amino acids play a beneficial role in forming the necks of a loop. Under this hypothesis, lysine and arginine should occur more frequently in the necks of loops than elsewhere. However, these amino acids are somewhat underrepresented in these regions: of the proteins in the Leszczynski and Rose data set, lysine and arginine account for 6.9% and 2.4% of the residues at loop borders and 7.1% and 3.1% of the total residues, respectively.

An alternative possibility is suggested by the dual-site model considered above. Under this hypothesis, binding of a processing endoprotease to a recognition site on the exposed portion of the loop would induce a conformational change that would bring the basic site into juxtaposition with the catalytic site on the enzyme. An interesting feature of this mechanism is that the conformational shift induced by enzyme binding could conceivably strain peptide bonds in the neck region, aiding in bond scission.

ACKNOWLEDGMENTS

We thank L. Douglas for advice and encouragement, Dr. A. Tesler for graphics, and Dr. R. Chisholm for database searches.

Registry No. Procalcitonin, 56645-65-9; procalcitonin gene-related peptide, 95077-07-9; procorticotropin-releasing factor, 99533-36-5; proenkephalin A, 88402-54-4; proglucagon, 55963-74-1; progonadotropin-releasing hormone, 123774-88-9; pro-growth-hormone-releasing factor, 96511-22-7; proinsulin, 9035-68-1; proinsulin-like growth factor II, 93928-21-3; proopiomelanocortin, 66796-54-1; prepara-thyroid hormone, 50815-54-8; prorelaxin, 87004-01-1; prorenin, 39364-01-7; prosomatostatin, 74315-46-1; prothyrotropin-releasing hormone, 98616-54-7; prourotensin, 123774-89-0; provasoactive intestinal peptide, 81032-42-0.

REFERENCES

Berry, R. W. (1981) *Biochemistry 20*, 6200–6205.
Berry, R. W., & Yates, M. E. (1986) *Peptides 7*, 637–643.
Creminon, C., Rholam, M., Bousetta, H., Marakchi, N., & Cohen, P. (1988) *J. Chromatogr. 440*, 439–448.
Gluschankof, P., & Cohen, P. (1987) *Neurochem. Res. 12*, 951–958.
Gluschankof, P., Gomez, S., Lepage, A., Creminon, C., Nyberg, F., Terenius, L., & Cohen, P. (1988) *FEBS Lett. 234*, 149–152.
Leszczynski, J. F., & Rose, G. D. (1986) *Science 234*, 849–855.
Mains, R. E., & Eipper, B. A. (1980) *Ann. N.Y. Acad. Sci. 343*, 94–108.
Ralph, W. W., Webster, T., & Smith, T. F. (1987) *Comput. Appl. Biosci. 3*, 211–216.
Rholam, M., Nicolas, P., & Cohen, P. (1986) *FEBS Lett. 207*, 1–6.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rose, G. D., Gierasch, L. M., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1–109.
Schwartz, T. W. (1986) *FEBS Lett. 200*, 1–10.

Sequential Resonance Assignment and Secondary Structure Determination of the *Ascaris* Trypsin Inhibitor, a Member of a Novel Class of Proteinase Inhibitors†

Angela M. Gronenborn,*,‡ Michael Nilges,‡,‖ Robert J. Peanasky,§ and G. Marius Clore*,‡

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, and Department of Biochemistry, School of Medicine, The University of South Dakota, 414 East Clark, Vermillion, South Dakota 57069-2390

Received May 2, 1989; Revised Manuscript Received July 20, 1989

ABSTRACT: The solution conformation of the *Ascaris* trypsin inhibitor, a member of a novel class of proteinase inhibitors, has been investigated by nuclear magnetic resonance spectroscopy. Complete sequence-specific assignments of the ^{1}H NMR spectrum have been obtained by using a number of two-dimensional techniques for identifying through-bond and through-space (<5-Å) connectivities. Elements of regular secondary structure have been identified on the basis of a qualitative interpretation of the nuclear Overhauser enhancement, coupling constant, and amide exchange data. These are two β-sheet regions. One double-stranded antiparallel β-sheet comprises residues 11–14 (strand 1) and 37–39 (strand 2). The other triple-stranded sheet is formed by two antiparallel strands comprising residues 45–49 (strand 4) and 53–57 (strand 5) connected by a turn (residues 50–52), and a small strand consisting of residues 20–22 (strand 3) that is parallel to strand 4.

Over the last two decades a large number of primary sequences of protein proteinase inhibitors from a variety of species have been determined [see Laskowski and Kato (1980) for a review]. On the basis of sequence comparisons and functional studies, the serine proteinases have been classified into at least 10 families among which the best known are (i) the pancreatic trypsin inhibitors (Kunitz type), (ii) the pancreatic secretory trypsin inhibitors (Kazal type), and (iii) the *Streptomyces* subtilisin inhibitor family (Laskowski & Kato, 1980). Classification into families is based on extensive amino acid sequence homology, in particular at and surrounding the reactive site, as well as on the topological relationship between disulfide bridges and the location of the reactive site loop. Most members of these families inhibit proteinases according to a common mechanism, forming a substrate-like enzyme–

†This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (A.M.G. and G.M.C.) and by the National Institute of Allergy and Infectious Diseases (R.J.P.).

‡National Institutes of Health.
‖Present address: The Howard Hughes Medical Institute and Department of Molecular Biochemistry and Biophysics, Yale University, 260 Whitney Ave., P.O. Box 6666, New Haven, CT 06511.
§The University of South Dakota.

