Native State EX2 and EX1 Hydrogen Exchange of *Escherichia coli* CspA, a Small β-Sheet Protein†

Hector M. Rodriguez,‡ Andrew D. Robertson,§ and Lydia M. Gregoret*,‡

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, and Department of Biochemistry, College of Medicine, The University of Iowa, Iowa City, Iowa 52242

Received June 27, 2001; Revised Manuscript Received October 23, 2001

ABSTRACT: *Escherichia coli* CspA is a small all-β-sheet protein that folds fast ($\tau = 4$ ms) via an apparent two-state mechanism. Our previous studies have shown that a large aromatic cluster on the surface of the protein participates in the rate-limiting step of folding and thus may be part of the folding nucleus of this protein. To obtain a more detailed picture of molecular events at the peptide backbone during unfolding and folding of CspA, we used native state hydrogen exchange and nuclear magnetic resonance spectroscopy (NMR). The experiments with native CspA were performed over a range of pH values from low pH, where exchange is governed by a rapid equilibrium before chemical exchange (EX2 exchange), to high pH, where exchange is dictated by the rate of unfolding (EX1 exchange). Rates of folding and unfolding were determined for 11 residues. The distribution of rates of folding within the structure of CspA suggests that hairpin turns, including one near the aromatic cluster, may nucleate the folding of CspA.

A number of key discoveries made in recent years bring us closer than ever to a full understanding of the relationship between a protein's primary sequence and its three-dimensional structure. The characterization of the folding mechanisms of a number of structurally diverse, single-domain proteins (*1*) has provided essential information about the relative importance of secondary structural elements to folding nucleation and about the relationship between topology and folding rate (*2*−*5*). Much of the work leading to these discoveries has been conducted using rapid mixing techniques coupled with optical spectroscopy. Although optical methods lack detailed structural specificity and rapid mixing methods are limited to approximately 1 ms dead times, when used in conjunction with the protein engineering method, energetic and structural information can be inferred at the residue level (*6*, *7*). These types of experiments have revealed key determinants to folding (*8*, *9*) and will continue to provide essential information about the role of side chains in folding mechanisms.

Nuclear magnetic resonance spectroscopy (NMR)[1] has the potential to provide structural information about folding processes at each residue simultaneously and on rapid time scales. A variety of methods have been applied to studying folding including pulse-labeling hydrogen exchange (*10*), magnetization transfer (*11*, *12*), line shape analysis (*13*), and native state hydrogen exchange (*14*). The latter is distinguished from the other techniques in providing access to the kinetics of unfolding and folding at specific hydrogen bonds under native solution conditions. Thus, proteins whose three-dimensional structures have been solved by NMR are ideal candidates for this approach. Here we report on the folding of *Escherichia coli* CspA using native state hydrogen exchange.

We have previously characterized the folding of CspA using stopped-flow fluorescence spectroscopy and concluded that it folds rapidly via an apparent two-state mechanism (*15*). However, the single tryptophan that accounts for the majority of the fluorescence signal of CspA is located on the surface of the protein in the nucleic acid binding site within a solvent-exposed aromatic cluster (Figure 1). It is possible that this probe reports only on the environment of this cluster during the folding process and not on the global folding of the protein. Using NMR and hydrogen exchange allows us to study the folding of multiple sites throughout the protein simultaneously, both near and away from the aromatic cluster.

Amide protons involved in backbone hydrogen bonds are proposed to exchange with the solvent by the following

† This work was supported by NIH Grant GM52885 (L.M.G.), including a minority student award supplement, and NIH Grant GM46869 (A.D.R.).

* To whom correspondence should be addressed. E-mail: gregoret@chemistry.ucsc.edu. Phone: 831-459-4002. Fax: 831-459-2935.

‡ University of California, Santa Cruz.

§ The University of Iowa.

[1] Abbreviations: $\alpha^{\ddagger}$, fractional solvent accessibility of the transition state ensemble of folding; ΔG, Gibbs free energy for protein unfolding (a subscript denotes whether values were obtained from equilibrium, kinetic, or hydrogen exchange experiments); D_2O, deuterium oxide; EX1, mechanism of hydrogen exchange in which k_{obs} equals k_{op}; EX2, mechanism of hydrogen exchange in which k_{obs} equals $k_{ch}(k_{op}/k_{cl})$; FID, free induction decay; HX, hydrogen exchange; I, observed peak intensity; I_0, initial peak intensity; k_{ch}, rate of exchange in model peptides; k_{cl}, the rate of closing or refolding; k_f, rate of folding; k_{obs}, observed exchange rate; k_{op}, rate of opening or unfolding; k_u, the rate of unfolding; m_{eq}, dependence of the free energy of unfolding on denaturant concentration; m_f, dependence of the rate of folding on denaturant concentration; m_u, dependence of the rate of unfolding on denaturant concentration; NMR, nuclear magnetic resonance; R, universal gas constant; T, temperature; T_1, longitudinal relaxation time constant; T_m, midpoint of thermal denaturation; TSP, sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$.

10.1021/bi011347x CCC: $22.00
Published on Web 01/19/2002

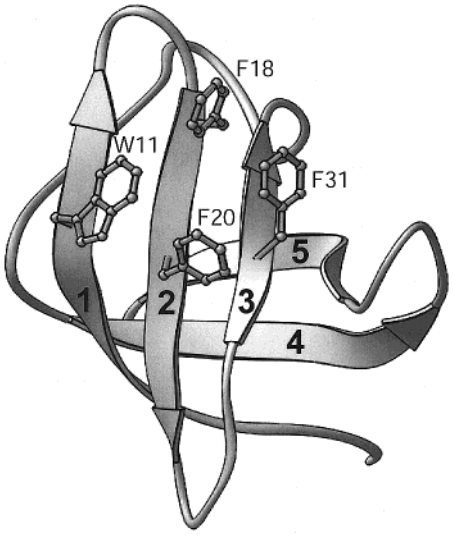

FIGURE 1: Ribbon diagram of *E. coli* CspA showing the aromatic cluster with the phenylalanine residues highlighted. This image was rendered with the Ribbons software (*48*) using the coordinates of Schindelin et al. (*49*).

scheme (*17*, *18*), where k_{op} is the rate of opening, k_{cl} is the rate of closing, and k_{ch} is the intrinsic exchange rate of the proton:

$$\text{closed} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{open} \xrightarrow{k_{ch}} \text{exchanged} \quad (1)$$

