INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

SOLUTION-STATE CONFORMATIONAL STUDIES OF ENDOTHELIN ANALOGS

by

Gregory Mitchell Lee

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

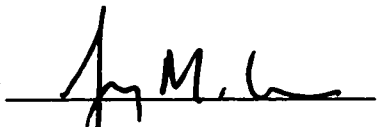
University of Washington

1999

Program Authorized to Offer Degree: Chemistry

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to UMI Dissertation Services, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date April 9, 1999

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

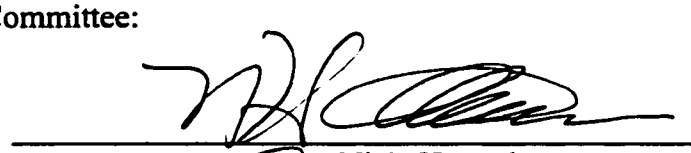
Gregory Mitchell Lee

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

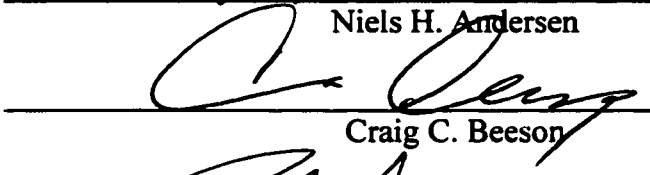
Chair of Supervisory Committee:

Niels H. Andersen

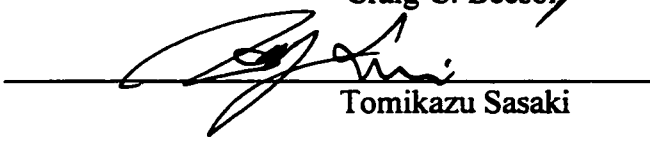
Reading Committee:



Niels H. Andersen



Craig C. Beeson



Tomikazu Sasaki

Date: Apr. 13, 1999

University of Washington

Abstract

SOLUTION-STATE CONFORMATIONAL STUDIES OF ENDOTHELIN ANALOGS

by Gregory Mitchell Lee

Chairperson of the Supervisory Committee:
Professor Niels H. Andersen
Department of Chemistry

Endothelin-1 (ET-1), a 21-residue peptide hormone containing two disulfide bridges, is the most potent endogenous mammalian vasoconstrictor known. Endothelin binds, with high specificity, to a G-protein coupled transmembrane receptor, designated ET_A. Although a consensus NMR structure of the bicyclic core has been reached, the conformation of the biologically important C-terminus remains a point of controversy. An ET-1 analog, [$Pen^{3,15}$-Nle^{7}]-ET-1, henceforth Pen-1, is a potent ET_A receptor *agonist*. An N-methyl scan of the C-terminal residues yields a potent ET_A receptor *antagonist*. This analog, known as Pen-2, retains the previous substitutions of Pen-1 but includes an $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ substitution.

NMR structure ensembles for the Pen-1 analog indicate that the bicyclic core (residues 1-15) is more rigid than that of ET-1, and thus is considered to be a better model of the bioactive state. Increased rigidity of the bicyclic core also stabilizes and lengthens the helical region (from 9-16 in ET-1 to 9-18 in Pen-1). Abnormal methyl group chemical shifts observed for the ET-1 analogs are attributed to hydrophobic clustering between the sidechains of Ile^{19} and Trp^{21}. N-methylation of Ile^{20} appears to create a

backbone conformation that enhances this hydrophobic interaction, restricts the positional space of Trp^{21}, and prevents ET_A signaling.

A number of C-terminal ET-1 and Pen-2 analogs were designed and synthesized to (1) determine if the Cys residues are essential for the helix formation, (2) examine the role of the Trp^{21} sidechain on the shift of the upfield 19γ2 peak, and (3) probe the N-methyl effect on the backbone conformation in the absence of other secondary structure. NMR studies indicate the presence of a stable 7-residue helical segment in a linear Pen-2 C-terminal fragment. This segment represents one of the shortest stable helices (2 turns) observed in a peptide. CD studies of the N-methylated fragments reveals an unusually large negative band at 225 → 230nm which is absent in the non-N-methylated analogs. The resulting CD difference spectra indicate the presence of a unique turn-like signature that is thermally stable in the absence of other secondary structure.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

General:

DAG. 1,2-Diacylglycerol

DSS. 2,2-Dimethyl-2-silapentane-5-sulfonic acid

EDCF. Endothelium-derived contracting factor

EDRF. Endothelium-derived relaxing factor

ET-1. Endothelin -1

HOAc. Acetic acid

HFIP. 1,1,1,3,3,3-Hexafluoroisopropanol

IP_3. Inositol-1,4,5-triphosphate

NO. Nitric oxide

Pen-1. [$Pen^{3,15}$-Nle^{7}]-Endothelin-1

Pen-2. [$Pen^{3,15}$-Nle^{7}-$^{(NMe)}Ile^{20}$]-Endothelin-1

SRTX. Sarafotoxin

TFA. 2,2,2-Trifluoroacetic acid

TFE. 2,2,2-Trifluoroethanol

TMS. Tetramethylsilane

TNBS. Trinitrobenzene sulfonic acid

TNP-aa. Trinitrophenyl-amino acid conjugate

TSP. 3-(Trimethylsilyl) propionate sodium salt

VIC. Mouse vasoactive intestinal contractor

Peptide Synthesis:

DCC. Dicyclohexylcarbodiimide

DMF. Dimethylformamide

FMOC. 9-Fluorenylmethoxycarbonyl

HATU. N-[(dimethylamino)-1H-1,2,3-triazol-[4,5-b]-pyridiylmethylene]-N-methyl methanaminium hexafluorophosphate-N-oxide (Uronium salt of HOAt)

HBTU. N-[(1-H-benzotriazol-1-yl)(dimethylamine)methylene]-N-methylmethan aminium hexafluorophosphate-N-oxide. (Uronium salt of HOBt)

HOAt. 1-Hydroxy-7-azabenzotriazole

HOBt. 1-Hydroxy-benzotriazole

NMP. N-methyl-2-pyrrolidinone

PyBrOP. Bromo-tris-pyrollidino phosphonium hexafluorophosphate

Spectroscopy:

CD. Circular Dichroism

COSY. Correlated spectroscopy

CSD. Chemical shift deviation

NMR. Nuclear magnetic resonance

NOE. Nuclear Overhauser effect

NOESY. Nuclear Overhauser effect spectroscopy

ROESY. Rotating frame Overhauser effect spectroscopy

TOCSY. Total correlated spectroscopy or Homonuclear Hartmann-Hahn spectroscopy

Amino Acids (1 and 3 letter codes):

A (Ala). Alanine

C (Cys). Cysteine

D (Asp). Aspartic acid

E (Glu). Glutamic acid

F (Phe). Phenylalanine

G (Gly). Glycine

H (His). Histidine

I (Ile). Isoleucine

K (Lys). Lysine

L (Leu). Leucine

M (Met). Methionine

N (Asn). Asparagine

P (Pro). Proline

Q (Gln). Glutamine

R (Arg). Arginine

S (Ser). Serine

T (Thr). Threonine

V (Val). Valine

W (Trp). Tryptophan

Y (Tyr) Tyrosine

Nle. Norleucine

Pen. Penicillamine

$^{(NMe)}$**Ile**. N^{α}-methyl isoleucine

ACKNOWLEDGMENTS

I wish to thank my advisor, Prof. Niels H. Andersen, for his guidance and support through the several years required for this project and during the preparation of this manuscript. Thanks are also given to my doctoral committee members, Professors Tomikazu Sasaki, Craig Beeson, Gary Drobny, Terry Lybrand and Heintz Floss for their contributions. In addition, acknowledgements are given to former and current members of the Andersen Group: Dr. Chinpan Chen, Dr. Scott Harris, Dr. Bolong Cao, Dr. Rob Palmer, Dr. John Cort, Charles Wandler, John Tomaszewski, Zhihong Liu, Hui Tong, Jonathan Neidigh, Matt Fesinmeyer and undergraduate assistant Yan Brodsky. I would also like to thank members of the Sasaki and Drobny groups are for the use of their instruments in the peptide syntheses, as well as Dr. Tom Pratum for his assistance in the NMR lab and Mr. Jim Roe for his help with the CD spectrometer.

For helping keep my sanity (when I didn't have broken fingers), I would like to acknowledge the Chemistry Department softball team(s). Finally, I wish to thank my family, specifically my parents Katherine and (the late) Wiley Lee, as well as my sister, Nicole Lee for the lifetime of love and support.

CHAPTER 1: INTRODUCTION

Macromolecules, specifically proteins and peptides, hold important roles in biological functions. Some peptides and proteins act as ligands and generally undergo a conformational change in order to effectively bind to receptor sites. Once bound to a specific site, the ligand may activate the receptor, initiating a biological process. In other cases, the ligand may bind to a different site on the same receptor, or a different receptor entirely, decreasing the signal producing the biological process. Occasionally an imbalance of these regulatory cycles occurs, resulting in a disease state. This is the stage that has garnered great interest in the pharmaceutical research communities. By understanding how the ligands adjust their conformation in order to activate or deactivate receptors, analogs that are either able to inhibit or help regulate these biological processes may be designed. This dissertation focuses on conformational studies one particular ligand, endothelin-1, and its analogs using multidimensional NMR, CD and molecular dynamics simulations.

1.1: PEPTIDE STRUCTURE

Proteins and peptides are amino acid polymers which can have four levels of structure: primary, secondary, tertiary and quaternary. Due to their limited number of amino acids and contact points, peptides have only primary and secondary structure. Proteins may contain all four subtypes. As a result, it is more difficult to establish a dominant conformation in peptides since they tend to be less structured and have more conformational averaging than proteins.

1.1.1: PRIMARY STRUCTURE

Primary structure is defined as the linear amino acid sequence and specifies only the covalent amide bonds between each residue (Figure 1.1). The peptide bond is planar

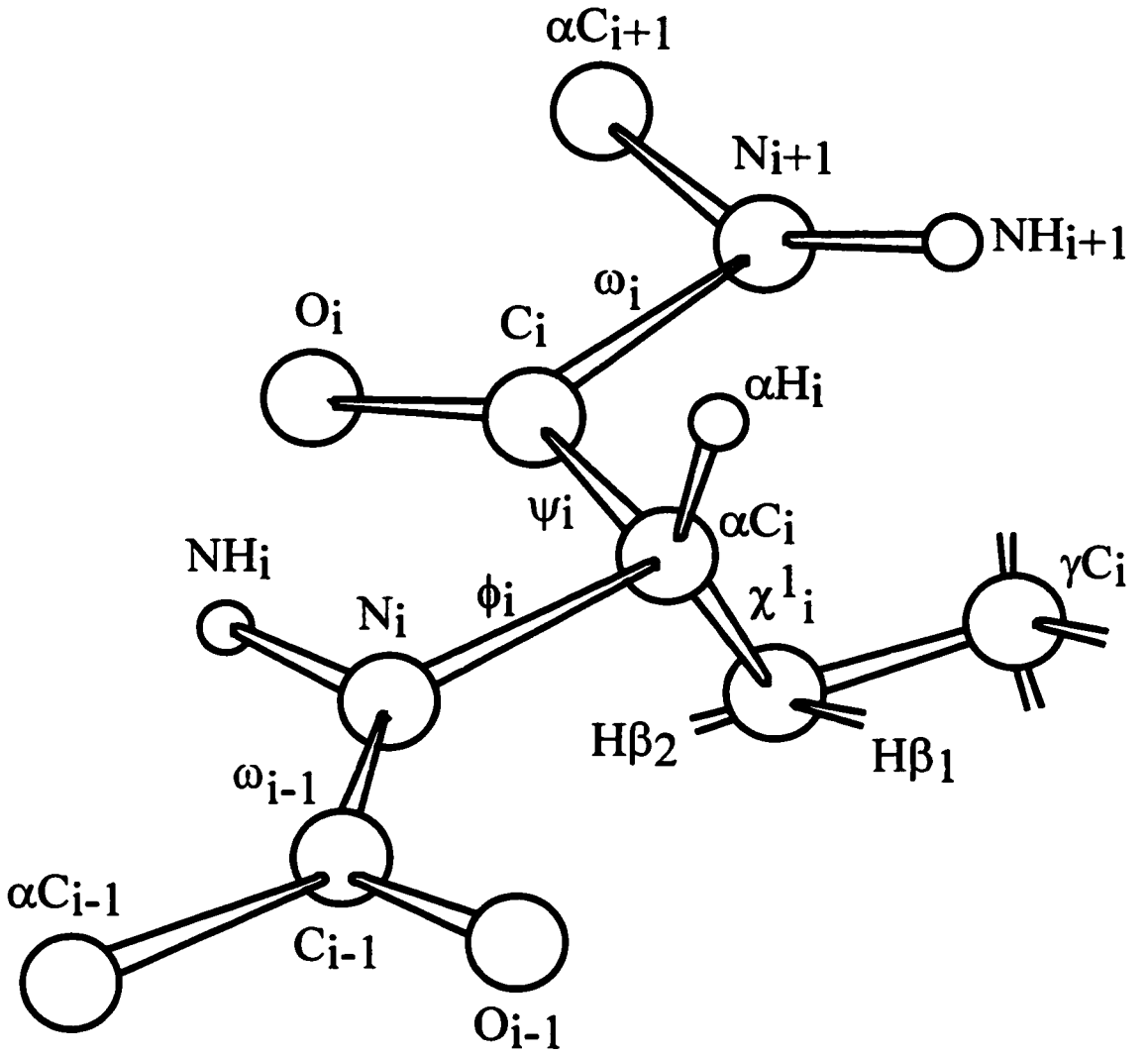


Figure 1.1: The peptide backbone.

and has some double bond character, as well as a dipole, due to resonance. The peptide unit is defined by six backbone atoms (N^{α}, amide H, Cα, αH, C' and O) with the side chain functional groups bonded to the Cα atom. The α-carbon is also the chiral center for most naturally occurring amino acids, usually of the L-configuration. As a result, the peptide backbone is inherently asymmetric.

By convention, the peptide sequence is listed from N-terminus to C-terminus with either a one or three letter code per residue. When describing interresidue relationships, amino acids *n* residues apart from each other in the sequence are often referred to as residues *i* and *i*+*n*. For example, an interaction between an atom of one amino acid and an atom of a neighboring amino acid three positions higher in the sequence are described as $x_{(i)} \rightarrow y_{(i+3)}$ (also denoted $x_{(i)}y_{(i+3)}$), where x and y correspond to specific atoms in the respective residues. For convenience, a peptide or protein containing *n* amino acids in its sequence is called an *n*-mer.

1.1.2: Secondary Structure Motifs

Secondary structure is the local folding of a short region of peptide backbone. The most commonly observed motifs are helices, β-strands (or extended conformations), turns, and β-hairpins (which combine the latter two features). Generally, short to intermediate range interresidue interactions such as hydrogen bonds or side chain salt bridges will help stabilize the conformation.

Three dihedrals define the conformation of the backbone (see Figure 1.1): ϕ to the $N_{(i)}$-$C\alpha_{(i)}$ bond, ψ to the $C\alpha_{(i)}$-$C'_{(i)}$ bond, and ω to the $C'_{(i)}$-$N_{(i+1)}$ bond. Sidechain conformations are defined by χ torsional angles. The ϕ and ψ angles are variable depending on the secondary structure formed; χ is dependent on the side chain packing. Due to steric factors, ω is fixed to either 180° (trans peptide bond) or 0° (cis peptide bond)

with the former favored by a 1000:1 ratio. This ratio reduces to 80:20 (trans:cis) when residue *i+1* is a proline. A "disordered" or "random coil" state usually applies to a backbone configuration that has no defined conformation. In this case, both ϕ and ψ dihedrals sample all regions of conformational space consistent with the steric requirements of the local sidechains.

Tables 1.1 and 1.2 (Creighton, 1984) list the typical backbone dihedrals for the major conformation types observed in protein and peptide structure. Dihedral angles of the various secondary structure motifs generally have rigid local restrictions, owing to the spatial arrangement of the backbone and side chain atoms. Ramachandran and co-workers (1968) used hard-sphere models and fixed geometries to map the allowed ϕ and ψ angles for various polypeptide structure types. These values are placed in a two-dimensional plane (ϕ vs. ψ), yielding the Ramachandran plot. A third dimension, corresponding to potential energy or interatomic distances may also be added to the plot. For most polypeptides that contain β-carbons, the allowed regions are rather narrow. In the absence of a β-carbon, such as the case for polyglycine, the permitted ϕ and ψ values are expanded dramatically.

Table 1.1: Secondary structure backbone dihedrals[1]

	ϕ (degrees)	ψ (degrees)	ω (degrees)
antiparallel β-sheet	-139	+135	-178
parallel β-sheet	-119	+113	+180
right-handed α-helix	-57	-47	+180
3_{10} helix	-49	-26	+180
poly(pro) I	-83	+158	0
poly(pro) II	-78	+149	+180

[1] Adapted from Creighton (1984), page 171.

1.1.2.1: α-helices

Right-handed α-helices, which contain 3.6 residues per turn and with a 1.5Å translation per residue, are stabilized by interresidue hydrogen bonds between the carbonyl oxygen of residue *i* and the amide hydrogen of residue *i+4*. These hydrogen bonding interactions stack head-to-tail and align with the axis of the helix, generating a net helix macrodipole (positively charged at the N-terminus, negative at the C-terminus). In addition, this geometry allows the side chains of residue *i* to be situated close together with those of residues *i+3* and *i+4*. Salt bridges or hydrophobic interactions between these residues will often help stabilize a helix. In the case of an amphipathic helix, charged residues will normally be aligned on one face of the helix and exposed to solvent. Hydrophobic residues tend to lie on the opposite side of the helix and in proteins are sometimes buried by a hydrophobic cluster formation.

1.1.2.2: 3_{10} helices

The 3_{10} helix is a more tightly coiled, longer version of the α-helix. There are three residues per turn extending to a length of 2.0Å per residue. Hydrogen bonding between the carbonyl oxygen of residue *i* and the amide hydrogen of residue *i+3* is the norm in a 3_{10} helix. Due to the tight backbone packing, 3_{10} helices are less stable than α-helices and tend to occur less frequently in peptide and protein structures (Thornton, 1992). In proteins, 3_{10} helices typically consist of 3 to 5 residues and appear at the termini of α-helices or as connections between β-strands in all-β proteins (Thornton, 1992, Millhauser, 1995). A survey of protein helices (Barlow and Thornton, 1988) had indicated that most 3_{10} helices averaged 3.3 residues in length, while α-helices were nearly four times longer (an average of 12.2 residues). In addition, α-helices are clearly favored to form if the peptide sequence contains more than six residues. Short peptides containing Aib (α-aminoisobutyric acid) residues favor formation of 3_{10} helices (Millhauser, 1995).

Millhauser (1995) has also suggested that the 3_{10} helices may act as an intermediate along the folding pathway of a coil to α-helix transition. Further studies from this group (Millhauser, 1995; Fiori and Millhauser, 1995; Millhauser et al., 1997) have focused on Ala-rich peptides containing nitroxide spin-labels spaced either three or four residues apart in the sequence. Taking advantage of through-space electron-electron coupling, ESR spectroscopy was used to probe whether the peptides favored the 3_{10} or α-helix. In the case of the 3_{10} helix, the d($i \rightarrow i+3$) is much shorter than d($i \rightarrow i+4$), while for the α-helix, both distances are nearly equivalent. Stronger coupling, indicative of shorter distances, are manifested in broader ESR peaks. The results from these studies demonstrated that the 3_{10} helix $\rightarrow$ α-helix transition is length dependent: a 16mer containing three repeats of the AAAAK fragment showed stronger $i \rightarrow i+3$ interactions, while a 21mer composed of 4 fragment repeats had equivalent $i \rightarrow i+3$ and $i \rightarrow i+4$ interactions.

1.1.2.3: β-strands

The β-strand, also known as extended, conformations aren't planar but have a slight right-handed twist to the backbone. Each residue in the extended conformation translates to 3.2 - 3.4Å. Side chains of adjacent residues lie on alternating sides of the strand. Because the side chains tend to line up well with each other, and conditions for intrastrand hydrogen bonding between backbone amide hydrogens and carbonyl oxygens are favored, aggregation may occur in the form of either parallel or antiparallel β-sheets. Parallel β-sheets are structures in which one or more β-strands align head-to-tail sequentially. In most parallel β-sheets, other long fragments of secondary structure, such as a helix, separate the strands. Antiparallel β-sheets consist of one or more β-strands which are separated by loops or small fragments, such as reverse turns. Here the strands align in the opposite fashion of the parallel β-sheets. That is, the N-terminus of one

strand will line up with the C-terminus of its partner and vice versa. Both forms are more commonly observed in proteins than in peptides. Individual strands of antiparallel β-sheets contain alternating hydrophilic and hydrophobic residues in the sequence. This would allow the former to be exposed to solvent while the later are buried in the hydrophobic core of the protein. Parallel β-sheets are commonly observed in the solvent-inaccessible interior of the protein. These strands are more likely to contain hydrophobic residues throughout the sequence. With the exception of β-hairpins, β-sheet formation for peptides is an intermolecular association that yields aggregates. This process often limits the solubility of peptides.

1.1.2.4: Reverse turns and β-hairpins.

Three major types of β-turns (also known as reverse turns) have been identified in proteins and peptides and are designated types I, II and III. The type I β-turn is slightly more open than its type II counterpart. The type III turn is the first turn of a 3_{10} helix. All of these forms are stabilized by a hydrogen bond between the carbonyl oxygen of residue *i* and the amide hydrogen of residue *i+3*. Although the β-turns consist of four residues, only the backbone dihedrals of the inner two are necessary to define the turn. The residue at position 3 is important for both type I and type II turns. A type I β-turn cannot contain a proline at residue 3, while a type II β-turn must have a glycine at position 3. All amino acids are allowed at position 3 in a type III β-turn, however. The type I′, II′ and III′ β-turns are analogous to their non-primed counterparts but have dihedrals of opposite sign. Table 1.2 lists the typical backbone dihedral values of the major β-turns. Other turn types, designated IV → VII, have also been defined but will not be discussed here.

Beta-hairpins are a simple form of antiparallel β-sheets which help stabilize β-turns. These structures are stabilized by intrastrand sidechain-sidechain interactions as

well as main chain (backbone) hydrogen bonding (Blanco et al., 1998). The β-hairpin class of secondary structure has gain attention as *de novo* models for peptide and protein

Table 1.2: Typical backbone dihedrals for β-turns[1]

	ϕ_2 (degrees)	ψ_2 (degrees)	ϕ_3 (degrees)	ψ_3 (degrees)
Type I	-60	-30	-90	0
Type II	-60	+120	+80	0
Type III	-60	-30	-60	-30

[1] Adapted from Creighton (1984), page 237.

folding (for example: deAlba et al., 1996; Ramiréz-Alvarado et al., 1996; Ramiréz-Alvarado et al., 1997). Many of the simplest β-hairpin models consist of a minimum of 10 residues. However, NMR studies have indicated that these short β-hairpins are in rapid equilibrium with a disordered conformation (Ramiréz-Alvarado et al., 1997). One recent publication (Kortemme et al., 1998) reports the structure of a 20 residue polypeptide which folds into a stable three-stranded β-sheet protein (two serial β-hairpins) in aqueous media.

The conformation most commonly found in proteins utilizes either a type β_I' or a type β_{II}' turn and is known as a 2:2 β-hairpin. Other major β-hairpin conformations are the 3:5 type, which includes a β_I turn and a bulge, and the 4:4 type, which consists of a β_I turn. The classification system (Sibanda et al., 1989; Sibanda and Thornton, 1991) used to describe the β-hairpins indicates the length of the β-strand segments and the extent of interstrand hydrogen bonding. The first number corresponds to the number of strand residues whose amide *or* carbonyl groups are involved in H-bonding. The second indicates the number of strand residues whose amide *and* carbonyl groups participate in the β-strand H-bonding network.

Another class of reverse turn is the γ-turn, which is tighter than its β counterpart, encompassing 3 residues in the sequence. A hydrogen bond between the carbonyl oxygen of residue *i* and the amide hydrogen of residue *i+2* stabilizes the conformation. As a result, the dihedral angles for residue *i+2* are the most critical for defining the turn. Average values for a series of tripeptides that adopt a γ-turn conformation (Ala_3 and Gly_3) are ϕ_2 = +69° and ψ_2 = -61° (Némethy and Printz, 1972). The inverse γ-turn contains dihedrals with opposite sign.

1.1.2.5: Poly(Pro) conformations

Proline has long been known to disrupt helices and, to a lesser extent, β-strands, primarily due to its lack of an amide proton and constraints to its backbone torsional angles. Proline polymers, however, can form two extended helical structures. These conformations are most often observed in proteins and peptides which have sequential proline residues. The first, known as $poly(Pro)_I$, consists entirely of cis peptide bonds and forms a right-handed helix (3.33 residues per turn, 1.9Å translation per residue). This conformation is generally observed in nonpolar media. The second, $poly(Pro)_{II}$, is the more frequently occurring poly(Pro) conformer and consists entirely of trans peptide bonds, forming a left-handed helix (3.0 residues per turn). This conformation is nearly as extended as a β-strand (3.12Å *vs.* 3.4Å). The $poly(Pro)_{II}$ conformation is primarily observed in aqueous media. The typical dihedrals observed for both $poly(Pro)_I$ (ϕ, -83°, ψ, +158°) and $poly(Pro)_{II}$ (ϕ, -78°, ψ, +149°) conformations indicate that these structures are local minima in the β-region of the Ramachandran plot (Creighton, 1984, Woody, 1992).

1.1.3: TERTIARY AND QUATERNARY STRUCTURE

Tertiary structure is distinguished from secondary structure in that the former describes the global folding in a protein while the latter describes local interactions which

make up protein or peptide conformations. Tertiary structure usually consists of secondary structure subunits that interact with each other to form the overall protein fold. The interior of these folding motifs is hydrophobic, while the solvent-exposed protein exterior contain the polar residues. Long range interresidue hydrogen bonds, hydrophobic interactions, or side chain salt and/or disulfide bridges between two residues far apart in the sequence stabilize tertiary structure. For example, bovine pancreatic trypsin inhibitor (BPTI), a small 58 residue protein, has three long range disulfide bridges (Cys^{5}-Cys^{55}, Cys^{14}-Cys^{38} and Cys^{30}-Cys^{51}) which help constrain the overall conformation (Creighton, 1975). Even in the absence of covalent linkages, hydrophobic clustering can yield stable tertiary folds for domains as small as 35-40 residues.

Quaternary structure involves non-covalent interactions between two or more subunits (or domains) of tertiary structure, which may be homologous. Hemoglobin is a protein that has quaternary structure, consisting of four separate domains. These domains are separated into two homologous dimers, designated $\alpha^1\beta^1$ and $\alpha^2\beta^2$, which are arranged as a rough tetrahedron known as a heterologous tetramer. In this model, there are no direct interactions between the $\alpha 1$ and $\alpha 2$ domains or the $\beta 1$ and $\beta 2$ domains. The $\alpha 1$ domain interacts solely with $\beta 1$ and $\beta 2$, primarily through hydrophobic contacts between each subunit. The $\alpha 2$ subunit has symmetrical contacts with the $\beta 2$ and $\beta 1$ chains.

1.1.4: Fluoroalcohol Effects on Peptide Folding and Stability

Fluoroalcohol (FA) addition is known to stabilize peptide structure, specifically helices, in aqueous media (see Harris, 1993 and Cort, 1997 for a brief historic overview). The most commonly used FAs in structure studies are 2,2,2-trifluoroethanol (TFE) and 1,1,1,3,3,3-hexafluoroisopropanol (HFIP). Several direct and indirect interaction models have been proposed to explain the helix-promoting effects of the fluorinated alcohols. The former generally involves a shift in the helix/coil equilibrium and reflects differences

in FA binding to the two states (Jasanoff and Fersht, 1994; Bodkin and Goodfellow, 1995). The latter models are indicative of medium effects on the peptide. Since addition of FA reduces the ability of the media to accept hydrogen bonds, the random coil state is destabilized (Llinás and Klein, 1975; Storrs et al., 1992; Cammers-Goodwin et al., 1996). A second indirect mechanism proposes the formation of micelle-like clusters of FAs upon addition to aqueous media (Kuprin et al., 1995; Hirota et al., 1997). A more recent article states that FA addition reduces the entropic cost of assembling a solvent shell around the solvated helix (Walgers et al., 1998). Previous reports from the Andersen group (Andersen et al., 1996) are consistent with various aspects of the indirect mechanisms, suggesting that the intramolecular H-bonded state evident in helices are favored by decreased stability of the coil state and by a solvophobic effect.

1.2: USE OF NMR IN PEPTIDE STRUCTURE ELUCIDATION

Although nuclear magnetic resonance (NMR) suffers from low sensitivity with respect to other spectroscopic methods, it remains a powerful tool in peptide and protein structure elucidation. With the advent of more powerful magnets, improved electronics and sophisticated pulse sequences, larger, more complex proteins can be studied. Incorporation of ^{15}N and ^{13}C labels, whether specific residues or uniformly over the entire sequence, in proteins and larger peptides has become common practice. In such cases, heteronuclear and 3D (and sometimes 4D) spectra are utilized in order to combat spectral overlap problems encountered in homonuclear 2D NMR, thereby elucidating a more accurate structure. However, with small to medium-sized peptides, where degenerate chemical shifts are generally less problematic than in proteins, standard 2D homonuclear methods are sufficient.

1.2.1: HOMONUCLEAR 2-DIMENSIONAL NMR TECHNIQUES

A general strategy for studying small to medium sized peptides *via* solution-state NMR has long been established (Wüthrich, 1986). COSY (*co*rrelated *s*pectroscop*y*) (Aue et al., 1976; Braunchweiler and Ernst, 1983) and TOCSY (*to*tal *c*orrelated *s*pectroscop*y* or homonuclear Hartman-Hahn) (Braunchweiler and Ernst, 1983; Bax and Davis, 1985b) experiments are used to identify and assign chemical shifts of individual amino acids. Both experiments rely on strong scalar coupling in order to effectively undergo polarization transfer through bonds. Each amino acid residue in a peptide chain can be considered an individual spin system due to weak scalar coupling through amide bonds. Scalar coupling within each residue is strong and, in principle, one should be able to easily progress down the side chain.

The primary difference between the COSY and TOCSY experiments is that in the former, crosspeaks arise from coherence transfer between coupled spins, while the latter uses a spin lock which induces isotropic mixing, to transfer magnetization between multiple spins that are strongly coupled. Commonly used spin-lock sequences are the MLEV-17 (Levitt et al., 1982; Bax and Davis, 1985b) and WALTZ-17 (Shaka et al., 1983). As a result, COSY crosspeaks are usually limited to three bond coupling, while TOCSY crosspeaks, given a long enough spin lock time, can be generated along the entire spin system. For example, a COSY spectrum of a leucine residue would only show crosspeaks between NH $\rightarrow \alpha$H, αH $\rightarrow \beta$H's, βH's $\rightarrow \gamma$H and γH $\rightarrow \delta$H's. In the TOCSY, one should observe crosspeaks between the resonances of the amide NH and the rest of the protons in the side chain. The side chain protons (α, β, γ and δ) will have a network of crosspeaks amongst themselves as well.

NOESY (*n*uclear *O*verhauser *e*ffect *s*pectroscop*y*) (Jeener et al., 1979) experiments rely on strong dipolar couplings in order to transfer magnetization through space. NOE cross relaxation, or growth, rates (σ_{IS}) are inversely proportional to the sixth

root of the distance between two spins (I and S) that experience dipolar coupling, as well as the product of the Larmor frequency (ω_o) and the correlation time (τ_c), or tumbling rate of the molecule. Therefore, for any given molecule (where ω_o and τ_c are constant), σ_{IS} is primarily dependent on the r^{-6} (where r is the distance between spins I and S). Spins which are closer together in space will have faster initial growth rates, and, therefore, stronger NOE intensities. Thus, the NOESY experiment is sensitive to the distance between the dipolar coupled spins, having an effective range < 5Å, and can provide insights regarding the secondary and tertiary structure of a peptide or protein.

However, NOE enhancements (η_{IS}) between the two spins are not dependent only upon r^{-6}, the $\omega_o\tau_c$ factor needs to be considered. Large molecules, which tumble slowly and have a long correlation time with respect to the Larmor frequency ($\omega_o\tau_c >> 1$), will experience *negative* NOE enhancements. Small molecules, which tumble at a faster rate, will have a short correlation time with respect to the Larmor frequency ($\omega_o\tau_c << 1$), leading to a *positive* NOE enhancement. In a typical NOESY experiment, molecules where $\omega_o\tau_c >> 1$ will have strong crosspeaks of the same sign as the diagonal peaks. Molecules where $\omega_o\tau_c << 1$ will have weak crosspeaks of opposite sign to the diagonal peaks.

Intermediate-sized molecules or small molecules in viscous solution ($\omega_o\tau_c \approx 1$) may have NOE enhancements that crossover from positive to negative. For these cases where the NOE intensities are effectively zero, ROESY (*r*otating frame *O*verhauser *e*ffect *s*pectroscop*y*, originally called CAMELSPIN) (Bothner-By et al., 1984; Bax and Davis, 1985a) experiments may be used. Whereas NOEs evolve during the mixing time (τ_m) portion of the NOESY pulse sequence, a spin lock is used to observe ROEs. In the spin lock condition, the effective Larmor frequencies are dependent on the much smaller spin locking field strength. This results in positive enhancements for all values of $\omega_o\tau_c$.

1.2.2: Secondary Structure Characteristics in NMR

Three important pieces of information harvested from NMR data can lead to insights regarding peptide secondary structure: NOEs, deuterium/hydrogen exchange rates and NH-αH coupling constants. Since NOEs grow at a rate proportional to r^{-6} and have an effective range of < 5Å, NOESY and ROESY experiments contain diagnostic crosspeak patterns corresponding to specific secondary structure motifs. Backbone NH/D_2O exchange studies are also useful tools in determining secondary structure. Those backbone amides that are solvent-exposed or disordered should display short exchange half-lives (for Ala-Ala: 1.3min at pD 6.0 and T = 280K or ~20min at pD 3.6 and T = 302K). Backbone amide NHs that exhibit longer exchange half-lives are involved in interresidue hydrogen bonding. The $^3J_{NH\text{-}\alpha H}$ coupling constants (when measurable) can lead to some insights regarding the backbone conformation. The coupling constants are dependent upon the torsion angle defined by NH-N-Cα-αH (θ). For an L-amino acid, θ is converted to φ by the following relationship, $\phi = \theta + 60$. Table 1.3 lists average distances and coupling constants for the major structure motifs.

Intense $i \rightarrow i+3$ crosspeaks, specifically the $\alpha_i N_{i+3}$ and $\alpha_i \beta_{i+3}$ interactions, are observed for α-helices. A medium-sized $\alpha_i N_{i+4}$ crosspeak may also be observed. Additionally, α-helices will have medium-to-strong $N_i N_{i+1}$ and often detectable medium $N_i N_{i+2}$ interactions, although the latter are due to secondary NOEs in most NOESY experiments. Due to the backbone dihedrals, intraresidue ($\alpha_i N_i$) NOEs will be much stronger than their sequential interresidue ($\alpha_i N_{i+1}$) counterparts. For the more tightly wound 3_{10} helix, medium $i \rightarrow i+2$ interactions, such as $\alpha_i N_{i+2}$ and $N_i N_{i+2}$, are more prevalent. Although the $\alpha_i N_{i+3}$ crosspeak remains strong, the $\alpha_i \beta_{i+3}$ interactions are less intense and the $\alpha_i N_{i+4}$ is very weak or non-existent. The sequential interresidue NOE ($\alpha_i N_{i+1}$) in this case becomes slightly more intense, but are still smaller than the intraresidue $\alpha_i N_i$. Since α-helices contain intra-strand hydrogen bonds between CO_i

and NH_{i+4}, all but the first four residues of the helix will exhibit long half-lives in a D_2O exchange study. In the case of the 3_{10} helix, where the intra-strand hydrogen bond exists between CO_i and NH_{i+3}, all but the first three residues of the helix will have long half-lives. Small $^3J_{NH-\alpha H}$ couplings are also diagnostic of helices. For α-helices, a series of couplings *circa* 3.9Hz are present, while a 3_{10} helix would suggest couplings of 4.2Hz (Wüthrich, 1986).

Table 1.3: Average interatomic distances and coupling constants[1]

crosspeak	α-helix	3_{10} helix	β-strand [2]	β_I turn [3]	β_{II} turn [3]
$\alpha_i N_i$	2.6 - 2.7 Å	2.6 - 2.7 Å	2.7 - 2.8 Å		
$\alpha_i N_{i+1}$	3.5 Å	3.4 Å	P: 2.3 Å	2/3: 3.4 Å	2/3: 2.2 Å
			AP: 3.0 Å	3/4: 3.2 Å	3/4: 3.2 Å
$\alpha_i N_{i+2}$		3.8 Å		3.6 Å	3.3 Å
$\alpha_i N_{i+3}$	3.4 Å	3.3 Å		3.1 - 4.2 Å[4]	3.8 - 4.7 Å[4]
$\alpha_i N_{i+4}$	4.2 Å				
$N_i N_{i+1}$	2.8 Å	2.6 Å	P: 3.3 Å	2/3: 2.6 Å	2/3: 4.5 Å
			AP: 4.0 Å	3/4: 2.4 Å	3/4: 2.4 Å
$N_i N_{i+2}$	4.4 Å	4.1 Å		3.8 Å	4.3 Å
$\alpha_i \beta_{i+3}$	2.5 - 4.4 Å[5]	3.5 - 5.1 Å[5]			
H-bonds	$CO_i \rightarrow NH_{i+4}$	$CO_i \rightarrow NH_{i+3}$	alt. residues	$CO_i \rightarrow NH_{i+3}$	$CO_i \rightarrow NH_{i+3}$
$^3J_{NH-\alpha H}$	3.9 Hz	4.2Hz	P: 8.9Hz	2/3: 4Hz	2/3: 4Hz
			AP: 9.7Hz	3/4: 9Hz	3/4: 5Hz

[1] Adapted from Wüthrich (1986). [2] P = parallel β-sheet, AP = antiparallel β-sheet.
[3] 2/3 and 3/4 correspond to residues 2 → 4 in the β-turns.
[4] Dependent on ψ. [5] Dependent on χ^1.

The β- or extended conformation doesn't have a large number of intermediate range NOEs within its own strand. Here intraresidue ($\alpha_i N_i$) crosspeaks are much less intense than the sequential interresidue ($\alpha_i N_{i+1}$) interactions and the intraresidue $N_i N_{i+1}$ crosspeaks tend to be weak. The only medium-to-long range NOEs observed for this

conformation are between the two (or more) strands of a parallel or antiparallel sheet. In this case, the hydrogen bonding network would involve alternating residues. The solvent-exposed strands would display alternating short and long NH half-lives during the D_2O/H_2O exchange study. When observable, extended structures also have larger than average $^3J_{NH\text{-}\alpha H}$ couplings. For parallel β-strands, the couplings average 8.9Hz, while in the antiparallel β-strands, the coupling is near the maximal value, 9.7Hz.

Hydrogen exchange and $^3J_{NH\text{-}\alpha H}$ information are less reliable for determining the presence of reverse turns. Since the turns generally appear between two sets of structural motifs, β-strands for example, any distinct hydrogen bond pattern would be difficult to establish. In the case of the β_I turn, the coupling constants for the corner residues (2 and 3) are 4Hz and 9Hz, respectively. The same residues in a β_{II} turn exhibit constants of 4Hz and 5Hz. Conversely, NOE interactions can yield important information in locating turn regions. Since the dihedrals of the corner residues (2 and 3) define the turn, a number of NOEs should be present from residues $2 \rightarrow 4$. In the β_I case, strong crosspeaks are observed for the N_iN_{i+1} of residues 2/3 and 3/4. The sequential α_iN_{i+1} crosspeaks are generally of medium intensity, while a number of intermediate range NOEs, for example α_iN_{i+2}, α_iN_{i+3} and N_iN_{i+2}, have medium to weak intensities. For a β_{II} turn, the N_iN_{i+1} peak for residues 2/3 becomes very weak, while that of residues 3/4 remains strong. Unlike the β_I turn, the type II version shows a strong α_iN_{i+1} NOE between residues 2 and 3. Additionally, the intermediate range NOEs, with the exception of the α_iN_{i+2} peak, tend to be very weak at best. Peak intensities for both N_iN_{i+2} and α_iN_{i+3} correspond to distances > 4Å.

1.2.3: NOE Ratios

Large differences in NOE crosspeak intensities, especially between intra- and sequential interresidue αN interactions, are diagnostic for α-helices and β-strands. For

example, in an α-helix, the intraresidue (α_iN_i) crosspeak is always more intense than the corresponding interresidue (α_iN_{i+1}) crosspeak. For β-strands, the reverse is true. A plot of these NOE "ratios" *versus* the sequence should, thus, be highly informative. These ratios are represented as a difference in contour levels (Δncl) between the sequential interresidue and intraresidue αH crosspeaks. Although no direct scanning of the sequence is implied, these NOE ratios are essentially comparisons of the inter/intra residue crosspeak intensities. The "N-line" examines the amide NH of residue *i* and its crosspeaks to the preceding residue (*i-1*) and intraresidue (*i*) α-methine proton [($N_i\alpha_{i-1}$) - ($N_i\alpha_i$)]. The "α-line" focuses on the α-methine proton of residue *i* and its crosspeak to the succeeding residue (*i+1*) and intraresidue (*i*) amide NH [(α_iN_{i+1}) - (α_iN_i)]. For example, if the N-line and α-line of residue 2 were examined, then one would calculate the Δncl of ($N_2\alpha_1$ - $N_2\alpha_2$) and (α_2N_3 - α_2N_2), respectively. Assuming the contour levels in the NOESY or ROESY spectrum were set to a geometric multiplier of 1.4, then:

$$\Delta \text{ncl} \approx \log_{1.4}\left[\frac{\text{inter NOE}}{\text{intra NOE}}\right] \qquad \text{[Eqn. 1.1]}$$

In the case of the α-line, secondary structure is easily identifiable: an uninterrupted string of negative Δncl values indicates an α-helix while a string a positive Δncl values denotes a β-strand. Capping residues display abrupt changes in the sign of the Δncl. A portion of the sequence that is considered to be "100%" helical should have a NOE ratio of -4, while a "100%" β-strand segment will have a value of +4. A string of Δncl values of approximately +1.5 would indicate a disordered or random coil conformation. The N-line of residue *i* and the α-line of residue *i-1* should also have the same sign since they both depend on the same backbone dihedral angle, ψ_{i-1}. However, the N-line and α-line of the same residue may have opposite signs, depending on the ϕ torsional angle (Bradley et al., 1990; Andersen et al., 1992b). Care must be taken not to

misinterpret peak intensities and/or volumes. For example, protons which have chemical shifts near that of the solvent may experience a decrease in peak intensity due to solvent suppression techniques.

1.2.4: Chemical Shift Deviations

Another powerful, yet simple, tool used in the determination of local secondary structure is the chemical shift deviation (CSD), which is based on the chemical shift index (CSI) developed by Wishart et al. (1991, 1992). Both the CSD and CSI are NOE-independent, and rely mostly on backbone dihedrals and H-bond interactions. CSDs are, in general, more informative than the CSIs since the former are "actual" chemical shift differences while the latter are limited to three values (-1, 0 and +1). As a result, CSDs are a more sensitive probe of secondary structure, especially when studying a peptide or protein under various conditions. In this procedure, a residue's "disordered" or "random coil" reference values (Wishart et al., 1995; Merutka et al., 1995; Andersen et al., 1997) is subtracted from the observed chemical shift. (For the current reference values used in the Andersen group, see Appendix A of this dissertation.) Usually these methods are applied to the backbone amide NH or the α-methine proton, but $^{13}C\alpha$, $^{13}C'$ and backbone ^{15}N values can also be used, although the ^{15}N chemical shifts need nearest neighbor corrections (Wishart, 1995). Upfield chemical shift deviations are observed as negative values ($\Delta\delta < 0$) while downfield chemical shift deviations are observed as positive values ($\Delta\delta > 0$).

Since the observed chemical shifts of the α-methine are primarily dependent on anisotropy contributions from the preceding backbone carbonyl (Wishart et al., 1991), and, therefore, on the backbone dihedrals, patterns that correlate with the major secondary structure motifs are evident in the αH-CSD plot. A continuous string of four or more negative αH-CSDs would indicate the presence of an α-helix, with a value of -0.4ppm

considered "100%" helical. A β-strand would exhibit a continuous string of three or more positive αH-CSDs. A maximum value of +0.25ppm is expected for the outwardly facing (i.e. exposed to solvent) α-methines or for stand-alone β-strands. Those α-protons which are inwardly facing in a β-sheet have αH-CSD values approaching +0.8ppm. The reverse turns, although undefined, normally display alternating positive and negative values at the N- and C-termini of the turn. In small to medium-sized peptides, capping residues generally have αH-CSDs which approach zero. Although the observed αH chemical shift values are primarily determined by the backbone dihedrals, they are also susceptible to ring current effects from neighboring aromatic residues.

However, less susceptible to neighboring ring current effects are the backbone amide NHs. The hydrogen bonding strength experienced by the NH is more of a determinant in the observed chemical shift than secondary structure (Wishart et al., 1991). The shorter (and stronger) the H-bond, the more deshielded the amide, and the more positive the observed NH-CSD. In the case of α-helices, amide NHs at the N-terminus of the helix were discovered to have more positive NH-CSDs than those at the C-terminus of the helix. Wishart et al (1991) proposed that this trend is due to the helix macrodipole. The N-terminus has a net positive charge, which leads to deshielding effects, and the C-terminus has a net negative charge, which leads to shielding effects.

1.2.5: NH-CSD vs. NH-Temperature Gradients

A more recent tool used extensively in the Andersen group is the NH-CSD vs. NH temperature gradient plots (Andersen et al., 1997). These were initially used as to determine how the strength of hydrogen bonds experienced by the backbone NHs are related to their temperature gradients. It was widely assumed in the peptide/protein NMR community that those amide NHs which had shallow temperature gradients (usually greater than -4ppb/°C) were always sequestered in a hydrogen bond, and therefore

exchange protected. Those backbone NHs exposed to solvent displayed steeper gradients (-6.0 → -8.5ppb/°C). While, in general, this is true for most proteins and highly constrained peptides, this presumption can not be applied to conformationally averaged or partially folded peptides.

This plot is presented such that increasing (downfield) NH chemical shift deviations from random coil values lie on the left side of the x-axis and decreases (i.e. moves upfield) towards the right. In all cases, the NH-CSD would correspond to the values calculated for the lowest temperature studied. The NH temperature gradients are arranged such that the largest negative values (i.e. those which display the largest upfield shifts upon warming) are placed in the top portion of the y-axis.

The slope and correlation coefficient (R^2) calculated from linear regression analysis provides some useful information regarding the peptide or protein studied. The slope is a measure of the decrease in the population of a structured state as the temperature increases. The steeper the regression line, the more facile the thermal "melting" of the folded state. The R^2 value is a measure of the cooperativity in the folding (or unfolding) of the system in question. Assuming a structure ↔ disorder transition, a large R^2 ($\geq$ 0.75) would indicate high cooperativity in folding as the peptide or protein fragment is cooled. Those amides with large CSDs and "abnormal" temperature gradients (i.e. > 0ppb/°C or < -9ppb/°C) would imply that partial structuring of the system occurs at lower temperatures. As the temperature increases, so does randomization of the conformation.

1.2.6: Translating NMR Data to Molecular Models

Once interatomic distances have been estimated from NOE intensities, they are used as constraints in high temperature molecular dynamics simulations. This method involves defining a set of coordinates, a randomized linear structure for instance, then

applying a force field based on energy terms and NOEs and solving Newton's equations of motion (Scheek et al., 1989). Throughout this procedure, NOE-derived distance constraints are treated as pseudo-bonds between the two atoms and the energy penalty (k_{NOE}) for violating the pseudo-bond constraint can be varied. The high temperature studies allow the molecule to sample all regions of conformational space. After several rounds, the temperature is slowly reduced, a process known as simulated annealing. During this process and the final energy minimization, the molecule is allowed to fall into energy wells. Several refinement protocols are required for the molecule to move from a local to the global energy well, which presumably represents the most highly populated conformation of the peptide or protein. Also included in separate constraint sets are backbone torsional angles derived from coupling constants and, to a lesser extent, αH-CSD and NOE connectivity data. Torsional constraints, however, are generally given a lower weight during structure calculations than the NOE distance constraints.

The program used for the present study, X-PLOR (Brünger, 1992), is based upon a modified version of CHARMm (Brooks et al., 1983). The molecular dynamics portion of X-PLOR solves the following equation:

$$m_i\left(\frac{\partial^2 x_i}{\partial t^2}\right)(t) = -\nabla x_i E_{TOTAL} \qquad \text{[Eqn. 1.2]}$$

where m_i is the defined atomic mass and x_i is the x-coordinate of the atom. Other Cartesian coordinates for the atom (y_i, z_i) are processed in a similar manner. The energy term is based on holonomic constraints which define characteristics that are inherent in bonds, specifically amide bond planarity, bond angles and lengths, impropers, and van der Waals interactions. The van der Waals terms are adjusted from purely repulsive (r^{-12}) terms during the initial energy minimizations to zero during the high temperature dynamics runs. More realistic attractive (r^{-6})/repulsive (r^{-12}) interactions are enforced during the final annealing stages. This model is known as a Lennard-Jones potential.

Since NOESY and ROESY crosspeaks reflect population weighted averaging of the accessible conformers, NOE intensities and interactions may be misinterpreted when converted to distance constraints. In regions of significant conformational averaging, for instance small, non-crosslinked peptides, multiple attributions could be misassigned to the major conformer. Eventually this could lead to sets of conflicting distance constraints or sets which produce false energy minima. To correct for this, ambiguous constraints may be placed in a low-weighted set or eliminated *via* an iterative approach which concentrates on the final reported energies (dihedral, van der Waals, or impropers) or NOE violations. An ensemble of structures is usually required to generate a sufficient population for sampling. If the majority of the distance and dihedral constraints are correct, the ensemble should be highly converged in the conformationally defined portions of the peptide or protein.

1.3: USE OF CD IN PEPTIDE STRUCTURE ELUCIDATION

Circular dichroism (CD) spectropolarimetry takes advantage of the chiral nature of peptide backbones. CD consists of linearly polarized light which is divided into two distinct components of equal intensity, left- and right-handed polarized light, which revolve about the propagation line, completing one revolution per wavelength. What the spectropolarimeter measures is the difference in absorbance (ΔA) between the left-handed (A_L) and right-handed (A_R) light, which are directly proportional to their respective molar extinction coefficients (ε_L and ε_R). When the Beer-Lambert law is considered, this gives the relationship,

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R = \frac{(A_L - A_R)}{bc} \quad \text{[Eqn. 1.3]}$$

where b is the cell path length and c is the sample concentration. The ΔA is then converted into a measurement of ellipticity (θ) by the following,

$$\theta = 32.98\Delta A \qquad \text{[Eqn. 1.4]}$$

However, most instruments report optical constants in terms of molar ellipticity, $[\theta]$,

$$[\theta] = \frac{100\theta}{bc} = 3298\Delta\varepsilon \qquad \text{[Eqn. 1.5]}$$

whose units are $deg \cdot cm^2 \cdot dmol^{-1}$. In order to convert the constant to terms of mean residue ellipticity, as it is most frequently reported in the literature, $[\theta]$ is divided by the total number of backbone amide bonds in the peptide or protein (Woody, 1995).

1.3.1: Secondary Structure Characteristics

Most backbone amides of peptides and proteins are located next to a chiral center (the α-carbon, which in most cases is the L-conformation). As a result, this is the chromophore which is most influenced by backbone configurations and is most easily observed through CD. All amides have a $n \rightarrow \pi^*$ transition at $\approx$220nm ($\varepsilon_{max} \approx 100$) and a $\pi \rightarrow \pi^*$ at $\approx$195nm ($\varepsilon_{max} \approx 7000$). The actual intensities and curve shapes are dependent upon secondary structure. At best, CD spectroscopy can only provide an estimate of the peptide's net folding; no specific local structuring information can be deduced. CD is essentially a snapshot of the mixture of conformers present with the more populated conformers dominating the spectrum. The upper panel of Figure 1.2 shows four major structure motifs.

The α-helix curve (Chen, 1974; Harris, 1993) contains an intense positive band at 192nm ($\pi \rightarrow \pi_{\perp}^*$) and two minima of equal intensity at 207nm ($\pi \rightarrow \pi_{\parallel}^*$) and 222nm ($n \rightarrow \pi^*$). The β-sheet (Cort et al., 1994) contains a maximum at 195nm and a minimum

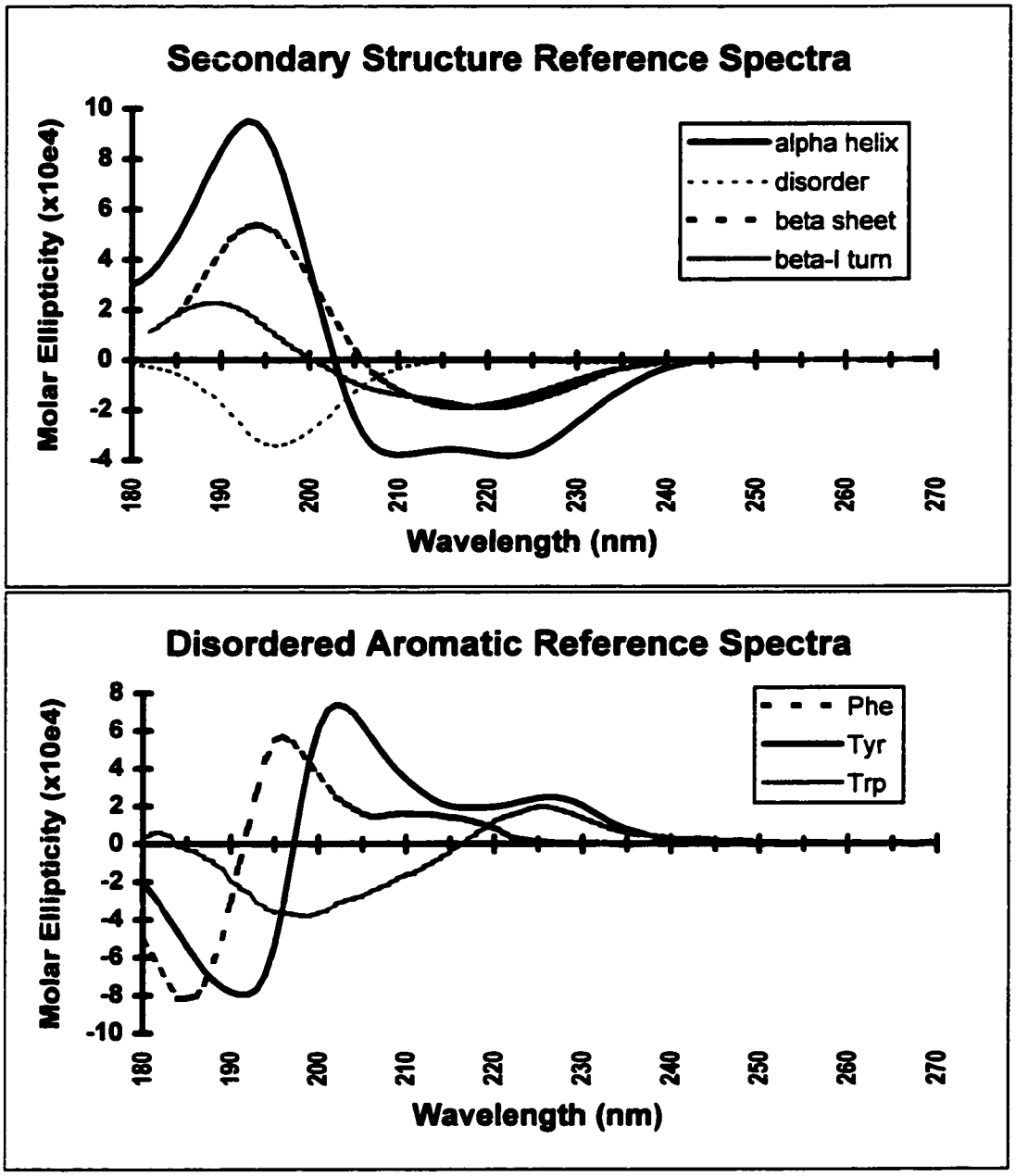


Figure 1.2: Reference CD spectra - (top) secondary structure, (bottom) disordered aromatic residues. Units are in molar ellipticity [(deg·cm^2·dmol^{-1})x10^4].

at 217nm. The intensities of the two bands have a ratio of approximately 2:1 (max:min). The β-turns generally resemble the β-sheet signatures, but are blue-shifted 5-10nm. The type-I β-turn (Brahms & Brahms, 1980) has a maximum at 190nm and a minimum at 217nm. Unlike the β-sheet, these two bands are of equal intensity. Type-II β-turns are believed to be red-shifted with respect to the β_I turns (Brahms & Brahms, 1980; Harris, 1993). The disordered, or random coil, curve contains a strong minimum at 197nm which is nearly as intense as the minima of the α-helix. A very weak positive band, which may be due to the poly(Pro)$_{II}$ conformation (Woody, 1995), is also present at 217nm.

Monitoring changes in backbone structure is straightforward using CD. Typically one can deduce significant changes by an initial survey of the CD spectra. For example, as conditions are varied, a set of spectra displaying an isodichroic near 203nm will usually indicate a two state equilibrium known as the helix-coil transition (Woody, 1995). Increases or decreases in a specific band with respect to changes in experimental conditions such as temperature or sample concentration are also measured. Addition of denaturants (guanidinium chloride, urea) or structure-stabilizing solvents (HFIP, TFE) are also used to affect the backbone conformation. In the case of the α-helix, the band at 220-222nm is the most commonly monitored feature. If an overall characterization of backbone conformational changes upon media or temperature change is required, difference (and sometimes double difference) CD spectra are calculated.

1.3.2: Aromatic Chromophores in CD Spectra

Problems in deducing secondary structure content from CD spectra can arise when the peptide contains aromatic residues. As in the case of the backbone amides, the aromatic side chains are linked to the nearby chiral α-carbon and, therefore, exhibit a CD signal. The bottom panel of Figure 1.2 displays the CD spectra for disordered aromatic

residue side chains. Both the Tyr and Trp curves (Brahms & Brahms, 1980; Harris, 1993; Andersen et al., 1995a) have a large positive band at 225 $\rightarrow$ 228nm. The Tyr spectrum also has a large negative band at 190nm and a large positive band at 200nm. The Trp signal contains a negative band at 198nm. The Phe spectrum (Brahms & Brahms, 1980) is blue shifted and slightly less intense with respect to the Tyr curve.

On a molar ellipticity (per aryl) scale, these signal intensities are larger than those of the secondary structure reference spectra. This may present complications in the spectral interpretation of the CD data if the peptide contains a significant number of aromatic residues in its sequence. Any aryl sidechain chromophores may interfere with the backbone signals. The primary influence on the signal intensities of these bands is the rotomeric preference of the aromatic sidechains. That is, the side chain torsion angle, specifically χ_1, will favor a particular conformation in an unstructured peptide (Alder et al., 1973; Kelly and Price, 1997).

1.4: ENDOTHELINS

Endothelin-1 (henceforth, ET-1, Table 1.4) is the most potent vasoconstrictive mammalian peptide hormone known. Consequently, the endothelin (ET) family of peptides has garnered a great deal of interest from the biochemical and medicinal chemistry communities. Although nearly 9200 journal articles (as of March, 1999) regarding the ET peptide family have been published since its characterization in 1988, only a small fraction deal with structural preferences of ET-1 and its analogs. Originally discovered in the culture supernatant of bovine aortic endothelial cells and termed an endothelium-derived constricting factor, or EDCF (Hickey et al., 1985), ET-1 was isolated, cloned and sequenced from porcine aortic endothelial cells (Yanagisawa et al., 1988). The endothelium-derived relaxing factor (EDRF), later identified as nitric oxide (NO), had been previously isolated (Furchgott and Zawadski, 1980; Palmer et al., 1988).

Table 1.4: Primary sequences of the endothelin isopeptides.

	1 5	10	15	20	
ET-1	C S C S S	L M D K E	C V Y F C	H L D I I	W
ET-2	C S C S S	**W** L D K E	C V Y F C	H L D I I	W
ET-3	C **T** C **F T**	**Y K** D K E	C V **Y Y** C	H L D I I	W
VIC	C S C **N** S	**W** L D K E	C V Y F C	H L D I I	W
SRTX-6a	C S C **K D**	**M T** D K E	C **L N** F C	H **Q** D V I	W
SRTX-6b	C S C **K D**	**M T** D K E	C **L** Y F C	H **Q** D V I	W
SRTX-6c	C **T** C **N D**	**M T** D **Q** E	C **L N** F C	H **Q** D V I	W
SRTX-6d	C **T** C **K D**	**M T** D K E	C **L** Y F C	H **Q** D I I	W
BTX	C S C **A D**	**M T** D K E	C **L** Y F C	H **Q** D V I	W
Apamin	C **N** C **K A**	**P E T A L**	C **A R R** C	**Q Q H**-NH_2	

ETs, VIC, STRXs, BTX: disulfides at Cys^{1}-Cys^{15} and Cys^{3}-Cys^{11}
Apamin: disulfides at Cys^{1}-Cys^{11} and Cys^{3}-Cys^{15}

Yanagisawa et al. (1988) reported that a feature of the endothelins, specifically a relatively short amino acid sequence containing multiple disulfide linkages, was novel for bioactive mammalian peptides. This feature, the authors noted, was similar to the configurations of a number of non-mammalian peptide toxins "such as apamin from bee venoms, conotoxins from the venoms of fish-hunting sea snails and neurotoxins of scorpion venom" (Yanagisawa et al., 1988). The endothelins also have high sequence homology and similar functionality to mouse vasoactive intestinal contractor (VIC). In addition, the endothelins are similar to the sarafotoxins (SRTX) and bibrotoxin (BTX), peptide families of potent cardiotoxins isolated from the venom of Israeli (*Attractaspis engaddensis*) and South African (*Attractaspis bibroni*) burrowing asps (Sokolovsky, 1992; Doherty, 1992; Cody and Doherty, 1995). Table 1.4 lists the primary sequences for these isopeptides.

1.4.1: Endothelin Receptors

Early *in vivo* and *in vitro* studies demonstrated that the three known endothelin isoforms, designated ET-1, ET-2 and ET-3, induce both vasoconstrictor and vasodilator responses. The vasoconstriction has a rank order of potency (ET-1 > ET-2 >> ET-3), while the vasodilation is essentially equipotent (ET-1 = ET-2 = ET-3), suggesting the presence of two distinct endothelin receptors (Spokes et al., 1989; Takayanagi et al., 1991). The selective vasoconstrictor receptor, which is located in vascular smooth muscle cells, was termed ET_A, while the nonselective vasodilator receptor, which is found in the endothelium, was termed ET_B (Rubanyi and Polokoff, 1994; Goto et al., 1996; Webb and Meek, 1997). An ET-3 selective receptor, termed ET_C, has also been discovered and cloned from the dermal melanophores of *Xenopus laevis* (Karne et al., 1993).

Cloning and the subsequent sequence homology studies of the ET_A and ET_B receptors indicated that they are members of the rhodopsin superfamily of transmembrane G-protein coupled receptors (Arai et al., 1990; Sakurai et al., 1990; Birnbaumer et al., 1990). Basic features of this superfamily include 1) a hydrophobic extracellular N-terminal domain consisting of 75-100 amino acids, 2) seven amphipathic transmembrane helices connected by three inter- and three extracellular loops, and 3) an intracellular C-terminal domain. Among the members of the superfamily, only the transmembrane helices have highly conserved sequences; the loops and both terminal domains have highly diversified sequences (Rubanyi and Polokoff, 1994).

1.4.2: Biological Actions of Endothelins

Early *in vitro* studies indicated that the endothelins, especially ET-1, cause a potent and long-lasting constrictive response in vascular and arterial cell cultures (Yanagisawa et al., 1988). One hypothesis of the signal transduction involve the indirect association of the endothelins to intracellular Ca^{2+} concentrations ($[Ca^{2+}]_{int}$), a major

signal of contractility in smooth and cardiac muscles (Simmonson and Dunn, 1990; Kochva et al., 1993). Binding to the ET_A receptor leads to G-protein mediated activation of phospholipase C, which hydrolyses phosphatidylinositol, forming inositol-1,4,5-triphosphate (IP_3) and 1,2-diacylglycerol (DAG). Accumulation of IP_3 and DAG causes a prompt transient, followed by a long-lasting sustained, increase of $[Ca^{2+}]_{int}$ in vascular smooth muscle cells (Yanagisawa et al., 1988; Masaki and Yanagisawa, 1992; Rubanyi and Polokoff, 1994; Kuwaki et al., 1997). Binding to the ET_B receptor activates the release of nitric oxide and prostacyclin (both vasodilators) in endothelial cells, which implies a possible role for endothelin in the regulation of vascular tone. An imbalance in this regulatory mechanism has been postulated to lead to vascular disorders such as hypertension (Doherty, 1992; Rubanyi, 1994; Ferro and Webb, 1997).

Biological actions of endothelins aren't limited to vascular smooth muscles. Various studies have also demonstrated increased contractility of the heart and constriction of non-vascular smooth muscles such as the intestines, trachea and uterus (Doherty, 1992). Endothelins are also synthesized by and bind to receptors in various internal organs such as the kidneys, liver and lungs, as well as the endocrine, nervous and reproductive systems (Rubanyi and Polokoff, 1994). In addition, ET-1 mRNA is expressed in human macrophage and cancer cell lines (Doherty, 1992). These pieces of evidence would suggest that the endothelins are local, rather than circulating, hormones and indicate that they have wide physiological significance in diseases other than hypertension.

1.4.3: Structure-Activity Relationships of Endothelins

The three known endothelin isoforms consist of 21 amino acids and contain a bicyclic core defined by two disulfide bonds, $Cys^{1\text{-}15}$ and $Cys^{3\text{-}11}$ (see Table 1.4). In addition, all the endothelin, VIC, bibrotoxin and sarafotoxin isoforms, with the exception of SRTX-6c, include a cluster of charged amino acids from residues 8 → 10 (Asp^8-Lys^9-

Glu[10]) and a hydrophobic C-terminus consisting of six highly conserved residues (Yanagisawa et al., 1988; Kimura et al., 1988; Cody and Doherty, 1995). Apamin, a peptide isolated from bee venom, also has limited sequence homology to the ET/SRTX family. In this case, the peptide contains 18 residues and the disulfide crosslinking is switched (Cys^{1-11}, Cys^{3-15}) (Freeman et al., 1986).

A number of ET-1 solution state NMR studies have appeared since the peptide was characterized in 1989 (for example: Endo et al., 1989; Saudek et al., 1989; Tamaoki et al., 1991; Reily and Dunbar, 1991; Krystek et al., 1991; Andersen et al., 1992a). These studies were performed using a variety of media (aqueous, organic, or mixtures thereof) and a general consensus was reached regarding the conformational preferences of the bicyclic core. Most of these studies indicate that in aqueous media (containing organic co-solvents to prevent aggregation at neutral pHs), the peptide adopts a reverse turn from $Ser^5 \rightarrow Asp^8$ and an irregular helix from $Lys^9 \rightarrow Cys^{15}/His^{16}$. In DMSO, the helical region was thought to encompass $Leu^6 \rightarrow Cys^{11}$ (Munro et al., 1991). An early CD study (Calas et al., 1992) suggested that this region wasn't helical, but instead contained a series of β-turns. Other CD studies, however, have suggested that the intact peptide is 30-35% helical in aqueous media (Andersen et al., 1995a). Monocyclic analogs display decreased helicity by both NMR (Coles et al., 1994) and CD (Harris, 1993; Andersen et al., 1995a), suggesting that the disulfides limit the conformational versatility of endothelins, with most accessible conformers being helical over the 9-15 span.

The conformation of the hydrophobic C-terminus ($His^{16} \rightarrow Trp^{21}$), however, has elicited a great deal of controversy. Although most NMR-derived models indicate that this region is conformationally averaged, interaction between the rigid bicyclic core and the C-terminus are unresolved. Any differences in conformational preferences are most likely due to the solvent system used to study the peptide. In DMSO media, Saudek et al. (1989) proposed that the tail turns back towards the core. Endo et al. (1989) and Munro

et al. (1991) observed no intermediate or long range NOE interactions between the C-terminus and the bicyclic core. Krystek et al. (1991) and Andersen et al. (1992a) have reported structures in which the C-terminus interconverts between two discreet conformations in aqueous glycol media. In acetonitrile, the C-terminus undergoes rapid conformational averaging (Reily and Dunbar, 1991). A crystal structure (Janes et al., 1994) suggests that the helical region of the peptide extends along the sequence, completely encompassing the C-terminus. A time-resolved fluorescence study of ET-1, focusing on Trp^{21}, failed to distinguish between the various conformations observed in solution or crystalline form (Cowley and Pelton, 1995).

Several groups have shown the importance of the two disulfide bonds for biological activity of endothelin-1 (Kimura et al., 1988; Nakajima et al., 1989; Hunt et al., 1993). Reduction and carboxamidomethylation of the four cysteines, as well as destruction of the loop region (Ser^5-Asp^8), led to endothelin analogs with notably decreased potency (Kimura et al., 1988). An alanine scan of the cysteine residues showed that the outer disulfide ($Cys^{1\text{-}15}$) was more important than the inner disulfide ($Cys^{3\text{-}11}$): the monocyclic $Ala^{3,11}$-ET-1 analog demonstrated the highest *in vivo* potency among these alanine mutants (Nakajima et al., 1989). This study also included an acyclic peptide, $Asu^{1,15}$-$Ala^{3,11}$-ET-1, which showed negligible activity, suggesting that flexibility, or a reduction in helicity, in the core greatly diminishes potency.

Regioisomers of the disulfide bridges were also probed (Hunt et al., 1993). In this study, two of the four cysteines were substituted with penicillamine (β,β-dimethylcysteine) to create three endothelin analogs: (1) [$Pen^{1,11}$-Nle^7]-ET-1, (2) [$Pen^{3,15}$-Nle^7]-ET-1, and (3) [$Pen^{1,15}$-Nle^7]-ET-1. Equilibrium constants indicated that steric factors favor a Pen-Cys disulfide bond over the Pen-Pen/Cys-Cys combination. For the first two analogs studied, the 1-15, 3-11 (or "endomeric") regioisomer was highly favored over the 1-11, 3-15 (or "apameric") and 1-3, 11-15 regioisomers. The third peptide

slightly favors the apameric over the endomeric form. The $Met^{7} \rightarrow Nle^{7}$ substitution had earlier been shown not to affect the structure or biological activity of ET-1 (Aumelas et al., 1991). Vasoconstrictor activity studies indicated that the apameric and 1-3, 11-15 regioisomers were less potent than the endomeric forms by at least a factor of two. Of the endomeric analogs, [$Pen^{3,15}$-Nle^{7}]-ET-1 (henceforth, Pen-1) was shown to be a full and potent agonist of the ET_A receptor — the EC_{50} was identical to ET-1 while its binding constant was five fold higher than the native peptide. The [$Pen^{1,11}$-Nle^{7}]-ET-1 and [$Pen^{1,15}$-Nle^{7}]-ET-1 analogs showed partial and very little agonism, respectively, although the apameric form of the latter was slightly more effective than its endomeric counterpart (Table 1.5).

Table 1.5: Peptide penicillamine substitution pattern, disulfide isomer and biological activity[1].

Peptide	Disulfide Isomer	EC_{50} (nM)	% of ET Contraction	K_d (nM)	1-15, 3-11/1-11,3-15 ratio
ET-1	1-15, 3-11	0.94		0.15	3:1
$Pen^{1,11}$	1-15, 3-11	7.50	30	4.50	≥ 12:1
$Pen^{3,15}$	1-15, 3-11	0.89	100	0.70	≥ 22:1
$Pen^{1,15}$	1-11, 3-15	880	100	87	1:4
$Pen^{1,15}$	1-15, 3-11		5*	396	1:4

[1] Adapted from Hunt et al., 1993.

* At 10μM, produced 4.8 ± 2.5% of a maximal ET-1 contraction

The charged and aromatic residues located in the helical region of the peptide also influence activity. Nakajima et al. (1989) noted that the side chain carboxylic acid groups of Asp^{8} and Glu^{10}, as well as the aromatic moiety of Phe^{14} are important. Mutation of $Asp^{8} \rightarrow Asn^{8}$, $Glu^{10} \rightarrow Gln^{10}$ and $Phe^{14} \rightarrow Ala^{14}$ greatly reduced the biological activity of the peptide. Another study (Panek et al., 1992) examined two ET-1 analogs: [Pro^{12}]-ET-1 and [Phe^{16}]-ET-1. The $His^{16} \rightarrow Phe^{16}$ mutation demonstrated good agonist activity and some selectivity for the ET_A receptor. The $Val^{12} \rightarrow Pro^{12}$ mutant showed slightly

decreased potency and selectivity for the ET_B receptor. Since prolines are known to disrupt helices and, based on numerous NMR studies, position 12 is the central residue of the helix, the authors concluded that the helical region is not critical for binding and (ET_B) activity. However, a study of monocyclic ET-1 analogs (Andersen et al., 1995a) indicated that helix propensity affects ET_A receptor binding affinity and biological activity.

