Extensive Misfolding in the Refolding Reaction of Alkaline Ferrocytochrome *c*†

D. Krishna Rao, N. Prakash Prabhu, and Abani K. Bhuyan*

School of Chemistry, University of Hyderabad, Hyderabad 500046, India

Received January 23, 2006; Revised Manuscript Received May 4, 2006

ABSTRACT: This work describes an extensively misfolded kinetic intermediate in the folding of horse ferrocytochrome *c*. Under absolute native conditions, the alkali-unfolded protein liganded with carbon-monoxide exhibits misfolding. The misfolded product, apparently an off-pathway intermediate, requires large-scale unfolding in order to have a chance to fold correctly to the native state. The rate of unfolding of the misfolded intermediate limits the overall rate of protein folding. The high level of observed misfolding possibly results from a failure of the polypeptide chain to achieve by stochastic search the transition state relevant for successful folding. Such misfolding may be analogous to the failure of a sizable set of proteins in the intracellular milieu to fold to the functionally active native state.

As explicable by the principles of chemical kinetics, the polypeptide chain has to pass through a relatively high-energy transition region to achieve the functionally active native state of the protein. Recent advances in theory and experiments with small single-domain proteins appear to indicate that the transition state for folding is topologically similar to the native state (*1*−*4*). The polypeptide chain finds the transition state presumably by a time-consuming search for certain crucial intrapolypeptide interactions that can support the required nativelike chain topology (*1*, *4*−*7*). The stochastic nature of the search mechanism, however, does not guarantee the full achievement of the relevant transition state. A failure could result in a wrong transition state with undesirable chain topology, the passage over which may then lead to partial or complete misfolding depending on the extent and location of the non-native interactions. Instances of formation of some non-native interactions in certain segments, and hence transient trapping of folding structures in local energy minima, are known (*8*−*10*). In the case of very high level misfolding, where overwhelming majority of non-native interactions are acquired, the misfolded state must unfold substantially in order to have a chance to refold correctly. In the absence of a parallel folding pathway, classical kinetics places such an intermediate at the dead-end of the folding pathway schematized by I ⇌ U ⇌ N, where I, U, and N, all three considered under the final refolding condition, refer to the off-pathway intermediate, the unfolded state, and the native state, respectively. The misfolded state is prone to aggregate or interact with other molecules, which seems to be the case for some proteins within the cell (*11*). One of the functions of cellular chaperones is to shield and protect the non-native polypeptides from misfolding and subsequent undesirable processes (*12*, *13*).

Here, we provide a direct demonstration of extensive misfolding of equine ferrocytochrome *c*, by observing denaturant dependence of refolding rates of the alkali-unfolded protein at a final pH of 7. Auxiliary kinetic experiments, analyses, and kinetic modeling have been carried out to determine the extent of misfolding and the pathway placement of the misfolded species. Both double-jump kinetics and NMR mapping of amide hydrogen exchange kinetics show a very high level of misfolding and no significant proportion of the population folding through a parallel "productive" route. Of the on-pathway and off-pathway mechanisms of folding, the results weigh in more for the latter.

MATERIALS AND METHODS

Cyt[1] *c* was from Sigma (Type VI). All experiments were done in strictly anaerobic atmosphere at 22 °C using 0.1 M sodium phosphate buffer for pH 7 conditions, and NaOH or NaOD (with or without 1 mM CAPS) for pH 12.7 conditions. Solutions contained 0.5−3 mM freshly prepared sodium dithionite, and experiments were completed within 2 h of exposing the protein to high pH. High-pH fluorescence changes were corrected by using NATA (*N*-acetyl tryptophanamide) fluorescence.

Equilibrium Unfolding. Cyt *c* solutions containing 1 mM CO, prepared in the 0−7 M range of GdnHCl or 7−13.25 range of pH, were deaerated and reduced under nitrogen with 1 mM sodium dithionite, and incubated in tightly capped quartz cuvettes or rubber-capped glass tubes for ∼30 min. Fluorescence emission spectra (ex: 280 nm) were taken in a FluoroMax-3 instrument (Jobin-Yvon, Horiba). Data were analyzed using the standard two-state equation for equilibrium unfolding (*14*).

Stopped-Flow Kinetics. Cyt *c* initially dissolved at pH 13 in the presence or absence of 4 M GdnHCl was reduced

† This work was supported by grants from Department of Biotechnology (BRB/15/227/2001), Department of Science and Technology (4/1/2003-SF), and University Grants Commission (UPE Funding), Government of India.

* Corresponding author. E-mail, akbsc@uohyd.ernet.in; fax, 91-40-2301-2460.

[1] Abbreviations: GdnHCl, guanidinium chloride; cyt, cytochrome; ferricyt, ferricytochrome; ferrocyt, ferrocytochrome; cyt-CO, carbon-monoxide adduct of ferrocytochrome *c*.

10.1021/bi060141z CCC: $33.50
Published on Web 06/15/2006

under nitrogen by adding a concentrated solution of sodium dithionite to a final concentration of ~3 mM. The solution was then saturated with CO under 1 atm of the gas. The resulting carbonmonoxycyt *c* is unfolded under the conditions used. Refolding was initiated by an 8-fold dilution of the protein solution with the refolding buffer. The final protein concentration in the refolding mixture was in the 5−50 μM range. For unfolding experiments, the native state of carbonmonoxycyt *c* (NCO) was prepared by manual dilution of the alkali-unfolded protein with the refolding buffer in 1:40 ratio. NCO was then unfolded by an 8-fold dilution into the unfolding buffer (0.1 M phosphate, 3 mM sodium dithionite, pH 7) containing different concentrations of GdnHCl. The final protein concentration in these experiments was ~4 μM.

For interrupted folding experiments, the alkali-unfolded protein solution (pH 13, with or without 4 M GdnHCl) was diluted 6-fold into the refolding buffer. Folding was allowed to proceed for variable time, t_i, before the solution was transferred to the appropriate unfolding medium. The amplitude of the unfolding reaction was analyzed as a function of t_i. A Bio-Logic SFM 400 instrument was used to study kinetics.