0006-2960/90/0429-183$02.50/0

inhibitor complex with the reactive site peptide bond located on the surface of the inhibitor bound in the active site of the cognate enzyme. Most of the protein inhibitors are small proteins of less than 120 residues. Detailed structural information on a number of inhibitors as well as on inhibitor–enzyme complexes has been obtained by X-ray crystallography (Huber et al., 1970, 1974; Rühlmann et al., 1973; Sweet et al., 1974; Deisenhofer & Steigemann, 1975; Sato et al., 1978; Mitsui et al., 1978; Weber et al., 1981; Papamokos et al., 1982; Fujinaga et al., 1982; Read et al., 1983; Hirono et al., 1984; Bode et al., 1985, 1989; McPhalen et al., 1985a,b). More recently, structures in solution by nuclear magnetic resonance (NMR)[1] spectroscopy of some proteinase inhibitors have become available (Williamson et al., 1985; Wagner et al., 1987; Clore et al., 1987a,b; Folkers et al., 1989; Heitz et al., 1989). The results of these studies indicate that while there is a clear structural homology between members of a family, no such homology exists between proteins of different families.

The newly established *Ascaris* trypsin inhibitor family (Babin et al., 1984; Laskowski, 1986; Peanasky et al., 1987a–c) contains several members with five chymotrypsin/elastase inhibitors and three trypsin inhibitors identified to date (Goodman et al., 1983; Babin et al., 1984). The primary sequences for the chymotrypsin/elastase inhibitors and one trypsin inhibitor have been recently determined (Babin et al., 1984; Peanasky et al., 1987c), allowing alignment of the linear sequences with respect to the reactive site and the 10 half-cystines. The presence of five disulfide bonds within the single-domain inhibitor represents the highest number of disulfides found in serine proteinase inhibitors to date. A further novel feature found for the members of this family is the presence of a half-cystine in the P2′ position of the reactive site loop [notation of Schecter and Berger (1967)].

The lack of structural information on any member of this novel family prompted us to embark on a NMR study of the trypsin inhibitor (ATI) of *Ascaris lumbricoides var. suum* with the final aim of determining its three-dimensional structure in solution. This paper presents the sequential assignment of the ^{1}H NMR spectrum of the inhibitor and the determination of its secondary structure.

Experimental Procedures

Purification and Sample Preparation. ATI was prepared from *Ascaris suum* obtained at the local abattoir from the intestinal contents of hogs. The worms were collected in a Dewar flask at 37 °C containing a balanced salt medium, one that is isosmotic with the parasites' pseudocoelomic fluid (Baldwin & Moyle, 1947). The urogenital tract of the worm was dissected away, and the remnants of the worms were packaged in 100-g lots and stored at −20 °C. The purification of ATI was performed as described by Goodman and Peanasky (1982). Chopped worms were diluted in 4 volumes of cold distilled H_2O and homogenized. The debris was removed and the sample clarified by centrifugation first at 11000*g* and then at 59000*g*. After treatment of the supernatant at pH 1.9 and 37 °C for 75 min, the precipitate that formed at pH 5.5 was removed. ATI in the clear supernatant was precipitated at pH 8.7 by raising the ammonium sulfate concentration to 0.9 saturation with salt. The ATI in the dissolved precipitate was removed at neutral pH by affinity chromatography using porcine trypsin liganded to Sepharose 4B. The inhibitor was released from the affinity gel complex by a sudden change to pH 1.70, removal of the affinity gel, then neutralization and lyophilization. ATI was desalted in 10 mM HCl on a Sephadex G-25 column and lyophilized. The isoforms were resolved on a CM-Sephadex column equilibrated and eluted with 30 mM borate buffer, pH 8.1. Following two small breakthrough peaks, 80% of the trypsin inhibiting activity eluted as a single isoform. Three more small inhibitor peaks, accounting for the remainder of the trypsin inhibiting activity, can be eluted with a linear salt gradient. For these studies, the major isoform was lyophilized and desalted on a Sephadex G-25 column in 10 mM HCl. The salt-free hydrochloride of the inhibitor was lyophilized. ATI is a single component in polyacrylamide gel electrophoresis at pH 9.3.

Samples for NMR spectroscopy contained approximately 3.3 mM ATI in either 90% H_2O/10% D_2O or 99.996% D_2O, pH 2.4.

NMR Spectroscopy. NMR measurements were carried out at 600 MHz on a Bruker AM-600 spectrometer equipped with digital-phase shifters, a "reverse"-mode ^{1}H probe, and an Aspect 3000 computer. All two-dimensional spectra were recorded in the pure-phase absorption mode using the time-proportional incrementation method (Redfield & Kunz, 1975) as described by Marion and Wüthrich (1983). NOESY spectra (Jeener et al., 1979; Macura et al., 1981) with mixing times of 100, 150, and 200 ms and HOHAHA spectra (Braunschweiler & Ernst, 1983; Davis & Bax, 1985) with a WALTZ17$_y$ mixing sequence (Bax, 1989) of 30–60-ms duration sandwiched between 1.5-ms trim pulses were recorded at 15, 27, and 40 °C.

For NOESY spectra recorded in H_2O, the water resonance was suppressed by using a semiselective "jump–return" pulse (Plateau & Gueron, 1982) in place of the last 90° pulse in the NOESY sequence. Water suppression in the HOHAHA spectra was achieved by using a 90° "flip-back" pulse, a fixed recovery delay set to half the mixing time to remove rotating-frame NOE effects in a manner analogous to that described by Griesinger et al. (1988), and the "jump–return" sequence after the WALTZ17$_y$ mixing sequence (Bax et al., 1987). The effects of radiation damping were minimized by cycling the phase of the preparation pulse 45° out of register with those of the evolution and detection periods (Driscoll et al., 1989). Optimization of the receiver phase to eliminate base-line distortions was carried out according to Marion and Bax (1988b), and suppression of t_2 ridges was achieved by zeroing the first point of each FID as well as linear base-line corrections of both the initial FIDs prior to Fourier transformation in F_2 and the F_2 cross sections prior to Fourier transformation in F_1 (Driscoll et al., 1989).

Typically, 800–1024 increments of 2K data points were collected for each NOESY and HOHAHA experiment resulting in 2D spectra with a digital resolution of the order of 6–8 Hz/point in each dimension after zero filling.

