UNNATURAL AMINO ACIDS IN THE SYNTHESIS AND SEMISYNTHESIS OF METALLOPROTEIN MOTIFS

Thesis by

Stewart L. Fisher

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1994

(December 17, 1993)

to Susan and my parents

Acknowledgments

I would like to thank Professor Barbara Imperiali for her guidance and support through my graduate career. The research that I have carried out under her supervision has been an inspiration for me and the many successes I have enjoyed over the past four years are a direct result of her remarkable insight. I would also like to thank Professor Harry Gray both for his strong support and his helpful advice. The collaborative efforts undertaken with him and Dr. Deborah Wuttke comprise some of the most satisfying work that I have done over the past four years.

I am greatly indebted to the members of the Imperiali group, both past and present, for their support and helpful discussions. Tom Prins, Dr. Rex Moats, Ranabir Sinha Roy, Mary Struthers, Tamara Hendrickson, Rahul Pathak, Keith Rickert, Richard Cheng, Grant Walkup, Robert Davis, and Dr. Lisa Wang deserve particular note as their ready willingness to lend a hand and provide valuable scientific input will always be remembered. Dr. Jeff Spencer played a heroic role during the final stages of my tenure at Caltech; his critical reviews of this manuscript and my seemingly endless practice talks are greatly appreciated. I would also like to thank the members of the Dougherty, Bercaw, Dervan, Grubbs, and Gray (particularly Bill Connick, Dr. Deborah Wuttke, Dr. Angelo DeBilio, and Dr. Mike Hill) groups for their helpful advice and support.

Over the past four years, I have made a number of long lasting friendships that will continue far beyond the confines of the Caltech campus. I could not have asked for a better roommate than Pete Beal, whose optimistic outlook on life has always been refreshing. Jack Hwang is a great friend and excellent skiing partner, as are Dr. Peter Green, Sakae Suzuki, and Chris Kenyon. The occasional weekend backcountry skiing and rock-climbing excursions with the "Caltech Backcountry Ski Team" made the weekday grind much more bearable. Think snow!

Finally, I would like to thank both my family and Susan for their unending support throughout my graduate studies. Thanks, Susan, for being my best friend.

Over the past four years I have had the good fortune to be supported by a number of predoctoral fellowships; these include a California Institute of Technology Predoctoral Fellowship, a Department of Education Predoctoral Fellowship, and the Amercian Chemical Society Division of Organic Chemistry (American Cyanamid) Predoctoral Fellowship. Funds for general supplies and analyses were provided by the National Science Foundation (CHE-9104445).

Abstract

The design, synthesis, and characterization of two novel metalloprotein motifs is presented. The first project involved the design and construction of a protein motif which was programmed to form a tetradentate metal complex upon the addition of metal cations. The overall structure of the motif was based on a ββ super-secondary structure consisting of a flexible peptide sequence flanked by metal binding regions located at the carboxy and amino termini. The metal binding region near the amino terminus was constructed from a reverse turn motif with two metal ligating residues, (2*R*, 3*R*)-β-methyl-cysteine and histidine. Selection of the peptide sequence for this region was based on the conformational analysis of a series of tetrapeptides designed to form reverse turns in solution.

The stereospecific syntheses of a series of novel bipyridyl- and phenanthrolyl-substituted amino acids was carried out to provide ligands for the carboxy terminus metal binding region. These residues were incorporated into peptide sequences using solid phase peptide synthesis protocols, and metal binding studies indicated that the metal binding properties of these ligands was dictated by the specific regioisomer of the heteroaromatic ring and the peptide primary sequence.

Finally, a peptide containing optimized components for the metal binding regions was prepared to test the ability of the compound to form the desired intramolecular peptide:metal cation complexes. Metal binding studies demonstrated that the peptide formed monomeric complexes with very high metal cation binding affinities and that the two metal binding regions act cooperatively in the metal binding process. The use of these systems in the design of proteins capable of regulating naturally occurring proteins is discussed.

The second project involved the semisynthesis of two horse heart cytochrome *c* mutants incorporating the bipyridyl-amino acids at position 72 of the protein sequence. Structural studies on the proteins indicated that the bipyridyl amino acids had a neglible

effect on the protein structure. One of the mutants was modified with $Ru(bpy)_2^{+2}$ to form a redox-active protein, and the modified protein was found to have enhanced electron transfer properties between the heme and the introduced metal site.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 2

Chapter 3

Chapter 5

Chapter 6

Chapter 7

List of Schemes

Chapter 3

Chapter 4

Appendix I

Abbreviations Used

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

Bn	:	Benzyl
Bmc	:	(2*R*, 3*R*) -β-Methyl-cysteine
4Bpa	:	2-amino-3-(2,2'-bipyrid-4-yl)-propanoic acid
5Bpa	:	2-amino-3-(2,2'-bipyrid-5-yl)-propanoic acid
6Bpa	:	2-amino-3-(2,2'-bipyrid-6-yl)-propanoic acid
2Fen	:	2-amino-3-(1,10-phenanthrol-2-yl)-propanoic acid
Bpy	:	2,2'-bipyridine
FMOC	:	9-fluorenylmethoxycarbonyl
HRMS	:	High Resolution Mass Spectrometry
CD	:	Circular Dichroism
NMR	:	Nuclear Magnetic Resonance
BOP	:	(Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
HOBT	:	1-Hydroxybenzotriazole
DBU	:	1,8-Diazabicyclo[5.4.0]undec-7-ene

General Methods

All reagents were obtained from commercial sources and used without further purification. Where anhydrous conditions were employed, solvents were dried before use following the general procedures reported by Perrin and Armarego.[1]

Routine ^{1}H and ^{13}C Nuclear magnetic resonance spectra were recorded at 300 MHz and 75 MHz respectively on a General Electric QE-300 Spectrometer, or at 500 MHz and 125 MHz respectively on a Bruker AM500 Spectrometer. All spectra were referenced using tetramethylsilane (TMS) or d_6-dimethyl sulfoxide (d_6-DMSO) as internal standards. Optical rotations were recorded at room temperature in a microcell, 1 dm path length at 589 nm (Na lamp) on a Jasco DIP-181 Digital Polarimeter. Mass spectra were taken on a ZAB mass spectrometer in fast atom bombardment (FAB) mode. Infrared spectra were collected on a Perkin-Elmer 1600-Series FTIR Spectrometer. Thin layer chromatography (TLC) was carried out using EM reagents hard TLC plates with fluorescence indicator (SiO_2 60, F-254) or Merck (Al_2O_3 60 F-254). TLC plates were visualized with UV lamps or ninhydrin 0.2% solution in ethanol (followed by heat). Flash column chromatography was carried out according to the procedure of Still *et al.*[2]using J. T. Baker (~40 μm) flash silica gel. All synthetic intermediates were stored at 4°C after purification.

Circular dichroism spectra were recorded on a Jasco J600 Circular Dichroism Spectrometer. The spectra were baseline corrected and noise reduced using the Jasco J600 System Software on a 486-AT IBM-compatible personal computer, then transferred to a MacIntosh IIci personal computer for further analysis using Kaleidagraph Version 3.0 (Synergy Software). UV/Vis spectra were recorded on a Shimadzu UV160U Spectrometer or an HP8451A Diode Array Spectrometer and subsequently processed on a MacIntosh IIci personal computer using Kaleidagraph Version 3.0.

Reversed-phase high pressure liquid chromatography was performed on a Beckman Series Gold HPLC, equipped with a dual wavelength UV detector.

(1) Perrin, D.D.; Armarego, W.L.F. *Purification of Laboratory Chemicals*; Pergamon Press: New York, 1988.

(2) Still, W.C.; Kahn, M.; Mitra, A. "Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution," *J. Org. Chem.* **1978**, *43*, 2923.

Chapter 1: Introduction

Naturally occurring proteins possess enviable properties in that their diverse functions, accompanied by well defined supramolecular structures, provide almost limitless opportunities for the construction of new devices. In the quest for designing new proteins, it has become apparent that stable, well-defined, 3-dimensional structures are crucial to the development of functional entities. To date, attempts to assemble constructs with enhanced stabilities or novel, pre-programmed functions have focused largely on one of three methods: the modification of existing, naturally occurring protein structures, the selective screening of monoclonal antibodies raised to organic templates designed to model reaction pathways (catalytic antibodies),[1-3] or the *de novo* synthesis of new proteins.[4-7] In the first approach, the existing protein structure is treated as a scaffold, upon which new functionalities can be incorporated as long as the introduced modifications do not compromise the overall protein superstructure. This approach to protein design has benefited from the wide range of techniques now available to introduce both natural and unnatural elements directly into the polypeptide backbone; these techniques include site-directed mutagenesis,[8] *in vitro* expression of chemically acylated suppressor *t*-RNA,[9] direct chemical synthesis,[10] and semi-synthesis.[11] In addition, the chemical modification of protein side-chains with functionalized reagents has provided an alternative method to generate proteins with novel functions.[11] Thus the modification of existing protein structures has provided an array of diverse, functional entities which have been excellent systems to study protein structure and function. One of the main drawbacks to this approach, however, is that the designed systems are inherently limited to those naturally occurring proteins that are amenable to the techniques listed above.

The area of *de novo* protein synthesis offers the most potential for the creation of truly novel protein structures with unprecedented functions. Simply stated, it involves the construction of proteinaceous products with rationally assembled structures and functions, all designed from first principles. The process of designing and testing new protein constructs provides the most rigorous test of the underlying principles of protein folding,

structure, and function. Many of the recent advances in the *de novo* protein design have focused on the construction of compact, stable, secondary structural motifs, which have then been assembled in a modular sense to form higher order structures that mimic native systems. This hierarchic condensation is based on the framework model of protein folding,[12] which proposes that proteins fold along a defined pathway and that the folding process is initiated and stabilized by small units of secondary structure (e.g., α-helix, β-sheet, β-turns). The design of small structural elements has relied heavily on the results of studies with native proteins and with protein fragments found to adopt stable secondary structures,[13-15] as well as synthetically prepared templates designed to induce α-helix[16] and β-sheet[17] conformations. The conclusions from these studies, in conjunction with those derived from surveys of crystal structures of natural proteins,[18-20] have provided a foundation for the selection of primary sequences designed to assume particular secondary structures (e.g., α-helix, β-sheet). Protein designs based on helical motifs have received considerable attention, primarily due to the prevalence of helical motifs in natural proteins,[21,22] and the ability to construct sequences with controlled aggregation properties.[23] Successful protein designs have been reported for the coiled-coil structural motif[24] and higher order helical bundles composed of 3 or 4 helical domains.[4,25,26]

In contrast, designs based on β-sheet structures have been hampered due to extensive solubility problems; the limited solubilities have been attributed to the use of bulky, hydrophobic amino acids which have high propensities for extended conformations[19,20] and to the potential for β-sheet structures to form infinite lattices.[27] This packing arrangement has been observed in the crystal structures of enkephalin peptide hormones[28,29] and has recently been implicated in the formation of amyloid plaque proteins associated with Alzheimer's disease.[30] However, globular β-structures are prevalent in nature (e.g., immunoglobin family of proteins), and several designs modeled after these systems have been reported. For example, an eight-stranded antiparallel β-barrel structure has been designed by Richardson[7] based on the immunoglobulin V_L domain, and the

design, synthesis, and characterization of a protein containing a novel β-sheet fold with an engineered metal binding site has recently been reported.[31] The latter protein, called the "minibody," was designed using β-sheet strands and hypervariable loop fragments derived from the immunoglobin V_H chain to form a 6-stranded β-sandwich protein. It should be noted that solubility problems hindered the characterization of these motifs; in both cases the solution structure was assessed using circular dichroism spectroscopy. The results from these models hold promise for the future design of stable, soluble protein motifs based solely on the β-sheet secondary structure.

Despite these advances, the *de novo* synthesis of functional proteins has remained largely elusive to date. While several systems have been reported to exhibit catalytic activity,[32-35] the lack of detailed structural analyses in these systems has prevented comprehensive characterization of the reaction mechanisms. As a result, the nature of the observed catalytic activities for most of these systems is not well understood.[6] Even in the most well characterized systems, which are composed of defined secondary structural elements (e.g., 3- and 4-helical bundles), it is very difficult to precisely predict the tertiary interactions between the structural motifs.[5] In fact, most of the designs reported to date which employ natural amino acid sequences mimic the so-called "molten-globule" protein folding state rather than the correctly folded proteins found in nature.[4] It is clear from these studies that more constraints on the structural framework are required for these systems to assume native-like protein structures.

One approach to this problem has employed the use of synthetic templates to induce and direct the correct folding of polypeptide chains. Mutter, using the template-assisted synthetic protein (TASP) approach, has assembled a series of 4-helix bundle proteins covalently attached to a branched cyclic decapeptide template.[36] The resulting synthetic proteins were more stable than the corresponding 4-helical bundle proteins constructed from linear peptide sequences as determined by CD spectroscopy, NMR spectroscopy, and

chemical denaturation studies. Similarly, 4-helical bundle motifs have been synthesized using porphyrin ring templates by a number of research groups.[4]

The incorporation of metal binding sites into the designed motifs has been attempted as another possible solution to this problem. Metal cations play central structural and functional roles in naturally occurring protein architectures.[37] In some cases, such as the "zinc-finger" regulatory proteins[38] and the protein hormone insulin,[39] metal cations are essential for correct folding and stabilization of the protein structure. In light of these examples, the construction of sites capable of binding metal cations selectively and in high affinity has received considerable attention in *de novo* protein design.[31,40-46] Indeed, the demonstration of selective binding of metal cations has been considered to be one of the first steps in the construction of functional protein designs.[4] However, the construction of sites capable of binding metal cations using the ribosomally encoded amino acids presents a formidable challenge, as the number of amino acids capable of ligating metal cations is fairly limited.

The approach taken by this laboratory has been to utilize unnatural amino acids in designed motifs to direct the polypeptide conformation or introduce functional groups capable of emulating the reactivities found in native enzyme systems. It is surprising that while the ability of non-encoded amino acids to constrain peptide conformation has precedent in a wide variety of naturally occurring peptide hormones and synthetically prepared peptides,[47] these moieties have not been utilized to a significant extent in *de novo* protein design.[4] With the wide variety of methods now available for the synthesis of optically pure α-amino acids containing complex functionalities,[48] it is now possible to prepare functionalized amino acids capable of directing the overall polypeptide conformation or introducing groups with enhanced reactivities. In addition, the recent advances in solid phase peptide synthesis[49] have provided methods to prepare relatively long polypeptide sequences (≈100 residues) suitable for the construction of small designed motifs or for the assembly of larger constructs using semisynthetic methodologies. As a

result, unnatural amino acids can be incorporated anywhere within a peptide sequence, and this versatility makes them ideal building blocks for the assembly of highly functionalized polypeptide motifs. The work presented in this dissertation describes the application of these concepts towards the development of two novel metal-binding protein motifs which display not only effective metal-binding properties, but also utilize metal cations to perform designated functions. The first motif concerns to the *de novo* design of proteins programmed to regulate the activity of peptide-binding enzymes while the second project involves the engineering of an existing protein structure through semisynthetic techniques to incorporate unnatural amino acids capable of binding metal cations with high affinities.

The first project involved the development of proteins which were designed to function as regulatory subunits for naturally occurring enzymes which process peptidyl substrates (e.g., kinases, proteases). The *de novo* design of these proteins was based on native enzyme systems where regulation of enzyme activity is often achieved through conformational changes in regulatory domains which are induced by chemical modulators.[50-53] For example, the *c*AMP dependent protein kinase complex contains a regulatory subunit which responds to physiological levels of *c*AMP; in the absence of *c*AMP it has a high affinity for the catalytic subunit active site and thus forms a tight complex between the catalytic and regulatory domains.[54-56] However, upon binding *c*AMP, the regulatory subunit undergoes a conformational change which results in the dissociation of the protein complex to release the catalytic subunit. This process is shown schematically in Figure 1-1.

The *de novo* designed regulatory protein is proposed to effect its action on a targeted enzyme in a similar fashion; however, the required conformational changes will be induced by metal cations. As shown in Figure 1-2, the overall structure of the protein is based on a ββ supersecondary structure (in the absence of metal cations) and it contains three distinct regions which are responsible for either metal cation binding or interacting with the catalytic pocket of the targeted enzyme. The first region of the protein, situated

Figure 1-1. Mechanism of *c*AMP dependent protein Kinase C regulation (C = catalytic subunit, R = regulatory subunit)

near the amino terminus, is designed to form a β–Type II reverse turn secondary structure with two metal binding ligands. The Type II β-turn conformation is expected to serve as a conformationally stable "platform" upon which the two ligands can form a bidentate metal binding site. The third region, situated near the carboxy terminus, is also

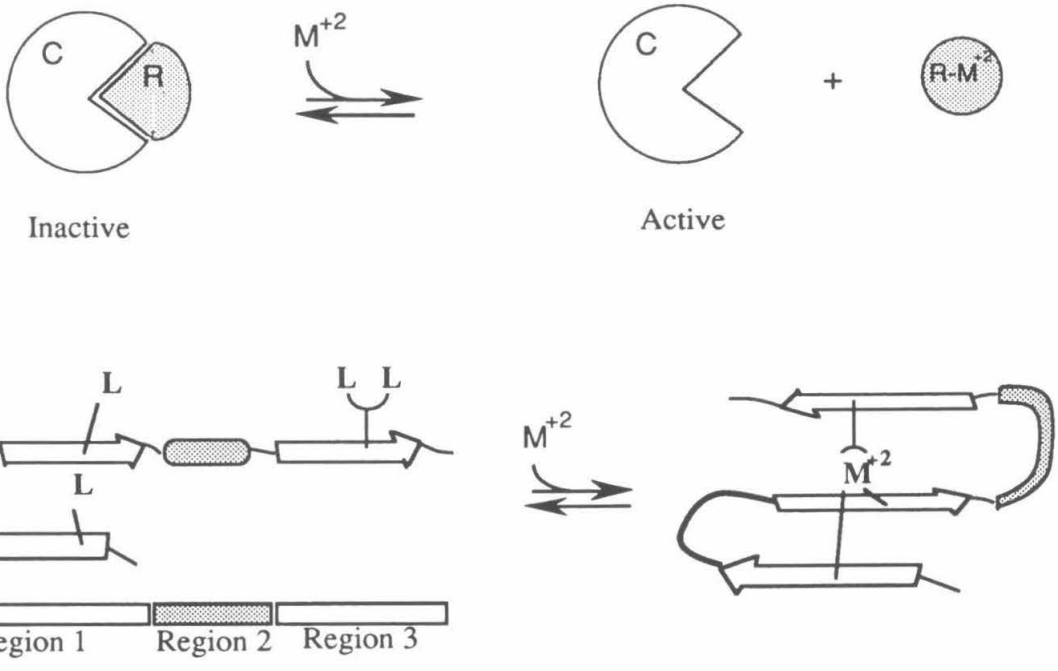


Figure 1-2. Schematic of proposed metal-dependent regulatory protein (M^{+2} = metal cations)

implicated in metal cation binding, as it is composed of an unnatural amino acid capable of bidentate metal cation coordination. Finally, the second region, located in the center of the peptide sequence, is designed to incorporate a consensus sequence capable of acting as a competitive inhibitor for the targeted enzyme (e.g., *c*AMP-dependent protein kinase C). In the absence of metal cations, this region of the peptide is expected to bind in the catalytic pocket of the targeted enzyme and effectively block the enzyme catalytic activity.

The regulatory action of the proposed motif is based on the fact that many enzymes that process peptidyl substrates (e.g., kinases, proteases) have binding domains which require the substrates to adopt an extended conformation. For example, analyses of the X-ray crystal structure of the *c*AMP dependent protein kinase containing a bound inhibitor[57,58] have shown that a five residue consensus sequence common to all known substrates and inhibitors of this enzyme adopt an extended conformation when bound in the catalytic pocket. It is believed that the region of the regulator motif containing the inhibitory consensus sequence targeted to the enzyme (e.g., *c*AMP dependent protein kinase) will have a high affinity for the catalytic pocket in the absence of divalent metal cations, since the peptide will be expected to be relatively conformationally free under these conditions. However, upon the addition of a divalent metal cation, the designed protein is proposed to form a tetradentate metal complex utilizing the metal binding ligands located in the first and third regions of the protein sequence. In this case the metal cation serves as a template, and monomeric complex formation involving these ligands would induce a conformational change in the central region of the protein motif causing it to change from an extended conformation to one which contains a constrained bulge or loop (See Figure 1-2). This restriction on the conformational flexibility of the central region is proposed to reduce the binding affinity of the regulatory motif for the catalytic pocket of the enzyme and therefore promote the dissociation of the regulatory motif from the enzyme active site.

Since the regulatory protein motif design contains three distinct regions, it was envisioned that overall design could be assembled *via* a modular approach utilizing

optimized peptide sequences and amino acids as determined by model studies performed for each region of the design. As a result, initial studies focused on the design of the first metal binding region which was proposed to contain a reverse turn motif capable of binding metal cations in aqueous solution. In order to determine the sequence requirements for stable reverse turn motifs in solution, a systematic study was carried out on a series of tetrapeptides as described in Chapter 2. These studies provided a series of guidelines for the selection of primary sequences which contain significant amounts of turn character in solution. The guidelines provided a basis for the work described in Chapter 3, which investigated the effects of incorporating functional groups at the first and fourth positions (positions *i* and *i+3* respectively) of the reverse turn motif. These studies indicated that a tetrapeptide incorporating the unnatural amino acid, (2*R*, 3*R*)-β-methyl cysteine, at the first position with a histidine residue at the fourth position adopted a significant amount of reverse turn character and was capable of binding metal cations in aqueous solution.

The third region of the designed motif, proposed to contain an unnatural amino acid capable of bidentate metal cation coordination, was investigated in Chapter 4 through the stereospecific syntheses of a series of novel bipyridyl- and phenanthrolyl-substituted amino acids. These amino acids were synthesized using a variety of routes to afford the optically pure enantiomers with N^{α}-FMOC protection to allow peptide incorporation using solid phase peptide synthesis protocols. The metal binding properties of peptides incorporating these amino acids was studied in Chapter 5, and the results from these investigations indicated that these amino acids were capable of selectively binding a variety of metal cations with affinities ranging over five orders of magnitude in some cases. The metal binding affinities and the stoichiometries of the metal complexes was found to be dependent both on the specific regioisomer of the bipyridyl/phenanthrolyl-amino acid employed as well as on the particular primary sequence of the peptide in which the amino acid was incorporated.

The studies described in Chapter 6 highlight the potential applications of the bipyridyl-substituted amino acids in larger systems through the semisynthesis of two cytochrome *c* mutants incorporating the residues at position 72 of the protein sequence. This project was undertaken as a collaborative effort with Professor Harry B. Gray and Deborah Wuttke, and it was proposed that the bipyridyl amino acids, when incorporated into the polypeptide backbone of the protein structure, could serve as effective anchors for redox-active metal centers. These redox-active centers could then be used in the study of intraprotein electron transfer processes between the introduced metal centers and the covalently bound heme group. The construction and structural characterization of these novel proteins was carried out and the results from these studies indicated that incorporation of the bipyridyl amino acids at position 72 of the protein sequence using semi-synthetic methodologies had a negligible effect on the overall protein structure. Further, one of the proteins was found to undergo modification of with $Ru(bpy)_2^{+2}$ to form the desired redox-active $Ru(bpy)_3^{+2}$ center on the protein surface. Enhanced electron transfer properties were observed with the modified protein, and these properties were applied towards the study of protein folding processes.

Finally, the work described in Chapter 7 represents a continuation of the regulatory protein design as it involves the construction of a prototype of the regulatory motif. The prototype was based on optimized components determined from the studies in Chapters 2-5, and it was prepared in an effort to test the ability to form intramolecular monomeric complexes with the proposed metal binding regions. The results from metal binding studies demonstrated that the peptide forms monomeric complexes with very high metal cation binding affinities. These results show the potential for the future design of functional, regulatory motifs targeted to naturally occurring enzymes.

The work described in this dissertation demonstrates the utility of unnatural amino acids in the development and construction of novel metal-binding protein motifs. In general, the results from these studies indicate that the metal complex coordination

environment can be influenced both by the choice of the metal binding ligands and the three-dimensional structure of the polypeptide framework. These attributes will prove useful in the future design of metalloprotein motifs containing metal coordination sites capable of complex reactivities.

References

(1) Stewart, J.D.; Benkovic, S.J. "Catalytic Antibodies - Mechanistic and Practical Considerations," *Chem. Soc. Rev.* **1993**, *22*, 213-219.

(2) Stewart, L.D.; Liotta, L.J.; Benkovic, S.J. "Reaction Mechanisms Displayed by Catalytic Antibodies," *Acc. Chem. Res.* **1993**, *26*, 396-404.

(3) Benkovic, S.J. "Catalytic Antibodies," *Ann. Rev. Biochem.* **1992**, *61*, 29-54.

(4) 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.

(5) DeGrado, W.F.; Raleigh, D.P.; Handel, T. "*De novo* protein design: what are we learning?" *Curr. Opin. Struct. Biol.* **1991**, *1*, 984-993.

(6) Regan, L. "Protein Design," *Curr. Opin. Biotechnology* **1991**, *2*, 544-550.

(7) Richardson, J.S.; Richardson, D.C. "The *de novo* design of protein structures," *Trends in Biol. Sci.* **1989**, *14*, 304-309.

(8) *Protein Engineering*; Oxender, D.L.; Fox, C.F. Eds.; A.R. Liss Inc.: New York, 1987; pp 47-109.

(9) Anthony-Cahill, S.J.; Griffith, M.C.; Noren, C.J.; Suich, D.J.; Schultz, P.G. "Site-specific mutagenesis with unnatural amino acids," *Trends in Biol. Sci.* **1989**, *14*, 400-403.

(10) Muir, T.W.; Kent, S.B.H. "The chemical synthesis of proteins," *Curr. Opin. Struct. Biol.* **1993**, *4*, 420-427.

(11) Humphries, J.; Offord, R.E.; Smith, R.A.G. "Chemical methods of protein synthesis and modification," *Curr. Opin. Biotechnology* **1991**, *2*, 539-543.

(12) Baldwin, R.L. "How does protein folding get started?" *Trends in Biol. Sci.* **1989**, *14*, 291-294.

(13) Kim, C.A.; Berg, J.M. "Thermodynamic β-sheet propensities measured using a zinc-finger host peptide," *Nature* **1993**, *362*, 267-270.

(14) 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.

(15) Marqsee, S.; Baldwin, R.L. "Helix stabilization by $Glu^{-}\cdots Lys^{+}$ salt bridges in short peptides of *de novo* design," *Proc. Natl. Acad. Sci., USA* **1987**, *84*, 8898.

(16) Kemp, D.S.; Boyd, J.G.; Muendel, C.C. "The Helical *s* Constant for Alanine in Water Derived From Template-Nucleated Helices," *Nature* **1991**, *352*, 451-454.

(17) Kemp, D.S.; Bowen, B.R.; Muendel, C.C. "Synthesis and Conformational Analysis of Epindolidione-Derived Peptide Models for β-Sheet Formation," *J. Org. Chem.* **1990**, *55*, 4650-4657.

(18) 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*, 211-245.

(19) Levitt, M. "Conformational Preferences of Amino Acids in Globular Proteins," *Biochemistry* **1978**, *17*, 4277-4285.

(20) Lifson, S.; Sander, C. "Antiparallel and parallel β-strands differ in amino acid residue preferences," *Nature* **1979**, *282*, 109-111.

(21) Presnell, S.R.; Cohen, F.E. "Topological Distribution of four-α-helix bundles," *Proc. Natl. Acad. Sci., USA* **1989**, *86*, 6592-6596.

(22) Weber, P.C.; Salemme, F.R. "Structural and Functional Diversity in 4-α-Helical Bundles," *Nature* **1980**, *287*, 82-84.

(23) Ho, S.P.; DeGrado, W.F. "Design of a 4-Helix Bundle Protein: Synthesis of Peptides Which Self-Associate into a Helical Protein," *J. Am. Chem. Soc.* **1987**, *109*, 6751-6758.

(24) Cohen, C.; Parry, D.A.D. "α-Helical Coiled Coils and Bundles: How to Design an α-Helical Protein," *Proteins: Struct. Func. Gen.* **1990**, *7*, 1-15.

(25) 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.

(26) Lovejoy, B.; Choe, S.; Cascio, D.; McCrorie, D.K.; DeGrado, W.F.; Eisenberg, D. "Crystal Structure of a Synthetical Triple-Stranded α-Helical Bundle," *Science* **1993**, *259*, 1288-1293.

(27) Chou, K-C.; Nemethy, G.; Rumsey, S.; Tuttle, R.W.; Scheraga, H.A. "Interactions Between Two β-Sheets: Energetics of β/β Packing in Proteins," *J. Mol. Biol.* **1986**, *188*, 641-649.

(28) Griffin, J.F.; Langs, D.A.; Smith, G.D.; Blundell, T.L.; Tickle, I.J.; Bedarkar, S. "The crystal structures of [Met^5]enkephalin and a third form of [Leu^5]enkephalin: Observations of a novel pleated β-sheet," *Proc. Natl. Acad. Sci., USA* **1986**, *83*, 3272-3276.

(29) Camerman, A.; Mastropaolo, D.; Karle, I.; Karle, J.; Camerman, N. "Crystal structure of leucine-Enkephalin," *Nature* **1983**, *306*, 447-450.

(30) Jarrett, J.T.; Lansbury, P.T. "Seeding One-Dimensional Crystallization of Amyloids: A Pathogenic Mechanism in Alzheimers-Disease and Scrapie," *Cell* **1993**, *73*, 1055-1058.

(31) 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.

(32) Atassi, M.Z.; Manshouri, T. "Design of peptide enzymes (pepzymes): Surface-simulation synthetic peptides that mimic the chymotrypsin and trypsin active sites

exhibit the activity and specificity of the respective enzyme," *Proc. Natl. Acad. Sci., USA* **1993**, *90*, 8282-8286.

(33) Sasaki, T.; Kaiser, E.T. "Helichrome: Synthesis and Enzymatic Activity of a Designed Hemeprotein," *J. Am. Chem. Soc.* **1989**, *111*, 380-381.

(34) Hahn, K.W.; Klis, W.A.; Stewart, J.M. "Design and Synthesis of a Peptide Having Chymotrypsin-Like Esterase Activity," *Science* **1990**, *248*, 1544-1547.

(35) Goraj, K.; Renard, A.; Martial, J.A. "Synthesis, purification and initial structural characterization of octarellin, a *de novo* polypeptide modelled on the α/β-barrel proteins," *Protein Eng.* **1990**, *3*, 259-266.

(36) Mutter, M.; Tuscherer, G.G.; Miller, C.; Altmann, K-H.; Carey, R.I.; Wyss, D.F.; Labhardt, A.M.; Rivier, J.E. "Template-Assisted Synthetic Proteins with Four-Helix-Bundle Topology. Total Chemical Synthesis and Conformational Studies," *J. Am. Chem. Soc.* **1992**, *114*, 1463-1470.

(37) Glusker, J.P. "Structural Aspects of Metal Liganding to Functional Groups in Proteins," *Adv. Prot. Chem.* **1991**, *42*, 1-76.

(38) Shi, Y.; Beger, R.D.; Berg, J.M. "Metal Binding Properties of Single Amino Acid Deletion Mutants of Zinc Finger Peptides: Studies Using Cobalt(II) as a Spectroscopic Probe," *Biophys. J.* **1993**, *64*, 749-753.

(39) Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D. "Insulin: the Structure in the Crystal and its Reflection in Chemistry and Biology," *Adv. Prot. Chem.* **1972**, *26*, 279-402.

(40) 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.

(41) Ghadiri, M.R.; Soares, C.; Choi, C. "A Convergent Approach to Protein Design: Metal Ion Assisted Spontaneous Self-Assembly of a Polypeptide into a Triple Helix Bundle Protein," *J. Am. Chem. Soc.* **1992**, *114*, 825-831.

(42) Ghadiri, R.M.; Soares, C.; Choi, C. "Design of an Artificial Four-Helix-Bundle Metalloprotein via a Novel Ruthenium(II)-Assisted Self-Assembly Process," *J. Am. Chem. Soc.* **1992**, *114*, 4000-4002.

(43) Lieberman, M.; Sasaki, T. "Iron(II) Organizes a Synthetic Peptide into Three-Helix Bundles," *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.

(44) Handel, T.; DeGrado, W.F. "De Novo Design of a Zn^{+2}-Binding Protein," *J. Am. Chem. Soc.* **1990**, *112*, 6710-6711.

(45) Regan, L.; Clarke, N.D. "A Tetrahedral Zinc(II)-Binding Site Introduced into a Designed Protein," *Biochemistry* **1990**, *29*, 10878-10883.

(46) Ruan, F.; Chen, Y.; Hopkins, P.B. "Metal Ion Enhanced Helicity in Synthetic Peptides Containing Unnatural, Metal-Ligating Residues," *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404.

(47) Hirschmann, R. "Medicinal Chemistry in the Golden Age of Biology: Lessons from Steroid and Peptide Research," *Angew. Chem. Int. Ed., Eng.* **1991**, *30*, 1278-1301.

(48) Williams, R.M. *Synthesis of Optically Active α-Amino Acids*; Pergamon Press: New York, 1989.

(49) Stewart, J.M.; Young, J.D. *Solid Phase Peptide Synthesis*; Pierce Chemical Co.: Rockford, IL, 1984.

(50) Banzon, J.A.; Kelly, J.W. "β-Sheet Rearrangements: serpins and beyond," *Protein Eng.* **1992**, *5*, 113-115.

(51) Crowther, D.C.; Evans, D.L.I.; Carrell, R.W. "Serpins: implications of a mobile reactive center," *Curr. Opin. Biotechnology* **1992**, *3*, 399-407.

(52) Phillips, M.A.; Fletterick, R.J. "Proteases," *Curr. Opin. Struct. Biol.* **1992**, *2*, 713-720.

(53) Storer, A.C. "Engineering of Proteases and Protease Inhibition," *Curr. Opin. Biotechnology* **1991**, *2*, 606-613.

(54) Beebe, S.J.; Corbin, J.D. "Cyclic Nucleotide-Dependent Protein Kinases," in *The Enzyme: Control by Phosphorylation, Part A*; Krebs, E.G.; Boyer, P.D., Eds.; Academic Press: New York, 1986.

(55) Bramson, H.N.; Kaiser, E.T.; Mildvan, A.S. "Mechanistic Studies of *c*AMP-Dependent Protein Kinase Action," *CRC Crit. Rev. Biochem.* **1984**, *15*, 93-124.

(56) Flockhart, D.A.; Corbin, J.D. "Regulatory Mechanisms in the Control of Protein Kinases," *CRC Crit. Rev. Biochem.* **1982**, *13*, 133-186.

(57) Knighton, D.R.; Zheng, J.; Ten Eyck, L.F.; Ashford, V.A.; Xuong, N-H.; 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.

(58) Knighton, D.R.; Zheng, J.; Ten Eyck, L.F.; Ashford, V.A.; Xuong, N-H.; Taylor, S.S.; Sowadski, J.M. "Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase," *Science* **1991**, *253*, 407-414.

Chapter 2. Investigation of the Sequence Requirements for β-Type II Reverse Turns

Introduction

Reverse turns play an important structural role in the compact globular architecture of native folded proteins.[1-3] However, although the reverse turn has been the subject of numerous experimental[4-9] and theoretical[10-12] investigations, the explicit role of the turn motif in the protein folding process in nature remains an open issue.[2,13-15] In fact it is probable that this small motif is simply a versatile element in the hierarchy of protein structure, and the specific responsibility in folding, be it active or passive, is dependent upon the local amino acid sequence.[16] A current predicament in the *de novo* design and synthesis of proteins concerns the ability to precisely predict and control the three-dimensional architecture of a protein-like biopolymer.[17] In this context, we reasoned that it would be advantageous to delegate a specific function to the reverse turn in the protein infrastructure. Therefore, our objective in this work has been to identify and construct a conformationally well-characterized turn motif, which could be used to actively nucleate[18] and stabilize folded protein structures. Progress in this area would complement the recent developments in α-helix and β-sheet stabilization and the utilization of those structural motifs in *de novo* protein design.[17]

At the present time there are 11 classes of reverse turns known;[19] however, only 3 of these classes are commonly found in protein structures.[1,19] We have chosen specifically to focus on the type II rather than type I turn. While the latter type of β-turn is most common in naturally occurring proteins which are composed of L-amino acids,[1] the former appears to represent a more useful target for optimization for use as a structural nucleation element for a number of reasons. First, the type II turn motif is compatible with a heterochiral sequence, which would render it "position specific" within a longer polypeptide chain since such a sequence cannot be integrated as part of either a regular α-helix or β-sheet structure.[16] This preference, first realized by Venkatachalam,[20] has been attributed to the elimination of steric interactions between the amide carbonyl oxygen and the side-chain of the D residue. The type I turn, however, is more compatible with an L-L

sequence based on similar arguments and the fact that the central amide bond adopts the opposite orientation. Second, the peptide backbone of the type II turn has been found to form a more planar conformation than a type I.[19] Such a conformation maximizes the interactions within the peptide backbone and minimizes those between the side-chain and backbone. This planar arrangement fixes the side-chains of the residues at positions *i* and *i+3* to lie in the same plane with the distance between the two C^{α} carbons of these residues approximately 7Å apart. These attributes could be exploited in *de novo* design motifs, as the type II reverse turn can be viewed as a structural platform with the potential for programming functionality at the *i* and *i+3* positions. In contrast, the type I turn has been found to adopt a more twisted conformation in protein structures and interactions between the side-chain of the residue at position *i* and the main chain atoms of position *i+3* are common.[19] Third, the dihedral angles for the residues involved in a type II β-turn allow for a very favorable hydrogen-bonding alignment between the carbonyl oxygen and the amide NH of the *i* and *i+3* residues which may be more inherently stable.[4]

Finally, the wide range of studies using cyclic peptides[4,7] and immunogenic peptide fragments[6] provide a basis for the investigation of reverse turn formation in aqueous solution. Despite this wealth of information, however, the linear peptide sequence requirements for the type II turn conformation have not been well defined. Many of the experimental studies reported to date have been performed under widely different conditions and characterized by a variety of techniques[3] thereby making direct comparisons of the results difficult to interpret. In addition, since the type II turn prefers a heterochiral sequence, probabilistic studies[1,19,21] based on amino acid occurrences in protein structures are of limited value for the determination of the primary sequence requirements for this motif. Indeed, the results of these studies have shown that there are very few statistically significant positional preferences in the type II turn beyond proline at the *i+1* position and glycine at the *i+2* position.[19] In order to determine the specific sequence requirements for the type II turn motif, a systematic study of a series of tetrapeptides was undertaken in both

dimethyl sulfoxide (DMSO) and aqueous environments. The goal of this work was to formulate general guidelines for the construction of reverse turn motifs under these conditions, similar to those published for α-helices[22] and β-sheets.[23] The systematic study was based on the general sequence Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH_2, with the following constraints: a proline or proline derivative was invariant at the *i+1* position, a D-amino acid residue was always incorporated at the *i+2* position, and blocking groups on the amino and carboxyl termini were employed to better mimic a long peptide chain (see Figure 2-1).

Conformational analysis of the tetrapeptides in solution in dimethyl sulfoxide (DMSO) and water was based on several experimental criteria. These include an

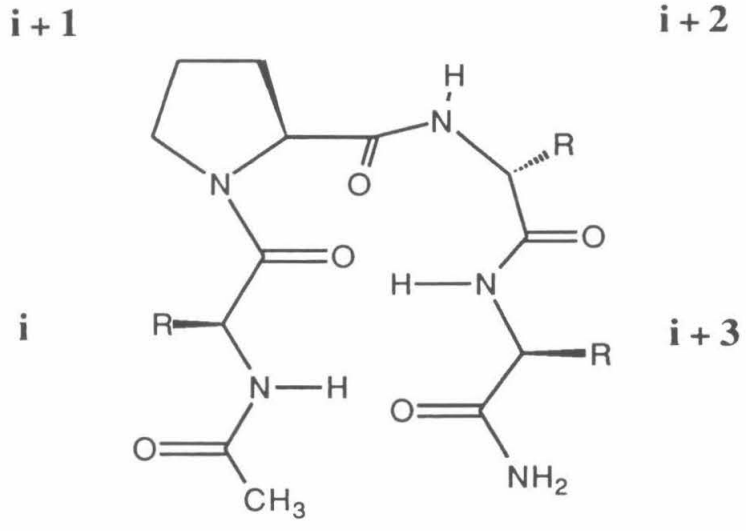


Figure 2-1. Diagram of the tetrapeptide sequence used for systematic studies

evaluation of nuclear Overhauser effects (NOEs), amide temperature coefficients and qualitative interpretation of circular dichroism (CD) spectra. NOE effects were observed using two-dimensional rotating-frame nuclear Overhauser effect spectroscopy[9,24-26] (ROESY) and results were compared with those predicted for protein secondary structure by Wuthrich.[27] NOE studies provide unequivocal evidence for through-space interactions between protons, which would be unique to a folded structure. Variable temperature (VT)

studies allow for an assessment of whether or not exchangeable protons are solvent shielded and hence potentially involved in hydrogen-bonding interactions.[3,6] Temperature coefficients can be influenced by many factors beyond backbone hydrogen-bonding; however, comparisons made between closely related peptides, such as those in the studies discussed herein, are very useful for providing auxiliary information pertaining to substitution trends in solution conformations. Finally, the CD spectra of certain reverse turn structures have been predicted[28,29] and observed[30,31] to be distinct from those of other secondary structural motifs and random chain conformers; therefore, the CD studies in water provide a further, *independent*, indicator of solution state conformations.

In general, the main limitation of these data is that they cannot provide a quantitative measure of the absolute population of a specific conformation in linear peptide models; however, the evidence under evaluation can afford an indication of the *existence* of a population of structures with a defined conformation. Furthermore, a comparison of relative populations within the series of peptides can be used to indicate which features of the sequence are important for the structural integrity of the motif.

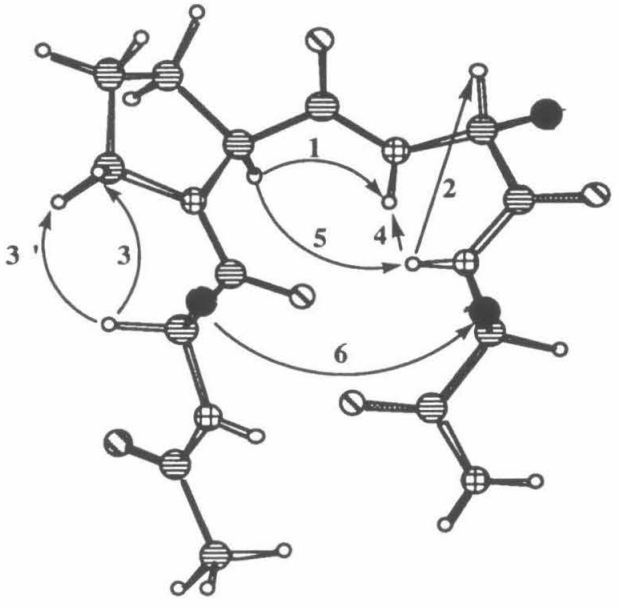
Through the course of these studies, several tetrapeptides, Ac-Val-Pro-D-Ser-His-NH_2 in particular, were found to adopt a significant amount of reverse turn (β-type II) character in both water and dimethyl sulfoxide (DMSO). The positive results from these peptides, in conjunction with the results from the other peptides of the systematic study, provide a basis for understanding the interactions that govern the reverse turn formation in linear peptide sequences. The conclusions from these studies can now be applied to design of viable, pre-programmed structural nucleation elements for incorporation into larger polypeptides. The key attributes of the motif are two-fold: firstly, the capped tetrapeptides emulate short, internal sequences of long, linear polypeptides. Secondly, the motif appears to owe its stability only to a complementary combination of local steric and hydrogen bonding effects and not to dominant interactions between oppositely charged groups or specific effects due to strong side chain-main chain interactions. These effects have been

shown to "lock" small peptides into local secondary structures and β-turn conformations.[32,33]

Results

Conformational Study of Ac-Val-Pro-D-Ser-His-NH_2 and Ac-Val-Pro-D-Ala-His-NH_2.

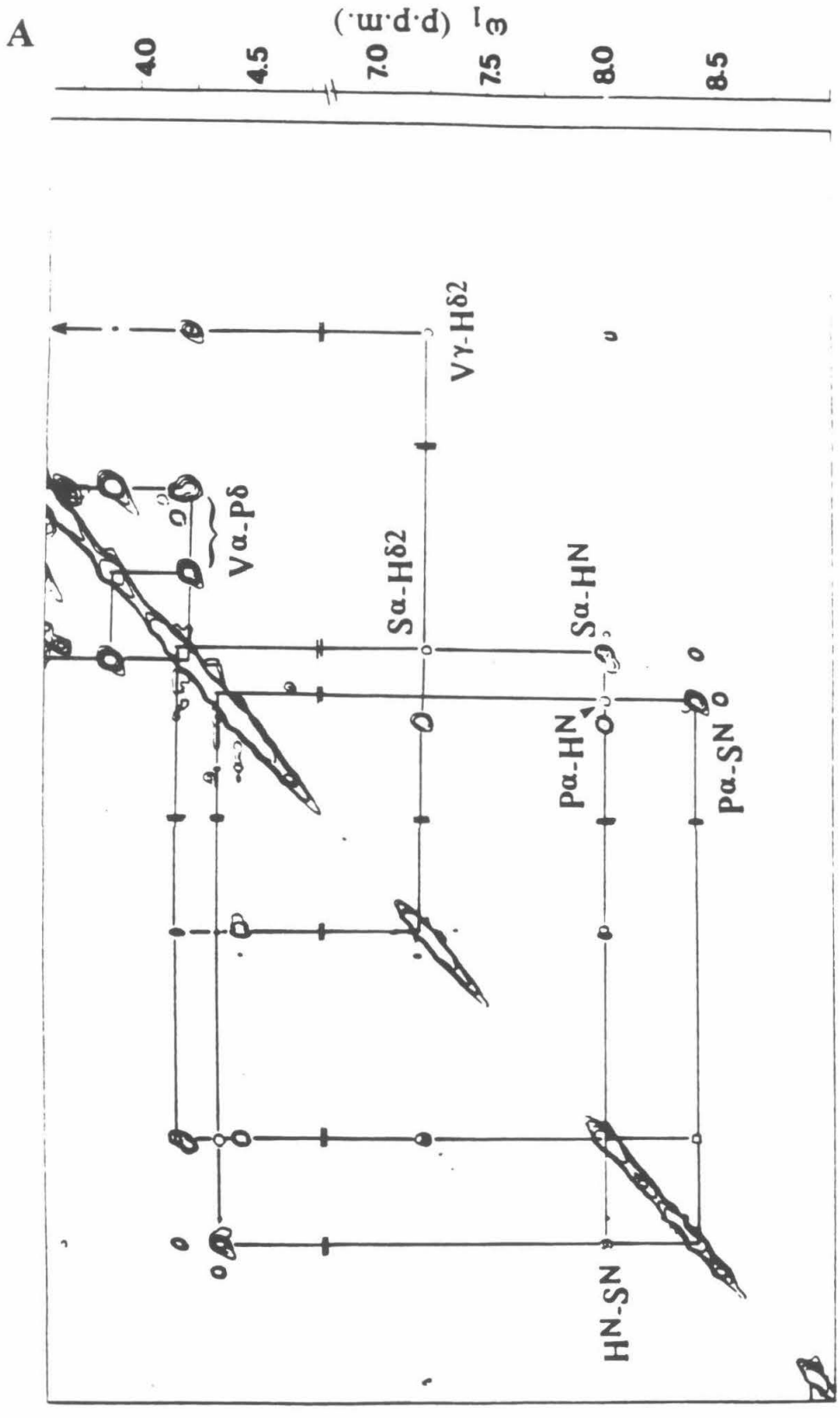
NOE studies on Ac-Val-Pro-D-Ser-His-NH_2 and Ac-Val-Pro-D-Ala-His-NH_2 in water and DMSO provide the strongest evidence to indicate the presence of a turn population. Figure 2-2 illustrates the through space connectivities which might be observed in a typical type II β-turn structure. Interactions **1**, **2**, and **3** are classified as "sequential" connectivities.[6]



(◦ , H; ⊖ , C; ⊘, O; ⊕, N; ●, R)

Figure 2-2. NOE interactions expected for a β-type II reverse turn motif: **1**, ProCα–(*i*+*2*)NH; **2**, (*i*+*2*)Cα–(*i*+*3*)NH; **3** and **3'**, (*i*)Cα–ProCδ, **4**, (*i*+*2*)NH-(*i*+*3*)NH; **5**, ProCα-(*i*+*3*)NH, **6**, (*i*) sidechain-(*i*+*3*) sidechain.

Figure 2-3. Sections of the 500-MHz ROESY spectra of Ac-Val-Pro-D-Ser-His-NH_2 in (A) DMSO (298 K) and (B) 90/10 H_2O/D_2O (278 K). The regions encompassing crosspeaks from the amide and alpha carbon protons are shown. Labeled peaks represent those important for conformational analyses and can be assigned to the following interactions: d_{NN}[His, Ser], $d_{\alpha N}$[Pro, Ser], $d_{\alpha N}$ [Pro, His], $d_{\alpha N}$ [Ser, His], $d_{\alpha \delta}$ [Val, Pro], $d_{\gamma \delta 2}$[Val,His] and $d_{\alpha \delta 2}$[Ser,His].


A
ω1 (p.p.m.)
4.0
4.5
7.0
7.5
8.0
8.5
Vγ-Hδ2
Vα-Pδ
Sα-Hδ2
Sα-HN
Pα-HN
Pα-SN
HN-SN
0.0
0.5
1.0
3.5
4.0
4.5
7.0
7.5
8.0
ω2 (p.p.m.)

B

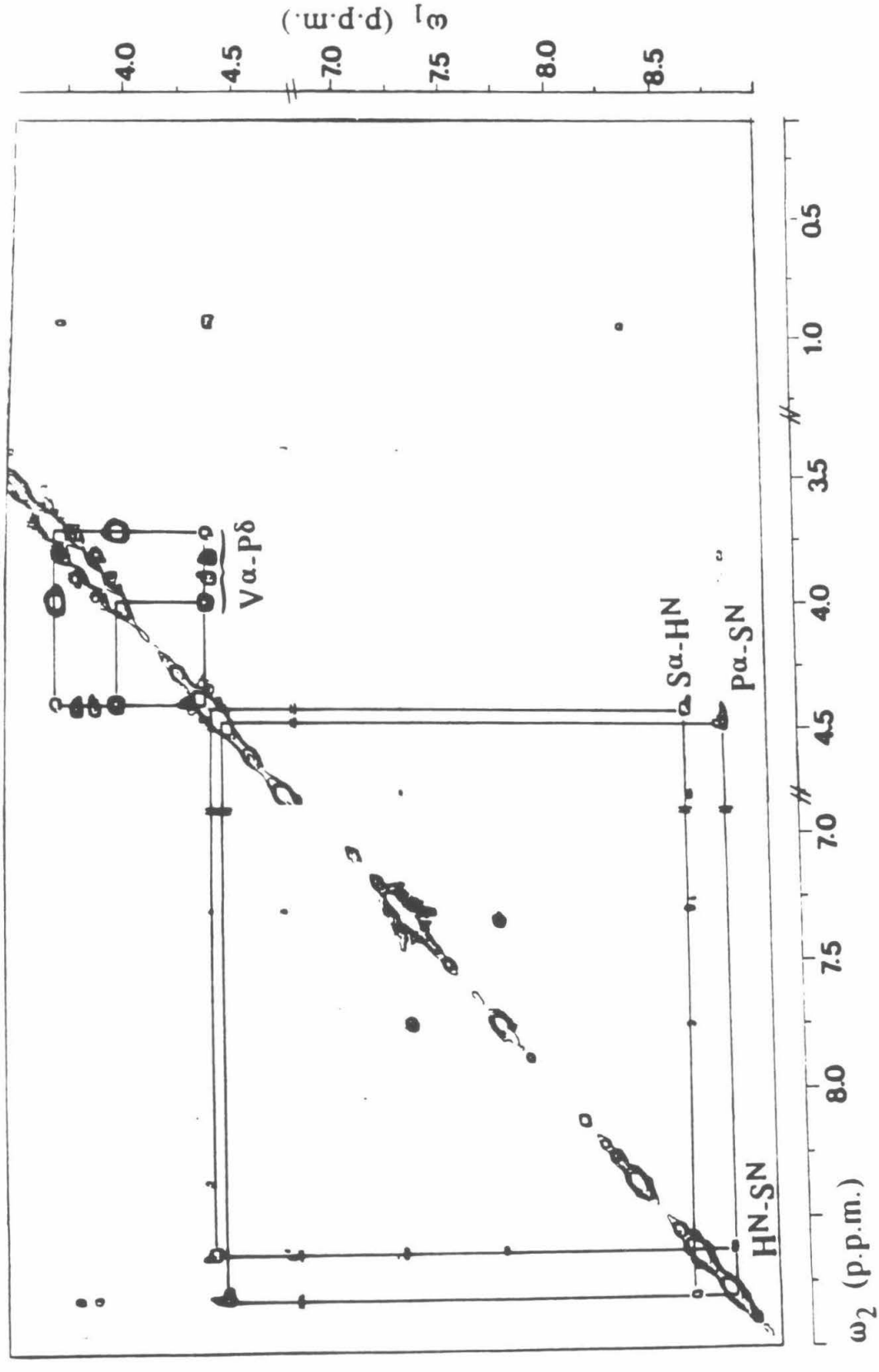
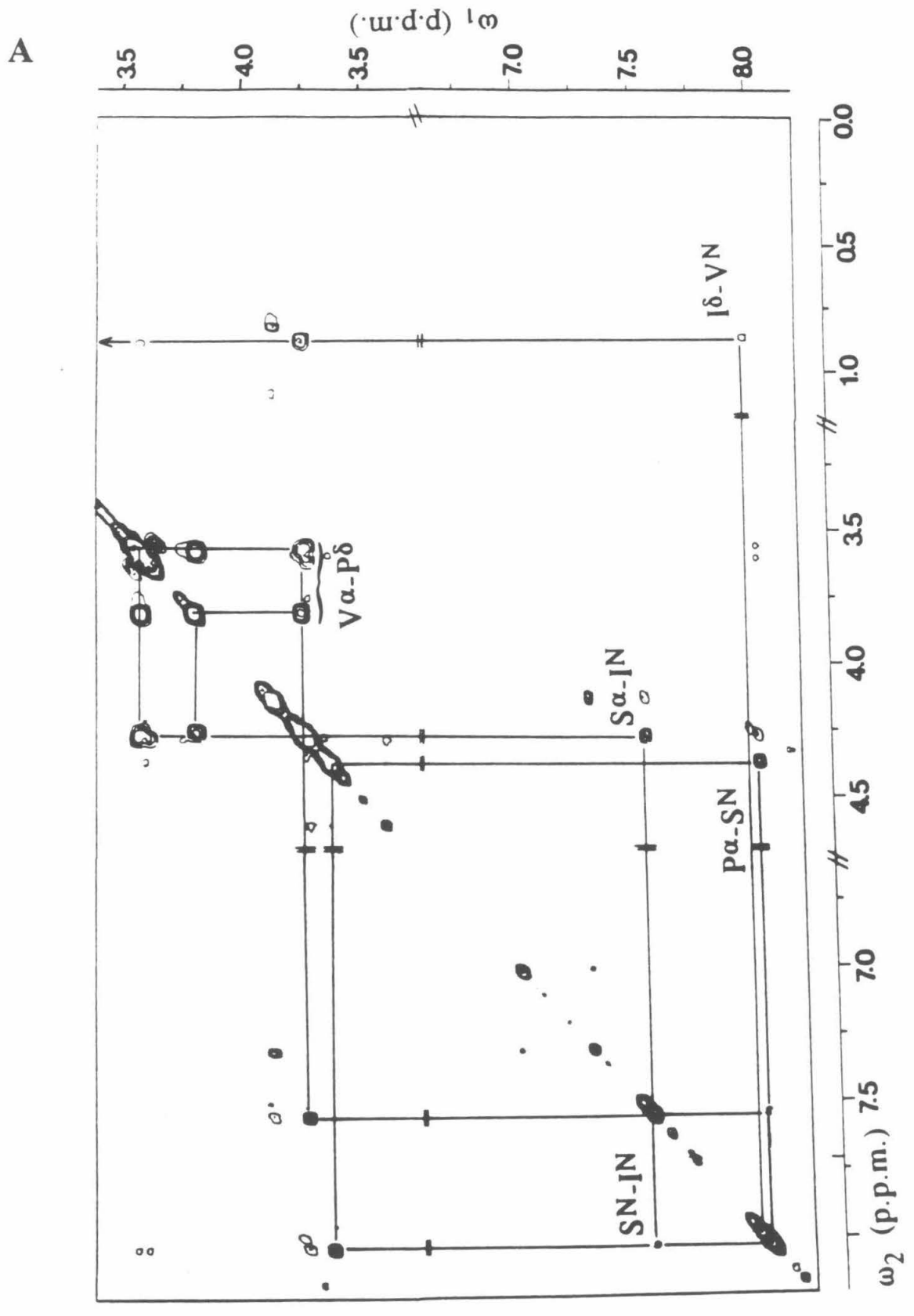
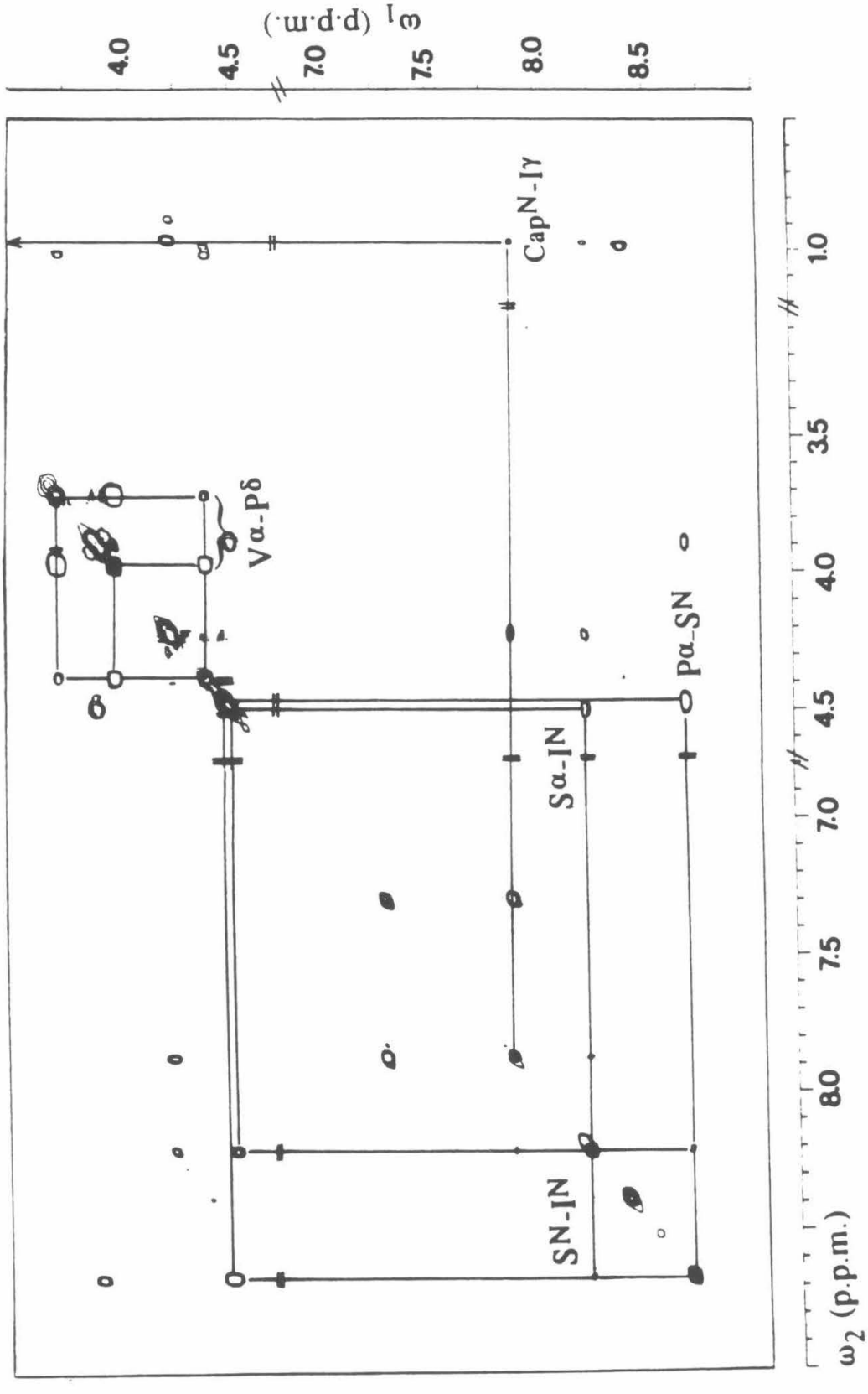

ω1 (p.p.m.)
4.0
4.5
7.0
7.5
8.0
8.5
Vα-Pδ
Sα-HN
Pα-SN
HN-SN
ω2 (p.p.m.)
0.5
1.0
3.5
4.0
4.5
7.0
7.5
8.0

Figure 2-4. Sections of the 500-MHz ROESY spectra of Ac-Val-Pro-D-Ser-Ile-NH_2 in (A) DMSO (298 K) and (B) 90% H_2O /10% D_2O (278 K). The regions encompassing crosspeaks from the amide and alpha carbon protons are shown. Labeled peaks represent those important for conformational analyses and can be assigned to the following interactions: d_{NN} [Ser, Ile], $d_{\alpha N}$ [Pro, Ser], $d_{\alpha N}$ [Ser, Ile], $d_{\alpha\delta}$ [Ile, Pro], $d_{\delta N}$ [Ile,Val] (DMSO only), and $d_{N\gamma}$ [Amide Cap, Ile].


A
ω1 (p.p.m.)
3.5
4.0
3.5
7.0
7.5
8.0
Iδ-VN
Vα-Pδ
Sα-IN
Pα-SN
SN-IN
ω2 (p.p.m.)
0.0
0.5
1.0
3.5
4.0
4.5
7.0
7.5

B


ω1 (p.p.m.)
4.0
4.5
7.0
7.5
8.0
8.5
CapN-Iγ
Vα-Pδ
Pα-SN
Sα-IN
SN-IN
ω2 (p.p.m.)
8.0
7.5
7.0
4.5
4.0
3.5
1.0

Although interactions **1**, **2** and **3** would be present, in varying intensity, in the spectra of many peptides, **4**, **5** and **6**, would be observed in a type II reverse turn folded structure.[27] The diagnostic regions of the ROESY spectra of Ac-Val-Pro-D-Ser-His-NH_2 in DMSO and water are shown in Figure 2-3. In DMSO, the sequential NOEs (**1**, **2** and **3**, Figure 2-2) are strong. In addition, the long range connectivities d_{NN} (His, Ser), $d_{\alpha N}$ (Pro, His) and $d_{\gamma\delta 2}$ (Val, His) interactions (**4**, **5** and **6**, Figure 2-2) are observed. In water the sequential NOEs and the d_{NN}(His,Ser) are present; however, the remaining long range interactions are not seen. The ROESY spectra of Ac-Val-Pro-D-Ala-His-NH_2 in both DMSO and water show the interactions **1**, **2**, **3** and **4**, with the long range interaction (**4**) being weak in water. The ROESY data for these and related peptides are summarized in Table 2-1. A small amount of a second, ordered conformation (<10% in all peptides) is evident in the 1D NMR spectra. This minor component could be due to the *cis* amide conformer of the proline residue; however, this conformation was not analyzed in detail.

Table 2-1. Observed ROESY Connectivities[a] in DMSO and Water

Peptide	DMSO	Water
Ac-Val-Pro-D-Ser-His-NH_2	**1**,**2**,**3**,**4**,**5** and **6**	**1**,**2**,**3** and **4**
Ac-Val-Pro-D-Ser-Phe-NH_2	**1**,**2**,**3**,**4** and **6**	**1**,**2**,**3** and **4**[b]
Ac-Val-Pro-D-Ser-Ile-NH_2	**1**,**2**,**3**,**4** and **6**	**1**,**2**,**3** and **4**
Ac-Val-NapPro-D-Ser-His-NH_2[c]	**ND**	**1**,**2**,**3** and **4**
Ac-Val-Pro-D-Ala-His-NH_2	**1**,**2**,**3** and **4**	**1**,**2**,**3** and **4**[b]
Ac-Val-Pro-D-Ala-Phe-NH_2	**1**,**2** and **3**	**1**,**2** and **3**

[a] See Figure 2-2 for definition of interactions; [b] Crosspeak relatively weak; [c] NapPro = *cis* -(2*S*,4*S*)-4-(naphth-2-yl)-amido-proline

VT studies on these two peptides provide data which is consistent with the NOE observations. In an ideal β-turn [*i+3* to *i* hydrogen bond] in the absence of interactions between the side chain functionality and the peptide backbone, the temperature coefficients would be predicted to show maximum solvent shielding of the *i+3* residue NH and minimum shielding of the *i+2* residue NH.[3] In a structure containing a γ-turn, [*i+2* to *i* hydrogen bond] involving the proline and D-amino acid residue, the converse would apply.[3] It should be noted that the absolute magnitudes of the temperature coefficients are influenced strongly by solvent effects; in DMSO values range from 0 to -8 ppb/K, and in water the values can range from -2 to -11 ppb/K. It is generally accepted that temperature coefficients > -4 ppb/K in H_2O (> -3 ppb/K in DMSO) is indicative of strong solvent shielding and coefficients from -4 to -6 ppb/K in H_2O (-3 to -4.5 ppb/K in DMSO) indicate moderate solvent shielding of the exchangeable proton. Protons with temperature coefficients < -6 ppb/K in H_2O (< -4.5 ppb/K in DMSO) are considered to be exposed to the solvent milieu. The amide proton temperature coefficients of the tetrapeptides examined are summarized in Tables 2-2 and 2-3. In both DMSO and water, the VT data for Ac-Val-Pro-D-Ser-His-NH_2 and the D-alanine-containing analog indicate some β-turn population: temperature coefficients for the *i+3* NH show of strong solvent shielding in DMSO and moderate shielding in water.

The CD spectra of Ac-Val-Pro-D-Ser-His-NH_2 and Ac-Val-Pro-D-Ala-His-NH_2 in water are shown in Figures 2-5a and 2-5b. CD spectra of reverse turns are distinct from those of other secondary structural motifs.[28-31] The spectral type associated with a type II β-turn, classified as a class B spectrum, is characterized by negative ellipticities >220 nm and <190 nm and a positive ellipticity in the 200-210 nm range. All spectra in Figures 2-5a and 2-5b show maxima and minima in the appropriate wavelength range for a type II β-turn; however, the absolute value for θ is weaker for Ac-Val-Pro-D-Ala-His-NH_2 than

Table 2-2. *Temperature* Coefficients for Amide Proton Chemical Shifts in DMSO ($\Delta\delta/\Delta T$)[a]

Peptide	i	i + 2	i + 3
Ac-Val-Pro-D-Ser-His-NH_2	-7.3	-7.8	-2.9
Ac-Val-Pro-D-Ala-His-NH_2	-5.5	-5.0	-2.7
Ac-Val-Pro-D-Ser-Phe-NH_2	-5.0	-5.1	-2.9
Ac-Val-Pro-D-Ser-Ile-NH_2	-6.2	-4.6	-3.5
Ac-Val-Pro-D-Ala-Phe-NH_2	-5.8	-4.2	-4.5
Ac-Ala-Pro-D-Val-His-NH_2	-4.6	-3.8	-4.4
Ac-Ala-Pro-D-Ala-Phe-NH_2	-5.2	-3.8	-4.5
Ac-*tert*-Leu-Pro-D-Ala-His-NH_2	-6.0	-4.8	-4.2
Ac-Ala-Pro-Gly-Phe-NH_2	-5.5	-3.9	-3.9
Ac-Ala-Pro-D-Leu-Phe-NH_2	-5.3	-3.2	-5.4
Ac-Ala-Pro-D-Val-Phe-NH_2	-5.3	-3.9	-5.4

[a] In parts per billion per degree Kelvin.

Table 2-3. *Temperature Coefficients for Amide Proton Chemical Shifts in Water* ($\Delta\delta/\Delta T$)[a]

Peptide	i	i + 2	i + 3
Ac-Val-Pro-D-Ser-His-NH_2	-8.7	-9.4	-5.0
Ac-Val-Pro-D-Ala-His-NH_2	-7.9	-8.1	-4.9
Ac-Val-Pro-D-Ser-Phe-NH_2	-8.7	-8.8	-5.8
Ac-Val-Pro-D-Ser-Ile-NH_2	-8.1	-9.3	-6.2
[b]Ac-Val-NapPro-D-Ser-His-NH_2	-8.9	-7.0	-4.8
Ac-Val-Pro-D-Ala-Phe-NH_2	-8.4	-9.1	-6.2
Ac-Ala-Pro-D-Val-His-NH_2	-7.6	-8.8	-7.8
Ac-Ala-Pro-D-Ala-Phe-NH_2	-9.3	-8.5	-6.8
Ac-*tert*-Leu-Pro-D-Ala-His-NH_2	-7.6	-9.0	-6.5
Ac-Ala-Pro-Gly-Phe-NH_2	-6.1	-7.3	-7.5
[c]Ac-Ala-Pro-D-Leu-Phe-NH_2	-8.2	-9.1	-8.9
Ac-Ala-Pro-D-Val-Phe-NH_2	-9.9	-9.9	-8.7

[a] In parts per billion per degree Kelvin; [b] NapPro = *cis*-(2*S*,4*S*)-4-(naphth-2-yl)amido-proline; [c] In 10% DMSO/H_2O.

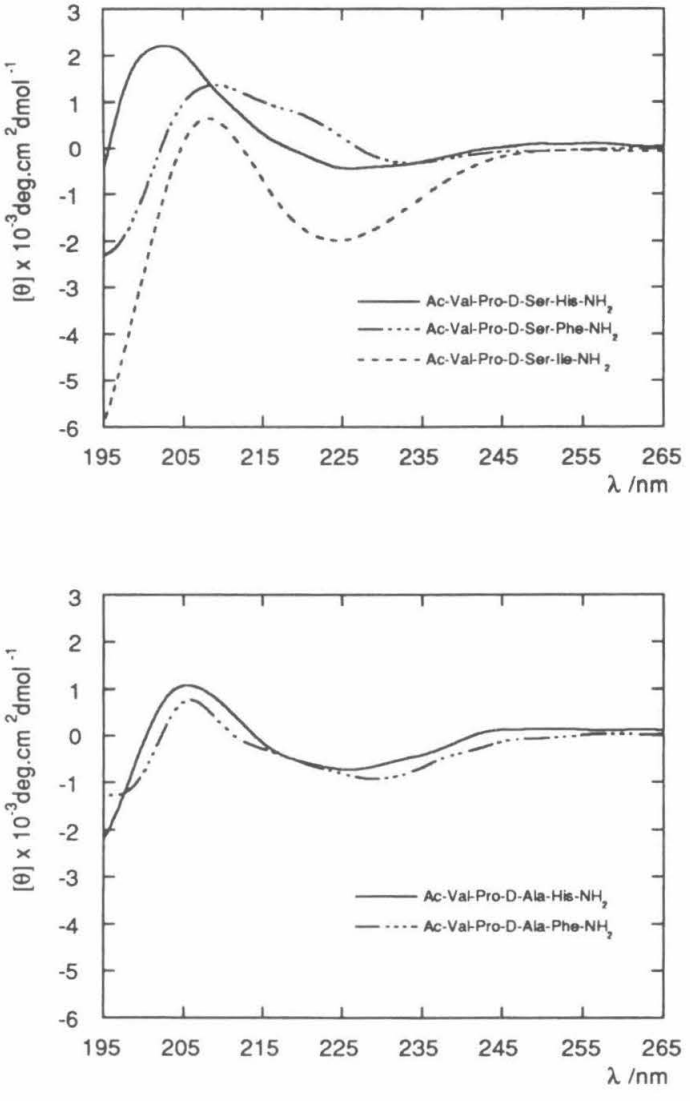


Figure 2-5. CD spectra in H_2O at 25° C of (A) Ac-Val-Pro-D-Ser-Xaa-NH_2 (Xaa = His, Phe and Ile); (B) Ac-Val-Pro-D-Ala-Xaa-NH_2 (Xaa = His and Phe)

Ac-Val-Pro-D-Ser-His-NH_2. The relative spectral intensities lend support to the NOE observations which suggested that the latter peptide was a stronger turn. The intensity of the positive ellipticity for the D-serine-containing peptide approaches a value for θ of 2300 $deg.cm^2\,dmol^{-1}$. (It should be noted that it is not possible to directly compare the CD spectra of peptides in this study with those calculated or in model systems, since the latter are based on only a core triamide unit.)

The CD spectrum of Ac-Val-Pro-D-Ser-His-NH_2 is invariant over the concentration range from 5μM-5mM; therefore, it is assumed that intermolecular interactions are not affecting the measurements. Also, the upper concentration limit for the CD studies approaches that used in the NMR experiments, so the observations made in the two experiments can be correlated.

Collectively, these spectroscopic observations indicate the presence of a population of an ordered conformation in Ac-Val-Pro-D-Ser-His-NH_2. Ac-Val-Pro-D-Ala-His-NH_2 also shows evidence of a similar structure, but to a lesser degree. In order to more fully understand the interactions that stabilize these particular sequences, a series of related peptides were studied based on the general structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH_2. Since turn formation is a complex process involving many interdependent effects, it is not feasible to carry out an exhaustive survey. However, these studies demonstrate that in this closely related series of peptides, the opportunity exists for a significant amount of conformational variability.

Conformational Consequences of Residue Substitution

Influence of Modifications at position (*i+3*)

In order to address the importance of histidine as the *i+3* residue, related peptides with single residue replacements were examined. All spectroscopic data in water was collected < pH 4.5, thus the histidine side chain would be ionized. Ac-Val-Pro-D-Ser-**Phe**-NH_2 and Ac-Val-Pro-D-Ser-**Ile**-NH_2 can be compared with Ac-Val-Pro-D-Ser-**His**-

NH_2, and Ac-Val-Pro-D-Ala-**Phe**-NH_2 can be compared with Ac-Val-Pro-D-Ala-**His**-NH_2.

NOE studies of these peptides in DMSO and water essentially indicate that a diminution of the turn population accompanies replacement of either D-serine at *i+2* or histidine at *i+3*. For example, ROESY spectra of Ac-Val-Pro-D-Ser-Ile-NH_2 in DMSO lack the $d_{\alpha N}$ [*i+1*, *i+3*] and long range connectivities observed in the corresponding histidine-containing peptide (See Figure 2-4). Also, the d_{NN} [*i+2*, *i+3*] interaction in Ac-Val-Pro-D-Ala-His-NH_2 is absent in Ac-Val-Pro-D-Ala-Phe-NH_2 in water. However, observation of the $d_{\alpha N}$ [*i+1*, *i+3*] connectivity may not be possible, even in a viable turn, as the predicted distance[27] for this interaction in a perfect type II turn is 3.3Å--a value which is close to the limits of detection in the ROESY experiment.[34] The VT data, presented in Tables 2-2 and 2-3, show similar trends, indicating that histidine at *i+3* has a positive, but not unique, effect on the conformation. In the unordered peptide, Ac-Ala-Pro-D-Val-His-NH_2 (also documented in Tables 2-2 and 2-3), the histidine NH temperature coefficient is typical of a solvent exposed proton. This latter observation rules out the possibility that the low coefficients are due only to a side chain specific effect.

CD studies for the D-serine and D-alanine containing peptides illustrated in Figures 2-5a and 2-5b, respectively, support the conclusions reached in NMR experiments. Essentially, the peptides which show NMR evidence of being better turns are characterized by a CD spectra with more Class B character.

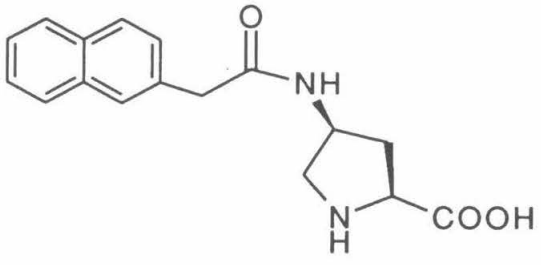
Influence of Modification at position (*i*)

The side chain substituent of the amino acid residue at the *i* position experiences steric crowding from the δ-methylene of the proline ring in a β-turn.[35] Therefore, the influence of the steric effects of residues was probed through a series of systematic variations of steric bulk in this position. Specifically, Ac-**Val**-Pro-D-Ala-Phe-NH_2 was compared with Ac-**Ala**-Pro-D-Ala-Phe-NH_2, and Ac-***tert*-Leu**-Pro-D-Ala-His-NH_2 was

compared with Ac-**Val**-Pro-D-Ala-His-NH_2. The ROESY spectrum of Ac-Val-Pro-D-Ala-Phe-NH_2 does not show any long range interactions indicative of a folded structure; therefore, it is likely that the population of folded structure(s) is low. However, the VT data for these two pairs of peptides in water and DMSO (Tables 2-2 and 2-3) and the CD spectra in water *collectively* provide an idea of the influence of the *i* residue substitutions. The CD spectra for each pair of peptides are illustrated in Figures 2-6a and 2-6b. In contrast to the spectra observed with the previous five peptides (in which CD spectra showed a distinct positive ellipticity maximum between 200 and 210 nm), Ac-Ala-Pro-D-Ala-Phe-NH_2 shows a very broad positive ellipticity ranging from 205-225 nm, and the *tert*-leucine-containing peptide exhibits no positive ellipticity. Thus it appears that either increasing or decreasing the steric bulk beyond valine perturbs the turn conformation in a negative way.

Effects of Proline Functionalization

Proline serves a key role in protein folding and structural stabilization due to the constrained ϕ dihedral angle imposed by the pyrrolidine ring.[3] In an effort to expand the utility of the proline residue to assume both structural and functional roles, a functionalized proline derivative was synthesized which contains a naphthyl group attached to the γ-carbon through an amide linker (See Appendix I). This residue, *cis*-(2*S*,4*S*)-4-(naphth-2-yl)-amido-proline (NapPro, see below) was designed to act primarily as a fluorescent probe in designed motifs.



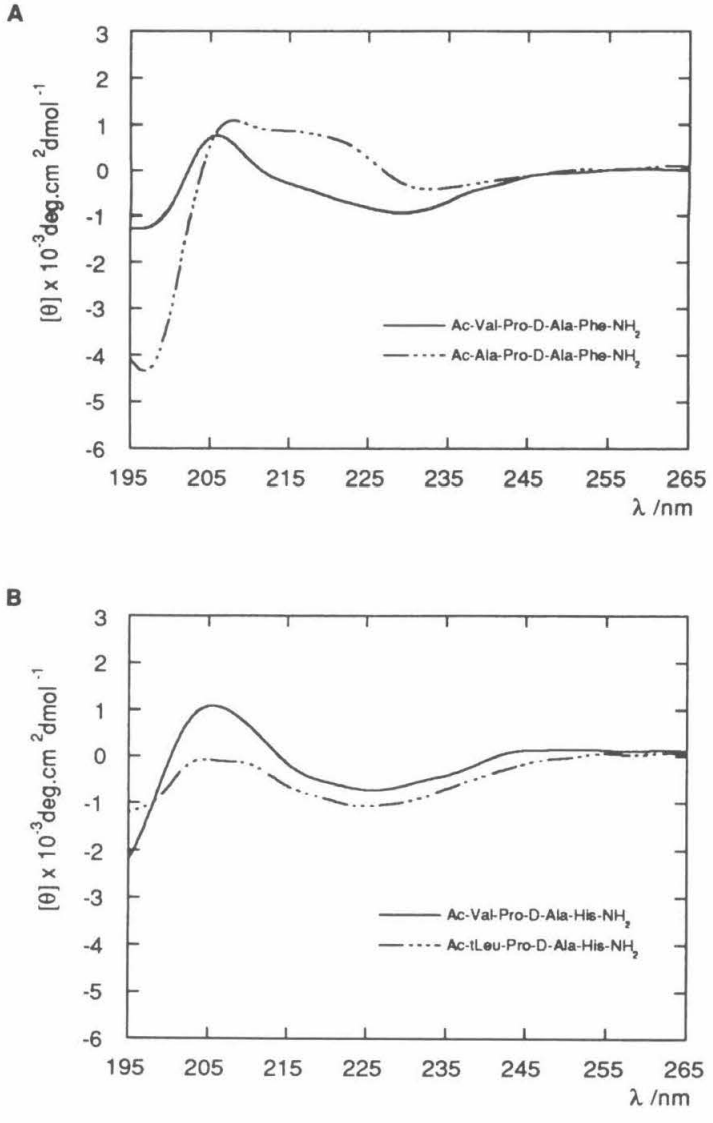


Figure 2-6. CD spectra in H_2O at 25° C of (A) Ac-Xaa-Pro-D-Ala-Phe-NH_2 (Xaa = Ala and Val); (B) Ac-Xaa-Pro-D-Ala-His-NH_2 (Xaa = *t*-Leu and Val)

The tetrapeptide Ac-Val-NapPro-D-Ser-His-NH_2 incorporating this unnatural residue was prepared and conformationally analyzed in water to test the effect of proline modification on the turn population.