Various ET-1 structure-activity studies have also demonstrated the importance of the C-terminus (Kimura et al., 1988; Nakajima et al., 1989; Olsen et al., 1993). Progressive deletion of the six C-terminal residues led to decreasing potency, with ET(1-16) essentially inactive. The hexapeptide, ET(16-21), itself had no biological activity in most mammalian cell cultures. Alteration or removal of Trp^{21} led to dramatic decreases in contractility with respect to the intact ET-1. ET(1-20) was three orders of magnitude less potent than ET-1, while the D-Trp^{21} analog was less active by a factor of a hundred. Successive deletions along the chain, however, led to less dramatic decreases in potency (Kimura et al., 1988). Olsen et al. (1993) demonstrated that Ile^{20} was important as well: hydrolysis of the peptide bond between Ile^{19} and Ile^{20} lowered binding affinities to negligible levels. Longer sequences, such as ET(1-39), known as big-ET-1, ET(1-25) and ET-NH_2 also had lower biological activities. Proteolytic cleavage of the Trp^{21}-Val^{22} bond in big-ET increases contractility by over 100 fold (Yanagisawa and Masaki, 1989; Nakajima et al., 1989).

Since the conformationally averaged C-terminus is known to be essential for ET_A receptor binding affinity and signaling, a backbone N-methyl scan was performed over positions 17 $\rightarrow$ 20 (Hunt, unpublished data; Harris, 1993), while keeping the Pen-1 sequence intact. Activity studies performed at the Bristol Myers Squibb Pharmaceutical Research Institute (Princeton, NJ) indicated that the $^{(NMe)}Leu^{17}$ and $^{(NMe)}Asp^{18}$ analogs had slightly reduced potency with respect to Pen-1. The $^{(NMe)}Ile^{19}$ analog was less active than

the N-methylated Leu^{17}/Asp^{18} peptides. N-methylation at Ile^{20} showed a dramatic change in potency. The [$Pen^{3,15}$-Nle^{7}-$^{(NMe)}Ile^{20}$]-ET-1 peptide, henceforth Pen-2, is a potent antagonist of the ET_A receptor. Pen-2 also had the most dramatic change in the CD data observed for the N-methylated analogs (Harris, 1993). Difference CD of the intact analogs suggested that N-methylation over residues 17 → 19 promotes the formation of a reverse turn, possibly of type β_I. The ΔCD curve for Pen-2 corresponds to a β_{II}-type turn. A recent article (Cody et al., 1997) has also confirmed the ET_A receptor antagonism of a 6-mer containing $^{(NMe)}Ile^{20}$. Although relatively unstructured, the N-methyl group had a large effect on the conformation. According to Cody et al. (1997), the peptide bond between Ile^{19} and $^{(NMe)}Ile^{20}$ adopted a *trans* conformation in aqueous media. When the 6mer was studied in DMSO at basic pH, the same bond was reported to exclusively adopt the *cis* configuration.

1.5: PROJECT GOALS

One of the primary goals of this doctoral project was to resolve major questions concerning the three-dimensional structure of the endothelins and how these relate to their biological activity. In order to study peptide hormones normally bound in transmembrane receptors, a more rigid analog was designed and studied in aqueous media containing organic co-solvents. This analog, [$Pen^{3,15}$-Nle^{7}]-ET-1 (Pen-1), not only displayed a more rigid bicyclic core with respect to the native peptide, but was also a potent ET_A receptor agonist. Studies involving a backbone N-methyl scan of the conformationally-averaged C-terminus gave rise to a second endothelin analog, [$Pen^{3,15}$-Nle^{7}-$^{(NMe)}Ile^{20}$]-ET-1 (Pen-2), which retained the rigid bicyclic core of Pen-1, but was discovered to be a potent ET_A receptor antagonist (Hunt, unpublished data).

These discoveries led to further studies in order to determine conformational requirements for recognition *versus* those for signaling (i.e. agonism vs. antagonism).

Specifically, questions regarding the structural preferences of ET-1, Pen-1 and Pen-2 led to studies of various ET-1 and Pen-2 C-terminal analogs in order to determine: 1) whether the two disulfide bridges of the ET-1 bicyclic core are required to stabilize helix formation, 2) the importance of the C-terminal hydrophobic clustering and its relationship to peptide signaling, and 3) the effect of the N-methyl group on the conformational preferences of Pen-2 and its analogs.

CHAPTER 2: INTACT ENDOTHELIN ANALOGS

2.1: INTRODUCTION: WHY STUDY PEN-1 AND PEN-2?

In their study of endothelin-1 regioisomer formation, Hunt et al. (1993) synthesized an "endomeric" (1-15, 3-11 crosslinked) analog that was later determined to be a potent ET_A receptor agonist. This analog, termed Pen-1 (Figure 2.1), had three significant mutations in the peptide sequence. The first, $Met^7 \rightarrow Nle^7$, was to help prevent oxidative reactions associated with the methionine side chain. Previous NMR studies had determined that this change did not significantly alter the overall structure of ET-1 (Aumelas et al., 1991). The two other mutations, $Cys^3 \rightarrow Pen^3$ and $Cys^{15} \rightarrow Pen^{15}$, were used to insure the correct endomeric regioisomer formed in high yields upon oxidation of the linear peptide (Hunt et al., 1993). Initial structure studies performed in the Andersen group (Chen, 1992; Harris, 1993) determined that the Cys → Pen mutations increased steric interactions within the bicyclic core. This created a more rigid core which helped stabilize and extend the helix from Lys^9 to Leu^{17}, or possibly Asp^{18}. As with most other ET-1 solution structures, Pen-1 displayed rapid conformational motion at the C-terminus.

Since the conformationally flexible C-terminus was known to be an important signaling factor for the native peptide, several more structured C-terminal analogs were studied. A backbone N-methyl scan of the C-terminal residues was performed at Bristol Myers Squibb (Hunt, unpublished data; Harris, 1993). One mutant, denoted Pen-2 (Figure 2.1), retained the same substitutions as the Pen-1 analog but contained a N-methylated isoleucine at position 20. Initial structural studies of the Pen-2 analog (Chen, 1992; Harris, 1993) showed that the bicyclic core retained the same general features as

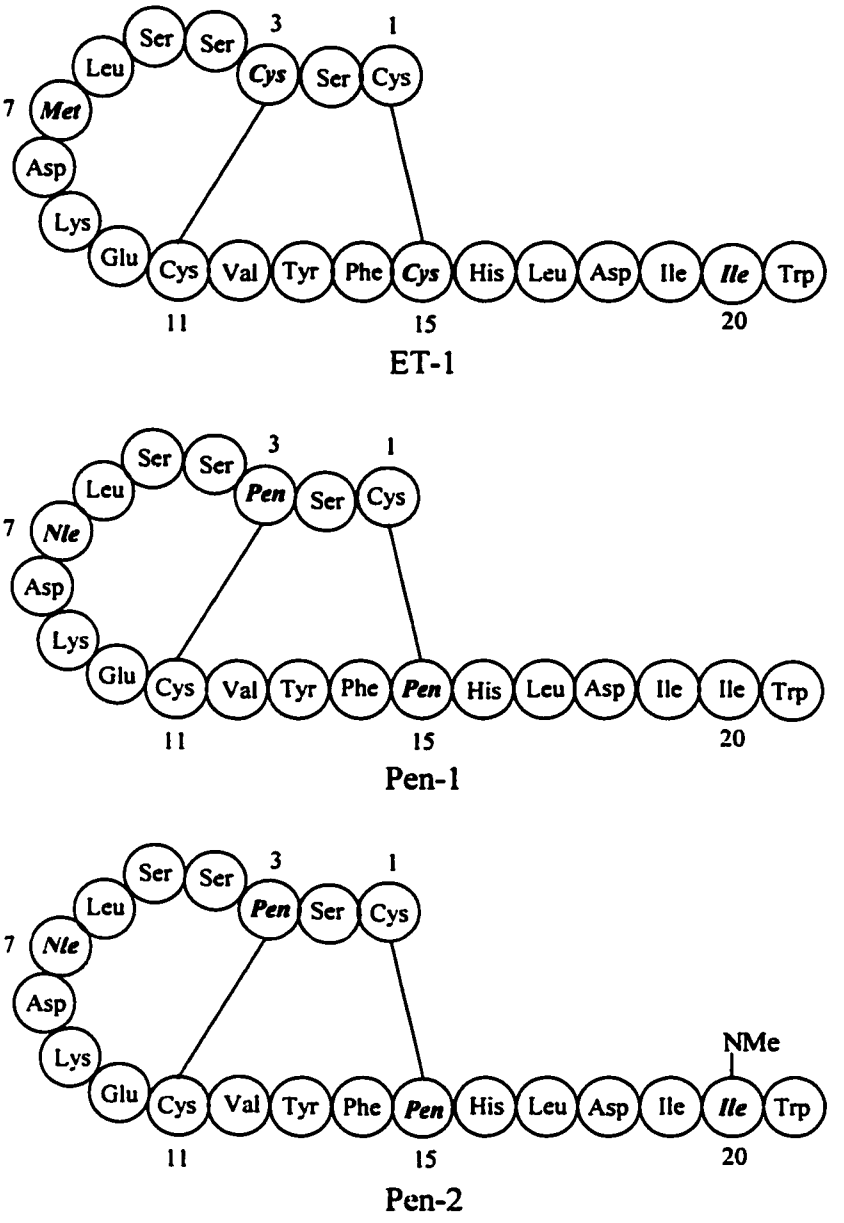


Figure 2.1: Snake diagrams of ET-1, Pen-1 and Pen-2.

that of Pen-1. However, the C-terminus was noted to be more rigid, possibly containing a conformation resembling a β_I type turn. Pharmacological studies indicated that the Pen-2 mutant was the most potent ET_A receptor *antagonist* of the N-methyl scan analogs (Hunt, unpublished data).

Two questions arise from these initial surveys. First, how does the addition of a single methyl group affect the conformation of the C-terminus? Second, how does the change in the structure affect the signaling? Specifically, how does the N-methylation of a single residue in the C-terminus transform a potent ET_A receptor *agonist* (Pen-1) to a potent ET_A receptor *antagonist* (Pen-2)? An ET-1 crystal structure (Janes et al., 1994), whose overall conformation widely contrasted those derived from solution-state NMR data was published after these preliminary analyses. This raised several questions regarding the conformation of the C-terminus. This chapter completes the structural studies of the intact ET-1 analogs started by Chinpan Chen (1992) and Scott Harris (1993), as well as discusses how the solution-state conformation differs from the one solid-state structure reported.

2.2: MATERIALS AND EXPERIMENTAL DETAILS

2.2.1: Peptide Synthesis and Sample Preparation

Isotope labeled [$Pen^{3,15}$-Nle^{7}]-ET-1 (7mg, FW = 2548.21) was acquired from the Bristol Myers Squibb Pharmaceutical Research Institute (Princeton, NJ) and was used without further purification. Synthesis is as previously described (Hunt et al., 1993) with the following exception: ^{15}N-Leu at positions 6 and 17, ^{15}N-Phe at position 14, d_9-Nle at position 7, and d_8-Val at position 12 were incorporated in the sequence during the automated peptide synthesis. [$Pen^{3,15}$-Nle^{7}-$^{(NMe)}Ile^{20}$]-ET-1 (FW = 2542.12) was also acquired from Bristol Myers Squibb and was used without further purification.

Freshly prepared MilliQ® water was used to prepare the phosphate buffers. Deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Due to the aggregation problem encountered in earlier NMR studies (Chen, 1992), organic co-solvents were required to solubilize the samples. For the isotope labeled Pen-1, the solvent mixture used for the initial NMR studies was a follows: 52% d_2-formate/phosphate buffer (pH 3.5), 42% d_6-ethylene glycol and 6% 80mM aqueous d-trifluoroacetic acid (TFA) by volume; the nominal peptide concentration was 2.56mM. Aliquots of d_2-hexafluoroisopropanol (HFIP) were added directly to the NMR tube during the titration study such that volume percentages ranged from 3 to 15%. In the deuterium exchange studies, the sample was lyophilized and reconstituted under the same conditions as the original sample (0% HFIP) using 99.96% D_2O during the preparation of the buffers.

The unlabeled Pen-2 sample used in this study was lyophilized powder recovered from initial NMR studies performed by Chinpan Chen (1992). These samples were reconstituted in 5% d_4-acetic acid (HOAc), 40% d_6-ethylene glycol and 55% H_2O by volume. The concentration of the resulting solution was approximately 2.45mM. The Pen-2 HFIP titration was performed as described for the isotope labeled Pen-1. However, due to solubility problems, no deuterium exchange study was performed.

2.2.2: NMR Experiments

Two Brüker Instruments AM-500 spectrometers (~11.74T, one tuned to 500.13 MHz, the second tuned to 499.87 MHz) equipped with Aspect 3000 computers were used to acquire and collect NMR data for the isotope labeled Pen-1 and Pen-2 analogs. Most spectra were acquired at temperatures ranging from 290 to 305K using 16K complex points for 1D spectra. During the isotope-labeled Pen-1 temperature studies, spectra were collected from 275 to 305K (280 to 310K for Pen-2) in 5K intervals. The temperature was allowed to equilibrate for 20-30 minutes prior to acquisition. Data points in the

deuterium exchange study were acquired at 300K. For the 2D experiments, 2048 complex t_2 points and > 500 t_1 increments were collected. In order to gain sufficient signal-to-noise resolution, 64 to 128 scans per t_1 increment were acquired. Probe temperatures were calibrated once every six months using degassed methanol in a sealed NMR tube.

Pulse sequences for the TOCSY (Griesinger et al., 1984; Bax & Davis, 1985) and NOESY (Bodenhausen et al., 1984) programs are the same as those found in the Ph.D. dissertations of Bolong Cao (1993) and Scott Harris (1993). All TOCSY and NOESY spectra were acquired using the time proportioned phase increment (TPPI) method. Typical MLEV-17 (Levitt et al., 1982; Bax and Davis, 1985b) spin lock times for the TOCSY experiments ranged from 45 to 50 ms while mixing times for the NOESY experiments ranged from 80 to 300ms. The residual water signal was suppressed during the recycle delay using a presaturation pulse sequence applied over an 8Hz range centered on the solvent peak.

All NMR data were transferred to a Silicon Graphics 4D/20 or 4D/25 Personal Iris computer for processing and analysis using Felix versions 2.0, 2.3 or 95.0 (Biosym Research Inc., San Diego, CA). Prior to Fourier transformation, all 2D data sets were zero filled to 2K data points and apodized with a skewed sine bell function. Where applicable, baseline correction using a cubic spline function was also performed on the 2D data sets.

2.2.3: Spectral Assignment

The chemical shifts for all Pen-1 and Pen-2 spectra were calibrated with respect to the trimethylsilylpropanoate (TSP) reference peak. Spectral assignment of the 2D spectral set follows the strategy outlined in Wüthrich (1986). Individual spin systems were located and assigned via the TOCSY, while sequential and intermediate range

interactions were derived from the NOESY. Chemical shift deviations for both backbone amide and α-methine protons were calculated using the disorder reference chemical shifts appearing in Appendix A. These reference values had recently been re-determined from several laboratories (Wishart et al., 1995; Merutka et al., 1995; Andersen et al., 1997). NOE ratios and peak intensities were determined directly from the NOESY spectra. Inter- *vs.* intraresidue αN NOE ratios (Lee et al., 1994) are shown as differences in crosspeak contour levels, with the N-line corresponding to $(N_i\alpha_{i-1})$ - $(N_i\alpha_i)$ and the α-line to (α_iN_{i+1}) - $(N_i\alpha_i)$.

2.2.4: Molecular Dynamics Simulations

Starting structures for modeling Pen-1 were created in the Sequence Builder module of Quanta version 3.3 (Molecular Simulations Inc., San Diego, CA). One β-sheet and one right handed α-helix starting structure were subjected to a 100 step CHARMm energy minimization. Next, five to six sets of phi (ϕ) and psi (ψ) dihedral angles were randomized at various positions to create 30 acyclic structures (10 derived from helices, 20 from strands) used for the initial refinement stages.

Molecular dynamics (MD) and simulated annealing (SA) experiments were performed using X-PLOR version 3.1 (Brünger, 1988) on Silicon Graphics Indigo workstations. Distance constraints were derived from NOESY data sets acquired in both protic and deuterated media. These constraints were placed in different categories based on the level of confidence in the NOESY crosspeak assignments and NOE intensities. A complete listing of the Pen-1 distance and dihedral constraints appear in Appendix D. In the case of the isotope labeled Pen-1 analog, more distance constraints could be used with higher confidence due to the lack of spectral overlap in the backbone amide and methyl regions. Distance constraints with high uncertainties were placed in a low-weighted "trial" category. Dihedral (ϕ and ψ) constraints were based on α-proton chemical shift

deviations (αH-CSD) and local NOE ratio data (Andersen et al., 1995c). Where the β-CH_2s are resolved and prochirality can be assigned, anti- and gauche- (αH or backbone NH to βHs) constraints were included. Pen α-methine to γ-methyl anti- and gauche- assignments were also included in the constraint set. Modified topology files which contained information pertaining to Pen and Nle residues were used to set bond lengths, bond angles, improper angles, atomic charges, atomic masses, aromatic ring and amide bond planarity, and force field energy constraints.

Several stages of X-PLOR calculations were required in order to allow the structures to settle in a global energy minimum. Each round, in turn, required several MD calculations prior to the simulated annealing. Typical X-PLOR protocols initially set distance constraints in "soft" potential energy wells with low asymptotic slope values. The structures were subjected to a 150 step Powell energy minimization, then randomization of velocity vectors before undergoing a short high temperature molecular dynamics calculation (100 steps, 1000.01K) which would enable the structures to sample all allowable configurations. The weights of the distance constraints were increased before a second MD run (2000 steps for the initial simulations, 8000 steps for the refinement stages). A third MD calculation (100 steps) was simulated in which the dihedral energy, van der Waals repulsion and asymptote slope values are ramped up. Prior to a fourth MD calculation in which the temperature was ramped down from 1000.01K to 300K in 50K intervals, the higher weighted category constraint sets were switched from a soft to square potential energy well. Simulated annealing allows the structures to slowly fall into an energy well before the final Powell minimization is performed. Timesteps of 2fs were used for all MD simulations with the exception of the initial rounds (1fs per step).

Initial disulfide patching runs were performed with Lennard-Jones potentials turned off. Long distance (nonsequential) and disulfide constraints were set to low

weights and kept in soft energy wells until after the first simulated annealing calculation. "Trial" category sets were kept in soft energy wells with shallow asymptotes at all times to prevent uncertain constraints from having a large influence in the initial simulations. This primary stage was performed twice prior to bridging the disulfide bonds in the model. The long distance and disulfide constraints were returned to their normal weights during the third round of refinement calculations. All resolvable β-CH_2 assignments were wildcarded while stereospecific γ-methyl assignments for Val^{12} were used throughout the first three stages. Since the sidechain configuration of the Cys and Pen residues have great influence over the conformation of the bicyclic core, determining the correct stereochemical assignments was imperative. An iterative approach was utilized during the refinement stages in order to establish prochiral assignments for the Pen methyl groups, resulting in four permutations (60 models each) after the third stage. Structures were discarded from ensembles based upon high NOE, improper, and/or van der Waals energies, as well as poor convergence and large rms deviations of the superimposed backbone over the helical portion of the molecule. However, convergence criteria were not allowed to eliminate structures that have fully acceptable energies and NOE violations.

Once a consensus distance constraint set was reached, the original 30 acyclic structures were subjected to several refinement stages where only the stereochemical assignments for the Pen^3, Cys^{11}, Val^{12} and Pen^{15} sidechains were enforced. All other resolvable β protons were wildcarded at this stage. Lennard-Jones potentials were applied during the last energy minimization step of the final refinement stages to more accurately model attractive-repulsive interactions.

Purely repulsive van der Waals potentials were replaced with standard Lennard-Jones potentials during the comparison study of the Pen-1 solution and ET-1 crystal structures. Where the helical portions of the structure were emphasized, narrowly defined

ϕ and ψ dihedral constraint ranges ($\pm < 10°$) were used to test NOE compatibility with a "regular" helical conformation. The NOE force constant (k_{noe}) was also decreased so that the E_{noe} contribution in the total energy term was less than that of the dihedrals (k_{cdih}). For the ensemble where the NOEs were emphasized, the dihedral constraints were allowed their normal ranges (see Appendix D) and force constants while k_{noe} was increased. Throughout this comparison, k_{noe} and k_{cdih} were adjusted so that their constraints contributed equally to E_{total}.

2.3: NMR STUDIES OF PEN-1 AND PEN-2

2.3.1: WHY USE ISOTOPE LABELS ON PEN-1?

The initial Pen-1 survey (Chen, 1992) provided a first generation NMR structure ensemble derived from the available TOCSY and NOESY spectra. However, a number of degenerate chemical shifts in the backbone amide region resulted in overlapped and indistinguishable NOESY crosspeaks. This lead to several misassignments or mis-interpretations that were later translated into highly uncertain distance constraints. This was especially problematic since these overlapping amides corresponded to those located in the turn and helical regions of the peptide. For example, in 50% aqueous glycol at pH 3.5 (Figure 2.2, top), a number of overlapping resonances are evident: 4NH with 6NH, 14NH with 18NH (and possibly 9NH and 21NH) and 17NH with a number of other amides (7NH, 8NH, 12NH, and 19NH). Since these amide proton chemical shifts could not be sufficiently resolved by temperature or pH changes, any distance constraints associated with these amides could not be used with high confidence during the structure calculations.

To more accurately probe the conformational preferences of the Pen-1 analog, isotope labels were placed in key regions of the peptide. Two ^{15}N-labeled leucines were

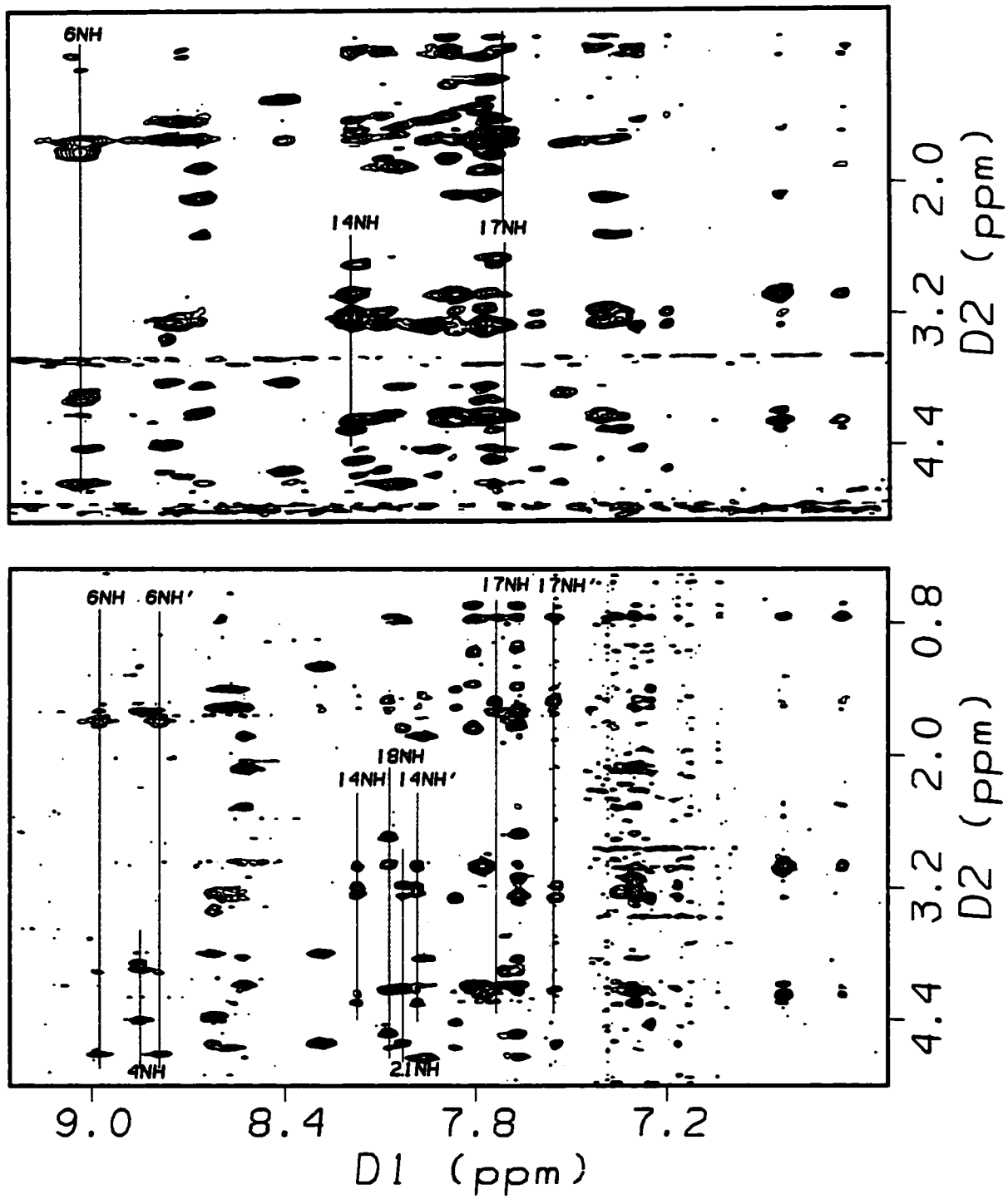


Figure 2.2: Top, the amide NH → αH/upfield crosspeak region of the Pen-1 NOESY, acquired at 500MHz (Chen, 1992). Bottom, the same region of the isotope-labeled Pen-1 NOESY. Large ^{15}N-^{1}H couplings (*circa* 90Hz) are indicated for Leu^{6}, Phe^{14} and Leu^{17}.

inserted into the sequence at positions 6 (the loop region) and 17 (the C-terminus), while a ^{15}N-Phe was placed at residue 14 (the helical region). Originally, ^{15}N dynamics studies to probe the mobility of these three regions and ^{15}N-edited NOESYs to further deconvolute the spectrum were planned. Unfortunately, these studies weren't feasible due to the low signal-to-noise resolution of the instrument at its then current configuration (1993-1994). However, a standard NOESY reveals strong ^{15}N-^{1}H couplings of *circa* 90Hz (Figure 2.2, bottom), which relieves some spectral overlap in the amide region.

Due to the high number of fatty alkyl residues in the Pen-1 sequence (2 Leu, 1 Nle, 2 Ile, and 1 Val), the methyl region of the NOESY spectrum was overcrowded with crosspeaks. Correct attribution of these upfield regions is critical since three of the aliphatics (Leu^{17}, Ile^{19} and Ile^{20}) are located in the conformationally averaged C-terminus. To help alleviate spectral crowding in the upfield region of the spectrum, two perdeuterated residues, d_9-Nle^{7} and d_8-Val^{12}, were inserted into the sequence. Since the deuteria in these two residues are unobserved in the proton spectrum, interactions between the C-terminus and the bicyclic core, if any, should be more easily observed.

2.3.2: Spectral Assignment of Pen-1.

2.3.2.1: Aqueous glycol media.

Spectral assignment of the isotope-labeled Pen-1 analog in 50% aqueous glycol (pH 3.5) closely followed that of Chinpan Chen's (1992) analysis. As expected, ^{15}N-^{1}H couplings of *circa* 90Hz and the effect of selective perdeuteration (in the form of absent Nle^{7} and Val^{12} resonances) are clearly observed (Figure 2.2, bottom). A peak due to a residual α-methine proton for the Nle^{7} is also present in the spectrum. Signals at *circa* 4.8ppm, such as the α-methine proton of Pen^{3}, were bleached out due to solvent suppression. Table 2.1 reports the final chemical shift assignments and backbone NH

Table 2.1: Chemical shift assignments for Isotope-labeled [$Pen^{3,15}$-Nle^{7}]-ET-1 in 50% aqueous ethylene glycol (pH 3.5) at 305K.

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Cys 1	exchanged	4.341	3.405, 3.258	
Ser 2	8.600 [-4.559]	4.582	3.794, 3.794	
Pen 3	8.248 [-7.293]	bleached		γ,γ′ 1.595, 1.212
Ser 4	8.814 [-6.823]	4.394	3.925, 3.866	
Ser 5	7.458 [+0.951]	bleached	?	
Leu 6 *	8.829 [-9.324] J = 92Hz	3.960	1.661, 1.583	γ 1.703, δ,δ′ 0.922, 0.813
Nle 7 ^	7.695 [-2.722]	4.080	n/a	γ,γ′ n/a, δ,δ′ n/a, ε,ε′ n/a
Asp 8	7.644 [-0.692]	bleached	3.263, 2.708	
Lys 9	7.972 [-2.937]	3.849	1.831, 1.831	γ,γ′ 1.572, 1.474, δ,δ′ 1.678, 1.678
				ε,ε′ 3.002, 3.002
Glu 10	8.485 [-7.656]	4.084	2.123, 2.123	γ,γ′ 2.466, 2.466
Cys 11	7.319 [+0.682]	4.135	3.247, 3.139	
Val 12 ^	7.646 [-0.939]	n/a	n/a	γ,γ′ n/a
Tyr 13	7.754 [-1.609]	4.161	2.993, 2.993	δ,δ′ 6.853, ε,ε′ 6.664
Phe 14 *	8.066 [-3.515] J = 92Hz	4.253	3.256, 3.200	δ,δ′ 7.311, ε,ε′ 7.373, ζ ~7.373
Pen 15	8.503 [-10.970]	bleached		γ,γ′ 1.566, 1.406
His 16	7.870 [-2.552]	4.412	3.291, 3.291	δ 7.256, ε 8.602
Leu 17 *	7.661 [-2.303] J = 91Hz	4.132	1.521, 1.597	γ 1.521, δ,δ′ 0.757, 0.757
Asp 18	8.078 [-0.491]	4.503	2.967, 2.739	
Ile 19	7.650 [-3.782]	4.076	1.734	γ1,γ1′ 1.371, 1.035
				γ2 0.649, δ 0.750
Ile 20	7.773 [-5.499]	4.115	1.755	γ1,γ1′ 1.361, 1.059
				γ2 0.763, δ 0.763
Trp 21	7.985 [-7.899]	bleached	3.283, 3.185	δ 7.173, ε1 ?, ε3 7.559
				ζ2 7.396, ζ3 7.050, η2 7.132

*** ^{15}N, ^ perdeuterated side chains**

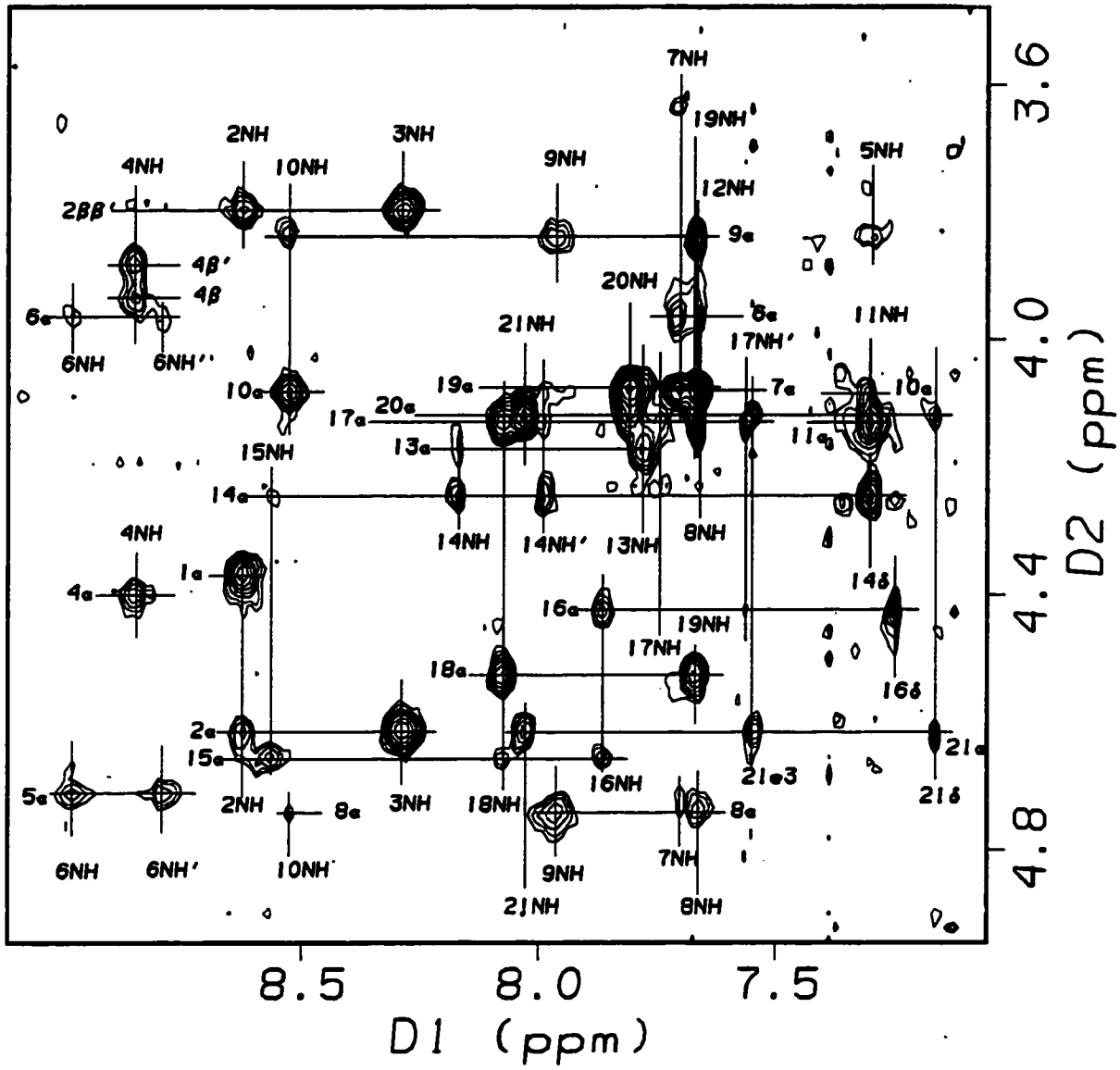


Figure 2.3: The amide NH → αH crosspeak region of the isotope-labeled Pen-1 NOESY, acquired at 500MHz. Signals *circa* 4.8ppm have been bleached out due to the residual HOD peak.

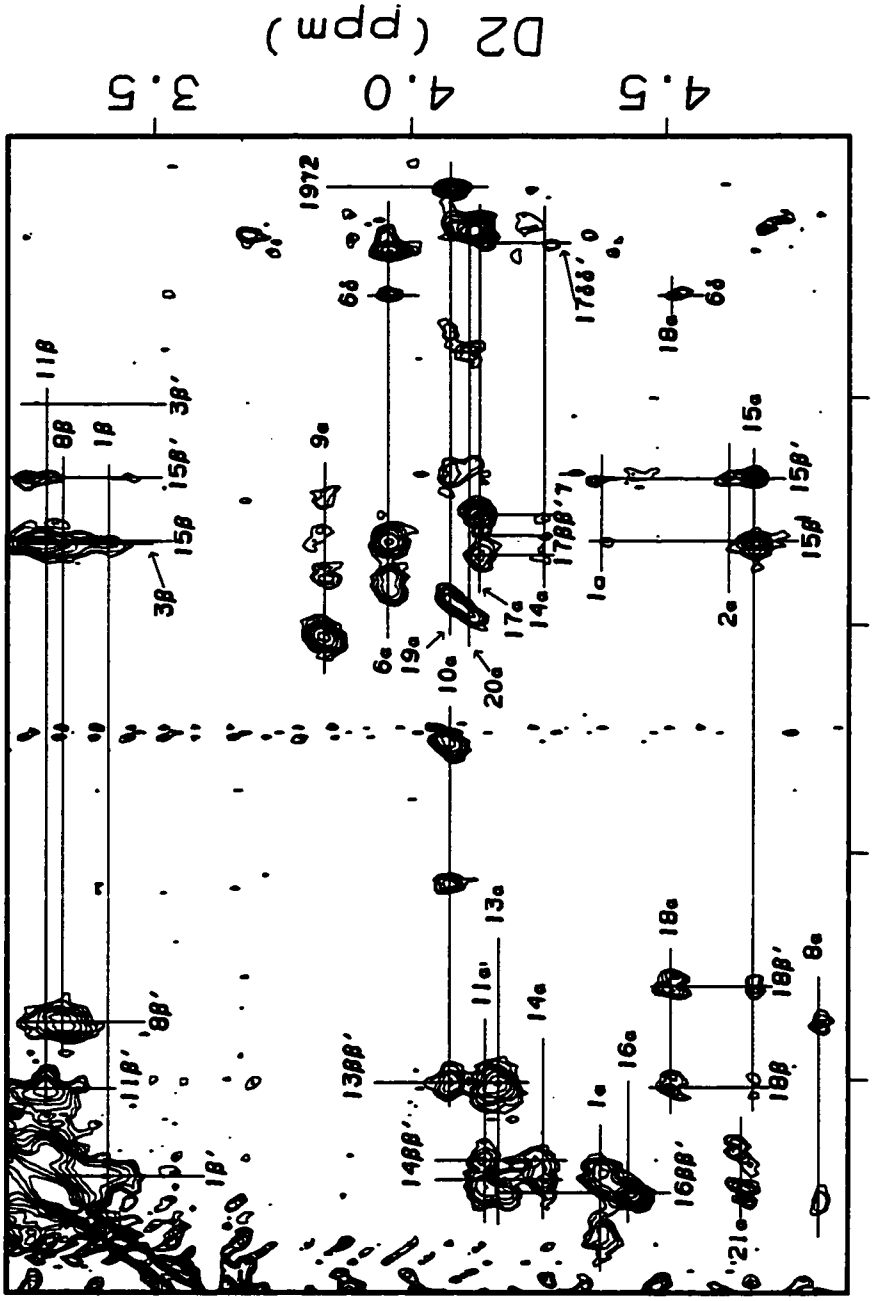


Figure 2.4: The αH → upfield crosspeak region of the isotope-labeled Pen-1 NOESY, acquired at 500MHz in 50% ethylene glycol/50% acidic D_2O buffer.

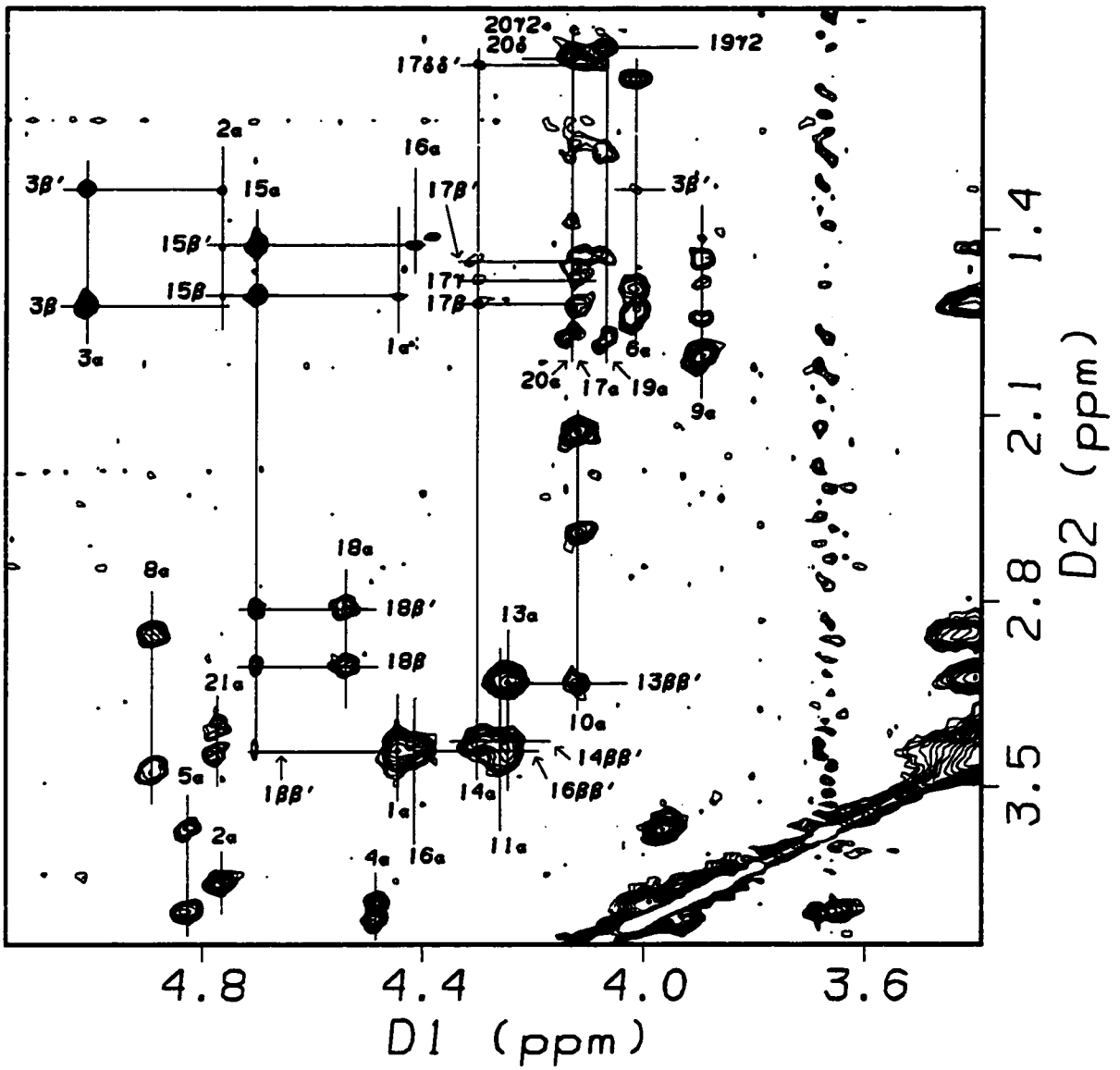


Figure 2.5: The αH → upfield crosspeak region of the isotope-labeled Pen-1 NOESY, acquired at 500MHz in 5% ethylene glycol/95% acidic D_2O buffer. This provided the stereoselective Pen methyl (3β, 3β′, 15β, 15β′) NOE connectivities.

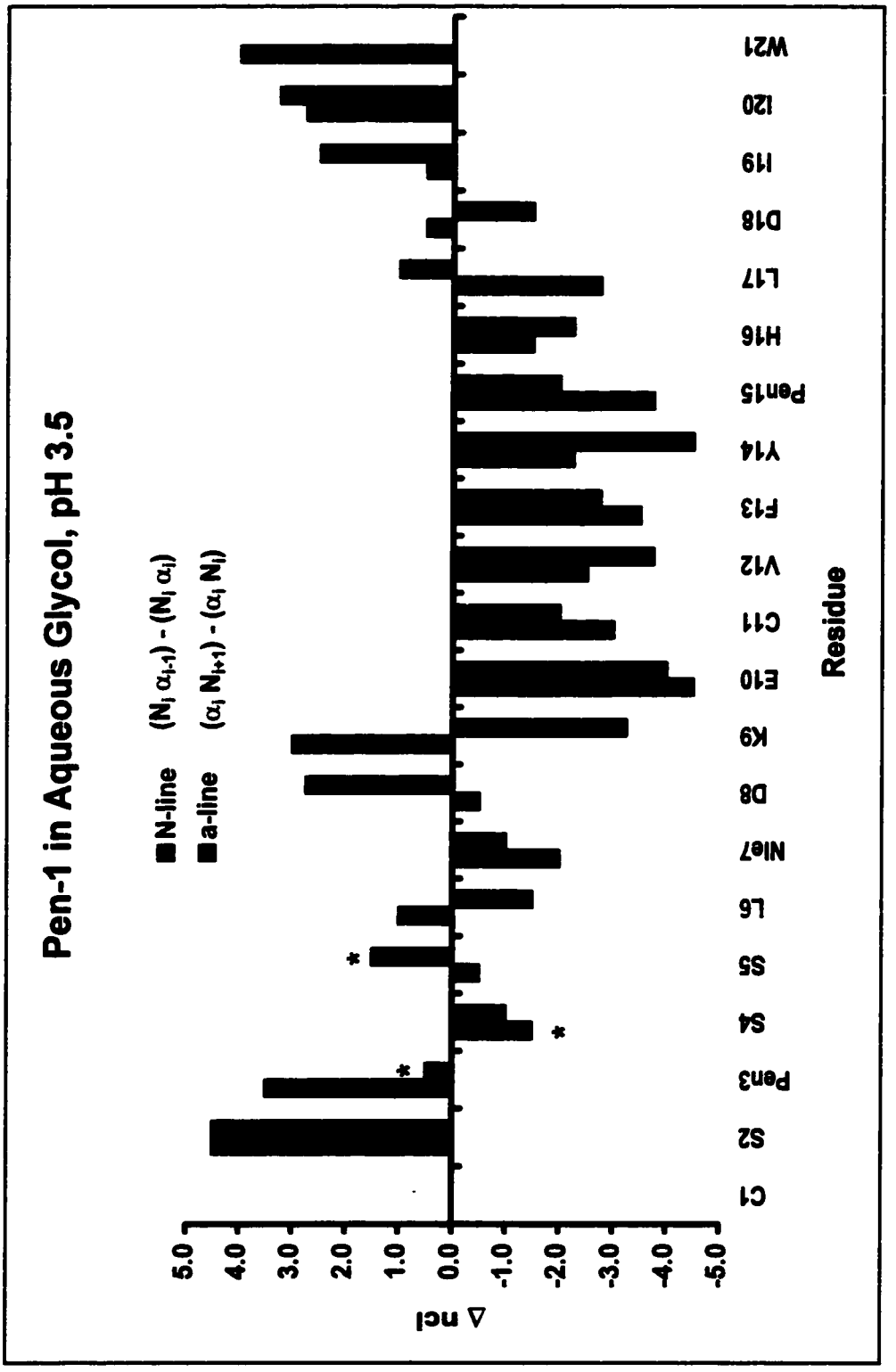


Figure 2.6: NOE ratio plot of Pen-1 in 50% aqueous glycol. Asterisks indicate less well-determined NOE intensities.

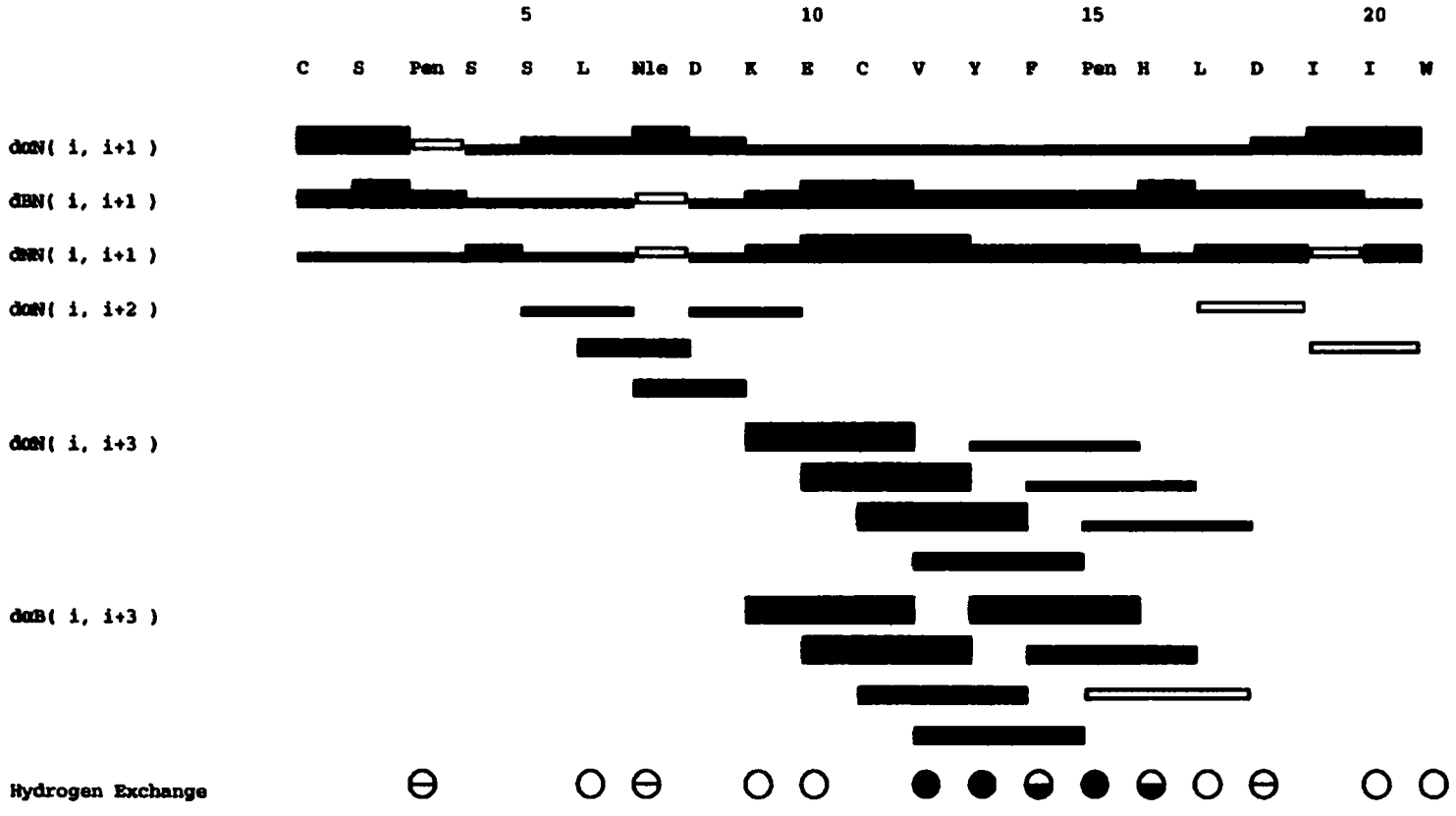


Figure 2.7: NOE connectivity and D_2O exchange plot for Pen-1 in 50% aqueous ethylene glycol. For the hydrogen exchange data: unlabeled points, indistinguishable or $t_{½} < 4$ minutes; open circles, $t_{½} = 5$-20 minutes; slashed circles, $t_{½} = 0.4$-1.2 hours; half-filled circles, $t_{½} = 2$-5 hours; filled circles, $t_{½} = 6$-16 hours.

temperature gradients. A number of sequential and intermediate-range crosspeaks are observed in the H_2O and D_2O NOESYs (Figures 2.3 - 2.5). These are summarized in the NOE ratio and connectivity plots (Figures 2.6 and 2.7).

Two major structure motifs are evident from these data sets: a turn encompassing $Ser^5 \rightarrow Asp^8$ and a helix from $Lys^9 \rightarrow Leu^{17}$, possibly extended to Asp^{18}. The turn, which caps the N-terminal portion of the helix, is primarily indicated by a number of weak-to-medium $\alpha_i \rightarrow N_{i+2}$ NOEs for the residues ranging from Ser^5 and Glu^{10}. Also observed in the NOESY, but not listed in the NOE connectivity plot, are weak crosspeaks between the 5β methylene protons and 7NH. Sequential interresidue NOEs also help locate the turn: medium $\alpha_i \rightarrow N_{i+1}$ and weak $N_i \rightarrow N_{i+1}$ crosspeaks are present for Ser^5, Leu^6 and Asp^8. The corresponding crosspeaks for Nle^7 are strong and unquantifiable, respectively. The NOE ratio plot for this region is less helpful. Ser^5 through Asp^8 display a number of alternating positive and negative Δncl values for both the N- and α-lines.

The helix shows several strong to medium $\alpha_i \rightarrow N_{i+3}$ interactions, extending from Lys^9 to Asp^{18}. These NOEs are strongest at the beginning of the helix and becomes weaker as the sequence progresses, indicating that fraying occurs at the C-terminal end. For example, an intense crosspeak is observed for 9α → 12NH, while 15α → 18NH shows medium intensity. The 12α → 15NH crosspeak, which is observed in the non-isotope case, is absent here. A crosspeak between 10α and 13NH is also present, although it lies near the 19α → 20NH crosspeak. Other crosspeaks are broadened due to ^{15}N-^1H coupling, specifically those corresponding to 11α → 14NH and 14α → 17NH. The same pattern emerges for the $\alpha_i \rightarrow \beta_{i+3}$ interactions. Other *i* → *i+3* interactions involving 10α and 11α and the aromatic side chains of residues 13 and 14 are also present. The former (10α → 13δδ′) is easily observed; the latter (11α → 14δδ′) is ambiguous due to the coincident chemical shifts of 11NH and the 14δ protons.

Sequential interresidue NOEs also help resolve this region: weak $\alpha_i \rightarrow N_{i+1}$ and medium-to-strong $N_i \rightarrow N_{i+1}$ crosspeaks are observed from Lys[9] to Asp[18]. The N- and α-line of the NOE ratio plot displays negative Δncl values approaching -4 for residues Glu[10] through Pen[15], indicating that this region is nearly 100% helical.

Further evidence of the helix N-cap is present in the NOE ratio histogram. The intense α-line of Asp[8] and N-line of Lys[9], which both represent ψ_8, indicates that the *inter*residue αN crosspeaks are more intense than their *intra*residue counterparts. However, the abrupt change in the sign of the NOE ratios beginning with the α-line of Lys[9] indicates the reverse is true for residues 9 through 16 and that Asp[8] is the N-capping unit of the helical region. The lack of any $i \rightarrow i+3$ crosspeaks prior to Lys[9] in the sequence (as observed in the NOE connectivity plot) also confirms this conclusion.

Strong $\alpha_i \rightarrow N_{i+1}$ and $\beta_i \rightarrow N_{i+1}$, but weak $N_i \rightarrow N_{i+1}$ crosspeaks suggest that the N-terminus (Cys[1] → Pen[3]) is extended. The NOE ratio plot shows strong positive Δncl values for the N- and α-lines of Ser[2] and Pen[3], although the Pen[3] α-line is less well-determined due to its proximity to the residual water peak. The C-terminus (Asp[18] through Trp[21]) is the least structured region of the peptide. The NOE ratios indicate an extended conformation, as shown by the positive values approaching +4 for both the N- and α-lines. However, sequential $\alpha_i \rightarrow N_{i+1}$ and $N_i \rightarrow N_{i+1}$ crosspeaks have strong and medium intensities, respectively. In addition, there are several $i \rightarrow i+2$ interactions, between Leu[17] and Ile[19], as well as Ile[19] and Trp[21], indicating the possibility of a turn.

2.3.2.2: Deuterium exchange.

Expectation amide exchange rate ratios can be calculated from the available data using Molday factors (Molday et al., 1972; Bai et al., 1995). The individual amide reference values are dependent upon its own sidechain and that of the previous residue.

Elevated temperatures, as well as acid or base catalysis, also accelerate exchange rates. The minimum (slowest) rates for Ala occurs near pH 3 (Bai et al., 1995). A positive Molday factor indicates an enhancement of exchange rates, relative to Ala; a negative Molday factor represents retardation. Bai et al. (1995) reports the procedure for calculating Molday factors and correcting disordered Ala exchange rates for both temperature and pH effects. Once the reference $t_{1/2}$ for the other residues in the peptide are calculated [(ref $t_{1/2}$)$\cdot 10^{-MF}$], protection factors, which are a measurement of the relative folding, are obtained from the quotient of the observed and predicted half-lives.

Table 2.2 reports the backbone amide exchange half-lives and protection factors for Pen-1 in aqueous 50% ethylene glycol (pH 3.5) at 302K. At this pH and temperature, base catalysis dominates the H-D exchange reaction. The chosen Molday factors (Bai et al., 1995) reflect this condition. Cys^{1} is omitted from the table since the N-terminal residue is a free amine, and due to normal fast exchange with the solvent, is not observed in the ^{1}H spectrum. The Ser^{5} and Cys^{11} amide peaks are indistinguishable from the aromatic protons. For Asp^{8} and Asp^{18} (pKa $\approx$ 3.9), Molday factors for both Asp° and Asp^{-} were calculated. The Asp^{-} values were used to determine the slowest predicted $t_{1/2}$, resulting in the *minimum* protection factor for residues 8/9 and 18/19.

The D_2O exchange data (Figures 2.8 and 2.9) confirms the location of the helical region. The largest protection factors range between Glu^{10} and Leu^{17} (all > 10). The unexpectedly large protection factor for His^{16} appears to be an anomaly. Under the current experimental conditions, the sidechain have been protonated, giving a predicted exchange half-life of 45 seconds. Chen (1992) reports a $t_{1/2}$ between 5-20min under comparable conditions; here it is nearly 2 hours. As expected, the N- and C-termini (residues 2-4 and 19-21) are relatively unprotected: all have protection factors less than 2. The turn (residues 5-8) appears to have some amide exchange protection, with Nle^{7} having the largest protection factor in this region.

Table 2.2: Pen-1 NH exchange half lives and estimated exchange protection factors (302K, pH 3.5) [1]

Residue	Molday Factor [2]	Est. Predicted $t_{1/2}$	Observed $t_{1/2}$	Protection Factor
S (2)	+0.83	0.50 min	< 5 min	<< 10.00
[3] Pen (3)	-0.40	8.52 min	10 min	1.17
S (4)	+0.23	2.00 min	< 5 min	≤ 2.50
S (5)	+0.67	0.72 min	unknown	unknown
L (6)	-0.28	6.46 min	10 min	1.55
[4] Nle (7)	-0.30	6.76 min	40 min	5.91
[5] $D^{o/-}$ (8)	+0.49/***-0.50***	10.72 min	30 min	≥ 2.80
K (9)	+0.56/***-0.22***	5.63 min	20 min	≥ 3.55
E° (10)	+0.36	1.48 min	15 min	10.14
C (11)	+0.94	0.39 min	unknown	unknown
V (12)	-0.24	5.89 min	8 hrs	81.49
Y (13)	-0.41	8.71 min	4 hrs	27.55
F (14)	-0.19	5.25 min	1 hr	11.43
[3] Pen (15)	-0.64	14.80 min	4 hrs	16.22
H^+ (16)	+0.66	0.74 min	2 hrs	*162.16?*
L (17)	+0.25	1.91 min	20 min	10.47
[5] $D^{o/-}$ (18)	+0.48/***-0.51***	10.97 min	10 min	≥ 0.91
I (19)	-0.13/***-0.91***	27.55 min	5 min	≥ 0.18
I (20)	-0.96	30.92 min	10 min	0.32
W (21)	-0.64	14.80 min	5 min	0.34

[1] Ala is expected to have an exchange $t_{1/2}$ of 3.39min in these conditions. Reference Ala exchange rates (at 280K) used for the calculations were: log k_A = 1.47, log k_B = 10.08, log k_W = -2.26, K_W = 15.65 (Rohl and Baldwin, 1994; Bai et al., 1995).

[2] Base catalyzed reaction dominates at these conditions (Bai et al., 1995).

[3] Val Molday factors used for Pen.

[4] Log average of Ala & Leu Molday factors used for Nle.

[5] Assuming slowest exchange rates for Asp.

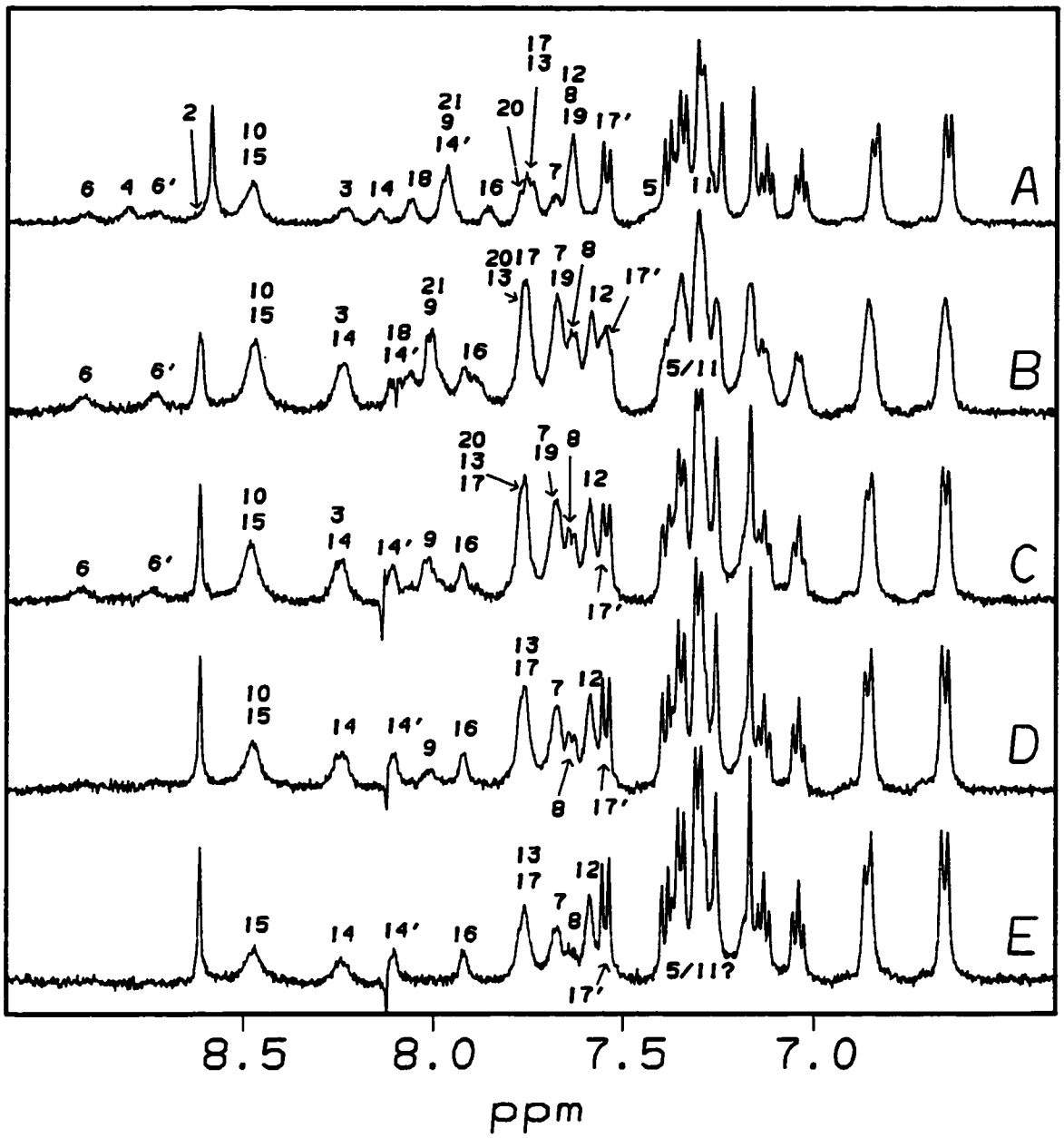


Figure 2.8: Deuterium exchange study for the isotope-labeled Pen-1 in 50% aqueous ethylene glycol at pH 3.5 (302K). Time points are 0 min, H_2O media (A), 0 min D_2O media (B), 10 min (C), 20 min (D) and 40 min (E).

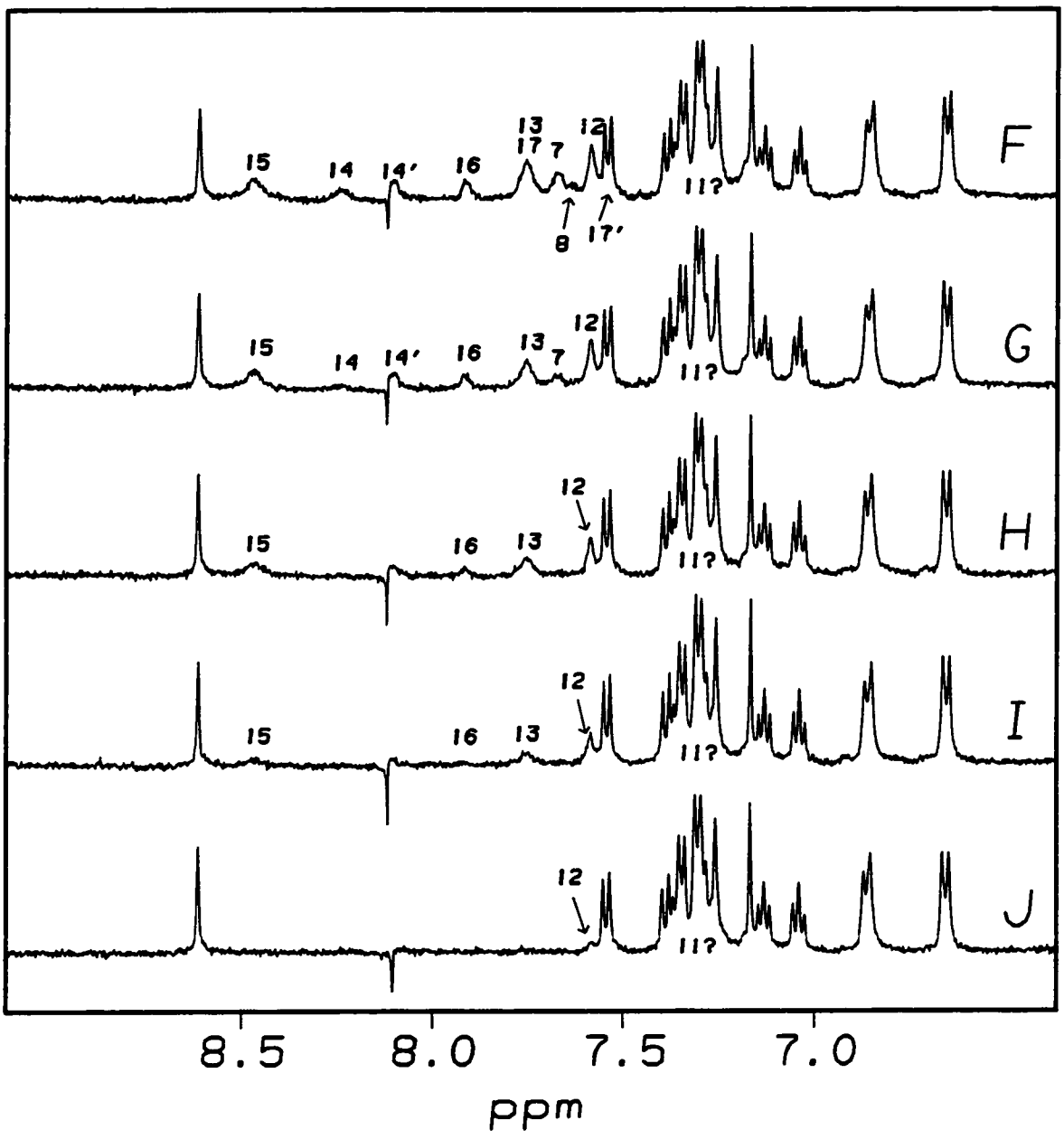


Figure 2.9: Part 2 of the deuterium exchange study for the isotope-labeled Pen-1 (pH 3.5, 302K). Time points are 1 hr (F), 2 hrs (G), 4 hrs (H), 8 hrs (I) and 16 hrs (J).

The amide proton of Val^{12} displays the longest exchange rate (and has the largest protection factor for the helical residues), indicating that it has the strongest (most persistent) hydrogen bond. Incidentally, both its NH and α-methine protons have the most upfield chemical shifts of the residues comprising the helical region. The upfield NH is a bit of a paradox (Wishart et al., 1995; Andersen et al., 1997), since it would suggest that the Val^{12} NH does not have a strong hydrogen bond (in theory, a strong hydrogen-bond should deshield the proton). The upfield chemical shift of the α-methine proton suggests that Val^{12} is the central residue of the helix (see also Section 2.3.4). Evidence of helical fraying is also supported by the hydrogen exchange data: the amides of Leu^{17} and Asp^{18} display shorter exchange half-lives than those of Val^{12} through His^{16} (see also, Figure 2.7). In contrast, residues 12 through 16 of ET-1 displayed much shorter half lives (Chen, 1992), indicating the helix population of ET-1 is lower than that of Pen-1. This is further evidence that the Pen-1 helix is more stable.

2.3.2.3: Information from D_2O TOCSY AND NOESYs.

Deuterated media allowed for a clearly defined set of crosspeak assignments. NOE intensities of the α-methines which were bleached due to solvent suppression, for example 3α and 8α, are present and interpretable. In cases where the β protons are resolvable, stereochemical assignments of the α-carbons were deduced from the NH $\rightarrow$ βHs and αH $\rightarrow$ βHs crosspeaks and placed into the distance constraint sets. Additional stereochemical insights were obtained from the D_2O TOCSY (Figure 2.10). In this case, several intraresidue $\alpha_i \rightarrow \beta_i$ crosspeaks appear to have doublet-like character. These would correspond to a large passive vicinal ($J_{\alpha\beta}$), comparable to the geminal ($J_{\beta\beta}$), coupling indicative of the presence of a gauche configuration between the α and β protons. (Driscoll et al., 1989). Those with the anti configuration would appear as normal TOCSY crosspeaks. Stereochemical assignments for Ser^5, Asp^8, Cys^{11} and Trp^{21} were

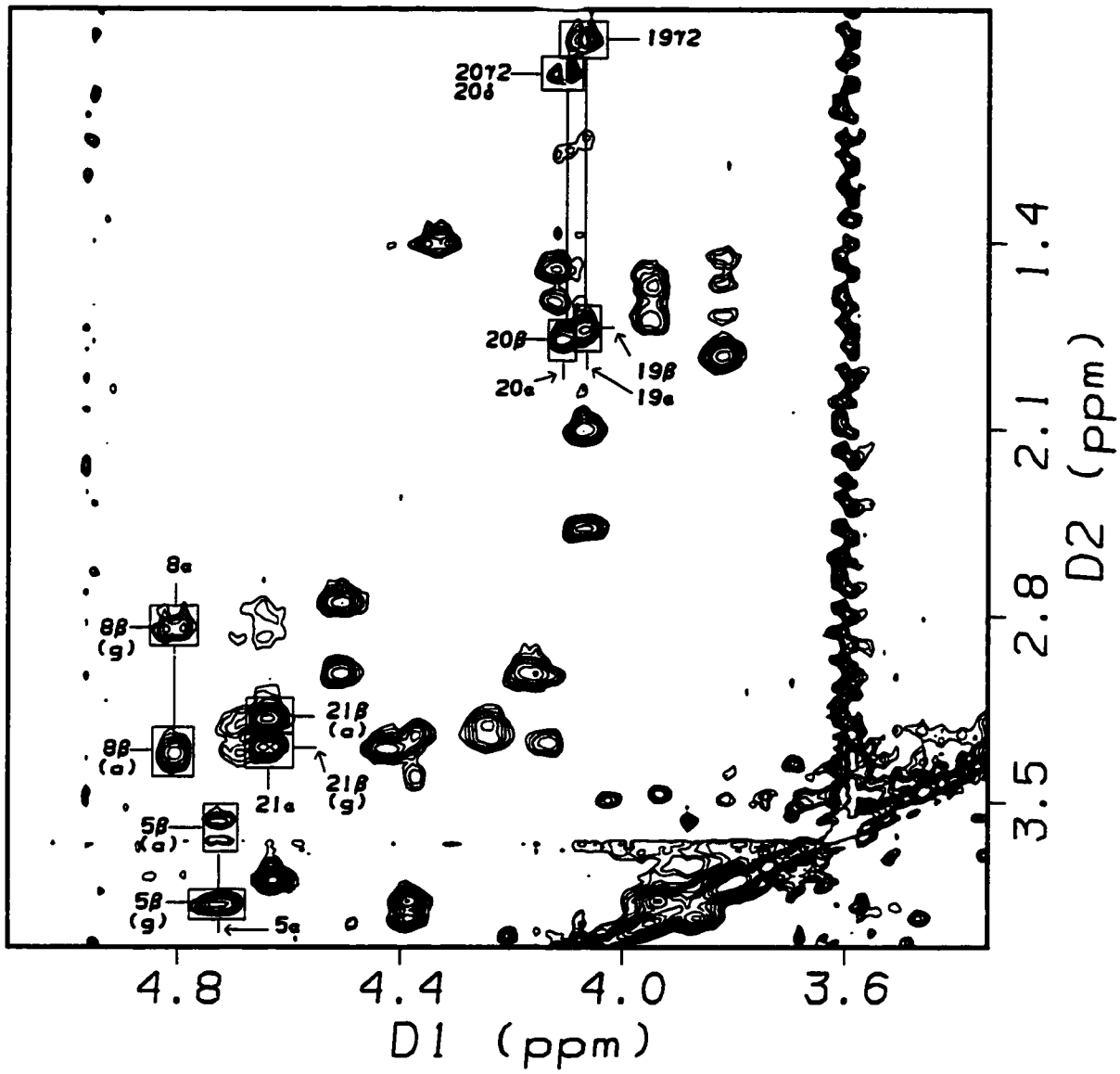


Figure 2.10: The αH → upfield region of the isotope-labeled Pen-1 D_2O TOCSY, acquired at 500MHz.

determined from the D_2O TOCSY in conjunction with the intraresidue αβ and αN NOESY crosspeak intensities. However, these stereospecific assignments, with the exception of Cys^{11}, were not used in the structure refinement since their corresponding side chains probably do not influence the overall conformation of the peptide backbone.

The more useful stereochemical assignments were those corresponding to the Pen residues since these side chains had more influence on the backbone configuration. One of the major problems with the protic media was that the Pen^3 α-methine chemical shift was nearly coincident with the solvent peak. As a result, the Pen^3 α proton was bleached when solvent suppression was applied. The D_2O NOESY (see Figure 2.5), although acquired in slightly different media (5% ethylene glycol and 95% acidic D_2O buffer) than the protic data, clearly shows the 3α crosspeaks, allowing for stereochemical assignments.

2.3.2.4: Fluoroalcohol addition.

HFIP addition (Figures 2.11 and 2.12) affects local structuring. Downfield NH shifts are observed (from 0% to 15% HFIP) in the helical region of the peptide, specifically for Val^{12} through Leu^{17}. The most dramatic deviation occurred at Tyr^{13} (+0.400ppm). These would suggest that the hydrogen bonding interactions within the helix strengthen upon fluoroalcohol addition. The amide chemical shift for Nle^7 also moves downfield (+0.099ppm), indicating that a hydrogen bond in the loop region may also be stabilized by HFIP. Final chemical shift assignments for the 15% HFIP data set are reported in Table 2.3, while Table 2.4 lists chemical shift differences for the backbone amides. The upfield movement of the His^{16} ε1 peak with increasing fluoroalcohol volumes (ranging from 8.587ppm at 0% HFIP to 8.444ppm at 15% HFIP) suggests that the solution becomes less acidic as HFIP is added.

