ELSEVIER

Biophysical Chemistry 64 (1997) 45–57

Biophysical Chemistry

Local fluctuations and global unfolding of partially folded BPTI detected by NMR [1]

Elisar Barbar [a], Vince J. LiCata [a], George Barany [b], Clare Woodward [a,*]

[a] *Department of Biochemistry, University of Minnesota, St. Paul, MN 55108 USA*
[b] *Department of Chemistry, University of Minnesota, Minneapolis, MN 55455 USA*

Received 22 May 1996; revised 19 July 1996; accepted 19 July 1996

Abstract

The protein $[14–38]_{Abu}$ is a chemically synthesized variant of bovine pancreatic trypsin inhibitor (BPTI) with the 14-38 disulfide bond intact and cysteines 5, 30, 51, and 55 replaced by α-amino-*n*-butyric acid (Abu). At 1–6°C and pH 4.5–6.5, $[14–38]_{Abu}$ is partially folded with a native-like core [1]. Heteronuclear NMR spectra contain two, and in a few cases three or four, exchange cross peaks for each ^{15}N-bound ^{1}H, reporting the presence of two or more conformations that interconvert on a time scale of $\geq$ milliseconds. Thermodynamic analysis of ^{15}N-^{1}H exchange peak volumes as a function of temperature in the range 1–35°C indicates that partially folded $[14–38]_{Abu}$ undergoes local segmental motions as well as cooperative global unfolding. The relative abundance of more folded versus disordered conformations changes throughout the molecule, indicating that various regions of the partially folded protein are disordered to different extents prior to onset of thermal denaturation. This system is unique in providing a measure of the populations of interconverting partially folded conformations, as well as a microscopic view of cooperative folding of a fluctuating ensemble. Although global thermal denaturation is cooperative, significant deviation from simple two-state behavior is reflected in several parameters, including the difference in T_m for thermal unfolding measured by NMR versus circular dichroism.

Keywords: Protein folding; Segmental motions; Chemical exchange; NMR; Equilibrium thermodynamics

1. Introduction

The BPTI analog $[14–38]_{Abu}$ is an ensemble of conformations in which the anti-parallel strands of the central β-sheet are stable, while other parts of the molecule are flexible and disordered [1]. This partially folded BPTI provides a model for the structure, stability and dynamics of an early folding intermediate in which the core is native-like while the rest of the protein fluctuates among non-native conformations. $[14–38]_{Abu}$ has characteristics of a highly ordered, β-sheet molten globule, and undergoes cooperative, temperature-induced denaturation with a

Abbreviations: Abu, α-amino-*n*-butyric acid; BPTI, bovine pancreatic trypsin inhibitor; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; CD, circular dichroism; HPLC, high-performance liquid chromatography; HSQC, ^{15}N-^{1}H heteronuclear single quantum coherence; DSS, 2,2-dimethyl-2-silapentane, sodium salt; ANS, 1-anilino-8-naphthalenesulphonic acid; PF, partially folded ensemble; P_f, more folded conformation of PF; P_d, more disordered conformation of PF

* Corresponding author.

[1] This work was supported by NIH grants GM 26242, GM 51628 and GM 17341.

PII S0301-4622(96)02210-7

midpoint at 19°C when monitored by CD; acid-induced denaturation and formation of an A-state in the presence of chloride anions at low pH are also observed [2]. Unlike native BPTI, partially folded $[14–38]_{Abu}$ binds the fluorescent dye ANS, indicating the presence of solvent-exposed, clustered hydrophobes. Hydrodynamic measurements at pH 4.5 show that partially folded $[14–38]_{Abu}$ is, on average, more extended than native BPTI, but less extended than fully unfolded (reduced) BPTI [3], which is itself collapsed relative to random coil (H. Pan, G. Barany and C. Woodward, manuscript submitted). The 14-38 single disulfide species was detected as a significant kinetic intermediate in disulfide-linked folding of BPTI [4].

The native-like portion of partially folded $[14–38]_{Abu}$ is *not* in the vicinity of the only disulfide, 14-38. Stabilization of partially folded BPTI is apparently a chain entropy effect: the 14-38 cross link eliminates the most extended and entropically stabilized species from the distribution of unfolded conformations; the resulting interconverting ensemble favors partially folded conformations in which the core is collapsed and native-like while the remainder of the protein is more or less extended. Spectra at low temperature show no indication of fully native protein, i. e., the sample is truly 'partially folded', and not a mixture of fully folded and fully unfolded conformations.

The NMR-detected structure of partially folded $[14–38]_{Abu}$ is schematically represented in the upper half of Fig. 1. Spectra contain 2, and in some cases 3