For the P.COSY (Marion & Bax, 1988a) and PE.COSY (Mueller, 1987) experiments, a reference one-dimensional FID was recorded with a 0° mixing pulse, $t_1 = 0$ μs, and 128 scans, 8 times the number of scans used per increment in the COSY experiment recorded with either a 70° (P.COSY) or 35° (PE.COSY) mixing pulse. By left shifting the data of this single reference FID, the time domain data for successive values of the 0° COSY experiment were obtained. The 0° COSY time domain data were then subtracted from the 70° or 35° COSY time domain data to generate the P.COSY or

[1] Abbreviations: NMR, nuclear magnetic resonance; ATI, trypsin inhibitor from *Ascaris suum*; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; COSY, two-dimensional correlated spectroscopy; P.COSY, primitive COSY; PE.COSY, primitive exclusive COSY; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy.

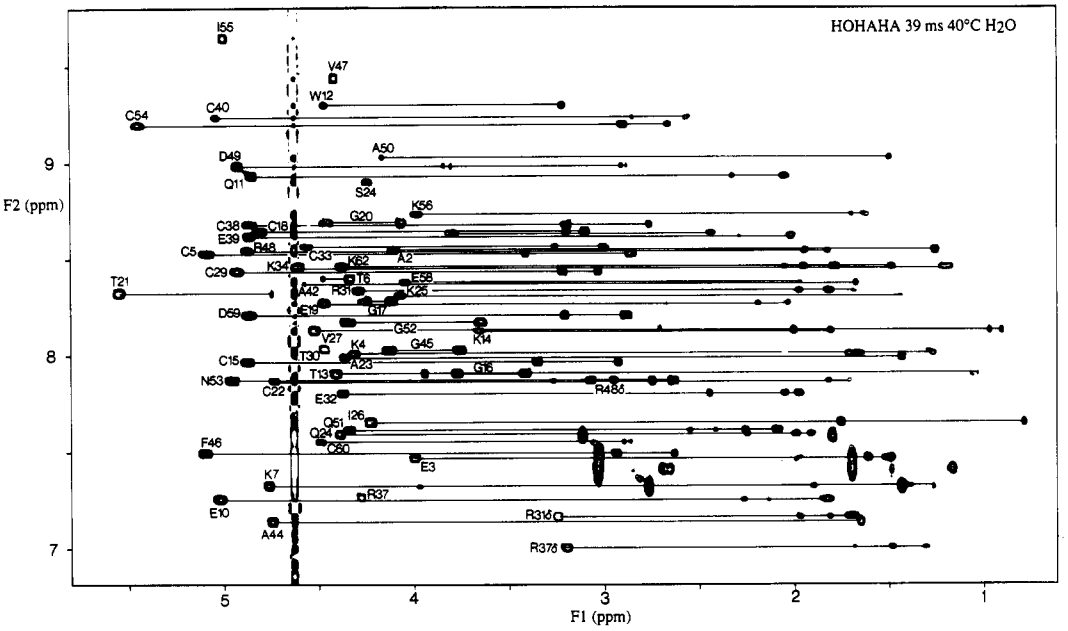

FIGURE 1: NH (F_2 axis)–aliphatic (F_1 axis) of the 39-ms HOHAHA spectrum of ATI in H_2O at 40 °C. Labels are at the direct NH–C$^\alpha$H connectivities, and spin systems are indicated by the continuous lines. Some connectivities from the Arg N$^\epsilon$H protons (F_2 axis) are also indicated with the label at the Arg C$^\delta$H proton (indicated by a δ after the residue name).

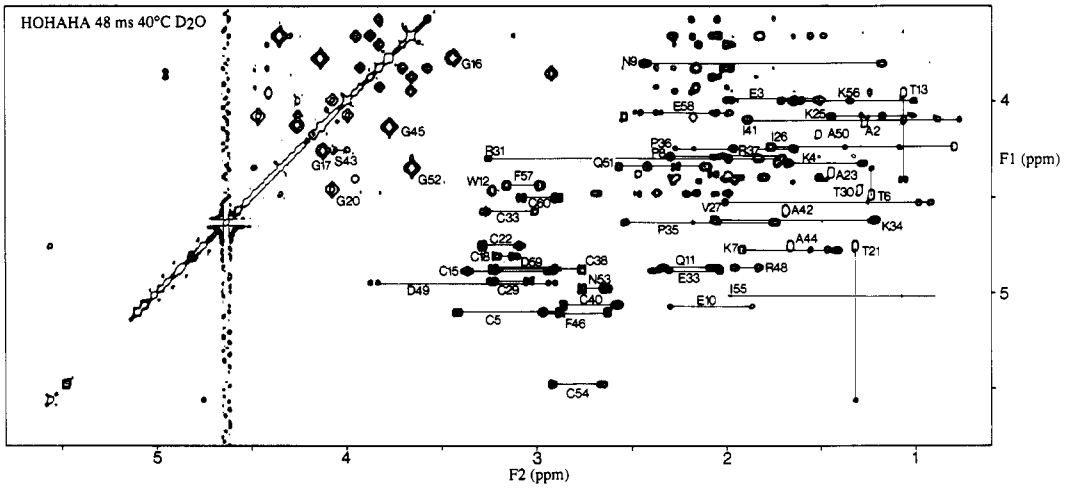

FIGURE 2: C$^\alpha$H (F_1 axis)–aliphatic (F_2 axis) region of the 48-ms HOHAHA spectrum of ATI in D_2O at 40 °C. Some spin systems are indicated.

PE.COSY time domain data, respectively. By this means the dispersive character of the diagonal in a regular COSY spectrum is purged. The sensitivity of the resulting P.COSY spectrum is approximately twice that of a conventional DQF-COSY spectrum recorded in the same measuring time (Marion & Bax, 1988a). PE.COSY spectra were recorded in both H_2O and D_2O, and a P.COSY spectrum at 40 °C was recorded in H_2O. For these experiments 1024 increments of 4K data points were collected, and the digital resolution in F_2 was 1.7 Hz/point.

