), as well as the kinetics of unfolding of the N state at pH 7 to the D form at pH 12, have been studied. The W38FW44F mutant form of barstar, having a single trp (Trp 53) in the core of the protein, and whose stability and activity are similar to that of the *wt* protein (*37*) has been used. It is first shown that not only is the D form very similar to the U form, with respect to spectroscopic properties and size, but also that residual structure is absent in the D form as it is in the U form. The high pH-induced unfolding transition does not populate any equilibrium intermediates, as has been shown from the overlapping pH titration curves obtained using different probes. The kinetics of folding from the D form are shown to be significantly more complex than those from the U form. Both refolding and unfolding reactions induced by pH-jumps lack any burst phase signal. Changes in the fluorescence of Trp 53 during folding from the D form to the N state as well as unfolding from the N state to the D form occur in four kinetic phases. It is shown that 85% of the protein molecules undergo fast folding from the D form to the N state, while the remaining 15% of the molecules fold via alternative pathways in which several intermediates accumulate. Similarly, 90−95% of the protein molecules unfold from the N state to the D form via a fast pathway, while the remaining 5−10% of the molecules appear to unfold via the alternative pathways populated by intermediate structures.

MATERIALS AND METHODS

The W38FW44F mutant form of barstar was purified using a protocol similar to that for *wt* barstar (*35*). All reagents used to make buffers were of the highest grade purity.

For equilibrium pH titrations, a universal buffer was used (10 mM sodium citrate, 10 mM sodium phosphate, 10 mM sodium borate) and the pH was adjusted with NaOH or HCl to different final values. Special care was taken to keep the variation in pH to within ±0.01 pH units. All the solutions contained 250 μm EDTA and 1 mM DTT. All experiments were carried out at 25 °C. For refolding experiments, the protein was unfolded to equilibrium (4 h) in 5 mM sodium borate solution at pH 12, whereas for unfolding experiments the protein was dissolved in 5 mM sodium phosphate solution at pH 7.

Equilibrium Denaturation Studies. Equilibrium unfolding studies were carried out between pH 2 and 12. Protein was incubated in solutions of different pH values for 3 h, after which fluorescence or CD measurements were carried out. Fluorescence measurements were recorded on a Spex DM 3000 Fluorolog spectrofluorimeter. The sample was excited at 295 nm, using a bandwidth of 5 nm, and the emission was monitored at 320 nm, using a bandwidth of 10 nm. CD measurements at 222 nm were carried out on a JASCO, J-700 spectropolarimeter, using a bandwidth of 1 nm and a response time of 1 s. In both cases, the final protein concentration in the cuvette was 2 μM. Absorbance spectra were collected on a Cary 100 spectrometer, from 240 to 320 nm, with a wavelength step of 0.2 nm, signal integration time 1 s, and

scan rate of 15 nm/min. The final concentration of protein was 10 μM in a 1 cm path-length cuvette.

Size-exclusion chromatography (SEC) experiments were carried out using a Superdex-75 column with a fractionation range of 1000−100000 Da. Native protein was made in 20 mM sodium phosphate (pH 7). The alkaline unfolded protein was made in 30 mM sodium borate (pH 12). For comparison, protein was also unfolded in 6 M GdnHCl, 20 mM sodium phosphate (pH 7). All solutions contained 250 μM EDTA, 1 mM DTT, and 200 mM KCl. In each case, the column was preequilibrated with two column volumes of the same buffer in which the protein is dissolved, following which the protein was loaded.

Dynamic light-scattering studies on the protein were carried out using a DynaPro-99 equipment (Protein Solutions Ltd.). The detection capability of the instrument (1 nm to 1 μm) required that the minimum protein concentration be 50 μM. Samples were incubated in buffers at pH 7, pH 12 and 6 M GdnHCl, pH 7 for 3 h. The samples were degassed, spun down at 14 000 rpm for 15 min, and filtered through a 0.02 μm filter. The data acquisition time was 3 s at a sensitivity of 90%. This was long enough to collect an adequate number of photons so as to obtain smoother correlation curves and thus a greater confidence in the experimental results and short enough to prevent any diffusing dust particle from entering the observation volume. All fluctuations in intensities greater than 15% were marked as excluded, and not used for data analysis. The DynaLs software (Protein Solutions Ltd.) was used to resolve the measurements into well-defined Gaussian distributions. The goodness of fit was verified by the residuals. The refractive indices of solutions were determined using an Abbe-type refractometer (Milton Roy); these were then used to determine the viscosities of the solutions using the refractive index table provided with the software. All readings were taken at 25 °C.

Kinetic Experiments. Initiation of mixing and observation of kinetic processes from the millisecond time scale onward was achieved using a Biologic SFM-4 mixing module (Biologic Inc. France). Mixing dead-times of the order of 1 ms were obtained, using a cuvette of 0.08 cm path length, with flow rates of 20 mL/s. The final protein concentration in the cuvette ranged from 1 to 20 μM. Excitation was at 295 nm, and fluorescence emission was measured using a 320 nm band-pass filter. Data were acquired in two time domains on different channels, with different sampling times for each domain. For refolding experiments, 30 μL of equilibrium-unfolded protein at pH 12 were diluted into 270 μL of refolding buffer, to different final pH values (±0.02). For unfolding experiments, 30 μL of equilibrated native protein at pH 7 were diluted into 270 μL of unfolding buffer at different pH values.

All buffers were filtered through 0.2 μm filters and degassed before adjusting the pH. Sodium phosphate (30 mM) was used for all buffers below pH 8, and 30 mM sodium borate was used for all buffers to obtain final pH values beyond 9. All buffers contained 250 μM EDTA and 1 mM DTT.

Double-Jump Assay for Formation of the N State during Folding at pH 7. Using a Biologic SFM-400 mixing module, 40 μL of equilibrium-unfolded protein at pH 12 were mixed with 80 μL of refolding buffer (30 mM phosphate, pH 4.65), such that the final pH was 7 ± 0.02. The refolding mixture was aged for different lengths of time in a delay loop of 90 μL volume (the inter-mixer volume was 116 μL). After different refolding times, the solution in the delay loop was mixed with unfolding buffer (60 mM sodium borate, pH 12.62) such that the final pH of the unfolding solution is 12 ± 0.02. The final protein concentration was 15 μM in a 0.08 cm path-length cuvette. Dead-times of ~1 ms were obtained using flow rates of −13 mL/s. To improve the quality of the stop, a hard stop was mounted on top of the cuvette. The hard stop was programmed so that its closure was synchronized with the end of the push, to avoid any turbulence in the fluid coming into the cuvette. The unfolding reaction after each time of refolding was monitored by measurement of fluorescence emission at 320 nm. Data points were collected at 200 μs intervals for a duration of 50 ms to capture the fast phase of the unfolding reaction. The amount of native protein formed at each time of refolding was determined from measurement of the amplitude of the fast phase of unfolding, which occurred at a rate of $350 \pm 20\ s^{-1}$. The fraction of native protein (N state) formed at time t was determined as the amount of protein formed at time t divided by the amount of native protein formed at $t = 100$ s.

Double-Jump Assay for Formation of the D Form during Unfolding at pH 12. Using a Biologic SFM-4 mixing module, 40 μL of equilibrated native protein solution (5 mM sodium phosphate, pH 7) was mixed with 80 μL of unfolding buffer (30 mM sodium borate at pH 12.52) such that the final pH was 12 (±0.02) The unfolding protein solution was aged in a delay loop of 90 μL (the intermixer volume was 116 μL). After different times of unfolding, the solution in the delay loop was mixed with refolding buffer (60 mM phosphate, pH 4.65) such that the final pH was 7 ± 0.02. The final protein concentration was 15 μM in a 0.08 cm path-length cuvette. Mixing dead-times were ~2 ms. For unfolding times between 9 and 26 ms, 30 μL of native protein at pH 7 were mixed with 60 μL of unfolding buffer at pH 12.5 (final pH 12), passed through a delay loop of 17 μL (the intermixer volume was 43 μL), before being mixed with refolding buffer. The unfolding time was varied by varying the flow rate through the delay loop. The refolding reaction after each time of unfolding was monitored by measurement of fluorescence emission at 320 nm. The amount of unfolded protein (D form) formed at each time of unfolding was determined from measurement of the amplitude of the fast phase of refolding, which occurred at a rate of $50 \pm 5\ s^{-1}$. The fraction of unfolded protein (D form) formed at time t was determined as the amount of D form present at time t divided by the amount of the D form present at $t = 100$ s.

Data Analysis. Raw equilibrium unfolding data of an optical property, Y, vs pH were converted to plots of fraction apparently unfolded, f_{app} versus pH by use of the following equation:

$$f_{app} = \frac{Y_N - Y}{Y_N - Y_D} \quad (1)$$

where Y_N is the signal of N at pH 7, and Y_D is the signal corresponding to the D form at pH 12.

