

Structural Templates for Protein Design: The Assembly of Peptidyl
Motifs with Defined Supersecondary Structure

Thesis by
Mary D. Struthers

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1996
(Submitted May 31, 1996)

© 1996

Mary Struthers

All rights Reserved

To Ranabir
and
To my parents

Acknowledgments

I would like to thank my advisor Professor Barbara Imperiali for creating a stimulating research environment in which diversity of approach and thought was encouraged. Barbara's enthusiasm and energy for science were infectious and under her guidance I've learned a lot about chemistry and biochemistry.

I thank all Imperiali group members past and present, for helping to create an atmosphere in which doing science was fun. A special thanks to Richard Cheng, Dr. Stewart Fisher, and Dr. Alicia Torrado my coworkers and collaborators on several projects discussed in this thesis. Thanks to Richard, Jen Ottesen, Ranabir Sinha Roy, Tamara Hendrickson, and Barbara for proofreading this manuscript. Special thanks to Grant Walkup and Jen Ottesen for their computer support during the writing of this thesis. Jen, thank you for understanding.

Dr. Keith Rickert, in addition to being an endless source of interesting trivia, you've also been a great friend and coworker. I already miss your input and advice.

To Dr. Jeff Spencer, thank you for your friendship and support (and your patience in teaching me racquetball). It's been a pleasure to have had the chance to work with you.

Mom and Dad, I've accomplished what I have today, in part, because I've always known that I have your support and encouragement in whatever I choose to do. The discipline and perseverance you have taught me has served me well. Thank You.

Tamara, without your support and friendship, the last four years would have been a lot harder and a lot less fun. Any lab I go to will seem a little empty without you.

To Ranabir, my best friend and fiancée, your companionship and support as well as your scientific insight and encouragement have made this possible. You've inspired and motivated me, and I couldn't have done it without you, thank you. WPT, Bahut aur humesha.

I would also like to acknowledge a National Science Foundation Predoctoral Fellowship and a Parsons Foundation Fellowship for financial support.

Abstract

An approach to the design of small (<40 residues), independently folded, monomeric motifs based on naturally occurring protein domains is presented. Natural peptidyl systems of this limited size generally require either disulfides or metal ligation for proper folding and stability. In order to compensate for the stability originally afforded by native cross-links, defined β -turns are employed as structural nucleation elements and non-standard amino acids are used to control structure and incorporate function.

The construction of a triple stranded β -sheet motif based on a 30-residue fragment of erabutoxin b, a member of a family of neurotoxins, was initially attempted. Sequences incorporating a defined type II β -turn and a single original disulfide were investigated by 2D NMR techniques. These experiments revealed that the designed polypeptides were flexible in solution, adopting primarily random coil conformations. Although the peptide did not fold into the intended conformation, these studies provided guidelines for the subsequent selection of natural templates.

The $\beta\beta\alpha$ motif exemplified by the zinc finger domains was chosen for the second design attempt. Using an iterative design process, sequences based on these domains were evaluated for their ability to fold, in the *absence* of metal ions, by both circular dichroism and NMR techniques. Type II and II' β -turns were investigated as structural nucleation elements and an unnatural metal binding amino acid, 3-(1,10-phenanthroline-2-yl)-L-alanine (Fen), was incorporated to serve as a reporter group. A well defined metal-independent motif was obtained in five iterations. These studies culminated in the design and synthesis of a 23-residue peptide, **BBA1**, which adopted a defined tertiary structure in the *absence* of metal ions. The solution structure of **BBA1** was

solved using NOE-restrained simulated annealing, and confirmed the desired $\beta\beta\alpha$ topology.

In a second project, unnatural amino acids and defined β -turns were used in the construction of metal-dependent inhibitors for cAMP-dependent protein kinase. Control of peptide conformation either through metal ligation or disulfide bond formation was shown to modulate the inhibitory properties of sequences based on the heat stable protein kinase inhibitor.

Table of Contents

Abstract.....	vi
List of Figures.....	xi
List of Tables.....	xvi
List of Abbreviations	xix
 Chapter 1: Introduction	
References.....	20
 Chapter 2: Natural peptide domains as templates for protein design: Evaluation of a designed $\beta\beta\beta$ motif patterned after snake neurotoxins	
Introduction.....	30
Results.....	36
TOX1	36
Design Rationale.....	36
Structural Analysis.....	40
NAT1 and TOX2.....	45
Structural Analysis.....	47
Discussion.....	53
Conclusions.....	57
Experimental.....	60
References.....	67
 Chapter 3: Evolution of a metal-independent $\beta\beta\alpha$ motif based on the zinc finger domains	
Introduction.....	75
Results and Discussion.....	79
ZFO1.....	81
Design.....	81
CD Analysis	83
ZFO2.....	85
Design.....	85

CD Analysis.....	85
NMR Analysis.....	87
ZFO3	90
Design.....	90
CD Analysis.....	91
NMR Analysis.....	94
Use of Model Peptides to Examine the <i>cis-trans</i>	
Isomerization of DPro in Type II' Turns.....	95
Conclusions.....	101
Experimental.....	104
References.....	121

Chapter 4: BBA1: Structural analysis of an independently folded $\beta\beta\alpha$ motif

Introduction.....	127
Results	132
ZFO4	132
CD Analysis.....	132
NMR Analysis.....	134
BBA1 (ZFO5)	140
CD Analysis	141
NMR Analysis.....	143
Structure Determination.....	147
Stability of BBA1	152
ANS Binding	155
ZFO6	157
Structural Analysis.....	157
ANS Binding	160
BBA2	161
Design and Modeling.....	162
NMR Analysis.....	164
Discussion.....	172
Analysis of BBA1 Structure.....	173
Role of the Type II' β -Turn.....	178
Role of the Fen Amino Acid.....	179
Conclusions.....	181

Experimental.....	183
References.....	206

Chapter 5: Design and evaluation of a metallopeptide regulatory motif for cAMP-dependent protein kinase

Introduction.....	211
Results.....	219
Design of an Inhibitory Linker Region	219
Metal Binding Properties of Peptide MSW1	222
Metal Binding Properties of Peptide DMW1	230
Metal Binding Properties of MSW2	234
Assay for Protein Kinase Activity	239
Characterization of the Inhibitory Properties of MSW1 , DMW1 , MSW2 with PK.....	241
PKCYS : Inhibitor Cyclization with a Disulfide.....	243
Discussion.....	245
Conclusion	250
Experimental.....	253
References.....	267

Appendix A: Electrostatic stabilization of sheet and turn structure

Introduction.....	272
Results and Discussion.....	273
Electrostatic Stabilization of a Type II b-Turn.....	273
<i>a</i> Maa: A non-standard Amino Acid to Stabilize Sheet and Turn Structure.....	285
Conclusion	292
Experimental.....	293
References.....	301

List of Figures

Chapter 1

1-1 Several examples of the unnatural amino acids exploited for the control of structure and function in the Imperiali Group.....	14
1-2 The dihedral angles (ϕ, ψ) of the central two residues of a four residue b-turn define the β -turn-type	16
1-3 The different topologies of the type I and type II β -turns.....	17

Chapter 2

2-1 Schematic representation of the proposed $\beta\beta\beta$ motif	33
2-2 The structure of erabutoxin b.....	34
2-3 Sequence of erabutoxin b residues 24-53.....	35
2-4 Design of TOX1	39
2-5 Conformation of a type II β -turn	42
2-6 Schematic illustration of the NOEs which define an ideal β -sheet structure.....	43
2-7 The $d_{\alpha N(i+1, i+2)}$ and $d_{NN(i+1, i+2)}$ distances in an ideal type I or type II β -turn.....	48
2-8 A sequential d_{NN} NOE for a peptide backbone in an extended conformation.....	49
2-9 The dihedral angles that defined disulfide geometry.....	53

Chapter 3

3-1 The consensus sequence and structure of natural zinc finger domains	76
3-2 The CD spectrum of a natural zinc finger, ADR1.....	78
3-3 Design of the first three generations of zinc finger peptides.....	80
3-4 Structure of 3-(1,10-phenanthrol-2-yl)-L-alanine or Fen	82
3-5 CD analysis of ZFO1	84
3-6 CD analysis of ZFO2	86
3-7 Similar regions of the TOCSY spectra of the ZFO2-Zn²⁺ complex and the uncomplexed peptide.....	88
3-8 The differing effects of type II and type II' turns on β -sheet formation.....	91
3-9 CD analysis of ZFO3	92

3-10	Section of the TOCSY spectrum of ZFO3	93
3-11	The amide region of the 1D ^1H spectra of HF1	96
3-12	NMR identification of the <i>cis</i> and <i>trans</i> proline amide isomers of HF1	97
3-13	Expected NOE interaction in a heterochiral type II' reverse turn.....	98
3-14	Upfield region of the 1D ^1H NMR spectrum of VF1	100
3-15	1D ^1H NMR spectrum of HF1	112
3-16	1D ^1H NMR spectrum of VF1	116
3-17	1D ^1H NMR spectrum of TF1	118
3-18	1D ^1H NMR spectrum of AF1	120

Chapter 4

4-1	Summary of the evolution of a metal independent $\beta\beta\alpha$ motif based on the zinc finger domains.....	131
4-2	CD analysis of ZFO4	133
4-3	NOEs indicative of secondary structure in uncomplexed ZFO4	135
4-4	NOESY spectrum of ZFO4	139
4-5	Measurement of the <i>cis-trans</i> isomer ratio for BBA1 by 1D NMR.....	140
4-6	CD Analysis of BBA1	142
4-7	NOEs indicative of secondary structure in BBA1	144
4-8	H/D exchange experiment for BBA1 monitored by NMR.....	145
4-9	Long range NOEs present in the NOESY spectrum of BBA1 indicative of tertiary structure.....	146
4-10	Schematic representation of a NOE-restrained molecular dynamics simulated annealing procedure.....	149
4-11	Overlay of the backbone atoms of the 29 NMR-derived structures of BBA1	150
4-12	Overlay of the heavy atoms of all 29 BBA1 structures.....	151
4-13	Thermal denaturation of BBA1	153
4-14	Chemical denaturation of BBA1	154
4-15	Titration of BBA1 with ANS.....	156
4-16	NOEs indicative of secondary structure in ZFO6	158
4-17	Similar regions of the NOESY spectra of BBA1 and ZFO6	159
4-18	Titration of ZFO6 with ANS.....	160
4-19	The position of the Fen residue within the NMR structure of BBA1	163

4-20	A portion of the model for BBA2 superimposed on the NMR structure of BBA1	164
4-21	The amide-alpha region of the TOCSY spectrum of BBA2 at 7 °C	166
4-22	The amide-alpha region of the TOCSY spectrum of BBA2 at 17 °C	167
4-23	The secondary structure of BBA2 at 17 °C.....	168
4-24	NOEs indicative of β -hairpin formation for BBA2	169
4-25	Long range NOEs for BBA2 indicative of tertiary structure.....	170
4-26	Stereo image of a representative NMR structure for BBA1	173
4-27	The solvent accessible surface of representative NMR structures of BBA1 and Xfin.....	176
4-28	Charged amino acids which could be employed for the electrostatic stabilization of β -sheet structure	181
4-29	Plot of amide proton chemical shift <i>vs</i> temperature for ZFO4	191

Chapter 5

5-1	Schematic representation of the proposed metal dependent regulatory peptide.....	212
5-2	Metal binding regions of the regulatory metallo-peptide design.....	214
5-3	The catalytic subunit of cAMP-dependent protein kinase phosphorylates a serine residue within a consensus peptide sequence.....	217
5-4	Schematic illustration of the regulation of an enzymatic cascade by a designed metallopeptide and the catalytic subunit of cAMP dependent protein kinase (PK).....	218
5-5	Titration of MSW1 with ZnCl_2	223
5-6	Proposed equilibria for the formation of MSW1 -metal complexes	224
5-7	Absorbance binding isotherm for the titration of MSW1 with ZnCl_2	225
5-8	Absorbance spectra for the titration of MSW1 with CoCl_2	227
5-9	Absorbance binding isotherm for the titration of MSW1 with CoCl_2	228
5-10	Molar extinction coefficient spectra of the MSW1 - Co^{2+} and MSW1 - Zn^{2+} complexes.....	229
5-11	Absorbance spectra for the titration of DMW1 with ZnCl_2	231
5-12	Absorbance spectra for the titration of DMW1 with CoCl_2	232

5-13	Comparison of the extinction coefficient spectra of the MSW1-Co²⁺ complex and the DMW1-Co²⁺ complex.....	233
5-14	Absorbance spectra of the titration of MSW2 with CoCl ₂ at low peptide concentration	235
5-15	Absorbance spectra of the titration of MSW2 with CoCl ₂ at high peptide concentration	236
5-16	Absorbance spectra of the titration of MSW2 with ZnCl ₂ at low peptide concentration	237
5-17	Comparison of a proposed conformation for PKCYS-O and the structure of the analogous region of PKI(5-24) bound in the active site of the catalytic subunit of c-AMP dependent protein kinase.....	250
5-18	Metal mediated control of c-AMP dependent protein kinase inhibition.....	251
5-19	An example of unnatural amino acids which place the metal chelating moiety within the peptide backbone	252
5-20	Inhibition of PK by both uncomplexed peptide and the Co ²⁺ complex of MSW2	262
5-21	Inhibition of PK by both uncomplexed peptide and the Co ²⁺ complex of DMW1	263
5-22	Inhibition of PK by both uncomplexed peptide and the Co ²⁺ complex of MSW1	264
5-23	Inhibition of PK by both reduced and oxidized forms of peptide PKCYS	265
5-24	PK activity in the presence of divalent metal cations	266

Appendix A

A-1	NOEs indicative of type II β -turn formation in tetrapeptides.....	274
A-2	Summary of the NOEs observed in the ROESY spectrum of peptide PA1	276
A-3	CD spectra of peptide PA1	278
A-4	Summary of the NOEs observed in the ROESY spectrum of GG1	280
A-5	The CD spectra of peptide GG1 at various pH.....	281
A-6	Comparative C β -C β distances (Å) for an α -helix, ₃₁₀ helix, and type II β -turn.....	284
A-7	Structure of <i>a</i> Maa (1) and the corresponding derivative, Fmoc-L- <i>a</i> Maa(Boc)-OH (2) protected for solid phase peptide synthesis	285

A-8 Schematic representation of an amphiphilic β -hairpin employing inter-strand <i>a</i> Maa-Asp salt bridges.....	286
A-9 Synthesis of Fmoc- <i>a</i> Maa(Boc)-OH (2)	288
A-10 Summary of NOEs observed in the ROESY spectrum of MD1	289

List of Tables

Chapter 1

1-1 The dihedral angles of the central two residues of the most commonly occurring β -turn-types.....	16
---	----

Chapter 2

2-1 Sequences of peptides based on erabutoxin b 24-53.....	36
2-2 The dihedral angles of the type I turn in the native sequence of erabutoxin b	37
2-3 Summary of $d_{NN}(i, i+1)$ NOEs observed for peptides TOX1 , NAT1 , and TOX2	50
2-4 The relative intensities of the sequential $d_{\alpha N}(i, i+1)$ NOEs for peptides TOX1 , NAT1 , and TOX2	51
2-5 Comparison of amide proton variable temperature coefficients for peptides TOX1 , NAT1 , and TOX2	52
2-6 Analysis of the geometries and relative energies of the disulfides present in erabutoxin b.....	56
2-7 NMR assignments for TOX1	64
2-8 NMR assignments for NAT1	65
2-9 NMR assignments for TOX2	66

Chapter 3

3-1 Conformational analysis of the <i>trans</i> isomer of HF1	99
3-2 Conformational analysis of model type II' tetrapeptides.....	101
3-3 Partial NMR assignment of ZFO2	109
3-4 Partial NMR assignment of the spin systems ZFO3	110
3-5 NMR assignment of the <i>cis</i> and <i>trans</i> DPro ³ amide isomers of HF1	113
3-6 Significant NOEs observed in the ROESY spectrum for the <i>trans</i> DPro ³ amide isomer of peptide HF1	114
3-7 NMR assignment of the <i>cis</i> and <i>trans</i> DPro ² amide isomers of VF1	115
3-8 Significant NOEs observed in the ROESY spectrum for the <i>trans</i> DPro ² amide isomer of peptide VF1	115
3-9 NMR assignment of the <i>cis</i> and <i>trans</i> DPro ² amide isomers of TF1	117
3-10 Significant NOEs observed in the ROESY spectrum for the <i>trans</i> DPro ² amide isomer of peptide TF1	117

3-11 NMR assignment of the <i>cis</i> and <i>trans</i> DPro ² amide isomers of AF1.....	119
3-12 Significant NOEs observed in the ROESY spectrum for the <i>trans</i> DPro ² amide isomer of peptide AF1.....	119

Chapter 4

4-1 Amide proton temperature coefficients for ZFO4.....	137
4-2 The average value of the RMSDs of all 29 NMR derived structures from the average structure.....	152
4-3 NMR assignment of ZFO4.....	187
4-4 NMR assignment of BBA1.....	188
4-5 NMR assignment of ZFO6.....	189
4-6 NMR assignment of BBA2.....	190
4-7 Pseudoatom corrections added to the upper bounds of restraints used in the simulated annealing calculations.....	194
4-8 List of restraints used in the structure generation process.....	201
4-9 The stages of NOE-restrained simulated annealing used for the structure generation process.....	203
4-10 Hydrogen bonding interactions in the 29 NMR-derived structures of BBA1.....	205

Chapter 5

5-1 The sequences of the peptides assembled and analyzed in the development of a metal dependent inhibitory peptide for cAMP- dependent protein kinase.....	219
5-2 The inhibitory constants of fragments of PKI.....	221
5-3 The pH dependence of the MSW1-Co ²⁺ dissociation constant.....	238
5-4 Summary of the results of PK inhibition by metal switch peptides.....	243
5-5 Inhibition constants for the oxidized and the reduced forms of the inhibitor PKCYS.....	245
5-6 Comparison of the metal binding properties of MSW2 and Ac- 4Bpa-Thr-Pro-DSer-Val-Phe-NH ₂	247

Appendix A

A-1 NOEs observed in the ROESY spectrum of peptide PA1.....	275
A-2 Amide proton temperature coefficients for peptide PA1.....	275

A-3	Summary of NOEs between amide protons in the ROESY spectrum of peptide GG1	279
A-4	Amide proton variable temperature coefficients for peptide GG1	282
A-5	Summary of NOEs indicative of turn and sheet structure observed in the ROESY spectrum of peptide MD1	290
A-6	Comparison of the NOEs observed in the ROESY experiments of peptides MD1 , KD1 , and VPDSF	291
A-7	Comparison of the amide proton temperature coefficients for peptides MD1 , KD1 , and VPDSF	291
A-8	Assignment table of the ^1H NMR spectrum of PA1	299
A-9	Assignment table of the ^1H NMR spectrum of GG1	300
A-10	Assignment table of the ^1H NMR spectrum of MD1	301

List of Abbreviations

Standard three-letter and one-letter codes are used for the naturally occurring amino acids.

1D	One-Dimensional
2D	Two-Dimensional
4Bpa	2-amino-3-(2,2'-bipyridy-4-yl)-propanoic acid
Abs	Absorbance
Ac	Acetyl
aMaa	(2S, 3S)- β -methyl- β -amino-alanine
amu	Atomic Mass Unit
ANS	8-anilino-1-naphthalenesulfonic acid
ATP	Adenosine 5'-triphosphate
Baa	β -Aminoalanine
Bmc	(2R, 3R)- β -methyl-cysteine
BME	beta-mercato ethanol
Boc	<i>tert</i> -Butoxycarbonyl
BOP	Benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate
calcd	Calculated
cAMP	Adenosine 3',5'-monophosphate
CD	Circular Dichroism
COSY	Correlation Spectroscopy
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	<i>N,N'</i> -dicyclohexylcarbodiimide
DCM	Dichloromethane
DEAD	Diethylazodicarboxylate
DIEA	Diisopropylethylamine
DIPCDI	Diisopropylcarbodiimide
DMAD	Dimethylazodicarboxylate
DMF	<i>N,N</i> -dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DQCOSY	Double Quantum filtered spectroscopy

ESI	Electrospray ionization Mass Spectroscopy
FAB	Fast Atom Bombardment Mass Spectroscopy
Fen	3-(1,10-Phenanthrol-2-yl)-L-alanine
FID	Free Induction Decay
Fmoc	9-Fluorenylmethyloxycarbonyl
FmocOSu	9-Fluorenylmethyloxycarbonylsuccinimidyl carbonate
HEPES	(N-[2-Hydroxyethyl]piperazine-N'-[2-ethane]-sulfonic acid]
HOBt	1-Hydroxybenzotriazole
HPLC	High Pressure Liquid Chromatography
K _i	Inhibition constant
LB	Lowerbound
LG	Linear Gradient
MALDI	Matrix-Assisted Laser Desorption Ionization
MES	1-(N-Morpholino)ethanesulfonic acid
mmCOSY	Magnitude mode correlation spectroscopy
MS	Mass Spectroscopy
MW	Molecular weight
ND	Not determined
NMM	N-methylmorpholine
NMP	N-methylpyrrolidinone
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
obsd	Observed
OPfp	Pentafluorophenyl
PDMS	Plasma Desorption Mass Spectroscopy
PK	Catalytic subunit of cAMP-dependent protein kinase
Pmc	2,2,3,7,8,-pentamethyl-chroman-6-6sulfonyl chloride
ppb	Parts per billion
ROESY	Rotating Frame Overhauser Spectroscopy
SA	Simulated Annealing
satd.	Saturated
SPPS	Solid Phase Peptide Synthesis
tBu	<i>tert</i> -Butyl
TEA	Triethylamine
tert	Tertiary
TFA	Trifluoroacetic acid

THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TMS	Tetramethylsilane
TOCSY	Total Correlation Spectroscopy
TPPI	Time Proportional Phase Increments
Trt	Trityl
UB	Upperbound
UV	Ultraviolet
Vis	Visible
VT	Variable Temperature

Chapter 1

Introduction

Protein structure is critical to function. In fact, the majority of the hundreds of amino acids that form a typical enzyme simply serve to afford a well defined structural scaffold to precisely position the handful of residues that are directly involved in catalysis. This complex protein architecture creates the precise array of functionality necessary to effect a particular transformation. The ability to design a polypeptide with well defined structure is therefore a prerequisite for the design of novel biocatalysts based on the protein architecture. For this reason, most protein design efforts have focused on the creation of peptidyl motifs with defined secondary and tertiary structure.

Protein structure can be considered as a hierarchical framework in which tertiary structure is composed of various elements of super-secondary and secondary structure, which are in turn determined by the primary amino acid sequence of a protein. The primary sequence of the biopolymer contains all the necessary information required to specify the complex three-dimensional architecture of many proteins. Solving the "protein folding problem", i.e., understanding how the primary sequence determines the three-dimensional structure of a protein, is one of the great challenges facing modern biochemistry. In this regard, protein design addresses the "inverse folding problem;" the objective is to design a primary sequence which will uniquely adopt a target three-dimensional structure. Thus, the design of well defined structural motifs tests and refines our understanding of the principles of protein folding and structure, in addition to providing stable scaffolds for the construction of functional polypeptides.

As the first step in tackling the protein folding problem, considerable effort has been expended to understand the rules that relate primary sequence to secondary structure. Early research focused on investigations of the crystal structures of proteins to determine the relative propensities of the naturally

occurring amino acids to be found in the various elements of secondary structure (α -helices, β -sheets, and β -turns).¹⁻⁵ Expanding on these rules, many researchers have used model systems to determine the thermodynamic basis for the α -helical propensities of amino acids.⁶⁻⁹ There is good agreement among the various scales developed;¹⁰ in general, unbranched residues (eg., alanine and leucine) have the highest helical propensities.¹¹ Several important stabilizing interactions present in both native and designed helices have been also investigated. N- and C-terminal residues can contribute significantly to helix stability either through the electrostatic stabilization of the helix macrodipole¹² or by participating in hydrogen bonding interactions with the "open" amide and carbonyl functionalities at the ends of helices (helix "caps").¹³⁻¹⁶ Exploiting the alignment of the side chains of residues separated by one turn of an α -helix, electrostatic interactions have been used to produce helices with increased structure and stability.¹⁷ Based on these efforts, the construction of α -helices that are inherently stable as monomers in aqueous solution is now routine.¹⁸

Investigations of β -sheet proteins have been considerably more challenging. In contrast to α -helices which involve "near neighbor" ($i \rightarrow i+4$) hydrogen bonds, β -sheet formation results from the interaction of residues which are at variable and often distal positions in the primary sequence. Sheet structure is therefore much less modular than that of helices, and relies strongly on tertiary interactions. Additionally, the alternating nature of the hydrogen bonding interactions of β -sheets leaves an exposed edge of unfulfilled hydrogen bond acceptors and donors, which can lead to intermolecular association and the formation of insoluble aggregates. Despite these difficulties several groups have developed β -sheet model systems based on natural β -sheet proteins to investigate the propensities of amino acids to form sheet structure.¹⁹⁻²¹ Unlike the numerous thermodynamic scales for α -helical structural preferences, which in

general are in good agreement, β -sheet propensities appear more context-dependent, relating to both secondary and tertiary structure interactions.^{10, 19} Although there is some variability among scales the qualitative trends observed are that β -branched amino acids (isoleucine, threonine, and valine) are and the aromatic amino acids (phenylalanine and tyrosine) have high sheet propensities.^{20, 22}

While considerable understanding of the determinants that relate a primary sequence to the adoption of secondary structure has been developed, much less is understood about how elements of secondary structure interact to form super-secondary and tertiary structure. For this reason the design of proteins with well defined tertiary structure is a challenging (and as yet unattained) goal.

Approaches to protein design have varied greatly.¹⁰ One strategy employs synthetic templates to direct the folding of the polypeptide chain. Template-assisted synthetic protein (TASP) is a concept introduced by Mutter and co-workers for the assembly of diverse protein structures.²³ In this approach sequences designed to form amphiphilic α -helices or β -strands are pre-organized for association by attachment to a branched cyclic peptide scaffold. A four helix bundle protein has been successfully assembled using this methodology.²⁴ Several groups have employed a similar strategy in the construction of three helix bundles, in which a template organized about a metal cation directs bundle formation.²⁵⁻²⁷ In these designs, a bidentate bipyridyl functionality is attached to the ends of amphiphilic helices and metal coordination is exploited to promote helix association. While these efforts have produced constructs with significant secondary structure, insufficient structural characterization prevents the assessment of tertiary interactions.^{25, 27} Additionally, the synthetic procedures

required to assemble these constructs limits an efficiency of the iterative synthesis and analysis required to generate improved designs.

DeGrado and co-workers have used a 'minimalistic' approach to design amphiphilic peptides consisting of simple repeating sequences which form helix bundles.^{28, 29} Their initial efforts included only four amino acids Leu, Gly, Lys, and Glu in the design of a 16-residue helix consisting of two seven-residue repeats. In addition to an individual helix (α_1 B), peptides consisting of two (α_2) or four (α_4) copies of the helix connected by three-residue loops were synthesized.^{28, 29} Association of the helices to form bundles was mediated by hydrophobic interactions between the leucine residues lining one face of each helix. While the designed proteins adopted the correct aggregation state and exhibited considerable secondary structure which was very stable, these molecules also had many of the properties of the "molten globule" state.³⁰ This state is characterized by a compact and globular structure which lacks well defined tertiary packing interactions and appears to have dynamic core. The molten globule state is experimentally characterized by a number of properties which reflect the poorly folded structure; the binding of a hydrophobic dye, (8-anilino-1-naphthalenesulfonic acid (ANS), non-cooperative loss of structure, and poor dispersity in the NMR spectrum, for example. Subsequent α_1 based designs introduced metal sites into the interior of the protein to confer specificity to the hydrophobic interactions that determine helix association and tertiary structure.^{31, 32} Most recently a helix-loop-helix peptide (α_2 D) based on α_2 B was designed.³³ This peptide forms the desired non-covalent dimer and exhibits many of the properties of native proteins (dispersity of the 1D NMR spectra and a cooperative thermal denaturation). Unfortunately, detailed structural analysis of this peptide has been hampered by the degeneracy of the NMR signals and the aggregation of the molecule.³³

In a more 'native-like' approach, Richardson and coworkers produced a 79 residue sequence consisting of nearly all the twenty natural amino acids, in the design of "Felix", a four helix bundle.³⁴ Similar to the α_1 based peptides, Felix exhibited some of the characteristics of a "molten globule", although a compact monomeric structure of high helical content was achieved. In a similar native-like approach, a β -sandwich protein, betadoublet, has been developed.³⁵ This 35-residue sequence was designed to form a 4-stranded, antiparallel β -sheet through disulfide-mediated homo-dimerization. The assembly of a similar disulfide-linked β -sandwich, betabellin, has been the target of an on-going design effort in the Richardson group. The latest design by Yan and Erickson, betabellin14D, incorporates a simple pattern of alternating polar and non-polar residues for the construction of the sheet region, and unnatural D-amino acids in the turn regions.³⁶ While both these proteins exhibit some of the properties of native proteins such as cooperative thermal denaturation, they also possess some of the characteristics of a poorly folded molten globule-like state (ANS binding and poor dispersity of the NMR signals).³⁷

Many design attempts have exploited the concept of "negative design"; the final design includes not only specific interaction intended to stabilize the desired structure but also contains features that would *disfavor* alternative conformations. This strategy has been frequently used in the design of loop or turn regions to promote solubility and prevent adverse intermolecular association of elements of secondary structure. For example "helix-breaking" residues such as glycine were incorporated at the ends of helices in the design of both Felix³⁴ and α_4 ²⁸ in order to promote loop formation and prevent the extension of the helices. Similarly, the residues comprising the turn regions in betabellin³⁶ and betadoublet³⁵ were chosen to specifically disfavor the continuation of the β -strands and favor type-I turn formation.

The analysis of the designed proteins described above has yielded insight into the complexity of native protein structure. In addition to hydrophobic interactions, specific well defined packing interactions, including electrostatic and hydrogen bonding interactions, are important to confer the conformational specificity found in the interior of native proteins. This fact has been further demonstrated by the analysis of coil-Ser, a coil-coiled protein, of *de novo* design.³⁸ This peptide which was designed to form a parallel coil-coiled dimer, was found in the crystal structure to form an anti-parallel trimer. In comparison, the sequence of a naturally occurring dimeric coiled coil protein, GCN4, contains a single hydrophilic asparagine residue at the helix-helix interface, which appears to confer the specificity needed for dimerization.³⁹ Based on this observation, asparagine was incorporated into the coil-Ser sequence, and resulted in sequences for which dimerization predominated.³⁹ Thus, it is not only the hydrophobic association of residues at the helix-helix interface that determine the aggregation state and structure but the defined electrostatic and hydrogen-bonding interactions among carefully positioned *hydrophilic* amino acids.

The design of coiled-coil proteins, such as coil-Ser, has been based on the heptad repeat found in natural coiled-coil proteins.^{40, 41} The peptide 'Velcro' was designed to form a heterodimeric coiled-coil.⁴² Electrostatic interactions were employed in a "negative design" approach to destabilize the homodimer and favor heterodimer formation. Additionally, an asparagine residue was included in the sequence to confer aggregation specificity for the dimeric form. Other approaches to the design of coiled-coil proteins include the formation of intramolecular cross-links through the use of disulfides⁴¹ or by linking the two helices *via* peptide loops.^{40, 43} Diverse helix-loop-helix proteins have also been designed⁴³⁻⁴⁵ and several have been partially characterized by NMR techniques.^{43, 44}

The *de novo* design of well defined packing interactions is extremely difficult given our limited understanding of these forces in natural systems. To overcome this problem, several groups have based designs on the primary sequences of natural proteins. Pessi and coworkers have used the heavy chain variable domain of immunoglobulin as a template for the construction of "minibody", a β -sheet protein which included a metal binding site.⁴⁶ The resulting 61-residue protein had significant β -sheet content, exhibited cooperative urea denaturation, and bound divalent metal ions. Unfortunately, the polypeptide had a solubility limit of 10 μ M in water, which precluded any detailed structural analysis. In another example of exploiting natural sequences in protein design, the C-terminal helix of cytochrome c served as the basis for the construction of a 35-residue disulfide linked helix-loop-helix peptide.⁴⁷ The solution state structure of this peptide, determined by NMR, revealed that while the area closest to the disulfide formed a reasonably well defined helix, the inter-helical loop and peptide termini were highly disordered.^{45, 47}

Recombinant DNA techniques have been used to design libraries of proteins with novel sequences. Hecht and coworkers used a simple binary pattern of amino acids in which the polarity but not the identity of residues was specified to create synthetic genes encoding for 48 unique 72-residue sequences.⁴⁸ Expression of 29 of the 48 genes produced soluble protease-resistant products. Characterization of several of the gene products indicated that the proteins had the desired helical secondary structure and exhibited cooperative chemical denaturation.⁴⁸ In a similar approach, random libraries of gene sequences encoding for proteins containing only three residues, glutamine, leucine and arginine, were created and expressed.^{49, 50} Folded proteins were selected by their resistance to proteolytic processing during expression, the demonstration of cooperative unfolding transitions, and high helicity as assessed

by CD. While these experiments demonstrate that limited sequence information is required for the generation of globally folded proteins, none of the designed proteins display all of the characteristics of native protein structures and significantly more structural characterization is needed to adequately assess the success of such efforts.

In general, most protein design efforts have focused on a single type of secondary structure in order to maintain the greatest simplicity in design. However, several groups have targeted the design of eight-stranded α/β -barrel protein motif exemplified by the structure of triosephosphate isomerase.^{51, 52} The best characterized of these efforts is a 207 amino acid sequence.⁵¹ Although this protein exhibits cooperative thermal denaturation, it also displays some of the characteristics of a molten globule-like state such as a poorly defined NMR spectra and no protection of amide protons from exchange with solvent.⁵¹

One of the fundamental drawbacks of most of the protein design efforts to date is the lack of conclusive structural characterization of the resulting constructs. One of the primary techniques used to probe the structural content of designs is circular dichroism. This technique provides only macroscopic secondary structural information, and generally yields no data on the presence of tertiary structure. Conclusions based solely on CD data can be misleading, as illustrated by the comparative NMR and CD analysis of a designed peptide presented in the Chapter 3. If a designed protein has a native-like conformation then determination of the structure by NMR or X-ray crystallographic techniques should be possible.³⁹ Unfortunately, detailed structural characterization of several designed proteins has been hampered by the large size, aggregation state, limited solubility, and the poorly defined conformations of these constructs.³⁹

In the absence of detailed structural information such as that provided by NMR or crystal structure data, the analysis of designed proteins has relied

heavily on auxiliary *indirect* methods to probe the presence of a defined tertiary structure. Several methods typically used to characterize the stability of natural systems are frequently employed to gauge the success of a *de novo* design.³⁹ Native protein structures undergo a cooperative unfolding transition, and a reproduction of this phenomenon in designed proteins has been used as a measure of "native-like" structure within the constructs. In these experiments, the unfolding of a peptide or protein with either heat or denaturant is monitored (typically by CD). If the secondary and tertiary structures are tightly coupled as they are in most natural globular proteins, then a cooperative transition should be observed in the unfolding of designed proteins. While cooperative denaturation is a characteristic of native folded structures, the analysis of several designed constructs reveals that poorly folded constructs can also exhibit this behavior.^{35, 36, 51}

Dispersity of the proton resonances in the 1-D NMR spectra of designed proteins has also been used as an indicator of tertiary structure.^{30, 33, 35, 36} If chemically similar protons exhibit significantly different chemical shifts, then this phenomenon can be attributed to side chain-side chain interactions that arise in tertiary structure. For example, the methyl protons of Ile, Val, and Leu residues within the interior of a well packed protein core should experience different micro-environments and therefore exhibit slightly different chemical shifts.

Many proteins have been shown to form what has been termed a "molten globule" state under partially denaturing conditions.^{53, 54} This conformational state retains a significant amount of secondary structure but loses the well defined tertiary interactions present in the fully folded protein. The hydrophobic dye, 8-anilino-1-naphthalenesulfonic acid (ANS), has been used to probe the molten globule state of proteins. A poorly folded or molten globule-like structure has exposed hydrophobic patches which bind ANS and alter the fluorescence of

the dye. In contrast, native fully folded proteins have buried hydrophobic side chains and will not bind the dye.^{53, 54} While many designed systems have successfully reproduced some of the characteristics of native proteins such as cooperative denaturation and dispersity in the NMR signals, many designs have "failed" the ANS-binding test.^{30, 35, 36} These systems exhibit a significant amount of secondary structure but have poorly defined tertiary packing interactions which lead to a molten globule state.

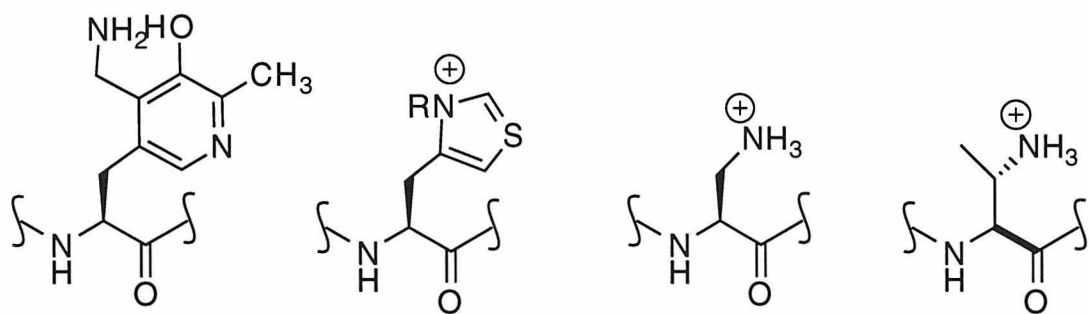
The lack of well defined structural templates has also hindered the exploitation of the protein architecture in the design of novel catalytic biomolecules. Very few examples of *de novo* designed functional polypeptides have been reported. The majority of examples utilize a leucine-rich four helix bundle motif similar to that described by DeGrado and coworkers as a structural template.³⁹ Dutton and coworkers have reported the synthesis of heme-binding four helix bundles for the development of synthetic electron-transfer systems.^{55, 56} These designs exhibit some of the desired spectroscopic and electrochemical properties of heme proteins. In addition, Benner and coworkers⁵⁷ and Sasaki and Kaiser⁵³ have reported using the helical bundle motif in the design of catalytic polypeptides. The decarboxylation of oxaloacetate *via* an imine intermediate was accomplished by Benner and coworker through the construction of a 14-residue helical peptide, "oxaldie", that forms helical bundles.⁵⁷ Although the secondary structure of the peptide was verified by 2-D NMR, no tertiary structure was reported. The decarboxylation reaction was studied in detail and helicity appears to be critical for function. Sasaki and Kaiser have attached amphiphilic helices to a porphyrin template in the construction of "helichrome", an artificial heme protein which demonstrated aniline hydroxylase activity.⁵³ One of the fundamental problems with all these

efforts is that the structural scaffold upon which the designs were based probably exists in a molten globule-like state.

To overcome structural ambiguity some groups have focused their efforts on the modification of existing enzymes. These efforts rely on a variety of techniques for the introduction of both natural and non-standard functionalities into the protein framework. Sequence changes involving the 20 naturally occurring amino acids can be easily accomplished by site directed mutagenesis while non-standard amino acids have been incorporated by *in vitro* expression using chemically acylated suppressor *t*-RNA.^{58, 59} Alternatively, unnatural amino acids can be introduced by protein semisynthesis.⁶⁰ In this approach a native protein is either chemically or proteolytically digested to produce two or more fragments. Chemical synthesis of one of the fragments is followed by reconstitution of native structure either through non-covalent association or chemical or enzymatic re-ligation of the fragments. This approach has been used for the introduction of a variety of unnatural amino acids into existing proteins both, to probe the mechanism of the native protein, and for the constructions of systems with altered function.⁶⁰⁻⁶² Chemical modification of the residue side chains (typically cysteine) of existing proteins presents yet another method for introduction of new functionality, such as coenzyme moiety.⁶³ This strategy can be combined with site directed mutagenesis to allow greater control in the location and environment of the reactive moiety. The architecture of antibodies and the specificity of the immune system can also be exploited in the creation of proteins which are selected to transition state analogs of certain transformations and in some cases, effect catalysis.^{64, 65} Unfortunately all of these approaches are limited to systems which are amenable to the specific demands of the various techniques.

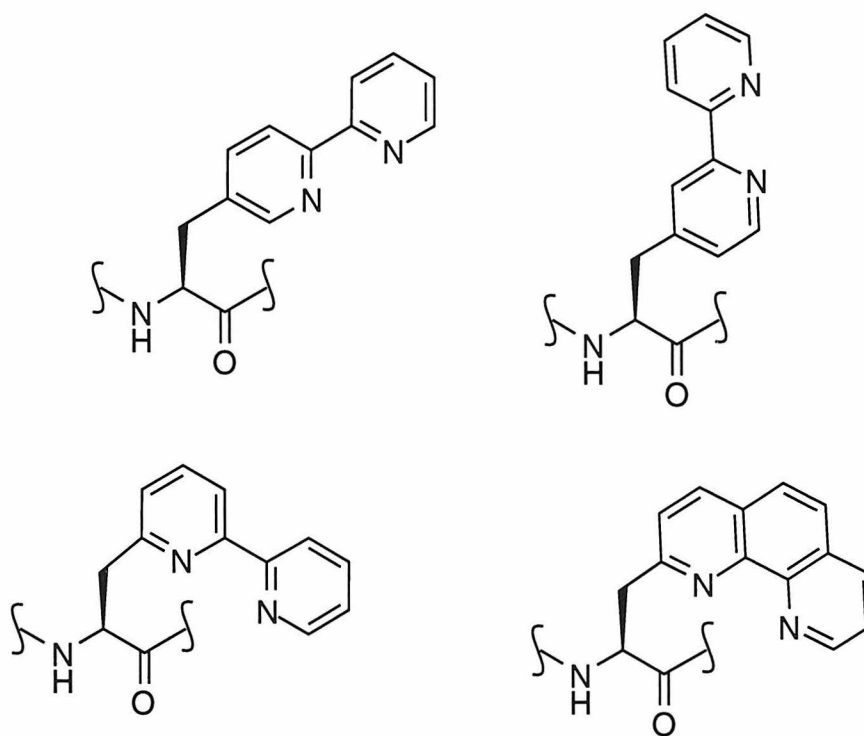
The goal in the Imperiali group has been to create totally synthetic motifs with well defined structure and preprogrammed function. Of course, as discussed earlier, the definition of structure must precede the introduction of function. Therefore, we have targeted the assembly of versatile structural templates in less than 40 residues. The design motifs should ideally be compact, soluble, and monomeric. Consequently we have chosen to base our designs of natural polypeptides or protein domains which exhibit these properties. Most naturally occurring polypeptide or protein domains of this size generally require the use of disulfides or metal ligation for proper folding and stability. These cross-links reduce the versatility of the motif for elaboration into functional constructs. Our approach is to modify these motifs to eliminate the need for such cross-links. By targeting motifs of less than 40 residues, the designs are accessible by solid phase peptide synthesis and are amenable to detailed structural analysis afforded by ^1H NMR techniques. The ability to readily obtain specific information regarding the presence of secondary and tertiary structure within designed peptides is critically important to the development of the desired motif. As described earlier, protein design efforts have generally been hindered by the inability to obtain sufficient structural characterization of the designed constructs. Iterative improvements of designed motifs can greatly benefit from the detailed information available through NMR analyses.

The utilization of unnatural amino acids both to control structure, and to introduce function forms one of the cornerstones of our design approach. A variety of unnatural amino acids have been developed for these purposes in the Imperiali group.⁶⁶⁻⁷⁰ Unnatural amino acids incorporating both bipyridyl^{66, 67} and phenanthrolyl⁷¹ functionalities have been used to modulate the metal binding properties of diverse synthetic polypeptides. "Coenzyme amino acid chimeras" and metal binding bipyridyl amino acids have been used to reproduce



**Incorporation of conenzyme
functionality**

**Electrostatic-stabilization of
 β -sheet and turn structure**



**Control of structure and function through
metal-coordination**

Figure 1-1 Several examples of the unnatural amino acids exploited for the control of structure and function in the Imperiali Group.

pyridoxal mediated chemistry^{68-70, 72} or introduce electron-transfer properties⁷³ in semisynthetic constructs. Further exploitation of these functional residues within synthetic motifs awaits the development of structural templates with the properties described above.

In addition to unnatural amino acids, a key feature of our design philosophy is the use of defined β -turns as structural nucleation elements.⁷⁴ Early investigations of the protein folding process have suggested the structural importance of the β -turn.⁷⁵⁻⁷⁸ The explicit role of reverse turns in native systems appears to be context-dependent. In some cases, turns may passively result from β -sheet formation, or in other systems the chain reversal may contribute more directly to the structural stability and proper folding of the proteins.⁷⁹⁻⁸¹ Regardless of the role of turns in natural systems, a well defined and structurally characterized β -turn could serve as a valuable tool in protein design; β -turns result in the reversal of the peptide backbone and bring residues distal in the primary sequence into close proximity and therefore could be exploited to promote the adoption of β -sheet structure within synthetic constructs. Several classes of β -turns have been identified by investigations of protein crystal structures.^{1, 2, 82, 83} Turns are generally considered to consist of four residues and the turn type is specified by the backbone dihedral angles of the central two residues (Figure 1-2). The most common β -turns are the type I and type II turns.⁸⁴ The type I' and II' are the "antipodes" of these turns, respectively and are formed by rotation of the central amide bond through 180° (see Table 1-1).²

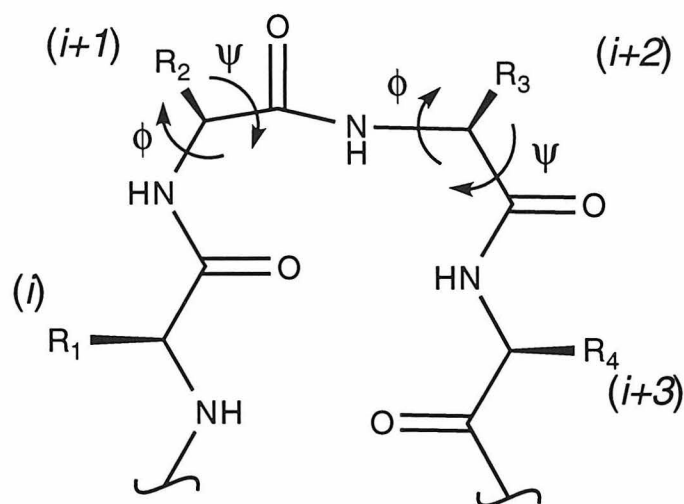


Figure 1-2 The dihedral angles (ϕ, ψ) of the central two residues of a four residue β -turn define the β -turn-type.

turn	(i+1)		(i+2)	
Type	ϕ_1 (°)	ψ_1 (°)	ϕ_2 (°)	ψ_2 (°)
I	-60	-30	-90	0
I'	60	30	90	0
II	-60	120	80	0
II'	60	-120	-80	0

Table 1-1 The dihedral angles of the central two residues of the most commonly occurring β -turn-types.

Efforts in the Imperiali group have focused on the development of type II β -turns rather than the more common type I turns. The type II turns are more planar than the type I turns permitting greater control over the relative positioning of the side chains of the residues at positions (i) and (i+3).⁸³ The structure of the type II turns also appear to be less reliant on side chain to main-

chain hydrogen bonds compared to the type I β -turns. Extensive studies of tetrapeptide sequences^{71, 74} revealed that stable type II turns could be formed in isolation by the incorporation of proline at ($i+1$) and non-standard D-amino acids at position ($i+2$). These studies further defined the "guidelines" governing the choice of residues at positions (i), ($i+2$), and ($i+3$) that were required to promote type II β -turn formation.⁷⁴

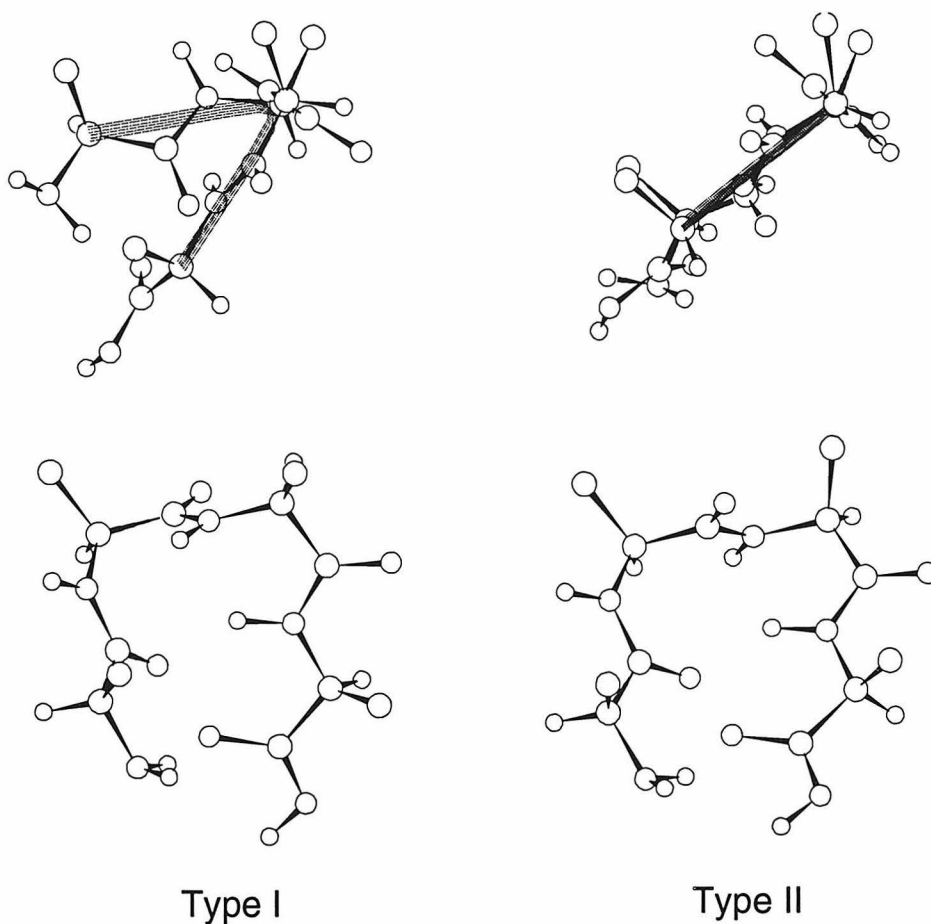


Figure 1-3 The different topologies of the type I and type II turns. The top structures are viewed from the side and illustrate the differences in planarity between type I and type II β -turns.

The design of a $\beta\beta\beta$ motif based on a domain from a family of disulfide-rich snake toxins described in Chapter 2 represents the first implementation of the design approach that was adopted. Type II turns were incorporated to promote anti-parallel sheet formation and complete structural analysis by NMR was performed. Based on this study, a more compact motif was targeted. The $\beta\beta\alpha$ motif, exemplified by the zinc finger domains, represented an appealing alternative.

Zinc finger sequences form the modular domains in many DNA binding proteins and require the coordination of divalent zinc by cysteine and histidine residues for the adoption of a defined three dimensional structure.⁸⁵ These small motifs have received considerable attention in recent years due to their wide spread occurrence in a variety of transcription factors.⁸⁵ These polypeptide domains have been the target of several design efforts, which have focused on exploiting the motif in the construction of DNA-binding proteins with preprogrammed specificity.^{86, 87} Structural studies indicate that the very little sequence information, besides the metal ligating residues and three key hydrophobic amino acids, is needed to specify the $\beta\beta\alpha$ architecture.⁸⁸ Chapter 3 presents the iterative design and analysis procedure used in the evolution of a metal-independent $\beta\beta\alpha$ motif based on the zinc finger domains including the incorporation of both type II and type II' β -turns. Information gained through detailed NMR and CD analysis of each design was used to determine the sequence modification of subsequent designs. These efforts culminated in the solution structure determination of a "zinc-less" finger described in Chapter 4. The explicit role of the β -turn in the adoption of the defined three-dimensional structure is also discussed in Chapter 4.

In an ancillary project, unnatural amino acids and a defined type II β -turn were used in the metal-dependent modulation of the conformation of a designed peptide. This project had been initiated by Dr. Stewart Fisher.⁷¹ The exploitation of this motif in the construction of a metal-dependent regulatory peptide for the control of cAMP-dependent protein kinase activity is described in Chapter 5.

Finally, Appendix-I describes the use of both non-standard and natural amino acids in the electrostatic stabilization of β -sheet and reverse turn structure. The synthesis of tetrapeptides incorporating salt bridges, comprising the naturally occurring amino acids lysine and glutamic acid demonstrates the potential utility of electrostatic interactions in protein design. The synthesis and peptide incorporation of an amino acid, (2*S*,3*S*)-diaminobutyric acid (β -methyl-amino alanine), designed to be more amenable to sheet and turn conformations, is also reported.

The work described in this dissertation demonstrates the feasibility of constructing structurally defined peptidyl templates in less than 40 residues. Significantly, sequences of fewer than 70 residues were considered to be unlikely to form stable monomeric proteins if folding is determined by hydrophobic forces alone. Through the use of unnatural amino acids and defined β -turns the successful construction of a stable folded, monomeric motif in 23 residues has been accomplished. The determination of the solution structure of the motif by NMR techniques lays the foundation for the rational design of functional polypeptides based on this structural template.

References

- (1) Sibanda, B.L.; Thornton, J.M. " β -Hairpin Families in Globular Proteins," *Nature* **1985**, 316, 170-174.
- (2) Rose, G. D.; Gierasch, L. M.; Smith, J. A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.* **1985**, 37, 1-105.
- (3) Levitt, M. "Conformational Preferences of Amino Acids in Globular Proteins," *Biochemistry* **1978**, 17, 4277-4285.
- (4) Chou, P. Y.; Fasman, G. D. "Conformational Parameters for Amino Acids in Helical, β -Sheet, and Random Coil Regions Calculated from Proteins," *Biochemistry* **1974**, 13, 211-222.
- (5) Chou, P. Y.; Fasman, G. D. "Prediction of Protein Conformation," *Biochemistry* **1974**, 13, 222-245.
- (6) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. "Side Chain Contributions to the Stability of Alpha-Helical Structure of Peptides," *Science* **1990**, 250, 669-673.
- (7) O'Neil, K. T.; DeGrado, W. F. "A Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occurring Amino Acids," *Science* **1990**, 250, 646-651.
- (8) Lyu, P.; Sherman, J. C.; Chen, A.; Kallenbach, N. R. " α -Helix Stabilization by Natural and Unnatural Amino Acids with Alkyl Side Chains," *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 5317-5320.
- (9) Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L. "Relative Helix-Forming Tendencies of Nonpolar Amino Acids," *Nature* **1990**, 344, 269-270.
- (10) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. "Protein Design: A Hierarchic Approach," *Science* **1995**, 270, 935-941.

- (11) Padmanabhan, S.; Baldwin, R. L. "Straight-chain Non-polar Amino acids are Good Helix-formers in Water," *J. Mol. Biol.* **1991**, 135-137.
- (12) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. "Tests of the Helix Dipole Model for the Stabilization of α -helices," *Nature* **1987**, 326, 563-567.
- (13) Lyu, P. C.; Wemmer, D. E.; Zhou, H. X.; Pinker, R. J.; Kallenbach, N. R. "Capping Interactions in Isolated α Helices: Position-Dependent Substitution Effects and Structure of a Serine-Capped Peptide Helix," *Biochemistry* **1993**, 32, 421-425.
- (14) Harper, E. T.; Rose, G. D. "Helix Stop Signals in Proteins and Peptides: The Capping Box," *Biochemistry* **1993**, 32, 7605-7609.
- (15) Richardson, J. S.; Richardson, D. C. "Amino Acid Preferences for Specific Locations at the Ends of α Helices," *Science* **1988**, 240, 1649-1652.
- (16) Presta, L. G.; Rose, G. D. "Helix Signals in Proteins," *Science* **1988**, 240, 1632-1641.
- (17) Marqusee, S.; Baldwin, R. L. "Helix Stabilization by $\text{Glu}^- \dots \text{Lys}^+$ Salt Bridges in short Peptides of *de novo* design.," *Proc. Natl. Acad. Sci. U.S.A.* **1987**, 84,
- (18) Muñoz, V.; Serrano, L. "Helix Design, Prediction and Stability," *Curr. Opin. Biotechnology* **1995**, 6, 382-386.
- (19) Otzen, D. E.; Fersht, A. R. "Side-Chain Determinants of β -Sheet Stability," *Biochemistry* **1995**, 34, 5718-5724.
- (20) Kim, C. A.; Berg, J. M. "Thermodynamic β -Sheet Propensities Measured Using a Zinc-finger Host Peptide," *Nature* **1993**, 362, 267-270.
- (21) Smith, C. K.; Regan, L. "Guidelines for Protein Design: The Energetics of β Sheet Side Chain Interactions," *Science* **1995**, 270, 980-982.
- (22) Minor, D. L.; Kim, P. S. "Measurement of the β -sheet-forming Propensities of Amino Acids," *Nature* **1994**, 367, 660-663.

- (23) Mutter, M. "The Construction of New Proteins and Enzymes- a Prospect for the Future," *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 639-653.
- (24) Mutter, M.; Tuchscherer, G. G.; Miller, C.; Altmann, K.-H.; Carey, R. I.; Wyss, D. F.; Labhardt, A. M.; Rivier, J. E. "Template-Assembled Synthetic Proteins with Four-Helix-Bundle Topology. Total Chemical Synthesis and Conformation Studies," *Journal of the American Chemical Society* **1992**, *114*, 1463-1470.
- (25) Ghadiri, M. R.; Case, M. A. "De Novo Design of a Novel Heterodinuclear Three-Helix Bundle Metalloprotein," *Angew. Chem. Int. Ed. Engl.* **1991**, *32*, 1594-1597.
- (26) Ghadiri, M. R.; Fernholz, A. K. "Peptide Architecture. Design of Stable α -Helical Metallopeptides via a Novel Exchange -Inert RuIII Complex," *J. Am. Chem. Soc.* **1990**, *112*, 9633-9635.
- (27) Lieberman, M.; Sasaki, T. "Iron(II) Organizes a Synthetic Peptide into Three-Helix Bundles," *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.
- (28) Regan, L.; DeGrado, W. F. "Characterization of a Helical Protein Designed from First Principles," *Science* **1988**, *241*, 976-978.
- (29) Ho, S. P. "Design of a 4-helix Bundle Protein: Synthesis of Peptides Which Self-Associate into a Helical Protein," *Journal of the American Chemical Society* **1987**, *109*, 6751-6758.
- (30) Betz, S.F.; Raleigh, D.P.; DeGrado, W.F. "De Novo Protein Design: from Molten Globules to Native-like States," *Curr. Opin. Struct. Biol.* **1993**, *3*, 601-610.
- (31) Handel, T.; DeGrado, W. F. "De Novo Design of a Zn^{2+} -Binding Protein," *Journal of the American Chemical Society* **1990**, *112*, 6710-6711.
- (32) Handel, T.M.; Williams, S.A.; DeGrado, W.F. "Metal Ion-Dependent Modulation of the Dynamics of a Designed Protein," *Science* **1993**, *261*, 879-885.

- (33) Raleigh, D. P.; Betz, S. F.; DeGrado, W. F. "A *de Novo* Designed Protein Mimics the Native State of Natural Proteins," *J. Am. Chem. Soc.* **1995**, *117*, 7558-7559.
- (34) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. "De Novo Design, Expression, and Characterization of Felix: A Four-Helix Bundle Protein of Native-Like Sequence," *Science* **1990**, *249*, 884-891.
- (35) Quinn, T. P. ; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C. "Betadoublet: *De novo* design, Synthesis, and Characterization of a β -sandwich Protein," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747-8751.
- (36) Yan, Y. B.; Erickson, B. W. "Engineering of Betabellin 14D: Disulfide-induced Folding of a β -sheet Protein," *Protein Sci.* **1994**, *3*, 1069-1073.
- (37) Hecht, M. H. "*De novo* design of β -sheet Proteins," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8729-8730.
- (38) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D.K.; DeGrado, W.F.; Eisenberg, D. "Crystal Structure of a Synthetic Triple-Stranded α -Helical Bundle," *Science* **1993**, *259*, 1288-1293.
- (39) Betz, S. F.; Bryson, J. W.; DeGrado, W. F. "Native-like and Structurally Characterized Designed α -helical Bundles," *Curr. Opin. Struct. Biol.* **1995**, *5*, 457-463.
- (40) Myszka, D. G.; Chaiken, I. M. "Design and Characterization of an Intramolecular Antiparallel Coiled Coil Peptide," *Biochemistry* **1994**, *33*, 2363-2372.
- (41) Zhou, N. E.; Kay, C. M.; Hodges, R. S. "Disulfide Bond Contribution to Protein Stability: Positional Effects of Substitution in the Hydrophobic Core of the Two-Stranded α -Helical Coiled-Coil," *Biochemistry* **1993**, *32*, 3178-3187.
- (42) O'Shea, E. K.; Lumb, K. J. ; Kim, P. S. "Peptide 'Velcro': Design of a Heterodimeric Coiled Coil," *Current Biology* **1993**, *3*, 658-667.

- (43) Fezoui, Y.; Weaver, D.L.; Osterhout, J.J. "De novo Design and Structural Characterization of an α -helical Hairpin Peptide: A Model System For the Study of Protein Folding Intermediates," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3675-3679.
- (44) Olofsson, S.; Johansson, G.; Baltzer, L. "Design Synthesis and Solution Structure of a Helix-loop-Helix Dimer- A Template for the Rational Design of Catalytically Active Polypeptides," *J. Chem. Soc. Perkin Trans. II* **1995**, 2047-2056.
- (45) Kuroda, Y.; Nakai, T.; Ohkubo, T. "Solution Structure of a *de Novo* Helical Protein by 2D-NMR Spectroscopy," *Journal Molecular Biology* **1994**, *236*, 862-868.
- (46) Pessi, A.; Bianchi, E.; Crameri, A.; Venturini, S.; Tramontano, A.; Sollazzo, M. "A Designed Metal-binding Protein with a Novel Fold," *Nature* **1993**, *362*, 367-369.
- (47) Kuroda, Y. "A Strategy for the *de novo* Design of Helical Proteins with Stable Folds," *Protein Eng.* **1995**, *8*, 97-101.
- (48) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. "Protein Design by Binary Patterning of Polar and Nonpolar Amino Acids," *Science* **1993**, *262*, 1680-1685.
- (49) Davidson, A. R.; Sauer, R. T. "Folded Proteins Occur Frequently in Libraries of Random Amino Acid Sequences," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2146-2150.
- (50) Davidson, A. R.; Lumb, K. J.; Sauer, R. T. "Cooperatively Folded Proteins in Random Sequence Libraries," *Nature Structural Biology* **1995**, *2*, 856-863.
- (51) Tanaka, T.; Kuroda, Y.; Kimura, H.; Kidokoro, S.-i. "Cooperative Deformation of a *de novo* Designed Protein," *Protein Eng.* **1994**, *7*, 969-976.
- (52) Goraj, K.; Renard, A.; Matrial, J. A. "Synthesis, Purification and Initial Structural Characterization of Octarellin, A *de novo* Polypeptide Modeled on the α/β -barrel Proteins," *Protein Eng.* **1990**, *3*, 259-266.

- (53) Sasaki, T.; Kaiser, E. T. "Helichrome: Synthesis and Enzymatic Activity of a Designed Hemeprotein," *J. Am. Chem. Soc.* **1989**, *111*, 380-381.
- (54) Kuwajima, K. "The Molten Globule State as a Clue for Understanding the Folding and Cooperativity of Globular-Protein Structure," *Proteins: Struct. Func. Gen.* **1989**, *6*, 87-103.
- (55) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. "Design and Synthesis of Multi-haem Proteins," *Nature* **1994**, *368*, 425-432.
- (56) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, L.; Robertson, D. E.; DeGrado, W. F. "Design of a Heme-Binding Four-Helix Bundle," *J. Am. Chem. Soc.* **1994**, *116*, 856-865.
- (57) Johnsson, K.; Allemann, R. K.; Widmer, H.; Benner, S. A. "Synthesis, Structure and Activity of Artificial, Rationally Designed Catalytic Polypeptides," *Nature* **1993**, *365*, 530-532.
- (58) Anthony-Cahill, S. J.; Griffith, M. C.; Noren, C. J.; Suich, D. J.; Schultz, P.H. "Site-specific Mutagenesis with Unnatural Amino Acids," *Trends in Biol. Sci.* **1989**, *14*, 400-403.
- (59) Nowak, M. W.; Kearney, P. C.; Sampson, J. R.; Saks, M. E.; Labaraca, C. G.; Silverman, S. K.; Zhong, W. G.; Abelson, J. N., *et al.* "Nicotinic Receptor-Binding Site Probed with Unnatural Amino-Acid Incorporation in Intact Cells," *Science* **1995**, *268*, 439-442.
- (60) Wallace, C. J. A. "Peptide Ligation and Semisynthesis," *Curr. Opin. Biotechnology* **1995**, *6*, 403-410.
- (61) Humphries, J.; Offord, R. E.; Smith, R. A. G. "Chemical Methods of Protein Synthesis and Modification," *Curr. Opin. Biotechnology* **1991**, *2*, 539-543.
- (62) Offord, R. E. *Semisynthetic Proteins*; John Wiley and Sons: New York, 1980.

- (63) Kaiser, E. T. "Catalytic Activity of Enzymes Altered at Their Active Site," *Angew. Chem. Int. Ed. Engl.* **1988**, 27, 913-922.
- (64) Schultz, P. G.; Lerner, R. A. "From Molecular Diversity to Catalysis-Lessons From the Immune System," *Science* **1995**, 269, 1835-1842.
- (65) Benkovic, S. J. "Catalytic Antibodies," *Ann. Rev. Biochem.* **1992**, 61, 29-54.
- (66) Imperiali, B.; Fisher, S. L. "Stereoselective Synthesis and Peptide Incorporation of (S)-alpha-Amino-(2,2'-bipyridine)-6-Propanoic Acid," *J. Org. Chem.* **1992**, 57, 757-759.
- (67) Imperiali, B.; Prinns, T. J.; Fisher, S. L. "Chemoenzymatic Synthesis of 2-Amino-3-(2,2'-bipyridinyl)-Propanoic Acids," *J. Org. Chem.* **1993**, 58, 1613.
- (68) Imperiali, B.; Sinha Roy, R. "Coenzyme-Amino Acid Chimeras: New Residues for the Assembly of Functional Proteins," *J. Am. Chem. Soc.* **1994**, 116, 12083-12084.
- (69) Imperiali, B.; Sinha Roy, R. "Stereoselective Synthesis and Peptide Incorporation of a Pyridoxal Coenzyme-Amino Acid Chimera," *J. Org. Chem.* **1995**, 60, 1891-1894.
- (70) Imperiali, B.; Sinha Roy, R.; Walkup, G. K.; Wang, L. "Unnatural Amino Acids for the Design of Functional Proteins: Biomimetic Catalysis Using Coenzyme Amino Acids," in *NATO Series on Biomimetic Catalysis*; Wilcox, C. S. Eds.; Kluwer Academic Publishing: Dordrecht, 1996; pp in press.
- (71) Fisher, S. L., "Unnatural Amino Acids in the synthesis and Semisynthesis of Metalloprotein Motifs," Ph.D. Thesis, California Institute of Technology, 1993.
- (72) Imperiali, B.; Sinha Roy, R.; Wang, L. "Peptidyl Models for Coenzyme Catalysis," in *Peptides: Chemistry, Structure, and Biology (Proceedings of the 14th American Peptide Symposium)*; Eds.; Escom Science Publisher: Leiden, 1996; pp 546.

- (73) Wuttke, D. S.; Gray, H. B.; Fisher, S. L.; Imperiali, B. "Semisynthesis of Bipyridyl-Alanine Cytochrome c Mutants: Novel Proteins with Enhanced Electron-Transfer Properties," *J. Am. Chem. Soc.* **1993**, *115*, 8455-8456.
- (74) Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide.," *J. Am. Chem. Soc.* **1992**, *114*, 3182.
- (75) Lewis, P. N.; Momany, F. A.; Scheraga, H. A. "Folding of Polypeptide Chains in Proteins: A Proposed Mechanism for Folding," *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 2293-2297.
- (76) Mantelione, G. T. ; Arnold, E.; Meinwald, Y. C.; Stimson, E. R.; Denton, J. B.; Huang, S.-G.; Clardy, J.; Scheraga, H. A. "Chain-Folding Initiation Structures in Ribonuclease A: Conformational Analysis of trans-Ac-Asn-Pro-Tyr-NHMe and trans-Ac-Tyr-Pro-Asn-NHMe in Water and in the Solid State," *J. Am. Chem. Soc.* **1984**, *106*, 7946-7958.
- (77) Wright, P. E.; Dyson, H. J.; Lerner, R. A. "Conformation of Peptide Fragments of Proteins in Aqueous Solutions: Implications for Initiation of Protein Folding," *Biochemistry* **1988**, *27*, 7167-7175.
- (78) Dyson, H. J.; Cross, K. J.; Houghten, R. A.; Wilson, I. A.; Wright, P. E.; Lerner, R. A. "The Immunodominant Site of a Synthetic Immunogen Has a Conformation Preference in Water for a Type-II Reverse Turn," *Nature* **1985**, *318*, 480-483.
- (79) Blanco, F. J.; Rivas, German; Serrano, Luis "A Short Linear Peptide that Folds into a Native Stable β -hairpin in Aqueous Solution," *Nature Structural Biology* **1994**, *1*, 584-589.

- (80) Viguera, A. R.; Blanco, F. J.; Serrano, L. "The Order of Secondary Structure Elements does not Determine the Structure of a Protein but does Affect its Folding Kinetics," *J. Mol. Biol.* **1995**, *247*, 670-681.
- (81) Stroup, A. N.; Gierasch, L. M. "Reduced Tendency To Form a β Turn in Peptides from the P22 Tailspike Protein Correlates with a Temperature-Sensitive Folding Defect," *Biochemistry* **1990**, *29*, 9765-9771.
- (82) Chou, P. Y.; Gasman, G. D. " β -Turns in Proteins," *J. Mol. Biol.* **1977**, *115*, 135-175.
- (83) Wilmot, C. M.; Thornton, J. M. "Analysis and Prediction of the Different Types of β -Turns in Proteins," *J. Mol. Biol.* **1988**, *203*, 221-232.
- (84) Hutchinson, E. G.; Thornton, J. M. "A Revised Set of Potentials for β -turn Formation in Proteins," *Protein Sci.* **1994**, *3*, 2207-2216.
- (85) Berg, J.M. "Zinc Finger Domains: Hypotheses and Current Knowledge," *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405-421.
- (86) Berg, J. M. "Zinc Finger Domains: From Prediction to Design," *Acc. Chem. Res.* **1995**, *28*, 14-19.
- (87) Choo, Y.; Klug, A. "Toward a Code for the Interactions of Zinc Fingers with DNA: Selection of Randomized Fingers Displayed on Phage," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11163-11167.
- (88) Michael, S. F.; Kilfoil, V. J.; Schmidt, M. H.; Amann, B. T.; Berg, J. M. "Metal Binding and Folding Properties of a Minimalist Cys2His2 Zinc Finger Peptide," *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4796-4800.

Chapter 2

Natural peptide domains as templates for protein design: Evaluation of a designed $\beta\beta\beta$ motif patterned after snake neurotoxins

Introduction

The protein architecture exhibits remarkable structural and functional diversity, and therefore represents an enticing target for synthetic manipulation and the potential for the creation of novel biocatalysts. Crucial to the design of functional polypeptides is the availability of synthetic motifs with defined tertiary structure. For this reason, considerable research has focused on the design of polypeptide sequences targeted to adopt a specific three-dimensional architecture.¹⁻³ Many design efforts have targeted helical structure, as nature provides several hints for the construction of motifs composed of this element of secondary structure. The discovery that the S-peptide of ribonuclease A^{4, 5} and sequences containing primarily alanine⁶ are helical has led several groups to develop simple helical model peptides patterned after these systems. These peptides have been used to establish thermodynamics scales for the helix propensities of natural and non-standard amino acids⁷⁻⁹ and to investigate the effects of helix dipole interactions,¹⁰ helix-stop signals,¹¹ and charged groups on helix stability.^{12, 13} The analysis of helix bundle proteins and the coiled coils found in fibrous proteins have also yielded insights into the packing interactions between these elements of secondary structure.¹⁴ Several natural peptides have also illustrated the potential of exploiting amphiphilicity in the creation of helical structure.¹⁴ Based in part on these observations, the *de novo* design of four helix bundles,¹⁵ coiled coils,^{16, 17} and helix-loop-helix¹⁸⁻²⁰ proteins has met with considerable success. As described in Chapter 1, although many designed constructs demonstrate some of the hallmarks of native protein structure such as cooperative thermal or chemical denaturation and protection of amide protons from exchange, few of these efforts have produced polypeptides for which a detailed NMR or crystal structure has been achieved.

Model systems for studying β -sheets have been more difficult to obtain; unlike α -helices, isolated β -sheets are generally not stable. Some of the difficulties in studying β -sheets arise from the extended nature of this element of secondary structure; β -sheets are typically formed by hydrogen bonds between residues distal in the primary sequence. The alternating nature of the hydrogen bonding interactions within β -sheet structure results in exposed amides on the edge of the sheet, which can lead to insoluble aggregates. For these reasons, design of proteins with β -sheet structure has proven considerably more challenging than the design of simple α -helices; many designs have suffered from insolubility and aggregation effects. Despite these difficulties, several groups have made significant strides toward the construction of novel β -sheet proteins through iterative design and analysis.²¹⁻²³ Biophysical measurements on these designs indicate that like many of their helical counterparts, these systems display some, but not all of the hallmarks of native protein structure.²⁴

While protein design has met with considerable progress in recent years, most successful designs are either large (>60 residues)^{15, 21, 22, 25} or multimeric.^{22, 23} Both of these factors contribute to the difficulty in obtaining detailed structural characterization of designed constructs. Iterative modeling, synthesis, and analysis are facilitated for smaller constructs. As described in Chapter 1, our approach to protein design has therefore been to target the assembly of small polypeptide motifs (< 40 residues) with defined tertiary structure based on naturally occurring polypeptides or protein domains. This small target size makes these motifs synthetically accessible, permitting the utilization of unnatural amino acids with novel steric and electronic properties to control secondary structure and incorporate functionality. Furthermore these constructs are amenable to detailed structural characterization by 2D ^1H NMR techniques. Most naturally occurring motifs of this limited size require extensive

cross-linking either in the form of disulfide bridges or metal ligation to define structure and impart stability. Our goal has been to modify these naturally occurring peptides and eliminate the need for multiple cross-links. A key concept of this approach is the use of small well defined β -turns as structural nucleation elements²⁶ to promote β -hairpin formation and aid in the development of tertiary structure in the absence of cross-links.

The first target of this approach is the design of a triple stranded antiparallel β -sheet, consisting of a strand-turn-strand-loop-strand topology in which a defined β -turn could be employed to join the first two strands. Considerable work in the Imperiali lab has focused on understanding the requirements for type II turn formation in model tetrapeptides,²⁶ and several short sequences which form defined turn structures in aqueous solution have been designed. It was envisioned that a type II turn could be utilized as a structural nucleation element to promote β -hairpin formation in the first two strands of the proposed $\beta\beta\beta$ motif (Figure 2-1, Turn A). In addition to functioning as a structural nucleation element, a well defined type II β turn at this position could function as the site for incorporation of catalytic activity in future designs. For example, a hydrophobic substrate binding pocket could be created between the sheet and loop regions and functionalized amino acids placed in the turn region at the edge of the pocket to create a "mini" active site. The proposed design is illustrated schematically in Figure 2-1.

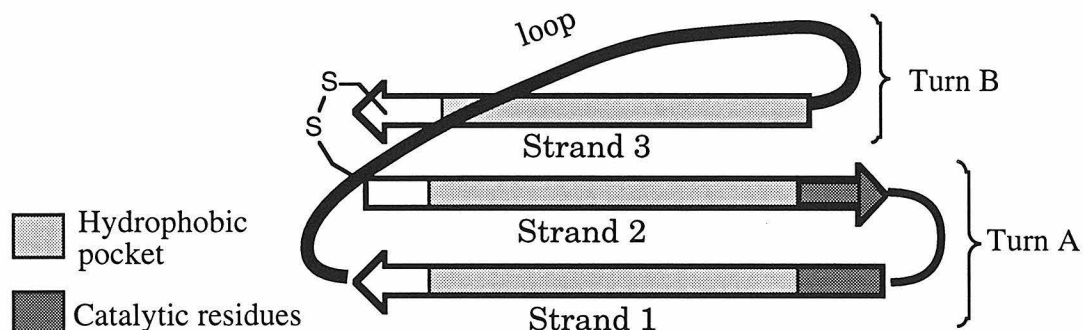


Figure 2-1 Schematic representation of the proposed $\beta\beta\beta$ motif.

The template for the design of a $\beta\beta\beta$ motif was derived from an investigation of the structures and sequences of a family of homologous snake neurotoxins.²⁷ This large family consists of very stable β -sheet proteins, 60-70 residues in length, which interact with the neuromuscular acetylcholine receptor. This family of proteins exhibits a high degree of structural and sequence homology and the crystal or NMR structures of several members have been solved.²⁸⁻³¹ Each protein contains three domains forming a five-strand antiparallel sheet which projects out from a core region incorporating four disulfide bonds (Figure 2-2). Contained within this five-stranded sheet is the desired $\beta\beta\beta$ motif (Figure 2-2, encircled region). Our plan was to excise this motif and engineer it to fold in isolation, using only one cystine cross-link, reverse turns of *de novo* design, and native hydrophobic interactions as structural determinants.

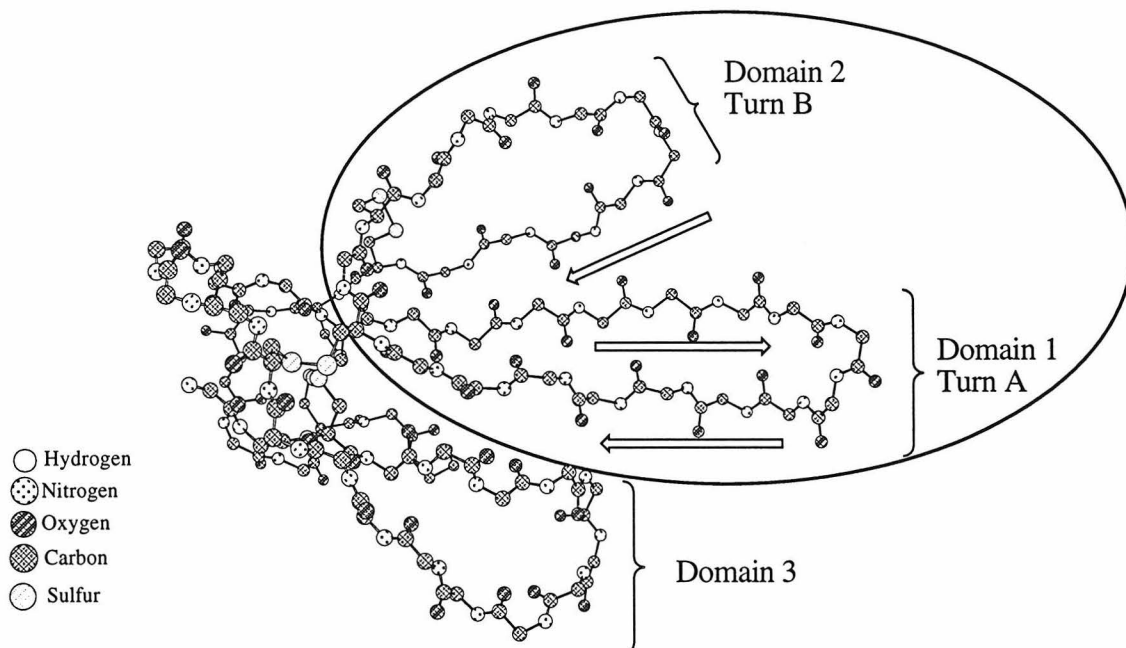


Figure 2-2 The structure of erabutoxin b.²⁸ The template used for the construction of a $\beta\beta\beta$ motif is encircled and arrows indicate the direction of the backbone.

The available structures of the neurotoxins in the protein database were explored²⁸⁻³¹ to determine the best target sequence for this design effort. Several factors were considered: the extracted $\beta\beta\beta$ motif would ideally contain only one internal disulfide bond, it would be as independent from the rest of the structure as possible, and comprise approximately 30 residues. The sequence which was found to be most compatible with these criteria was a 30 residue fragment (residues 25-54) of Erabutoxin b.²⁸ This region is approximately 50-60% β -sheet, consisting of an antiparallel sheet-turn-sheet domain which is connected to a third strand by a loop formed by residues 41-47. Two of the three domains which extend out from the disulfide core are included in this sequence (Figure 2-2, Domains 1 and 2). The third antiparallel sheet domain (Figure 2-2, Domain 3)

does not appear to interact significantly with the other two. In fact, distance measurements in the crystal structure indicate that the third domain participates in only one hydrogen bond to Domain 1. This region also contains a single internal disulfide bridge between residue 43 and 54, which connects the end of the third strand with the loop. The sequence does contain one additional cysteine (Cys⁴¹) which forms a disulfide with Cys¹⁷, connecting the triple stranded sheet to the cystine rich core. Based on this analysis, residues 25-54 of erabutoxin b were chosen as the template for our attempts to design a $\beta\beta\beta$ motif (Figure 2-3).

