Native-like Interactions Favored in the Unfolded Bovine Pancreatic Trypsin Inhibitor Have Different Roles in Folding†

Renhao Li,‡ John L. Battiste,§ and Clare Woodward*

Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, Minnesota 55108

Received August 21, 2001; Revised Manuscript Received November 27, 2001

ABSTRACT: Folding kinetics of a series of bovine pancreatic trypsin inhibitor (BPTI) variants with similar stabilities and structures have been measured. All are strongly destabilized relative to WT. In Y21A, F22A, Y23A, G37A, and F45A, the three native disulfide bonds are retained. In RM(14−38), Cys14 and Cys38 thiols are methylated while C30−C51 and C5−C55 disulfides remain intact. At pH 2 and 20 °C, relaxation rate constants of the major kinetic phase range from ∼10 ms to 0.71 s in the absence of denaturant. All mutants except G37A exhibit standard two-state behavior. Y21A, F22A, and Y23A fold much more slowly than other mutants. The experiments were designed to test the hypothesis that native-like structure detected in the unfolded BPTI is important in folding. Two native-like contacts are implied by NOEs in reduced and unfolded BPTI, between residues Tyr23 and Ala25, and between Gly37 NH and the Tyr35 ring. The results support an earlier hypothesis that formation of the central β-hairpin, monitored by a local native interaction between Tyr23 and Ala25, is crucial to initiation of BPTI folding. The second native-like contact is important, not in folding initiation, but in preventing a kinetic trap later in the process. Evidence for this comes from mutant G37A, which behaves very differently from the others in displaying a phenomenon called rollover. G37A is, to our knowledge, the first reported case in which a single-site replacement causes rollover, while the wild type and all other known mutants of the same protein show typical two-state chevron plots. The best explanation is that the G37A mutation introduces a kinetic trap of the type described by Chan and Dill [(1998) *Proteins 30,* 2−33]. In native BPTI, there is an unusual polar interaction between the ring of Tyr35 and the backbone NH of Gly37. Our results suggest that the NH−aromatic interaction between residues 37 and 35 is important throughout folding in stabilizing native-like loop conformations and in preventing the flexible loops from being trapped in nonfunctional conformations during later stages of folding.

Good candidates for folding initiation sites are local, native-like interactions detected in the unfolded ensemble of a protein. In reduced and unfolded bovine pancreatic trypsin inhibitor (BPTI),[1] two native-like contacts involving aromatic rings are implied by NMR-detected NOEs (*1*). One is an aromatic−aliphatic contact between side chains of Tyr23 and Ala25. The other, between Gly37 and the Tyr35 ring, is part of an unusual network of interactions involving two NH groups and an aromatic side chain (*2*). Figure 1 shows these and other side chains in the native BPTI. The NOEs between residues 23 and 25, and between residues 35 and 37, are also observed in peptide models of BPTI fragments (*3*), implying that the native-like interactions reported by NOEs in the full-length, reduced BPTI are entirely local and not dependent on the presence of the complete polypeptide chain. A 25−28 turn (residues 25−28) connects the antiparallel strands of the slow exchange core and is observed in the partially folded BPTI; this turn produces the interaction between residues 23 and 25 and has been proposed as a nucleation site for folding (*4*, *5*). In contrast, the interaction between residues 35 and 37 (35−37 interaction) involves a flexible loop and is not otherwise implicated in folding initiation; in a partially folded BPTI, the native-like configuration of the interaction between residues 35 and 37 (part of an NH−aromatic−NH network) is the most minor of three slowly interconverting local conformations (*4*, *5*).

A series of mutants in and around residues 23−25, 35−37, and 45 were constructed, and their folding kinetics and thermodynamics in GdmCl-induced denaturation were characterized. The results support the hypothesis that the interaction between residues 23 and 25 is diagnostic of the 25−28 turn which initiates folding. The 35−37 interaction, however, has a different role; it prevents a kinetic trap at later stages of folding. Kinetic studies reported here were carried out under oxidizing conditions. The disulfide bonds are intact throughout the folding−unfolding reaction, which distinguishes this study from investigations of disulfide-linked folding of BPTI (*6*−*10*).

† This work was supported by NIH Grant GM26242. J.L.B. was supported by NIH Postdoctoral Fellowship CA73104.

* To whom correspondence should be addressed. Telephone: (435) 946-2949. Fax: (435) 946-8251. E-mail: clare@biosci.cbs.umn.edu.

‡ Present address: Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104.

§ Present address: 3M Corporate Analytical Technology Center, 3M Center, Building 201-BS-05, St. Paul, MN 55144.

[1] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; CD, circular dichroism; DSC, differential scanning calorimetry; GdmCl, guanidium chloride; NMR, nuclear magnetic resonance.

10.1021/bi0116947 CCC: $22.00
Published on Web 01/22/2002

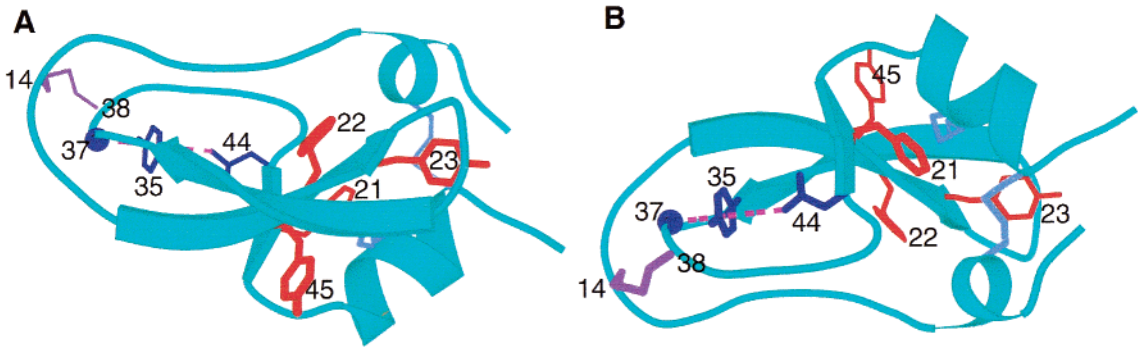

FIGURE 1: Ribbon diagrams of BPTI. The view in panel A is rotated 180° about the *x*-axis relative to the view in panel B. Side chains of relevant residues are shown as stick models and labeled by residue number. Side chains of Tyr21, Tyr22, Tyr23, and Phe45 are red. The three disulfide bonds are purple, and the C14−C38 disulfide bond is labeled and distinguished as darker purple. Side chains of Tyr35 and Asn44 and the backbone amide of residue Gly37 are blue; a pink dashed line represents the polar network of interactions between the NH of residue 37, the ring of residue 35, and the $N_\delta H$ of residue 44. The side chain of Tyr23 is in contact with Ala25 (not labeled) in the turn between the antiparallel strands.

