Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

STRUCTURAL BIOLOGY DIVISION

Structures and Related Properties of Helical, Disulfide-Stabilized Peptides

M.D. Pagel
(Ph.D. Thesis)

October 1993

JAN 14 1994

OSTI

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. 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 liability or 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 and shall not be used for advertising or product endorsement purposes.

Lawrence Berkeley Laboratory is an equal opportunity employer.

LBL-34903
UC-408

Structures and Related Properties
of Helical, Disulfide-Stabilized Peptides

by

Mark David Pagel
Ph.D. Thesis

Department of Chemistry
of the University of California at Berkeley

and

Structural Biology Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

November, 1993

This work was supported, in part, by the Director, Office of Energy Research, Office
of Health Effects Research, of the U.S. Department of Energy under Contract No.
DE-AC03-76F00098

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

870

**Structures and Related Properties of
Helical, Disulfide-Stabilized Peptides**

Copyright 1993

by

Mark D. Pagel

**The Government reserves for itself and
others acting on its behalf a royalty free,
nonexclusive, irrevocable, world-wide
license for governmental purposes to publish,
distribute, translate, duplicate, exhibit,
and perform any such data copyrighted by
the contractor.**

Abstract

Structures and Related Properties of Helical, Disulfide-Stabilized Peptides

by

Mark D. Pagel

Doctor of Philosophy in Chemistry

University of California at Berkeley

Professor David E. Wemmer

The three dimensional structure of several peptides were determined by NMR spectroscopy and distance geometry calculations. Each peptide formed a predictable, rigid structure, consisting of an α -helix, a "scaffold" region which packed along one face of the helix, and two disulfide bridges which covalently connect the helix and scaffold regions.

The peptide Apa-M5 was designed to constrain the M5 peptide from MLCK in a helical geometry using the apamin disulfide scaffold. This scaffold constrains the N-terminal end of the helix with two disulfide bridges and a reverse turn. Like the M5 peptide, Apa-M5 was found to bind calmodulin in a Ca^{2+} -dependent 1:1 stoichiometry. However, the dissociation constant of the (Apa-M5)-calmodulin complex, 107 nM, was 100-fold higher than the dissociation constant of the M5-calmodulin complex. This difference was due to a putative steric overlap between the Apa-M5 scaffold and calmodulin.

The peptide Apa-Cro was designed to replace the large structural protein matrix of λ Cro with the apamin disulfide scaffold. However, Apa-Cro did not bind the consensus DNA operator half-site of λ Cro, probably due to a steric overlap between the Apa-Cro disulfide framework and the DNA.

The amino acid sequence of the scaffold-disulfide bridge arrangement of the peptide Max was derived from the core sequence of scyllatoxin, which contains an α -helix constrained at the C-terminal end by two disulfide bridges and a two-stranded β -sheet

scaffold. Max was shown to fold with >84% yield to form a predictable, stable structure that is similar to scyllatoxin. The folding and stability properties of Max make this scaffold and disulfide bridge arrangement an ideal candidate for the development of hybrid sequence peptides.

The dynamics of a fraying C-terminal end of the helix of the peptide Apa-Ala^N was determined by analysis of ¹⁵N NMR relaxation properties. The analysis demonstrated that the stable helix extends three residues beyond the last disulfide bridge of the apamin disulfide scaffold, while the C-terminal end of the helix exhibited dynamic transitions between helix and random coil states.

Chair

Date

11/22/93

This work is dedicated
to my father, Paul J. Pagel,
and to my grandfather, Clarence A. Pagel
both of whom showed by example
how accomplishments arise
from hard work and long hours,

and to my mother, Doris J. Pagel,
who showed that the greatest skill
is understanding.

Table of Contents

List of Figures.....	vi
List of Tables.....	ix
Acknowledgements.....	x
Chapter 1. Introduction	
1.1 Protein Engineering.....	1
1.2 The α -helix.....	3
1.3 NMR Spectroscopy.....	9
Chapter 2. General Methods and Materials	
2.1 Solid-Phase Peptide Synthesis.....	12
2.2 High-Performance Liquid Chromatography.....	18
2.3 NMR Spectroscopy.....	30
2.4 Distance Geometry.....	49
Chapter 3. Apa-M5	
3.1 Introduction.....	55
3.2 Peptide Design.....	56
3.3 Peptide Structure.....	56
3.4 Peptide Function.....	65
3.5 Molecular Modeling.....	72
Chapter 4. Apa-Cro	
4.1 Introduction.....	79
4.2 Peptide and DNA Design.....	81
4.3 Peptide Structure.....	86
4.4 DNA Structure.....	90
4.5 Peptide Function.....	93
4.6 Molecular Modeling.....	96

Chapter 5. Max

5.1 Introduction.....	103
5.2 Peptide Design.....	105
5.3 Peptide Folding.....	109
5.4 Peptide Structure.....	111
5.5 Molecular Modeling.....	120
5.6 Hybrid Sequence Peptides.....	122

Chapter 6. Apa-AlaN

6.1 Introduction.....	127
6.2 Peptide Design.....	129
6.3 Peptide Folding.....	131
6.4 Peptide Structure.....	134
6.5 Theory of Peptide Dynamics.....	141
6.6 Experimental Peptide Dynamics.....	147

References.....	153
-----------------	-----

Appendices

1. Peptide Synthesis.....	177
2. NMR Pulse Programs.....	182
3. Distance Constraints.....	186
4. Distance Constraints.....	189

List of Figures

Chapter 1

Figure 1.1. The ideal α -helix.....	4
--	---

Chapter 2

Figure 2.1. General protocol for solid-phase peptide synthesis.....	13
Figure 2.2. Theoretical yield from incomplete coupling.....	14
Figure 2.3. Theoretical yield of all polymeric lengths.....	16
Figure 2.4. Experimental yield from cleavage of PRGRP-NH ₂	17
Figure 2.5. Schematic of a HPLC chromatogram.....	19
Figure 2.6. Dependence of resolution upon relative retention.....	21
Figure 2.7. Examples of solute migration along column in gradient HPLC.....	24
Figure 2.8. HPLC chromatograms of Apa-Cro.....	27
Figure 2.9. Schematic of DQFCOSY, RELAY, and TOCSY spectra.....	35
Figure 2.10. Methyl region of 2QCOSY spectrum of Max.....	37
Figure 2.11. Schematic representation of NOESY-HMQC spectrum.....	38
Figure 2.12. One dimensional spectrum of Apa-M5.....	42
Figure 2.13. Chemical shifts of helical residues.....	43
Figure 2.14. Distances within helices which give rise to NOE intensities.....	45
Figure 2.15. NH chemical shift temperature coefficients of Apa-Ala.....	46
Figure 2.16. J _{NH-αH} of residues of Apa-Ala.....	47

Chapter 3

Figure 3.1. Amino acid sequences of Apa-M5, apamin, M5, and M13.....	57
Figure 3.2. Edmundson helical wheel projection of Apa-M5.....	58
Figure 3.3. Circular dichroism spectrum of Apa-M5.....	59
Figure 3.4. Temperature dependence of Apa-M5 CD intensity.....	61
Figure 3.5. NH region of NOESY spectrum of Apa-M5.....	62
Figure 3.6. Summary of observed connectivities of Apa-M5.....	63

Figure 3.7. 20 distance geometry structures of Apa-M5.....	66
Figure 3.8. Comparison of structures of Apa-M5 and apamin.....	67
Figure 3.9. CD spectra of (Apa-M5)-calmodulin complex.....	68
Figure 3.10. Fluorescence intensity of Apa-M5 with calmodulin.....	70
Figure 3.11. Fluorescence intensity of Apa-M5 with calmodulin and CBD.....	71
Figure 3.12. Structures of Apa-M5 and the M13-calmodulin complex.....	74
Figure 3.13. Amino acid sequences of M5, Apa-M5, and Apa-M60.....	75
Figure 3.14. Edmundson helical wheel projection of Apa-M60.....	77
Figure 3.15. Structures of Apa-M60 and the M13-calmodulin complex.....	78

Chapter 4

Figure 4.1. Model of the complex between λ -Cro and the OR3 5' half-site.....	80
Figure 4.2. Sequences of apamin, Apa-Cro, and HTH protein recognition helices	82
Figure 4.3. Illustration of interactions between λ -Cro and the OR3 5' half-site.....	84
Figure 4.4. Edmundson helical wheel projection of Apa-Cro.....	85
Figure 4.5. Nucleic acid sequences of OR3/2 and OR3.....	87
Figure 4.6. NH region of NOESY spectrum of Apa-Cro.....	88
Figure 4.7. Summary of observed connectivities of Apa-Cro.....	89
Figure 4.8. 20 distance geometry structures of Apa-Cro.....	91
Figure 4.9. Comparison of structures of Apa-Cro, Apa-M5, and apamin.....	93
Figure 4.10. Imino region of one dimensional spectrum of OR3/2.....	94
Figure 4.11. Imino and aromatic regions of OR3/2 spectrum with Apa-Cro.....	95
Figure 4.12. Structures of Apa-Cro and the (434-Cro)-OR1 complex.....	97
Figure 4.13. Sequences of Apa-Cro, Apa-Crotate, and λ Cro recognition helix.....	99
Figure 4.14. Edmundson helical wheel projection of Apa-Crotate.....	100
Figure 4.15. Structures of Apa-Crotate and the (434-Cro)-OR1 complex.....	102

Chapter 5

Figure 5.1. Structures of apamin, Max, and Apa-Max.....	105
---	-----

Figure 5.2. Molecular model of nimapa.....	107
Figure 5.3. Amino acid sequences of Max, ScyTx, ChTx, and <i>CsEv3</i>	108
Figure 5.4. HPLC chromatograms of Max, reduced Max, and reoxidized Max.....	111
Figure 5.5. CD spectrum of Max.....	113
Figure 5.6. Temperature dependence of Max CD intensity.....	114
Figure 5.7. NH region of NOESY spectrum of Max.....	115
Figure 5.8. Summary of observed connectivities of Max.....	116
Figure 5.9. chemical shift differences of Max and ScyTx amide protons.....	119
Figure 5.10. 20 distance geometry structures of Max.....	120
Figure 5.11. Comparison of Max and <i>CsE v3</i> structures.....	122
Figure 5.12. Helical net diagrams of Max and <i>CsE v3</i>	124
Figure 5.13. View of Max structure along helix axis.....	126

Chapter 6

Figure 6.1. Sequences of bee venom peptides and hybrid sequence peptides.....	131
Figure 6.2. HPLC chromatograms of folded forms of Apa-AlaN analogs.....	133
Figure 6.3. CD spectra of Apa-Ala.....	136
Figure 6.4. Temperature dependence of Apa-Ala CD intensity.....	138
Figure 6.5. NH region of NOESY spectrum of Apa-AlaN.....	139
Figure 6.6. Strip plot of NOESY-HMQC spectrum of Apa-AlaN.....	140
Figure 6.7. Summary of observed connectivities of Apa-AlaN.....	141
Figure 6.8. Model of motions on time scales τ_f and τ_s	146
Figure 6.9. Dependence of T ₁ , T ₂ , and NOE upon molecular tumbling.....	147
Figure 6.10. ¹⁵ N T ₁ and T ₂ data and relaxation times of Apa-AlaN.....	149
Figure 6.11. The relationship between T ₁ /T ₂ and τ_R	150
Figure 6.12. S _f ² and (S _{f,i} /S _{f,i-1}) ² of Apa-AlaN.....	152

List of Tables**Chapter 1**

Table 1.1. Helix propensities of amino acids.....	7
---	---

Chapter 2

Table 2.1. Optimized HPLC conditions.....	29
---	----

Chapter 5

Table 5.1. Peptide backbone RMS deviations of Max DGII structures.....	121
--	-----

Chapter 6

Table 6.1. Folding products of analogs of Apa-AlaN.....	134
---	-----

Acknowledgements

I would like to thank David Wemmer for his objective guidance and motivation throughout the research presented in this dissertation. I would like to thank Mark Kubinec and Brian Volkman for many helpful discussions concerning this research. I would like to thank my colleagues associated with the Laboratory of Chemical Biodynamics NMR Spectroscopy Facility, especially Dave, Mark, Brian, Valerie Copie, Fred Damberger, Tammy Dwyer, Patty Fagan, Bernhard Geierstanger, Debbie Kallick, Andrew Lee, Young Lee, Ping-Chiang Lyu, Joe Pease, Jeff Pelton, Beth Snowden-Ifft, Pete Spielman, Rick Storrs, Milan Tomic, and Milton Werner, for their assistance with NMR spectroscopy, and for the opportunities to assist them in return.

I would like to thank Laurie Princiotto for her encouragement, support, and for proofreading this dissertation.

I would like to thank Prof. R. E. Klevit for providing the *Drosophisia* calmodulin, Prof. P. G. Schultz for access to his fluorimeter, Prof. I. Tinoco for access to his CD spectropolarimeter, Dr. A. Bax for access to the calmodulin-M13 NMR structure coordinates Dr. D. King for synthesis of Apa-Alan, and Dr. M. Tomic for helpful discussions regarding Apa-Cro. This work was supported by the National Science Foundation, 88 15998, and through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134. In part supported by the Director, Office of Health Effects Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Finally, I would like to thank Café Strada, which provided the grounds for many stimulating conversations related to this work, as summerized below.

Café Rendezvous

I haven't heard from liver, kidney, or spleen,
but the rest of my body needs the caffeine
from the brown bitter brew of the coffee bean
to keep me from the desire to dream.

I may be jittery, the lights seem bright,
my physical condition may not be right,
but a café rendezvous offers respite
from sleepily wasting the day and night.

Chapter 1: Introduction

1.1 Protein Engineering

Protein engineering involves the application of relationships between protein functions, structures, and primary amino acid sequences towards the development of new and improved proteins (Wetzel, 1986). These relationships between primary sequences, structures, and functions are poorly understood. At the most fundamental level, protein function is directly determined by structure, although other provisions, such as the presence of ligands, may also be required (Fersht, 1977). At the same fundamental level, protein structure is directly determined by primary sequence, although other conditions, such as the chemistry of the surrounding medium, also influence protein structure (Schultz & Schirmer, 1979). Yet a variety of simple experiments have been conducted which result in the engineering of new and improved proteins.

Each scientific experiment requires a hypothesis, in which an antecedent predicts a conditional. In protein engineering experiments, the primary sequence, structure, and function may serve as the antecedent or conditional, making possible six types of hypotheses. However, only four of these types of hypotheses are based upon the fundamental relationships of protein engineering. Since function is not directly dependent upon primary sequence, these non-fundamental experiments are of less value to protein engineering. For example, an experiment was conducted in which an improvement in function, subtilisin activity in the presence of oxidants, was hypothesized from a change in primary sequence, the substitution of an aliphatic amino acid for an oxidatively sensitive methionine residue (Estell, et al., 1985). As hypothesized, this new protein exhibited function in the presence of hydrogen peroxide, an improvement over native subtilisin. However, the fundamental basis for this improvement was not determined. Improvement in function may have occurred by stabilization of the structure, or by a change in the

chemical interactions at the protein's active site. Thus, the applicability of this improvement towards engineering other proteins is of limited value.

Experiments based upon one of the types of hypotheses, in which the conditional is determined by the antecedent at the fundamental level of relationships, usually yield one conclusion. These experiments often use the hypothesis that one and only one structure will be determined from a known primary sequence, and that one and only one function will be determined from a known structure. These experiments often lead to the improvement of the function of existing proteins, but rarely create new proteins. These experiments are similar to the chemical processes of biological genetics, in which small changes of single residues, insertions and deletions of several residues, gene doubling, and gene fusion most often result in small improvements in proteins.

Conversely, experiments using one of the types of hypotheses in which the antecedent is determined by the conditional may reach more than one conclusion. These experiments use the hypothesis that a primary sequence may form a known structure or the hypothesis that a structure may perform a known function. These experiments usually do not use the hypothesis that the conditional primary sequence or structure is the only possible solution. Indeed, it is usually accepted before these experiments are conducted that a wide variety of structures may perform a known function, or that many primary sequences may form a known structure.

The experiments which can yield more than one conclusion are the most valuable towards the engineering of new proteins. The results of these experiments are the same as the results of quasi-random biological evolution of proteins, the results of numerous natural experiments which have given rise to an awesome diversity of conclusions. The possibility of more than one conclusion from evolution and protein engineering experiments does not imply that all conclusions are equal. For example, evolution has selected a primary sequence for hemoglobin which resists malaria, but which also results in sickle cell anemia. The selection of the best conclusion drives evolution and protein engineering.

Mutagenesis and expression of oligonucleotides and automated solid-phase peptide synthesis are applied to prepare large quantities of a variety of proteins. X-ray crystallography and NMR spectroscopy are powerful methods to determine protein structures at atomic resolution. These technological conveniences often predicate the antecedent of an experiment's hypothesis. However, protein engineering should not be limited by technological convenience, since all four fundamental experiments provide important information.

This dissertation involves the two types of protein engineering experiments in which the antecedent of the hypothesis is the peptide structure. Chapters 3 and 4 hypothesize function based upon a known structure. Chapters 5 and 6 hypothesize primary sequences which form a known structure. Thus, both of the fundamental relationships in protein engineering are studied, each of these experiments takes advantage of the powerful techniques of automated peptide synthesis and NMR spectroscopy, and the engineering of both new and improved proteins is investigated.

1.2 The α -Helix

Each of the peptide structures discussed in this dissertation contains an α -helix. The α -helix is the most abundant secondary structure in proteins (Schultz & Schirmer, 1979), and more than one-third of all residues in known protein structures are considered to be part of an α -helix (Creighton, 1983). The ideal α -helix, first postulated by Linus Pauling and Robert Corey, consists of 3.65 residues per right-handed turn, a pitch of 5.5 Å per turn, a diameter of 4.6 Å, and ϕ and ψ backbone angles of -48 degrees and -57 degrees, respectively (Pauling, et al., 1951) (Figure 1.1). This postulation assumed that the peptide bond is planar, and that the α -helix axis is linear. Furthermore, the stability of the helical structure was based upon the hydrogen bond between amide carbonyl of each residue and the amide proton which is four residues closer to the carboxylate terminus (the

Figure 1.1. Stereoview of backbone structure of the ideal α -helix.

$i, i+4$ hydrogen bond). While the enthalpy of this hydrogen bond, approximated at 5 to 7 kcal/mol for vapor-phase model systems, exceeds the conformational energies for rotation about single bonds, about 1 kcal/mol, Pauling and Corey speculated that the ideal α -helix should contain the strongest hydrogen bonds to compensate for destabilizing enthalpic and entropic energies. Thus, the postulation of the α -helix also assumed that the helix backbone hydrogen bonds are as short and linear as possible.

These assumptions, upon which the ideal helix is based, are not exact in α -helices of known protein structures. Helical residues that are exposed to solvent have carbonyl groups tilted an average of 0.2 Å outwards from the helix axis, relative to the more linear carbonyl groups of solvent-shielded residues (Glover, et al., 1983). While this lengthens and bends the hydrogen bond between the exposed carbonyl and helix amide proton, a second hydrogen bond can be made between the exposed carbonyl and the solvent or side chains (Watson, 1969; Glover, et al., 1983). Furthermore, the side chains of solvent-shielded helix residues must pack with the rest of the protein (Chothia, 1984). The packing of the α -helix with a second helix or a β -sheet necessitates systematic repositioning of α carbons, if the helix packing continues beyond two turns. The net effect of these distortions is reflected in the differences of average backbone angles of solvent-exposed helix residues ($\phi_{i+1} = -66$ degrees, $\psi_i = -41$ degrees) versus solvent-shielded helix residues ($\phi_{i+1} = -59$ degrees, $\psi_i = -44$ degrees) (Barlow & Thornton, 1988). This translates to a curvature of the helix axis with an average radius of curvature of 60 Å, corresponding to approximately one degree of arc per residue. Thus, if linear and curved three-turn helices are superposed at the amino terminus, the axis-to-axis separation at the carboxylate terminus approaches 5 Å.

This variability in helix geometry allows for different definitions of a helical residue, based upon ϕ and ψ backbone angles (IUPAC-IUB Commission, 1970), hydrogen bonding (IUPAC-IUB Commission, 1970; Kabsch & Sander, 1983), or α carbon positions (Levitt & Greer, 1977). It is not often recognized that the correct criteria

should be chosen to suit an application's purpose, rather than choosing a universal definition for an α -helix. For instance, when studying the stability of helical structures, hydrogen bonding more accurately defines a helix; when studying side chain packing or chemical functions, α carbon positions are more important. Both of these definitions will be used to describe helices in the following chapters.

Both Zimm-Bragg theory (Zimm & Bragg, 1959) and host-guest data (Sueki, et al., 1984) predict that the random coil, or unfolded state, is thermodynamically favored by 2 to 5 kcal/mole over the folded α -helix in peptides less than 20 residues in length in water.

The stabilizing enthalpy of the $i, i+4$ hydrogen bonds within the helix backbone is compensated by the enthalpy of hydrogen bonds between water and the backbone amide proton or carbonyl groups in the random coil. This competition by solvent hydrogen bonds is demonstrated by the formation of helices by short peptides in solvents containing trifluoroethanol (TFE). This solvent with a lower dielectric constant reduces the enthalpy of TFE-backbone amide proton hydrogen bonds (Presta & Rose, 1988; Storrs, et al., 1992; Sönnichsen, et al., 1992), although the stabilizing effects of TFE are entropic as well (Conio, et al., 1970). The unfavorable loss in entropy upon folding further contributes to the destabilization of short α -helices. The presence of tertiary contacts in larger proteins compensate for this loss of entropy. For example, short peptide fragments from sperm whale myoglobin show very little helicity in water, while the intact protein is almost entirely α -helix (Epand & Scheraga, 1968).

Several sequence-dependent properties of peptides have recently been reported which may stabilize the α -helical conformation of short peptides in water. These properties are used to design the peptides discussed in chapter 6. Designs of all short helical peptides must use amino acids with high helix propensity (Table 1.1). The frequencies of occurrence of each amino acid within secondary structures in proteins have been tabulated, and these frequencies have been interpreted as a conformational propensity (Chou & Fasman, 1974; Levitt, 1978). The effect of substituting a "guest" residue into several

Table 1.1. Helix propensities of amino acids

Residue, P_{α}^1	Residue P_{α}^2	Residue, P_{α}^3	Residue, P_{α}^4	Residue, $P_{\alpha x}/P_{\alpha A}^5$
E 1.53	M 1.47	E 1.35	A 1.60	A 1.00
A 1.45	E 1.44	M 1.20	L 1.50	L 0.93
L 1.34	L 1.30	I 1.14	F 1.45	M 0.95
H 1.24	A 1.29	L 1.14	M 1.44	Q 0.81
M 1.20	Q 1.27	W 1.11	W 1.34	I 0.71
Q 1.17	K 1.23	F 1.09	I 1.31	V 0.70
W 1.14	H 1.22	A 1.07	R 1.25	N 0.56
V 1.14	C 1.11	R 1.03	Q 1.22	T 0.56
F 1.12	F 1.07	Y 1.02	E 1.18	S 0.52
K 1.07	D 1.04	Q 0.98	V 1.09	G 0.45
I 1.00	W 0.99	V 0.95	K 1.05	
D 0.98	I 0.97	K 0.94	D 1.03	
T 0.82	R 0.96	C 0.92	H 0.97	
S 0.79	V 0.91	H 0.85	N 0.80	
R 0.79	N 0.90	T 0.82	R 0.79	
C 0.77	S 0.82	N 0.78	C 0.66	
N 0.73	T 0.82	D 0.78	Y 0.61	
Y 0.61	Y 0.72	S 0.76	G 0.59	
P 0.59	G 0.56	P 0.66	S 0.44	
G 0.53	P 0.52	G 0.59	P 0.19	

1. From Chou & Fasman, 1974.
2. From Levitt, 1978.
3. From Sueki, et al., 1984; Padmanabhan, et al., 1990.
4. From O'Neil & DeGrado, 1990.
5. From Lyu, et al., 1990.

different "host" peptides has been interpreted as propensities (Sueki, et al., 1984; Padmanabhan, et al., 1990; Lyu, et al., 1990; O'Neil & DeGrado, 1990). These propensities are in qualitative agreement, and correlate well with amino acid chemical structure. Amino acids with high helical propensity do not have branched β carbon or short polar side chains.

To determine the effects of charged termini upon helix destabilization, the charge of the end group of the helical C-peptide of ribonuclease A was varied in a series of new peptides (Fairman, et al., 1989). A free amino N-terminus or carboxylate C-terminus was found to destabilize the helix, while an acylated N-terminus or amidated N-terminus stabilized the helix. A methylester C-terminus was also found to stabilize the helix, indicating that the stabilization depends upon electrostatic interactions instead of hydrogen

bonding. This simplifies the design of stable helices, since the electrostatic interactions require a less stringent geometry than hydrogen bonds.

Salt bridges between charged side chains of $i, i+4$ residues also stabilize α -helices, due to the stereochemistry of salt bridge formation and (side chain)-(side chain) van der Waals contacts (Perutz & Fermi, 1988). Helix stabilization is greatest if the negatively charged residue of the salt bridge is placed nearer the N-terminus (Marqusee & Baldwin, 1987). A salt bridge between glutamate and arginine provides the most stabilization, and a glutamate-lysine salt bridge was the next-best stabilizer (Merkuta, et al., 1991). A minimum of three salt bridges is required to stabilize a helical peptide (Merkuta & Stellwagen, 1991).

Native proteins often form disulfide bridges, which decrease the loss of entropy upon folding. Non-native disulfide bridges have been substituted into proteins to stabilize helices (Pace & Creighton, 1986; Sauer, et al., 1986; Mitchinson & Wells, 1989). In this dissertation, two similar, general strategies are presented which may be used to design α -helical peptide stabilized by disulfide bridges. The design of the helix is not limited to specific terminal groups, or amino acids with high helix propensity or side chains which form salt bridges. Thermal denaturation studies indicate that the first two helical turns of these disulfide-stabilized peptides are extremely stable. This is in contrast to all short helical peptides without stabilizing disulfide bridges, which are helical only at low temperatures.

While the structure of the α -helix has been studied extensively, the dynamics of the helix are not well characterized. In particular, slower motions of helices, such as conformational exchange between helix and random coil conformations, occur more rapidly than a millisecond, a time scale of motion that is difficult to characterize (Jardetzky & Roberts, 1981). Recently developed NMR experiments have measured the fast and intermediate dynamics of helices within protein tertiary structures. Chapter 6 describes the application of these NMR experiments towards the study of α -helix dynamics. Unlike

previous NMR experiments, the α -helix of this study is independent of tertiary contacts normally present in proteins. The results of chapter 6 provide the first description of peptide α -helix dynamics using natural amino acids in water.

1.3 NMR Spectroscopy

NMR spectroscopy was first applied towards the study of peptides, proteins, and other biomolecules in an attempt to measure hydration of DNA (Jacobson, et al., 1954). The first successful direct observation of a biomolecule by NMR was the proton spectrum of ribonuclease at 40 MHz (Saunders, et al., 1957). However, the first investigations of biomolecular structures and dynamics were difficult due to limitations in radio-frequency electronics, computer hardware, and suitable samples for NMR experimentation. The development of Fourier transform NMR spectroscopy in the late 1960's (Ernst & Anderson, 1966; Farrar & Becker, 1971) and two dimensional NMR experiments in the 1970's (Jeener, 1971; Aue, et al., 1976a,b) provided a revolution in sophistication of NMR spectroscopy techniques. These techniques were quickly adapted for studies of biomolecules (Nagayama, et al., 1977; Wider, et al., 1984). Improvements in NMR experiments throughout the 1980's has developed NMR spectroscopy into a powerful, routine methodology for studying the structures and dynamics of a variety of biomolecules (Wüthrich, 1986; Clore & Gronenborn, 1989).

NMR spectroscopy offers several advantages over other techniques used to study structure. Most importantly, NMR spectroscopy can be used to determine the three-dimensional structures of proteins at atomic resolution. Structures are studied in solution and the technique can accommodate a wide variety of solvent compositions and conditions. Interactions between biomolecules can be studied easily and quickly. Dynamic events can also be studied, from high frequency motions, such as rotational tumbling of the biomolecule (Kay, et al., 1989; Peng, et al., 1991), through intermediate motions

associated with the breaking and formation of hydrogen bonds (Englander & Kallenbach, 1984), to slow motions, such as exchange between different conformations (Clare, et al., 1990). All of these advantages of NMR spectroscopy are employed in this dissertation.

The abundance of protons within proteins, the close proximity of these protons to each other, and the high NMR sensitivity of protons relative to other nuclei make proton NMR spectroscopy particularly useful for the study of proteins. Each of these protons give rise to a magnetic resonance, which must be resolved to measure spectroscopic parameters and to assign each resonance to specific protons within the protein. NMR spectroscopy of all proteins is inherently limited by the number of proton resonances that can be resolved within a narrow range of resonance frequencies. Thus, the size of a monomeric protein that can be studied by proton NMR spectroscopy is limited to roughly 100 residues or less (Wemmer & Reid, 1985). Several solutions have been successfully implemented to circumvent or reduce the large the number of proton resonances present in a typical protein spectrum.

Recent advances in molecular biology can be used to construct proteins with the uncommon isotopes ^{13}C and/or ^{15}N (McIntosh & Dahlquist, 1990). These isotopes may be incorporated uniformly throughout the protein, or may be isolated to specific residues or atoms. NMR experiments can be performed in which only protons attached to specific isotopes are observed (Griffey & Redfield, 1987; Bax, et al., 1990). Three dimensional NMR experiments can be performed in which protons at the same resonance frequency in a two dimensional proton spectrum are separated along a third dimension by the frequencies of their attached ^{13}C or ^{15}N nuclei (Griesinger, et al., 1989; Fesik & Zuiderweg, 1990). Both of these types of NMR experiments are used in chapter 6 to separate unresolved amide proton resonances of specifically labeled residues.

In many proteins, the information of interest to the protein engineer may be limited to a small region of a large structure. Much effort is spent in the attempt to isolate regions of proteins which have a size acceptable for NMR spectroscopy while still retaining

necessary structural or functional characteristics. Chapters 3 through 6 of this dissertation describe two peptide motifs used to stabilize isolated, α -helical regions of proteins. The small size of these peptide motifs allow for convenient study of structure, function, and dynamics by NMR spectroscopy. The routine, systematic application of these peptide motifs towards the stabilization of other α -helical regions for study by NMR spectroscopy is also presented.

Chapter 2: General Methods and Materials

2.1 Solid Phase Peptide Synthesis

Synthesis of peptides was conducted using the established techniques of tBoc, Fmoc (HOBT), and Fmoc (HBTU) solid phase peptide synthesis (Model 430A Peptide Synthesizer Manual). A general protocol for solid phase peptide synthesis is outlined in Figure 2.1. The chemical reagents used in the various chemistries, details of reaction conditions, and the specific protocol for each peptide synthesis are listed in Appendix 1.

Solid phase peptide synthesis is designed to produce large quantities of pure peptide. Yields from synthesis, purification, and folding of peptides in this dissertation ranged from 10 mg to 30 mg with purity greater than 95%. The final peptide yield for each residue depends upon the coupling step yield. As the average percent yield of coupling for all residues decreases, the overall yield of the final peptide quickly diminishes (Figure 2.2). Coupling can be especially difficult for residues 5 to 20 and for the coupling of amino acids to residues which both have a large side chain and protecting group (Clark-Lewis & Kent, 1985). To increase the average coupling efficiency, the synthesis protocol was modified to include a double coupling step (Kent, 1980). After performing a coupling reaction, the peptide chain was washed, and a second coupling reaction was performed with no intermediate amino-deprotection step. If the efficiency of the first and second coupling steps for each amino acid coupling is equal, then the overall peptide yield is significantly improved (Figure 2.2). In practice, the yield of the second coupling is usually much less than the first coupling (D. King, personal communication). Thus, double coupling substantially improves overall peptide yield for only those couplings of low yield. In Fmoc chemistry, the types of amino acids which couple with low yield were identified by mass spectrometry for several apamin hybrid-sequence peptides; these amino acids were double coupled for all Fmoc apamin hybrid-sequence peptide syntheses except Apa-Ala^N. In tBoc

Figure 2.1. An outline of the general solid phase synthesis protocol. The established technique is outlined by solid arrows, while modifications are outlined by dashed arrows.

Figure 2.2. The theoretical yield from incomplete coupling of a 27-residue peptide. The yield of both coupling steps in double coupling are equal.

chemistry, amino acid coupling yields were not analyzed, so all amino acids were double coupled.

A capping reaction was also added to the synthesis protocol to increase the peptide purity (Mitchell, et al., 1978). After a coupling reaction, remaining uncoupled peptide chains were reacted with acetic anhydride to form an acetyl peptide terminus. In subsequent coupling and deprotection reactions, this methylamide terminus is unreactive, leading to truncated peptide chains after final cleavage. This capping step does not affect the overall yield of the fully synthesized peptide sequence, since a fully synthesized peptide cannot exist as an uncoupled peptide chain following a coupling reaction. Rather, capping affects the distribution of yields of peptide sequences which are missing at least one residue (Figure 2.3). Most importantly, the yield of peptide sequences which are close in length (and thus similar in sequence) to the fully synthesized peptide is decreased. Since the quality of peptide purification is inversely related to the similarity of the desired peptide to undesired peptide contaminants (Hearn, 1983), capping during synthesis facilitates purification of the fully synthesized peptide. In this dissertation, a capping step was added to the tBoc synthesis protocol, since side reactions using tBoc synthesis and HF cleavage chemistries typically lead to difficulties in peptide purification. Capping was not performed during Fmoc chemistry, since purification of peptides synthesized with Fmoc chemistry was not problematic.