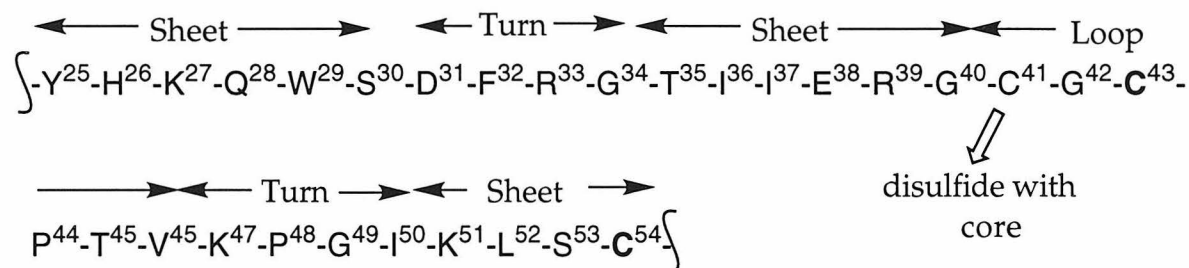


Figure 2-3 Sequence of erabutoxin b residues 24-53. The secondary structure of the sequence as observed in the crystal structure²⁸ is given. Residues involved in the internal disulfide bridge are in bold.

The design contains four distinct structural determinants. The defined type II β -turn (Figure 2-1, Turn A) serves as a structural nucleation element to promote β -hairpin formation between the first and second strands. The single disulfide forms a structural restraint on the loop and the third strand (Figure 2-1). In the crystal structure there appears to be significant hydrophobic interactions between the loop and the "top" of the triple stranded sheet,²⁸ which could

mediate interactions between the two regions (Figure 2-2 Domains 1 and 2). Additionally this design places the N- and C-termini in close proximity, which creates a potentially stabilizing electrostatic interaction between the charged termini in the intended $\beta\beta\beta$ topology.

Results

Three peptides based on the erabutoxin fragment were synthesized and analyzed. The sequences of these peptides are shown in Table 4-1. The design, synthesis and analysis of each peptide is discussed.

TOX1:																														
1				5						10					15					20					25				30	
Y	H	K	Q	F	S	V	P	D	S	H	T	I	I	E	R	G	S	G	C	P	T	V	K	P	G	I	K	L	S	C
NAT1:																														
1				5						10					15					20					25				30	
Y	H	K	Q	W	S	D	F	R	G	T	I	I	E	R	G	S	G	C	P	T	V	K	P	G	I	K	L	S	C	
TOX2:																														
1				5						10					15					20					25				30	
Y	H	K	Q	W	S	V	P	D	S	H	T	I	I	E	R	G	S	G	C	P	T	V	K	P	G	I	K	L	S	C

Table 2-1 Sequences of peptides based on erabutoxin b 24-53. All sequences have free amino and carboxylic acid termini.

TOX1 : Incorporation of a type II turn

Design Rationale

In the first design attempt (TOX1) the turn connecting the first two stands of "Domain 1" (Figure 4-2, Turn A) was re-engineered to include a defined type II β -turn. This region is poorly defined in the solution structures of several neurotoxins that have been solved by NMR,^{30, 32} implying that this segment of the peptide does not serve an explicit role in creating the β -sheet structure of the

toxin family. A type II turn was incorporated in this region to act as a structural nucleation element and play an *active* role to restore some of the stability lost by the removal of the disulfide cross-linked core of the native erabutoxin molecule.

A reverse turn consists of four residues and the dihedral angles of the central two amino acids [positions ($i+1$) and ($i+2$)] determine the turn type.³³ In the native erabutoxin b structure the sequence -Asp³¹-Phe-Arg-Gly³⁴- corresponds a type I β -turn.²⁸ Type I β -turns have a slightly twisted geometry³⁴ which often results in unusual dihedral angles at residues in positions (i) and ($i+3$) (Chapter1). The unusual conformations of these residues are required to accommodate the twist imparted by the turn and yet allow for the formation of β -sheet structure. The ideal dihedral angles of type I and type II β -turns as well as those of residues 31-34 in the erabutoxin b crystal structure²⁸ are shown for comparison in Tables 1-1 and 2-2.

Residue	position	ϕ (°)	ψ (°)
Asp 31	(i)	-135	-176
Phe 32	($i+1$)	-62.3	-18.1
Arg 33	($i+2$)	-86.1	-27.5
Gly 34	($i+3$)	154.3	-158.5

Table 2-2 The dihedral angles of the type I turn (Turn A) in the native sequence of erabutoxin b. The ideal dihedral angles for an antiparallel β -sheet are ($\phi = -139^\circ$, $\psi = 135^\circ$)

In **TOX1** the original type I turn of erabutoxin b was replaced with the sequence -Val-Pro-DSer-His- previously determined to adopt a defined type II β -turn in aqueous solution.²⁶ A heterochiral sequence, -Pro-DSer- was used to direct type II β -turn formation. The pyrrolidine ring of the proline residue

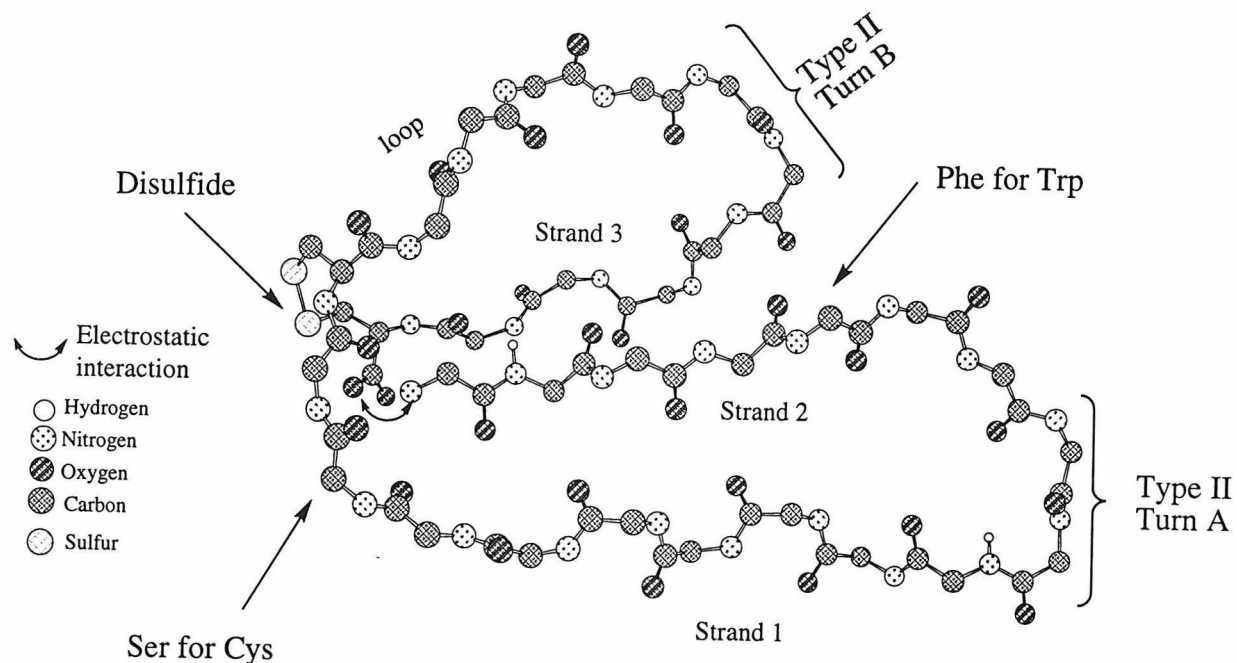
"kinks" the peptide backbone and initiates turn formation, while the use of a non-standard D-amino acid at position ($i+2$) favors the dihedral angles required at this position for a type II β -turn. Valine and histidine at positions (i) and ($i+3$), respectively, have been shown to favor turn formation in sequences incorporating this heterochiral sequence. It was envisioned that this rigid turn would help nucleate and stabilize sheet formation. Furthermore, a type II β -turn in the native structure (-Lys-Pro-Gly-Ile-, residues 23-26 in the **TOX1** sequence) connecting the crossover loop and the third strand (Figure 2-2, Turn B) was retained in **TOX1**.

In order to confirm that the incorporation of the engineered type II β -turn was compatible with the motif structure, modeling studies investigated the most common types of turns as potential replacements of Turn A (Figure 2-2) in the 24-54 fragment of the crystal structure of erabutoxin b. As discussed earlier, these four residues form a type I turn in which the (i) and ($i+3$) dihedral angles deviate considerably from ideal antiparallel sheet values to accommodate the type I β -turn geometry (Table 2-2). The dihedrals of the native tetrapeptide sequence, residues 31-34, were set to those of the ideal turns at positions ($i+1$) and ($i+2$) for the major turn types (Table 1-1) and ideal dihedrals for antiparallel sheet formation were used at positions (i) and ($i+3$). The results of this modeling study indicated that an ideal type II β -turn was as compatible with a $\beta\beta\beta$ topology as any of the other turns, and furthermore caused the least perturbation to the native structure.

Two other modifications were made to the primary sequence of the erabutoxin b fragment in the initial design, **TOX1**. Although the disulfide bond connecting the loop with the third strand was retained, the single cystine that connects this loop with the disulfide bonded core in the erabutoxin b sequence was replaced by serine to prevent asymmetric disulfide-bond formation among

three cysteine residues during synthesis. In addition, since tryptophan can present difficulties in solid phase peptide synthesis due to the acid lability and sensitivity to alkylation of the indole ring, phenylalanine was used at position five. These changes are summarized in Figure 2-4. As discussed earlier, the structure of the design places the N- and C-termini of the peptide in close proximity and an attempt was made to exploit this interaction between the charged amine and acid functionalities in the design of TOX1.

A.



B.

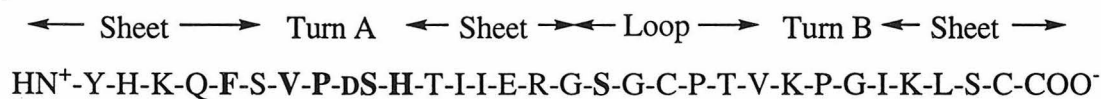


Figure 2-4 Design of TOX1. A. The proposed structure of TOX1. B. The sequence of TOX1. The changes to the original erabutoxin b sequence are indicated.

In summary, the initial design, **TOX1**, included three structural constraints: the disulfide bridge restricts the conformational mobility of the loop and third strand, the potential electrostatic interaction between the C- and the N-termini serve to place strands 2 and 3 in proximity for sheet formation, and the engineered type II β -turn serves as a structural nucleation site for sheet formation between strands 1 and 2. It was envisioned that these structural constraints would act in concert with the "native" hydrophobic interactions and allow this peptide to fold in isolation. The sequence of **TOX1** is shown in Table 2-1.

The peptide was synthesized by standard Fmoc based solid phase peptide synthesis. Mild oxidation of the cysteine residues in 20% aqueous dimethyl sulfoxide³⁵ to form the intramolecular disulfide bridge followed by HPLC purification yielded sufficient quantities of this peptide for NMR analysis.

Structural Analysis

Structural analysis of **TOX1** was performed with a number of two-dimensional NMR techniques. Spin systems were assigned by analysis of COSY and TOCSY spectra.³⁶ Most spin systems were identified from data acquired at a temperature of 281K, and the remaining ambiguous assignments were clarified with spectra acquired at 300K. The resonance of the C_{α} proton of Cys¹⁹ was obscured by the water signal and could not be identified in spectra taken using 90:10 H₂O:D₂O as solvent, but was observed in TOCSY spectra acquired in D₂O. The NH and β -proton resonances were confirmed by observation of the amide-beta cross-peak in the TOCSY spectrum. Sequential assignments were made from the observation of $d_{\alpha N}(i, i+1)$ cross-peaks in the ROESY and NOESY spectra. Spectral overlap was a significant problem in the analysis of this peptide. With the exception of one outlying amide resonance (Ile²⁶, $\delta_{NH} = 7.81\text{ppm}$), the amide resonances were tightly clustered in only a 0.54 ppm range

(δ : 8.18 - 8.72 ppm). Since amide proton chemical shifts are temperature dependent, the sequential assignment of every spin system could only be accomplished with reasonable confidence by acquisition of spectra at several temperatures. Additionally, the synthesis and comparison of other designs listed in Table 2-1 and described in subsequent sections, aided in correct identification and sequential assignment of spin systems for **TOX1**.

Analysis of the NOESY spectrum taken at 281 K was performed to determine the presence of secondary or tertiary structure for **TOX1**. In particular the spectrum was inspected for the presence of a number of key NOEs that might indicate sheet or turn formation. Observation of both a $d_{NN}(i+2, i+3)$ and a strong $d_{\alpha N}(i+1, i+2)$ NOEs, although not conclusive, are indicative of type II turn formation (Figure 2-5).³⁶ These NOEs were observed for both the designed and the natural type II turns in **TOX1**. All three $d_{NN}(\text{Ser}^9, \text{His}^{10})$, $d_{\alpha N}(\text{Ser}^9, \text{His}^{10})$ and strong $d_{\alpha N}(\text{Pro}^8, \text{Ser}^9)$ NOEs were present in the NOESY spectrum of **TOX1** and suggested turn formation in this region.³⁶ The absence of the expected $d_{\alpha N}(\text{Pro}^8, \text{His}^{10})$ NOE indicated that this sequence did not form a tight β -hairpin. Two of the analogous NOEs ($d_{NN}(\text{Gly}^{25}, \text{Ile}^{26})$ and $d_{\alpha N}(\text{Pro}^{24}, \text{Gly}^{25})$) were also observed for residues 23-26 (Figure 2-4, Turn B) in the NOESY spectrum indicating that some population of the peptide contained a turn-like conformation in this region.

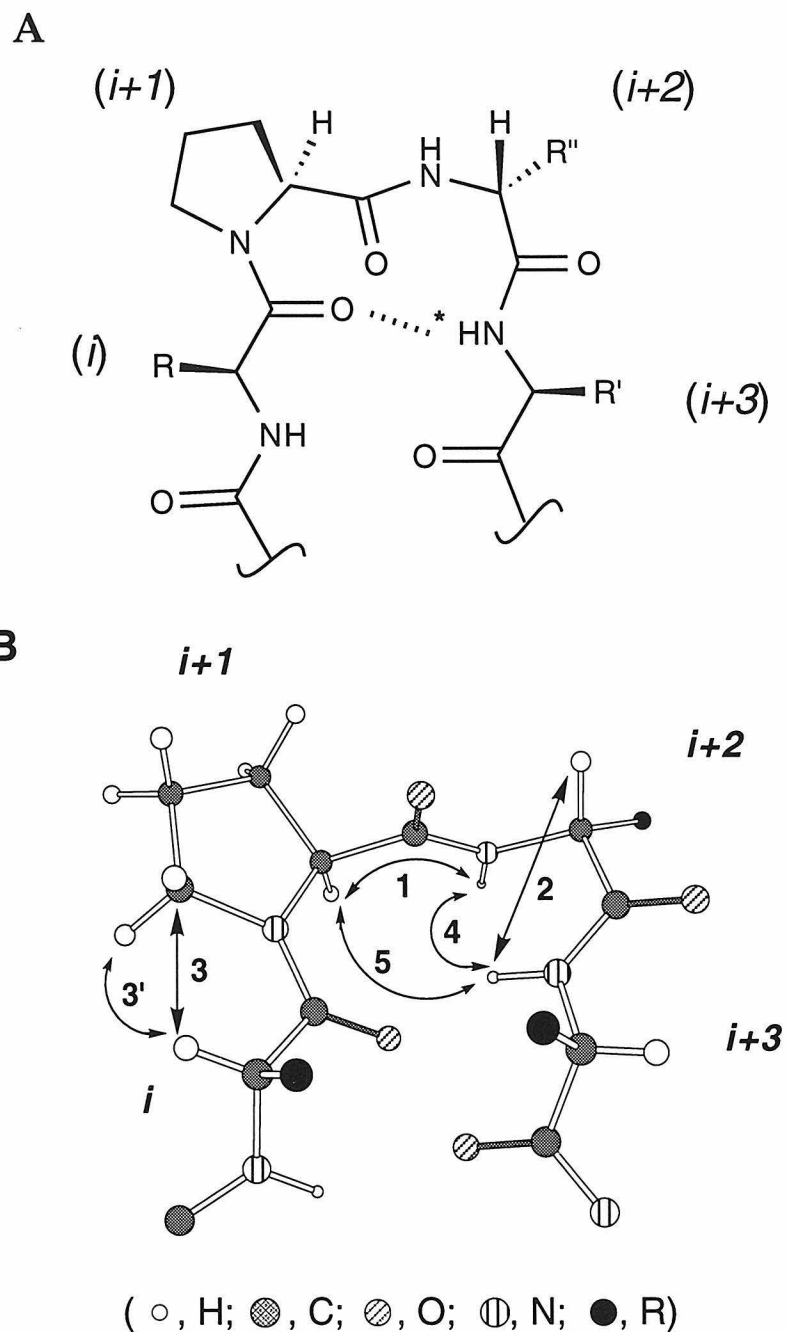


Figure 2-5 Conformation of a type II β -turn. **A.** Tetrapeptide (with proline at position $i+1$) in a type II turn formation. **B.** NOEs indicative of type II turn formation. 1 $d_{\alpha N}(i+1, i+2)$; 2 $d_{\alpha N}(i+2, i+3)$; 3,3' $d_{\alpha \delta}(i, i+2)$; 4 $d_{NN}(i+2, i+3)$; 5 $d_{\alpha N}(i+1, i+3)$. The residue shown at position $(i+2)$ is a D-amino acid. Glycine was present at this position in Turn B (residues 23-26).

The turn sequences of TOX1:

Turn A: -Val⁷-Pro⁸-DSer⁹-His¹⁰-

Turn B: -Lys²³-Pro²⁴-Gly²⁵-Ile²⁶-

In addition to the NOEs specified, a low amide proton temperature coefficient ($-\Delta\delta/\Delta T < 6.0$ ppb/K) for the amide at position ($i+3$) is expected for a turn-like conformation (Figure 2-5).²⁶ In a β -turn structure this amide participates in a hydrogen bonding interaction with the carbonyl at position (i). Such an interaction shields the amide proton from exchanging with solvent and lowers the amide temperature coefficient. Assignment of the 2D TOCSY spectra at three temperatures, permitted the determination of amide temperature coefficients for most amide protons. His¹⁰ and Ile²⁶ at position ($i+3$) of the designed and native type II β -turn, respectively, have significantly lowered temperature coefficients ($-\Delta\delta/\Delta T < 4.0$ ppb/K) in agreement with the NOEs observed for these region .

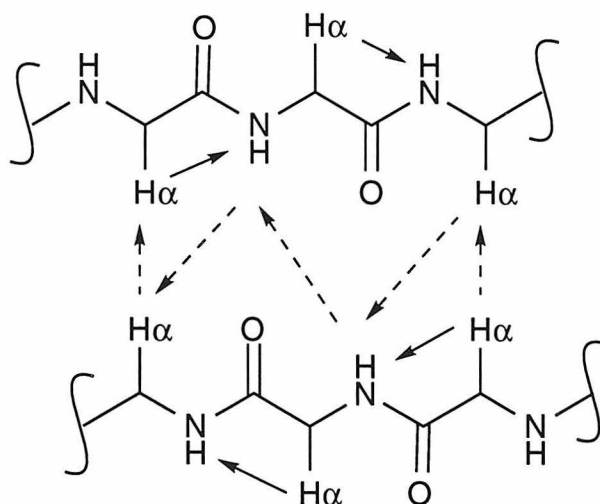


Figure 2-6 Schematic illustration of the NOEs which define an ideal β -sheet structure. Intra-strand $d_{\alpha N}(i, i+1)$ are shown as solid arrows, while the inter-strand $d_{\alpha\alpha}(i, j)$, $d_{NN}(i, j)$ and $d_{\alpha N}(i, j)$ are shown by dashed arrows.

β -sheet structure in proteins is characterized by both short range intra-strand and long range inter-strand NOEs (Figure 2-6).³⁶ In an extended, sheet-like structure, the $d_{\alpha N}(i, i+1)$ distance is $< 3.0 \text{ \AA}$, and should result in strong $d_{\alpha N}(i, i+1)$ NOEs in NOESY ($\sim 5.0 \text{ \AA}$ cutoff) and ROESY ($\sim 3.5 \text{ \AA}$ cutoff) experiments. Strong $d_{\alpha N}(i, i+1)$ NOEs were observed in a ROESY experiment for residues in the ranges 4-7, 11-15, and 26-30. These regions were designed to form the triple stranded β -sheet in **TOX1**. Although $d_{\alpha N}(i, i+1)$ NOEs are not unique to β -sheet structure, observation of these NOEs did suggest an extended backbone conformation.³⁶ While the remaining residues showed strong to medium $d_{\alpha N}(i, i+1)$ NOEs in the NOESY spectrum, the corresponding NOEs were weak or absent from the ROESY spectrum, indicating a less extended conformation in the loop region. Typically, in addition to strong intra-strand $d_{\alpha N}(i, i+1)$ sequential NOEs, the observation of several $d_{\alpha\alpha}(i, j)$, $d_{NN}(i, j)$ and $d_{\alpha N}(i, j)$ NOEs uniquely identifies sheet structure (Figure 2-6).³⁶ None of these inter-strand NOEs were observed for **TOX1** indicating that despite the presence of the two turns and an extended conformation for much of the sequence, the peptide did not adopt the desired secondary or tertiary structure. In addition, several weak $d_{NN}(i, i+1)$ NOEs were observed for **TOX1**. In a purely extended or sheet like conformation the amide protons on adjacent residues should be on the opposite side of the strand and would not be expected to show strong d_{NN} NOEs. Therefore, observation of these NOES indicates that although the backbone of much of the peptide is extended several regions probably adopt transient non-extended conformations.

Variable temperature analysis and H/D exchange studies confirmed the lack of a defined structure in **TOX1**. Reduced amide temperature coefficients ($-\Delta\delta/\Delta K < 4.0 \text{ ppb/K}$) were observed only for His¹⁰ and Ile²⁶ in the turn regions, and no other amide had a significantly lowered amide temperature coefficient.

Additionally, no slow exchanging amide protons ($t_{\text{exchange}} > 20\text{min}$) were seen in deuterium exchange experiments performed on **TOX1**. Therefore, although some amide protons exhibited low temperature coefficients suggestive of hydrogen bonding in the turn regions, these interactions were not sufficient to prevent rapid deuterium exchange with the solvent. This observation was also consistent with the absence of the $d_{\alpha\text{N}}(\text{Pro}^8, \text{His}^{10})$ or $d_{\alpha\text{N}}(\text{Pro}^{24}, \text{Ile}^{26})$ NOEs for these turns.

Our first attempt at the design of a $\beta\beta\beta$ motif on the erabutoxin b sequence met with limited success. Significantly, despite the absence of a distinct tertiary structure, several NOEs and amide variable temperature analysis revealed that the designed type II β -turn was indeed present to some degree in the large peptide. Two additional peptides were synthesized to aid in the overall analysis of these results: the native sequence of erabutoxin b 25-54 with only the $\text{Cys}^{42} \rightarrow \text{Ser}$ substitution (**NAT1**), and **TOX2**, the $\text{Phe}^5 \rightarrow \text{Trp}$ derivative of **TOX1**. The first of these peptides, **NAT1**, served as a control to investigate the effect of the engineered type II turn on the inherent structure of the fragment. The native sequence has a tryptophan residue at position 29 (position 5 in the **TOX1** numbering scheme). This residue is invariant in the toxin family and evidence suggests that it plays both a structural and functional role.³⁸ Trp^{29} appears to be central to the hydrophobic cluster formed by the residues of strands 2 and 3 and the loop (Figure 2-2, Domains 1 and 2), and might have a large impact on tertiary structure.³⁸ Therefore peptide **TOX2**, with the same sequence as **TOX1** but retaining this tryptophan at position 5, was synthesized and analyzed.

NAT1 and TOX2:

Both peptides **TOX2** and peptide **NAT1** were synthesized by standard Fmoc solid phase peptide synthesis. Some difficulty was initially encountered in

the synthesis of these peptides. Previously, standard protocols used 20% piperidine in DMF for Fmoc deprotection during solid phase peptide synthesis (SPPS).³⁷ Subsequent to the synthesis of **TOX1**, the successful use of 2% DBU in DMF for Fmoc deprotection during SPPS was reported.³⁸ Since piperidine is toxic and a controlled chemical, the deprotection protocol was changed to use DBU. Mass spectral analysis of HPLC purified material from the initial synthesis attempts of **NAT1** and **TOX2** revealed a mixture of two peptides: a product with the correct molecule weight and a product with a molecular weight less by 34 amu ($\Delta MW -34$). These two products co-eluted during HPLC and capillary zone electrophoresis (CZE) analysis, complicating the identification of the side product. The ratio of the side product to the desired peptide appeared to vary in different syntheses ranging from 10:1 to 2:1 in favor of the side product. It was initially suspected that this species was an artifact of the ionization techniques used in the mass spectral analysis, but data obtained using both electrospray ionization and plasma desorption ionization techniques yielded similar results. It was subsequently discovered that the use of 2% DBU for Fmoc deprotection resulted in the elimination of the side chain trityl-protected sulfur from the C-terminal cysteine to produce dehydroalanine. Since the peptide design incorporated a free C-terminus, all peptides were assembled *via* an ester linkage to the resin in order to provide a free C-terminal acid upon cleavage.³⁷ This ester linkage labilized the alpha proton of the C-terminal cysteine residue and facilitated elimination. It was also apparent from these synthesis attempts that although elimination occurs to only a small extent during each deprotection cycle, the synthesis of a 30 residue peptide ultimately produced significant amounts of dehydroalanine-containing peptides. Due to these problems, peptide **NAT1** was obtained immobilized on the solid support as a gift from

MilliGen/Biosearch (Millipore, Burlington, MA). A solution of 20% piperidine in DMF was used successfully for Fmoc deprotection during SPPS of **TOX2**.

Structural Analysis

The NMR results for both peptides **NAT1** and **TOX2** are qualitatively similar to those obtained for **TOX1**. The native sequence (**NAT1**) was found to be considerably more insoluble than either of the sequences containing the designed type II turns (**TOX1** and **TOX2**). While the solubility limit appeared to be ~ 8mg/ml (2.7mM), this concentration was sufficient for NMR analysis.

In the **NAT1** peptide, the original type I β -turn connecting strands 1 and 2 (residues Asp⁷-Gly¹⁰), and the type II β -turn between the loop and the third strand (residue Lys²³-Ile²⁶) were maintained (Figure 2-2, Turn A and B, respectively). NMR analysis revealed that, just as was the case for **TOX1**, several NOEs indicative of type II turn formation for residues Lys²³-Ile²⁶ were observed in the NOESY spectrum of **NAT1** (strong $d_{\alpha N(i+1, i+2)}$ and $d_{NN(i+2, i+3)}$ NOEs). Additionally, the amide of Ile²⁶ exhibited a reduced amide temperature coefficient as would be expected for turn formation (Figure 2-5). Similarly, NMR analysis indicated that residues 7-10 formed a turn, although definitive assignment of a turn type to this region was not possible. In all four of the major β -turn types (Table 1-1) the amide of position ($i+3$) participates in an (i) CO $\rightarrow$ ($i+3$) NH hydrogen bond, and would therefore be expected to exhibit a reduced variable temperature coefficient. Additionally, all turn types would be expected to exhibit a $d_{NN(i+2, i+3)}$ NOE, although they can be differentiated by the strength of the $d_{\alpha N(i+1, i+2)}$ NOE. Specifically, a strong $d_{\alpha N(i+1, i+2)}$ NOE is expected for a type II β -turn, while this NOE is very weak in a type I β -turn which exhibits a strong $d_{NN(i+1, i+2)}$ NOE instead (Figure 2-7). In the case of the type II turn containing the heterochiral Pro-DSer sequence in **TOX1**, the presence of

the D-amino acid precluded the adoption of any other turn type, since a D-amino acid at the ($i+2$) position is only compatible with a type II β -turn.²⁶ The native sequence of **NAT1** for residues 7-10 consists of all natural L-amino acids and would therefore be expected to favor a type I turn. Unfortunately the amide proton resonances of Phe⁸ and Arg⁹ had very similar chemical shifts which prevented observation of the $d_{\text{NN}}(i+1, i+2)$ NOE characteristic of a type I turn. The strength of the $d_{\alpha\text{N}}(i+1, i+2)$ NOE between Phe⁸ and Arg⁹ could also not be accurately ascertained due to spectral overlap. Therefore, although the reduced amide temperature coefficient of Gly¹⁰ and the $d_{\text{NN}}(\text{Arg}^9, \text{Gly}^{10})$ NOE suggested turn formation, the type of turn formed could not be established.

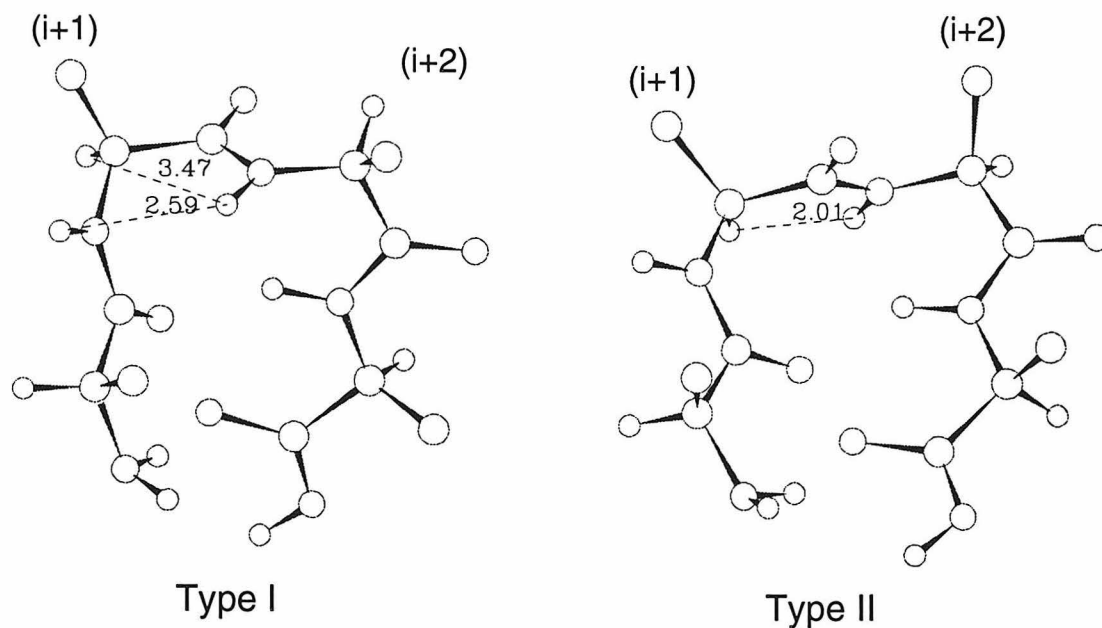


Figure 2-7 The $d_{\alpha\text{N}}(i+1, i+2)$ and $d_{\text{NN}}(i+1, i+2)$ distances in an ideal type I or type II β -turn.

Sequential $d_{\alpha N}(i, i+1)$ NOEs were observed for many residues of NAT1 (Table 2-4). Additionally $d_{NN}(i, i+1)$ NOEs were observed for several residues within the peptide that were not expected to be in a turn conformation (Table 2-3). These NOEs should be weak or absent in an ideal extended sheet-like conformation (Figure 2-8), yet they were present and stronger than expected for the native sequence, especially for the residues adjacent to the turn sequence (Asp⁷-Gly¹⁰). As mentioned earlier the type I β -turn at this position in the structure of erabutoxin b determined by crystallographic analysis²⁸ is twisted, and the twist is compensated for by the unusual dihedrals of the residues at positions (i) and ($i+3$). Due to this twist it is possible that the surrounding residues adopt a non-extended conformation resulting in stronger $d_{NN}(i, i+1)$ NOEs in this region.

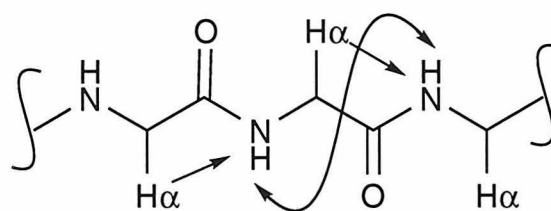


Figure 2-8 A sequential d_{NN} NOE is illustrated by the curved arrow for a peptide backbone in an extended conformation. The expected sequential $d_{\alpha N}$ NOE is also shown.

Peptide **TOX2** differs from **TOX1** only in the Phe⁵→Trp substitution. As with **TOX1**, NOEs for both the engineered and the native type II β -turns were observed for peptide **TOX2**. A strong sequential $d_{\alpha N}(i+1, i+2)$ NOE and the $d_{NN}(i+2, i+3)$ NOE together with a reduced amid proton temperature coefficient

for His¹⁰ at position ($i+3$) indicated type II turn formation for residues 7-10. In this case the inclusion of the heterochiral sequence, Pro⁸-DSer⁹, disfavored formation of any of the other turn types. Similar NOEs and amide temperature coefficients are observed for residues 23-26, indicating type II turn formation in the disulfide linked loop. Although, this peptide displays sequential $d_{\alpha N}(i, i+1)$ NOEs similar to NAT1, the $d_{NN}(i, i+1)$ observed for residues surrounding the turn region are considerably weaker (relative to the turn $d_{NN}(\text{Arg}^9, \text{Gly}^{10})$ NOE) or absent in this peptide, highlighting the differing effects of the type I and type II β -turns on the conformation of the surrounding residues. A comparison of the relative intensities of both $d_{\alpha N}(i, i+1)$, the $d_{NN}(i, i+1)$ NOEs, and the results of the amide variable temperature analysis are given in tables 2-4, 2-5, and 2-3 respectively.

(i)	(i+1)	TOX1	NAT1	TOX2
5	6	not observed	weak	not observed
9	10	medium	medium	medium
10	11	medium	medium	weak
11	12	overlap	medium	overlap
13	14		weak	weak
17	18	strong	strong	strong
25	26	weak	medium	medium
29	30	weak	weak	weak

Table 2-3 Summary of $d_{NN}(i, i+1)$ NOEs observed for all three peptides. The relative intensities were determined by qualitative inspection of the 200 ms NOESY at 281K The NOEs indicative of turn formation are highlighted.

Sequential $d_{\alpha N}(i, i+1)$ NOEs (200 ms NOESY, 281 K)				
α -proton	amide	TOX1	NAT1	TOX2
2	3	weak	absent	
3	4	not observed	strong	strong
4	5	medium	medium	medium
5	6	medium	not observed	strong
6	7	strong	medium	strong
7	8	^a	medium	^a
8	9	strong	overlap	strong
9	10	medium	medium	medium
10	11	^b	medium	weak
11	12	strong	medium	weak
12	13	strong	overlap	medium
13	14	strong	strong	not observed
14	15	not observed	overlap	not observed
15	16	not observed	overlap	not observed
16	17	medium	strong	medium
17	18	very weak	weak	weak
18	19	medium	medium	medium
20	21	strong	not observed	strong
21	22	strong	strong	strong
22	23	strong	overlap	medium
24	25	medium	medium	medium
25	26	not observed	not observed	very weak
26	27	strong	medium	medium
27	28	medium	weak	medium
28	29	medium	weak	medium
29	30	medium	weak	strong

Table 2-4 The relative intensities of the sequential $d_{\alpha N}(i, i+1)$ NOEs for each peptide. Relative intensities are based on qualitative inspection of the NOESY spectrum and serve only for comparison within a peptide. ^a Proline at the $i+1$ position. ^b Alpha proton resonance obscured due to water signal.

Residue	TOX1	NAT1	TOX2
His 2			
Lys 3	-7.72	-7.755	-7.67
Gln 4	-7.57		
Trp 5	-8.32 (Phe)	-9.91	-9.45
Ser 6	-6.10	-5.22	-5.84
Val 7	-8.68	-4.82 (Asp)	-7.57
Pro 8		-6.37 (Phe)	
DSer 9	-9.30	-5.48 (Arg)	-8.47
His 10	-3.93	-1.81(Gly)	-4.59
Thr 11	-6.30	-6.45	-6.46
Ile 12	-7.15	-8.67	-9.22
Ile 13	-9.63	-8.68	-8.54
Glu 14	-7.55	-6.14	
Arg 15	-8.12	-8.5	-6.54
Gly 16	-6.44	-6.52	
Ser 17	-7.22	-7.41	-7.65
Gly 18	-6.68	-6.75	
Cys 19		-5.16	-5.33
Thr 21	-9.07	-10.0	-9.48
Val 22	-8.70	-9.08	-9.46
Lys 23	-8.92	-9.35	-9.36
Gly 25	-7.10	-7.205	-7.24
Ile 26	-3.45	-3.34	-3.22
Lys 27	-6.78	-7.67	7.58
Leu 28	-8.26	-7.47	-7.77
Ser 29	-8.13	-8.36	-8.35
Cys 30	-4.82	-5.197	-5.48

Table 2-5 Comparison of amide proton variable temperature coefficients ($\Delta\delta/\Delta T$, ppb/K) for the three peptides. The sequence listed is for **TOX2**. Sequence differences in peptides **NAT1** and **TOX1** are noted by the appropriate residue listed after the amide VT coefficient.

Discussion

On the basis of the structural studies, it was concluded that all three peptides lacked a defined secondary or tertiary structure. Although there is some evidence for turn formation for residues 7-10 and 23-26 in all peptides, this was insufficient to direct the adoption of a discrete structure, and all peptides existed primarily in random coil conformations. Based on these results a careful analysis of the single covalent restraint retained within the excised $\beta\beta\beta$ motif (the disulfide between residues 19 and 30) was performed.

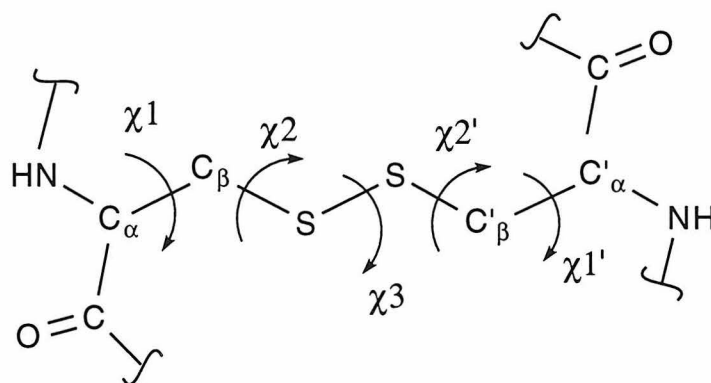


Figure 2-9 The dihedral angles that defined disulfide geometry.

The geometry of a disulfide is described by five dihedral angles as illustrated in Figure 2-9. These five angles also determine the torsional strain energy associated with a particular disulfide bond.³⁹ A disulfide is generally considered to stabilize the folded conformation of proteins by affecting the energetics of the unfolded and folded states, although the exact nature of the stabilization is still a matter of uncertainty.⁴⁰ Engineering cystine bridges into

existing proteins has been the focus of several protein design efforts.⁴¹⁻⁴³ Interestingly, not all efforts have resulted in increased stability.⁴⁴ The failure of some efforts has been attributed, in part, to the strict geometrical constraints on the formation of disulfides with low torsional energies.⁴⁵ Early analysis of disulfide conformations in the crystal structures of natural proteins indicated that there was a strong preference for either a left-handed "spiral" conformation ($\chi^3 = -90^\circ$) or a right-handed "hook" conformation ($\chi^3 = +90^\circ$) although considerable variance was noted.³⁹ Recently, Katz and Kossiakoff solved the crystal structure of several variants of subtilisin which contained engineered disulfides.⁴⁶ They also performed a more comprehensive analysis of the conformations of both engineered and native disulfides in proteins.⁴⁶ Their results indicate that disulfides can have very different torsional energies, and the conformation of the disulfides engineered into subtilisin were of "atypical" geometries and high calculated energies. Disulfides from all proteins studied were classified into five general families based on the dihedral values and geometries, and disulfide energies were estimated for each class based on the average dihedral angles.⁴⁶ The theoretical torsional strain energy of a disulfide is determined primarily by the χ^3 dihedral angle, which exhibits an energy minimum at $\pm 83^\circ$.⁴⁶ The other side chain dihedral angles (χ^1, χ^2) have less influence on the energy but do exhibit definite conformational preferences.⁴⁶ While Katz and Kossiakoff concluded that even high energy strained disulfides are tolerated in engineered systems,⁴⁶ it is not clear that such cross-links contribute to protein stability.⁴⁷ In fact, the high energy disulfides engineered into subtilisin do not increase its stability towards irreversible inactivation.⁴⁴ Additional experiments point to the importance of disulfide geometry in determining the potential stabilizing effects of engineered cystine cross-links. Multiple disulfide bridges were engineered into coiled-coil proteins of *de novo* design.⁴⁸ Analysis of the resulting constructs

indicated that most effective cross-links (i.e. those that afforded the greatest increase in stability) were achieved with cystine bridges for which near-optimal geometries could be modeled into the designed coiled-coil structure.⁴⁸ Clearly, the geometrical preferences of disulfides must be included in estimating the potential stability afforded by the cross-link. For our purposes, a disulfide bond should serve as a structural constraint, to stabilize and promote the folding of the desired tertiary structure. Therefore it is important to design a disulfide with the near-optimum geometry in future designs in order to afford the most effective structural determinant.

Analysis of the dihedral angles associated with all four cystines found in the erabutoxin b structure²⁸ revealed that the disulfide retained in the **TOX1** design was of somewhat high energy. In fact, this disulfide has the highest torsional energy of all the disulfides found in the erabutoxin b crystal structure (Table 2-6). Thus the disulfide which was initially anticipated to be a powerful constraint on the structure of the motif was probably not the optimal choice for the desired structure. While this disulfide conformation is clearly tolerated in the erabutoxin b structure, there are many other factors which determine the structure of the toxin, including three additional disulfides which create a hydrophobic core. If the strain of this particular disulfide is high then the incorporation of this cross-link may actually disfavor the desired $\beta\beta\beta$ structure. As a covalent cross-link a disulfide can be a very powerful structural determinant. However, by retaining the 43-54 disulfide (erabutoxin b numbering scheme) we may not be using the potential of this cross-link to the fullest possible extent.

Disulfide	χ_1	χ_2	χ_3	χ_2'	χ_1'	Ca-Ca Dist. (Å)	Class	Predicted energy ⁴⁶
54-43 30-19	-61.5	178.5	83.6	66.7	-174.3	5.60	4 RH	2.83 kcal/mol
17-41 -7-17	-47.41	-54.1	-82.7	-91.8	-74.1	5.70	1 LH	0.54 kcal/mol
3-24 -22-(-1)	-65.2	-67.8	-85.4	-59.8	-60.7	5.68	1 LH	0.54 kcal/mol
55-60 +1-+6	53.5	88.9	85.5	88.1	63.1	6.16	unk	

Table 2-6 Analysis of the geometries and relative energies of the disulfides present in erabutoxin b.²⁸ Erabutoxin b numbering is in bold, and the position relative to the **TOX1** sequence indicated below. If feasible, each disulfide has been classified (according to Katz and Kossiakoff)⁴⁶ on the basis of dihedral angle values. The predicted torsional energy of each class is listed for comparison.

Analysis of the conformation of the erabutoxin b disulfides led to the consideration of relocating the disulfide constraint, possibly creating a cross-link not present in the native system. An alternative position for the cystine bridge is linking the polypeptide termini. If the two β -turns promote the reversal of the peptide backbone and the ends of the peptide are covalently linked *via* a cystine bridge, then the ensemble of structures compatible with these three constraints should include the desired $\beta\beta\beta$ topology. This design would require lengthening the peptide by several residues on either end in order to place the disulfide external to the sheet region of the motif and allow for considerable flexibility in disulfide bond formation, permitting low energy dihedral angles. Initial modeling studies, including simulated annealing molecular dynamics, were performed to investigate this possibility. The results were encouraging; if the residues 7-10 and 23-26 are fixed in a β -turn conformation and the ends covalently cross-linked with cystine, the structure of the peptide is largely

maintained even during dynamics at 1500K. Dynamics at this temperature allow the molecule to sample many conformations; yet with the simple constraints of the two turns and the terminal cross-link, the overall fold retains the desired $\beta\beta\beta$ topology.

In addition to cystine, non-natural thiol containing amino acids could be considered for disulfide bond formation, including β -methyl cysteine (Bmc)⁴⁹ and homocysteine. Bmc is a β -branched amino acid and would therefore favor an extended sheet like conformation,⁵⁰ while homocysteine would allow for additional flexibility in disulfide bond formation potentially reducing any torsional strain. An asymmetrical pairing of these two residues might also be profitably exploited; while Bmc is potentially too conformationally constrained to form low energy symmetrical disulfides, cross-linking with homocysteine might compensate for this drawback and allow the benefits of the β -branched Bmc to be reaped without paying any dues in torsional strain.

Conclusions

The construction of a well defined structural template in less than 50 residues is a difficult task; Nature typically employs cross-links such as disulfides or metal ligation to achieve this goal. However, these constraints reduce the versatility of these natural peptides for future elaboration into functional constructs. In addressing this problem, our approach to protein design involves the utilization of defined β -turns and unnatural amino acids to achieve a defined three-dimensional structure containing only the minimal cross-linking interactions. The work presented in this chapter represents the first attempt at the implementation of this approach. While assembly of the target $\beta\beta\beta$ motif was not achieved, several valuable lessons were learned. Caution must be exercised in the use of a disulfide bridge to control peptide structure; the

torsional energy of a disulfide must be considered when analyzing its potential effects on conformation. Encouragingly NMR analysis indicated that the engineered turn motif adopted a folded conformation even in the presence of largely random coil structure. This observation underscores the potential utility of these small motifs to form elements of local secondary structure that might be properly exploited as structural constraints if located in the appropriate environment.

The $\beta\beta\beta$ motif is an appealing motif for the design of a catalytic polypeptides; the accessibility of a hydrophobic binding pocket comprising the β -sheet and the ability to position several functional residues in close proximity around a β -turn are desirable characteristics for the long term goal of functional peptidyl constructs. While reassessment of the designed $\beta\beta\beta$ motif did suggest some design alternatives for future constructs, the "open" nature of the snake neurotoxins, and the conformational dependence of the excised motif on both auxiliary protein structure and multiple cross-links may have made the snake neurotoxins a poor candidate for the basis of a first design attempt. Based on these results, two factors will need to be considered in choosing future design targets. First, the chosen natural template should form an independently folded peptide or protein domain and not be dependent on a supporting protein framework. This desired characteristic would restrict our task to the elimination of the requirement for cross-links in the adoption of the desired tertiary structure. Second, the target motif should have a well defined hydrophobic core creating a compact globular structure. In this regard, there are several proteins with substantial β -sheet character which may serve as more appropriate natural templates. Charybdotoxin forms a compact $\beta\alpha\beta\beta$ motif in only 37 residues, including 3 disulfides.⁵¹ The helix packs against a triple stranded antiparallel β -sheet, linked by two disulfides, while a third disulfide connects the first strand to

the C-terminal β -hairpin. In addition to the disulfides, two groups of amino acids form hydrophobic clusters within the motif and appear to be structurally important.⁵²

The C-terminal domain of cellobiohydrolase represents the "reverse" of the $\beta\beta\beta$ motif described here, consisting of a strand-loop-strand-turn-strand topology.⁵³ Unlike the "toxin" motif investigated in this chapter, this protein domain forms an independently folded structure comprising only 36 residues, including 2 disulfides. The solution structure has been solved by 2D NMR techniques⁵³ and this system presents several opportunities to exploit the β -turn structures developed in this lab.

A third and very appealing possibility is presented by the zinc finger peptides.^{54, 55} Sequential arrays of these small domains (25-30 residue) form the DNA binding regions in many transcription factor proteins.^{54, 55} These polypeptide fragments form compact mini-globular structures with a $\beta\beta\alpha$ topology through the use of metal ligation. In addition to metal coordination, these motifs include a defined hydrophobic core of three residues, which are critical for defined three-dimensional structure in these peptides.⁵⁵ The following chapters describe the second attempt at the design of a versatile structural template from naturally occurring peptide sequences based on the zinc finger motifs. Key concepts presented in this chapter such as the incorporation of unnatural amino acids to control structure and incorporate functionality and the utilization of defined β -turns as structural nucleation elements are elaborated upon in the next two chapters.

Acknowledgments

The initial synthesis and analysis of TOX1 was done in collaboration with Tarun Kapoor and Ranabir Sinha Roy. Craig Dobbs of Millipore provided

peptide NAT bound to the solid support. Jane Sanders provided advice concerning the mass spectral analysis of NAT1 and TOX2 samples.

Experimental

Peptide Synthesis

TOX1 was synthesized by solid phase methods on a Milligen 9050 automated peptide synthesizer using *N* α -9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids and BOP/HOBT or DIPCDI/HOBT amide coupling chemistry.⁵⁶ Peptide TOX1 was assembled on Fmoc-L-Cys(Trt)-KA resin (Millipore, MA) of 0.12 mmole/gram substitution. Fmoc deprotection was accomplished by washing the resin with 20% piperidine in DMF. Typical coupling cycles were 1 to 1.5 hours using a 3- or 4-fold excess of amino acid. Double coupling cycles were performed for amino acids which have been reported to be difficult to couple during SPPS.⁵⁷ Simultaneous removal of side chain protecting groups and cleavage from the resin was accomplished with Reagent K (82.5:2.5:5:5:5 TFA/ethanedithiol/H₂O/ anisole/phenol)⁵⁸ in 2-3 hours. The resin was filtered and the filtrate collected and concentrated. The peptide was precipitated from the TFA mixture with 2:1 ether/hexane. The supernatant was decanted after centrifugation and trituration with 2:1 ether/hexane was repeated until the peptide formed a fluffy white solid. The peptide was dissolved in a minimal amount of water and lyophilized. The crude peptide sample was oxidized by addition of 20% dimethyl sulfoxide (DMSO)³⁵ to an aqueous solution of the peptide at pH 7.0 or by a air oxidation of dilute solutions. Treatment of the oxidized peptide with Ellman's reagent⁵⁹ revealed that the free thiol content was < 2%. The crude oxidized sample was passed through a Centricon membrane filter with a molecular weight cutoff of 3000 daltons (slightly less than the expected molecular weight of the peptide). All

peptidyl components of the crude mixture passed through the membrane indicating the absence of intermolecular disulfide formation. The oxidized sample was then purified by HPLC. Electrospray ionization (ESI) mass spectrometry confirmed that the peptide was a monomer with a molecular weight of 3270.8 daltons (MW calc. 3270.8). The identity of the polypeptide was also confirmed by amino acid analysis.

NAT1 was obtained as the fully protected resin immobilized peptide from MilliGen/Biosearch (MA). The incorporation of the tryptophan residue *via* solid phase peptide synthesis has been known to produce significant side products due to facile alkylation of the indole ring by the cations produced during side chain deprotection and cleavage of the peptide from the resin.⁶⁰ Tryptophan is also acid labile and sensitive to oxidation, therefore minimal peptide cleavage time is desirable.⁶¹ For this reason the Pmc (2,2,5,7,8,-pentamethyl-chroman-6-sulfonyl choride) group was used for the side chain protection of arginine, since this group is more quickly removed during cleavage than alternatives (such as the Mtr, methoxytrimethylbenzene sulfonyl, protecting group).⁶¹ Studies were performed to determine the minimal cleavage time necessary for both the complete side chain deprotection and removal from the resin. Approximately 150 mgs of resin was cleaved using Reagent K. Approximately equal aliquots of resin and cleavage mixture were removed at various time points from 15 minutes to 3 hours and treated as described for **TOX1**. Cleavage times of less than one hour appeared to afford lower yields of crude peptide than did longer time points. HPLC analysis indicated that the purity of the crude peptide was comparable for all samples cleaved for at least 1 hour. ESI mass spectral analysis of purified samples from 45 minute and 3 hour cleavage times confirmed that the major product in each case was the correct peptide ($C_{149}H_{235}O_{42}N_{43}S_2$, MW calc. 3364.9 obsvd 3364.6).

Large scale cleavage was accomplished by treatment with Reagent K for 2.5 hours. The cleavage mixture was freshly prepared and sparged briefly with N₂ and the reaction mixture blanketed with N₂. The crude unoxidized peptide was purified by semipreparative (10 mm x 25 cm) reverse phase C₁₈ HPLC using acetonitrile:H₂O (0.1% TFA) solvent mixtures (t_R = 12 minutes, 26% acetonitrile isocrat). Disulfide oxidation of the purified peptide was accomplished by addition of 20% DMSO to an aqueous solution of the peptide at pH 7.0 - 8.0 (unbuffered) with stirring at room temperature.³⁵ The progress of the oxidation reaction was monitored by HPLC analysis. Aliquots of the DMSO/water mixture were directly injected on to an analytical (4.6 mm x 15 cm) C₁₈ reverse phase column (t_R = 17.1 minutes, reduced; t_R = 9.4 minutes, oxidized) using a 25% acetonitrile isocrat. The reaction was complete after 9 hours and yielded the oxidized peptide in excellent purity. After removal of the solvent by lyophilization, no additional purification was required. The identity and purity of the oxidized product was confirmed by ESI mass spectrometry (MW calculated 3362.8 amu, observed 3362.9 amu). Treatment of the crude peptide under similar conditions resulted in a much slower oxidation, with complete oxidation requiring in excess of 20 hours, potentially due to the presence of residual thiols from the cleavage reagent.

Peptide **TOX2** was synthesized essentially in the manner described for peptide **TOX1**. The cleavage protocols developed for peptide **NAT1** were adopted. The peptide was purified in the reduced form and the disulfide formed by oxidation with 20% DMSO in H₂O, pH 7.5 - 8.0. Concentration of the peptide from this mixture yielded pure oxidized peptide. The residual DMSO was removed by repeated dissolution in water and lyophilization. The purity and identity of the resulting oxidized peptide was verified by plasma desorption mass spectrometry (C₁₄₅H₂₃₂O₄N₄₂S₂, MW calc. 3307.9 obsvd. 3308±1).

NMR Analysis

Two-dimensional ^1H NMR spectra of **TOX1**, **NAT1**, and **TOX2** were obtained on a 500 MHz AMX Bruker spectrophotometer. In each case purified peptide was dissolved in 440 μl of 90:10 $\text{H}_2\text{O}/\text{D}_2\text{O}$ and the pH adjusted to pH 4.2-4.3 with small aliquots of NaOH or HCl. Spectra of **TOX1** were acquired at 277 K, 281 K, and 300 K, at a peptide concentration of 6.9 mg/440 μl . Peptide **NAT1** was found to be only partially soluble in water and the NMR spectra were acquired at a concentration of 4.4 mg/440 μl (2.7 mM) at 277 K, 281 K, and 300 K. Spectra of **TOX2** were acquired at a concentration of 9.9 mg/464 μl at 277 K, 281 K, 290 K, and 300 K. Complete sequence specific assignments³⁶ (see Tables 2-8, 2-9, and 2-10) for backbone and side chain protons were determined using total correlation spectroscopy (TOCSY),⁶² magnitude mode correlation spectroscopy (mmCOSY), nuclear Overhauser spectroscopy (NOESY),⁶³ and rotating frame nuclear Overhauser spectroscopy (ROESY) at 281 K. Spectral ambiguities at 281 K were resolved by analysis of spectra taken at different temperatures and by comparison of spectra from each peptide. All spectra were processed using Felix or Felix 2.0 software (Biosym, CA). Free induction decays (FID) were multiplied by a phase shifted sine bell apodization function and zero filled (2048 points) prior to Fourier transform. Suppression of the water signal was accomplished by presaturation during the relaxation delay (1.3 s). Amide proton variable temperature coefficients were obtained by determining the spin system and sequential assignments at multiple temperatures. Plots of amide proton chemical shift versus temperature were linear and the slopes corresponded to the amide proton temperature coefficient.

residue	alpha	amide	side chain
Tyr ¹	4.19		3.05, 3.12, 6.81, 7.09
His ²	4.61	8.60	3.14, 3.32, 7.24, 8.61
Lys ³	4.16	8.59	1.70, 1.83, 1.35, 1.45, 2.94
Gln ⁴	4.28	8.60	1.89, 6.98, 7.61
Phe ⁵	4.65	8.54	3.02, 3.12, 7.33, 7.23
Ser ⁶	4.45	8.35	3.76
Val ⁷	4.40	8.37	2.02, 0.94, 0.97
Pro ⁸	4.44		2.31, 1.93, 2.08, 1.99 d 3.93, 3.72
DSer ⁹	4.55	8.44	3.86
His ¹⁰	4.81	8.54	3.14, 3.22, 7.24, 8.61
Thr ¹¹	4.28	8.35	4.10, 1.17
Ile ¹²	4.16	8.44	1.83, 1.17, 1.48, 0.858, 0.84
Ile ¹³	4.14	8.40	1.83, 1.17, 1.48, 0.86, 0.84
Glu ¹⁴	4.34	8.62	1.94, 2.05, 2.32
Arg ¹⁵	4.34	8.62	1.87, 1.79, 1.65, 1.66, 3.20, 7.27
Gly ¹⁶	4.01	8.65	
Ser ¹⁷	4.45	8.44	3.87, 3.92
Gly ¹⁸	3.98	8.63	
Cys ¹⁹	5.00	8.42	2.91, 3.21
Pro ²⁰	4.50		1.94, 2.30, 2.02, 2.08 d 3.72, 3.85
Thr ²¹	4.27	8.25	4.13, 1.21
Val ²²	4.19	8.30	2.09, 0.91, 0.93
Lys ²³	4.63	8.30	1.83, 1.86, 1.45, 1.68, 2.99, 7.61
Pro ²⁴	4.34		1.92, 2.30, 2.00, 2.09. 3.63, 3.84
Gly ²⁵	4.17, 3.74	8.71	
Ile ²⁶	4.25	7.83	1.97, 1.20, 1.39, 0.91, 0.90
Lys ²⁷	4.35	8.55	1.75, 1.43, 1.68, 2.99
Leu ²⁸	4.44	8.41	1.63, 1.65, 0.90, 0.93
Ser ²⁹	4.41	8.72	3.80, 3.90
Cys ³⁰	4.50	8.20	3.13, 3.25

Table 2-7 NMR assignments for TOX1. 500 MHz, 90:10 H₂O:D₂O, pH 4.25, 281K, 6.9mg/440μl, referenced to DMSO (δ 2.72ppm).

residue	alpha	amide	side chain
Tyr ¹	4.176		3.038, 3.124, 6.802, 7.065
His ²	4.597	8.597	3.066, 3.119, 8.515, 7.186
Lys ³	4.628	8.297	2.983, 1.68, 1.448, 1.835, 7.61
Gln ⁴	4.312	8.65	1.986, 1.916
Trp ⁵	4.662	8.489	3.28, 7.242, 10.19, 7.581, 7.096, 7.192, 7.442
Ser ⁶	4.316	8.144	3.745, 3.596
Asp ⁷	4.543	8.226	2.642
Phe ⁸	4.536	8.251	3.081, 7.312, 7.236
Arg ⁹	4.21	8.286	3.085, 1.465, 1.791, 1.655, 7.280
Gly ¹⁰	3.902	7.811	
Thr ¹¹	4.315	8.150	4.162, 1.186
Ile ¹²	4.15	8.37	1.44, 1.156, 1.824, 0.826, 0.844
Ile ¹³	4.14	8.37	1.145, 1.44, 0.850, 0.805
Glu ¹⁴	4.335	8.545	2.30, 1.93, 2.03
Arg ¹⁵	4.303	8.57	1.861, 1.768, 1.61, 1.65, 3.146, 7.237
Gly ¹⁶	4.00	8.60	
Ser ¹⁷	4.446	8.415	3.873, 3.922
Gly ¹⁸	3.975	8.617	
Cys ¹⁹	4.988	8.408	2.911, 3.202
Pro ²⁰	4.49		1.95, 2.29, 2.00, 3.825, 3.704
Thr ²¹	4.267	8.241	4.126
Val ²²	4.18	8.290	2.093, 0.925
Lys ²³	4.628	8.297	1.835, 1.448, 1.680, 2.983, 7.61
Pro ²⁴	4.335		1.91, 2.01, 2.29, 2.00, 3.84, 3.631
Gly ²⁵	4.163, 3.729	8.694	
Ile ²⁶	4.255	7.821	1.96, 1.381, 1.214, 0.907, 0.828
Lys ²⁷	4.34	8.550	1.758, 1.735, 1.409, 1.672, 2.983, 7.58
Leu ²⁸	4.44	8.411	1.641, 1.606, 1.636, 0.883, 0.9332
Ser ²⁹	4.540	8.473	3.863
Cys ³⁰	4.502	8.201	3.245, 3.119

Table 2-8 NMR assignments for peptide **NAT1**. 500 MHz, 281 K, 4.4mg/490µl
90:10 H₂O:D₂O, pH 4.2., referenced to DMSO (δ 2.72 ppm)

residue	alpha	amide	side chain
Tyr ¹	4.185		3.033, 3.101, 6.806, 7.070
His ²	4.599	8.60	3.06, 3.129
Lys ³	4.144	8.579	1.643, 1.265, 1.592, 2.803, 7.538
Gln ⁴	4.292	8.607	1.925, 2.03
Trp ⁵	4.697	8.4664	3.278, 3.207, 7.224, 10.172, 7.595, 7.105, 7.167, 7.432
Ser ⁶	4.424	8.217	3.886, 3.797
Val ⁷	4.345	8.193	2.000, 0.963, 0.914
Pro ⁸	4.419		2.316, 1.93, 2.087, 1.991, 3.93, 3.70
DSer ⁹	4.403	8.7163	3.886, 3.797
His ¹⁰	4.78	8.53	3.10, 3.274, 8.53, 7.189
Thr ¹¹	4.282	8.330	4.089, 1.163
Ile ¹²	4.153	8.423	1.826, 1.469, 1.156, 0.8552, 0.8465
Ile ¹³	4.145	8.385	1.813, 1.452, 1.162, 0.8606, 0.8052
Glu ¹⁴	4.312	8.616	2.04, 1.936, 2.29
Arg ¹⁵	4.329	8.585	1.868, 1.782, 1.653, 3.187, 7.257
Gly ¹⁶	4.01	8.647	
Ser ¹⁷	4.453	8.438	3.875, 3.928
Gly ¹⁸	3.973	8.647	
Cys ¹⁹	4.997	8.422	2.914, 3.206
Pro ²⁰	4.505		2.305, 1.923, 2.01, 3.825, 3.714
Thr ²¹	4.129	8.257	4.129, 1.207
Val ²²	4.190	8.310	2.091, 0.926
Lys ²³	4.633	8.304	1.840, 1.469, 1.689, 2.985, 7.61
Pro ²⁴	4.339		2.229, 1.904, 1.9968, 2.09, 3.84, 3.633
Gly ²⁵	4.168, 3.735	8.703	
Ile ²⁶	4.255	7.828	1.961, 1.378, 1.214, 0.904, 0.823
Lys ²⁷	4.344	8.563	1.775, 1.426, 1.672, 2.992, 7.58
Leu ²⁸	4.437	8.421	1.641, 1.606, 1.635, 0.895, 0.937
Ser ²⁹	4.543	8.483	3.861
Cys ³⁰	4.502	8.206	3.245, 3.128

Table 2-9 NMR assignments for peptide **TOX2**. 500 MHz, 281 K, 9.9mg/490µl 90:10 H₂O:D₂O, pH 4.2., referenced to DMSO (δ 2.72 ppm)

References

- (1) Betz, S.F.; Raleigh, D.P.; DeGrado, W.F. "De Novo Protein Design: from Molten Globules to Native-like States," *Curr. Opin. Struct. Biol.* **1993**, 3, 601-610.
- (2) Betz, S. F.; Bryson, J. W.; DeGrado, W. F. "Native-like and Structurally Characterized Designed α -helical Bundles," *Current Opin in Structural Biology* **1995**, 5, 457-463.
- (3) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. "Protein Design: A Hierarchic Approach," *Science* **1995**, 270, 935-941.
- (4) Shoemaker, K. R.; Kim, P. S.; Brems, D. N.; Marquese, S.; York, E. J.; Chaiken, I. M.; Stewart, J. M.; Baldwin, R. L. "Nature of the Charged-group Effect on the Stability of the C-peptide Helix," *Proc. Natl. Acad. Sci. U.S.A.* **1985**, 82, 2349-2353.
- (5) Kim, P. S.; Baldwin, R. L. "A Helix Stop Signal in the Isolated S-peptide of Ribonuclease A," *Nature* **1984**, 307, 329.
- (6) Marquese, S.; Robbins, V. H.; Baldwin, R. L. "Unusually Stable Helix Formation in Short Alanine-Based Peptides," *Proc. Natl. Acad. Sci. U.S.A.* **1986**, 86, 5286-5290.
- (7) Padmanabhan, S.; Marquese, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L. "Relative Helix-Forming Tendencies of Nonpolar Amino Acids," *Nature* **1990**, 344, 268-270.
- (8) Lyu, P.; Sherman, J. C.; Chen, A.; Kakkenback, N. R. " α -Helix Stabilization by Natural and Unnatural Amino Acids with Alkyl Side Chains," *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 5317-5320.
- (9) O'Neil, K. T.; DeGrado, W. F. "A Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occurring Amino Acids," *Science* **1990**, 250, 646-651.

- (10) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. "Tests of the Helix Dipole Model for Stabilization of α -helices," *Nature* **1987**, 326, 563-567.
- (11) Nelson, J. W.; Kallenbach, N. R. "Persistence of the α -helix Stop Signal in the S-Peptide in Trifluoroethanol Solutions," *Biochemistry* **1989**, 28, 5256-5261.
- (12) Marqusee, S.; Baldwin, R. L. "Helix Stabilization by $\text{Glu}^- \dots \text{Lys}^+$ Salt Bridges in short Peptides of *de novo* Design," *Proc. Natl. Acad. Sci. U.S.A.* **1987**, 84, 5527-5531.
- (13) Merutka, G.; Stellwagen, E. "Effect of Amino Acid Ion Pairs on Peptide Helicity," *Biochemistry* **1991**, 30, 1591-1594.
- (14) DeGrado, W. F. "Design of Peptides and Proteins," *Adv. Prot. Chem.* **1988**, 39, 51-124.
- (15) Osterhout, J. J.; Handel, T.; Na, G.; Toumadje, A.; Long, R. C.; Connolly, P. J.; Hoch, J. C.; Johnson Jr., W. C.; Live, D.; DeGrado, W. F. "Characterization of the Structural Properties of $\alpha 1\text{B}$, a Peptide Designed To Form a Four-Helix Bundle," *J. Am. Chem. Soc.* **1992**, 114, 331-337.
- (16) O'Shea, E. K.; Lumb, K. J. ; Kim, P. S. "Peptide 'Velcro': Design of a Heterodimeric Coiled Coil," *Current Biology* **1993**, 3, 658-667.
- (17) Myszka, D. G.; Chaiken, I. M. "Design and Characterization of an Intramolecular Antiparallel Coiled Coil Peptide," *Biochemistry* **1994**, 33, 2363-2372.
- (18) Fezoui, Y.; Weaver, D.L.; Osterhout, J.J. "*De novo* Design and Structural Characterization of an α -helical Hairpin Peptide: A Model System For the Study of Protein Folding Intermediates," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 3675-3679.
- (19) Kuroda, Y. "A Strategy for the *de novo* Design of Helical Proteins with Stable Folds," *Protein Eng.* **1995**, 8, 97-101.

- (20) Olofsson, S.; Johansson, G.; Baltzar, L. "Design, Synthesis, and Solution Structure of a Helix-Loop-Helix Dimer- a Template of the Rational Design of Catalytically Active Polypeptides," *J. Chem. Soc. Perkin Trans. II* **1995**, 2047-2057.
- (21) Pessi, A.; Bianchi, E.; Crameri, A.; Venturini, S.; Tramontano, A.; Sollazzo, M. "A Designed Metal-binding Protein with a Novel Fold," *Nature* **1993**, 362, 367-369.
- (22) Quinn, T. P. ; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C. "Betadoublet: *De novo* Design, Synthesis, and Characterization of a β -sandwich Protein," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 8747-8751.
- (23) Yan, Y.; Tropsha, A.; Hermans, J.; Erickson, B. W. "Free Energies for Refolding of the Common β -turn into the Inverse-common β -turn: Simulation of the Role of D/L Chirality.," *Proc. Natl. Acad. Sci. U.S.A.* **1993**, 90, 7898-7901.
- (24) Hecht, M. H. "*De novo* Design of β -sheet Proteins," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 8729-8730.
- (25) Smith, D.D.S.; Pratt, K.A.; Sumner, I.G.; Henneke, C.M. "Greek Key Jellyroll Protein Motif Design: Expression and Characterization of a First-generation Molecule," *Protein Eng.* **1995**, 8, 13-20.
- (26) Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide," *J. Am. Chem. Soc.* **1992**, 114, 3182.
- (27) Menez, A.; Bontems, F.; Roumestand, C.; Gilquin, B.; Toma, F. "Structural Basis for Functional Diversity of Animal Toxins," *Proceedings - Royal Society of Edinburgh* **1992**, 99, 83-103.
- (28) Low, B. W.; Cornfield, P. W. R. "Erabutoxin b-Structure/Function Relationships Following Initial Refinement at 0.140-nm Resolution," *Eur. J. Biochem.* **1986**, 161, 579-587.

- (29) Love, R. A.; Stroud, R. M. "The Crystal Structure of α -Bungarotoxin at 2.5Å resolution: Relation to Solution Structure and Binding to Acetylcholine Receptor," *Protein Eng.* **1986**, *1*, 37-46.
- (30) Le Goas, R.; LaPlante, S. R.; Kikou, A.; Delsuc, M.-A.; Guittet, E.; Robin, M.; Charpentier, I.; Lallemand, J.-Y. " α -Cobratoxin: Proton NMR Assignments and Solution Structure," *Biochemistry* **1992**, *31*, 4867-4875.
- (31) Walkinshaw, M. D.; Saenger, W.; Maelicke, A. "Three-dimensional Structure of the "long" Neurotoxin from Cobra Venom," *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 2400-2404.
- (32) Zinn-Justin, S.; Roumestand, C.; Gilquin, B.; Bontems, F.; Menéz, A. "Three-Dimensional Solution Structure of a Curaremimetic Toxin From *Naja nigricollis* Venom: A Proton NMR and Molecular Modeling Study," *Biochemistry* **1992**, *31*, 11335-11347.
- (33) Rose, G. D.; Gierasch, L. M.; Smith, J. A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.* **1985**, *37*, 1-109.
- (34) Wilmot, C. M.; Thornton, J. M. "Analysis and Prediction of the Different Types of β -Turns in Proteins," *Journal of Molecular Biology* **1988**, *203*, 221-232.
- (35) Tam, J. P.; Wu, C.-R.; Liu, W.; Zhang, J.-W. "Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide Scope and Applications," *J. Am. Chem. Soc.* **1991**, *113*, 6657-6662.
- (36) Wüthrich, D. *NMR of Proteins and Nucleic Acids*; John Wiley and sons: New York, 1986.
- (37) 9050 Plus Operators Manual 1.0
- (38) Wade, J. D.; Bedford, J.; Sheppard, R. C.; Tregear, G. W. "DBU as an N^α -Deprotecting Reagent for the Fluorenyl-methoxycarbonyl Group in Continuous Flow Solid-Phase Peptide Synthesis," *Peptide Research* **1991**, *4*, 184-199.

- (39) Richardson, J. S. "The Anatomy and Taxonomy of Protein Structure," *Adv. Prot. Chem.* **1981**, 34, 167-339.
- (40) Betz, S. F. "Disulfide Bonds and the Stability of Globular Proteins," *Protein Sci.* **1993**, 2, 1551-1558.
- (41) Matsumura, M.; Signor, G.; Matthews, B. W. "Substantial Increase of Protein Stability by Multiple Disulphide Bonds," *Nature* **1989**, 342, 291-293.
- (42) Matsumura, M.; Becktel, W. J.; Levitt, M.; Matthews, B. W. "Stabilization of Phage T4 Lysozyme by Engineered Disulfide Bonds," *Proc. Natl. Acad. Sci. U.S.A.* **1989**, 86, 6565-6566.
- (43) Villafranca, J. E.; Howell, E. E.; Oatley, S. J.; Xuong, N.; Kraut, J. "An Engineered Disulfide Bond in Dihydrofolate Reductase," *Biochemistry* **1987**, 26, 2182-2189.
- (44) Mitchinson, C.; Wells, J. A. "Protein Engineering of Disulfide Bonds in Subtilisin BPN'," *Biochemistry* **1989**, 28, 4807-4815.
- (45) Creighton, T. E. "Disulphide Bonds and Protein Stability," *BioEssays* **1988**, 8, 57-63.
- (46) Katz, B.; Kossiakoff, A. "The Crystallographically Determined Structures of Atypical Strained Disulfides Engineered into Subtilisin," *J. Biol. Chem.* **1986**, 261, 15480-15485.
- (47) Wetzel, R. "Harnessing Disulfide Bonds Using Protein Engineering," *Trends in Biol. Sci.* **1987**, 12, 478-482.
- (48) Zhou, N. E.; Kay, C. M.; Hodges, R. S. "Disulfide Bond Contribution to Protein Stability: Positional Effects of Substitution in the Hydrophobic Core of the Two-Stranded α -Helical Coiled-Coil," *Biochemistry* **1993**, 32, 3178-3187.
- (49) Fisher, S. L., "Unnatural Amino Acids in the Synthesis and Semisynthesis of Metalloprotein Motifs," Ph.D. Thesis, California Institute of Technology, 1993.

- (50) Fasman, G.D. "The Development of the Prediction of Protein Structure," in *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G.D. Eds.; Plenum Press: New York, 1989; pp 193-316.
- (51) Bontems, F.; Roumestand, C.; Boyot, P.; Gilquin, B.; Doljansky, Y.; Menéz, A.; Toma, F. "Three-Dimensional Structure of Natural Charybdotoxin in Aqueous Solution by ^1H -NMR," *Eur. J. Biochem.* **1991**, *196*, 19-28.
- (52) Bontems, F.; Gilquin, B.; Roumestand; Menéz, A.; Toma, F. "Analysis of Side-Chain Organization on a Refined Model of Charybdotoxin: Structural and Functional Implications," *Biochemistry* **1992**, *31*, 7757-7764.
- (53) Kraulis, R. J.; Clore, G. M.; Nigles, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. "Determination of the Three-Dimensional Solution Structure of the C-Terminal Domain of Cellobiohydrolase From *Trichoderma reesei*. A Study Using Magnetic Resonance and Hybrid Distance Geometry-Dynamical Simulated Annealing," *Biochemistry* **1989**, *28*, 7241-7257.
- (54) Klug, A.; Schwabe, J. W. R. "Zinc Fingers," *FASEB J.* **1995**, *9*, 597-604.
- (55) Berg, J.M. "Zinc Finger Domains: Hypotheses and Current Knowledge," *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405-421.
- (56) Bodanszky, M. *Principles of Peptide Synthesis*; Springer-Verlag: Berlin, 1993.
- (57) Van Woerkom, W. J.; Van Nispen, J. W. "Difficult Couplings in Stepwise Solid Phase Peptide Synthesis: Predictable or Just a Guess?," *Int. J. Pep. Protein Res.* **1991**, *38*, 103-113.
- (58) King, D. S.; Fields, C. G.; Fields, G. B. "A Cleavage Method which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis," *Int. J. Pep. Protein Res.* **1990**, *36*, 255-266.
- (59) Ellman, G. L. "Tissue Sulfhydryl Groups," *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.

- (60) Johnson, T.; Sheppard, R. C. "Resin Effects in Solid-phase Peptide Synthesis. Enhanced Purity of Tryptophan-containing Peptides through Two-step Cleavage of Side Chain Protecting Groups and Peptide-Resin Linkage," *J. Chem. Soc. Chem. Comm.* **1991**, 1653-1655.
- (61) Van Wandelen, C.; Zeikus, R.; Tsou, D., "Cleavage, Deprotection, and Isolation of Peptides after Fmoc Synthesis," Chemistry Update, MilliGen/Biosearch, Division of MILLIPORE
- (62) Bax, A.; Davis, D. G. "MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy," *J. Magn. Res.* **1985**, 65,
- (63) States, D. J.; Haberkorn, R. A.; Ruben, D. J. "A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants," *J. Magn. Res.* **1982**, 48, 286-292.

Chapter 3

Evolution of a metal-independent $\beta\beta\alpha$ motif based on
the zinc finger domains

Introduction

Most naturally occurring small peptides and protein domains (<40 residues) require either disulfide bridges or metal ligation to form a defined three-dimensional architecture in aqueous solution. Our approach to protein design, introduced in the previous chapter, is to modify these existing natural systems and eliminate the need for cross-links, in order to create versatile structural templates for the design of catalytic biomolecules. One appealing target for this approach is the $\beta\beta\alpha$ motif exemplified by the zinc finger domains.¹ These polypeptides form small independent folding units which adopt a well defined tertiary structure *only* in the presence of divalent metal cations; the peptides are unstructured in the absence of metal.^{2, 3} Sequential arrays of these domains are present in many DNA binding proteins, and are responsible for DNA binding affinity and specificity.¹ Each domain has the approximate sequence given in Figure 3-1A, and binds one zinc atom in a tetrahedral geometry.^{1, 4} Metal binding by zinc fingers is mediated by the conserved metal binding cysteine and histidine residues (Figure 3-1A, residues in bold). In addition to the metal coordinating residues, three hydrophobic amino acids are conserved among zinc finger sequences. These residues (Figure 3-1A, underlined residues), form a hydrophobic cluster in the final folded motif and are critical to the adoption of a defined three dimensional structure.

A
 -(Tyr, Phe) -X-Cys-X_{2,4}-Cys-X₃-Phe-X₅-Leu-X₂-His-X_{3,4}-His-X-

B.

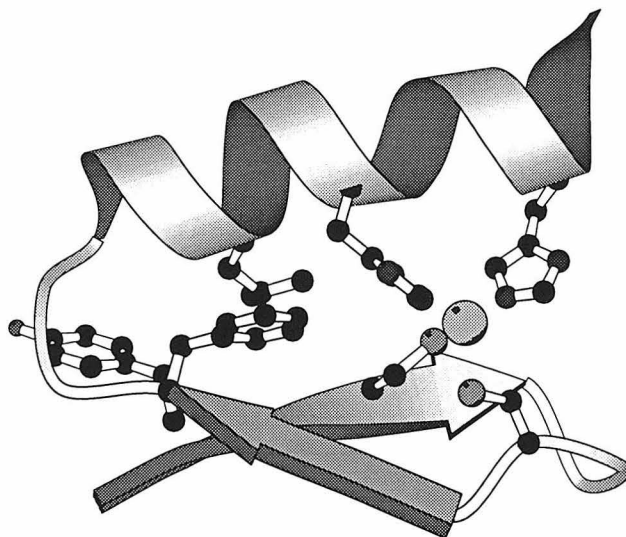


Figure 3-1 A. The conserved sequence of the zinc finger domains. B. The $\beta\beta\alpha$ architecture of a natural zinc finger domain of the protein Zif268.⁵ The crystal structure of a fragment of Zif268 containing three zinc finger domains bound to DNA has been solved. The amino acids which form the metal binding site and the hydrophobic cluster are shown.

Several individual zinc finger sequences have been chemically synthesized and the solution state structures solved by ^1H nuclear magnetic resonance spectroscopy (NMR) in the presence of zinc.^{6, 7} Additionally the crystal structures of arrays of zinc finger domains (2-5) co-crystallized with DNA have been determined.^{5, 8, 9} The structure of a zinc finger peptide-metal complex consists of a $\beta\beta\alpha$ topology; two beta strands connected by an atypical loop form the β -hairpin which packs against the alpha helix. Interactions between the helix

and the sheet are mediated at one end by the metal site and at the other by the hydrophobic cluster. The metal binding site is formed from two cysteines at the end of the β -hairpin and two histidine residues from the α -helix, while the hydrophobic cluster consists of two aromatic residues (typically phenylalanine and tyrosine) in the β -hairpin and a hydrophobic amino acid (typically leucine) from the helix. This architecture is represented schematically for a natural zinc finger domain in Figure 3-1B.

The zinc finger motif has several features which make it an ideal design target. Unlike the erabutoxin b fragment presented in the last chapter, zinc fingers are small and compact "mini-globular" proteins, which fold independently of any other protein structure, requiring only metal ligation. The presence of an anti-parallel β -sheet permits the use of structurally defined β -turns to act as structural nucleation elements and promote β -hairpin formation in the absence of metal. The small size of these domains makes them amenable to solid phase peptide synthesis which allows for the incorporation of unnatural amino acids to control structure, implement function, or incorporate a chromophore which can act as a reporter group during the design process. Metal binding by these domains can be used as a sensitive probe of the pre-existing structure. The circular dichroism (CD) signature of several zinc fingers has been reported;^{2, 10-12} therefore CD can be used as a rapid and sensitive method for probing the effect of metal binding on secondary structure in early designs (Figure 3-2). Additionally, the limited size and monomeric nature of this motif facilitates detailed structural analysis using 2D ^1H nuclear magnetic resonance (NMR) experiments.