MATERIALS AND METHODS

Proteins and Materials. The wild-type BPTI was purchased from Novo Industries and used without further purification. RM(14−38) was prepared on the basis of the method of Kress and Laskowski (*11*) and purified by reverse-phase HPLC. Other BPTI mutants were expressed from *Escherichia coli* (*12*). The purification procedures were described previously (*13*), although on a larger scale. GdmCl was purchased from GIBCO-BRL (ultrapure grade), and its concentration was determined from the refractive index as described previously (*14*).

Equilibrium CD Measurements. Equilibrium denaturation was monitored by the ellipticity change at 220 nm as a function of GdmCl concentration, using a Jasco J-715 spectropolarimeter fitted with a 0.1 cm water-jacketed cuvette. The concentration of the protein stock solution was determined from the absorbance at 280 nm. Before the measurement, an aliquot from a BPTI stock solution (3−3.5 mg/mL) was diluted 20 times with 20 mM glycine (pH 2.0) containing varying amounts of GdmCl; the sample was incubated at 20 °C for at least 2 h. The data were fit to a two-state transition, and the protein stability in the absence of denaturant, ΔG^0, was obtained by linear extrapolation of free energy changes from the transition region (*14*).

Stopped-Flow Kinetic Measurements. Rapid kinetic experiments were carried out on a BioLogic SFM-3 stopped-flow apparatus attached to a Jasco J-715 CD spectrometer. The dead time of the instrument is ~5 ms. The time course of folding or unfolding was followed by the change in ellipticity at 220 nm with a 5 nm bandwidth. The path length of the observation chamber was 0.2 cm. A good measurement generally requires 50−60 repetitions. To minimize the effects of proline isomerization, the denatured protein solution was incubated for 2 h before the folding reaction was initiated. Both folding and unfolding experiments were carried out at 20 °C in 20 mM glycine with the desired concentration of GdmCl at pH 2.0. For both types of measurements, the final protein concentration was ~0.2 mg/mL.

Kinetic Data Analysis. In folding experiments, at least two kinetic phases were observed. In the unfolding experiment, due to complications related to back flow, it is not possible to follow the reaction long enough to collect data for possible slower phases. For both types of reaction, only the fast phase was fit to a single-exponential function with KaleidaGraph version 3.0.5 (Abelbeck Software).

The kinetic data were fit to a two-state transition as described by Jackson and Fersht (*15*). Briefly, the observed relaxation rate constant, k_{obs}, is

$$k_{obs} = k_f + k_u \tag{1}$$

where k_f and k_u are the folding and unfolding rate constants, respectively, for BPTI mutants at various concentrations of GdmCl. Since the logarithm of k_f and the logarithm of k_u are linearly dependent on the concentration of GdmCl, the chevron plot was fit to the following equation:

$$\ln k_{obs} = \ln(k_f^0 e^{m_f C} + k_u^0 e^{m_u C}) \tag{2}$$

where C is the molar concentration of GdmCl, k_f^0 and k_u^0 are the folding and unfolding rate constants, respectively, in the absence of denaturant, and m_f and m_u describe their GdmCl dependence.

Subsequently, the equilibrium Gibbs free energy change for the folding−unfolding reaction, ΔG^0_{calc}, was calculated from the folding and unfolding rate constants:

$$\Delta G^0_{calc} = -RT \ln \frac{k_u^0}{k_f^0} \tag{3}$$

where R is the gas constant and T is the absolute temperature. The equilibrium m value was calculated from m_f and m_u:

$$m_{calc} = RT(m_u - m_f) \tag{4}$$

NMR Analysis of Temperature Coefficients. Lyophilized WT and G37A BPTI were dissolved in a 90% H_2O/10% D_2O mixture and adjusted to pH 4.5 with HCl. The assignments and complete structure determination of G37A are described in the preceding paper (*16*). The chemical shifts of amide protons were determined from two-dimensional TOCSY experiments carried out at 10 °C intervals from 10 to 80 °C. Spectra were processed with the program nmrPipe (*17*) and analyzed with the program XEASY (*18*).

RESULTS

The GdmCl-dependent folding−unfolding reactions of six BPTI mutants [Y21A, F22A, Y23A, G37A, F45A, and RM-

Table 1: Thermodynamic Parameters[a] of BPTI Mutants at pH 2 and 20 °C, Measured by GdmCl Equilibrium Denaturation Experiments

	RM(14−38)	Y21A	F22A[b]	Y23A	G37A	F45A
C_m (M)	3.77 ± 0.05	5.86 ± 0.12	8.19 ± 0.21	3.53 ± 0.09	5.94 ± 0.05	2.70 ± 0.08
m (kcal mol^{-1} M^{-1})	1.26 ± 0.11	0.88 ± 0.07	0.97 ± 0.04	0.85 ± 0.09	1.06 ± 0.05	0.83 ± 0.08
$\Delta G°$ (kcal/mol)	4.7 ± 0.4	5.1 ± 0.4	8.0 ± 0.3	3.0 ± 0.3	6.3 ± 0.4	2.2 ± 0.3
$\Delta\Delta G°_{WT\rightarrow mut}$ (kcal/mol)[c]	4.3	3.9	1.0	6.0	2.7	6.8
$\Delta\Delta G°_{DSC}$ (kcal/mol)[d]	−	4.8	1.2	5.9	4.9	6.9

[a] C_m is the transition midpoint, in GdmCl molar concentration. $\Delta\Delta G^0_{WT\rightarrow mut}$ is the difference in ΔG^0 between the mutant and WT. $\Delta\Delta G^0_{DSC}$ is the difference in ΔG^0 between the mutant and WT, measured by DSC experiments. [b] The data for F22A are fit assuming that its denatured state has a molar ellipticity of −800 deg cm^2 dmol^{-1}, the same as the other mutants. [c] The difference in ΔG^0 between the mutant and WT is calculated using a value of 9.0 kcal/mol for ΔG^0_{WT}, obtained by DSC at pH 2.0 (*19*, *20*). [d] The stability difference measured by DSC (*20*, *26*).

