NMR Analysis and Sequence of Toxin II from the Sea Anemone *Radianthus paumotensis*†

David E. Wemmer,*,‡ N. Vasant Kumar,§ Robert M. Metrione,∥,⊥ Michel Lazdunski,# Gary Drobny,○ and Neville R. Kallenbach§

Department of Chemistry, University of California, and Chemical Biodynamics Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, Departments of Biochemistry and Chemistry, University of Washington, Seattle, Washington 98195, and Centre de Biochimie du CNRS, Faculté des Sciences, Université de Nice, Parc Valrose, 06034 Nice Cedex, France

Received April 30, 1986; Revised Manuscript Received August 19, 1986

ABSTRACT: Toxin II from *Radianthus paumotensis* (Rp_{II}) has been investigated by high-resolution NMR and chemical sequencing methods. Resonance assignments have been obtained for this protein by the sequential approach. NMR assignments could not be made consistent with the previously reported primary sequence for this protein, and chemical methods have been used to determine a sequence with which the NMR data are consistent. Analysis of the 2D NOE spectra shows that the protein secondary structure is comprised of two sequences of β-sheet, probably joined into a distorted continuous sheet, connected by turns and extended loops, without any regular α-helical segments. The residues previously implicated in activity in this class of proteins, D8 and R13, occur in a loop region.

Recently, four new sea anemone toxins, isolated from *Radianthus paumotensis* (Rp), were reported (Schweitz et al., 1985) that act on Na^+ channels by slowing the channel inactivation process. These proteins represent a distinct class of sea anemone toxins, possessing both similarities to and differences from related toxins that have been isolated from *Anemonia sulcata* (As) and *Anthopleura xanthogrammica* (Ax). The action of these toxins is related to that of toxin II from the scorpion *Androctonus australis* (Aa), in that they compete for the same binding sites. However, the *R. paumotensis* toxins form a distinct immunological class, antibodies to Rp_{III} recognizing all of the other Rp toxins but not the As or Ax proteins, though there are significant sequence homologies among these. We have undertaken structural studies of these toxins using high-resolution ^1H NMR[1] methods to try to clarify the relationship between the structures of these toxins, their modes of action, and their immunological reactivity. In this work, we describe the NMR assignments, corrected sequence, and basic structural features of the Rp_{II} toxin. Structural characterization of a related family of molecules such as this will allow us to better understand the relationship of the primary sequence and structure, as well as the conformational features that are important for activity. Comparison of these toxins with other protein toxins that compete for the same binding sites, such as those from scorpions, will be particularly valuable in defining structural rather than sequence features that are necessary for activity. NMR studies of a related toxin, As_I, have been reported (Gooley et al., 1984; Gooley & Norton, 1986a), and recently (Gooley & Norton, 1986b) a sufficient number of assignments have been obtained to evaluate some aspects of the secondary structure in this protein.

The development of two-dimensional NMR (2D NMR) methods for systematic assignment of ^1H resonances (Wüthrich et al., 1982) has made NMR a powerful approach for the determination of the solution structure of proteins and nucleic acids (Wemmer & Reid, 1985). We feel that this approach is particularly well suited for analysis of a family of related proteins, such as the sea anemone toxins. Since it is very probable that the structures of these toxins are related, it is likely that many of the patterns of sequential connectivities will be repeated, simplifying analysis of the spectra. Initial analysis of the secondary structure can be carried out by using backbone NOEs, as discussed previously (Wüthrich et al., 1984; Wemmer & Kallenbach, 1983). In addition, once the structure of one of the toxins has been determined, it will likely provide a good starting point for successive structural refinements, whether by use of distance geometry (Havel & Wüthrich, 1985; Williamson et al., 1985) or molecular mechanics (Kaptein et al., 1985) approach.

Materials and Methods

NMR. The toxin Rp_{II} was isolated from the sea anemone *Radianthus paumotensis* as described previously (Schweitz et al., 1985). NMR studies were performed on a single sample of 14 mg of the purified toxin dissolved in ca. 0.4 mL of water, at pH 4.6. Studies were carried out in both H_2O and D_2O solutions, changing between them by lyophilization in a microcentrifuge tube and dissolving in pure H_2O or D_2O as

†This work was supported by NIH Grants GM-331861 (N.R.K.) and GM 15731 (K. Walsh), Centre National de la Recherche Scientifique and the Ministère de l'Industrie et de la Recherche (Grant 83.C.0918) (M.L.), and instrumentation grants from the Murdock Foundation, NIH GM 2874-01S1, and NSF PCM 80-18053 (University of Washington) with NSF CHE 82-08994 and NIH RR02428A-01 (University of California).

*Author to whom correspondence should be addressed.
‡University of California.
§University of Pennsylvania.
∥Department of Biochemistry, University of Washington.
⊥Permanent address: Department of Biochemistry, Thomas Jefferson University, Philadelphia, PA 19107.
#Université de Nice.
○Department of Chemistry, University of Washington.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TSP, sodium (trimethylsilyl)propionate; ppm, parts per million; 2D, two dimensional; COSY, 2D correlated spectroscopy; NOESY, 2D nuclear Overhauser effect spectroscopy; RELAY, 2D relayed coherence transfer spectroscopy.

0006-2960/86/0425-6842$01.50/0

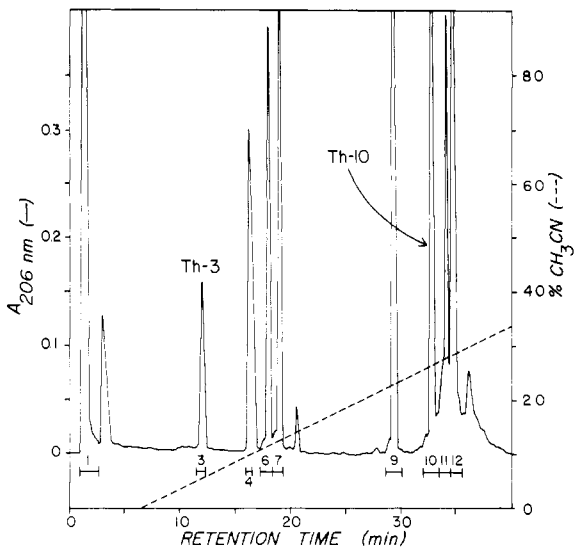

FIGURE 1: HPLC chromatogram of the thermolysin derived peptides from the Rp_{II} toxin. The peptides Th-3 and Th-10 were collected and sequenced, along with the intact protein, to give the full protein sequence shown in Figure 2.

