Extensive Nonrandom Structure in Reduced and Unfolded Bovine Pancreatic Trypsin Inhibitor†

Hong Pan,‡ Elisar Barbar,‡ George Barany,§ and Clare Woodward*,‡

Department of Biochemistry, University of Minnesota, St. Paul, Minnesota 55108, and Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received July 6, 1995; Revised Manuscript Received August 15, 1995®

ABSTRACT: Two-dimensional ^{1}H NMR spectra of an analog of reduced BPTI at pH 4.5, 1 °C, have been assigned. Spectra indicate considerable conformational averaging, as expected for a flexible, unfolded protein. The presence of extensive nonrandom structure is detected by the presence of NH_i-NH_{i+1} and aromatic−aliphatic NOEs. Sequential amide−amide NOEs indicate that turn-like conformations are significantly populated at 18 pairs of residues along the chain. Many of these are located in a turn, loop, or helix in native BPTI, but six are observed for contiguous pairs in the segment composed of residues 29−35, which in native BPTI constitute a strand of extended sheet. A novel finding for unfolded proteins is our observation of NOEs implying non-native hydrophobic interactions. Multiple aromatic−aliphatic NOEs are observed for pairs of residues that are within 1−3 residues of each other. Most are non-native and involve residues in both strands of the central antiparallel strand−turn−strand of native BPTI comprised of residues 18−35. All NOEs reported for oligopeptides spanning the BPTI sequence [Kemmink, J., & Creighton, T. (1993) *J. Mol. Biol. 234*, 861−878] are observed in reduced BPTI, but many others are present as well. Similar spectra are obtained for naturally occurring BPTI reduced by dithiothreitol, BPTI with cysteines replaced by α-amino-*n*-butyric acid, and BPTI mutant F45A reduced by dithiothreitol. The indications of numerous turn-like conformations and of hydrophobic interactions are consistent with earlier reports that reduced BPTI is a molten coil which is collapsed to some extent but not as much as native, and which has exposed, clustered hydrophobes [Ferrer, M., Barany, G., & Woodward, C. (1995) *Nature Struct. Biol. 2*, 211−217]. Comparison of the NOEs in reduced BPTI to those in a model for early BPTI folding intermediates suggests a significant role for non-native interactions in initial steps of BPTI folding. A variant of reduced BPTI, in which all cysteines are replaced with *S*-[^{13}C]methylcysteine, also displays chemical shift dispersion in HMQC-detected resonances of the [^{13}C]methyl protons. The dispersion is lost by addition of guanidine hydrochloride, indicating that nonrandom structure in reduced BPTI is disrupted by the denaturant.

Structures of protein unfolded states have important implications for the protein folding problem (Dill & Shortle, 1991; Dobson, 1992; Shortle, 1993). Nonrandom structure in an unfolded protein indicates which sequence-specific contacts are favored at the onset of folding and identifies likely sites for nucleation of early folding events. Apolar contacts in unfolded states may identify sites of hydrophobic collapse, a process proposed to be fundamental in protein folding (Dill, 1990). Structure in unfolded states may help determine the role of sequence-specific secondary structural propensities in folding. Non-native structure in unfolded states implies that initial folding steps involve loss of interactions in unfolded states as well as gain of interactions in early intermediates. Structural differences among unfolded states of the same protein produced by alternative denaturants or mutations demonstrate that full interpretation of protein folding thermodynamics and kinetics must take into consideration the variability of unfolded states.

Atomic level characterization of nonrandom structure in unfolded states of several proteins in the 57−110 amino acid size range has been accomplished recently by NMR. In the urea-denatured 434 repressor N-terminal domain, NOEs within a stretch of hydrophobic amino acids indicate formation of a native-like hydrophobic cluster (Neri et al., 1992). In chemically denatured FK506 binding protein, defined regions of secondary structure are indicated by NOEs. While most correspond to similar secondary structure in the folded protein, a C-terminal region forms different secondary structure in the folded and unfolded states (Logan et al., 1994). In an SH3 domain, the unfolded state studied in equilibrium with the native state has a number of residues with NOEs indicating turn-like conformations; in the native state a stretch of these residues are in stable β-strands (Zhang & Forman-Kay, 1995). Preliminary studies of barnase unfolded at low pH suggest sequential amide−amide NOEs between residues located in early folding regions (Arcus et al., 1994).

Here we report full assignments of ^{1}H NMR spectra of unfolded bovine pancreatic trypsin inhibitor (BPTI) at pH 4.5 and 1 °C. BPTI is unfolded by cleavage of all three disulfides, either by replacement of cysteines with α-amino-*n*-butyric acid (Abu), by reduction, or by reduction followed by methylation. We find evidence of extensive non-native structure in unfolded BPTI, including sequential amide−amide NOEs indicative of turn-like conformations and aromatic−aliphatic, side chain−side chain, and side chain−backbone NOEs. Some aromatic−aliphatic NOEs are native-

† This work is supported by NIH Grants GM26242 and GM51628.
* Author to whom correspondence should be addressed.
‡ University of Minnesota, St. Paul.
§ University of Minnesota, Minneapolis.
® Abstract published in *Advance ACS Abstracts,* October 1, 1995.

0006-2960/95/0434-13974$09.00/0

like, but many others indicate specific, non-native hydrophobic interactions in unfolded BPTI. From a consideration of the NMR-detected structures of reduced BPTI and of $[14-38]_{Abu}$, a model for BPTI early folding intermediates (Ferrer et al., 1995; Barbar et al., 1995), we suggest that early events in BPTI folding have the following features. (1) Reduced BPTI is collapsed enough to have turn-like conformations all along the backbone. (2) Non-native hydrophobic interactions in reduced BPTI are primarily in the 18−35 antiparallel strand−turn−strand. They are specific and apparently function to maintain residues 18−35 in unpacked conformations; residues that fold into strand 18−24 are stabilized in extended conformations while residues that fold into strand 29−35 are stabilized in turn-like conformations. (3) When any single native disulfide bond is formed, the 18−35 antiparallel strand−turn−strand folds first, and the non-native hydrophobic interactions are replaced with native-like interactions involving the same aromatic residues. (4) Native-like interactions of residues 25−28 in reduced BPTI point to this turn as a folding initiation site. (5) These early events occur no matter which native disulfide is formed first and no matter what the subsequent order of disulfide bond maturation.