The cleavage and final deprotection of synthesized peptides using tBoc chemistry was performed with hydrogen fluoride by Applied Biosystems, Inc. (Sakakibara & Yasutsugu, 1965; Tam, et al., 1983). The cleavage and final deprotection of synthesized peptides using Fmoc chemistry was conducted using a trifluoroacetic acid cocktail (King, et al., 1990). The kinetic rate of the TFA cleavage reaction was investigated for a small peptide with the sequence PRGRP-NH₂. A reaction time of six hours was found to produce the highest yield of cleaved, deprotected peptide (Figure 2.4). This optimal reaction time was (unsystematically) observed for cleavage reactions of the peptides

Figure 2.3. The theoretical yield of all polymeric lengths from the synthesis of a 27-residue peptide. The yield per single coupling step was set to 94%, the experimental yield per coupling step of the Apa-M5 synthesis.

Figure 2.4. The experimental yield from the cleavage of the peptide with amino acid sequence PRGRP-NH₂. The same amount of resin and peptide was cleaved separately using identical reaction conditions. The percent yield of the peptide was determined from the weights of lyophilized peptide and resin before cleavage, lyophilized resin after cleavage, and lyophilized peptide after cleavage, and the molecular weights of protected and deprotected peptide.

described in this dissertation.

The crude product of each peptide synthesis in this dissertation was purified using high performance liquid chromatography (HPLC). To form the disulfide bridges, purified peptide was dissolved to a concentration of 2 mM in a 16 mM solution of dithiothreitol (pH < 2.0) to ensure that all disulfides were initially reduced. After four hours, the solution was diluted with Tris buffer (pH=8.0) to a final concentration of 0.1 mM. The low peptide concentration guaranteed that the formation of intermolecular disulfide bridges was insignificant. This mixture was allowed to air oxidize for four days, was then lyophilized to dryness, and was repurified using HPLC.

2.2 High-Performance Liquid Chromatography

The overall separation of two sample components by general chromatography methods is represented by the chromatographic resolution (Introduction to the Theory of Chromatography, 1977) (Figure 2.5). The dependence of resolution upon experimental variables using constant mobile and stationary phases can be most readily understood by considering resolution as a function of relative retentions k_i and the number of theoretical plates of the chromatographic interaction:

$$R = \frac{\sqrt{N}}{4} \left(\frac{k_j}{k_i} - 1 \right) \left(\frac{k_i}{1 + k_i} \right)$$

where $k_i = \frac{t_{r,i} - t_0}{t_0}$ = relative retention of sample component i (Figure 2.5)

k_j = relative retention of sample component j

$N = 16 \left(\frac{t_{r,i}}{t_0} \right)^2$ = number of theoretical plates

The relative retention k_i can also be related to the thermodynamic equilibrium constant established for sample distribution between stationary (S) and mobile (M) phases (Hearn, 1983):

Figure 2.5. A schematic of a HPLC chromatogram, with the experimental parameters discussed in the text:

t_0 = time from injection to first elution of solvent associated with the injection, also known as the "dead time" of the HPLC system

$t_{r,i}$ = time from injection to maximum peak amplitude of solute i

$t_{r,j}$ = time from injection to maximum peak amplitude of solute j

$\Delta t = t_{r,i} - t_{r,j}$

W = peak width at amplitude equal to noise level

$W_{1/2}$ = peak width at half amplitude

W_{lead} = time from start of peak (at amplitude of noise level) to maximum peak amplitude

W_{lag} = time from maximum peak amplitude to end of peak (at amplitude of noise level)

$$k_i = \Phi \frac{[i \cdot S]}{[i][S]} = \Phi \exp\left(\frac{-\Delta G_{\text{assoc } i}}{RT}\right)$$

where $\Phi = \frac{[S]}{[M]}$ = phase ratio of the system.

$$\frac{[i \cdot S]}{[i][S]} = \text{equilibrium constant}$$

In reverse phase HPLC of peptides, the reversible sorption of peptides to alkyl-bonded hydrophobic stationary phases can occur by a physical partitioning and a chemical adsorption mechanism, but both types of equilibria can be included in the same equilibrium constant. However, both mechanisms must be considered when comparing the retentions of two different peptides or the resolution of two different chromatographic experiments.

It is clear that resolution can occur only if k_j is not equal to k_i , or that $\Delta(\Delta G_{\text{assoc}})_{i,j}$ is not equal to zero. It is also clear that resolution increases with increasing k_i , although the improvement in resolution becomes increasingly smaller at large k_i (Figure 2.6). In practice, the solute band width, W , is minimized for handling convenience once the sample is collected:

$$W = \left(\frac{4t_o}{\sqrt{N}}\right)(1 + k_i)$$

Since solute band width is dependent upon k_i , an upper limit of five is empirically chosen for k_i for most HPLC (Guiochon, 1983) (Figure 2.6). In most reverse-phase HPLC, including all chromatography in this dissertation, the ratio of $\frac{k_j}{k_i}$ does not significantly change upon a change in chromatographic conditions, relative to the change in k_i itself.

Thus, a change in resolution is most dependent upon a change in k_i . The resolution is also dependent upon the efficiency of the chromatographic system, represented by the number of theoretical plates. System efficiency is dependent of the multipath effect, molecular diffusion, and resistance to mass transfer, and is described by the Van Deemeter equation:

$$N = \frac{L}{H}$$

Figure 2.6. The dependence of resolution, R (left abscissa), or resolution relative to solute band width, R/W (right abscissa), upon relative retention, k_i . $t_0 = 6$ min., $\frac{k_2}{k_1} = 1.5$, and $N = 50,000$.

where L = column length

$$H = \left[\frac{2}{F} \right] \left[\gamma_M D_M + k_1 \gamma_s D_s \right] + A \left\{ \sqrt{\frac{F d_p^4}{D_M}} + \frac{K F d_p^2}{D_M} \right\}$$

where F = flow rate

γ = tortuosity coefficient

D = diffusion coefficient

A = column packing coefficient

$K = \frac{[i \cdot S]}{[i][S]}$ = equilibrium constant

d_p = particle diameter

However, the efficiency of the HPLC systems and reverse phase columns used in this dissertation is dominated by the resistance of mass transfer, or sorption-desorption kinetics.

$$H \approx \frac{K F d_p^2}{D_M}$$

While several HPLC systems were used in this dissertation, the HPLC purification or analysis of each experiment was conducted using the same HPLC system, reverse-phase column, and flow rate, so that the effect of system efficiency upon resolution was not altered.

A similar expression for resolution can be derived for chromatography using a linear gradient mobile phase (Snyder, 1983):

$$R = \frac{\sqrt{N}}{4} \left(\frac{k_{M,j}}{k_{M,i}} - 1 \right) \left(\frac{\overline{k_{M,i}}}{1 + k_{M,i}} \right)$$

where $\overline{k_{M,i}}$ = median of $k_{M,i}$ during the chromatography.

In these expressions, $k_{M,i}$ is defined as the retention expected from isocratic chromatography of the solute starting at the column inlet at solvent composition M . $\overline{k_{M,i}}$ can be considered to be $k_{M,i}$ when the solute reaches the midpoint of the column length in gradient chromatography. This important difference between isocratic and gradient

chromatography causes resolution to be dependent upon the rate of change of retention, or gradient steepness, instead of retention itself. Figure 2.7 illustrates the relationship between retention and $k_{M,i}$, and shows how $\overline{k_{M,i}}$ depends upon the initial ($M_{t=0}$) and final ($M_{t=final}$) solvent conditions and gradient steepness, G. This relationship can be empirically described:

$$\overline{k_{M,i}} = \frac{(M_{t=0} - M_{t=final})}{115 \cdot t_0 G} \log \left(\frac{k_{M, \text{solvent A}}}{k_{M, \text{solvent B}}} \right)$$

where $M_{t=0}$ = solvent composition at start of gradient

$M_{t=final}$ = solvent composition at end of gradient

G = gradient steepness, change in percent of solvent
composition per minute

This relationship assumes that the solvent strength profile, the relationship between the equilibrium constant K and the solvent composition, is linear for a linear gradient. In practice, the solvent strength profile is not linear for macromolecules, since sorption equilibria depend upon more than one mechanism (physical partitioning and chemical adsorption). Thus, the relationship between relative retention and the gradient is approximate. In this dissertation, solvent A is composed of 0.1% trifluoroacetic acid (Sigma, protein sequencing grade) in water (filtered through Millipore Milli-Q Water Purification System followed by Millipore 0.45 μm HV filter) and solvent B is composed of 0.1% trifluoroacetic acid in acetonitrile (Fisher, HPLC grade). In HPLC of most peptides and proteins using these solvents,

$$\log \left(\frac{k_{M, 0.1\% \text{ TFA in water}}}{k_{M, 0.1\% \text{ TFA in acetonitrile}}} \right) = 3.0$$

Figure 2.7. Examples of solute migration along the column in gradient HPLC. $k_{M,i}$ is displayed at each point along the column, and $\overline{k_{M,i}}$ and k_f are indicated along the right abscissas.

A. $M_{t=0} = 100\%$ solvent A, $M_{t=final} = 100\%$ solvent B, $G = 2\%$ per minute, $t_0 = 6$ minutes.

B. $M_{t=0} = 75\%$ solvent A, $M_{t=final} = 55\%$ solvent A, all other parameters are the same as in 2.7.A. The change in initial and final solvent compositions results in slightly different values of $k_{M,i}$, $\overline{k_{M,i}}$, and k_f in 2.7.A and 2.7.B.

C. $G = 1\%$ per minute, all other parameters are the same as in 2.7.B. The change in gradient steepness results in large differences between $k_{M,i}$ and $\overline{k_{M,i}}$ in figures 2.7.B and 2.7.C, but k_f does not change.

A.

Figure 2.7 (continued).

B.

C.

Sample band width depends upon $\overline{k_{M,1}}$ in gradient HPLC:

$$W = \left(\frac{t_0}{\sqrt{N}} \right) \left[1 + \frac{k_0 k_{M,1}}{2k_0 + k_{M,1}} \right] \left[\frac{\sqrt{k_{M,2} + 2k_{M,1} + \frac{4}{3}}}{k_{M,1} + 2} \right]$$

An upper limit of seven is imposed upon $\overline{k_{M,1}}$ for practical considerations of sample handling (Figure 2.6). Sample band width can be approximated as a function of relative retention using a solvent composition in which the solute is finally eluted:

$$W = \left(\frac{4t_0}{\sqrt{N}} \right) (1 + k_{M, \text{final}})$$

From these functions, it is clear that solute band width and solute resolution are dependent upon different expressions of relative retention (Figure 2.7). Thus, it is possible to increase resolution without a detrimental increase in band width in gradient HPLC. This is not possible in isocratic HPLC, since resolution and band width are both directly related to the same expression of relative retention, k_i . Almost all HPLC separations of peptides and proteins use gradients because of this partial decoupling of resolution and band width.

To attain an acceptable $\overline{k_{M,1}}$ for adequate resolution, the gradient steepness or the initial and final solvent conditions may be changed. Since $t_0=6$ minutes in this dissertation, the change in solvent conditions ($M_{t=0} - M_{t=\text{final}}$) should be no more than sixteen times greater than the gradient steepness, G , to limit $\overline{k_{M,1}}$ to seven. In practice, this ratio of gradient steepness to percent change in solvent composition was not found to be stringent, since the function of R/W vs. k_i contains a broad maximum. The non-linearity of the solvent strength profile also makes this ratio approximate.

The first HPLC chromatographic experiment for each peptide was performed using a gradient with a large change in solvent conditions and a corresponding large gradient steepness. The conditions of this initial chromatographic experiment were kept constant for almost all peptides (Figure 2.8.A). The separation was then optimized by choosing initial and final solvent conditions which bracket solvent composition M of $k_{M,f}$ of the desired

Figure 2.8.A. Initial HPLC chromatogram of Apa-Cro, performed with the following conditions:

$M_{t=0}$ = 0.1% TFA, 6% acetonitrile

$M_{t=final}$ = 0.1% TFA, 42% acetonitrile

G = 0.72% per minute

detector wavelength = 220 nm

detector sensitivity = 0.2 AUFS

injection = 70 μ l of ~5 mM solution of Apa-Cro

flow rate = 12 ml/min

HPLC = IBM LC/9533 Liquid Chromatograph with a Waters Delta Pak preparative column (C18, 300 angstroms, 19 mm x 30 cm column)

Figure 2.8.B. Optimized HPLC chromatogram of Apa-Cro, performed with the same conditions as in 2.8.A, except the following:
 $M_{t=0}$ = 0.1% TFA, 9% acetonitrile
 $M_{t=final}$ = 15% acetonitrile
 G = 0.4% per minute

peptide. Initial solvent conditions were typically 3% less in acetonitrile concentration than M of $k_{M,f}$, and final solvent conditions were typically 3% to 6% greater in acetonitrile concentration than M of $k_{M,f}$. The gradient steepness was then adjusted, as described above, and typically ranged from 0.4% per minute to 0.75% per minute. Further HPLC chromatography was conducted using these optimized conditions (Table 2.1) (Figure 2.8.B).

Table 2.1. Optimized HPLC Conditions

<u>Peptide</u>	<u>Initial %B</u>	<u>Final %B</u>	<u>time (min)</u>
Apa-M5	9	30	35
Apa-Cro	9	12.6	12
Max	24	30	20
Apa-AlaN	18	24	20

When purifying a peptide, the amount of the injected solute was then increased beyond the capacity of the column. This column overloading increased solute band width and reduced the resolution, but the resolution still was adequate for purification in all cases. Increasing the amount of injected solute decreased the number of necessary injections, reducing time and waste generation. The identity of the major peptide peak was verified by electrospray mass spectrometry.

Before and after experimentation with a series of sample injections, the HPLC column was washed with a minimum of three column volumes of 100% acetonitrile (Fisher Scientific), the column was equilibrated with three column volumes of the starting solvent, and a "blank" experiment was performed without injection of a sample and at high detection sensitivity (Runser, 1981). Washing with three column volumes of acetonitrile cleaned the column of all contaminants from peptide and DNA purification. When contaminants from protein purification were encountered, the column was washed with several gradients of 0.1% trifluoroacetic acid (TFA) (photometric grade, Aldrich Chemicals) in water (filtered through 0.45 mm HV filters, Millipore) to 100% acetonitrile, 0.1% TFA water to 100% isopropanol ("Baker Analyzed" photometric grade, JT Baker), and the reverse of these

gradients. The "blank" experiment ensured that the HPLC system and column were adequately cleaned before experimentation with samples was conducted. Between experiments of a series, the column was washed with a minimum of one column volume of acetonitrile, which cleaned the column of all but the most hydrophobic contaminants. Since these most hydrophobic contaminants elute only in 100% acetonitrile, the presence of these contaminants during a series of experiments reduced the column capacity, but did not interfere with purification during the gradient.

Column performance was checked after approximately 100 sample injections by injecting 50 ml of a standard solution [0.108% uracil, 0.16% phenol, 0.16 % methylbenzene, and 0.53% toluene in methanol. All chemicals were spectrophotometric grade, all percents by weight]. The resolution of toluene (longest retention) from uracil (next longest retention), and the relative retention, peak symmetry, and column efficiency (number of theoretical plates) of toluene were measured and compared to the values of these measurements recorded when the column was first used.

$$\text{peak symmetry} = \left(\frac{W_{\text{lead}}}{W_{\text{lag}}} \right) \times 100\%$$

The performance of each column did not change appreciably during experimentation described in this dissertation.

2.3 NMR Spectroscopy

The density operator of a nuclear spin system can be separated into the spin density operator of nuclear spin variables, and the "lattice" of the remaining degrees of freedom (Levitt, 1988). The spin density matrix may be expressed by an orthonormal basis set of product operators (Sorensen, et al., 1983):

$$\sigma_s = 2^{(q-1)} \prod_k^N I_{kv}^a$$

where σ_s = spin density operator

N = total number of $I=1/2$ nuclei in spin system

k = index of nucleus

$v = x, y, \text{ or } z$

q = number of single spin operators in product

$a = 1$ for q nuclei

$a = 0$ for $N-q$ nuclei

The time evolution of the product operators is described by the Liouville-von Neumann equation applied to the spin density operator (Slichter, 1980):

$$d\sigma_S(t)/dt = -i[H_S(t), \sigma_S(t)] - \hat{\Gamma}\{\sigma_S(t) - \sigma_S(t_0)\}$$

where $H_S(t)$ = nuclear spin Hamiltonian

$\hat{\Gamma}\{\sigma\}$ = relaxation superoperator

The nuclear spin Hamiltonian includes interactions with time-dependent radio-frequency (r.f.) fields, the Zeeman interactions with the local static magnetic field, and electron-mediated interactions with other nuclear spin moments.

$$H_S(t) = H_{r.f.}(t) + H_Z + H_J$$

While the nuclear spin Hamiltonian also includes direct dipolar interactions and electric quadrupole interactions, these interactions only contribute to relaxation and hence need not be considered in the formalism of product operators for most NMR experiments.

Instead, dipolar interactions are included in the relaxation superoperator, along with chemical shielding anisotropy of the Zeeman interaction:

$$\hat{\Gamma}\{\sigma(t)\} = (0.5)\sum J(\omega_q)[H^{-q}(t)[H^q(t), \sigma(t)]]$$

where $J(\omega_q)$ = power spectral densities

$H^q(t)$ = operators of dipolar interactions and chemical
shielding anisotropy

where q = irreducible tensor components and different types of fluctuating
interactions.

Both dipolar interactions and chemical shielding anisotropies of Zeeman interactions are

modulated by random molecular processes of the lattice. This relationship assumes that the correlation times of the random processes are much less than t . In the case of isotropic random motion with correlation time τ_c , the spectral densities are most simply described as:

$$J(\omega) = \frac{2\tau_c}{1 + (\omega\tau_c)^2}$$

For a weakly coupled two spin system in high field, dipolar interactions are also simplified:

$$H_D(t) = \frac{\mu_0 \gamma_k \gamma_l \hbar}{8\pi^2 r_{kl}^3(t)} [1 - 3\cos^2\Theta_{kl}(t)] I_{kz} I_{lz}$$

where μ_0 = permeability of free space

γ_k = gyromagnetic ratio of nucleus k

γ_l = gyromagnetic ratio of nucleus l

r_{kl} = distance between nuclei k and l

$\Theta_{kl}(t)$ = angle of internuclear vector between nuclei k and l and the static magnetic field

The Zeeman interaction is modulated through the anisotropy of chemical shielding:

$$H_{CSA}(t) = \sum \gamma_k B_k(t) I_k$$

where $B_k(t)$ = spherical components of fluctuating random fields of the lattice.

All other relaxation mechanisms, including quadrupole interactions, are not included in the relaxation superoperator, since these interactions are negligible in this dissertation. In the absence of cross-correlation between dipolar and chemical shift anisotropy relaxation mechanisms (Boyd, et al., 1991) evaluation of the relaxation superoperator leads to the following rates of longitudinal relaxation (T_1), transverse relaxation (T_2), and the nuclear Overhauser effect (NOE):

$$(1/T_1)_k = d^2 [J(\omega_k - \omega_l) + 3J(\omega_k) + 6J(\omega_k + \omega_l)] + c^2 J(\omega_k)$$

$$(1/T_2)_k = (d^2/2) [4J(0) + J(\omega_k - \omega_l) + 6J(\omega_k) + 3J(\omega_l) + 6J(\omega_k + \omega_l)] \\ + (c^2/6) [3J(\omega_k) + 4J(0)]$$

$$NOE_{kl} = 1 + \{ \gamma_k / \gamma_l \} d^2 [6J(\omega_k + \omega_l) - J(\omega_k - \omega_l)]$$

$$\text{where } d^2 = 0.1 \gamma_k^2 \gamma_l^2 \hbar^2 / (4\pi^2) < 1/r_{kl}^3 >^2$$

$$c^2 = (2/15) \gamma_k^2 H_0^2 (\sigma_{\parallel} + \sigma_{\perp})^2$$

where $\sigma_{\parallel}$ = chemical shift tensor parallel to static magnetic field

$\sigma_{\perp}$ = chemical shift tensor perpendicular to static magnetic field

In Fourier NMR spectroscopy, pulsed radio frequency fields are applied to nuclear spin systems. The time-domain responses of the spin systems are detected and recorded as the spin systems evolve under the influence of Zeeman interactions, scalar coupling interactions, and relaxation processes. Frequency-domain spectra are obtained by Fourier transformation of these responses. These spectra are used to measure chemical shifts, scalar couplings, and rates of relaxation. The application of each of these procedures towards the research of this dissertation is outlined below.

Interactions with strong, nonselective r.f. fields are approximated to be time-independent within a finite time segment by transformation into a coordinate frame rotating at the Larmor frequency:

$$R_{Hr.f.} = -B_1 \sum \gamma_k \{ I_{kx} \cos \phi + I_{ky} \sin \phi \}$$

where B_1 = r.f. field strength

γ_k = gyromagnetic ratio of nucleus k

ϕ = pulse flip angle

Pulse programs were applied using a standard 500-MHz General Electric GN-500 spectrometer or a standard 600-MHz Bruker AMX-600 spectrometer. All pulse experiments are outlined in Appendix 2. For ^{15}N -T₂ and ^{15}N - ^1H NOE experiments using the AMX-600 spectrometer, the power of the ^{15}N pulses was controlled by linking in series the B-SV 10 X-nucleus low power amplifier output of the spectrometer to a 3 dB power attenuator (Telonic model TAB-50C), a 50 watt linear RF power amplifier (ENI model 550L), and the probe X-nucleus transmitter input of the spectrometer probe. The BS-V 10 X-nucleus high power amplifier, the X fast switch unit, and the 3 dB attenuator at the probe transmitter input of the spectrometer were disconnected.

Correlations between coupled proton pairs were identified by performing double

quantum filtered correlation spectroscopy (DQFCOSY) experiments (Muller, et al., 1986) (Figure 2.9). Elimination of zero quantum and odd quanta coherences was accomplished by phase cycling (Bodenhausen, et al., 1984). Artifacts due to incomplete relaxation between scans were reduced by careful selection of dummy scan phases and the order in which the phase cycle was performed (Derome & Williamson, 1990). The resolution of the DQFCOSY spectra was superior to COSY spectra due to suppression of single-quantum diagonal resonances and the absorption-mode character of the diagonal in DQFCOSY spectra (Rance, et al., 1983).

Correlations between uncoupled proton pairs which are coupled to a third "central" proton were obtained by performing relayed correlation spectroscopy (RELAY) experiments (Eich, et al., 1982). A 60 msec mixing time was used to optimize the intensity of relayed magnetization and minimize the intensity of COSY and diagonal resonances in helical residues (Bax & Drobny, 1985). A composite 180° pulse was used to more accurately remove chemical shift effects from the mixing time (Levitt, 1986). These experiments were particularly useful for assigning spin systems in which an α proton was degenerate with another α proton (Wagner, 1983) (Figure 2.9).

Coupled proton networks were mapped using total correlation spectroscopy (TOCSY) experiments (Braunschweiler & Ernst, 1983). The MLEV-17 isotropic mixing sequence was used with the GN-500 spectrometer (Bax & Davis, 1985). The "clean" MLEV-17 mixing sequence was used with the AMX-600 spectrometer, which suppressed cross relaxation caused by non-ideal pulses (Griesinger, 1988). The isotropic mixing time was set to 100 msec to optimize the coherence transfer for helical residues in peptides (Cavanaugh, et al. 1990). Complete assignments of individual spin systems were straightforward (Figure 2.9).

Double quantum coherences can be mapped by performing double quantum correlated spectroscopy (2QCOSY) experiments (Braunschweiler, et al., 1983). A composite 180° pulse was used to more completely remove chemical shift

Figure 2.9. Schematic representation of DQFCOSY, RELAY, and TOCSY spectra of typical alanine, serine residues above the diagonal, and a typical lysine residue below the diagonal. The α proton resonances of the serine and alanine residues are degenerate. DQFCOSY cross peaks are shown as white squares, RELAY cross peaks are shown as black squares, and TOCSY cross peaks are represented by white, black, and gray squares.

effects from the mixing time (Levitt, 1986). The 45° read pulse attenuated responses from passive spins (Mareci & Freeman, 1983). Since responses from directly coupled spins (Type I responses) are observed at the sum of their chemical shifts in the evolution dimension, two uncoupled, overlapped resonances are separated by the chemical shift differences of their coupled spins (Martin & Zektzer, 1988). This experiment was particularly useful for resolving resonances in the crowded aliphatic region of the spectrum of Max (Wagner & Zuiderweg, 1983) (Figure 2.10).

Correlations between directly bonded ^1H and ^{15}N nuclei were identified by heteronuclear multiple quantum correlation (HMQCJR) spectroscopy experiments (Bax, et al., 1983). A 180° refocusing pulse during the evolution period was incorporated into the pulse sequence to suppress signals from protons that are not attached to ^{15}N (Sklenar & Bax, 1987). The phase cycle also served to suppress signals from protons not attached to ^{15}N (Cavanaugh & Keeler, 1988).

Nuclear Overhauser effects between pairs of protons closer than $5 \text{ \AA}$ were measured using two dimensional NOE spectroscopy (NOESY and NOESYJR) experiments (Kumar, et al., 1980). Mixing times between 250 and 350 msec resulted in small, negative NOE cross peaks, indicating that the overall correlation time of the peptides described in this dissertation was greater than $(\omega_0)^{-1}$. Spectra were recorded at temperatures at or below 25°C , so that the overall correlation time of this small peptide remained greater than $(\omega_0)^{-1}$, producing spectra with greater negative NOE intensities than in spectra recorded at higher temperatures. Zero-quantum artifacts were negligible at these mixing times (Macura, et al., 1981) and did not require a suppression scheme.

Degenerate protons can be resolved by performing a NOESY-HMQC spectroscopy experiment (Marion, et al., 1989). This three dimensional experiment combines the two dimensional NOESY and HMQC experiments by superimposing the NOESY detection HMQC evolution periods. NOESY responses in the 2D (ω_1, ω_3) planes are separated in the ω_2 dimension by the frequency of the ^{15}N attached to the proton of the

Figure 2.10. Methyl region of 2QCOSY spectrum of Max. Methyl resonances and resonances correlated with these methyl groups are equidistant from the $F_1 = 2\omega_2$ diagonal.

Figure 2.11. Schematic representation of NOESY-HMQC spectrum, along with the relationship between the complete three dimensional spectrum and the two-dimensional strip plot of the 3D matrix. A schematic representation of the homonuclear ^1H 2D spectrum is shown for comparison. See also Figure 6.6.

ω_3 dimension (Figure 2.11). This experiment was preferred over the HMQC-NOESY experiment (Zuiderweg, et al., 1990; Ikura, et al., 1990) for macromolecules, since NH- α H NOE connectivities are easily detected in the presence of a large water signal (Fesik & Zuiderweg, 1990). The initial t_2 evolution period was optimized to eliminate a linear phase correction in this dimension (Bax, et al., 1991).

^{15}N relaxation rates were measured by ^{15}N - T_1 and ^{15}N - T_2 spectroscopy experiments (Barbato, et al., 1992). Two refocused INEPT-type sequences are employed to transfer magnetization from the directly bound proton to the backbone amide nitrogen and back to the bound proton for detection. The amplitude of the response is attenuated by the T_1 or T_2 CPMG relaxation period between the magnetization transfers. Scrambling pulses and ^1H 180° refocusing pulses during the relaxation period are employed to reduce the cross-correlation between dipolar and chemical shift anisotropy relaxation mechanisms of T_2 relaxation rates (Boyd, et al., 1991). The experiments are independent of the delay time between scans (Sklenar, et al., 1987). The ^{15}N - ^1H NOE was measured using the ^{15}N - ^1H NOE spectroscopy experiment (Barbato, et al., 1992). Spectra were recorded in the presence and absence of ^1H saturation during the relaxation delay. A 20 msec scrambling pulse was applied at the end of the relaxation delay in both experiments to suppress the water signal. While the scrambling pulse led to a truncated NOE effect in the experiment recorded in the absence of ^1H saturation, this effect is negligible when compared to ^{15}N - ^1H cross relaxation times. The relaxation delay was set to 7 seconds to ensure that the ^{15}N magnetization had reached equilibrium between scans.

All 1D spectra were recorded with a sample spinning along the axis of the static magnetic field to average field inhomogeneities perpendicular to the static magnetic field (Chmurny & Hoult, 1990). All 2D spectra were collected with a non-spinning sample to reduce t_1 noise in the evolution dimension (Mehlkopf, et al., 1984). Phase-sensitive two-dimensional (2D) spectra may be recorded using time-proportional-phase-incrementation (TPPI) (Marion & Wüthrich, 1983) or may be obtained from hypercomplex data sets (States,

et al., 1982). All 2D spectra were recorded using TPPI, since TPPI places axial artifacts at the edge of the spectrum {at the Nyquist frequency (ω_N)}, and TPPI places resonances (ω_k) outside the spectral window ($\omega_k > \omega_N \pm \omega_{r.f.}$) at $2\omega_N - \omega_k$, simplifying assignment of these resonances (Keeler & Neuhaus, 1985). Suppression of the water signal was most often achieved by using continuous low-power irradiation at the chemical shift of water during the relaxation delay, and during the NOESY and NOESY-HMQC mixing time. However, signal intensities of peptide protons exchanging with water protons are attenuated by irradiation at the water chemical shift (Hosur, et al., 1983). HMQCJR spectra and additional NOESY spectra were recorded using a 1-1 jump-return detection pulse to suppress the water signal without attenuating intensities of exchangeable peptide protons (Hore, 1983). Sigmoidal baseline curvatures of these 1-1 jump-return spectra were reduced by doubling the sweep width of the acquisition dimension (Stejskal & Schaefer, 1974; B. Wimberly, personal communication). Scrambling pulses were employed in the ^{15}N -T₁, ^{15}N -T₂, and ^{15}N - ^1H NOE pulse sequences to suppress the water signal (Messerle, et al., 1989).

Homonuclear spectra were processed using FTNMR (Hare Research, Inc.) on a Micro Vax II computer or FELIX (Biosym Technologies, Inc.) on a Silicon Graphics Iris UNIX computer. Two-dimensional data sets were apodized with squared sine bell filters, phase shifted by $\pi/3$ to $2\pi/3$, and skewed by 0.5 in both dimensions (Lindon & Ferrige, 1980; Ernst, 1966). DQFCOSY and RELAY data sets were apodized with unshifted filters (Bax & Davis, 1985). All data sets were zero-filled to double the resolution in the evolution dimension (Bartholdi & Ernst, 1973) and to produce symmetrical matrices. Baselines were corrected using a fourth order polynomial or by convolving baseline points with a moving average and applying a straight-line correction to non-baseline intervals. The first data point in t_1 was multiplied by $\left(\frac{1}{2} + \frac{\text{initial } t_1}{t_1 \text{ dwell}}\right)$ to suppress t_1 ridges (Otting, et al., 1986).

Heteronuclear 2D and 3D spectra were similarly processed, with the application of a linear prediction algorithm (Marion, 1989) in the evolution dimension(s) before the matrix was zero-filled. This algorithm served to double the spectral resolution along the evolution

dimension(s) by extrapolating additional points from decaying end of the existing time-domain data.

The static magnetic field was defined to be the product operator z-axis. This reduced the complexity of the Zeeman interaction to scalar terms:

$$H_Z = \sum \Omega_k I_{kz}$$

where $\Omega = \omega_k - \omega_{r.f.}$ = chemical shift

where ω_k = frequency of spin k

$\omega_{r.f.}$ = radio frequency

^1H chemical shifts were referenced to TSP. ^{15}N chemical shifts were indirectly referenced to TSP (Live, et al., 1984). Specific types of protons usually resonate in specific regions of the ^1H NMR spectrum (Wüthrich, 1986), as demonstrated in the one dimensional spectrum of Apa-M5 (Figure 2.12). Disregarding ring-current effects, α protons within a helix have been shown to resonate about 0.4 ppm upfield from their random coil chemical shifts (Szilágyi, & Jardetzky, 1989). Also, α protons near the helix N-terminus have been shown to resonate upfield from helix C-terminal α protons (Williamson, 1990). For amide protons, the C-terminal end resonates upfield. The distributions of helical chemical shifts of peptides studied in this dissertation qualitatively agreed with these analyses (Figure 2.13). α protons resonated an average of 0.27 ± 0.13 ppm upfield from random coil chemical shifts. Disregarding the first few N-terminal residues, the α protons closer to the helix N-terminus resonated upfield from helix C-terminal α protons (relative to random coil chemical shifts), while the opposite trend was observed for amide protons. However, the pattern of chemical shift differences for the first few helix N-terminal residues was significantly different from the pattern for the rest of the peptide. Similar patterns of chemical shift differences were observed in the peptides with an N-terminal disulfide framework and a peptide with a C-terminal disulfide framework, indicating that the patterns are independent of the framework. The patterns are most likely due to the absence of hydrogen bonds for the first few helix N-terminal residues.

Figure 2.12. One dimensional spectrum of Apa-M5. Chemical shift regions are labeled in the spectrum.

Figure 2.13. Chemical shifts of helical residues of peptides studied in this dissertation relative to random coil chemical shifts. Residues were determined to be helical based upon α carbon positions of energy-minimized average coordinates of distance geometry ensembles (a structural interpretation of observed helical NOE patterns), hydrogen bonds from slow amide exchange rates. Helical residues were as follows: Apa-M5, residues 9-17; Apa-Cro, residues 9-18; Max, residues 2-9; Apa-AlaN, residues 9-19. The ordinate reflects the position of the residue relative to the helix N-terminus. Random coil values, obtained from Wüthrich, 1986, were measured from spectra recorded at 20°C and pH=7.0. Chemical shifts of lysine were used for ornithine. Peptide chemical shifts were measured from spectra recorded at slightly different temperatures and pHs (Appendix 3). A 3-point smoothing routine was applied to the data to suppress effects of the disulfide framework.

A. α proton chemical shifts. B. Amide proton chemical shifts.

Degenerate protons were assigned as a pseudoatom (Wüthrich, et al., 1983).