Equilibrium pH-induced denaturation data, measured using fluorescence or far-UV CD, in the pH range 7−12 were fitted

to a two-state model, where the equilibrium constant K_{app} that characterizes the transition between N and D, is given by

$$K_{app} = \frac{Y_N - Y}{Y - Y_D} \quad (2)$$

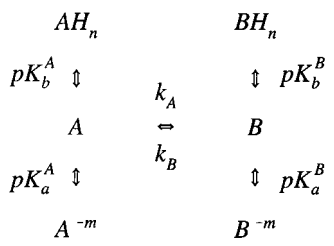
In the narrow transition zone, log K_{app} can be approximated to be linearly dependent on pH:

$$\log K_{app} = n(\mathrm{pH} - \mathrm{pH}_m) \quad (3)$$

where, pH_m is the midpoint of the structural transition in the high pH range, and n is the number of protons lost in the high pH-induced unfolding transition. Thus, eq 1 can be rewritten as

$$f_{app} = \frac{10^{n\,(\mathrm{pH}-\mathrm{pH}_m)}}{1 + 10^{n\,(\mathrm{pH}-\mathrm{pH}_m)}} \quad (4)$$

The pH dependence of folding and unfolding kinetics for an A $\rightleftharpoons$ B transition can be explained on the basis of the following simple model, which is based on models commonly used to explain the pH dependence observed in enzyme reactions:



Mechanism 3

In this model, the structural transitions between A and B are coupled to protonation and deprotonation reactions. It is assumed that there are m equivalent, noninteracting, proton-binding sites on A and B, with pK values of pK_a^A and pK_a^B in A and B, respectively; as well as n equivalent, noninteracting, proton-binding sites on A as well as B, with pK values of pK_b^A and pK_b^B in A and B, respectively. pK_a as well as pK_b values may or may not be different in A and B, depending on structural differences between A and B. For simplicity, it is assumed, as in the Hill equation, that all members of each class of sites are either all protonated or all deprotonated; m and n therefore correspond to Hill coefficients.

For mechanism 3, with the assumption that all protonation and deprotonation steps are fast with respect to the structural transitions, the observed relaxation rate, λ, is given by

$$\lambda = k'_A + k'_B = \frac{k_A}{1 + \left(\frac{H^+}{K_a^A}\right)^m + \left(\frac{K_b^A}{H^+}\right)^n} + \frac{k_B}{1 + \left(\frac{H^+}{K_a^B}\right)^m + \left(\frac{K_b^B}{H^+}\right)^n} \quad (5)$$

The pH dependence of the equilibrium constant K_{AB} is given by

$$K_{AB} = \frac{B}{A} = \frac{k'_A}{k'_B} = \frac{k_A}{k_B}\frac{1 + \left(\frac{H^+}{K_a^B}\right)^m + \left(\frac{K_b^B}{H^+}\right)^n}{1 + \left(\frac{H^+}{K_a^A}\right)^m + \left(\frac{K_b^A}{H^+}\right)^n} = \frac{k_A}{k_B}\frac{1 + 10^{m(\mathrm{pK_a^B}-\mathrm{pH})} + 10^{n(\mathrm{pH}-\mathrm{pK_b^B})}}{1 + 10^{m(\mathrm{pK_a^A}-\mathrm{pH})} + 10^{n(\mathrm{pH}-\mathrm{pK_b^B})}} \quad (6)$$

All kinetic data were analyzed using the Biologic Biokine software. All refolding as well as unfolding data fit to the sum of four exponentials. The goodness of fit was determined from a plot of residuals and the χ-square values obtained. The refolding traces were fit to eq 7:

$$Y(t) = at + b - c_1 e^{-\lambda_1 t} - c_2 e^{-\lambda_2 t} - c_3 e^{-\lambda_3 t} - c_4 e^{-\lambda_4 t} \quad (7)$$

Similarly, the unfolding traces were fit to eq 8:

$$Y(t) = at + b + c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t} + c_3 e^{-\lambda_3 t} + c_4 e^{-\lambda t} \quad (8)$$

where a is the slope of the baseline; b is the $t = \infty$ value; λ_1, λ_2, λ_3 and λ_4 are the apparent rate constants of the fast, intermediate, slow, and very slow phases, respectively; and c_1, c_2, c_3, and c_4 correspond to the amplitudes of these four observable phases. The relative amplitudes α_1, α_2, α_3, and α_4 were determined by dividing the observable amplitude of each phase by the total amplitude change ($c_1 + c_2 + c_3 + c_4$) observed.

RESULTS

High pH-Induced Unfolding of Barstar. Figure 1 compares pH-induced structural transition as monitored by three different optical probes. Far UV−CD at 222 nm has been used as a probe for secondary structure and gives an estimate of the helical content of the protein (Figure 1a). The fluorescence intensity of Trp 53 has been used to estimate the extent of solvation of the protein core and serves as a probe of gross tertiary structure (Figure 1b). UV-absorbance spectra show the emergence of a peak at 295 nm, at pH values greater than 10, which monitors the formation of the tyrosinate anion (Figure 1c).

All optical probes show that the protein stays native from pH 6 to 10, with no change in the signal of any probe (Figure 1). Below pH 6, the protein unfolds partially to a molten globule form that has been described earlier for the *wt* protein (*35*). All three optical probes show that the protein undergoes an unfolding transition from pH 10 to 12, whose midpoint is at pH 10.85. Figure 1d demonstrates clearly that the fluorescence and far-UV CD-monitored structural transitions at high pH are identical and overlapping. Analysis of the pH dependence of the fluorescence and far-UV CD data in Figure 1d, by the use of eq 4, indicates that two protons are lost from the protein in the unfolding transition, whose midpoint occurs at pH 10.85. The absorbance monitored transition is coincident with that measured with the other two probes till the midpoint of transition but deviates considerably beyond that, following a biphasic behavior.

In a control experiment, the fluorescence of *N*-acetyl l-tryptophan amide (NATA) was shown not to depend on pH in the range 10−12 (data not shown). Thus, the pH

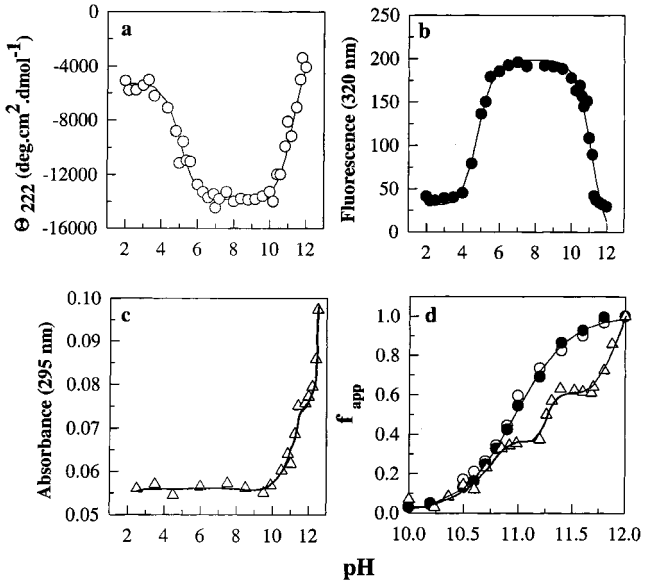


FIGURE 1: Equilibrium pH-induced unfolding of the W38FW44F mutant of Barstar at 25 °C. The structural transitions were followed by monitoring the (a) mean residual ellipticity at 222 nm (○), (b) intrinsic tryptophan fluorescence at 320 nm upon excitation at 295 nm (●), (c) absorbance at 295 nm (△), (d) The raw data in panels a, b, and c were converted to fraction apparently unfolded, f_{app} using eq 1, and subsequently plotted against pH. The solid lines through the data points in panels a and b have been obtained by fitting the data to eq 2, which yields midpoint values of pH 4.5 (for low pH denaturation) and 10.85 (for high pH unfolding). The solid line through the fluorescence and far-UV CD data in panel d is a nonlinear least-squares fit of the data to eq 4 and yield values for n and pH_m of 2 and 10.9, respectively. Line through absorbance data is drawn by inspection only.

dependence of the fluorescence change seen in Figure 1b cannot be attributed to the dependence of the spectroscopic properties of the chromophore on solvent pH but must arise from structural transitions that affect the quantum yield of Trp53 fluorescence. Above pH 12, the indole group of Trp53 begins to titrate, with a resultant decrease in fluorescence. This, together with the possibility of irreversible chemical modification of the protein above pH 12, such as the formation of a cyclic amide linkage between the side-chain primary amine group of a lysine or arginine residue and a backbone carbonyl group, has restricted the upper limit of working in the alkaline range to pH 12. Two criteria suggest, however, that the protein is completely unfolded at pH 12. The fluorescence intensity as well as the wavelength of maximum emission (355 nm) match that of the protein in 6 M GdnHCl, and the mean residue ellipticity at 222 nm of -4000 deg cm^2 dmol^{-1} matches that of the protein in 6 M GdnHCl (see below).