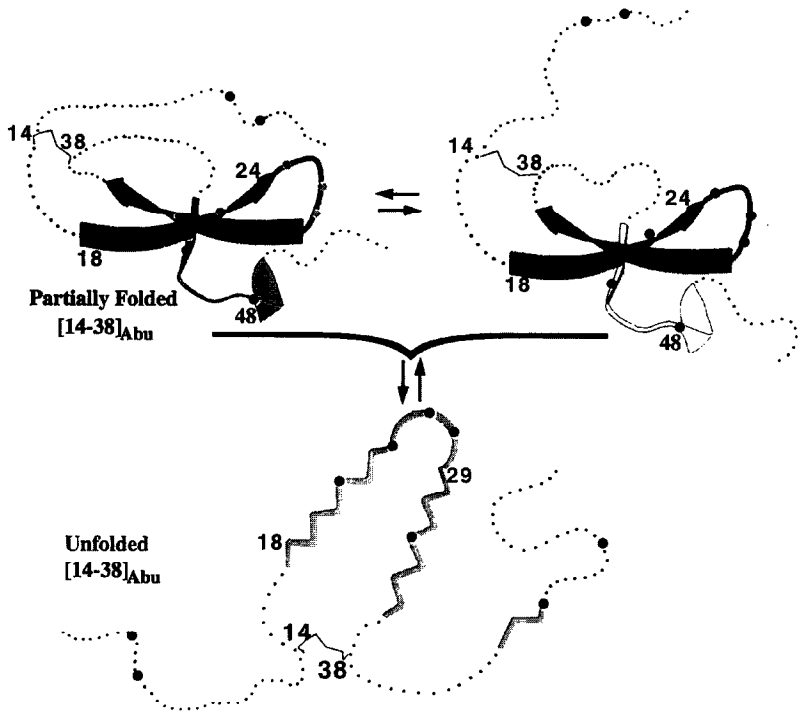

Fig. 1. Schematic representation of the NMR-detected structures of $[14–38]_{Abu}$ in partially folded and unfolded states. The partially folded ensemble of structures of $[14–38]_{Abu}$ (top) is detailed in [1]; there are local, non-cooperative fluctuations between conformations in slow exchange. The most stable region in partially folded $[14–38]_{Abu}$ includes the 18–24, 29–35 antiparallel strands of β-sheet (darkest shaded ribbons). The partially folded protein undergoes a temperature-induced cooperative, global unfolding to an ensemble of denatured conformations (lower). Unfolded $[14–38]_{Abu}$ is characterized in [5]. Shaded circles indicate the sequence positions of ^{15}N backbone probes: Phe4 and Leu6 are in a mostly disordered region. In the more stable, native-like regions, Phe22 and Phe33 are in the anti-parallel strands; Ala25, Ala27, and Gly28 are in the turn between anti-parallel strands; Phe45 is in the β-bridge; and Ala48 is in the first turn of the C-terminal helix. Shading indicates segmental stability; more stable regions are shaded darker. The dotted lines correspond to the more flexible part of the protein for which both conformations are non-native.

or 4, slow exchange cross peaks for each ^{15}N-bound ^{1}H probe; that is, one ^{1}H reports a separate peak for each of 2 or more conformations that interconvert on a time scale longer than milliseconds. Each conformation may contain numerous conformers in fast or intermediate exchange (more rapid equilibrium). The favored conformation of residues 18–35 is a native-like anti-parallel β-sheet, unambiguously indicated by long range NOEs between strands, short range NOEs within the strands, and slowed hydrogen isotope exchange of peptide NHs [1]. In the rest of the molecule, the more abundant conformations are disordered. The C-terminal helix does not behave as a unit, and is more disordered after the first turn.

The only side chains in partially folded $[14–38]_{Abu}$ that show significant, multiple NOEs are Tyr21 and Tyr23, suggesting that interactions of these side chains stabilize the core. When either Tyr21 or Tyr23 is replaced by alanine, $[14–38]_{Abu}$ does not partially fold, but rather is fully denatured and similar, but not identical, to reduced BPTI [5]. This suggests that the 25–28 turn, also sampled in reduced BPTI [3] and apparently stabilized by 23–25 interactions in partially folded $[14–38]_{Abu}$, is the nucleation site for BPTI folding; this hypothesis is being tested in folding kinetics experiments.

Peptide amide nitrogens were specifically labelled with ^{15}N at nine positions distributed along the backbone; their sequence positions are indicated by shaded circles in Fig. 1. This provides ^{15}N-bound ^{1}H reporters whose cross peaks are clearly resolved in heteronuclear NMR spectra of partially folded $[14–38]_{Abu}$. The relative populations of conformations in slow chemical exchange, and their temperature dependence over the range 1–35°C, are reported here. Interconversion rates between conformations are also being estimated, and will be included in future reports. The system is unique in revealing details of conformational fluctuations arising from both segmental motions and global unfolding.

2. Materials and methods

2.1. Peptide synthesis

The protein $[14–38]_{Abu}$, labelled with ^{15}N at positions 4, 6, 22, 25, 27, 28, 33, 45, and 48, was prepared by automated Fmoc solid-phase synthesis as described elsewhere [6]. Purity of $> 98\%$ was confirmed by analytical C-4 reversed-phase HPLC and capillary zone electrophoresis. Amino acid analysis and ion electrospray mass spectrometry of the purified protein were in good agreement with the theoretical values. Mass calculated, 6451.60; mass found 6452.76 ± 0.38 amu.

2.2. CD spectroscopy

CD spectra were measured on a JASCO 710 spectropolarimeter at pH 5.0 in 50 mM acetate buffer and at pH 6.4 without added buffer. The sample in a 10 mm water-jacketed cell was equilibrated 10 min before ellipticity at 220 nm was recorded. Temperatures were measured by direct insertion of a thermistor into the cell. Reversibility was $> 95\%$, as determined by comparison of measurements taken at low temperature before and after the experiment. The reliability of the thermistor was verified by comparing the temperature readings of the thermistor and several high quality mercury thermometers.

2.3. NMR spectroscopy

NMR samples were at a concentration of 0.5 mM in 50 mM deuterated sodium acetate buffer at pH 5.0 or in water without added buffer at pH 6.4. Spectra were acquired on a Bruker AMX-500 spectrometer. Typically, 180–200 t_1 values were obtained, and free induction decays for t_2 were recorded in 2048-point blocks, summing 256 acquisitions each. The water resonance was suppressed by high-power ^{1}H purge pulses [7]. Solution conditions and NMR acquisition parameters for all ^{15}N-^{1}H HSQC spectra [8] were exactly the same, except that temperatures were varied between 1–35°C (initial temperature equilibration time of 5 min). A relaxation delay time of 3 s was used; this is sufficiently longer than T_1 values to permit quantitative volume integration. At 4°C, amide nitrogen T_1 values of (f) and (u) conformations are approximately 0.5 s and 0.7 s, respectively (unpublished results). ^{15}N decoupling during acquisition employed the WALTZ-16 sequence [9]. Proton chemical shifts were measured from an internal DSS standard at 0 ppm. Reversibility of thermal denaturation was verified by comparison of low temperature

spectra acquired before and after unfolding. The temperature of NMR samples was determined in two ways. First, before and/or after the experiment, the protein sample was removed from the spectrometer and replaced by a sealed methanol sample. The temperature of the equilibrated methanol was then determined from the difference between CH_3 and OH proton chemical shifts, which is linearly related to temperature [10]. Second, temperatures of the aqueous protein NMR samples were measured directly by means of a thermistor inserted after the sample was pulsed with the usual sequence for 1 or 5 h. The sample tube was removed from the probe, and the temperature readings were recorded every 10–15 s for 3 min; readings were extrapolated to zero time to give the actual temperature. Both types of measurement gave the same temperature.

2.4. Data processing and analysis

Data were processed and analyzed on a Silicon Graphics work station using the program FELIX 95.0 (Biosym, San Diego) and/or Bruker UXNMR. Data points were weighted with Gaussian resolution enhancement window function with line broadening of −20 Hz and 0.08 degree of Gaussian character in each dimension and zero filled to form 2K x 1K real matrices. Resolution in ω1 was increased by using linear prediction to extend the data from 200 to 400 points. Cross peak intensities were quantified by measuring peak-heights of cross peak slices along the ω2 axis. The population of each conformation was measured from peak volumes after baseline correction.