Resonances were assigned to each proton by the main-chain directed method (Englander & Wand, 1987). The patterns of NOEs of helices and reverse turns were particularly useful in the application of this assignment method (Billeter, et al., 1982) (Figure 2.14). The sequential assignment method (Wüthrich, et al., 1982) was then applied to the TOCSY experiment (Davis & Bax, 1985) to resolve ambiguities of the main-chain directed method. Slight differences in chemical shifts of Apa-M5 at 0° and 25° served to resolve overlapped chemical shifts, making possible full chemical shift assignments. The dependence of amide proton chemical shifts upon temperature may be indicative of hydrogen-bonds (Dyson, et al., 1988). However, the temperature dependence of amide proton chemical shifts of the peptide Apa-Ala were not correlated with hydrogen bonds deduced from slow amide proton exchange rates (Figure 2.15). This agrees with similar analysis of peptide Apa-S25 (Storrs, et al., 1992). ¹⁵N resonances were assigned to peptide backbone amide nitrogens based upon couplings to ¹H chemical shifts in the NOESY-HMQC spectrum. All chemical shifts are listed in Appendix 3.

Electron-mediated interactions with other nuclear spins, restricted to weakly coupled spin systems, results in a simplified Hamiltonian describing scalar coupling:

$$H_J = \sum 2\pi J_{kl} I_k I_l$$

where J_{kl} = scalar coupling constant

The value of J_{kl} is dependent upon the dihedral angle between the coupled nuclei k and l (Karplus, 1959). For peptides, the $J_{NH-\alpha H}$ coupling constant is dependent upon the Φ backbone angle (Pardi, et al., 1984). However, this function is not single-valued, and J_{kl} below 7 Hz cannot be interpreted as an unambiguous distance constraint. Scalar coupling constants of amide protons of the peptide Apa-Ala were measured from one dimensional spectra apodized with unshifted, skewed squared sine bell filters (Figure 2.16). Since almost all coupling constants were less than 7 Hz, distance constraints were not derived from coupling constants. Similar results were observed for the other peptides of this dissertation.

Figure 2.14. Distances within helices which give rise to NOE intensities.

$dN\alpha$ = intraresidue distance between amide proton and α proton

$dN\beta$ = intraresidue distance between amide proton and β proton

$d\alpha N$ = distance between α proton and amide proton of next residue

$d\beta N$ = distance between β proton and amide proton of next residue

dNN = distance between amide proton and amide proton of next residue

$d\alpha N_{i+3}$ = distance between α proton of residue i and amide proton of residue $i+3$

$d\alpha\beta_{i+3}$ = distance between α proton of residue i and β proton of residue $i+3$

$d\alpha N_{i+4}$ = distance between α proton of residue i and amide proton of residue $i+4$

Figure 2.15. Amide proton chemical shift temperature coefficients of Apa-Ala residues. Coefficients were determined from spectra collected over the temperature range 10°C - 40°C. Solid bars represent amide protons in hydrogen bonds (deduced from slow amide proton exchange rates), while hatched bars represent amide protons that are not part of hydrogen bonds. Bars are missing for those residues which could not be evaluated due to overlapping resonances. Coefficients below 5 ppb/K (horizontal line) have been considered as evidence that these amide protons are hydrogen bonded.

Figure 2.16. The scalar coupling constant $J_{\text{NH}-\alpha\text{H}}$ of the residues of Apa-Ala. Bars are missing for residues which could not be evaluated due to overlapping resonances.

Analysis of NOESY experiments led to the identification of sequential and nonsequential interresidue NOEs. In the description of the NOE relaxation rate, the spectral densities were approximated by the rigid-body motion at a single correlation time. The influence of internal motions faster than the Larmor frequency introduces an average deviation of 10% in distances derived from the rigid-body approximation (Lane, 1988; Keepers & James, 1984). Internal motions slower than the Larmor frequency may introduce larger distance deviations. Also, the description of the NOE was limited to dipolar interactions between pairs of spins. If additional spins are present, NOE intensities may not be accurately correlated with distance due to spin diffusion involving the additional spins (Noggle & Schirmer, 1971). In particular, the effect of spin diffusion at longer mixing times leads to systematic underestimation of distances longer than 4 Å (Madrid, et al., 1989; Keepers & James, 1984). To avoid inaccuracies in distance measurements from NOE intensities, the intensities were qualitatively categorized as strong, medium, or weak by comparing intensities within the same NOESY spectrum (Havel & Wüthrich, 1984). The strong, medium, and weak NOE intensities were interpreted as short ($1.8 \text{ Å} \leq d \leq 2.5 \text{ Å}$), medium ($1.8 \text{ Å} \leq d \leq 3.5 \text{ Å}$), and long ($2.5 \text{ Å} \leq d \leq 5.0 \text{ Å}$) distance constraints, respectively. The categorization of intensities within the 250 msec NOESY spectrum was identical to the categorization within the 350 msec NOESY spectrum. This indicated that spin diffusion was not problematic for the qualitative categorization of NOE intensities for this peptide of small size and short rotational correlation time. Furthermore, spin diffusion is unlikely to lead to false identification of secondary structure elements, the principle purpose of the NMR spectroscopy of this dissertation (Lane, 1988; Wüthrich, et al., 1984). NOE cross peak intensities involving pseudoatoms were scaled by the number of protons within the pseudoatom (Ying, 1990).

Amide exchange rates were determined by dissolving protonated peptide in 99.9% D₂O, pD=2.8 (uncorrected for the isotope effect), below 25°C. One dimensional spectra were sequentially recorded over 1 hour, and decreases in resonance intensities were determined.

2.4 Distance Geometry

The distance geometry algorithm can be used to translate distance information from NMR spectroscopy data into structural information (Crippen & Havel, 1988). However, the distance constraints from NMR data are not sufficient to fully characterize a peptide structure. An ensemble of peptide structures can be calculated using distance geometry which are consistent with the NMR data. Provided that the ensemble is sufficiently large and properly randomized (Metzler, et al., 1989), the geometric properties displayed in all members of the ensemble must be a necessary consequence of the NMR data (Crippen, 1977). The local geometric properties of the ensemble are generally more accurate than the global geometric properties, since the NMR data consists of short-range distances (Patel, et al., 1987). Thus, the conclusions drawn from distance geometry structures in this dissertation concentrate upon well-defined local secondary structures.

The distance geometry algorithm can also be used to assess whether the NMR data correctly represents a single structure. The distance constraints that are consistently violated in an ensemble of distance geometry structures are identified as being geometrically inconsistent with the remainder of the NMR data. These violated distance constraints can be reevaluated to determine if the constraints are in error, or if the constraints are derived from more than one stable conformation of the peptide (Kim & Prestegard, 1989).

The DG-II program package (Biosym Technologies), based upon the EMBED distance geometry algorithm (Havel, et al., 1983), was used to calculate peptide structures and verify the geometric consistency of the NMR data. The first step of the algorithm is the construction of a bounds matrix, a matrix representing the upper distance bound and lower distance bound between each atom of the peptide. Many upper and lower distance bounds are derived from the local covalent structure of the peptide. Protons which are degenerate in the NMR spectrum are defined as a pseudoatom (Wüthrich, et al., 1983). Prochiral pairs, such as nondegenerate methylene protons and geminal methyl groups, are also

defined as pseudoatoms, and the longer upper bound and shorter lower bound are used in the matrix. Distances between directly bonded atoms and geminal atoms are exact, and the upper and lower bounds are equal. Chiral atoms and distances between vicinal atoms are also used to determine upper and lower bounds. In the distance geometry performed in this dissertation, upper and lower bounds between the β -carbons of the disulfide bridges were also included. These bounds hold the $C\beta-S-S-C\beta$ dihedral angle at a standard conformation of approximately $\pm (90 \pm 10 \text{ degrees})$ (Richardson, 1981). All other lower bounds are set to the sum of the hard sphere radii of each pair of atoms. All other upper bounds are set to 1000 Å, essentially providing no upper constraint.

Upper and lower distance bounds are also derived from the NMR data. Hydrogen bonds were deduced from slowly exchanging amides. The small size of the peptides in this dissertation allows solvent accessibility to most residues, and therefore slow amide exchange is not likely due to burial of the amide proton within the peptide fold (Englander & Kallenbach, 1984). The distance between the amide proton and carbonyl oxygen of each hydrogen bond was set to a lower bound of 1.8 Å and an upper bound of 2.0 Å (Chothia, 1984). The strong, medium, and weak NOE intensities were interpreted as short ($1.8 \text{ Å} \leq d \leq 2.5 \text{ Å}$), medium ($1.8 \text{ Å} \leq d \leq 3.5 \text{ Å}$), and long ($2.5 \text{ Å} \leq d \leq 5.0 \text{ Å}$) distance constraints, respectively.

The disparity between upper bounds and lower bounds is especially large for long distances, since short-range distances are used to construct the bounds matrix. Smoothing of the bounds matrix improves the accuracy of the bounds, especially the bounds of longer distances. Bounds smoothing can also identify geometric inconsistencies within the NMR data. The triangle inequality is exhaustively applied to the bounds matrix using Floyd's shortest-paths algorithm, which reduces computation time to a value proportional to N^3 (where N is equal to the number of atoms) (Dress & Havel, 1988). Tetrangle inequalities are then applied to the bounds (Easthope & Havel, 1988). While the tetrangle inequalities provide only slight improvements to the accuracy of the bounds matrix, these inequalities

are much more effective in determining inconsistent bounds than triangle inequalities. Since an exhaustive application of tetrahedron inequalities requires computation time proportional to N^4 and $16N^2$ bytes of memory, tetrahedron inequalities were limited to quadruples of atoms in sequential residues only.

Distances between each pair of atoms are determined through the procedure of metrization. A pair of upper and lower bounds is arbitrarily chosen, a random distance between these bounds is selected, and this value is substituted for the chosen upper and lower bounds in the matrix. Triangle inequalities are then reevaluated. This procedure is repeated until all distances have been selected. Since triangle inequalities, proportional to N^3 , must be evaluated $N(N-1)/2$ times, a variation of this procedure, termed prospective metrization, was performed in this dissertation. A single atom was arbitrarily chosen, a random value between the upper and lower bounds is selected for each distance between the chosen atom and all other atoms, and these values are substituted for their upper and lower bounds. This procedure changes the values only along the row and column of the bounds matrix for the chosen atom. Triangle inequalities are reevaluated, and metrization is then repeated until all distances have been selected. The prospective version of metrization reduces computation time by a factor proportional to $(N-1)$.

Scalar distances are then converted into four-dimensional tensors by the process of embedding. The bounds matrix is diagonalized and the four largest eigenvalues and corresponding eigenvectors define the four-dimensional space. Atomic coordinates of the structure are generated using these eigenvectors. A fourth dimension was used in the determination of the three-dimensional structures of peptides in this dissertation. The fourth dimension allowed more than one atom to exist at the same three-dimensional coordinates in this early stage of structure calculation. In latter stages of refinement, the dimensionality is reduced to three.

These initial coordinates do not satisfy all of the original constraints. The process of majorization refines these coordinates by minimizing the distances between coordinates

to the randomly chosen distances of the bounds matrix. This minimization was not weighted to improve the speed of the calculation.

The coordinates are further optimized using a simulated annealing protocol and the following error function:

$$F = \sum_{i,j} A^2 + \sum_{i,j} B^2 + \sum_{i=1}^N w_i^2$$

$$A_{ij} = 0, \text{ if } \|p_i - p_j\|^2 > l_{ij}^2$$

$$A_{ij} = l_{ij}^2 - \|p_i - p_j\|^2, \text{ if } \|p_i - p_j\|^2 < l_{ij}^2$$

$$B_{ij} = 0, \text{ if } \|p_i - p_j\|^2 < u_{ij}^2$$

$$B_{ij} = \|p_i - p_j\|^2 - u_{ij}^2, \text{ if } \|p_i - p_j\|^2 > u_{ij}^2$$

where $\|p_i - p_j\|$ = root mean square distance between the three dimensional coordinates of atoms i and j

l_{ij} = lower distance bound

u_{ij} = upper distance bound

w_i = coordinate of fourth dimension

Only distance bounds from the disulfide bridges, hydrogen bonds, and NOE distance constraints are evaluated in this error function. The system is heated to 200K in approximately 10 steps and cooled in 25000 steps, with a step size of 3 psec and atomic masses of 1000 Daltons for all atoms. The use of very heavy masses allows for a very long step size, which causes the annealing to sample larger regions of conformation space and follow the large-scale contours of the error function instead of local fluctuations (Brünger, 1990). The fourth dimension allows atoms to exist at or near the same three-dimensional coordinates during the early steps of annealing. The coordinate of the fourth dimension is eventually reduced towards zero.

To obtain the final structure, the four dimensional coordinates are projected into three dimensions, and the error function is minimized using 250 iterations of conjugate gradients, with early termination only if the RMS gradient norm falls below 0.001. Atomic

distances are compared to the bounds matrix to identify upper or lower bounds which are consistently violated in all structures of the ensemble, or to determine if the maximum violations of the constraints consistently exceed 0.5 Å. The distance constraints of all apamin hybrid sequence peptides were found to be geometrically consistent using these criteria. The upper bounds of five distance constraints of Max were consistently violated in an ensemble of calculated structures. Each of these distance constraints was derived from an NOE cross peak with intensity that was very close to the minimum required for classification as representing medium-range distances. These five NOE cross peaks were reclassified as representing long-range distances, and the corresponding distance constraints were changed. No upper or lower bounds were consistently violated in the ensemble of structures calculated with the new distance constraints.

Many applications of molecular modeling in this dissertation required a single structure which best represented the NMR data. An average structure was calculated from the ensemble of distance geometry structures by calculating the average of the coordinates of each atom in the ensemble. However, this average structure contains non-physical bond lengths and angles for those atoms that were not constrained by the bounds matrix, such as the hydrogens of methyl groups. The average structures were improved by performing energy minimization using a physical forcefield. Energy minimization was also used to improve crude constructions of other molecular models, especially those of biomolecular interactions (Mackay, et al., 1989).

The CVFF89 forcefield and InsightII and Discover program packages (Biosym Technologies) were used for molecular modeling. The physical forcefield includes penalty functions for bond lengths, angle, dihedral angles, and chirality. Cross terms between these penalty functions were not included, since the improvement in accuracy did not justify the 120% increase in computation time (Ermer, 1976). A nonbonded electrostatic term was also not included, since the calculation was performed *in vacuo* (Warwicker, 1986). A nonbonded van der Waals term was included, using a Morse potential, an

approximation of the Lennard-Jones 6-12 potential that is also computationally efficient. Evaluation of the van der Waals term was limited to atom pairs separated by less than 8.5 Å in the initial structure, and the term was linearly damped to zero using a quintic spline function for distances between 6.5 Å and 8.5 Å (Brooks, et al., 1985). The first 100 steps of each minimization used the steepest descents algorithm, while the conjugate gradients algorithm was used until the largest gradient was less than $0.1 \text{ kcal mol}^{-1} \text{ Å}^{-1}$. Minimization was performed with all non-hydrogens fixed, followed by minimization with all non-hydrogen backbone atoms fixed, and concluded with minimization with no atoms fixed (Roberts, et al., 1986).

Chapter 3: Apa-M5

3.1 Introduction

Domains from a variety of proteins have been discovered which retain the ability to interact with other proteins, even when separated from the remainder of the parent molecule (Cox, et al., 1985). In some cases, in spite of the fact that they retain the ability to make sequence specific contacts, it appears that these domains are unfolded when present alone in solution. This makes it difficult to establish the conformation of the "active", bound form of the domain. One such domain is the M5 peptide from skeletal muscle light chain myosin kinase, which is known to bind tightly to calmodulin in a Ca^{+2} dependent way (Kennelly, et al., 1987). Previous NMR and CD studies showed that the M13 peptide, an analog of M5 containing an additional nine C-terminal residues, behaves as a random coil alone in solution, but a large increase in helical CD upon complex formation suggested that it forms an a helix when bound to calmodulin (Klevit, et al., 1985). Furthermore, the solution structure of the complex formed by calmodulin and M13 has been determined recently using 3D NMR spectroscopy (Ikura, et al., 1992). The results of this chapter establish that the M5 peptide also binds tightly to calmodulin as a helix in a Ca^{+2} dependent fashion.

It has been shown that a disulfide framework, taken from the bee venom peptide apamin, can be used to make hybrid sequence peptides with defined conformations and high stability (Pease, et al., 1990). This chapter describes the application of this approach to make a hybrid sequence peptide consisting of the apamin and M5 peptide amino acid sequences. The structure of residues 1 to 15 of this hybrid sequence peptide are shown to be very similar to other hybrid sequence peptides of similar design. The remaining residues are shown to be in an equilibrium between a helix and a random coil. The results of this chapter show that the hybrid sequence peptide approach can be used to define the

bound conformation of biologically active peptides.

3.2 Peptide Design

The amino acid sequences of apamin, M13, M5, and the hybrid Apa-M5 are shown in Figure 3.1. The hybrid peptide numbering scheme will be used to refer to residues in the remainder of this chapter. The hybrid peptide was made by fusing the eight amino terminal residues of the bee venom peptide apamin with the sequence of M5, modified to insert two cysteines required to make stabilizing disulfide bridges. The sequence fusion chosen for the Apa-M5 hybrid was based upon the known requirement of W13 and F17 residues in the peptide, which has been previously established by the sequence homology of amphiphilic peptides which bind calmodulin (Cox, et al., 1985). The postulated formation of a helix by M5 upon binding, and the requirement for hydrophobic amino acids W13 and F17, led to the hypothesis that M5 forms an amphipathic helix in the bound state. For this reason, the cysteine residues required for disulfide bridge formation, were placed into the peptide sequence at positions 11 and 15. In other hybrid sequence peptides, these disulfide bridges and additional residues from the apamin framework are proximal to residues 11, 12, and 15 of the helix and do not contact residues 13 and 17. Drawn as an Edmundson helical wheel projection, this then puts W13 and F17 fully exposed on one face of the helix, while the disulfide bridges and apamin framework are positioned along the face of the helix that was previously hydrophilic (Figure 3.2).

3.3 Peptide Structure

The CD spectrum of the oxidized hybrid peptide is very similar to that of apamin, with a minimum at 208 nm and a shoulder at 222 nm (Figure 3.3) (Hider & Ragnarrson, 1981). This spectrum has the general character of that from an α -helix, but must have

Figure 3.1. Amino acid sequences of Apa-M5, apamin, the M5 peptide, and the M13 peptide, along with the corresponding numbering scheme of Apa-M5. Disulfide bridges are indicated above the sequences of Apa-M5 and apamin.

Figure 3.2. An Edmundson helical wheel projection of Apa-M5. This crude representation views the helix along the helix axis from the N-terminal end. The helical wheel depicts the radial distribution of side chains along the cylindrical helix surface. W13 and F17, which have been shown to be necessary for binding, are shown in bold. C11 and C15, which form disulfide bridges with the apamin framework, are shown in outline.

Figure 3.3. The circular dichroism spectrum of Apa-M5 shown as a solid line. CD measurements were carried out on a Jasco J500 spectropolarimeter using a 0.2 cm pathlength cell. Samples were approximately 40.8 mM, in 1.0 mM MOPS buffer and 100 mM NaCl at pH=7.0 and 25°C. Scans were taken from 250 to 205 nm in 0.2 nm steps with a 1.0 nm bandwidth. Sixteen scans were averaged for each spectrum, the spectrum was further processed with a three-point smoothing routine, and an identically-processed spectrum of the solvent was subtracted. The first derivative of the spectrum is shown as a dotted line.

contributions from the β -turn and loop in apamin as well, since the 208 nm / 220 nm ratio and the shape of the CD spectrum are significantly different from purely helical peptides (Johnson, 1990). However, the spectrum shape is very reproducible among the hybrid sequence peptides, and seems to be indicative of the folded core structure (Pease, et al., 1990). The Θ_{222} observed corresponds to approximately 42% of the peptide being helical at 25°C (Johnson & Tinoco, 1972). In light of the NMR data presented below, it appears more likely that there are some residues which are fully helical, but others which are in an equilibrium between helix and coil. The CD spectrum shows a slight linear decrease in intensity upon raising the temperature to 75°C (Figure 3.4). The similarity between the CD spectra at 25°C and 75°C and the display of no cooperative, global unfolding between 25°C and 75°C indicates that the Apa-M5 structure is stable towards thermal denaturation. The CD spectrum shows a non-linear increase in intensity upon lowering the temperature to 0°C, indicating that a different stable conformation exists at low temperature (Figure 3.4). Since the CD spectra of similar hybrid sequence peptides show a linear dependence upon temperature between 0°C and 25°C, the low-temperature conformation Apa-M5 must involve the side chain of at least one residue unique to Apa-M5 among hybrid sequence peptides. Stabilizing interactions involving side chains almost always occur between $i, i+4$ residues, as discussed in chapter 1.2 and evidenced by other results (Ghardiri & Choi, 1990). Based upon NMR spectroscopy (discussed below) the conformations at 0°C and 25°C are identical for the first 17 residues; therefore, the stabilizing interaction occurs beyond residue 17. While the stabilizing interaction in Apa-M5 was not characterized, candidates for an $(i, i+4)$ side chain interaction beyond residue 17 are limited to (S21, R25) and (N24, C-terminus"28").

The NMR assignments for the peptide were straightforward to obtain, taking advantage of the temperature dependent chemical shifts (Appendix 3). The tracing of the assignments of the amide region is shown in Figure 3.5. A summary of the sequential connectivities observed, and longer range contacts in the helix are shown in Figure 3.6.

Figure 3.4. The temperature dependence of the Apa-M5 circular dichroism intensity. CD measurements were carried out as described in Figure 3.3. Error bars represent the standard deviation of five measurements at each temperature. The temperature was varied between each measurement, and the same sample was used for all measurements. The linear least squares fit of the data from 25 $^{\circ}\text{C}$ to 75 $^{\circ}\text{C}$ is also shown.

3.5. NOESY spectrum of Apa-M5. The sample was 10.8 mM in 90% H₂O/10% D₂O, at 0°C. Data were collected and processed as described in chapter 2.3. Only the downfield region of the symmetrical 5208 Hz spectrum is displayed. Sequential NOEs within the helix and reverse turn are labeled.

Figure 3.6. A summary of the sequential and long-range NOE connectivities observed in the NOESY spectrum of Figure 3.5. The sizes of the bars reflect the relative NOE intensities, and question marks reflect indeterminable NOE intensities due to resonance degeneracies.

The types of connectivities observed are exactly the same as those in both apamin (Pease & Wemmer, 1988) and the Apa-S hybrid (Pease, et al., 1990) for the first 16 residues at 0°C and 25°C. These include the α H-NH-NH-NH connectivity characteristic of the type I β -turn, and $i,i+2$ connectivity at the beginning of the helix, and the continued NH-NH connectivities in the α -helix from residue 9 through residue 16. The presence of the helix is verified by a number of $i,i+3$, α H-NH, and α H- β H connectivities, which are normally seen in this secondary structure. Additional tertiary NOEs are observed which indicate that the fold of the first 16 residues is very similar to apamin (Pease & Wemmer, 1988) and the Apa-S hybrid (Pease, et al., 1990). These include contacts between residues C3 & P6, P6 & R12, and C3 & C15.

Partial NOE connectivities in the NMR spectra of the Apa-M5 hybrid indicate that the helix continues towards the end of the peptide. Due to overlap in chemical shifts, there are some residues for which the helical NOE connectivities cannot be observed. Given the number of interspersed helical NOE connectivities, it seems most likely that the helix propagates in a cooperative fashion from the section nucleated by the formation of disulfide bonds (residues 9-15). However, beginning at residue 16 immediately after the last cystine residue, there is an increase in the intensity of the NH- α H NOEs. This indicates that there is an equilibrium between helical and random coil states from residue 16 through 27.

Although the chemical shift overlap makes it difficult to assess the helical state of several residues near the end of the sequence, the presence of the $i,i+3$ connectivity between residues 20 and 23 strongly supports the idea that there is a continuous helix at least through residue 23. If the helix propagates to the end of the peptide, residues 24-27 would form the last helical turn. While the backbone amides of these residues would be involved in the helix hydrogen bond network, the backbone carbonyls of these residues would not be involved in backbone or side chain hydrogen bonds, and thus the stability of this last turn of helix is not expected to be as great as the interior turns of helix. The lack of structural stability of the last turn has been observed in a variety of other short helical

peptides (Finkelstein, et al., 1991). Thus, it appears likely that the 42% helical content at 25°C corresponds in reality to residues 9-15 which are fully helical and residues 16-27 which are helical in 40% of the molecules.

This behavior is somewhat different from what was seen in the Apa-S hybrid (Pease, et al., 1990). In that case there was a high helical content near the disulfide stabilized section, but then a clear loss of helix well before the C-terminus of the peptide, with a dynamic disordering leading to a complete loss of observable NOEs for the C-terminal residues. This behavior correlates well with the helix stop signal characterized in this segment of the molecule (Rico, et al., 1983; Kim & Baldwin, 1984). It has recently been shown that the helix can be pushed through this stop signal by the presence of high TFE concentrations in the solvent (Storrs, et al., 1992). In the Apa-M5 hybrid, the helix seems to propagate further into the sequence, perhaps not terminating until the last residue. However, the fraction of molecules with a full helical length is approximately 40%.

Twenty distance geometry structures were calculated using distance constraints for the first 17 residues. NOEs from residues 18 to 27 were not included since the equilibrium between helical and random coil states from residue 18 to 27 indicated that the NOEs were not derived from a single structure. The 20 structures display a well-defined backbone conformation for the first 17 residues (Figure 3.7). The backbone RMS difference between residues 1 to 15 of Apa-M5 and apamin, 1.35 Å, is greater than the backbone RMSDs of the first 15 residues of the Apa-M5 ensemble, $0.56 \text{ Å} \pm 0.25 \text{ Å}$. This difference between the Apa-M5 and apamin structures is primarily due to the displacement of residues 1, 2, 5, and 6 of Apa-M5 relative to apamin (Figure 3.8). However, the backbones of the helices, from residues 9 to 15, of Apa-M5 and apamin are indistinguishable within the limits of resolution of the NMR data.

3.4 Peptide Function

CD studies were conducted to examine the bound conformation of Apa-M5

Figure 3.7. Stereoview of the 20 distance geometry structures of the first 18 residues of Apa-M5 showing the backbone and disulfide bridges. All structures were superimposed by a separate least-squares fit of all backbone atoms of each of 19 structures to one structure.

A.

Apa-M5

apamin

B.

Apa-M5

apamin

Figure 3.8. Comparison of energy minimized average coordinates of the Apa-M5 distance geometry ensemble and the structure of apamin. A. The backbones of residues 1-17 were superimposed, followed by a horizontal translation of apamin, so that both peptides have the same orientation. B. A second view of the structures, rotated 90° relative to the view of A. The N-termini of the helices are closest to the viewer.

Figure 3.9. The CD spectrum of the (Apa-M5)-calmodulin complex, and the sum of separate spectra of Apa-M5 and calmodulin. CD measurements were carried out using the same conditions as used in Figure 3.3 with the addition of 1.0 mM CaCl_2 . The concentration of *Drosophila* calmodulin (kindly provided by R. Klevit) ranged from 0 to 200 μM .

(Figure 3.9). Upon addition of the peptide to calmodulin in the presence of Ca^{+2} , the ellipticity of the 222 nm helical band increased beyond the sum of the ellipticities of Apa-M5 and calmodulin. It is not possible to assign the increase in CD ellipticity of the (Apa-M5)-calmodulin complex to either peptide or calmodulin alone. Calmodulin has a highly folded structure with high helix content, even in the absence of Ca^{+2} (Ikura, et al., 1992). If the increase in the CD signal is interpreted as coming from the peptide, then this corresponds to an increase in peptide helicity from 30% to 60% helicity corresponds to the helix extending from residue 9 to 23 and an equal distribution of helical and random coil conformations for residues 24 to 27 (the last turn of helix) in each molecule. Thus Apa-M5 is essentially a fully helical peptide when bound to calmodulin. Similar analysis of the CD spectrum of the complex formed by calmodulin and M13 indicated that M13 is essentially a fully helical peptide when bound (Kleitman, et al., 1985). This indicates that the M13 peptide and Apa-M5 bind to calmodulin in an identical way.

The fluorescence signal of Apa-M5 was measured with incremented concentrations of *Drosophila* calmodulin (Figure 3.10). Under the conditions used in these experiments, the only fluorophor contributing to signal is the single tryptophan of the peptide (Blumenthal, et al., 1985). As expected, there is enhancement of fluorescence, with a clear stoichiometry of Ca^{+2} -dependent binding of one Apa-M5 per calmodulin, paralleling the behavior of the M13 peptide. To measure the binding affinity of Apa-M5 for calmodulin, concentrations would be required of Apa-M5 that are near the detection limit for the relatively weakly fluorescing tryptophan. Thus the binding affinity was measured by titrating a competitor for the calmodulin binding site into the (Apa-M5)-calmodulin solution. The Calmodulin Binding Domain (Calbiochem) peptide competitor contains no fluorophor, has been determined to be a potent calmodulin antagonist with an IC_{50} of 52 nM for inhibition of Ca^{+2} /calmodulin-dependent protein kinase II (Payne, et al., 1988). This measure of calmodulin inhibition was assumed to be the binding constant of Calmodulin Binding Peptide to calmodulin. The fluorescence signal of Apa-M5 was

Figure 3.10. The fluorescence intensity of Apa-M5 with calmodulin. Fluorescence measurements were done on a Perkin Elmer LS-5B fluorimeter using excitation at 295 nm and detection at 355 nm at 25°C. Apa-M5 was at a concentration of 1 μ M in a 1 mM CaCl_2 solution, with calmodulin concentrations ranging from 0 to 5 μ M.

Figure 3.11. The fluorescence intensity of Apa-M5 with calmodulin and Calmodulin Binding Domain. Fluorescence measurements were done using the same conditions listed in figure 3.8. The concentrations of Apa-M5 and calmodulin (Calbiochem) were 1 mM in a 1 mM CaCl_2 solution, and the concentration of Calmodulin Binding Domain (a 20 residue peptide containing no fluorophore, purchased from Calbiochem) ranged from 0 to 5 mM. The dissociation constants are shown as follows: $K_{Apa-M5} = 102$ nM, solid line; $K_{Apa-M5} = 103$ nM, dotted line; $K_{Apa-M5} = 116$ nM, dashed line.

measured with incremented concentrations of inhibitor (Figure 3.11). Three values for the (Apa-M5)-calmodulin dissociation constant are obtained from the second-order polynomial derived from equilibrium theory:

$$[A \cdot C][CBD]_0 = \frac{K_{CBD}}{K_A} + [A \cdot C] \left(1 - \frac{2K_{CBD}}{K_A} \right) + [A \cdot C]^2 \left(\frac{K_{CBD}}{K_A} - 1 \right)$$

where $[A \cdot C]$ = concentration of (Apa-M5)-calmodulin complex

$[CBD]_0$ = total concentration of Calmodulin Binding Peptide

K_{CBD} = binding constant of Calmodulin Binding Peptide for calmodulin

= 52 nM

K_A = binding constant of Apa-M5 for calmodulin

= 102 nM, from $\frac{K_{CBD}}{K_A}$

= 103 nM, from $[A \cdot C] \left(1 - \frac{2K_{CBD}}{K_A} \right)$

= 116 nM, from $[A \cdot C]^2 \left(\frac{K_{CBD}}{K_A} - 1 \right)$

$\overline{K_A} = 107 \text{ nM}$

Each of the dissociation constants fits the data equally well, within experimental error.

Thus, the dissociation constant of (Apa-M5)-calmodulin is approximately 107 nM, which is 100-fold greater than the dissociation constant of the complex formed by calmodulin and M5, which is approximately 1 nM (Kennelly, et al., 1987).

3.5 Molecular Modeling

The 100-fold decrease in the Apa-M5 affinity for calmodulin, relative to the affinity of M5, suggests that the addition of the disulfide framework may interfere with binding. A molecular model of Apa-M5 was constructed using the energy minimized average structure of the distance geometry ensemble for residues 1 to 16, and setting residues 17 to 27 to ideal helical dihedral angles. By superimposing part of the backbone of the Apa-M5 helix

(residues 13 to 27) upon the corresponding portion of the M13 helix backbone (residues 4-18) in the M13-calmodulin complex, it is evident that the portion of the M13 helix that is shielded by the disulfide framework in Apa-M5 does not contact calmodulin. (Figure 3.12) Thus, the framework of Apa-M5 does not shield a portion of the surface of the helix which makes binding contacts to calmodulin. In particular, residues R11 and K15 in M5, which are replaced by cysteines in Apa-M5, do not contact calmodulin and appear to be exposed to solvent. However, the first helix of calmodulin, which is not in contact with the M13 peptide, interferes sterically with the framework of Apa-M5. While the interaction of Apa-M5 and calmodulin is unknown, this model would suggest that the structure of the (Apa-M5)-calmodulin complex must differ from the structure of the M13-calmodulin complex to alleviate this steric overlap. It is possible that the orientation of calmodulin and Apa-M5 may change by approximately 40° about the Apa-M5 helix axis, changing the orientation of the binding surfaces. It is also possible that the axis of the first helix of calmodulin may adjust about 20° away from the Apa-M5 disulfide framework to accommodate Apa-M5, without necessitating a change in the relative orientation of Apa-M5 and calmodulin. This adjustment of the first helix may cause further changes in the conformation of the calmodulin binding surface, or may sterically interfere with the translation of Apa-M5 to the calmodulin binding surface. Adjustments in the structure of Apa-M5 are less likely, since the disulfide bridges hold Apa-M5 in a constrained, rigid conformation.

An additional hybrid sequence peptide was designed which places the disulfide framework and the hydrophobic residues W13 and F17 on opposite sides of a stable amphiphilic helix (Figure 3.13). Apa-M60 shifts the register of the M5 peptide relative to the disulfide framework by 3 residues. This results in a clockwise 300° rotation, or a counterclockwise 60° rotation, of the disulfide framework around the cylindrical surface of the helix, as seen in an Edmundson helical wheel projection (Figure 3.14). Apa-M60 does not contain the first two residues of the M5 peptide. These residues appear to be exposed to solvent and are not part of the M13 helix in the calmodulin-M13 complex. The

Figure 3.12. Structures of Apa-M5 and the M13-calmodulin complex (coordinates of the M13-calmodulin complex provided by A. Bax). The backbone atoms of residues 13 to 20 of Apa-M5 were superimposed by a least-squares fit upon the backbone atoms of residues 5 to 12 of M13. Only α carbon traces are shown, and Apa-M5, M13, and the first helix of calmodulin are emphasized by ribbons. The disulfide framework of Apa-M5 overlaps with the first helix of calmodulin.