Comparison of the Optical Properties of the pH 12 Unfolded Form (D form) with the 6 M GdnHCl Unfolded Form (U form). Figure 2 compares and contrasts the optical properties of the pH 7 and 12 forms of barstar upon being subjected to increasing concentrations of GdnHCl. Both secondary and tertiary structure probes were used to assess the titration behavior of the protein at pH 7 and 12. The fluorescence monitored data in Figure 2a show that the native protein at pH 7 undergoes the expected cooperative unfolding transition observed earlier (*37*), while the form at pH 12 does not exhibit any cooperative unfolding transition. A similar result is seen when far-UV CD is used to monitor the

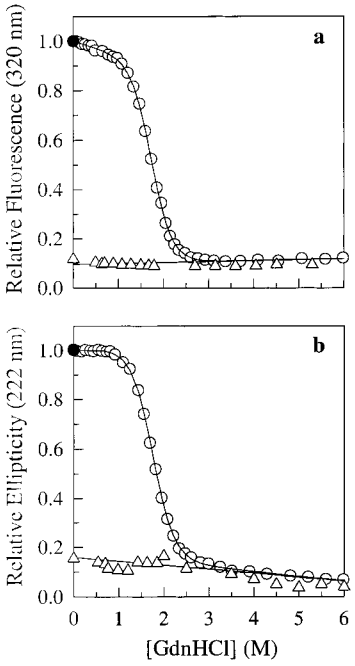


FIGURE 2: Equilibrium GdnHCl-induced denaturation curves of the W38FW44F mutant of barstar at 25 °C. (a) Fluorescence at 320 nm was used to monitor unfolding (b) Ellipticity at 222 nm was used to monitor unfolding. All values have been normalized to a value of 1 for native protein at pH 7 in the absence of any GdnHCl. The open circles (○) represent GdnHCl titration data for the protein at pH 7, while the open triangles (△) represent GdnHCl titration data for the protein at pH 12. The solid lines through the data points at pH 7 in panels a and b represent nonlinear least-squares fits to a two-state N ⇌ U unfolding model (*65*). The lines through the △ represent the extrapolated unfolded baseline, for the pH 7 data, into the pre-transition and transition region. In panels a and b, the solid cirlce (●) represents the fluorescence and ellipticity, respectively, of protein that had been returned to native conditions (pH 7) after equilibration at pH 12 in the absence of GdnHCl.

GdnHCl dependence of mean residue ellipticity. At pH 7, the midpoint of the equilibrium unfolding transition, C_m, is at 1.75 M GdnHCl, whether measured by fluorescence or by far-UV CD. The optical properties of the pH 12 form are barely affected by the presence of GdnHCl, and at high concentrations of GdnHCl, the fluorescence as well as the mean residue ellipticity of the protein at pH 7 correspond to the fluorescence and mean residue ellipticity of the protein at pH 12. Thus, the D form appears to be as unfolded as the U form, and the values of either optical property at low GdnHCl concentrations at pH 12 define the linearly-extrapolated unfolded protein baseline at pH 7.

On transferring protein, that had been unfolded to equilibrium at pH 12 in the absence of any GdnHCl, back to pH 7, the fluorescence as well as the mean residue ellipticity of the native protein is fully recovered (Figure 2). Thus, unfolding at pH 12 appears to be completely reversible. Moreover, mass spectrometric analysis (data not shown) indicates that the mass of the protein is not changed upon transfer from pH 7 to 12 and back to pH 7, indicating that no chemical modification of the protein occurs at pH 12.

Comparison of the Size of the D Form with That of the U Form. Dynamic light scattering (DLS) as well as size-exclusion chromatography (SEC) were used to compare the relative sizes of the D and U forms (Table 1). In DLS measurements, the hydrodynamic radius is determined from

Table 1: Characterization of the Sizes of Different Forms of Barstar in Solution at 25 °C

solution condition	dynamic light-scattering radius (nm)[a]	size exclusion chromatography elution volume (mL)[b]
pH 7	2.02 (±0.06)	13.4
pH 12	3.85 (±0.13)	10.2
6 M GdnHCl, pH 7	3.98 (±0.16)	9.8

[a] The number in brackets represents the standard deviation in the measurement. [b] A Superdex 75 column of void volume 8 mL was used.

measurement of the diffusion coefficient and use of the Stokes−Einstein equation, assuming a spherical shape for the protein. The native state (N state) of barstar at pH 7 has a radius of 1.95 nm, whereas the D form with a radius of 3.85 nm and the U form with a radius of 3.98 nm are considerably expanded in size. The hydrodynamic radius of 1.95 nm at pH 7 is similar to the radius of 1.9 nm calculated from the rotational correlation time that had been determined from time-resolved fluorescence anisotropy decay measurements (*38*).

The larger and similar sizes for the D form and the U form are also evident from SEC measurements of these forms and of the pH 7 native form (Table 1). Whereas the N state elutes out at a volume of 13.4 mL, the D form and the U form elute out at 10.2 and 9.8 mL, respectively, on a Superdex-75 gel filtration column that is preequilibrated with the specific buffer in which the protein has been dissolved.

In summary, the equilibrium studies show that unfolding at pH 12 is fully reversible and leads to a form (the D form) whose fluorescence and far-UV CD properties, as well as size, are identical to those of the protein unfolded in 6 M GdnHCl at pH 7 (the U form). The structural unfolding transition at high pH is the same, whether measured by far-UV CD or fluorescence, suggesting that high pH unfolding is highly cooperative.

Folding and Unfolding Kinetics Are Complex. Figure 3 illustrates the complex kinetics of the folding reactions, observed upon jumping completely across the high pH-induced equilibrium folding transition from pH 12 to 7, or to the midpoint of the transition, as well as the complex kinetics of the unfolding reactions, observed upon jumping the pH from 7 to 12 or to the midpoint of the unfolding transition. Both the refolding and the unfolding kinetics fit to 4 exponentials, with time constants spanning six decades in time from 3 ms to 3000 s.

Figure 3a shows a representative kinetic trace of refolding, from the D form to the N state, initiated by a pH-jump from 12 to 7. The entire folding process is observable: the $t = 0$ point of the kinetic trace coincides with the signal for the D form, and the $t = \infty$ point overlaps with the equilibrium signal of the N state, suggesting that the reaction has reached completion (only the first 100 ms of folding is shown). The fast phase accounting for 85% of the signal change occurs at a rate of 50 (±5 s^{-1}), whereas the remaining 15% of the signal change occurs in three phases, an intermediate phase, slow phase, and a very slow phase, with observable rates of 8 (±3) s^{-1}, 0.08 (±0.01) s^{-1}, and 0.008 (±0.001) s^{-1}, respectively. The slower phases of folding are more evident in Figure 3b, in which the kinetics of folding of the D form are shown, following a transfer from pH 12 to 11, close to the midpoint of the folding transition.

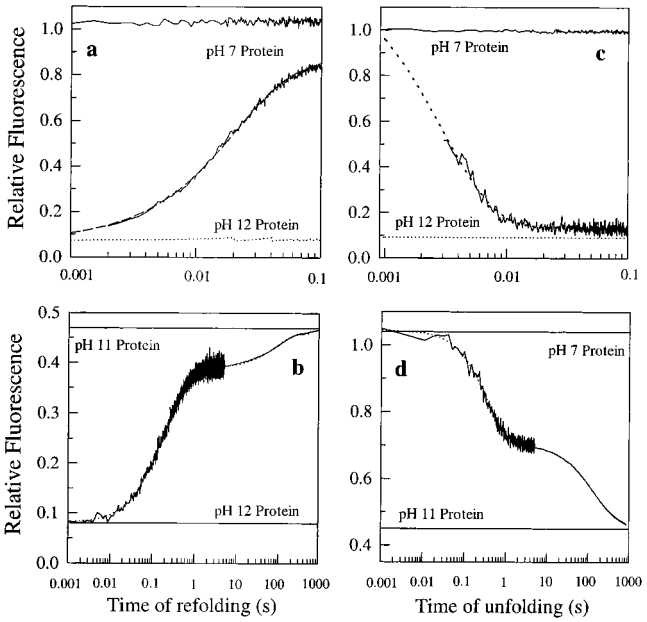


FIGURE 3: Kinetics of folding and unfolding monitored by intrinsic tryptophan fluorescence. All values have been normalized to a value of 1 for the fluorescence of the native protein at pH 7. (a) Protein equilibrated at pH 12 was refolded at pH 7. The data have been fitted to eq 7. The fitted line extrapolates down to the unfolded baseline indicating the lack of any burst phase signal. The signal change corresponding only to the fast phase has been shown. (b) Protein equilibrated at pH 12 was refolded at pH 11. The increase in intrinsic tryptophan fluorescence occurs in four kinetic phases, starting from the value of the unfolded protein baseline and reaching the value corresponding to equilibrium unfolded protein at pH 11, and has been fit to eq 7. (c) Protein equilibrated at pH 7 was unfolded at pH 12. The dotted line through the data points was obtained by fitting the data to eq 8. The solid line indicates the native baseline, and the dashed line represents the unfolded baseline. (d) Protein equilibrated at pH 7 was unfolded at pH 11. The decrease in fluorescence relaxation starting from the signal for native protein and ending at the equilibrium unfolded baseline at pH 11, occurs in four kinetic phases, and has been fit to eq 8.

Figure 3c represents the kinetics of unfolding of the N state to the D form, following a pH-jump from 7 to 12. As in the case of folding experiments from the D form, the entire change in fluorescence is observable. Approximately 90−95% of the unfolding reaction is over within 10 ms. The fast phase occurring at a rate of 350 (±20) s^{-1} is followed by an intermediate phase of rate 35 (±5) s^{-1}, a slow phase of 0.3 (±0.05)s^{-1} and a very slow phase of 0.025 (±0.004) s^{-1}, which together account for the remaining 5−10% of the signal change. The multiphasic nature of the unfolding reaction is more evident in Figure 3d, in which the kinetics of unfolding of the N state are shown following a transfer from pH 7 to 11.