2.5. Nonlinear analysis

Data were fit nonlinearly to several unfolding models using the program NONLIN [11]. The data set for each residue consists of a curve for the disappearance and appearance with increasing temperature, respectively, of (f) cross peak volumes, '(f)-curves', and (u,u') peak volumes, '(u)-curves'. Fits were performed on the (f)-curves alone, and on the (f)- and (u)-curves simultaneously and are shown along with the data in Figs. 4 and 5. This procedure

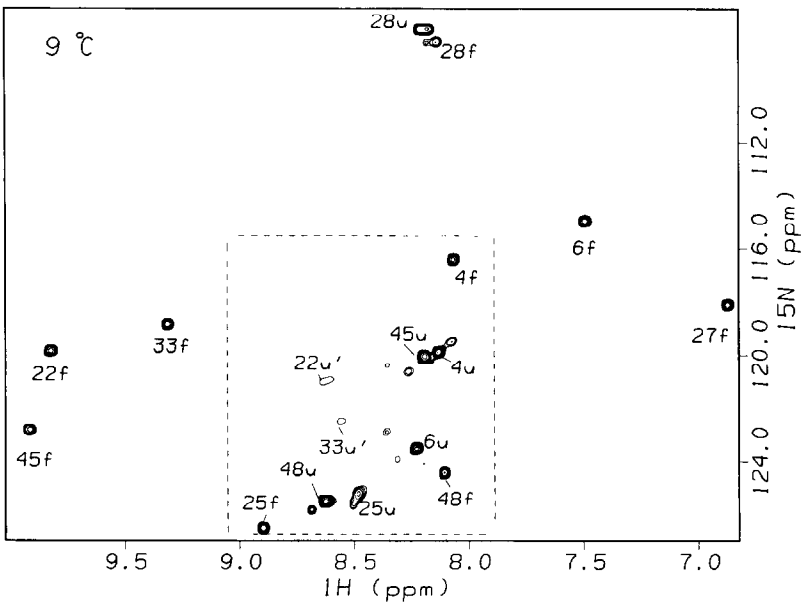

Fig. 2. ^{15}N-^{1}H HSQC spectrum of specifically labelled, partially folded $[14–38]_{Abu}$ at 9°C and pH 5.0 in ^1H_2O. Cross peaks are labelled by residues number for (f), (u) and (u') conformations, as described in Table 1. The region outlined by dashed lines is shown in Fig. 3 at different temperatures.

was used because the shape of the (u)-curve neither mirrors the (f)-curve nor appears to have a normal cooperative shape.

The (f)-curves were fit both individually and simultaneously to a two state transition, N ⇔ D

$$f = yl^{f} + (yh^{f} - yl^{f})1/1 + K \quad (1)$$

where $K = \exp((-\Delta H(1 - T/T_m))/RT$, f is the cross peak relative volume, yh^f and yl^f are the high and low plateau values, respectively, of cross peak relative volumes. (The high plateaus of (f) and (u) peaks are at low and high temperature, respectively.) This equation is modified from Ramsey and Eftink [12]. Fitting the endpoints effectively normalizes the curves. Simultaneous fits of all (f)-curves are constrained to minimize to the same value of ΔH and T_m, while their plateaus are fit individually. Parameter values from individual and simultaneous fits are given in Table 2. The pairs of (f)- and (u)-curves were fit simultaneously to equation (1) for the (f)-curves, and for the (u)-curves to the equation

$$u = yl^{u} + (yh^{u} - yl^{u})K/1 + K \quad (2)$$

where K is as defined above, u is the cross peak relative volume, yh^u and yl^u are the high and low plateaus. This requires that the f- and u-curves have the same ΔH and T_m, but allows each curve to have different endpoints. Data were also fit to several models incorporating additional thermodynamic intermediates, but as discussed below, these more complex models did not improve fits.

The initial plateau values for each pair of curves defines a pre-existing equilibrium at low temperature between (f) and (u) for that particular residue, with an equilibrium constant, K_p, given by

$$K_p = [u]/[f] = yl^{u}/yh^{f} \quad (3)$$

The variation of K_p with residue position indicates the existence of local, or segmental, motion at temperatures below global unfolding.

3. Results

3.1. Each ^{15}N-^{1}H probe reports 2 or more slowly interconverting conformations

Fig. 2 shows an ^{15}N-^{1}H HSQC spectrum at 9°C of partially folded $[14–38]_{Abu}$ with backbone amide nitrogens labelled on residues 4, 6, 22, 25, 27, 28, 33, 45, and 48. The ^{15}N labels report different sections of partially folded $[14–38]_{Abu}$ (Fig. 1). Selective labelling permits clear resolution of exchange peaks for ^{15}N-bound ^{1}H. Slow chemical exchange arises when the same proton has a different chemical shift in two or more conformations that interconvert on the time scale of milliseconds or longer, thereby giving rise to a distinct signal for each conformation [13]. Chemical shift assignments of exchange cross peaks are listed in Table 1. Exchange cross peaks of the same ^{1}H are assigned to 'more folded' conformations, (f), or to 'more disordered' conformations, (u) or (u'), as described previously [1] and in footnotes to Table 1. The (f) conformation of one ^{1}H is not the

Table 1
Amide ^{1}H and ^{15}N resonance assignments of selectively labeled $[14–38]_{Abu}$ at pH 5.0 and 1°C

Residue	(f) cross peaks [a]		(u) cross peaks [a]	
	N (ppm)	H (ppm)	N (ppm)	H (ppm)
Phe4	116.7	8.17	119.9 [b]	8.11 [b]
			119.6 [b] (u_1)	8.06 [b] (u_1)
Leu6	114.9	7.54	123.8	8.32
			122.5 [b] (u_1)	8.19 [b] (u_1)
			123.9 [b] (u_2)	8.29 [b] (u_2)
Phe22	120.0	9.88	119.0	8.67 (u')
Ala25	126.5	9.01	125.2	8.56
Ala27	118.2	6.91	121.3	7.90 (u')
Gly28	108.5	8.19	107.9	8.25
Phe33	119.0	9.35	122.5	8.62 (u')
Phe45	122.9	9.98	120.1	8.17
Ala48	124.5	8.17	125.6	8.72