A. The view is along the average Apa-M5 and M13 helix axes from the N-terminal end.
 B. The view is along the helix axis of the first helix of calmodulin from the N-terminal end. This view differs from the view in Figure 3.15.A by a 72° rotation about the third axis of the Cartesian coordinate frame specified by the axes of the Apa-M5 helix and the first helix of calmodulin.

contribution of these residues towards the formation of the calmodulin-M5 complex are unknown.

A model of Apa-M60 was constructed using the backbone coordinates of the Apa-M5 model. By superimposing part of the backbone of the Apa-M60 helix (residues 10 to 24) upon the corresponding portion of the M13 helix backbone (residues 4-18) in the M13-calmodulin complex, it is evident that the portion of the M13 helix that is shielded by the disulfide framework in Apa-M60 does not contact calmodulin. (Figure 3.15) Most importantly, no steric contacts exist between Apa-M60 disulfide framework and calmodulin. This suggests that Apa-M60 may bind calmodulin without steric interference from the disulfide framework, which may decrease the dissociation constant of the (Apa-M60)-calmodulin complex relative to the (Apa-M5)-calmodulin complex.

Figure 3.14. An Edmundson helical wheel projection of Apa-M60. W10 and F14, which have been shown to be necessary for binding, are shown in bold. C11 and C15, which form the disulfide bridges with the apamin framework, are shown in outline.

A.

B.

Figure 3.15. Structures of Apa-M60 and the M13-calmodulin complex. The backbone atoms of residues 10-17 of Apa-M5 were superimposed by a least-squares fit upon the backbone atoms of residues 5-12 of M13. Only α carbon traces are shown, and Apa-M60, M13, and the first helix of calmodulin are emphasized by a ribbon. A. The view is along the average Apa-M60 and M13 helix axes from the N-terminal end. B. The view is along the helix axis of the first helix of calmodulin from the N-terminal end. This view differs from the view in Figure 3.15.A by a 72° rotation about the third axis of the Cartesian coordinate frame specified by the axes of the Apa-M60 helix and the first helix of calmodulin.

Chapter 4. Apa-Cro

4.1 Introduction

The functions of many proteins depend upon small, discrete motifs within larger structures. In some cases, a motif can be inserted into another protein, providing the other protein with the function of the motif (Marks, et al., 1986; Wharton, et al., 1984; Neuberger, et al., 1984). A variety of evidence suggests that similar shufflings of motifs have been conducted during protein evolution (De Sanctis, et al., 1986). One such motif is the recognition helix of helix-turn-helix (HTH) DNA-binding proteins, which is the primary determinant of the sequence specificity of the HTH protein for a DNA operator site (Jordan & Pabo, 1988; Anderson, et al., 1981; Brennan, et al., 1986; Mondragón, et al., 1989). Residues in a HTH recognition helix which contact DNA may replace the corresponding residues in a second HTH protein. The sequence specificity of the second HTH protein with a "swapped helix" is switched to the specificity of the first HTH protein (Wharton, et al., 1984; Wharton & Ptashne, 1985). However, the swap of the helix alone appears to work only in a few specific cases, indicating that sequence specificity is not limited to the recognition helix.

This chapter describes the characterization of a hybrid sequence peptide in which the sequence of the prokaryotic λ Cro recognition helix is "swapped" into the helix of the bee venom peptide apamin. The disulfide framework of the hybrid sequence peptide is required to stabilize the conformation of the recognition helix, since the conformation of this motif is not a stable helix without the surrounding HTH protein matrix (Mayer, et al., 1983). The helix and C-terminal tail of the hybrid sequence peptide accommodates the DNA sequence-specific contacts contained within the recognition helix and C-terminal arm of λ Cro (Brennan, et al., 1986) (Figure 4.1). Of the HTH proteins, the λ Cro sequence was chosen for the sequence hybridization since the hybrid sequence peptide cannot

Figure 4.1. Stereoview of a model of λ Cro and the 5' half site of the OR3 operator, derived from the α carbon and phosphate coordinates of the x-ray crystal structure of the (λ Cro)-OR3 complex (Entry 4CRO, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987)). The resolution of the structure (3.9 Å along two axes, 5.1 Å along the third axis) precludes the identification of sequence-specific contacts. Therefore, this model only contains the α carbons of λ Cro. The recognition helix and C-terminal arm of λ Cro are shown in bold lines.

accommodate sequence-specific contacts contained within the N-terminal arm or loops of other HTH proteins. Almost all HTH proteins have significant dimer contacts when bound to DNA (Anderson, et al., 1987; Wolberger, et al., 1988), and independent binding of a wild-type HTH monomer to a DNA half-site has not been observed (Mondragón, et al., 1989; Aggarwal, et al., 1988). The hybrid sequence peptide does not include the residues of λ Cro involved in dimerization, and only monomeric binding of the hybrid peptide to DNA is investigated in this chapter.

Sequence specificity of HTH protein recognition helices is often investigated by determining the effects of amino acid substitutions upon DNA binding affinity or protection of DNA from chemical modification (Nelson, et al., 1983). However, such a determination must also consider the interactions of the new residues with the flanking amino acid sequence, especially those interactions which affect folding and stability of the recognition helix. Hybrid sequence peptides have been shown to fold into a predictable structure to form a helix with high stability. Residues in the recognition helix of the hybrid sequence peptide may be substituted to determine the effects upon sequence specificity without an effect upon the conformation or stability of the helix. Furthermore, the disulfide framework of hybrid sequence peptides replaces the flanking amino acid sequences of HTH proteins, allowing for the study of an isolated recognition helix. This provides for a more direct comparison of amino acid substitutions upon sequence specificity.

4.2 Peptide and DNA Design

The amino acid sequences of apamin, the hybrid Apa-Cro, and the recognition helix of λ Cro are shown in Figure 4.2. The hybrid peptide numbering scheme will be used to refer to residues in the remainder of this chapter. The hybrid peptide was made by fusing the eight amino terminal residues of apamin with the sequence of the λ Cro recognition

Figure 4.2. Amino acid sequences of apamin, the Apa-Cro hybrid peptide, and the recognition helices of λ Cro, λ Repressor, 434 Cro, and 434 Repressor, along with the corresponding numbering scheme of Apa-Cro. Disulfide bridges are indicated above the sequences of apamin and Apa-Cro.

helix, modified to insert two cysteines required to make stabilizing disulfide bridges. The sequence fusion chosen for the Apa-Cro hybrid was based upon the known sequence specificity of residues 9, 10, 13, 14, and 17 in λ Cro (Ohlendorf, et al., 1982) (Figure 4.3). The cysteine residues required for disulfide bridge formation were placed into the peptide sequence at positions 11 and 15. In other hybrid sequence peptides, these disulfide bridges and additional residues from the apamin framework are proximal to residues 11, 12, and 15 of the helix. Drawn as an Edmundson helical wheel projection, this puts residues required for sequence specificity fully exposed on one face of the helix, while the disulfide bridges and apamin framework are positioned along the opposite face of the helix (Figure 4.4).

In HTH proteins, the first helix is packed along residues 12, 15, and 16 of the recognition helix to form part of an enclosed hydrophobic core (Jordan & Pabo, 1988; Anderson, et al., 1981; Brennan, et al., 1986; Mondragón, et al., 1989). A similar hydrophobic core is not formed in Apa-Cro, and residues 12 and 16 were replaced to decrease the helical hydrophobic moment in Apa-Cro (Eisenberg, et al., 1982). Positively charged residues were substituted for residues 12 and 16 to provide possible electrostatic interactions with the negatively charged phosphate DNA backbone. Ornithine was chosen for this substitution, since the shorter length of the ornithine side chain relative to arginine and lysine facilitates NMR assignments.

A five residue C-terminal "arm" of λ Cro wraps around the double helix to interact with the minor groove of base pairs G10 and C11 (Ohlendorf, et al., 1982). This arm is disordered in the free protein (Anderson, et al., 1981) and is not constrained by tertiary contacts in λ Cro bound to DNA. The design of Apa-Cro includes a similar arm consisting of the three sequential positively-charged residues of the λ Cro arm. To allow the Apa-Cro arm to extend to the minor groove of base pair 11, one additional residue must be inserted between the recognition helix and the arm. This residue was chosen to be a glycine to provide the greatest amount of conformational flexibility for the Apa-Cro arm.

Figure 4.3. Schematic illustration of the presumed interactions between λ Cro and the 5' half site of OR3 (Ohlendorf, et al., 1982). The numbering scheme refers to residues of Apa-Cro.

Figure 4.4. An Edmundson helical wheel projection of Apa-Cro. This crude representation views the helix along the helix axis from the N-terminal end. The helical wheel depicts the radial distribution of side chains along the cylindrical helix surface. Q9, S10, N13, K14, and H17, which have been shown to be necessary for binding, are shown in bold. C11 and C15, which form disulfide bridges with the apamin framework, are shown in outline.

The nucleic acid sequences of OR3/2 and OR3 are shown in Figure 4.5. The numbering scheme of OR3/2 will be used to refer to nucleic acids in the remainder of this chapter. The OR3 sequence was selected to investigate DNA-binding activity of λ Cro, since the binding affinity of λ Cro is strongest for OR3 relative to the other five known operator sites (Johnson, et al., 1979; Johnson, et al., 1981). The 5' half-site of the OR3 operator was selected since this sequence fully represented the "consensus" half-site (Benson, et al., 1989; Benson & Youderian, 1989). Two GC base pairs were added to each end of the OR3/2 sequence to stabilize the operator half-site (Saenger, 1984). The order of the G and C end base pairs was different for the ends of each DNA strand to facilitate NMR assignments. The sequence of the 5' end was chosen to be CG to match residue C11 of OR3. It has been shown that the binding of λ Repressor to a 5' half-site requires this base pair at the dyad axis (Beamer & Pabo, 1992), although the sequence-specificity of λ Cro for C11 is unknown. Also, the duplex of OR3/2 is expected to be 2.60 kcal/mol more stable than the duplex of CGTATCACCGGC due to base stacking interactions (Saenger, 1984).

4.3 Peptide Structure

The NMR assignments for the peptide were straightforward to obtain (Appendix 3). The tracing of the assignments of the amide region is shown in Figure 4.6. A summary of the sequential connectivities observed, and longer range contacts in the helix are shown in Figure 4.7. The types of connectivities observed are exactly the same as those in both apamin (Pease & Wemmer, 1988), the Apa-S hybrid (Pease, et al., 1990), and the Apa-M5 hybrid (chapter 3.3) for the first 16 residues at 15°C. These include the α H-NH-NH-NH connectivity characteristic of the type I β -turn, $i, i+2$ connectivity at the beginning of the helix, and the continued NH-NH connectivities in the α -helix from residue 9 through residue 16. The presence of the helix is verified by a number of $i, i+3$, α H-NH, and α H-

OR3/2: **GCTATCACCGCG**
CGATAGTGGCGC

OR3: **TATCACCGCAAGGGATA**
ATAGTGGCGTTCCCTAT

Figure 4.5. The nucleic acid sequences of OR3/2 and the OR3 operator binding site of λ Cro. The putative Apa-Cro binding region of OR3/2 and the 5' half site of OR3 are shown in bold.

Figure 4.6. NOESY spectrum of Apa-Cro. The sample was 12 mM in 90% H₂O/10% D₂O, at 15°C. Data were collected and processed as described in chapter 2.3. Only the downfield region of the symmetrical 6250 Hz spectrum is displayed. Sequential NOEs within the helix and reverse turn are labeled.

β H connectivities, which are normally seen in this secondary structure. Additional tertiary NOEs are observed which indicate that the fold of the first 16 residues is very similar to apamin (Pease & Wemmer, 1988) the Apa-S hybrid (Pease, et al., 1990), and the Apa-M5 hybrid (chapter 3.3). These include contacts between residues C3 & P6, P6 & R12, and C3 & C15.

NOE connectivities in the NMR spectra and amide proton exchange rates of the Apa-Cro hybrid strongly indicate that the helix propagates in a cooperative fashion from the section nucleated by the formation of disulfide bonds (residues 9-15) through residue H17. Although the chemical shift overlap of the H17 and A18 amide protons makes it difficult to assess the helical state of residues beyond H17, the slower exchange rate of the amide protons from O12 through A18 supports the idea that there is a continuous helix to A18. A lack of α -helicity beyond residue 18 is indicated by the strong intensities of the α H-NH NOEs beginning at residue A18, and supported by the lack of α -helical NOEs beyond A18. This conclusion is further supported by evidence that glycine is the most common residue found one residue beyond the C-terminus of a helix (Richardson & Richardson, 1988). Also, if the helix propagates to the end of the peptide, residues 19-22 would form the last helical turn. The stability of this last turn of helix is not expected to be as great as the interior turns of helix (as discussed in chapter 3.3), and has been observed in a variety of other short helical peptides (Finkelstein, et al., 1991). This termination of the Apa-Cro helix agrees with the peptide design, since the last four residues were designed to be conformationally flexible.

20 distance geometry structures were calculated using distance constraints for the first 18 residues. The NOEs from residues 19 to 22 were not included since the pattern of NOEs suggested that these NOEs were derived from a random coil, a geometrically ill-defined state. The 20 structures display a well-defined backbone conformation for the first 18 residues (Figure 4.8). The backbone RMS difference between residues 1 to 15 of Apa-Cro and apamin, 1.36 Å, and the backbone RMS difference between residues 1 to 15 of

Figure 4.8. Stereoview of the 20 distance geometry structures of the first 18 residues of Apa-Cro, showing the backbone and disulfide bridges. All structures were superimposed by a separate least-squares fit of all backbone atoms of each of 19 structures to one structure.

Apa-Cro and Apa-M5, 1.32 Å, are greater than the backbone RMSDs of the first 15 residues of the Apa-Cro ensemble, $0.57 \text{ Å} \pm 0.25 \text{ Å}$. This difference between the Apa-Cro, apamin, and Apa-M5 structures is primarily due to the displacement of residues 1, 2, 5, and 6 of Apa-M5 relative to apamin (Figure 4.9). However, the backbones of the helices, from residues 9 to 15, of Apa-Cro, apamin, and Apa-M5 are indistinguishable within the limits of resolution of the NMR data.

4.4 DNA Structure

The presence of 10 imino resonances in the spectrum of OR3/2 at 15°C indicated that OR3/2 was duplex DNA at pH 7.5 (Figure 4.10). No population of single strand DNA was detectable in the NMR sample. The four downfield and six upfield imino chemical shifts suggest that four AT and six GC base pairs exist in the duplex, although resonances were not assigned to individual protons. Upon raising the temperature to 45°C, only the resonances at 12.93 and 13.09 ppm decreased in intensity. This indicated that these resonances correspond to the C2 and C11 base pairs, and that only the other base pairs, corresponding to the OR3 operator half-site, are stable towards thermal denaturation. Similar duplex stability was observed at pH 5.5.

4.5 Peptide Function

Increments of 0.25 fractional equivalents Apa-Cro in 50 mM dibasic potassium phosphate, pH=7.5 were added to 6.75 mM of OR3/2 in an identical solvent at 15°C (Figure 4.11). No change was observed in the intensities and chemical shifts of OR3/2 imino or aromatic protons at either pH. The amide proton chemical shifts of Apa-Cro were also invariant. The chemical shift of each proton in a biomolecule in fast exchange between free and bound forms is the population-weighted average of chemical shift of each proton

Figure 4.9. Comparison of energy minimized average coordinates of the Apa-Cro and Apa-M5 distance geometry ensembles and the structure of apamin. A. The backbones of residues 1-17 were superimposed, followed by a translation of Apa-M5 and apamin, so that all peptides have the same orientation. B. A second view of the structures, rotated 90° relative to the view of A. The N-termini of the helices is closest to the viewer.

Figure 4.10. The downfield imino region of the one dimensional spectrum of OR3/2, pH 7.5, at 15°C and 45°C. While the chemical shifts of each resonance systematically change with a change in temperature, the order of the resonances does not change. Resonances above 13.3 ppm are assigned as thymidine imino protons, and the other resonances are assigned as guanine imino protons.

Figure 4.11. The imino and aromatic regions of the spectrum of Apa-Cro and OR3/2 at pH 7.5 and 15°C. The scales of the imino regions are equal, and the scales of the aromatic regions are equal, but the scales between the imino and aromatic regions are different.

in the free and bound forms. Since no changes in chemical shifts were detected, no association between Apa-Cro and OR3/2 was observed by these NMR titration experiments. The titration of OR3/2 with Apa-Cro was repeated with new samples in 50 mM monobasic potassium phosphate at pH=5.5. Again, no association between Apa-Cro and OR3/2 was observed, indicating that the lack of association is not a result of the protonation state of the histidine side chain, since titrations were conducted above and below the imidazole pK_a . Thus, the dissociation constant of the putative (Apa-Cro)-OR3/2 complex must be greater than 55 mM, assuming that a 10% population of complex is the minimum threshold of detection at the concentrations used in the NMR titration. This is in contrast to OR3 containing 5-fluorodeoxyuridine titrated with λ Cro, which displayed changes in fluorine chemical shifts with increasing λ Cro concentration (Metzler & Lu, 1989).

4.6 Molecular Modeling

A model of Apa-Cro was constructed using the energy minimized average structure of the distance geometry ensemble for residues 1 to 18, and setting residues 19 to 22 to ideal extended strand dihedral angles. This model was compared to the structure of the (λ Repressor)-OR1 complex determined by x-ray crystallography (Entry 1LRD, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987) (Figure 4.12). While the structures of λ Cro and λ Repressor differ beyond the HTH motif (Jordan & Pabo, 1988; Brennan, et al., 1986), the locations of recognition helices within the major groove of operators are expected to be similar, based upon sequence homologies of recognition helices and operators (Lewis, et al., 1983; Hochschild & Ptashne, 1986). Operators bound to λ Cro and λ Repressor are bent at the junction of the two half-sites, but each half-site conforms to B-form DNA when bound to these repressors (Kirpichnikov, et al., 1984; Metzler & Lu, 1989; Evertsz, et al., 1991; Brennan, et al., 1986; Jordan & Pabo, 1988). A comparison

Figure 4.12. The backbone and disulfide bridges of Apa-Cro, the backbone of the recognition helix of λ Repressor, and the structure of the OR1 5' half site. Structures of λ Repressor and OR1 were obtained from the (λ Repressor)-OR1 crystal structure (Entry 1LRD, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987). The backbone atoms of residues 9-17 of Apa-Cro were superimposed by a least-squares fit upon the backbone atoms of residues 44-52 of λ Repressor in the crystal structure. The view is along the average Apa-Cro and λ Repressor recognition helix axes from the N-terminal end.

of the Apa-Cro model with the structure of the (λ Cro)-OR3 complex led to conclusions with dubious validity due to the low resolution of the (λ Cro)-OR3 structure determined by x-ray crystallography (Entry 4CRO, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987).

Part of the backbone of the Apa-Cro helix (residues 8 to 17) was superimposed upon the corresponding portion of the λ Repressor helix backbone (residues 28 to 37) in the (λ Repressor)-OR1 complex to investigate the structure of the putative Apa-Cro-OR3/2 interaction. It is evident that the framework of Apa-Cro interferes sterically with the backbone of the operator half-site in this putative biomolecular association. To alleviate this steric overlap, the conformation of the half-site may change from a B-form conformation, and/or the relative orientation of Apa-Cro and OR3/2 may be adjusted. Both of these possible changes alter the interaction of the binding surfaces, and are the primary cause of the experimentally observed poor binding affinity. Adjustments within the conformation of Apa-Cro are unlikely, since the disulfide bridges hold Apa-Cro in a constrained, rigid conformation.

An additional hybrid sequence peptide, Apa-Crotate, was designed which meets the same peptide design requirements of Apa-Cro. Apa-Crotate shifts the register of the λ Cro recognition helix relative to the disulfide framework by 3 residues (Figure 4.13). This results in a clockwise 300° rotation, or a counterclockwise 60° rotation, of the disulfide framework around the cylindrical surface of the helix. As seen in an Edmundson helical wheel projection, residues required for sequence specificity and the disulfide framework are positioned along different faces of the helix (Figure 4.14). Apa-Crotate adds two residues to the N-terminus of the recognition helix. Glutamine was selected for residue 9 to model the interaction of the glutamine in the first helix of λ Cro. Lysine was selected for residue 10 to increase peptide solubility and to provide putative electrostatic interactions with the negatively charged phosphate DNA backbone.

A model of Apa-Crotate was constructed using the backbone coordinates of the

Figure 4.13. Amino acid sequences of Apa-Cro, Apa-Crotate, and the recognition helix of λ Cro, along with the corresponding numbering scheme of Apa-Cro. Disulfide bridges are indicated above the sequences of Apa-Cro and Apa-Crotate.

Figure 4.14. An Edmundson helical wheel projection of Apa-Crotate. This crude representation views the helix along the helix axis from the N-terminal end. The helical wheel depicts the radial distribution of side chains along the cylindrical helix surface. Q12, S13, N16, K17, and H20, which have been shown to be necessary for binding, are shown in bold. C11 and C15, which form disulfide bridges with the apamin framework, are shown in outline.

Apa-Cro model, setting residues 18 through 21 to ideal helical dihedral angles and residues 22 through 25 to ideal extended dihedral angles. By superimposing part of the backbone of the Apa-Crotate helix (residues Q12 to H20) upon the corresponding portion of the λ Repressor recognition helix backbone (residues 28-37) in the (λ Repressor)-OR1 complex, it is evident that the disulfide framework of Apa-Cro does not contact DNA (Figure 4.15). This suggests that Apa-Crotate may bind OR3 without steric interference from the disulfide framework, which may decrease the dissociation of the (Apa-Crotate)-OR3 complex.

Both Apa-Cro and Apa-Crotate do not include the residues of λ Cro involved in dimerization. The dimer interface of λ Cro consists of a β -strand from the carboxyl-terminus of each monomer, which pair to form an antiparallel β -ribbon (Anderson, et al., 1981; Ohlendorf, et al., 1982). Without this dimerization motif, the structure of λ Cro is not stable and λ Cro does not bind DNA (Ohlendorf, et al., 1982). Stable, monomeric variants of λ Cro have been constructed by replacing the β -ribbon dimerization motif with a designed carboxyl-terminus β -hairpin sequence (Mossing & Sauer, 1990). Binding of one of these variants, Cro.mDG, to an OR3 half-site was detected, and the dissociation constant is estimated to be 180 μ M. This indicates that dimerization is required to stabilize the HTH motif of λ Cro, but is not directly required by the HTH motif to bind an operator half-site. This suggests that a monomeric variant of λ Cro may be constructed using the disulfide framework in Apa-Crotate to stabilize the recognition helix, resulting in monomeric binding to an operator half-site.

Both Apa-Cro and Apa-Crotate do not include the residues of λ Cro involved in non-specific contacts of HTH proteins with the operator. Most importantly, a glutamine initiates the first helix of HTH proteins and participates in a hydrogen-bond network linking the first residue of the recognition helix with the second phosphate of the operator. Mutation of this glutamine to serine in λ Cro decreases operator affinity of the dimer by 4.5 kcal/mol (Pakula, et al., 1986). Since the new serine side chain may introduce unfavorable contacts in the mutant, the absence of glutamine may decrease operator affinity of a dimer

Figure 4.15. The backbone and disulfide bridges of Apa-Crotate and the structure of the OR1 5' half site in the (λ Repressor)-OR1 crystal structure (Entry 1LRD, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987). The backbone atoms of residues 12-20 of Apa-Crotate were superimposed by a least-squares fit upon the backbone atoms of residues 44-52 of λ Repressor in the crystal structure. The view is along the Apa-Crotate helix axis from the N-terminal end.

by less than 4.5 kcal/mol. Assuming that this residue is not involved in dimerization, the decrease in operator affinity of the monomer Apa-Crotate without the glutamine is less than 2.25 kcal/mol, which would increase the dissociation constant by less than 9.4 relative to an Apa-Crotate monomer with a glutamine. Residue nine of Apa-Crotate was chosen to be a glutamine to provide a crude model of this non-specific interaction.

Chapter 5. Max

5.1 Introduction

Many biologically active peptides, such as lipid-associating amphipathic helices and peptide hormones, have helix lengths longer than 14 residues (Segrest, et al., 1990). The mean helix length in known globular protein structures is 10.2 residues, with a standard deviation of 5.0 residues (Barlow & Thornton, 1988). Chapters 3 and 4 discuss the disulfide framework of apamin, which may be used to design hybrid sequence peptides to define the bound conformation of biologically active peptides and to determine the biological activity of individual protein helices. The structures of all apamin hybrid sequence peptides contain a stable helix from residues 9 to 15, while an equilibrium between helical and random coil states exists beyond residue 15. Thus, only a maximum two turns of stable helix exist in all apamin hybrid sequence peptide molecules. Helix stabilization of the C-terminus will improve the general application of hybrid sequence peptides towards the more thorough characterization of longer peptide and protein helices.

In chapter 1.2, several peptide designs were discussed which may stabilize the helix C-terminus. Only one design strategy, the application of a disulfide framework, provides helical stability to all peptide molecules. Combining such a C-terminal disulfide framework with the apamin disulfide framework in the design of a single peptide would increase the length of a stable helix from the "apamin-imum" of two turns to an "apa-maximum" of 5 turns (Figure 5.1). The C-terminal disulfide framework designed and characterized in this chapter is termed "Max" to emphasize the helical length of a proposed Apa-Max hybrid sequence peptide. The design of Apa-Max towards the study of longer peptide and protein helices is also presented.

Figure 5.1. The structures of apamin, Max, and a molecular model of Apa-Max. The structure of Apa-Max was created by fusing the C-terminus of apamin with the N-terminus of Max. The disulfide frameworks of Apa-Max are not in van der Waals contact.

5.2 Peptide Design

It is clear from simple modeling studies that the inverse of the apamin sequence, the peptide nimapa, will not form a C-terminal disulfide framework with a conformation similar to the apamin disulfide framework (Figure 5.2). The chirality of nimapa's reverse turn and the chirality of the C-terminal cysteine α -carbon causes the side chain of this C-terminal residue to extend away from the helix. To form the 8,18 disulfide bridge, residues 14 to 18 must change conformation. Thus, the effect of a conformational change in the reverse turn residues upon the structure and stability of nimapa cannot be estimated from apamin, since the reverse turn is the only prominent secondary structural element in the apamin disulfide framework. However, the model of nimapa indicates that other structural characteristics of apamin, such as the trans-proline and the putative threonine helix cap may be accommodated by a C-terminal disulfide framework.

Several other families of proteins, especially protein toxins, show strong conservation of disulfide paired cysteine residues (Kumar, et al., 1988). Such conservation of three disulfide cross-links exists among a family of many scorpion toxins and several insect defensins, in which the only additional conserved residue is a glycine (Bontems, 1991a,b) (Figure 5.3). The conservation of this sequence appears to be related to a well-defined structural requirement as inferred from the high-resolution structures of Charybdotoxin from *Leiurus quinquestriatus Hebraeus* (LqH ChTx) (Bontems, et al., 1991a) and toxin II of *Androctonus australis Hector* (AaH II) (Fontecilla-Camps, et al., 1988), and from the coordinates for variant 3 of *Centruroides sculpturatus Ewing* (CsE v3) (Fontecilla-Camps, et al., 1980) obtained from the Protein Data Bank at Brookhaven National Laboratory (Entry 1SN3, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987). Specifically, the disulfide bridges serve to define the relative position of the α -helix and β -sheet, the only two secondary structural elements found in this toxin family. Also, at the point of closest proximity between the helix and sheet, there is no room to

A.

B.

Figure 5.2. A. A molecular model of nimapa, constructed from the dihedral angles of apamin. Distances between sulfur atoms in this model which participate in disulfide bridges in apamin are displayed. B. The model of nimapa, rotated 90° from the view in A. The helix N-terminus is closest to the viewer.

Figure 5.3. The amino acid sequences of the Max peptide, *Leiurus quinquestriatus Hebraeus* scyllatoxin (LqH ScyTx), *Leiurus quinquestriatus Hebraeus* Charybdotoxin (LqH ChTx), and *Centruroides sculpturatus Ewing* variant 3 (CsE v3), along with the numbering scheme of Max. The sequences have been aligned with respect to conserved residues, and disulfide bridges are indicated above each sequence. The secondary structure elements are also indicated (alpha helix underlined, beta sheet in bold, and hairpin turn italicized).

accommodate a side chain, thus requiring a glycine residue at this site. In contrast, the regions between and beyond these secondary structures contain variable numbers and types of residues. In the high-resolution structures listed above, these sequentially variable regions also appear to be structurally variable.

The relationships between sequence and structure of this toxin family suggested that a core peptide sequence containing the conserved glycine and conserved disulfide bridges could be devised which would fold with only these conserved secondary structural elements. Such a core peptide sequence is nearly represented by scyllatoxin of *Leiurus quinquestriatus Hebraeus* (ScyTx), also known as leiurotoxin I, which is the shortest known sequence of the toxin family, containing only 31 residues (Bontems, et al., 1991a). Nuclear magnetic resonance spectroscopy studies of ScyTx have shown that the ScyTx secondary structure matches the known structures of this toxin family (Martins, et al., 1990). To obtain the desired core peptide sequence which contains the regular secondary structure elements of this toxin family, the roles of three regions of residues that were not assigned as part of secondary structure in ScyTx were evaluated. This was accomplished by using the three-dimensional structures of *LqH* ChTx (Bontems, 1991a) and *CsE* v3 (Entry 1SN3, version Jan. 1992; Bernstein, et al., 1977; Abola, et al., 1987).

The first of these regions, the C-terminal residue, clearly did not contact the secondary structures, and hence the C-terminal residue was replaced with tyrosine as a spectroscopic aid. The previous analysis of ScyTx showed that the helix and sheet may be connected only by a single residue. However, this conclusion that one and only one "unstructured" residue connects the α -helix with the β -sheet in ScyTx differs from the 3D structure of *LqH* ChTx and *CsE* v3, which contain four and five "unstructured" residues between the helix and sheet. It was clear from simple molecular modeling studies that the distance between the ends of the sheet and the helix is greater than the distance spanned by a single residue. A minimum of two residues is required to span this distance. From this analysis, G16 and L17 are expected to comprise the minimum necessary "unstructured"

region between the helix and sheet in ScyTx. This new conclusion is further supported by evidence that glycine is the most common residue found one residue beyond the C-terminus of a helix (Richardson & Richardson, 1988). Furthermore, the helix cannot extend beyond three residues after the last cysteine in the helix without necessitating additional "unstructured" residues. Thus, the sequence of ScyTx in this region was left unchanged. Finally, C3 forms a disulfide bridge with C21 in ScyTx, and the five N-terminal residues, which are not part of a regular secondary structural element, contact the β -sheet. Therefore, the influence of the five N-terminal residues upon the folding and stability of the ScyTx secondary structure was unknown. These five N-terminal residues were removed to obtain the core peptide sequence. C21 was replaced with alanine, since this residue's partner in a disulfide bridge, residue C3, was removed.

5.3 Peptide Folding

The HPLC chromatograms of oxidized Max, the reduced peptide, and the reoxidized peptide are shown in Figure 5.4. Air oxidation over several months at 4°C resulted in oxidation of 59% of methionine (as determined by electrospray mass spectrometry), leading to two chromatographic forms. This differs from the NMR experiments, which were performed immediately after peptide synthesis and purification, and used peptide with reduced methionine only. While the presence of a second form complicates the chromatograms, the retention time difference and integration ratio between the two forms aids in the identification of the peptide peaks. The HPLC chromatogram of the oxidation products of reduced peptide shows that ~99% of this peptide was oxidized. 82% of the oxidized product folded into the desired Max structure, while 14% of the oxidized product folded into one additional, unidentified peptide structure.

Upon oxidation, a peptide containing four cysteine residues can form three disulfide bridge isomers. In Max, one of these arrangements would contain a disulfide

Figure 5.4. The HPLC chromatograms of A) purified, oxidized Max, B) reduced peptide of chromatogram A, and C) reoxidized peptide of chromatogram B. The chromatography was performed using a Vydak semi-preparative column (C18, 300 Å, 10 mm x 25 cm) using a linear 18% - 36% acetonitrile gradient in 0.1% TFA over 30 minutes. The chromatogram was quantitated using a Hewlett-Packard 3392A integrator. The chromatograms were normalized with respect to the integrated intensity of peaks corresponding to Max. Each chromatogram displays an injection artifact at 6.0 min., and chromatogram B and C contain a peak at 8.7 min. corresponding to dithiothreitol. The identity of Max with and without oxidized methionine was confirmed by electrospray mass spectrometry. Other smaller peaks in these chromatograms do not correspond to the intensity ratio of Max with reduced methionine and oxidized methionine, indicating that these peaks do not correspond to Max.

bridge between C21 and C23. The equilibrium constant for the oxidized form relative to the reduced form of this type of disulfide bridge, in which only one residue intercedes between the bridged cystine pair, is much lower compared to the equilibrium constants of disulfide bridges with more interceding residues (Zhang & Snyder, 1988). While the equilibrium constants between oxidized and reduced forms of the possible disulfide bridge arrangements were not determined for the specific amino acid sequence of Max, these constants are only slightly dependent upon the nature of the interceding residue(s) (Zhang & Snyder, 1989). Thus, it is expected that this disulfide bridge is thermodynamically disfavored relative to the other disulfide combinations. It seems most likely that the unidentified peptide structure is the (3-23, 7-21) disulfide bridge arrangement of Max.