Dependences of the Refolding and the Unfolding Kinetics on pH. Figure 4 shows the dependence on pH of each of the four refolding rates observed when the D form is transferred from pH 12 to different pH values between 6 and 11 and the dependence on pH of each of the four unfolding rates observed when the N state is transferred from pH 7 to different pH values between 10.2 and 12. For each kinetic phase, the folding and unfolding rates display a 'V'-shaped chevron. The midpoint of each of the four chevrons, where the folding and unfolding rates are equal as well as the slowest, occurs at pH 10.85, which is the same as the midpoint of the equilibrium transition. From the midpoint

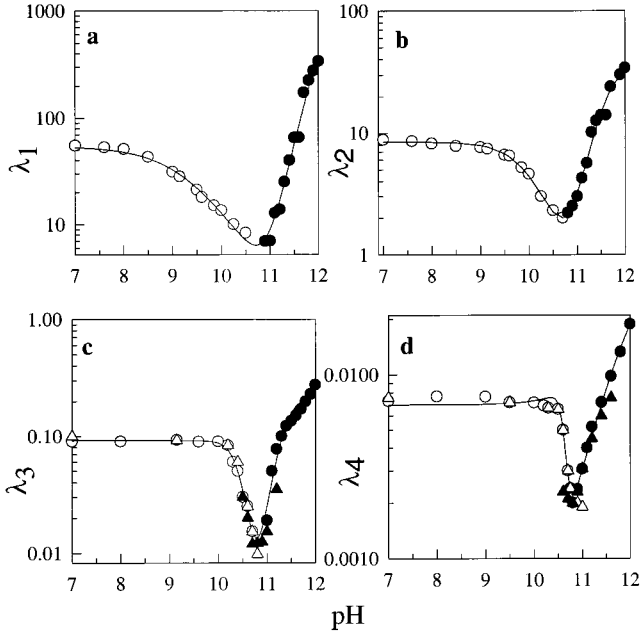


FIGURE 4: pH dependence of the apparent rate constants for the four observable kinetic phases. Protein that has been unfolded to equilibrium at pH 12 was refolded at different pH values, and the apparent refolding rate was determined for each pH value (○, △). Native protein that had been equilibrated at pH 7 was unfolded at different pH values, and the apparent unfolding rate was determined for each pH value (●, ▲). Open and closed circles (○, ●) denote data from stopped flow mixing measurements, carried out using protein concentration of 10 μM; closed and open triangles (▲, △) denote data from manual-mixing experiments, carried out at 10-fold lower protein concentrations. Each data point represents the average of three independent experiments, with a standard deviation of ±5%. The solid lines through the data represent nonlinear least-squares fits to eq 5, and values obtained for various parameters are listed in Table 2.

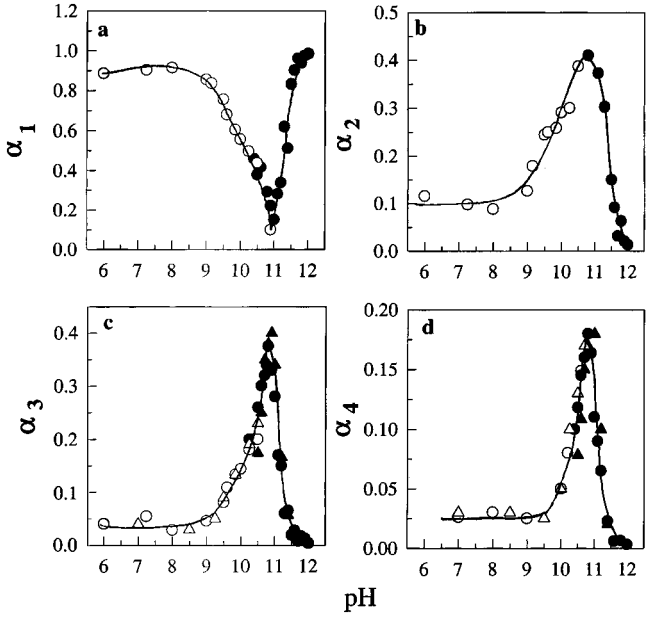


FIGURE 5: pH dependence of the relative amplitudes of the four observable kinetic phases. Protein that had been unfolded to equilibrium was refolded at different pH values, and the relative amplitude was determined for each pH value (○). Native protein that had been equilibrated at pH 7 was unfolded at different pH values, and the relative amplitude was determined for each pH value (●). △, ▲ Represent manual mixing fluorescence and CD data. The relative amplitude of each phase at any pH was determined by dividing the amplitude by the total amplitude of fluorescence change. Each data point represents the average of three independent experiments, with a standard deviation of ±5%. The solid lines have been drawn by inspection only.

of the chevron, the fast refolding rate increases as the pH is decreased to pH 8, below which it is independent of pH. In contrast, the intermediate refolding rate is independent of pH below pH 9, while the slow and very slow refolding rates are independent of pH between pH 6 and 10. Thus, the two slowest rate constants are independent of pH in the pH range where the protein retains native structure (Figure 1). In contrast to this, the four unfolding rates show a steep dependence on pH from pH 11 to 12. All four chevrons are seen to fit well to eq 5.

Figure 5 shows the dependence on pH of the relative amplitudes of each of the four folding, as well as of each of the four unfolding phases. For the fast phase, the minimum relative amplitude (10%) occurs at pH 11, which is also the pH at which the maximum relative amplitudes are seen for the intermediate (40%), slow (30%), and the very slow (20%) phases. The relative amplitude of the fast phase of folding is maximum and invariant for jumps from pH 12 to pH values between 6 and 8, and the relative amplitude of the intermediate phase is the smallest in the same pH range. The relative amplitudes of the slow and very slow phases are minimum and invariant between pH 6 and 9. The decrease in relative amplitude of the fast refolding phase between pH 8 and 11 occurs at the cost of increases in the relative amplitudes of the other three phases, and the increase in the relative amplitude of the fast unfolding phase occurs at the expense of decreases in the relative amplitudes of the three slower phases.

The two slowest kinetic phases of folding and unfolding were also measured by manual-mixing far-UV CD measurements at 222 nm as well as manual-mixing fluorescence measurements. The rates and amplitudes measured were identical to those observed in the stopped-flow measurements (Figures 4 and 5). Thus, the kinetics of the two slowest phases are independent of the mode of measurement. The kinetics of all four phases were also found to be independent of protein concentration in stopped-flow experiments carried out in the concentration range 1−20 μM (data not shown), ruling out any aggregation artifacts.

Formation of N Measured by a Double-Jump (D → N → D) Unfolding Assay. The extent of formation of N at any time during folding from the D form, following a jump in pH from 12 to 7, was assayed for in a double-jump assay in which the rate as well as amplitude of the fast phase of unfolding of any N present at that time were measured. At any time of refolding, the rate of 350 ± 30 s^{-1} (Figure 4a) for the fast phase of unfolding identifies the unfolding form as N or N-like, and the amplitude of the fast unfolding phase signified the amount of N or N-like forms present. In Figure 6, the fraction of N or N-like forms is seen to form from D in two kinetic phases. A total of 85% of the N or N-like molecules is formed with an apparent rate constant of 50 (±5) s^{-1}, while the remaining 15% is formed at a rate of 0.008 s^{-1}. The fast phase corresponds in rate as well as amplitude to the fast phase of fluorescence change, and the slow phase corresponds in rate to the very slow phase of fluorescence change in direct refolding experiments.

Formation of D Measured by a Double-Jump (N → D → N) Refolding Assay. The extent of formation of D at any

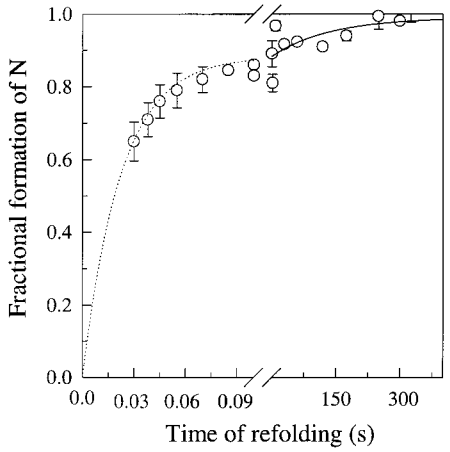


FIGURE 6: Kinetics of formation of the N state from the D form, following a jump in pH from 12 to 7. The fraction of N, F_N, formed is plotted against time of refolding. The amount of N formed at each time of refolding, from 30 ms to 300 s, was determined from a double-jump unfolding assay, as described in the Materials and Methods. The fraction of N formed at any time of folding was determined as the amount of N formed at that time divided by the amount of N formed after folding was complete. The dotted line describes the formation of N according to a two-exponential process and is described by $F_N = 1 - 0.85 \exp(-50t) - 0.15 \exp(-0.008t)$. Each data point represents the average of two independent determinations from different experiments, and the error bars represent the spreads in the values.