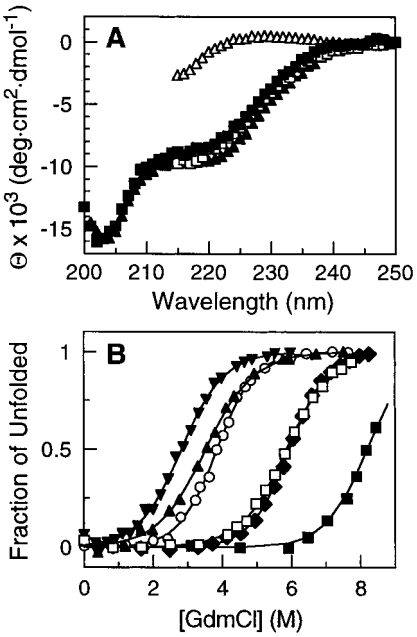

FIGURE 2: Equilibrium denaturation of BPTI mutants. (A) CD spectra of WT and RM(14−38) in the presence and absence of 8.1 M GdmCl at pH 2.0 and 20 °C: WT (■), WT in GdmCl (□), RM(14−38) (▲), and RM(14−38) in GdmCl (△). Other mutants have very similar CD spectra. In the presence of GdmCl, spectra are shown above 215 nm due to the large absorbance of GdmCl below 215 nm. (B) GdmCl equilibrium denaturation of BPTI mutants. The denaturation was followed by ellipticity at 220 nm, and plotted as the fraction unfolded vs GdmCl concentration. From left to right, the curves are for F45A (▼), Y23A (▲), RM(14−38) (○), Y21A (□), G37A (◆), and F22A (■). The curves are fits to a two-state transition. F22A was fit with an assumption that its denatured state has the same molar ellipticity (−800 deg cm^2 dmol^{-1} at 220 nm) as other mutants.

(14−38)] were characterized. The location of the replaced side chains in native BPTI is shown in Figure 1. The high stability of WT BPTI makes it difficult to measure its folding kinetics because unfolding is not detectable even in 8 M GdmCl and at very low pH (Figure 2A). Studies were carried out at pH 2 for several reasons. First, the decreased stability at pH 2 (*19*, *20*) facilitates monitoring the folding reaction. In WT BPTI, there is a stability increase between pH 2 and 4 (*19*), likely because titration of the carboxyl groups around pH 3.5 disrupts the ionic interactions between the N- and C-termini of the protein (*21*). Second, calorimetric studies indicate that BPTI and its mutants have better folding reversibility at pH 2 than at higher pH (*19*, *20*). Finally, an acidic environment ensures that the disulfide bonds are intact throughout the experiments. Folding is monitored by CD rather than by fluorescence, as BPTI has low fluorescence emission from four tyrosines and no tryptophan, and fluorescence intensity is similar for both native and denatured states (*22*).

Equilibrium Denaturation of BPTI Mutants. GdmCl equilibrium unfolding at pH 2.0 and 20 °C (Figure 2B) is ≥90% reversible for all mutants. Table 1 lists the thermodynamic parameters derived from two-state fits to the ellipticity change at 220 nm. Thermodynamic values for F22A were estimated assuming that, like the other mutants, its denatured state has a molar ellipticity of −800 deg cm^2 dmol^{-1} at 220 nm. F45A is the least stable mutant, while F22A is the most stable. The Gibbs free energy change for RM(14−38) is 4.7 kcal/mol, consistent with measured values for similar 14−38 derivatives, C14A/C38A (*23*, *24*) and RCAM(14−38) (*25*). The CD-determined stabilities reported here are comparable to those measured for the same mutants by DSC (*20*, *25*), with the exception of that of G37A (see Table 1). Its ΔG^0 in GdmCl denaturation is 6.3 ± 0.4 kcal/mol, which is 2.2 kcal/mol higher than in the DSC experiment (*26*).

The equilibrium m value is thought to be a measure of the difference in solvent accessible surface area (ASA) between native and denatured states (*27*, *28*). Core mutants Y21A, Y23A, and F45A all have equilibrium m values around 0.85 kcal mol^{-1} M^{-1}, reflecting the similarity in their native and denatured ensembles. A somewhat higher m value is estimated for F22A (Table 1), but the significance of this is clouded because the denatured state baseline is not obtained for F22A, due to its high stability (Figure 2). RM(14−38), with one cleaved and two intact disulfide bonds, has the highest m value (1.26 kcal mol^{-1} M^{-1}). As suggested for other proteins (*28* and references therein), it is likely that the observed increase in the m value for RM(14−38) arises from a decrease in the number of configurational constraints in the denatured state due to removal of a disulfide bond. The m value for RM(14−38) is similar to m values reported for C14A/C38A (*24*) and a number of C30−C51 disulfide mutants (*29*). An intermediate m value of 1.06 kcal mol^{-1} M^{-1} was obtained for G37A. This m value cannot be explained by simply invoking ASA changes, since folding of G37A is not two-state (see below).

Folding Kinetics of BPTI Mutants. In folding experiments, at least two kinetic phases are observed. Typical folding data are shown in Figure 3. In 7.2 M GdmCl, RM(14−38) is fully unfolded, while in 0.8 M GdmCl, it is fully folded. Within less than 150 ms, RM(14−38) recovers ∼85% of the total ellipticity difference between the native and denatured states. Then it slowly evolves to the fully folded state in ∼45 min. A time constant of 19.6 ms was obtained for the fast phase (Figure 3A), the same order of magnitude as the time constant of 40 ms reported for RCAM(14-38) (*22*). We think