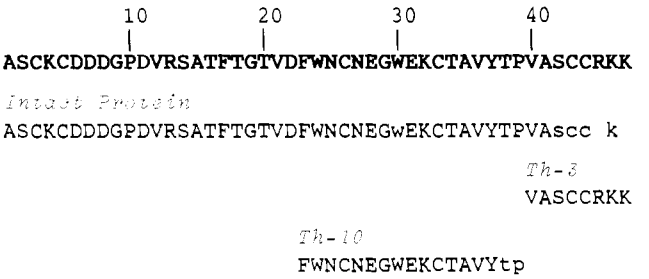

FIGURE 2: Sequence of the Rp_{II} toxin determined with an automated sequencer. The peptides Th-3 and Th-10 from thermolysin digestion were selected by amino acid composition to complete the sequence determined from the intact protein. Residues shown in lower case letters had relatively weak amino acid peaks compared to background, and were verified by analysis of Th-3 and Th-10.

needed. ^1H NMR spectra were collected on Bruker WM-500 and AM-500 spectrometers, equipped with Aspect computers. To obtain sequential assignments, standard COSY and NOESY spectra (Wider et al., 1984), plus RELAY and double RELAY (Eich et al., 1982; Bax & Drobny, 1985) two-dimensional spectra, were collected at temperatures between 20 and 40 °C. It was necessary to vary temperature to shift the solvent resonance relative to the protein in order to observe α-proton resonances close to the water frequency. In addition, small changes in amide and α-proton chemical shifts with temperature helped by removing a number of degeneracies, clarifying spin system connectivities. For all spectra the solvent signal was suppressed by irradiation at the solvent frequency at all times except during data acquisition. A decoupler power of 40–45 L (Bruker units) was used in D_2O, while 25–30 L was required for spectra in H_2O. NOESY spectra were collected with mixing times of 200–300 ms; no sign of spin diffusion was observed. For the RELAY experiments, total delays of 25–40 ms during the mixing period were used, the shorter being appropriate for observation of α to methyl connectivities while the latter is better for general amide to β-proton connectivities. In general, approximately 300 t_1 values were sampled, a collection of 1024 complex data points for each being used. Zero filling was applied to give a final spectrum of 1024 × 1024 real data points, corresponding to a digital resolution of 5 Hz/pt in D_2O (spectral width 5000 Hz), and 6 Hz/pt in H_2O (spectral width 6024 Hz). Data processing was carried out with the FTNMR program (D. Hare, unpublished) on a VAX 11/780. A skewed sine bell apodization (Hare et al., 1985) was used in both dimensions, typically with phase shift of 0° and skew of 0.7 for magnitude spectra (most COSY spectra) and a phase shift of 20° and skew of 0.5 for phase-sensitive spectra (all NOESY spectra). Phase-sensitive spectra were collected either with separate real and imaginary parts (States et al., 1982) or with TPPI (Bodenhausen et al., 1984) with equal success.

Thermolysin Digestion. Protein was prepared for sequencing by dissolving 500 μg of protein in 0.4 mL of 6 M guanidine, and 4.8 mg of dithiothreitol was added. This mixture was incubated at 37 °C for 1 h, after which 13.4 mg of iodoacetic acid dissolved in 0.1 mL of 0.5 M NaOH was added. After 30 min in the dark, the mixture was centrifuged to remove salts and stored frozen. The reduced and carboxymethylated protein, after acidification with formic acid at 9%, was separated from guanidine by reversed-phase HPLC on an Ultrapore-RPSC column (4.6 × 75 mm, Altex), with a gradient of acetonitrile in 0.1% trifluoroacetic acid. There was a single protein peak that emerged at 23% acetonitrile. This material was dried on a Speed-Vac concentrator (Savant), and the residue (15.5 nmol) was dissolved in 180 μL of 0.1 M NH_4HCO_3, pH 7.8. Thermolysin in the same buffer was added to give a protein to enzyme ratio of 100:1 (w/w). This solution was incubated for 1 h and 45 min at 37 °C and then lyophilized.

Separation of Peptides. The sample was dissolved in 100 μL of 10% acetic acid and applied to a Synchropak RP-P column, 4.1 × 250 mm (SynChrom, Inc., Linden, IN), in 0.1% trifluoroacetic acid (solution A). Solution B was acetonitrile containing 0.08% trifluoroacetic acid. A linear gradient of solution B into solution A was begun after 5 min, reaching 40% B after 40 min at a flow rate of 2 mL/min. The effluent was monitored by absorption at 206 nm. The apparatus was a Varian Model 5000 with a Varichrom detector. The effluent fractions containing 206 nm absorbing material were collected (Figure 1), dried on a Speed-Vac, and redissolved in 250 μL of 10% acetic acid. An aliquot (50 μL) of each fraction was removed for amino acid analysis.

Amino Acid Analysis. Amino acid compositions were determined with the Pico-tag system (Waters), after hydrolysis in the dry state over 6 N HCl containing 1% phenol in a nitrogen-purged vacuum. The samples were derivatized in phenyl isothiocyanate (PITC) solution (methanol–water–triethylamine–PITC, 7:1:1:1) for 20 min. Derivatized amino acids were separated and analyzed on a NovaPak C_{18} column (Waters), with a linear gradient from 0 to 46% solvent B in 10 min (solvent A = 0.135 M sodium acetate, pH 6.4, containing 0.55 mL of triethylamine per liter; solvent B = 60% acetonitrile in water).

Sequence Determination. The intact protein, peptide Th-3, and peptide Th-10 (Figure 2) were subjected to automated sequence analysis on a Beckman Model 890C sequencer in the presence of polybrene (Pierce) by using the method of Edman and Begg (1967) with the 0.16 M Quadrol program of Brauer et al. (1975). Phenylthiohydantoin derivatives of amino acids were identified by reversed-phase HPLC in two complementary systems (Ericsson et al., 1977; Glajch et al., 1985).