Hydrogen Exchange Pulse Labeling and NMR Spectroscopy. Cyt *c* with fully deuterated backbone amide sites was dissolved in NaOD at pD 12.9 with or without 3 M GdnHCl, reduced with sodium dithionite (3 mM), and unfolded by adding1 atm CO. Initial protein concentration was 6 mM. Refolding was initiated by a 6-fold dilution into a H_2O buffer containing 75 mM phosphate, pH 6.4. The final pH of the refolding medium was 7.3. After variable times of refolding, the solution was mixed with an equal volume of the H_2O pulse buffer containing of 50 mM CAPS, pH 10.4. The pH of the mixture at this stage was 9.6. After 50 ms of pulse time, the solution was combined with 0.7 vol of an H_2O quench buffer containing 0.5 M sodium acetate and 0.35 M ascorbate, pH 5. The pH at this stage was 5.05. The quenched solution was washed and concentrated at 5 °C using a D_2O buffer consisting of 7 mM citrate and 7 mM ascorbate, pH 5.05. The samples were then subjected to NMR analysis by recording magnitude COSY spectra in a 400 MHz Bruker spectrometer. Proton occupancies at individual amide sites were calculated by using the cross-peak intensities of a control unfolded sample whose amide sites were labeled maximally. Rates were extracted from plots of proton occupancies as a function of refolding time. Pulse-labeled samples were prepared using the Bio-Logic instrument in the quench-flow mode.

Preparation of the Flash-Frozen Intermediate and EPR. Cyt *c* was dissolved in 45% glycerol at pH 13, reduced with sodium dithionite, and liganded with nitric oxide. The refolding buffer, consisting of 45% glycerol buffered with 50 mM phosphate at pH 6.6, was placed in a quartz EPR tube and cooled to −11 °C. The NO-liganded unfolded protein was rapidly diluted 10-fold into the precooled refolding buffer and transferred immediately into a bath of liquid nitrogen. The final protein concentration was ~125 μM. The control sample was prepared by direct nitrosylation of the native protein at pH 7. Spectra were recorded at −150 °C using a JEOL (X-band) spectrometer with a microwave frequency of 9.18 GHz.

Computational. For numeric calculation of the microscopic rate constants, k_j (j = IU, UI, UN, and NU) in the scheme I ⇌ U ⇌ N, a 3 × 3 rate matrix was set up from the three-coupled linear differential equations for I, U, and N. The observable rates, λ_i (i = 1, 2), computed by diagonalizing the rate matrix are functions of k_js, the logarithm of which are assumed to have linear dependence on GdnHCl concentration (*15*). The equilibrium unfolding transitions were calculated from the eigenvectors by assigning expected values of fluorescence to I, U, and N in the range 0−1.

RESULTS AND DISCUSSION

Ferrocytochrome *c* (cyt *c*) unfolds in alkaline aqueous medium when carbonmonoxide is allowed to bind to its ferrous heme iron (Figure 1A). The pH-midpoint for the unfolding transition is ≈12.5. Unfolding of carbonmonoxy-cytochrome *c* (cyt-CO) is substantial even in the aqueous medium, because the addition of the denaturant guanidinium hydrochloride (GdnHCl) to the protein solution at pH 13 or higher does not produce any considerable change in the fluorescence emission by the lone tryptophan (W59). We do not however suggest this as an evidence for a random coil-like state. Completeness of unfolding is a controversial issue and cannot be shown unequivocally with a number of techniques.

The Reverse-Denaturant Effect and Protein Misfolding. For a quantitative kinetic study, we prepared unfolded cyt-CO by adding 1 atm CO to the protein solution held at pH 13 with or without 4 M GdnHCl and measured the millisecond refolding rate constants by stopped-flow dilution of the unfolded protein solution into refolding buffers containing different concentrations of GdnHCl, pH 7, at 22 °C. Figure 1B reports single-exponential fits of two representative refolding traces monitored by fluorescence. The observed refolding rate in the presence of 0.1 M GdnHCl (k_{obs} = 11.5 s^{-1}) is 10-fold slower than that for 2 M GdnHCl (k_{obs} = 120 s^{-1}). Both traces also show burst loss of fluorescence signals, larger for the former. The excellent quality of data allowed us to quantify the folding-unfolding rates and kinetic fluorescence amplitudes as a function of GdnHCl, pH 7 (Figure 1C,D). Interestingly, under strongly nativelike conditions (<2M GdnHCl), the folding rate decreases linearly with decreasing concentration of the denaturant, producing a sharp inversion in the folding limb of the chevron (Figure 1C). This reverse-denaturant effect does not depend on the content of GdnHCl in the initial unfolded protein solution at pH 13, although the refolding rates in the inversion zone are marginally faster when the denaturant is included (Figure 1C). We take the accentuated inversion of the rate-denaturant profile as an evidence for extensive misfolding: the misfolded species is increasingly stabilized and hence accumulates as the denaturant concentration in the refolding medium falls lower. The misfolded product cannot, however, achieve forward folding because of the preponderance of non-native interactions. In this sense, it is a dead-end intermediate, and it needs to unfold extensively in order to have a chance to refold correctly. The gradient of the rate-denaturant plot, $m^{\ddagger}$, is generally calculated from $\ln k_{obs} = \ln k^{\circ} + (m^{\ddagger}/RT)[D]$, where k_{obs} is the observed rate, $\ln k^{\circ}$ is the rate in the absence of the denaturant, and [D] is the concentration of the denaturant (*16*). The positive slope of the inverted part of the refolding