Under native conditions, there is an equilibrium preceding chemical exchange, with the open conformation being susceptible to exchange (*19*, *20*). This behavior is described by eq 2, where k_{obs} is the observed rate of exchange:

$$k_{obs} = \frac{k_{op}k_{ch}}{k_{cl} + k_{ch}} \quad (2)$$

At low pH, the intrinsic exchange rate is slow ($k_{cl} \gg k_{ch}$), leading to EX2 exchange. Equation 2 simplifies to

$$k_{obs} = \frac{k_{op}}{k_{cl}}k_{ch} \quad (3)$$

At pH >4, the intrinsic rates of exchange, k_{ch}, increase 10-fold with each increase in pH and are sensitive to the neighboring residues. Eventually, at high pH, $k_{ch} \gg k_{cl}$ (EX1 exchange), and the observed rate is equal to the rate of opening (unfolding):

$$k_{obs} = k_{op} \quad (4)$$

Thus, the combination of EX2 and EX1 exchange at the slowest exchanging amide protons should allow the estimation of the rates of folding and unfolding (*21*, *22*). Arrington and Robertson demonstrated the utility of this method in characterizing the folding of the turkey ovomucoid third domain (*14*), and more recent studies of ubiquitin have demonstrated the accuracy of this approach (*23*).

Hydrogen exchange experiments are typically carried out by resuspending lyophilized protein in D_2O and monitoring the replacement of protons by deuterons as a function of time. This approach is not practical with CspA because it is marginally stable, with an unfolding free energy of only 3 kcal/mol. Consequently, all amide protons exchange fully within 2 h of dissolving CspA in D_2O at pH 6.0 (*24*). Moreover, CspA unfolds and refolds rapidly, and at higher pH, we thus expect the amide protons to exchange too rapidly to measure by hydrogen−deuterium exchange. We have thus used saturation transfer to follow slow amide exchange in native CspA at alkaline pH (*25*, *26*).

The transfer of saturated (magnetized) spins from water to an amide proton is observed as a decrease in peak intensity, I, which depends on the exchange rate of the amide proton (k_{obs}; eq 2) and its longitudinal relaxation time constant T_1 according to

$$f = \frac{I}{I_0} = \frac{1}{1 + k_{obs}T_1} \quad (5)$$

These fractional intensities, f, will vary for each amide proton resonance as a function of pH. By measuring the peak intensities over the pH range encompassing EX2 and EX1 conditions, estimates of the closing and opening rates for individual amide protons can be obtained by substituting eq 2 into eq 5. This only requires that the intrinsic rates of exchange, k_{ch}, are known. The dependence of the rate of chemical exchange on pH is described by eq 6, where k_H and k_{OH} are the rates of acid and base catalysis and k_{H_2O} is the rate of water catalysis:

$$k_{ch} = k_H[H^+] + k_{OH}[OH^-] + k_{H_2O} \quad (6)$$

Values of k_{ch} have been determined for model peptides as a function of pH (*27*, *28*) and are obtained from this literature.

We are able to follow the exchange of 11 amide protons in CspA using this technique. These amide protons are well distributed throughout four of the five β-strands of CspA and now provide us with a more comprehensive picture of the folding of this protein. As we hypothesized earlier (*15*, *16*, *29*), the β-strands that include the aromatic cluster residues and the RNA binding RNP1 and RNP2 motifs are the first to come together in the folding of CspA. A turn at the end of the loop between strands 3 and 4 forms early as well.

MATERIALS AND METHODS

Analytical grade urea was purchased from ICN. Buffer components and other reagents were of analytical grade and were purchased from Fisher Scientific. All isotopic chemicals were purchased from Cambridge Isotope Laboratories. CspA was expressed and purified as described previously (*15*), and the concentration of CspA was determined by absorbance measurements at 280 nm using the calculated extinction coefficient 8437 M^{-1} cm^{-1} (*30*). pH measurements were made using an Orion Research model 611 pH meter equipped with a 3 mm Ingold electrode or a Corning pH meter, model 430, with a Corning pH combination electrode. The pH meters were calibrated with two standards (VWR Scientific).

Protein samples to be used for NMR analysis were lyophilized to constant weight. To ensure that CspA was properly folded after lyophilization, one of the lyophilized samples was resuspended in buffer and analyzed by circular dichroism spectroscopy.

Urea-Induced Denaturation Monitored by Fluorescence. Urea titrations were performed using a Perkin-Elmer LS-50B fluorescence spectrometer as described previously (*29*). The stock urea solutions were prepared daily, and the concentrations were calculated by their index of refraction using a Bausch & Lomb refractometer (*31*).

Thermal Denaturation. Thermal denaturation of CspA was followed using circular dichroism spectroscopy by measuring the change in ellipticity at 222 nm with increasing temperature. A protein concentration of ~50 μM in 50 mM potassium phosphate and 100 mM potassium chloride was used, and the pH was adjusted with potassium hydroxide to give the desired final pH. The temperature was increased at a rate of 0.33 °C/min using a programmable water bath, and the experimental data were analyzed as described previously to obtain the midpoint of thermal denaturation (T_m) (*15*, *32*).

Stopped-Flow Fluorescence. All folding and unfolding reactions were performed using an Applied Photophysics SX.18MV stopped-flow spectrometer. The temperature was controlled to ±0.5° of the desired temperature with a circulating water bath. Unfolding was initiated by rapid dilution of the native protein in 50 mM potassium phosphate and 100 mM KCl at pH 7.0 or 11.0 with the varying concentrations of buffered urea at the same pH. The refolding reactions were performed by rapidly diluting CspA in 7 M urea with buffer to which urea was added to give the final desired concentration of urea. Excitation was at 280 nm, and the emission was measured through a 305 nm cutoff filter. Final protein concentrations were between 2 and 7 μM, and at least 10 transients were collected and averaged for each urea concentration. The average was fit to a monoexponential time course using Kaleidagraph (*33*). The rate constants for refolding (k_f) and unfolding (k_u) in the absence of denaturant and the dependence of these rates on the concentration of denaturant (m_f and m_u) were determined as described previously (*15*).

Nuclear Magnetic Resonance Spectroscopy. NMR samples were prepared by resuspending lyophilized CspA with buffer [50 mM potassium phosphate, 100 mM potassium chloride, 10% D_2O, and ~1 mM TSP [sodium 3-(trimethylsilyl)-propionate-*2,2,3,3-d*$_4$]] preadjusted to the desired experimental pH with potassium hydroxide (the pH ranged from 6 to 11). The pH of the samples was measured prior to each experiment, and the samples were filtered through a 0.2 μm Millipore syringe filter. Final sample volume was 600 μL at a concentration of ~1 mM CspA. Amide proton assignments are based on published assignments of CspA (BMRB 4296) (*24*, *36*). Amide proton chemical shifts were referenced to TSP (0 ppm), and peak intensities were determined using the VNMR software (Varian).