Incorporation of the functionalized proline derivative does not substantially affect the turn integrity when compared to the parent sequence Ac-Val-Pro-D-Ser-His-NH_2. In water, the observed NOE connectivities are consistent with a significant population of type II β-turn structure and all of the sequential NOEs are present (**1**, **2**, and **3**) in addition to the long range connectivity d_{NN} (Ser, His) (**4**). In addition to these NOE connectivities, several long range connectivities were observed between the naphthyl ring protons and the side chain groups of positions *i* and *i+2* .[36] These connectivities suggest that the naphthyl group resides above the proline ring, between the side chains of the *i* and *i+2* residues. The variable temperature coefficients for Ac-Val-NapPro-D-Ser-His-NH_2 support the results from the ROESY experiment as the *i+3* NH (see Table 2-3) shows moderate solvent shielding as compared to the other amide protons within the peptide. However, the CD spectra for this peptide does not show the class B structure that would be expected from the nuclear magnetic resonance studies (See Figure 2-7). This marked difference can be attributed to the presence of the naphthyl group chromophore in the peptide; aromatic chromophores have been shown to affect circular dichroism spectra, especially when the group resides within a reverse turn.[3,37]

Influence of *(i+2)* Residue.

The effect of varying the steric bulk of the *i+2* residue was also investigated. For this purpose, a homologous series of four peptides, of general structure Ac-Ala-Pro-D-**Xaa**-Phe-NH_2 (**Xaa** = Gly, D-Ala, D-Val and D-Leu), were studied to see if a change in the *i+2* residue might alter the turn population. The temperature coefficients for the amide protons of these peptides in both DMSO and water are documented in Tables 2-2 and 2-3

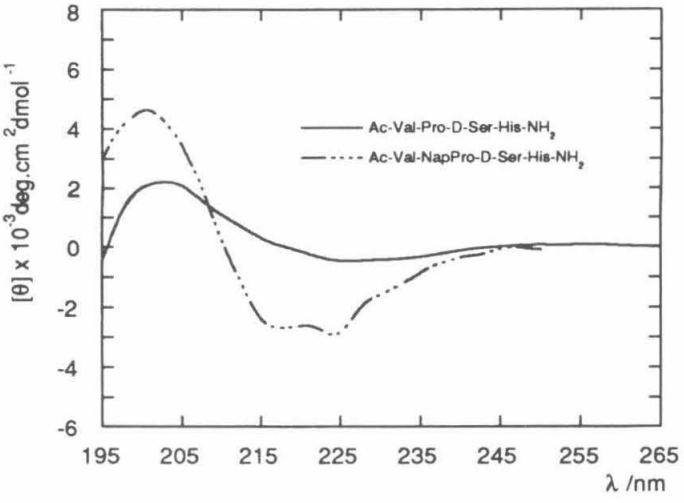


Figure 2-7. CD spectra in H_2O at 25° C of Ac-Val-NapPro-D-Ser-His-NH_2 and Ac-Val-Pro-D-Ser-His-NH_2

respectively. In DMSO there appears to be no type II β-turn population and, in fact, in some cases, the *i+2* NH coefficients show some solvent shielding (e.g., for Ac-Ala-Pro-D-**Leu**-Phe-NH_2; (-)3.2 ppb/K). This observation suggests the presence of an alternate folded structure such as a γ-turn or a reverse turn in which the amide NH of the D-leucine is hydrogen bonded to the carbonyl group of the acetyl cap. In water, changes in the steric demand of the *i+2* residues to the small, flexible residue glycine, or more sterically demanding residues D-valine or D-leucine, weaken the turn; the D-alanine-containing peptide has the lowest coefficient for the *i+3* NH. In water there is no indication of an alternate folded structure.

Figure 2-8 shows the CD spectra for this series of peptides in water. Following a similar trend to the water VT data in which Ac-Ala-Pro-D-Ala-Phe-NH_2 demonstrates a slight turn population, the CD spectrum shows a negative ellipticity at 197 nm and a broad positive ellipticity from 208 nm to 220 nm. The spectra of the glycine, D-valine and D-leucine analogs are quite varied and lack any positive ellipticity in the region characteristic of the type II turns.

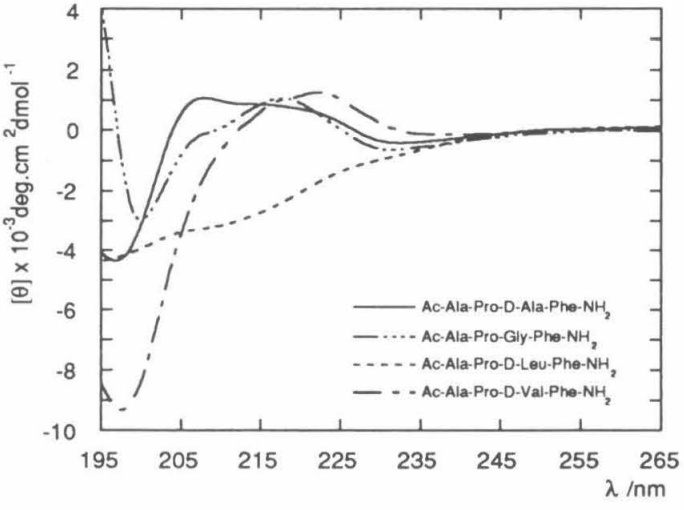


Figure 2-8. CD spectra in H_2O at 25° C of Ac-Ala-Pro-Xaa-Phe-NH_2 (Xaa = Gly, D-Ala, D-Val and D-Leu)

Discussion

The short linear peptides examined in this study could potentially adopt many conformations. However, the prolyl-D-amino acid dipeptide core of the sequence prevents adoption of either ordered α-helical or β-sheet structures; thus, it is reasonable to limit discussion to the three predominant types of interconverting conformers.[38] These would be represented by: (i) "open" (unfolded) structures with no hydrogen-bonding interactions; (ii) a γ-turn structure [*i+2* to *i* hydrogen bond]; and (iii) β-turn structures [*i+3* to *i* hydrogen bond]. The β-turn could be either a Type I or II structure; however, the latter is expected in the heterochiral sequence studied, since this conformation results in reduced steric interactions between the carbonyl oxygen of the central amide and the side chain of the *i+2* residue.[20]

The combined spectroscopic data for the peptide Ac-Val-Pro-D-Ser-His-NH_2 show evidence of a significant population of a folded conformation and allow identification of the likely conformation. The NOE experiments indicate the presence of a Type II β-turn structure. In both solvents, the observed NOE $d_{\alpha N}$ (Pro, Ser) (**1**, Figure 2) is very strong, indicating that these two protons are on the same face of the molecule as is characteristic of a type II turn (typical proton-proton distances in turns; type II 2.2Å, type I 3.4Å). In addition, the interaction between the δ-protons on the proline ring with the α-CH of the (i) residue is characteristic of a *trans* Xaa(*i*)-Pro(*i+2*) amide (in the *trans* isomer the $d_{\alpha\delta}$(Val, Pro) distance is 2.0-3.9Å, whereas in the *cis* isomer it is 4.3-5.0Å[27]). A small amount of a second conformer, assigned as the *cis* amide rotomer to proline (based on chemical shift data) was noted in the 1D ^{1}H NMR spectra. It is surprising that the contribution of the *cis* rotomer is as low as it is, particularly with a bulky residue such as valine in the *i* position. The *cis* amide rotamer often arises in those situations where steric congestion becomes an issue and the backbone cannot unfold to relieve the added strain. For example, the distribution of conformers for the peptide *cyclo*-(L-Val-L-Pro-Gly)$_2$ is biased very strongly in favor (>80% in D_2O or DMSO) of the two-*cis* amide isomer.[4]

Also, studies on linear, uncapped pentapeptides of general structure Tyr-Pro-L-Xaa-L-Yaa-L-Zaa (uncapped terminii) also show high proportions of the cis amide rotomer.[6] It appears that the balance of steric bulk at the *i*, *i+2* and *i+3* positions around the turn, coupled with the turn type, may be responsible for the shift in equilibrium between the two rotomers.

In the peptides examined, the side chain of the D *i+2* residue would be removed from the crowded face of the molecule in the Type II conformation. Both VT and CD studies are in agreement with the turn identification. In the VT experiments, the *i+3* amide NH is solvent shielded in both solvents. Furthermore, with regards to the CD studies, while either an open conformation or a Type I turn would be characterized by negative ellipticities in the 200-210 nm range, and the γ-turn would show a maximum at longer wavelengths[2] (approximately 230 nm), the Type II turn would be expected to show a CD pattern similar to that illustrated in Figure 2-5a. The intensity of the positive ellipticity for the D-serine-containing peptide approaches a value for θ of 2300 deg.cm^{2}dmol^{-1} which is not strong. However, the absolute intensity of the CD spectra for reverse turns is always much lower than that observed for the regular α-helix and β-sheet, since it is composed of a "non-recurring" arrangement of chromophores. In addition, low band intensities can be attributed to equilibrium between the turn conformation and unordered structures which show an intense negative band around 200 nm[39] and can therefore cancel out the intensity of the positive maximum.

The role of D-serine in stabilizing the type II β-turn motif is unclear. Although many reports have discussed Type I turn stabilization by L-serine,[40,41] there is only one discussion of the Pro-D-Ser dipeptide influence on conformation. Studies of Boc-Pro-D-Ser-NHMe in solution in carbon tetrachloride and in the solid state have suggested that the serine hydroxyl group may provide additional stabilization to the turn conformation through a hydrogen bonding interaction from the γ-O (Ser) to the amide NH;[38] however, the VT data for the blocked tetrapeptides in these studies do not indicate that this is the case in

water or DMSO. In homochiral sequences, residues such as serine and cysteine exhibit unique conformational properties which have been correlated with advantageous side chain interactions.[42] While the prolyl-D-serine dipeptide core of the sequence does seem to be a strong turn determinant, it is noteworthy that even in the absence of the extra hydroxyl functionality, the turn is present, albeit to a lesser degree.

The results from the peptide containing the functionalized proline derivative, NapPro, stand as a testament to the turn nucleation properties of the tetrapeptide sequence Ac-Val-Pro-D-Ser-His-NH_2. Incorporation of the bulky naphthyl group on the proline ring had significant β-turn character based on the observed variable temperature coefficients and the observed NOE connectivities; in fact, the results of these studies are comparable to those observed for Ac-Val-Pro-D-Ala-His-NH_2. While the effects of incorporation of more reactive functionalities on the proline ring will need to be examined, the results from this model system hold promise for the generation of designed motifs where proline derivatives can play active roles in both structural stabilization and pre-programmed functions.

The observed steric effects at the *i+2* position are consistent with the predicted dihedral angles for each residue, as derived from Ramachandran plots.[43] The ideal dihedral angles for the *i+2* residue in the type II turn are $\phi = 80°$ and $\psi = 0°$;[3] β-branching, or incorporation of sterically demanding groups at the *i+2* amino acid residue, would be expected to restrict the conformational space open to that residue, and hence prohibit assumption of the ideal torsional angles. The spectroscopic data (both CD and VT studies) indicate that the tetrapeptides with either D-leucine or D-valine at the *i+2* position exist largely in an open conformation in water. In contrast, VT studies in DMSO suggest some population of an alternate folded structure (Ac-Ala-Pro-D-Leu-Phe-NH_2 [$\Delta\delta/\Delta T$ = -3.2 ppb/K for *i+2* NH]). In the latter solvent, however, the ROESY studies did not provide further insight on the conformation of these peptides as no diagnostic NOEs were observed. Conformational analysis of simple dipeptides, of general structure *tert*-BuCO-Pro-D-Xaa-NHMe,[5] have indicated that the heterochiral dipeptides showed a high β-turn

ratio in aprotic solvents, regardless of the steric bulk of the *i+2* residue. The results of these studies suggest that the full effect of steric bulk at the *i+2* position in the type II turn may not be felt in the absence of complete amino acids at the *i* and *i+3* location around the motif.

The optimum residue at the *i* position, in the folded motif, should be one which favors an extended β-sheet conformation. In these investigations, we have focused principally on the influence of steric effects. It appears, from the limited study carried out thus far, that valine at the *i* position results in the highest turn population based on both CD and VT data. Increasing the steric demand of this residue to *tert*-leucine, or decreasing it to alanine, weakens the integrity of the turn. These steric effects are rather pronounced as would be anticipated with residues preceding proline. Schimmel and Flory[35] have noted that although the (ϕ,ψ) dihedral angles of residue *i* are independent of the neighboring residues (*i-1* and *i+1*), this is not true for residues adjacent to proline. In this case, interactions with the δCH_2 of proline affects the (ϕ,ψ) conformational space accessible to the preceeding residue significantly.

With regards to the influence of amino acid replacements at the *i+3* position, three amino acids have been considered. In this case, in contrast to the *i* residue effect, β-branching, (e.g., in the peptide, Ac-Val-Pro-D-Ser-Ile-NH_2) does not stabilize turn formation relative to the two unbranched *i+3* amino acid-containing peptides. While the spectroscopic data shows that histidine at the *i+3* position stabilizes the turn motif, the origin of this enhanced stability is still unknown. It is possible that the histidine imidazole ring forms a weak hydrogen bond with the central amide carbonyl oxygen atom; both the protonated Nδ and Nε atoms are within hydrogen bonding distance (depending on the value of $\chi^2 = \pm 90°$) when the histidine resides in the *trans* χ_1 rotamer. In support of this hypothesis, analyses of peptide crystal structures[44] have shown that the aromatic amino acids often adopt the *trans* rotamer for χ^1 and generally assume a planar orientation with respect to the peptide main chain ($\chi^2 = \pm 90°$). Further investigation will be necessary to

clarify the actual mechanism for the enhanced stability of the turn motif when histidine is incorporated at position *i+3*.

Conclusions

In general short linear peptides, in which both amino and carboxyl terminii are derivatized, are considered to be conformationally unstructured in solution in polar aprotic and protic media. This instability is attributed to the fact that competition for intermolecular hydrogen-bonding interactions by solvent tends to diminish any advantageous stabilization which might be gained from intramolecular hydrogen bond formation which accompanies a folded structure. This study has presented spectroscopic evidence to indicate that the short linear peptide Ac-Val-Pro-D-Ser-His-NH_2 shows an observable population of a folded conformation in both DMSO and water. From the NMR and CD studies, this conformation has been identified as a type II β-turn. In this peptide, we suggest that an adventitious combination of steric effects influences the conformational space accessible to the peptide backbone in such a way as to promote intramolecular hydrogen-bonding and folding of the turn motif. The specific contribution of the hydroxyl and imidazole side chains in stabilizing this reverse turn structure have yet to be elucidated.

The demonstration that a short peptide, which is terminal capped to emulate an internal sequence within a long polypeptide, can adopt a sufficiently stable folded structure to be observed in aqueous solution now sets the stage for utilizing this short peptide as an engineered structural nucleation site for *de novo* protein design and synthesis. The heterochiral sequence provides an additional advantage in this regard, since the prolyl-D-amino acid dipeptide pair cannot adopt any other ordered secondary structure such as the α-helix or β-sheet. In addition, the demonstration that the turn stability is observed with polar residues such as D-serine and histidine and to some degree with less polar residues such as D-alanine and phenylalanine also means that the turn motif could be used as an exposed turn on the surface of a globular, folded structure or as an internal turn element. Finally,

the fact a peptide containing a functionalized proline derivative still retains significant β-turn character in aqueous solution is particularly interesting, as it provides direct evidence that the derived turn sequences are amenable to functionalization without loss of structural integrity. The ability to incorporate reactive groups into the turn sequence is crucial to the design of functional motifs. The research described in Chapter 3 provides an extension to the work outlined herein, as it involves an examination of the effects of incorporation of functionality at position *i* of the turn sequence.

Acknowledgment

This work was performed in collaboration with Dr. Rex A. Moats and Thomas J. Prins. The main body of this work appeared in a full paper published in the *Journal of the American Chemical Society*:

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-NH_2: Spectroscopic Evidence for a Peptide with Significant β-Turn Character in Water and Dimethyl Sulfoxide," *J. Am. Chem. Soc.* **1992**, *114*, 3182-3188.

Experimental

Peptide Synthesis and Characterization

Commercially available starting materials and reagents for the synthesis of all tetrapeptides described were purchased from Milligen Biosearch or Sigma Chemical Co. The unnatural amino acid, NapPro, was synthesized as described in Appendix I. Peptides were synthesized on a 0.1-0.2 mmol scale by solid phase methods using N^{α}-9-fluorenylmethyloxycarbonyl (FMOC) protected amino acids[45,46] and benzotriazol-1-yloxy-tris(dimethylamino)phosphonium hexafluorophosphate/1-hydroxybenzotriazole (BOP/HOBt)[47] mediated amide coupling chemistry on a Milligen 9050 automated peptide synthesizer. PAL resin[48] (Milligen GEN077483 substituted at ca. 0.30 meq/g) was used to afford amides at the carboxyl terminus. In most cases, 2.5 amino acid equivalents were used per coupling; however, in cases where low coupling efficiency was expected (as in the incorporation of the unnatural amino acid, NapPro), 4.0 equivalents were employed. Activated esters were formed *in situ* using BOP, HOBt and 0.451 M *N*-methyl-morpholine in dimethylformamide (DMF). Coupling times varied from 0.5-2.0 hr depending upon coupling efficiency of the particular amino acid. Deprotection of FMOC-protected amine groups was performed using a 7 min. 20% piperidine/DMF wash. Peptides were *N*-acetylated on the resin using 20 equivalents of acetic anhydride and 5 equivalents of triethylamine in 3 mL DMF. After shaking for 2 hr, the resin was washed with dichloromethane and air dried.

Peptides were deprotected and cleaved from the resin by treatment with trifluoroacetic acid (TFA)/thioanisole/ethanedithiol/anisole (90:5:3:2) for 2 hr. After filtration of the resin, the combined filtrates were concentrated to ca. 2 mL volume and precipitated with ether/hexane 2:1. The supernatant was decanted and the peptides were washed with ether/hexane 2:1 (3x20 mL). The peptides were repeatedly lyophilyzed until no traces of the cleavage mixture remained. Purification of the peptides was accomplished

using reverse phase HPLC (C_{18}) Gradient elution with 0.08% TFA in acetonitrile added to 0.1% TFA in water when necessary.

The purity of each peptide was assessed to be >95% by thin layer chromatography (TLC) and High Pressure Liquid Chromatography (HPLC). TLC was carried out using the following eluent systems: A. *n*-butanol: acetic acid: water (4:1:1) and B: *n*-butanol: acetic acid: water: ethyl acetate (1:1:1:1). TLC plates were visualized with I_2 vapors, UV lamps, or by exposure to Cl_2 vapors followed by treatment with a solution of *o*-toluidine in acetic acid with KI.[49] In some peptides, the presence of a small quantity (<10%) of an alternate conformational isomer was apparent by ^1H NMR.

Chemical shifts of NMR spectra are reported as δ units ppm relative to DMSO-d_6 (2.49 ppm) for ^1H and (39.5 ppm) for ^{13}C in DMSO-d_6 as solvent and DMSO (2.73 ppm) as an internal standard in 10% D_2O/90% H_20 solutions at 5 °C, pH= 3.6 ± 0.6.

Spectroscopic Studies

CD Studies

Peptide solutions were prepared in double-distilled H_2O (4 mM concentration) and degassed using four freeze-pump-thaw cycles and stored under argon. Standard peptide spectral samples were prepared by serial dilution with degassed double-distilled H_2O (degassed by argon sparging for 1 hr) to 50 mM and analyzed in a 1.0 cm path length quartz cell. For aggregation studies, CD spectra were collected at concentrations ranging from 5 μM to 5 mM in quartz cells varying in path length from 0.01-1.0 cm. Both the optics and sample chamber were flushed continuously with dry N_2 throughout each experiment. All spectra were obtained from 280 nm to 195 nm at a scan speed of 50 nm/min, a time constant of 0.5 sec, and a bandwidth of 1 nm. In all cases a minimum of 8 scans were taken. Data was processed on a Macintosh IIci computer with KaleidaGraph software, version 2.1.1. Values of θ were calculated in ellipticity per mole of peptide residue with the blocked tetrapeptides treated as pseudohexapeptides.

NMR Studies

^1H resonance peak assignments were made either in one-dimension (1-D) by homonuclear single frequency decoupling experiments or by the use of two-dimensional correlation spectroscopy (2D COSY) experiment. NOEs were detected using the 2D spin-locked ROESY experiment, which is necessary for the detection of NOE effects in relatively small systems (MW 400-1500).[24-26,49] It should be noted that the ROESY cross peaks, while useful in obtaining qualitative proximity information (<3 Å) cannot be easily used to derive quantitative distance information in the same way as 2D nuclear Overhauser effect spectroscopy (NOESY) data.[34]

2D COSY experiments were run in either magnitude mode or phase sensitive mode and the FID's were multiplied by a phase shifted sine bell apodization function prior to Fourier transformation. 2D ROESY experiments were run in phase sensitive mode using TPPI (Time Proportional Phase Increments). For the ROESY experiments the spinlock power setting was calculated by determining the power of a 125 microsecond 90 degree pulse. The experimental parameters, such as mixing time, water suppression parameters, spin-lock power and offset frequency were varied to optimize amide cross peak intensity and to minimize HOHAHA artifacts[50] using the model peptide Tyr-Pro-Gly-Asp-Val which had previously been shown to exhibit the NOEs of interest in this study.[6] In ROESY experiments on the AM500 spectrometer the spinlock was applied through the decoupler channel using normal mode with O1/O2 coherence. The FIDs were multiplied by a phase shifted squared sine bell apodization function prior to Fourier transformation. In a typical COSY or ROESY experiment the relaxation delay was 1.3 s and the transmitter offset was positioned in the center of the spectrum or on the water resonance and the data matrix consisted of 512 t_1 increments containing 2K complex points. All COSY and ROESY experiments run in 90/10 H_20:D_2O, solvent suppression were obtained using presaturation at 278K. (Note: the reported 90/10 H_20:D_2O proton chemical shifts were also recorded at this temperature.)

The temperature dependence of the assigned amide proton shifts was determined between 300 K and 325 K in DMSO and 278 K and 310 K in H_2O. In all cases the chemical shifts were found to vary linearly with temperature. A minimum of six temperature steps were recorded in each experiment. Temperature calibration of the actual probe temperature was performed using either an ethylene glycol or methanol standard.

Ac-Val-Pro-D-Ser-His-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.95 His-ε (s, 1H), 8.53 Ser-NH (d, 1H), 8.05 Val-NH (d, 1H), 8.01 His-NH (d, 1H), 7.35 Cap-NH (s, 1H), 7.27 His-δ & Cap-NH (m, 2H), 4.42 His-α (m, 1H), 4.34 Pro-α (m, 1H), 4.20 Val-α (m, 1H), 4.17 Ser-α (m, 1H), 3.86 Pro-δ (m, 1H), 3.71 Ser-β (m, 1H), 3.65 Ser-β (m, 1H), 3.60 Pro-δ (m, 1H), 3.22 His-β (m,1H), 3.00 His-β (m, 1H), 2.05 Pro-β (m, 1H), 1.98 Pro-γ (m,1H), 1.82 Pro-γ (m, 1H), 1.82 Ac-cap (s, 3H), 1.80 Pro-β (m, 1H), 1.75 Val-β (m, 1H), 0.90 Val-γ (m, 6H).

^{1}H NMR (10% D_2O/ 90% H_2O) δ: 8.70 Ser-NH (d, 1H), 8.50 His-NH & His-ε (m, 2H), 8.25 Val-NH (d, 1H), 7.52 Cap-NH (s, 1H), 7.21 Cap-NH (s, 1H), 7.15 His-δ (s, 1H), 4.60 His-α (m, 1H), 4.38 Ser-α (m, 1H), 4.32 Pro-α (m, 1H), 4.23 Val-α (m, 1H), 3.85 Pro-δ (m, 1H), 3.75 Ser-β (m, 1H), 3.67 Ser-β (m, 1H), 3.55 Pro-δ (m, 1H), 3.22 His-β (m,1H), 3.00 His-β (m, 1H), 2.21 Pro-β, (m, 1H), 2.00 Pro-γ (m, 1H), 1.90 Pro-γ (m, 1H), 1.88 Ac-cap (s, 3H), 1.85 Pro-β (m, 1H), 1.82 Val-β (m, 1H), 0.83 Val-γ (d, 3H), 0.80 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.2, 171.8, 171.5, 170.7, 169.1, 133.0, 129.9, 116.6, 61.1, 60.0, 56.1, 55.3, 51.8, 47.5, 30.0, 29.1, 26.3, 24.5, 22.1, 18.7.

HRMS: Calcd. for $[MH]^+$ $C_{21}H_{34}N_7O_6$ [480.2570]; Obsd. [480.2586].

TLC (B) R_f 0.44.

Ac-Val-Pro-D-Ala-His-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.95 His-ε (s, 1H), 8.37 Ala-NH (d, 1H), 7.95 His-NH (d, 1H), 7.90 Val-NH (d, 1H), 7.29 Cap-NH (s, 1H), 7.27 His-δ (s, 1H), 7.23 Cap-NH (s, 1H), 4.44 His-α (m, 1H), 4.22 Val-α (m, 1H), 4.20 Pro-α (m, 1H), 4.17 Ala-α (m, 1H), 3.84 Pro-δ (m, 1H), 3.56 Pro-δ (m, 1H), 3.21 His-β (m,1H), 2.96 His-β (m, 1H), 2.08 Pro-β (m, 1H), 1.95 Pro-γ (m,1H), 1.84 Pro-γ (m, 1H), 1.82 Ac-cap (s, 3H), 1.78 Pro-β & Val-β (m, 2H), 1.16 Ala-β (d, 3H), 0.82 Val-γ (d, 6H).

^{1}H NMR (10% D_2O/ 90% H_2O) δ: 8.71 Ala-NH (d, 1H), 8.62 His-ε (s, 1H), 8.56 His-NH (d, 1H), 8.36 Val-NH (d, 1H), 7.71 Cap-NH (s, 1H), 7.33 Cap-NH (s, 1H), 7.30 His-δ (s, 1H), 4.67 His-α (m, 1H), 4.38 Pro-α (m, 1H), 4.36 Val-α (m, 1H),4.27 Ala-α (m, 1H), 3.96 Pro-δ (m, 1H), 3.69 Pro-δ (m, 1H), 3.35 His-β (m, 1H), 3.13 His-β (m, 1H), 2.31 Pro-β (m, 1H), 2.09 Pro-γ (m, 1H), 2.00 Pro-γ, Val-β, & Ac-cap (m, 5H), 1.90 Pro-β (m, 1H), 1.30 Ala-β (d, 3H), 0.97 Val-γ (d, 3H), 0.93 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 171.86, 171.66, 171.64, 170.68, 169.22, 133.48, 129.87, 116.63, 60.09, 56.05, 51.65, 48.25, 47.41, 29.65, 29.07, 26.60, 24.62, 22.12, 18.75, 18.64, 17.25.

MS: Calcd. for $[MH]^+$ (Low Res.) $C_{21}H_{34}N_7O_5$ [464]; Obsd. [464].

TLC (B) R_f 0.35.

Ac-Val-Pro-D-Ser-Phe-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.24 Ser-NH (d, 1H, J = 8.1 Hz), 8.07 Val-NH (d, 1H, J = 8.3 Hz), 7.90 Phe-NH (d, 1H, J = 8.8 Hz), 7.28 Cap-NH (s, 1H), 7.18 Phe-ring & Cap-NH (m, 6H), 4.33 Phe-α & Pro-α (m, 2H), 4.23 Val-α (m, 1H), Ser-α 4.17 (m, 1H), 3.85 Pro-δ (m, 1H), 3.52 Pro-δ & Ser-β (m, 3H), 3.07 Phe-β (m, 1H), 2.85 Phe-β (m, 1H), 2.04 Pro-β (m, 1H), 1.95 Pro-γ (m, 1H), 1.86 Pro-β, Pro-γ & Val-β (m, 3H), 1.83 Ac-cap (s, 3H), 0.86 Val-γ (d, 3H), 0.82 Val-γ (d, 3H).

^{1}H NMR (10% D_2O/90% H_2O) δ: 8.46 Ser-NH (d, 1H, J = 8.1 Hz), 8.28 Phe-NH (d, 1H, J = 7.3 Hz), 8.14 Val-NH (d, 1H, 7.8), 7.59 Cap-NH (s, 1H), 7.33 Phe-ring (m, 5H), 7.09 Cap-NH (s, 1H), 4.58 Phe-α (m, 1H), 4.45 Pro-α (m, 1H), 4.39 Val-α & Ser-α (m, 2H), 3.94 Pro-δ (m, 1H), 3.71 Pro-δ & Ser-β (m, 3H), 3.23 Phe-β (m, 1H), 3.01 Phe-β (m, 1H), 2.30 Pro-β (m, 1H), 2.05 Pro-γ & Val-β (m, 3H), 2.02 Ac-cap (s, 3H), 1.95 Pro-β (m, 1H), 1.00 Val-γ (d, 3H), 0.94 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.7, 171.8, 170.6, 169.5, 169.2, 138.3, 129.1, 127.9, 126.0, 61.2, 59.8, 56.0, 54.9, 54.3, 47.3, 37.2, 29.8, 29.0, 24.6, 22.1, 18.9, 18.7.

HRMS: Calcd. for $[MH]^+$ $C_{24}H_{36}N_5O_6$ [490.2666]; Obsd. [490.2674].

TLC (B) R_f 0.50.

Ac-Val-Pro-D-Ser-Ile-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.01 Ser-NH (d, 1H, J = 7.7 Hz), 7.99 Val-NH (d, 1H, J = 7.3 Hz), 7.24 Ile-NH (d, 1H) 7.16 Cap-NH (s, 1H), 7.00 Cap-NH (s, 1H), 4.18 Pro-α (m, 1H) 4.16 Ser-α (m, 1H), 4.15 Val-α (m, 1H), 4.04 Ile-α (m, 1H), 3.4 Pro-δ (m, 1H) 3.30 Pro-δ & Ser-β (m, 3H), 2.03 Pro-β (m, 1H), 1.95 Pro-γ (m, 2H), 1.92 Ac-cap (s, 3H), 1.90 Val-β & Pro-β (m, 2H) 1.87 Ile-β (m, 1H), 1.18 Ile-γ1 (m, 1H), 1.03 Ile-γ1 (m, 1H), 0.94-90 Val-γ, Ile-γ2, & Ile-δ (m, 12H).

^{1}H NMR (10% D_2O/90% H_2O) δ: 8.20 Ile-NH (d, 1H), 7.85 Cap-NH (s, 1H), 7.65 Ser-NH (d, 1H), 7.35 Val-NH (d, 1H), 7.27 Cap-NH (s, 1H), 4.49 Pro-α (m, 1H), 4.47 Ser-α (m, 1H) 4.38 Val-α (m, 1H), 4.22 Ile-α (m, 1H), 3.99 Pro-δ (m, 1H), 3.90 Ser-β (m, 2H), 3.70 Pro-δ (m, 1H), 2.32 Pro-β (m, 1H), 2.05 Pro-γ (m, 2H), 2.02 Ac-cap (s, 3H), 1.95 Pro-β, Val-β, & Ile-β (m, 3H), 1.42 Ile-γ1 (m, 1H), 1.21 Ile-γ1 (m, 1H), 1.00 Val-γ (d, 3H), 0.98 Val-γ & Ile-γ2 (m, 6H), 0.88 Ile-δ (m, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.7, 171.7, 170.4, 169.5, 169.2, 61.5, 59.6, 57.0, 56.0, 54.9, 47.2, 36.3, 29.8, 29.1, 24.5, 24.0, 22.1, 19.0, 18.6, 15.4, 10.9.

HRMS: Calcd. for $[MH]^+$ $C_{21}H_{38}N_5O_6$ [456.2822]; Obsd. [456.2820].

TLC (A) $R_f = 0.65$.

Ac-Val-Pro-D-Ala-Phe-NH_2

^1H NMR (DMSO-d_6) δ: 8.11 Ala-NH (d, 1H, J = 7.8 Hz), 8.08 Val-NH (d, 1H, J = 8.4 Hz), 8.03 Phe-NH (d, 1H, J = 8.6 Hz), 7.37 Cap-NH (s, 1H), 7.26 Phe-ring (m, 5H), 7.16 Cap-NH (s, 1H), 4.35 Phe-α (m, 1H), 4.27 Val-α (m, 1H), 3.84 Pro-α (m, 1H), 3.11 Ala-α (m, 1H), 3.12 Pro-δ (m, 1H), 2.83 Pro-δ (m, 1H), 2.12 Phe-β (m, 1H), 1.98 Phe-β (m, 1H), 1.87 Ac-cap, Pro-β, Pro-γ, & Val-β (m, 8H), 1.05 Ala-β (d, 3H, J = 7.0 Hz), 0.91 Val-γ (d, 3H, J = 6.5 Hz), 0.87 Val-γ (d, 3H, J = 6.6 Hz).

^1H NMR (10% D_2O/90% H_2O) δ: 8.54 Ala-NH (d, 1H, J = 6.5 Hz), 8.37 Phe-NH (d, 1H, J = 7.3 Hz), 8.30 Val-NH (d, 1H, J = 7.3 Hz), 7.65 Cap-NH (s, 1H), 7.40 Phe-ring (m, 2H), 7.30 Phe-ring (m, 3H), 7.20 Cap-NH (s, 1H), 4.57 Phe-α (m, 1H), 4.38 Val-α (m, 1H), 4.36 Pro-α (m, 1H), 4.22 Ala-α (m, 1H), 3.96 Pro-δ (m, 1H), 3.71 Pro-δ (m, 1H), 3.25 Phe-β (m, 1H), 3.00 Phe-β (m, 1H), 2.29 Pro-β (m, 1H), 2.1-1.95 Pro-γ & Val-β(m, 3H), 2.0 Ac-cap (s, 3H), 1.91 Pro-β (m, 1H) 1.17 Ala-β (d, 3H), 0.99 Val-γ (d, 3H), 0.91 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.9, 171.5, 171.2, 170.6, 169.2, 138.3, 129.1, 127.9, 126.1, 59.9, 55.9, 54.1, 47.9, 47.32, 37.3, 29.8, 29.1, 24.6, 22.2, 19.0, 18.7, 17.8;

HRMS: Calcd. for $[MH]^+$ $C_{24}H_{36}N_5O_5$ [474.2716]; Obsd. [474.2706].

TLC (A) $R_f = 0.57$.

Ac-Ala-Pro-D-Val-His-NH_2

^1H NMR (DMSO-d_6) δ: 8.95 His-ε (s, 1H), 8.22 Ala-NH (d, 1H, J = 7.3 Hz), 8.09 His-NH (d, 1H, J = 7.0 Hz), 8.0 Val-NH (d, 1H J = 7.3 Hz), 7.28 Cap-NH (s, 1H), 7.25

His-δ (s, 1H), 7.22 Cap-NH (s, 1H), 4.50 Ala-α (m, 1H), 4.49 His-α (m, 1H), 4.35 Pro-α (m, 1H), 3.78 Val-α (m, 1H), 3.65 Pro-δ (m, 1H), 3.52 Pro-δ (m, 1H), 3.20 His-β (m, 1H), 2.95 His-β (m, 1H), 2.05 Pro-β (m, 1H), 1.92 Val-β & Pro-γ (m, 2H), 1.85 Pro-γ (m, 1H), 1.80 Ac-cap (s, 3H), 1.75 Pro-β (m, 1H), 1.11 Ala-β (d, 3H), 0.7 Val-γ (d, 3H), 0.62 Val-γ (d, 3H).

^{1}H NMR (10% D_2O/ 90% H_2O) δ: 8.85 His-NH (d, 1H, J = 8.1 Hz), 8.54 His-ε (s, 1H), 8.46 Val-NH (d, 1H, J = 7.4 Hz), 8.39 Ala-NH (d, 1H, J = 5.3 Hz), 7.67 Cap-NH (s, 1H), 7.36 Cap-NH (s, 1H), 7.33 His-δ (s, 1H), 4.76 His-α (m, 1H), 4.53 Ala-α (m, 1H), 4.49 Pro-α (m, 1H), 4.08 Val-α (m, 1H), 3.85 Pro-δ (m, 1H) 3.68 Pro-δ (m, 1H), 3.40 His-β (m, 1H), 3.17 His-β (m, 1H), 2.32 Pro-β (m, 1H), 2.07 Pro-γ (m, 2H) 1.99 Ac-cap (s, 3H), 1.90 Val-β & Pro-β (m, 2H), 1.32 Ala-β (d, 3H) 0.8 Val-γ (d, 3H), 0.71 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.2, 171.9, 171.2, 170.8, 168.7, 133.7, 130.1, 117.1, 59.7, 58.2, 51.7, 46.8, 46.0, 29.3, 29.1, 27.0, 24.6, 22.2, 18.9, 17.6, 16.8.

HRMS: Calcd. for $[MH]^+$ $C_{21}H_{34}N_7O_5$ [464.2621]; Obsd. [464.2598].

TLC (A) R_f = 0.30.

Ac-*tert*-Leu-Pro-D-Ala-His-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.07 Ala-NH (d, 1H, J = 7.0 Hz), 7.98 His-NH (d, 1H, J = 8.2 Hz), 7.92 *t*-Leu-NH (d, 1H, J = 8.9 Hz), 7.53 His-ε (s, 1H), 7.25 Cap-NH (s, 1H), 7.07 Cap-NH (s, 1H), 6.76 His-δ (s, 1H), 4.53 *t*-Leu-α (d, 1H, J = 8.9 Hz), 4.36 Pro-α & His-α (m, 2H), 4.26 Ala-α (m, 1H), 3.76 Pro-δ (m, 1H), 3.65 Pro-δ (m, 1H), 3.00 His-β (m, 1H), 2.83 His-β (m, 1H), 2.06 Pro-β (m, 1H), 1.87 Pro-β, Pro-γ & Ac-cap (m, 6H), 1.14 Ala-β (d, 3H, J = 7.0 Hz), 0.97 *t*-Leu-γ (s, 9H).

^{1}H NMR (10% D_2O/ 90% H_2O) δ: 8.63 His-ε (s, 1H), 8.59 Ala-NH (d, 1H), 8.56 His-NH (d, 1H), 8.16 *t*-Leu-NH (d, 1H), 7.70 Cap-NH (s, 1H), 7.36 His-δ (s, 1H), 7.31

Cap-NH (s, 1H), 4.61 His-α (m, 1H), 4.50 *t*-Leu-α (d, 1H), 4.40 Pro-α (m, 1H), 4.25 Ala-α (m, 1H), 3.95 Pro-δ (m, 1H), 3.66 Pro-δ (m, 1H), 3.37 His-β (m, 1H), 3.17 His-β (m, 1H), 2.10 Pro-β (m, 1H), 2.08 Pro-γ (m, 1H), 2.06 Ac-cap (s, 3H), 1.98 Pro-γ (m, 1H), 1.92 Pro-β (m, 1H), 1.29 Ala-β (d, 3H), 1.01 *t*-Leu-γ (s, 9H).

^{13}C NMR (DMSO-d_6) δ: 171.9, 171.4, 168.9, 168.6, 168.1, 129.5, 116.2, 57.3, 51.3, 50.4, 47.9, 47.3, 46.1, 33.8, 28.8, 25.9, 25.1, 24.1, 21.8, 17.2, 16.2.

MS: Calcd. for $[MH]^+$ (Low Res.) $C_{22}H_{36}N_7O_5$ [478]; Obsd. [478].

TLC (A) R_f = 0.36.

Ac-Ala-Pro-D-Ala-Phe-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.15 L-Ala-NH (d, 1H, J = 7.3 Hz), 8.09 D-Ala-NH (d, 1H, J = 7.6 Hz), 8.04 Phe-NH (d, 1H, J = 8.7 Hz), 7.36 Cap-NH (s, 1H), 7.25 Phe-ring (m, 5H), 7.12 Cap-NH (s, 1H), 4.53 L-Ala-α (m, 1H), 4.38 Phe-α (m, 1H), 4.27 Pro-α (m, 1H), 4.18 D-Ala-α (m, 1H), 3.69 Pro-δ (m, 1H), 3.58 Pro-δ (m, 1H), 3.11 Phe-β (m, 1H), 2.84 Phe-β (m, 1H), 2.08 Pro-β (m, 1H), 1.96 Pro-γ (m, 1H), 1.84 Ac-Cap & Pro β,γ (m, 5H), 1.18 L-Ala-β (d, 3H, J = 6.9 Hz), 1.05 D-Ala-β (d, 3H, J = 7.8 Hz).

^{1}H NMR (10% D_2O/ 90% H_2O) δ: 8.57 D-Ala-NH (d, 1H), 8.40 L-Ala-NH (d, 1H), 8.39 Phe-NH (d, 1H) 7.66 Cap-NH (s, 1H), 7.37 Phe-ring (m, 2H), 7.31 Phe-ring (m, 1H), 7.27 Phe-ring (m, 2H), 7.23 Cap-NH (s, 1H), 4.56 Phe-α (m, 1H), 4.54 L-Ala-α (m, 1H), 4.40 Pro-α (m, 1H), 4.22 D-Ala-α (m, 1H), 3.85 Pro-δ (m, 1H), 3.65 Pro-δ (m, 1H), 3.25 Phe-β (m, 1H), 3.00 Phe-β (m, 1H), 2.30 Pro-β (m, 1H), 2.08 Pro-γ (m, 2H), 1.98 Ac-cap (s, 3H), 1.90 Pro-β (m, 1H) 1.34 D-Ala-β (d, 3H), 1.18 L-Ala-β (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.9, 171.6, 171.5, 171.2, 168.8, 138.3, 129.2, 127.9, 126.1, 59.9, 54.1, 48.2, 46.8, 46.1, 37.3, 28.9, 24.6, 22.3, 17.6, 16.8.

HRMS: Calcd. for $[MH]^+$ $C_{22}H_{32}N_5O_5$ [446.2403]; Obsd. [446.2408].

TLC (A) R_f = 0.51.

Ac-Ala-Pro-Gly-Phe-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.21 Gly-NH (t, 1H, J = 5.6 Hz), 8.16 Phe-NH (d, 1H, J = 7.4 Hz), 7.90 Ala-NH (d, 1H, J = 8.5 Hz), 7.36 Cap-NH (s, 1H), 7.25 Phe-ring (m, 5H), 7.18 Cap-NH (s, 1H), 4.55 Phe-α (m, 1H), 4.41 Ala-α (m, 1H), 4.28 Pro-α (m, 1H), 3.75-3.42 Pro-δ & Gly-α (m, 4H), 3.09 Phe-β (m, 1H), 2.85 Phe-β (m, 1H), 2.07 Pro-β (m, 1H), 1.95 Pro-γ (m, 1H), 1.84 Ac-cap & Pro-γ, β (m, 5H), 1.18 Ala-β (d, 3H, J = 6.87 Hz).

^{1}H NMR (10% D_2O/90% H_2O) δ: 8.60 Ala-NH (d, 1H), 8.40 Gly-NH (dd, 1H, J = 5.5 Hz), 8.22 Phe-NH (d, 1H, J = 7.3 Hz), 7.70 Cap-NH (s, 1H), 7.40 Phe-ring (m, 2H), 7.32 Phe-ring (m, 1H), 7.22 Phe-ring (m, 2H), 7.18 Cap-NH (s, 1H), 4.59 Ala-α (m, 1H), 4.57 Phe-α (m, 1H), 4.40 Pro-α (m, 1H), 3.90 Gly-α (dd, 1H), 3.86 Pro-δ (m, 1H), 3.84 Gly-α (dd, 1H), 3.68 Pro-δ (m, 1H), 3.20 Phe-β (m, 1H), 3.05 Phe-β (m, 1H), 2.30 Pro-β (m, 1H), 2.06 Pro-γ (m, 2H), 2.01 Ac-cap (s, 3H), 1.85 Pro-β (m, 1H), 1.33 Ala-β (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 172.7, 172.0, 171.0, 168.7, 168.4, 138.0, 129.0, 127.9, 126.1, 59.9, 53.9, 46.7, 45.9, 42.0, 37.3, 28.9, 24.4, 22.1, 16.5.

HRMS: Calcd. for $[MH]^+$ $C_{21}H_{30}N_5O_5$ [432.2247]; Obsd. [432.2250].

TLC (A) R_f = 0.70.

Ac-Ala-Pro-D-Val-Phe-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.24 Phe-NH (d, 1H, J = 8.7 Hz), 8.14 Ala-NH (d, 1H, J = 7.5 Hz), 7.72 Val-NH (d, 1H, J = 8.6 Hz), 7.36 Cap-NH (s, 1H), 7.26 Phe-ring (m, 4H), 7.18 Phe-ring (m, 1H), 7.14 Cap-NH (s, 1H), 4.55 Ala-α (m, 1H), 4.46 Phe-α (m, 1H), 4.40 Pro-α (m, 1H), 4.09 Val-α (m, 1H), 3.68 Pro-δ (m, 1H), 3.59 Pro-δ (m, 1H), 3.14 Phe-β (m, 1H), 2.78 Phe-β (m, 1H), 2.04 Pro-β (m, 1H), 1.84 Ac-cap, Pro-β, Pro-γ & Val-β (m, 7H), 1.20 Ala-β (d, 3H, J = 6.9 Hz), 0.58 Val-γ (dd, 6H).

^{1}H NMR (10% D_2O/90% H_2O) δ: 8.68 Phe-NH (d, 1H, J = 7.9 Hz), 8.41 Ala-NH (d, 1H, J = 5.4 Hz), 8.30 Val-NH (d, 1H, J = 7.9 Hz), 7.68 Cap-NH (s, 1H), 7.33 Phe-ring (m, 2H), 7.30 Phe-ring (m, 3H), 7.26 Cap-NH (s, 1H), 4.65 Phe-α (m, 1H), 4.54 Ala-α (m,1H), 4.46 Pro-α (m, 1H), 4.08 Val-α (m, 1H), 3.83 Pro-δ (m, 1H), 3.66 Pro-δ (m, 1H), 3.30 Phe-β (m, 1H), 2.94 Phe-β (m, 1H), 2.32 Pro-β (m, 1H), 2.02 Pro-γ (m, 2H), 1.99 Ac-cap (s, 3H), 1.92 Pro-β (m, 1H), 1.88 Val-β (m, 1H), 1.35 Ala-β (d, 3H), 0.69 Val-γ (d, 3H), 0.60 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 173.0, 171.5, 171.0, 170.5, 168.7, 138.2, 129.0, 127.9, 126.0, 59.5, 57.6, 54.0, 46.6, 45.9, 37.2, 29.9, 29.0, 24.4, 22.1, 19.0, 17.3, 17.0.

HRMS: Calcd. for $[MH]^+$ $C_{24}H_{36}N_5O_5$ [474.2716]; Obsd. [474.2716].

TLC (A) R_f = 0.75.

Ac-Ala-Pro-D-Leu-Phe-NH_2

^{1}H NMR (DMSO-d_6) δ: 8.15 Ala-NH (d, 1H, J = 7.6 Hz), 8.12 Phe-NH (d, 1H, J = 8.8 Hz), 7.95 Leu-NH (d, 1H, J = 8.0 Hz), 7.34 Cap-NH (s, 1H), 7.24 Phe-ring (m, 5H), 7.16 Cap-NH (s, 1H), 4.55 Ala-α (m, 1H), 4.39 Phe-α (m, 1H), 4.27 Pro-α (m, 1H), 4.17 Leu-α (m, 1H), 3.63 Pro-δ (m, 2H), 3.12 Phe-β (m, 1H), 2.81 Phe-β (m, 1H), 2.04 Pro-β (m, 1H), 1.94 Pro-γ (m, 1H), 1.88 Pro-γ (m, 1H), 1.83 Ac-cap (s, 3H), 1.75 Pro-β (m, 1H), 1.30 Leu-β (m, 2H), 1.20 Ala-β (d, 3H, J = 7.0 Hz), 0.85 Leu-γ (m, 1H), 0.78 Leu-δ (d, 3H, J = 5.7 Hz), 0.74 Leu-δ (d, 3H, J = 5.6 Hz).

^{1}H NMR (10% D_2O/90% H_2O) δ: 8.63 Phe-NH (d, 1H, J = 8.3 Hz), 8.44 Leu-NH (d, 1H), 8.42 Ala-NH (d, 1H), 7.73 Cap-NH (s, 1H), 7.41 Phe-ring (m, 2H), 7.36 Phe-ring (m, 1H), 7.32 Phe-ring & Cap-NH (m, 3H) 4.68 Phe-α (m, 1H), 4.60 Ala-α (m, 1H), 4.43 Pro-α (m, 1H), 4.25 Leu-α (m, 1H), 3.85 Pro-δ (m, 1H), 3.70 Pro-δ (m, 1H), 3.40 Phe-β (m, 1H), 2.95 Phe-β (m, 1H), 2.30 Pro-β (m, 1H) 2.05, Pro-γ (m, 2H), 2.02 Ac-

cap (s, 3H), 1.90 Pro-β (m, 1H), 1.38 Ala-β (d, 3H), 1.32 Leu-β (m, 2H), 1.25 Leu-γ (m, 1H), 0.82 Leu-δ (d, 3H), 0.79 Leu-δ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 173.0, 171.7, 171.5, 171.1, 168.7, 138.2, 129.1, 127.9, 126.0, 59.8, 54.1, 51.0, 46.8, 46.0, 37.2, 29.0, 24.5, 23.9, 22.8, 22.2, 21.6, 16.9.

HRMS: Calcd. for $[MNa]^+$ $C_{25}H_{37}NaN_5O_5$ [510.2692]; Obsd. [510.2690].

TLC (A) $R_f = 0.60$.

Ac-Val-NapPro-D-Ser-His-NH_2

^1H NMR (10% D_2O/90% H_2O) δ: 8.52 Ser-NH (d, 1H), 8.29 His-ε (s, 1H), 8.08 His-NH & Val-NH (d, 2H), 8.03 Nap-NH (d, 1H), 7.66 Nap-Ring (m, 3H), 7.54 Nap-Ring (s, 1H), 7.42 Cap-NH (s, 1H), 7.30 Nap-Ring (m, 2H), 7.16 Nap-Ring (d, 1H), 7.00 Cap-NH (s, 1H), 6.98 His-δ (s, 1H), 4.41 His-α (m, 1H),4.30 Pro-γ (m, 1H), 4.24 Pro-α (m, 1H), 3.95 Val-α (m, 1H), 3.88 Ser-α (m, 1H), 3.83 Pro-δ (m, 1H), 3.55 Ser-β (m, 1H), 3.50 Nap-β (s, 2H), 3.42 Ser-β (m, 1H), 3.35 Pro-δ (m, 1H), 3.06 His-β (m, 1H), 2.82 His-β (m, 1H), 2.33 Pro-β (m, 1H), 1.73 Acetyl-Cap & Pro-β (m, 4H), 1.57 Val-β (m, 1H), 0.61 Val-γ (d, 3H), 0.55 Val-γ (d, 3H).

^{13}C NMR (DMSO-d_6) δ: 180.0, 171.2, 171.1, 170.4, 169.4, 168.8, 133.2, 133.0, 132.4, 131.2, 127.1, 126.8, 125.5, 125.0, 60.4, 58.3, 55.7, 54.8, 52.3, 51.5, 47.8, 29.1, 25.9, 21.7, 18.2.

Calcd. for $[MH]^+$ $C_{33}H_{43}N_8O_7$ [663.3254]; Obsd. [663.3228]; RP-HPLC (Eluents, A: 0.1% TFA in H_2O; B: 0.08% TFA in CH_3CN: 5 mins - 100% A, 25 mins - Linear Gradient 0-60 % B, 5 mins - 60% B, 2 mins - Linear Gradient, 60-0% B; UV Detection 228, 280 nm) RT: 30.1 mins.

References

(1) Chou, P.Y.; Fasman, G.D. "β-Turns in Proteins," *J. Mol. Biol.* **1977**, *115*, 135-175.

(2) Smith, J.A.; Pease, L.G. "Reverse Turns in Peptides and Proteins," *CRC Crit. Rev. Biochem.* **1980**, *8*, 315-399.

(3) Rose, D.; Gierasch, L.M.; Smith, J.A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.* **1985**, *37*, 1-109.

(4) Gierash, L.M.; Deber, C.M.; Madison, V.; Niu, C.-H.; Blout, E.R. "Conformations of $(X\text{-}L\text{-}Pro\text{-}Y)_2$ Cyclic Hexapeptides. Preferred β-Turn Conformations and Implications for β-Turns in Proteins," *Biochemistry* **1981**, *20*, 4730-4738.

(5) Aubry, A.; Cung, M.T.; Marraud, M. "βI and βII Conformations in Model Dipeptides with the Pro-Xaa Sequences," *J. Am. Chem. Soc.* **1985**, *107*, 7640.

(6) 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.

(7) Stradley, S.J.; Rizo, J.; Bruch, M.D.; Stroup, A.N.; Gierasch, L.M. "Cyclic Peptides as Models for Reverse Turns - Determination of the Equilibrium Distribution Between Type I and Type II Conformations of Pro-Asn and Pro-Ala β-Turns," *Biopolymers* **1990**, *29*, 263-287.

(8) Narasinga Rao, B.N.; Kumar, A.; Balaram, H.; Ravi, A.; Balaram, P. "Nuclear Overhauser Effects and Circular Dichroism as Probes of β-Turn Conformations in Acyclic and Cyclic Pro-X Sequences," *J. Am. Chem. Soc.* **1983**, *105*, 7423-7428.

(9) Kopple, K.D.; Schamper, T.J.; Go, A. "Conformation of Cyclic Peptides VIII. Cyclic Peptides containing the L-Pro-D-Phe Sequence," *J. Am. Chem. Soc.* **1974**, *96*, 2597-2605.

(10) Nishikawa, K.; Momany, F.A.; Scheraga, H.A. *Macromolecules* **1974**, *7*, 797.

(11) Zimmerman, S.S.; Scheraga, H.A. "Influence of Local Interactions on Protein Structure IV. Conformational Energy Studies of N-Acetyl-N'-Methylamides of Ser-X and X-Ser Dipeptides," *Biopolymers* **1978**, *17*, 1885-1890.

(12) Tobias, D.J.; Sneddon, S.F.; Brooks, C.L., III "Reverse Turns in Blocked Dipeptides are Intrinsically Unstable in Water," *J. Mol. Biol.* **1990**, *216*, 783-796.

(13) Lewis, P.N.; Momany, F.A.; Scheraga, H.A. "Folding of Polypeptide Chains in Proteins: A Proposed Mechanism for Folding," *Proc. Natl. Acad. Sci., USA* **1971**, *68*, 2293-2297.

(14) Hilter, W.A.; Walton, A.G. "Energetics of Folding a Lysozyme β-bend," *J. Mol. Biol.* **1975**, *92*, 567-572.

(15) Montelione, G.T.; Arnold, E.; Meinwald, Y.C.; Stimson, E.R.; Denton, J.B.; Huang, J.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.

(16) Schultz, G.E.; Schirmer, R.H. *Principles of Protein Structure*; Cantor, C.R., Ed.; Springer-Verlag: New York, **1978**; pp 74-75.

(17) 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.

(18) Wetlaufer, D.B. "Nucleation in protein folding - confusion of structure and process," *Trends in Biol. Sci.* **1990**, *15*, 414-415.

(19) 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.

(20) Venkatachalam, C.M. "Stereochemical Criteria for Polypeptides and Proteins. V. Conformation of a System of Three Linked Peptide Units," *Biopolymers* **1969**, *6*, 1425-1436.

(21) Levitt, M. "Conformational Preferences of Amino Acids in Globular Proteins," *Biochemistry* **1978**, *17*, 4277-4285.

(22) O'Neil, K.T.; DeGrado, W.F. "Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occurring Amino Acids," *Science* **1990**, *250*, 646-651.

(23) Kim, C.A.; Berg, J.M. "Thermodynamic β-sheet propensities measured using a zinc-finger host peptide," *Nature* **1993**, *362*, 267-270.

(24) Bothner-By, A.A.; Stephens, R.L.; Lee, J.; Warren, C.D.; Jeanloz, R.W. "Structure Determination of a Tetrasaccharide: Transient Nuclear Overhauser Effects in the Rotating Frame," *J. Am. Chem. Soc.* **1984**, *106*, 811-813.

(25) Bax, A. "Correction of Cross-Peak Intensities in 2D Spin-Locked NOE Spectroscopy for Offset and Hartmann-Hahn Effects," *J. Magn. Res.* **1988**, *77*, 134-147.

(26) Bax, A.; Davis, D.G. "Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy," *J. Magn. Res.* **1985**, *63*, 207-234.

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

(28) Woody, R.W. *Peptides, Polypeptides, and Proteins*; Blout, E.R.; Bovey, F.A.; Goodman, M.; Lotan, N., Eds.; John Wiley and Sons: New York, **1974**; pp 338.

(29) Manning, M.C.; Illangasekare, M.; Woody, R.W. "Circular Dichroism Studies of Distorted α-Helices, Twisted β-Sheets, and β-Turns," *Biophys. Chem.* **1988**, *31*, 77-86.

(30) Kawai, M.; Fasman, G.D. "A Model β-Turn. Circular Dichroism and Infrared Spectra of a Tetrapeptide," *J. Am. Chem. Soc.* **1978**, *100,* 3630.

(31) Bush, C.A.; Sarkar, S.K.; Kopple, K.D. "Circular Dichroism of β-Turns in Peptides and Proteins," *Biochemistry* **1978**, *17,* 4951-4954.

(32) Abbadi, A.; Mcharfi, M.; Aubry, A.; Premilat, S.; Broussard, G.; Marraud, M. "Involvement of Side Functions in Protein Structures: The Asx Turn - Occurrence and Conformational Aspects," *J. Am. Chem. Soc.* **1991**, *113,* 2729-2735.

(33) 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.

(34) Bauer, C.J.; Frenkiel, T.A.; Lane, A.N. "A Comparison of the ROESY and NOESY Experiments for Large Molecules, with Application to Nucleic Acids," *J. Magn. Res.* **1990**, *87,* 144-152.

(35) Schimmel, P.R.; Flory, P. "Conformational Energies and Configurational Statistics of Copolypeptides Containing L-Proline," *J. Mol. Biol.* **1968**, *34,* 105-120.

(36) NOE connectivities were observed between the naphthyl ring protons and the Ser-β and Val-γ protons.

(37) Manning, M.C.; Woody, R.W. "Theoretical Study of the Contribution of Aromatic Side Chains to the Circular Dichroism of Basic Bovine Pancreatic Trypsin Inhibitor," *Biochemistry* **1989**, *28,* 8609-8613.

(38) Boussard, G.; Marraud, M. "β-Turns in Model Dipeptides. An Infrared Quantitative Analysis with NMR Correlation," *J. Am. Chem. Soc.* **1985**, *107,* 1825-1828.

(39) Greenfield, N; Fasman, G.D. "Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation," *Biochemistry* **1969**, *8,* 4108-4116.

(40) Hollosi, M.; Kover, K.E.; Holly, S.; Fasman, G.D. "β-Turns in Serine-Containing Linear and Cyclic Models," *Biopolymers* **1987**, *26,* 1527-1553.

(41) Perczel, A.; Hollosi, M.; Fulop, V.; Kalman, A.; Sandor, P.; Fasman, G.D. "Environment Dependent Conformation of Boc-Pro-Ser-$NHCH_3$," *Biopolymers* **1990**, *30*, 763.

(42) Gorbitz, C.H. "Conformational Properties of the Amino Acid Residues L-Cysteine, L-Serine, and L-Cystine," *Acta Chem. Scand.* **1990**, *44*, 584-590.

(43) Ramachandron, G.N.; Sasisekharan, V. "Conformation of Polypeptides and Proteins," *Adv. Prot. Chem.* **1968**, *23*, 283-437.

(44) Benedetti, E.; Morelli, G.; Nemethy, G.; Scheraga, H.A. "Statistical and energetic analysis of side-chain conformations in oligopeptides," *Int. J. Pept. Protein Res.* **1983**, *22*, 1-15.

(45) Atherton, E.; Fox, H.; Harkiss, D.; Sheppard, R.C. "Application of Polyamide Resins to Polypeptide Synthesis: An Improved Synthesis of β-Endorphin using Fluorenylmethoxycarbonyl amino acids," *J. Chem. Soc. Chem. Comm.* **1978**, 539-540.

(46) Carpino, L.A.; Han, G.Y. "The 9-Fluorenylmethoxycarbonyl Amino Protecting Group," *J. Org. Chem.* **1972**, *37*, 3404-3409.

(47) Hudson, D. "Methodological Implications of Simultaneous Solid-Phase Peptide Synthesis. 1. Comparison of Different Coupling Procedures," *J. Org. Chem.* **1988**, *53*, 617-624.

(48) Biosearch Technical Bulletin No. 9000-02.

(49) Reindel, F.; Hoppe, W. "Uber eine Farbemethode zum Anfarben von Aminosauren, Peptiden, und Proteinen auf Papierchromatogrammen und Papierelektropherogrammen," *Chem. Ber.* **1954**, *87*, 1103-1107.

(50) Cavanagh, J.; Keeler, J. "Suppression of HOHAHA and "False" NOE Cross Peaks in CAMELSPIN Spectra," *J. Magn. Res.* **1988**, *80*, 186-194.

Chapter 3: Investigations of Functional Group Incorporation at Position *i* of β-Type II Reverse Turns

Introduction

Reverse turns have been implicated in a number of essential protein functions, from acting as glycosylation or phosphorylation sites for effective post-translational protein processing to providing a template in the construction of enzyme active sites.[1] Reverse turns sequences in native proteins are known to have a strong bias toward polar residues and these amino acids often act to stabilize and direct the overall folding of the turn structure through specific side-chain interactions with the polypeptide backbone. For example, most Type I turns are stabilized by hydrogen bonds between polar groups at position *i* (Ser, Asp, Asn) and the main chain atoms of the residue at position *i+2*. In fact, over 48% of all Type I turns contain one of these residues at position *i*, of which over 57% participate in a long range hydrogen bond with the *i+2* peptide bond.[2] The Type I turn has been shown to have a fairly twisted, non-planar topography due to the relative orientation of the central amide bond, and it has been postulated that non-backbone hydrogen bonding networks are often required to stabilize the overall motif.

In contrast, such interactions are rare for Type II turns in protein structures, as the more planar arrangement of the peptide backbone discourages side-chain - main chain interactions. Therefore, the stability of the Type II turn is not expected to be as dependent on interactions at long range which involve side chain groups. These predictions were borne out by the results of the studies in Chapter 2, which found that peptides could adopt the reverse turn motif in the absence of any side-chain based electrostatic or hydrogen bonding interactions. As a result, the Type II turn is particularly well-suited for protein design, since the addition or removal of functional groups on the turn structure would not be expected to have deleterious effects on the overall structural stability. In addition, the relatively flat topography of the Type II turn provides a structural "platform" where the side-chains at positions *i* and *i+3* are brought into close proximity [$C_\alpha(i)$-$C_\alpha(i+3) \approx 7$Å] with the same relative orientation with respect to the polypeptide backbone (See Figure 3-1). These characteristics provide an excellent template for the design of pre-programmed,

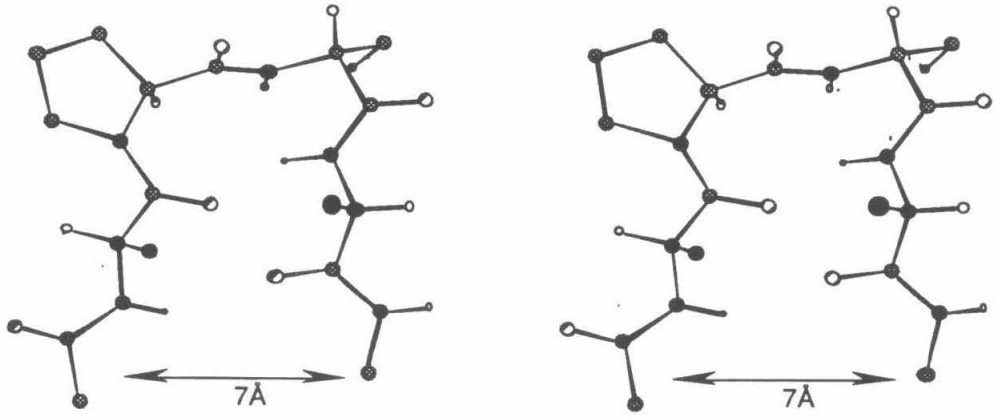


Figure 3-1. Stereodiagram of a Type II reverse turn

functional protein motifs through the selective incorporation of reactive groups at the *i* and *i+3* positions of the reverse turn sequence.

The results of the studies presented in Chapter 2 provided a set of guidelines for selecting primary sequences which adopt stable Type II reverse turn structures in aqueous solution. While the primary focus of that work was to evaluate the effects of *steric* interactions on Type II reverse turn formation, a strong preference was observed for histidine at position *i+3*. This observation is particularly useful for designing functional motifs, as histidine plays a central role in many aspects of native protein function and structure. The imidazole side-chain is a powerful functional group capable of acting as both a hydrogen bond acceptor and donor through the two nitrogen atoms in the heteroaromatic ring. The ability to serve a dual role in hydrogen bonding interactions has been exploited in the construction of "relay" systems, where the imidazole ring serves as a bridging group between two distal hydrogen bonding groups. These *indirect* hydrogen bonding arrangements have been implicated in the catalytic function of a number of enzymes, including members of the serine protease family and more recently, a family of lipases.[3-8] In these enzymes, the histidine is involved in a hydrogen bonding network with

a nucleophilic serine hydroxyl group (or water). This interaction provides a mechanism to further activate the serine hydroxyl group by facilitating the removal of the hydroxyl proton. Histidine has also been implicated in other functional roles of catalysis, as the imidazole ring can act either as a nucleophile or as an efficient base under physiological conditions; in fact, it is one of the strongest bases known at neutral pH.[9] Finally, imidazole can also serve as an excellent ligand for a wide variety of metal cations[10] and histidine is frequently employed as a ligand in metal binding sites of metalloproteins; in fact, histidine is the most common ligand for Zn^{+2} metalloenzymes.[11]

It is envisioned that the diverse functional properties of histidine could be exploited in the design of novel protein motifs based on the Type II reverse turn. In an effort to realize this goal, two functional protein motifs were designed using the structural guidelines presented in Chapter 2 to provide a basis for the incorporation of reactive groups at position *i* capable of interacting with a histidine at position *i+3* . In the first motif, the potential for a hydrogen bonding network between the histidine side chain and hydroxy amino acids incorporated at postion *i* was investigated in an attempt to mimic part of the active site machinery present in native serine proteases. In the second design, the Type II reverse turn was studied as a potential bidentate metal binding site through the incorporation of groups capable of binding metal cations at position *i*.

Serine Protease Mimic

The serine protease family of enzymes has been the subject of intense investigation for over 50 years, and is probably the most well studied class known to date.[12] The key step in the catalytic mechanism of these enzymes is believed to involve a nucleophilic attack by the serine hydroxyl group on the substrate amide carbonyl carbon to produce a tetrahedral intermediate, which then collapses to form an acyl-enzyme intermediate.[12] Subsequent hydrolysis of the acyl-enzyme intermediate releases the cleaved substrate and regenerates the active site (See Figure 3-2). Activation of the serine hydroxyl group is

accomplished through the so-called "catalytic triad," a hydrogen bonding relay system composed of an aspartatic acid, histidine, and the nucleophilic serine residues.[3-8] A large number of model systems have been assembled to mimic the hydrogen bonding network, many of which are based on a cyclic peptide backbone.[13-18] The cyclic peptide backbone provides a rigid constraint on the overall topography of the peptide, and therefore allows the coordination of long range interactions between various functional groups and the imidazole side chain of histidine. Attempts to model the hydrogen bonding network with linear peptides have presented largely ambiguous results,[15,16,18] due to the lack of structural rigidity peptide backbone.

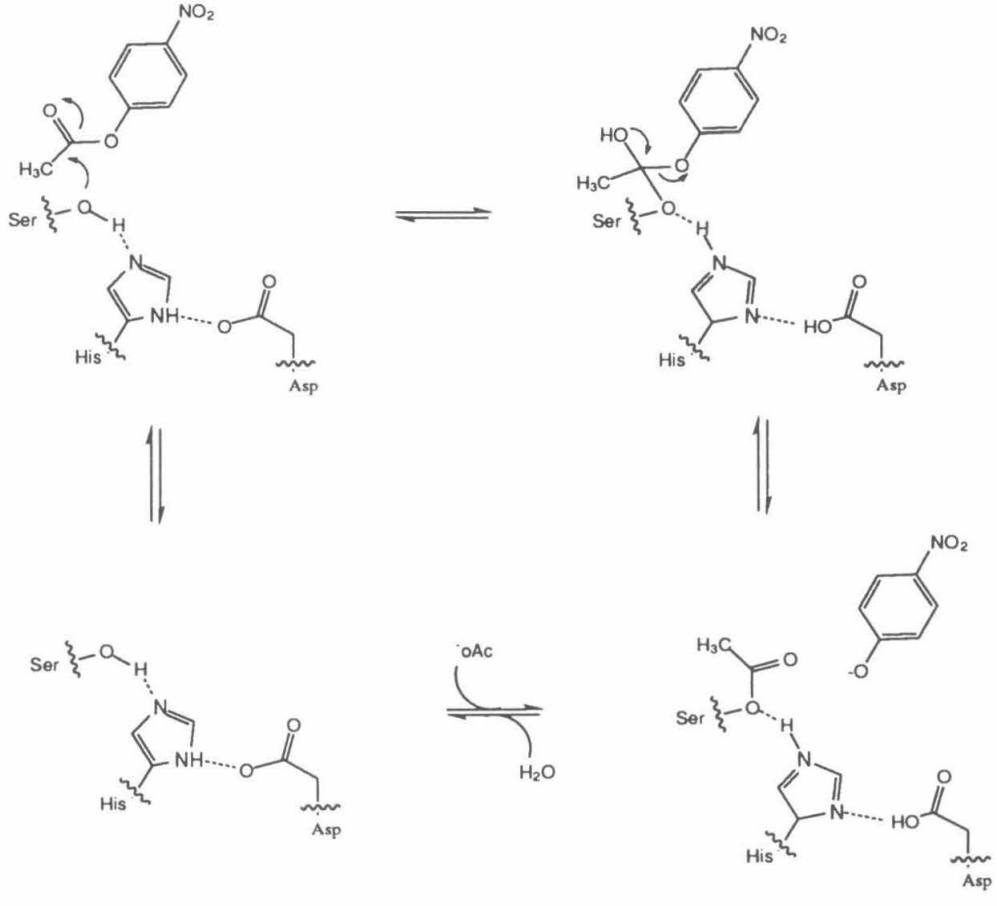


Figure 3-2. Mechanism of esterase hydrolysis proposed for the serine protease family of enzymes

In contrast, the reverse turn motif studied in Chapter 2 was structurally well-characterized and found to be stable in aqueous solution. In order to test whether the reverse turn motif could be used to construct part of the long range hydrogen bonding interactions similar to that found in native serine proteases, a series of peptides was synthesized with the general sequence Ac-Xaa-Pro-D-Ser-His-NH_2, where Xaa = Ser, Thr, (*allo*)Thr. The three peptides were designed to test various aspects of the turn motif; the serine peptide provides a sensitive test of the requirement of β-branched amino acids at the first position of the turn when the potential for long range interactions exists. In the absence of any side-chain interactions, the peptide would not be expected to assume the reverse turn structure. In contrast, the threonine analogs are more conformationally restrained and would be expected to nucleate and stabilize the reverse turn motif based on the conclusions obtained from the earlier studies. The pair of diastereomers differ only in the stereochemistry at the β-center, and therefore provide a means to probe the effect of the orientation of the side-chain hydroxyl group on the overall structure. All three of the peptides were structurally characterized in aqueous solution using nuclear magnetic resonance techniques and circular dichroism spectroscopy as discussed in Chapter 2. These peptides were further analyzed for long range hydrogen bonding interactions through the measurement of the histidine imidazole side chain pK_a values and subsequent comparison to known values. Finally these peptides were tested for esterase activity to evaluate the potential function of the overall designs.

Metal-Binding Site Motif

Metal binding sites centered on reverse turns are well-precedented in native proteins; for example, both superoxide dismutase and carboxypeptidase contain Zn^{+2} metal-binding sites constructed on a Type I β-turn. In both of these proteins, the metal ligands are two histidine residues located at the *i* and *i+3* positions of the reverse turn motif. The metal-binding motifs studied in this chapter were based on this overall design,

with thiolate ligands placed at postion *i*. The choice of a thiolate ligand, rather than an imidazole ligand, was based on several criteria. First, thiolate groups (cysteine) are often used as metal ligands in native protein structures for the construction of both structural and functional metal binding sites. In fact, "structural" Zn^{+2} binding sites found in many proteins (e.g., alcohol dehydrogenase and aspartate transcarbamylase) are constructed almost exclusively from thiolate ligands (4 cysteines).[11] In addition, while non-ligating cysteine residues are often found in regular secondary structures (α-helix, β-sheet), metal-binding cysteine residues are most often located in turn and loop regions of protein structures.[19] Second, based on the guidelines presented in Chapter 2, it is known that turn formation is highly dependent on the steric bulk of the residue placed at position *i*. Beta-branched amino acids (e.g., valine) were found to be essential for efficient turn formation. Incorporation of a neutral thiol group, either with cysteine or the unnatural amino acid (2*R*,3*R*)-*β*-methyl-cysteine (Bmc), was believed to be a fairly conservative substitution of a β-branched amino acid such as valine.

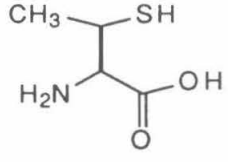


Figure 3-3. (2*R*,3*R*)-β-Methyl-cysteine (**1**)

In order to test whether a metal-binding site could be assembled using the Type II reverse turn motif as a structural template, two peptides were prepared containing thiolate ligands at position *i* and histidine at position *i+3*; Ac-Cys-Pro-D-Ser-His-NH_2 and Ac-Bmc-Pro-D-Ser-His-NH_2. The sequences were chosen based on the guidelines obtained from Chapter 2 using a Pro-D-Ser dipeptide core, as well as acyl and amide capping groups to eliminate electrostatic interactions between the main chain terminii. In both cases, the

structural conformation of each peptide was determined using 2D-NMR and circular dichroism spectroscopy. In addition, the metal-binding characteristics and the peptides were analyzed by circular dichroism spectroscopy for both Zn^{+2} and Co^{+2} cations.

Results and Discussion

Serine-Protease Mimic

Conformational Analysis of Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-Ser-Pro-D-Ser-Phe-NH_2

The observed NOE connectivities for the two peptides which incorporate serine at position *i* indicate that Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-Ser-Pro-D-Ser-Phe-NH_2 have very different turn propensities (See Table 3-1). All of the sequential NOEs (**1**, **2**, **3**) and two of the long range NOEs (**4** and **5**) expected for type II reverse turns are observed for Ac-Ser-Pro-D-Ser-His-NH_2, indicating that this peptide has a significant population of Type II reverse turn conformation in solution. In contrast, the Ac-Ser-Pro-D-Ser-Phe-NH_2 peptide has very little reverse turn character as indicated by the fact that only the sequential NOEs (**1**, **2**, **3**) are observed. These generalizations are borne out by the variable temperature coefficients for the peptides (See Table 3-2); Ac-Ser-Pro-D-Ser-His-NH_2 has a low coefficient for the *i+2* proton (-6.1 ppb/K) and a high coefficient for the *i+3* proton (-8.6 ppb/K), both of which are consistent with Type II turn formation. However, Ac-Ser-Pro-D-Ser-Phe-NH_2 shows almost no difference between the three amide proton coefficients, indicating that the peptide is essentially disordered in solution. Finally, the circular dichroism spectra for the two peptides is consistent with the NMR studies; the CD spectrum for Ac-Ser-Pro-D-Ser-His-NH_2 is a typical Class B spectra (See Figure 3-4), with a maximum ellipticity at 200 nm near 2000 deg·cm^2·dmol^{-1}. The corresponding spectra for Ac-Ser-Pro-D-Ser-Phe-NH_2 does not contain any features characteristic of Type II turns.

Table 3-1. Observed ROESY Connectivities[a] of Peptides Containing Functionalized Residues at Position *i* in Water

Peptide	Observed NOE Connectivities
Ac-Cys-Pro-D-Ser-His-NH_2	**1**,**2**[b], and **3**
Ac-Bmc-Pro-D-Ser-His-NH_2[c]	**1**,**2**,**3**,**4** and **5**
Ac-Ser-Pro-D-Ser-His-NH_2	**1**,**2**,**3**,**4**[b], **5**[b] and **6**[d]
Ac-Ser-Pro-D-Ser-Phe-NH2	**1**,**2**, and **3**
Ac-(*allo*)Thr-Pro-D-Ser-His-NH2	**1**,**2**,**3**, and **4**
Ac-Thr-Pro-D-Ser-His-NH_2	**1**,**2**,**3**, and **4**

[a] See Figure 2-2. for definition of interactions; [b] Crosspeak is relatively weak; [c] Bmc : (*2R*,*3R*)-β-Methyl-cysteine; [d] Connectivites observed between His δ ring proton and Ser(*i*+*2*) α, His δ ring proton and His α proton.

Table 3-2. Temperature Coefficients for Amide Proton Chemical Shifts of Peptides Containing Functionalized Residues at Position *i* in Water ($\Delta\delta/\Delta T$)[a]

Peptide	i	i + 2	i + 3
Ac-Cys-Pro-D-Ser-His-NH_2	-8.4	-9.3	-6.6
Ac-Bmc-Pro-D-Ser-His-NH_2[b]	-7.7	-9.0	-6.1
Ac-Thr-Pro-D-Ser-His-NH_2	-7.8	-8.7	-5.3
Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2	-8.2	-6.5	-6.7
Ac-Ser-Pro-D-Ser-His-NH_2	-8.0	-8.6	-6.1
Ac-Ser-Pro-D-Ser-Phe-NH_2	-7.7	-8.3	-7.6

[a] In parts per billion per degree Kelvin. [b] Bmc : (*2R*,*3R*)-β-Methyl-cysteine

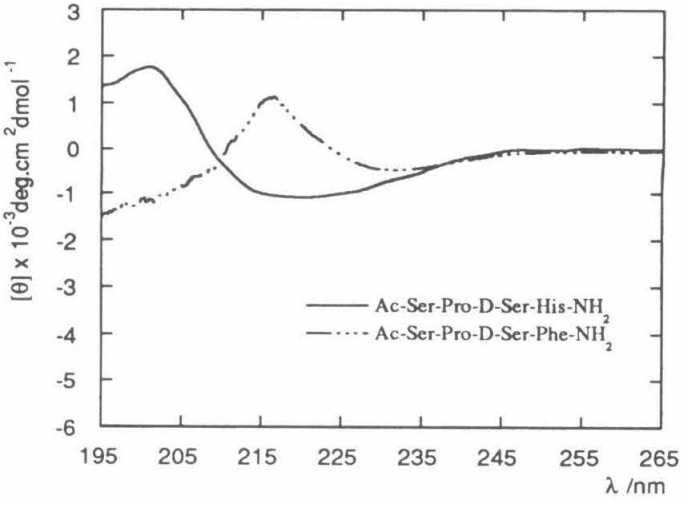


Figure 3-4. Circular dichroism spectra for Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-Ser-Pro-D-Ser-Phe-NH_2 at pH=4.0, 25° C.