Results and Discussion

The assignment of the ^{1}H NMR spectrum of ATI was achieved by using two-dimensional NMR spectroscopy (Ernst et al., 1987) to identify intraresidue through-bond connectivities in combination with sequential interresidue through-space (<5-Å) connectivities following well-established procedures [see Wüthrich (1986) and Clore and Gronenborn (1987) for reviews]. Direct through-bond connectivities were unambiguously established by using the P.COSY and PE.COSY spectra, while the HOHAHA spectra were used to identify direct, single and multiple relayed connectivities as the mixing time was successively increased. Examples of HOHAHA spectra in H_2O and D_2O illustrating a number of spin systems are shown in Figures 1 and 2, respectively. Side-chain resonances of the various spin systems were predominantly identified in the D_2O HOHAHA spectrum, and the H_2O spectrum served to identify exchanging side-chain amide protons of arginine, lysine, glutamine, and asparagine residues. NOESY spectra displaying a range of sequential connectivities are shown in Figure 3–5. C$^\alpha$H(i)–NH(i+1) and C$^\beta$H(i)–NH(i+1) NOE connectivities are shown in Figure 3, NH(i)–NH(i+1) NOE connectivities in Figure 4, and C$^\alpha$H(i)–ProC$^\delta$H(i+1) NOE connectivities in Figure 5. The complete list of assignments is given in Table I.

Several sets of spectra at three temperatures, 15, 27, and 40 °C, were recorded to resolve problems associated with spectral overlap. Further, the higher temperature of 40 °C proved necessary to sharpen up peaks which were too broad to be observed at 27 °C. At 40 °C all amide proton resonances could be observed, except for that of the N-terminal glutamine. We found only one position where the identification of an amide resonance was not straightforward, namely, Ile-41, which exhibited cross peaks between the amide resonance and

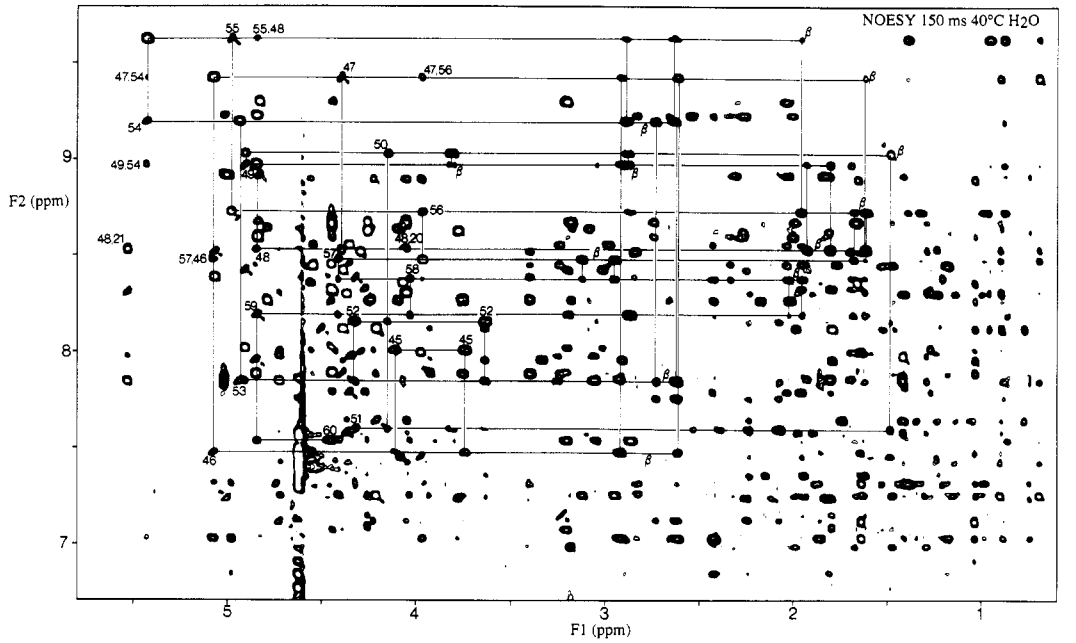

FIGURE 3: NH/aromatic (F_2 axis)–aliphatic (F_1 axis) region of the 150-ms NOESY spectrum in H_2O at 40 °C. Some $C^{\alpha}H(i)$–$NH(i+1)$ and $C^{\beta}H(i)$–$NH(i+1)$ sequential NOE connectivities are indicated by continuous lines with the labels at the position of the intraresidue $C^{\alpha}H(i)$–$NH(i)$ and $C^{\beta}H(i)$–$NH(i)$ cross peaks. (The latter are denoted by the letter β.) Also indicated are some long-range $C^{\alpha}H(i)$–$NH(j)$ NOEs (labeled in the order i,j).

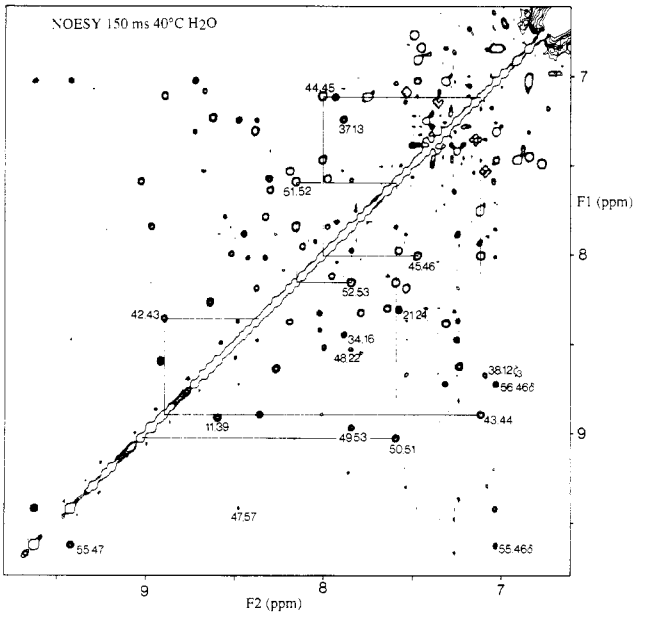

FIGURE 4: NH (F_1 axis)–NH (F_2 axis) region of the 150-ms NOESY spectrum in H_2O at 40 °C. Two stretches of sequential $NH(i)$–$NH(i+1)$ connectivities in the turn region from residues 42 to 46 and 50 to 53 are indicated by continuous lines. Also shown are a number of long-range $NH(i)$–$NH(j)$ NOEs.

