

Formation of a Hydrophobic Cluster in Denatured Bovine Pancreatic Trypsin Inhibitor

Kevin J. Lumb and Peter S. Kim†

Howard Hughes Medical Institute
Whitehead Institute for Biomedical Research
Department of Biology
Massachusetts Institute of Technology
Nine Cambridge Center
Cambridge, MA 02142, U.S.A.

Bovine pancreatic trypsin inhibitor (BPTI) unfolds upon reduction of its three disulfide bonds. A recombinant model of the reduced state of BPTI, called [R]_{Ala}, in which all six Cys residues are replaced with Ala, has been expressed in *Escherichia coli*. ¹H nuclear magnetic resonance spectroscopy shows that [R]_{Ala} does not contain stable secondary structure. Some chemical shift dispersion exists, however, suggesting the existence of non-random interactions in [R]_{Ala}. In particular, the side-chain protons of Ile19 resonate upfield of those of Ile18. This observation was investigated using an eight residue peptide model, P₁₇₋₂₄, corresponding to residues 17 to 24 of BPTI. The non-random chemical shift dispersion of the Ile residues observed in [R]_{Ala} also occurs in P₁₇₋₂₄, indicating that P₁₇₋₂₄ contains interactions that are similar to those found in the corresponding region of [R]_{Ala}. The only interresidue nuclear Overhauser effects observed in P₁₇₋₂₄ are between the ring protons of Tyr21 and the γ CH₃ group of Ile19, indicating that these protons are in close proximity. Substitution of Tyr21 by Ala in P₁₇₋₂₄ results in the loss of the chemical shift dispersion of the Ile resonances, suggesting that the upfield shifts of the Ile19 resonances are due to ring current shifts arising from the proximity of Tyr21. Collectively, these results suggest that the side-chain of Ile19 is positioned at least some of the time above the plane of the aromatic ring of Tyr21. We conclude that these two residues participate in a hydrophobic cluster in P₁₇₋₂₄ and in the denatured state of BPTI.

Keywords: BPTI; NMR; alanine scanning; protein folding; unfolded state

The denatured states of proteins have traditionally been assumed to be ensembles of structures in which conformational preferences are determined by steric restraints, the "random coil" state (Tanford, 1968). Recent studies of protein folding and stability, however, have prompted renewed interest in the denatured state, and evidence is emerging that this state is not necessarily as "random" as once thought (e.g. Dill & Shortle, 1991; Dobson, 1992; Shortle, 1993). The existence of structure in the denatured state has implications for understanding the early events in protein folding, and for rationalizing protein stability, which is determined by the difference in free energies of the unfolded and folded states.

BPTI† is a small (58 residue), highly characterized protein that has a melting temperature at neutral pH in excess of 100°C. The stability of BPTI is linked, however, to the formation of the

three native disulfide bonds; reduction of all of these disulfide bonds unfolds the protein. Fluorescence energy transfer experiments suggest that reduced BPTI can exist in a compact conformation (Amir & Haas, 1988; Amir *et al.*, 1992; Gottfried & Haas, 1992). Circular dichroism and optical rotary dispersion spectroscopy suggest that reduced BPTI is

† Abbreviations used: BPTI, bovine pancreatic trypsin inhibitor; 1D, one-dimensional; 2D, two-dimensional; DMSO, dimethylsulfoxide; DQF COSY, 2D double quantum filtered correlation spectroscopy; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; p.p.b., parts per billion; p.p.m., parts per million; P₁₇₋₂₄, synthetic peptide RIIRYFYFN with an acetylated N terminus and amidated C terminus, corresponding to residues 17 to 24 of BPTI; P₁₄₋₂₈, synthetic peptide CKARIIRYFYNAKAG corresponding to residues 14 to 28 of BPTI in which the Cys residue is carboxymethylated; [R]_{Ala}, recombinant BPTI mutant in which the six Cys residues are replaced by Ala; ROESY, 2D rotating frame NOE spectroscopy.

† Author to whom all correspondence should be addressed.