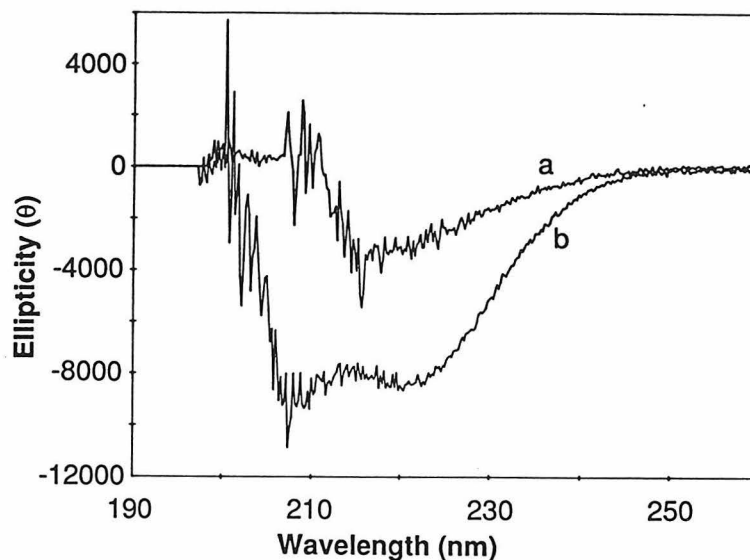


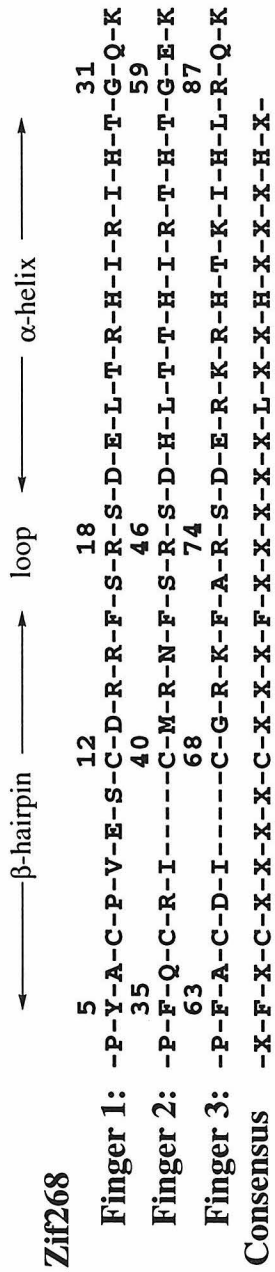
Figure 3-2 The CD spectrum of a natural zinc finger, ADR1, is shown both as the free peptide (a) and the complex with Zn²⁺ (b).¹⁰ Reprinted with permission from Parraga, G.; Horvath, S.; Eisen, A.; Taylor, W.; Hood, L.; Young, E.; Klevit, R.E. "Zinc-Dependent Structure of a Single-Finger Domain of Yeast ADR1," *Science* **1988**, 241, 1489-1492. Copyright 1988 American Association for the Advancement of Science.

An initial investigation of this motif as a potential target for the development of catalytic metalloprotein models was initiated by Richard Cheng in the Imperiali group. In addition to representing a target for the design of metalloenzyme mimics, it was envisioned that a metal *independent* version of this motif would form a general structural template for the construction of a variety functional polypeptides. Therefore, using an iterative design process in a joint project with Richard Cheng, the evolution of the zinc finger motif into a metal independent structure was undertaken.

Results and Discussion

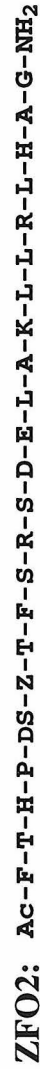
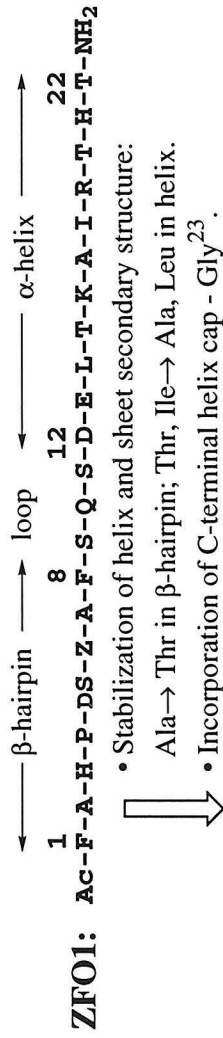
The sequential design of three zinc finger based peptides is outlined in Figure 3-3. The sequences are based on the consensus zinc finger sequence¹ and Zif268.⁵ In this process, each design was structurally analyzed by CD and NMR (of the later two designs) and the results obtained from these analyses were used to determine the sequence modifications incorporated into subsequent generations.

Figure 3-3 (following page) Design of the first three generations of zinc finger peptides. The consensus zinc finger sequence¹ and that of Zif268, a DNA binding protein consisting of three domains which has been co-crystallized with DNA, are shown for comparison.⁵ (Standard one-letter abbreviations for the natural amino acids; Z=Fen)



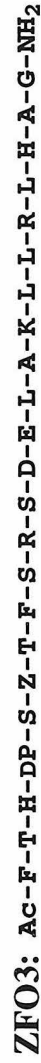
$\Downarrow$

- Redesign of β -hairpin to incorporate type II turn, shortening loop by 2-4 residues.
- Manipulation of ligand sphere to include reporter functionality
(Cys₂His₂ $\rightarrow$ His₂Fen).



$\Downarrow$

- Replace type II with type II' turn: promote β -hairpin formation
Pro⁴DSer⁵ $\rightarrow$ DPro⁴-Ser⁵.



ZFO1: Incorporation of a Reporter Group and a Type II Reverse Turn.*Design*

The first design incorporated two important modifications into the zinc finger sequence; the β -hairpin and metal binding site were redesigned to include a type II β -turn and an unnatural metal binding amino acid, respectively. In natural zinc finger peptides, an atypical loop connects the β -strands. This loop appears to be flexible and poorly defined in the first finger of the crystal structure of Zif268,⁵ as well as the NMR structure of the single zinc finger, Xfin.⁷ A key feature of our design process was to delegate an active role to this part of the polypeptide sequence. Conformational studies of proteins and polypeptides have indicated that in some instances the β -turn may serve as a nucleating element in the folding process.¹³⁻¹⁸ Accordingly, several recent computational¹⁹⁻²¹ and spectroscopic studies²²⁻²⁴ have focused on establishing the stability of the various β -turn types. In particular, heterochiral type II turns have exhibited an inherent stability in theoretical and experimental investigations.^{19, 22, 23} Thus, in the first generation design, **ZFO1**, the β -hairpin region was engineered to incorporate a heterochiral type II β -turn forming sequence, Pro-DSer, while retaining only the two key conserved hydrophobic aromatic amino acids of the natural system at either end of the β -hairpin. It has been shown that with the exception of the metal binding ligands and the three highly conserved hydrophobic core residues (Phe¹, Phe⁸, and Leu¹⁴ in our system), alanine substitutions are tolerated at nearly every position in the natural zinc finger motif.²⁵ Therefore for the initial design, peptide **ZFO1**, the central residues of the redesigned β -strands (positions 2 and 7) were alanines. The helix of **ZFO1** was based on the sequence of the Zif268 protein.⁵

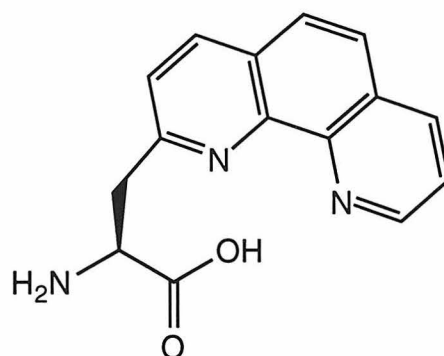


Figure 3-4 3-(1,10-phenanthrol-2-yl)-L-alanine or Fen.

The metal binding site of natural zinc fingers was also targeted for modification in the first generation design. For natural zinc fingers, assessment of divalent zinc coordination to the imidazole and thiolate ligands is complicated by the spectroscopically silent d^{10} Zn^{2+} ion.²⁶ Therefore the elucidation of metal coordination for these motifs generally relies on observation of the tetrahedral coordination of Co^{2+} to these peptides.^{2, 27} Metal binding by natural zinc fingers is spectroscopically observable because of the d-d transition of the peptide- Co^{2+} complex. Thus, information on Zn^{2+} binding is only determined indirectly by competitive titration of the Co^{2+} complex with Zn^{2+} .²⁷ To overcome this limitation and enable the facile assessment of metal binding properties for peptides generated during the design process, the metal coordination site of the natural zinc fingers was modified to include an unnatural bidentate amino acid [3-(1,10-phenanthrol-2-yl)-L-alanine (Fen)] (Figure 3-4).^{28, 29} The metal binding affinities and stoichiometries for a variety of divalent metal ions can be determined independent of coordination geometry by monitoring the UV spectrum of the phenanthroline chromophore. CD studies of the Fen moiety also allow the assessment of metal ion induced structural changes. The native

Cys₂His₂ site was therefore converted to a His₂Fen site containing exclusively nitrogen-based ligands. This modification required the replacement of the two Cys ligands within the β -hairpin of Zif268 with His and Fen respectively. The central histidine residue of the α -helix was replaced with alanine (position 17) to retain a four coordinate site. The removal of the sulfhydryl-containing amino acids also eliminates the need for special handling techniques normally required to prevent the oxidation of these residues to the corresponding disulfide.³⁰ Therefore, the Fen residue acts a reporter group during the design process allowing for the facile assessment of metal binding by designed peptides.

CD Analysis

ZFO1 was analyzed by circular dichroism and UV-Vis spectroscopy. CD spectroscopy was used to investigate the metal binding behavior of **ZFO1**. These studies indicated that the structure of **ZFO1** was metal dependent (Figure 3-5 A). Although the CD signature of the peptide in the absence of metal was mainly random coil,³¹ the addition of Zn²⁺ increased the overall secondary structure content. The resulting spectroscopic signature of the peptide-Zn²⁺ complex, although considerably weaker, is reminiscent of natural zinc fingers.^{10, 11} Furthermore, metal-binding titrations monitoring the UV signal of the phenanthroline chromophore, demonstrated the formation of a 1:1 **ZFO1**-Zn²⁺ complex ($K_D = 0.65 \pm 0.60 \mu\text{M}$). Concurrent with the increase in secondary structure, the addition of Zn²⁺ also produced a significant CD signal in the near-UV (250-300 nm) region. This signal was indicative of the formation of a chiral metal complex involving the Fen residue (Figure 3-5B). The CD spectrum of uncomplexed peptide **ZFO1** was invariant with pH (4.0-7.0) and peptide concentration (14 μM -1.3 mM) demonstrating the absence of aggregation behavior.

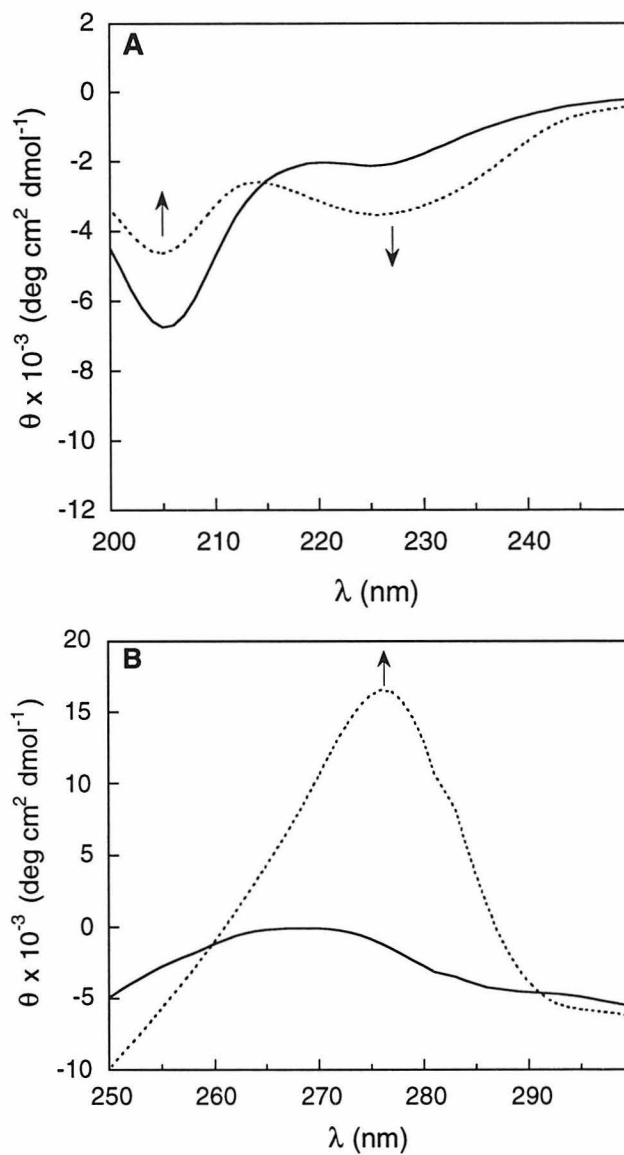


Figure 3-5 CD analysis of **ZFO1**. **A.** Far UV CD spectra of **ZFO1**. The solid line represents the peptide in the absence of metal, the dashed line the presence of Zn^{2+} . The spectra are plotted in mean residue ellipticity. **B.** Near UV CD spectra of **ZFO1** in the presence (solid line) and absence (dashed line) of Zn^{2+} . The spectra are normalized to the concentration of Fcn chromophore.

ZFO2: Increasing the Inherent Secondary Structure.

Design

Towards the goal of a preorganized metal independent structural motif, the second generation design, **ZFO2**, included amino acid substitutions selected to increase inherent secondary structure.³² The alanine residues in the β -hairpin region (positions 2, and 7) were replaced with the β -branched amino acid threonine, which has a higher propensity to adopt an extended or sheet-like conformation.³² These positions are located on the solvent-exposed exterior of the sheet and threonine also imparts hydrophilicity to this face. In addition, the β -branched amino acids in the α -helix were replaced with residues exhibiting higher helical propensities (Leu or Ala), and the sequence was lengthened by one residue to include a glycine at the end of the helix. This residue is often found at the C-terminus of helices and is thought to serve as a helix “stop signal” or helix “cap”.^{33, 34}

CD Analysis

CD analysis of **ZFO2** revealed that the secondary structure of **ZFO2** was considerably more preorganized than that of **ZFO1**, although it was still metal dependent. The increased negative ellipticity of this peptide at 222 nm was consistent with greater inherent helicity in the absence of metal (Figure 3-6A). Nevertheless, addition of metal resulted in spectral changes consistent with the formation of a more structured peptide. The formation of a 1:1 complex with Zn^{2+} was confirmed by UV-Vis spectroscopy ($K_D = 0.89 \pm 0.36 \mu\text{M}$). The CD signature was not dependent on peptide concentration (15 μM -1.5 mM) or pH (4.5-8.1). Appearance of a distinct CD signal in the near-UV range (250-350 nm) upon addition of Zn^{2+} indicated the formation of a chiral complex involving the Fen residue similar to peptide **ZFO1** (Figure 3-6B).

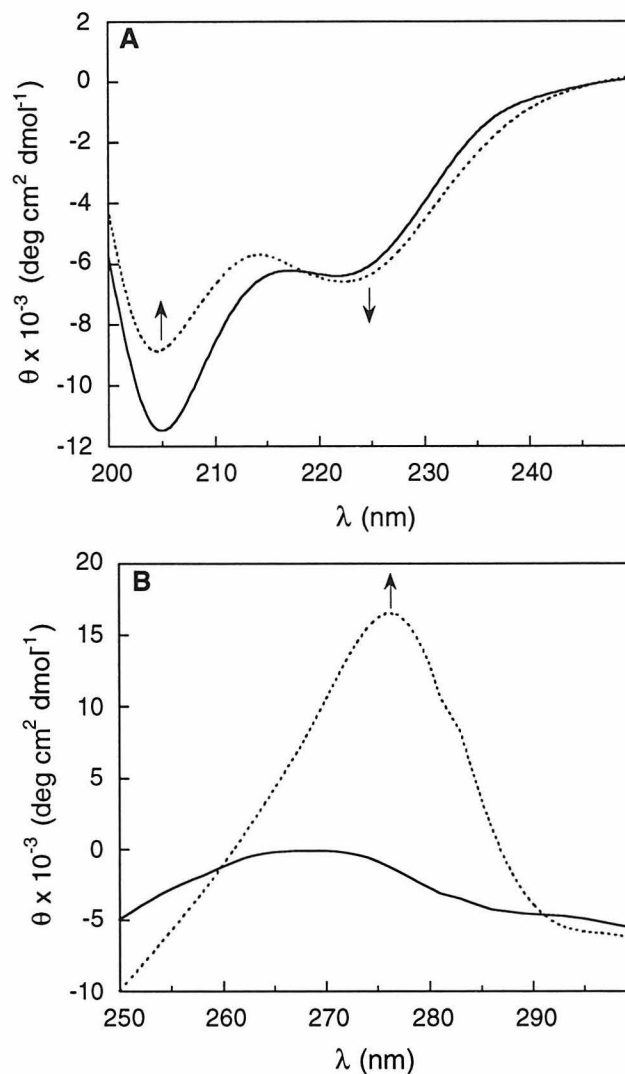
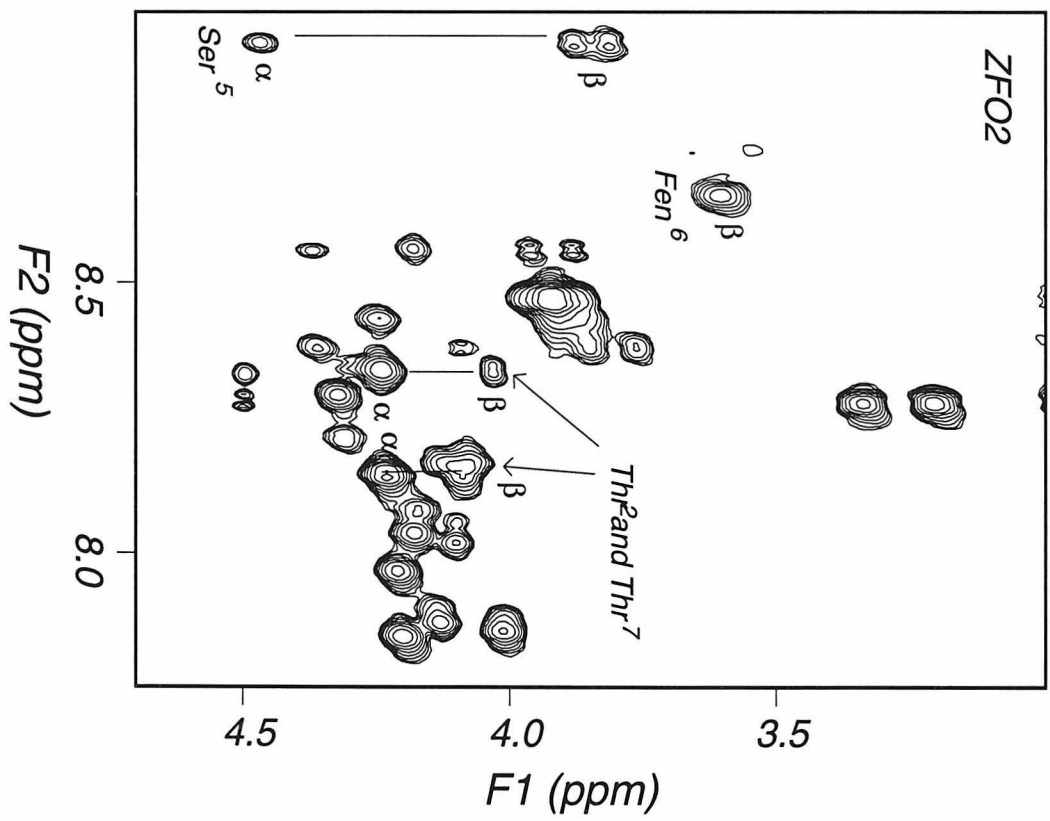
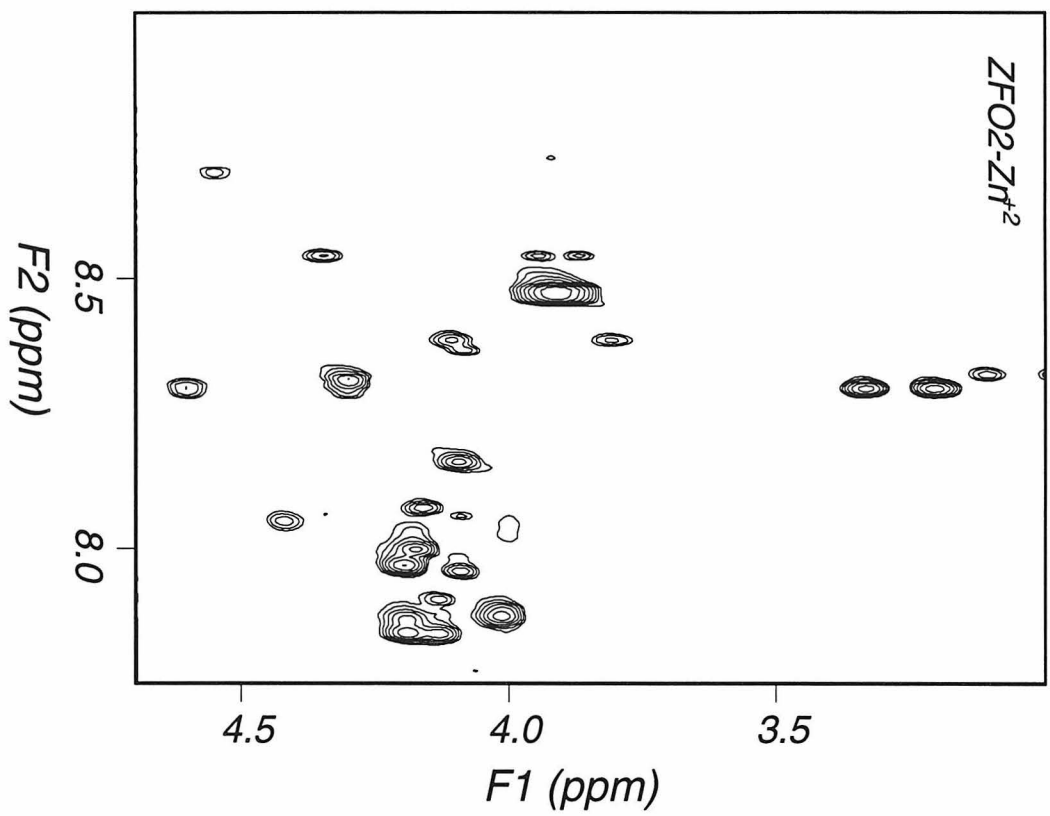


Figure 3-6 CD analysis of **ZFO2**. **A.** Far-UV CD spectra of **ZFO2**. The solid line represents the peptide in the absence of metal, the dashed line after the addition of Zn^{2+} . The spectra are plotted in the mean residue ellipticity. **B.** Near-UV CD spectrum of **ZFO2** in the presence (solid line) and absence of (dashed line) Zn^{2+} . The spectra are normalized to the concentration of Fcn chromophore.

NMR Analysis

The CD spectrum of **ZFO2** strongly resembled those of native zinc fingers, therefore the structure of **ZFO2** was investigated in the presence and the absence of metal by NMR techniques. Analysis of the 2D NMR spectra of **ZFO2** in the presence of metal revealed that although spin systems corresponding to residues in the helix were visible, most residues in the turn, surrounding Fen⁶, were not observed. This presumably resulted from severe signal broadening due to the rapid exchange of Zn²⁺ with the nitrogen containing ligands, especially the Fen residue, of the peptide metal site at the pH required for NMR studies (pH 4.5 - 5.0). In contrast, these spin systems are clearly visible in spectra taken before the addition of metal. Figure 3-7 illustrates the differences in the 2D TOCSY spectra of **ZFO2** taken in the presence and the absence of zinc.

Figure 3-7 (following page) Similar regions TOCSY spectra of the **ZFO2**-Zn²⁺ complex and the uncomplexed peptide are shown. The spectra show crosspeaks between the amide and the alpha or beta protons of all amino acid spin systems. Resonances for several residues which are visible in the uncomplexed peptide but absent from the Zn²⁺ complex are highlighted.



Although the analysis of the structure of the **ZFO2**-Zn⁺² complex by NMR was not feasible, the CD spectra of **ZFO2** indicated the presence of considerable structure in this peptide before the addition of metal. Therefore, the investigation of the structure of the uncomplexed peptide by NMR was initiated. These studies revealed that the structure of uncomplexed **ZFO2** was not adequately well defined, which resulted in insufficient dispersity in the NMR resonances on a 500 MHz instrument. Thus, a complete sequence specific assignment of the peptide was not feasible. The sequential assignment of the helical region (11-22) containing four leucines was particularly difficult. Sequential $d_{\alpha N}(i, i+1)$ are usually weak or absent within this element of secondary structure,³⁵ and sequential assignment therefore relies on the observation of $d_{NN}(i, i+1)$ which should be strong *if* the residues were in a helical conformation. Although several d_{NN} NOEs were observed, there was considerable overlap in the amide region of the spectra for these residues, complicating definitive assignment of these NOEs to particular residues. Assignments based on $d_{NN}(i, i+1)$ NOEs within a helix can be verified by the observation of $d_{\alpha N}(i, i+3)$ NOEs, although observation of such NOEs is more critically dependent on the presence of a defined helix.³⁵ These NOEs were very weak or absent in the NOESY spectrum of **ZFO2** which resulted in some ambiguity in the sequence specific assignment of these residues. Additionally sequential assignments within in the sheet region were difficult due to degeneracy in the sequence and partial overlap of the water signal with several alpha protons for the aromatic amino acids (Phe¹, His³, Fen⁶, Phe⁸). Despite the lack of complete sequential assignment, the chemical shifts for the ring protons of the two Phe residues could be determined although not attributable to a specific spin system. If the peptide adopted the tertiary structure expected for this motif then interaction between the side chain of a leucine residue (Leu¹⁴) and *both* the

Phe rings should be observed (due to the formation of the hydrophobic cluster), despite the lack of definitive sequential assignment. The absence of such NOEs indicated that this peptide did not have the desired tertiary structure.

ZFO3: Incorporation of a type II' turn

Design

While the results of the CD studies for the second generation design were encouraging, it was evident from the CD and the NMR analysis that the **ZFO2** did not have the desired tertiary structure in the absence of metal. In order to address this issue, the third design examined the β -hairpin region. Specifically the incorporation of a type II' β -turn was investigated. Studies of the protein database³⁶ have found that tight β -hairpins primarily involve type I' or type II' β -turns, despite a general prevalence of type I and type II β -turns in protein structures.³⁷ In other words, type I' and type II' turns more frequently lead directly into the hydrogen bonding pattern associated with β -sheets than other turn types. To further investigate the difference between type II and type II' turns with respect to the geometry of the resulting sheet, molecular modeling studies were performed with ideal turn and sheet dihedral angles. The analysis revealed that the overall topologies of the turns are similar. However, the more common type II turn affords planar but divergent strands, while the type II' turn leads to convergent β -strands (Figure 3-8). Thus, a type II' turn might be more effective in promoting β -hairpin formation within the $\beta\beta\alpha$ motif. Therefore, the third generation design **ZFO3** incorporated a type II' reverse turn.

The (ϕ , ψ) dihedral angles associated with the central two residues of the type II' β -turn [(60°, -120°), (-80°, 0°)] are the inverse of those defining type II β -turn formation [(-60°, 120°), (80°, 0°)].³⁶ This fact leads to the opposite chirality preference for the type II' β -turn relative to the type II β -turn. Specifically, the

LPro⁴-DSer⁵ sequence present in the first and second generation designs was replaced with DPro⁴-LSer⁵ for **ZFO3**.

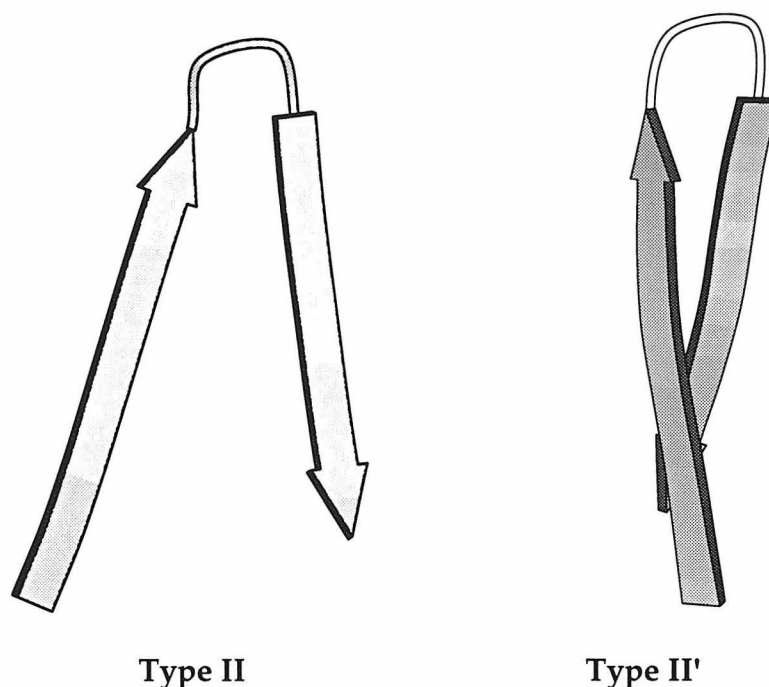


Figure 3-8 Results of modeling studies illustrating the differing effects of type II (on the left) and type II' (on the right) turns on β -sheet formation.

CD Analysis

The CD spectrum of uncomplexed **ZFO3** exhibited a signature similar to that seen for peptide **ZFO2**, however, in this case the spectrum changed very little upon addition of Zn²⁺, indicating that the secondary structure for **ZFO3** was not metal dependent (Figure 3-9A). As with previous designs, the CD spectrum in the absence of metal was invariant with changes in pH (pH 4.5-8.0) and peptide concentration (14 μ M-1.4 mM). UV studies conducted in parallel confirmed metal binding by the Fen residue ($K_D = 0.99 \pm 0.73 \mu$ M). However, examination of the Zn²⁺ induced changes in the near-UV CD spectrum suggested that the structural changes encountered by the phenanthroline chromophore as a

result of metal binding were less dramatic than those observed for peptides **ZFO1** and **ZFO2**.

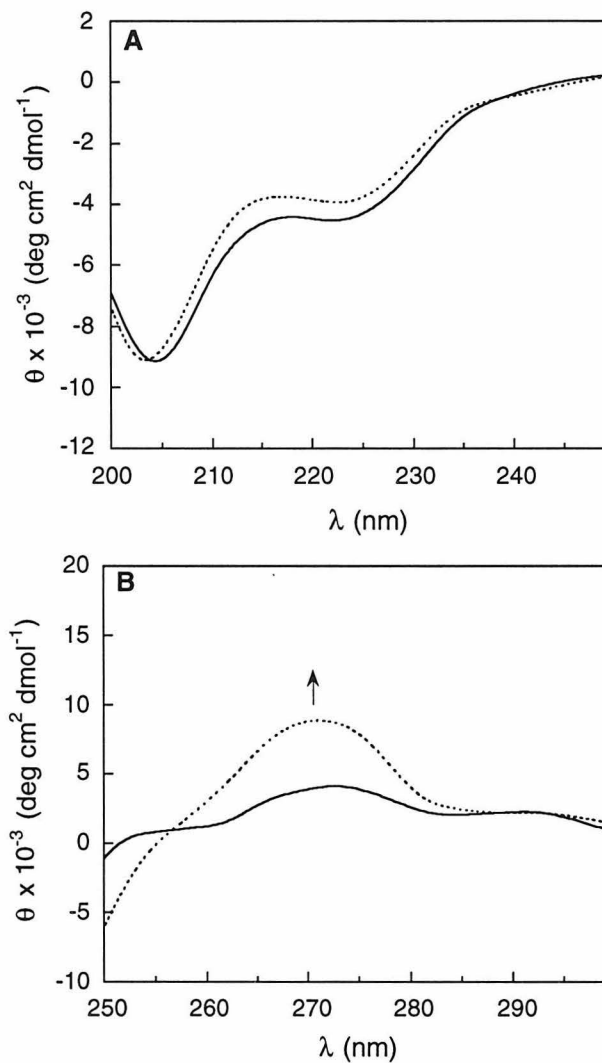


Figure 3-9 CD analysis of **ZFO3**. **A.** Far-UV CD spectra of **ZFO3**. The solid line represents the peptide in the absence of metal, the dashed line the presence of Zn²⁺. The spectra are plotted in mean residue ellipticity. **B.** Near-UV CD spectrum of **ZFO3** in the presence (solid line) and absence (dashed line) of Zn²⁺. The spectra are normalized to the concentration of Fen chromophore.

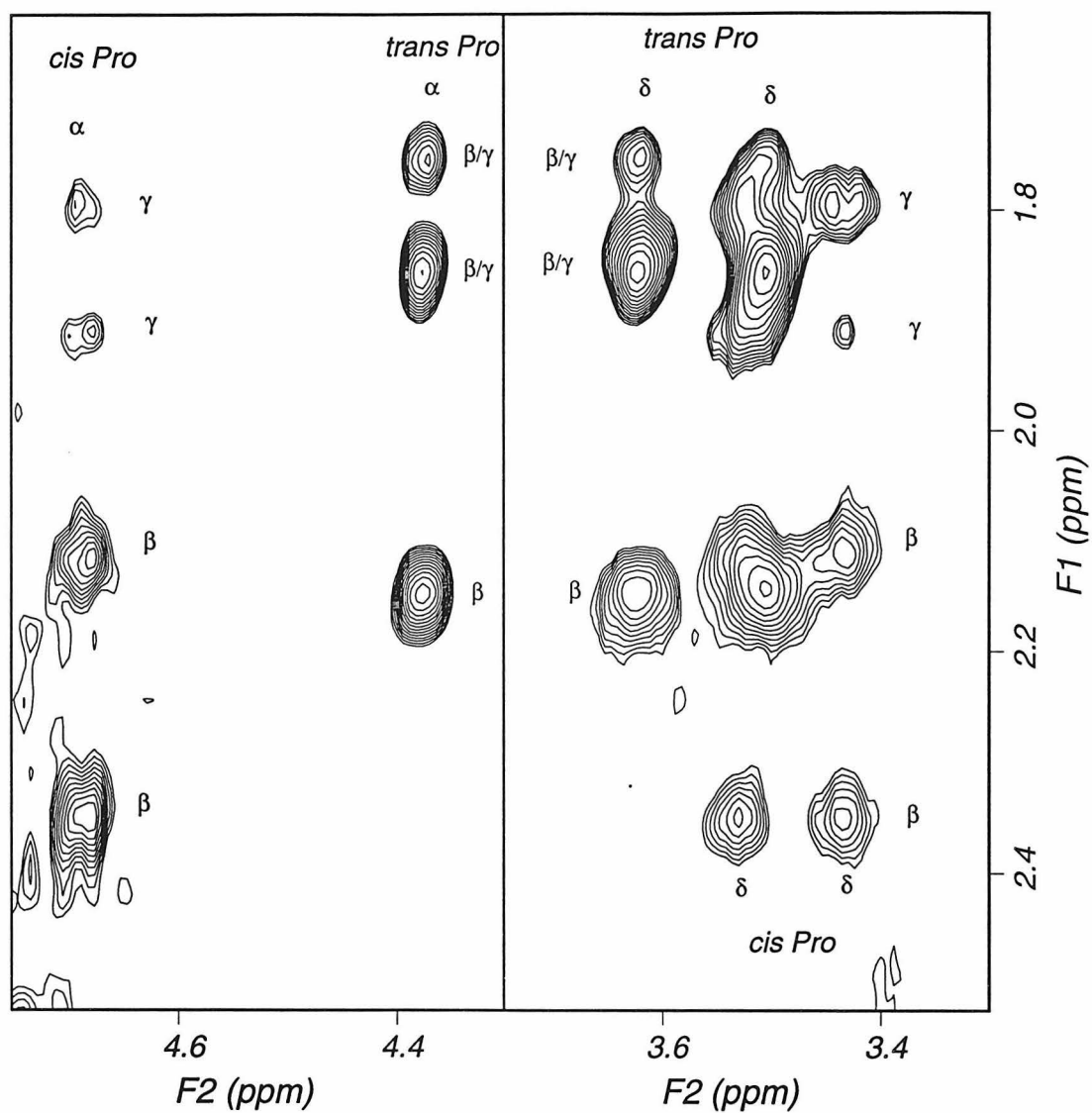


Figure 3-10 Region of the TOCSY spectrum of ZFO3, illustrating the two alternative conformations observed for this peptide for the residue DPro⁴. The assignment of *cis* and *trans* Pro⁴ amide isomers were obtained from the analysis of model peptides.

NMR Analysis

Based on the CD results, **ZFO3** was further analyzed by 2D NMR spectroscopy in the absence of metal cations. Observation of more than 30 spin systems, rather than the 23 expected for this peptide suggested the presence of slowly interconverting conformers. The purity and identity of the peptide was verified by mass spectrometry, HPLC and capillary zone electrophoresis (CZE) analysis. The presence of more than the expected number of spin systems therefore suggested that the peptide existed as an equilibrium mixture of slowly interconverting conformers. The additional spin systems greatly complicated the spectral analysis, and precluded complete sequence specific assignment and conformational analysis. Partial assignment indicated the possible identities of the conformers. The dual spin systems that exhibited vastly different chemical shifts were observed for residues in the β -hairpin region containing DPro⁴ (Figure 3-10). In contrast, residues in the helical region (Leu¹⁴-Gly²³) exhibited dual spin systems with essentially overlapping chemical shifts. These results suggested that the two conformers were the *cis-trans* isomers of the amide bond between His³ and DPro⁴. While this peptide exhibited a CD spectrum that changed very little upon metal addition, the more detailed analysis afforded by NMR spectroscopy revealed the presence of approximately a 40/60 ratio of slowly interconverting conformers.

This unexpected result highlights the importance of complementary CD and NMR studies in the analysis of designed polypeptides. CD analysis is a very sensitive and rapid technique for the investigation of the secondary structural content of peptides or proteins, but it is restricted to the observation of bulk effects and provides no detailed structural information, i.e. no data on the conformation of particular residues. Many protein design efforts to date have relied solely on circular dichroism for the assessment of structure within the

designed construct, probably due in part to limitations that the size and aggregation state of many designs place on NMR analysis. Such an approach is restrictive, in that very little specific data can be obtained to direct future design efforts. In the example presented here, analysis of solely the CD data would have resulted in a false sense of success and presented no avenues for future investigation. Based on the NMR analysis of **ZFO3**, the conformation of the β -hairpin region of the motif was investigated through the use of model peptides.

Use of Model Peptides to Examine the *cis-trans* Isomerization of DPro in Type II' Turns

In order to examine the isomerism of DPro⁴ of **ZFO3** in greater detail, the peptide Ac-Thr-His-DPro-Ser-Phe-Thr-NH₂ (**HF1**) corresponding to the turn region of **ZFO3** was synthesized. The 1D NMR spectra was considerably more complicated than expected for a simple hexapeptide, reminiscent of the larger peptide construct, **ZFO3**. This is illustrated in Figure 3-11 which depicts the downfield region of the 1D spectra. Spin system assignment by total correlation spectroscopy (TOCSY) demonstrated that **HF1** was present as an equilibrium mixture of two conformers. Analysis by rotating frame nuclear Overhauser spectroscopy (ROESY) enabled the spectroscopic assignment of the two conformers. Observation of $d_{\alpha\alpha}(\text{His}^3, \text{DPro}^4)$ and $d_{\alpha\delta}(\text{His}^3, \text{DPro}^4)$ NOE confirmed the identity of the *cis* (~40%) and *trans* (~60%) Pro⁴ amide isomers respectively (Figure 3-12).

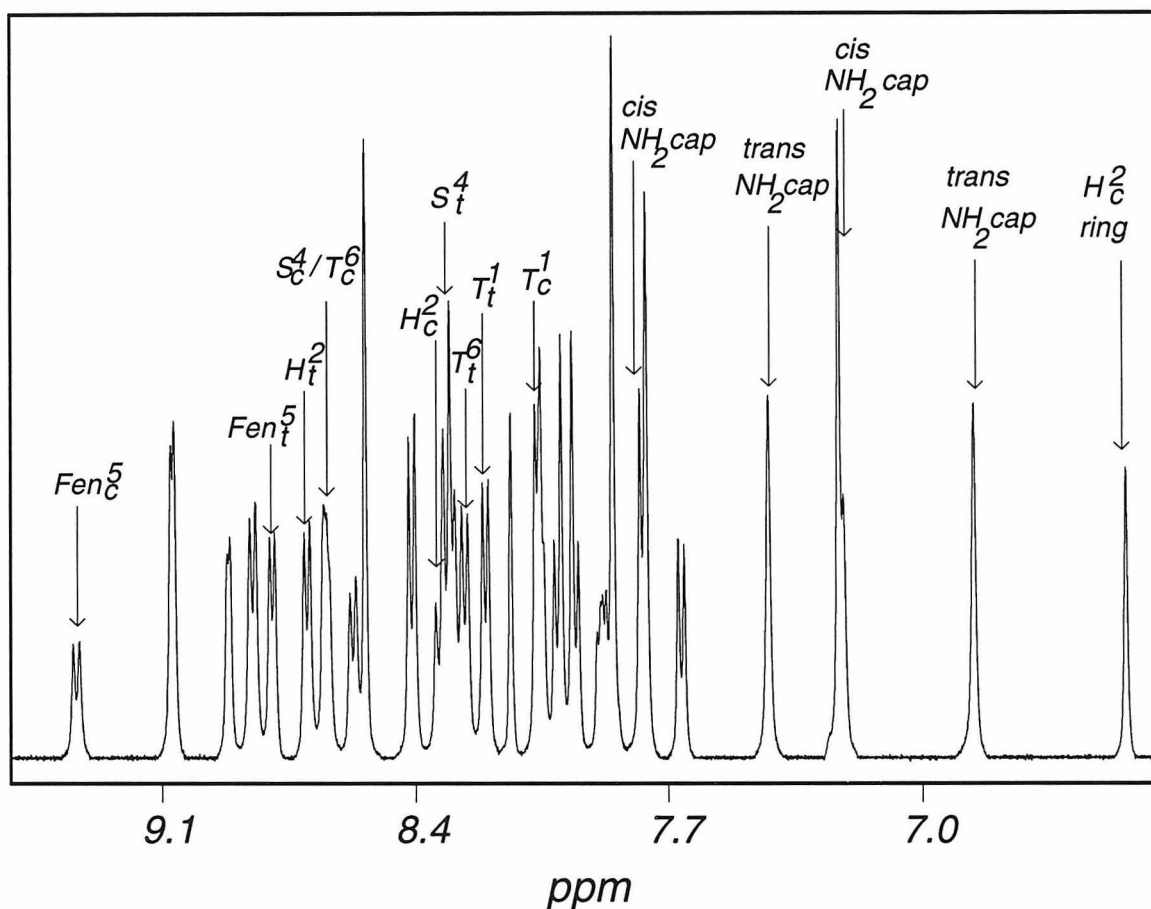


Figure 3-11 The amide region (δ 6.0-9.4 ppm) of the 1D ^1H spectra of HF1, indicating the resonances arising from the *cis* and the *trans* proline amide isomers. The amide peak of each residue is labeled by the one letter code indicating the amino acid with the residue number in the upper right and the Pro⁴ amide isomer (c=*cis*, t=*trans*) in the lower right. The peaks arising from the ring protons of the histidine or phenanthroline rings are not labeled except for the ϵ proton of His² ring of the *cis* amide isomer, due to its extreme upfield shift.

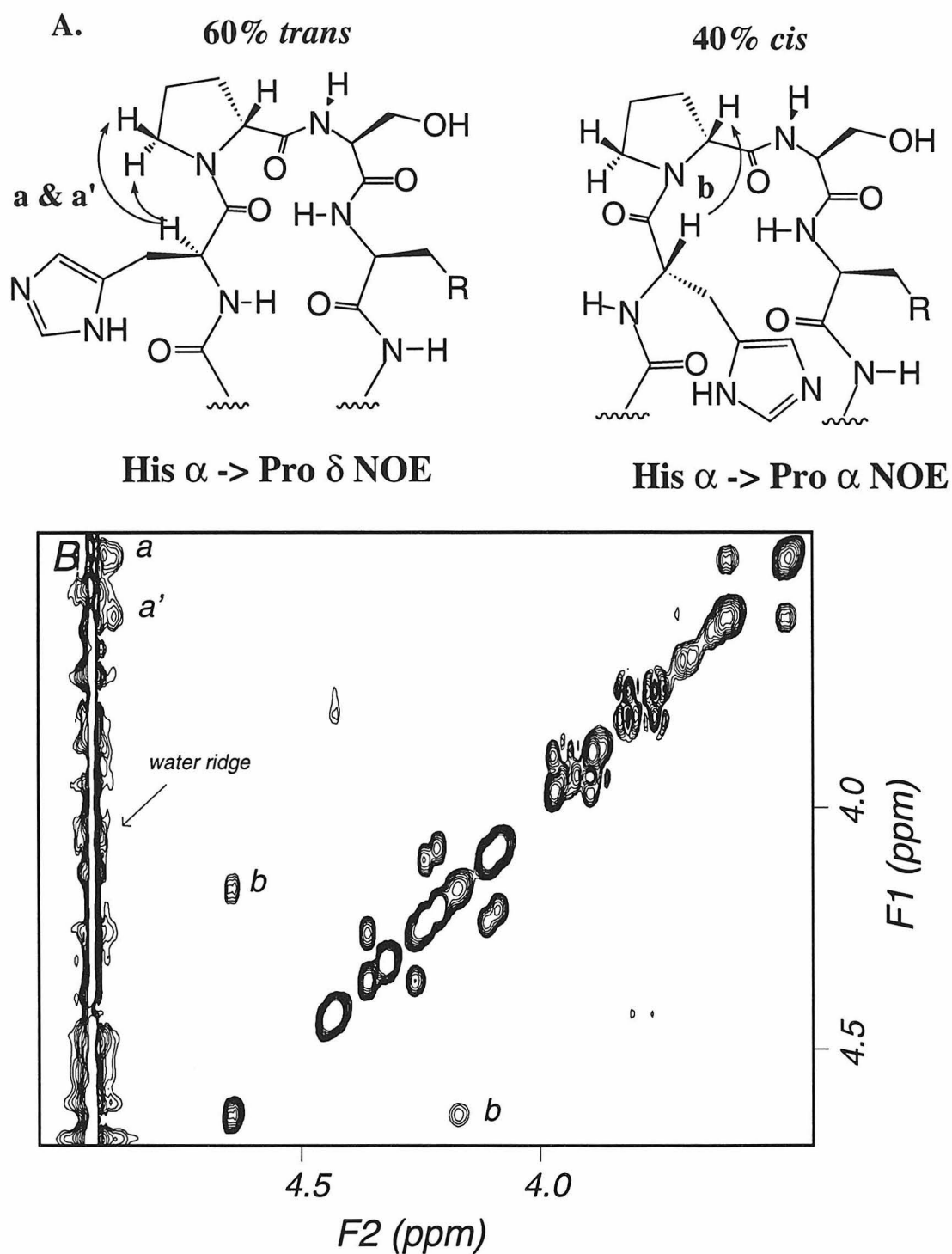


Figure 3-12 NMR identification of the *cis* and *trans* proline amide isomers of **HF1**. **A.** Schematic representation of the *cis* and *trans* proline amide isomer of **HF1**, illustrating the NOEs indicative of the proline amide geometry. **B.** A section of the ROESY spectra of **HF1**. The NOEs described in **A** are highlighted.

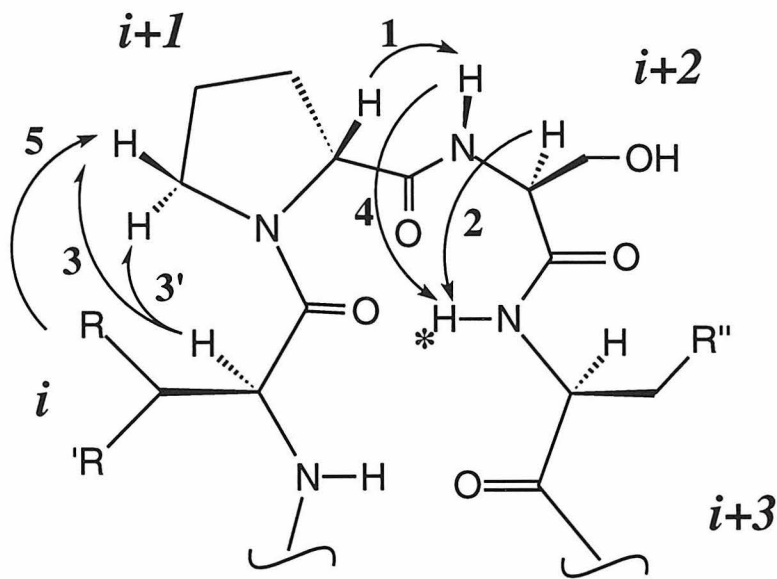


Figure 3-13 Expected NOE interaction in a heterochiral type II' reverse turn.: 1, $d_{\alpha,N}(\text{Pro, Ser})$; 2, $d_{\alpha,N}(\text{Ser, Fen})$; 3 and 3', $d_{\alpha,\delta}(\text{Xaa, Pro})$; 4, $d_{NN}(\text{Ser, Fen})$; 5, $d_{\beta\delta}(\text{Xaa, Pro})$. The $(i)-(i+1)$ amide is shown in the *trans* conformation.

The sequence of **HF1** was designed to adopt a type II' turn (in the *trans* conformation) within the $\beta\beta\alpha$ motif. The 2D NMR spectra of **HF1** were therefore examined to determine the conformation of the *trans* isomer in solution. The results of this analysis are summarized in Table 3-1. The spectroscopic signature of type II' turn formation includes the NOEs³⁵ illustrated in Figure 3-13 and a low VT coefficient ($-\Delta\delta/\Delta T < 6$ ppb/K) for the $(i+3)$ amide proton. In particular, NOE #4 [$d_{NN}(\text{Ser, Fen})$] would be expected in the desired conformation. All of the NOEs indicated in Figure 3-13 were observed for **HF1**. Additionally variable temperature analysis of the amide proton chemical shifts by 1D ^1H NMR revealed that the Fen amide was protected in the *trans* proline amide isomer. Thus, a significant population of the *trans* amide isomer adopted a type II' turn-like conformation in aqueous solution. With the modeling results presented

earlier, these NMR data indicate the potential utility of the type II' turn, to promote β -hairpin formation within the motif. In order to exploit the type II' turn in the design, the control of the isomerization of the DPro residue to favor the *trans* geometry is essential.

Peptide sequence	NOEs ^a	$-\Delta\delta/\Delta T$ (ppb/K)				
		Thr ¹	His	Ser	Fen	Thr ⁶
Ac-Thr-His-DPro-Ser-Fen-Thr-NH ₂	1, 2, 3, 4, 5	nd	6.0	nd	3.52	nd

Table 3-1 Conformational analysis of the *trans* isomer of **HF1**. ^aNOEs present in ROESY experiment for the *trans* conformation. NOE identities are as presented in Figure 3-8. The amide proton temperature coefficients for Thr¹, Thr⁶, and Ser⁴ could not be determined due to overlap in the 1D ¹H NMR spectra.

In order to manipulate the *cis-trans* proline amide isomer ratio, a series of tetrapeptides was prepared in which the residue immediately preceding DPro was varied. Since the nature of this residue has been shown to influence the proline amide geometry,^{38, 39} peptides containing the sequence DPro-LSer, positions (*i*+1) and (*i*+2) respectively, and various amino acids (Thr, Val, and Ala) at position (*i*) were synthesized and analyzed. The peptides studied are listed in Table 3-2. These sequences were chosen to probe the influence of a β -branched amino acid (**VF1** vs. **AF1**) and the affect of a more functionalized hydroxy amino acid (**VF1** vs. **TF1**) at position (*i*). Since these peptides served as prototypes for the next generation design, all the peptides included the Fen residue at position (*i*+3).

All three peptides were analyzed by 1D and 2D ^1H NMR. The *cis* amide bond between residue (*i*) and Pro (*i*+1) results in a large upfield shift in the resonances for protons associated with residue (*i*). The *cis-trans* isomer ratio could therefore be quantified by integration of the 1D NMR signals of the side chain methyl protons of residue (*i*). This is illustrated in Figure 3-14 for peptide VF1. The *trans* amide isomer dominated by greater than 90% for all sequences investigated.

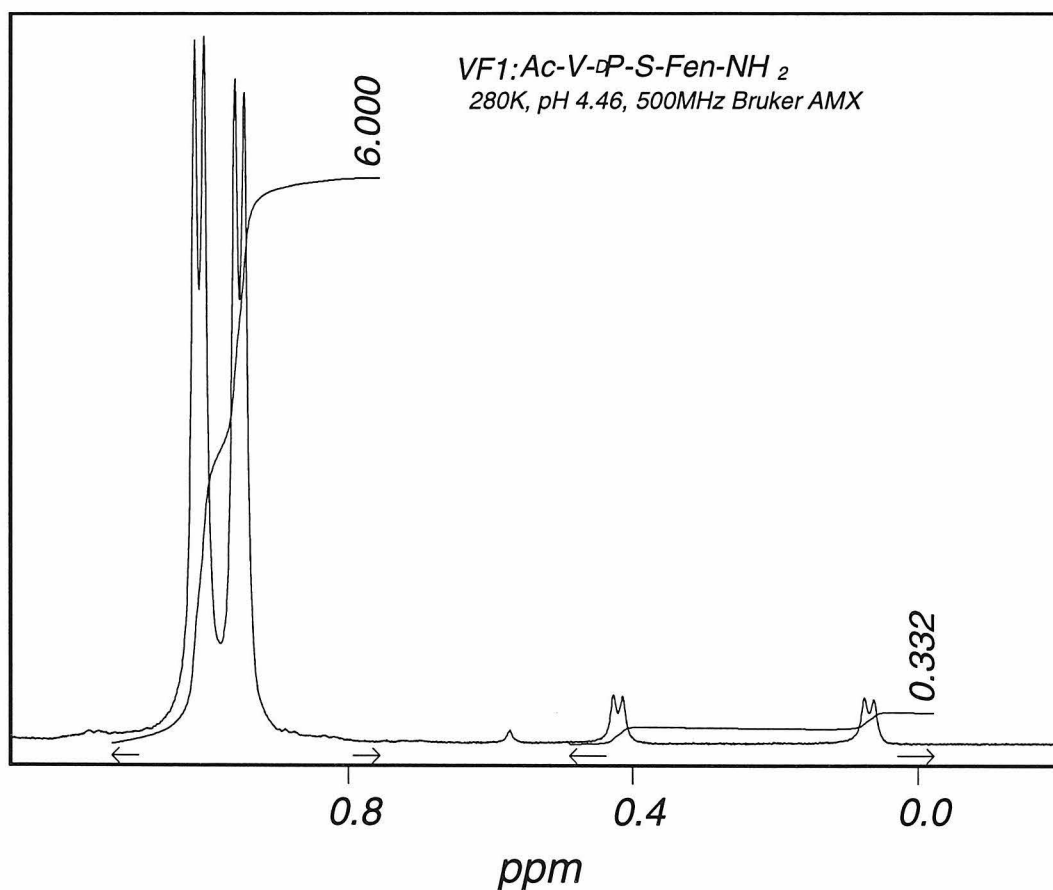


Figure 3-14 Upfield region of the 1D ^1H NMR spectrum of VF1, illustrating the resonances of the methyl proton of Val¹. Integration of these peaks for the *cis* and *trans* proline amide isomer yielded the isomer ratio (18:1 or 5.5% *cis*)

The peptides were further examined by ROESY and amide proton variable temperature (VT) NMR analyses to determine the propensity of each sequence to adopt a type II' turn conformation in solution. The results of these studies are summarized in Table 3-2. As expected for turn conformation, the amide temperature coefficient for Fen at position (*i*+3) in each peptide was very low ($-\Delta\delta/\Delta T < 5.0$ ppb/K). While NOEs #1-3 and #5 were observed for all peptides, VF1 also exhibited NOE #4 which is particularly indicative of a turn-like conformation in small peptides (Figure 3-13).^{23, 24} In the series of tetrapeptides studied, peptide VF1 exhibited the lowest Fen amide temperature coefficient, as well as the most NOEs characteristic of turn formation. Therefore, this sequence was chosen as the basis for the fourth generation design described in the next chapter.

Peptide	% <i>trans</i>	NOEs ^a	$-\Delta\delta/\Delta T$ (ppb/K)		
			Xaa	Ser	Fen
VF1: Ac-Val-DPro-Ser-Fen-NH ₂	94 %	1, 2, 3, 4, 5	9.36	8.03	3.52
TF1: Ac-Thr-DPro-Ser-Fen-NH ₂	94 %	1, 2, 3, 5	8.18	9.51	3.59
AF1: Ac-Ala-DPro-Ser-Fen-NH ₂	92 %	1, 2, 3, 5	9.10	8.83	4.53

Table 3-2 Conformational analysis of model type II' tetrapeptides. ^a NOEs present in ROESY experiment for the *trans* amide conformation of DPro². NOE identities are as presented in Figure 3-8.

Conclusions

The design efforts described in this chapter focused on the evolution of the $\beta\beta\alpha$ structure of the zinc finger domains into a metal *independent* motif. The inclusion of an unnatural metal binding amino acid, Fen,²⁹ into designed

peptides allowed for the rapid and facile determination of metal binding affinities, stoichiometries, and the effects of metal binding on structure. Type II and type II' β -turns were used as structural nucleation elements within the synthetic "zinc finger" motif. Considerable previous work in the Imperiali group established the stability of type II β -turns in small peptides,²³ therefore the first two designs, **ZFO1** and **ZFO2**, included this turn type. Additionally, simple rules for the propensity of certain amino acids for elements of secondary structure³² were successfully used to create a more preorganized $\beta\beta\alpha$ motif in the second generation design. Despite the presence of considerable preorganized secondary structure within this design, NMR and CD analysis indicated that tertiary structure of **ZFO2** was still lacking in the absence of metal.

Based on analysis of β -turns the protein database,³⁷ modeling studies were performed to compare the effects of type II and type II' β -turns on sheet geometry. These studies suggested the potential utility of the type II' turn for promoting β -hairpin formation. The third generation design, **ZFO3**, therefore included a type II' turn. While CD analysis indicated the presence of secondary structure that was metal independent, NMR analysis revealed that this peptide existed as a mixture of slowly interconverting conformers. The identity of the two conformers was confirmed by the NMR analysis of a short hexapeptide encompassing the turn region of the sequence to be the *cis* and *trans* isomers of proline. A series of tetrapeptides was synthesized to probe the effect of the residue preceding proline on the isomer ratio. The specific sequences changes required to both increase the *trans* proline amide isomer ratio and promote the adoption of a turn like structure in solution were determined from the NMR analyses of these peptides. The optimal sequence determined by these efforts, -Val-DPro-Ser-Fen-, will therefore form the basis of future design efforts described in the following chapter.

The results of the analysis of **ZFO3** demonstrates the importance of using both NMR and CD techniques in the analysis of protein structure, especially for designed constructs. Although CD is a rapid sensitive technique for the estimation of the macroscopic structural content of peptides and proteins, requiring very little material and permitting structural investigations under a variety of conditions, it yields no detailed information and conclusions based solely on CD data can be misleading. On the other hand, NMR requires considerably more material and effort, but can yield the detailed information needed to guide further design efforts. Both types of experiments were used successfully in the work presented here in an iterative protein design process. The continuation of this effort is described in the next chapter.

Acknowledgments

This work was performed in collaboration with Richard Cheng. Richard developed the coupling methods used for the Fen amino acid, synthesized peptides **ZFO1**, **ZFO2**, **ZFO3**, and **HF1**, performed CD and UV studies on these peptides, and assisted in the synthesis of **VF1**, **AF1**, and **TF1**. Ranabir Sinha Roy assisted in the comparative modeling of the type II and type II' turns. The Dorthy and Camilla Chandler - Frost Laboratory in Chemical Biology is gratefully acknowledged for the use of the 600 MHz spectrometer. Part of the work presented in this chapter appeared in a full paper published in the *Journal of the American Chemical Society*:

Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Economy in Protein Design: Evolution of a Metal-Independent $\beta\beta\alpha$ Motif Based on the Zinc Finger Domains," *J. Am. Chem. Soc.* **1996**, *118*(13), 3073-3081.

Experimental

Peptide Synthesis

The peptides were synthesized by solid phase methods using N^{α} -9-fluorenylmethyloxycarbonyl (Fmoc) amino acid pentafluorophenyl esters on a Milligen 9050 automated peptide synthesizer. Commercially available starting materials and reagents were purchased from Milligen Biosearch, EM Science, and Aldrich Chemical Company. Fmoc-PAL-PEG-PS resin (substitution 0.17 - 0.21 mmol/g) was used to afford C-terminal primary amides. The Fen residue was coupled off-line with three equivalents of Fmoc-Fen-OH (N^{α} -fluorenylcarboxymethyl-3-(1,10-phenanthrol-2-yl)-L-alanine) amino acid²⁸ and four equivalents of 1-hydroxy-7-azabenzotriazole (HOAT) and *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU)⁴⁰ and 8 equiv of diisopropylethylamine in *N*-methylpyrrolidinone for 2.5 hours. Peptides were *N*-acylated on the resin using 7 equivalents of acetic anhydride and 7 equivalents of triethylamine in 3 mL of DMF for 2 hours.

The peptides were deprotected and cleaved from the resin by treatment with Reagent K (trifluoroacetic acid/phenol/H₂O/thioanisole/ethanedithiol, 82.5:5:5:5:2.5)⁴¹ for 2 hours. Trituration with 1:1 ether/hexane afforded the crude peptide. The peptides were purified by reverse-phase HPLC on a C₁₈ semipreparative column. Peptide purity confirmed by analytical HPLC and/or capillary zone electrophoresis (CZE). Identity of all peptides was confirmed by MALDI or plasma desorption mass spectroscopy (PDMS). Peptide **ZFO1** MH⁺ calcd 2635.9, obsd 2635.9 (PDMS). Peptide **ZFO2** MH⁺ calcd 2775.2, obsd 2775.7 (PDMS). Peptide **ZFO3** MH⁺ calcd 2775.2, obsd 2775.7 (MALDI). Peptide **VF1** MH⁺ calcd 592.67 obsd 592.29 (FAB). Peptide **TF1** cacd 594.64 obsd 594.27 (FAB). Peptide **AF1** calcd 564.62, obsd 564.26 (FAB).

Absorption Spectroscopy (UV-Vis) Studies

Absorption spectra were obtained on a Shimadzu UV-110 spectrophotometer using 1.0 cm path length quartz cuvettes. The concentrations of the peptide solutions were determined spectrophotometrically using the characteristic absorption band of the phenanthrolyl moiety ($\lambda_{\text{max}} = 268 \text{ nm}$, $\epsilon_{\text{max}} = 20,890 \text{ M}^{-1}\text{cm}^{-1}$). The extinction coefficient of the Fen amino acid at 268 nm was calculated from UV spectra of a Fen containing peptide. The concentration of which was determined by quantitative amino acid analysis. Aqueous solutions of known peptide concentrations (1-20 μM) were titrated with standardized metal cation solutions at room temperature in 0.05 M HEPES buffer at pH 8.25. The distinct red-shift of the phenanthrolyl π - π^* transition upon metal chelation was monitored. The dissociation constants were determined the calculations outlined below. Data analysis was performed using KaleidaGraph version 3.0 (Abelbeck Software). A minimum of three titrations were performed for each peptide-metal cation complex. Binding isotherms were assumed to follow two-state equilibria; well defined isosbestic points were observed in all cases. The dissociation constants (K_D) were obtained using an iterative process based on the Scott equation⁴²:

$$\frac{b \cdot [M]}{\Delta A} = \frac{[M]}{[\text{Pep}]_t \cdot \Delta \epsilon} + \frac{K_D}{[\text{Pep}]_t \cdot \Delta \epsilon} \quad (1)$$

Where :	b	=	cell path length
	$[M]$	=	free metal cation concentration
	ΔA	=	change in absorbance due to complex formation
	$[\text{Pep}]_t$	=	total peptide in solution

$\Delta\epsilon$ = change in molar absorptivity due to complex formation

The dissociation constants (K_D) were calculated from plots of $[M]$ vs $[M]/\Delta A$ by the following equation:

$$K_D = \frac{\text{y - intercept}}{\text{slope}} \quad (2)$$

Initially, data was considered only at high $[M]_{\text{total}}$ so that the approximation $[M]_{\text{total}} \approx [M]$ could be used. The resultant values of K_D and $\Delta\epsilon$ obtained by least squared analysis were fitted to absorbances using equation 1 to obtain better estimates of $[M]$. This was repeated in an iterative manner until the values of K_D and $\Delta\epsilon$ converged.

Circular Dichroism Spectroscopy Studies

Circular dichroism spectra were recorded on a Jasco J600 Circular Dichroism Spectrometer. The concentration of the peptide solutions was determined spectrophotometrically using the characteristic absorption band of the phenanthrolyl moiety. The spectra were processed using the Jasco J600 System Software, and subsequently analyzed using KaleidaGraph (Abelback Software). Metal cation titrations were performed with peptide concentrations ranging from 13 μM to 170 μM in 0.01 M TRIS buffer (pH 8.25) in either a 1.0 cm or a 0.1 cm path-length quartz cell. Aggregation properties were investigated with peptide concentrations ranging from 10 μM to 2.6 mM at pH 4.5 in cuvettes of appropriate path-length.

Modeling Studies

Modeling studies were performed using the Biopolymer module of Insight II (Biosym Technologies, CA). Anti-parallel β -strands were generated by using the dihedral angles $\phi = -139^\circ$, $\psi = 135^\circ$. A β -hairpin connecting these strands was created by adjusting the dihedrals of the central two turn residues: type II ($\phi_1 = -60^\circ$, $\psi_1 = 120^\circ$, $\phi_2 = 80^\circ$, $\psi_2 = 0^\circ$), type II' ($\phi_1 = 60^\circ$, $\psi_1 = -120^\circ$, $\phi_2 = -80^\circ$, $\psi_2 = 0^\circ$).

NMR Studies

ZFO2

Two-dimensional NMR experiments (TOCSY,^{43, 44} DQFCOSY,⁴⁵ ROESY,⁴⁶ and NOESY⁴⁷) for the peptide in the absence of metal were performed in 90:10 H₂O:D₂O, pH 4.5 on a 500 MHz Bruker AMX spectrometer. Zn²⁺ (1.1 equivalents of ZnCl₂) was added to the same NMR sample. 2D NMR experiments in the presence of zinc were performed on both 500 MHz Bruker AMX and 600 MHz Varian Unity Plus spectrometers. Spin systems for both systems (with and without Zn²⁺) were assigned where possible using total correlation spectroscopy (TOCSY),^{35, 43, 44} magnitude mode correlation spectroscopy (mmCOSY), and nuclear Overhauser spectroscopy⁴⁷ (NOESY, 200ms mixing time). Suppression of the water signal was accomplished by presaturation during the relaxation delay (1.3 s, Bruker; 1.5 s, Varian). Processing and analysis of the spectra were done using Felix 2.3 software (Biosym Technologies, CA).

Assignment of the NMR of the **ZFO2-Zn²⁺** was complicated by the presence of zinc. While many of the spin systems for residues in the helix (alanine, lysine, and leucine) were assignable, spin systems in the turn region (Pro⁴, D-Ser,⁵ Fen⁶) were not observed. Therefore the peptide could not be completely assigned. The absence of resonances arising from these residues is

presumably due to the broadening of the signals that would result from the rapid exchange of the metal into the metal site. In contrast, these spin systems (Pro⁴, DSer⁵, Fen⁶) were clearly observable in the spectra taken before metal addition. Unfortunately, complete sequence specific assignment of the metal free peptide was not possible due both to the degeneracy in the sequence, and the lack of sufficient spectral dispersity in the amide and alpha regions of the 500MHz spectra. The partial assignment is given in Table 3-3. The spin systems resulting from the aromatic (Phe¹, His³, Phe⁸, and His²¹) residues and Asp¹² could not be definitively assigned to a particular residue type or sequence number, and are therefore listed under the residue heading of Xaa. The resonance of the ring protons of the Phe residues could not be assigned to a particular spin system and are listed at the end of the table.

ZFO3

Two-dimensional NMR experiments for **ZFO3** were performed in the absence of metal on either a 600 MHz Varian Unity Plus (NOESY)⁴⁵ or a 500 MHz Bruker AMX (TOCSY⁴³, mmCOSY, and ROESY⁴⁶) spectrometer, in 90:10 H₂O:D₂O, pH 4.5. Spin systems assignments were accomplished using TOCSY, mmCOSY, and NOESY (200 ms mixing time)³⁵ experiments. Suppression of the water signal was accomplished by presaturation during the relaxation delay (1.3 s, Bruker; 1.5 s, Varian). Processing and analysis of the spectra were done using Felix 2.3 software (Biosym Technologies, CA).

Residue	alpha	amide	side chain
Glu ¹³	4.077	8.164	2.086, 2.346
Fen ⁶	5.266	8.665	3.605
Thr	4.229	8.146	4.088, 1.063
Thr	4.241	8.366	4.033, 1.062
Pro ⁴	4.454		1.919, 2.277, 2.00, 3.750, 3.511
DSer ⁵	4.466	8.937	3.884, 3.816
Ser	4.363	8.375	3.763, 3.825
Ser	4.374	8.566	3.935, 3.877
Arg	4.199	7.848	1.627, 1.87, 1.726, 3.209, 7.300
Arg	4.178	8.032	1.763, 1.597, 3.126, 7.338
Lys ¹⁶	4.013	7.855	1.893, 1.555, 1.665, 2.925, 7.65
Leu	4.132	7.87	1.822, 1.711, 1.613, 0.863
Leu	4.128	8.159	1.688, 1.568
Leu	4.212	7.954	1.689, 1.605, 1.502, 0.902, 0.839
Leu	4.171	8.077	1.764, 1.575, 1.669,
Ala	4.106	8.018	1.470
Ala	4.328	8.287	1.429
Gly ²³	3.92	8.469	
Phe	4.314	8.214	2.764, 2.670
Xaa	nd ^a	8.479	2.899, 2.938
Xaa	4.504	8.278	2.96, 2.81
Xaa	4.677	8.275	3.20, 3.345
Xaa	4.49	8.333	2.65
Phe rings			7.19, 7.0277, 7.025
Phe rings			7.09, 7.19

Table 3-3 Partial assignment of ZFO2. Most spin systems could not be assigned sequence numbers. ^a The alpha proton of this spin system was obscured by the water resonance.

Residue	alpha	amide	side chain
Asp	4.49	8.31	2.757
Asp	4.51	8.335	2.667
Fen	5.29	8.750	3.632, 3.533, 7.720
Fen	5.171	8.600	3.51, 3.38, 7.639
Thr	4.122	8.101	3.944, 1.015
Thr	4.34	8.418	4.032, 1.06
Thr	4.294	8.607	4.087, 1.122
Thr	4.35	8.223	4.088, 1.108
Pro	4.377		1.753, 1.853, 3.624, 3.503
Pro	4.688		1.810, 1.905, 2.103, 1.786, 3.552, 3.436
Ser	4.388	8.456	3.86, 3.78
Ser	4.410	8.393	3.827, 3.767
Ser	4.237	8.456	3.845, 3.90
Ser	4.259	8.498	3.858, 3.91
Ser	4.483	8.598	3.955, 3.853
Ser	4.499	8.807	3.963, 3.896
Arg	4.203	7.843	1.625, 1.883, 1.723, 3.209, 7.305
Arg	4.249	nd	1.701, 1.845, 1.937, 3.266, 7.414
Arg	4.212	nd	1.617, 1.759, 1.884, 3.125, 7.374
Lys ¹⁶	4.011	7.854	1.892, 1.552, 1.659, 1.364, 2.916
Leu	4.129	7.877	1.822, 1.716, 1.613, 0.86
Leu	4.144	8.160	1.688, 1.716, 1.568,
Leu	4.171	8.044	1.764, 1.591, 0.858, 0.896
Leu	4.213	7.97	1.689, 1.502, 1.605, 0.902, 0.842
Ala	4.100	8.015	1.48
Ala	4.322	nd	1.34
Gly	3.92	8.47	
Glu	4.075	8.179	2.341, 2.095
Xaa	4.23	8.50	2.066, 2.456
Xaa	4.884	8.583	3.084, 3.218, 7.077
Xaa	4.63	8.396	3.008, 3.129
Xaa	4.670	8.280	3.343, 3.202, 7.303
Xaa	4.337	8.287	2.599, 2.674
Xaa	4.474	8.173	2.693, 2.800, 6.847
Xaa	4.462	nd	2.886, 2.808, 7.118

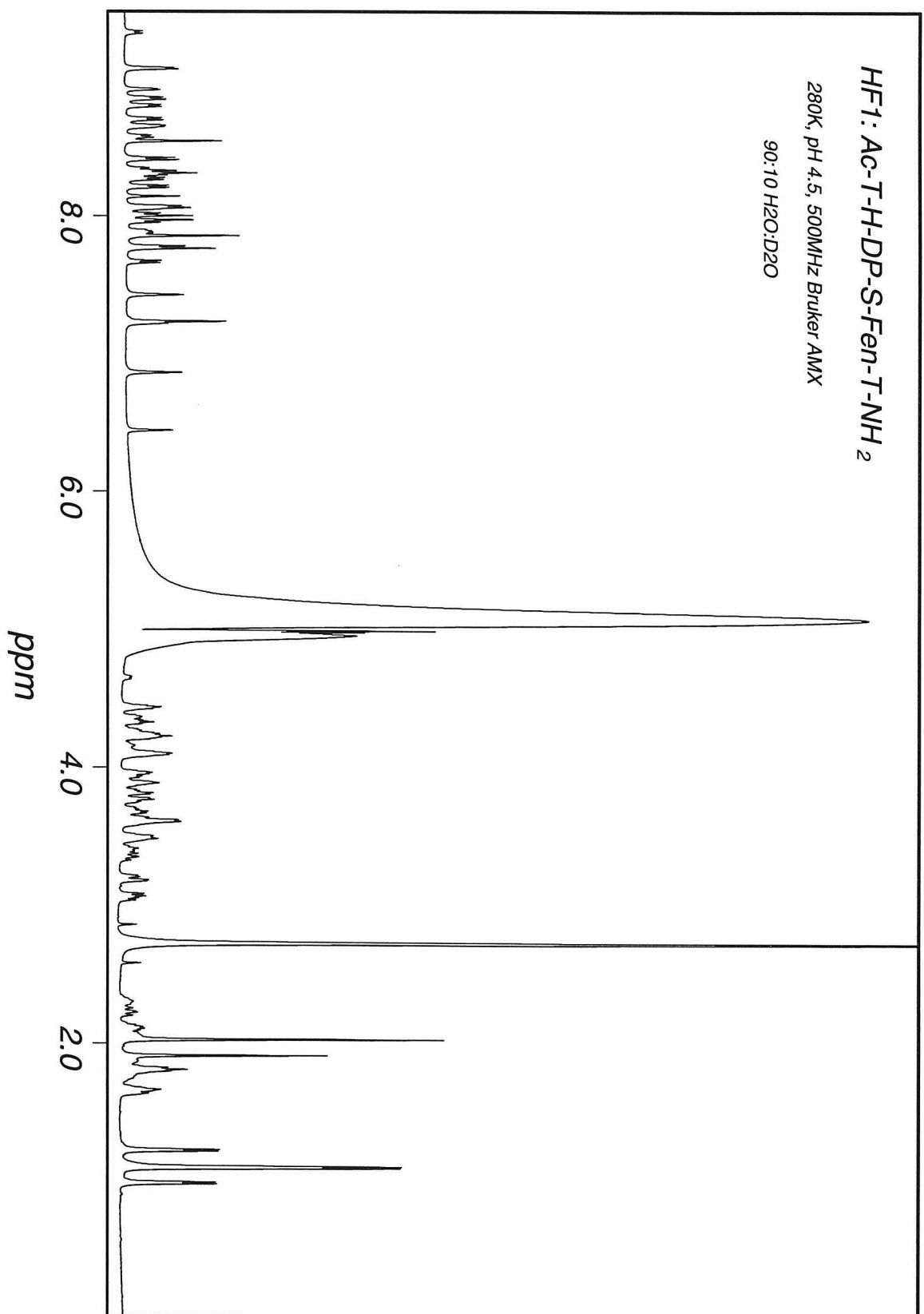
Table 3-4 Partial assignment of the spin systems **ZFO3** (both conformations). Xaa indicates an aromatic amino acid, glutamic acid or aspartic acid. Most spin systems could not be assigned sequence numbers, and assignments are not complete. nd = Not determinable due to overlap.

Two-dimensional NMR spectroscopy of **HF1**, **VF1**, **AF1**, and **TF1** were performed on a Bruker Instruments AMX 500MHz spectrometer (**VF1**, **AF1**, or **TF1**) or a Varian Unity Plus 600 MHz spectrometer (**HF1**). Complete sequence specific assignments were accomplished using total correlation spectroscopy (TOCSY)⁴³ and rotating-frame nuclear Overhauser effect spectroscopy (ROESY).⁴⁶ The temperature dependence of amide proton shifts were determined between 7 °C and 37 °C, in 5 °C increments. In all cases the chemical shifts varied linearly with temperature.

HF1

Complete sequence specific assignments were possible for both the *cis* and *trans* Pro³ amide isomers of **HF1** and are listed in Table 3-5. The NMR spectra were further analyzed to determine the conformation of the *trans* amide isomer, and a list of NOEs observed in the ROESY spectrum (Varian 600 MHz Unity Plus) relevant to the assessment of conformation are listed in Table 3-6. The entire ¹H 1D spectrum for **HF1** is shown in Figure 3-15.

Figure 3-15 (following page) 1D ¹H NMR spectrum of **HF1**.



Residue	α proton	amide proton	others
trans isomer:			
Ac	2.02		
Thr ¹	4.21	8.18	β 4.08; γ 1.09
His ²	4.89	8.67	β 3.18, 3.06; ring 8.53, 7.22
Pro ³	4.33		β 2.10; γ 1.80, 1.65; δ 3.60, 3.49
Ser ⁴	4.43	8.26	β 3.81; 3.76
Fen ⁵	5.22	8.78	β 3.67; 3.59; ring 7.78, 8.02, 8.86, 8.42, 8.06, 8.86
Thr ⁶	4.24	8.25	β 4.11; γ 1.09
NH ₂		6.85, 7.40	
cis isomer:			
Ac	2.05		
Thr ¹	3.97	8.034	β 3.88; 0.98
His ²	4.17	8.32	β 2.24; ring 6.46, 8.32
Pro ³	4.64		β/γ 2.29, 2.04, 1.86, 1.75; δ 3.47, 3.40
Ser ⁴	4.45	8.62	β 3.88, 3.93
Fen ⁵	nd	9.33	β 3.63, 3.35; 8.33, 8.59, 8.93, 7.89, 7.87, 7.68
Thr ⁶	4.36	8.60	β 4.26, γ 1.21
NH ₂		7.18, 7.73	

Table 3-5 NMR assignment of the *cis* and *trans* Pro³ amide isomers of **HF1**. Spectra acquired on 600MHz Varian Unity Plus (13.8 mM in 700 μ l 90:10 H₂O:D₂O, pH 4.5, 280K) spectrometer and referenced to DMSO in water; 2.72 ppm.

Residue	Residue	relative intensity
Ser ⁴ α	Fen ⁵ NH	strong
Thr ¹ α	His ³ NH	strong
Thr ¹ β	His ³ NH	medium
Fen ⁵ α	Thr ⁶ NH	strong
Pro ³ α	Ser ⁴ NH	strong
Ac cap CH ₃	Thr ¹ NH	medium
His ² NH	Thr ¹ NH	weak
Ser ⁴ NH	Fen ⁵ NH	weak
His ² α	Pro ³ δ	medium ^a
His ² α	Pro ³ δ'	medium ^a

Table 3-6 Significant NOEs observed in the ROESY spectrum for the *trans* Pro³ amide isomer of peptide **HF1** (does not include intra-residue NOEs). ^a Assignment of relative intensity of this peak was complicated by proximity of the water signal.

VF1, AF1, TF1

Complete sequence specific assignments were possible for the *trans* Pro² amide isomers and partial assignment for the *cis* isomers and are listed in Tables 3-7, 3-9, and 3-11. The NMR spectra were further analyzed to determine the conformations of the *trans* amide isomers, and lists of NOEs observed in the ROESY spectra (Bruker 500 MHz) relevant to the assessment of the conformations of these peptides are listed in Tables 3-8, 3-10, and 3-12. The one-dimensional spectra of the peptides are shown in Figures 3-16, 3-17, 3-18.

