Rationally Designing the Accumulation of a Folding Intermediate of Barnase by Protein Engineering

Jesús M. Sanz and Alan R. Fersht*

MRC Unit for Protein Function and Design, Cambridge Centre for Protein Engineering, Medical Research Council Centre, Cambridge CB2 2QH, U.K.

Received July 2, 1993; Revised Manuscript Received September 15, 1993®

ABSTRACT: A method for the stabilization of transient folding intermediates is presented. Barnase folds and unfolds via such an intermediate. Mutations that destabilize the folded state relative to the folding intermediate had been previously identified from the free energy profiles for the unfolding of mutant proteins. It is predicted that the accumulation of such mutations should lead to the intermediate being the most stable species at certain concentrations of denaturant. Mutants were prepared that contained combinations of such mutations. The behavior of these mutants on urea denaturation was studied by probes for tertiary structure (fluorescence, near-UV CD), secondary structure (far-UV CD), and hydrodynamic volume (size-exclusion chromatography). Whereas wild-type shows a two-state transition in all cases, with the same thermodynamic values being found by all probes, some of the mutants show different transitions with different structural probes. On increasing concentration of denaturant, the tertiary structure of these mutants is lost before all the secondary structure and before the protein shows the maximum expanded volume that is characteristic of the unfolded state. These mutants thus accumulate an intermediate state at equilibrium under certain urea concentrations. The intermediate state retains some degree of secondary structure but has a disrupted tertiary structure, and its degree of compactness is intermediate between the folded and the unfolded forms, probably expanding with increasing concentration of denaturant. The accumulation of the intermediate should allow its direct characterization by spectroscopy, especially NMR.

Elucidation of the mechanisms that drive the acquisition of the native three-dimensional structure of a protein from its unfolded state (the protein folding problem) is one of the major aims of protein chemistry [see Creighton (1990) for a review]. Solving the pathway of protein folding requires characterizing the structures of all species involved on the progression from the unfolded state to the final product, determining their energy levels and the rate constants for their interconversion. Since a completely random search mechanism of protein folding can be ruled out because of the huge number of possible conformations (Levinthal, 1968), definite folding pathways must be present. There is now evidence for partly folded intermediates on folding pathways, and several models of folding have been proposed to accommodate these accordingly (Dill, 1985, Ptitsyn, 1987; Oas & Kim, 1988; Kuwajima, 1989; Kim & Baldwin, 1990). These species are usually very short-lived and less stable than the fully-folded form, and so they are not significantly populated at equilibrium. Conversely, stable conformations of proteins that are not fully folded or fully unfolded can be found under several nonnative conditions [high ionic strength, extreme pHs, organic solvents, mild denaturant concentrations; see Kuwajima (1989) for a review]. Most of these states share the feature of retaining considerable native-like secondary structure but having a disrupted tertiary structure. In view of their physical properties, they have been called "molten globules" (Ptitsyn, 1987; Goto & Fink, 1989; Kuwajima, 1989) or "compact intermediates" (Creighton, 1990). It has been postulated that these compact intermediates could be folding intermediates (Kuwajima, 1989; Creighton, 1990; Ptitsyn et al., 1990).

The folding of the ribonuclease of *Bacillus amyloliquefaciens* (barnase) does not follow a simple two-state process but involves at least one significant intermediate that accumulates transiently on the folding pathway (Matouschek et al., 1990, 1992):

$$\text{unfolded} \overset{\text{fast}}{\rightleftarrows} \ldots \rightleftarrows \text{intermediate} \underset{k_u}{\overset{k_{-u}}{\rightleftarrows}} \text{folded} \qquad (1)$$

(where k_{-u} and k_u represent the refolding and unfolding rate constants, respectively). The structure of the intermediate has been mapped in considerable detail from a kinetic analysis of a large number of barnase mutants (Matouschek et al., 1992). It is desirable to extend the analysis by studying the intermediate state directly. The major problem in doing so is that only about 0.1% of the protein is present as this state at equilibrium in water.

Strategy for Engineering the Accumulation of the Folding Intermediate. The free energies of unfolding of proteins have been found to vary linearly with the concentration of urea or GuHCl[1] (Pace 1986):

$$\Delta G_{U\text{-}F} = \Delta G_{U\text{-}F}^{H_2O} - m_{U\text{-}F}[D] \qquad (2)$$

where $\Delta G_{U\text{-}F}$ is the observed free energy of unfolding at a particular concentration of denaturant D, $\Delta G_{U\text{-}F}^{H_2O}$ is the free energy of unfolding in water, and $m_{U\text{-}F}$ is a constant that is proportional to the increase in solvent exposure of the protein on unfolding. A similar equation should account for the free energy of formation of partly unfolded folding intermediates

* To whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* November 15, 1993.

[1] Abbreviations: GuHCl, guanidine hydrochloride; ANS, 1-anilinonaphthalene-8-sulfonic acid; CD, circular dichroism; UV, ultraviolet; SEC, size-exclusion chromatography; $NaDodSO_4$–PAGE, electrophoresis in polyacrylamide gels in the presence of sodium dodecyl sulfate; ΔG(Fl), free energy associated with the 315 nm spectrofluorimetric transition; NMR, nuclear magnetic resonance spectroscopy.

0006-2960/93/0432-13584$04.00/0

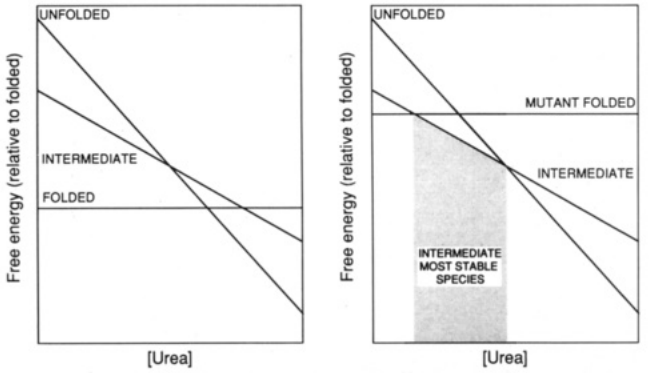

FIGURE 1: Free energy dependence with urea concentration of the folded, intermediate, and unfolded states of wild-type barnase (left) and of a very unstable mutant with $\phi = 0$ (right).

[e.g., as found by Ikeguchi et al. (1992)]:

$$\Delta G_{I\text{-}F} = \Delta G_{I\text{-}F}^{H_2O} - m_{I\text{-}F}[D] \quad (3)$$

where definitions corresponding to those for the unfolded state apply to the free energy of formation of I from F. Note that $m_{U\text{-}F} > m_{I\text{-}F}$ since there must be a greater increase in exposure of buried groups on the complete unfolding of a protein than there is on the formation of a partly unfolded intermediate. Equations 2 and 3 are illustrated in Figure 1. For wild-type barnase (Figure 1, left), the free energy of the intermediate is never sufficiently low to allow this state to accumulate. However, a selective destabilization of the folded state should allow the intermediate to be the most stable state at moderate concentrations of denaturant (Figure 1, right). Suitable mutations do exist. The mutations Ile → Ala4, Ile → Val25, Ile → Val51, Ile → Val76, and Tyr → Phe78, for example, all have essentially no effect on the stability of the intermediate compared with the unfolded state, but each raises the energy of the folded state by 0.9–1.5 kcal mol^{-1}. The mutations at positions 25, 51, and 78 are in hydrophobic core$_2$, a subdomain that forms a self-contained unit with very few contacts with the rest of the protein (core$_2$; Serrano et al., 1992a) and is completely disrupted in the intermediate (Matouschek et al., 1990, 1992).