5.4 Peptide Structure

The CD spectrum of oxidized Max shows minimum at 202.8 nm and has the general character of a combination of helix and antiparallel β -sheet (Johnson, 1990) (Figure 5.5). The CD spectrum shows a slight linear decrease in intensity upon raising the temperature to 75°C, at which temperature the average intensity of the spectrum is 82.3% of the intensity at 15°C. The similarity between the CD spectra at 15°C and 75°C and the display of no cooperative, global unfolding between 15°C and 75°C indicates that the Max structure is very stable towards thermal denaturation (Figure 5.6). This is in contrast to the CD spectrum of reduced peptide at 15°C, which shows a minimum at 199.0 nm, and has the general character of a random coil.

The NMR assignments for the peptide were straightforward to obtain (Appendix 3). The assignment of the 14 methyl groups was greatly facilitated by analysis of the multiple quantum COSY spectrum. The tracing of the assignments in the amide-to-amide region is shown in Figure 5.7, and a summary of the observed sequential connectivities and longer range contacts is shown in Figure 5.8. These observed NOE connectivities

Figure 5.5. The circular dichroism spectrum of Max at 15°C (solid line), Max at 75°C (dashed line), and reduced peptide at 15°C (dotted line). CD measurements were conducted using a Jasco J500 spectropolarimeter, using a 0.2 cm path length cell, at 15°C. Samples were 84.7 μ M, and the pH was adjusted to 2.8 using 0.01 M HCl. Scans were taken from 250 to 195 nm in 0.2 nm steps with a 1.0 nm bandwidth. Sixteen scans were averaged for each spectrum, the spectrum was further processed with a three-point smoothing routine, and an identically-processed spectrum of the solvent was subtracted.

Figure 5.6. The temperature dependence of the Max circular dichroism intensity. CD measurements were carried out as described in Figure 5.5. A linear least-squares fit of the data is also shown.

Figure 5.7. NOESY spectrum of Max. Data were collected and processed as described in chapter 2.3, with a mixing time of 350 msec. Only the downfield region of spectrum is displayed. Sequential NOEs within the helix, β -sheet interstrand NOEs, and the reverse turn NOE are labeled.

Figure 5.8. A. Sequence and numbering scheme of Max with a summary of the NOE connectivities observed. The heights of the bars represent the qualitative classification of the NOE as being strong, medium, or weak. The amide protons which exchange slowly are indicated by a "+" in the NH_{ex} row.

B.

Figure 5.8. B. The antiparallel beta sheet, reverse turn, and other selected residues of Max. Observed interstrand NOEs are indicated by arrows. Observed NOEs which help determine the relative orientation of the helix and sheet are also shown. The disulfide bridges between Cys-3 and Cys-21, and Cys-7 and Cys-23, are also indicated.

and amide exchange rates for Max were found to be similar to those observed for ScyTx, indicating that the structure of Max contains the same α -helix and antiparallel β strand observed in ScyTx and the other proteins of this toxin family. It was also observed that the chemical shifts of Max and ScyTx amide proton resonances were very similar for most residues, which served as an additional indication that the structures of Max and ScyTx were similar (Figure 5.9). The chemical shift differences of the amide protons of residue 1, 2, 16, 17, and 22 are most likely due to the presence of the additional five N-terminal residues in ScyTx and the substitution of C21 in ScyTx for an alanine in Max.

The backbone and disulfide bridges of the 20 3D-structures of Max derived from distance geometry are displayed in Figure 5.10. Max has a compact structure consisting of a helix, a small, double-stranded, antiparallel sheet, and a type I reverse turn. The disulfide bridges connect the helix and sheet, and both bridges conform to a left-handed corkscrew configuration. The helix and sheet are connected by G11 and L12, which do not belong to these secondary structures (defining secondary structures either through hydrogen bonding or α carbon position). This agrees with the peptide design analysis of this region as described above.

The average RMS deviations of the backbone of the helix, sheet, and reverse turn, when each individual secondary structural element is superimposed, indicate that each individual region of the peptide is well-defined. However, the RMS deviations of these backbone segments when the entire backbone is superimposed clearly shows that the position of the reverse turn relative to the entire backbone varies to a significantly greater extent than the other regions of the peptide, as can also be seen in Figure 5.10.A. This indicates that the turn is a well-defined structural element, but the orientation of the turn relative to the rest of the peptide, especially the sheet, is variable. This variability in the distance geometry structures has several possible explanations. First, the NOE distance constraints which define the orientation of the turn to the rest of the peptide may not have been identified in this analysis. Since the NMR spectra were straightforward to

Figure 5.9. The chemical shift differences (in ppm) of amide proton chemical shifts of Max relative to ScyTx.

Figure 5.10. A. Stereoview of the 20 distance geometry structures showing only the backbone and disulfide bridges, and the average structure of this ensemble. All structures were superimposed by a separate least-squares fit of all backbone atoms of each of 19 structures to one structure. B. Same as in 5.10.A with a 90 degree rotation parallel to the helix axis.

assign, and the spectral sensitivity was excellent, this possibility appears unlikely. A second possibility is that a rigid conformation of the turn relative to the peptide exists, but that distances between protons of the turn and the rest of the peptide (which would define this rigid conformation) are greater than approximately 5 Å, and thus do not give rise to an observable NOE. However, protons from residue 17 in the turn give rise to observable NOEs to protons within the helix, which help to define the relative orientation of these secondary structural elements. Lastly, it is possible that the distance geometry structures accurately reflect that the connection between the turn and the sheet serves as a hinge, allowing the turn to concertedly move relative to the sheet. This possibility best explains how NOEs can be generated between protons of residue 17 and protons of the helix, while the conformation of the turn can appear to be variable in the distance geometry structures.

Table 5.2. Average RMS deviations of the backbones of the 20 Max distance geometry structures.

Backbone Region	Residues	Superposition of Backbone Region	Superposition of Entire Backbone	Difference Between Superpositions Using Different Regions
Helix	1-10	0.485 Å	0.861 Å	0.377 Å
Loop	11-12	0.383	1.096	0.713
Sheet	13-16, 21-24	0.658	0.858	0.201
Turn	17-20	0.465	1.780	1.316
Entire	1-26	1.171	1.171	0

5.5 Molecular Modeling

A comparison of the α carbon trace of Max and CsE v3 shows that the dimensions of the helix backbones and sheet backbones are very similar between these peptides (Figure 5.11.A). Furthermore, three more residues connect the helix and sheet in CsE v3 than in Max. The comparison of the α carbon traces shows that residues 31, 35, and 36 of CsE v3 superimpose well upon the backbone of residues 10, 11, and 12 of Max. Since the comparison of the α carbon traces clearly shows that the conformation of residues 32 to 34 in CsE v3 do not correspond to the conformation of residues in Max,

Figure 5.11. Comparison of the α carbon traces of Max (solid line) and CsE v3 (dashed line). A. The structures were superimposed by a least-squares fit of all backbone atoms. B. The structures were superimposed by a least-squares fit of the backbone atoms of the Max helix (residues 1 to 8) and CsE v3 helix (residues 23-30).

residues 32 to 34 in *CsE v3* seem not to be a structural requirement. When only the helix backbones of these peptides are superimposed, it is evident that the intersection of the helix and sheet in each peptide is different (Figure 5.11.B). Specifically, the axis of the sheet forms a -16 degree angle with the axis of the helix in Max, which differs from the corresponding -45 degree and -42 degree angles in *LqH ChTx* and *CsE v3*, respectively (Bontems, et al., 1991a). This difference can be attributed to the presence of the third disulfide bridge in *LqH ChTx* and *CsE v3*. To accommodate this disulfide bridge, the end of the sheet which is connected by the reverse turn must be extended above and away from the amino-terminal end of the helix. Since this end of the sheet has been shown to be antigenic in *AaH II*, the third disulfide bridge may be an important structural element, preventing the end of the sheet from packing against the helix and providing for more solvent exposure of this antigenic peptide segment (Fourquet, et al., 1988). The absence of the third disulfide bridge in Max allows the sheet to cross the helix in a more parallel fashion, which increases the surface area in contact between the helix and sheet (Figure 5.12) (Chothia, 1984). This change in the crossing angle also results in a shift in the register of the "knobs into holes" packing of the helix and sheet side chains. Most importantly, the conservation of G26 has been rationalized for the structure of *LqH ChTx*, but the different packing arrangement in Max does not seem to necessitate the conservation of this residue. Based upon modeling the glycine in Max may be replaced with alanine, with no steric clash of the side chain with other parts of the peptide.

5.6 Hybrid Sequence Peptides

By comparing the sequences of these toxin proteins, it is clear that a variety of amino acid sequences can be accommodated within the folded, stabilized three dimensional structure, with the requirement of only two conserved disulfide cross-links. These folding and stability properties of Max's disulfide framework are similar to those of the disulfide

Figure 5.12. A. The helical net diagram of Max, showing the relative position of the α carbons of the helix and sheet. Gray circles represent the α carbons of the unwound helix which lays flat in the plane of the page. The β -sheet lays above the helix, and the side chains alternately point away from the helix (white circles) or towards the helix (black circles). The scale used to define the distance between the α carbons of the sheet reflects the average scale of the projection of these atoms of the sheet onto the helix plane. B. The helical net diagram of CsE v3, using the same circles and similar orientation as in A. The scale used to define the distance between the α carbons of the sheet of CsE v3 again reflects the average scale of the projection of these atoms onto the CsE v3 helix plane. This scale differs slightly from the scale used for the sheet in A.

framework of apamin, which was used to make hybrid sequence peptides in chapters 3 and 4. It may be possible to use the disulfide framework of Max to make similar hybrid sequence peptides. The sequence of a helical region of Max may be substituted by the helical sequence of an unrelated protein. A different type of hybrid sequence peptide may be designed by substituting the sequence of an antiparallel sheet region of an unrelated peptide for the sheet sequence in Max. Since the packing arrangement of Max seems not to require the conservation of the glycine residue in the sheet sequence, this residue may also be substituted. Lastly, the sequence of the reverse turn in Max may be substituted by the reverse turn sequence of an unrelated peptide. In all three types of hybrid sequence peptides, the cysteine residues in the Max sequence must be retained to make the same disulfide framework as in Max. It is postulated that such hybrid sequences will fold with high yield to form stable structures that are similar to the structure of Max.

The sequences of apamin hybrid peptides contain 3 to 12 residues beyond the last cysteine. The sequences of Max hybrid peptides may also contain more than two residues prior to the first cysteine. However, if these N-terminal residues are helical, the backbone of the residue third residue prior to the first cysteine sterically interferes with the side chain of I17 (Figure 5.13). To alleviate this steric overlap with the putative N-terminal helix, the side chain of I17 must move away from the helix or must be removed. The concerted motion of the reverse turn relative to the helix and sheet in Max suggests that I17 may move away from the helix. If I17 cannot move, only the substitution of glycine for I17 results in no steric contacts with the putative helix.

If the Max hybrid sequence peptide can accommodate an N-terminal extension of the helix, then the disulfide frameworks of apamin and Max may be combined to design an Apa-Max hybrid sequence peptide. A model of Apa-Max was constructed using the energy-minimized average structures of the Apa-M5 and Max distance geometry ensembles, and setting residues 17 to 27 to ideal helical dihedral angles (Figure 5.1). No steric contacts exist between the N-terminal and C-terminal disulfide frameworks. 105

A.

B.

Figure 5.13. Stereoview of the backbone and the side chain of I17 of the energy-minimized structure of Max. A. The view is parallel to the helix axis. B. The view is along the helix axis from the N-terminus.

disulfide bridge isomers may be formed by the oxidation of the eight cysteine residues in Apa-Max. While the folding of apamin and Max occur with high yield, the folding of Apa-Max to form the designed disulfide bridge isomer is unknown.

Chapter 6: Apa-AlaN

6.1 Introduction

The venom of the honeybee, *Apis mellifera*, contains a set of structurally stable peptides, including MCD, tertiapin, and apamin (Kumar, et al., 1988). These peptides share the same (1-11, 3-15) disulfide bridge arrangement and some homology in amino acid sequence. In particular, asparagine is conserved immediately after the first cysteine, and proline is conserved between the second and third cysteines. Both of these residues may direct the folding of honeybee venom peptides to form the (1-11, 3-15) disulfide bridge arrangement. In the distance geometry structures of apamin (Pease & Wemmer, 1988), Apa-M5 (chapter 3.3), and Apa-Cro (chapter 4.3), the asparagine is the first position of the reverse turn, and the side chain of this asparagine is in a position to form a hydrogen bond with the amide proton of K4. This stabilization of a Type 1 reverse turn has been observed in other proteins (Richardson, 1981). In the same distance geometry structures, P6 contributes to the hydrophobic contacts between the helix and disulfide framework. The rigidity of proline also restricts the conformational freedom of this residue and the preceding residue during folding. In contrast, the family of endothelin and sarafotoxin peptides, which contain cysteine residues at positions 1, 3, 11, and 15, do not share the same homology of honeybee venom peptides beyond the cysteines, and do not form the (1-11, 3-15) disulfide bridge arrangement, (Kitazumi, et al., 1990). The first experiments of this chapter investigate the sequence requirements for forming the (1-11, 3-15) disulfide bridge arrangement during the folding of honey bee venom peptides and apamin hybrid sequence peptides.

The elucidation of protein dynamics has lead to an understanding of the special dynamic features required for the performance of specific functions (Heinz, et al., 1992). Furthermore, regional differences in protein dynamics may be correlated with the sequence

of protein folding events, although direct correlations are not currently known (Scholtz & Baldwin, 1992). Finally, a direct correlation between dynamics and resolution of three dimensional NMR structures has been established for a variety of proteins (Clare, et al., 1990; Stone, et al., 1992; Barbato, et al., 1992; Powers, et al., 1992; Stone, et al., 1993).

Structural comparisons in chapters 3 and 4 indicate that the structures of all apamin hybrid sequence peptides contain a stable helix from residues 9 to 15, while an equilibrium between helical and random coil states exists beyond residue 15 to the C-terminus. In this chapter, the dynamics of the C-terminal residues are measured and interpreted. This description of dynamics contributes to the structural characterization of apamin hybrid sequence peptides. These dynamics measurements may also contribute to the design of hybrid sequence peptides with special dynamic features required for the performance of specific functions.

The results of this chapter provide the first description of α -helix dynamics of a peptide consisting of natural amino acids in water, using NMR relaxation measurements. A peptide incorporating unnatural α -aminoisobutyric acids in methanol has been shown to form a rigid rod with anisotropic dynamics using NMR relaxation (Kelsh, et al., 1992). The dynamics of peptide α -helices have also been investigated using other techniques. The distribution of helicity in two different peptides was investigated by measuring the CD signal of two series of peptides in which residues were systematically replaced with glycine (Chakrabarty, et al., 1991). These investigations revealed a non-monotonic decrease in fractional helicity of each residue from the center to the ends of the peptide, as expected from Lifson-Roig theory (Lifson & Roig, 1961). Recent studies by ESR spectroscopy revealed that the 4 C-terminal residues of alanine-based peptides have dynamics which monotonically increase with decreasing distance from the C-terminus (Miick, et al., 1993). The measurement of NMR spectroscopic α carbon chemical shifts and $J_{NH-\alpha H}$ in a helical peptide stabilized by a series of EK salt bridges revealed a gradual decrease in helicity of the eight C-terminal residues (Liff, et al., 1991). The description of apamin hybrid

sequence peptide dynamics is compared with these results.

The dynamic rates and amplitudes of spectral density functions have been analyzed from heteronuclear NMR relaxation rates in other proteins and peptides. To within the accuracy of the data, dynamic rates and amplitudes of helices, sheets, turns, and loops are indistinguishable (Kay, et al., 1989; Schneider, et al., 1992). However, helical dynamics from 10^8 to 10^3 sec^{-1} have narrower distribution compared to other secondary structures, indicating that motions of this frequency range are more greatly correlated in helices (Kay, et al., 1989). Furthermore, a strong correlation exists between the amplitude of motion and hydrogen bonding of the amide hydrogen or its peptide bond associated carbonyl (Schneider, et al., 1992). The protein C-terminal residues which are not involved in tertiary contacts have been shown to rotate with increasing mobility toward the C-terminus, but the C-terminal residue fails to reach free rotation (Schneider, et al., 1992). The correlated motion of the C-terminal residues in apamin hybrid sequence peptides is also compared with these results.

6.2 Peptide Design

The amino acid sequences of apamin and the hybrid Apa-Ala_N are shown in Figure 6.1. The hybrid peptide numbering scheme will be used to refer to residues in the remainder of this chapter. Except for the two cysteines, which are required to make stabilizing disulfide bridges, the sequence of the helix in the hybrid peptide (residues 9 to 22) was designed based upon sequence-dependent properties that stabilize α -helical conformations (as discussed in chapter 1.2). The selected amino acids each have high helical propensities. E10 and K14 are designed to form a stabilizing $i,i+4$ salt bridge, with the negatively-charged glutamate side chain placed nearer the N-terminal helix dipole. Residue 19 was chosen to be lysine to decrease the hydrophobicity of the C-terminal region. The amidated C-terminus serves to stabilize the helix by decreasing the helix

MCD:	IKCNCKRHVIKPHICRKICGKN
tertiapin:	ALCNCNR II IPHMCWKKCGKK
apamin:	CNCK AP ETALCARRCQQH
Apa-Ala N :	CNCK AP ETAECAYK CAA AKAAA-NH ₂
Apa-2A:	CACK AP ETAECAYK CAA AKAAA-NH ₂
Apa-6A:	CNCK AA ETAECAYK CAA AKAAA-NH ₂
Apa-Ala:	CNCK AP ETAACKYE CAA AKAAA-NH ₂
endothelin-1:	CSCS SLMDKECVYFCHLDI IW
sarafotoxin-6b:	CSCK DMTDKECLYFCHQDVIW

Figure 6.1. Amino acid sequences of MCD, tertiapin, apamin, Apa-Ala**N**, Apa-2A, Apa-6A, Apa-Ala, endothelin-1, and sarafotoxin-6b along with the corresponding numbering scheme of Apa-Ala**N**. ¹⁵N-labeled alanine residues are shown in bold.

dipole. Finally, Y13 serves as a spectroscopic aid.

The NMR frequencies of the alanine residues in Apa-AlaN were expected to be very similar, based upon the two dimensional spectrum of a related peptide, Apa-Ala. These resonance degeneracies lead to problems with assignment and quantitation of homonuclear two dimensional NMR spectra. To circumvent the degeneracies of alanine resonances, ^{15}N -labeled Fmoc alanine was used in the synthesis of Apa-AlaN at positions 16, 17, 18, 20, 21, and 22. ^{15}N -labeled Fmoc alanine was also used in the synthesis at position 12, to serve as a reference for the measurement of helix dynamics. A second synthesis of Apa-AlaN was performed with no ^{15}N -labeled amino acids for folding studies.

6.3 Peptide Folding

The folding of Apa-AlaN was compared to two analogs, Apa-2A and Apa-6A, which substitute an alanine residue for the asparagine or proline residues (Figure 6.1). Peptides were oxidized and folded and the disulfide isomers were quantitated by HPLC as described in chapter 2.2. HPLC experiments were repeated during each folding reaction until three final experiments produced similar results. The third chromatogram of the final folded forms of each peptide appear in Figure 6.2. Relevant parameters of the three final chromatograms of each peptide appear in Table 6.1. The standard deviations of % isomeric ratios were well correlated with the value of the % isomeric ratios, indicating that the uncertainty of the chromatography was greater than the change in these values over time. This indicates that the folding reaction reached equilibrium, assuming that no folding intermediate is stable for longer than 80 minutes (the duration of the first two of the final three HPLC experiments). The high quality of the HPLC was reflected in the low values of these standard deviations, and in the standard deviations of the retention times (0.2 to 0.4 min), which was much less than the difference between retention times of different isomers.

Figure 6.2. The HPLC chromatograms of oxidized A) Apa-AlaN B) Apa-2A C) Apa-6A. The chromatography was performed using a Vydak semi-prepartive column (C18, 300 Å, 10 mm x 25 cm) using a linear 15% - 26.4% acetonitrile gradient in 0.1% TFA over 22 minutes, followed by a 26.4% - 100% acetonitrile gradient in 0.1% TFA over 3 minutes. The chromatogram was quantitated using a Hewlett-Packard 3392A integrator. Each chromatogram displays an injection artifact from 7 to 8 minutes, and hydrophobic material eluted during the acetonitrile wash starting at 29 minutes.

Table 6.1. Folding Products of Alanine Monosubstituted Analogs of Apa-AlaN

Peptide	% Isomeric Ratio¹				HPLC Retention Time (min.)			
	0	1	2	3	0²	1	2	3
Apa-AlaN	n.d. ³	2.9±0.5	93.6±0.2	3.4±0.6	14.00	15.58	16.74	18.50
Apa-2A	n.d.	23.5±0.4	73.2±0.4	3.3±0.5	17.22	18.04	19.79	23.66
Apa-6A	2.4±0.6	15.4±0.4	80.7±0.2	1.4±0.6	15.63	20.93	22.30	25.76

1. Isomer assignments: 0 = reduced peptide
 1 = (1-15, 3-11) disulfide bridge arrangement
 2 = (1-11, 3-15) disulfide bridge arrangement
 3 = (1-3, 11-15) disulfide bridge arrangement
2. Retention times measured from HPLC chromatogram of reduced peptide.
3. n.d. = not detected

Upon oxidation, a peptide containing four cysteine residues can form three disulfide bridge isomers. In Apa-AlaN, one of these arrangements would contain a disulfide bridge between C1 and C3. The equilibrium constant for the oxidized form relative to the reduced form of this type of disulfide bridge, in which only one residue intercedes between the bridged cystine pair, is much lower compared to the equilibrium constants of disulfide bridges with more interceding residues (Zhang & Snyder, 1988). While the equilibrium constants between oxidized and reduced forms of the possible disulfide bridge arrangements were not determined for the specific amino acid sequence of Apa-AlaN, these constants are only slightly dependent upon the nature of the interceding residue(s) (Zhang & Snyder, 1989). Thus, it is expected that this disulfide bridge is thermodynamically disfavored relative to the other disulfide combinations. It seems most likely that isomer three is the (1-3, 11-15) disulfide bridge isomer of Apa-AlaN, since folding Apa-AlaN, Apa-2A, and Apa-6A results in a low relative population of this isomer.

The HPLC purification and NMR analysis of peptides similar to Apa-AlaN, Apa-2A, and Apa-6A, including Apa-Ala, have shown that a large majority of these peptides fold to form the (1-11, 3-15) disulfide bridge arrangement. Since isomer two comprises a large majority of the folded forms of Apa-AlaN, Apa-2A, and Apa-6A, this isomer is the

(1-11, 3-15) disulfide bridge isomer. Finally, isomer two is the (1-15, 3-11) disulfide bridge isomer. The order of HPLC retentions is the same for these disulfide bridge isomers and the isomers of endothelin-1, further supporting the assignments of these disulfide bridge isomers of Apa-Ala_N (Kumagaye, et al., 1988).

The % isomeric ratios indicate that both N2 and P6 help direct the folding of honeybee venom peptides and apamin hybrid sequence peptides to form the (1-11, 3-15) disulfide bridge arrangement. If the effects of these residues upon folding are noncooperative, these results suggest that only 61.0 % of an analog of Apa-Ala_N with alanine substitutions for both N2 and P6 would fold to form the (1-11, 3-15) disulfide bridge arrangement. Such a result would indicate that the asparagine and proline residues are almost completely responsible for the folding of honeybee venom peptides and apamin hybrid sequence peptides to form the (1-11, 3-15) disulfide bridge arrangement.

6.4 Peptide Structure

The structures of the peptides Apa-Ala_N and a related peptide Apa-Ala, are expected to be the same, based upon the similarity of the structures of residues 1-15 of apamin hybrid sequence peptides, and the similar C-terminal sequences of Apa-Ala_N and Apa-Ala (Figure 6.1). The CD spectrum of the oxidized hybrid peptide Apa-Ala is very similar to that of apamin, with a minimum at 208 nm and a shoulder at 222 nm (Figure 6.3) (Hider & Ragnarrson, 1981). This spectrum has the general character of that from an α -helix, but must have contributions from the β -turn and loop in apamin as well, since the 208 nm / 220 nm ratio and the shape of the CD spectrum are significantly different from purely helical peptides (Johnson, 1990). However, the spectrum shape is very reproducible among the hybrid sequence peptides, and seems to be indicative of the folded core structure (Pease, et al., 1990). The Θ_{222} observed corresponds to approximately 47% of the peptide being helical at 10°C (Klevit, et al., 1985). If this helicity is attributed only to the helical region

Figure 6.3. The circular dichroism spectrum of Apa-Ala. CD measurements were carried out on a Jasco J500 spectropolarimeter using a 0.2 cm pathlength cell. Samples were approximately 30.6 mM, in 0%, 30%, and 40% trifluoroethanol and 15°C. Scans were taken from 250 to 205 nm in 0.2 nm steps with a 1.0 nm bandwidth. Sixteen scans were averaged for each spectrum, the spectrum was further processed with a three-point smoothing routine, and an identically-processed spectrum of the solvent was subtracted.

of Apa-Ala, then 73%, or 11 of the 15 residues, of this region is helical. If the helix propagates to the end of the peptide, residues 19-22 would form the last helical turn. While the backbone amides of these residues would be involved in the helix hydrogen bond network, the backbone carbonyls of these residues would not be involved in backbone or side chain hydrogen bonds, and thus the stability of this last turn of helix is not expected to be as great as the interior turns of helix. The lack of structural stability of the last turn has been observed in a variety of other short helical peptides (Finkelstein, et al., 1991). Addition of 30% trifluoroethanol increases the helicity to 66% (Figure 6.3). If this helicity is attributed only to the helical region of Apa-Ala, then 99% of this region is helical. This is similar to the effect of trifluoroethanol upon helix propagation in a similar peptide, Apa-S (Storrs, et al., 1992). The CD spectrum shows a slight linear decrease in intensity upon raising the temperature to 70°C (Figure 6.4). The similarity between the CD spectra at 10°C and 70°C and the display of no cooperative, global unfolding between 10°C and 70°C indicates that the Apa-Ala structure is stable towards thermal denaturation.

The NMR assignments for Apa-Ala_N were straightforward to obtain (Appendix 3). The tracing of the assignments of the amide region is shown in the homonuclear NOESY spectrum (Figure 6.5) and in the strip plot of the NOESY-HMQC experiment (Figure 6.6). A summary of the sequential connectivities observed and longer range contacts in the helix are shown in Figure 6.7. The types of connectivities observed are exactly the same as those in both apamin (Pease & Wemmer, 1988), the Apa-S hybrid (Pease, et al., 1990), Apa-M5 (chapter 3.3), Apa-Cro (chapter 4.3), and Apa-Ala for the first 16 residues. These include the α H-NH-NH-NH connectivity characteristic of the type I β -turn, and $i,i+2$ connectivity at the beginning of the helix, and the continued NH-NH connectivities in the α -helix from residue 9 through residue 16. The presence of the helix is verified by a number of $i,i+3$, α H-NH, and α H- β H connectivities, which are normally seen in this secondary structure. Additional tertiary NOEs are observed which indicate that the fold of the first 16 residues is very similar to apamin (Pease & Wemmer, 1988), Apa-S (Pease, et

Figure 6.4. The temperature dependence of the Apa-Ala circular dichroism intensity. CD measurements were carried out as described in Figure 6.2. The temperature was varied between each measurement, and the same sample was used for all measurements. The linear least squares fit of the data is also shown.

Figure 6.5. NOESY spectrum of Apa-Ala_N. The sample was 10 mM in 90% H₂O/10% D₂O, at 20°C. Data were collected and processed as described in chapter 2.3. Only the downfield region of the symmetrical 6250 Hz spectrum is displayed. Sequential NOEs within the helix and reverse turn are labeled.

Figure 6.6. The strip plot of the NOESY-HMQC spectrum of Apa-Ala_N. Sample conditions are described in Figure 6.5. Data were collected and processed as described in chapter 2.3. Only the downfield region is displayed. Sequential NOEs within the helix are labeled. See also Figure 2.11.

Figure 6.7. A summary of the sequential and long-range NOE connectivities observed in the NOESY spectrum of Figures 6.5 and 6.6. The sizes of the bars reflect the relative NOE intensities, and question marks reflect indeterminable NOE intensities due to resonance degeneracies.

al., 1990), Apa-M5 (chapter 3.3), Apa-Cro (chapter 4.3), and Apa-Ala. These include contacts between residues C3 & P6, P6 & R12, and C3 & C15.

Partial NOE connectivities in the NMR spectra of the Apa-M5 hybrid indicate that the helix propagates in a cooperative fashion from the section nucleated by the formation of disulfide bonds (residues 9-15). However, beginning at residue 15, there is an increase in the intensity of the NH- α H NOEs. This indicates that there is an equilibrium between helical and random coil states from residue 16 through 22. Since long range and short range helical NOEs start to decrease in intensity at different residues, no single residue appears to serve as the point of transition from helix to random coil. Rather, the data indicate that this transition occurs between residues 18 and 20. If the helix propagates to the end of the peptide, residues 20 to 22 and the amide C-terminus would form the last helical turn. While the backbone amides of these residues would be involved in the helix hydrogen bond network, the backbone carbonyls of these residues would not be involved in backbone or side chain hydrogen bonds, and thus the stability of this last turn of helix is not expected to be as great as the interior turns of helix. The lack of structural stability of the last turn has been observed in a variety of other short helical peptides (Finkelstein, et al., 1991).

6.5 Theory of Peptide Dynamics

The relaxation superoperator discussed in chapter 2.3 can be described as:

$$\hat{\Gamma}\{\sigma\} = (0.5)SJ(\omega_q)[H-q[Hq, \sigma]] = \text{relaxation superoperator}$$

where $J(\omega_q)$ = power spectral densities

Hq = eigenoperators acting only on spin system

where q = irreducible tensor components and different types of fluctuating interactions

In the case of isotropic random motion with correlation time τ_c , the spectral densities are

described as:

$$\mathcal{J}(\omega) = \frac{2\tau_c}{1 + (\omega\tau_c)^2}$$

For Apa-AlaN, the spectral density of ^{15}N nuclear spins depends upon the overall motions of the macromolecule (τ_R) and on the internal motions of the covalently bonded proton relative to the ^{15}N nucleus (τ_f). The minimum number of unique motional parameters can be evaluated using a model-free approach, which again assumes that the overall tumbling of Apa-AlaN is isotropic (Lipari & Szabo, 1982):

$$J(\omega) = \frac{S_f^2 \tau_R}{1 + (\omega\tau_R)^2} + \frac{(1 - S_f^2) \tau_{Rf}}{1 + (\omega\tau_{Rf})^2}$$

where S_f^2 = generalized order parameter for τ_f

$$\frac{1}{\tau_{Rf}} = \frac{1}{\tau_R} + \frac{1}{\tau_f}$$

if $\tau_f \ll \tau_R$, then $\frac{1}{\tau_{Rf}} \approx \frac{1}{\tau_R}$

$$J(\omega) \approx \frac{S_f^2 \tau_R}{1 + (\omega\tau_R)^2}$$

An additional term may be included to determine additional dynamic motion on a time scale (τ_s and S_s^2) between τ_f and τ_R in the model-free formalism of an isotropically tumbling Apa-

AlaN (Powers, et al., 1992):

$$J(\omega) = \frac{S_f^2 S_s^2 \tau_R}{1 + (\omega\tau_R)^2} + \frac{(1 - S_f^2) \tau_{Rf}}{1 + (\omega\tau_{Rf})^2} + \frac{S_f^2 (1 - S_s^2)}{1 + (\omega\tau_{Rs})^2}$$

where $\frac{1}{\tau_{Rs}} = \frac{1}{\tau_R} + \frac{1}{\tau_s}$

if $\tau_f \ll \tau_s$, then $J(\omega) \approx \frac{S_s^2 \tau_R}{1 + (\omega\tau_R)^2} + \frac{S_f^2 (1 - S_s^2) \tau_{Rs}}{1 + (\omega\tau_{Rs})^2}$

Anisotropic motion on the τ_R and τ_f time scales is approximated by a similar formalism for ellipsoids (Woessner, 1962):

$$J(\omega) = S_f^2 \left\{ \frac{A_1 \tau_1}{1 + (\omega\tau_1)^2} + \frac{A_2 \tau_2}{1 + (\omega\tau_2)^2} + \frac{A_3 \tau_3}{1 + (\omega\tau_3)^2} \right\} + \frac{(1 - S_f^2) \tau_{Rf}}{1 + (\omega\tau)^2}$$

where $A_1 = 0.75 \sin^4 \Theta$

$$A_2 = 3 \sin^2 \Theta \cos^2 \Theta$$

$$A_3 = (1.5 \cos^2 \Theta - 0.5)^2$$

Θ = angle between ^{15}N - ^1H bond and major ellipsoid axis

$$\tau_1 = (4D_{\parallel} + 2D_{\perp})^{-1}$$

$$\tau_2 = (D_{\parallel} + 5D_{\perp})^{-1}$$

$$\tau_3 = (6D_{\parallel})^{-1}$$

$$1/\tau_{\text{Rf}} = 1/\tau_{\text{R}} + 1/\tau_{\text{f}}$$

$$\approx (2D_{\parallel} + 4D_{\perp})^{-1}$$

where $D_{\parallel}$ = rate of diffusion parallel to major ellipsoid axis

$D_{\perp}$ = rate of diffusion perpendicular to major ellipsoid axis

This formalism is described by seven parameters, while only three relaxation rates are determined in this analysis. Assuming that $D_{\parallel}$ and $D_{\perp}$ are not correlated, this formalism can be interpreted as (Lipari & Szabo, 1982):

$$J(\omega) = 0.4 \left\{ S_f^2 \left[\frac{A \tau_{\text{R}\parallel}}{(1 + \omega^2 \tau_{\text{R}\parallel}^2)} + \frac{(1 - A) \tau_{\text{R}\perp}}{(1 + \omega^2 \tau_{\text{R}\perp}^2)} \right] + (1 - S_f^2) \left[\frac{A \tau_{\text{Rf}\parallel}}{(1 + \omega^2 \tau_{\text{Rf}\parallel}^2)} + \frac{(1 - A) \tau_{\text{Rf}\perp}}{(1 + \omega^2 \tau_{\text{Rf}\perp}^2)} \right] \right\}$$

where A = anisotropic coefficient

$\tau_{\text{R}\parallel}$ = τ_{R} parallel to major ellipsoid axis

$\tau_{\text{R}\perp}$ = τ_{R} perpendicular to major ellipsoid axis

$\tau_{\text{Rf}\parallel}$ = τ_{Rf} parallel to major ellipsoid axis

$\tau_{\text{Rf}\perp}$ = τ_{Rf} perpendicular to major ellipsoid axis

$$\text{If } \tau_{\text{f}} \ll \tau_{\text{R}}, \text{ then } J(\omega) \approx \frac{S_f^2 A \tau_{\text{R}}}{1 + (\omega \tau_{\text{R}})^2} + \frac{S_f^2 (1 - A) \tau_{\text{R}}}{1 + (\omega \tau_{\text{R}})^2}$$

This interpretation of the anisotropic model is equivalent to the model-free formalism which incorporates the additional terms τ_{s} and S_{s}^2 .

