Native-like Conformations Are Sampled by Partially Folded and Disordered Variants of Bovine Pancreatic Trypsin Inhibitor†

Judit Tulla-Puche,‡ Irina V. Getun,‡ Clare Woodward,*,§ and George Barany*,‡

Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, Minnesota 55455, and Department of Biochemistry, Molecular Biology, and Biophysics, 1479 Gortner Avenue, University of Minnesota, St. Paul, Minnesota 55108

Received July 23, 2003; Revised Manuscript Received October 25, 2003

ABSTRACT: Partially folded conformational ensembles of bovine pancreatic trypsin inhibitor (BPTI) are accessed by replacing Cys 5, 30, 51, and 55 by α-amino-*n*-butyric acid (Abu) while retaining the disulfide between Cys 14 and 38; the resultant variant is termed $[14-38]_{Abu}$. Two new analogues with modifications in the β-turn, P26D27$[14-38]_{Abu}$ and N26G27K28$[14-38]_{Abu}$, are compared to partially folded $[14-38]_{Abu}$, as well as to $[R]_{Abu}$, the unfolded protein with all six Cys residues replaced by Abu. Structural features of the new analogues of $[14-38]_{Abu}$ have been determined by circular dichroism (CD), one-dimensional ^{1}H NMR, and 8-anilino-1-naphthalenesulfonic acid (ANS) fluorescence experiments. Both analogues are more disordered than the parent $[14-38]_{Abu}$, but while P26D27$[14-38]_{Abu}$ has a small population of native-like conformations observed by NMR, no ordered structure is detected for N26G27K28-$[14-38]_{Abu}$. Trypsin inhibition assays were carried out using a modified rat trypsin, C191A/C220A, that minimizes cleavage of unfolded peptides. Both $[14-38]_{Abu}$ and P26D27$[14-38]_{Abu}$ significantly inhibit modified trypsin. N26G27K28$[14-38]_{Abu}$ has low but measurable inhibitor activity, while $[R]_{Abu}$ has no activity even when in very high molar excess relative to trypsin. ANS fluorescence is enhanced by $[14-38]_{Abu}$ and by both variants but not by $[R]_{Abu}$. We conclude that partially folded ensembles of BPTI, even those with little or no CD- or NMR-detectable structure, contain minor populations of native-like conformations. Partially folded $[14-38]_{Abu}$ and both variants, as well as $[R]_{Abu}$, have enhanced negative ellipticity in CD spectra acquired in the presence of the osmolyte trimethylamine *N*-oxide (TMAO). TMAO-induced structure is formed cooperatively, as indicated by thermal unfolding curves. Inhibitor activity as a function of TMAO concentration implies that the osmolyte-induced structure is native-like for $[14-38]_{Abu}$ and P26D27$[14-38]_{Abu}$ and is probably native-like for N26G27K28$[14-38]_{Abu}$. $[R]_{Abu}$ also shows increased CD-detected structure in the presence of TMAO, but such structure is likely to be collapsed and non-native.

Partially folded proteins are postulated as models of metastable forms transiently populated during folding. Their properties may also be relevant to understanding natively denatured proteins (*1*, *2*) and/or pathogenic, aggregating proteins (*3*). We have developed a system to access partially folded analogues of bovine pancreatic trypsin inhibitor (BPTI)[1] by maintaining the disulfide bridge between Cys 14 and Cys 38 but replacing the remaining four Cys by α-amino-*n*-butyric acid (Abu); the parent protein is called $[14-38]_{Abu}$. The fully unfolded analogue, $[R]_{Abu}$, has all Cys residues replaced by Abu (*4*, *5*).

The ensemble structure and dynamics of partially folded $[14-38]_{Abu}$ and its variants, and of fully unfolded $[R]_{Abu}$, form the basis of our working model of protein folding (*6*, *7*), which is based on the concepts of folding by multiple parallel pathways in a funnel-shaped energy landscape (*8*). During folding, conformational ensembles with varying degrees of order are populated. At every stage, even when the average structure is highly disordered, the partially folded ensemble is an equilibrating mixture of disordered conformations and more ordered conformations that sample native-like structure, especially in core regions. Core regions of partially folded and unfolded proteins are sequences that in the native protein compose core elements, the elements of mutually packed secondary structure that contain the slowest exchanging backbone amide groups (*9*). There are site-specific energy barriers between more ordered and more disordered conformations. At early stages of folding, conformations with native-like structure are minor populations. As folding proceeds, more native-like structure develops in the core while the rest of the molecule favors more disordered species. Late folding steps involve packing noncore elements.

† This work was supported by NIH Grant GM 51628 (to G.B. and C.W.).

* To whom correspondence should be addressed. C.W.: phone, 435-946-2949; e-mail, clare@cbs.umn.edu. G.B.: phone, 612-625-1028; fax, 612-626-7541; e-mail, barany@umn.edu.

‡ Department of Chemistry.

§ Department of Biochemistry, Molecular Biology, and Biophysics.

[1] Abbreviations: Abu, α-amino-*n*-butyric acid; ANS, 8-anilino-1-naphthalenesulfonic acid; BPTI, bovine pancreatic trypsin inhibitor; CD, circular dichroism; Fmoc, 9-fluorenylmethoxycarbonyl; HPLC, high-performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time of flight (mass spectrometry); NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PAC, peptide acid linker; PEG−PS, poly(ethylene glycol)−polystyrene (graft resin support); SBTI, soybean trypsin inhibitor; TMAO, trimethylamine *N*-oxide; t_R, retention time. Amino acid symbols denote the L-configuration.

10.1021/bi035301a CCC: $27.50
Published on Web 01/23/2004

Before onset of folding, the conformational ensembles of fully unfolded protein have some detectable order in core elements (*7*, *10*), but this order is not necessarily native-like.

The core β-sheet (18−35) in native BPTI is flanked by sequences that form the trypsin binding loops connected by the 14−38 disulfide. Residues 11−18 interact with trypsin and include the specificity group of Lys 15, while the 34−39 loop provides a key contact between the main chain carboxyl group of Gly 36 in the inhibitor and His 57 at the trypsin active site. The partially folded ensemble of $[14-38]_{Abu}$ is, on average, more native-like in core regions and far more disordered in regions that make up the trypsin binding loops in native BPTI.

As a probe of our model that populations of native-like conformations are in equilibrium with more disordered conformations during folding, we have measured inhibitor activity of partially folded and unfolded BPTI analogues. For these experiments, two new analogues of $[14-38]_{Abu}$ were constructed with perturbations in the hairpin turn. Both new analogues are more disordered than the parent $[14-38]_{Abu}$. Inhibitor activity is observed for $[14-38]_{Abu}$ and both of its analogues, including one with no structure evident by CD or NMR.