Table 2.3: Chemical shift assignments for isotope-labeled [Pen^{3,15}-Nle^{7}]-ET-1 in 15% HFIP, 42.5% aqueous ethylene glycol at 292K.

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Cys 1	exch.	4.382	3.356, 3.356	
Ser 2	8.735 [-7.152]	4.708	3.871, 3.871	
Pen 3	8.211 [-6.337]	5.086		γ,γ′ 1.722, 1.246
Ser 4	9.015 [-4.517]	4.497	4.019, 3.979	
Ser 5	7.350 [-0.468]	?	4.002, 3.619	
Leu 6 *	8.854 [-8.076] J = 94Hz	4.044	1.705, 1.643	γ 1.772, δ,δ′ 0.963, 0.861
Nle 7 ^	7.849[-2.653]	4.097	n/a	γ,γ′ n/a, δ,δ′ n/a, ε,ε′ n/a
Asp 8	7.710 [-3.094]	4.789	3.427, 2.704	
Lys 9	7.876 [-1.852]	3.892	1.878, 1.878	γ,γ′ 1.574, 1.511, δ,δ′ 1.750, 1.750
				ε,ε′ 3.895, 3.895
Glu 10	8.524 [-5.444]	4.051	2.164, 2.164	γ,γ′ 2.505, 2.505
Cys 11	7.270 [-0.075]	4.207	3.427, 3.160	
Val 12 ^	7.794 [-4.185]	n/a	n/a	γ,γ′ n/a
Tyr 13	8.271 [-4.629]	4.189	3.146, 3.146	δ,δ′ 6.911, ε,ε′ 6.728
Phe 14 *	8.356 [-4.884] J = 93Hz	4.305	3.347, 3.347	δ,δ′ 7.353, ε,ε′ 7.379, ζ ~7.379
Pen 15	8.939 [-7.700]	4.695		γ,γ′ 1.679, 1.492
His 16	8.092 [-0.678]	4.334	3.358, 3.358	δ 7.240, ε 8.444
Leu 17 *	7.799 [-2.278] J = 94Hz	4.069	1.717, 1.613	γ 1.520?, δ,δ′ 0.806, 0.806?
Asp 18	8.006 [-5.459]	4.507	3.026, 2.792	
Ile 19	7.696[-3.375]	4.081	1.861	γ1,γ1′ 1.534, 1.150
				γ2 0.794, δ 0.794
Ile 20	7.659 [-5.602]	4.146	1.851	γ1,γ1′ 1.428, 1.140
				γ2 0.789, δ 0.789
Trp 21	7.992 [-7.396]	4.691	3.374, 3.304	δ 7.230, ε1 ?, ε3 7.644
				ζ2 7.446, ζ3 7.126, η2 7.211

*** ^{15}N, ^ perdeuterated side chains**

Table 2.4: Amide chemical shifts for the HFIP titration of the isotope-labeled [$Pen^{3,15}$-Nle^{7}]-ET-1 at 305K.

Residue	0% HFIP	15% HFIP	Δδ (15% - 0%)
Cys 1	exchanged	exchanged	n/a
Ser 2	8.595	8.649	+0.054
Pen 3	8.240	8.118	-0.122
Ser 4	8.807	8.912	+0.105
Ser 5	7.439	7.343	-0.096
Leu 6 *	8.827	8.737	-0.090
Nle 7	7.683	7.782	***+0.099***
Asp 8	7.639	7.654	+0.015
Lys 9	7.973	7.863	-0.110
Glu 10	8.481	8.454	-0.027
Cys 11	7.313?	7.290?	-0.023
Val 12	7.639	7.727	***+0.088***
Tyr 13	7.741	8.141	***+0.400***
Phe 14 *	8.061	8.279	***+0.218***
Pen 15	8.492	8.825	***+0.333***
His 16	7.859	8.039	***+0.180***
Leu 17 *	7.659	7.759	***+0.100***
Asp 18	8.063	7.971	-0.092
Ile 19	7.639	7.654	+0.015
Ile 20	7.650	7.588	-0.062
Trp 21	7.973	7.892	-0.081

Note: $\Delta\delta > 0$ and $\Delta\delta < 0$ indicate downfield and upfield chemical shift deviations with increasing HFIP concentrations, respectively. Units are in ppm.

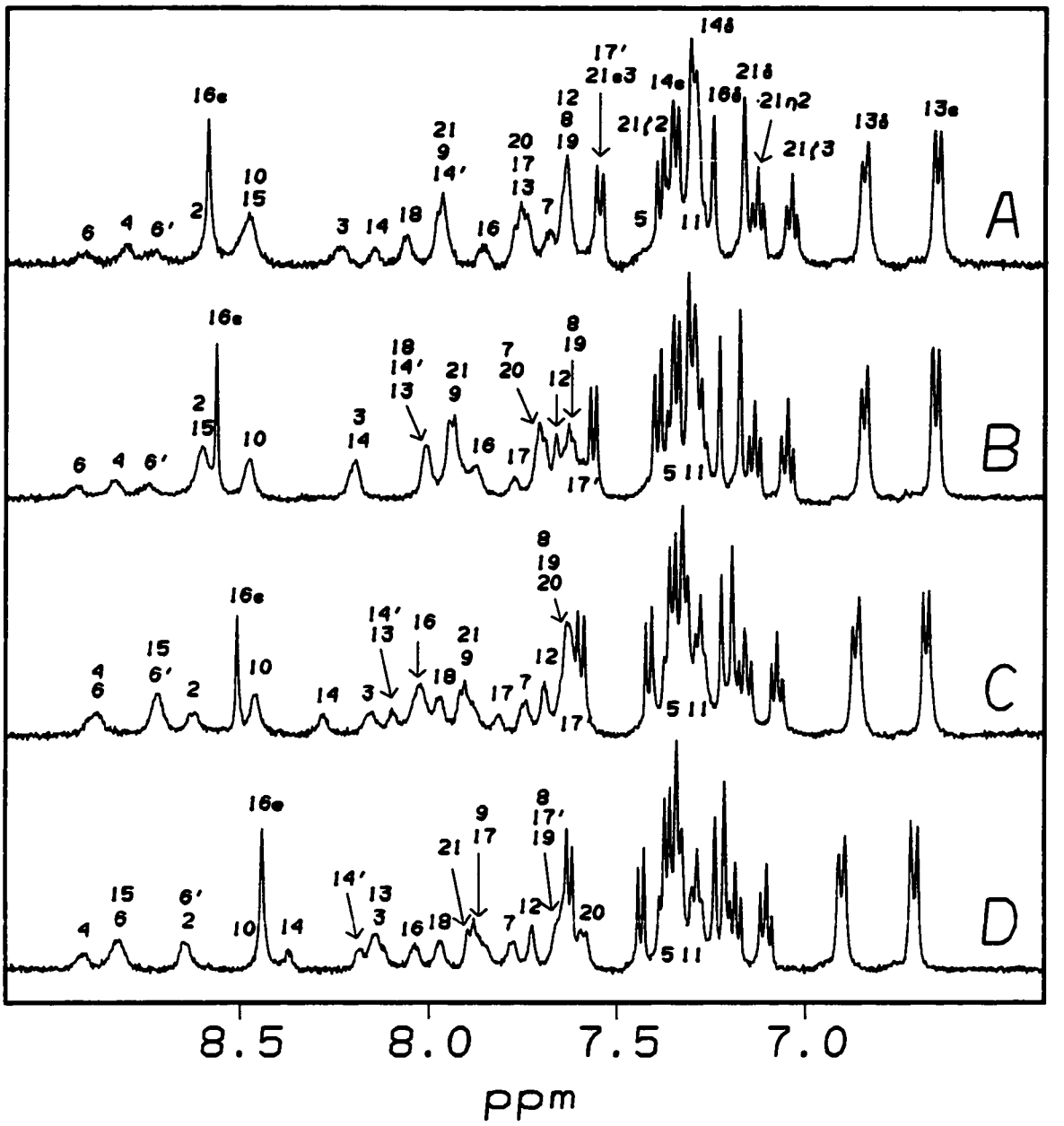


Figure 2.11: Amide NH region of the isotope-labeled Pen-1 HFIP titration, starting from 50% aqueous ethylene glycol, pH 3.5. Points in the titration are 0% (A), 3% (B), 8% (C), and 15% HFIP (D) by volume.

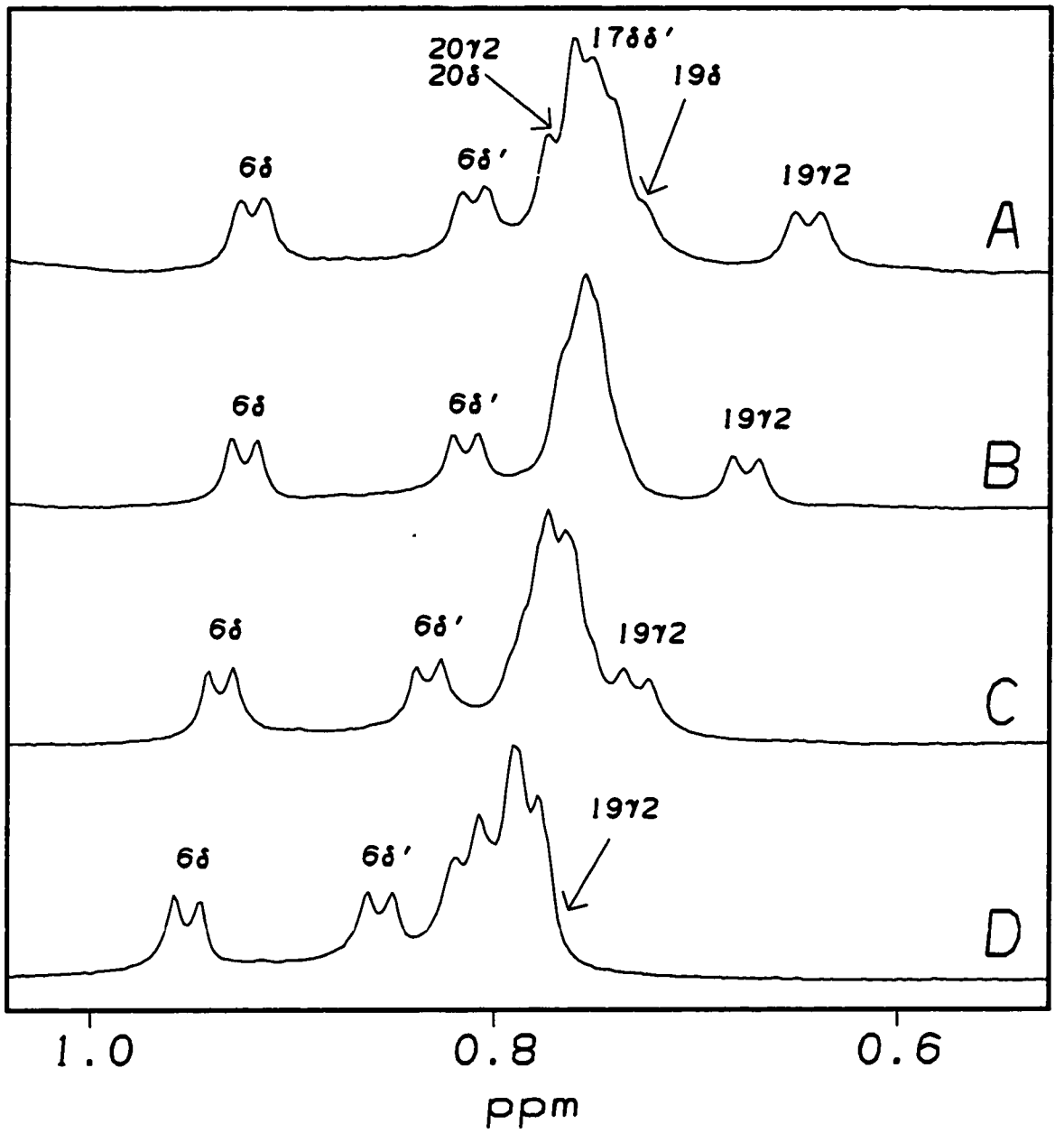


Figure 2.12: Methyl region of the isotope-labeled Pen-1 HFIP titration, starting from 50% aqueous ethylene glycol, pH 3.5. Points in the titration are 0% (A), 3% (B), 8% (C), and 15% HFIP (D) by volume.

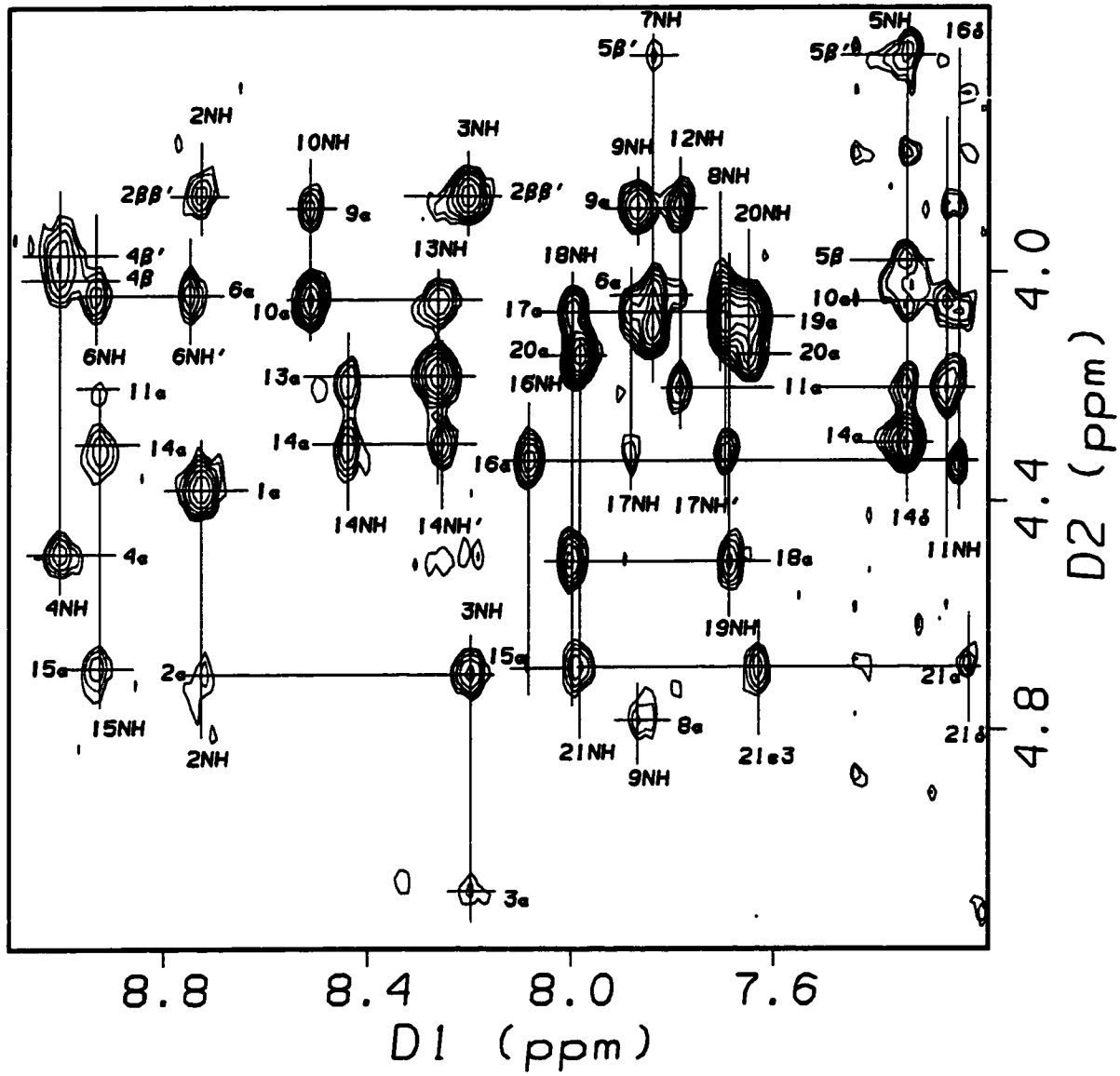


Figure 2.13: The amide → αH crosspeak region for the isotope-labeled Pen-1 NOESY, acquired at 500MHz in aqueous 15% HFIP/42.5% ethylene glycol.

The NOESY spectrum (Figure 2.13) also shows that helix stability is enhanced. Medium-range crosspeaks are more intense at 15% HFIP than in the aqueous glycol mixture alone. A number of stronger $i \rightarrow i+3$ NOEs are evident in this case, ranging from the 9α → 12NH (very strong) to the 15α → 18NH (medium-weak) crosspeaks. Also present in the spectrum are medium-to-weak $\alpha_i \rightarrow N_{i+4}$ NOEs: one observed between 9α and 13NH, another between 11α and 15NH. Strong 10α → 13δδ′ and 11α → 14δδ′, as well as 10α → 14δδ′ crosspeaks are also observed in the NOESY.

The C-terminal residues are also affected by the HFIP titration, as shown by the upfield region of the spectrum (Figure 2.12). The most obvious case, the Ile^{19} β-methyl (γ2) group, shows a diminution of the upfield shift deviation with increasing HFIP concentrations. At 0% HFIP, the 19γ2 peak has a chemical shift of 0.649ppm, while at 15% HFIP, it has a shift of 0.794ppm (a change of +0.145ppm). The upfield shifts in the more aqueous case are most likely due to a hydrophobic interaction that positions the Ile^{19} aliphatic sidechain in the shielding cone of the Trp^{21} indole ring. In Pen-1, this clustering is strong in 50% aqueous glycol, but collapses when the solution becomes less polar.

2.3.3: Spectral Assignment of Pen-2.

2.3.3.1: Differences between Pen-1 and Pen-2.

Although significant solubility problems occurred when the recovered Pen-2 sample was reconstituted in aqueous 5% acetic acid/40% ethylene glycol, a set of NOESY and TOCSY spectra were acquired and analyzed. Table 2.5 lists the final chemical shift assignments for Pen-2 in this solvent system. Assignment of the Pen-2 NOESY spectra was similar to that of Pen-1: the bicyclic core displayed the same connectivity patterns and relative intensities for both peptides.

Table 2.5: Chemical shift assignments for [Pen^{3,15}-Nle^{7}-$^{(NMe)}$Ile^{20}]-ET-1 in 5% HOAc, 40% ethylene glycol and 55% H_2O at 290K.

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Cys 1	exch.	4.388	unknown	
Ser 2	8.651 [-2.974]	4.614	3.814, 3.814	
Pen 3	8.325 [-5.301]	bleached		γ,γ′ 1.596, 1.217
Ser 4	8.898 [-6.857]	4.423	3.932, 3.874	
Ser 5	7.442? [?]	4.871	3.884, 3.618	
Leu 6	8.933 [-6.327]	3.980	1.677, 1.575	γ 1.718, δ,δ′ 0.923, 0.811
Nle 7	7.741 [-2.189]	4.106	1.866, 1.610	γ,γ′ 1.292, 1.273, δ,δ′ 0.879, 0.879
				ε ?
Asp 8	7.725 [-1.770]	4.796	3.308, 2.777	
Lys 9	8.009 [-2.026]	3.842	1.834, 1.834	γ,γ′ 1.552, 0.887, δ,δ′ 3.018, 3.018
				ε,ε′ 1.693, 1.693
Glu 10	8.582 [-7.449]	4.101	2.140, 2.140	γ,γ′ 2.489, 2.489
Cys 11	7.262 [?]	4.201	3.301, 3.019	
Val 12	7.716 [-1.847]	3.350	2.054	γ,γ′ 1.024, 0.820
Tyr 13	7.865 [-0.235]	4.200	3.015, 3.015	δ,δ′ 6.883, ε,ε′ 6.665
Phe 14	8.099 [-1.454]	4.300	3.303, 3.204	δ,δ′ 7.319, ε,ε′ 7.352, ζ ~7.352
Pen 15	8.567 [-8.378]	4.634		γ,γ′ 1.582, 1.398
His 16	7.893 [-4.929]	4.483	3.312, 3.312	δ 7.317, ε 8.604
Leu 17	7.630 [+0.663]	4.215	1.598, 1.403	γ 1.510, δ,δ′ 0.771, 0.771
Asp 18	8.140 [-2.847]	4.533	2.994, 2.643	
Ile 19	7.724 [-2.372]	4.461	1.558	γ1,γ1′ 1.389, 0.924
				γ2 0.290, δ 0.698
$^{(NMe)}$Ile 20	n/a	4.699	2.003	γ1,γ1′ 1.219, 0.918, γ2 0.867
				δ 0.779, NMe 2.975
Trp 21	8.106 [-6.806]	4.524	3.300, 3.105	δ 7.165, ε1 10.062, ε3 7.553
				ζ2 7.391, ζ3 7.062, η2 7.150

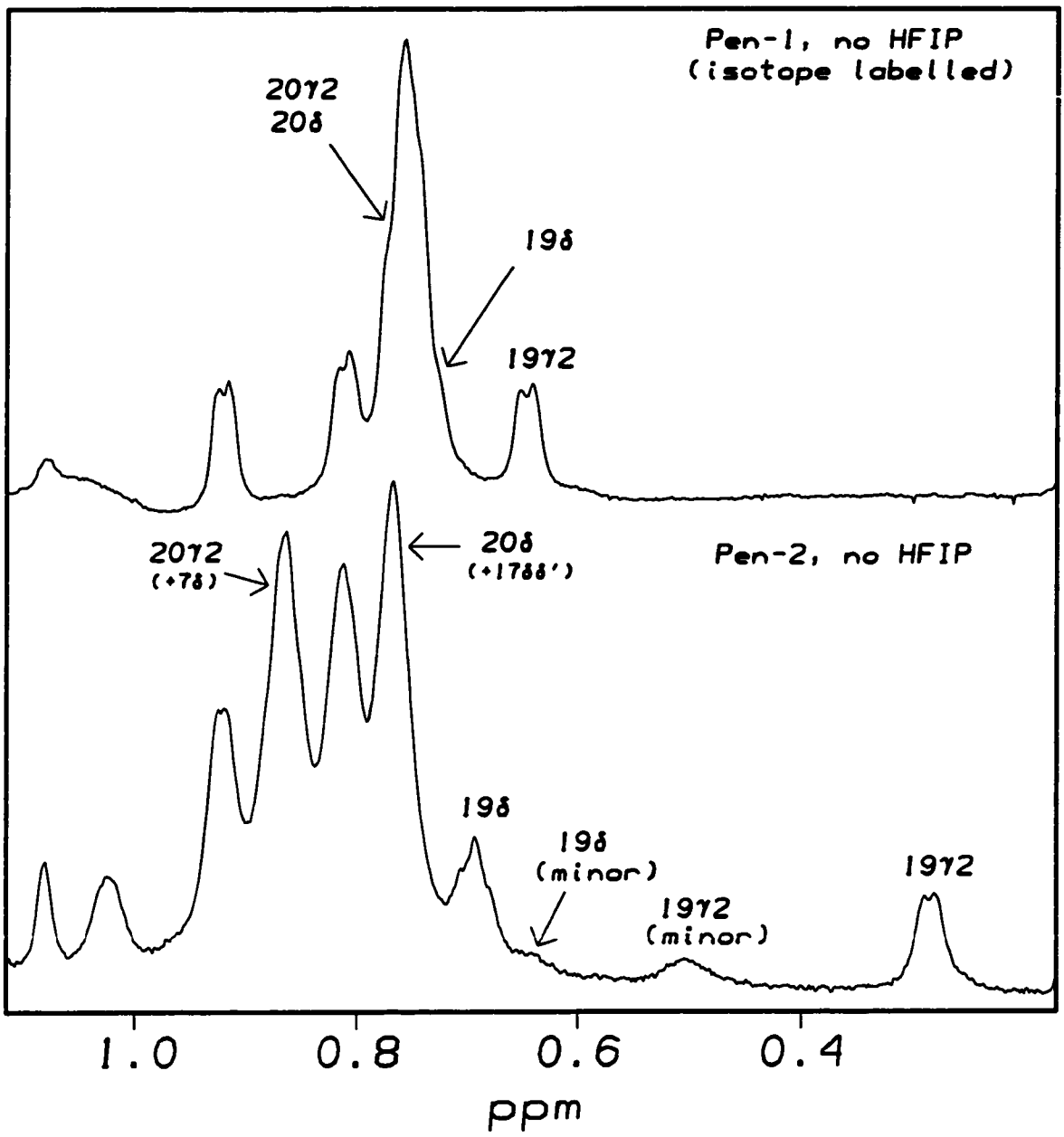


Figure 2.14: The methyl region of the isotope-labeled Pen-1 (top) and Pen-2 (bottom), both in 50% aqueous ethylene glycol at pH 3.5.

As expected, notable differences are observed for the C-terminal residues. This is readily apparent in Figure 2.14, which displays the methyl region of the isotope-labeled Pen-1 and the Pen-2 proton spectra. The most obvious difference between the two peptides is the Ile^{19} β-methyl (19γ2) peak, which has a chemical shift of 0.290ppm, an *upfield* deviation of 0.359ppm with respect to Pen-1. This upfield methyl shift would suggest the hydrophobic clustering observed in the ET-1 and Pen-1 is even more dominant in the case of Pen-2. The β-methyl group for residue 20 moves in the opposite direction. N-methylation shifts the 20γ2 resonance from 0.763ppm (Pen-1) to 0.867ppm (Pen-2).

Dramatic downfield chemical shift deviations of the 19 and 20 α-methine protons upon N-methylation of Ile^{20} are readily observed. In the Pen-1 analog, 19αH has a chemical shift of 4.076ppm, while in Pen-2, it is 4.461ppm, a change of +0.385ppm. The difference between the 20α methines is even more dramatic: 4.115ppm for Pen-1 and 4.699ppm for Pen-2, a difference of +0.584ppm. The β-methine chemical shifts for these two residues show opposite tendencies: 19βH moves upfield (1.734ppm to 1.558ppm) while 20βH moves downfield (1.755ppm to 2.003ppm). The other aliphatic sidechain protons (γ1, γ1′, δ) of residues 19 and 20 don't display dramatic shift differences between Pen-1 and Pen-2.

An artifact of N-methylation is the emergence of a minor conformer (Figure 2.14), which, based on signal intensities, represents 15% of the population. The minor conformer can be explained by the N-methyl group at Ile^{20}, which has the same effect on the peptide backbone as a proline (Figure 2.15). Prolines are known to lower the energy difference between the cis/trans peptide bond isomers. In most peptides and proteins without prolines, the peptide bond favors the trans conformation ($\omega = 180°$) by a factor of

1000. When prolines are incorporated into the sequence, the cis isomer (ω = 0°) is observed approximately 20% of the time.

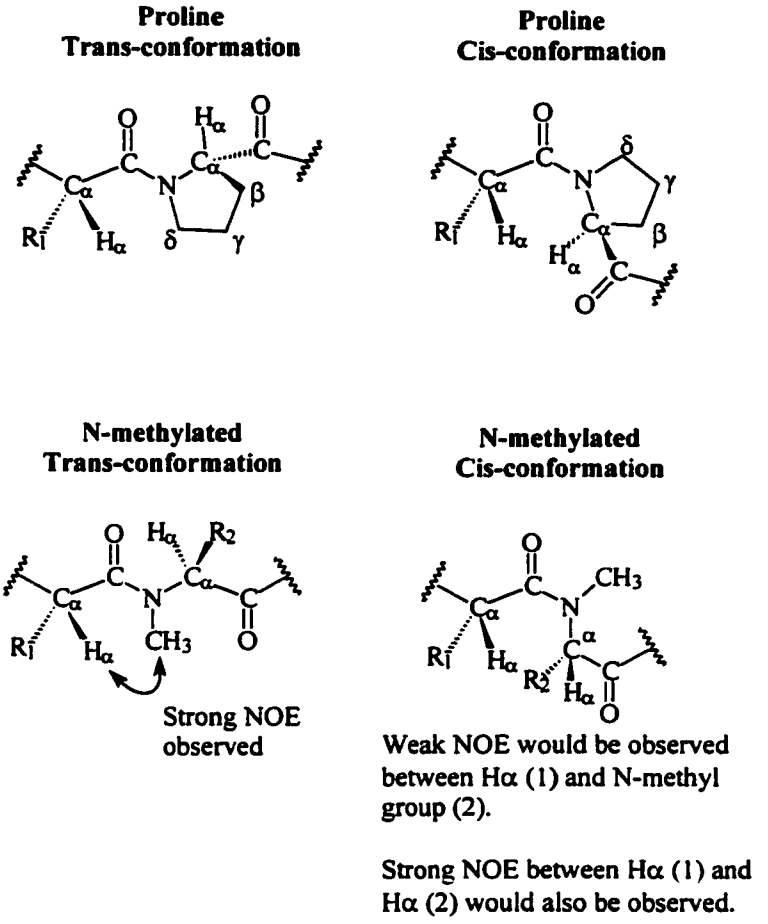


Figure 2.15: Proline and N-methyl effects on the peptide backbone.

2.3.3.2: HFIP titration of Pen-2.

The HFIP titration study of Pen-2 (Figures 2.16 and 2.17) shows similar effects in the bicyclic core as for Pen-1. Final chemical shift assignments for the peptide in 15% HFIP are listed in Table 2.6. Fluoroalcohol effects on the Pen-2 amide chemical shifts for residues 13 through 17 (Table 2.7) are observed, suggesting that the hydrogen

bonding network is stabilized within the helical core. As with Pen-1, the Pen-2 Tyr^{13} amide displays the largest downfield shift deviation (+0.369ppm from 0% → 15% HFIP) of the residues in the helical region. The methyl region of the proton spectrum also illustrates the increasing helicity: the γ-methyl peaks of Val^{12}, which were not observed in the isotope-labeled Pen-1 spectra, move downfield upon increasing HFIP concentrations.

Evidence of the more rigid C-terminal hydrophobic cluster is observed for the HFIP titration study. The Ile^{19} β-methyl shifted 0.145ppm downfield when 15% HFIP was added to the solution in the case of Pen-1. Although a larger upfield CSD is evident, the same titration performed on the Pen-2 sample resulted in only a 0.054ppm downfield chemical shift deviation. Assuming the reference value for the Ile β-methyl group is 0.95ppm (Wüthrich, 1986), the CSD values for Pen-1 and Pen-2 in 15% aqueous HFIP are -0.156 (a 48% loss of the original CSD) and -0.606ppm (an 8% loss), respectively. At the higher fluoroalcohol levels, the hydrophobic interactions between the Ile^{19} and the Trp^{21} sidechains are expected to be less favorable since HFIP addition makes the media less lipophobic. This is reflected in the deviation observed for Pen-1. However, the deviation observed for Pen-2 is smaller than expected, suggesting that N-methylation of Ile^{20}, possibly through steric interactions, increases the rigidity of the Pen-2 C-terminal backbone. As a result, the Ile^{19} β-methyl group is less likely to move out of the shielding cone of the Trp^{21} indole ring.

The NOESY spectrum (Figures 2.18 and 2.19) displays much stronger crosspeak intensities when HFIP is added to Pen-2. However, this is primarily an effect of the peptide's higher solubility in a less polar medium. At 0% HFIP, the peptide concentration was low enough so as to make interpretation of crosspeak volumes difficult due to insufficient signal-to-noise resolution. However, at 15% HFIP, the peptide concentration was high enough to quantitate crosspeaks. As in the Pen-1 case, Pen-2 in

Table 2.6: Chemical shift assignments for [Pen^{3,15}-Nle^{7}-$^{(NMe)}$Ile^{20}]-ET-1 in 15% HFIP, 4.3% HOAc, 34% ethylene glycol and 46.7% H_2O at 290K.

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Cys 1	exch.	4.422	not observed	
Ser 2	8.685 [-4.490]	4.736	3.879, 3.879	
Pen 3	8.229 [-5.031]	5.035		γ,γ′ 1.712, 1.255
Ser 4	8.973 [-5.704]	4.514	4.020, 3.965	
Ser 5	7.367 [?]	4.818	4.000, 3.698	
Leu 6	8.839 [-6.194]	4.048	1.704, 1.642	γ 1.772, δ,δ′ 0.958, 0.857
Nle 7	7.822 [-3.378]	4.136	1.905, 1.669	γ,γ′ 1.340, 1.333, δ,δ′ 0.909, 0.909 ε,ε′ ?, ?
Asp 8	7.705 [-2.633]	4.857	3.442, 2.755	
Lys 9	7.904 [+1.153]	3.898	1.871, 1.871	γ,γ′ 3.053, 3.055, δ,δ′ 1.751, 1.751 ε,ε′ 3.047, 3.047
Glu 10	8.519 [-4.469]	4.072	2.170, 2.170	γ,γ′ 2.509, 2.509
Cys 11	7.253 [?]	4.232	3.427, 3.154	
Val 12	7.770 [-3.633]	3.481	2.247	γ,γ′ 1.117, 0.924
Tyr 13	8.143[-7.776]	4.213	3.118, 3.118	δ,δ′ 6.936, ε,ε′ 6.727
Phe 14	8.304 [-3.526]	4.379	3.369, 3.369	δ,δ′ 7.350, ε,ε′ 7.294, ζ ~7.294
Pen 15	8.851 [-8.296]	4.682		γ,γ′ 1.681, 1.475
His 16	8.055 [-2.408]	4.455	3.400, 3.352	δ 7.322, ε 8.478
Leu 17	7.683 [+0.857]	4.203	1.754, 1.539	γ 1.641, δ,δ′ 0.841, 0.792
Asp 18	7.970 [-0.352]	4.601	3.058, 2.708	
Ile 19	7.647 [-2.469]	4.487	1.586	γ1,γ1′ 1.432, 1.003 γ2 0.344, δ 0.745
$^{(NMe)}$Ile 20	n/a	4.702	2.061	γ1,γ1′ 1.310, 1.310, γ2 0.898 δ 0.745, NMe 3.016
Trp 21	7.899 [-9.954]	4.633	3.348, 3.196	δ 7.226, ε1 9.874, ε3 7.611 ζ2 7.422, ζ3 7.109, η2 7.189

Table 2.7: Amide chemical shifts for the HFIP titration of [Pen^{3,15}-Nle^{7}-$^{(NMe)}$Ile^{20}]-ET-1 at 305K.

Residue	0% HFIP	15% HFIP	Δδ (15% - 0%)
Cys 1	exchanged	exchanged	n/a
Ser 2	8.647	8.704	+0.057
Pen 3	8.323	8.229	-0.094
Ser 4	8.880	8.992	+0.112
Ser 5	unknown	unknown	unknown
Leu 6	8.921	8.832	-0.089
Nle 7	7.734	7.835	***+0.101***
Asp 8	7.717	7.716	-0.001
Lys 9	7.994	7.900	-0.094
Glu 10	8.570	8.505	-0.065
Cys 11	7.268	7.241	-0.027
Val 12	7.705	7.778	***+0.073***
Tyr 13	7.860	8.229	***+0.369***
Phe 14	8.087	8.316	***+0.229***
Pen 15	8.556	8.832	***+0.276***
His 16	7.871	8.063	***+0.192***
Leu 17	7.615	7.668	***+0.053***
Asp 18	8.131	7.963	-0.168
Ile 19	7.702	7.659	-0.043
$^{(NMe)}$Ile 20	n/a	n/a	n/a
Trp 21	8.093	7.990	-0.103

Note: $\Delta\delta > 0$ and $\Delta\delta < 0$ indicate downfield and upfield chemical shift deviations with increasing HFIP concentrations, respectively. Units are in ppm.

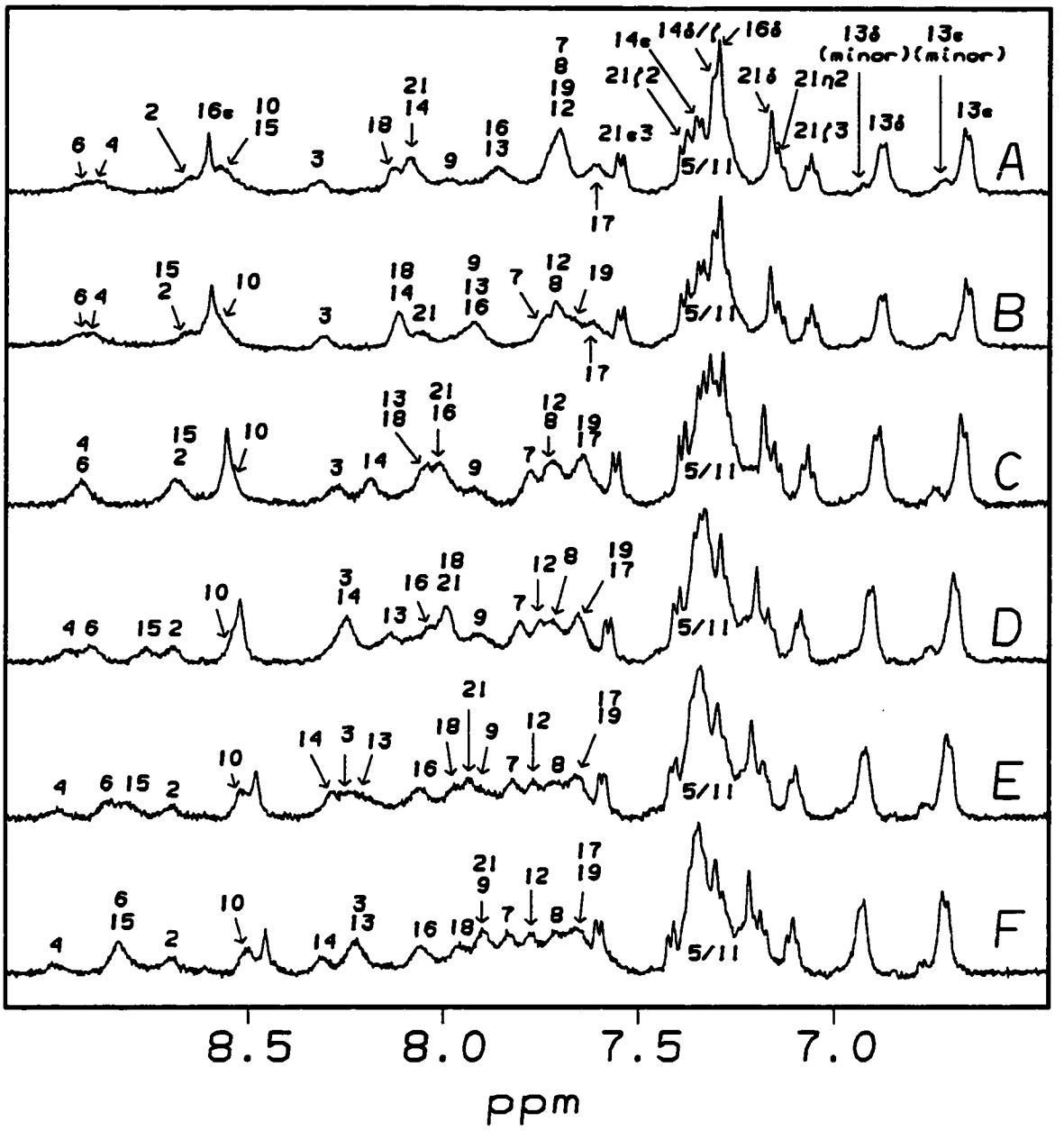


Figure 2.16: The amide NH region of the Pen-2 HFIP titration, starting from 5% aqueous acetic acid/40% ethylene glycol. Points in the titration are: 0% (A), 2% (B), 5% (C), 8% (D), 12% (E) and 15% (F) by volume.

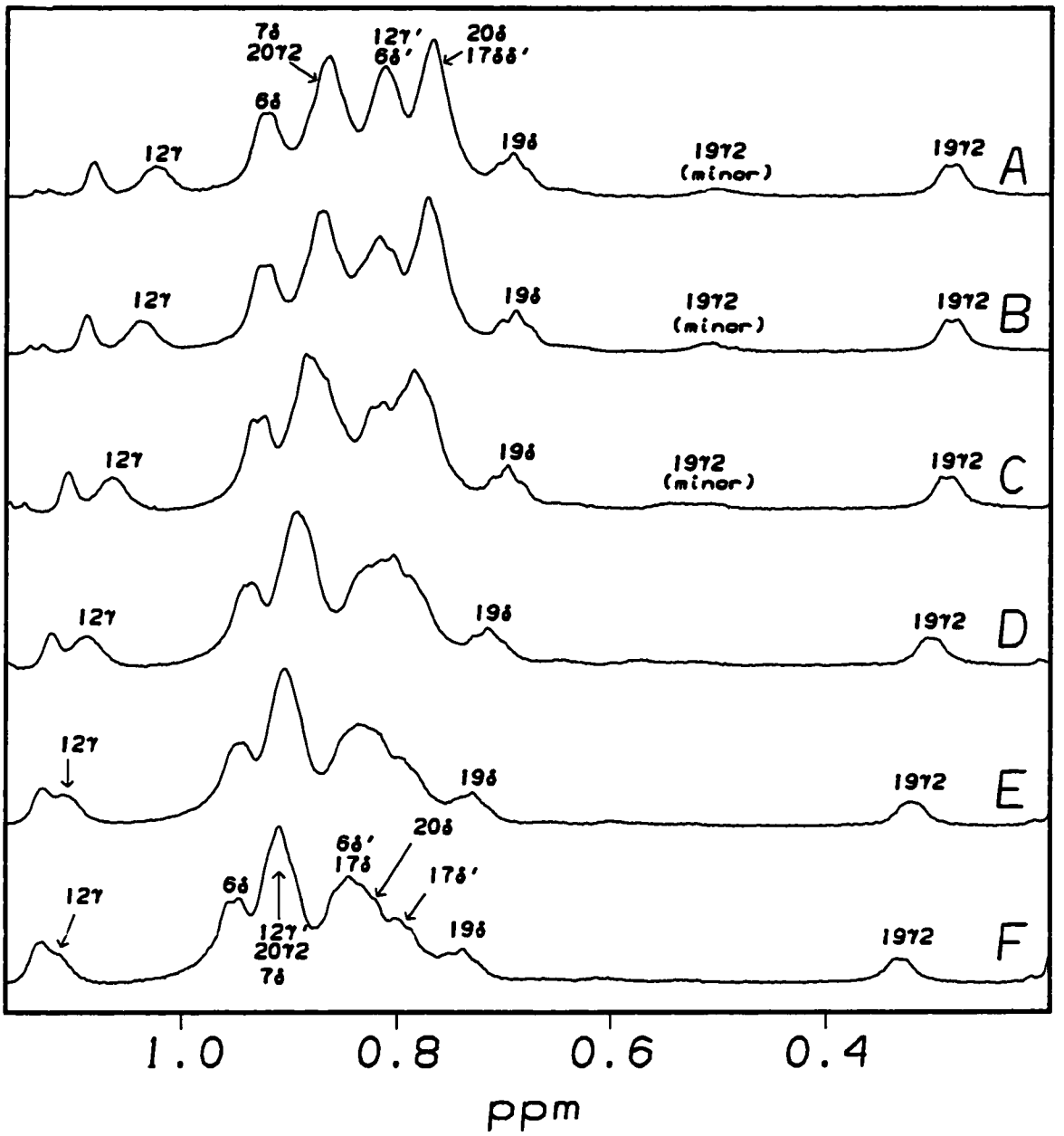


Figure 2.17: The methyl region of the Pen-2 HFIP titration, starting from 5% aqueous acetic acid/40% ethylene glycol. Points in the titration are: 0% (A), 2% (B), 5% (C), 8% (D), 12% (E) and 15% (F) by volume.

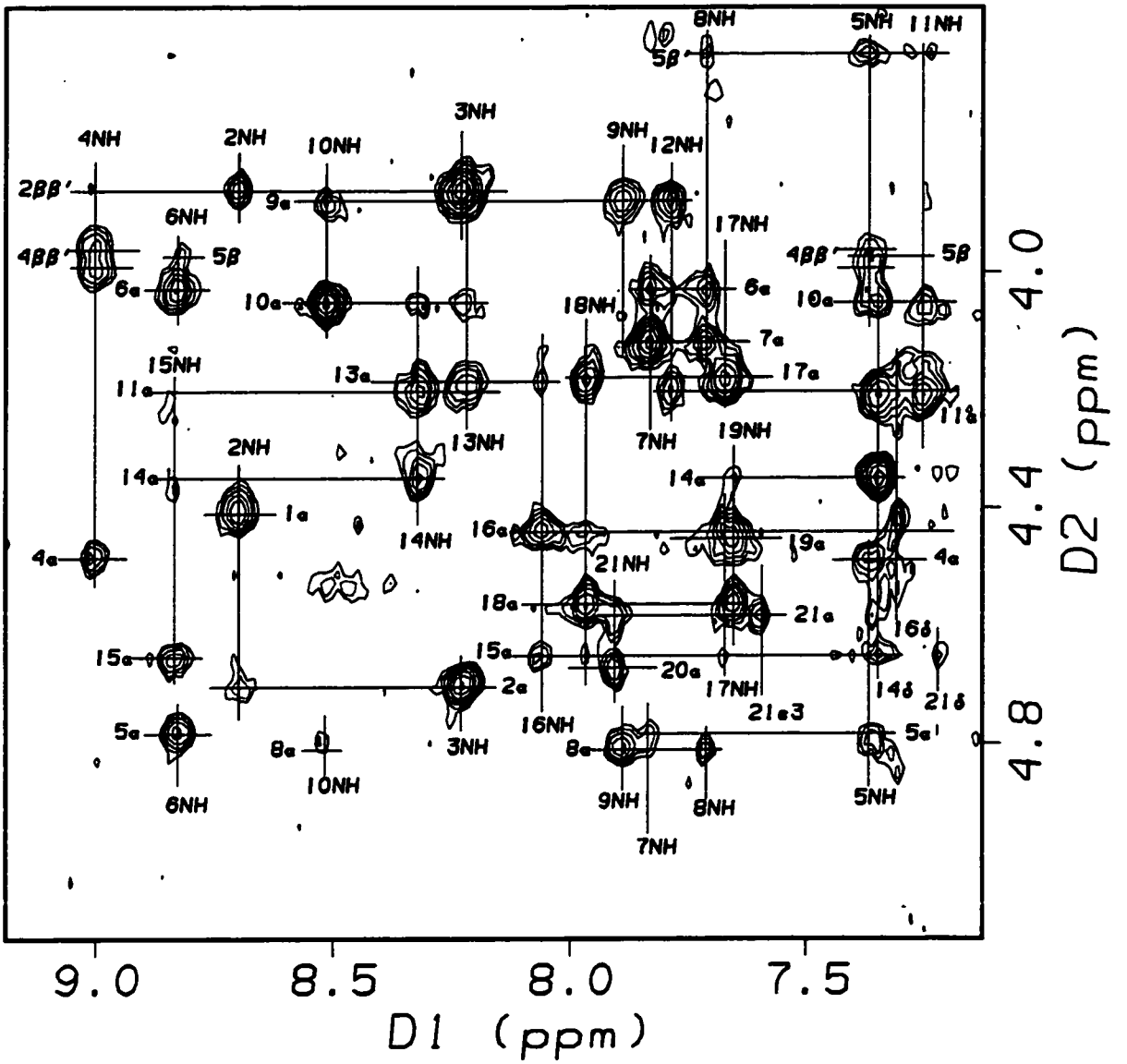


Figure 2.18: The amide NH → αH crosspeak region of the Pen-2 NOESY, acquired at 500MHz in aqueous 15% HFIP/4.3% acetic acid/34% ethylene glycol.

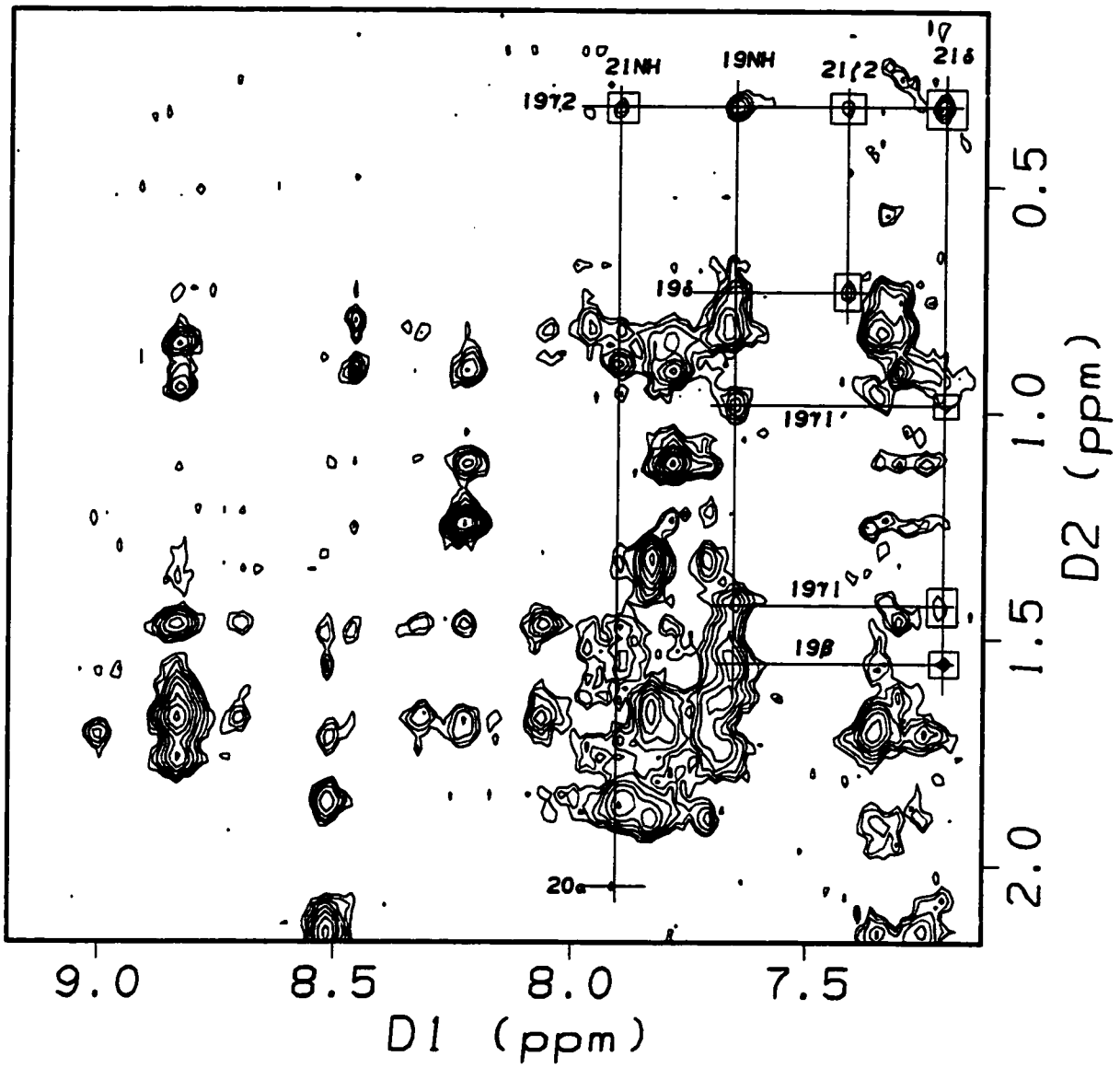


Figure 2.19: The amide NH → upfield region of the Pen-2 NOESY, acquired at 500MHz in aqueous 15% HFIP/4.3% acetic acid/34% ethylene glycol. Key $i \rightarrow i+2$ NOEs between Ile^{19} and Trp^{21} are indicated.

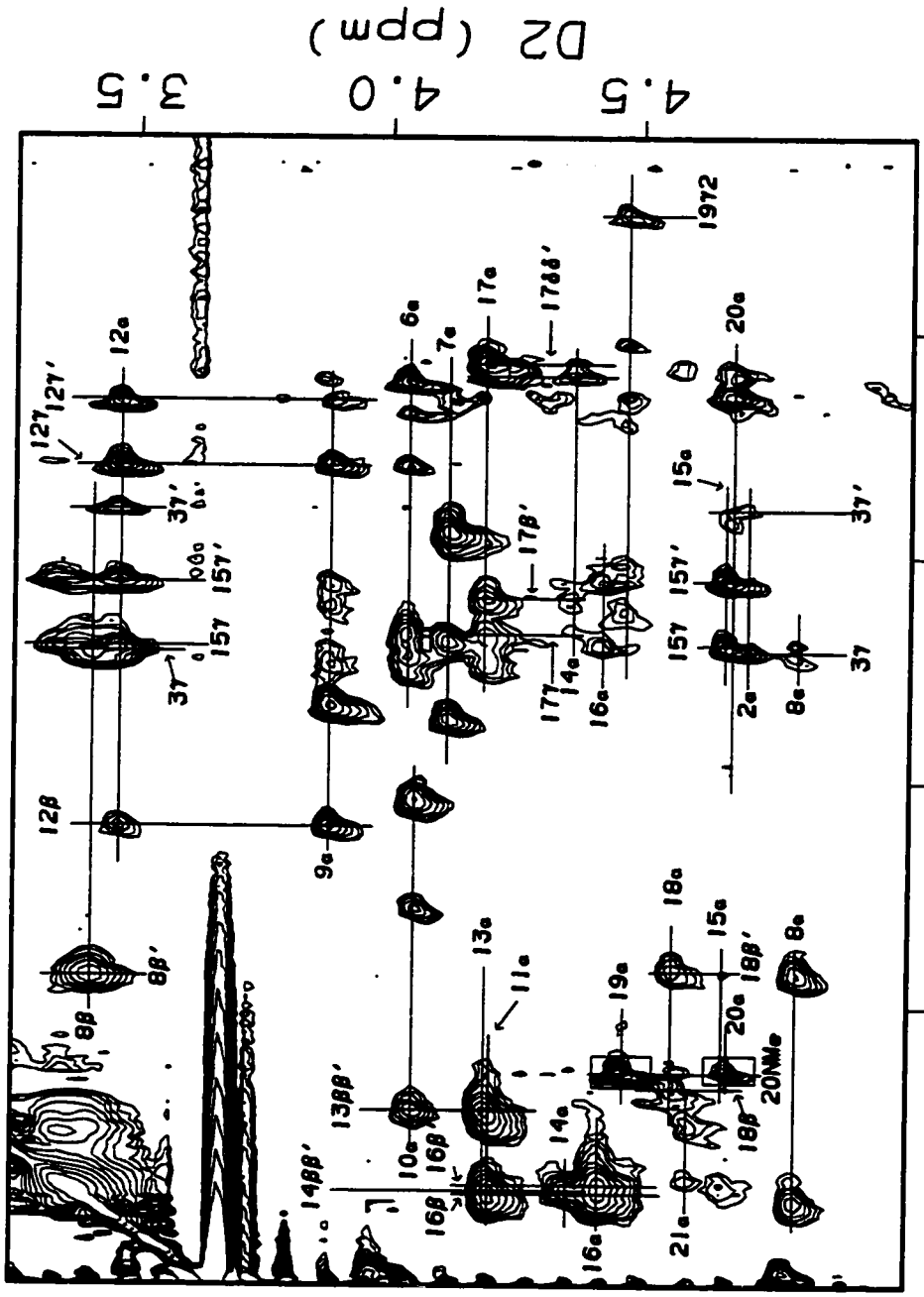


Figure 2.20: The αH → upfield crosspeak region for the Pen-2 D_2O NOESY, acquired at 500MHz in aqueous 15% HFIP/4.3% acetic acid/34% ethylene glycol. The 19α → 20NMe crosspeak is indicated.

15% HFIP displays some medium-to-weak $i \rightarrow i+2$ crosspeaks in the loop region ($Ser^5 \rightarrow Asp^8$) and strong $i \rightarrow i+3$ interactions for Lys^9 through Asp^{18}.

Stronger $i \rightarrow i+2$ interactions are observed between Ile^{19} and Trp^{21}, mostly between the aliphatic sidechain protons the isoleucine and the indole ring protons of the tryptophan. For example, 19γ2 → 21δ and 19δ → 21ζ2 crosspeaks are present in the NOESY spectrum (Figure 2.19). Weaker interactions between 19γ2 and 21NH, as well as 19β and 21δ are also observed, indicating the close proximity between the two residues in Pen-2. Two prominent peaks in the αH → upfield region of the NOESY spectrum (Figure 2.20) verify that the conformation of the Ile^{19}-$^{(NMe)}Ile^{20}$ peptide bond. If the bond adopts the *trans* configuration, then a strong NOE between 19α and the 20NMe group should be present. However, if the bond had the *cis* conformation, then the 19α → 20NMe crosspeak should be weak, while a strong crosspeak between 19αH and 20αH would be present. Figure 2.20 clearly shows a strong interaction between 19α and 20NMe group. Not shown in this figure is the clear absence of any strong NOEs between the α-methine protons of residues 19 and 20.

2.3.4: Comparison of α-Methine Chemical Shift Deviations.

ET-1 is, as expected, structurally similar to both Pen-1 and Pen-2. A quick survey of the αH-CSD histogram (Figure 2.21) shows positive CSD values for residues 1 through 3, indicating that all three peptides contain a relatively extended region at the N-terminus. A number of alternating positive and negative CSDs for Ser^4/Ser^5/Leu^6 and Nle^7/Asp^8/Lys^9 implies the presence of a turn encompassing residues 5 → 8. The C-terminus ($Ile^{19} \rightarrow Trp^{21}$) appears to be disordered for all three peptides. In the case of Pen-2, new reference values were used for Ile^{19} and $^{(NMe)}Ile^{20}$ which reflect the N-methylation effect as observed in smaller peptides (see Section 3.6.4).

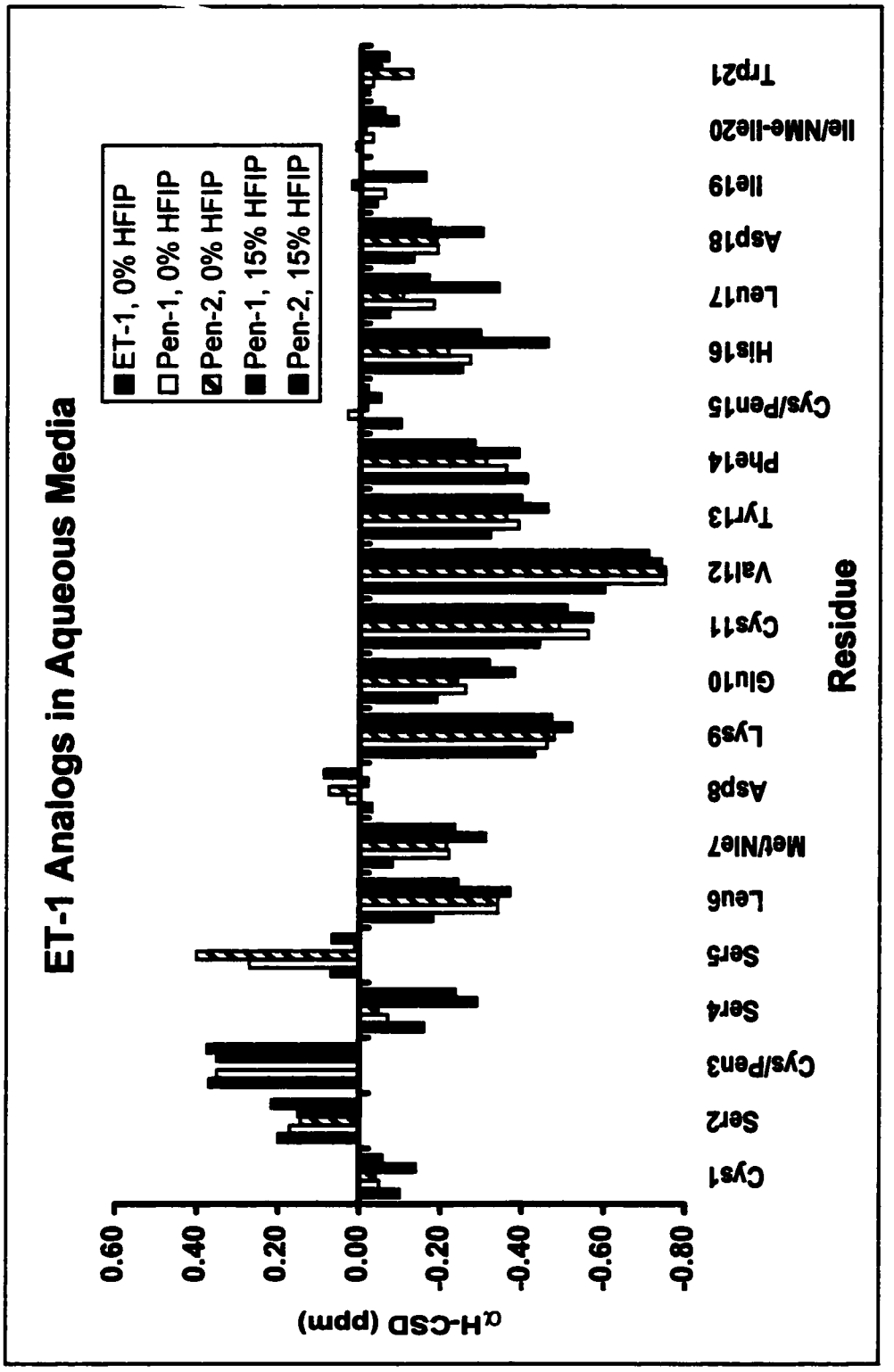


Figure 2.21: The αH-CSD plot for the intact ET-1 analogs.

Several differences, however, are apparent in the αH-CSD plot. Overall, the data indicate the presence of a stable helical region which, based on NOEs, is capped by Asp^{8}. In the case of ET-1, the helical region encompasses Lys^{9} through His^{16}, possibly Leu^{17}. The same residues of Pen-1 and Pen-2 display more negative αH-CSD values. The helical regions for both mutants also appear to be extended out to Asp^{18}. The loop region shows remarkable differences in the αH-CSD values of residues 5 → 7. The value for Ser^{5} is three times as positive, while Leu^{6} and Nle^{7} are nearly twice as negative for Pen-1, relative to the native peptide. This may suggest that increased stability of the helix and turn regions is coupled. That is, the fully helical conformer is also more structured at residues 5 → 8. Comparable patterns are observed for Pen-2, although the αH of Ser^{5} appears to be further downfield (by nearly 0.100ppm) than Pen-1.

According to the αH-CSD data, fluoroalcohol addition doesn't greatly effect either the N-terminal or the C-terminal regions of Pen-1. Structural enhancements mainly occur in the helix, with some effects in the turn region. Most of the turn residues are affected by HFIP: for Pen-1, the αHs of residues 4 → 8 move upfield, with Ser^{4} and Ser^{5} demonstrating the most dramatic adjustments. The same patterns for Pen-2 are also observed, although the α-methines of Leu^{6} and Asp^{8} actually move downfield. For all residues located in the helical region of the peptide, the αH-CSD becomes more negative, indicating an upfield chemical shift deviation upon HFIP addition. The effect is more dramatic at the C-terminal end of the helix: the αH-CSDs for residues 16 → 18 nearly double in size. This indicates that HFIP not only stabilizes the helical region, but extends it as well. An alternative explanation is that HFIP decreases the fraying at the C-terminus by increasing the structural stability of the peptide. This pattern, however, isn't as dramatic for the same residues of Pen-2, suggesting that a structuring effect of the N-methyl group limits the HFIP-induced structural changes in this span. The α-methine proton of Pen^{15} has a nearly disordered CSD value. This is most likely an artifact of

nearby aromatic residue sidechains which could deshield 15αH, rather than a kink in the helical structure (as discussed in Section 2.3.2, several strong *i* → *i+3* NOEs are present in this segment of the peptide).

2.3.5: Information from Pen-1 and Pen-2 NH-CSD/Temperature Gradient Plots.

NH-CSD/temperature gradient plots (Figure 2.22; see also Andersen et al., 1997) provide useful information about fluoroalcohol effects on Pen-1 and Pen-2. Trendlines of the helical regions for both 0% (solid line) and 15% (dashed lines) are presented. Table 2.8 reports the calculated slopes, which represent thermal stability, and correlation coefficients, which indicate the level of folding cooperativity, for the entire peptide (residues 1 → 21), as well as different helical lengths (residues 9 → 16 and 9 → 17). Also included in the table are the non-helical residues (1-8, 17-21).

Table 2.8: Pen-1 and Pen-2 Amide NH temperature gradient/NH-CSD.

	Residues	0% HFIP Slope [R^2] (ppt/K)	15% HFIP Slope [R^2] (ppt/K)
Pen-1	1 – 21 (all)	-5.955 [0.811]	-3.569 [0.561]
	9 – 16 (helix)	-7.337 [0.902]	-4.512 [0.864]
	9 – 17 (helix)	-7.018 [0.883]	-4.516 [0.869]
	1-8, 17-21	-5.765 [0.694]	-3.029 [0.391]
Pen-2	1 – 21 (all)	-4.823 [0.629]	-4.325 [0.412]
	9 – 16 (helix)	-6.847 [0.613]	-7.210 [0.673]
	9 – 17 (helix)	-7.269 [0.690]	-7.564 [0.754]
	1-8, 17-21	-4.386 [0.698]	-3.500 [0.294]

Upon first inspection, the entire peptide appears to be more thermally stable than the helical regions: slope values for 9 → 16 and 9 → 17 are steeper than for 1 → 21 for

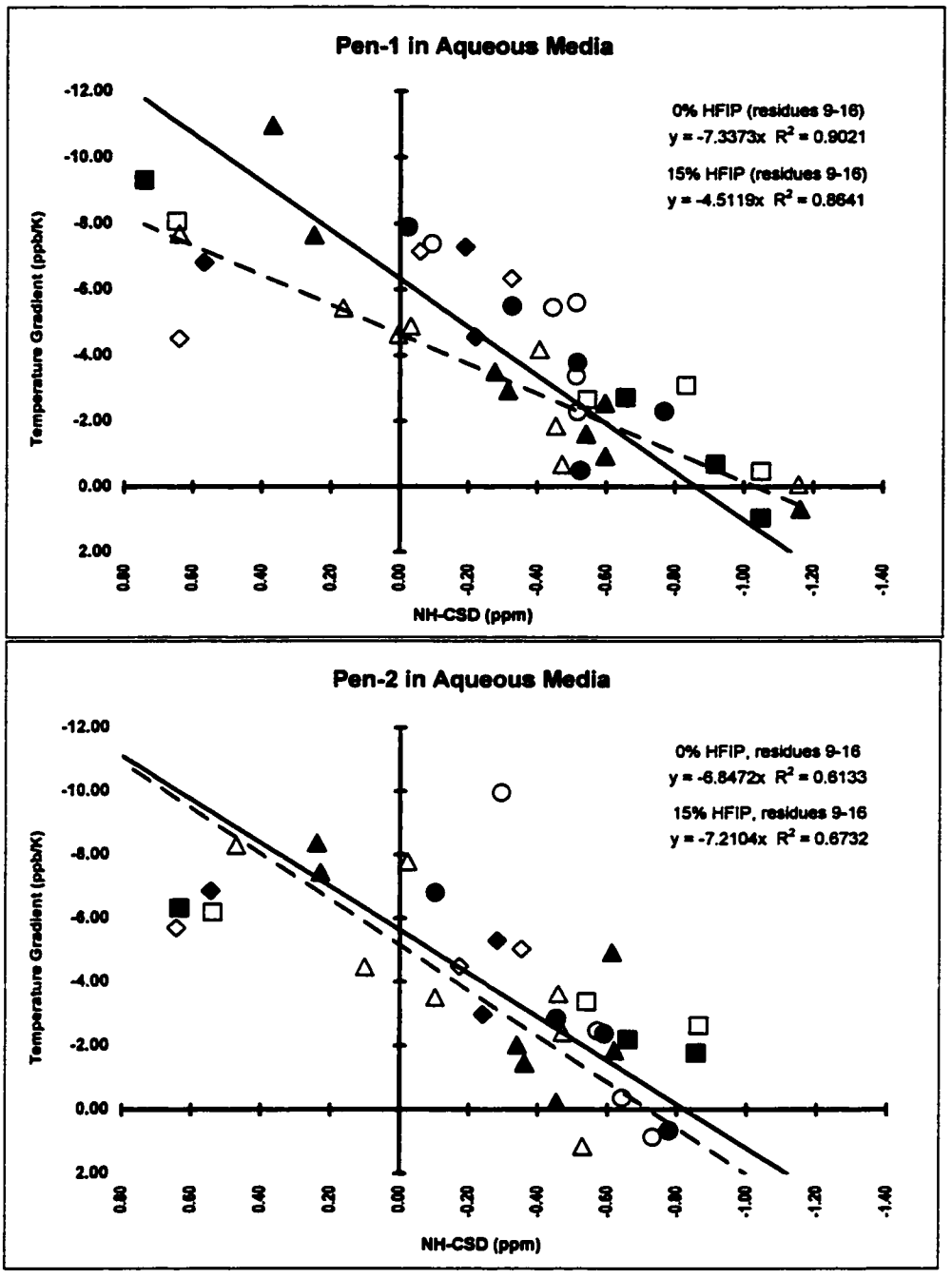


Figure 2.22: NH-CSD vs. temperature gradient plot of Pen-1 (top) and Pen-2 (bottom) in 0% (filled symbols) and 15% HFIP (open symbols). Symbol legend: diamonds (residues 1-4), squares (5-8), triangles (9-16), circles (17-21).

both Pen-1 and Pen-2 at 0% and 15% HFIP. Differences between residues 9-16 and 9-17 don't appear to be statistically significant. In both media, the helical region (9-16) has a steeper slope and higher R^2 values than the non-helical residues (1-8, 17-21). This would suggest that the helix is less thermally stable, but has higher folding cooperativity, than the N- and C-termini combined. The only exception is Pen-2 (0% HFIP), in which the correlation coefficient for the non-helical residues ($R^2 = 0.698$) is slightly larger than that of residues 9-16 ($R^2 = 0.613$).

The shallower slopes for Pen-2 at 0% HFIP indicate that the major folded state of the peptide is more stable than for Pen-1, although at 15% HFIP, the opposite is indicated. The lower correlation coefficients imply that Pen-2 demonstrates less cooperative folding in both fluoroalcohol concentrations. The Pen-2 helical region is initially comparable to that of Pen-1, although it demonstrates opposite trends when HFIP is added: the slopes become slightly steeper and the R^2 values increase at 15% HFIP, indicating that the helix becomes less thermally stable, but has higher cooperative folding. This may imply that in an aqueous glycol medium, the N-methyl group at the C-terminus initially stabilizes the fold of the peptide in the form of a hydrophobic cluster. When the solution becomes less polar, as when fluoroalcohols are added, this cluster begins to destabilize and appears to affect the helical region in terms of the thermal stability and folding cooperativity.

2.4: PEN-1 SOLUTION STRUCTURE *VS.* ET-1 CRYSTAL STRUCTURE

2.4.1: WHY COMPARE PEN-1 TO ET-1?

In most solution structures, ET-1 adopts a fairly regular helical conformation, capped by Asp^8 and significantly frayed by Leu^{17}. However, the conformation of the C-terminal region has generated a considerable amount of controversy since a consensus structure for the solution state has not yet been reached. The biologically critical C-

terminus adopts diverse conformations and has different interactions with the bicyclic core in structure ensembles derived for different solvent systems (for example: Saudek et al., 1989; Endo et al., 1989; Andersen et al., 1992a). The structure displayed in the upper panel of Figure 2.23 (Andersen et al., 1992a) emphasizes the bicyclic core. Due to rapid motion in solution, which generated multiple conformers, the C-terminal residues (Asp^{18} $\rightarrow$ Trp^{21}) are not displayed in this structure.

The authors of the X-ray structure (Janes et al., 1994) hoped to better define the ET-1 C-terminal configuration by more accurately modeling the peptide in physiological conditions (i.e. crystallized at neutral pH without the presence of organic co-solvents). Since the ligand binding site is located within the transmembrane helical bundle of the ET_A receptor, ET-1 is believed to be more rigid in its bound-state than in solution alone. Their conclusions, which contrasted those derived from NMR structures, further contributed to the debate. The crystal structure (Figure 2.23, bottom; Janes et al., 1994) displays an extended region at the N-terminus and a bulge between Ser^5 and Met^7. However, unlike the solution structure, the crystalline form has an irregular helix extending from Val^{12} to Trp^{21}. A loop defined by a $CO_i \rightarrow NH_{i+4}$ hydrogen bond between Met^7 and Cys^{11} helps to distort the helix at its N-terminus. In the crystal structure, the N-terminus of the helix appears to be either Val^{12} or Tyr^{13}. In solution-state NMR ensembles, the helical region encompasses residues 9 $\rightarrow$ 15 (or 18), not residues 12 $\rightarrow$ 21.

Steric interactions caused by the Pen β-methyl groups restricted conformational motion of the Pen-1 bicyclic core. As a result, the Pen-1 solution structure is less dynamic (Chen, 1992). Given that Pen-1 is a fully potent agonist, its solution structure may be more representative of the bound state than ET-1. For this reason, the Pen-1 analog was chosen as a model for a comparison study between the solution and crystal state conformations.