We now present evidence for the accumulation of the main folding intermediate of barnase in mutants which were constructed by combining the above single mutations that destabilize selectively the folded state. The existence of intermediates at equilibrium can be revealed by either finding a biphasic transition upon denaturation using a single structural probe or noncoincidence of denaturation transitions monitored by different structural probes (Kim & Baldwin, 1990). We have followed the changes in secondary (by far-UV CD) and tertiary (by near-UV CD and intrinsic fluorescence) structure, as well as the degree of compactness (by size-exclusion chromatography) of the multiple mutants of barnase upon urea denaturation.

MATERIALS AND METHODS

Materials. The buffer used in the equilibrium denaturation studies, unless otherwise stated, was 50 mM 2-(*N*-morpholino)-ethanesulfonic acid (Mes), pH 6.3 (Sigma) (19.35 mM acid form and 30.65 mM sodium salt). Urea was enzyme grade from Bethesda Research Laboratories, MD. SP-Trisacryl was obtained from IBF, Villeneuve la Garenne, France. *Escherichia coli* BL21(DE3) [pLysS] was a generous gift from Dr. F. W. Studier. 1-Anilinonaphthalene-8-sulfonic acid (ANS) was from Molecular Probes, Inc.

Site-Directed Mutagenesis. Oligonucleotide-directed mutagenesis was performed as described elsewhere (Serrano et al., 1992b) using the kit supplied by Amersham. The mutant Ile76 → Thr (I76T) was prepared using the oligonucleotide 5′-GATGTATAGTTA*G*T*ATCCGCTT-3′ (an asterisk follows the mismatched bases). The multiple mutants were constructed using the single-stranded DNA of existing single or double mutants as templates. Mutations were identified by DNA sequencing of the whole barnase gene.

Expression and Purification of Barnase. Wild-type barnase and its mutants were expressed in *E. coli* BL21(DE3) [pLysS] cells harboring a pTZ18U plasmid containing the barnase gene and were purified as previously described (Serrano et al., 1990). Owing to an unusual, nonspecific affinity to the Trisacryl resin of the mutants described in this work, only up to 30% of the protein could be released using high ionic strength buffer. The remaining protein had to be eluted with the same buffer plus 6 M urea. Salts and urea were removed after dialysis against water at 4 °C (25 L in five changes). Since the proteins show a similar affinity for the Mono-S FPLC column (Pharmacia), we changed the chromatographic step. The samples were lyophilized after dialysis, subsequently redissolved in a small volume of 50 mM acetate buffer, pH 4.5, plus 0.15 M NaCl, and then applied to a Superdex G-75 FPLC gel filtration column (10 × 300 mm) (Pharmacia) equilibrated in the same buffer. Chromatography was run at room temperature at a rate of 0.6 mL/min. Barnase was purified in this way to electrophoretic homogeneity, as determined by NaDodSO$_4$–PAGE analysis. The purified proteins were dialyzed against water, flash-frozen, and stored in aliquots. No appreciable differences in kinetic and thermodynamic behavior were found in the proteins purified with either the standard or the new procedure.

Urea Denaturation Followed by Fluorescence Spectroscopy. Changes in the intrinsic fluorescence of barnase upon urea denaturation were followed in an Aminco Bowman Series 2 spectrofluorimeter (American Instruments Co.) at 25 °C with excitation at 290 nm (bandpass = 2 nm) and emission at 315 nm (bandpass = 8 nm), as previously described (Kellis et al., 1989).

Urea Denaturation Followed by Circular Dichroism. CD measurements were made in a Jasco J-720 spectropolarimeter fitted with a thermostated cell holder and interfaced with a Neslab RTE-110 water bath at 25 °C. Spectra were acquired at a scan speed of 20 nm/min and averaging six scans. The response time was 2 s. In the far-UV region (250–200 nm), the concentration of barnase was 0.135 mg mL^{-1}, and a cuvette of 0.1-cm path length was used. In the near-UV (320–250 nm), the sample concentration was 0.45 mg mL^{-1}, and the path length of the cuvette was 1.0 cm. Ellipticities are expressed in deg·cm^2·dmol^{-1}, using the mean residue concentration in the far-UV or the protein concentration in the near-UV experiments. The urea-denaturation experiments were performed as follows. A sample with the corresponding urea concentration was placed in the cuvette, and the CD spectrum was recorded. Then aliquots from an 8 M urea stock solution were added stepwise, and the spectrum was recorded after 5 min to allow equilibration. When the cuvette was full, a new sample with a concentration of urea somewhat lower than the last point was prepared, and the process was repeated again. The overlapping of different parts of the curve ensured the detection of any possible systematic error that could originate in a titration segment. For every spectrum, the buffer baseline was subtracted and the ellipticity corrected for protein dilution.

Table I: Free Energy of Unfolding Determined Spectrofluorimetrically[a]

mutant	m	$[urea]_{1/2}$ (M)	$\Delta G^{H_2O}_{U-F}$	$\Delta\Delta G^{[urea]_{1/2}}_{U-F}$	$\Delta\Delta G^{[urea]_{1/2}}_{U-F}$ expected[b]
wild-type	1.92 ± 0.03	4.58 ± 0.01	8.79 ± 0.14		
I4A/I76V	1.92 ± 0.01	3.61 ± 0.02	6.93 ± 0.05	1.86 ± 0.15	2.18
I76T	1.87 ± 0.08	3.26 ± 0.02	6.10 ± 0.26	2.53 ± 0.15	
I4A/Y78F	1.95 ± 0.15	3.21 ± 0.02	6.26 ± 0.48	2.63 ± 0.15	2.70
I4A/I51V	2.11 ± 0.12	2.92 ± 0.06	6.16 ± 0.37	3.19 ± 0.19	3.14
I4A/I25V/I51V/Y78F	1.95 ± 0.05	1.84 ± 0.04	3.59 ± 0.12	5.23 ± 0.16	5.61

[a] All energies are expressed in kcal mol^{-1}. [b] Calculated from summing the values of the single mutations determined by Serrano et al. (1992b).

Urea Denaturation Followed by Size-Exclusion Chromatography (SEC). Changes in hydrodynamic volumes in the presence of urea were monitored by gel filtration chromatography in a Superdex G-75 FPLC gel filtration column (10 × 300 mm) (Pharmacia), equilibrated in a buffer containing 50 mM Mes, pH 6.3, plus 0.15 M NaCl and the corresponding concentration of urea. Samples of barnase (0.1 mg/mL) were prepared in the same buffer as present in the column, allowed to equilibrate for at least 30 min, and then applied to the column. The chromatography was run at a flow rate of 0.75 mL min^{-1}.

ANS Binding. Barnase samples (2 μM) were prepared in 50 mM Mes, pH 6.3, at 25 °C, with the corresponding concentration of urea. ANS aliquots were added to give a concentration that ranged from 1 to 1000 times that of the protein. A similar solution lacking the protein was also prepared as a blank. Binding to the folded state at pH 2.7 (20 mM glycine-HCl buffer) was followed at 4 °C. ANS fluorescence was measured on excitation at 380 nm and emission at 480 nm. The possibility of aggregation was determined by centrifuging the samples after addition of ANS (13 000 rpm, 5 min) and comparing the fluorescence of the protein and ANS before and after the centrifugation.