All NMR experiments were performed on a 500 MHz Varian INOVA spectrometer using a triple resonance probe (located in the University of Iowa College of Medicine NMR facility) or a 500 MHz Varian UNITY using a triple resonance probe (Varian) (at the University of California, Santa Cruz). The spectrometers' temperature controllers were calibrated to 38 ± 0.5 °C using a methanol standard (*34*).

Hydrogen exchange experiments were performed by acquiring 1D ^{1}H spectra of native CspA at varying pH using the PRESAT pulse sequence included with the VNMR software. All parameters were identical for each sample, and the peak heights were normalized to a nonexchanging aliphatic signal. Each FID was the sum of 256 transients consisting of 8000 complex data points. The spectral width was 6000 Hz, and the recycle time was 4.4 s. Total acquisition time for each FID was 22 min.

The identities of the most slowly exchanging amide protons determined by the hydrogen exchange experiments (I8, K10, G19, I21, T22, V30, V32, Q49, V51, S52, and F53) were verified by performing a TOCSY experiment at pH 11 and 38 °C, the most extreme conditions used in this study. TOCSY spectra consisted of 128 blocks of 12 summed transients, each of which consisted of 2048 complex data points with a spectral width of 6000 Hz. Total acquisition time for each spectrum was 2 h.

Longitudinal relaxation time constants (T_1) were determined for the slowly exchanging amide protons using an inversion recovery pulse sequence (*35*) with presaturation of the solvent, and analysis was conducted using the VNMR software. Each FID was the sum of 16 transients collected at 8000 complex data points, a spectral width of 6000 Hz, and a recycle time of 6 s. The pH range and temperature of the samples were identical to that used in the hydrogen exchange experiments.

Data Analysis. The rates of closing (k_{cl}) and opening (k_{op}) were determined by simultaneously fitting the data in Figure 5 (fractional intensities versus pH) to eqs 2, 5, and 6 with Kaleidagraph (*33*). Values for the rate constants in eq 6 were obtained from published values (*27*, *28*). The T_1 values for the slowly exchanging protons were found to be invariant with pH and protein sequence. As such, an average T_1 value for all amide protons was used in eq 5 ($\langle T_1 \rangle = 0.7$ s). Proton and hydroxide concentrations in eq 6 were computed from the pH of the samples. Substitution of eqs 6 and 2 into eq 5 with all known variables results in a new equation with two unknown values, k_{cl} and k_{op}. Confidence intervals were computed using the standard errors of the curve fitting.

RESULTS

Our intent was to monitor hydrogen exchange rates as a function of pH by NMR and observe EX2 exchange and, at high pH, the switch to EX1. Once in the EX1 limit, a direct measurement of the opening rate (k_{op}) can be obtained (*14*, *18*, *21*−*23*). To accomplish this, we acquired 1D ^{1}H NMR spectra at seven pH values between 6 and 11 at a temperature of 38 °C. This is the lowest temperature at which we observed saturation transfer. Simulations of the exchange behavior using the global folding and unfolding rates determined by stopped-flow fluorescence spectroscopy (*15*) suggest that CspA should be in EX1 exchange at pH >9.

Folding Kinetics and Stability of CspA at Elevated Temperature and pH. To ensure that CspA was structured and that its folding pathway had not been altered at alkaline pH, the stability and folding kinetics were characterized at pH 7.0 (38 °C) and pH 11.0 (25 and 38 °C). The thermal stability of CspA at pH 11 was determined by monitoring the temperature dependence of the change in ellipticity at 222 nm (Figure 2). The midpoint of thermal denaturation (T_m) under these conditions was found to be 53 °C, which is 5° less than that of CspA at pH 7.0 (T_m = 58 °C). A single transition was observed in both cases (pH 7 and 11). At both pH values, the signal at 222 nm was completely recovered on return to 5 °C, indicating that CspA refolds reversibly

Table 1: Thermodynamic and Kinetic Parameters of Wild-Type CspA Determined by Fluorescence Spectroscopy[a]

	ΔG_{eq} (kcal mol^{-1})	ΔG_{kin} (kcal mol^{-1})	k_f (s^{-1})	m_f (M^{-1})	k_u (s^{-1})	m_u (M^{-1})	m_{eq} (kcal mol^{-1} M^{-1})	m_{kin} (kcal mol^{-1} M^{-1})	$\alpha^{\ddagger}$
pH 7.0, 25 °C	3.0 ± 0.2	2.9 ± 0.1	266 ± 16	−0.87 ± 0.02	1.9 ± 0.2	0.14 ± 0.07	0.71 ± 0.04	0.60 ± 0.07	0.86 ± 0.07
pH 7.0, 38 °C	2.4 ± 0.1	2.0 ± 0.1	565 ± 12	−1.01 ± 0.03	23.7 ± 2.9	0.14 ± 0.02	0.67 ± 0.01	0.71 ± 0.03	0.88 ± 0.04
pH 11.0, 25 °C	2.6 ± 0.2	2.9 ± 0.1	201 ± 3	−0.93 ± 0.02	1.5 ± 0.4	0.23 ± 0.07	0.70 ± 0.01	0.69 ± 0.06	0.80 ± 0.05
pH 11.0, 38 °C	1.5 ± 0.2	1.6 ± 0.1	250 ± 11	−1.02 ± 0.06	18.3 ± 2.4	0.14 ± 0.02	0.96 ± 0.20	0.72 ± 0.05	0.88 ± 0.07

[a] The reported errors of k_f, k_u, m_f, and m_u are the standard errors of the curve fits. Abbreviations: ΔG_{eq}, unfolding free energy obtained from an equilibrium experiment; ΔG_{kin}, unfolding free energy obtained using kinetically determined unfolding and refolding rates; k_f, refolding rate; m_f, slope of refolding rate dependence on denaturant concentration; k_u, unfolding rate; m_u, slope of unfolding rate dependence on denaturant concentration; m_{eq}, slope of unfolding free energy dependence on denaturant concentration; $m_{kin} = RT(m_u - m_f)$; $\alpha^{\ddagger}$, fractional solvent accessibility of the transition state ensemble of folding [$\alpha^{\ddagger} = -m_f/(m_u - m_f)$].