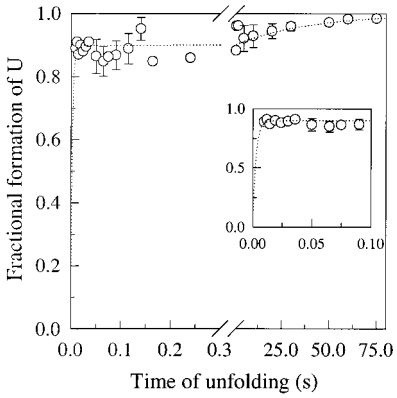


FIGURE 7: Kinetics of formation of the D form from the N state, following a jump in pH from 7 to 12. The fraction of D, F_D, formed is plotted against time of refolding. The amount of D formed at each time of unfolding, from 10 ms to 100 s, was determined from a double-jump refolding assay, as described in the Materials and Methods. The fraction of D formed at any time of unfolding was determined as the amount of D formed at that time divided by the amount of D formed after folding was complete. The dotted line describes the formation of D according to a two-exponential process, and is described by $F_D = 1 - 0.9 \exp(-350t) - 0.10 \exp(-0.025t)$. The inset shows data obtained for the first 100 ms of unfolding. Each data point represents the average of two independent determinations from different experiments, and the error bars represent the spreads in the values.

time during unfolding from the N state, following a jump in pH from 7 to 12, was assayed for in a double-jump assay in which the rate as well as amplitude of the fast phase of refolding of any D present at that time were measured. At any time of unfolding, the rate of 50 ($\pm$10) s^{-1} (Figure 4a) for the fast phase of refolding identifies the refolding form as D, and the amplitude of the fast refolding phase signifies the amount of D present at that time. In Figure 7 is shown the fraction of molecules present as D at different times of unfolding, following a jump in pH from 7 to 12. D appears to form from N in two kinetic phases. A total of 90% of the molecules forms D within 10 ms after initiation of unfolding, suggesting that the rate of formation of D is greater than 330 s^{-1}. The remaining 10% of the molecules forms D at a rate of 0.025 s^{-1}. The fast phase of formation of D corresponds in rate as well as amplitude to the fast phase of fluorescence change, and the slow phase corresponds in rate to the very slow phase of fluorescence change seen in direct unfolding experiments. When protein that had been unfolded for 10 ms is refolded, the refolding kinetics are identical to equilibrium unfolded protein with respect to the rates as well as the relative amplitudes of the four kinetic phases (data not shown).

DISCUSSION

High pH Unfolding Transition. As described earlier (*36*), barstar retains its native structure as the pH is increased from 6 to 10. The fluorescence, far-UV CD, or absorbance properties do not change in this range of pH. GdnHCl-induced unfolding studies in this range of pH show, however, that the stability of the protein decreases with increasing pH. This decrease in stability could be attributed to ionization of the thiols of the two cysteine residues present, because it did not occur in a mutant protein which had the two cysteines replaced by alanines (*36*).

The fluorescence of Trp 53 decreases when the pH is increased from 10 to 12. This decrease is not because of ionization of the indole side chain, because that is seen to occur only above pH 12 in the case of NATA (see Results). The side chains of Tyr 29, 30, and 47 are expected to ionize in this pH range, as indeed they appear to do (Figure 1c), and although a tyrosinate anion can quench Trp phosphorescence, which it can do so only when certain very specific symmetry requirements are satisfied between the orientation of the dipoles of the indole and the tyrosinate groups (*39*), its effect on Trp fluorescence is not known. That the decrease in fluorescence of Trp 53 is indeed caused by a structural unfolding transition is confirmed by the observation that it is overlapped by a decrease in mean residue ellipticity at 222 nm (Figure 1d), which suggests simultaneous breakdown of tertiary interactions and loss in secondary structure. The midpoint of the high pH-induced unfolding transition is pH 10.85, as obtained from both the optical probes, hinting at the titration of tyrosine or lysine residues.

In unfolded proteins, the side chains of tyrosine (pK_a = 10.5), lysine (pK_a = 10.8), and arginine (pK_a = 12.5) are expected to titrate in the pH range 10−12, and the protein studied here has three, six, and three of these residues, respectively. Unfolding at high pH could occur in two ways. All arginine and lysine residues are on the protein surface. Ionization of surface-charged residues at high pH would increase the net negative charge on the protein, and mutual repulsion between like charges would destabilize the protein. Alternatively, if any one of the Tyr or Lys residues has a higher pK_a in the native state than in the unfolded form, then the residue in the unfolded form would ionize at a lower pH than it would in the native state. Thermodynamic coupling would then lead to conversion of molecules to the unfolded form as the pH is increased.

At present, the mode of action of alkaline pH (above pH 10) in unfolding barstar is unclear. It is known that mutual

repulsion between negatively charged glutamate and aspartate residues clustered in the barnase-binding site of barstar has a destabilizing effect (*40*). It is, however, unclear whether the increase in negative charge by 8 charge units that is caused by further ionization between pH 10 and 12 has a similar effect, because the lysine and arginine residues are all fully solvent exposed and do not appear to be clustered on the protein surface.

The six lysines in barstar have high solvent accessibilities, project outward, and do not appear to be involved in any charge−charge interactions because they are more than 6.5 Å away from each other or from any negative charge. It is therefore unlikely that any one of them has a higher pK_a in the N state than in the unfolded state. The three tyrosines are the other residues expected to ionize in the pH range 10−12. Of the three, only Tyr 30 appears to be involved in a specific interaction, hydrogen bonding with His 17 (*41*) in the N state. Tyr 30 might therefore be expected to have a higher pK_a value in the N state than in the unfolded state. The increase in absorbance at 295 nm, which accompanies formation of tyrosinate anions, occurs in at least two steps (Figure 1, panels c and d) with an increase in pH from 10 to 12. It is likely that the second step at higher pH reflects the ionization of Tyr 30. Mutagenesis experiments currently in progress should make it possible to determine whether the unfolding of barstar at high pH might be occurring because of the specific effect of the altered pK_a of Tyr 30.

Denatured State at pH 12. According to several criteria, the D form of barstar at pH 12 appears to be as unfolded as the U form in 6 M GdnHCl, pH 7 (Figures 1 and 2, Table 1). (1) Both have identical fluorescence properties, and show emission maxima at 355 nm. (2) Both show the same value for the mean residue ellipticity at 222 nm, which is also similar to that expected for a random coil (*35*). (3) No additional unfolding transition is seen when GdnHCl is added to the D form, suggesting that no residual structure is present in the D form. The fluorescence at pH 12 is that expected from extrapolation of the unfolded protein baseline determined in high concentrations of GdnHCl at pH 7. (4) Both have similar sizes, whether measured by DLS or by SEC, with radii nearly twice the radius of the N state at pH 7. Several other proteins, including creatine kinase (*17*), staphylococcal nuclease (*16*), and β-lactamase (*15*) have also been reported to be fully unfolded at high pH.

The apparent similarity between the D and U forms is surprising because the mechanism of unfolding by chemical denaturants is expected to be very different from the mechanism by which alkaline pH might unfold a protein. The two commonly used denaturants, urea and GdnHCl, interact with the amide backbone of the polypeptide chain and also with the hydrophobic groups of the side chains of various amino acids (*1*, *42*) and thereby solubilize the polypeptide chain by weakening hydrophobic interactions. As discussed above, proteins unfold at high pH because of the effect of ionization of specific side chains.

The criteria by which the D and U forms appear similar is based on global properties of the polypeptide chain. Small differences between the two forms may be missed out because of the inadequate sensitivity of the gross structural probes used. The macroscopically observable quantity (fluorescence, far-UV CD, size) measured using these probes is in fact an average over many different configurations of individual molecules (*43*−*45*). It is one of the goals of the present work to compare and contrast the properties of the D and the U forms, as well as their refolding kinetics.

Thus, the existence of residual structure in either form cannot be ruled out. The sequence of the polypeptide chain is a significant determinant of local interactions among neighboring residues in the chain and can thus contribute to structure in denatured states (*1*, *9*). Even "good" solvents such as 6 M GdnHCl have been shown to be unable to break these local, fluctuating centers of microscopic structures under strongly destabilizing conditions (*46*−*48*), in case of other proteins. An aromatic residue that flanks a proline residue can interact with it through stacking interactions between the two cyclic structures, as observed by ring current shifts in NMR (*49*, *50*), and these very local interactions are known to occur in unstructured peptides and also in the presence of GdnHCl. Substantial hydrophobic clustering of residues can occur even in the presence of very high concentration of denaturant (*51*). As opposed to chemical denaturation, the fraction of buried surface exposed upon denaturation is increased by electrostatic repulsion when the polypeptide chain is highly charged (*9*, *52*), as at low or high pH. Studies on the acid and thermally denatured forms of barstar (*5*) have indicated that these forms are not extended coils but compact forms with intramolecular charge repulsion between side chains.

Thus, it is possible that the D and U forms may differ in the nature and extent of persistent native or nonnative interactions that cannot be detected by the gross structural probes used. The presence of residual native (*53*, *54*) or nonnative (*55*, *56*) interactions in the unfolded forms of other proteins is known to affect folding kinetics. Hence, it was of interest to compare the refolding kinetics of the D and U forms.

Complex Kinetics of the D → N Folding Transition and the N → D Unfolding Transition. Changes in the fluorescence of Trp 53 have been used to monitor the folding and unfolding transitions. These changes occur in four exponential phases for folding as well as unfolding. The entire folding and unfolding reactions can be observed. The kinetics have been shown to be independent of protein concentration, suggesting that the complexity does not arise from transient protein aggregation during the folding and unfolding processes (*57*, *58*). The fluorescence of Trp 53, which is present in the core of the N state, changes when the core gets dehydrated during folding or hydrated during unfolding; hence, it is a gross measure of tertiary structure. The slow and very slow phases of folding and unfolding have, however, identical kinetics when measured by far-UV CD, suggesting that secondary and tertiary structure change concurrently in these kinetic phases.