Ac-Thr-Pro-D-Ser-His-NH_2 and Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2

The NOE data for Ac-Thr-Pro-D-Ser-His-NH_2 suggests that this peptide has some turn character in solution, as the long range amide-amide NOE, **4**, is observed in addition to strong NOEs for the sequential connectivities (**1**,**2**, and **3**). This is further supported by the variable temperature data, as the *i+3* amide has the lowest coefficient of the three amide protons. Finally, the CD spectra shows a weak but definite amount of Class B character (See Figure 3-5). These results are all consistent with the conclusions derived from the studies in Chapter 2 which found that β-branched amino acids help stabilize turn formation at when placed at position *i*.

The *allo*-threonine peptide has spectroscopic signatures which suggest the presence of a type II β-turn conformation. The observed NOE connectivities are all consistent with

type II turn formation as the long range d_{NN}[Ser-NH, His-NH] (**4**) connectivity is observed in addition to all of the expected sequential NOEs (**1**,**2**,**3**). The variable temperature data show that the *i+3* amide proton is shielded from solvent as expected, but the data also shows a similar effect for the residue at position *i+2*. As discussed in Chapter 2, the amide proton at the *i+2* position would be expected to show a large temperature coefficient in an ideal type II turn; it is possible that the increased solvent shielding of the *i+2* amide proton is due to another turn population, such as a γ-turn.[20] However, the CD spectrum for this peptide, while it does not contain elements of the traditional Class B spectra, is not consistent with the experimentally derived spectrum for a γ-turn (maximum positive ellipticity near 230 nm).[20] Alternatively, the reduced amide coefficient for the *i+2*

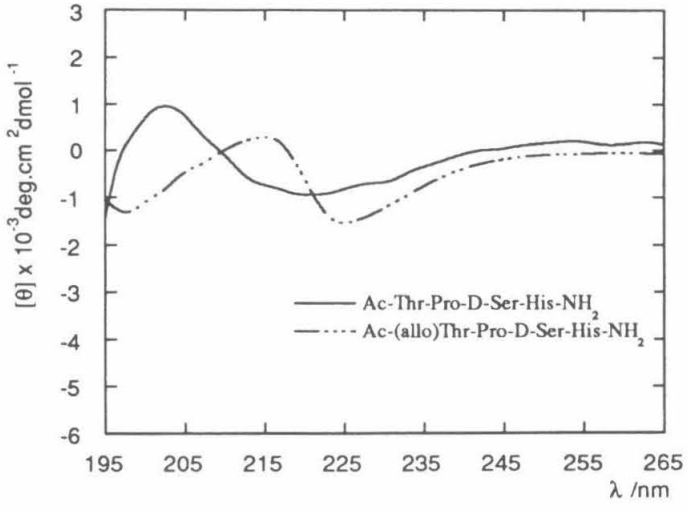


Figure 3-5. Circular dichroism spectra for Ac-Thr-Pro-D-Ser-His-NH_2 and Ac-(*allo*)Thr-Pro-D-Ser-Phe-NH_2 at pH=4.0, 25° C.

position could be attributed to a hydrogen bonding interaction between the *allo*-threonine side-chain and the amide proton. This hypothesis would be consistent with the observed NOE connectivities, as the peptide could still assume a type II reverse turn conformation with the central amide bond slightly rotated to accomodate the hydrogen bonding interaction with the threonine side-chain.

Measurement of Histidine pK_a Values

The fact that Ac-Ser-Pro-D-Ser-His-NH_2 shows a signficant amount of turn character in solution is surprising in light of the earlier studies outlined in Chapter 2, which demonstrated that β-branched amino acids were strongly preferred over non-branched amino acids in the *i* position of the turn sequence. Since Ac-Ser-Pro-D-Ser-Phe-NH_2 demonstrated very little turn character in solution, this suggests that the serine at position *i* and the histidine at position *i+3* act cooperatively to stabilize the reverse turn conformation. It was postulated that this enhanced turn stability could arise from an intramolecular hydrogen bond between the serine hydroxyl side chain and one of the histidine imidazole ring nitrogen atoms. Hydrogen bonding interactions between serine side-chains and a wide number of functional groups are well precedented in protein and peptide structures; in fact it is estimated that over 80% of all serine or threonine residues are involved in some form of side-chain hydrogen bonding interactions in protein structures.[21,22] Similarly, histidine is often involved in hydrogen-bonding relay systems between hydroxyl and carboxylate groups in nature;[5,9] these interactions have been known to influence the relative basicity of the ring system over a wide range of pK_a values.[9]

In order to test whether hydrogen bonding interactions were present in the peptides, incorporating serine, and the threonine diastereomers at position *i*, the histidine imidazole pK_a were determined using ^1H NMR monitored pH titrations.[23] The pK_a value was also determined for Ac-Val-Pro-D-Ser-His-NH_2, since this peptide has been found to have a significant amount of Type II reverse turn character. (See Chapter 2.)

As can be seen in Table 3-3, all of the peptides and the control compound, Ac-His-NHMe, have Hill coefficients near unity so it can be assumed that the histidine ring is the only ionizing species in the titration over the pH range studied (pH= 2.5-10). The pK_a values for Ac-Val-Pro-D-Ser-His-NH_2 and Ac-Thr-Pro-D-Ser-His-NH_2 are identical and are within experimental error (± 0.02 pK_a units) to the pK_a found for Ac-His-NHMe,

Table 3-3. Histidine Imidazole pK_a Values for Selected Tetrapeptides

Peptide	Hill Coefficient (*n*)	pK_a
Ac-His-NHMe	0.95	6.53
Ac-Val-Pro-D-Ser-His-NH_2	1.00	6.52
Ac-Thr-Pro-D-Ser-His-NH_2	1.00	6.52
Ac-Ser-Pro-D-Ser-His-NH_2	0.96	6.65
Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2	1.09	6.70

thereby implying that the histidine ring protons are relatively free in solution. However, Ac-Ser-Pro-D-Ser-His-NH_2 shows an elevated pK_a value, which supports the hypothesis that the histidine ring protons are involved in a hydrogen bonding interaction. In addition, Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2 exhibits an even larger pK_a shift from the control compound (Ac-His-NHMe) in contrast to the results obtained for Ac-Thr-Pro-D-Ser-His-NH_2. The elevated pK_a value for the (*allo*)Thr containing compound could be explained through a hydrogen bonding interaction similar to that proposed for Ac-Ser-Pro-D-Ser-His-NH_2. The observed changes in pK_a are relatively small in magnitude; however, similar shifts have been observed in mutants of bovine pancreatic ribonuclease when specific hydrogen bonding interactions were altered through site-specific mutagenesis.[24] It has been noted that the stereochemistry of β-branched amino acids can drastically affect the orientation of the side-chains in both solution and the solid state.[25,26] In light of this, it is possible that the (2*S*,3*S*) stereochemistry of *allo*-threonine allows the hydroxyl group to

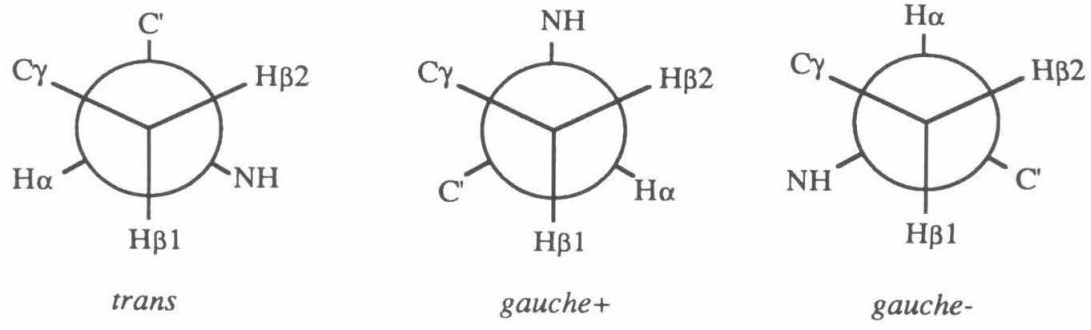
assume the correct orientation to undergo hydrogen bond formation with the histidine ring protons across the turn, whereas the (2*S*,3*R*) stereochemistry of threonine directs the hydroxyl away from the histidine ring protons. (For a stereodiagram of the peptide in this conformation, see Figure 3-7.) This arrangement could also account for the discrepancy between the two threonine-containing peptides in terms of the amide variable temperature coefficients (see above). It is possible that the hydroxyl group of the *allo*-threonine could form an additional hydrogen bond to the *i+2* amide bond, thereby shielding the proton from solvent. Further support for this hypothesis can be obtained through analysis of the χ_1 rotamer distributions of these residues.

Side-chain χ_1 Rotamer Conformer Analysis

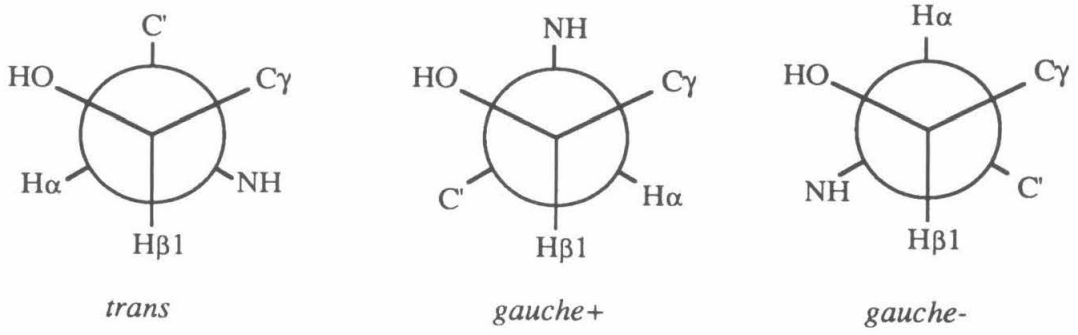
Side-chain χ_1 distributions of amino acids in peptides and proteins have been found to exist almost exclusively in one of three staggered rotamer populations in both the solution[27] and solid states.[21,22] The strong preference of residues to adopt these conformations has been attributed to the reduction of unfavorable steric interactions between the β-substituents and the main atoms in the eclipsed conformations. The three conformations have been classified as *gauche-* (g^-, χ_1= -60°), *gauche+* (g^+, χ_1= 60°), and *trans* (t, χ_1= 180°) based on the relative position of the largest β-substituent and the α-carbon NH group (see Figure 3-6). Information regarding the relative populations of these rotamer conformations in solution can be obtained through analysis of the ^{1}H NMR $^3J_{\alpha\beta}$ vicinal coupling constants. (See Experimental Section.) These calculations were performed for those residues in the tetrapeptides at positions *i* and *i+3* where spectral overlap did not pose a problem in determining the coupling constants. The full results from these calculations are presented in Table 3-4.

The rotamer population calculations indicate that the aromatic amino acids at position *i+3* have a strong preference for the g^- rotamer conformation (>55% of the time on average) in all the peptides studied. This rotamer population places the aromatic side-chain

A. *Unbranched L*-Amino Acid



B. L-Threonine (2*S*, 3*R*)



C. L-(*allo*)Threonine (2*S*, 3*S*)

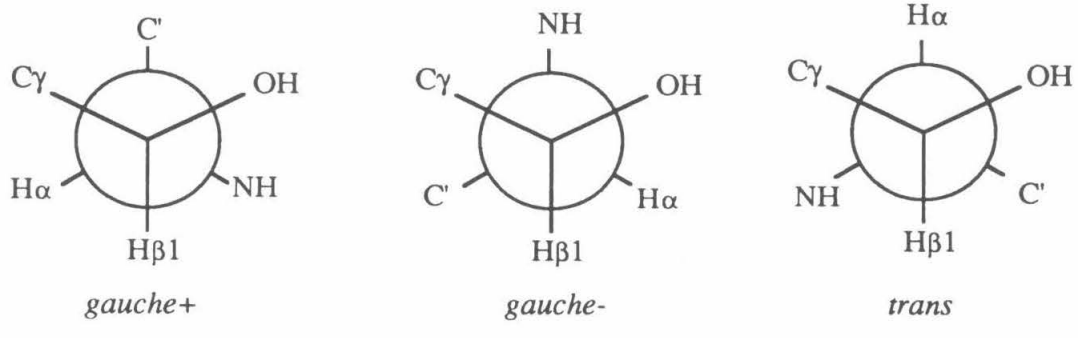


Figure 3-6. Three-state model for χ_1 rotamer conformers of amino acids

over the *i+2* residue, and in support of these calculations, several weak, but significant, long-range NOEs have been observed between the aromatic ring protons (either Phe-ring or the histidine imidazole-Cδ protons) and the *i+2* serine α- and β-protons. The strong preference for the g$^-$ rotamer population is well precedented for aromatic amino acids in peptides, as a survey of oligopeptide crystal structures found that ≈45% of the aromatic amino acids assumed this rotamer conformation.[21]

Unfortunately, detailed studies of the rotamer conformations of the residues at position *i* were not possible, as the coupling constants for the β-protons could not be obtained for many of the residues due to extensive peak overlap in the ^{1}H NMR 1D spectra. However, the coupling constants were determined for the two peptides containing the threonine analogs. While one rotamer population can be determined for threonine (it is not possible to calculate more than one rotamer population for β-branched amino acids based on the $^3J_{\alpha\beta}$ ^{1}H coupling constant alone; see Experimental Section), several interesting conclusions can be drawn from the data about the effects of the threonine residues on the overall turn conformation. The rotamer population calculated for Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2 indicates that the threonine exists primarily in the *trans* conformer (48% of the time on average).[28] This conformation places the methyl group in the g$^-$ conformation and orients the hydroxyl group towards the central amide bond thereby bringing it in close contact with the histidine ring Nε or Nδ (depending on the value of χ_2 for the histidine residue). This conformation would be expected to be energetically favored overall, as neither β-substituent would experience unfavorable *gauche* steric interactions with the main chain atoms. A stereodiagram of the peptide, with *allo*-threonine in the preferred *trans* conformation and the histidine side-chain in the *gauche*$^-$ rotamer, is shown in Figure 3-7. Numerous hydrogen-bonding interactions are possible when the *allo*-threonine side-chain resides in this conformation; the hydroxyl group of the side-chain is within ≈2.6 Å of the histidine ring Nε or Nδ protons, and it has the correct orientation for

Table 3-4. $^3J_{\alpha\beta}$ ^{1}H Vicinal Coupling Constants and Calculated Side-Chain χ_1 Rotamer Populations for Peptides Containing Functionalized Residues at Position *i*

Peptide	Position *i*	Position *i*+3[a]
Ac-Ser-Pro-D-Ser-His-NH_2	ND[b]	$^3J_{\alpha\beta1}$= 9.60 Hz; $^3J_{\alpha\beta2}$= 4.40 Hz f(g$^-$) = 0.59 f(g$^+$) = 0.10 f(t) = 0.31
Ac-Ser-Pro-D-Ser-Phe-NH_2	ND[b]	$^3J_{\alpha\beta1}$= 9.70 Hz; $^3J_{\alpha\beta2}$= 5.56 Hz f(g$^-$) = 0.60 f(g$^+$) = 0.21 f(t) = 0.19
Ac-Thr-Pro-D-Ser-His-NH_2	$^3J_{\alpha\beta}$ = 5.5 Hz f(g$^-$) = 0.27	$^3J_{\alpha\beta1}$= 9.54 Hz; $^3J_{\alpha\beta2}$= 4.40Hz f(g$^-$) = 0.58 f(g$^+$) = 0.08 f(t) = 0.34
Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2	$^3J_{\alpha\beta}$ = 7.7 Hz f(g$^-$) = 0.48	$^3J_{\alpha\beta1}$= 9.84 Hz; $^3J_{\alpha\beta2}$= 4.35 Hz f(g$^-$) = 0.58 f(g$^+$) = 0.08 f(t) = 0.34
Ac-Bmc-Pro-D-Ser-His-NH_2	ND[b]	$^3J_{\alpha\beta1}$= 9.73 Hz; $^3J_{\alpha\beta2}$= 4.72 Hz f(g$^-$) = 0.60 f(g$^+$) = 0.29 f(t) = 0.11
Ac-Cys-Pro-D-Ser-His-NH_2	ND[b]	$^3J_{\alpha\beta1}$= 9.41 Hz; $^3J_{\alpha\beta2}$= 4.77 Hz f(g$^-$) = 0.57 f(g$^+$) = 0.31 f(t) = 0.12

[a] $^3J_{\alpha\beta1}$ and $^3J_{\alpha\beta2}$ refer to the vicinal coupling constants for Hβ1 and Hβ2 respectively as defined in Figure 3-6a; [b] Coupling constants could not be determined due to extensive overlapping peaks in the ^{1}H 1D NMR spectra.

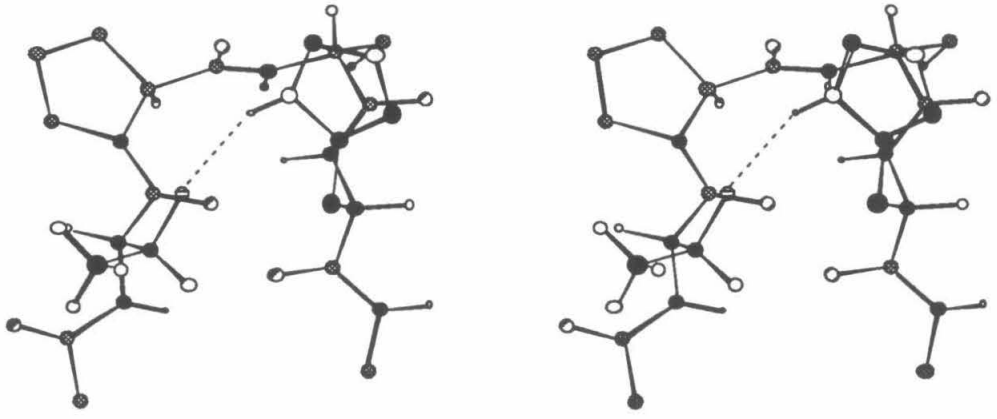


Figure 3-7. Stereodiagram of Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2 with sidechains in the preferred χ_1 rotamer conformations [(*allo*)Thr χ_1= t, His χ_1= g^-; dashed line represents proposed hydrogen bond].

favorable hydrogen bonding interactions. In addition, while the central amide bond proton is too distant (≈4.8 Å) to interact with the side-chain hydroxyl group when the peptide has ideal β-turn dihedral angles, it is possible that these two groups could be involved in a hydrogen bonding interaction if some distortion of the turn dihedrals occurred.[29] These interactions, while not rigorously proven, are consistent with both the anomalous variable temperature coefficient data and the elevated pK_a values for the histidine imidazole protons observed for this peptide. By analogy, it is possible that similar hydrogen bonding interactions are involved in Ac-Ser-Pro-D-Ser-His-NH_2, as this peptide also shows an elevated histidine pK_a. While the rotamer populations for the serine at position *i* could not be calculated in this case, it has been noted that serine is capable of assuming any one of the rotamer populations with equal propensity in protein structures.[22] As a result, it is likely that the serine hydroxyl could reside in the trans conformer (as was found for (*allo*)Thr) and participate in a hydrogen bond with the histidine imidazole.

In contrast to the peptide containing (*allo*)Thr, the preferred rotameric state for threonine in Ac-Thr-Pro-D-Ser-His-NH_2 is more difficult to determine, as the calculated

population of g^- is too low for reasonable analysis (only 27% on average). However, surveys of oligopeptide crystal structures have shown that threonine residues frequently adopt the g+ or g- conformations but not the t rotamer; [28] indeed, no structures have been found with the hydroxyl group in the t conformation.[21] This selectivity has been attributed to unfavorable steric interactions experienced by the threonine-methyl group when the residue assumes the t conformation (hydroxyl group in trans, methyl-group in g^+). In the t rotamer, the methyl group of threonine would experience two gauche steric interactions between the main chain amide groups. (See Figure 3-6b.) Based on the low percentage of the g^- conformer, the predominant rotamer conformation would be expected to be g^+ (g^+ hydroxyl, g^- for methyl). In support of this conformer, a strong NOE is observed between the threonine β-proton and the proline δ-protons. While it is possible that this NOE would be observed in any case due to conformational averaging, only the g^+ rotamer places the β-proton close enough in space to the proline-δ protons to give rise to a significant NOE in the ROESY experiments[30] ($d_{\beta\delta}$[Thr,Pro] in g^+, 2.60 Å; in g^-, 4.23 Å; in t, 4.31 Å).[29] A stereodiagram of the peptide, Ac-Thr-Pro-D-Ser-His-NH_2, is shown in Figure 3-8, with the threonine in the predicted g^+ conformation and the histidine in the preferred g^- conformation. Unlike the (*allo*)-threonine-containing peptide, the hydroxyl group is not in the correct orientation for efficient hydrogen bonding with the histidine side-chain when the residue assumes the predicted conformation. The fact that the peptide did not show an elevated histidine imidazole pK_a is also consistent with the proposed conformation. As a result, while the preferred conformation of the threonine residue cannot be determined rigorously, the g^+ conformation is most consistent with all of the data collected on the peptide.

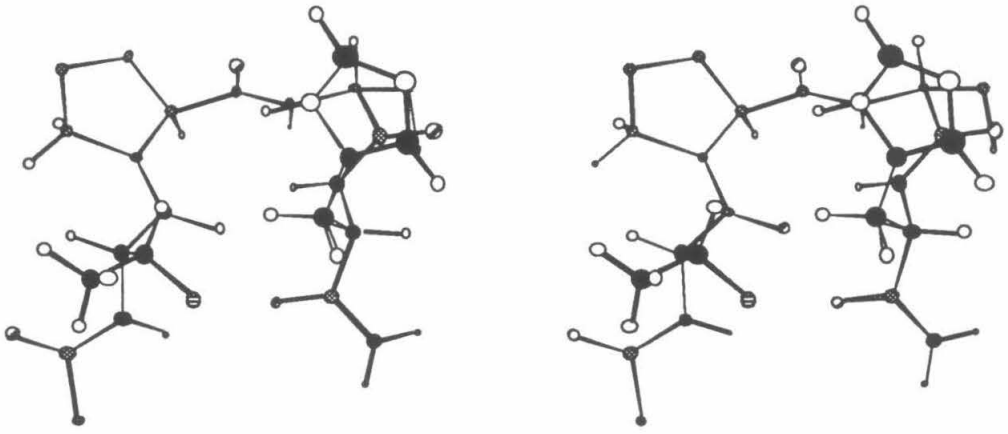


Figure 3-8. Stereodiagram of Ac-Thr-Pro-D-Ser-His-NH_2 with the sidechains in the preferred χ_1 rotamer populations (Thr $\chi 1 = g^+$, His $\chi 1 = g^-$)

Esterase Actvity of Serine and Threonine Containing Peptides

The data for Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2 suggests that a hydrogen bonding relay system is in operation between the Ser-Oγ-His(Im)-NH hydrogen bond. In order to test whether these interactions are sufficient to promote catalytic activity, the rates of esterase hydrolysis were measured for the series of peptides using *p*-nitrophenyl-acetate (*p*-NPA) as the ester substrate.

Rates of *p*-NPA hydrolysis have been used to measure the catalytic efficiencies for a number of model enzyme systems, [14-16,18] primarily because it is relatively simple to assay and the results can be directly correlated to native proteolytic enzymes which show esterase activity (e.g., chymotrypsin).[31] The *p*-NPA esterase activity of the peptides was measured under pseudo-first order conditions, and the observed rate constants (k_1) were then converted to second order rates (k_{cat}) to correct for the background hydrolysis rate and differences in peptide concentrations. (See Experimental section of this chapter for methods of calculation.) These conversions allow a direct comparison of the esterase activity both

Table 3-5. *p*-Nitrophenol-acetate Hydrolysis Rates for Peptides in H_2O at pH=7.60[a]

Peptide	pK_a	α[b]	k_{cat} ($M^{-1} \cdot min^{-1}$)[c]
Imidazole	7.20 (H_2O)	0.71	32.1 ± 0.3
Ac-Ser-Pro-D-Ser-Phe-NH_2	-	-	0
Ac-Thr-Pro-D-Ser-His-NH_2	6.52	0.92	2.92 ± 0.2
Ac-Val-Pro-D-Ser-His-NH_2	6.52	0.92	3.58 ± 0.1
Ac-Ser-Pro-D-Ser-His-NH_2	6.65	0.90	5.42 ± 0.3
Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2	6.70	0.89	4.96 ± 0.1

[a] Rates were measured in 0.07 M NaP_i with 5% v/v 1,4-dioxane; [b]The apparent degree of dissociation; see Experimental Section for calculations; [c]Second order hydrolysis rates were calculated using a background hydrolysis rate of $k_w = 1.1 \times 10^{-2}$ min^{-1}. (See Experimental Section.)

among the peptides within the current study and also to the results of the existing protease models and native enzymes. The results of these studies are presented in Table 3-5.

It is well established that imidazole is capable of efficient *p*-NPA hydrolysis[32] and the second order rate of hydrolysis for imidazole obtained in these studies is identical with the reported value.[33] In general, the peptides show reduced catalytic activities when compared to the imidazole rate. This phenomenon has been noted with derivatives of histidine[16] and other peptide models;[13-15,17] the reduced activity has been attributed to unfavorable steric interactions arising from the substitution of the imidazole ring. Even with the reduced rates, however, several trends emerge from the kinetic data. First, both peptides with the elevated histidine imidazole pK_a values, Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2, show higher rates of esterase activity than the other peptides. This effect cannot be attributed to a higher concentration of neutral imidazole resulting from the elevated pK_a values, since calculation of the second order rate constant (k_{cat}) is based on the effective imidazole concentration determined from the pK_a value for

each peptide (see Experimental Section). Further, the fact that Ac-Thr-Pro-D-Ser-His-NH_2 does not exhibit an enhanced rate demonstrates that the elevated esterase activities present in the Ser(*i*) and (*allo*)Thr(*i*) peptides are not simply due to the presence of a hydroxyl group at position *i*, but are more dependent upon the presence of the hydrogen-bonding interaction between the hydroxyl side-chain at position *i* and the histidine ring imidazole. Similar effects have been noted with highly constrained cyclic peptides, where the extent of hydrogen bonding interactions can be rigidly controlled.[14] In these studies, the peptides which were capable of side-chain interactions demonstrated higher catalytic efficiencies when compared to those where these interactions were disfavored.

There are two mechanisms proposed for the imidazole-catalyzed hydrolysis of *p*-NPA; either general-base catalyzed or nucleophilic attack by the imidazole ring nitrogen atoms. The mechanism proposed for the native proteolytic enzymes essentially follows the general-base mechanism, where the histidine ring nitrogens activate a serine hydroxyl (or water molecule) which can then attack the ester bond. (See Figure 3-2.) In contrast, in the nucleophilic attack mechanism, the imidazole molecule attacks the ester carbonyl directly to form an acyl-imidazole intermediate which then hydrolyzes to regenerate the free imidazole molecule. (See Figure 3-9.) It is now generally accepted that the imidazole catalyzed hydrolysis of *p*-NPA follows the direct nucleophilic attack mechanism.[33]

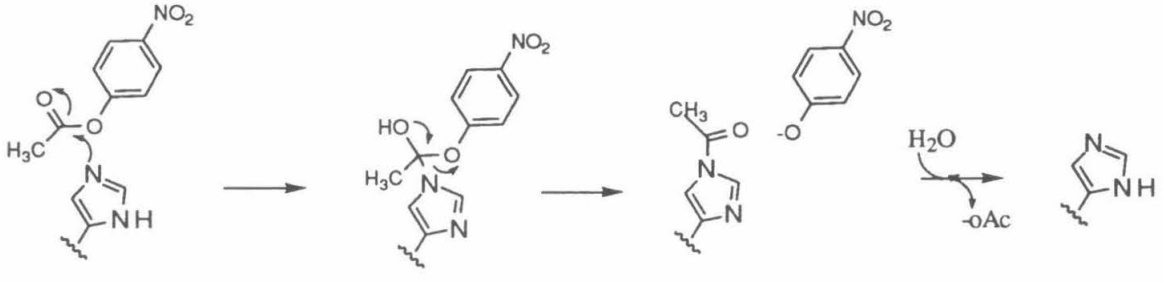


Figure 3-9. Imidazole-catalyzed ester hydrolysis by the nucleophilic attack mechanism

One experiment which can be used to distinguish the mechanistic pathways is through measurement of a solvent isotope effect on the hydrolysis rate when the reactions are run in deuterium oxide rather than in water. A large solvent isotope effect ($k_H/k_D > 2$) would be expected for the general base mechanism, as this pathway requires that a hydrogen bond be broken during the rate limiting step of the nucleophilic attack on the acyl substrate.[34] In contrast, the direct nucleophilic attack mechanism (see Figure 3-9) does not require any hydrogen bonds to be broken in the rate limiting step, and therefore a negligible solvent isotope effect would be expected ($k_H/k_D = 1$).[34] In order to distinguish which mechanism was favored for the peptides in the present study, the rates of hydrolysis were measured for the peptides in 99.9% D_2O and compared to the rates obtained in water. These rates are presented in Table 3-6, along with the pK_a values for the peptides in 99.9% D_2O.

Both peptides, Ac-Ser-Pro-D-Ser-His-NH_2 and Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2, do not show any appreciable differences in the esterase activity between the two solvent systems. The negligible solvent isotope effect direct nucleophilic attack mechanism as

Table 3-6. *p*-Nitrophenol-acetate Hydrolysis Rates for Peptides in D_2O at pD=7.60[a]

Peptide	pK_a	k_{cat} ($M^{-1} \cdot min^{-1}$)[b]	k_H/k_D
Imidazole	7.20 (H_2O)	32.1 ± 0.3	≈ 1.0
	7.63 (D_2O)	33.5 ± 2.3	
Ac-Ser-Pro-D-Ser-His-NH_2	6.65 (H_2O)	5.42 ± 0.3	1.11
	7.17 (D_2O)	5.11 ± 0.4	
Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2	6.70 (H_2O)	4.96 ± 0.1	1.22
	7.25 (D_2O)	4.05 ± 0.1	

[a] pD = pH + 0.4; 0.07 M NaPi, 5% v/v 1,4-dioxane; [b] Second order hydrolysis rates for D_2O were calculated using a background hydrolysis rate of $k_w = 3.2 \times 10^{-3}$ min^{-1}

observed for free imidazole. The enhanced rates of esterase activity for these peptides is likely to be due to efficient acyl transfer of the acetate group from the acylated imidazole intermediate to the Oγ of the Ser(*i*) and (*allo*)Thr(*i*) residues. These processes have been well-documented in other peptide models containing both hydroxyl and thiol groups.[14,35] The results from these mechanistic studies have implications for the design of future models targeted to emulate native proteolytic enzymes, as the observed nucleophilic mechanism is very different from the reaction pathway proposed for the native enzymes. The data suggests that the Ser-His motif, while contributing to enhanced rates of catalysis, is not sufficient to mimic the full enzyme activity (chymotrypsin, k_{cat}=10,000 $M^{-1} \cdot min^{-1}$).[16] It is apparent that at least the third component, an aspartatic acid residue, is required for an accurate model of the proteolytic enzymes. The full Asp-His-Ser triad would be expected to disfavor the nucleophilic attack mechanism, as both nitrogens of the central histidine ring would be occupied in hydrogen bonding interactions and therefore unable to play a nucleophilic role. Future protease mimics based on the reverse turn motif could incorporate a carboxylate group in the turn sequence, possibly through the use of a modified proline such as that described in Appendix I, or through a modified D-residue at position *i+2*, to provide the third component of the catalytic triad.

Metal-Binding Motif

Synthesis and Peptide Incorporation of N^{α}-(9-Fluorenylmethoxycarbonyl)-*S*-tritylmethyl-β-methyl-cysteine (**8**)

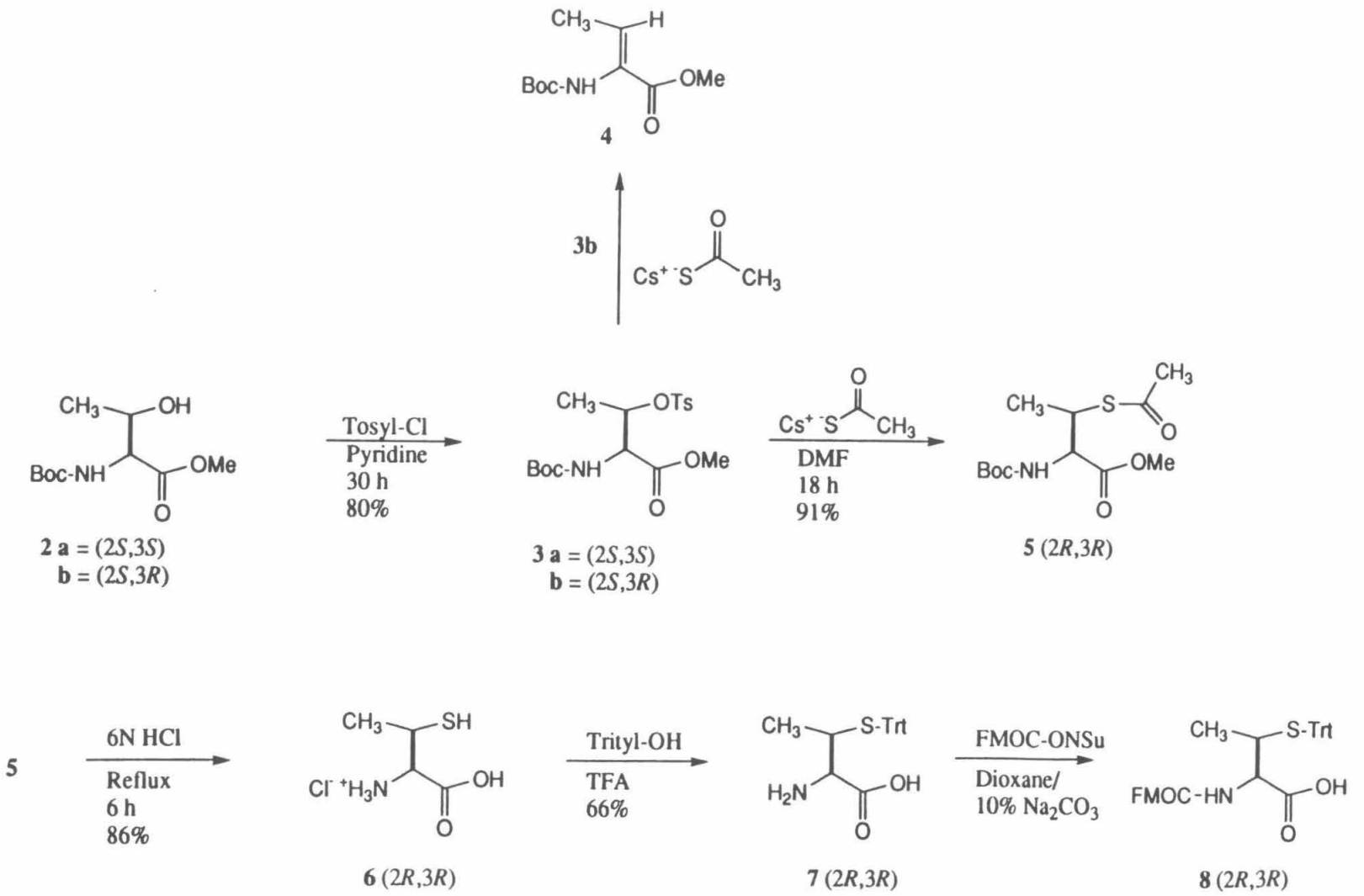
Synthesis of (2*R*,3*R*)-N^{α}-Fmoc-*S*-trityl-β-methyl-cysteine was accomplished along a stereoselective synthetic route (Scheme 3-1) following the methods described by Morell, Fleckenstein, and Gross.[36] The key step of the synthetic route involves a nucleophilic attack of a protected thiolate (thioacetate anion) on a suitably-protected activated threonine analog, **3**. Interestingly, the nucleophilic substitution was found to be extremely sensitive to the configuration of the stereocenter on the β-carbon of the threonine analog **3**. When the naturally-occurring threonine diastereomer (2*S*,3*R*) **3b** was employed in the synthesis,

the reaction proceeded to yield the corresponding elimination product **4** as the predominant product; this was found to be the case over a wide range of solvents (DMF, THF, DMSO) and temperatures (-78° C to 40° C). However, when the corresponding protected (*allo*)threonine-tosylate **3a** was used in the substitution reaction, a quantitative yield of the substitution product **5** was obtained with complete inversion as determined by inspection of the 1D NMR spectra. Evidently the combination of the bulky protecting groups on the amino (BOC) and carboxyl terminii (methyl ester), in conjunction with the large tosyl group on the γ-hydroxyl group, force the threonine analog **3b** to assume a rotamer conformation that promotes α,β elimination rather than substitution. Extensive hydrolysis of the protected β-methyl cysteine analog **5** in 6N HCl, followed by thiol protection with triphenylmethanol in trifluoracetic acid as described by Photaki et al.,[37] yielded the *S*-trityl protected β-methyl cysteine, which was subsequently protected with Fmoc-ONSu to provide the target β-methyl-cysteine analog **8** ready for solid phase peptide synthesis.

Incorporation of **1** into a tetrapeptide, Ac-Bmc-Pro-D-Ser-His-NH_2, was accomplished using standard solid phase peptide synthesis protocols. The resulting peptide was found to be fairly clean, with only one other peptidyl impurity detected by RP-HPLC.

Conformational Analysis of Ac-Cys-Pro-D-Ser-His-NH_2 and Ac-Bmc-Pro-D-Ser-His-NH_2

The spectroscopic data collected on the β-methyl cysteine peptide, Ac-Bmc-Pro-D-Ser-His-NH_2, shows that it adopts a significant population of reverse turn character in solution. As can be seen in Table 3-1, all of the sequential NOEs (**1**, **2**, and **3**; see Figure 2-2 for assignment) and the two long range interactions that are diagnostic of Type II reverse turn formation, **4** d_{NN}(Ser-NH, His-NH) and **5** $d_{\alpha N}$(Pro-α, His-NH), are observed. The NOE connectivity **5** is particularly informative, as this long range interaction is only observed in those cases where the peptide conformation is highly constrained. As mentioned in Chapter 2, this connectivity is at the sensitivity limits of the



Scheme 3-1. Synthetic Route Employed for (2*R*,3*R*)-N$^{\alpha}$-FMOC-S-trityl-β-methyl cysteine

ROESY experiment (<3.5Å) and would be expected to have a weak intensity even if the peptide assumed an ideal reverse turn conformation. Since the NOE connectivity observed for the β-methyl-cysteine peptide has a moderate intensity, it is possible that the peptide assumes a slightly distorted turn conformation, as result of the steric bulk of the thiol and methyl groups of the β-methyl-cysteine residue. This possibility is also supported by the circular dichroism data (see below). The results from the variable temperature studies (see Table 3-2) on the β-methyl-cysteine peptide further support the NOE observations as low coefficients are observed for the *i+3* amide (-6.6 ppb/K), while high coefficients are observed for the *i+2* amide (-9.3 ppb/K). Interestingly, the circular dichroism spectrum (see Figure 3-10) of the β-methyl-cysteine peptide has a positive ellipticities at ≈200 nm

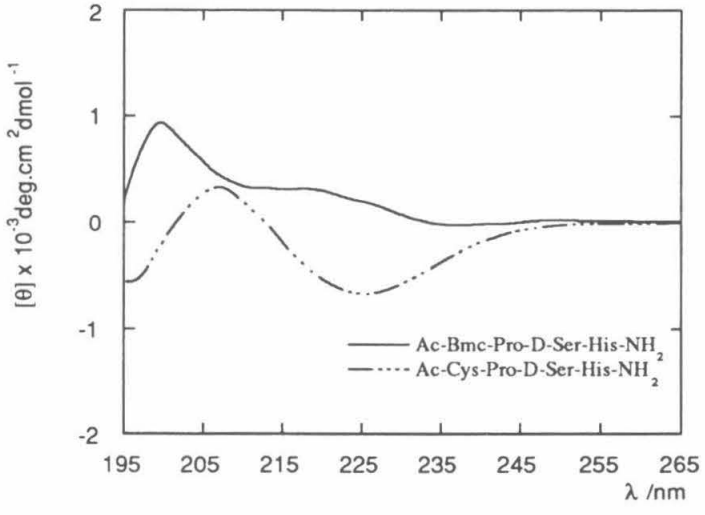


Figure 3-10. Circular dichroism spectra for Ac-Cys-Pro-D-Ser-His-NH_2 and Ac-Bmc-Pro-D-Ser-His-NH_2 at pH=4.0, 25°C.

and ≈220 nm, along with a negative ellipticity below 190 nm, unlike the ideal type II turn which is characterized by a Class B spectrum (see Chapter 2). This unexpected CD signature could be due to a distorted reverse turn conformation, as slight distortions of peptide conformations have been predicted to have dramatic effects on the overall CD spectrum.[38]

In contrast, the spectroscopic data collected for the cysteine-containing peptide indicate that the peptide has a reduced population of reverse turn conformation in solution. Results from the ROESY experiments indicate that the peptide is not constrained in the reverse turn conformation, as only the sequential NOE connectivities (**1**, **2**, and **3**; see Table 3-1) are observed in the ROESY spectrum. However, results from the variable temperature NMR experiments (see Table 3-2) show a low coefficient for the *i+3* amide and a correspondingly high *i+2* amide coefficient, which is suggestive of reverse turn formation in solution. Similarly, the CD spectrum of the cysteine-containing peptide exhibits weak Class B character characteristic of a reduced, but observable, population of the type II reverse turn.

In summary, the spectroscopic data collected for the two thiol-containing peptides, Ac-Cys-Pro-D-Ser-His-NH_2 and Ac-Bmc-Pro-D-Ser-His-NH_2, is consistent with the general conclusions from the earlier peptide studies discussed in Chapter 2, which indicated that sterically hindered amino acids at the *i* position can serve to constrain the peptide conformation significantly.

Metal-Binding Titrations

The metal binding properties of the two thiol-containing peptides were evaluated with metal cation titrations monitored with circular dichroism spectroscopy. The two peptides were titrated with $ZnCl_2$ and $CoCl_2$, and the spectra from typical titrations are shown in Figure 3-11 and Figure 3-12. For both peptides dramatic changes in the spectra are observed upon addition of metal cations, thus indicating that the peptides bind metal

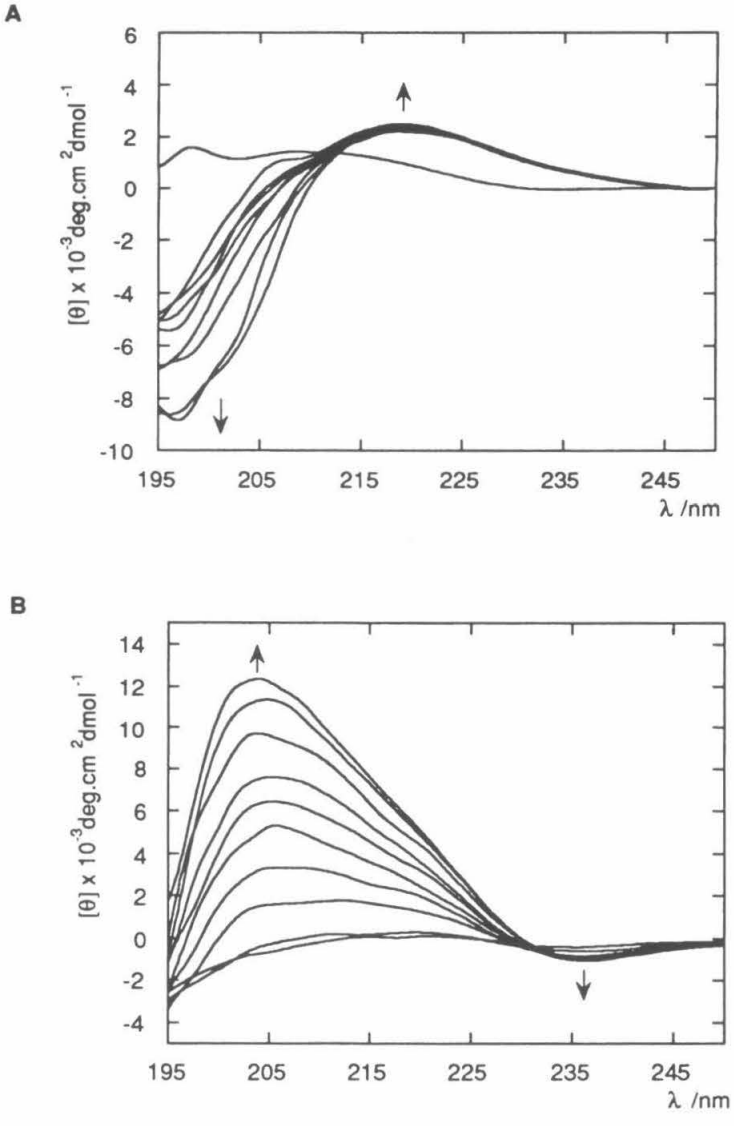


Figure 3-11. (A) Circular dichroism spectra of the titration of Ac-Cys-Pro-D-Ser-His-NH_2 (50 μM) with $ZnCl_2$ (pH=7.4, 100 mM PIPES, 25° C; $ZnCl_2$ added in 20 μM aliquots). (B) Circular dichroism spectra of the titration of Ac-Bmc-Pro-D-Ser-His-NH_2 (50 μM) with $ZnCl_2$ (pH=7.4, 100 mM PIPES, 25° C; $ZnCl_2$ added in 40 μM aliquots).

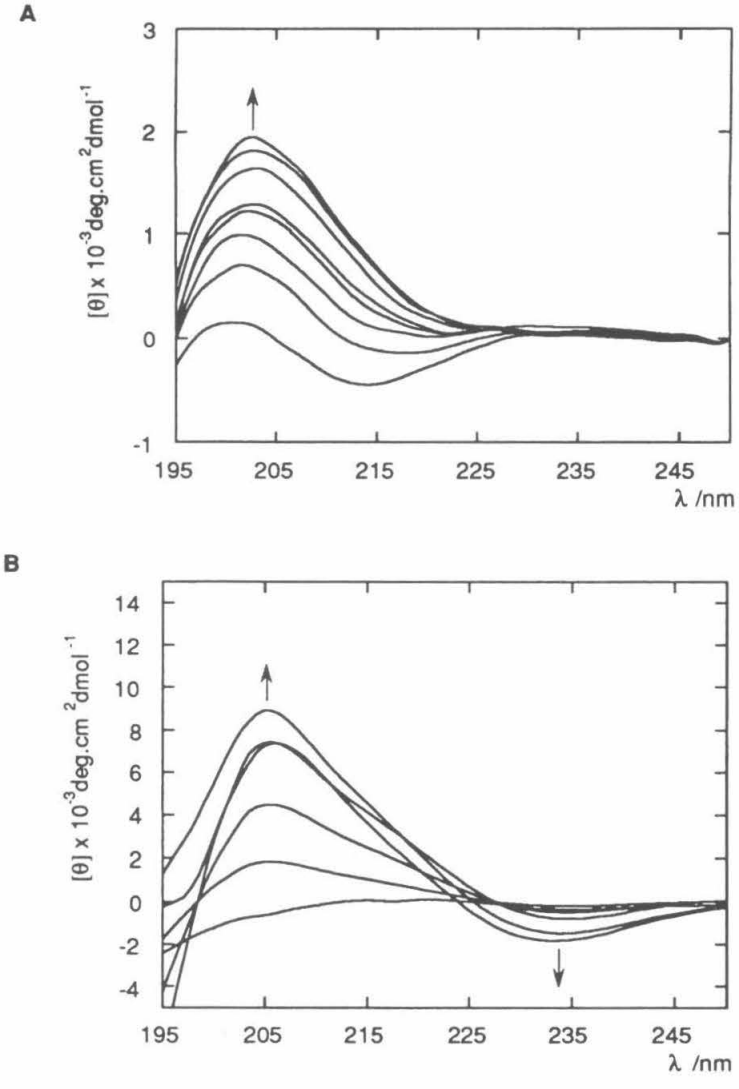


Figure 3-12. (A) Circular dichroism spectra of the titration of Ac-Cys-Pro-D-Ser-His-NH_2 (50 μM) with $CoCl_2$ (pH=7.4, 100 mM PIPES, 25° C; $CoCl_2$ added in 40 μM aliquots). (B) Circular dichroism spectra of the titration of Ac-Bmc-Pro-D-Ser-His-NH_2 (25 μM) with $CoCl_2$ (pH=7.4, 100 mM PIPES, 25° C; $CoCl_2$ added in 10 μM aliquots).

cations and that a conformational change occurs upon metal coordination. Unfortunately, while the spectral changes follow a consistent pattern upon increasing metal cation concentration for both peptides, well-defined isodichroic points were not observed in any of the titrations. The lack of isodichroic points suggests that the peptides undergo multiple equilibrium processes upon addition of metal cations.[39] Attempts were made to discourage intermolecular complex formation by variation of the peptide concentration; however, these effects were observed over the full range of concentrations employed (5 μM - 100 μM). Since the peptides contain two metal ligating groups (thiol, imidazole), the peptide could potentially form a wide range of intermolecular complexes. The number and identity of the various metal complexes could not be ascertained directly from the spectra and, as a result, it was not possible to determine the dissociation constants for the monomeric peptide:metal complexes.

Conclusions

The type II reverse turn provides an ideal template for *de novo* design of functional protein constructs. As discussed in Chapter 2, the propensity for linear tetrapeptides to form stable reverse turn conformations is highly dependent upon the primary sequence, and the overall conformational stability arises from a series of cooperative interactions between neighboring residues. Since the studies in the previous chapter focused primarily upon the effects of steric interactions on turn formation, incorporation of functional groups at position *i* capable of long range intramolecular interactions could have dramatic effects on the overall peptide conformation. The results described in this chapter demonstrate that tetrapeptides incorporating either hydroxyl or thiol groups at position *i* in the consensus sequence Ac-Xaa-Pro-D-Ser-His-NH_2 can adopt stable reverse turn conformations in aqueous solution. While the studies in Chapter 2 indicated that steric bulk at position *i* was essential for effective reverse turn formation, the results from the current studies suggest that several additional long range interactions may contribute to the overall turn stability

when hydroxyl-containing amino acids are incorporated at position *i*. For example, the peptide containing serine at position *i* exhibited a significant amount of reverse turn character in solution, as detected by NMR and CD spectroscopy, despite the fact that this residue has limited steric bulk. Further, pH titrations on the peptide revealed that the imidazole side chain of the histidine residue has an elevated pK_a value. The elevated pK_a value was attributed to a long range hydrogen bonding interaction between the hydroxyl side-chain and the imidazole ring. Similar observations were noted with a peptide containing (2*S*,3*S*)-L-threonine, but not with the peptide containing the (2*S*,3*R*)-L-threonine diastereomer. The contrast in reactivities of the two threonine diastereomers was ascribed to differences in the χ_1 rotamer preferences; in the (2*S*,3*S*)-L-threonine-containing peptide, the hydroxyl group is preferentially directed towards the imidazole ring, whereas in the (2*S*,3*R*)-L-threonine diastereomer the hydroxyl group is directed towards the solvent milieu.

Since similar serine-histidine hydrogen bond interactions are an integral component of the active site catalytic machinery of the serine protease family, the peptides were tested for esterase activity. While the peptides displayed reduced reactivities relative to imidazole, enhanced reactivities were observed for the serine and (2*S*,3*S*)-L-threonine containing peptides relative to the peptide containing the (2*S*,3*R*)-L-threonine residue. The mechanism of the observed catalytic activities, determined by deuterium isotope effects, was found to be the result of an imidazole-based nucleophilic attack mechanism. In contrast, the native enzymes carry out esterase activity through a general base mechanism. These findings suggest that the peptide model systems, while they contain elements of native protease active site machinery, do not contain the full complement of the native enzyme active site topographies. It is possible that through additional functional group incorporation that more accurate representations of the protease mimics could be achieved with these model systems.

The results from peptides incorporating thiol-containing amino acids at position *i* were found to be consistent with the guidelines derived in Chapter 2 for reverse turn formation. The peptide containing a β-branched, thiol-containing amino acid, (2*R*, 3*R*)-β-methyl-cysteine, constrained the peptide conformation to adopt a conformation consistent with a distorted type II reverse-turn character in solution. In contrast, the corresponding peptide containing a cysteine residue exhibited a small amount of reverse turn population in solution as observed by CD and NMR spectroscopy. It was believed that a thiol group at position *i* and a histidine residue at position *i+3* would assemble to form a bidentate metal binding site when the compound adopted a reverse turn conformation. Metal cation titrations performed on the peptides using circular dichrosim spectroscopy indicated that both peptides were capable of effective metal binding activity. However, a critical analysis of the metal binding properties of the two peptides was complicated by the formation of multiple metal cation:peptide complexes. It is envisioned that a more quantitative analysis of these motifs could be accomplished through studies of larger peptide systems which are designed favor the formation of stable intramolecular, monomeric metal complexes. An example of such a system is described in Chapter 7, where a peptide containing the Bmc-Pro-D-Ser-His sequence was incorporated in a dodecapeptide containing 4-bipyridyl-alanine (4Bpa, see Chapter 4). This peptide formed well-defined monomeric complexes for all of the metal cations studied (See Chapter 7).

Experimental Section

Preparation of N^{α}-(9-Fluorenylmethoxycarbonyl)-S-tritylmethyl-β-methyl-cysteine

Preparation of (2*R*,3*R*)-β-Methyl-cysteine (**1**, See Figure 3-3) was carried out using the method of Gross[36] using a protected derivative of L-*allo*-threonine. The (*allo*)-threonine was prepared from *N*-Benzoyl-L-threonine (Aldrich Chemical Co.) as described by Elliot.[40] Tritylation of the cysteine analog was performed as reported by Photaki, *et al.*[37]

N^{α}-(9-Fluorenylmethoxycarbonyl)-*S*-tritylmethyl-β-methyl-cysteine (**1**) A solution of (2*R*,3*R*)*S*-tritylmethyl-β-methyl cysteine (0.9939 g, 2.52 mmol) in 24 mL 1:1 (v:v) dioxane:10% Na_2CO_3 was stirred at room temperature and a solution of FMOC-ONSu (0.8568 g, 2.54 mmol) in 10 mL dioxane was added dropwise. After 1.5 h, the mixture was extracted with ethyl ether (3 x 50 mL). The ether fractions were dried over Na_2SO_4 and concentrated *in vacuo* to ~5 mL. The ether solution was transferred to a centrifuge tube and triturated with hexanes (3 x 30 mL). The resulting precipitate was dissolved in 20 mL ethyl acetate and concentrated *in vacuo* to yield 1.3669 g (90%) of a pale yellow foam. mp = 85-87° C; TLC (SiO_2: 5:1 (v:v) $CHCl_3$:MeOH; UV detection) R_f : 0.70; HRMS: Calc. for $[MH]^+$ $C_{38}H_{32}NO_4S$ [598.2052]; Obs. [598.2079]; ^1H NMR ($CDCl_3$) δ: 7.90 (d, 2H, J=6.8Hz), 7.3-7.8 (m, 22H), 5.50 (d, 1H, J=8.5 Hz), 4.55 (d, 2H, J=7.0 Hz), 4.45 (m, 1H), 4.38 (m, 1H), 3.01 (m, 1H), 1.12 (d, 3H, J=6.8 Hz); ^{13}C NMR ($CDCl_3$) δ: 146.3, 130.5, 129.0, 128.9, 128.7, 128.6, 128.0, 127.2, 126.3, 121.0, 66.4, 61.7, 47.7, 43.1, 17.5; IR (thin film) cm^{-1} : 3395, 3056, 1708, 1610, 1500, 1446, 1400, 1328, 1251, 1077, 1036, 908, 733, 697.

Peptide Synthesis

All peptides were prepared using standard solid phase peptide synthesis protocols as described in Chapter 2. Crude peptides were purified using reversed-phase high pressure liquid chromatography (RP-HPLC) on a C18 semi-preparative column (1.0 cm

x 25 cm Beckman Instruments) using water/acetonitrile gradients. Purified peptides were lyophilized and stored at -70° C. Free thiol concentrations of the sulfur-containing peptides were measured spectrophotometrically after modification with bis(p-nitrophenyl)-disulfide (DTNB) as described by Ellman.[41]

HPLC Methods:

A : Flow rate 1 ml/min; Analytical C18 column, detection @ 228, 254 nm. Gradient: 5 mins. H_2O (0.1% TFA), Linear gradient 0-25% CH_3CN (0.08 % TFA) over 25 mins, 5 mins 25% CH_3CN (0.08% TFA), 25-0% CH_3CN (0.08% TFA) over 2 mins.

B : Flow rate 1 ml/min; Analytical C18 column, detection @ 228, 254 nm. Gradient: 5 mins. H_2O (0.1% TFA), Linear gradient: 0-40% CH_3CN (0.08 % TFA) over 25 mins, 5 mins 25% CH_3CN (0.08% TFA), 40-0% CH_3CN (0.08% TFA) over 5 mins.

Spectroscopic Studies

^1H and ^{13}C NMR spectra for synthetic intermediates were collected at 300 MHz and 75.5 MHz respectively on a General Electric QE-300 NMR Spectrometer. ^1H and ^{13}C NMR data for the petpides were collected at 500.14 MHz and 125.7 MHz respectively on a Bruker AMX500 NMR Spectrometer. All 2D-NMR spectra (ROESY, TOCSY) and variable temperature experiments were recorded as described in Chapter 2. In cases where spectral overlap was a concern, the temperature was elevated or lowered by 5°C to avoid difficulties in spectral analysis. Circular dichroism spectra were recorded at ambient temperature (25-28° C) at 25-50 uM in a 1.0 cm quartz cell. All solutions were prepared in fully degassed (N_2 sparged) Millipore water and the pH was adjusted using small amounts (1-10uL) of 1.0 N NaOH or 1.0 N HCl solutions.

Metal-Binding Titrations of Thiol-Containing Peptides

Circular dichroism metal titrations were performed in 1 cm path-length quartz cuvettes with peptide concentrations ranging from 25-50 μM at pH=7.4 in 100 mM PIPES (25° C). Spectra were collected after each addition of metal over the wavelength range 195-250 nm using a scan speed of 50 nm/min, a time constant of 0.5 s, and a bandwith of 1 nm. A minimum of 8 scans were taken for each spectra. The spectra are reported in ellipticity/residue. Metal cation solutions (Zn^{+2}, Co^{+2}) were prepared from the corresponding anhydrous chloride salt, and the stock solutions were standardized by ethylene-diamine tetraacetic acid complexometric titrations (0.5 M standardized EDTA solution, Aldrich) using murexide as an indicator.[42] Peptide concentrations were determined using the Ellman's test for free thiol groups.[41]

Calculation of χ_1 Rotamer Populations from $^3J_{\alpha\beta}$ ^{1}H NMR Coupling Constants

Calculation of the χ_1 rotamer populations for L-amino acids with two β-protons Hβ1 and Hβ2 in peptide and protein structures can be accomplished using the following equations:

$$f(g^-) = (J_{H\alpha\text{-}H\beta 1} - J_G) / (J_T - J_G) \qquad (3\text{-}1a)$$

$$f(t) = (J_{H\alpha\text{-}H\beta 2} - J_G) / (J_T - J_G) \qquad (3\text{-}1b)$$

$$f(g^+) = 1 - f(t) - f(g^-) \qquad (3\text{-}1c)$$

The standard values for J_G and J_T have been found to be dependent upon the side-chain substituent; Pachler[43] has derived the values of J_T = 13.56 Hz and J_G = 2.60 Hz for non-aromatic amino acids and Cung and Marraud[44] have found that J_T = 13.85 Hz and J_G = 3.55 Hz provide accurate estimations of side-chain rotamer populations for aromatic amino acids (His, Phe, Trp, etc.).

In order to determine all three rotamer populations, the β-methylene protons must be diastereotopically assigned. This can be done most easily through inspection of the relative intraresidue NOE intensities between the α- and β-protons, and the amide and β-protons. For example, if the residue resides in the *trans* conformer, strong NOEs would be expected between the α-proton and the Hβ1 proton, whereas a weaker NOE would be expected for the Hβ2 proton since this proton is *anti* to the α-proton (see Figure 3-6a for clarification). In addition, equal intensity NOE connectivities would be expected for both Hβ1 and Hβ2 protons to the amide proton. Similar arguments can be used assign the stereochemistry of the β-methylene protons for the two gauche conformations. Once the protons are diasterotopically assigned and the corresponding vicinal coupling constants are determined, the rotamer populations of the residue can be calculated using equations 3-1a-c.

In cases where the amino acid has only one β-proton (e.g., threonine), only one population can be assessed, as the other two rotamer populations are indistinguisable in terms of the coupling constant. For example, for L-threonine (2*S*,3*R*) only the f(g^-) population can be calculated, as the f(g^+) and f(t) are indistinguishable (see Figure 3-6b). Conversely, when the β-center is in the *S*-configuration (e.g., *allo*-threonine), only the f(t) population can be calculated (Figure 3-6c). The calculation of the rotamer population follows a similar equation to those above (Equations 3-1a-c) for the unsubstituted residues, and is shown below (for a residue with *R*-configuration at the β-center):

$$f(g^-) = (J_{H\alpha\text{-}H\beta} - J_G) / (J_T - J_G) = 1 - f(t) - f(g^+) \qquad (3\text{-}2)$$

The fractional distributions of rotamer populations were calculated using the equations derived by Pachler[43] (for aromatic residues) or those derived by Cung and Marraud[44] (for non-aromatic residues) using the observed $^3J\alpha\beta$ vicinal coupling constants. The side-chain rotamer populations were assumed to follow a three-state model where low-

energy populations were associated with $\chi 1$ in *gauche*$^-$, *gauche*$^+$, and *trans* conformations. For non-β-branched amino acids, diastereotopic assignment of the β-protons was accomplished through evaluation of intra-residue NOE's between the β-protons and the amide and α-protons.

The vicinal coupling constants ($^3J_{\alpha\beta}$) for selected residues in the tetrapeptide sequences were obtained by inspection of the 1D NMR spectra or through ^{1}H homonuclear decoupling experiments. Diastereotopic assignment of the β-protons of non-β-branched amino acids were performed as described above using NOE connectivities derived from 2D-ROESY experiments.

Histidine Imidazole pK_a Titrations

Histidine imidazole titrations were collected as a series of 1D NMR spectra in 90/10 H_2O/D_2O from pH*=3.5 to pH*=9.0 in increments of 0.2 pH* units. Solution pH* was adjusted using 1.0 N NaOH or 1.0 N HCl and determined with a standardized Beckman Model ϕ32 pH meter using an NMR-probe microelectrode (MI-412) from Microelectronics, Inc. (Londonderry, NH). For the determination of histidine pK_a values in D_2O, spectra were collected in 99.9% D_2O in increments of 0.2 pD units at ambient temperature. Solution pD was adjusted using freshly prepared 1.0 N NaOD and determined using corrected pH readings (pD = pH + 0.40 units).[45]

The chemical shift changes of the imidazole ring protons (δ and ε) were plotted as a function of pH which resulted in a series of sigmoidal curves with inflection points observed at the pK_a of the hisitidine imidazole ring. A typical titration curve is shown in Figure 3-13. While an inspection of the curves can provide an estimate of the pK_a value, a more accurate value can be obtained by applying nonlinear least-squares analysis to a modified form of the Hill equation (Equation 3-3).

$$(\delta_{H+} - \delta_{obs}) / (\delta_{H+} - \delta_{H\circ}) = K_a^n / (K_a^n + [H+]^n) \quad (3\text{-}3)$$

Where:

δ_{H+} = chemical shift of the protonated imidazole ring

δ_{obs} = chemical shift of the imidazole ring at a given pH

$\delta_{H\circ}$ = chemical shift of the neutral imidazole ring

n = Hill coefficient

K_a = equilibrium constant

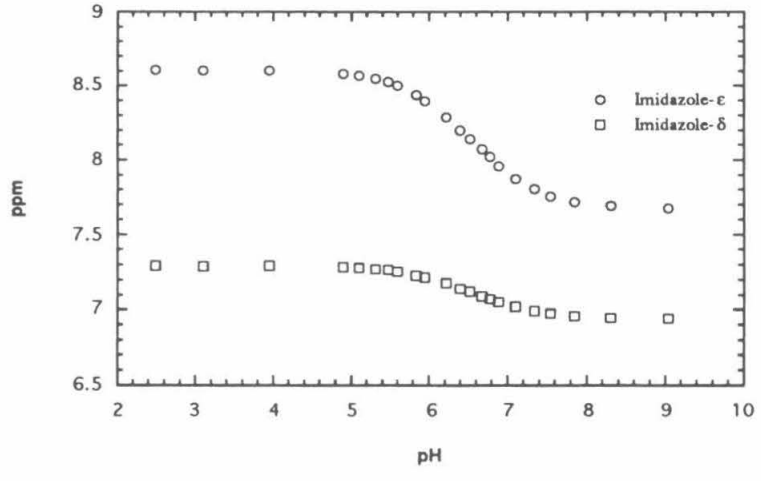


Figure 3-13. Results from a Typical Histidine Imidazole pH Titration Monitored through ^{1}H NMR Spectroscopy (Data shown for Ac-His-NHMe)

This equation can be further simplified to a linear equation which is a direct function of the observed pH:

$$\log[\xi / (1 - \xi)] = n\ \mathrm{pH} - n\ \mathrm{pK_a} \qquad \xi = (\delta_{H^+} - \delta_{obs}) / (\delta_{H^+} - \delta_{H^\circ})$$

Therefore, a linear plot of $\log[\xi / (1 - \xi)]$ vs pH will provide the Hill coefficent n as the slope and $-n\mathrm{pKa}$ as the y-intercept. The Hill coefficient provides a convenient indicator of the cooperativity of the dissociation. Ideally, if only one species is dissociating under the observed conditions, the Hill coefficient will be unity; any deviation from this value (beyond experimental error) implies that other dissociating groups are participating in the titration. All histidine $\mathrm{pK_a}$ values reported in this study were processed in this manner.

Measurement of Esterase Activity

The catalytic rate of *p*-nitrophenyl-acetate (*p*-NPA) hydrolysis for each peptide was obtained through measurement of the change in absorbance at 410 nm (due to *p*-nitrophenol release) over time using the conditions described by Katchalski *et al.*[46] (0.07 M phosphate buffer, pH=7.60, 5% v/v 1,4-dioxane). The reactions were run under pseudo-first order conditions with the peptide in large excess over the *p*-NPA concentration ([Pep] = 1-3 mM, [*p*-NPA] = 50 μM). In a typical experiment, a 25 μL aliquot of a 1.0 mM *p*-NPA (recrystallized twice from EtOH) in 1,4-dioxane (HPLC grade, Aldrich) was added to a thermostatted (25.5° C) solution of the peptide (1-3 mM) in buffer solution (350 μL) in a 1.0 cm quartz UV cell. The mixture was thoroughly mixed and the reaction was monitored continously at 410 nm for at least 10 half-lives on a Shimadzu UV160U Spectrometer equipped with a TCC-240A temperature-controlled cell holder. All buffer solutions were prepared from either Ultrapure water from a Milli-Q filtration system or commercially available HPLC grade water (Aldrich). The imidazole was recrystallized from ethanol two times prior to use.

The observed first order rate constants were calculated using graphical analysis of the data collected over the first three half-lives. The first order plots ($\ln(A_\infty - A_{obs})$ vs time) followed Equation 3-4, where A_∞ corresponds to the absorbance at infinite time (>10 half-lives), A_{obs} is the absorbance at any time t , and A_0 represents the absorbance at time t=0 min. In all cases the plots were found to be strictly linear for up to 90% of the reaction conversion, and the rate constant k_1 was

$$\ln(A_\infty - A_{obs}) = -k_1 (t) + \ln(A_\infty - A_0) \qquad (3\text{-}4)$$

obtained from the slope derived from linear regression analyses on the data. Second order rate constants k_{cat} were calculated by correction for the background hydrolysis rate (k_w) and normalization by the effective imidazole concentration ([Im_{eff}]) as shown in Equation 3-5. The effective imidazole concentration (the amount of neutral imidazole present in solution) was obtained through multiplication of the peptide concentration by the apparent degree of dissociation, α (See Equation 3-6, pK' = pK_a of histidine imidazole).[46]

$$k_{cat} = (k_1 - k_w) / [Im_{eff}] \qquad (3\text{-}5)$$

$$pH - pK' = \log [\, \alpha / (1 - \alpha)]$$

$$[Im_{eff}] = (\alpha)\,[Pep] \qquad (3\text{-}6)$$

Spectroscopic Data

Ac-Cys-Pro-D-Ser-His-NH_2

^{1}H NMR (10% D_2O/90% H_2O, 282K) δ: 8.62 (Ser-NH, 1H, d ,J=7.0 Hz), 8.55 (His ε, 1H, s), 8.53 (His-NH, 1H, d, J=9.0 Hz), 8.41 (Cys-NH, 1H, d, J=6.5 Hz), 7.67 (Amide Cap, s, 1H), 7.25 (His δ, Amide Cap, s, 2H), 4.86 (Cys α, m, 1H), 4.69 (His α, m, 1H), 4.46 (Pro α, m, 1H), 4.37 (Ser α, m, 1H), 3.80 (Pro δ, Ser β, m, 2H), 3.75

(Pro δ, Ser β, m, 2H), 3.33 (His β, m, 1H), 3.15 (His β, m, 1H), 2.82 (Cys β, m, 2H), 2.30 (Pro β, m, 1H), 1.90 (Pro β, γ(2), Acyl Cap, m, 6H).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 175.7, 175.2, 172.9, 171.5, 175.0, 134.7, 129.9, 118.5, 62.2, 62.1, 56.8, 54.9, 53.6, 48.9, 48.8, 30.4, 27.4, 25.8, 22.6.

HRMS : Calc. for $[MH]^+$ $C_{19}H_{30}N_7O_6S$ [484.1978]; Obs. [484.1984].

RP-HPLC RT : (A) 15.98 mins.

Ac-β-MeCys-Pro-D-Ser-His-NH_2

^1H NMR (10% D_2O/90% H_2O, 282K) δ: 8.83 (Ser-NH, 1H, d, J=7.2 Hz), 8.65 (His-NH, 1H, d, J=8.40 Hz), 8.63 (His ε, 1H, s), 8.47 (β-MeCys-NH, 1H, d, J=7.64 Hz), 7.74 (Amide Cap, s, 1H), 7.35 (His δ, Amide Cap, s, 2H), 4.73 (β-MeCys α, His α, m, 2H), 4.48 (Pro α, m, 1H), 4.41 (Ser α, m, 1H), 3.88 (Pro δ, Ser β, m, 2H), 3.80 (Pro δ, Ser β, m, 2H), 3.38 (β-MeCys β, His β, m, 2H), 3.19 (His β, m, 1H), 2.40 (Pro β, m, 1H), 2.10 (Pro γ(2), Acyl Cap, m, 5H), 2.00 (Pro β, m, 1H), 1.38 (β-MeCys γ, 3H, d, J=6.8 Hz).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 175.1, 174.5, 173.4, 172.4, 168.5, 133.8, 129.6, 129.1, 126.5, 117.6, 61.4, 61.3, 56.2, 52.8, 50.5, 29.7, 26.3, 25.1, 21.7, 11.9.

HRMS : Calc. for $[MH]^+$ $C_{20}H_{32}N_7O_6S$ [498.2134]; Obs. [498.2156].

RP-HPLC RT : (A) 21.20 mins.

Ac-Thr-Pro-D-Ser-His-NH_2

^1H NMR (10% D_2O/90% H_2O, 282K) δ: 8.72 (Ser-NH, 1H, d, J=6.2 Hz), 8.63 (His ε, 1H, s), 8.55 (His-NH, 1H, d, J=7.81 Hz), 8.30 (Thr NH, 1H, d, J=6.2 Hz), 7.69 (Amide Cap, s, 1H), 7.28 (His δ, Amide Cap, s, 2H), 4.71 (His α, m, 1H), 4.57 (Thr α, m, 1H), 4.45 (Pro α, m, 1H), 4.36 (Ser α, m, 1H), 4.10 (Thr β, m, 1H), 3.90 (Pro δ, m, 1H), 3.85 (Ser β, m, 1H), 3.77 (Ser β, m, 1H), 3.70 (Pro δ, m, 1H), 3.36 (His β,

m, 1H), 3.14 (His β, m, 1H), 2.30 (Pro β, m, 1H), 2.03 (Pro γ(2), Acyl Cap, m, 5H), 1.93 (Pro β, m, 1H), 1.22 (Thr γ, 3H, d, J=6.8 Hz).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 175.7, 175.1, 175.0, 172.8, 171.4, 134.5, 129.8, 118.4, 70.6, 68.0, 62.0, 58.2, 56.6, 53.5, 49.4, 30.3, 27.2, 25.8, 22.7, 19.9.

HRMS Calc. for $[MH]^+$ $C_{20}H_{32}N_7O_7$ [482.2392]; Obs. [482.2363].

RP-HPLC RT : (B) 12.38 mins.

Ac-(*allo*)Thr-Pro-D-Ser-His-NH_2

^{1}H NMR (10% D_2O/90% H_2O, 282K) δ: 8.64 (His-NH, 1H, d, J=7.6 Hz), 8.60 (His ε, 1H, s), 8.49 (His-NH, 1H, d, J=6.8 Hz), 8.41 (Thr NH, 1H, d, J=8.4 Hz), 7.73 (Amide Cap, s, 1H), 7.28 (His δ, Amide Cap, s, 2H), 4.72 (His α, m, 1H), 4.65 (Thr α, m, 1H), 4.47 (Pro α, m, 1H), 4.37 (Ser α, m, 1H), 4.04 (Thr β, m, 1H), 3.90 (Pro δ, m, 1H), 3.80 (Pro δ, Ser β, m, 2H), 3.74 (Ser β, m, 1H), 3.36 (His β, m, 1H), 3.14 (His β, m, 1H), 2.30 (Pro β, m, 1H), 2.03 (Pro β, γ(2), Acyl Cap, m, 6H), 1.25 (Thr γ, 3H, d, J=6.4 Hz).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 207, 206.5, 206.4, 203.6, 202.9, 156.6, 150.7, 136.7, 78.0, 75.3, 67.2, 62.5, 60.6, 51.8, 28.4, 24.5, 22.4, 18.9, 15.3.

HRMS : Calc. for $[MH]^+$ $C_{20}H_{32}N_7O_7$ [482.2363]; Obs. [482.2392].

RP-HPLC RT : (A) 13.41 mins.

Ac-Ser-Pro-D-Ser-His-NH_2

^{1}H NMR (10% D_2O/90% H_2O, 287K) δ: 8.66 (Ser(*i*+2)-NH, 1H, d, J=7.0 Hz), 8.60 (His ε, 1H, s), 8.58 (His-NH, 1H, d, J=8.96 Hz), 8.39 (Ser(*i*) NH, 1H, d, J=7.0 Hz), 7.69 (Amide Cap, s, 1H), 7.28 (His δ, Amide Cap, s, 2H), 4.74 (Ser(*i*) a, m, 1H), 4.71 (His α, m, 1H), 4.48 (Pro α, m, 1H), 4.36 (Ser(*i*+2) α, m, 1H), 3.90 (Pro δ(2), Ser(*i*,*i*+2) β(2), m, 6H), 3.36 (His β, m, 1H), 3.14 (His β, m, 1H), 2.30 (Pro β, m, 1H), 2.03 (Pro γ(2), Acyl Cap, m, 5H), 1.93 (Pro β, m, 1H).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 207.2, 206.5, 206.4, 203.7, 201.9, 156.5, 150.8, 136.8, 78.0, 67.4, 66.7, 60.8, 58.4, 56.7, 51.4, 28.3, 24.5, 22.7, 18.9.

HRMS : Calc. for $[MH]^+$ $C_{19}H_{30}N_7O_7$ [468.2207]; Obs. [468.2181].

RP-HPLC RT : (A) 10.80 mins.

Ac-Ser-Pro-D-Ser-Phe-NH_2

^1H NMR (10% D_2O/90% H_2O, 282K) δ: 8.51 (Ser(*i*+2)-NH, 1H, d, J=7.7 Hz), 8.42 (Ser(*i*) NH, 1H, d, J=7.7 Hz), 8.39 (Phe-NH, 1H, d, J=6.71 Hz), 7.69 (Amide Cap, s, 1H), 7.30 (Phe Ring, m 5H), 7.20 (Amide Cap, s, 2H), 4.74 (Ser(*i*) α, m, 1H), 4.60 (Phe α, m, 1H), 4.47 (Pro α, m, 1H), 4.35 (Ser(*i*+2) α, m, 1H), 3.86 (Pro δ, Ser(*i*+2) β, m, 2H), 3.75 (Pro δ, Ser(*i*,*i*+2) β, m, 5H), 3.23 (Phe β, m, 1H), 3.00 (Phe β, m, 1H), 2.30 (Pro β, m, 1H), 2.03 (Pro γ(2), Acyl Cap, m, 5H), 1.93 (Pro β, m, 1H).

^{13}C NMR (10% D_2O/90% H_2O, 282K) δ: 208.4, 207.0, 206.5, 202.0, 203.2, 160.5, 151.4, 150.7, 148.8, 78.0, 67.3, 66.9, 60.4, 59.8, 58.4, 51.4, 37.8, 28.3, 22.7, 18.9.

HRMS : Calc. for $[MH]^+$ $C_{22}H_{32}N_5O_7$ [478.2302]; Obs. [478.2330].

RP-HPLC RT : (A) 28.53 mins.

References

(1) Rose, D.; Gierasch, L.M.; Smith, J.A. "Turns in Peptides and Proteins," *Adv. Prot. Chem.* **1985**, *37*, 1-109.

(2) 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.

(3) Blow, D. "More of the Catalytic Triad," *Nature* **1990**, *343*, 694-695.

(4) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. "How Do Serine Proteases Really Work?" *Biochemistry* **1989**, *28*, 3629-3637.

(5) Christianson, D.W.; Alexander, R.S. "Carboxylate-Histidine-Zinc Interactions in Protein Structure and Function," *J. Am. Chem. Soc.* **1989**, *111*, 6412-6419.

(6) Craik, C.S.; Roczniak, S.; Largman, C.; Rutter, W.J. "The Catalytic Role of the Active Site Aspartic Acid in Serine Proteases," *Science* **1987**, *237*, 909-913.

(7) Derewanda, Z.S.; Sharp, A.M. "News from the interface: the molecular structures of triacylglyceride lipases," *Trends in Biol. Sci.* **1993**, *18*, 20-25.

(8) Sprang, S.; Standing, T.; Fletterick, R.J.; Stroud, R.M.; Finer-Moore, J.; Xuong, N-H.; Hamlin, R.; Rutter, W.J.; Craik, C.S. "The Three-Dimensional Structure of Asn^{102} Mutant of Trypsin: Role of Asp^{102} in Serine Protease Catalysis," *Science* **1987**, *237*, 905-909.

(9) Schneider, F. "Histidine in Enzyme Active Centers," *Angew. Chem. Int. Ed., Eng.* **1978**, *17*, 583-592.

(10) Sundburg, R.J.; Martin, R.B. "Interactions of Histidine and Other Imidazole Derivatives with Transition Metal Ions in Chemical and Biological Systems," *Chem. Rev.* **1973**, *74*, 471-517.

(11) Vallee, B.L.; Auld, D.S. "Zinc: Biological Functions and Coordination Motifs," *Acc. Chem. Res.* **1993**, *26*, 543-551.

(12) Walsh, C. *Enzymatic Reaction Mechanisms*; W.H. Freeman and Co.: New York, **1979**; pp 53-107.