Table 2: Folding Kinetic Parameters[a] of BPTI Mutants at pH 2 and 20 °C

	RM(14−38)	Y21A	F22A[b]	Y23A	G37A[c]	F45A
m_f (M^{-1})	−1.20 ± 0.04	−0.92 ± 0.10	−0.81 ± 0.10	−1.04 ± 0.12	−0.73 ± 0.06	−0.91 ± 0.21
m_u (M^{-1})	0.99 ± 0.08	0.57 ± 0.05	−	0.49 ± 0.08	1.65 ± 0.24	0.51 ± 0.08
m_{calc} (kcal mol^{-1} M^{-1})	1.28	0.87	−	0.89	1.38	0.83
k_f^0 (s^{-1})	110 ± 10	16 ± 2	19 ± 2.4	1.4 ± 0.2	48 ± 6	63 ± 5
k_u^0 ($\times 10^{-3}$ s^{-1})	9.26 ± 0.84	2.51 ± 0.85	−	5.76 ± 0.31	0.035 ± 0.006	984 ± 52
ΔG^0_{calc} (kcal/mol)	5.5	5.1	na	3.2	8.2	2.4
ϕ [RM(14−38)][d]	0	0.30	1.03	0.43	0.19	0.05
ϕ (F45A)[e]	−0.07	0.21	0.69	0.37	0.06	0

[a] m_f and m_u are the GdmCl concentration dependencies of the logarithms of folding and unfolding rate constants, respectively. k_f^0 and k_u^0 are the folding and unfolding rate constants in the absence of denaturant, respectively. m_{calc} and ΔG^0_{calc} are the m value and ΔG^0 calculated from kinetic measurements, respectively. ϕ values are calculated using the equation $\phi = \Delta\Delta G^{\ddagger}/\Delta\Delta G^0 = -RT \ln(k^0_{f,WT}/k^0_{f,mut})/\Delta\Delta G^0_{WT\rightarrow mut}$, where k_f^0 is the folding rate in the absence of denaturant. [b] The folding kinetics of F22A were fit to $\ln k_{obs} = \ln k_f^0 + m_f C$ (see the text). [c] The data for G37A were fit to eq 2 ignoring the rollover part of the chevron plot. [d] Assume that the folding rate constant for WT is the same as that for RM(14−38). [e] Assume that the folding rate constant for WT is the same as that for F45A.

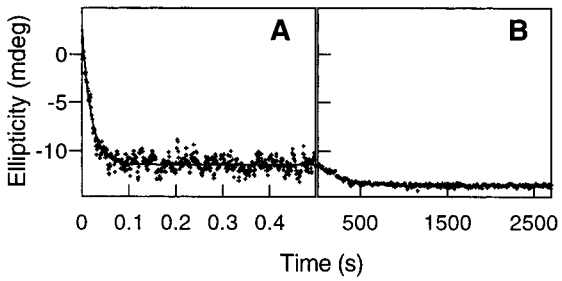

FIGURE 3: Folding of RM(14−38) at pH 2.0 and 20 °C. The initial and final concentrations of GdmCl are 7.2 and 0.8 M, respectively. (A) The fast phase was fit to a single-exponential function with a τ of 19.6 ms, which is consistent with earlier studies on folding of RCAM(14−38) (*22*). (B) The slower phase is attributed to proline isomerization and was not fit.

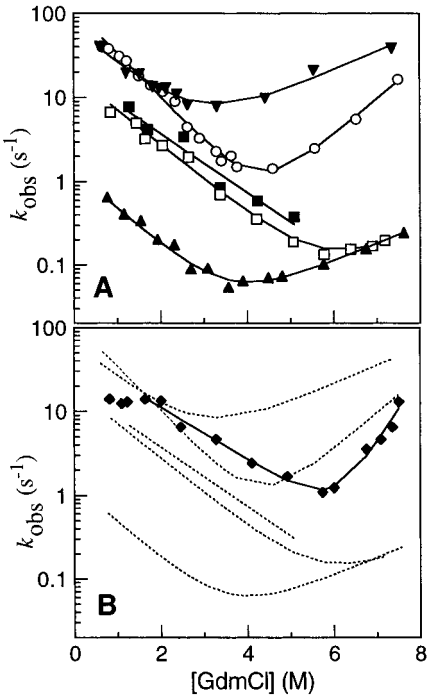

FIGURE 4: Dependence of the folding and unfolding rate constants of BPTI mutants on GdmCl concentration, at pH 2.0 and 20 °C. (A) Chevron plots of mutants other than G37A. Curves are two-state fits for, from top to bottom, F45A (▼), RM(14−38) (○), F22A (■), Y21A (□), and Y23A (▲). (B) Rollover in G37A (◆). The data points above 1.8 M GdmCl were fit with eq 2 (solid line). The dotted lines are curves for the other mutants shown in panel A.

that the fast kinetic phase reflects the direct folding reaction. Due to technical limitations at the time in following fast kinetic phases, earlier studies of denaturant-dependent folding of BPTI disulfide mutants focused on slower kinetic phases attributed to proline isomerization (*22*, *23*). The slow folding phases observed here are likewise taken to arise from proline isomerization. Three points support this. First, these slow phases account for 15−20% of the total amplitude change, consistent with earlier measurements. Second, although we did not obtain an accurate time constant(s) to describe slow folding phases because of the uncertainty in the number of the slow phases, our overall time is consistent with earlier measurements. Third, the slow phases do not change noticeably with different mutations. A burst phase is not observed within the dead time, and we can account for all the signal change during the reaction (Figure 3). This is not surprising because the denatured states of mutants studied here have two or three disulfide bonds and are more compact than fully extended denatured states. In *unfolding* experiments, the large density difference between the solutions before and after mixing causes a back-flow problem in the stopped-flow instrument ~2−3 min following mixing (data not shown). Thus, it is not possible to follow unfolding long enough to collect data for slower unfolding phases, if there are any. Therefore, we focused on the fast step in both folding and unfolding experiments, and on the mutational effects on this phase.

For both folding and unfolding, the fast phase fits well to a single-exponential decay process at all GdmCl concentrations. The effect of GdmCl concentration on the observed rate constant, k_{obs}, is summarized in the chevron plots of Figure 4. The chevron plots are fit to eq 2, and Table 2 lists the calculated kinetic parameters. The folding rate constants of F22A were derived from fits of the data to eq 2 because the unfolding rate is negligible at GdmCl concentrations of <6 M. In the absence of GdmCl, RM(14−38) has the fastest folding rate, k_f^0, followed by F45A. F22A folds ~1 order of magnitude slower than F45A and slightly faster than Y21A. The folding rate is slowest for Y23A, and on the same order as the slowest folding times reported for simple two-state proteins, e.g., acylphosphatase (*30*), the SH3 domain of PI3 kinase (*31*), and the histidine-containing phosphocarrier protein HPr (*32*).

