Lawrence Berkeley National Laboratory

Recent Work

Title

Structures of Peptide Families by Nuclear Magnetic Resonance Spectroscopy and Distance Geometry

Permalink

<https://escholarship.org/uc/item/63g1496k>

Author

Pease, J.H.

Publication Date

1989-12-01

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

CHEMICAL BIODYNAMICS DIVISION

**Structures of Peptide Families by Nuclear Magnetic
Resonance Spectroscopy and Distance Geometry**

J.H. Pease
(Ph.D. Thesis)

December 1989

For Reference

Not to be taken from this room

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

**Structures of Peptide Families by Nuclear Magnetic
Resonance Spectroscopy and Distance Geometry**

by

Joseph Henry Pease

Ph.D. Thesis

December 1989

**Chemical Biodynamics Division, Lawrence Berkeley Laboratory
and Department of Chemistry, University of California
Berkeley, California 94720**

This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy, under contract number DE-AC03-76SF00098, by the National Science Foundation under grant number DMB 88-15998, and through instrumentation grants from the U.S. Department of Energy, DE-FG05-86ER75281, and the National Science Foundation, DMB 86-09035.

Acknowledgements

I am very grateful to my research director Dave Wemmer for his support, help, and enthusiasm during my stay at Berkeley. Vasant Kumar and Jeff Pelton provided a lot of advice and were always willing to listen. I would like to acknowledge David Koh and Enrique Dalmasso for their help with the peptide synthesizer. Finally I would like to thank my wife Ann and the rest of my family for all of their support over the years.

Table of Contents

Table of Figures	iv
Abbreviations	vi
Introduction	1
Chapter I	
NMR and Distance Geometry	3
Chapter II	
The Structure of Apamin	33
Chapter III	
Analogues of Apamin and the Structure of Apa-S25	58
Chapter IV	
Sea Anemone Peptides:	
NMR Assignments and Secondary Structure	76
Chapter V	
Sea Anemone Peptides: Tertiary Structure	98
Appendices	
Appendix I: NMR Pulse Programs	122
Appendix II: Distance Constraints	131
Appendix III: α -Carbon Coordinates of Peptides	151
References	156

Structures of Peptide Families by Nuclear Magnetic
Resonance and Distance Geometry

by

Joseph Henry Pease

Abstract

The three dimensional structures of several small peptides were determined using a combination of ^1H nuclear magnetic resonance (NMR) and distance geometry calculations. These techniques were found to be particularly helpful for analyzing structural differences between related peptides since all of the peptides' ^1H NMR spectra are very similar. The structures of peptides from two separate classes are presented.

Peptides in the first class are related to apamin, an 18 amino acid peptide toxin from honey bee venom. The ^1H NMR assignments and secondary structure determination of apamin were done previously [Wemmer, D. & Kallenbach, N. R. (1983) *Biochemistry* 22, 1901-1906]. Quantitative NMR measurements and distance geometry calculations were done to calculate apamin's three dimensional structure. The structure contains an N-terminal β -turn (residues 2 to 5) and a C-terminal α -helix (residues 9 to 18). Two analogs of apamin were chemically synthesized, combining secondary structure elements derived from apamin and the S-peptide from ribonuclease-A. Residues from apamin's C-terminal α -helix were replaced by those from the S-peptide except for the cystines, which are kept to conserve apamin's disulfide pattern. The peptides fold spontaneously into a structure essentially identical to that of native apamin. Apamin's N-

terminal β -turn and two disulfide bonds provided a framework which stabilized the α -helix of the S-peptide. The utility of using native disulfide bonds to stabilize structural elements from large proteins is discussed.

Peptides in the second class are 48 amino acid toxins from the sea anemone *Radianthus paumotensis*. The ^1H NMR assignments of toxin II were done previously (Wemmer, D. E., Kumar, N. V., Mettrione, R. M., Lazdunski, M., Drobny, G., & Kallenbach, N. R. (1986) *Biochemistry* 25, 6842-6849]. The ^1H NMR assignments of toxin III and the distance geometry calculations for both peptides are presented. Both peptides contain a highly twisted four stranded β -sheet core connected by loops of irregular structure. One particular difference between them is the positioning of the N-terminal strand of the β -sheet relative to its neighboring strand. These two structures were found to be similar to structures of related peptides from other sea anemones.

Table of Figures

Chapter I

- 1.1 Nuclear Spin Energy Level Diagram
- 1.2 Graph of $^3J_{\text{NH}-\alpha\text{H}}$ and Distance Constraints versus ϕ angle
- 1.3 Generic Two Dimensional NMR Experiment
- 1.4 Two Dimensional NMR Experiments
- 1.5 Theoretical COSY, RELAY, and TOCSY Spectra of Leucine
- 1.6 Aromatic Distance Constraints for Phenylalanine and Tyrosine
- 1.7 Amino Acids
- 1.8 One Dimensional NMR Spectrum of a Peptide
- 1.9 Sequential Distances used for Peptide Assignments
- 1.10 Short Distances for Secondary Structure Assignment
- 1.11 α -helix Showing Amide to Amide NOE Connectivities

Chapter II

- 2.1 Apamin Sequence with Observed Sequential Connectivities
- 2.2 Chemical Shifts of Assigned Resonances for Apamin
- 2.3 Distance Constraints used to Generate Apamin Structures
- 2.4 Upfield Region of a NOESY Spectrum of Apamin
- 2.5 NOE Buildup Curves for Apamin
- 2.6 N-Terminal β -turn in Apamin
- 2.7 Amide to Amide Region of a NOESY Spectrum
- 2.8 Distance Geometry Structures of Apamin
- 2.9 Table of Structure Analysis
- 2.10 Comparison of Apamin with other Structural Models

Chapter III

- 3.1 Sequences of Apamin and Hybrids of Apamin
- 3.2 Amide to Amide Region of a NOESY Spectrum of Apa-S25
- 3.3 Upfield Region of a NOESY spectrum of Apa-S25
- 3.4 Fingerprint Region of a COSY of Apa-S25
- 3.5 Observed Sequential Connectivities for Apa-S25
- 3.6 Chemical Shifts of Assigned Resonances for Apa-S25

- 3.7 Distance Geometry Structures of Apa-S25
- 3.8 Table of Structure Analysis
- 3.9 Standard Deviation for each Residue of Apa-S25
- 3.10 Comparison of Apamin and Apa-S25 Structures
- 3.11 RMS Deviations of Backbone Atom Positions of Apamin and Apa-S25

Chapter IV

- 4.1 Sequences of Sea Anemone Peptides
- 4.2 Upfield Region of a COSY Spectrum of Rp III
- 4.3 Upfield Region of a RELAY Spectrum of Rp III
- 4.4 Upfield Region of a NOESY Spectrum of Rp III
- 4.5 Fingerprint Region of a COSY Spectrum of Rp III
- 4.6 Fingerprint Region of a NOESY Spectrum of Rp III
- 4.7 Fingerprint Region of a D₂O COSY Spectrum of Rp III
- 4.8 Observed Sequential Connectivities for Rp III
- 4.9 Amide to Amide Region of a NOESY Spectrum for Rp III
- 4.10 Chemical Shifts of Assigned Resonances for Rp III
- 4.11 β -sheet Structure of Rp III
- 4.12 Chemical Shift Comparison of Rp II, Rp III, and ATX Ia

Chapter V

- 4.1 Distance Geometry Structure of Rp III
- 4.2 Distance Geometry Structure of Rp II
- 4.3 Structure of Rp II Superimposed on Rp III
- 4.4 Table of Structure Analysis for Rp II and Rp III
- 4.5 Standard Deviation for each Residue of Rp III
- 4.6 Regions of Rp II and Rp III

Abbreviations

Amino acid one and three letter code: A, Ala - alanine; C, Cys - cysteine; D, Asp - aspartic acid; E, Glu - glutamic acid; F, Phe - Phenylalanine; G, Gly - glycine; H, His - histidine; I, Ile - isoleucine; K, Lys - lysine; L, Leu - leucine; M, Met - methionine; N, Asn - asparagine; P, Pro - proline; Q, Gln - glutamine; R, Arg - arginine; S, Ser - serine; T, Thr - threonine; V, Val - valine; W, Trp - tryptophan; and Y, Tyr - tyrosine.

Å - Ångstrom

AA II - toxin II from *Androctonus australis*

AP-A - anthopleuran-A from *Anthopleura xanthogrammica*

ATX Ia - toxin Ia from *Anemonia sulcata*

BPTI - basic pancreatic trypsin inhibitor

CD - circular dichroism

cm - centimeter

COSY - correlated spectroscopy experiment

CPU - central processing unit

$d_{\alpha N}$ - alpha proton to neighboring amide proton distance

$d_{\beta N}$ - beta proton to neighboring amide proton distance

d_{NN} - amide proton to neighboring amide proton distance

DTT - dithiothreitol

γ - magnetogyric ratio

H - magnetic field strength

HPLC - high performance liquid chromatography

Hz - hertz

$^3J_{\alpha H-\beta H}$ - alpha proton to beta proton vicinal coupling constant

$^3J_{NH-\alpha H}$ - amide proton to alpha proton vicinal coupling constant

kD - kilodalton

M - molar

MHz - megahertz

mg - milligrams

m_I - nuclear spin quantum number

min - minute

ml - milliliter

mM - millimolar
ms - milliseconds
mw - molecular weight
NMR - nuclear magnetic resonance
NOE - nuclear Overhauser effect
NOESY - NOE spectroscopy experiment
ppm - parts per million
RELAY - relayed coherence transfer spectroscopy experiment
r. m. s. or RMS - root mean square
RNase-A - ribonuclease A
Rp II - toxin II from *Radianthus paumotensis*
Rp III - toxin III from *Radianthus paumotensis*
Sh I - toxin I from *Stichodactyla helianthus*
S-peptide - first 20 amino acids from RNase-A
 τ_c - correlation time
TFA - trifluoroacetic acid
 τ_m - mixing time
TOCSY - total correlation spectroscopy experiment
11NOESY - NOESY with a 11-echo detection pulse

Introduction

In this thesis I describe the three dimensional structures of several small proteins which I determined using nuclear magnetic resonance (NMR) spectroscopy and distance geometry techniques. NMR was used for two reasons: First, small peptides (< 10kD) frequently do not crystallize thus making X-ray crystallography impossible; and second, NMR is a much faster structure determining technique than X-ray crystallography. NMR is particularly useful for studying classes of proteins since all of the NMR spectra are very similar. I analyzed two separate classes of proteins and thus I divided this thesis into two parts.

The first class of proteins I looked at were related to apamin, a very small polypeptide toxin from honey bee venom. The NMR assignments and secondary structure determination of apamin were done previously (Wemmer & Kallenbach, 1983). I did further NMR measurements and calculated apamin's three dimensional structure. I then chemically synthesized two analogs of apamin, combining secondary structure elements derived from apamin and the S-peptide from ribonuclease-A. I did a complete structural analysis for one of these peptides.

In the second part of this thesis, I will describe the three dimensional structure of a class of sea anemone polypeptide toxins. I determined the structure of *Radianthus paumotensis* III (Rp III) and compared its structure with Rp II's structure, determined with Dr. Vasant Kumar in our group, a similar protein from the same sea anemone. I also compare these two structures to the structures of ATX Ia from *Anemonia sulcata* characterized

by two groups (Widmer et al., 1988; Gooley & Norton, 1986), AP-A from *Anthopleura xanthogrammica* (Torba et al., 1988), and Sh I from *Stichodactyla helianthus* (Norton et al., 1989; Fogh et al., 1989).

Chapter I

NMR and Distance Geometry

The NMR methodology is described in detail in several different books (Wüthrich, 1986; Ernst et al., 1987). I will quickly describe basic NMR principles and then describe the pulse sequences I used with any modifications made from the basic experiments. Distance geometry techniques, on the other hand, have been rapidly changing due to the recent availability of very powerful computers and better software. Basic distance geometry theory is presented in several places (Crippen, 1981; Havel & Wüthrich, 1984; Havel et al., 1983) and thus I will only describe the techniques that I used.

NMR

Nuclear magnetic resonance is the observation of the splitting of nuclear spin energy levels in the presence of a magnetic field. This splitting is directly proportional to the applied magnetic field and to the magnetogyric ratio (γ), a constant for each nucleus which represents the sensitivity of the energy levels to an applied magnetic field.

$$\Delta E = \gamma H \quad (1-1)$$

Protons (^1H) are the most sensitive nucleus to study since it has the second largest γ (^3H is the largest) and has a 100% natural abundance (^3H , 0%). This is the primary reason I used proton NMR for all of my experiments. Protons are a spin 1/2 nucleus and thus give rise to two energy levels in

the presence of a magnetic field (figure 1.1). The energy difference measured depends on the magnetic field felt by the individual proton, called the effective field. This effective field is very sensitive to the proton's environment, and differences are termed chemical shifts (δ). Shifts are primarily due to electronic shielding of the protons from the applied magnetic field. In the presence of a magnetic field, electrons circulate such that a secondary magnetic field is generated which opposes the applied field. Resonance frequencies are usually referenced to a standard compound with a known frequency (ν) and are reported in parts per million (ppm).

$$\delta = \frac{10^6 [\nu_{\text{sample}} - \nu_{\text{standard}}]}{\nu_{\text{standard}}} \quad (1-2)$$

Here ν is the frequency in Hz and is on the order of 500MHz for my experiments. Proton resonances in peptides have a frequency range of approximately 6,000 Hz (12 ppm) for a 500 MHz NMR spectrometer. One of the largest environmental sources of chemical shift is the ring current effect which can shift protons near aromatic rings up to about one ppm. Since chemical shifts depend on the local environment surrounding a proton, they can be used for structural analysis. The chemical shifts of the sea anemone peptides were analyzed in detail (see chapter IV) to look for structural similarities between the different peptides. There are two other important NMR effects which can be used to obtain structural information: scalar coupling and the nuclear Overhauser effect (NOE). These are discussed below.

$$E = -\gamma m_I H(1-\delta)$$

Figure 1.1 The splitting of the proton nuclear spin energy levels as a function of the magnetic field strength. γ is the magnetogyric ratio, m_I is the nuclear spin quantum number, and δ is the chemical shift. The splitting depends on the field felt by the proton, and hence depends on the proton's environment.

As protons are sensitive to the presence of electrons through the chemical shift, protons close to one another also interact. The electrons in the bonds (typically three or fewer) connecting the two protons are polarized, reflecting the spin states, splitting the resonances. This is called spin-spin or scalar coupling, and the coupling constant (J) is independent of magnetic field. The splitting is usually on the order of 0 to 20 Hz. It is important to note that this effect is only transmitted through bonds. The value of J , like the chemical shift, depends on the environment and is often used to determine dihedral angles using a Karplus relationship (Karplus, 1959). The most important of these for peptide conformational analysis is the amide to alpha proton vicinal coupling ($^3J_{\text{NH}-\alpha\text{H}}$). The splitting value is given by

$$^3J_{\text{NH}-\alpha\text{H}} = 6.4\cos^2\theta - 1.4\cos\theta + 1.9 \quad (1-3)$$

where J is in Hz and $\theta = |\phi - 60^\circ|$ is the dihedral angle (Pardi et al., 1984). In order to be useful for distance geometry this angle information is converted into distance information using figure 1.2. It is unfortunate that for many values of $^3J_{\text{NH}-\alpha\text{H}}$ there are multiple values of ϕ . In fact, the only time this relationship is really useful is for extended conformations ($^3J_{\text{NH}-\alpha\text{H}} > 7$ Hz) where reasonably strong distance constraints can be obtained. Stronger constraints are obtained if only regular secondary structure is present in peptides. For BPTI, there was only one exception to having all values of ϕ consistent with the common secondary structures (figure 1.2) (Wüthrich, 1986). This exception shows that ϕ can have values outside of

	J (Hz)	ϕ
parallel β -sheet	9.7	-119°
antiparallel β -sheet	8.9	-139°
α -helix	3.9	-57°
3^{10} -helix	4.2	-60°

Figure 1.2 The scalar coupling constant (J) as a function of ϕ angle. Distance constraints are determined using a measured J value. Note: one value of J can lead to several values of ϕ , making weak distance constraints. Only large values of J (> 7 Hz) give strong distance constraints.

the common secondary structure region. In the case of apamin, the information from coupling was found, in general, to be redundant to that from NOEs (see chapter II). The most important aspect of scalar coupling is that it allows one to connect proton resonances within amino acid residues using through bond information.

The second important interaction arises through dipolar relaxation and is called the nuclear Overhauser effect (NOE). The magnitude depends on both the correlation time for tumbling of the internuclear vector between two protons, and the internuclear distance. Protons which are fairly close in space ($< 4 \text{ \AA}$) will experience this interaction. Since this is a through space interaction, it is used to obtain direct distance information. This interaction is described in more detail later in this chapter when the two dimensional NOE experiment is discussed. The two dimensional NMR experiments I used select for either the through bond (scalar coupling) or through space (NOE) interaction.

Almost all of the proton NMR experiments I have done are two dimensional (2D). A schematic of a generic 2D NMR experiment is shown in figure 1.3. The preparation period is for preparation of spins in a specific nonequilibrium state. The evolution period (t_1) is the time when the individual spins are frequency labeled so that the origin of the signal (cross peak) is known. This time is incremented to give the second frequency axis and is usually incremented so that a square matrix is obtained. The mixing period (τ_m) is the time when the desired transfer of information takes place. As was stated above, this transfer can be either

Figure 1.3 Generic 2D NMR experiment. Preparation, preparation of spins in a specific nonequilibrium state; evolution (t_1), protons are frequency labeled so that the origin of the cross peaks is known; mixing (τ_m), desired transfer of information takes place, either through bond or through space; and detection (t_2), time when the signal is recorded, contains information from t_1 and τ_m . The value of t_1 is incremented to obtain a square matrix.

through bond or through space. The detection period (t_2) is the time when the signal, which contains information from t_1 and τ_m , is recorded. The raw data consists of a square matrix with two time parameters: $S(t_1, t_2)$. The data is first Fourier transformed with respect to t_2 , and then with respect to t_1 to yield a frequency matrix: $S(f_1, f_2)$. Diagonal peaks ($f_1=f_2$) represent the one dimensional spectrum and cross peaks ($f_1 \neq f_2$) represent correlations between the protons. Spectra are usually analyzed as a topographical map, with contour lines to show the different peak intensities. The 2D NMR experiments I used for assigning NMR spectra of proteins are shown diagrammatically in figure 1.4 (the actual pulse programs are listed in appendix I). These experiments are described below.

The first 2D NMR experiment developed was COrelated Spectroscopy (COSY) (Jeener, 1971; Aue et al., 1976), which has cross peaks between pairs of scalar (J) coupled protons. Sufficiently strong scalar couplings occur for protons that are two or three bonds apart, and occasionally extending to four bonds: e.g. between His $\epsilon_1\text{H}$ and $\delta_2\text{H}$ protons. A RELAYed coherence transfer experiment (RELAY) is essentially the combination of two COSY experiments. Cross peaks are observed between pairs of protons that have a common coupling partner whether they are directly coupled or not. The TOtal Correlation Spectroscopy (TOCSY) experiment (Davis & Bax, 1985; Bax & Davis, 1985; Braunschweiler & Ernst, 1983) has cross peaks between all pairs of protons in a spin system. A spin system is a set of protons which are connected through some coupling pathway. Figure 1.5 shows the cross peaks obtained from each of these experiments for a Leucine residue.

Figure 1.4 2D NMR experiments used for protein structure determination. COSY, COrrrelated SpectroscopY; RELAY, RELAYed coherence transfer spectroscopy; TOCSY, TOtal Correlation SpectroscopY; and NOESY, Nuclear Overhauser Effect SpectroscopY. Experiments are described in the text.

Figure 1.5 COSY, RELAY, and TOCSY spectra for a Leu residue. Diagonal peaks are denoted by black dots and cross peaks by C, R, and T for COSY, RELAY, and TOCSY cross peaks respectively.

The Nuclear Overhauser Effect Spectroscopy (NOESY) experiment (Jeener et al., 1979; Kumar et al., 1980) has cross peaks between pairs of protons which are close in space (via the NOE), less than 4Å apart. The NOESY peak intensities are given by

$$a_{kl}(\tau_m) = e^{-R\tau_m} \frac{n_l}{N} M_0 \quad (1-4)$$

where R is the relaxation matrix, τ_m is the mixing time, n_l is the number of equivalent nuclei, and M_0 is the total equilibrium magnetization of the N nuclei (Macura & Ernst, 1980). For short mixing times, the exponential can be simplified using a Taylor series

$$e^{-R\tau_m} \approx \delta_{kl} - R_{kl}\tau_m \quad (1-5)$$

where R_{kl} is the relaxation rate constant between protons k and l . The cross peak intensity ($k \neq l$) is then given by

$$a_{kl}(\tau_m) = -R_{kl}\tau_m \frac{n_l}{N} M_0 \quad (1-6)$$

The result is that the cross peak intensity is directly proportional to the cross relaxation rate (σ_{kl} , which equals $n_l R_{kl}$) between two protons when short mixing times are used. The cross relaxation rate for a pair of protons, i and j , is given by

$$\sigma_{ij} = 2 \left| W_2^{ij} - W_0^{ij} \right| \quad (1-7)$$

where W_2 and W_0 represent double and zero quantum transition probabilities respectively. These are usually written in terms of spectral densities (J_{ij})

$$W_2^{ij} = 6q_{ij} J_{ij}(2\omega_0) \quad (1-8)$$

$$W_0^{ij} = q_{ij} J_{ij}(\emptyset) \quad (1-9)$$

where

$$q_{ij} = \frac{1}{10} \gamma_H^4 \left(\frac{\mu_0 h}{8\pi} \right)^2 r_{ij}^{-6} \quad (1-10)$$

and

$$J_{ij}(\omega) = \frac{\tau_c^{ij}}{1 + (\omega\tau_c^{ij})^2} \quad (1-11)$$

The γ_H is the magnetogyric ratio of the proton, μ_0 is the permeability of free space, h is Planck's constant, r_{ij} is the internuclear distance, τ_c is the correlation time for the internuclear vector between i and j , and ω is the resonant frequency ($2\pi \cdot 500$ MHz). For the peptides I studied, τ_c is long enough such that the spin diffusion limit applies, $\omega\tau_c \gg 1$. In this case only W_0 contributes significantly to cross relaxation:

$$\sigma_{ij} = 2 q_{ij} \tau_c^{ij} \quad (1-12)$$

The final assumption is that the molecule of interest rotates as a rigid body making all the internuclear correlation times the same. This expression can then be written as

$$\sigma_{ij} = k \langle r_{ij}^{-6} \rangle \quad (1-13)$$

where k is a constant. Unknown distances are determined using known distances and the ratio of NOESY cross peak intensities.

$$r_{ij} = r_{ab} \left[\frac{\sigma_{ab}}{\sigma_{ij}} \right]^{\frac{1}{6}} \quad (1-14)$$

Here r_{ij} is the unknown distance, r_{ab} is a known distance, and σ is directly proportional to the NOESY cross peak intensity (in the short mixing time approximation, eq. 1-6). Different fixed distances are used depending on the NOE involved. The geminal proton NOEs (1.75 Å), usually between the beta protons of amino acids, are used to calibrate NOEs between single protons, Ala alpha to methyl proton NOEs (2.38 Å) are used to calibrate NOEs between methyl groups and single protons, and Tyr aromatic NOEs (2.5 Å) to calibrate longer NOEs between single protons.

There are two special cases for deriving distance constraints from NOEs. The first is for NOEs involving the aromatic side chains of Tyr and Phe. In all the peptides I studied, these aromatic rings were rapidly rotating about the $\beta\text{C} - \gamma\text{C}$ bond. The NOEs observed are then actually an average for the

two sites occupied by the proton involved, and constraints must be modified to account for this (Wüthrich et al., 1983). NOEs observed to the δH or ϵH were converted into distance constraints to the γC and ζC of the aromatic ring (figure 1.6). The second case is for NOEs to nondegenerate βHs . The βC of these residues is defined to be not chiral so that stereo specific assignments are not required (Weber et al., 1988). This allows the protons to interchange positions without generating a chirality error (see distance geometry section). In the case of degenerate βHs , the constraint is defined to the βC by adding the appropriate $\beta\text{C}-\beta\text{H}$ bond length ($1.0\text{\AA}$) to the NOE constraint (Wüthrich et al., 1983).