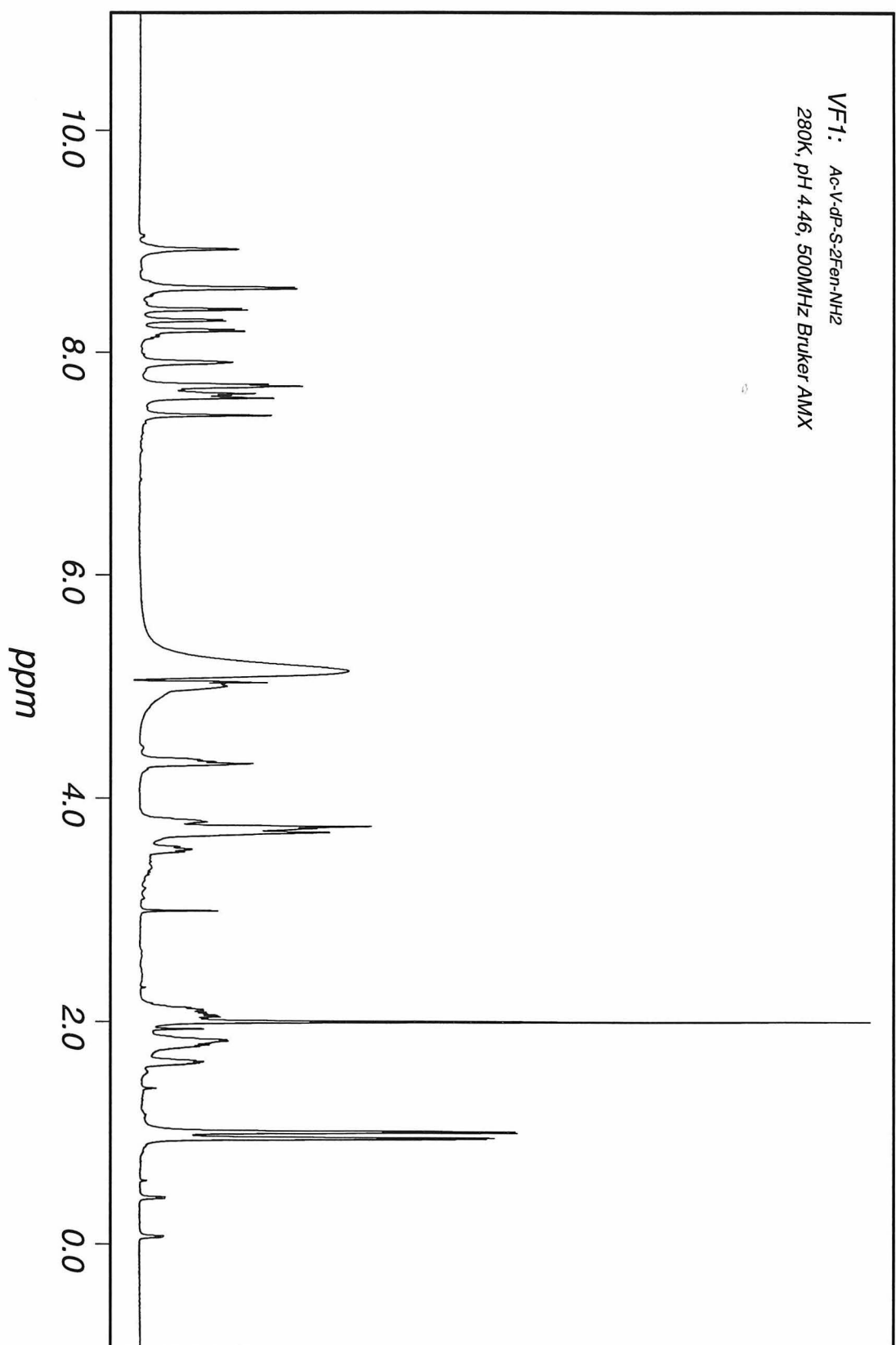
Residue	α proton	amide proton	side chain
<i>trans</i> isomer:			
Ac	2.00 ^a		
Val ¹	4.31	8.37	β 2.05; γ 1.00
Pro ²	4.34		β/γ 2.105 1.815, 1.636; δ 3.685, 3.794,
Ser ³	4.30	8.28	β 3.68; 3.74
Fen ⁴	5.13	8.57	β 3.71, 3.54
NH ₂		7.416, 6.859	
<i>cis</i> isomer:			
Ac	2.00 ^a		
Val ¹	3.72	8.127	β 1.559; γ 0.423, 0.069
Pro ²	4.81		β/γ 2.147, 1.66, 1.89; δ 3.32
Ser ³	4.45	8.50	β 3.380
Fen ⁴	4.97	9.045	β 3.39, 3.64
NH ₂			

Table 3-7 NMR assignment of the *cis* and *trans* Pro² amide isomers of **VF1**. Spectra acquired on 500 MHz Bruker AMX spectrometer (507 μ l 90:10 H₂O:D₂O, pH 4.46, 280K) and ^areferenced to the acetyl cap at 2.00 ppm.

Residue	Residue	relative intensity
Ser ³ α	Fen ⁴ NH	strong
Pro ² α	Ser ³ NH	strong
Ac cap CH ₃	Val ¹ NH	strong
Ser ³ NH	Fen ⁴ NH	weak
Val ¹ α	Pro ² δ	strong
Val ¹ α	Pro ² δ'	medium
Val ¹ γ	Pro ² δ	weak

Table 3-8 Significant NOEs observed in the ROESY spectrum for the *trans* Pro² amide isomer of peptide **VF1** (does not include intra-residue NOEs).

Figure 3-16 (following page) 1D ¹H NMR spectrum of **VF1**.



Residue	α proton	amide proton	side chain
<i>trans</i> isomer:			
Ac	1.98		
Thr ¹	4.49	8.31	β 4.08; γ 1.179
Pro ²	4.27		β/γ 2.091, 1.80, 1.63; δ 3.627, 3.713
Ser ³	4.27	8.39	β 3.62; 3.67
Fen ⁴	5.09	8.52	β 3.54, 3.54
NH ₂			
<i>cis</i> isomer:			
Thr ¹	3.907		β 3.42, γ 0.566
Pro ²	4.67		β/γ 1.891, 1.504, 1.74
Ser ³	4.35	8.46	β 3.79, 3.64
Fen ⁴	nd ^a	8.95	β 3.64, 3.62

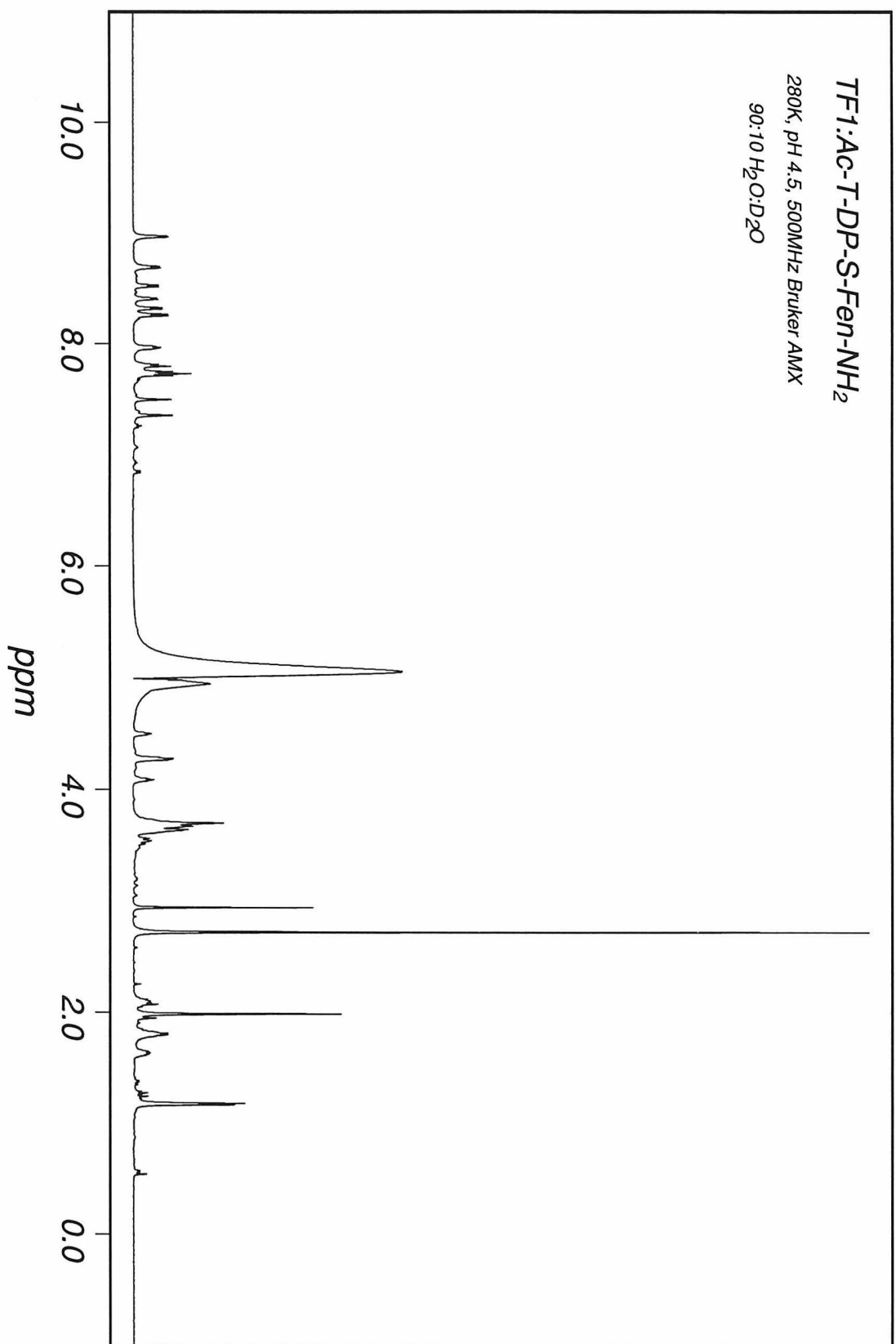
Table 3-9 NMR assignment of the *cis* and *trans* Pro² amide isomers of TF1.

Spectra acquired on 500 MHz Bruker AMX spectrometer (484 μ l 90:10 H₂O:D₂O, pH 4.50, 280K) and referenced to DMSO at 2.72 ppm. ^a The alpha resonance for this residue is obscured by the water signal.

Residue	Residue	relative intensity
Ser ³ α	Fen ⁴ NH	strong
Pro ² α	Ser ³ NH	strong
Ac cap CH ₃	Thr ¹ NH	weak
Thr ¹ α	Pro δ	strong
Thr ¹ β	Pro δ'	strong
Thr ¹ γ	Pro δ'	weak

Table 3-10 Significant NOEs observed in the ROESY spectrum for the *trans* Pro² amide isomer of peptide TF1 (does not include intra-residue NOEs).

Figure 3-17 (following page) 1D ¹H NMR spectrum of TF1.



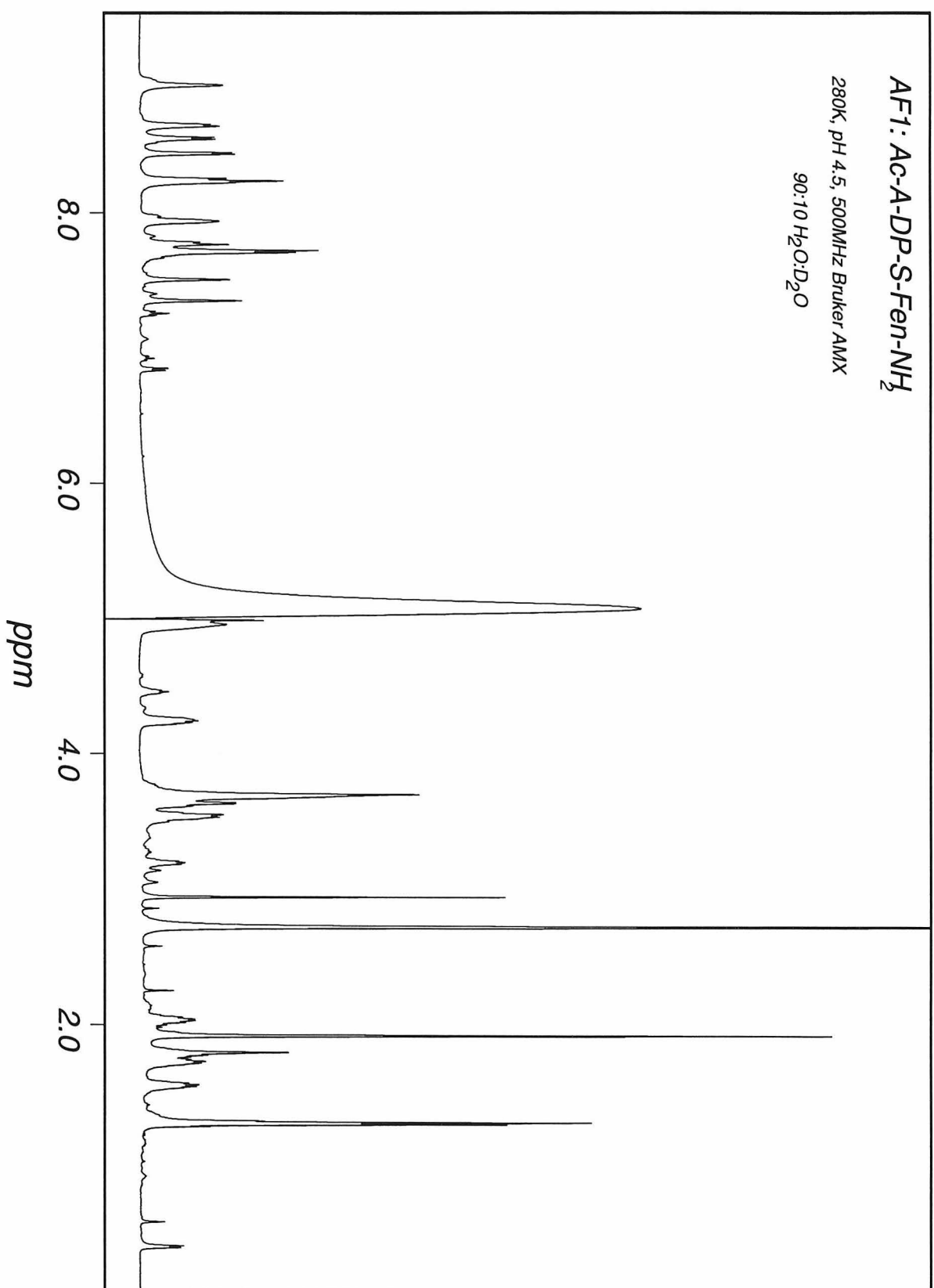
Residue	α proton	amide proton	side chain
<i>trans</i> isomer:			
Ac	1.914		
Ala ¹	4.45	8.44	β 1.26
Pro ²	4.23		β, γ 2.035, 1.55, 1.77, 1.74; δ 3.39, 3.54,
Ser ³	4.25	8.22	β 3.68; 3.62
Fen ⁴	5.10	8.55	β 3.70, 3.52
NH ₂		7.25, 6.84	
<i>cis</i> isomer:			
Ala ¹	3.64	7.97	β 0.352
Pro ²	4.57		β, γ 2.137, 1.50, 1.71, 1.70; δ 3.18, 3.29
Ser ³	4.33	8.52	β 3.75
Fen ⁴		8.99	β 3.382, 3.642

Table 3-11 NMR assignment of the *cis* and *trans* Pro² amide isomers of **AF1**. Spectra acquired on 500 MHz Bruker AMX spectrometer (478 μ l 90:10 H₂O:D₂O, pH 4.46, 280K) and referenced to the DMSO at 2.72 ppm.

Residue	Residue	relative intensity
Ser ³ α	Fen ⁴ NH	medium
Pro ² α	Ser ³ NH	strong
Ac cap CH ₃	Thr ¹ NH	medium
Ala ¹ α	Pro δ	strong
Ala ¹ β	Pro δ'	strong

Table 3-12 Significant NOEs observed in the ROESY spectrum for the *trans* Pro² amide isomer of peptide **AF1** (does not include intra-residue NOEs).

Figure 3-18 (following page) 1D ¹H NMR spectrum of **AF1**.



References

- (1) Berg, J.M. "Zinc Finger Domains: Hypotheses and Current Knowledge," *Annu. Rev. Biophys. Biophys. Chem.* **1990**, 19, 405-421.
- (2) Frankel, A.D.; Berg, J.M.; Pabo, C.O. "Metal-dependent Folding of a Single Zinc Finger From Transcription Factor IIIA," *Proceedings of the National Academy of Sciences* **1987**, 84, 4841-4845.
- (3) Eis, P.S.; Lakowicz, J.R. "Time-Resolved Energy Transfer Measurements of donor-Acceptor Distance Distributions and Intramolecular Flexibility of a CCHH Zinc Finger Peptide," *Biochemistry* **1993**, 32, 7981-7993.
- (4) Berg, J. M. "Zinc Finger Domains: From Predictions to Design," *Acc. Chem. Res.* **1995**, 28, 14-19.
- (5) Pavletich, N. P.; Pabo, C. O. "Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 Å," *Science* **1991**, 252, 809-817.
- (6) Omichinski, J. G.; Clore, G. M.; Appella, E.; Sakaguchi, K.; Gronenborn, A. M. "High-Resolution Three-Dimensional Structure of a Single Zinc Finger from a Human Enhancer Binding Protein in Solution," *Biochemistry* **1990**, 29, 9324-9334.
- (7) Lee, M. S. ; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. "Three-Dimensional Solution Structure of a Single Zinc Finger DNA-Binding Domain," *Science* **1989**, 245, 635-637.
- (8) Pavletich, N. P.; Pabo, C. O. "Crystal Structure of a Five-Finger GLI-DNA Complex: New Perspectives on Zinc Fingers," *Science* **1993**, 261, 1701-1707.
- (9) Fairall, L.; Schwabe, J. W. R.; Chapman, L.; Finch, J. T.; Rhodes, D. "The Crystal Structure of a Two Zinc-Finger Peptide Reveals an Extension to the Rules for Zinc-finger/DNA Recognition," *Nature* **1993**, 366, 483-487.
- (10) Parraga, G.; Horvath, S.; Eisen, A.; Taylor, W.; Hood, L.; Young, E.; Klevit, R.E. "Zinc-Dependent Structure of a Single-Finger Domain of Yeast ADR1," *Science* **1988**, 241, 1489-1492.

- (11) Weiss, M.A.; Mason, K.A.; Dahl, C.E.; Keutmann, H.T. "Alternating Zinc-Finger Motifs in the Human Male-Associated Protein ZFY," *Biochemistry* **1990**, *29*, 5660-5664.
- (12) Weiss, M. A.; Keutmann, H. T. "Alternating Zinc Finger Motifs in the Male-Associated Protein ZFY: Defining Architectural Rules by Mutagenesis and Design of an "Aromatic Swap" Second-Site Revertant," *Biochemistry* **1990**, *29*, 9808-9813.
- (13) Blanco, F. J.; Rivas, German; Serrano, Luis "A Short Linear Peptide that Folds into a Native Stable β -hairpin in Aqueous Solution," *Nature Structural Biology* **1994**, *1*, 584-589.
- (14) Viguera, A. R.; Blanco, F. J.; Serrano, L. "The Order of Secondary Structure Elements does not Determine the Structure of a Protein but does Affect its Folding Kinetics," *J. Mol. Biol.* **1995**, *247*, 670-681.
- (15) Stroup, A. N.; Gierasch, L. M. "Reduced Tendency To Form a β Turn in Peptides from the P22 Tailspike Protein Correlates with a Temperature-Sensitive Folding Defect," *Biochemistry* **1990**, *29*, 9765-9771.
- (16) Mantelione, G. T. ; Arnold, E.; Meinwald, Y. C.; Stimson, E. R.; Denton, J. B.; Huang, S.-G.; Clardy, J.; Scheraga, H. A. "Chain-Folding Initiation Structures in Ribonuclease A: Conformational Analysis of trans-Ac-Asn-Pro-Tyr-NHMe and trans-Ac-Tyr-Pro-Asn-NHMe in Water and in the Solid State," *J. Am. Chem. Soc.* **1984**, *106*, 7946-7958.
- (17) Wright, P. E.; Dyson, H. J.; Lerner, R. A. "Conformation of Peptide Fragments of Proteins in Aqueous Solutions: Implications for Initiation of Protein Folding," *Biochemistry* **1988**, *27*, 7167-7175.
- (18) Blanco, F. J.; Jiménez, A.; Rico, M.; Santoro, J.; Heranz, J.; Nieto, L. "Tendamistat (12-26) Fragment," *Eur. J. Biochem.* **1991**, *200*, 345-351.

- (19) Yan, Y.; Erickson, B. W.; Tropsha, A. "Free Energies for Folding and Refolding of Four Types of β turns: Simulation of the Role of Chirality," *J. Am. Chem. Soc.* **1995**, *117*, 7592-7599.
- (20) Chalmers, D. K.; Marshall, G. R. "Pro-D-NMe-Amino Acid and D-Pro-NMe-Amino Acid: Simple, Efficient Reverse-Turn Constraints," *J. Am. Chem. Soc.* **1995**, *117*, 5927-5937.
- (21) Yan, Y.; Tropsha, A.; Hermans, J.; Erickson, B. W. "Free Energies for Refolding of the Common β -turn into the Inverse-Common β -turn: Simulation of the Role of D/L Chirality," *Proceedings of the National Academy of Sciences, U.S.A.* **1993**, *90*, 7898-7902.
- (22) Aubry, A.; Cung, M. T. ; Marraud, M. " β I and β II-Turn Conformations in Model Dipeptides with the Pro-Xaa Sequences," *J. Am. Chem. Soc.* **1985**, *105*, 7640-7647.
- (23) Imperiali, B; Fisher, S. L.; Moats, R. A.; Prins, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide.," *J. Am. Chem. Soc.* **1992**, *114*, 3182.
- (24) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. "Folding of Immunogenic Peptide Fragments of Proteins in Water Solution I. Sequence Requirements for the Formation of a Reverse Turn," *J. Mol. Biol.* **1988**, *201*, 161-200.
- (25) Michael, S. F.; Kilfoil, V. J.; Schmidt, M. H.; Amann, B. T.; Berg, J. M. "Metal Binding and Folding Properties of a Minimalist Cys₂His₂ Zinc Finger Peptide," *Proceedings of the National Academy of Sciences* **1992**, *89*, 4796-4800.
- (26) Maret, W.; Vallee, B. L. "Cobalt as Probe and Label of Proteins," *Meth. Enzym.* **1993**, *226*, 52-71.

- (27) Berg, J. M.; Merkle, D. L. "On the Metal Ion Specificity of "Zinc Finger" Proteins," *J. Am. Chem. Soc.* **1989**, *111*, 3759-3761.
- (28) Fisher, S. L.; Imperiali, B. *33rd National Organic Symposium*, **1993**,
- (29) Fisher, S. L., "Unnatural Amino Acids in the Synthesis and Semisynthesis of Metalloprotein Motifs," Ph.D. Thesis, California Institute of Technology, 1993.
- (30) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. "A Consensus Zinc Finger Peptide: Design, High-Affinity Metal Binding, a pH-Dependent Structure, and a His to Cys Sequence Variant," *J. Am. Chem. Soc.* **1991**, *113*,
- (31) Johnson, Jr., W. C. "Protein Secondary Structure and Circular Dichroism: A Practical Guide," *Proteins: Struct. Func. Gen.* **1990**, *7*, 205-214.
- (32) Fasman, G.D. "The Development of the Prediction of Protein Structure," in *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G.D. Eds.; Plenum Press: New York, 1989; pp 193-316.
- (33) Richardson, J. S.; Richardson, D. C. "Amino Acid Preferences for Specific Locations at the Ends of α Helices," *Science* **1988**, *240*, 1649-1652.
- (34) Presta, L. G.; Rose, G. D. "Helix Signals in Proteins," *Science* **1988**, *240*, 1632-1641.
- (35) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley and Sons: New York, 1986.
- (36) Wilmot, C.M.; Thornton, J.M. "Analysis and Prediction of the Different Types of β -Turn in Proteins," *J. Mol. Biol.* **1988**, *203*, 221-232.
- (37) Sibanda, B.L.; Thornton, J.M. " β -Hairpin Families in Globular Proteins," *Nature* **1985**, *316*, 170-174.
- (38) Yao, J.; Feher, V. A.; Espejo, B. F.; Reymond, M. T.; Wright, P. E.; Dyson, H. J. "Stabilization of a Type VI Turn in a Family of Linear Peptides in Water Solution," *J. Mol. Biol.* **1994**, *243*, 736-753.

- (39) Grathwohl, C.; Wüthrich, K. "The X-Pro Peptide Bond as an Nmr Probe for Conformational Studies of Flexible Linear Peptides," *Biopolymers* **1976**, *15*, 2025-2041.
- (40) Carpino, L. A. ; El-Faham, A.; Minor, C. A.; Albericio, F. "Advantageous Applications of Azabenzotriazole (Triazolopyridine)-based Coupling Reagents to Solid-phase Peptide Synthesis," *J. Chem. Soc. Chem. Comm.* **1994**, 201-203.
- (41) King, D. S.; Fields, C. G.; Fields, G. B. "A Cleavage Method Which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis," *Int. J. Pep. Protein Res.* **1990**, *36*, 255-266.
- (42) Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*; John Wiley and Sons: New York, 1987.
- (43) Bax, A.; Davis, D.G. "MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy," *J. Magn. Res.* **1985**, *65*, 355-360.
- (44) Griesinger, C.; Otting, G.; Wüthrich, D.; Ernst, R. R. "Clean TOCSY for ^1H Spin System Identification in Macromolecules," *J. Am. Chem. Soc.* **1988**, *110*, 7870-7872.
- (45) Piantini, U.; Sørensen, O.W.; Ernst, R.R. "Multiple Quantum Filters for Elucidating NMR Coupling Networks," *J. Am. Chem. Soc.* **1982**, *104*, 6800-6801.
- (46) Bax, A.; Davis, D. G. "Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy," *J. Magn. Res.* **1985**, *63*, 207-213.
- (47) States, D.J.; Haberkorn, R.A.; Ruben, D.J. "A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants," *J. Magn. Res.* **1982**, *48*, 286-292.

Chapter 4

BBA1: Structural analysis of an independently
folded $\beta\beta\alpha$ motif

Introduction

The protein biopolymer represents nature's most diverse biomolecule. Proteins span an incredible range of sizes and shapes:¹ From proteins such as viral coat proteins or large enzyme complexes which reach fantastic sizes comprising thousands of amino acids, to the smallest enzymes consisting of a mere hundred or less residues.^{1, 2} The smallest size necessary for independently folded tertiary structure within a natural protein or protein domain is generally considered to be on the order of 70-100 amino acids.^{1, 2} Exceptions to this limit are found in a variety of proteins or protein domains including peptidyl toxins such as charybdotoxin (37 amino acids),³ protease inhibitors such as BPTI (basic pancreatic trypsin inhibitor, 56 amino acids), or transcription factor domains such as the zinc finger peptides (25-30 amino acids).⁴ These proteins or peptides are able to circumvent the apparent size limitation through the utilization of extensive cross-linking in the form of metal binding sites or disulfides to create a stable and defined structure. While many naturally occurring small peptides have been shown to have extensive secondary structure, especially helices, very few exhibit the defined tertiary interactions that are the hallmark of native protein structure in the absence of cross-links.

A key issue in protein design is the minimal size of an independently folded peptidyl motif. Iterative modeling, synthesis, and structural characterization are facilitated for smaller constructs, allowing for the efficient testing of design principles and strategies. NMR is one of the most convenient methods for the detailed structural characterization of designed polypeptides. This technique is restricted to smaller polypeptides, especially if one wants to avoid the use of expensive labeled amino acids or 3D NMR. The large size and aggregation state of many protein design efforts has limited the ability to

adequately characterize the resulting peptidyl constructs by NMR.⁵ In the absence of the detailed structural information provided by NMR or crystal structure data, the analysis of designed proteins has relied heavily on auxiliary, indirect methods to probe tertiary structure (Chapter 1).^{6,7} Circular dichroism (CD) is most frequently used to investigate the *secondary* structural content, while dispersity in the 1D NMR spectra, cooperative denaturation (usually monitored by CD) and the absence of ANS (8-anilino-1-naphthalenesulfonic acid) binding, are used as indicators of tertiary structure.⁶⁻⁸ While these techniques do yield information on the *presence* of tertiary structure, they do not reveal the nature of the interactions responsible for such behavior. For these reasons, our efforts have focused on designing systems which are more amenable to complete and detailed structural analysis by NMR techniques.

The project presented in the previous chapter targeted the design and assembly of a $\beta\beta\alpha$ motif based on the zinc finger domains. These peptides are the DNA binding units in a large class of transcription factor proteins.^{4,9} Although zinc finger peptides adopt a well defined three-dimensional architecture in approximately 30 residues, the structure of these peptides is critically dependent on ligation of a divalent metal ion.^{10,11} Zinc coordination by zinc fingers and other zinc dependent transcription factors has been referred to as the "intracellular version of the disulfide,"⁴ forming cross-links that are absolutely required for the formation of structure in these small protein domains. The reproduction of this defined structure in the *absence* of metal cations is therefore a challenging goal. If small motifs, such as the $\beta\beta\alpha$ structure of the zinc fingers, are to be used as versatile structural templates for future functional design efforts then the requirement for cross-links must be eliminated. In the zinc finger domains, four out of 25-30 residues (~10 - 15%) are involved in forming the

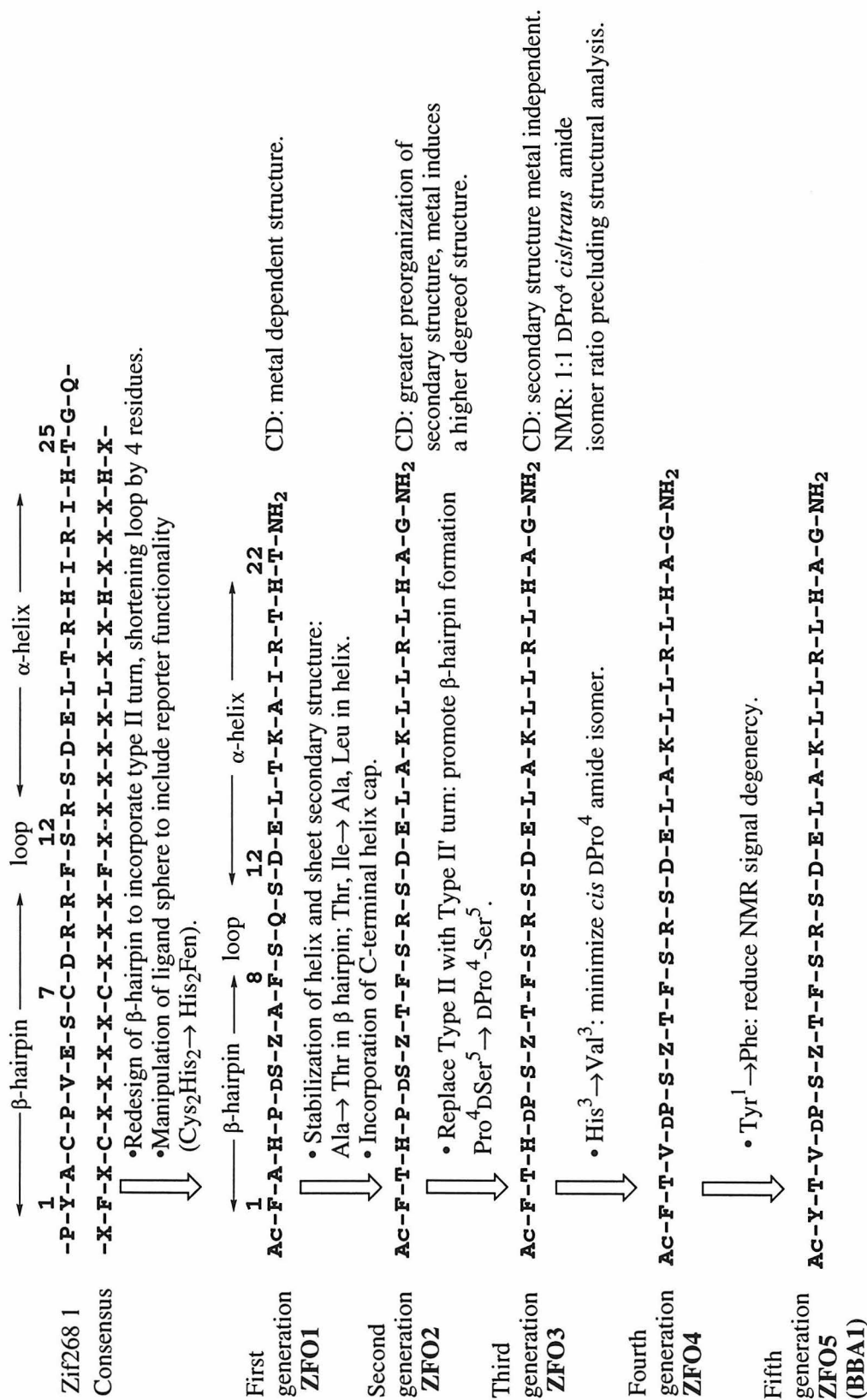
necessary metal binding site,⁹ severely limiting the versatility of this system for the development of functional polypeptides.

Building on the results described in the last chapter a metal independent $\beta\beta\alpha$ motif was achieved in only two additional design iterations (for a total of five), demonstrating that the $\beta\beta\alpha$ architecture of the zinc finger peptides can be reproduced in as few as 23 amino acids in the *absence* of metal. These sequences along with those described in the last chapter are summarized in Figure 4-1. The sequence, -Val-DPro-Ser-Fen-, was found to form a β -turn in solution with a *cis:trans* proline amide isomer ratio of 1:18. Therefore this sequence was incorporated into the motif, to produce the fourth generation design, **ZFO4**. Although NMR analysis of this peptide revealed the presence of defined secondary and tertiary structure, some degeneracy in the NMR resonances necessitated the synthesis of an additional peptide, **ZFO5** with a minor sequence alteration (Phe¹→Tyr). Structural analysis of this peptide by 2D NMR confirmed the presence of defined secondary and tertiary structure for **ZFO5**, and NOE restrained simulated annealing was used to solve the three-dimensional solution structure. Several methods traditionally used to probe designed peptides in the absence of an NMR or crystal structure such as chemical and thermal denaturation, ANS binding, and H/D exchange studies were also performed on **ZFO5**. As the first successfully designed metal independent $\beta\beta\alpha$ motif, this peptide was renamed **BBA1**.

Additional peptides were synthesized and analyzed to probe two unique features of **BBA1**: the type II' turn and the role of the unnatural amino acid Fen. A peptide, **ZFO6**, in which the type II' turn of **BBA1** was replaced with a type II turn was synthesized and analyzed. The analysis of this peptide highlighted the importance of the type II' turn in the generation of both secondary and tertiary structure in **BBA1**. Inspection of the NMR derived solution structure, indicated

that Fen⁶ did not play a direct structural role within the motif. Nevertheless, analysis of a peptide with a Fen⁶→Tyr replacement, **BBA2**, was performed to determine the tolerance of the motif to substitutions at this position. The analysis of peptides **ZFO6** and **BBA2** helped define the interactions which contribute to the adoption of a defined three-dimensional structure in only 23-residues in **BBA1**.

Figure 4-1 Summary of the evolution of a metal independent $\beta\beta\alpha$ motif based on the zinc finger domains. The design of the five peptides evaluated is detailed. The analysis of the first three generations is summarized on the right and described in Chapter 3.



Results

ZFO4

This fourth generation peptide, **ZFO4**, contained a His³→Val substitution to control the *cis/trans* proline amide isomerization problem encountered in the previous generation design, **ZFO3** (Chapter 3). CD and NMR studies were performed to investigate the structure.

CD Analysis

The CD signature of uncomplexed **ZFO4** resembles those of natural zinc finger peptides. Unlike previous designs the addition of Zn²⁺ resulted in a slight increase of the ellipticity at 222 nm, indicating a small loss of secondary structure (Figure 4-2A). The CD spectrum (190 nm - 250 nm) was invariant with peptide concentration (11 μM - 0.95 mM) and pH (4.5 - 8.0). In contrast to earlier designs (**ZFO1** and **ZFO2**), there was very little change in the near-UV region of the CD spectrum which indicated that the phenanthroline moiety does not experience a chiral environment in the metal bound state (Figure 4-2B). UV studies monitoring the phenanthroline chromophore confirmed the formation of a 1:1 **ZFO4**-Zn²⁺ complex involving the Fen amino acid. These results suggested that metal binding by the peptide no longer had a strong influence on secondary or tertiary structure, and may have involved solely the Fen residue.

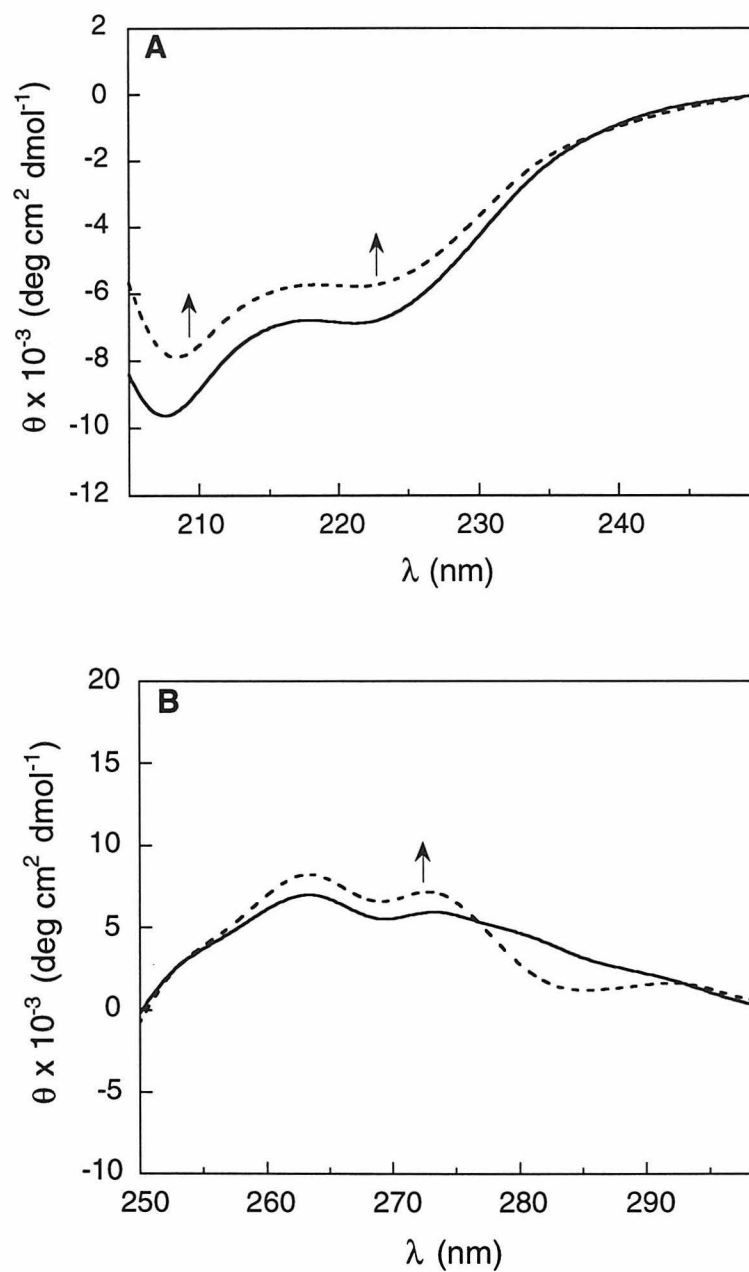


Figure 4-2 CD analysis of **ZFO4**. **A.** Far-UV CD spectrum of **ZFO4** uncomplexed (solid line) and after the addition of 3 equivalents of Zn²⁺ (dashed line) **B.** Near-UV CD spectrum of **ZFO4** uncomplexed (solid line) and after the addition of 3 equivalents of Zn²⁺ (dashed line). The spectra have been normalized to the concentration of the phenanthroline moiety.

NMR Analysis

The structure of uncomplexed **ZFO4** was also investigated by 2D NMR techniques. These studies confirmed that the His³→Val substitution resulted in a peptide that had an acceptable level of the *trans* proline amide isomer (80 - 90 %). The *cis:trans* isomer ratio was estimated from the integration of the methyl protons of Val³ in the *cis* isomer and compared to the integration of the corresponding peaks in the *trans* isomer. Although observable in the 2D NMR spectra, cross-peaks resulting from the *cis* isomer were very weak and did not complicate assignment. Complete sequence specific assignments of the *trans* isomer were possible, which permitted the investigation of the solution conformation of **ZFO4** through the analysis of 2D NOESY spectra.

NOESY analysis in 90:10 H₂O:D₂O indicated that the desired elements of secondary structure were present in **ZFO4**. Figure 4-3 summarizes the secondary structural content of **ZFO4**. The NOESY signature for an alpha helix includes $d_{NN}(i, i+1)$, $d_{\alpha N}(i, i+3)$, and $d_{\alpha\beta}(i, i+3)$.¹² All three of these characteristic NOEs were present for residues 12-22 of **ZFO4**, indicating a well formed alpha helix in this region. The sheet region of the peptide was also defined by a number of key NOEs. These NOEs are illustrated and numbered in Figure 4-3B. Specifically the long-range $d_{NN}(\text{Val}^3, \text{Fen}^6)$ (1), $d_{\alpha N}(\text{Thr}^7, \text{Val}^3)$ (2), $d_{\alpha N}(\text{Thr}^2, \text{Phe}^8)$ (4) and the strong sequential $d_{\alpha N}(i, i+1)$ (9-12) indicated that residues 1-8 an antiparallel β -sheet. Observation of $d_{NN}(\text{Ser}^5, \text{Fen}^6)$ (7) and $d_{\alpha N}(\text{Pro}^4, \text{Fen}^6)$ (8), as well as strong sequential NOEs, $d_{\alpha N}(\text{Pro}^4, \text{Ser}^5)$ (5) $d_{\alpha N}(\text{Pro}^4, \text{Ser}^5)$ (6) indicated the proper formation of a type II' turn.¹² Although assignment of NOE #3 [$d_{\alpha\alpha}(\text{Thr}^2, \text{Thr}^7)$] in the NOESY spectrum acquired in 90:10 H₂O:D₂O was not possible due to interference from the water signal, this NOE was easily identified in spectra taken in 100% D₂O.

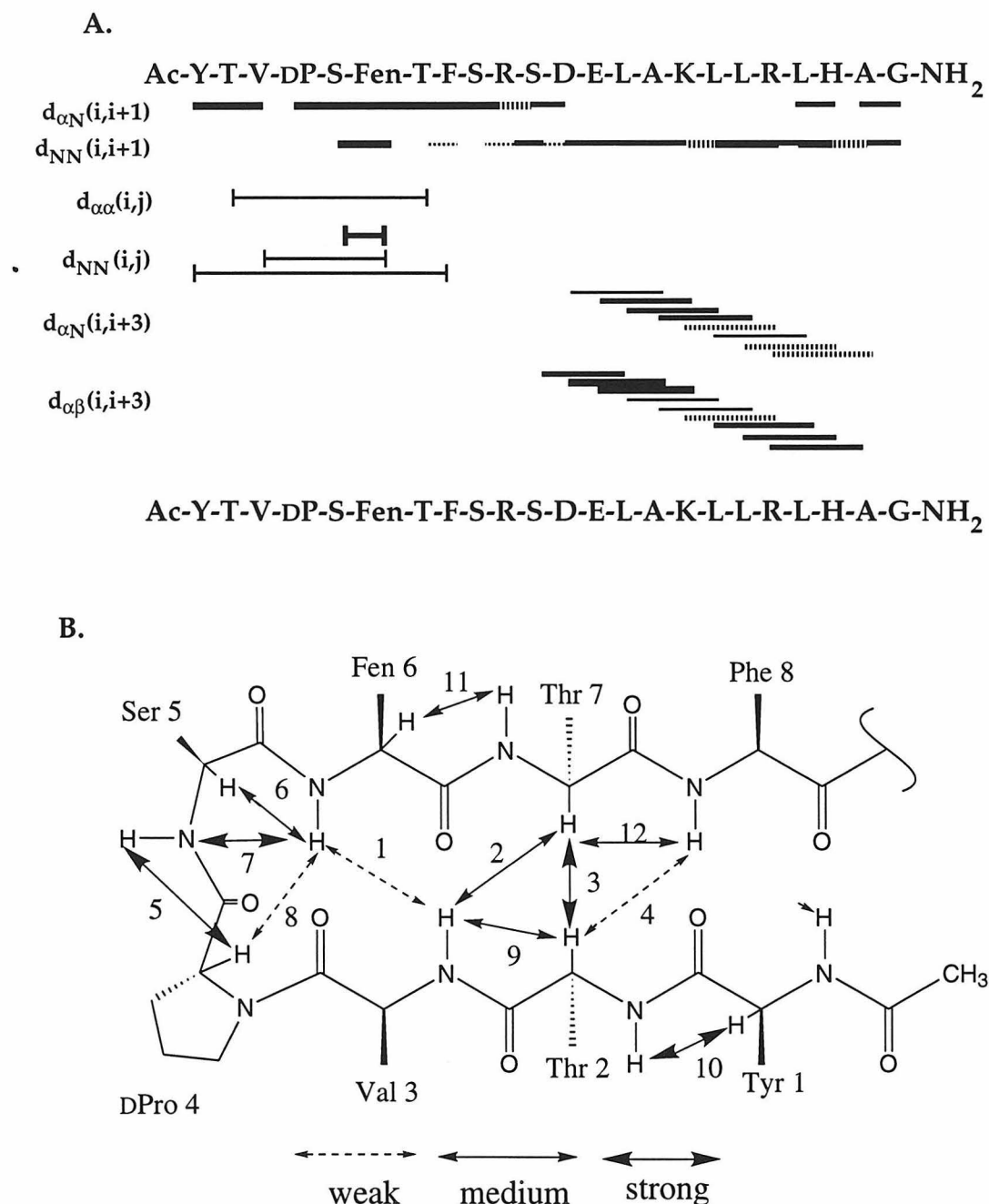


Figure 4-3 **A.** NOEs indicative of secondary structure in uncomplexed **ZFO4** are illustrated. The width of the lines represent the relative strength of the NOEs, a dashed line indicates that the NOE could not be identified due to overlap. **B.** Schematic representation of key NOEs defining the β -hairpin. The relative intensities of the NOEs were determined from inspection of the cross-peaks in the NOESY spectrum (200 ms mixing time).

The β -sheet and α -helical secondary structures were also supported the amide temperature coefficients for **ZFO4**. A low amide temperature coefficient ($-\Delta\delta/\Delta T < 6.0$ ppb/K) is an indication of the shielding of the amide proton from solvent. One probable cause of solvent shielding is hydrogen bonding interactions¹³ such as those found in elements of secondary structure. Therefore, the pattern of amide proton shielding can be a confirmation of the secondary structure. Based on the original sequence specific assignments at 7 °C, additional TOCSY experiments performed at various temperatures (17 °C, 27 °C, and 37 °C) were used to assign amide chemical shifts and determine the amide temperature coefficients. Although the majority of residues could be assigned for all four temperatures, the amide temperature coefficients of several residues could not be accurately determined. The results of this study are summarized in Table 4-1. Several residues in the sheet have amide temperature coefficients consistent with the hydrogen bonding pattern of this element of secondary structure: Phe¹ and Fen⁶ have low values ($-\Delta\delta/\Delta T < 5.0$ ppb/K), while Thr² and Thr⁷ have higher values ($-\Delta\delta/\Delta T \sim 6.0$ ppb/K). The shielding pattern of residues 12-22 was also consistent with the expected structure in this region. The NOEs illustrated in Figure 4-3 identify an α helix consisting of residues 12 - 22. In this case, the first amide proton predicted to participate in a hydrogen bond would be residue Ala¹⁵ (with additional ($i, i+4$) hydrogen bonds expected for residues 15-22). Indeed, the amide proton temperature coefficients were generally low ($-\Delta\delta/\Delta T < 4.0$ ppb/K) for most residues in this region. Unexpectedly, two additional positions had low amide proton temperature coefficients, Ser¹¹ and Glu¹³. Ser¹¹ is in the loop connecting the helix and the sheet, while Glu¹³ is at the N-terminus of the helix. Based on the NOE analysis, these residues should not be involved in backbone-backbone hydrogen bonds. However, other interactions such as a side chain to backbone hydrogen bonds or hydrophobic shielding by neighboring

residues might protect these protons from solvent, resulting in depressed amide temperature coefficients.

Residue	$-\Delta\delta/\Delta T$ (ppb/K)	expected solvent shielding
Phe1	4.76	yes (sheet)
Thr2	6.07	no (sheet)
Val3	nd	yes (turn)
Ser5	nd	no (turn)
Fen6	3.36	yes (turn)
Thr7	5.95	no (sheet)
Phe8	nd	yes (sheet)
Ser9	5.6	no
Arg10	9.78	no
Ser11	3.62	no
Asp12	nd	no (helix)
Glu13	3.65	no (helix)
Leu14	6.83	no (helix)
Ala15	3.19	yes (helix)
Lys16	4.73	yes (helix)
Leu17	3.42	yes (helix)
Leu18	3.98	yes (helix)
Arg19	2.48	yes (helix)
Leu20	4.22	yes (helix)
His21	2.98	yes (helix)
Ala22	nd	yes (helix)
Gly23	5.73	no

Table 4-1 Amide proton temperature coefficients for **ZFO4**. nd = not determinable. The right-hand column indicates if shielding of this amide proton is expected within the element of secondary structure determined for this residue from the NOE analysis. A low amide temperature coefficient ($-\Delta\delta/\Delta T < 5.0$ - 6.0 ppb/K) is an indication of the shielding of the amide proton from solvent.

In addition to the desired secondary structure, analysis of the NOESY spectrum revealed a number of long range NOEs between the α -helix and the β -sheet, which indicated a defined tertiary structure for **ZFO4**. Several of these NOEs are shown in Figure 4-4. In the absence of a metal binding, interactions between the helix and the sheet should be mediated by Phe¹, Phe⁸, and Leu¹⁴; the analogous residues in folded natural zinc fingers make up the hydrophobic cluster in the presence of metal cations. The NOEs observed for **ZFO4** indicated the involvement of Phe¹, Phe⁸ and Leu¹⁴ in the formation of a similar hydrophobic cluster in the absence of metal. Assignment of several of these key NOEs was hindered by overlapping resonances for the aromatic rings of Phe⁸ and Phe¹. Although there was considerable evidence for defined secondary and tertiary structure in **ZFO4**, the spectral overlap of the Phe resonances made translating the NMR data into a three-dimensional structure difficult. Tyrosine is often found at the analogous position in natural zinc fingers and its ring protons have distinctive chemical shifts that are well separated from those of phenylalanine. Therefore, the peptide, **ZFO5**, containing a conservative Phe¹→Tyr replacement, was synthesized.

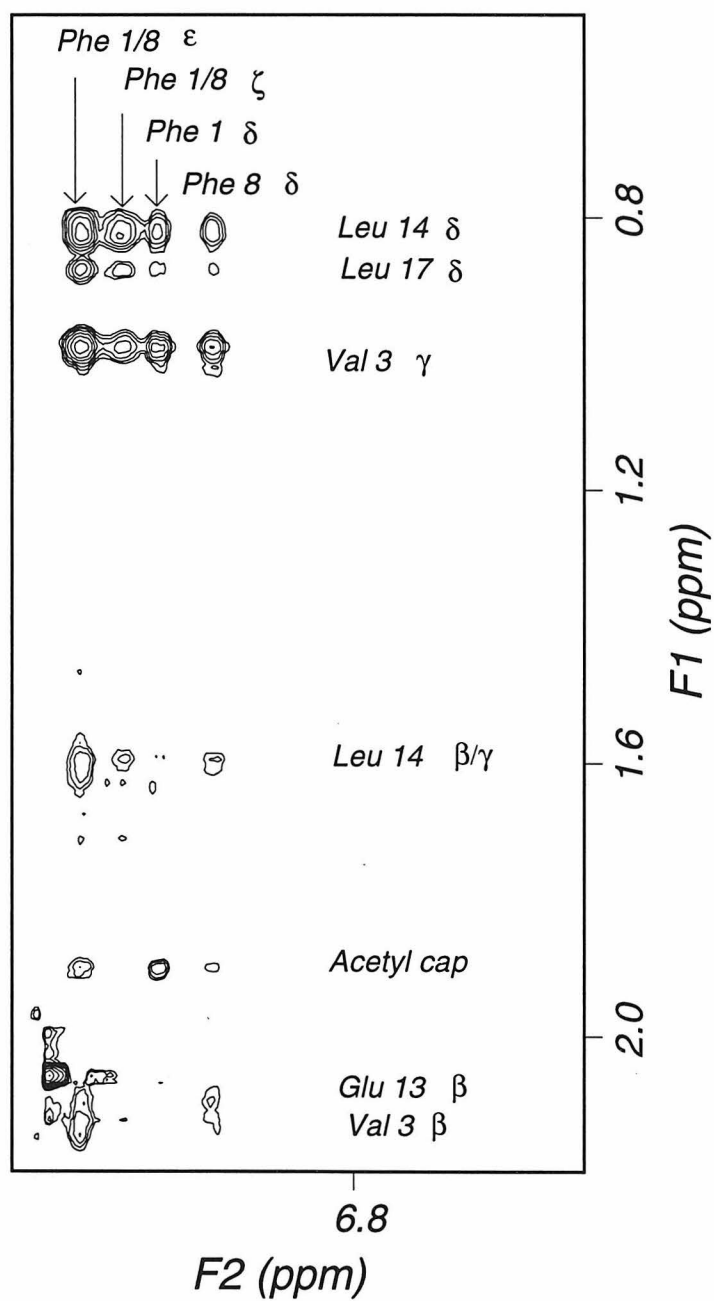


Figure 4-4 NOESY spectrum of ZFO4. Long range NOEs between residues in the helix (*Leu*¹⁴, *Leu*¹⁷) and the sheet (*Phe*¹ and *Phe*⁸) are indicated.

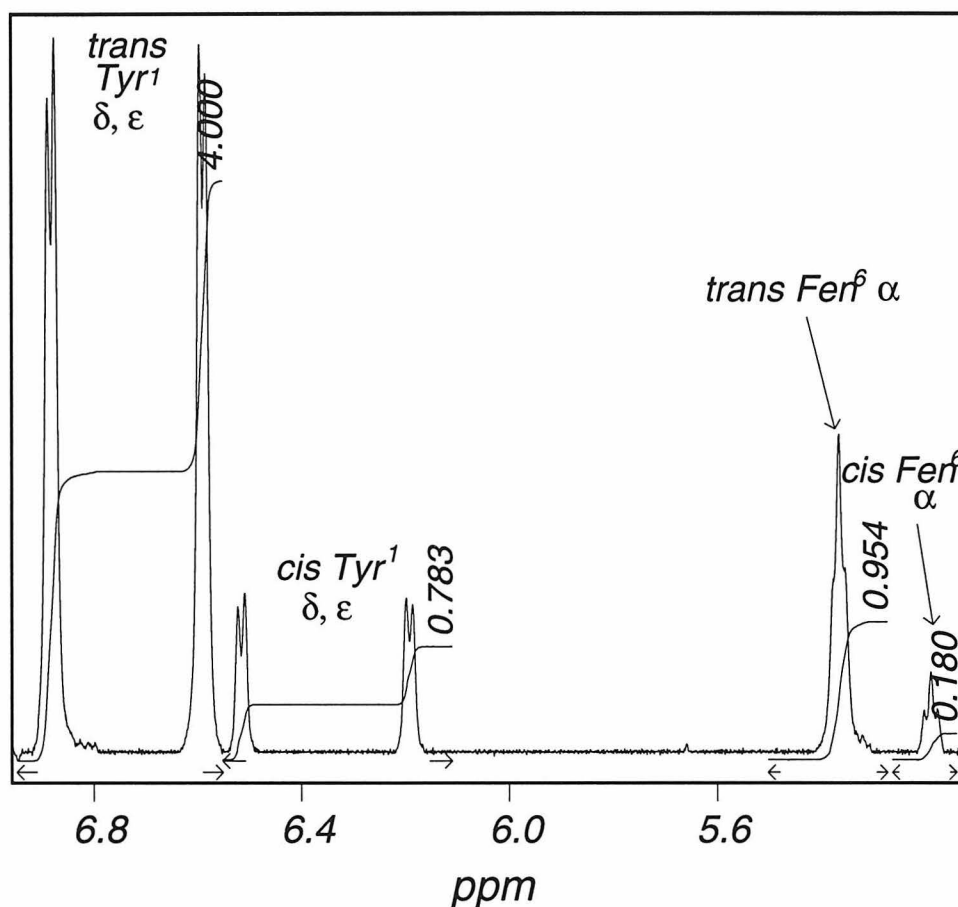
BBA1 (ZFO5): A Metal Independent $\beta\beta\alpha$ Motif

Figure 4-5 Measurement of the *cis-trans* isomer ratio for **BBA1** by 1D NMR. A portion of the 1D spectrum of uncomplexed **BBA1** acquired in D₂O. Average of 1:5.2 *cis/trans* or 84% *trans* proline amide isomer.

The fifth generation design, **ZFO5** (or **BBA1**), included the optimized type II' β -turn and a Phe¹→Tyr replacement to decrease NMR signal degeneracy and to allow complete structure determination. (This peptide will hereafter be referred to as **BBA1**). A number of proton resonances of the *cis* and *trans* proline amide isomers of **BBA1** were sufficiently well resolved in the 1D ¹H NMR

spectra to allow for the quantification of the amide isomer ratio (Figure 4-5). Integration of these signals indicated ~16% ($\pm 3\%$ depending on the peaks used) of the *cis* conformer. While this value is higher than expected based on the study of the model peptide presented in Chapter 3, cross-peaks arising from the *cis* conformer were easily identified and sufficiently weak, and therefore did not interfere with the assignment and analysis of the 2D NMR spectra.

CD Analysis

The CD spectrum of **BBA1** strongly resembles those of the natural zinc fingers.^{11, 14} Similar to **ZFO4**, the addition of Zn^{2+} resulted in only a slight change in the CD spectrum in the 190-250 nm range, which indicated that secondary structure is metal *independent* (Figure 4-6A). In the first two designs, concomitant with the increase of the secondary structure upon addition of Zn^{2+} , a strong CD signal was observed in the near UV region indicating the formation of a chiral Fe-metal complex. In contrast a much weaker signal was observed for **BBA1**, consistent with the idea that metal binding no longer dominated secondary and tertiary structure (Figure 4-6B). The formation of a 1:1 **BBA1**- Zn^{2+} complex ($K_D = 4.56 \pm 1.78 \mu\text{M}$) was confirmed by UV-Vis spectroscopy. The CD spectrum of uncomplexed **BBA1** was invariant with pH (4.5-8.5) and peptide concentration (8.2 μM -2.6 mM), indicating the absence of aggregation effects.

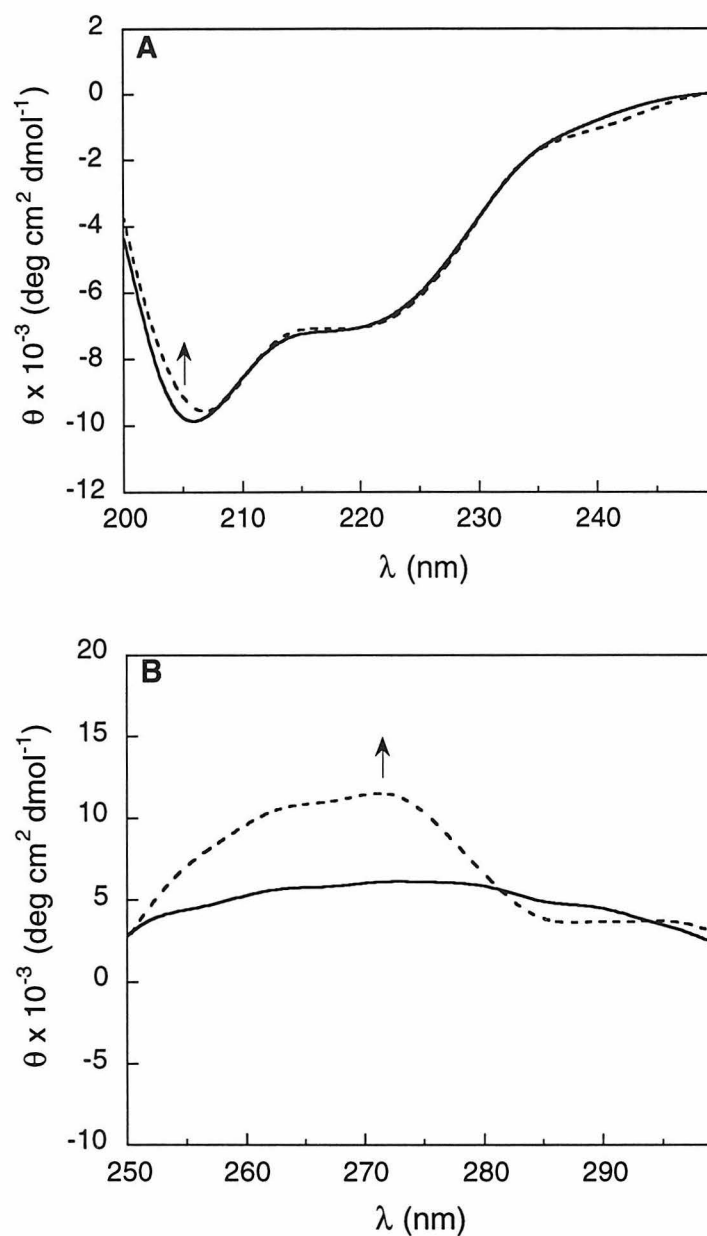


Figure 4-6 CD Analysis of **BBA1**. **A.** Far-UV CD spectrum of uncomplexed **BBA1** (solid line) and after the addition of Zn^{2+} (dashed line). The spectra are plotted in mean residue ellipticity. **B.** Near-UV CD spectrum of **BBA1** uncomplexed (solid line) and after the addition of Zn^{2+} (dashed line). The spectra are normalized to the concentration of the phenanthroline chromophore.

NMR Analysis

The complete sequence specific assignment was performed for **BBA1**.¹² Analysis of the NOESY spectra, acquired in both 90:10 H₂O:D₂O and D₂O, for the uncomplexed peptide revealed the presence of well defined secondary structure in good agreement with the design and similar to the results obtained for **ZFO4**. NOEs indicative of helix formation [$d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$] were observed for residues 12-22. Formation of the β -hairpin between residues 1 and 8 was indicated by an extensive network of NOEs (Figure 4-7).

The assignment of secondary structure was also confirmed by H/D exchange studies. In these experiments the exchange of the amide protons with deuterium from the solvent (D₂O) was monitored by 1D NMR. The difference 1D NMR spectra used in this analysis are illustrated in Figure 4-8. These spectra were obtained by subtracting the 1D ¹H NMR spectra acquired ~ 3 hrs in which all exchangeable protons were assumed to be completely deuterated. This was necessary in order to eliminate the peaks arising from the aromatic protons of the Phe amino acid, and allow the partial identification of protected amide protons. These one-dimensional NMR experiments indicated that many amide protons of **BBA1** were protected from solvent exchange. While some protons exchanged rapidly (~6 min), several amide proton signals were still visible after one hour. The exchange half time for an amide proton within a peptidyl in a random coil conformation at this temperature and pH is less than three minutes.¹⁵ Therefore the slow exchange observed for **BBA1** is evidence of considerable protection, due to secondary and tertiary structure. Although not all of the protected protons could be assigned in the 1D spectra, at least one amide proton in the β -sheet (Val³) and several in the α -helix (Leu¹⁷, Leu¹⁸, and Ala¹⁵), exhibited considerably slower exchange, in agreement with the predicted structure.

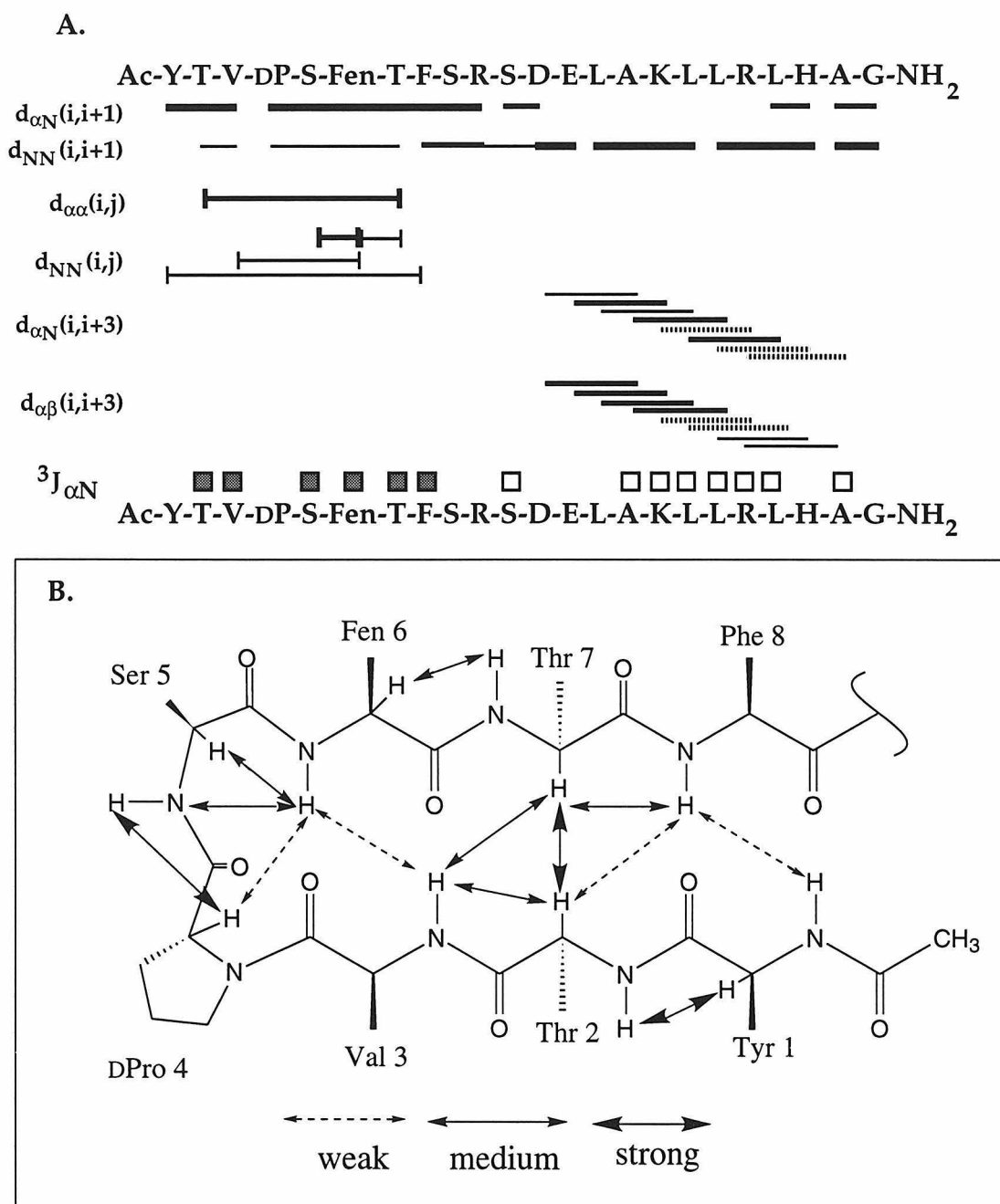


Figure 4-7 **A.** NOEs indicative of secondary structure in **BBA1**. The width of the lines represent the relative strength of the NOEs, a dashed line indicates that the NOE could not be identified due to overlap. Solid squares indicate $^3J_{\alpha N} > 8$ Hz, and open squares indicate $^3J_{\alpha N} < 6.8$ Hz. **B.** Schematic representation of key NOEs defining the β -hairpin. The relative intensities of the NOEs are determined from observation in the NOESY spectrum (200 ms mixing time).

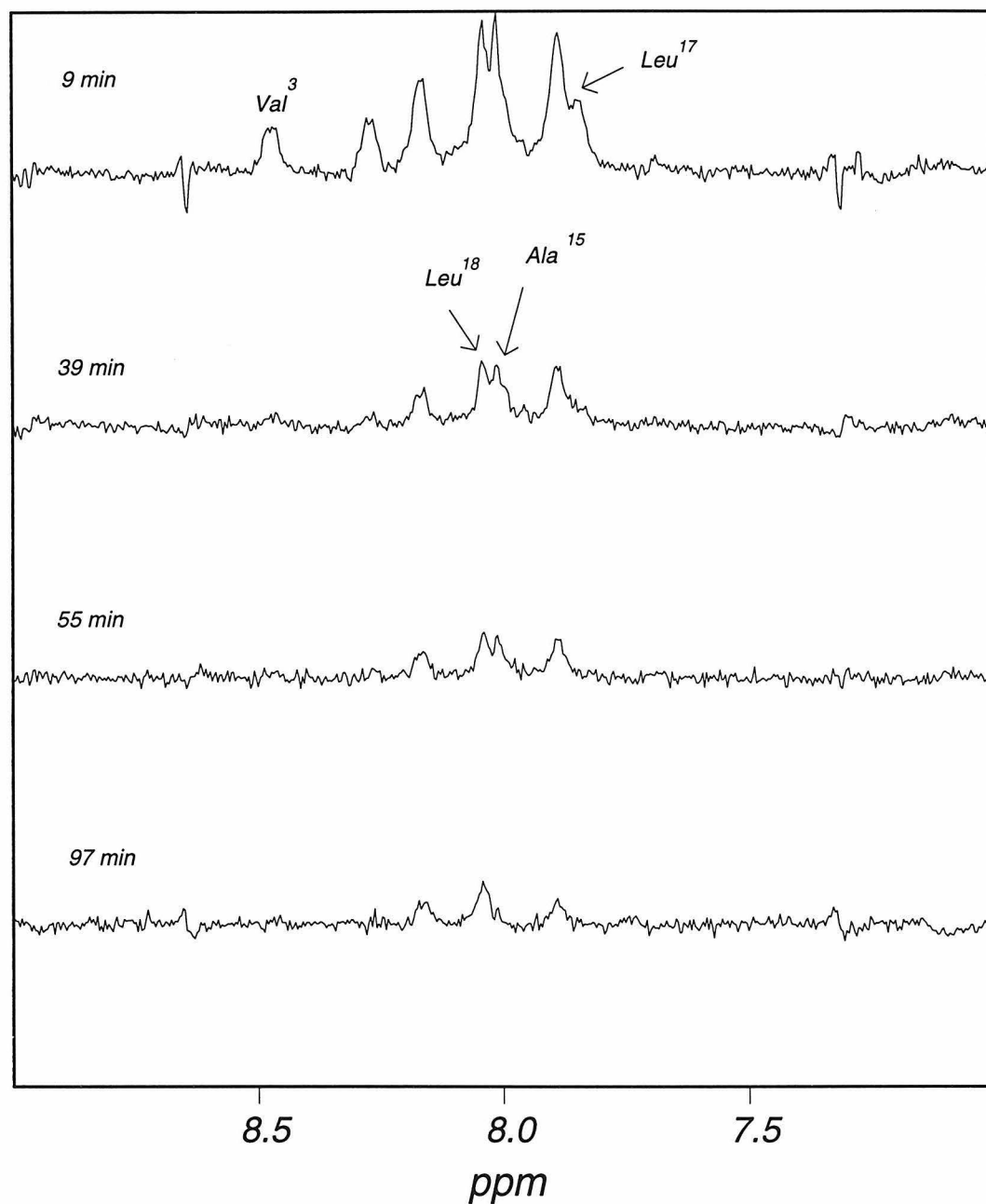


Figure 4-8 H/D exchange experiment for **BBA1** monitored by NMR. The difference spectra are produced by subtracting the 1D spectra after ~ 3 hrs of exchange with D₂O. The sample was freshly dissolved (t=0 min) in 50 mM acetate-*d*₆, pD (uncorrected meter reading) 4.5, 3.2 mM peptide, 7 °C.

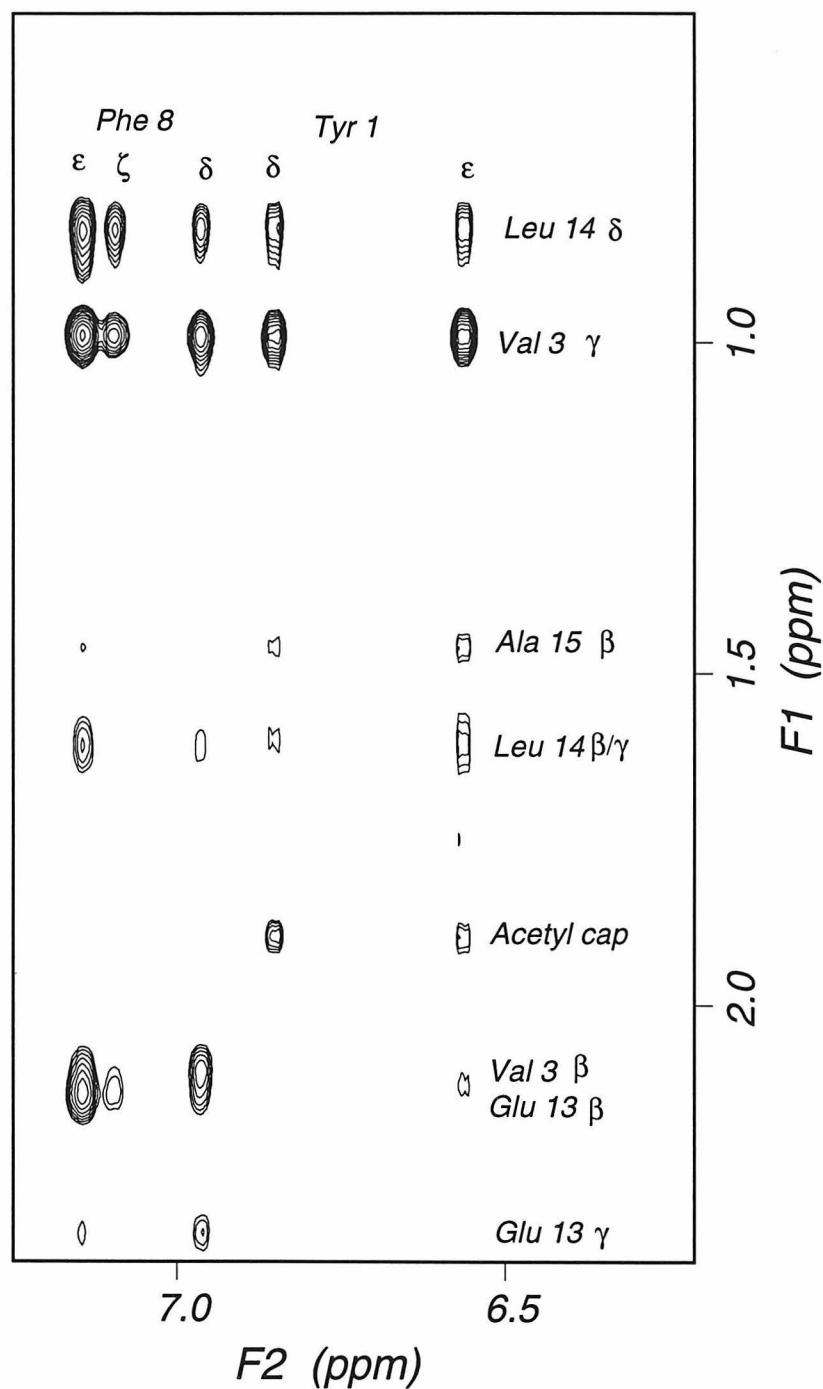


Figure 4-9 Long range NOEs present in the NOESY (300 ms mixing time) spectrum of **BBA1** indicative of tertiary structure. Data acquisition in D₂O, pD (uncorrected) 4.5, 600 MHz, 7 °C.

In addition to the NOESY analysis and H/D exchange studies, the alpha amide ($C_{\alpha}H - NH$) coupling constants for **BBA1** were determined. This data further confirmed the NOE based determination of secondary structure. The splitting of cross-peaks in the double quantum filter COSY spectrum were used to measure the $^3J_{\alpha N}$ coupling constants.¹² Unlike the NOESY and H/D exchange experiments this data was acquired at 25 °C due to the peak broadening observed at 7 °C. The coupling constant data is summarized in Figure 4-7.

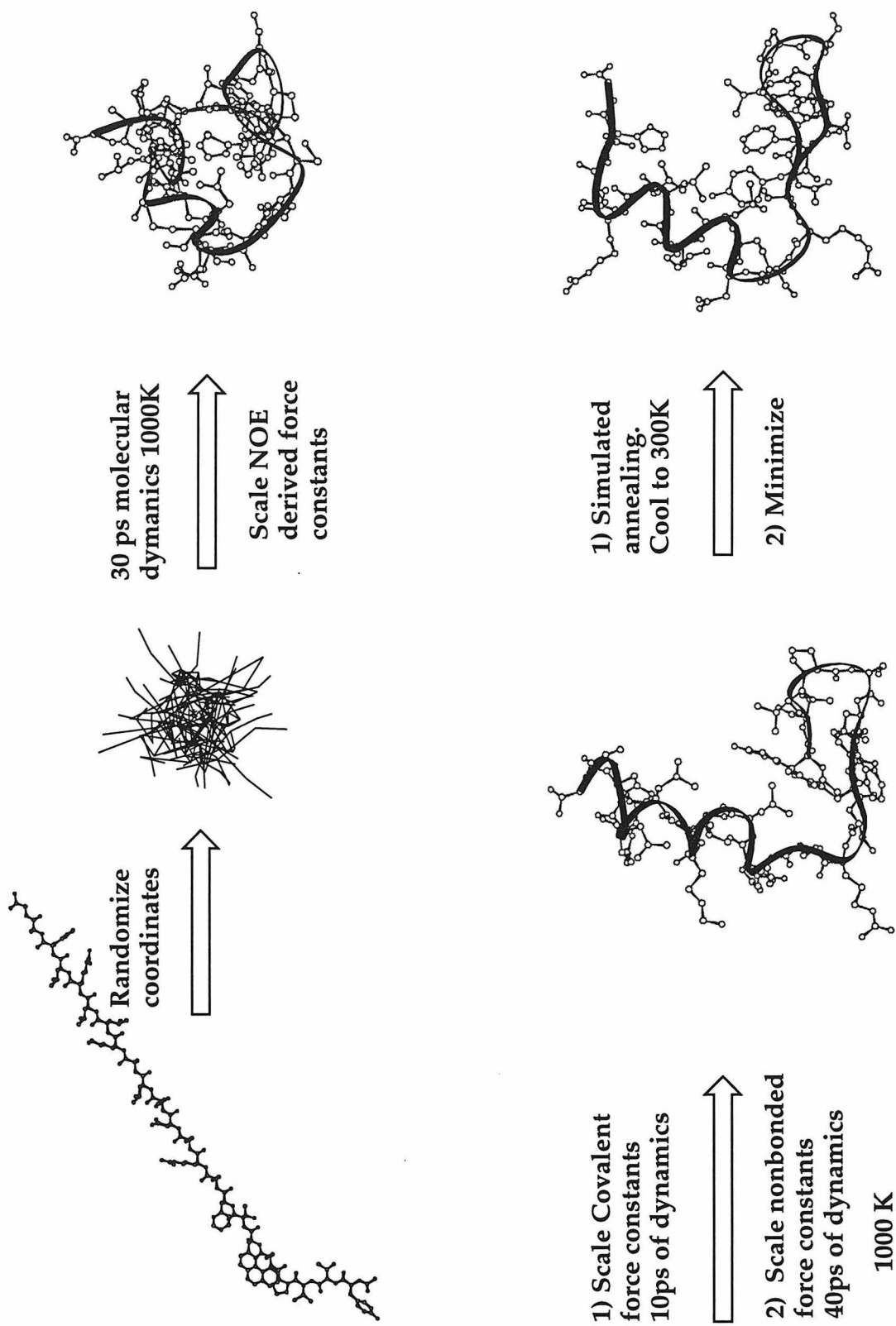
In addition to adopting the desired secondary structure, 2D NMR analysis indicated that **BBA1** folds with a defined tertiary structure. Specifically, long range NOEs between protons of the side chains of residues Glu¹³, Leu¹⁴, Ala¹⁵, Leu¹⁷, and Leu¹⁸ within the helix, and residues Tyr¹, Val³, and Phe⁸ in the β -hairpin revealed the formation of a discrete hydrophobic core. A portion of the 2D NOESY spectrum illustrating several of the diagnostic long range NOEs is shown in Figure 4-9.

Structure Determination

The three-dimensional structure of **BBA1** was determined using an NOE-restrained simulated annealing molecular dynamics protocol,¹⁶ employing a total of 135 intra-residue and 166 inter-residue distance restraints derived from 367 NOEs. In this procedure, the solution structure of a peptide or protein is solved through a molecular dynamics protocol that involves energy terms for the NOE restraints derived from the measurement of cross-peak volumes in NOESY spectra of various mixing times (50 - 300 ms). The initial steps involve the randomization of the coordinates of the starting structure to prevent any bias from the starting conformation. The force constants for the NMR restraints and the covalent and non-bonded energy terms are scaled up in separate steps of molecular dynamics performed at 1000 K. The structure is then cooled from 1000

K to 300 K in five steps over 20 ps. A series of minimizations produce the final structure. The procedure which was used for the structure determination of **BBA1** is illustrated schematically in figure 4-10. Forty-five structures were generated for **BBA1**, 29 of which converged to a common fold.

Figure 4-10 (following page) Schematic representation of a NOE-restrained molecular dynamics simulated annealing procedure. The structures shown are taken from the trajectory files of a single dynamics run used to generate one of the structures shown in Figure 4-11.



The 29 structures generated for **BBA1** agree very well with the intended design (Figure 4-11 and Figure 4-12). The α -helix and the β -hairpin are well defined, and are associated *via* interactions between the side chains of residues Tyr¹, Val³, Phe⁸, Leu¹⁴, Leu¹⁷, and Leu¹⁸. The three residues Tyr¹, Phe⁸, and Leu¹⁴, which were involved in the majority of long-range NOE contacts, form the hydrophobic core of the molecule. Superimposition of the 29 structures revealed that the conformations of the side chains of these residues are as well ordered as the backbone atoms, indicating a well defined interior (Figure 4-12). The conformation of Tyr¹ was the most disordered of the core residues, probably reflecting its terminal location. In contrast to the residues of the hydrophobic cluster, the C-terminal glycine and the side chains of the arginine and lysine residues appeared rather flexible and disordered in the NMR structures.