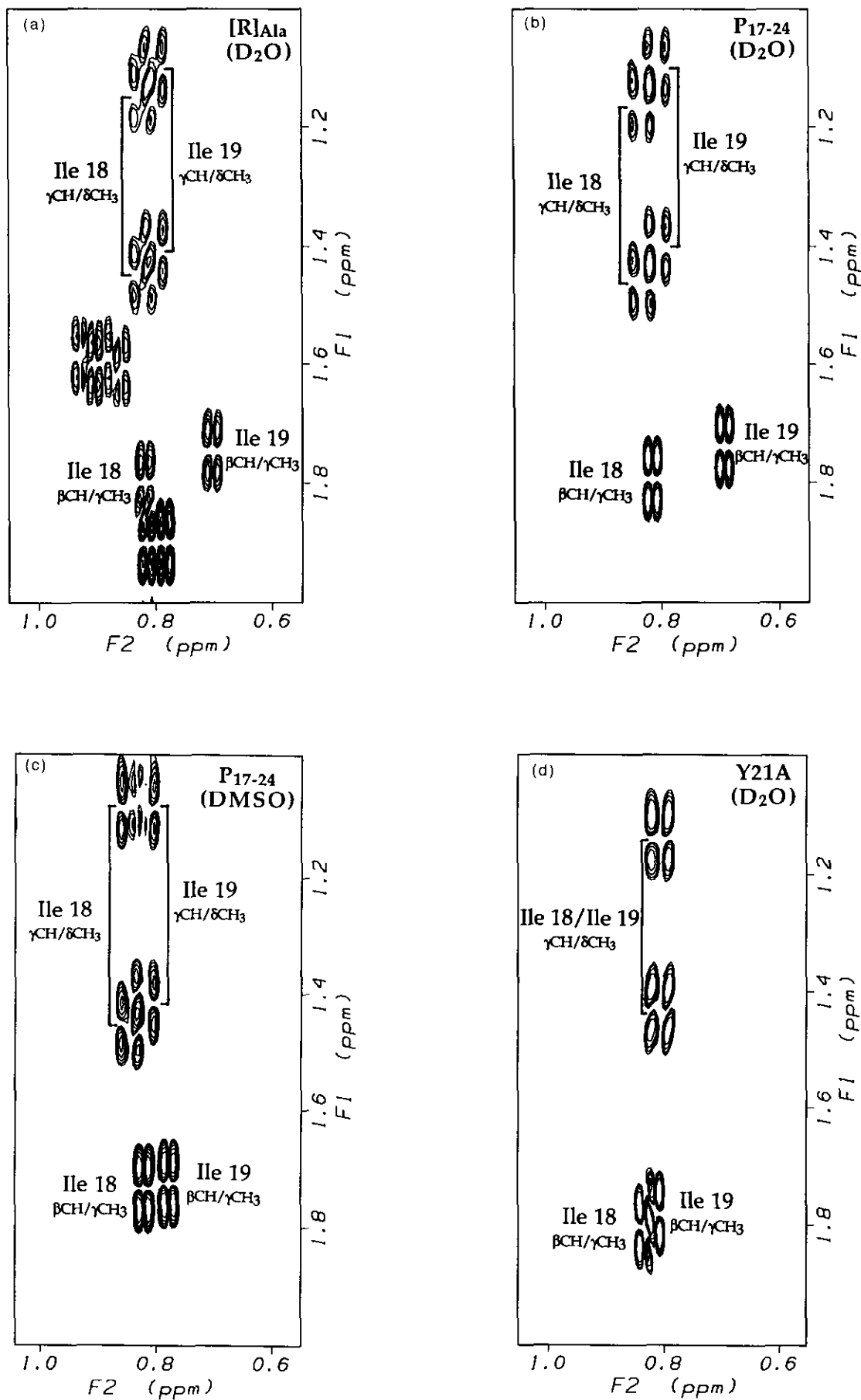


Fig. 1.

devoid of any elements of secondary structure (Kosen *et al.*, 1981; Gussakovskiy & Haas, 1992), and mutational studies suggest that the average properties of this state are relatively insensitive to amino acid replacements (Goldenberg & Zhang, 1993). Chemical shift dispersion in the ^1H NMR spectrum suggests that some residual structure exists in denatured BPTI (Roder, 1988). The anomalous chemical shift of the amide proton of Gly12 in unfolded BPTI has been investigated in detail, and has been attributed to an interaction between the aromatic ring of Tyr10 and the amide proton of Gly12 (Kemink *et al.*, 1993; Staley, 1993).

We describe here a recombinant model of the reduced unfolded state of BPTI, called $[\text{R}]_{\text{Ala}}$, in which the six Cys residues are replaced with Ala. ^1H NMR studies indicate that $[\text{R}]_{\text{Ala}}$ is predominantly unfolded. Nonetheless, significant chemical shift dispersion exists in $[\text{R}]_{\text{Ala}}$, similar to that found in reduced BPTI (Roder, 1988). In particular, the resonances of Ile18 and Ile19 are resolved. Using synthetic peptides, we demonstrate that this chemical shift dispersion results from a specific interaction between Ile19 and Tyr21.

NMR studies of $[\text{R}]_{\text{Ala}}$

The ^1H NMR spectrum of $[\text{R}]_{\text{Ala}}$ is similar, in general, to that expected of an unfolded protein.

Interresidue NOEs that are characteristic of stable secondary structure (Wüthrich, 1986) are not observed in NOESY experiments. The spectrum of $[\text{R}]_{\text{Ala}}$ freshly dissolved in $^2\text{H}_2\text{O}$ at 20°C and pH 3.0 (where the half-life for hydrogen exchange is approximately 10 minutes; Englander *et al.*, 1972) does not contain resonances corresponding to labile protons after approximately 40 minutes. These observations suggest that $[\text{R}]_{\text{Ala}}$ does not contain stable elements of secondary structure, in accord with previous studies by circular dichroism and optical rotary dispersion spectroscopy of the reduced state of BPTI (Kosen *et al.*, 1981; Gussakovskiy & Haas, 1992) and the low stability of the helix formed by a peptide corresponding to the α -helix of BPTI (Goodman & Kim, 1989).

The chemical shifts of many of the resonances of $[\text{R}]_{\text{Ala}}$ in the DQF COSY spectrum are close to those expected from studies of unstructured model peptides (Bundi & Wüthrich, 1979; Wüthrich, 1986). There are, however, several deviations from the chemical shifts expected of a "random coil" peptide that may reflect the existence of non-random interactions in $[\text{R}]_{\text{Ala}}$. In particular, the resonances of the two Ile residues (Fig. 1A) and of the aromatic residues (Fig. 2A) are resolved from one another. This was also found in reduced BPTI (Roder, 1988).