RESULTS

The Rp_{II} toxin has a well-resolved NMR spectrum, with line widths generally consistent with the protein existing as a monomer of molecular mass about 5 kDa. Assignment of

Table I: Amino Acid Composition[a] of Major Peptides Isolated after Cleavage with Thermolysin

	Th-3	Th-4	Th-6	Th-7	Th-9	Th-10	whole protein
Asp/Asn		0.95 (1)	3.86 (4)	4.01 (4)	2.40 (2)	2.25 (2)	5.4 (7)
Glu/Gln					2.45 (2)	2.12 (2)	1.9 (2)
Ser	0.85 (1)		1.68 (2)	1.73 (2)			2.3 (3)
Gly		0.80 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.05 (1)	2.8 (3)
Arg	1.00 (1)		0.89 (1)	0.99 (1)			1.9 (2)
Thr		1.70 (2)		0.83 (1)	0.97 (1)	1.89 (2)	4.4 (5)
Ala	0.91 (1)		0.95 (1)	1.69 (2)	1.00 (1)	1.00 (1)	4.0 (4)
Pro			0.97 (1)	0.87 (1)		0.92 (1)	1.8 (2)
Tyr						0.83 (1)	1.0 (1)
Val	0.89 (1)	1.00 (1)	0.92 (1)	0.82 (1)		0.96 (1)	3.8 (4)
Phe		0.71 (1)			0.77 (1)	0.91 (1)	1.9 (2)
Lys	2.44 (2)		0.59 (1)	0.55 (1)	0.57 (1)	0.60 (1)	5.8 (4–5)
Cys[b]	1.95 (2)		1.94 (2)	1.71 (2)	2.36 (2)	1.79 (2)	5.5 (6)
Trp					ND (2)	ND (2)	(2)
peptide[c]	40–47	17–22	1–14	1–16	23–35	23–39	(47–48)

[a] Residues per peptide by amino acid analysis. Values in parentheses are those in the corresponding sequence. Tryptophan was not determined. Fractions 1, 11, and 12 (Figure 1) were impure and are not included. [b] Determined as *S*-(carboxymethyl)cysteine. [c] The position of the peptide in the sequence of Rp_{II}, as indicated by amino acid composition.

the ^1H NMR spectrum was undertaken by the sequential approach of Wüthrich and co-workers (1982, 1983), with combinations of two-dimensional spectra (COSY, NOESY, and RELAY) taken both in H_2O and D_2O solution. From these spectra it was possible to identify the spin systems of several residues, including the four valines, the five threonines, the four alanines, the three glycines, the two glutamates, the tyrosine, one of the phenylalanines, and one of the lysines. These were used as starting points and markers during the assignment process. The numbers of each type of residue identified agreed precisely with the composition and sequence that have been published for this protein (Schweitz et al., 1985). The assignments in the N-terminal half of the protein proceeded normally, with only one significant gap remaining, at aspartate-6. However, in the C-terminal half of the protein many connectivities were found that could not be reconciled with the published primary sequence. In particular, the sequence XEGWEKXTAVYT (X corresponds to any AMX_2 spin system, e.g., C, S, D, etc.) was found from the sequential connectivities, while only the end of this; AVYT, occurred in the published sequence. After the failure of several attempts to develop alternate explanations for the observed connectivities, we undertook the direct chemical sequencing of a sample of the protein that was used for the NMR studies.

Automated sequence analysis on the intact carboxymethylated protein yielded a single amino acid at every position except 45 and 47. Sequence analysis of peptide Th-3 (Figure 1) showed that it corresponded to positions 40–47 and completed the structure of the protein. The sequence of peptide Th-10 (Figure 1) was shown to correspond to residues 23–39 and confirmed the sequence in this part of the intact molecule. The stepwise yield for the first 35 residues of the intact protein was 93.5%.

The positions of the other major peptides generated by the action of thermolysin on RP_{II} are indicated in Table I. This correlation is based on amino acid composition, although pooled fractions 3 and 10 were also sequenced. These peptides account for the entire sequence of the protein. The reported sequence of the protein corresponds to the amino acid composition that we have obtained, accounting for all residues. The sequence obtained by NMR is consistent with the primary sequence that we obtained by chemical sequencing for this protein, except for the C-terminal residue, but is significantly different from that reported previously in the C-terminal half of the molecule.

Using the corrected sequence, we were able to complete

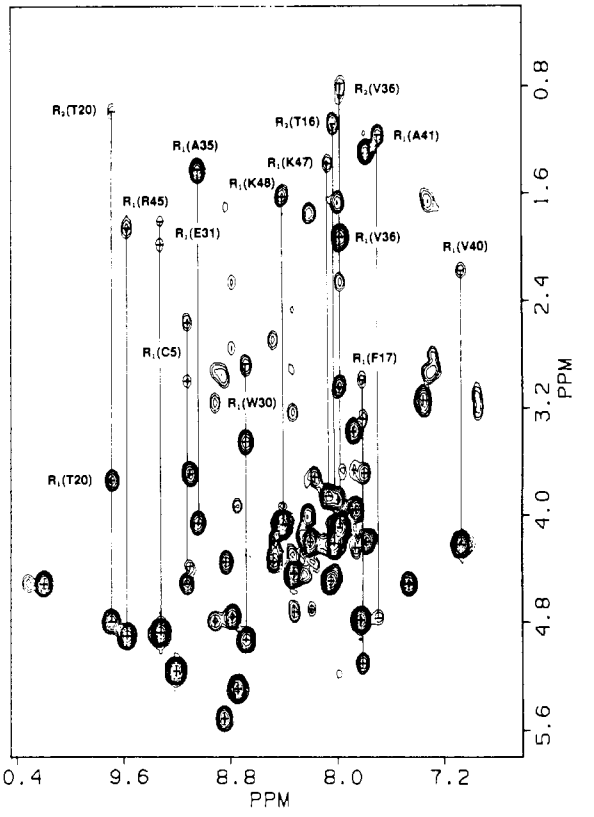

FIGURE 3: Amide to α, β, γ, and methyl region of the double RELAY spectrum taken in 90% H_2O solution. The spectrum was collected off-resonance; hence, the digital resolution was limited, giving rise to apparently broader peaks than in the same region of the COSY spectrum (Figure 7). The direct connection of amide to β and methyl peaks was important for correct identification of spin systems, particularly in regions of the sequence that were incorrect in the initial determination. Comparison with the COSY spectrum shows that most peaks above 3.2 ppm are RELAY or double RELAY cross peaks. Some of these are labeled with R_n(residue), where n indicates single or double RELAY peak and residue identifies the position of the spin system in the primary sequence of the molecule as derived from sequential assignments. The direct COSY cross peaks are labeled with a (+), and vertical lines are drawn from COSY peaks to the associated RELAY peaks.