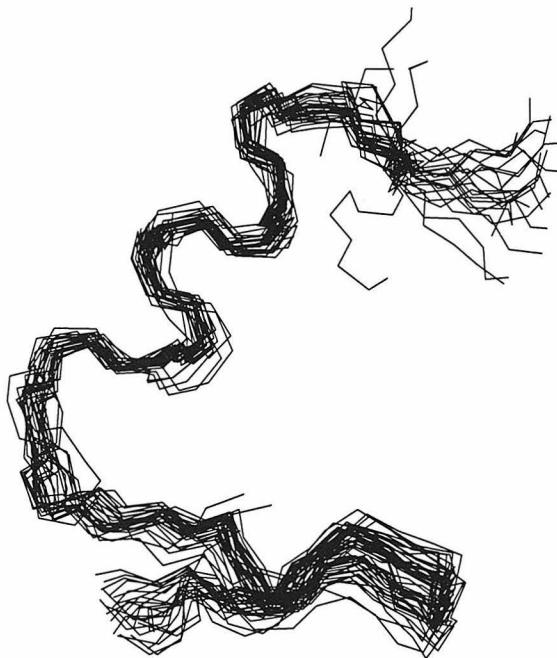


Figure 4-11 Overlay of the backbone atoms of the 29 NMR-derived structures of **BBA1**.



Figure 4-12 Overlay of the heavy atoms of all 29 **BBA1** structures. The structures were superimposed based on the alignment of the backbone atoms of residues 1-22.

The quality of the BBA1 structure compares well with that of Xfin,¹⁷ a natural zinc finger whose structure was determined under similar NMR acquisition conditions: **BBA1** (6.9mM, 7°C, pH 4.5, 29 structures) - RMSD 1.51Å for all heavy atoms, 0.90Å for the backbone atoms of residues 1-22; Xfin (5.8mM, 5°C, pH 5.5, 1.5eq ZnCl₂, 37 structures) - RMSD 1.81Å for all heavy atoms, 0.81Å for backbone atoms. The RMSD values for **BBA1** are given in Table 4-2.

Residues	RMSD backbone atoms	RMSD all heavy atoms
Tyr ¹ -Ala ²²	0.9 (±0.3) Å	1.4 (±0.3) Å
All residues	1.2 (±0.5) Å	1.5 (±0.4) Å

Table 4-2 The average value of the RMSDs of all 29 NMR derived structures from the average structure. Standard deviations are in parentheses.

Stability of BBA1: CD Analysis of Chemical and Thermal Denaturation of BBA1

Chemical and thermal denaturation experiments for **BBA1** were monitored by CD spectroscopy. The secondary structure of **BBA1**, as monitored at 222 nm, is rather stable; thermal denaturation does not appear to be complete, even at 80 °C (Figure 4-13) Complete denaturation is observed only in 8M urea (Figure 4-14). These studies were performed at pH 4.5 to parallel the conditions of the NMR structure.

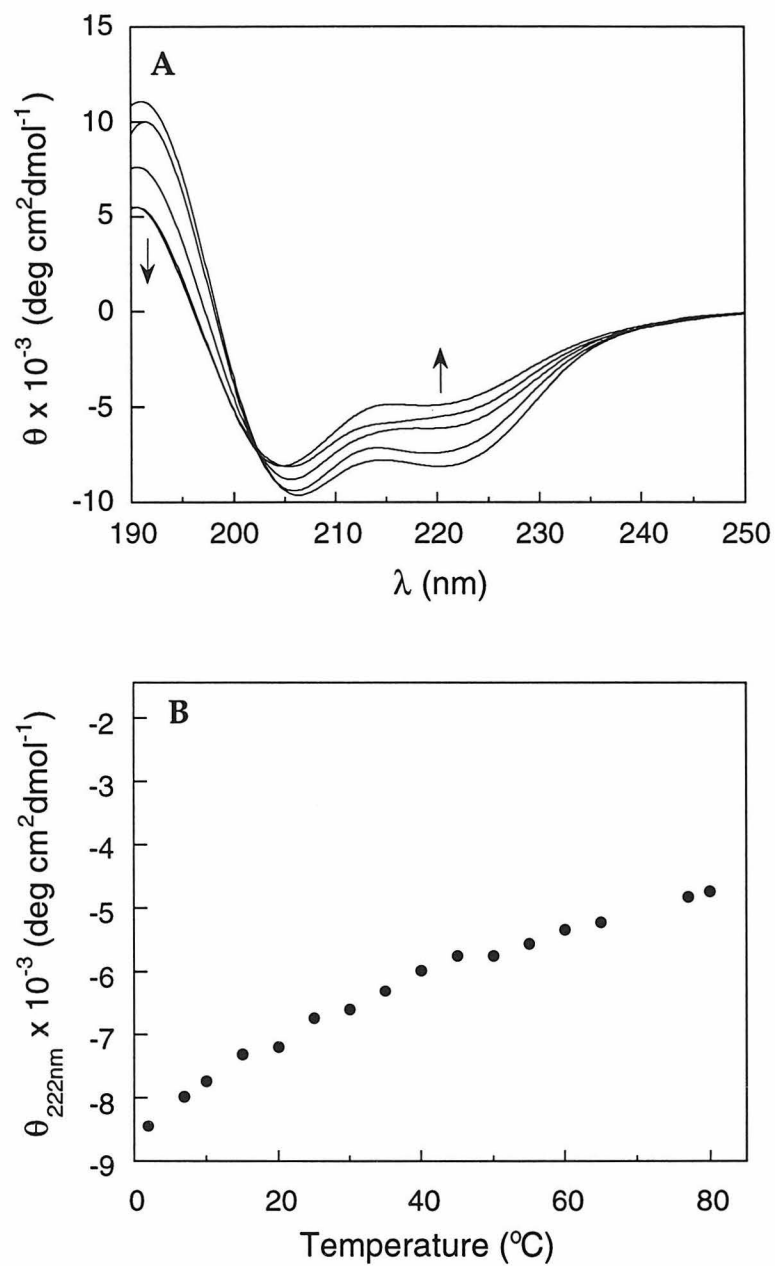


Figure 4-13 Thermal denaturation of **BBA1**. **A**. CD spectra of **BBA1** taken at 7 $^{\circ}\text{C}$, 20 $^{\circ}\text{C}$, 40 $^{\circ}\text{C}$, and 80 $^{\circ}\text{C}$ are shown. **B**. The temperature dependence of mean residue ellipticity at 222 nm. (458 μM , 10mM acetate buffer, pH 4.5, 0.01cm pathlength cell)

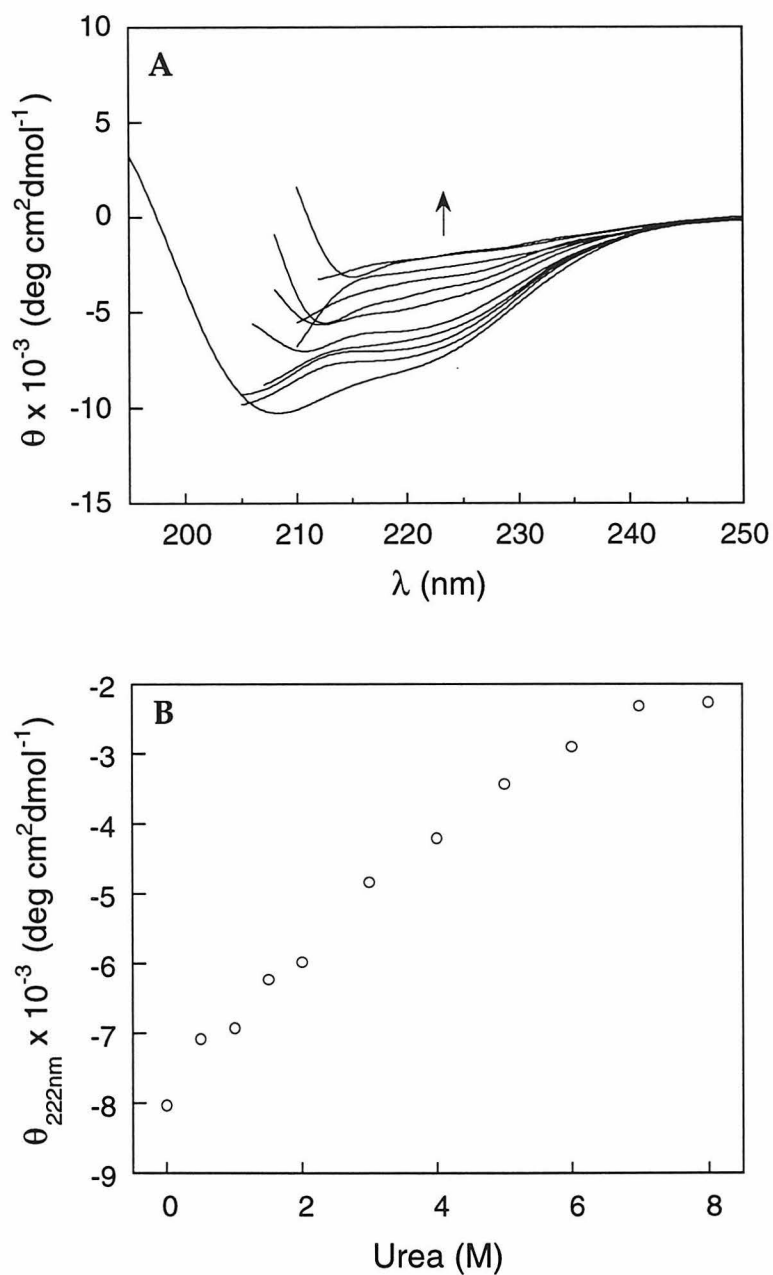


Figure 4-14 Chemical denaturation of **BBA1**. **A**. CD spectra of **BBA1** taken in 0, 0.5, 1.0, 1.5, and 2.0 - 8.0 M urea (1.0 M increments) are shown. The arrow indicates the trend with increasing urea concentration. **B**. The dependence of mean residue ellipticity at 222 nm on urea concentration. (83.9 μM **BBA1**, 1 mM acetate buffer, pH 4.5, 0.1 cm pathlength cell)

ANS Binding

The formation of a discrete hydrophobic core within **BBA1** was further investigated by analysis of 8-anilino-1-naphthalenesulfonic acid (ANS) binding. The presence of well defined tertiary structure is often probed by studying the interaction of this hydrophobic fluorescent dye with proteins.¹⁸ This technique has been used extensively to probe the tertiary structure of designed peptides for which there is no detailed structural information such as a crystal or NMR structure.^{19, 20} While ANS will bind to the apolar regions of partially folded or “molten globule” states of proteins, resulting in a large increase (50 - 100 fold) in fluorescence of the dye, native folded proteins do not bind ANS.¹⁸ Many designed proteins to date have demonstrated significant ANS binding which is attributed to poor tertiary packing interactions and the adoption of a molten globule rather than a native-like state.^{7, 19, 20} Although the well defined NMR structure of **BBA1** argues against a molten globule state for this peptide, ANS binding experiments were performed to support this conclusion. In order to conserve the peptide supplies, experiments were performed by titration of ANS into a buffered solution of peptide and compared to a control ANS titration performed in the absence of peptide. **BBA1** did not enhance the fluorescence of ANS, even at ANS concentrations in large excess of the peptide (Figure 4-15), indicating that **BBA1** does not bind ANS under conditions⁵ previously shown to establish the presence of poorly folded states in other designed proteins. In light of the NMR structure of **BBA1**, this result implies that the apolar side chains of **BBA1** are sufficiently buried in the folded structure to preclude interaction with the dye. However, there remains the possibility that this motif is too small to bind ANS, even in a molten globule-like state. Since neither unfolded nor fully folded proteins exhibit ANS binding, a molten globular $\beta\beta\alpha$ motif is needed as a positive control in order to differentiate between these two possibilities.

Unfortunately, it is extremely difficult to design a motif with molten globular properties *a priori*. ANS binding experiments performed with uncomplexed **BBA1** (pH 4.5, and pH 8.25) and the **BBA1-Zn²⁺** complex (pH 8.25) yielded similar results.

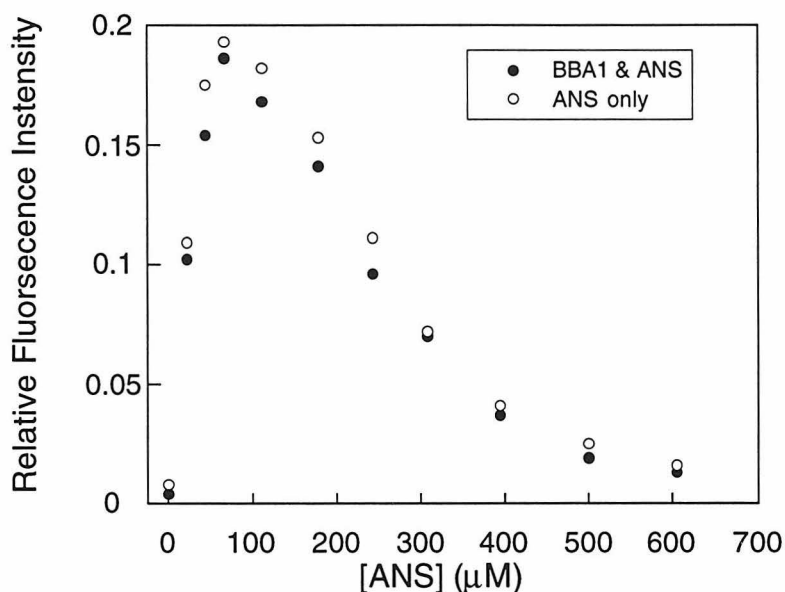


Figure 4-15 Titration of **BBA1** with ANS. (20.3μM peptide, 20mM acetate, pH 4.5). At ANS concentration greater than 100 μM, self quenching of ANS fluorescence becomes evident. Emission was monitored at 530 nm, with excitation at 370 nm.

The experiments described above clearly demonstrate the presence of a well defined secondary and tertiary structure for **BBA1** in the absence of metal ligation. One of the key ideas utilized in the design of this structure was the concept of a type II' β-turn as a structural nucleation element. In order to

establish the explicit role of the type II' β -turn in the successful generation of the metal independent structure exhibited by **BBA1**, a control peptide was synthesized and analyzed (**ZFO6**).

ZFO6: A Type II β -Turn Control

To verify the importance of the type II' β -turn in the generation of the defined secondary and tertiary structure observed for **BBA1** a peptide in which the type II' β -turn was modified to a type II β -turn was synthesized. **ZFO2** (described in the previous chapter) contained a type II β -turn and was shown by CD and NMR to lack the presence of well defined secondary or tertiary structure. However, several sequence changes, not directly involving the turn type, were made during the design process. Therefore in order to create a peptide directly analogous to **BBA1** but differing only in turn type, an additional peptide was synthesized. Peptide **ZFO6** has an identical primary sequence to **BBA1**, except the chirality of the central turn residues (DPro⁴-Ser⁵ $\rightarrow$ Pro⁴-DSer⁵) were altered to produce a type II rather than a type II' β -turn. The sequences of **BBA1** and **ZFO6** are shown below; the central two residues of the turn are in bold.

BBA1: Ac-Y-T-V-**DP-S**-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

ZFO6: Ac-Y-T-V-**P-DS**-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

Structural Analysis

The structure of peptide **ZFO6** was evaluated in the absence of coordinating metal ions. NMR analysis revealed that while the helical structure was largely maintained, both the β -hairpin and the tertiary structure observed in **BBA1** were absent. Although weaker than those observed for **BBA1**, several

$d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$ NOEs, diagnostic of helical conformation, were present for residues 12-21 of **ZFO6** (Figure 4-16). In contrast, the NOEs that would be indicative of sheet formation were not observed.

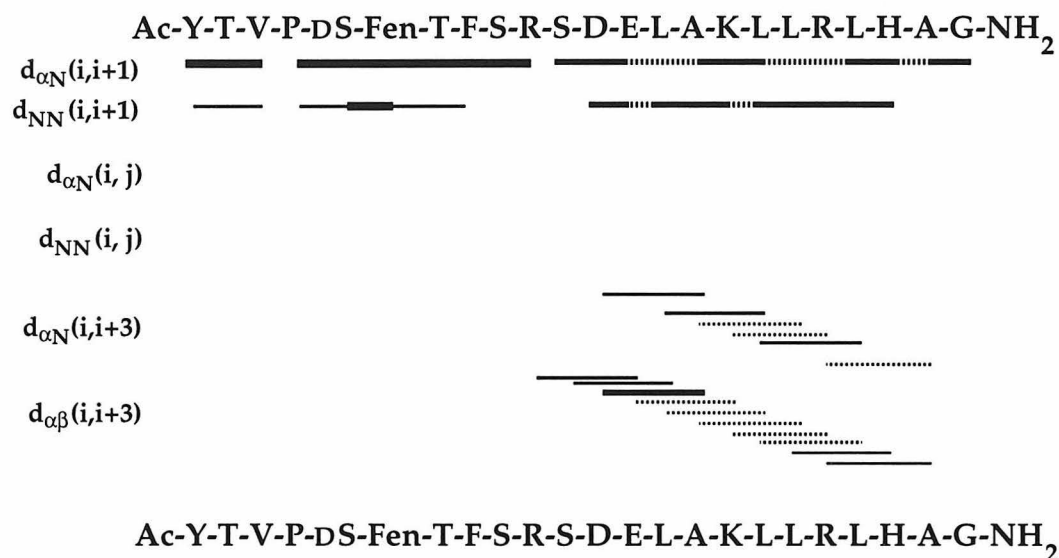


Figure 4-16 NOEs indicative of secondary structure in **ZFO6**. The width of the lines represent the relative strength of the NOEs, a dashed line indicates that the NOE could not be identified due to overlap. NOE strengths were determined from inspection of the NOESY spectrum (200 ms mixing time).

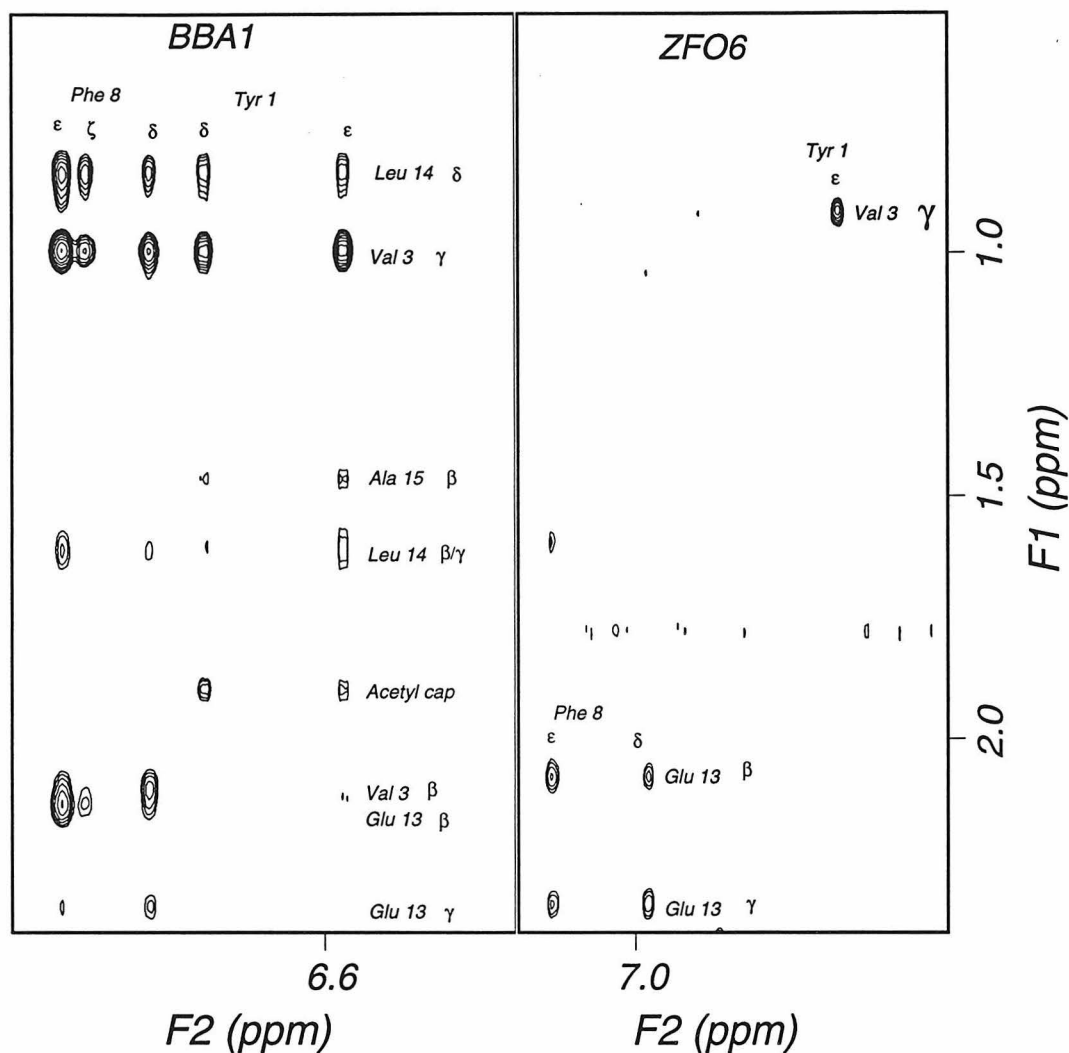


Figure 4-17 Similar regions of the NOESY spectra of **BBA1** and **ZFO6**.

In addition, none of the NOEs diagnostic of tertiary structure were visible. Comparison of the identical regions of NOESY spectra (300 ms mixing time, D₂O) of **ZFO6** and **BBA1** underscores the contrasting structural behavior of the two peptides (Figure 4-17). In **BBA1**, this region contains a number of long range NOEs between the aromatic side chain protons of Phe⁸ and Tyr¹ and residues

(Leu¹⁴, Ala¹⁵, and Glu¹³) in the helix. These NOEs are absent from the spectrum of **ZFO6**. Thus, although **ZFO6** and **BBA1** differ only in the type of β -turn employed, the differences in secondary and tertiary structure are dramatic.

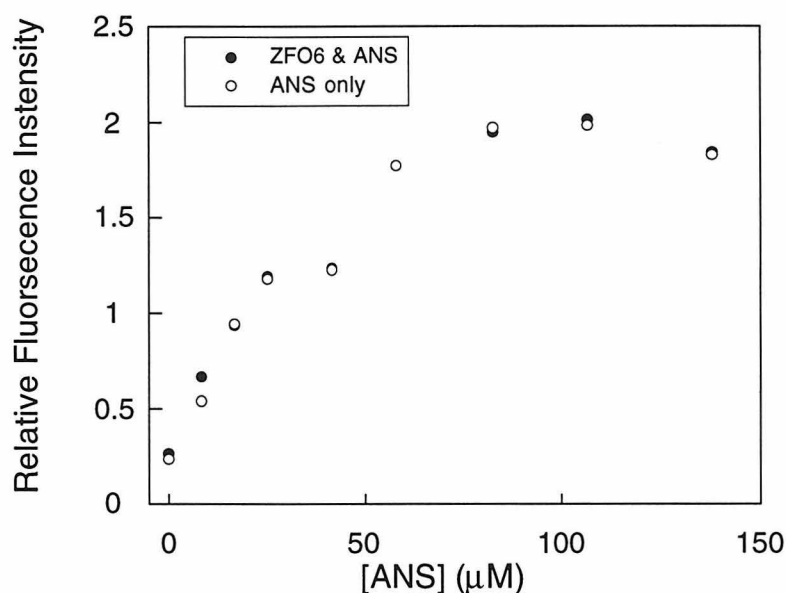


Figure 4-18 Titration of **ZFO6** with ANS. (18.4μM peptide, 20mM HEPES, pH 8.25). At ANS concentrations greater than 100 μM self quenching of ANS fluorescence becomes evident. Emission was monitored at 530 nm, with excitation at 370 nm.

ANS Binding

The structure of **ZFO6** was also evaluated by ANS binding to probe whether this motif exhibited the properties of a molten globule, and could therefore possibly serve as the positive control for the experiment performed on **BBA1**. Similar to the behavior exhibited by **BBA1**, **ZFO6** did not bind ANS even at very high concentrations (Figure 4-18), despite the NMR evidence which

indicated a lack of defined tertiary structure in the presence of some secondary structure (one of the characteristics of a molten globule). There are two possible interpretations of these results. One possibility is that the motif is too small to bind ANS in any state. Although this may be true, a very narrow set of protein structural states exhibit ANS binding.¹⁸ Investigations of ANS binding to poly-lysine under conditions which favor different elements of secondary structure indicates that the dye does not bind to isolated α -helical structure but does have some affinity for β -like secondary structure.¹⁸ For globular proteins, ANS binding results from the association of the dye with exposed patches of hydrophobic amino acids of protein structures with poorly *defined* tertiary interactions. Tertiary interactions are not absent from these states; they are simply non specific and fluctational, which results in the hydrophobic association of the dye with exposed regions of the protein. Thus, lack of ANS binding by **BBA1** and **ZFO6** may result from two very different states. **ZFO6** appears to be largely unfolded lacking *any* tertiary interactions, while **BBA1** is entirely folded with well-defined tertiary interactions. It is possible that ANS binding would only be a useful tool for the analysis of protein design efforts if an earlier design exhibited ANS binding behavior which was eliminated by modifications in later designs.

BBA2: The Structural Role of the Fen Amino Acid

The unnatural metal binding amino acid, Fen, was used to great advantage as a reporter group during the design process, permitting the direct detection of metal coordination affinities and stoichiometries for intermediates in the design pathway. Although this amino acid was extremely useful during the early design stages, the evolution of the metal independent motif represented by **BBA1** obviates the need for its reporter functionality. In addition, this amino

acid is difficult to couple (requiring off line coupling techniques, see Chapter 3) and preparation of the residue requires a multi-step synthesis. Therefore the next logical step is to create a more versatile structural template by the elimination of the residue from the sequence of the motif. While the NMR structure of **BBA1** does not implicate a direct structural role for this amino acid, one cannot rule out more subtle effects on structure, especially for such a bulky residue. Therefore, **BBA2** which included a Fen⁶→Tyr substitution was synthesized and analyzed.

Design and Modeling

In order to determine the best natural amino acid to substitute for the Fen residue, modeling based on the NMR structure of **BBA1** was performed. Of the 29 structures generated the system with a low NMR-derived energy and the lowest backbone RMSD from the average structure was chosen to serve as the basis of modeling efforts. Analysis of the NOEs observed for **BBA1** involving the Fen residue reveals that the majority of the interactions of the Fen side chain with other residues involves either the β protons or the protons on the proximal ring of the phenanthrolyl side chain (Figure 4-19). Therefore phenylalanine and tyrosine were investigated in this position, by direct substitution, using similar side chain dihedral angles. While both substitutions could fit the first ring of the NMR structure well, replacement with tyrosine was chosen for the new design **BBA2**. Based on the conformation of the Fen side chain, the phenolic hydroxyl of the side chain of a tyrosine residue at this position would point out of the motif towards solvent (Figure 4-20). Additionally it was anticipated that tyrosine at position 6 might lead to less degeneracy in NMR spectra; replacement of Fen⁶ with phenylalanine would place two phenylalanines in close proximity. Accordingly, **BBA2** was synthesized.

BBA1: Ac-Y-T-V-DP-S-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

BBA2: Ac-Y-T-V-DP-S-Tyr-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

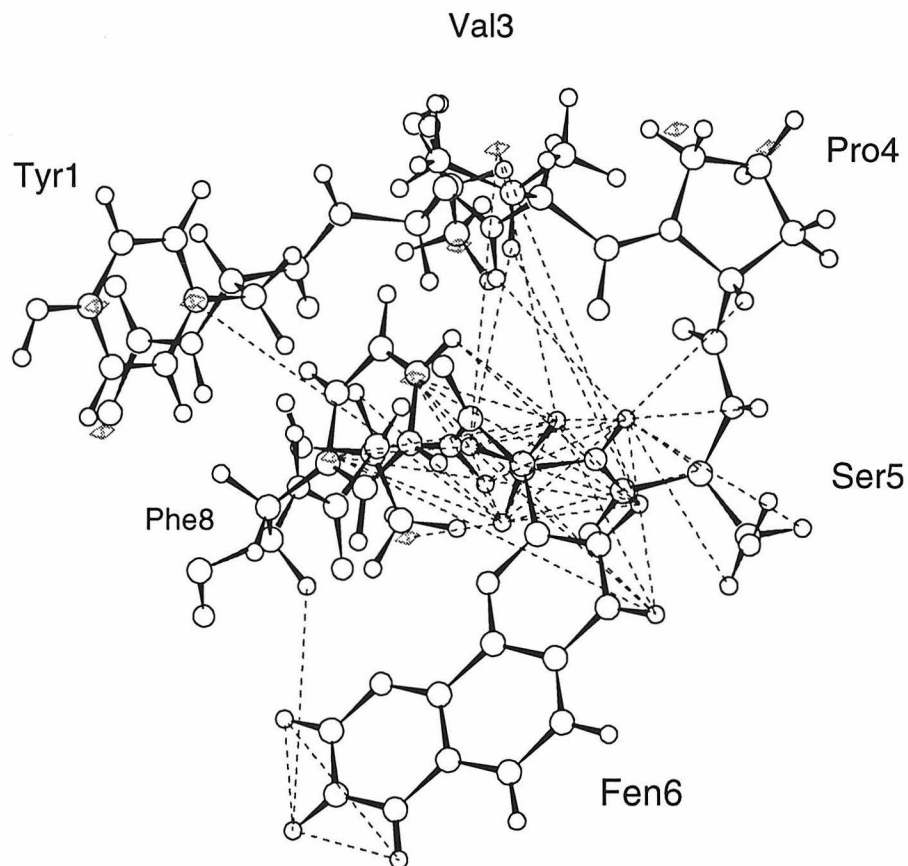


Figure 4-19 The position of the Fen residue within the NMR structure of **BBA1**. The NOE derived distance restraints (only for the Fen residue) used for the structure calculation are shown as dashed lines, pseudo atoms are shown as light diamonds. The helix has been deleted for clarity.

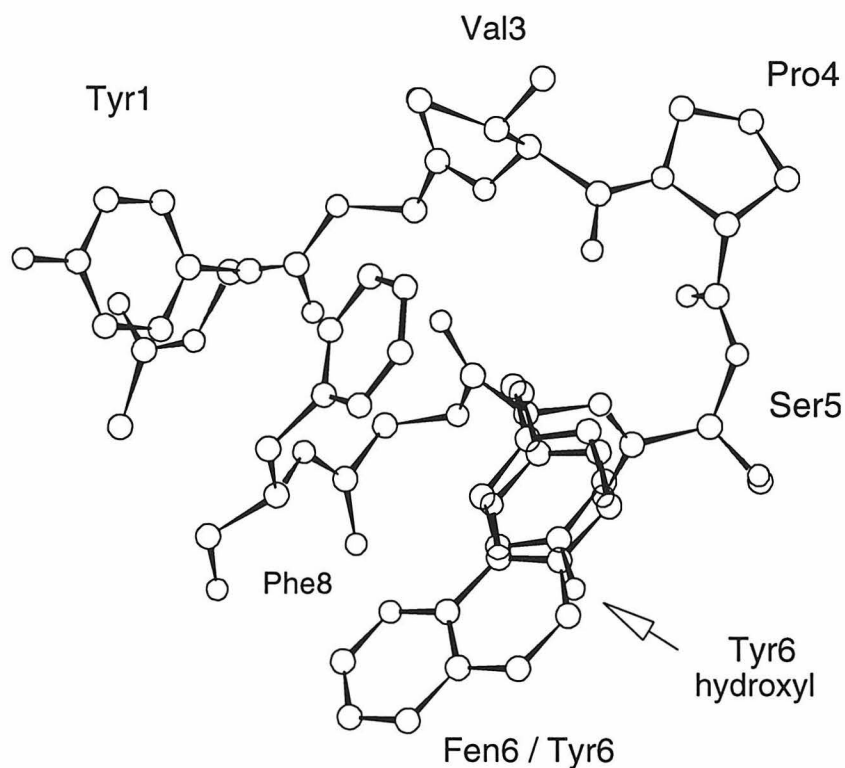


Figure 4-20 A portion of the model for **BBA2** superimposed on the NMR structure of **BBA1**. The helix has been deleted for clarity.

NMR Analysis

NMR analysis of **BBA2** revealed several interesting structural features which apparently result from the Fen⁶→Tyr replacement. Initial NMR analysis (TOCSY, ROESY, NOESY) was performed at 7 °C. Unfortunately, assignment at this temperature proved difficult. In addition to the *cis* proline conformer which was also present in **BBA1**, several spin systems with exceedingly broad proton resonances were observed. Cross-peaks involving the amide proton of these residues were sufficiently diffuse at 7 °C that identification of these spin systems required inspection of the spectra at very low contour (intensity) levels. Cross-peaks arising from the *cis* proline amide isomer were also visible at this lower level, (Figure 4-20) and complicated spectral assignment. A small impurity that

was not removed in the HPLC purification was also visible at this low contour level. CZE analysis indicated that this impurity was present at less than 5% the total concentration and therefore, this minor contaminant did not hinder the assignment of most spin systems. Unfortunately identification of the spin systems which exhibited the weak and broadened resonances described above was more difficult. Significantly less broadening was observed in spectra acquired at 17 °C (compare Figure 4-21 and Figure 4-22) and complete sequence specific assignment could be accomplished at this temperature. Based on these assignments the three residues with unusually broad resonances especially for the amide protons were identified. Cross-peaks observed in the TOCSY spectrum involving the amide protons of Thr⁷, Phe⁸, and Arg¹⁰ (and Ser⁹ to a lesser extent) were very shallow and diffuse (Figure 4-21). Although not conclusive, this suggests that there is some conformational flexibility of the backbone in this region of the peptide chain. Broad NMR signals can result when the molecule undergoes dynamic conformational averaging on a time scale that approximates the time scale of the NMR experiment. Fast interconversion of two conformations leads to one sharp peak and slow exchange results in the observation of two distinct peaks. While intermediate exchange produces a broad signal whose position is the average of the signals that would result from the separate conformations. The temperature dependence of these phenomena in the NMR spectra of **BBA2** supports this explanation. Higher temperatures could either favor a particular conformation or speed up the interconversion of the conformers. Interestingly, these residues are all relatively close to the original position of the Fen amino acid (position 6) implying that the Fen residue may have had a subtle effect on structure and imparted a conformational rigidity to this region of the peptide that was not duplicated by the Tyr substitution .

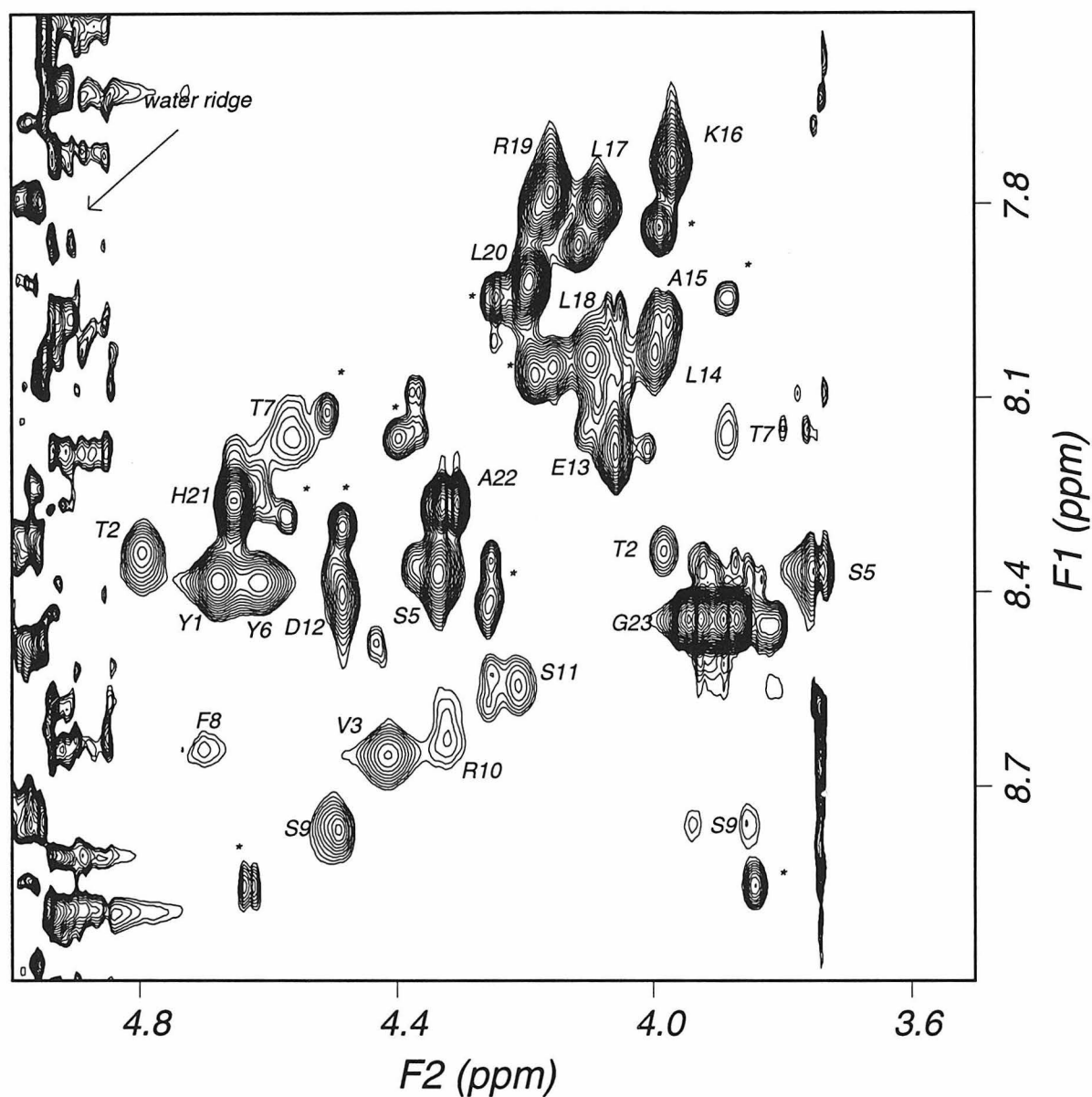


Figure 4-21 The amide-alpha region of the TOCSY spectrum of **BBA2** at 7 °C. Alpha-amide and amide-beta cross-peaks for all 23 residues are labeled. A very low contour level of the plot is required for the observation of the alpha-amide cross-peaks for residues Phe⁸, Thr⁷, Ser⁹ and Arg¹⁰. At this level cross-peaks arising from the Pro⁴ *cis* amide isomer and an impurity (<5%) are visible and denoted by (*) symbols. Sequential assignments at this temperature were based on the those obtained at 17 °C.

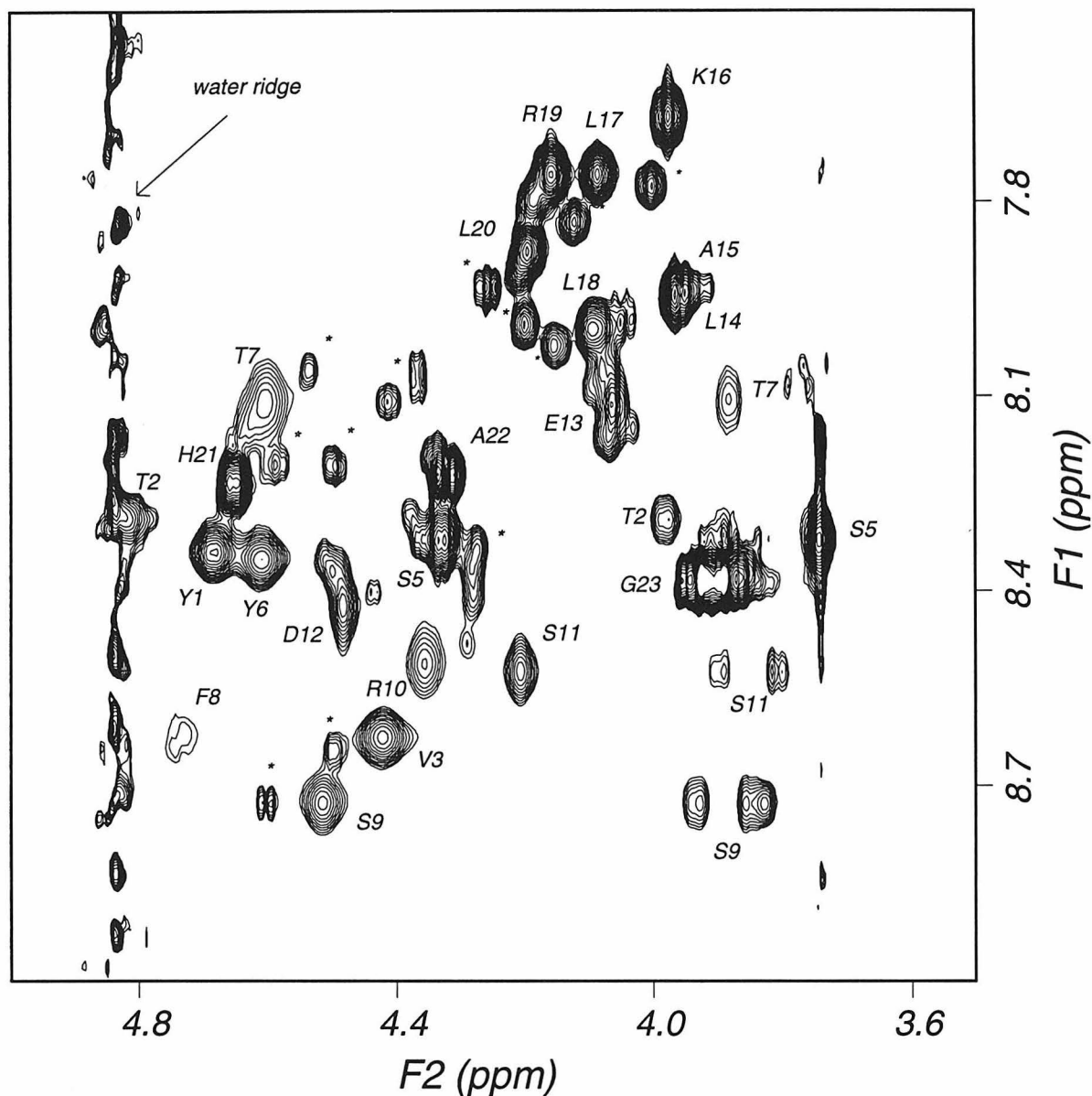


Figure 4-22 The amide-alpha region of the TOCSY spectrum of **BBA2** at 17 °C. Alpha-amide and amide-beta cross-peaks for all 23 residues are labeled. A very low contour level of the plot is required for the observation of the alpha-amide cross-peaks for residues Phe⁸, Thr⁷, and Arg¹⁰. At this level cross-peaks arising from the Pro⁴ *cis* amide isomer and an impurity (<5%) are visible and denoted by (*) symbols.

Despite the conformational flexibility implied by the broad signals observed for several residues, the complete sequence specific assignment enabled the investigation of the secondary and tertiary structure within the peptide. Similar to **BBA1**, **BBA2** exhibited a series of NOEs indicative of the desired secondary structure (Figure 4-23), although several of the important NOEs in the sheet region were obscured at 17 °C by the overlap of the Thr² alpha proton signal with the water (Figure 4-24). Additionally a similar set of long range NOEs were observed among the residues involved in the hydrophobic cluster (Tyr¹, Phe⁸, and Leu¹⁴) (Figure 4-25).

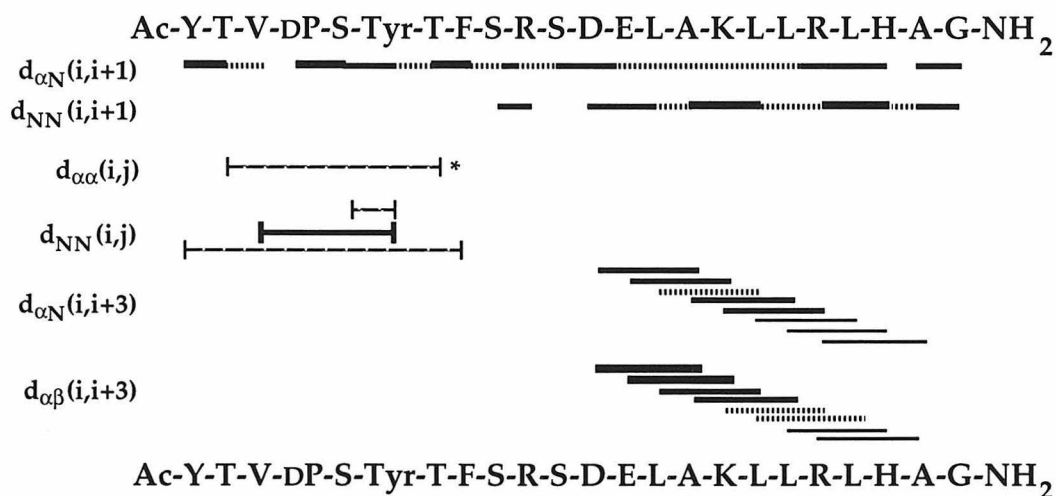


Figure 4-23 The secondary structure of **BBA2** at 17 °C. The relative strength of the NOE is indicated by the width of the line based on the intensity of cross-peaks observed in the NOESY spectrum (300 ms mixing time). A dashed line indicates that the NOE was obscured by overlap. * Observed in the 7 °C NOESY spectra taken in D₂O.

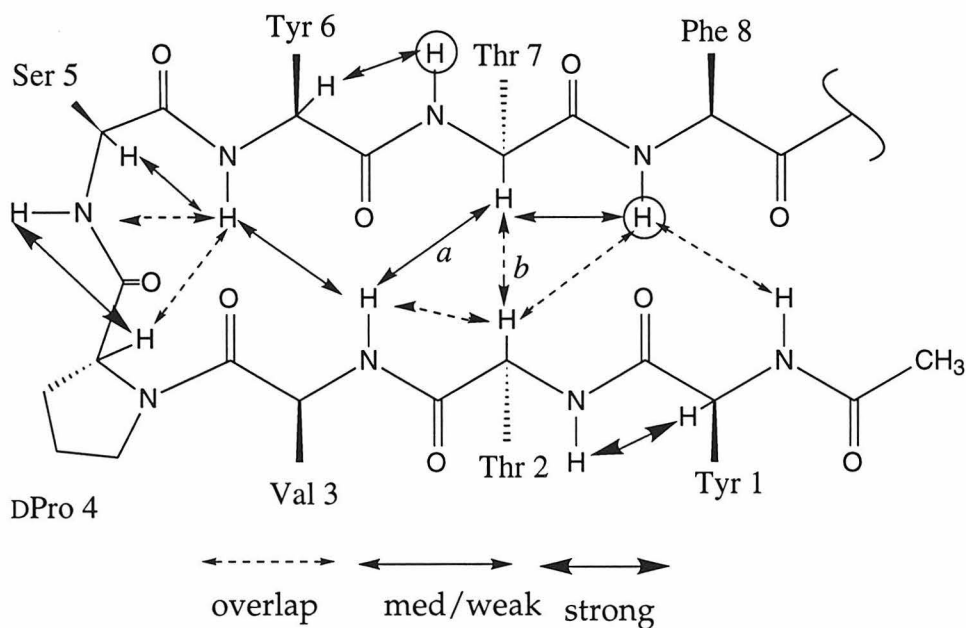


Figure 4-24 NOEs indicative of β -hairpin formation for **BBA2**. Relative intensities based on cross-peaks observed in the NOESY spectrum (300ms mixing time). *a* This NOE is very weak. *b* The $d_{\alpha\alpha}(\text{Thr}^2, \text{Thr}^7)$ NOE could only be detected in the NOESY spectrum taken in D_2O at 7 °C. The amide protons with broad NMR resonances are encircled.

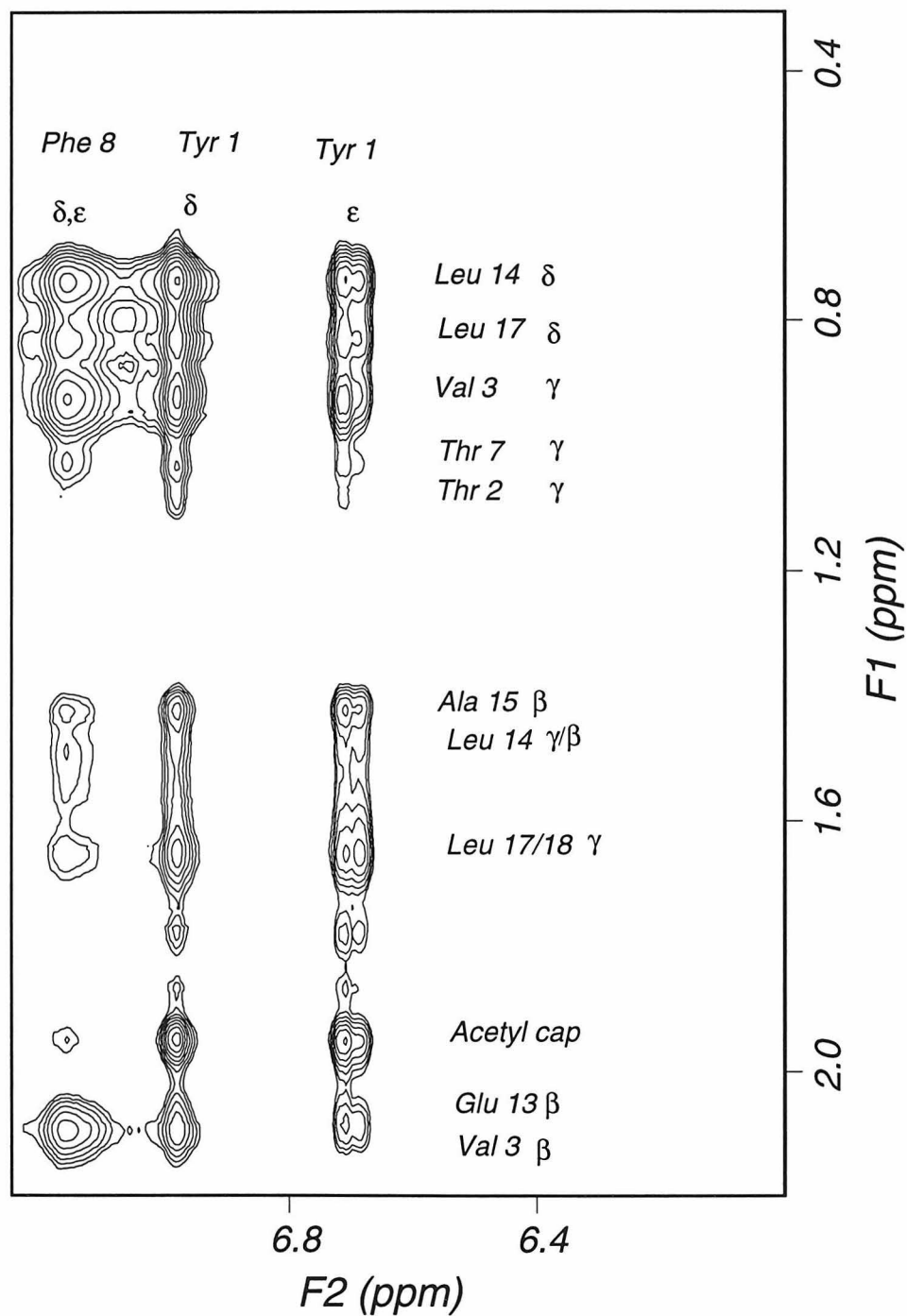


Figure 4-25 Long range NOEs for **BBA2** indicative of tertiary structure. NOESY spectrum (300 ms mixing time) was acquired at 17 °C, 600 MHz, pD (uncorrected) 4.5, D₂O.

In addition to the expected long range NOEs, several unexpected very weak long range NOEs were observed. These NOEs are between the side chain protons (β and γ) of the two threonine residues (especially from Thr⁷) in the sheet and residues in the hydrophobic cluster (both the aromatic ring protons (δ and ϵ) of Phe⁸ and Tyr¹ and the side chain protons (δ) of Leu¹⁴) (see Figure 4-25). Although these NOEs are weak, they are clearly visible and their assignments are supported both by observation in several NOESY spectra taken at different mixing times and by the presence of NOEs to multiple protons of the threonine side chains. The threonine side chains are solvent exposed in the designed motif, and on the *exterior* of the sheet. Therefore, these side chains would not be expected to show NOEs to residues found on the *interior* of the motif. Despite the presence of a number of NOEs indicative of a β -hairpin within this peptide [$d_{NN}(\text{Tyr}^6, \text{Val}^3)$ and $d_{\alpha\alpha}(\text{Thr}^2, \text{Thr}^7)$], it appears that there may be some twisting of the strands such that the side chains of the threonine residues have partial exposure to the interior of the motif. It is notable that Thr⁷ is both involved in the majority of these unexpected NOEs and exhibits broadened resonances. Thus, the cause of both the signal broadening and these unexpected NOEs might be an increased flexibility in the region at the end of second strand (residues 7-10) such that this portion of the peptide exists in a conformational flux between an extended sheet-like conformation and a more twisted geometry.

One other effect of the Fen residue in **BBA1** was observed in the comparative analysis of the NMR spectra of **BBA2**. The presence of the large extended ring system of the phenanthroline moiety in Fen, had a significant effect on the chemical shift of the neighboring residues. In general, a single amino acid substitution in a structured peptide would not be expected to have a large effect on the chemical shift of neighboring residues, unless the substitution greatly perturbed the structure. While there were definitely some structural

effects of the Fen⁶→Tyr substitution, a large number of NOEs point to the same overall structure albeit with potentially altered stability. Therefore, the likely cause of the differing chemical shifts is the unusual nature of the Fen side chain. This was particularly evident in the assignment of the aromatic amino acids (Phe⁸ and Tyr¹) and the threonine residues (Thr² and Thr⁷). The alpha resonances of these amino acids were shifted downfield (relative to **BBA1**) and closer to the water signal in the absence of the Fen amino acid. For Thr⁷, the alpha signal was almost overlapping with that of the water at 17 °C, which prevented the observation of NOEs involving this proton. Additionally, the aromatic protons of Tyr¹ in the *cis* and the *trans* DPro⁴ conformations were nearly overlapping in **BBA2**, preventing determination of the *cis:trans* Pro⁴ amide isomer ratio by simple integration of the 1-D signals as was possible for **BBA1**.

Discussion

The design of the fourth peptide in the series (the first three are described in Chapter 3) was based on an analysis of *cis-trans* proline amide isomerism in a series of tetrapeptides incorporating the type II' β -turn. The optimal sequence, -Val-DPro-Ser-Fen-, was introduced into the turn region of peptide **ZFO4** (residues 3-6). Replacement of the residue immediately preceding proline successfully limited the population of the *cis* proline amide isomer to approximately 15%. Although NMR analysis revealed the presence of defined secondary and tertiary structure in this peptide, overlap in the resonances for the ring protons of Phe¹ and Phe⁸ hindered complete structure generation based on the NMR data. A Phe¹→Tyr substitution in peptide **BBA1** resolved this spectral ambiguity. Detailed NMR analysis, including the determination of the solution state structure by NOE restrained simulated annealing (SA) confirmed that this peptide exhibited the desired three-dimensional architecture (Figure 4-26).

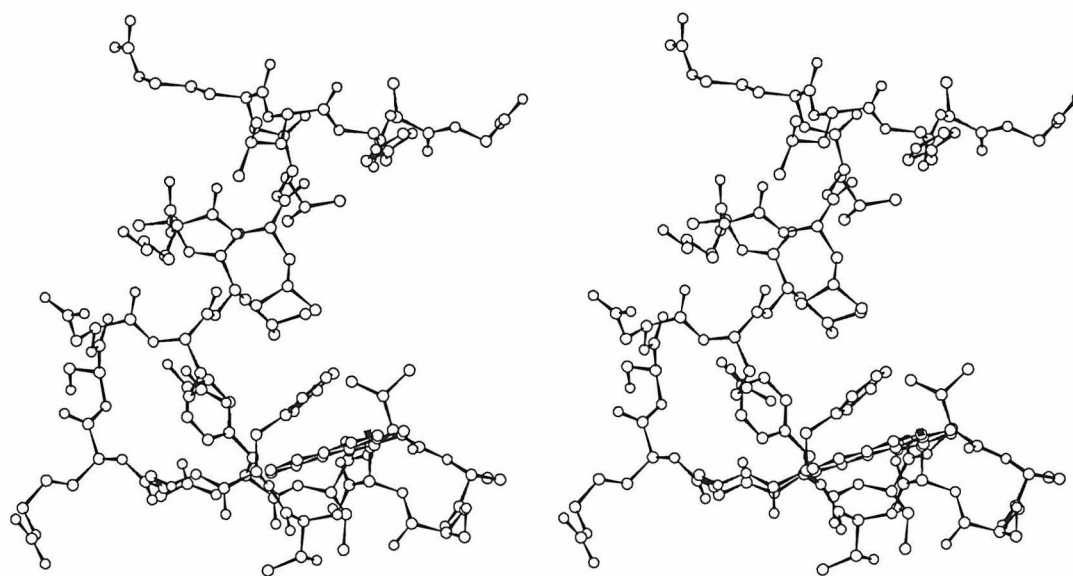


Figure 4-26 Stereo image of a representative NMR structure for **BBA1**.

Analysis of BBA1 Structure

NMR and CD analysis demonstrated that **BBA1** adopts a well-defined, monomeric tertiary structure in aqueous solution. The 29 SA NMR structures generated for this peptide confirm that the α -helix and β -hairpin regions of the motif are in agreement with the intended design. Analysis of this ensemble revealed that 26 out of 29 structures (>90%) exhibit characteristic type II' β -turn dihedral angles for DPro⁴-Ser⁵ although significant deviation from the ideal value was observed for the ψ value of Ser⁵. The remaining three structures differed in the orientation of the central amide bond and did not conform to a defined turn type. Analysis of the hydrogen bonding interactions within the 29

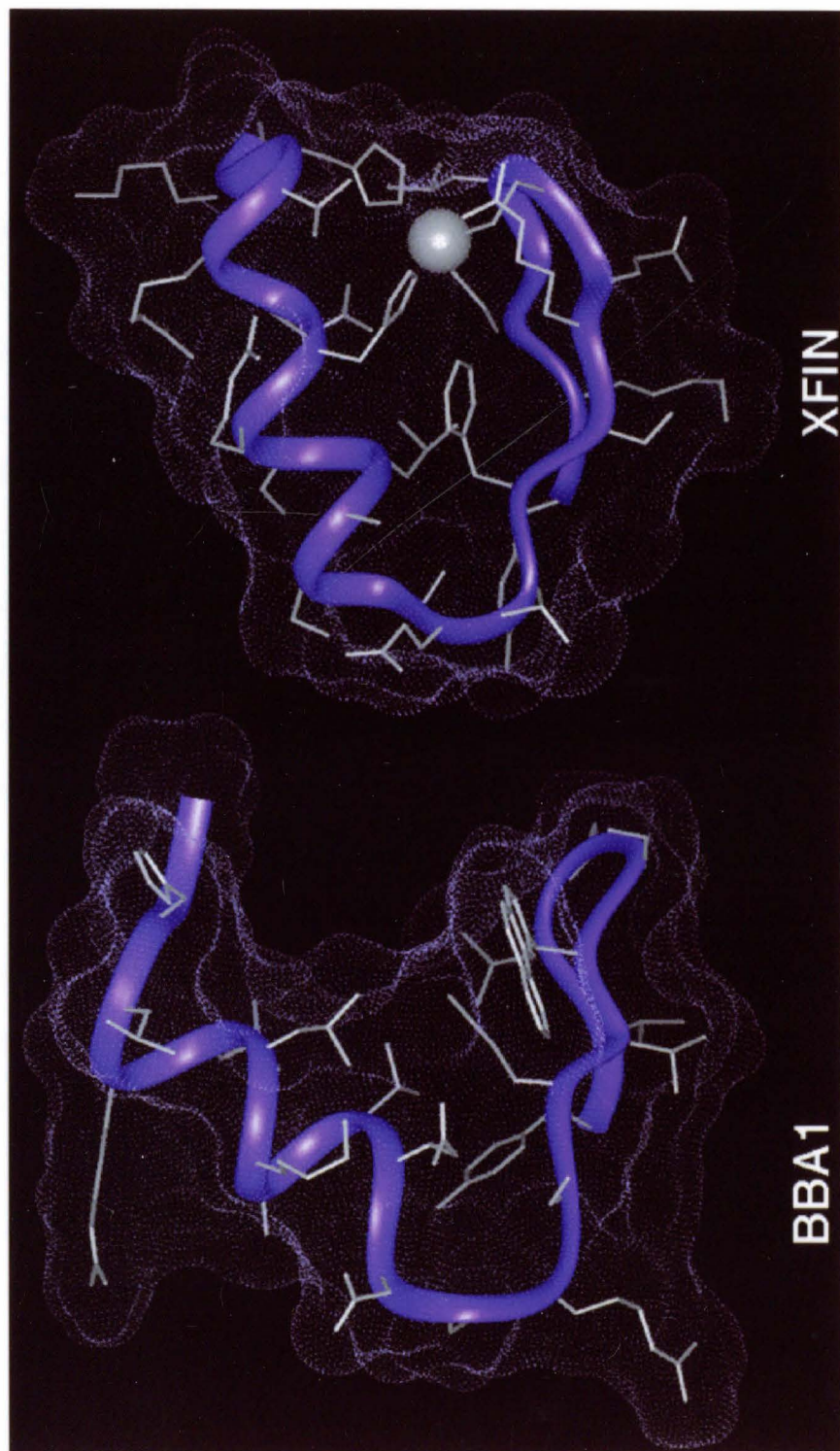
NMR structures was also performed. Although 28 of the 29 structures had at least one of the hydrogen bonds within the turn region [3(CO)→6(NH) or 6(CO)→3(NH)] very few had the remaining two hydrogen bonds which would be expected in an extended β -sheet. This is not very surprising given the terminal location of Tyr¹. Presence of the $d_{\alpha\alpha}$ (Thr², Thr²) and extended dihedral angles for these residues indicated that a conformation similar to the intended β -sheet is formed between residues 1 and 8.

Similar analyses were performed for the residues within the helix. Residues 13 through 20 adopt typical α -helix backbone dihedrals in most structures. Residues Asp¹² and His²¹ appear to serve as “helix cap” residues.²¹ The dihedral angles at these positions deviate significantly from ideal α -helix values ($\phi=-60^\circ, \psi=-40^\circ$) although these residues do participate in the expected 12(CO)→16(NH) and 17(CO)→21(NH) hydrogen bonds in most of the structures. In the same family of structures, it is apparent that the C-terminal glycine is considerably disordered, as are the side chains of surface-exposed residues such as Arg¹⁰, Lys¹⁶, and Arg¹⁹. While the latter is not surprising, the flexibility of the C-terminus could be a target in future designs.

In natural zinc finger peptides, the conserved hydrophobic core residues are unable to drive folding without the assistance of metal coordination. In **BBA1**, the formation of the hydrophobic core has been assisted by the inclusion of a heterochiral β -type II' turn as a structural nucleation element, together with a judicious selection of residues to enhance the appropriate secondary structure throughout the peptide. Although the NMR derived structure of **BBA1** resembles that of the natural zinc finger domains,^{17, 22} the designed motif is more “open” than its native counterpart. Figure 4-26 compares the structure of **BBA1** with that of Xfin,¹⁷ a natural zinc finger the NMR structure of which has been solved under similar conditions (except for the presence of zinc). In the natural

zinc finger peptides, a metal binding site cross-links the end of the α -helix to the β -hairpin, and it is this interaction that defines the angle between the helix and the sheet. It is therefore not surprising that the removal of such a cross-link would result in a more open structure.

Figure 4-27 (following page) The solvent accessible surface of representative NMR structures of **BBA1** and Xfin.¹⁷



Chemical and thermal denaturation experiments monitored by CD demonstrate that **BBA1** has considerable stability. The motif does not completely unfold even at temperatures as high as 80 °C and requires 8 M urea for complete denaturation. Considerable thermal stability has been observed with a metal-complexed natural zinc finger.²³ Although cooperative unfolding is often viewed as one of the hallmarks of native protein structure, non-cooperative chemical denaturation has been observed for natural zinc fingers despite the presence of a well-defined tertiary structure.²⁴ Additionally, the unfolding studies employ CD spectroscopy which primarily measures the secondary, and not the tertiary structure content of the peptide. Therefore, a peptide such as **BBA1**, which has been designed to possess a high degree of inherent secondary structure, might not be expected to manifest cooperative unfolding of secondary structure as monitored by this method. This is especially relevant to our system since the helix of **BBA1** is inherently stable in solution (as indicated by the NMR characterization of the control peptide **ZFO6**), and the CD signal is dominated by the helical signal.

Two other techniques are commonly used to analyze the structure of protein design efforts; ANS binding and protection of amide protons from exchange with solvent. **BBA1** was analyzed by both these methods. Binding of the hydrophobic dye ANS, by the molten globule state of proteins results in a large increase (on the order of 50 - 100 fold) in fluorescence intensity.¹⁸ Unlike many design efforts to date, **BBA1** did not bind ANS even at large excesses of dye. Although this may be due to the native-like state of **BBA1**, this motif may also be too small to bind the large dye molecule. The similar behavior of **ZFO6** points to some of the uncertainties that are inherent in the analysis of protein structure by ANS binding. H/D exchange experiments revealed the presence of several slowly exchanging protons whose signals persist for over one hour (pH

4.5, 7° C). While the rates of amide exchange for **BBA1** are somewhat more rapid than that found in large globular proteins, they are comparable to rates for peptides derived from natural zinc fingers^{25, 26} and are considerably slower than those expected for random coil peptides.¹⁵

Role of the Type II' β -Turn

One of the key concepts presented in the design of **BBA1** was the use of a reverse turn as a structural nucleation element. Two different β -turn types (II and II') were investigated for their ability to promote β -hairpin formation within the motif (See chapter 3). The structural integrity of **BBA1** was attributed to the use of a type II' β -turn. The explicit role of the turn in the final structure was confirmed by the analysis of the peptide **ZFO6**, in which the type II' turn of **BBA1** was replaced with a type II β -turn. This peptide exhibits none of the NOEs indicative of β -hairpin formation which are present in **BBA1** (compare Figure 4-7 and 4-16). In addition, the long range NOEs indicative of tertiary structure were absent (Figure 4-17). While both type II and type II' β -turns exhibit inherent stability in aqueous solution as evidenced by spectroscopic studies of short model peptides,¹³ the type II' β -turn clearly has a greater ability to promote formation of secondary and consequently, tertiary structure in this system.

The β -turns have been extensively studied because of their structural and functional roles in peptides and proteins. These structures have been shown to be critical in the bioactivity of small naturally occurring peptides and have been implicated in the process of protein folding.²⁷ This study highlights the potential use of these small motifs as structural nucleation elements in protein design.

Role of the Fen Amino Acid

With the demonstration of the well defined tertiary structure of **BBA1** and the determination of its solution structure by NMR, the stage is set for the exploitation of this motif in the design of catalytic polypeptides. In order to facilitate the design and synthesis of such constructs, the removal of the Fen amino acid from the sequence is desirable. This amino acid is obtained through a multi-step synthetic protocol, and is therefore expensive both in terms of cost and effort. Additionally, the poor solubility and coupling efficiencies of this amino acid require the use of special off-line coupling procedures for successful peptide incorporation. Therefore, the explicit role of the Fen amino acid in the structure of **BBA1** was investigated by the design and synthesis of a Fen⁶→Tyr substitution (**BBA2**). While this peptide did exhibit a number of NOEs indicative of both secondary and tertiary structure, there were several undesirable features associated with **BBA2** that warrant further investigation. The presence of several spin systems with very broad resonances, and a number of unexpected weak long range NOEs point to the potential that a region of this peptide (Thr⁷-Arg¹⁰) exists in a conformational flux. The dynamics of this peptide may result from the twisting of the second β -strand. In fact, the structure of **BBA2** could only be evaluated at 17 °C, due to the severe broadening of several amides at the lower temperature (7 °C) employed for the analysis of previous designs. Despite the fact that the NMR structure of **BBA1** indicated no direct structural role for the Fen amino acid, this residue does appear to play a more subtle role by influencing the conformational flexibility of the adjacent residues. In addition to these structural effects, analysis of **BBA2** revealed that the Fen amino acid indirectly facilitated solution structure determination by NMR. The extended aromatic ring system had a favorable effect on the chemical shift of the surrounding residues of **BBA1**, serving to differentiate similar residues (such as

Thr² and Thr⁷) and prevent the loss of information by shifting the C_α proton resonances upfield of the water peak. Furthermore, these effects were different in the *cis* and *trans* Pro⁴ amide isomers which allowed for the facile quantification of the *cis/trans* amide isomer ratio in **BBA1**.

Despite the ambiguities concerning the structure of the β-hairpin, tertiary interactions were, encouragingly, still maintained in **BBA2**, which points out the potential utility of this motif in future design efforts. Several strategies could be employed to reduce the conformational flexibility of the hairpin region. Incorporation of a bulkier natural amino acid such as tryptophan, could supply some of the conformational rigidity originally supplied by the Fen amino acid in **BBA1**. However, this would limit the positions available for modification in future designs. Alternatively, the threonine residues on the exterior of the sheet could be replaced with charged amino acids which might create an electrostatic stabilization of the desired, planar β-hairpin structure. In addition to creating a stabilizing effect in the desired conformation, use of a charged residue at these positions (2 and 7) would also disfavor an undesirable twisted conformation which places these residues near the hydrophobic core residues. Several unnatural amino acids have been investigated within the Imperiali Group specifically for this purpose. (S,S)-α,β-Diaminobutyric acid (*allo*-β-methyl-β-amino alanine, or *aMaa*) and α,β-diaminopropanoic acid (β-amino alanine or *Baa*) have been synthesized and incorporated into peptides *via* Fmoc based solid phase peptide synthesis (Appendix A). The side chains of these amino acids are significantly shorter than those of the naturally available lysine or arginine amino acids and therefore more suited to form salt bridge interactions across a β-sheet. Analysis of several peptides incorporating these residues demonstrated their potential utility in stabilizing sheet structure when paired across a sheet with an aspartic acid (see appendix A for details). In addition to being of the appropriate

length, *a*Maa is also β -branched which would favor an extended sheet-like conformation.

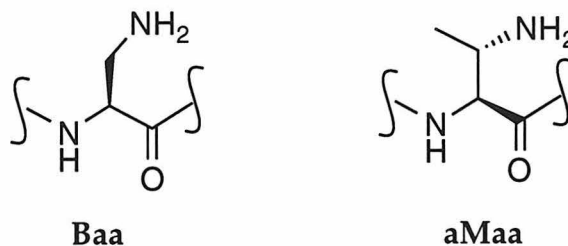


Figure 4-28 Charged amino acids which could be employed for the electrostatic stabilization of β -sheet structure.

Conclusions

The structure of small naturally occurring protein domains generally relies on either disulfide bridges or metal binding sites for stability. Our goal has been the construction of stable polypeptide motifs, based on natural systems, which are designed to fold in the absence of these cross-links. In our approach the utilization of appropriate structural nucleation elements compensates for the stability originally afforded by the cross-links. The success of this approach has been illustrated by the development of a metal independent $\beta\beta\alpha$ motif, **BBA1**. While the $\beta\beta\alpha$ topology of **BBA1** is reminiscent of the zinc finger domains, elimination of the zinc binding site results in a more open structure. The absence of a metal-mediated cross-link creates an accessible cleft between the helix and the sheet, which may be exploited as a small molecule substrate binding site in future designs. The template resulting from these efforts therefore provides a well-defined structural scaffold for the development of functional polypeptides. To increase the versatility of this scaffold, the peptide **BBA2** was synthesized in

which the Fen amino acid was replaced with a naturally occurring amino acid, tyrosine. Although this change appeared to have some effect on the conformational flexibility of the motif, tertiary structure was largely maintained and the detailed NMR analysis suggested some future avenues for improvement.

The **BBA1** motif represents one half of the solution to the design of polypeptides with novel functions: A well defined structural template. The development within the Imperiali group of novel functional residues, such as the coenzyme-amino acid chimeras²⁸⁻³¹ and the metal-binding amino acids^{32, 33} represents the other: The opportunity for the integration of function. The pieces are now in place for the construction of catalytic peptides and for the exploration of enzyme based catalysis within a peptidyl model system. The insight gained by such efforts will aid in the understanding of the elements of protein catalysis so skillfully crafted in natural systems.

Acknowledgments

This project was done in collaboration with Richard Cheng. Richard assisted in the synthesis of peptides **ZFO4**, **BBA1**, and **ZFO6**. He also performed some of the Zn^{2+} titrations of **BBA1** monitored by CD and UV.

A portion of this work described in this chapter was presented in the following publications:

Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Economy in Protein Design: Evolution of a Metal-Independent $\beta\beta\alpha$ Motif Based on the Zinc Finger Domains," *J. Am. Chem. Soc.* **1996**, *118* (13), 3073 - 3081

Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Design of Monomeric 23-Residue Polypeptide with Defined Tertiary Structure," *Science* **1996**, *271*, 342 - 345

Experimental

Peptide Synthesis

Peptides were synthesized essentially according to the protocol reported in Chapter 3. All peptides were purified by HPLC and analyzed by mass spec. All HPLC purifications were performed with H_2O :acetonitrile solvent mixtures with 0.1% TFA added as a carrier on C_{18} reverse phase columns. Linear gradients (LG) are given in the percentage of acetonitrile. Mass spectrometry was performed using either electrospray ionization (ESI) or matrix assisted laser desorption ionization (MALDI) techniques. Analytical equilibrium centrifugation experiments were performed on **ZFO4** by AMGEN (Thousand Oaks, CA) from a stock solution (48 μM) in 90:10 H_2O : D_2O at pH 4.5. Several concentrations were run with monitoring at 280 nm. All solutions could be reasonably described as a single monomeric molecular system, or as a monomer-dimer equilibrium, in which the dimeric form was not concentration dependent.

In the latter case the detected dimer was very small ranging from 0-8%, indicating that the peptide existed as a predominately monomeric species in solution. It should be noted that these experiments were performed using an extinction coefficient for the Fen residue within a peptide that was determined by comparing the analytical weight of a Fen containing peptide to the absorbance spectra. This analysis yielded a value for the extinction coefficient of $13,950 \text{ M}^{-1}\text{cm}^{-1}$ (268 nm). Subsequently the extinction coefficient was more accurately determined ($20,890 \text{ M}^{-1}\text{cm}^{-1}$ at 268 nm) from absorbance spectra of a Fen-containing peptide, the concentration of which was determined by quantitative amino acid analysis.

ZFO4: Ac-F-T-V-DP-S-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

Analytical HPLC: t_R = 27.7 min; LG 0-70% over 25 min.

Purification : t_R = 14.94 min: Semiprep column LG 36-41% over 15 min.

MW (ESI) calc. 2736; observed. 2735.9

BBA1 (ZFO5): Ac-Y-T-V-DP-S-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

Analytical HPLC: t_R = 26.3 min; LG 0-70% over 25 min.

Purification : t_R = 15.40 min: Semiprep column LG 33-38% over 15 min.

MW (MALDI) calc. 2753.1; observed. 2752.8

ZFO6: Ac-Y-T-V-P-DS-Fen-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

Analytical HPLC: t_R = 28.4 min; LG 0-70% over 25 min.

Purification : t_R = 11.9 min: Semiprep column LG 33-38% over 15 min.

MW (MALDI) calc. 2753.1; observed. 2752.8

BBA2: Ac-Y-T-V-DP-S-Y-T-F-S-R-S-D-E-L-A-K-L-L-R-L-H-A-G-NH₂

Analytical HPLC: t_R = 28.5 min; LG 0-70% over 25 min.

Purification : t_R = 24 min: Preparative column LG 25-50% over 35 min.

MW (ESI) calc. 2666.0; observed. 2666.1

CD and UV Analyses

CD and UV analyses were performed as reported in Chapter 3. Zinc titrations monitored by CD for **ZFO4** (9.9 μ M, 13.5 μ M) in a 1.0 cm pathlength cell 10 mM Tris, pH 4.5. Similar titrations were performed for **BBA1** (13-170 μ M) in either a 0.1 cm or a 1.0 cm pathlength cell. UV titration of **ZFO4** was performed in 50 mM HEPES, 9.2 μ M peptide (**ZFO4**) and 1-20 μ M (**BBA1**) and analyzed as reported in Chapter 3.

NMR Acquisition and Analysis

Two-dimensional ¹H NMR spectra of **ZFO4**, **BBA1**, **ZFO6**, and **BBA2** were obtained in both 90:10 H₂O:D₂O (6.2 mM **ZFO6** and 6.9 mM **BBA1**) and 99.98% D₂O (5.9 mM **ZFO6** and 5.2 mM **BBA1**) at 7 °C on a Varian 600 MHz instrument. Additional spectra were recorded for **BBA2** at 17 °C. All spectra were collected at pH 4.5, or pD* 4.5 (uncorrected meter reading). Complete sequence specific assignments (see Tables 4-3, 4-4, 4-5, 4-6) for backbone and side chain proton were determined using TOCSY (mixing time 15 ms and 70 ms),³⁴ double-quantum filtered correlation spectroscopy (DQF-COSY),³⁵ nuclear Overhauser spectroscopy (NOESY).^{12, 36} NOESY spectra were obtained for **BBA1** (7 °C) and **BBA2** (17 °C) in H₂O (50, 100, 200, and 300 ms mixing times) and in D₂O (100, 200, and 300 ms mixing times) and for **ZFO6** in H₂O (100, 200, and 300 ms mixing times) and in D₂O (300 ms mixing time). Suppression of the water signal was accomplished by presaturation during the relaxation delay

(1.5 s). Amide variable temperature coefficients were obtained for **ZFO4**, by the assignment of the amide chemical shift from examination of TOCSY experiments acquired at various temperatures (280, 290, 300, 310 K). The variation of δ (ppm) was linear with temperature for residues reported in Table 4-1 (Figure 4-29). The $^3J_{\alpha N}$ coupling constants for **BBA1** were obtained from cross peak measurements in a DQF-COSY spectra acquired at 25 °C to minimize the line broadening observed at 7 °C.¹² The exchange of amide protons for **BBA1** with D₂O was monitored by dissolving a sample of **BBA1** (4.3 mg, dried from H₂O) in freshly prepared deuterated acetate buffer (50 mM), pD 4.5 (uncorrected meter reading at room temperature) which was stored on ice. The sample was then placed in the NMR (500MHz, Bruker AMX) at 7 °C which had been shimmed on a similarly prepared sample of **BBA1** and 1D spectra were immediately acquired. A watergate pulse sequence was used to suppress the residual water signal.³⁷ This technique is not suppressive and therefore should not contribute to the diminution of the amide proton signals. Each one-dimensional spectrum consisted of 64 transients with an acquisition time of 2 minutes. Processing and analysis of all spectra were done using Felix 2.3 (**ZFO4**, **BBA1**, and **ZFO6**) or Felix 95 (**BBA2**) software (Biosym Technologies, CA).

Residue	alpha	amide	side chain
Ac	1.90		
Phe ¹	4.654	8.282	2.966, 2.727, 7.05, 7.156, 7.223, 7.117
Thr ²	4.657	8.325	1.071, 3.98
Val ³	4.428	8.579	0.985, 2.137
DPro ⁴	4.437		1.817, 2.188, 1.930, 3.729
Ser ⁵	4.384	8.424	3.723, 3.788
Fen ⁶	5.307	8.687	3.522, 3.566, 7.577, 8.125, 7.89, 8.003, 8.754, 9.069
Thr ⁷	4.49	8.437	1.023, 3.913
Phe ⁸	4.368	8.429	2.592, 2.677, 6.986, 7.10, 7.158
Ser ⁹	4.421	8.580	3.889, 3.797
Arg ¹⁰	4.263	8.577	1.903, 1.609, 1.753, 3.145, 7.42
Ser ¹¹	4.245	8.482	3.917, 3.830
Asp ¹²	4.505	8.408	2.759, 2.675
Glu ¹³	4.072	8.197	2.085, 2.34
Leu ¹⁴	4.057	8.113	1.60, 0.8115
Ala ¹⁵	4.033	8.005	1.457
Lys ¹⁶	4.003	7.790	1.881, 1.651, 1.549, 1.351, 2.918, 7.661
Leu ¹⁷	4.115	7.840	1.1458, 1.795, 1.583, 0.822, 0.873
Leu ¹⁸	4.141	8.039	1.71, 1.75, 1.55, 0.8125, 0.804
Arg ¹⁹	4.182	7.812	1.872, 1.705, 1.612, 3.191, 7.289
Leu ²⁰	4.205	7.940	1.48, 1.670, 1.596, 0.830, 0.880
His ²¹	4.649	8.27	3.316, 3.162
Ala ²²	4.317	8.29	1.418
Gly ²³	3.92	8.462	
NH ₂		7.28, 8.610	

Table 4-3 NMR assignment of **ZFO4**. 11.8 mg/700 μ l, 90:10 H₂O:D₂O, pH 4.5, 7 °C, referenced to dimethyl sulfoxide (δ 2.72), 600 MHz.

Residue	alpha	amide	side chain
Ac	1.890		
Tyr ¹	4.54	8.260	2.807, 2.621, 6.843, 6.560
Thr ²	4.593	8.230	1.055, 3.970
Val ³	4.391	8.479	0.970, 2.114
DPro ⁴	4.435		1.806, 2.167, 1.907, 3.734
Ser ⁵	4.388	8.365	3.728, 3.796
Fen ⁶	5.31	8.707	3.569, 3.494, 7.571, 8.117, 7.88, 8.10, 8.77, 9.05
Thr ⁷	4.483	8.408	1.020, 3.916
Phe ⁸	4.346	8.417	2.674, 2.61, 6.958, 7.125, 7.108
Ser ⁹	4.408	8.543	3.783, 3.881
Arg ¹⁰	4.241	8.584	1.888, 1.596, 1.740, 3.137, 7.388
Ser ¹¹	4.238	8.456	3.826, 3.903
Asp ¹²	4.496	8.369	2.679, 2.733
Glu ¹³	4.066	8.193	2.074, 2.33
Leu ¹⁴	4.057	8.128	1.59, 0.825, 0.804
Ala ¹⁵	4.035	7.989	1.450
Lys ¹⁶	3.995	7.784	1.876, 1.650, 1.540, 1.360, 2.907, 7.698
Leu ¹⁷	4.106	7.841	1.785, 1.69, 1.575, 0.8677, 0.816
Leu ¹⁸	4.138	8.033	1.741, 1.550, 0.810, 0.827
Arg ¹⁹	4.177	7.808	1.867, 1.702, 1.611, 3.188, 7.278
Leu ²⁰	4.19	7.933	1.475, 1.67, 1.596, 0.868, 0.820
His ²¹	4.65	8.267	3.172, 3.314, 8.609, 7.288
Ala ²²	4.317	8.277	1.411
Gly ²³	3.926, 3.889	8.453	
NH ₂		7.51, 7.20	

Table 4-4 NMR assignment of **BBA1**. 13.8 mg/700 μ l, 90:10 H₂O:D₂O, pH 4.5, 7 °C, referenced to dioxane (δ 3.74), 600 MHz.