Kinetic Experiments. All measurements were carried out at 25 °C. Refolding experiments were performed by pH-jump using an Applied Photophysics SF.17MV stopped-flow apparatus, as previously described (Matouschek et al., 1992). Excitation wavelength was 290 nm. Emission was monitored at wavelengths greater than 315 nm using a cut-off filter. The monochromator slits were set to 2-nm bandpass. Data were analyzed using the software supplied by Applied Photophysics. Refolding follows a triple-exponential process, where the fastest of the phases (70–80% of the total amplitude) represents the conversion from the folding intermediate to the folded form (k_{-u}, thereafter). The two slow phases have been related to *cis–trans* proline isomerization events. A linear relationship of log k_{-u} vs [urea] was calculated with points obtained at very low concentrations of urea, where the concentration of intermediate outweighs the concentration of unfolded form.

Unfolding was initiated by dilution of 1 volume of protein plus 10 volumes of the corresponding urea solution, using a Perkin Elmer MPF 44B fluorescence spectrophotometer, as described elsewhere (Serrano et al., 1992c). Calculation of unfolding constants in water was performed using the equation:

$$\log k_u = \log k_u(H_2O) + m^*_{ku}[urea] - 0.014[urea]^2 \quad (4)$$

which fits the unfolding data better than the usual linear approximation of $\log k_u = \log k_u(H_2O) + m[urea]$ (Matouschek & Fersht, 1993; A. Matouschek, J. Matthews, and A. R. Fersht, unpublished results).

RESULTS

Description of Mutants. The single mutant I76T, the double mutants I4A/I51V, I4A/I76V, and I4A/Y78F, and the quadruple mutant I4A/I25V/I51V/Y78F were obtained and analyzed. The structural characteristics of the single mutations had previously been described by Serrano et al. (1992b). The mutation I4A deletes the interactions with α-helix$_1$ and loop$_1$; I25V deletes the van der Waals interactions with α-helix$_2$, α-helix$_3$, and β-strand$_1$; I51V deletes the van der Waals interactions with α-helix$_2$, α-helix$_3$, loop$_1$, and loop$_2$; I76V deletes the van der Waals interactions with α-helix$_1$, loop$_1$, and β-strand$_3$; finally, Y78F deletes the hydrogen bonds of the tyrosine side-chain OH group with the main chain NH and CO groups of Gly81. The mutant I76T was used as a control mutant with a degree of stability comparable to the double mutants but which has the intermediate destabilized.

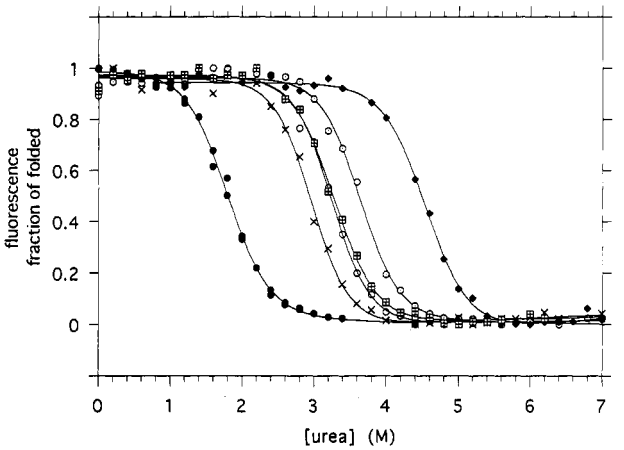

FIGURE 2: Urea denaturation of barnase mutants followed by intrinsic fluorescence at 315 nm. (●) I4A/I25V/I51V/Y78F; (×) I4A/I51V; (○) I4A/Y78F; (⊞) I76T; (⊙) I4A/I76V; (◆) wild-type. The raw data are presented, including the slopes in the baselines, except that the values are normalized to 1 and 0 for the fluorescence of the native (at 0 M urea) and unfolded proteins, respectively.

Yields of mutants (0.7 mg/L of culture) were usually much lower than that of wild-type and showed a great tendency to stick to the Trisacryl-SP resin as well as to the mono-S column by nonionic interactions. For this reason, in order to increase the recovery of the proteins, we modified the standard purification strategy. As an additional procedure, the remaining barnase in the resin was eluted adding 1 M NaCl and 6 M urea to the 50 mM acetate buffer, pH 4.5. Salts and urea were subsequently removed by dialysis. After this step, we purified the proteins by size-exclusion chromatography using a Superdex G-75 gel filtration column. The CD spectra, as well as the kinetic and denaturation equilibrium parameters of the mutants thus purified and of the same proteins purified with the standard procedure, are indistinguishable (data not shown).

Fluorescence Studies. In order to calculate the free energy of unfolding, equilibrium constants for urea-induced denaturation (Table I) were determined spectrofluorimetrically at 315 nm (Figure 2) using eq 2 (Kellis et al., 1989; Serrano et al., 1992b; Clarke & Fersht, 1993) where m_{U-F} is found to be close to 1.92 kcal mol^{-2} for wild-type and most mutants