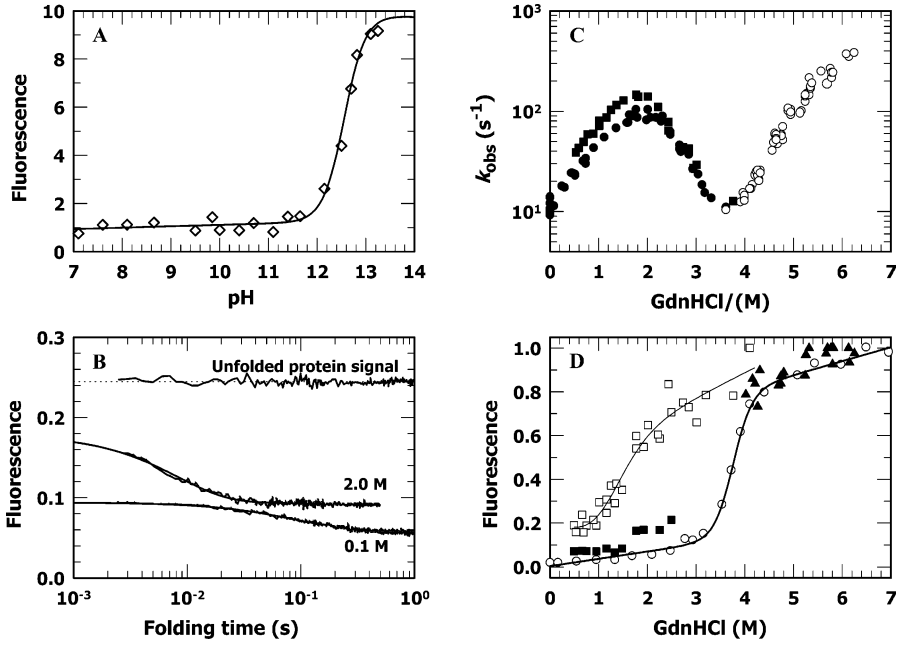

FIGURE 1: The reverse denaturant effect and extensive misfolding of alkali-unfolded carbonmonoxycytochrome *c* at a final pH of 7, 22 °C. (A) Alkaline pH-induced unfolding of ferrocytochrome *c* in the presence of CO. The pH-midpoint for the transition is ≈12.5. (B) Representative one-exponential refolding traces in the millisecond window showing the burst loss of fluorescence with reference to the unfolded signal. With increasingly nativelike condition, the burst amplitude increases but the refolding slows down (k_f is 120 and 11 s^{-1} in the presence of 2.0 and 0.1 M GdnHCl, respectively). (C) Observed relaxation rate constants versus GdnHCl. The symbols ● and ■ represent refolding rates when the alkali-unfolded CO-bound protein solution contained 0 and 4 M GdnHCl, respectively. Open symbols (○) are results from unfolding experiments. (D) Denaturant distribution of burst fluorescence (□), and equilibrium fluorescence extracted from kinetic traces (■ and ▲). The circles represent the unfolding transition measured by an independent equilibrium experiment. The lines are two-state transition fits (*14*).

limb, $m_f^{\ddagger}$ (≈0.85 kcal mol^{-1} M^{-1}), is pretty close to the slope defined by the unfolding limb of the main chevron ($m_u^{\ddagger}$ ≈1 kcal mol^{-1} M^{-1}). This comparison reflects the enormity of the unfolding of the misfolded species (Figure 1C). Since the intermediate unfolds in milliseconds, its rate of formation must be faster. We assign the burst-associated submillisecond phase to the formation of the intermediate. The denaturant dependence of the burst signal, normalized relative to the fluorescence of the unfolded protein, describes the titration of the misfolded intermediate whose energetic stability is ∼2 ± 0.1 kcal mol^{-1} (Figure 1D). Thus, we consider the I ⇌ U ⇌ N mechanism, where I is the dead-end product and U is the initial unfolded chain under the refolding conditions employed.

No Aggregation of the Protein. We needed to ascertain that the observed misfolding was not related to protein dimerization or aggregation at any stage, since short-lived aggregates may cause a downward curvature in the folding limb of the chevron plot (*17*). Several control experiments were done. First, by titrating the unfolded protein with CO in the presence or absence of GdnHCl, we find that only one CO binds to the protein molecule (Figure 2A), suggesting that the unfolded state is monomeric. Second, the refolding rate is invariant to the concentration of protein used (Figure 2B), indicating that the structural attribute of the intermediate whose accumulation produces the inversion in the folding chevron is unlikely to be a protein aggregate or a nonmonomeric form. Third, the EPR spectrum of the flash-frozen intermediate prepared from the nitrosyl derivative of the unfolded protein shows neither dipolar broadening nor a significant reduction in intensity (Figure 2C), implying no close proximity of the unpaired electron spins due to the NO ligands bound to the hemes of two different chains. The intermediate rather exhibits a hyperfine pattern (Figure 2C) characteristic of monomeric nitrosylcytochrome *c* (*18*), and demonstrates that only one NO binds to the ferrous heme. Fourth, the NMR spectrum of alkaline ferrocyt *c* (pH 12.9) does not change at all even after tens of hours of sample preparation (Figure 2D). Furthermore, the refolding rates of alkaline ferrocyt *c* respond little to the variation of the ionic strength of the medium (unpublished result), and the time-of-flight mass spectra of the alkali-unfolded protein show no protein modification (*7*). These results eliminate the possibility of protein dimerization, aggregation, or chemical modification at different stages of folding.

Rate of Formation of Native Molecules Derived from Interrupted Refolding. To eliminate the possibility that a considerable proportion of the unfolded population folds through a parallel productive pathway, we performed interrupted refolding experiments (*19*, *20*). The idea here is to resolve the rate of formation of N. If an unfolding reaction is initiated at some point of time while folding is still in progress, then only those molecules that have already achieved the native structure will undergo unfolding, and the total fluorescence change observed for this unfolding process will be proportional to the amount of native molecules that were present at the time of interruption. A plot of total fluorescence change for unfolding as a function of time of folding interruption is essentially a plot of native molecules formed as a function of time of folding, and it provides the rate of formation of N in a straightforward manner. In our experiments, unfolded cyt-CO was allowed to refold in the presence of 0.5 M GdnHCl for different times, t_i, and then driven back to the initial unfolded state. The amplitude of the unfolding reaction plotted as a function of t_i produces an exponential time course described by a rate