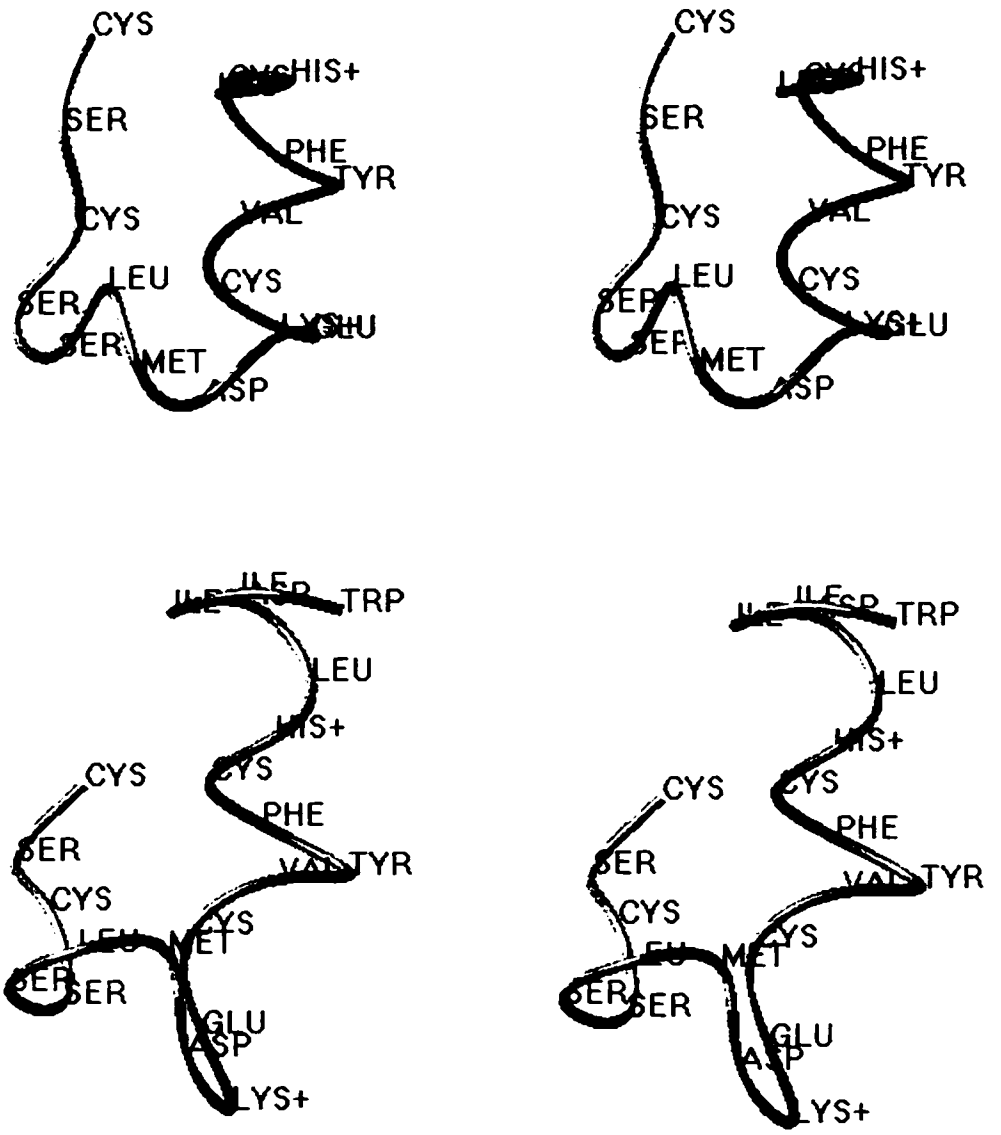


Figure 2.23: Stereoview of the ET-1 structures. Top, NMR solution structure (PDB code: 1edp; Andersen et al., 1992a), bottom, X-ray crystal structure (PDB code: 1edn; Janes, et al., 1994).

2.4.2: Initial Molecular Dynamics Studies of Pen-1.

Several major differences between this structure calculation and that performed by Chinpan Chen (1992) should be noted. First, ^{15}N-labeling and selective perdeuteration produced less complicated NOESY spectra. This allowed for more crosspeak assignments. As a result, many distance constraints, such as those involving the ^{15}N-labeled and perdeuterated residues, could be used with higher confidence in the molecular dynamics simulations. Second, a larger number of structures were used in the initial calculations, generally 60 models per refinement stage. This would help insure that a sufficient population was available for sampling. Third, to probe the viability of the available distance and dihedral constraints in the formation of the helix, no hydrogen bond constraints were used in this calculation. Chen's structures were calculated using hydrogen bond constraints, which may have created an artificially rigid and well-defined helix. Fourth, Chen's initial Pen-1 models were highly refined bicyclic ET-1 structures. Using a single family of starting structures as the initial set could, in effect, severely limit sampling. Randomized starting structures from different classes, such as helices and strands, would be better models from which to gage the effectiveness and accuracy of the distance constraints. In addition, intact disulfide bonds could also have restricted conformational sampling. Starting structures in this study were acyclic and randomized from helices and strands. The latter two differences may explain the high backbone convergence and rmsd (0.34Å, over residues 9 through 17) reported in Chen's dissertation. The final constraint sets used for this study, which represents the peptide in 50% aqueous glycol at pH 3.5, appear in Appendix D.

2.4.3: Effects of Altered Distance and Dihedral Constraint Weights.

The major difference between the solution-state and crystal-state endothelin structures is the locus of the helical region. The NMR ensemble structures (Andersen et

al., 1992) displays a helix from $Lys^{9} \rightarrow Cys^{15}$ (or Asp^{18}), while the crystal structure indicates a shift in its position, encompassing $Val^{12} \rightarrow Trp^{21}$. Why is the $9 \rightarrow 15$ helix not observed in the crystal structure? One possibility is that steric and covalent constraints preclude the formation of a stable α-helix from residues $9 \rightarrow 15$ in either the native peptide or the Pen-1 analog. If this were the case, the NMR data may reflect a "dynamic helix" in which a non-helical segment appears at different loci in the resulting solution-state conformational mixture. It is possible that one of these less helical conformations is present in the crystal structure. Alternatively, the solid-state structure may be the result of interactions not present in a monomeric solution-state conformational mixture. In order to resolve this question, distance and torsional angle weights for the Pen-1 models were probed to determine if a regular helix for residues 9 through 18 is consistent with the NOEs. To do so, force constants for the NOE (k_{NOE}) and dihedral (k_{cdih}) were adjusted so as to yield the same relative contributions to the total energy (E_{TOTAL}) during the molecular dynamics simulation. Structure ensembles that resulted from these calculations are displayed in Figure 2.24. Figures 2.25 and 2.26 display the torsional angles for both ensembles. Energies and backbone overlay rms deviations are reported in Table 2.9.

Any structural distortions caused by these constraints would appear exclusively in the impropers and van der Waals energies (E_{impr} and E_{vdw}, respectively). Rmsd values for the bond angles (± 2.5°) and bond lengths (± 0.007Å) were nearly identical between the two sets, indicating no significant deviation from normal geometry. The higher NOE rms values observed for the constrained helix indicates that mutually inconsistent NOEs are present due to residual dynamics in the solution state. The higher E_{impr} observed for the weighted NOE ensemble indicates that some structural distortions are created as k_{NOE} is raised. The fit to the NOE data remains adequate at lower k_{NOE} values. These results suggest that a regular helix from $9 \rightarrow 18$, which is not observed in the crystal structure,

can explain all of the observed NOEs and does not introduce any steric problems for endothelins.

Table 2.9: Effects of high helical torsion *vs.* NOE weights on Pen-1 structure calculations.

	Helical torsions	High NOE weights
Energies and NOE violations		
E_{TOTAL} (kcal)	622.812 ± 6.986	632.825 ± 19.305
E_{IMPR} (kcal)	0.390 ± 0.155	1.603 ± 0.415
E_{VDW} (kcal)	-81.895 ± 5.255	-74.330 ± 5.388
E_{NOE} (kcal)	29.226 ± 3.919	72.633 ± 13.510
E_{CDIH} (kcal)	56.293 ± 0.281	7.505 ± 6.735
rms (NOE)	0.245 ± 0.033Å	0.160 ± 0.020Å
RMSD of key region overlays		
rms (1-15)	1.377 ± 0.337Å	1.459 ± 0.371Å
rms (5-8)	1.123 ± 0.301Å	1.151 ± 0.433Å
rms (9-16)	0.448 ± 0.151Å	0.646 ± 0.156Å
rms (17-21)	1.828 ± 0.464Å	1.204 ± 0.460Å

The ensemble set in which the dihedral constraints are emphasized, as expected, show better convergence over the helical region than the set with highly weighted NOEs. In both ensembles the "source" structure used was the one which had the lowest E_{NOE}. Presumably this would represent the most stable model. The "constrained to a helix" set, superimposed over residues 9 through 16, has a rms deviation of 0.448 ± 0.151Å with respect to the structure with the lowest E_{NOE}. The comparable set for the "highly weighted NOEs" has a rmsd of 0.646 ± 0.156Å. Overlays of the intact bicyclic core (residues 1 through 15) and the loop region (residues 5 through 8) don't demonstrate significant statistical differences in the backbone rmsd values. Both sets indicate high conformational flexibility at the C-terminus.

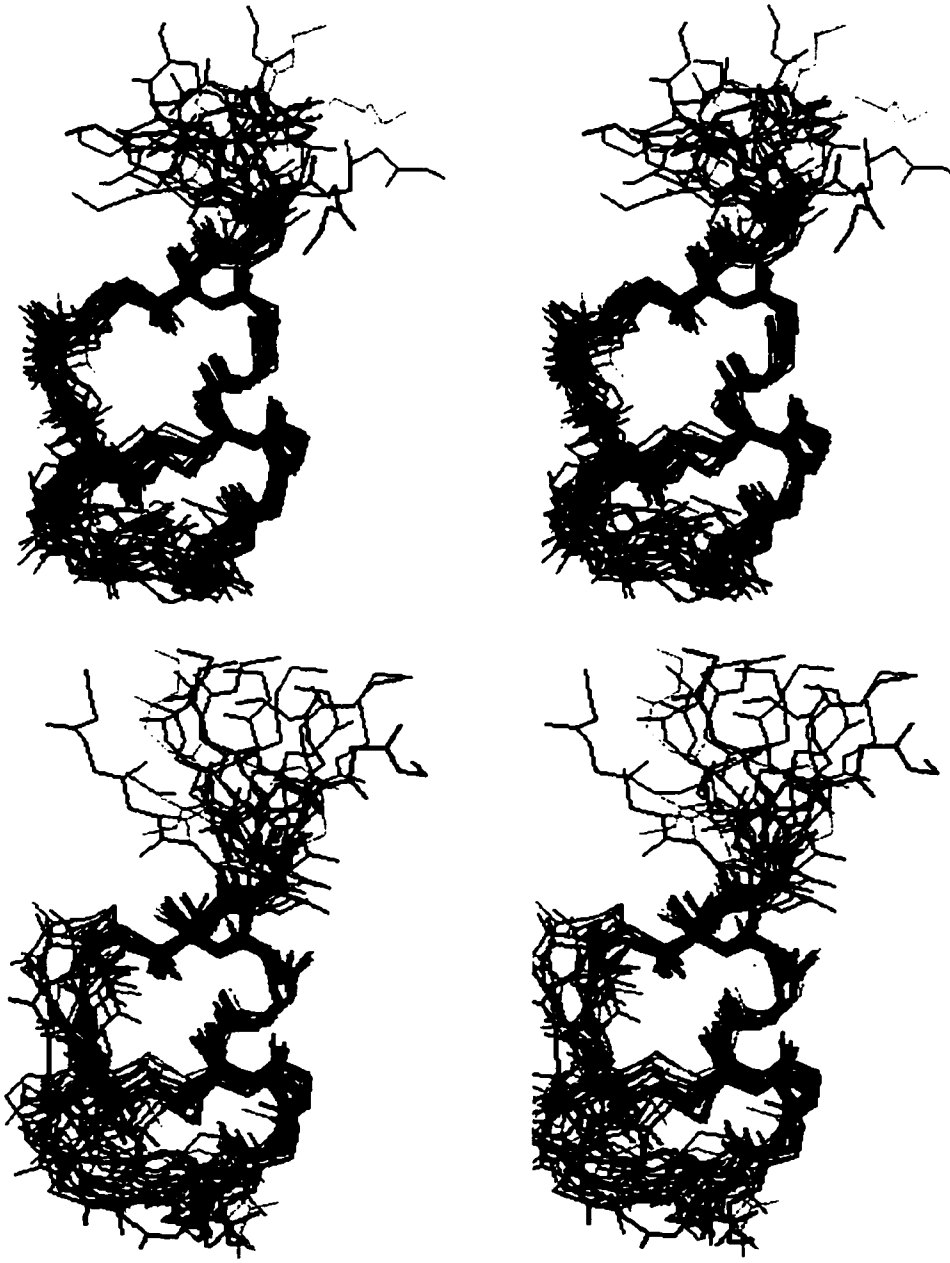


Figure 2.24: Structure ensembles superimposed over the backbones of residues 9 through 16. Top, highly weighted helix torsion angles (rms = 0.448 ± 0.151Å). Bottom, highly weighted NOE constraints (rms = 0.646 ± 0.156Å).

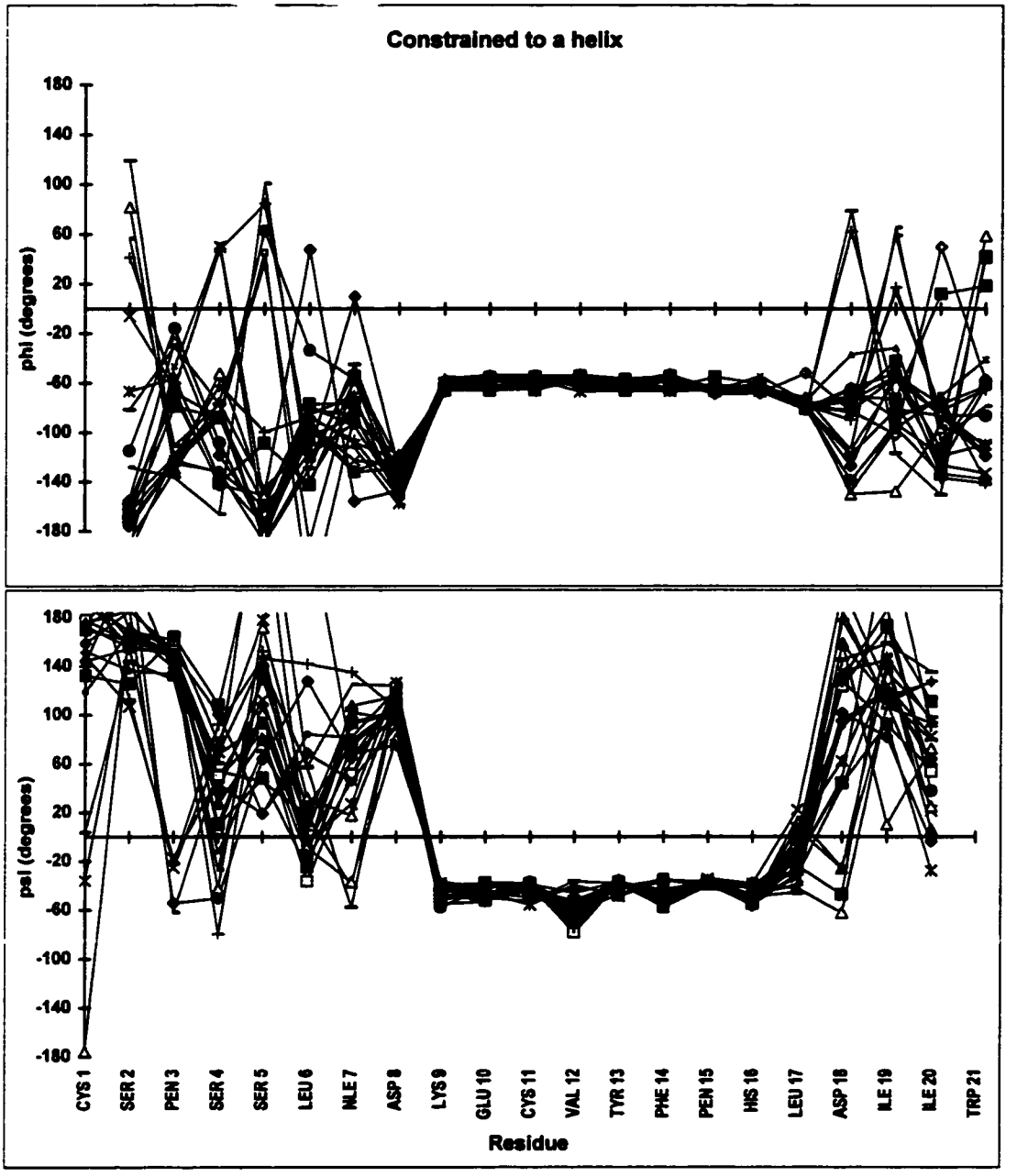


Figure 2.25: The phi/psi values of the best 21 structures from the Pen-1 NMR ensemble, corresponding to the highly constrained helix set.

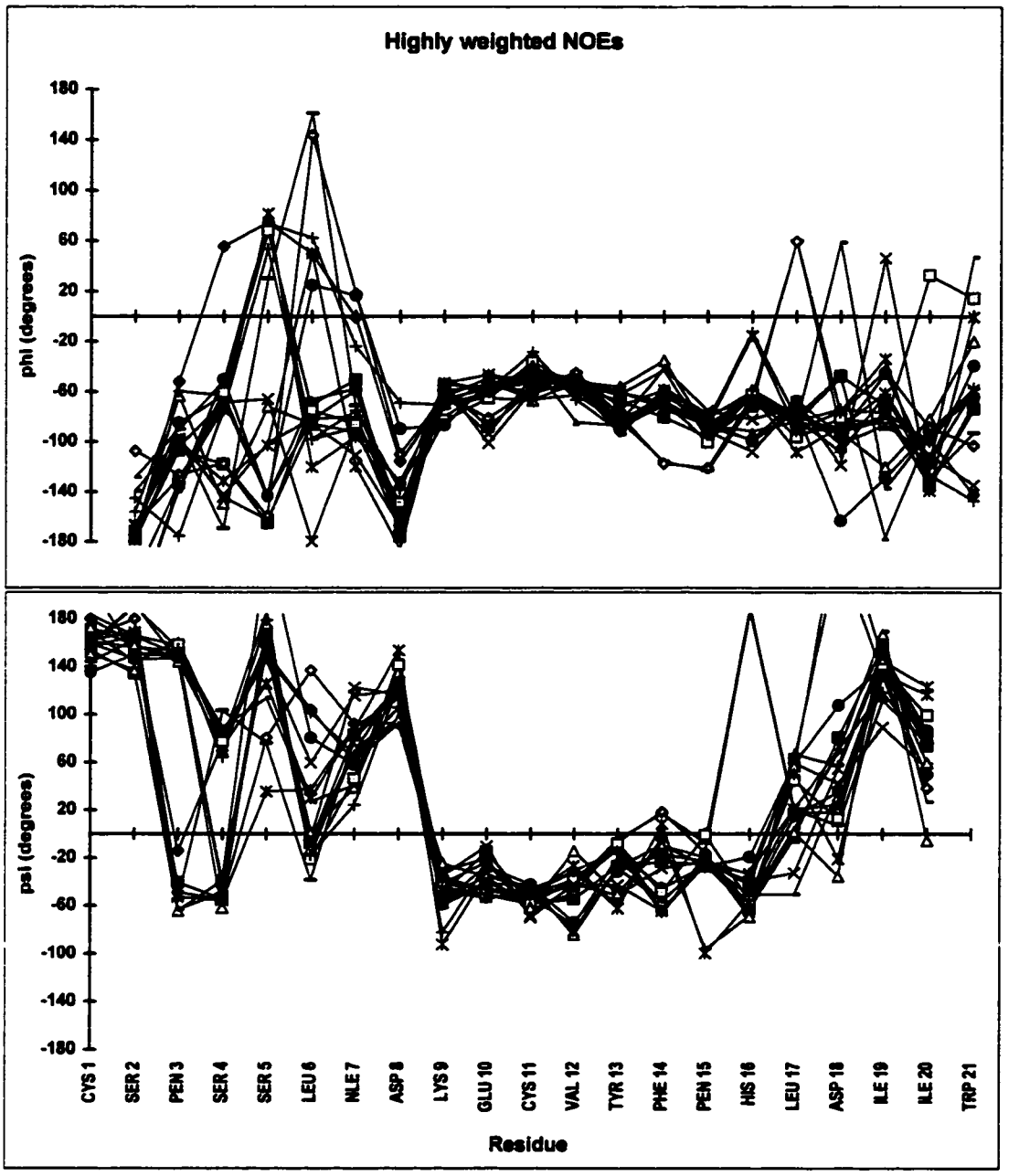


Figure 2.26: The phi/psi values of the best 21 structures from the Pen-1 NMR ensemble, corresponding to the highly weighted NOE set.

2.4.4: Evidence of the Helical Preference in Other ET-1 Analogs.

Janes et al. (1994) suggested that key differences between the ET-1 crystal structure and the various solution structures were due to the organic co-solvents used for the NMR studies. However, α-methine CSDs histograms (Figure 2.27) reveal that ET-1, as well as several monocyclic and intact mutants, adopts a stable helical conformation from Lys^9 through His^{16} in a variety of media. The top panel displays ET-1, the Pen-1 analog, and the monocyclic [$Ala^{3,11,17}$-Nle^7]-ET-1 analog in aqueous media. The peptides are arranged in order of decreased helicity. The αH-CSDs also show that a lower helix predominance over residues 9 through 15 correlates with a decreased structure in the loop region. The monocyclic analog in which the 3-11 disulfide bond was removed shows more disorder in *both* the loop and helical region.

Dalgarno et al. (1992) have reported the chemical shifts of two mutants, [Ala^7]-ET-1 (M7A) and [Ala^8]-ET-1 (D8A) in 40% aqueous acetic acid. Coles et al. (1994) studied a monocyclic ET-1 analog, [(α-aminobutyric acid)^{1,15}]-ET-1, in 10% aqueous acetonitrile. These are displayed in the lower panel of Figure 2.27. The Pen-1 analogs (with and without HFIP) appear to have an extended and less dynamic helical region than ET-1. Although both monocyclic analogs indicate that removal of the 3-11 or 1-15 disulfide bond increases flexibility of the peptides, both are helical from residues 9 through 15. The M7A and D8A mutants also demonstrate the same tendencies. The results for the D8A mutant are consistent with an important capping role for Asp^8: all αH-CSDs become less negative when the $Asp^8 \rightarrow Ala^8$ substitution takes place. In addition, the Ala^8 substitution also results in changes in the αH-CSD for residues in the loop region comparable to those observed for the monocyclic analogs.

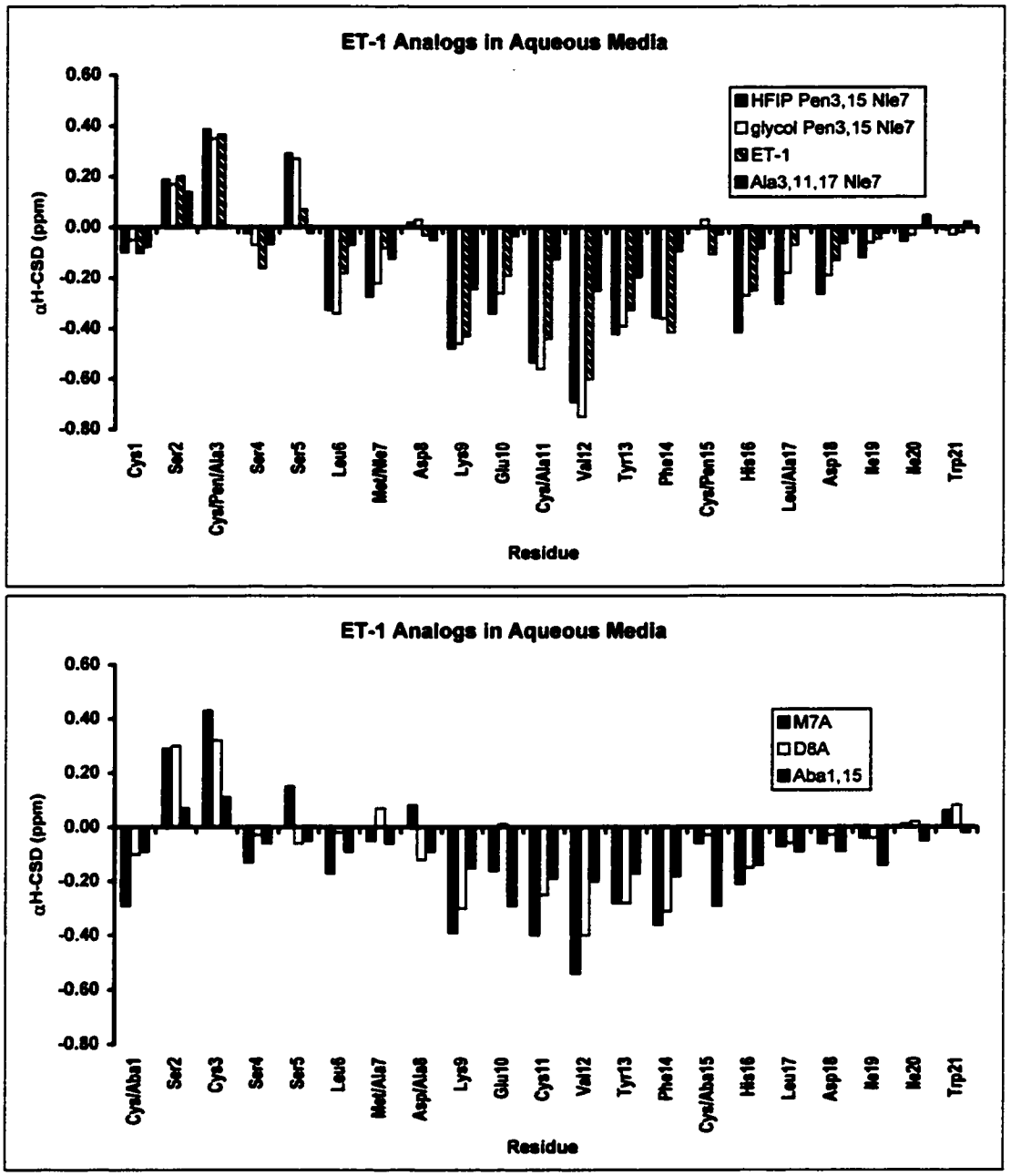


Figure 2.27: Alpha-CSD plots for ET-1 mutants. [Ala^{3,11,17}-Nle^{7}]-ET-1 from Andersen et al., 1995a. [Ala^{7}]-ET-1 (M7A) and [Ala^{8}]-ET-1 (D8A) from Dalgarno et al., 1992. [Aba^{1,15}]-ET-1 from Coles et al., 1994.

2.4.5: The ET-1 Crystal Structure can not Explain the Solution-State Conformational Equilibrium.

A quick survey of the calculated structures reveal that the crystal structure is a poor fit to the ET-1 and Pen-1 solution structures. Table 2.10 compares rmsd values for the ET-1 solution structure (PDB code: 1edp; Andersen et al., 1992a), ET-1 crystal structure (PDB code: 1edn; Janes et al., 1994), and the two Pen-1 structure ensembles (pnhx, constrained to a helix; apnhx, highly weighted NOEs). An overlay of the peptide backbone for the bicyclic core ($Cys^1 \rightarrow Cys^{15}$) generates high rmsd values for all models. However, if the helical regions are superimposed, the fits dramatically improve. In the case of the ET-1 and Pen-1 NMR ensembles, the rmsd values are nearly cut in half. This would indicate that the N-terminal and/or the loop region experience high mobility in solution. Moving the starting residue from Asp^8 to Lys^9 in the procedure does not significantly alter the fit between the ET-1 and Pen-1 structures. With the exception of the C-terminal helical region (residues 13 $\rightarrow$ 16), the X-ray structure is significantly different than the solution structures: over the key regions (residues 1-15, 8-16 and 9-16), the crystal structure has rmsd values two to three times higher with respect to the solution structures.

Several NOE contacts that are routinely observed in most ET-1 NMR studies are not predicted by the crystal structure. Table 2.11 compares key distances observed by NMR with those measured from the crystal structure. The most dramatic differences are observed at the N- and C-terminal residues of the helix. An extensive web of strong $i \rightarrow i+3$ NOEs ranging from Lys^9 to His^{16} are present in the ET-1 solution state. However, due to the bulge present in the crystal structure, larger interatomic distances are predicted between the residues at the N-terminus of the helical region. In addition, a number of close contacts at the C-terminus (residues 19 $\rightarrow$ 21) that are predicted by NMR, have

Table 2.10: RMSD values for ET-1 analog model overlays.

Reference model	Target models: 1edn	Target models: pnhx	Target models: apnhx	overlay regions
1edp	2.495Å	1.996 ± 0.286Å	1.959 ± 0.277Å	1-15
	1.950Å	1.080 ± 0.135Å	1.062 ± 0.156Å	8-16
	2.001Å	1.056 ± 0.153Å	1.030 ± 0.165Å	9-16
	1.308Å	1.074 ± 0.150Å	1.073 ± 0.186Å	12-16
	1.107Å	1.025 ± 0.123Å	1.084 ± 0.198Å	13-16
1edn		2.896 ± 0.173Å	2.964 ± 0.185Å	1-15
		2.019 ± 0.129Å	1.948 ± 0.162Å	8-16
		2.013 ± 0.131Å	1.960 ± 0.153Å	9-16
		1.335 ± 0.099Å	1.288 ± 0.135Å	12-16
		1.027 ± 0.082Å	1.028 ± 0.175Å	13-16
		2.202 ± 0.425Å	2.293 ± 0.323Å	17-21
pnhx7		1.377 ± 0.337Å	1.535 ± 0.262Å	1-15
		0.563 ± 0.171Å	0.782 ± 0.222Å	8-16
		0.448 ± 0.151Å	0.685 ± 0.209Å	9-16
		0.257 ± 0.112Å	0.572 ± 0.187Å	12-16
		0.204 ± 0.094Å	0.509 ± 0.211Å	13-16
		1.828 ± 0.464Å	1.778 ± 0.324Å	17-21
apnhx4		1.706 ± 0.232Å	1.459 ± 0.371Å	1-15
		0.788 ± 0.113Å	0.766 ± 0.206Å	8-16
		0.632 ± 0.114Å	0.646 ± 0.156Å	9-16
		0.470 ± 0.093Å	0.531 ± 0.118Å	12-16
		0.398 ± 0.073Å	0.491 ± 0.149Å	13-16
		1.399 ± 0.370Å	1.206 ± 0.461Å	17-21

1edp: ET-1 NMR structure (Andersen et al., 1992a). *Note: 1edp missing 18-21.*
1edn: ET-1 X-ray structure (Janes et al., 1994).
pnhx: Pen-1 NMR structure ensemble, constrained to a helix (Lee et al., 1994).
apnhx: Pen-1 NMR structure ensemble, highly weighted NOEs (Lee et al., 1994).

Table 2.11: NMR-detected intermediate range contacts absent in the X-ray structure.

Connectivity	NOE distance constraint[1] (Å)	X-ray structure distance[2] (Å)
Helical region		
9α → 12β	2.40 – 2.75	8.71
9α → 12γ1*	2.55 – 3.25	9.19 – 10.58
9α → 12γ2*	3.20 – 4.40	10.35 – 11.21
9α → 12NH	2.90 – 3.70	6.50
9α → 13NH	3.65 – 5.50	10.14
10α → 13NH	3.20 – 3.70	7.21
10α → 13β*	2.90 – 3.65	5.77 – 7.33
10α → 14NH	3.55 – 4.35	6.51
11α → 14NH	3.20 – 4.80	5.08
12α → 17NH	3.20 – 3.80	5.77
13α → 15NH	3.50 – 4.60	5.76
13α → 16NH	3.35 – 3.90	6.02
Other regions		
5NH → 11β*	3.10 – 3.90	8.46 – 8.79
5β* → 7NH	3.20 – 4.25	6.40 – 7.03
6α → 12γ2*	2.60 – 3.50	7.29 – 8.65
12γ1* → 16δ2	3.00 – 4.00	5.23 – 6.95
19α → 20NH	2.05 – 2.55	3.55
19γ2* → 21NH	3.35 – 4.75	6.29 – 7.75
Short distances in the X-ray structure		
6α → 7NH	2.80 – 3.45	2.13
7NH → 11β*	none (> 5.00)	3.45 – 4.12
7α → 11α	none (> 5.00)	4.36
7α → 11β*	none (> 5.00)	1.49 – 2.78
8NH → 10NH	none (> 5.00)	4.73
12α → 14NH	3.60 – 4.70	2.09
12α → 13NH	3.30 – 3.65	***1.84***
13NH → 14NH	2.75 – 3.11	***1.17***

[1] Andersen et al., 1992a. [2] Janes et al., 1994.

larger distances in the crystal structure. There are also a number of short distances observed in the crystal structure which are not observed as NOEs or as NOEs that reflect longer distances in the NMR ensembles. This is most noticeable in the loop/bulge region of the crystal structure, for example, between Met^{7} and Cys^{11}.

These major structural differences between the two states may be due to the fact that the ET-1 crystals were obtained from water without the presence of any organic co-solvents, surfactants or buffers. Previous reports have indicated ET-1 aggregates in pure water at concentrations exceeding 22μM (Bennes et al., 1990; Calas et al., 1992). This aggregation effect was manifested in the intermolecular contacts observed in the crystal structure. One extensive "dimeric contact", which buries 800Å^2 of surface area, partially explains discrepancies in the helical region. The major feature of this ET-1 dimer was an intermolecular hydrogen bond between the sidechain hydroxyl group of Tyr^{13} and the backbone carbonyl oxygen of Asp^{8}. The structures presented by Janes et al. also suggest that the aliphatic sidechains of Met^{7} and Val^{12} contribute to the contact region of the dimer. This hydrophobic cluster may provide sufficient stabilization in the structure so as to overcome the helical propensity of the $Lys^{9} \rightarrow Tyr^{13}$ span observed in the monomeric state.

In summary, the crystal structure is a poor model for the solution state. However, this is not to say that either the crystal or solution structures are *the* correct conformations. The ET-1 peptide may require certain flexibility for effective receptor binding and signaling.

CHAPTER 3: NMR STUDIES OF ET-1 AND PEN-2 C-TERMINAL ANALOGS

3.1: INTRODUCTION: WHY STUDY C-TERMINAL ANALOGS?

Numerous studies in a variety of media (Saudek et al., 1989; Endo et al., 1989; Andersen et al., 1992a) have determined the solution-state structure of the bicyclic core of ET-1, but failed to reach a consensus regarding the conformationally flexible C-terminus. Both regions, however, are known to be important ET_A receptor binding and signaling factors. Several monocyclic ET-1 analogs (Coles et al., 1994; Andersen et al., 1995a) have also been analyzed as to the extent of their structuring in aqueous media. Most of the monocyclic analogs lacked either the Cys^1-Cys^{15} or Cys^3-Cys^{11} disulfide bond. Both reports conclude that even with one of the disulfides missing, the 21 residue analogs still retained some helicity. However, several questions remain regarding the disulfide bridges and induced helicity. Specifically, are the disulfides an absolute requirement in the formation and stabilization of the helical segment? In addition, is the antiparallel extended region (residues 1 $\rightarrow$ 3), which forms one side of the hydrophobic core of the intact analogs, also a significant factor?

As discussed in the previous chapter, the $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ mutation significantly alters the conformational preferences of the C-terminus, as well as the biological activity of the peptide. Large upfield chemical shift deviations of the Ile^{19} γ-methyl group and a turn-like CD signature (Harris, 1993) suggests that presence of a local hydrophobic cluster at the C-terminus, which may be further stabilized by the bicyclic core. These initial conclusions led to further questions regarding the major conformation of the agonist *vs.* that of the antagonist, as well as to the extent and importance of hydrophobic interactions between the bicyclic core and the C-terminus, if any exist.

To answer some of these questions, several C-terminal fragment analogs were examined. Although studies of C-terminal fragments have recently appeared in the literature (for example, Cody et al., 1997; Katahira et al., 1998), none have expressly dealt with longer acyclic sequences in aqueous media. Cody et al. (1997) primarily discuss a 6-residue ET_A receptor antagonist in both aqueous and DMSO media. One significant difference in their $^{(NMe)}Ile^{20}$ analog is the presence of a large unnatural aromatic residue at position 16 which, the authors claim, greatly affects the peptide conformation in DMSO: in aqueous media, the 19/20 peptide bond has the *trans* configuration, while in DMSO, it isomerizes to the *cis* configuration. Katahira et al. (1998) present a number of 14-residue ET_B receptor agonists studied in dipalmitoyl phosphatylcholine vesicles. In that report, several ET_B receptor agonists, all spanning residues 8 through 21 and N-capped with either acetyl or succinyl groups, are examined in DPPC micelles. The authors note that most of these peptides are helical through residue 15 or 16, and that the conformation of the C-terminus ($Ile^{19} \rightarrow Trp^{21}$) is similar to a γ-turn. In all the fragments examined in this chapter, the bicyclic core is disrupted by both truncation of the sequence (the largest of these analogs encompass residues $8 \rightarrow 21$) and incorporate $Cys \rightarrow Ala$ mutations at residues 11 and 15. Fragments with and without the $^{(NMe)}Ile^{20}$ moiety are also examined herein.

3.2: MATERIALS AND METHODS

3.2.1: Peptide Synthesis and Sample Preparation

Two C-terminal ET-1 peptide analogs (see Table 3.1, below) were received from the Bristol Myers Squibb Pharmaceutical Research Institute (Princeton, NJ) and were used without further purification. These samples were prepared using standard t-Boc solid phase methods on a Bioresearch 9600 peptide synthesizer and characterized *via* FAB mass spectrometry.

Table 3.1: BMSQ Peptides

Sequence	Name	FW	FAB-MS Data
HLDIIW	6mer (16-21)	794.93	none reported
VYFAHLDIIW	10mer (12-21)	1276.51	$[M+H]^+$ 1277

All longer sequences were synthesized on an Applied Biosystems Model 430A peptide synthesizer using the standard Fmoc solid phase coupling procedures (HOBt as the coupling reagent). Wang resin (Wang, 1973) preloaded with Fmoc-Trp was purchased either from Advanced ChemTech, Inc. (Louisville, KY) or Nova Biochem Corp. (San Diego, CA). Coupling reagents and protected amino acids were purchased from Advanced ChemTech. Peptides were purified *via* reversed-phase HPLC using a Waters DeltaPak® semi-preparative C18 column and water/acetonitrile gradients. Major HPLC fractions were collected and characterized on a Kratos HV-4 mass spectrometer equipped with an electrospray ionization source. Further HPLC analysis using a Waters μBondapak® analytical C18 column was required to determine sample purity prior to NMR and CD studies.

Residues 19 to 21 of the peptides containing the $I^{(NMe)}IW$ moiety were manually coupled using PyBrOP (Coste et al., 1994) as the coupling reagent. The procedure used is slightly modified from that of Spencer et al. (1992) wherein the reaction time of the Ile-$^{(NMe)}$Ile coupling was expanded from 24 hours to approximately 72 hours. To confirm that coupling was complete, 50 to 100mg of the peptide-resin complex were subjected to another coupling cycle, either with Fmoc-Ala or protected Fmoc-Asp, to yield Ac-$AI^{(NMe)}IW$ or Ac-$DI^{(NMe)}IW$ after N-capping and resin cleavage. These two peptides were characterized *via* ESI-MS (Ac-$AI^{(NMe)}IW$, $[M+H]^+$ 558.3; Ac-$DI^{(NMe)}IW$, $[M+H]^+$ 602.6) and ^{1}H-NMR. Once the coupling was confirmed, approximately 0.25mmol portions (300-400mg) of the Fmoc-$I^{(NMe)}IW$-resin fragment were placed on the ABI 430A peptide synthesizer to complete the sequences. Longer fragments used for this portion of the

study are listed below in Table 3.2. All peptides lyophilize as fluffy white powder, with purity > 90% as assessed by analytical HPLC.

Table 3.2: Peptides – Long Fragments

Sequence	Name	FW	ESI-MS Data	
GSHLDI IW	8mer (14-21)	939.99	M^+ $[M+Na]^+$ $[M+Na]^{2+}$	940.2 (Base) 962.1 (24.8%) 481.9 (38.1%)
GSHLDI$^{(NMe)}$IW	NMe-8mer (14-21)	954.45	M^+ M^{2+} $[M-Trp]^{2+}$	954.5 (2.3%) 478.1 (25.5%) 375.6 (Base)
Ac-GSHLDI$^{(NMe)}$IW	AcNme-8mer (14-21)	996.45	M^+ $[M+H]^+$ $[M+Na]^{2+}$	996.5 (56.9%) 997.8 (19.7%) 509.9 (54.6%)
EAVYFAHLD	9mer (10-18)	1063.51	M^+ M^{2+}	1107.2 (Base) 554.0 (40.4%)
EAVYFAHLDI IW	12mer (10-21)	1477.71	M^{2+}	738.5 (69.7%)
DAEAVYFAHLDI IW	14mer (8-21)	1664.54	M^{2+}	831.3 (Base)
DAEAVYFAHLDI$^{(NMe)}$IW	NMe-14mer (8-21)	1678.54	M^{2+} $[M+Na]^{2+}$ $[M-Trp]^{2+}$	839.3 (Base) 850.4 (3.2%) 737.2 (29.1%)
Ac-DAEAVYFAHLDI$^{(NMe)}$IW	AcNMe-14mer (8-21)	1718.83	M^{2+} M^{3+}	858.8 (39.8%) 573.3 (Base)

Typical peptide concentrations for NMR studies ranged from 2 to 5mM. Samples for initial NMR studies were either dissolved in neat d_6-DMSO or MilliQ® water containing nominally 5 to 40% d_4-acetic acid by volume. DMSO samples were further titrated with water (2 to 5%) and d-TFA (up to 3 equivalents) to insure charged sidechains were fully protonated. This sharpened the backbone amide signals to such a degree so that $^3J_{NH/\alpha H}$ coupling constants could be accurately determined from the one-dimensional spectra. Aqueous samples were titrated with d_6-ethylene glycol (up to 30%) and/or d_2-HFIP (up to 40%). The longer peptides, specifically the NMe-14mer analogs required up to 40% d_4-acetic acid in order to solvate sufficient quantities for NMR studies. No studies of the 14mer in aqueous media were performed due to peptide

aggregation at concentrations greater than 1mM. During the AcNMe-14mer D_2O exchange studies, the aqueous sample was lyophilized in the NMR tube and reconstituted in a solvent mixture containing 24% d_4-HOAc/36% D_2O/40% d_2-HFIP. For the ^1H reincorporation, the same procedure was followed except where water replaces D_2O in the solution mixture.

3.2.2: NMR EXPERIMENTS

Except where otherwise noted, all NMR experiments were acquired on one of two Brüker Instruments AM-500 spectrometers (one tuned to 500.13MHz, the other to 499.87MHz). Residual water peaks were suppressed using a presaturation pulse sequence applied over an 8Hz range centered on the water signal. One- and two-dimensional NMR experiments on samples in DMSO media were used to verify primary sequences, while samples dissolved in aqueous media were used to analyze conformational preferences.

The major NMR experiments performed are the same as those noted in Section 2.4.2 of this dissertation. Due to the relatively smaller sizes of the fragment peptides, tppi COSY (Aue et al., 1976; Braunchweiler and Ernst, 1983) and tppi ROESY (Bax and Davis, 1985) experiments were used to assign chemical shifts and determine conformational preferences. The pulse sequences for the tppi COSY and tppi ROESY are the same as those found in the Ph.D. dissertations of Bolong Cao (1993) and Scott Harris (1993). Typical spin lock times for the ROESY experiments ranged from 150 to 350ms. Temperature gradients were obtained from 1D data collected in 5K intervals over a range of 295 to 320K (DMSO media) or 285 to 320K (aqueous media). NMR spectra for the BMSQ-synthesized peptides were acquired with TSP as the internal reference standard, while DSS was used for the C-terminal fragment peptides synthesized in house.

For both the NMe-14mer and AcNMe-14mer, tppi NOESY (Bodenhausen et al., 1984) experiments (τ_m of 200ms and 240ms) were also acquired at 285K and 310K on a

Brüker Instruments DMX-750 spectrometer (17.4T, tuned to 750.13MHz) equipped with a Silicon Graphics Indigo2 workstation. D_2O exchange and ^1H reincorporation studies of AcNMe-14mer in its most helical state (40% d_2-HFIP/24% d_4-HOAc/36% water) were performed on a Brüker Instruments DRX-500 spectrometer (11.74T, tuned to 499.87MHz) equipped with a Silicon Graphics O_2 workstation. For these experiments, the residual water signal was suppressed using a pulse field gradient watergate (Piotto et al., 1992; Sklenar et al., 1993) sequence centered on the solvent peak. All exchange data was collected at ambient temperatures (~300K).

3.3: AQUEOUS MEDIA NMR STUDIES – THE 14MERS

Several ET-1 fragments were synthesized in order to determine the role of the disulfides in helix formation and stability in aqueous media. All of these peptide lack the rigidity of the bicyclic core: the longest of the sequences begin at Asp^8 and include Cys → Ala substitutions at positions 11 and 15. Asp^8 was chosen as the starting position due to its role as the N-capping residue of the ET-1 helical region (see also section 2.3.2). Structural effects of the $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ mutation are also examined by inserting the N-methylated Ile^{20} in the sequence of the linear fragments.

3.3.1: Unexpected Aggregation Problems.

Due to the hydrophobic nature of the peptides, the 14mer aggregated in aqueous media at NMR concentrations. Addition of organic co-solvents such as acetic acid or HFIP did not noticeably improve the solubility of this peptide. However, contrary to expectations, the NMe-14mers (both with and without the acetyl N-cap) did not suffer from the same aggregation problems. The hydrophobic cluster that forms in aqueous media may partially explain the differences in solubility between these peptides. The cluster is less compact in the case of the 14mer, resulting in a more flexible structure than

the N-methylated 14mer analogs. This may allow the C-terminal region of the molecules to more easily stack upon each other at high concentrations, producing aggregates.

3.3.2: SPECTRAL ASSIGNMENT OF THE 14MERS.

Due to a number of degenerate αH chemical shifts, as well as low signal-to-noise resolution of the spectra acquired at 500MHz (prior to the spectrometer console and probe upgrade), NOESYs were acquired at 750MHz. The higher field strength increased the shift dispersion of the α-methine protons, while the newer spectrometer allowed for higher overall signal-to-noise resolution in the spectra. Final chemical shift assignments for NMe-14mer and AcNMe-14mer in aqueous 40% acetic acid are listed in Tables 3.3 and 3.4. One striking observation is that even in the absence of fluorinated alcohols, both NMe-14mers display NOESY crosspeak patterns consistent with secondary structure. A number of intermediate range NOESY crosspeaks are present for the NMe-14mer in 40% aqueous acetic acid (Figure 3.1). For example, several strong $\alpha_i \rightarrow NH_{i+3}$ interactions are observed, beginning with 10α → 13NH and ranging to 15α → 18NH. Also evident are weak 10α → 13δδ′ and 11α → 14δδ′ crosspeaks. Several $\alpha_i\beta_{i+3}$ crosspeaks are present in the upfield region of the NOESY (Figure 3.2) ranging from 9α → 12β through 12α → 15β. NOE interactions between 14α and the 17βs may be also present; the α-methine chemical shifts of Phe[14] and Ile[19] are degenerate. The most prominent peak in this region is one between 19α and 20NMe. This and the lack of any 19α → 20α interaction indicates that the major conformer of the 19/20 peptide bond is *trans*.

As expected, the AcNMe-14mer in 40% aqueous acetic acid displays many of the same features in the NOESY (Figures 3.3-3.5) as the non-capped analog. However, capping the peptide with an acetyl group stabilizes the structure remarkably. Whereas many $NH_i \rightarrow NH_{i+1}$ crosspeaks are weak in the case of the NMe-14mer (not shown), the corresponding crosspeaks are medium-to-strong over the 9NH through 19NH region for

the capped peptide (Figure 3.3). Greater α-methine shift dispersion is observed, presumably due to the increased helix stability. As a result, $i \rightarrow i+3$ crosspeaks are more clearly observed, especially in the upfield region of the NOESY spectrum (Figure 3.5) where the $\alpha_i\beta_{i+3}$ crosspeaks are observed from residues 8 through 17. Many of the $^3J_{NH\text{-}\alpha H}$ coupling constants (see Table 3.5) also provide evidence of helicity in the fragments: in 40% aqueous acetic acid for both peptides, residues 9 through 15 display coupling constants less than 6Hz. More importantly, all of the NH/αH coupling constants decrease with N-terminal acetylation, indicating a higher population of helical states in the ensemble.

Table 3.3: Chemical shift assignments for DAEAVYFAHLDI$^{(NMe)}$IW in 40% acetic acid, 60% water at 285K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	exchanged	4.428	3.196, 3.196	
Ala 9	8.771 [-6.238]	4.367	1.457	
Glu 10	8.279 [-6.340]	4.316	2.150, 2.069	γ,γ′ 2.518, 2.518
Ala 11	8.073 [-3.204]	4.317	1.435	
Val 12	7.800 [-4.534]	3.963	2.037	γ,γ′ 0.926, 0.815
Tyr 13	7.987 [-6.681]	4.466	2.955, 2.921	δ,δ′ 6.893, ε,ε′ 6.717
Phe 14	8.058 [-7.704]	4.503	3.204, 2.994	δ,δ′ 7.272, ε,ε′ 7.343, ζ 7.281
Ala 15	8.057 [-6.204]	4.268	1.402	
His 16	8.151 [-2.795]	4.697	3.375, 3.235	δ 7.363, ε 8.659
Leu 17	7.982 [-1.773]	4.380	1.655, 1.568	γ 1.581, δ,δ′ 0.874, 0.835
Asp 18	8.349 [-3.681]	4.746	2.927, 2.774	
Ile 19	7.902 [-5.488]	4.506	1.601	γ1,γ1′ 1.437, 1.015
				γ2 0.276, δ 0.755
$^{(NMe)}$Ile 20	n/a	4.761	2.026	γ1,γ1′ 1.261, 0.942
				γ2 0.885, δ 0.791, NMe 3.025
Trp 21	7.949 [-7.056]	4.647	3.373, 3.190	δ 7.221, ε1 9.920, ε3 7.599
				ζ2 7.412, ζ3 7.092, η2 7.166

Table 3.4: Chemical shift assignments for Ac-DAEAVYFAHLDI$^{(NMe)}$IW in 40% acetic acid, 60% water at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	8.304 [-5.556]	4.775	2.991, 2.991	Ac-Me 2.081
Ala 9	8.445 [-5.795]	4.212	1.452	
Glu 10	8.207 [-5.693]	4.252	2.171, 2.110	γ,γ′ 2.520, 2.520
Ala 11	7.896 [-2.557]	4.270	1.484	
Val 12	7.877 [-7.158]	3.867	2.067	γ,γ′ 0.952, 0.847
Tyr 13	7.942 [-6.067]	4.392	2.991, 2.963	δ,δ′ 6.818, ε,ε′ 6.685
Phe 14	8.143 [-8.385]	4.417	3.252, 3.048	δ,δ′ 7.310, ε,ε′ 7.368, ζ ?
Ala 15	8.052 [-5.965]	4.269	1.437	
His 16	8.033 [-1.432]	4.670	3.398, 3.240	δ 7.375, ε 8.648
Leu 17	7.819 [+0.307]	4.324	1.652, 1.565	γ 1.566, δ,δ′ 0.842, 0.813
Asp 18	8.259 [-2.591]	4.703	2.951, 2.772	
Ile 19	7.861 [-4.943]	4.489	1.598	γ1,γ1′ 1.435, 1.002
				γ2 0.270, δ 0.751
$^{(NMe)}$Ile 20	n/a	4.746	2.020	γ1,γ1′ 1.246, 0.927
				γ2 0.879, δ 0.787, NMe 3.025
Trp 21	7.954 [-6.067]	4.643	3.349, 3.180	δ 7.213, ε1 9.900, ε3 7.592
				ζ2 7.402, ζ3 7.087, η2 7.159

Table 3.5: Amide NH/α-methine coupling constants at 290K

Residue	NMe-14mer, 40% HOAc $^3J_{NH-\alpha H}$ (Hz)	AcNMe-14mer, 40% HOAc $^3J_{NH-\alpha H}$ (Hz)
Asp^{8}	n/a	6.9
Ala^{9}	4.3	~ 4.0
Glu^{10}	5.7	~ 4.7
Ala^{11}	~ 5.2	5.0
Val^{12}	6.1	5.9
Tyr^{13}	unknown	4.6
Phe^{14}	~ 5.2	~ 4.7
Ala^{15}	~ 5.2	unknown
His^{16}	7.0	unknown
Leu^{17}	unknown	6.0
Asp^{18}	7.5	6.0
Ile^{19}	7.8	6.6
$^{(NMe)}Ile^{20}$	n/a	n/a
Trp^{21}	unknown	6.4

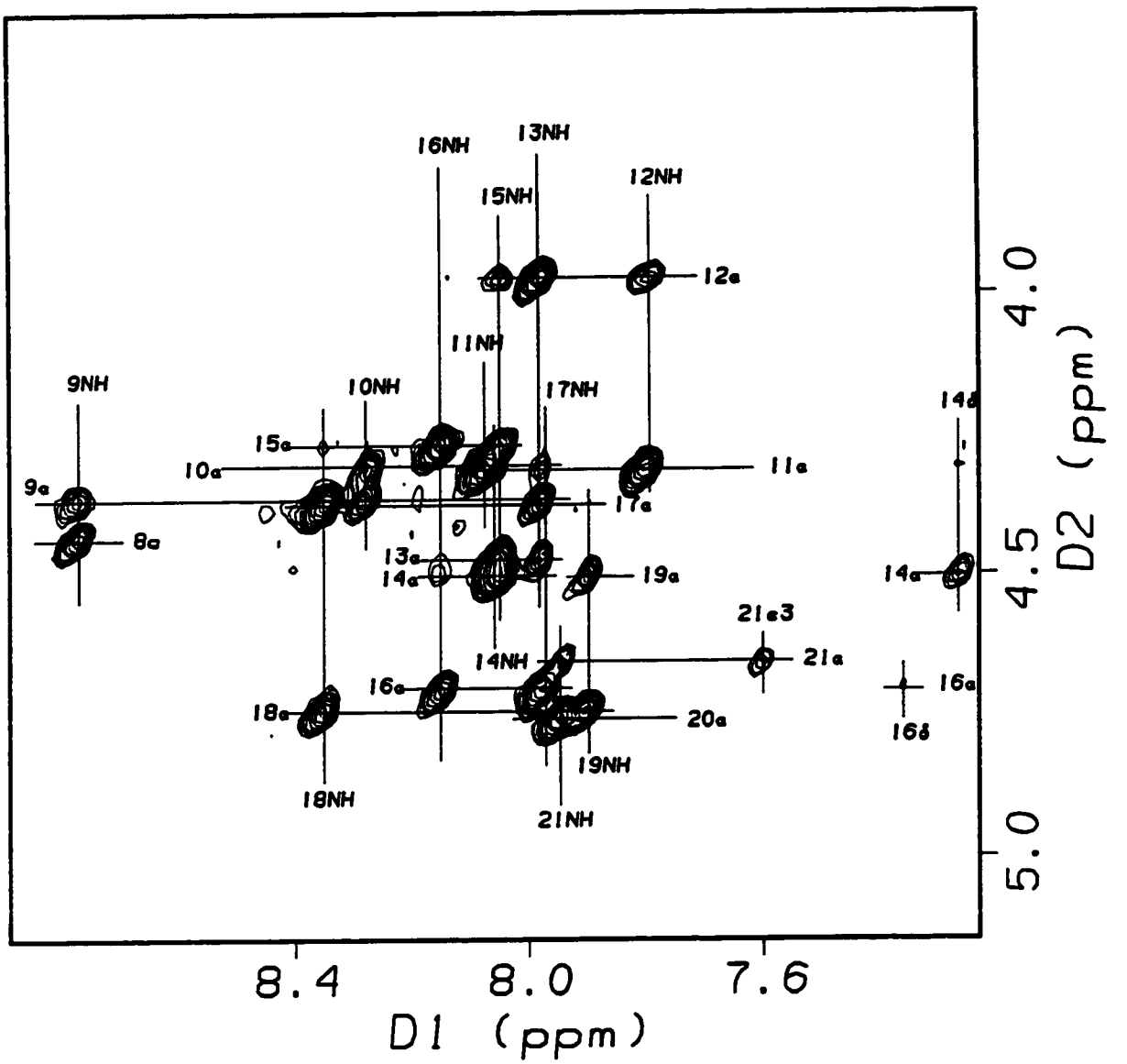


Figure 3.1: The amide NH → αH crosspeak region of the NMe-14mer NOESY in 40% aqueous acetic acid, acquired at 750MHz.

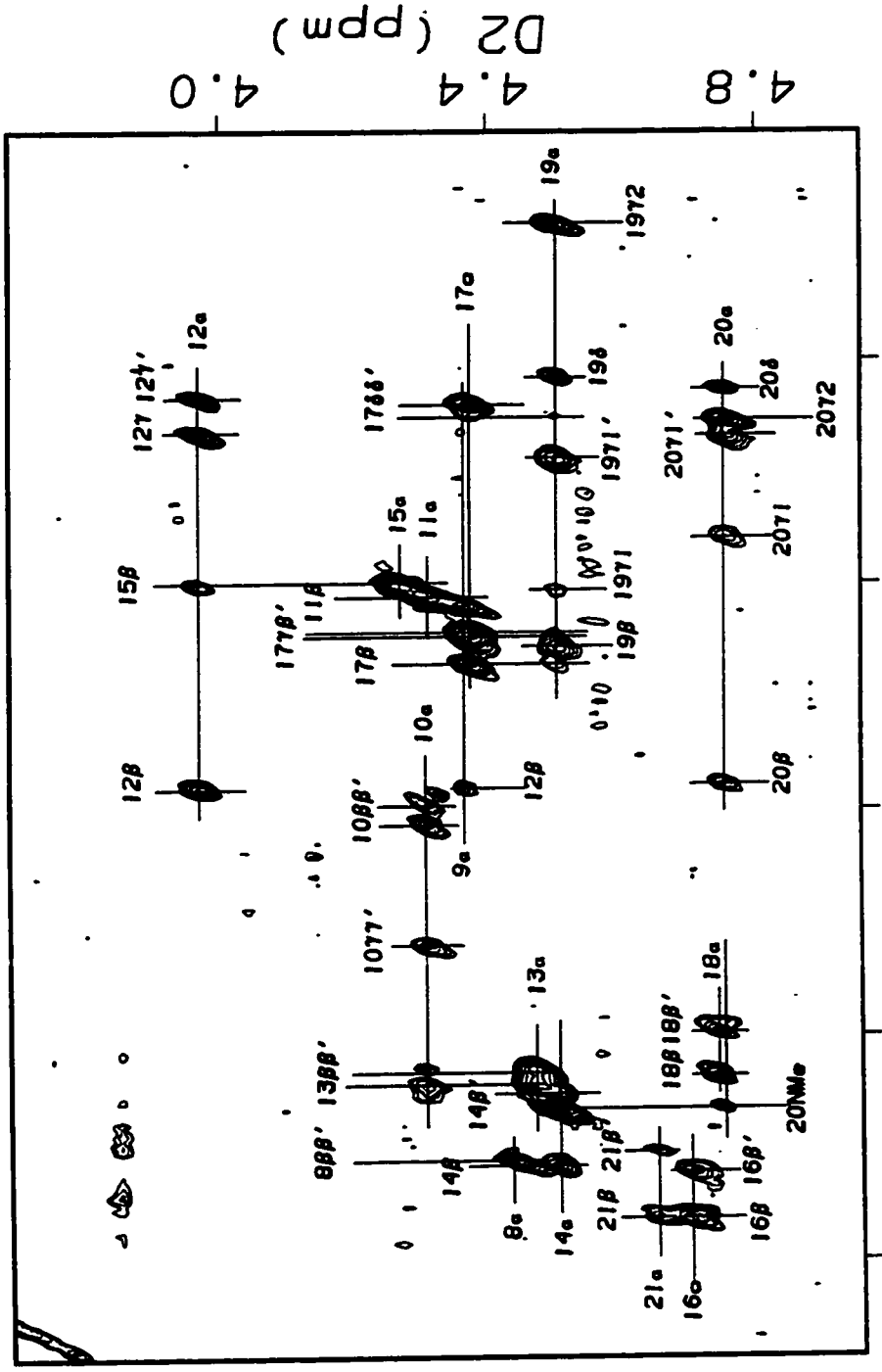


Figure 3.2: The αH → upfield crosspeak region of the NMe-14mer NOESY in 40% aqueous acetic acid, acquired at 750MHz. Note the strong 19α → 20NMe crosspeak.

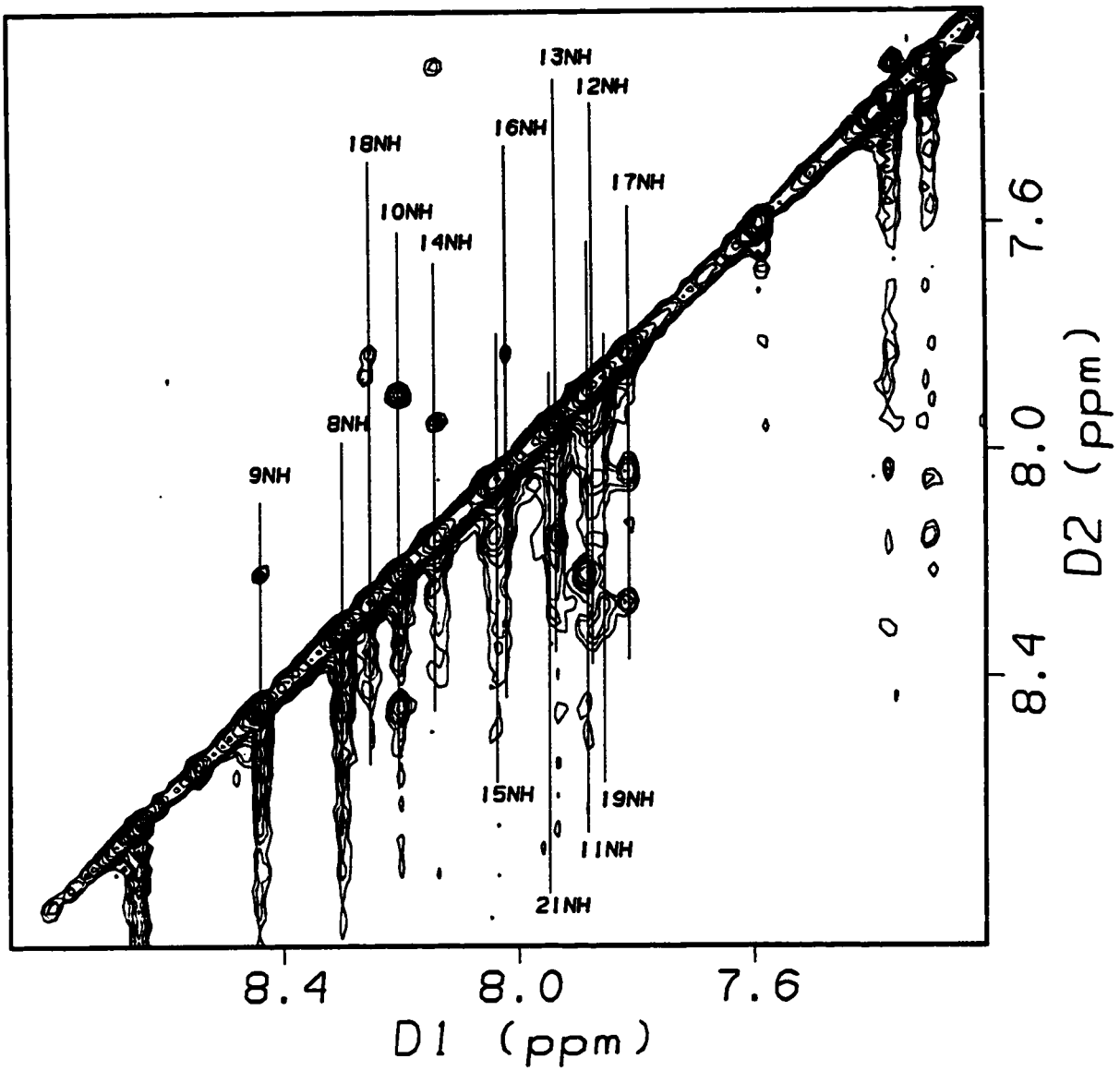


Figure 3.3: The amide NH crosspeak region of the AcNMe-14mer NOESY in 40% aqueous acetic acid, acquired at 750MHz. The aromatic proton assignments are omitted for clarity.

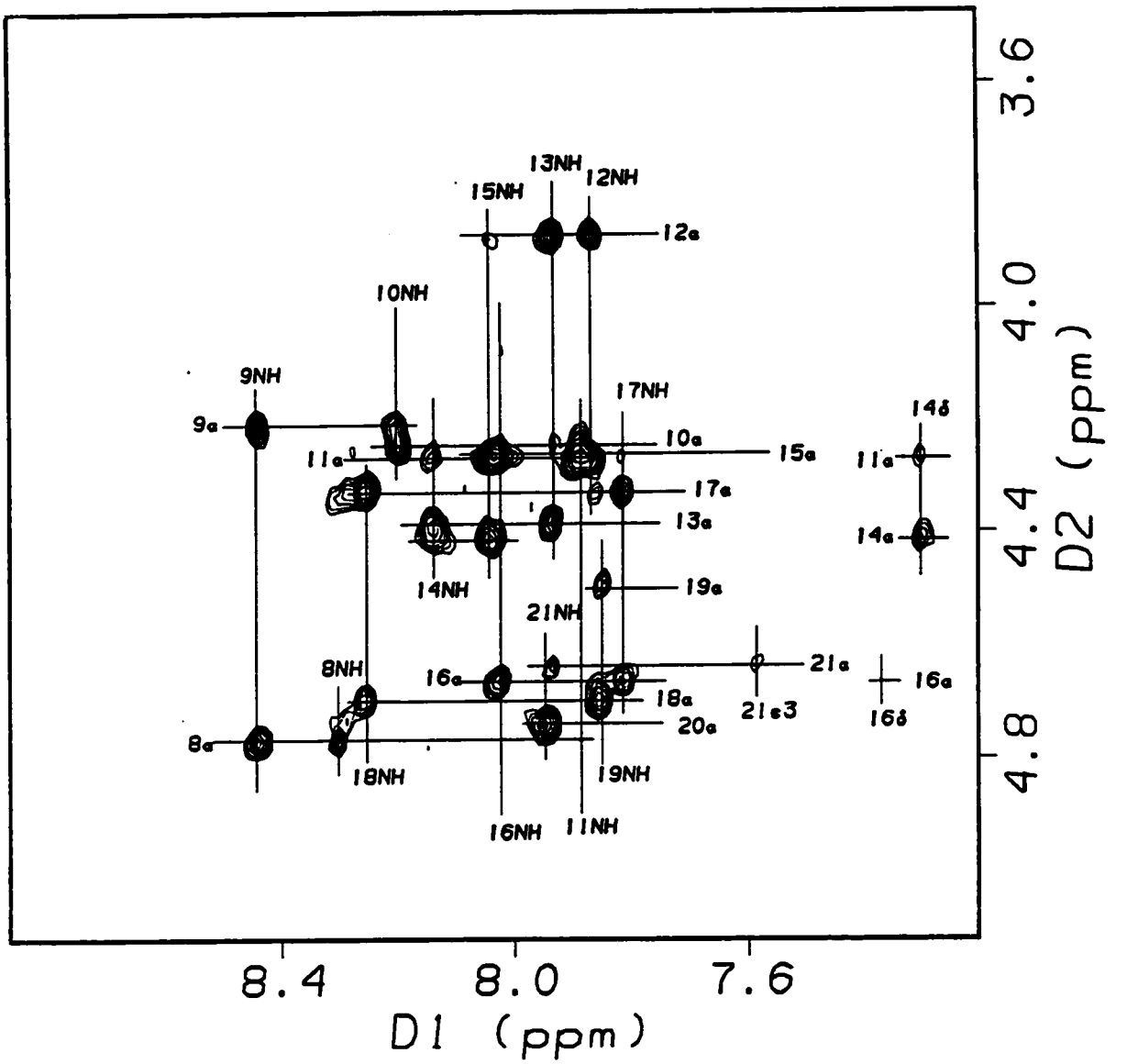


Figure 3.4: The NH → αH crosspeak region of the AcNMe-14mer NOESY in 40% aqueous acetic acid, acquired at 750MHz.

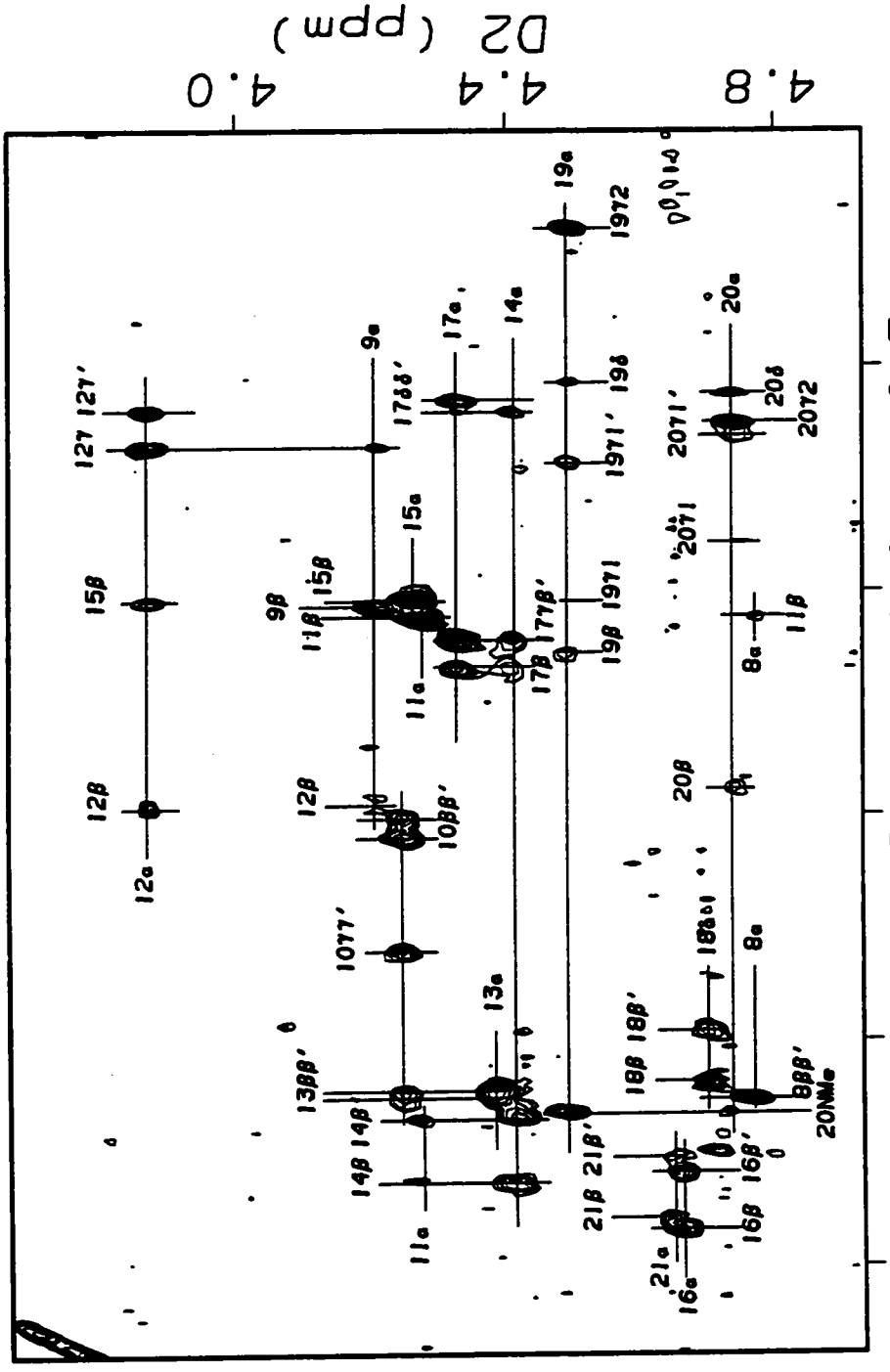


Figure 3.5: The αH → upfield crosspeak region of the AcNMe-14mer NOESY in 40% aqueous acetic acid, acquired at 750MHz. Note the strong 19α → 20NMe crosspeak.

3.3.3: HFIP TITRATION.

3.3.3.1: Effects on backbone amide and methyl group chemical shifts.

For both the NMe-14mer and AcNMe-14mer, addition of HFIP significantly stabilizes the overall structure. These spectra displayed higher signal-to-noise resolution due to increased solubility of the peptides in a less polar medium. This allowed for some previously weak crosspeaks to be more easily observed. Final chemical shift assignments for the NMe-14mer and AcNMe-14mer in aqueous 40% HFIP/24% acetic acid are listed in Tables 3.6 and 3.7. The amide region of the 1D-NMR spectra for both analogs (Figures 3.6 and 3.7) display net downfield chemical shift deviations for residues 13 through 15 (Table 3.8), indicating that interresidue hydrogen bonding within the helical region is strengthened. However, in the case of the NMe-14mer, the Tyr^{13} amide initially moves upfield prior to moving downfield at the higher HFIP concentrations. This artifact is not observed for the N-capped analog and would account for the overall smaller downfield deviation observed in the NMe-14mer (+0.045ppm from 0 $\rightarrow$ 40% HFIP, essentially a negligible effect) *versus* the AcNMe-14mer (+0.207ppm). The central residues of the helical region ($Ala^{11} \rightarrow Leu^{17}$) also display larger downfield (or smaller upfield) deviations in the case of the AcNMe-14mer. In both cases the NHs of Val^{12} move upfield with increasing HFIP concentration. Another curious feature appears during the HFIP titration: in the case of the NMe-14mer, the chemical shifts of 9NH and 10NH initially move downfield (from 0 $\rightarrow$ 20% HFIP), but then move upfield (from 20 $\rightarrow$ 40% HFIP). This may suggest that structuring effects, at least for the helix N-terminus reach their maximum at 20% HFIP. However, since the intrinsic effect of HFIP is to move chemical shifts upfield, it is more likely that HFIP-induced structuring effects initially dominates. At the high HFIP concentrations, solvent effects on the 9NH and 10NH chemical shifts is the dominant factor. This trend is less dramatic for the corresponding amides of the AcNMe-14mer. In this case, the N-terminus is already more

structured than its non-capped analog in aqueous medium. As a result, HFIP addition would have smaller structuring effects, at least at the N-terminus.