Residue	alpha	amide	side chain
Ac			
Tyr ¹	4.406	8.211	2.773, 6.883, 6.582
Thr ²	4.270	8.040	1.072, 4.082
Val ³	4.313	8.124	2.008, 0.910
Pro ⁴	4.429		2.272, 1.883, 2.053, 1.966, 3.667, 3.938
DSer ⁵	4.416	8.738	3.786, 3.852
Fen ⁶	5.235	8.619	3.618, 3.576, 7.678, 8.308, 8.022, 8.18, 8.797, 9.072
Thr ⁷	4.211	8.252	1.038, 3.998
Phe ⁸	4.259	8.147	2.698, 2.619, 6.980, 7.192, 7.139
Ser ⁹	4.352	8.334	3.833, 3.753
Arg ¹⁰	4.191	8.52	1.726, 1.855, 1.594, 1.574, 3.106, 3.148, 7.352
Ser ¹¹	4.239	8.400	3.893, 3.832
Asp ¹²	4.491	8.323	2.688, 2.635
Glu ¹³	4.071	8.153	2.075, 2.336
Leu ¹⁴	4.111	8.133	1.692, 1.672, 1.586, 0.858, 0.899
Ala ¹⁵	4.089	7.998	1.458
Lys ¹⁶	4.003	7.827	1.866, 1.888, 1.543, 1.650, 1.646, 2.912, 7.638
Leu ¹⁷	4.126	7.858	1.807, 1.712, 1.601, 0.849, 0.899
Leu ¹⁸	4.169	8.026	1.745, 1.577, 0.886, 0.842
Arg ¹⁹	4.194	7.829	1.614, 1.868, 1.706, 1.59, 3.195, 7.276
Leu ²⁰	4.201	7.944	1.488, 1.673, 0.891, 0.836
His ²¹	4.658	8.263	3.190, 3.326, 8.611, 7.298
Ala ²²	4.314	8.273	1.417
Gly ²³	3.937, 3.879	8.449	
NH ₂		7.50, 7.193	

Table 4-5 NMR assignment of **ZFO6**. 12.2 mg/700 μ l, 90:10 H₂O:D₂O, pH 4.5, 7 °C, referenced to dioxane (δ 3.74), 600 MHz.

Residue	alpha	amide	side chain
Ac	1.95		
Tyr ¹	4.684	8.340	2.940, 2.749, 6.71, 6.991
Thr ²	4.81	8.29	1.087, 3.98
Val ³	4.421	8.627	0.937, 2.096
DPro ⁴	4.440		1.951, 2.24, 2.037, 3.695
Ser ⁵	4.321	8.32	3.738
Tyr ⁶	4.606	8.354	2.805, 2.916, 6.69, 6.98
Thr ⁷	4.599 ^a	8.11 ^a	1.039, 3.879
Phe ⁸	4.72	8.627 ^a	2.895, 3.095, 7.165 ^a , 7.032
Ser ⁹	4.511	8.721	3.848, 3.933
Arg ¹⁰	4.355	8.51	1.788, 1.968, 1.651, 1.633, 3.20, 7.48
Ser ¹¹	4.201	8.527	3.894, 3.799
Asp ¹²	4.485	8.426	2.72
Glu ¹³	4.059	8.108	2.085, 2.357, 2.319
Leu ¹⁴	3.961	7.946	1.55, 1.50, 1.516, 0.736, 0.757
Ala ¹⁵	3.956	7.936	1.426
Lys ¹⁶	3.974	7.667	1.865, 1.532, 1.376, 1.646, 2.91
Leu ¹⁷	4.082	7.759	1.756, 1.551, 1.662, 0.783, 0.833
Leu ¹⁸	4.089	7.998	1.69, 1.658, 0.742
Arg ¹⁹	4.158	7.777	1.874, 1.610, 1.680, 3.180, 7.237
Leu ²⁰	4.190	7.875	1.460, 1.65, 1.59, 0.809, 0.850
His ²¹	4.653	8.236	3.186, 3.321, 8.582, 7.284
Ala ²²	4.322	8.223	1.407
Gly ²³	3.906	8.388	
NH ₂		7.142, 7.480	

Table 4-6 NMR assignment of **BBA2**. 14 mg/700 μ l, 90:10 H₂O:D₂O, pH 4.5, 17 °C, referenced to dioxane (δ 3.74), 600 MHz.

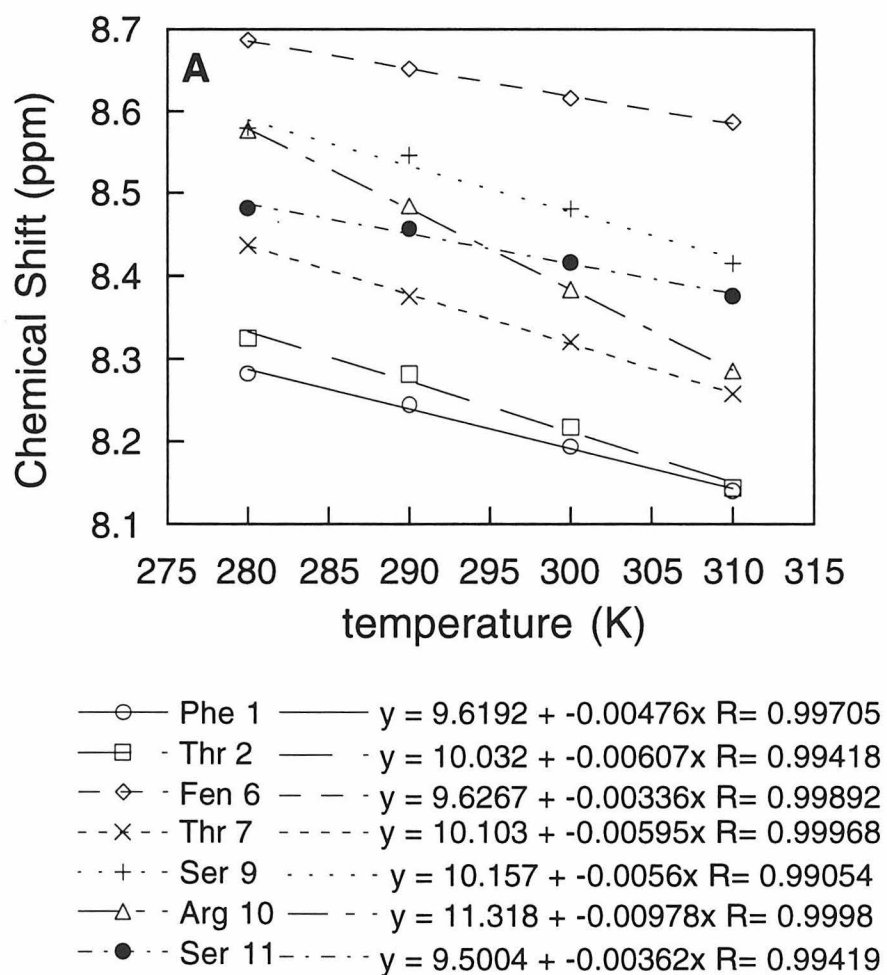
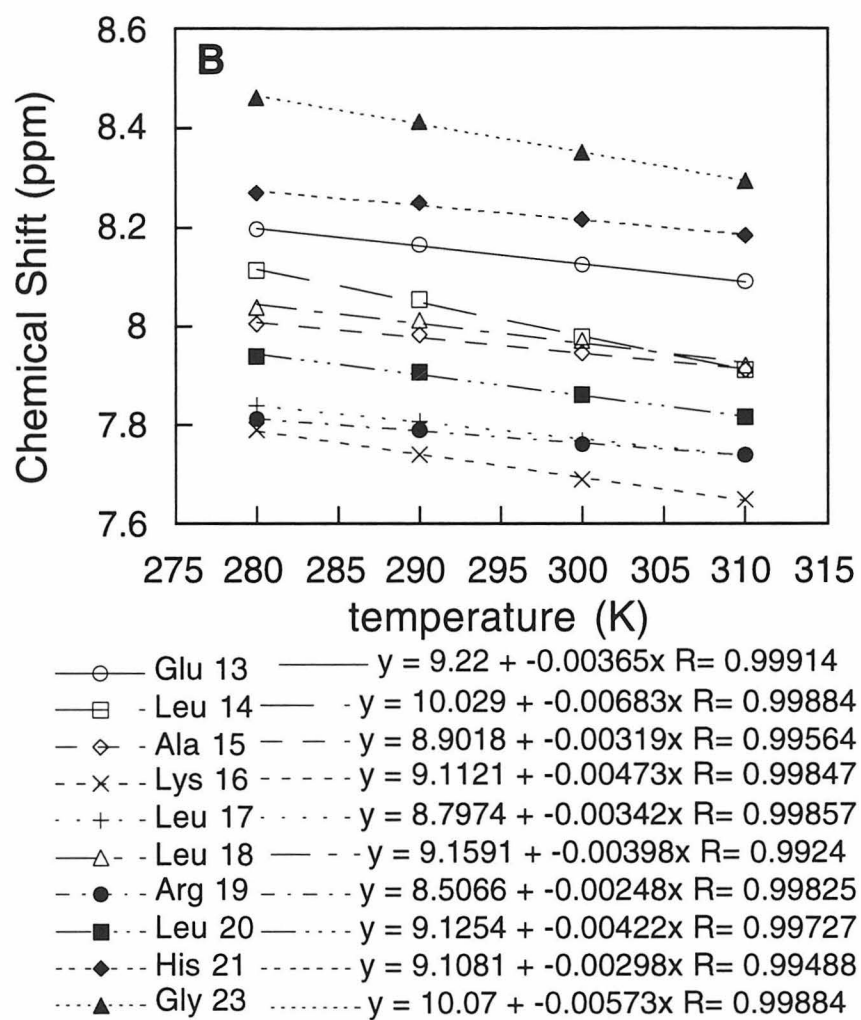


Figure 4-29 Plot of amide proton chemical shift *vs* temperature for **ZFO4**. **A.** The residues in the β -hairpin and loop regions are shown. **B.** (following page) The residues in the helix are shown.



Structure Determination

The structure of **BBA1** was determined using a simulated annealing protocol (SA)¹⁶ implemented through InsightII NMRchitect (Biosym Technologies, San Diego, CA). Cross peaks in the NOESY spectra (H₂O and D₂O) were classified as strong, medium, weak, or very weak based on the integrated volumes at various mixing times (50, 100, and 200 ms mixing times for H₂O spectra, 100 and 200 ms mixing times for D₂O spectra). The volumes were fit to a linear relationship with mixing time and calibrated to volume measurements for cross peaks corresponding to known inter-proton distances derived from a number of well separated β protons and from several of the aromatic protons of the Fen ring. Each NOE category was assigned a distance range: strong (1.0-2.5 Å), medium (1.8-3.5 Å), weak (1.8-5.0 Å), very weak (1.8-5.5 Å). A maximum force constant of 25 kcal mol⁻¹ Å⁻² was used for all NOE distance restraints. Pseudoatoms were used when necessary and standard corrections were applied to inter-proton distances restraints involving pseudoatoms.¹² These corrections are listed in Table 4-7. The upper bound of restraints for pseudoatoms defined for protons which were spectroscopically differentiable but not stereospecifically assigned were corrected in one of two manners. If both protons of the prochiral pair have restraints to the same atom then the restraint is collapsed to one pseudo atom and the upper bound is calculated exactly. If only one atom of the prochiral pair shows a restraint to a third atom, then a generic correction is added to the upper bound of the restraint which is transferred to the pseudoatom.

The use of pseudoatoms resulted in the generation of 301 restraints from 367 NOE volume measurements. The NOE derived restraints which were used in the structure generations procedure and obtained from both the D₂O and H₂O NOESY data, including the pseudoatom substitutions and corrections are listed

in Table 4-8. The *cis* proline amide conformation, accounting for approximately 15% of the total signal was visible in the NMR spectra. NOESY cross peaks attributable to this conformation were omitted for the generation of restraints. Dihedral angle restraints were included in the SA analysis to enforce *trans* geometry of the amide bond between Val³ and DPro⁴.

Group type	Residues	Correction
methyl (CH ₃)	Thr (γ), Ala (β), Leu (δ)	1.0 Å
methylene (CH ₂)	Glu, Asp, Pro, Lys, Arg, Leu, Phe, Tyr, Fen (β , γ , etc)	1.0 Å
dimethyl ((CH ₃) ₂)	Val (γ , γ'), Leu (δ , δ')	2.4 Å
aromatic (meta or para protons)	Tyr, Phe (ϵ), (δ)	2.0 Å

Table 4-7 Pseudoatom corrections added to the upper bounds of restraints used in the simulated annealing calculations.

Table 4-8 Proton 1	Proton 2	LB	Corrected UB	UB
1:PHN2_6:H3	1:PHN2_6:H1	1.800	5.000	5.000
1:SER_5:HN	1:PHN2_6:HN	1.800	3.500	3.500
1:ALA_22:HN	1:GLY_23:HN	1.800	3.500	3.500
1:ASP-_12:HN	1:GLU-_13:HN	1.800	3.500	3.500
1:LEU_17:HN	1:LEU_18:HN	1.800	3.500	3.500
1:PHN2_6:H3	1:PHN2_6:H2	1.800	3.500	3.500
1:PHN2_6:H1	1:PHN2_6:H2	1.800	3.500	3.500
1:LEU_14:HN	1:ALA_15:HN	1.800	3.500	3.500
1:LYS+_16:HN	1:ALA_15:HN	1.800	3.500	3.500
1:HIS+_21:HN	1:LEU_20:HN	1.800	3.500	3.500
1:ARG+_19:HN	1:LEU_20:HN	1.800	3.500	3.500
1:PHN2_6:HN	1:PHN2_6:H7	1.800	5.000	5.000
1:PHN2_6:H8	1:PHN2_6:H7	1.800	3.500	3.500
1:PHE_8:HE*	1:PHN2_6:H7	1.800	7.000	5.000
1:GLY_23:HN	1:NH2_24:HN1	1.800	5.000	5.000
1:NH2_24:HN2	1:NH2_24:HN1	1.000	2.500	2.500
1:HIS+_21:HN	1:HIS+_21:HD2	1.800	5.000	5.000
1:PHE_8:HD*	1:PHE_8:HE*	1.000	6.500	2.500
1:SER_9:HN	1:PHE_8:HD*	1.800	7.000	5.000
1:PHE_8:HN	1:PHE_8:HD*	1.800	5.500	3.500
1:TYR_1:HN	1:TYR_1:HD*	1.800	5.500	3.500
1:TYR_1:HE*	1:TYR_1:HD*	1.000	6.500	2.500
1:VAL_3:HN	1:PHN2_6:HN	1.800	5.000	5.000
1:THR_7:HN	1:PHN2_6:HN	1.800	3.500	3.500
1:ARG+_19:HN	1:LEU_18:HN	1.800	3.500	3.500
1:THR_2:HN	1:TYR_1:HD*	1.800	7.000	5.000
1:PHN2_6:HN	1:PHN2_6:HA	1.800	5.000	5.000
1:HIS+_21:HN	1:HIS+_21:HA	1.800	3.500	3.500
1:VAL_3:HN	1:THR_2:HA	1.800	3.500	3.500
1:THR_2:HN	1:TYR_1:HA	1.800	3.500	3.500
1:ALA_15:HN	1:ASP-_12:HA	1.800	5.000	5.000
1:THR_7:HN	1:THR_7:HA	1.800	3.500	3.500
1:SER_5:HN	1:PRO_4:HA	1.800	3.500	3.500
1:ARG+_10:HN	1:SER_9:HA	1.800	3.500	3.500
1:PHN2_6:HN	1:SER_5:HA	1.800	3.500	3.500
1:VAL_3:HN	1:VAL_3:HA	1.800	3.500	3.500
1:PHN2_6:H2	1:PHE_8:HA	1.800	5.000	5.000
1:SER_9:HN	1:PHE_8:HA	1.800	3.500	3.500
1:PHE_8:HN	1:PHE_8:HA	1.800	3.500	3.500
1:GLY_23:HN	1:ALA_22:HA	1.800	3.500	3.500
1:ALA_22:HN	1:ALA_22:HA	1.800	3.500	3.500
1:ARG+_10:HN	1:ARG+_10:HA	1.800	3.500	3.500
1:SER_11:HN	1:SER_11:HA	1.800	3.500	3.500

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:ASP-12:HN	1:SER_11:HA	1.800	3.500	3.500
1:HIS+21:HN	1:LEU_20:HA	1.800	3.500	3.500
1:LEU_20:HN	1:LEU_20:HA	1.800	3.500	3.500
1:ARG+19:HN	1:ARG+19:HA	1.800	3.500	3.500
1:LEU_18:HN	1:LEU_18:HA	1.800	3.500	3.500
1:LEU_17:HN	1:LEU_17:HA	1.000	2.500	2.500
1:LEU_14:HN	1:LEU_14:HA	1.800	3.500	3.500
1:ALA_15:HN	1:ALA_15:HA	1.800	3.500	3.500
1:LYS+16:HN	1:LYS+16:HA	1.800	3.500	3.500
1:THR_2:HN	1:THR_2:HB	1.800	3.500	3.500
1:SER_5:HN	1:SER_5:HB*	1.800	3.388	3.388
1:PHN2_6:HN	1:SER_5:HB*	1.800	4.922	4.922
1:PHN2_6:HN	1:PHN2_6:HB*	1.800	3.388	3.388
1:THR_7:HN	1:PHN2_6:HB*	1.800	4.923	4.923
1:SER_11:HA	1:ALA_15:HN	1.800	5.000	5.000
1:THR_7:HN	1:PHN2_6:HA	1.800	3.500	3.500
1:ARG+10:HN	1:SER_9:HB*	1.800	3.388	3.388
1:SER_9:HN	1:SER_9:HB*	1.800	3.388	3.388
1:SER_11:HN	1:SER_11:HB*	1.800	4.500	3.500
1:THR_7:HN	1:THR_7:HB	1.800	3.500	3.500
1:ASP-12:HN	1:SER_11:HB*	1.800	4.500	3.500
1:PHE_8:HN	1:THR_2:HA	1.800	5.000	5.000
1:PRO_4:HA	1:PHN2_6:HN	1.800	5.000	5.000
1:GLU-13:HN	1:ASP-12:HA	1.800	3.500	3.500
1:PHN2_6:H7	1:PHN2_6:HA	1.800	5.000	5.000
1:ASP-12:HN	1:ASP-12:HA	1.800	3.500	3.500
1:SER_9:HN	1:SER_9:HA	1.800	3.500	3.500
1:VAL_3:HN	1:THR_7:HA	1.800	3.500	3.500
1:SER_5:HA	1:SER_5:HN	1.800	3.500	3.500
1:LEU_18:HN	1:ALA_15:HA	1.800	3.500	3.500
1:LEU_20:HN	1:LEU_17:HA	1.800	3.500	3.500
1:GLU-13:HN	1:GLU-13:HA	1.800	3.500	3.500
1:GLU-13:HA	1:LYS+16:HN	1.800	3.500	3.500
1:LEU_18:HA	1:ALA_22:HN	1.800	5.000	5.000
1:PHN2_6:H7	1:PHN2_6:HB*	1.800	3.388	3.388
1:THR_2:HB	1:VAL_3:HN	1.800	3.500	3.500
1:ARG+19:HN	1:ARG+19:HD*	1.800	6.000	5.000
1:HIS+21:HN	1:HIS+21:HB*	1.800	3.387	3.387
1:TYR_1:HN	1:TYR_1:HB*	1.800	3.390	3.390
1:SER_9:HN	1:PHE_8:HB*	1.800	3.390	3.390
1:GLU-13:HN	1:GLU-13:HG*	1.800	4.500	3.500
1:VAL_3:HN	1:VAL_3:HB	1.800	3.500	3.500
1:GLU-13:HN	1:GLU-13:HB*	1.800	4.500	3.500
1:SER_5:HN	1:PRO_4:HG*	1.800	6.000	5.000

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:TYR_1:HN	1:ACE_0:HA*	1.800	4.500	3.500
1:ARG+_10:HN	1:ARG+_10:HB*	1.800	4.500	3.500
1:SER_11:HN	1:ARG+_10:HB*	1.800	6.000	5.000
1:LEU_20:HN	1:ARG+_19:HB*	1.800	4.500	3.500
1:SER_5:HN	1:PRO_4:HB*	1.800	4.910	4.910
1:LEU_18:HN	1:LEU_18:HB*	1.000	3.500	2.500
1:ARG+_10:HN	1:ARG+_10:HG*	1.800	3.387	3.387
1:SER_11:HN	1:ARG+_10:HG*	1.800	4.922	4.922
1:LEU_20:HN	1:LEU_20:HG	1.800	3.500	3.500
1:HIS+_21:HN	1:LEU_20:HG	1.800	5.500	5.000
1:LEU_17:HN	1:LEU_17:HG	1.800	3.500	3.500
1:LEU_18:HN	1:LEU_18:HG	1.800	3.500	3.500
1:LEU_14:HN	1:ALA_15:HB*	1.800	6.000	5.000
1:ALA_15:HN	1:ALA_15:HB*	1.800	4.500	3.500
1:LYS+_16:HN	1:ALA_15:HB*	1.800	4.500	3.500
1:GLY_23:HN	1:ALA_22:HB*	1.800	6.000	5.000
1:ALA_22:HN	1:ALA_22:HB*	1.800	4.500	3.500
1:LYS+_16:HN	1:LYS+_16:HD*	1.800	6.000	5.000
1:VAL_3:HN	1:THR_2:HG2*	1.800	4.500	3.500
1:THR_2:HN	1:THR_2:HG2*	1.800	4.500	3.500
1:THR_7:HN	1:THR_7:HG2*	1.800	4.500	3.500
1:TYR_1:HN	1:THR_7:HG2*	1.800	6.000	5.000
1:VAL_3:HN	1:VAL_3:HG*	1.800	5.900	3.500
1:PHN2_6:H7	1:VAL_3:HG*	1.800	7.400	5.000
1:HIS+_21:HN	1:LEU_18:HD*	1.800	7.400	5.000
1:LEU_20:HD*	1:HIS+_21:HN	1.800	7.400	5.000
1:LEU_14:HD*	1:LEU_14:HN	1.800	5.058	5.058
1:LEU_18:HD*	1:LEU_18:HN	1.800	4.208	4.208
1:LEU_20:HD*	1:LEU_20:HN	1.800	5.058	5.058
1:LEU_17:HD*	1:LEU_17:HN	1.800	4.208	4.208
1:HIS+_21:HN	1:LEU_20:HB*	1.800	3.389	3.389
1:ALA_15:HN	1:LEU_14:HB*	1.800	4.500	3.500
1:LYS+_16:HN	1:LYS+_16:HG*	1.800	4.500	3.500
1:LYS+_16:HN	1:LYS+_16:HB*	1.000	3.500	2.500
1:ARG+_19:HN	1:ARG+_19:HG*	1.800	3.387	3.387
1:LEU_17:HD*	1:LEU_18:HN	1.800	7.400	5.000
1:PHE_8:HN	1:PHE_8:HB*	1.800	3.390	3.390
1:ASP-_12:HN	1:ASP-_12:HB*	1.800	3.389	3.389
1:GLU-_13:HN	1:ASP-_12:HB*	1.800	3.389	3.389
1:TYR_1:HD*	1:TYR_1:HA	1.800	5.500	3.500
1:PHE_8:HD*	1:THR_7:HA	1.800	7.000	5.000
1:PHE_8:HD*	1:PHE_8:HA	1.800	5.500	3.500
1:PHN2_6:HA	1:PHN2_6:HB*	1.800	3.388	3.388
1:PHE_8:HE*	1:VAL_3:HG*	1.800	7.900	3.500

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:PHE_8:HE*	1:LEU_14:HD*	1.800	7.900	3.500
1:ARG+_10:HD*	1:ARG+_10:HE	1.800	4.500	3.500
1:ARG+_10:HB*	1:ARG+_10:HE	1.800	6.000	5.000
1:ARG+_10:HG*	1:ARG+_10:HE	1.800	4.225	4.225
1:ARG+_19:HB*	1:ARG+_19:HE	1.800	6.000	5.000
1:ARG+_19:HG*	1:ARG+_19:HE	1.800	4.922	4.922
1:LEU_20:HD*	1:HIS+_21:HD2	1.800	5.784	5.784
1:VAL_3:HB	1:PHE_8:HE*	1.800	5.500	3.500
1:PHE_8:HB*	1:PHE_8:HD*	1.800	5.431	5.431
1:GLU-_13:HB*	1:PHE_8:HD*	1.800	8.000	5.000
1:VAL_3:HG*	1:PHE_8:HD*	1.800	9.400	5.000
1:LEU_14:HD*	1:PHE_8:HD*	1.800	7.839	7.839
1:TYR_1:HB*	1:TYR_1:HD*	1.800	5.431	5.431
1:ACE_0:HA*	1:TYR_1:HD*	1.800	8.000	5.000
1:VAL_3:HG*	1:TYR_1:HD*	1.800	9.400	5.000
1:VAL_3:HG*	1:TYR_1:HE*	1.800	9.400	5.000
1:LEU_14:HD*	1:TYR_1:HD*	1.800	7.839	7.839
1:PHE_8:HZ	1:VAL_3:HG*	1.800	7.400	5.000
1:LEU_17:HD*	1:PHE_8:HZ	1.800	7.400	5.000
1:LEU_17:HD*	1:PHE_8:HE*	1.800	9.400	5.000
1:PHE_8:HZ	1:LEU_14:HD*	1.800	5.900	3.500
1:ARG+_10:HA	1:SER_9:HA	1.800	5.000	5.000
1:THR_2:HA	1:THR_2:HB	1.800	3.500	3.500
1:THR_7:HA	1:THR_7:HB	1.800	3.500	3.500
1:SER_9:HA	1:SER_9:HB*	1.800	4.500	3.500
1:SER_11:HA	1:SER_11:HB*	1.800	3.388	3.388
1:LEU_18:HA	1:HIS+_21:HB*	1.800	6.000	5.000
1:LYS+_16:HA	1:ARG+_19:HD*	1.800	6.000	5.000
1:ARG+_19:HA	1:ARG+_19:HD*	1.800	4.500	3.500
1:HIS+_21:HA	1:HIS+_21:HB*	1.800	3.387	3.387
1:ARG+_10:HA	1:ARG+_10:HD*	1.800	6.000	5.000
1:GLU-_13:HA	1:LYS+_16:HE*	1.800	6.000	5.000
1:TYR_1:HA	1:TYR_1:HB*	1.800	3.390	3.390
1:ASP-_12:HA	1:ASP-_12:HB*	1.800	3.389	3.389
1:GLU-_13:HA	1:GLU-_13:HG*	1.800	4.500	3.500
1:VAL_3:HA	1:VAL_3:HB	1.800	3.500	3.500
1:SER_11:HA	1:GLU-_13:HB*	1.800	6.000	5.000
1:GLU-_13:HA	1:GLU-_13:HB*	1.000	3.500	2.500
1:PRO_4:HA	1:PRO_4:HG*	1.800	4.500	3.500
1:PRO_4:HD*	1:PRO_4:HG*	1.000	4.500	2.500
1:SER_9:HA	1:ARG+_10:HB*	1.800	4.500	3.500
1:GLU-_13:HA	1:LYS+_16:HB*	1.800	4.500	3.500
1:LYS+_16:HA	1:LYS+_16:HB*	1.000	3.500	2.500
1:ARG+_19:HA	1:ARG+_19:HB*	1.000	3.500	2.500

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:PRO_4:HA	1:PRO_4:HB*	1.800	3.370	3.370
1:PRO_4:HD*	1:PRO_4:HB*	1.800	4.400	4.400
1:LEU_17:HA	1:LEU_17:HB*	1.000	2.341	2.341
1:LEU_20:HA	1:LEU_20:HG	1.000	2.500	2.500
1:LEU_18:HA	1:LEU_18:HG	1.000	2.500	2.500
1:LEU_20:HA	1:LEU_20:HB*	1.800	3.389	3.389
1:ASP-12:HA	1:ALA_15:HB*	1.800	4.500	3.500
1:ALA_15:HA	1:ALA_15:HB*	1.000	3.500	2.500
1:ALA_22:HA	1:ALA_22:HB*	1.800	4.500	3.500
1:LYS+16:HA	1:LYS+16:HD*	1.800	4.500	3.500
1:PRO_4:HD*	1:THR_2:HG2*	1.800	7.000	5.000
1:THR_2:HA	1:THR_2:HG2*	1.800	4.500	3.500
1:VAL_3:HA	1:THR_2:HG2*	1.800	6.000	5.000
1:THR_2:HB	1:THR_2:HG2*	1.800	4.500	3.500
1:THR_7:HA	1:THR_7:HG2*	1.800	4.500	3.500
1:THR_7:HB	1:THR_7:HG2*	1.800	4.500	3.500
1:VAL_3:HA	1:VAL_3:HG*	1.800	5.900	3.500
1:PRO_4:HD*	1:VAL_3:HG*	1.800	6.900	3.500
1:LEU_20:HD*	1:LEU_20:HA	1.000	3.702	3.702
1:PHE_8:HA	1:PHE_8:HB*	1.800	3.390	3.390
1:ARG+10:HA	1:ARG+10:HB*	1.800	4.500	3.500
1:ARG+10:HA	1:ARG+10:HG*	1.800	3.387	3.387
1:LEU_17:HA	1:LEU_17:HG	1.000	2.500	2.500
1:LEU_14:HA	1:LEU_17:HB*	1.800	4.500	3.500
1:ALA_15:HA	1:LEU_18:HB*	1.800	4.500	3.500
1:LYS+16:HA	1:LYS+16:HG*	1.800	4.500	3.500
1:LEU_14:HD*	1:LEU_14:HA	1.800	4.208	4.208
1:LEU_18:HD*	1:LEU_18:HA	1.000	4.900	2.500
1:LEU_17:HD*	1:LEU_17:HA	1.800	4.208	4.208
1:GLU-13:HG*	1:GLU-13:HB*	1.000	4.500	2.500
1:PRO_4:HB*	1:PRO_4:HG*	1.800	5.500	3.500
1:ARG+19:HD*	1:ARG+19:HB*	1.800	5.500	3.500
1:ARG+10:HB*	1:ARG+10:HG*	1.000	4.500	2.500
1:ARG+19:HD*	1:ARG+19:HG*	1.000	3.934	3.934
1:LYS+16:HE*	1:LYS+16:HG*	1.800	4.413	4.413
1:ARG+19:HB*	1:ARG+19:HG*	1.000	4.500	2.500
1:LYS+16:HB*	1:ALA_15:HB*	1.800	5.500	3.500
1:ASP-12:HB*	1:ALA_15:HB*	1.800	7.000	5.000
1:LYS+16:HB*	1:LYS+16:HD*	1.800	5.500	3.500
1:LYS+16:HG*	1:LYS+16:HD*	1.000	3.934	3.934
1:LYS+16:HE*	1:LYS+16:HD*	1.800	5.500	3.500
1:VAL_3:HB	1:VAL_3:HG*	1.000	4.900	2.500
1:HIS+21:HB*	1:LEU_20:HD*	1.800	6.900	3.500
1:LEU_20:HD*	1:LEU_20:HB*	1.800	5.263	5.263

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:VAL_3:HG*	1:LEU_18:HD*	1.000	7.300	2.500
1:ARG+_10:HN	1:SER_11:HN	1.800	3.500	3.500
1:SER_9:HN	1:PHE_8:HN	1.800	3.500	3.500
1:VAL_3:HN	1:PRO_4:HD*	1.800	6.000	5.000
1:LEU_14:HN	1:GLU-_13:HB*	1.800	4.500	3.500
1:LEU_20:HN	1:LEU_20:HB*	1.800	3.389	3.389
1:LEU_14:HN	1:GLU-_13:HG*	1.800	6.000	5.000
1:ALA_15:HN	1:GLU-_13:HB*	1.800	6.000	5.000
1:PHE_8:HE*	1:PHN2_6:HB*	1.800	6.945	6.945
1:PHE_8:HD*	1:ACE_0:HA*	1.800	8.000	5.000
1:TYR_1:HE*	1:ACE_0:HA*	1.800	8.000	5.000
1:TYR_1:HE*	1:TYR_1:HB*	1.800	6.946	6.946
1:SER_5:HA	1:SER_5:HB*	1.000	2.341	2.341
1:GLU-_13:HB*	1:ASP-_12:HA	1.800	6.000	5.000
1:VAL_3:HB	1:PRO_4:HD*	1.800	4.500	3.500
1:THR_2:HA	1:VAL_3:HG*	1.800	7.400	5.000
1:GLU-_13:HA	1:LYS+_16:HD*	1.800	6.000	5.000
1:LEU_14:HD*	1:PHE_8:HB*	1.800	6.816	6.816
1:LEU_17:HD*	1:GLU-_13:HB*	1.800	6.816	6.816
1:PHN2_6:HA	1:TYR_1:HD*	1.800	7.000	5.000
1:HIS+_21:HB*	1:HIS+_21:HD2	1.800	4.922	4.922
1:LEU_14:HD*	1:TYR_1:HE*	1.800	9.400	5.000
1:THR_7:HA	1:THR_2:HA	1.000	2.500	2.500
1:ASP-_12:HA	1:GLU-_13:HA	1.800	5.000	5.000
1:ARG+_19:HA	1:ALA_22:HB*	1.800	6.000	5.000
1:LEU_18:HD*	1:HIS+_21:HB*	1.800	8.400	5.000
1:LYS+_16:HG*	1:GLU-_13:HB*	1.800	7.000	5.000
1:PHE_8:HD*	1:PHN2_6:H7	1.800	7.000	5.000
1:PHE_8:HE*	1:PHN2_6:H8	1.800	7.000	5.000
1:PHE_8:HD*	1:PHN2_6:H8	1.800	7.000	5.000
1:PHE_8:HE*	1:TYR_1:HE*	1.800	9.000	5.000
1:PHE_8:HD*	1:TYR_1:HE*	1.800	9.000	5.000
1:PHE_8:HE*	1:TYR_1:HD*	1.800	9.000	5.000
1:PHE_8:HD*	1:TYR_1:HD*	1.800	9.000	5.000
1:PHE_8:HZ	1:PHN2_6:H7	1.800	5.000	5.000
1:PHN2_6:HA	1:PHE_8:HD*	1.800	7.000	5.000
1:PHN2_6:HA	1:PHE_8:HE*	1.800	7.000	5.000
1:PHN2_6:HB*	1:PHN2_6:H8	1.800	4.923	4.923
1:PHE_8:HZ	1:PHN2_6:HB*	1.800	5.000	5.000
1:PHE_8:HD*	1:PHN2_6:HB*	1.800	6.945	6.945
1:PHE_8:HD*	1:THR_2:HA	1.800	7.000	5.000
1:TYR_1:HD*	1:THR_2:HA	1.800	7.000	5.000
1:PHE_8:HE*	1:PHE_8:HA	1.800	7.000	5.000
1:PHE_8:HE*	1:LEU_14:HA	1.800	7.000	5.000

Table 4-8 continued Proton 1	Proton 2	LB	Corrected UB	UB
1:HIS+_21:HA	1:HIS+_21:HD2	1.800	5.000	5.000
1:LEU_17:HA	1:HIS+_21:HD2	1.800	5.000	5.000
1:THR_2:HA	1:PHE_8:HE*	1.800	7.000	5.000
1:THR_7:HA	1:PHE_8:HE*	1.800	7.000	5.000
1:LEU_14:HA	1:PHE_8:HD*	1.800	7.000	5.000
1:TYR_1:HA	1:TYR_1:HE*	1.800	7.000	5.000
1:THR_7:HA	1:PHN2_6:HA	1.800	5.000	5.000
1:SER_5:HA	1:PHN2_6:HA	1.800	5.000	5.000
1:PHE_8:HE*	1:GLU-_13:HG*	1.800	8.000	5.000
1:PHE_8:HD*	1:GLU-_13:HG*	1.800	8.000	5.000
1:PHE_8:HB*	1:PHE_8:HE*	1.800	8.000	5.000
1:VAL_3:HB	1:PHN2_6:H7	1.800	5.500	5.000
1:ALA_15:HB*	1:TYR_1:HE*	1.800	8.000	5.000
1:LEU_20:HA	1:HIS+_21:HA	1.800	5.000	5.000
1:SER_11:HA	1:ASP-_12:HA	1.800	5.000	5.000
1:PHE_8:HA	1:THR_7:HA	1.800	5.000	3.500
1:ALA_22:HA	1:HIS+_21:HA	1.800	5.000	5.000
1:PRO_4:HD*	1:THR_2:HA	1.800	6.000	5.000
1:VAL_3:HA	1:THR_2:HB	1.800	5.000	5.000
1:SER_9:HB*	1:ASP-_12:HB*	1.800	5.537	5.537
1:GLU-_13:HA	1:ASP-_12:HB*	1.800	6.000	5.000
1:LYS+_16:HE*	1:LYS+_16:HA	1.800	6.000	5.000
1:ARG+_10:HA	1:GLU-_13:HG*	1.800	6.000	5.000
1:ACE_0:HA*	1:TYR_1:HA	1.800	6.000	5.000
1:SER_11:HA	1:ALA_15:HB*	1.800	6.000	5.000
1:VAL_3:HG*	1:PHN2_6:HB*	1.800	7.348	7.348
1:VAL_3:HB	1:PHN2_6:HB*	1.800	5.000	5.000
1:LYS+_16:HB*	1:LYS+_16:HE*	1.800	7.000	5.000
1:VAL_3:HG*	1:PHE_8:HB*	1.800	8.400	5.000
1:LEU_17:HD*	1:GLU-_13:HG*	1.800	6.089	6.089
1:PRO_4:HG*	1:THR_2:HG2*	1.800	7.000	5.000
1:PRO_4:HG*	1:VAL_3:HG*	1.800	8.400	5.000
1:ACE_0:HA*	1:THR_7:HG2*	1.800	7.000	5.000
1:LEU_14:HA	1:PHE_8:HZ	1.800	5.500	5.000
1:THR_7:HG2*	1:PHN2_6:HA	1.800	6.000	5.000
1:PHE_8:HA	1:GLU-_13:HB*	1.800	6.000	5.000
1:TYR_1:HD*	1:SER_9:HA	1.800	7.000	5.000
1:PHE_8:HD*	1:TYR_1:HB*	1.800	8.000	5.000

Table 4-8 List of restraints used in the structure generation process. The upper bound (UB) and lower bound (LB) are given. The corrected UB results from the use of pseudo atom corrections.

NOE restrained molecular dynamics and simulated annealing were carried out in four stages; 1. Preparation (including randomization of the initial coordinates), 2. Folding, 3. Annealing, and 4. Final Minimization. The folding stage consist of 3 steps of dynamics at 1000 K. In each step, the force constants are scaled to a fraction of their total value. The first step involves scaling the NMR derived force constant to 1.0 times its total value over 30 ps of dynamics, while the force constants for the non-bonded (NB) and the covalent (Cov.) interactions are maintained at a small fraction of their total value. In the second stage (10 ps) the covalent force constants are similarly scaled. In the last step (40 ps) the non-bonded force constants are scaled to a quarter of their total value while the NMR derived force constants are doubled. The folding stages are followed by a gradual cooling or annealing of the structure. The temperature is cooled in a geometric progression consisting of 5 steps of 4 ps dynamics. The final stage of the procedure in a series of four minimizations. In the last two of these minimizations the force constants for all terms of the energy equation are maintained at their total values These stages are delineated in Table 4-9 and described graphically in Figure 4-9.

Stage	Step	Procedure ^a	Force constants ^b	Length	Temp (K)
Preparation					
1	1	randomize coordinates	all 0.000001	100 iterations	1000
1	2	minimization	all 0.000001	500 iterations	1000
Folding					
2	1	Dynamics Scale: NMR	NMR 0.001→1.0 Chiral 0.001→0.15 Cov. 0.0002→0.15	30 ps	1000
2	2	Covalent (Cov.)	Cov. 0.15→1.0 Chiral 0.15→1.0 NB 0.001→0.015	10 ps	1000
2	3	Non-bonds (NB)	NB 0.015→0.25 NMR 1.0→2.0	40 ps	1000
Annealing					
3	1	dynamics	2.0 NMR 1.0 Cov. 0.25 NB	4 ps	650
3	2	dynamics	same as above	4 ps	475
3	3	dynamics	same as above	4 ps	385
3	4	dynamics	same as above	4 ps	340
3	5	dynamics	same as above	4 ps	300
4	1	minimization	same as above	100 iterations	300
4	2	minimization	same as above	500 iterations	300
4	3	minimization	1.0 NMR 1.0 Cov. 1.0 NB	100 iterations	300
4	4	minimization	same as above	1500 iterations	300

Table 4-9 The stages of NOE-restrained simulated annealing used in the structure generation process. ^a The force constant which is scaled in the step listed for stage 2 is given. ^b The values given are the coefficients of the total force constants used in the step of the protocol. For dynamics runs in step 2 the value is scaled from the initial to the final (initial→final) value over the duration of the dynamics. In minimizations and simulated annealing dynamics the value is constant.

Using this procedure, 45 structures were generated, 29 of which converged to a common fold and were considered for further analysis. The remaining 16 structures had high energies and significant restraint violations. A success rate of ~70% is typical for such a procedure.¹⁶

RMSD analysis of these final structures was accomplished by superimposing all 29 structures to a structure with the lowest NMR derived energy based on the backbone atoms of residues 1-22. The N-terminal acetyl cap and the C-terminal glycine were left out of the calculation since they appear disordered in the NMR structure. The RMSD for the backbone atoms or all heavy atoms from the average structure was then calculated for each structure and averaged. A similar analysis was performed using the backbone or heavy atoms of residues 1-22.

All structures were analyzed for the presence of hydrogen bonding interactions. A backbone - backbone hydrogen bond was defined if the amide proton was within 2.5 Å of the oxygen of an amide carbonyl, and the angle was greater than 120°. The results of this analysis are summarized in Table 4-10. Although there appears to be a break in the H-bonding pattern of the helix at position 15 inspection of several structures indicated that for many structures the amide proton was just outside the range (2.5Å) defined for an H-bonding interaction. Partial overlap of the amide resonances for Lys¹⁶ and Arg¹⁹ prevented the identification a few crucial NOEs ($d_{\alpha N}(i, i+3)$ and $d_{\alpha \beta}(i, i+3)$) indicative of helix formation, although the surrounding NOEs adequately define the length of the helix.

acceptor	donor	fraction (of 29)
Helix:		
12 (CO)	16 (NH)	15/29
13 (CO)	17 (NH)	22/29
14 (CO)	18 (NH)	28/29
15 (CO)	19 (NH)	2/29
16 (CO)	20 (NH)	6/29
17 (CO)	21 (NH)	27/29
18 (CO)	22 (NH)	2/29
Sheet/Turn:		
3 (CO)	6 (NH)	27/29
6 (CO)	3 (NH)	6/29

Table 4-10 Hydrogen bonding interactions in the 29 NMR-derived structures of **BBA1**.

ANS Binding experiments

ANS titration experiments were performed in either 50 mM HEPES, pH 8.25 or 20 mM acetate pH 4.5. Aliquots of ANS in water (1-6 μ l) were added to the 700 μ l sample, and the fluorescence spectrum acquired between additions. Emission spectra were recorded with excitation at 370 nm. An identical experiment was performed in the absence of peptide. ANS titrations of the BBA1-Zn²⁺ complex were performed at pH 8.25, after the addition of 2 equivalents of ZnCl₂.

References

- (1) Goodsell, D. S.; Olson, A. J. "Soluble Proteins: Size, Shape and Function," *Trends in Biol. Sci.* **1993**, *18*, 65-68.
- (2) Srere, P. A. "Why are Enzymes so Big?," *Trends in Biol. Sci.* **1984**, *9*, 387-390.
- (3) Bontems, F.; Rounestand, C.; Boyot, P.; Gilquin, B.; Doljansky, Y.; Menez, A.; Toma, F. "Three-dimensional Structure of Natural Charybdotoxin in Aqueous Solution by ^1H -NMR," *Eur. J. Biochem.* **1991**, *196*, 19-28.
- (4) Klug, A.; Schwabe, J. W. R. "Zinc Fingers," *FASEB J.* **1995**, *9*, 597-604.
- (5) Raleigh, D. P.; Betz, S. F.; DeGrado, W. F. "A *de Novo* Designed Protein Mimics the Native State of Natural Proteins," *J. Am. Chem. Soc.* **1995**, *117*, 7558-7559.
- (6) Betz, S. F.; Bryson, J. W.; DeGrado, W. F. "Native-like and Structurally Characterized Designed α -helical Bundles," *Curr. Opin. Struct. Biol.* **1995**, *5*, 457-463.
- (7) Betz, S.F.; Raleigh, D.P.; DeGrado, W.F. "De Novo Protein Design: from Molten Globules to Native-like States," *Curr. Opin. Struct. Biol.* **1993**, *3*, 601-610.
- (8) Hecht, M. H. "De Novo Design of β -sheet Proteins," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8729-8730.
- (9) Berg, J.M. "Zinc Finger Domains: Hypotheses and Current Knowledge," *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405-421.
- (10) Eis, P.S.; Lakowicz, J.R. "Time-Resolved Energy Transfer Measurements of Donor-Acceptor Distance Distributions and Intramolecular Flexibility of a CCHH Zinc Finger Peptide," *Biochemistry* **1993**, *32*, 7981-7993.
- (11) Parraga, G.; Horvath, S.; Eisen, A.; Taylor, W.; Hood, L.; Young, E.; Klevit, R.E. "Zinc-Dependent Structure of a Single-Finger Domain of Yeast ADR1," *Science* **1988**, *241*, 1489-1492.

- (12) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley and Sons: New York, 1986.
- (13) Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide.," *J. Am. Chem. Soc.* **1992**, *114*, 3182.
- (14) Weiss, M. A.; Keutmann, H. T. "Alternating Zinc Finger Motifs in the Male-Associated Protein ZFY: Defining Architectural Rules by Mutagenesis and Design of an "Aromatic Swap" Second-Site Revertant," *Biochemistry* **1990**, *29*, 9808-9813.
- (15) Englander, S. W.; Downer, N. W.; Teitelbaum, H. "Hydrogen Exchange," *Ann. Rev. Biochem.* **1972**, *41*, 901-924.
- (16) Nilges, M.; Clore, G. M.; Gronenborn, A. M. "Determination of Three-dimensional Structures of Proteins from Interproton Distance Data by Dynamical Simulated Annealing from a Random Array of Atoms," *FEBS Lett.* **1988**, *239*, 129-136.
- (17) Lee, M. S. ; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. "Three-Dimensional Solution Structure of a Single Zinc Finger DNA-Binding Domain," *Science* **1989**, *245*, 635-637.
- (18) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. "Study of the "Molten Globule" Intermediate State in Protein Folding by a Hydrophobic Fluorescent Probe," *Biopolymers* **1991**, *31*, 119-128.
- (19) Quinn, T. P. ; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C. "Betadoublet: *De novo* Design, Synthesis, and Characterization of a β -sandwich Protein," *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747-8751.

- (20) Handel, T.M.; Williams, S.A.; DeGrado, W.F. "Metal Ion-Dependent Modulation of the Dynamics of a Designed Protein," *Science* **1993**, 261, 879-885.
- (21) Presta, L. G.; Rose, G. D. "Helix Signals in Proteins," *Science* **1988**, 240, 1632-1641.
- (22) Pavletich, N. P.; Pabo, C. O. "Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 Å," *Science* **1991**, 252, 809-817.
- (23) Frankel, A.D.; Berg, J.M.; Pabo, C.O. "Metal-dependent Folding of a Single Zinc Finger From Transcription Factor IIIA," *Proc. Natl. Acad. Sci. U.S.A.* **1987**, 84, 4841-4845.
- (24) Weiss, M.A.; Mason, K.A.; Dahl, C.E.; Keutmann, H.T. "Alternating Zinc-Finger Motifs in the Human Male-Associated Protein ZFY," *Biochemistry* **1990**, 29, 5660-5664.
- (25) Kochoyan, M.; Havel, T. F.; Nguyen, D. T.; Dahl, C. E.; Keutmann, H. T.; Weiss, M. A. "Alternating Zinc Fingers in the Human Male Associated Protein AFY: 2D NMR Structure of an Even Finger and Implication for "Jumping-Linker" DNA Recognition," *Biochemistry* **1991**, 30, 3371-3386.
- (26) Jasanoff, A.; Weiss, M. A. "Aromatic-Histidine Interactions in the Zinc Finger Motif: Structural Inequivalence of Phenylalanine and Tyrosine in the Hydrophobic Core," *Biochemistry* **1993**, 32, 1423-1432.
- (27) Rose, G. D.; Gierasch, L. M.; Smith, J. A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.* **1985**, 37, 1-105.
- (28) Imperiali, B.; Sinha Roy, R. "Coenzyme-Amino Acid Chimeras: New Residues for the Assembly of Functional Proteins," *J. Am. Chem. Soc.* **1994**, 116, 12083-12084.
- (29) Imperiali, B.; Sinha Roy, R. "Stereoselective Synthesis and Peptide Incorporation of a Pyridoxal Coenzyme-Amino Acid Chimera," *J. Org. Chem.* **1995**, 60, 1891-1894.

- (30) Imperiali, B.; Sinha Roy, R.; Wang, L. "Peptidyl Models for Coenzyme Catalysis," in *Peptides: Chemistry, Structure, and Biology (Proceedings of the 14th American Peptide Symposium)*; Escom Science Publisher: Leinden, 1996; pp 546.
- (31) Imperiali, B.; Sinha Roy, R.; Walkup, G. K.; Wang, L. "Unnatural Amino Acids for the Design of Functional Proteins: Biomimetic Catalysis Using Coenzyme Amino Acids," in *NATO Series on Biomimetic Catalysis*; Wilcox, C. S. Eds.; Kluwer Academic Publishing: Dordrecht, 1996; pp in press.
- (32) Imperiali, B.; Fisher, S. L. "Stereoselective Synthesis and Peptide Incorporation of (S)- α -Amino-(2,2'-bipyridine)-6-Propanoic Acid," *J. Org. Chem.* **1992**, *57*, 757-759.
- (33) Imperiali, B.; Prinns, T. J.; Fisher, S. L. "Chemoenzymatic Synthesis of 2-Amino-3-(2,2'-bipyridinyl)-Propanoic Acids," *J. Org. Chem.* **1993**, *58*, 1613.
- (34) Bax, A.; Davis, D.G. "MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy," *J. Magn. Res.* **1985**, *65*, 355-360.
- (35) Bax, A.; Davis, D. G. "Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy," *J. Magn. Res.* **1985**, *63*, 207-213.
- (36) States, D.J.; Haberkorn, R.A.; Ruben, D.J. "A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants," *J. Magn. Res.* **1982**, *48*, 286-292.
- (37) Piotto, M. ; Saudek, V.; Sklenár, J. "Gradient-Tailored Excitation for Single-Quantum NMR-Spectroscopy for Aqueous-Solutions," *Biomolecular NMR* **1992**, *2*, 661-665.

Chapter 5

Design and evaluation of a metallopeptide regulatory
motif for cAMP-dependent protein kinase

Introduction

Enzymes turnover large amounts of substrate with unsurpassed efficiency and selectivity. Several enzymes operate under the allosteric control of small molecule metabolites, allowing natural systems to respond quickly to slight variations in metabolite levels. This "bio-amplification" could be exploited for the development of a metallo-biosensor. By targeting enzymes which operate on peptidyl substrates, metal dependent control of enzyme activity could be achieved through the creation of a regulatory metallo-peptide of *de novo* design. If the metal binding event could be coupled to the inhibition of the target enzyme then a sensitive detection method could be developed.

The basic approach to protein design adopted by this lab involves the use of defined β -turns and unnatural amino acids to modulate the structural and functional properties of synthetic polypeptides. The previous chapters discussed the use of these concepts in the creation of an independently folded polypeptide motif, the structure of which was *not dependent* on metal ligation. Using these same concepts but in a converse approach the construction of a small peptide the folding of which is metal *dependent*, was undertaken. The architecture of a type II β -turn was used in conjunction with an unnatural, bidentate metal-ligating amino acid to create a multidentate metal binding site within a peptide of *de novo* design. This motif could then be programmed to regulate the activity of enzyme systems that bind peptidyl substrates. Regulatory function would be accomplished through the metal-dependent folding of a enzyme inhibitor into a conformation not recognized by the enzyme. The system was geared towards enzymes which bind peptidyl substrates (and inhibitors) in an extended conformation. The target inhibitor motif was anticipated to adopt (or be able to adopt) an extended conformation in the absence of metal, thereby inhibiting the enzyme (Figure 5-1). Addition of metal to the system would constrain the

inhibitor into an non-extended conformation, thereby down regulating the inhibitory properties of the peptide.

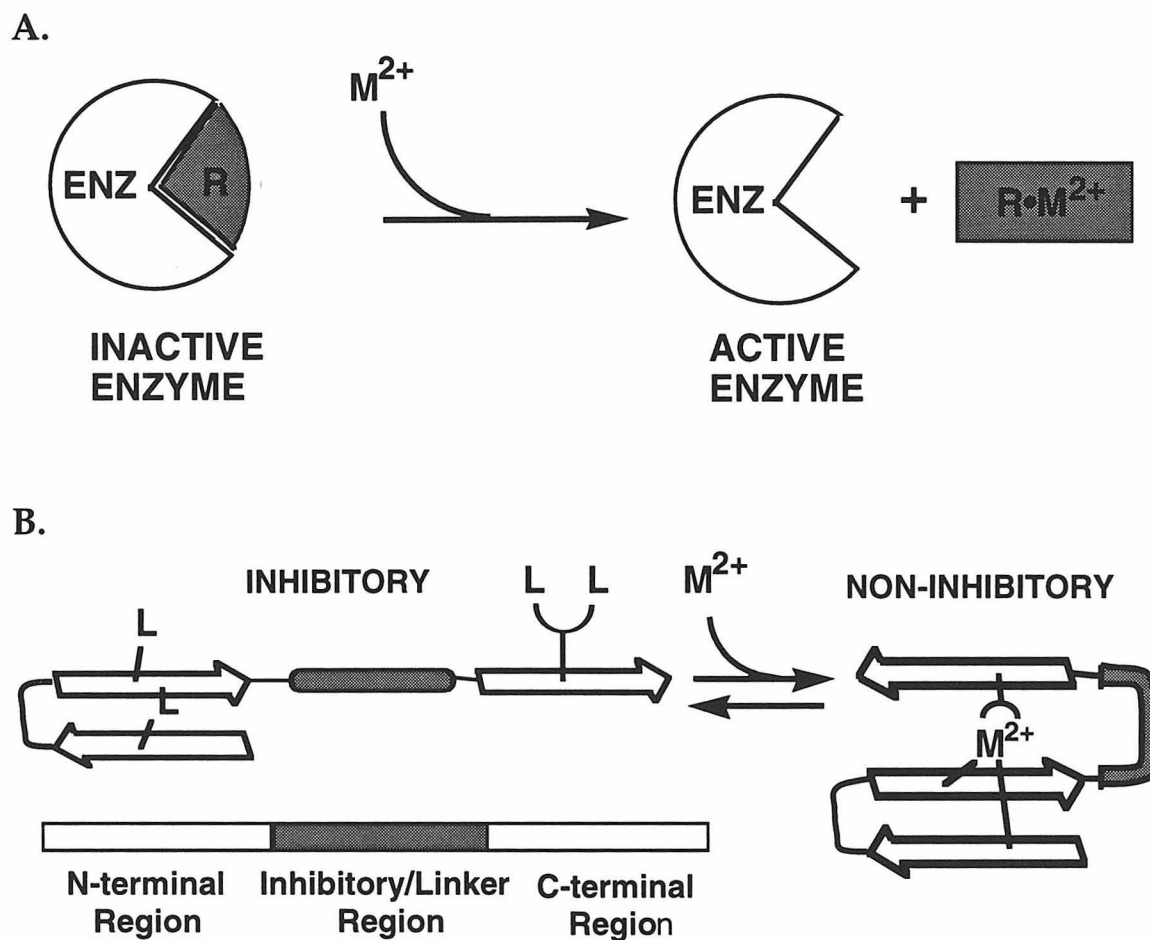


Figure 5-1 Schematic representation of the proposed metal dependent regulatory peptide. **A.** Inhibition of the target enzyme is metal dependent. **B.** The regulatory peptide utilizes a $\beta\beta$ topology for the creation of a multidentate metal ligation site. Metal binding modulates its conformation and inhibitory properties.

The versatility of the peptide architecture enables the utilization of a modular approach in the construction of the desired motif. The ideal motif was envisioned to consist of three modules: A N-terminal bidentate metal binding region, the linker or inhibitory region, and the C-terminal metal binding residue. The N-terminal region would consist of a bidentate site is created from two monodentate ligands through the architecture of a type II β -turn, while a single unnatural metal binding region would supply bidentate functionality at the C-terminus. The linker region would contain an inhibitory sequence for the target enzyme. The regulatory peptide (Figure 5-1B) would therefore be adaptable to a variety enzyme systems and metal ions, by varying the inhibitor linker region or the three amino acids involved in metal ligation. Initial work on this project by Dr. Stewart Fisher in the Imperiali group involved the assembly of the metal ligating "pieces" of the regulatory peptide: synthesis of the unnatural bidentate amino acid and the development of stable type II β -turns incorporating monodentate amino acids.¹ The C-terminal unnatural bidentate metal ligating residue chosen was (*S*)-2- α -amino-3-(2,2'-bipyrid-4-yl) propanoic acid (4Bpa).² This residue provides sterically unconstrained access to the *N,N*-chelation moiety and has been successfully incorporated into peptides (Figure 5-2B).¹ The sequence -Bmc-Pro-DSer-His- (Bmc = β -methyl cysteine) was used to create the N-terminal metal binding region. This sequence has been previously shown to adopt a solution conformation consistent with type II turn formation.¹ The use of the β -branched, unnatural thiol-containing amino acid was necessitated by the steric preferences for a β -branched amino acid at position (*i*) of a type II reverse turn.³ The turn motif was used to place the thiol and imidazole moieties of the Bmc and His residues, respectively, in close proximity and create a metal binding site with high metal binding affinity (Figure 5-2A).

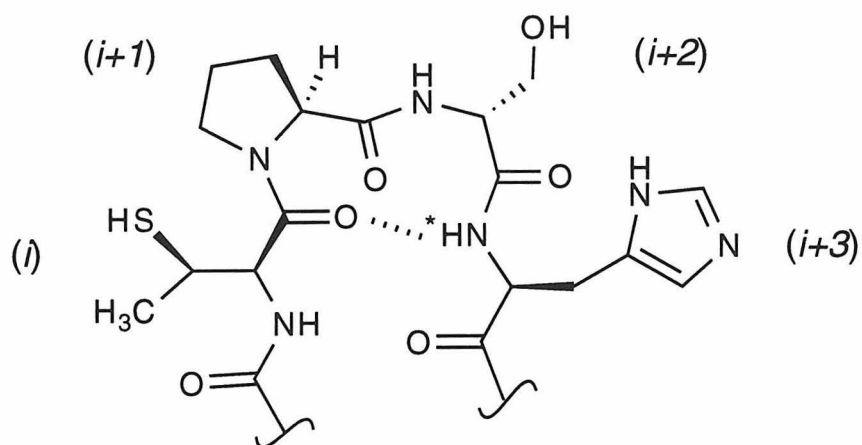
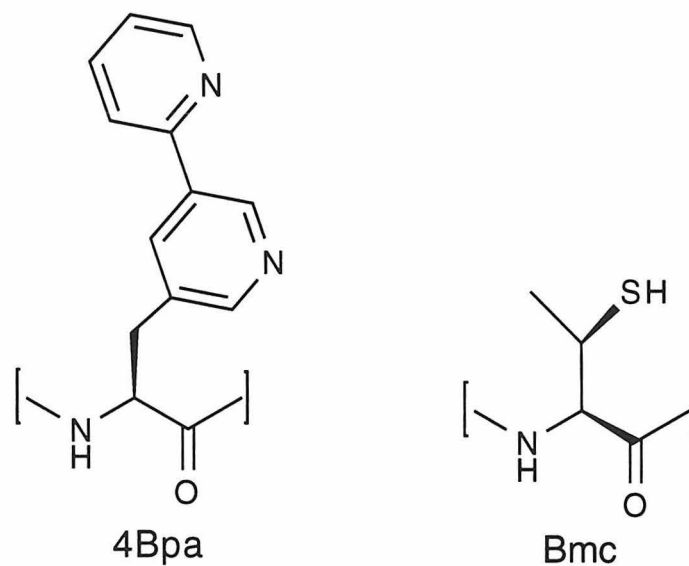
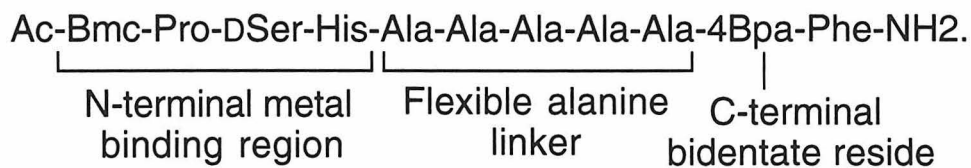
A**B****C****LSW1:**

Figure 5-2 Metal binding regions of the regulatory metallo-peptide design. **A.** The N-terminal metal binding region consisting of β -methyl cysteine at position (*i*) and histidine at position (*i*+3) of a type II β -turn. **B.** The C-terminal metal binding ligand, 4Bpa and the structure of β -methyl cysteine (Bmc). **C.** The sequence of the metal switch prototype **LSW1**.

The feasibility of the proposed motif was demonstrated in the design and synthesis of a prototype for the desired metal dependent regulatory "switch" peptide, **LSW1**. This peptide incorporated the N- and C-terminal metal binding regions (Figure 5-2) connected by a flexible alanine linker connecting the N- and C-terminal metal ligating residues (Figure 5-2 C).

The metal binding properties of **LSW1** were evaluated by Dr. Stewart Fisher to determine if the cooperative formation of a metal binding site involved both the N- and C-terminal ligating residues as desired.¹ Data on metal binding for peptides incorporating 4Bpa is easily obtained by monitoring the changes in the π - π^* transition of the chromophore in the UV spectrum of the bipyridyl moiety.^{1, 4} This peptide binds metals, Co^{2+} , Cu^{2+} , and Zn^{2+} with sub-micromolar dissociation constants and the peptide-metal complex appears monomeric at peptide concentrations below 100 μM .¹ Additional evidence for the participation of the Bmc group in metal ligation within this peptide was provided by the observation of a ligand-to-metal charge transfer band ($\lambda_{\text{max}}=385\text{nm}$) in a UV titration of the peptide with Co^{2+} .¹ A combination of UV and EPR studies suggested peptide **LSW1** formed a high spin octahedral complex involving the Bmc and 4Bpa amino acids. Although no direct evidence for the involvement of the histidine residue could be obtained, it was assumed that this residue was also involved in metal coordination.¹ The prototype peptide **LSW1** demonstrated the feasibility of creating the desired metal site within a peptide. Now it remained to design the linker region to include an inhibitory sequence for an enzyme which has been shown to bind inhibitors and substrates in an extended conformation.

The enzyme system chosen for regulation was cAMP dependent protein kinase (PK). PK catalyses the transfer of a phosphoryl group from ATP to the serine of many protein substrates (Figure 5-3A) and has been well

characterized.^{5,6} The enzyme consists of two associated subunits in a tetrameric protein and is inactive until the binding of cAMP by the regulatory subunit causes the dissociation of the catalytic subunits (Figure 5-3B).^{5,6} The catalytic subunit can be successfully purified and retains activity.⁷ The consensus substrate sequence for phosphorylation comprises five residues. The phosphorylated serine must be separated by one residue from a pair of arginines (Figure 5-3C).⁶ Peptidyl inhibitors for this enzyme are easily generated by the substitution of alanine for the serine in the consensus substrate sequence (Figure 5-3D).⁸

A crystal structure of the catalytic subunit of cAMP dependent protein kinase bound to a peptidyl inhibitor reveals that the enzyme binds inhibitors and substrates in an extended conformation in the region of the consensus sequence.^{9,10} The inhibitor bound to the active site in this structure is a fragment of a very potent natural inhibitor for this enzyme, residues 5-24 of PKI (heat stable protein kinase inhibitor). This fragment exhibits inhibitory constants (K_i) in the nanomolar range.¹¹ If this peptide is truncated further, inhibitors with K_i s ranging from nanomolar to high micromolar can be generated.¹¹

Metal-dependent regulation of this enzyme system presents exciting opportunities for signal amplification. Several enzymes require phosphorylation by cAMP-dependent PK to be active.⁶ The use of a metal-dependent inhibitor of the catalytic subunit of c-AMP dependent protein kinase could be employed in the control of an enzymatic cascade which could lead to significant signal amplification. This metal dependent regulation of this cascade is illustrated schematically in Figure 5-4.

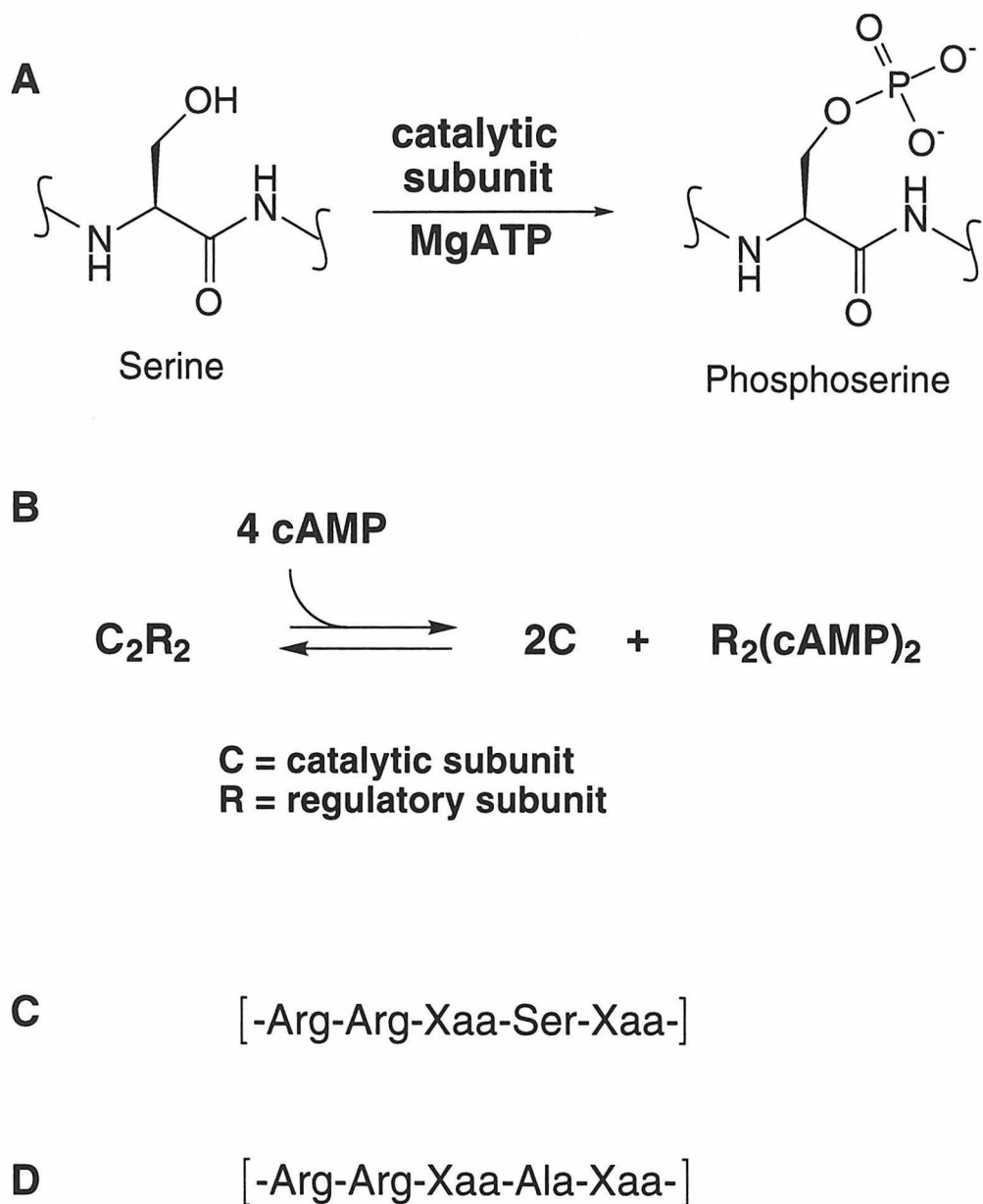


Figure 5-3 **A.** The catalytic subunit catalyses the phosphorylation of a serine residue within a protein. **B.** Binding of the regulatory subunit of cAMP dependent protein kinase releases the catalytic subunit. **C.** The consensus substrate sequence for cAMP dependent protein kinase. **D.** The consensus inhibitor sequence for cAMP dependent protein kinase.

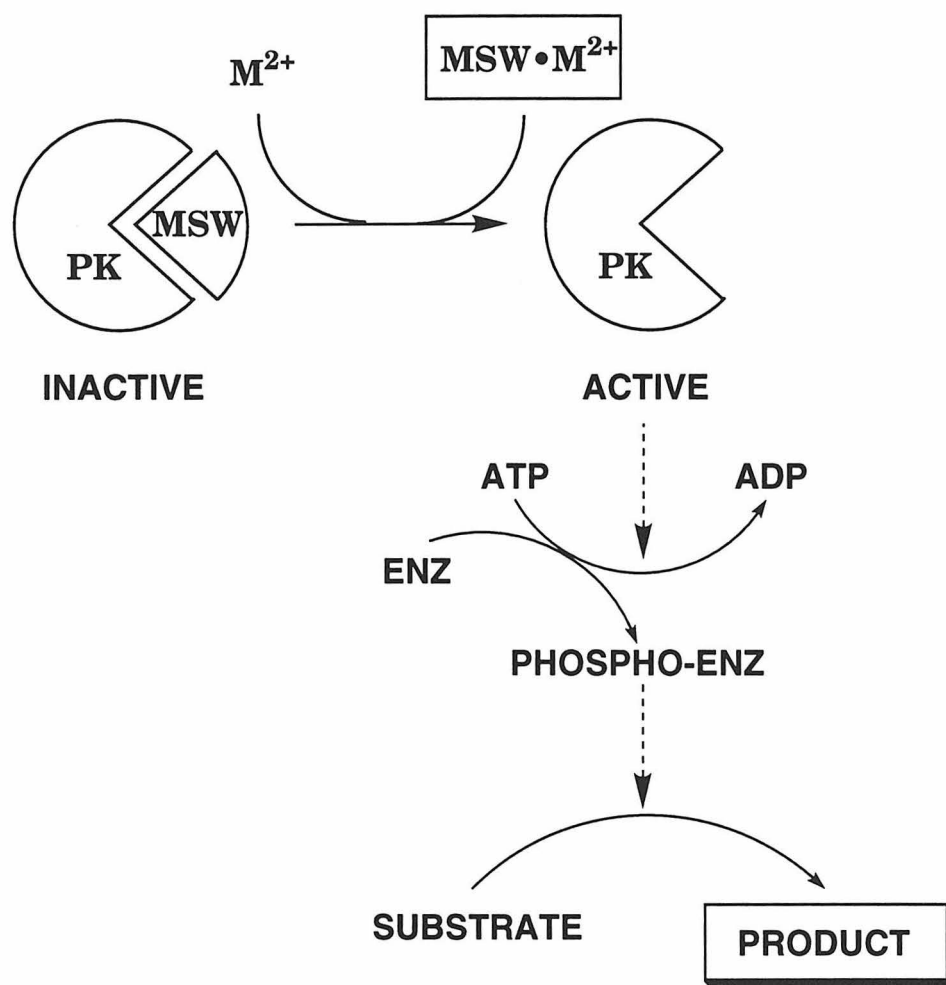


Figure 5-4 Schematic illustration of the regulation of an enzymatic cascade by a designed metalloprotein "metal switch" (MSW) and the catalytic subunit of cAMP dependent protein kinase (PK).

Results

Two different metal binding peptides were synthesized as inhibitors for the catalytic subunit of cAMP-dependent protein kinase (PK). The initial design (**MSW1**) included a seven residue inhibitory linker region between the N- and C-terminal metal binding regions. A side product in the synthesis of **MSW1** was peptide **DMW1** in which the N-terminal Bmc residue was missing. This peptide served as a useful control to gauge the effect of the Bmc residue on the metal binding and inhibitory properties of **MSW1**. Based on the analysis of the metal binding and inhibitory properties of these peptides an additional construct in which the linker region was shortened to five residues was synthesized. These peptides are listed in Table 5-1

Peptide	Sequence
MSW1	Ac-Bmc-P-DS-H- T-G-R-R-N-A-I -4Bpa-F-NH ₂
DMW1	Ac-P-DS-H- T-G-R-R-N-A-I -4Bpa-F-NH ₂
MSW2	Ac-Bmc-P-DS-H- R-R-N-A-I -4Bpa-F-NH ₂
LSW1 ¹	Ac-Bmc-P-DS-H- A-A-A-A-A -4Bpa-F-NH ₂

Table 5-1 The sequences of the peptides assembled and analyzed in the development of a metal dependent inhibitory peptide for cAMP-dependent protein kinase. The inhibitory-linker regions are highlighted.

Design of an Inhibitory Linker Region

The first design of a metal-dependent inhibitor (metal switch) for cAMP-dependent protein kinase employed the metal binding regions previously optimized by Stewart Fisher.¹ The inhibitory (linker) region was chosen on the

basis of data reported for various fragments of the naturally occurring inhibitor PKI.

The prototype peptide **LSW1** exhibited monomeric binding with Zn^{2+} and Co^{2+} at peptide concentrations below $100\ \mu\text{M}$. An inhibition constant for the metal switch in the $1\text{-}20\ \mu\text{M}$ range would allow the use of a inhibitor concentrations below the $100\ \mu\text{M}$ limit for monomeric metal binding, and produce significant inhibition. Therefore, the regulatory peptide should ideally have a inhibition constant (K_i) of approximately $1\text{-}20\ \mu\text{M}$. A 5-8 residue linker was desirable for construction of the metal switch peptide. It was anticipated that a linker of this length would allow for sufficient constraint of peptide conformation upon metal binding to disrupt enzyme inhibition. In this respect, a study by Glass *et al.* determined the inhibition constants of a variety of peptides of diverse lengths based on the inhibitor PKI.¹¹ Although all of the fragments studied were longer than 5-8 residues, the K_i s of shorter fragments could be estimated from the data reported. Based on this analysis a seven residue fragment of PKI (T-G-R-R-N-A-I) was chosen to replace the (Ala)₅ linker used in the **LSW1** prototype. The K_i of this sequence was estimated from the data presented in Table 1 to be between $1\text{-}10\ \mu\text{M}$.

Inhibitor sequence	K _i (nm)
T-T-Y-A-D-F-I-A-S-G-R-T-G- R-R-N-A-I -H-D-NH ₂ (PKI5-24)	2.3
A-S-G-R-T-G- R-R-N-A-I -H-D-NH ₂	119
T-G- R-R-N-A-I -H-D-NH ₂	1250
G-R-T-G- R-R-N-A-I -NH ₂	36
G-R-T-G- R-R-N-A-I -H-D-NH ₂	57
-T-G- R-R-N-A-I -	1-10 μM ?

Table 5-2. The inhibitory constants of fragments of PKI, adapted from Glass et al.¹¹ The consensus inhibition sequence is highlighted.

To test the inhibitory potential of the chosen sequence (-T-G-R-R-N-A-I-) the corresponding acetyl- and amide-capped kinase inhibitor **KI1** (Ac-T-G-R-R-N-A-I-NH₂) was synthesized. The K_i of this peptide for the catalytic subunit of c-AMP dependent protein kinase (PK) was estimated from the fractional inhibition¹² at two different inhibitor concentrations to be $6.7 \pm 1.3 \mu\text{M}$. The value obtained for K_i is within the range (1 - 10 μM) estimated on the basis of values for peptides containing this sequence (see Table 5-2).

The first design of a kinase metal switch peptide (**MSW1**) incorporated this seven residue sequence of **KI1** in place of the five residue alanine linker found in prototype **LSW1**. As in the prototype sequence, the C-terminal metal binding region consisted of a 4Bpa-Phe sequence. These residues replace the His-Asp sequence in the native PKI inhibitor. Since these C-terminal residues are not essential for inhibition by PKI fragments,¹¹ (Table 5-2) and are rather poorly defined in the crystal structure, this substitution was not expected to interfere

with the inhibitory potential of this sequence. The N-terminal region consisted of the two metal binding amino acids Bmc¹ and His³ flanking the type II turn inducing sequence Pro-DSer, as was present in the prototype **LSW1**.

MSW1: Ac-Bmc-Pro-DSer-His-Thr-Gly-Arg-Arg-Asn-Ala-Ile-4Byp-Phe-NH₂

Metal Binding Properties of Peptide MSW1

The metal binding properties of **MSW1** were characterized with two divalent transition metals, Zn²⁺ and Co²⁺. Initial attempts to characterize the binding to Cu²⁺ were complicated by interfering UV absorbance that was attributed to the presence of Cu²⁺ hydroxide complexes. Metal titrations were performed by monitoring the red shift in the π - π^* absorbance of the bipyridinyl moiety of the 4Bpa residue, at pH 7.5, 12 mM KCl (conditions which match those used in the PK assay). In particular the decrease of absorbance at 275 nm and the concurrent appearance of a new maximum at 310 nm indicated metal chelation by the 4Bpa residue (Figure 5-5A).¹ Titration of **MSW1** with ZnCl₂ did not yield well defined isosbestic points, indicating the presence of multiple equilibria in the metal titration (Figure 5-5B).

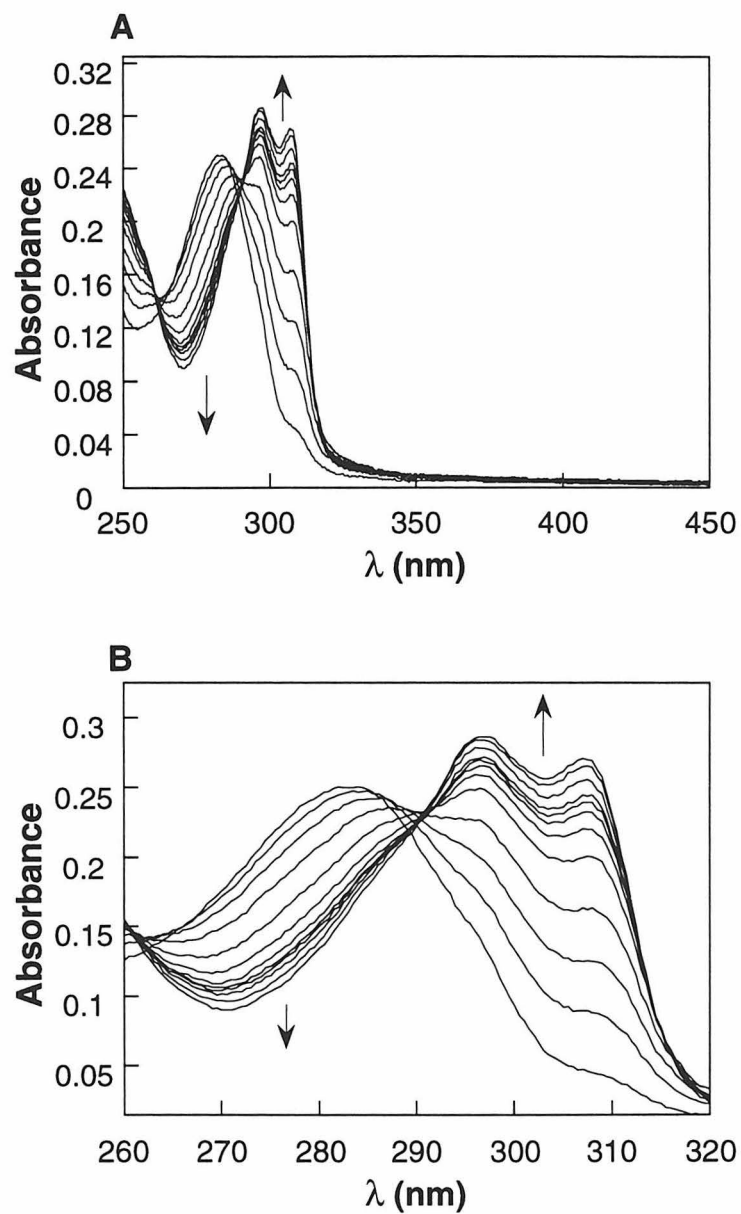


Figure 5-5 Titration of **MSW1** with **ZnCl₂**. **A.** Absorbance spectra for the titration of **MSW1** with **Zn²⁺**. **B.** Expansion of the titration plot illustrating the lack of isosbestic points. (18.2 μ M **MSW1**, 20 mM HEPES, pH 7.5, 12 mM KCl)

Three potentially competing metal binding equilibria are illustrated in Figure 5-6. Equilibrium 1 represents the desired 1:1 complex formation. The second equilibrium (2) represents a net 1:1 stoichiometry due to formation of a 2:2 complex. Formation of the latter would involve dimerization of a pair of peptides binding two metal cations. However, the formation of a 2:2 dimeric complex is unlikely, as electrospray mass spectral analysis of the prototype LSW1 indicated only the formation of a monomeric 1:1 species, and a very small amounts of 2:1 complexes at low peptide concentrations (20 μM). A third possible equilibria involves formation of a 2:1 peptide: Zn^{2+} complex. The presence of multiple species due to competing equilibria (1:1, 2:1, and 2:2) could be responsible for the lack of isosbestic points observed in the Zn^{2+} titration.