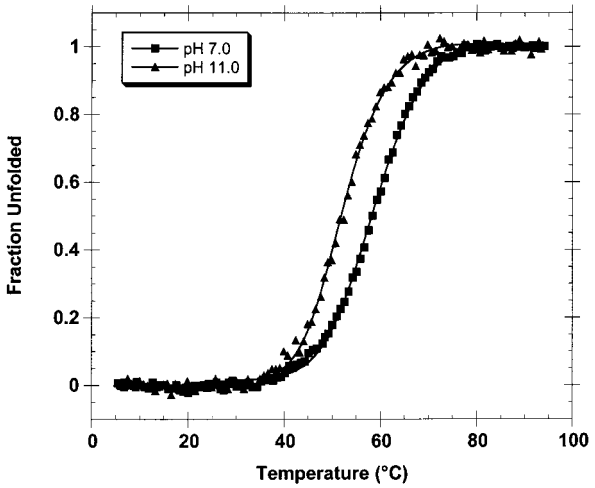

FIGURE 2: Thermal stability of CspA expressed as the fraction unfolded as a function of temperature. CspA, pH 7.0, is shown as squares and CspA, pH 11.0, is shown in triangles. The solid lines represent the fit to the data. The concentration of CspA was 50 μM in 50 mM potassium phosphate and 100 mM KCl at the appropriate pH.

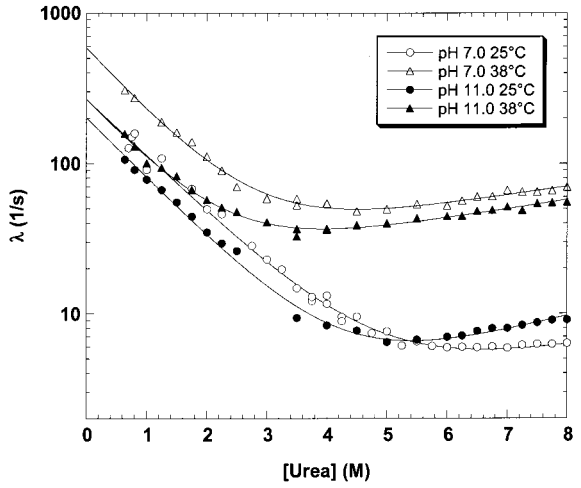

FIGURE 3: Apparent rate constants (λ) of CspA plotted as a function of urea concentration. Rates were determined by stopped-flow fluorescence spectroscopy at varying pH and temperature. Emission intensities were measured through a 305 nm cutoff filter.

following thermal denaturation (not shown). Moreover, the CD scans of the samples before thermal denaturation and after returning to 5 °C were identical.

Stability was also assessed by chemical denaturation. The fluorescence emission intensity was monitored as a function of urea at pH 7.0 (38 °C) and pH 11.0 (25 and 38 °C). A single transition between the folded and unfolded states was observed under all conditions. The unfolding free energy at pH 11.0 and 38 °C, the most extreme conditions under which hydrogen exchange was performed, is 1.0 kcal/mol lower than at pH 7 and 38 °C (Table 1). Nonetheless, CspA undergoes cooperative unfolding at the higher pH, and the circular dichroism scans at both pH values are identical. This indicates that CspA is still structured under these conditions. Additionally, Petrosian and Makhatadze have shown that the stability of CspA is independent of pH in the pH range 6.5−8.5 (*37*).

At pH 7.0 and 38 °C the folding rate (k_f) is 564 s^{-1} (Table 1 and Figure 3). At pH 11.0 and 38 °C, the refolding rate is 250 s^{-1}. The unfolding rates (k_u) for both pH 7 and pH 11 at 38 °C are virtually identical, though both increase by an order of magnitude relative to rates at 25 °C. The unfolding rates still exhibit the very slight dependence on denaturant, which is characteristic of the cold shock proteins (*15*, *38*, *39*). This slight dependence implies that the transition and the folded states have similar interactions with the solvent.

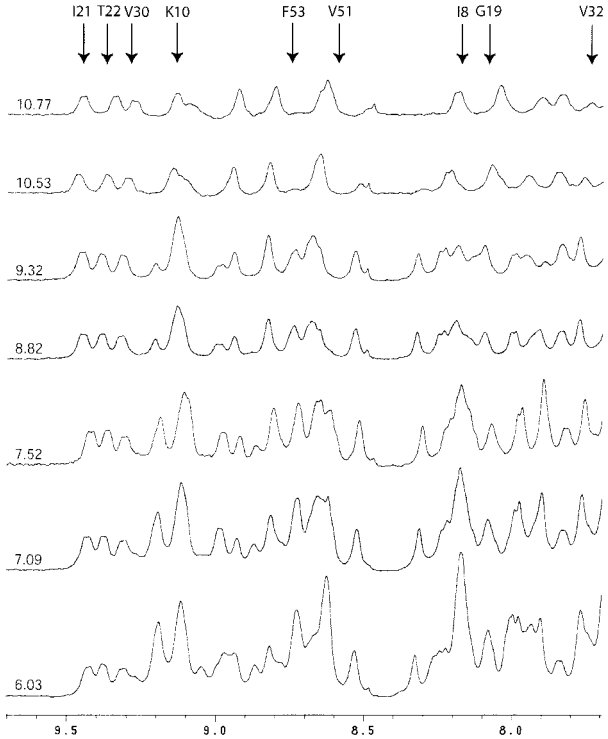

FIGURE 4: NMR spectra of the amide region of CspA shown at varying pH with temperature held constant at 38 °C. All spectra are to scale and referenced to TSP (0 ppm).