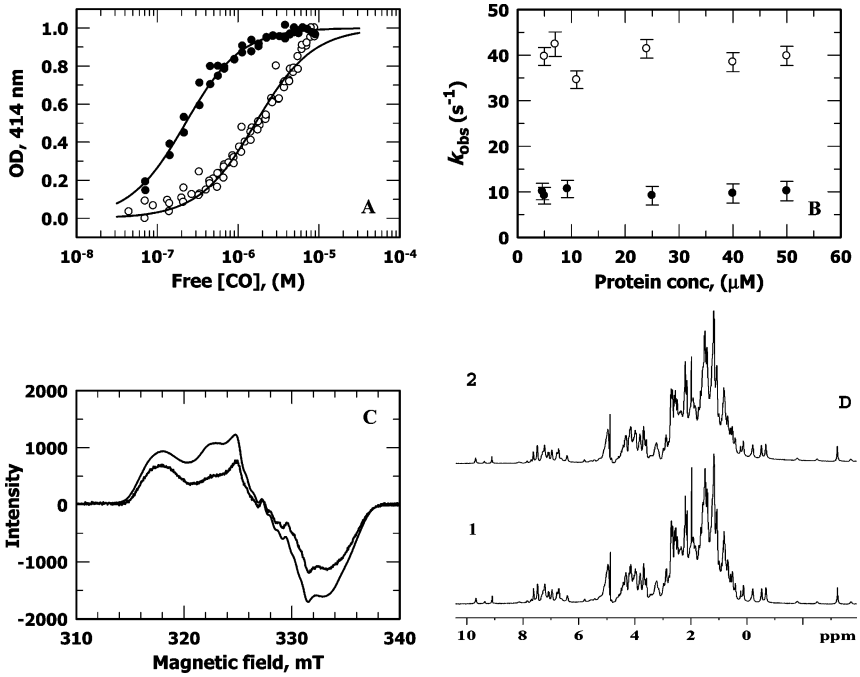

FIGURE 2: The protein does not aggregate or dimerize at different stages of folding. (A) CO titration of alkaline ferrocyt *c* with (●) and without (○) 3 M GdnHCl, pH 13. Fits to data are according to the equation, $y(x) = 1/[1 + 10^{n(x-K)}]$. Value of *n*, the number of CO bound, is unity in each case. Apparent dissociation constant, *K*, is 1.64 and 22 μM in the presence and absence of the denaturant, respectively. (B) Protein concentration independence of the observed folding rate in the presence of 0 M (●) and 0.55 M GdnHCl (○). (C) EPR spectrum of the frozen nitrosylated intermediate in 45% glycerol, pH 7. The control spectrum (thin line) was generated by direct nitrosylation of ferrocyt *c* at pH 7. Both spectra are characterized by the "triplet of triplet" hyperfine structure, typical of NO-bound monomeric ferrocyt *c* (*18*). (D) NMR spectra of ferrocyt *c* at pH 12.9: soon after dissolving the protein (spectrum 1), and 38 h later (spectrum 2).

constant, $k \approx 33\ s^{-1}$ (Figure 3A), which is equal to the rate that is obtained in direct refolding experiments (Figure 1C), suggesting that the formation of N is rate-limited by the unfolding of the misfolded I, and that the burst phase cannot be associated with a parallel folding pathway.

Reverse Sodium Sulfate Effect under Strongly Refolding Conditions. We also checked the refolding rate in 0.5 M GdnHCl as a function of molar concentration of Na_2SO_4. The rationale was that Na_2SO_4, which is known to stabilize intermediate species (*20*, *21*), would retard the conversion of I to N via U if I were an off-pathway intermediate, and it did (Figure 3B). In the case of lysozyme refolding, for which a triangular parallel mechanism has been established, the rate of I → N conversion becomes faster in the presence of Na_2SO_4 as the GdnHCl concentration in the refolding medium is decreased (*20*). But Na_2SO_4 is expected to stabilize N as well. This also happens and is detectable in the range of GdnHCl concentrations where the folding mechanism is essentially two-state, U ⇌ N (Figure 3B). As moderate to strongly nativelike conditions are approached and the mechanism shifts from U ⇌ N to I ⇌ U ⇌ N, the intermediate accumulates, and the refolding kinetics are not affected by the stabilization of N by Na_2SO_4. This is because the unfolding of the misfolded intermediate rate-limits the formation of N according to the I ⇌ U ⇌ N off-pathway mechanism. To illustrate this point, we present chevron data for refolding in the presence and the absence of 0.4 M Na_2SO_4 (Figure 3C). In the presence of denaturing concentrations of GdnHCl (>2 M), where the folding mechanism is U ⇌ N, addition of Na_2SO_4 accelerates folding, indicating the stabilization of N by Na_2SO_4. Under two-state conditions, the folding rate is directly controlled by U → N conversion. When I is populated (<2 M GdnHCl), the mechanism shifts to I ⇌ U ⇌ N. Now we see the stabilizing effect of Na_2SO_4 on I. It is this stabilization that decelerates the unfolding of I to U, and since I → U conversion rate-limits the overall folding, the stabilization of N does not affect the folding rate so long as the I ⇌ U ⇌ N mechanism operates.

Simulation to an Off-Pathway Mechanism. We then simulated the experimental data by numeric calculation of the denaturant dependences of the eigenvalues, λ_i, and the concentrations of each of the species corresponding to the I ⇌ U ⇌ N mechanism. Figure 4A shows one of the solutions. The two chevrons resulting from the two eigenvalues (λ_1 and λ_2) produce a saddle point region (~ 2 M GdnHCl). In higher concentrations of the denaturant, the refolding mechanism is effectively U ⇌ N. As progressively refolding conditions are achieved (<2 M GdnHCl), I is increasingly stabilized and hence accumulates. Consequently, the refolding limb of the λ_2 chevron inverts instead of continuing linearly to the ordinate. The positive slope of the inverted region of the folding limb is due to the unfolding of I to U, and since no kinetic phase is observed at longer times, λ_2 (which approximates the microscopic rate constant for the I → U process under strongly refolding conditions) limits the conversion of U to N. GdnHCl dependences of the four microscopic rate constants that provide the eigenvalue solution are shown in Figure 4B. Values of these rate constants in water, k_j° (s^{-1}), and the kinetic *m*-values, $m_j^\ddagger$ (kcal mol^{-1} M^{-1}), are listed in Table 1. The saddle region contains the transition midpoint of I. This is also shown by the simulation of the denaturant distribution of the burst amplitude according to the I ⇌ U ⇌ N scheme (Figure 4C). These results suggest that the dead-end product is fully formed in the submillisecond regime.

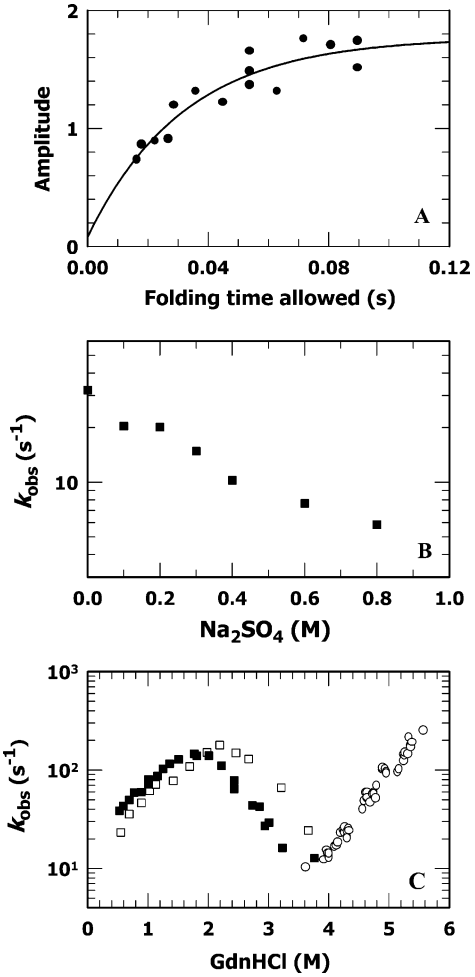

FIGURE 3: The rate of formation of N is limited by the rate of unfolding of the intermediate, I. (A) Record of the amplitude of unfolding traces as a function of variable folding time, t_i. Folding was allowed in the presence of 0.5 M GdnHCl, and unfolding was initiated by transferring the folding mixture to 5 M GdnHCl. Single-exponential fit of the amplitude versus refolding time gives a rate constant of 33 s^{-1}. (B) Variation of the rate of folding in the presence of 0.5 M GdnHCl as a function of Na_2SO_4. Folding is retarded because I is stabilized by Na_2SO_4. (C) Denaturant dependence of refolding rate constants in the absence (■) and the presence (□) of 0.4 M Na_2SO_4. The same alkali-unfolded CO-bound protein solution containing 4 M GdnHCl was used for both sets of refolding experiments.