Table 3.6: Chemical shift assignments for DAEAVYFAHLDI$^{(NMe)}$IW in 40% HFIP, 24% acetic acid, 36% water at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	exchanged	4.444	3.371, 3.252	
Ala 9	8.768 [-5.897]	4.295	1.491	
Glu 10	8.251 [-5.420]	4.162	2.161, 2.161	γ,γ′ 2.580, 2.541
Ala 11	7.837 [-1.057]	4.291	1.553	
Val 12	7.556 [-2.829]	3.839	2.202	γ,γ′ 1.055, 0.939
Tyr 13	8.032 [-9.681]	4.335	3.076, 3.030	δ,δ′ 6.848, ε,ε′ 6.729
Phe 14	8.287 [-11.624]	4.403	3.307, 3.157	δ,δ′ 7.334, ε,ε′ 7.370, ζ 7.319
Ala 15	8.185 [-8.522]	4.270	1.514	
His 16	7.943 [-2.011]	4.603	3.450, 3.270	δ 7.365, ε 8.414
Leu 17	7.763 [-0.170]	4.336	1.726, 1.578	γ 1.606, δ,δ′ 0.878, 0.825
Asp 18	8.116 [-2.454]	4.742	2.937, 2.799	
Ile 19	7.746 [-4.636]	4.534	1.595	γ1,γ1′ 1.452, 1.025
				γ2 0.323, δ 0.791
$^{(NMe)}$Ile 20	n/a	4.685	2.064	γ1,γ1′ 1.293, 0.952
				γ2 0.890, δ 0.815, NMe 3.043
Trp 21	7.605 [-3.750]	4.659	3.377, 3.219	δ 7.209, ε1 9.487, ε3 7.613
				ζ2 7.401, ζ3 7.118, η2 7.190

Table 3.7: Chemical shift assignments for Ac-DAEAVYFAHLDI$^{(NMe)}$IW in 40% HFIP, 24% acetic acid, 36% water at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	7.780 [-2.625]	4.836	3.060, 3.007	Ac-Me 2.077
Ala 9	8.332 [-5.113]	4.132	1.478	
Glu 10	8.117 [-3.681]	4.121	2.189, 2.189	γ,γ′ 2.560, 2.560
Ala 11	7.688 [-1.227]	4.275	1.581	
Val 12	7.744 [-5.829]	3.776	2.203	γ,γ′ 1.073, 0.960
Tyr 13	8.149 [-8.726]	4.272	3.089, 3.089	δ,δ′ 6.813, ε,ε′ 6.714
Phe 14	8.435 [-9.885]	4.348	3.322, 3.202	δ,δ′ 7.380, ε,ε′ 7.389, ζ 7.316
Ala 15	8.344 [-8.556]	4.255	1.540	
His 16	7.914 [-2.148]	4.578	3.460, 3.267	δ 7.378, ε 8.382
Leu 17	7.704 [-0.307]	4.290	1.727, 1.572	γ 1.590, δ,δ′ 0.851, 0.798
Asp 18	8.073 [-2.863]	4.712	2.947, 2.804	
Ile 19	7.732 [-4.329]	4.531	1.605	γ1,γ1′ 1.457, 1.026
				γ2 0.320, δ 0.791
$^{(NMe)}$Ile 20	n/a	4.667	2.063	γ1,γ1′ 1.308, 0.965
				γ2 0.891, δ 0.820, NMe 3.045
Trp 21	7.615 [-3.681]	4.638	3.367, 3.213	δ 7.208, ε1 9.475, ε3 7.614
				ζ2 7.417, ζ3 7.120, η2 7.192

Table 3.8: HFIP Titration – Amide NH chemical shift differences

Residue	NMe-14mer Δδ (40% HFIP - 0% HFIP)	AcNMe-14mer Δδ (40% HFIP – 0% HFIP)
Asp^{8}	n/a	-0.524
Lys^{9}	-0.003	-0.113
Glu^{10}	-0.028	-0.090
Ala^{11}	-0.236	-0.208
Val^{12}	-0.244	-0.133
Tyr^{13}	+0.045	+0.207
Phe^{14}	+0.229	+0.292
Ala^{15}	+0.128	+0.292
His^{16}	-0.208	-0.119
Leu^{17}	-0.219	-0.115
Asp^{18}	-0.233	-0.186
Ile^{19}	-0.156	-0.129
$^{(NMe)}Ile^{20}$	n/a	n/a
Trp^{21}	-0.344	-0.339

Table 3.9: HFIP Titration – Other reporter groups

Residue	NMe-14mer Δδ (40% HFIP - 0% HFIP)	AcNMe-14mer Δδ (40% HFIP – 0% HFIP)
Ala 9β	+0.034	+0.026
Ala 11β	***+0.118***	***+0.097***
Val 12γγ′	***+0.129, +0.124***	***+0.121, +0.113***
Val 15β	***+0.112***	***+0.103***
Leu 17δδ′	+0.004, -0.010	+0.009, -0.015
Ile 19γ2	+0.047	+0.050
Ile 19δ	+0.036	+0.040
$^{(NMe)}$Ile 20γ2	+0.005	+0.012
$^{(NMe)}$Ile 20δ	+0.024	+0.033

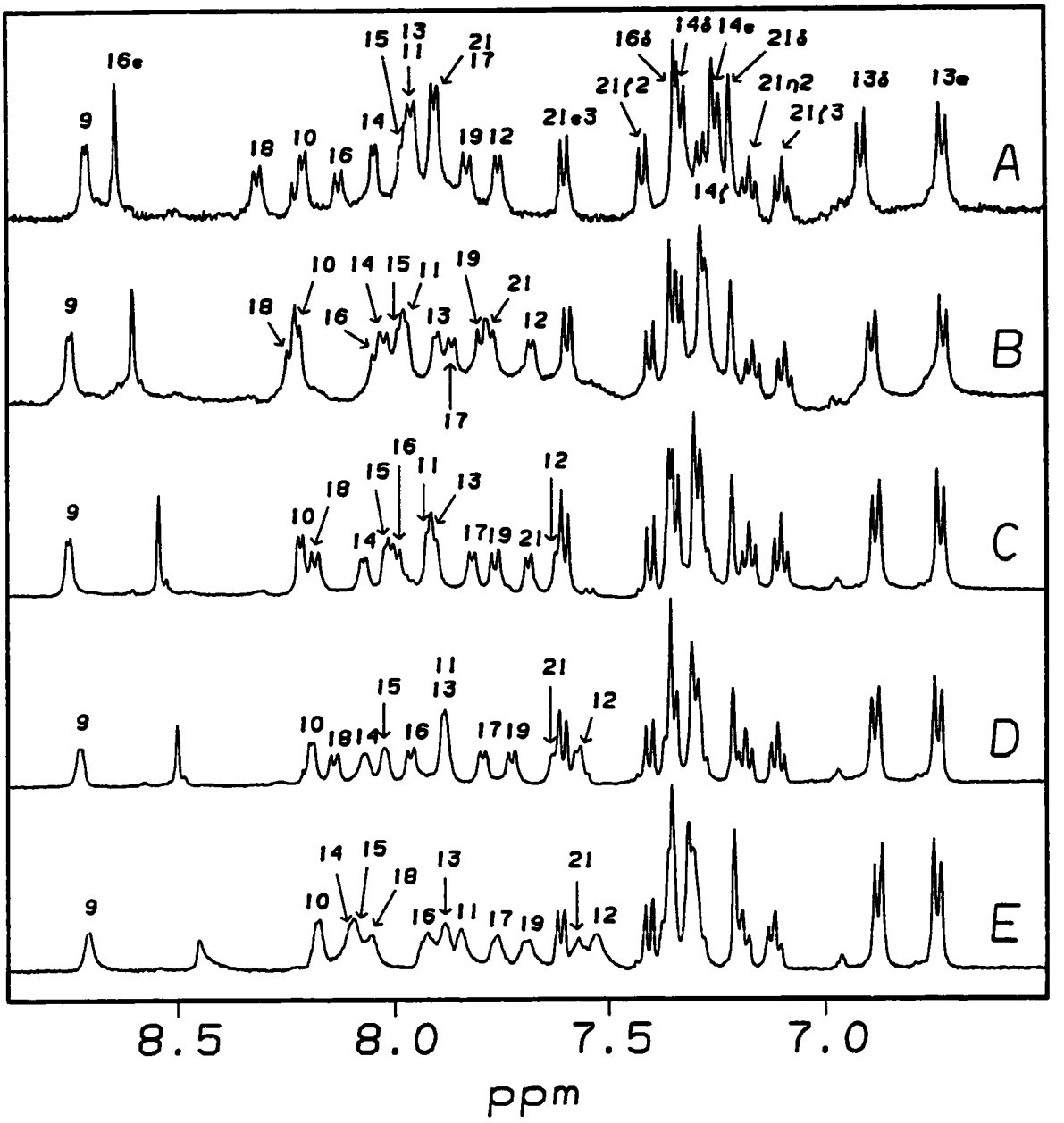


Figure 3.6: The amide NH region of the NMe-14mer HFIP titration. Points in the titration are 0% (A), 10% (B), 20% (C), 30% (D) and 40% HFIP (E) by volume.

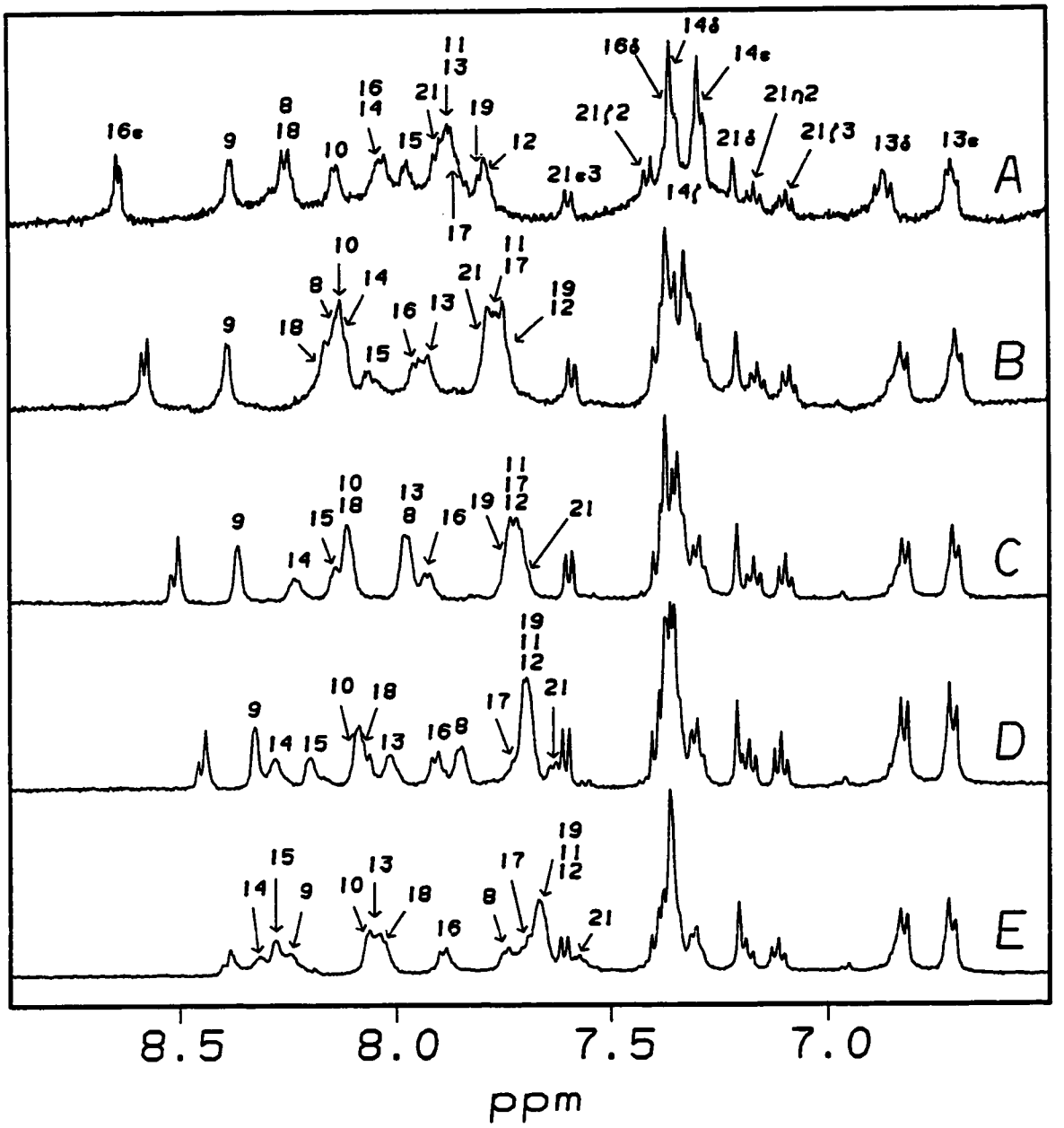


Figure 3.7: The amide NH region of the AcNMe-14mer HFIP titration. Points in the titration are 0% (A), 10% (B), 20% (C), 30% (D) and 40% HFIP (E) by volume.

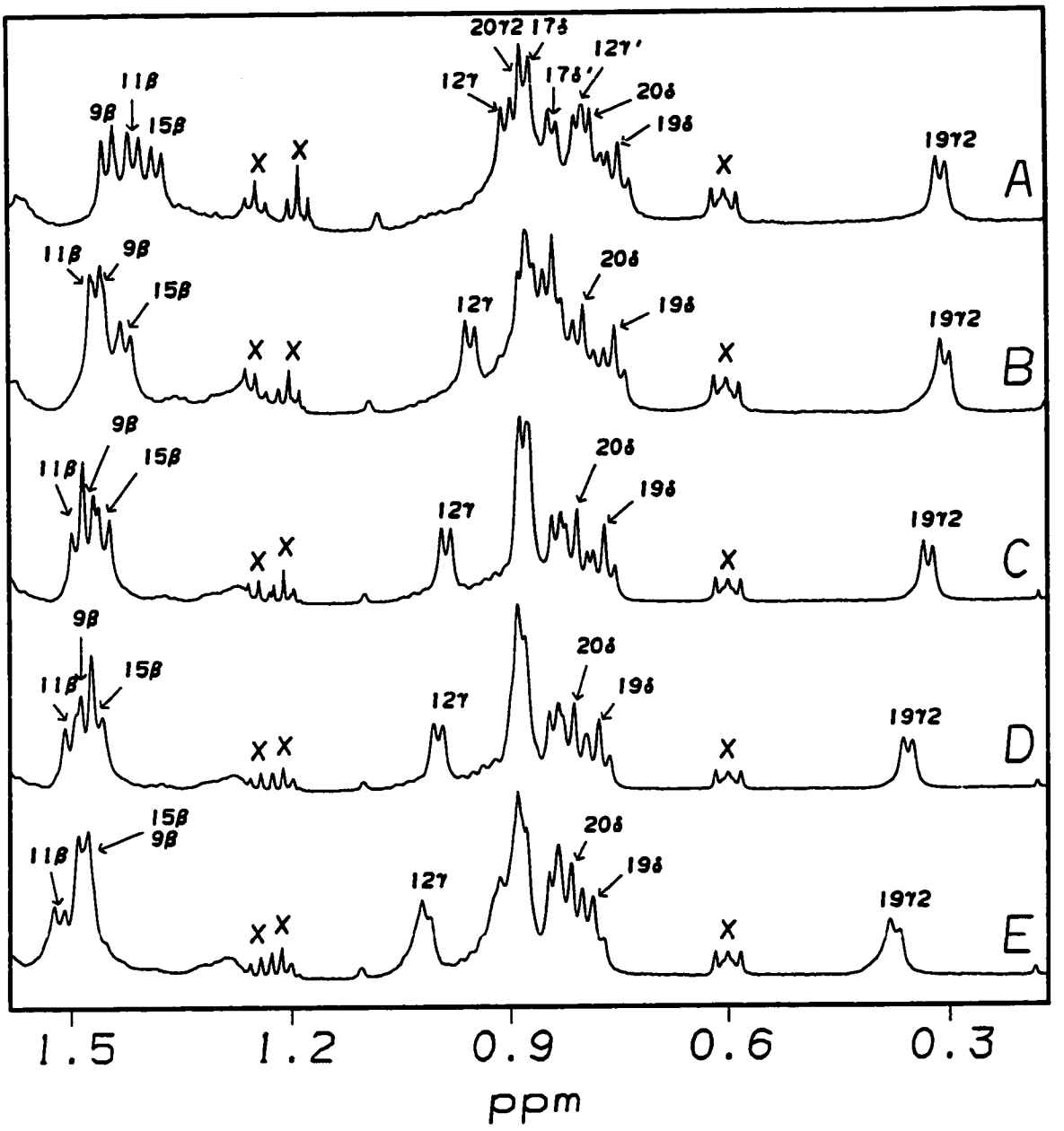


Figure 3.8: The methyl region of the NMe-14mer HFIP titration. Points in the titration are 0% (A), 10% (B), 20% (C), 30% (D) and 40% HFIP (E) by volume.

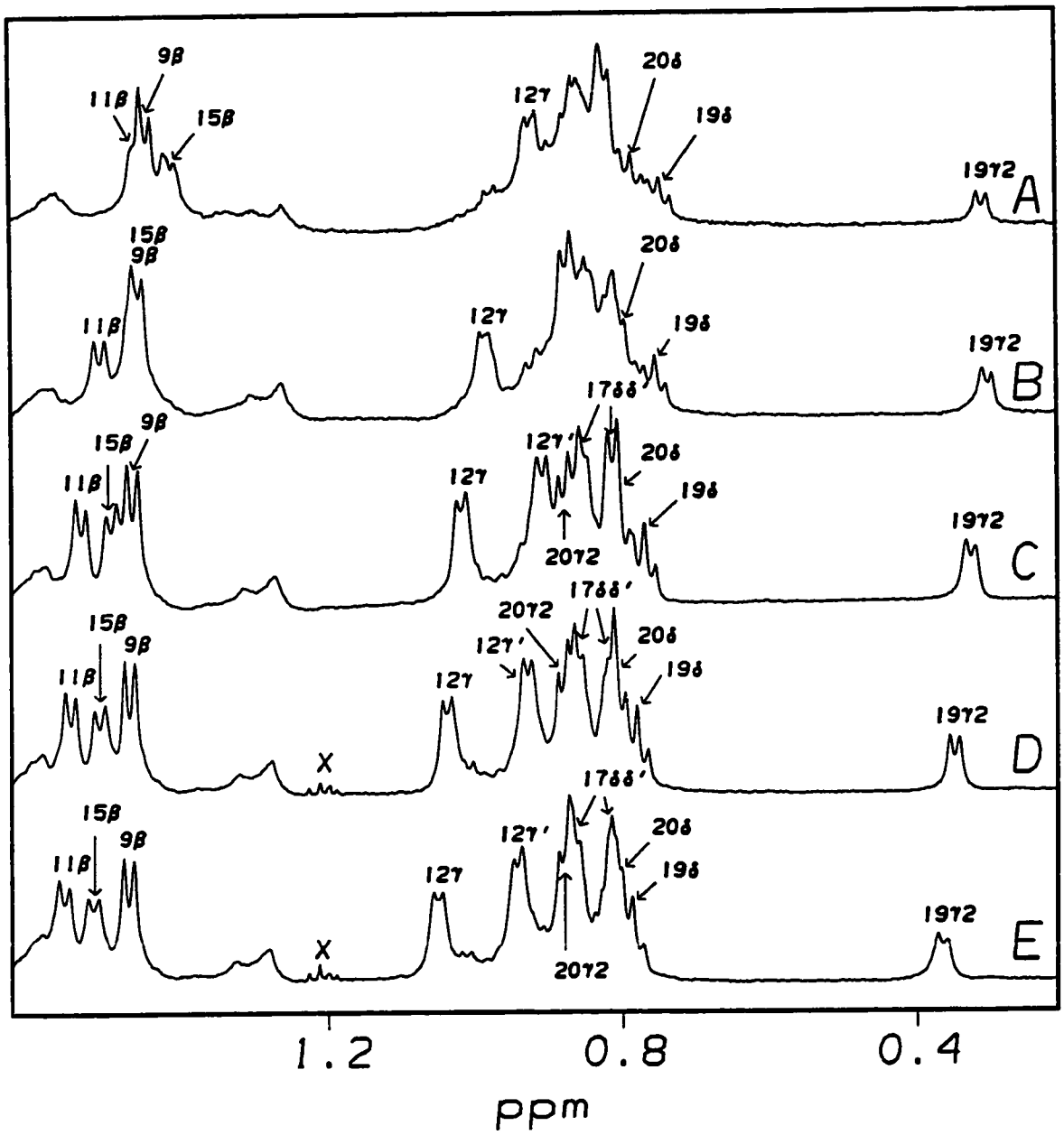


Figure 3.9: The methyl region of the AcNMe-14mer HFIP titration. Points in the titration are 0% (A), 10% (B), 20% (C), 30% (D) and 40% HFIP (E) by volume.

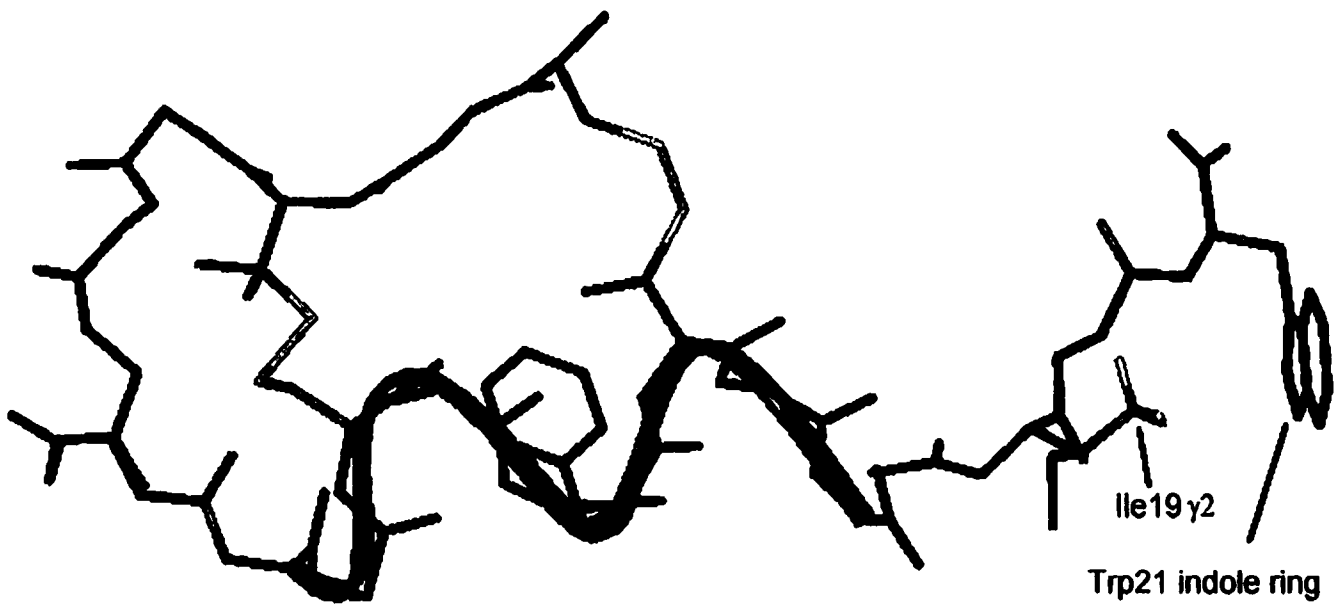


Figure 3.10: A member of the Pen-1 NMR structure ensemble which displays the C-terminal hydrophobic cluster. The sidechains of Ile^{19} and Trp^{21} are shown.

Also evident in the amide region of the spectra are the "2-effect" and "3-effect" of N-capping. The 9NH is the furthest downfield of the backbone amides in the NMe-14mer, with a chemical shift of 8.771ppm in 0% HFIP. The same proton has a chemical shift of 8.332ppm in 0% HFIP, an upfield deviation of 0.439ppm in the N-capped version. The chemical shifts of 10NH in 0% HFIP for the corresponding peptides are 8.279 and 8.207ppm, respectively, an upfield deviation of 0.072ppm. In 40% HFIP, the "2-effect" is comparable (0.436ppm upfield with N-capping), while the "3-effect" nearly doubles (0.134ppm upfield with N-capping).

The upfield region of the spectra (Figures 3.8 and 3.9) also clearly demonstrates the HFIP effect on peptide structure. Those residues in the central portion of the helix are more greatly affected by HFIP than either the N- or C-terminal regions of the peptide. This can be probed using the methyl peaks, most of which move downfield with increasing fluoroalcohol concentrations. As shown in Table 3.9, the sidechain methyl groups of residues 11, 12 and 15 all have downfield chemical shift deviations greater than or equal to 0.10ppm. Other residues on the periphery of the helix and the C-terminal region do not have as dramatic adjustments in their methyl group chemical shifts. None of the methyl groups for Ala^{9}, Leu^{17}, Ile^{19} and $^{(NMe)}Ile^{20}$ shift more than 0.050ppm downfield when the solution contains 40% HFIP.

In both NMe-14mer analogs, the hydrophobic clustering of the C-terminus is still present at 40% HFIP. The slight downfield chemical shift deviation observed suggests that the cluster slightly destabilizes upon increasing the concentration of the fluorinated alcohols. As the medium becomes less "lipophobic", hydrophobic interactions between the Ile^{19} aliphatic sidechain and the Trp^{21} indole ring become less important. As the sidechain protons of Ile^{19}, specifically those of the β-methyl group, move away from the shielding cone of the Trp^{21} indole ring, their chemical shifts move downfield towards nominal methyl group values. The relative positions of the Ile^{19} and Trp^{21} sidechains are

shown in Figure 3.10, which displays one member of the Pen-1 NMR structure ensemble. Although this structure doesn't include the $^{(NMe)}$Ile^{20}, it does show that the 19γ2 methyl group lies within the shielding cone of the Trp^{21} indole ring. In the case of the N-methylated analog, this interaction is stronger (i.e. the sidechains are positioned closer together), which is manifested by the unusual upfield chemical shift of 19γ2.

3.3.3.2: HFIP effects on the NOESY spectra.

At 40% HFIP, a number of interresidue crosspeaks become more intense over a greater portion of the sequence for the N-methylated analogs. Strong $NH_i \rightarrow NH_{i+1}$ crosspeaks from Ala^9 through Asp^{18} develop in the NOESY spectrum (Figures 3.11 and 3.12) for both peptides. Intermediate range NOEs also are apparent in the case of the NMe-14mer (Figure 3.13), for example, weak $i \rightarrow i+2$ (10α → 12NH, 12α → 14NH and 14α → 16NH) crosspeaks appear. Over the entire helical region, which is extends from Lys^9 through Asp^{18}, intraresidue α_iN_i crosspeaks are more intense than their sequential α_iN_{i+1} counterparts. In addition, $i \rightarrow i+4$ (9α → 13NH and 10α → 14NH), as well as $i \rightarrow i+3$ interactions between αH and sidechain aromatic protons (10α → 13δδ′, 11α → 14δδ′ and 13α → 16δ) are present.

The upfield region of the NOESY (Figure 3.14) display a number of intense $\alpha_i\beta_{i+3}$ crosspeaks. Also notable in this portion of the spectrum is the strong 19α → 20NMe crosspeak, which is more easily observed in 40% HFIP since the 14α and 19α chemical shifts are no longer degenerate. Another prominent crosspeak between 19γ2 and 21α, provides further evidence of C-terminal hydrophobic cluster formation. Several crosspeaks between 19γ2 and the sidechain protons of Trp^{21} are also present (Figure 3.15).

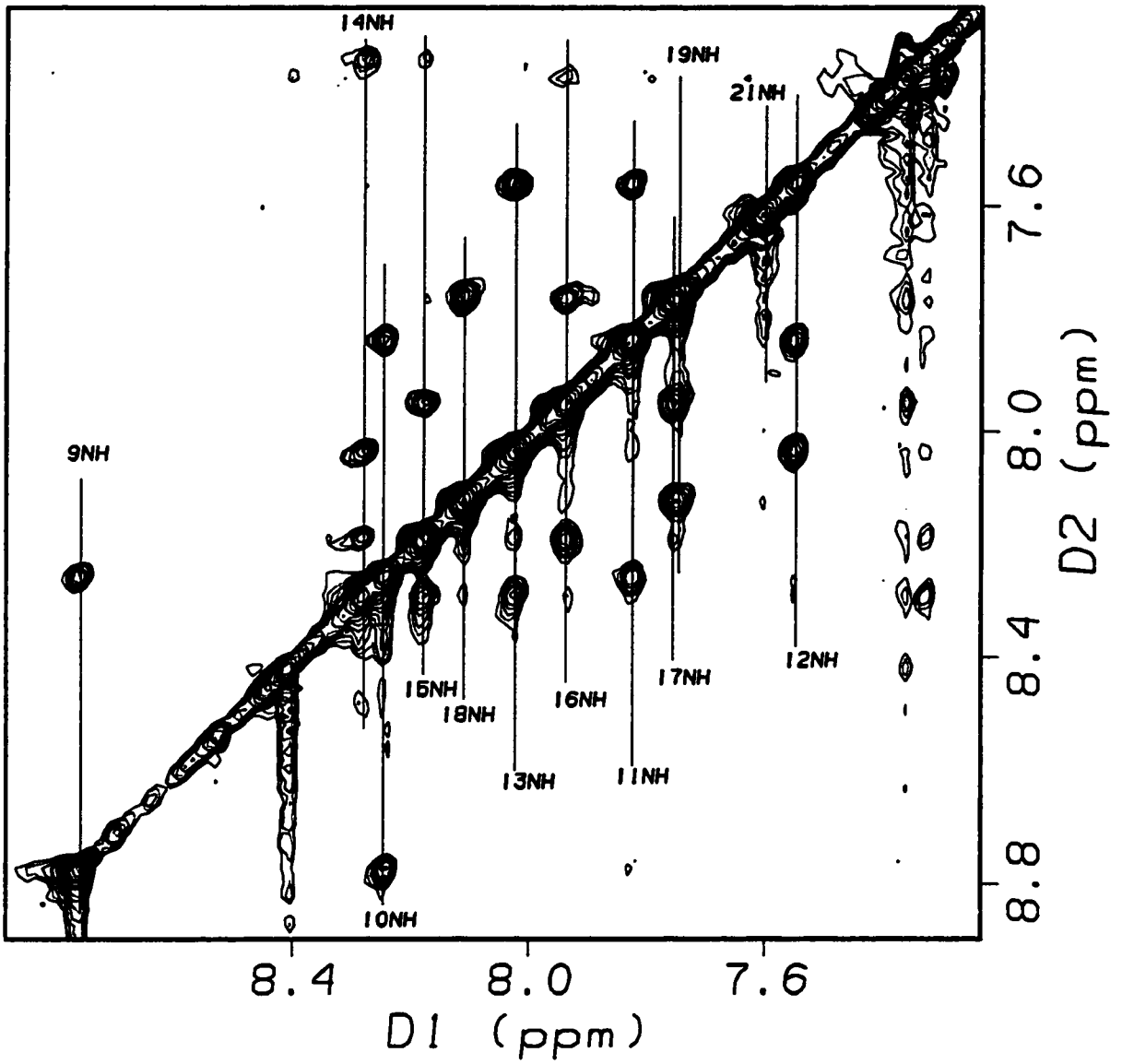


Figure 3.11: The amide NH crosspeak region of the NMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz. The aromatic protons are omitted for clarity.

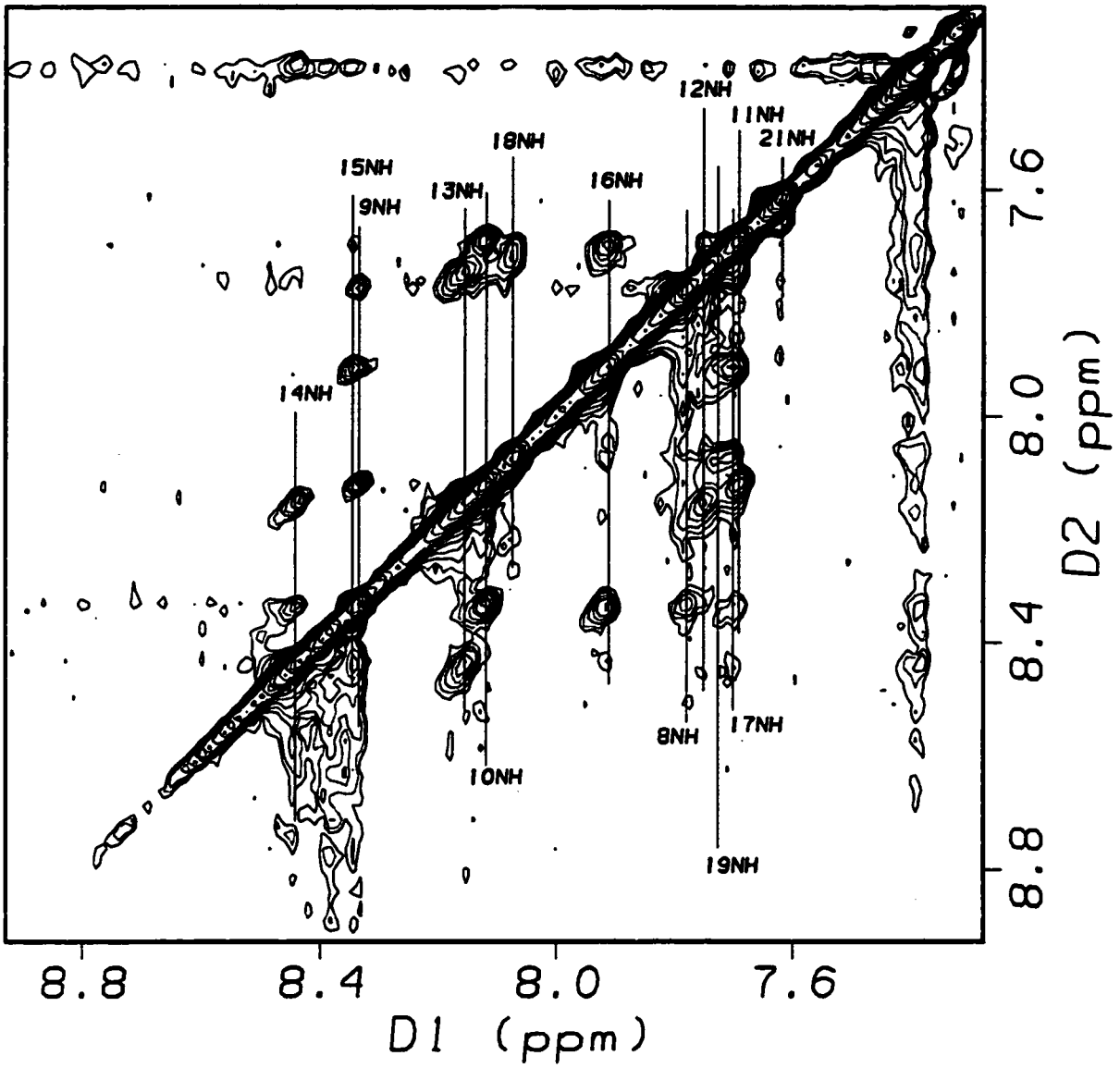


Figure 3.12: The amide NH crosspeak region of the AcNMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz. The aromatic protons are omitted for clarity.

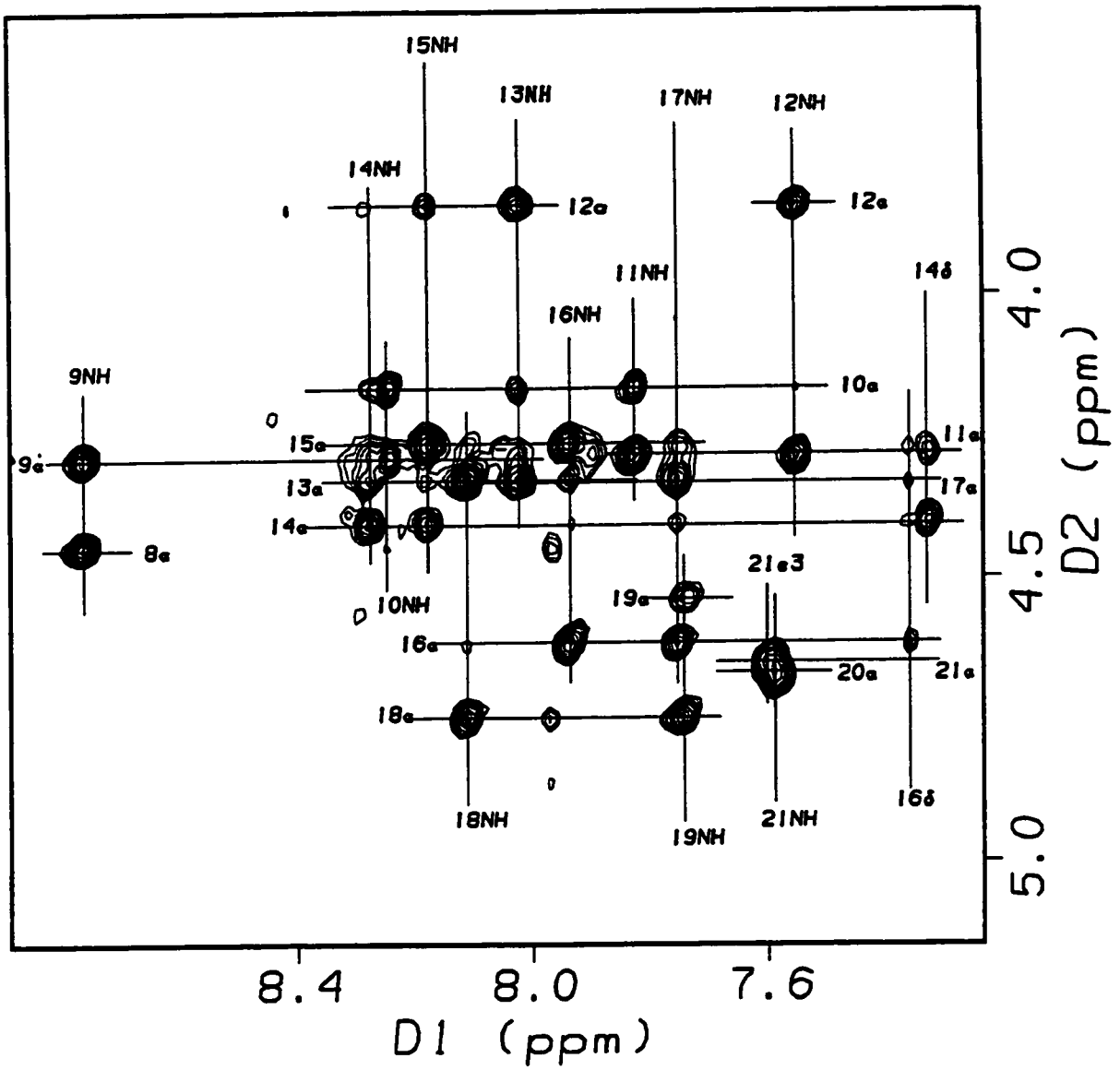


Figure 3.13: The amide NH → αH crosspeak region of the NMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz.

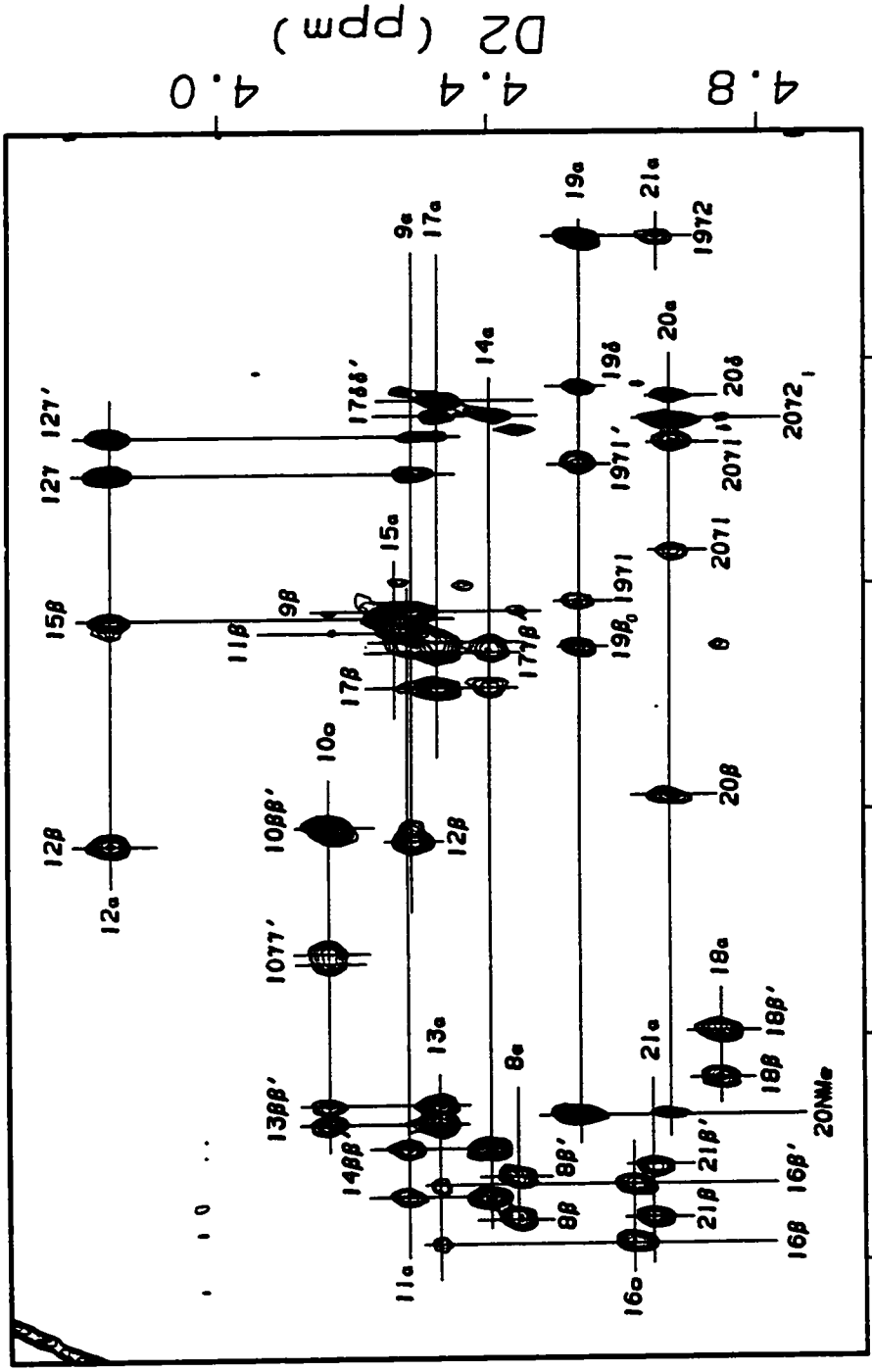


Figure 3.14: The αH → upfield crosspeak region of the NMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz. Note the strong 19α → 20NMe and 19γ2 → 21α crosspeaks.

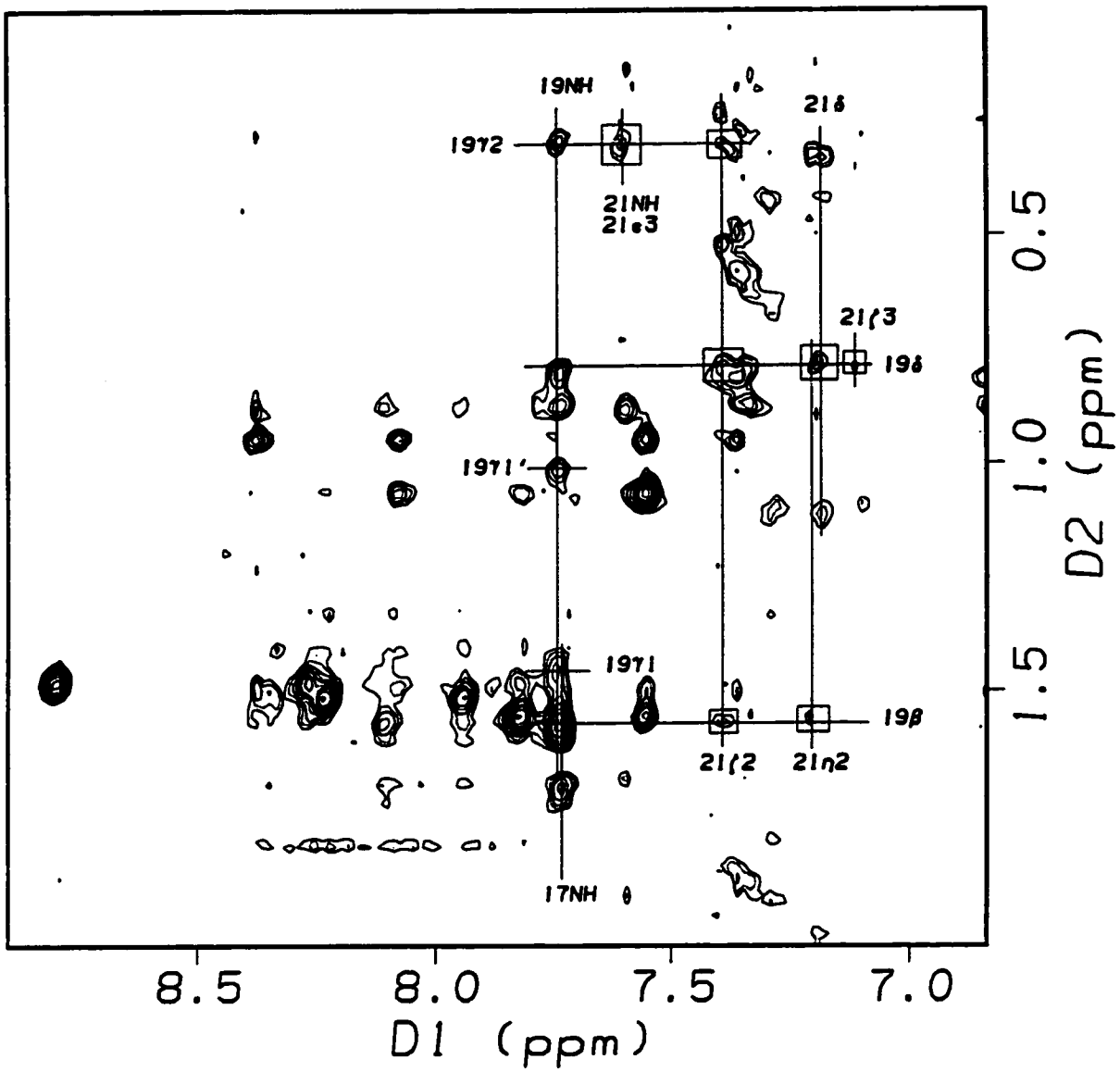


Figure 3.15: The amide-NH/aromatic → upfield crosspeak region of the NMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 499MHz. Note the crosspeaks between the Ile^{19} and Trp^{21} sidechains.

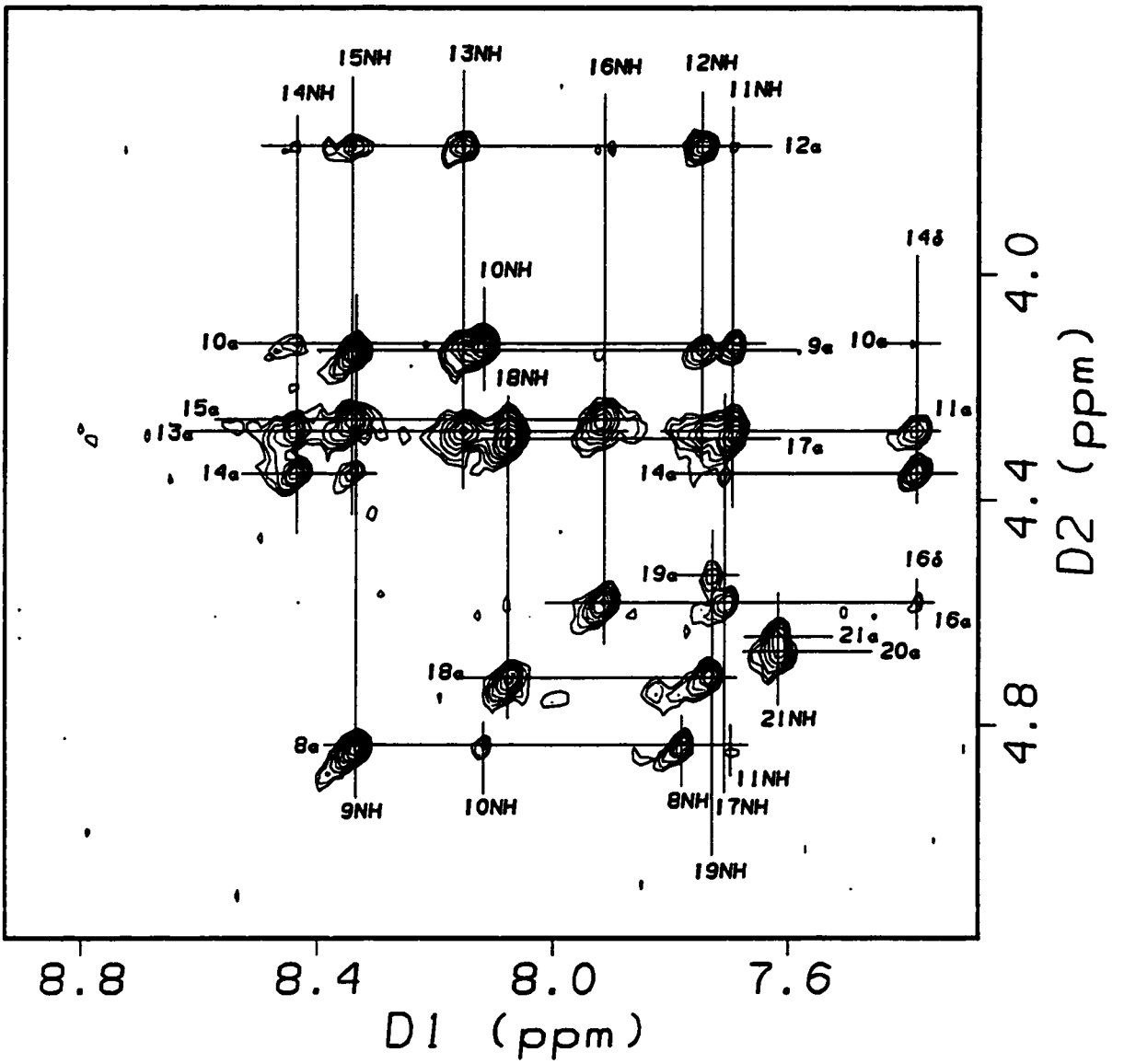


Figure 3.16: The amide NH → αH crosspeak region of the AcNMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz.

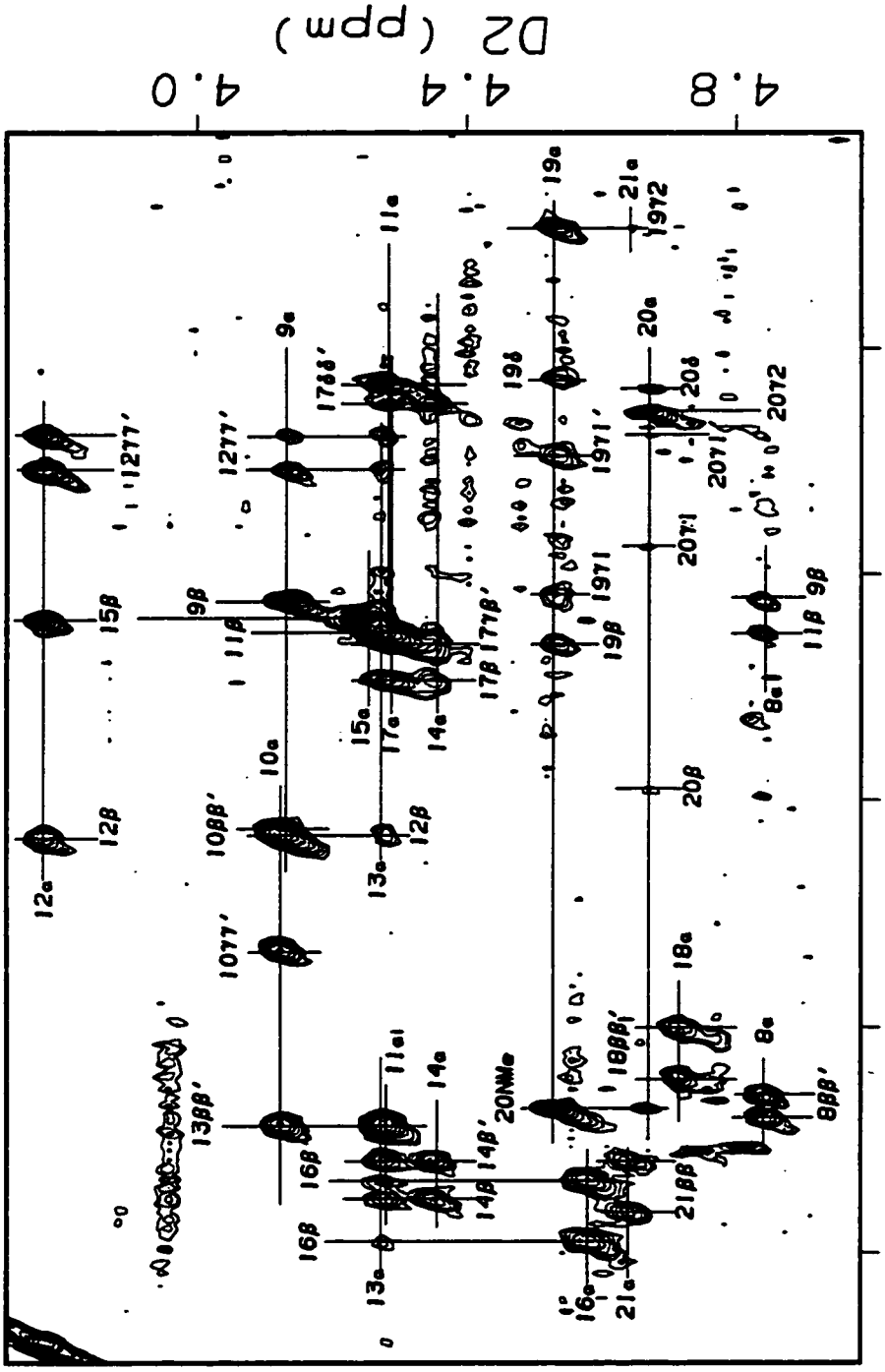


Figure 3.17: The αH → upfield crosspeak region of the AcNMe-14mer NOESY in aqueous 40% HFIP/24% acetic acid, acquired at 750MHz. Note the strong 19α → 20NMe and weaker 19γ2 → 21α crosspeaks.

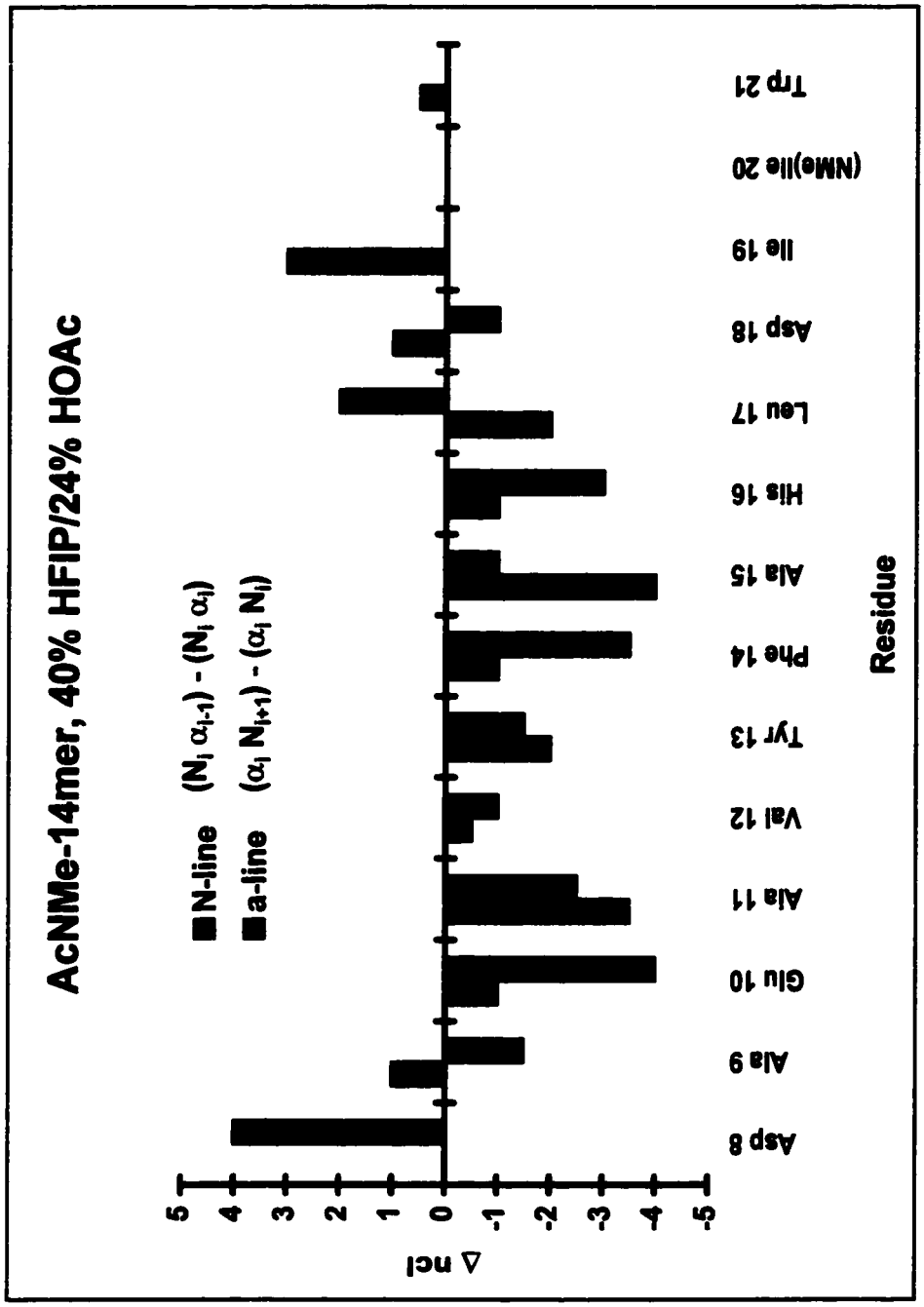


Figure 3.18: The NOE ratio plot of the AcNMe-14mer in aqueous 40% HFIP/24% acetic acid. Units are in contour level differences (Δncl) for the N- and α-lines.

Figure 3.19: The NOE connectivity and hydrogen exchange plot for the AcNMe-14mer in aqueous 40% HFIP/24% acetic acid. Asterisks correspond to the 20NMe line. Exchange rates are as follows: open circle, < 10min, slashed circle, 10min, half-filled circle, 20min and filled circle, 40-60min.

The acetylated NMe-14mer analog displays a more stable helical region, with relatively stronger, and more extensive web of, $i \rightarrow i+3$ and $i \rightarrow i+4$ crosspeaks (Figures 3.16 and 3.17). As in the non-capped analog, intraresidue $\alpha_i N_i$ crosspeaks are more intense than their sequential $\alpha_i N_{i+1}$ counterparts over residues 9 through 18. NOE ratios and connectivity data are summarized in Figures 3.18 and 3.19. Many of the same NOE ratios are less intense, that is, the difference in contour levels between the intra- and sequential interresidue αN crosspeaks are smaller, in the case of the NMe-14mer (data not shown). Both plots clearly indicate the presence of a stable helix spanning Ala^9 through Leu^{17}, capped on either end by aspartic acid residues.

3.3.3.3: HFIP induced structuring effects evidenced by αH-CSDs.

A comparison of the αH-CSDs (Figure 3.20, top) indicates that the NMe-14mer fragments aren't as structured as the intact ET-1 analogs. Although the helical region of the peptides range from Lys^9 through His^{16} in the fragments, most of the αH-CSD values are less than half of those observed for the same residues of Pen-1 and Pen-2. In addition, the helical region is more extended in the intact peptides ($Lys^9 \rightarrow Asp^{18}$). This would suggest that the disulfide bonds help stabilize, but are not essential for, helix formation. The large difference in structuring may also be due to solvent effects. The C-terminal residues (Ile^{19} through Trp^{21}) aren't as greatly affected by sequence truncation: with the exception of 21α, the αH-CSD values for the N-methylated analogs are completely within experimental error. Also observed in this plot is evidence that the acetyl N-cap increases stability of the helical region of the fragments, as indicated by the more negative αH-CSD values of the AcNMe-14mer. Corrections for local N-methylation shift effects were employed in Figure 3.20. Determination of these N-methyl correction factors is described in Section 3.6.4.

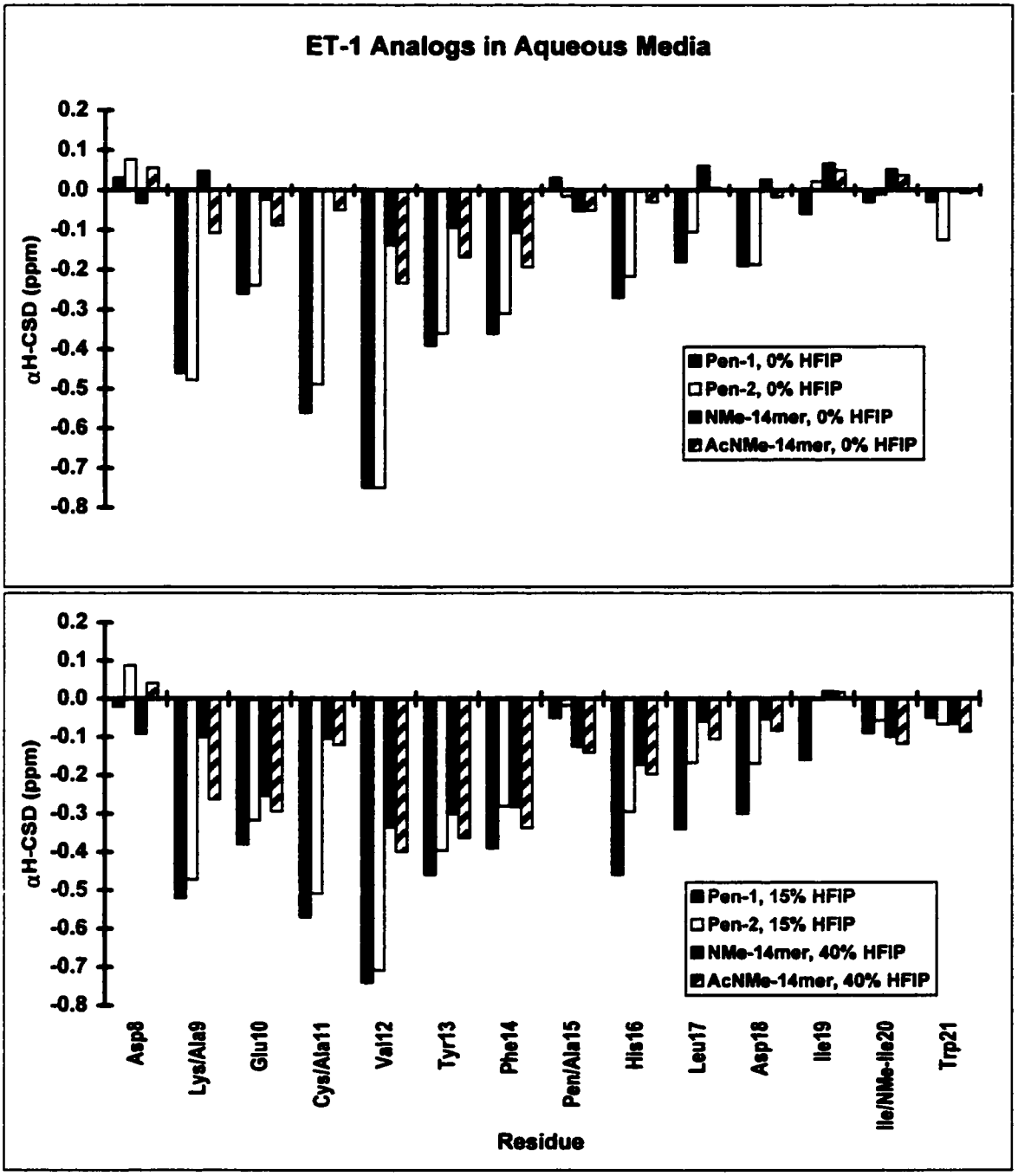


Figure 3.20: The αH-CSD plot of the intact and large fragment ET-1 analogs with (top panel) and without (bottom panel) the presence of fluoroalcohols. Residues 1 through 7 are removed for clarity.

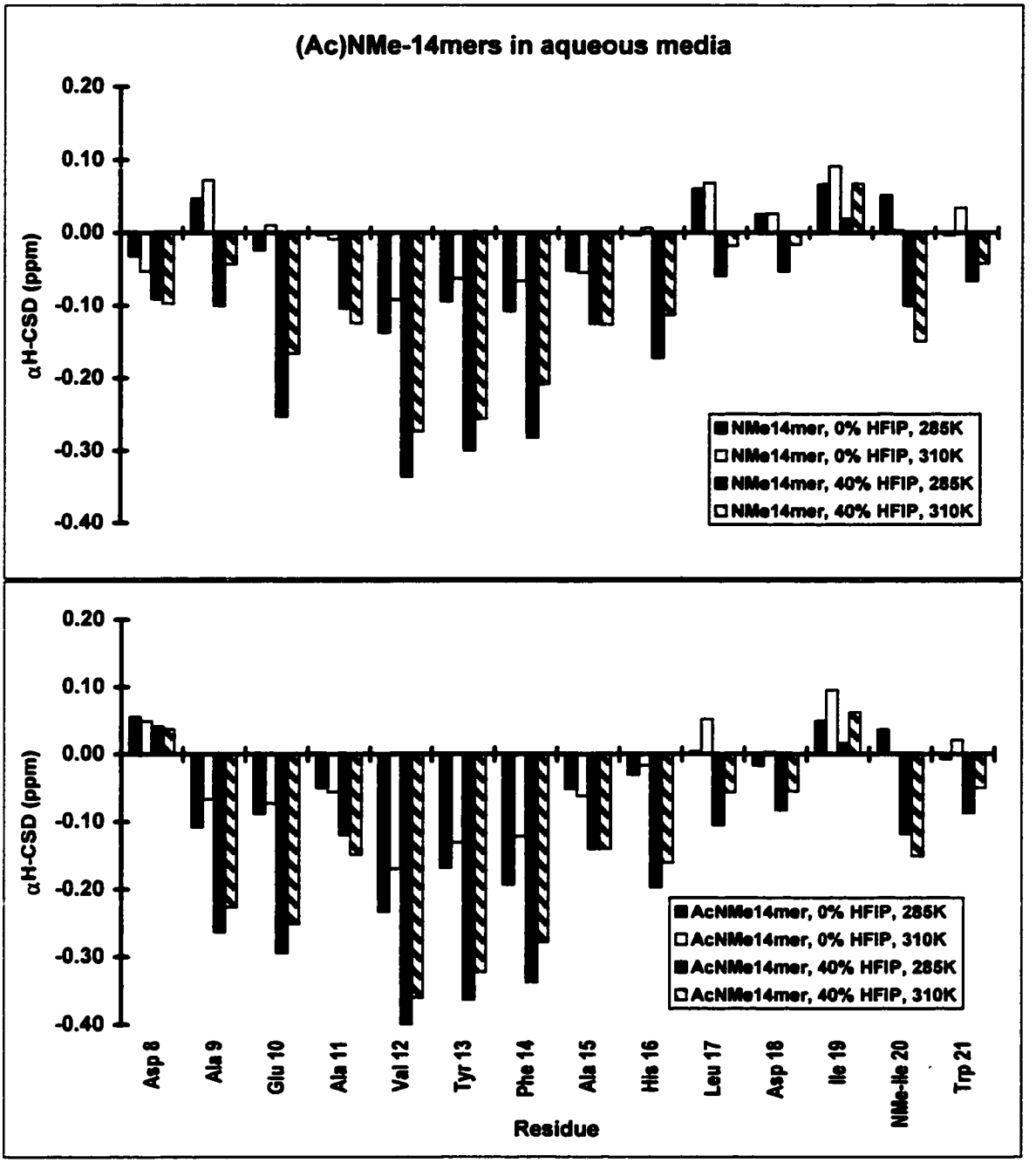


Figure 3.21: Temperature and fluoroalcohol effects on the αH-CSD values of the NMe-14mer (top panel) and AcNMe-14mer (bottom panel).

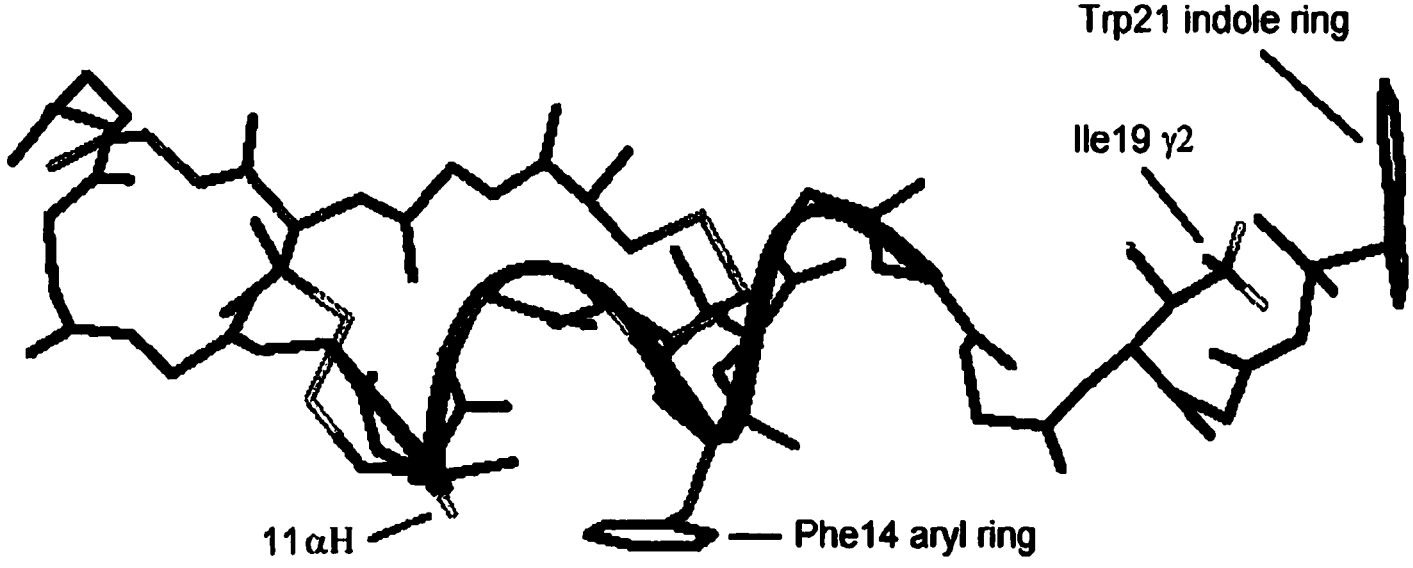


Figure 3.22: A member of the Pen-1 NMR ensemble which displays the C-terminal hydrophobic cluster. The sidechains of Ile^{19} and Trp^{21} are shown. The α-methine proton of residue 11 is shown to be co-planar with the sidechain aromatic ring of Phe^{14}.

Notable differences are apparent in the αH-CSD values when the peptides are studied in the presence of fluorinated alcohols (Figure 3.20, bottom). As with the intact peptides, high concentrations of HFIP increase the stability of the helix, as well as extending it to Asp^{18}. However, as in the 0% HFIP case, the NMe-14mer still appears to be less structured than the AcNMe-14mer: for residues 9 through 18, the αH-CSDs are more negative for the acetylated fragment. The large differences in the Asp^{8} and Ala^{9} αH-CSD values between the NMe-14mer and AcNMe-14mer are solely due to the acetyl N-cap. Residues in the center of the helical region of the fragments (13 → 15) now have comparable values with respect to the intact peptides. With the exception of Glu^{10}, the residues at the ends of the helix (Ala^{9} → Val^{12} and His^{16} → Asp^{18}) have αH-CSDs nearly half of those corresponding to Pen-2. This would suggest that the ends of the helix display greater fraying in the fragments.

A comparison of the temperature effects on the αH-CSD plots reveals an interesting pattern. For both the $^{(NMe)}Ile^{20}$ analogs, most of the αH-CSDs move towards their disorder reference values (Figure 3.21). However, the opposite trend is observed for 11α, 19α and 20α. The latter two are artifacts of the 20NMe group, which appears to enhance the hydrophobic interactions at the C-terminus. The larger CSD values observed for 19αH and 20αH may reflect an increase in the hydrophobic effect, which is known to increase with temperature (Privalov, 1992). That is, the hydrophobic cluster, whose backbone resembles a turn-like structure (Harris, 1993; Chapter 4, this dissertation), is enhanced when the temperature is raised.