In general, τ_{f} and τ_{s} are superpositions of the products of internal correlation times and geometric factors. Therefore, detailed interpretations of these correlation times may not be derived for the complex structures of peptides. The order parameter S_{f} of each of these spectral density functions is related to the amplitude of the internal motions. For free

diffusion of the NH bond vector in a cone (Kay, et al., 1989) (Figure 6.8.A):

$$S_f^2 = [0.5 \cos \theta (1 + \cos \theta)]^2$$

where θ = semiangle between NH bond vector at equilibrium and during
dynamic motion

Motions related to τ_s may be interpreted as a discrete jump of the NH bond vector between two states, a model which retains the independence of the interpretation of S_f^2 (Clare, et al., 1990) (Figure 6.8.A):

$$S_s^2 = P_i P_j (1 + 3 \cos^2 \phi_s) / 4$$

where P_i = normalized population of NH bond vectors in state i

P_j = normalized population of NH bond vectors in state j

ϕ_s = angle between NH bond vectors of states i and j

The order parameters of successive residues in peptides are correlated through the motions of the ψ backbone angle (Schneider, et al., 1992) (Figure 6.8.B):

$$S_{c,i}^2 = \frac{3 \sin^2 \theta \sin^2 \psi_{i-1}}{\psi_{i-1}^2} (\cos^2 \theta + 0.25 \sin^2 \theta \cos^2 \psi_{i-1})$$

where ψ_{i-1} = ψ backbone angle of residue i-1 in radians

Assuming that the peptide bond is rigid with sp^2 geometry and that residue i is completely free to rotate about ψ , it can be shown that

$$\frac{S_{f,i}^2(\text{observed})}{S_{f,i-1}^2(\text{observed})} = \frac{S_{f,i}^2(\text{theoretical}) S_{c,i}^2}{S_{f,i-1}^2(\text{theoretical})} = 0.111$$

The relaxation superoperator can be evaluated for each of these representations of the spectral densities. Only dipolar and chemical shift anisotropy relaxation mechanisms significantly contribute to the relaxation of ^{15}N in Apa-Ala N . In the absence of cross-correlation between dipolar and chemical shift anisotropy relaxation mechanisms (Boyd, et al., 1991) the rates of longitudinal relaxation (T_1), transverse relaxation (T_2), and nuclear Overhauser effect (NOE) can be determined from this evaluation (Figure 6.9):

$$1/T_1 = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \text{ in sec rad}^{-1}$$

A.

B.

Figure 6.8. A. Model of the motions on time scales τ_f and τ_s . The faster motion is represented as free diffusion within an axially symmetric cone with semiangle θ . The slower motion is represented by a discrete jump between two states i and j separated by angle ϕ_s . B. The rotation axes and angles which define the correlated motions of the amide nitrogens of two sequential residues.

Figure 6.9. The T_1 , T_2 , and NOE relaxation times for ^{15}N relaxed by ^1H - ^{15}N dipolar and chemical shift anisotropy interactions at a field strength of 14089 G. The ^{15}N was assumed to be tumbling isotropically, $S^2 = 1.0$, and $r = 1.02 \text{ \AA}$.

$$1/T_2 = (d^2/2)[4J(0) + J(\omega_H - \omega_N) + 6J(\omega_N) + 3J(\omega_H) + 6J(\omega_H + \omega_N)] \\ + (c^2/6)[3J(\omega_N) + 4J(0)] \text{ in sec rad}^{-1}$$

$$\text{NOE} = 1 + \{\gamma_H/\gamma_N\}d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)] \text{ in sec rad}^{-1}$$

$$\text{where } \omega_H = 3.771 \times 10^9 \text{ rad sec}^{-1}$$

$$\omega_N = 3.821 \times 10^8 \text{ rad sec}^{-1}$$

$$d^2 = 0.1 \gamma_H^2 \gamma_N^2 h^2 / (4\pi^2) \langle 1/r_{HN}^3 \rangle^2 = 5.202 \times 10^8 \text{ rad}^2 \text{ sec}^{-2} \text{ at } 14089 \text{ G}$$

$$c^2 = (2/15) \gamma_N^2 H_0^2 (\sigma_{||} + \sigma_{\perp})^2 = 4.981 \times 10^8 \text{ rad}^2 \text{ sec}^{-2} \text{ at } 14089 \text{ G}$$

$$\text{where } \gamma_H = 2.675 \times 10^4 \text{ rad sec}^{-1} \text{ G}^{-1}$$

$$\gamma_N = -2.712 \times 10^3 \text{ rad sec}^{-1} \text{ G}^{-1}$$

$$h = 1.055 \times 10^{-27} \text{ G}^2 \text{ cm}^3 \text{ sec rad}^{-1}$$

$$r_{HN} = 1.02 \text{ \AA}$$

$$H_0^2 = 14089 \text{ G}$$

$$(\sigma_{||} + \sigma_{\perp}) = -160 \text{ ppm (Hiyama, et al., 1988).}$$

6.6 Experimental Peptide Dynamics

The T_1 , T_2 , and NOE relaxation times were measured using NMR experiments described in chapter 2.3. The T_1 and T_2 data fit well to a monoexponential function, displaying a monotonic increase in relaxation times that was correlated with the position of the residue from the core helix of Apa-Ala_N (residues 9-15) (Figure 6.10). The T_1 and T_2 relaxation times were applied to the evaluation of the model-free formalism, using the assumption that $\tau_f \ll \tau_R$, so that the contribution of τ_f to T_1 and T_2 through $J(\omega)$ was negligible. Using this assumption, the T_1/T_2 ratio was solely dependent upon τ_R , and was used to graphically determine τ_R (Figure 6.11). The rotational correlation time displayed a monotonic decrease which was also dependent upon the position of the residue. In particular, residues 16 through 18, the first three residues beyond the core helix of Apa-Ala_N, displayed nearly identical rotational motion. Residues 20 through 22, the

A.

B.

Figure 6.10. A. The T₁ data of Apa-Ala_N from two dimensional experiments. B. The T₁ and T₂ relaxation times of Apa-Ala_N. Two dimensional data are represented by hollow squares (T₁) and circles (T₂) connected by a line, while one dimensional data are represented by filled squares (T₁) and circles (T₂).

Figure 6.11. The relationship between T_1/T_2 and τ_R . The values of the ^{15}N nuclei of Apa-Ala N are indicated by the dotted lines. Residues 16 and 18 display identical T_1/T_2 ratios and identical τ_R .

last three C-terminal residues, also displayed similar rotational motion. However, residues 16 through 18 displayed different rotational motion relative to residues 20 through 22. The rotational correlation time of residue 12, which is most indicative of the rotational motion of the folded core of Apa-Ala_N, exceeds the τ_R anticipated for a 22 residue globular peptide. This agrees with similar analyses of proteins, in which τ_R calculated from NMR relaxation was found to be roughly 10-20% greater than τ_R determined using other techniques, and has been attributed to the rotation of the associated solvation shell (Powers, et al., 1992).

S_f^2 was determined from T_1 (Figure 6.12), since the quality of the two dimensional T_1 experimental data was superior to the T_2 and NOE data (Figure 6.10). Again, S_f^2 exhibited a monotonic decrease that was related to residue position, with a clear difference between residues 16 through 18 and 20 through 22. The S_f^2 parameters of sequential residues indicated that the motions of residues 12 through 18 are extremely well correlated, assuming that S_f^2 of residues 13 to 15 are similar to residue 12. The S_f^2 parameters of residues 20 through 22 indicated that these residues share less correlation with the rest of the peptide. Thus, this analysis provides further evidence that residues 9 through 18 constitute the stable helix of Apa-Ala_N, while residues 20 through 22 display increased mobility. However, residues 20 through 22 do not reach the values predicted for free rotation.

The increased mobility of the C-terminal residues agrees with the peptide structure derived from NOE connectivities. Both of these analyses conclude that the C-terminal residues conform to an equilibrium between a helix and a mobile random coil configuration relative to the more rigid helix preceding these residues. It is possible to apply formalisms which incorporate anisotropic motion or motion on a time scale τ_s to the analysis of the C-terminal residues of Apa-Ala_N. However, both of these formalisms require the definition of axes about which the ^{15}N - ^1H pair evolve according to the additional anisotropic or moderately fast motions. The random coil conformation of the C-terminal residues clearly lacks a definitive axis. Rather, it is most likely that the values of τ_R for each ^{15}N

Figure 6.12. Order parameters S_f^2 , the correlation of these order parameters, and the difference between observed and predicted NOE intensities of Apa-Ala_N. S_f^2 and $\frac{S_{f,i}^2}{S_{f,i-1}^2}$ may range from 0 to 1 dimensionless units, while predicted NOE intensities are as follows (in sec rad⁻¹): residue 12, 0.864; residue 16, 0.828; residue 17, 0.821; residue 18, 0.828; residue 20, 0.615; residue 21, 0.379; residue 22, 0.175.

accurately reflect a more continuous distribution of random coil motions of the C-terminal residues relative to the rigid helix, including additional anisotropic and moderately fast motions. Thus, analysis of the motion of the C-terminal residues with a model more complex than the model-free formalism may not appropriately interpret the the random coil motion of these residues.

References

- Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., Weng, J. (1987). "Protein Data Bank," In "Crystallographic Databases," Allen, F.H., Bergerhoff, G., & Sievers, R., eds., Data Commission of the International Union of Crystallography, Cambridge, 107-132.
- Aggarwal, A.K., Rogers, D.W., Droltar, M., Ptashne, M., & Harrison, S.C. (1988). Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science*, **242**, 899-907.
- Anderson, W.F., Ohlendorf, D.H., Takeda, Y., & Matthews, B.W. (1981). Structure of the Cro repressor from bacteriophage λ and its interaction with DNA. *Nature*, **290**, 754-758.
- Anderson, J.E., Ptashne, M., & Harrison, S.C. (1987). Structure of the repressor-operator complex of bacteriophage 434. *Nature*, **326**, 846-852.
- Aue, W.P., Bartholdi, E., & Ernst, R.R. (1976a). Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.*, **64**, 2229-2246.
- Aue, W.P., Karhan, J., & Ernst, R.R. (1976b). Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *J. Chem. Phys.*, **64**, 4226-4227.
- Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W., & Bax, A. (1992). Backbone dynamics of calmodulin-studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry*, **31**, 5269-5278.
- Bartholdi, E. & Ernst, R.R. (1973). Fourier spectroscopy and the causality principle. *J. Magn. Reson.*, **11**, 9-26.
- Barlow, D.J. & Thornton, J.M. (1988). Helix geometry in proteins. *J. Mol. Biol.*, **201**, 601-619.
- Bax, A., Griffey, R.H., & Hawkins, B.L. (1983). Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.*, **55**, 301-315.

- 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.
- Bax, A., Ikura, M., Kay, L.E., Torchia, D.A., & Tschudin, R. (1990). Comparison of different modes of two-dimensional reverse-correlation NMR experiments for the study of proteins. *J. Mag. Reson.*, **86**, 304-318.
- Bax, A., Ikura, M., Kay, L.E., & Zhu, G. (1991). Removal of F₁ baseline distortion and optimization of folding in multidimensional NMR spectra. *J. Magn. Reson.*, **91**, 174-178.
- Beamer, L.J. & Pabo, C.O. (1992). Refined 1.8 Å crystal structure of the lambda Repressor-operator complex. *J. Mol. Biol.*, **227**, 177-196.
- Benson, N., Sugiono, P., & Youderian, P. (1989). DNA sequence determinants of λ Repressor binding *in vivo*. *Genetics*, **118**, 21-29.
- Benson, N. & Youderain, P. (1989). Phage lambda Cro protein and cI Repressor use two different patterns of specific protein-DNA interactions to achieve sequence specificity *in vivo*. *Genetics*, **121**, 5-12.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112**, 535-542.
- Beyermann, M., Granitza, D., Bienert, M., Haussner, M., & Carpino, L. (1989). In "Peptides 1988. Proceedings of the 20th European Peptide Symposium," Jung, G. & Beyer, E., eds., Walter de Gruyter, Berlin, 28.
- Beyermann, M., Bienert, M., Niedrich, H., Chao, H., Carpino, L. (1990). In "Peptides: Chemistry and Biology. Proceedings of the 11th American Peptide Symposium," Rivier, J. & Marshall, G., eds., ESCOM, Leiden, 925.

- 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.
- Blumenthal, D.K., Takio, K., Edelman, A.M., Charbonneau, H., Titani, K., Walsh, K.A., & Krebs, E.G. (1985). Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase. *Proc. Nat. Acad. Sci. USA*, **82**, 3187-3191.
- Bodenhausen, G., Kogler, H., & Ernst, R.R. (1984). Selection of coherence-transfer pathways in NMR pulse experiments. *J. Magn. Reson.*, **58**, 370-388.
- Bontems, F., Roumstead, C., Gilquin, B., Menez, A., & Toma, F. (1991a). Refined structure of charybdotoxin: common motifs in scorpion toxins and insect defensins. *Science*, **254**, 1521-1523.
- Bontems, F., Roumstead, C., Boyot, P., Gilquin, B., Doljansky, Y., Menez, A., & Toma, F. (1991b). Three-dimensional structure of natural charybdotoxin in aqueous solution by ^1H -NMR. Charybdotoxin possesses a structural motif found in other scorpion toxins. *Eur. J. Biochemistry*, **196**, 19-28.
- Boyd, J., Hommel, U., & Krishnan, V.V. (1991). Influence of cross-correlation between dipolar and chemical shift anisotropy relaxation mechanisms upon the transverse relaxation rates of ^{15}N in macromolecules. *Chem. Phys. Lett.*, **187**, 317-324.
- Braunschweiler, L., Bodenhausen, G., & Ernst, R.R. (1983). Analysis of networks of coupled spins by multiple quantum N.M.R. *Mol. Phys.*, **48**, 535-560.
- Braunschweiler, L. & Ernst, R.R. (1983). Coherence transfer by isotropic mixing: application to proton experiments. *J. Magn. Reson.*, **58**, 370-388.
- Brennan, R.G., Takeda, Y., Kim, J., Anderson, W.F., & Matthews, B.W. (1986). Crystallization of a complex of Cro repressor with a 17-base pair operator. *J. Mol. Biol.*, **188**, 115-118.
- Brooks, C.L.III, Pettitt, B.M., & Karplus, M. (1985). Structural and energetic effects of

- truncating long ranged interactions in ionic and polar fluids. *J. Chem. Phys.*, **83**, 5897-5908.
- Brünger, A.T. (1990). Refinement of three-dimensional structures of proteins and nucleic acids. In "Topics in Molecular Biology," Goodfellow, J. ed., MacMillan, London, 137-178.
- Cavanaugh, J., & Keeler, J. (1988). Improvement of carbon-13 satellite spectra by double difference spectroscopy. *J. Magn. Reson.*, **77**, 356-361.
- Cavanaugh, J., Chazin, W.J., & Rance, M. (1990). The time dependence of coherence transfer in homonuclear isotropic mixing experiments. *J. Magn. Reson.*, **87**, 110-131.
- Chakrabarty, A. Schellman, J.A., & Baldwin, R.L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature*, **351**, 586-588.
- Chmurny, G.N. & Hoult, D.I. (1990). The ancient and honourable art of shimming. *Concepts Magn. Reson.*, **2**, 131-150.
- Chothia, C., Levitt, M., & Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol.*, **145**, 215-250.
- Chothia, C. (1984). Principles that determine the structure of proteins. *Ann. Rev. Biochem.*, **53**, 537-572.
- Chou, P.Y. & Fasman, G.D. (1974). Prediction of protein conformation. *Biochemistry*, **13**, 222-244.
- Clark-Lewis, I. & Kent, S. (1985). "Modern methods for the chemical synthesis of biologically active peptides." In "Synthetic peptides in Biology and Medicine," Alitalo, K., Partanen, P., & Vaheri, A., eds., Elsevier, Amsterdam. 1-28.
- Clore, G.M. & Gronenborn, A.M. (1989). Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Molec. Biol.*, **24**, 479-564.
- Clore, G.M., Driscoll, P.C., Wingfield, P.T., & Gronenborn, A.M. (1990). Analysis of

- backbone dynamics of Interlukin-1 β using two-dimensional inverse detected heteronuclear ^{15}N - ^1H NMR spectroscopy. *Biochemistry*, **29**, 7387-7401.
- Cohen, C. & Parry, D.A.D. (1990). α -helical coiled coils and bundles: how to design an α -helical protein. *Proteins*, **7**, 1-15.
- Conio, G., Patrone, E., & Brighetti, S. (1970). The effect of aliphatic alcohols on the helix-coil transition of poly-L-ornithine and poly-L-glutamic acid. *J. Biol. Chem.*, **245**, 3335-3340.
- Cox, J.A., Comte, M., Fitton, J.E., and DeGrado, W.F. (1985). The interaction of calmodulin with amphiphilic peptides. *J. Biol. Chem.*, **260**, 2527-2534.
- Creighton, T.E. (1983). "Proteins," Freeman, New York.
- Crippen, G.M. (1977). A Novel Approach to Calculation of Conformation: Distance Geometry. *J. Comput. Phys.*, **24**, 96-107.
- Crippen, G.M., Havel, T.F. (1988). "Distance Geometry and Molecular Conformation," Research Studies Press, U.K.
- Davis, D.G. & Bax, A. (1985) Assignment of complex ^1H NMR spectra via two-dimensional homonuclear Hartman-Hahn spectroscopy. *J. Am. Chem. Soc.*, **107**, 2820-2821.
- Dempsey, C.E. (1986). pH dependence of hydrogen exchange from backbone peptide amides in apamin. *Biochemistry*, **25**, 3904-3911.
- Derome, A.E. & Williamson, M.P. (1990). Rapid-pulsing artifacts in double-quantum filtered COSY. *J. Magn. Reson.*, **88**, 177-185.
- Distance Geometry. (1988). In "NMRchitect User Guide, Version 2.0," Biosym Technologies, San Diego, 2.3-2.18.
- Dress, A.W.M. & Havel, T.F. (1988). Shortest paths problems and molecular conformation. *Discrete Applied Math.*, **19**, 129-160.
- Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A., & Wright, P.E. (1988). Folding of immunogenic peptide fragments of proteins in water solution. 1.

- Sequence requirements for the formation of a reverse turn. *J. Mol. Biol.*, **201**, 161-200.
- Easthope, P. & Havel, T.F. (1988). Computational experience with an algorithm for tetrahedron inequality bound smoothing. *Bull. Math. Biol.*, **51**, 173-194.
- Eich, G., Bodenhausen, G., & Ernst, R.R. (1982). Exploring nuclear spin systems by relayed magnetization transfer. *J. Am. Chem. Soc.*, **104**, 3731-3732.
- Eisenberg, D., Weiss, R.M., & Terwilliger, T.C. (1982). The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature*, **299**, 371-374.
- Englander, S.W. & Kallenbach, N.R. (1984). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quart. Rev. Biophys.*, **16**, 521-655.
- Epand, R.F. & Scheraga, H.A. (1968). The influence of long-range interactions on the structure of myoglobin. *Biochemistry*, **7**, 2864-2872.
- Ermer, O. (1976). Calculation of molecular properties using force fields. Applications in organic chemistry. *Structure and Bonding*, **27**, 161-211.
- Ernst, R.R. (1966). Sensitivity enhancement in magnetic resonance. *Adv. Magn. Reson.*, **2**, 1-135.
- Ernst, R.R. & Anderson, W.A. (1966). Application of Fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instrum.*, **37**, 93-102.
- Ernst, R.R., Bodenhausen, G., & Wokaun, A. (1987). "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," Clarendon Press, Oxford.
- Estell, D.A., Graycar, T.P., & Wells, J.A. (1985). Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.*, **260**, 6518-6521.
- Evertsz, E.M., Thomas, G.A., & Peticolas, W.L. (1991). Raman spectroscopic studies of the DNA cro binding site conformation, free and bound to cro protein. *Biochemistry*, **30**, 1149-1155.
- Fairman, R., Shoemaker, K.R., York, E.J., Stewart, J.M., & Baldwin, R.L. (1989).

- Further studies of the helix dipole model: effects of a free $\alpha\text{-NH}_3^+$ or $\alpha\text{-COO}^-$ group on helix stability. *Proteins Struct. Funct. Gen.*, **5**, 1-7.
- Farrar, T.C. & Becker, E.D. (1971). "Pulse and Fourier Transform NMR," Academic Press, New York.
- Fersht, A. (1977). "Enzyme Structure and Mechanism," W.H. Freeman, New York.
- Fesik, S.W. & Zuiderweg, E.R.P. (1990). Heteronuclear three-dimensional NMR spectroscopy of isotopically labelled biological macromolecules. *Quart. Rev. Biophys.*, **23**, 97-131.
- Finkelstein, A.V., Badretdinov, A.Y., & Ptitsyn, O.B. (1991). Physical reasons for secondary structure stability: α -helices in short peptides. *Proteins*, **10**, 287-299.
- Fontecilla-Camps, J., Almassy, R.J., Suddath, F.L., Watt, D.D., & Bugg, C.E. (1980). Three-dimensional structure of a protein from scorpion venom: a new structural class of neurotoxins. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 6496-6500.
- Fontecilla-Camps, J.C., Habersetzer-Rochat, C., & Rochat, H. (1988). Orthorhombic crystals and three-dimensional structure of the potent toxin II from the scorpion *Androctonus australis* Hector. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 7443-7447.
- Fourquet, P., Bahraoui, E., Fontecilla-Camps, J.C., van Rietschoten, J., Rochat, H., & Granier, C. (1988). Immunochemistry of scorpion toxins. *Int. J. Peptide Protein Res.* **32**, 81-88.
- Ghadiri, M.R. & Choi, C. (1990). Secondary structure nucleation in peptides. Transition metal ion stabilized α -helices. *J. Am. Chem. Soc. USA*, **112**, 1630-1632.
- Glover, I., Haneef, I., Ptis, J., Wood, S., Moss, D., Tickle, I., & Blundell, T. (1983). Conformational flexibility in a small globular hormone: X-ray analysis of avian pancreatic polypeptide at 0.98 Å resolution. *Biopolymers*, **22**, 293-304.
- Goldman, M. (1988). "Quantum Description of High-Resolution NMR in Liquids," Clarendon Press, Oxford.
- Griesinger, C., Otting, G., Wuthrich, K., & Ernst, R.R. (1988). Clean TOCSY for ^1H

- spin system identification in macromolecules. *J. Am. Chem. Soc.*, **110**, 7870-7872.
- Griesinger, C., Sorensen, O.W., & Ernst, R.R. (1989). Three-dimensional Fourier spectroscopy. Application to high-resolution NMR. *J. Magn. Reson.*, **84**, 14-63.
- Griffey, R.H. & Redfield, A.G. (1987). Proton-detected heteronuclear edited and correlated nuclear magnetic resonance and nuclear Overhauser effect in solution. *Quart. Rev. Biophys.*, **19**, 51-82.
- Guiochon, G. (1983). Optimization in Liquid Chromatography. In "High-Performance Liquid Chromatography," Hearn, M.T.W., Regnier, F.E., & Wehr, C.T., eds., Academic Press, New York, 1-55.
- Harrison, S.C. & Aggarwal, A.K. (1990). DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem.*, **59**, 933-969.
- Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. *Nature*, **353**, 715-719.
- Havel, T.F., Kuntz, I.D., & Crippen, G.M. (1983). The theory and practice of distance geometry. *Bull. Math. Biol.*, **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 intermolecular ^1H - ^1H proximities in solution. *Bull. Math. Biol.*, **46**, 673-698.
- Hearn, M.T.W. (1983). High-Performance Liquid Chromatography of Peptides. In "High-Performance Liquid Chromatography," Hearn, M.T.W., Regnier, F.E., & Wehr, C.T., eds., Academic Press, New York, 87-155.
- Heinz, D.W., Hybertz, S.G., Peng, J.W., Priestle, J.P., Wagner, G., & Grütter, M.G. (1992). Changing the inhibitory specificity and function of the proteinase inhibitor Eglin c by site-directed mutagenesis: functional and structural investigation. *Biochemistry*, **31**, 8755-8766.

- Hider, R.C. & Ragnarsson, U. (1981). A comparative structural study of apamin and related bee venom peptides. *Biochim. Biophys. Acta.*, **667**, 197-208.
- Ho, S.P. & DeGrado, W.F. (1987). Design of a 4-helix bundle protein: synthesis of peptides which self-associate into a helical protein. *J. Am. Chem. Soc.*, **109**, 6751-6758.
- Hochschild, A., Douham, J.III, & Ptashne, M. (1986). How λ Repressor and λ Cro distinguish between O_R1 and O_R3. *Cell*, **47**, 925-933.
- Hochschild, A. & Ptashne, M. (1986). Homologous interactions of lambda repressor and lambda Cro with the lambda operator. *Cell*, **44**, 925-933.
- Hore, P.J. (1983). Solvent suppression in Fourier transform nuclear magnetic resonance. *J. Magn. Reson.*, **55**, 283-300.
- Hosur, R.V., Ernst, R.R., & Wüthrich, K. (1983). A simple two-dimensional measurement of the decoupler power during continuous homonuclear irradiation for the correction of Bloch-Siegert shifts. *J. Magn. Reson.*, **54**, 142-145.
- Hull, W.R. (1987). "Experimental Aspects of Two-Dimensional NMR." In "Two-Dimensional NMR Spectroscopy--Applications for Chemists and Biochemists," Croasmun, W.R. & Carlson, R.M.K., eds., VCH, New York, 67-258.
- Ikura, M., Kay, L.E., Tschudin, R., & Bax, A. (1990). Three-dimensional NOESY-HMQC spectroscopy of a ¹³C-labeled protein. *J. Magn. Reson.*, **86**, 204.
- Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B., & Bax, A. (1992). Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science*, **256**, 632-638.
- Introduction to the Theory of Chromatography. (1977). In "Basics of Liquid Chromatography," Third Edition. Spectra-Physics, Cnicago, 1-8.
- IUPAC-IUB Commision on Biochemical Nomenclature (1970). *J. Biol. Chem.*, **245**, 6489.
- Jacobson, B., Anderson, W.A., & Arnold, J.T. (1954). A proton magnetic resonance

- study of the hydration of deoxyribonucleic acid. *Nature*, **173**, 772-773.
- Jardetzky, O. & Roberts, G.C.K. (1981). "NMR in Molecular Biology," Academic Press, New York.
- Jeener, J. (1971). Ampère International Summer School, Basko Polje, Yugoslavia.
- Johnson, W.C., Jr., & Tinoco, I., Jr., (1972). Circular dichroism of polypeptide solutions in the vacuum ultraviolet. *J. Am. Chem. Soc.*, **94**, 4389-4390.
- Johnson, A.D., Meyer, B.J., & Ptashne, M. (1979). Interaction between DNA-bound repressors govern regulation by the λ phage Repressor. *Proc. Nat. Acad. Sci. USA*, **76**, 5061-5065.
- Johnson, A.D., Poteete, A.R., Laure, G., Sauer, R.T., Ackers, G.K., & Ptashne, M. (1981). λ Repressor and Cro—components of an efficient molecular switch. *Nature*, **294**, 217-223.
- Johnson, W. (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins: Struct. Funct. Genet.* **7**, 205-214.
- Jordan, S.R. & Pabo, C.O. (1988). Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions. *Science*, **242**, 893-899.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.
- Karplus, M. (1959). Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.*, **30**, 11-15.
- Kay, L.E., Torchia, D.A., & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972-8979.
- Kelsh, L.P., Ellena, J.F., & Cafiso, D.S. (1992). Determination of the molecular dynamics of Alamethicin using ^{13}C NMR: implications for the mechanism of gating of a voltage-dependent channel. *Biochemistry*, **31**, 5136-5144.

- Kennelly, P.J., Edelman, A.M., Blumenthal, D.K., & Krebs, E.G. (1987). Rabbit skeletal muscle myosin light chain kinase. *J. Biol. Chem.*, **262**, 11958-11963.
- Kent, S.B. (1980). New Aspects of Solid State Phase Peptide Synthesis. In "Biomedical Polymer, Polymeric Materials and Pharmaceuticals for Biomedical Use." New York, Academic Press, 213-237.
- Keepers, J.W. & James, T.L. (1984). A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.*, **57**, 404-426.
- Kim, P.S. & Baldwin, R.L. (1984). A helix stop signal in the isolated S-peptide of ribonuclease A. *Nature*, **307**, 329-334.
- Kim, Y. & Prestegard, J.H. (1989). A dynamic model for the structure of acyl carrier protein in solution. *Biochemistry*, **28**, 8792-8797.
- King, D.S., Fields, C.G., Fields, G.B. (1990). A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int. J. Peptide Protein Res.*, **36**, 255-266.
- Kirpichnikov, M.P., Hahn, K.D., Buck, F., Ruterjans, H., Chernov, B.K., Kurochkin, A.V., Skryabin, K.G., & Bayev, A.A. (1984). ¹H NMR study of the interaction of bacteriophage lambda Cro protein with the OR3 operator. Evidence for a change of the conformation of the OR3 operator upon binding. *Nucleic Acids Res.*, **12**, 3551-3561.
- Kitazumi, K., Shiba, T., Nishiki, K., Furukawa, Y., Takasaki, C., & Tasaka, K. (1990). Structure-function relationship in vasoconstrictor effects of sarafotoxins and endothelin-1. *FEBS Lett.*, **260**, 269-272.
- Klevit, R.E., Blumenthal, D.E., Wemmer, D.E., and Krebs, E.G. (1985). Interaction of calmodulin and a calmodulin-binding peptide from myosin light chain kinase: major spectral changes in both occur as the result of complex formation. *Biochemistry*, **24**, 8152-8157.

- Kugamaye, S., Kuroda, H., Nakajima, K., Watanabe, T.X., Kimura, T., Masaki, T., & Sakakibara, S. (1988). Synthesis and disulfide structure determination of porcine endothelin: an endothelium-derived vasoconstricting peptide. *Int. J. Peptide Protein Res.*, **32**, 519-526.
- 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. *Proc. Natl. Acad. Sci. Res. Comm.*, **95**, 1-6.
- Kumar, N.V., Wemmer, D.E., Kallenbach, N.R. (1988). Structure of α -melanocyte stimulating hormone (mast cell degranulating peptide) in solution. *Biophys. Chem.*, **31**, 113-119.
- Lane, A.N. (1988). The influence of spin diffusion and internal motions on NOE intensities in proteins. *J. Magn. Reson.*, **78**, 425-439.
- Levitt, M. & Greer, J. (1977). Automatic identification of secondary structure in globular proteins. *J. Mol. Biol.*, **114**, 181-193.
- Levitt, M. (1978). Conformational preferences of amino acids in globular proteins. *Biochemistry*, **17**, 4277-4282.
- Levitt, M.H. (1986). Composite Pulses. *Prog. in NMR Spectro.*, **18**, 61-122.
- Levitt, M.H. (1988). Density-operator theory of pulses and precession. In "Pulse Methods in 1D and 2D Liquid-Phase NMR," Academic Press, New York, 111-147.
- Lewis, M., Jeffrey, A., Wang, J., Ladner, R., Ptashne, M., & Pabo, C.O. (1983). Structure of the operator-binding domain of bacteriophage lambda repressor: implications for DNA recognition and gene regulation. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 435-440..
- Liff, M.I., Lyu, P.C., & Kallenbach, N.R. (1991). Analysis of asymmetry in the distribution of helical residues in peptides by ^1H nuclear magnetic resonance. *J. Am. Chem. Soc.*, **113**, 1000-1008.