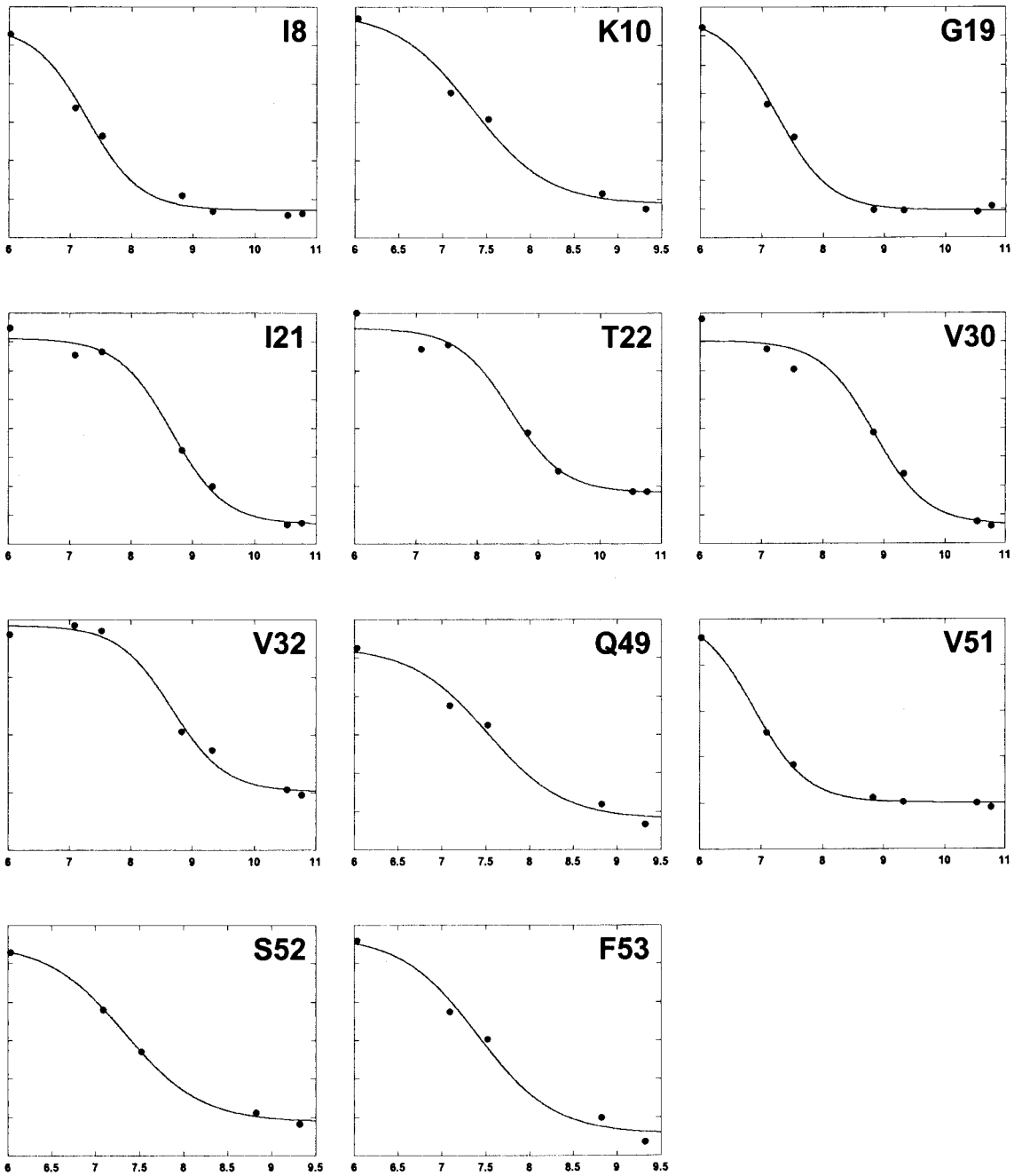

FIGURE 5: Peak intensities of all slowly exchanging amide protons plotted as a function of pH. The solid lines indicate the fits to the data (see Materials and Methods).

In other words, the transition state ensembles appear to be quite compact under all conditions examined.

The equilibrium unfolding and kinetic data are consistent with a two-state mechanism at 38 °C and pH 11. In all of the conditions studied, there is a good correspondence between the Gibbs free energy obtained from urea titrations and the folding kinetics. The folding and unfolding kinetics can be modeled well by a monoexponential time course, also supporting a two-state mechanism. Hence, a two-state hydrogen exchange mechanism is a reasonable model for our studies.

Hydrogen Exchange via Saturation Transfer. Hydrogen exchange experiments were conducted by acquiring 1D NMR spectra of native CspA in water at varying pH. All spectra were acquired under identical conditions where saturation of the solvent (water) was achieved by preirradiation. Though initially complex, the spectrum of CspA simplifies as the pH increases (Figure 4). This results from an increase in the intrinsic exchange rate of the amide protons. The amide protons still visible at high pH are those that exchange most slowly and are a subset of those observed previously to be the most slowly exchanging (I8, K10, G19, I21, T22, V30, V32, Q49, V51, S52, and F53) (*24*, *40*). The identity of these protons was verified by performing a TOCSY experiment at pH 11 and 38 °C using published assignments (*36*). These amide protons are located on strands 1, 2, 3, and 4 of CspA (Figure 1) and provide independent probes of the structure in these regions of the protein. As has been reported previously (*36*, *40*), there are no slowly exchanging amide protons on strand 5 of CspA.

Hydrogen Exchange Rates. Signal intensities for the slowly exchanging amide protons vary significantly with pH. In a

Table 2: Thermodynamic and Kinetic Parameters of CspA Determined by NMR[a]

residue	k_{cl} (s^{-1})	k_{op} (s^{-1})	ΔG_{HX} (kcal mol^{-1})
Ile 8	206 ± 65	9.7 ± 2.1	1.9 ± 0.2
Lys 10	493 ± 302	8.2 ± 3.6	2.5 ± 0.5
Gly 19	489 ± 78	3.3 ± 0.2	3.1 ± 0.1
Ile 21	1172 ± 324	2.0 ± 0.2	3.9 ± 0.2
Thr 22	2235 ± 993	2.2 ± 0.4	4.3 ± 0.3
Val 30	1251 ± 612	2.5 ± 0.5	3.8 ± 0.3
Val 32	1323 ± 451	2.0 ± 0.2	4.0 ± 0.2
Gln 49	995 ± 579	3.6 ± 1.1	3.5 ± 0.4
Val 51	52 ± 6	5.8 ± 0.4	1.4 ± 0.1
Ser 52	488 ± 70	2.3 ± 0.1	3.3 ± 0.1
Phe 53	646 ± 370	4.7 ± 1.5	3.0 ± 0.4

[a] The reported errors are at the 70% confidence interval. Experimental conditions were 38 °C using identical NMR parameters, with CspA samples in 50 mM potassium phosphate, 100 mM potassium chloride, 10% D_2O, and 1 mM TSP at varying pH (see Materials and Methods).

purely EX2 exchange, the signal intensity is expected to decrease with increasing pH with the eventual disappearance of the signal intensity at higher pH. In the EX1 limit, peak intensities are expected to plateau to nonzero values when the rate of exchange approaches the rate of opening. Figure 5 shows the dependence of signal intensity on pH for each of the observed protons. In most cases, the signal intensity decreases with pH and then begins to plateau at pH >7. A plateau in signal intensity can be explained into two ways: EX1 exchange is occurring and k_{obs} is approaching k_{op} or regions within CspA are becoming more stable with increasing pH. The second explanation is unlikely since the stability of CspA begins to decrease at pH values above 11 (H. Rodriguez, unpublished results). Thus, the plateau almost certainly results from a switch from EX2 to EX1 exchange.