The trend observed for the Ala^{11} αH is most likely due to the Phe^{14} sidechain aromatic ring. As previously noted, 11α → 14δδ′ NOESY crosspeaks are present at both 0% and 40% HFIP. The relative positions of the Ala^{11} αH and the Phe^{14} phenyl ring are shown in Figure 3.22 (this is the same structure presented in Figure 3.10). In this side

view, 11α is co-planar with the phenyl ring, resulting in a strongly deshielded methine proton. As the temperature is raised, the Phe[14] side chain may experience increased mobility, and the 11α proton becomes less deshielded. If Phe[14] were substituted with a non-aromatic residue, the Ala[11] αH would most likely be further upfield than currently observed.

3.3.4: D_2O Exchange and H_2O Re-incorporation.

A deuterium exchange study (Figure 3.23; see also Figure 3.19) was performed on the most structured of the 14mer analogs, AcNMe-14mer in 40% aqueous HFIP/24% acetic acid at 300K. Table 3.10 reports the exchange half-lives and protection factors.

Table 3.10: AcNMe-14mer NH exchange half lives and estimated exchange protection factors (300K, pH ~ 2.3) [1]

Residue	Molday Factor [2]	Correction Factor	Est. Predicted $t_{1/2}$	Observed $t_{1/2}$	Protection Factor [3]
A (9)	-0.12	1.00 (ref)	~ 3.0 min	~ 3 min	1.00
E (10)	-0.60	3.02	9.1 min	~ 5 min	0.55
A (11)	-0.27	1.41	4.2 min	~7-10 min	1.67 – 2.38
V (12)	-0.74	4.17	12.5 min	25 min	2.00
Y (13)	-0.71	3.89	11.7 min	18 min	1.54
F (14)	-0.89	5.89	17.7 min	15 min	0.85
A (15)	-0.43	2.04	6.1 min	12 min	1.97
H^+ (16)	-0.80	4.79	14.4 min	< 2 min	0.14
L (17)	-1.08	9.12	27.4 min	~ 7 min	0.26
D° (18)	-1.03	8.13	24.4 min	< 2 min	0.08
I (19)	-1.03	8.13	24.4 min	< 2 min	0.08

[1] See also Table 2.2 for some parameters.

[2] Acid catalyzed reaction dominates at these conditions (Bai et al., 1995).

[3] This is the most conservative possible estimate; the actual protection factors could be as much as 4 times as large.

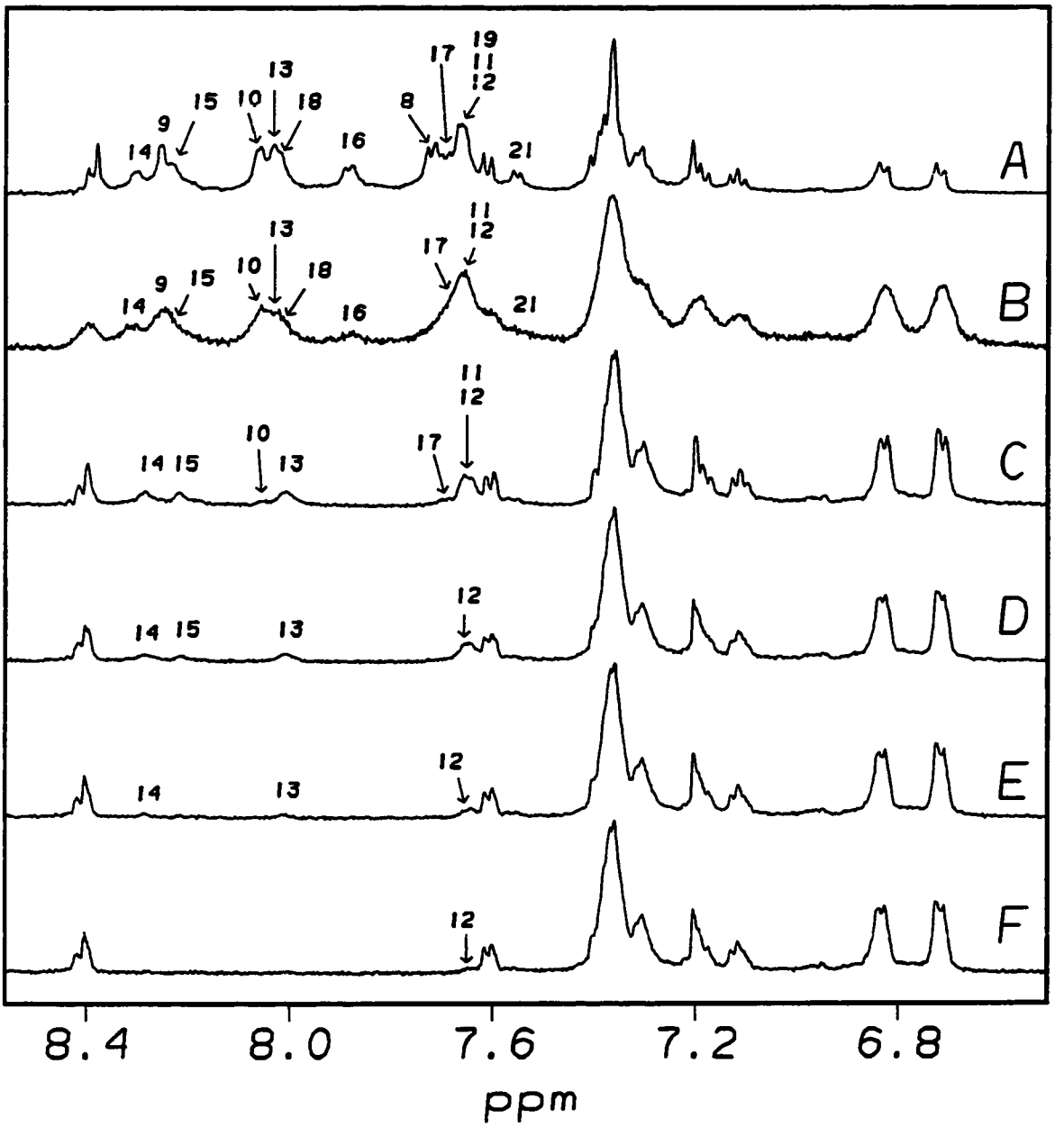


Figure 3.23: Deuterium exchange study of the AcNMe-14mer in aqueous 40% HFIP/24% acetic acid (pH ~ 2.3, 300K). Time points are 0 min, H_2O media (A), 0 min, D_2O media (B), 10 min (C), 20 min (D), 40 min (E) and 1 hour (F). Aromatic proton assignments are omitted for clarity.

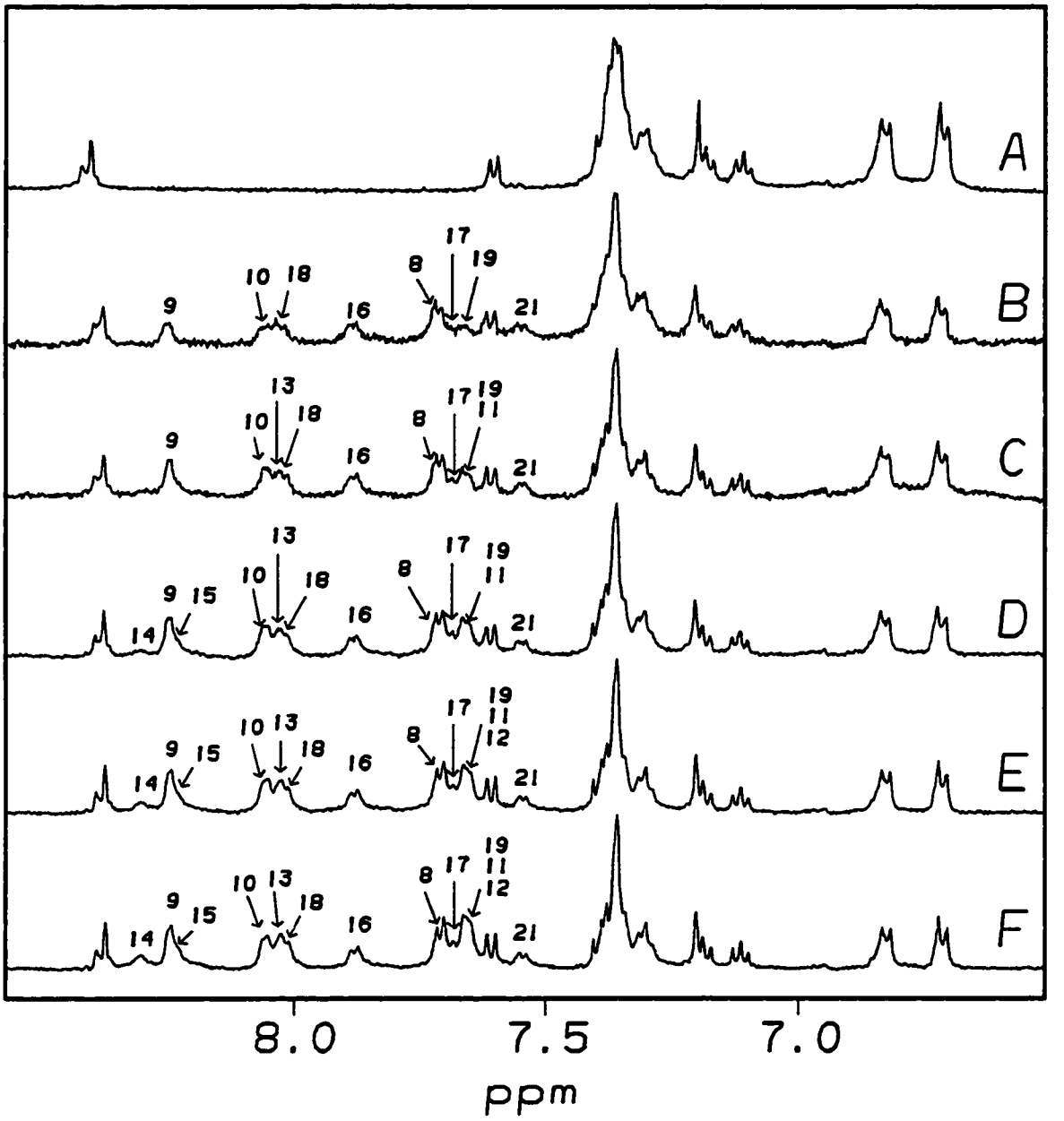


Figure 3.24: Hydrogen re-incorporation study of the AcNMe-14mer in aqueous 40% HFIP/24% acetic acid (pH ~2.3, 300K). Time points are 0 min, D_20 media (A), 0 min, H_20 media (B), 10 min (C), 20 min (D), 40 min (E) and 1 hour (F). Aromatic proton assignments are omitted for clarity.

The estimated pH of the solution is 2.3, which indicates that acid catalysis dominates the reaction. Since Bai et al. (1995) did not include N-capping effects on the Molday factors, Asp^8 is not included in the table. Although the peptide is relatively short, it does display several exchange-protected backbone amides, further evidence that the peptide contains a short helix that extends to at least Ala^{15}. The backbone amide of Val^{12} is the most protected of the peptide, suggesting that it is the least frayed residue in the helical region. Assuming the standard $CO_i \rightarrow NH_{i+4}$ hydrogen bond of an α-helix, Asp^8 is, thus, an effective helix N-cap. Hydrogen re-incorporation (Figure 3.24) also helps verify the location of the helical residue. As in the previous study, the exchange-protected backbone amides correspond to Val^{12} through Ala^{15}.

Overall, exchange rates are much faster for the 14mers than for the intact ET-1 and Pen-1 peptides. For example, residues $Val^{12} \rightarrow Ala^{15}$ exchanged in little over an hour for the AcNMe-14mer. However, $Val^{12} \rightarrow Cys^{15}/Pen^{15}$ of the intact 21mers required between 6 and 16 hours to fully exchange. In addition, the protection factors for residues 9-15 of AcNMe-14mer are much smaller than those of Pen-1 (see also Table 2.2), suggesting that the helical region is more stable in the intact 21mer. This is not unexpected; a 6 to 8 residue helix with no covalent cross-links to stabilize it would not be expected to be 95+% folded. The data implies a helix formation constant of 2 or more.

3.3.5: Comparison of NH-CSD/NH temperature gradient plots.

The NH-CSD/NH temperature gradient diagrams (Figure 3.25) provide further evidence of the stabilizing effects of both N-capping and fluoroalcohol. The backbone NH-CSDs of NMe-14mer at 0% HFIP are clustered together, suggesting little, if any, structure is present. Regression analysis is meaningless in this case with the correlation coefficient of 0.043. In higher concentrations of HFIP, the numbers significantly improve: the regression line has a slope of -14.166ppt/K and a correlation coefficient of

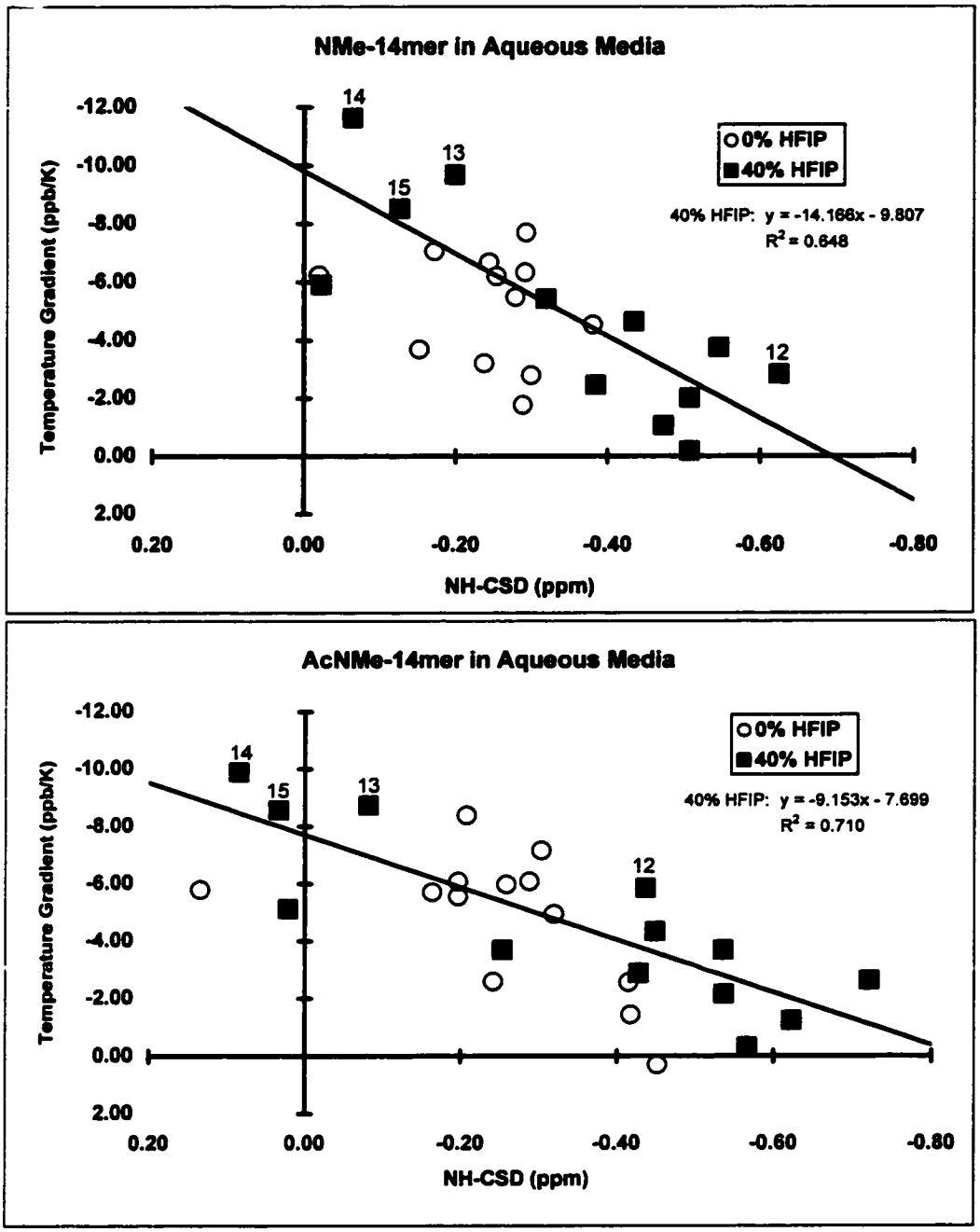


Figure 3.25: The amide NH-CSD/temperature gradient plot for the NMe-14mer (top panel) and AcNMe-14mer (bottom panel), collected in 0% (open circles) and 40% HFIP (filled boxes). The regression line for the fluoroalcohol case is displayed. Also shown are the data points corresponding to amides of $Val^{12} \rightarrow Ala^{15}$.

0.648. This would suggest that in the absence of HFIP, facile melting of structured states occurs at low temperatures and that this process is not cooperative. However, when fluorinated alcohols are added, both structural stability and peptide folding cooperativity improve significantly.

In the case of the AcNMe-14mer, addition of HFIP doesn't have as great affect on the slope and correlation coefficients. In the absence of fluoroalcohol, the slope of the regression line is -9.221ppt/K, with an $R^2 = 0.313$, while at aqueous 40% HFIP, the slope is -9.153ppt/K ($R^2 = 0.710$). These data sets would indicate that the structure of the AcNMe-14mer is more thermally stable than the non-capped analog. The AcNMe-14mer displays significant folding cooperativity both with and without the presence of HFIP.

The location of the points corresponding to the deuterium exchanged-protected backbone amides ($Val^{12} \rightarrow Ala^{15}$) is a notable feature of these diagrams. Residue 12, which has the longest $t_{½}$, is further upfield than would be expected for the most strongly hydrogen bonded amide. Residues $13 \rightarrow 15$ demonstrate much larger temperature gradients than would be expected. This is further evidence that temperature gradients cannot accurately predict deuterium exchange protection factors in relatively structured peptides (Andersen et al., 1997). In proteins, residues with long exchange half-lives would be expected to display temperature gradients less negative than -4ppb/K.

3.3.6: Evidence of Structuring in DMSO?

Although DMSO was originally used to verify the sequences of the 14mers, some evidence of structure is present in this denaturing solvent. Addition of 2-5% H_2O and 3 equivalents of TFA were required to sufficiently sharpen the backbone amide peaks. Final chemical shift assignments of the 14mers are reported in Appendix C. The αH-CSD histogram of the 14mers (Figure 3.26, upper panel) display limited helicity from Glu^{10} to Leu^{17}. The central residue of the helix is Val^{12}. The maximum negative value is

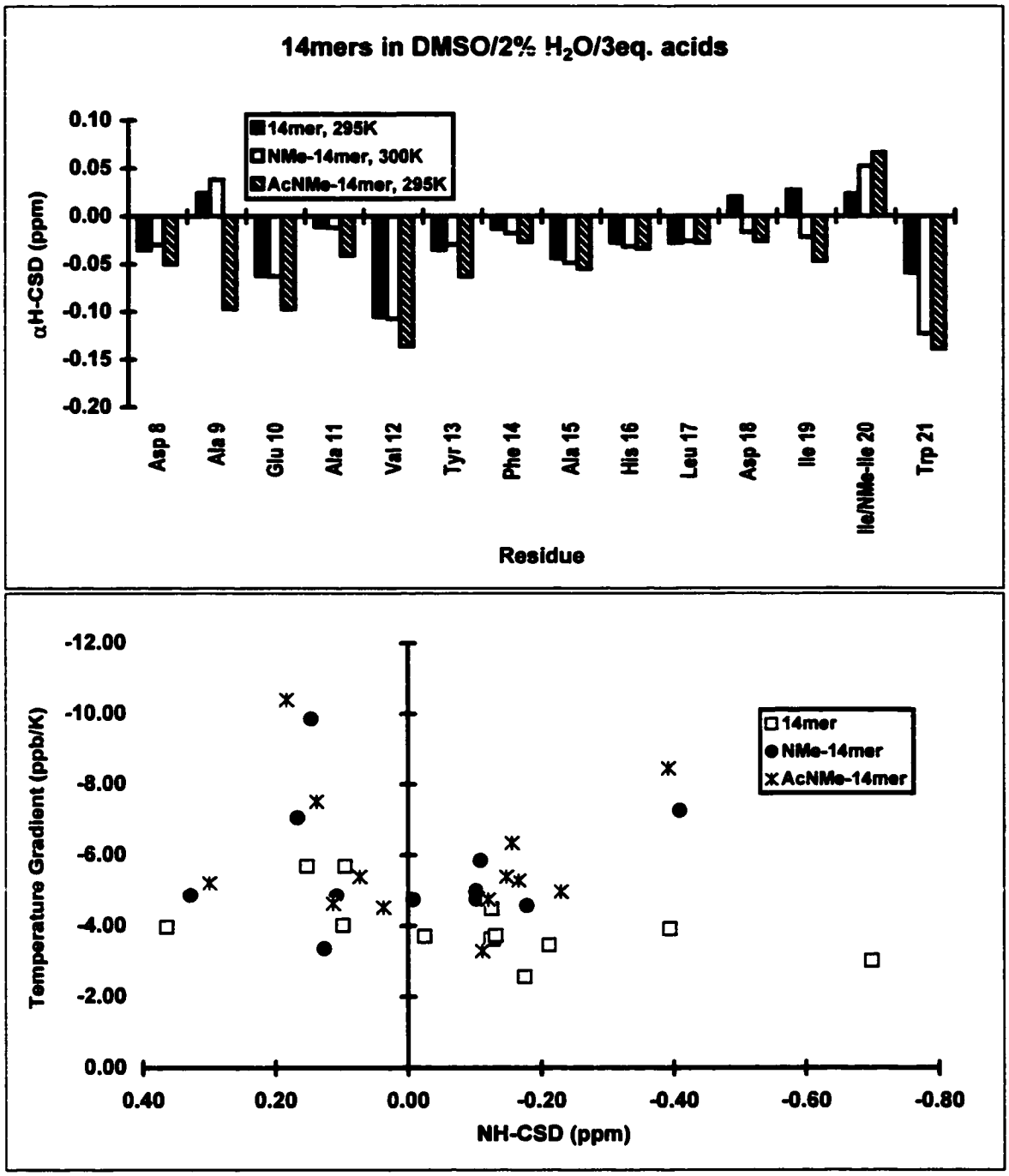


Figure 3.26: The 14mers in DMSO, 2% H_2O, 3 equivalents TFA. Top panel: αH-CSD histogram. Bottom panel: NH-CSD vs. NH temperature gradient plot.

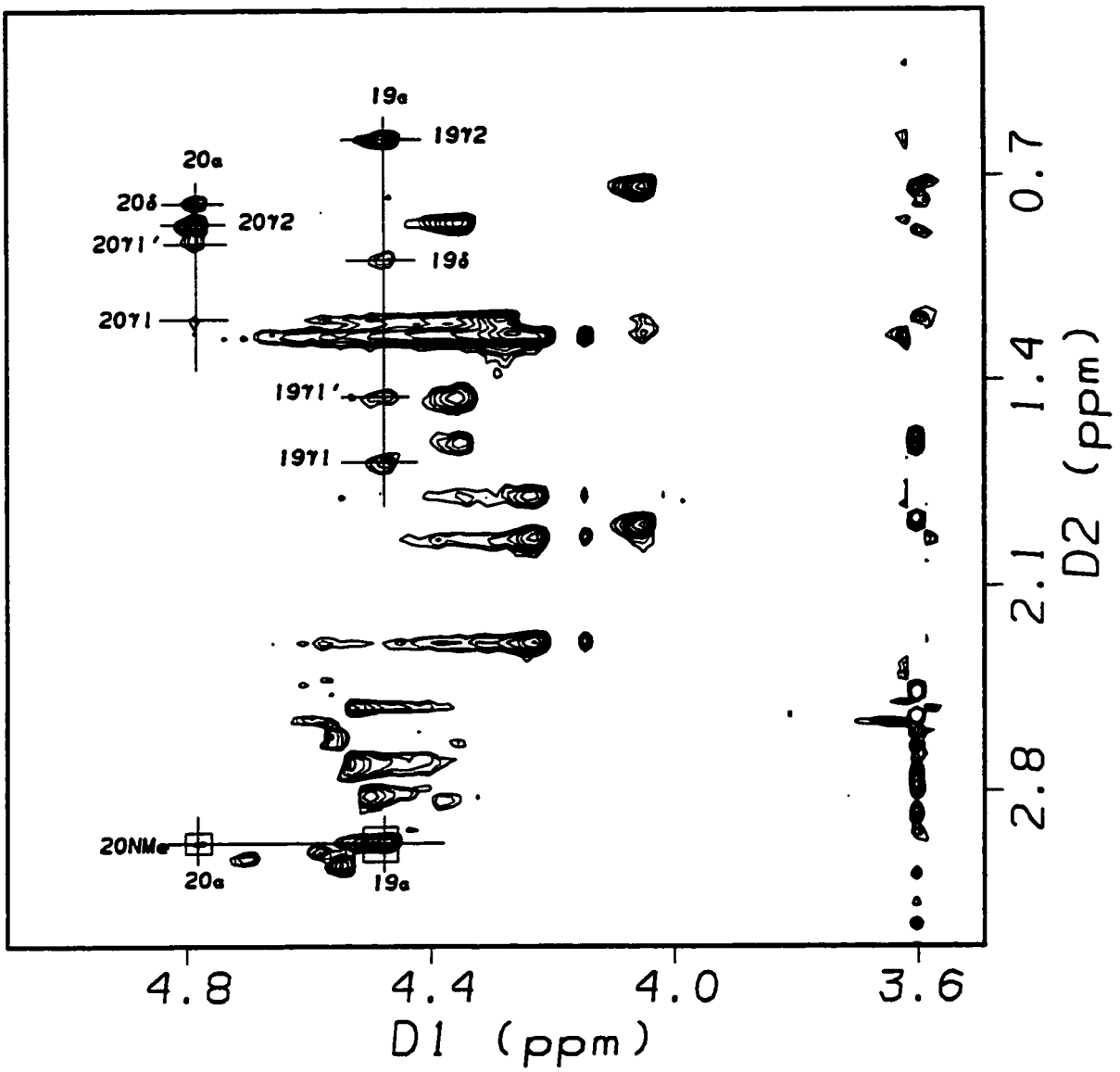


Figure 3.27: The αH → upfield crosspeak region of the AcNMe-14mer NOESY in DMSO, 2% H_2O, 3 equivalents TFA. The spectrum was acquired at 499MHz. Note the strong 19α → 20NMe crosspeak.

-0.15ppm over this span, suggesting that the helical conformation is not highly populated. In the N-methylated analogs, this helix appears to extend to Ile^{19}, however the alternating signs for the αH-CSD values of residues 19-21 suggests a turn-like feature. The acetylated version has more negative values over a longer range of residues, indicating that N-capping stabilizes the helix, even in DMSO.

The NH-CSD/NH temperature gradient plot (Figure 3.26, lower panel) suggests neither folding cooperativity nor a highly structured state for any of these three peptides in DMSO media. However, slight improvements in the correlation coefficients are observed over the helical region (residues 9 through 16) in all cases. This would suggest that more cooperative folding is present for the helical region of the peptides.

Further insights regarding the C-terminal hydrophobic cluster can also be gained from the DMSO data using the 19γ2 peak as a probe. The cluster is more stable when the peptide contains the $^{(NMe)}Ile^{20}$ residue: the chemical shift of 19γ2 is further downfield for the 14mer (0.742ppm) relative to the NMe-14mer (0.567ppm). A comparison of data acquired in both aqueous and DMSO media also indicates that the media affects the stability of the cluster. In 40% aqueous acetic acid, the 19γ2 peak of the NMe-14mer has a chemical shift of 0.276ppm, an upfield deviation of 0.276ppm relative to that of the peptide in DMSO. This would suggest that the hydrophobic cluster is more stable the more polar (lipophobic) the medium. Unfortunately, the aggregation problems encountered with the 14mer prevented a direct comparison of the aqueous and DMSO media chemical shifts. The NMe-14mer and AcNMe-14mer display no significant chemical shift differences over the C-terminal residues.

The ROESY and $^3J_{NH\text{-}\alpha H}$ data would also suggest no significant structuring in DMSO media for 14mer. The interresidue $\alpha_i \rightarrow NH_{i+1}$ crosspeaks are stronger than their intraresidue $\alpha_i \rightarrow NH_i$ counterparts over the entire peptide. In addition, there are no α_i

$\rightarrow NH_{i+3}$ or $\alpha_i \rightarrow \beta_{i+3}$ crosspeaks in the spectrum. There are, however, a few small-to-medium sized N_iN_{i+1} crosspeaks for residues 9/10, 13/14, 14/15, 19/20 and 20/21. With few exceptions, the $^3J_{NH\text{-}\alpha H}$ coupling constants are greater than 7Hz, indicating an averaged structure.

Both N-methylated analogs display the same lack of intermediate range ROESY interactions as for the 14mer. However, one significant set of ROESY crosspeaks corresponding to the 20NMe line is present for both the NMe-14mer and AcNMe-14mer (Figure 3.27). The large 19α → 20NMe, as well as the absent 19α → 20α crosspeak indicate that the trans conformation is present for the 19/20 peptide bond in DMSO. This is contrary to the observations of the Parke-Davis group (Cody et al., 1997) who reported the NMR structure of a six residue ET_A/ET_B receptor antagonist containing an unnatural amino acid residue (D-Bhg[16]) and an N-methylated Ile[20]. The authors noted that in aqueous media, the 19/20 peptide bond is 100% trans. However, when the peptide was studied in DMSO media, the 19/20 peptide bond purportedly isomerizes to the cis conformation. This discrepancy may be partially explained by the presence of the D-Bhg residue, which contains a large aromatic moiety in the side chain. One author (Reily, personal communication) suggests that a hydrophobic interaction between the D-Bhg[16] and Trp[21] side chains may allow the 19/20 peptide bond to adopt the cis conformation.

3.4: AQUEOUS MEDIA NMR STUDIES: SHORT FRAGMENTS

A number of peptide fragments were originally synthesized at Bristol-Myers Squibb to more fully understand the conformational preferences of the ET-1 C-terminal region. To limit structural constraints induced by the bicyclic core, these fragments were truncated so as to either remove the N-capping residues of the helix, or the helical region itself. The analogs examined in this section, ET-1[16-21] and ET-1[12-21] (henceforth, the 6mer and 10mer, respectively), contain the HLDIIW moiety (see also, Table 3.1).

Both peptides retain free amine N-termini to insure that any structuring preferences influenced by N-capping are eliminated.

However, other than an N-methyl scan of the Pen-1 C-terminus (Harris, 1993), no other Pen-2 analogs were produced at Bristol-Myers Squibb. As a result, several $^{(NMe)}Ile^{20}$ fragments (and their non-N-methylated counterparts were synthesized for the present study. The initial syntheses resulted in the 8mer (GSHLDIIW) and the 12mer (VYFAHLDIIW). In the case of the 8mer, the $^{(NMe)}Ile^{20}$ analog, with and without an acetyl N-cap (the NMe-8mer and AcNMe-8mer, respectively) were also synthesized. The major product isolated from the initial 12mer synthesis yielded an unexpected, yet useful, deletion product, the 9mer. This was fortunate as this peptide lacked the three C-terminal residues ($Ile^{19} \rightarrow Trp^{21}$), and, as a result, a portion of the helical region could be probed without influences from the hydrophobic cluster. Final chemical shift assignments for the peptides are listed in Appendix B.

3.4.1: BMSQ C-Terminal Analog Studies

Given their short sequences, the 6mer and 10mer chemical shifts were unambiguously assigned from their COSY spectra. In DMSO, both structures appear to be fully disordered. For the 6mer, the average αH-CSD value over the entire sequence is -0.065 ± 0.112ppm. (His^{16} has a large negative deviation due to its proximity to the terminal ammonium ion. When only residues 17-21 are considered, the average deviation is –0.015 ± 0.050ppm.) The average αH-CSD in the case of the 10mer over residues 16 through 21 is –0.030 ± 0.040ppm. As expected, both the 6mer and 10mer DMSO NOESY spectra (not shown) displays mostly intraresidue and sequential NOEs (where $\alpha_i N_i << \alpha_i N_{i+1}$). No significant intermediate-range NOEs are observed for either peptide.

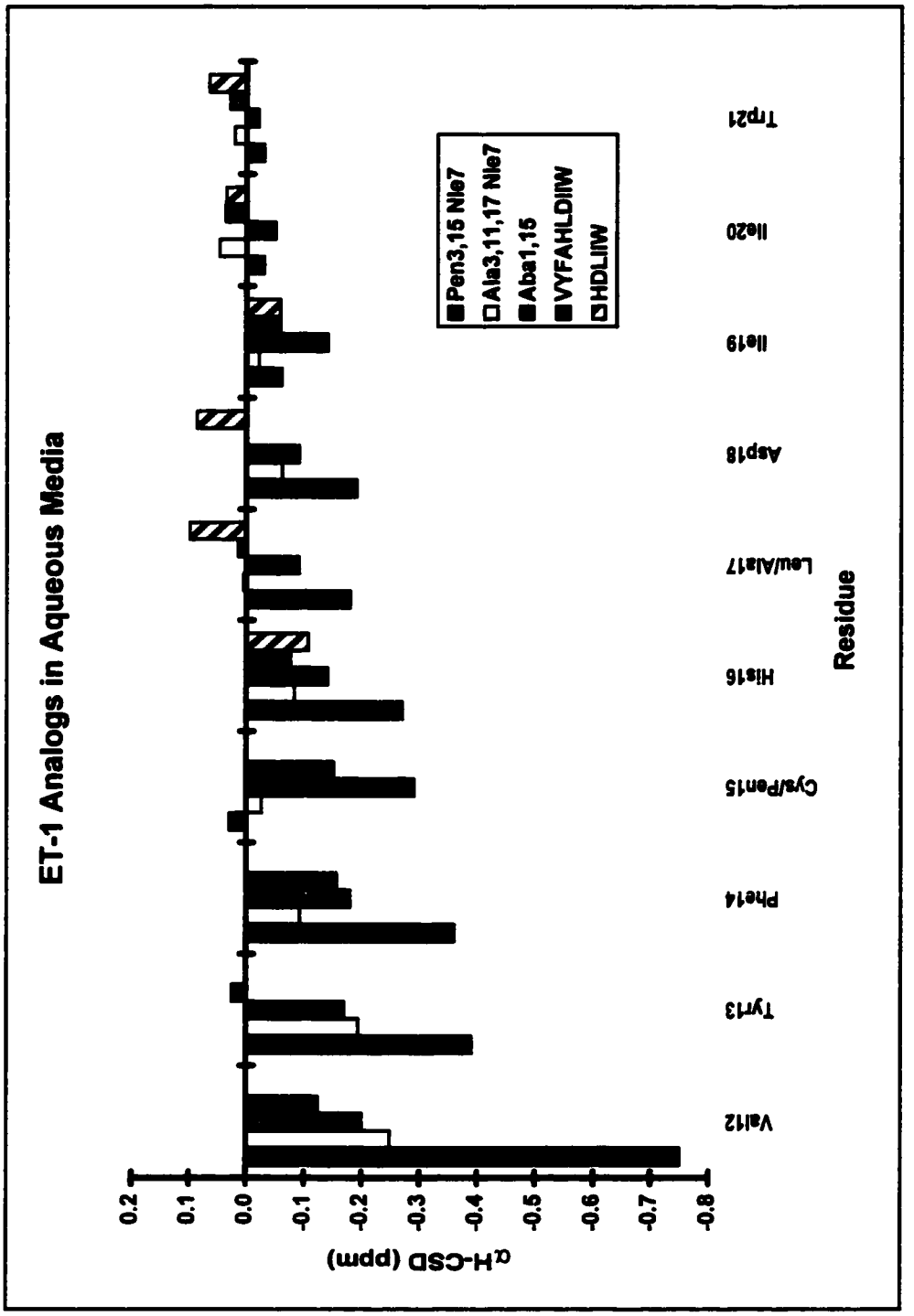


Figure 3.28: The αH-CSD histogram of the 6mer and 10mer, compared with intact and monocyclic ET-1 analogs.

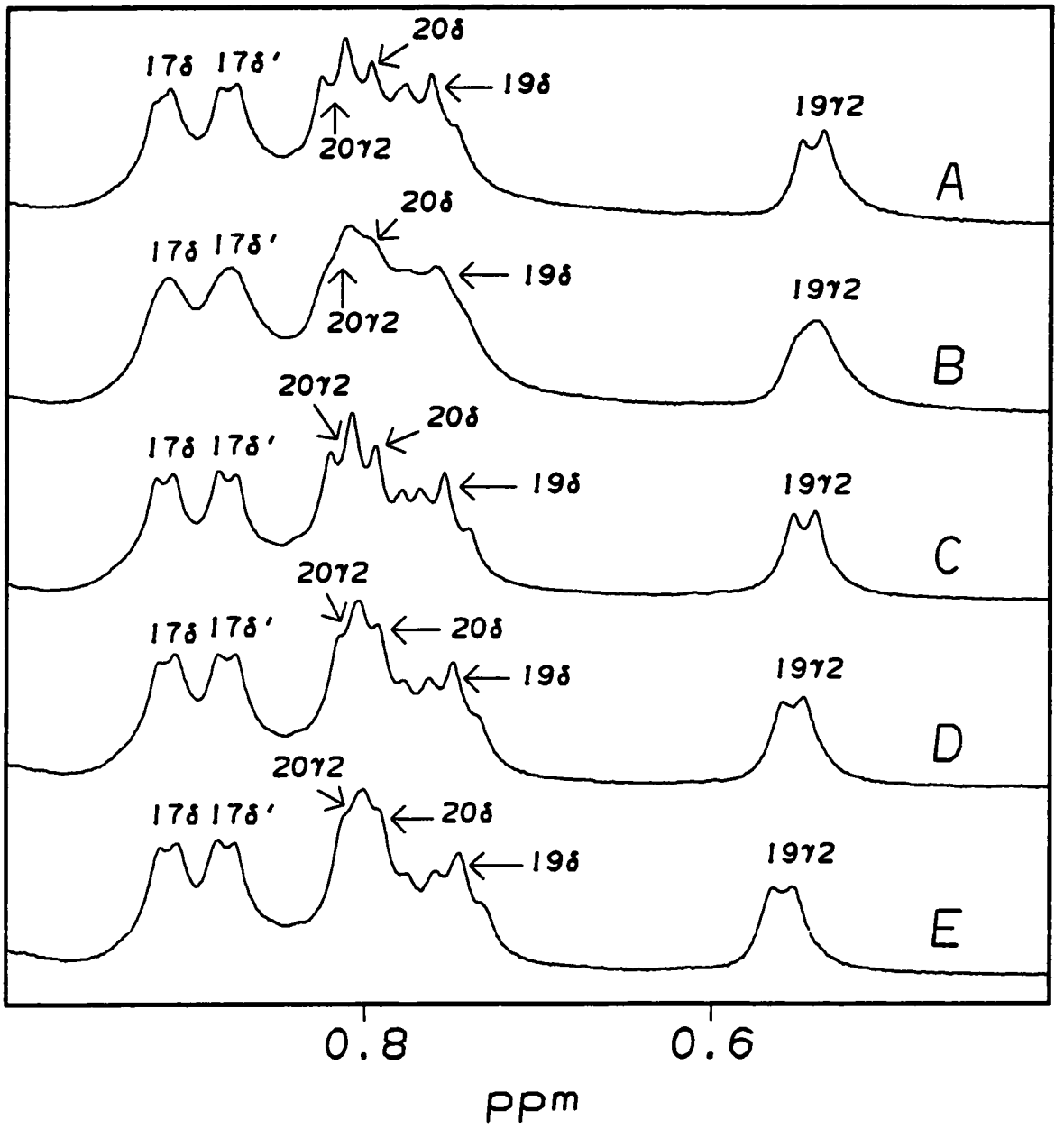


Figure 3.29: The methyl group region of the HLDIIW (6mer) ethylene glycol titration, acquired at 285K. Points in the titration are 10% (A), 20% (B), 30% (C), 40% (D) and 50% glycol (E) by volume.

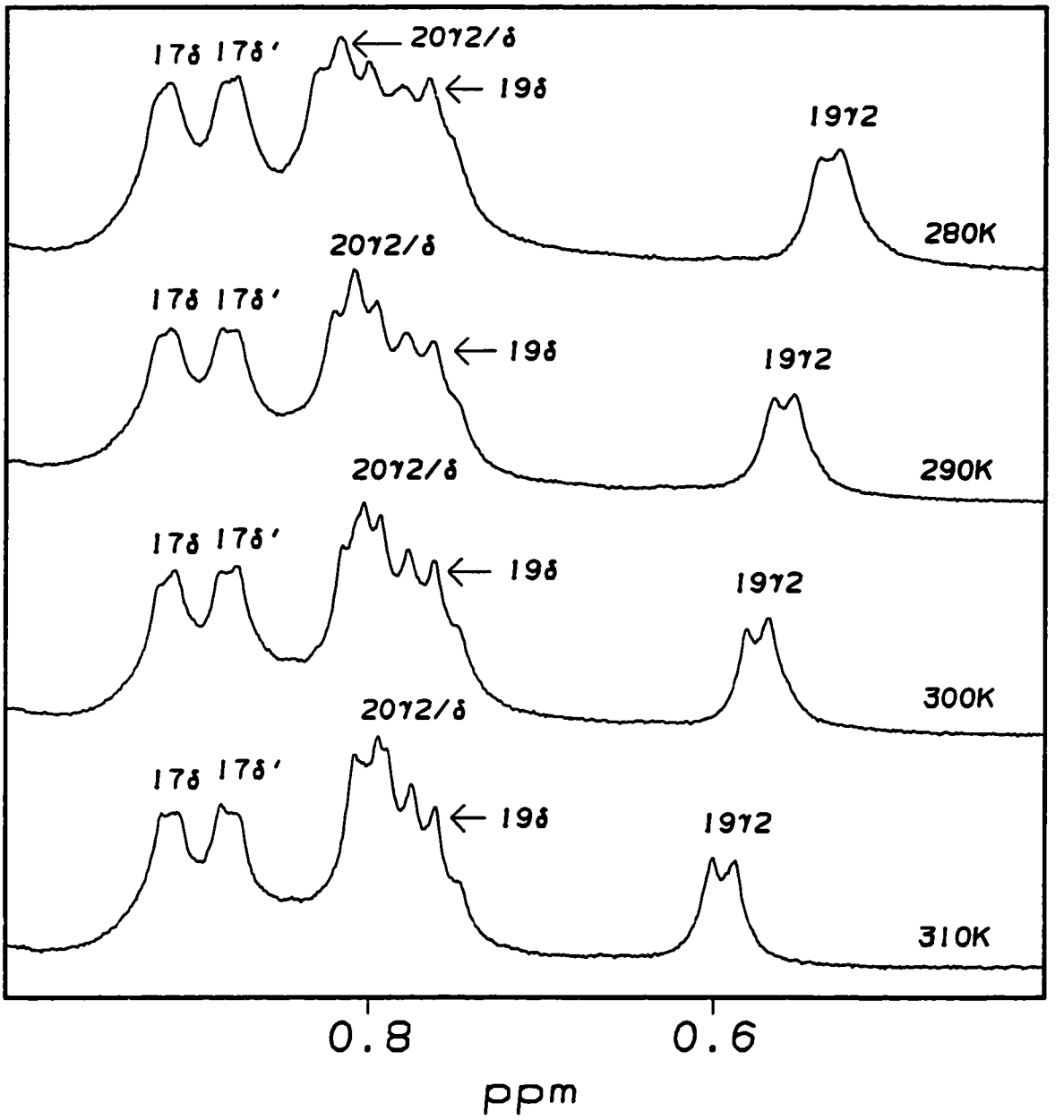


Figure 3.30: The methyl group region of the 6mer temperature study, acquired in 10% ethylene glycol.

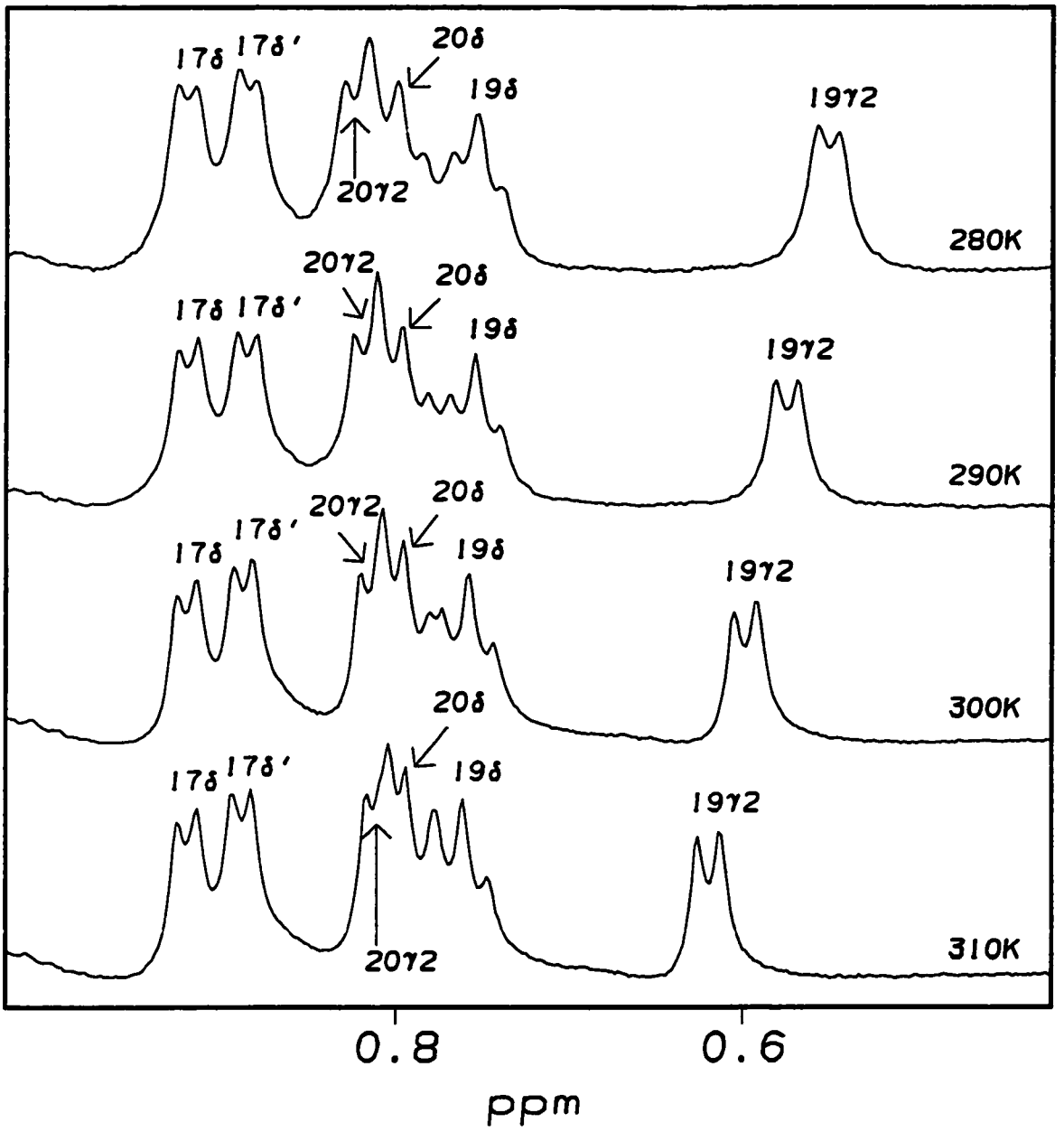


Figure 3.31: The methyl group region of the 6mer temperature study, acquired in 30% ethylene glycol.

The αH-CSD histogram (Figure 3.28) indicates that, as expected, none of the fragments are as structured as Pen-1. Negative αH-CSDs over residues 14 → 16 suggest that the 10mer may be slightly helical. However, the small values (maximum negative values of -0.15ppm) and the fact that the 10mer does not display any NOESY crosspeak patterns indicative of a helical structure suggests that only a small portion of the conformer population samples local α_R backbone torsional angles. The C-terminal residues (Ile^{19} → Trp^{21}) for the peptides display similar small αH-CSD values throughout, indicating that this region is unstructured, or experiences extensive conformational averaging in aqueous solution. Overall, the 10mer has comparable deviations with respect to the monocyclic analogs, [$Ala^{3,11,17}$-Nle^{7}]ET-1 (Andersen et al., 1995a) and [$Aba^{1,15}$]ET-1 (Coles et al., 1994), although values are smaller at the N-terminus of the helix due to sequence truncation.

As with the intact ET-1 analogs, both the 6mer and 10mer Ile^{19}-γ2 peak displays unusual upfield chemical shifts due to shielding effects of the Trp^{21} sidechain indole ring. In the case of the 6mer, addition of ethylene glycol doesn't appear to greatly affect the stability of the hydrophobic cluster (Figure 3.29). At 0% aqueous glycol, the 19γ2 peak has a chemical shift of 0.545ppm, while at 50% glycol, it shifts downfield to 0.561ppm, a difference of +0.016ppm. The other methyl groups, specifically those of Leu^{17} and Ile^{20}, as well as 19δ, have a slight upfield deviation with increasing glycol concentrations. Temperature studies at 10% and 30% glycol (Figures 3.30 and 3.31) demonstrate dramatic effects on the cluster: although the 19γ2 peak moves downfield in both cases, it has a steeper gradient in the higher glycol levels. At 10% aqueous glycol, the 19γ2 chemical shift is 0.545ppm at 285K, with a gradient of +1.578ppb/K, while in 30% aqueous glycol, the values are 0.549ppm and +2.054ppb/K, respectively. In 50% glycol (data not shown), the chemical shift is 0.561ppm, with a gradient of +2.076ppb/K. However, spectra were acquired at only two temperatures, so the latter value carries a

high uncertainty. This data would suggest that increasing glycol concentrations partially destabilizes the hydrophobic cluster. That is, the interaction between the 19γ2 and the Trp^{21} sidechain becomes weaker (and more likely to unfold at higher temperatures) with decreasing "lipophobicity" of the solution.

3.4.2: THE 8MERS.

The ROESY spectrum of the 8mer in 40% aqueous ethylene glycol (data not shown) displays patterns consistent with a disordered structure. Strong $\alpha_i N_{i+1}$ crosspeaks are present, while those corresponding to other intraresidue αN NOEs have very weak intensities. Most of the $^3J_{NH/\alpha H}$ coupling constants in this medium are greater than 6.5Hz at ambient temperatures. The αH-CSD histogram (Figure 3.32) also indicates a disordered structure: the values for $His^{16} \rightarrow Trp^{21}$ are less than ± 0.05ppm. The larger positive values for Gly^{14} and Ser^{15} suggest a partially extended conformation. Addition of HFIP does not seem to have much influence of the backbone conformation. At 15% HFIP, the coupling constants are still greater than 6.5Hz and the αH-CSDs aren't significantly altered. However, an increase in the chemical shift of the 19γ2 peak from 0.586ppm to 0.686ppm is observed as the hydrophobic clusters becomes less dominant.

The NMe-8mers in 30% glycol (with and without HFIP) do not display any dramatic structural differences with respect to the 8mer for the N-terminal portion of the sequence. The αH-CSDs for residues 14 through 18 are comparable for the three peptides. The difference between the two $^{(NMe)}Ile^{20}$ analogs at Gly^{14} is solely due to the N-acetylation effect. The αH deviations are more negative for residues 20 and 21, and become more so as HFIP concentrations increase. The coupling constants also have similar values with respect to the 8mer ($^3J_{NH/\alpha H} > 6.0$Hz), although those corresponding to residue 21 are smaller in the N-methylated peptides. The ROESY spectra of the $^{(NMe)}Ile^{20}$

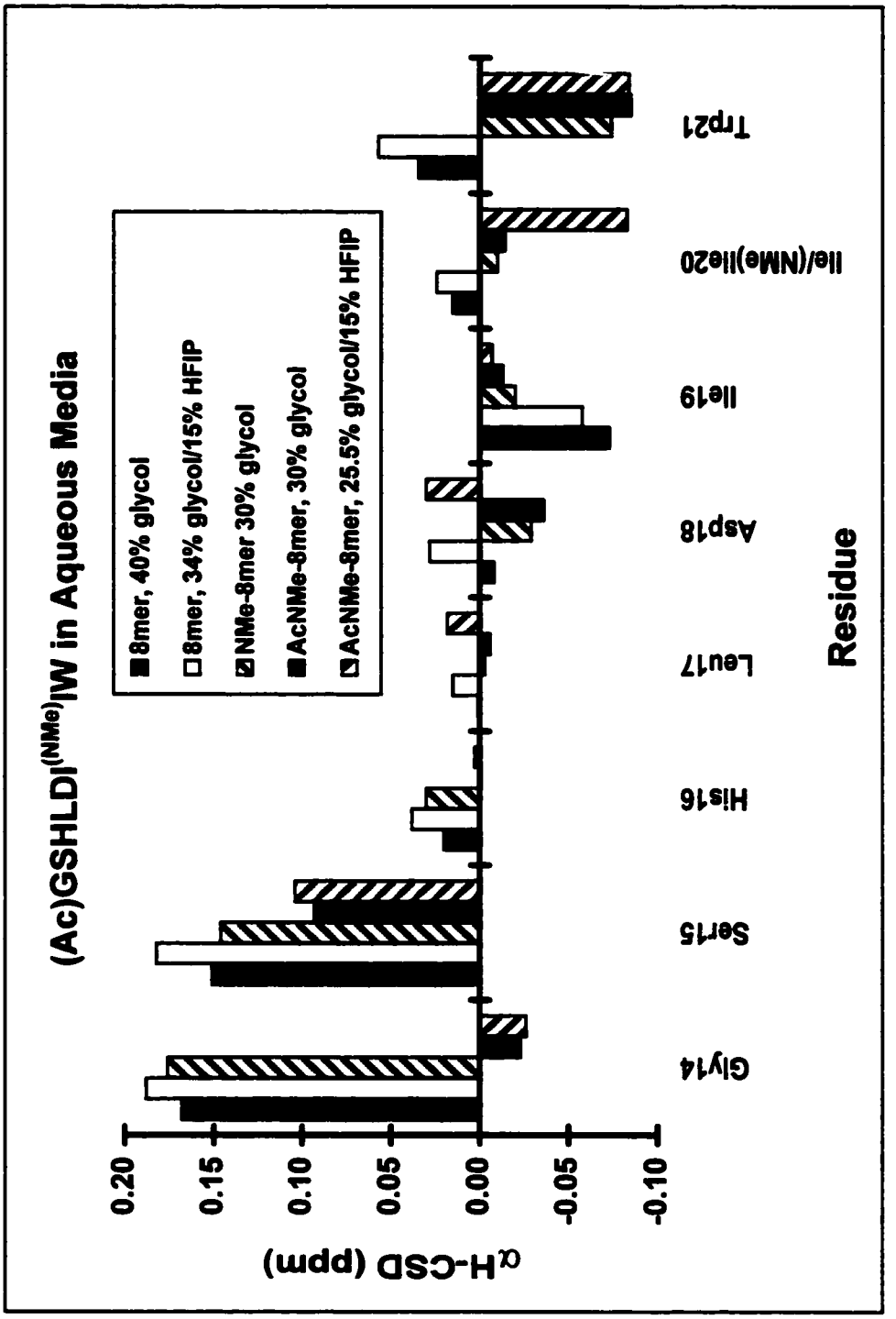


Figure 3.32: The αH-CSD histogram of the 8mers in aqueous media.

analogs (data not shown) display similar patterns to their Ile^{20} counterpart, especially in the amide region. The α_iN_i crosspeak intensities are always smaller than those corresponding to the α_iN_{i+1} interaction. However, the C-terminus, as expected, displays a number of differences. For both NMe-8mers, the 19γ2 peak is shifted further upfield than the 8mer. Similar crosspeak patterns between 19α → 20NMe and 19γ2 → 21α, which are observed in the larger $^{(NMe)}Ile^{20}$ analogs, are present. This is further evidence that the hydrophobic cluster can form without the presence of a stable helix.

3.4.3: THE 9MER AND 12MER: EVIDENCE OF STRUCTURE?

Contrary to the 8mers, both the 12mer and 9mer display evidence of limited helicity. The αH-CSD (Figure 3.33) values are negative between Val^{12} and His^{16}. Values for the C-terminal residues of the helix are comparable to those of the 10mer. As expected, increasing concentrations of HFIP appears to slightly stabilize the helix. However, the relatively small upfield chemical shift deviations (-0.15ppm, maximum) compared to the longer fragments (see Figure 3.34) suggest that the helical region isn't as stable in the 9mer and 12mer. Even at high HFIP concentrations, the upfield deviations for residues 12 through 14 of the 9mer and 12mer are only a quarter that of the AcNMe-14mer. This isn't surprising considering that in the intact peptides the helical region is capped by Asp^8, while sequences of the 9mer and 12mer begin with Glu^{10}. Since neither fragment is N-capped, the terminal region is expected to have higher mobility.

The $^3J_{NH/\alpha H}$ coupling constants also support the presence of a helical structure that is stabilized by increasing concentrations of HFIP. At 0% HFIP, the 9mer displays average coupling constant (J > 7.0Hz), with the exceptions of Ala^{11} and Ala^{15}, which have values less than 6.0Hz at 285K. When the peptide is studied at 25% HFIP, the coupling constants for residues 11 through 15 are less than 6.5Hz at 285K. Coupling constants for the 12mer were more difficult to measure due to a number of degenerate amide chemical

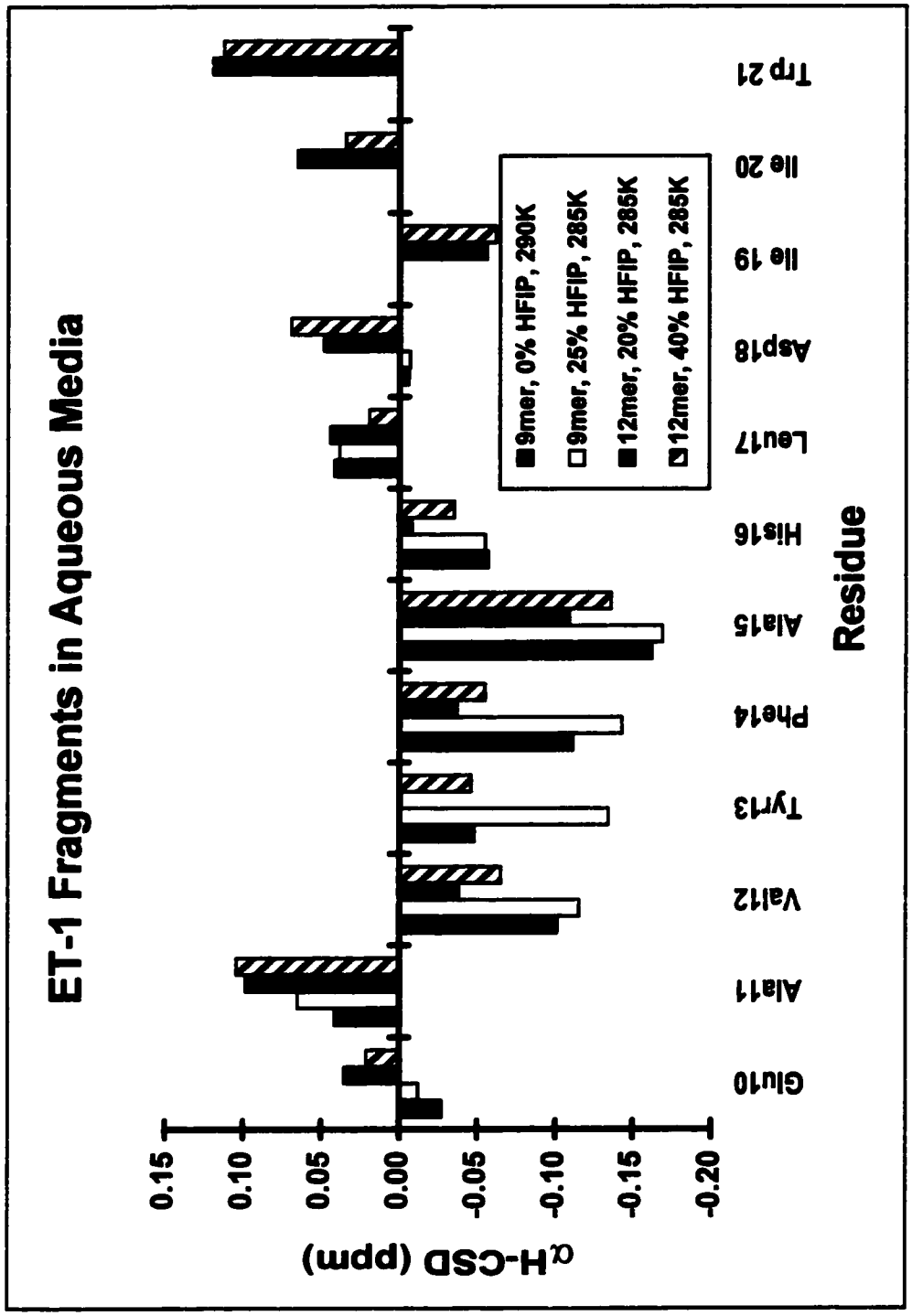


Figure 3.33: HFIP effects on the αH-CSD histograms of the 9mer and 12mer.

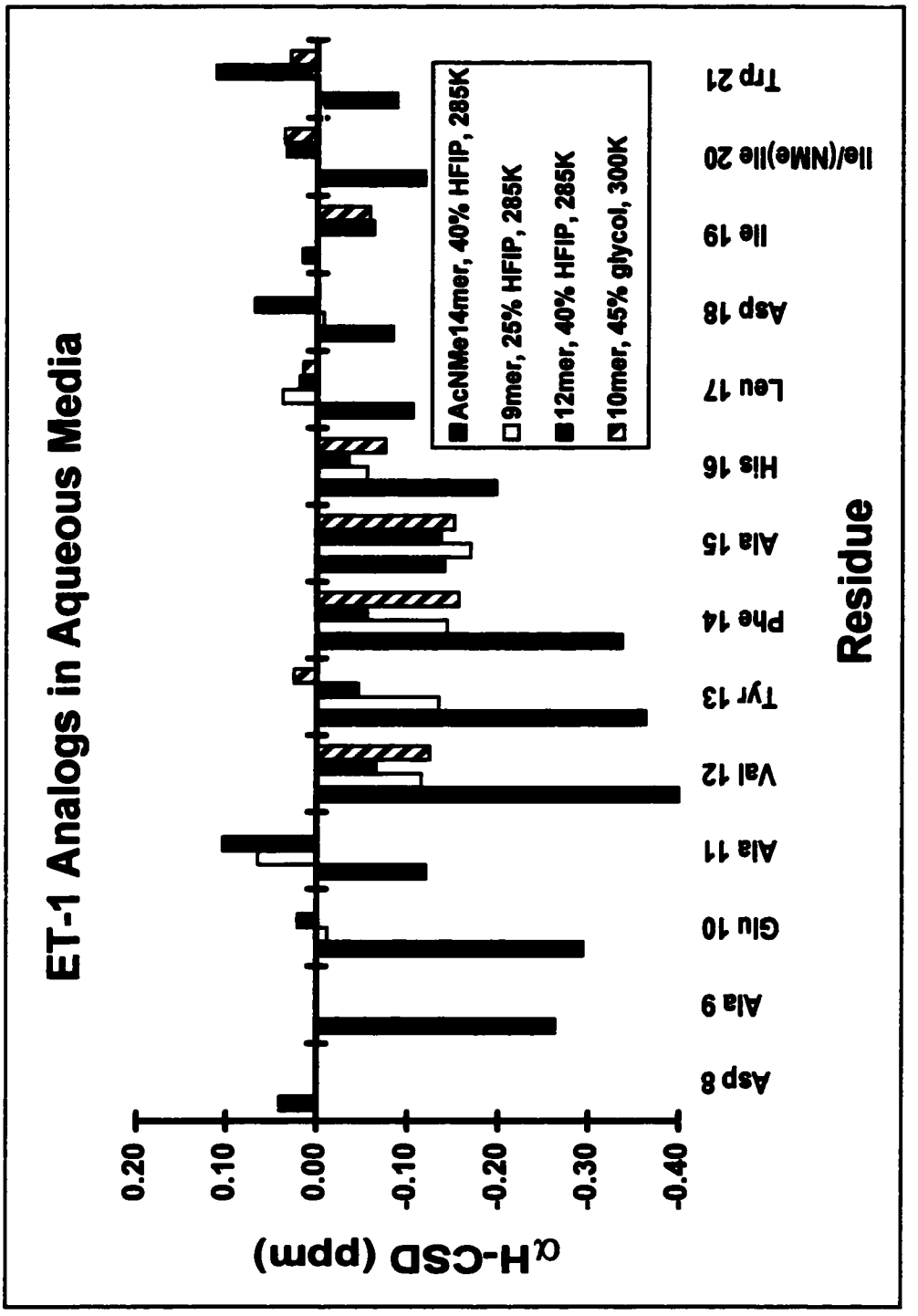


Figure 3.34: The aH-CSD histogram of the 9mer, 10mer and 12mer in aqueous HFIP media, compared with the AcNMe-14mer.

shifts in both 20% and 40% HFIP. However, at near ambient temperatures (295-300K), residues 11, 16 and 17 display values less than 6.5Hz. Values for $Val^{12} \rightarrow Ala^{15}$, unfortunately, could not be accurately determined due to the aforementioned signal overlap.

However, the NOESY spectra (data not shown) for both the 9mer and 12mer (even at high fluoroalcohol concentrations) do not display any significant intermediate range NOEs, suggesting that any structuring is only local. In both cases, the intraresidue $\alpha_i N_i$ crosspeak is less intense than their sequential $\alpha_i N_{i+1}$ counterparts. A few $N_i N_{i+1}$ crosspeaks are present in the 12mer NOESY, which is consistent for a helical conformation, but no unambiguous $i \rightarrow i+3$ crosspeaks are observed. Returning to the αH-CSD histograms, the helical region of the 12mer appears to be less structured than that of the 9mer, even in higher levels of fluoroalcohol. This suggests that the C-terminal hydrophobic cluster may effect helix stability, independent of media.

3.5: UNUSUAL ILE[19] BETA-METHYL CHEMICAL SHIFTS

The 19γ2 (β-methyl) group is the most easily recognizable peak of the ^{1}H-NMR methyl region for the ET-1 peptide series. Due to its location within the shielding cone of the neighboring Trp^{21} sidechain indole ring, the 19γ2 peak has a (sometimes) dramatic upfield chemical shift. As a result, it can be effectively used as a probe for identifying structural preferences of the conformationally averaged C-terminus. Table 3.11 reports the 19γ2 chemical shifts and temperature gradients for a series of ET-1 peptide analogs in a variety of aqueous media. The data are arranged in order of decreasing 19γ2 CSDs and separated by groups (Ile^{20} or $^{(NMe)}Ile^{20}$ analogs). Table 3.12 summarized the average deviations and temperature gradients according to species and conditions. General trends are evident in this data set.

Table 3.11: Ile^{19}-γ2 chemical shift deviations (normalized to 285K) and temperature gradients in aqueous media [1,2]

Peptide	Conditions	19γ2 CSD (ppm)	Δδ/ΔT (ppb/K)	(Δδ/ΔT)/CSD (ppt/K)
HLDIIW (6mer)	10% glycol	-0.405	+1.578	-3.896
VYFAHLDIIW	8% HOAc/45% glycol	-0.398	+1.829	-4.595
ET-1 [3]	50% glycol	-0.396	~+2.480	-6.263
GSHLDIIW (8mer)	5% HOAc/40% glycol	-0.392	+1.643	-4.191
HLDIIW (6mer)	30% glycol	-0.385	+2.054	-5.335
EAVYFAHLDIIW	32% HOAc/20% HFIP	-0.326	+2.188	-6.712
Pen-1 [4]	50% glycol	-0.321	+0.935	-2.913
GSHLDIIW (8mer)	34% glycol/15% HFIP	-0.284	+1.286	-4.528
Ac-AIIY	40% HOAc	-0.273	+2.053	-7.520
EAVYFAHLDIIW	24% HOAc/40% HFIP	-0.270	+1.843	-6.826
Ac-AIIY	32% HOAc/20% HFIP	-0.197	+1.712	-8.690
AcNMe-8mer	3.5% HOAc/30% glycol	-0.727	+3.351	-4.609
NMe-8mer	3.5% HOAc/30% glycol	-0.719	+3.402	-4.732
AcNMe-8mer	5% HOAc	-0.695	+2.725	-3.921
Pen-2	50% glycol	-0.694	+2.173	-3.131
NMe-14mer	40% HOAc	-0.693	+3.566	-5.146
AcNMe-14mer	40% HOAc	-0.690	+3.494	-5.064
NMe-14mer	32% HOAc/20% HFIP	-0.689	+4.473	-6.492
NMe-8mer	5% HOAc	-0.689	+2.640	-3.832
AcNMe-14mer	32% HOAc/20% HFIP	-0.683	+4.617	-6.760
AcNMe-8mer	25.5% glycol/15% HFIP	-0.675	+3.886	-5.757
NMe-8mer	25.5% glycol/15% HFIP	-0.662	+3.830	-5.785
Pen-2	15% HFIP	-0.657	+2.519	-3.834
NMe-14mer	24% HOAc/40% HFIP	-0.646	+4.557	-7.054
AcNMe-14mer	24% HOAc/40% HFIP	-0.637	+4.736	-7.435

[1] All data acquired at 500MHz.

[2] For CSD calculations, using Ile γ2 reference = 0.95ppm (Wüthrich, 1986).

[3] Data from Chen, 1992.

[4] The 19γ2 chemical shifts of Pen-1 in 15% HFIP could not be accurately determined.

Table 3.12: Average Ile^{19}-γ2 chemical shifts and temperature gradients from Table 3.11

	δ (ppm)	CSD (ppm)	Δδ/ΔT (ppb/K)	(Δδ/ΔT)/CSD
Ile^{20} analogs (n = 11)				
Overall	0.618 ± 0.069	-0.332 ± 0.069	+1.782 ± 0.428	-5.588 ± 1.754
HOAc & glycol	0.583 ± 0.050	-0.367 ± 0.050	+1.796 ± 0.485	-4.959 ± 1.551
HFIP	0.681 ± 0.054	-0.269 ± 0.054	+1.757 ± 0.373	-6.689 ± 1.702
$^{(NMe)}Ile^{20}$ analogs (n = 14)				
Overall	0.267 ± 0.025	-0.683 ± 0.025	+3.569 ± 0.858	-5.254 ± 1.341
HOAc & glycol	0.249 ± 0.015	-0.701 ± 0.015	+3.050 ± 0.536	-4.348 ± 0.741
HFIP	0.286 ± 0.019	-0.664 ± 0.019	+4.088 ± 0.779	-6.160 ± 1.198

3.5.1: ILE^{20} *VS.* $^{(NMe)}ILE^{20}$ ANALOGS.

The 19γ2 peak of the $^{(NMe)}Ile^{20}$ analogs are shifted further upfield and have steeper temperature gradients than their Ile^{20} counterparts. Assuming that the Ile β-methyl random coil chemical shift is 0.95ppm (Wüthrich, 1986), the CSD values of the 19γ2 methyl groups are -0.683 ± 0.025ppm (^{NMe}Ile analogs) and -0.332 ± 0.069ppm (Ile analogs). The corresponding temperature gradients, as with the CSDs, are twice as large for the $^{(NMe)}Ile^{20}$ analogs. Upon first inspection, this might suggest that the N-methylated peptides have a more compact hydrophobic cluster (i.e., the 19γ2 methyl group aligns with the shielding cone of the Trp^{21} indole ring more efficiently) whose structure more readily undergoes thermal "melting." However, when the overall temperature gradient-CSD correlations are examined, no significant differences arise between the two sets of analogs. This suggests that, on average, the hydrophobic cluster of both the Ile^{20} and $^{(NMe)}Ile^{20}$ peptides experience the same level of thermal melting.

3.5.2: MEDIA EFFECTS.

The presence of fluorinated alcohols influences the 19γ2 chemical shift, although this effect appears to be more pronounced for the $^{(NMe)}$Ile^{20} analogs. A clear delineation for both the Ile^{20} and $^{(NMe)}$Ile^{20} series is observed: the 19γ2 peaks are shifted further downfield in the presence of HFIP. The solvent effects on the average temperature gradients and Δδ/ΔT-CSD correlations are statistically less significant, although in both the Ile^{20} and $^{(NMe)}$Ile^{20} analog set, the correlations are larger (more negative) in the presence of HFIP. In the case of the N-methylated peptides, these correlations are more clearly separated if only the fragments are considered (without HFIP: -4.551 ± 0.560ppt/K, with HFIP: -6.547 ± 0.678ppt/K). Trends in the glycol titration are less predictable for the 19γ2 CSDs. For the 6mer (HLDIIW), the 19γ2 methyl group shifts downfield at higher ethylene glycol levels. In the case of the (Ac)NMe-8mer, glycol addition shifts the 19γ2 peak upfield. However, in all three peptide systems, increasing the glycol concentration results in larger (Δδ/ΔT)/CSD ratios. This would suggest that the C-terminal hydrophobic cluster is more likely to form (and less likely to unfold with increasing temperatures) the more polar (or "lipophobic") the media.