MATERIALS AND METHODS

Naturally occurring BPTI was purchased from Novo Industries, Copenhagen (apotinin Novo, ultrapure grade). BPTI mutant F45A was produced by site-directed mutagenesis in an expression system developed by Goldenberg (1988) and described in Housset et al. (1991). $[R]_{Abu}$ was chemically synthesized by automated Fmoc solid-phase synthesis and purified as described in Ferrer et al. (1992) and Barany et al. (1995). These papers also give procedures for production and purification of reduced naturally occurring BPTI ($[R]_{SH}$) and reduced BPTI mutant F45A ($[R,F45A]_{SH}$). $[R]_{Smc}$ with all cystines converted to *S*-[^{13}C]methylcysteine and a variant of $[R]_{Smc}$ with the ^{13}C label only in methyl groups attached to cysteines 14 and 38 were prepared by methods detailed in Barany et al. (1995). Protein purities were determined by analytical HPLC and capillary zone electrophoresis. Ion electrospray mass spectrometry of purified proteins is in good agreement with calculated values. The secondary structure of residues in the native BPTI crystal structure is defined by the rules of Kabsch and Sander (1983).

NMR samples were prepared by dissolving lyophilized, pure protein in 90% ^1H_2O/10% ^2H_2O or 99% ^2H_2O preadjusted to pH 3; the dissolved sample was then adjusted to pH 4.5. The pH adjustments were made with HCl or ^{2}HCl. For $[R]_{SH}$ and $[R,F45A]_{SH}$ samples, a 20-fold excess of deuterated β-mercaptoethanol was included. In all NMR samples, 0.02% azide was added. NMR spectra were obtained at pH 4.5, 1 °C; the protein concentration was usually 0.6 mM, but a few samples were at 0.2 mM. Spectra obtained with 0.6 and 0.2 mM samples show no differences in line widths, chemical shifts, or NOEs. To rule out aggregation or any other deterioration of the samples, 2D spectra of the same sample taken several days apart were compared and found to be identical; likewise, 1D spectra of samples at 0.6 and 0.06 mM protein concentrations were identical. Samples taken before and after the NMR experiments behaved identically on reverse-phase HPLC. NMR experiments were performed on Bruker AMX-500 and -600 spectrometers with spectral windows of 6024 and 7246 Hz, respectively, and water suppression was accomplished by presaturation during the relaxation delay of 1 s. FIDs were acquired with 2048 complex data points accumulating at least 128 scans. Routinely 400−512 t_1 were obtained using TPPI for quadrature detection in this dimension (Marion & Wüthrich, 1983). Clean TOCSY (Griesinger et al., 1988) spectra were acquired with mixing times between 60 and 85 ms and a spin-lock power of 11 kHz. NOESY (Jeener et al., 1979) spectra with mixing times of 200, 300, or 400 ms were collected. Assignments and/or NOEs were determined from TOCSY and NOESY spectra with different mixing times and in ^1H_2O as well as in ^2H_2O. ^{1}H chemical shifts are relative to internal (trimethylsilyl)propanesulfonic acid (TSP) at 0 ppm. 1D HMQC spectra (Bax et al., 1983) were obtained at 500 MHz with a ^{13}C frequency of 125.76 MHz. ^{13}C decoupling during acquisition employed the WALTZ-16 sequence (Shaka et al., 1983). In guanidine hydrochloride unfolding experiments, ^{1}H chemical shifts were referenced to external TSP in a coaxial stem insert.

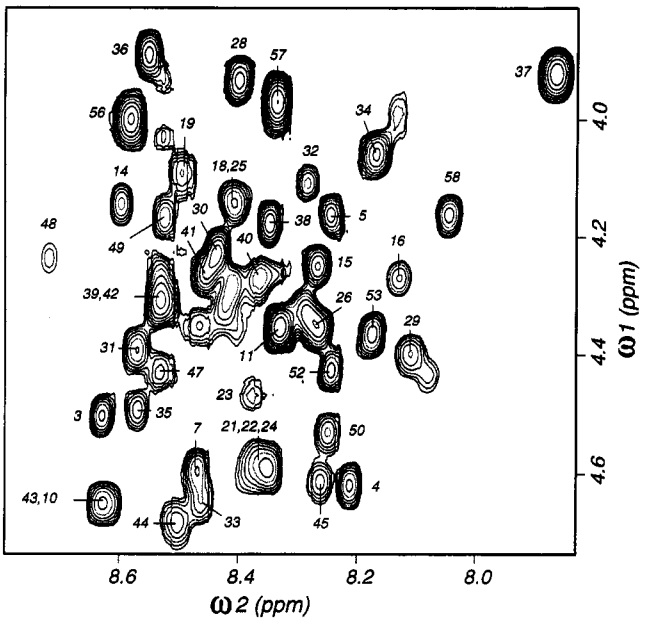

FIGURE 1: TOCSY spectrum of $[R]_{Abu}$ in ^1H_2O at pH 4.5, 1 °C. The NH−CαH region of the spectrum taken at 600 MHz is shown. Full assignments are available as supporting information.

RESULTS

An analog of reduced BPTI, $[R]_{Abu}$, with all Cys residues replaced by Abu, was prepared by chemical synthesis (Ferrer et al., 1992). Abu has the same number of heavy atoms as Cys and a methyl group in place of sulfur. $[R]_{Abu}$ is a good model for reduced BPTI with all cystines in the thiol form, the predominant titration state at physiological pH. In 2D ^{1}H NMR spectra of $[R]_{Abu}$ at pH 4.5 and 1 °C, most peaks cluster in and near the random-coil region, but they are sufficiently resolved to permit assignments from the combined information of spin systems in TOCSY spectra and CαH$_i$−NH$_{i+1}$ connectivities in NOESY spectra. Many assignments are given in Figure 1, and a table of complete backbone and side chain assignments is available as supporting information. Strong CαH$_i$−NH$_{i+1}$ connectivities are observed for almost all residues although for a few residues CαH$_i$−NH$_{i+1}$ NOEs are either absent or overlapped with other peaks. There are no NOEs indicating stable secondary