A few resonance assignments of reduced BPTI

Figure 1. Comparison of DQF COSY spectra at 37°C showing the Ile18 and Ile19 side-chain resonances. A, $[\text{R}]_{\text{Ala}}$, showing chemical shift dispersion of the Ile18 and Ile19 resonances. B, The resonances of the Ile residues of P_{17-24} have essentially identical chemical shifts to those of $[\text{R}]_{\text{Ala}}$. C, The chemical shift dispersion of the Ile side-chain resonances of P_{17-24} is much reduced in $\text{DMSO}-d_6$. The random coil chemical shifts are different in aqueous solution and in DMSO (Bundi *et al.*, 1975; Bundi & Wüthrich, 1979), complicating a direct comparison of the spectra in B and C. It is evident, however, that the chemical shift dispersion is much reduced in DMSO, and in particular the γCH_3 resonances of the Ile residues are much closer in chemical shift in DMSO than in water. D, The Y21A variant of P_{17-24} . Substitution of the Tyr21 side-chain with Ala results in essentially complete loss of chemical shift dispersion between the Ile resonances. $[\text{R}]_{\text{Ala}}$ was expressed in *Escherichia coli* strain BL21 (DE3) pLysS using the T7 system (Studier *et al.*, 1990). The plasmid encoding $[\text{R}]_{\text{Ala}}$ was prepared by oligonucleotide directed mutagenesis (Kunkel *et al.*, 1987) of the plasmid for $[\text{R}]_{5-55}$ (Staley & Kim, 1992) by mutating Cys5 and Cys55 to Ala. Cells were grown from overnight cultures in LB media to $A_{600} = 0.6$ and induced with isopropyl- β -D-thiogalactopyranoside (final concentration 0.5 mM). After 3 h the cells were harvested by centrifugation and lysed by freezing followed by sonication in 0.2 M NaOH, 1% SDS. An equal volume of detergent buffer (0.2 M NaCl, 20 mM Tris-HCl (pH 8.0), 2 mM EDTA, 1% deoxycholic acid, 1% Nonidet P-40) was added and the mixture sonicated. The resulting mixture was centrifuged and the inclusion body pellet was resuspended in 6 M guanidine hydrochloride (pH 8.0) and dialyzed against 5% acetic acid. The dialysate was centrifuged to remove insoluble impurities and $[\text{R}]_{\text{Ala}}$ was purified from the supernatant by reverse phase high performance liquid chromatography on a Vydac C_{18} column using a linear acetonitrile/water gradient containing 0.1% trifluoroacetic acid. Peptides were synthesized using solid phase t-boc methods as described previously (Goodman & Kim, 1989) and purified by high performance liquid chromatography as described above. The identity of the peptides were confirmed by laser desorption mass spectrometry on a Finnigan Lasermat and in all cases the expected and observed molecular weights agreed to within 2 Da. ^1H NMR spectroscopy was performed at 500.1 MHz on a Bruker AMX spectrometer. A spectral width of 5208.33 Hz, a recycle delay of 1 s and solvent presaturation was used. The residual solvent resonance was suppressed by convolution of the time-domain data (Marion *et al.*, 1989). The receiver phase was adjusted to minimize baseline distortion (Marion & Bax, 1988). DQF COSY (Piantini *et al.*, 1982; Rance *et al.*, 1983; Shaka & Freeman, 1983), NOESY (Jeener *et al.*, 1979; Kumar *et al.*, 1980; Macura *et al.*, 1981) and z-filtered ROESY (Bothner-By *et al.*, 1984; Rance, 1987) data sets consisted of 512 increments defined by 32 to 80 transients and 1024 complex points. Phase discrimination in F1 was achieved using time-proportional phase incrementation (Marion & Wüthrich, 1983). Data were resolution enhanced using a Gaussian function in t_2 and a 50° shifted squared sine bell in t_1 and zero filled prior to Fourier transformation to give a final digital resolution of 2.5 Hz/point in both dimensions. Samples were typically 1 mM in $[\text{R}]_{\text{Ala}}$ or 2 to 4 mM in peptide, except the R17A and R20A variants of P_{17-24} , which had solubility limits in water at pH 3 of approximately 0.3 mM. Peptide concentration was determined by tyrosine absorbance in 6 M guanidine hydrochloride assuming an extinction coefficient of $1500 \text{ M}^{-1} \text{ cm}^{-1}$ at 276 nm (Edelhoch, 1967). Samples were prepared in either 90% $\text{H}_2\text{O}/10\%$ $^2\text{H}_2\text{O}$ or $^2\text{H}_2\text{O}$ at pH 3, where the quality of the spectrum of $[\text{R}]_{\text{Ala}}$ was greatest, and were referenced to zero p.p.m. with internal trimethylsilylpropionic acid.