3.5.3: EFFECT OF PEPTIDE LENGTH

A general trend emerges from the fragments in comparable media. The Δδ/ΔT-CSD correlation of the 8mer (40% glycol, -4.191ppt/K) is smaller than that of the 10mer (45% glycol, -4.595ppt/K). The 6mer in 30% glycol (-5.335ppt/K) and Ac-AIIY in 40% HOAc (-7.520ppt/K) have larger correlations than either the 8mer or 10mer. However, this may reflect incomplete cluster formation in the shorter peptides. A more dramatic difference is observed when HFIP is added. The 8mer (15% HFIP) and 12mer (20% HFIP) have correlations of -4.528ppt/K and -6.712ppt/K, respectively. The slight difference in HFIP concentrations is not expected to have a dramatic influence in the

$\Delta\delta/\Delta T$-CSD ratios: the correlation of the 12mer in 40% HFIP is -6.826ppt/K. The $^{NMe}Ile^{20}$ analogs also demonstrate the same trends in both aqueous glycol/HOAc and aqueous HFIP conditions. Specifically, the (Ac)NMe-8mers have smaller (less negative) temperature gradient-CSD correlations than the (Ac)NMe-14mers. This may suggest that, in the absence of a rigid bicyclic core, an increase of the peptide length to include the helical region decreases the thermal stability of the C-terminal hydrophobic cluster.

3.5.4: ET-1 *vs.* PEN-1 *vs.* PEN-2.

A cooperative structure-stabilizing effect is also observed between the helical and C-terminal regions of the intact bicyclic 21mers. The penicillamine-containing peptides (Pen-1 and Pen-2) have the shallowest temperature gradients of their respective groups. Conversely, ET-1 has the largest gradient in the Ile^{20} series. In comparable conditions (50% ethylene glycol), the 19γ2 CSDs rank as follows: Pen-2 (-0.657ppm) >> ET-1 (-0.396) > Pen-1 (-0.321). The relative temperature gradients, however, have a different order: ET-1 (+2.480ppb/K) > Pen-2 (+2.173) >> Pen-1 (+0.935). On first inspection, there doesn't appear to be a correlation between the two data sets. The 19γ2 CSD would suggest that Pen-2 has, by far, the strongest hydrophobic interaction between Ile^{19} and Trp^{21}, but its temperature gradient ranks in the middle of the series. The temperature gradients indicate that ET-1 has the most thermally unstable hydrophobic cluster, but its CSD falls between Pen-1 and Pen-2. In the case of Pen-1, the CSD and temperature gradient appears to be contradict each other: the former indicates the weakest hydrophobic interactions in the series, but the latter indicates the most thermally stable cluster.

An examination of the $(\Delta\delta/\Delta T)$/CSD ratio is more enlightening. Here, ET-1 has the largest value (-6.712ppt/K), followed by Pen-2 (-3.131) and Pen-1 (-2.913). The Pen β-methyl groups have already been shown to increase the rigidity of the bicyclic core (see

Section 2.3.4), which also enhances and extends the helical region, regardless of media effects. The αH-CSD data (see Section 2.3.4) also shows that the Pen-2 helix is less stable than that of Pen-1 in both 0% and 15% HFIP. This would indicate that increased rigidity of the bicyclic core, not necessarily the helix, enhances the thermal stability of the C-terminal hydrophobic cluster.

3.6: N-METHYLATION AND TRP CHEMICAL SHIFT EFFECTS.

A notable problem was evident during the survey of the $^{(NMe)}Ile^{20}$ analogs: how to correct α-methine proton shifts of residues *i* and *i-1* with respect to the NMe group. Proline correction factors (see Appendix A) were initially used to estimate the effect of an N-methyl group on 19α and 20α. However, unusually large positive αH-CSD values resulted even when these corrections were applied to residues 19 and 20. For example, the NMe-14mer analogs displayed αH-CSDs of approximately +0.300ppm and +0.600ppm for 19α and 20α, respectively. This led to several questions. Previous circular dichroism studies (Harris, 1993) had suggested the presence of a turn-like feature at the C-terminus of Pen-2. Could these large downfield chemical shift deviations be due to this turn or is this an artifact of inaccurately determined reference values? Also, what is the true effect of the Trp^{21} aromatic ring on the chemical shifts of Ile^{19} aliphatic sidechain protons? Are upfield chemical shifts observed without the presence of the tryptophan residue? Finally, how does the N-methylated isoleucine influence hydrophobic cluster formation?

No study regarding the backbone NH and αH chemical shift reference values of N-methylated peptides has yet appeared in the literature. In order to more accurately determine the N-methyl and Trp chemical shift effects on the C-terminal residues, several smaller peptide fragments were synthesized (see also, Table 4.1). These peptides were either 3 or 4 residues in length, small enough such that no stable conformation should be

present. These small peptides would, however, retain ϕ/ψ torsion limits imposed by N-methylation alone. In addition, the 3- and 4mers were N-capped with an acetyl group to negate 2- and 3-effects on the amide chemical shifts. The proton chemical shifts of residues 19 and 20 were used to probe structural effects of the various mutations.

On a side note, minor conformers (*circa* 15% of the population) are readily observed in the 1D spectra of the N-methylated 4mers. These would correspond to the 19/20 *cis* peptide bond. The identity of the major (*trans*) configuration was verified, throughout, by the ROESY spectra (not shown), which displays the same intense 19α → 20NMe crosspeak observed in the larger fragments.

3.6.1: Effects of the Asp^{18} → Ala^{18} Mutation.

In neat DMSO, the Asp^{18} amide peaks of the various fragments were broadened, but sharpened considerably when molar equivalents of TFA were added to the medium. This change in ionization heavily influenced the amide and α-methine proton chemical shifts of the C-terminal residues. To negate sidechain ionization effects on chemical shifts, an Asp → Ala mutation at position 18 was performed. An initial survey of residues 19 and 20 (Table 3.13) indicate that the Asp → Ala substitution has negligible effects: most of the differences in chemical shifts are no greater than ±0.050ppm. In the case of the $^{(NMe)}$Ile^{20} analog, the deviation for 19NH is slightly larger than experimental error.

3.6.2: N-Methyl Effects on C-terminal Reporter Groups.

Table 3.14 lists the chemical shift deviations resulting from the Ile^{20} → $^{(NMe)}$Ile^{20} mutation. For comparative purposes, the intact ET-1 analogs (Pen-1 and Pen-2) and 3mers are included. Chemical shift deviations are generally larger for the intact peptides than for the fragments. This is especially noticeable for the 19γ2, 20α and 20β protons,

and may suggest that the C-terminal hydrophobic cluster is partially stabilized by the bicyclic core. In the case of the fragments, the deviations observed for the Ile^{19} protons are less than ±0.100ppm. The 19α and 19γ2 are the exceptions to this trend, but are greatly influenced by the presence of the tryptophan residue at position 21. For the Trp^{21} analogs, the CSD for the Ile^{19} α-methine is +0.408 ± 0.011ppm, while the Ala^{21} mutant displayed a larger difference (+0.476ppm). The large downfield deviation of the 19α proton appears to be inherent in the Ile-$^{(NMe)}$Ile moiety. The 3mer, containing Ile^{19} → Gly^{19} and Trp^{21} → Lys^{21} substitutions, shows less dramatic downfield shifts (+0.184ppm). The smaller difference observed for the 3mer 19NH is possibly due to increased mobility in the backbone.

Table 3.13: Asp^{18} → Ala^{18} substitution effects on the chemical shifts of Ile^{19} and Ile^{20} in aqueous media.

		Ac-DIXW (1) *vs.* Ac-AIXW (2)	
residue		$^{(NMe)}$Ile20	Ile20
		Δδ (1-2)	Δδ (1-2)
19	NH	-0.071	-0.033
	α	+0.000	+0.022
	β	-0.005	+0.016
	γ1	-0.033	-0.026
	γ1′	-0.020	-0.018
	γ2	-0.012	+0.001
	δ	+0.037	-0.009
20	NH	n/a	+0.007
	α	-0.010	-0.013
	β	-0.007	+0.001
	γ1	-0.002	+0.010
	γ1′	-0.005	-0.004
	γ2	+0.002	-0.003
	δ	+0.005	+0.002

Table 3.14: Effects of N-methylation on the chemical shifts of Ile^{19} and Ile^{20} in aqueous media.

residue		(Pen-2) – (Pen-1)	Ac-XI$^{(NMe)}$IX (1) *vs.* Ac-XIIX (2) D18, W21 Δδ (1-2)	A18, W21 Δδ (1-2)	A18, A21 Δδ (1-2)	[Ac-G$^{(NMe)}$IK] – [Ac-GIK]
19	NH	+0.074	+0.057	+0.095	+0.051	+0.019
	α	***+0.385***	***+0.397***	***+0.419***	***+0.476***	***+0.184, +0.184***
	β	-0.176	-0.093	-0.072	+0.021	
	γ1	+0.018	+0.039	+0.046	+0.082	
	γ1′	-0.111	-0.006	-0.004	+0.027	
	γ2	***-0.359***	***-0.276***	***-0.263***	***-0.008***	
	δ	-0.052	-0.008	-0.054	+0.021	
20	NH	n/a	n/a	n/a	n/a	n/a
	α	***+0.584***	***+0.552***	***+0.549***	***+0.536***	***+0.510***
	β	***+0.248***	***+0.231***	***+0.239***	***+0.228***	***+0.190***
	γ1	-0.142	-0.165	-0.153	-0.203	-0.110
	γ1′	-0.141	-0.181	-0.180	-0.186	-0.161
	γ2	-0.141	+0.048	+0.043	-0.011	-0.009
	δ	+0.016	-0.022	-0.025	-0.046	-0.012

The 19γ2 peaks have large upfield CSDs (-0.270 ± 0.007ppm) for the 4mers containing Trp^{21}. However, the Ala^{21} analog shows negligible N-methylation difference effects on the CSD. This indicates that the large upfield chemical shifts observed for 19γ2 in the N-methylated analogs is primarily due to anisotropy effects from the tryptophan sidechain indole ring. Once the Trp residue is replaced with Ala, the Ile^{19} β-methyl group displayed nominal chemical shift values.

In the case of residue 20, the N-methyl group had its largest effect on the protons closest to the backbone (i.e., the α, β, and, to a lesser extent, the γ1 protons). The Trp^{21}

$\rightarrow$ Ala^{21} substitution doesn't have as great an impact on the chemical shift deviations of the 4mers. The CSDs for 20α (including the Ala^{21} analog) is +0.546 ± 0.007ppm, for 20β, +0.233 ± 0.005ppm. These deviations are somewhat smaller for the 3mer. These results indicate that the N-methyl group, not deshielding effects from neighboring residues, has the greatest influence on the chemical shifts of residue 20.

3.6.3: Tryptophan Effects on C-terminal Reporter Groups.

The Trp^{21} effect on the C-terminal reporter groups is also displayed in Table 3.15. In the case of the $^{(NMe)}Ile^{20}$ analogs, the $Trp^{21} \rightarrow Ala^{21}$ substitution has its greatest influence on the Ile^{19} chemical shifts: the deviations for all the mainchain and sidechain protons are greater than ±0.100ppm. Conversely, the chemical shifts of the residue 20 protons do not vary by more than ±0.090ppm between the two peptides. When the same substitution is applied to the Ile^{20} analogs, smaller CSDs are observed for the residue 19 protons, while those of residue 20 display comparable deviations within experimental error. The $Tyr^{21} \rightarrow Ala^{21}$ mutation displays even smaller deviations for the Ile^{19} protons. The deviations observed for Ile^{20} are comparable with respect to the two previous cases. An examination of the $Trp^{21} \rightarrow Tyr^{21}$ difference indicates that the substitution has a greater effect on the Ile^{19} chemical shifts (upfield deviations > 0.050ppm, except for 19NH and 19γ1'), but doesn't greatly effect the chemical shifts of Ile^{20} (all deviations = ± 0.010ppm).

This is further evidence that the hydrophobic cluster observed in the longer peptide fragments is still formed in the 4mers and is enhanced by the N-methyl group. The upfield chemical shift deviations observed in the Tyr^{21} analog suggests that the 19γ2 methyl group is located in the shielding cone of the tyrosine phenyl ring. The smaller CSDs would indicate that the cluster is not as stable as in the case of the Trp^{21} analogs. However, a more reasonable explanation is that the Trp sidechain has larger ring current

effect on the chemical shift relative to that of Tyr. For the structure displayed in Figures 3.10 and 3.22, MOLMOL (Koradi et al., 1996) predicts an average upfield Trp contribution of 0.535ppm to the 19γ2 chemical shift. If the Trp residue is replaced with Tyr, the predicted chemical shift contribution from the aryl ring is -0.330ppm.

Table 3.15: Residue 21 substitution effects on the chemical shifts of Ile^{19} and Ile^{20} in aqueous media.

		Ac-AIXX (1) *vs.* Ac-AIIA (2)			
residue		$^{(NMe)}$I20, W21	I20, W21	I20, Y21	Trp → Tyr effect
		Δδ (1-2)	Δδ (1-2)	Δδ (1-2)	
19	NH	+0.105	+0.061	+0.060	+0.001
	α	***-0.238***	***-0.181***	***-0.097***	***-0.084***
	β	-0.264	-0.171	-0.099	***-0.072***
	γ1	-0.126	-0.090	-0.011	***-0.079***
	γ1′	-0.129	-0.098	-0.071	-0.027
	γ2	***-0.592***	***-0.337***	***-0.189***	***-0.148***
	δ	-0.151	-0.076	-0.023	-0.053
20	NH	n/a	-0.019	-0.012	-0.007
	α	***-0.076***	***-0.089***	***-0.102***	+0.013
	β	-0.087	-0.098	-0.097	-0.001
	γ1	-0.078	-0.128	-0.118	-0.010
	γ1′	-0.064	-0.070	-0.074	+0.004
	γ2	-0.057	-0.111	-0.114	+0.003
	δ	-0.044	-0.065	-0.063	-0.002

3.6.4: NME DISORDER REFERENCE VALUES AND CORRECTION FACTORS.

Local structuring effects induced by the N-methyl group were more accurately probed by comparing the NMR data from the several short C-terminal analogs and their respective NH and αH "disorder" reference values. To verify that these 3 and 4 residue peptides are relatively unstructured, the NH and αH-CSDs for the non-N-methylated

analogs are also included. The NH-CSD results for both DMSO and aqueous acetic acid media are listed in Table 3.16, while the αH-CSDs are reported in Table 3.17. Based on their αH and NH-CSDs, the Ile^{20} analogs appear to be unstructured in both aqueous 5 to 10% acetic acid. The peptides in acidic DMSO also appear to have a disordered conformation: the αH-CSDs are less than ±0.100ppm. However, large deviations in the amide chemical shifts are observed in DMSO media.

Table 3.16: NMe effects on NH-CSDs of residues 19 → 21.

	DMSO/2% H_2O/3 equiv. TFA			5-10% HOAc/H_2O		
Peptide	19NH	20NH	21NH	19NH	20NH	21NH
Ac-$G^{(NMe)}$IK	-0.115	—	+0.150	-0.132	—	+0.301
Ac-DI$^{(NMe)}$IW	-0.344	—	+0.136	-0.006	—	+0.161
Ac-AI$^{(NMe)}$IW	-0.178	—	+0.138	+0.065	—	+0.155
Ac-AI$^{(NMe)}$IA	-0.170	—	+0.168	-0.040	—	-0.037
average [1]	***-0.202***	—	***+0.148***	***-0.028***	—	***+0.145***
std. dev.	±0.086		±0.013	±0.071		±0.120
average [2]	***-0.231***	—	***+0.147***	***+0.006***	—	***+0.093***
std. dev.	±0.080		±0.015	±0.044		±0.092

Peptide	19NH	20NH	21NH	19NH	20NH	21NH
Ac-GIK	-0.013	-0.447	+0.214	-0.037	-0.015	+0.313
Ac-DIIW	-0.573	-0.433	+0.099	-0.063	-0.045	+0.100
Ac-AIIW	-0.414	-0.548	+0.078	-0.030	-0.052	+0.112
Ac-AIIA	-0.387	-0.564	+0.169	-0.091	-0.033	+0.038
Ac-AIIY	-0.426	-0.580	+0.043	-0.145	-0.137	-0.046
average [1]	***-0.363***	***-0.514***	***+0.121***	***-0.073***	***-0.056***	***+0.103***
std. dev.	±0.186	±0.062	±0.062	±0.042	±0.042	±0.119
average [2]	***-0.450***	***-0.531***	***+0.097***	***-0.082***	***-0.067***	***+0.051***
std. dev.	±0.072	±0.058	±0.046	±0.042	±0.041	±0.063

[1] Average + standard deviation *including* Ac-GXK.
[2] Average + standard deviation *excluding* Ac-GXK.

Table 3.17: NMe effects on αH-CSDs of residues 19 → 21.

	DMSO/2% H_2O/3 equiv. TFA			5-10% HOAc/H_2O		
Peptide [1]	19α	20α	21α	19α	20α	21α
Ac-G$^{(NMe)}$IK	+0.262	+0.508	-0.194	+0.147	+0.551	-0.039
	+0.126			+0.147		
Ac-DI$^{(NMe)}$IW	+0.227	+0.529	-0.136	+0.304	+0.557	-0.029
Ac-AI$^{(NMe)}$IW	+0.222	+0.575	-0.140	+0.304	+0.567	-0.015
Ac-AI$^{(NMe)}$IA	+0.289	+0.535	-0.183	+0.542	+0.643	+0.066
average [2]	***+0.225***	***+0.537***	***-0.163***	***+0.289***	***+0.580***	***-0.004***
std. dev.	±0.055	±0.024	±0.026	±0.145	±0.037	±0.041
average [3]	***+0.246***	***+0.546***	***-0.153***	***+0.383***	***+0.589***	***+0.007***
std. dev.	±0.030	±0.020	±0.021	±0.112	±0.038	±0.042

Peptide	19α	20α	21α	19α	20α	21α
Ac-GIK	+0.008	+0.034	-0.122	-0.037	+0.041	+0.052
	-0.041			-0.037		
Ac-DIIW	-0.018	+0.022	-0.072	-0.093	-0.035	+0.083
Ac-AIIW	-0.069	+0.011	-0.093	-0.115	+0.018	bleached
Ac-AIIA	-0.031	+0.011	-0.130	+0.066	+0.107	+0.089
Ac-AIIY	-0.078	-0.015	-0.111	-0.031	+0.075	+0.104
average [2]	***-0.038***	***+0.013***	***-0.106***	***-0.029***	***+0.041***	***+0.082***
std. dev.	±0.029	±0.016	±0.021	±0.064	±0.049	±0.019
average [3]	***-0.049***	***+0.007***	***-0.102***	***-0.043***	***+0.041***	***+0.092***
std. dev.	±0.025	±0.014	±0.021	±0.070	±0.054	±0.009

[1] Ile αH reference values were used for the CSD calculations of the $^{(NMe)}$Ile analogs.
[2] Average + standard deviation *including* Ac-GXK.
[3] Average + standard deviation *excluding* Ac-GXK.

The $^{(NMe)}$Ile^{20} analogs, in contrast, display large NH and αH-CSD values over residues 19 through 21. In DMSO, where the C-terminal hydrophobic cluster is less

likely to be observed, the effect on 19NH is especially dramatic for the Ile-Ile peptides. If the 3mers are excluded from the data set, the average Ile^{19} NH-CSD are –0.231 ± 0.080ppm and –0.450 ± 0.072ppm for the $^{(NMe)}Ile^{20}$ and Ile^{20} analogs, respectively. The averages in aqueous acetic acid are +0.006 ± 0.044ppm [$^{(NMe)}Ile^{20}$] and –0.082 ± 0.042ppm [Ile^{20}]. The differences between the Gly^{19} and Ile^{19} analogs are primarily due to the increased mobility inherent in the glycine backbone. In both media, the N-methyl group shifts the 19NH peak upfield and the 21NH peak downfield with respect to disorder values. The CSDs for the 21NH, 19α and 20α peaks are similar in the two media. In the case of the 19α proton, the CSD change approaches the currently used proline correction factor (α_{i-1}: reference + 0.29ppm). This suggests that the N-methyl group imparts similar structural constraints to the backbone as a proline residue. In contrast, the 21α experiences an upfield deviation in both media, although the uncertainty is higher in the aqueous acetic acid case. As expected, the N-methyl effect is greatest on its own αH. The observed CSD in both media are greater than +0.500ppm (+0.537 ± 0.024ppm in acidic DMSO, +0.580 ± 0.037ppm in aqueous acetic acid). These deviations are far stronger than those caused by a "normal" proline. For example, the difference in reference values for a proline α-methine proton and an amino acid with an unbranched, extended sidechain such as lysine is +0.070ppm in DMSO and +0.110ppm in water.

Can accurately determined correction factors be derived and applied to the longer fragments? The answer is yes, but with a few caveats. The correction to 20α might only be applicable to the N-methylated isoleucines. The case of the 19α appears to be sequence dependent, at least in aqueous media: the CSD for the Gly^{19} analog is half of those corresponding to the Ile^{19} peptides. This may be partially explained by the increased backbone mobility of Ac-G$^{(NMe)}$IK at residue 19.

Tentatively, the current reference values for the α-methine of $^{(NMe)}Ile$ is 4.75ppm in DMSO and 4.73ppm in aqueous media. The correction factor used for the αH of a

residue preceding a proline appears to be sufficient for a residue immediately preceding an $^{(NMe)}$Ile. However, this value may be slightly overestimated. General tendencies can also be observed for the amide of the preceding residue, as well as the amide and α-methine of the succeeding residue. In the case of the preceding amide, the chemical shift moves upfield (-0.202 ± 0.086ppm in DMSO, -0.028 ± 0.071ppm in aqueous media). The succeeding amide is downfield with respect to disorder values (+0.148 ± 0.013ppm in DMSO, +0.145 ± 0.120ppm in water). The α-methine proton for the same residue has a negative shift correction (-0.163 ± 0.026ppm in DMSO, -0.004 ± 0.041ppm in water). Noticeable differences in the chemical shift deviations are also observed between the I$^{(NMe)}$I and $^{(NMe)}$I moieties, although this most likely an artifact of increased backbone mobility inherent in the Gly^{19} residue and is more dramatic in aqueous media. For example, both amides of the residues preceding and succeeding to the N-methylated Ile experience larger deviations when residue 19 is a glycine (Tables 3.16 and 3.17). In addition, the deviation for the 19αH is much smaller for the Gly^{19} series.

CHAPTER 4: CD STUDIES OF ET-1 AND PEN-2 C-TERMINAL ANALOGS

4.1: INTRODUCTION: UNUSUAL CD SIGNATURES

The initial survey of intact and fragment ET-1 analogs (Harris, 1993) yielded unusual CD signatures. This was observed in a series of intact N-methylated analogs, which displayed larger than expected negative bands from *circa* 225 → 232nm. These bands were less intense in the non-N-methylated counterparts. At the time, these bands were attributed to a "β_I turn-like" conformation at the C-terminus in the case of Pen-1 and a "β_{II} turn-like" conformation for Pen-2. However, this region of the CD also corresponds to the aromatic side chain transitions, specifically those of Trp and Tyr (Brahms and Brahms, 1980). Given that, on a per residue basis, there are a relatively high percentage of aromatic residues in the sequence (the 14mers contain 3 aromatic residues, while the 4mers contain 1 aromatic residue), the sidechain moieties could significantly effect the resulting CD signatures.

Although Brahms and Brahms (1980) had earlier determined the CD signatures of random coil aromatic residues, no temperature studies (with or without fluorinated alcohols) were completed. As a result, correction factors for use in calculating fractional helicities were unavailable. Because of this, the true nature of the C-terminal backbone conformation was unresolved. That is, there was high uncertainty whether the unusual CD signatures are indicative of the N-methyl group or of the $I^{(NMe)}I$ moiety, and whether aromatic residues at position 21 enhance this CD signature.

4.2: MATERIALS AND METHODS

4.2.1: PEPTIDE SYNTHESIS – SHORT SEQUENCES

All 3mers and 4mers (see Table 4.1) not containing the $^{(NMe)}$Ile residue were synthesized using standard solid phase "Fast-moc" (HBTU as the coupling reagent) protocols on an Applied Biosystems 433A peptide synthesizer. All peptides containing the Ile-$^{(NMe)}$Ile moiety were manually synthesized using PyBrOP (Coste et al., 1994) as the coupling reagent with reaction conditions as previously described in Chapter 3. HATU (Carpino, 1993) was used as the reagent for the Ac-G$^{(NMe)}$IK synthesis, with reaction times of 2 hours for both coupling cycles. HATU was purchased from PerSeptive Biosystems (South San Francisco, CA). All purified peptides lyophilized as white fluffy powder, except Ac-G$^{(NMe)}$IK which appears as an oily bead.

Table 4.1: Peptides – Short Fragments

Sequence	Name	FW	ESI-MS Data	
Ac-G$^{(NMe)}$IK	3mer #1 (19-21)	372.44	$[M+H]^+$ $[M+H]^{2+}$	373.1 (Base) 187.3 (3.1%)
Ac-G IK	3mer #2 (19-21)	358.44	$[M+H]^+$	359.6 (Base)
Ac-AI$^{(NMe)}$IW	4mer #1 (18-21)	557.32	$[M+H]^+$ $[M+Na]^+$	558.3 (16.9%) 580.1 (67.5%)
Ac-AI IW	4mer #2 (18-21)	543.67	$[M+Na]^+$ $[M+Na]^{2+}$	566.2 (Base) 283.1 (10.0%)
Ac-DI$^{(NMe)}$IW	4mer #3 (18-21)	601.31	$[M+H]^+$	602.6 (8.1%)
Ac-DI IW	4mer #4 (18-21)	587.68	M^{2+}	294.6 (7.9%)
Ac-AI$^{(NMe)}$IA	4mer #5 (18-21)	442.26	$[M+Na]^+$ $[M+H]^+$ $[M\text{-Ala-OH}]^+$	465.0 (34.7%) 442.8 (6.8%) 354.1 (Base)
Ac-AI IA	4mer #6 (18-21)	428.53	$[M+K]^+$ $[M+Na]^+$ $[M+H]^+$	467.7 (Base) 451.6 (92.8%) 429.7 (60.8%)
Ac-AI IY	4mer #7 (18-21)	520.29	$[M+H]^+$	524.9 (53.5%)

Peptides used for the determination of random coil Tyr and Trp signatures (KYK, KWK, TYS, WA and APGW-NH_2) were purchased from Sigma Pharmaceuticals (St. Louis,

MO). Three other peptides (GHKW, GHKF and GHK) used in the early CD studies (Harris, 1993) were gifts from ProCyte Corp. (Redmond, WA).

4.2.2: Sample Preparation

Peptides were dissolved in 5mM formate/acetate buffers to yield stock solutions, with nominal concentrations between 0.5 to 1.5mM and pHs of 2.0 to 6.0. CD sample concentrations were determined via UV spectroscopy, using the following extinction coefficients: ε_{274} Tyr = 1394 ± 6 and Trp = 5341 ± 8; ε_{278} Tyr = 1260 ± 2 and Trp = 5554 ± 12. Due to the hydrophobic nature of the peptides, all samples were prepared in conical microfuge tubes and centrifuged. The supernatant was transferred to another microfuge tube and used for further study.

Two other peptides (Ac-GIK and Ac-G$^{(NMe)}$IK) were also synthesized to determine temperature effects on the N-methyl group CD signature without interference from the sidechains of aromatic residues. A modified version of the trinitrobenzene sufonic acid (TNBS) assay (Satake et al., 1960; Fields, 1972; Sashidar et al., 1994), which is specific for primary amines, was used to more accurately quantify these CD sample concentrations. This procedure required a 0.1% TNBS (by weight) solution in 0.5M borate/phosphate buffers (pH 8.5). Small aliquots of the CD stock solution (50-100μL) in 3mL TNBS buffer were incubated in a 45°C water bath for 45min to 1.5hr in the dark, then immediately quenched with 3M phosphoric acid, bringing the nominal pH of the solution to 2.2-2.4. The trinitrophenyl-amine (TNP-amine) conjugate solution was allowed to equilibrate at ambient conditions prior to quantitative analysis in a dual beam UV spectrophotometer. Absorbance values were recorded at 340 ($\varepsilon \approx 11498 \pm 1137$/TNP-amine) and 420nm ($\varepsilon \approx 5029 \pm 463$/TNP-amine). The value at 600nm was used as the baseline absorbance value. Extinction coefficients at 340 and 420nm were calibrated using several peptide and protein samples containing both Lys and aromatic (Trp or Tyr)

residues. However, due to the persistence of picric acid, which has a strong absorbance at 340nm ($\varepsilon \approx 8000$), the 420nm band was used exclusively for the determination of peptide concentrations.

4.2.3: CD Spectroscopy

Aliquots of stock solutions were diluted 10 fold with either buffer, or buffers containing various percentages of fluoroalcohol. CD sample concentrations were also adjusted in order to make the solutions sufficiently transparent at wavelengths less than 190nm. All data whose corresponding dynode voltage (also known as the HT voltage) was less than 650 volts were treated with high confidence. Cut off points for the buffers and HFIP titration studies were generally 180-185nm.

All CD data was acquired on a JASCO Instruments Model 720 spectropolarimeter. Typical parameters for the CD experiments are as follows: range 185-270nm, step resolution 0.2nm, scan rate 100nm/min, number of scans 20, sensitivity 20-100mdeg, response time 0.25s, bandwidth 1nm. All temperature studies were acquired using a square quartz UV cell with a 1mm pathlength, in circa 10K intervals, starting from the low temperature and ranging from 273K to 333K. Temperatures were adjusted using a VWR temperature controller filled with a 50:50 water/ethylene glycol cryogen. The cryogen was pumped into a jacketed metal block that holds square quartz UV cuvettes of various pathlengths. Temperatures were measured directly from a thermometer placed in the metal block, located approximately 1cm away from the UV cell. The system was allowed to equilibrate for 5 - 10 minutes after reaching the desired temperature prior to spectral acquisition. After data acquisition, the curves were noise reduced, then converted to their corresponding optical constants and normalized at 260nm. To calculate the mean residue ellipticity, the curves were divided by the total number of amide bonds in the peptide.

An estimate of a peptide's fractional helicity, $\langle f_H \rangle$, can be calculated from the available CD data, using equation 4.1a or 4.1b (Andersen & Tong, 1997):

$$\langle f_H \rangle_{CD} = \frac{(\theta_{obs} - \theta_C)}{(\theta_H - \theta_C)} \qquad \text{[Eqn. 4.1a]}$$

$$\langle f_H \rangle_{CD} = \left(\frac{N-1}{n-1}\right) \cdot \frac{(\theta_{obs} - \theta_C)}{(\theta_H - \theta_C)} \qquad \text{[Eqn. 4.1b]}$$

where θ_{obs} is the observed mean residue ellipticity at 221nm. In cases where the helix spans only *n* of total of *N* residues in the sequence, a ratio of (N-1)/(n-1) is applied to the calculation (equation 4.1b). Typically, *n* is determined by NMR. The coil and 100% helix reference values at 273K are represented by θ_C and θ_H, respectively. Aromatic residues can influence the intensity of the 221nm band and do so differently depending on whether they are in a helical or disordered segment (Andersen and Tong, 1997). Phe and Tyr contribute positive increments while Trp contributes a negative increment within helices (Chakrabarty et al., 1993). As a result, correction factors are applied to the reference values. For a peptide with *N* residues,

$$\theta_C = +200 + \frac{1}{(N-1)} \cdot \sum \Delta\theta_C(\text{aryl}) \qquad \text{[Eqn 4.2]}$$

$$\theta_H = -42000 + \frac{1}{(N-1)} \cdot \sum \Delta\theta_H(\text{aryl}) \qquad \text{[Eqn 4.3]}$$

where $\Delta\theta_C$ and $\Delta\theta_H$ represent the molar coil and 100% helix values of the individual aromatic residues. The θ_H correction is applied only when the aryl amino acid is located in the helical span. Current values, which incorporates the coil reference values of Brahms and Brahms (1980), are listed in Table 4.2 (see also Section 4.4.5). Prolines are also included in Table 4.2 since the poly(Pro)$_{II}$ CD signature has a positive contribution to

the 221nm band. The disorder value corrections for Phe, Tyr and Trp incorporate the CD signatures determined in the present study.

Intensities of the CD curves, and therefore fractional helicities, are dependent on accurately determined sample concentrations. Another method used to estimate helicity from CD data is concentration-independent and relies on the relative intensities of the maximum and double minima (Bruch et al., 1991; Muñoz et al., 1995). The first, known as the R_1 parameter, is a ratio of $\theta_{max}/\theta_{min}$, where θ_{max} is the maximum (generally between 191 and 194nm) and the θ_{min} is the global minimum (between 196 and 208nm). A second, the R_2 parameter, is the ratio of $\theta_{221}/\theta_{min}$, where θ_{221} is the mean residue ellipticity at 221nm. Consensus values of the R_1 ratio range between +0.7 (disorder) and –2.4 (100% helicity). The corresponding values for R_2 fall between –0.2 (0% helicity) and +1.05 (100% helicity). Muñoz et al. (1995) note that the relationship between the ratios and percent helicity is non-linear and provide a third-order polynomial curve fit for R_1. Unfortunately, they don't provide a similar relationship for the R_2 ratio.

Table 4.2: Molar correction factors for aromatic residues at 221nm.

Residue	$\Delta\theta_C$ (0% HFIP)	$\Delta\theta_C$ (30% HFIP)	$\Delta\theta_H$ (est.) [1]
Trp	+24400	+22100	-44000
Tyr	+21200	+13500	+66000
Phe [2]	+14200	(+15300)	+36000
Pro	+9600 [1]		+15000

[1] From Andersen & Tong, 1997.

[2] Data for phenylalanine is less well-determined.

4.3: CD STUDIES OF THE LONGER FRAGMENTS

4.3.1: THE 14MERS.

Due to the hydrophobic nature of the peptide, the 14mer (DAEAVYFAHLDIIW) aggregated in purely aqueous solutions at NMR conditions. Addition of organic co-solvents such as acetic acid and fluorinated alcohols did not significantly affect the solubility of the sample (see also Chapter 3). Fortunately, the C-terminal hydrophobic cluster of the $^{(NMe)}Ile^{20}$ analogs prevented sample aggregation in aqueous media during the NMR studies. Although the N-methylated analogs displayed NOESY crosspeaks diagnostic of an α-helix, it was unclear what effects the N-methyl group had on the stability of the helical region relative to the 14mer. [A comparison of the Pen-1 and Pen-2 αH-CSD histograms (Figure 2.21) had suggested that the N-methyl group disrupts the C-terminal region of the α-helix.] As a result, CD studies were attempted to ascertain the extent of structuring in aqueous media. Unfortunately, the 14mer and AcNMe-14mer aggregated in purely aqueous media, while the NMe-14mer was sparingly soluble. Addition of HFIP helped alleviate the aggregation problems observed in the N-methylated analogs. The 14mer, however, was only soluble in aqueous 30% HFIP. The initial CD studies were performed at pH 4.06 using a 5mM formate-acetate buffer titrated with dilute NaOH or KOH. The N-methylated peptides were also studied at pH 2.6 (comparable to NMR conditions) and at pH 6.0 (the point at which the intact bicyclic 21mers start to aggregate in aqueous media).

The structuring effect of capping Asp^8 with an N-acetyl group is clearly evident: the CD curve of the AcNMe-14mer is more intense (top panel, Figure 4.1), suggesting that the NMe-14mer is less helical than its capped analog. As observed in the earlier Pen-2 CD studies (Harris, 1993), the N-methylated 14mers display an unusually large negative band between 225-230nm that is not observed in the 14mer. Since this artifact masks the 220-222nm band of helices, a straightforward comparison of the relative

helicities of the three peptides is not available. An additional complication is the Trp^{21} residue. According to previous CD studies (Brahms and Brahms, 1980; Harris, 1993), the sidechain of a random coil Trp residue has an intense positive CD absorption betwcen 225-230nm. To help deconvolute the θ_{221} band, difference spectra were calculated. The spectra of Ac-DIIW and Ac-AIIW (see also Section 4.4) were subtracted from the 14mer, while Ac-AI$^{(NMe)}$IW was subtracted from both N-methylated analogs. This effectively truncated the peptide such that the CD signatures represented residues 8 through 17 (lower panel, Figure 4.1). Table 4.3 reports the mean residue ellipticity values at 194, 207 and 221nm, as well as their respective initial temperature gradients for the corrected spectra. Table 4.4 lists the fractional helicities calculated from Eqn. 4.1, as well as R_1 and R_2 ratios, of the three 14mer analogs.

Table 4.3: CD optical constants and initial temperature gradients for residues 8-17 of the 14mers in aqueous 30% HFIP/5mM buffer at 273K. The signature of the terminal four residues was removed by subtracting the spectrum of Ac-XIXW.

Peptide	Mean Residue Ellipticity θ_{221nm}	θ_{207nm}	θ_{194nm}
14mer			
pH 4.06 (*vs.* A^{18})	-11100 + 197T	-14500 + 129T	+48200 – 408T
pH 4.06 (*vs.* D^{18})	-11700 + 192T	-14300 + 136T	+51100 – 434T
NMe14mer			
pH 2.60	-19900 + 278T	-22400 + 171T	+88800 – 787T
pH 4.06	-9630 + 154T	-13700 + 118T	+55500 – 472T
pH 6.00	-6700 + 118T	-13500 + 114T	+46500 – 437T
AcNMe-14mer			
pH 2.60	-21500 + 185T	-26200 + 182T	+85400 – 554T
pH 4.06	-13400 + 113T	-18300 + 74T	+61300 – 427T
pH 6.00	-11400 + 118T	-19000 + 103T	+52800 – 432T

Table 4.4: Fractional helicity, R_1 and R_2 values for the 14mers in aqueous 30% HFIP/5mM buffer at 273K, based on the data in Table 4.3.

Peptide	$\langle f_H \rangle_{CD}$ (Eqn. 4.1)	R_1 ($\theta_{max}/\theta_{min}$)	R_2 ($\theta_{221}/\theta_{min}$)
14mer			
pH 4.06 [‡]	54.21 ± 1.10%	-3.45 ± 0.13%	+0.80 ± 0.03%
NMe14mer			
pH 2.60	85.35%	-3.96	+0.89
pH 4.06	47.73%	-4.05	+0.70
pH 6.00	37.00%	-3.44	+0.50
AcNMe-14mer			
pH 2.60	85.58%	-3.26	+0.82
pH 4.06	57.38%	-3.35	+0.73
pH 6.00	50.42%	-2.78	+0.60

[‡] Average values and standard deviations *vs.* the D^{18} and A^{18} analogs.

With the exception of a more intense maximum at 194nm, subtraction of the Asp^{18} analog from the 14mer (thin solid line) was comparable to that of the Ala^{18} peptide (thin dashed line). Initial temperature gradients (Figure 4.2) between the two sets are also within experimental error. Based on the θ_{221} ellipticity values and the calculated $\langle f_H \rangle$, the NMe-14mer (thick dashed line, Figure 4.1) appears to be less helical than the 14mer, lending further evidence that the N-methyl group disrupts the C-terminal region of the helix. As expected, capping the NMe-14mer with an N-acetyl group (thick solid line, Figure 4.1) increases helicity, as well as the thermal stability of the helical region: θ_{221} and $\langle f_H \rangle$ values are larger, while the temperature gradient is smaller for the AcNMe-14mer. The R_1 and R_2 ratios also fit the pattern, although the absolute R_1 values are much larger than expected. This may indicate that there is incomplete subtraction at the 194nm band or too much subtraction at the 207nm band. That is, the maximum at 194nm is too intense and/or the minimum at 207nm is too weak. Other factors to consider are the remaining aromatic residues, Tyr^{13} and Phe^{14}. Although the CD spectra of disordered Tyr

and Phe have large negative molar-residue ellipticity values in this region (Phe, $\theta < 0$ for $\lambda < 191$nm; Tyr, $\theta < 0$ for $\lambda < 194$nm) their ellipticity values at the far UV extrema to the blue of 221nm are not known. On a molar scale, the two remaining aromatic residues (out of 10) may have a significant collective effect on the 194nm CD band (and possibly the 207nm band).

The CD signatures (Figure 4.3) also appear to be pH dependent in the case of the NMe-14mers: the helix-like signal becomes more intense as the medium becomes more acidic. The initial temperature gradients also appear to be steeper in the more acidic conditions. However, in terms of percent signal lost per °C, some variance is observed for the two peptides. In the case of the NMe-14mer, these values increase from 1.40% (at pH 2.6) to 1.60% (at pH 4.06) to 1.76% (at pH 6.0) for the θ_{221} band. The CD signature for the AcNMe-14mer is more thermally stable at the same pH conditions: 0.86%, 0.84% and 1.04% loss in signal, respectively. These changes may be attributed to the Asp^8 ionization state. At pH 2.6, the Asp^8 sidechain is expected to be the neutral carboxylic acid, while at pH 4.06 and 6.0, it would be a carboxylate anion. The former may be a better helix capping unit (Andersen & Tong, 1997) than the latter. However, Asp^- should stabilize the helix macrodipole more readily than Asp°. An examination of the fractional helicities reveals the importance of the N-acetyl cap and the ionization state of the acidic residue sidechains. At pH 2.6, both peptides have, within experimental error, the same relative helicities (85.35% for the NMe-14mer *vs.* 85.58% for the AcNMe-14mer). The NMe-14mer loses a greater amount of its helicity than the AcNMe-14mer as the medium becomes less acidic. This suggests that in the more acidic medium, the backbone N-acetyl group has a greater effect on helicity than sidechain ionization states.

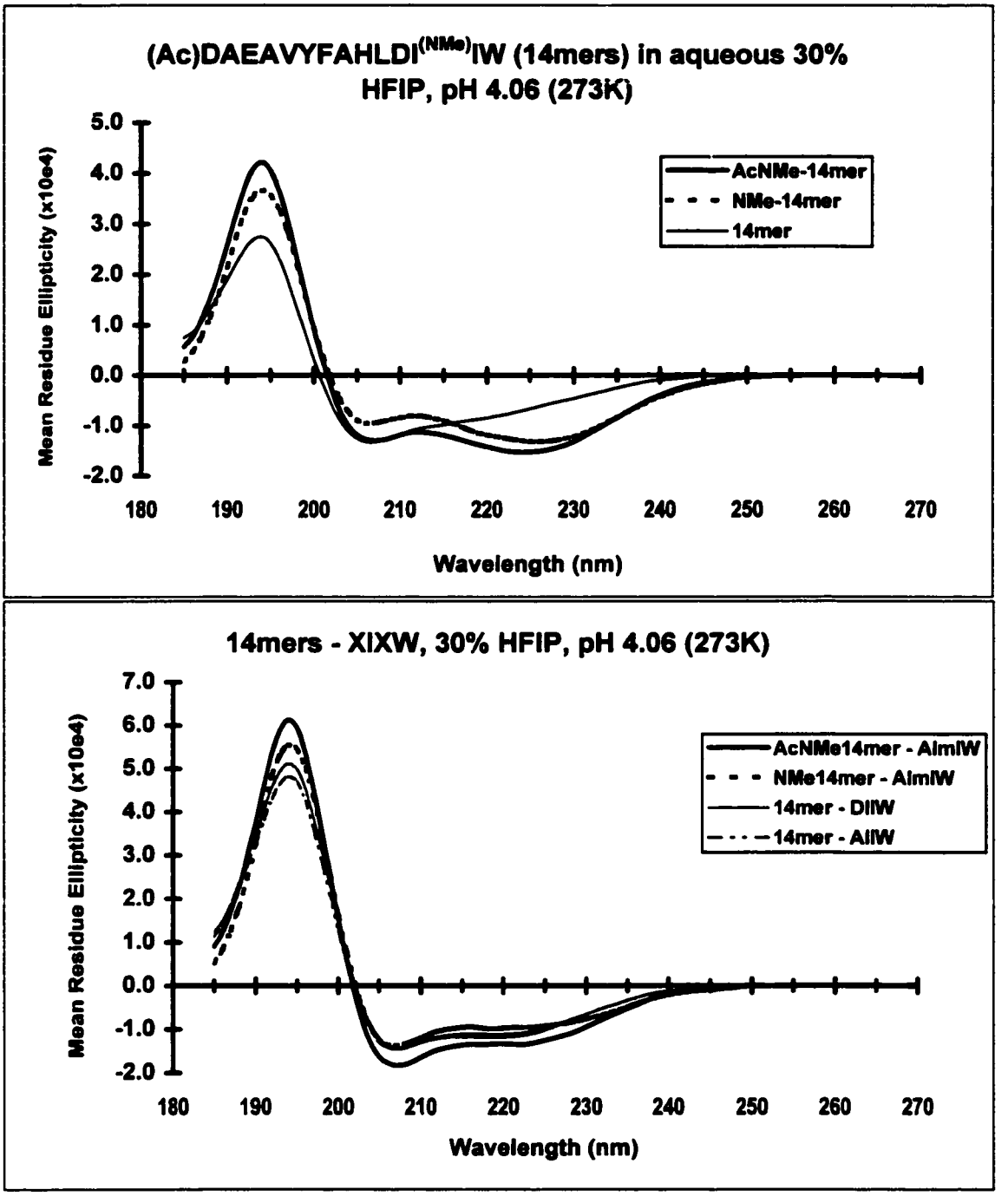


Figure 4.1: CD signatures of the 14mers in aqueous 30% HFIP, pH 4.06 (273K). Top panel, uncorrected curves. Bottom panel, difference spectra (14mers – XIXW).

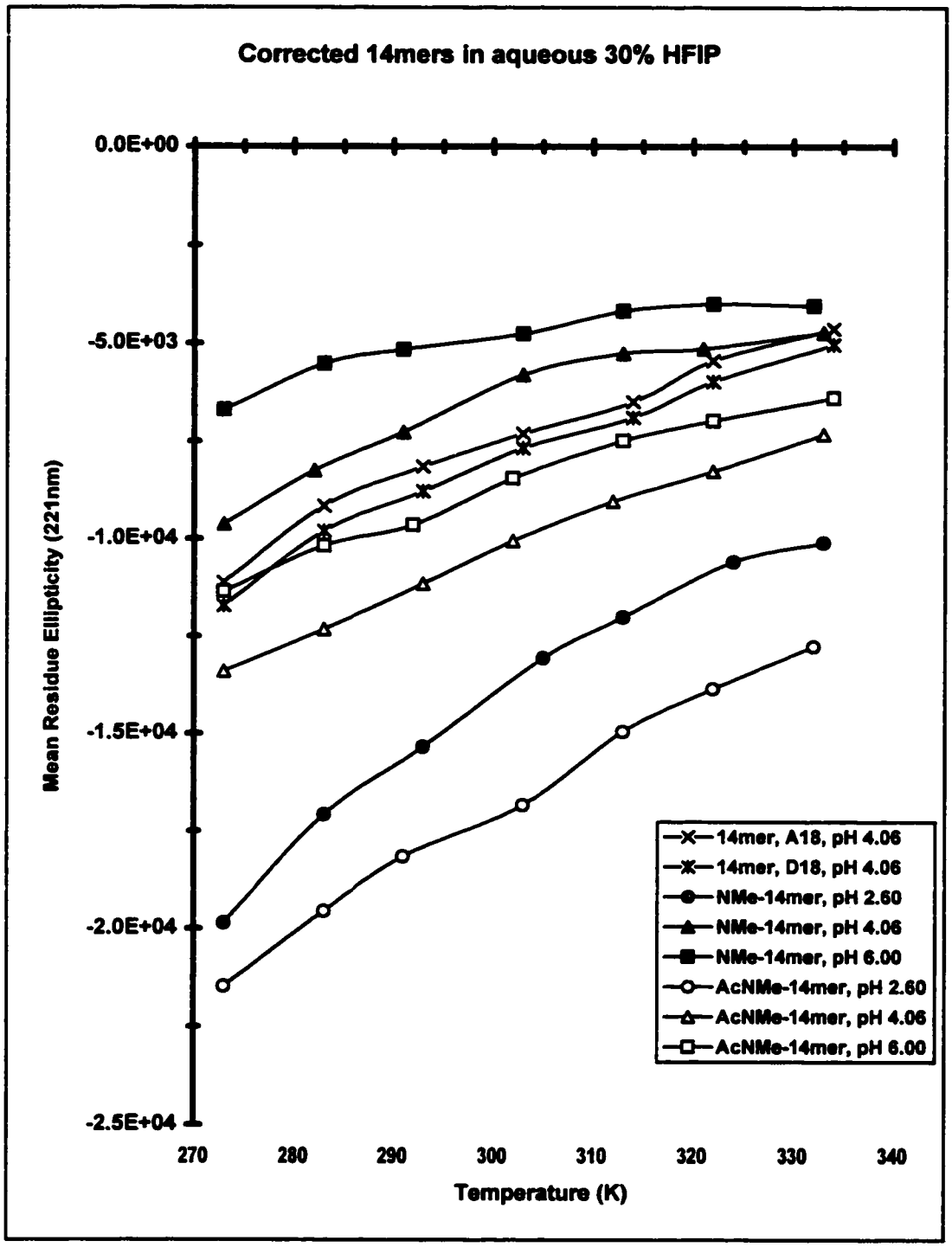


Figure 4.2: Temperature and pH dependence on the θ_{221} band of the corrected 14mers in aqueous 30% HFIP.

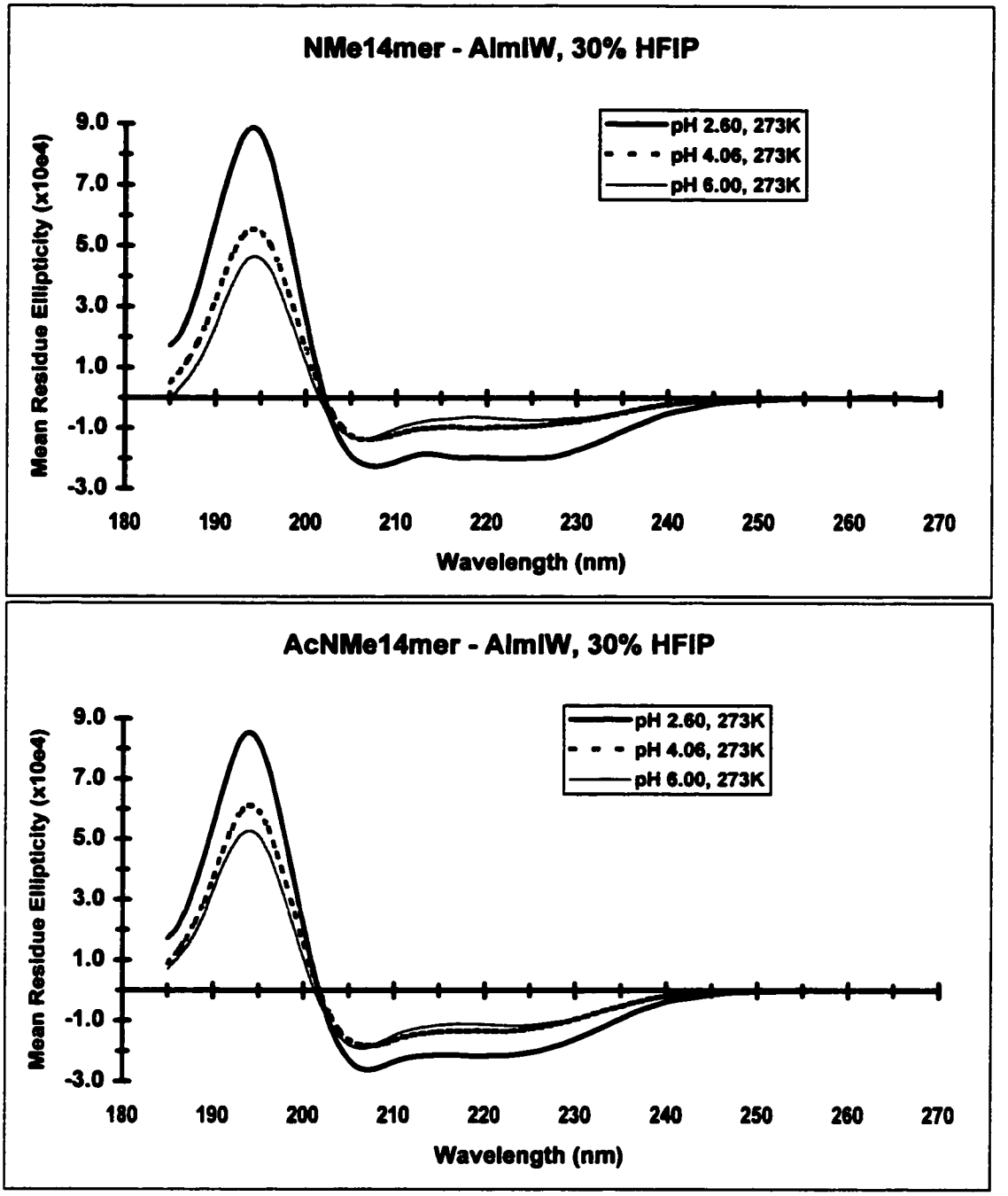


Figure 4.3: The pH dependence of the corrected NMe-14mer CD signatures in aqueous 30% HFIP, pH 4.06 (273K). Top panel: NMe-14mers; bottom panel: AcNMe-14mer.

4.3.2: THE 8MERS.

The 8mer (GSHLDIIW) displays a standard disordered coil CD spectrum in purely aqueous media (Figure 4.4). Table 4.5 lists the mean residue ellipticity values and initial temperature gradients of the 8mers. Addition of fluoroalcohols decreases the intensities of the random coil CD signature (minimum at 197.2nm). The small maximum at 225.6nm also decreases in intensity and experiences a red shift with 30% HFIP. The NMe-8mer displays an unusual CD spectrum that doesn't correspond to any of the standard secondary structure motifs. In this case, an intense double minimum, centered at the 198.0 and 230.0nm bands, is observed. Addition of HFIP decreases the intensity of the minimum at 198nm; however, the second minimum at 230nm increases in intensity and has a slight blue shift to 229nm. In terms of percent signal lost per °C, the CD signatures of the 8mers are more thermally stable in the presence of HFIP.

Table 4.5: CD optical constants and initial temperature gradients of the 8mers in aqueous HFIP/5mM buffer at 273K.

		Mean Residue Ellipticity		
Peptide	θ_{221nm}	θ_{max1}	θ_{min}	θ_{other}
8mer				
0% HFIP	67 – 66T	1760 – 42T (225.6nm)	-25300 + 154T (197.0nm)	
30% HFIP	-1200 – 12T	-107 – 11T (226.4nm)	-16100 + 16T (197.0nm)	
NMe-8mer				
0% HFIP	-6310 – 84T		-18100 + 129T (198.0nm)	-10800 – 29T **(230.0nm)**
30% HFIP	-8780 – 26T		-10100 + 75T (198.0nm)	-12300 + 2T **(229.0nm)**

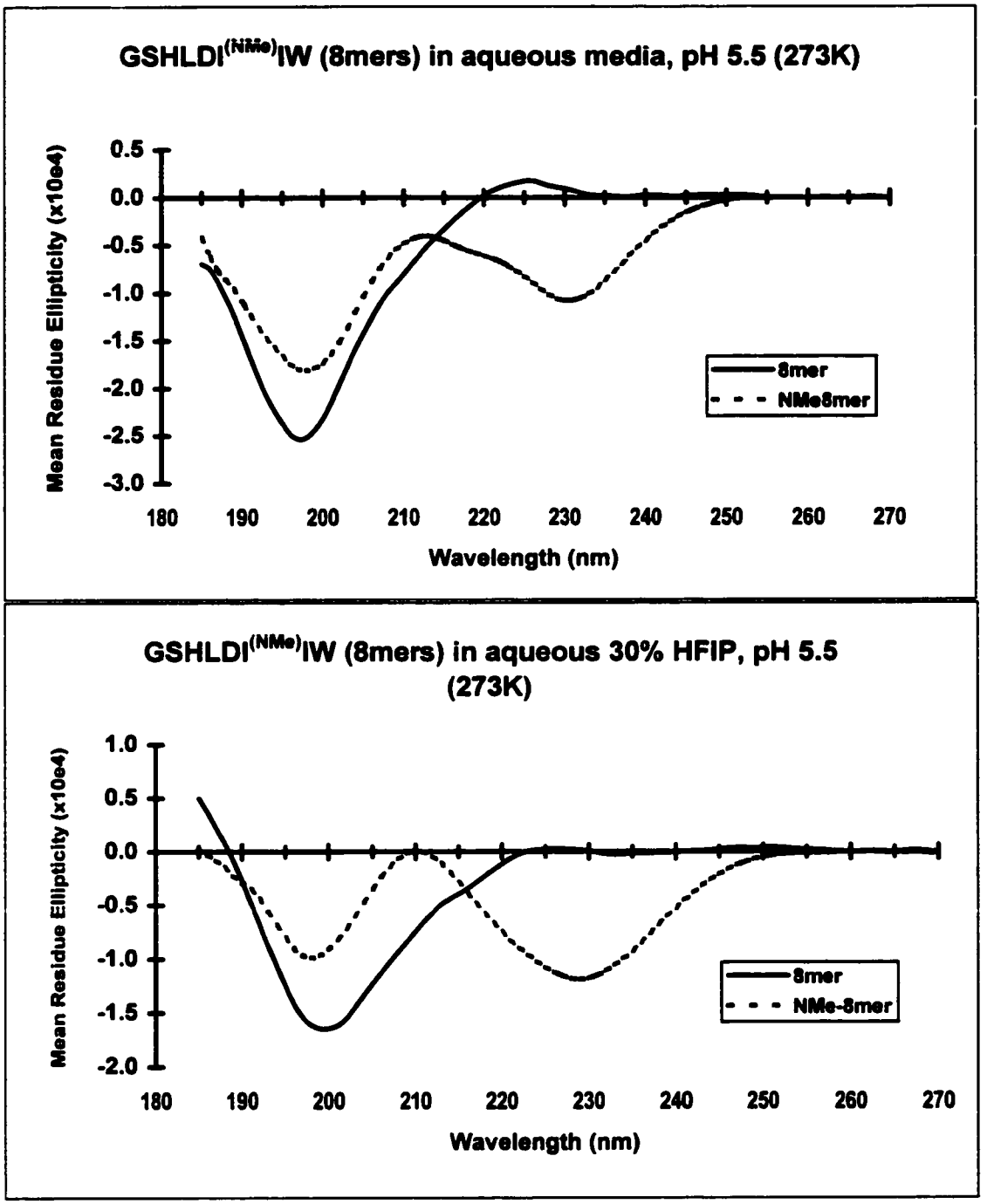


Figure 4.4: CD spectra of the NMe-8mer and 8mer in aqueous media (top panel) and aqueous 30% HFIP (lower panel) at pH 5.5.

4.3.3: INTERMEDIATE PEPTIDES: 9MER, 10MER AND 12MER.

Other C-terminal fragments of various lengths were also available for study. The 10mer (VYFAHLDIIW) was a gift from Bristol Myers Squibb, while the 12mer (EAVYFAHLDIIW) was synthesized for the current study. The initial 12mer synthesis yielded a major deletion product, henceforth called the 9mer (EAVYFAHLD). Table 4.6 reports the mean residue ellipticity values and initial temperature gradients for the three peptides. The 9mer has certain advantages over the 12mer. Since the Trp and two Ile residues have been truncated from the peptide, the 9mer is less hydrophobic than the 12mer (and 14mer). The missing Trp residue is also beneficial to CD studies since the large signals associated with a disordered Trp sidechain are eliminated. Fortunately, the concentrations of the CD samples could be accurately quantified using the sidechain chromophores of Tyr^{13} and Phe^{14}. The CD spectrum of the 9mer in a purely aqueous medium (pH 5.5) is a standard disorder signature (Figure 4.5, see also Table 4.6) at low HFIP levels. At 25% HFIP, the coil spectrum is virtually gone. The spectral intensity in 25% HFIP was very weak (at 273K, $\theta_{221} = -1590$, corresponding to a $\langle f_H \rangle$ of 12.77%).

The 12mer and the 14mer displayed aggregation problems in purely aqueous media. As a result, serial dilutions of the stock solution, which contained 30% HFIP, was attempted. The spectra at the lower HFIP levels (data not shown) suggest the presence of an aggregate, specifically a high level of spectral noise and a weak CD signature resembling a β-conformation. With the exception of a slight reduction in signal intensity, as well as a slight red shift in the minimum, no significant changes in the spectra were observed between 0% and 9% HFIP. In 30% HFIP, the 12mer displays a CD signature (Figure 4.6) which fully resembles a β conformation. The initial temperature gradient of the minimum (218nm) is shallow, indicating a thermally stable (presumably) oligomer conformation. After 20°C, the 12mer undergoes rapid thermal melting or disaggregation.

Table 4.6: CD optical constants and initial temperature gradients of the intermediate peptides in aqueous HFIP/5mM buffer at 273K.

Peptide	Mean Residue Ellipticity θ_{221nm}	θ_{max}	θ_{min}	θ_{other}
10mer				
0% HFIP	+136	+2210 (228.0nm)	-4260 (213.8nm)	
3% HFIP	-384	+2130 (228.6nm)	-3180 (213.0nm)	
8% HFIP	-3520	+12600 (228.8nm)	-21900 (212.0nm)	
16% HFIP	-2570	+7980 (228.4nm)	-16100 (212.0nm)	
32% HFIP	-3070	-190 (229.2nm)	-6050 (214.2nm)	
50% HFIP	-3300	-715 (~229.2nm)	-5710 (~214.2nm)	
9mer				
0% HFIP	1990 – 77T	2360 – 87T (224.4nm)	-21600 + 177T (195.0nm)	
10% HFIP	1170 – 113T	1730 – 92T (226.2nm)	-17400 + 275T (196.0nm)	
25% HFIP	-1590 – 67T		-4300 – 10T (204.0nm)	
12mer				
30% HFIP	-6580 + 6T	11600 – 46T (194.4nm)	-7380 – 7T (218.0nm)	1190 – 3T (235.4nm)

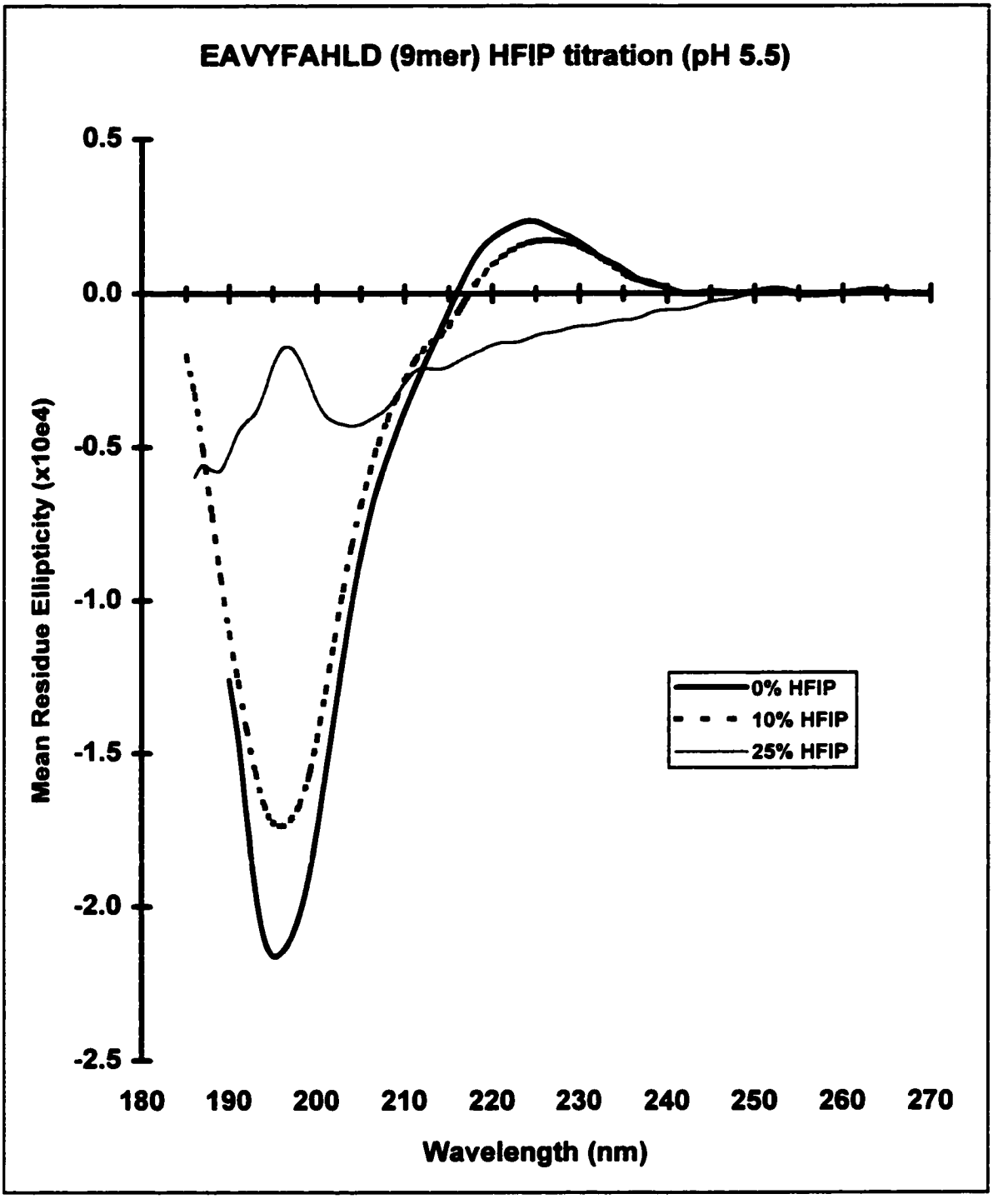


Figure 4.5: The HFIP titration of the 9mer in aqueous buffer (pH 5.5) at 273K.

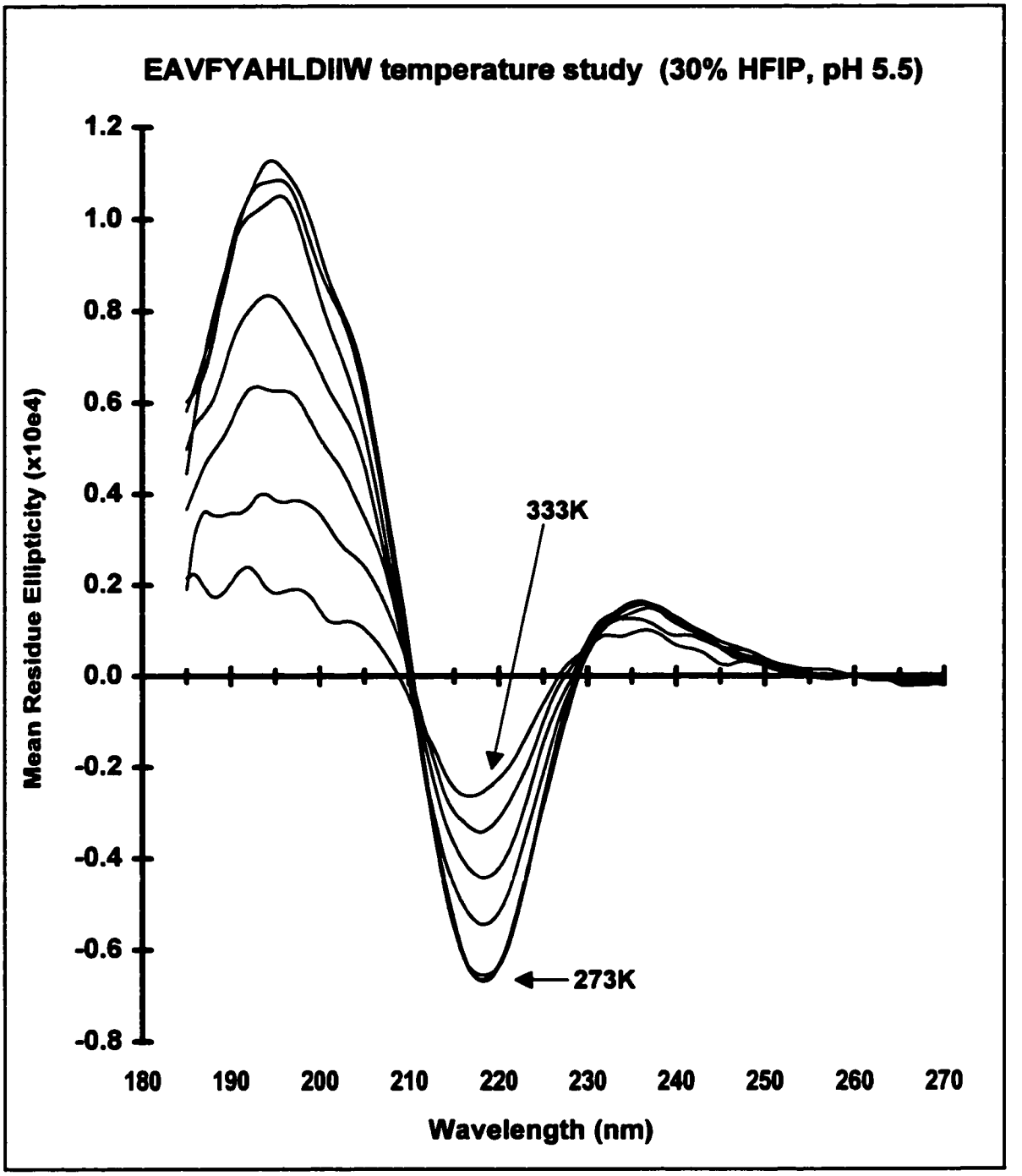


Figure 4.6: The temperature CD study of the 12mer in aqueous 30% HFIP (pH 5.5).

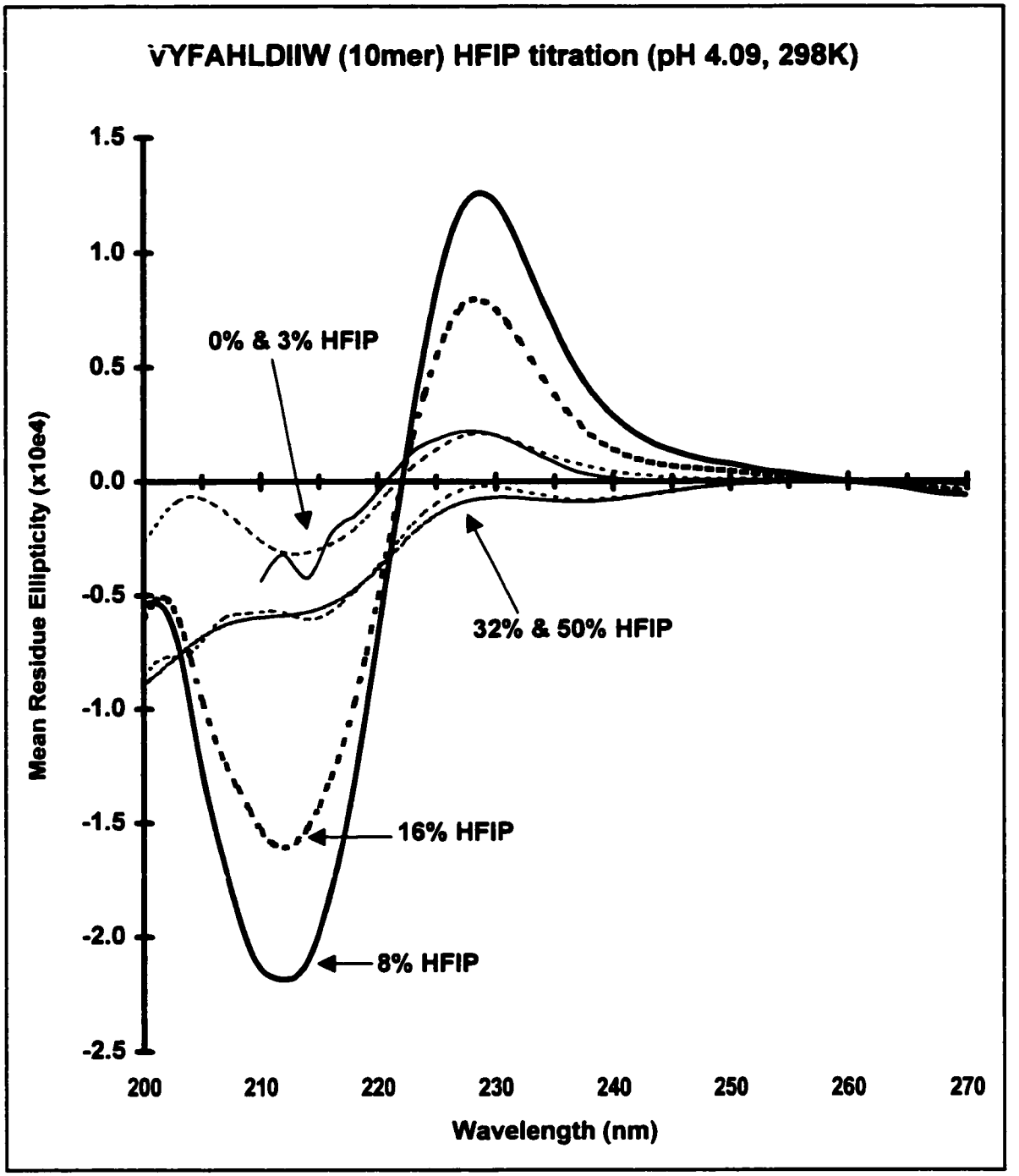


Figure 4.7: The HFIP titration of the 10mer in aqueous buffer (pH 4.09) at 298K.

The 10mer displays an unusual pattern during the HFIP titration (Figure 4.7). At the end points of the titration (32% and 50% HFIP), the spectrum is "coil-like" with a slightly more intense "minimum" at *circa* 214nm. The intermediate points of the titration (8% and 16% HFIP) display the most unusual feature: a maximum at 228.4nm (θ = +12600 and +7980 at 8% and 16% HFIP, respectively). This value is much greater than one-tenth of the molar ellipticity of a Trp residue at 226nm. In contrast, the "maxima" of the high HFIP conditions are located at *circa* 229.2nm and have small *negative* ellipticity values (-190 and -715 at 32% and 50% HFIP, respectively). The minimum has also blue shifted to 212nm (θ = -21900 at 8% HFIP -16100 at 16% HFIP). These minima are 2.5-3.5 times larger than those observed for the peptide at the high HFIP concentrations. Because the 10mer is a hydrophobic peptide, aggregation of the β-strands is likely to occur. At the intermediate HFIP levels, it appears that the Trp^{21} sidechains of the individual strands stack upon each other, leading to a dramatic increase in the signal intensity. The blue shifted minimum may represent a combination of the β-strand and disordered Trp sidechain absorptions. High concentrations of HFIP appear to destabilize the intermolecular Trp sidechain-sidechain interactions, leading to a decrease in signal intensity.