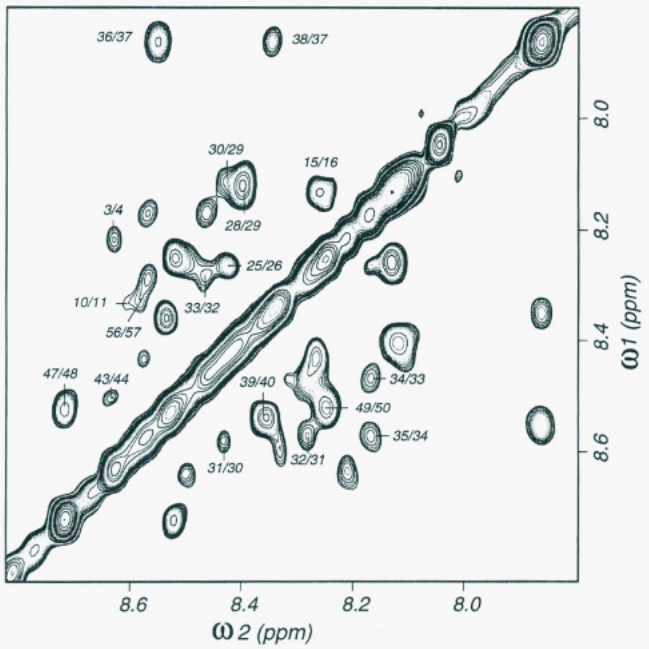

FIGURE 2: NOESY spectrum of $[R]_{Abu}$ in ^1H_2O at pH 4.5, 1 °C, taken at 600 MHz, with a mixing time of 200 ms. Cross peaks show NH_i-NH_{i+1} NOEs.

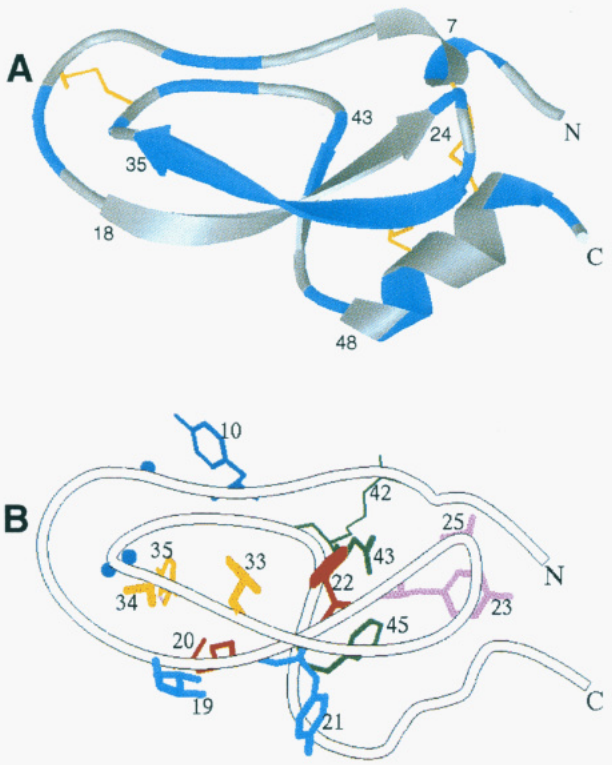

FIGURE 3: Crystal structure of native BPTI. In (A), backbone segments in dark blue are those with NH_i-NH_{i+1} NOEs in $[R]_{Abu}$. The NOEs are identified in Figure 2. Yellow bonds show locations of the three disulfide bonds in native BPTI, between cysteines 5−55, 14−38, and 30−51. In (B), side chains in the same color have interresidue NOEs between them in $[R]_{Abu}$; the NOEs are listed in Table 1. Dark blue spheres show the locations of Gly 12, Gly 36, and Gly 37, which make NOEs with Tyr 10 or Tyr 35 ring atoms.

structure. However, a substantial number of NH_i-NH_{i+1} NOEs as well as aromatic−aliphatic NOEs indicate considerable persistent structure in $[R]_{Abu}$.

The NOESY spectrum in Figure 2 shows NH_i-NH_{i+1} NOEs in $[R]_{Abu}$. Amide−amide $(i,i+1)$ NOEs occur between residues which, in native BPTI, are in the aperiodic loops

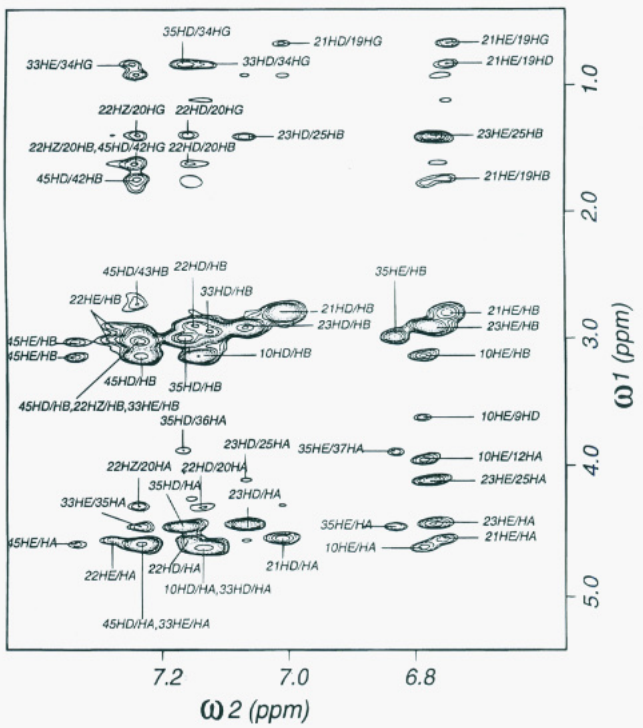

FIGURE 4: NOESY spectrum of $[R]_{Abu}$ in ^2H_2O at pH 4.5, 1 °C, taken at 500 MHz, with a mixing time of 400 ms. Cross peaks show NOEs between atoms on aromatic residues (ω_2) and aliphatic atoms (ω_1). Unlabeled cross peaks with chemical shifts at 0.95 pmm in the ω_1 dimension arise from CγH of Abu side chains.

(10/11, 15/16, 36/37, 37/38, 39/40, 43/44), one strand of the antiparallel β-sheet (29/30, 30/31, 31/32, 32/33, 33/34, 34/35), the turn between the two strands of antiparallel β-sheet (25/26 and 28/29), the N-terminal helix (3/4), the bend preceding the C-terminal helix (47/48), the C-terminal helix (49/50), and the C-terminus (56−57). The strongest NH_i-NH_{i+1} NOEs are between residues that are in a loop, turn, or bend in native BPTI. However, a clear set of sequential amide−amide NOEs are observed for all contiguous pairs of residues in the sequence 29−35, which in native BPTI is the antiparallel β-strand following the turn. Figure 3A shows the crystal structure of BPTI, with the residues involved in NH_i-NH_{i+1} NOEs in $[R]_{Abu}$ indicated in dark blue.