NMR assignments of nearly all residues in the protein. The assignment process, particularly in initial stages where assignments were not in agreement with the primary sequence, was aided by use of RELAY and double RELAY spectra, in

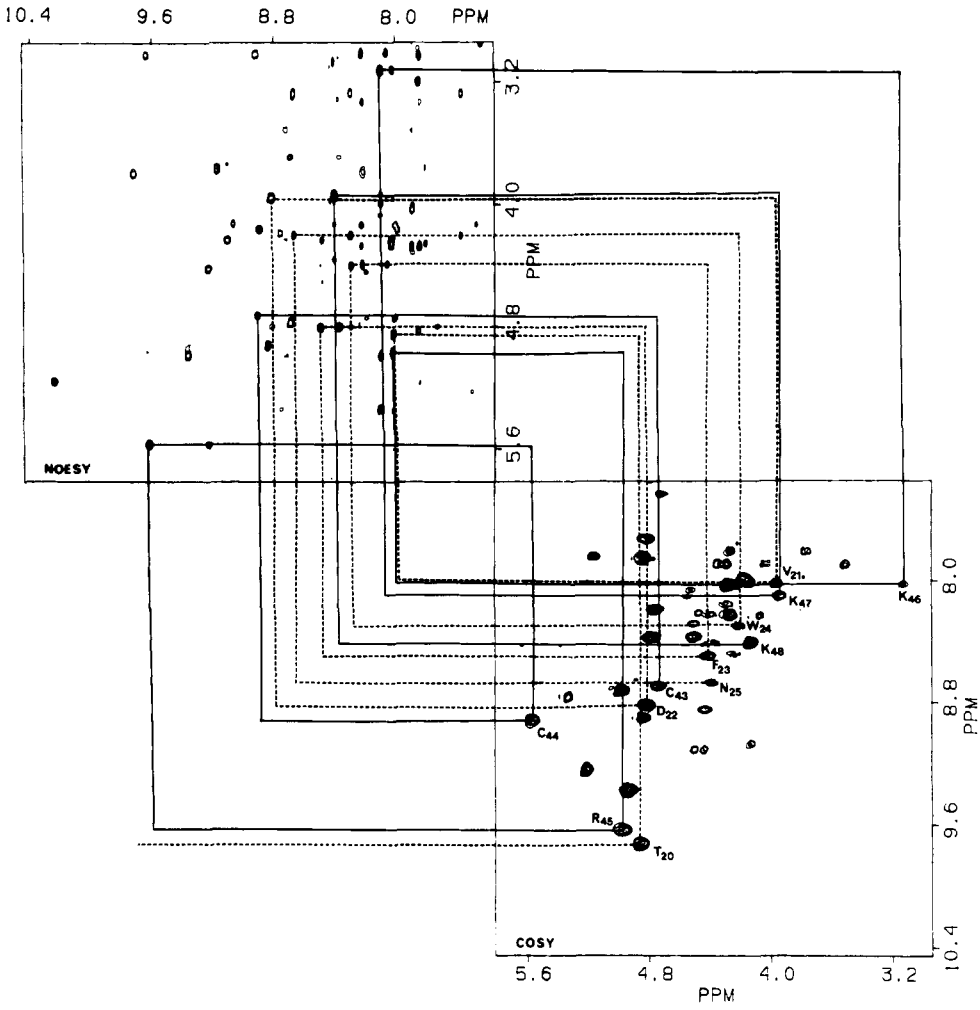

FIGURE 4: Regions of the COSY and NOESY spectra of Rp_{II}, showing sequential connectivities in two of the extended segments of the molecule. These correspond to two opposite strands of β-sheet, whose connections were established through cross-strand α–α and amide–α NOEs.

both H_2O and D_2O solutions. As discussed by Weber et al. (1985), RELAY cross peaks between α-H and γ-CH_3s are particularly useful for the unambiguous identification of the valine and threonine spin systems. In our case, the absence of leucine and isoleucine simplified analysis of the methyl region; however, the RELAY spectrum helped in distinguishing the threonine and alanine spin systems. In addition, it provided verification of the connectivities to the γ-protons in the two glutamate side chains. The single and double RELAY spectra taken in H_2O solution (Figure 3) allowed us to clearly identify the amide resonances of most of the alanine, threonine, and valine residues and, additionally, connect many other β- and γ-protons with the appropriate amide resonance, avoiding the crowded α-proton region. In the RELAY spectrum taken in D_2O, an unusual feature was observed in the aromatic region. A RELAY cross peak appeared between two doublet aromatic peaks corresponding to the protons at the ϵ_3 and ζ_2 positions of a tryptophan ring. This is unusual because these two protons do not have a common coupling partner, as is the general basis for RELAY peak. However, examination of the direct COSY cross-peak positions, together with that of the RELAY peak, showed that the chemical shifts of the ζ_3- and η-proton resonances were almost identical. Thus, these two protons form a strongly coupled pair, which act as one "super-spin" in the RELAY experiment, acting as the common partner of the ϵ_3 and ζ_2 protons. This "extra" cross peak was helpful in understanding the absence of one of the expected COSY cross peaks, in this tryptophan system, and thus completed identification of this spin system. This phenomenon is related to the presence of virtual coupling, which has been described in a tryptophan ring spin system from another anemone toxin (Gooley et al., 1984). A similar RELAY has been observed through a strongly coupled α–β proton pair of threonine in the S-peptide of ribonuclease A (D. E. Wemmer and P. S. Kim, unpublished results).