Origin of the Multiple Phases of Folding. The origin of multiple kinetic phases during refolding is explained classically by the existence of multiple U forms, which have similar spectroscopic properties, but differ in their refolding kinetics. The unfolded state is an ensemble of many conformations, interconversion among which is silent spectroscopically. The best characterized reason for heterogeneity in the U form arises from the slow isomerization of the Xaa-Pro peptide bond (*59*) resulting in an equilibrium mixture of cis and trans conformers. Thus, barstar unfolded in 6 M GdnHCl or 8 M urea at pH 7 (the U form) comprises of

30% U_F molecules with a nativelike cis isomer of the *Tyr 47−Pro 48* peptide bond and 70% U_S molecules with the nonnative trans isomer. The relatively high fraction of U_F molecules in the U form of barstar compared to other proteins is probably due to the presence of a Tyr preceding the Pro. Studies on unstructured model peptides (*49*, *50*) have established that a favorable stacking interaction can occur between the planar ring of an aromatic residue preceding proline and the cyclic proline side chain. Thus, many of the cis Xaa−Pro peptide bonds found in proteins have Tyr preceding the Pro (*60*, *61*).

pH Dependence of the Ratio of Fast to Slow-Refolding Unfolded Forms. In the case of the U form at pH 7, the ratio of 30:70 for U_F to U_S was determined from double jump (U → N → U) experiments which assayed for the formation of N at different times of folding. In those experiments, 30% of the molecules were found to fold fast to N, and 70% of the molecules were found to fold slow to N. Similar double-jump (D → N → D) experiments have been carried out here to determine the ratio of fast to slow folding forms in the D unfolded ensemble. In Figure 6, it is seen that 85% of the D molecules fold fast to N (or to a N-like form that can unfold at the same rate as N), at a rate of 50 (±5) s^{-1}, and 15% fold slowly to N at a rate of 0.008 s^{-1}. This suggests a ratio of 85:15 for fast-folding D molecules (D_F) to slow-folding D molecules (D_S) at pH 12.

Such a large shift in the ratio of fast-folding to slow-folding unfolded forms from 30/70 at pH 7 or 8 (*27*, *28*) to 85/15 at pH 12 is only possible if the $D_S \rightleftharpoons D_F$ equilibrium is coupled to the protonation/deprotonation of at least one group whose pK_a is different and varies between 9 and 11 in D_S and D_F. Then, the pH dependence of the $D_F \rightleftarrows D_S$ equilibrium can be explained on the basis of a simplified version of mechanism 3 (see Methods), with $n = 0$ and $m = 1$:

$$\begin{array}{ccc} & K_{SF} & \\ D_S & \Leftrightarrow & D_F \\ K_a^S \updownarrow & & \updownarrow K_a^F \\ & K_{SF}^- & \\ D_S^- & \Leftrightarrow & D_F^- \end{array}$$

Mechanism 4

If, for example, the pK_a of a titrating group has a value of 10.5 (=pK_a^S) in the D_S form (see below) and a value of 9.4 (=pK_a^F) in the D_F form and if K_{SF} (=D_F/D_S) is equal to 30/70 at pH 7, then according to the thermodynamic cycle in mechanism 4, K_{SF}^- (=D_F^-/D_S^- = $K_{SF} \cdot K_a^F/K_a^S$) will have a value of 85/15 at pH 12. A similar ratio for D_F:D_S is obtained at pH 12 by use of eq 6. Thus, while it is clearly possible for the fraction of the cis isomer in the unfolded form to be 85/15 at pH 12 and 30/70 at pH 7 or 8, this possibility rests on at least one amino acid side chain having different pK_a values in the cis and trans isomers of the unfolded form and can be considered seriously only if the amino acid side chain can be identified.

Two types of side chains can titrate in the relevant pH range. The tyrosine side chain in an unfolded protein has a pK_a of 10.5. Measurements of the cis to trans prolyl isomerization rate of *Tyr−Pro* containing peptides demonstrate that it is 3-fold retarded at a pH above the pK_a of the phenolic hydroxyl and that the cis isomer is favored more strongly at this pH (*64*), presumably because the negative charge of the tyrosinate can interact favorably with the transition state of rotation or with other nearby residues, than can an uncharged tyrosine side chain (*60*, *64*). If, with the Tyr 47−Pro 48 bond in a cis conformation, the side chain of Tyr 47 in unfolded barstar shows more favorable local interactions when deprotonated and charged than when uncharged, then its pK_a value will be lower in D_F than in D_S, because in the latter the stabilizing interaction will be absent. The involvement of Tyr 47 or any other Tyr in determining the ratio of D_F to D_S, would, however, result in it having a pK_a of 9.4 in D_F, but absorbance measurements (Figure 1d) show that the tyrosines titrate only above pH 10, thereby ruling out any role of Tyr 47 in determining the pH dependence of the D_F to D_S ratio.

The second candidate side chain is that of Cys40. It is possible that the thiol side chain of Cys40 has an abnormally high pK_a value of 10.5 in D_S, but a more normal pK_a value of 9.4 in D_F, if the −SH interacts with residual structure adjacent to Pro48 in D_S but not in D_F. There is, however, no evidence in support for this suggestion.

Since there is no evidence that the abnormal titration of any amino acid side chain might cause the ratio of fast to slow refolding D forms at pH 12 to be any different from the ratio of fast to slow refolding U forms at pH 7, it is assumed that the ratio of D_F to D_S is also 30/70.

Mechanism of Folding and Unfolding. Kinetic simulations show that mechanism 5, which is similar to mechanism 1 that accounts for the refolding of GdnHCl or urea-unfolded protein (*27*, *28*), and in which the ratio of D_S to D_F at equilibrium at pH 12 is 70:30, is the simplest mechanism that can account for the data:

$$\begin{array}{ccccccc} D_F & \Leftrightarrow & I_F & \leftrightarrow & N \\ \updownarrow & & & & \\ D_S & \Leftrightarrow & I_{M1} & \leftrightarrow & I_L \\ \updownarrow & & & & \updownarrow \\ I_{M2} & \leftrightarrow & I_1 & \leftrightarrow & I_2 & \leftrightarrow & N \end{array}$$

Mechanism 5

In mechanism 5, as in mechanism 1, I_F, I_{M1}, and I_{M2} are very early intermediates, whose stability and formation are strongly pH dependent at high pH. λ_1 has principal contributions from the reactions $D_F \rightleftharpoons N$ and $D_S \rightleftharpoons I_L$ because both have similar rates. This is similar to what is seen for folding from urea or GdnHCl unfolded protein (*27*, *28*). λ_2 has principal contributions from the $I_{M2} \rightleftharpoons I_1$ reaction, and λ_3 from the $I_1 \rightleftharpoons I_2$ reaction. λ_4 is expected to represent principally the $I_2 \rightleftharpoons N$, $I_L \rightleftharpoons N$, and $D_S \rightleftharpoons D_F$ reactions, all of which are expected to occur at the same rate, corresponding to the rate of cis to trans proline isomerization.

Mechanism 5 will explain all the observed data if it is assumed that I_L is sufficiently nativelike that it unfolds at the same rate as N at pH 12, so that the fast phase of fluorescence change would correspond to the $D_F \rightleftharpoons N$ reaction as well as the $D_S \rightleftharpoons I_L$ reaction. According to mechanism 5, of the 85% of the molecules that form N or I_L at the same rate as that of the fast phase of fluorescence change at pH 7 (Figure 6), 30% fold in the $D_F \rightarrow I_F \rightarrow N$

reaction, and 55% in the $D_S \rightarrow I_{M1} \rightarrow I_L$ reaction. The remaining 15% fold via the $D_S \rightarrow I_{M2} \rightarrow I_1 \rightarrow I_2 \rightarrow N$ route and are responsible for the observation that 15% of the molecules form N at a rate corresponding to λ_4 at pH 7.

According to mechanism 5, the increase in α_2, α_3, and α_4 at the expense of a decrease in α_1, with increase in the pH of folding from 7 to 10.85, is expected if the stabilities of I_{M1} and I_F decrease with increase in pH. The dependence of stability of any intermediate on pH is predicted by eq 6 to depend on the number of protons released when it unfolds (*5*, *6*). If I_{M1} becomes very unstable at pH 10.85, then the $D_S \rightarrow I_{M1} \rightarrow I_L \rightarrow N$ pathway will effectively not operate at that pH. All 70 D_S molecules will instead fold along the $D_S \rightarrow I_{M2} \rightarrow I_1 \rightarrow I_2 \rightarrow N$ route. Similarly, if I_F becomes less stable at pH 10.85, fewer molecules will use the $D_F \rightarrow I_F \rightarrow N$ route. According to the data in Figure 5, it appears that at pH 10.85 the stability of I_F has decreased such that only 10 of 30 molecules originally present as D_F follow the $D_F \rightarrow I_F \rightarrow N$ route (Figure 5a), while the remaining 20 molecules follow the $D_F \rightarrow D_S \rightarrow I_{M2} \rightarrow I_1 \rightarrow I_2 \rightarrow N$ route. Thus, the incorporation of two competing pathways for the folding of D_S in mechanism 5, as well as the differential stabilities of the early intermediates I_{M1} and I_{M2} on these two pathways, allow mechanism 5 to account for the pH dependence of the amplitudes of the four kinetic phases.