The two new analogues of $[14-38]_{Abu}$, as well as parent $[14-38]_{Abu}$ and control $[R]_{Abu}$, were examined for osmolyte-induced structure, using trimethylamine *N*-oxide (TMAO). Organic osmolytes are reported to protect proteins from denaturation, induce folding in otherwise destabilized proteins, and restore biological activity of partially folded proteins (*11*−*14*). Some of these osmolyte effects are observed for the BPTI system of partially folded and unfolded proteins.

EXPERIMENTAL PROCEDURES

Materials, solvents, instrumentation, and general methods, as well as procedures for peptide synthesis, CD, NMR, trypsin expression, and trypsin inhibition assays, followed the literature from our laboratory (*4*, *5*, *15*−*17*) or others (*18*−*22*). Brief overviews follow below, and full experimental descriptions are provided in the Supporting Information.

Variants of BPTI were prepared by automated Fmoc solid-phase synthesis, followed by appropriate purification, as has been described previously (*4*, *5*, *15*−*17*); reversed-phase analytical HPLC and MALDI-TOF, respectively, confirmed that final products were pure (>97%) and had the expected structures. CD spectra were acquired on a JASCO 710 spectropolarimeter with samples at 1 °C in 10 mM phosphate buffer, pH 4.6, and at protein concentrations of 20 μM for far-UV spectra and 100 μM for near-UV spectra. For CD spectra of samples in the presence of TMAO, the final concentration of osmolyte was 1.0 or 2.0 M. NMR spectra were determined using Varian 600 MHz Inova instruments with samples at 1 °C, in 50 mM deuterated sodium acetate buffer in 90:10 $H_2O/^2H_2O$, at pH 4.60, and at a final protein concentration of 4 mM. Fluorescence emission spectra were measured on a Perkin-Elmer LS-50B spectrofluorometer. 8-Anilino-1-naphthalenesulfonic acid (ANS) samples contained 1.2 mM ANS in 10 mM phosphate buffer at pH 4.6.

Recombinant rat trypsinogen II was produced as a fusion protein in a *Saccharomyces cerevisiae* expression system (*18*, *19*). The C191A/C220A trypsin mutant was obtained by activating the corresponding trypsinogen with 1 μL of enterokinase (Biozyme, activity 300 units/μL). Assay mixtures (0.4 mL) contained 100 mM NaCl, 10 mM $CaCl_2$, and 50 mM Hepes, pH 8.0, plus 0.1% Triton X-100, to stabilize the enzyme, plus the substrate Z-Lys-SBzl (100 μM to 1 mM, depending on the BPTI analogue) plus 4,4′-dithiodipyridine (25 μM) (*19*). Before the assay was carried out, C191A/C220A trypsin (0.2 μM) and the BPTI variant (0.2−10 μM) were preincubated together at 4 °C for 1−3 h. Enzyme reactions were initiated by adding 20 μL aliquots to the assay mixture (0.4 mL) at 25 °C. Hydrolysis of Z-Lys-SBzl was monitored spectrophotometrically at 324 nm (*23*, *24*). Inhibition constants were determined graphically from Dixon plots. Inhibition assays in the presence of TMAO followed essentially the same general procedure but with C191A/C220A trypsin (0.2 μM) and the BPTI analogue (0.08−0.1 μM for WT BPTI, 0.2−0.4 μM for $[14-38]_{Abu}$, 0.4−1.8 μM for P26D27$[14-38]_{Abu}$, 1.7−3.8 μM for N26G27K28$[14-38]_{Abu}$, and 3.8−4.4 μM for $[R]_{Abu}$) prepared separately in buffer containing 0−2 M TMAO. The complex was then preincubated at 4 °C for 1 h, after which time the assay was carried out at 25 °C.

RESULTS

New $[14-38]_{Abu}$ Variants. To further characterize the properties of partially folded BPTI ensembles, we constructed by continuous-flow Fmoc solid-phase synthesis two variants of $[14-38]_{Abu}$ in which the turn between core strands of the antiparallel sheet is perturbed. In P26D27$[14-38]_{Abu}$, replacements are by amino acids with higher propensity for a type I β-turn (*25*). In N26G27K28$[14-38]_{Abu}$, replacements in the turn favor a type I′ turn, the most common in β-hairpins (*26*, *27*). Ala 25 is unaltered, the rationale being to preserve strong side chain−side chain interactions between Ala 25 and the Tyr 23 ring that are indicated by NOE data in partially folded $[14-38]_{Abu}$ (*15*) and also detected in fully unfolded $[R]_{Abu}$, an unfolded analogue with all six Cys replaced by Abu (*10*). Replacement of Tyr 23 by Ala in $[14-38]_{Abu}$ results in a species with no obvious CD or NMR structure (*28*).

Far-UV CD spectra of partially folded $[14-38]_{Abu}$ (Figure 1) show negative ellipticity at 220 nm arising from secondary structure and a strong minimum around 202 nm characteristic of BPTI and attributed to tertiary contacts of aromatic groups (*29*). $[R]_{Abu}$ has a CD spectrum typical of denatured proteins. Both $[14-38]_{Abu}$ analogues are more disordered than the parent compound. Spectra of P26D27$[14-38]_{Abu}$ show loss of negative ellipticity at 220 nm as compared to $[14-38]_{Abu}$, consistent with a higher population of more unfolded species, but slightly more negative ellipticity than $[R]_{Abu}$, suggesting some secondary structure. For the other analogue, N26G27K28$[14-38]_{Abu}$, no local minimum is detected at 220 nm. Near-UV CD spectra of partially folded $[14-38]_{Abu}$ (Figure 1) have a minimum at 277 nm, indicative of tertiary structure. Both analogues have >75% reduction of signal at 277 nm, indicating that most tertiary structure is lost, compared to $[14-38]_{Abu}$.

One-dimensional ^1H NMR spectra of partially folded $[14-38]_{Abu}$ (Figure 2C), under the same pH and temperature conditions as the CD experiments, contain resolved downfield resonances that are characteristic of the antiparallel β-sheet strands in the core of native BPTI. Similar spectra

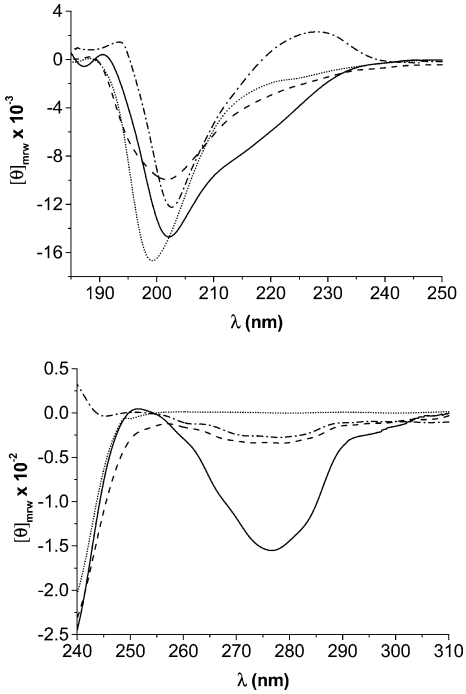

FIGURE 1: Far-UV (upper) and near-UV (lower) CD spectra of partially folded and unfolded BPTI variants at 1 °C in 10 mM phosphate, pH 4.6: $[14-38]_{Abu}$ (solid line); $[R]_{Abu}$ (dotted line); P26D27$[14-38]_{Abu}$ (dashed line); N26G27K28$[14-38]_{Abu}$ (dashed−dotted line).