The time constants for U → I (190 μs) and U → N (220 μs) predicted by the simulation appear quite small. We did not see a time as fast as 220 μs for the U → N process in any of the previous studies (*7*, *22*−*24*). The reason for this could be the use of the linear *k*−denaturant relationship. Stopped-flow chevrons of ferrocyt *c* and cyt-CO under different conditions of initial and/or final pH, except for the present ones, invariably show pronounced curvatures in the folding limbs. Those results bind one to use nonlinear or quadratic *k*−denaturant interpretations, and the folding time in water is found in the 1−2 ms range (*7*, *22*−*24*). Because of the inversion of the chevron in the present case (Figure 4A), the curvature is not detected, so we chose to simulate the data using the linear *k*−denaturant relation. Thus, the simulated values for the rate constants for U → I and U → N, especially the latter, are substantially larger.

Hydrogen Exchange Pulse Labeling. We used a pulse hydrogen exchange labeling strategy (*25*, *26*) to probe

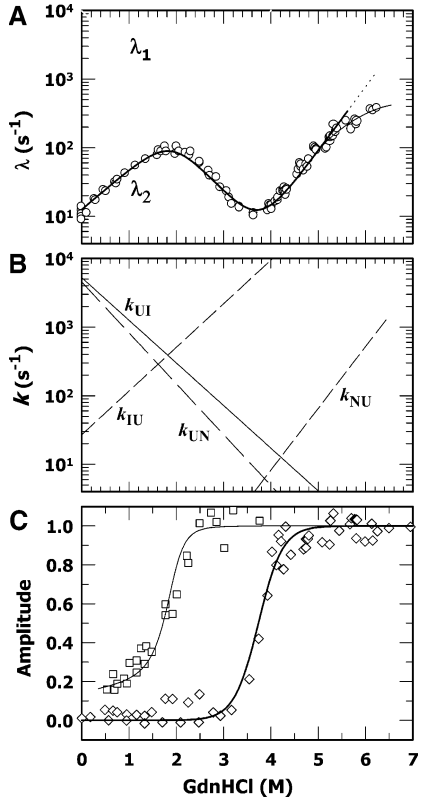

FIGURE 4: Simulation of observed rates and amplitudes according to the I ⇌ U ⇌ N mechanism. (A) Variation of the eigenvalues (λ_1 and λ_2) as a function of the denaturant. The inverted dependence of λ_2 on GdnHCl in the folding limb suggests the unfolding of I to U. The reason for the slight rollover in the unfolding limb of the chevron is unknown. (B) The k_j-denaturant space corresponding to denaturant dependences of the two eigenvalues. (C) GdnHCl-induced transitions of I and N predicted according to the dead-end mechanism.

Table 1: Numeric Calculations of Kinetic Parameters for Off-Pathway and On-Pathway Models[a]

parameters	I ⇌ U ⇌ N	U ⇌ I ⇌ N
k_{UI}°	5250	15000
$m_{UI}^{\ddagger}$	−0.83	−0.9
k_{IU}°	27	10
$m_{IU}^{\ddagger}$	0.85	1.2
k_{UN}°	4550	−
$m_{UN}^{\ddagger}$	−0.99	−
k_{NU}°	0.002	−
$m_{NU}^{\ddagger}$	1.2	−
k_{IN}°	−	11
$m_{IN}^{\ddagger}$	−	0.85
k_{NI}°	−	0.02
$m_{NI}^{\ddagger}$	−	0.95

[a] Values are given in s^{-1} for k_j° and kcal mol^{-1} M^{-1} for $m_j^{\ddagger}$.

structure formation during folding in the presence of 0.5 M GdnHCl (Figure 5). For all NMR-resolved NH sites, only moderate protection against proton labeling was seen at ∼14 ms of refolding time, although within ∼55 ms the proton occupancy dropped to about 60−70% in a lone kinetic phase corresponding to the I → N transition via U. The rate extracted from the time course of protection against H-exchange labeling at four representative NH sites is 30 ± 2 s^{-1} (Figure 5D), consistent with those obtained from direct stopped-flow refolding and interrupted refolding protocols. The single rate for protection suggests that no transient structure accumulates during the folding of I to N.