(13) Masuda, Y.; Tanihara, M.; Imanishi, Y.; Higashimura, T. "Hydrolysis of Active Esters of Aliphatic Carboxylic Acids with Cyclic Dipeptide Catalysts Consisiting of L-Histidine and Different Aliphatic α-Amino Acids," *Bull. Chem. Soc. Jpn.* **1985**, *58*, 497-504.

(14) Kopple, K.D.; Nitecki, D.E. "Reactivity of Cyclic Peptides. II. *cyclo*-L-Tyrosyl-L-Histidyl and *cyclo*-L-Tyrosyltriglycyl-L-histidylglycyl," *J. Am. Chem. Soc.* **1962**, *84*, 4457-4464.

(15) Sheehan, J.C.; McGregor, D.N. "Synthetic Peptide Models of Enzyme Active Sites. I. *cyclo*-Glycyl-L-Histidyl-L-serylglycyl-L-histidyl-L-seryl," *J. Am. Chem. Soc.* **1962**, *84*, 3000-3005.

(16) Sheehan, J.C.; Bennett, G.B.; Schneider, J.A. "Synthetic Peptide Models of Enzyme Active Sites. III. Stereoselective Esterase Models," *J. Am. Chem. Soc.* **1966**, *88*, 3455-3457.

(17) Imanishi, Y.; Tanihara, M.; Sugihara, T.; Higashimura, T. "Histidine-Containing Cyclic Dipeptides as Catalysts in the Hydrolysis of Carbonic Acid *p*-Nitrophenyl Esters," *Biopolymers* **1977**, *16*, 2203-2215.

(18) Cruikshank, P.; Sheehan, J.C. "Synthetic Peptide Models of Enzyme Active Sites. II. L-Threonyl-L-alanyl-L-seryl-L-histidyl-L-aspartic Acid, an Esterase Model," *J. Am. Chem. Soc.* **1964**, *86*, 2070-2071.

(19) Chakrabarti, P. "Geometry of Interaction of Metal Ions with Sulfur-Containing Ligands in Protein Structures," *Biochemistry* **1989**, *28*, 6081-6085.

(20) Smith, J.A.; Pease, L.G. "Turns in Peptides and Proteins," *CRC Crit. Rev. Biochem.* **1980**, *8*, 315-399.

(21) Benedetti, E.; Morelli, G.; Nemethy, G.; Scheraga, H.A. "Statistical and energetic analysis of side-chain conformations in oligopeptides," *Int. J. Pept. Protein Res.* **1983**, *22*, 1-15.

(22) Janin, J.; Wodak, S.; Levitt, M.; Maigret, B. "Conformation of Amino Acid Side-chains in Proteins," *J. Mol. Biol.* **1978**, *125*, 357-386.

(23) Markley, J.L. "Observation of Histidine Residues in Proteins by Means of Nuclear Magnetic Resonance Spectroscopy," *Acc. Chem. Res.* **1975**, *8*, 70-80.

(24) Cederholm, M.T.; Stuckley, J.A.; Doscher, M.S.; Lee, L. "Histidine pKa shifts accompanying the inactivating Asp121-Asn substitution in a semisynthetic bovine pancreatic ribonuclease," *Proc. Natl. Acad. Sci., USA* **1991**, *88*, 8116-8120.

(25) He, Y-B.; Huang, Z.; Raynor, K.; Reisine, T.; Goodman, M. "Syntheses and Conformations of Somatostatin-Related Cyclic Hexapeptides Incorporating Specific α- and β-Methylated Residues," *J. Am. Chem. Soc.* **1993**, *115*, 8066-8072.

(26) Hruby, V.J.; Toth, G.; Gehrig, C.A.; Kao, L.-F.; Knapp, R.; Lui, G.K.; Yamamura, H.I.; Kramer, T.H.; Davis, P.; Burks, T.F. "Topologically Defined Analogues of [D-Pen, D-Pen^5]enkephalin," *J. Med. Chem.* **1991**, *34*, 1823-1830.

(27) Yamazaki, T.; Abe, A. "Conformational Studies of *N*-Acetyl-*N'*-Methylamide Derivatives of α-Aminobutyric Acid, Norvaline, and Valine. I. Preferred Conformations in Solution as Studied by ^{1}H-NMR Spectroscopy," *Biopolymers* **1988**, *27*, 969-984.

(28) In the case of threonine and *allo*-threonine, the hydroxyl group is taken as the largest substituent for defining the rotamer population; for example, threonine in the g^+ rotamer refers to the hydroxyl group in the g^+ position, with the methyl group residing in the t position.

(29) Molecular modelling studies were performed on a Silicon Graphics Personal Iris 4D/25TG using Insight II v.2.1.0 Modeling Package (Biosym Technologies). In

all cases, the peptide fixed in a Type II reverse turn conformation using ideal dihedral angles for ϕ and ψ.

(30) Bauer, C.J.; Frenkiel, T.A.; Lane, A.N. "A Comparison of the ROESY and NOESY Experiments for Large Molecules, with Application to Nucleic Acids," *J. Magn. Res.* **1990**, *87*, 144-152.

(31) Bender, M.L.; Kezdy, F.J.; Wedler, F.C. "α-Chymotrypsin: Enzyme Concentration and Kinetics," *J. Chem. Ed.* **1967**, *44*, 84-88.

(32) Bruice, T.C.; Schmir, G.L. "Imidazole Catalysis. I. The Catalysis of the Hydrolysis of Phenyl Acetates by Imidazole," *J. Am. Chem. Soc.* **1957**, *79*, 1663-1667.

(33) Anderson, B.M.; Cordes, E.H.; Jencks, W.P. "Reactivity and Catalysis in Reactions of the Serine Hydroxyl Group and of *O*-Acyl Serines," *J. Biol. Chem.* **1961**, *236*, 455-463.

(34) Bender, M.L.; Pollock, E.J.; Neveu, M.C. "Deuterium Oxide Solvent Isotope Effects in the Nucleophilic Reactions of Phenyl Esters," *J. Am. Chem. Soc.* **1962**, *84*, 595-599.

(35) Heller, M.J.; Walder, J.A.; Klotz, I.M. "Intramolecular Catalysis of Acylation and Deacylation in Peptides Containing Cysteine and Histidine," *J. Am. Chem. Soc.* **1977**, *99*, 2780-2785.

(36) Morell, J.L.; Fleckenstein, P.; Gross, E. "Stereospecific Synthesis of (2*S*,3*R*)-2-Amino-3-mercaptobutyric Acid - an Intermediate for Incorporation into β-Methyllanthionone-Containing Peptides," *J. Org. Chem.* **1977**, *42*, 355-356.

(37) Photaki, I.; Taylor-Papadimitriou, J.; Sakarellos, C.; Mazarakis, P.; Zervas, L. "On Cysteine and Cystine Peptides. Part V. *S*-Trityl- and *S*-Diphenyl-methyl-cysteine and -cysteine Peptides," *J. Chem. Soc.* **1970**, 2683-2687.

(38) Manning, M.C.; Illangasekare, M.; Woody, R.W. "Circular Dichroism Studies of Distorted α-Helices, Twisted β-Sheets, and β-Turns," *Biophys. Chem.* **1988**, *31*, 77-86.

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

(40) Elliot, D.F. "Preparation of L-Threonine. Interconversions of the Four Stereoisomeric α-Amino-β-Hydroxybutyric Acids," *J. Chem. Soc.* **1950**, 62-68.

(41) Ellman, G.L. "Tissue Sulfhydryl Groups," *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.

(42) Schwarzenbach, G.; Flashka, H. *Complexometric Titrations*; Interscience: New York, **1969**.

(43) Pachler, K.G.R. "Nuclear Magnetic Resonance Study of Some α-Amino Acids - II. Rotational Isomerism," *Spectrochimica Acta* **1964**, *20*, 581-587.

(44) Cung, M.T.; Marraud, M. "Conformational Dependence of the Vicinal Proton Coupling Constant for the C^{α}-C^{β} Bond in Peptides," *Biopolymers* **1982**, *21*, 953-967.

(45) Fife, T.H.; Bruice, T.C. *J. Phys. Chem.* **1961**, *65*, 1079.

(46) Katchalski, E.; Fasman, G.D.; Simons, E.; Blout, E.R. "Synthetic Histidine-Containing Polypeptides as Catalysts for the Hydrolysis of *p*-Nitrophenyl Acetate," *Arch. Biochem. Biophys.* **1960**, *88*, 361-365.

Chapter 4. Synthesis of Unnatural Amino Acids with 2,2'-Bipyridine and 1,10-Phenanthroline Side Chains

Introduction

Protein-bound metal cations play central roles in almost all aspects of cellular function. In fact, it has been estimated that over 30% of all enzymes are dependent upon metal cations for efficient activity.[1] Metal cations have been implicated in the nucleation and stabilization of protein supermolecular structure (e.g., zinc fingers [2,3]) and have been implicated as templates for macromolecular assembly.[4] Metal cations have also been shown to play active roles in protein function, from electron transfer processes and enzymatic catalysis to genetic and metabolic regulation.[5] The ability of metal ions to serve both structural and functional roles in metalloproteins has prompted an intense effort in the area of *de novo* protein design to incorporate metal cation binding sites within polypeptide frameworks to create new three-dimensional structures with defined structural and potentially functional properties.

In the design of new protein constructs, it has become apparent that the number and diversity of the ribosomally-encoded metal binding amino acids is surprisingly limited, particularly when compared to the wide variety of synthetic ligands which are available for selective complexation of metal ions in aqueous media.[6] The majority of metal-binding sites found in proteins are constructed from just four functional groups: imidazole (histidine), sulfur (cysteine, methionine), side chain carboxylate (aspartate, and glutamate), and the main chain carbonyl oxygen.[1] Thus, a current emphasis has been to expand the repertoire of protein building blocks through the design and construction of *unnatural* metal-binding amino acids that would enhance metal cation selectivities as well as widen the range of metal coordination geometries over that which is currently available.[7-9]

Multidentate ligands are especially interesting in terms of *de novo* design, as they generally afford high binding affinities and selectivities of metal cations, while simultaneously reducing the total number of residues needed to construct a full metal coordination sphere. This economy of residues is especially appealing in *de novo* protein designs reliant upon solid phase peptide synthesis for production of the polypeptide chain,

since this technique limits the total number of residues that can be employed in the overall design (<100 residues). Several synthetic amino acids capable of multi-dentate coordination have been reported which are based on the well-known dicarboxylate ligands, ethylenediaminetetraacetic acid (EDTA)[8] and iminodiacetic acid (IDA).[7,9] In one documented case, a *de novo* designed polypeptide which incorporated the IDA-derived amino acid demonstrated enhanced helical structure and structural stability upon metal cation binding.[9]

Another series of well-studied bidentate ligands are the heteroaromatic 1,10-phenanthroline and 2,2'-bipyridine ring systems. These ligands have been used extensively in a wide range of chemical applications, including supramolecular chemistry for the design of metal-directed self-assemblies.[10] Lehn and co-workers[11] have recently designed the so-called "deoxyribonucleohelicates" from a bipyridyl-ether polymer with covalently attached deoxyribonucleoside sidechains. These structures self assemble in the presence of various metal cations to form a helical core of bipyridyl-metal cation complexes. Similar helicate structures have been reported for substituted phenanthroline analogs.[10] Most recently, the bipyridine ligand has been used in *de novo* protein design of α-helical bundle motifs by two independent research groups.[12,13] In both cases, the bipyridine ligands were covalently attached to the *N*-terminus through an amide linker, and the modified peptides self-assembled to form three helical bundles upon metal cation binding. While these studies demonstrate the potential power of the bipyridyl amino acids in *de novo* design for the development of novel protein motifs, the fact that the ligands have only been covalently attached to the amino terminus (or through a pendant side chain) puts a severe limitation on the utility and applications of the incorporated ligand.

A more general and practical approach would be to incorporate these moieties as α-amino acid derivatives, as this would allow for placement of the ligand anywhere in the primary sequence of a synthetic peptide. Toward this goal, the work presented in this chapter describes the stereoselective syntheses of a series bipyridyl and phenanthrolyl-

substituted α-amino acids, and the subsequent N^{α}-FMOC amino protection to make them suitable for solid phase peptide synthesis. The amino acids, (*S*)-2-α-amino-3-(2,2'-bipyrid-6-yl)propanoic acid (**1**, See Figure 4-1), (*S*)-2-α-amino-3-(2,2'-bipyrid-4-yl)propanoic acid (**2**), (*S*)-2-α-amino-3-(2,2'-bipyrid-5-yl)propanoic acid (**3**), and (*S*)-2-α-amino-3-(1,10-phenanthrol-2-yl)propanoic acid (**4**) represent the three sterically accessible regioisomers of the 2,2'-bipyridine ring system and the *ortho*-substituted analog of the 1,10-phenanthroline ring family.

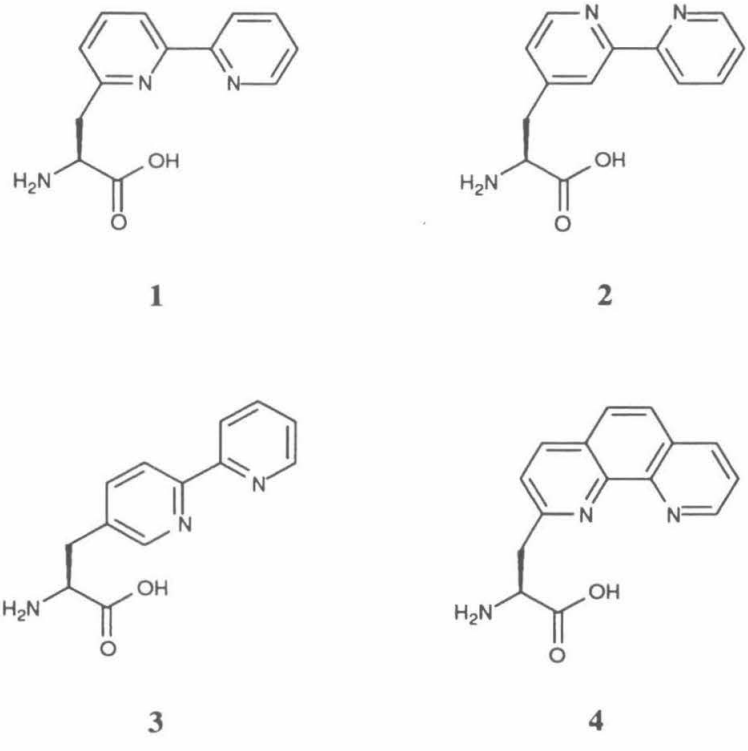


Figure 4-1. Heteroaromatic metal-chelating amino acids

The four amino acids provide a complementary set of metal ligands that allow for the design of an almost unlimited number of metal binding site geometries with controlled binding affinities. The specific regioisomers of the bipyridine ring are known to have widely different metal cation binding affinities and metal coordination geometry

preferences. It is well known that substitutions in the 6 or 6' positions (as in the bipyridyl-amino acid **1**) have reduced metal binding affinities due to the steric crowding near the metal-ligating nitrogen atoms.[14,15] In addition, the substitutions at these positions tend to disfavor those metal coordination geometries where steric interactions can become an issue, such as square planar or octahedral geometries. The 1,10-phenanthroline analog, also substituted at the *ortho* position, would be expected to have similar metal coordination properties; however, the metal-binding affinities of the phenanthroline series are known be higher (approximately 10-100 fold) than the corresponding 6-substituted-bipyridine ligands.[16] This has been attributed to the reduced conformational flexibility and the rigid amine orientation present in the phenanthroline ring system due to the extended aromatic ring system.[15] As a result, the phenanthroline substituted amino acid can serve as a convenient alternative to the 6-bipyridyl amino acid, when strong metal cation binding and sterically constrained metal geometries are desired.

In contrast, the 4- and 5-substituted bipyridyl amino acids, which contain an unobstructed N-N' chelation moiety, are complementary to the 6-bipyridyl and 2-phenanthrolyl-substituted amino acids. These amino acids are expected to have metal binding properties comparable in both metal-cation binding affinities and coordination geometry preferences to 2,2'-bipyridine.[16] However, the subtle difference between these two amino acids allows for specific control of the ligand orientation with respect to the polypeptide backbone. These attributes could be used to "fine-tune" a programmed metal binding site to achieve the desired metal binding orientation and geometry within a given designed motif.

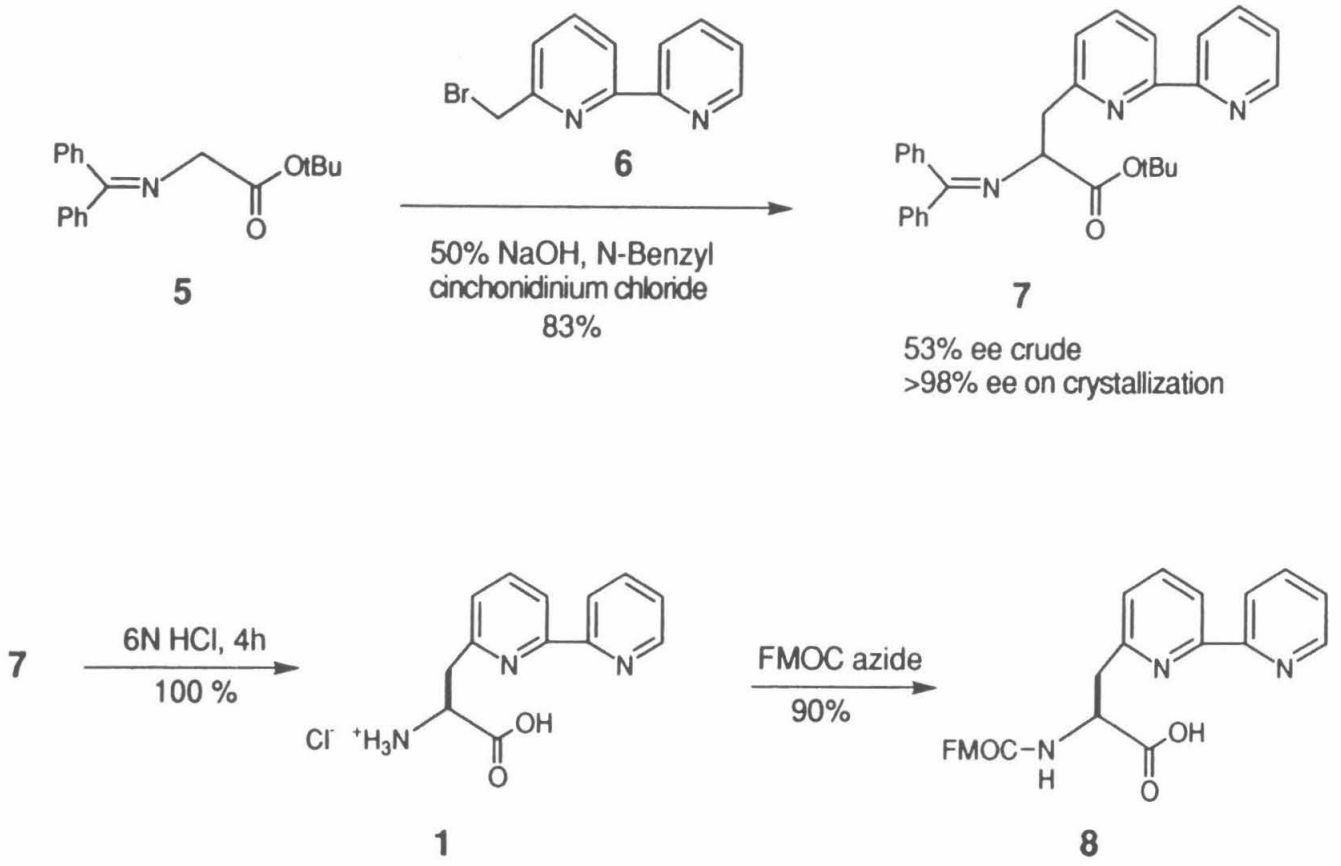
In addition to the varied metal cation binding properties, these ligands have several other qualities that make them particularly useful in terms of *de novo* design and peptide incorporation. First, since the nitrogen atoms are a constitutive part of the aromatic ring system, the ligands do not contain any nucleophilic functionalities that would interfere with peptide synthesis. As a result, the unnatural amino acids (**1-4**) can be directly incorporated

into peptides using solid phase peptide synthesis protocols without side-chain protecting group manipulations. Secondly, the heteroaromatic ring systems of 1,10-phenanthroline and 2,2'-bipyridine are stable to a wide variety of synthetic conditions, including those used for the cleavage of peptides from solid resin supports after peptide synthesis (trifluoroacetic acid, thiol scavengers, etc.). Finally, since the stereospecific attachment of the ligands to the amino acid framework is in the *S* configuration, the amino acids (**1-4**) are compatible with all of the commonly-occurring residues. As a result, these amino acids can be used in conjunction with any of the ribosomally-encoded amino acids in the construction of homochiral polypeptide sequences.

Results and Discussion

*Synthesis of (S)-2-Amino-3-(bipyrid-6-yl)-propanoic Acid (6Bpa , **1**)*

The stereoselective synthesis of **1** was accomplished by the asymmetric alkylation of commercially available N-(diphenyl-methylene)glycine *tert*-butyl ester (**5**) with 6-(bromomethyl)-2,2'-bipyridine[17] (**6**) using the phase transfer catalyst, (8*S*,9*R*)-(-)-N-benzylchinchonidinium chloride (Scheme 4-1) according to the method of O'Donnell *et al.*[18] The alkylation proceeded with modest asymmetric induction, 53% ee in favor of the *S* enantiomer, and affords an 80% chemical yield of pure product, **7** after purification.[19] Enantiomerically pure material (>99% ee) was obtained, as reported for several other derivatives,[18] after crystallization of the racemate from a hexane solution of enantiomerically enriched **7** at 0 °C. Quantitative hydrolysis of **7** was then effected by refluxing in 6N hydrochloric acid for 4 hr to provide (*S*)-**1** in 40% overall yield from **5**. The enantiomeric excess of the amino acid was evaluated either by ^{1}H NMR or HPLC analysis of the (*R*)-(+)-α-methoxy-α-(trifluoromethyl)phenylacetic acid amide (Mosher's Amide)[20] prepared from the methyl ester of **1**. Of the two methods, HPLC provided a more accurate, independent measure of the enantiomeric excess, as the NMR analysis was complicated due to overlapping resonances in the ^{1}H NMR spectrum.



Scheme 4-1. Stereoselective Synthetic Route to (*S*)-2-Amino-3-(2,2'-bipyrid-6-yl)propanoic Acid

It should be noted that while the absolute stereochemical outcome of the asymmetric alkylation was not determined rigorously, the alkylation products have been found to be quite predictable with a wide variety of electrophiles.[18] In addition, the change in the direction of the optical rotation of **1** in going to an acidic aqueous solution ($[\alpha]_D$ (-)18.6° (c=1, H_2O, 25°C); $[\alpha]_D$ (-) 12.9° (c=1, 5N HCl, 25°C)] is in keeping with the Clough-Lutz-Jirgenson rule,[21] and the two enantiomers follow the predicted order of elution from a chiral CrownPak CR (-) HPLC column. (See Experimental section this chapter.) Since this amino acid was designed for incorporation into polypeptide sequences using solid phase peptide synthesis methods, the corresponding amino acid was converted to the corresponding 9-fluorenylmethoxycarbonyl (FMOC) derivative **8** through treatment with FMOC-azide in 50% aqueous dioxane containing 5% sodium carbonate (70% yield).[22]

The overall synthetic route provides a convenient method to obtain large scale quantities of the amino acid starting from 6-bromomethyl bipyridine. In fact, the most difficult aspect of the synthetic route has involved generation of large quantities of the bromomethyl-bipyridine derivative **6**. The reported route[17] employs a radical initiated bromination of the 6-methyl-2,2'-bipyridine[23] with *N*-bromo-succinimide and benzoyl peroxide. This method provided only low yields (<30%) of the desired bromomethyl derivative as the reaction conditions had to be compromised to avoid the competing reaction to form the ω,ω-dibromo compound. Subsequent purification of the reaction mixture also proved difficult, as the monobromo- and dibromo-methyl derivatives had very similar chromatographic properties to the starting material on both silica gel and alumina chromatographic supports. However, the recent report of modifications to these procedures, as well as an alternative, high-yielding synthetic route employing CBr_4 and PPh_3 holds promise for the facile generation of large scale quantities of the 6-bromomethyl derivative.[24]

Unfortunately, the synthetic route employed for the 6-bipyridyl-alanine was not found to be general for the other regioisomers of the bipyridine amino acids. In particular,

attempts to prepare the 4-bipyridine-substituted amino acid through the asymmetric alkylation were hindered with problems similar to those found for the 6-bromomethyl-bipyridine in the preparation and purification of large scale quantities of the starting material, 4-bromomethyl-bipyridine. In addition, difficulties were encountered in the resolution of the amino acid imine derivative.[25] While the asymmetric alkylation of the 4-bromomethyl bipyridine proceeded smoothly in moderate ee (60%), all attempts to resolve the amino acid through selective crystallization of the racemic mixture of protected imine failed, making the subsequent resolution of the amino acid impossible. As a result of these problems, a number of enantioselective enzymatic hydrolysis reactions were investigated as an alternate route to obtain the optically pure 4-bipyridyl-substituted amino acid.

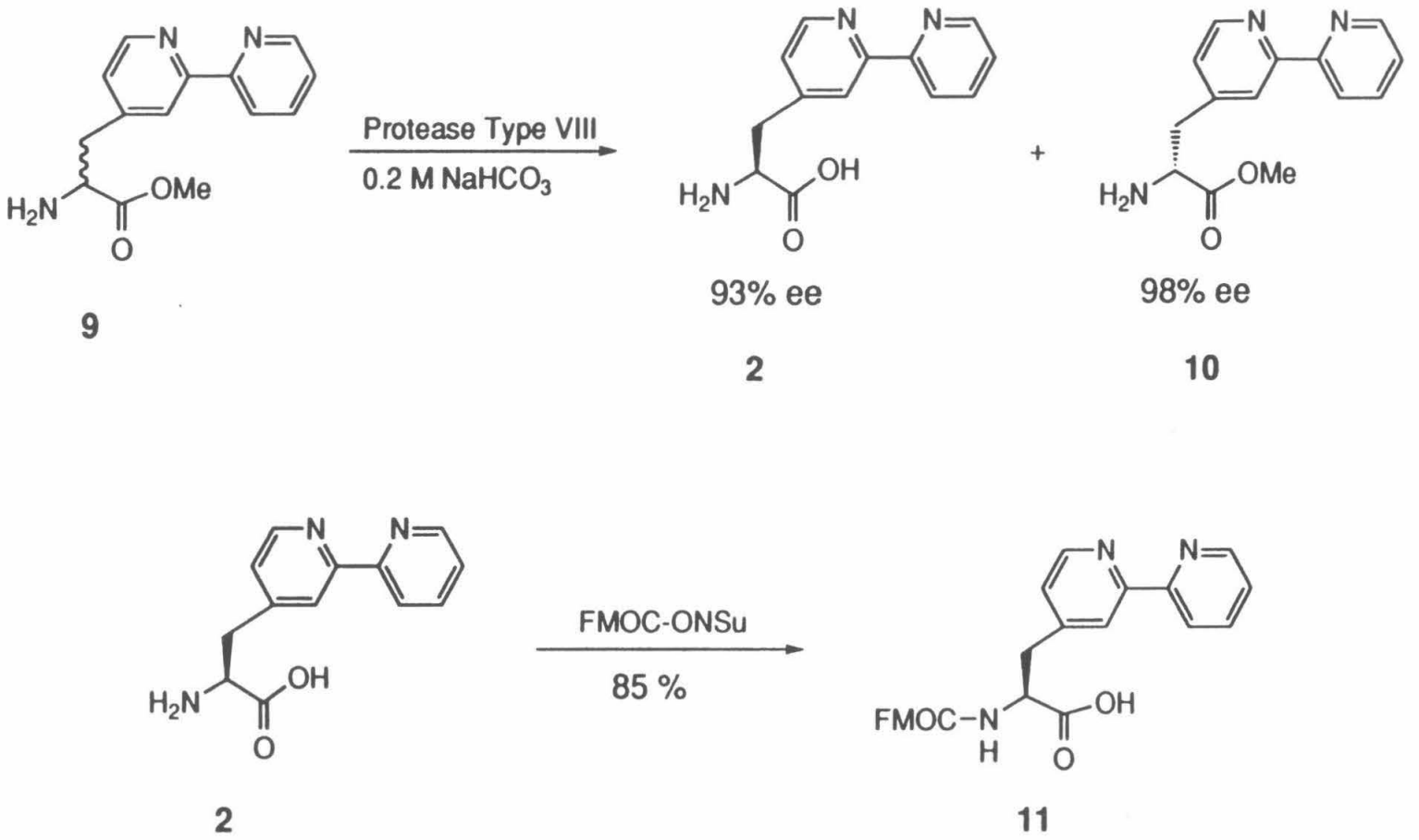
Enzymatic Resolution of Heteroaromatic Metal-Chelating Amino Acids

A number of commercially available enzymes have been shown to effectively resolve racemic mixtures of both natural and unnatural amino acid derivatives, several of which are specific for large hydrophobic sidechains.[26] One example, Acylase I isolated from *Aspergillus*, has been shown to be effective in resolving a wide variety of aromatic and β-branched amino acids to near enantiomeric purity.[27] Hydrolyses catalyzed by this enzyme generally afford near quantitative hydrolysis of the L-amino acid derivative, while leaving the corresponding *N*-acyl D-amino acid derivative intact. Separation of the two enantiomers is then accomplished by an extraction of the remaining *N*-acyl derivatives from the reaction mixture under acidic conditions. Unfortunately, attempts to resolve enantiomeric mixtures of *N*-acylated 4-bipyridyl-amino acid using the conditions of Chenault, Dahmer, and Whitesides[27] were unsuccessful, and no reaction was observed even after incubation for several days. This lack of activity could be attributed to the steric bulk of the bipyridine ring system; however, since Acylase I is known to be a zinc metalloenzyme, it is likely that the strong metal-binding affinities of the bipyridyl ligand effectively inhibited the enzyme through chelation of the essential zinc cation.

Another enzyme which has been shown to effectively resolve enantiomeric mixtures of hydrophobic/aromatic amino acid derivatives is the alkaline protease Subtilisin Carlsberg.[28] This protease was expected to be ideal for the bipyridyl-substituted amino acids since it is not dependent upon metal cations for activity and inspection of the crystal structure coordinates[29,30] indicated that the substrate binding pocket was compatible with the large aromatic side chains of **1-4**. Unlike Acylase I, subtilisin was found to be extremely selective in the hydrolysis of the 4-bipyridyl-alanine methyl esters. As shown in Scheme 4-2, the L-methyl ester was selectively hydrolyzed from the optically impure mixture **9** using Subtilisin Carlsberg, using the conditions outlined by Chen et al.[28] The hydrolysis proceeded smoothly at room temperature and was generally complete within two hours to afford the L-amino acid **2** in high optical purity (93% ee). Separation of the two enantiomers was accomplished through a simple extraction of the D-amino acid ester **10** (98% ee) with $CHCl_3$.

Finally, as our main goal is to incorporate the bipyridyl within a peptide backbone, the amino acids were *N*-protected as the 9-fluorenylmethyl carbamate (FMOC) in preparation for solid phase peptide synthesis. Since the enzymatic resolution reactions and the N^{α}-FMOC protection steps are run under similar conditions, the crude amino acid from the enzymatic resolution was used without further purification[31] and was treated with FMOC-succinimidyl carbonate[32] in 30% aqueous dioxane containing 10% sodium carbonate (85% yield based on L-amino acid methyl ester) to form the 9-fluorenylmethyl carbamate derivative **11**.

The protease-catalyzed hydrolysis proved to be not only simple and convenient, but also allowed for the *quantitative retrieval* of both amino acid enantiomers in high optical purity on preparative scales (>3g). In addition, the enzymatic hydrolysis was found to be quite general, as all of the regioisomers of the bipyridine L-amino acid methyl esters, as well as the phenanthroline L-amino acid methyl ester, were found to be selectively hydrolyzed by the enzyme in high enantiomeric excess, while the D-amino acid esters



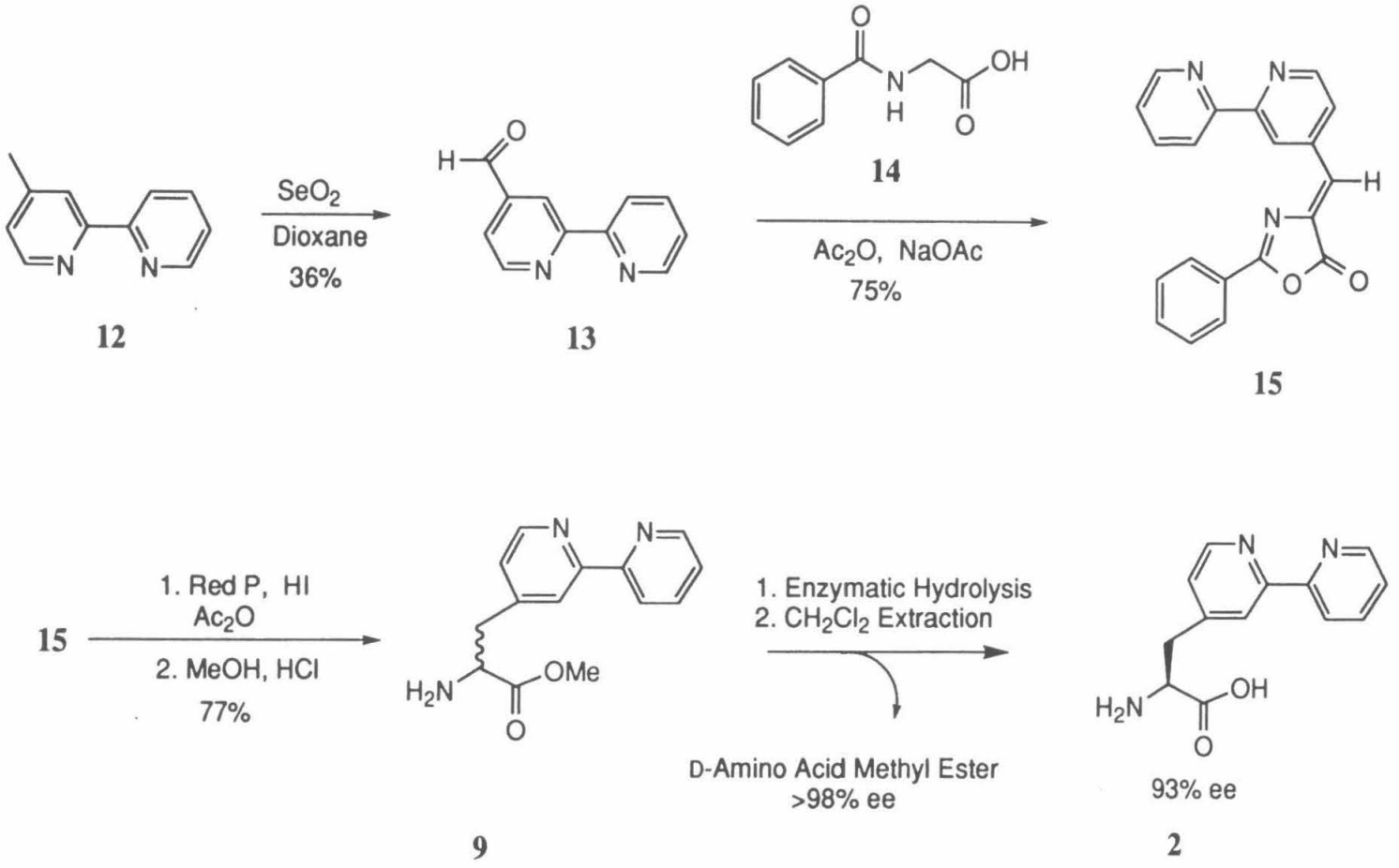
9 2 10 2 11

Scheme 4-2. Enzymatic Hydrolysis Resolution of 2-Amino-3-(2,2'-bipyrid-4-yl)propanoic Acid

remained intact. (See below for details.) Finally, the enzyme was found to be robust and very stable to a wide variety of reaction conditions, even in the presence of organic solvents. Also, even though the commercially available preparation of the enzyme contains a mixture of proteases (Subtilisin Carlsberg is the major component), the enzyme catalyzed hydrolyses were found to be consistent with different batches of the enzyme preparation.

Synthesis of 2-Amino-3-(2,2'-bipyrid-4-yl)propanoic Acid (4Bpa, **2***)*

The efficiency of the enzymatic resolution, coupled with the difficulties encountered in the preparation of 4-bromomethyl-bipyridine, prompted a re-evaluation of the synthetic route to the 4-bipyridyl-amino acid. Since the enzymatic resolution provided a means to quantitatively retrieve the L-enantiomer, it was no longer necessary to follow a stereoselective synthetic route. As a result, the synthesis of the 4-bipyridyl amino acid ester **9** was accomplished using a synthetic route originally described by Erlenmeyer for the generation of racemic, aromatic amino acids.[33,34] As shown in Scheme 4-3, oxidation of 4-methyl-bipyridine[35] (**12**) with selenium dioxide in dioxane afforded the carboxaldehyde **13** (36%) which was subsequently condensed with hippuric acid **14** in acetic anhydride containing one equivalent of sodium acetate at 90°C to form the bipyridyl-azlactone **15** (75 %). Reduction of the azlactone with red phosphorous and hydriodic acid in acetic anhydride,[34] followed by esterification of the amino acid in methanol saturated with anhydrous HCl, afforded the racemic methyl ester **9**. Subsequent enzymatic resolution of the methyl ester as described above yielded the L-amino acid **2** in >93% ee and the D-amino acid methyl ester in >98% ee. Conversion of the L-amino acid to the N^{α}-FMOC derivative **11** with FMOC-ONSu proceeded smoothly to afford the protected amino acid in high yield (90% yield, ~10% overall yield starting from 4-methyl-bipyridine).

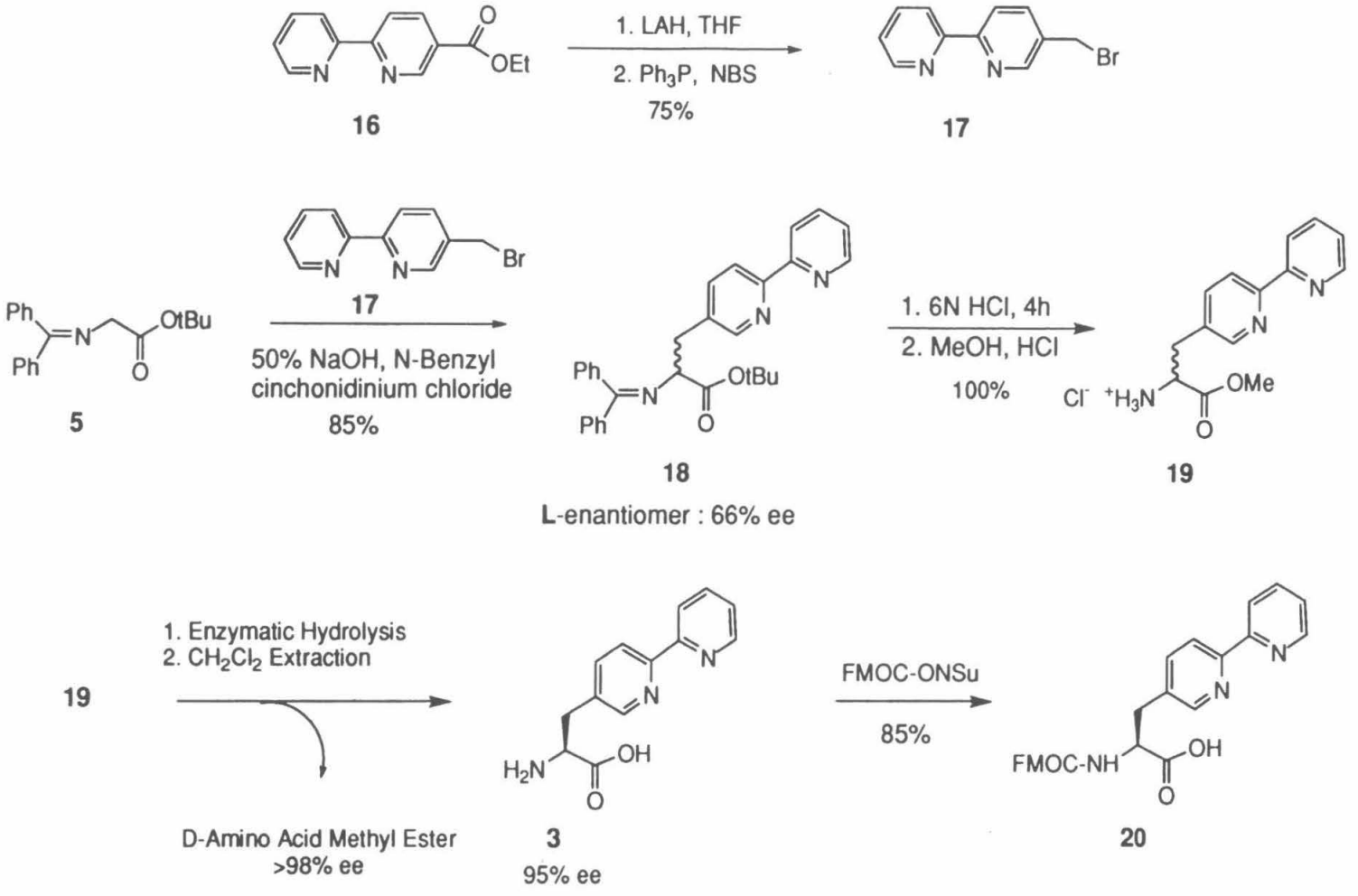


Scheme 4-3. Synthetic Route to (*S*)-2-Amino-3-(2,2'-bipyrid-4-yl)propanoic Acid

Synthesis of 2-Amino-3-(2,2'-bipyrid-5-yl)propanoic Acid (5Bpa, 3)

Synthesis of the 5-substituted-2,2'-bipyridyl derivative **3** (Scheme 4-4) followed the stereoselective route employed for the 6-bipyridyl-substituted amino acid (**1**, See Scheme 4-1), the key step involving an asymmetric alkylation of commercially available N-(diphenyl-methylene)glycine *tert*-butyl ester (**5**) with 5-(bromomethyl)-2,2'-bipyridine (**17**) using the phase transfer catalyst, (8*S*,9*R*)-(-)-*N*-benzylcinchonidinium chloride.[18] The alkylation proceeded with modest asymmetric induction (40-60% ee) with an 85% overall yield of the protected amino acid **18**. For reasons that are still not fully understood, the asymmetric induction was found to be sensitive to preparative scale-up. Upon increasing the scale of the reaction above 1mM in reactants, the asymmetric induction dropped precipitously (40-60% to <20%). As a result, when preparative scales were performed, the reaction was run in small batches (with reactants <1mM) which were subsequently pooled for purification. Recent work in this laboratory has shown that the asymmetric induction is dependent upon efficient mixing of the two phases, and high enantiomeric excesses have been achieved on preparative scales for related systems.[36] Rather than attempt to resolve the amino acid-imine derivatives through selective crystallization of the racemic mixture, the protected amino acid derivative **18** was subsequently hydrolyzed in 6N HCl, followed by esterification of the amino acid in acidic methanol afforded the optically enriched methyl ester **19** which was ready for the protease-catalyzed enzymatic resolution. Alkaline protease catalyzed hydrolysis of the enantiomerically enriched methyl ester was found to extremely stereoselective, resulting in quantitative retrieval of the L-amino acid **3** in >96%ee. The amino acid was subsequently N^{α}-FMOC protected as described for the 4-bipyridyl-amino acid to afford the protected amino acid derivative **20** in high yield (85%).

Unlike the preparations of 4-bromomethyl-bipyridine and 6-bromomethyl-bipyridine, generation of the 5-bromomethyl bipyridine was relatively facile and could be accomplished on a preparative scale (> 3 g). Preparation of the 5-bromomethyl-bipyridine



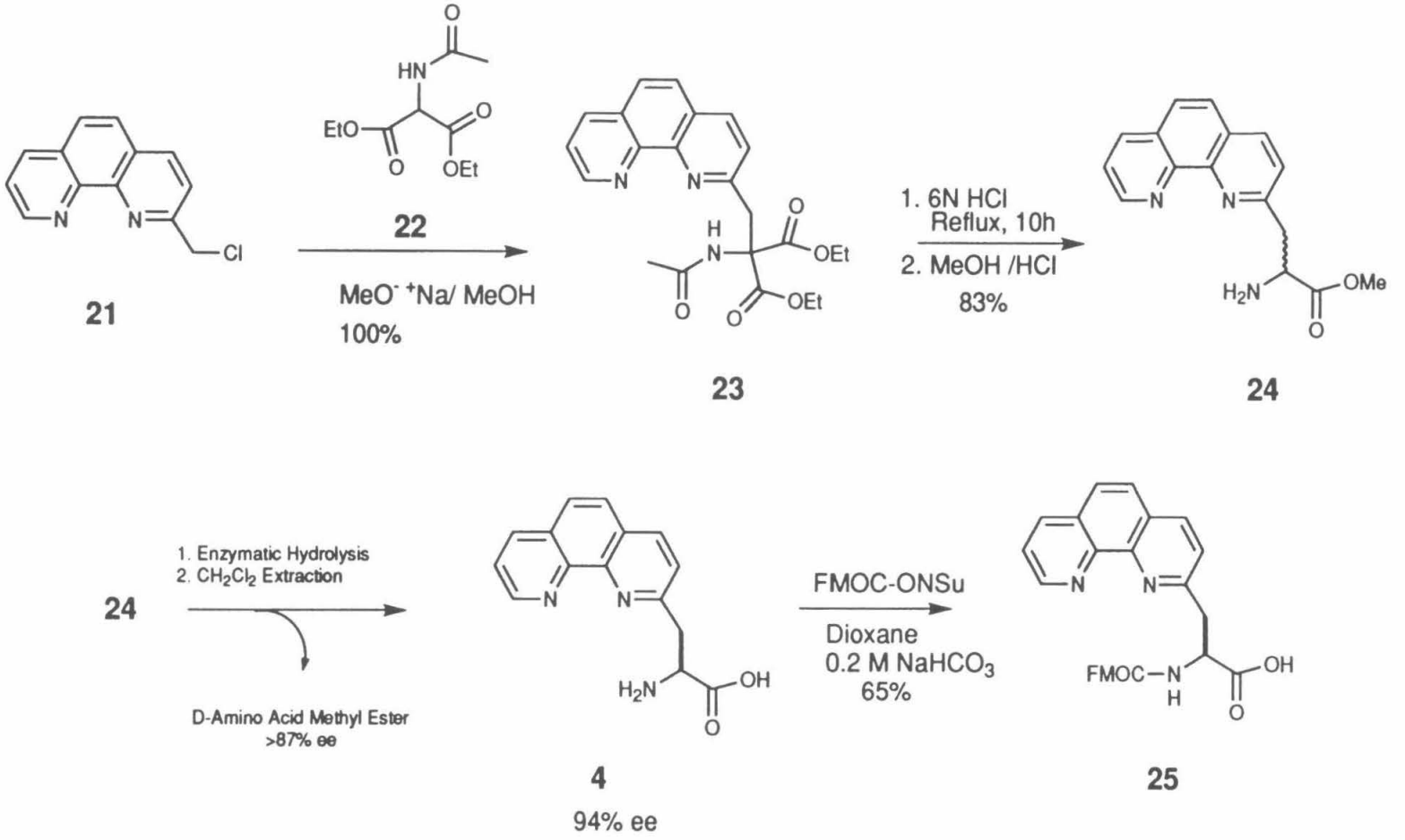
Scheme 4-4. Synthetic Route to (*S*)-2-Amino-3-(2,2'-bipyrid-5-yl)propanoic Acid

(**17**) was accomplished by reduction of the ester **16** to the corresponding alcohol using lithium aluminum hydride and subsequent bromination using triphenylphosphine and N-bromo-succinimide in CH_2Cl_2 to afford 5-(bromomethyl)-2,2'-bipyridine (**17**) in 75% yield (over two steps). Preparation of the ester **16** was accomplished in moderate yield (50-70%), on preparative scales (5-10g) as reported.[12]

Synthesis of 2-Amino-3-(1,10-phenanthrol-2-yl)propanoic Acid (2Fen, **4***)*

Synthesis of the 2-substituted-1,10-phenanthrolyl amino acid **4** was accomplished using a modified version of the Sorenson method[37] to prepare racemic amino acids. The key step involved alkylation of the glycine equivalent diethyl acetamidomalonate (**22**) with 2-chloro-1,10-phenanthroline[38] (**21**) in sodium ethoxide/ethanol solution (Scheme 4-5). The alkylation proceeded in a quantitative yield (yield based on diethyl acetamidomalonate) and no subsequent purification of the desired alkylated product **23** was required. Exhaustive hydrolysis of the racemic material **23** in refluxing 6N HCl, followed by esterification of the amino acid in acidic methanol produced the racemic amino acid methyl ester **24** ready for enzyme-catalyzed resolution.

Alkaline protease catalyzed hydrolysis of the 2-phenanthroline-substituted amino acid methyl ester **24** was found to be highly stereoselective, as expected from the positive results obtained from the bipyridyl substituted amino acid esters (**9**, **19**). However, unlike the bipyridyl-amino acid methyl esters (**9**, **19**), the phenanthroline methyl ester **24** was only sparingly soluble in the 10% sodium bicarbonate buffer employed for the enzymatic hydrolysis. As a result, a number of mixed organic/aqueous solvent conditions were screened for enzyme activity and stereoselectivity. It was found that while 10% (v/v) *t*-butanol/10% sodium bicarbonate had no effect on the catalytic efficiency or stereoselectivity compared to neat buffer, both methanol and dimethylsulfoxide (10% v/v) not only slowed the hydrolysis rate by ~2 fold, but the stereoselectivity of the enzyme was compromised as well (ee ~ 85% in these solvent systems). As a result of the modified



Scheme 4-5. Synthetic Route to (*S*)-2-Amino-3-(1,10-phenanthrol-2-yl)propanoic Acid

reaction conditions, the phenanthroline methyl ester enzymatic resolution hydrolysis required more units of enzyme (~2 fold more enzyme required) and a shorter reaction time (40 min vs. 2h) than had been employed in the corresponding bipyridyl-amino acid enantiomeric resolutions to maintain a sufficiently high ee (>93%) of the L-enantiomer **4**. Subsequent FMOC protection of the L-amino acid **4** proceeded smoothly (65%) with FMOC-ONSu to provide the fully-protected phenanthroline amino acid derivative **25**.

Synthesis of the starting material, 2-chloromethyl-1,10-phenanthroline (**21**), was accomplished through halogenation of 2-hydroxymethyl-1,10-phenanthroline with thionyl chloride as described by Engerbersen and co-workers.[38] This method was found to proceed smoothly in high yield (>95%) and could be performed on preparative scales (>10g). In addition, all of the steps of the synthetic route to the 2-substituted-phenanthroline amino acid were found to be amenable to reaction scale-up. As a result, the 2-phenanthroline amino acid (2Fen) can be prepared readily on a preparative scale with little difficulty.

Conclusions

In conclusion, the synthetic routes to the four heteroaromatic, metal chelating amino acids allow for the facile generation of gram scale quantities in high optical purity, suitable for peptide incorporation using solid phase peptide synthesis protocols. The four amino acids provide a powerful array of metal ligands, each of which is capable of creating a unique metal coordination environment. The accessibility of these amino acid isomers for polypeptide synthesis will aid in the *de novo* design of metalloproteins where correct programming of the metal cation selectivity, coordination geometry of the metal center, and orientation of the metal with respect to the polypeptide backbone is paramount to success. The unique metal binding properties of these ligands, when incorporated into polypeptide sequences, is examined in Chapter 5.

Acknowledgment

The work in this chapter was performed in collaboration with Thomas J. Prins (M.S., Caltech, 1992) and Professor Barbara Imperiali. Tom was responsible for synthesis of the 4-bipyridyl-substituted amino acid and helped in the preparation of the 5-substituted analog. Barbara worked out the synthetic procedures for the 6-bipyridyl substituted amino acid, and developed the protocols for the enzymatic resolution of the amino acid enantiomers. The bulk of the work in this chapter appeared in two publications:

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.

Imperiali, B.; Prins, T.J.; Fisher, S.L. "Chemoenzymatic Synthesis of 2-Amino-3-(2,2'-bipyridinyl)propanoic Acids,"*J. Org. Chem.* **1993**, *58*, 1613-1616.

Experimental

All reagents were obtained from Aldrich Chemical Co. and used without further purification. The enzymes, Acylase I, (*Asperigillus melleus*, EC 3.5.1.14, 0.45 units/mg) and Protease Type VIII (Subtilisin Carlsberg, Bacterial, CA#[9014-01-1], 11.6 units/mg) were purchased from Sigma Chem. Co., and stored at 4°C. All synthetic intermediates were stored at 4°C after purification.

Determination of Enantiomeric Purity of Amino Acids

2-Amino-3-(2,2'-bipyrid-6-yl)propanoic Acid (**1**)

The asymmetric induction in the phase transfer catalyzed reaction was determined by derivatizing the unprotected amino acid as the methyl ester (by treatment with excess

diazomethane) followed by conversion to the Mosher's amide[20] (by reaction with excess (*R*)-α-methoxy-α-trifluoromethylphenyl acetyl chloride). Purification of the final product was achieved by preparative thin layer chromatography (eluent, chloroform:methanol, 10:1). HPLC analysis to evaluate the enantiomeric excess of crude and recrystallized products was carried out on a DNBPG covalent column, 5mm (4.6 x 250 mm) eluent 3% *i*-PrOH/hexanes, 1.2 ml/min.

CrownPak CR Column Analysis

Characterization was obtained through HPLC analysis using a CrownPak CR (-) or a CrownPak CR (+) column (J.T. Baker), on an Isco Model 2350 HPLC equipped with an Isco Model V[4] Absorbance detector using the following conditions:

2-Amino-3-(2,2'-bipyrid-6-yl)propanoic acids: (Crownpak CR (-)) Flow rate 0.9 mL/min, 25 °C, 285 nm, eluent 12% MeOH/ 1 M $HClO_4$: Amino Acid: L 11.1 min; D 17.15 min.

2-Amino-3-(2,2'-bipyrid-4-yl)propanoic acids:(Crownpak CR (+)) Flow rate 0.7 mL/min, 25 °C, 285 nm, eluent 0.1 M $HClO_4$: Amino Acids: D 8.15 min; L 9.73 min.

2-Amino-3-(2,2'-bipyrid-5-yl)propanoic acids :(Crownpak CR (+)) Flow rate 0.7 mL/min, 25 °C, 285 nm, eluent 0.1 M $HClO_4$: Amino Acids: D 9.55 min; L 11.92 min.

2-Amino-3-(1,10-phenanthrol-2-yl)propanoic acids and methyl esters:(Crownpak CR (+)) Acids: Flow rate 1.0 mL/min, 25 °C, 278 nm, eluent 10% MeOH / 0.1 M $HClO_4$: Methyl esters: D 11.50 min, L 12.25 min; Amino Acids: D 7.75 min, L 8.25 min.

N-(Diphenylmethylene)-2-amino-3-(2,2'-bipyrid-6-yl)propanoic Acid *tert*-Butyl Ester (**7**) *N*-(Diphenylmethylene) glycine *tert*-butyl ester (**5**) (1.3g, 4.4 mmol), *N*-benzylcinchonidinum chloride (0.31g, 0.736 mmol), and **6** (0.95g, 3.68 mmol) were added to a suspension of 20 mL CH_2Cl_2 and 7.04 mL 50% aqueous NaOH, and stirred vigorously for 1hr at room temperature at which time the reaction was complete by TLC. The organic phase was separated and the aqueous phase was extracted with CH_2Cl_2 (20 mL). The combined organic phases were concentrated *in vacuo*. The residue was redissolved in ether (40 mL) and water (20 mL). The organic phase was separated and washed with water (2 x 10 mL), dried (Na_2SO_4), and concentrated to afford the crude product. The product was purified by flash chromatography (eluent hexane:ethyl acetate, 8:1.5 containing 0.5% triethylamine - addition of the triethylamine was found to be necessary in order to avoid decomposition of the acid labile imine product on silica) to yield 1.37 g (80%) of pure **7**. At this stage, 65 mg of material was removed for stereochemical analysis. The product was crystallized from hexane at 4°C. The solid product (0.69g) was removed and the filtrate concentrated to afford an oil (0.58g). The chemical yield of optically pure material is approximately 40%, mp (DL racemate) 90-91 °C, (L) oil; $[\alpha]_D^{25}$ = (-) 327° (c = 1, EtOAc); ^{1}H NMR ($CDCl_3$) δ:1.46 (s, 9H), 3.40(dd, 1H, J = 9.2 and 13.4 Hz), 3.52 (dd, 1H, J = 4.3 and 13.4 Hz), 4.63 (dd, 1H, J = 4.3 and 9.2 Hz), 7.16 (m, 4H), 7.27 (m, 5H), 7.32 (t, 1H, J = 6.1 Hz), 7.52 (d, 2H, J = 8.1 Hz), 7.65 (t, 1H, J = 7.9 Hz), 7.69 (dd, 1H, J = 1.5 and 7.6 Hz), 8.09 (d, 1H, J = 7.9 Hz), 8.18 (d, 1H, J = 7.9 Hz), 8.64 (dd, 1H, J = 1.5 and 3.5 Hz); ^{13}C NMR ($CDCl_3$) δ: 28.0, 41.8, 66.5, 81.1, 118.4, 121.4, 123.4, 124.4, 127.8, 128.0, 128.1, 128.7, 130.0, 136.3, 136.5, 136.9, 139.6, 148.9, 155.4, 156.2, 158.0, 170.6, 170.9, 203.3. Anal. Calcd for $C_{30}H_{29}N_3O_2$: C, 77.72; H, 6.30; N, 9.06. Found: C, 78.08; H, 6.15; N, 9.09.

2-Amino-3-(2,2'-bipyrid-6-yl)propanoic Acid (**1**) A suspension of optically pure **7** (0.5g) was refluxed in 6N HCl (10 mL) for 4 hr. The hydrolyzed reaction mixture was then cooled, extracted with ether (3 x 5 mL) and concentrated to dryness. The residual material was lyophilized several times from water to afford (0.3g, 100%) of the amino acid hydrochloride **1**. mp 220 °C (decomp); $[\alpha]_D^{25}$ = (-) 18.6° (c = 1, H_2O; pH 7.0), (-) 12.9° (c = 1, 5N HCl); MS (M^+) 244; ^1H NMR (D_2O) δ: 3.30 (dd, 1H, J = 6.6 and 12.5 Hz), 3.35 (dd, 1H, J = 5.0 and 12.5 Hz), 4.09 (dd, 1H, J = 5.0 and 6.6 Hz), 7.29 (d, 1H, J = 7.4 Hz), 7.49 (t, 1H, J = 6.1 Hz), 7.79 (t, 1H, J = 7.6 Hz), 7.82 (t, 1H, J = 7.7 Hz), 7.98 (t, 1H, 7.7 Hz), 8.07 (d, 1H, J = 7.9 Hz), 8.50 (d, 1H, J = 4.1 Hz); ^{13}C NMR (D_2O) δ: 37.8, 54.9, 121.7, 123.9, 126.1, 140.2, 141.5, 148.2, 154.0, 154.6, 157.2, 174.1.

N-(9-Fluorenylmethoxycarbonyl)-2-amino-3-(2,2'-bipyrid-6-yl)propanoic Acid (**8**)

A solution of FMOC-azide (0.2615 g, 0.9862 mmol) in 2.80 mL 1,4-dioxane was added dropwise with stirring to a solution of **1** (0.24 g, 0.99 mmol) in 2.80 mL 10% Na_2CO_3/H_2O at 0°C, over 2h. The reaction was then allowed to warm to room temperature and stirred for 36h. The mixture was diluted with 50 mL distilled H_2O and extracted 3 times with 30 mL 2:1 ethyl acetate:hexane. The aqueous phase was cooled in an ice bath, brought to pH=2 with conc. HCl, and concentrated to 5 mL. The suspension was then centrifuged (3000 rpm) for 20 min. The aqueous phase was decanted and the solid washed with H_2O (2 x 10 mL), centrifuged, and decanted. The solid was taken up in MeOH and concentrated *in vacuo* to yield 0.32g (70%) of white powder; $[\alpha]_D^{25}$ = (-) 61.3° (c = 1, MeOH); HRMS [MH^+]: Calcd for $C_{28}H_{24}N_3O_4$: [466.1767]; Obs. [466.1781]; ^1H NMR (CD_3OD) δ: 3.34 (m, 1H), 3.62 (m, 1H), 4.02 (t, 1H, J=7.2Hz), 4.28 (m, 2H), 5.13 (m, 1H), 7.07 (m, 1H), 7.20 (t, 1H, J=7.4 Hz), 7.26 (m, 1H), 7.32 (t, 1H, J=7.3 Hz), 7.44 (d, 1H, J=7.4 Hz), 7.46 (d, 1H, J=7.4 Hz), 7.59 (t, 1H, J=7.9 Hz), 7.65 (d, 1H, J=7.4 Hz), 7.69 (d, 1H, J=7.5

Hz), 8.00 (t, 1H, J=5.2 Hz), 8.05 (t,1H, J=7.8 Hz), 8.28 (d, 1H, J=7.8 Hz), 8.65 (d, 1H, J=7.5 Hz), 8.69 (d, 1H, J=8.0 Hz), 8.78 (d, 1H, J=4.9 Hz); ^{13}C NMR (CD_3OD) δ: 39.7, 39.8, 48.2, 54.0, 67.8, 120.9, 122.1, 125.4, 126.0, 128.0, 128.1, 128.3, 128.7, 128.8, 128.9, 140.7, 142.4, 143.4, 144.8, 145.0, 145.1, 146.8, 148.3, 148.4, 149.3, 158.8, 158.9, 160.0, 160.1, 173.8, 175.1.

4-Formyl-2,2'-bipyridine (**13**) A slurry of 4-methyl-2,2'-bipyridine[25] (**12**, 7.68g, 45.1 mmol), SeO_2 (7.5g, 68 mmol), and water (1.2 mL, 67 mmol) in 285 mL dioxane was refluxed for 2 hr. After cooling slightly additional SeO_2 (7.5g) and water (1.2 mL) were added and refluxing was continued overnight. The mixture was filtered while still warm and the filtrate was dried onto silica gel (~10cc). The resulting solid mixture was purified using flash column chromotography (eluent : CCl_4/EtOAc 2:1) yielding 3.02 g (36%) of a cream colored solid: mp 84.8-86.0 °C; HRMS [MH^+], calcd. for $C_{11}H_9N_2O$ [185.0715], obsd. [185.0715]; ^1H NMR ($CDCl_3$) δ: 7.37 (m, 1H), 7.73 (m, 1H), 7.86 (m, 1H), 8.44 (d, 1H), 8.73 (d, 1H), 8.83 (s, 1H), 8.90 (d, 1H), 10.18 (s, 1H); ^{13}C NMR ($CDCl_3$) δ: 120.8, 121.0, 121.2, 124.4, 137.1, 142.6, 149.3, 150.4, 154.8, 158.0, 191.7; IR (thin film) cm^{-1} : 2860, 1704, 1585, 1560, 1458, 1446, 1410, 1388, 1267, 1251, 1231, 1172, 1094, 1061, 988, 908, 839, 788, 739, 664, 649; UV (MeOH) λ_{max} = 283 (13,500), 236 (10,900).

4-[(2,2'-Bipyridyl)-4-methylene]-2-phenyl-5(4H)-oxazolone(**15**) 4-formyl-2,2'-bipyridine (**8**, 6.86g, 37.2 mmol), hippuric acid (**14**, 6.86g, 38.3 mmol), and sodium acetate (3.09g, 37.7 mmol) were slurried in acetic anhydride (70 mL). Upon heating to 90 °C, the mixture thickened and more acetic anhydride (28 mL) was added. After stirring at 90 °C for 1 hr, the mixture was cooled to room temperature, and the solid was filtered and washed with 12 mL acetic anhydride. This solid was dried *in vacuo* yielding 12.6 g of tan solid contaminated with sodium acetate. This solid was stirred in 400 mL chloroform and

filtered. The filtrate were concentrated *in vacuo* to yield 9.1 g (75%) of a pale tan solid: mp 179.9-180.5 °C; HRMS [MH^+], calcd. for $C_{20}H_{14}N_3O_2$ [328.1086], obsd. [328.1077]; ^1H NMR ($CDCl_3$) δ : 7.26 (s, 1H), 7.35 (m, 1H), 7.56 (m, 3H), 7.66 (m, 1H), 7.85 (m, 1H), 8.23 (d, 2H), 8.32 (m, 1H), 8.42 (m, 1H), 8.74 (m, 1H), 8.82 (m, 1H), 8.85 (s, 1H); ^{13}C NMR ($CDCl_3$) δ: 121.4, 123.4, 124.3, 124.8, 125.3, 128.4, 129.1, 129.3, 134.3, 137.2, 137.3, 141.3, 149.6, 150.1, 155.8, 157.3, 165.6, 166.8; IR (thin film) cm^{-1} : 1803, 1794, 1654, 1560, 1451, 1397, 1325, 1298, 1230, 1202, 1158, 1112, 1095, 1068, 984, 921, 878, 846, 781, 730, 695, 681, 621; UV (MeOH, nm) λ_{max} = 354 (20,500), 278 (19,300), 238 (26,000).

2-Amino-3-(2,2'-bipyrid-4-yl)propanoic Acid Methyl Ester (**9**)

Reductive cleavage of the oxazolone: The oxazolone **15** (3.0 g, 9.2 mmol) and red phosphorous (1.84 g, 59.4 mmol) were stirred in acetic anhydride (15 mL). Hydriodic acid (47%, 12.8 mL, 70.5 mmol) was added dropwise with stirring over 40 min. After refluxing for 3.5 hr the mixture was allowed to cool to room temperature. Unreacted phosphorous was removed by filtration and washed with 8 mL acetic acid. The filtrate was concentrated under reduced pressure at ~50 °C. Water (10 mL) was added and removed. Water (15 mL) and ether (15 mL) were added to the residue and shaken until all was dissolved. The aqueous layer was isolated and washed with ether (3 x 10 mL) and concentrated *in vacuo* to yield a syrupy residue.

Esterification of the amino acid: Residual water in the above residue was removed by toluene azeotrope. The resulting residue was dissolved in MeOH (39 mL, dried over and distilled from CaH_2). The solution was cooled to 0°C and hydrogen chloride was bubbled through it for ~3 min. After standing overnight, the solution was concentrated *in vacuo*. The residue was rubbed with Et_2O to yield a solid. This solid was stirred with ether for 2 hr, collected by centrifugation, washed with ether, 2 x 25 mL and dried *in vacuo* over KOH. The solid residue (in 1 g portions) was partitioned between chloroform and

0.4 M $NaHCO_3$ (100 mL/100 mL). The organic phase was isolated and the aqueous phase was washed with chloroform, 4 x 100 mL. The combined organic phases were dried over Na_2SO_4 and concentrated *in vacuo* to yield 1.81 g (77%) of a yellow oil; HRMS [MH^+], calcd. for $C_{14}H_{16}N_3O_2$ [258.1243], obsd. [258.1250]; ^1H NMR ($CDCl_3$) δ: 1.69 (s, br, 2H), 2.89 (m, 1H), 3.16 (m, 1H), 3.72 (s, 3H), 3.82 (m, 1H), 7.16 (m, 1H), 7.28 (m, 1H), 7.79 (m, 1H), 8.27 (s, 1H), 8.37 (d, 1H, J=8.0 Hz), 8.58 (d, 1H, J=5.0 Hz), 8.65 (m, 1H); ^{13}C NMR ($CDCl_3$) δ: 40.4, 52.0, 55.0, 121.1, 121.6, 123.6, 124.5, 136.8, 147.5, 148.9, 149.1, 155.8, 156.1, 174.9; IR (thin film) cm^{-1} : 3372, 3308, 3057, 3006, 2950, 1736, 1604, 1585, 1556, 1460, 1438, 1401, 1271, 1201, 1177, 1017, 992, 793, 746; UV (MeOH, nm) λ_{max} = 282 (13,300), 238 (10,600).

2-Amino-3-(2,2'-bipyrid-5-yl)propanoic Acid Methyl Ester (**19**) Toluene (20 mL) was added to the amino acid hydrochloride (1.0 g, 3.2 mmol, 40 % ee) and removed under reduced pressure. The residue was slurried in dry MeOH (15 mL) and cooled to 0°C. Anhydrous hydrogen chloride was passed through the mixture for ~3 min. After standing at room temperature overnight, the solvent was removed under reduced pressure. MeOH (~10 mL) was added and removed. The oily residue was rubbed with Et_2O until the residue formed a fine powder. This powder was isolated by centrifugation and washed twice with Et_2O (20 mL). The residual solid was dried under reduced pressure, transferred to a separatory funnel and then partitioned between $CHCl_3$ (100 mL) and 0.2 M $NaHCO_3$ (100 mL). The organic phase was isolated and the aqueous phase was washed with $CHCl_3$ (3 x 100 mL). The combined organic phases were dried over Na_2SO_4 and concentrated *in vacuo* to yield 591 mg (73 %) of an oil which partially solidified upon cooling overnight; HRMS [MH^+], calcd. for $C_{14}H_{16}N_3O_2$ [258.1250], obsd. [258.1239]; ^1H NMR ($CDCl_3$) δ: 1.87 (s, br, 2H), 2.70 (m, 1H), 2.86 (m, 1H), 3.46 (s, 3H), 3.54 (m, 1H), 7.03 (m, 1H), 7.44 (m, 1H), 7.54 (m, 1H), 8.14 (m, 2H), 8.31 (m, 1H), 8.43 (m, 1H); ^{13}C NMR ($CDCl_3$) δ: 37.3, 51.4, 54.8, 120.1, 120.3, 123.0, 132.5, 136.2, 137.1,

148.5, 149.3, 154.1, 155.2, 174.3; IR (thin film) cm^{-1} : 3368, 3052, 3003, 2949, 1736, 1588, 1573, 1555, 1460, 1436, 1200, 1025, 850, 796, 751; UV (MeOH) λ_{max} = 286 (17,000), 240 (12,600).

N-(9-Fluorenylmethoxycarbonyl)-(*S*)-α-amino-(2,2'-bipyrid-4 and 5-yl)-propanoic acids (**11**, **20**) and *N*-(9-Fluorenylmethoxy)carbonyl)-(*S*)-2-amino-3-(1,10-phenanthrol-2-yl)-propanoic acid(**25**)

General method for resolution of enantiomers and isolation of Bipyridyl-substituted L-amino acids: A solution of the protease (30 mg, 200 - 500 units) in 120 mL of 0.2 M $NaHCO_3$ was added to the amino acid methyl ester (**9** or **19**, 8 mmol) in a 250 mL Erlenmeyer flask. The mixture was rotated at 120-200 rpm on an orbital shaker and monitored by HPLC at 20 min. intervals. When the ratio of the D-amino acid methyl ester to the L-amino acid methyl ester was >50:1 (about 2 hours), the mixture was extracted with chloroform, 6 x 110 mL. The aqueous phase was reduced in volume to remove any chloroform and then lyophilized yielding a mixture of amino acid and carbonate salts. The combined organic layers were dried over Na_2SO_4 and the solvent was evaporated. HPLC analysis of the aqueous phase generally shows a 92-94% ee of the L-amino acid (**2** or **3**), while analysis of the organic residue generally showed 98% ee of the D-amino acid methyl ester.

Method for Enzymatic Resolution of 2-Phenanthrolyl-substituted amino acid: A solution of the racemic methyl ester (**24**, 3.6587g, 16.0 mmol) was dissolved in 45 mL *t*-butanol and added to 445 mL 10% $NaHCO_3$. A sample of alkaline protease (125 mg, 800 - 2000 units) in 120 mL of 0.2 M $NaHCO_3$ was added to the amino acid methyl ester (8 mmol) in a 1L Erlenmeyer flask. The mixture was rotated at 120-200 rpm and monitored by HPLC at 20 minute intervals. When the ratio of the D-amino acid methyl ester to the L-amino acid methyl ester was >50:1 (about 40 min.), the mixture was extracted with chloroform, 6 x 150 mL. The aqueous phase was reduced in volume to remove any

chloroform and then lyophilized yielding a mixture of amino acid and carbonate salts. The combined organic layers were dried over Na_2SO_4 and the solvent was evaporated. HPLC analysis of the aqueous phase generally shows a 92-94% ee of the L-amino acid, while analysis of the organic residue generally showed 87% ee of the D-amino acid methyl ester.

General method for amine protection of the L-amino acid following enzymatic resolution: The amino acid/carbonate salt mixtures from above were dissolved in 25 mL of 10% Na_2CO_3. 9-Fluorenylmethyl succinimidyl carbonate (1.2 eq) was dissolved in 15 mL dioxane and added dropwise to the amino acid solution. The reaction mixture was shaken periodically for 1.5 hours and then transferred to a separatory funnel and diluted with 100 mL water. This mixture was washed with ether, 4 x 60 mL, and transferred to a 250 mL Erlenmeyer flask. After cooling to 0°C and adjusting to pH <2 with conc. HCl, the precipitate was isolated (by centrifugation if necessary) and washed with water, 3 x 100 mL. The residue was transferred to a round bottom flask with methanol (~200 mL). The methanol was removed under reduced pressure and replaced with toluene (~80 mL). The resulting toluene/water mixture was azeotroped under reduced pressure. The residue was rubbed with ether (if necessary) to yield a solid which was dried *in vacuo* to yield approximately 85% (yield based on the amount of L-amino acid methyl ester) of the FMOC-protected amino acid.

N-(9-Fluorenylmethoxycarbonyl)-(*S*)-2-amino-3-(2,2'-bipyrid-4-yl) propanoic acid (**11**)

mp ~167°C dec. (over broad range); HRMS [MH^+], calcd. for $C_{28}H_{24}N_3O_4$ [466.1779], obsd. [466.1783]; ^1H NMR (DMSO-d_6) δ: 3.03 (m, 1H), 3.24 (m, 1H), 4.15 (m, 3H), 4.35 (m, 1H), 7.17 (m, 2H), 7.33 (m, 2H), 7.47 (m, 2H), 7.55 (m, 2H), 7.81 (d, 2H), 7.88 (m, 1H), 7.97 (m, 1H), 8.39 (d, br, 1H), 8.42 (s, 1H), 8.59 (d, br, 1H), 8.66 (d, br, 1H); ^{13}C NMR (DMSO-d_6) δ: 25.2, 36.0, 46.5, 54.5, 65.6, 120.1, 121.1, 122.0, 124.7, 125.1, 125.2, 125.4, 127.0, 127.6, 138.0, 140.6, 140.6, 143.6, 143.7, 148.2, 148.9, 149.9, 153.7, 156.0, 172.7; IR (thin film) cm^{-1} : 3315, 3060, 2957, 1731, 1715,

1706, 1696, 1540, 1449, 1255, 1219, 1182, 1152, 1104, 1079, 1046, 1018, 994, 792, 758, 737; $[\alpha]^{28}_D$ -4.7° (c = 1, DMSO); UV (MeOH) λ_{max} = 299 (7300), 265 (19,900).

N-(9-Fluorenylmethoxycarbonyl)-(*S*)-2-amino-3-(2,2'-bipyridin-5-yl)propanoic Acid (**20**) mp 117°C dec.; HRMS [MH^+], calcd. for $C_{28}H_{24}N_3O_4$ [466.1767], obsd. [466.1779]; ^1H NMR (DMSO-*d*$_6$) δ: 3.01 (m, 1H), 3.24 (m, 1H), 4.17 (m, 3H), 4.32 (m, 1H), 7.25 (m, 4H), 7.58 (m, 2H), 7.65 (m, 1H), 7.84 (m, 3H), 8.07 (m, 1H), 8.18 (m, 1H), 8.44 (d, 1H), 8.49 (d, 1H), 8.70 (s, 1H), 8.75 (d, br, 1H); ^{13}C NMR (CD_3OD) δ: 35.5, 55.7, 68.0, 120.9, 123.5, 124.0, 126.1, 127.3, 128.1, 128.7, 138.4, 142.4, 142.7, 143.4, 145.0, 147.2, 148.8, 149.3, 150.6, 158.3, 174.0; IR (thin film) cm^{-1} : 1711, 1531, 1448, 1246, 1045, 759, 739; $[\alpha]^{26}_D$ -1.4° (c = 0.5, MeOH); UV (MeOH, nm) λ_{max} = 299 (12,500), 288 (16,900), 265 (22,400), 228 (10,700).

N-(9-Fluorenylmethoxycarbonyl)-(*S*)-2-amino-3-(1,10-phenanthrol-2-yl)propanoic Acid (**25**) mp = 187°C (dec.); HRMS $[MH]^+$, calcd. for $C_{30}H_{24}N_3O_4$ [490.1767]; Obs. [490.1749]; TLC (SiO_2; 4:1 $CHCl_3$:MeOH; UV); R_f = 0.55; ^1H NMR (DMSO-*d6*) δ: 3.45 (m, 1H), 3.65 (m, 1H), 4.10 (m, 3H), 4.40 (m, 1H), 4.80 (d, 2H, J=5.8 Hz), 7.05 (m, 2H), 7.35 (m, 2H), 7.50 (m, 2H), 7.68 (d, 1H, J=7.2Hz), 7.80 (d, 1H, J=7.6 Hz), 7.93 (m, 2H), 8.10 (m, 2H), 8.64 (d, 1H, J=8.3Hz), 9.00 (d, 1H, J=4.6 Hz), 9.25 (d, 1H, J=4.85 Hz); ^{13}C NMR (DMSO-*d6*) δ: 47.5, 55.1, 66.6, 121.0, 124.3, 125.0, 126.0, 126.1, 127.2, 127.3, 127.4, 128.0, 128.6, 128.8, 129.6, 137.2, 137.6, 141.6, 144.7, 144.8, 150.9, 159.4, 174.2; IR (mineral oil mull) cm^{-1} : 3307, 1690, 1601, 1537, 1454, 1372, 1337, 1261, 1150, 1102, 1078, 1037, 850, 761, 732; $[\alpha]^{25}_D$ -67.9° (c = 0.35, DMSO); UV (MeOH, nm) λ_{max} = 299 (11,000), 268 (32,000), 224 (38,900).

5-(Bromomethyl)-2,2'-Bipyridine (**17**) To a solution of 5-(ethoxycarbonyl)-2,2'-bipyridine[12] (**16**, 2.25 g, 9.9 mmol) in anhydrous THF (50 mL) at -78 °C was added 10 mL (10 mmol) of a 1M solution of lithium aluminum hydride in THF. The reaction mixture was allowed to warm to -20°C at which temperature the solution became completely homogenous. The reaction was stirred at this temperature for 0.5 hr and then cooled back down to -78 °C and quenched slowly with 30 mL 10% aqueous THF. After warming to room temperature, the reaction mixture was stirred with dry Celite for 15 minutes and then filtered. The solution was then concentrated under reduced pressure to afford 1.8 g of 5-(hydroxymethyl)-2,2'-bipyridine as a thick orange oil which was used in the subsequent reaction without purification.

The product from the preceding reaction was dissolved in dichloromethane (40 mL) and cooled to 0 °C. To this was added triphenylphosphine (2.75 g, 10.5 mmol) and N-bromosuccinimide (1.95 g, 11 mmol). After stirring for 0.5 hr, the reaction mixture was concentrated to approximately one quarter of the original volume and applied directly to a flash silica gel column (eluent: hexanes/diethylether 1:1). Concentration of the pure fractions provided 1.85g (75% yield over two step) of 5-(bromomethyl)-2,2'-bipyridine as a pale yellow solid. mp : 70.8-71.3°C; HRMS [MH^+]: Calcd for $C_{11}H_9N_2Br$: [247.9949]; Obs. [247.9952]; ^1H NMR ($CDCl_3$) δ: 4.52 (s, 2H); 7.29 (dd, 1H, J = 4.8, 7.4 Hz); 7.83 (m, 2H); 8.39 (d, 2H, J = 8.1 Hz); 8.67 (d, 2H, J = 2.3 Hz); ^{13}C NMR ($CDCl_3$) δ: 29.9, 121.2, 121.5, 124.2, 133.9, 137.2, 137.8, 149.4, 149.5, 155.7, 156.2; IR (thin film) cm^{-1} : 610, 650, 750, 798, 834, 858, 992, 1025, 1039, 1061, 1127, 1203, 1253, 1392, 1434, 1461, 1556, 1574, 1597, 1713, 2967, 3050; UV : (MeOH, nm) λ_{max} = 289 (26,700), 244 (18,600).