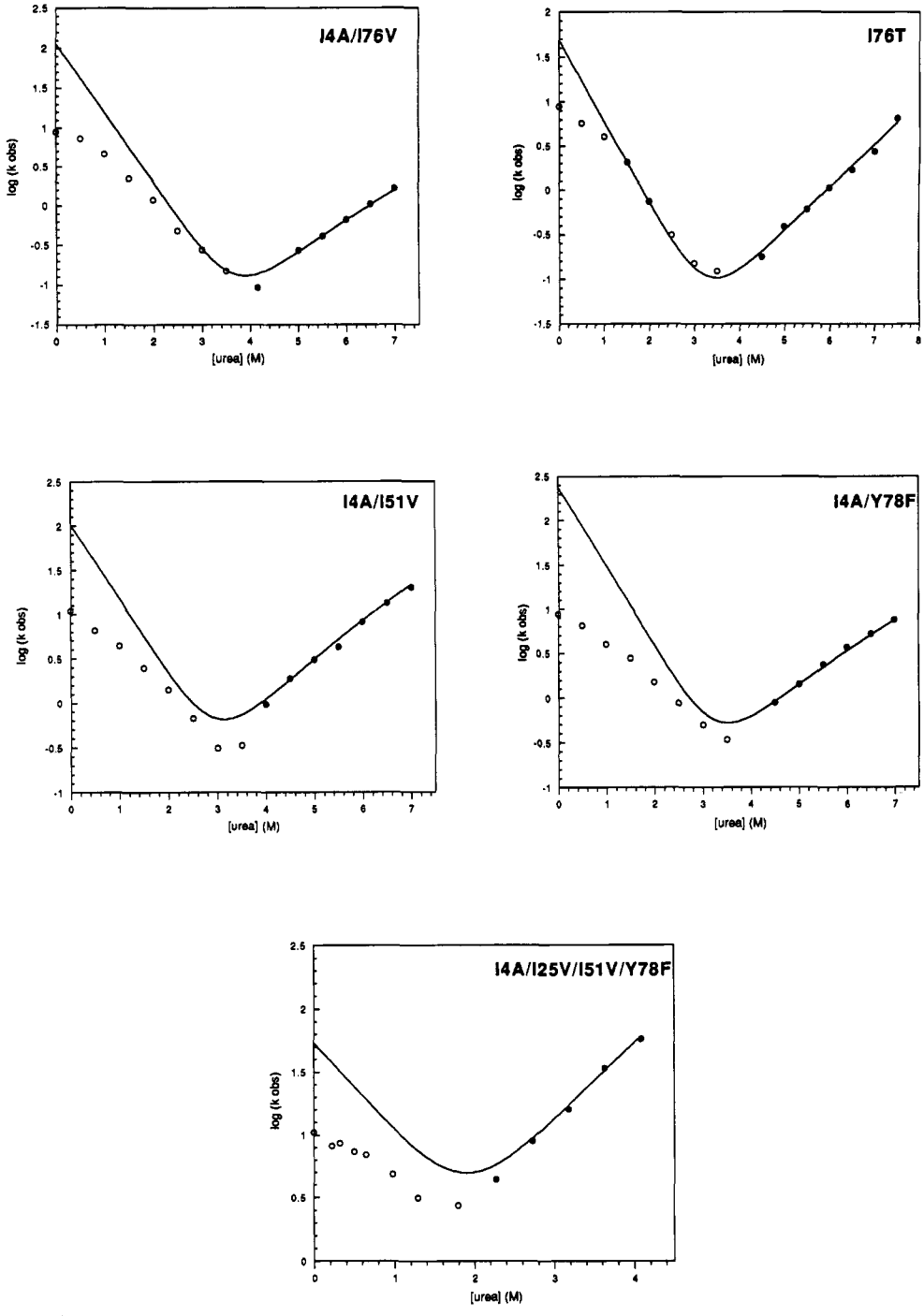

FIGURE 3: Urea dependence of rate constants for unfolding (●) and refolding (○) of wild-type and mutants. The major of the three phases of the refolding process is analyzed. Rate constant units are s^{-1}. Scatter points represent the experimental observed rate constants, and the solid line represents the expected behavior for a two-state transition (Matouschek et al., 1990).

(Serrano et al.,1992b). In previous studies on barnase, the concentration of intermediate states at equilibrium is negligible, and the loss of tertiary structure (monitored by fluorescence) is concomitant with the loss of the overall structure. All fluorimetric transitions of the multiple mutants of this work fit well to a two-state process with an m value within the range found for wild-type and its single mutants (Figure 2, Table I). The most reliable way of measuring changes in stability is to use the changes measured at the mean of the concentrations of urea for 50% denaturation, $\Delta\Delta G_{U\text{-}F}([\text{urea}]_{1/2})$, (Serrano et al., 1992b). Also listed in Table I are the values of $\Delta\Delta G_{U\text{-}F}([\text{urea}]_{1/2})$ expected from the simple addition of the destabilization energies of the single mutants (Serrano et al.,1992b), assuming that there is no cooperativity in the multiple mutant. The experimental values of $\Delta\Delta G_{U\text{-}F}$ agree quite nicely with the expected values when $\Delta G_{U\text{-}F}([\text{urea}]_{1/2})$ is used. The quadruple mutant I4A/I25V/I51V/Y78F has the highest deviation (0.38 kcal/mol).

Kinetic Studies. The logarithms of the refolding and unfolding rate constants together with the expected behavior for a two-state transition, calculated as described by Matouschek et al. (1990), are shown in Figure 3. The mutant I4A/I76V shows a behavior similar to that of wild-type (Matouschek et al., 1990) presumably because the selective destabilization of the folded state is insufficient to allow a significant accumulation of intermediate. The mutants I4A/I51V, I4A/Y78F, and I4A/I25V/I51V/Y78F show, however, very distorted kinetics, even in the transition region. I76T is as unstable as I4A/I51V and I4A/Y78F, but it deviates much less from the two-state behavior. Kinetic parameters were

Table II: Kinetic Quantities for Refolding and Unfolding[a]

mutant	$\log k_{-u}^{H_2O}$	$m_{k\text{-}u}$[b]	$\log k_{u}^{H_2O}$	m_{ku}^{*}	$\Delta G_{I\text{-}U}^{H_2O}$	$\Delta G_{I\text{-}F}^{H_2O}$	m_i[c]	Φ_i (4 M urea)[d]
wild-type	1.08 ± 0.01	−0.39 ± 0.02	−4.59 ± 0.01	0.65 ± 0.01	−1.06 ± 0.14	7.73 ± 0.02	1.41 ± 0.02	
I4A/I76V	0.95 ± 0.01	−0.30 ± 0.07	−3.04 ± 0.05	0.56 ± 0.01	−1.51 ± 0.16	5.42 ± 0.07	1.15 ± 0.08	0.27
I76T	1.00 ± 0.01	−0.33 ± 0.02	−3.46 ± 0.04	0.67 ± 0.01	−0.31 ± 0.15	6.06 ± 0.06	1.34 ± 0.02	0.41
I4A/Y78F	0.95 ± 0.01	−0.34 ± 0.05	−2.15 ± 0.08	0.53 ● 0.01	−1.95 ± 0.16	4.21 ± 0.08	1.18 ± 0.06	0.00
I4A/I51V	1.02 ± 0.01	−0.39 ± 0.03	−2.11 ± 0.09	0.59 ± 0.01	−1.35 ± 0.16	4.25 ± 0.07	1.33 ± 0.04	0.00
I4A/I25V/I51V/Y78F	1.05 ± 0.01	−0.30 ± 0.02	−0.89 ± 0.05	0.71 ± 0.01	−0.93 ± 0.16	2.63 ± 0.07	1.37 ± 0.03	0.06

[a] Rate constants are expressed in s^{-1} and energies in kcal mol^{-1}. Data were determined at 25 °C and pH 6.3. [b] The slope of the logarithm of the refolding rate constants k_{-u} against [urea] at very low [urea]. $k_{-u}^{H_2O}$ is the value of k_{-u} in water. [c] Calculated from kinetics as $m_i = 2.303RT(m_{ku}^{*} - m_{k\text{-}u})$. [d] Φ_i is defined by $[1 - (\Delta\Delta G_{I\text{-}F}/\Delta\Delta G_{U\text{-}F})]$; see text.

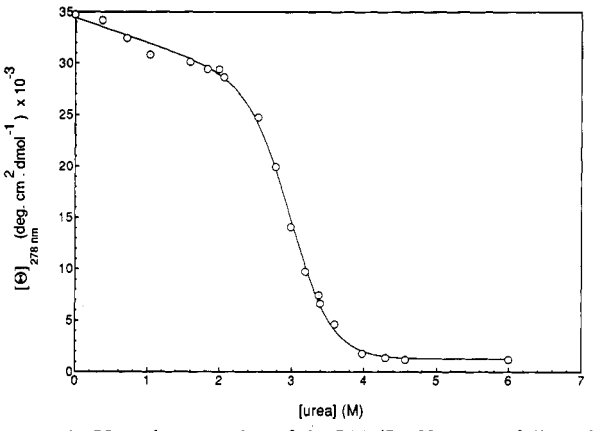

FIGURE 4: Urea denaturation of the I4A/I51V mutant followed by CD in the near-UV (278 nm). Fitting to a two-state transition is shown in solid line.

calculated for all these mutants (Table II). The ϕ values [$\phi = 1 - (\Delta\Delta G_{I\text{-}F}/\Delta\Delta G_{U\text{-}F})$], where $\Delta\Delta G_{I\text{-}F}$ is the change in $\Delta G_{I\text{-}F}$ on mutation and $\Delta\Delta G_{U\text{-}F}$ is the change in $\Delta G_{U\text{-}F}$, relative to wild-type (Matouschek et al., 1990), were calculated in 4 M urea instead of water, since the extrapolation is not so large and the data are more accurate. I76T has a high ϕ value, consistent with a high destabilization of the intermediate. I4A/I76V has a moderate value. On the other hand, the mutants I4A/I51V, I4A/Y78F, and I4A/I25V/I51V/Y78F have a ϕ value of 0, so the intermediate should not be destabilized in these mutants.