F45A has the highest unfolding rates. At high GdmCl concentrations, both RM(14−38) and G37A unfold at least 1 order of magnitude faster than Y21A and Y23A, which have the slowest unfolding rates. However, the unfolding rate constant in the absence of denaturant, k_u^0, for RM(14−38) is comparable to those of Y21A and Y23A, mainly because the k_u of RM(14−38) has a higher dependence on GdmCl, that is, a higher m_u value (Table 2).

Estimation of ϕ Values. Values of the ϕ parameter (*33*−*35*) are often taken as indicators of the critical involvement of individual residues, particularly their side chains, in the folding process. Since the calculation of ϕ requires knowledge of $k_{f,WT}^0$, the folding rate constant for the WT protein in the absence of GdmCl, it is not feasible to obtain precise ϕ values for the BPTI mutants characterized here. Nevertheless, with an assumption that mutant RM(14−38) closely resembles WT in folding kinetics, we can estimate ϕ values for the other mutants by substituting $k_{f,WT}^0$ with $k_{f,RM(14-38)}^0$ (Table 2). Of these, F22A has the highest value (∼1). Although ϕ values of Y21A and Y23A are less than 1, they are significantly higher than those of RM(14−38) and F45A. The G37A ϕ value is difficult to assess since this mutant does not fold via a two-state mechanism. We also estimated the ϕ values with $k_{f,F45A}^0$ in place of $k_{f,WT}^0$, and obtained similar results (Table 2). F22A has the highest ϕ value, closely followed by Y23A and Y21A; F45A and RM(14−38) have ϕ values close to 0.

All Mutants Except G37A Fold by an Apparent Two-State Mechanism. As expected for two-state folding, the values of ΔG_{calc}^0, the free energy change derived from kinetic parameters, and ΔG^0, the free energy change measured from equilibrium denaturation experiments, are similar for mutants RM(14−38), Y21A, Y23A, and F45A. The average difference between ΔG_{calc}^0 and ΔG^0 for these mutants is 0.3 kcal/mol, with the largest difference being 0.8 kcal/mol for RM(14−38). Further, the m_{calc} value (eq 4) for each of these mutants also equals the equilibrium *m* value, consistent with two-state folding. The average difference between m_{calc} and equilibrium *m* values is 0.02 kcal mol^{-1} M^{-1}, with the largest difference being 0.04 kcal mol^{-1} M^{-1} for Y23A. For F22A, the comparison of ΔG_{calc}^0 to ΔG^0 cannot be made since only the folding part of the chevron plot is measurable, but it appears that F22A has an m_f value similar to those of Y21A and Y23A, and is likely to fold by a two-state mechanism.

Contrary to the other mutants, G37A does not fold by a two-state mechanism. Values for ΔG_{calc}^0 and m_{calc}, obtained from fits to eq 2 of rate constants at >1.8 M GdmCl, are 8.2 kcal/mol and 1.38 kcal mol^{-1} M^{-1}, respectively (Table 2). In comparison, ΔG^0 and *m* values for G37A derived from equilibrium experiments are 6.3 kcal/mol and 1.06 kcal mol^{-1} M^{-1}, respectively (Table 1). The differences between kinetic and thermodynamic measurements, 1.9 kcal/mol for ΔG^0 and 0.32 kcal mol^{-1} M^{-1} for *m*, are much greater than for other mutants (ΔG_{calc}^0 and m_{calc} obtained from fits to the G37A chevron plot over the entire GdmCl concentration range are even larger, resulting in a greater difference between the kinetic and thermodynamic measurements). Further, the rate constants are relatively independent of denaturant concentration below 2 M GdmCl (Figure 4B), a phenomenon termed rollover (*36*). We do not observe rollover on the *unfolding* arm of the chevron plot for G37A; this could be due to its high stability. As far as we are aware, this is the first case of rollover being introduced by a single-amino acid replacement in a protein whose WT variant and other single-site mutants show standard two-state folding behavior.

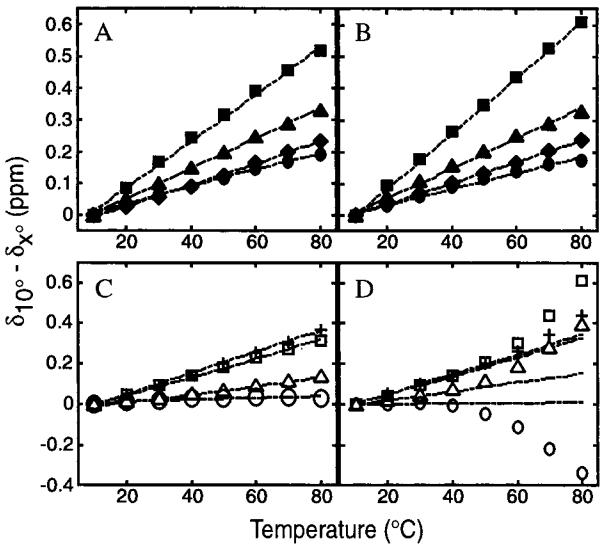

FIGURE 5: Plots of chemical shift vs temperature for WT and G37A HN protons. Representative HN protons that have linear plots for both WT (A) and G37A (B) are shown for residues 5 (●), 21 (◆), 25 (■), and 50 (▲). Representative protons that have linear plots for WT (C) but nonlinear plots for G37A (D) are shown for residues 10 (+), 33 (△), 36 (□), and 43 (○). Dotted lines are linear fits of the data, except in panel D, which is a linear fit only for the temperatures from 10 to 30 °C.

There are cautionary indications that rollover can result from transient aggregation of the denatured protein during folding (*37*, *38*). A way to check for such a possibility is to monitor the same folding kinetics at lower protein concentration. G37A folding was assessed at two protein concentrations (30 and 60 μM); rollover was observed at both (data not shown). The protein concentrations used for G37A are comparable to those for CI2 at which no transient aggregation is observed (*38*). We did not follow G37A folding kinetics at protein concentrations lower than 30 μM due to the requirements of the stopped-flow CD experiment.