N-(Diphenylmethylene)-2-amino-3-(2,2'-bipyrid-5-yl)propanoic Acid *tert*-Butyl Ester (**18**) A solution of *N*-(diphenylmethylene)glycine *tert*-butyl ester (**5**) (150 mg, 0.5 mmol) and (8*S*,9*R*)-(-)-*N*-benzylcinchonidinium chloride (21 mg, 0.05 mmol) was dissolved in

dichloromethane (10 mL) and cooled to 0 °C. To this was added 0.8 mL 50% aqueous sodium hydroxide and 5-(bromomethyl)-2,2'-bipyridine (**17**, 160 mg, 0.64 mmol) in solution in dichloromethane (2 mL). The reaction mixture was then stirred at 0 °C for 24 h. At this time the organic phase was separated and the remaining aqueous phase washed with an additional aliquot (5 mL) of dichloromethane. The pooled organic phases were then concentrated under reduced pressure and purified by flash chromatography (eluent hexane/ethylacetate 5:1 containing 0.5% triethylamine) to afford 200 mg (85% yield) of pure product. The optical purity of the product was determined to be 55% ee. HRMS [MH^+] : Calcd for $C_{30}H_{30}N_3O_2N$: [464.2338]; Obs. [464.2333]; ^1H NMR ($CDCl_3$) δ: 1.45 (s, 9H); 3.26 (m, 2H); 4.19 (dd, 1H, J = 5.34, 8.09 Hz); 6.78 (d, 2H, J = 6.64 Hz); 7.2-7.6 (m, 9H); 7.77 (t, 1H, J = 7.9 Hz); 8.23 (d, 1H, J = 8.1 Hz); 8.31 (d, 1H, J = 8.0 Hz); 8.42 (d, 1H, J = 9.5 Hz); 8.64 (d, 1H, J = 4.1 Hz); ^{13}C NMR ($CDCl_3$) δ: 28.3, 36.9, 67.4, 81.7, 120.7, 121.2, 123.7, 127.7, 128.3, 128.5, 128.7, 129.0, 130.6, 134.4, 136.4, 137.1, 138.5, 139.4, 149.3, 150.6, 154.4, 170.6, 171.1; IR (thin film) cm^{-1} : 650, 668, 696, 751, 849, 1029, 1150, 1220, 1253, 1288, 1368, 1392, 1458, 1558, 1575, 1589, 1622, 1732, 2933, 2977, 3056; UV: (MeOH, nm) λ_{max} = 243 (18,900); 285 (18,050).

2-Amino-3-(2,2'-bipyrid-5-yl)-propionic Acid (**19**) A suspension of pure N-(diphenylmethylene)-2-amino-(2,2'-bipyridin-5-yl)propanoic acid *tert*-butyl ester (**18**, 1.2g, 2.6 mmol) was refluxed in 6N hydrochloric acid (30 mL) for 4h. The hydrolyzed reaction mixture was then cooled, extracted with ether, (3 x 15 mL), and concentrated to dryness. The residual material was lyophilized several times from water to afford 0.72g (100%) of the amino acid hydrochloride: mp : 153 °C (dec.); HRMS [MH^+]: Calcd for $C_{13}H_{13}N_3O_2$: [244.1086]; Obs. [244.1098]; ^1H NMR (D_2O) δ: 3.33 (t, 2H, J = 7 Hz); 4.37 (t, 1H, J = 6.8 Hz); 7.86 (t, 1H, J = 6.8 Hz); 7.86 (t, 1H, J = 5.8 Hz); 8.08 (d, 1H, J = 7 Hz); 8.15 (d, 1H, J = 8.15 Hz); 8.3-8.5 (m, 2H); 8.61 (s, 1H); 8.66 (d, 1H, J = 5.6

Hz); ^{13}C NMR (D_2O) δ: 32.9, 53.4, 123.5, 124.5, 127.6, 134.7, 141.6, 143.2, 145.8, 146.8, 147.2, 149.8, 170.6; UV: (H_2O, nm) λ_{max} = 238 (8440); 287 (9960).

2-Acetamido-2-ethoxycarbonyl-3-(1,10-phenanthrol-2-yl)propanoic Acid Ethyl Ester (**23**) : A portion of diethyl acetamidomalonate (**22**, 0.6074 g, 2.80 mmol) was added to a solution of sodium (0.0641 g, 2.80 mmol) in 3 mL absolute ethanol and stirred for 30 min. A solution of 2-chloromethyl-1,10-phenanthroline[38] (**21**, 0.6710g, 2.93 mmol) in 3 mL absolute ethanol was added dropwise to the mixture. After 8 hr at reflux, the mixture was cooled, concentrated to a thick red oil, and suspended in water (50 mL), and extracted with $CHCl_3$ (3 x 50 mL). The organic fractions were washed with 50 mL each of water, 10% Na_2CO_3, and brine, dried over Na_2SO_4, and concentrated to an oily red solid. The crude product was judged to be >95% pure by TLC and NMR and was taken on to the next step without further purification. An analytical sample was purified using a short neutral Al_2O_3 column (2:1 CH_2Cl_2:EtOAc) to yield a pale yellow solid; mp: 112.4-113.2°C; TLC (Al_2O_3; 2:1 CH_2Cl_2:EtoAc, UV detection) R_f = 0.43; HRMS Calc. for [MH^+] $C_{22}H_{24}N_3O_5$ [410.1715]; Obs. [410.1694]; ^1H NMR (CD_3OD) δ: 1.20 (m, 6H), 2.01 (s, 3H), 4.10 (s, 2H), 4.24 (m, 4H), 7.62 (d, 2H, J=8.3Hz), 7.92 (m, 2H), 8.01 (d, 2H, J=1.8 Hz), 8.42 (d, 2H, J=8.3 Hz), 8.70 (d, 2H, J=8.2 Hz), 9.15 (d, 2H, J=3.1 Hz); ^{13}C NMR (CD_3OD) δ: 13.4, 21.5, 41.3, 62.9, 124.2, 125.9, 126.3, 127.8, 137.5, 140.1, 148.2, 171.9, 175.9; IR (thin film) cm^{-1} : 3377, 3268, 1660, 1507, 1484, 1396, 1366, 1290, 1202, 1090, 1055, 1014, 855.

Racemic 2-amino-3-(1,10-phenanthrol-2-yl)propanoic Acid Methyl Ester (**24**) : A solution of **23** (1.1463 g, 2.8 mmol) in 100 mL 6N HCl was heated to 90 °C with stirring for 10 hr. The reaction was cooled and concentrated *in vacuo* to a thick red oil. The oil was then resuspended in water (50 mL) and concentrated *in vacuo* twice to remove excess HCl, after which the residue was dissolved in 20 mL water and brought

to pH=9.5 with concentrated NaOH solution. The mixture was then washed with $CHCl_3$ (3 x 50 mL) and concentrated *in vacuo* to yield a white solid. The mixture of amino acid and sodium salts was dried over P_2O_5 under vacuum overnight. The solid mixture was then suspended in 50 ml of freshly distilled MeOH, and anhydrous HCl was passed through the mixture for 10 min. After stirring overnight the reaction was complete by chiral-HPLC analysis, the solid suspension was concentrated *in vacuo*. The solid was resuspended in MeOH (100 mL) and concentrated *in vacuo* twice to remove excess acid. The solid residue was carefully neutralized with saturated Na_2CO_3 and extracted with $CHCl_3$ (3 x 50mL). The organic fractions were combined, dried over Na_2SO_4, and concentrated to yield a yellow oil (0.5337 g, 83%). HRMS : Calc. for $[MH]^+$ $C_{15}H_{14}N_3O_2$ [268.1086]; Obs. [268.1101]; ^1H NMR ($CDCl_3$) δ: 3.43 (m, 1H), 3.72 (m, 4H), 4.26 (m, 1H), 7.53 (m, 2H), 7.67 (s, 2H), 8.08 (d, 1H, J=8.1 Hz), 8.15 (d, 1H, J=8.0 Hz), 9.08 (d, 1H, J=3.8 Hz); ^{13}C NMR (CD_3OD) δ: 43.8, 52.8, 55.3, 119.5, 123.4, 124.3, 126.6, 126.9, 127.8, 129.3, 136.5, 137.0, 146.3, 150.7, 159.1, 175.6; IR (thin film) cm^{-1} : 855, 1015, 1139, 1206, 1277, 1394, 1454, 1494, 1506, 1556, 1590, 1732, 2944, 3373.

References

(1) Tainer, J.A.; Roberts, V.A.; Getzoff, E.D. "Metal Binding Sites in Proteins," *Curr. Opin. Biotechnology* **1991**, *2*, 582-591.

(2) 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.

(3) Micheal, S.F.; Kilfoil, V.J.; Schmidt, M.H.; Amann, B.T.; Berg, J.M. "Metal Binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide," *Proc. Natl. Acad. Sci., USA* **1992**, *89*, 4796.

(4) Hughes, M.N. *The Inorganic Chemistry of Biological Processes*; John Wiley and Sons: New York, **1981**.

(5) Otsuka, S.; Yamanaka, T. *Metalloproteins: Chemical Properties and Biological Effects*; Elsevier: New York, **1988**.

(6) Hancock, R.D.; Martell, A.E. "Ligand Design for Selective Complexation of Metal Ions in Aqueous Solution," *Chem. Rev.* **1989**, *89*, 1875.

(7) Kazmierski, W.M. "Metal Chelating Amino Acids in the Design of Peptides and Proteins. Synthesis of N^{α}-FMOC/Bu^{t} Protected Amino Acids Incorporating Aminodiacetic Acid Moiety," *Tet. Let.* **1993**, *34*, 4493-4496.

(8) Rana, T.M.; Ban, M.; Hearst, J.E. "Synthesis of a Metal-Ligating Amino Acid Suitable for Solid Phase Assembly of Peptides," *Tet. Let.* **1992**, *33*, 4521-4524.

(9) Ruan, F.; Chen, Y.; Hopkins, P.B. "Metal Ion Enhanced Helicity in Synthetic Peptides Containing Unnatural, Metal-Ligating Residues," *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404.

(10) Constable, E.C. "Helices, Supramolecular Chemistry, and Metal-directed Self Assembly," *Angew. Chem. Int. Ed., Engl.* **1991**, *30*, 1450-1451.

(11) Koert, U.; Harding, M.M.; Lehn, J.-M. "DNH deoxyribonucleohelicates self assembly of oligonucleosidic double-helical metal complexes," *Nature* **1990**, *346*, 339-342.

(12) Ghadiri, M.R.; Soares, C.; Choi, C. "A Convergent Approach to Protein Design: Metal Ion Assisted Spontaneous Self-Assembly of a Polypeptide into a Triple Helix Bundle Protein," *J. Am. Chem. Soc.* **1992**, *114*, 825-831.

(13) Lieberman, M.; Sasaki, T. "Iron(II) Organizes a Synthetic Peptide into Three-Helix Bundles," *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.

(14) Reedjik, J. *Comprehensive Coordination Chemistry*; Wilkinson, G.; Gillard, R.D.; McCleverty, J.A., Eds.; Pergamon Press: New York, **1987**; pp 73-98.

(15) McKenzie, E.D. "The Steric Effect in Bis(2,2'-Bipyridyl) and Bis(1,10-Phenathroline) Metal Compounds," *Coord. Chem. Rev.* **1971**, *6*, 187-216.

(16) Smith, R.M.; Martell, A.E. *Critical Stability Constants, Vol. 2, Amines*; Plenum Press: New York, **1975**; pp 235-262.

(17) Newkome, G.R.; Gupta, V.K.; Fronczek, F.R. "Molecular Tweezer-Type. A Trigonal-Bipyramidal Copper(II) [5.8.5] Complex: Synthesis and Single-Crystal X-ray Structure Determination," *Inorg. Chem.* **1983**, *22*, 171.

(18) O'Donnell, M.J.; Bennett, W.D.; Wu, S. "The Stereoselective Synthesis of α-Amino Acids by Phase Transfer Catalysts," *J. Am. Chem. Soc.* **1989**, *111*, 2353.

(19) Triethylamine (0.5% v/v) was used in all eluents to avoid product decomposition on the silica gel medium.

(20) Dale, J.A.; Mosher, H.S. "Nuclear Magnetic Resonance Enantiomer Reagents. Configurational Correlations *via* Nuclear Magnetic Resonance Chemical Shifts of Diastereomeric Mandelate, *O*-Methylmandelate, and α-Methoxy-α-trifluoromethylphenylacetate (MTPA) Esters," *J. Am. Chem. Soc.* **1973**, *95*, 512-519.

(21) Greenstein, J.P.; Winitz, M. *Chemistry of the Amino Acids*; John Wiley and Sons: New York, **1961**; p 85.

(22) Carpino, L.A.; Han, G.Y. "The 9-Fluorenylmethoxycarbonyl Amino-Protecting Group," *J. Org. Chem.* **1972**, *37*, 3404-3408.

(23) Kauffman, T.; Konig, J.; Woltermann, A. *Chem. Ber.* **1976**, *109*, 3864.

(24) Uenishi, J.; Tanaka, T.; Nishiwaki, K.; Wakabayashi, S.; Oae, S.; Tsukube, H. "Synthesis of ω-(Bromomethyl)bipyridines and related ω,ω-(Bromomethyl)pyridine Heteroaromatics: Useful Functional Tools for Ligands in Host Molecules," *J. Org. Chem.* **1993**, *58*, 4382-4388.

(25) Prins, T.J., M.S. Thesis, "Synthesis of Unnatural Amino Acids with Chelating Side Chains and the Characterization of Peptides Incorporating Them," California Institute of Technology, **1992**.

(26) Williams, R.M. *Synthesis of Optically-Active α-Amino Acids*; Pergamon Press: New York, **1989**; pp 257-280.

(27) Chenault, H.K.; Dahmer, J.; Whitesides, G.M. "Kinetic Resolution of Unnatural and Rarely Ocurring Amino Acids: Enantioselective Hydrolysis of *N*-Acyl Amino Acids Catalyzed by Acylase I," *J. Am. Chem. Soc.* **1989**, *111*, 6354-6364.

(28) Chen, S.-T.; Wang, K.T.; Wong, C.-H. "Chirally Selective Hydrolysis of D,L-Amino Acid Esters by Alkaline Protease," *J. Chem. Soc. Chem. Comm.* **1986**, 1514.

(29) Kraut, J.; Robertus, J.D.; Birktoft, J.J.; Alden, R.A.; Wilcox, P.E.; Powers, J.C. "The Aromatic Substrate Binding Site in Subtilisin BPN and its Resemblance to Chymotrypsin," *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *36*, 117-123.

(30) Alden, R.A.; Birktoft, J.J.; Kraut, J.; Robertus, J.D.; Wright, C.S. "Atomic Coordinates for Subtilisin BPN (or Novo)," *Biochim. Biophys. Res. Comm.* **1971**, *45*, 337.

(31) The resolved amino acids were contaminated with carbonate salts.

(32) del. Milton, R.C.; Becker, E.; Milton, S.C.F.; Baxter, J.E.J.; Elsworth, J.F. "Improved Purities for FMOC-Amino Acids From FMOC-ONSu," *Int. J. Pept. Protein Res.* **1987**, *30*, 431-432.

(33) Erlenmeyer, E. *Ann.* **1893**, *275*, 1.

(34) Blatt, A.H. *Organic Syntheses*; John Wiley and Sons: New York, 1943; pp 489.

(35) Huang, T.L.; Brewer, D.G. "An investigation of the ^{1}H nmr isotropic shifts for some methyl-substituted bipyridine complexes with Ni(II) and Co(II)," *Can. J. Chem.* **1981**, *59*, 1689.

(36) Personal communication with R. Singha Roy.

(37) Albertson, N.F. "The Synthesis of Amino Acids from Ethyl Acetamidomalonate and Ethyl Acetamidocyanoacetate. III. The Use of Primary Halides," *J. Am. Chem. Soc.* **1946**, *68*, 450-453.

(38) Weijnen, J.G.; Koudijs, A.; Schellekens, G.A.; Engbersen, J.F.J. "Functionalized 1,10-Phenanthroline Metallocatalysts as Models for Hydrolytic Metalloenzymes," *J. Chem. Soc., Perkin Trans. 2* **1992**, 829-834.

Chapter 5. Characterization of the Metal-Binding Properties of Peptides Incorporating Unnatural Amino Acids with 2,2'-Bipyridyl and 1,10-Phenanthrolyl Side Chains

Introduction

The design of metalloproteins requires a combined knowledge of the factors that contribute to the stabilization of protein structure and those that govern effective metal ligation and coordination geometry preferences. This is especially pertinent in the construction of catalytic metal centers, as the metal binding sites in these proteins are often characterized by distorted geometries which are dictated by the overall protein structure.[1,2] Even in the more simple structural metal binding sites, where the metal cation can assume a thermodynamically stable coordination geometry, the choice and orientation of the metal ligating groups must be carefully controlled to promote effective metal cation binding.[1,3] Factors such as the size of the metal cation, the polarizabilities of the ligands and metal cation, and the relative spatial distribution of the ligating groups all play important roles in metal binding and are primarily responsible for the remarkable metal cation selectivities that are observed in naturally-occurring metalloproteins.[4] The functional groups commonly found in the metal-binding sites of native proteins is largely limited to just four classes; imidazole, thiol, carboxylate, and main chain carbonyl. Despite this limitation, a wide variety of complex structures constructed from these moieties are known and the binding properties of each of these functional groups has been studied in detail. [5-8]

In an effort to expand the repertoire of metal binding ligands available for metalloprotein design, a series of novel, metal-cation binding amino acids based on the well-known bipyridine and phenanthroline ligands were prepared (see Chapter 4). These ligands were anticipated to be complementary to the ribosomally encoded metal-binding amino acids due to the bidentate N-N' ligating atoms and the electronic π-accepting properties of the heteroaromatic ring systems.[9] In addition to providing enantiomerically pure products, the synthetic routes included steps to afford suitable N^{α}-protection for incorporation into larger peptide sequences using solid phase peptide synthesis. The focus of this chapter is to examine the overall metal binding properties of these ligands when they are incorporated into polypeptide sequences. The studies described herein provide a survey

of the overall binding properties of these novel ligands and establish a foundation for the assembly of larger, functional metalloprotein constructs. As mentioned in Chapter 4, the four amino acids (**1-4**) are expected to have widely different metal binding affinities and coordination geometry preferences. For example, the two amino acids with substitutions *ortho* to the chelating nitrogen atoms, 2Fen and 6Bpa, are expected to have reduced binding affinities and more restricted coordination geometries than the 4- and 5-substituted bipyridine amino acids which have unobstructed N-N' chelation moieties.[9,10] In order to delineate these effects, a variety of peptides were synthesized which incorporate these amino acids and the relative metal-binding properties were analyzed using UV/Vis absorption and circular dichroism spectroscopies.

The first set of peptides (**1-4**) was designed to measure the relative binding affinities of each of the respective ligands when incorporated into a polypeptide backbone in the absence of any other strong metal binding ligands, as well as provide a means to determine the relative propensity for these ligands to undergo intermolecular metal cation ligation.

Ac-Xaa-Thr-Pro-D-Ala-Val-Phe-NH_2

Xaa = 4Bpa (**1**)
5Bpa (**2**)
6Bpa (**3**)
2Fen (**4**)

The primary sequence includes a central Pro-D-amino acid dipeptide core, which is known to induce β-turn formation (see Chapter 2). The central tetrapeptide, Thr-Pro-D-Ala-Val has been conformationally analyzed in this lab,[11] and it was found to have a limited amount of Type II reverse turn character in aqueous solution. This sequence was used to induce a moderate amount of steric bulk near the metal binding ligand to more fully mimic a larger protein structure. In addition, the same tetrapeptide sequence was used in the design

of a peptide predicted to have enhanced intramolecular metal binding capabilities; Ac-6Bpa-Thr-Pro-D-Ala-Val-6Bpa-NH_2 (**5**). Since the two bipyridyl-amino acids are juxtaposed on either side of the turn sequence, it was anticipated that the two ligands would be brought into close proximity and therefore provide a chelating effect. These effects were investigated further with another set of peptides, Ac-2Fen-Thr-Pro-D-Ala-Val-His-NH_2 (**6**), Ac-2Fen-Val-Pro-D-Ser-Phe-His-NH_2 (**7**), and Ac-6Bpa-Thr-Pro-D-Ala-Val-His-NH_2 (**8**) which incorporate either 2Fen or 6Bpa in conjunction with a histidine residue. Peptide **6**, like the bis(6Bpa)-peptide **5**, was designed to investigate the ability of histidine to participate in intramolecular coordination of metal cations with the 2-phenanthrolyl-amino acid. Further, the effect of the central tetrapeptide sequence on the ability to induce intramolecular coordination was examined with peptide **7**, as the central tetrapeptide sequence, Val-Pro-D-Ser-Phe, has been shown to exhibit a significant Type II reverse turn character in solution (see Chapter 2). As a result, it was anticipated that peptide **7** would have enhanced metal binding properties in comparison with peptide **6**. Finally, in order to test the relative binding properties of the similarly substituted 6Bpa and 2Fen in a multifunctional peptide, Ac-6Bpa-Val-Pro-D-Ser-His-NH2 (**8**) was also synthesized and investigated.

A wide range of analytical methods have been employed to study the metal binding properties of the bipyridine and phenanthroline ligand series. The most popular method is based upon differences between the solubility properties of the metal complexes relative to that of the free ligand.[12] The technique relies upon partitioning of the ligand and metal-ligand complexes in a biphasic system and the stepwise formation constants are calculated from the relative concentrations of the metal-bound components found in each phase. This method was not employed for the peptide models described above, as these systems were expected to remain soluble in an aqueous environment regardless of the coordination state of the bipyridine or phenanthroline ligand. Alternatively, potentiometry is a popular method for the determination metal-cation binding affinities for peptidyl systems.[13]

Unfortunately, this technique has been shown to produce inaccurate results for systems containing bipyridine and phenanthroline.[12] This inaccuracy has been attributed to the large differential between the ligand pK_a and the respective metal cation equilibrium constants. Apparently the proton cannot effectively compete for the bipyridyl- or phenanthrolyl-nitrogen atoms in the presence of metal cations, even under low pH conditions. Finally, UV/Vis absorption spectroscopy has been used to determine the binding affinities of both bipyridine and phenanthroline metal cation complexes.[14,15] It is well known that the bipyridine ligands undergo a significant red-shift in the π-π^* transition ($\approx$10-15 nm depending upon the metal cation) upon metal binding, and a similar, but smaller, effect is observed in the phenanthroline ligand series. These characteristic shifts provide a means to quantify the unbound and bound states of the ligand directly.[12] This technique is ideally suited for the peptide systems developed in this chapter, as it provides a sensitive, convenient method to determine the metal binding affinities of the peptides under physiological conditions.

In addition to providing a method for determining the binding affinities of the complexes, absorption spectra of some metal complexes may provide information about the coordination sphere geometry. Complexes with Co^{+2} can give rise to distinctive absorptions in the visible and near IR regions due to transitions in the metal *d-d* orbitals.[16,17] Since these transitions are highly dependent upon the ligand field arrangement about the metal cation, both the wavelengths of the absorptions and the relative molar absorbtivities of the transitions can be used to evaluate the metal-coordination environment. Indeed, for many Co^{+2} complexes it is generally possible to tentatively assign both the ligand geometries (tetrahedral, penta-coordinate, octahedral) as well as the ligand composition (N_3S, N_2S_2, etc.), by inspection of the visible wavelength regions of absorption spectra.[16] These transitions have been proved to be invaluable in the characterization of a wide variety of metalloprotein metal-binding site environments, including carbonic anhydrase,[18] the protein hormone insulin,[19] and the "zinc-fingers."[20]

Circular dichroism spectroscopy provides an additional measure of the metal binding properties of the peptide-metal complexes. Since the process of metal cation binding is likely to affect the overall structure of the peptide, measurable changes in the CD spectra would be expected for the peptide backbone upon metal cation binding. In addition, the relative disposition of ligand moieties around the metal cation center may provide a chiral environment with measurable Cotton effects. These effects are often used to distinguish Λ and Δ enantiomers of symmetrical metal-cation complexes.[21] As a result, CD can be used both as an alternative method for monitoring the progress of a metal cation titration, as well as a sensitive probe of the overall peptide-metal complex structure. For the peptides studied in this chapter, the CD spectra of the peptide:metal cation complexes provide a basis for comparison of the effects of various metal cations on the overall complex structure for a given peptide sequence, and within a series of peptides, the effects of metal cation binding can be delineated from the structural perturbations arising from changes in the peptide sequence.

Results and Discussion

Peptide Synthesis

The heteroaromatic, metal binding amino acids (**1-4**, See Figure 4-1) readily undergo peptide incorporation using N^{α}-FMOC protection-based solid phase peptide synthesis strategies and BOP/HOBt activated ester mediated couplings (See Experimental section of this chapter). While the amino acids were incorporated into a variety of peptides using the standard acylation and deprotection protocols, better coupling efficiencies were achieved when the amino acid activated ester solutions were prepared and filtered just prior to use and injected manually into the column containing the solid-phase resin. In addition, a double coupling protocol was used for all four of the amino acids to ensure that high coupling efficiencies were realized. In a typical amino acid coupling, 2.5 equivalents of the amino acid activated ester were loaded onto the column for a 1hr acylation period, after which a fresh sample of activated ester (1.5 equivalents) was applied for another 1hr. acylation period. Peptides synthesized using these protocols exhibited near quantitative coupling of the amino acids (> 95%, as determined by RP-HPLC and ^{1}H-NMR spectra of the cleaved peptides) regardless of position of the amino acid in the peptide primary sequence.

The resulting peptides were found to be quite stable to both the TFA cleavage conditions and subsequent workup and purification steps. While many of the peptides were sufficiently pure for analysis after synthesis, impure peptides were further purified using either semi-preparative RP-HPLC or P2 gel filtration chromatography. Due to the high metal cation binding affinities of 4Bpa and 5Bpa, peptides containing these amino acids were handled with great care to avoid metal contamination; doubly-distilled, de-ionized water and polypropylene or Teflon labware was used for all peptide manipulations.

Metal-Binding Properties of Peptides With the General Sequence:

*Ac-Xaa-Thr-Pro-D-Ala-Val-Phe-NH_2 (Peptides **1-4**)*

Metal cation titrations were performed on each of the four peptides, **1-4** with $ZnCl_2$, $CoCl_2$, and $CuCl_2$ metal salts. As can be seen in Figure 5-1a, the peptides containing bipyridyl-amino acids show an absorption at 285 nm due to the π-π^* transition of the bipyridyl moiety. A similar absorption is observed for the peptide containing the phenanthroline amino acid at 268 nm as shown in Figure 5-1b. Upon addition of metal cations, these absorptions decrease intensity with the concomitant appearance of new absorptions at ≈240 nm and ≈310 nm for the bipyridyl-containing peptides and ≈280 nm and ≈300 nm for the phenanthrolyl-containing peptides. The fact that clearly defined isosbestic points were observed upon addition of metal cations to the peptide solutions, regardless of the metal cation or the particular peptide, suggests that the transformations follow a two state equilibrium model.[22] As a result, the metal binding affinities and respective stoichiometry of the metal-cation:peptide complexes can be obtained directly from binding isotherms extracted from the absorption spectra.[22]

Binding isotherms were analyzed at those wavelengths that exhibited the largest change in absorption upon metal cation addition and data from a typical titration are shown in Figure 5-2a-b. The metal-cation binding process follows a square hyperbola relation as expected for a two-state binding process.[22] Calculation of the binding constants was performed through fitting isotherms collected from a series of wavelengths to the *y*-reciprocal form of the square hyperbola (Scott equation, see Experimental section of this chapter); sample plots are shown in Figure 5-3a-b. The stoichiometries of the metal:peptide complexes were determined directly from binding isotherms collected under saturating conditions ([Pep] >> K_d) by evaluation of the number of equivalents of metal cations (based on the concentration of peptide) needed to reach saturation. The overall results from these calculations are presented in Table 5-1.

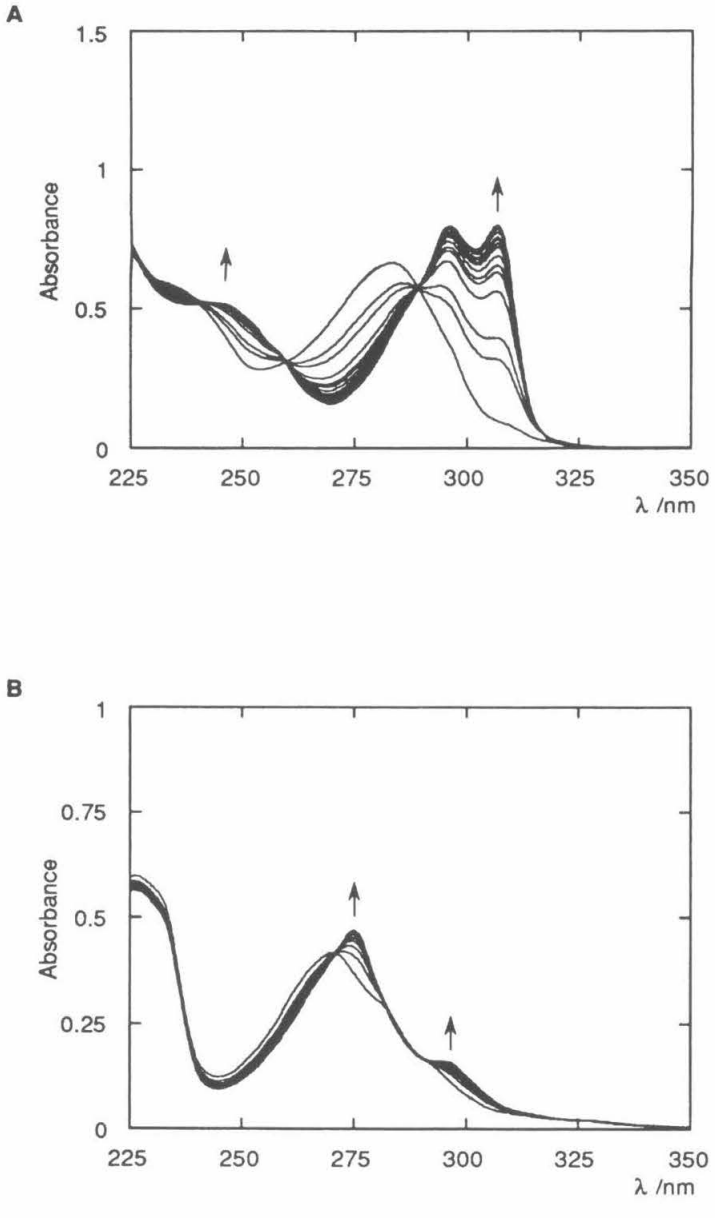


Figure 5-1. (A) Absorption spectra of the titration of **1** (52 μM) with $ZnCl_2$, (pH=6.8-7.0, 0.2M NaCl, 25° C; $ZnCl_2$ was added in 10 μM increments); (B) Absorption spectra of the titration of **4** (14 μM) with $ZnCl_2$, (pH=6.8-7.0, 0.2M NaCl, 25° C, $ZnCl_2$ was added in 100 μM increments).

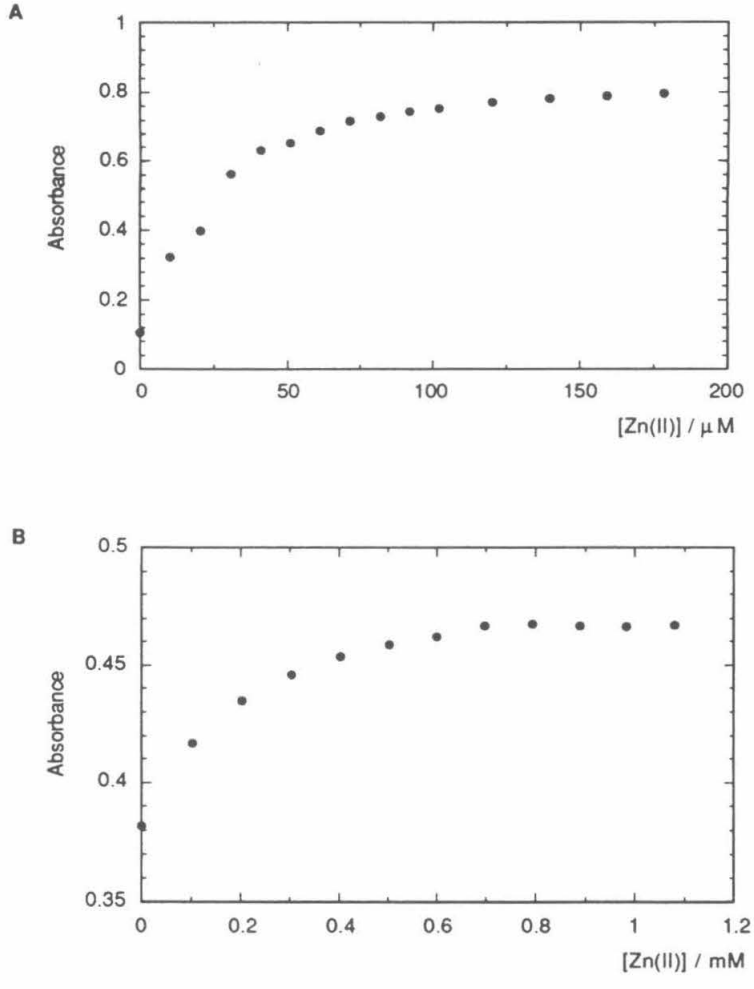


Figure 5-2. (A) Binding isotherm from the titration of **1** (52 μM) with $ZnCl_2$ (data taken at 306 nm); (B) Binding isotherm from the titration of **4** (14 μM) with $ZnCl_2$ (data taken at 276 nm).

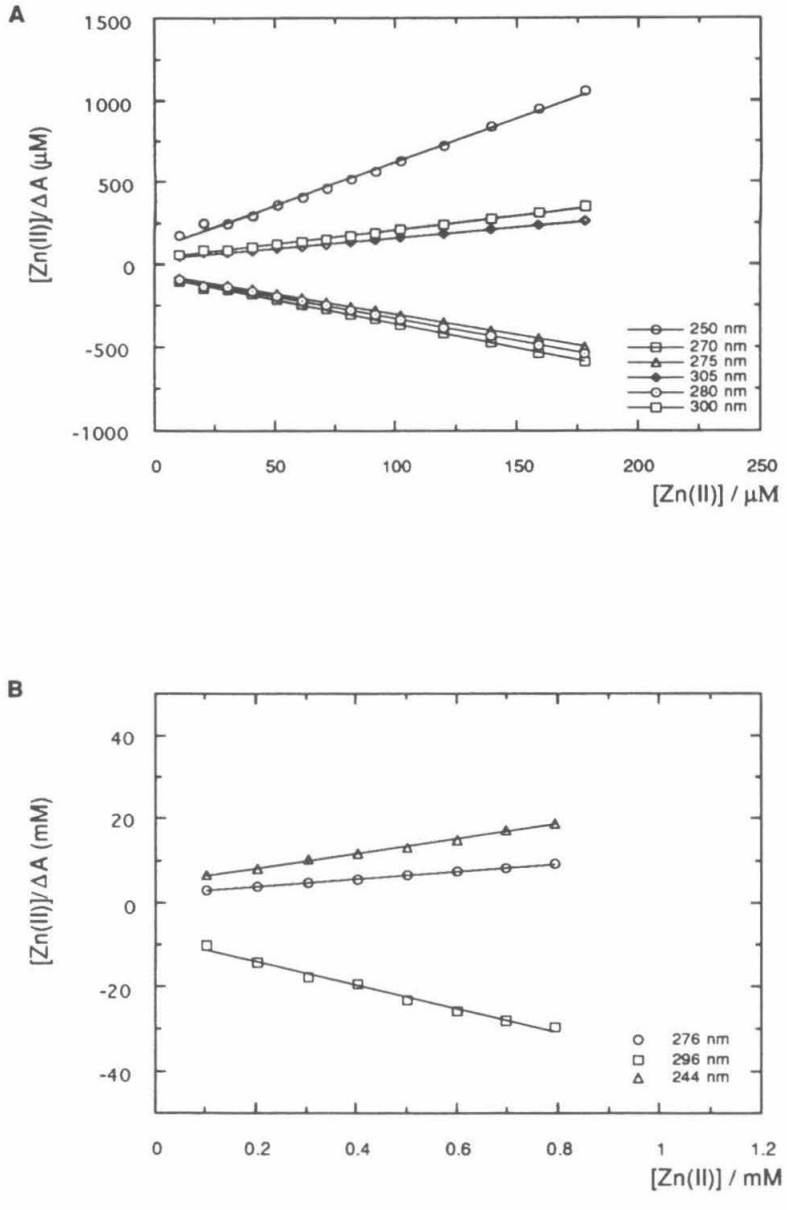


Figure 5-3. (A) *y*-Reciprocal plot of data from titration of **1** (52 μM) with $ZnCl_2$; (B) *y*-Reciprocal plot of data from the titration of **4** (14 μM) with $ZnCl_2$.

Table 5-1. Dissociation Constants and Metal Complex Stoichiometries for Bipyridyl- and Phenanthrolyl-Containing Peptides with the General Sequence Ac-Xaa-Thr-Pro-D-Ala-Val-Phe-NH_2.

Metal Cation	K_d, M (Peptide:Metal Stoichiometry)			
	4Bpa (**1**)	5Bpa (**2**)	6Bpa (**3**)	2Fen (**4**)
Zn^{+2}	$2.2 \pm 0.2 \times 10^{-5}$ (1:1)	$1.9 \pm 0.3 \times 10^{-5}$ (1:1)	$2.7 \pm 0.3 \times 10^{-3}$ (1:1)[b]	$2.6 \pm 0.3 \times 10^{-4}$ (1:1)
Co^{+2}	$3.8 \pm 0.4 \times 10^{-6}$ (2:1)[a]	$4.0 \pm 0.6 \times 10^{-6}$ (2:1)[a]	$1.1 \pm 0.1 \times 10^{-3}$ (1:1)[b]	$1.1 \pm 0.4 \times 10^{-5}$ (1:1)
Cu^{+2}	$< 1 \times 10^{-8}$[c] (1:1)	$< 1 \times 10^{-8}$[c] (1:1)	$0.8 \pm 0.2 \times 10^{-7}$ (1:1)	$< 1 \times 10^{-8}$[c] (1:1)

[a] The binding constants were calculated assuming a 1:1 relationship between the metal cation and the peptide; these binding constants represent the overall dissociation $K_d = [M(II)][2L]^2 / [M(II)L_2]$. [b]Metal complex stoichiometry was assumed to be 1:1.[23] [c] The sensitivity of the UV method limited assessment at lower concentrations.

The dissociation constants for the peptides vary over 5 orders of magnitude, depending the particular unnatural amino acid and the nature of the metal cation. In all cases, the tightest binding is observed with Cu^{+2} and the weakest for Zn^{+2}. The relative affinities for the metal cations follow that predicted by the Irving-Williams order of stabilities as expected from the similar trends observed in the parent 2,2'-bipyridine and 1,10-phenanthroline ligands as shown in Table 5-2.[24] However, the absolute magnitudes of the binding affinities are smaller than that observed with the parent ligands, particularly in the cases where the substitutions are *ortho* to the chelating nitrogen atoms, 6Bpa (**3**) and 2Fen (**4**). The diminished affinities for **3** and **4** is due in part to the steric hindrance of the substitution near the nitrogen atoms;[9,10] however, the lower overall affinities observed for

the peptides are probably the result of steric interactions arising from the polypeptide chain. The higher affinities of the 2Fen peptide (**4**) relative to the 6Bpa peptide (**3**) are consistent with the parent ligands (see Table 5-2), and these effects have been attributed to the more rigid structure of the 1,10-phenanthroline ring system.[9] The selectivity of the metal cations by the various peptides complements that reported for the polypeptide constructs containing unnatural oxygen-containing bidentate ligands (IDA and EDTA)[25,26] or the "zinc-finger" proteins ($K_d^{Zn} = 2.8 \pm 0.9 \times 10^{-9}$ M).[20]

Table 5-2. Dissociation Constants for Several Bipyridine and Phenanthroline Ligands[a]

	K_d, M		
Metal Cation	2,2'-Bipyridine	1,10-Phenanthroline	2-MePhen[b]
Zn^{+2}	7.4×10^{-6}	4.0×10^{-7}	1.1×10^{-5}
Co^{+2}	1.6×10^{-6}	8.3×10^{-8}	7.9×10^{-6}
Cd^{+2}	6.6×10^{-5}	1.6×10^{-6}	7.1×10^{-6}
Ni^{+2}	9.1×10^{-8}	2.5×10^{-9}	1.1×10^{-6}
Cu^{+2}	4.7×10^{-7}	4.0×10^{-8}	4.0×10^{-8}

[a]Values are based on the first dissociation step; K_d = [M(II)][L] / [M(II)L].[24] [b]2-MePhen = 2-Methyl-1,10-phenanthroline

All of the metal cations, with the exception of Co^{+2}, form 1:1 complexes with the peptides **1** and **2**. The Co^{+2} complexes for these peptides have 2:1 peptide:metal cation stoichiometry; the tendency to undergo intermolecular metal binding for these complexes can be attributed to the unobstructed N-N' chelation moieties which can assemble around the metal center efficiently. In fact, it is surprising that the intermolecular aggregation did not result in a 3:1 peptide:metal complex, as the $Co(Bpy)_3$ species readily forms when

excess 2,2'-bipyridine is incubated with Co^{+2} cations.[27] The lower stoichiometries observed for these peptides, as well as that exhibited by all of the peptide:metal complexes is most likely due to the steric bulk introduced by the polypeptide backbone; evidently the peptide chain restricts the metal coordination from forming higher order supermolecular structures. While these peptides provide useful information about the relative propensity for the peptides to undergo intermolecular metal coordination, it is not possible to evaluate the potential for intrapeptide ligation with these systems. These effects were studied with peptide **5**, which contains two bipyridyl-ligands in the primary sequence.

Metal-Binding Properties of Ac-6Bpa-Thr-Pro-D-Ala-Val-6Bpa-NH_2 (5)

The metal binding affinities of peptide **5** were measured using UV/Vis absorption spectroscopy as described above. The calculated metal binding constants, as shown in Table 5-3, were found to be at least an order of magnitude higher than the corresponding control peptide **3**. In addition, the metal-binding stoichiometries of the metal complexes were found to be 1:1 peptide/metal-cation for peptide **5**; this metal cation complex stoichiometry would not be expected if the bipyridine ligands acted independently. These observations, coupled with the fact that the peptide exhibited well-defined isosbestic points over the entire course of the titrations for all of the metal cations studied, strongly suggests that both bipyridyl ligands in the peptide sequence act cooperatively to ligate the metal-cation.

These conclusions are supported by metal titrations monitored with circular dichroism spectroscopy. Comparison of the circular dichroism titration studies for **3** and **5** with Cu^{+2} cations (see Figure 5-4a-b) demonstrates that the two complexes are quite distinct. The complex associated with peptide **5** exhibits an intense, positive ellipticity at 315 nm (Figure 5-4a) which is attributed to the bipyridyl chromophore. In contrast, the CD spectra for the control peptide (**3**, Figure 5-4b) shows only weak and broad ellipticities in this wavelength range (310-330nm). These spectra indicate that the Cu^{+2} cation induces

Table 5-3. Dissociation Constants and Metal Complex Stoichiometries for 6Bpa-Containing Peptides of the General Sequence: Ac-6Bpa-Thr-Pro-D-Ala-Val-Xaa-NH_2.

	K_d, M (Peptide:Metal Stoichiometry)	
Metal Cation	6Bpa (**5**)	Phe (**3**)
Zn^{+2}	$9.6 \pm 0.6 \times 10^{-4}$ (1:1)	$2.7 \pm 0.3 \times 10^{-3}$ (1:1)
Co^{+2}	$1.1 \pm 0.2 \times 10^{-4}$ (1:1)	$1.1 \pm 0.1 \times 10^{-3}$ (1:1)
Cd^{+2}	$1.1 \pm 0.2 \times 10^{-4}$ (1:1)	$4.3 \pm 0.3 \times 10^{-3}$ (1:1)
Ni^{+2}	$1.1 \pm 0.2 \times 10^{-6}$ (1:1)	$5.0 \pm 0.5 \times 10^{-5}$ (1:1)
Cu^{+2}	$< 1.0 \times 10^{-8}$ (1:1)	$0.8 \pm 0.2 \times 10^{-7}$ (1:1)

the formation of a chiral complex involving both bipyridyl ligands of peptide **5**. A schematic of the peptide complex is shown in Figure 5-5.

The corresponding Co^{+2}:peptide **5** complex was analyzed by visible absorption spectroscopy. As shown in Figure 5-6a, addition of Co^{+2} to the peptide is accompanied by a weak absorption at 450 nm ($\varepsilon_{max} < 100$), which is suggestive of a five-coordinate complex.[18] Interestingly, however, the Co^{+2}:peptide **5** complex did not show any evidence for chiral complex in the circular dichroism spectra (see Figure 5-6b); indeed, the spectra shows a negative ellipticity at ≈310 nm in direct contrast to the Cu^{+2} complex.

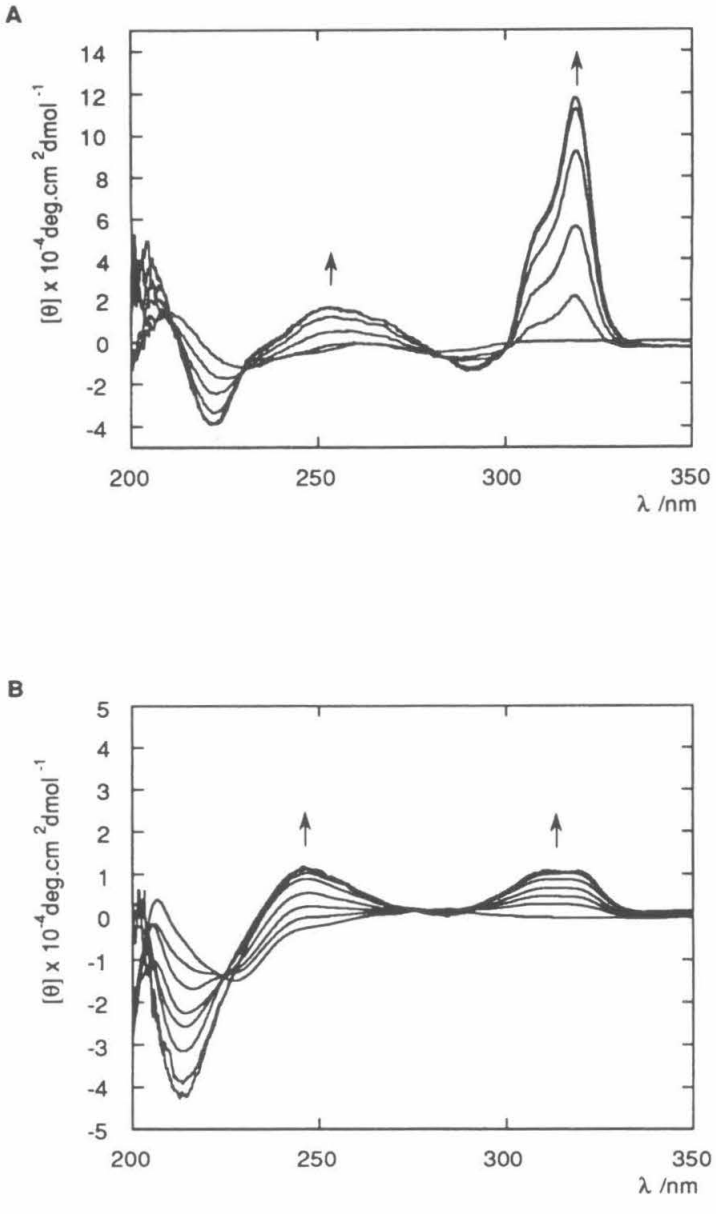


Figure 5-4. (A) Circular dichroism spectra of the titration of **5** (40 μM) with $CuCl_2$, (pH=6.6, 0.2M NaCl, 25° C, $CuCl_2$ was added in 7.5 μM increments); (B) Circular dichroism spectra of the titration of **3** (40 μM) with $CuCl_2$ (pH=6.6, 0.2M NaCl, 25° C, $CuCl_2$ was added in 7.5 μM increments).

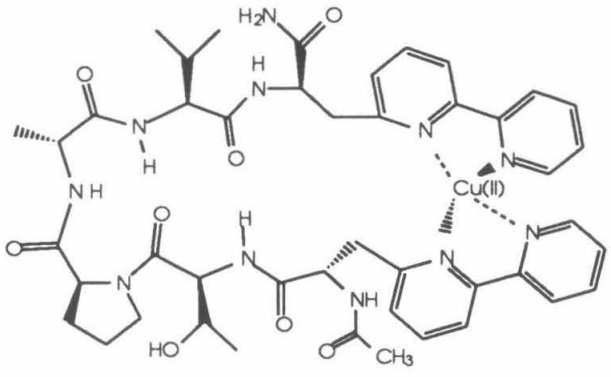


Figure 5-5. Schematic of the Cu(II):peptide **5** complex. (No indication of the coordination geometry is intended.)

However, the other regions of the CD spectra (200-280 nm) for the Co^{+2} complex show a great deal of similarity to the spectra collected for the Cu^{+2}:peptide **5** complex. As a result, the data suggests that the peptide undergoes similar conformational changes upon binding the two metal cations; however, the respective coordination geometries of the two complexes are not identical.

The results from these clearly indicate that the metal binding properties of peptides containing the bipyridyl amino acids can act cooperatively to bind metal cations, provided that the primary sequence is carefully selected to promote these interactions. However, in all of these cases, the two bipyridine ligands provide four ligation moieties. Many metalloenzymes, particularly the catalytically active Zn^{+2} enzymes,[3,28] contain tetrahedral coordination geometries where three of the ligands are derived from the protein structure. Based on the results from peptide **5**, peptides **6-8** were designed to form a tridentate metal binding site comprised of a bipyridyl or phenanthrolyl ligand with a histidine imidazole group.

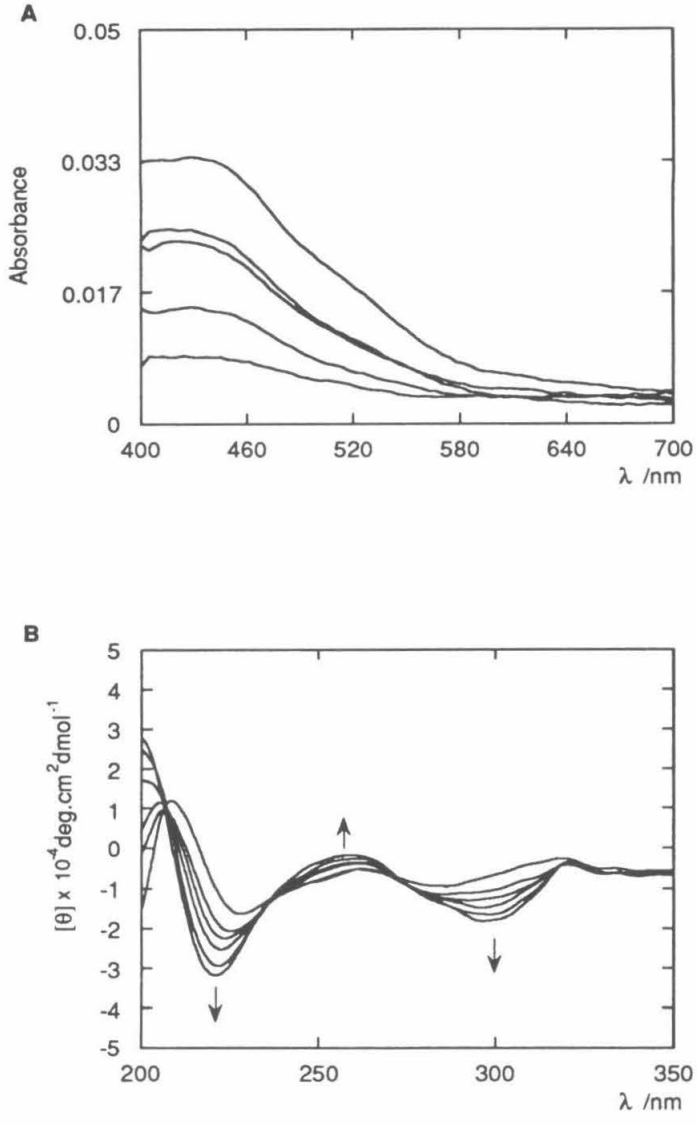


Figure 5-6. (A) Visible absorption spectra of the titration of **5** (300 μM) with $CoCl_2$ (pH=7.0, 0.2 M NaCl, 25° C); (B) Circular dichroism spectra of the titration of **5** (300 μM) with $CoCl_2$ (pH=7.0, 0.2 M NaCl, 25° C). $CoCl_2$ concentrations were as follows: 100 μM, 200 μM, 300 μM, 600 μM, and 1.2 mM.

Metal Binding Properties of Peptides Containing Phenanthrolyl or Bipyridyl ligands and Histidine - Peptides 6-8

The metal binding affinities of peptides **6-8** were determined for Co^{+2} and Zn^{+2} cations using absorption spectroscopy titrations as described above, and the results of these experiments are shown in Table 5-4. All three of the mixed-ligand peptides were found to have 1:1 peptide/metal ion stoichiometries for both Co^{+2} and Zn^{+2} cations and well-defined isosbestic points were observed throughout the metal cation titrations. These observations, along with the 25-fold enhancement in the Zn^{+2} binding affinities of peptide **6** over the corresponding control peptide **4** clearly suggests that the histidine and phenanthroline ligands act cooperatively to bind the metal cation, in a similar fashion to that observed for peptide **5**. Similar effects are assumed to be in operation for peptides **7** and **8**, since these peptides also exhibit enhanced binding affinities over the corresponding control peptides **3** or **4**.

Table 5-4. Dissociation Constants and Metal Complex Stoichiometries for Peptides containing 6Bipyridyl- or 2-Phenanthrolyl-Alanine and Histidine

	K_d, M (Peptide:Metal Stoichiometry)	
Peptide	Zn^{+2}	Co^{+2}
Ac-2Fen-Thr-Pro-D-Ala-Val-His-NH_2 (**6**)	$9.4 \pm 0.2 \times 10^{-6}$ (1:1)	$< 5 \times 10^{-8}$ (1:1)
Ac-2Fen-Val-Pro-D-Ser-Phe-His-NH_2 (**7**)	$2.0 \pm 0.8 \times 10^{-6}$ (1:1)	$< 5 \times 10^{-8}$ (1:1)
Ac-6Bpa-Thr-Pro-D-Ala-Val-His-NH_2 (**8**)	$9.3 \pm 0.2 \times 10^{-5}$ (1:1)	$1.0 \pm 0.3 \times 10^{-5}$ (1:1)

The circular dichroism spectra for the metal cation titrations of peptide **6** are shown in Figure 5-7a-b. It is clear that the peptide undergoes some spectral changes upon metal cation addition; however, unlike the spectra observed for Cu^{+2}:peptide **5** complex, the near UV absorption region of the spectra for both metal cations shows only broad, weak ellipticities. In addition, comparison of the spectral changes for the two metal cation titrations indicates that the two metals have complexes that are distinct from one another.

The two peptides containing the central tetrapeptide sequences known to promote turn formation (**7** and **8**) show large changes in the UV region of the circular dichroism spectra upon addition of metal binding cations (see Figures 5-8a-b, 5-9a-b). The results from the titrations with both Co^{+2} and Zn^{+2} are very similar for both of the peptides, indicating that they undergo similar changes upon binding the metal ions. The resulting metal-bound peptide complexes exhibit CD spectra which resemble that predicted for Type II reverse turns and β-sheets;[29] however, the presence of the bipyridyl and phenanthrolyl chromophores could lead to effects on the overall spectra which are hard to characterize. As a result, it is not possible to deduce unambiguous structural information from the spectra based on the polypeptide chromophores.

The relative contribution of the central tetrapeptide conformation on the overall metal binding properties can be determined through comparison of the metal binding affinities of peptides **6** and **7**. These peptides differ only in the central tetrapeptide sequence; peptide **6** contains the central tetrapeptide used in all the above studies, Thr-Pro-D-Ala-Val, which is known to have only moderate reverse turn character in solution.[11] In contrast, the central tetrapeptide incorporated in peptide **7**, Val-Pro-D-Ser-Phe, is known to have a large population of reverse turn character in aqueous solution (see Chapter 2). The 5-fold increase in binding affinity observed for peptide **7** over peptide **6** for Zn^{+2} indicates that the reverse turn plays an active role in the metal binding properties of the peptide, presumably through pre-organization of the two metal binding ligands. These results are in agreement with the results from similar studies investigating the effects of turn forming

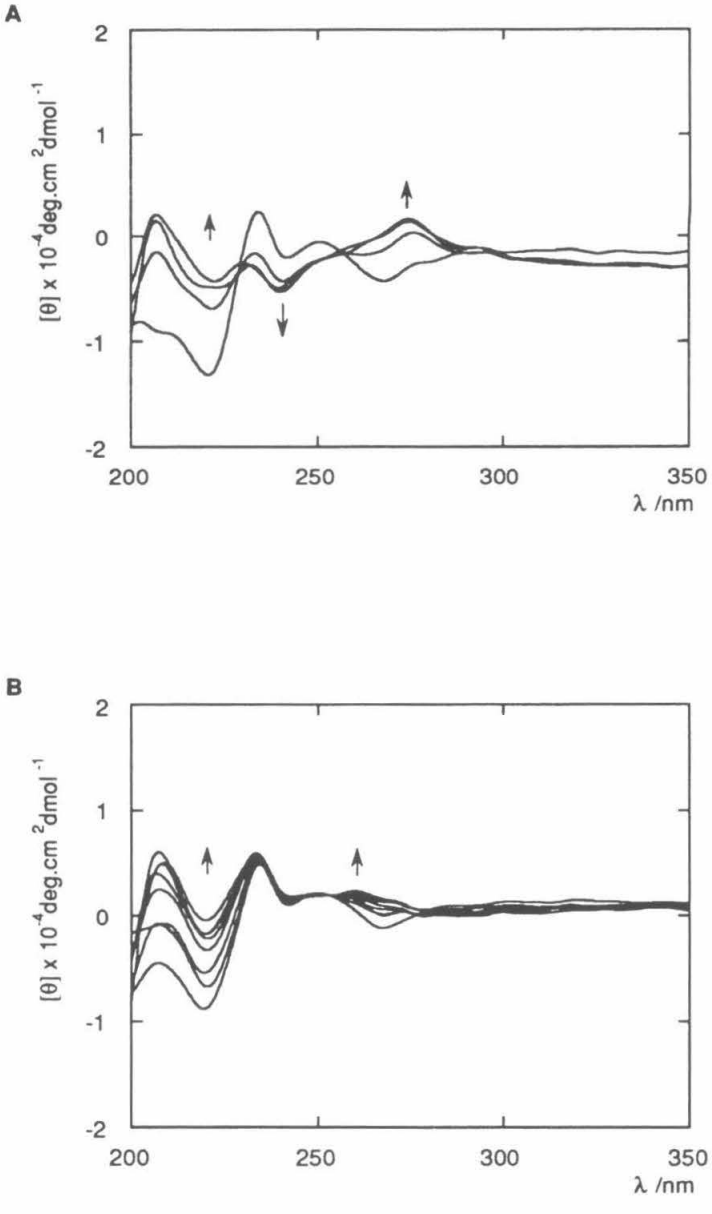


Figure 5-7. (A) Circular dichroism spectra of the titration of **6** (30 μM) with $ZnCl_2$, (pH=8.25, 250 μM HEPES, 25° C, $ZnCl_2$ was added in 8 μM increments); (B) Circular dichroism spectra of the titration of **6** (30 μM) with $CoCl_2$ (pH=8.25, 250 μM HEPES, 25° C, $CoCl_2$ was added in 50 μM increments).

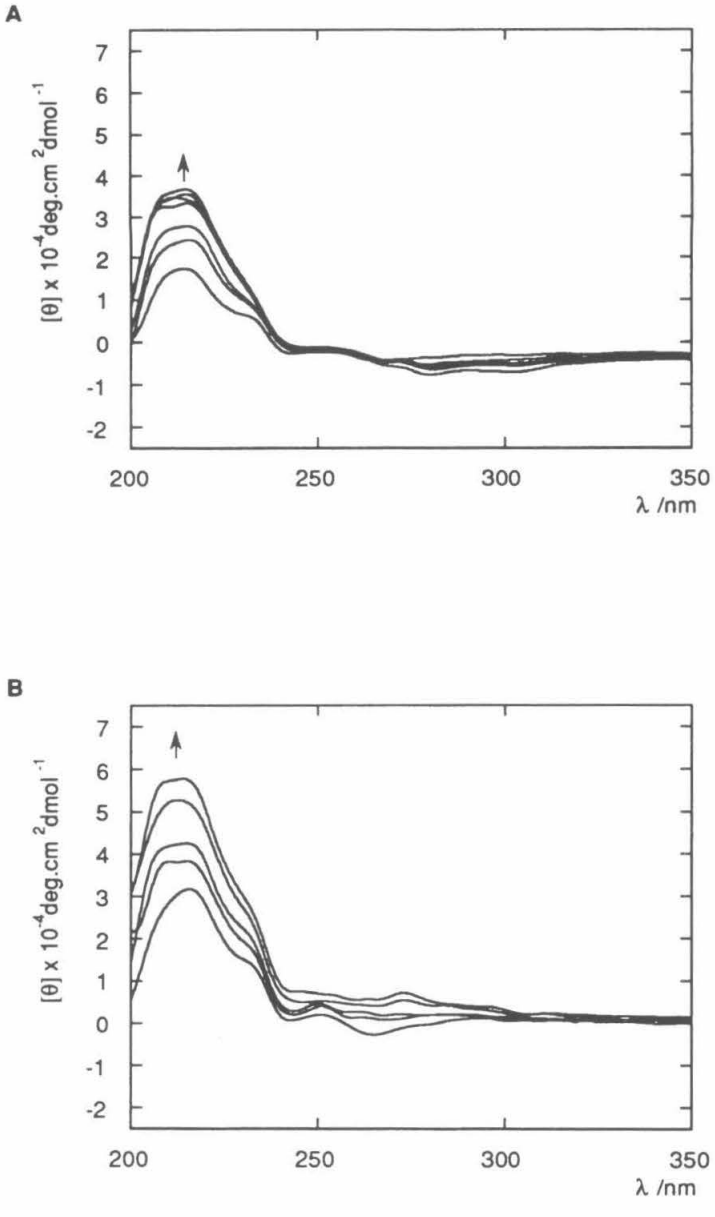


Figure 5-8. (A) Circular dichroism spectra of the titration of **7** (30 μM) with $ZnCl_2$, (pH=8.25, 250 μM HEPES, 25° C, $ZnCl_2$ was added in 5 μM increments); (B) Circular dichroism spectra of the titration of **7** (20 μM) with $CoCl_2$ (pH=8.25, 250 μM HEPES, 25° C, $CoCl_2$ was added in 5 μM increments).

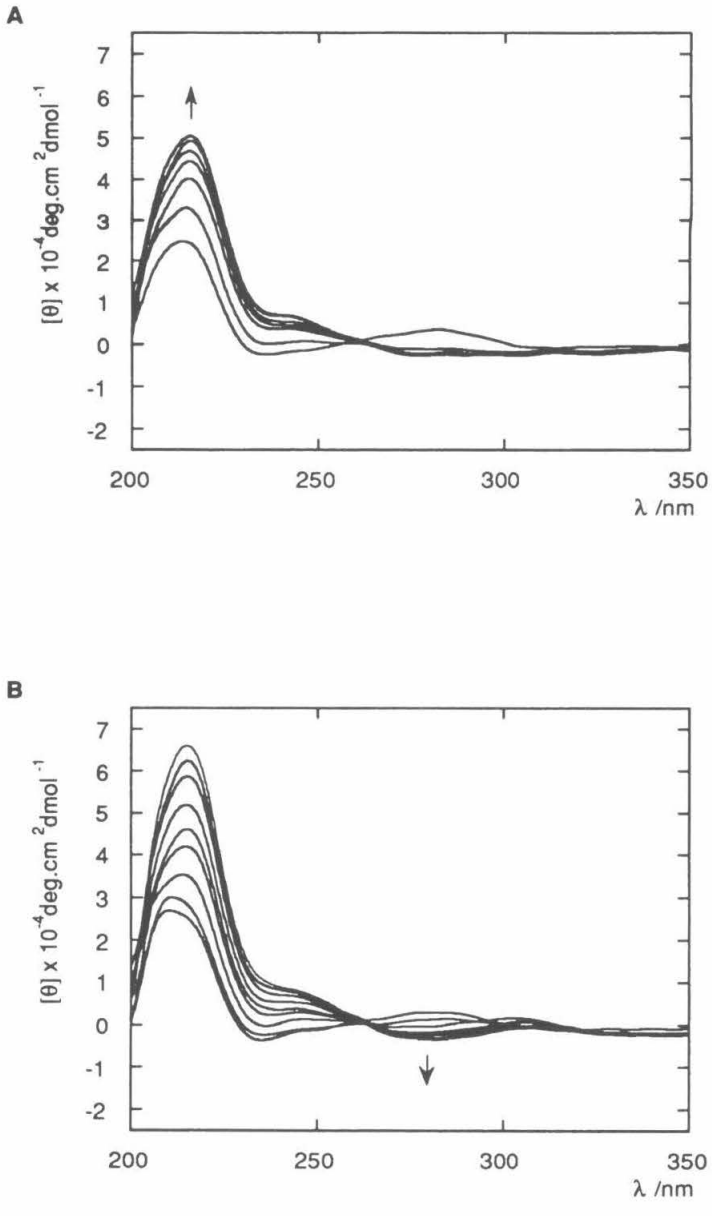


Figure 5-9. (A) Circular dichroism spectra of the titration of **8** (40 μM) with $ZnCl_2$, (pH=8.25, 250 μM HEPES, 25° C, $ZnCl_2$ was added in 50 μM increments); (B) Circular dichroism spectra of the titration of **8** (40 μM) with $CoCl_2$ (pH=8.25, 250 μM HEPES, 25° C, $CoCl_2$ was added in 20 μM increments).

sequences on the binding affinities of hexapeptides incorporating two bipyridyl groups (6Bpa).[11] Similar correlations were observed between the metal-cation binding affinities and the turn nucleation potential of the central tetrapeptide sequence; in the extreme case, a 15-fold differential in binding affinities was observed. Similarly, the turn sequence has been shown to be essential for effective metal cation binding with octapeptides containing histidine residues.[23] In these studies, a series of octapeptides were prepared which incorporated a central tetrapeptide sequence flanked by two dipeptide sequences containing a total of three histidine residues. Efficient metal cation binding was observed using circular dichroism spectroscopy for the peptides which contained sequences that promoted reverse turn nucleation, whereas the sequences that were predicted to have poor reverse turn character did not have well-defined metal binding characteristics.

Finally, the fact that peptide **8**, which contains a 6Bpa residue, shows a ≈50 fold reduction in the binding affinity for Zn^{+2} and a full 200 fold reduction in the Co^{+2} binding affinity relative to the 2Fen peptide (**7**) illustrates the ability to precisely modulate the metal binding properties of peptides containing mixed-ligand systems. These results demonstrate the ability to select for various metal cations based on both the choice of the metal binding amino acid, in conjunction with the carefully designed primary sequences.

Conclusions

The bipyridyl- and phenanthrolyl-based unnatural amino acids provide ideal components for the construction of peptides with unique metal binding properties. The amino acids can be readily incorporated into peptide sequences with little modification to the standard solid phase peptide synthesis protocols and are stable to the conditions needed for peptide cleavage from the resin and subsequent purification. Peptides incorporating these amino acids bind a wide variety of metal cations with affinities only slightly lower than the parent 2,2'-bipyridine and 1,10-phenanthroline ligands. Interestingly, while peptides containing these amino acids did not readily undergo intermolecular metal

chelation, the peptides which were designed to promote intramolecular metal cation binding had greatly enhanced metal binding affinities and peptide:metal cation stoichiometries consistent with cooperative metal binding. In addition, the cooperative metal binding properties were found to be dependent on the primary sequence of the peptide. Several peptides were designed to provide enhanced chelation abilities through the incorporation of reverse turn character in the polypeptide sequence, and the expected correlation between the enhanced metal binding affinities and the amount of reverse turn character in the peptide sequence was observed.

The results of the studies discussed in this chapter provide a foundation for the construction of novel metalloprotein motifs containing the four bipyridyl- and phenanthrolyl amino acids. The ability to modulate the metal binding affinities through both the choice of the metal binding amino acid and the careful selection of the primary sequence should allow the construction of tailored metal binding sites programmed for either structural or functional roles. An example of the use of these amino acids in a functional role is discussed in Chapter 6, where 4Bpa and 6Bpa were used to create mutants of horse heart cytochrome *c* which had enhanced electron transfer properties.

The metal binding studies on the 6Bpa containing peptides **3** and **5** was published as a communication in the *Journal of the American Chemical Society*:

Imperiali, B.; Fisher, S.L.; "(*S*)-α-Amino-2,2'-bipyridine-6-propanoic Acid: A Versatile Amino Acid for de Novo Metalloprotein Design," *J. Am. Chem. Soc.*, **1991**, *113*, 8527-8528.

Experimental

Peptide Synthesis:

All peptides were synthesized on a 0.1-0.2 mmol scale using solid-phase FMOC-amino protection and BOP/HOBT activated-ester chemistry on a Milligen 9050 automated peptide synthesizer using the methods described in Chapter 2 with some modifications. PAL resin was used to afford amides at the carboxy terminus. For the commercially available residues 4.0 equivalents of amino acid were used per coupling. However, for the bipyridyl and phenanthrolyl amino acids a double coupling protocol was employed, where the first coupling used 2.5 equivalents of amino acid, followed by a similar acylation with 1.5 equivalents of residue. Activated esters were formed *in situ* using BOP, HOBt, and 0.451 M *N*-methylmorpholine in *N*, *N*-dimethylformamide (DMF). Due to the poor solubility of the bipyridyl and phenanthrolyl amino acid activated esters, the amino acid ester solutions were filtered manually through a 0.45 μm nylon syringe filter (Gelman Scientific) prior to manual injection onto the resin column. Acylation times varied from 30-90 mins depending upon the coupling efficiency of the particular amino acid; however, a minimum of 60 min was employed for any of the unnatural amino acid acylation steps. Deprotection of FMOC-protected amine groups was performed using a 7-min 20% piperidine/DMF wash or a 5-min 2% DBU/DMF wash. All peptides were acetyl capped on the resin using 21 equivalents of acetic anhydride and 5 equivalents of triethylamine in 3 mL DMF. The resin was shaken for 2 h, filtered, washed with CH_2Cl_2, and dried *in vacuo* overnight. The peptides were cleaved from the resin and lyophilized from water as described in Chapter 2. The purity of the peptides was assessed by RP-HPLC and ^1H NMR. Impure peptides were purified using P2 gel filtration chromatography with 50 mM acetic acid as eluent, or through semi-preparative RP-HPLC using CH_3CN/H_2O gradients. Fractions containing pure peptide were combined and concentrated by lyophilization, then stored at -20 °C.

Metal Cation Binding Titrations:

Equilibrium binding constants were determined through titrations of aqueous solutions of known peptide concentration with additions of standardized metal solutions at 25 °C. The titrations were followed through monitoring the red-shift of the bipyridyl or phenanthrolyl π-π^* transition upon metal binding using UV spectroscopy. Metal cation solutions (Zn^{+2}, Co^{+2}, Cu^{+2}, Cd^{+2}, Ni^{+2}) were prepared from the corresponding anhydrous chloride salt and the Zn^{+2}, Co^{+2} and Cu^{+2} solutions were standardized by ethylene-diamine tetraacetic acid complexometric titrations (0.5 M standardized EDTA solution, Aldrich) using murexide as an indicator.[30] Concentrations of peptide were determined spectrophotometrically using the absorption of the unbound-bipyridyl ligand transition at 285 nm (ε_{max} = 12,800) or the corresponding phenanthrolyl transition at 268 nm (ε_{max} = 13,950). For peptides containing only phenanthrolyl or bipyridyl ligands, the pH of the solution was monitored by a glass electrode and held within the range pH=7.0-7.5 through small additions of 0.1 N NaOH or 0.1 N HCl in the presence of 200 mM NaCl. For peptides containing histidine ligands, the titrations were performed in 50 mM HEPES buffer, pH=8.25. The peptide concentrations for each titration ranged from 0.5 μM-0.5 mM and were varied according to the particular metal cation affinity to ensure that the titrations were not performed under saturating conditions. Titrations of dilute peptide solutions (< 2 μM) were carried out in a 10 cm-path length quartz cuvette, all other titrations were performed in a 1 cm cuvette. A miminum of three titrations were collected for each metal cation-peptide complex.

Binding constants were determined from the data between 20-80% of the complex formation using graphical analysis. Binding isotherms were obtained from those wavelengths where maximal changes in the absorptions were observed upon metal binding (typical wavelengths: 250, 270, 275, 280, 300, 305 nm; the actual wavelengths used for each determination was based on the particular peptide-metal cation pair). The data from each binding isotherm was then plotted as a function of metal concentration divided by the

change in absorbance (ΔA = observed absorbance - absorbance of unbound ligand) versus the metal concentration in accordance with the *y*-reciprocal form of the rectangular hyperbola (Scott equation):[31,32]

$$\frac{b[M(II)]}{\Delta A} = \frac{[M(II)]}{[Pep]_{tot}\,\Delta\varepsilon_{11}} + \frac{1}{[Pep]_{tot}\,K_{11}\,\Delta\varepsilon_{11}}$$

Where : b = cell pathlength

$\Delta\varepsilon_{11}$ = change in molar absorbtivities due to complex formation

The resulting plots were subjected to linear least-squares regression analysis and the binding constant was determined from the corresponding slope and *y*-intercept according to the relation:

$$K_d = (y\text{-intercept})/(\text{slope})$$

In all cases, the data for each of the wavelengths was found to be strictly linear, and the correlation constant (R) from least-squares linear regression analysis was always greater than 0.99. The reported metal binding constants represent the average value resulting from the multiple determinations and the standard deviation was taken to be the error in the binding constant value. In all cases, the binding isotherms were assumed to follow a two-state equilibrium when several well-defined isosbestic points associated with the addition of metal cations were observed.[22]

The stoichiometry of the metal-cation complexes was determined from titrations performed under saturating conditions ([Pep] >> K_d). Under these conditions, the stoichiometry of the complex was reflected by the number of equivalents of metal needed to reach saturation.

Circular dichroism metal titrations were performed in 1 cm path-length quartz cuvettes with peptide concentrations ranging from 25-50 μM using the same conditions as listed above for the UV/Vis metal titrations for the peptides in unbuffered solutions; peptides containing histidine residues were performed at pH=8.25 in 250 μM HEPES (25° C). Spectra were collected after each addition of metal over the wavelength range 200-350 nm using a scan speed of 50 nm/min, a time constant of 0.5 s, and a bandwith of 1 nm. A minimum of 8 scans were taken for each spectra. The spectra are reported in ellipticity/mole peptide.

RP-HPLC Gradients Used:

(A) 100% H_2O (0.1% v/v TFA) 5 min., Linear Gradient 0-60% CH_3CN (0.08% TFA)/ H_2O (0.1 % v/v TFA) 25 min., 60 % CH_3CN (0.11% v/v TFA) 5 min., Linear Gradient 60-0% CH_3CN (0.08% v/v TFA)/ H_2O (0.1 % v/v TFA) 2 min.

(B) Linear Gradient 10-60% CH_3CN (0.08% TFA)/ H_2O (0.1 % v/v TFA) 30 min., 60 % CH_3CN (0.11% v/v TFA) 5 min., Linear Gradient 60-0% CH_3CN (0.08% v/v TFA)/ H_2O (0.1 % v/v TFA) 5 min.

Ac-2Fen-Thr-Pro-D-Ala-Val-Phe-NH_2: ^1H NMR (90/10 H_2O/D_2O) δ: 9.12 (d, 1H, J=11.0 Hz), 9.08 (d, 1H, J= 19Hz), 8.60 (d, 1H, J=16Hz), 8.44 (d, 1H, J=13Hz), 8.27 (m, 2H), 8.17 (m, 1H), 8.09 (m, 1H), 8.00 (d, 2H), 7.90 (d, 1H, J=17Hz), 7.80 (d, 1H, J=20Hz), 7.20 (m, 4H), 7.10 (m, 2H), 6.94 (s, 1H), 5.25 (m, 1H), 4.54 (m, 1H), 4.41, (m, 1H), 4.24 (m, 1H), 3.94 (m, 1H), 3.85 (m, 1H), 3.58 (m, 4H), 3.50 (m, 1H), 3.05 (m, 1H), 2.78 (m, 1H), 2.00 (m, 8H), 1.27 (d, 3H, J=17 Hz), 1.50 (d, 3H,J=16Hz), 0.69 (m, 6H); ^{13}C NMR (90/10 H_2O/D_2O) δ: 175.6, 175.2, 174.4, 173.7, 173.3, 173.0, 169.6, 159.0, 146.7, 143.0, 138.3, 137.7, 136.6, 136.0, 129.9, 129.7, 129.3, 129.2, 128.9, 128.8, 128.7, 127.2, 127.1, 125.1, 124.8, 69.8, 67.3, 60.3, 56.8, 54.8, 48.3,

39.6, 39.0, 37.0, 36.9, 30.0, 29.5, 24.7, 22.0, 18.9, 18.4, 17.9, 17.2. RP-HPLC (A) : RT = 25.0 min.; HRMS [MH^+]: calcd. for $C_{43}H_{54}N_9O_8$ [824.4095] obsd. [824.4117].

Ac-2Fen-Thr-Pro-D-Ala-Val-His-NH_2: ^1H and ^{13}C NMR in (90/10 H_2O/D_2O) - peptide exists in two well defined conformations, see reproduced spectra; RP-HPLC (A) : RT = 18.8 min.; HRMS [MH^+], calcd. for $C_{40}H_{52}N_{11}O_8$ [814.4000] obsd. [814.3983].

Ac-2Fen-Val-Pro-D-Ser-Phe-His-NH_2: ^1H and ^{13}C NMR in (90/10 H_2O/D_2O) - peptide exists in several distinct conformations, see reproduced spectra. RP-HPLC (A) : RT = 27.0 min.;HRMS [MH^+], calcd. for $C_{45}H_{54}N_{11}O_8$ [876.4156] obsd. [876.4140].

Ac-6Bpa-Val-Pro-D-Ser-Phe-His-NH_2: ^1H NMR (D_2O) δ: 8.68 (d, 1H, J=5.0 Hz), 8.51 (s, 1H), 8.29 (m, 2H), 8.04 (d, 1H, J=7.8 Hz), 7.97 (t, 1H, J=7.8 Hz), 7.77 (m, 1H), 7.41 (d, 1H, J=7.5 Hz), 7.15 (m, 5H), 7.05 (s, 1H), 5.0 (m, 1H), 4.46 (m, 1H), 4.38 (m, 3H), 4.29 (m, 1H), 3.75 (m, 3H), 3.60 (m, 1H), 3.25 (m, 2H), 3.15 (m, 2H), 3.02 (m, 2H), 2.15 (m, 1H), 2.10 (m, 2H), 1.92 (m, 4H), 0.96 (d, 3H, J=6.8 Hz), 0.88 (d, 3H, J=6.7 Hz); ^{13}C NMR (90/10 H_2O/D_2O) δ: 175.5, 175.0, 174.3, 173.9, 173.8, 172.5, 172.2, 145.5, 142.1, 140.6, 136.6, 134.5, 130.0, 129.7, 128.2, 127.5, 127.3, 124.7, 128.2, 122.0, 118.5, 118.2, 62.2, 61.8, 57.9, 57.0, 56.3, 54.5, 53.0, 49.2, 30.9, 30.4, 27.2, 25.8, 22.6, 19.4, 18.2, 16.1; RP-HPLC (B) : RT = 21.8 min.; HRMS [MH^+], calcd. for $C_{43}H_{54}N_{11}O_8$ [852.4157] obsd. [852.4160].