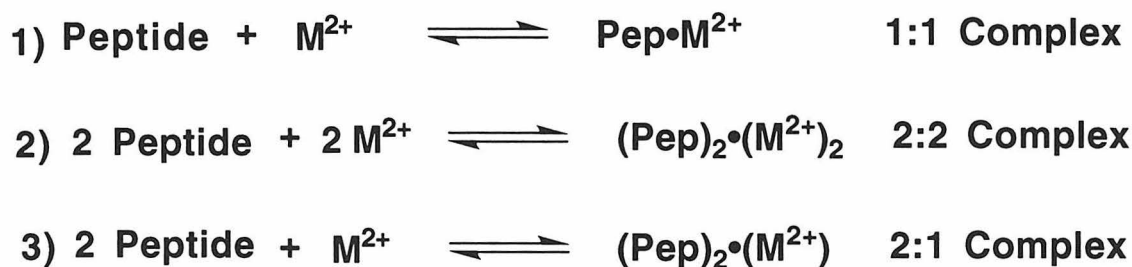


Figure 5-6 Proposed equilibria for the formation of MSW1-metal complexes.

While the binding isotherm indicated that overall stoichiometry of complex formation was 1:1, unusual behavior was noted at low metal concentrations (Figure 5-7). The absorbance of the bipyridyl moiety in the presence of 0.5 equivalents of Zn^{2+} was significantly greater than one half the total absorbance change. This result suggested that a 2:1 peptide: Zn^{2+} complex (Figure 5-6, 3) was formed at low metal concentrations. A 1:1 complex may have

formed upon the addition of an equivalent of metal (Figure 5-6, 1). The lack of isosbestic points in the Zn^{2+} titrations probably resulted from the presence of both 1:1 and 2:1 peptide: Zn^{2+} complexes at sub-equimolar Zn^{2+} concentrations. An estimate of the dissociation constant of **MSW1** for Zn^{2+} was obtained using an iterative method based on the Scott equation¹³ ($K_D = 0.35 \mu\text{M}$). This value is only an estimate since the calculation is based on a two-state equilibrium, which appears to be only partially valid for this peptide.

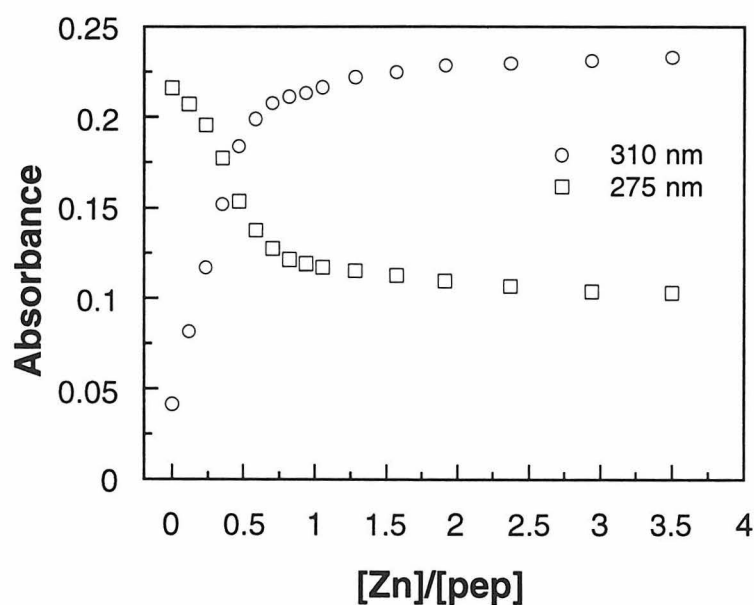


Figure 5-7 Absorbance binding isotherm for the titration of **MSW1** ($18.7 \mu\text{M}$) with ZnCl_2 . (pH 7.5, 20 mM HEPES, 12 mM KCl)

Titration of **MSW1** with CoCl_2 also did not exhibit well defined isosbestic points (Figure 5-8). Unlike the Zn^{2+} titration (Figure 5-5B), the spectra exhibited an increase in the absorbance of the 4Bpa residue at *both* 275 nm and 310 nm, upon the initial additions of Co^{2+} (Figure 5-8). Further additions of Co^{2+} resulted in the expected decrease in absorbance at 275 - 280 nm and a concurrent increase at 310 nm. Additionally, an isosbestic point at 290 nm was observed only *after* the initial addition of 0.15 equivalents. The absorbance at 310 nm was approximately 85% of the saturation level after the addition of 0.5 equivalents of CoCl_2 . These data suggested that the peptide initially formed a 2:1 peptide- Co^{2+} complex which equilibrated to a mixture of 1:1 and 2:2 peptide- Co^{2+} complexes at higher metal concentrations.

A weak broad absorbance (340 - 380 nm) was observed in the Co^{2+} titration of **MSW1** (Figure 5-9). A similar band ($\lambda_{\text{max}} = 385 \text{ nm}$) was observed in the Co^{2+} titration of the prototype **LSW1**, and was attributed to a ligand-to-metal charge transfer (LMCT) band between Co^{2+} and the thiol ligand (Bmc¹).¹ The binding isotherms at this wavelength resembled those obtained by monitoring the red-shifted bipyridyl absorbance at 310 nm (Figure 5-9). In contrast, no absorbance at this wavelength was detected in the absorbance spectrum of the **MSW1-Zn**²⁺ complex (Figure 5-10). This result provides direct evidence for the participation of Bmc in the coordination sphere of Co^{2+} , bound to peptide **MSW1**. These data also suggested that the peptide forms a 2:1 complex, which involved *both* the 4Bpa and the Bmc residues. This complex may have equilibrated to a 1:1 complex as multiple equivalents of Co^{2+} levels were added. A 2:1 peptide: Co^{2+} complex involving both the bipyridyl (4Bpa) and thiol (Bmc) functionalities would have a saturated coordination sphere, eliminating the possibility of the participation of histidine in complex formation.

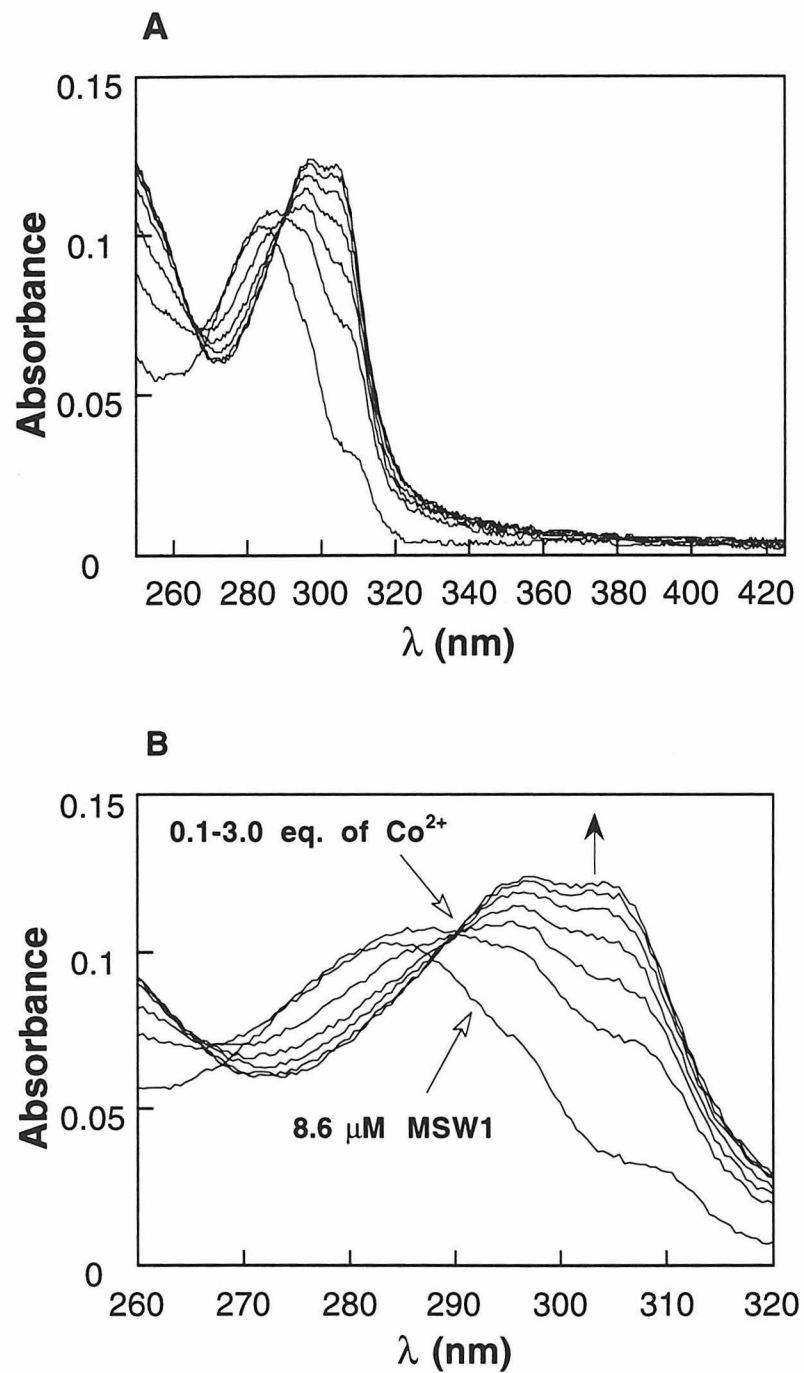
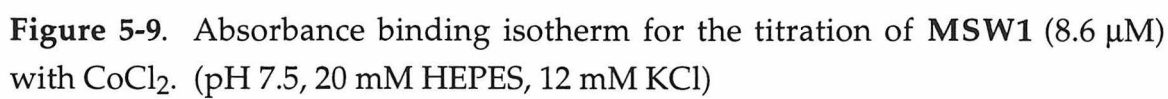


Figure 5-8 A. Absorbance spectra of the titration of **MSW1** with CoCl_2 . B. Expansion plot of the titration of **MSW1** (8.6 μM) with CoCl_2 . (pH 7.5, 20 mM HEPES, 12mM KCl)



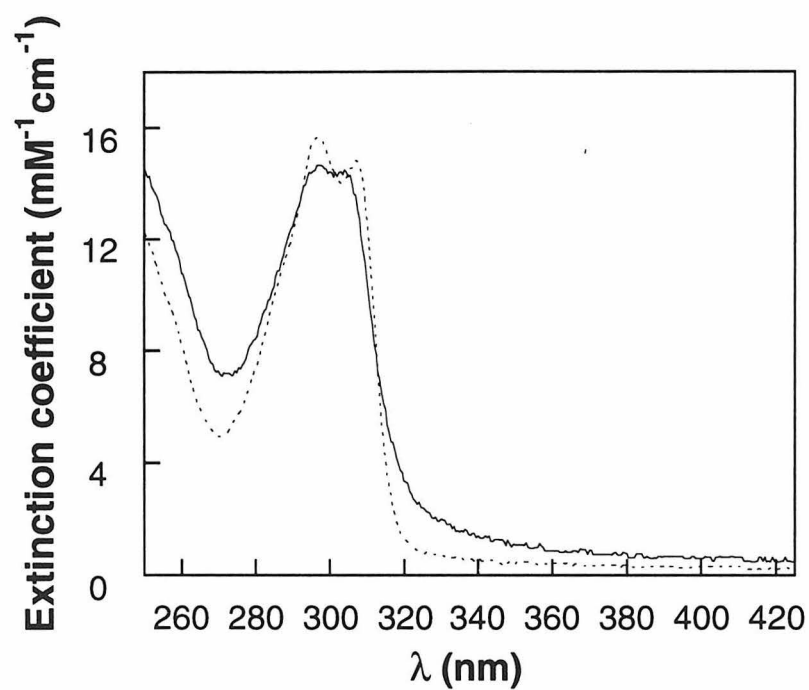


Figure 5-10 Molar extinction coefficient spectra of the **MSW1-Co²⁺** (solid line) and **MSW1-Zn²⁺** (dashed line) complexes. (3 equivalents of ZnCl₂ or CoCl₂, pH 7.5, 20 mM HEPES, 12 mM KCl)

Characterization of Metal Binding Properties of Peptide DMW1

Poor coupling of the β -branched unnatural amino acid during the synthesis of **MSW1**, resulted in the production of a N-terminal deletion peptide **DMW1**. Mass spectral analysis confirmed that the N-terminal Bmc residue was not present in this peptide. **DMW1**, therefore, contained only the 4Bpa and His metal ligating residues. Peptide **DMW1** therefore served as a fortuitous control peptide to determine the specific effects of the Bmc residue on the metal binding properties of **MSW1**.

DMW1: Ac-Pro-DSer-His-Thr-Gly-Arg-Arg-Asn-Ile-4Bpa-Phe-NH₂

Unlike the multiple equilibria observed for **MSW1**, Zn²⁺ and Co²⁺ titrations of **DMW1** exhibited well defined isosbestic points indicative of a two-state equilibrium. Binding isotherms for both metal titrations suggested 1:1 complex formation (Figure 5-11 and 5-12). Notably, the Co²⁺ titration of peptide **DMW1** did not exhibit the broad absorbance band in the 340 - 380 nm range observed for the **MSW1**-Co²⁺ complex (Figure 5-12). The extinction coefficient for **DMW1** at 340 nm was independent of Co²⁺ concentration ($\epsilon \sim 500 \text{ M}^{-1}\text{cm}^{-1}$, Figure 5-13). In contrast, the extinction coefficient of **MSW1** at 340 nm increased from $\sim 500 \text{ M}^{-1}\text{cm}^{-1}$ to $1250 \text{ M}^{-1}\text{cm}^{-1}$ upon formation of the **MSW1**-Co²⁺ complex, which further supported the assignment of this absorbance to a thiol LMCT band. Dissociation constants were determined for both Zn²⁺ ($K_D = 10 \text{ }\mu\text{M}$) and Co²⁺ binding ($K_D = 5.5 \text{ }\mu\text{M}$).

The values for the dissociation constants of the **DMW1**-Zn²⁺ and Co²⁺ complexes were very similar to those previously obtained for peptides which contained the 4Bpa residue as the only metal ligation functionality.¹ This comparison implied that metal binding in peptide **DMW1** did not involve coordination of the N-terminal histidine residue.

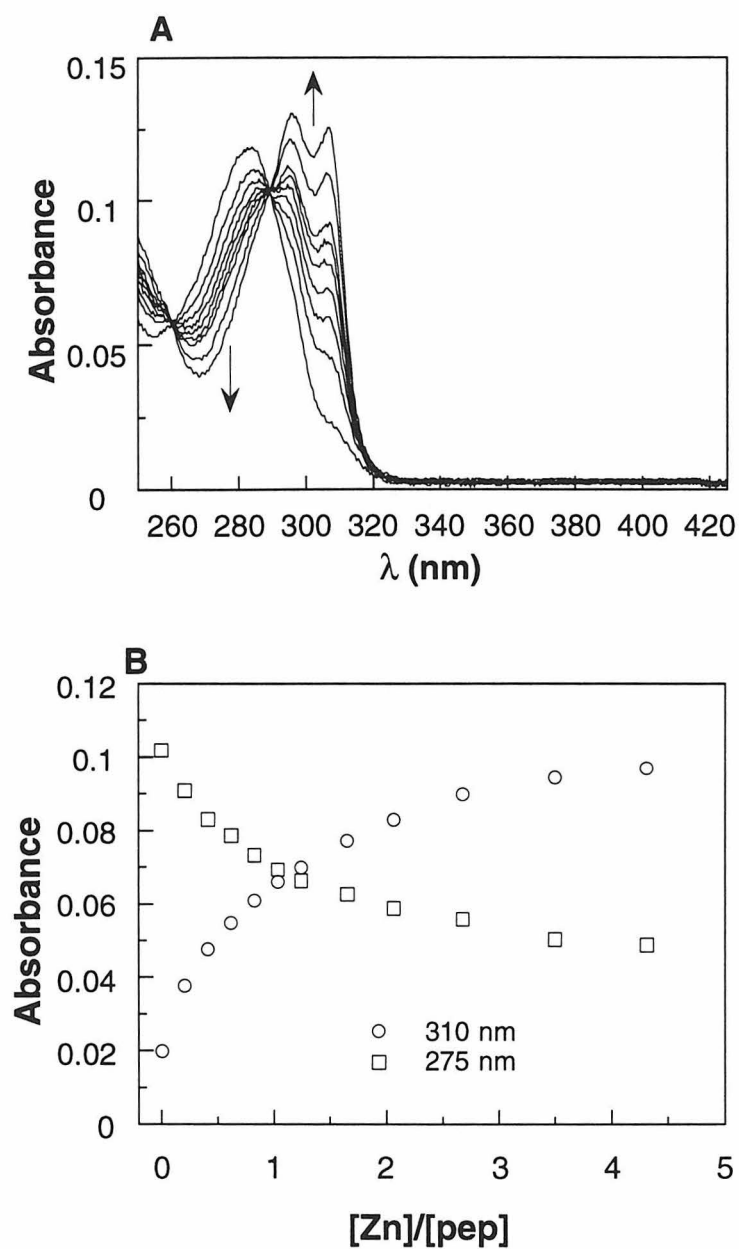


Figure 5-11 A. Absorbance spectra for the titration of **DMW1** with ZnCl₂. B. Binding isotherm for the titration of **DMW1** (9 μ M) with ZnCl₂. (pH 7.5, 20 mM HEPES, 12mM KCl)

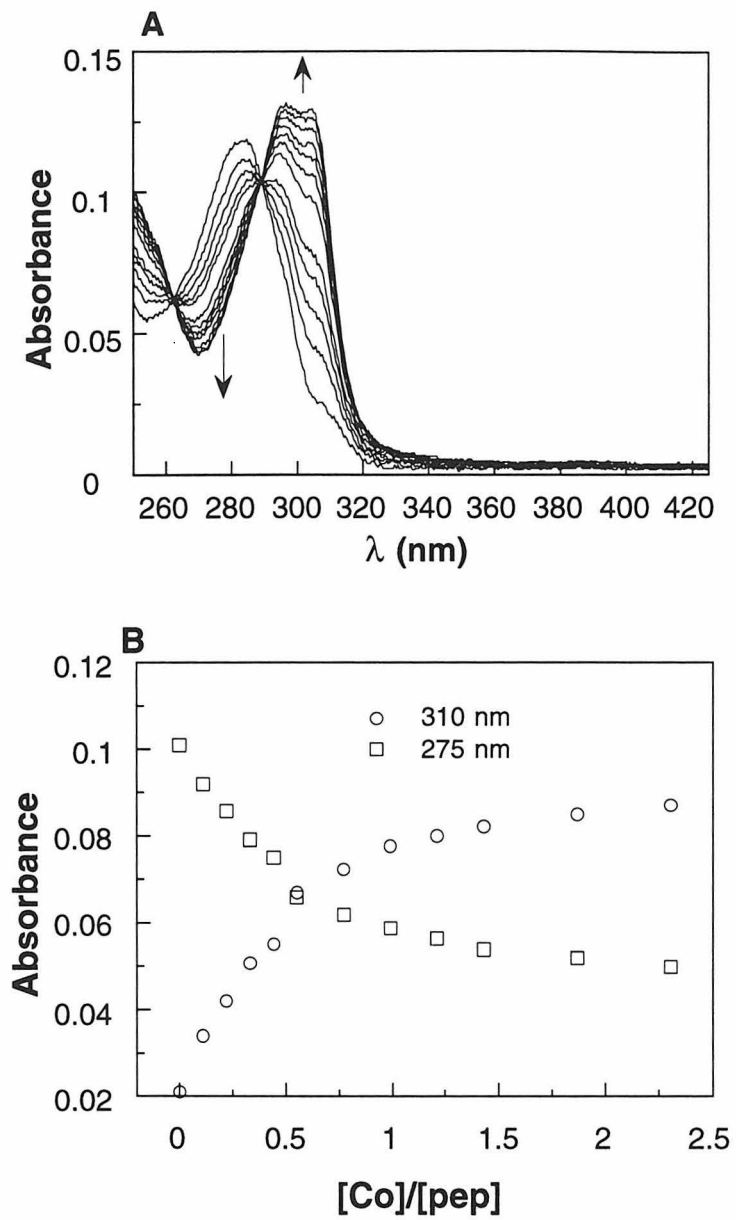


Figure 5-12 A. Absorbance spectra for the titration of **DMW1** with CoCl_2 . B. Binding isotherm for the titration of **DMW1** ($9\mu\text{M}$) with CoCl_2 . (pH 7.5, 20 mM HEPES, 12mM KCl)

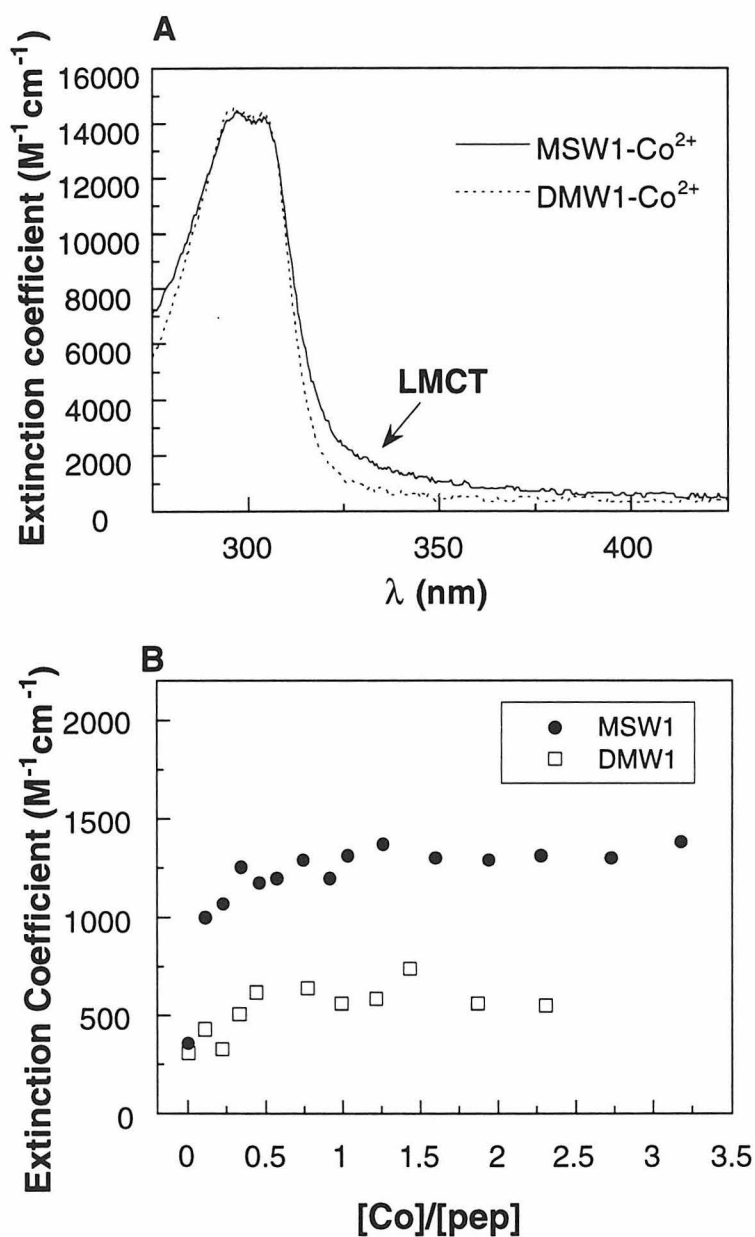


Figure 5-13 **A.** Comparison of the extinction coefficient spectra of the **MSW1- Co^{2+}** complex (8.6 μM **MSW1**, 3 eq. $CoCl_2$) and the **DMW1- Co^{2+}** complex (9.0 μM **DMW1**, 2.8 eq. $CoCl_2$). **B.** Isotherms at 340 nm (LMCT band) for titrations of **MSW1** and **DMW1** with $CoCl_2$. (pH 7.5, 20 mM HEPES, 12 mM KCl).

Characterization of the Metal Binding Properties of MSW2

In the second metal switch design (**MSW2**), the linker region was shortened by two residues to include only the essential five residue consensus sequence (R-R-X-A-X, Figure 5-3 D). The asparagine and isoleucine residues of the PKI sequence were retained for optimal inhibitory capacity.

MSW2: Ac-Bmc-Pro-DSer-His-Arg-Arg-Asn-Ala-Ile-4Bpa-Ile-NH₂

The metal binding characteristics of this peptide more closely match those observed for the prototype **LSW1**, which also contained a five residue linker.¹ Titrations of **MSW2** at low peptide concentrations (10-12 μM) with Co^{2+} indicated a two-state cooperative equilibrium with a dissociation constant of 11 nM (Figure 5-14). Binding isotherms at both 310 nm (bipyridyl moiety from Bpa4) and 340 nm (LMCT band from Bmc) indicated a 1:1 stoichiometry. In contrast, CoCl_2 titrations of **MSW2** at peptide concentrations above 40 μM revealed a cooperative two-state equilibrium with a stoichiometry of 2:1 peptide-metal (Figure 5-15). Isotherms obtained from absorbances at *both* the 310 nm (4Bpa) and 340 nm (LMCT) were consistent with this stoichiometry.

Similar stoichiometries, which were dependent on peptide concentration, were observed in the Zn^{2+} titrations (Figure 5-16), except that a slight deviation from isosbestic behavior was observed at lower **MSW2** concentrations (7-9 μM). The dissociation constant for the 1:1 Zn^{2+} -complex (peptide concentrations < 40 μM) was calculated to be 0.19 μM .

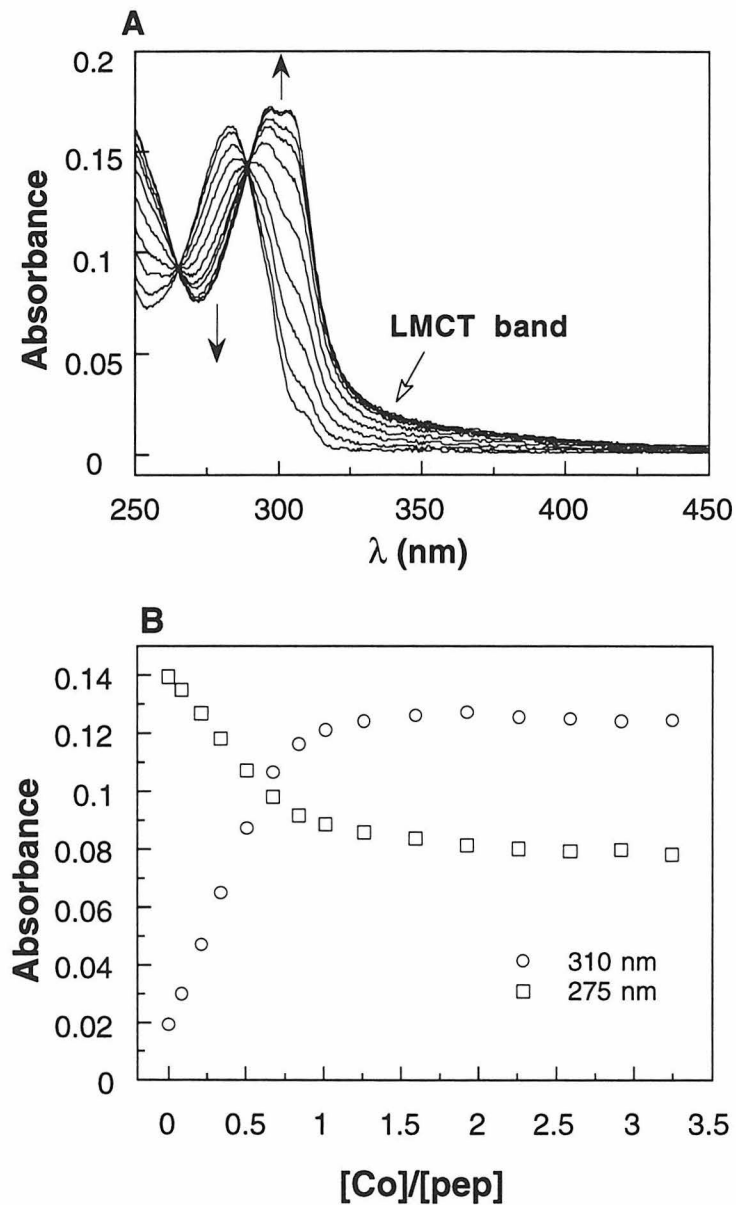


Figure 5-14 **A.** Absorbance spectra of the titration of **MSW2** (12 μ M) with CoCl_2 . **B.** Binding isotherm for the titration of **MSW2** with CoCl_2 . (pH 7.5 20 mM HEPES, 12 mM KCl)

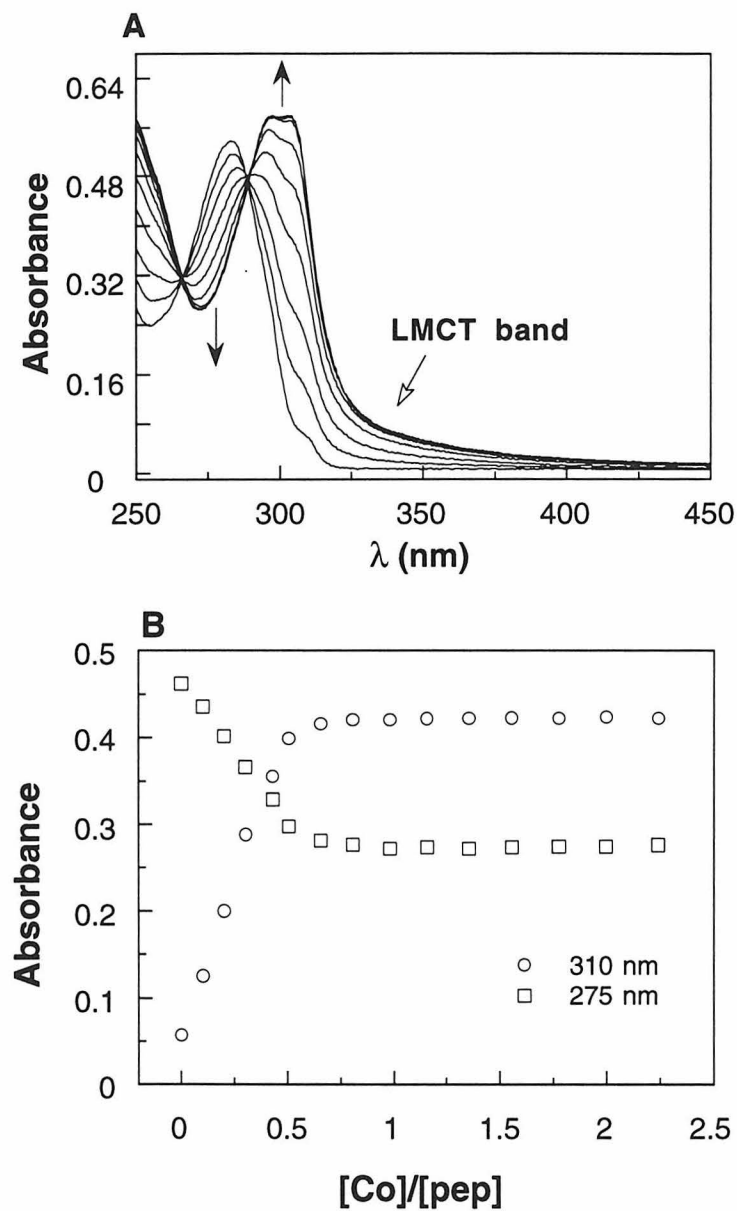


Figure 5-15 A. Absorbance spectra of the titration of **MSW2** (41 μM) with CoCl_2 . B. Binding isotherm for the titration of **MSW2** with CoCl_2 . (pH 7.5 20 mM HEPES, 12 mM KCl)

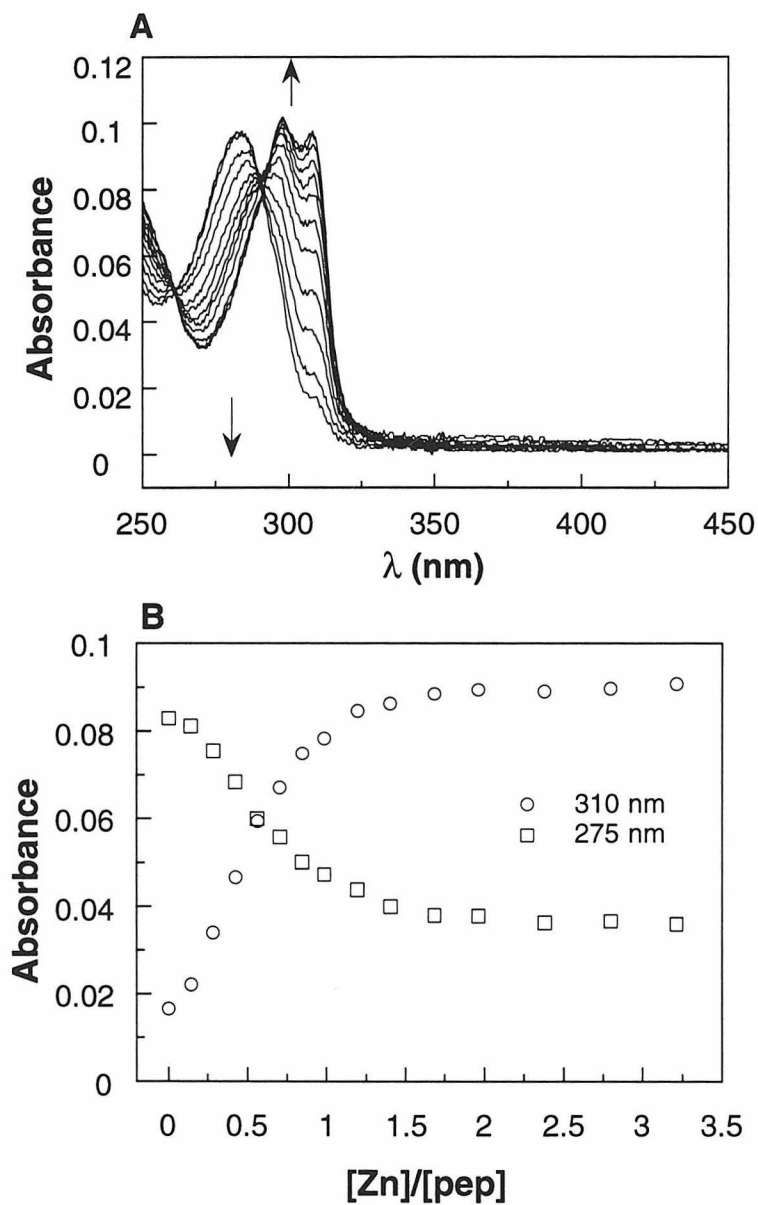


Figure 5-16 A. Absorbance spectra of the titration of **MSW2** (7 μM) with ZnCl_2 . B. Binding isotherm for the titration of **MSW2** with ZnCl_2 . (pH 7.5 20 mM HEPES, 12 mM KCl)

The pH dependence of the binding of Co^{2+} to **MSW2** was determined to investigate the role of the His residue in metal coordination. If the histidine amino acid was involved in the 1:1 complex formation observed at low peptide concentrations then higher dissociation constants could be expected at acidic pH. Protonation of the His imidazole ring at low pH would affect participation in metal binding and the affinity of the peptide for metal ions should decrease. Titration data was acquired at pH 8.0, 7.5, 7.0, and 6.0 and is summarized in Table 5-3.

pH	Peptide Concentration (μM)	Metal complex Stoichiometry	MSW2 :metal complex K_D (μM)
8.0	9.7	2:1	ND
7.5	8.1	1:1	0.01
7.0	8.9	1:1	0.09
6.0	9.6	1:1	1.4

Table 5-3 The pH dependence of the **MSW1**- Co^{2+} dissociation constant.

Co^{2+} binding by **MSW2** did exhibit a definite pH dependence, with a significant decrease in affinity at lower pH. The 15-fold decrease in affinity of **MSW2** for Co^{2+} at pH 6.0 versus pH 7.0, implies the involvement of an ionizable group in with a pK_a in this range. In this respect, the pK_a of Ac-His-NHMe is 6.5¹ which suggested that His could participate in complex formation. Bipyridine has a pK_a of approximately 4.5¹⁴ and the pK_a of the thiol of the β -methyl cysteine can be estimated to be ~ 8.5 (based on the pK_a of cysteine).¹⁵ While suggestive of participation of His in the coordination sphere of Co^{2+} , these data are not conclusive due to the presence of other ionizable ligands (4Bpa and Bmc). Interestingly, the stoichiometry of the Co^{2+} complex at higher pH appears

to be 2:1, potentially reflecting the greater affinity of the thiolate form of β -methyl cysteine.

Metal binding by **MSW1** and **MSW2** was anticipated to result in a significant change in the conformation of the linker/inhibitory region. Metal titrations were performed to determine the nature of the metal complex formed. In summary, at peptide concentrations below 30 μ M peptide and at least one equivalent of CoCl_2 , the metal complex of **MSW2** appeared to be monomeric with coordination by the thiolate ligand (Bmc^1) and the bipyridyl functionality (4Bpa^{10}). As discussed earlier, the identity of the Co^{2+} complex of **MSW1** is less certain, although titration data suggested that complex formation involved both the Bmc and 4Bpa residues. Although the participation of histidine in metal coordination could not be conclusively verified for either peptide, pH studies suggested involvement of this ligand in Co^{2+} chelation for monomeric complexes of **MSW2**. At higher peptide concentrations and one equivalent or less of Co^{2+} , the metal complex of **MSW2** was dimeric. In this case the coordination sphere of the metal ion involved both the Bmc and 4Bpa residues, which precluded histidine from participating in metal ligation.

Assay for Protein Kinase Activity

In order to test the inhibitory properties of the regulatory peptide designs described in previous sections, several assay systems were investigated. For ease and simplicity, the purified catalytic subunit for cAMP protein kinase (PK) from bovine heart,¹⁶ which is commercially available was chosen as the enzyme source. The abbreviation PK is used hereafter to refer to only the catalytic subunit of this enzyme.

The most common assays for kinase activity monitor the transfer of a radioactive phosphate from γ - ^{32}P -ATP to the peptide.¹⁷⁻¹⁹ Although these

assays are typically very sensitive and fast, a non-radioactive continuous assay method was desired. Therefore, the continuous fluorescent-based assay for kinase activity which was reported by Wright *et al.* was therefore initially investigated.²⁰ In this assay, the phosphorylation of serine in the sequence of the kinase substrate Leu-Arg-Arg-Trp-Ser-Leu-Gly modulates the fluorescence of the adjacent tryptophan amino acids resulting in an increase in the fluorescence emission intensity of the tryptophan. This substrate sequence (**KS1**) was synthesized and this assay was tested with commercially available PK. Unfortunately, treatment of **KS1** with the PK under assay conditions similar to those reported²⁰ resulted in a *decrease* in the fluorescence of the tryptophan group. Additionally while HPLC analysis of the assay mixture suggested that the change in the signal resulted from phosphorylation of **KS1**, the fluorescent analysis did not yield reproducible results. Therefore an alternative method of monitoring phosphoryl transfer was sought.

The phosphorylation of a peptide introduces a negative charge which should alter the retention time in C₁₈ reverse phase HPLC analysis. Several HPLC based assays for other kinases (tyrosine protein kinase²¹ and myosin light chain kinase²²) based on this principle have been reported. Initial velocities can be determined by quantifying the substrate and product peaks. An HPLC based assay was therefore investigated as an alternative to the above assay. Using peptide substrate **KS1** a similar assay for cAMP kinase was developed but with limited success. Despite repeated attempts at optimization, velocities for identical runs varied as much as 10-20%. Additionally, the analysis of inhibitory peptides with similar sequences to the substrate peptide **KS1** was hindered by overlap of peaks corresponding to the inhibitors with substrate or product peaks in the HPLC chromatograms. Therefore a radioactive assay was re-investigated.^{18, 19} In this system the transfer of a radioactive phosphate group

from [γ - ^{32}P]-ATP to the substrate is monitored. The radiolabeled phosphorylated product is isolated from excess [γ - ^{32}P]-ATP by absorption of the peptidyl product containing two arginines to phosphocellulose paper. Phosphoric acid is used to quench the enzyme reaction and remove excess [γ - ^{32}P]-ATP from the phosphocellulose paper.¹⁹ The substrate used for these assays was "Kemptide" ($\text{NH}_2\text{-L-R-R-A-L-G-COOH}$) which is commercially available. Substrate turnover was quantified by scintillation counting of the radiolabeled product immobilized on the phosphocellulose paper.¹⁹

Several modifications to the standard assay protocols were required to accommodate the analysis of inhibition by peptide-metal complexes (see Experimental section for details). The concentration of Mg^{2+} was reduced from 3 mM to 120 μM (1.2 equivalents with respect to ATP) in the assay mixture to reduce competitive binding of Mg^{2+} to the designed regulatory peptides. Additionally the enzyme was reconstituted using the minimum concentration (20 μM in the assay mixture) of a thiol reducing agent (β -mercaptoethanol, BME) in order to minimize the binding of divalent transition metal ions. Generally a dithiol (dithiothreitol) is used, however it was found to complex divalent transition ions.²³

Characterization of the Inhibitory Properties of MSW1, DMW1, MSW2 with PK

Investigation of the inherent effects of the divalent transition metal cations Co^{2+} , Zn^{2+} , and Cu^{2+} on the activity of the catalytic subunit of c-AMP dependent protein kinase (PK) revealed that both Zn^{2+} and Cu^{2+} had a significant inhibitory effect at the concentrations required for metal binding by the designed peptides (20 - 100 μM). For this reason, all reported inhibition studies were performed on the Co^{2+} complexes of **MSW1**, **MSW2**, and **DMW1**. Analysis of the inhibitory properties of all three peptides (Table 5-1) both uncomplexed and in the presence

of Co^{2+} were performed. Peptide **DMW1** acted as a control for these experiments. Metal titrations of this peptide suggested that metal binding was mediated exclusively through the 4Bpa functionality. The conformation of the linker region was therefore anticipated to be unaffected by metal ligation, and the inhibitory properties of this peptide were expected to be independent of Co^{2+} binding.

The approximate inhibitory constants for peptides **MSW1**, **MSW2**, and **DMW1** (calculated as described for **KI1**) were estimated from the fractional inhibition in PK assays with and without the addition of CoCl_2 . While this method¹² of calculating K_i has limited accuracy, it does reveal general trends in the *relative* potencies of inhibitors. The results of these experiments are summarized in Table 5-4. Inhibition by the **MSW1** and **MSW2** was metal dependent but not in the manner expected. Rather than diminishing inhibition, formation of the Co^{2+} complexes of these peptides resulted in *better* inhibitors of PK. The calculated K_i of the **MSW2**- Co^{2+} complex was three to four fold lower than the K_i of the uncomplexed peptide, for both the 1:1 and 2:1 peptide-metal complexes. For experiments performed on **MSW1** there was only a 2-fold decrease in K_i upon metal binding. Significantly, inhibition by **DMW1** was *not* metal dependent.

Peptide	Stoichiometry	concentration	equivalents of Co ²⁺	K _i (μM)
MSW1		12, 31 μM	0	1.2 (± 0.4)
MSW1•Co²⁺	1:1	12 μM	1	0.6
	^a	31 μM	1	0.6
MSW2		10-44 μM	0	2.9 (± 0.9)
MSW2•Co²⁺	2:1	42, 44 μM	1.0	0.6 (± 0.3) ^b
	2:1	42, 44 μM	0.5	0.9 (± 0.2) ^b
	1:1	11 μM	1.0	0.9
	1:1	24 μM	1.4	0.9
M2W		10 μM		0.86
M2W•Co²⁺	1:1	10 μM	3.5	0.83

Table 5-4 Summary of the results of PK inhibition by metal switch peptides. Average values determined from multiple experiments are listed if appropriate with the standard deviation or the range given in parentheses. ^a Stoichiometry assumed to be 1:1. ^b Average of two experiments.

PKCYS: Inhibitor Cyclization With a Disulfide

Analysis of the inhibitory properties of the designed metal switch peptides indicated that cyclization *via* metal ligation actually produced a *better* inhibitor. This result was unexpected based on the extended binding of the inhibitor PKI(5-24) in the active site of PK in the crystal structure of the enzyme-inhibitor complex.¹⁰ In this regard, some ambiguity about the conformation of the linker region in the Co²⁺ complexes of the designed metal switch peptides existed, due to uncertainties regarding the exact nature of the metal complex and the involvement of the histidine in the coordination sphere. As an alternative to the metal mediated cyclization presented in peptides **MSW1** and **MSW2**, a disulfide could serve as an unambiguous covalent constraint. A peptide containing the five residue inhibitory sequence of **MSW2** flanked by two

cysteines was therefore synthesized (**PKCYS**). Formation of a disulfide bond between these two cysteines produced a peptide with a well defined covalent constraint between the N- and C-termini. HPLC analysis revealed that the oxidized peptide was not reduced under the conditions of the assay. Inhibition assays of both the reduced (**PKCYS-R**) and oxidized (**PKCYS-O**) forms of the peptide helped determine the amenability of the PK system to modulation of inhibition through inhibitor cyclization.

PKCYS-R: Ac-Cys-Arg-Arg-Asn-Ala-Ile-Cys-NH₂

PKCYS-O: Ac-Cys-Arg-Arg-Asn-Ala-Ile-Cys-NH₂

$$\begin{array}{c} | \qquad \qquad \qquad | \\ \text{S} \text{-----} \text{S} \end{array}$$

These studies indicated that cyclization of **PKCYS** to the disulfide *decreased* the affinity of the inhibitor for PK, although the effect was not large. The calculated inhibition constant for the oxidized peptide (**PKCYS-O**) was approximately 2.5-fold higher than that of the reduced form (**PKCYS-R**) (Table 5-5). This trend is the opposite of that seen for the metal switch peptides: cyclization of peptides **MSW1** and **MSW2** through metal ligation produced inhibitors with *greater* affinity for PK. .

Peptide	K _i (μM)
PKCYS-R	5.7 (± 1.9)
PKCYS-O	2.2 (± 0.4)

Table 5-5 Inhibition constants for the oxidized (**PKCYS-O**) and the reduced forms (**PKCYS-R**) of the inhibitor **PKCYS**.

Discussion

Based on the prototype **LSW1** developed by Dr. Stewart Fisher, two metal switch peptides were designed as regulatory units for c-AMP dependent protein kinase. Both designs included identical metal binding residues; metal ligation was anticipated to occur *via* cooperative coordination of a bidentate unnatural amino acid, 4Bpa, at the C-terminus and two monodentate residues, His and Bmc at the N-terminus of the peptide. The metal binding properties of these peptides were dependent on the linker spacing between the N- and C-terminal metal binding regions. **MSW1** incorporated a seven residue linker and metal binding by this peptide involved multiple equilibria probably resulting from a mixture of different peptide-metal stoichiometries. In contrast, metal titrations of **MSW2**, which included a five residue linker, exhibited well defined isosbestic points and defined peptide-metal stoichiometries. The stoichiometry of both Zn²⁺ and Co²⁺ complex formation with **MSW2** was dependent on peptide concentration; a 2:1 peptide-metal complex was formed at peptide concentrations greater than 40 μM peptide while a 1:1 stoichiometry was observed at lower peptide concentrations (5-20 μM). The formation of dimeric metal complexes at higher concentrations of **MSW2**, suggests that the multiple equilibria observed in the metal titrations of

peptide **MSW1** may have resulted from the formation of a mixture of monomeric and dimeric species in this system.

UV-Vis absorbance spectroscopy confirmed the participation of both the 4Bpa and Bmc ligands in metal coordination by **MSW1** and **MSW2**. Metal binding by the 4Bpa residue is indicated by the red shift of the bipyridyl $\pi\text{-}\pi^*$ absorbance upon metal addition ($\lambda_{\text{max}} = 305 \text{ nm}$). Additionally the observation of a broad absorbance band (340-380 nm) in the spectrum of the Co^{2+} complex of both peptides provided direct evidence for the participation of the side chain functionality of the β -methyl cysteine amino acid in complex formation (Figure 5-15A). This absorbance band was not observed in either the Zn^{2+} complexes of **MSW1** and **MSW2**, or in the Co^{2+} complex of **DMW1**, which does not contain the Bmc residue (compare Figure 5-14A and Figure 5-12A). Sulfur ligands bound to Co^{2+} can undergo ligand-to-metal charge transfer (LMCT) processes which typically produce absorbance bands in the near UV region (300-600 nm).²⁴ The broad absorbance present in the Co^{2+} complexes of **MSW1** and **MSW2** was therefore attributed to a LMCT band between the thiol functionality of β -methyl cysteine and Co^{2+} . A similar band was observed for the prototype **LSW1**.¹ The LMCT transfer band exhibited a dependence on metal cation concentration similar to that observed for the $\pi\text{-}\pi^*$ absorbance of the bipyridyl chromophore. The observation of both the $\pi\text{-}\pi^*$ and LMCT bands in CoCl_2 titrations of **MSW2** which also exhibited well defined isosbestic points suggested cooperative formation of a Co^{2+} complex involving both Bmc and 4Bpa in 2:1 and 1:1 **MSW2**- Co^{2+} complexes. This conclusion is supported by comparison of the dissociation constants for this peptide with a model peptide which contains 4Bpa as the only metal ligating amino acid. The dissociation constants of **MSW2** complexes with Co^{2+} and Zn^{2+} were 100- to 200-fold lower than those values for the model peptide (Table 5-6). The enhanced affinity of **MSW2** for Co^{2+} and Zn^{2+} can be

attributed to the cooperative participation of the auxiliary ligands supplied by the β -methyl cysteine or histidine residues in metal ligation.

Peptide	K_D (M)	
	Zn^{2+}	Co^{2+}
Peptide MSW2 ^a	1.9×10^{-7}	1.1×10^{-8}
	(1:1)	(1:1)
Ac-4Bpa-Thr-Pro-DSer-Val-Phe-NH ₂ ^b	22×10^{-6}	3.8×10^{-6}
	(1:1)	(2:1) ^c

Table 5-6 Comparison of the dissociation constants and stoichiometries for metal complexes of **MSW2** and Ac-4Bpa-Thr-Pro-DSer-Val-Phe-NH₂. ^a K_D was measured at pH 7.5 and at **MSW2** concentrations below 10 μ M. ^b K_D was measured at pH 8.25. ^c Calculation of K_D assumed a 1:1 stoichiometry.

Although clear evidence indicated the participation of Bmc¹ and 4Bpa¹⁰ in metal coordination, the involvement of His⁴ metal ligation for 1:1 **MSW2**-Co²⁺ complexes was more difficult to prove. Investigation of the pH dependence of the K_D of the 1:1 **MSW2**-Co²⁺ complex suggested the involvement of His⁴, but these experiments were not conclusive.

Unlike the 1:1 peptide-metal complex, there is no ambiguity concerning the ligands participating in complex formation in the 2:1 **MSW2**-metal complexes observed at higher peptide concentrations (>40 μ M). The 2:1 peptide-metal complexes cannot accommodate all the residues capable of metal ligation (4Bpa¹⁰, Bmc¹, and His⁴) in an octahedral coordination geometry. As described earlier, the Co²⁺-complex formed under these conditions involved both the Bmc

and 4Bpa residues. In this complex the octahedral coordination sphere of Co^{2+} was saturated and excluded the involvement of histidine.

Unlike Zn^{2+} and Cu^{2+} , PK activity was not affected by the presence of Co^{2+} and this metal was therefore used to study the metal-dependency of inhibition by **MSW1** and **MSW2**. In contrast to **MSW1**, **MSW2** exhibited well behaved metal binding characteristics and the majority of these studies focused on this peptide.

As expected, uncomplexed peptides **MSW1** and **MSW2** were good inhibitors of PK activity. The seven residue inhibitory linker region of **MSW1** resulted in slightly better inhibition ($K_i=1.2\ \mu\text{M}$) than the five residue sequence in peptide **MSW2** ($K_i=2.9\ \mu\text{M}$). Unexpectedly, however, PK inhibition by **MSW1** and **MSW2** *increased* upon addition of CoCl_2 ; the peptide- Co^{2+} complexes were *better* inhibitors of PK activity than the uncomplexed peptides. Inhibition by both the 1:1 and 2:1 **MSW2**- Co^{2+} complexes ($K_i=0.9\ \mu\text{M}$) was approximately 3-fold lower than that of the uncomplexed peptide ($K_i=2.9\ \mu\text{M}$). This affect was attributed to the formation of a metal binding site that involves both the C- and N-terminal regions of the peptide. Peptide **DMW1** does not contain the Bmc residue at the N-terminus and metal ligation by this peptide involved only the C-terminal 4Bpa residue. The linker region was therefore unconstrained in the **DMW1**- Co^{2+} complex and inhibition of PK by **DMW1** was not affected by the addition of Co^{2+} . Thus, cyclization of peptide **MSW2** through the formation of the Co^{2+} complex apparently constrained the inhibitory linker region of this peptide into a *more* bioactive conformation, exactly the opposite of the intended design.

One possible reason for this observation was that a constraint formed was too "long" and did not sufficiently constrain the peptide backbone. In **MSW2**, the cross-link between the N- and C-termini was formed from the coordination of the

side chains of amino acids and was mediated by a metal ion; seven atoms separate the C α atoms of Bmc¹ and Bpa¹⁰ in the metal complexes. If metal ligation involved only Bmc¹ and 4Bpa¹⁰ (as was observed in the 2:1 **MSW2-Co²⁺** complex), the constraint would be even less effective. A "long" cross-link might actually result in the peptide termini being held distal from each other, constraining the intervening region into an extended conformation and resulting in improved PK inhibition. If this was the case, then the design of a "shorter" cross-link might remedy the problem.

In order to evaluate the effect of a short cross-link between the N- and C-terminus of the five residue inhibitory sequence, the peptide **PKCYS** was synthesized. This peptide contained a short, four atom cyclization (a disulfide bridge) between the C α atoms of the cysteine residues flanking the -R-R-N-A-I-sequence. The peptide was tested in its reduced (uncyclized) and oxidized (cyclized) form, and the inhibitory constants were determined. As expected introduction of a constraint on the peptide structure produced a poorer inhibitor. Unfortunately, cyclization had only a small effect on the inhibitory properties [$K_i(\text{ox})/K_i(\text{red}) = 2.5$]. However a basic premise of this project was that if the inhibitory sequence could be constrained into a non-extended conformation, then inhibition of the enzyme would *drastically* decrease. The modest decrease in inhibition despite the utilization of the shortest conceivable side chain-to-side chain cyclization (cystine) points to the unsuitability of this enzyme system for the modulation of inhibition through a control of peptide conformation. The enzyme active site appears flexible enough to accommodate inhibitors in non-extended conformations. Subsequent modeling of **PKCYS-O** revealed that even in the cyclized form, the inhibitor region could adopt a conformation which presented the side chains of the amino acids in an array similar to that observed for the inhibitor (PKI) bound to the active site of PK in crystal structure.¹⁰

Although the peptide inhibitor backbone was not in an identical extended conformation, the overall disposition of functionality was very similar (Figure 5-17).

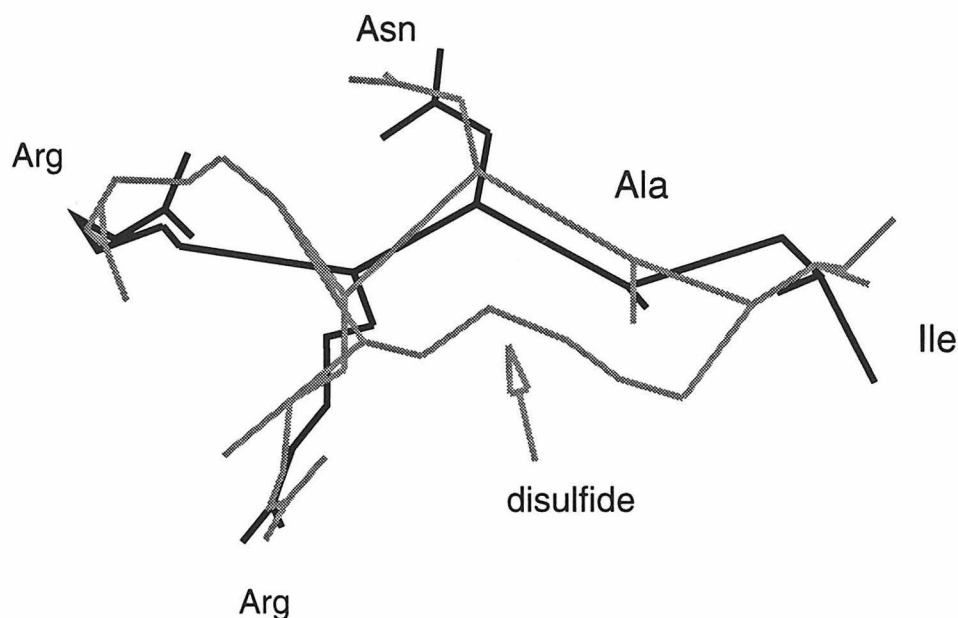


Figure 5-17 Comparison of a proposed conformation for **PKCYS-O** (gray) obtained by molecular modeling, with the crystal structure of the analogous region of PKI(5-24) bound in the active site of the catalytic subunit of c-AMP dependent protein kinase (PK) (black).¹⁰ The side chains and C α trace is shown.

Conclusion

The allosteric control of some enzyme cascade systems produces enormous amplification of the signals initiated by small molecule metabolites. The goal of this project was to utilize the enviable signal sensitivity and amplification properties found in these natural systems for the creation of new metallosensors. The substitution of a regulatory protein under the allosteric control of small molecule metabolites, with a regulatory metallo-peptide of *de*

novo design, would allow us to harness the "bio-amplification" potential of these systems for the detection of transition metal cations. The design was intended to *reduce* the inhibitory properties of a peptide through the metal mediated control of peptide conformation. Unfortunately the metal complexes of the designed regulatory peptides exhibited greater inhibition than the uncomplexed species, potentially due to the "length" of the metal-mediated cross-link. Cyclization of the inhibitor *via* formation of a metal complex apparently constrained the peptides into a more bio-active conformation (Figure 5-18). While the objectives of this project were not entirely met, the study does highlight the potential to modulate enzymatic activity through the metal-dependent folding of small peptidyl motifs.

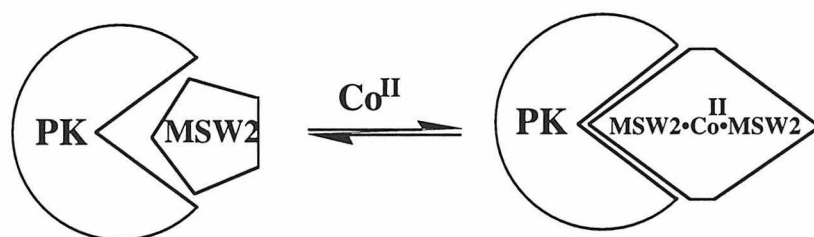


Figure 5-18 Metal mediated control of c-AMP dependent protein kinase inhibition.

The construction of a disulfide-cyclized inhibitor indicated that insufficient conformational constraint was obtained *via* a side chain-to-side chain cyclization. In this regard, unnatural amino acids which place the metal chelating moiety within the backbone of a peptide (Figure 5-19) have been synthesized recently in the Imperiali group²⁵ and could be envisioned to provide

greater control of backbone conformation, while still retaining the desired modularity of the peptidyl framework.

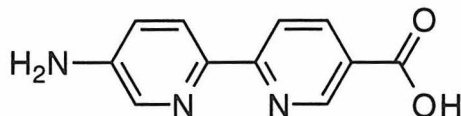


Figure 5-19 An example of the unnatural amino acids which place the metal chelating moiety within the peptide backbone and are currently under development in the Imperiali group.

Unfortunately the c-AMP dependent protein kinase system was not the ideal system for the creation of a metallo-sensor. Kinase activity was inherently sensitive to the presence of several divalent metal cations and enzyme stability required the use of thiol containing reducing agents that could interfere with metal binding by designed peptides. Although the PK system is not very amenable to the concept of metal-dependent regulation described in this chapter any enzyme which binds peptidyl substrates (and inhibitors) in an extended conformation could be a potential target for regulation, if sufficient metal-mediated control over peptide conformation can be achieved.

Acknowledgement

Dr. Alicia Torrado synthesized the Fmoc-Bmc(Trt)-OH used in the synthesis of **MSW2**. She also assisted in the synthesis of **MSW2** and performed some of the UV-Vis metal titrations of this peptide.

Experimental

Peptide Synthesis

Peptides **MSW1** and **MSW2** were synthesized by standard solid phase methods on a Milligen 9050 automated peptide synthesizer using *N*^α-9-fluorenylmethoxycarbonyl (Fmoc) protected amino acid pentafluorophenyl esters. Fmoc-Bmc(Trt)-OH, Fmoc-Arg(Pmc)-OH, and Fmoc-4Bpa-OH were used as the free acid derivatives and coupled using HOBT/DIPCDI chemistry.²⁶ The Enhanced Monitoring Option (Millipore, MA) was used to monitor coupling efficiencies for the standard amino acids.²⁷ Peptides were assembled on PAL-PEG-PS resin (Millipore, MA) to yield the C-terminal amide peptide. Fmoc deprotection was accomplished by washing the resin with 20% piperidine in DMF. Typical coupling cycles were 45 minutes to an hour using 3 fold excess of standard amino acids. A 2.5 fold excess of unnatural amino acids (Fmoc-Bmc(Trt)-OH and Fmoc-4Bpa-OH) was employed for the initial coupling cycle, followed by a second "double-coupling" cycle using 1.5 equivalents. The Fmoc-4Bpa-OH amino acid was manually injected (19-20 mg/ml concentration) onto the column due to poor solubility in DMF. Residual terminal amines were at were *N*-acetylated using a 0.3 M acetic anhydride/HOBT solution in 9:1 DMF/DCM the end of each coupling cycle. Peptides were *N*-acetylated on the resin after assembly, using 21 equivalents of acetic anhydride and 6 equivalents of triethyl amine in 4 ml of DMF for 2 hours. Peptide **MSW1** was cleaved using Reagent R (90:5:3:2 TFA/thioanisole/ethanedithiol/anisole)²⁸ and peptide **MSW2** was cleaved using Reagent K (82.5:2.5:5:5:5 TFA/ethanedithiol/anisole/phenol/H₂O)²⁹ for 2.5 hours. Peptides were precipitated from the cleavage mixture with 2:1 ether/hexane. The supernatant was decanted and trituration with 2:1 ether/hexane was repeated. The peptides were lyophilized from water. The crude peptides were analyzed and purified using reverse phase (C₁₈) HPLC

on a dual pump system (Beckman Instruments) with of 0.1% TFA in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). All HPLC linear gradients (LG) are listed as the increase in the percentage of Solvent B in the mobile phase. Care was taken to prevent intermolecular oxidation of the thiol residue (Bmc) to the corresponding disulfide. Purified HPLC samples were immediately frozen after removal of the acetonitrile under a gentle stream of N₂. The frozen samples were then lyophilized. The purified dried peptide was transferred as a solid to a weighed microcentrifuge tube, which was blanketed with N₂ and stored at -80 °C until ready for use. Stock solutions of peptides were prepared by dissolution in hydrogen sparged H₂O and were frozen at -80 °C between assays. The free thiol content was frequently verified using Ellman's reagent³⁰ and generally agreed to within 5-15 % of the values obtained by quantification of the bipyridyl chromophore of the 4Bpa residue by UV spectroscopy ($\epsilon_{286\text{nm}}=12,800$). Analysis of stock solutions of cysteine indicates that a 10 % variance is within the accuracy of the Ellman's test.

MSW1: Ac-Bmc-Pro-DSer-His-Thr-Gly-Arg-Arg-Asn-Ala-Ile-4Bpa-Phe-NH₂

HPLC: t_R = 16.06 min. Semi preparative column, LG 22-35% acetonitrile in H₂O, 0.1% TFA. C₇₃H₁₀₇O₁₇N₂₅S₁; PDMS MH⁺ calculated 1639.8 observed 1639.9

DMW1: Ac-Pro-DSer-His-Thr-Gly-Arg-Arg-Asn-Ala-Ile-4Bpa-Phe-NH₂

HPLC: t_R = 15.35 min. Semi preparative column, LG 22-35 % acetonitrile in H₂O, 0.1% TFA. C₆₉H₁₀₀O₁₆N₂₄; PDMS MH⁺ calculated 1522.7 observed 1523.1

MSW2: Ac-Bmc-Pro-DSer-His-Arg-Arg-Asn-Ala-Ile-4Bpa-Phe-NH₂

HPLC: t_R = 16.82 min Semi preparative column, LG 25-30 % acetonitrile in H₂O, 0.1% TFA. C₆₇H₉₇O₁₅N₂₃S₁; PDMS MH⁺ calculated 1481.7 observed 1481.0

Peptide **DMW1** was a major side product in the synthesis of **MSW1**. Mass spectral analysis of this side product (MH^+ 1523 amu) corresponded to the peptide sequence lacking the N-terminal Bmc residue. Furthermore, in the synthesis of both **MSW1** and **MSW2**, an additional side product was observed. HPLC analysis of crude **MSW1** indicated that this side product was produced in approximately equal quantities to the desired peptide (**MSW1**) and had a slightly lower $\text{Abs}_{280\text{nm}}/\text{Abs}_{228\text{nm}}$ ratio compared to the desired product. Mass spectral analysis revealed that this product had a molecular weight corresponding to the loss of H_2S (PDMS, MH^+ 1605.7). This product apparently results from α,β elimination of the side chain of the β -methyl cysteine amino acid to produce dehydrobutyrine during peptide synthesis or cleavage. A similar product was observed in the synthesis of **MSW2**.

Peptides **KS1** was synthesized on a 0.125 mmole scale using commercially available N^{α} -9-fluorenylmethoxycarbonyl protected amino acids (3 equivalents), coupled using HOBT/DIPCDI chemistry.²⁶ Coupling efficiencies were monitored using the Enhanced Monitoring Option (Millipore).²⁷ Fmoc-Gly-PAC-PEG-PS resin (0.22 meq/g, Millipore, MA) was used to produce the free C-terminal acid. The peptide was cleaved using Reagent K²⁸ for 2.25 hours under dilute conditions to prevent alkylation of the tryptophan residue. Peptide **KS1** was purified using a SepPak reverse phase (C_{18}) cartridge (Millipore, MA). The peptide was loaded in H_2O , the column was washed with 3 ml of H_2O , and elution accomplished by washing the column with 3 ml fractions of acetonitrile/water mixtures from 10% to 50% in 10% increments. The fractions were analyzed by analytical HPLC (t_R = 27 min, linear gradient 0-40 % acetonitrile in H_2O , 0.1% TFA), and the pure fractions were combined and lyophilized. The identity of the peptide was confirmed by FAB mass spectral analysis ($\text{C}_{40}\text{H}_{66}\text{O}_9\text{N}_{14}$, MH^+ calculated 888.1 observed 888)

Peptide **KI1** was synthesized in the same manner as described for **KS1** on Pal-Peg-PS resin (0.21 meq/g, Millipore, MA), to yield the C-terminal amide. Commercially available standard Fmoc protected amino acids were used. The peptide was purified by reverse phase (C₁₈) chromatography using a SepPak cartridge (Millipore, MA). The peptide was loaded in H₂O, the cartridge was washed with 5 ml of H₂O, and the peptide eluted with 30 % acetonitrile. The purity of the peptide was verified by analytical HPLC analysis (t_R = 19 min, linear gradient 0-40 % acetonitrile in H₂O, 0.1 % TFA). The identity of the peptide was confirmed by amino acid analysis; relative to glycine content: 0.72 ± 0.14 Asx (hydrolysis of Asn); 1.0 Gly; 2.07 ± 0.22 Arg; 0.77 ± 0.02 Thr; 1.11 ± 0.02 Ala; 1.06 ± 0.06 Ile.

The synthesis of **PKCYS** was accomplished using the same protocols described for the synthesis of **KI1** with the exception that the OPfp ester of asparagine (Fmoc-Asn-OPfp, Millipore, MA) was used to eliminate the need for side chain protection. Precautions were taken to ensure that **PKCYS** was not oxidized during purification. The identity of the reduced peptide (**PKCYS-R**) was confirmed by ESI mass spectral analysis. The presence of free sulfhydryls was verified using Ellman's reagent.³⁰ Stock solutions of **PKCYS-R** were monitored by HPLC analysis to verify the absence of oxidation products.

The oxidation of **PKCYS-R** to the cyclic cystine peptide **PKCYS-O** was accomplished by treatment of dilute solutions of purified peptide (1.0 mg/ml) with aqueous DMSO (20%) at pH 7.5.³¹ The reaction was monitored by HPLC, and proceeded to completion at room temperature with stirring in 15-24 hours. The reaction mixture was diluted 5-10 fold into 0.1% TFA in H₂O and directly purified by reverse phase (C₁₈) chromatography on a Sep-Pak cartridge (Millipore, MA). After washing the cartridge with water, the peptide was eluted with a 5 ml of 10% acetonitrile in H₂O (0.1% TFA). The peptide did not need

further purification and was obtained in approximately 50% yield from the purified **PKCYS-R**. HPLC analysis indicated that the peptide was >90% pure and that no reduced peptide remained. The identity of the oxidized peptide was verified by mass spectral analysis and quantitative amino acid analysis.

PKCYS-R: Analytical HPLC: t_R = 26.9 min LG 0-30% acetonitrile in H₂O, 0.1% TFA. Molecular formula C₃₃H₆₁O₉N₁₅S₂; ESI-MS MW calculated 876.1 observed 876.2.

PKCYS-O: Analytical HPLC: t_R = 25.9 min LG 0-30% acetonitrile in H₂O, 0.1% TFA. Molecular formula C₃₃H₅₉O₉N₁₅S₂; ESI-MS MW calculated 874.1 observed 874.1.

UV-Vis Spectroscopic Studies

Absorption spectra were obtained on a Shimadzu UV-110 spectrophotometer using 1.0 cm path length quartz cuvettes. The concentrations of the peptide solutions were determined spectrophotometrically using the characteristic absorption band of the bipyridyl moiety ($\lambda_{\max}$ = 286 nm, $\epsilon_{\max}$ = 12,800 M⁻¹cm⁻¹). Aqueous solutions of known peptide concentrations (7-45 μ M) were titrated with standardized metal cation solutions at room temperature in 0.02 M HEPES buffer, 0.012 M KCl at pH 7.5. Metal titrations of **MSW2** were also performed at various pH values using either HEPES buffer (pH 7.0 and 8.0) and MES buffer (pH 6.0). The distinct red-shift of the bipyridyl π - π^* transition upon metal chelation was monitored. The dissociation constants were determined with the calculations outlined below. Data analysis was performed using KaleidaGraph version 3.0 (Abelbeck Software). Binding isotherms were assumed to follow two-state

equilibria. In most cases well defined isosbestic points were observed (see Results section). The dissociation constants (K_D) were obtained using an iterative process based on the Scott equation¹³ using absorbance data at 310 nm:

$$\frac{b \cdot [M]}{\Delta A} = \frac{[M]}{[Pep]_t \cdot \Delta \epsilon} + \frac{K_D}{[Pep]_t \cdot \Delta \epsilon} \quad (1)$$

Where :	b	=	cell path length
	$[M]$	=	free metal cation concentration
	ΔA	=	change in absorbance due to complex formation
	$[Pep]_t$	=	total peptide in solution
	$\Delta \epsilon$	=	change in molar absorptivity due to complex formation

The dissociation constants (K_D) were calculated from plots of $[M]$ vs $[M]/\Delta A$ by the following equation:

$$K_D = \frac{\text{y - intercept}}{\text{slope}} \quad (2)$$

Initially, data was considered only at high $[M]_{\text{total}}$ so that the approximation $[M]_{\text{total}} \approx [M]$ could be used. The resultant values of K_D and $\Delta \epsilon$ obtained by least squared analysis were fitted to absorbances using equation 1 to obtain better estimates of $[M]$. This was repeated in an iterative manner until the values of K_D and $\Delta \epsilon$ converged.

Standard Assay Protocols

The assay used was adapted from previously published protocols.¹⁹ All assays were performed at room temperature in a final volume of 250 μ l and consisted of the following components:

Buffer: The buffer solution (25mM HEPES, 15 mM KCl, pH 7.5) was sparged with hydrogen for 15 minutes prior to use (to prevent oxidation of the thiol containing inhibitors **PKCYS-R**, **MSW1**, and **MSW2**).

ATP and Mg²⁺: Stock solutions of [γ -³²P]ATP•Mg were prepared by diluting [γ -³²P]ATP (10 Ci/ mmol, NEN) with unlabeled MgATP (Sigma) to the desired specific activity (75-150 cpm/pmol). The concentration of ATP in all assays was 100 μ M, and the total concentration of Mg²⁺ was 124 μ M (98 μ M from the Mg•ATP stock and the remaining from MgCl₂).

Substrate: Stock solutions of the substrate "Kemptide" (L-R-R-A-S-L-G) (Sigma) were quantified by quantitative amino acid analysis prior to use. Substrate concentrations varied from 16 μ M to 40 μ M depending on inhibitor levels.

Enzyme: The catalytic subunit of cAMP dependent Protein Kinase from bovine heart was obtained from Sigma as a lyophilized powder containing approximately 0.11% protein and a balance of sucrose and phosphate buffer salts. The enzyme was prepared for use by dissolution in 2.5 mM BME to a protein concentration of 0.1 mg/ml. The enzyme was allowed to warm to room temperature for 5 minutes before use.

Inhibitors: Stock solutions of all inhibitors were quantified either by quantitative amino acid analysis for peptides **KI1**, **PKCYS-R**, and **PKCYS-O** or by UV spectrometry of the bipyridyl chromophore (ϵ =12,800, $\lambda_{\max}$ = 286 nm) for peptides **MSW1**, **MSW2**, or **DMW1**. In some experiments CoCl₂ was added to the assay to form the inhibitor-Co²⁺ complex.

Final concentrations in the assay mixture were 12 mM KCl, 20 mM HEPES, 100 μ M [γ - 32 P]ATP, 124 μ M Mg^{2+} , and 20 μ M β -mercapto ethanol. Substrate, inhibitor, and CoCl_2 concentrations varied. The reaction was initiated by addition of the enzyme (200-250 ng in 2.0-2.5 μ l) and gently shaken at room temperature for the duration of the assay (8 min).

The initial velocity of the reaction was monitored essentially by the method of Roskoski.¹⁹ Briefly aliquots of the reaction mixture (25 μ l) were spotted onto strips of phosphocellulose paper (1 mm x 3 mm) (Whatman, P81) at four 2-minute intervals after the start of the reaction. The paper was immediately suspended in 75 mM of phosphoric acid (100 ml per 16 strips). The paper strips were washed four times in phosphoric acid. The strips were then transferred to 7 ml scintillation tubes suspended in scintillation fluid and the radioactivity counted.

Several modifications to previously reported kinase assays were necessary. Typically the assay protocols described previously utilized 3 - 10 mM MgCl_2 as the Mg^{2+} source.¹⁸ Addition of 3 mM MgCl_2 to peptide **MSW2** resulted in the distinctive red shift of the bipyridyl absorbance which indicated that the peptide bound Mg^{2+} . Studies indicated that the MgCl_2 concentrations could be lowered to 124 μ M (1.24 equivalents with respect to ATP) in the assay mixture without adversely affecting enzyme activity. CoCl_2 titrations of **MSW2** in the presence of 124 μ M MgCl_2 exhibited isotherms identical to titrations performed in the absence of Mg^{2+} . Additionally the recommended procedure for reconstitution of the catalytic subunit of PK purchased from Sigma utilized DTT. Since this di-thiol has significant metal binding properties²³ it was replaced with an alternative mono-thiol reducing agent β -mercapto ethanol (BME).³² BME levels could be reduced to 20 μ M in the assay mixture (enzyme reconstitution in 2.5 mM BME) without adversely affecting the activity of the enzyme over the

time course of the assay (<12 min). The stability of the enzyme did suffer from the diminished levels of reducing agent, and for this reason all enzyme solutions were prepared immediately prior to use.

For all experiments involving the use of an inhibitor the fraction inhibition was determined by the following equation:

$$i = 1 - (v_i/v_c)$$

v_i represents the initial velocity of the assay mixture (in cpm/min) containing the inhibitor and v_c represents the initial velocity of the control assay. The fraction inhibition was used to estimate the K_i by use of the following equation¹²:

$$i = [I] / ([I] + K_i(1 + [S]/K_m))$$

or

$$K_i = ([I] - i[I]) / (i + [S]i/K_m)$$

[S] = Kemptide concentration.

$K_m = 4.7 \mu\text{M}$ (for Kemptide)³³

The fractional inhibition is given by i . Examples of the inhibition of PK activity by **MSW2**, **DMW1**, **MSW1**, **PKCYS-O**, and **PKCYS-R** are given in Figures 5-20, 5-21, 5-22, 5-23.

The inhibitory effect of divalent transition metal cations Zn^{2+} , Co^{2+} and Cu^{2+} on PK activity was determined. Although Zn^{2+} and Cu^{2+} resulted in inhibition, Co^{2+} did not have an effect on enzyme activity at concentrations up to $50 \mu\text{M}$ (Figures 5-24). Since Co^{2+} can restore partial activity to PK in the absence of Mg^{2+} ,³² inhibition of $50 \mu\text{M}$ CoCl_2 was tested at lower MgCl_2 levels. The addition of $50 \mu\text{M}$ CoCl_2 in the presence of a total of $124 \mu\text{M}$ Mg^{2+} and $300 \mu\text{M}$ Mg^{2+} cations had no effect on activity.