[a] Exchange cross peaks of the same ^{1}H are designated (f), (u) or (u') as described in [1]. Those assigned to a more folded (f) conformer have chemical shifts outside the random coil envelope and/or NOEs indicating native-like structure. Cross peaks with no medium or long range NOEs are assigned to a fully disordered (u) conformation if their chemical shifts are different from equivalent hydrogens in random coils by ≤ 0.3 ppm, or to an unfolded, but not fully disordered, conformation (u') if these differences are > 0.3 ppm. In this paper, we use an updated reference for random coil chemical shifts at pH 5 and 25°C [17]. The ^{15}N-^{1}H probes in this table which have (u') chemical shifts with the greatest deviation from random coil are Ala27 (0.52 ppm for ^{1}H and 4.3 ppm for ^{15}N), Phe22 (0.37 ppm for ^{1}H), and Phe33 (0.32 ppm for ^{1}H).
[b] Chemical shifts are obtained at 9°C due to overlap or line broadening at 1°C. The designations (u_1, u_2) are tentative assignments of cross peaks observed at temperatures ≥ 9°C. These most likely arise from *cis–trans* isomerism of nearby prolyl peptide bonds, as discussed in the text.

same as the (f) conformation of another; rather (f) is relative to (u,u′) for the same ^{1}H. For example, the (f) cross peaks of 22 and 33 do not represent the same conformation as the (f) cross peaks of 4 and 6, as discussed below. The sequence positions of ^{15}N labels were chosen to permit us to sample representative sections of the partially folded structure, and are somewhat different from the positions labelled in [1] in order to minimize peak overlap.

The relative volumes of (f) and (u,u′) exchange peaks in HSQC spectra are a measure of the relative populations of that ^{1}H in more folded *versus* more disordered conformations. The volumes and, in some cases the number, of peaks change with temperature. As illustrated in Fig. 2, at 9°C the intensities of (u,u′) cross peaks vary dramatically with the residue, e. g., (u′) peaks of Phe22 and Phe33 are weak while (u) peaks of Leu6 and others are strong. As the temperature is raised, (u,u′) cross peak intensities increase for all residues, and for two residues, new cross peaks appear. Fig. 3 compares spectra at 4 and 27°C of the region enclosed by dashed lines in Fig. 2. At lower temperature, Fig. 3a, the (u′) peaks of 22 and 33 are small and there is only one (u) peak for Phe4 and for Leu6. At higher temperature, Fig. 3b, all (f) peaks disappear, and Phe4 has a second unfolded cross peak, labelled (u_1), while Leu6 has two additional unfolded cross peaks, labelled (u_1 and u_2).

The most likely origin of the additional (u) cross peaks of 4 and 6 are *cis* prolyl peptide isomers populated in thermally unfolded conformations. There are four prolines, at positions 2, 8, 9, and 13, all *trans* in the native BPTI crystal structure. The *cis-trans* isomerism is a slow reaction and although the *trans* isomer is favored, a *cis* content of 10–30% is common in unfolded proline-containing peptides [14]. Examples of multiple exchange cross peaks for residues near a prolyl peptide bond have been reported [15].

3.2. Populations of slowly exchanging conformations vary with residue and temperature.

The variation of cross peak volume with temperature is shown in Figs. 4 and 5; volumes of (f) peaks are indicated by circles, and (u,u′) peaks are indicated by open symbols. The sum of the volumes of both exchange peaks for the same probe at the lowest temperature, 0.6°C, is taken as one; volumes at higher temperatures are normalized to the peaks at 0.6°C. Even though errors are large at low volumes, we report volumes, rather than peak intensities at full or half height, in order to take into account differences in peak broadening. For example, the line widths of (u′) peaks of 22 and 33 are larger than the rest, and this accounts for their very low intensity,

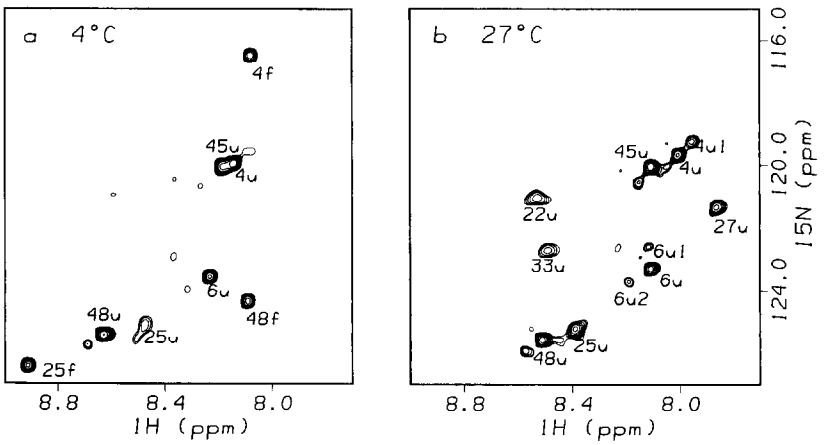

Fig. 3. ^{15}N-^{1}H HSQC spectra of $[14–38]_{Abu}$ at pH 5.0 in $^{1}H_2O$. Spectra at 4°C (a) and at 27°C (b) correspond, respectively, to partially folded and unfolded $[14–38]_{Abu}$. At high temperatures, (f) peaks are not observed while additional (u) peaks, labelled (u_1) and (u_2), appear for unfolded conformations of F4 and L6. The same region at 9°C is enclosed by dashed lines in Fig. 2.