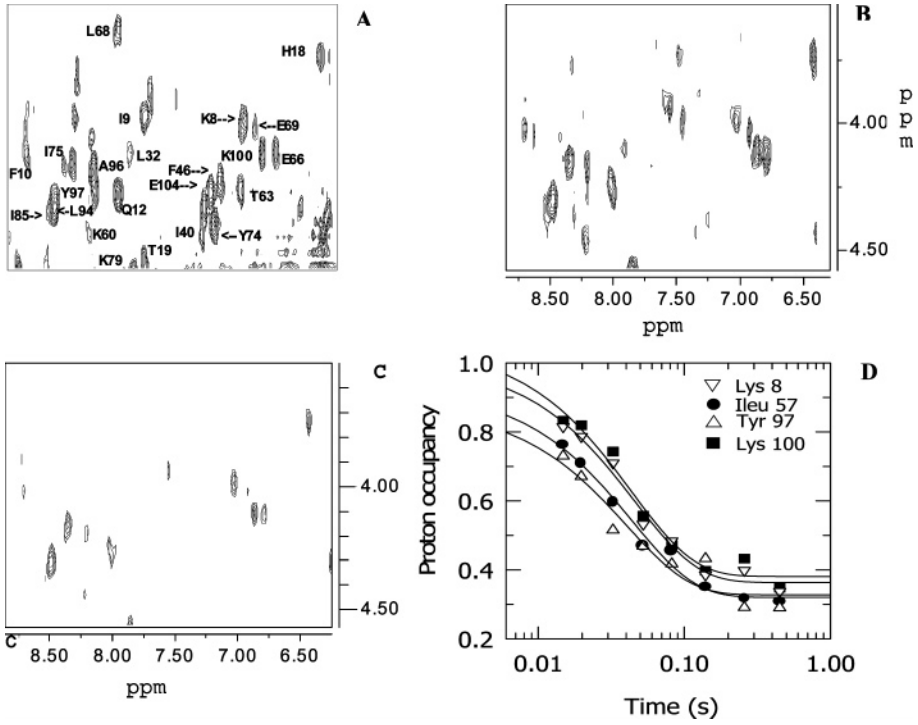

FIGURE 5: NMR analysis of representative samples prepared by hydrogen exchange pulse labeling. NH−C_αH region of the magnitude COSY (*J*-correlated) spectrum of ferrocyt *c*: (A) native protein solution at pH 5.1; (B) pulse-labeled sample at 14 ms of refolding time; (C) pulse-labeled sample at 55 ms of refolding time. (D) Time course of protection against hydrogen exchange for a set of four representative amide sites. Single-exponential fits to the data yield 30 ± 2 s^{-1} for the refolding rate.

Consideration of the On-Pathway Model. The phenomenon of deviation from the predicted linear dependence of logarithm of folding rate on denaturant can be both qualitatively interpreted and quantitatively modeled in terms of off-pathway as well as on-pathway placement of an intermediate (*20*, *27*). Several kinetic studies have considered the nonlinear denaturant dependence of the logarithm of folding rates, and the burst signal changes observed in stopped flow or quenched-flow experiments in terms of a sequential on-pathway model of protein folding (*28*−*30*; reviewed in ref *31*). For the current description of data, we now consider the on-pathway mechanism. The commonly used starting point in the classical discussion of the on-pathway model is that the interconversion rates between U and I are much faster than those between I and N, so the folding rate is equal to the rate of the I → N transition modulated by the population of I: $\lambda_2 \sim k_f = k_{IN} \cdot f_I$, where f_I is the fractional population of the intermediate (*31*). Accordingly, one can examine if the on-pathway model, U ⇌ I → ‡ → N, can explain the data described here. The two assumptions of the model are (1) U and I are in rapid equilibrium on the time-scale of the experiment, and (2) the transition state (‡, TS) is extensively unfolded with respect to I. The latter assumption will be essential to account for the chevron inversion under strongly refolding conditions.

Figure 1C shows that the reverse denaturant effect already sets in at ~2 M GdnHCl, but the burst phase transition plotted in Figure 1D registers ~1.5 M GdnHCl for the C_m of unfolding of the intermediate. This may imply that U and I are not in rapid equilibrium. When a rapid U ⇌ I equilibrium exists, the on-pathway model would predict that the overall folding rate should decrease as the equilibrium is shifted toward U or I is destabilized with increasing concentration of the denaturant, unless the denaturant acts to lower the effective energy barrier. The assumption of extensively unfolded nature of the TS also presents a difficulty, since accumulating evidences for a large number of small proteins strongly indicate that the TS is substantially structured, and resembles the native state to an extent of ~63%, suggesting that roughly two-thirds of the surface area that is buried in the native state becomes buried in the transition state (*32*−*34*). This is also consistent with theoretical studies (*35*, *36*). Thus, even if the experiments presented in this paper cannot entirely exclude an on-pathway model, it appears that more evidence will be required to examine the on-pathway mechanism. Nonetheless, we compare the numeric calculation results for the on-pathway scheme (U ⇌ I ⇌ N) with those for the dead-end mechanism in Table 1. The on-pathway calculations require that I unfolds extensively before achieving N, so that the relevant TS is substantially disordered. The $k_{UI}°$ value appears large for both mechanisms. While the value predicted by the off-pathway scheme ($\tau \sim 220$ μs) is still within the range observed for the faster folding proteins, it is much larger ($\tau \sim 67$ μs) for the on-pathway mechanism. The concern is whether this time is enough for such a massive surface burial ($m_{UI}^{\ddagger} = -0.9$).

Although the TS is generally placed closer to the native state along the folding coordinate, the operational assumption of extensive unfoldedness of the TS with respect to I affords an alternative interpretation for the observed Na_2SO_4-induced decrease in the folding rate at low GdnHCl concentrations, irrespective of on-pathway and off-pathway models. This explanation has been kindly offered by a reviewer of this paper. Figure 6 shows the free energies of the different states for the on-pathway U → I → ‡ → N model, where the TS is extensively unfolded with respect to I. The free energies have been calculated from the observed equilibrium and kinetic data and scaled by arbitrarily setting to zero the free

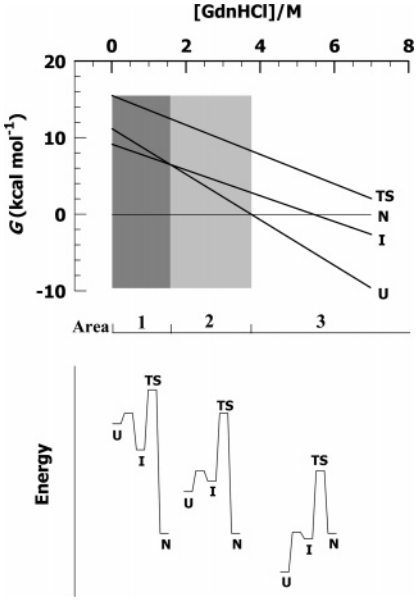

FIGURE 6: Graphical representation of the free energies of the different states as a function of GdnHCl for an on-pathway model. The dependence of the free energy of the N state is set to zero arbitrarily. Areas 1 and 2 are shaded in dark and light gray, respectively. Area 3 is not shaded. The lower panel illustrates the energy level diagrams corresponding to the three areas.

energy of the N state with GdnHCl. As strongly nativelike conditions are approached (area 1), I is populated, and the rate of the reaction decreases as the energy gap between the TS and I states increases. The addition of Na_2SO_4 will stabilize the I state with respect to the TS because the latter is more unfolded. Consequently, the observed folding rates will decrease. Under moderately folding conditions (area 2), I is no longer populated, and the rate of the reaction decreases commensurate with the decrease in the energy gap between the TS and U states as the concentration of GdnHCl is decreased. The addition of Na_2SO_4 will now stabilize the TS with respect to the U state, and therefore, the rate of folding will increase as seen in Figure 3C. Under unfolding conditions (area 3) where the U state is more stable, the rate of the unfolding reaction increases as the energy gap between the TS and N states diminishes with increase in GdnHCl concentration. Here, the presence of Na_2SO_4 will stabilize the N state with respect to the TS such that the rate of unfolding will increase.