The major assumption that needs to be made to explain the data according to mechanism 5 is that I_L must unfold at the same rate as N, at pH 12. I_L probably corresponds to the late, nativelike intermediate, I_N, on the folding of urea unfolded protein (mechanism 1).The rate of unfolding of N at pH 12 is very fast, more than 200 times faster than the unfolding of N in 8 M urea at pH 7, which suggests that the stability of N at pH 12 is greatly reduced compared to its stability in 8 M urea at pH 7. It therefore appears that at pH 12 the stability of N approaches that of I_L, and both therefore unfold at the same fast rate.

The assumption that I_L unfolds at the same rate as N has two important implications: (1) I_L and N have the same stability at pH 12, even though N is more stable than I_L (I_N) at pH 7, and (2) since I_L is an on-pathway intermediate which, unlike N, possesses the nonnative trans conformation of the Tyr 47−Pro 48 bond, the proline isomerization reaction accompanying the N to I_L transition must occur in less than 10 ms at pH 12. Proline isomerization reactions are known to be accelerated very significantly when the polypeptide chain is constrained by structure (*66*, *67*). It appears that, at pH 12, both I_L and N are destabilized similarly with respect to D and sufficiently so that the structural transformation coupled to a proline isomerization reaction, which separates them, has an activation energy similar to the activation energy of unfolding. This provides an explanation for the observation that protein that had been unfolded at pH 12 for only 10 ms refolds with essentially the same kinetics as protein that had been unfolded to equilibrium at pH 12.

To confirm the validity of mechanism 5, kinetic simulations were carried out, using the program KINSIM (*68*). Figure 8 shows that, when folding is initiated from the D forms, the kinetics of fluorescence change accompanying folding at pH 7, the kinetics of formation of N and I_L at pH 7, and the kinetics of fluorescence change accompanying folding at pH 10.85 are all well described by the simulations

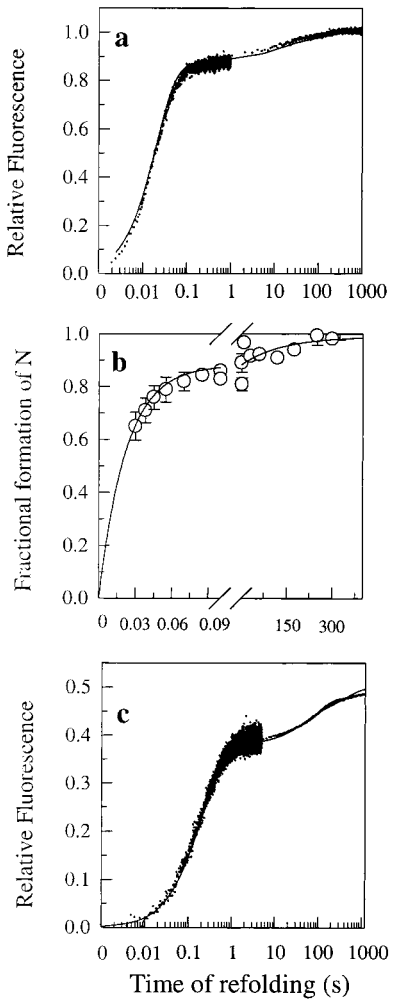


FIGURE 8: Kinetic simulations of folding kinetics. (a) Kinetics of folding at pH 7, following a jump in pH from 12 to 7 overlaid on an experimental refolding trace. (b) Kinetics of formation of N upon folding at pH 7, following a jump in pH from 12 to 7, fit to the D → N → D double jump data (open circles). (c) Kinetics of folding at pH 10.85, following a jump in pH from 12 to 10.85 overlaid on an experimental refolding trace. In each panel, the solid line indicates a simulation of the experimental data to mechanism 5. The rates used in the simulation for the $D_F \rightleftharpoons I_F$, $I_F \rightleftharpoons N$, $D_F \rightleftharpoons D_S$, $D_S \rightleftharpoons I_{M1}$, $I_{M1} \rightleftharpoons I_L$, $I_L \rightleftharpoons N$, $D_S \rightleftharpoons I_{M2}$, $I_{M2} \rightleftharpoons I_1$, $I_1 \rightleftharpoons I_2$, $I_2 \rightleftharpoons N$ reactions are 50 000; 0.0005; 50; 0.0005; 0.002; 0.001; 50;000; 0.0005; 50; 0.005; 0.008; 0.000 08; 8;000; 0.0008; 8; 0.0008; 0.08; 0.0008; 0.008; 0.000 08 s^{-1}, respectively, in panels a and b and 1, 10 000; 7, 0.0005; 0.002, 0.001; 1, 10 000; 7, 0.0005; 0.0015, 0.0015; 10, 10; 1.5, 1.5; 0.015, 0.015; 0.0015, 0.0015 s^{-1}, respectively, in panel c.

to mechanism 5. Several mechanisms in which only one pathway was available for the folding of D_S were also extensively simulated, but found to be inadequate in explaining the data in Figure 7, even if the initial ratio of D_F to D_S was changed from 30:70, thereby confirming that the inclusion of two competing pathways for the folding of D_S is critically important for mechanism 5 to account for the data.

pH Dependence of the $D_F \rightleftharpoons N$ Reaction. The agreement between the rate of the fast phase of formation of N (Figure 6) and fast phase of formation of D (Figure 7) with the rate (λ_1) of the fast phase of fluorescence change (Figure 5) at pH 7 and 12, respectively, suggests that the predominant contribution to λ_1 arises from the microscopic rate constants describing the $D_F \rightleftharpoons N$ reaction of mechanism 5. For this

Table 2: Parameters Obtained from Fitting the pH Dependence of Folding and Unfolding Kinetics to Mechanisms 3 and 5

	observed rate constant for folding and unfolding			
parameter	λ_1	λ_2	λ_3	λ_4
k_A (s^{-1})	54	9	0.09	0.007
k_B (s^{-1})	1000	37	0.22	0.04
m	1	1	3	4
pK_a^A	9.4	10	10.4	10.6
pK_a^B	14	14	14	14
n	2	2	2.5	1
pK_b^A	5	5	5	5
pK_b^B	12	11.6	11.3	12

phase, the rates of folding ($D_F \rightarrow N$) and unfolding ($N \rightarrow D_F$) are identical and slowest at pH 10.85. To explain the pH dependence of λ_1, the $D_F \rightleftharpoons N$ reaction was assumed to be coupled to n protonation reactions at a set of n equivalent, noninteracting sites, and m deprotonation reactions at a set of m, equivalent, noninteracting sites (mechanism 3). Accordingly, the pH dependence of λ_1 is described by eq 5. Figure 4 shows λ_1 fits reasonably well to this eq 5, and the values for the parameters of eq 5 are listed in Table 2. It should be noted that eq 5 is not strictly valid over the entire pH range studied because of the assumptions made in its derivation (see Materials and Methods) but, nevertheless, to a first approximation, is seen to describe adequately the data. For instance, the values of parameters obtained (Table 2) indicate that a net number of two groups on the protein get additionally protonated upon folding from pH 12 to 7, which is in agreement with the analysis of the equilibrium high pH-induced unfolding transition (Figure 1d).

Although λ_2, λ_3, and λ_4 also fit well to eq 5 (Figure 4, Table 2), no attempt has been made in this study to carry out a quantitative treatment of all the data. Even though the principal contributions to λ_2, λ_3, and λ_4 are expected to come from the $I_{M2} \rightleftharpoons I_1$, $I_1 \rightleftharpoons I_2$, and $I_2 \rightleftharpoons N$ reactions, respectively, the observed rates are less well separated, and each λ is expected to have contributions from several steps. This is reflected, for instance in the values of m and n obtained form fitting λ_3 and λ_4 to eq 5 being substantially larger than those obtained for the $D_F \rightleftharpoons N$ reaction.

It is difficult to rule out, at this stage of the work, the possibility that observed rates λ_2, λ_3, and λ_4 represent the average rates of parallel folding reactions of the protein in different protonation states or average rates of the structural transitions coupled to protonation-deprotonation reactions, which will be very slow at pH 11 to 12. It is emphasized that the mechanism proposed is the minimal mechanism which takes into account cis−trans proline isomerization. To do so, certain assumptions have had to be made about the relative stabilities of the late intermediate I_L and the N state at pH 12 and about the relative stability of the transition state separating them. While it has been argued that these assumptions are plausible, it is clear that they need to be tested in future experiments.

Folding and Unfolding Reactions Induced by Urea or GdnHCl-Jumps vs Those Induced by pH-Jumps. The principle reason for carrying out the kinetic studies reported here was to compare the refolding at pH 7 of protein unfolded by high pH, to the refolding at pH 7 of protein unfolded by urea or GdnHCl. Such studies are important, because chemical denaturant is always present during refolding in the latter case, and it is not clear how the presence of low concentrations of chemical denaturant can affect the energy landscape of folding. It is possible for example, that small regions of residual structure can bias the refolding along a particular route, and the presence of nonnative interactions in the denatured state can slow the folding reaction, and the presence of chemical denaturant may have significant effects on both. In this study, folding induced by a pH-jump from pH 12 to 7 appears to be more complex than folding induced by a dilution of urea or GdnHCl concentration. The former is characterized by four exponential processes while the latter by two (*27*). Nevertheless, when refolding is carried out in any GdnHCl concentration in the range 0.6−1.5 M, the fast refolding rates are identical whether refolding was commenced from the U form in 6 M GdnHCl at pH 7 or from the D form at pH 12 (unpublished results).