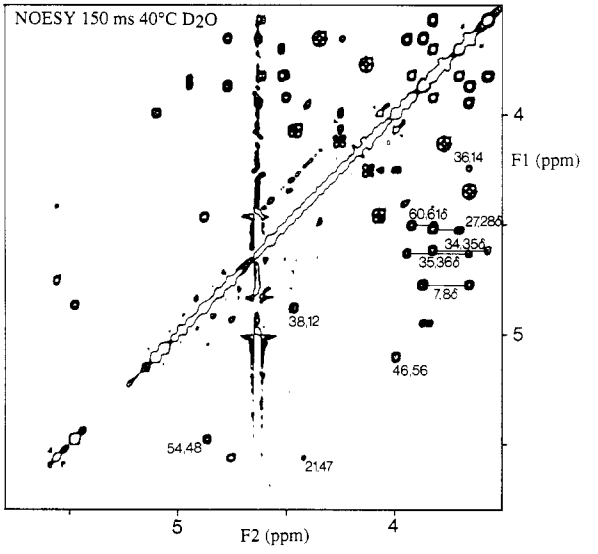

FIGURE 5: $C^{\alpha}H/ProC^{\delta}H$ (F_1 axis)–$C^{\alpha}H/ProC^{\delta}H$ (F_2 axis) region of the 150-ms NOESY spectrum in D_2O at 40 °C. The sequential $C^{\alpha}H(i)$–$ProC^{\delta}H(i+1)$ NOEs together with some long-range $C^{\alpha}H(i)$–$C^{\alpha}H(j)$ NOEs are indicated.

other intraresidue resonances that were either broadened and very weak or had a double-lobed appearance. Increasing the temperature to 50 °C resolved this ambiguity, since at the elevated temperature the amide resonance sharpened up and cross peaks appeared as single peaks. The simplest explanation for the behavior of the Ile 41 amide resonance is the possibility that at the lower temperatures this proton resides in two conformations, resulting in exchange broadening of the resonance.

Regular protein secondary structure elements can be deduced from an analysis of the backbone NOEs (Wüthrich et al., 1984). A summary of the short-range ($i - j \leq 3$) NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons, as well as the $C^{\delta}H$ protons of proline, is shown in Figure 6. As can be deduced from the large number of strong consecutive $C^{\alpha}H(i)$–$NH(i+1)$ NOE connectivities, the protein consists mainly of β-strands, interrupted by turns or loops. Indeed, $C^{\alpha}H(i)$–$NH(i+1)$ or $C^{\alpha}H(i)$–$ProC^{\delta}H(i+1)$ connectivities are observed along the entire polypeptide chain, starting from Ala-2. Corroborating evidence for the predominance of β-strand structures can be drawn from the values of the $^3J_{HN\alpha}$ coupling constants, which are also listed in Figure 6. They were classified into three ranges, $^3J_{HN\alpha} < 6$ Hz, 6 Hz $< {}^3J_{HN\alpha} < 8$ Hz, and $^3J_{HN\alpha} \geq 8$ Hz. Large coupling constants are found for ϕ angles around −140°, characteristic of a β-strand structure, and in the present case approximately 50% of the $^3J_{HN\alpha}$ coupling constants that could be measured (23 out of 49) fall in the last category. There is no evidence for any α-helical segments, which would be characterized by a stretch of consecutive strong $NH(i)$–$NH(i+1)$ NOE connectivities in combination with a set of $C^{\alpha}H(i)$–$NH(i+2,3,4)$ and $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ connectivities