An expansion of the fingerprint region for this protein, from the COSY and NOESY spectra taken in H_2O solution, is shown in Figure 4. The connectivities of two segments of the protein are indicated in the figure, drawn in the usual spiral pattern. Many other $d_{\alpha N}$ connectivities were seen in sequences throughout the protein, as shown in Figure 6, but are omitted from Figure 4 for clarity [connectivities between neighboring residues are indicated with the notation of Stassinopoulou (1984), where the subscripts α, N, and β indicate the protons on each residue involved in the observed NOE]. In addition to the interference of the residual solvent signal, there were several amide proton chemical shift degeneracies that presented problems for assignments. Some of the resonances involved were sufficiently temperature dependent that a change of 15 °C removed the degeneracy and clarified the correct connectivities. In Figure 5 the amide to amide cross-peak region of the NOESY spectrum is shown, again together with several connectivities. Since there were no particular irregularities or difficulties with the assignments, once the correct sequence was determined, they will not be discussed in detail; the connectivities used in sequential assignments are summarized schematically in Figure 6, and chemical shifts for the assigned proton resonances are given in Table II, with assignments for

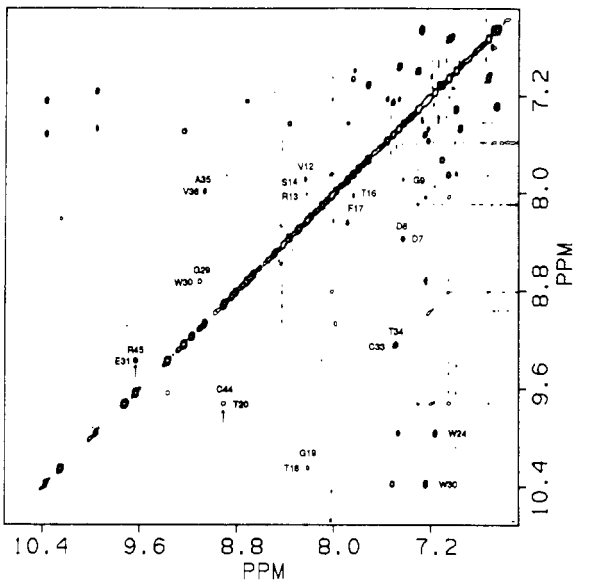

FIGURE 5: Part of the NOESY spectrum, in 90% H_2O solution, showing amide–amide connectivities. The cross peaks are labeled with the two residues involved. Although many of these are weak, all were seen in at least three different NOESY spectra taken under different conditions. The cross-strand NOEs observed are indicated with a vertical arrow.

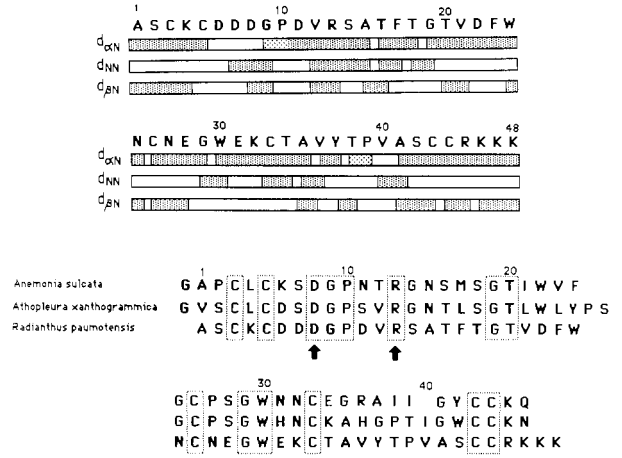

FIGURE 6: Summary of the sequential resonance assignments for the Rp_{II} toxin. These connectivities are $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ between residues $i-1$ and i, with α, N, and β designating the protons on these residues. The bar between two residues is shaded if that connectivity was observed between those two residues. The relationship of the Rp_{II} sequence to other anemone toxins is also shown. The positions of residues that have been previously identified as crucial for activity are shown with vertical arrows.

peaks in the fingerprint region presented in Figure 7. A break in the sequential connectivities occurs at D6, in addition to the two prolines, which have no amide protons. The only point at which there is any ambiguity between the chemical sequencing described here and NMR is in the C-terminal residue, the NMR connectivities suggesting that the sequence ends CCRKKK while chemical sequencing did not find the last lysine. Is is possible that due to the repeat of lysine at this point, and the hydrophobicity of the blocked lysine derivative used in sequencing, that the last of them was missed in the automated sequencing. It should be noted that the amino acid compositions of both the intact protein and the Th-3 are basically consistent with a terminus of three lysine residues. The presence of this lysine seems quite clear from the NMR, and we will pursue further chemical verification of its presence. The analysis of the structure and assignment of the spectrum are not affected by this uncertainty.

As can be seen from Figure 6, the predominant connectivity observed was of the $d_{\alpha N}$ type, with no contiguous sequence containing *only* d_{NN} and $d_{\beta N}$ connections. This indicates that the structure is comprised primarily of extended chains, which are connected together with a variety of turns, without any significant helical structure. In NOESY spectra taken in D_2O solution (Figure 8), several α-proton to α-proton cross peaks were observed, which are generally indicators of antiparallel β-sheet structure (Wüthrich et al., 1984). On the basis of the assignments of these cross peaks, together with several cross-strand amide to α-proton cross peaks, we have assigned two segments of β-sheet, one with residues 30–33 across from residues 43–46 and the other placing residues 2–5 across from 18–21. Two additional weak NOEs were observed, one between the α-H of 21 and the α-H of 43 and the other between the amides at 20 and 44. These suggest that the two β-sheet segments identified above may actually form a larger continuous sheet. Amide proton exchange experiments have verified that the protons at the centers of these β-sheet structures are protected against exchange relative to the majority of other amides in the protein, as expected for amides involved in hydrogen bonds. Residues that exchange slowly enough to observed in COSY spectra taken in D_2O include 5, 19, 20, 21, 22, 30, 31, 33, 44, and 45. The protection of residues 20, 22, and 44 argues in favor of a four-strand structure; however as discussed below, it must be quite highly twisted. The interstrand contacts and protected amides that we have observed correspond precisely to the structure observed for the analogous residues in the anthopleurin A and As_I toxins (Gooley & Norton, 1986b), except that the strand involving the C-terminal residues is extended by one amino acid.

We have assumed that the cystine pairing in the three disulfides of this protein is the same as that determined for the related toxin Ax_{II} (Wunderer, 1978). With this assumption, and the placement of the strands described, there must be a sharp turn in the backbone between residues 26 and 30, in order to accomodate the 26–44 disulfide. This is in fact supported by the NMR data, with a change in the connectivities from $d_{\alpha N}$ to d_{NN} at Gly-29. In addition, the cystines at residues 33 and 43 have correct relative positions and are on the correct side of the β-sheet to pair with cystine-5 and -3, respectively, as required in this pairing scheme. In order for these disulfides to form, the end of the outside strand of the four proposed to comprise a single sheet must curve tightly over the central strands, and there must be significant distortion of the sheet to accomplish this. From the weak NOEs between the α-protons of T20 and C43, and between the amides of V21 and C44, it appears that the central section of the sheet is most distorted. The β-sheet forms a fairly compact core for the protein, while two of the connecting segments, including residues 6–17 and 34–42, form loops. It is again clear from the NOE patterns observed that there are turns and extended segments in each of these regions, which include a proline near the end of each loop. These loops do not, however, appear to be fully extended or more flexible than the core. Tyrosine-37, near the end of the second loop, shows NOEs to a number of residues, including side chains on the core segments. It is additionally interesting to note that lysine-46 is relatively unusual in that the resonances for most protons on the side chain are shifted significantly upfield from their random coil positions and the two protons at the β- and γ-carbons are inequivalent in chemical shift. This suggests that the side chain of this lysine is relatively immobile and might occupy a site near an aromatic group, though we have not yet assigned direct NOEs to verify this.