The NOESY spectrum of $[R]_{Abu}$ in Figure 4 shows NOEs between aromatic (ω_2) and aliphatic (ω_1) ^1H. In the lower two-thirds of the spectrum ($\omega_1 > 2.0$), many peaks arise from intraresidue NOEs of aromatic residues, and these are labeled with one number and two atom designators. In addition, 10 cross peaks in this region arise from interresidue NOEs, and these are labeled with two numbers and two atom designators. The peaks in the top third of Figure 4 are all side chain−side chain NOEs, and most are between aromatic and aliphatic groups. All aromatic−aliphatic interresidue NOEs are listed in the first column of Table 1, and a (+) sign in column 2 or 3 indicates whether they are side chain−backbone or side chain−side chain.

The NOEs observed for $[R]_{Abu}$ are also observed for reduced, naturally occurring BPTI, $[R]_{SH}$, and for reduced BPTI mutant F45A, $[R,F45A]_{SH}$. In the mutant, Phe 45 is replaced by alanine. There are no significant differences between two-dimensional spectra of $[R]_{Abu}$, $[R]_{SH}$, and $[R,F45A]_{SH}$ at pH 4.5, 1 °C, except for the extra Abu peaks in $[R]_{Abu}$ ($\omega_1 = 0.95$ ppm) and the missing Phe 45 peaks in

Table 1: Interresidue Aromatic−Aliphatic NOEs in $[R]_{Abu}$

$[R]_{Abu}$ NOEs[a]				side chain−backbone	side chain−side chain	native BPTI	BPTI peptides	$[14-38]_{Abu}$
Y10	HD	G12	HA	+		−	+[b]	−
	HE	G12	HA	+		−	+[b]	−
	HE	P9	HD		+	−	+	−
Y21	HD	I19	HG(CH_3)		+	−	+	−
	HE	I19	HD		+	−	+	−
	HE	I19	HB		+	−	+	−
F22	HD	R20	HA	+		−	−	−
	HD	R20	HB		+	−	−	−
	HD	R20	HG		+	−	−	−
	HZ	R20	HA	+		−	−	−
	HZ	R20	HB		+	−	−	−
	HZ	R20	HG		+	−	−	−
Y23	HD	A25	HA	+		+	+	+
	HD	A25	HB		+	+	+	+
	HE	A25	HA	+		+	+	+
	HE	A25	HB		+	+	+	+
F33	HD	V34	HG		+	−	−	−
	HE	Y35	HA	+		−	−	−
	HE	V34	HG		+	−	−	−
Y35	HE	G37	HA	+		+	+[b]	+
	HD	G36	HA	+		−	−	−
	HD	V34	HG		+	−	−	−
F45	HD	N43	HB		+	−	−	−
	HD	R42	HB		+	−	−	−

[a] Interresidue NOEs observed for $[R]_{Abu}$ (Figure 4). In the last three columns, + indicates that the same NOEs are observed in native BPTI by Berndt et al. (1992), in BPTI peptides by Kemmink and Creighton (1993), and/or in $[14-38]_{Abu}$ by Barbar et al. (1995). [b] In BPTI peptides the corresponding NOE is to NH instead of HA.

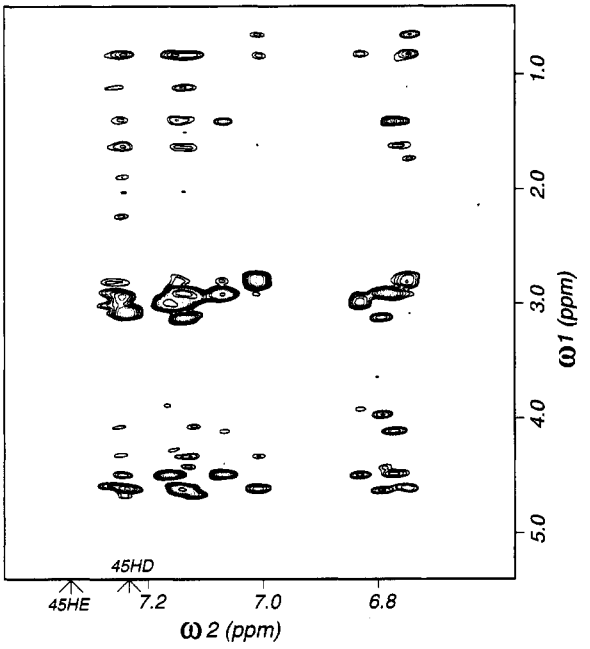

FIGURE 5: NOESY spectrum of $[R,F45A]_{SH}$ in ^2H_2O at pH 4.5, 1 °C, taken at 500 MHz, with a mixing time of 400 ms. Cross peaks show NOEs between atoms on aromatic residues (ω_2) and aliphatic atoms (ω_1). Arrows on the ω_2 axis indicate chemical shifts of Phe 45 peaks observed in $[R]_{Abu}$ (Figure 4) and in $[R]_{SH}$ but missing in $[R,F45A]_{SH}$.

$[R,F45A]_{SH}$. The NOESY spectrum of $[R,F45A]_{SH}$ in the aromatic−aliphatic region is shown in Figure 5. Comparison to equivalent NOESY spectra of $[R]_{Abu}$ (Figure 4) shows that the only difference is the absence of Phe 45 peaks in Figure 5 (in the region ω_2 = 7.2−7.4 ppm). When the same NOESY spectra of reduced BPTI in ^2H_2O are acquired at pH 3.5, many fewer cross peaks are observed and interresidue NOEs are observed only for 19/21 and 23/25 (data not shown). Although many NOEs are lost when the pH is lowered from 4.5 to 3.5, NOESY spectra taken at pH 6 are indistinguishable from those at pH 4.5. The NOESY spectra in Figures 4 and 5 were taken with a mixing time of 400 ms; similar spectra were obtained with a mixing time of 200 ms.