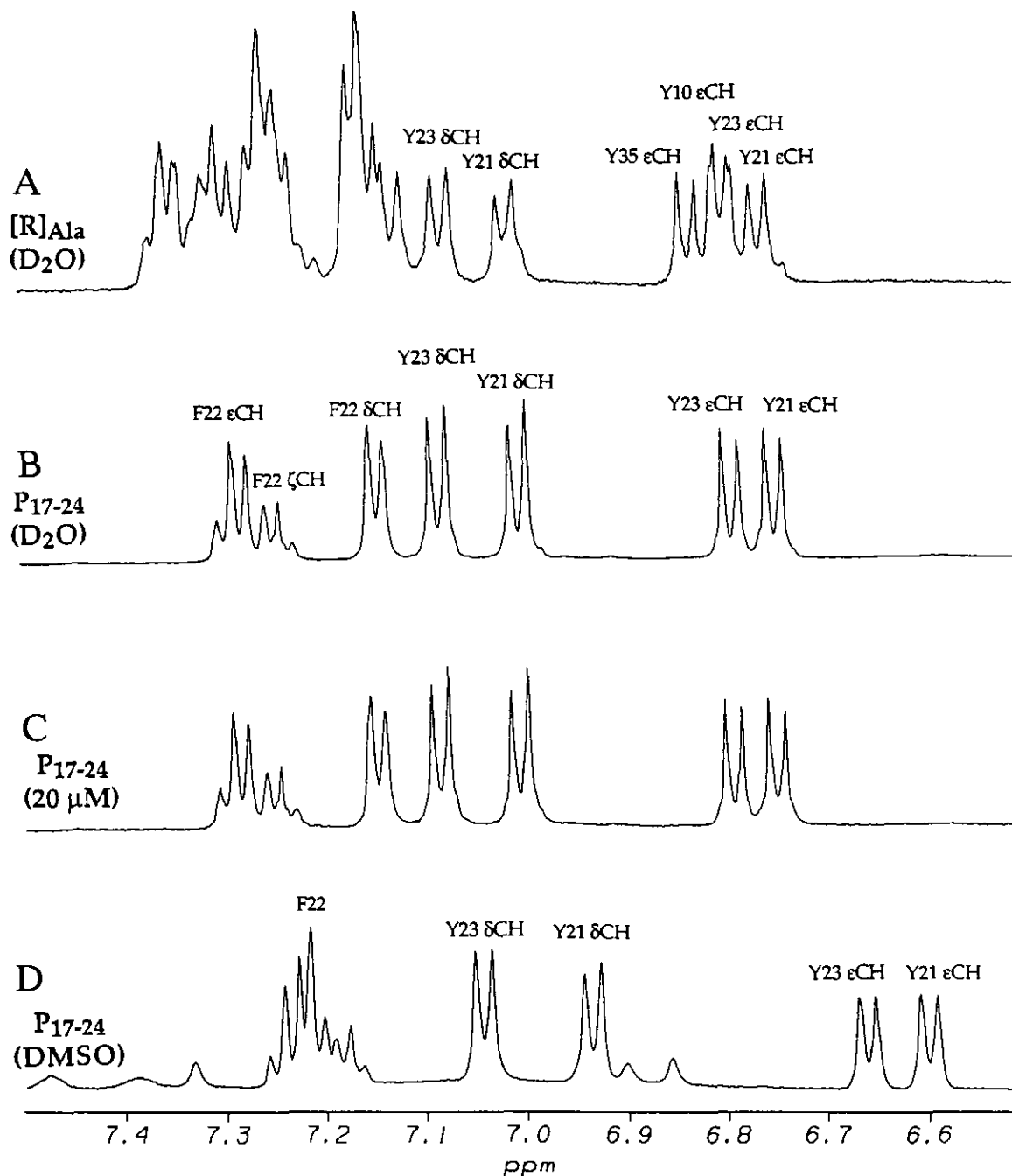


Figure 2. Comparison of the aromatic regions of 1D ^1H NMR spectra in $^2\text{H}_2\text{O}$ at 37°C . A, $[\text{R}]_{\text{Ala}}$, showing chemical shift dispersion of the aromatic resonances. B, P_{17-24} . The chemical shifts of the Tyr21, Phe22 and Tyr23 aromatic resonances are very similar in $[\text{R}]_{\text{Ala}}$ and P_{17-24} . C, A 100-fold dilution of P_{17-24} in B to $20\ \mu\text{M}$ does not affect the chemical shifts or linewidths of the peptide resonances. D, P_{17-24} in $\text{DMSO}-d_6$. The chemical shift dispersion of the aromatic resonances in DMSO is similar to that in water. Note that the random coil chemical shifts are different in aqueous solution and in DMSO (Bundi *et al.*, 1975; Bundi & Wüthrich, 1979). 1D NMR data sets consist of either 512 (Fig. 1A, B and D) or 20000 (Fig. 1C) transients defined by 4096 complex points and were zero-filled once but not apodized prior to Fourier transformation. Other experimental details are given in the legend to Fig. 1.

have been reported previously (Roder, 1988). The similarity between the spectra of $[\text{R}]_{\text{Ala}}$ and reduced BPTI allowed the transfer of assignments for a number of resonances from reduced BPTI to $[\text{R}]_{\text{Ala}}$, including the aromatic side-chain resonances of the four Tyr residues and of the Ile19 γCH_3 resonance. The assignments of the remaining ^1H resonances of

these residues in $[\text{R}]_{\text{Ala}}$ were then completed from DQF COSY and NOESY data. Since BPTI contains only two Ile residues it was possible to assign Ile18. One striking feature of the DQF COSY spectrum of $[\text{R}]_{\text{Ala}}$ is the substantial chemical shift dispersion of the two Ile residues (Fig. 1A). In a random conformation these two residues would be expected to

have indistinguishable chemical shifts. That this is not the case suggests that some residual structure exists in this region of $[R]_{Ala}$.

Characterization of a hydrophobic cluster involving Ile 19 and Tyr 21

To investigate the existence of non-random structure in the vicinity of the Ile residues, an eight residue peptide model, P_{17-24} , corresponding to residues 17 to 24 of BPTI (i.e. RIIRYFYN) was characterized. In addition to the two Ile residues, the peptide includes three aromatic residues Tyr21, Phe22 and Tyr23 (ring current effects from these residues are possible origins of the chemical shift dispersion of the Ile resonances). Arg17 and Asn24 were included to extend the termini beyond Ile18 and Tyr23. The N terminus was acetylated and the C terminus was amidated to avoid the introduction of additional charges.

The amide NH resonances of P_{17-24} were assigned by identifying $d_{\alpha N}(i, i+1)$ connectivities in NOESY spectra (Fig. 3A), and the assignments of the remaining resonances were then completed from DQF COSY spectra. The resonances of the Ile, Phe and Tyr residues of P_{17-24} have very similar chemical shifts to those observed in $[R]_{Ala}$ (compare Fig. 1A with 1B and Fig. 2A with 2B). This indicates that the different Ile chemical shifts in $[R]_{Ala}$ are largely a consequence of local interactions, rather than a result of some more global property such as compactness. We conclude that P_{17-24} contains structure that is essentially identical to that found in the corresponding region of $[R]_{Ala}$.

The chemical shifts and line widths of the resonances of the peptide are independent of concentration over the range of 20 μ M to 2 mM (compare Fig. 2B and 2C), indicating that the spectral dispersion is not the result of aggregation. At millimolar concentrations, however, there are a number of very weak resonances that are not present at micromolar concentrations (apparent in the 1D spectra in Fig. 3B), suggesting that there is a small degree of aggregation of the peptide at millimolar concentrations. Since the aggregated and monomeric species are in slow exchange on the chemical shift timescale (i.e. the two species give rise to discrete resonances) it is possible to distinguish between the properties of the monomeric and aggregated species. Furthermore, NOEs observed in the monomeric peptide will not be the result of NOEs that are transferred from the aggregated to the monomeric species, since transfer of NOEs between species in slow exchange on the chemical shift timescale occurs only if the species are in fast exchange on the spin-lattice timescale (Clare & Gronenborn, 1982). The difference in chemical shifts of the two species indicate that the rate of interconversion, k , is much less than 10 s^{-1} and the spin-lattice relaxation rates, R_1 , of the P_{17-24} protons measured by inversion-recovery range from approximately 0.4 to 1.5 s^{-1} (data not shown). For the two species to be in

fast exchange on the spin-lattice timescale, k must be much greater than R_1 , and this is not the case.