The values of m_{ku}^{*} from eq 4 are strongly dependent on mutation, the highest being the value for the quadruple mutant. This and the I4A/I51V and I4A/Y78F mutants seem to be suitable for accumulating sufficient intermediate (Tables I and II). Nevertheless, I4A/I51V was the easiest to express, and most of the experiments were carried out with this mutant.

Circular Dichroism Studies. (1) Near-UV CD. The titration of the CD signal at 278 nm of the I4A/I51V mutant with urea fits a two-state equation with sloping baselines (Figure 4). The values of both $[\text{urea}]_{1/2}$ (the value of [urea] at 50% denaturation) and m are similar to those obtained by spectrofluorimetric titrations (Table III).

(2) Far-UV CD. The far-UV CD spectrum of wild-type barnase is relatively weak and does not resemble typical spectra of mixtures of secondary structures (Sancho & Fersht, 1992). Very low spectral intensities indicate extensive cancelling of positive and negative CD bands. The region around 230 nm is a local minimum and gives the best signal for monitoring conformational changes. However, this may not be a pure secondary structure probe because it also reports changes on the contribution of Trp94 to the spectrum (S. Vuilleumier, unpublished results). Changes in $[\theta]_{230}$ of wild-type barnase on urea titration fit well to a two-state equation and give the same values of m and $[\text{urea}]_{1/2}$ found by other techniques (Table III). The I4A/I51V mutant, however, shows a less steep transition for $[\theta]_{230}$ than those monitored by fluorescence or $[\theta]_{278}$ (Table III). The I4A/Y78F and I4A/I25V/I51V/Y78F mutants exhibit similar behavior monitored at 230 nm (not shown). The deviations from those found for the pure tertiary structure transitions become much more obvious when measuring CD changes at 222 nm, a wavelength that gives a poorer signal but that reports primarily changes in secondary structure (Figure 5, Table III). It is evident here that the transition becomes much shallower, consistent with a very low m value, and that the midpoint of the transition is no longer coincident with that measured from fluorescence, but occurs at a higher value of [urea]. Data recorded at 215 nm exhibited the same tendency (not shown). Wild-type barnase, however, displays virtually no difference with the fluorescence curve (Figure 5). This result strongly suggests that unfolding of wild-type barnase does not involve the accumulation of species other than the fully folded and the fully unfolded states.

It seems clear, from the curves shown in Figure 5, that some species lacking the characteristic tertiary structure, but with significant secondary structure, does exist at moderate urea concentrations above the midpoint measured from fluorescence urea titrations for some mutants, for example, at about 4 M urea for the I4A/I51V mutant (1 M above the fluorescence midpoint, where very little amounts of folded state should be present). The CD spectra of the folded and the unfolded forms of I4A/I51V, together with the spectra recorded at 4 M urea, are shown in Figure 6 (the spectrum below 215 nm is unreliable due to the high noise produced by the absorption of the denaturant). Whereas the near-UV spectrum at 4 M urea is indistinguishable from that of the unfolded form, the far-UV spectrum shows that not all the secondary structure is lost. A similar result is found for the I4A/Y78F and the I4A/I25V/I51V/Y78F mutants, whereas the extent of remaining structure for wild-type or I76T is almost negligible in those conditions (data not shown).

Size-Exclusion Chromatography. The hydrodynamic volumes of the folded and the unfolded forms of the proteins can be distinguished by measuring their elution volumes by size-exclusion chromatography at different concentrations of urea (Corbett & Roche, 1984). The unfolded form, since it is less compact, should be more excluded from the column than the folded form. If the equilibration is on a slower time scale than the time of chromatography, both species can be separated as individual elution peaks under the appropriate denaturant concentration. If, on the other hand, the equilibration is very fast, only one peak is observed, with an elution volume that is a weighted mean value of that for the population of the folded and unfolded forms present at equilibrium. The accumulation of a third different form should modify the elution profile obtained for a simple two-state transition. Samples of the wild-type and the double mutant I4A/I51V were chromatographed on a gel filtration column equilibrated at different urea concentrations. It was found necessary to

Table III: Values of m and $[\text{urea}]_{1/2}$ for Equilibrium Denaturation of Wild-Type and the I4A/I51V Mutant with Different Structural Probes[a]

probe	structure probed	wild type		I4A/I51V	
		m	$[\text{urea}]_{1/2}$ (M)	m	$[\text{urea}]_{1/2}$ (M)
fluorescence (315 nm)	tertiary structure	1.92 ± 0.03	4.58 ± 0.01	2.11 ± 0.12	2.92 ± 0.06
[Θ] (278 nm)	tertiary structure	1.99 ± 0.04	4.51 ± 0.01	1.96 ± 0.08	2.95 ± 0.01
[Θ] (230 nm)	secondary + tertiary structures	1.90 ± 0.09	4.63 ± 0.02	1.50 ± 0.11	3.03 ± 0.04
[Θ] (222 nm)	secondary structure	1.95 ± 0.15	4.66 ± 0.07	0.61 ± 0.09	3.62 ● 0.03
size-exclusion chromatography (+ 150 mM NaCl)	hydrodynamic volume	1.80 ± 0.20	5.16 ± 0.07	1.44 ± 0.18[b]	3.40 ± 0.06[b]
fluorescence (315 nm) (+ 150 mM NaCl)	tertiary structure	1.99 ± 0.15	5.27 ± 0.08	1.93 ± 0.13	3.40 ± 0.03

[a] Values obtained assuming a two-state model. [b] The data were fitted to a two-state model but more likely arise from a multistate equilibrium.