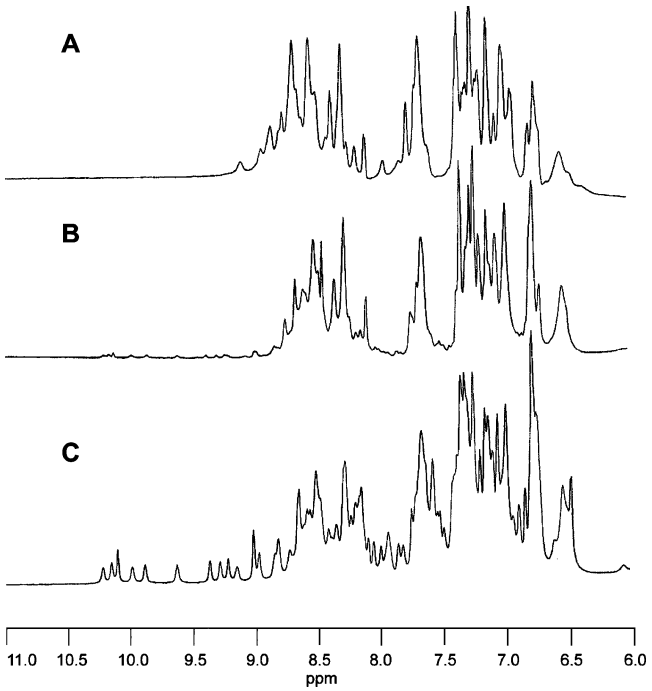

FIGURE 2: One-dimensional ^{1}H NMR spectra (600 MHz) of partially folded $[14-38]_{Abu}$ and two variants, recorded at 1 °C. Panels: A, N26G27K28$[14-38]_{Abu}$; B, P26D27$[14-38]_{Abu}$; C, $[14-38]_{Abu}$. NMR samples were 0.4 mM in protein in 50 mM deuterated sodium acetate buffer in 90:10 $H_2O/^2H_2O$, pH 4.6. Solvent peak suppression was achieved with the WATERGATE sequence (*54*).

of P26D27$[14-38]_{Abu}$ (Figure 2B) have the same native downfield peaks, but their diminished intensity indicates a greater population of unfolded conformations, consistent with the CD results (Figure 1). Spectra of N26G27K28$[14-38]_{Abu}$ (Figure 2A) have no downfield native-like resonances,

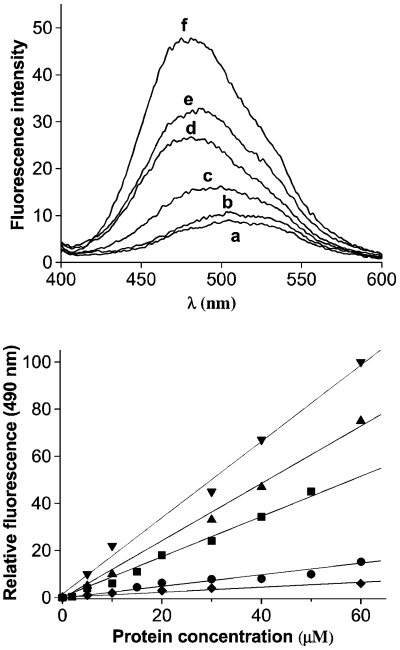

FIGURE 3: Effects of BPTI and analogues on ANS fluorescence. Upper: Fluorescence intensity spectra of ANS (60 μM) in the absence or presence of BPTI and variants (protein concentration 30 μM) at pH 4.6 and 4 °C: (a) no protein; (b) BPTI; (c) $[R]_{Abu}$; (d) $[14-38]_{Abu}$; (e) N26G27K28$[14-38]_{Abu}$; (f) P26D27$[14-38]_{Abu}$. Lower: Enhancement of ANS fluorescence intensity at 490 nm as a function of protein concentration: native BPTI (◆); $[R]_{Abu}$ (●); $[14-38]_{Abu}$ (■); N26G27K28$[14-38]_{Abu}$ (▲); P26D27$[14-38]_{Abu}$ (▼). Protein concentrations ranged from 4 to 60 μM, except for $[R]_{Abu}$ in which the concentration range extended to 100 μM (last data point not shown; it fits on the drawn straight line). The relative intensity is the difference between fluorescence of free ANS and fluorescence of ANS in the presence of protein.

implying mainly or all unfolded structures, also consistent with the CD data.

The dye 8-anilino-1-naphthalenesulfonic acid (ANS) is thought to increase fluorescence when bound to hydrophobic surfaces and thereby to provide a fluorescence diagnostic of partially folded proteins (*30*). Neither fully folded proteins nor unfolded proteins bind ANS. Figure 3 (upper) shows the emission spectrum of 60 μM ANS, by itself or in the presence of 30 μM natural BPTI or the various analogues. Both $[14-38]_{Abu}$ analogues exhibit higher apparent binding for ANS than the parent partially folded protein and considerably higher binding than unfolded $[R]_{Abu}$. In accord with common observations, the native protein does not appreciably affect ANS fluorescence. Compared to $[14-38]_{Abu}$, P26D27$[14-38]_{Abu}$ exhibits a blue shift in the wavelength of maximum intensity from 520 to 490 nm, suggesting a more exposed hydrophobic surface. However, the maximum emission wavelength in the presence of N26G27K28$[14-38]_{Abu}$ is close to that of $[14-38]_{Abu}$.

Figure 3 (lower) shows the dependence of ANS fluorescence emission intensity at 490 nm on protein concentration. Binding is not saturated, and intensity increases linearly with increasing protein levels. The slopes of the curves give a qualitative indication of the strength of ANS−protein binding. For natural BPTI, as well as for $[R]_{Abu}$, the small slope confirms weak ANS binding. Higher slopes for other analogues suggest that their partially folded ensembles have, on average, greater exposed hydrophobic surface area and that P26D27$[14-38]_{Abu}$ has the highest. The linearity of the

Table 1: Affinities of BPTI and Analogues for the C191A/C220A Trypsin Mutant[a]

protein	K_i (nM)
BPTI	0.32 ± 0.07[b]
$[14-38]_{Abu}$	4.46 ± 0.19
$P26D27[14-38]_{Abu}$	16.9 ± 1.26
$N26G27K28[14-38]_{Abu}$	127 ± 4
$[R]_{Abu}$	not detectable

[a] Assays were performed at 25 °C in 50 mM Hepes, 100 mM NaCl, and 10 mM $CaCl_2$, pH 8.0. Values reported are the average of at least two independent experiments. [b] K_i for wild-type trypsin is, as reported by Wang et al. (*34*), 0.46 nM.

curves indicates the absence of aggregation in the concentration range of the experiment. When present, aggregates have high affinity for ANS and cause a nonlinear increase in ANS fluorescence with increasing protein concentration (*30*).