There are some problems associated with NOE derived distance constraints. The first is the assumption that the correlation time is the same for all proton pairs. This is clearly not the case for apamin or the hybrids of apamin, where the C-terminal end of the helix frays: NOEs for fixed distances, such as Ala alpha to methyl protons are much weaker at the end of the helix (see chapter III). Fixed distance NOE intensities from throughout the molecule can be used to help interpret local motion. The second problem occurs when there are multiple conformations of a peptide present. In this case conformers with shorter interproton distances are more heavily weighted due to the NOE dependence on the inverse sixth power of the distance. All of the peptides I studied have only one set of resonances (one conformation), but there does appear to be some motion in regions of the peptides. A final problem is a secondary NOE effect called spin diffusion. This occurs when two protons are very close in space, such as geminal protons ($1.75\text{\AA}$), and there is a direct NOE to one of these proton

Constraint to $\gamma\text{C} = 3.3\text{\AA} + \text{NOE}$

Constraint to $\zeta\text{C} = 2.1\text{\AA} + \text{NOE}$

(R = H for Phe and OH for Tyr).

Figure 1.6 Distance constraint for NOEs to either a Tyr or Phe aromatic proton (δH or ϵH). NOE to ϵH is pictured here. NOE to δH would have a similar constraint: to $\gamma\text{C} = 2.1\text{\AA} + \text{NOE}$ and to $\zeta\text{C} = 3.3\text{\AA} + \text{NOE}$. NOE is the distance constraint determined by the NOESY cross peak intensity.

but not the other. A secondary NOE will be seen to the other proton, the effect having been transmitted by the intermediate proton. This problem is minimized at very short mixing times. These problems will be discussed in more detail for each peptide that I studied.

Hydrogen bond donors are identified by looking for slowly exchanging amide protons. Amide protons slowly exchange due to either hydrogen bonding or from being buried within the peptide (protected from solvent)(Englander & Kallenbach, 1984). For small peptides, most residues are solvent accessible and therefore, slow amide exchange is predominantly from hydrogen bonding. This is done by dissolving a peptide sample freshly lyophilized from H₂O, in D₂O (Wüthrich & Wagner, 1979; Wagner & Wüthrich, 1982). Amide proton exchange rates are then determined by observing the disappearance of the amide proton signals as a function of time. The most slowly exchanging amide protons are easily determined by running a COSY spectrum of the D₂O sample; those cross peaks observed between amide and alpha protons are from slowly exchanging amide protons. One problem with this technique is that cross peaks which are weak even in a COSY in H₂O will be less sensitive, since fewer amide protons need to exchange for the COSY cross peak to disappear. Hydrogen bond acceptors are derived from the observed secondary structure (NOE assignment pattern, see below). Often some slowly exchanging amides can not be assigned to hydrogen bonds since no unambiguous acceptor can be determined.

I assigned the proton NMR spectra of several peptides using the sequential assignment algorithm developed by Wüthrich's group (for review see Wüthrich, 1986). The first step is to assign proton resonances to spin systems using COSY and RELAY (or TOCSY) spectra taken in D₂O solution. D₂O solution is used to obtain better sensitivity since the strong water resonance (110 M water signal compared to a 5 mM peptide signal) is not present. The amide resonances, however, must be observed in H₂O. Spin systems are assigned to amino acid groups as shown in figure 1.7. Some amino acids are easily identified since they have unique spin systems. Specific types of protons usually resonate in specific regions of the proton NMR spectrum as is shown in figure 1.8, aiding in the assignment of spin systems to particular amino acids. Once a large number of spin systems are identified, the peptide sequence (chemically determined) is analyzed for connectivities using a NOESY spectrum in H₂O. Three types of NOE connectivities are particularly important for sequential assignments: NH(i) to NH(i+1), α H(i) to NH(i+1), and β H(i) to NH(i+1) (figure 1.9 and 1.10). At least one of these NOEs will be short enough to be observed (Billeter et al., 1982) and the NOEs observed depend on the secondary structure present (Wemmer & Kallenbach, 1983; Wüthrich et al., 1984). Strong NOEs of the NH(i) to NH(i+1) type are found in helical regions (figure 1.10 and 1.11), including turns; strong α H(i) to NH(i+1) NOEs are observed in regions of extended structure like that found for β sheets and in regions of random coil; and β H(i) to NH(i+1) NOEs are found in all kinds of secondary structure but are found regularly for helices. These initial secondary structure assignments can be confirmed by longer range (between residues far apart in sequence) NOEs, as shown in figure 1.10, and also by the observed slowly

Figure 1.7 Amino acids with assigned spin systems. Both the one letter and three letter amino acid codes are shown. Protons are labelled with Greek letters.

Figure 1.8 One dimensional proton NMR spectrum of RpIII, a 48 amino acid peptide. Chemical shift regions are labelled in the spectrum. The water resonance has been zeroed (peak at 4.8 ppm).

Figure 1.9 Distances important for sequential assignments. d_{NN} , amide to amide interproton distance which is short for helical regions including turns; $d_{\alpha\text{N}}$, alpha to amide interproton distance which is short for extended regions such as β -sheet and for regions of random coil; and $d_{\beta\text{N}}$, beta to amide interproton distance which can be short for most types of secondary structure, but is usually found regularly in helices.

Figure 1.10:
Short Distances for Secondary Structure Determination

α -helix

sequential	Long Range	Intraresidue
$d_{NN} = 2.8$	$d_{\alpha\beta}(i, i+3) = 2.5 - 4.4, 4.0, 3.1^*$	$d_{N\alpha} = 2.7$
$d_{\alpha N} = 3.5$	$d_{\alpha N}(i, i+3) = 3.4$	$d_{N\beta} = 2.0 - 3.4$
$d_{\beta N} = 2.5 - 4.1, 3.8, 3.0^*$		

Type I β -turn: 4 residues, NH(4) hydrogen bonds to O(1).

$d_{NN}(2,3) = 2.6$	$d_{\alpha N}(1,4) = 3.1 - 4.2$	$d_{N\alpha}(2,2) = 2.7$
$d_{NN}(3,4) = 2.4$		$d_{N\alpha}(3,3) = 2.8$
$d_{\alpha N}(2,3) = 3.4$		$d_{N\beta}(2,2) = 2.0 - 3.4$
$d_{\alpha N}(3,4) = 3.2$		$d_{N\beta}(3,3) = 2.1 - 3.5$
$d_{\beta N}(2,3) = 2.9 - 4.4, 4.1, 3.4^*$		

Type II β -turn: 4 residues, NH(4) hydrogen bonds to O(1) and 3rd residue is Gly.

$d_{NN}(3,4) = 2.4$	$d_{\alpha N}(2,4) = 3.3$	$d_{N\alpha}(2,2) = 2.7$
$d_{\alpha N}(2,3) = 2.2$		$d_{N\alpha}(3,3) = 2.2$
$d_{\alpha N}(3,4) = 3.2$		$d_{N\beta}(2,2) = 2.0 - 3.4$
		$d_{N\beta}(3,3) = 3.2 - 4.0$

β -sheet

$d_{\alpha N} = 2.2$		$d_{N\alpha} = 2.8$
$d_{\beta N} = 3.2 - 4.5, 4.2, 3.6$ (anti)*		$d_{N\beta} = 2.6 - 3.8$ (anti)
	$d_{\alpha\alpha}(i,j) = 2.3$ (anti)	$d_{N\beta} = 2.4 - 3.7$ (para)
	$d_{\alpha N}(i,j) = 3.2$ (anti)	
	$d_{NN}(i,j) = 3.3$ (anti)	
	$d_{\alpha N}(i,j) = 3.0$ (para)	

Pro - X

$d_{\delta N} = 2.8 - 5.7$	$d_{\delta N} = 2.9$ (α -helix)
----------------------------	---

X - Pro

$d_{N\delta} = 1.9 - 4.8$	$d_{\alpha\delta} = 2.1 - 3.8$	$d_{\alpha\delta} = 2.1$ (β -sheet)
		$d_{N\delta} = 2.1$ (α -helix)

* distance to 1, 2, and 3 β -protons respectively; anti - antiparallel β -sheet; para - parallel β -sheet. All distances in Å and only distances < 3.5 Å are shown.

Figure 1.11 The α -helix from apamin showing the amide to amide proton NOE connectivities. The amide protons are the dots and the NOEs are shown by the arrows.

exchanging amides, which are presumed to be due to hydrogen bonding. Tertiary structure is determined by either model building or distance geometry. Model building has the inherent problem that it is biased by the builder of the model. Distance geometry, on the other hand, generates a randomized set of conformers, which are refined to be consistent with a set of distance constraints. I therefore used distance geometry for all of the structures I have generated.

Distance Geometry

Distance geometry techniques involve converting a set of distances between atoms into three dimensional coordinates, a process called embedding. It is a general method for computing a random sampling of molecular conformations which are consistent with a set of input distance constraints. The calculations I did were done using the DSPACE program (D. Hare, unpublished) on a Micro Vax II or Micro Vax III computer. The program uses the metric matrix approach as described by Crippen (1981). The basic protocol for obtaining embedded structures is generally similar to that in the program DISGEO (Havel & Wüthrich, 1984), but DSPACE uses all the atoms when embedding, rather than the subembedding step as is used in DISGEO.

The first step for structure determination using distance geometry is to generate a bounds matrix. Each element of this matrix contains a distance between pairs of atoms (i and j) in the peptide. The diagonal is zero, one side of the diagonal contains lower bounds (l_{ij}) and the other side contains

upper bounds (u_{ij}). Upper and lower bounds are used to take into account uncertainties in distances. This matrix is initially constructed by entering the chemical sequence into the program. The program then uses templates derived from crystallographic data to set bond lengths, geminal distances (bond angles), chirality constraints, and the planarity of peptide bonds and aromatic rings. These covalent constraints are in the form of exact distances ($u_{ij} = l_{ij}$). This matrix also includes constraints from the NOE data, hydrogen bonds deduced from the slowly exchanging amides, and the disulfide bonds. NOE distance constraints have a distance range corresponding to the uncertainty of the mixing time dependence of the NOE, hydrogen bonds have upper and lower bounds corresponding to the range of hydrogen bond lengths observed in crystal structures of proteins, and the distance between the β -carbons in a disulfide bond is defined so that it adopts a standard conformation, making the $C\beta$ -S-S- $C\beta$ dihedral angle of approximately $\pm(90^\circ \pm 10^\circ)$ (Richardson, 1981). All atoms of the molecule are included explicitly, except for methyl protons which are defined to be a pseudo atom since each methyl group gives rise to only one NMR resonance, due to the rapid rotation of the methyl group (Wüthrich et al., 1983). Unknown upper bounds are set to 1000 Å (effectively no constraint) and unknown lower bounds to the sum of van der Waals radii. The bounds are improved by applying a triangle inequality, making a smoothed bounds matrix. The triangle inequality is a simple geometric constraint. For example, for three atoms a, b, and c the triangle inequalities are

$$u_{ac} \leq u_{ab} + u_{bc} \quad (1-15)$$

$$l_{ac} \geq l_{ab} - u_{bc} \quad (1-16)$$

In words, the length of the third side of a triangle must be less than or equal to the sum of the lengths of the other two sides (eq. 1-15), and must also be greater than or equal to the difference of the lengths of the other two sides (eq. 1-16). The triangle inequality is applied to the upper bounds first, reducing all upper bounds until no further alterations can be made; and then applied to the lower bounds, increasing lower bounds until no further alterations can be made. There are also higher order geometric constraints, such as a quadrangle inequality, but they do not improve the bounds significantly, they require much larger amounts of CPU time, and hence, DSPACE does not use them. This smoothed bounds matrix is used for generating coordinates and for refinement of the embedded structures.

The next step is to generate a symmetric trial matrix. This matrix is generated by either randomly picking distances between the upper and lower bounds for each pair of atoms, or by selecting one random distance, then setting other distances, within their allowed range, to be consistent with the smoothed bounds matrix and the initially chosen distance. This second method generates what are called correlated trial distances. Although the second process takes a considerable amount of time, the embedded structures are clearly better than those with all distances chosen randomly. This increase in time, however, is larger than the extra time needed to refine a random trial distance structure (for later versions of DSPACE, used after the apamin structure was determined). Different random seeds are used to generate different trial matrices and hence

different initial coordinates. It is at this stage that the range of allowed conformations is explored at random.

A metric matrix is generated from the symmetric matrix, resulting in a "center of mass" coordinate system where the center of mass is the origin. The "center of mass" is calculated by assuming a mass of one for each atom. The distance to the center of mass is calculated using the formula (Crippen, 1981)

$$d_{i0}^2 = \frac{1}{n} \sum_{j=1}^n d_{ij}^2 - \frac{1}{2} \sum_{j=2}^n \sum_{k=1}^{j-1} d_{jk}^2 \quad (1-17)$$

where d_{i0} is the distance from atom i to the origin and n is the total number of atoms in the matrix. Each element of the metric matrix is given by (Crippen, 1981)

$$g_{ij} = \frac{1}{2} (d_{i0}^2 + d_{j0}^2 - d_{ij}^2) \quad (1-18)$$

This matrix is diagonalized and the three eigenvectors associated with the three largest eigenvalues of this matrix are taken to be the three orthogonal axes: x , y , and z . When four dimensional refinement was used, the four largest eigenvalues and their corresponding eigenvectors were used to generate four axes. The coordinates are generated using these eigenvectors. The initial coordinates generated in this way do not satisfy all of the original constraints, and subsequently refined using a penalty function generated from the smoothed bounds matrix.

The penalty function used during refinement includes independently weighted terms which correspond to chirality; planarity; geminal, bond, van der Waals distances; experimental distances; hydrogen bonds; linearity of hydrogen bonds; constraints derived from smoothing; and constraints which minimize the coordinate corresponding to the fourth dimension. The atoms in the bounds matrix are flagged for chirality and planarity when the atoms are first input into the matrix. Chirality and planarity use signed volumes to generate penalty values. For example, an α -carbon is chiral and has four atoms around it N, β C, α H, and C. The α H is the atom which always will be moved if the chirality is wrong. The other three atoms define a plane from which the signed volume is generated. To test the signed volume: first, two vectors are calculated from the coordinates of C and N and C and β C; then the cross product of these vectors generates a vector perpendicular to the plane which is subsequently normalized; finally a third vector is calculated from the coordinates of C and α H and the scalar product with the normalized vector yields the distance from the plane. This process is also used for atoms in the templates to determine the correct value of the signed volume for the known chirality and distance from the plane. The difference between this reference value and actual value at any point in the refinement is used to generate the penalty value. All of the other terms come directly from the bounds matrix except for the fourth dimension term, which just minimizes the values of the coordinate in the fourth dimension. The penalty function is parabolic for all terms. For upper bounds violations it takes the form

$$\text{penalty} = \text{weight} (d - u)^2 \quad (1-19)$$

and for lower bounds violations

$$\text{penalty} = \text{weight} (d - l)^2 \quad (1-20)$$

where d is the distance between a pair of atoms and u and l are the smoothed upper and lower bounds respectively. When the distance is between the upper and lower bounds there is no penalty value. The gradient used for refinement is the derivative of this expression.

This penalty function is less than ideal in several ways. First, the Van der Waals repulsion is only quadratic with respect to distance. Also, there is no attractive term for dispersion forces. This is far from the accepted Leonard Jones 6-12 potential for structure calculations (Schultz & Shirmer, 1979). This gives rise to distance geometry structures being more loosely packed (side chains extended etc.) than structures refined using a function including a full Leonard-Jones potential. The overall fold, however, is essentially the same.

I used two different refinement algorithms: conjugate gradient and simulated annealing. I also used four dimensional refinement to generate most of the peptide structures (except apamin). Four dimensional refinement is more efficient since it allows atoms to pass through one another in three dimensional space, preventing the peptide from becoming tangled. The conjugate gradient routine minimizes the penalty

function using an adjustable cutoff distance for testing violations. This was done either locally, for specific groups or residues in the molecule, or for the entire molecule at once with a cutoff of 5 Å. The simulated annealing method uses a molecular dynamics approach, assigning a random velocity term $v(i)$ to each atom i with adjustable average value. At each step the new velocity is calculated using:

$$v(i) = [v(i) - k * \text{gradient}(i)] / (1 + \text{damp}) \quad (1-21)$$

where k is the force constant for maintaining bounds. The system evolves dynamically using these essentially classical equations of motion, maintaining constant total error instead of energy. It is not a true dynamics refinement since it does not use real potential energy terms, but it serves a very similar function in moving the molecule out of local minima in the error function. When desired a damping term (damp) is added, it gradually reduces the velocities, and hence the total error. A cycle of assigning random velocities, letting the molecule evolve dynamically, then damping to gradually dissipate the velocity, was repeated several times with lower initial average velocity in each cycle, worked well to refine the structures. When these cycles were no longer effective at further reducing the total error, conjugate gradient minimization was done until no further reduction of error could be achieved. The ability to independently weight each element of the penalty function (and to change these weights during refinement) also provides a mechanism for avoiding local minima during refinement with either method.

The final result is a conformation of the molecule which satisfies the input constraints. Several structures are generated taking advantage of the randomization intrinsic to the process with the hope that the structural features in each will be similar. If this isn't the case, then more distance constraints are needed to better define the structure. These can be found in part by analyzing the resulting structures for NOEs which should be present, and then examining the NOESY spectrum to see if they are present. Another possibility is to use negative NOE distance constraints: if no NOE is observed between a particular pair of protons, then the distance between them is greater than 3Å. The use of negative data is questionable and extreme care must be taken when it is used. None of the structure refinements described here used negative information.

Chapter II

The Structure of Apamin

Apamin is a small neurotoxic peptide component of honey bee venom. Like many other peptide neurotoxins apamin has a high cystine content and a high basicity, but apamin is different from most peptide toxins in its unusual ability to cross the blood brain barrier and act on the central nervous system (Habermann, 1972). Apamin is known to block calcium dependent potassium fluxes, possibly by binding to a Ca^{+2} dependent potassium channel (Banks et al., 1979). Additionally apamin serves as a model for understanding various aspects of peptide folding, and amide proton exchange. To determine the structural basis of its activity, and for interpretation of other experimental results, it is important to have a good understanding of the structure of this peptide. Apamin contains 18 amino acids with two disulfides (see figure 2.1), which give it an extremely stable structure with respect to temperature, pH, and denaturants, not unfolding completely even at 70°C in 6M guanidinium hydrochloride (Miroshnikov et al., 1978). During the last few years there have been a number of studies aimed at determination of apamin's structure, using energy refinements (Freeman et al., 1986; Hider & Ragnarsson, 1980 and 1981), circular dichroism (CD) (Miroshnikov et al., 1978), and nuclear magnetic resonance spectroscopy (NMR) (Wemmer & Kallenbach, 1983; Okhanov et al., 1980; Bystrov et al., 1978 and 1980). Here I use two dimensional proton NMR data with distance geometry to obtain an improved picture apamin's solution structure.

Figure 2.1 Sequence of apamin with observed slowly exchanging amides (+) and observed secondary structure NOEs. There are two disulfides: Cys 1 to Cys 11 and Cys 3 to Cys 15. NOEs observed where the size depends on the NOE intensity. Ala 5 to Pro 6 δ -protons; and ? possible NOE but chemical shift degeneracy makes it impossible to tell. The lines indicate that the long range NOE was observed.

Previously, two dimensional NMR was used to obtain almost complete assignments of apamin's proton NMR spectrum (Wemmer & Kallenbach, 1983). Additional assignments were obtained and some were corrected. Figure 2.2 shows the chemical shifts of apamin's assigned proton resonances. As with many other peptides, the sequential assignment pattern and identification of slowly exchanging amides gave information about the secondary structure of the molecule. Using this information Wemmer and Kallenbach (1983) derived a model which was consistent with the previously reported CD data (Miroshnikov et al., 1978) and ϕ angles from $^3J_{\text{NH}-\alpha\text{H}}$ coupling constants (Bystrov et al., 1980). Freeman et al. (1986) have also developed structural models for apamin based on energy refinement methods, and Zell and coworkers (1987) have used coordinates from their crystal structure of a scorpion toxin, for which one section seemed to be structurally similar to apamin, together with modelling and energy refinement to develop a structural model. Here I describe the detailed structure obtained from further NMR measurements and distance geometry calculations, and compare the previously derived structural models with it.

Materials and Methods

NMR Spectroscopy. All two dimensional nuclear Overhauser effect (NOESY) spectra (Kumar et al., 1980; Jeener et al., 1979) were recorded on a 500MHz General Electric GN-500 spectrometer using time proportional phase incrementation (TPPI) to obtain phase sensitive spectra (Drobny et al., 1979; Bodenhausen et al., 1984). Each NOESY spectrum had 512 t_1 points, a

Figure 2.2
Chemical Shifts of Assigned Resonances, pH 2.0, 25°C^{a,b}

residue	amide	α -H	β -H	other
Cys-1	c	4.55	3.12, 2.74	
Asn-2	9.11	4.84	3.05, 2.74	
Cys-3	9.00	4.63	3.30, 2.75	
Lys-4	8.02	4.20	1.85, 1.85	γ 1.45; δ 1.68; ϵ 2.96; ϵ N7.50
Ala-5	7.24	4.52	1.16	
Pro-6	---	4.66	1.90, 1.75	γ 2.14, 1.98; δ 3.55, 3.44
Glu-7	9.05	4.39	2.25, 2.08	γ 2.66, 2.57
Thr-8	7.43	4.60	4.66	CH ₃ : 1.25
Ala-9	8.93	4.14	1.44	
Leu-10	8.28	4.07	1.62	γ 1.58; CH ₃ : 0.92, 0.88
Cys-11	7.75	4.52	3.16, 2.74	
Ala-12	8.44	3.81	1.43	
Arg-13	7.96	4.08	1.92	γ 1.70; δ 3.15; ϵ N7.20
Arg-14	8.21	4.06	1.98	γ 1.74; δ 3.16; ϵ N7.22
Cys-15	8.35	4.38	2.99	
Gln-16	7.65	4.20	2.15	γ 2.44
Gln-17	7.92	4.22	2.04	γ 2.36
His-18	8.31	4.63	3.23, 3.13	δ 7.32; ϵ 8.61

^a Chemical shifts in ppm (parts per million) indirectly referenced to TSP [sodium (trimethylsilyl)propionate] using the water peak at 4.85 ppm. ^b Most resonances were previously assigned by Wemmer & Kallenbach (1983). ^c Resonance that could not be assigned in the spectrum.

spectral width of 5000 Hz, and 1024 data points giving 5 Hz digital resolution in t_2 . Suppression of the residual water peak was achieved by using continuous low power irradiation during the relaxation delay for the D_2O spectra and during both the relaxation delay and the mixing time for the H_2O spectrum. A skewed sine bell apodization (Hare et al., 1985) was used in both dimensions with a skew of 0.6 and a phase shift of 50° . Data were zero filled in the second dimension to yield 1024×1024 point real matrices. The first row of the (t_1, ω_2) matrix was multiplied by 0.5 before the t_1 transform to suppress t_1 ridges (Otting et al., 1986). One D_2O buildup curve was done at $30^\circ C$ with mixing times of 150, 300, and 450 milliseconds. NOE intensities were significantly higher upon lowering the temperature to $5^\circ C$, due to the lengthening of the correlation time, moving away from $\omega\tau_c \approx 1$. Mixing times of 200 and 450 milliseconds were used at this lower temperature. The $5^\circ C$ spectra also had the advantage that the residual water peak moved down field of all of the α -protons. One $30^\circ C$ 90% H_2O , 10% D_2O NOESY was collected with a mixing time of 450 milliseconds. A recycle delay of 2 seconds was used for all spectra. All spectra were obtained from an approximately 20 mM sample of apamin at $pH=2$ in 400 microliters of either D_2O or 90% H_2O , 10% D_2O to include amide resonances. All data processing was done using the FTNMR program (D. Hare, unpublished) on either a Micro Vax II or a Vax 11/785 computer. NOE cross peaks on both sides of the diagonal were integrated using the volume integration routine in FTNMR. The average intensity of the cross peaks on either side of the diagonal was used except when a streak was present, in which case the peak without the streak was used. Most cross peak intensities were calibrated using the geminal distance between the β -protons of residues 3 and 7, assuming a distance of

1.75 Å between them. Some methyl to proton NOEs were calibrated using α -proton to methyl intraresidue NOEs of Ala 5 and Ala 12, assuming a distance of 2.38 Å between them. The uncertainty in all calculated distances was ± 0.25 Å, estimated from uncertainties in slope of the mixing time dependence of NOEs. Figure 2.5 shows sample buildup curves for both short and long distance NOEs. Distances derived from the water spectrum were treated as semi-quantitative: 1.9-2.5 Å for strong cross peaks, 1.9-3.0 Å for medium cross peaks, and 2.5-3.5 Å for weak cross peaks. The maximum proton-proton distance which had an observable cross peak was < 3.5 Å, based upon NOE intensities of intraresidue α H to NH peaks for Asn 2 and Cys 3 which are in an extended conformation. This is confirmed by the absence of interresidue NOEs between the α -proton and amide proton of residues in the helix (the maximum value of $d_{\alpha N}$ is 3.5 Å). Stereo specific assignments were used for the side chain protons of Pro 6. The assignments were established using the short distances between the protons on either side of the ring, beginning with the α H. The α H is closer to β H than to β' H and is also close to γ H; γ H is closer to δ H than to δ' H. This puts α H, β H, γ H, and δ H on the same side of the proline ring, and with a similar argument for β' H, γ' H, and δ' H on the opposite side.