Similarly, unfolding induced by pH 7 → 12 jumps is characterized by four kinetic phases in comparison to unfolding induced by jumps to high concentrations of GdnHCl or urea, which is characterized by one or two kinetic phases. During urea or GdnHCl-induced unfolding, a 5−8 ms burst phase change in fluorescence may be seen under some conditions. The fast phase of unfolding at pH 12 has a rate of 350 (±30) s^{-1}, and if a similar rate were present during unfolding in high urea or GdnHCl, it would account for the change in fluorescence during the burst phase of the earlier urea-induced unfolding experiments (*33*).

The mechanism of high pH-jump-induced folding and unfolding that is proposed here (mechanism 5) is very similar to the mechanism proposed earlier (*27*, *28*) for folding and unfolding induced by jumps in urea or GdnHCl concentrations (mechanism 1). The present study emphasizes the similarities in the energy landscapes of folding and unfolding of barstar, whether folding and unfolding are initiated by jumps in pH or jumps in chemical denaturants. Future studies will be targeted toward obtaining an understanding of the differences seen.

ACKNOWLEDGMENT

We thank R. Varadarajan and M. K. Mathew for discussions.

REFERENCES

1. Tanford, C. (1970) *Adv. Protein. Chem. 24*, 1−95.
2. McPhie, P. (1975) *Biochemistry 11*, 879−883.
3. Anderson, D. E., Becktel, W. J., and Dahlquist, F. W. (1990) *Biochemistry 29*, 2403−2408.
4. Sancho, J., Serrano, L., and Fersht, A. R. (1992). *Biochemistry 31*, 2253−2258.
5. Oliveberg, M., Arcus, V. L., and Fersht, A. R. (1995) *Biochemistry 34*, 9424−9433.
6. Tan, Y. J., Oliveberg, M., Davis, B., and Fersht, A. R. (1995) *J. Mol. Biol. 254*, 980−992.
7. Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., and Baldwin R. L. (1987) *Nature 326*, 563−567.
8. McNutt, L., Mullins, S. L., Rauchel, F. M., and Pace, C. N. (1990) *Biochemistry 29*, 7572−7576.
9. Pace, C. N., Laurents, D. V., and Thomson, J. A. (1990) *Biochemistry 29*, 2564−2572.
10. Pace, C. N., Laurents, D. V., and Erikson, R. E. (1992) *Biochemistry 31*, 2728−2734.

11. Nall, B. T., Osterhout, J. J., and Ramdas, L. (1988) *Biochemistry 27*, 7310−7314.
12. Oliveberg, M., and Fersht, A. R. (1996) *Biochemistry 35*, 2726−2737.
13. Waldburger, C. D., Jonsson, T., and Sauer, R. T (1996) *Proc. Natl. Acad. Sci. 93*, 2629−2634.
14. Luisi, D. L., and Raleigh, D. P. (2000) *J. Mol. Biol. 299*, 1091−1100.
15. Goto, Y., and Fink, A. L. (1989) *Biochemistry 28*, 945−952.
16. Chen, H. M, Markin, V. S., and Tsong, T. Y. (1992) *Biochemistry 31*, 1483−1491.
17. Hai-Peng, Y., Hai-Ning, Z., Li, S., and Hai-Meng, Z. (1997) *Biochem. Mol. Biol. Int. 41*, 257−267.
18. Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) *Nature 340*, 122−126.
19. Go, N. (1976) *Adv Biophys.* 65−113.
20. Bryngelson, J. D., and Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. 84*, 7524−7528.
21. Honeycutt, J. D., and Thirumalai, D. (1990) *Proc. Natl. Acad. Sci. 87*, 3526−3529.
22. Shakhnovich, E., Farztdinov, G., Gutin, A. M., and Karplus, M. (1991) *Phys. Rev. Lett. 67*, 1665−1668.
23. Dill, K. A and Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.
24. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
25. Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
26. Baldwin, R. L. (1995) *J. Biomol. NMR. 5*, 103−109.
27. Schreiber, G., and Fersht, A. R. (1993) *Biochemistry 32*, 11195−11203.
28. Shastry, M. C. R., Agashe, V. R., and Udgaonkar, J. B. (1994) *Protein Sci. 3*, 1409−1417.
29. Shastry, M. C. R., and Udgaonkar, J. B. (1995) *J. Mol. Biol. 247*, 1013−1027.
30. Agashe, V. R., and Udgaonkar, J. B. (1995) *Biochemistry 34*, 3286−3299.
31. Nolting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997a) *Proc. Natl. Acad. Sci. 94*, 826−830.
32. Bhuyan, A. K., and Udgaonkar, J. B. (1999) *Biochemistry 38*, 9158−9168.
33. Zaidi, F. N., Nath, U., and Udgaonkar, J. B. (1997) *Nat. Struct. Biol. 4*, 1016−1024.
34. Ramachandran, S., Rami, B. R., and Udgaonkar, J. B. (2000) *J. Mol. Biol. 297*, 733−745.
35. Khurana, R., and Udgaonkar J. B (1994) *Biochemistry 33*, 106−115.
36. Khurana, R., Hate, A., Nath, U., and Udgaonkar, J. B. (1995) *Protein Sci. 4*, 1133−1144.
37. Nath, U., and Udgaonkar, J. B. (1997) *Biochemistry 36*, 8602−8610.
38. Swaminathan, R., Nath, U., Udgaonkar, J. B., Periasamy, N., and Krishnamoorthy, G. (1996) *Biochemistry 35*, 9150−9157
39. Strambini, G. B., Gabellieri, E., Gonnelli, M., Rahuel-Clermont, S., and Branlant, G. (1998) *Biophys. J. 74*, 3165−3172.
40. Schreiber, G., Buckle, A. M., and Fersht, A. R. (1994) *Structure 2*, 945−951.
41. Nath, U., and Udgaonkar, J. B. (1995) *Biochemistry 34*, 1702−1713.
42. Schellman, J. A. (1978) *Biopolymers 7*, 1305−1322.
43. Kronman, M. J., Hoffman, W. B., Jeroszko, J., and Sage, G. W. (1974) *Biochim. Biophys. Acta 285*, 124−144.
44. Kuwajima, K., Ogawa, Y., and Sugai, S. (1979) *Biochemistry 18*, 878−882.
45. Shortle, D., and Meeker, A. K. (1986) *Proteins: Struct. Funct. Genet. 1*, 81−89.
46. Bierzynski, A., and Baldwin, R. L. (1982) *J. Mol. Biol. 162*, 173−186.
47. McWherter, C. A., Haas, E., Leed, A. R., and Scheraga, H. A (1986) *Biochemistry 25*, 1951−1963.
48. Amir, D., and Haas, E. (1988) *Biochemistry 27*, 8889−8893
49. Grathwohl, C., and Wuthrich, K. (1976) *Biopolymers 15*, 2025−2041.
50. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) *J. Mol. Biol. 201*, 161−200.
51. Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) *Science 257*, 1559−1563.
52. Pace, C. N., and Vandenburg, K. E (1979) *Biochemistry 18*, 288−292.
53. Nolting, B., Golbik, R., Soler-Gonzalez, A. S., and Fersht, A. R (1997b) *Biochemistry 36*, 9899−9905.
54. Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R., and Daggett, V. (2000) *J. Mol. Biol. 296*, 1257−1282.
55. Smith, L. N., Fiebeg, K. M., Schwalbe, H and Dobson, C. M. (1996) *Folding Des. 1*, R95−106.
56. Thelma, A. P., Daizo, H., Lorna, J. S., Fabrizio, C., Niccolo, T., Massimo, S., and Dobson, C. M (2000) *Protein Sci. 9*, 1466−1473.
57. Silow, M., and Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. 94*, 6084−6086.
58. Bai, Y. (1999) *Proc. Natl. Acad. Sci. 96*, 477−480.
59. Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
60. Kemmink, J., and Creighton, T. E (1995) *J. Mol. Biol. 245*, 251−260.
61. Wu, W. J., and Raleigh, D. P. (1998) *Biopolymers 45*, 381−394.
62. Juminaga, D., Wedemeyer, W. J., Garduno-Juarez., McDonald, M. A., and Scheraga, H. A. (1997) *Biochemistry 36*, 10131−10145.
63. Steinberg, I. Z., Harrington, W. F., Berger, A., Sela, M., and Katchalski, E. (1960) *J. Am. Chem. Soc. 82*, 5263−5279.
64. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) *J. Mol. Biol. 279*, 449−460.
65. Agashe, V. R., Shastry, M. C. R., and Udgaonkar, J. B. (1995) *Nature 377*, 754−757.
66. Cook, K. H., Schmid, F. X., and Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157−6161.
67. Schmid, F. X. (1986) *FEBS Lett. 198*, 217−220.
68. Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983) *Anal. Biochem. 130*, 134−145.

BI011701R