As another probe of nonrandom structure in unfolded BPTI, naturally occurring BPTI was reduced and reacted with [^{13}C]iodomethane to give the ^{13}C-labeled, S-methylated cysteine derivative, $[R]_{Smc}$. In one-dimensional HMQC spectra of $[R]_{Smc}$, ^{1}H bound to [^{13}C]methyl groups are selectively observed. The clear chemical shift dispersion of ^{13}C-bound, ^{1}H resonances in HMQC spectra at pH 5 and 1 °C (Figure 6A) indicates that the methyl groups on different cysteines have different average local environments. Since there are six *S*-methyl groups, the envelope centered around 2.1 ppm in Figure 6A contains 18 proton resonances. When the experiment is repeated with 14 and 38 *S*-methylcysteines labeled with ^{13}C but the other four *S*-methyl groups labeled with ^{12}C, only two peaks are observed (data not shown). One has the chemical shift of the lowest field peak in the ^{1}H envelope in Figure 6A, and the other has the chemical shift of the tallest peak in the envelope, allowing a partial assignment of the peaks. The chemical shift dispersion in [^{13}C]methyl protons verifies the presence of significant nonrandom structure in reduced BPTI and provides a simple way to monitor loss of nonrandom structure with increasing denaturant concentration. As guanidine hydrochloride (Gdn·HCl) concentration is increased (Figure 6B−D), the ^{13}C-bound ^{1}H resonances merge, indicating that with added denaturant the protein becomes increasingly more disordered and magnetic environments of the methyl groups become more uniform. At concentrations of Gdn·HCl >1.5 M the chemical shifts of the methyl protons are equivalent, implying that the protein is fully unfolded.

Differences in chemical shift for the same CαH in $[R]_{Abu}$ and in hexapeptides in 1 M urea at 25 °C (Wishart & Sykes, 1994) were determined (available as supporting information). Five CαH chemical shifts deviate by >0.2 ppm from random

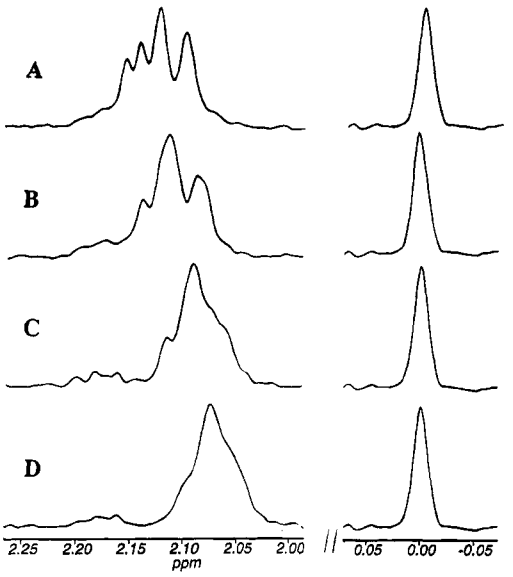

FIGURE 6: One-dimensional HMQC spectra of $[R]_{Smc}$ at pH 5.0, 1 °C. $[R]_{Smc}$ is an analog of reduced BPTI with cysteines replaced by *S*-[^{13}C]methylcysteine. The envelope of peaks centered around 2.1 ppm contains resonances of methyl protons attached to ^{13}C. Spectra A–D are for samples in 0, 0.4, 1.0, and 1.5 M guanidine hydrochloride, respectively. The peak at 0.0 ppm is the external TSP reference.

coil values, in residues 3, 7, 8, 25, and 50. Residues 3 and 7 are in the 3_{10} helix, and there is also an NH_i–NH_{i+1} NOE between residues 3 and 4, consistent with a turn-like conformation. Residue 25 may be influenced by the proximity of the Tyr 23 ring with which it makes an NOE (Table 1); also there is an NH_i–NH_{i+1} NOE between residues 25 and 26. The perturbed chemical shift of 50 CαH may be associated with a helix capping interaction with Ser 47. The presence of a hydrogen bond "cap" between the side chain oxygen of residue 47 and the NH of residue 50 is most likely responsible for the NOE between 50 HB and 47 NH in $[R]_{Abu}$ and in BPTI peptides. The deviation from random-coil chemical shift of Pro 8 is probably due to its interaction with Tyr 10 when the 8–9 peptide bond is in the *cis* form. In BPTI peptides, Kemmink and Creighton (1993) report a significant fraction of *cis* peptide bond preceding Pro 8 or Pro 9; they also report an alternative interaction of Tyr 10 with Gly 12 reflected in an anomalous chemical shift of Gly 12 NH (6.70 ppm). We observe the same evidence for these interactions in $[R]_{Abu}$, including a similar chemical shift for Gly 12 NH (6.76 ppm).

DISCUSSION

Turn-like Conformations. NMR analysis indicates that reduced BPTI has significant nonrandom structure at pH 4.5 and 1 °C. Most of the results presented here are for an analog of reduced BPTI, $[R]_{Abu}$, but very similar spectra are obtained for $[R]_{SH}$, naturally occurring BPTI with disulfides reduced to thiols, and for $[R,F45A]_{SH}$, mutant F45A with disulfides reduced to thiols. NMR spectra of unfolded proteins indicate extensive conformational averaging of flexible, unpacked species. NOEs are particularly sensitive NMR reporters of nonrandom structure, more so than, for example, chemical shifts or coupling constants (Zhang & Forman-Kay, 1995). Complementary information about nonrandom structure in reduced BPTI was obtained from two types of NOEs, sequential amide–amide NOEs and aromatic–aliphatic NOEs. NH_i–NH_{i+1} NOEs imply that a significant population of the rapidly interconverting, unfolded molecules sample turn-like conformations (Zhang & Forman-Kay, 1995; Dyson &Wright, 1991). In $[R]_{Abu}$, 18 sequential amide–amide NOEs are observed; these are assigned in Figure 2 and highlighted in dark blue in Figure 3A, a ribbon drawing of the crystal structure of native BPTI. Of the 18 NOEs, 4 are observed in native BPTI, in a helix (3/4, 49/50), turn (25/26), or loop (39/40). An additional 8 NOEs are between residues located in a loop, turn, or bend in native BPTI, and their propensity to populate turn-like conformations in the unfolded state seems congruent. The remaining 6 NH_i–NH_{i+1} NOEs in $[R]_{Abu}$, however, are particularly interesting in that they occur in contiguous residues 29–35 which in native BPTI fold as one strand of the central antiparallel β-sheet and which are predicted by the rules of Chou and Fasman (1978) to be sheet. In the other antiparallel strand (residues 18–24), amide–amide NOEs are clearly absent; this, taken together with the strong $C\alpha H_i$–NH_{i+1} NOEs for this stretch, suggests that residues 18–24 are highly extended. In summary, in $[R]_{Abu}$, the residues which ultimately fold into two contiguous strands of antiparallel sheet behave differently; the first is extended in reduced BPTI and in native BPTI, while the second has extensive turn-like conformation in reduced BPTI but is extended in native BPTI. As discussed in the next section, these results are intriguing because extensive non-native aromatic–aliphatic NOEs are observed for both strands in reduced BPTI.