Ac-4Bpa-Thr-Pro-D-Ala-Val-Phe-NH_2: ^1H NMR (DMSO-$_{d6}$) δ: 8.70 (d, 1H, J=5.0 Hz), 8.57 (d, 1H, J=6.0 Hz), 8.38 (m, 2H), 8.18 (m, 2H), 8.00 (t, 1H, J=3.0 Hz), 7.80 (m, 4H), 7.50 (m, 1H), 7.40 (m, 1H), 7.22 (m, 4H), 7.15 (m, 1H), 7.07 (m, 2H), 4.75 (m, 1H), 4.54 (m, 1H), 4.34 (m, 2H), 4.28 (m, 2H), 4.00 (m, 2H), 3.90 (m, 1H), 3.72 (m, 1H), 3.66 (m, 1H), 3.10 (m, 2H), 2.85 (m, 2H), 1.80-1.70 (m, 8H), 1.65 (d, 3H, J=8.0

Hz), 1.10 (d, 3H, J=6.0Hz), 0.74 (d, 3H, J=8.9 Hz), 0.66 (d, 3H, J=7.0 Hz); ^{13}C NMR (90/10 H_2O/D_2O) δ: 176.5, 176.1, 175.1, 174.8, 174.2, 173.1, 170.6, 155.0, 148.7, 148.3, 146.8, 144.1, 137.6, 130.2, 129.8, 128.6, 128.3, 125.3, 124.7, 68.1, 61.8, 61.0, 57.9, 55.7, 54.7, 50.7, 49.4, 38.2, 31.0, 30.4, 25.7, 22.7, 19.7, 18.8; RP-HPLC (A) : RT = 20.04 min.; HRMS [MH^+], calcd. for $C_{41}H_{54}N_9O_8$ [800.4095] obsd. [800.4055].

Ac-5Bpa-Thr-Pro-D-Ala-Val-Phe-NH_2: (NMR spectra show evidence of aggregation in solution) ^1H NMR (DMSO-*d6*) δ: 8.68 (m, 1H), 8.51 (m, 1H), 8.38 (m, 1H), 8.30 (m, 1H), 8.17 (m, 1H), 8.13 (m, 1H), 8.0 (m, 1H), 7.80 (m, 4H), 7.47 (m, 1H), 7.22 (m, 4H), 7.21 (m, 1H), 7.15 (m, 2H), 4.71 (m, 1H), 4.54 (m, 1H), 4.35 (m, 2H), 4.28 (m, 2H), 4.00 (m, 2H), 3.88 (m, 1H), 3.73 (m, 2H), 2.84 (m, 2H), 1.80-1.70 (m, 8H), 1.18 (d, 3H, J=8.0 Hz), 1.08 (d, 3H, J=7.0 Hz), 0.75 (d, 3H, J=9.0 Hz), 0.66 (d, 3H, J=6.9 Hz); ^{13}C NMR (90/10 H_2O/D_2O) δ: 175.8, 175.7, 175.3, 175.2, 174.2, 174.0, 173.3, 170.0, 169.2, 136.7, 129.3, 129.0, 127.3, 67.5, 67.2, 61.0, 60.1, 57.2, 54.9, 49.9, 48.6, 37.0, 30.1, 29.5, 24.9, 21.9, 18.5, 18.0; RP-HPLC (A) : RT = 25.1 min.; HRMS [MH^+], calcd. for $C_{41}H_{54}N_9O_8$ [800.4095], obsd. [800.4057].

Ac-6Bpa-Thr-Pro-D-Ala-Val-Phe-NH_2: ^1H NMR (DMSO-*d6*) δ: 8.75 (d, 1H, J=4.4 Hz), 8.52 (d, 1H, J=7.9 Hz),8.30 (m, 2H), 8.09 (t, 1H, J=7.1 Hz), 7.91 (m, 5H), 7.58 (t, 1H, J=6.6 Hz), 7.40 (d, 1H, J=7.6 Hz), 7.24 (m, 7H), 4.98 (m, 1H), 4.58 (t, 1H, J=6.3 Hz), 4.40 (m, 1H), 4.32 (m, 2H), 4.04 (m, 2H), 3.73 (m, 1H), 3.63 (m, 1H), 3.32 (m, 1H), 3.11 (m, 2H), 2.86 (m, 1H), 2.04 (m, 2H), 1.81 (m, 6H), 1.20 (d, 3H, J=7.0 Hz), 1.11 (d, 3H, J=6.2 Hz), 0.77 (d, 3H, J= 6.8 Hz), 0.70 (d, 3H, J=6.8 Hz); HRMS [M^+], calcd for $C_{41}H_{54}N_9O_8$ [800.4095], obsd. [800.4135].

Ac-6Bpa-Thr-Pro-D-Ala-Val-6Bpa-NH_2: ^1H NMR (D_2O) δ: 8.29 (s, 2H), 7.63 (m, 8H), 7.21 (m, 2H), 7.10 (m, 2H), 4.30 (d, 1H, J=5.8 Hz), 3.99 (d, 1H, J=7.14 Hz), 3.86 (t, 1H, J=7.52 Hz), 3.77 (t, 1H, J= 6.1 Hz), 3.66 (d, 1H, J=7.23 Hz), 3.36 (m, 2H), 3.16 (m, 1H), 2.93 (m, 4H), 1.64 (m, 13H), 0.99 (d, 3H, J=7.2 Hz), 0.36 (d, 3H, J=6.8 Hz); HRMS [M^+], calcd for $C_{45}H_{56}N_{11}O_8$ [878.4313], obsd. [878.4285].

References

(1) Glusker, J.P. "Structural Aspects of Metal Liganding to Functional Groups in Proteins," *Adv. Prot. Chem.* **1991**, *42*, 1-76.

(2) Tainer, J.A.; Roberts, V.A.; Getzoff, E.D. "Metal Binding Sites in Proteins," *Curr. Opin. Biotechnology* **1991**, *2*, 582-591.

(3) Vallee, B.L.; Auld, D.S. "Zinc: Biological Functions and Coordination Motifs," *Acc. Chem. Res.* **1993**, *26*, 543-551.

(4) Frausto da Silva, J.J.R.; Williams, R.J.P. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*; Clarendon Press: Oxford, **1991**; pp 23-68.

(5) Sigel, H.; Martin, R.B. "Coordinating Properties of the Amide Bond. Stability and Structure of Metal Ion Complexes of Peptides and Related Ligands," *Chem. Rev.* **1982**, *82*, 385-426.

(6) Chakrabarti, P. "Geometry of Interaction of Metal Ions with Sulfur-Containing Ligands in Protein Structures," *Biochemistry* **1989**, *28*, 6081-6085.

(7) Chakrabarti, P. "Geometry of Interaction of Metal Ions with Histidine Residues in Protein Structures," *Protein Eng.* **1990**, *4*, 57-63.

(8) Chakrabarti, P. "Interaction of Metal Ions with Carboxylic and Carboxamide Groups in Protein Structures," *Protein Eng.* **1990**, *4*, 49-56.

(9) McKenzie, E.D. "The Steric Effect in Bis(2,2'-Bipyridyl) and Bis(1,10-Phenathroline) Metal Compounds," *Coord. Chem. Rev.* **1971**, *6*, 187-216.

(10) Newkome, G.R.; Gupta, V.K.; Fronczek, F.R. "Molecular Tweezer-Type. A Trigonal-Bipyramidal Copper(II) [5.8.5] Complex: Synthesis and Single-Crystal X-ray Structure Determination," *Inorg. Chem.* **1983**, *22*, 171.

(11) Prins, T.J., M.S. Thesis, "Synthesis of Unnatural Amino Acids with Chelating Side Chains and the Characterization of Peptides Incorporating Them," California Institute of Technology, **1992**.

(12) Irving, H.; Mellor, D.H. "The Stability of Metal Complexes of 1,10-Phenanthroline and its Analogues. Part I. 1,10-Phenanthroline and 2,2'-Bipyridyl," *J. Chem. Soc.* **1962**, 5222-5237.

(13) Martell, A.E.; Motekaitis, R.J. *Determination and Use of Stability Constants*; VCH Publishers: New York, **1988**.

(14) Miller, R.R.; Brandt, W.W. "Metal-Amine Coordination Compounds. III. Manganese Complexes with 2,2'-Bipyridine and 1,10-Phenanthroline," *J. Am. Chem. Soc.* **1955**, *77*, 1384-1385.

(15) Sone, K.; Krumholtz, P.; Stammreich, H. "Studies on the Coordinate Bond. III. Absorption Spectra of Mono-α,α'-dipyridyl and Mono-*o*-phenanthroline Complexes," *J. Am. Chem. Soc.* **1955**, *77*, 777-780.

(16) Bertini, I.; Luchinat, C. "High Spin Cobalt(II) as a Probe for the Investigation of Metalloproteins," *Adv. Inorg. Biochem.* **1984**, *6*, 71-111.

(17) Cotton, F.A.; Wilkinson, G. *Advanced Inorganic Chemistry*; John Wiley and Sons: New York, **1988**; pp 726-732.

(18) Bertini, I.; Luchinat, C. "Cobalt(II) as a Probe of the Structure and Function of Carbonic Anhydrase," *Acc. Chem. Res.* **1983**, *16*, 272-279.

(19) Brader, M.L.; Kaarsholm, N.C.; Lee, R.W-K.; Dunn, M.F. "Characterization of the R-State Hexamer and its Derivatives. The Hexamer is Stabilized by Heterotropic Ligand Binding Interactions," *Biochemistry* **1991**, *30*, 6636-6645.

(20) Shi, Y.; Beger, R.D.; Berg, J.M. "Metal Binding Properties of Single Amino Acid Deletion Mutants of Zinc Finger Peptides: Studies Using Cobalt(II) as a Spectroscopic Probe," *Biophys. J.* **1993**, *64*, 749-753.

(21) Huheey, J.E. *Inorganic Chemistry*; Harper and Row: New York, **1983**; pp 499-501.

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

(23) Imperiali, B.; Kapoor, T.K. "The Reverse Turn as a Template for Metal Coordination," *Tetrahedron* **1993**, *49*, 3501-3510.

(24) Smith, R.M.; Martell, A.E. *Critical Stability Constants, Vol. 2, Amines*; Plenum Press: New York, **1975**; pp 235-262.

(25) Ruan, F.; Chen, Y.; Hopkins, P.B. "Metal Ion Enhanced Helicity in Synthetic Peptides Containing Unnatural, Metal-Ligating Residues," *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404.

(26) Lieberman, M.; Sasaki, T. "Iron(II) Organizes a Synthetic Peptide into Three-Helix Bundles," *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.

(27) Huang, T.L.J.; Brewer, D.G. "An investigation of the ^{1}H nmr isotropic shifts for some methyl-substituted bipyridine complexes with Ni(II) and Co(II)," *Can. J. Chem.* **1981**, *59*, 1689-1700.

(28) Argos, P.; Gravatto, R.M.; Eventoff, W.; Rossman, M.G.; Branden, C.I. "Similarities in the Active Center Geometries of Zinc-containing Enzymes, Proteases, and Dehydrogenases," *J. Mol. Biol.* **1978**, *126*, 141-158.

(29) Greenfield, N.; Fasman, G.D. "Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation," *Biochemistry* **1969**, *8*, 4108-4116.

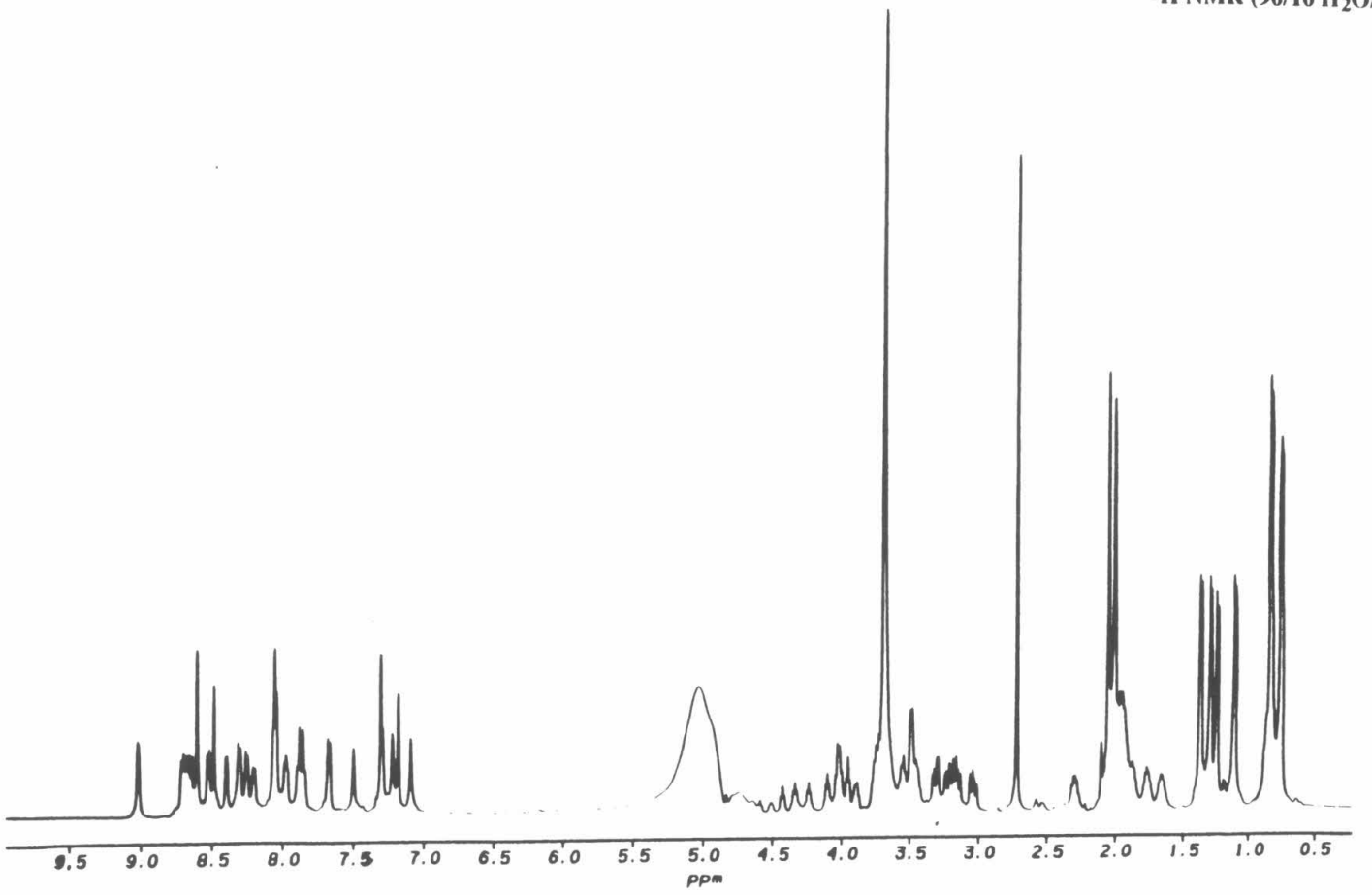
(30) Schwarzenbach, G.; Flashka, H. *Complexometric Titrations*; Interscience: New York, **1969**.

(31) Scott, R.L. "Some Comments on the Benesi-Hildebrand Equation," *Recueil* **1956**, *75*, 787-789.

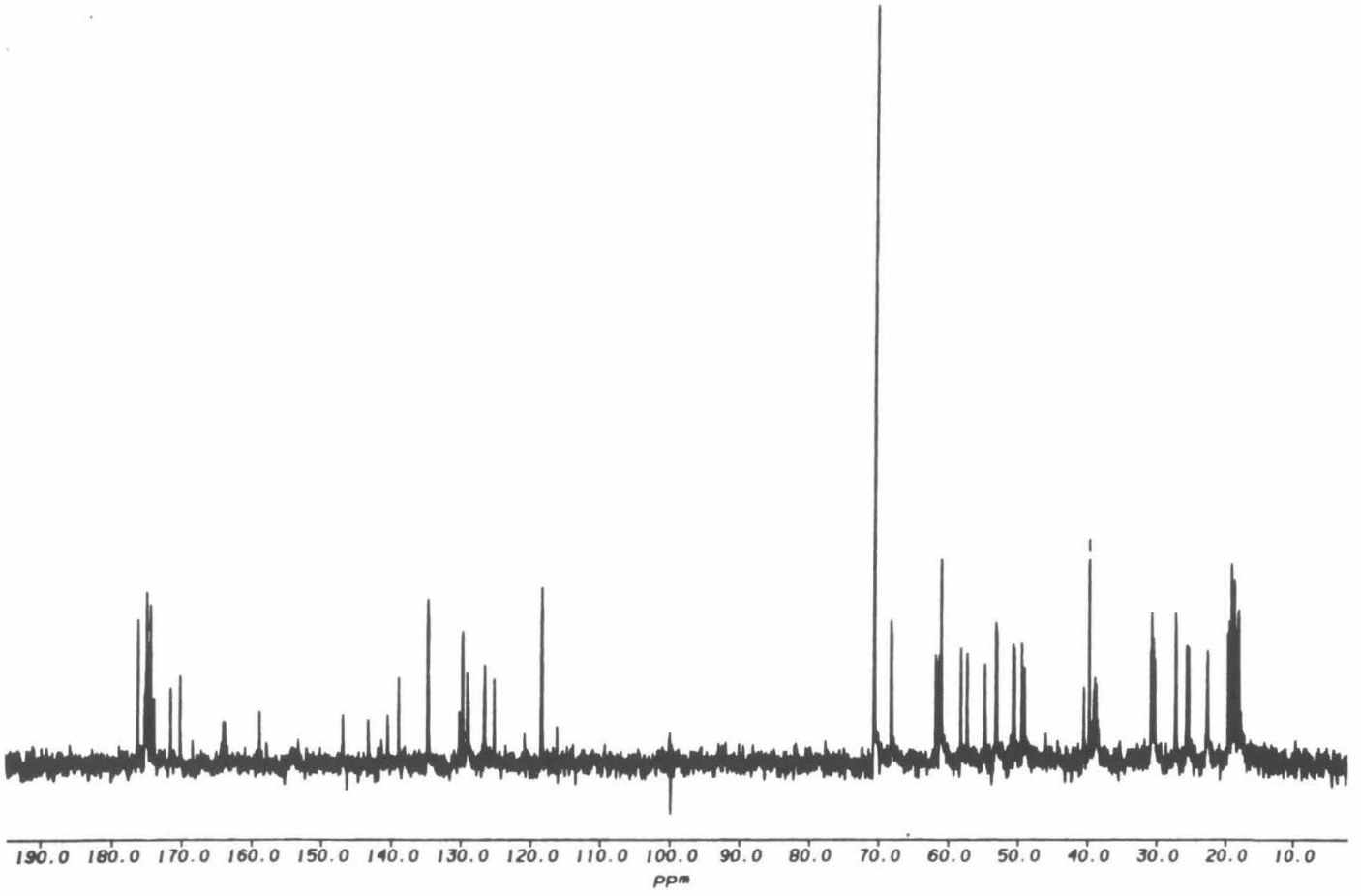
(32) Benesi, H.; Hildebrand, J.H. "A Spectrophotometric Investigation of the Interaction of Iodine with Aromatic Hydrocarbons," *J. Am. Chem. Soc.* **1949**, *71*, 2703.

Ac-2Fen-Thr-Pro-(D)Ala-Val-His-NH_2

^1H NMR (90/10 H_2O/D_2O)

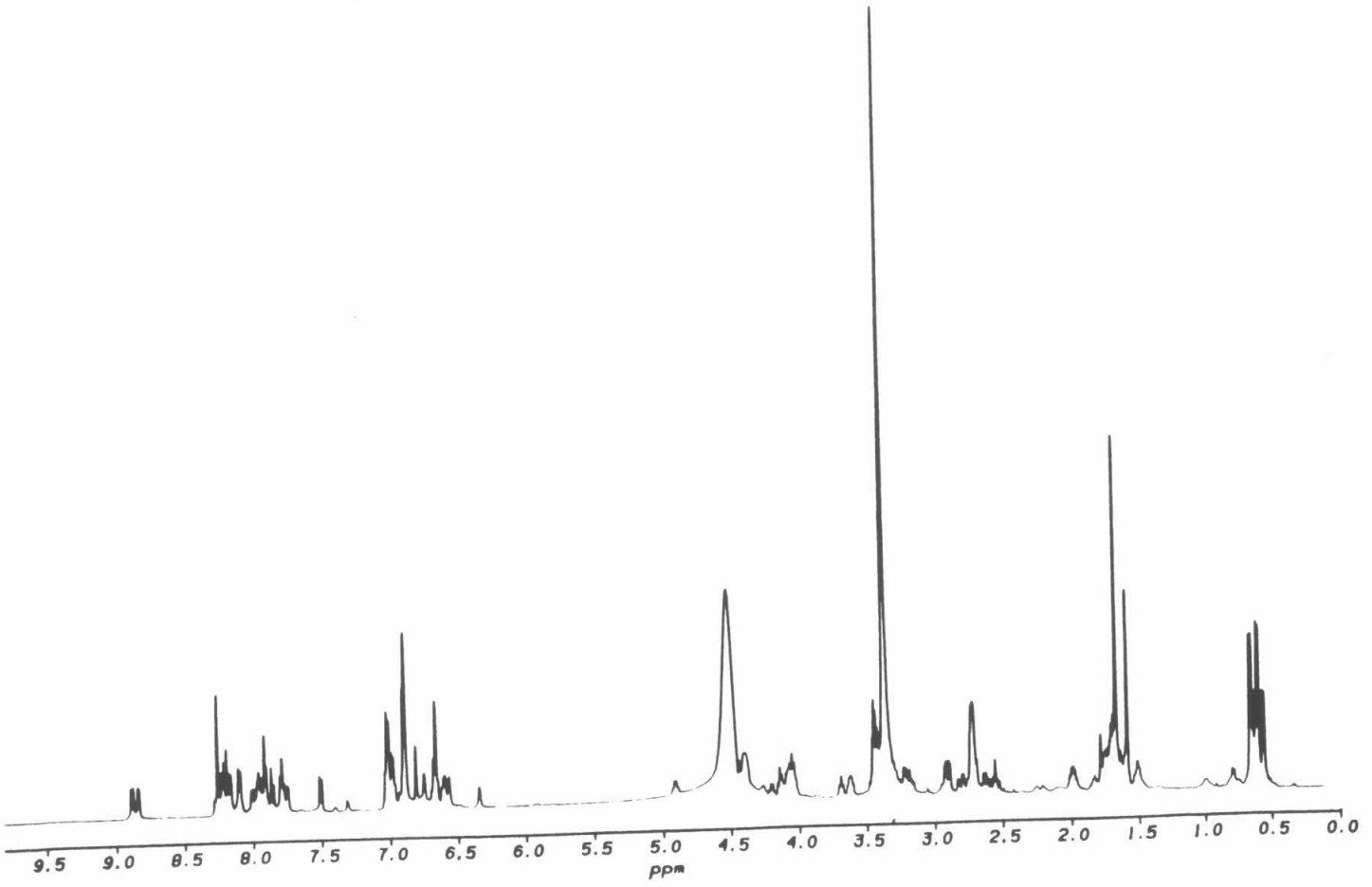


Ac-2Fen-Thr-Pro-(D)Ala-Val-His-NH_2
^{13}C NMR (90/10 H_2O/D_2O)



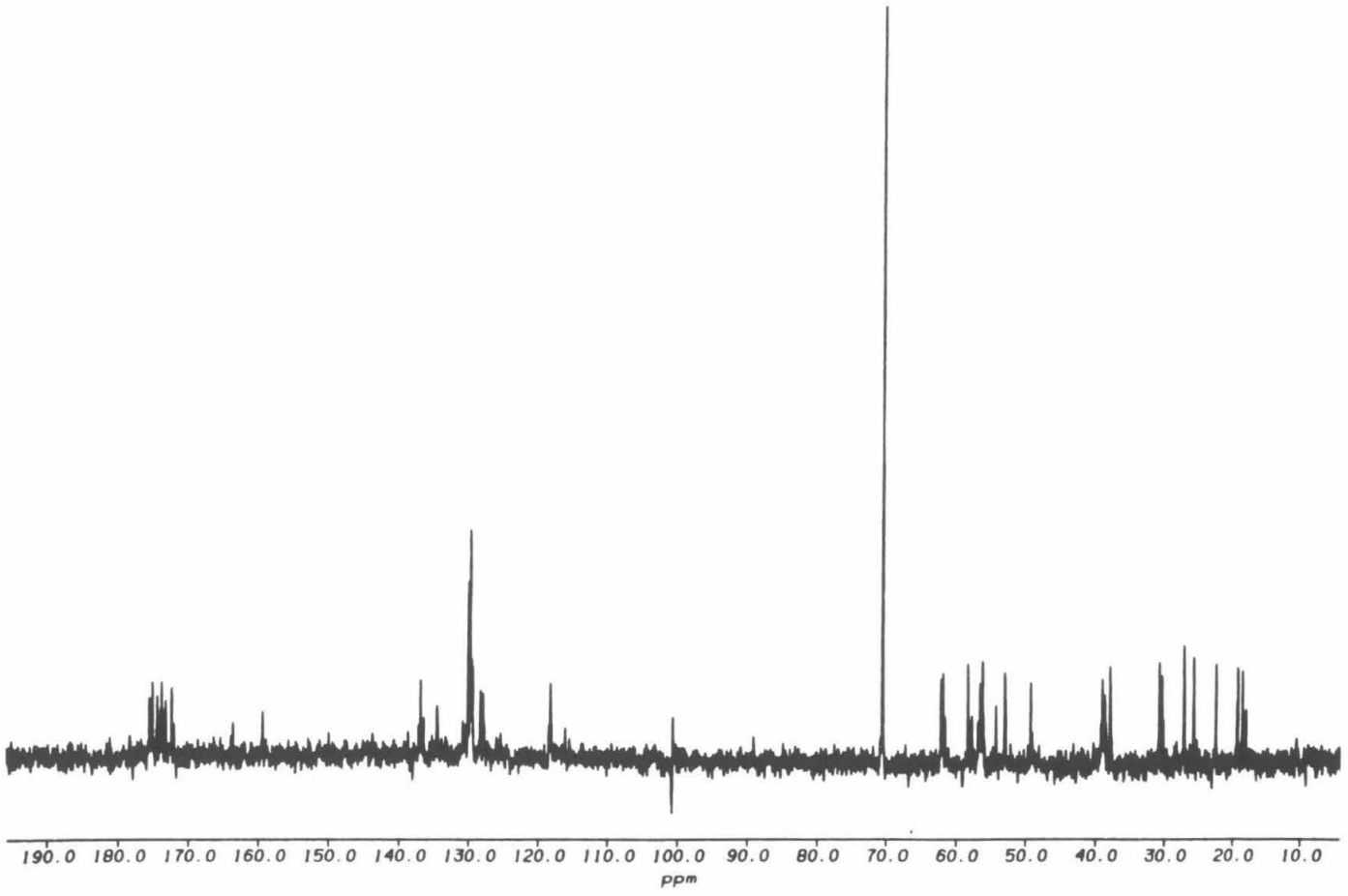
Ac-2Fen-Val-Pro-(D)Ser-Phe-His-NH_2

^1H NMR (90/10 H_2O/D_2O)

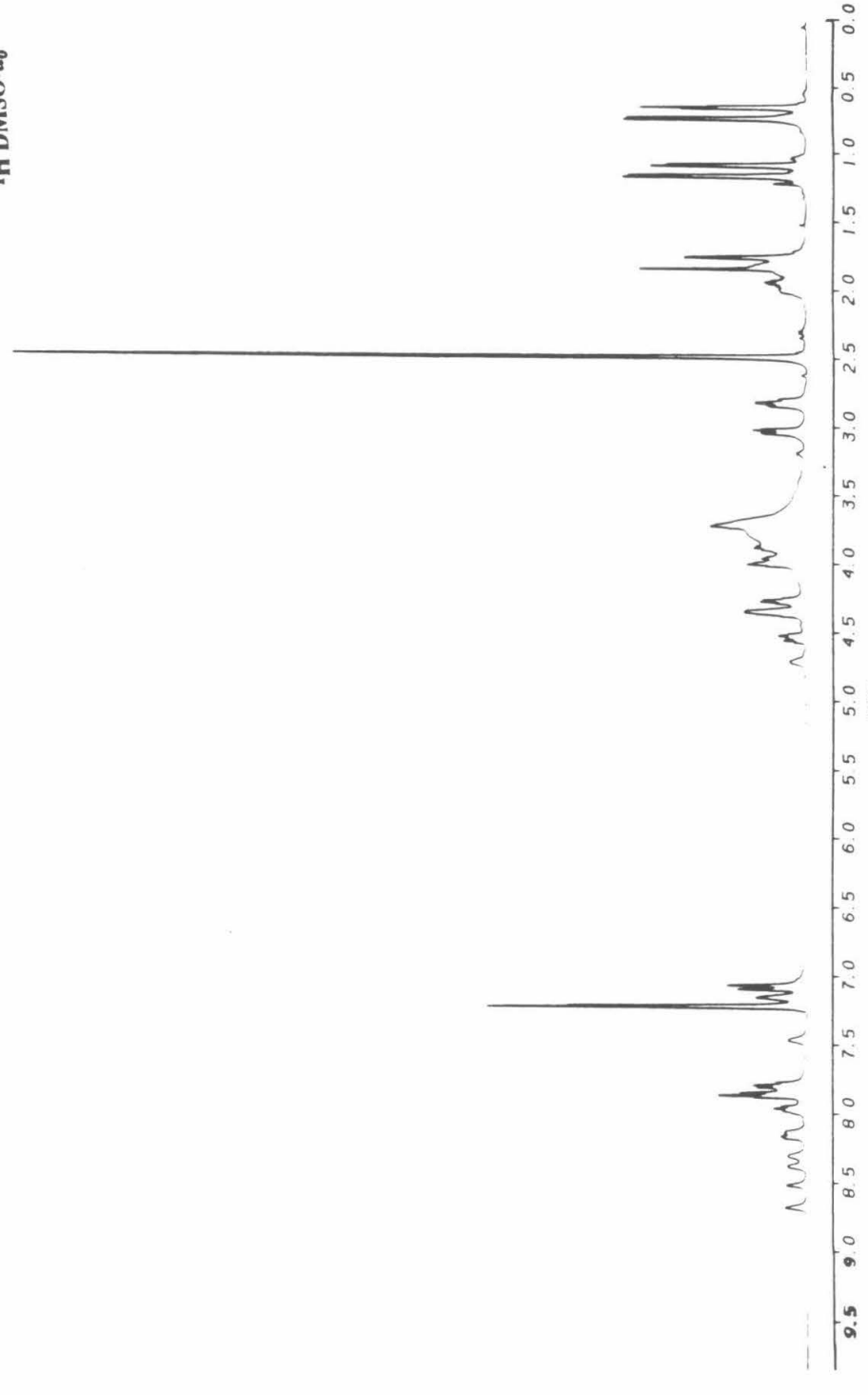


Ac-2Fen-Val-Pro-(D)Ser-Phe-His-NH_2

^{13}C NMR (90/10 H_2O/D_2O)



Ac-5Bpa-Thr-Pro-(D)Ala-Val-Phe-NH_2
^1H DMSO-d_6



Chapter 6. Semisynthesis of Bipyridyl-Alanine Cytochrome *c* Mutants: Novel Proteins with Enhanced Electron Transfer Properties

Introduction

Recent advances in protein engineering have allowed the construction of modified proteins with novel properties ranging from modified substrate specificities[1-3] or metal-dependent regulated activities[4,5] to enhanced structural stabilities resulting from metal cation binding at sites engineered on the protein surface.[6] Although site-directed mutagenesis has allowed specific alteration of protein architecture, it is a technique that relies upon the cell's biosynthetic machinery, and is therefore limited to the encoded amino acids. It would be advantageous to include unnatural amino acids into protein ensembles to augment existing structures, or to introduce new functionalities altogether. Both *in vitro* expression of chemically acylated suppressor *t*RNA[7-11] and total chemical synthesis[12] methods have been employed to incorporate unnatural amino acids into protein sequences. However, the former technique suffers from low throughput of protein, while the latter is limited to sequence lengths that can be built by employing solid phase peptide synthesis technology.

In contrast, the semisynthetic methodology,[13-15] in which fragments of native protein are fused with synthetically prepared peptides to form complete protein sequences, can provide a means for facile generation of proteins containing unnatural amino acids in large quantities. To date, the semisynthetic technique has been limited to those proteins that are naturally amenable to facile cleavage and religation; however, the use of genetic engineering should provide a means for more general application of semisynthesis in the future. In addition, the recent developments in synthetic methods for production of novel, optically pure amino acids[16] and advances in solid phase peptide synthesis [17] now allow the assembly of peptides in essentially infinite variety. With these tools in hand, the stage is now set for the general manipulation of proteins well beyond the constraints set by nature.

Semisynthetic methods have been employed with great success for the creation of novel cytochrome *c* (cyt. *c*) mutants in relatively large (milligram) quantities. The

preparation of these mutants is generally accomplished through the coupling of two protein segments representing residues 1-65 and 66-104 of the complete horse heart cyt. *c* primary sequence as shown in Figure 6-1.[18,19] The *N*-terminal fragment (residues 1-65), which contains the covalently bound heme, is generated from the native protein through cyanogen bromide cleavage of native horse heart cyt. *c* at Met65 under denaturing conditions, followed by purification using cation-exchange chromatography to afford a homoserine (Hse) lactone at the carboxy terminus.[20] The corresponding polypeptide consisting of residues 66-104 can be prepared by solid phase peptide synthesis, thereby allowing the incorporation of both natural and unnatural amino acids anywhere within this region of the protein sequence, with the caveat that the protein must be able to correctly fold to the native structure in the presence of these mutations. Reconstitution of the protein is accomplished through incubation of the two fragments under neutral reducing conditions, which promotes the reaction of the 66-104 amino terminus with the Hse65 lactone to form a peptide bond. The fully-formed protein can then be purified using standard chromatographic techniques. These methods have been used to prepare a wide variety of cytochrome *c* mutants which contain both natural and unnatural amino acids; several notable examples include a mutant containing *p*-fluorophenylalanine at position 67[21,22] and mutants incorporating a variety unnatural metal ligating residues at position 80 to study the effects of axial ligation on the protein-bound heme group.[23,24]

To date, Gray and co-workers have used cytochrome *c* mutants modified with redox-active metal centers to study the factors that govern intraprotein electron transfer (ET).[25,26] These studies have relied on the semisynthetic technique for site-specific incorporation of histidine residues on the protein surface, which were subsequently modified with $Ru(bpy)_2CO_3$ to afford $Ru(bpy)_2(im)$(His)-modified proteins. These proteins are capable of efficient photoinduced ET from an electronically excited ($^*Ru^{2+}$) site to the heme-bound Fe-center. The rates of intraprotein ET between the two metal sites have been measured for these mutants[27] and the results from these studies, in conjunction

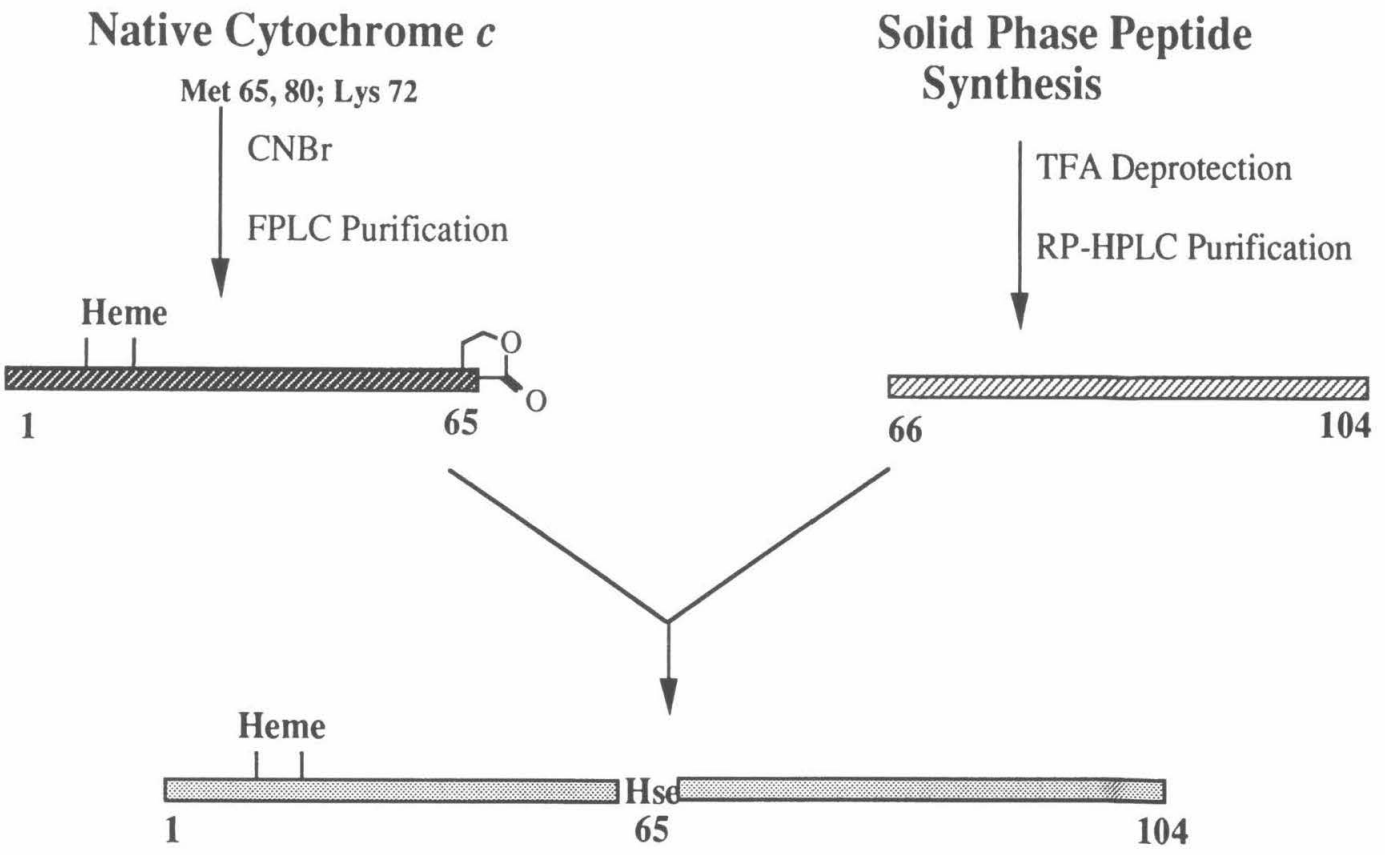


Figure 6-1. Schematic for the semisynthetic preparation of horse heart cytochrome *c* mutants

with electron transfer measurements taken for a number of similarly prepared Ru-modified proteins,[26,28,29] have been used to correlate the observed electron transfer rates with the effective electronic coupling term (H_{ab})[30] between the metal centers using a tunneling-pathway dependent model.[31]

In an effort to demonstrate the ability of the bipyridyl-substituted amino acids (4Bpa and 6Bpa, See Chapter 4) to play *functional* roles in designed protein ensembles, a collaborative project was undertaken with Deborah S. Wuttke and Professor Harry Gray which involved the semisynthetic construction of two horse heart cyt. *c* mutants incorporating 6Bpa and 4Bpa directly into polypeptide backbone at position 72 of the native protein sequence. The bipyridyl amino acid mutations allow for the assembly of a $Ru(bpy)_3^{2+}$ center[32] at a defined site on the protein surface, which is expected to increase markedly the efficiency of photoinduced ET from an electronically excited ($^*Ru^{2+}$) surface site to the protein metal center relative to that of a $Ru(bpy)_2$(im)(His)-modified protein.[27,28]

The $Ru(bpy)_3^{2+}$ redox center has been used extensively in electron transfer studies, including models based on polypeptide[33-35] and protein systems.[36-41] In general, these models have been prepared through chemical modification of naturally occurring amino acids with reactive side chain functional groups (e.g., lysine, cysteine). As a result, these model systems are limited not only in the selection of suitable sites for label incorporation, but also hampered by extensive purification problems arising from indiscriminate modification of the protein surface sites. In contrast, the bipyridyl amino acids are ideal for building versatile electron-transfer (ET) model systems, since they can be placed site-specifically anywhere on the polypeptide backbone using solid-phase peptide synthesis (see Chapter 5). The choice of position 72 of the cytochrome *c* protein sequence as the site for the bipyridyl-alanine mutations was based on the successful construction of the $Ru(bpy)_2$(im)(His72)-modified cyt. *c*.[42] Since no structural perturbations were observed with this mutant,[43] it was envisioned that incorporation of the bipyridyl sidechain of the amino acid, as well as the subsequent $Ru(bpy)_3^{+2}$ center formed after metal modification,

would be well tolerated by the protein structure. In addition, the two bipyridyl-substituted mutant proteins prepared differ only in the regioisomer of the bipyridyl amino acid employed, thereby allowing a sensitive probe of the effects of ligand orientation on metal coordination.

Results and Discussion

Semisynthesis of 6Bpa 72 and 4Bpa72 cyt. c

The C-terminal fragments (residues 66-104) of horse heart cytochrome *c* incorporating the bipyridyl-alanine mutations at position 72 were synthesized using standard solid phase peptide synthesis protocols with N^{α}-FMOC protection strategies with either HOBT/BOP or pentafluorophenyl OPFP (or ODhbt)/HOBT mediated amide bond coupling chemistries. In order to maximize the yield of the full length peptide, all of the amino acids were run in high molar excess (4.0-8.0 amino acid equivalents per coupling) and in cases where the coupling efficiencies were found to be poor (as determined by ninhydrin tests performed on the resin)[44] the amino acids were subjected to double or triple acylation cycles. Further, in an effort to prevent the generation of deletion peptides, the resin was subjected to an acetyl capping protocol with an acetic anhydride/HOBT solution

NH_2–Glu–Tyr–Leu–Glu–Asn–*Pro*–*Bpa*–Lys–Tyr–*Ile*–Pro–Gly–*Thr*–Lys–
66 70 75

Met–*Ile*–Phe–Ala–Gly–*Ile*–*Lys*–**Lys**–*Lys*–**Thr**–Glu–*Arg*–Glu–Asp–Leu–
80 85 90

Ile–Ala–Tyr–Leu–Lys–Lys–Ala–*Thr*–Asn–Glu–COOH
95 100

Figure 6-2. Schematic of 4(6)Bpa72 cyt. *c* 66-104 protein fragment synthesis strategy - (Amino acid residues shown in italics represent double coupled residues, those shown in bold were triple coupled, numbering corresponds to horse heart cyt. *c* sequence.)

prior to the N^{α}-deprotection step of each cycle. The overall synthetic strategy for the peptide synthesis is shown schematically in Figure 6-2.

Cleavage of the peptides from the resin and subsequent workup proceeded smoothly, and although the crude material exhibited an envelope of peaks, two major products of approximately equal size were clearly resolved in the mixtures of both the 6Bpa and 4Bpa peptide syntheses (see Figure 6-3, data shown for the 4Bpa72 peptide). Attempts to purify the crude materials using gel filtration chromatography resulted in a substantial enhancement of the overall peptide purity; however, the two major peaks in each case eluted at the same void volume suggesting that they were of similar molecular weight. The two major products were further purified to homogeneity using semi-preparative RP-HPLC, and mass spectral analysis of the purified peptides indicated that one of the products had the correct molecular weight for the full-length peptide (MW = 4670 amu) and the other was 16 amu higher than the expected molecular weight. (See Figure 6-4, data shown for the 4Bpa72 peptide reflect the results obtained for the 6Bpa72 peptide.) This difference was attributed to the oxidation of the Met80 thioether to the corresponding sulfoxide during the peptide cleavage step. Methionine sulfoxide formation is well precedented in solid phase peptide synthesis,[45] and there is evidence that the oxidation can be catalyzed by metal cations.[46] It is possible that the bipyridyl ligands act as a carrier for trace amounts of metal cations which in this case promoted the formation of the sulfoxide form. Metal chelation by the peptides is evident in the mass spectra of the purified 4Bpa containing peptide fragments (Figure 6-4b,d), as a higher molecular weight component (corresponding to ≈60 amu)[47] is observed for both the reduced and oxidized forms of this peptide.

A number methods are known that selectively reduce alkyl sulfoxides to the corresponding thioether; for example, mercaptoethanol,[48,49] dithiothreitol,[48,49] and dimethylsulfide in HCl[50] have all been shown to selectively reduce methionine sulfoxide residues in peptides to methionine. Attempts to reduce the 66-104 protein fragments using

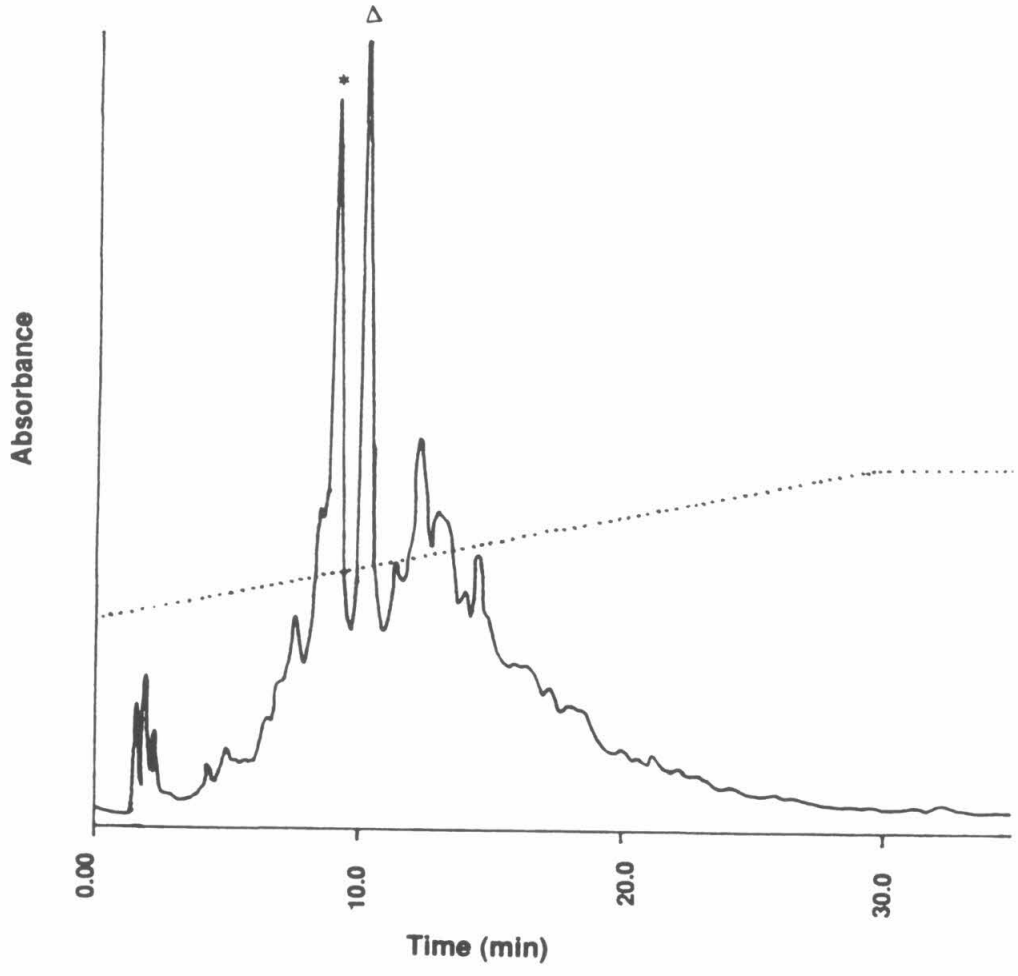


Figure 6-3. Reversed-phase high pressure liquid chromatogram of the crude 4Bpa72 cyt. *c* protein fragment 66-104. (* = Peak at 9.05 min, Met80 sulfoxide peptide, Δ = peak at 10.2 min, Met80 thioether; dotted line represents the percentage of CH_3CN (0.08 % TFA) in eluent (H_2O 0.1% TFA). The chromatogram was collected on a C18 Analytical Column with 228 nm detection at a flow rate of 1 mL/min).

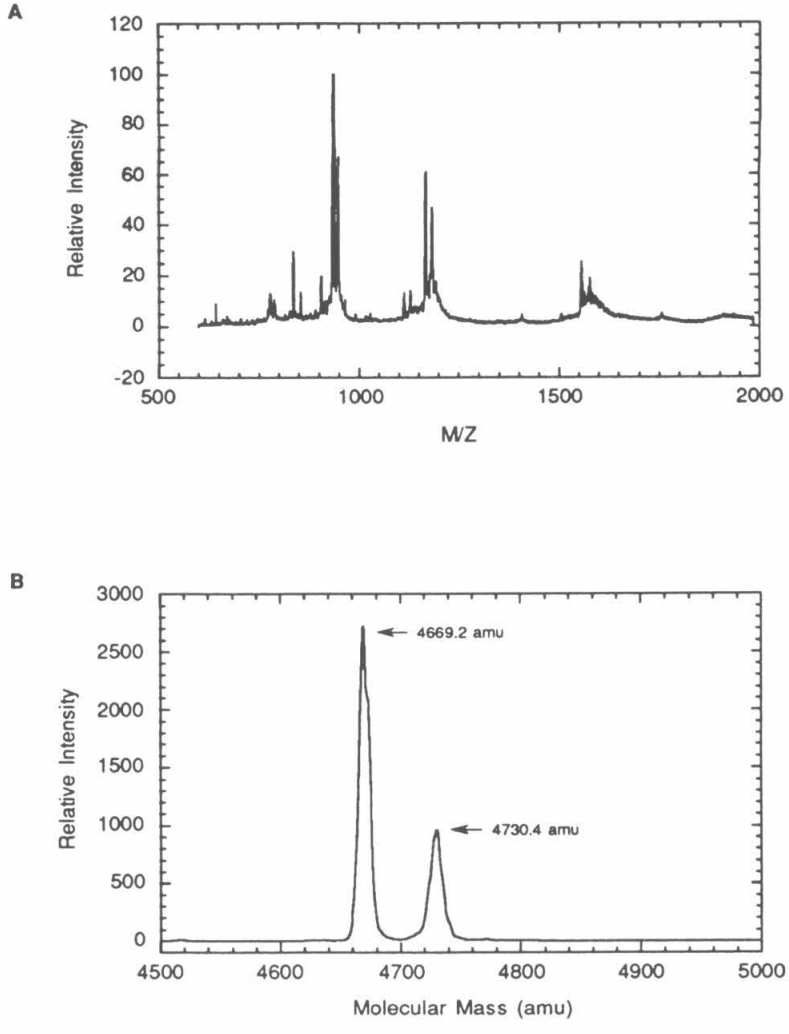
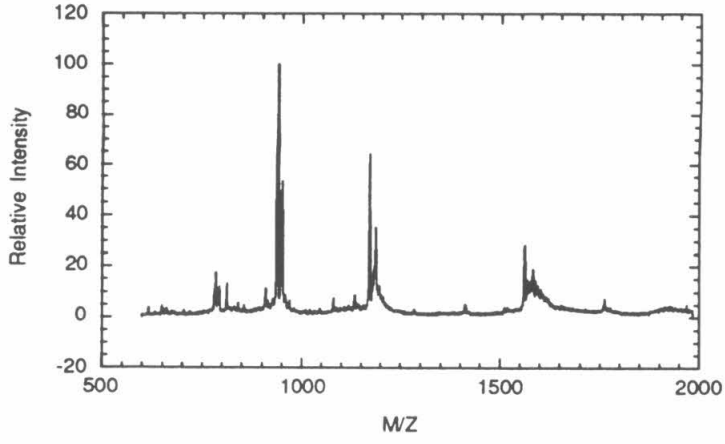
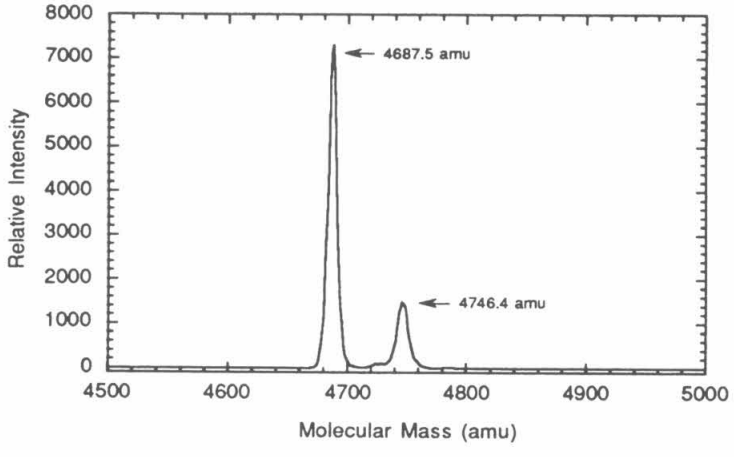


Figure 6-4. Electrospray mass spectra for 4Bpa72 cyt. *c* protein fragment 66-104 Met80 thioether (raw (A) and deconvoluted (B) data) and Met80 sulfoxide (raw (C) and deconvoluted (D) data). Calculated molecular weight for the reduced, full-length peptide is 4671.4 amu.

C


120
100
80
60
40
20
0
-20
Relative Intensity
500
1000
1500
2000
M/Z

D


8000
7000
6000
5000
4000
3000
2000
1000
0
Relative Intensity
4687.5 amu
4746.4 amu
4500
4600
4700
4800
4900
5000
Molecular Mass (amu)

β-mercaptoethanol or dimethylsufide were largely successful as monitored by RP-HPLC; the peak corresponding to the sulfoxide peptide decreased over time with a concomitant increase in peak assigned to the reduced peptide. The extent of the sulfoxide conversion depended upon the reagent used and the conditions employed, but optimal results were found with excess dimethylsulfide in 6.0 N HCl, which quantitatively reduced the sulfoxide peptide to the thioether of the 4Bpa-containing peptide within 1 h. Similarly, complete conversion was observed under neutral conditions using excess β-mercaptoethanol (20 molar equivalents) over 4 days at 37°C.

Even though it was possible to convert the sulfoxide form to the methionine thioether, the purification of the full length peptides was plagued with solubility problems associated with the purified peptide fragments. While the crude peptide readily dissolved in both aqueous and mixed organic/aqueous systems (H_2O/CH_3CN, H_2O/MeOH, DMF/H_2O), the peptides were largely insoluble under any of these conditions after subsequent lyophilization of the purified peptide fractions. It was assumed that the peptides formed higher order aggregates upon concentration under neutral conditions, but attempts to prevent aggregate formation by acidification of the peptidyl fractions prior to lyophilization met with little success. Concentration of peptide fragment solutions using ultrafiltration membranes (Centricon-3) was also investigated; however, peptide flow-through presented problems in recovery of the peptide. In light of these difficulties, reconstitution reactions to form the native protein were performed using crude peptide material (see below).

Automated Edman degradation peptide sequencing[51] was attempted to further characterize the peptides. Sequencing of the peptides followed typical protocols for all of the residues preceding the bipyridyl amino acids; however, efficient amide bond cleavage was not observed for either of the 4Bpa or 6Bpa residues. Instead, a slow, prolonged release of the phenylthiohydantoin (PTH) derivative of the bipyridyl amino acids was observed for several cycles following position 72, and interpretation of the results for

subsequent residues was complicated by the degradation products arising from multiple peptide chain lengths. As a result, characterization of the peptides relied upon the mass spectral analysis and by inference from the extensive characterization of the semisynthetic proteins generated from the peptide fragments (see below).

Preparation and purification of the native protein fragment (residues 1-65) was performed by Deborah S. Wuttke in the laboratories of Professor Harry B. Gray.[43] The fragment was prepared using cyanogen bromide cleavage of commercially available horse heart cytochrome *c*, and subsequently purified using cation-exchange FPLC chromatography, to provide the purified heme-containing fragment with the homoserine lactone (Hse65) intact.

Reconstitution of the protein was effected through incubation of the two purified fragments under neutral reducing conditions, thereby allowing reaction of the Hse65 lactone and the amino terminus of 66-104, to form a peptide bond. The renatured, fully-formed protein was then purified using cation-exchange chromatography. The reconstitution reaction was found to be remarkably selective for the native protein sequence of the 66-104 fragment. Attempts to react the crude peptide 66-104 with purified fragment 1-65 resulted in similar overall yields of the fully formed protein in slightly lower purity. It is noteworthy that attempts to couple the purified Met80-sulfoxide 66-104 fragment with the purified 1-65 fragment did not result in any detectable protein formation. These results demonstrate that the religation reaction is so specific that protein products are not obtained when the primary sequences differ only by the oxidation state of one residue. This selectivity suggests that preorganization of the two peptide fragments through the formation of favorable secondary structure interactions is required for efficient amide bond formation.

Characterization of 4(6)Bpa72 cyt. c Mutants

The mutant proteins were characterized by several independent methods. Absorption spectra of the 4(6)Bpa72 cyt. *c* overlay nearly identically with the native cyt. *c*

spectra (see Figure 6-5a) for both the ferric and ferrous states. An increased absorption at 290 nm was observed as expected with the addition of a bipyridine absorption (ε_{bipy}=12800 M^{-1} cm^{-1}, $\varepsilon_{cyt\ c}$= 19500 M^{-1} cm^{-1}). The spectra of the 4(6)Bpa72 cyt. *c* proteins exhibit an absorption at 695 nm identical to that of the native protein indicating that the Met80 thioether has the correct ligation to the heme (see Figure 6-5b). The circular dichroism spectra of the mutants and the native protein in the far UV region are superposable (see Figure 6-6a), and the melting points of the proteins, determined from thermal denaturation curves derived using variable temperature CD studies, are virtually identical for the native and semisynthetic proteins (See Figure 6-6b). The reduction potentials of the mutant proteins, measured by differential pulse polarography, were found to be nearly identical to the native protein ($E°_{native}$= 0.265(5) V, $E°_{4Bpa72cyt.c}$= 0.265(5) V, $E°_{6Bpa72cyt.c}$= 0.258(5) V vs NHE)[43] and the *g* values measured from the EPR spectra of the mutant proteins are similar to those measured for the native protein,[43] indicating that the heme environments and ligand geometries in the 4(6)Bpa72 cyt. *c* proteins are structurally equivalent to the native protein. High resolution electrospray mass spectrometry of the mutant proteins produced peaks at ≈12445 amu, as opposed to the native protein which produced a peak at 12385 amu (a nominal difference of 67 amu is expected for the bipyridyl mutation and homoserine65). Both the raw and the deconvoluted mass spectral data are shown in Figure 6-7; the presence of the smaller molecular weight peaks is attributed to deamidated species. The proteins were also characterized using two electrophoretic techniques, isoelectric-focusing (IEF) and SDS-PAGE. IEF gel analysis shows the mutant proteins to have a slightly lower isoelectric point relative to the native protein, consistent with the substitution of bipyridine for lysine at position 72. The gel shifts from SDS-PAGE were consistent with the proposed molecular weight and charge. In summary, the results from these experiments demonstrate that incorporation of the bipyridyl amino acids does not perturb the three-dimensional structure or the thermodynamic stability of the protein. A model of the bipyridyl-substituted proteins,

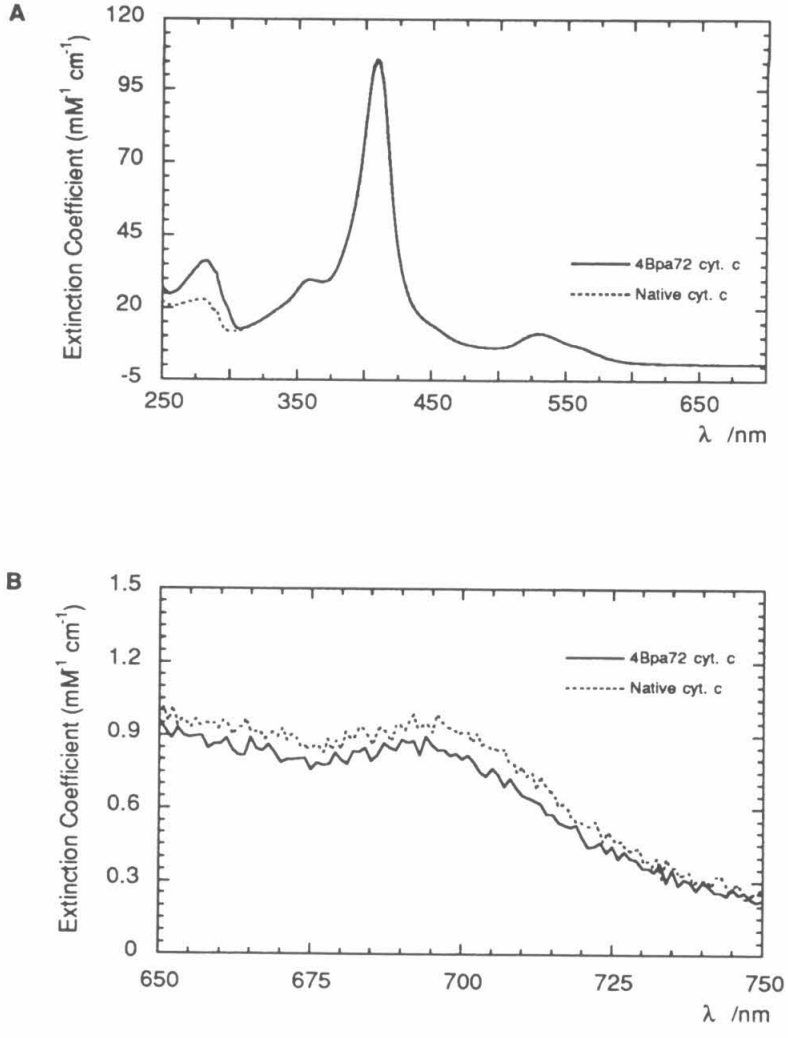
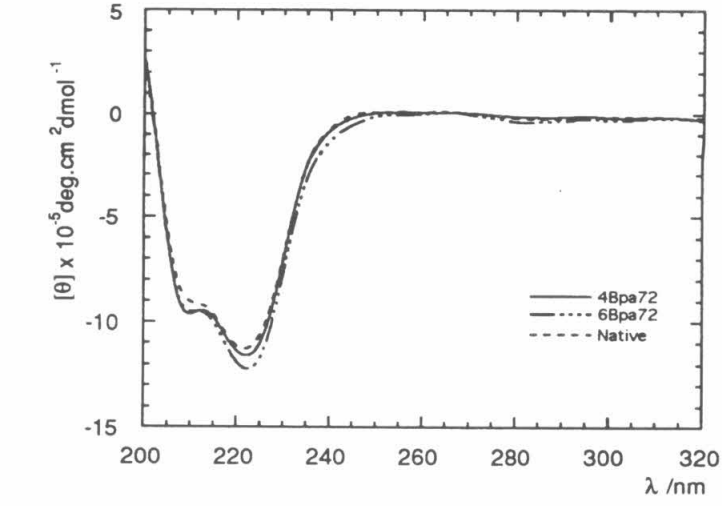


Figure 6-5. Electronic absorption spectra of native and 4Bpa72 cyt. *c*. The increased absorbance at 290 nm for the 4Bpa72 cyt. c corresponds to the bipyridine absorption. The absorption spectrum for the 6Bpa72 cyt. c was identical to that observed for the 4Bpa72 cyt. c. (B) Expansion plot demonstrating the 695 nm absorption band arising from Met80 ligation to the heme group.



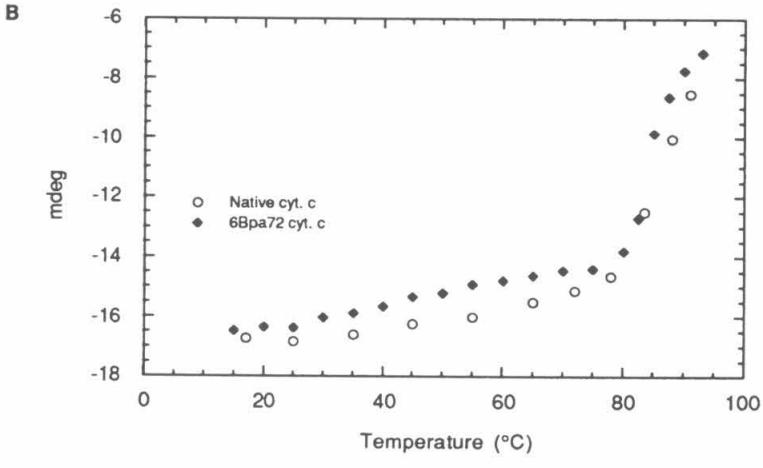


Figure 6-6. (A) Circular dichroism spectra of native and semisynthetic cyt. *c* proteins. (B) Thermal denaturation curves of the native and 6Bpa72 cyt. *c* from variable temperature CD studies monitored at 222 nm. Spectra were recorded in a 0.10 cm water-jacketed cell in 50 mM sodium phosphate buffer (pH 7.0) at 25°C.

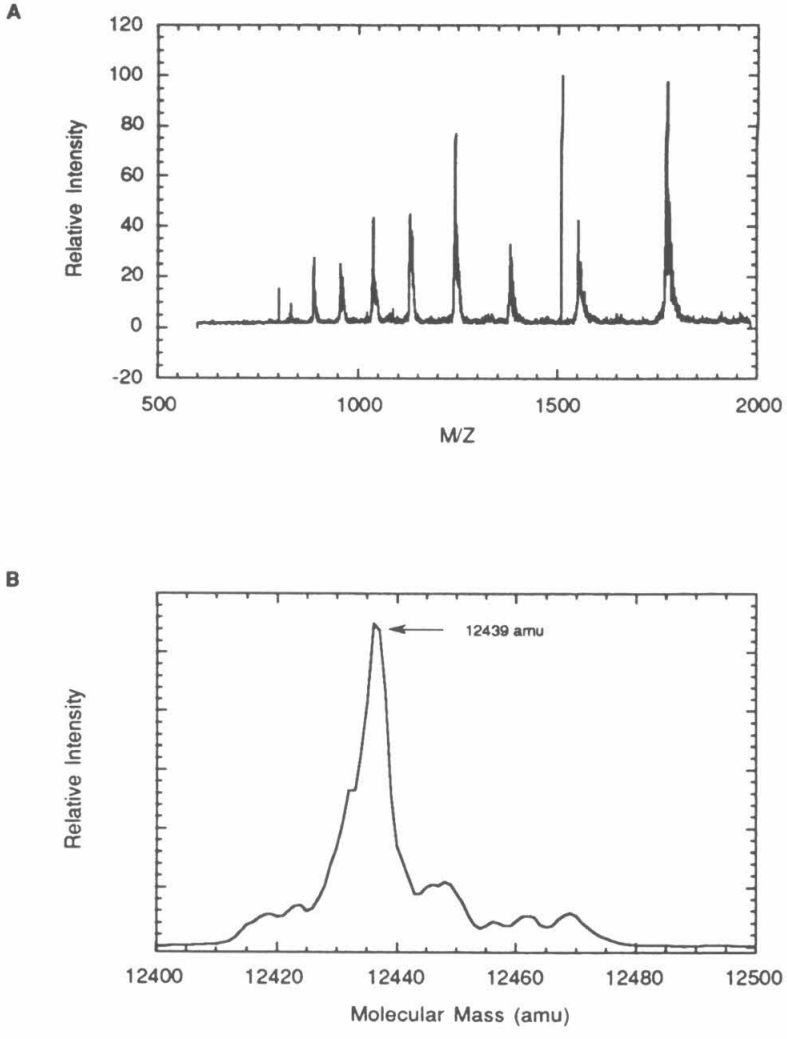
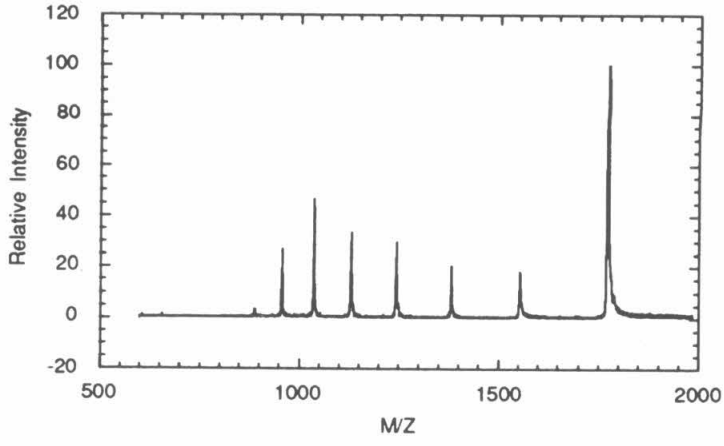


Figure 6-7. Electrospray mass spectra for the 4Bpa72 cyt. *c* (raw data (A), deconvoluted data (B)) and the 6Bpa72 cyt. *c* (raw data (C), deconvoluted data (D)). Spectra were collected in 20 mM ammonia bicarbonate buffer (pH=7.6) diluted with 1N acetic acid.

C


120
100
80
60
40
20
0
-20
Relative Intensity
500
1000
1500
2000
M/Z

D

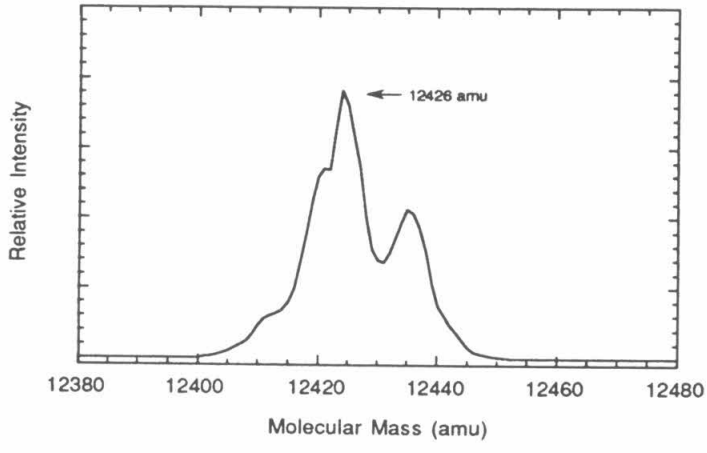
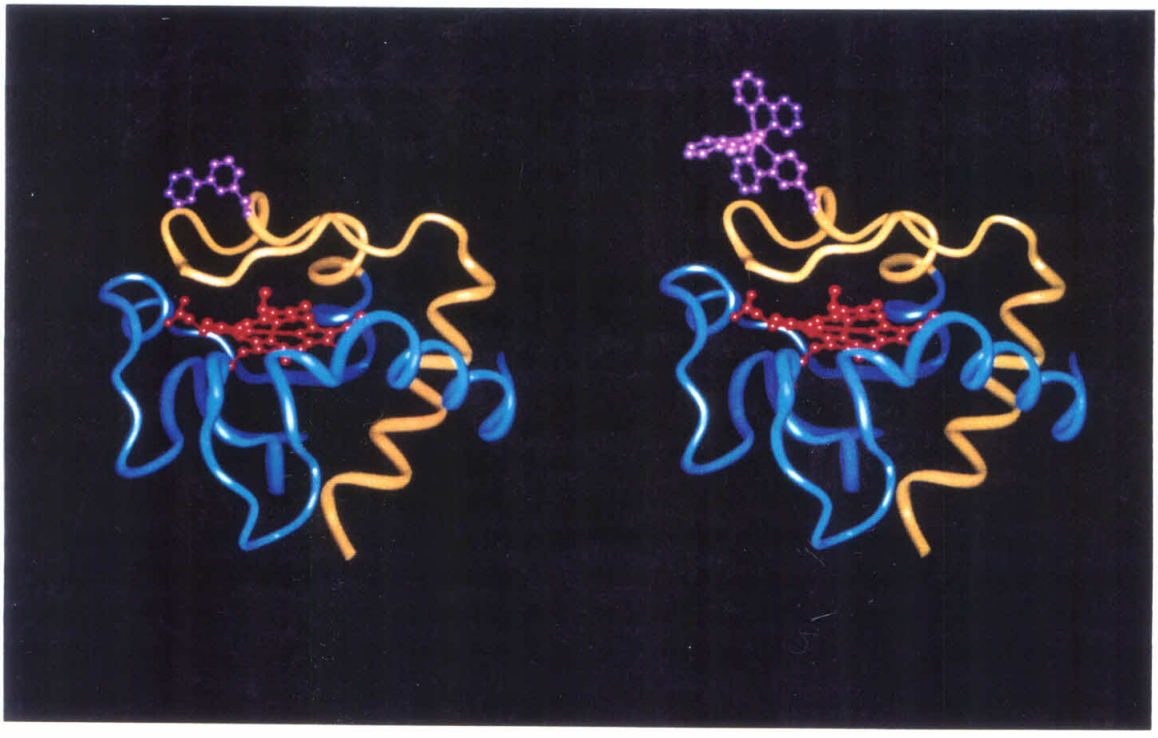

12426 amu
Relative Intensity
12380
12400
12420
12440
12460
12480
Molecular Mass (amu)

Figure 6-8. (A, left) Model of 4Bpa72 cyt. *c* based on the crystal structure of ferric horse heart cyt. *c* .[52] The blue tube is the 1-65 native fragment backbone; the yellow tube is the 66-104 backbone constructed by solid phase peptide synthesis; red indicates the heme group; and purple is the 4(6)Bpa72 residue. (B, right) Model of $Ru(bpy)_2$(4Bpa72)cyt. *c* with purple indicating the $Ru(bpy)_3^{2+}$ unit.



based on the crystal structure coordinates, is shown in Figure 6-8.

$Ru(bpy)_2^{2+}$ Modification of 4Bpa72 cyt c

Modification of the 4Bpa72 protein with a redox-active unit was achieved via incubation with excess $Ru(bpy)_2CO_3$.[28,43,53] The modification reaction was monitored by absorption spectroscopy and terminated by addition of excess imidazole and gel filtration when a $Ru(bpy)_3^{2+}$ to heme ratio of 1:1 was obtained. The yield of Ru(II)$(bpy)_3$-4Bpa72 cyt. *c* was fairly low (≈20 %) and the reaction proceeded slower than anticipated; as a result, the reaction mixture produced a variety of singly and doubly modified proteins, with the singly modified His33 protein (identified by FPLC cation-exchange retention time)[28] representing the predominant side product. These products were purified by cation-exchange chromatography, and the 4Bpa72-modified protein was unambiguously distinguished from the His33-modified form[28] by absorption and emission spectroscopy. The absorption spectrum of the Ru(II)$(bpy)_3$-4Bpa72 cyt. *c* corresponds to the sum of the native protein and the $Ru(bpy)_3^{2+}$ chromophore as demonstrated by the enhanced absorbances (280, 492 nm) relative to the native protein (see Figure 6-9a). Similarly the emission spectra correspond to the $Ru(bpy)_2(4Bpa)^{+2}$ complex [$Ru(bpy)_2(4Bpa)^{2+}$, λ_{max}(abs) = 452 nm; λ_{max}(em) = 615 nm (uncorrected); $Ru(bpy)_2(im)(His)^{2+}$, λ_{max}(abs) = 492 nm; λ_{max}(em) = 670 nm (uncorrected)].[43] The CD spectra of $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c* is nearly identical to the native protein in the far UV region, indicating that the Ru-modification has a negligible effect on the overall protein structure (See Figure 6-9b). The ellipticity observed in the near UV region is consistent with the Λ-isomer; evidently the protein environment surrounding the metal center is more compatible with this isomer of the metal complex. Based on these results, a model of the $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c* was generated based on the crystal structure coordinates of the native protein and is shown in Figure 6-8b.

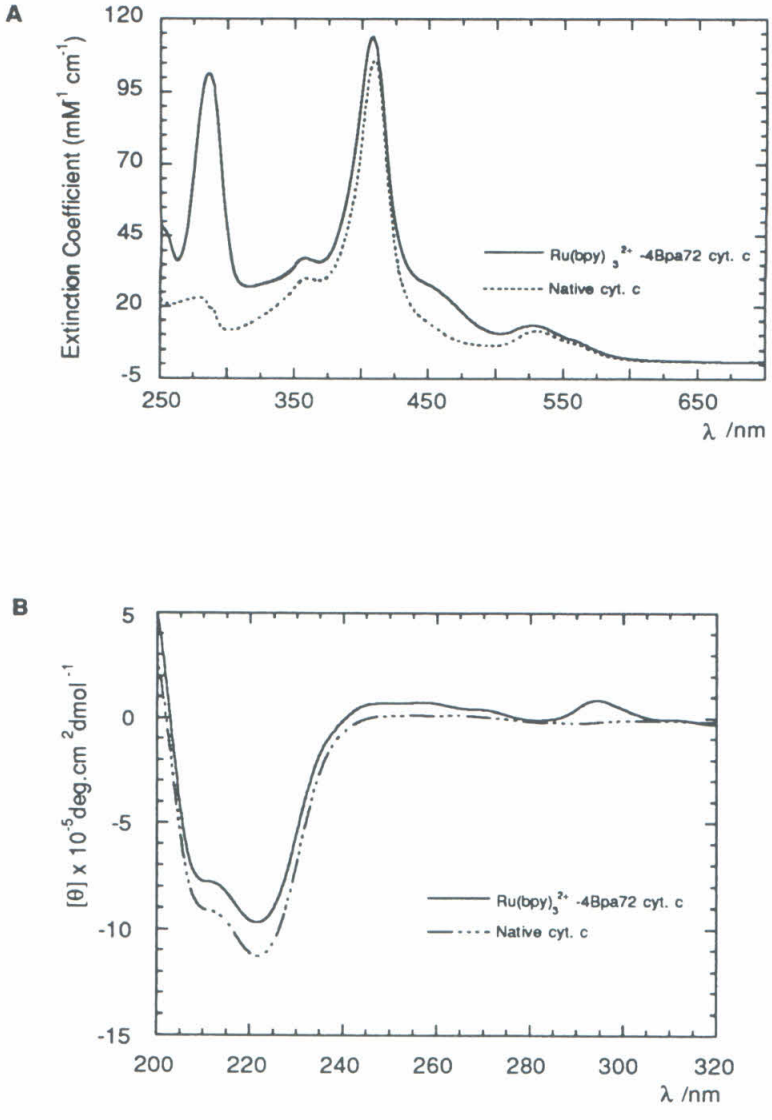
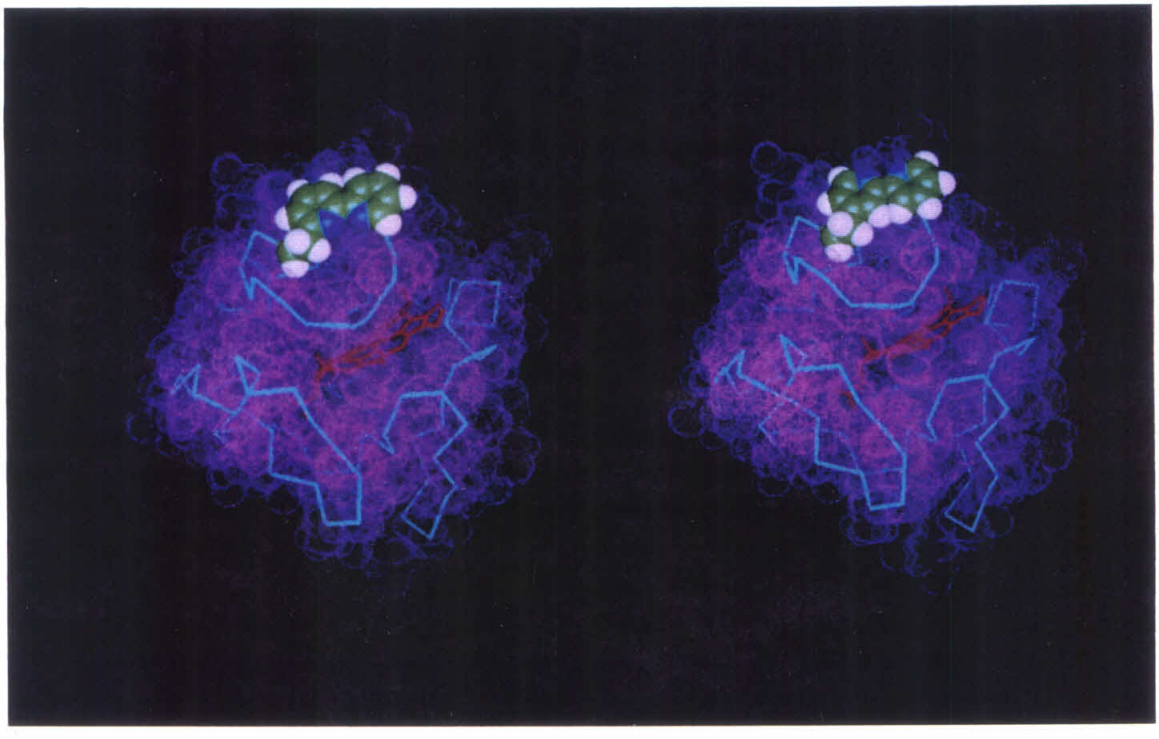


Figure 6-9. Electronic absorption (A) and circular dichroism (B) spectra of the $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c* complex. (Spectra of the native protein is shown for comparison.) Spectra were recorded as above for the unmodified proteins.