Distance Geometry. The NMR constraints used are given in figure 2.3. NOE buildup rates were determined to obtain more accurate distances than would be obtained from a single NOESY spectrum, and to determine whether spin diffusion was occurring (Kumar et al., 1981). A sample NOESY spectrum is shown in figure 2.4. Although relatively long mixing times were used, the small size of apamin makes its rotational correlation time quite short

Figure 2.3.1
Distance Constraints for Apamin Obtained from NOESY Spectra

Sequential backbone		Other	
Cys 1			
αH[1]-NH[2]	1.90 3.00		
Asn 2			
αH[2]-NH[3]	1.90 3.00	NH[2]-βH[2]	1.90 3.00
		NH[2]-β'H[2]	1.90 3.00
		αH[2]-βH[2]	2.41 2.86 ¹
		αH[2]-β'H[2]	2.41 2.86 ¹
Cys 3			
NH[3]-NH[4]	2.50 3.50	NH[3]-βH[3]	1.90 3.50
		NH[3]-β'H[3]	1.90 3.50
		αH[3]-γH[6]	2.50 6.00 ²
		αH[3]-β'H[6]	2.50 6.00
		βC[3]-βH[15]	2.20 5.45 ³
		βC[3]-β'H[15]	2.20 5.45 ³
		αH[3]-βH[15]	2.50 5.25 ²
		αH[3]-β'H[15]	2.50 5.25 ²
Lys 4			
NH[4]-NH[5]	1.90 2.50		
Ala 5			
NH[5]-δH[6]	2.50 3.50		
αH[5]-δH [6]	2.25 2.75		
αH[5]-δ'H[6]	1.90 2.46		
Pro 6			
αH[6]-NH[7]	1.90 3.00	γH[6]-αH[12]	2.55 6.10
		γ'H[6]-αH[12]	3.03 3.53
		βH[6]-αH[12]	3.04 3.54
		βH'[6]-αH[12]	2.55 6.10
Glu 7			
NH[7]-NH[8]	2.50 3.50		
Thr 8			
αH[8]-NH[9]	1.90 3.00	αH[8]-γM[8]	2.71 3.73 ⁴
βH[8]-NH[9]	1.90 3.00	βH [8]-NH[10]	1.90 3.00
γM[8]-NH[9]	1.95 5.20 ⁵		
Ala 9			
NH[9]-NH[10]	1.90 3.00	NH[9]-βM[9]	1.95 2.90 ⁴
		αH[9]-βM[12]	1.95 5.50 ⁵
		αH[9]-NH[12]	2.50 3.50
Leu 10			
NH[10]-NH[11]	2.50 3.50		

Figure 2.3.2

Cys 11				
NH[11]-NH[12]	2.50 3.50	α H[11]- β H[14]	2.43 4.68	
		α H[11]- β' H[14]	2.43 4.68	
		α H[11]-NH[14]	2.50 3.50	
Ala 12				
NH[12]-NH[13]	2.50 3.50	NH[12]- β M[12]	1.95 3.00 ⁴	
		α H[12]- β H[15]	2.55 4.80	
		α H[12]- β' H[15]	2.55 4.80	
Arg 13				
NH[13]-NH[14]	2.50 3.50	α H[13]-NH[16]	2.50 3.50	
Arg 14				
NH[14]-NH[15]	2.50 3.50			
Cys 15				
NH[15]-NH[16]	2.50 3.50			
Gln 16				
NH[16]-NH[17]	2.50 3.50	α H[16]- β H[16]	2.44 2.87 ¹	
		α H[16]- β' H[16]	2.44 2.87 ¹	
Gln 17				
α H[17]-NH[18]	1.90 3.00			
Hydrogen Bonds		Disulfide constraints ⁶		
NH[4] δ O[2]	2.00	β C[1]- β C[11]	3.72 3.99	
NH[5] O[2]	2.00	β C[3]- β C[15]	3.72 3.99	
NH[12] O[8]	2.00			
NH[13] O[9]	2.00			
NH[14] O[10]	2.00			
NH[15] O[11]	2.00			
NH[16] O[12]	2.00			
NH[17] O[13]	2.00			

All distances in Å.

Atom names: α H= α -proton, NH=amide proton, β H and β' H= β -protons, γ H and γ' H= γ -protons, δ H and δ' H= δ -protons, γ M and β M=methyl pseudo atoms, β C= β -carbon, and O=carbonyl oxygen.

Figure 2.3.3

Note: Not all NOESY peaks have a constraint listed since the resulting constraint covers the entire range of possible values. Constraints with amide protons are broken up into three categories: strong 1.90-2.50 Å, medium 1.90-3.0 Å, and weak 2.5-3.5 Å. Van der Waals distance between two protons is 1.90 Å and between a methyl group and a proton is 1.95 Å.

¹largest possible value.

²semiquantitative.

³constraint to βC is used since the NOE is to one β -proton. Upper bound= NOE distance constraint + $\beta\text{H}-\beta\text{C}$ bond length (1.0 Å).

⁴NOE distance constraint referenced to $\alpha\text{H}-\beta\text{M}$ distance from Ala 5 and Ala 9.

⁵treated as pseudo atom constraint: added 1.0 Å to upper bound (Wüthrich et al., 1983).

⁶distance constraint to give a $\chi(\beta\text{C}-\text{S}-\text{S}-\beta\text{C})$ angle of $\pm(90^\circ \pm 10^\circ)$ (Richardson, 1981).

Figure 2.4 Upfield region of a NOESY spectrum of apamin. Mixing time of 450 ms, pH=2, and 30°C. Some important long range NOEs are labelled.

and thus the NOE buildup rather slow (see figure 2.5), minimizing the importance of spin diffusion except for the longest mixing times at low temperature.

Several specific hydrogen bonds were defined corresponding to regions of defined secondary structure in apamin (figure 2.1). Apamin's small size allows solvent accessibility to most of its residues and therefore slow amide exchange is most likely due to hydrogen bonding and not from the amide being buried within the peptide (Englander & Kallenbach, 1984). The amides of residues 4, 5, 8, and 11 through 16 are relatively slowly exchanging (Bystrov, et al., 1980; Wemmer & Kallenbach, 1983; Dempsey, 1986). Residue 5's amide hydrogen bonds with the backbone carbonyl of 2 forming a type I β -turn (figure 2.6). This is supported by the NOE connectivity pattern NH - NH - NH for residues 3, 4 and 5 (figure 2.7), by the slow exchange of residue 5's amide, and the values of coupling constants for these residues. Residue 4's amide is in a position where it can form a hydrogen bond with the side chain carbonyl of Asn 2, which would stabilize the β -turn (figure 2.6). The preference for asparagine at this position in a β -turn, and its ability to hydrogen bond to the backbone amide are well known (Richardson, 1981). This hydrogen bond has little effect on the position of the backbone atoms, and if not included the side chain of Asn 2 would take on a more random conformation. Studies on the apamin hybrids, described in the next chapter, support the existence of this hydrogen bond. The amide protons of residues 12 through 17 are in an α -helix, which frays significantly beyond residue 15. The connectivities for residues 9 through 17 are all NH to NH (figure 2.7) and NH to β H type,

Figure 2.5 Buildup curves for selected NOEs in apamin. A: fixed short distance NOEs, βH to $\beta'\text{H}$, 1.75Å and βH to methyl in Ala, 2.38Å. B: much longer distance NOEs. Note the large difference in vertical scales.

Figure 2.6 The N-terminal β -turn in apamin. The lines represent hydrogen bonds and the arrows represent the observed NH to NH NOEs. The alpha carbons of each residue are labelled using the standard one letter code.

Figure 2.7 Amide to amide region of a NOESY spectrum of apamin. Mixing time of 450 ms, pH=2, and 30°C. The sequential NOEs are labelled using the one letter amino acid code.

consistent with a helix. This helix is further confirmed by tertiary NOE contacts of the $\alpha\text{H}(i)$ to $\beta\text{H}(i+3)$ and $\alpha\text{H}(i)$ to $\text{NH}(i+3)$ types (Wüthrich et al., 1984). The observed NOEs were between $\alpha\text{H}(9) - \beta\text{H}(12)$, $\alpha\text{H}(10) - \beta\text{H}(13)$, $\alpha\text{H}(11) - \beta\text{H}(14)$, $\alpha\text{H}(12) - \beta\text{H}(15)$ (figure 2.4); and $\alpha\text{H}(9) - \text{NH}(12)$, $\alpha\text{H}(11) - \text{NH}(14)$, and $\alpha\text{H}(13) - \text{NH}(16)$. The connectivity between the $\alpha\text{H}(10)$ and the $\beta\text{H}(13)$ was clearly present as a shoulder on the $\alpha\text{H}(13)$ to $\beta\text{H}(13)$ cross peak in some spectra, but could not be quantified due to near degeneracy of the αH chemical shifts of residues 10 and 13. This degeneracy also makes it hard to tell if the $\alpha\text{H}(10)$ to $\text{NH}(13)$ NOE is present. Although the exchange of the amide protons of residues 8 and 11 is somewhat slower than amide protons which are fully exposed to solvent, there were no clear hydrogen bond acceptors which could be identified, and hence no hydrogen bond constraints were used for these residues. The amide of residue 11, however, appears to be hydrogen bonded to the side chain of Thr 8 in apa-S25 and most likely does the same for apamin (next chapter). Hydrogen bonds were included for the helix and the turn as listed in figure 2.3, and were defined to be 2.0 Å from amide proton to acceptor with the donor to acceptor distances defined by the program so that the hydrogen bonds are nearly linear. There is clearly extensive fraying of the last two residues and the lack of NOE constraints leaves this region relatively poorly defined.

Twenty starting structures were generated using correlated trial distances between the upper and lower bounds. Full refinements were done with two different approaches. The first initially refined the local structure using the conjugate gradient routine on individual amino acids, and then subsequently the whole molecule using both conjugate gradient

and dynamics routines. The alternate approach refined the whole molecule at once, occasionally using local refinements to fix chirality. The dynamics refinement algorithm (also called simulated annealing) was both faster to converge, and better at getting out of local minima than the conjugate gradient routine. Both refinement algorithms gave similar final results. Of the twenty embedded structures six refined to very low total error (figure 2.8), giving total residual errors (sum of the absolute value of all deviations) of 2.68 to 3.66 Å. Figure 2.9 shows the r.m.s. differences in nonhydrogen backbone atom positions of the six structures. These r.m.s. differences do not correlate with the total error, i.e. a higher total error does not necessarily mean a larger r.m.s. difference. Approximately 24 hours of CPU time was required to refine each structure to very low residual error.

Results and Discussion

Although apamin is rather small, $mw \approx 2$ kD giving an $\omega\tau_c$ value near 1, negative NOEs are observed, and by using somewhat longer than typical mixing times quantitative estimates of distances were obtained. There is some uncertainty in these distances due to the uncertainty in slope of the mixing time dependence of NOEs. Some of the distances derived from NOE cross peaks, particularly those which were intraresidue, were not used for the refinement since after adding experimental uncertainties they did not additionally constrain the molecule. NOEs which contribute significantly to the definition of the structure include: sequential NOEs which define local backbone conformation; secondary structure dependent NOEs, such as

Figure 2.8 Stereo view of the six best distance geometry structures of apamin. The top view is all heavy atoms and the bottom view is just the backbone atoms. The N-terminal β -turn is on the right and the C-terminal α -helix is on the left. The structures are superimposed on the lowest error structure using the backbone atoms: carbonyl carbon, amide nitrogen, and alpha carbon for all residues.

Figure 2.9
Apamin Distance Geometry Structure Analysis

Dspace structure	error (Å)	RMS difference (Å) for DSPACE structure				
		2	3	4	5	6
1	2.679	0.74	0.77	0.54	0.81	0.65
2	3.662		0.86	0.77	1.15	0.75
3	2.780			0.59	0.78	0.50
4	2.681				0.76	0.48
5	2.912					0.88
6	3.357					

Error is the sum of the absolute value of all residual violations. RMS is the root mean square difference between the backbone atoms: carbonyl carbon, amide nitrogen, and alpha carbon of all the residues. The average RMS difference is 0.74 Å.

the $\alpha\text{H}(i)$ - $\beta\text{H}(i+3)$ in the helix; and tertiary NOEs between sidechains, the most significant of which for apamin is the contact between the Pro 6 sidechain and the αH of Ala 12 (figure 2.4).

The structures obtained from distance geometry, using the constraints listed in figure 2.3, are illustrated in figure 2.8. The structures were superimposed by choosing the lowest error structure as the reference, and then least squares minimizing deviations between backbone nonhydrogen atom positions. Total errors and r. m. s. differences of the six structures are in figure 2.9. None of the structures contained individual residual violations of greater than 0.05 Å for any of the constraints. It is clear that the backbone of the peptide is fairly well defined especially within the α -helix. This helix is highly constrained by the combination of NOEs, hydrogen bonds and disulfide bonds. The remaining backbone is not as precisely defined, but all of the structures have very similar structural features. There is significant variation in the geometry of the disulfide bonds, which leads to rearrangement of the position of the backbone atoms in the β -turn region as well. The degeneracy of the Cys 15 β -protons makes the NOEs to the Cys 3 α - and β -protons relatively weak constraints. The chemical shifts of the β -protons of Cys 1 and Cys 11 are so similar that even if NOEs between them were present they could not be assigned. Thus the exact conformations of the disulfides are relatively poorly determined. The positions of other amino acid side chains could not be determined accurately since there were few NOEs involving them. This is not surprising since the side chains are most likely mobile in the solution. More NOEs were seen at low temperature, possibly both because of an

increase in rotational correlation time for the whole molecule and from a decrease in conformational averaging, but all NOEs were still within single sidechains.

The structures in figure 2.8 are generally consistent with the previously proposed model (Wemmer & Kallenbach, 1983), with the largest differences occurring between the β -turn and the α -helix, residues 6 through 8. It was previously suggested that the amide of residue 8 was hydrogen bonded to the carbonyl of residue 5, explaining the slow exchange of 8. While this was stereochemically reasonable from models, the new constraints involving residues 6 and 12 cannot be met and simultaneously have this hydrogen bond be present. Attempts to include both constraints (the NH(8)-O(5) hydrogen bond was defined the same as the other hydrogen bonds) during refinement always led to much higher residual error, specifically arising from violations in this segment of the molecule. When the hydrogen bond constraint was removed, the error was significantly reduced, and was approximately equally distributed throughout the molecule. Dempsey (1986) has suggested that the amide of residue 11 is included in the α -helix hydrogen bonding at low pH, thus explaining its relatively low exchange rate. However the NOEs involving residue 7 (the hydrogen bond acceptor required to extend the helix to include the amide of 11), and the NOEs in the 6-8 loop, cannot be satisfied if this hydrogen bond is present. It appears more likely that the sidechain of Thr 8 is in a position to interact with the amide of 11, and could slow its exchange. This is supported by the fact that threonines are commonly found at the beginning of helices (Richardson, 1981) and by the exchange behavior of

Cys 11, which behaves differently at higher pH from the rest of the residues in the helix (Dempsey, 1986). In the case of the apamin hybrids discussed in the next chapter, there are several NOEs supporting the existence of this hydrogen bond. On the other hand, some interaction with the sidechain of Glu 7 could explain the exchange behavior of this residue. Dempsey also suggested that a different structure exists at higher pH (pH > 3.5). I tested this idea by running a NOESY at pH 4 under similar conditions as those used at pH 2 (30°C, τ_m =450ms). The sequential connectivity pattern was the same and the side chain of Pro 6 was still in close proximity to α H of Ala 12. This suggests that the structure is generally the same as at pH 2 and no major structural change occurs at high pH. However, the amide protons of Glu 7 and Asn 2 shift downfield dramatically at the higher pH, 1.37 ppm and 0.98 ppm respectively, indicating some sort of change in their environment. The salt bridge proposed by Okhanov and coworkers (1980) between the sidechain carboxylate of Glu 7 and the α -amino group could be present in the calculated structures since the two groups are in fairly close proximity. However there are not a sufficient number of NOEs involving protons on the sidechain of Glu 7 and Cys 1 to definitively demonstrate that it is present. The strongest evidence for it is the titration behavior of amide chemical shifts for residues 2 and 7, as described by Bystrov et al. (1980).

I have also compared the distance geometry structures with the models based on energy refinement proposed by Freeman et al. (1986) (see figure 2.10). The coordinates for these models were calculated from the reported angles. These structures were developed using energy refinement of a

Figure 2.10 Comparison of energy refined structures (Freeman et al., 1986) to the lowest error DSPACE structure. Hider and Ragnarsson's (1980) model is closest to the DSPACE structures and is superimposed over the best DSPACE structure (bottom) using the backbone atoms of residues 9 through 15 (α -helical region).

number of different starting models. The fact that these models converge locally, but do not evolve into the same structure is a demonstration of the well known problem of finding the global energy minimum in such calculations. Among the structures proposed, the refined model of Hider and Ragnarsson (1980) is in best agreement with the distance geometry structure (3.42 ± 0.18 Å r.m.s. difference between backbone atom positions), but several differences remain. The largest differences occur in N-terminal half of the molecule. Their model has a turn which involves residues 1 through 4, instead of 2 through 5, and then an extended region which puts the Pro 6 side chain too far away from the α H of Ala 12 to be consistent with our data. The C-terminal α -helix is very similar to our structure (see figure 2.10).

Arginines 13 and 14, which have been found to be important for activity (Granier et al., 1978; Cosand & Merrifield, 1977), are both in the helix but again the lack of NOEs involving sidechain protons made it impossible to determine the exact conformation of their side chains. In fact the lack of NOEs strongly suggests that both sidechains are mobile in solution. The lack of activity for the reduced peptide with blocked disulfides (Vincent et al., 1975) suggests that some aspect of the positioning of these groups is important.

The distance geometry models were also checked against the previously reported $^3J_{\text{NH}-\alpha\text{H}}$ and $^3J_{\alpha\text{H}-\beta\text{H}}$ coupling constants (Bystrov et al., 1980). This was done for $^3J_{\text{NH}-\alpha\text{H}}$ by defining distance constraints between the amide- and α - protons, and between the amide-proton and carbonyl-carbon

consistent with the ϕ angles derived from the Karplus relationship (figure 1.2; eq. 1-3). The distance geometry structures which had been previously refined without these constraints essentially did not change upon their addition. There were only very minor refinements required to lower the residual error to its previous value. This is not surprising since the sequential assignments define helical (small $^3J_{\text{NH}-\alpha\text{H}}$) and extended regions (large $^3J_{\text{NH}-\alpha\text{H}}$) very well making coupling constant information redundant. The $^3J_{\alpha\text{H}-\beta\text{H}}$ coupling constant was only useful for residues 2, 3, and 6 since the β -protons of all the other residues are degenerate or near degenerate. Distances were calculated between the α -proton and β -protons using a Karplus relationship (De Marco et al., 1978) and all were found to be consistent with the J values.

As has been found with other distance geometry structures of peptides, it is the NOEs between residues far apart in the sequence that have the largest effect in determining the global fold of the peptide (Havel & Wüthrich, 1985). The secondary structure of peptides can be determined quite accurately using sequential assignment patterns and observation of slowly exchanging amides (Wüthrich, 1986; Wüthrich et al., 1984; Wemmer & Kallenbach, 1983). Information from coupling constants has little effect on the global structure, but can be used to help define local structure. In this case, as has been found for other proteins, it is unlikely that more accurate distances would significantly improve the structure. Havel and Wüthrich (1985) have found that the global fold of BPTI can be defined using a large number of relatively inaccurate medium and long range distance constraints. However in this case, the small size of apamin means

that there are relatively few sidechain contacts, probably all have already been identified.

The structure of apamin defined here through extensive NMR measurements, and distance geometry calculations should be quite useful in testing a variety of modelling calculations, understanding the amide exchange from segments of secondary structure, and in understanding the stabilization of such structures by disulfide bonds.

Chapter III

Analogues of Apamin and the Structure of Apa-S25

Apamin's highly stable structure is due primarily to the presence of its two disulfide cross links. Proteins which have Cys positions conserved often have the same disulfide pattern and similar tertiary structure. Two peptides were designed which take advantage of apamin's stable framework. A combination of apamin's N-terminal β -turn and its disulfide pattern is used to stabilize the α -helix from the S-peptide of ribonuclease-A (RNase-A): This peptide was chosen since it is known to form a partial helix in solution and is well characterized. The hybrid peptide I concentrated on was called apa-S25, which is a combination of apamin and the full S-peptide. The second peptide (apa-S) has only the N-terminal 13 residues of S-peptide, to be the same length as apamin. The three sequences are shown in figure 3.1 together with the S-peptide. In this chapter I describe the detailed three dimensional structure of apa-S25 and compare it to apamin's structure.

RNase-A when cleaved with subtilisin forms two peptides: a 20 residue N-terminal peptide called the S-peptide and the C-terminal S-protein. The S-peptide forms a partial helix in aqueous solution comprised of residues 3 through 12 (Kim & Baldwin, 1984). The average length of α -helices in natural proteins is 11 residues, whereas in synthetic polymers the average length is closer to 100 residues (Schultz & Schirmer, 1979). This suggests that there is a helix stop signal present in proteins. Kim and Baldwin (1984) proposed that a helix stop signal exists in the S-peptide at approximately residue 14 (residue 19 using numbering from figure 3.1). In

	1	5	10	15	20	25																				
Apamin	C	N	C	K	A	P	E	T	A	L	C	A	R	R	C	Q	Q	H								
Apa-S25	C	N	C	K	A	P	E	T	A	A	C	K	F	E	C	Q	H	M	D	S	S	T	S	A	A	
Apa-S	C	N	C	K	A	P	E	T	A	A	C	K	F	E	C	Q	H	M								
RNase-A																										

Figure 3.1 Sequences of apamin, the two hybrids, and the S-peptide (the first 20 amino acids of RNase-A). The standard one letter code was used to label the amino acids. The bars indicate the disulfide bonds present in apamin and the two hybrids.

the crystal structure of RNase-A (Wlodawer et al., 1982), the amide of Asp 14 hydrogen bonds to the carbonyl oxygen of Val 47, and therefore is not in the helix. Even though Asp 14 can participate in the helix in the isolated S-peptide, since it no longer has its natural hydrogen bonding partner, the helix still terminates in the same place as the naturally occurring protein. This helix stop signal persists even in the presence of 2,2,2-trifluoroethanol, a helix stabilizing reagent, where the last NH to NH NOE observed is between Met 13 and Asp 14 (Nelson & Kallenbach, 1989). Short helices are highly cooperative and effective nucleation of a helix promotes its formation. The apa-S25 hybrid provides a framework for essentially perfect nucleation of the S-peptide helix to allow study of the helix termination signal.