Table II: Chemical Shifts of Assigned Resonances, pH 4.6, 35 °C

residue	amide	α-H	β-H	other
Ala-1	*b*	4.20	1.55	
Ser-2	8.83	5.36	4.05	
Cys-3	8.12	4.57	2.93, 2.69	
Lys-4	7.86	3.77	1.47, 1.33	γ, 0.24
Cys-5	9.16	4.58	3.08, 2.64	
Asp-6	*b*	*b*	*b*	
Asp-7	8.39	4.53	3.04, 2.56	
Asp-8	7.40	4.59	3.08, 2.64	
Gly-9	7.94	4.32, 4.03		
Pro-10		*b*	*b*	γ, 2.04; δ, 3.72, 3.61
Asp-11	8.30	4.78	2.89, 2.70	
Val-12	8.03	4.18	2.33	CH_3, 1.04, 1.02
Arg-13	8.26	4.29	1.83	
Ser-14	7.97	3.39	3.76, 3.66	
Ala-15	7.87	4.27	1.37	
Thr-16	8.07	4.30	4.27	CH_3, 1.15
Phe-17	7.85	5.17	3.37, 3.05	ring: δ 7.02; ε, 7.32; ζ, 7.26
Thr-18	10.25	4.59	4.21	CH_3, 1.21
Gly-19	8.26	4.47, 3.79		
Thr-20	7.76	4.86	3.78	CH_3, 1.03
Val-21	8.02	3.96	1.71	CH_3, 0.59, 0.55
Asp-22	8.83	4.86	2.84, 2.50	
Phe-23	8.47	4.38	2.75	ring: δ, 7.18; ε, 7.26; ζ, ca. 7.3
Trp-24	8.39	4.14	3.36	ring: ϵ_3, 7.63; ζ_3, 7.22; η, 7.24; ζ_2, 7.49; δ, 7.18; NH, 10.01
Asn-25	8.81	4.39	2.46, 1.86	
Cys-26	8.30	4.35	2.80	
Asn-27	8.09	4.56	2.65, 1.98	
Glu-28	8.58	4.26	2.07	γ, 2.45
Gly-29	9.19	4.49, 3.76		
Trp-30	8.79	5.00	3.54, 2.96	ring: ϵ_3, 7.25; ζ_3, 7.06; η, 7.26; ζ_2, 7.55; δ, 7.28; NH, 10.44
Glu-31	9.40	4.97	1.94, 2.08	γ, 2.26
Lys-32	8.95	4.42	1.80	γ, *b*
Cys-33	9.32	5.24	3.08	
Thr-34	7.53	4.59	4.52	CH_3, 1.30
Ala-35	9.11	4.14	1.51	
Val-36	8.00	4.18	2.01	CH_3, 0.90, 0.85
Tyr-37	8.97	4.86	3.22, 3.00	ring: δ, 7.06; ε, 6.74
Thr-38	7.87	4.85	4.50	CH_3, 0.68
Pro-39		*b*	*b*	γ, 2.04, 1.95; δ, 3.79, 3.72
Val-40	7.11	4.26	2.19	CH_3, 0.83, 0.74
Ala-41	7.69	4.80	1.17	
Ser-42	8.45	4.79	3.71, 3.30	
Cys-43	8.78	4.74	2.98, 3.89	
Cys-44	8.96	5.60	3.03	
Arg-45	9.65	4.99	1.80, 1.94	γ, *b*
Lys-46	8.05	3.12	1.13, 0.43	γ, 0.63, 0.35; δ, 1.24; ε, 2.85
Lys-47	8.14	3.96	3.77	γ, *b*
Lys-48	8.44	4.15	1.73	γ, 1.35

[a] Chemical shifts in ppm (parts per million) from internal TSP [sodium (trimethylsilyl)propionate]. [b] Resonances that could not be assigned in the spectrum.

A doubling of resonances has been observed previously in the spectra of toxins As_I and As_{II} (also called ATX-I and ATX-II, Gooley et al., 1984; Widmer et al., 1984), interpreted as arising from two different conformations of the molecules coexisting in solution. In Rp_{II} we also see some evidence for a second conformation, but the population of the second conformation is small (≤15%). In addition, the doubling is resolved for relatively few resonances in the spectrum, being clear for the amide–α cross peak of T18, and probably the β-methyl peak of T16, and a few other fingerprint region peaks that cannot clearly be assigned. This region of the molecule is the same as that in which doubling was identified in As_I (Gooley et al., 1984) and suggests that the doubling arises from a particular structural feature that is conserved in these sequences. We are now in the process of quantifying NOE intensities to generate more rigorous distance constraints and will describe a more detailed structure in a future publication.

Discussion

The sequence of Rp_{II} has previously been reported by Schweitz et al. (1985). However, the sequence reported here (Figure 2) differs substantially from that in the previous report. It should be noted again that our results with NMR and chemical sequence studies are consistent with each other (except the C-terminal residue) and both disagree with the previously reported sequence from residue 24 to residue 48. The cause of these differences is not evident at this time; however, it is unlikely that the sequences represent different proteins, as the proteins were prepared in the same laboratory and have identical composition and amino-terminal sequences. In the region from the N-terminus to residue 23, the sequence reported here is identical with that of Schweitz et al. (1985). From residue 24 to the C-terminus there is essentially no sequence identity, although there are small stretches of equivalent sequence. For example, our sequence for residues 35–39 is identical with the sequence previously reported for residues 30–34. There is also marked similarity near the C-terminus. In the sequence reported previously, it was surprising to find that the positions of two of the cysteines (corresponding to one of the disulfide pairs) were in different positions than in the other homologous anemone toxins. In our corrected sequence there is now complete conservation of