Inhibitor Activity of BPTI Analogues. The inhibitor activity of partially folded and unfolded BPTI analogues is difficult to measure, as disordered proteins may be digested by the biological ligand, a serine proteinase. Earlier ways to circumvent this problem (*20*, *31*−*33*) included using anhydrotrypsin with a modified Ser residue in the active center or taking measurements at low pH where trypsin activity is diminished by several orders of magnitude. These approaches have drawbacks, such as loss of trypsin activity in both cases and considerable decrease of inhibitor binding affinity in the second. Here, we use modified rat trypsin missing the conserved C191−C220 disulfide bond. As shown by Wang et al. (*34*), C191A/C220A trypsin has 20−200-fold decreased trypsin activity in assays of the hydrolysis of amide substrates. However, ester hydrolysis activity is decreased by less than 10-fold. Importantly, C191A/C220A trypsin retains a binding affinity for BPTI similar to that for wild-type trypsin [K_i(mut)/K_i(wt) = 0.7], although affinity for small-sized inhibitors is considerably decreased (*34*). Thus, C191A/C220A rat trypsin is well suited for assays of trypsin inhibitor activity of partially folded and unfolded variants of BPTI.

Binding constants of BPTI analogues for C191A/C220A trypsin (Table 1) were determined from Dixon plots (*35*), in which initial rates of trypsin substrate hydrolysis are evaluated (Figure 4). Since BPTI is a slow-binding inhibitor (*19*), the enzyme and the BPTI analogue were preincubated at 4 °C for at least 1 h, following which the complex was added to the substrate-containing assay mixture at 25 °C. Assay mixtures had final concentrations of 5 nM enzyme and 5−250 nM inhibitor. Control experiments show that preincubation is necessary to ensure that equilibrium is attained; the enzyme is entirely stable over the protocol time period. Further, when a 1:1 molar mixture of C191A/C220A trypsin with $[R]_{Abu}$ is incubated overnight, there is no cleavage of the unfolded analogue.

Partially folded $[14-38]_{Abu}$ has 14-fold weaker affinity for trypsin than wild-type BPTI but completely inhibits trypsin at a molar ratio of 6:1 (Table 1 and Figure 4D). Relative to parent $[14-38]_{Abu}$, affinity for trypsin is decreased by only 4-fold for $P26D27[14-38]_{Abu}$ and 28-fold for $N26G27K28[14-38]_{Abu}$. Incubation with $P26D27[14-38]_{Abu}$ reduces enzyme activity to 20% at a molar ratio of 20:1 (Figure 4D), whereas incubation with $N26G27K28$ $[14-38]_{Abu}$ reduces residual tryptic activity to about 40% at a molar ratio of 45:1 (data not shown). With $[R]_{Abu}$, no inhibition is detected even at very high excess (up to 500 molar equivalents) relative to trypsin.

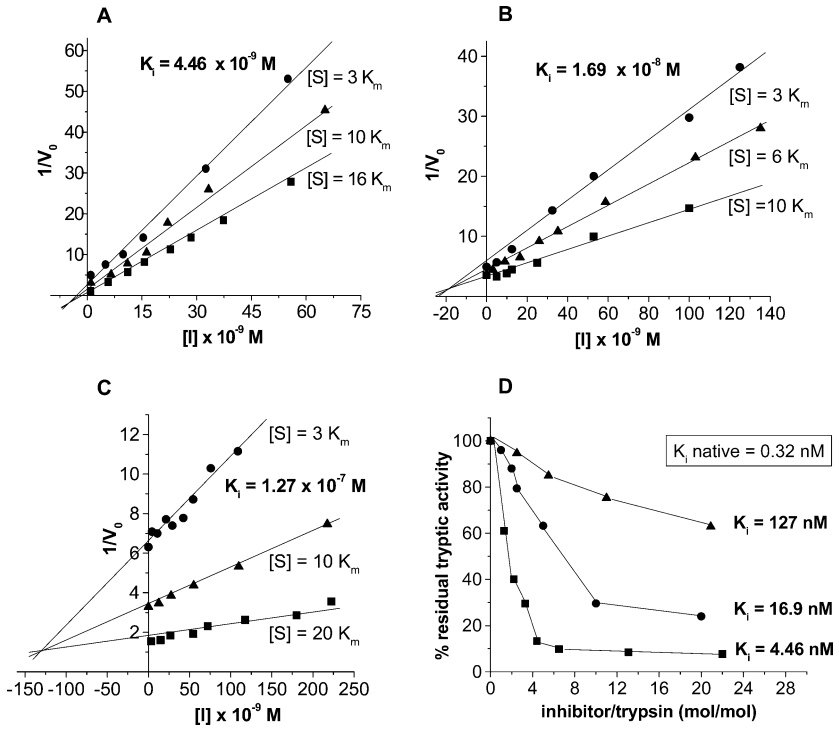

FIGURE 4: Trypsin inhibition assays of $[14-38]_{Abu}$ and two variants. Dixon plots for the inhibition of C191A/C220A trypsin by (A) $[14-38]_{Abu}$, (B) $P26D27[14-38]_{Abu}$, and (C) $N26G27K28[14-38]_{Abu}$. (D) Residual C191A/C220A trypsin activity in the presence of increasing amounts of $[14-38]_{Abu}$. Measurements were performed in 100 mM NaCl, 10 mM $CaCl_2$, and 50 mM Hepes, pH 8.0. The K_m of C191A/C220A trypsin for the Z-Lys-SBzl substrate is 31 μM (*34*). The experimental procedure is sketched in the text, and further details are given in Supporting Information.