Interestingly, only 4Bpa72 cyt. *c* was modified at the surface bipyridyl residue; attempts to modify the 6Bpa72 cyt. *c* under the conditions used for the 4Bpa72 protein did not result in any detectable formation of the $Ru(bpy)_3^{+2}$-6Bpa72 protein. These results were supported by molecular modeling studies of the two regioisomers based on the crystal structure coordinates of the native protein.[52] Ramachandran plots of the χ_1 and χ_2 dihedral angles indicated that the minimum energy conformations of the bipyridyl sidechains oriented the chelating nitrogens towards the protein core in the 6Bpa protein, whereas in the 4Bpa72 cyt. *c* the metal-ligating groups of the bipyridyl moiety were more accessible to solvent (see Figure 6-10). Further studies using peptide models confirmed that the difference in reactivities observed for the two proteins could not be attributed to an intrinsic property of the bipyridyl ligand regioisomers.

Two peptides of the general sequence Ac-Xaa-Thr-Pro-D-Ala-Val-Phe-NH_2, where Xaa is either 4Bpa or 6Bpa, were synthesized (see Chapter 5) and modified by $Ru(bpy)_2CO_3$ in high yield to afford derivatives with electronic-absorption and steady-state-emission properties characteristic of $Ru(bpy)_3^{2+}$.[43] Mass spectral analyses on the modified peptides were consistent with the expected molecular weights and the ^1H NMR of the 4Bpa-peptide complex contains peaks aromatic peaks consistent with the $Ru(bpy)_3^{2+}$ complex. While both peptides were readily modified to form the $Ru(bpy)_3^{+2}$ complex, the $^*Ru^{2+}$ lifetimes measured for the two complexes were very different. The $^*Ru^{2+}$ lifetime of the $Ru(bpy)_2$(4Bpa) peptide was nearly identical with that of $Ru(bpy)_3^{2+}$ (τ=610 ns for $Ru(bpy)_2$(4Bpa) peptide, τ=640 ns for $Ru(bpy)_3^{2+}$ in aqueous solution[32]); however, the $^*Ru^{2+}$ lifetime of the $Ru(bpy)_2$(6Bpa) peptide is significantly shorter (τ< 6 ns). The deactivation of the excited state in the 6Bpa derivative is probably the result of a sterically induced increase in the Ru-N(6Bpa) bond length that places a ligand field excited state in the same energy region as the metal-to-ligand charge-transfer (MLCT, Ru^{2+} to bpy) state.[53]

Figure 6-10. Comparison of the 6Bpa72 cyt. *c* (left, A) and 4Bpa72 cyt.*c* (right, B) models based on the crystal structure of ferric horse heart cyt. *c* .[52] The van der Waals surface of the protein is shown as purple dots and the backbone trace is depicted as the teal colored line. The 4(6)Bpa72 residues are colored by atom; carbon atoms are shown in green, hydrogen atoms are shown in white, and the metal-chelating nitrogen atoms are shown in blue. Note that in the 6Bpa protein the chelating bipyridyl nitrogen atoms are facing into the protein; in the 4Bpa protein the chelating bipyridyl nitrogens are likely to be solvent accessible.



Direct-photoinduced (DP) and flash-quench (FQ) techniques[28,42,43] were employed to obtain the rates of electron transfer in $Ru(bpy)_2$(4Bpa72)cyt. *c*. The rates of both Fe^{2+} to Ru^{3+} ET and DPET ($^*Ru^{2+}$ to Fe^{3+}) are much higher than in the corresponding $Ru(bpy)_2$(im)-modified His72 protein[27] (Table 6-1). Since the intrinsic decay rates of the His72 and 4Bpa72 modified proteins are comparable, the DPET rates are proportional to ET product yield. The increased yield of DPET products (~30%) in the 4Bpa72 protein relative to the His72 (2.3%)[27]protein may be understood in terms of the nature of the DPET reaction. The electron transfer originates from MLCT excitation, and the effective electron donor, a bipyridyl-based anion radical, is built directly into the polypeptide backbone in $Ru(bpy)_2$(4Bpa72)cyt. *c*, in contrast to being at some undefined location with respect to the protein surface in $Ru(bpy)_2$-modified His72 cyt. *c*. Therefore, it was concluded that the bipyridyl ligand enhances the distant donor-acceptor electronic coupling

Table 6-1. Electron-Transfer Parameters for $Ru(bpy)_2$(4Bpa72)cyt. *c*.

	Excited-state lifetimes	ET Reactions	
	$^*Ru^{2+}$ (ns)	$^*Ru^{2+}$-Fe^{3+} (s^{-1})	Fe^{2+}-Ru^{3+} (s^{-1})
$Ru(bpy)_2$(4Bpa72) cyt *c*	Fe^{2+}: 62	6(2) x 10^6	6.5(5) x 10^6
(pH=7.0)	Fe^{3+}: 52		
$Ru(bpy)_2$(4Bpa72) cyt *c*		9(3) x 10^6	1.2(5) x 10^7
(pH=11.0)	Fe^{3+}: 51		1.0(5) x 10^6
$Ru(bpy)_2$(His72) cyt *c*	Fe^{2+}: 70	3.4(7) x 10^5	9.0(3) x 10^5
(pH 7.0)	Fe^{3+}: 70		

by effectively shortening the tunneling pathway between the Ru and heme redox units in position-72 modified proteins.[27,31] High DPET rates also have been observed in cyt. *c* derivatives with a $Ru(bpy)_2$(dicarboxybipyridine) tethered directly to a surface lysine.[38,39,41]

Protein Folding Experiments

The enhanced ET properties of $Ru(bpy)_2$(4Bpa72)cyt. *c* were used to probe conformationally perturbed states of the protein. The structural conformation of cytochrome *c* is highly dependent upon pH; at high pH (p*Ka* ~9.3),[54] ferric cyt. *c* exists in a low-potential state (the midpoint potential is roughly -200 mV) with altered ligation of the heme group,[55,56] while at identical conditions the ferrous state has a heme ligation and overall protein structure resembling the neutral pH state. Due to the large quantum yield and enhanced electron-transfer rates of $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c*, DPET experiments at pH=11 allow the rapid (< 60 ns, 480 nm-excitation) generation of ferrous protein in the ferric alkaline conformation. The subsequent Fe^{2+} to Ru^{3+} ET kinetics are comprised of at least two components (~1 x 10^7, ~1 x 10^6 s^{-1}),[43] both of which are much faster than estimates of the Met80 religation rate (40 to 1 s^{-1}).[55,57-61] The finding of two distinct decay pathways for the intramolecular oxidation of the photogenerated high-pH ferrous protein accords with other evidence indicating that alkaline ferricytochrome *c* is a mixture of at least two different conformers.[55,56]

Conclusions

The successful incorporation of 4Bpa and 6Bpa into large protein motifs was clearly demonstrated through the semisynthetic construction of two horse heart cytochrome *c* variants with these unnatural amino acids at position 72. While the semisynthetic method was found to be extremely selective for the correct peptide sequence, the proteins were readily constructed using standard protocols. Characterization of the mutant proteins by a

variety of independent methods indicated that negligible perturbation of the protein structure resulted from the introduction of the unnatural amino acids. The potential for these amino acids to play a functional role in protein designs was investigated through the formation of a $Ru(bpy)_3^{2+}$ redox center on the unnatural amino side chain. Interestingly, Ru(2,2'-bipyridine)$_2^{2+}$ was found to bind to 4Bpa72 cyt *c* but not to the 6Bpa72 protein, even though peptide models indicated that both regioisomers were capable of efficient metal modification. The differences in reactivity were attributed to the relative orientations of the ligating groups with respect to the protein surface specific for each regioisomer. The $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c* was found to undergo efficient electron transfer between the Ru-metal site and the heme group; electron transfer rates measured for $Ru(bpy)_3^{2+}$-4Bpa72 cyt. *c* were approximately tenfold higher than that observed for the corresponding His72-modified protein. The enhanced electron transfer properties were used to probe conformationally altered states of the protein, and these experiments demonstrate that laser-induced electron transfer may be employed to study submicrosecond protein-folding events.

Experimental

The work described in this chapter was carried out as a collaborative project with Deborah S. Wuttke and Professor Harry B. Gray. Deborah Wuttke was responsible for the generation of the purified protein fragment 1-65 and the protein reconstitution reactions, as well as the subsequent metal modification of the 4Bpa72-cyt. *c*.[43] Absorption spectra were recorded in a 1.0 cm cell in 50 mM phosphate buffer (pH=7.0) at 25°C on a Cary 14 Spectrometer. Circular dichroism spectra were recorded at 15-20 μM (50 mM Phosphate, pH=7.0) from 190-280 nm in a 0.10 cm cell using a bandwidth of 1 nm, a scan rate of 50 nm/min, a 20 mdeg sensitivity, and 0.5 s time constant. Variable temperature circular dichroism spectra were recorded in a 0.10 jacketed cell under the same conditions; temperatures were regulated using a thermostated water/ethylene glycol bath. Protein

concentrations were determined from the absorption of the oxidized protein measured at 410 nm ($\varepsilon_{cyt\ c}$ = 1.07 x 10^5 M^{-1} cm^{-1}). Reduction potentials were measured by differential pulse poloragraphy at a 4,4'-bipyridine-modified gold electrode.[43] EPR spectra and isoelectric-focusing gel electrophoresis experiments were performed by D.S. Wuttke according to standard protocols.[43] Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) was carried out by T.J. Mizoguchi. Mass spectra of the 4Bpa72 protein fragment 66-104 and 4(6)Bpa72 cyt. *c* proteins were performed in 20 mM ammonium bicarbonate buffer (pH=7.6) diluted with 1N acetic acid. Spectra were recorded on a Vestec Thermospray LC-Mass Spectrometer and analyzed using algorithms developed in the laboratories of Prof. Curtis Monig at the University of California, Riverside. Low resolution mass spectra analyses for the 6Bpa72 66-104 fragment were performed by Jane Sanders on a Bio-Ion 20R Biopolymer Mass Analyzer at the California Institute of Technology Biomolecular Synthesis and Analysis Resource Center. All steady-state luminescence and excitation spectra, as well as the time-resolved luminescence decay and transient absorption measurements were performed and analyzed by D.S. Wuttke.[43]

HPLC Methods:

A, B : Refer to Chapter 3 Experimental Section

C : Flow Rate : 1 mL/min; Analytical C18 column, UV detection : 228, 254 nm. Linear Gradient : 25-45% CH_3CN (0.8% TFA) /H_2O (0.1 % TFA) over 30 min, 5 min at 45% CH_3CN/H_2O 0.1 % TFA.

D : Flow Rate : 2 mL/min; Semi-preparative C18 column, UV detection : 228, 254 nm. Isocratic elution : 31% CH_3CN (0.8% TFA) /H_2O (0.1 % TFA) over 15 min.

Semisynthesis of (4)6Bpa-cytochrome c Mutants

Synthesis of Protein Fragment 66-104 (*Bpa72*)

All peptides were synthesized on a 0.1 mmol scale using solid phase FMOC-amino protection and BOP/HOBT activated-ester chemistry on a Milligen 9050 peptide synthesizer. Pepsyn KA resin (Milligen) was used to afford free carboxylic acid groups on the peptide C-terminus. For the naturally occurring amino acids, 4.0-8.0 equivalents were used per coupling; the double coupling protocol described in Chapter 5 was employed for the bipyridyl amino acids. Activated esters were prepared *in situ* using either BOP and HOBT or with pre-activated pentafluorophenyl (oPFP) or benzotriazole esters (oDhbt for serine and threonine) with HOBT in 0.451 M N-methyl-morpholine/DMF. Functionalized amino acids were used with the following sidechain protection schemes : Ser, Thr, Tyr - *t*-butyl ether (*t*-Bu); Lys - *t*-butoxycarbamate (Boc); Asp, Glu - *t*-butyl ester (o*t*-Bu); Arg - 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc). Double and triple couplings were employed when needed as determined by ninhydrin tests on the peptide resin.[44] Acylation times varied from 0.75 to 2 hours depending upon the coupling efficiency of the particular amino acid and a 10 min acetyl capping wash (0.3 M acetic anhydride, 0.3 M HOBT in 3:1 DMF/CH_2Cl_2, 3 mL/min) was used after the maximum coupling yields had been achieved to facilitate purification of the full length peptide. Deprotection of FMOC-protected amine groups was performed using a 7 minute 20% piperidine/DMF wash (3 mL/min). Cleavage of the peptides followed the protocols outlined in Chapter 2. Purity was assessed reversed-phase HPLC (H_2O:CH_3CN mixtures; UV 256, 228 nm detection). The crude peptide mixtures were purified using P6 gel filtration chromatography with 50 mM acetic acid eluent, and semi-preparative reversed-phase HPLC. Peptide containing fractions were identified by TLC and HPLC methods and concentrated by lyophilization. Pure peptides were stored at -20°C.

Peptide fragment 66-104 : 6Bpa72:

RP-HPLC : Met80 sulfoxide, Method D, RT : 6.30 min.

Met80 thioether, Method D, RT : 8.70 min.

LRMS : Met80 sulfoxide : Calc. for [MH^+] $C_{216}H_{340}N_{53}O_{61}S$ [4687.3] Obs. [4688.4]

Met80 thioether : Calc. for [MH^+] $C_{216}H_{340}N_{53}O_{60}S$ [4671.4] Obs. [4670.8]

Peptide fragment 66-104 : 4Bpa72:

RP-HPLC : Met80 sulfoxide, Method C, RT : 9.08 min.

Met80 thioether, Method C, RT : 10.19 min.

LRMS : Met80 sulfoxide : Calc. for [MH^+] $C_{216}H_{340}N_{53}O_{61}S$ [4687.3] Obs. [4687.5]

Met80 thioether : Calc. for [MH^+] $C_{216}H_{340}N_{53}O_{60}S$ [4671.4] Obs. [4669.2]

Reconstitution of Protein Fragments[43]

4(6)Bpa72 cyt. *c* was synthesized by using 0.15 to 0.35 mM 1-65 lactone with 0.8 equivalents 66-104 (4(6)Bpa72) peptide in 25 mM sodium phosphate buffer, pH 6.8 to 7.0. The reaction solution was thoroughly degassed and the reaction carried out under anaerobic conditions. Reduction was achieved using 1 to 1.5 equivalents of sodium dithionite and the reaction allowed to proceed for ~ 40 hours. The products were purified by cation-exchange chromatography (Pharmacia FPLC 10/10 Mono S column) in 25 mM sodium phosphate buffer, pH 7.0 with a 0.1 to 0.4 M NaCl gradient. The semisynthetic proteins eluted at similar volumes as native cyt. *c*. Concentrations of all protein samples were carried out using ultrafiltration devices (Amicon YM-3 and Centricon-3). Buffer exchange was by Sephadex G-25 gel filtration chromatography. Using this methodology ~20 milligrams of 4(6)Bpa72 cyt. *c* was isolated.

$Ru(bpy)_2^{2+}$ Modification and Characterization of 4(6)Bpa Peptide Models

The $Ru(bpy)_2^{2+}$-modification reaction was performed with 1.0 mM peptide and 3.8 mM $Ru(bpy)_2CO_3$ in 25 mM phosphate buffer (sodium) pH = 7.0, under anaerobic conditions in the dark. After 5.5 hours, the reaction was judged to be complete by absorption spectroscopy and subsequently quenched with the addition of excess imidazole. The peptides were initially purified by gel filtration (BioGel P2 200-400 mesh column, 50 mM acetic acid eluent) to yield a yellow oil. Similar yields of $Ru(bpy)_2^{2+}$-4Bpa and 6Bpa peptide complexes were obtained. Further purification of the 4Bpa-peptide complex was accomplished with RP-HPLC. The modified peptides were characterized by ^1H NMR, absorption spectroscopy, mass spectrometry, and luminescence decay measurements. The mass spectra for the peptides exhibited an envelope of peaks that was consistent with the theoretically calculated isotopic distribution.

$Ru(bpy)_3^{2+}$-4Bpa-peptide Complex:

RP-HPLC : Method A, RT : 20.04 min.

LRMS : Calc. for $[MH^+]$ $C_{61}H_{68}N_{13}O_8Ru$ [1213] Obs. [1213].

$Ru(bpy)_3^{2+}$-6Bpa-peptide Complex:

RP-HPLC : Method A, RT : 16.88 min.

LRMS : Calc. for $[MH^+]$ $C_{61}H_{68}N_{13}O_8Ru$ [1213] Obs. [1213].

$Ru(bpy)_2^{2+}$ Modification of 4Bpa72 cyt. c[43]

Modification of 4Bpa72 cyt. *c* with $Ru(bpy)_2CO_3$ was accomplished under similar conditions to that described for the Ru-modification of the model peptides. The modified protein was purified to homogeneity with cation-exchange chromatography.

Molecular Modeling Calculations on 4(6)Bpa72 cyt. c

A qualitative appraisal of the low energy conformers was performed for the unnatural amino acid side chain at position 72 in horse heart cyt. *c*. The calculations were based on the native crystal structure [52] with mutations for the bipyridyl amino acids at position 72 using Insight II v.2.1.0 Modeling Package (Biosym Technologies) on a Silicon Graphics Personal Iris 4D/25TG. In each case, the bipyridyl side chain was rotated through 360° for both χ_1 and χ_2 using 10 degree increments. Energy minima were calculated for each conformation, keeping all other protein atoms fixed to facilitate the computation. Low energy conformers were selected from the Ramachandran map generated for χ_1 and χ_2. In all cases the chelating nitrogens of 6Bpa72 cyt *c* were directed towards the protein structure.

Acknowledgment

David Goodin, Curtis Monnig, Jane Sanders, and Jay Winkler provided assistance and helpful discussions. The work described in this chapter was published as a communication in the *Journal of the American Chemical Society*:

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.

References

(1) Hedstrom, L.; Graf, L.; Stewart, C-B.; Rutter, W.J.; Phillips, M.A. "Modulation of Enzyme Specificity by Site-Directed Mutagenesis," *Meth. Enzym.* **1991**, *202*, 671.

(2) Bone, R.; Agard, D.A. "Mutational Remodeling of Enzyme Specificity," *Meth. Enzym.* **1991**, *202*, 643-671.

(3) Storer, A.C. "Engineering of proteases and protease inhibition," *Curr. Opin. Biotechnology* **1991**, *2*, 606-613.

(4) McGrath, M.E.; Haymore, B.L.; Summers, N.L.; Craik, C.S.; Fletterick, R.J. "Structure of an Engineered, Metal-Actuated Switch in Trypsin," *Biochemistry* **1993**, *32*, 1914-1919.

(5) Higaki, J.N.; Fletterick, R.J.; Craik, C.S. "Engineered metalloregulation in enzymes," *Trends in Biol. Sci.* **1992**, *17*, 100-104.

(6) Muheim, A.; Todd, R.J.; Casimiro, D.R.; Gray, H.B.; Arnold, F.H. "Ruthenium-Mediated Protein Cross-Linking and Stabilization," *J. Am. Chem. Soc.* **1993**, *115*, 5312-5313.

(7) Bain, J.D.; Switzer, C.; Chamberlin, A.R.; Benner, S.A. "Ribosome-Mediated Incorporation of a Non-Standard Amino Acid into a Peptide Through Expansion of the Genetic Code," *Nature* **1992**, *356*, 537-539.

(8) Bain, J.D.; Glabe, C.G.; Dix, T.A.; Chamberlin, A.R.; Diala, E.S. "Biosynthetic Site-Specific Incorporation of a Non-Natural Amino Acid into a Polypeptide," *J. Am. Chem. Soc.* **1989**, *111*, 8013-8014.

(9) Anthony-Cahill, S.J.; Griffith, M.C.; Noren, C.J.; Suich, D.J.; Schultz, P.G. "Site-specific mutagenesis with unnatural amino acids," *Trends in Biol. Sci.* **1989**, *14*, 400-403.

(10) Chung, H.H.; Benson, D.R.; Schultz, P.G. "Probing the Structure and Mechanism of Ras Protein with an Expanded Genetic Code," *Science* **1993**, *259*, 806-809.

(11) Noren, C.J.; Anthony-Cahill, S.J.; Griffith, M.C.; Schultz, P.G. "A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins," *Science* **1989**, *244*, 182-188.

(12) Muir, T.W.; Kent, S.B.H. "The chemical synthesis of proteins," *Curr. Opin. Biotechnology* **1993**, *4*, 420-427.

(13) Humphries, J.; Offord, R.E.; Smith, R.A.G. "Chemical methods of protein synthesis and modification," *Curr. Opin. Biotechnology* **1991**, *2*, 539-543.

(14) Chaiken, I.M. "Semisynthetic Peptides and Proteins," *CRC Crit. Rev. Biochem.* **1981**, 255-301.

(15) Offord, R.E. *Semisynthetic Proteins*; John Wiley and Sons: New York, **1980**.

(16) Williams, R.M. *Synthesis of Optically Active α-Amino Acids*; Pergamon Press: New York, **1989**.

(17) Stewart, J.M.; Young, J.D. *Solid Phase Peptide Synthesis*; Pierce Chemical Co.: Rockford, IL, **1984**.

(18) Corradin, G.; Harbury, H.A. "Reconstitution of Horse Heart Cytochrome *c*: Interaction of the Components Obtained upon Cleavage of the Peptide Bond following Methionine Residue 65," *Proc. Natl. Acad. Sci., USA* **1971**, *68*, 3036-3039.

(19) ten Kortenaar, P.B.W.; Adams, P.J.H.M.; Tesser, G.I. "Semisynthesis of horse heart cytochrome *c* analogues from two or three fragments," *Proc. Natl. Acad. Sci., USA* **1985**, *82*, 8279-8283.

(20) Raphael, A.L.; Gray, H.B. "Semisynthesis of Axial-Ligand (Position 80) Mutants of Cytochrome *c*," *J. Am. Chem. Soc.* **1991**, *113*, 1038-1040.

(21) Koul, A.K.; Wasserman, G.F.; Warme, P.K. "Semisynthetic Analogs of Cytochrome *c* at Positions 67 and 74," *Biochem. Biophys. Res. Comm.* **1979**, *69*, 1253-1259.

(22) Frauenhoff, M.M.; Scott, R.A. "The Role of Tyrosine 67 in the Cytochrome *c* Crevice Structure Studied by Semisynthesis," *Proteins* **1992**, *14*, 202-212.

(23) Wallace, C.J.; Clark-Lewis, I. "Functional Role of Heme Ligation in Cytochrome *c*," *J. Biol. Chem.* **1992**, *267*, 3852-3861.

(24) Wasserman, G.F.; Nix, P.T.; Koul, A.K.; Warme, P.K. "Semisynthetic Analogs of Cytochrome *c* : Substitutions for Methionine at Position 80," *Biochim. Biophys. Acta* **1980**, *623*, 457-460.

(25) Bowler, B.E.; Raphael, A.L.; Gray, H.B. "Long-Range Electron Transfer in Donor (Spacer) Acceptor Molecules and Proteins," in *Progress in Inorganic Chemistry: Bioinorganic Chemistry*; Lippard, S.J., Ed.; John Wiley and Sons: New York, **1990**; pp 259-322.

(26) Winkler, J.R.; Gray, H.B. "Electron Transfer in Ruthenium-Modified Proteins," *Chem. Rev.* **1992**, *92*, 369-379.

(27) Wuttke, D.S.; Bjerrum, M.J.; Winkler, J.R.; Gray, H.B. "Electron-Tunneling Pathways in Cytochrome *c*," *Science* **1992**, *256*, 1007-1009.

(28) Chang, I-J.; Gray, H.B.; Winkler, J.R. "High-Driving-Force Electron Transfer in Metalloproteins: Intramolecular Oxidation of Ferrocytochrome *c* by Ru(2,2'-bpy)$_2$(im)(His33)$^{3+}$," *J. Am. Chem. Soc.* **1991**, *113*, 7056-7057.

(29) Bowler, B.E.; Meade, T.J.; Mayo, S.L.; Richards, J.H.; Gray, H.B. "Long-Range Electron Transfer in Structurally Engineered Pentaammineruthenium (Histidine-62)cytochrome *c*," *J. Am. Chem. Soc.* **1989**, *111*, 8757-8759.

(30) Marcus, R.A.; Sutin, N. "Electron transfers in chemistry and biology," *Biochim. Biophys. Acta* **1985**, *811*, 265-322.

(31) Onuchic, J.N.; Beratan, D.N.; Winkler, J.R.; Gray, H.B. "Pathway Analysis of Protein Electron-Transfer Reactions," *Ann. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 349-377.

(32) Juris, A.; Balzani, V.; Barigelletti, F.; Campagna, S.; Belser, P.; Von Zelewsky, A. "Ru(II) Polypyridine Complexes: Photophysics, Photochemistry, Electrochemistry, and Chemiluminescence," *Coord. Chem. Rev.* **1988**, *84*, 85-277.

(33) Mecklenberg, S.L.; Peek, B.M.; Schoonover, J.R.; McCafferty, D.G.; Wall, C.G.; Erickson, B.W.; Meyer, T.J. "Photoinduced Electron Transfer in Amino Acid Assemblies," *J. Am. Chem. Soc.* **1993**, *115*, 5479-5495.

(34) Schoonover, J.R.; Strouse, G.F.; Chen, P.; Bates, W.D.; Meyer, T.J. "Application of Time-Resolved Resonance Raman Spectroscopy to Intramolecular Electron Transfer," *Inorg. Chem.* **1993**, *32*, 2618-2619.

(35) Mecklenburg, S.L.; Peek, B.M.; Erickson, B.W.; Meyer, T.J. "Photoinduced Electron Transfer in Redox-Active Lysines," *J. Am. Chem. Soc.* **1991**, *113*, 8540-8542.

(36) Pan, L.P.; Hibdon, S.; Liu, R-Q.; Durham, B.; Millett, F. "Intracomplex Electron Transfer between Ruthenium-Cytochrome *c* Derivatives and Cytochrome *c* Oxidase," *Biochemistry* **1993**, *32*, 8492-8498.

(37) Willie, A.; McLean, M.; Liu, R-Q.; Hilgen-Willis, S.; Saunders, A.J.; Pielak, G.J.; Sligar, S.G.; Durham, B.; Millett, F. "Intracomplex Electron Transfer between Ruthenium-65-Cytochrome b_5 and Position-82 Variants of Yeast Iso-1-cytochrome *c*," *Biochemistry* **1993**, *32*, 7519-7525.

(38) Pan, L.P.; Durham, B.; Wolinska, J.; Millett, F. "Preparation and Characterization of Singly Labeled Ruthenium Polypyridine Cytochrome *c* Derivatives," *Biochemistry* **1988**, *27*, 7180-7184.

(39) Hahm, S.; Durham, B.; Millett, F. "Photoinduced Electron Transfer between Cytochrome *c* Peroxidase and Horse Cytochrome *c* Labeled at Specific Lysines with (Dicarboxybipyridine)(bisbipyridine)ruthenium(II)," *Biochemistry* **1992**, *31*, 3472-3477.

(40) Geren, L.; Hahm, S.; Durham, B.; Millett, F. "Photoinduced Electron Transfer between Cytochrome *c* Peroxidase and Yeast Cytochrome *c* Labeled at Cys 102 with (4-Bromomethyl-4'-methylbipyridine)[bis(bipyridine)]ruthenium^{2+}," *Biochemistry* **1991**, *30*, 9450-9457.

(41) Durham, B.; Pan, L.P.; Long, J.E.; Millett, F. "Photoinduced Electron-Transfer Kinetics of Singly Labeled Ruthenium Bis(bipyridyl) dicarboxybipyridine Cytochrome *c* Derivatives," *Biochemistry* **1989**, *28*, 8659-8665.

(42) Wuttke, D.S.; Bjerrum, M.J.; Chang, I-J.; Winkler, J.R.; Gray, H.B. *Biochim. Biophys. Acta* **1992**, *1101*, 168.

(43) Wuttke, Deborah S., Ph.D. Thesis, California Institute of Technology, **1993**.

(44) Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. "Color Test for Detection of Free Terminal Amino Groups in the Solid Phase Synthesis of Peptides," *Anal. Biochem.* **1970**, *34*, 595-598.

(45) Norris, K.; Halstrom, J.; Brunfeldt, K. "Stability of Methionyl Residues Towards Oxidation During Solid Phase Peptide Synthesis," *Acta Chem. Scand.* **1971**, *25*, 945-954.

(46) Stadtman, E.R. "Oxidation of Free Amino Acids and Amino Acid Residues in Proteins by Radiolysis and by Metal-Catalyzed Reactions," *Ann. Rev. Biochem.* **1993**, *62*, 797-821.

(47) The observed mass increase is consistent with either Co, Fe, or Ni cation binding.

(48) Madesclaire, M. "Reduction of Sulfoxides to Thioethers," *Tetrahedron* **1988**, *44*, 6537-6580.

(49) Houghten, R.A.; Li, C.H. "Reduction of Sulfoxides in Peptides and Proteins," *Anal. Biochem.* **1979**, *98*, 36-46.

(50) Schecter, Y. "Selective Oxidation and Reduction of Methionine Residues in Peptides and Proteins by Oxygen Exchange between Sulfoxide and Sulfide," *J. Biol. Chem.* **1986**, *261*, 66-70.

(51) Peptide sequencing experiments were carried out at the California Insititute of Technology Biomolecular Synthesis and Analysis Resource Center.

(52) Bushnell, G.W.; Louie, G.V.; Brayer, G.D. "High-resolution Three-dimensional Structure of Horse Heart Cytochrome *c*," *J. Mol. Biol.* **1990**, *214*, 585-595.

(53) Fujita, E.; Milder, S.J.; Brunschwig, B.S. "Photophysical Properties of Covalently Attached $Ru(bpy)_3^{2+}$ and $Mcyclam^{2+}$ (M= Ni, H_2) Complexes," *Inorg. Chem.* **1992**, *31*, 2079-2085.

(54) Thoerell, H.; Åkesson, Å. "Studies on Cytochrome *c*. II. The Optical Properties of Pure Cytochrome *c* and Some of Its Derivatives," *J. Am. Chem. Soc.* **1941**, *63*, 1812-1818.

(55) Barker, P.D.; Mauk, A.G. "pH-Linked Conformational Regulation of a Metalloprotein Oxidation-Reduction Equilibrium: Electrochemical Analysis of the Alkaline Form of Cytochrome *c*," *J. Am. Chem. Soc.* **1992**, *114*, 3619-3624.

(56) Ferrer, J.C.; Guillemette, J.G.; Bogumil, R.; Inglis, S.C.; Smith, M.; Mauk, A.G. "Identification of Lys79 as an Iron Ligand in One Form of Alkaline Yeast Iso-1-Ferricytochrome *c*," *J. Am. Chem. Soc.* **1993**, *115*, 7507-7508.

(57) Hasumi, H.; "Kinetic Studies on Isomerization of Ferricytochrome *c* in Alkaline and Acid pH Ranges by the Circular Dichroism Stopped-Flow Method," *Biochim. Biophys. Acta* **1980**, *626*, 265-276.

(58) Hong, X.; Dixon, D.W. "NMR study of the alkaline isomerization of ferricytochrome *c*," *FEBS Lett.* **1989**, *246*, 105-108.

(59) Lambeth, D.O.; Campbell, K.L.; Zand, R.; Palmer, G. "The Appearance of Transient Species of Cytochrome *c* upon Rapid Oxidation or Reduction at Alkaline pH," *J. Biol. Chem.* **1973**, *248*, 8130-8136.

(60) Land, E.J.; Swallow, A.J. "One-Electron Reactions in Biochemical Systems as Studied by Pulse Radiolysis V. Cytochrome *c*," *Arch. Biochem. Biophys.* **1971**, *145*, 365-372.

(61) Pecht, I.; Faraggi, M. "Electron Transfer to Ferricytochrome *c*: Reaction with Hydrated Electrons and Conformational Transitions Involved," *Proc. Natl. Acad. Sci., USA* **1972**, *69*, 902-906.

Chapter 7: The Design, Synthesis, and Preliminary Characterization of a Prototype of the Metal-Dependent Regulatory Protein Motif

Introduction

A major objective of *de novo* protein design is the assembly of protein structures with defined structure and function. As discussed in Chapter 1, a goal of the work described in this thesis has been the design and synthesis of a protein construct which can serve as a metal-dependent regulator for a naturally occurring enzyme such as *c*AMP dependent protein kinase. The motif is based on a ββ super-secondary structure, with a pair of metal ligating groups situated on either end of a variable region containing a consensus sequence capable of inhibiting the targeted enzyme. (See Figure 1-2.) The metal ligating region near the amino terminus is composed of two residues with side chains capable of binding metal cations which are brought into close proximity by the formation of a Type II reverse turn, while the metal binding ligands at the carboxy terminus are provided by an unnatural amino acid capable of bidentate metal coordination. The motif is designed to have an extended conformation in solution in the absence of metal cations; however, upon addition of metal cations, the peptide can undergo a conformational change to afford a monomeric, tetradentate metal complex if the ligation properties of the two regions of the motif are matched. (See Figure 1-2.) The overall topology of the metal binding sites in the proposed motif resembles the "short" and "long" spacer distribution of metal ligating groups found in naturally occurring Zn^{+2} metalloenzymes.[1] In these systems, two metal ligating residues are often separated by 1-3 residues, and the third metal binding residue is generally separated from the former by a long stretch of amino acid residues (5-196 residues). It should be noted, however, that the proposed metal-binding motif differs from the Zn^{+2} metalloenzymes in the number of ligands afforded by the protein; the proposed motif is designed to form a tetradentate metal binding site, whereas the metal sites found in naturally-occurring systems are invariably composed of three protein-bound ligands, with the vacant coordination sites occupied by coordinated water molecules.

In an effort to test the efficacy of the designed motif, a dodecapeptide (**1**) was designed, synthesized, and characterized in terms of its metal binding properties. The

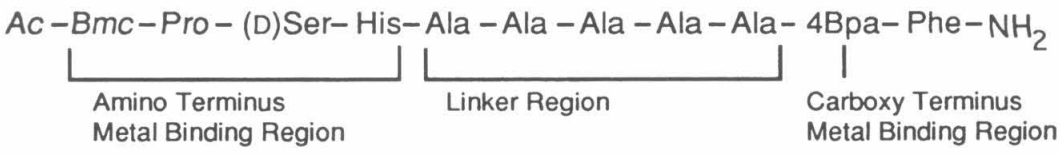


Figure 7-1. Sequence of the metal-dependent regulatory motif prototype (**1**)

constituents of the metal binding regions of the peptide was based on results from the preceding chapters, and the peptide sequence is shown in Figure 7-1. The 4Bpa residue was chosen for the carboxy terminus metal binding ligands because the results from preliminary modeling experiments[2] suggested that the binding site was fairly sterically constrained and that the unobstructed N,N-chelation moiety of this residue produced the best fit of the available regioisomers of the bipyridyl-amino acids. The selection of the amino terminus metal binding region of the motif was based on the results from Chapter 3, which found that the tetrapeptide sequence Bmc-Pro-D-Ser-His adopted a conformation consistent with a type II reverse turn in aqueous solution, and it was anticipated that the strong turn-forming propensity of this sequence would place the two metal ligation moieties of the histidine and β-methyl-cysteine residues in close proximity when incorporated into the larger polypeptide sequence.

The selection of the sequence for the variable linker region was based on several criteria. Alanine was chosen because the methyl side chain was expected to impart a limited amount conformational constraint on the overall peptide conformation. Recent work by Berg and co-workers,[3] in which a "minimalistic Zn^{+2} finger" protein was designed and synthesized, provides an excellent precedent for the use of alanine for limited conformational constraints in designed motifs. In this case, alanine residues were substituted for all but the core hydrophobic and metal binding residues of a zinc finger protein consensus sequence (26 residues total, 16 were converted to Ala). Despite these drastic changes, the peptide bound metal cations with affinities approaching that observed

for the native proteins, and the peptide was found to adopt a native-like structure upon metal cation binding as determined by UV/Vis absorption and ^{1}H NMR spectroscopy. In addition, the polyalanine sequence was not expected to compete with the incorporated metal ligands for the metal cations, because this sequence does not contain any side chain functional groups capable of effective metal coordination.

The length of the linker region sequence was also selected to accommodate the proposed conformational changes. Based on modeling studies[2] of the peptide:metal cation complex, a five residue sequence was anticipated to be the optimal spacer length between the two metal binding regions. Attempts were made to minimize the length of the linker region to further promote the intramolecular metal cation:peptide complex formation; however, sequences containing less than five residues were found to be too constrained to undergo the necessary folding to form the proposed tetradentate complex. These conclusions are supported by the fact this spacer length has precedent in naturally occurring Zn^{+2} metalloenzymes (e.g., astacin).[1] Finally, the five residue linker region was designed to be compatible with future designs of metal-dependent regulatory motifs targeted to *c*AMP protein kinase. Analysis of the crystal structure containing a bound peptide inhibitor demonstrated that inhibitors containing a five residue consensus sequence reside in the catalytic pocket in an extended chain conformation.[4]

The peptide **1** was synthesized and the metal binding properties were characterized for both Zn^{+2} and Co^{+2} metal cations. Zinc(II) was chosen for study since it is known to play an important regulatory role in biological systems; indeed the native regulatory domain of *c*AMP protein kinase has been found to contain four Zn^{+2} metal binding sites. On the other hand, Co^{+2} was used to characterize the metal binding properties of the peptide complex, as complexes with this metal cation generally afford a wide range of spectroscopic signatures amenable to a variety of techniques (see Chapter 5).[5]

Results and Discussion

Peptide Synthesis

Synthesis of the peptide, including *N*-terminal acetylation, was performed using standard solid-phase peptide synthesis protocols as outlined in the Experimental sections of the previous chapters (see Chapters 2, 3, and 5). Double coupling acylation protocols were employed for the β-methyl-cysteine and 4-bipyridyl-alanine residues as described in Chapter 3 and Chapter 5, respectively. For the β-methyl-cysteine residue 2.0 equivalents were employed on the first coupling cycle, and 0.5 equivalents were used in the second acylation cycle. The synthesis proceeded smoothly and the subsequent cleavage/workup of the peptide was carried out using the methods described in Chapter 2 with no difficulties. The peptide was fairly pure by RP-HPLC analysis; only two minor side products were detected. Subsequent purification of the peptide was performed using semi-preparative RP-HPLC and purified peptide fractions were lyophilized and stored at -80° C. The peptide was characterized by high resolution mass spectrometry, ^{1}H NMR, Ellman's test for free thiol groups,[6] and UV/Vis absorption spectroscopy.

Metal-Binding Titrations with $CoCl_2$ and $ZnCl_2$

The metal binding properties of peptide **1** were assessed using metal cation titrations monitored by UV/Vis absorption spectroscopy for both $ZnCl_2$ and $CoCl_2$ using the protocols described in Chapter 5. In all cases, the bipyridyl π-π* band undergoes a red shift upon addition of metal cations, indicating that the bipyridyl participates in the metal binding process (see Figure 7-2). Further, the titrations for both metal cations at low peptide concentrations ([Pep] < 100 μM) exhibit well-defined isosbestic points (265, 290 nm) which suggests that the metal binding process follows a two state process comprised of free peptide and the metal-cation:peptide complex. As a result, it was possible to deduce the peptide:metal cation complex stoichiometry and evaluate the metal binding affinities

Table 7-1. Dissociation Constants and Metal Complex Stoichiometries for peptide **1** and Ac-4Bpa-Thr-Pro-D-Ser-Val-Phe-NH_2

	K_d, M (Peptide:Metal Stoichiometry)	
Peptide	Zn^{+2}	Co^{+2}
peptide **1**	$< 5 \times 10^{-8}$ (1:1)	$< 5 \times 10^{-8}$ (1:1)
4Bpa-Thr-Pro-D-Ser-Val-Phe-NH_2	$2.2 \pm 0.2 \times 10^{-5}$ (1:1)	$3.8 \pm 0.4 \times 10^{-6}$ (2:1)

from binding isotherms extracted from the titration spectra at wavelengths where significant absorbance changes were observed (see Chapter 5). Typical binding isotherms for both the Zn^{+2} and Co^{+2} metal cations are shown in Figure 7-3, and the results from these studies are presented in Table 7-1.

Peptide **1** has an extremely high affinity for both Zn^{+2} and Co^{+2} cations that is well beyond the sensitivity limit of the UV/Vis absorption assay employed for the binding constant determinations. Saturating conditions were observed ([Pep] >> K_d) for all of the titrations, even when the measurements were performed at extremely dilute peptide concentrations. ([Pep] ≈ 0.5 μM, see Figure 7-4. Data is shown for a titration with $CoCl_2$ and similar results were obtained with $ZnCl_2$.) As a result, the dissociation constants for **1** with these metal cations could not be determined accurately using these methods, but they are assumed to be at or below the mid-nanomolar range (see Table 7-1). It is evident, through comparison of the dissociation constants for **1** with those obtained for the control peptide Ac-4Bpa-Thr-Pro-D-Ser-Val-Phe-NH_2 (see Chapter 5), that the enhanced metal binding properties of peptide **1** arise from a cooperative binding process involving the bipyridyl amino acid and the other residues capable of ligating metal cations, β-methyl-cysteine and histidine. This is particularly apparent in the observed binding affinities of **1** for Zn^{+2} cations; the dissociation constants for the Zn^{+2}:peptide **1** complex are at least three

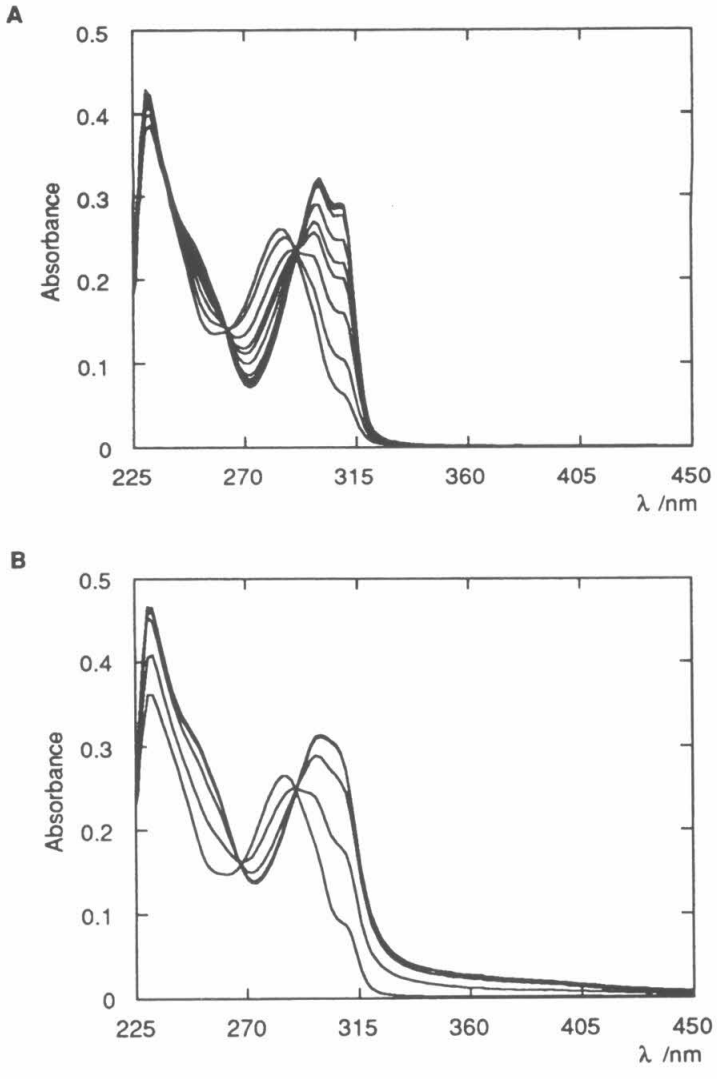


Figure 7-2. (A) Absorption spectra of the titration of **1** (20.3 μM) with $ZnCl_2$, (pH=8.25, 50 mM HEPES buffer) $ZnCl_2$ was added in 3 μM aliquots; (B) Absorption spectra of the titration of **1** (20.6 μM) with $CoCl_2$, (pH=8.25, 50 mM HEPES buffer) $CoCl_2$ was added in 8 μM aliquots.

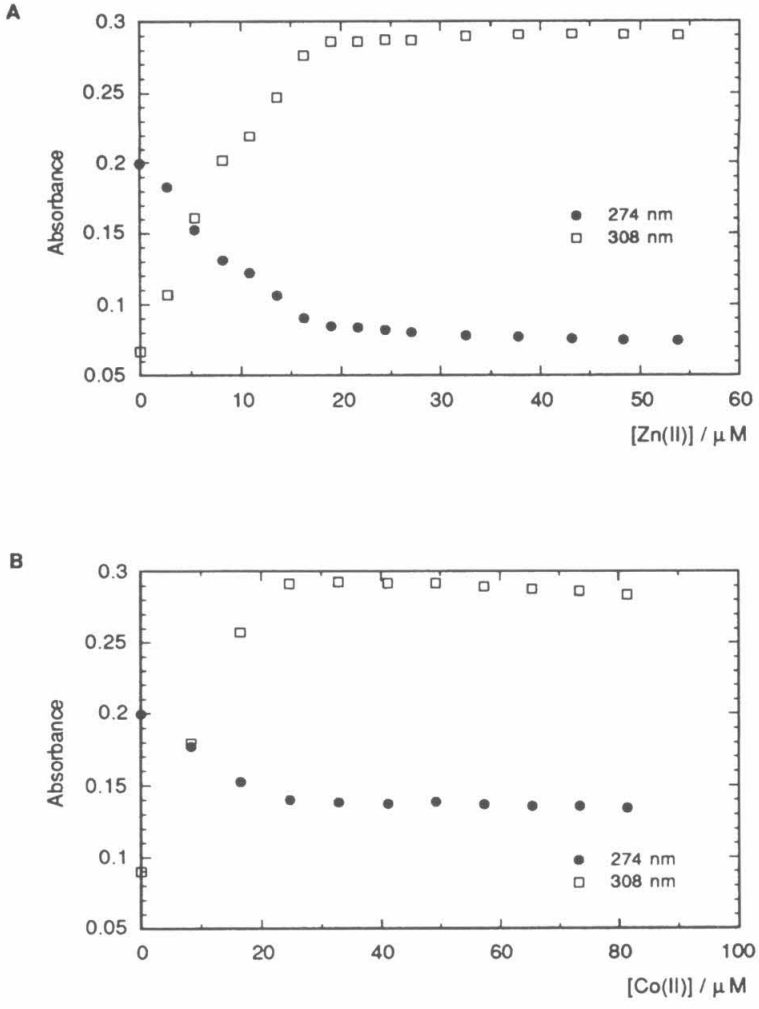


Figure 7-3. (A) Binding isotherms from the titration of **1** (20.3 μM) with $ZnCl_2$ (pH=8.25, 50 mM HEPES buffer); (B) Binding isotherms from the titration of **1** (20.6 μM) with $CoCl_2$, (pH=8.25, 50 mM HEPES buffer).

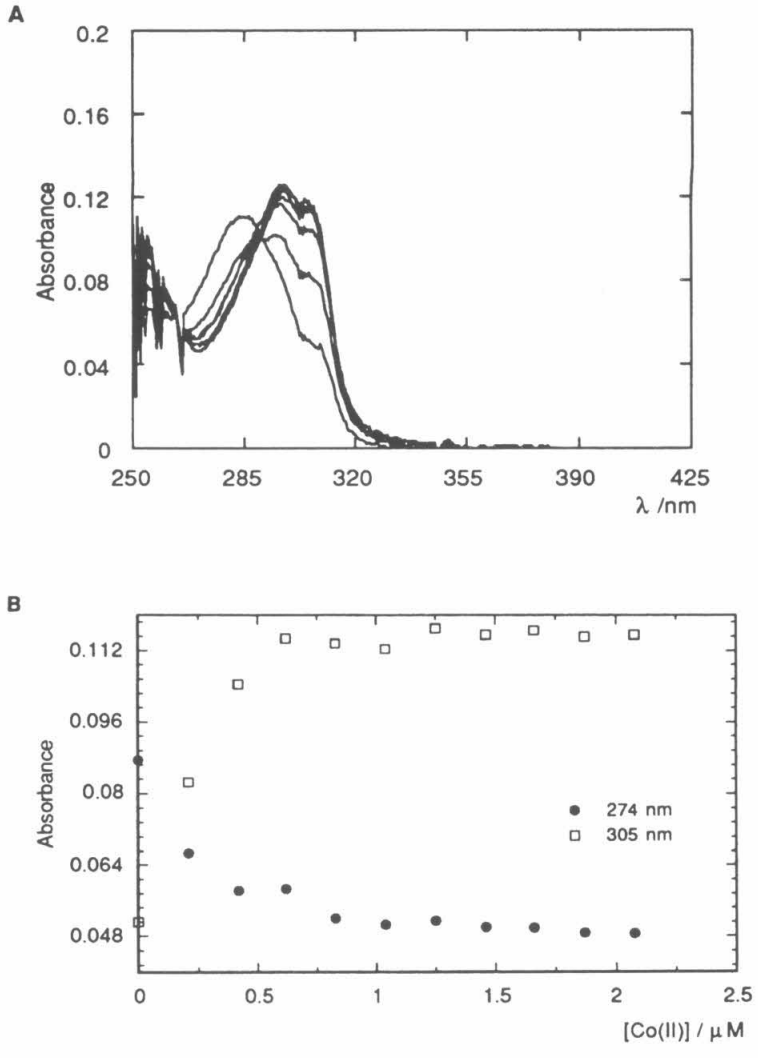


Figure 7-4. (A) Absorption spectra of the titration of **1** (0.86 μM) with $CoCl_2$, (pH=8.25, 50 mM HEPES buffer) $CoCl_2$ was added in 0.23 μM aliquots; (B) Binding isotherms from the titration of **1** (0.86 μM)with $CoCl_2$, (pH=8.25, 50 mM HEPES buffer).

orders of magnitude greater than would be expected if the bipyridyl amino acid were solely responsible for the complex formation. Similar inferences can be made on the basis of the observed stoichiometries of the metal complexes. The results from the control peptide, Ac-4Bpa-Thr-Pro-D-Ser-Val-Phe-NH_2, demonstrated that the 4Bpa residue preferentially formed 2:1 peptide:metal cation complexes with $CoCl_2$. However, the observed 1:1 stoichiometry for the peptide **1**:Co^{+2} complex suggests that the bipyridyl amino acid does not act independently in the formation of the complex.

Conclusive evidence for participation of the β-methyl-cysteine residue in the complex formation can be obtained from the near UV Co^{+2} complex absorption spectrum (see Figure 7-5). Upon addition of Co^{+2} cations, a new absorption is observed on the shoulder of the bipyridyl π-π* band ($\lambda_{max} \approx 385$ nm). Interestingly, this effect is only observed in the Co^{+2}:peptide **1** complex; the corresponding Zn^{+2}:peptide **1** complex and the Co^{+2} metal cation complex formed with Ac-6Bpa-Val-Pro-D-Ser-Phe-His-NH_2 (see Chapter 5) do not show enhanced absorptions in this region (see Figure 7-2). Since sulfur ligands are known to undergo ligand-to-metal charge transfer (LMCT) processes when bound to Co^{+2} cations which typically produce bands in the near UV region (16,000-30,000 cm^{-1}),[7] the absorption band at $\lambda \approx 385$ nm observed for the Co^{+2}:peptide **1** complex can be assigned to an LMCT band arising from coordination of the β-methyl-cysteine residue to the Co^{+2} metal cation. The presence of the isosbestic points in the titration spectra, and the fact that the LMCT band shows a similar dependence upon the metal cation concentration as that observed for the red shift of the bipyridyl π-π* band, strongly suggests that the bipyridyl ligand and the β-methyl-cysteine residue are both involved in the formation of the peptide **1**:Co^{+2} complex.

An alternative method using ^1H NMR was investigated to confirm the role of the histidine residue in the metal binding interaction. It is well known that protons covalently attached to ligands bound to Co^{+2} experience large paramagnetic shifts in the ^1H NMR spectrum due to contact and dipolar interactions with the metal cation.[5] In addition,

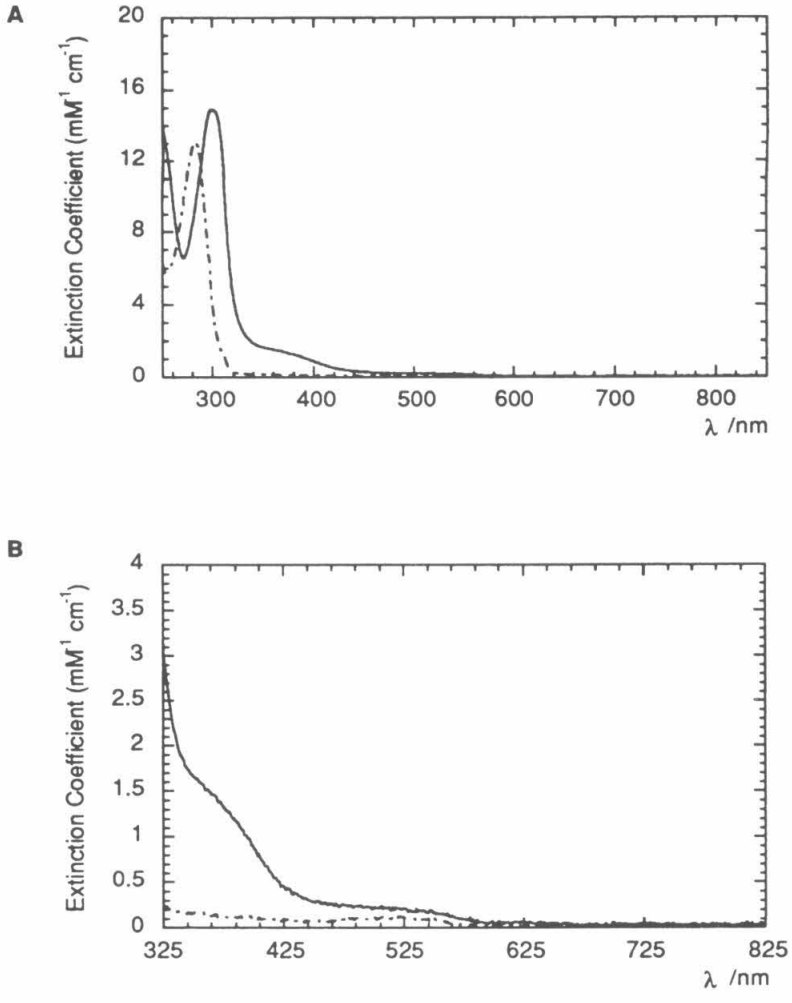


Figure 7-5. (A) UV/Visible absorption spectrum of the Co^{+2}:peptide **1** complex; (B) Near UV/Vis absorption spectrum of the Co^{+2}:peptide **1** complex. (Spectra were collected in 50 mM HEPES, pH=8.25 on a Cary-14 UV/Vis absorption spectrometer).

since the Co^{+2} cation has a fast electron spin relaxation time, the line widths of the shifted protons are generally fairly narrow. This technique has been used to characterize a wide variety of low molecular weight Co^{+2} complexes,[7] as well as provide a means to identify metal-bound histidine residues in proteins such as carbonic anhydrase.[8] At the current time, attempts to confirm the role of the histidine imidazole ring in the complex formation of peptide **1** with Co^{+2} using this method have been suspended, due to the concentration requirements of the monomeric complex formation. While the complex forms readily at dilute peptide concentrations, multiple equilibrium processes were observed in titrations where higher peptide concentrations ([Pep] > 100 μM) were employed as evidenced by the lack of clearly defined isosbestic points in the titration spectra (see Figure 7-6). These effects are likely to be due to the formation of intermolecular peptide:metal cation aggregates. As a result, the ^1H NMR experiments had to be performed at low concentrations of the peptide **1**:Co^{+2} complex ([Pep] = 100 μM; [Co^{+2}] = 90 μM); this concentration was below the sensitivity limit of the technique as no peptide peaks were observed even after extensive signal averaging on a high field spectrometer (500.14 MHz, 5 mm probe). Data acquisition in a 10 mm probe and variation of the parameters will be attempted at a later date. Despite these results, the histidine imidazole side chain is assumed to play a role in the metal complex formation, as a result of the strong propensity for the tetrapeptide sequence Bmc-Pro-D-Ser-His to adopt a type II reverse turn (see Chapter 3). As discussed in Chapter 3, the reverse turn conformation would bring the histidine side chain in close proximity to the β-methyl-cysteine residue, and since the β-methyl-cysteine is known to bind to the metal cation, it is likely that the imidazole side chain of the histidine also participates in the metal complex formation.

Information about the Co^{+2} metal geometry can be inferred from the visible spectrum of the peptide:Co^{+2} complex. As discussed in Chapter 5, the absorption spectra of Co^{+2} complexes give rise to distinctive peaks in the visible and near IR regions due to

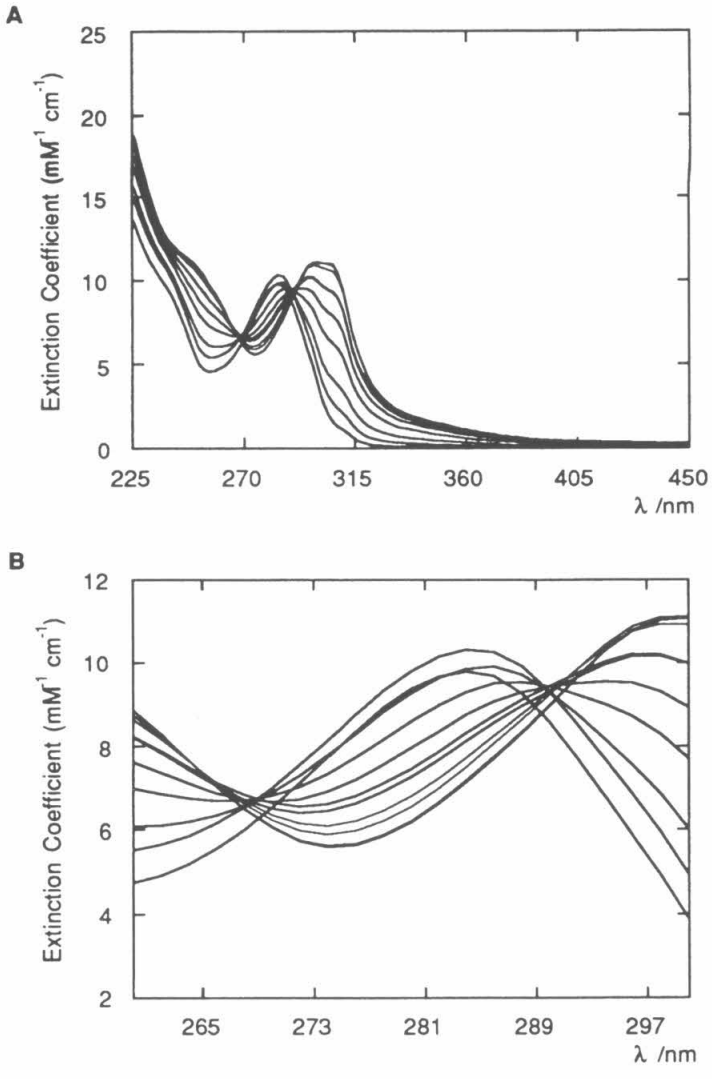


Figure 7-6. (A) Absorption Spectra of the titration of **1** (121 mM) with $CoCl_2$ (pH=8.25, 50 mM HEPES buffer) $CoCl_2$ was added in 20 μM aliquots. (B) Expansion plot of the titration spectra.

d-d orbital transitions. In general, tetrahedral complexes give rise to fairly intense bands in visible region ($\varepsilon > 250\ M^{-1}\ cm^{-1}$), pentacoordinate complexes display weaker, but readily observable, bands in the same region ($\varepsilon < 225\ M^{-1}\ cm^{-1}$), and octahedral complexes are characterized by very weak absorptions in both the visible and near IR regions ($\varepsilon < 30\ M^{-1}\ cm^{-1}$). The absorption spectrum of the Co^{+2}:peptide **1** complex is relatively featureless (see Figure 7-5b) over the entire visible region; no bands distinctive for either the pentacoordinate or tetrahedral geometries were observed. The absence of distinctive bands in the visible region supports the assignment of an octahedral geometry for the metal cation.

In addition, electron paramagnetic resonance spectroscopy (EPR) measurements on the complex provided further information regarding the spin state of the complex. Octahedral complexes of Co^{+2} can be either high (S=3/2) or low spin (S=1/2). Low spin octahedral Co^{+2} complexes often give well-defined EPR spectra in solution at relatively high temperatures (77K-298K) since these complexes generally have long-lived electron spin lifetimes. In contrast, high spin octahedral Co^{+2} complexes generally give broad, poorly defined spectra at these temperatures due to the large manifold of potential spin states in the ground state and the ability of the unpaired electrons to participate in spin relaxation mechanisms through spin orbit coupling.[9] As a result, EPR measurements for high spin octahedral Co^{+2} complexes require extremely low temperatures (2K) for detection. The EPR spectrum of the Co^{+2}:peptide **1** complex was measured at 77K; however, no peaks were observed at this temperature. These results suggest that the peptide **1**:Co^{+2} complex is high spin with octahedral coordination geometry. Alternatively, the sample concentrations employed in these experiments may be below the threshold of the EPR technique for transition metal complexes.[10] Futher experiments to confirm the spin state through measurements of the Co^{+2}:peptide complex paramagnetic susceptibility[11] will be performed at a later date. The proposed hexacoordinate geometry of the complex could be accommodated through coordination of the three metal binding residues in the peptide

(4Bpa, His, and Bmc) along with two water molecules (i.e., $Co(Pep)(H_2O)_2^{+}$]. A computer generated model of the proposed complex is shown in Figure 7-7.

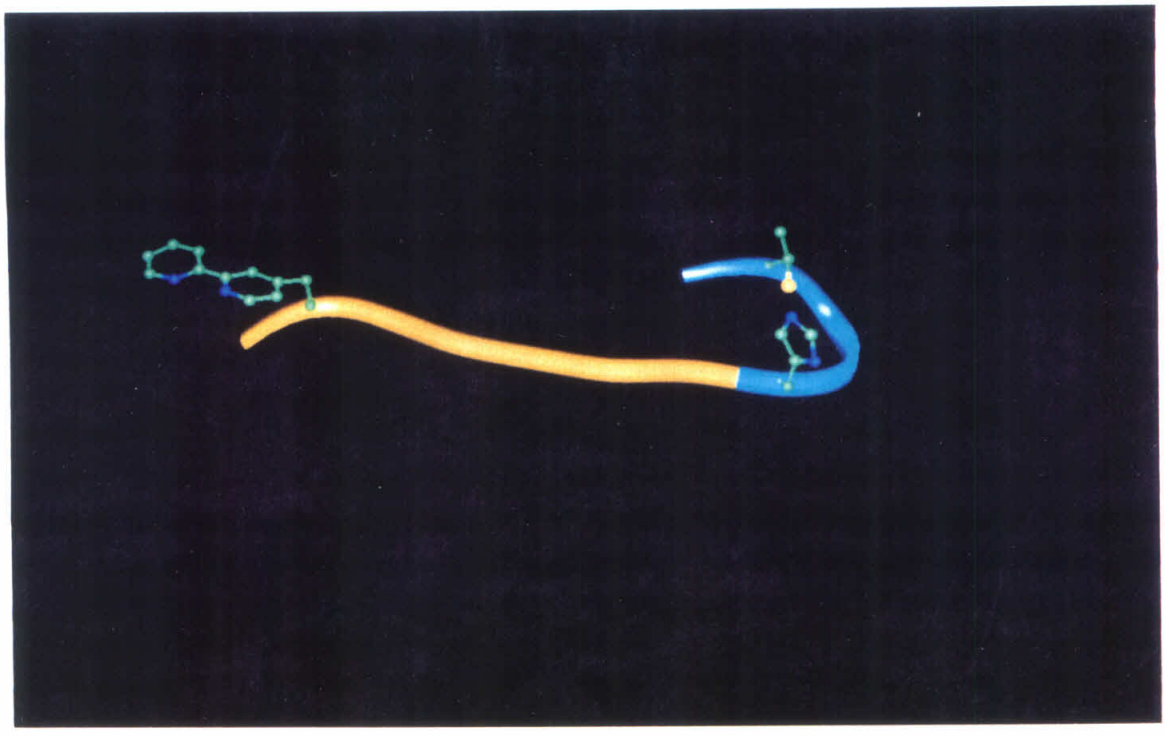
Finally, although the results from the metal binding titrations support the formation of a monomeric, intramolecular complex involving single metal cation and one peptide, the data is also consistent with the formation of an intermolecular homodimer complex composed of two metal cations bound by a pair of peptides. However, the latter complex is considered to be unlikely since the titrations were found to be independent of peptide concentration ([Pep] = 0.5-100 μM) and pH (pH = 6.5-8.5). If the intermolecular complex were the predominant species, the metal binding titrations would be expected to be dependent upon the peptide concentration, especially when the spectra were collected under extremely dilute conditions ([Pep]≈0.5 μM; see Figure 7-4).

These conclusions were supported by results from electrospray mass spectrometry analysis of the peptide metal complexes. Electrospray mass spectrometry has been shown to be an extremely sensitive technique for investigating peptide:metal interactions, and it has been used with success on peptides containing sulfur or imidazole groups in the presence of divalent cations (Cu(II), Zn(II)).[12-14] The technique is ideal for analyzing mixtures of peptide:metal complexes, as multiple species can be delineated through distinctive mass/charge (*m/z*) distributions. In addition, since this technique requires dilute solutions for accurate determinations,[12-14] it is ideally suited for the study of the peptide **1**: metal complexes investigated herein. Mass spectra were collected for both the Zn^{+2} and Co^{+2} complexes of peptide **1**, prepared under dilute conditions where the two-state equilibrium model is followed (20 μM peptide). The results from these studies are shown in Figure 7-8 and are presented in Table 7-2. The results from the spectra collected for the peptide without metal cations is shown for comparison.

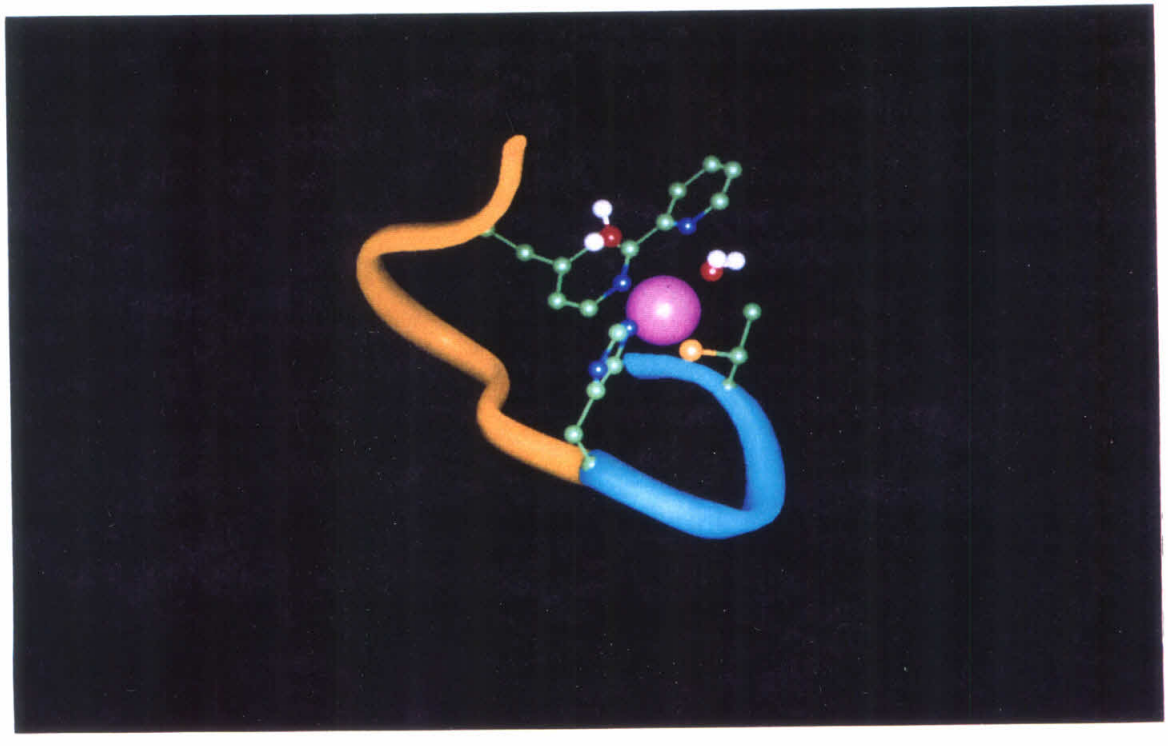
All of the peptide samples show evidence of contamination by sodium cations, as the corresponding singly- and doubly-charged sodium:peptide complexes are observed. Despite this impurity, inspection of the mass spectra for both metal:cation complexes

Figure 7-7. Proposed models of (A) peptide **1** in an extended conformation in the absence of metal cations and (B) the peptide **1**:metal cation complex. The teal tube denotes region 1, which incorporates the reverse turn containing the thiol and imidazole ligands (β-methyl-cysteine, histidine), and the yellow tube indicates the polyalanine linker region (region 2). The ligating groups are colored by atom with the following color code: red denotes oxygen atoms, green represents carbon atoms, blue corresponds to nitrogen atoms, white represents hydrogen, and the sulfur atom is yellow. The metal cation [Co(II), Zn(II)] is shown in purple.

A



B



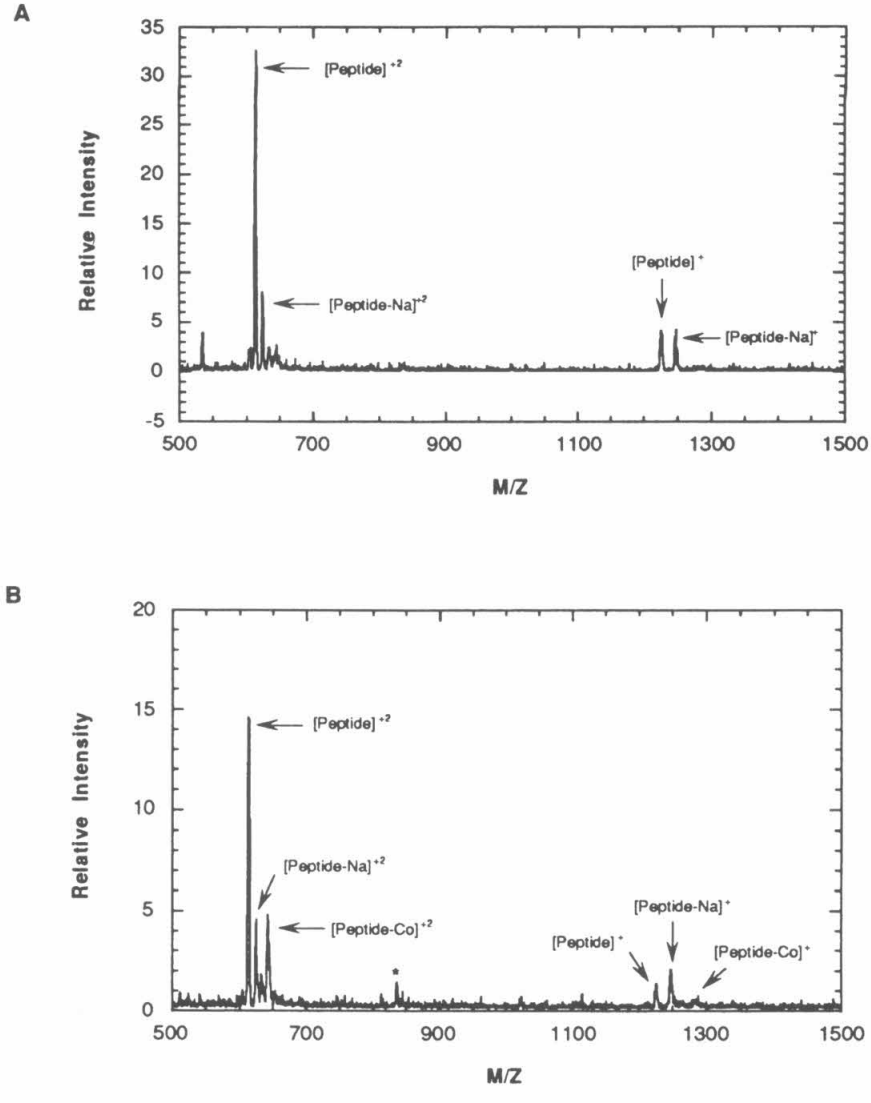


Figure 7-8. Electrospray mass spectra of (A) peptide **1**; (B) the peptide **1**:Co^{+2} complex; and (C) the peptide **1**:Zn^{+2} complex. (* and Δ represent triply-charged dimer products--see text for explanation.)

C

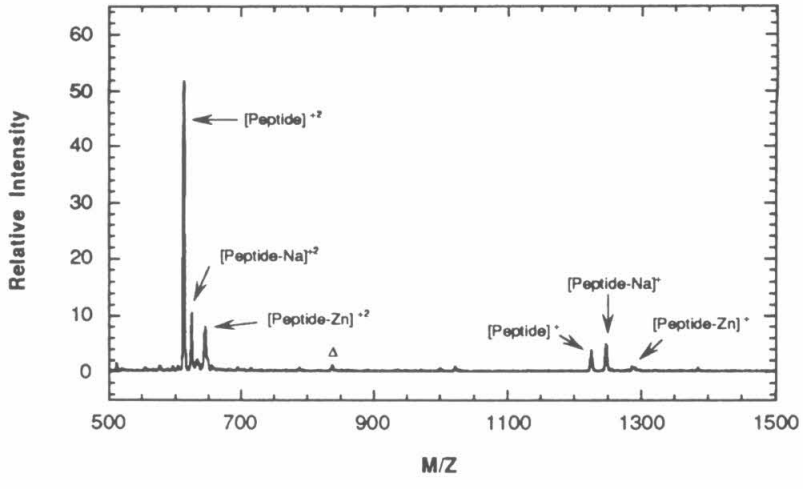

Relative Intensity
60
50
40
30
20
10
0
[Peptide] +2
[Peptide-Na]+2
[Peptide-Zn] +2
Δ
[Peptide] +
[Peptide-Na]+
[Peptide-Zn] +
500
700
900
1100
1300
1500
M/Z

Table 7-2. Determined Mass Values for the Stable Metal Cation Complexes of Peptide **1** Observed by Electrospray Mass Spectrometry

Complex	*m/z* Observed; [Calculated][a,b] (Charge state)	
peptide **1**	1225.6; [1225.5] (+1)	613.5; [613.3] (+2)
peptide **1**-Na	1247.5; [1247.5] (+1)	624.3; [624.3] (+2)
peptide **1**-Zn	1288.5; [1289.9] (+1)	645.00; [645.4] (+2)
peptide **1**-Co	1282.8; [1283.4] (+1)	641.7; [642.2] (+2)

[a] Mass values used in calculations: peptide : 1224.5 amu; Na(I) atom : 23.00 amu; Zn(II) atom : 65.39; Co(II) : 58.93 amu.

demonstrates that a monomeric, 1:1 complex is favored under the conditions employed for the mass spectrometry analysis. In both cases, the doubly-charged species exhibits the strongest signal; however, singly-charged species also display weak, but observable, peaks. Surprisingly, the observed masses of the complexes do not correspond to those calculated for an octahedral geometry containing two auxiliary coordinated water molecules. This may be attributed to the electrospray method of ionization, as the elevated temperatures employed in this technique (block temperatures ≈225 °C) ensure that excess solvent molecules are removed from the ionized solute species. It is possible that the peptide:metal complexes dehydrated as a result of these conditions.

The formation of dimers and higher-order aggregates can be distinguished from the monomeric species from the peaks arising from higher charge states. For example, dimer formation can be detected through the presence of a peak corresponding to the triply-charged species. For the peptide **1**:metal complexes studied herein, the homodimer

consisting of two metal cations and two peptides would be expected to have a triply-charged peak at ≈855 amu. The mass spectra of the complexes do not show any evidence for the expected; however, weak signals are observed in both spectra (denoted * or Δ) which correspond to dimeric complexes with the general formula: $M(II)[Pep]_2^{+3}$. The formation of these complexes could be due to an artifact of the electrospray technique since the ionization step requires evaporation of the solvent to form detectable ions. As a result, the sample would be expected to increase in concentration and therefore promote the formation of intermolecular complexes. However, the weak intensity of the observed signals suggests that these processes represent a minor component of the overall sample composition.

Conclusions

The results from this chapter demonstrate the ability to program high affinity metal binding sites into polypeptide constructs using both the reverse turn motif and the unnatural amino acid, 4Bpa. The dodecapeptide **1**, designed to bind metal cations in a tetradentate fashion through the thiol group of the β-methyl-cysteine side chain and the nitrogen atoms of the 4Bpa and histidine residues, bound Co^{+2} and Zn^{+2} metal cations in a 1:1 stoichiometry with affinities well beyond that predicted for the individual ligands based on metal cation titrations monitored with UV/Vis absorption spectroscopy. In fact, the metal binding dissociation constants were found to be at or below the mid-nanomolar range. Binding constants of this magnitude have precedent in native biological systems; dissociation constant for metal ions in the Zn^{+2} finger proteins have been found to be in this range.[3] At low peptide concentrations the complex formation followed a two state equilibrium model as metal cation titrations contained well-defined isosbestic points. The bipyridyl residue was clearly active in the formation of the metal complex, as distinctive red shifts of the bipyridyl π-π* absorption were observed for both Zn^{+2} and Co^{+2} metal cations. Evidence for the participation of the β-methyl-cysteine residue in the binding

process was obtained from the absorption spectra of the Co^{+2} complex. Upon addition of Co^{+2} cations, a new band was observed on the shoulder of the bipyridyl absorption. This band was found to be unique to the Co^{+2}:peptide **1** complex; the corresponding Zn^{+2} complex did not show an enhanced absorption in this region. The new absorption band was assigned to a ligand-to-metal charge transfer interaction between the thiol group and the Co^{+2} cation. The bipyridyl ligand and the β-methyl-cysteine are believed to act cooperatively in the formation of the metal peptide complex, as the absorptions attributed these ligands share the same metal concentration dependence.

Formation of the peptide **1**:Co^{+2} complex was found to be highly concentration dependent. At peptide concentrations higher than 100 μM, multiple equilibria were observed in the metal cation titrations as evidenced by the loss of clearly defined isosbestic points. As a result, preliminary attempts to confirm the histidine ligation to the Co^{+2} metal cation using ^1H NMR were unsuccessful, as the sensitivity of this technique was not sufficiently high enough to detect the peptide **1**:Co^{+2} complex under the conditions required to form the 1:1 peptide:metal cation species. However, it is likely that the histidine is active in the complex formation as a result of the reverse turn propensity of sequence containing the histidine residue. Work in this lab using octapeptides containing histidine residues has shown that the reverse turn can actively nucleate metal binding motifs in aqueous solution, by directing the imidazole side chain groups to reside in close proximity with other ligands to form a multi-dentate binding site.[15] It is believed that a similar process is involved in the formation of the peptide **1**:metal complex.