Mitchinson and Baldwin (1986) have studied S-peptide analogs to determine what interactions stabilize the helix. When the analogs were combined with the S-protein, analogs with increased helical stability led to a more stable complex. The melting temperature of the complex increased by up to 6°C, the association constant increased by up to 10 times that of the natural S-peptide (10^6 M^{-1}), and the endoribonuclease activity of RNase-A was completely restored. Studies have been done on apa-S25 to see if the endoribonuclease activity is restored (Storrs, unpublished).

Materials and Methods

I synthesized the two peptides on an Applied Biosystems 430A peptide synthesizer using standard t-boc chemistry. The peptides were HF cleaved

by the Applied Biosystems custom peptide synthesis group. 1.25 grams of crude apa-S25 was obtained from the synthesis. 180 mg of crude apa-S25 was placed in 100 ml of 5 mM Tris, pH 8.0, and a molar excess of DTT (1 liter flask). The peptide was allowed to air oxidize for one week at room temperature. The peptide solution was then lyophilized to 10 ml. HPLC was done on a Waters instrument using a Water's Delta Pak preparative column (C18-300 Å, 1.9 X 30 cm). A two buffer system was used: buffer A, 0.1% TFA/H₂O and buffer B, 0.1% TFA/60% CH₃CN/40% H₂O. The run started at 100% A for five minutes, followed by a linear gradient to 30% B over 10 minutes, and finally a linear gradient to 70% B over 80 minutes. The flow rate was 4 ml/min. The peptide eluted at 58% B and was the major peak detected at 235 nm. Mass spectrometry was used to confirm that the correct peptide was purified. Approximately 60 mg of purified peptide was obtained from the 180 mg of starting material. 40 mg of this material was sent to Dr. Peter Kim (Whitehead Institute) for antibody studies.

Approximately 20 mg of pure apa-S25 was dissolved into 0.5 ml of H₂O/10% D₂O, 100 mM NaCl, pH 2.0 for NMR studies. NOESY and COSY spectra were run at 30°C to obtain proton resonance assignments and NOE distance constraints for distance geometry calculations. The peptide was rapidly assigned since most of the resonances did not shift significantly from apamin's resonances and the sequential assignment pattern was nearly identical. As with apamin, the dominant sequential connectivity was amide to amide (figure 3.2), consistent with the C-terminal α -helix. Many of the same long range NOEs were observed (figure 3.3) and also the same amides were found to be slowly exchanging (figure 3.4). The sequential

Figure 3.2 Amide to amide region of a NOESY spectrum of apa-S25 at 30°C and pH=2. Cross peaks are labelled using the standard one letter code for each amino acid.

Figure 3.3 Upfield region of a NOESY of apa-S25 at 30°C and pH=2. Important NOEs are labelled. Note the lack of NOEs between the alpha and methyl protons of Ala 24 and 25, due to the motion of these residues.

Figure 3.4 Finger print region of a double quantum filtered COSY spectrum of apa-S25. Amide to alpha cross peaks present correspond to those amides which are slowly exchanging. The one letter code is used to label the peaks.

assignments, secondary structure NOEs, and the slowly exchanging amides are summarized in figure 3.5. The chemical shifts of the assigned resonances are shown in figure 3.6.

NOE constraints for distance geometry were approximated from the NOESY cross peak intensities (30°C, 450 ms). NOE cross peaks were divided into three groups: strong (1.9 to 2.5 Å), medium (1.9 to 3.0 Å), and weak (1.9 to 3.5 Å). Floating stereo specific assignments were used for nondegenerate beta protons to obtain stronger distance constraints, with the upfield proton of a geminal pair labelled with a prime. Embedding was done into four dimensions so that four dimensional refinement could be used. A newer version of DSPACE (relative to that used for apamin calculations) was used which allowed for some play in the hydrogen bond lengths, being set to 1.8 to 2.0 Å in this case. The DSPACE constraints are listed in appendix II. Twenty starting structures were generated using random trial distances and were refined using primarily the simulated annealing algorithm. The structures refined much faster than the apamin structures due to both a faster computer (Micro Vax III) and the four dimensional refinement algorithm. Initial calculations used all 25 residues of apa-S25 and it was found that residues beyond Met 18 were disordered. This is hardly surprising since there are few observed NOEs in this region. In fact, NOEs between the alpha and methyl protons of Ala 24 and Ala 25 were not observed, showing that the C-terminal residues are dynamically disordered (the correlation time of these residues is smaller than it is for the rest of the molecule). The final refinements, therefore, used only the first 18 residues of apa-S25. These structures refined to a very low residual error

Figure 3.5 The sequential assignments, slowly exchanging amides (+), and secondary structure NOEs. sequential NOE where the size is related to NOE intensity; ? possible NOE, but chemical shift degeneracy makes it uncertain; NOE to Pro δ -proton(s); and lines indicate long range NOE was observed.

Figure 3.6
Chemical Shifts of Assigned Resonances, pH 2.0, 30°C^a

residue	amide	α -H	β -H	other
Cys-1	b	4.48	2.96, 2.70	
Asn-2	9.15	4.88	3.09, 2.75	δ N7.77, 7.19
Cys-3	9.07	4.70	3.31, 2.72	
Lys-4	8.02	4.20	1.82, 1.82	γ 1.48, 1.39; δ 1.66, 1.66; ϵ 2.96, 2.96
Ala-5	7.26	4.52	1.15	
Pro-6	---	4.70	2.02, 1.63	γ 2.07, 2.02; δ 3.54, 3.40
Glu-7	8.99	4.44	2.29, 2.11	γ 2.60, 2.55
Thr-8	7.57	4.63	4.63	CH ₃ : 1.27
Ala-9	9.11	4.01	1.46	
Ala-10	8.76	4.22	1.42	
Cys-11	7.91	4.72	3.13, 2.75	
Lys-12	8.63	3.73	1.89, 1.79	γ 1.39, 1.24; δ 1.53, 1.53; ϵ 2.83, 2.83
Phe-13	8.01	4.29	3.23, 3.20	δ 7.22; ϵ 7.22; ζ 7.06
Glu-14	8.60	4.01	2.36, 2.28	γ 2.69, 2.56
Cys-15	8.64	4.34	2.99, 2.94	
Gln-16	7.68	4.13	1.94, 1.87	γ 2.41, 2.24; ϵ N7.46, 6.87
His-17	7.77	4.54	3.27, 2.77	δ 6.97; ϵ 8.30
Met-18	7.92	4.40	2.08, 2.04	γ 2.05, 2.05; CH ₃ : 2.05
Asp-19	8.46	4.73	2.97, 2.84	
Ser-20	8.26	4.46	3.91, 3.87	
Ser-21	8.39	4.51	3.92, 3.88	
Thr-22	8.18	4.38	4.26	CH ₃ : 1.18
Ser-23	8.28	4.45	3.83, 3.83	
Ala-24	8.32	4.31	1.36	
Ala-25	8.21	4.27	1.38	

^a Chemical shifts in ppm (parts per million) indirectly referenced to TSP [sodium (trimethylsilyl)propionate] using the water peak at 4.80 ppm. ^b Resonance that could not be assigned in the spectrum.

in about 6 hours of CPU time. Eight structures refined to very low total error and contained no residual violations greater than 0.2 Å.

Results and Discussion

The apa-S25 structures generated are better than the apamin structures discussed in the previous chapter. This is due to two factors: First, the NOEs from apa-S25 were more intense than apamin's due primarily to the increase in correlation time (larger molecular weight), and second, many chemical shift degeneracies, which occurred for apamin, disappeared. This also led to more observable NOEs, and hence more distance constraints. The sequential connectivity pattern is essentially the same as that found for apamin (figure 2.1 & figure 3.5). The helix is formed, as expected, and terminates at Ser 20 (the last NH to NH NOE). The helix stop signal appears to be spread over four or five residues, indicated by the decrease in intensity of the d_{NN} NOE and increase in the $d_{\alpha N}$ NOE in the region Gln 16 to Ser 20 (figure 2.1). The helix frays significantly beyond residue 20, confirmed by the lack of observed NOEs between Ala alpha and methyl protons of residues 24 and 25. The lack of NOEs is due to the decrease in the correlation time for those residues relative to the rest of the molecule. The structures generated from DSPACE are shown in figure 3.7 and the residual error and r. m. s. differences between backbone atom positions are shown in figure 3.8. Figure 3.9 shows the deviations for each residue to give some feel for what regions of the molecule are well defined. The largest deviations are at the ends of the molecule and for the side chain of Lys 4. This is hardly surprising since there are very few NOEs constraining these

Figure 3.7 Stereo view of the eight DSPACE structures. Top, showing all the atoms except hydrogen; and bottom, the backbone heavy atoms. The N-terminal β -turn is on the right and the C-terminal α -helix on the left.

Figure 3.8
Apa-S25 Distance Geometry Structure Analysis

Dspace structure	error (Å)	violations (Å)		RMS difference (Å) for DSPACE structure							
		<0.1	>0.1	2	3	4	5	6	7	8	
1	7.908	18	2	0.29	0.39	0.23	0.45	0.28	0.91	0.43	
2	4.707	0	0		0.25	0.31	0.44	0.25	1.01	0.37	
3	4.866	0	0			0.36	0.44	0.32	0.98	0.32	
4	8.189	20	4				0.49	0.24	0.98	0.36	
5	7.145	8	1					0.55	0.79	0.56	
6	6.105	9	2						1.06	0.34	
7	4.951	0	0							1.12	
8	7.095	14	2								

Error is the sum of the absolute value of all residual violations; violations, number of violations <0.1 Å and ≥ 0.1 Å; and RMS, difference between backbone atoms (alpha carbon, amide nitrogen, and carbonyl oxygen) of residues 1 through 18. Average RMS difference is 0.52 Å and there are no violations > 0.16 Å.

Apa-S25 Deviations

Figure 3.9 Standard deviation from the average structure for each residue. Left hand scale for all atoms in a residue and the right scale for the α -carbons. Note the large difference in scales.

residues. The hydrogen bonds between the side chain carbonyl oxygen of Asn 2 and the amide proton of Lys 4 and between the side chain hydroxyl group of Thr 8 and the amide proton of Cys 11 were confirmed by several NOEs which were not observed in apamin. NOEs were observed between the side chain amide protons of Asn 2 and both the β -protons of Lys 4 and the methyl protons of Ala 5, positioning the side chain carbonyl of Asn 2 such that it can hydrogen bond with the amide proton of Lys 4. NOEs were also observed between the α -, β -, and methyl protons of Thr 8 and several protons on residues 9 through 11, positioning Thr 8's side chain hydroxyl group such that it can participate in a hydrogen bond with the amide proton of Cys 11. The Cys 3 to Cys 15 disulfide has a preferred conformation in apa-S25 whereas it was more random in apamin (figure 3.10). The backbones of apamin and apa-S25 were superimposed using the first 16 residues as is shown in figure 3.10 and the r. m. s. values are in figure 3.11.

The apa-S25 structures generated have an average r. m. s. difference of 0.52 Å between non hydrogen backbone atom positions. This is a little better than apamin which had a 0.74 Å average difference. The structures of apamin and apa-S25 are very similar with an average r. m. s. value of 1.04 Å. In fact, structure 7 of apa-S25 is closer to the apamin structures than to the other apa-S25 structures.

The structure of apa-S25 was superimposed on the S-peptide part of the RNase-A crystal structure. The helices superimposed well and the N-terminal β -turn region is away from the rest of the protein. Complementation studies with the S-protein were done by Richard Storrs.

Figure 3.10 Comparison of apa-S25 and apamin backbones. Top, six apamin structures; middle, eight apa-S25 structures; and bottom, two best apamin on two best apa-S25. Superimposing carbonyl carbon, alpha carbon, and amide nitrogen of residues 1 to 16. Disulfide bonds are also pictured.

Figure 3.11
RMS Difference Between Backbone Atoms
of Apamin and Apa-S25

		RMS difference (Å) superimposing C, C α , and N of residues 1 - 16							
apa-S25:		1	2	3	4	5	6	7	8
apamin									
1		1.01	1.09	1.06	1.05	0.91	1.15	0.90	1.06
2		1.36	1.47	1.43	1.42	1.27	1.53	0.90	1.48
3		0.77	0.95	0.91	0.82	0.86	0.93	0.71	0.96
4		0.96	1.09	1.07	1.03	0.92	1.13	0.70	1.09
5		0.84	0.92	0.90	0.84	0.98	0.94	1.14	0.82
6		1.01	1.16	1.12	1.08	0.96	1.18	0.70	1.18

It was found that the endoribonuclease activity of S-protein was restored by addition of the peptide.

In collaboration with Dr Peter Kim, antibodies were raised against apa-S25. Preliminary results show that these antibodies recognize both apa-S25 and RNase-A. Peptide toxins, in general, possess a high stability *in vivo* so that they can retain activity under a variety of conditions. Generating peptide hybrids using a peptide toxin framework could be used to develop new vaccines. This could be done by transferring the desired part of a pathogen protein sequence into a peptide toxin hybrid. The hybrid has the stability to elicit an immune response *in vivo* before degradation. This would minimize the amount of material needed for a vaccine.

Chapter IV

Sea Anemone Peptides:

NMR Assignments and Secondary Structure

Sea anemones produce a variety of small peptide toxins to use as part of their feeding and powerful defense systems (Beress, 1982). These toxins are generally around 50 amino acids in length and have several conserved residues, including all cystine residues (see figure 4.1). Sea anemone toxins are known to bind to Na^+ channels and alter their ion conducting characteristics by slowing down the inactivation step and hence prolonging the action potential (Romey et al., 1976). Because of their specificity and high affinity for Na^+ channels, these toxins have been used as tools in the study of excitable membranes (Rathmayer, 1979). Determination of structures for these toxins would help to further characterize conformational features required for binding to the Na^+ channels. Recently, resonance assignments and the secondary structures of several sea anemone toxins have been determined using two dimensional nuclear magnetic resonance spectroscopy (2D-NMR). Gooley and Norton studied toxin I from *Anemonia sulcata* (ATX I) (Gooley et al., 1984; Gooley & Norton, 1986b) and Anthopleurin-A from *Anthopleura xanthogrammica* (AP-A) (Gooley and Norton, 1985; 1986a). More recently Widmer and coworkers further purified ATX I and obtained complete proton NMR assignments for ATX Ia (Widmer et al., 1988). Our laboratory, on the other hand, is studying the structures of toxin II (Rp II) (Wemmer et al., 1986) and toxin III (Rp III) from the sea anemone *Radianthus paumotensis* (Pease et al., 1989).

Figure 4.1 Sequences of Anemone Toxins

Radianthus Paumotensis II (Rp II)	-ASCKCDDDGPDVRSATFTGTVDFTWN--CNEGWEKCTAVYTPVASCCRKKK
Radianthus Paumotensis III (Rp III)	-GNCKCDDEGPNVRTAPLTGYVDLGY--CNEGWEKCASYYSPIAECCRKKK
Stichodactyla helianthus I (Sh I)	-AACKCDDEGPDIRTAPLTGTVDLGS--CNAGWEKCASYTTIADCCRKKK
Anemonia Sulcata Ia (ATX Ia)	GAACLCKSDGPNTRGNSMSGTIWVF--GCPSGWNNCEGRA-IIGYCCKQ
Anemonia sulcata II (ATX II)	GVPCLCDSDGPSVRGNTLSGIIWLA--GCPSGWHNCKKHGPTIGWCCCKQ
Anthopleura xanthogrammica (AP-A)	GVSCLCDSDGPSVRGNTLSGTLWLYPGCPSGWHNCKAHGPTIGWCCCKQ

Other known variants include Radianthus macrodactylus III (differs from Rp III in N->Y at position 11); Anemonia sulcata Ib (differs from ATX Ia in A->P at position 3); Anemonia sulcata V (differs from ATX II in G->K at 37 and deletion of the C-terminal Q); Anthopleura xanthogrammica B (differs from A in S->P at 3; SV->RP at 11,12; T->I at 20; L->F at 23; T->N at 39 and Q->K at 46). References: Rp II (Wemmer et al., 1986), Rp III (Metrione et al., 1988), Sh I (Kem, 1987), ATX Ia and ATX Ib (Wunderer & Eulitz, 1978), ATX II (Wunderer et al., 1976), AP-A (Tanaka et al., 1977), Rm III (Zykova et al., 1986), ATX V (Scheffler et al., 1982), AP-B (Reimer et al., 1985).

The purification and the immunological and pharmacological properties of the toxins from *R. paumotensis* have been discussed previously (Schweitz et al., 1985). It was found that toxins from *R. paumotensis* form a distinct immunological class, even though they have significant sequence homology with the toxins from *A. sulcata* and *A. xanthogrammica* (figure 4.1). Antibodies to Rp III recognize the other *R. paumotensis* toxins but not toxins from *A. sulcata* or *A. xanthogrammica*, and conversely antibodies against toxins from *A. sulcata* (ATX II, ATX V) and *A. xanthogrammica* (AP-A, and AP-B) do not recognize the *R. paumotensis* toxins. It is also interesting that toxins from *R. paumotensis* do not affect the binding of the toxins from *A. sulcata* to the Na⁺ channels. Instead the *R. paumotensis* toxins compete with a toxin from the scorpion *Androctonus australis* (toxin AA II) for binding sites on the Na⁺ channels even though AA II has no sequence homology with the *R. paumotensis* toxins (it is longer, 64 residues vs. 48, and contains four disulfides instead of three; Rochat et al., 1972). Hopefully the determination of the structures of these toxins will lead to a better understanding of this behavior.

Our laboratory recently reported the assignment of Rp II's proton NMR spectrum (Wemmer et al., 1986). The sequential assignment pattern and the observation of certain kinds of long range NOEs (between residues far apart in sequence, Wüthrich et al., 1984) led to the determination of Rp II's secondary structure. This structure contains a core of antiparallel β -sheet connected by loops of irregular structure, with no helical content. The reported secondary structures of toxins from other sea anemones also contain the same general pattern (Gooley & Norton, 1986a; Widmer et al.,

1988). Here I describe the proton NMR assignments and the three dimensional structure of Rp III and compare its structure with these toxins.

Materials and Methods

NMR Spectroscopy. The purification (Schweitz et al., 1985) and primary sequence (Metrione et al., 1987) of Rp III have been described previously. A sample containing ca. 14 mg of peptide was dissolved in 400 ml of either D₂O or 90% H₂O/10% D₂O at pH 4.6. All two dimensional spectra were recorded on either a General Electric GN-500 or a Bruker AM-500 spectrometer operating at 500MHz ¹H frequency. COSY (Wider et al., 1984), RELAY (Eich et al., 1982; Bax & Drobny, 1985), and NOESY (Wider et al., 1984) spectra were recorded at temperatures between 10°C and 37°C to obtain sequential assignments. Different temperatures were used to shift the water resonance relative to the protein resonances so that α-proton resonances close to the water peak could be observed. Some chemical shifts in the protein change slightly with temperature, aiding assignments of near degenerate proton frequencies. All COSY spectra were recorded in magnitude mode except for one phase sensitive double quantum filtered COSY spectrum in D₂O. All RELAY spectra were done in magnitude mode with total mixing delays between 25 and 40 milliseconds. Phase sensitive NOESY spectra were used with mixing times from 100 to 300 ms. Phase sensitive spectra were obtained using time proportional phase incrementation (TPPI) (Drobny et al., 1979; Bodenhausen et al., 1984). Each spectrum had 512 t₁ points and 1024 complex t₂ data points. Data were zero filled in the t₁ dimension to yield final 1024 x 1024 point real matrices.

Spectral widths of 5952 Hz (6 Hz/ pt.) and 5000 Hz (5 Hz/ pt.) were used for H₂O and D₂O spectra respectively. Suppression of the residual water peak was achieved by using continuous low power irradiation during the relaxation delay and the mixing time (NOESY). One NOESY used a 11-echo detection pulse instead of presaturation to suppress the water resonance (appendix I). This experiment has the advantage that the alpha protons under the water resonance are not saturated, hence all amide to alpha NOEs are observed. Sine bell apodization with 0° phase shift was used in both dimensions of the magnitude COSY and RELAY spectra. A skewed sine bell apodization (Hare et al., 1985) was used in both dimensions with a skew of 0.6 and a phase shift of 30° for NOESY spectra and 10° for the double quantum filtered COSY spectrum. The first row of the (t₁, w₂) matrix was multiplied by 0.5 before the t₁ transform to suppress t₁ ridges (Otting et al., 1986). All data processing was done using FTNMR program (D. Hare, unpublished) on either a Micro Vax II or a Vax 11/785.

Results and Discussion

The assignment of Rp III's ¹H NMR spectrum was carried out using the sequential approach developed by Wüthrich and coworkers (1982; 1983). Since this process has been described in detail in the literature for general cases, and specifically for Rp II (Wemmer et al., 1986), I will not repeat the details of the assignment process here. Sample upfield COSY, RELAY, and NOESY spectra are shown in figures 4.2 - 4.4. The fingerprint region, with labeled amide to α-proton intraresidue COSY cross peaks, is shown in figure 4.5. The sequential assignments for the C-terminal residues are shown in

Figure 4.2 Upfield region of a double quantum filter COSY spectrum taken at 20°C and pH=4.6. The Val 21 spin system is traced out.

Figure 4.3 Upfield region of a double quantum filtered RELAY spectrum at 35°C and pH=4.6. Several RELAY peaks are labelled. Note the α -proton of Lys 46 and the β' -proton of Leu 17 are both shifted upfield dramatically due to ring current effects.

Figure 4.4 Upfield region of a NOESY spectrum of Rp III at 20°C and pH 4.6. Alpha to alpha proton cross sheet NOEs are labelled.

Figure 4.5 COSY of the fingerprint region of Rp III showing all of the amide to alpha intraresidue cross peaks. The spectrum was acquired at 20°C and pH=4.6. The one letter code is used to label cross peaks. Ala 34 alpha proton resonance was saturated so that no cross peak was observed. The Ala 34 cross peak is therefore indicated by the black dot.

figure 4.6 and the slowly exchanging amides were determined from figure 4.7. Figure 4.8 is the amide to amide region of a NOESY spectrum showing some important NOEs. There were no unusual problems associated with determining the assignments, and the only question remaining concerns the reported sequence (Metrione et al., 1987). The presence of a third lysine at the C-terminus was ambiguous in the chemical sequencing of both Rp II and Rp III, but the NMR data indicates that an extra residue (with a long side chain) is clearly present in this position. While the NMR cannot easily distinguish between lysine and arginine in this case, from the compositions reported (Schweitz et al., 1985) it is almost certainly lysine. A summary of the observed sequential connectivities is given in figure 4.9 and the chemical shifts of all assigned protons are listed in figure 4.10. Some of the long side chain proton resonances were unassignable due to the chemical shift degeneracy of many of these protons.

The dominant sequential connectivity is of the $d_{\alpha N}$ type, showing that the majority of residues are in an extended conformation, such as that found in regions of β -sheet (Wüthrich et al., 1984). There are only a few scattered NOEs of the d_{NN} type with no stretches of more than four consecutive residues, thus the peptide has no regular helix. The observed α -proton to α -proton NOEs, in figure 4.4, indicate the presence of the four stranded antiparallel β -sheet structure shown in figure 4.11. This sheet structure is further confirmed by the presence of several other long range NOEs, and by the presence of hydrogen bonds deduced from slow amide exchange. Slowly exchanging amides were determined by first

Figure 4.6 NOESY of the fingerprint region of Rp III showing sequential connectivities of residues 43 through 48. The spectrum was acquired at 20°C and pH=4.6. The one letter code is used to label intraresidue amide to alpha proton cross peaks.

Figure 4.7 Double quantum filtered COSY of the fingerprint region of Rp III showing the slowly exchanging amide protons. The sample was prepared as described in chapter I. The spectrum was acquired at 20°C and pH=4.6. The one letter code is used to label cross peaks.

Figure 4.8 Amide to amide region of a 11NOESY spectrum of Rp III. Mixing time of 150 ms, pH=4.6, and 20°C. Some important NOEs are labelled using the one letter code.