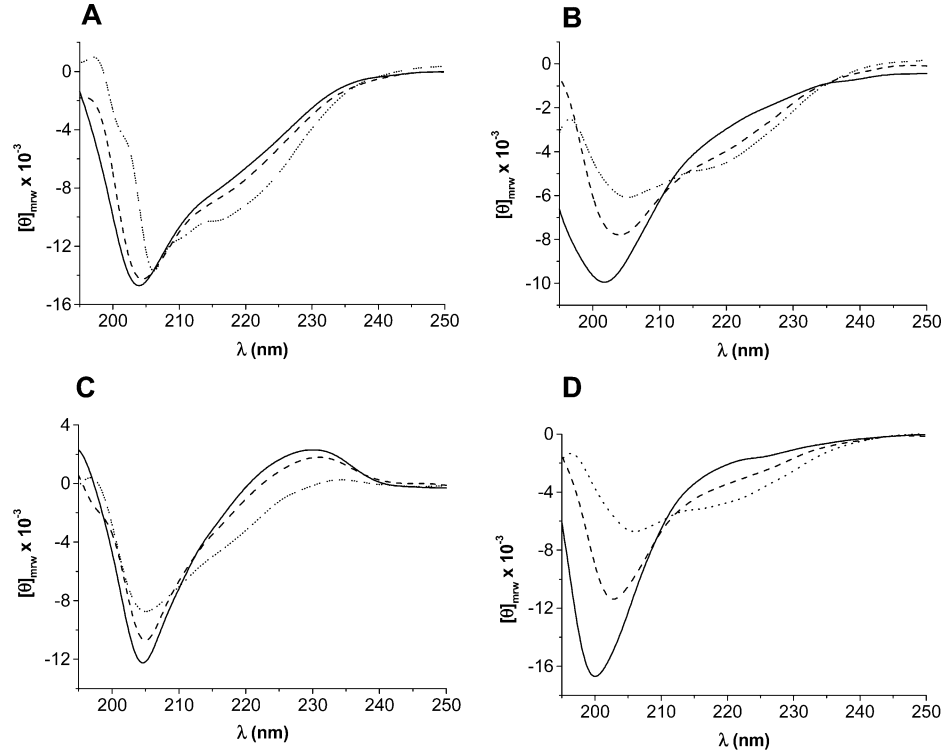

FIGURE 5: Effect of TMAO on the far-UV CD spectra of BPTI variants: (A) $[14-38]_{Abu}$; (B) P26D27$[14-38]_{Abu}$; (C) N26G27K28-$[14-38]_{Abu}$; (D) $[R]_{Abu}$. Spectra were recorded at 1 °C in 10 mM phosphate, pH 4.6, in the absence of TMAO (solid line), in the presence of 1 M TMAO (dashed line), and in the presence of 2 M TMAO (dotted line).

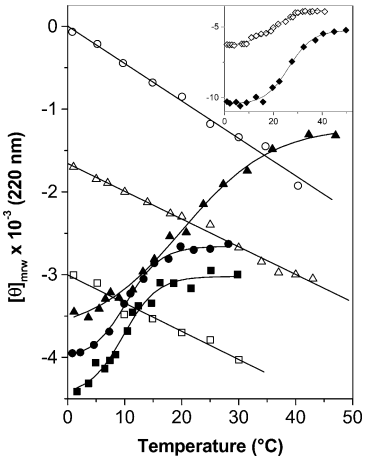

FIGURE 6: Temperature dependence of molar ellipticity at 220 nm in the absence (open symbols) and presence of 2 M TMAO (closed symbols), monitored at pH 4.6. Variants examined: $[R]_{Abu}$ (▲), P26D27$[14-38]_{Abu}$ (■), N26G27K28$[14-38]_{Abu}$ (●), and $[14-38]_{Abu}$ (inset, ◆).

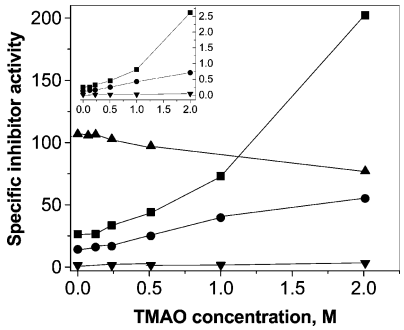

FIGURE 7: Effect of TMAO on inhibition of C191A/C220A trypsin by native BPTI (▲), $[14-38]_{Abu}$ (■), P26D27$[14-38]_{Abu}$ (●), and N26G27K28$[14-38]_{Abu}$ (▼). Specific activity is computed as (units/nmol) $\times$ 4 $\times$ 10^{-4}. Inset: Data are normalized with respect to the effect of TMAO on the activity of native BPTI. All measurements and calculations are detailed in Supporting Information.

CD-Detected Ordering by an Osmolyte. CD spectra of BPTI analogues as recorded in 1 and 2 M TMAO (Figure 5) show an increase in molar negative ellipticity at 220 nm, relative to the spectrum in the absence of TMAO. There is also a decrease in the minimum, which shows a progressive shift from 200 to 205 (depending on the analogue) to about 206 nm for all four analogues in 2 M TMAO. These observations indicate increased secondary structure in the presence of TMAO. Experiments in the presence of higher concentrations of TMAO were not feasible due to protein precipitation. In 2 M TMAO, the ellipticity at 220 nm follows sigmoidal thermal unfolding curves (Figure 6, closed symbols); midpoint temperatures are 9.7 °C for P26D27-$[14-38]_{Abu}$, 10.3 °C for N26G27K28$[14-38]_{Abu}$, 20.1 °C for $[R]_{Abu}$, and 26.3 °C for $[14-38]_{Abu}$. In the absence of TMAO, the temperature-dependent ellipticity at 220 nm (Figure 6, open symbols) for P26D27$[14-38]_{Abu}$, N26G27K28-$[14-38]_{Abu}$, and $[R]_{Abu}$ shows no evidence of cooperative unfolding but rather a linear decrease as is characteristic of peptides with a high degree of disorganization (*36*). In contrast, $[14-38]_{Abu}$ without TMAO shows a moderately cooperative transition with a midpoint temperature of 19 °C (Figure 6, inset). For $[14-38]_{Abu}$, T_m increases by 7 °C in 2 M TMAO, with respect to controls with no TMAO. For the other analogues, a difference in T_m cannot be obtained since in the absence of TMAO the temperature dependence does not indicate cooperative unfolding.

Trypsin inhibitor activity of $[14-38]_{Abu}$ and its two analogues was measured as a function of TMAO concentration (Figure 7). The molar ratio of inhibitor to trypsin in these experiments was varied from 0.4:1 to 22:1, depending

on the analogue. Trypsin inhibitor activity increases for $[14-38]_{Abu}$ and P26D27$[14-38]_{Abu}$ as TMAO concentration is increased to 2 M. For N26G27K28$[14-38]_{Abu}$, a small increase (3-fold) of inhibitor activity is observed in 2 M TMAO, but this is at the limit of detection and is not evident on the scale of Figure 7. $[R]_{Abu}$ lacks detectable activity both in the absence of TMAO as well as in the presence of 2 M TMAO.

DISCUSSION

Native-like Conformations in Disordered Proteins. A key feature of our folding model, summarized in the introduction, is the proposal that partially folded ensembles have site-specific energy barriers between interconverting conformations, some of which are native-like. In partially folded $[14-38]_{Abu}$, these are evident in NMR spectra as separate slow-exchange peaks for the same $^{15}N^{1}H$ group (*37*); the peaks report different conformations that interconvert on a time scale longer than milliseconds. The conformations, designated P_f and P_d, are site-specific because they are detected by a microscopic probe, a single NH. Each conformation consists of numerous rapidly interconverting conformers. The notation refers to the partially folded ensemble, P, containing a denatured conformation (P_d), in which the NH has a chemical shift and sequential NOEs indistinguishable from "random coil", and a more ordered conformation (P_f), in which the NH has a chemical shift and short- and long-range NOEs more similar to folded protein. For partially folded $[14-38]_{Abu}$ labeled with ^{15}N at selected backbone amides, NMR spectra show P_f and P_d conformations for each $^{15}N^{1}H$ group; for core residues the dominant population is P_f, while outside the core the dominant population is P_d. For residues in core segments, P_f peaks report native-like amide properties. For residues in noncore segments, P_f peaks report amide groups that are not native-like but are more ordered than a random coil.