$d_{\alpha N}(i, i+1)$ NOEs, but not $d_{NN}(i, i+1)$ NOEs, were observed for all the residues of P_{17-24} (Fig. 3A) indicating that the peptide exists predominantly in the β region of ϕ/ψ space (Dyson & Wright, 1991). The only interresidue NOEs observed were between the monomeric side-chain resonances of Ile19 and Tyr21 (Fig. 3B). These NOEs are negative in the rotating frame and thus do not result from spin diffusion (Bothner-By *et al.*, 1984) and are also observed in NOESY spectra of $[R]_{Ala}$ recorded with a 400 ms mixing time. Thus, for at least some of the time, the Ile19 and Tyr21 side-chain protons are within 5 Å of each other. In native BPTI, the Ile19 γCH_3 protons are between 5.2 and 7.8 Å from the Tyr21 γCH and δCH protons (Fig. 3C). Thus, the NOE data imply that these protons are significantly closer in the denatured than in the native state.

The proximity of the Ile19 and Tyr21 side-chains might be thought to be a natural consequence of an extended random conformation of P_{17-24} , in which the side-chains of every other residue are brought close to one another randomly but sufficiently often to result in the NOEs and anomalous chemical shifts. If this were the case, then such effects would be expected to be observed throughout the peptide. Instead, the NOE data and Ala substitutions (see below and Table 1) indicate that the majority of the side-chains of P_{17-24} do not interact with those of alternating residues. Moreover, the Ile19 γCH_3 resonance shifts linearly downfield towards the random coil chemical shift as the temperature is increased with a temperature coefficient of +1 p.p.b./deg.C, such that the chemical shift dispersion between the Ile γCH_3 resonances is reduced by about 0.08 p.p.m. at 85°C relative to that seen at 5°C (data not shown). In contrast, the chemical shifts of the carbon bound protons of the other residues are not significantly affected by temperature, and have temperature coefficients in the range 0.1 to 0.5 p.p.b./deg.C (data not shown). In addition, the dispersion of the aliphatic proton chemical shifts of P_{17-24} is much reduced in DMSO, a solvent that has been observed to produce a very disordered state in other proteins (Broadhurst *et al.*, 1991; Evans *et al.*, 1991). In particular, the chemical shift dispersion of the Ile resonances is much lower in DMSO (Fig. 1C) than in water (Fig. 1B). This is not a general effect, since DMSO did not affect significantly the dispersion of the aromatic side-chain resonances (Fig. 2D). Collectively, these results suggest that the chemical shift dispersion of the Ile resonances results from specific interactions. A truly random ensemble of all available conformations of the Ile19 γCH_3 group would not be expected to have the observed temperature-dependent chemical shift or to be affected by DMSO.

Intramolecular hydrogen bonds involving amide NH protons result in a reduced temperature dependence of the amide NH chemical shift (reviewed by Rose *et al.*, 1985). The temperature coefficients of the amide NH resonances, except Asn24, were -8