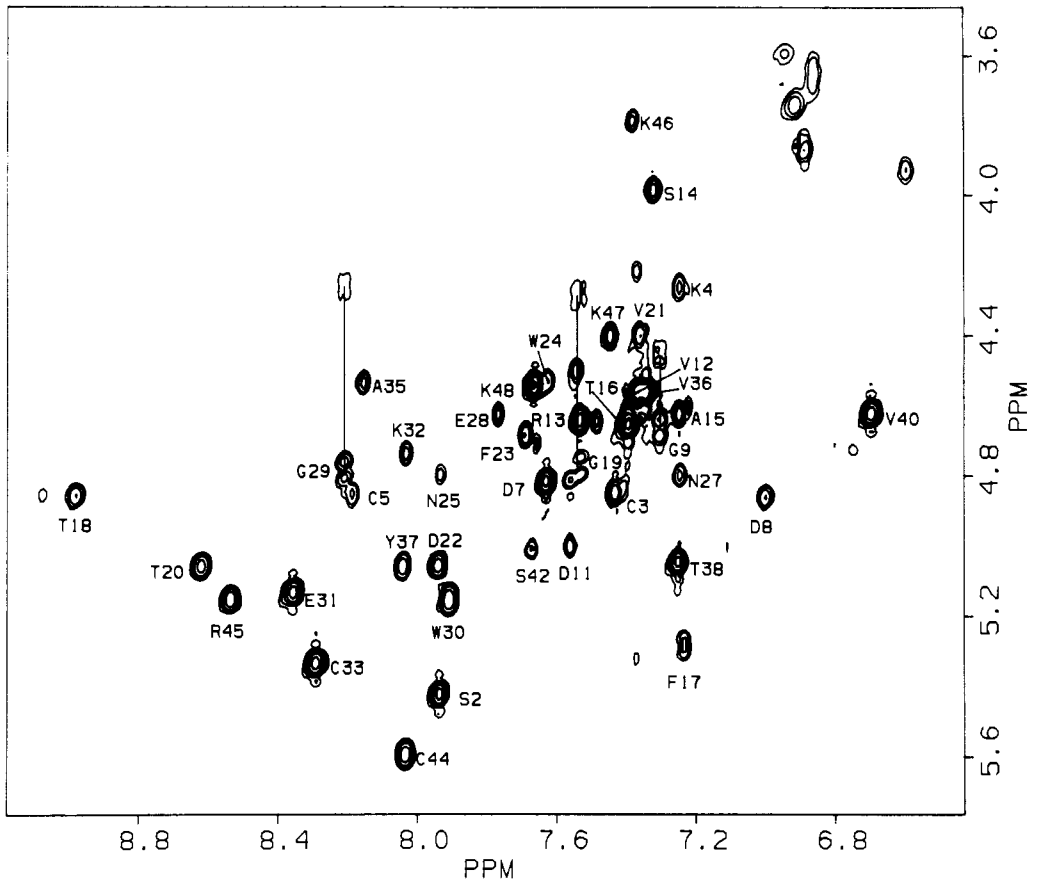

FIGURE 7: Fingerprint region of the Rp_{II} toxin, with the peak assignments derived from sequential connectivities shown next to the peaks. Due to irradiation of the solvent, some of the α-proton resonances were saturated, and hence, the corresponding amide–α cross peaks were not observed under these conditions. Using temperature to shift the water resonance, lines saturated in this spectrum could be seen.