Figure 4.9 Summary of observed sequential NOEs and slowly exchanging amide protons (+) for Rp III. The structure is clearly extended with no regions of helix.

Figure 4.10**Chemical Shifts of Assigned Resonances, pH 4.6, 20°C^a**

residue	amide	α -H	β -H	other
Gly-1	b	3.49, 2.54		
Asn-2	8.78	4.70	3.00, 2.64	
Cys-3	7.98	4.72	2.88, 2.77	
Lys-4	8.65	4.47	1.81, 1.76	γ 1.47, 1.44
Cys-5	9.25	4.54	3.21, 2.69	
Asp-6	8.96	4.41	2.86, 2.71	
Asp-7	8.66	4.37	2.99, 2.60	
Glu-8	7.70	4.19	2.30, 1.97	γ 2.36, 2.33
Gly-9	7.94	4.34, 3.96		
Pro-10	---	4.40	2.22, 1.96	γ 2.02, 1.94; δ 3.69, 3.55
Asn-11	8.44	4.74	2.95, 2.68	
Val-12	8.24	3.93	2.15	CH ₃ : 0.96, 0.95
Arg-13	8.48	4.23	1.88, 1.88	
Thr-14	7.51	4.34	4.38	CH ₃ : 1.08
Ala-15	7.61	4.60	1.47	
Pro-16	---	4.28	2.22, 1.77	γ 2.00, 1.90; δ 3.92, 3.78
Leu-17	8.87	4.16	2.04, 0.69	γ 1.48; CH ₃ : 0.68, 0.62
Thr-18	9.86	4.34	4.31	CH ₃ : 1.15
Gly-19	8.31	4.67, 3.48		
Tyr-20	9.78	4.95	2.55, 2.55	δ 7.33; ϵ 6.90
Val-21	8.37	3.88	1.57	CH ₃ : 0.42, 0.24
Asp-22	8.98	4.79	2.55, 2.55	
Leu-23	8.45	3.98	1.54, 1.47	γ 1.49; CH ₃ : 0.86, 0.79
Gly-24	8.99	4.07, 3.03		
Tyr-25	7.86	4.52	3.10, 2.74	δ 7.10; ϵ 6.81
Cys-26	8.20	4.67	2.91, 2.55	
Asn-27	7.87	4.44	2.70, 1.94	
Glu-28	8.65	4.24	2.39, 2.09	
Gly-29	9.32	4.44, 3.71		
Trp-30	8.92	4.94	3.56, 2.93	δ 7.25; ϵ ₃ 7.06; ζ ₂ 7.36; ζ ₃ 7.36; η 6.77; HN: 10.45
Glu-31	9.64	4.92	2.03, 1.87	γ 2.22, 2.11
Lys-32	9.07	4.24	1.69, 1.64	
Cys-33	9.86	5.24	3.28, 3.00	
Ala-34	7.34	4.81	1.48	
Ser-35	9.76	4.07	3.96, 3.96	
Tyr-36	8.46	4.46	3.01, 2.64	δ 7.04; ϵ 6.63
Tyr-37	8.69	4.34	2.88, 2.78	δ 7.09; ϵ 6.68
Ser-38	7.51	4.49	4.05, 3.30	
Pro-39	---	4.63	2.59, 1.93	γ 2.08, 2.08; δ 3.96, 3.68
Ile-40	7.41	4.71	2.15	γ 1.05, 0.90; CH ₃ : β 0.85; γ 0.80
Ala-41	7.60	4.78	1.19	
Glu-42	8.25	4.73	1.92, 1.40	γ 2.54, 2.24
Cys-43	9.11	4.67	2.78, 2.31	
Cys-44	9.30	5.54	3.00, 3.00	
Arg-45	9.73	5.18	1.91, 1.82	
Lys-46	8.26	3.04	1.11, 0.37	γ 0.55, 0.25; δ 1.12, 1.12; ϵ 2.76, 2.76; HN ₃ : 7.46
Lys-47	8.19	3.91	1.69, 1.40	γ 1.28, 1.10; δ 1.54, 1.49; ϵ 2.88, 2.88
Lys-48	8.54	4.08	1.71, 1.71	γ 1.35, 1.29

^a Chemical shifts in ppm (parts per million) indirectly referenced to TSP [sodium (trimethylsilyl)propionate] using the water peak at 4.85 ppm. ^b Resonance that could not be assigned in the spectrum.

Figure 4.11 The four stranded β -sheet present in Rp III. Similar sheet structure is found in all of the sea anemone peptides studied thus far. The alpha carbons are labelled with the one letter code. The dashed lines indicate NOEs observed and the black dots indicate slowly exchanging amide protons.

lyophilizing the sample from H_2O , then redissolving it in D_2O , and immediately running a COSY experiment, shown in figure 4.7. Amide to α -proton cross peaks present in this spectrum indicate those amide protons most protected against exchange with solvent, presumably through hydrogen bonding. These include residues 20, 21, 22, 30, 31, 33, and 41 through 45 for Rp III. The cystine pairing to form the disulfides: Cys 3 with 43; Cys 5 with 33; and Cys 26 with 44, forces the β -sheet core to be highly twisted. The disulfides are presumed to be in the same location as in AP-A (Norton, 1981) and toxin II from *A. sulcata* (Wunderer, 1978) due to the conservation of cystine residues and similarity of secondary structure among these toxins. The disulfide pattern was confirmed by several NOEs. The NOE connectivity patterns in the loops connecting the β -sheet strands, comprised of residues 6-16, 24-27, and 34-42, show that there are no other regions of regular secondary structure, except for a type II β -turn involving residues 27 through 30. This turn is clearly indicated by the slow exchange of Trp 30's amide (hydrogen bonding to Asn 27's backbone carbonyl) and the NOE intensity pattern: $d_{\alpha N}(28-29)$ is strong, $d_{NN}(29-30)$ is strong, and $d_{\alpha N}(28-30)$ is weak (Wüthrich et al., 1984). The basic structural pattern of the β -sheet core and type II β -turn are present in all the other anemone toxins studied thus far: Rp II, AP-A (Gooley & Norton, 1986a), ATX Ia (Widmer et al., 1988), and Sh I (Norton et al., 1989; Fogh et al., 1989).

In comparing the secondary structures of these proteins we do find some differences in structure. The $\alpha H(2)$ to $\alpha H(20)$ NOESY cross peak is much weaker for Rp III than for Rp II, while the $\alpha H(4)$ to $\alpha H(17)$ cross peak is very weak in Rp II but is quite strong in Rp III (it was identified in Rp II

only after its clear presence in Rp III was noted). These two NOESY cross peaks in combination with an NH to NH NOE between residues 18 and 19 indicates the presence of a β -bulge in both Rp II and Rp III (figure 4.11). This region of ATX Ia also contains a β -bulge, but with the two α H - α H NOE peaks apparently of normal intensity (Widmer et al., 1988). The differences in NOE intensities in this region between Rp II and Rp III indicate that the N-terminal strand of the β -sheet in Rp III is translated quite significantly (the order of 1.5 Å) relative to its neighboring strand. This strand in ATX Ia is probably translated to a lesser extent than in Rp III. The region containing this β -bulge is quite interesting. The Gly residue at position 19 is one of very few non-cystine residues which is completely conserved among the anemone toxins, figure 4.1. This suggests that the unusual structure formed is functionally important in the protein and requires glycine at this position, probably to prevent a clash between the sidechain of this residue with the neighboring strand of β -sheet. This β -bulge does not fit any of the categories described by Richardson (1981). The residue preceding this Gly is always Thr or Ser, however a variety of residues are observed following it. Interactions of the hydrophobic sidechain of residue 17 (Leu, Phe or Met) may play a role in determining the conformation of the loop containing residues 6 to 16. In Rp II the aromatic ring protons of the Phe 17 are close to the methyls of Val 12, but in Rp III no cross peaks are seen between the methyls of Leu 17 and those of Val 12. In both peptides clear contacts between several of the other aromatic and methyl containing residues are present and very similar in intensity, showing that the overall tertiary structure is in fact quite similar. The complete three dimensional structures of these proteins have been

determined by distance geometry calculations (next chapter), where the differences in long range NOEs are analyzed.

Comparing the chemical shifts of Rp II, Rp III and ATX Ia, strong similarities were found, both for residues within the β -sheet core and in loops. This is hardly surprising for Rp II vs. Rp III since 31 of 48 amino acids are identical, however it is also true for Rp III vs. ATX Ia even though only 14 residues match. This is diagramed schematically in figure 4.12, in which the deviations in chemical shift from the random coil value are shown (Bundi & Wüthrich, 1979). There are a number of differences in aromatic residues among the three toxins Rp II, Rp III and ATX Ia. Some of the differences in chemical shift are attributable to ring currents, even though coordinates for the structure are not yet accurately known. Residue 17 is Phe in Rp II, but Leu in Rp III and Met in ATX Ia. NOEs observed between the side chain of Phe 17 and residues 4 and 14 suggest that ring currents are responsible for the upfield shifts of the α H resonances of these two amino acid residues (figure 4.12), chemical shifts are normal for the other two proteins. The residues between 22 and 25 have several changes in aromatics, which must contribute to the spread of both α H and NH chemical shifts for amino acids in this region. Similarly changes in residues near 36 and 37 may stem from changes in tyrosines: Rp III YY $\rightarrow$ Rp II VY $\rightarrow$ ATX Ia RA. The dramatic shifts (both upfield and downfield) of residues near the C terminus (44, 45, & 46) in all of the proteins stem from ring currents arising from the conserved Trp at position 30. This is demonstrated both from the secondary structure (figure 4.11) and through NOEs involving this aromatic sidechain and residue 45.

Figure 4.12 Chemical shift comparison of the alpha and amide protons from Rp II, Rp III, and ATX Ia. Plotted as the difference from random coil values (Bundi & Wüthrich, 1979).

In the case of inhibitors (of known structure) studied by Pardi et al. (1983) it was clear that conserved structural features often gave rise to similar chemical shifts, although the basis for shifts away from random coil values are uncertain. It is interesting to note the strong similarity in chemical shifts for residues 6 to 17 for all of the proteins discussed here. In this region the deviations from random coil values are all quite small, probably reflecting the lack of regular secondary structure. This suggests (though it certainly does not prove) that the structure (or lack there-of) of the 6-17 loop is really quite similar for these proteins. This is of interest since sequential connectivities show it to be irregular in secondary structure, but it contains several residues implicated as important in toxicity. Residues Asp 8 and Arg 13 (Rp II sequence) which have been found to be important for activity in other anemone toxins (Barhanin et al., 1981) are found in this loop. Arg 13 is the only completely conserved residue known to be involved in binding of toxin to its receptor. Rp III and Rm III from *Radianthus macrodactylus* have a Glu at position 8 instead of an Asp like all of the other toxins. This substitution may explain Rp III's greater toxicity than Rp II (Schweitz et al., 1985).

The resonance assignments for toxin III from the anemone *Radianthus paumotensis* were presented. From the pattern of sequential connectivities, slowly exchanging amide protons, and $\alpha\text{H}-\alpha\text{H}$ NOEs it is clear that its secondary structure is very similar to that of previously studied toxins Rp II and ATX Ia, including a twisted four stranded β -sheet, and a type II β -turn connected by loops of irregular secondary structure. The position of Rp III's outside strand of the sheet involving residues 1 through

5, with respect to the neighboring strand, appears to be intermediate between ATX Ia and Rp II even though the primary sequences of Rp III and ATX Ia are less similar than Rp II and Rp III. Distance geometry calculations are being carried out to determine the three dimensional structure of ATX Ia (Widmer et al., 1988). The three dimensional structures of the two Rp toxins are presented in the next chapter. A crystal structure of AP-A is also being determined (Smith et al., 1984). Together these structures will lead to a better understanding of the differences between the toxins and give some insight into their mode of action on Na⁺ channels.

Chapter V

Sea Anemone Peptides:

Tertiary Structure

The previous chapter described the secondary structure of the sea anemone peptides studied to date. This chapter, on the other hand, is concerned with the calculation and analysis of the three dimensional structures of Rp II and Rp III. Since the structure calculations for these peptides are much more time consuming, I concentrated on only one of these peptides. I chose Rp III since it has a better resolved NMR spectrum and it was the peptide I assigned. There are far fewer proton resonance degeneracies in Rp III, allowing use of much stronger distance constraints.

Materials and Methods

All the structures were calculated using the latest version of the DSPACE program (D. Hare, unpublished). The distance constraints were derived from both NOE and amide exchange data. For Rp III, two 11NOESY spectra in H₂O (20°C, τ_m = 150, 250 ms), and two NOESY spectra in D₂O (same conditions) were analyzed. Regular NOESY spectra were used for Rp II (25°C, τ_m = 300 ms). The NOE distance constraints were semi quantitative:

0 to 2.5 Å	strong NOE
0 to 3.0 Å	medium NOE
0 to 3.5 Å	weak NOE (short mixing time)
0 to 4.0 Å	weak NOE (long mixing time)

All slowly exchanging amides were included as hydrogen bonds, except Ala 41 since it has no clear hydrogen bond acceptor. Rp II had one additional

slowly exchanging amide, Cys 3, which was included as a hydrogen bond in the N-terminal strand of the β -sheet (figure 4.11). Both ATX Ia and Sh I have exactly the same slowly exchanging amides (Widmer et al., 1988; Norton et al., 1989). All of the hydrogen bonds defined are in the sheet region (figure 4.11) and were defined as 1.8 to 2.0 Å amide proton to acceptor. Floating stereo specific assignments were used for both peptides, with the upfield proton labelled with a prime. The initial constraint list for Rp II was prepared by Dr. Vasant Kumar and I added the floating stereo specific assignments. The disulfides were included as before (chapter I). The DSPACE constraint lists for both proteins are in appendix II.

DSPACE was used to generate 50 structures of Rp III and only the best 25 were analyzed. For Rp II, 25 structures were generated of which 11 refined to very low error. Both four dimensional embedding and four dimensional refinement were used. Each structure refinement took approximately 24 hours of CPU time on a micro Vax III.

Results and Discussion

As expected, the structures of both peptides contain a twisted, four strand antiparallel β -sheet connected by loops of irregular structure, forming the clover leaf structure pictured in figure 5.1. The backbones of both peptides are relatively well defined except for loop-I which has very few long range NOE constraints. Rp II has only one NOE between loop-I and the rest of the molecule, between Val 12 and Phe 17, whereas Rp III doesn't have any. There are 321 distance constraints and 10 hydrogen bonds for Rp III

Figure 5.1 Superposition of the backbone atoms of the four best Rp III structures. Top, superimposing carbonyl carbon, amide nitrogen, and alpha carbon of residues 1 to 5 and 17 to 48. Bottom, schematic of the backbone labelling residue positions and the three loops.

and 186 distance constraints and 11 hydrogen bonds for Rp II. This led to a better defined structure for Rp III. The two structures are very similar (see figure 5.1, 5.2, and 5.3). The distance geometry errors, residual violations, and average r. m. s. values are in figure 5.4. The standard deviations for each residue of Rp III are in figure 5.5. The peptides were analyzed in 8 segments: residues 1 to 5, 6 to 16 (loop I), 17 to 24, 25 to 32, 33 to 40, and 41 to 48 (see figure 5.6).

The first segment (residues 1 to 5) is the N-terminal strand of the β -sheet. This region is better defined for Rp II since there are NOEs between Ala 1 and Val 21 (across the β -sheet) and the amide of Cys 3 is slowly exchanging in Rp II, defining a hydrogen bond to carbonyl oxygen of Gly 19 (figure 5.6). Gly 1 of Rp III, on the other hand, has no NOEs to the neighboring strand. Also, the sequential connectivity is different in this region: for Rp II, it is all $d_{\alpha N}$ and for Rp III, it is also predominantly $d_{\alpha N}$ except for residue 2 to 3 which is d_{NN} . The second segment (residues 6 to 16) is the least well defined region of both peptides. This is unfortunate since the only residue shown to be important for binding of toxin to receptor is Arg 13 (Barhanin et al., 1981). It may be that this region needs to be away from the rest of the protein so that it can bind to its receptor. The third segment (residues 17 to 24) is the second strand of the β -sheet (figure 4.11). This is a very well defined region of the peptide with NOEs to many other parts of the protein. The fourth segment (residues 25 to 32) contains a Gly - Trp dipeptide conserved in all anemone toxins studied thus far (figure 4.1). In Rp III, the Trp 30 indole ring has a well defined conformation (positioned in the middle of loop-II). NOEs are observed

Figure 5.2 Stereo view of the backbone atoms of the four best Rp II structures. Superimposing same residues and atoms as in figure 5.1.

Figure 5.3 Stereo view of the two best Rp II structures superimposed on the two best Rp III structures using the same atoms and residues as in figure 5.1. All structures are superimposed on the lowest error Rp III structure.

Figure 5.4
Distance Geometry Structure Analysis

RpII				RpIII			
Mol.	Error	0.1-0.3	0.3-0.5	Mol.	Error	0.1-0.3	0.3-0.5
1	11.57	3	0	1	32.22	9	2
2	11.60	2	0	2	38.69	14	0
3	11.74	2	0	3	39.49	15	0
4	12.46	2	0	4	40.03	17	0
5	13.78	3	0	5	41.38	18	0
6	13.83	2	0	6	41.72	16	1
7	16.34	7	1	7	42.41	22	2
8	16.51	5	0	8	43.17	16	1
9	17.59	10	0	9	43.43	26	0
10	18.14	4	0	10	43.53	20	0
11	18.99	10	0	11	45.43	25	2
				12	45.71	24	2
				13	45.72	23	0
				14	46.05	23	0
				15	46.38	19	0
				16	46.54	22	2
				17	47.45	20	2
				18	48.01	26	0
				19	48.29	28	2
				20	48.50	30	0
				21	48.77	32	0
				22	48.94	26	1
				23	49.08	10	0
				24	50.04	22	2
				25	51.69	30	2

Average	
	RMS Difference
Rp II	2.22 Å
Rp III	1.33 Å
Both	2.41 Å

Mol., distance geometry structures listed in order of increasing total error; Error, the sum of all residual violations in Å; 0.1-0.3, number of violations between 0.1 and 0.3 Å; 0.3-0.5, number of violations between 0.3 and 0.5 Å. Also shown is the average RMS difference superimposing the backbone atoms of residues 1-5 and 17-48 (all residues except those in loop I) of Rp II, Rp III, and Rp II on Rp III (Both).

Rp III Deviations

Figure 5.5 Standard deviation from the average structure for each residue in Rp III. Left hand scale for all atoms in a residue and right hand scale for the alpha carbons. Note that the loop I region (residues 6 to 16) is the least well defined.

Figure 5.6
Regions of Rp II and Rp III

Region	Average RMS Difference (Å)	
	Rp II	Rp III
1 to 5	0.85	1.07
6 to 16	3.60	3.59
17 to 24	1.82	1.01
25 to 32	1.14	1.07
33 to 40	1.84	1.15
41 to 48	1.45	1.27

Superposition of backbone atoms of specific regions of Rp II and Rp III. The carbonyl carbon, carbonyl oxygen, amide nitrogen, α -carbon, and β -carbon of each region were used for superimposing. The ten best Rp II and ten best Rp III structures were analyzed. The next few pages show the local superpositions for each of these regions.

Rp III residues 1 to 5. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp III residues 6 to 16. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp III residues 17 to 24. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp III residues 25 to 32. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp III residues 33 to 40. Top, all atoms except hydrogens and bottom, backbone atoms.

48

Rp III residues 33 to 40. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 1 to 5. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 6 to 16. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 17 to 24. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 25 to 32. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 33 to 40. Top, all atoms except hydrogens and bottom, backbone atoms.

Rp II residues 41 to 48. Top, all atoms except hydrogens and bottom, backbone atoms.

between Trp 30 and residues 20, 25, 27, 28, 29, 31, 44, 45, 46, and 47, defining the position of the indole ring. In Rp II the position of Trp 30 is less well defined due to chemical shift degeneracies (less strong NOE constraints). The fifth segment (residues 33 to 40) is loop-III. This region is less well defined. The final segment (residues 41 to 48) is the final strand of the β -sheet and is thus fairly well defined except for the final two residues.

The three dimensional structures of Rp II and Rp III show that the overall fold is essentially the same. Small changes in local conformation, such as the positioning of the N-terminal strand of the β -sheet, may lead to large changes in potency. Rp III, for example, is up to 100 times more potent than Rp II (Schweitz et al., 1985), even though the overall fold is essentially the same. Alternatively, these changes may be due to differences in sequence, as was discussed in the previous chapter. Once the structures of the scorpion toxin AA II and of the other sea anemone toxins become available differences in binding and potency may become more clearly understood.

Closing Remarks

The combination of nuclear magnetic resonance spectroscopy and distance geometry techniques provide a powerful and rapid method for determining peptide structures. The techniques are extremely useful for studying families of peptides since the NMR spectra are all similar. Comparing individual NOE intensities provides a method for analyzing small changes in local conformation, whereas the distance geometry algorithm allows one to analyze the overall fold of peptides. Peptide synthesis in combination with these other techniques gives one the opportunity to study various aspects of peptide folding. In the future these techniques will be applied to much larger peptides if the NMR spectrometers and computers continue to improve as they have over the past decade.

Appendices
and
References

Appendix I

NMR Pulse Programs Written for General Electric GN500

NCOSY

Magnitude N-type COSY with presaturation
during the recycle delay.

PULSE SEQUENCE:

D5,D,C,L1,G

P2,G Phase: A

D+8,G

P2,G Phase: B

A,G Phase: C

D6 Jump to # 1

PARAMETERS:

QP+Ø

D8 = 5 µsec

NA = 16*

DG = 2 or more

L1 = Ø to 45

AB-

I8 = DW

PHASES:

A = S/4

B = S + (S/4)

C = (2*S) + (S/4)

MOD 4

MOD 4

MOD 4

DQFCSY

Phase sensitive (TPPI) double quantum filtered COSY with
presaturation during the recycle delay.

PULSE SEQUENCE:

D5,D,C,L1,G
P2,G Phase: A
DW/2
D+8,G
P2,G Phase: B
D9,G
P2,G Phase: C
A,G Phase: D
D6 Jump to # 1

PARAMETERS:

QP+Ø AB-
D8 = 5 µsecs l8 =

D9 = 5 µsecs
NA = 16°
L1 = Ø to 45

PHASES:

A = 1 + S + (S/4) + #	MOD 4
B = S + (S/4)	MOD 4
C = S/4	MOD 4
D = (2*S) + (S/4)	MOD 4

NOTE: This experiment works best with a long relaxation delay. All peaks will
phase absorptive and antiphase. Acquire a reference spectrum for phasing
parameters. The first block should have little or no signal.

Reference: Rance et al., 1983.

RELAY

Magnitude homonuclear relay with presaturation
during the recycle delay.

PULSE SEQUENCE:

D5,D,C,L1,G Dec Phase: +1

P2,G Phase: A

D+8,G

P2,G Phase: B

D2,G

PC4,G Phase: B

D2,G

P2,G Phase: C

A,G Phase: D

D6 Jump to #1

PARAMETERS:

QP+Ø AB-

D8 = 5 µsec NA = 32*

D2 = (mixing time)/2

L1 = Ø to 45

PHASES:

A = (S/8) MOD 4

B = (S/2) + (S/8) MOD 4

C = ((2*S) + (S/2)) + (S/8) MOD 4

D = (2*(S/2)) + (S/8) MOD 4

NOTE: The mixing time can be estimated by looking at Bax & Drobny (1985).

2QFRLY

Phase sensitive (TPPI) double quantum filtered relay
with presaturation during the recycle delay.

PULSE SEQUENCE:

D5,D,C,L1,G Dec Phase: +1
P2,G Phase: A
D+8,G
P2,G Phase: B
D10,G
P2,G Phase: C
D11,G
P4,G Phase: C
D11,G
P2,G Phase: C
D10,G
P2,G Phase: D
A,G Phase: E
D6 Jump to #1

PARAMETERS:

QP+Ø AB-
D8 = 5 µsec D10 = 5 µsec
D11 = (mixing time)/2
NA = 16*
L1 = Ø to 45

PHASES:

$A = S + 1 + (S/4) + \#$	MOD 4
$B = S + (S/4)$	MOD 4
$C = \emptyset$	MOD 4
$D = S/4$	MOD 4
$E = (2 \cdot S) + (S/4)$	MOD 4

NOTE: The mixing time can be estimated by looking at Bax & Drobny (1985).