A difference between secondary structure preference in unfolded *versus* folded states, such as we observed for residues 29–35, is reported for two other β-sheet proteins and may be a general feature of folding of β-sheet proteins. For FK506 binding protein unfolded in urea or Gdn·HCl, Logan et al. (1994) found $C\alpha H_i$–NH_{i+1} NOEs for all residues and NH_i–NH_{i+1} NOEs for many residues, plus a number of $C\alpha H_i$–NH_{i+2} and $C\alpha H_i$–NH_{i+3} NOEs. The latter are strongly indicative of turns and helices and mostly occur between residues which in native FK506 binding protein are located at or near the end of a β-strand or in a turn, loop, or helix. However, some are observed for residues that fold into a stable strand of β-sheet, and interestingly, statistical rules (Chou & Fasman, 1978) predict that this segment should be helix. For an unfolded SH3 domain, studied at equilibrium with the native conformation, Zhang and Forman-Kay (1995) observe $C\alpha H_i$–NH_{i+1} NOEs for most, but not all residues. They also observe NH_i–NH_{i+1} NOEs, diagnostic of turn-like conformations, for several noncontiguous pairs of residues and for a stretch of residues which, in native SH3 domains, are in stable β-sheet strands. This stretch of residues loses the NH_i–NH_{i+1} NOEs when the SH3 domains are denatured by Gdn·HCl.

Aromatic–Aliphatic NOEs. In addition to turn-like conformations all along the backbone, extensive hydrophobic NOEs are observed in reduced BPTI. These are side chain–side chain and side chain–backbone NOE peaks in the aromatic–aliphatic region of the spectrum (Figure 4 and Table 1). They occur between $i,i\pm1-3$ residues and primarily involve aromatic residues 21, 22, 23, 33, and 35, all of which are in the two strands of antiparallel β-sheet in native BPTI. Two aromatic–aliphatic NOEs in $[R]_{Abu}$ spectra are observed in native BPTI, as noted in the fourth column of Table 1. These suggest an interaction of the Tyr

23 ring with Ala 25 in the turn between the strands and an aromatic−NH interaction of the Tyr 35 ring with Gly 37 NH (Tüchsen & Woodward, 1987). However, all other aromatic−aliphatic NOEs in $[R]_{Abu}$ are non-native, a point illustrated in Figure 3B, which shows the side chains involved in aromatic−aliphatic NOEs. Residues in the same color contain atoms that make NOEs with each other in reduced BPTI; most are clearly too distant in native BPTI to give rise to NOEs. Apparently, in residues 18−24 non-native hydrophobic interactions stabilize extended conformations, while in residues 29−35 non-native hydrophobic interactions stabilize turn-like conformations. In addition, turn-like conformations are stabilized by native-like hydrophobic interactions in residues 25−28, which fold into the turn between the strands.

Comparison to Nonrandom Structure in BPTI Peptides. In comprehensive peptide studies, Creighton and associates have characterized local, NMR-detected interactions at pH 4.6 and −2 °C in peptides of 9−16 residues with sequences spanning the entire length of BPTI. The interactions they report for BPTI peptides are among those observed in our study, as indicated by a plus sign in the fifth column of Table 1. From their peptide work, Kemmink and Creighton (1993) conclude that residues 18−24 have the characteristics of an extended peptide, that there are native-like interactions between the Tyr 23 ring and turn residue Ala 25 and between the Tyr 35 ring and Gly 37 NH, and that an NOE between atoms of Ser 47 and Asp 50 is consistent with native-like helix capping. BPTI peptides further indicate non-native interactions between the Tyr 10 ring and Gly 12 NH and between side chains of Ile 19 and Tyr 21 and a significant fraction of *cis* peptide bond preceding Pro 8 or Pro 9. Lumb and Kim (1994) also reported the Ile 19−Tyr 21 interaction in BPTI peptides. We observe all these interactions in reduced BPTI but observe many others as well (Table 1 and Figure 2).

Comparison to Other Studies of Reduced BPTI. For some time, the lack of CD-detectable, stable secondary structure, along with other features of reduced BPTI, seemed consistent with the possibility that reduced BPTI lacks significant nonrandom structure (Creighton, 1992; Darby & Creighton, 1993). Interpretations of disulfide-linked folding of BPTI often carry an implicit assumption that reduced BPTI has no nonlocal structure and little, if any, local structure. However, Haas and associates concluded in 1988 from their fluorescence energy transfer studies that reduced BPTI samples collapsed conformations, and their more recent work confirms this view (Amir & Haas, 1988; Ittah & Haas, 1995). Daggett and Levitt (1992) reported calculations which indicate that reduced BPTI has molten globule characteristics. Recently, we reported evidence, from electrophoresis and ANS dye binding, that reduced BPTI is more collapsed than fully denatured BPTI and forms hydrophobic clusters (Ferrer et al., 1995), and we proposed that reduced BPTI is appropriately referred to as a "molten coil". The extensive hydrophobic NOEs in $[R]_{Abu}$, $[R]_{SH}$, and $[R,F45A]_{Abu}$ corroborate our ANS-binding results and identify specific hydrophobic contacts favored in reduced BPTI. Also, the numerous NOEs reported here are consistent with our electrophoresis results, indicating that while reduced BPTI is less compact than native BPTI, it is also less extended than fully denatured BPTI. Although in rapid equilibrium between flexible and extended conformations, reduced BPTI is sufficiently compact, on the average, to populate turn-like conformations at 18 pairs of residues along the chain and to give multiple aromatic−aliphatic NOEs between six $i,i\pm1-3$ pairs of residues.