HN Temperature Coefficients Exhibit Noncooperative Structural Transitions in G37A. The rollover behavior of G37A is intriguing since the other mutants apparently fold by a two-state mechanism, and since G37A has the same average structure as WT (*16*). To explore the structural basis for the deviation from two-state folding for G37A, a detailed NMR investigation of the anomalous changes in amide proton (HN) chemical shift as a function of temperature was carried out. In NMR spectra of G37A, many amide protons are broadened beyond detection at higher temperatures (>50 °C). Plots of chemical shift versus temperature (Figure 5) show that many amide NH have large deviations from linearity, especially at high temperatures. Normally, over the temperature range where a protein is folded, plots of HN chemical shift versus temperature are linear with a slope (temperature coefficient) qualitatively proportional to the strength of hydrogen bonds (*39*). Changes in chemical shift are upfield with increasing temperature, and are thought to arise from thermal fluctuations that slightly increase the average hydrogen bond distance and hence HN chemical shift. When a protein and/or peptide enters a folding−

unfolding transition, the temperature behavior may become nonlinear, since the chemical shift becomes a weighted average of the folded and unfolded states. G37A exhibits very interesting anomalous behavior in that portions of the molecule show large deviations from linearity (Figure 5D), while other regions are unaffected and behave like WT (Figure 5B). Since not all resonances show the altered chemical shift behavior, it is unlikely to be from global unfolding. Additional lines of evidence further strongly suggest that the nonlinear dependence of the chemical shift for some, but not all, of the NH groups is not global unfolding. The transition temperatures are not consistent with DSC results obtained for G37A at pH 2 (*26*). At 80 °C, amides in the core of the protein with linear temperature coefficients have slow exchange cross-peaks at 80 °C to protons with chemical shifts consistent with the unfolded form, implying that the global unfolding transition begins to occur at this temperature (data not shown).

DISCUSSION

Although BPTI is a well-studied model protein, the folding−unfolding kinetics of oxidized form, with all disulfides intact, are not characterized due to the very fast folding associated with its high stability. To circumvent this, we chose variants that have large stability changes, and minimal structural deviation from WT. The six BPTI mutations are in and around either the central β-sheet (Y21A, F22A, Y23A, and F45A) or the active site loops [RM(14−38) and G37A]. The mutated side chains are shown in Figure 1. Residues 21−23 are in the folding core. The side chain of Tyr23 is involved in a contact between residues 23 and 25, one of only two native-like contacts observed in the reduced and unfolded BPTI. Phe45 is located on the β-bridge; its side chain interacts with the C-terminal α-helix. RM(14−38) lacks the C14−C38 disulfide bond linking the flexible loops. Gly37 is located on the flexible loop between the second β-strand and the β-bridge, close to the C14−C38 disulfide bond. Its amide group is involved in an unusual set of polar interactions, an NH−aromatic−NH network composed of the NH of residue 37, the ring of residue 35, and the $N_\delta H$ of residue 44. G37A was constructed to perturb the 37 NH−35 ring part of the network, as the ϕ and ψ backbone angles for residues 36 and 37 required by the geometry of this interaction are energetically favorable only for glycine (in WT, residues 36 and 37 are glycines).

All six mutants and WT have very similar structures, the availability of which allows us to analyze structural implications of the folding results. In the crystal structures of Y21A, F22A, Y23A, and F45A, the average rearrangement of atoms around the mutation site is less than 0.5 Å (*40*). In the recently determined solution structure of G37A, the average structure is essentially the same as WT, although local changes in internal motions are apparent (*16*). While a detailed structure of RM(14−38) is not available, the cleavage of this disulfide bond results in no significant structural changes (*41*). Judged by the chemical shift perturbation (*42*), the NH−aromatic interaction between residues 37 and 35 is largely preserved in RM(14−38).

Core Mutants Have the Greatest Effect on the Rate-Limiting Step in Folding. The kinetic parameters for the direct folding step of BPTI mutants at pH 2.0 and 20 °C are summarized in Table 2. The largest decreases in folding rates are observed for Y23A, followed by Y21A and F22A. Folding rates of Y23A are comparable to some of the slowest reported for simple two-state proteins (*30*−*32*). Residues 21−23 are on the first strand of the central β-sheet. The side chain of Tyr21 and Phe22 interact with residues from the second β-strand, while the side chain of Tyr23 apparently interacts with the side chain of turn residue Ala25. Our results support the hypothesis that local interactions favoring formation of the β-turn hairpin (strand of residues 18−24, turn of residues 25−28, and strand of residues 29−35) are important in folding initiation (*1*). In WT, the side chain of Tyr23 is in the proximity of the backbone and side chain of Ala25 (*43*). This contact, indicative of the presence of the turn, is also indicated by NOEs in NMR spectra of the reduced BPTI (*1*) and BPTI model peptides (*3*), implying that the turn is significantly populated in the globally unfolded conformational ensemble. Consistently, $[14-38]_{Abu}$, a BPTI variant retaining only the C14−C38 disulfide and aminobutyric acid in place of the other cysteines, and which is thought to model early folding intermediates, also has strong, native-like NOEs between residues 23 and 25 (*4*). The Y23A mutation perturbs the strand−turn contacts, and slows folding by 100-fold compared to F45A and RM(14−38) (Figure 4), both of which show the smallest effects on kinetics relative to WT.

In WT, Phe45 is in the β-bridge, not the central β-sheet. F45A has a hydrophobic cavity corresponding to the atoms lost in the Phe → Ala substitution (*40*). Although F45A is the most destabilized of our panel of mutants, it folds and unfolds rapidly.

G37A Introduces a Kinetic Trap(s) into the Energy Landscape of Folding. Rollover, a deviation from the standard chevron plot at low denaturant concentrations, is generally interpreted as indicative of a folding intermediate (*36*). In lysozyme, rollover is attributed to a "fast track" for folding (*44*). In cytochrome *c*, it is attributed to nonspecific molecular collapse under both strongly folding and unfolding conditions (*45*). In barnase, it is attributed to the presence of an off-pathway intermediate (*46*). In RNase H*, it is attributed to the presence of a burst-phase A-state-like intermediate (*47*). In ubiquitin, it is proposed to arise from an on-pathway intermediate (*48*), although disagreements have been raised (*49*). An alternative explanation, proposed by Chan and Dill (*50*) in terms of folding on a funnel-shaped energy landscape, seems more relevant to our results. The data reported herein, and other information about Gly37, are most consistent with the Chan and Dill model, in which rollover is explained as a kinetic trap(s). Described as a dent on a rugged energy landscape, a kinetic trap has a local energy minimum. Although the conformation favored in the kinetic trap is close to the native state, it is different in that a molecule that has fallen into the trap must cross significant energy barriers before continuing a downhill trajectory to the native state. Normally, kinetic traps slow folding.