TOCSY

Phase sensitive (TPPI) total correlation spectroscopy experiment
with presaturation during the recycle delay

PULSE SEQUENCE:

PARAMETERS:

D5,D,C,L1,G

P2,G PHASE: C

D+8,G

DO I=1, ILIM

PC4,G PHASE: A $\text{PHASE A} = (S/2) + 3 \quad \text{MOD } 4$

PC4,G PHASE: B $\text{PHASE B} = (S/2) + 1 \quad \text{MOD } 4$

PC4,G PHASE: B $\text{PHASE C} = (S/2) + (2*S) + \# \quad \text{MOD } 4$

PC4,G PHASE: A $\text{PHASE D} = (S/2) \quad \text{MOD } 4$

$\text{PHASE E} = (S/2) + (2*S) \quad \text{MOD } 4$

PC4,G PHASE: B

PC4,G PHASE: B

PC4,G PHASE: A

PC4,G PHASE: A

MLEV 17

PC4,G PHASE: B

PC4,G PHASE: A

PC4,G PHASE: A

PC4,G PHASE: B

PC4,G PHASE: A

PC4,G PHASE: A

PC4,G PHASE: B

PC4,G PHASE: B

P4,G NEXT I PHASE: D

A,G PHASE: E

D6 JUMP TO # 1

NOTE: The spectrometer needs to be reconfigured to run this experiment.

- 1) Move decoupler input to xmtr input at console.
- 2) Connect xmtr cable to attenuator; then connect attenuator output to dec input at console.
- 3) Switch xmtr and dec cables at probe interface (box next to magnet).
 - the decoupler line should now go to xmtr in.
 - the xmtr line should be connected to dec in.
- 4) Set L1 to 82; OC to LOW power (high=0, low approx 800).
- 5) Use 1PULSE experiment to measure 90° pulse.
- 6) Adjust attenuator (back of console) to give 35 μ sec 90° pulse (about 3db).
- 7) Use PRESAT to adjust water saturation. Adjust low power OC to change saturation power.
- 8) You should now be ready to go.

Note: To calculate the mixing time = (90° pulse length)*(66)*(ILIM). ILIM should be an even integer. Typical values: ILIM=32, 90° pulse length= 35 μ secs.

Other parameters: QP+0, AB-, NA=8*, DW=VDW, CB=2K

NOESY

2D phase sensitive (TPPI) NOE experiment with presaturation
during both the mixing time and recycle delay.

PULSE SEQUENCE:

D5,D,C,L1,G
P2,G Phase: A
DW/8
D+8,G
D+8,G
D+8,G
D+8,G
P2,G Phase: B
D9,D,C,L1,G
D+8,G
D9,D,C,L1,G
P2,G Phase: C
A,G Phase: D
D6 Jump to # 1

PARAMETERS:

QP+Ø AB-
D8 = 5 µsecs l8 =
D9 = (Mixing time)/2
NA = 16°
DG = 2 or more
L1 = Ø to 45

PHASES:

A= (S/4) + (2*S) + #	MOD 4
B= S/4	MOD 4
C= (S/4) + (2*(S/2))	MOD 4
D= (S/4) + (2*((S+1)/2))	MOD 4

11NOESY

2D phase sensitive (TPPI) NOE experiment with a one-one
echo detection pulse.

PULSE SEQUENCE:

D5,G
P2,G Phase: A
DW/8
D+8,G
D+8,G
D+8,G
D+8,G
P2,G Phase: B
D9,G
D+8,G
D9,G
P2,G Phase: C
D2,G
P2,G Phase: C+2
D10,G
P2,G Phase: D
D4,G
P2,G Phase: D+2
D10,G
A,G Phase: E
D6 Jump to # 1

PARAMETERS:

QP+Ø AB-
D8 = 5 µsecs l8 =

D9 = (Mixing time)/2
NA = 16*
DG = 2 or more
D10 = 125 µsecs
D2 = 1/(4*(offset)) offset = max.
P2 = 25 µsecs

PHASES:

$A = (2*(S/4)) + (S/16) + \#$	MOD 4
$B = S/16$	MOD 4
$C = 3 + (3*(S/16)) + (2*(S/8))$	MOD 4
$D = S$	MOD 4
$E = (S/16) + (2*S) + (2*(((S/4)+1)/2))$	MOD 4

Idea came from Sklenar & Bax (1987).

Appendix II

Distance Geometry Constraints

Apa-S25

Rp III

Rp II

Constraints are listed sequentially and are written in DSPACE format. There are four types of constraints which are used to help define the structure:

- 1) H-bonds - hydrogen bonds, amide proton to acceptor distance is defined.
- 2) Disulfide Constraints - distance between β carbons defined such that $\beta\text{C} - \text{S} - \text{S} - \beta\text{C}$ dihedral angle is $\pm(90^\circ \pm 10^\circ)$.
- 3) Eliminate Chirality - eliminate chirality from atom so that protons can interchange positions.
- 4) Distance Constraints - distance constraints in the form of lower then upper values. Lower bounds are always the sum of van der Waals radii.

Atom names are as follows (also see figure 1.7):

ca, cb, cg, cd, cz - α , β , γ , δ , ζ carbons.
 har, has - α protons on a Gly (r and s label chirality).
 ha - α proton all other residues.
 hn - amide proton.
 c - backbone carbonyl carbon.
 o - backbone carbonyl oxygen.
 hbr, hbs - β , β' protons.
 hgr, hgs - γ , γ' protons.
 hdr, hds - δ , δ' protons.
 hd2e, hd2z - δ protons in the ϵ and ζ positions respectively.
 mb, mg, md - β , γ , δ methyls.
 mgr, mgs, mdr, mds - γ and δ methyls (Val and Leu).
 mg2 - γ methyl on Ile.
 he3, hz2, hz3, hh2, hd1 - ϵ 3, ζ 2, ζ 3, η 2, δ 1 protons on Trp.

Apa-S25

```

;
;   H-Bonds
;
def/hbond hn[4] od1[2] 1.8 2.0
def/hbond hn[5] o[2] 1.8 2.0
def/hbond hn[11] og[8] 1.8 2.0
def/hbond hn[12] o[8] 1.8 2.0
def/hbond hn[13] o[9] 1.8 2.0
def/hbond hn[14] o[10] 1.8 2.0
def/hbond hn[15] o[11] 1.8 2.0
def/hbond hn[16] o[12] 1.8 2.0
;
;   Disulfide Constraints
;
def/bou cb[1] cb[11] 3.72 3.99
def/bou cb[3] cb[15] 3.72 3.99
;
;   Eliminate Chirality
;
def/nochir cb[1]
def/nochir cb[3]
def/nochir cg[4]
def/nochir cb[7]
def/nochir cg[7]
def/nochir cb[11]
def/nochir cg[12]
def/nochir cb[12]
def/nochir cb[14]
def/nochir cb[15]
def/nochir cb[17]
def/nochir cb[18]
;
;   Distance Constraints
;
def/bou hn[2] ha[1] 1.9 2.5
def/bou hn[2] hbs[1] 1.9 3.0
;
def/bou hn[2] hbr[2] 1.9 3.0
def/bou hn[2] hbs[2] 1.9 3.0
def/bou hn[3] ha[2] 1.9 2.5
def/bou hn[4] ha[2] 1.9 4.0
def/bou hd2e[2] hbr[2] 1.9 3.5
def/bou hd2e[2] hbs[2] 1.9 3.5
def/bou hd2e[2] cb[4] 1.9 4.6
def/bou hd2e[2] mb[5] 1.9 4.2
def/bou hd2z[2] hbr[2] 1.9 3.5
def/bou hd2z[2] hbs[2] 1.9 3.5
def/bou hd2z[2] mb[5] 1.9 4.2
def/bou hd2z[2] cb[4] 1.9 4.6
def/bou hn[2] ha[6] 1.9 3.5
;

```

```

def/bou hn[3] hn[4] 1.9 3.0
def/bou hn[3] hbr[3] 1.9 3.0
def/bou hn[3] hbs[3] 1.9 3.0
def/bou ha[3] hbs[6] 1.9 6.56
def/bou ha[3] hgr[6] 1.9 6.56
def/bou ha[3] hgs[6] 1.9 6.56
def/bou ha[3] hdr[6] 1.9 3.5
def/bou hn[4] hbs[3] 1.9 3.5
def/bou hbr[3] hbr[15] 1.9 3.0
def/bou hn[4] ha[3] 1.9 3.5
def/bou hbr[3] ha[15] 1.9 3.5
;
def/bou hn[4] hn[5] 1.9 2.5
def/bou hn[5] ha[4] 1.9 3.5
def/bou hn[4] cb[4] 1.9 3.5
def/bou hn[4] hgs[4] 1.9 4.0
def/bou hn[5] cb[4] 1.9 4.0
;
def/bou hn[5] hds[6] 1.9 3.5
def/bou hn[5] hdr[6] 1.9 3.0
def/bou ha[5] hds[6] 1.9 2.5
def/bou ha[5] hdr[6] 1.9 2.5
def/bou hn[5] ha[6] 1.9 3.5
;
def/bou hn[7] ha[6] 1.9 2.5
def/bou hbs[6] ha[12] 1.9 6.0
def/bou hgr[6] ha[12] 1.9 6.0
def/bou hgs[6] ha[12] 1.9 6.0
def/bou hbr[6] ha[12] 1.9 3.0
def/bou hbs[6] hn[7] 1.9 7.0
def/bou hgr[6] hn[7] 1.9 7.0
def/bou hgs[6] hn[7] 1.9 7.0
def/bou hbs[6] hbr[11] 1.9 7.0
def/bou hgr[6] hbr[11] 1.9 7.0
def/bou hgs[6] hbr[11] 1.9 7.0
def/bou hbs[6] hn[12] 1.9 7.0
def/bou hgr[6] hn[12] 1.9 7.0
def/bou hgs[6] hn[12] 1.9 7.0
def/bou hn[7] hbr[6] 1.9 3.5
def/bou hn[12] hbr[6] 1.9 3.5
def/bou hbr[11] hbr[6] 1.9 4.0
def/bou hbs[11] hbr[6] 1.9 4.0
def/bou cb[6] cb[15] 1.9 9.0
def/bou cg[6] cb[15] 1.9 9.0
;
def/bou hn[7] hn[8] 1.9 3.0
def/bou hn[7] hbs[7] 1.9 3.0
def/bou hn[7] hbr[7] 1.9 3.5
def/bou hn[7] hgs[7] 1.9 3.5
def/bou hn[8] ha[7] 1.9 3.5
;
def/bou hn[9] ha[8] 1.9 5.6

```

;+3.0 for hgs to hbs upper bound

```

def/bou hn[9] hb[8] 1.9 5.6
def/bou hn[11] ha[8] 1.9 6.6
def/bou hn[11] hb[8] 1.9 6.6
def/bou hn[8] hbr[11] 1.9 3.5
def/bou hn[8] hbs[11] 1.9 3.0
def/bou hn[9] mg[8] 1.9 4.2
def/bou hn[10] mg[8] 1.9 4.8
def/bou hn[11] mg[8] 1.9 4.8
def/bou hn[8] mg[8] 1.9 4.2
def/bou ha[8] mb[9] 1.9 7.1
def/bou hb[8] mb[9] 1.9 7.1
def/bou hn[10] ha[8] 1.9 5.6
def/bou hn[10] hb[8] 1.9 5.6
;
def/bou hn[9] mb[9] 1.9 3.0
def/bou hn[9] hn[10] 1.9 3.0
def/bou hn[10] ha[9] 1.9 3.5
def/bou hn[12] ha[9] 1.9 3.5
def/bou hn[13] ha[9] 1.9 3.5
def/bou hn[10] mb[9] 1.9 3.6
def/bou ha[9] hn[11] 1.9 3.5
def/bou ha[9] hbr[12] 1.9 3.5
def/bou ha[9] hbs[12] 1.9 3.0
def/bou ha[9] cd[12] 1.9 5.0
;
def/bou hn[10] hn[11] 1.9 3.0
def/bou hn[14] ha[10] 1.9 4.0
def/bou hn[11] mb[10] 1.9 3.6
def/bou ha[10] cb[13] 1.9 4.0
def/bou hn[11] ha[10] 1.9 3.5
;
def/bou hn[11] hn[12] 1.9 3.0
def/bou hn[14] ha[11] 1.9 3.5
def/bou hn[15] ha[11] 1.9 3.5
def/bou hn[12] hbr[11] 1.9 3.5
def/bou hn[12] hbs[11] 1.9 3.5
def/bou hn[11] hbr[11] 1.9 3.0
def/bou hn[11] hbs[11] 1.9 3.0
def/bou ha[11] hbr[14] 1.9 3.0
def/bou ha[11] hbs[14] 1.9 3.0
;
def/bou hn[12] hn[13] 1.9 3.0
def/bou hn[15] ha[12] 1.9 3.5
def/bou hn[13] ha[12] 1.9 3.5
def/bou hn[12] hbr[12] 1.9 3.0
def/bou hn[12] hbs[12] 1.9 3.0
def/bou ha[12] hbs[15] 1.9 4.0
def/bou ha[12] hbr[15] 1.9 3.5
def/bou hn[12] hgs[12] 1.9 3.5
;
def/bou hn[13] hn[14] 1.9 3.0
def/bou hn[14] cb[13] 1.9 4.0

```

```

def/bou hn[14] ha[13] 1.9 3.5
def/bou hn[13] cb[13] 1.9 3.5
def/bou ha[13] hbr[16] 1.9 3.5
def/bou ha[13] hbs[16] 1.9 3.5
;
def/bou hn[15] ha[14] 1.9 3.5
def/bou hn[15] hbr[14] 1.9 3.5
def/bou hn[15] hbs[14] 1.9 3.0
def/bou hn[14] hbr[14] 1.9 3.5
def/bou hn[14] hbs[14] 1.9 3.0
def/bou ha[14] hn[17] 1.9 3.5
def/bou ha[14] hbr[17] 1.9 4.0
def/bou ha[14] hbs[17] 1.9 4.0
;
def/bou hn[15] hn[16] 1.9 3.0
def/bou hn[15] hbr[15] 1.9 3.0
def/bou hn[15] hbs[15] 1.9 3.0
def/bou hn[16] hbr[15] 1.9 3.5
def/bou hn[16] hbs[15] 1.9 4.0
def/bou hn[16] ha[15] 1.9 3.5
;
def/bou hn[16] hn[17] 1.9 3.0
def/bou hn[17] ha[16] 1.9 3.0
def/bou hn[16] hbr[16] 1.9 3.0
def/bou hn[16] hbs[16] 1.9 3.5
def/bou hn[16] hgr[16] 1.9 3.5
def/bou hn[16] hgs[16] 1.9 3.5
;
def/bou hn[17] hn[18] 1.9 3.0
def/bou hn[18] ha[17] 1.9 3.0
def/bou hn[18] hbr[17] 1.9 4.0
def/bou hn[17] hbr[17] 1.9 3.0
def/bou hn[17] hbs[17] 1.9 3.0

```

RpIII

```

;
;      H-Bonds
;
def/hbond hn[20] o[44] 1.80 2.00
def/hbond hn[21] o[1] 1.80 2.00
def/hbond hn[22] o[42] 1.80 2.00
def/hbond hn[30] o[27] 1.80 2.00
def/hbond hn[31] o[45] 1.80 2.00
def/hbond hn[33] o[43] 1.80 2.00
def/hbond hn[42] o[22] 1.80 2.00
def/hbond hn[43] o[33] 1.80 2.00
def/hbond hn[44] o[20] 1.80 2.00
def/hbond hn[45] o[31] 1.80 2.00
;
;      Disulfide Constraints
;
def/bou cb[3] cb[43] 3.72 3.99
def/bou cb[5] cb[33] 3.72 3.99
def/bou cb[26] cb[44] 3.72 3.99
;
;      Eliminate Chirality
;
def/nochiral ca[1]
def/nochiral cb[3]
def/nochiral cb[4]
def/nochiral cg[4]
def/nochiral cb[5]
def/nochiral cb[6]
def/nochiral cb[7]
def/nochiral cg[8]
def/nochiral cb[8]
def/nochiral ca[9]
def/nochiral cd[10]
def/nochiral cb[11]
def/nochiral cb[16]
def/nochiral cg[16]
def/nochiral cd[16]
def/nochiral cb[17]
def/nochiral cg[17]
def/nochiral ca[19]
def/nochiral cb[21]
def/nochiral cb[23]
def/nochiral cg[23]
def/nochiral ca[24]
def/nochiral cb[25]
def/nochiral cb[26]
def/nochiral cb[27]
def/nochiral cb[28]
def/nochiral ca[29]
def/nochiral cb[30]

```

```

def/nochiral cg[31]
def/nochiral cb[32]
def/nochiral cb[33]
def/nochiral cb[36]
def/nochiral cb[38]
def/nochiral cd[39]
def/nochiral cb[42]
def/nochiral cg[42]
def/nochiral cb[43]
def/nochiral cb[45]
def/nochiral cd[45]
def/nochiral cb[46]
def/nochiral cb[47]
def/nochiral cg[47]
def/nochiral cb[48]
;
;      Distance Constraints
;
def/bou har[1] hn[2] 1.9 3.0
;def/bou has[1] hn[2] 1.9 3.5
;def/bou hn[3] har[1] 1.9 3.5
;
def/bou ha[2] ha[20] 1.9 4.0
def/bou hn[2] hn[3] 1.9 2.5
def/bou hn[2] hn[19] 1.9 4.0
def/bou hbr[2] hn[34] 1.9 4.0
def/bou hbs[2] hn[34] 1.9 4.0
def/bou hn[2] mdr[17] 1.9 4.8
def/bou hn[3] hbs[2] 1.9 4.0
def/bou hn[2] ha[20] 1.9 4.0
;
def/bou hn[3] hn[19] 1.9 4.0
def/bou hn[4] hbr[3] 1.9 4.0
def/bou ha[3] mgs[21] 1.9 4.8
def/bou ha[3] mdr[17] 1.9 4.8
def/bou hbr[3] hbs[43] 1.9 4.0
def/bou hbr[3] mgs[21] 2.12 3.6
def/bou hbs[3] mgs[21] 2.12 3.0
def/bou ha[3] hn[4] 1.9 2.5
def/bou ha[3] hgr[4] 1.9 4.0
def/bou ha[3] mds[17] 2.12 3.6
def/bou hn[3] mdr[17] 2.12 4.2
def/bou cz[37] hbr[3] 2.45 4.6
def/bou cz[37] hbs[3] 2.45 4.6
def/bou cz[37] ha[3] 2.45 5.1
def/bou cg[37] hbr[3] 2.45 5.8
def/bou cg[37] hbs[3] 2.45 5.8
def/bou cg[37] ha[3] 2.45 6.3
;
def/bou hn[4] mdr[17] 1.9 4.8
def/bou hn[4] mds[17] 1.9 4.8
def/bou hn[4] cz[37] 1.9 6.1

```

```

def/bou hn[4] cg[37] 1.9 7.3
def/bou ha[4] mds[17] 2.12 3.6
def/bou ha[4] ha[17] 1.9 3.0
def/bou ha[4] hgr[4] 1.9 3.0
def/bou ha[4] hgs[4] 1.9 3.0
def/bou ha[4] hn[5] 1.9 2.5
def/bou hn[4] hbs[4] 1.9 3.0
def/bou hbs[4] hn[5] 1.9 3.5
;
def/bou hn[5] hbr[5] 1.9 4.0
def/bou hn[5] hbs[5] 1.9 4.0
def/bou hbr[5] hn[6] 1.9 4.0
def/bou hn[5] ha[32] 1.9 4.0
def/bou hbr[5] ha[33] 1.9 4.0
def/bou ha[5] hn[6] 1.9 2.5
def/bou ha[5] mb[34] 2.12 3.60
def/bou hn[5] mds[17] 2.12 4.20
def/bou hbs[5] hn[6] 1.9 3.5
def/bou hbs[5] ha[33] 1.9 4.0
def/bou ha[5] ha[33] 1.9 4.0
;
def/bou hn[6] hn[35] 1.9 4.0
def/bou hn[6] hn[7] 1.9 3.0
def/bou hn[6] ha[34] 1.9 3.0
def/bou hn[6] hbr[6] 1.9 3.5
def/bou hn[6] hbs[6] 1.9 3.0
def/bou hbr[6] hn[7] 1.9 3.0
def/bou hbs[6] hn[7] 1.9 3.0
;
def/bou hn[7] hn[8] 1.9 3.0
def/bou ha[7] hn[8] 1.9 3.0
def/bou hn[7] hbr[7] 1.9 3.0
def/bou hn[7] hbs[7] 1.9 3.0
;
def/bou hbr[8] hn[9] 1.9 4.0
def/bou ha[8] hgr[8] 1.9 3.0
def/bou ha[8] hgs[8] 1.9 3.0
def/bou hn[8] hn[9] 1.9 3.0
def/bou ha[8] hn[9] 1.9 3.0
def/bou hn[8] hbr[8] 1.9 3.0
def/bou hn[8] hbs[8] 1.9 3.0
;
def/bou har[9] hdr[10] 1.9 3.0
def/bou has[9] hdr[10] 1.9 3.0
def/bou har[9] hds[10] 1.9 3.0
def/bou has[9] hds[10] 1.9 3.0
;
def/bou ha[10] hn[11] 1.9 3.0
def/bou hbs[10] hn[11] 1.9 4.0
;
def/bou hn[11] hbr[11] 1.9 3.0
def/bou hn[11] hbs[11] 1.9 3.0

```

```

def/bou ha[11] hn[12] 1.9 3.0
;
def/bou hn[12] hn[13] 1.9 4.0
def/bou ha[12] hn[13] 1.9 3.0
def/bou hn[12] hb[12] 1.9 3.0
def/bou hn[13] hb[12] 1.9 3.5
;
def/bou cb[13] mg[14] 2.68 5.60
def/bou hn[13] hn[14] 1.9 3.0
def/bou ha[13] hn[14] 1.9 3.0
;
def/bou ha[14] hn[15] 1.9 3.0
def/bou hn[14] hn[15] 1.9 4.0
def/bou m[14] hn[15] 1.9 4.8
;
def/bou hn[15] mb[15] 1.9 4.8
def/bou ha[15] hdr[16] 1.9 2.5
def/bou ha[15] hds[16] 1.9 2.5
def/bou mb[15] hds[16] 2.12 3.6
def/bou mb[15] hdr[16] 2.12 3.6
;
def/bou hbr[16] hn[17] 1.9 4.0
def/bou hbs[16] hn[17] 1.9 4.0
def/bou hbs[16] mg[18] 2.12 4.8
def/bou hgs[16] mg[18] 2.12 4.8
def/bou ha[16] hn[17] 1.9 2.5
def/bou hgr[16] mg[18] 2.12 4.8
;
def/bou hn[17] hn[18] 1.9 3.0
def/bou ha[17] hn[18] 1.9 3.0
def/bou ha[17] mds[17] 2.12 3.0
def/bou hn[17] hbr[17] 1.9 3.0
;
def/bou hn[18] m[18] 1.9 4.0
def/bou ha[18] hdr[45] 1.9 4.0
def/bou ha[18] hds[45] 1.9 4.0
def/bou ha[18] hn[19] 1.9 3.0
def/bou hn[18] hn[19] 1.9 3.0
def/bou hn[18] hbs[45] 1.9 3.5
;
def/bou hn[19] hn[20] 1.9 4.0
def/bou har[19] hn[20] 1.9 3.0
def/bou has[19] hn[20] 1.9 3.0
def/bou cg[20] har[19] 2.45 6.6
def/bou cg[20] has[19] 2.45 6.6
def/bou cz[20] has[19] 2.45 6.6
def/bou cz[20] har[19] 2.45 7.8
def/bou har[19] ha[45] 1.9 4.0
def/bou has[19] ha[45] 1.9 4.0
;
def/bou ha[20] hn[21] 1.9 2.5
def/bou ha[20] mgs[21] 2.12 4.8