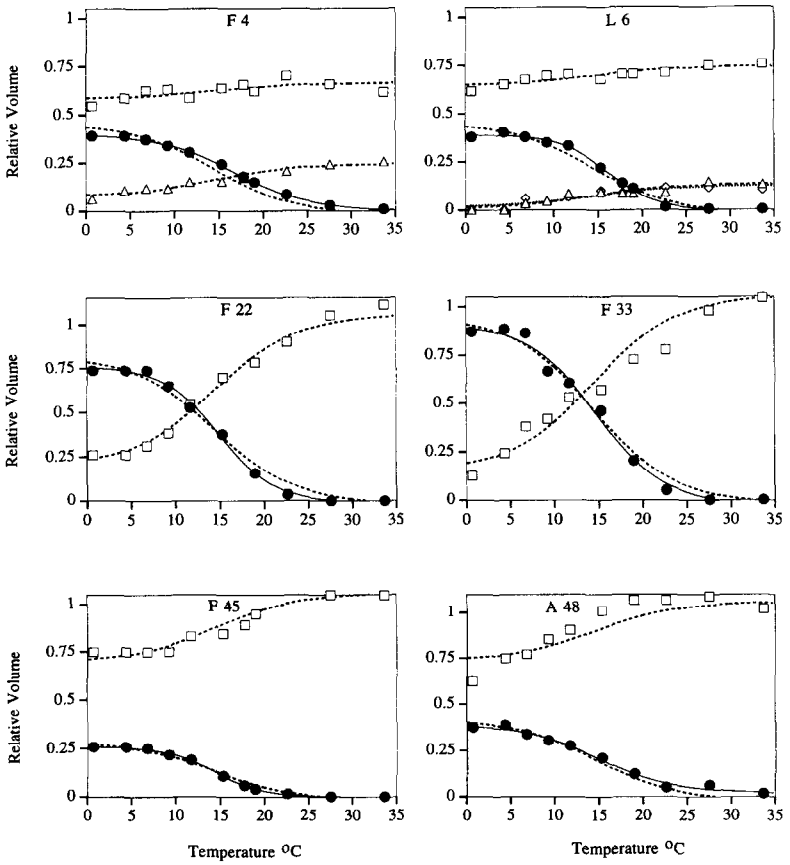

Fig. 4. Temperature dependence of cross peak volumes for six probes. Residue numbers are given at the top. ^{15}N-^{1}H HSQC spectra were taken at pH 5.0 and a sample concentration of 0.5 mM. For each residue, the '(f)-curve' is the variation of (f) cross peak volumes with temperature (circles) and the '(u)-curve' is the variation of (u,u') peak volumes with temperature (squares). Additional (u) cross peaks of residues 4 and 6 are indicated by triangles and diamonds. Solid lines are fits to individual (f)-curves. Dashed lines represent the simultaneous fit to all (f)- and (u)-curves in Fig. 4Fig. 5.

but significant volume, at low temperature (e.g., Fig. 2). This difference in line broadening among (u,u') peaks is itself interesting, and suggests that the 22 and 33 (u') peaks represent a distribution of unfolded conformers in intermediate exchange.

Two aspects of the data in Figs. 4 and 5 have a critical influence on fits to specific thermodynamic models: the shape of the (u)-curves, and the fact that the low temperature populations (plateaus) of the (f)- and (u)-curves are not 1 and 0. The deviation of the (u)-curve shape from the expected mirror image of the (f)-curve might qualitatively be interpreted as diagnostic of the presence of an intermediate in the global thermal denaturation. Two denaturation models incorporating a distinct thermodynamic intermediate I were fit to both the (u)-curves and to the combination of (f)- and (u)-curves

where PF, I and D stand for the partially folded ensemble, an intermediate and denatured state, respectively. In both models, the (u)-curve was constrained to contain the combined signal from I and D. Although exhaustive fits and simulations were performed, no set of parameter values for either of

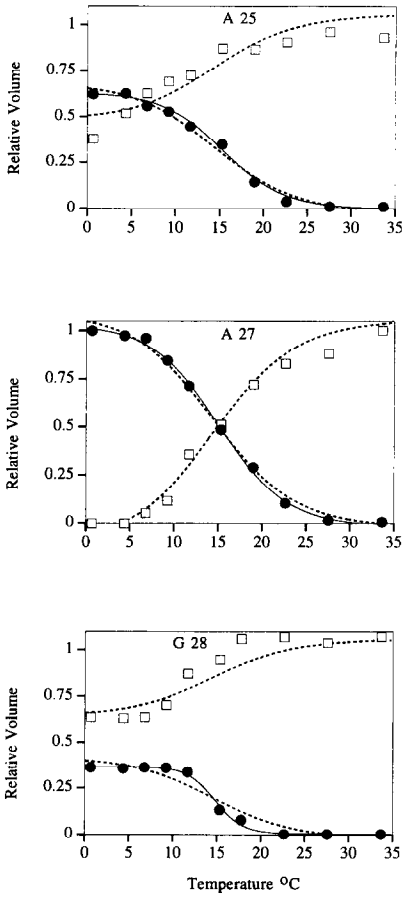

Fig. 5. Temperature dependence of cross peak volumes for residues in the turn. Ala27 is the only residue among the 9 probes that reports typical two-state behavior. NMR spectra, symbols, and solid or dashed lines are the same as in Fig. 4.

these models could be found to reproduce the shape of the (u)-curves. We conclude that the presence of additional intermediate states is unlikely. The deviation from a conventional S-shaped curve may be due to conformational drifting of D, as described in Dill and Shortle's variable two-state model [16].

The clear plateaus at low temperature of the (f)-curves indicate that we are not simply starting the low temperature measurements in the middle of the denaturation, i.e., 0.6°C is on the low temperature baseline of global unfolding. This is consistent with the CD unfolding curves ([2] and below). We interpret the initial (f)-curve plateaus as indicative of a pre-existing equilibrium in partially folded protein.

For *each* probe, the partially folded ensemble, PF, consists of two conformations, one more folded, P_f, and one more disordered, P_d. Further, P_f and P_d interconvert slowly, $\geq$ milliseconds. Restated schematically

$$PF = P_f + P_d \quad \text{and} \quad P_f \rightleftharpoons P_d \qquad \text{Scheme 1}$$

Thermal denaturation experiments, monitored by CD, also show that the partially folded ensemble undergoes global unfolding

$$PF \rightleftharpoons D \qquad \text{Scheme 2}$$

The magnetic environments, and therefore the chemical shifts, of ^1H are expected to be similar in P_d conformations of the partially folded state and in all conformations of the D state. Therefore, while the (f)-curves monitor only the population of P_f conformations, the (u)-curves monitor the combined populations of P_d and D conformations. The fact that each positional monitor (each ^{15}N-^1H) sees different initial populations of P_f and P_d indicates that varying segments of the molecule are disordered to different extents prior to the onset of thermal denaturation.

The interpretation represented by Schemes 1 and 2 was modeled both by fits to (f)-curves and by simultaneous fits to (f)- and (u)-curves. The fitting parameters with the confidence intervals for thermal unfolding of individual residues are given in Table 2. Since (f) cross peaks represent only the signal of P_f species (Scheme 1), (f)-curves should monitor only the disappearance of PF during global unfolding. The (f) data are well described by simple two-state unfolding, performed as described in Methods, and shown as solid lines in Figs. 4 and 5. When all (f)-curves are fit simultaneously, the T_m is 15.1°C and the apparent $\Delta H(T_m)$ is 47 kcal mol^{-1}. The similarity of the fitted values for T_m and ΔH, and the overlap of most of the confidence intervals, support the interpretation that the disappearance of any (f) cross peak reports the same global unfolding transition. Simultaneous fits of (f)- and (u)-curves, also described in Methods, are shown in Figs. 4 and 5 by dashed lines. They provide reasonably good descriptions of the (u)- as well as the (f)-curves. The