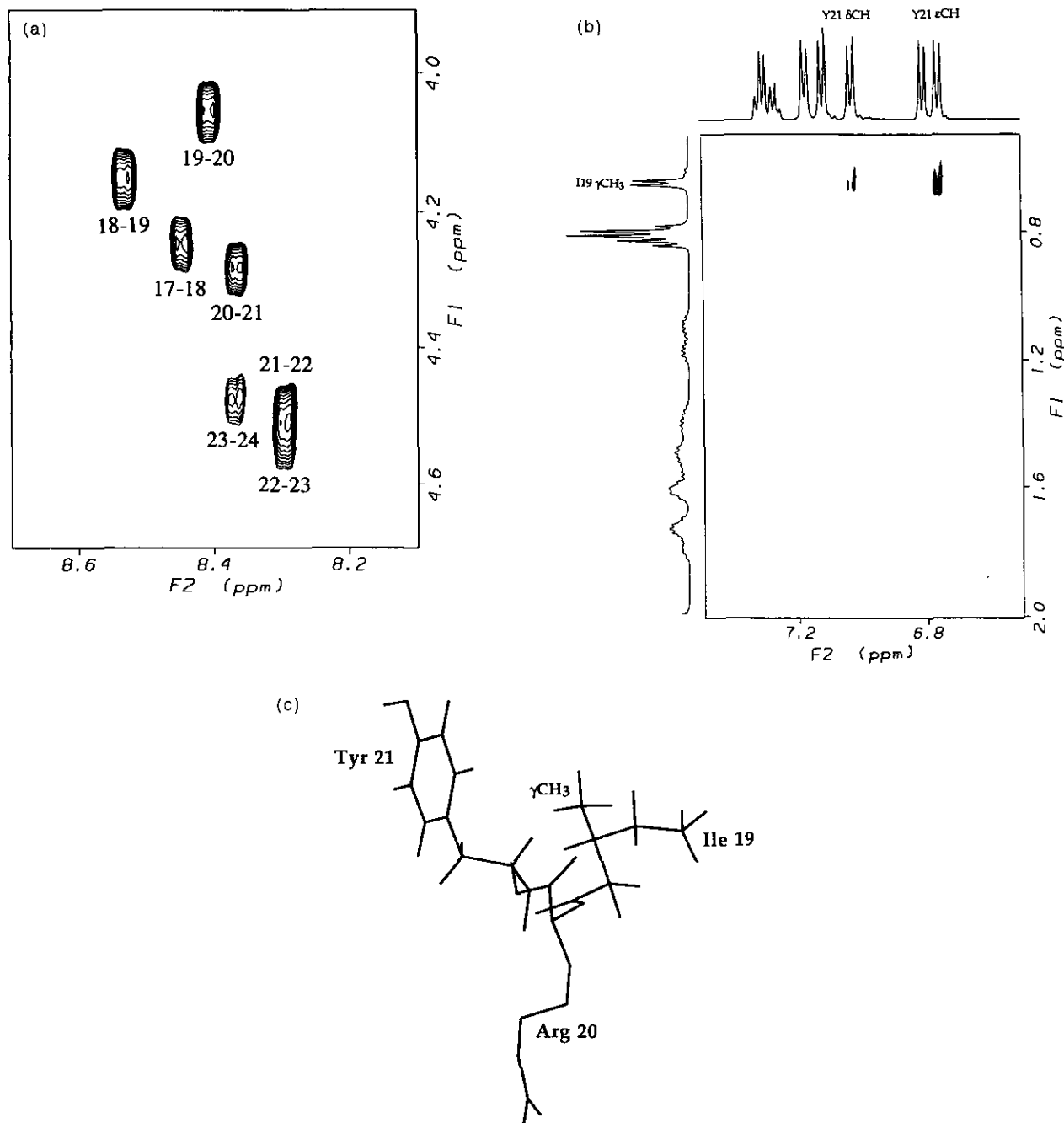


Figure 3. A, $d_{\alpha N}(i, i+1)$ but not $d_{NN}(i, i+1)$ NOEs are observed for P_{17-24} . This indicates that the peptide exists predominantly in the β region of ϕ/ψ space and provides the basis for assignment of the peptide. NOESY data were acquired at 5°C using a mixing time of 200 ms. B, Interresidue NOEs are observed between Ile19 γCH_3 and the Tyr21 aromatic side-chain protons of P_{17-24} . NOESY data were acquired at 5°C using a mixing time of 400 ms. C, The relative positions of Ile19 and Tyr21 in the crystal structure of BPTI. Ile19 is positioned above the plane of the aromatic ring of Tyr21 in the native state. Also shown is Arg20. The Figure was drawn using the Brookhaven Protein Data Bank coordinates 5PTI (Wlodawer *et al.*, 1987). Experimental details are given in the legend to Fig. 1, except data shown in B was sensitivity-enhanced using exponential multiplication in F1, rather than resolution-enhanced.

to -9 p.p.b./deg.C (data not shown), suggesting that these protons are exposed to solvent and are not involved in intramolecular hydrogen bonds. Asn24 NH has a slightly reduced temperature coefficient of -6 p.p.b./deg.C, both in P_{17-24} and in the longer peptide P_{14-28} (indicating that the reduced temperature coefficient of Asn24 NH in P_{17-24} is not the result of an end effect). Although

in general temperature coefficients smaller than -6 p.p.b./deg.C are observed for protons involved in intramolecular hydrogen bonds, the reduced temperature coefficient of Asn24 NH may reflect the existence of some degree of hydrogen bonding. However, NOEs characteristic of turn formation (Wüthrich, 1986) were not observed for P_{17-24} or for P_{14-28} , nor did substitution of the aromatic residues

Table 1
Effect of Ala substitutions in P_{17-24} on chemical shift

	R17A†	I18A	I19A	R20A†	Y21A	F22A	Y23A	N24A†
R17	N/A		NH +0.06					
I18		N/A	NH -0.06					
I19		NH -0.07 αCH -0.08	N/A		NH +0.06 γCH ₃ -0.18			
R20				N/A		NH +0.11 αCH +0.05 γCH +0.10 εNH +0.07	NH +0.08 εNH +0.05	
Y21		δCH -0.05	NH -0.17 δCH -0.06		N/A	NH +0.10 βCH +0.09	βCH +0.05, +0.10	
F22		αCH -0.05 βCH -0.08 εCH -0.07	αCH -0.06 βCH -0.08 δCH -0.06		βCH -0.11	N/A		
Y23		αCH -0.05 βCH +0.05	αCH -0.07		βCH +0.05		N/A	
N24		βCH -0.14 +0.06	βCH -0.06		βCH -0.14, +0.07		αCH +0.05 βCH +0.09, +0.10	N/A

The Table lists resonances that have a change in chemical shift resulting from the indicated Ala substitution of at least 0.05 p.p.m., which is approximately 3 times the average change in chemical shift observed as a result of the Ala substitutions. Changes in chemical shifts were measured at 5°C and pH 3.0 and are accurate to 0.02 p.p.m. (if measured from DQF COSY data) or 0.01 p.p.m. in the case of the R17A and R20A variants (measured from 1D NMR data). A negative change in chemical shift corresponds to a downfield shift as a result of the Ala substitution.

† For these substitutions, changes in chemical shift were 0.02 p.p.m. or smaller. The low solubility of the R17A and R20A variants precluded 2D NMR analysis, and chemical shift data could be obtained only from 1D spectra for the Ile18 and 19γCH, Ile19 γCH₃, and aromatic side-chain protons.

by Ala (see below) result in changes of the Asn24 NH chemical shift (Table 1), suggesting that an NH-aromatic interaction (reviewed by Burley & Petsko, 1988) involving Asn24 NH does not exist. Thus, the origins of the slightly reduced temperature coefficient of Asn24 NH remain ambiguous.

Ala scanning of P_{17-24}

To map the interresidue interactions within P_{17-24} , single Ala substitutions were made throughout the peptide (Cunningham & Wells, 1989). Changes in chemical shift as a result of the Ala substitution were then used to probe the existence of interactions between the side-chains of the substituted residue and others in the peptide. Significant changes in the chemical shift of the peptide protons result from Ala substitutions of Ile18, Ile19, Tyr21, Phe22 and Tyr23 (Table 1). A significant change is taken to be at least 0.05 p.p.m., which is approximately three times the average change in chemical shift as a result of the Ala substitutions.

The most striking change in chemical shift arises from the Y21A substitution, which results in essentially complete loss of chemical shift dispersion between the resonances of the Ile residues (Fig. 1D). This result suggests strongly that Ile19 and Tyr21 interact, in accord with the NOE data, and that the upfield shifts of the Ile19 side-chain resonances are due predominantly to a ring current interaction with Tyr21. The Ile19 γCH₃ resonance is shifted upfield of Ile18 γCH₃ in P_{17-24} , implying that this methyl group is positioned at least some of the time above the plane of the aromatic ring of Tyr21 (Perkins, 1982). The I18A substitution affects

significantly the chemical shifts of the aromatic side-chains and Ile19 resonances, and the I19A substitution affects the chemical shifts of Ile18, suggesting that Ile18 is also involved in the cluster (Table 1).