The slow conformational interchange phenomena in partially folded BPTI variants are, we propose, general for partially folded proteins (*6*, *7*). This is consistent with results obtained for other proteins, including α-lactalbumin, *staphylococcal* nuclease, and thioredoxin (*38*−*40*). In α-lactalbumin, NMR spectra cross-peaks are observed under unfolding conditions but not under molten globule conditions; this is explained by line broadening in the molten globule due to intermediate exchange (*38*). In denatured *staphylococcal* nuclease, some residues are apparently obscured by broadening; intermediate chemical exchange is inferred from the observation of their cross-peaks in spectra of fully unfolded nuclease (*39*). In thioredoxin, NMR spectra of reconstituted fragments have peaks that are present under denaturing conditions but missing under milder conditions, implying that, in the latter case, the protein is partially folded and the peaks are broadened by intermediate exchange (*40*). We suggest that, in partially folded proteins, core backbone atoms are, in general, in slow−intermediate exchange between native-like and disordered conformational families that interconvert by *local* motions. Thus, families of locally fluctuating conformations are separated by an energy barrier, the basis for which may be important in understanding early events in protein folding.

The underlying process responsible for slow exchange in core segments in BPTI variants is not *cis*−*trans* isomerism of proline-containing peptide bonds. The core hairpin consists of residues 18−35, while the four prolines in BPTI are at positions 2, 8, 9, and 13. The turn between amino acids 25−28 is apparently an initiation site for formation of the core strands of antiparallel β-sheet (*6*), a plausible candidate for the process governing slow exchange. For P_f conformations outside the core in the C-terminal region, proline isomerism is ruled out as the underlying process for P_f-P_d interconversion by the absence of proline. For P_f conformations outside the core in the N-terminal region, proline isomerism cannot be ruled out, but it is unlikely. Residues 4 and 6 have more than one major peak representing P_d conformations, and they are assigned to *cis*- and *trans*-proline isomers of denatured conformations; however, residues 4 and 6 have only one minor peak representing a P_f (*37*).

The special case of Gly 37, which has three slow-exchange peaks, is especially relevant to this activity study of partially folded $[14-38]_{Abu}$. Three slow-exchange peaks for Gly 37 mean that for this site there are three slowly interconverting conformations, one P_d and two P_f (*6*); P_d is most populated, the native-like P_f is least populated, and the other P_f is intermediate. Assignment of the minor native-like P_f peak is unambiguous, as its chemical shift is far upfield and similar to 37NH in native BPTI (*41*). It is reasonable to assume that when Gly 37 NH is native-like, the 11−18 active site loop is also native-like, as it packs with the loop containing Gly 37. Plausible candidates for the process(es) governing conversion to native-like P_f for 37NH are formation of the strained Gly 36−Gly 37 dihedral angles to bring Gly 37 into proper orientation for an NH−aromatic contact with the Tyr 35 ring and/or packing of the 11−18 loop with the 34−39 loop. The existence of a minor native-like conformation of the active site loops in partially folded $[14-38]_{Abu}$ and in its more disordered analogues is strongly supported by results on their inhibition of trypsin, as described in the next section.

Partially Folded $[14-38]_{Abu}$ Has Inhibitor Activity. Partially folded $[14-38]_{Abu}$ has 10-fold weaker affinity for trypsin than wild-type BPTI and completely inhibits trypsin at a molar ratio of 6:1 (Table 1 and Figure 4D). This is very strong binding, as trypsin−BPTI is among the most tightly bound complexes known, with a K_{assoc} for rat trypsin approaching 10^{10} at neutral pH (*19*, *42*). Although native-like loops for Gly 37 NH may be present in a very small proportion of the population, a strong ligand binds native conformations and effectively removes them from the pool of free conformers, thereby converting a significant fraction of the ensemble to bound, native-like conformations. This is further confirmed by crystal structures recently solved for a modified rat trypsin complexed with $[14-38]_{Abu}$ (*43*). The indication from the NMR data of a significant energy barrier between a minor native-like loop conformation and two other conformations might explain why long equilibration prior to assay is most effective.

Highly Disordered Variants of $[14-38]_{Abu}$ Sample Native-like Conformations. Both new analogues of $[14-38]_{Abu}$ are more disordered than the parent $[14-38]_{Abu}$. P26D27$[14-38]_{Abu}$ has much less native-like core structure than $[14-38]_{Abu}$, while N26G27K28$[14-38]_{Abu}$ has no detectable ordered structure (Figures 1 and 2). Both analogues exhibit higher apparent binding for ANS than the parent protein and considerably higher binding than unfolded $[R]_{Abu}$. Both $[14-38]_{Abu}$ analogues inhibit trypsin (Table 1).

Affinity for trypsin is decreased by only 4-fold for P26D27-$[14-38]_{Abu}$ relative to parent $[14-38]_{Abu}$ and by 50-fold relative to native BPTI. Incubation with enzyme reduces trypsin activity to 20% at a molar ratio of 20:1. Affinity for trypsin is decreased 28-fold for N26G27K28$[14-38]_{Abu}$ relative to parent $[14-38]_{Abu}$ and 400-fold relative to native BPTI. Incubation of trypsin with N26G27K28$[14-38]_{Abu}$ reduces residual tryptic activity to about 40% at a molar ratio of 45:1. For $[R]_{Abu}$, no inhibition is detectable even at very high concentration relative to trypsin. Thus, even though P26D27$[14-38]_{Abu}$ has considerably more disorder than $[14-38]_{Abu}$, and N26G27K28$[14-38]_{Abu}$ has no obvious ordered structure, both convert to some extent to native-like conformations in the presence of trypsin. The crystal structure of the complex of modified rat trypsin and P26D27-$[14-38]_{Abu}$ (*43*) confirms these ideas. The detection of inhibitor activity in $[14-38]_{Abu}$ and in its analogues implies that native-like ligand binding conformations are populated in their partially folded ensembles.