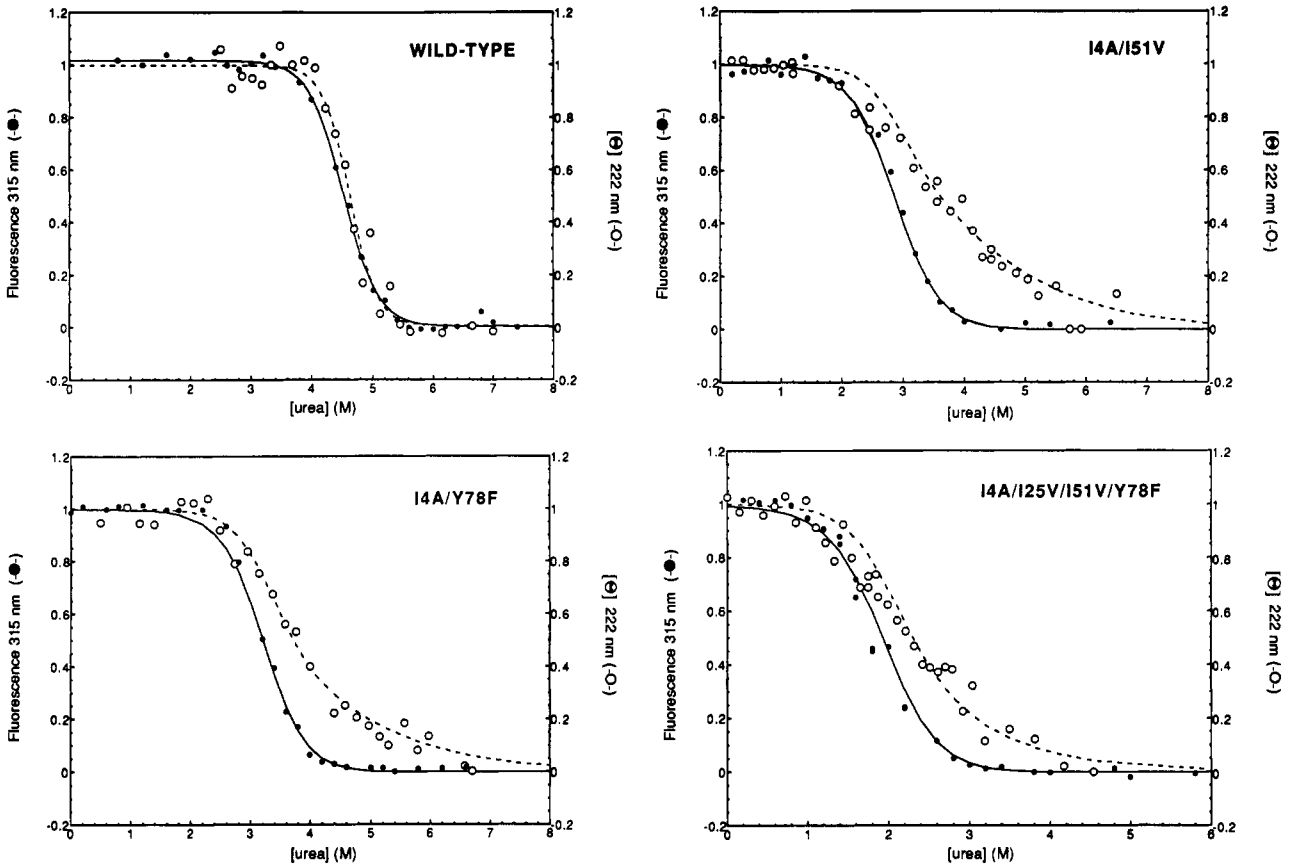

FIGURE 5: Urea denaturation profiles of wild-type barnase (top left) and the mutants I4A/I51V (top right), I4A/Y78F (bottom left), and I4A/I25V/I51V/Y78F (bottom right) as measured by fluorescence at 315 nm (●) and ellipticity at 222 nm (○). Results of the simulated fittings for the changes in fluorescence and ellipticity calculated for a three-state transition (eqs 5, 6, and 7) are shown in solid and dashed lines, respectively, except for wild-type, where the data fit a two-state transition for both fluorescence and CD. The raw data are normalized as in Figure 2.

add 150 mM NaCl to the Mes buffer to screen ionic interactions with the support. The stability of barnase is known to increase with increasing ionic strength (Serrano et al., 1990), which leads to an increase of about 10–15% on the $[\text{urea}]_{1/2}$ in 150 mM NaCl, but with negligible changes on the m value. The data on size exclusion must, therefore, be compared with other data at the higher ionic strength. Data for elution of wild-type and the double I4A/I51V mutant, together with the fluorescence titrations in the presence of the same ionic strength, are shown in Figure 7 and Table III. There is an artifact with both wild-type and the mutant that occurs on chromatography at low concentrations of urea which we interpret as a result of a nonspecific hydrophobic interaction with the support that is weakened in the presence of a relatively small concentration of urea. There are no conformational changes detectable by CD at such urea concentrations (data not shown). These data were ignored for curve fitting. The elution volumes for native wild-type and mutant proteins are identical, as are those for the fully denatured proteins at 8 M urea. However, whereas the data for the unfolding of wild-type enzyme measured by fluorescence and elution volume are superimposable when normalized (Figure 7A), those for the mutant I4A/I51V are not (Figure 7B). The elution data for the mutant can be fitted to a two-state equation assuming that there is a highly sloping baseline at high urea concentration, but this "slope" more likely results from a multistate transition.

The results of the fits to two-state transitions are summarized in Table III. The size-exclusion chromatography data for wild-type barnase are consistent with all the other data. As expected, the value of $[\text{urea}]_{1/2}$ is increased for wild-type and the I4A/I51V mutant corresponding to a higher ionic strength, being 3.40 M for the mutant. Fluorescence denaturation studies of this protein in the presence of 150 mM NaCl give an m value of 1.93 and an $[\text{urea}]_{1/2}$ of 3.40 M. However, the m value obtained by size-exclusion chromatography for this

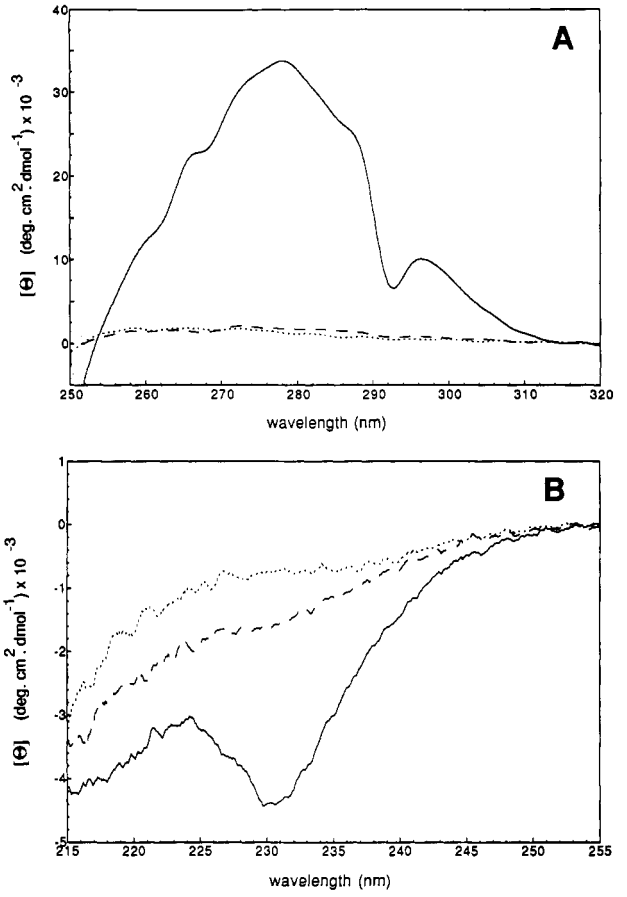

FIGURE 6: Near-UV (A) and far-UV (B) CD spectra of the I4A/I51V mutant in the presence of 0 M (solid line), 4 M (dashed line), and 7 M urea (dotted line).

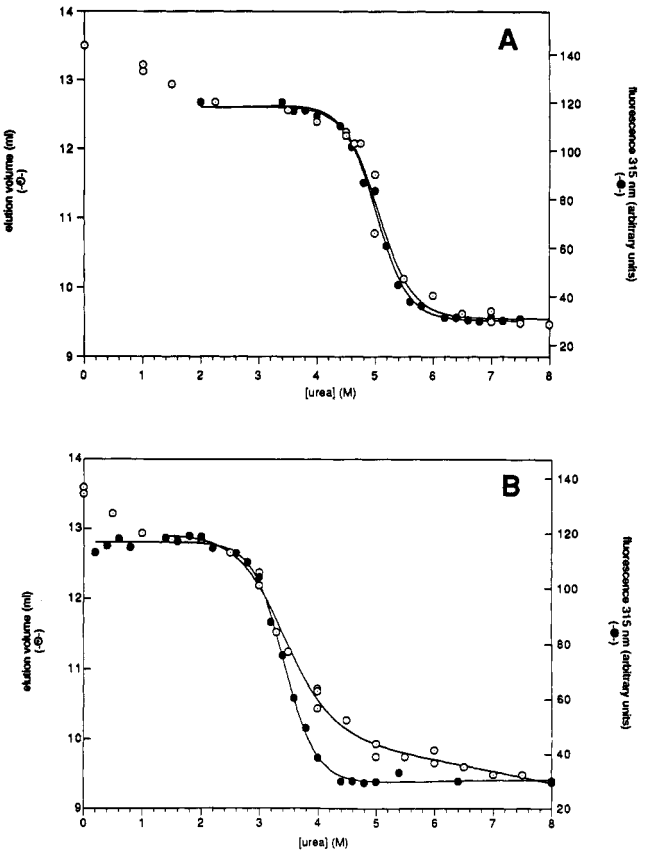

FIGURE 7: Urea denaturation of wild-type barnase (A) and the I4A/I51V mutant (B) in the presence of 150 mM NaCl: (●) intrinsic fluorescence at 315 nm; (○) elution volume measured by gel filtration in a Superdex G-75 column equilibrated in the same concentration of urea. The solid lines are the fits to a two-state transition.