The rollover observed for G37A is, to our knowledge, the first induced by a single-site mutation in a protein that gives normal chevron plots for WT and other mutants. The placement, in one of the large flexible loops at the active site, perturbs an unusual NH−aromatic−NH interaction. The flexible active site loops in WT can undergo large structural rearrangements illustrated, for example, in Y35G (*13*). The C14−C38 disulfide cross-link connecting the loops is highly

flexible in WT (*51*). G37A retains the 37 NH−35 ring polar interaction and a native-like, average backbone conformation at residues 36 and 37 (*16*). However, in G37A, the geometry required for the interaction, with the NH centered over and pointing to the middle of the ring, requires a strained configuration of the backbone of residues 36 and 37. This strain is likely reflected in the high level of destabilization of G37A relative to WT, and in increased mobility of segments of the protein in the vicinity of the mutation (*16*). Apparently, a kinetic trap is introduced when the 37 NH−35 ring interaction is less favorable, as in G37A, but not when the C14−C38 disulfide is lost as in RM(14−38). Chan and Dill suggest that kinetic traps result when the protein is close to native, but not fully native (*50*). It is easily envisioned that, during later stages of folding, G37A samples more non-native conformations in which the core is essentially packed, aperiodic loops are collapsed, and the backbone angles for Gly36 and Ala37 are unstrained, but the geometry required for the 37 NH−35 ring interaction is not favorable. These conformations would be trapped until energy barriers to rearrangement to native-like packing are overcome. In trapped conformations, the active site loop probably does not have a functional structure that can bind tightly to trypsin without being cleaved by the enzyme. That is, the absence of a kinetic trap may be functionally important for this inhibitor.

Our proposal that the conformations trapped in G37A are in species with a native-like core and alternately packed loops is consistent with data from partially folded $[14-38]_{Abu}$, in which alternative conformations of the *same* NH are shown to interconvert slowly (approximately milliseconds), that is, through relatively high energy barriers (*4*, *5*, *52*). The slowly interconverting conformations in partially folded $[14-38]_{Abu}$ are indicated by the presence of more than one "slow exchange" NMR cross-peak for each backbone NH. Most NH groups throughout the molecule have two conformations in slow exchange, but Gly37 has three (*4*). The unusual upfield chemical shift of one of the peaks representing alternative conformations of Gly37 indicates that it arises from the native-like 37 NH−35 ring interaction. Significantly, this native-like conformation is the least populated, demonstrating that in partially folded $[14-38]_{Abu}$, an alternative non-native, and ordered, conformation is favored (in this case, conformation means a family of rapidly interconverting conformers). This conformation could readily trap molecules in later steps in folding if favorability of the interaction between residues 37 and 35 is significantly diminished, as it is in G37A.

Another line of evidence indicates that the conformations trapped during G37A folding are in species with a native-like core and alternately packed loops. Apparently, the structure of G37A in the vicinity of the mutation can become locally disordered at temperatures well below those for global thermal unfolding. Amides exhibiting detectable nonlinearity in chemical shift versus temperature plots are mapped onto the backbone of BPTI in Figure 6. They cluster on the side of the protein surrounding the site of mutation, consistent with a local structural perturbation. Apparently, G37A undergoes local or noncooperative unfolding above 50 °C, which is well below the temperature required for global unfolding. The data are consistent with a collective change, above 50 °C, in the average structure of the ensemble of conformations sampled by the backbone in the vicinity of the mutation. Other data (*16*) show that although the average structure at 30 °C is not detectably changed by the G37A replacement, there is a local increase in internal mobility around the mutation site. Since the direction of the chemical shift changes is toward random coil, a loss of tertiary contacts with increasing temperature most likely occurs. Regardless of the specific nature of the structural changes, it is clear that G37A has unusual nonglobal transitions in and around the mutation site at higher temperatures. Also consistent with the notion of kinetic traps involving alternatively arranged loops is the observation that, when the C14−C38 disulfide of Y35G is reduced, it does not re-form when the reducing agent is removed, as does WT (*42*).

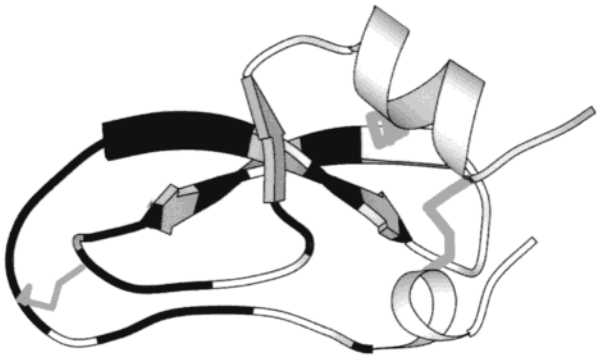

FIGURE 6: Location of HN protons with detectable deviation from linearity in plots of chemical shift vs temperature ($R < 0.99$ in Figure 5). Residues 7, 10−12, 14−20, 22, 24, 31, 33, 35, 36, 38−40, 43, and 44 are shown in black on a ribbon diagram of WT.

In conclusion, our data indicate that late in the folding of G37A there is a kinetic trap of accumulated conformations that have an ordered core and locally rearranged loops.

Implications for WT Folding. Although it is not presently feasible to examine the folding kinetics of WT directly, comparison of mutant behaviors suggests that WT folds by a two-state mechanism, consistent with earlier calorimetric studies (*19*). We take the data presented here, along with previous reports, to indicate that formation of the central β-sheet and hairpin turn is an early, rate-limiting step. The local interaction between residues 23 and 25, a native contact resulting from β-hairpin formation, is detected in the denatured BPTI. The favorability of this turn even in the unfolded chain permits rapid initiation of folding to the strand−turn−strand hairpin. The packing of the small β-bridge to the central β-sheet rapidly follows formation of the central strand−turn−strand structure. Acquisition of the native conformation of the flexible loops, away from the protein core, occurs in later steps. The G37 NH−Y35 aromatic−N44 NH polar network prevents the protein from falling into kinetic traps that slow folding.