```



```

def/bou hn[20] hn[44] 1.9 3.0
def/bou hh2[30] cb[20] 2.45 5.0
def/bou cg[20] ha[20] 2.45 5.1
def/bou cz[20] ha[20] 2.45 6.3
def/bou cb[20] hh2[30] 1.9 5.4
;
def/bou ha[21] ha[43] 1.9 2.5
def/bou ha[21] hn[44] 1.9 4.0
def/bou hn[21] hb[21] 1.9 3.5
def/bou hn[21] mgs[21] 2.12 3.6
def/bou mgr[21] mb[41] 2.34 3.40
def/bou ha[21] hn[22] 1.9 2.5
def/bou mgr[21] hn[22] 2.12 3.60
def/bou cz[37] mgr[21] 2.67 5.10
def/bou cz[37] mgs[21] 2.67 5.10
def/bou cg[37] mgr[21] 2.67 6.30
def/bou cg[37] mgs[21] 2.67 6.30
def/bou mgr[21] ha[22] 2.12 4.8
;
def/bou ha[22] hn[23] 1.9 2.5
def/bou hn[22] ha[43] 1.9 4.0
def/bou ha[22] hn[42] 1.9 4.0
def/bou hn[22] hn[42] 1.9 4.0
def/bou cb[22] hn[23] 1.9 5.0
;
def/bou hbr[23] hn[24] 1.9 4.0
def/bou hbs[23] hn[24] 1.9 4.0
def/bou ha[23] ha[41] 1.9 2.5
def/bou ha[23] mb[41] 2.12 3.0
def/bou ha[23] mdr[23] 2.12 3.6
def/bou ha[23] mds[23] 2.12 3.0
def/bou ha[23] hn[24] 1.9 2.5
def/bou hn[23] hbr[23] 1.9 3.0
def/bou hn[23] hbs[23] 1.9 3.0
def/bou mds[23] ha[41] 2.12 4.8
def/bou mds[23] ha[40] 2.12 3.6
def/bou mds[23] mb[41] 2.34 4.0
;
def/bou har[24] hn[25] 1.9 2.5
def/bou has[24] hn[25] 1.9 3.0
def/bou hn[24] mb[41] 2.12 3.6
def/bou hn[24] ha[41] 1.9 3.0
def/bou hn[24] hn[42] 1.9 3.5
def/bou har[24] cg[25] 2.45 6.1
def/bou har[24] cz[25] 2.45 7.3
;
def/bou ha[25] hgr[42] 1.9 4.0
def/bou ha[25] hgs[42] 1.9 4.0
def/bou ha[25] hn[26] 1.9 2.5
def/bou hh2[30] hbs[25] 1.9 6.4
def/bou hel[30] hbs[25] 1.9 3.5
def/bou cg[25] ha[25] 2.45 5.1

```

```

def/bou cz[25] ha[25] 2.45 6.3
def/bou hn[26] hbr[25] 1.9 3.0
def/bou hn[26] hbs[25] 1.9 3.0
;
def/bou hn[26] mb[34] 1.9 4.8
def/bou ha[26] hn[27] 1.9 2.5
def/bou hn[26] hbs[26] 1.9 3.0
def/bou hn[26] hbr[26] 1.9 3.0
def/bou hn[27] hbr[26] 1.9 3.0
def/bou hn[27] hbs[26] 1.9 3.5
def/bou ha[44] hbr[26] 1.9 4.0
def/bou hn[26] cb[39] or hbs[7]           possible
;
def/bou hn[28] hbs[27] 1.9 4.0
def/bou hn[27] hbr[27] 1.9 3.5
def/bou hn[27] hbs[27] 1.9 3.0
def/bou hn[27] hbr[30] 1.9 3.5           possible
def/bou hd1[30] hbr[27] 1.9 3.5
def/bou hd1[30] hbs[27] 1.9 3.5
;
def/bou ha[28] hn[30] 1.9 3.0
def/bou ha[28] hn[29] 1.9 2.5
def/bou hn[28] hbs[28] 1.9 3.0
def/bou hn[28] hbr[28] 1.9 3.5
def/bou hn[29] hbr[28] 1.9 3.0
def/bou hn[29] hbs[28] 1.9 4.0
;
def/bou hn[29] hn[30] 1.9 3.0
def/bou has[29] hn[30] 1.9 3.5
def/bou har[29] hn[30] 1.9 3.5
def/bou hn[29] hd1[30] 1.9 3.5
def/bou hd1[30] has[29] 1.9 4.0
def/bou hd1[30] har[29] 1.9 3.5
;
def/bou ha[30] ha[46] 1.9 2.5
def/bou ha[30] hgs[47] 1.9 4.0
def/bou ha[30] hn[31] 1.9 3.0
def/bou hbr[30] hn[31] 1.9 3.5
def/bou hbs[30] hn[31] 1.9 3.5
def/bou hn[30] hd1[30] 1.9 3.0           ;added 0.5 to hd1[30] constraints
def/bou ha[30] hn[47] 1.9 2.5           ;since it is a singlet.
def/bou he3[30] ha[44] 1.9 4.0
def/bou he3[30] ha[45] 1.9 3.0
def/bou he3[30] hbr[30] 2.01 4.0
def/bou he3[30] hbs[30] 2.01 4.0
def/bou hz3[30] ha[45] 1.9 3.0
def/bou hz2[30] hbr[46] 1.9 4.0
def/bou hd1[30] ha[46] 1.9 4.5
def/bou hd1[30] ha[30] 1.9 4.5
def/bou hd1[30] hbr[30] 1.9 4.5
def/bou hd1[30] hbs[30] 1.9 4.5
def/bou he3[30] ha[30] 1.9 4.0

```

```

def/bou hbr[30] cb[44] 1.9 5.0
def/bou hn[30] hbr[30] 1.9 3.0
def/bou hn[30] hbs[30] 1.9 3.0
def/bou hz3[30] hn[46] 1.9 3.5
def/bou he3[30] hn[46] 1.9 3.5
def/bou he3[30] hn[45] 1.9 3.5
def/bou he3[30] hn[44] 1.9 4.0
def/bou he3[30] hn[31] 1.9 4.0
;
def/bou hbr[31] hn[32] 1.9 4.0
def/bou hbs[31] hn[32] 1.9 4.0
def/bou hn[31] hn[47] 1.9 4.0
def/bou hn[31] ha[44] 1.9 4.0
def/bou hn[31] ha[46] 1.9 4.0
def/bou hn[31] hgr[31] 1.9 3.5
def/bou hn[31] hgs[31] 1.9 4.0
def/bou hn[31] hn[45] 1.9 3.5
def/bou ha[31] hn[32] 1.9 2.5
;
def/bou hn[32] cb[45] 1.9 5.0
def/bou hbr[32] hn[33] 1.9 4.0
def/bou hbs[32] hn[33] 1.9 4.0
def/bou ha[32] ha[44] 1.9 2.5
def/bou ha[32] hn[33] 1.9 2.5
def/bou ha[32] hn[45] 1.9 3.5
def/bou hn[32] hbr[32] 1.9 3.0
def/bou hn[32] hbs[32] 1.9 3.0
;
def/bou hn[33] hbr[33] 1.9 4.0
def/bou hbr[33] hn[45] 1.9 4.0
def/bou hn[33] ha[44] 1.9 3.0
def/bou hn[33] hbs[33] 1.9 3.0
def/bou hbs[33] hn[45] 1.9 3.0
def/bou hn[33] hn[34] 1.9 3.0
def/bou ha[33] hn[34] 1.9 3.0
;
def/bou hn[34] mb[34] 1.9 4.8
def/bou mb[34] hbr[43] 2.12 3.0
def/bou mb[34] hbs[43] 2.12 3.6
def/bou hn[35] mb[34] 2.12 4.20
def/bou cg[37] mb[34] 2.67 6.1
def/bou cz[37] mb[34] 2.67 7.3
def/bou hn[35] ha[34] 1.9 2.5
;
def/bou ha[35] hn[40] 1.9 3.5
def/bou cb[35] hn[36] 1.9 5.0
def/bou hn[35] hn[36] 1.9 3.0
def/bou ha[35] hn[36] 1.9 3.0
def/bou ha[35] hn[40] 1.9 3.5
;
def/bou ha[36] hn[37] 1.9 2.5
def/bou cz[36] ha[39] 2.45 6.1

```

```

def/bou cg[36] ha[39] 2.45 7.3
def/bou cg[36] ha[36] 2.45 4.6
def/bou cz[36] ha[36] 2.45 5.8
def/bou cz[36] cg[39] 3.00 7.1
def/bou cg[36] cg[39] 3.00 8.3
def/bou hn[36] hbr[36] 1.9 3.0
def/bou hn[36] hbs[36] 1.9 3.0
;
def/bou hbr[37] hn[38] 1.9 4.0
def/bou hbs[37] hn[38] 1.9 4.0
def/bou cz[37] hn[43] 1.9 6.1
def/bou cg[37] hn[43] 1.9 7.3
def/bou hn[37] hn[38] 1.9 2.5
def/bou cg[37] hbs[43] 2.45 6.1
def/bou cg[37] mb[41] 2.67 5.10
def/bou cz[37] hbs[43] 2.45 6.6
def/bou cz[37] mb[41] 2.67 5.10
def/bou cg[37] ha[42] 2.45 5.1
def/bou cz[37] ha[42] 2.45 6.3
def/bou cg[37] ha[37] 2.45 4.6
def/bou cz[37] ha[37] 2.45 5.8
def/bou cg[37] hbs[38] 2.45 6.1
def/bou cz[37] hbs[38] 2.45 6.1
;
def/bou hn[38] hbs[38] 1.9 4.0
def/bou ha[38] hbr[39] 1.9 3.0
def/bou ha[38] hds[39] 1.9 3.0
;
def/bou hbr[39] mg2[40] 2.12 4.8
;
def/bou mg2[40] hn[41] 1.9 4.8
def/bou md[40] hn[41] 1.9 4.8
def/bou hn[40] hn[41] 1.9 3.0
def/bou ha[40] mg2[40] 2.21 3.0
def/bou hn[40] mg2[40] 2.12 3.60
def/bou hn[40] md[40] 2.12 3.60
;
def/bou hn[41] mb[41] 1.9 4.8
def/bou hn[42] mb[41] 2.12 3.60
;
def/bou hn[42] hgr[42] 1.9 4.0
def/bou ha[42] hgr[42] 1.9 4.0
def/bou ha[42] hgs[42] 1.9 3.0
def/bou hn[42] hbr[42] 1.9 3.5
def/bou hn[42] hbs[42] 1.9 3.5
def/bou hn[43] hbr[42] 1.9 3.5
def/bou ha[42] hn[43] 1.9 2.5
;
def/bou ha[43] hn[44] 1.9 2.5
def/bou hn[43] hbr[43] 1.9 3.5
def/bou hn[43] hbs[43] 1.9 3.5
;

```

possible

```
def/bou cb[44] hn[45] 1.9 5.0
def/bou ha[44] hn[45] 1.9 2.5
;
def/bou hn[45] cb[45] 1.9 5.0
def/bou cb[45] hn[46] 1.9 5.0
def/bou ha[45] hn[46] 1.9 2.5
;
def/bou hn[46] hgs[46] 1.9 4.0
def/bou hn[47] hgs[46] 1.9 4.0
def/bou ha[46] hn[47] 1.9 2.5
def/bou hn[46] hbr[46] 1.9 3.0
def/bou hn[46] hbs[46] 1.9 3.0
def/bou hbr[46] hn[47] 1.9 3.0
;
def/bou ha[47] hn[48] 1.9 2.5
def/bou hn[47] hbr[47] 1.9 3.5
def/bou hn[47] hbs[47] 1.9 3.0
def/bou hn[48] hbs[47] 1.9 3.5
def/bou hn[48] hbr[47] 1.9 4.0
;
def/bou hn[48] hbr[48] 1.9 3.0
def/bou hn[48] hbs[48] 1.9 3.0
```

RpII

```

;
;      H-bonds
;
def/hbond hn[3]  o[19] 1.8 2.0
def/hbond hn[20] o[44] 1.8 2.0
def/hbond hn[21] o[1]  1.8 2.0
def/hbond hn[22] o[42] 1.8 2.0
def/hbond hn[30] o[27] 1.8 2.0
def/hbond hn[31] o[45] 1.8 2.0
def/hbond hn[33] o[43] 1.8 2.0
def/hbond hn[42] o[22] 1.8 2.0
;def/hbond hn[43] o[33] 1.8 2.0
def/hbond hn[44] o[20] 1.8 2.0
def/hbond hn[45] o[31] 1.8 2.0
;
;      Disulfide Constraints
;
def/bound cb[3]  cb[43] 3.72 3.99
def/bound cb[5]  cb[33] 3.72 3.99
def/bound cb[26] cb[44] 3.72 3.99
;
;      Eliminate Chirality
;
def/nochir cb[5]
def/nochir cb[21]
def/nochir cb[36]
;
;      Distance Constraints
;
def/bound ha[1] hn[2] 1.9 2.5
def/bound mb[1] hn[2] 1.9 3.6
def/bound mb[1] mgs[21] 1.9 4.0
def/bound mb[1] hb[21] 1.9 3.6
;
def/bound ha[2] hn[3] 1.9 2.5
def/bound hbr[2] hn[3] 1.9 4.0
def/bound hbs[2] hn[3] 1.9 4.0
def/bound ha[2] ha[20] 1.9 2.5
;
def/bound ha[3] hn[4] 1.9 2.5
def/bound hbr[3] hn[4] 1.9 4.6
def/bound hbs[3] hn[4] 1.9 4.6
def/bound hbr[3] mgs[21] 1.9 3.6
def/bound hbs[3] mgs[21] 1.9 3.6
def/bound ha[3] mgs[21] 1.9 4.8
def/bound cz[37] hbr[3] 2.5 4.6
def/bound cz[37] hbs[3] 2.5 4.6
def/bound cg[37] hbr[3] 2.5 5.8
def/bound cg[37] hbs[3] 2.5 5.8
;

```

```

def/bound ha[4] hgr[4] 1.9 4.0
def/bound ha[4] hgs[4] 1.9 4.0
def/bound hbr[4] hn[4] 1.9 4.6
def/bound hbs[4] hn[4] 1.9 4.6
def/bound ha[4] hn[5] 1.9 2.5
def/bound hbr[4] hn[5] 1.9 4.1
def/bound hbs[4] hn[5] 1.9 4.1
def/bound ha[4] ha[17] 1.9 3.0
def/bound cg[17] ha[4] 2.5 5.1
def/bound ha[4] cz[17] 2.5 6.3
;
def/bound hbs[5] ha[33] 1.9 4.0
def/bound ha[5] ha[33] 1.9 4.0
;
def/bound hn[7] hn[8] 1.9 4.0
;
def/bound hbr[8] hn[8] 1.9 4.6
def/bound hbs[8] hn[8] 1.9 4.6
def/bound hn[8] hn[9] 1.9 4.0
def/bound hbr[8] hn[9] 1.9 5.6
def/bound hbs[8] hn[9] 1.9 5.6
;
def/bound har[9] hdr[10] 1.9 3.0
def/bound has[9] hdr[10] 1.9 3.0
def/bound har[9] hds[10] 1.9 3.0
def/bound has[9] hds[10] 1.9 3.0
;
def/bound ha[11] hn[12] 1.9 2.5
;
def/bound hb[12] hn[12] 1.9 3.0
def/bound ha[12] hn[13] 1.9 2.5
def/bound hn[12] hn[13] 1.9 4.0
def/bound hb[12] hn[13] 1.9 4.0
def/bound cb[12] hn[13] 1.9 5.45
def/bound cg[17] cb[12] 2.5 7.55
def/bound cz[17] cb[12] 2.5 7.55
;
def/bound ha[13] hgr[13] 1.9 4.1
def/bound ha[13] hn[14] 1.9 3.0
def/bound hn[13] hn[14] 1.9 4.0
;
def/bound ha[14] hn[15] 1.9 3.0
def/bound hn[14] hn[15] 1.9 4.0
;
def/bound mb[15] hn[15] 1.9 3.0
def/bound mb[15] hn[16] 1.9 4.8
def/bound mb[15] mg[16] 1.9 4.0
def/bound cz[17] mb[15] 2.5 5.7
def/bound cg[17] mb[15] 2.5 6.9
;
def/bound ha[16] hn[17] 1.9 3.0
def/bound hn[16] hn[17] 1.9 4.0

```

```

def/bound ha[17] cg[17] 1.9 4.6
def/bound cg[17] hbr[17] 1.9 5.1
def/bound cg[17] hbs[17] 1.9 5.1
def/bound ha[17] hn[18] 1.9 2.5
def/bound cg[17] ha[18] 2.5 6.1
def/bound cz[17] ha[18] 2.5 6.1
;
def/bound ha[18] mg[18] 1.9 3.0
def/bound mg[18] hn[18] 1.9 4.8
def/bound hn[18] hn[19] 1.9 4.0
def/bound ha[18] hdr[45] 1.9 3.5
def/bound ha[18] hds[45] 1.9 3.5
;
def/bound har[19] hn[20] 1.9 4.0
def/bound has[19] hn[20] 1.9 4.0
;
def/bound ha[20] mg[20] 1.9 3.6
def/bound hb[20] hn[20] 1.9 3.0
def/bound ha[20] hn[21] 1.9 2.5
def/bound mg[20] hn[21] 1.9 4.8
def/bound hn[20] hn[44] 1.9 4.0
;
def/bound ha[21] mgr[21] 1.9 3.0
def/bound ha[21] mgs[21] 1.9 3.0
def/bound hb[21] hn[21] 1.9 4.0
def/bound mgr[21] hn[21] 1.9 3.6
def/bound mgs[21] hn[21] 1.9 3.6
def/bound ha[21] hn[22] 1.9 2.5
def/bound cz[37] mgr[21] 2.5 5.1
def/bound cz[37] mgs[21] 2.5 5.1
def/bound mb[41] mgr[21] 1.9 4.0
def/bound ha[43] ha[21] 1.9 2.5
def/bound cg[37] mgr[21] 1.9 6.3
def/bound cg[37] mgs[21] 1.9 6.3
;
def/bound ha[22] hn[23] 1.9 2.5
;
def/bound ha[23] hn[24] 1.9 3.0
def/bound ha[23] ha[41] 1.9 3.0
;
def/bound ha[24] he3[24] 1.9 4.0
def/bound ha[24] hn[25] 1.9 3.0
def/bound hbr[24] hn[25] 1.9 4.6
;
def/bound ha[25] hn[26] 1.9 4.0
;
def/bound hbr[26] hn[26] 1.9 4.0
def/bound hbs[26] hn[26] 1.9 4.0
def/bound ha[26] hn[27] 1.9 3.0
def/bound hbr[26] hn[27] 1.9 4.6
;
def/bound hbr[27] hn[27] 1.9 4.0

```



```

def/bound hbs[27] hn[27] 1.9 4.0
def/bound ha[27] hn[28] 1.9 3.0
def/bound hbr[27] hn[28] 1.9 4.6
;
def/bound ha[28] hgr[28] 1.9 3.0
def/bound ha[28] hn[29] 1.9 2.5
;
def/bound har[29] hn[30] 1.9 5.6
def/bound has[29] hn[30] 1.9 5.6
def/bound hn[29] hn[30] 1.9 4.0
def/bound hd1[30] har[29] 1.9 3.5
def/bound hd1[30] has[29] 1.9 3.5
;
def/bound he3[30] hbr[30] 1.9 4.0
def/bound ha[30] hgs[47] 1.9 4.0
def/bound he3[30] hbs[30] 1.9 4.0
def/bound ha[30] hn[31] 1.9 2.5
def/bound hbr[30] hn[31] 1.9 4.6
def/bound hbs[30] hn[31] 1.9 4.6
def/bound ha[30] ha[46] 1.9 2.5
def/bound ha[30] hn[47] 1.9 4.0
def/bound he3[30] ha[45] 1.9 3.0
def/bound hz3[30] ha[45] 1.9 3.0
;
def/bound ha[31] hn[32] 1.9 2.5
def/bound hn[31] hn[45] 1.9 4.0
;
def/bound ha[32] hn[33] 1.9 2.5
def/bound ha[32] ha[44] 1.9 2.5
;
def/bound ha[33] hn[34] 1.9 4.0
def/bound hn[33] hn[34] 1.9 4.0
;
def/bound ha[34] mg[34] 1.9 3.0
def/bound ha[34] hn[35] 1.9 3.0
def/bound mg[34] hbr[43] 1.9 3.0
def/bound mg[34] hbs[43] 1.9 3.0
def/bound cg[37] mg[34] 2.67 6.1
def/bound cz[37] mg[34] 2.67 7.3
;
def/bound mb[35] hn[35] 1.9 3.6
def/bound mb[35] hn[36] 1.9 4.8
def/bound hn[35] hn[36] 1.9 4.0
def/bound mb[35] mgr[36] 1.9 4.7
;
def/bound mgr[36] hn[36] 1.9 4.8
def/bound mgs[36] hn[36] 1.9 4.8
def/bound ha[36] mgr[36] 1.9 3.6
def/bound ha[36] mgs[36] 1.9 3.6
def/bound ha[36] hn[37] 1.9 2.5
;
def/bound ha[37] cg[37] 1.9 4.6

```

```

def/bound cg[37] hbr[37] 2.5 5.1
def/bound cg[37] hbs[37] 2.5 5.1
def/bound hbr[37] hn[38] 1.9 4.6
def/bound hbs[37] hn[38] 1.9 4.6
def/bound cg[37] mg[38] 2.5 5.1
def/bound cz[37] mg[38] 2.5 5.1
;
def/bound ha[38] mg[38] 2.5 3.0
def/bound ha[38] hbr[39] 1.9 2.5
def/bound ha[38] hds[39] 1.9 2.5
def/bound mg[38] hn[38] 2.5 4.8
;
def/bound ha[40] mgr[40] 1.9 3.0
def/bound ha[40] mgs[40] 1.9 4.2
def/bound hn[40] hn[41] 1.9 4.0
;
def/bound mb[41] hn[41] 1.9 4.8
def/bound ha[41] hn[42] 1.9 3.0
def/bound mb[41] hn[42] 1.9 3.6
;
def/bound hbr[42] hn[42] 1.9 5.6
def/bound hbs[42] hn[42] 1.9 5.6
def/bound ha[42] hn[43] 1.9 3.0
def/bound hbr[42] hn[43] 1.9 5.6
def/bound hbs[42] hn[43] 1.9 5.6
;
def/bound hbr[43] hn[43] 1.9 4.0
def/bound hbs[43] hn[43] 1.9 4.0
def/bound ha[43] hn[44] 1.9 4.0
;
def/bound hbr[44] hn[44] 1.9 4.6
def/bound hbs[44] hn[44] 1.9 4.6
def/bound ha[44] hn[45] 1.9 2.5
def/bound hbr[44] hn[45] 1.9 4.6
def/bound hbs[44] hn[45] 1.9 4.6
;
def/bound ha[45] hn[46] 1.9 2.5
;
;def/bound hbr[46] hn[46] 1.9 3.0
def/bound ha[46] hn[47] 1.9 2.5
def/bound hbr[46] hn[47] 1.9 5.6
def/bound hbs[46] hn[47] 1.9 5.6
;
def/bound ha[47] hgr[47] 1.9 3.0
def/bound hbr[47] hn[47] 1.9 4.6
def/bound hbs[47] hn[47] 1.9 4.6
def/bound ha[47] hn[48] 1.9 2.5
def/bound hbr[47] hn[48] 1.9 5.8
def/bound hbs[47] hn[48] 1.9 5.8
;
;def/bound ha[48] hn[48] 1.9 3.0