Table I: ^{1}H NMR Chemical Shifts of ATI at 40 °C and pH 2.4 (in ppm)[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
E1		4.08	2.17, 2.17	$C^{\gamma}H$ 2.57, 2.57
A2	8.54	4.10	1.26	
E3	7.46	3.99	1.61, 1.51	$C^{\gamma}H$ 1.98, 1.98
K4	8.00	4.30	1.67, 1.71	$C^{\gamma}H$ 1.28, 1.28; $C^{\delta}H$ 1.43, 1.43; $C^{\epsilon}H$ 2.77; $N^{\zeta}H_3^+$ 7.31
C5	8.52	5.08	3.41, 2.85	
T6	8.39	4.33	4.48	$C^{\gamma 2}H_3$ 1.23
K7	7.32	4.76	1.42, 1.92	$C^{\gamma}H$ 1.56, 1.47; $C^{\gamma}H$ 1.80, 1.69; $C^{\epsilon}H$3.05; $N^{\zeta}H_3^+$ 7.56
P8		4.27	2.28, 1.83	$C^{\gamma}H$ 2.07, 2.03; $C^{\delta}H$ 3.87, 3.66
N9	8.63	3.79	2.43, 1.19	$N^{\delta}H_2$ 7.03, 6.85
E10	7.24	5.02	1.84, 1.82	$C^{\gamma}H$ 2.27, 2.12
Q11	8.92	4.84	2.05	$C^{\gamma}H$ 2.33
W12	9.30	4.46	3.22, 3.22	$C^{\delta 1}H$ 7.08; $C^{\zeta 2}H$ 7.38; $C^{\eta}H$ 7.16; $C^{\zeta 3}H$ 7.10; $C^{\epsilon 3}H$ 7.55; $N^{\epsilon}H$ 10.06
T13	7.90	4.40	3.94	$C^{\gamma}H_3$ 1.04
K14	8.13	3.65	1.82, 1.82	$C^{\gamma}H$ 1.55, 1.49; $C^{\delta}H$ 1.81; $C^{\epsilon}H$ 3.12; $N^{\zeta}H_3^+$ 7.58
C15	7.96	4.86	3.35, 2.92	
G16	7.89	3.76, 3.42		
G17	8.27	4.25, 4.12		
C18	8.64	4.80	3.19, 3.10	
E19	8.27	4.46	2.18, 2.04	$C^{\gamma}H$ 2.67, 2.44
G20	8.68	4.46, 4.07		
T21	8.32	5.54	4.74	$C^{\gamma}H_3$ 1.30
C22	7.86	4.74	3.26, 3.07	
A23	7.98	4.35	1.43	
Q24	7.59	4.38	1.99, 1.91	$C^{\gamma}H$ 2.26, 1.82; $N^{\epsilon}H_2$ 7.46, 6.85
K25	8.31	4.06	1.43	$C^{\gamma}H$ 1.01; $C^{\delta}H$ 1.17; $C^{\epsilon}H$ 2.70, 2.66; $N^{\zeta}H_3^+$ 7.41
I26	7.65	4.22	1.75	$C^{\gamma}H_3$ 0.79; $C^{\gamma}H_2$ 1.36, 1.08; $C^{\delta}H_3$ 0.79
V27	8.13	4.52	2.00	$C^{\gamma}H_3$ 0.97, 0.90
P28		4.40	2.27, 1.97	$C^{\gamma}H$ 2.14, 2.00; $C^{\delta}H$ 3.82, 3.70
C29	8.43	4.92	3.21, 3.03	
T30	8.03	4.46	4.46	$C^{\gamma}H_3$ 1.28
R31	8.33	4.29	1.97, 1.81	$C^{\gamma}H$ 1.73, 1.69; $C^{\delta}H$ 3.25; $N^{\epsilon}H$ 7.16
E32	7.80	4.37	2.04, 1.97	$C^{\gamma}H$ 2.44
C33	8.56	4.56	3.25, 2.99	
K34	8.45	4.59	2.05, 1.20	$C^{\gamma}H$ 1.51, 1.49; $C^{\delta}H$ 1.71; $C^{\epsilon}H$ 3.03; $N^{\zeta}H_3^+$ 7.40
P35		4.62	2.54, 1.74	$C^{\gamma}H$ 2.18, 2.04; $C^{\delta}H$ 3.83, 3.57
P36		4.22	1.96, 1.64	$C^{\gamma}H$ 2.27, 2.11; $C^{\delta}H$ 3.93, 3.66
R37	7.26	4.27	2.00, 1.69	$C^{\gamma}H$ 1.48, 1.30; $C^{\delta}H$ 3.21; $N^{\epsilon}H$ 6.99
C38	8.68	4.85	3.19, 2.76	
E39	8.61	4.85	2.02, 2.02	$C^{\gamma}H$ 2.35, 2.27
C40	9.23	5.03	2.84, 2.56	
I41	7.86	4.08	1.87	$C^{\gamma}H_3$ 1.06; $C^{\gamma}H_2$ 1.44, 0.88; $C^{\delta}H_3$ 0.77
A42	8.37	4.57	1.67	
S43	8.90	4.24	4.06, 3.98	
A44	7.13	4.74	1.65	
G45	8.01	4.12, 3.75		
F45	7.49	5.09	2.94, 2.63	$C^{\delta}H$ 7.04; $C^{\epsilon}H$ 7.32; $C^{\zeta}H$ 7.25
V47	9.43	4.41	1.63	$C^{\gamma}H_3$ 0.91, 0.71
R48	8.54	4.86	1.94, 1.82	$C^{\gamma}H$ 1.82, 1.71; $C^{\delta}H$ 3.07, 2.94; $N^{\epsilon}H$ 7.86
D49	8.98	4.92	3.82, 2.90	
A50	9.03	4.16	1.50	
Q51	7.60	4.34	2.25, 2.09	$C^{\gamma}H$ 2.55, 2.41
G52	8.17	4.34, 3.64		
N53	7.85	4.95	2.74, 2.64	$N^{\delta}H_2$ 7.76, 7.13
C54	9.20	5.44	2.90, 2.65	
I55	9.63	4.99	1.97	$C^{\gamma}H_3$ 0.89; $C^{\gamma}H_2$ 1.40, 0.97; $C^{\delta}H_3$ 0.88
K56	8.73	3.98	1.69, 1.63	$C^{\gamma}H$ 1.34, 1.00; $C^{\delta}H$ 1.35; $C^{\epsilon}H$ 2.85; $N^{\zeta}H_3^+$ 7.40
F57	8.49	4.42	3.14, 2.96	$C^{\delta}H$ 7.26; $C^{\epsilon}H$ 7.44; $C^{\zeta}H$ 7.38
E58	8.38	4.04	2.04, 1.97	$C^{\gamma}H$ 2.44, 2.34
D59	8.20	4.85	3.20, 2.88	
C60	7.54	4.47	3.06, 2.88	
P61		4.47	2.37, 1.99	$C^{\gamma}H$ 2.15; $C^{\gamma}H$ 3.93, 3.82
K62	8.45	4.36	1.95, 1.79	$C^{\gamma}H$ 1.49; $C^{\gamma}H$ 1.62; $C^{\epsilon}H$ 3.03; $N^{\zeta}H_3^+$ 7.47

[a] Chemical shifts are reported with respect to 4,4-dimethyl-4-silapentane-1-sulfonate.

as well as regions with small $^3J_{HN\alpha}$ coupling constants. All residues exhibiting coupling constants smaller than 6 Hz are in isolated positions throughout the sequence, lending strong support to the notion that no α-helix is present in ATI. Turns between the β-strands can be inferred from medium to strong NH(i)–NH(i+1) and isolated $C^{\alpha}H(i)$–NH(i+2), NH(i)–NH(i+2), and $C^{\alpha}H(i)$–NH(i+3) connections. In the present case we were unable to identify any classical type turns on the basis of a qualitative interpretation of the data. The existence of several irregular turns or loops, however, can be clearly inferred. These are centered around positions 14, 23, 31, 44, 51, and 59. The turn from residues 22–25 is characterized by two $C^{\alpha}H(i)$–NH(i+2) NOEs from Thr-21 to Ala-23 and from Cys-22 to Gln-24, as well as a NH(i)–NH(i+3) NOE between Cys-22 and Lys-25, suggesting that this is a helical turn.