The presence of NMR-detected nonrandom structure in reduced BPTI is sensitive to pH below pH 4.5, and this accounts for why we observe many more NOEs than Kemmink and Creighton (1993), who report that NMR spectra of intact reduced BPTI at pH 3.5 and −2 °C, though not assigned, appeared to be essentially the sum of the spectra of large peptide fragments. Our studies are done at pH 4.5 because there is a cooperative acid denaturation for partially folded $[14-38]_{Abu}$, a model for BPTI early folding intermediates (Ferrer et al., 1995; Barbar et al., 1995). $[14-38]_{Abu}$ retains the 14−38 disulfide bond while other cysteines are replaced by Abu. Between pH 4 and 2.5, $[14-38]_{Abu}$ unfolds, presumably from titration of the four carboxyl groups. The 10 positively charged residues (Arg + Lys) on this highly basic protein of only 58 amino acids are likely to undergo significant repulsive interactions in the absence of anionic groups. We reasoned that reduced BPTI may also be further denatured by titration of the carboxyls, and this is consistent with the differences we observe between spectra of reduced BPTI at pH 3.5 and 4.5. NOESY spectra acquired under conditions exactly equivalent to those of Figure 4, but at pH 3.5, have clear interresidue NOE cross peaks only for 19/21 and 23/25, suggesting that these are the most stable interactions in reduced BPTI. Although their peptide spectra were taken at pH 4.6, Kemmink and Creighton (1993) state that reduced BPTI was measured at pH 3.5 for solubility reasons. Their final concentration of reduced BPTI was 3 mM; ours is 0.6 mM or less (see Materials and Methods).

Denaturation of Nonrandom Structure in Reduced BPTI. Consistent with the indications that nonrandom structure in reduced BPTI undergoes acid denaturation, we also find that nonrandom structure is lost by addition of a chemical denaturant. The chemical shift dispersion of resonances of ^1H bound to ^{13}C-methylated cysteines decreases with increasing guanidine hydrochloride (Figure 6). At Gdn·HCl concentrations > 1.5 M the [^{13}C]methyl ^1H resonances merge into a single peak, implying loss of nonrandom structure and denaturation to a more fully extended unfolded state. Fluorescence-detected binding of ANS also is eliminated by addition of Gdn·HCl at < 1 M denaturant concentration (Ferrer et al., 1995). The loss of ANS fluorescence at Gdn·HCl concentrations less than those required to fully unfold reduced BPTI may occur because Gdn·HCl competes with ANS binding or otherwise interferes with the ANS fluorescence signal. Another possible explanation is that Gdn·HCl disrupts the hydrophobic clusters thought to be responsible for ANS binding at concentrations less than it takes to eliminate all nonrandom structure.

Comparison to $[14-38]_{Abu}$, a Model for BPTI Early Folding Intermediates. $[14-38]_{Abu}$ is a cooperatively formed, highly ordered β-sheet molten globule (Ferrer et al., 1995). It is an ensemble of partially folded conformations interconverting slowly enough to be detected separately by NMR (Barbar et al., 1995). Although the 14−38 single disulfide species was long considered inconsequential in BPTI disulfide-coupled folding, it has recently been detected and suggested to be a significant early kinetic intermediate (Dadlez & Kim, 1995). We take $[14-38]_{Abu}$ as a model for early folding intermediates of BPTI, regardless of which

native-like disulfide bond forms first. In this context, it is of interest to examine whether the interactions observed in reduced BPTI persist in $[14-38]_{Abu}$. As columns 4 and 6 in Table 1 show, the only interactions in reduced BPTI also observed in $[14-38]_{Abu}$ are found in native BPTI as well. This implies that when the 14−38 disulfide is formed, the 14−38 single disulfide species loses a number of interactions present in reduced BPTI, retains and/or strengthens three interactions (Tyr 23 with turn residue Ala 25, the aromatic−NH interactions of Tyr 35 with Gly 37 NH, and helix capping of Ser 47 with Asp 50), and gains other interactions, including nonlocal hydrophobic interactions and β-sheet interactions, neither of which are in the vicinity of the 14−38 disulfide cross-link [described in Barbar et al. (1995)]. If the 14−38 single disulfide species loses a number of interactions present in reduced BPTI, then the unfolded state of $[14-38]_{Abu}$ should have many fewer NOE-detected interactions than reduced BPTI. This is observed for $[14-38]_{Abu}$ unfolded by a destabilizing mutation (E. Barbar, G. Barany, and C. Woodward, unpublished results).

NOEs indicating that numerous hydrophobic contacts are sampled in reduced BPTI are consistent with the models of Dill and associates in which formation of secondary structure is preceded by hydrophobic collapse, which may thereby reduce the conformational space to be searched in a thermodynamic model of folding (Dill, 1990). Our results are also consistent with the view that, rather than being nonspecific, such hydrophobic collapse has sequence specificity and is encoded in the linear message of protein sequence. In biology, important functions are often guaranteed by highly redundant, though specific, processes; removal of one or several specific hydrophobic interactions is not expected to prevent collapse and folding but may well slow it. Observation of numerous non-native interactions in reduced BPTI also underscores the point emphasized by Dill and Shortle (1991) that understanding the structural basis of protein stability requires understanding the structure and thermodynamics of the unfolded as well as the folded state.

CONCLUSIONS

Reduced BPTI at pH 4.5 consists of rapidly interconverting, flexible, unfolded conformations without stable secondary structure; on the average, it is collapsed and has exposed hydrophobic clusters. NMR results are consistent with previous characterization of reduced BPTI as a molten coil (Ferrer et al., 1995). Extensive nonrandom structure is reflected in NH_i-NH_{i+1} NOEs, which indicate turn-like conformations, and in multiple aromatic−aliphatic NOEs, which imply hydrophobic interactions. The NH_i-NH_{i+1} NOEs occur between 18 pairs of adjacent residues spread through the molecule; 12 involve residues which in native BPTI are located in a turn, loop, or bend, and of these, 4 give native amide−amide NOEs. The remaining 6 are contiguous in residues 29−35, which in native BPTI is a strand of stable antiparallel β-sheet. Multiple aromatic−aliphatic NOEs are observed between ring hydrogens of 5 aromatic residues (21, 22, 23, 33, 45) and apolar atoms of the side chain and backbone of the $i\pm1-3$ residue. In some cases the aliphatic ^1H is on a polar side chain (Arg 20, Arg 42, Asn 43). Two additional aromatic groups (10, 35) make alternative contacts, one with a Gly in the $i+2$ position and the other with apolar atoms in the $i\pm1$ residue.