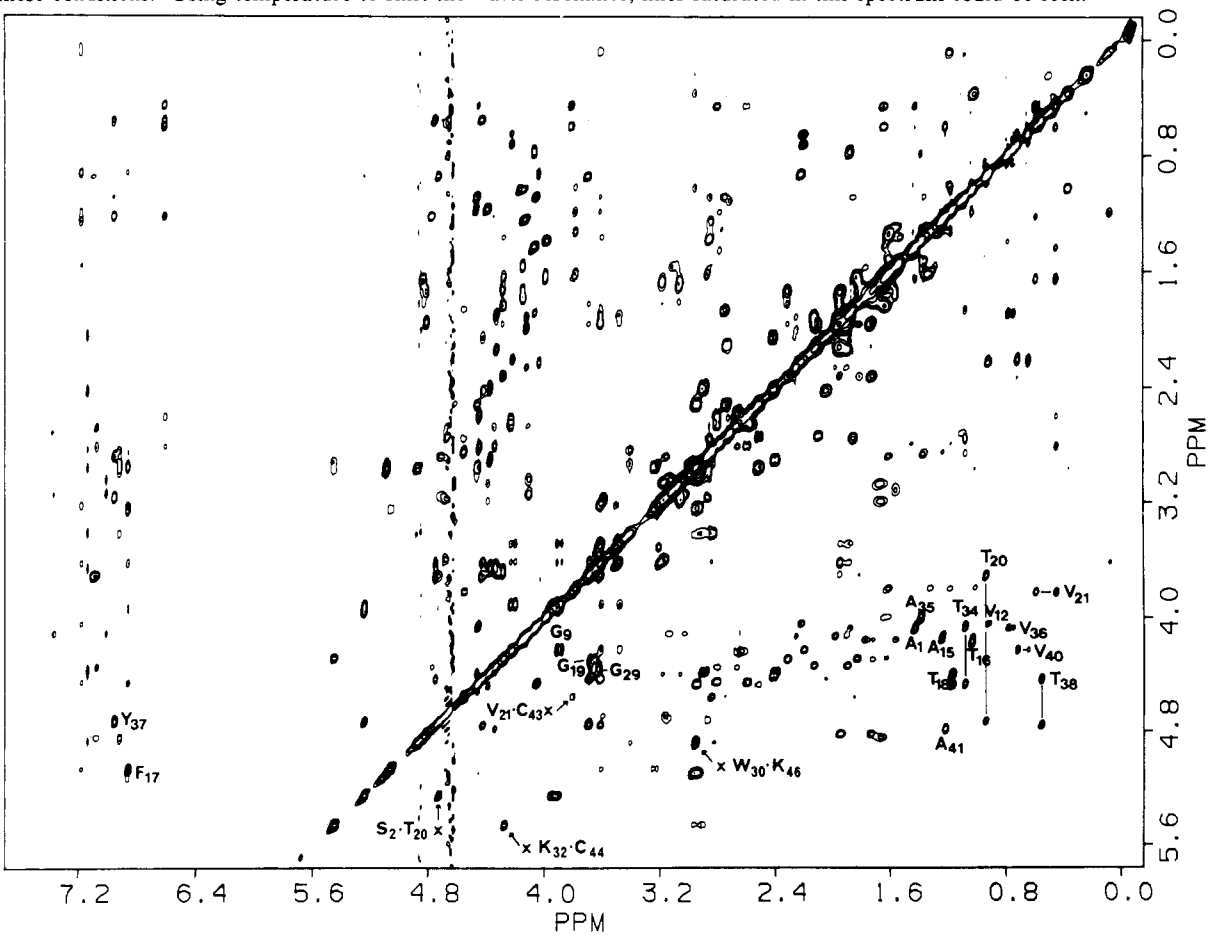

FIGURE 8: NOESY spectrum of Rp_{II} taken in D_2O solution. The cross peaks designated with an x indicate α–α NOEs between opposite strands of β-sheet and are labeled with the two residues involved. The methyl resonances of alanine, threonine, and valine are also labeled with the assignment derived from the sequential connectivities.

the cysteine positions, as well as for the GW dipeptide at position 29–30, Figure 2.

The structural characteristics that we have identified are in agreement with conclusions reached by other workers using circular dichroism, laser Raman, fluorescence spectroscopy, and ^{13}C NMR on other sea anemone toxins (Ishizake et al., 1979; Norton & Norton, 1979; Norton et al., 1982). All of these studies suggested that this class of proteins has numerous β-bends and some β-sheet regions, though there are also suggestions of regions of irregular secondary structure (Prescott et al., 1976), all generally consistent with the features found through our NMR studies. The most detailed description of secondary structure in As_I comes from the recent NMR work, similar to that described here, of Gooley and Norton (1986b). These workers also find evidence for β-sheet structure involving most of the same residues as we find in Rp_{II}. In our experiments, done at higher field, assignments are more complete, but the overall picture of the structure is the same.

The structure of this toxin has some features in common with other toxins whose structures are known, e.g., snake venom neurotoxins. These are comprised of a central core, which has several disulfide cross-links that are in proximity. From this core there extend several loops of somewhat irregular structure. It appears that the active sequences in these proteins often occur in these loop regions. In the anemone toxins, for example, it has been determined that conserved aspartate at position 6 or 8 and the arginine at position 13 (numbering taken from Rp_{II}) are necessary for activity (Barhanin et al., 1981) and are not in a region of regular secondary structure. From analysis of active peptides from α-cobratoxin, it also appears that the active sequence is in a loop extending from the core (Martin et al., 1983a), which again does not have regular secondary structure (Martin et al., 1983b). A better picture of the active regions in the anemone toxins may emerge as the structures of more of the related proteins are determined.

Acknowledgments

We particularly thank Dr. H. Schweitz, who prepared the toxin, and Prof. Kenneth Walsh (University of Washington) for providing laboratory space for R.M.M. during a sabbatical stay and for help with sequencing methodology. We also thank Drs. Rachel Klevit and Paul Weber for many helpful discussions.

References

Barhanin, J., Hughues, M., Schweitz, H., Vincent, J. P., & Lazdunski, M. (1981) *J. Biol. Chem. 256*, 5764.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306.

Brauer, A. W., Margolies, M. N., & Huber, E. (1975) *Biochemistry 14*, 3209.

Billiter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80.

Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731.

Ericsson, L. H., Wade, R. D., Gagaon, J., McDonald, R. M., Granberg, R. R., & Walsh, K. A. (1977) in *Solid Phase Methods in Protein Sequence Analysis* (Previero, A., & Coletti-Previero, M. A., Eds.) p 137, Elsevier/North-Holland, New York.

Glajch, J. L., Gluckman, J. C., Charikofsky, J. G., Minor, J. M., & Kirkland, J. J. (1985) *J. Chromatogr. 318*, 23.

Gooley, P. R., & Norton, R. S. (1986a) *Biopolymers 25*, 489.

Gooley, P. R., & Norton, R. S. (1986b) *Biochemistry 25*, 2349.

Gooley, P. R., Beress, L., & Norton, R. S. (1984) *Biochemistry 23*, 2144.

Hare, D. R., Ribeiro, N. S., Wemmer, D. E., & Reid, B. R. (1985) *Biochemistry 24*, 4300.

Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 281.

Ishizake, H., McKay, R. H., Norton, T. R., Yasunobu, K. T., Lee, J., & Tu, A. (1979) *J. Biol. Chem. 254*, 9651.

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179.

Martin, B. M., Chibber, B. A., & Maelicke, A. (1983a) *Toxicon 3*(Suppl), 1983.

Martin, B. M., Chibber, B. A., & Maelicke, A. (1983b) *J. Biol. Chem. 258*, 8714.

Norton, R. S., & Norton, T. R. (1979) *J. Biol. Chem. 254*, 10220.

Norton, R. S., Norton, T. R., Sleigh, R. W., & Bishop, D. G. (1982) *Arch. Biochem. Biophys. 213*, 87.

Prescott, B., Thomas, G. J., Beress, L., Wunderer, G., & Tu, A. T. (1976) *FEBS Lett. 64*, 144.

Schweitz, H., Bidard, J.-N., Frelin, C., Pauron, D., Vijverberg, H. P. M., Mahasneh, D. M., Lazdunski, M., Vilbois, F., & Tsugita, A. (1985) *Biochemistry 24*, 3554.

Stassinopoulou, C., Wagner, G., & Wüthrich, K. (1984) *Eur. J. Biochem. 145*, 423.

States, D. J., Haberkron, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.

Weber, P. L., Drobny, G., & Reid, B. R. (1985) *Biochemistry 24*, 4549.

Wemmer, D. E., & Kallenbach, N. R. (1983) *Biochemistry 22*, 1901.

Wemmer, D. E., & Reid, B. R. (1985) *Annu. Rev. Phys. Chem. 36*, 105.

Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1982) *J. Magn. Reson. 56*, 207.

Widmer, H., Wagner, G., & Wüthrich, K. (1984) *Proceedings of the 22nd Congress Ampere*, Zürich, p 498, Zürich Ampere Committee, Zürich.

Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295.

Wunderer, G. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 1193.

Wüthrich, K. (1983) *Biopolymers 22*, 131.

Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 319.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715.