Table 2
Fitting parameters for thermal unfolding of $[14–38]_{Abu}$ at pH 5.0. Fits for (f) curves of individual residues [a]

Residue	Fraction (f)	T_m (f) (°C)	ΔH (kcal/mol)	σ [b]
F4	0.40 (0.45) [c]	16.9 (16.4–17.4) [d]	39 (34–43) [d]	3×10^{-5}
L6	0.39 (0.45)	16.2 (15.6–17.0)	60 (48–73)	2×10^{-4}
F22	0.76 (0.82)	15.0 (14.0–16.0)	52 (41–63)	6×10^{-4}
A25	0.62 (0.68)	15.4 (14.0–16.6)	48 (34–62)	7×10^{-4}
A27	1.0 (1.1)	15.1 (14.0–16.0)	43 (36–49)	4×10^{-4}
G28	0.36 (0.42)	14.8 (14.2–15.3)	104 (77–152)	2×10^{-3}
F33	0.91 (0.95)	14.8 (13.0–16.6)	42 (28–57)	2×10^{-3}
F45	0.26 (0.28)	14.3 (14.0–14.6)	59 (53–65)	2×10^{-5}
A48	0.39 (0.42)	15.0 (13.0–17.0)	39 (25–53)	3×10^{-4}
Simultaneous fits				
All (f)		15.1 (14.6–15.7)	47 (40–54)	7×10^{-4}
All data sets [e] (f) and (u)		14.6 (13.3–15.8)	36 (25–44)	2×10^{-3}

[a] Fits are shown by solid lines in Fig. 4Fig. 5. [b] σ is the variance of the fit (sum of residuals squared/degrees of freedom). [c] Values in parentheses are from the simultaneous fits to all the data. Confidence intervals for the fraction (f) were all $< \pm 10\%$ for individual fits and $< \pm 14\%$ for the fits to all (u) and (f) data. [d] Numbers in parentheses are the 67% confidence intervals from the fits. [e] Fits are shown by dashed lines in Fig. 4Fig. 5.

fits are consistent with the interpretation that (f) cross peaks report the more folded conformations of the partially folded ensemble at all temperatures, and that (u) cross peaks at low temperatures represent the more disordered conformations of the partially folded ensemble, but at higher temperatures report both conformations of the partially folded ensemble and conformations of the globally denatured state.

3.3. T_m values are not the same for thermal unfolding monitored by CD and NMR

Thermal unfolding monitored by circular dichroism at pH 6 is reported in [2] to be 19°C. This value is higher by about 4°C (depending on the amide group) than T_m obtained from 2-state fits to the (f)-curves obtained at pH 5.0 (Figs. 4 and 5, and Table 2). To investigate this further, CD folding/unfolding experiments were carefully repeated at pH 5.0 in 50 mM sodium acetate and at pH 6.4 in water. The CD data are well fit by a 2-state model and T_m values of 18.7°C at pH 5.0 and 19.2°C at pH 6.4 were obtained (data not shown), in close agreement with the earlier data in [2].

In the CD experiments, temperature is regulated by water circulating around a jacketed cell, and measured directly in the cell by means of a thermistor. In the NMR instrument, temperature is controlled by circulating warm air mixed with precooled N_2 gas. The temperature of NMR samples was determined in two ways, as described in Methods, from a methanol sample inserted just before and/or after the NMR run, and from direct insertion of a thermistor into a duplicate NMR sample after 1 and 5 h of acquisition pulsing. Both types of measurement gave the same, highly reproducible, temperature reading, excluding the possibility that the high power spinlock pulses used in the NMR experiment result in local heating in the NMR sample, which is not measured by the methanol temperature determination. We therefore consider it highly likely that the difference between the CD and NMR-detected T_m values is real.

4. Discussion

4.1. Segmental motions

Conformations of partially folded $[14–38]_{Abu}$ interconvert on time scales longer than milliseconds. Two types of slow fluctuations of partially folded states are apparent, cooperative global unfolding, and segmental motions. The latter corresponding to the process in Scheme 1, are monitored by exchange cross peaks at low temperature (1–6°C). Nine strate-

gically placed ^{15}N residues provide microscopic probes of relative populations of more folded (f), and more disordered (u,u'), conformations that interconvert by local motions at low temperature.

For ^{1}H's in residues in the anti-parallel strands, the β-bridge and the first turn of the C-terminal helix, the (f) conformations are native-like, and exhibit numerous NOEs also observed in native BPTI; these regions are shown as ribbons in Fig. 1. However, it is clear that they do *not* fluctuate as a group. Probes in the core (22 and 33) report 75–95% native-like and 5–25% disordered conformation, while those in the β-bridge and first turn of the helix (45 and 48) report 25–40% native-like conformations and the rest disordered (Table 3). For 4 and 6, which in native BPTI are in the N-terminal helix, the (f) conformations of the partially folded state have no indication of native-like structure; that is, *both* (f) and (u) conformations of these residues are more disordered than native [1].

The three turn residues vary widely in the percentage of (f) and (u,u') conformations at low temperature (Table 3). The major conformation for Ala25 is (f), but for Gly28, it is (u), implying that the turn is flexible. Ala27 is unusual in being the only residue that reports canonical two-state behavior; it is 100% (f) at low temperature, and exhibits symmetrical and cooperative thermal denaturation for both the (f)- and (u)-curve (Fig. 5). In native BPTI, the type I β-turn of 25–28 has a hydrogen bond between the carbonyl of Asn24 and Ala27 NH, which has a chemical shift of 6.9 ppm. The (f) cross peak of Ala27 NH in partially folded $[14–38]_{Abu}$ is also shifted upfield to 6.9 ppm, suggesting that the reason for the 100% population of Ala27 (f) is a hydrogen bond between 24 and 27. Further, the chemical shift of the (u') cross peak of Ala27, 7.9 ppm, is also significantly shifted upfield from 8.5 ppm, the value in random peptides [17]. This implies that the 24-27 hydrogen bond persists in a significant fraction of thermally unfolded conformers.

We conclude that in different parts of the molecule, populations of folded versus disordered conformations vary. Thus, partially folded $[14–38]_{Abu}$ undergoes segmental motions, i.e., local fluctuations that are not components of its global cooperative unfolding [1]. Characterization of multiple partially folded conformations in slow exchange has not before been possible. Slow exchange has been reported for the N-terminal SH_3 domain of drk [18], but in this case global unfolding, not segmental motion, was observed. We view partially folded $[14–38]_{Abu}$ as a model for an early folding intermediate in BPTI (regardless of which disulfide is formed first [1,5]). The dynamic structure of the partially folded protein, represented schematically in Fig. 1, is fully consistent with the conclusions of Hilser and Freire [19] on the order in which sections of BPTI fold.