On the basis of a qualitative interpretation of the short-range NOE patterns in combination with the identification of strategic long-range NOEs between adjacent β-strands, some of which are indicated in the NOESY spectra shown in Figures

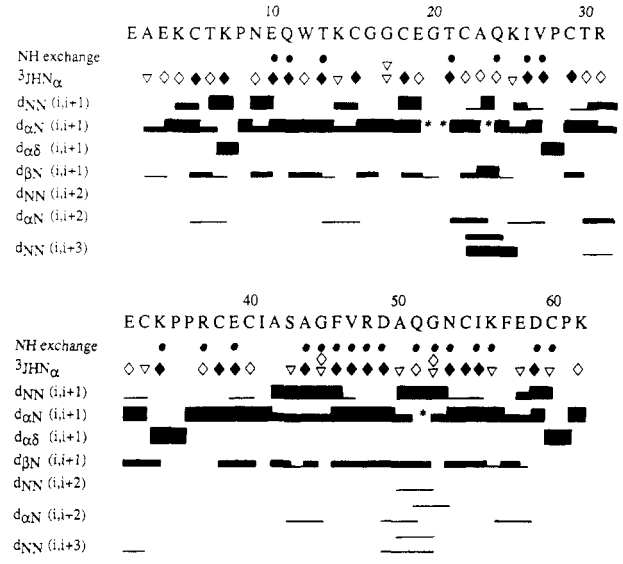

FIGURE 6: Summary of short-range NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons as well as the $C^\delta H$ protons of proline, the $^3J_{HN\alpha}$ coupling constant data, and the slowly exchanging amide protons. The relative intensities of the NOEs, classified into strong, medium, and weak, are indicated by the thickness of the lines. The asterisks indicate potential sequential connectivities that are obscured by cross-peak overlap. The slowly exchanging amide protons that are still present in a HOHAHA spectrum recorded 24 h after dissolving the protein in D_2O are indicated by filled-in circles (●). The symbols for the coupling constant data are as follows: (▽) $^3J_{HN\alpha} < 6$ Hz; (◊) 6 Hz $< {}^3J_{HN\alpha} < 8$; (♦) $^3J_{HN\alpha} > 8$ Hz.

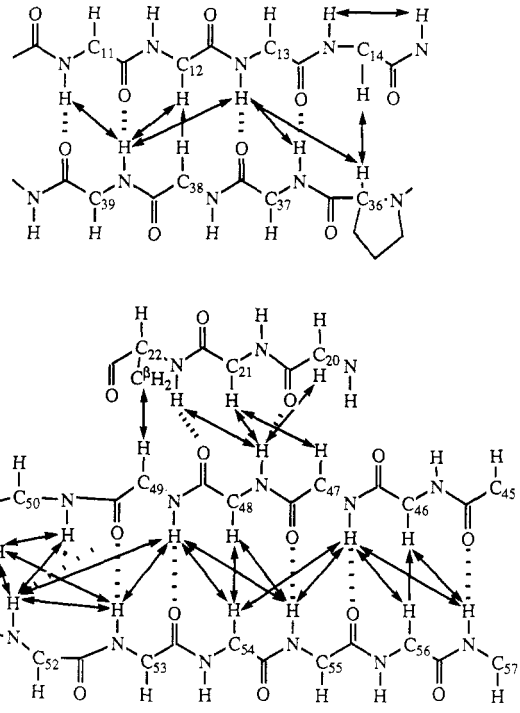

FIGURE 7: Regular secondary structure elements of ATI deduced from a qualitative interpretation of the NOE, $^3J_{HN\alpha}$ coupling constant, and amide exchange data. The arrows represent NOEs and the dashed lines hydrogen bonds.

3–5, and information about amide proton exchange, it is possible to identify two regions of regular secondary structure in ATI. The first comprises residues 11–14 (strand 1) and residues 37–39 (strand 2) which form an antiparallel β-sheet. The second β-sheet consists of two β-strands comprising residues 45–48 (strand 3) and 54–57 (strand 4). The third strand of this sheet is very short and comprises only residues 20–22 (strand 5), which are arranged parallel to strand 3. Strands 3 and 4 are arranged in an antiparallel fashion and connected by a turn from residues 49 to 53. The pattern of NOEs and slowly exchanging amide protons suggests that this is a 3:5 β-hairpin with a type I β-turn and a G1 β-bulge (Sibanda & Thornton, 1985) stabilized by a backbone hydrogen bond from Asn-53(NH) to the carbonyl oxygen atom of Asp-49 with the possibility of further stabilization by hydrogen bonding involving Gly-52(NH). A schematic illustration of these secondary structure elements is presented in Figure 7 together with the NOEs that led to their identification.

Additional information about slowly exchanging amide proton resonances which is also listed in Figure 6 supports the delineation of the secondary structure. Thus the antiparallel β-sheet formed by strands 1 and 2 involves four backbone hydrogen bonds associated with slowly exchanging amide protons of Gln-11, Thr-13, Arg-37, and Glu-39. In the antiparallel β-sheet formed by strands 3 and 4, as well as in the β-hairpin connecting these two strands, all the backbone hydrogen bonds, with the exception of that between the NH of Phe-57 and the carbonyl oxygen atom of Gly-45, are associated with slowly exchanging amide protons. The faster exchange rate of the NH of Phe-57 is probably due to the fact that the associated postulated hydrogen bond is situated at the end of the sheet and hence is more accessible to exchange with solvent. Interestingly, the NH of Gln-51 and Gly-52 within the β-hairpin are also slowly exchanging, although both clearly cannot be involved in a backbone hydrogen bond within the turn. This slow exchange behavior is probably due to hydrogen bonding between the amide protons of Gln-51 and Gly-52 and a side-chain acceptor group. The most likely candidate for this interaction is the side chain of Asp-49, since it has been shown that Asp side chains in a type I turn stabilize the turn structure by hydrogen bonding between the carbonyl oxygen atom and backbone amide protons. Such arrangements have been found in several protein crystal structures (Thornton et al., 1987). Finally, only one of the two hydrogen bonds between strands 3 and 5 is associated with a slowly exchanging amide proton, namely, that of Arg-48. The NH of Cys-22 is not slowly exchanging, again presumably due to the reduced stability of the associated hydrogen bond as the parallel β-sheet is extremely short and probably irregular. It should be noted, however, that the orientation of strand 5 with respect to strand 3 can be unambiguously determined on account of the NOEs between Arg-48(NH) and Thr-21($C^\alpha H$) and between Asp-49($C^\alpha H$) and Cys-22($C^\beta H$).