Comparison of N26G27K28$[14-38]_{Abu}$ with $[R]_{Abu}$ illustrates interesting differences between partially folded and fully unfolded ensembles of the same protein. Both variants show no obvious indication of ordered structure in NMR or CD spectra. While no inhibitor activity is measured even at very high ratios of $[R]_{Abu}$ to trypsin, activity is observed for N26G27K28$[14-38]_{Abu}$, albeit at a reduced level relative to $[14-38]_{Abu}$ and P26D27$[14-38]_{Abu}$. The fact that any level of activity is observed for N26G27K28$[14-38]_{Abu}$ implies that the ensemble includes a small population of native-like conformations in equilibrium with the dominant population of disordered forms. This is consistent with the significant enhancement of ANS fluorescence by N26G27K28-$[14-38]_{Abu}$ but not by $[R]_{Abu}$ (Figure 3), suggesting that the former contains, on average, more hydrophobic clustering for ANS binding. It is interesting to note in passing that $[R]_{Abu}$ has some detectable non-native order in core residues (*44*).

We conclude from activity and ANS binding data that while N26G27K28$[14-38]_{Abu}$ and $[R]_{Abu}$ are both primarily extended and flexible, the disulfide cross-link in the $[14-38]_{Abu}$ analogue favors formation of minor populations of collapsed conformations that sample native-like structure. It is reported that 14−38 is the first, transiently formed, disulfide in disulfide-linked refolding (*45*, *46*). Since the partially folded ensemble of $[14-38]_{Abu}$ is, on average, more native-like in core regions and more disordered in the vicinity of the 14−38 disulfide, the ordering influence of the cross-link is primarily an entropic effect (*6*). That is, elimination by the cross-link of more extended conformations results in an ensemble with a higher population of more collapsed conformations, among which species that are native-like in the slow exchange core are sampled. We expect that any single, native-like disulfide would do the same, suggesting that order of disulfide bond formation is not critical in BPTI folding.

Osmolyte-Induced Folding. Addition of TMAO induces CD-detected structure in partially folded and unfolded analogues of BPTI. Organic osmolytes are small molecules that accumulate in the blood of some animals and prevent proteins from unfolding, aggregating, and losing functional activity when exposed to environmental stresses such as extreme temperatures and high salt concentration (*47*). The organic osmolyte TMAO has been reported to fold a number of proteins in vitro (*48*−*51*). CD spectra of the three $[14-38]_{Abu}$ analogues, and of $[R]_{Abu}$, have greater negative ellipticity at 220 nm in 1 and 2 M TMAO, as compared to spectra of samples without osmolyte (Figure 5). The structure induced by TMAO is cooperatively formed, as shown by the shapes of their thermal unfolding curves in 2 M osmolyte (Figure 6, closed symbols**)**. In the absence of TMAO, only $[14-38]_{Abu}$ shows cooperative thermal unfolding (Figure 6, inset), while the other analogues show a progressive decrease in ellipticity at 220 nm (Figure 6) typical of disordered polypeptides (*36*).

The inhibitor activities of $[14-38]_{Abu}$ and P26D27-$[14-38]_{Abu}$ are enhanced by TMAO (Figure 7); the specific inhibitor activity, normalized to WT BPTI activity, increases by about an order of magnitude upon going from the absence of TMAO to 2 M TMAO (Figure 7, inset). For the N26G27K28$[14-38]_{Abu}$ variant, a small (3-fold) increase of inhibitor activity is observed in 2 M TMAO, but this is too small a change to show on the scale of Figure 7. These results imply that TMAO-induced structure is native-like in $[14-38]_{Abu}$ and P26D27$[14-38]_{Abu}$, consistent with the results of Bolen and associates on other proteins (*12*, *13*, *52*). The small increase in activity of N26G27K28$[14-38]_{Abu}$ in the presence of 2 M TMAO suggests that its TMAO-induced structure is also native-like, but the data are not conclusive. Since the triple mutant is more disordered that the other $[14-38]_{Abu}$ analogues, it very likely requires higher TMAO concentrations to induce refolding. At 2 M osmolyte, it may well be that for N26G27K28$[14-38]_{Abu}$ the CD-detected refolding, indicated by the data in Figures 5 and 6, is not enough to give clear indication of increased trypsin inhibitor activity under the assay conditions used here. $[R]_{Abu}$ shows no trypsin inhibitor activity in 2 M TMAO, although it has some CD-detected changes that indicate increased collapsed structure (Figure 7). This is likely to be similar to the TMAO effect on reduced and carboxamidated RNase A, often taken as a model for a random coil protein. In this case, TMAO causes the random coil to contract, but no folding takes place (*53*).

CONCLUSIONS

Three partially folded analogues of BPTI have inhibitor activity. The analogues differ widely in the degree to which they are folded, and inhibitor activity increases with detectable order. The parent $[14-38]_{Abu}$ favors native-like structure in the core β-hairpin and has minor populations of ordered conformations in regions outside the core. Analogue P26D27-$[14-38]_{Abu}$ has minor native-like structure detectable for core residues, while analogue N26G27K28$[14-38]_{Abu}$ has no ordered structure observable by CD or NMR. The observation of inhibitor activity implies that in partially folded proteins, even those with very small levels of ordered structure, there are native-like conformations in equilibrium with the dominant disordered species. This supports our NMR interpretation of slow-exchange peaks in $[14-38]_{Abu}$ and the related model of partially folded protein ensembles.

The osmolyte TMAO induces CD-detected structure in the three partially folded analogues of $[14-38]_{Abu}$, as well as in fully unfolded $[R]_{Abu}$. In all four species, the TMAO-induced structure is formed cooperatively, but to varying degrees. Inhibitor activity assays in the presence of TMAO indicate that osmolyte-induced structure is native-like in $[14-38]_{Abu}$

and P26D27[14−38]$_{Abu}$ and also may be native-like in N26G27K28[14−38]$_{Abu}$. In fully unfolded [R]$_{Abu}$, the CD-detected changes likely arise from non-native structure.

Although neither N26G27K28[14−38]$_{Abu}$ nor [R]$_{Abu}$ have CD- or NMR-detected structure, the former can be distinguished as partially folded. N26G27K28[14−38]$_{Abu}$ has ANS fluorescence enhancement similar to [14−38]$_{Abu}$ and P26D27-[14−38]$_{Abu}$, as well as measurable inhibitor activity. Apparently, the triple mutant sufficiently favors collapsed conformations to give an ANS response similar to a partially folded protein and sufficiently favors native-like conformations to have measurable (albeit low) inhibitor activity. In these respects its ensemble of conformations, although mostly disordered, is clearly different from fully unfolded [R]$_{Abu}$.

ACKNOWLEDGMENT

We thank Professor Lizbeth Hedstrom of Brandeis University (Waltham, MA) for providing the trypsin expression system and for many useful suggestions and discussions, Dr. Frank Bordusa from the University of Leipzig (Germany) for encouragement and helpful discussions on trypsin mutant expression, and Dr. Mark F. Reynolds, formerly from the University of Minnesota, for help with enzyme purifications. NMR experiments were performed at the University of Minnesota High Field NMR Laboratory.