Segmental fluctuations, not linked to global unfolding, are observed for folded proteins by native-state hydrogen isotope exchange kinetics [20–22]. One interesting distinction between local conformational equilibria observed in native proteins versus those in partially folded $[14–38]_{Abu}$ is that the latter

Table 3
Relative populations of low temperature conformations reported by exchange cross peak volumes of ^{15}N-^{1}H probes [a]

Probe	Structure in native BPTI	% Native-like	% Not r.c., but not native-like [b]	% Disordered
4,6	N-terminal helix		40 (f)	60 (u)
22,33	Anti-parallel β-sheet	75–95 (f)	5–25 (u')	
37 [c]	Polar interaction with Y35 ring	20 (f)	60 (u')	20 (u)
45,48 [d]	β-Bridge, first turn of C-terminal helix	25–40 (f)		60–75 (u)
25	Turn		60 (f)	40 (u)
27	Turn		100 (f)	
28	Turn		40 (f)	60 (u)
56 [c]	End of C-terminal helix			100 (u)

[a] Percentages are approximate relative volumes of exchange cross peaks at 0.6°C taken from Table 2. The designation of the cross peaks as (f), (u) and (u') is explained in the text and in footnotes to Table 1. [b] Random coil is abbreviated r.c. [c] Gly37 and Gly56 data are from earlier studies [1]. Their NMR data are not reported here. [d] The relative (f) and (u) populations of 48 reported here replace those in [1]; in the earlier work 48 NH overlaps other peaks, a problem eliminated by this set of ^{15}N labels.

apparently have large energy barriers between 2 or 3 conformations, which are therefore in slow exchange. In NMR spectra of native proteins, there is no indication of locally fluctuating conformations in slow exchange; this may be because local mobility is between conformers that are in rapid equilibrium and therefore give an averaged NMR signal and/or because the disordered conformations represent < 5–10% of the native species.

4.2. Global unfolding

Global unfolding is shown by thermal denaturation monitored by CD and by (f) cross peaks in NMR spectra. As temperature is raised above about 9°C, the (u,u′) cross peaks contain intensity both from more disordered conformations of the partially folded ensemble, corresponding to P_d in Scheme 1, and from the conformations of the unfolded state, corresponding to D in Scheme 2. There is an increasing contribution from D to the (u,u′) cross peak as the temperature is increased. However, at all temperatures, the (f) cross peaks contain intensity only from the more folded conformations, corresponding to P_f in Scheme 1, and therefore provide a microscopic probe of cooperative thermal unfolding.

Interestingly, there is a difference between thermal unfolding curves measured by NMR and equivalent curves monitored by a macroscopic property, in this case, CD. The CD-derived T_m is 19°C, while fits to the NMR-derived T_m values are around 15°C for most residues. Further, two residues, 4 and 6, have NMR T_m values in the range 16–17°C (Table 2). The higher NMR T_m values of 4 and 6, compared to the other residues, suggest that the N-terminus of partially folded $[14–38]_{Abu}$ unfolds at higher temperature than the core, a reasonable possibility since 4 and 6 are more disordered than the other residues measured. Significant discrepancies in T_m of this type are indicative of a deviation from strict two-state behavior. One explanation is that thermal denaturation of partially folded $[14–38]_{Abu}$ is a variable two-state transition, of the type described by Dill and Shortle [16], in which the D state ensemble varies with temperature. It is possible that in thermal unfolding of $[14–38]_{Abu}$, the D state and/or the PF state vary with temperature during the transition.

4.3. Residues with more than two conformations in slow exchange

We have shown previously that in partially folded $[14–38]_{Abu}$, Gly37 NH has 3 exchange peaks, one minor and native-like, one major and clearly non-random but also non-native, and a third minor peak whose chemical shift is indistinguishable from fully disordered (Table 3, and [1]). The Gly37 case is particularly interesting because 37 NH in native BPTI has a very unusual upfield chemical shift arising from the polar interaction of 37 NH with the center of the Y35 ring [23].

The use of selective labelling also reveals other ^{1}H's for which there are 3 or 4 slow exchange peaks, but these are different from Gly37, as they apparently arise only in the fully unfolded state. As global unfolding proceeds, more than one slowly interconverting, disordered conformation is populated. Near their major (u) peaks, Phe4 and Leu6 NH show additional peaks, which we designate (u_1) and (u_2) (Fig. 3); their change in volume with temperature is shown in Fig. 4 by triangles and diamonds. These most likely arise from the *cis-trans* proline isomerism of the prolyl peptide bonds of residues 2, 8, 9, and 13, the only prolines in the molecule. Conformational heterogeneity due to *cis-trans* isomerism is only found in D conformations of the N-terminal residues and does not explain the slow exchange at low temperature.

4.4. Positional energetics

The data of Figs. 4 and 5 have novel thermodynamic and statistical mechanical character. Standard thermodynamic analysis of denaturation profiles concentrates on the transition region of the unfolding curves. As in this study, denaturation profiles are generally fit to specific baselines, often ones with significant slopes; however, unlike conventional unfolding experiments, here we get information from the baselines. (Differential scanning calorimetry is an example of another method in which information is also derived from the baselines.) Because the NMR experiments separately monitor the relative populations of the two endstates of the unfolding reaction, baseline positions before and after thermal melting

provide thermodynamic information about the relative populations of the endstates before and after the thermal denaturation. Further, the microscopic nature of the NMR signal means that the thermodynamic information gained by the position of the baselines is specific to each particular positional monitor. If the protein behaves as a completely cooperative unit, each probe would always report the same fractional population distribution. The fact that this is not the case is indicative of segmental motion within the protein, as discussed above and in [1].

From a statistical mechanical point of view, segmental motion results in each positional monitor reporting a biased view of the population distribution of the conformational ensemble of the protein. To take a very simple example, consider Q as the partition function describing all individual molecular states of a protein. If the protein has two segments and each segment can separately occupy different conformations, then the ensemble of all molecular states is

$$Q = pAB + pAb + paB + pab$$

where A and a (and B and b) are the two conformations of each segment, pAB is the population of molecules with segments in conformations A and B, and similarly for the other terms. A probe in segment A(a) would monitor the same ensemble from a subjective perspective, $Q_{(A)} = pA + pa$. A probe in segment B(b) would likewise sample its own specific population distribution. Any particular probe reports only its own conformational distribution, but the combined reports of an array of monitors provide detailed sub-molecular information about the molecule. As the precision of such data increases, the positional thermodynamic information they can provide, and the manner in which it sums to the macroscopic thermodynamic behavior, may lead to new insight into the energetics of protein folding.