The individual closing and opening rates of the slowly exchanging protons were determined by simultaneously fitting the data in Figure 5 to eqs 2, 5, and 6 (see Materials and Methods) and are listed in Table 2. The closing rates vary by 2 orders of magnitude, ranging from 52 to 2200 s^{-1}. The values of k_{op} vary less than an order of magnitude and range from 2 to 10 s^{-1}. The solid lines in Figure 5 denote the curve fits to the data and, in most cases, describe the data well. Equation 5 assumes that the opening and closing rates remain constant as a function of pH. Since, by stopped flow, we found a 2-fold difference in the refolding rates at pH 7 vs 11 (though no difference in the unfolding rates; Table 1), there may be some error in the absolute k_{cl} and k_{op} values. However, the pH dependence of the refolding rates is unlikely to explain the 10−40-fold differences seen in the individual closing rates (k_{cl}).

The free energy differences between the open and closed conformations can be computed from $\Delta G_{HX} = -RT \ln(k_{op}/k_{cl})$. The computed ΔG_{HX} for the observed protons range from 1.4 to 4.3 kcal/mol, though 9 of 11 values are between 1.9 and 4.0 kcal/mol. These values are comparable to those of Jaravine et al. (*40*), who found values in the range 2.6−3.9 kcal/mol (with one outlier at 1.2 kcal/mol), although the different solution conditions used in their study preclude a more precise comparison. Again, as observed and noted by Jaravine et al. (*40*), all of the amide protons on strand 5 exchange too rapidly to observe, even though many of the amides on strand 4 to which strand 5 is hydrogen bonded in an antiparallel interaction are slowly exchanging. This

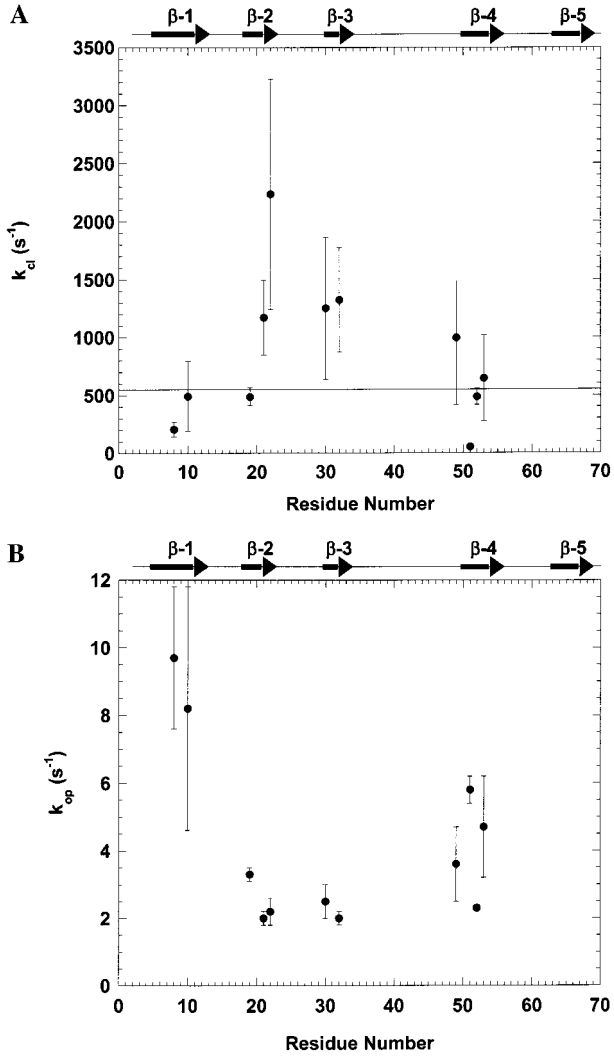

FIGURE 6: Measured rates of (A) folding and (B) unfolding plotted for all observed amide protons. The solid horizontal line in (A) represents the extrapolated refolding rate observed by stopped-flow experiments at pH 7.0 (38 °C).

puzzling observation is difficult to explain but similar to results observed in the β-sheets of ovomucoid third domain and ubiquitin (*14*, *23*).

DISCUSSION

We have studied the folding of CspA by native state hydrogen exchange. We were able to follow 11 amide protons on strands 1, 2, 3, and 4 of CspA, giving us a more comprehensive picture than before of the folding of this model protein. Although there is some variation particularly in the closing rates, we find that the opening and closing rates are similar to the global folding and unfolding rates measured by stopped-flow fluorescence spectroscopy.

Unfolding of CspA. The unfolding rate constants measured by hydrogen exchange (Table 2) are slower, by factors of 2−10, than the unfolding rate constants obtained by stopped-flow fluorescence spectroscopy (Table 1). The possible sources of experimental error in determining k_{op} values are the T_1 values and the intensity measurements, particularly at the low pH extrema which is the denominator in determining fractional intensity. The average T_1 value for amide protons in CspA, 0.7 s, is similar to T_1 values observed