To summarize this section, examination of the on-pathway mechanism is hindered by the lack of evidence for both faster interconversion rates between U and I, and an extensively unfolded transition state.

Off-Pathway Intermediate. Whether intermediates are on the pathway or are merely side products is one of the longest-running controversies (*37*). Here, experimental record of a fundamental kinetic criterion, namely, inversion of the denaturant dependence of the observed folding rate reveals the off-pathway dead-end mechanism. This direct observation has obviated the need for checking out the other off-pathway kinetic criteria based on ϕ-value analysis (*38*), pulse-chase competition experiments (*39*), and the evaluation by experiments of various microscopic rate constants (*40*).

Thus, proteins can nearly completely misfold in the test tube as some do in the cell (*11*). Clearly, not all kinetic intermediates play a productive role. The vulnerability of the abortive product especially in the cytosolic milieu adds on to the inefficiency of folding. It is not surprising that the cell engages a variety of molecular chaperones to achieve folding efficiency (*12*, *13*).

The Misfolding of Alkaline Ferrocytochrome c Is Not Directly Related to Heme Ligation. Dead-end intermediates associated with intrinsically slow steps in protein folding have been reported for the oxidative folding pathway of BPTI (*41*), and for an intermediate with non-native prolyl isomers in RNAse T1 (*42*). In the present case, one can consider such a possibility by examining the influence of the axial non-native heme ligation at different stages of folding of cytochrome *c*.

Refolding of oxidized cytochrome *c* (ferricyt *c*) at neutral pH has been shown to proceed through an intermediate, I_{NC}, in which the N- and C-terminal helices dock against each other in an orthogonal geometry (*26*, *43*). There have, however, been conflicting viewpoints regarding the label of I_{NC}, whether off-pathway or on-pathway. The native-state axial ligation of M80 to the oxidized heme is replaced by H26 and H33 under denaturing conditions (*44*). Even though non-native, the $Fe^{3+}-H26$ and $Fe^{3+}-H33$ contacts are so stable that the dissociation rates of these two ligands from the iron are far smaller than the protein folding rate. A quantitative kinetic model of these processes has been described earlier (*22*, *45*). As a result, the ferricyt polypeptide collapses with persisting $Fe^{3+}-H26$ and $Fe^{3+}-H33$ contacts. This act of trapping of the non-native histidines displaces the histidine-resident chain segments from the normal distal side of the heme to the proximal side of M80. The misconfiguration arrests folding because the frustrated chain organization stays in a kinetically frozen intermediate state, I_{NC}, until thermal dissociation of non-native histidines from the heme iron gives way to the ligation of M80 and, hence, polypeptide reconfiguration. In this picture, an unfolding event disperses the incorrect chain organization in order to facilitate correct folding, so the intermediate could be viewed as a trapped misfolded intermediate (*22*, *24*, *45*, *46*). On the other hand, the intermediate may be described, perhaps inappropriately, just at the level of the relevant N- and C-terminal helices without paying any attention to wrong chain organizations in other parts of the molecule. I_{NC} then carries the on-pathway label (*47*). The problem of kinetic trap or chain misfolding in ferricyt *c* can be nearly fully eliminated by allowing refolding at low pH or in the presence of an extrinsic heme ligand that blocks non-native histidine ligation. Folding kinetics are then modeled as a two-state reaction (*5*).

These conflicts hardly surface in the case of refolding of virtuous ferrocytochrome *c* (ferrocyt *c*) under normal conditions of refolding, since non-native heme ligands do not interfere with inherent accelerated folding of this redox state of cyt *c*. We have shown earlier that 45% of the unfolded ferrocyt molecules have no equilibrium occupancy at the sixth coordination site of the heme, and ligands do nothing to their folding (*45*). For the remaining 55%, a liganded form will enter the folding pathway only if the rate of dissociation of the ligand from the heme iron is slower than the folding rate (*48*). The dissociation time constants for all four ligands of ferrocytochrome *c* are smaller than the polypeptide refolding time (*45*). Thus, heme ligation in the unfolded state of ferrocytochrome *c* does not interfere with refolding.

The exact interaction(s) that facilitate misfolding of alkaline ferrocytochrome *c* liganded with CO is hard to figure out with the available data. It is possible to invoke a change in the average chain configuration of cyt-CO in the initial alkaline medium, or a charge-dipole interaction between the CO electric dipole (0.112 D) and one of a number of possible charged groups in the proximity. But these could be mere speculations. More remains to be done to obtain a mechanistic description of the observed misfolding at the level of intramolecular interactions.

ACKNOWLEDGMENT

We thank Walter Englander for discussions and suggestions. This research was supported by the Departments of Biotechnology (BRB/15/227/2001) and Science and Technology (4/1/2003-SF). A.K.B. is the recipient of a Swarnajayanti Fellowship from the Department of Science & Technology, Government of India.