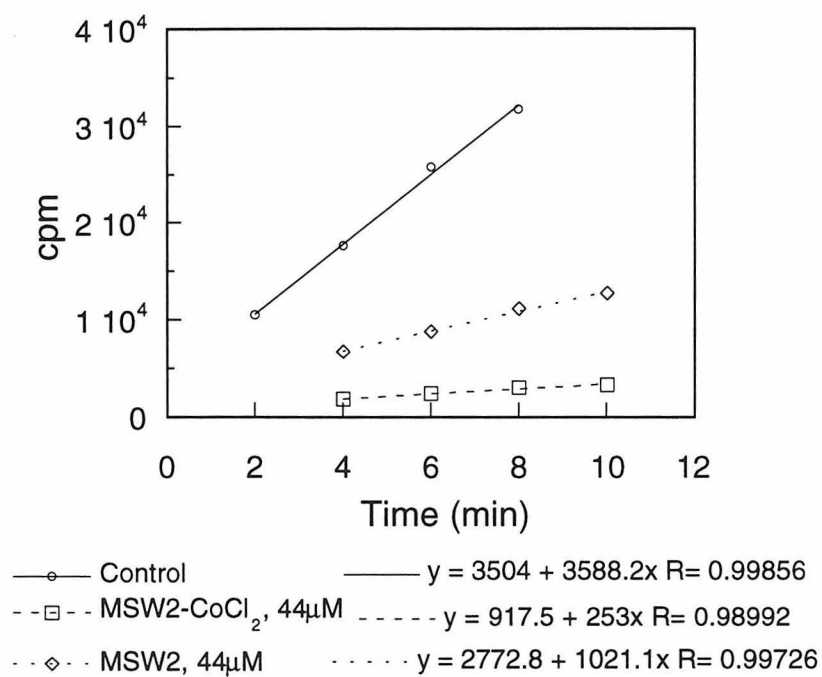


Figure 5-20 Inhibition of PK by both uncomplexed peptide and the Co²⁺ complex of MSW2. (24 μM MgCl₂, 100 μM Mg[γ-³²P]ATP (100cpm/pmole), 250 ng enzyme, 20 μM BME, 39.5 μM Kemptide)

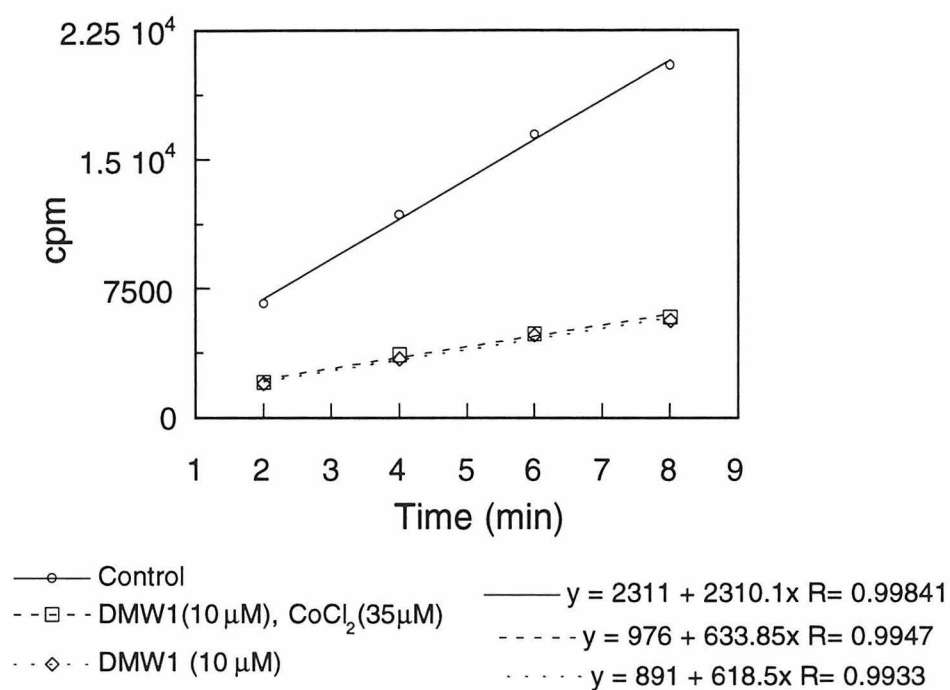


Figure 5-21 Inhibition of PK by both uncomplexed peptide and the Co²⁺ complex of **DMW1**. (24 μM MgCl₂, 100 μM Mg[γ-³²P]ATP (84 cpm/pmole), 250 ng enzyme, 20 μM BME, 15.8 μM Kemptide)

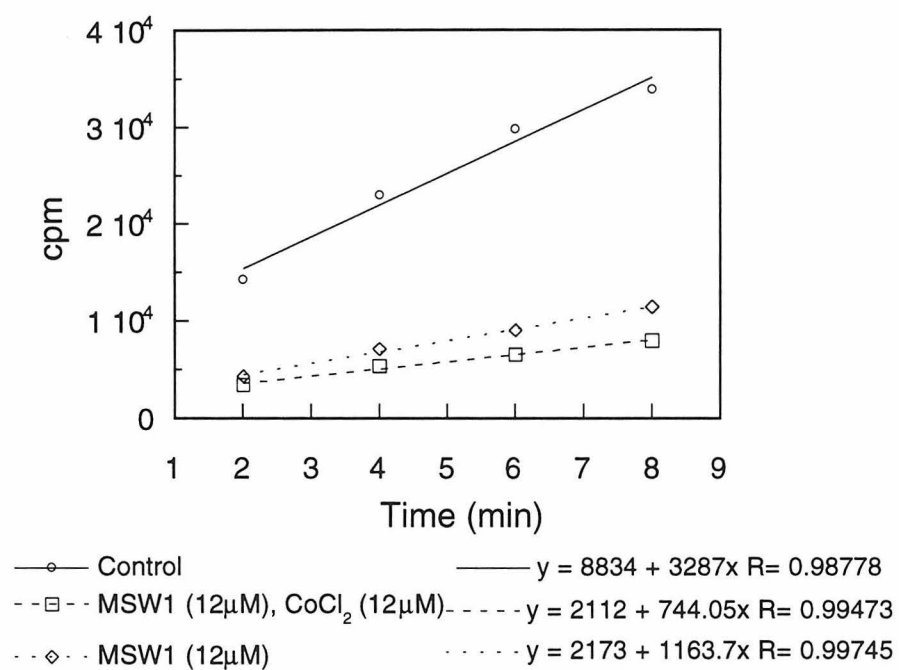


Figure 5-22 Inhibition of PK by both uncomplexed peptide and the Co²⁺ complex of MSW1. (24 μM MgCl₂, 100 μM Mg[γ-³²P]ATP (84 cpm/pmole), 250 ng enzyme, 20 μM BME, 15.8 μM Kemptide)

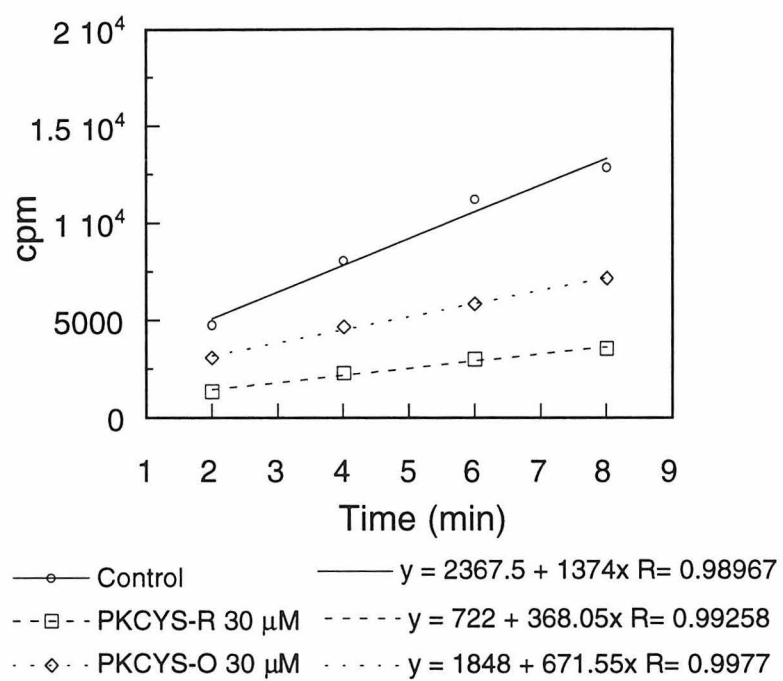


Figure 5-23 Inhibition of PK by both reduced (**PKCYS-R**) and oxidized (**PKCYS-O**) forms of peptide **PKCYS**. (24 μM MgCl_2 , 100 μM $\text{Mg}[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (98 cpm/pmole), 250 ng enzyme, 20 μM BME, 18.8 μM Kemptide)

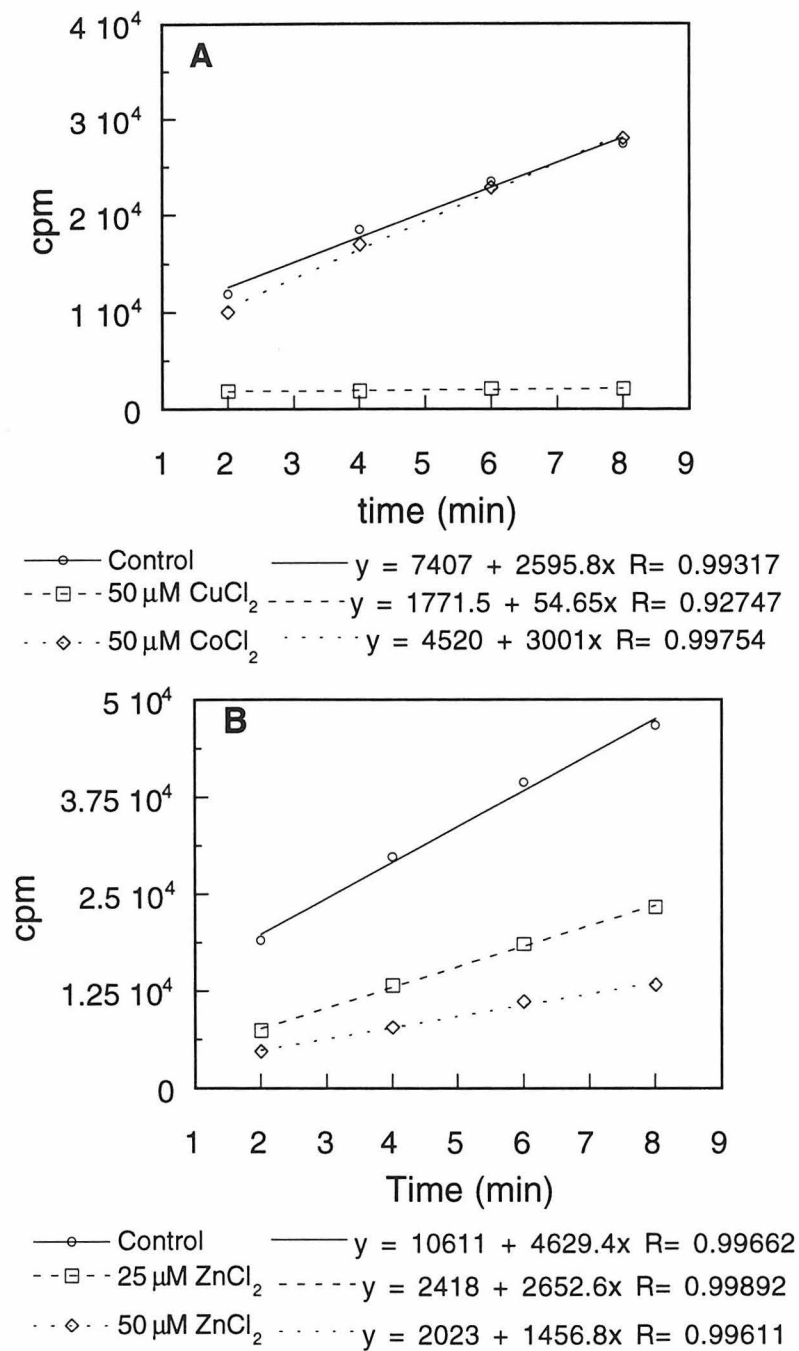


Figure 5-24 PK activity in the presence of divalent metal cations. **A.** Inhibition by Cu^{2+} . Co^{2+} does not inhibit. **B.** Inhibition by Zn^{2+} (3.2 mM MgCl_2 , 20 μM BME, 200 ng enzyme, 23.7 μM Kemptide, 100 μM $\text{Mg}[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (108-150 cpm/pmole))

References

- (1) Fisher, S. L., "Unnatural Amino Acids in the Synthesis and Semisynthesis of Metalprotein Motifs," Ph.D. Thesis, California Institute of Technology, 1993.
- (2) Imperiali, B.; Prinns, T. J.; Fisher, S. L. "Chemoenzymatic Synthesis of 2-Amino-3-(2,2'-bipyridinyl)-Propanoic Acids," *J. Org. Chem.* **1993**, *58*, 1613.
- (3) Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prinns, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide.," *J. Am. Chem. Soc.* **1992**, *114*, 3182.
- (4) Imperiali, B.; Fisher, S. L. "Stereoselective Synthesis and Peptide Incorporation of (S)- α -Amino-(2,2'-bipyridine)-6-Propanoic Acid," *J. Org. Chem.* **1992**, *57*, 757-759.
- (5) Taylor, S. S.; Buechler, J. A.; Yonemoto, W. "cAMP-Dependent Protein Kinase: Framework for a Diverse Family of Regulatory Enzymes," *Ann. Rev. Biochem.* **1990**, *59*, 971-1005.
- (6) Bramson, H.N.; Kaiser, E.T.; Mildvan, A.S. "Mechanistic Studies of cAMP-Dependent Protein Kinase Action," *CRC Crit. Rev. Biochem.* **1984**, *15*, 93-124.
- (7) Bechtel, P.J.; Beavo, J. A.; Krebs, E. G. "Purification and Characterization of Catalytic Subunit of Skeletal Muscle Adenosine 3':5'-monophosphate-dependent Protein Kinase," *J. Biol. Chem.* **1977**, *252*, 2691.
- (8) Feramisco, J. R.; Krebs, E. G. "Inhibition of Cyclic AMP-dependent Protein Kinase by Analogues of a Synthetic Peptide Substrate," *J. Biol. Chem.* **1978**, *253*, 8968-8971.
- (9) Kington, D. R.; Zheng, J.; Ten Eyck, L.F.; Ashford, V.A.; Xuong, N.; Taylor, S.S.; Sowadski, J.M. "Crystal Structure of the Catalytic Subunit of Cyclic

Adenosine Monophosphate-Dependent Protein Kinase," *Science* **1991**, 253, 407-414.

(10) Kinghton, D.R.; Zheng, J.; Ten Eyck, L.F.; Xuong, N.; Taylor, S.S.; Sowadski, J.M. "Structure of a Peptide Inhibitor Bound to the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase," *Science* **1991**, 253, 414-420.

(11) Glass, D.B.; Cheng, H.; Mende-Mueller, L.; Reed, J.; Walsh, D. "Primary Structural Determinants Essential for Potent Inhibition of cAMP-dependent Protein Kinase by Inhibitory Peptides Corresponding to the Active Portion of the Heat-stable Inhibitor Protein," *J. Biol. Chem.* **1989**, 264, 8802-8810.

(12) Segel, I. H. *Enzyme Kinetics*; John Wiley and Sons: New York, 1975.

(13) Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*; John Wiley and Sons: New York, 1987.

(14) Smith, R. M.; Martell, A. E. *Critical Stability Constants, Vol 2, Amines*; Plenum Press: New York, 1974.

(15) Dawson, R. M. C.; Elliot, D. C.; Elliot, W. H.; Jones, K. M. *Data For Biochemical Research, 3rd Edition*; Clarendon Press: Oxford, 1986.

(16) Beavo, J. A.; Bechtel, P. J.; Krebs, E. G. "Preparation of Homogenous Cyclic AMP-dependent Protein Kinases(s) and It's Subunits from Rabbit Skeletal and Bovine Heart Muscle," *Meth. Enzym.* **1974**, 38C, 299.

(17) Witt, J. J.; Roskoski Jr., R. "Rapid Protein Kinase Assay Using Phosphocellulose-Paper Absorption," *Analytical Biochem* **1975**, 66, 253-258.

(18) de la Houssaye, B.A.; Masaracchia, R.A. "Standardization of the Assay for the Catalytic Subunit of Cyclic AMP-Dependent Protein Kinase Using a Synthetic Peptide Substrate," *Anal. Biochem.* **1983**, 128, 54-59.

(19) Roskoski, Rober, Jr. "Assays of Protein Kinase," *Meth. Enzym.* **1983**, 99, 3-6.

- (20) Wright, D.E.; Noiman, E.S.; Chock, P.B.; Chau, V. "Fluorometric Assay for Adenosine 3',5'-cyclic Monophosphate-Dependent Protein Kinase and Phosphoprotein Phosphatase Activities," *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 6048-6050.
- (21) Ferry, G.; Ernould, A.-P.; Genton, A.; Boutin, J. A. "Assay of Tyrosine Protein Kinase Activity from HF-60 by High-Performance Liquid Chromatography for Specificity Studies," *Anal. Biochem.* **1990**, *190*, 32-38.
- (22) Nakanishi, S.; Kase, H.; Matsuda, Y. "Assay of Myosin Light Chain Kinase Activity by High-Performance Liquid Chromatography Using a Synthetic Peptide as Substrate," *Anal. Biochem.* **1991**, *195*, 313-318.
- (23) Cornell, N. W.; Crivaro, K. E. "Stability Constant for the Zinc-Dithiolthreitol Complex," *Anal. Biochem.* **1972**, *47*, 203-208.
- (24) Thompson, J. S.; Sorrell, R.; Marks, T. H.; Ibers, J. A. "Synthesis Structure and Spectroscopy of Pseudotetrahedral $\text{Co}^{\text{II}}\text{N}_2(\text{SR})$ Complexes. Active Site Approximations to the $\text{Co}(\text{II})$ -Substituted Type I Copper Proteins," *J. Am. Chem. Soc.* **1979**, *101*, 4193-4200.
- (25) Torrado, A.; Imperiali, B., unpublished results
- (26) Bodanszky, M. *Principles of Peptide Synthesis*; Springer-Verlag: Berlin, 1993.
- (27) Young, S. C.; White, P. D.; Davies, J. W.; Owen, D. E. I. A.; Salisbury, S. A.; Tremeer, E. J. "Counterion Distribution Monitoring - A Novel Method for Acylation Monitoring in Solid-Phase Peptide Synthesis," *Biochem. Soc. Trans.* **1990**, *18*, 1311-1312.
- (28) Biosearch Technical Bulletin No. 9000-02
- (29) King, D. S.; Fields, C. G.; Fields, G. B. "A Cleavage Method which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis," *Int. J. Pep. Protein Res.* **1990**, *36*, 255-266.

- (30) Ellman, G. L. "Tissue Sulfhydryl Groups," *Arch. Biochem. Biophys.* **1959**, 82, 70-77.
- (31) Tam, J. P.; Wu, . C.-R.; Liu, W.; Zhang, J.-W. "Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide Scope and Applications," *J. Am. Chem. Soc.* **1991**, 113, 6657-6662.
- (32) Rubin, C. S.; Erlichman, J.; Rosen, O. M. "Cyclic AMP-Dependent Protein Kinase from Bovine Heart Muscle," *Meth. Enzym.* **1974**, XXXVII, 308-315.
- (33) Whitehouse, S.; Feramisco, F. R.; Casnelli, J. E.; Krebs, E. G.; Walsh, D. A. "Studies on the Kinetic Mechanism of the Catalytic Subunit of the cAMP-dependent Protein Kinase," *J. Biol. Chem.* **1983**, 258, 3693-3701.

Appendix A

Electrostatic stabilization of sheet and turn structure

Introduction

Electrostatic interactions such as salt bridges and ion pairs have been shown to stabilize the formation of helices, both as isolated polypeptides,^{1, 2} and as part of four-helix bundles.³ The electrostatic stabilization of reverse turn structures has also been reported,⁴⁻⁶ although most conformational studies of these peptides have been performed in non-aqueous solutions or with a non-aqueous co-solvent; a non-aqueous environment enhances electrostatic interactions. In order to assess the potential for utilizing electrostatic or "salt bridge" interaction to stabilize a type II β -turn formation in water, structural analyses of several peptides incorporating both natural and non-standard charged amino acids have been performed using 2D NMR and CD techniques.

The type II β -turn has been extensively investigated within the Imperiali group as a useful motif for protein design⁷ (Chapter 1). A heterochiral Pro-DXaa sequence (Xaa = a small unbranched amino acid) at positions ($i+1$) and ($i+2$) of a tetrapeptide has been shown to induce type II β -turn formation.⁷ The pyrrolidine ring system of proline at position ($i+1$) 'kinks' the peptide backbone initiating a reversal of the peptide backbone. Analysis of type II β -turns within proteins of known structure has revealed that the frequent occurrence of glycine at position ($i+2$) of native type II β -turns is due to the stereochemical restrictions at this position.⁸ Large side chains of L-amino acids cannot be accommodated without an inversion of chirality. The use of a D-amino acid at this position is therefore appropriate and allows for the "position specific" incorporation of a type II β -turn^{7, 9} since D-amino acids are not compatible with other elements of secondary structure. The electrostatic stabilization of a heterochiral Pro-DAla type II β -turn was investigated by incorporating the charged amino acids glutamic acid and lysine at the (i) and ($i+3$) positions of the turn, respectively.

Since previous studies^{7, 9} have indicated that turn formation within tetrapeptides can be very sensitive to the identity of the residues at positions (*i*) and (*i*+3), the electrostatic stabilization of a β -hairpin motif incorporating charged residues at positions (*i*-1) and (*i*+4) was also investigated. In this regard, an unnatural amino acid, (2*S*,3*S*)-diaminobutyric acid (*allo*- β -methyl- β -amino alanine, or *a*Maa), was considered due to its unique electronic and conformational properties. Synthesis of this amino acid, appropriately protected for solid phase peptide synthesis, and peptide incorporation is reported.

Results and Discussion

Electrostatic Stabilization of a Type II β -Turn

As discussed, the heterochiral Pro-DAla sequence has previously been demonstrated to promote type II β -turn formation in isolated tetrapeptides.^{7, 9} Furthermore, electrostatic stabilization of helical structure has exploited interactions between glutamic acid and lysine residues separated by one turn of an α -helix [at positions (*i*, *i*+3) and (*i*, *i*+4)].² Peptide **PA1** (Ac-Glu-Pro-DAla-Lys-NH₂) was synthesized, wherein glutamic acid and lysine were included at positions flanking the Pro-DAla sequence, to test the ability of these charged amino acids to stabilize type II β -turn formation *via* formation of a salt bridge.

Analysis of the solution conformation of **PA1** by two-dimensional NMR and circular dichroism (CD) revealed a significant population of a type II β -turn character. The presence of diagnostic NOEs in the ROESY experiment (Figure A-1) can be used to identify a type II turn in tetrapeptides.⁷ Significantly a $d_{\beta\epsilon}(i,i+3)$ NOE was observed for **PA1** between the side chains at positions (*i*) and (*i*+3) indicating a reversal of the peptide backbone, and direct interaction of the side chains of Glu¹ and Lys⁴. Other NOEs characteristic of a type II reverse

turn^{7, 10} were observed for **PA1** and are listed in Table A-1 and shown in Figure A-2.

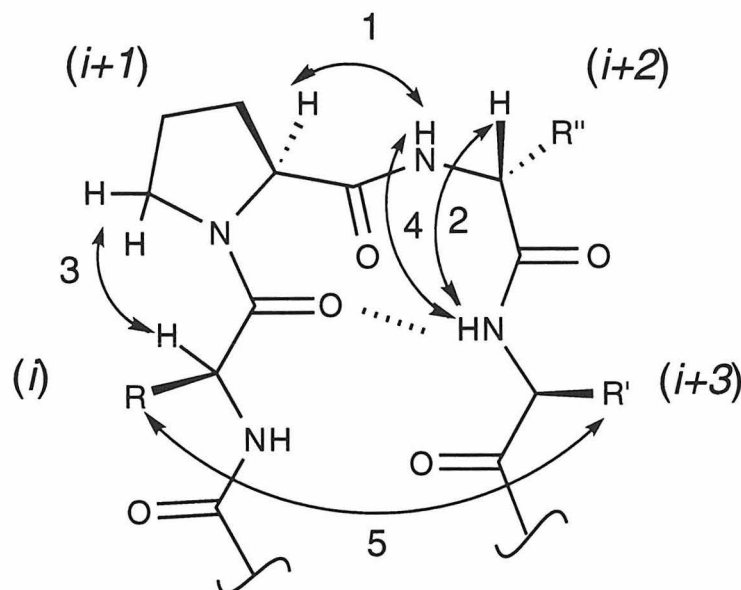


Figure A-1 NOEs indicative of type II β -turn formation in tetrapeptides. 1, $d_{\alpha N}(i+1, i+2)$; 2, $d_{\alpha N}(i+2, i+3)$; 3, $d_{\alpha, \delta}(i, i+1)$; 4, $d_{NN}(i+2, i+3)$; 5, $d_{sc}(i, i+3)$ (sc = side chain proton).

The temperature dependence of the chemical shifts of the amide protons of peptide **PD1** were also determined. All amide proton temperature coefficients in this peptide were unusually high ($-\Delta\delta/\Delta T > 8$ ppb/K); this may be due to interaction of the charged groups with the peptide backbone. Typically, depressed amide proton temperature coefficients ($-\Delta\delta/\Delta K < 6$ ppb/K, in water) are an indication that the amide proton is protected from solvent, potentially through a hydrogen bonding interaction.⁷ Although all the temperature coefficients were too high to make a definite conclusion on the extent of

hydrogen bonding, the amide proton of lysine at position ($i+3$) exhibits a temperature coefficient that is 3 ppb/K lower than the coefficients of the other two amide protons. This observation would be consistent with a type II β -turn in which the amide at the ($i+3$) position is involved in a hydrogen bond with the carbonyl oxygen of the residue at position (i)¹¹ (Figure A-1).

NOE #	Interaction	Strength
1	Pro ² α - DAla ³ NH	Strong
2	Lys ⁴ NH - DAla ³ α	Obscured by overlap
3, 3'	Glu ¹ α - Pro ² δ	Strong
4	Lys ⁴ NH - DAla ³ NH	Weak
5	Glu ¹ β - Lys ⁴ ϵ	Weak

Table A-1 NOEs observed in the ROESY spectrum of peptide **PA1** (90:10 H₂O/D₂O, pH 4.4). The NOE numbers correspond to those illustrated in Figure A-1.

Position	Residue	$-\Delta\delta/\Delta T$ (ppb/K)
(i)	Glu ¹	12.2
($i+2$)	DAla ³	11.3
($i+3$)	Lys ⁴	8.5

Table A-2 Amide proton temperature coefficients for peptide **PA1** in 90:10 H₂O/D₂O, pH 4.4.

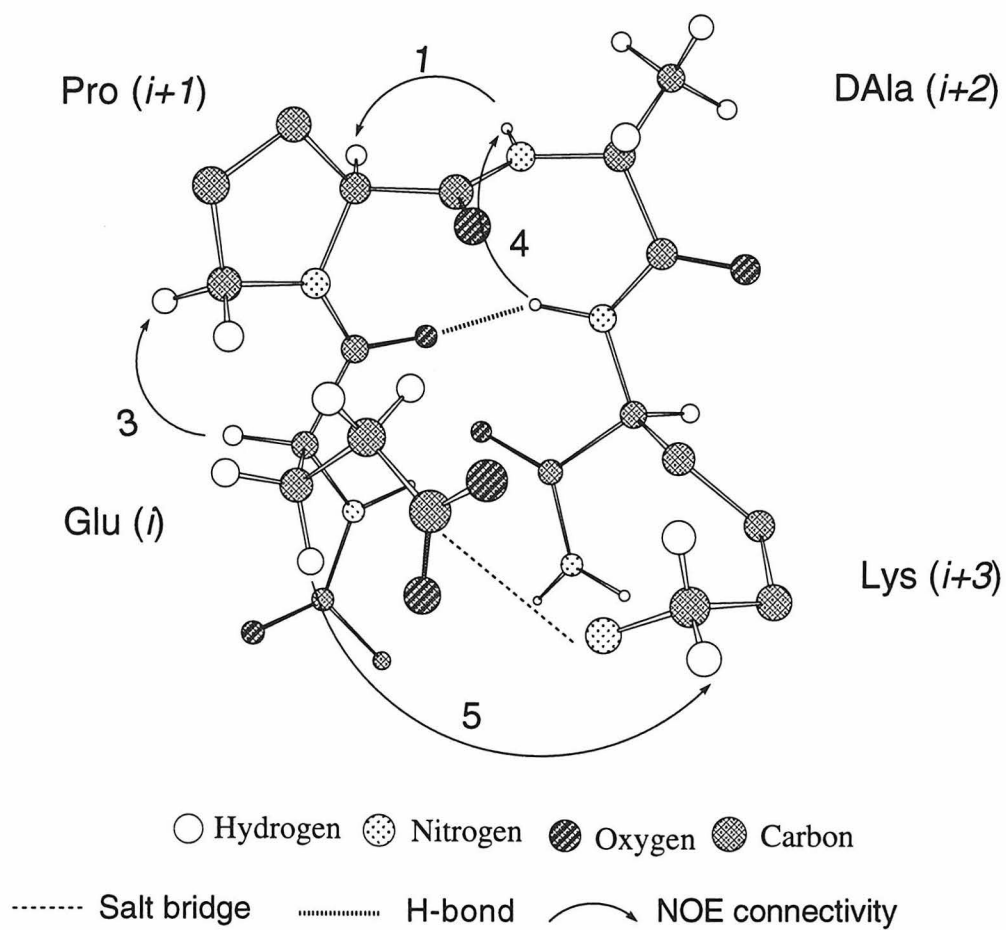


Figure A-2 Summary of NOEs observed in the ROESY spectrum of peptide PA1. Potential electrostatic and hydrogen bonding interactions are indicated.

CD analysis of **PA1** independently verified a type II β -turn solution conformation for this peptide in solution. The positive ellipticity at 200 - 215 nm and the negative ellipticity observed below 195 nm and above 220 nm are indicative of a type II β -turn (Figure A-3). The folded conformation of **PA1** was quite stable; the CD spectrum of the peptide was invariant between 5 °C and 50 °C (Figure A-3, plot A). In contrast, the CD spectrum exhibited a marked pH dependence. While the turn conformation appeared stable between pH 4.1 and pH 7.9, a sharp decrease in turn character was observed at pH 3 (Figure A-3, plot B). The dependence of the CD spectrum on pH suggested that the electrostatic interaction between Glu¹ and Lys⁴ stabilized the turn structure at pH values at which both the residues were ionized. Protonation of the side chain of Glu¹ below pH 4.0 disrupted this stabilizing interaction, and resulted in reduced turn character.

The CD and NMR analysis of peptide **PA1** indicated the presence of a folded type II β -turn conformation in solution. Two distinct structural features were included in this peptide to stabilize the folded structure: The heterochiral Pro²-DAla³ sequence, and the electrostatic interaction between residues Glu¹ and Lys⁴. The Pro-DAla sequence has been demonstrated to promote type II β -turn formation in tetrapeptides, although the turn propensity was dependent on the identity of the residues at positions (*i*) and (*i*+3).^{7, 9} To determine the relative importance of the electrostatic interaction between Glu¹ and Lys⁴ in the promotion of a type II β -turn conformation in peptide **PA1**, peptide Ac-Glu-Gly-Gly-Lys-NH₂ (**GG1**) was synthesized. In this peptide the turn-inducing heterochiral sequence of **PA1** was replaced by a flexible glycine linker (Gly²-Gly³).

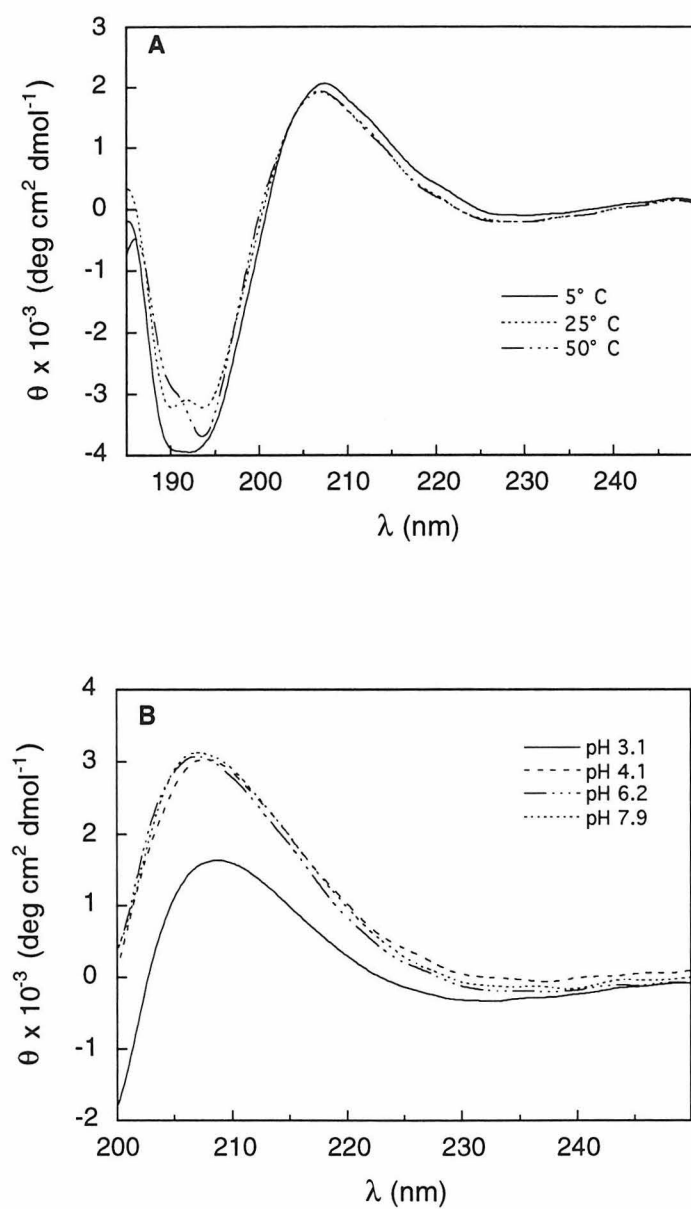


Figure A-3 CD spectra of peptide PA1. **A.** The temperature-dependence of CD spectrum. **B.** The pH-dependence of the CD spectrum.

NMR analysis of peptide **GG1** indicated that this peptide adopts a folded conformation in solution. Analysis of the ROESY spectrum for **GG1** revealed $d_{NN}(i, i+1)$ and $d_{NN}(i+1, i+2)$ NOEs, which would not be expected in an unstructured peptide. Although these NOEs are sequential, the short distance cutoff of the ROESY experiment ($< 3.5\text{\AA}$), makes observation of these NOEs indicative of a folded conformation.¹¹ Due to the small size of tetrapeptide **GG1**, the folded conformation can be most likely attributed to a reverse turn or a short 3_{10} helix. Observation of a sequential $d_{NN}(i, i+1)$ NOE is compatible with a type III reverse turn, which is similar in geometry to a single turn of a 3_{10} helix.^{10, 12} Three of the four possible $d_{NN}(i, i+1)$ NOEs were observed for peptide **GG1** (Table A-3) and are illustrated in Figure A-4. Significantly, these NOEs are not anticipated in peptides adopting the more common type I or type II reverse turn conformations.¹⁰ In addition to $d_{NN}(i, i+1)$ NOEs, observation of a $d_{\alpha N}(i, i+3)$ NOE would be expected in a 3_{10} helix. Unfortunately only one such NOE is possible in the tetrapeptide **GG1** [$d_{\alpha N}(\text{Glu}^1, \text{Lys}^4)$], and this NOE was obscured in the ROESY spectrum due to the overlapping C_{α} resonances of Glu^1 and Lys^4 . The results of the 2D NMR analysis are summarized in Table A-3 and Figure A-4.

NOE	Assignment	Intensity
1	$d_{NN}(\text{Glu}^1, \text{Gly}^2)$	medium
2	$d_{NN}(\text{Gly}^2, \text{Gly}^3)$	weak
3	$d_{NN}(\text{Lys}^4, \text{amide cap})$	weak

Table A-3 Summary of NOEs between amide protons in the ROESY spectrum of peptide **GG1**. The relative intensities of the NOEs are indicated.

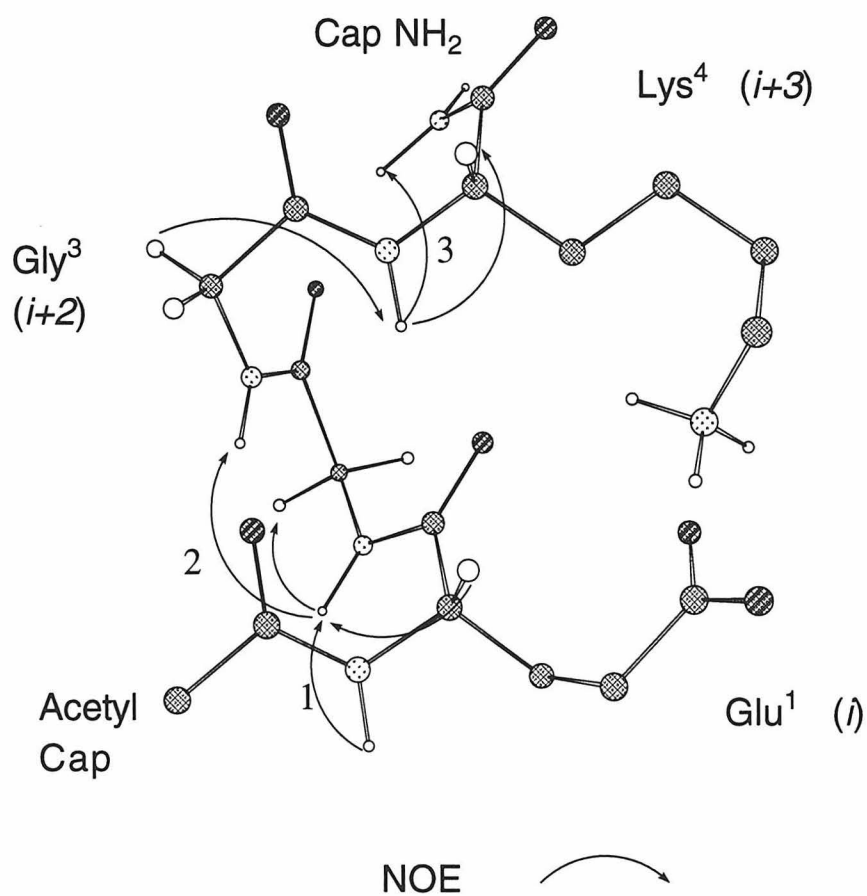


Figure A-4 Summary of the NOEs observed in the ROESY spectrum of **GG1**. The peptide is shown in the proposed 3_{10} helical conformation.

Analysis of the CD spectrum of **GG1** also suggested that the solution conformation of this peptide was a type III turn. Although the ideal CD spectrum for a type III reverse turn is not well characterized, there are several features which support this conformation over alternative reverse turn conformations. The CD spectra of **GG1** at various values of pH exhibited a small positive ellipticity at 205-215 nm as well as a slight negative ellipticity at 225 nm; CD spectra of type I turns are characterized by much larger ellipticities.¹³ The CD spectrum of **GG1** also correlates with the theoretically calculated spectrum for a type III turn by Woody.¹³ As expected the solution conformation of **GG1** was dependent on pH. Below pH 4, the CD spectrum indicated a reduced population of structured peptide (Figure A-5), indicating that an electrostatic interaction between Glu¹ and Lys⁴ is most likely responsible for the folded conformation at higher pH.

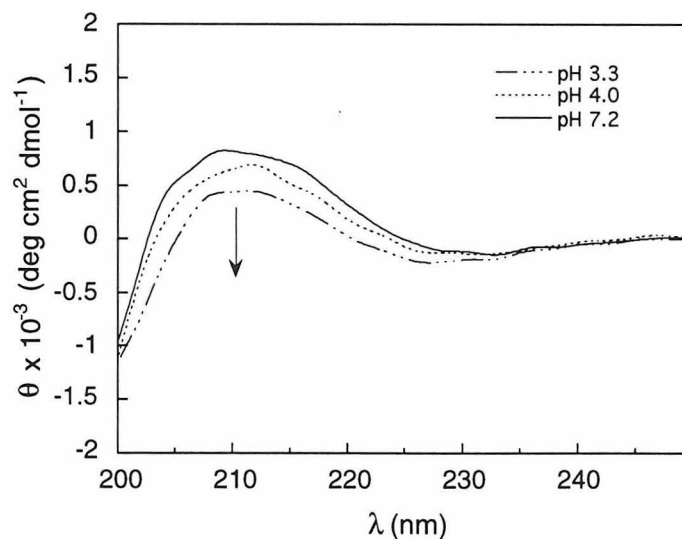


Figure A-5 The CD spectra of peptide **GG1** at various pH.

A type III turn is equivalent to a single turn of a 3_{10} helix,¹⁰ which exhibits a $C=O(i) \rightarrow NH(i+3)$ hydrogen bonding pattern. Reduced amide proton temperature coefficients are therefore anticipated for position $(i+3)$ and perhaps the $(i+2)$ amide protons (due to interaction with the carbonyl of the acetyl cap). In contrast, the amide protons at the (i) and $(i+1)$ positions are expected to be solvent exposed and have correspondingly higher amide proton temperature coefficients. Similar to the data obtained for peptide **PA1**, all of the amide proton temperature coefficients ($-\Delta\delta/\Delta T$) for peptide **GG1** are high (> 7 ppb/K), potentially due to the presence of charged residues.⁹ Although the amide proton temperature coefficients of **GG1** were too large to be conclusive of hydrogen bonding interactions ($-\Delta\delta/\Delta T > 6$ ppb/K), the *relative* values for the residues at positions $(i+2)$ and $(i+3)$ supported the hypothesis that this peptide adopts a type III or 3_{10} helical conformation in solution. Coefficients for the $(i+2)$ and $(i+3)$ amide protons were 1-2 ppb/K lower than those observed for residues at positions (i) and $(i+1)$ (Table A-4). This data, in conjunction with the two - dimensional NMR and CD analyses, indicated that a significant population of **GG1** adopted a folded conformation in water which resembled a type III reverse turn or a 3_{10} helix.

Position	Residue	$-\Delta\delta/\Delta T$ (ppb/K)
(i)	Glu ¹	9.8
$(i+1)$	Gly ²	9.9
$(i+2)$	Gly ³	7.4
$(i+3)$	Lys ⁴	8.6

Table A-4 Amide proton variable temperature coefficients for peptide **GG1**.

Although conformational analysis indicated that a significant population of **PA1** does adopt a type II β -turn conformation, tetrapeptide sequences which exhibit stronger type II β -turn forming tendencies have been examined.⁷ Although lysine and glutamic acid have been shown to be effective in forming an efficient salt bridge across the turn of an α -helix in the $(i, i+4)$ positions, the side chains of these residues are probably too long, and have too much conformational freedom to be ideal in stabilizing a turn structure. Significantly, studies of model α -helical peptides indicated that peptides incorporating a Glu(-ve)-Lys(+ve) salt bridge separated by only two residues [i.e., at $(i, i+3)$], were less helical than similar peptides in which these residues were separated by three residues [at $(i, i+4)$].² Analysis of peptide **GG1** indicated that glutamic acid and lysine at positions (i) and $(i+3)$, respectively, was more compatible with a 3_{10} helical conformation than a type II β -turn in the absence of the Pro-DAla sequence, further supporting the conclusion that these residues are not of the ideal side chain length to stabilize a turn conformation. These results also highlight the influence of a Pro-DXaa sequence in directing type II β -turn formation. The obvious preference for the Glu(-ve)-Lys(+ve) pair is to span a distance of approximately 6 Å,^{1, 2} which approximates the pitch of an α -helix or 3_{10} helix, as demonstrated both by the analysis of model α -helical peptides² and by the data presented for peptide **GG1**. Nevertheless if these residues flank a Pro-DXaa sequence, then a type II β -turn is preferred.

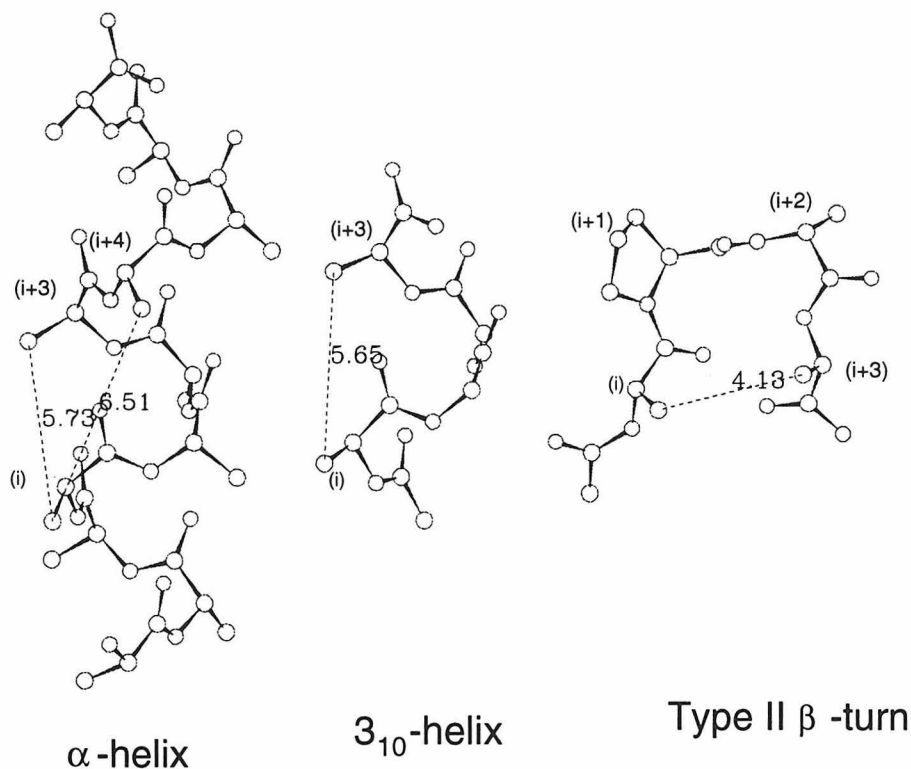


Figure A-6 Comparative C β -C β distances (Å) for an α -helix, 3_{10} helix, and type II β -turn.

The conformational analyses of peptides **PA1** and **GG1** emphasize the potential utility of salt bridge interactions in directing the folding of the peptide backbone, at least for small peptides. Significantly, tetrapeptide **GG1** exhibited NMR and CD evidence of a folded conformation despite the presence of two flexible glycine residues. Unfortunately, the naturally occurring ionizable amino acids are probably not of the ideal side chain length to stabilize sheet and turn structure through inter-strand salt bridges. Therefore, the synthesis of an unnatural amino acid which might be more amenable to this application was undertaken.

***α*Maa: A Non-Standard Amino Acid to Stabilize Sheet and Turn Structure**

The approach to protein design adopted by the Imperiali group and described in Chapter 1 utilizes non-standard and unnatural amino acids to control polypeptide structure and incorporate function. As an extension of this concept, the potential utility of (2*S*,3*S*)-diaminobutyric acid (*allo*-β-methyl-β-amino-alanine or *α*Maa, **1**) to stabilize sheet and turn structures through steric and electrostatic interactions was investigated.

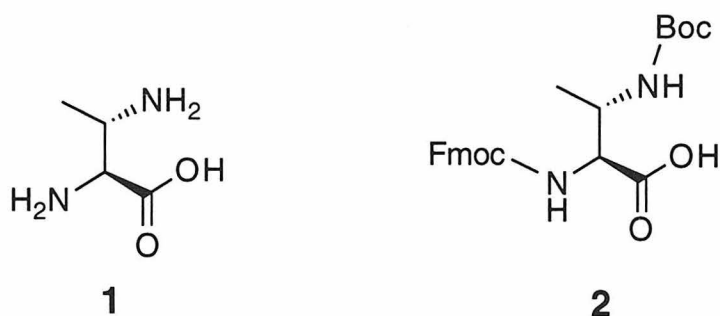


Figure A-7 Structure of *α*Maa (**1**) and the corresponding derivative, Fmoc-L-*α*Maa(Boc)-OH (**2**) protected for solid phase peptide synthesis.

The propensity of β-branched amino acids to form β-sheet structure has been demonstrated by both investigations of the crystal structures of proteins,¹⁴ and experimental investigations of model β-sheet systems.¹⁵⁻¹⁷ The 20 naturally occurring amino acids present a limited choice in this regard, consisting of mainly hydrophobic residues (Val, Ile). Utilization of the natural β-branched residues in the design of stable sheet structures can result in poor solubility and aggregation of the resulting peptides due to the hydrophobic nature of these residues. The non-standard, ionizable amino acid *α*Maa was synthesized to overcome these problems. Since this amino acid is both β-branched and

hydrophilic, it was envisioned to be useful in the design of structures incorporating amphiphilic sheets. Alternating with hydrophobic residues in a β -sheet structure, *a*Maa would create the hydrophilic face of an amphiphilic sheet, increasing the solubility of the peptide in water. Furthermore, the length of the *a*Maa side chain makes it ideal for inter-strand salt bridge formation with appropriately placed aspartate residues on an adjacent strand (Figure A-8). When used in conjunction with a nucleating turn, these salt bridges could serve to 'zip' together two strands of a sheet, propagating sheet formation in the creation of a β -hairpin structure.

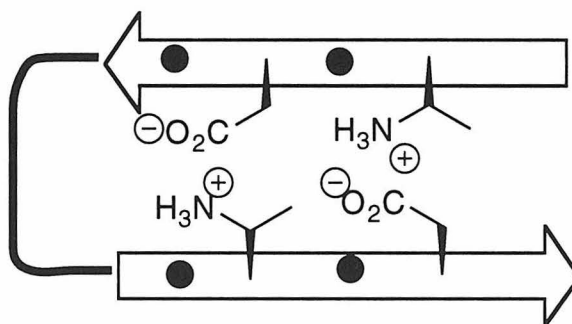


Figure A-8 Schematic representation of an amphiphilic β -hairpin employing inter-strand *a*Maa-Asp salt bridges; • indicates a hydrophobic residue.

The synthesis of N^{α} -9-fluorenylmethoxycarbonyl- N^{β} -*tert*-butyloxycarbonyl-(2*S*,3*S*)-diaminobutyric acid, **2**, has already been published as a step in the synthesis of Lavandomycin, an antibiotic hexapeptide by Schmidt *et al.*¹⁸ Some problems were encountered in reproducing the reported synthetic protocols and a modified synthesis scheme was pursued (Figure A-9). Following protection of the amine and carboxylate groups of threonine, a Mitsunobu

reaction¹⁹ was used to convert the β -alcohol of **4** to the azide (**5**). Attempts to reduce azide **5** by the method reported by Schmidt¹⁸ (hydrogenation with Lindlar catalyst)²⁰ were unsuccessful. Instead reduction of azide **5** to the amine **6** was effected with H₂S in water and pyridine.²¹ After removal of the carboxylate protecting group, the free side chain amine is protected with a *tert*-butoxycarbonyl (Boc) protecting group to yield the suitably protected amino acid, **2**. This residue was successfully incorporated into a peptide by solid phase peptide synthesis.

To probe the turn and sheet-stabilizing properties of *a*Maa, the peptide Ac-Asp-Val-Pro-D-Ser-Phe-*a*Maa-NH₂ (**MD1**) was synthesized. In this design *a*Maa and Asp were incorporated flanking the type II β -turn-forming sequence, Val-Pro-D-Ser-Phe,¹² at positions (*i*-1) and (*i*+4). It was anticipated that salt-bridge formation between these residues would serve to propagate the sheet formation initiated by the type II β -turn.

2-D NMR analysis of peptide **MD1** revealed a number of NOEs characteristic of a type II β -turn.^{7, 10} The results of this study is summarized in Table A-5 and Figure A-10. The amide proton temperature were consistent with the NOE data. A depressed amide proton temperature coefficient for the (*i*+3) amide proton was observed ($-\Delta\delta/\Delta T < 6.0$ ppb/K). Additionally, strong $d_{\alpha N}$ (*i*, *i*+1) NOEs for residues Asp¹-Val² and Phe⁵-*a*Maa⁶ suggested an extended or sheet-like conformation for these residues.¹⁰

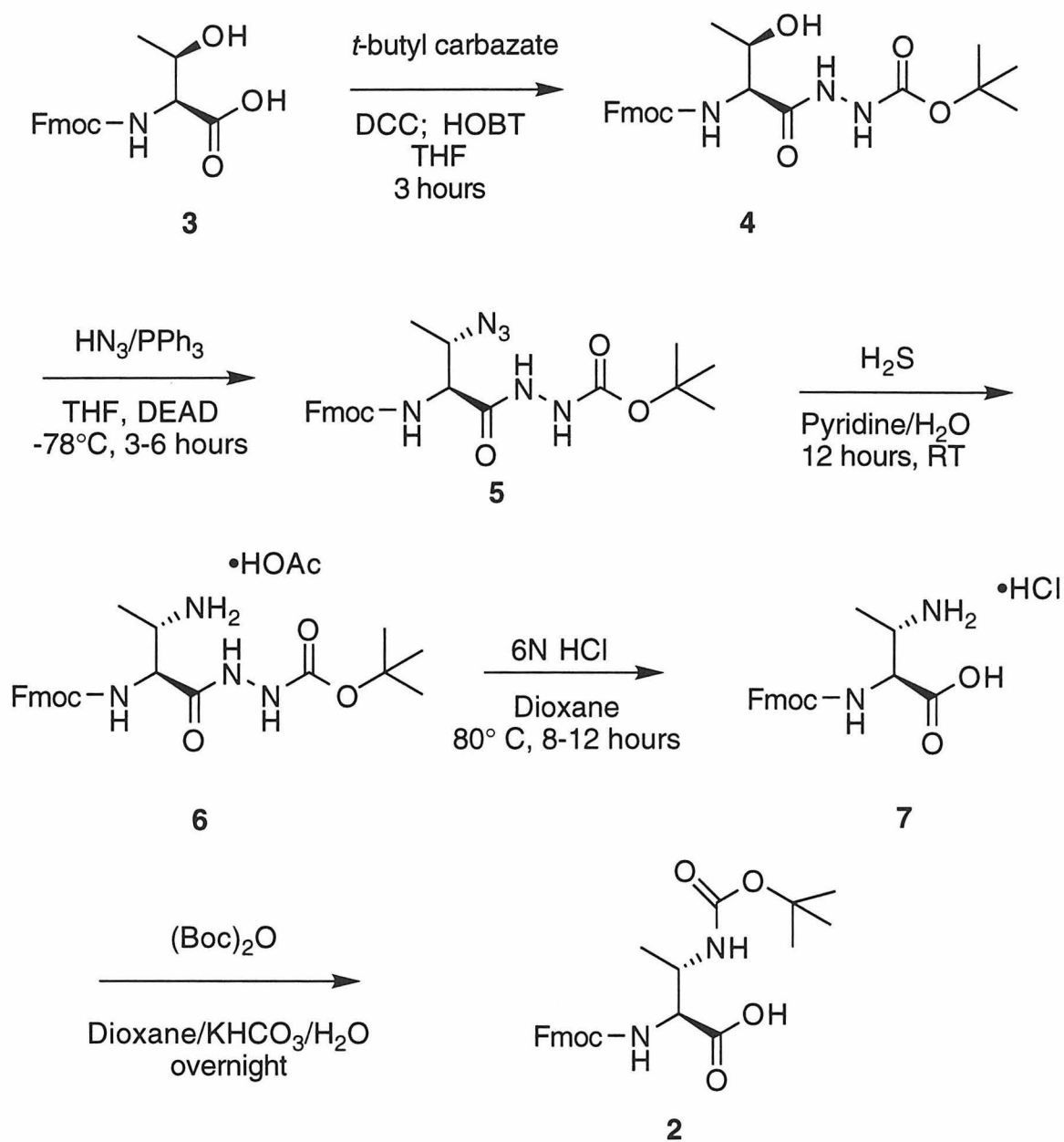


Figure A-9 Synthesis of Fmoc- α Maa(Boc)-OH (2).

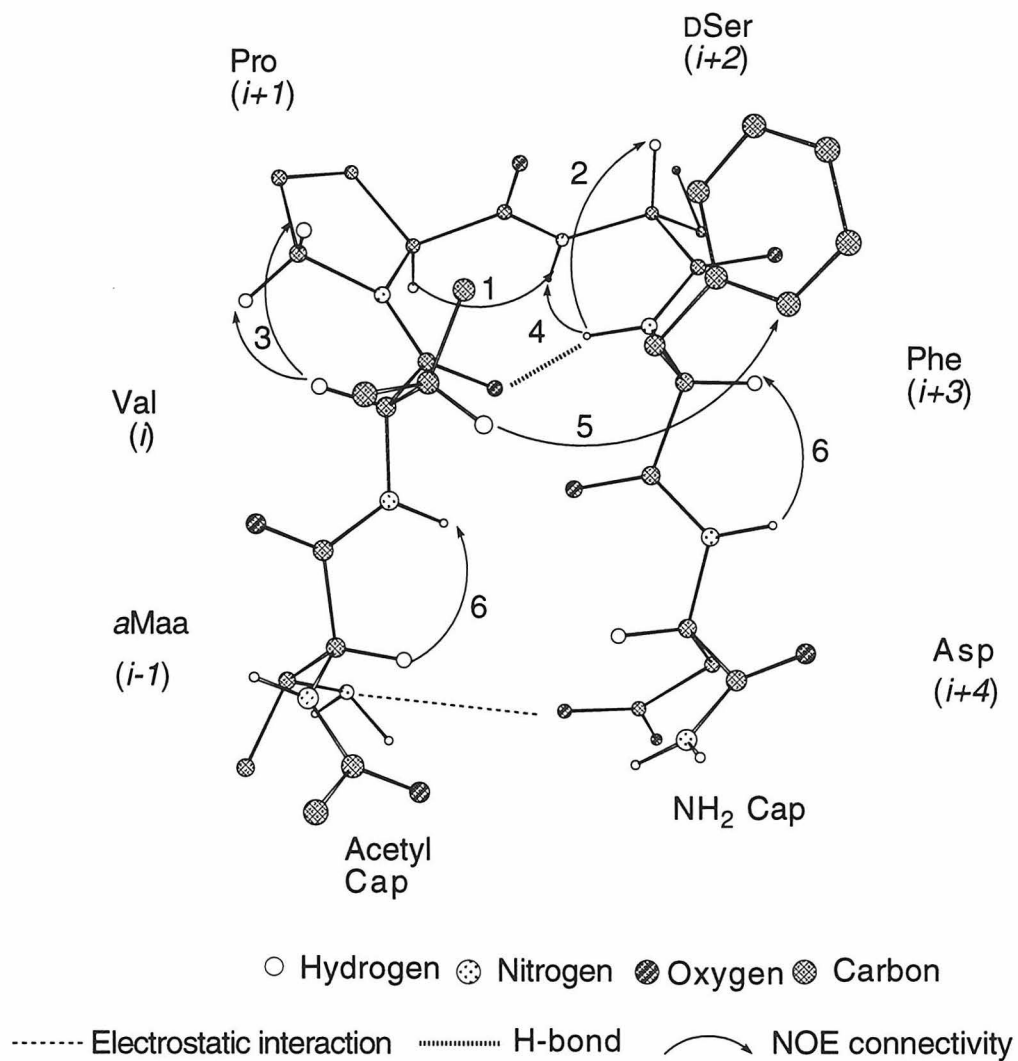


Figure A-10 Summary of NOEs observed in the ROESY spectrum of **MD1**. The peptide is illustrated in the proposed β -hairpin conformation. The NOEs are listed in Table A-5.

NOE #	Assignment	Relative intensity
1	$d_{\alpha N}(\text{Pro}^3, \text{Ser}^4)$	strong
2	$d_{\alpha N}(\text{Ser}^4, \text{Phe}^5)$	strong
3	$d_{\alpha\delta}(\text{Val}^2, \text{Pro}^3)$	medium
4	$d_{NN}(\text{Ser}^4, \text{Phe}^5)$	weak
5	$d_{\gamma,\delta/\epsilon}(\text{Val}^2, \text{Phe}^5)$	very weak
6, 6'	$d_{\alpha N}(i, i+1)$	strong

Table A-5 Summary of NOEs indicative of turn and sheet structure observed in the ROESY spectrum of peptide **MD1**.

The potential role of the unnatural amino acid, *a*Maa, in stabilizing the type II β -turn structure observed for peptide **MD1** was investigated by comparing the results of the NMR experiments performed on **MD1** with those obtained for the tetrapeptide Ac-Val-Pro-D-Ser-Phe-NH₂⁷ (VPDSF, Table A-6 and A-7). **VPDSF** exhibited fewer NOEs indicative of turn formation and a higher (*i*+3) amide proton temperature coefficient compared to peptide **MD1** (Table A-5, Table A-6).⁷ These differences indicate that the inclusion of Asp and *a*Maa residues, flanking the turn sequence, resulted in a more compact folded turn structure. In particular, peptide **MD1** exhibited an NOE interaction between the side chain protons of Val² and Phe⁵ which was not observed in the tetrapeptide sequence. The more compact turn structure of peptide **MD1** might result from a stabilizing electrostatic interaction between the charged residues (*a*Maa¹ and Asp⁶).

Peptide	Sequence	NOEs
MD1	Ac- <i>a</i> Maa-Val-Pro-DSer-Phe-Asp-NH ₂	1,2,3,4,5
KD1	Ac-Lys-Val-Pro-DSer-Phe-Asp-NH ₂	1,2,3,5
VPDSF	Ac-Val-Pro-DSer-Phe-Asp-NH ₂	1,2,3,4

Table A-6 Comparison of the NOEs observed in the ROESY experiments of peptides **MD1**, **KD1**, and **VPDSF**. NOEs correspond to those illustrated in Figure A-1.

Peptide	Sequence	$-\Delta\delta/\Delta T$ (ppb/K)				
		(<i>i</i> -1)	(<i>i</i>)	(<i>i</i> +2)	(<i>i</i> +3)	(<i>i</i> +4)
MD1	Ac- <i>a</i> Maa-Val-Pro-DSer-Phe-Asp-NH ₂	7.0	8.2	10.9	5.6	5.8
KD1	Ac-Lys-Val-Pro-DSer-Phe-Asp-NH ₂	7.8	10.5	9.5	5.9	6.3
VPDSF	Ac-Val-Pro-DSer-Phe-Asp-NH ₂	--	8.7	8.8	5.8	--

Table A-7 Comparison of the amide proton temperature coefficients for peptides **MD1**, **KD1**, and **VPDSF**.

The NMR data obtained for peptide **MD1** was also compared to that obtained for peptide **KD1** (Ac-Asp-Pro-DSer-Phe-Lys-NH₂).²² Significantly, the $d_{NN(i+2, i+3)}$ NOE which is particularly indicative of type II β -turn formation in short peptides^{7, 10} is *not* observed in the ROESY spectrum of peptide **KD1**. Surprisingly NOE #5 [$d_{\beta, \delta/\epsilon}(\text{Val}^2, \text{Phe}^5)$] was observed for **KD1**, indicating association of the side chains of the residues at positions (*i*) and (*i*+3). This

peptide differs from **MD1** only in the substitution of lysine for the non-standard amino acid *a*Maa. Peptide **KD1** is still capable of forming the $(i-1) - (i+4)$ electrostatic interaction, although it was anticipated that this electrostatic interaction would be less favorable than the shorter salt bridge of the **MD1** design. The observation of NOE #5 [$d_{\beta,\delta/\epsilon}(\text{Val}^2, \text{Phe}^5)$] in the absence of NOE #4 [$d_{\text{NN}}(\text{Ser}^4, \text{Phe}^5)$] suggests that although the Lys-Asp salt-bridge may be stabilizing a reversal of the peptide backbone for **KD1**, the peptide *does not* adopt an ideal type II β -turn geometry. Additionally, the amide proton temperature coefficient for position $(i+3)$ is higher for peptide **KD1** than for peptide **MD1**. These results indicate that *a*Maa residue is more appropriate than the naturally occurring alternatives for electrostatic stabilization of a β -hairpin containing the desired type II β -turn.

Conclusion

Conformational analyses of the peptides **PA1**, **GG1**, and **MD1** demonstrate the possibility of stabilizing folded conformations with electrostatic interactions. In particular the ability of a non-standard amino acid, *a*Maa, to stabilize type II turn formation was highlighted by a comparison of the solution conformations of peptides **MD1** and **VPDSF**. The ability of defined β -turns to nucleate structure within designed polypeptides was demonstrated in previous chapters and has been extensively exploited within the Imperiali group. These studies point to the potential for augmenting the utilization of these turns with stabilizing electrostatic interactions. Therefore, the exploitation of salt-bridges represents yet another protein design tool for the creation of polypeptides with preprogrammed structure.

Experimental

The synthesis of N^{α} -fluorenyloxycarbonyl- N^{β} -*tert*-butoxycarbonyl-(2*S*,3*S*)-diaminobutyric acid [Fmoc-*a*Maa(Boc)-OH] was accomplished by an adaptation of the protocols reported by Schmidt and coworkers¹⁸ (Scheme A-1).

N^{α} -9-Fluorenylmethoxycarbonyl-L-threonine (3)

L-Threonine (42.2 mmoles) was dissolved in 45 ml of 10% NaCO₃(w/v) and stirred at 0 °C. 9-fluorenylmethyloxycarbonyl succinimidyl carbonate (Fmoc-OSu) was added to 55 ml of dioxane to form a slurry which was then added dropwise to the cooled (0 °C) threonine solution. After complete addition of the Fmoc-OSu the reaction mixture was allowed to warm to room temperature, and stirred for 2hrs. The dioxane was removed *in vacuo* and the reaction mixture was diluted with water and washed with ether (3x 400ml). The aqueous phase was cooled to 0 °C and the pH adjusted to 3.0 with 6 N HCl. The resulting suspension was extracted with ethyl acetate (3x 200ml). The ethyl acetate phase was dried to yield 10.64 grams of white solid (74% yield).

¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.2-7.8 (m, 8H), 6.7 (s, broad), 6.4 and 6.1 (both doublets, 1H, alternate conformations), 4.4-4.5 (m, 4H), 4.2 (t, 1H), 1.0-1.25 (set of doublets, 3H, alternate conformations)

N^{α} -9-Fluorenylmethoxycarbonyl-(2*S*,3*R*)-threonine- N' -*tert*-butyl-oxycarbonyl-hydrazide (Fmoc-Thr-NHNHBoc) (4).

Fmoc-amino acid 3 (1.46 mmole), dicyclohexyl carbodiimide (1.46 mmole), and *t*-butyl carbazate were dissolved in dry THF and stirred at room temperature for 4 hours. A white solid (dicyclohexyl urea, DCU) which precipitated from the reaction mixture was filtered and the solution concentrate *in vacuo*. Small scale reactions were purified by flash chromatography in 1:1 CHCl₃/EtOAc in a 92 %

yield. Large scale reactions were purified by repeated precipitation from EtOAc to recover 40-60 % of pure **4**. The remaining product in the concentrated supernate was purified by flash chromatography in 1:1 CHCl₃/EtOAc.

Recovered yield: 60-92% depending on scale.

R_f=0.31, 1:1 CHCl₃/EtOAc; R_f=0.36, 20:1 CHCl₃/MeOH

¹H-NMR (300 MHz, CDCl₃) δ (ppm): 8.6 (s, 1H, broad), 7.3-7.75 (m, 8H), 6.9 (s, 1H), 6.1 (d, 1H), 4.25-4.5 (m, 4H), 4.1 (m, 1H), 1.45 (s, 9H), 1.2 (d, 3H)

¹³C-NMR (300 MHz, CDCl₃) δ (ppm): 18.5, 28.3, 47.3, 59, 67.5, 82.5, 120.2, 125.3, 127.3, 127.9, 141.5, 143.8, 158, 169, 171.4

***N*^α-9-Fluorenylmethoxycarbonyl-(2*S*,3*S*)-2-amino-3-azido-butyryl-*N*'-tert-butylloxycarbonyl-hydrazide (**5**).**

To a cooled solution (-78 °C) of triphenylphosphine (PPh₃) (2.8 mmole) the following were added in order: 89.5 μl DEAD (diethyl azodicarboxylate, 2.8 mmole), 6 ml HN₃ in C₆H₆ (0.48 M), and **4** (2.6 mmole in 1ml dry THF). The reaction was warmed to room temperature and stirred for 3-6 hours. The crude reaction mixture was concentrated in *vacuo* and purified by flash chromatography in 35:65 EtOAc/Hexane. The presence of the azide was confirmed by observation of a peak at 2100 cm⁻¹ in the IR spectra.

Yield 30 - 60 %

R_f=0.65, 1:1 EtOAc/CHCl₃; R_f=0.31 35:65, EtOAc/Hexane

¹H-NMR (300MHz, CDCl₃) δ (ppm): 1.2 (d, 3H), 1.35 (s, 9H), 3.9 (m, 1H), 4.2 (m, 1H), 4.3 (m, 1H), 4.5 (m, 2H), 6.0 (d, 1H), 7.2-7.4 (m, 4H) 7.6 (d, 2H), 7.8 (d, 2H), 8.9 (s, 1H).

***N*^α-Fluorenylmethoxycarbonyl-(2*S*,3*S*)-diamino-butyryl-*N*'-tert-butylloxycarbonyl-hydrazide (**6**).**

Protected amino acid **5** (2.5 mmole) was dissolved in pyridine (48 ml) and 48 ml of water added. H₂S was bubbled slowly through the stirred reaction mixture over 12 hours. The reaction mixture was neutralized with acetic acid, concentrated in *vacuo*, and lyophilized to dryness. The resulting solid was dissolved in methanol and filtered through celite. The solution was concentrated in *vacuo* to yield the product **6**. The reduction of the azide was verified by IR spectroscopy.

Yield: 100 %

¹H-NMR (300MHz, CD₃OD) δ (ppm): 1.2 (d, 3H), 1.4 (s, 9H), 1.9 (s, 3H) 3.6 (m, 1H), 4.2 (m, 1H), 4.4-4.5 (m, 3H), 6.0 (d, 1H), 7.2-7.4 (m, 4H) 7.64 (m, 2H), 7.75 (m, 2H), 8.5 (m, 1H).

N^α-9-Fluorenylmethoxycarbonyl-(2S,3S)-diamino-butyric acid, (Fmoc-Maa-OH) (7)

Protected amino acid **6** (2.5 mmole) was dissolved in 4.6 mls of dioxane and 9.3 mls of 6N HCl was added. The reaction mixture was stirred at 80 °C for 8-12 hours. The acid was removed from the reaction mix by repeated dilution with water followed by concentration in *vacuo*. Yield on this reaction was quantitative.

¹H NMR (300MHz, CD₃OD) δ (ppm): 1.25 (d, 3H); 3.8 (m, 1H); 4.3 (d, 1H); 4.5 (m, 2H); 4.65 (m, 1H); 7.2-7.4 (m, 4H); 7.7-7.8 (m, 4H)

¹³C NMR (CD₃OD) δ (ppm): 23.5, 55.0, 56.1, 67.7, 120.1, 125.4, 127.3, 128.0, 141.7, 144.3

***N*^αFluorenylmethoxycarbonyl-*N*^β-*tert*.-bytryloxycarbonyl-(2*S*,3*S*)-diaminobutyric acid (2).**

Amino acid **7** (1.3 mmoles) was brought up in 14 ml dioxane/water/1*N* KHCO₃ (2:1:1) to form a slurry. Boc₂O (di-*tert*-butylpyrocarbonate, 1.3 mmole) was added dropwise to the reaction mixture. The reaction was stirred at room temperature overnight. The dioxane was removed *in vacuo* and the reaction mixture diluted with water and washed with ether. Addition of ether produces a gel-like aqueous phase which is washed 3 times with ether. The aqueous phase is acidified with citric acid and the product extracted into ethyl acetate. The ethyl acetate phase was dried over sodium sulfate and concentrated *in vacuo* to yield a slightly beige product.

Yield 37 %

R_f = 0.36, 5:1 CHCl₃/MeOH

¹H NMR (300MHz, DMSO-*d*₆) δ (ppm): 1.0 (d, 3H); 1.33 (s, 9H); 3.9 (m, 1H); 4.1-4.3 (m, 4H); 6.7 (d, 1H); 7.2-7.4 (m, 4H), 7.5 (d, 1H); 7.7 (d, 2H); 7.9 (d, 2H)

¹³C NMR (CD₃OD) δ (ppm): 15.4, 27.9, 43.0, 58.0, 67.4, 73.3, 120.1, 125.4, 127.3, 127.9, 141.7, 144.4, 172.7

Peptide Synthesis

Peptides **PD1**, **GG1**, and **MD1** were synthesized by solid phase methods on a Milligen 9050 automated peptide synthesizer using *N*^α-9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids and BOP/HOBT or DIPCDI/HOBT amide coupling chemistry.²³ Syntheses were performed on a 0.126 mmole of peptide using Fmoc-Pal-PS resin (0.31 meq/gram) and glass beads. Activated esters were formed *in situ* using either BOP, HOBT, and 0.451 M *N*-methyl-morpholine in DMF or DIPCDI and HOBT in 1:1 DMF:DCM. Fmoc deprotection was accomplished by washing the resin for 7 min with either 20%

piperidine in DMF or 2%(v/v) 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU).²⁴ Typical recycle times were 1 to 1.5 hours using a 3 or 4 fold excess of amino acid. Double coupling procedures were used for amino acids which have been reported to be difficult to couple.²⁵ The amino terminal of the peptide was N-acetylated while on the resin by treatment with 20 equivalents of acetic anhydride and 5 equivalents of triethylamine in DMF. After shaking for 3 hours the resin was washed with DMF, dichloromethane, and methanol, and vacuum dried.

Deprotection and cleavage of the peptides was accomplished by treatment with 5-10 mls of Reagent R²⁶ (trifluoroacetic acid/thioanisole/ethane dithiol (EDT)/anisole (90:5:3:2)) for 2-3 hours. The resin was filtered, washed with TFA and the filtrate collected and concentrated under a stream of nitrogen. The peptides were precipitated and triturated with ether/hexane 2:1. The peptides were then dissolved in a minimum quantity of water and lyophilized to dryness. Peptide **PA1** and **GG1** did not require further purification. Peptide **MD1** was purified by size exclusion chromatography (P2) using 50 mM acetic acid as the running solvent. The peptide eluted in 12 minutes, under gravity

NMR Studies

All NMR data were taken on either a Bruker AM or a Bruker AMX 500 MHz instrument. An initial 1-D spectra were taken in DMSO-*d*₆ to access purity, all other spectra including two-dimensional experiments were taken in 90:10 H₂O/D₂O with typical peptide concentrations of 10mg/0.5ml. Phase sensitive COSY (peptide **PA1**) and TOCSY²⁷ experiments were used to assign spin systems.¹⁰ Sequence specific assignments and conformational analysis were accomplished on the basis of NOE information obtained from a ROESY²⁸ experiment. Presaturation of the water signal was accomplished during the

relaxation delay (1.3 s). All 2D spectra were recorded at 282 K or 285.5 K. The temperature dependence of the amide proton chemical shifts was determined by acquisition of a 1D experiment at 285, 290, 295, 300, 305, and 310 K. All spectra were processed on Felix. FIDs were multiplied by a phase shifted (90°, TOCSY & ROESY) (0°, COSY) sine bell apodization function prior to Fourier transformation.

CD studies

CD studies were performed on a Jasco J-600 Spectropolarimeter. Peptide samples were prepared in distilled water at an initial concentration of 10 - 51 μ M, and spectra recorded in 1cm pathlength cells. For the variable temperature study of peptide **PA1**, the sample was degassed by sparging with nitrogen for 30 min. Variable temperature studies were performed in a water jacketed cell using a circulating temperature controlled bath and CD spectra were acquired from 185 nm to 250 nm on the same sample in 5 °C increments from 5 °C to 50 °C. The pH dependence of the CD spectra of peptides **PA1** and **GG1** was determined using a 10 - 50 μ M peptide sample. The pH of the sample was adjusted with small additions of dilute NaOH and HCl. CD spectra were taken from 200 nm to 250 nm at room temperature. Eight scans were obtained for all spectra at a scan speed of 50 nm/min, a bandwidth of 1.0 nm and a time constant of 0.5 sec. The concentration of the sample was adjusted for each peptide depending on signal strength. All spectra were normalized to molar ellipticity. Data was processed with KaleidaGraph software on a Macintosh IIfx.

Peptide PA1

^1H NMR (500MHz, d_6 -DMSO) δ (ppm): 8.2 (d, 1H), 8.06 (d, 1H), 7.8 (d, 1H), 7.7 (broad singlet), 7.26 (s, 1H), 7.04 (s, 1H), 4.5 (m, 1H), 4.2 (m, 2H), 4.1 (m, 1H), 3.5-3.7 (m, 4H), 2.7 (broad singlet), 2.3 (m, 2H), 1.9 (m, 1H), 1.8 (m, 1H), 1.5-1.75 (m, 6H), 1.4 (m, 3H), 1.2 (m, 5H)

Residue	alpha	amide	side chain
Acetyl cap	2.00		
Glu ¹	4.58	8.48	2.03, 1.87, 2.38
Pro ²	4.40		3.88, 3.70, 2.06, 1.94, 1.30
DAla ³	4.31	8.72	1.41
Lys ⁴	4.31	8.27	1.79, 1.90, 1.66, 1.41, 2.99
Amide cap		7.29, 7.71	

Table A-8 Assignment table of the ^1H NMR spectrum of **PA1** (δ , ppm). Spectra were referenced to the acetyl cap at 2.0 ppm (water signal at 5.0 ppm) (500 MHz, 90:10 $\text{H}_2\text{O}/\text{D}_2\text{O}$, pH 4.4)

Peptide GG1

^1H NMR (500 MHz, DMSO- d_6) δ (ppm): 8.25 (t, 1H), 8.10 (m, 2H), 7.8 (d, 1H), 7.7 (broad singlet), 7.3 (s, 1H), 7.0 (s, 1H), 4.2 (m, 1H), 4.13, (m, 1H), 3.70, (m, 4H), 2.75, (s, broad, 2H), 2.24 (t, 2H), 1.8 (m, 4H), 1.7 (m, 2H), 1.5 (m, 4H), 1.25 (m 2H)

Residue	alpha	amide	side chain
Acetyl cap	2.00		
Glu ¹	4.24	8.49	2.05, 1.91, 2.33
Gly ²	3.91	8.70	
Gly ³	3.91	8.36	
Lys ⁴	4.26	8.20	1.83, 1.72, 1.39, 1.62, 2.94
Amide cap		7.16, 7.64	

Table A-9 Assignment table of the ^1H NMR spectrum of **GG1** (δ , ppm). Spectra were referenced to the acetyl cap at 2.0 ppm (water signal at 5.0 ppm) (500 MHz, 285.5, 90:10 H₂O/D₂O, pH 4.4)

Peptide MD1

Residue	alpha	amide	Side chain
Acetyl cap	2.00		
Maa ¹	4.80	8.59	β 3.8; γ 1.18
Val ²	4.40	8.52	β 2.0; γ 0.88, 0.95
Pro ³	4.39		β 1.86, 2.26; γ 2.04, 1.94; δ 3.89, 3.6
DSer ⁴	4.33	8.64	β 3.75, 3.69
Phe ⁵	4.54	8.28	β 2.99, 3.09; ring 7.20
Asp ⁶	4.50	8.39	β 2.61, 2.56

Table A-10. Assignment table of the ^1H NMR spectrum **MD1** (δ , ppm). Spectra were referenced to the acetyl cap at 2.0 ppm (water signal at 5.0 ppm) (500 MHz, 282 K, 90:10 $\text{H}_2\text{O}/\text{D}_2\text{O}$, pH 4.4)

References

- (1) Merutka, G.; Stellwagen, E. "Effect of Amino Acid Ion Pairs on Peptide Helicity," *Biochemistry*, **1991**, 30, 1591-1594.
- (2) Marqusee, S.; Baldwin, R. L. "Helix Stabilization by $\text{Glu}^- \dots \text{Lys}^+$ Salt Bridges in Short Peptides of *de novo* design.," *Proc. Natl. Acad. Sci. U.S.A.*, **1987**, 84,
- (3) Regan, L.; DeGrado, W. F. "Characterization of a Helical Protein Designed from First Principles," *Science*, **1988**, 241, 976-978.
- (4) Otter, A.; Scott, P. G.; Liu, X.; Kotovych, G. "A ^1H and ^{13}C NMR Study on the Role of Salt-Bridges in the Formation of a Type I β -Turn in N-Acetyl-L-Asp-L-Gly-L-Lys-L-Ser- NH_2 ," *Biomol. Struct. Dynam.*, **1989**, 7, 455-476.

- (5) Liu, X.; Scott, P. G.; Otter, A.; Kotovych, G. "A Sequence-Dependent ^1H -NMR Study of the Formation of β -Turns in Tetrapeptides Containing Charged Residues," *Biopolymers*, **1992**, 32, 119-130.
- (6) Mayer, R.; Lancelot, G. "Conformational Study of the Tetrapeptide Boc-Arg-Ala-Gly-Glu-NHEt. A β Turn Locked by a Salt Bridge," *J. Am. Chem. Soc.*, **1981**, 103, 4738-4742.
- (7) Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: Spectroscopic Evidence for a Peptide with Significant β -Turn Character in Water and Dimethyl Sulfoxide," *J. Am. Chem. Soc.*, **1992**, 114, 3182.
- (8) Chou, P. Y.; Gasman, G. D. " β -Turns in Proteins," *J. Mol. Biol.*, **1977**, 115, 135-175.
- (9) Fisher, S. L., "Unnatural Amino Acids in the Synthesis and Semisynthesis of Metalloprotein Motifs," Ph.D. Thesis, California Institute of Technology, 1993.
- (10) Wüthrich, D. *NMR of Proteins and Nucleic Acids*; John Wiley and sons: New York, 1986.
- (11) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. "Folding of Immunogenic Peptide Fragments of Proteins in Water Solution," *J. Mol. Biol.*, **1988**, 201, 161-200.
- (12) Rose, G. D.; Gierasch, L. M.; Smith, J. A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.*, **1985**, 37, 1-109.
- (13) Woody, R. W. "Studies of Theoretical Circular Dichroism of Polypeptides: Contributions of β Turns," in *Peptides, Polypeptides, and Proteins*, E.R. Blout, F. A. Bovey, M. Goodman, N. Lotan Eds.; Wiley, New York, 1974; pp 338-350.
- (14) Fasman, G.D. "The Development of the Prediction of Protein Structure," in *Prediction of Protein Structure and the Principles of Protein Conformation*, Fasman, G.D. Eds.; Plenum Press, New York, 1989; pp 193-316.

- (15) Otzen, D. E.; Fersht, A. R. "Side-Chain Determinants of β -Sheet Stability," *Biochemistry*, **1995**, 34, 5718-5724.
- (16) Minor, D. L.; Kim, P. S. "Measurement of the β -sheet-forming Propensities of Amino Acids," *Nature*, **1994**, 367, 660-663.
- (17) Kim, C. A.; Berg, J. M. "Thermodynamic β -Sheet Propensities Measured Using a Zinc-finger Host Peptide," *Nature*, **1993**, 362, 267-270.
- (18) Schmidt, U.; Mundinger, K.; Mangold, R.; Lieberknecht, A. "Lavendomycin: Total Synthesis and Assignment of Configuration," *J. Chem. Soc. Chem. Comm.*, **1990**, 1216-1219.
- (19) Mitsunobu, O. "The Use of Diethyl Azodicarboxylate and Triphenylphosphine in Synthesis and Transformation of Natural Products," *Synthesis*, **1981**, 1-28.
- (20) Corey, E. J.; Nicolaou, K. C.; Blanson, R. D.; Machida, Y. "A Useful Method for the Conversion of Azides to Amines," *Synthesis*, **1975**, 590-591.
- (21) Adachi, R.; Yamada, Y.; Inoue, I. "An Alternative Method for the Selective Reduction of Unsaturated Nucleoside Azides to the Amines," *Synthesis*, **1977**, 45-46.
- (22) Sinha Roy, R.; Imperiali, B., unpublished results
- (23) Bodanszky, M. *Principles of Peptide Synthesis*; Springer-Verlag: Berlin, 1993.
- (24) Wade, J. D.; Bedford, J.; Sheppard, R. C.; Tregear, G. W. "DBU as an N^α -Deprotecting Reagent for the Fluorenyl-methoxycarbonyl Group in Continuous Flow Solid-Phase Peptide Synthesis," *Peptide Research*, **1991**, 4, 184-199.
- (25) Van Woerkom, W. J.; Van Nispen, J. W. "Difficult Couplings in Stepwise Solid Phase Peptide Synthesis: Predictable or Just a Guess?," *Int. J. Pep. Protein Res.*, **1991**, 38, 103-113.
- (26) Biosearch Technical Bulletin No. 9000-02

- (27) Bax, A.; Davis, D. G. "MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy," *J. Magn. Res.*, **1985**, 65,
- (28) Bax, A.; Davis, D. G. "Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy," *J. Magn. Res.*, **1985**, 63, 207-213.