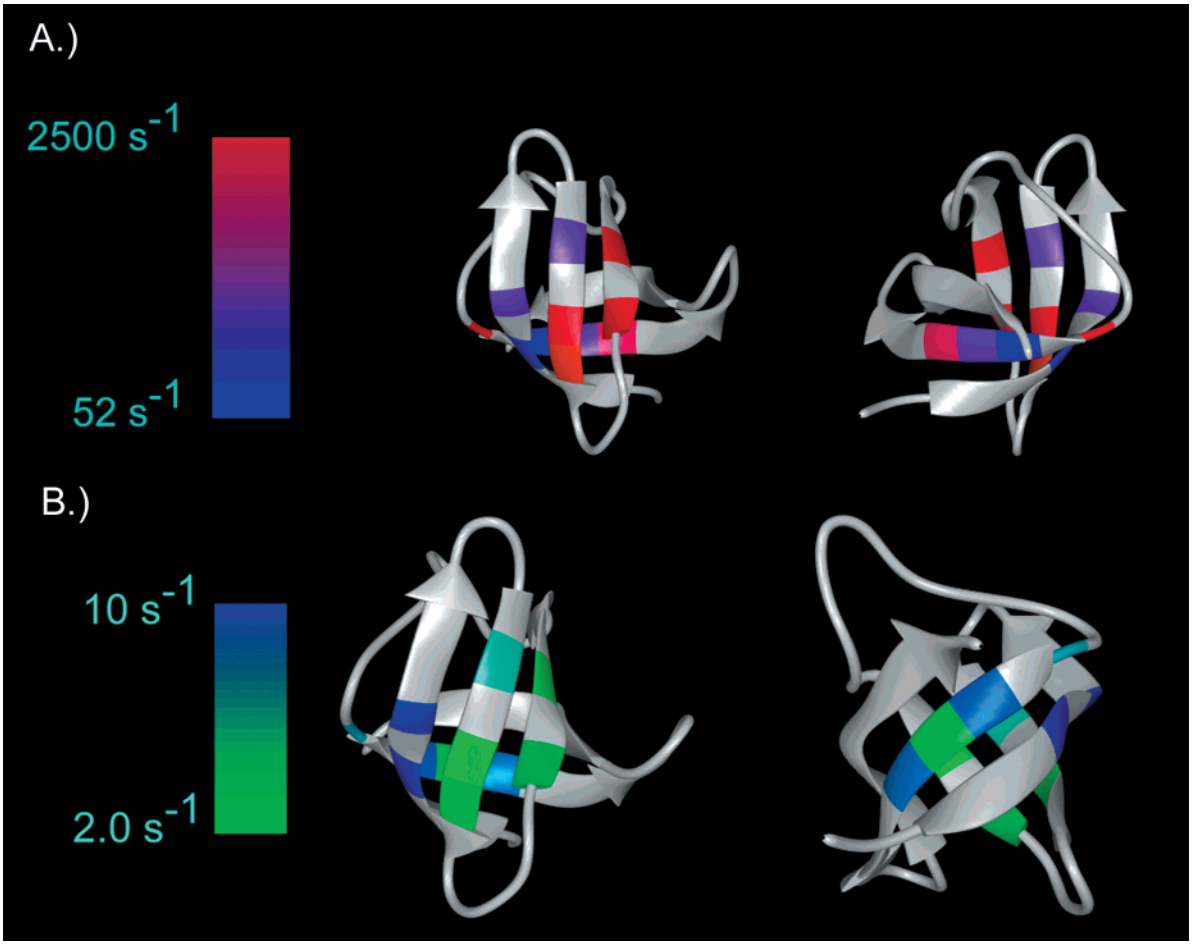

FIGURE 7: Ribbon diagram of CspA with the rates of (A) closing and (B) opening of individual amide protons mapped onto the structure. This figure was rendered with UCSF MidasPlus (*50*).

in ovomucoid third domain and ubiquitin (A. D. Robertson, unpublished results), two proteins similar in size to CspA. Assuming that the intensity measurements are accurate, T_1 would have to decrease by a factor of 10 in order to arrive at k_{op} values of about 20 s^{-1}, and an error of this magnitude is highly unlikely. A similar argument applies to the intensity measurements. In fact, at low pH some of the intensities may be overestimated because of possible overlap with other resonances. This type of error will only lead to overestimates in k_{op}. Overall, k_{op} values are indeed significantly less than k_u values.

Interestingly, k_{op} values of about 10 s^{-1} are found at I8 and K10, very close to W11, the fluorophore in stopped-flow experiments yielding k_u values of about 20 s^{-1}. The modest 2-fold difference in rate constants may reflect the difference in the experimental methods (*41*). Moreover, while the uncertainties in most k_{op} values are significant, a pattern in the structural distribution of k_{op} values suggests real differences in the kinetics of unfolding for different regions of native CspA (Figure 6). The slowest opening rate constants are located on strands 2 and 3 where most of the aromatic cluster residues are located (Figures 6 and 7). More rapid opening occurs at the more peripheral strands 1 and 4.

Folding of CspA. A similar pattern is observed in the structural distribution of k_{cl} values. Stopped-flow experiments at pH 7 and 11 and at 38 °C yield rate constants for folding ranging from about 300 to 600 s^{-1}, and given the uncertainties in k_{cl} determinations, most of the k_{cl} values measured by exchange fall within this range or are very close to it. Nevertheless, the distribution in k_{cl} values is nonrandom with respect to structure, with the larger values located on strands 2 and 3. Assuming that only nativelike hydrogen bonds form, this suggests that a possible scenario for the folding of CspA could be one in which strands 2 and 3 form a two-stranded sheet that then serves as a template for the folding of the remainder of the protein. Only one result does not fit with this scenario: the residue with the fastest closing rate of all, Thr 22, is on strand 2 but is hydrogen bonded to Ile 8 on strand 1. The closing rate of Ile 8 is 10-fold slower. Given the similarity of the closing rates of residues 8 and 10, both on strand 1, we suspect this discrepancy may be due to an inaccuracy in the k_{cl} of Thr 22, whose intensity which is just beginning to plateau at pH >10, or the result of intraresidue hydrogen bonding between the side chain hydroxyl group acting as the acceptor at high pH (the pK_a of the −OH of free threonine is 13) and the amide proton of the same residue acting as the donor.

Residues 8 and 10 on strand 1 have closing rates of 200−500 s^{-1}, very similar to the rate measured by stopped-flow fluorescence spectroscopy based on the fluorescence of Trp 11, the sole tryptophan of CspA. This suggests that Trp 11 is indeed reporting accurately on the folding of the protein but that it misses a slightly more rapid event—the faster closing of residues on strands 2 and 3. This event, which would occur within the dead time of mixing in the stopped-flow instrument, would not appear as missing amplitude in

the stopped-flow fluorescence traces because the probe does not lie on either of these strands. This result is also consistent with our previous findings that aromatic residues F18, F20, and F31 on these strands are important for stability as well as folding (*16*, *29*). (It should be kept in mind, however, that other parts of CspA may also be involved in nucleating folding, including those parts for which we have no information.)

A requirement for the preassociation of strands 2 and 3 may also explain the observation from stopped-flow experiments that the transition state ensemble of CspA is nativelike in its accessibility to solvent (*15*): If the hydrogen-bonding pattern and extensive hydrophobic contacts between strands 2 and 3 are established *prior* to the step observed by stopped-flow fluorescence (the docking of strand 1 with Trp 11 against strands 2−3), and if strands 4 and (quite possibly) 5 join the structure at approximately the same time or just before strand 1, then the structure of the CspA "transition state" ensemble may indeed appear to be quite nativelike.

Implications for Protein Structure Prediction. It has been proposed previously that turns may nucleate folding of β-sheets (*42*−*45*). Results of extensive ϕ-value analysis on the all-β SH3 domains, the α and β IgG binding domain, and protein L indicate that hairpins are formed in the transition state of folding and are required for proper folding (*4*, *5*, *45*−*47*). In the case of CspA, the antiparallel interaction of strands 2 and 3 may serve as a template to folding. Taken together, these results suggest that turn formation may be a general nucleating event in the folding of β-sheet proteins. Therefore, better β-sheet structure prediction may be achieved through improved prediction of turn signals and regions that form hairpins. Such regions may also be involved in other interactions that lead to the propagation of β-sheets. Progress toward accurate prediction of turn and hairpin formation could therefore lead to a better understanding of the sequence determinants of pathological protein aggregation.