4.5. Conclusions

For $[14–38]_{Abu}$ BPTI, ^{15}N-bound amide hydrogens at sites distributed along the backbone monitor two or more conformations that interconvert on a time scale $\geq$ milliseconds. Relative populations of interchanging conformations at temperatures between 1–35°C were determined from ^{15}N-^{1}H HSQC cross peak volumes, and fit to a thermodynamic model in which segmental fluctuations are a pre-equilibrium distinguished from global unfolding. The partially folded state undergoes both local internal motions at low temperature, as well as global, cooperative unfolding at higher temperature (Schemes 1 and 2). At 1–6°C, conformational fluctuations of partially folded $[14–38]_{Abu}$ are segmental, and different regions of the protein vary in the extent to which they are disordered. As the temperature is raised, the partially folded ensemble undergoes global unfolding with a T_m around 15°C, and the conformations sampled include globally unfolded species.

$[14–38]_{Abu}$ BPTI provides a rare system for characterization of multiple partially folded conformations in slow exchange. Each region of the molecule undergoes independent motions which result in an ensemble in which residues 18–24 and 29–35 are in a native-like antiparallel sheet $> 75\%$ of the time, residues 44–48 are in a native-like conformation 25–40% of the time and residues 1–17 and 48–57 fluctuate between two non-native conformations. Gly37 NH samples three slowly interconverting conformations.

While the thermal unfolding of partially folded [14–38] is clearly global and displays features of a cooperative transition, it does not behave in a standard two-state manner. The availability of heterologous, microscopic probes of global denaturation offers new positional insight into the unfolding thermodynamics. Deviation from typical two state folding/unfolding is indicated by a difference in T_m measured by CD *versus* NMR, by the difference in NMR T_m values for different residues, and by the observation that (u)-curves are not simple reciprocals of (f)-curves. These data are consistent with the variable two-state models described in Dill and Shortle [16] and may arise from temperature-dependent conformational drift of the D state and/or the partially folded state.

References

[1] E. Barbar, G. Barany and C. Woodward, Dynamic structure of a highly ordered β-sheet molten globule: Multiple conformations with a stable core, Biochemistry, 34 (1995) 11423–11434.

[2] M. Ferrer, G. Barany and C. Woodward, Partially folded, molten globule and molten coil states of bovine pancreatic trypsin inhibitor, Nature Struct. Biol., 2 (1995) 211–218.
[3] H. Pan, E. Barbar, G. Barany and C. Woodward, Extensive non-random structure in reduced and unfolded bovine pancreatic trypsin inhibitor, Biochemistry, 34 (1995) 13974–13981.
[4] M. Dadlez and P. Kim, Detection of a third native one-disulfide intermediate in the folding of BPTI, Nature Struct. Biol., 2 (1995) 6–12.
[5] E. Barbar, G. Barany and C. Woodward, Unfolded BPTI variants with a single disulfide bond have diminished non-native structure distant from the crosslink, Folding and Design 1 (1996) 65–76.
[6] G. Barany, C.M. Gross, M. Ferrer, E. Barbar, H. Pan and C. Woodward, Optimized methods for chemical synthesis of bovine pancreatic trypsin inhibitor analogues, Techniques in Protein Chemistry VII, (Marshak, D., ed) Academic Press, San Diego, 1996, pp. 503–514.
[7] B. A. Messerle, G. Wider, G. Otting, C. Weber and K. Wüthrich, Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with H_2O solutions, J. Magn. Reson., 85 (1989) 608–613.
[8] G. Bodenhausen and D. Ruben, Natural abundance Nitrogen-15 NMR by enhanced heteronuclear spectroscopy, Chem. Phys. Lett., 69 (1980) 185–189.
[9] A. J. Shaka, J. Keeler, T. Frenkiel and R. Freeman, An improved sequence for broadband decoupling: WALTZ-16, J. Magn. Reson., 52 (1983) 335–338.
[10] A.L. Van Geet, Calibration of the methanol and glycol nuclear magnetic resonance thermometers with a static thermistor probe, Anal. Chem., 40 (1968) 2227–2229.
[11] M.L. Johnson and S.G. Frasier, Nonlinear least-squares analysis, Methods Enzymol., 117 (1985) 301.
[12] M.R. Eftink and G.D. Ramsay, Analysis of multidimensional spectroscopic data to monitor unfolding of proteins, Methods Enzymol., 240 (1994) 615–645.
[13] J. Sandström, in Dynamic NMR, Academic Press, New York, 1982.
[14] H.N. Cheng and F.A. Bovey, *Cis-trans* equilibrium and kinetic studies of acetyl-L-proline and glycyl-L-proline, Biopolymers, 16 (1977) 1465–1472.
[15] J. Kemmink, C.P.M van Mierlo, R.M. Scheek and T.E. Creighton, Local structure due to an aromatic-amide interaction observed by ^1H NMR spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor, J. Mol. Biol., 230 (1993) 312–322.
[16] K.A. Dill and D. Shortle, Denatured states of proteins, Annu. Rev. Biochem., 60 (1991) 795–825.
[17] D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges, and B. D. Sykes, ^1H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigation of nearest-neighbor effects, J. Biomol. NMR, 5 (1995) 67–81.
[18] O. Zhang and J. Forman-Kay, Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, Biochemistry, 34 (1995) 6784–6794.
[19] V. Hilser and E. Freire, Structure based calculation of the equilibrium folding pathway of proteins. Correlations with hydrogen exchange protection factors, J. Mol. Biol., 262 (1996) 756–772.
[20] B. Hilton, K. Trudeau and C. Woodward, Hydrogen exchange rates in pancreatic trypsin inhibitor are not correlated to thermal stability in urea, Biochemistry, 20 (1981) 4697–4703.
[21] K.-S. Kim, J. Fuchs and C. Woodward, Hydrogen exchange identifies native-state motional domains important in protein folding, Biochemistry, 32 (1993), 9600–9608.
[22] K.-S. Kim and C. Woodward, Internal flexibility and global stability of proteins: Effect of urea on hydrogen exchange rates of BPTI, Biochemistry, 32 (1993), 9609–9613.
[23] E. Tüchsen, and C. Woodward, Assignment of Asn44 side chain primary amide ^1H NMR resonances and the peptide amide N^1H resonance of Gly37 in BPTI, Biochemistry, 26 (1987) 1918–1925.