mutant (1.44) is lower than for fluorescence, indicating a less cooperative transition, probably by accumulation of intermediate forms. The results from size-exclusion chromatography are not, however, as accurate as those from other techniques.

ANS Binding. The dye ANS binds to solvent-accessible hydrophobic regions of proteins, with generally an appreciable increase in fluorescence quantum yield (Semisotnov et al., 1991). The so-called "molten globules" can be detected this way because their flexible tertiary structure allows the access of the solvent (and small molecules such ANS) to the uncovered hydrophobic cores of the protein. Since this could be the state of the I4A/I51V mutant in 4 M urea (Figure 5), we assayed the binding of ANS to the mutant in such conditions as well as in the absence of denaturant. Nevertheless, in the range of ANS used (1–1000-fold the protein concentration), no difference was found with a control without the mutant. However, an increase in fluorescence could be found at pH 2.7, 4 °C, in the absence of urea, upon addition of a 10-fold excess of ligand, although no differences in secondary or tertiary structure with the folded state at pH 6.3 could be found by CD (data not shown).

Calculation of Free Energies. Kinetic and Thermodynamic Models. The evidence above implies that loss of tertiary structure does not appear to be necessarily linked to loss of secondary structure upon urea denaturation when there are intermediates that accumulate (Figure 5). This implies that the unfolding energy as measured from fluorescence techniques [i.e., ΔG(Fl)] in these cases does not reflect the global unfolding energy of the molecule, since changes in secondary structure are not measured. Thus, ΔG(Fl) is an underestimate of the true $\Delta G_{U\text{-}F}$. It is important to note that the far-UV CD data are generally of lower quality because of the weak signal from barnase in this region, so the calculations based on CD data have a higher error than those from fluorescence. We have analyzed first the fluorescence equilibrium data of the urea denaturation of I4A/I51V, I4A/Y78F, and I4A/I25V/I51V/Y78F mutants using a two-state model that does not necessarily apply. However, the values of m derived from the fluorescence data are in the expected range and the changes in energy do correspond with those expected from the sums of the single mutations.

Analysis Using a Three-State Model. Defining the equilibrium $[U]/[F] = K_{U\text{-}F}$, and $[I/[F] = K_{I\text{-}F}$, the fraction or the folded state at equilibrium is given by $f_F = [F]/([F] + [I] + [U]) = 1/(1 + K_{I\text{-}F} + K_{U\text{-}F})$, so that

$$f_F = 1/[1 + \exp(-\Delta G_{I\text{-}F}/RT) + \exp(-\Delta G_{U\text{-}F}/RT)] \quad (5)$$

Similarly, the fractions of intermediate and unfolded states are given by

$$f_I = [\exp(-\Delta G_{I\text{-}F}/RT)]/[1 + \exp(-\Delta G_{I\text{-}F}/RT) + \exp(-\Delta G_{U\text{-}F}/RT)] \quad (6)$$

$$f_U = 1 - (f_F + f_I) \quad (7)$$

The dependence of each equation on urea concentration is found from including eqs 2 and 3 in the above expressions. Equation 5 thus modified should describe the loss of tertiary interactions monitored by fluorescence on denaturation whereas eq 7 should describe the loss of secondary structure if both the folded and intermediate forms have similar CD signals.

Table IV: Free Energy Calculations of Barnase Mutants Using a Three-State Equilibrium Model[a]

mutant	$\Delta G_{U\text{-}F}$(Fl)[b]	$\Delta G_{I\text{-}F}$(kin)[c]	$\Delta G_{U\text{-}F}$(3st)[d]	$\Delta G_{I\text{-}F}$(3-st)[d]
I4A/I51V	5.60 – 1.92[urea]	4.25 – 1.33[urea]	6.15 – 1.99[urea]	4.50 – 1.51[urea]
I4A/Y78F	6.16 – 1.92[urea]	4.21 – 1.18[urea]	6.50 – 1.90[urea]	5.10 – 1.45[urea]
I4A/I25V/I51V/Y78F	3.56 – 1.92[urea]	2.63 – 1.37[urea]	4.30 – 2.10[urea]	3.20 – 1.40[urea]

[a] All energies are expressed in kcal mol^{-1}. [b] Values obtained with eq 2. The mean value of m of 1.92 is used, and the value of $\Delta G_{U\text{-}F}^{H_2O}$ is that calculated from using this value and the experimental value of $[\text{urea}]_{1/2}$. [c] Values obtained from kinetic data (Table II). [d] Values simulated with the three-state model (eqs 5, 6, and 7).

There are too many variables to apply these equations directly to the denaturation curves. We have, therefore, simulated the fluorescence and CD experimental data using eqs 5 and 7, which monitor the disappearance of the folded state and the appearance of the unfolded form. In a first approximation, $m_{U\text{-}F}$ can be calculated from the fluorescence denaturation curves (about 1.9), and a two-state fitting of the $[\theta]_{222}$ titration curve would closely report the intermediate to unfolded transition, so $m_{I\text{-}F}$ could be firstly guessed as m(fluorescence) – m(CD) (about 1.3). We then varied both the m and the ΔG values to obtain the fittings shown in Figure 5 and Table IV. The simulated data fit nicely with the experimental fluorescence data and reasonably well with the CD data. In the latter case, the major discrepancies can be seen in the low urea region, probably due to a slight contribution from the Trp-94 CD band even at 222 nm, which is more linked to the loss of tertiary structure in that urea concentration range.

DISCUSSION

Partly folded states have been characterized for several proteins. In a few cases, there is some evidence that these are genuine folding intermediates rather than alternative equilibrium states (Kuwajima, 1989; Ptitsyn et al., 1990). For example, stopped-flow CD data of the refolding of staphylococcal nuclease A (Sugawara et al., 1991) and α-lactalbumin (Kuwajima et al., 1985) suggest that the far-UV CD spectrum of its refolding intermediate is very similar to that of the forms found at low pH. Moreover, NMR studies on the trifluoroethanol-denatured state of hen egg-white lysozyme show that, in this state, the pattern of protection of several backbone amides from hydrogen exchange with the solvent resembles that found in an early kinetic intermediate detected in the refolding of this protein (Buck et al., 1993). The acid form of apomyoglobin seems also to be an intermediate in folding pathway (Barrick & Baldwin, 1993). On the other hand, stabilization of kinetic intermediates has also been performed by covalently trapping them (Creighton, 1986), although this approach is limited to disulfide-bridged proteins. In the present study, we have used an alternative approach to investigate the importance of intermediate states, i.e., energetically stabilize by site-directed mutagenesis an intermediate that has already been found as a transient in folding studies so that a significant amount of this state may be present at equilibrium. Protein engineering studies on mutants of barnase have allowed not only the mapping of the structure of its main folding intermediate but also have given clues for its stabilization. A combination of known mutations that decrease the stability of the folded state relative to the folding intermediate should allow its accumulation under certain concentrations of urea (Figure 1).