- Chem. Soc.*, **113**, 1014-1019.
- Lifson, S. & Roig, A. (1961). On the theory of helix-coil transitions in polypeptides. *J. Chem. Phys.*, **34**, 1963-1974.
- Lindon, J.C. & Ferrige, A.G. (1980). Digitization and data processing in Fourier transform NMR. *Prog. N.M.R. Spectro.*, **14**, 27-66.
- Lipari, G. & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, **104**, 4546-4559.
- Live, D.H., Davis, D.G., Agosta, W.C., & Cowburn, D. (1984). Long range hydrogen bond mediated effects in peptides: ^{15}N NMR study of gramicidin S in water and organic solvents. *J. Am. Chem. Soc.*, **106**, 1939-1941.
- Lyu, P.C., Liff, M.I., Marky, L.A., & Kallenbach, N.R. (1990). Side chain contributions to the stability of alpha-helical structure in peptides. *Science*, **250**, 669-673.
- Lyu, P.C., Wang, P.C., Liff, M.I., & Kallenbach, N.R. (1991). Local effect of glycine substitution in a model helical peptide. *J. Am. Chem. Soc.*, **113**, 3568-3572.
- Mackay, D.H., Cross, A.J., & Hagler, A.T. (1989). The role of energy minimization in simulation strategies of biomolecular systems. In "Prediction of protein structure and the principles of protein conformation," G. Fasman, ed., Plenum Press, New York, 317-358.
- Macura, S., Ernst, R.R. (1980). Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys.*, **41**, 95-117.
- Macura, S., Huang, H., Suter, D., & Ernst, R.R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.*, **43**, 259-281.
- Madrid, M., Mace, J.E., Jardetzky, O. (1989). Consequences of magnetization transfer on the determination of solution structures of proteins. *J. Mag. Reson.*, **83**, 267-278.
- Mareci, T.H., Freeman, R. (1983). Mapping proton-proton coupling via double-quantum

- coherence. *J. Magn. Reson.*, **51**, 531-535.
- Marion, D. & Wuthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Comm.*, **113**, 967-974.
- Marion, D, Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., & Clore, G.M. (1989). Overcoming the overlap problem in the assignment of the ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry*, **28**, 6150-6156.
- Marion, D., Ikura, M., & Bax, A. (1989). Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. *J. Magn. Reson.*, **84**, 425-30.
- Marks, C.B., Vasser, M., Ng, P., HEnzel, W., & Anderson, S. (1986). Production of native, correctly folded bovine pancreatic trypsin inhibitor by *Escherichia coli*. *J. Biol. Chem.*, **261**, 7115-7118.
- Marqusee, S. & Baldwin, R.L. (1987). Helix stabilization by $\text{Glu}^- \cdots \text{Lys}^+$ salt bridges in short peptides of de novo design. *Proc. Nat. Acad. Sci. USA*, **84**, 8898-8902.
- Martin, G.E. & Zektzer, A.S. (1988). "Two-Dimensional NMR Methods for Establishing Molecular Connectivity." VCH Publishers, New York.
- Martins, J.C., Zhang, W., Tartar, A., Lazdunski, M., & Borremans, F.A.M. (1990). Solution conformation of leiurotoxin I (scyllatoxin) by ^1H nuclear magnetic resonance. *FEBS Lett.*, **260**, 249-253.
- Mayer, R., Lancelot, G., & Helene, C. (1983). Interactions of the 26-39 fragment of the cro protein from lambda bacteriophage with nucleic acids. *FEBS Lett.*, **153**, 339-344.
- McIntosh, L.P. & Dahlquist, F.W. (1990). Biosynthetic incorporation of ^{15}N and ^{13}C

for assignment and interpretation of nuclear magnetic resonance spectra of proteins.

Quart. Rev. of Biophys., **23**, 1-38.

Mehlkopf, A.F., Korbee, D., Tiggelman, T.A. (1984). Sources of t_1 noise in two-dimensional NMR. *J. Magn. Reson.*, **58**, 315-323.

Merkuta, G. & Stellwagen, E. (1990). Positional independence and additivity of amino acid replacements on helix stability in monomeric peptides. *Biochemistry*, **29**, 894-898.

Merkuta, G., Shalongo, W., & Stellwagen, E. (1991). A model peptide with enhanced helicity. *Biochemistry*, **30**, 4245-4248.

Merkuta, G. & Stellwagen, E. (1991). Effect of amino acid ion pairs on peptide helicity. *Biochemistry*, **30**, 1591-1594.

Messerle, B.A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989). Solvent suppression using spin lock in 2D and 3D NMR spectroscopy with H₂O solutions. *J. Magn. Reson.*, **85**, 608-613.

Metzler, W.J., Hare, D.R., & Pardi, A. (1989). Limited sampling of conformational space by the distance geometry algorithm: Implications for structures generated from NMR data. *Biochemistry*, **28**, 7045-7052.

Metzler, W.J. & Lu, P. (1989). Lambda cro repressor complex with OR3 operator DNA. ¹⁹F nuclear magnetic resonance observations. *J. Mol. Biol.*, **205**, 149-164.

Miick, S.M., Casteel, K.M., & Millhauser, G.L. (1993) Experimental molecular dynamics of an alanine-based helical peptide determined by spin label electron spin resonance. *Biochemistry*, **32**, 8014-8021.

Mitchell, A.R., Kent, S.B.H., Engelhard, M., & Merrifield, R.B. (1978). tert-butylloxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl-resin, an improved support for solid-phase synthesis. *J. Org. Chem.*, **43**, 2845-2852.

Mitchinson, C. & Wells, J.A. (1989). Protein engineering of disulfide bonds in subtilisin BPN'. *Biochemistry*, **28**, 4807-4815.

- Mondragón, A., Wolberger, C., & Harrison, S.C. (1989). Structure of phage 434 Cro protein at 2.35 Å resolution. *J. Molec. Biol.*, **205**, 179-188.
- Muller, N., Ernst, R.R., & Wüthrich, K. (1986). Multiple-quantum filtered two-dimensional correlated NMR spectroscopy of proteins. *J. Am. Chem. Soc.*, **108**, 6482-6492.
- Nagayama, K., Wüthrich, K., Bachmann, P., & Ernst, R.R. (1977). Two-dimensional J-resolved ^1H n.m.r. spectroscopy for studies of biological macromolecules. *Biochem. Biophys. Res. Comm.*, **78**, 99-105.
- Nelson, H.C.M., Hecht, M., & Sauer, R.T. (1983). Interaction of mutant lambda Repressor with operator DNA. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 441-449.
- Nelson, J.W. & Kallenbach, N.R. (1986). Stabilization of ribonuclease S peptide by trifluoroethanol. *Proteins*, **1**, 211-217.
- Nelson, H.C.M. & Sauer, R.T. (1986). Interaction of mutant λ Repressors with operator and non-operator DNA. *J. Molec. Biol.*, **192**, 27-38.
- Neuberger, M.S., Williams, G.T., & Fox, R.O. (1984). Recombinant antibodies possessing novel effector functions. *Nature*, **312**, 604-608.
- Noggle, J.H. & Schirmer, R.E. (1971). "The Nuclear Overhauser Effect," Academic Press, New York.
- Ohlendorf, D.H., Anderson, W.F., Fisher, R.G., Takeda, Y., & Matthews, B.W. (1982). The molecular basis of DNA-protein recognition inferred from the structure of the Cro repressor. *Nature*, **298**, 718-723.
- Ohlendorf, D.H., Anderson, W.F., Lewis, M., Pabo, C.O., & Mathews, B.W. (1983). Comparison of the structures of cro and lambda repressor proteins from bacteriophage lambda. *J. Molec. Biol.*, **169**, 757-769.
- O'Neil, K.T. & DeGrado, W.F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science*, **250**, 646-651.

- Otting, G., Widmer, H., Wagner, G., Wüthrich, K. (1986). Origin of t₁ and t₂ ridges in 2D NMR spectra and procedures for suppression. *J. Magn. Reson.*, **66**, 187-193.
- Pace, C.N. & Creighton, T.E. (1986). The disulfide folding pathway of ribonuclease T₁. *J. Mol. Biol.*, **188**, 477-486.
- Pakula, A.A., Young, V.B., & Sauer, R.T. (1986). Bacteriophage lambda cro mutations: effects on activity and intracellular degradation. *Proc. Nat. Acad. Sci. USA*, **83**, 8829-8833.
- Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M., & Baldwin, R.L. (1990). Relative helix-forming tendencies of nonpolar amino acids. *Nature*, **344**, 268-270.
- 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.
- Pauling, L., Corey, R.B., & Branson, H.R. (1951). The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Nat. Acad. Sci. USA*, **37**, 205-211.
- Patel, D.J., Shapiro, L., & Hare, D. (1987). Nuclear magnetic resonance and distance geometry studies of DNA structures in solution. *Ann. Rev. Biophys. Biophys. Chem.*, **16**, 423-454.
- Payne, M.E., Fong, Y.-L., Ono, T., Colbran, R.J., Kemp, B.E., Soderling, T.R., & Means, A.R. (1988). Calcium/calmodulin-dependant protein kinase II. *J. Biol. Chem.* **263**, 7190-7195.
- 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., 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.

- Pease, J., Storrs, R., & Wemmer, D. E. (1990). Folding and activity of hybrid sequence, disulfide-stabilized peptides. *Proc. Nat. Acad. Sci. U.S.A.*, **87**, 5643-5647.
- Peng, J.W., Thanabal, V., & Wagner, G. (1991). 2D heteronuclear NMR measurements of spin-lattice relaxation times in the rotating frame of X nuclei in heteronuclear HX spin systems. *J. Mag. Reson.*, **94**, 82-100.
- Perutz, M.F. & Fermi, G. (1988). Stereochemistry of Salt-Bridge Formation in α -helices and β -strands. *Proteins Struct. Funct. Gen.*, **4**, 294-295.
- Powers, R., Clore, G.M., Stahl, S.J., Wingfield, P.T., & Gronenborn, A. (1992). Analysis of the backbone dynamics of the Ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using ^{15}N relaxation measurements. *Biochemistry*, **31**, 9150-9157.
- Presta, L.G. & Rose, G.D. (1988). Helix signals in proteins. *Science*, **240**, 1632-1641.
- Rance, M., Sorenson, 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.
- Regan, L. & DeGrado, W.F. (1988). Characterization of a helical protein design from first principles. *Science*, **241**, 976-978.
- Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem.*, **34**, 167-339.
- Richardson, J.S. & Richardson, D.C. (1988). Amino acid preferences for specific locations at the ends of α -helices. *Science*, **240**, 1648-1652.
- Rico, M., Neito, J.L., Santoro, J., Bermejo, F.J., Herranz, J., & Gallego, E. (1983). Low-temperature ^1H -NMR evidence of the folding of isolated ribonuclease S peptide. *FEBS Lett.*, **162**, 314-319.
- Roberts, V.A., Dauber-Osguthorpe, P., Osguthorpe, D.J., Levin, E., & Hagler, A.T. (1986). A comparison of the binding of the ligand trimethoprim to bacterial and vertebrate dihydrofolate reductases. *Israel J. Chem.*, **27**, 198-210.

- Runser, D.J. (1981). "Maintaining and Troubleshooting HPLC Systems--A Users Guide." John Wiley & Sons, New York.
- Saenger, W. (1984). "Principles of Nucleic Acid Structure." Springer-Verlag, New York.
- Sakakibara, S. & Yasutsugu, S. (1965). A new method for releasing oxytocin from fully-protected nona-peptides using anhydrous hydrogen fluoride. *Bull. Chem. Soc. Japan*, **38**, 1412-1413.
- Sauer, R.T., Hehir, K., Stearman, R.S., Weiss, M.A., Tietler-Nillson, A., Suchanek, E.G., & Pabo, C.O. (1986). An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of λ Repressor. *Biochemistry*, **25**, 5992-5998.
- Saunders, M., Wishnia, A., & Kirkwood, J.G. (1957). The nuclear magnetic resonance spectrum of ribonuclease. *J. Am. Chem. Soc.*, **79**, 3289-3290.
- Schneider, D.M., Dellwo, M.J., & Wand, A.J. (1992). Fast internal main-chain dynamics of human ubiquitin. *Biochemistry*, **31**, 3645-3652.
- Scholtz, J.M. & Baldwin, R.L. (1992). The mechanism of alpha-helix formation by peptides. *Ann. Rev. Biophys. Biomol. Struct.*, **21**, 95-118.
- Schultz, G.E. & Schirmer, R.H. (1979). "Principles of Protein Structure." Springer-Verlag, New York.
- Seeholzer, S.H. & Wand, A.J. (1989). Structural characterization of the interactions between calmodulin and skeletal muscle myosin light chain kinase: effect of peptide (576-594)G binding on the Ca^{2+} -binding domains. *Biochemistry*, **28**, 4011-4020.
- Segrest, J.P., De Loof, H., Dohlman, J.G., Brouillette, C.G., & Anantharamaiah, G.M. (1990). Amphipathic helix motif: classes and properties. *Proteins Struct. Funct. Gen.*, **8**, 103-117.
- Sklenar, V. & Bax, A. (1987). Two-dimensional heteronuclear chemical shift correlation in proteins at natural abundance ^{15}N and ^{13}C levels. *J. Magn. Reson.*, **71**, 379-383.

- Sklenar, V., Torchia, D.A., & Bax, A. (1987). Measurement of carbon-13 longitudinal relaxation using ^1H detection. *J. Magn. Reson.*, **73**, 375-379.
- Slichter, C.P. (1980). "Principles of Magnetic Resonance," Springer-Verlag, New York.
- Snyder, L.R. (1983). Gradient Elution. In "High-Performance Liquid Chromatography," Hearn, M.T.W., Regnier, F.E., & Wehr, C.T., eds., Academic Press, New York, 208-315.
- Solid Phase Peptide Synthesis Chemistry (1987). In "Model 430A Peptide Synthesizer Manual," Applied Biosystems, Foster City, CA.
- Sönnichsen, F.D., Van Eyk, J.E., Hodges, R.S., & Sykes, B.D. (1992). Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry*, **31**, 8790-8798.
- Sørensen, O.W., Eich, G.W., Levitt, M.H., Bodenhausen, G., & Ernst, R.R. (1983) Product operator formalism for the description of NMR pulse experiments. *Prog. NMR Spectro.*, **16**, 163-192.
- States, D.J., Haberkorn, R.A., & Ruben, D.J. (1982). A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.*, **48**, 286-292.
- Stejskal, E.O. & Schaefer, J. (1974). Comparisons of quadrature and single-phase Fourier transform NMR. *J. Magn. Reson.*, **14**, 160-169.
- Stone, M.J., Fairbrother, W.J., Palmer, A.G.III, Reizer, J., Saier, M.H., Jr., & Wright, P.E. (1992). Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from ^{15}N NMR relaxation measurements. *Biochemistry*, **31**, 4394-4406.
- Stone, M.J., Chandrasekhar, C., Holmgren, A., Wright, P.E., & Dyson, H.J. (1993). Comparison of backbone and tryptophan side-chain dynamics of reduced and oxidized *Escherichia coli* thioredoxin using ^{15}N NMR relaxation measurements. *Biochemistry*, **32**, 426-435.

- Storrs, R.W., Truckses, D., & Wemmer, D.E. (1992). Helix propagation in trifluoroethanol in solutions. *Biopolymers*, **32**, 1695-1702.
- Sueki, M., Lee, S., Powers, S.P., Denton, J.B., Konishi, Y., & Scheraga, H.A. (1984). Helix-coil stability constants for the naturally occurring amino acids in water. 22. Histidine parameters from random coil poly[(hydroxybutyl)glutamine-co-L-histidine]. *Macromolecules*, **17**, 148-155.
- Szilágyi, L. & Jardetzky, O. (1989). α -proton chemical shifts and secondary structure in proteins. *J. Magn. Reson.*, **83**, 441-449.
- Tam, J.P., Heath, W.F., & Merrifield, R.B. (1983). S_N2 deprotection of synthetic peptides with a low concentration of HF in dimethyl sulfide: evidence and application in peptide synthesis. *J. Am. Chem. Soc.*, **105**, 6442-6455.
- Warwicker, J. (1986). Continuum dielectric modelling of the protein-solvent system, and calculation of the long-range electrostatic field of the enzyme phosphoglycerate mutase. *J. Theor. Biol.*, **121**, 199-210.
- Wagner, G. (1983). Two-dimensional relayed coherence transfer spectroscopy of a protein. *J. Magn. Resn.*, **55**, 151-156.
- Wagner, G. & Zuiderweg, E. (1983). Two dimensional double quantum ^1H NMR spectroscopy of proteins. *Biochem. Biophys. Res. Comm.*, **113**, 854-860.
- Watson, H.C. (1969). *Progr. Stereochem.* **4**, 299-333.
- Waugh, J.S. (1970) Sensitivity in Fourier transform NMR spectroscopy of slowly relaxing systems. *J. Molec. Spectro.*, **35**, 298-305.
- Wemmer, D.E. & Reid, B.R. (1985). High Resolution NMR Studies of Nucleic Acids and Proteins. *Ann. Rev. Phys. Chem.*, **36**, 105-137.
- Wetzel, R. (1986). What is protein engineering? *Protein Engin.*, **1**, 3-6.
- Wharton, R.P., Brown, E.L., & Ptashne, M. (1984). Substituting an α -helix switches the sequence-specific DNA interactions of a repressor. *Cell*, **38**, 361-369.
- Wharton, R.P. & Ptashne, M. (1985). Changing the binding specificity of a repressor by

- redesigning an α -helix. *Nature*, **316**, 601-605.
- Wider, G., Macura, S., Kumar, A., Ernst, R.R., & Wüthrich, K. (1984). Homonuclear two-dimensional ^1H NMR of proteins. Experimental procedures. *J. Mag. Reson.*, **56**, 207-234.
- Williamson, M.P. (1990). Secondary-structure dependent chemical shifts in proteins. *Biopolymers*, **29**, 1423-1431.
- Woessner, D.E., (1962). Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. *J. Chem. Phys.*, **37**, 647-654.
- Wolberger, C., Dong, Y.C., Ptashne, M., & Harrison, S.C. (1988). Structure of a phage 434 Cro/DNA complex. *Nature*, **335**, 789-795.
- 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., 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," John Wiley & Sons, New York.
- Ying, P.F. (1990), Scaling NOESY cross peaks involving methyl protons. *J. Magn. Reson.*, **90**, 382-383.
- Zhang, R. Snyder, G.H. (1988). Kinetics of disulfide exchange reactions of monomer and dimer loops of cysteine-valine-cysteine peptides. *Biochemistry*, **27**, 3785-3794.
- Zhang, R. Snyder, G.H. (1989). Dependence of formation of small disulfide loops in two-

cysteine peptides on the number and types of intervening amino acids. *J. Biol. Chem.*, **264**, 18472-18479.

Zimm, B.H. & Bragg, J.K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.*, **31**, 526-535.

Zuiderweg, E.R.P., McIntosh, L.P., Dahlquist, F.W., & Fesik, S.W. (1990). Three-dimensional ^{13}C -resolved proton NOE spectroscopy of uniformly ^{13}C -labeled proteins for the NMR assignment and structure determination of larger molecules. *J. Magn. Reson.*, **86**, 210-216.

Appendix 1. Peptide Synthesis

Non-Amino Acid Synthesis Reagents

<u>Chemistry</u>	<u>Reactant</u>	<u>Chemical Moiety</u>
tBoc	Resin	polystyrene, 1% divinylbenzene crosslink, 200-400 mesh
tBoc	linker (L)	phenylacetimidomethyl (PAM)
tBoc	amino protecting group (PN)	tert-Butyloxycarbonyl (tBoc)
tBoc	amino deprotector (D)	trifluoroacetic acid (TFA)
tBoc	amino acid activator (A)	dicyclohexylcarbodiimide (DCC)
tBoc	Cap	acetic anhydride
tBoc	cleavage & deprotection (C)	hydrogen fluoride (HF)
Fmoc	Resin	polystyrene, 1% divinylbenzene crosslink, 200-400 mesh
Fmoc	linker (L), amide C-terminus	RINK AM
Fmoc	linker (L), free acid	WANG
Fmoc	amino protecting group (PN)	9-Fluorenylmethoxycarbonyl (Fmoc)
Fmoc	amino deprotector (D)	piperidine
Fmoc (HOBT)	amino acid activator (A)	N-hydroxybenzotriazole (HOBT)+ dicyclohexylcarbodiimide (DCC)
Fmoc (HBTU)	amino acid activator (A)	O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU) + dicyclohexylcarbodiimide (DCC)
Fmoc	Cap	acetic anhydride
Fmoc	cleavage & deprotection (C)	trifluoroacetic acid (TFA)

Amino Acid Synthesis Reagents

<u>Amino Acid (R_n)</u>	<u>Chemistry</u>	<u>Side Chain Protecting Group (PR_n)</u>
Alanine	All	None
Arginine	tBoc	Tosyl(p-Toluenesulfonyl) (Tos)
Arginine	Fmoc	2,2,5,7,8-pentamethylchroan-6-sulpholyl (PMC)
Asparagine	tBoc	None
Asparagine	Fmoc	Trityl (Trt)
Cysteine	tBoc	Paramethylbenzyl (4CH ₃ -Bzl)
Cysteine	Fmoc	Trityl (Trt)
Glutamine	tBoc	None
Glutamine	Fmoc	Trityl (Trt)
Glutamic acid	tBoc	Benzyl ester (OBzl)
Glutamic acid	Fmoc	tert-Butyl ester (OtBu)
Glycine	All	None
Histidine	tBoc	Dinitrophenyl (DNP)
Histidine	Fmoc	Trityl (Trt)
Isoleucine	All	None
Leucine	All	None
Lysine	tBoc	2-Chlorobenzyloxycarbonyl (Cl-Z)
Lysine	Fmoc	tert-Butyloxycarbonyl (tBoc)
Methionine	tBoc	none
Methionine	Fmoc	Sulfoxide (Sox)
Ornithine	Fmoc	tert-Butyloxycarbonyl (tBoc)
Phenylalanine	All	None
Proline	All	None

Amino Acid Synthesis Reagents (continued):

Amino Acid (R_n)	Chemistry	Side Chain Protecting Group (PR_n)
Serine	tBoc	Benzyl (Bzl)
Serine	Fmoc	tert-Butyl (tBu)
Threonine	tBoc	Benzyl (Bzl)
Threonine	Fmoc	tert-Butyl (tBu)
Tryptophan	tBoc	acetyl (CHO)
Tryptophan	Fmoc	none
Tyrosine	tBoc	ortho-Bromobenzyloxycarbonyl (Br-Z)
Tyrosine	Fmoc	tert-Butyl (tBu)
Valine	All	None

tBoc Solid Phase Peptide Synthesis Reaction Conditions

Reaction	Solvent	Time	Reactants
1. Amino Deprotection	I. DCM	1.33 min	0.25 mmol protected peptide chain 33% TFA
	II. DCM	18.5 min	0.25 mmol protected peptide chain 50% TFA
	III. DMF	2 min	0.25 mmol protected peptide chain 10% DIEA
2.A. Amino Acid Activation for Ala, Asp, Cys, Glu, Gly, Ile, Lys, Met, Phe, Pro, Ser, Thr, Tyr, and Val	5.0 ml DCM	8 min	2.0 mmol amino acid 1.0 mmol DCC
2.B. Amino Acid Activation for Leu, Trp	4.5 ml DCM 0.3 ml DMF	8 min	2.0 mmol amino acid 1.0 mmol DCC
2.C. Amino Acid Activation for Asn, Gln	4.0 ml DMF 4.3 ml DCM	33 min	2.0 mmol amino acid 2.0 mmol HOBT 2.0 mmol DCC
2.D. Amino Acid Activation for Arg	4.0 ml DMF 4.5 ml DCM	33 min	2.0 mmol amino acid 2.0 mmol HOBT 2.0 mmol DCC
3. Coupling	8 ml DMF	18.5 min	0.25 mmol amino-deprotected peptide chain 1.0 mmol activated amino acid
3A. Double Coupling	DCM	18.5 min	0.25 mmol amino-deprotected peptide chain 1.0 mmol activated amino acid
3B. Capping	DMF	6 min	0.25 mmol synthesized peptide chain 10% acetic anhydride
4. Cleavage & Deprotection	I. DMS	2 hours	100 mg peptide/resin 2.5 ml HF 6.5 ml dimethyl sulfide
	II. HF	30 min	1.0 ml p-cresol 9.0 ml HF 1.0 ml p-cresol

Fmoc (HOBT) Solid Phase Peptide Synthesis Reaction Conditions

Reaction	Solvent	Time	Reactants
1. Amino Deprotection	NMP	20 min	20% piperidine
2.A. Amino Acid Activation for Asn, His, Lys, Orn and Trp	0.4 ml DCM 3.2 ml NMP	51 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
2.B. Amino Acid Activation for Arg, Gln	0.8 ml DCM 3.2 ml NMP	51 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
2.C. Amino Acid Activation for Ala, Asp, Cys, Glu, Gly, Ile, Leu, Met, Phe, Pro, Ser, Thr, and Tyr	0.4 ml DCM 2.0 ml NMP	36.5 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
2.D. Amino Acid Activation for Val	0.4 ml DCM	61.5 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
3. Coupling	NMP	120 min	0.25 mmol amino-deprotected peptide chain 1.0 mmol activated amino acid
3A. Double Coupling	NMP	120 min	0.25 mmol amino-deprotected peptide chain (treated once with 1.0 mmol activated amino acid) 1.0 mmol activated amino acid
3B. Capping	NMP	8 min	0.25 mmol synthesized peptide chain 10% acetic anhydride
4. Cleavage & Deprotection	TFA	6-8 hrs	125 mg peptide/resin 1.66 ml TFA 0.088ml phenol 0.100 water 0.100 ml thioanisole 0.050 ml ethanedithiol

Fmoc (HBTU) Solid Phase Peptide Synthesis Reaction Conditions

Reaction	Solvent	Time	Reactants
1. Amino Deprotection	NMP	20 min	0.1 mmol protected peptide chain 20% piperidine
2.A. Amino Acid Activation for Asn, His, Lys, Orn and Trp	0.4 ml DCM 3.2 ml NMP	51 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
2.B. Amino Acid Activation for Arg, Gln	0.8 ml DCM 3.2 ml NMP	51 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC

Fmoc (HBTU) Solid Phase Peptide Synthesis Reaction Conditions (continued)

<u>Reaction</u>	<u>Solvent</u>	<u>Time</u>	<u>Reactants</u>
2.C. Amino Acid Activation for Ala, Asp, Cys, Glu, Gly, Ile, Leu, Met, Phe, Pro, Ser, Thr, and Tyr	0.4 ml DCM 2.0 ml NMP	36.5 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
2.D. Amino Acid Activation for Val	0.4 ml DCM	61.5 min	1.0 mmol amino acid 1.0 mmol HOBT 1.0 mmol DCC
3. Coupling (repeated a second time in Double Coupling)	NMP	120 min	0.05 mmol amino-deprotected peptide chain 0.5 or 1.0 mmol activated amino acid
3B. Capping	NMP	8 min	0.25 mmol synthesized peptide chain 10% acetic anhydride
4. Cleavage & Deprotection	TFA	6-8 hrs	125 mg peptide/resin 1.66 ml TFA 0.088 ml phenol 0.100 ml water 0.100 ml thioanisole 0.100 ml ethanedithiol

Coupling protocols: (SC = Single Coupling; DC = Double coupling)

<u>Apa-M5 protocol</u>	<u>Apa-Cro protocol</u>	<u>Max protocol</u>	<u>Apa-AlaN protocol</u>
tBoc Fmoc (HOBT)	Fmoc (HOBT)	Fmoc (HOBT)	
G on resin	Wang resin	Wang resin	Rink AM resin
F DC, capped	K SC	Y SC	A SC, 0.5 mmol
R DC, capped	K SC	K SC	A SC, 0.5 mmol
N DC, capped	R SC	V DC	A SC, 0.5 mmol
A DC, capped	G SC	C DC	K SC, 1.0 mmol
A DC, capped	A SC	E SC	A SC, 0.5 mmol
S DC, capped	H SC	C DC	A SC, 0.5 mmol
V DC, capped	O SC	K SC	A SC, 0.5 mmol
A DC, capped	C DC	D SC	C SC, 1.0 mmol
I DC, capped	K SC	G SC	K SC, 1.0 mmol
F DC, capped	N SC	I DC	Y SC, 1.0 mmol
N DC, capped	O SC	A SC	A SC, 0.5 mmol
C DC, capped	C DC	K SC	C SC, 1.0 mmol
K DC, capped	S SC	G SC	E SC, 1.0 mmol
W DC, capped	Q DC	L DC	A SC, 1.0 mmol
R DC, capped	T DC	L DC	T SC, 1.0 mmol
C DC, capped	E DC	G SC	E SC, 1.0 mmol
K DC, capped	P SC	L DC	P SC, 1.0 mmol
A DC, capped	A SC	S SC	A SC, 1.0 mmol
T DC, capped	K SC	R SC	K SC, 1.0 mmol
E DC, capped	C DC	C DC	C SC, 1.0 mmol
P DC, capped	N DC	S SC	N SC, 1.0 mmol
A DC, capped	C DC	L DC	C SC, 1.0 mmol
K DC, capped		Q DC	

Coupling Protocols (continued):Apa-M5 protocol

C DC, capped
N DC, capped
C DC, capped

Max protocol

C DC
M SC
R SC

Appendix 2. NMR Pulse Programs

Pulse programs and phase cycles of NMR spectroscopy experiments. Pulses not filled with a pattern are scaled to represent 45° , 90° , and 180° pulses. Pulses filled with a pattern represent:

- = low-power irradiation at water chemical shift to suppress water signal for 1.5 sec
- = MLEV spin-lock sequence
- = scrambling pulse
- = ^{15}N GARP composite pulse decoupling

A. DOFCOSY

$\tau_m = 5 \mu\text{sec}$ $\phi_1 = 4(y), 4(-x)$
 $\phi_2 = 4(x), 4(y), 4(-x), 4(-y)$
 $\phi_3 = (y), 2[(-x), (-y), (x)], (y), (-y), 2[(x), (y), (-x)], (-y)$
 $\phi_R = 2[(x), (-y), (-x), 2(y), (x), (-y), (-x)]$

B. RELAY

$\tau_m = 60 \text{ msec}$ $\phi_1 = 8(x), 8(y), 8(-x), 8(-y)$
 $\phi_2 = 2(x), 2[2(y), 2(-x), 2(-y)], 2(x), 2(-x), 2[2(-y), 2(x), 2(y)], 2(-x)$
 $\phi_3 = 2[(x), (-x), (y), (-y)], 2[(y), (-y), (x), (-x)]$
 $\phi_R = 2[2(x), 2(-x)], 2[2(y), 2(-y)]$

C. TOCSY

$\tau_m = 100 \text{ msec}$ $\phi_1 = (x), 2(-x), (x), (y), 2(-y), (y)$

D. 2D-COSY

$$\tau_m = 1/4J_{HH}$$

$$= 42 \text{ msec}$$

$$\phi_1 = 2[(x), (-x)], 2[(-x), (x)]$$

$$\phi_2 = (x), (y), (-x), (-y)$$

$$\phi_R = 2[(x), (-x)], 2[(y), (-y)]$$

E. HMOCJR

$$\Delta = 1/2J_{NH} \quad \phi_1 = 2(x), 2(y), 2(-x), 2(-y)$$

$$= 5.3 \text{ msec} \quad \phi_2 = (x), (-x)$$

$$\delta = 140 \text{ } \mu\text{sec} \quad \phi_3 = 2(-x), 2(-y), 2(x), 2(y)$$

$$\phi_R = (x), 2(-x), 2(x), 2(-x), (x), (-x), 2(x), 2(-x), 2(x), (-x)$$

F. NOESY

$$\tau_m = 250-350 \text{ msec}$$

$$\phi_1 = (x), (-x)$$

$$\phi_2 = 8(x), 8(-x)$$

$$\phi_3 = 2(x), 2(-x), 2(y), 2(-y)$$

$$\phi_R = (x), 2(-x), (x), (y), 2(-y), (y), (-x), 2(x), (-x), (-y), 2(y), (-y)$$

G. NOESYJR $\tau_m = 250-350 \text{ msec}$ $\phi_1 = (x), (-x)$ $\delta = 140 \mu\text{sec}$ $\phi_2 = 8(x), 8(-x)$ $\phi_3 = 2(x), 2(-x), 2(y), 2(-y), 2(-x), 2(x), 2(-y), 2(y)$ $\phi_4 = 2(-x), 2(x), 2(-y), 2(y), 2(x), 2(-x), 2(y), 2(-y)$ $\phi_R = (x), 2(-x), (x), (y), 2(-y), (y), (-x), 2(x), (-x), (-y), 2(y), (-y)$ H. NOESY-HMOC $\Delta = 1/2J_{NH}$ $\phi_1 = 4(x), 4(-x)$ $= 5.3 \text{ msec}$ $\phi_2 = (x), (-x)$ $\phi_3 = 2(x), 2(-x)$ $\phi_R = (x), 2(-x), (x), (-x), 2(x), (-x)$ I. ¹⁵N-¹H NOE $\Delta = 1/4J_{NH}$ $= 2.6 \text{ msec}$ $\delta = 10 \text{ msec}$ $d = 1 \text{ msec}$ $N = 0 \text{ or } 350$ $\phi_3 = 2(x), 2(-x)$ $\phi_R = (x), 2(-x), (x), (-x), 2(x), 2(-x), 2(x), (-x), (x), 2(-x), (x)$ $\phi_1 = (x), (-x)$ $\phi_2 = 8(x), 8(y), 8(-x), 8(-y)$ $\phi_4 = 4(y), 4(-y)$ $^{15}\text{N:}$

L_15N-T1

$\Delta = 1/4JNH$ $\phi_1 = 4(x), 4(-x)$
 $= 2.6 \text{ msec}$ $\phi_2 = 8(x), 8(y), 8(-x), 8(-y)$

$\delta = 7.5 \text{ msec}$ $\phi_3 = (y), (-y)$

$\boxed{1} = 1 \text{ msec}$ $\phi_4 = 2(x), 2(-x)$

$\boxed{2} = 10 \text{ msec}$ $\phi_R = (x), 2(-x), (x), (-x), 2(x), (-x)$

$N = \text{Int}[T1(\text{msec})/15]$

L_15N-T2

$\Delta = 1/4JNH$
 $= 2.6 \text{ msec}$

$d = 400 \mu\text{sec}$

$\boxed{1} = 1 \text{ msec}$

$\boxed{2} = 10 \text{ msec}$

$N = \text{Int}[T2(\text{msec})/1.7]$

$\phi_1 = 4(y), 4(-y)$

$\phi_2 = 2(x), 2(-x)$

$\phi_3 = 16(x), 16(-x)$

$\phi_4 = 8(y), 8(-y)$

$\phi_5 = 4(x), 4(y)$

$\phi_R = 2(x), 4(-x), 2(x), 2(-x), 4(x), 2(-x)$

Appendix 3. Chemical Shifts

All proton chemical shifts are referenced to TSP. All nitrogen chemical shifts are indirectly referenced to TSP (Live, et al., 1984).

Apa-M5 proton chemical shifts at 0°C and 10 mM.