A comparison of the secondary structural features of known proteinase inhibitor structures with the secondary structure of ATI supports the notion that *Ascaris* inhibitors belong to a new family. Members of the pancreatic trypsin, pancreatic secretory trypsin, and potato I inhibitor families all contain at least one pronounced α-helical segment in addition to a two to four stranded β-sheet (Read & James, 1986), whereas no such helix is found in ATI. There is also no homology in terms of secondary structure with the recently determined structure of cucumber trypsin inhibitor, which lacks any regulary secondary structure elements (Bode et al., 1989; Heitz et al., 1989) and belongs to the potato II inhibitor family. It therefore seems an interesting and important goal to determine the complete three-dimensional structure of ATI, and efforts to that effect are currently ongoing in our laboratory.

REFERENCES

Babin, D. R., Peanasky, R. J., & Goos, S. M. (1984) *Arch. Biochem. Biophys. 232*, 143–161.
Baldwin, E., & Moyle, V. (1947) *J. Exp. Biol. 23*, 277.
Bax, A. (1989) *Methods Enzymol. 176*, 151–168.

Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511–6513.
Bode, W., Epp, O., Huber, R., Laskowski, M., & Ardelt, W. (1985) *Eur. J. Biochem. 147*, 387–395.
Bode, W., Greyling, H. J., Huber, R., Otlewski, J., & Wilusz, T. (1989) *FEBS Lett. 242*, 285–292.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, *275*, 288.
Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987a) *Biochemistry 26*, 8012–8023.
Clore, G. M., Gronenborn, A. M., Kjaer, M., & Poulsen, F. M. (1987b) *Protein Eng. 1*, 313–318.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.
Deisenhofer, J., & Steigemann, W. (1975) *Acta Crystallogr. B31*, 238–250.
Driscoll, P. C., Clore, G. M., Beress, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178–2187.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Folkers, P. J. M., Clore, G. M., Driscoll, P. C., Dodt, J., Köhler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601–2617.
Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.
Goodman, R. B., & Peanasky, R. J. (1982) *Anal. Biochem. 120*, 387–393.
Goodman, R. B., Martzen, M. R., & Peanasky, R. J. (1983) *Acta Biochim. Pol. 30*, 233–244.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Heitz, A., Chiche, L., Le-Nguyen, D., & Castro, B. (1989) *Biochemistry 28*, 2392–2398.
Hirono, S., Akagawa, H., Mitsui, Y., & Iitaka, Y. (1984) *J. Mol. Biol. 178*, 389–413.
Huber, R., Kukla, D., Rühlmann, A., Epp, O., & Formanek, H. (1970) *Naturwissenschaften 57*, 389–392.
Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., & Steigemann, W. (1974) *J. Mol. Biol. 89*, 73–101.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Laskowski, M. (1986) in *Proteinase Inhibitors* (Barret, A. J., & Salvesen, G., Eds.) pp 1–17, Elsevier Science Publishers, Amsterdam.
Laskowski, M., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–626.
Macura, C., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–282.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marion, D., & Bax, A. (1988a) *J. Magn. Reson. 79*, 352–356.
Marion, D., & Bax, A. (1988b) *J. Magn. Reson. 80*, 528–533.
McPhalen, C. A., Svendsen, I., Jonassen, I., & James, M. N. G. (1985a) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7242–7246.
McPhalen, C. A., Schnebli, H. P., & James, M. N. G. (1985b) *FEBS Lett. 188*, 55–58.
Mitsui, Y., Satow, Y., Watanabe, Y., & Iitaka, Y. (1979) *J. Mol. Biol. 131*, 697–724.
Mueller, L. (1987) *J. Magn. Reson. 72*, 191–196.
Papamakos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., & Laskowski, M. (1982) *J. Mol. Biol. 158*, 515–537.
Peanasky, R. J., Bentz, Y., Paulson, B., Graham, D. L., & Babin, D. R. (1987a) *Arch. Biochem. Biophys. 232*, 127–134.
Peanasky, R. J., Bentz, Y., Homandberg, G. A., Minor, S. T., & Rabin, D. R. (1987b) *Arch. Biochem. Biophys. 232*, 135–142.
Peanasky, R. J., Martzen, M. R., Homandberg, G. A., Cash, J. M., Babin, D. R., & Litweiler, B. (1987c) in *Paradigms for Eradicating Helminthic Parasites* (Macinnis, A. J., Ed.) pp 349–366, Alan R. Liss, New York.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Read, R. J., & James, M. N. G. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G. S., Eds.) pp 310–336, Elsevier Science Publishers, Amsterdam.
Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420–4433.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Rühlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417–436.
Satow, Y., Mitsui, Y., & Iitaka, Y. (1978) *J. Biochem. 84*, 897–906.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.
Sibanda, B. L., & Thornton, J. M. (1985) *Nature 316*, 170–174.
Sweet, R. M., Wright, H. T., Janin, J., Clothia, C. H., & Blow, D. M. (1974) *Biochemistry 13*, 4212–4228.
Thornton, J. M., Barlow, D. J., Edwards, M. S., Sibanda, B. L., Singh, J., & Wilmot, C. M. (1987) in *Impact of Knowledge of Amino Acid Sequence Data on Biophysical Research* (Tsugita, A., Ed.) pp 1–22, Science University of Tokyo, Noda, Japan.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Weber, E., Papamakos, E., Bode, W., Huber, R., Kato, I., & Laskowski, M. (1981) *J. Mol. Biol. 149*, 109–123.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.