Results from EPR experiments, in conjunction with analysis of the visible spectrum of the Co^{+2}:peptide **1** metal complex, suggest that the coordination geometry of the complex is octahedral and that the metal is high spin. The complexes are assumed to be monomeric in solution as the complex formation does not show a concentration dependence under very dilute conditions. This assumption was supported by the results from electrospray mass spectrometry analysis of the peptide:metal cation complexes, as the

predominant product observed in these studies was a complex corresponding to the 1:1 metal cation:peptide species. In summary, the results presented in this chapter hold promise for the successful generation of motifs programmed to regulate naturally occurring enzymes which modulate their action through metal-induced conformational changes. Subsequent designs could test the effects of added functionality in the linker region on metal complex formation, as well as further investigate the role of the reverse turn in stabilizing and nucleating the metal binding motif near the amino terminus.

Acknowledgment

Helpful discussions with Prof. Harry Gray and his research group, Michael Murray, Jeff Hagen, and Prof. Curtis Monig are greatly appreciated.

Experimental

All buffer solutions were prepared with Ultrapure H_2O that was freshly distilled under an inert atmosphere (N_2). These solutions were subsequently sparged with H_2/N_2 gas mixtures prior to use to avoid intermolecular disulfide formation. Similarly, preparation of peptide solutions was performed under an inert atmosphere via syringe whenever possible, and stock solutions of the peptide were stored frozen under N_2 at -80° C at all times. The free thiol content in the stock solutions was checked regularly using the Ellman's test,[6] and the extent of oxidation was evaluated through comparison of the peptide concentrations determined spectrophotometrically from the bipyridyl π-π^* absorbance (λ_{max} = 286 nm, ε = 12,800 M^{-1} cm^{-1}).

All metal cation titrations were performed as described in Chapter 5, using 50 mM HEPES buffer. Absorption spectra of the Co^{+2} complex was obtained in a 1 cm pathlength cell on a Cary-14 UV/Vis spectrometer in the research laboratories of Professor Harry Gray. EPR measurements were performed in the laboratories of Dr. Sam Kim at the Jet Propulsion Laboratory, Caltech, on a Bruker ESP-300 Electron Spin Resonance spectrometer at 77 K. The EPR samples were prepared at 20 μM peptide, 15 μM Co^{+2} in

1:3 MeOH/buffer (50 mM HEPES, pH=8.25) and frozen in liquid N_2 just prior to analysis.

*Molecular Modeling of Proposed peptide **1**:Co^{+2} Complex*

The illustration of the peptide **1**:Co^{+2} complex (See Figure 7-7) was performed on a Silicon Graphics Personal Iris 4D/25TG computer using Insight II v.2.1.0 Modeling Package (Biosym Technologies). The peptide **1**:Co^{+2} complex structure was constructed using ideal Type II dihedral angles for the sequence Ac-Bmc-Pro-D-Ser-His. The dihedral angles for residues 6-10 (Ala-Ala-Ala-Ala) were initially set at the values for an ideal Type I turn to bring the bipyridyl residue in close proximity to the Bmc (residue 1) and His (residue 4) side chains. The metal ion was represented as a pseudoatom at the center of mass between the Bmc-Sγ, His-Nε, and bipyridine nitrogen atoms. Generic distance constraints were then generated between the metal ligating atoms and the pseudoatom with the following parameters: the generic distances were bounded between 2.0-2.5 Å, the upper and lower distance force constants were set at 10.00, and the maximum force per iteration was set at 1000. The complex was then subjected to 10000 iterations of a steepest descent minimization to normalize the structure. The water molecules of the complex were then added and these molecules were constrained to the pseudoatoms using identical distance constraints identical to those listed above. The complex was then subjected to 10000 iterations of a conjugate gradient minimization. The proposed model is intended to represent a possible peptide conformation for the metal complex; in reality, the complex is expected to assume an ensemble of similar structures.

*Electrospray Mass Spectrometry Analysis of the peptide **1**: Metal Cation Complexes*

Electrospray mass spectrometry analysis of the peptide **1**:metal cation complexes was performed by Jeff Hagen (Graduate Student) in the laboratories of Prof. Curtis Monig at the University of California, Riverside Campus. The spectra were collected on a Vestec

Model 201 single-quadrupole mass spectrometer (Vestec Corporation, Houston, TX), fitted with a 10 kV conversion dynode and a 2000 *m/z* mass range. The samples were ionized using a modified electrospray interface; a flat plate was used for the nozzle. The samples were infused at a rate of 1.83 μL/min using a syringe pump (Orion Research, Boston, MA). Typical operating conditions for the samples were as follows: electrospray ionization interface : 3.86 kV; repellor voltages ranged from 10V to 20V (no change in the spectrum was observed), nozzle voltage : 200 V; spray current : ≈ 200 μAmps. The temperatures of the electrospray chambers were as follows: spray chamber : 52 °C; block temperature : 238 °C; ionizing chamber : 148 °C. The second stage vacuum chamber was maintained at a pressure of 0.13 torr. Typical spectra were collected with a scan range of 100-2000 *m/z* range using a rate of 3 scans/sec. The mass spectrometer had been calibrated over the entire mass range to within 1 amu using a polyethylene glycol standard. Peptide samples were prepared in ultrapure H_2O (20 μM Pep), with or without metal cations (20 μM M^{+2}; metal solutions were prepared from the chloride salt), and diluted with an equal volume of methanol. The samples were prepared just prior to analysis to avoid ambiguities due to disulfide formation. Interestingly, upon the addition of acetic acid, even under dilute conditions (< 1%), no spectra were observed for either the peptide or the peptide:metal complexes.

Spectroscopic Data for peptide ***1***

RP-HPLC Gradient Used:

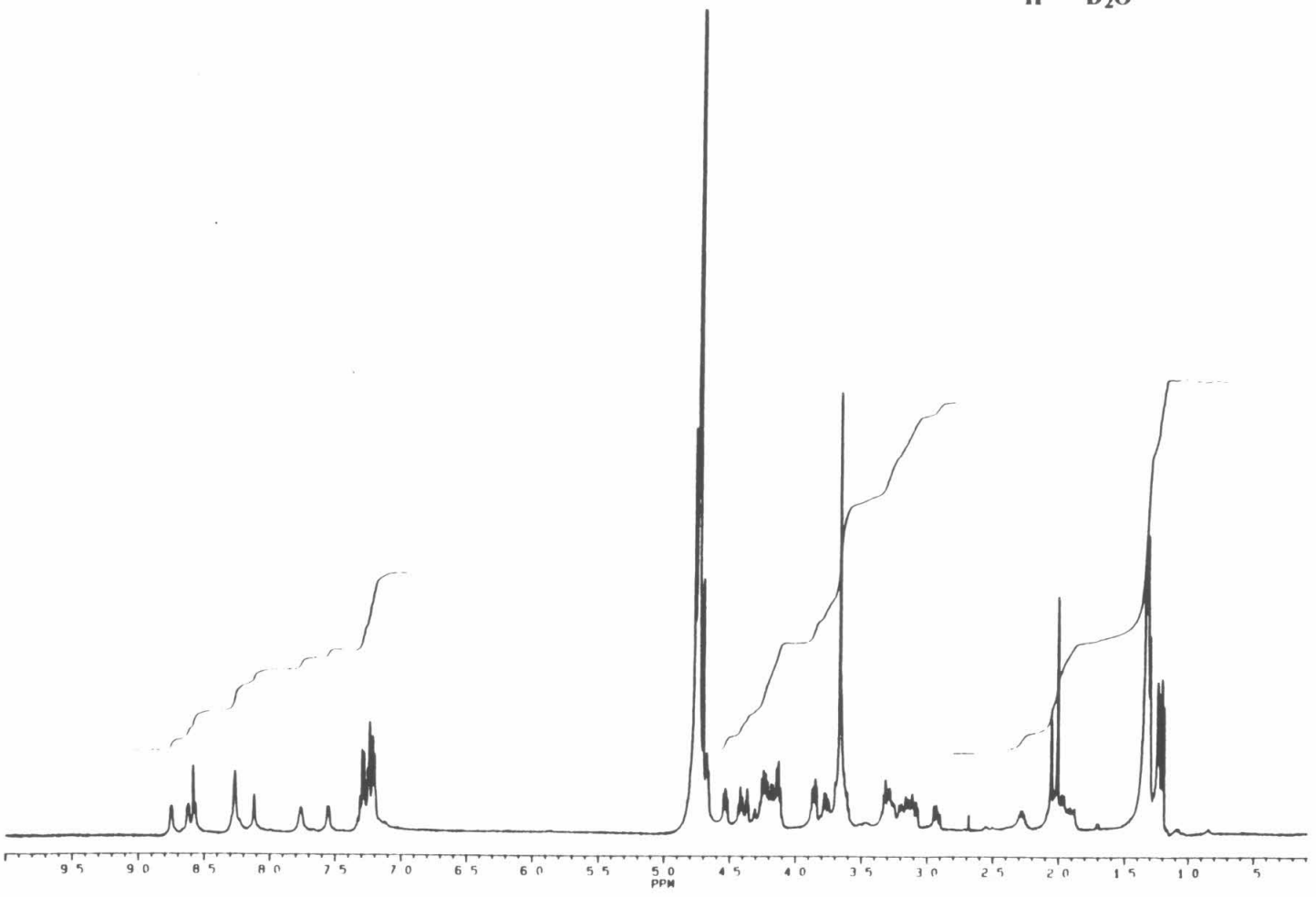
(E) 20% CH_3CN (0.08% TFA)/H_2O (0.1% TFA) 5 min, Linear Gradient 20-50% CH_3CN (0.08% TFA)/H_2O (0.1% TFA) 25 min, 50% CH_3CN (0.08% TFA)/H_2O (0.1% TFA) 2 min, Linear Gradient 50-20% CH_3CN (0.08% TFA)/H_2O (0.1% TFA) 2 min.

peptide **1**: RP-HPLC (E): RT = 20.2 min.

HRMS: Calcd. for $[MNa]^+$ $C_{57}H_{76}NaN_{16}O_{13}S$ [1247.5396]; Obsd. [1247.5430].

Ac-Bmc-Pro-(D)-Ser-His-[Ala]$_5$-4Bpa-Phe-NH_2

^1H D_2O



References

(1) Vallee, B.L., Auld, D.S. "Zinc: Biological Functions and Coordination Motifs," *Acc. Chem. Res.* **1993**, *26,* 543-551.

(2) Selection of the optimal sequence length and metal binding ligands for the designed motif was based on examination of hand-held CPK models of the peptide:metal cation complex.

(3) Micheal, S.F.; Kilfoil, V.J.; Schmidt, M.H.; Amann, B.T.; Berg, J.T. "Metal binding and folding properties of a minimalistic Cys_2His_2 zinc finger peptide," *Proc. Natl. Acad. Sci., USA* **1992**, *89,* 4796-4800.

(4) Knighton, D.R.; Zheng, J.; Ten Eyck, L.F.; Xuong, N-H.; 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.

(5) Bertini, I.; Luchinat, C. "High Spin Cobalt(II) as a Probe for the Investigation of Metalloproteins," *Adv. Inorg. Biochem.* **1984**, *6,* 71-111.

(6) Ellman, G.L. "Tissue Sulfhydryl Groups," *Arch. Biochem. Biophys.* **1959**, *82,* 70-77.

(7) Thompson, J.S.; Sorrell, T.; Marks, T.J.; Ibers, J.A. "Synthesis, Structure, and Spectroscopy of Pseudotetrahedral $Co^{II}N_3(SR)$ Complexes. Active Site Approximations to the Cobalt(II)-Substituted Type I Copper Proteins," *J. Am. Chem. Soc.* **1979**, *101,* 4193-4200.

(8) Bertini, I.; Luchinat, C. "Cobalt(II) as a Probe of the Structure and Function of Carbonic Anhydrase," *Acc. Chem. Res.* **1983**, *16,* 272-279.

(9) Drago, R.S. *Physical Methods in Chemistry*; W.B. Saunders Co.: Philadelphia, PA, **1977**; pp 467-509.

(10) Personal communcation with Dr. Angelo De Bilio in the laboratories of Professor Harry Gray, California Institute of Technology.

(11) Carlin, R.L. *Magnetochemistry*; Springer-Verlag: New York, **1986**.

(12) Hutchens, T.W.; Allen, M.H. "Differences in the Conformational State of a Zinc-finger DNA-binding Protein Domain Occupied by Zinc and Copper Revealed by Electrospray Ionization Mass Spectrometry," *Rapid Comm. Mass Spect.* **1992**, *6*, 467-473.

(13) Hutchens, T.W.; Nelson, R.W.; Allen, M.H.; Li, C.M.; Yip, T-T. "Peptide-Metal Ion Interactions in Solution: Detection by Laser Desorption Time-of-flight Mass Spectrometry and Electrospray Ionization Mass Spectrometry," *Biol. Mass Spect.* **1992**, *21*, 151-159.

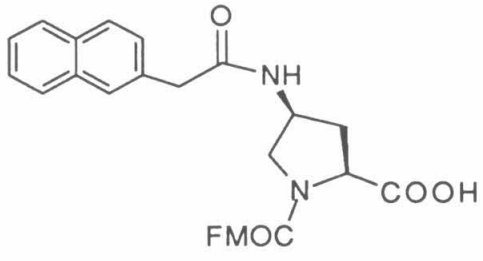
(14) Allen, M.H.; Hutchens, T.W. "Electrospray-ionization Mass Spectrometry for the Detection of Discrete Peptide/Metal-ion Complexes Involving Multiple Cysteine (Sulfur) Ligands," *Rapid Comm. Mass Spect.* **1992**, *6*, 308-312.

(15) Imperiali, B.; Kapoor, T.K. "The Reverse Turn as a Template for Metal Coordination," *Tetrahedron* **1993**, *49*, 3501-3510.

Appendix I: Synthesis of N^{α}-(9-fluorenylmethoxycarbonyl)-(2*S*,4*S*)-4-(naphth-2-yl)-amido-proline

Introduction

In nature, proline is believed to play a critical role in the formation and stabilization of the complex folded structures of proteins.[1] This ability to constrain and direct protein structure has been attributed to the restricted main chain ϕ dihedral angle of the proline residue due to the cyclic pyrrolidine ring side-chain. However, while the proline residue clearly serves a key *structural* role in native proteins, it cannot play an active part in protein function, as the alkyl side-chain does not contain any reactive functional groups. Ideally, proline derivatives with a functionalized side-chain could be employed in protein constructs to both induce and stabilize the overall structure, as well as play an active role in a pre-programmed function. In an effort to realize this goal, the synthesis of a proline derivative containing a fluorescent probe was undertaken.



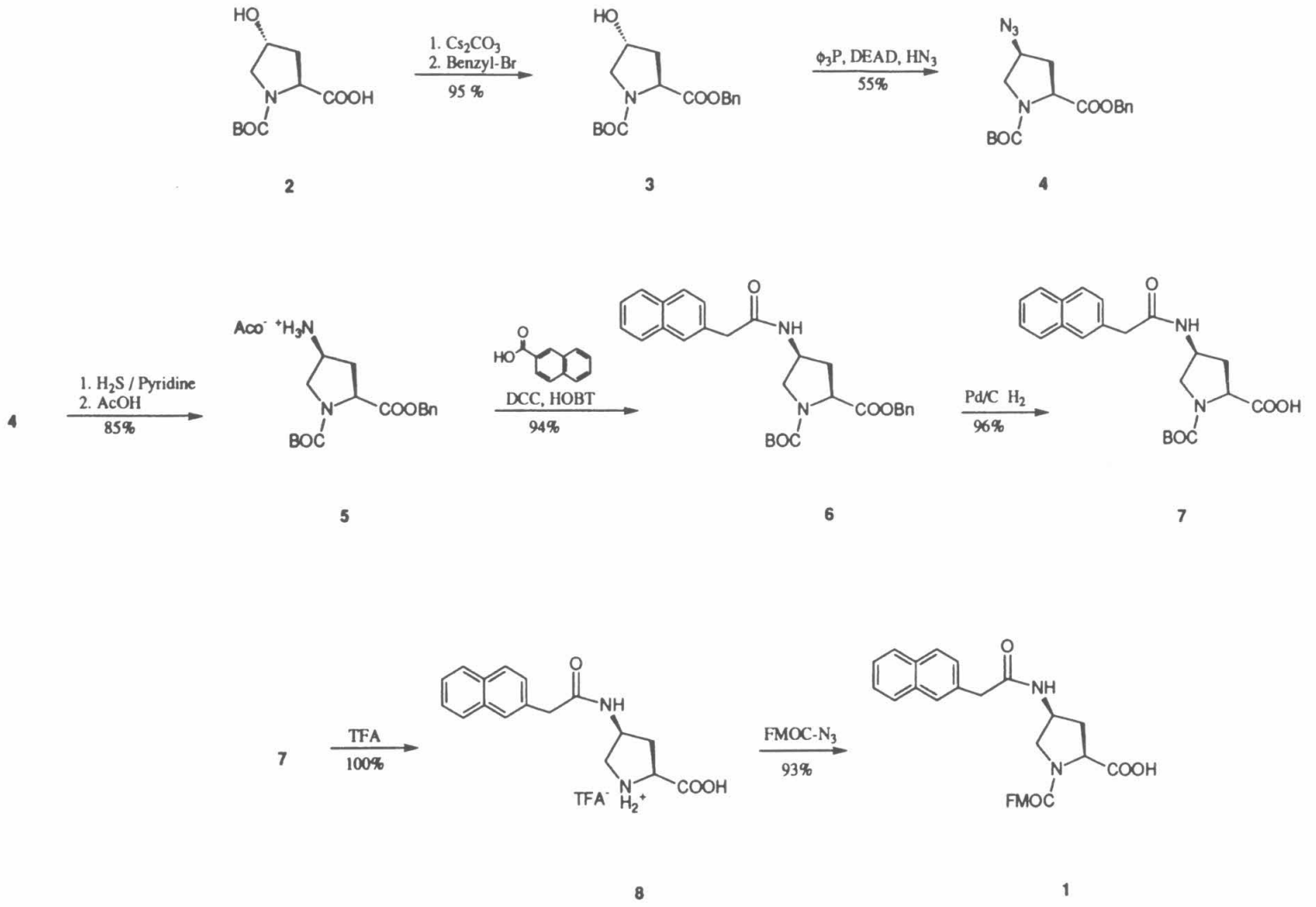
1

The target, N$^{\alpha}$-(9-fluorenylmethoxycarbonyl)-*cis*-(2*S*, 4*S*)-4-(naphth-2-yl)-amido-proline (FMOC-NapPro, **1**) contains a naphthyl group stereospecifically attached to the C-4 of the pyrrolidine ring through an amide linker. The amino acid derivative was designed to incorporate functionalities on the γ-carbon to minimize with the main chain atoms when incorporated into larger polypeptides; substitutions at the β and δ carbons of proline have been shown to dramatically affect the overall polypeptide conformation through interactions with the main chain amide groups.[2] In addition, the naphthyl group was chosen as it is a relatively simple fluorescent probe consisting of only two fused aromatic rings with no

additional reactive functionality which could interact with the main chain atoms and would require extensive protecting group strategies in the synthetic preparation. The stereospecific linkage of the reactive groups served to properly orient the introduced functionality when the proline derivative was incorporated into the Type II reverse turn (see Chapter 2). After inspection of the reverse turn structure, it was observed that attachment of the functional group with the *cis* (2*S*,4*S*)configuration would place the introduced groups on the same planar face of the reverse turn as the side-chains of the residues at positions *i* and *i+3*. It was anticipated that in this orientation, the naphthyl group could be targeted to participate in long-range interactions with the residues at positions *i* and *i+3*.

Results and Discussion

The synthesis of **1** was accomplished using an amino protected derivative of the *trans* (2*S*,4*R*) diastereomer of 4-hydroxyproline. Since this amino acid is a constitutive element of collagen, it is both relatively inexpensive and readily available in diastereomerically pure form. The key step of the synthetic route to **1** involves conversion of the alcohol to an azide under Mitsunobu conditions. The fully protected *trans*-(2*S*,4*R*)-hydroxy-proline derivative **3**, readily prepared from the commercially available free acid **2** using the conditions described by Wang *et al.*[3] for synthesis of aromatic esters under neutral conditions, was found to undergo facile azide displacement in the presence of triphenylphosphine, diethylazodicarboxylate (DEAD), and freshly prepared hydrazoic acid solution at low temperatures (see Scheme AI-1). The displacement was assumed to be stereospecific, as the Mitsunobu reaction mechanism is known to proceed through an S_N2 mechanism,[4,5] and similar secondary alcohols have been found to proceed with complete inversion and similar reaction product yields.[6] In addition, inspection of the product by thin layer chromatography and the ^1H NMR spectrum indicated only one diastereomer was present. While the yield for the conversion is relatively low, the reaction can be scaled to produce preparative quantities (>5g) of the desired azide **4**, with the upper limit being



Scheme A1-1. Synthetic Route to (2*S*,4*S*)-N$^{\alpha}$-FMOC-*cis*-4-(naphth-2-yl)amido-proline

restricted only by the constraints of the flash column chromatography technique.[7] Reduction of the azide **4** to the desired amine salt **5** was accomplished using hydrogen sulfide in pyridine, as described by Adachi, Yamada, and Inoue.[8] The reduction proceeded smoothly in high yield (85-90%) and the reaction product was judged to be >95% pure by TLC and NMR following reaction workup.

Since the amino-salt **5** is fully protected, it can be acylated using standard peptide synthesis protocols.[9] Acylation of the amine with the hydroxybenzotriazole-activated ester of 2-naphthylacetic acid afforded the protected naphthylated derivative **6** in >90% yield after purification by flash chromatography. Subsequent deprotection of the benzyl ester by catalytic hydrogenation and removal of the *tert*-butoxycarbonyl amine protecting group with acid produced the functionalized proline amino acid in near quantitative yield from the amine salt **5**. Finally, since this amino acid was designed for incorporation into larger polypeptide sequences, the amino acid was subsequently protected with FMOC-azide in 50% aqueous dioxane under basic conditions (93% yield). It should be noted that while FMOC-azide was employed for the N^{α}-protection step, it is likely that commercially available FMOC-succinimide[10] would provide similar, if not improved, yields of the target compound.

Conclusions

The synthetic route to the amine-protected, naphthylated-proline derivative **1** has several notable points. First, since the reactions employed in the synthesis are amenable to preparative scales and purification does not pose a problem for many of the intermediates, it is possible to prepare the target amino acid in gram scale quantities. The ability to generate large quantities of FMOC-protected unnatural amino acids is a prerequisite for the use of these materials in *de novo* protein design, as typical solid-phase peptide synthesis protocols require multi-milligram quantities (>100mg) per residue coupling. Second, the synthetic route provides a number of useful intermediates which could be exploited for the

production of a wide range of highly functionalized proline derivatives. The amine derivative **5**, in particular, could be used to form a wide array of acylated intermediates with virtually unlimited functionality using conditions similar to those employed for the production of **6**. However, difficulties in the preparation of different analogs may arise as a result of incompatibilities of the introduced groups with the subsequent deprotection steps (catalytic hydrogenation, acid). Subsequent to this work, a synthetic route was reported for the preparation of an intermediate similar to **5**[11] along an analogous route. In this case, the protecting groups for the carboxylic acid group was different than the one employed here (methyl ester), and this route may offer some additional scope in the preparation of new proline derivatives.

Experimental

N$^{\alpha}$-(*t*-Butoxycarbonyl)-*trans*-(2*S*,4*R*)-4-hydroxy-proline-benzyl ester (**3**): A solution of N$^{\alpha}$-(*t*-butoxycarbonyl)-*trans*-(2*S*,4*R*)-4-hydroxy-proline **2** (Bachem Biosciences, Inc.) (0.904 g, 3.89 mmol) in 18 mL 5:1 MeOH/H_2O was brought to pH = 7 with 20% Cs_2CO_3 solution, and concentrated *in vacuo* to a clear oil. Residual water was removed azeotropically with toluene. The oil was then dissolved in 9.5 mL DMF and 0.5 mL (0.7 g, 4.2 mmol) of benzyl bromide was added. After stirring at room temperature for 5 hr the solution was concentrated *in vacuo*, suspended in 20 mL H_2O and extracted ethyl acetate (3 x 20 mL). The organic fractions were washed with 50 mL H_2O, dried over Na_2SO_4, and concentrated *in vacuo* to yield 1.3 g (>95%) of a clear oil. TLC (7:3 ethyl acetate:hexanes; R_f = 0.47 ninhydrin after HCl); ^{1}H NMR ($CDCl_3$) δ: 1.28-1.40 (s, 9H), 2.00 (m, 1H), 2.20 (m, 1H), 3.3-5 (m, 2H), 5.08 (m, 2H), 7.27 (m, 5H); ^{13}C NMR ($CDCl_3$) δ: 28.4, 38.5, 39.1, 54.7, 57.7, 66.8, 69.5, 128.3; IR (neat, cm^{-1}) : 3438, 2976, 1746, 1694, 1498, 1478, 1455, 1415, 1367, 1258, 1157, 1086, 1005, 981, 773, 751, 698.

N^{α}-(*t*-Butoxycarbonyl)-*cis*-(2*S*,4*S*)-4-azido-proline-benzyl ester (**4**): To a 10 ml round bottom flask was added, in this order, triphenyl phosphine (0.9039 g, 3.43 mmol) in 2.5 ml freshly distilled THF, diethyl-azodicarboxylate (0.51 mL, 3.43 mmol) , 3.5 mL (3.43 mmol) freshly prepared 1M HN_3 in benzene,[12] and a solution of **3** (1.1051 g, 3.43 mmol) in 1.2 mL freshly distilled THF at -78 °C, with stirring. The solution was then allowed to warm to room temperature over 1hr. After 4 hr the mixture was concentrated *in vacuo* to yield a white amorphous solid. The crude product was purified using flash chromatography (2:1 hexane:ether eluent, 4 cm x 12 cm column), to yield 0.6362 g (1.83 mmol, 55%) of a clear oil. TLC (SiO_2, short λ UV; 2:1 ether:hexane) R_f = 0.28; ^{1}H NMR ($CDCl_3$) δ: 1.20-1.40, (s, 9H), 2.10 (m, 1H), 2.40 (m, 1H), 3.44 (m, 1H), 4.06 (m, 1H), 4.28 (m, 1H), 5.12 (m, 2H), 7.28 (m, 5H); ^{13}C NMR ($CDCl_3$) δ: 14.4, 35.1, 36.2, 47.1, 51.1, 51.5, 57.6, 57.9, 58.2, 59.2, 62.3, 67.2, 67.6, 67.7, 119.9, 125.0, 125.1 ,127.1, 127.7, 128.4, 128.5, 141.3, 143.6, 144.0, 170.8; IR (neat, cm^{-1}) : 2977, 2105, 1754, 1704, 1498, 1458, 1393, 1366, 1259, 1214, 1158, 1118, 1054, 1004, 968, 890, 751, 698.

N^{α}-(*t*-Butoxycarbonyl)-*cis*-(2*S*,4*S*)-4-amino-proline-benzyl ester (**5**): A sample of **4** (0.3957 g, 1.14 mmol) was dissolved in 30 mL of 1:1 H_2O/pyridine and a vigorous stream of H_2S gas was passed through the solution for 3hr. (Caution! H_2S is extremely toxic. Handle with care and pass outlet gas stream through a saturated KOH/ H_2O trap. Wash all glassware with $NaClO_4$ solution (bleach) after use). The orange solution was concentrated *in vacuo* to remove excess pyridine and neutralized with 6N acetic acid. The solution was then frozen and lyophilized overnight to a yield a yellow, amorphous solid which was taken up in 20 mL MeOH, filtered and concentrated to yield 0.3713 g (1.0 mmol, 85%) of a clear yellow oil. TLC (SiO_2, ninhydrin, short λ UV; 4:1:1 butanol: acetic acid:water) R_f = 0.55; ^{1}H NMR (CH_3OD; spectrum showed two distinct conformations) δ: 1.28, 1.43 (s, 9H); 1.90 (m, 4H), 2.64 (m, 1H), 3.29 (m, 1H), 3.78 (m, 2H), 4.33 (t, 1H, J=7.6Hz),

5.18 (m, 2H), 7.36 (m, 5H); ^{13}C NMR (CH_3OD) δ: 26.9, 27.2, 35.4, 51.0, 57.8, 57.9, 67.1, 80.9, 127.9, 128.0, 128.2, 128.3, 128.5, 181.1; IR (neat, cm^{-1}) : 2976, 1693, 1573, 1479, 1408, 1258, 1157, 1012, 922, 884, 751, 698, 653, 619.

N$^{\alpha}$-(*t*-Butoxycarbonyl)-*cis*-(2*S*,4*S*)-4-(naphth-2-yl)-amido-proline-benzyl ester (**6**): A solution of 2-naphthylacetic acid (0.2684 g, 1.44 mmol), dicyclohexylcarbodiimide (0.3146g, 1.52 mmol), and 1-hydroxybenzotriazole hydrate (0.2030g, 1.50 mmol) were dissolved in 3 mL CH_2Cl_2. After 45 min. the mixture was filtered into a solution of **5** (0.4944g, 1.3 mmol) and triethylamine (0.2 mL, 1.44 mmol) in 3 mL CH_2Cl_2. After stirring 2 hr at room temperature the mixture was concentrated *in vacuo* and purified using flash chromatography, (1:1 hexanes:ethyl acetate eluent, 2 cm x 12 cm column) to yield 0.6362 g (1.22 mmol, 94%) of a clear oil. TLC (SiO_2, ninhydrin, short λ UV; 1:1 ethyl actate:hexanes) $R_f = 0.33$; ^1H NMR ($CDCl_3$; spectrum showed two distinct conformations) δ: 1.27, 1.42 (s, 9H); 1.80 (m, 1H), 2.37 (m, 1H), 3.46 (m, 1H), 3.60 (m, 3H), 4.20 (m, 1H), 4.80 (m, 3H), 6.60 (d, 1H, J=8.8Hz), 7.18 (m, 2H), 7.33 (m, 4H); 7.44 (m, 2H), 7.70 (s, 1H), 7.80 (d,3H, J=7.4Hz); ^{13}C NMR ($CDCl_3$; spectrum showed two distinct conformations) δ: 28.1, 28.4, 35.7, 36.8, 44.0, 47.9, 48.9, 52.9, 53.3, 57.8, 67.0, 80.7, 125.9, 126.3, 127.3, 127.6, 127.7, 127.8, 128.2, 128.4, 128.6, 128.7, 132.2, 133.7, 135.1, 153.5, 170.6, 173.9; IR (neat, cm^{-1}) : 2976, 1693, 1573, 1479, 1408, 1258, 1157, 1012, 922, 884, 751, 698, 653, 619.

N$^{\alpha}$-(*t*-Butoxycarbonyl)-*cis*-4-(naphth-2-yl)-amido-proline (**7**): A solution of **6** (0.44g, 0.84 mmol) in 2 mL MeOH was degassed with N_2 and a sample of 10% Pd/carbon catalyst (120 mg) was added with stirring. The reaction mixture was charged with an atmosphere of H_2 and stirred at room temperature for 5 hr. The solution was then filtered through celite and concentrated to yield 0.3505g (0.82 mmol, 96%) of a clear oil. TLC (SiO_2, short λ UV; 4:1:1 butanol: acetic acid:water) $R_f = 0.75$; ^1H NMR ($CDCl_3$; spectrum

showed two distinct conformations) δ: 1.2-1.3 (s, 9H), 1.80, 2.20 (m, 1H), 2.40, 2.72 (m, 1H), 3.40 (m, 1H), 3.60 (m, 3H), 4.25, 4.35 (m, 1H), 4.50 (m, 1H), 6.95 (m, 1H), 7.3-7.5 (m, 3H), 7.6-7.9 (m, 4H); ^{13}C NMR ($CDCl_3$; spectrum showed two distinct conformations) δ: 23.6, 23.7, 28.8, 29.7, 34.3, 37.0, 44.1, 48.8, 49.8, 52.9, 54.0, 58.2, 58.7, 81.5, 82.5, 126.4, 126.8, 127.8 128.2, 128.5, 128.7, 129.2, 130.1, 132.6, 133.0, 134.1, 154.3, 156.4, 172.0, 175.4, 177.0, 181.7; IR (neat, cm^{-1}) : 3382, 2931, 1650, 1557, 1407, 1366, 1160, 853, 764.

cis-4-(Naphth-2-yl)-amido-proline, triflouroacetic acid salt (**8**): A solution of **7** (0.3505g, 0.82 mmol) in 2 mL of 1:1 CH_2Cl_2/ triflouroacetic acid was stirred at room temperature for 30 min. The solution was then concentrated under an N_2 sparge, then taken up in 5 mL MeOH and concentrated *in vacuo* to yield 0.3505g (0.80 mmol, 97%) of a white amorphous solid. TLC (SiO_2, short λ UV; 4:1:1 butanol: acetic acid:water) $R_f = 0.25$; ^1H NMR (CD_3OD; spectrum showed two distinct conformations) δ: 1.20 (m, 1H), 1.80 (m, 1H), 2.15 (m, 1H), 2.75 (m, 1H), 3.80 (m, 3H), 4.41 (m, 1H), 7.23 (m, 3H), 7.80 (m, 4H); ^{13}C NMR (CD_3OD) δ: 24.6, 32.9, 33.3, 42.1, 49.4, 58.3, 125.4, 125.8, 126.7, 127.2, 127.4, 127.8.

N^{α}-(9-fluorenylmethoxycarbonyl)-*cis*-4-(naphth-2-yl)-amido-proline (**1**): A solution of FMOC-N_3 (0.2758g, 1.04 mmol) in 9 mL 1,4-dioxane was added dropwise to a suspension of **8** (0.3505g, 0.8 mmol) in 9 mL 10% Na_2CO_3 and 1 mL MeOH at 0°C with stirring. The solution was allowed to warm to room temperature. After 30 hr the mixture was diluted with 30 mL H_2O and washed with 3 x 10 mL ether. The aqueous layer was cooled in an ice bath, brought to pH=2.0 (pH paper) with 6N HCl, and extracted with 3 x 25 mL EtoAc. The organic fractions were combined, dried over Na_2SO_4, and concentrated to yield 0.4107g (0.75 mmol, 93%) of a clear oil which formed a solid foam under vacuum. Calc. for $[MH]^+$ (Low Res.) $C_{32}H_{29}N_2O_5$ [521]; Obs. [521]; TLC (SiO_2, short λ UV; 10:1) $R_f = 0.25$; ^1H NMR ($CDCl_3$) δ: 2.1-2.3 (m, 2H), 3.35 (m, 1H), 3.5-3.7

(m, 3H), 4.0 (m, 1H), 4.2-4.35 (m, 3H), 4.5 (m, 1H), 6.6-6.7 (m, 1H), 7.1-7.4 (m, 9H), 7.5-7.8 (m, 6H); ^{13}C NMR ($CDCl_3$) δ: 43.4, 47.0, 49.5, 53.2, 67.0, 68.6, 120.0, 120.1, 124.9, 126.1, 126.4, 127.0, 127.2, 127.7, 127.9, 128.3, 128.9, 131.1, 141.3, 143.2, 172.1; IR (neat, cm^{-1}) : 3307, 3048, 2953, 1704, 1538, 1422, 1353, 1171, 1125, 1102, 1008, 967, 908, 758, 737.

References

(1) Brandon, C.; Tooze, J. *Introduction to Protein Structure*; Garland Publishing, Inc.: New York, **1991**; p 259.

(2) Delaney, N.G.; Madison, V. "Novel Conformational Distributions of Methylproline Peptides," *J. Am. Chem. Soc.* **1982**, *104*, 6635-6641.

(3) Wang, S-S.; Gisin, B.F.; Winter, D.P.; Makofske, R.; Kuleshe, I.D.; Tzougraki, C.; Meienhofer, J. "Facile Synthesis of Amino Acid and Peptide Esters under Mild Conditions via Cesium Salts," *J. Org. Chem.* **1977**, *42*, 1286-1290.

(4) Hughes, D.L.; Reamer, R.A.; Bergan, J.J.; Grabowski, E.J.J. "A Mechanistic Study of the Mitsunobu Esterification Reaction," *J. Am. Chem. Soc.* **1988**, *110*, 6487-6491.

(5) Varasi, M.; Walker, K.A.M.; Maddox, M.L. "A Revised Mechanism of the Mitsunobu Reaction," *J. Org. Chem.* **1987**, *52*, 4235-4238.

(6) Schweng, J.; Zbiral, E. "Synthese von Cyclishen Vinylaziden," *Justus Liebigs Ann. Chem.* **1978**, 1089-1095.

(7) Still, W.C.; Kahn, M.; Mitra, A. "Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution," *J. Org. Chem.* **1978**, *43*, 2923.

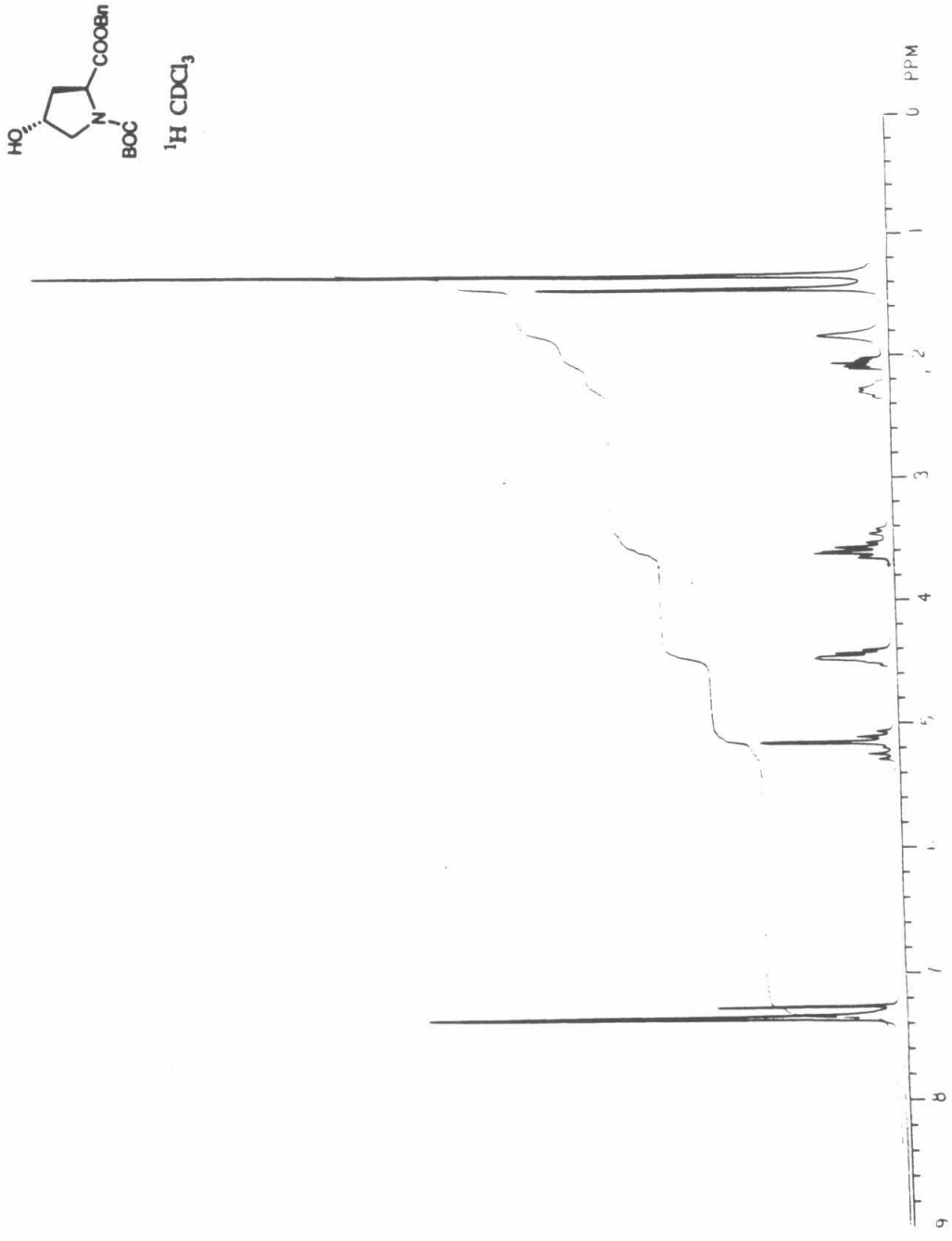
(8) Adachi, T.; Yamada, K.; Inoue, I. "An Alternative Method for the Selective Reduction of Unsaturated Nucleoside Azides to Amines," *Synthesis* **1977**, 45-46.

(9) Bodanszky, M. *Peptide Chemistry*; Springer-Verlag: New York, **1988**.

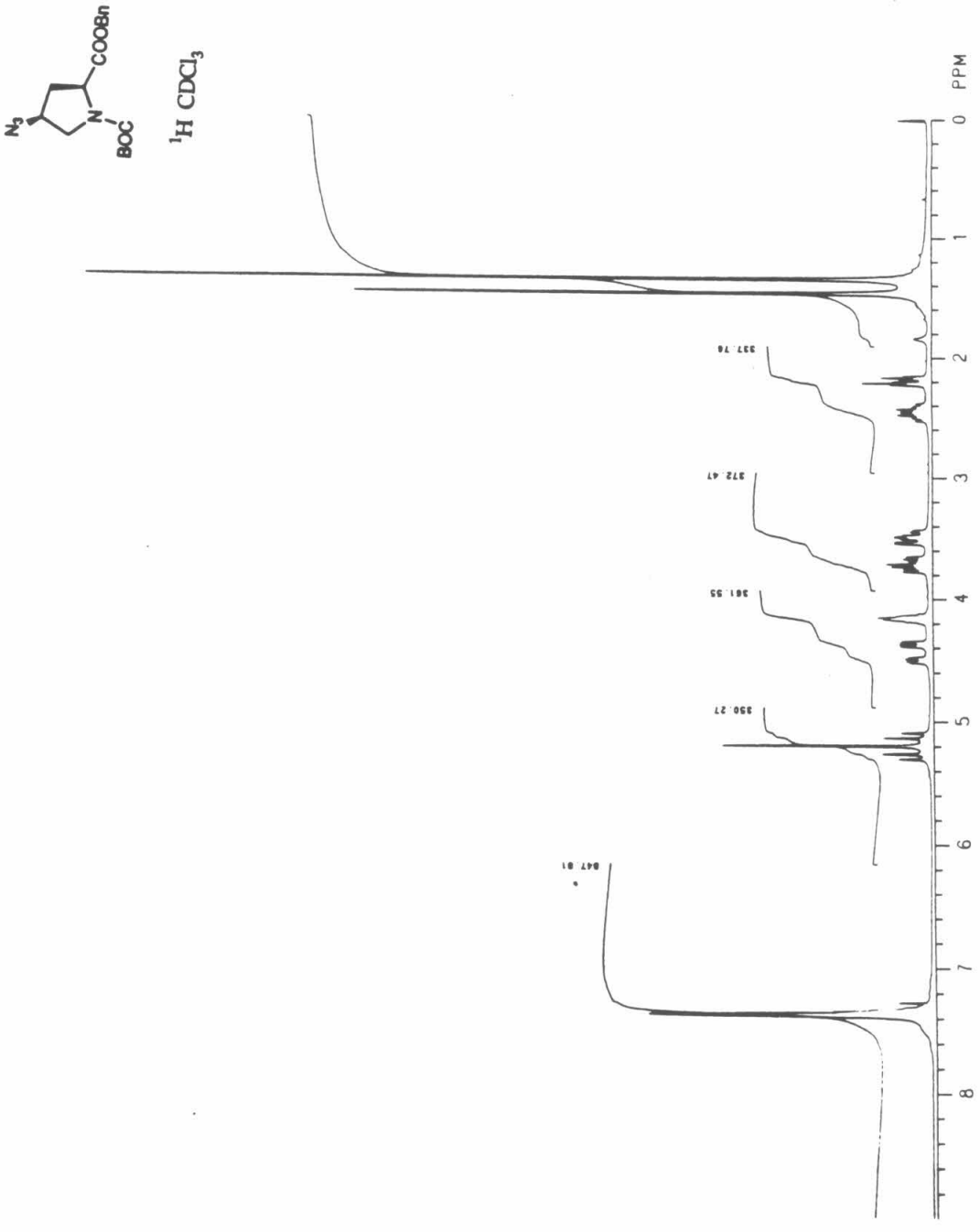
(10) deL. Milton, R.C.; Becker, E.; Milton, S.C.F.; Baxter, J.E.J.; Elsworth, J.F. "Improved purities for Fmoc-amino acids from FMOC-ONSu," *Int. J. Pept. Protein Res.* **1987**, *30*, 431-432.

(11) Webb, T.R.; Eigenbrot, C. "Conformationally Restricted Arginine Analogues," *J. Org. Chem.* **1991**, *56*, 3009-3016.

(12) Feiser, L.F.; Fieser, M. *Reagents For Organic Synthesis*; John Wiley and Sons: New York, **1967**; p 446.

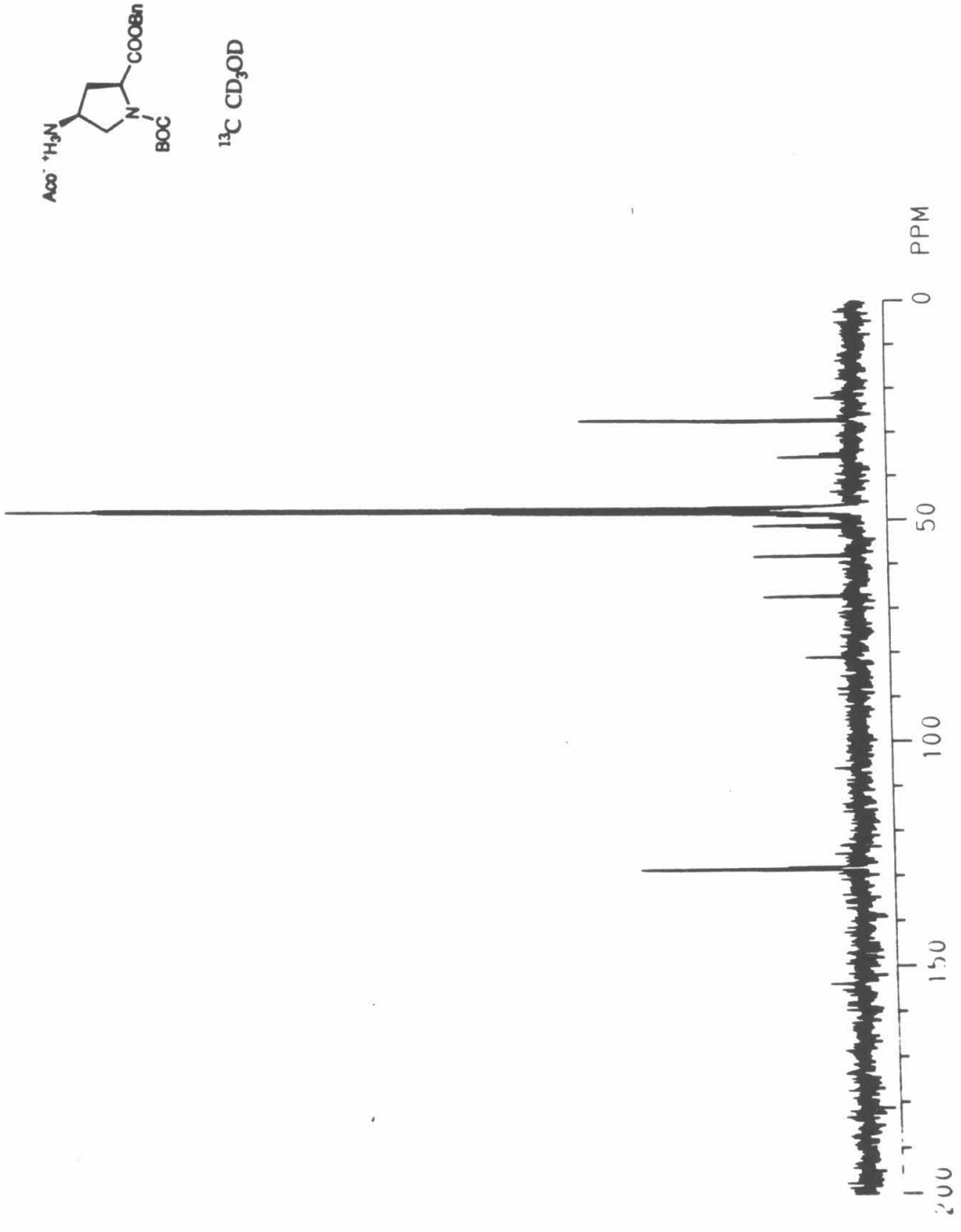

HO
COOBn
N
BOC
^1H $CDCl_3$
PPM
0
1
2
3
4
5
7
8
9

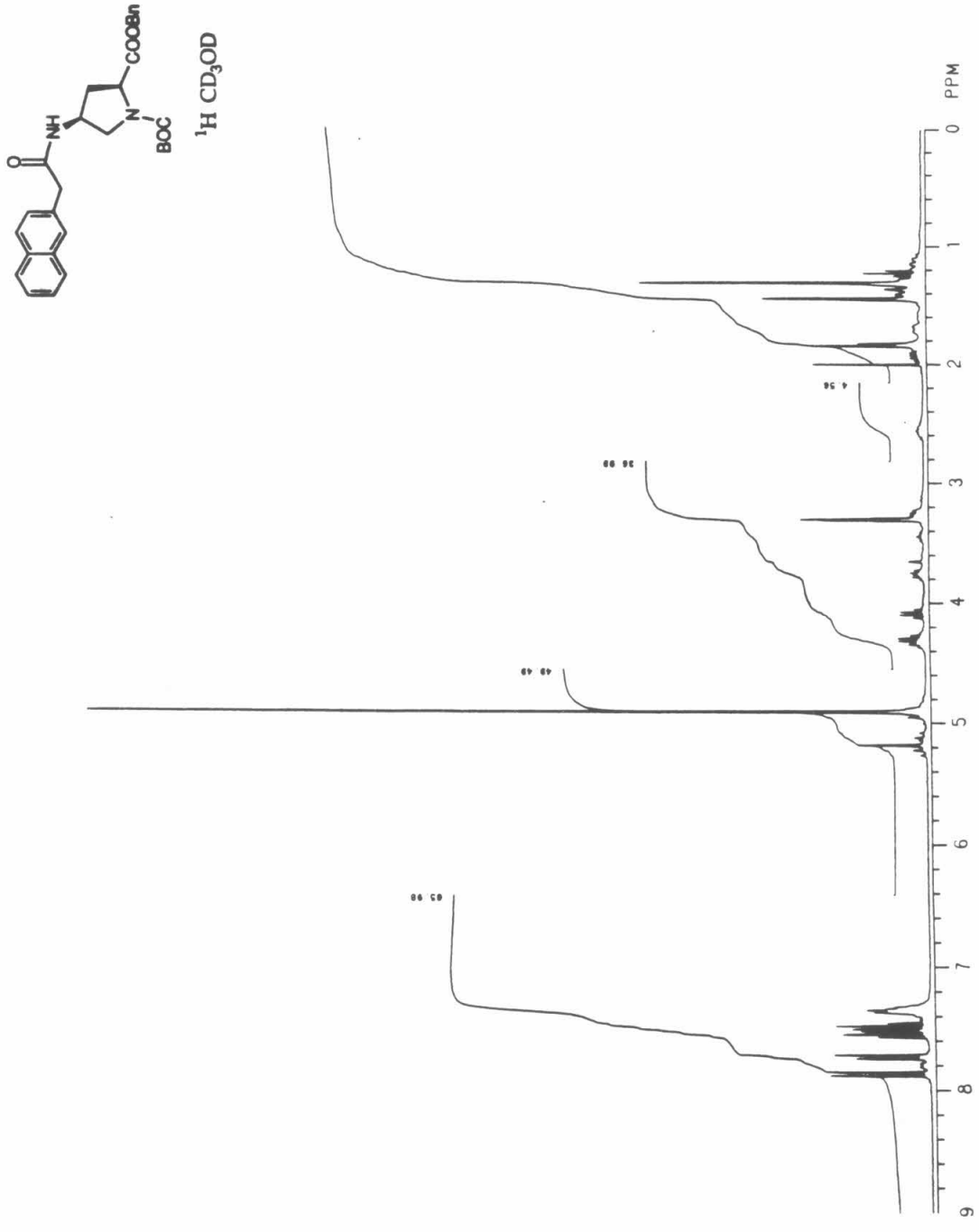
HO
COOBn
N
BOC
^{13}C $CDCl_3$
PPM

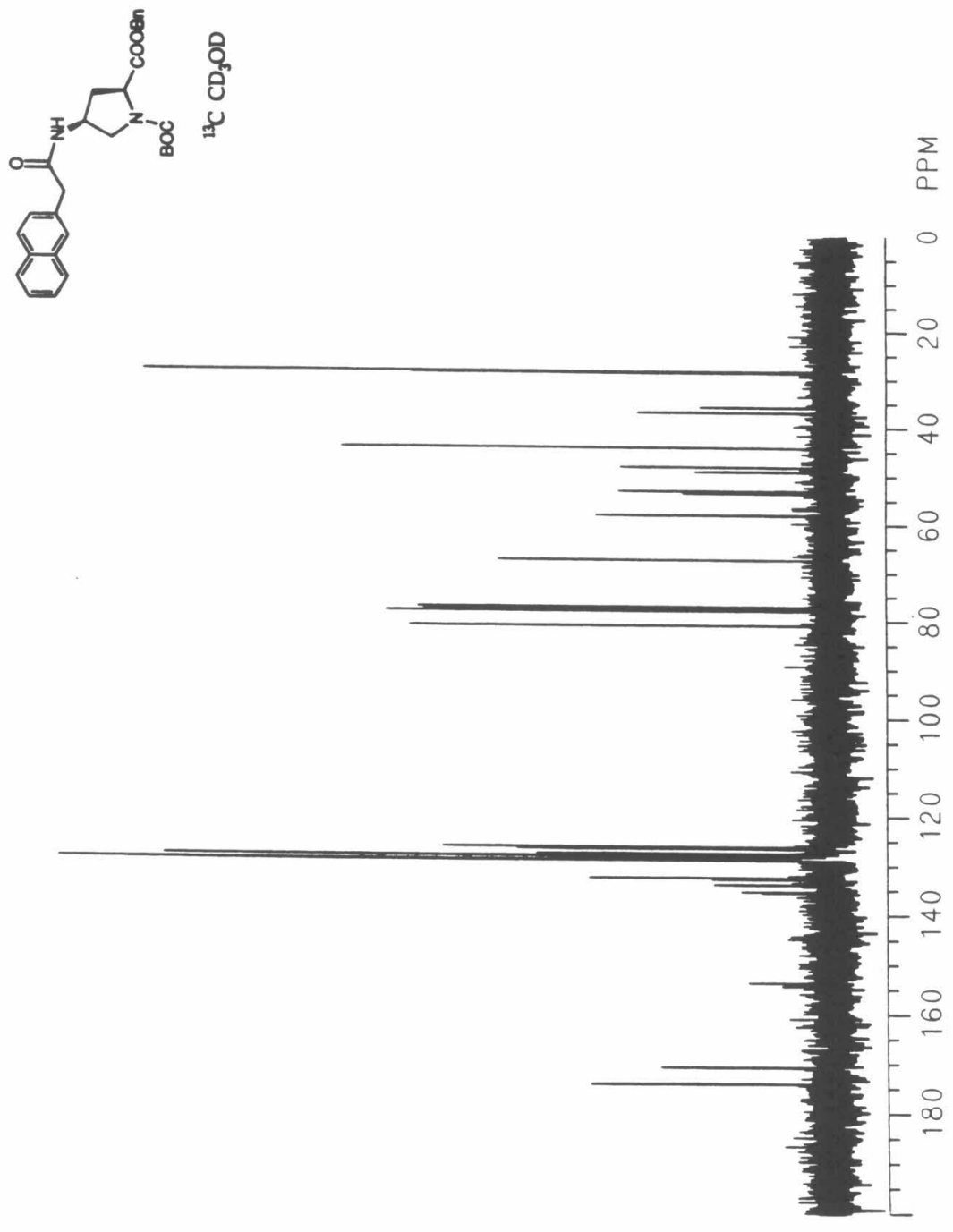

N3
COOBn
N
BOC
1H CDCl3
847.81
350.27
361.55
372.47
337.76
PPM
0
1
2
3
4
5
6
7
8

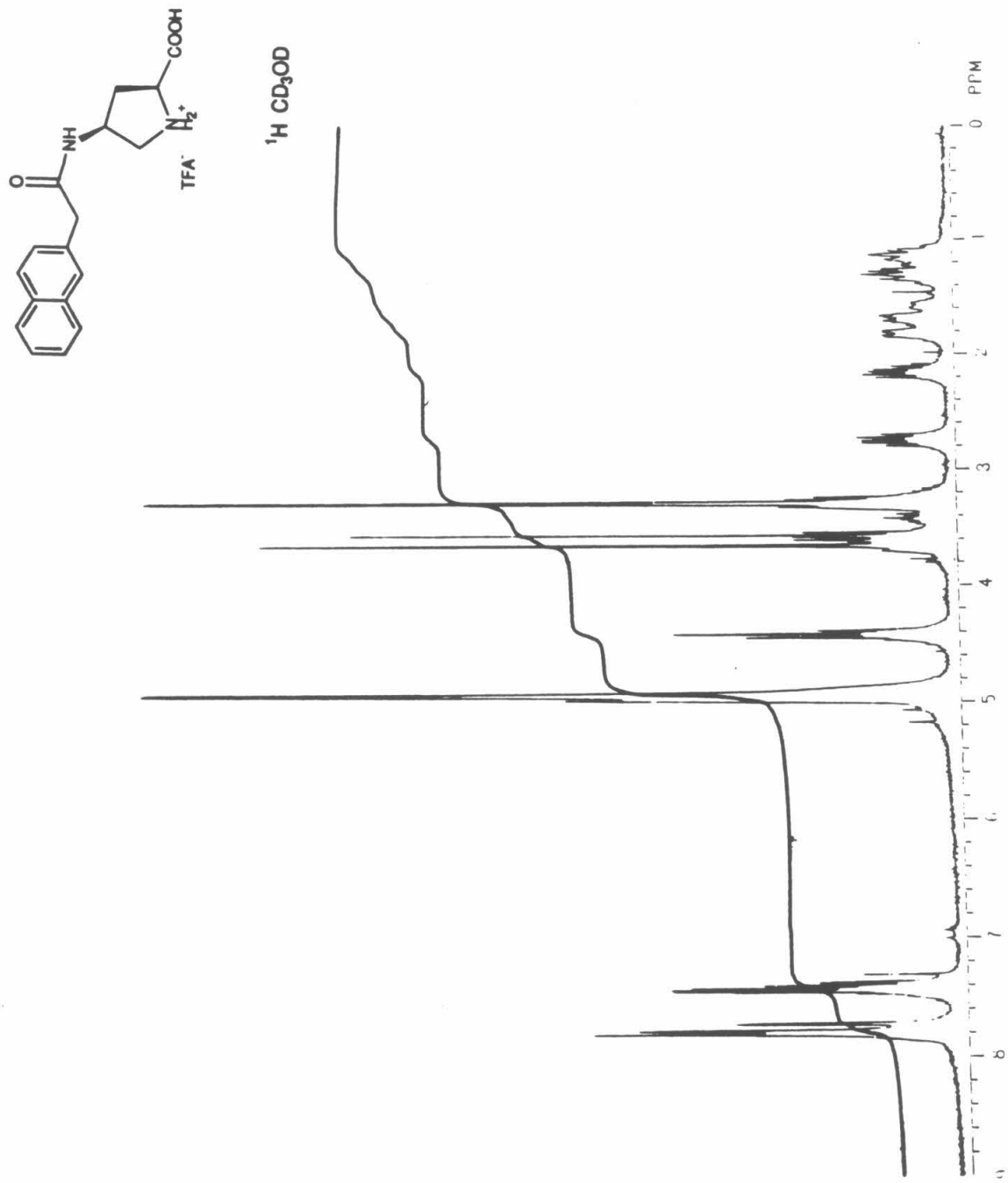
^{13}C $CDCl_3$

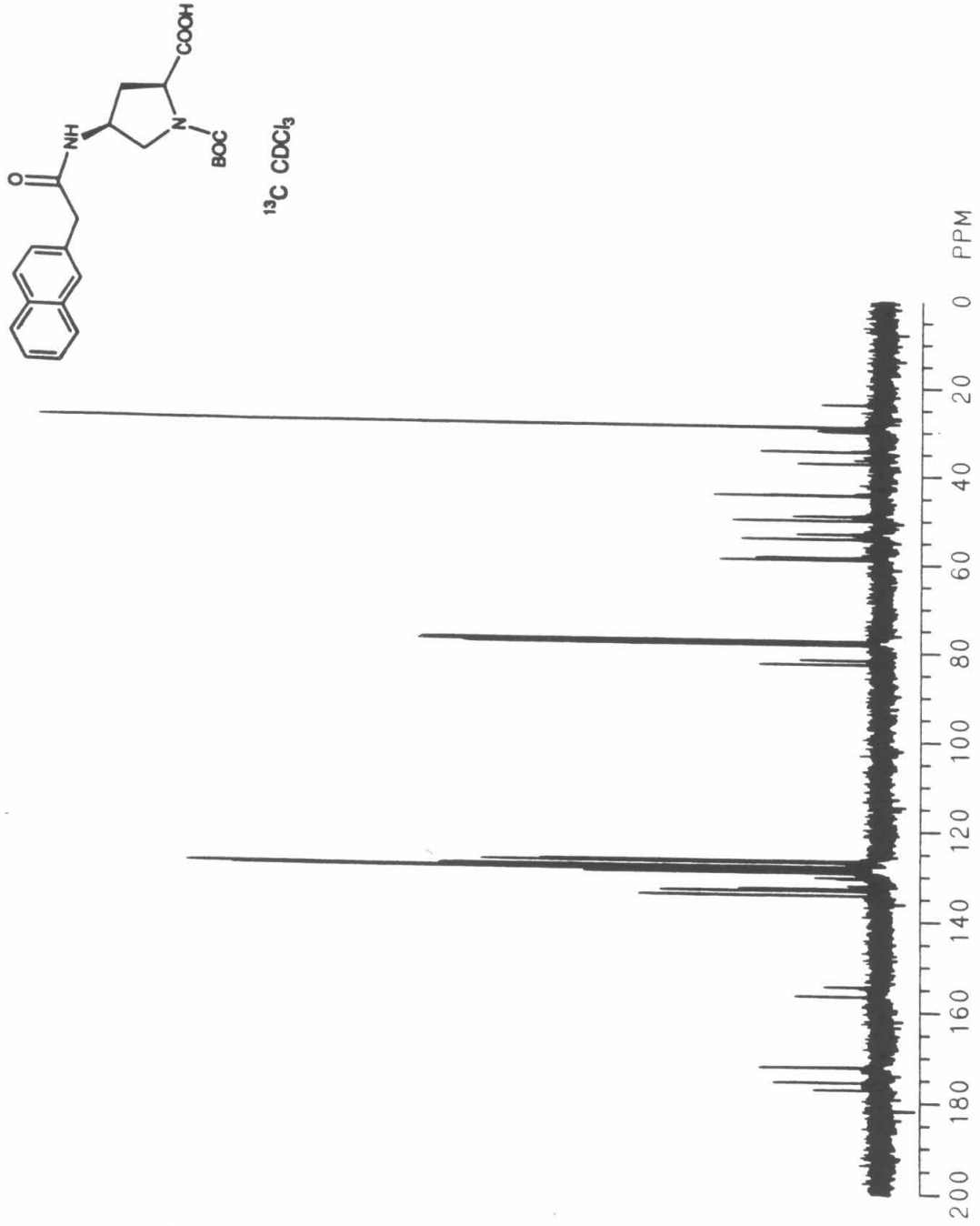
^1H CD_3OD

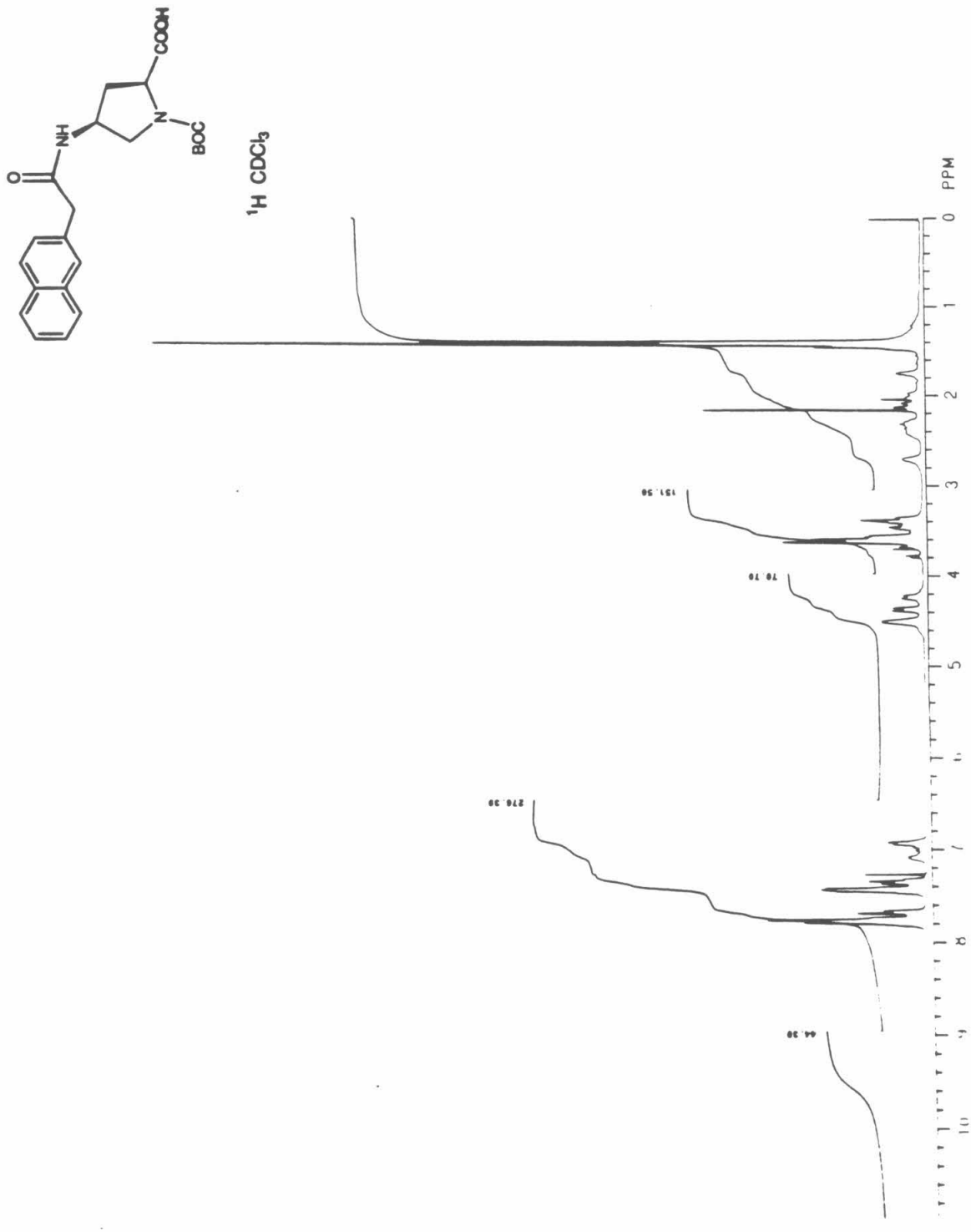

AcO⁻ ⁺H₃N
COOBn
N
BOC
^{13}C CD_3OD
0
50
100
150
200
PPM

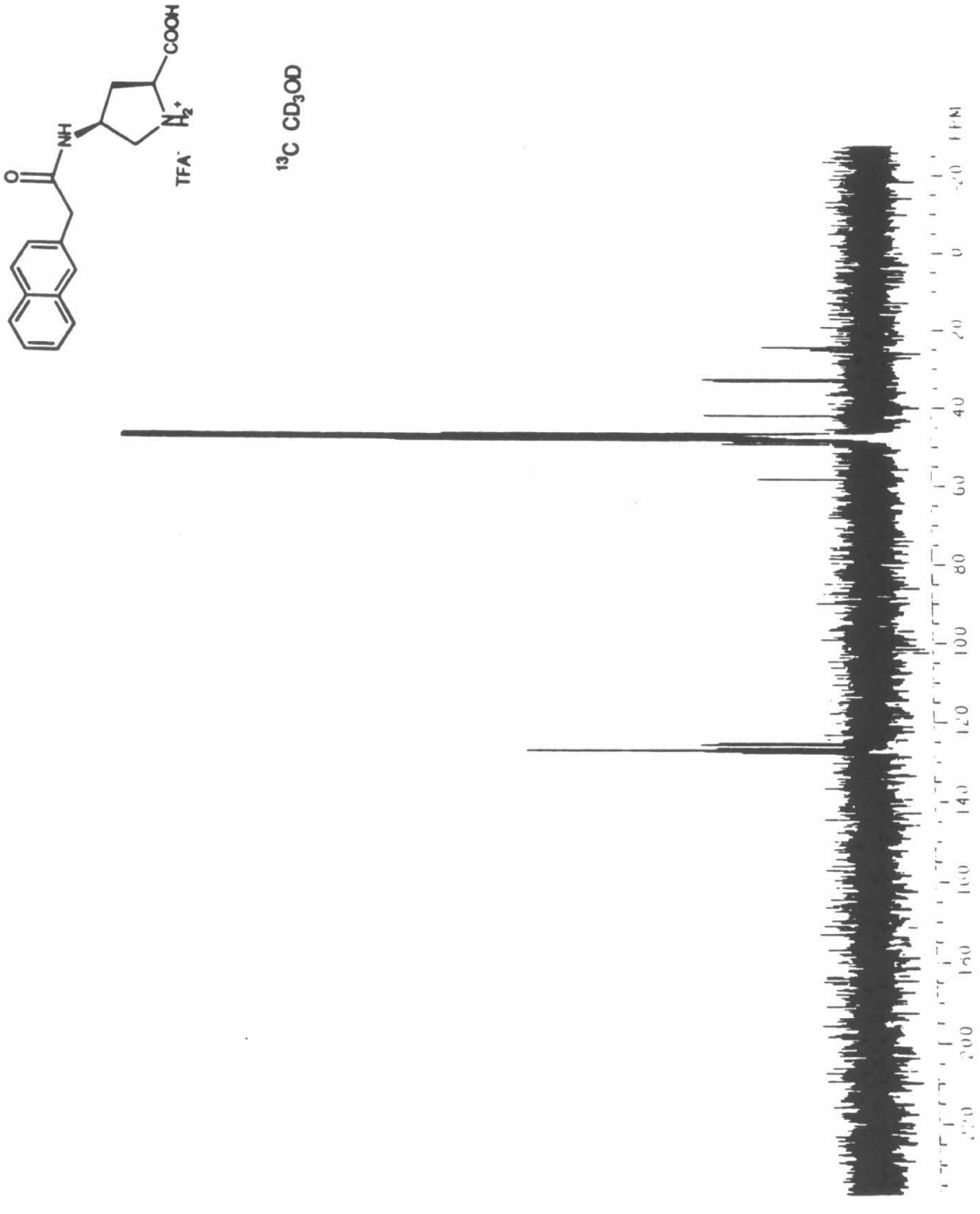

COOBn
NH
BOC
^{1}H CD$_3$OD
PPM
0
1
2
3
4
5
6
7
8
9

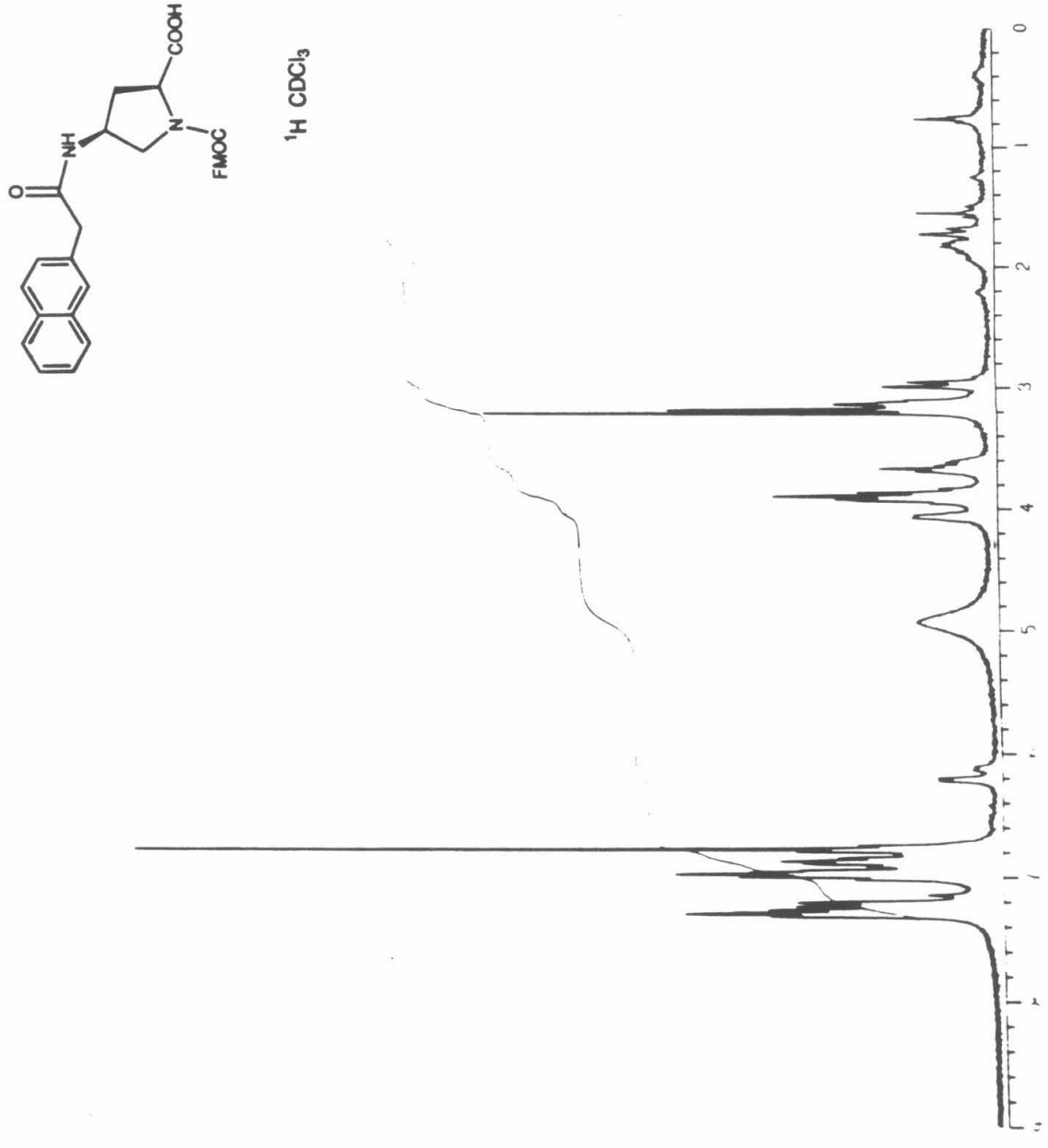

COOBn
NH
BOC
^{13}C CD_3OD
180
160
140
120
100
80
60
40
20
0
PPM

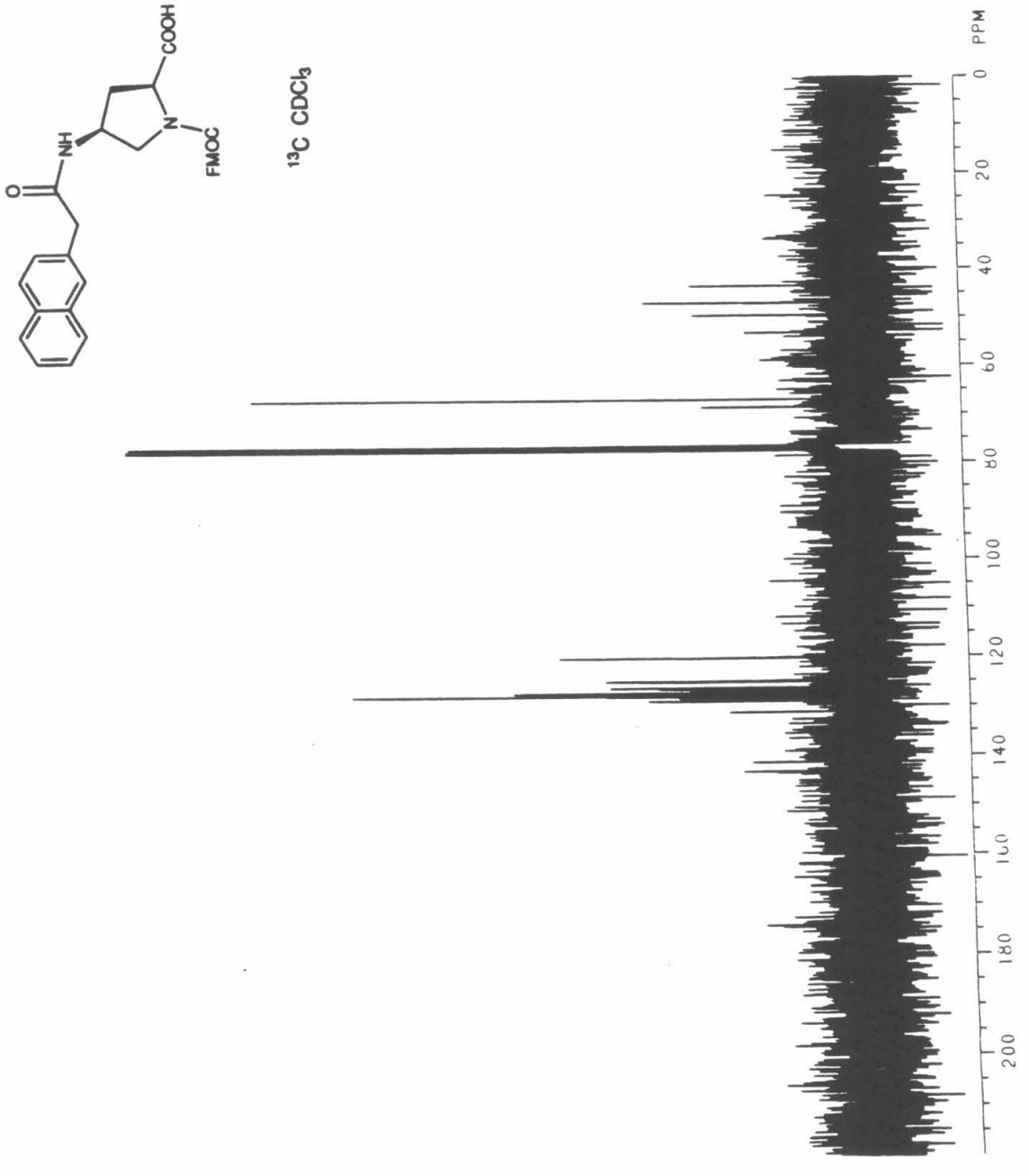

COOH
NH
O
TFA⁻
^{1}H CD$_3$OD
PPM


COOH
NH
N
BOC
O
^{13}C $CDCl_3$
200
180
160
140
120
100
80
60
40
20
0
PPM


COOH
NH
N
BOC
O
^{1}H CDCl$_3$
PPM
0
1
2
3
4
5
6
7
8
9
10
276.30
151.50
70.70
64.30


COOH
NH
TFA
^{13}C CD_3OD


COOH
NH
N
FMOC
O
^1H $CDCl_3$
0 PPM


COOH
NH
N
FMOC
O
^{13}C $CDCl_3$
PPM
0
20
40
60
80
100
120
140
160
180
200