The proximity of the Ile19 γCH₃ proton to the ring of Tyr21 indicated by the NOE data would be expected to result in large ring current contributions to the chemical shift of the Ile19 resonances, probably of the order of a p.p.m. (Perkins, 1982). The difference in chemical shifts between the Ile18 and Ile19 resonances is an order of magnitude less than this, suggesting that the chemical shifts of the Ile19 resonances are averaged over both the conformation that gives rise to the NOE and other conformations. While it is difficult to estimate the population of the conformation that gives rise to the NOE between Ile19 and Tyr21, Wüthrich and co-workers expect that such an NOE would be detectable if these residues were in close proximity at least approximately 10% of the time (Neri *et al.*, 1992).

The low solubility of the R17A and R20A variants precluded detailed analysis by 2D NMR. 1D spectra indicate, however, that the Arg to Ala substitutions do not affect within experimental error (± 0.01 p.p.m.) the chemical shifts of either the aromatic side-chain or Ile19 γCH₃ resonances. The F22A and Y23A substitutions affect the resonances of Arg20, Tyr21 and Asn24 only, while the N24A substitution did not affect within experimental error the chemical shifts of the other residues in the peptide (Table 1). These changes in chemical shift are not reciprocal. For example, the I18A substitution affected the chemical shifts of Phe22 and Tyr23, but the Ala substitutions of these two aro-

matic residues did not perturb Ile 18. Thus, while these results suggest that Phe22, Tyr23 and Asn24 may have non-random interactions with other residues in P₁₇₋₂₄, it is not possible to describe the conformational behavior of these three residues on the basis of chemical shift data alone.

We conclude that Ile19 and Tyr21 form at least one non-random conformation in which the Ile γ CH₃ group is positioned above the plane of the Tyr21 aromatic ring. Ala substitutions suggest that Ile18 is also likely to be involved in an interaction with Ile19 and/or Tyr21 (Table 1). However, NOEs are not observed to define further the nature of the interaction.

Implications for protein folding

Interactions between hydrophobic residues also exist in the unfolded states of other proteins (Garvey *et al.*, 1989; Evans *et al.*, 1991; Broadhurst *et al.*, 1991; Neri *et al.*, 1992; Alexandrescu *et al.*, 1993), in peptide fragments of proteins (Dyson *et al.*, 1992a, b), in hormone polypeptides (Boesch *et al.*, 1978; Bundi *et al.*, 1978) and in peptide mimetics (Wiley & Rich, 1993). It has been proposed that small hydrophobic clusters are important in the earliest stages of protein folding, either by reducing the conformational freedom of the polypeptide chain, by excluding solvent (Dyson *et al.*, 1988; Evans *et al.*, 1991) or by acting as folding nucleation sites (Garvey *et al.*, 1989; Neri *et al.*, 1992). There is, however, no direct evidence at present that such clusters are important for the acquisition of native structure. Nevertheless, the inferred conformation of the main-chain of the Trp62-Trp63 cluster of denatured lysozyme (Evans *et al.*, 1991), the Ile27-Phe29 hydrophobic cluster in a peptide fragment of plastocyanin (Dyson *et al.*, 1992b) and the residual structure in urea denatured 434-repressor (Neri *et al.*, 1992) resemble the structure found in the corresponding folded proteins. Similarly, in denatured BPTI, we find that the side-chains of Ile19 and Tyr21 are positioned, at least qualitatively, in a similar orientation to that observed in the native state (Fig. 3C), suggesting that in the denatured state of BPTI a tendency may also exist to form native like structure.

We are especially indebted to L. P. McIntosh for many valuable discussions in all aspects of this work. We thank M. A. Milla for preparation of the [R]_{Ala} plasmid, M. W. Burgess, S. Britt, M. A. Milhollen and R. Rutkowski for peptide synthesis and mass spectrometry and D. J. Lockhart, C. J. McKnight, D. L. Minor, B. A. Schulman and J. P. Staley for many helpful discussions. P.S.K. is a Pew Scholar in the Biomedical Sciences. This research was supported by grant GM41307 from the National Institutes of Health.

References

Alexandrescu, A. T., Evans, P. A., Pitkeathly, M., Baum, J. & Dobson, C. M. (1993). Structure and

- dynamics of the acid-denatured molten globule state of α -lactalbumin: a two-dimensional NMR study. *Biochemistry*, **32**, 1707-1718.
- Amir, D. & Haas, E. (1988). Reduced bovine pancreatic trypsin inhibitor has a compact structure. *Biochemistry*, **27**, 8889-8893.
- Amir, D., Krausz, S. & Haas, E. (1992). Detection of local structures in reduced unfolded bovine pancreatic trypsin inhibitor. *Proteins: Struct. Funct. Genet.* **13**, 162-173.
- Boesch, C., Bundi, A., Oppliger, M. & Wüthrich, K. (1978). ¹H nuclear-magnetic-resonance studies of the molecular conformation of monomeric glucagon in aqueous solution. *Eur. J. Biochem.* **91**, 209-214.
- Bothner-By, A. A., Stephens, R. L. & Lee, J. (1984). Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame. *J. Amer. Chem. Soc.* **106**, 811-813.
- Broadhurst, R. W., Dobson, C. M., Hore, P. J., Radford, S. E. & Rees, M. L. (1991). A photochemically induced dynamic nuclear polarization study of denatured states of lysozyme. *Biochemistry*, **30**, 405-412.
- Bundi, A. & Wüthrich, K. (1979). ¹H-NMR parameters of the common amino acid residues measured in aqueous solution of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers*, **18**, 285-297.
- Bundi, A., Grathwohl, C., Hochmann, J., Keller, R. M., Wagner, G. & Wüthrich, K. (1975). ¹H-NMR of the protected tetrapeptides [trifluoroacetyl]-Gly-Gly-L-X-L-Ala-OCH₃, where X stands for one of the 20 common amino acids. *J. Magn. Reson.* **18**, 191-198.
- Bundi, A., Andreatta, R. H. & Wüthrich, K. (1978). Characterisation of a local structure in the synthetic parathyroid hormone fragment 1-34 by ¹H nuclear-magnetic-resonance techniques. *Eur. J. Biochem.* **91**, 201-208.
- Burley, S. K. & Petsko, G. A. (1988). Weakly polar interactions in proteins. *Advan. Protein Chem.* **39**, 125-189.
- Clore, G. M. & Gronenborn, A. M. (1982). Theory and applications of the transferred nuclear Overhauser effect to the study of the conformations of small ligands bound to proteins. *J. Magn. Reson.* **48**, 402-417.
- Cunningham, B. C. & Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, **244**, 1081-1085.
- Dill, K. A. & Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795-825.
- Dobson, C. M. (1992). Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* **2**, 6-12.
- Dyson, H. J. & Wright, P. E. (1991). Defining solution conformations of small linear peptides. *Annu. Rev. Biophys. Chem.* **20**, 519-538.
- Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988). Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. *J. Mol. Biol.* **201**, 161-200.
- Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. & Wright, P. E. (1992a). Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin. *J. Mol. Biol.* **226**, 795-817.
- Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A. & Wright, P. E. (1992b). Folding of peptide fragments comprising the complete sequence