In summary, we have investigated the roles in folding of two native-like interactions implied by NOEs observed in unfolded BPTI. Both are important for folding, but at different stages. The local interaction between side chains of Tyr23 and Ala25 is diagnostic for formation of the β-hairpin of residues 18−35, the rate-limiting folding step. The 37 amide−35 ring, NH−aromatic polar interactions are important in stabilizing native-like loop structure and in preventing kinetically trapped, nonfunctional conformations.

REFERENCES

1. Pan, H., Barbar, E., Barany, G., and Woodward, C. (1995) *Biochemistry 34*, 13974−13981.
2. Tüchsen, E., and Woodward, C. (1987) *Biochemistry 26*, 1918−1925.
3. Kemmink, J., and Creighton, T. E. (1993) *J. Mol. Biol. 234*, 861−878.
4. Barbar, E., Barany, G., and Woodward, C. (1995) *Biochemistry 34*, 1142−1134.
5. Barbar, E., Hare, M., Daragan, V., Barany, G., and Woodward, C. (1998) *Biochemistry 37*, 7822−7833.
6. Creighton, T. E. (1977) *J. Mol. Biol. 113*, 275−293.
7. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386−1393.
8. Dadlez, M., and Kim, P. S. (1996) *Biochemistry 35*, 16153−16164.
9. Zhang, J. X., and Goldenberg, D. P. (1997) *Protein Sci. 6*, 1563−1576.
10. Bulaj, G., and Goldenberg, D. P. (2001) *Nat. Struct. Biol. 8*, 326−330.
11. Kress, L. F., and Laskowski, M. (1967) *J. Biol. Chem. 242*, 4925−4929.
12. Goldenberg, D. P. (1988) *Biochemistry 27*, 2481−2489.
13. Housset, D., Kim, K.-S., Fuchs, J., Woodward, C., and Wlodawer, A. (1991) *J. Mol. Biol. 220*, 757−770.
14. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) in *Protein structure: a practical approach* (Creighton, T. E., Ed.) pp 311−330, IRL Press, Oxford, U.K.
15. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10428−10435.
16. Battiste, J. L., Li, R., and Woodward, C. (2002) *Biochemistry 41*, 2237−2245.
17. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
18. Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.
19. Makhatadze, G. I., Kim, K. S., Woodward, C., and Privalov, P. L. (1993) *Protein Sci. 2*, 2028−2036.
20. Kim, K. S., Tao, F., Fuchs, J., Danishefsky, A. T., Housset, D., Wlodawer, A., and Woodward, C. (1993) *Protein Sci. 2*, 588−596.
21. Tüchsen, E., and Woodward, C. (1985) *J. Mol. Biol. 185*, 405−419.
22. Jullien, M., and Baldwin, R. L. (1981) *J. Mol. Biol. 145*, 265−280.
23. Hurle, M. R., Marks, C. B., Kosen, P. A., Anderson, S., and Kuntz, I. D. (1990) *Biochemistry 29*, 4410−4419.
24. Ma, L.-C., and Anderson, S. (1997) *Biochemistry 36*, 3728−3736.
25. Schwarz, H., Hinz, H.-J., Mehlich, A., Tschesche, H., and Wenzel, H. R. (1987) *Biochemistry 26*, 3544−3551.
26. Kim, K. S., Fuchs, J. A., and Woodward, C. K. (1993) *Biochemistry 32*, 9600−9608.
27. Shortle, D., Chan, H. S., and Dill, K. A. (1992) *Protein Sci. 1*, 201−215.
28. Shortle, D. (1995) *Adv. Protein Chem. 46*, 217−247.
29. Liu, Y., Breslauer, K., and Anderson, S. (1997) *Biochemistry 36*, 5323−5335.
30. van Nuland, N. A., Chiti, F., Taddei, N., Raugei, G., Ramponi, G., and Dobson, C. M. (1998) *J. Mol. Biol. 283*, 883−891.
31. Guijiarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D., and Dobson, C. M. (1998) *J. Mol. Biol. 276*, 657−667.
32. Van Nuland, N. A., Meijberg, W., Warner, J., Forge, V., Scheek, R. M., Robillard, G. T., and Dobson, C. M. (1998) *Biochemistry 37*, 622−637.
33. Serrano, L., Matouschek, A., and Fersht, A. R. (1992) *J. Mol. Biol. 224*, 805−818.
34. Fersht, A. R. (1993) *FEBS Lett. 325*, 5−16.
35. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) *J. Mol. Biol. 254*, 260−288.
36. Baldwin, R. L. (1996) *Folding Des. 1*, R1−R8.
37. Silow, M., and Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
38. Silow, M., Tan, Y.-J., Fersht, A. R., and Oliveberg, M. (1999) *Biochemistry 38*, 13006−13012.
39. Baxter, N. J., and Williamson, M. P. (1997) *J. Biomol. NMR 9*, 359−369.
40. Danishefsky, A. T., Housset, D., Kim, K.-S., Tao, F., Fuchs, J., Woodward, C., and Wlodawer, A. (1993) *Protein Sci. 2*, 577−587.
41. Stassinopoulou, C. I., Wagner, G., and Wüthrich, K. (1984) *Eur. J. Biochem. 145*, 423−430.
42. Beeser, S. A., Oas, T. G., and Goldenberg, D. P. (1998) *J. Mol. Biol. 284*, 1581−1596.
43. Berndt, K. D., Guntert, P., Orbons, L. P. M., and Wüthrich, K. (1992) *J. Mol. Biol. 227*, 757−775.
44. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9029−9033.
45. Sosnick, T. R., Mayne, L., and Englander, S. W. (1996) *Proteins: Struct., Funct., Genet. 24*, 413−426.
46. Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) *Nature 346*, 440−445.
47. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.
48. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) *Nat. Struct. Biol. 3*, 193−205.
49. Krantz, B. A., and Sosnick, T. R. (2000) *Biochemistry 39*, 11696−11701.
50. Chan, H. S., and Dill, K. A. (1998) *Proteins: Struct., Funct., Genet. 30*, 2−33.
51. Otting, G., Liepinsh, E., and Wüthrich, K. (1993) *Biochemistry 32*, 3571−3582.
52. Barbar, E., Hare, M., Makokha, M., Barany, G., and Woodward, C. (2001) *Biochemistry 40*, 9734−9742.

BI0116947