SUPPORTING INFORMATION AVAILABLE

Details about general materials and methods, instrumentation, and experimental procedures for peptide synthesis, CD, NMR, ANS binding, osmolyte experiments, C191A/C220A trypsin expression and purification, and trypsin inhibition assays. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) *Biochemistry 41*, 6573−6582.
2. Gunasekaran, K., Tsai, C.-J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) *Trends Biochem. Sci. 28*, 81−85.
3. Dobson, C. M. (1999) *Trends Biochem. Sci. 24*, 329−332.
4. Ferrer, M., Woodward, C., and Barany, G. (1992) *Int. J. Pept. Protein Res. 40*, 194−207.
5. Ferrer, M., Barany, G., and Woodward, C. (1995) *Nat. Struct. Biol. 2*, 211−217.
6. Barbar, E., Hare, M., Daragan, V., Barany, G., and Woodward, C. (1998) *Biochemistry 37*, 7822−7833.
7. Barbar, E., Hare, M., Makokha, M., Barany, G., and Woodward, C. (2001) *Biochemistry 40*, 9734−9742.
8. Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.
9. Li, R., and Woodward, C. (1999) *Protein Sci. 8*, 1571−1591.
10. Pan, H., Barbar, E., Barany, G., and Woodward, C. (1995) *Biochemistry 34*, 13974−13981.
11. Liu, Y., and Bolen, D. W. (1995) *Biochemistry 34*, 12884−12891.
12. Baskakov, I. V., Kumar, R., Srinivasan, G., Ji, Y.-S., Bolen, D. W., and Thompson, E. B. (1999) *J. Biol. Chem. 274*, 10693−10696.
13. Bolen, D. W., and Baskakov, I. V. (2001) *J. Mol. Biol. 310*, 955−963.
14. Song, J.-L., and Chuang, D. T. (2001) *J. Biol. Chem. 276*, 40241−40246.
15. Barbar, E., Barany, G., and Woodward, C. (1995) *Biochemistry 34*, 11423−11434.
16. Barany, G., Gross, C. M., Ferrer, M., Barbar, E., Pan, H., and Woodward, C. (1996) in *Techniques in Protein Chemistry* (Marshak, D., Ed.) Vol. VII, pp 503−513, Academic Press, San Diego, New York, Boston, London, Sidney, Tokyo, and Toronto.
17. Angell, Y. M., Alsina, J., Albericio, F., and Barany, G. (2002) *J. Pept. Res. 60*, 292−299.
18. Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992) *Science 255*, 1249−1253.
19. Hedstrom, L., Lin, T.-Y., and Fast, W. (1996) *Biochemistry 35*, 4515−4523.
20. Krokoszynska, I., Dadlez, M., and Otlewski, J. (1998) *J. Mol. Biol. 275*, 503−513.
21. Pasternak, A., Liu, X., Lin, T.-Y., and Hedstrom, L. (1998) *Biochemistry 37*, 16201−16210.
22. Wang, A., and Bolen, D. W. (1997) *Biochemistry 36*, 9101−9108.
23. Grassetti, D. R., and Murray, J. F. (1967) *Arch. Biochem. Biophys. 119*, 41−49.
24. Green, G. D. J., and Shaw, E. (1979) *Anal. Biochem. 93*, 223−226.
25. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci. 3*, 2207−2216.
26. Chou, P. Y., and Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135−175.
27. Chou, K.-C. (2000) *Anal. Biochem. 286*, 1−16.
28. Barbar, E., Barany, G., and Woodward, C. (1996) *Folding Des. 1*, 65−76.
29. Manning, M. C., and Woody, R. W. (1989) *Biochemistry 28*, 8609−8613.
30. Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991) *Biopolymers 31*, 119−128.
31. Ako, H., Foster, R. J., and Ryan, C. A. (1972) *Biochem. Biophys. Res. Commun. 47*, 1402−1407.
32. Finkenstadt, W. R., Hamid, M. A., Mattis, J. A., Schrode, J., Sealock, R. W., and Laskowski, M. (1974) in *Bayer Symposium V* (Fritz, H., Tschesche, H., and Green, L. J., Eds.) pp 389−411, Springer-Verlag, Berlin.
33. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Huber, R. (1983) *Acta Crystallogr. B39*, 480−490.
34. Wang, E. C. W., Hung, S.-H., Cahoon, M., and Hedstrom, L. (1997) *Protein Eng. 10*, 405−411.
35. Roberts, D. V. (1977) *Enzyme Kinetics*, Cambridge University Press, London, New York, and Melbourne.
36. Privalov, P. L., Tiktopulo, E. I., Venyaminov, S. Y., Griko, Y. V., Makhatadze, G. I., and Khechinashvili, N. N. (1989) *J. Mol. Biol. 205*, 737−750.
37. Barbar, E., LiCata, V. J., Barany, G., and Woodward, C. (1997) *Biophys. Chem. 64*, 45−57.
38. Schulman, B. A., Kim, P. S., Dobson, C. M., and Redfield, C. A. (1997) *Nat. Struct. Biol. 4*, 630−634.
39. Wang, Y., and Shortle, D. (1995) *Biochemistry 34*, 15895−15905.
40. Tasayco, M. L., Fuchs, J., Yang, X.-M., Dyalram, D., and Georgescu, R. E. (2000) *Biochemistry 39*, 10613−10618.
41. Tüchsen, E., and Woodward, C. (1987) *Biochemistry 26*, 1918−1925.
42. Vincent, J.-P., and Lazdunski, M. (1972) *Biochemistry 11*, 2967−2977.
43. Getun, I., Brown, C. K., Tulla-Puche, J., Ohlendorf, D., Woodward, C., and Barany, G. (2004) in *Innovation and Perspectives in Solid-Phase Synthesis & Combinatorial Libraries: Collected Papers, Eighth International Symposium* (Epton, R., Ed.) Mayflower Scientific, Kingswinford, England, and London, England, U.K. (in press).
44. Pan, H., Barany, G., and Woodward, C. (1997) *Protein Sci. 6*, 1985−1992.
45. Dadlez, M., and Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 674−679.
46. Dadlez, M., and Kim, P. S. (1996) *Biochemistry 35*, 16153−16164.
47. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) *Science 217*, 1214−1222.
48. Gursky, O. (1999) *Protein Sci. 8*, 2055−2064.
49. Jaravine, V. A., Rathgeb-Szabo, K., and Alexandrescu, A. T. (2000) *Protein Sci. 9*, 290−301.
50. Henkels, C. H., Kurz, J. C., Fierke, C. A., and Oas, T. G. (2001) *Biochemistry 40*, 2777−2789.
51. Uversky, V. N., Li, J., and Fink, A. L. (2001) *FEBS Lett. 509*, 31−35.
52. Baskakov, I., and Bolen, D. W. (1998) *J. Biol. Chem. 273*, 4831−4834.
53. Qu, Y., Bolen, C. L., and Bolen, D. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9268−9273.
54. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
55. Houghten, R. A., and Li, C. H. (1979) *Anal. Biochem. 98*, 36−46.
56. Sherman, F., Fink, G. F., and Hicks, J. B. (1986) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

BI035301A