```

Appendix III

α -Carbon Coordinates

Apamin

Apa-S25

Rp II

Rp III

Apamin

ca	2.605	0.775	-4.847
ca	4.628	-1.920	-3.134
ca	4.130	-3.938	0.025
ca	7.861	-3.857	0.594
ca	7.849	-0.120	1.176
ca	4.756	1.700	2.161
ca	5.757	4.609	-0.036
ca	2.748	6.429	1.347
ca	-0.520	6.069	3.220
ca	-1.616	5.023	-0.238
ca	0.503	1.897	-0.275
ca	-0.073	1.357	3.417
ca	-3.831	1.260	3.104
ca	-3.744	-0.584	-0.195
ca	-1.431	-3.297	1.075
ca	-3.657	-3.558	4.103
ca	-6.808	-3.566	2.024
ca	-7.894	-5.063	-1.265

Apa-S25

ca	0.197	-6.618	1.442
ca	-1.913	-5.259	4.261
ca	-2.616	-1.647	3.399
ca	-3.046	-1.399	7.152
ca	0.610	-2.154	7.576
ca	1.738	-2.066	3.979
ca	4.766	-4.187	3.179
ca	6.233	-1.039	1.692
ca	5.431	2.491	0.608
ca	5.148	0.902	-2.801
ca	2.420	-1.082	-1.094
ca	0.832	1.772	0.757
ca	1.099	2.909	-2.839
ca	-0.409	-0.291	-4.163
ca	-3.131	0.593	-1.750
ca	-3.095	4.124	-3.081
ca	-2.651	3.390	-6.764
ca	-5.403	0.862	-7.367

Rp II

ca	3.393	8.826	-1.105
ca	2.902	5.261	-2.275
ca	4.990	2.206	-1.534
ca	5.804	-1.172	-3.033
ca	4.945	-4.422	-1.321
ca	7.672	-6.543	0.182
ca	9.829	-5.671	-2.791
ca	12.707	-7.549	-4.369
ca	13.026	-7.104	-8.095
ca	11.330	-5.811	-11.197
ca	7.923	-7.205	-12.068
ca	4.483	-5.641	-12.164
ca	2.594	-4.212	-15.110
ca	-0.625	-2.533	-14.215
ca	1.371	-1.154	-11.372
ca	1.434	1.422	-8.670
ca	3.014	0.006	-5.529
ca	-0.601	0.403	-5.022
ca	-0.549	1.261	-1.344
ca	-0.779	4.945	-0.559
ca	1.593	5.691	2.264
ca	-1.337	7.026	4.237
ca	-2.646	7.277	7.769
ca	-3.166	4.246	9.971
ca	-6.589	2.652	10.174
ca	-6.591	0.468	7.084
ca	-8.640	-2.639	7.730
ca	-10.624	-5.371	6.039
ca	-11.921	-3.075	3.343
ca	-8.517	-2.122	2.014
ca	-5.201	-3.367	0.689
ca	-1.964	-2.936	2.601
ca	1.560	-3.939	1.694
ca	3.520	-2.426	4.519
ca	2.713	-0.090	7.381
ca	5.999	0.430	9.184
ca	7.902	2.212	6.462
ca	8.748	5.725	7.536
ca	7.771	8.649	9.701
ca	4.025	8.344	10.039
ca	2.154	6.350	7.436
ca	0.689	2.977	6.577
ca	1.253	1.602	3.096
ca	-2.267	0.963	1.887
ca	-4.882	1.864	-0.682
ca	-8.536	0.957	-0.825
ca	-9.716	-2.273	-2.371
ca	-8.235	-4.620	-4.934

Rp III

ca	-1.946	1.273	9.969
ca	-2.527	-0.949	6.983
ca	-3.318	2.250	5.129
ca	-5.207	2.603	1.880
ca	-3.267	3.016	-1.328
ca	-3.005	6.328	-3.140
ca	-1.545	6.980	-6.539
ca	-0.759	3.727	-8.273
ca	-3.697	1.385	-8.623
ca	-7.425	0.950	-8.299
ca	-8.212	3.741	-5.869
ca	-10.289	3.171	-2.768
ca	-13.289	0.880	-2.764
ca	-11.406	-1.159	-5.305
ca	-9.007	-3.768	-3.961
ca	-8.805	-3.857	-0.182
ca	-6.898	-2.013	2.504
ca	-5.659	-5.038	4.305
ca	-1.906	-4.934	4.678
ca	0.751	-2.872	6.337
ca	2.410	0.455	6.860
ca	5.960	1.140	7.771
ca	8.233	3.952	6.970
ca	7.977	1.077	4.618
ca	9.800	-0.661	1.787
ca	7.038	-1.773	-0.369
ca	7.656	-2.526	-3.945
ca	7.314	-5.192	-6.478
ca	6.991	-7.643	-3.680
ca	3.484	-6.346	-3.267
ca	0.767	-4.072	-4.377
ca	2.660	-1.209	-2.780
ca	0.286	1.243	-1.299
ca	0.841	4.882	-0.690
ca	3.615	7.374	-0.079
ca	1.549	10.448	-0.232
ca	0.357	8.876	2.961
ca	0.854	11.816	5.218
ca	4.025	11.309	3.316
ca	7.243	9.551	4.198
ca	6.052	6.222	5.073
ca	5.003	3.304	2.919
ca	1.850	1.233	2.819
ca	1.311	-2.217	1.397
ca	-0.548	-5.285	0.634
ca	1.870	-8.000	-0.432
ca	1.761	-9.986	-3.630
ca	1.950	-13.509	-4.956

References

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) "Two-Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance," *J. Chem Phys.* **54**, 2229-2246.

Banks, B. E. C., Brown, C., Burgess, G. M., Burnstock, G., Claret, M., Cocks, T. M., & Jenkinson, D. H. (1979) "Apamin Blocks Certain Neurotransmitter-induced Increases in Potassium Permeability," *Nature* **282**, 415-417.

Barhanin, J., Hugues, M., Schweitz, H., Vincent, J. P., & Lazdunski, M. (1981) "Structure-Function Relationships of Sea Anemone Toxin II from *Anemonia sulcata*," *J. Biol. Chem.* **256**, 5764-5769.

Bax, A. & Davis, D. G. (1985) "MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy," *J. Magn. Reson.* **65**, 355-360.

Bax, A. & Drobny, G. (1985) "Optimization of Two-Dimensional Homonuclear Relayed Coherence Transfer NMR Spectroscopy," *J. Magn. Reson.* **61**, 306-320.

Beress, L. (1982) "Biologically Active Compounds from Coelenterates," *Pure & Appl. Chem.* **54**, 1981-1994.

Billeter, M., Braun, W., & Wüthrich K. (1982) "Sequential Resonance Assignments in Protein ^1H Nuclear Magnetic Resonance Spectra, Computation of Sterically Allowed Proton-Proton Distances and Statistical Analysis of Proton-Proton Distances in Single Crystal Protein Conformations," *J. Mol. Biol.* **155**, 321-346.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) "Selection of Coherence Transfer Pathways in NMR Pulse Experiments," *J. Magn. Reson.* **58**, 370-388.

Braunschweiler, L. & Ernst, R. R. (1983) "Coherence Transfer by Isotropic Mixing: Application to Proton Correlation Spectroscopy," *J. Magn. Reson.* **53**, 521-528.

Bundi, A. & Wüthrich, K. (1979) " ^1H -NMR Parameters of the Common Amino Acid Residues Measured in Aqueous Solutions of the Linear tetrapeptides H-Gly-Gly-X-L-Ala-OH," *Biopolymers* **18**, 285-297.

Bystrov, V. F., Arseniev, A. S., & Gavrilov, Yu. D. (1978) "NMR Spectroscopy of Large Peptides and Small Proteins," *J. Magn. Reson.* **30**, 151-184.

Bystrov, V. F., Okhanov, V. V., Miroshnikov, A. I., & Ovchinnikov, Y. A. (1980) "Solution Spatial Structure of Apamin as Derived from NMR Study," *FEBS Lett.* **119**, 113-117.

Chou, P. Y. & Fasman, G. D. (1978) "Prediction of the Secondary Structure of Proteins from their Amino Acid Sequence," *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45-148.

Cossand, W. L. & Merrifield, R. B. (1977) "Concept of Internal Structural Controls for Evaluation of Inactive Synthetic Peptide Analogs: Synthesis of [Orn-13,14] Apamin and its Guanidination to an Apamin Derivative with Full Neurotoxic Activity," *Proc. Natl. Acad. Sci. USA* **74**, 2771-2775.

Crippen, G. (1981) *Distance Geometry and Conformational Calculations*, Research Studies (Wiley), New York.

Davis, D. G. & Bax, A. (1985) "Assignment of Complex ^1H NMR Spectra via Two-Dimensional Homonuclear Hartmann-Hahn Spectroscopy," *J. Am. Chem. Soc.* **107**, 2820-2821.

De Marco, A., Llimas, M., & Wüthrich, K. (1978) "Analysis of the ^1H -NMR Spectra of Ferrichrome Peptides. II. The Amide Resonances," *Biopolymers* **17**, 637.

Dempsey, C. E. (1986) "pH Dependence of Hydrogen Exchange from Backbone Peptide Amides in Apamin," *Biochemistry* **25**, 3904-3911.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D., & Wemmer, D. (1979) "Fourier Transform Multiple Quantum Nuclear Magnetic Resonance," *Symp. Faraday Soc.* **13**, 49-55.

- Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) "Exploring Nuclear Spin Systems by Relayed Magnetization Transfer," *J. Am. Chem. Soc.* **104**, 3731-3732.
- Englander, S. W. & Kallenbach, N. R. (1984) "Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids," *Q. Rev. Biophys.* **16**, 521-655.
- Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, New York.
- Fogh, R. H., Mabbutt, B. C., Kem, W. R., & Norton, R. S. (1989) "Sequence Specific Assignments and Secondary Structure in the Sea Anemone Polypeptide *Stichodactyla helianthus* Neurotoxin I," *Biochemistry* **28**, 1826-1834.
- Freeman, C. M., Catlow, A. M., Hemmings, A. M., & Hider, R. C. (1986) "The Conformation of Apamin," *FEBS Lett.* **197**, 289-296.
- Gooley, P. R., Beress, L., & Norton, R. S. (1984) "¹H Nuclear Magnetic Resonance Spectroscopic Study of the Polypeptide Toxin I from *Anemonia sulcata*," *Biochemistry* **23**, 2144-2152.
- Gooley, P. R. & Norton, R. S. (1985) "Specific Assignment of Resonances in the ¹H Nuclear Magnetic Resonance Spectrum of the Polypeptide Cardiac Stimulant Anthopleurin-A," *Eur. J. Biochem.* **153**, 529-539.
- Gooley, P. R. & Norton, R. S. (1986a) "Secondary Structure in Sea Anemone Polypeptides: A Proton Nuclear Magnetic Resonance Study," *Biochemistry* **25**, 2349-2356.
- Gooley, P. R. & Norton, R. S. (1986b) "Specific Assignment of Resonances in the ¹H-NMR Spectrum of the Polypeptide Toxin I from *Anemonia sulcata*," *Biopolymers* **25**, 489-506.
- Granier, C., Pedroso Muller, E., & van Rietschon, J. (1978) "Use of Synthetic Analogs for a Study on the Structure-Activity Relationship of Apamin," *Eur. J. Biochem.* **82**, 293-299.

- Habermann, E. (1972) "Bee and Wasp Venoms," *Science* **177**, 314-322.
- Hare, D. R., Ribeiro, N. S., Wemmer, D. E., & Reid, B. R. (1985) "Complete Assignment of the Imino Protons of Escherichia coli Valine Transfer RNA: Two-Dimensional NMR Studies in Water," *Biochemistry* **24**, 4300-4306.
- Havel, T., Kuntz, I. D., & Crippen, G. M. (1983) "The Theory and Practice of Distance Geometry," *Bulletin of Mathematical Biology* **45**, 665-720.
- Havel, T. & Wüthrich, K. (1984) "A Distance Geometry Program for Determining the Structures of Small Proteins and Other Macromolecules from Nuclear Magnetic Resonance Measurements of Intramolecular ^1H - ^1H Proximities in Solution," *Bull. Math. Biol.* **46**, 673-698.
- Havel, T. & Wüthrich, K. (1985) "An Evaluation of the Combined Use of Nuclear Magnetic Resonance and Distance Geometry for the Determination of Protein Conformations in Solution," *J. Mol. Biol.* **182**, 281-294.
- Hider, R. C. & Ragnarsson, U. (1980) "A Proposal for the Structure of Apamin," *FEBS Lett.* **111**, 189-193.
- Hider, R. C. & Ragnarsson, U. (1981) "A Comparative Structural Study of Apamin and Related Bee Venom Peptides," *Biochem. Biophys. Acta* **667**, 197-208.
- Jeener, J. (1971) *Ampere International Summer School*, Basko polje, Yugoslavia.
- 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.
- Karplus, M. (1959) "Contact Electron-Spin Coupling of Nuclear Magnetic Moments," *J. Chem. Phys.* **30**, 11-15.

Kem, W. R. (1987) presented at the Natural Toxins from Aquatic and Marine Environments Meeting at Woods Hole, MA.

Kim, P. S. & Baldwin, R. L. (1984) "A Helix Stop Signal in the Isolated S-peptide of Ribonuclease-A," *Nature* **307**, 329-334.

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.

Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) "Buildup Rates of the Nuclear Overhauser Effect Measured by Two-Dimensional Proton Magnetic Resonance Spectroscopy: Implications for Studies of Protein Conformation," *J. Am. Chem. Soc.* **103**, 3654-3658.

Macura, S. & Ernst, R. R. (1980) "Elucidation of Cross Relaxation in Liquids by Two-Dimensional NMR Spectroscopy," *Mol. Phys.* **41**, 95-117.

Melnikov, P. N. & Popov, E. M. (1979) "An a priori Calculation of the Three-Dimensional Structure of Apamin. II. Fragment Cys¹-Leu¹⁰," *Bioorg. Khim.* **5**, 1011-1024.

Melnikov, P. N. & Popov, E. M. (1980) "An a priori Calculation of the Three-Dimensional Structure of Apamin. IV. Conformational Possibilities of the Side Chains of the Molecule," *Bioorg. Khim.* **6**, 21-30.

Metrione, R. M., Schweitz, H., & Walsh, K. A. (1987) "The amino acid sequence of toxin RplII from the sea anemone, *Radianthus paumotensis*," *FEBS Lett.* **218**, 59-62.

Miroshnikov, A. I., Elyakova, E. G., Kudelin, A. B., & Senyavina, L. B. (1978) "Physiochemical Characteristics of Apamin, a Neurotoxin from Honey Bee *Apis Mellifica*," *Bioorg. Khim.* **4**, 1022-1028.

Mitchinson, C. & Baldwin, R. L. (1986) "The Design and Production of Semisynthetic Ribonucleases with Increased Thermostability by Incorporation of S-peptide Analogs with Enhanced Helical Stability," *Proteins* 1, 23-33.

Nelson, J. W. & Kallenbach, N. R. (1989) "Persistence of the α -Helix Stop Signal in the S-peptide in Trifluoroethanol Solutions," *Biochemistry* 28, 5256-5261.

Norton, R. S., Cossins, A. I., & Kem, W. R. (1989) " ^1H NMR Study of the Solution Properties of the Polypeptide Neurotoxin I from the Sea Anemone *Stichodactyla helianthus*," *Biochemistry* 28, 1820-1826.

Norton, T. R. (1981) "Cardioactive Polypeptides from *Anthopleura xanthogrammica* and *A. elegantissima*," *Fed. Proc., Fed. Am. Chem. Soc. Exp. Biol.* 40, 21-25.

Okhanov, V. V., Afanas'ev, V. A., Gurevich, A. Z., Elyakova, E. G., Miroshnikov, A. I., Bystrov, V. F., & Ovchinnikov, Y. A. (1980) "Spatial Structure of Apamin in Solution. Conformational NMR Analysis," *Sov. J. Biorg. Chem.* 6, 417-437.

Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1986) "Origin of t_1 and t_2 Ridges in 2D NMR Spectra and Procedures for Suppression" *J. Magn. Reson.* 66, 187-193.

Pardi, A., Wagner, G., & Wüthrich, K. (1983) "Protein Conformation and Proton Nuclear Magnetic Resonance Chemical Shifts," *Eur. J. Biochem.* 137, 445-454.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) "Calibration of the Angular Dependence of the Amide Proton- C^α Proton Coupling Constants, $^3J_{\text{HN}\alpha}$, in a Globular Protein," *J. Mol. Biol.* 180, 741-751.

Pease, J. H. B. & Wemmer, D. E. (1988) "Solution Structure of Apamin Determined by Nuclear Magnetic Resonance and Distance Geometry," *Biochemistry* 27, 8491-8498.

Pease, J. H. B., Kumar, N. V., Schweitz, H., Kallenbach, N. R., & Wemmer, D. E. (1989) "NMR Studies of Toxin III from the Sea Anemone *Radianthus paumotensis*

and Comparison of Its Secondary Structure with Related Toxins," *Biochemistry* **28**, 2199-2204.

Popov, E. M. & Melnikov, P. N. (1979) "An a priori Calculation of the Three-Dimensional Structure of Apamin. I. Fragment Cys¹-Pro⁶," *Bioorg. Khim.* **5**, 828-847.

Popov, E. M. & Melnikov, P. N. (1979) "An a priori Calculation of the Three-Dimensional Structure of Apamin. III. Structure of the Molecule," *Bioorg. Khim.* **5**, 1471-1593.

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.

Rathmayer, W. (1979) "Sea Anemone Toxins: Tools in the Study of Excitable Membranes," *Adv. in Cytopharm.* **3**, 335-344.

Reimer, N. S., Yasunobu, C. L., Yasunobu, K. T., & Norton, T. R. (1985) "Amino Acid Sequence of the *Anthopleura xanthogrammica* Heart Stimulant, Anthopleurin-B," *J. Biol. Chem.* **260**, 8690-8693.

Richardson, J. S. (1981) "The Anatomy and Taxonomy of Protein Structure," *Adv. Protein Chem.* **34**, 167-339.

Romey, G., Abita, J. P., Schweitz, H., Wunderer, G., & Lazdunski, M. (1976) "Sea Anemone Toxin: A Tool to Study Molecular Mechanisms of Nerve Conduction and Excitation-Secretion Coupling," *Proc. Natl. Acad. Sci. USA* **73**, 4055-4059.

Rochat, H., Rochat, C., Sampieri, F., Miranda, F., & Lissitzky, S. (1972) "The Amino Acid Sequence of Neurotoxin II of *Androctonus australis* Hector," *Eur. J. Biochem.* **28**, 381-388.

Scheffler, J.-J., Tsugita, A., Linden, G., Schweitz, H., & Lazdunski, M. (1982) "The Amino Acid Sequence of Toxin V from *Anemonia sulcata*," *Biochem. Biophys. Res. Commun.* **107**, 272-278.

Schultz, G. E. & Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.

Schweitz, H., Bidard, J., Frelin, C., Pauron, D., Vijverberg, H. P. M., Mahasneh, D. M., Lazdunski, M., Vilbois, F., & Tsugita, A. (1985) "Purification, Sequence, and Pharmacological Properties of Sea Anemone Toxins from *Radianthus paumotensis*. A New Class of Sea Anemone Toxins Acting on the Sodium Channel," *Biochemistry* **24**, 3554-3561.

Sklenar, V. & Bax, A. (1987) "Spin-Echo Water Suppression for the Generation of Pure Phase Two Dimensional NMR Spectra," *J. Magn. Reson.* **74**, 469-479.

Smith, C. D., DeLucas, L., Ealick, S. E., Schweitz, H., Lazdunski, M., & Bugg, C. E. (1984) "Crystallization and Preliminary X-ray Investigation of a Protein Neurotoxin from the Sea Anemone *Anthopleura xanthogrammica*," *J. Biol. Chem.* **259**, 8010-8011.

Tanaka, M., Haniu, M., Yasunobu, K. T., & Norton, T. R. (1977) "Amino Acid Sequence of the *Anthopleura xanthogrammica* Heart Stimulant, Anthopleurin-A," *Biochemistry* **16**, 204-208.

Torba, A. E., Mabbutt, B. C., von Gunsteren, W. F., & Norton, R. S. (1988) "Backbone Folding of the Polypeptide Cardiac Stimulant Anthopleurin-A Determined by Nuclear Magnetic Resonance, Distance Geometry, and Molecular Dynamics," *FEBS Lett.* **239**, 266-270.

Vincent, J. P., Schweitz, H., & Lazdunski, M. (1975) "Structure-Function Relationships and Site of Action of Apamin, a Neurotoxic Polypeptide of Bee Venom with an Action on the Central Nervous System," *Biochemistry* **14**, 2521-2525.

- Wagner, G. & Wüthrich, K. (1982) "Amide Proton Exchange and Surface Conformation of the Basic Pancreatic Trypsin Inhibitor in Solution. Studies with Two-Dimensional Nuclear Magnetic Resonance," *J. Mol. Biol.* **160**, 343-361.
- Weber, P. L., Morrison, R., & Hare, D. (1988) "Determining Stereo-specific ^1H Nuclear Magnetic Resonance Assignments from Distance Geometry Calculations," *J. Mol. Biol.* **204**, 483-487.
- Wemmer, D. & Kallenbach, N. R. (1983) "Structure of Apamin in Solution: A Two-Dimensional Nuclear Magnetic Resonance Study," *Biochemistry* **22** 1901-1906.
- Wemmer, D. E., Kumar, N. V., Metrione, R. M., Lazdunski, M., Drobny, G., & Kallenbach, N. R. (1986) "NMR Analysis and Sequence of Toxin II from the Sea Anemone *Radianthus paumotensis*," *Biochemistry* **25**, 6842-6849.
- Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1984) "Homonuclear Two-Dimensional ^1H NMR of Proteins. Experimental Procedures," *J. Magn. Reson.* **56**, 207-234.
- Widmer, H., Wagner, G., Schweitz, H., Lazdunski, M., & Wüthrich, K. (1988) "The Secondary Structure of the Toxin ATX Ia from *Anemonia sulcata* in Aqueous Solution Determined on the Basis of Complete Sequence-Specific ^1H -NMR Assignments," *Eur. J. Biochem.* **171**, 177-192.
- Wlodawer, A., Bott, R., & Sjölin, L. (1982) "The Refined Crystal Structure of Ribonuclease A at 2.0 Å Resolution," *J. Biol. Chem.* **257**, 1325-1332.
- Wunderer, G., Fritz, H., Wachter, E., & Machleidt, W. (1976) "Amino Acid Sequence of a Coelenterate Toxin: Toxin II from *Anemonia sulcata*," *Eur. J. Biochem.* **68**, 193-198.
- Wunderer, G. & Eulitz, M. (1978) "Amino Acid Sequence of Toxin I from *Anemonia sulcata*," *Eur. J. Biochem.* **89**, 11-17.

Wüthrich, K. & Wagner, G. (1979) "Nuclear Magnetic Resonance of Labile Protons in the Basic Pancreatic Trypsin Inhibitor," *J. Mol. Biol.* **130**, 1-18.

Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) "Sequential Resonance Assignments as a Basis for Determination of Spatial Protein Structures by High Resolution Proton Nuclear Magnetic Resonance," *J. Mol. Biol.* **155**, 311-319.

Wüthrich, K. (1983) "Sequential Individual Resonance Assignments in the ^1H -NMR Spectra of Polypeptides and Proteins," *Biopolymers* **22**, 131-138.

Wüthrich, K., Billeter, M., & Braun, W. (1983) "Pseudo-Structures for the 20 Common Amino Acids for use in Studies of Protein Conformations by Measurements of Intramolecular Proton-Proton Distance Constraints with Nuclear Magnetic Resonance," *J. Mol. Biol.* **169**, 949-961.

Wüthrich, K., Billeter, M., & Braun, W. (1984) "Polypeptide Secondary Structure Determination by Nuclear Magnetic Resonance Observation of Short Proton-Proton Distances," *J. Mol. Biol.* **180**, 715-740.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Zell, A., Carson, M., Epps, M., Hermans, J., & Bugg, C. E. "Scorpion Toxins and Apamin have Similar Three-Dimensional Structures," personal communication, paper submitted.

Zykova, T. A., Vinokurov, L. M., Kozlovskaya, E. P., & Elyakov, G. B. (1986) "Amino Acid Sequence of Neurotoxin III from the Actina *Radianthus macrodactylus*," *Sov. J. Bioorg. Chem.* **11**, 159-165.

*LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720*