ACKNOWLEDGMENT

We thank Kevin Plaxco for use of his laboratory's stopped-flow fluorometer and Jim Loo and John Diener for NMR assistance in the initial stages of this study. H. Rodriguez thanks A. Robertson for invaluable assistance and guidance in this study. NMR instrumentation at UCSC was supported in part by the Elsa U. Pardee Foundation and a grant from the W. M. Keck Foundation. Acquisition of equipment in the University of Iowa College of Medicine NMR Facility has been supported by the University of Iowa College of Medicine, the Howard Hughes Medical Institute, and the NIH.

REFERENCES

1. Jackson, S. E. (1998) *Folding Des. 3*, R81−R91.
2. Plaxco, K., Simmons, K., and Baker, D. (1998) *J. Mol. Biol. 277*, 985−994.
3. Plaxco, K. W., Simons, K. T., Ruczinski, I., and Baker, D. (2000) *Biochemistry 39*, 11177−11183.
4. Martinez, J. C., and Serrano, L. (1999) *Nat. Struct. Biol. 6*, 1010−1016.
5. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I., and Baker, D. (1999) *Nat. Struct. Biol. 6*, 1016−1024.
6. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) *J. Mol. Biol. 224*, 771−782.
7. Fersht, A. R. (1995) *Curr. Opin. Struct. Biol. 5*, 79−84.
8. Brockwell, D. J., Smith, D. A., and Radford, S. E. (2001) *Curr. Opin. Struct. Biol. 10*, 16−25.
9. Grantcharova, V., Aim, E. J., Baker, D., and Horwich, A. L. (2001) *Curr. Opin. Struct. Biol. 11*, 70−82.
10. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature 335*, 694−699.
11. Dobson, C. M., Evans, P. A., and Williamson, K. L. (1984) *FEBS Lett. 168*, 331−334.
12. Roder, H., and Wüthrich, K. (1986) *Proteins 1*, 34−42.
13. Huang, G. S., and Oas, T. G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6878−6882.
14. Arrington, C. B., and Robertson, A. D. (1997) *Biochemistry 36*, 8686−8691.
15. Reid, K. L., Rodriguez, H. M., Hillier, B. J., and Gregoret, L. M. (1998) *Protein Sci. 7*, 470−479.
16. Rodriguez, H. M., Vu, D. M., and Gregoret, L. M. (2000) *Protein Sci. 9*, 1993−2000.
17. Linderstrøm-Lang, K. (1955) *Chem. Soc. (London) Spec. Publ. 2*, 1−20.
18. Hvidt, A. A., and Neilsen, S. O. (1966) *Adv. Protein Chem. 21*, 287−386.
19. Woodward, C. K., and Hilton, B. D. (1980) *Biophys J. 32*, 561−575.
20. Woodward, C. K., Simon, I., and Tüchsen, E. (1982) *Mol. Cell. Biochem. 48*.
21. Roder, H., Wagner, G., and Wüthrich, K. (1985) *Biochemistry 24*, 7396−7407.
22. Pedersen, T. G., Thomsen, N. K., Anderson, K. V., Madsen, J. C., and Poulsen, F. M. (1993) *J. Mol. Biol. 230*, 651−660.
23. Sivaraman, T., Arrington, C. B., and Robertson, A. D. (2001) *Nat. Struct. Biol. 8*, 331−333.
24. Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S. D., Inouye, M., and Montelione, G. T. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5114−5118.
25. Forsen, S., and Hoffman, R. A. (1964) *J. Chem. Phys. 40*, 1189−1196.
26. Led, J. J., Gesmar, H., and Abildgard, F. (1989) *Methods Enzymol. 176*, 311−329.
27. Molday, R. S., Englander, S. W., and Kallen, R. G. (1972) *Biochemistry 11*, 150−158.
28. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins 17*, 75−86.
29. Hillier, B. J., Rodriguez, H. M., and Gregoret, L. M. (1998) *Folding Des. 3*, 87−93.
30. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−23.
31. Warren, J. R., and Gordon, J. A. (1966) *J. Phys. Chem. 67*, 1524−1527.
32. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8062−8068.
33. Kaleidagraph (1994) Synergy Software, Reading, PA.
34. Van Geet, A. L. (1968) *Anal. Chem. 40*, 2227−2229.
35. Hahn, E. L. (1949) *Phys. Rev. 76*, 145−146.
36. Feng, W., Tejero, R., Zimmerman, D. E., Inouye, M., and Montelione, G. T. (1998) *Biochemistry 37*, 10881−10896.
37. Petrosian, S. A., and Makhatadze, G. I. (2000) *Protein Sci. 9*, 387−394.
38. Perl, D., Welker, C., Schindler, T., Schroeder, K., Maraheil, M. A., Jaenicke, R., and Schmid, F. X. (1998) *Nat. Struct. Biol. 5*, 229−235.
39. Schindler, T., Graumann, P. L., Perl, D., Ma, S., Schmid, F. X., and Marahiel, M. A. (1999) *J. Biol. Chem. 274*, 3407−3413.
40. Jaravine, V. A., Rathgeb-Szabo, K., and Alexandrescu, A. T. (2000) *Protein Sci. 9*, 290−301.
41. Kiefhaber, T., and Baldwin, R. L. (1996) *Biophys. Chem. 59*, 351−356.
42. Searle, M., Williams, D., and Packman, L. (1995) *Nat. Struct. Biol. 2*, 999−1006.
43. Muñoz, V., Thompson, P. A., Hofrichter, J., and Eaton, W. A. (1997) *Nature 390*, 196−199.
44. Muñoz, V., Henry, E. R., Hofrichter, J., and Eaton, W. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 5872−5879.

45. McCallister, E. L., Alm, E., and Baker, D. (2000) *Nat. Struct. Biol. 7*, 669−673.
46. Grantcharova, V. P., Riddle, D. S., Santiago, J. V., and Baker, D. B. (1998) *Nat. Struct. Biol. 5*, 714−720.
47. Schönbrunner, N., Koller, K. P., and Kiefhaber, T. (1997) *J. Mol. Biol. 268*, 526−538.
48. Carson, M. (1991) *J. Appl. Crystallogr. 24*, 958.
49. Schindelin, H., Jiang, W., Inouye, M., and Heinemann, U. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5119−5123.
50. Jarvis, L., Huang, C., Ferrin, T., and Langridge, R. (1988) *J. Mol. Graphics 6*, 2−27.

BI011347X