REFERENCES

1. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) Understanding protein folding via free-energy surfaces from theory and experiment, *Trends Biochem. Sci. 25*, 331−339.
2. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins, *J. Mol. Biol. 277*, 985−994.
3. Baker, D. A. (2000) A surprising simplicity to protein folding, *Nature 405*, 39−42.
4. Makarov, D. E., and Plaxco, K. W. (2003) The topomer search model: a simple, quantitative theory of two-state protein folding kinetics, *Protein Sci. 12*, 17−26.
5. Sosnick, T. R., Mayne, L., and Englander, S. W. (1996) Molecular collapse: the rate-limiting step in two-state cytochrome *c* folding, *Proteins 24*, 413−426.
6. Rumbley, J., Hoang, L., Mayne, L., and Englander, S. W. (2001) An amino acid code for protein folding, *Proc. Natl. Acad. Sci. U.S.A. 98*, 105−112.
7. Bhuyan, A. K., Rao, D. K., and Prabhu, N. P. (2005) Protein folding in classical perspecive: folding of horse cytochrome *c*, *Biochemistry 44*, 3034−3040.
8. Hamada, D., Segawa, S.-I., and Goto, Y. (1996) Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein, *Nat. Struct. Biol. 3*, 868−873.
9. Troullier, A., Reinstädler, D., Dupont, Y., Naumann, D., and Forge, V. (2000) Transient non-native secondary structures during the refolding of alpha-lactalbumin detected by infrared spectroscopy, *Nat. Struct. Biol. 7*, 78−86.
10. Capaldi, A. P., Kleanthous, C., and Radford, S. E. (2002) Im7 folding mechanism: misfolding on a path to the native state, *Nat. Struct. Biol. 9*, 209−216.
11. Dobson, C. M. (2004) Principles of protein folding, misfolding and aggregation, *Semin. Cell Dev. Biol. 15*, 3−16.
12. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded proteins, *Science 295*, 1852−1858.
13. Gething, M.-J., and Sambrook, J. (1992) Protein folding in the cell, *Nature 355*, 33−45.
14. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants, *Biochemistry 27*, 8063−8068.
15. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Adv. Protein Chem. 24*, 1−95.
16. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) Extremely rapid protein folding in the absence of intermediates, *Nat. Struct. Biol. 2*, 663−673.
17. Silow, M., Tan, Y.-J., Fersht, A. R., and Oliveberg, M. (1999) Formation of short-lived protein aggregates directly from the coil in two-state folding, *Biochemistry 38*, 13006−13012.
18. Yonetani, T., Yamamoto, H., Erman, J. E., Leigh, J. S., and Reed, G. H. (1972) Electromagnetic properties of hemoproteins. V. Optical and electron paramagnetic resonance characteristics of nitric oxide derivatives of metalloporphyrin-apohemoprotein complexes, *J. Biol. Chem. 247*, 2447−2455.
19. Schmid, F. X. (1986) Fast-folding and slow folding forms of unfolded proteins, *Methods Enzymol. 131*, 70−82.
20. Wildegger, G., and Kiefhaber, T. (1997) Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate, *J. Mol. Biol. 270*, 294−304.
21. von Hippel, P. H., and Wong, K.-Y. (1965) On the conformational stability of globular proteins. The effect of various electrolytes and nonelectrolytes on the thermal ribonuclease transition, *J. Biol. Chem. 240*, 3909−3923.
22. Bhuyan, A. K., and Kumar, R. (2002) Kinetic barriers to the folding of horse cytochrome *c* in the reduced state, *Biochemistry 41*, 12821−12834.
23. Prabhu, N. P., Kumar, R., and Bhuyan, A. K. (2004) Folding barrier in horse cytochrome *c*: support for a classical folding pathway, *J. Mol. Biol. 337*, 195−208.
24. Kumar, R., and Bhuyan, A. K. (2005) Two-state folding of horse ferrocytochrome *c*: analyses of linear free energy relationship, chevron curvature, and stopped-flow burst relaxation kinetics, *Biochemistry 44*, 3024−3033.
25. Udgaonkar, J. B., and Baldwin, R. L. (1988) NMR evidence for an early intermediate in the folding pathway of ribonuclease A, *Nature 335*, 694−699.
26. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR, *Nature 335*, 700−704.
27. Creighton, T. E. (1997) How important is the molten globule for correct protein folding, *Trends Biochem. Sci. 22*, 6−10.
28. Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) Transient folding intermediates characterized by protein engineering, *Nature 346*, 440−445.
29. Houry, W. A., Rothwarf, D. M., and Scheraga, H. A. (1995) The nature of the initial step in the condormational folding of disulfide-intact ribonuclease A, *Nat. Struct. Biol. 2*, 495−503.
30. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) Evidence for a three-state model of protein foding from kinetic analysis of ubiquitin variants with altered core residues, *Nat. Struct. Biol. 3*, 193−205.
31. Baldwin, R. L. (1996) On-pathway versus off-pathway folding intermediates, *Folding Des. 1*, R1−R8.
32. Prabhu, N. P., and Bhuyan, A. K. (2006) Prediction of folding rates of small proteins: empirical relations based on length, secondary structure content, residue type, and stability, *Biochemistry 45*, 3805−3812.
33. Fersht, A. R. (1997) Nucleation mechanism in protein folding, *Curr. Opin. Struct. Biol. 7*, 3−9.
34. Paci, E., Lindorff-Larsen, K., Dobson, C. M., Karplus, M., and Vendruscolo, M. (2005) Transition state contact orders correlate with protein folding rates, *J. Mol. Biol. 352*, 495−500.
35. Socci, N. D., Onuchic, J. N., and Wolynes, P. G. (1998) Protein folding mechanisms and the multidimensional folding funnel, *Proteins 32*, 136−158.
36. Wolynes, P. G. (2004) Latest folding game results: protein A barely frustrates computationalists, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6837−6838.
37. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossman, J. G., Sato, S., Jas, G. S., Freund, S. M. V., Alonso, D. A. V., Daggett, V., and Fersht, A. R. (2003) The complete folding pathway of a protein from nanoseconds to microseconds, *Nature 421*, 863−867.
38. Fersht, A. R. (1993) Protein folding and stability: the pathway of folding of barnase, *FEBS Lett. 325*, 5−16.
39. Laurents, D. V., Bruix, M., Jamin, M., and Baldwin, R. L. (1998) A pulse-chase-competition experiment to determine if a folding intermediate is on or off-pathway: application to ribonuclease A, *J. Mol. Biol. 283*, 669−678.
40. Bai, Y. (1999) Kinetic evidence for an on-pathway intermediate in the folding of cytochrome *c*, *Proc. Natl. Acad. Sci. U.S.A.* 96, 477−480.
41. Weisman, J. S., and Kim, P. S. (1991) Reexamination of the folding of BPTI: predominance of native intermediates, *Science 253*, 1386-1393.
42. Kiefhaber, T., and Schmid, F. X. (1992) Kinetic coupling between protein folding and prolyl isomerization. II. Folding of ribonuclease A and ribonuclease T1, *J. Mol. Biol. 224*, 231−240.
43. Elöve, G. A., Bhuyan, A. K., and Roder, H. (1994) Kinetic mechanism of cytochrome *c* folding: involvement of the heme and its ligands, *Biochemistry 33*, 6925−6935.

44. Muthukrishnan, K., and Nall, B. T. (1991) Effective concentrations of amino acid side chains in an unfolded protein, *Biochemistry 30*, 4706−4710.

45. Bhuyan, A. K., and Udgaonkar, J. B. (2001) Folding of horse cytochrome *c* in the reduced state, *J. Mol. Biol. 312*, 1135−1160.

46. Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994) The barriers in protein folding, *Nat. Struct. Biol. 1*, 149−156.

47. Colón, W., Elöve, G. A., Wakem, L. P., Sherman, F., and Roder, H. (1996) Side chain packing of the N- and C-terminal helices plays a critical role in the kinetics of cytochrome *c* folding, *Biochemistry 35*, 5538−5549.

48. Zwanzig, R. W. (1997) Two-state models of protein folding kinetics, *Proc. Natl. Acad. Sci. U.S.A.* 94, 148−150.

BI060141Z