RESIDUE	NH	α H	β H	β' H	γ H	γ' H	OTHER
1	C	---	4.38	3.85	3.79	---	---
2	N	9.34	4.88	3.02	2.76	---	amine=7.81,7.19
3	C	9.26	4.71	2.72	3.32	---	---
4	K	8.09	4.22	1.82	---	1.49 1.40	δ =1.64 ϵ =2.94 amine=7.60
5	A	7.35	4.51	1.15	---	---	---
6	P	---	4.80	2.03	1.63	2.02 1.96	δ =3.40,3.49
7	E	9.12	4.39	2.04	2.27	2.54 2.52	---
8	T	7.62	4.62	4.62	---	1.23	---
9	A	9.17	4.02	1.41	---	---	---
10	K	8.69	4.13	1.77	---	---	---
11	C	7.87	4.72	2.76	3.17	---	---
12	R	8.36	3.79	2.04	---	1.86	δ =3.19,3.03
13	W	7.57	4.29	3.47	3.36	2=7.25 4=7.26	5,6=6.99 7=7.33 NH=10.19
14	K	8.24	3.74	2.34	---	1.62 1.42	δ =1.86 ϵ =3.04 amine=6.905
15	C	8.61	4.31	2.97	2.91	---	---
16	N	7.96	4.49	2.68	---	---	---
17	F	7.94	4.15	2.91	2.57	---	2,6=7.03 3,5=7.19
18	I	8.04	3.77	1.82	---	1.65 1.15	γ =0.86, δ =0.85
19	A	7.94	4.17	1.38	---	---	---
20	V	8.28	3.90	1.98	---	0.91 0.85	---
21	S	8.17	4.23	3.78	3.69	---	---
22	A	8.22	4.16	1.39	---	---	---
23	A	8.05	4.61	1.35	---	---	---
24	N	8.12	4.16	2.75	2.69	---	---
25	R	8.05	4.13	1.57	---	1.43 1.29	δ =3.01
26	F	8.30	4.66	3.21	2.92	---	2,6=7.23 3,5=7.29
27	G	8.14	3.90	---	---	---	α' =3.90

Apa-Cro chemical shifts at 15°C and 12 mM.

RESIDUE	NH	α H	β H	β' H	γ H	γ' H	OTHER
1	C	---	4.51	2.91	2.96	---	---
2	N	9.13	4.37	3.15	2.72	---	---
3	C	9.20	4.69	3.32	2.77	---	---
4	K	8.13	4.24	1.85	1.81	1.41	δ =1.52 ϵ =2.89
5	A	7.37	4.58	1.19	---	---	---
6	P	---	4.51	2.11	2.03	2.11 2.03	δ =3.57,3.46
7	E	8.77	4.41	2.28	2.11	2.55 2.60	---
8	T	7.75	4.60	4.60	---	1.28	---
9	Q	9.27	3.93	2.08	2.17	2.43	---
10	S	8.76	4.23	3.85	3.95	---	---
11	C	7.80	4.69	2.81	3.19	---	---
12	O	8.59	3.75	1.92	1.90	1.75 1.59	δ =2.91

Apa-Cro chemical shifts at 15°C and 12 mM (continued).

RESIDUE	NH	α H	β H	β' H	γ H	γ' H	OTHER
13 N	8.30	4.43	2.93	2.91	---	---	---
14 K	8.14	4.12	2.00	2.03	1.60	1.49	$\delta=1.91, 1.75$ $\epsilon=3.93$
15 C	8.29	4.42	3.07	2.94	---	---	---
16 O	7.73	4.24	1.88	1.79	1.59	---	$\delta=2.81$
17 H	8.23	4.68	3.22	3.35	---	---	$4=7.13$ $2=7.77$
18 A	8.25	4.35	1.44	---	---	---	---
19 G	8.43	3.91	---	---	---	---	$a'=4.03$
20 R	8.24	4.35	1.83	1.76	1.64	---	$\delta=3.22$ $\epsilon=7.21$
21 K	8.51	4.35	1.84	1.78	1.37	---	$\delta=1.68$ $\epsilon=2.98$
22 K	8.50	4.24	1.84	1.78	1.37	---	$\delta=1.68$ $\epsilon=2.98$

Max proton chemical shifts at 15°C and 8 mM.

RESIDUE	NH	α H	β H	β' H	γ H	γ' H	OTHER
1 R	---	4.07	1.93	---	1.64	---	$\delta=3.18$ $\epsilon=7.22$
2 M	9.19	4.63	2.43	2.27	3.08	2.99	$\epsilon=2.14$
3 C	9.05	4.70	3.05	2.83	---	---	---
4 Q	8.98	4.05	2.14	---	2.47	---	$\delta=1.52$ $\epsilon=2.89$
5 L	7.98	4.23	1.77	---	1.64	---	$\delta=0.93$, $\delta'=0.88$
6 S	8.29	4.35	3.93	---	---	---	---
7 C	8.08	4.60	2.62	2.48	---	---	---
8 R	8.30	4.38	2.08	2.02	1.77	---	$\delta=3.25$ $\epsilon=7.28$
9 S	8.15	4.40	4.00	---	---	---	---
10 L	7.13	4.42	---	---	1.71	---	$\delta=0.95$ $\delta'=0.86$
11 G	8.31	4.11	---	---	---	---	$\alpha'=3.98$
12 L	7.51	4.74	---	---	1.38	---	$\delta=0.85$
13 L	8.76	4.44	1.48	---	1.48	---	$\delta=1.06$ $\delta'=0.78$
14 G	8.60	4.99	---	---	---	---	$\alpha'=4.40$
15 K	9.03	4.45	1.76	---	1.35	1.26	$\delta=1.63$ $\epsilon=2.90$ amine=7.65
16 A	8.62	4.92	1.38	---	---	---	---
17 I	8.38	4.29	1.76	---	1.40	1.09	$\gamma=0.80$ $\delta=0.76$
18 G	9.07	3.98	---	---	---	---	$\alpha'=3.73$
19 D	8.87	4.76	3.02	---	---	---	---
20 K	8.15	4.61	1.92	1.90	1.44	1.35	$\delta=1.67$ $\epsilon=2.98$ amine=7.59
21 C	8.71	5.19	2.85	2.90	---	---	---
22 E	9.03	4.72	2.05	1.92	2.38	2.33	---
23 C	8.81	5.56	2.94	2.73	---	---	---
24 V	9.42	4.64	2.19	---	0.82	0.76	---
25 K	8.44	4.33	1.79	1.69	1.25	---	$\delta=1.63$ $\epsilon=2.90$ amino=7.63
26 Y	8.35	4.38	3.03	2.93	---	---	$2,6=6.79$ $3,5=7.21$

Apa-AlaN proton chemical shifts at 20°C and 10 mM.

RESIDUE	NH	α H	β H	β' H	γ H	γ' H	OTHER
1 C	---	4.25	3.32	2.75	---	---	---
2 N	9.96	4.49	2.72	2.61	---	---	amine= 7.79, 7.21
3 C	9.06	4.72	3.30	2.76	---	---	---
4 K	8.04	4.22	1.85	---	1.43	1.39	δ = 1.53 ϵ =2.90 amine=7.62
5 A	7.21	4.53	1.17	---	---	---	---
6 P	---	4.80	2.08	2.02	2.01	1.88	δ =3.57, 3.43
7 E	9.83	4.62	2.75	---	3.14	---	---
8 T	7.45	4.67	4.62	---	1.26	---	---
9 A	9.12	4.17	1.46	---	---	---	---
10 E	8.68	4.22	2.18	2.00	2.53	---	---
11 C	7.98	4.68	3.26	2.80	---	---	---
12 A	8.45	3.86	1.50	---	---	---	---
13 Y	8.25	4.31	3.20	---	---	---	2,6=7.13, 3,5=6.83
14 K	8.34	3.96	2.07	1.99	1.73	---	δ =1.87, 1.65 ϵ =3.27 amine=7.63
15 C	8.74	4.31	3.03	---	---	---	---
16 A	7.92	4.18	1.48	---	---	---	---
17 A	8.04	4.12	1.32	---	---	---	---
18 A	7.81	4.21	1.49	---	---	---	---
19 K	7.82	4.15	2.16	1.88	1.54	---	δ =1.77 ϵ =3.21 amine=7.65
20 A	7.97	4.19	1.44	---	---	---	---
21 A	7.96	4.21	1.44	---	---	---	---
22 A	7.88	4.22	1.44	---	---	---	---
23 NH ₂	7.80, 7.10						

Apa-AlaN nitrogen chemical shifts at 20°C and 10 mM.

RESIDUE	N
12 A	123.84
16 A	122.36
17 A	121.86
18 A	123.71
20 A	122.72
21 A	121.34
22 A	120.96

Appendix 4. Distance Geometry Constraints

Hydrogen Bond Constraints of Apa-M5

Atom 1	Atom 2	Lower	Upper
ASN_2:O	ALA_5:HN	1.800	2.000
THR_8:O	ARG_12:HN	1.800	2.000
ALA_9:O	TRP_13:HN	1.800	2.000
CYS_11:O	CYS_15:HN	1.800	2.000
ARG_12:O	ASN_16:HN	1.800	2.000
TRP_13:O	PHE_17:HN	1.800	2.000
LYS_14:O	ILE_18:HN	1.800	2.000
CYS_15:O	ALA_19:HN	1.800	2.000
ASN_16:O	VAL_20:HN	1.800	2.000

Disulfide Bridge Constraints of Apa-M5

Atom 1	Atom 2	Lower	Upper
CYS_1:CB	CYS_11:CB	3.720	3.990
CYS_3:CB	CYS_15:CB	3.720	3.990

NOE Constraints of Apa-M5

Atom 1	Atom 2	Lower	Upper
ASN_2:OD1	LYS_4:HN	1.800	2.000
ASN_2:HA	CYS_3:HN	1.900	2.500
PRO_6:HA	GLU_7:HN	1.900	3.000
ARG_12:HA	TRP_13:HN	2.500	3.500
TRP_13:HA	LYS_14:HN	2.500	3.500
LYS_14:HA	CYS_15:HN	2.500	3.500
CYS_15:HA	ASN_16:HN	2.500	3.500
ASN_16:HA	PHE_17:HN	2.500	3.500
ALA_19:HA	VAL_20:HN	1.900	3.000
CYS_3:HN	LYS_4:HN	1.900	3.000
LYS_4:HN	ALA_5:HN	1.900	2.500
GLU_7:HN	THR_8:HN	2.500	3.500
ALA_9:HN	LYS_10:HN	1.900	3.000

NOE Constraints of Apa-M5 (continued)

Atom 1	Atom 2	Lower	Upper
LYS_10:HN	CYS_11:HN	1.900	3.000
ARG_12:HN	TRP_13:HN	1.900	3.000
TRP_13:HN	LYS_14:HN	1.900	3.000
CYS_15:HN	ASN_16:HN	1.900	3.000
PHE_17:HN	ILE_18:HN	1.900	3.000
ILE_18:HN	ALA_19:HN	1.900	3.000
ALA_19:HN	VAL_20:HN	1.900	3.000
LYS_4:HB*	ALA_5:HN	1.900	3.000
PRO_6:HB1	GLU_7:HN	1.900	3.000
PRO_6:HB2	GLU_7:HN	2.500	3.500
ALA_9:HB*	LYS_10:HN	1.900	2.500
LYS_10:HB*	CYS_11:HN	2.500	3.500
ARG_12:HB*	TRP_13:HN	1.900	3.000
TRP_13:HB1	LYS_14:HN	1.900	3.000
TRP_13:HB2	LYS_14:HN	1.900	3.000
LYS_14:HB*	CYS_15:HN	1.900	3.000
CYS_15:HB1	ASN_16:HN	2.500	3.500
CYS_15:HB2	ASN_16:HN	2.500	3.500
ASN_16:HB*	PHE_17:HN	1.900	3.000
PHE_17:HB1	ILE_18:HN	2.500	3.500
PHE_17:HB2	ILE_18:HN	2.500	3.500
ILE_18:HB	ALA_19:HN	2.500	3.500
ALA_19:HB*	VAL_20:HN	1.900	4.500
THR_8:HB,HA	CYS_11:HN	2.500	3.500
LYS_10:HA	TRP_13:HN	1.900	2.500
CYS_11:HA	LYS_14:HN	2.500	3.500
ARG_12:HA	CYS_15:HN	2.500	3.500
TRP_13:HA	ASN_16:HN	2.500	3.500
LYS_14:HA	PHE_17:HN	2.500	3.500
CYS_15:HA	ILE_18:HN	1.900	2.500
ASN_16:HA	ALA_19:HN	2.500	3.500
CYS_11:HA	CYS_15:HN	1.900	2.500
TRP_13:HA	PHE_17:HN	2.500	3.500
LYS_14:HA	ILE_18:HN	2.500	3.500

NOE Constraints of Apa-M5 (continued)

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_15:HA	ALA_19:HN	2.500	3.500
ASN_16:HA	VAL_20:HN	2.500	3.500
LYS_10:HA	TRP_13:HB1	2.400	5.000
LYS_10:HA	TRP_13:HB2	2.400	5.000
CYS_11:HA	LYS_14:HB*	2.400	4.000
TRP_13:O	LYS_14:HN	1.800	5.000
ARG_12:HA	CYS_15:HB1	2.400	5.000
ARG_12:HA	CYS_15:HB2	2.400	5.000
TRP_13:HA	ASN_16:HB*	2.400	4.000
LYS_14:HA	PHE_17:HB1	2.400	5.000
LYS_14:HA	PHE_17:HB2	2.400	5.000
ASN_16:HA	ALA_19:HB*	1.900	4.000
PHE_17:HA	VAL_20:HB	2.400	5.000
CYS_1:HA	ASN_2:HN	1.900	3.000
ASN_2:HN	ASN_2:HB2	1.900	3.500
ASN_2:HN	ASN_2:HB1	1.900	3.000
CYS_3:HN	CYS_3:HB1	1.900	3.000
CYS_3:HN	CYS_3:HB2	1.900	3.000
GLU_7:HN	GLU_7:HB1	1.900	3.000
GLU_7:HN	GLU_7:HB2	1.900	3.000
ALA_9:HN	ALA_9:HB*	1.900	4.000
LYS_10:HN	LYS_10:HB*	1.900	4.000
CYS_11:HN	CYS_11:HB1	1.900	3.000
CYS_11:HN	CYS_11:HB2	1.900	3.000
ARG_12:HN	ARG_12:HB*	1.900	4.000
TRP_13:HN	TRP_13:HB1	1.800	5.000
TRP_13:HN	TRP_13:HB2	1.800	5.000
LYS_14:HN	LYS_14:HB*	1.800	6.000
CYS_15:HN	CYS_15:HB1	1.800	5.000
CYS_15:HN	CYS_15:HB2	1.800	5.000
ASN_16:HN	ASN_16:HB*	1.800	6.000
PHE_17:HN	PHE_17:HB1	1.800	5.000
PHE_17:HN	PHE_17:HB2	1.800	5.000
ILE_18:HN	ILE_18:HB	1.800	5.000

NOE Constraints of Apa-M5 (continued)

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_1:HB1	CYS_11:HB1	1.900	5.000
ASN_2:HA	LYS_4:HN	1.900	4.000
ASN_2:HN	ALA_5:HB*	1.900	5.000
CYS_3:HA	PRO_6:HD1	1.900	4.000
CYS_3:HB1	CYS_15:HB1	1.900	5.000
ALA_5:HB*	PRO_6:HD1	1.900	5.000
ALA_5:HB*	PRO_6:HD2	1.900	5.000
PRO_6:HB1	ARG_12:HA	1.900	4.000
PRO_6:HG1	ARG_12:HA	1.900	4.000
GLU_7:HA	CYS_11:HB1	1.900	5.000
GLU_7:HA	CYS_11:HB2	1.900	5.000
THR_8:HB,HA	LYS_10:HN	1.900	5.400
THR_8:HB,HA	ALA_9:HB*	1.900	6.000
THR_8:HG2*	ALA_9:HN	1.900	5.000
THR_8:HG2*	LYS_10:HN	1.900	6.000
THR_8:HG2*	CYS_11:HN	1.900	6.000
ALA_9:HA	ARG_12:HG*	1.900	5.000
ARG_12:HG*	CYS_15:HB1	1.900	5.000
ARG_12:HG*	CYS_15:HB2	1.900	6.000
TRP_13:HZ3	1:LYS_14:HG1	1.900	5.000
LYS_14:HG1	PHE_17:HB1	1.900	4.000
LYS_14:HG2	PHE_17:HB1	1.900	4.000
LYS_14:HG1	PHE_17:HD*	1.900	7.000
ASN_16:HB*	ALA_19:HB*	1.900	7.000
ASN_16:HB*	VAL_20:HG2*	1.900	6.000
ASN_16:HB*	VAL_20:HG1*	1.900	7.000
PHE_17:HN	VAL_20:HG2*	1.900	6.000
PHE_17:HD*	ILE_18:HG11	1.900	7.000
PHE_17:HD*	VAL_20:HG2*	1.900	8.000
ILE_18:HG11	ALA_19:HN	1.900	5.000
ALA_5:HA	PRO_6:HD1	1.800	3.000
ALA_5:HA	PRO_6:HD2	1.800	3.000
ASN_16:HD2*	VAL_20:HG2*	1.900	7.000
TRP_13:HE3	LYS_14:HG1	1.900	4.000

Hydrogen Bond Constraints of Apa-Cro

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
ASN_2:OD1	LYS_4:HN	1.800	2.000
ASN_2:O	ALA_5:HN	1.800	2.000
THR_8:O	LYS_12:HN	1.800	2.000
GLN_9:O	ASN_13:HN	1.800	2.000
SER_10:O	LYS_14:HN	1.800	2.000
CYS_11:O	CYS_15:HN	1.800	2.000
LYS_12:O	LYS_16:HN	1.800	2.000
ASN_13:O	HIS_17:HN	1.800	2.000

Disulfide Bridge Constraints of Apa-Cro

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_1:CB	CYS_11:CB	3.770	3.920
CYS_3:CB	CYS_15:CB	3.770	3.920

NOE Constraints of Apa-Cro

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_1:HA	ASN_2:HN	1.900	2.500
LYS_4:HA	ALA_5:HN	1.900	3.500
PRO_6:HA	GLU_7:HN	1.900	2.500
GLU_7:HA	THR_8:HN	1.900	3.500
GLN_9:HA	SER_10:HN	1.900	3.500
SER_10:HA	CYS_11:HN	1.900	3.500
LYS_12:HA	ASN_13:HN	1.900	3.500
ASN_13:HA	LYS_14:HN	1.900	3.500
LYS_14:HA	CYS_15:HN	1.900	3.500
CYS_15:HA	LYS_16:HN	1.900	3.500
LYS_16:HA	HIS_17:HN	1.900	3.000
HIS_17:HA	ALA_18:HN	1.900	3.000
ALA_18:HA	GLY_19:HN	1.900	2.500
GLY_19:HA1	ARG_20:HN	1.900	2.500
GLY_19:HA2	ARG_20:HN	1.900	2.500
CYS_3:HN	LYS_4:HN	1.900	3.000
LYS_4:HN	ALA_5:HN	1.900	2.500
GLU_7:HN	THR_8:HN	1.900	3.000

NOE Constraints of Apa-Cro (continued)

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
GLN_9:HN	SER_10:HN	1.900	3.000
SER_10:HN	CYS_11:HN	1.900	3.000
CYS_11:HN	LYS_12:HN	1.900	3.000
LYS_12:HN	ASN_13:HN	1.900	3.000
ASN_13:HN	LYS_14:HN	1.900	3.000
LYS_14:HN	CYS_15:HN	1.900	3.000
CYS_15:HN	LYS_16:HN	1.900	3.000
LYS_16:HN	HIS_17:HN	1.900	3.000
CYS_1:HB1	ASN_2:HN	1.900	3.500
CYS_1:HB2	ASN_2:HN	1.900	3.500
CYS_3:HB1	LYS_4:HN	1.900	3.500
LYS_4:HB1	ALA_5:N	1.900	3.500
LYS_4:HB2	ALA_5:HN	1.900	3.500
PRO_6:HB1	GLU_7:HN	1.900	3.500
GLN_9:HB1	SER_10:HN	1.900	3.000
GLN_9:HB2	SER_10:HN	1.900	3.500
SER_10:HB1	CYS_11:HN	1.900	3.500
SER_10:HB2	CYS_11:HN	1.900	3.500
CYS_11:HB1	LYS_12:HN	1.900	3.000
CYS_11:HB2	LYS_12:HN	1.900	3.500
LYS_12:HB*	ASN_13:HN	1.900	4.000
ASN_13:HB*	LYS_14:HN	1.900	4.000
LYS_14:HB*	CYS_15:HN	1.900	4.000
CYS_15:HB1	LYS_16:HN	1.900	3.500
CYS_15:HB2	LYS_16:HN	1.900	3.000
GLN_9:HA	LYS_12:HA	1.900	3.500
SER_10:HA	ASN_13:HA	1.900	3.500
CYS_11:HA	LYS_14:HA	1.900	3.500
LYS_12:HA	CYS_15:HA	1.900	3.500
ASN_13:HA	LYS_16:HA	1.900	3.500
LYS_14:HA	HIS_17:HA	1.900	3.500
GLN_9:HA	LYS_12:HB*	1.900	5.000
SER_10:HA	ASN_13:HB*	1.900	5.000
CYS_11:HA	LYS_14:HB*	1.900	5.000

NOE Constraints of Apa-Cro (continued)

Atom 1	Atom 2	Lower	Upper
LYS_12:HA	CYS_15:HB1	1.900	4.000
LYS_12:HA	CYS_15:HB2	1.900	4.000
LYS_14:HA	HIS_17:HB1	1.900	5.000
CYS_15:HA	ALA_18:HB*	1.900	5.000
GLN_9:HA	ASN_13:HN	1.900	3.500
SER_10:HA	LYS_14:HN	1.900	4.000
CYS_11:HA	CYS_15:HN	1.900	3.500
LYS_12:HA	LYS_16:HN	1.900	3.500
CYS_15:HA	GLY_19:HN	1.900	3.500
ASN_2:HN	ASN_2:HB1	1.900	3.000
ASN_2:HN	ASN_2:HB2	1.900	3.500
CYS_3:HN	CYS_3:HB1	1.900	3.000
CYS_3:HN	CYS_3:HB2	1.900	3.000
LYS_4:HN	LYS_4:HB1	1.900	3.000
LYS_4:HN	LYS_4:HB2	1.900	3.000
ALA_5:HN	ALA_5:HB*	1.900	4.000
GLU_7:HN	GLU_7:HB1	1.900	3.000
GLU_7:HN	GLU_7:HB2	1.900	3.500
GLN_9:HN	GLN_9:HB1	1.900	3.000
GLN_9:HN	GLN_9:HB2	1.900	3.000
SER_10:HN	SER_10:HB1	1.900	3.000
SER_10:HN	SER_10:HB2	1.900	3.000
CYS_11:HN	CYS_11:HB1	1.900	3.000
CYS_11:HN	CYS_11:HB2	1.900	3.000
LYS_12:HN	LYS_12:HB*	1.900	4.000
ASN_13:HN	ASN_13:HB*	1.900	4.000
LYS_14:HN	LYS_14:HB*	1.900	4.000
CYS_15:HN	CYS_15:HB1	1.900	3.500
CYS_15:HN	CYS_15:HB2	1.900	3.000
HIS_17:HN	HIS_17:HB1	1.900	3.000
HIS_17:HN	HIS_17:HB2	1.900	3.000
ALA_18:HN	ALA_18:HB*	1.900	4.000
ARG_20:HN	ARG_20:HB1	1.900	3.500
ARG_20:HN	ARG_20:HB2	1.900	3.500

NOE Constraints of Apa-Cro (continued)

Atom 1	Atom 2	Lower	Upper
LYS_21:HN	LYS_21:HB1	1.900	3.500
LYS_21:HN	LYS_21:HB2	1.900	3.500
CYS_1:HB1	CYS_11:HB1	1.900	5.000
ASN_2:HA	LYS_4:HN	1.900	4.000
CYS_3:HA	PRO_6:HB1	1.900	5.000
CYS_3:HA	PRO_6:HB2	1.900	4.000
CYS_3:HA	PRO_6:HG1	1.900	5.000
CYS_3:HA	PRO_6:HG2	1.900	4.000
CYS_3:HB1	CYS_15:HB1	1.900	5.000
ALA_5:HN	PRO_6:HD1	1.900	3.500
ALA_5:HN	PRO_6:HD2	1.900	3.500
ALA_5:HA	PRO_6:HD1	1.900	3.000
ALA_5:HA	PRO_6:HD2	1.900	3.000
PRO_6:HB1	LYS_12:HA	1.900	4.000
PRO_6:HG1	LYS_12:HA	1.900	4.000
PRO_6:HB1	LYS_12:HN	1.900	5.000
THR_8:HG2*	GLN_9:HN	1.800	6.000
THR_8:HG2*	SER_10:HN	1.800	6.000
THR_8:HG2*	CYS_11:HN	1.800	6.000
THR_8:HG2*	THR_8:HN	1.900	5.000
LYS_12:HN	LYS_12:HG1	1.900	5.000
LYS_12:HN	LYS_12:HG2	1.900	5.000
LYS_16:HB*	HIS_17:HN	1.900	4.000
ALA_18:HB*	GLY_19:HN	1.900	4.500
ASN_13:HA	LYS_16:HB*	1.900	5.000
LYS_16:HN	LYS_16:HB*	1.900	4.000
LYS_16:HN	LYS_16:HG2	1.900	5.000
LYS_16:HN	LYS_16:HG1	1.900	5.000
ARG_20:HN	ARG_20:HD*	1.900	6.000
LYS_14:HN	LYS_14:HG2	1.900	5.000
LYS_4:HN	LYS_4:HG*	1.900	6.000
THR_8:HB,HA	GLN_9:HN	1.900	2.500
THR_8:HB,HA	SER_10:HN	1.900	6.400

Hydrogen Bond Constraints of Max

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
ARG_1:O	LEU_5:HN	1.80	2.00
MET_2:O	SER_6:HN	1.80	2.00
CYS_3:O	CYS_7:HN	1.80	2.00
GLN_4:O	ARG_8:HN	1.80	2.00
LEU_5:O	SER_9:HN	1.80	2.00
SER_6:O	LEU_10:HN	1.80	2.00
CYS_7:O	GLY_11:HN	1.80	2.00
LEU_13:HN	VAL_24:O	1.80	2.00
GLU_22:O	LYS_15:HN	1.80	2.00
ILE_17:HN	LYS_20:O	1.80	2.00
LYS_20:HN	ILE_17:O	1.80	2.00
LYS_15:O	GLU_22:HN	1.80	2.00
VAL_24:HN	LEU_13:O	1.80	2.00

Disulfide Bridge Constraints of Max

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_3:CB	CYS_21:CB	3.72	3.99
CYS_7:CB	CYS_23:CB	3.72	3.99

NOE Constraints of Max

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
CYS_3:HN	1:GLN_4:HN	1.90	3.00
GLN_4:HN	1:LEU_5:HN	1.90	3.00
LEU_5:HN	1:SER_6:HN	1.90	3.00
SER_6:HN	1:CYS_7:HN	1.90	3.00
CYS_7:HN	1:ARG_8:HN	1.90	3.00
SER_9:HN	1:LEU_10:HN	1.90	3.00
LEU_10:HN	1:GLY_11:HN	1.90	3.50
LEU_13:HN	1:VAL_24:HN	1.90	3.50
LYS_15:HN	1:GLU_22:HN	1.90	3.50
ILE_17:HN	1:LYS_20:HN	1.90	3.50
ASP_19:HN	1:LYS_20:HN	1.90	3.50
ALA_16:HA	1:CYS_21:HA	1.90	3.00
CYS_3:HA	1:SER_6:HN	1.90	3.50

NOE Constraints of Max (continued)

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
GLN_4:HA	1:CYS_7:HN	1.90	3.50
LEU_5:HA	1:ARG_8:HN	1.90	3.50
CYS_7:HA	1:LEU_10:HN	1.90	3.50
ARG_8:HA	1:GLY_11:HN	1.90	3.50
CYS_3:HA	1:CYS_7:HN	1.90	3.50
GLN_4:HA	1:ARG_8:HN	1.90	3.50
LEU_5:HA	1:SER_9:HN	1.90	3.50
SER_6:HA	1:LEU_10:HN	1.90	4.00
CYS_3:HA	1:SER_6:HB*	1.90	5.00
GLN_4:HA	1:CYS_7:HB1	1.90	4.00
CYS_7:HB2	1:GLN_4:HA	1.90	5.00
LEU_5:HA	1:ARG_8:HB1	1.90	4.00
ARG_8:HB2	1:LEU_5:HA	1.90	4.00
SER_6:HA	1:SER_9:HB*	1.90	5.00
MET_2:HA	1:CYS_3:HN	1.90	3.00
CYS_3:HA	1:GLN_4:HN	1.90	3.50
GLN_4:HA	1:LEU_5:HN	1.90	3.50
LEU_5:HA	1:SER_6:HN	1.90	3.50
SER_6:HA	CYS_7:HN	1.90	3.50
CYS_7:HA	ARG_8:HN	1.90	3.50
ARG_8:HA	SER_9:HN	1.90	3.50
SER_9:HA	LEU_10:HN	1.90	3.50
LEU_10:HA	GLY_11:HN	1.90	3.50
LEU_13:HA	GLY_14:HN	1.90	2.50
LYS_15:HA	ALA_16:HN	1.90	2.50
ALA_16:HA	ILE_17:HN	1.90	2.50
ILE_17:HA	GLY_18:HN	1.90	2.50
LYS_20:HA	CYS_21:HN	1.90	2.50
CYS_21:HA	GLU_22:HN	1.90	2.50
GLU_22:HA	CYS_23:HN	1.90	2.50
CYS_23:HA	VAL_24:HN	1.90	2.50
VAL_24:HA	LYS_25:HN	1.90	4.00
ALA_16:HA	GLU_22:HN	1.90	3.00
LYS_15:HA	CYS_23:HN	1.90	3.00

NOE Constraints of Max (continued)

<u>Atom 1</u>	<u>Atom 2</u>	<u>Lower</u>	<u>Upper</u>
ILE_17:HA	CYS_21:HN	1.90	3.00
MET_2:HB1	CYS_3:HN	1.90	3.50
MET_2:HB2	CYS_3:HN	1.90	3.50
CYS_3:HB1	GLN_4:HN	1.90	3.00
CYS_3:HB2	GLN_4:HN	1.90	3.50
SER_6:HB*	CYS_7:HN	1.90	4.50
CYS_7:HB1	ARG_8:HN	1.90	3.50
CYS_7:HB2	ARG_8:HN	1.90	3.00
ARG_8:HB1	SER_9:HN	1.90	3.50
ARG_8:HB2	SER_9:HN	1.90	3.50
SER_9:HB*	LEU_10:HN	1.90	4.50
ALA_16:HB*	ILE_17:HN	1.90	5.00
CYS_21:HB2	GLU_22:HN	1.90	3.50
CYS_21:HB1	GLU_22:HN	1.90	3.50
GLU_22:HB1	CYS_23:HN	1.90	3.50
GLU_22:HB2	CYS_23:HN	1.90	3.50
CYS_23:HB1	VAL_24:HN	1.90	3.50
CYS_23:HB2	VAL_24:HN	1.90	4.00
VAL_24:HB	LYS_25:HN	1.90	4.00
MET_2:HN	MET_2:HB1	1.90	3.50
MET_2:HN	MET_2:HB2	1.90	3.50
CYS_3:HN	CYS_3:HB1	1.90	3.00
CYS_3:HN	CYS_3:HB2	1.90	3.00
SER_6:HN	SER_6:HB*	1.90	4.50
CYS_7:HN	CYS_7:HB2	1.90	3.00
CYS_7:HN	CYS_7:HB1	1.90	3.00
ARG_8:HN	ARG_8:HB1	1.90	3.00
ARG_8:HN	ARG_8:HB2	1.90	3.50
SER_9:HN	SER_9:HB*	1.90	4.50
ALA_16:HN	ALA_16:HB*	1.90	4.50
CYS_21:HN	CYS_21:HB1	1.90	3.50
CYS_21:HN	CYS_21:HB2	1.90	3.50
GLU_22:HN	GLU_22:HB1	1.90	3.50
GLU_22:HN	GLU_22:HB2	1.90	3.50

CYS_23:HN	CYS_23:HB2	1.90	3.50
CYS_23:HN	CYS_23:HB1	1.90	3.50
VAL_24:HN	VAL_24:HB	1.90	3.50
GLY_14:HA2	CYS_23:HA	1.90	3.00
CYS_23:HA	GLY_14:HA1	1.90	3.00
MET_2:HA	LEU_5:HB*	1.90	5.00
CYS_7:HA	LEU_10:HB*	1.90	5.00
GLY_14:HA2	LYS_15:HN	1.90	2.50
GLY_14:HA1	LYS_15:HN	1.90	2.50
GLY_14:HA2	VAL_24:HN	1.90	3.00
VAL_24:HN	GLY_14:HA1	1.90	3.00
ARG_1:HB*	MET_2:HN	1.90	4.50
GLN_4:HB*	LEU_5:HN	1.90	4.50
LEU_5:HB*	SER_6:HN	1.90	4.50
LEU_13:HB*	GLY_14:HN	1.90	4.50
LYS_15:HB*	ALA_16:HN	1.90	5.00
ILE_17:HG2*	GLY_18:HN	1.90	4.50
LYS_20:HB*	CYS_21:HN	1.90	4.50
GLN_4:HN	GLN_4:HB*	1.90	4.50
LEU_5:HN	LEU_5:HB*	1.90	4.50
LEU_10:HN	LEU_10:HB*	1.90	4.50
LEU_13:HN	LEU_13:HB*	1.90	4.50
LYS_15:HN	LYS_15:HB*	1.90	4.50
ILE_17:HN	ILE_17:HG2*	1.90	4.50
LYS_20:HN	LYS_20:HB*	1.90	4.50

END

DATE

FILMED

2/23/94