- of proteins. Models for initiation of protein folding. II. Plastocyanin. *J. Mol. Biol.* **226**, 819–835.
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.
- Englander, S. W., Downer, N. W. & Teitelbaum, H. (1972). Hydrogen exchange. *Annu. Rev. Biochem.* **41**, 903–924.
- Evans, P. A., Topping, K. D., Woolfson, D. N. & Dobson, C. M. (1991). Hydrophobic clustering in nonnative states of a protein: interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins: Struct. Funct. Genet.* **9**, 248–266.
- Garvey, E. P., Swank, J. & Matthews, C. R. (1989). A hydrophobic cluster forms early in the folding of dihydrofolate reductase. *Proteins: Struct. Funct. Genet.* **6**, 259–266.
- Goldenberg, D. P. & Zhang, J. (1993). Small effects of amino acid replacements on the reduced and unfolded state of pancreatic trypsin inhibitor. *Proteins: Struct. Funct. Genet.* **15**, 322–329.
- Goodman, E. M. & Kim, P. S. (1989). Folding of a peptide corresponding to the α -helix in bovine pancreatic trypsin inhibitor. *Biochemistry*, **28**, 4343–4347.
- Gottfried, D. S. & Haas, E. (1992). Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. *Biochemistry*, **31**, 12353–12362.
- Gussakovsky, E. E. & Haas, E. (1992). The compact state of reduced bovine pancreatic trypsin inhibitor is not the compact molten globule. *FEBS Letters*, **308**, 146–148.
- Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553.
- Kemmink, K., van Mierlo, C. P. M., Scheek, R. U. & Creighton, T. E. (1993). Local structure due to an aromatic-amide interaction observed by ^1H -nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **230**, 312–22.
- Kosen, P. A., Creighton, T. E. & Blout, E. R. (1981). Circular dichroism spectroscopy of bovine pancreatic trypsin inhibitor and five altered conformational states. Relationship of conformation and the refolding pathway of the trypsin inhibitor. *Biochemistry*, **20**, 5744–5754.
- Kumar, A., Ernst, R. R. & Wüthrich, K. (1980). A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1–6.
- Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.
- Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **43**, 259–281.
- Marion, D. & Bax, A. (1988). Baseline distortion in real-Fourier-transform NMR spectra. *J. Magn. Reson.* **352**–356.
- Marion, D. & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967–974.
- Marion, D., Ikura, M. & Bax, A. (1989). Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. *J. Magn. Reson.* **84**, 425–430.
- Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*, **257**, 1559–1563.
- Perkins, S. J. (1982). Application of ring current calculations to the proton NMR of proteins and transfer RNA. *Biol. Magn. Reson.* **4**, 193–336.
- Piantini, U., Sørensen, O. W. & Ernst, R. R. (1982). Multiple quantum filters for elucidating NMR coupling networks. *J. Amer. Chem. Soc.* **104**, 6800–6801.
- Rance, M. (1987). Improved techniques for homonuclear rotating-frame and isotropic mixing experiments. *J. Magn. Reson.* **74**, 557–564.
- Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983). Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479–485.
- Roder, H. (1988). Structural characterisation of protein folding intermediates by proton magnetic resonance and hydrogen exchange. *Methods Enzymol.* **176**, 446–473.
- Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985). Turns in peptides and proteins. *Advan. Protein Chem.* **37**, 1–109.
- Shaka, A. J. & Freeman, R. (1983). Simplification of NMR spectra by filtration through multiple-quantum coherence. *J. Magn. Reson.* **51**, 169–173.
- Shortle, D. (1993). Denatured states of proteins and their roles in folding and stability. *Curr. Opin. Struct. Biol.* **3**, 66–74.
- Staley, J. P. (1993). Structural studies of early intermediates in the folding pathway of bovine pancreatic trypsin inhibitor. Ph.D. Thesis, Massachusetts Institute of Technology, U.S.A.
- Staley, J. P. & Kim, P. S. (1992). Complete folding of bovine pancreatic trypsin inhibitor with only a single disulfide bond. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 1519–1523.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
- Tanford, C. (1968). Protein denaturation. *Advan. Protein Chem.* **23**, 121–282.
- Wiley, R. A. & Rich, D. H. (1993). Peptidomimetics derived from natural products. *Med. Res. Rev.* **13**, 327–384.
- Wlodawer, A., Deisenhofer, J. & Huber, J. (1987). Comparison of two highly refined structures of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **193**, 145–156.
- Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. Wiley, New York.

Edited by P. E. Wright

(Received 13 September 1993; accepted 29 October 1993)