The hydrophobic contacts in unfolded, reduced BPTI implied by the aliphatic−aromatic NOEs corroborate our earlier reports of ANS binding (Ferrer et al., 1995). Apolar interactions in unfolded BPTI also support the proposal that hydrophobic collapse precedes formation of secondary structure (Dill, 1990). The apparent hydrophobic interactions in this case are not within an amorphous "oil drop" but rather are between specific pairs of side chains that are 1−3 residues apart. Further, they are mostly non-native, and they involve side chains in segments that fold first. The NOEs involving aromatic rings 21 and 23 are observed in peptides, but those involving rings 22, 33, and 45 are not; this suggests that the latter are to some extent dependent on nonlocal interactions.

There are two other reports of unfolded proteins having turn-like conformations in segments that, in the native state, fold into a stable, extended β-strand (Logan et al., 1994; Zhang & Forman-Kay, 1995), and this may be a general feature of folding in β-sheet proteins. We suggest that this property is based on non-native hydrophobic interactions which in unfolded BPTI, and perhaps in other unfolded proteins, function to stabilize unpacked conformations until conditions are favorable for folding and thereby to minimize misfolding. That is, the protein sequence "message" for a fold may be evolutionarily optimized to include non-native interactions in unfolded states.

Comparison of the structure of reduced BPTI to the structure of $[14-38]_{Abu}$, an equilibrium model for early folding intermediates of BPTI (Barbar et al., 1995), strongly suggests that folding is initiated in the antiparallel strand−turn−strand of residues 18−35, which also contains the slow exchange core (Kim et al., 1993; Woodward, 1993). A notable result of these studies is that residues in the stretch 18−35 also are the ones involved in non-native hydrophobic NOEs. Initial steps in folding, following the formation of *any* native disulfide bond, apparently involve breaking non-native interactions, consolidating several native-like interactions (e.g., Tyr 23−Ala 25, Tyr 35 ring−Gly 37 NH), and establishing native-like interactions in the slow exchange core. In this view, any single native disulfide bond excludes the most stable unfolded states and thereby favors formation of more folded forms (Ferrer et al., 1995; Barbar et al., 1995). A feasible scheme consistent with our data for $[R]_{Abu}$ and $[14-38]_{Abu}$ is as follows. Unfolded conformations are stabilized by non-native hydrophobic interactions in segments 18−24 and 29−35 in reduced BPTI. Formation of a native disulfide bond destabilizes these interactions but not the interactions in turn residues 25−28 which nucleate folding. In initial folding steps, the strand−turn−strand (18−24−25−28−29−35) folds in a cooperative process involving formation of β-sheet hydrogen bonds and nonlocal hydrophobic contacts of Tyr 21 and Tyr 35. The small 45−46 β-bridge is also implicated in early folding steps. Though not as stable as the antiparallel β-sheet strands in the folding intermediate, the 45−46 bridge is the next most stable region, and it is significant that, in reduced BPTI, Phe 45 also shows non-native hydrophobic NOEs.

SUPPORTING INFORMATION AVAILABLE

A table of ^1H NMR assignments of reduced BPTI and a table of the differences in chemical shift for the same CαH in $[R]_{Abu}$ and hexapeptides in urea (Wishart & Sykes, 1994)

(4 pages). Ordering information is given on any current masthead page.

REFERENCES

Amir, D., & Haas, E. (1988) *Biochemistry 27*, 8889−8893.
Arcus, V., Vuilleumier, S., Freund, S., Bycroft, M., & Fersht, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9412−9416.
Barany, G., Gross, C., Ferrer, M., Barbar, E., Pan, H., & Woodward, C. (1995) in *Techniques in Protein Chemistry* (Marshak, D., Ed.) Vol. VII (in press).
Barbar, E., Barany, G., & Woodward, C. (1995) *Biochemistry 34*, 11423−11434.
Bax, A., Griffey, R., & Hawkins, B. (1983) *J. Magn. Reson. 55*, 301−315.
Chou, P. Y., & Fasman, G. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45−147.
Creighton, T. (1992) in *Protein Folding* (Creighton, T., Ed.) pp 301−351, W. H. Freeman and Co., New York.
Dadlez, M., & Kim, P. (1995) *Nature Struct. Biol. 2*, 674−679.
Daggett, V., & Levitt, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5142−5146.
Darby, N., & Creighton, T. (1993) *J. Mol. Biol. 232*, 873−896.
Dill, K. (1990) *Biochemistry 29*, 7133−7155.
Dill, K., & Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
Dobson, C. (1992) *Curr. Opin. Struct. Biol. 2*, 6−12.
Dyson, H. J., & Wright, P. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519−538.
Ferrer, M., Woodward, C., & Barany, G. (1992) *Int. J. Pept. Protein Res. 40*, 194−207.
Ferrer, M., Barany, G., & Woodward, C. (1995) *Nature Struct. Biol. 2*, 211−217.
Goldenberg, D. (1988) *Biochemistry 27*, 2481−2489.
Griesinger, C., Otting, K., Wüthrich, K., & Ernst, R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
Housset, D., Kim, K.-S., Fuchs, J., Woodward, C., & Wlodawer, A. (1991) *J. Mol. Biol. 220*, 757−770.
Ittah, V., & Haas, E. (1995) *Biochemistry 34*, 4493−4506.
Jeener, J., Meier, B., Bachmann, P., & Ernst, R. (1979) *J. Chem. Phys. 71*, 4546−4553.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Kemmink, J., & Creighton, T. (1993) *J. Mol. Biol. 234*, 861−878.
Kim, K.-S., Fuchs, J., & Woodward, C. (1993) *Biochemistry 32*, 9600−9608.
Logan, T., Theriault, Y., & Fesik, S. (1994) *J. Mol. Biol. 236*, 637−648.
Lumb, K., & Kim, P. (1994) *J. Mol. Biol. 236*, 412−420.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Neri, D., Billeter, M., Wider, G., & Wüthrich, K. (1992) *Science 257*, 1559−1562.
Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) *J. Magn. Reson. 52*, 335−338.
Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
Tüchsen, E., & Woodward, C. (1987) *Biochemistry 26*, 1918−1925.
Wishart, D., & Sykes, B. (1994) *Methods Enzymol. 239*, 363−392.
Woodward, C. (1993) *Trends Biochem. Sci. 18*, 359−360.
Zhang, O., & Forman-Kay, J. (1995) *Biochemistry 34*, 6784−6794.

BI9515230