The loss of tertiary structure is most easily monitored by the change in the fluorescence of the tryptophan residues. Analysis of the denaturation of some of these mutants monitored by fluorescence, together with the refolding and unfolding kinetic studies (Tables I and II), indicates that the destabilization free energy ($\Delta\Delta G_{U\text{-}F}$) of the single mutations is generally additive. The quadruple mutant I4A/I25V/I51V/Y78F shows a small deviation of 0.38 kcal mol^{-1} for $\Delta\Delta G_{U\text{-}F}$. In all cases, although the kinetic studies indicate that there should be a certain amount of intermediate at equilibrium for some mutants, fluorescence denaturation studies still show a two-state transition and so does the near-UV CD titration (Figures 2 and 4). However, far-UV CD shows appreciable differences for the I4A/I51V, I4A/Y78F and I4A/I25V/I51V/Y78F mutants with respect with wild-type (Figure 5). Whereas in the latter case no differences can be found irrespective of the wavelength used to monitor the unfolding (Table III), which indicates that the secondary and the tertiary structures of wild-type denature simultaneously, this does not apply to the mutants. In the case of the double I4A/I51V mutant, when $[\theta]_{230}$ is measured, the m value drops to about 1.50; and when the signal at 222 nm is used, the transition becomes even shallower (m = 0.61) and occurs in a much wider range of urea concentration (Figure 5). Although both wavelengths belong to the far-UV CD region, the value of $[\theta]_{230}$ has not only contributions from the secondary structure but also a major contribution from Trp94 that depends on its tertiary environment. Thus, at this wavelength we may be seeing conformational changes in both the secondary and the tertiary structures of the protein. On the other hand, $[\theta]_{222}$ should reflect primarily changes in secondary structure, particularly from α-helices. The loss of secondary structure appears to be a noncooperative process, which would suggest that the intermediate is not a unique species but rather an ensemble of conformations that progressively unfold upon increasing urea concentration. This is similar to the model recently proposed for the unfolding of DnaK that incorporates an intermediate that expands with increasing concentrations of denaturant (Palleros et al., 1993). An increasingly less compact intermediate (but still more compact than the unfolded state) could explain the linear dependence with urea concentration found in the size-exclusion chromatography experiment for the double mutant (Figure 7) that is not found in the wild-type case.

Aggregation of barnase can occur at acid pH's under certain conditions. However, at neutral pH, native gels, dependence of CD signal upon concentration, and differential scanning calorimetry show no sign of detectable aggregation (data not shown).

We have modeled the titration curves by assuming a three-state transition, simulating the curves for the loss of tertiary interactions monitored by fluorescence using eqs 5 and 7 and trial values of $m_{I\text{-}F}$ and $m_{U\text{-}F}$ and reasonable values of equilibrium constants. Using this procedure, reasonably good fits to the experimental data could be achieved (Figure 5, Table IV). Under these conditions, $\Delta G_{U\text{-}F}$ and $m_{U\text{-}F}$ as measured by fluorescence and a two-state equation approximate well to the values for the three-state system although, as expected, the former underestimates the true energy. The theoretical curves for the loss of secondary structure, monitored by $[\theta]_{222}$, are more poorly simulated due mainly to the lower

quality of the experimental data but generally they provide data for the energy of the intermediate within the range of the expected according to the kinetic data (Table IV). The equilibrium data could be fitted to a more complex model in which the intermediate is not a unique structure but unfolds further with increasing concentration of denaturant, as proposed for DnaK (Palleros et al., 1993).

The three-state model using the data in Table IV predicts that the maximum accumulation of intermediate for the I4A/I51V mutant is 43% of the total and is at 3.4 M urea. The maximum for the quadruple mutant I4A/I25V/I51V/Y78F is 23% at 2.2 M urea. Although accumulating more and more single mutations should, in theory, give rise to a greater accumulation of intermediate, this is only a small effect: if all the mutations were strictly additive and the values of m_{U-F} and m_{I-F} did not change on mutation, the intermediate would account for 56% of the total at 2.8 M in I4A/I25V/I51V/Y78F. Small effects from nonadditivity and a change in m_{I-F} cause I4A/I51V to be more suitable for further study on the accumulation of the intermediate than is I4A/I25V/I51V/Y78F. A higher value of m_{I-F} for the double mutant than for the quadruple is expected for an intermediate that gradually expands with increasing urea. In this case, m_{I-F} should increase with increasing urea concentration as the protein unfolds, and the less destabilized double mutant unfolds at a higher concentration of urea than does the quadruple.

In conclusion, we have proved the feasibility of a procedure to stabilize transient kinetic folding intermediates and to accumulate them at equilibrium, thus bringing together the kinetic and thermodynamic approaches to the protein folding problem. Direct characterization of the intermediate of barnase will be carried out in the future by other techniques, such as NMR.

ACKNOWLEDGMENT

J.M.S. is supported by EMBO.

REFERENCES

Barrick, D., & Baldwin, R. L. (1993) *Protein Sci. 2*, 869–876.
Buck, M., Radford, S. E., & Dobson, C. M. (1993) *Biochemistry 32*, 669–678.
Clarke, J., & Fersht, A. R. (1993) *Biochemistry 32*, 4322–4329.
Corbett, R. J. T., & Roche, R. S. (1984) *Biochemistry 23*, 1888–1894.
Creighton, T. E. (1986) *Methods Enzymol. 131*, 83–106.
Creighton, T. E. (1990) *Biochem. J. 270*, 1–16.
Dill, A. K. (1985) *Biochemistry 24*, 1501–1509.
Goto, Y., & Fink, A. L. (1989) *Biochemistry 28*, 945–952.
Ikeguchi, M., Sugai, S., Fujino, M., Sugawara, T., & Kuwajima, K. (1992) *Biochemistry 31*, 12695–12700.
Kellis, J. J., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–22.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.
Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87–103.
Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.
Levinthal, C. (1968) *J. Chim. Phys. 85*, 44–45.
Matouschek, A., & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., & Fersht, A. R. (1990) *Nature 346*, 440–445.
Matouschek, A., Serrano, L., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 819–835.
Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42048.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266–279.
Palleros, D. R., Shi, L., Reid, K. L., & Fink, A. (1993) *Biochemistry 32*, 4314–4321.
Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 12*, 20–24.
Sancho, J., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 741–747.
Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., & Gilmanshin, R. I. (1991) *Biopolymers 31*, 119–128.
Serrano, L., Horovitz, A., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *Biochemistry 29*, 9343–9352.
Serrano, L., Matouschek, A., & Fersht, A. R. (1992a) *J. Mol. Biol. 224*, 847–859.
Serrano, L., Kellis, J. T., Jr., Cann, P., Matouschek, A., & Fersht, A. R. (1992b) *J. Mol. Biol. 224*, 783–804.
Serrano, L., Matouschek, A., & Fersht, A. R. (1992c) *J. Mol. Biol. 224*, 805–818.
Sugawara, T., Kuwajima, K., & Sugai, S. (1991) *Biochemistry 30*, 2698–2706.