One question arises from these intermediate-sized peptides: why is there a dramatic change in conformational preferences between comparable peptides of various lengths? The 8mer was expected to be mostly disordered since it only contained three residues whose positions placed them in the helical region of the intact 21mers. The 10mer would be more likely to form aggregates than the 8mer due to the addition of two residues (Val^{12} and Tyr^{13}) and two substitutions ($Gly^{14} \rightarrow Phe^{14}$ and $Ser^{15} \rightarrow Ala^{15}$) which made the peptide more hydrophobic. Differences between the 9mer and 12mer are solely due to the deletion of the three hydrophobic C-terminal residues. As a result, the 9mer is less likely to form β-aggregates. The dramatic changes in conformation between the

14mer and 12mer may be linked to their respective capping units. Not only is the 14mer two residues longer, and therefore incorporates a longer segment of the 21mer helical region, but it contains an Asp rather than a Glu as its N-terminal residue. It has been well established (for example: Andersen & Tong, 1997) that Asp, either in its ionic or neutral form, is a better helix N-capping unit than Glu. In the absence of helicity, the hydrophobic peptides appear to form oligomers.

An important note regarding the 12mer and 14mer: there is a slight pH difference between the two samples. The 14mer was acquired in pH 4.06 buffer, while the 12mer was solvated in pH 5.5 buffer. However, if the CD data for the corrected 14mer and NMe-14mer spectra are comparable for all pHs, the 14mer should retain its helical CD signature, although signal intensities would presumably be smaller.

4.4: DETERMINING THE N-METHYL EFFECT

4.4.1: ASP[18] TO ALA[18] MUTATION.

As observed with the 14mers, CD signal intensity appears to be dependent upon the pH of the media. To avoid ionization effects in the 4mers, Asp^{18} was replaced with an Ala residue. The CD data (data not shown) indicate that the two analogs display similar disordered curve shapes. The difference spectrum between the Ac-DIIW and Ac-AIIW analogs (data not shown) shows no significant bands that would indicate structural changes in the backbone. Any differences in the spectral intensity between the two analogs may reflect ionization at residue 18. Addition of HFIP (30% by volume) leads to a decrease in the poly$(Pro)_{II}$ signal for both analogs: the minimum (195nm) and maximum (225nm) decrease in intensity and have a slight red shift.

4.4.2: TRP[21] TO ALA[21] MUTATION.

The initial CD studies of the NMe-14mers yielded spectra with large negative signals between 225 and 230nm. Since disordered Trp sidechain signals are located in this region (albeit with the opposite sign), it was originally hypothesized that the N-methyl group induces a conformational change in the backbone which effected the Trp sidechain conformation. On a molar-residue basis, sidechains of aromatic residues have intense signals which may overwhelm those caused by the backbone chromophores. In order to completely rule out a Trp sidechain contribution, the 4-residue Ile^{20} and $^{(NMe)}Ile^{20}$ analogs were also synthesized with a $Trp^{21} \rightarrow Ala^{21}$ mutation, yielding Ac-AIIA and Ac-AI$^{(NMe)}$IA. The major disadvantage in studying the two Ala^{21} peptides is that there are no intense UV chromophores, such as a conjugated aromatic ring, that could be used to quantify the concentrations of the stock solutions. Although the curve shapes are accurate, the spectral intensities have a degree of uncertainty.

Qualitatively, there are dramatic differences between the spectra of Ac-AIIW and Ac-AIIA (Figure 4.8, Table 4.7) even though both peptides are essentially disordered. The most noticeable feature of the Trp analog is a positive band, with a maximum at ~ 225nm. This feature is most likely due to the disordered Trp indole ring. In the case of the Ala^{21} mutant, this region shows only a negative shoulder. The two $^{(NMe)}Ile^{20}$ analogs have similar features, specifically the large negative band centered at 225-230nm, and a "maximum" between 205 and 210nm. This would indicate that the backbone conformation represented by this CD signature is not dependent on residue 21. Any significant difference in the overall curve shape is solely the result of the disordered Trp residue. For instance, the minimum (225-230nm) of the Ala^{21} analog appears to be more intense than that of the Trp^{21} peptide. The maximum (205-210nm) is positive for Ac-AI$^{(NMe)}$IA, but negative for Ac-AI$^{(NMe)}$IW. Addition of fluoroalcohol (up to 30% HFIP by volume) to the samples appears to have the same effect observed in the 8mers: the

minimum at 225-230nm becomes more intense with a loss of disorder, as indicated by a decrease in the signal intensity at 195nm (which may reflect a loss of poly(Pro)$_{II}$ signals).

Table 4.7: CD optical constants and initial temperature gradients for Ac-XIIX in 5mM buffer, pH 5.5

Peptide	Mean Residue Ellipticity θ_{221nm}	θ_{min}	θ_{max}
Ac-AIIW			
0% HFIP	3400 – 126T	-42100 + 276T (195.0nm)	4940 – 112T (224.6nm)
30% HFIP	-1510 – 56T	-21000 + 123T (196.0nm)	573 – 25T (226.4nm)
Ac-DIIW			
0% HFIP	4820 – 106T	-40000 + 262T (195.0nm)	5820 – 66T (224.2nm)
30% HFIP	-162 – 53T	-27200 + 187T (196.0nm)	1380 – 28T (225.6nm)
Ac-AIIY			
0% HFIP	3950 – 147T	-56000 + 467T (190.0nm)	11800 – 179T (201.4nm)
30% HFIP	-248 – 149T	-35400 + 488T (188.0nm)	16600 – 113T (200.4nm)

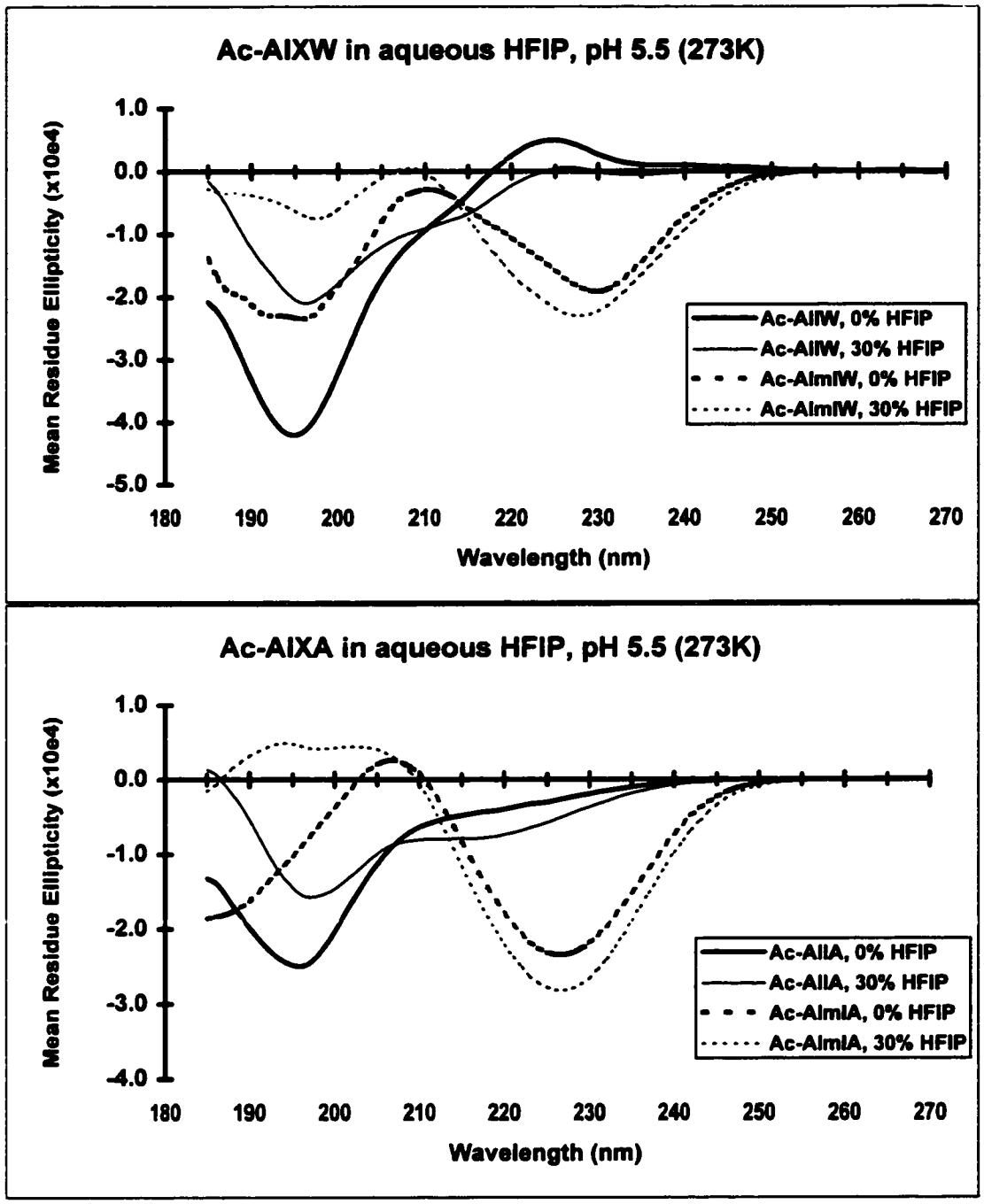


Figure 4.8: CD signatures of the 4mers in aqueous HFIP (pH 5.5) at 273K. Top panel: Ac-AIXW (^{NMe}Ile & Ile). Bottom panel: Ac-AIXA (^{NMe}Ile & Ile). Note that the intensities of the A^{21} peptide CD curves are estimated.

4.4.3: THE 3MERS.

Two three-residue peptides were designed to determine whether the large negative band between 225-230nm is found in all $^{(NMe)}$Ile analogs or is specific to an Ile-$^{(NMe)}$Ile moiety. Their sequences, Ac-GXK (where X = Ile or $^{(NMe)}$Ile) takes advantage of a Gly residue at position 19, which would allow the peptides to more freely sample conformational space. In addition, the peptides lack aromatic residues, which in their disordered state, would greatly affect the CD signals. As a result, any unusual CD signatures would be solely due to the N-methyl effect on the backbone chromophore. The CD stock solution concentrations were quantified (to *circa* ±10% error) by the TNBS assay (Satake et al., 1960; Fields, 1972; Sashidar et al., 1994).

Table 4.8: CD optical constants and initial temperature gradients for the Ac-GXK peptides in aqueous HFIP/5mM buffer (pH 5.5).

Peptide	Mean Residue Ellipticity θ_{221nm}	θ_{other}	θ_{min}
Ac-GIK			
0% HFIP	-1520 – 34T	-12300 + 151T (191.6nm)	-2470 – 41T (215.0nm)
30% HFIP	-2480 – 23T	-7930 + 58T (194.2nm)	-3370 – 18T (215.0nm)
Ac-G$^{(NMe)}$IK			
0% HFIP	-19000 + 16T	12900 + 32T (197.2nm)	-19800 + 9T **(223.0nm)**
30% HFIP	-19700 + 78T	4570 + 28T (192.0nm)	-21400 + 96T **(225.0nm)**

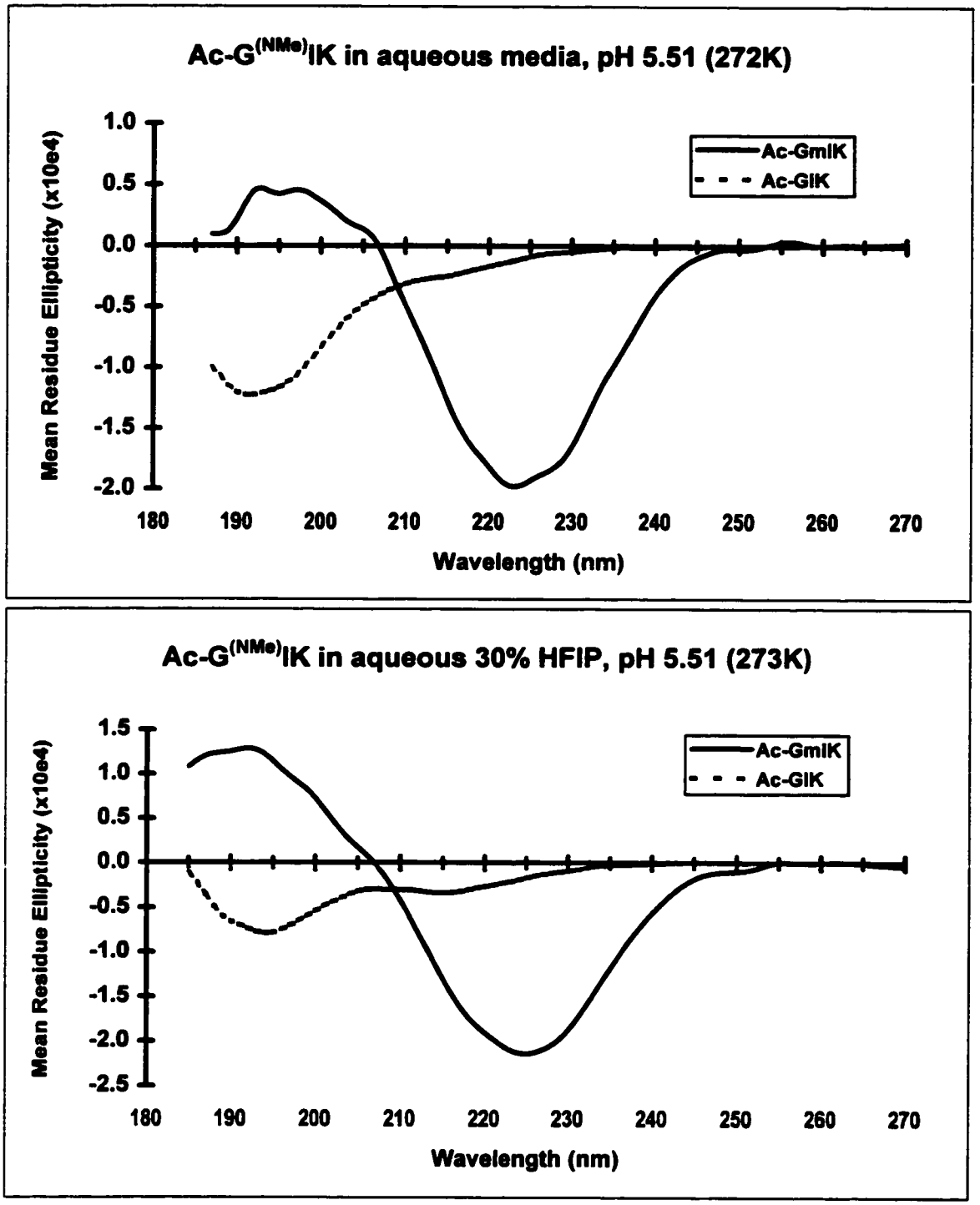


Figure 4.9: The CD spectra of the Ac-GXK peptides (^{NMe}Ile and Ile) in aqueous HFIP (pH 5.5) at 273K.

As expected, Ac-GIK displays a disordered spectrum (Figure 4.9, Table 4.8) with a minimum at 191.6nm which decreases upon addition of HFIP (to 30%). The θ_{191} decreases in intensity and red shifts to 194nm. Both the θ_{221} and the shoulder at 215nm increase intensity. The CD spectrum of the Ac-G$^{(NMe)}$IK peptide shows the same features as all the other $^{(NMe)}$Ile^{20} analogs, an intense minimum between 220-230nm.

4.4.4: ILE^{20} TO $^{(NME)}$ILE^{20} MUTATION.

Difference CD spectra comparing the $^{(NMe)}$Ile^{20} and Ile^{20} analogs were used to quantitate the N-methyl effect on the CD (Figures 4.10 and 4.11). Based on their general features (a positive lobe centered between 195-215nm and a negative lobe centered between 225-230nm), it is clear that the unusual CD signatures inherent in the $^{(NMe)}$Ile analogs are not greatly affected by the presence of a Trp residue at position 21. With the exception of a distinct shoulder at 207nm and slight differences in signal intensities, the CD difference spectra of the Trp^{21} and Ala^{21} 4mers were comparable. The shoulder present in the Trp^{21} analogs may represent incomplete subtraction of a Trp sidechain contribution. In the case of the $^{(NMe)}$Ile^{20} mutant, the stable hydrophobic cluster may impart some structuring to the Trp residue, while for the Ile^{20} peptide, the Trp should be completely disordered. However, the Ala^{21} mutants also contain a small bulge near 207nm. Differences in signal intensities are mostly due to errors in the calculation of the Ala^{21} sample concentrations. Increasing the length of the peptide to 8 residues also doesn't significantly effect the overall shape of the difference spectra (Figure 4.12). The difference spectra of the 8mers resemble that of the Ac-AIXW series. The top panel of Figure 4.13 displays the averaged CD spectrum for these three families. The lower panel of Figure 4.13 displays the averaged CD spectrum of the peptides in aqueous 30% HFIP. Table 4.9 lists the molar optical constants and temperature gradients for the major CD bands.

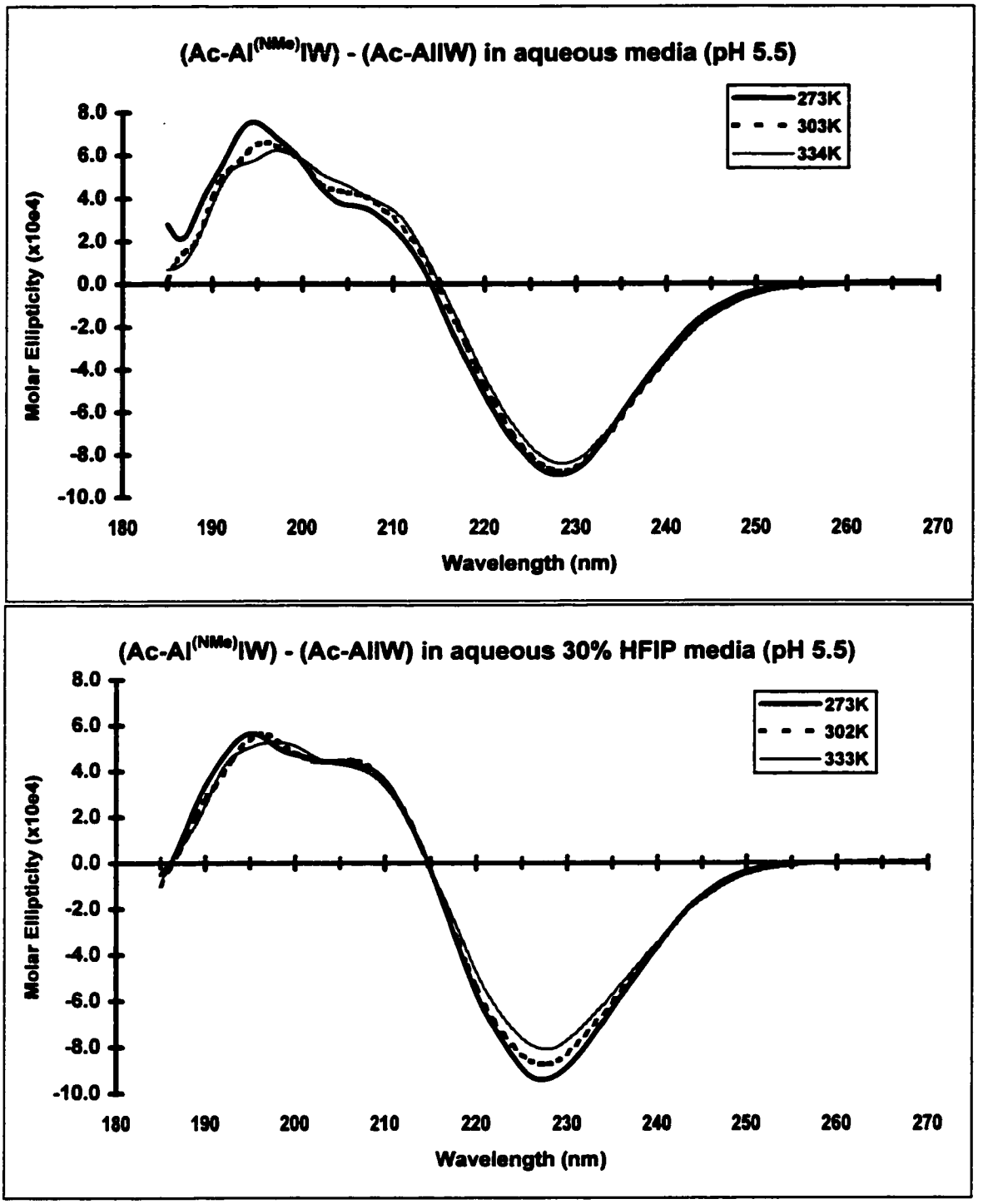


Figure 4.10: Difference CD spectra of the Ac-AIXW peptides in purely aqueous (top panel) and aqueous 30% HFIP (lower panel) media.

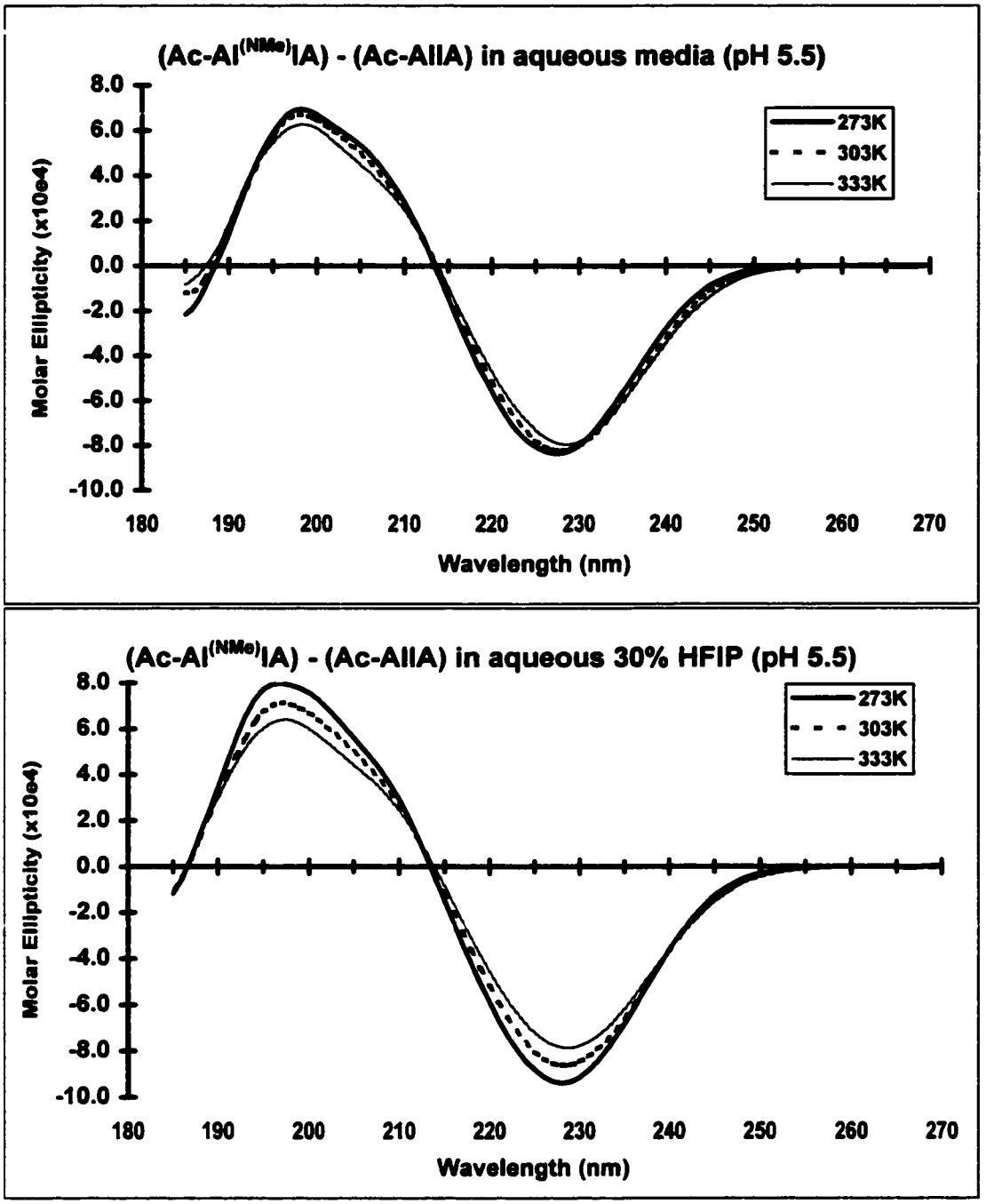


Figure 4.11: Difference CD spectra of the Ac-AIXA peptides in purely aqueous (top panel) and aqueous 30% HFIP (lower panel) media. Note that spectral intensities are estimated.

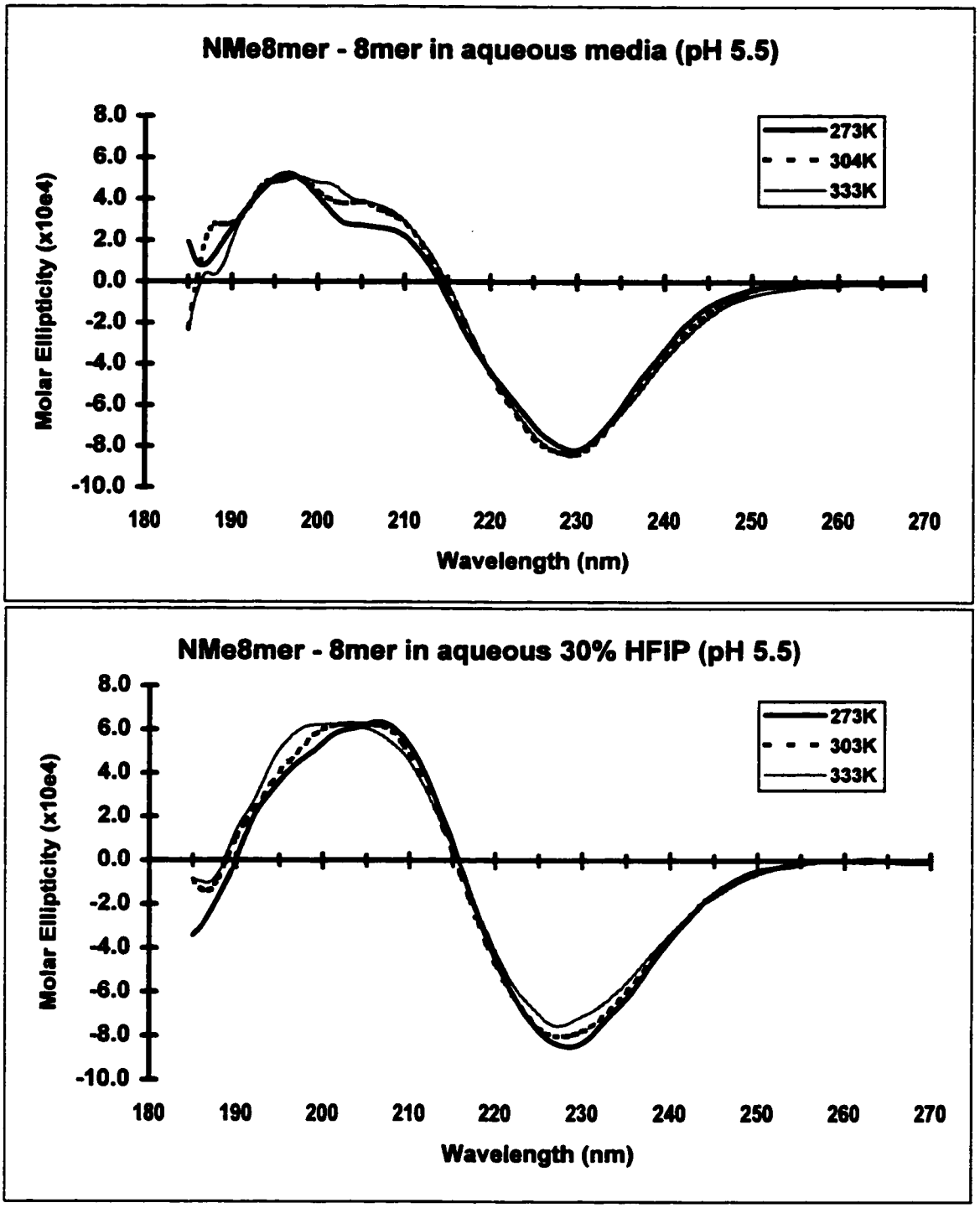


Figure 4.12: Difference CD spectra of the 8mers in purely aqueous (top panel) and aqueous 30% HFIP (lower panel) media.

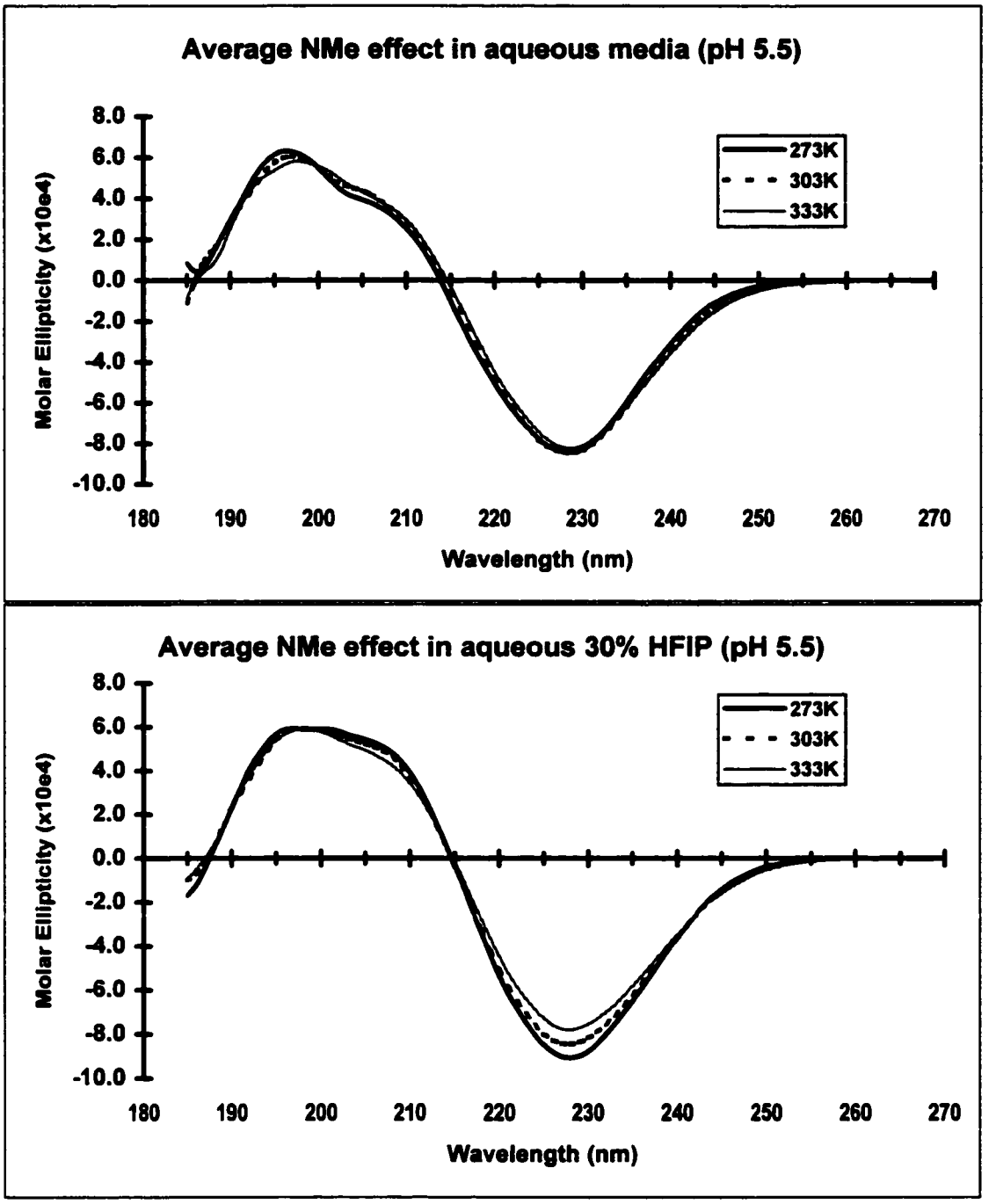


Figure 4.13: Average CD difference spectrum of the N-methyl backbone ($I^{NMe}I$) effect in purely aqueous media (top panel) and aqueous 30% HFIP (lower panel).

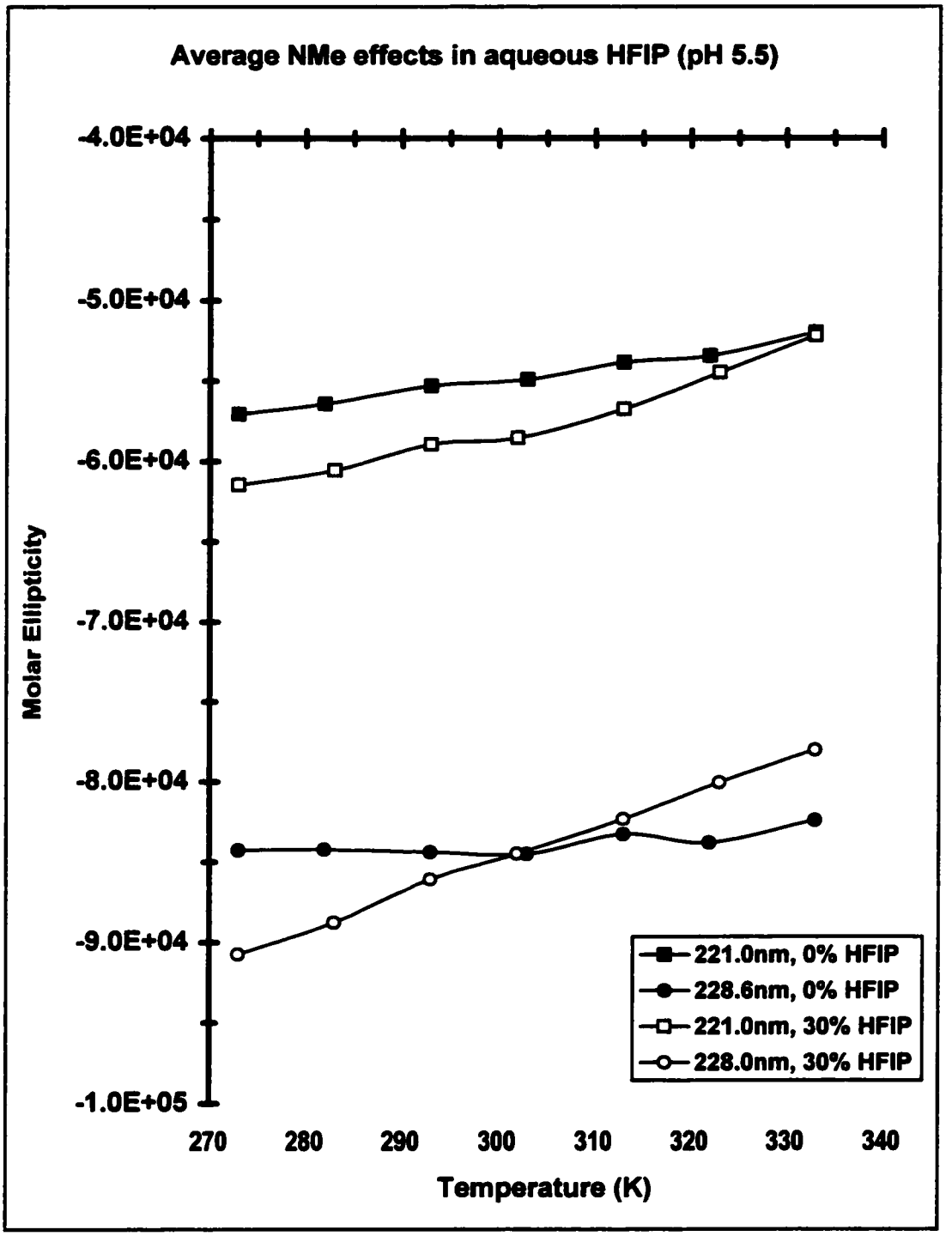


Figure 4.14: The temperature dependence of θ_{221} and θ_{min} for the N-methyl difference spectra.

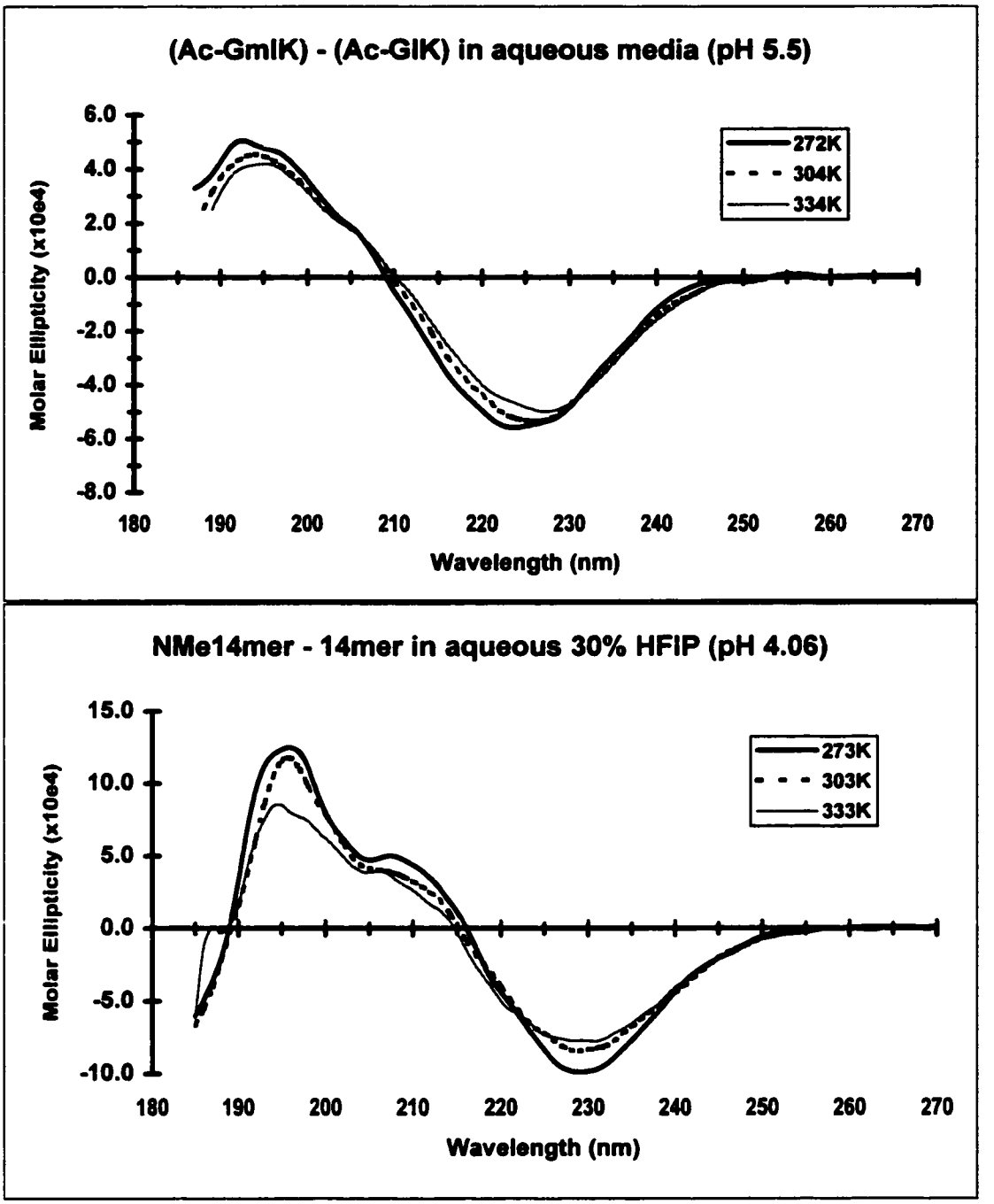


Figure 4.15: Top panel, difference CD spectra of the Ac-GXK peptides (^{NMe}Ile - Ile) in aqueous media (pH 5.5). Lower panel, difference CD spectra of the 14mers (NMe14mer – 14mer) in aqueous 30% HFIP (pH 4.06)

Table 4.9: CD optical constants for the average NMe-effect.

Peptide	Molar Residue Ellipticity θ_{221nm}	θ_{max}	θ_{min}	θ_{other}
$^{(NMe)}$I [1]				
0% HFIP	-52200 + 159T	+50700 – 198T (192.6nm)	-55800 + 123T (223.4nm)	
30% HFIP	-51900 + 201T	+61200 – 219T (193.0nm)	-59400 + 237T (226.0nm)	
I$^{(NMe)}$I [2]				
0% HFIP	-57100 + 81T	+63100 - 72T (196.4nm)	-84300 + 27T (228.6nm)	+35700 - 54T (207.0nm)
30% HFIP	-61500 + 151T	+59200 - 4T (197.0nm)	-90800 + 212T (228.0nm)	+50900 - 95T (207.0nm)
NMe14mer – 14mer				
30% HFIP	-51400 - 52T	+125000 – 720T (196.0nm)	-98800 + 359T (229.0nm)	+49900 - 212T (207.4nm)

[1] Data from the Ac-GXK difference spectra.
[2] Data acquired from the averaged difference spectra of the Ac-AIXX (W^{21} & A^{21}) and 8mer peptides.

Difference spectra of all the systems in aqueous 30% HFIP leads to overall increases in signal intensities, most notably at the minimum. However, the most dramatic fluoroalcohol effects are observed on the temperature gradients (Figure 4.14). In a purely aqueous medium (pH 5.5), the I$^{(NMe)}$I peptides exhibit very shallow thermal gradients (for example, at the minimum at 228.6nm, *circa* +27°/K, or 0.03%/K on a molar residue basis), suggesting that the structure or feature responsible for the CD signature is temperature invariant. However, in 30% HFIP, the minimum has a slight blue shift (to 228.0nm) and has a much larger temperature gradient (*circa* +212°/K, or 0.23%/K). This

would indicate that the "backbone conformation" induced by the N-methyl group thermally melts more readily in the presence of fluoroalcohols.

Examination of the Ac-GXK difference spectra (top panel of Figure 4.15 and Table 4.9), indicating that the $I^{(NMe)}I$ moiety is not required for the appearance of a large negative 225-230nm band. However, these difference spectra do lack the shoulder at 207nm that appears in the Ac-AIXW and, to a lesser extent, the Ac-AIXA series. The Ile-Ile moiety may experience more limited conformational sampling due to the bulky, β-branched sidechains. The intensities at the minimum appear to be smaller for $X^{(NMe)}I$ *vs.* $I^{(NMe)}I$. However, this could be due to an error in the concentration calculations using the TNBS assay (accuracy of results, ± 10-15%). The difference spectra of the 3mers in 30% HFIP (data not shown) display the same patterns shown by the 4mers and 8mers. Specifically, signal intensities increase and temperature gradients (at the minimum) become steeper in the presence of the fluorinated alcohol.

The 14mer difference spectra series (lower panel of Figure 4.15 and Table 4.9) is an interesting case: the maximum at ~196nm is far more intense than those of the shorter peptides by a factor of two. The maximum also experiences an extremely steep temperature gradient compared to the corresponding wavelengths of the shorter peptides in comparable conditions. This probably reflects the difference in helicity that occurs with N-methylation. Another factor could be an increase in the disorder of the Phe^{13} and Tyr^{14} sidechains (both have negative bands between 185 and 195nm).

The N-methylation effect difference spectra coincidentally resemble the CD signature of a type II β-turn (Brahms & Brahms, 1980; Harris, 1993). However, sidechain signatures of disordered aromatic residues cannot explain the discrepancies since the Ac-GXK peptides display the same patterns. Katahira et al. (1998) suggest that the C-termini of several 14 residue Ile^{20} analogs may adopt γ-turn conformations in DPPC vesicles. The authors make this assertion from the available NMR data: the unusual

upfield 19γ2 chemical shift and several NOEs between the Ile^{19} γ2 methyl group and the Trp^{21} sidechain protons. N-methylation at Ile^{20} could enhance such a conformational preference, and the CD spectrum may represent a stable γ-turn. The few examples of γ-turn CD studies that have been reported in the literature (for example: Lelj et al., 1992; Jois et al., 1992; Li et al., 1998) suggest that the γ-turns are red-shifted β_{II} turn signatures. However, all of these studies involved highly constrained cyclic peptides (disulfide bonds or cyclic backbones) and contained other CD signatures as well.

4.5: RE-DETERMINATION OF DISORDERED TRP, TYR AND PHE SIGNALS.

A re-determination of the aromatic residue (Trp, Tyr and Phe) sidechain signals was necessary for this study since the current reference set only contained two spectral sets [Brahms and Brahms (1980), which consisted of uncapped 2-residue peptides, and Harris (1993), which made questionable assumptions in the calculation of the difference spectra]. As previously mentioned, sidechains of disordered aromatic residues exhibit large CD signals. On a molar scale, these signals are far more intense than those of the backbone amides. Better defined reference spectra would therefore be useful in removing signals representing disordered aromatic residues, as well as for calculating more accurate $\langle f_H \rangle$ values.

The disordered aromatic residue signatures were primarily determined using the 3- and 4mers in aqueous media with various levels of added fluorinated alcohols. These were calculated using several methods. The first method involves a direct calculation of difference spectra. For example, subtraction of Ac-AIIA from Ac-AIIW would yield the signature for a disordered Trp sidechain (i.e. a Trp → Ala substitution), assuming that the sample concentration of the Ala^{21} mutant was accurately determined. Likewise, subtraction of Ac-AIIA from Ac-AIIY would yield a Tyr sidechain spectrum. Since the concentrations of the Ac-AIIA samples were estimated, these difference spectra would

have some uncertainty. A second method is related to the first. Since the peptides are of different lengths, the spectra are initially converted to units of molar residue ellipticity prior to subtraction. This was primarily used for the GHKX peptides. In this case, GHK was subtracted from GHKW and GHKF to yield one molar unit (backbone + sidechain) of Trp and Phe.

The third method yields more accurate results. In this case, one unit of a standardized "random coil" spectrum, which is corrected for temperature, is subtracted from those of the 3mers. For example, KYK acquired at 273K was corrected with one unit of coil, resulting in the one complete Tyr residue (backbone + sidechain). A fourth method incorporates the latter two. Here, the standardized Tyr spectrum is subtracted from Ac-AIIY to yield the spectrum of Ac-AII, which, in turn, is subtracted from Ac-AIIW. The difference spectrum represents a disordered Trp sidechain. In order to limit uncertainty in the reference spectra, the various Tyr, Trp and Phe spectra, including those reported by Brahms and Brahms (1980), are averaged. The CD difference curves with higher confidence have higher weights in the calculation. Note that all resulting CD signatures are in units of molar residue ellipticity.

The resulting averages representing both 0% and 30% HFIP are displayed in Figure 4.16 and Table 4.10. The Trp (with and without fluorinated alcohols) is the set with the highest confidence. Uncertainties for the Tyr and Phe spectra are higher due to the limited population sampled. The error may also be larger for the Phe since CD sample concentrations for GHKF and GHK were determined using the TNBS assay, which would form TNP conjugates to both the Lys ε-NH_2 and the N-terminal amine. High fluoroalcohol levels don't greatly effect the Trp spectrum: the extrema have the same relative intensities and experience a slight blue shift. Both the Tyr and Phe spectra appear to be more influenced by addition of HFIP or TFE.

Table 4.10: CD optical constants and initial temperature gradients for average disordered aromatic residues in aqueous HFIP/5mM buffer (pH 5.5) at 273K

	Molar Residue Ellipticity			
Residue	θ_{221nm}	θ_{max1}	θ_{max2}	θ_{min}
Trp				
0% HFIP	+24400 - 204T	+31000 - 173T (225.0nm)		-50900 + 179T (198.6nm)
30% HFIP	+22100 – 127T	+27800 – 121T (224.6nm)		-50700 + 182T (197.6nm)
Tyr				
0% HFIP	+21200 - 177T	+77300 - 328T (200.8nm)	+26900 - 180T (226.0nm)	-65700 + 308T (188.0nm)
30% HFIP	+13500 – 117T	+75500 – 294T (199.8nm)	+17700 – 108T (225.2nm)	-41800 + 356T (187.4nm)
Phe				
0% HFIP [1]	+14200 - 509T	+66600 - 188T (196.6nm)	+24500 - 123T (~215.0nm)	-93900 (~185.0nm)
30% HFIP [2]	+15300 – 944T	+82000 – 292T (196.6nm)	+34300 – 231T (~213.0nm)	unknown

[1] Second Phe maximum is a shoulder, minimum is an estimate.
[2] Data for Phe in 30% HFIP less well-determined.

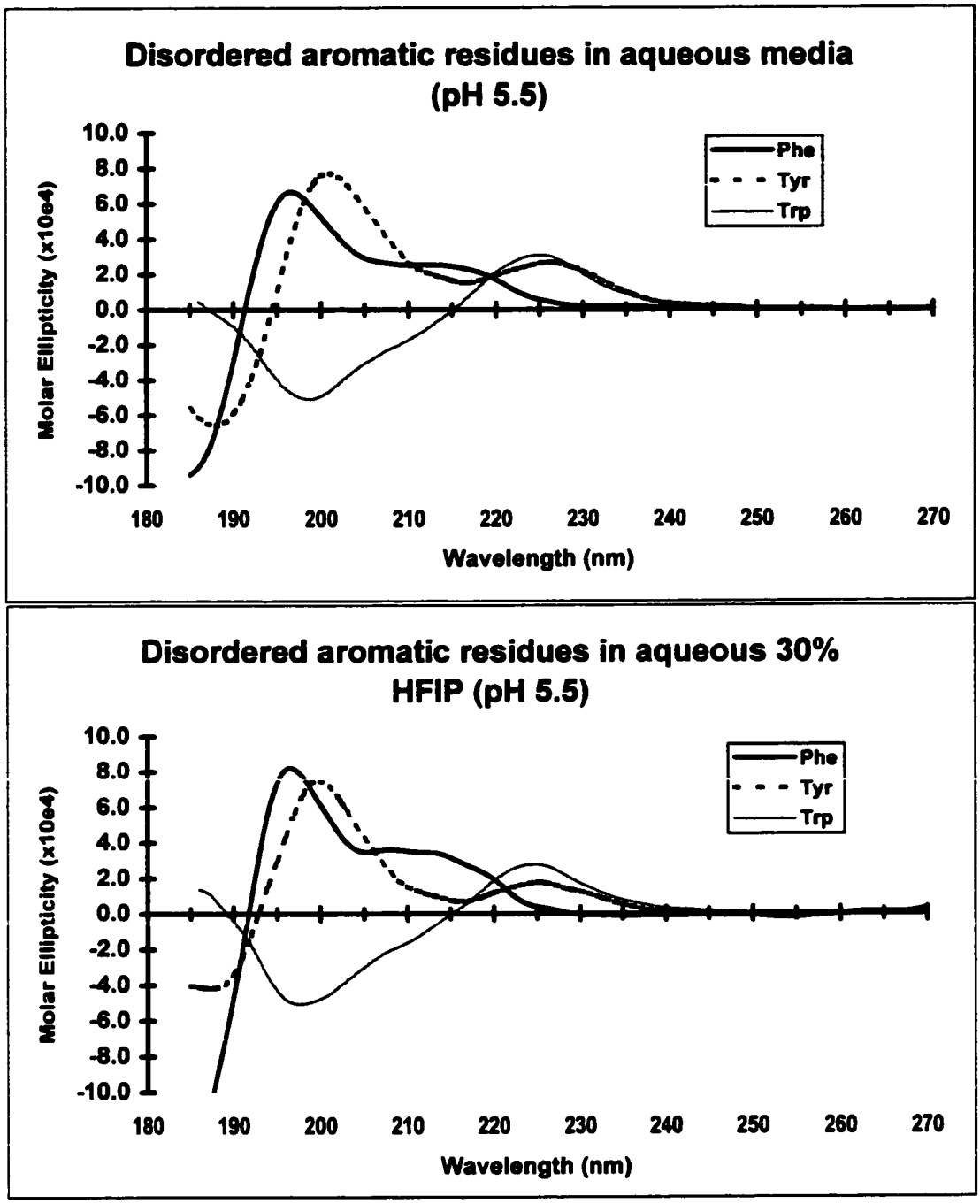


Figure 4.16: Average CD signatures of disordered aromatic residue sidechains in purely aqueous media (top panel) and aqueous 30% HFIP (lower panel) at pH 5.5.

CHAPTER 5: CONCLUSIONS AND OTHER INSIGHTS

5.1: REASSESSMENT OF PEN-1 AND PEN-2 STRUCTURES

Several conclusions regarding the solution-state structures of ET-1, Pen-1 and Pen-2 can be ascertained from these studies. To recap: ET-1 displays a relatively extended N-terminus ($Cys^1 \rightarrow Cys^3$), a loop ($Ser^5 \rightarrow Asp^8$) and a helix ($Lys^9 \rightarrow Cys^{15}$), which is capped by Asp^8. The six residue C-terminal region ($His^{16} \rightarrow Trp^{21}$) is conformationally averaged. However, the Ile^{19} β-methyl group of both ET-1 and Pen-1 exhibits an upfield chemical shift deviation in aqueous media, suggesting the presence of weak hydrophobic interactions involving the sidechains of Ile^{19} and Trp^{21}. In the case of Pen-1, which is a potent ET_A receptor *agonist*, the Cys → Pen substitutions at residues 3 and 15 added four methyl groups to the hydrophobic interior of the bicyclic core, allowing it to become more rigid. The Pen-1 solution-state structure may, therefore, be a better representative of the ET-1 bound state than the native peptide. The increased rigidity of the core also extends and increases the structural stability of the helical region.

The Pen-2 analog, which differs from the Pen-1 mutant only by an $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ substitution, is a potent ET_A receptor *antagonist*. The bicyclic core of Pen-2 retains the same increased rigidity present in the Pen-1 analog. However, the N-methyl group appears to restrict helix propagation into the C-terminal region. The Pen-2 mutant displays a dramatic upfield Ile^{19} β-methyl chemical shift deviation, as well as stronger NOE interactions between residues 19 and 21, suggesting the presence of a hydrophobic cluster. However, addition of ethylene glycol and/or HFIP diminishes the 19γ2 upfield chemical shift deviation, suggesting that the C-terminal hydrophobic interactions aren't

as strong in less "lipophobic" media. No significant NOE interactions between the bicyclic core and the C-terminus were observed for any of the endothelins examined.

5.2: EFFECTS OF THE N-METHYL AND AROMATIC RESIDUES

5.2.1: CONCLUSIONS FROM NMR DATA.

A previous study (Chen, 1992) did not account for the N-methyl effect on chemical shifts. As a result, the early Pen-2 CSD calculations included significant errors for the α-methines of Ile^{19} and $^{(NMe)}Ile^{20}$. Both protons have large downfield chemical shift deviations relative to reference Ile values, resulting in large positive CSDs if left uncorrected. A series of short 3 and 4 residue peptides (Ac-GXK, Ac-XIIX, Ac-$XI^{(NMe)}IX$) were used to determine the overall N-methyl correction factors for the αH of the preceding (reference +0.29ppm, for aqueous media) and own residue (reference +0.58ppm). These correction factors may only be applicable to N-methylated isoleucines. General tendencies for the amide of residue *i-1* and the amide and α-methine of residue *i+1* have also been observed, although correction factors for these groups have not been applied in this study. Table 5.1 summarizes the correction factors, separated into two sections. The first (A) reports the values for the Ile^{19}, Trp/Ala^{21} variant. The second (B) lists the deviations observed for the Gly^{19}, Lys^{21} analog.

The αH-CSD values observed for the $^{(NMe)}Ile$ in either DMSO or aqueous media were greater (more downfield) than +0.500ppm, suggesting that the residues *i-1* and *i+1* doesn't significantly influence the α-methine chemical shift of the N-methylated residue. The amide and α-methine of the preceding residue display more dramatic chemical shift deviations in both media. Analogs where residue i-1 is an isoleucine display large upfield amide CSDs in DMSO (-0.231ppm), but a negligible value in aqueous media. The α-methines have large downfield shifts: +0.246ppm in DMSO, +0.194ppm in aqueous

media. Smaller shift deviations with respect to reference values are observed for the case where residue *i-1* is a glycine. This reflects differences in backbone conformational sampling between the $G^{(NMe)}I$ and the $I^{(NMe)}I$ moieties. Due to steric factors inherent in β-branched aliphatic residues, the $I^{(NMe)}I$ moiety is expected to have more conformational restrictions in its backbone.

Table 5.1: N-methyl effects on neighboring NH and αH chemical shifts

	residue *i-1*		NMe-Ile		residue *i+1*	
	A	B	A	B	A	B
DMSO						
NH	-0.231 ±0.080	-0.115	——	——	+0.147 ±0.015	+0.150
αH	+0.246 ±0.030	+0.194 ±0.068	+0.546 ±0.020	+0.508	-0.153 ±0.021	-0.194
aqueous						
NH	+0.006 ±0.044	-0.132	——	——	+0.093 ±0.092	+0.301
αH	+0.383 ±0.112	+0.147 ±0.000	+0.589 ±0.038	+0.551	+0.007 ±0.042	-0.039

A: residue *i-1* = Ile, residue *i+1* = Trp or Ala
B: residue *i-1* = Gly, residue *i+1* = Lys

The unusual Ile^{19} β-methyl (γ2) group chemical shift has also been examined in detail. The 3 and 4-residue peptides indicated that a hydrophobic interaction with the Trp^{21} indole ring causes the upfield 19γ2 chemical shift. The $Trp^{21} \rightarrow Ala^{21}$ substitution resulted in an upfield 19γ2 shift change of -0.337ppm between the Ile^{20} analogs. This effect is enhanced upon N-methylation of Ile^{20}: the same difference becomes more negative (-0.592ppm). A comparison of the $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ substitution also demonstrates the same general trends: the Trp^{21} analogs display upfield shift differences more negative than -0.250ppm for the 19γ2 peak while the Ala^{21} analog had a negligible

value. Molecular models of the Pen-1 analog indicated that the Ile^{19} γ2 group sits within the shielding cone of the Trp^{21} indole ring. Although the NOESY and ROESY data for this peptide displayed limited contacts between the sidechains of residues 19 and 21, groups in the Ile^{20} analogs, the Pen-2 analogs show stronger 19 → 21 interactions. This suggests that the N-methyl group produces a backbone conformational preference which enhances the hydrophobic interaction the 19γ2 group and the aromatic sidechain of residue 21. In addition, a decrease in the stability of the cluster, as measured by a downfield chemical shift deviation of the 19γ2 peak, with decreasing solvent "lipophobicity" is also observed for the short C-terminal fragments.

An examination of the 19γ2 (Δδ/ΔT)/NH-CSD correlations reveal several trends. First, the Ile^{20} → $^{(NMe)}Ile^{20}$ substitution does not significantly alter the ease of hydrophobic cluster unfolding. Second, the more polar (or "lipophobic") the media, the more likely a thermally stable hydrophobic cluster forms. Third, in the case of the C-terminus of the larger fragments, incorporation of a helical region disrupts the thermal stability of the cluster. However, a rigid bicyclic core, such as those observed for Pen-1 and Pen-2, enhances the cluster stability.

Finally, the NMR data indicates that the N-methyl group has similar effects on a peptide bond (between residues *i* and *i-1*) as a proline. A strong 19αH → 20NMe NOE (and the lack of a significant 19αH → 20αH NOESY crosspeak) indicates that the major conformer of the 19/20 peptide bond is the *trans* configuration in both aqueous and DMSO media. In the case of Ac-G$^{(NMe)}$IK, the ratio was approximately 85% trans:15% cis for both media. The larger fragments and Pen-2, exhibit the same features. This conclusion disagrees with that reported by Cody et al. (1997), who note that a six residue ET_A receptor antagonist, which includes an $^{(NMe)}$Ile at residue 20, contains a 19/20 peptide bond that isomerizes from 100% trans in aqueous media to 100% cis in DMSO. However, this peptide also contained an unnatural aromatic amino acid (D-Bhg), which

may also promote the formation of a hydrophobic cluster, allowing the 19/20 peptide bond to favor the cis conformer.

5.2.2: CONCLUSIONS FROM CD DATA.

The CD data of the longer fragments, specifically the 14mers, suggest the presence of a helix. However, both the NMe-14mer and its N-acetyl capped analog display larger than expected negative bands between 225 and 230nm. The large negative 225-230nm band isn't solely a feature of the NMe-14mers. Difference spectra calculated from the 8mers and 4mers all contain the same feature and resemble a type II β-turn: positive bands between 195 and 210nm and negative bands between 225 and 230nm, both of nearly equal intensity. The CD signatures are slightly affected by residue 21. The difference spectra of the Trp^{21} analogs display a maximum at ~195nm and a pronounced shoulder at *circa* 207nm. Although the Ala^{21} analog CD spectra were calculated using estimated concentrations, the overall N-methyl effect remains the same. However differences in the intensities of the maximum at 195nm and the shoulder at 207nm is less dramatic. Any differences between the two series are likely due to a disordered Trp signal, which has negative bands at 195 and 207nm. The N-methyl effect on the 3mers, whose spectra were more accurately quantified using the TNBS assay, also demonstrate that this signature is not attributed solely to the $I^{(NMe)}I$ moiety. However, this signature lacks the shoulder at 207nm, indicating that the CD signatures of the $Ile^{(NMe)}Ile$ and $Xxx^{(NMe)}Ile$ moieties are different. This is not surprising since the backbone of two sequential β-branched aliphatic residues (isoleucines) is expected to have more limited conformational sampling than that of a single Ile residue. The signatures generated by the three-residue peptides also suggest the formation of a γ-turn. However, most published reports of γ-turn CD signatures (Lelj et al., 1992; Jois et al., 1992; Li et al., 1998) also include other secondary structures.

5.3: OTHER SIGNALING FACTORS

5.3.1: HELICAL REGION.

NMR data indicate that the 14mers exhibit a surprisingly large helicity for a 7 residue span ($Ala^9 \rightarrow Ala^{15}$) in aqueous media. This is among the shortest stable helices observed in a peptide. This indicates that the two disulfide bridges present in the intact endothelins are not required for helix formation, but rather help stabilize it. HFIP addition and N-capping of the disulfide-free peptide enhances the helix stability, although not to the level of the intact 21mers. The NMe-14mer displays several $i \rightarrow i+3$ NOEs, even in the absence of fluoroalcohol. The AcNMe-14mer displays a more extensive web of strong $i \rightarrow i+3$ NOEs. Deuterium exchange data indicate that the $Ala^{11} \rightarrow Ala^{15}$ span has exchange protected NHs, with Val^{12} the most protected backbone amide. The α-methine chemical shift deviation histograms for all three peptides also display values consistent with a helix in *both* aqueous and DMSO media (although in the case of DMSO, these values were very small).

Fractional helicities calculated from the difference spectra in which the CD signatures of the 4mers (Ac-XIIX and Ac-$XI^{(NMe)}IX$) were subtracted from the 14mers indicated that N-methylation of residue 20 reduces the helicity, while capping the N-terminus with an acetyl group enhances helix stability. In 30% HFIP (pH 4.06, 273K), the fractional helicities rank as follows: AcNMe-14mer (57.38%) > 14mer (54.21%) > NMe-14mer (47.73%). In general, the helix is more stable the more acidic the media, although more dramatic differences in $\langle f_H \rangle$ are observed for the NMe-14mer *vs.* the AcNMe-14mer. This would suggest that the ionization state of Asp^8 (the capping unit of the helical region in both the intact and fragment peptides) and acetyl N-capping influences helix stability, although the latter is more important.

5.3.2: Implications of Structural Changes to Signaling.

Endothelin-1 and its two intact bicyclic analogs, Pen-1 and Pen-2, present an interesting case where slight changes in the amino acid sequence lead to significant differences in both structure and biological activity. Both mutants have the same overall structural characteristics and have the same relative binding affinities to the ET_A receptor (Andersen et al., 1995a; Hunt, unpublished data) as the native peptide. In both Chen's (1992) and this study, the bicyclic core of Pen-1 was determined to be more rigid than that of ET-1. This would indicate steric factors induced by the Pen residue β-methyl groups have a direct effect on the hydrophobic interior of the bicyclic core and would partially explain the increased binding affinity and biological activity. The residues in the helical region of the peptide which are known to be important for receptor binding and signal transduction (Asp^8, Glu^{10} and Phe^{14}) may form better contacts to the binding site when this bicyclic core becomes more rigid.

N-methylation of Ile^{20} transforms the potent ET_A receptor agonist into a potent ET_A receptor antagonist. NMR data suggests that the $Ile^{20} \rightarrow {}^{(NMe)}Ile^{20}$ substitution enhances the hydrophobic clustering observed in ET-1 and Pen-1. The resulting backbone conformational preferences produces a turn-like CD signature, which appears to be unique for the N-methyl effect. Since residue 21 is known to be an important factor in ET_A receptor binding and biological activity (Kimura et al., 1988), the spectral data suggests that the position of the Trp indole ring is vital to signaling. In both ET-1 and Pen-1, the backbone of the C-terminal residues exhibits more conformational averaging than Pen-2. An increase in the rigidity of the C-terminal region, which limits the positional space of the Trp^{21} sidechain, could explain the switch to antagonism observed for the ${}^{(NMe)}Ile^{20}$ analog.

5.4: NEW DISORDERED AROMATIC RESIDUE CD SIGNATURES

Since fractional helicity calculations are based on θ_{221} ellipticity values, the newly determined disordered aromatic residue CD signatures (which, on a molar scale, have large positive contributions at 221nm) are essential for accurate helicity calculations when aryl amino acids are present. Other studies (Brahms & Brahms, 1980; Harris, 1993) had also obtained these CD signatures, but had not determined thermal and fluoroalcohol effects. In addition, the earlier calculations performed in this group (Harris, 1993) had used the peptides GHKX (X = Trp, Phe) and GHK. At the time, concentrations of the 3mer were estimated, so intensities of the difference spectra had high uncertainties. Later calculations (this study) used a series of short peptides (for example: WA, KXK, GHKX, Ac-AIIX, Ac-AI$^{(NMe)}$IX) which were quantified by sidechain aromatic and/or TNP-Lys chromophores. The new values for disordered Phe, Tyr and Trp residues are reported in Table 5.2.

Table 5.2: Molar ellipticity values of the 221nm band and extrema for disordered aromatic residue sidechains [1]

Residue		θ_{221nm}	θ_{max}	θ_{min}
Trp	0% HFIP	+24000	+31000 (225.0nm)	-50900 (198.6nm)
	30% HFIP	+22100	+27800 (224.6nm)	-50700 (197.6nm)
Tyr	0% HFIP	+21200	+77300 (200.8nm)	-65700 (188.0nm)
	30% HFIP	+13500	+75500 (199.8nm)	-41800 (187.4nm)
Phe [2]	0% HFIP	+14200	+66600 (196.6nm)	-93900 (185.0nm)
	30% HFIP	+15300	+82000 (196.6nm)	unknown

[1] See also, Table 4.10 for initial temperature gradients.
[2] Data for Phe in 30% HFIP is less well-determined.

These values should be of general validity, particularly when the aromatic amino acids are not centrally located in a helical span.

REFERENCES

Alder, A.J., Greenfield, N.J. and Fasman, G.D. *Meth. Enzymol.* **1973**, *27*, 675-735.

Andersen, N.H., Chen, C., Marschner, T.M., Krystek, S.R., Jr. and Bassolino, D.A. *Biochemistry.* **1992a**, *31*, 1280-1295.

Andersen, N.H., Cao, B. and Chen, C. *Biochem. Biophys. Res. Commun.* **1992b**, *184*, 1008-1014.

Andersen, N.H., Harris, S.M., Lee, V.G., Liu, E.C.-K., Moreland, S. and Hunt, J.T. *Bioorg. Med. Chem.* **1995a**, *3*(2), 113-124.

Andersen, N.H., Chen, C., and Lee, G.M. *Prot. Pept. Lett.* **1995b**, *1*(4), 215-222.

Andersen, N.H., Cort, J.R., Harris, S.M., Lee, G.M., Liu, Z., Neidigh, J.W., Tomaszewski, J.W. and Tong, H. *36th Experimental NMR Conference (Boston, MA).* **1995c**, poster #134.

Andersen, N.H., Cort, J.R., Liu, Z., Sjoberg, S.J. and Tong, H. *J. Am. Chem. Soc.* **1996**, *118*, 10309-10310.

Andersen, N.H., Neidigh, J., Harris, S.M., Lee, G.M., Liu, Z. and Tong, H. *J. Am. Chem. Soc.* **1997**, *119*, 8547-8561.

Andersen N.H. and Tong, H. *Prot. Sci.* **1997**, *6*, 1920-1936.

Arai, H., Hori, S., Aramori, I., Ohkubo, H. and Nakanishi, S. *Nature* (London). **1990**, *348*, 730-732.

Aue, W.P., Bartholdi, E. and Ernst, R.R. *J. Chem. Phys.* **1976**, *64*, 2229-2246.

Aumelas, A., Chiche, L., Mahe, E., Le-Nguyen, D., Sizum, P., Berthault, P. and Perly, B. *Int. J. Pept. Prot. Res.* **1991**, *37*, 315-324.

Aumelas, A., Chiche, L., Kubo, S., Chino, N., Tamaoki, H. and Kobayashi, Y. *Biochemistry.* **1995**, *34*, 4546-4561.

Aumelas, A. Kubo, S., Chino, N., Chiche, L., Forest, E., Roumestand, C. and Kobayashi, Y. *Biochemistry.* **1998**, *37*, 5220-5230.

Bai, Y., Milne, J.S., Mayne, L. and Englander, S.W. *Proteins· Struct. Funct. Gen.* **1993**, *17*, 75-86.

Barlow, D.J. and Thornton, J.M. *J. Mol. Biol.* **1988**, *201*, 601-609.

Bax, A. and Davis, D.G. *J. Magn. Reson.* **1985a**, *63*, 207-213.

Bax, A. and Davis, D.G. *J. Magn. Reson.* **1985b**, *65*, 355-360.

Bennes, R., Calas, B., Chabrier, P.-E., Demaille, J. and Heitz, F. *FEBS Lett.* **1990**, *276*, 21-24.

Bigaud, M. and D'Orchymont, H. *Endothelins in Biology and Medicine*, Huggins, J.P. and Pelton, J.T., Eds., CRC Press, New York, **1997**, 103-147.

Birnbaumer, L., Abramowitz, J. and Brown, A.M. *Biochim. Biophys. Acta.* **1990**, *1031*, 163-224.

Blanco, F., Ramiréz-Alvarado, M. and Serrano, L. *Curr. Opin. Struct. Biol.* **1998**, *8*, 107-111.

Bodenhausen, G., Kogler, H. and Ernst, R.R. *J. Magn. Reson.* **1984**, *58*, 370-388.

Bothner-By, A.A., Stephens, R.L., Lee, J.-M., Warren, C.D. and Jeanloz, R.W. *J. Am. Chem. Soc.* **1984**, *106*, 811-813.

Bodkin, M.J. and Goodfellow, J.M. *Biopolymers.* **1996**, *39*, 43-50.

Bradley, E.K., Thomason, J.F., Cohen, P.A. and Kuntz, I.D. *J. Mol. Biol.* **1990**, *215*, 607-622.

Brahms, S. and Brahms, J. *J. Mol. Biol.* **1980**, *138*, 149-178.

Braunschweiler, L. and Ernst, R.R. *J. Magn. Reson.* **1983**, *53*, 521-523.

Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. and Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187-217.

Bruch, M.D., Dhingra, M.M. and Giesrasch, L.M. *Prot. Struct. Funct. Gen.* **1991**, *10*, 130-139.

Brünger, A.T. *XPLOR: A System for X-Ray Crystallography and NMR*, Yale University Press. **1988**.

Cao, B. Ph.D. Dissertation, University of Washington, **1993**.

Calas, B., Harricane, M.-C., Gulmard, L., Heitz, F., Mendre, C., Chabrier, P.E. and Bennes, R. *Peptide Res.* **1992**, *5*, 97-101.

Cammers-Goodwin, A., Allen, T.J., Oslick, S.L., McClure, K.F., Lee, J.H. and Kemp, D.S. *J. Am. Chem. Soc.* **1996**, *118*, 3082-3090.

Carpino, L.A. *J. Am. Chem. Soc.* **1993**, *115*(10), 4397-4398.

Chakrabarty, A., Kortemme, T., Padmanabhan, S. and Baldwin, R.L. *Biochemistry.* **1993**, *32*, 5560-5565.

Chen, C. Ph.D. Dissertation, University of Washington, **1992**.

Chen, Y.-H., Yang, J.T. and Chau, K.H. *Biochemistry.* **1974**, *13*, 3350-3359.

Cody, W.L., He, J.X., Reily, M.D., Haleen, S.J., Walker, D.M., Reyner, E.L., Stewart, B.H. and Doherty, A.M. *J. Med. Chem.* **1997**, *40*, 2228-2240.

Cody, W.L. and Doherty, A.M. *Biopolymers.* **1995**, *37*, 89-104.

Coles, M., Munro, S.L.A. and Craik, D.J. *J. Med. Chem.* **1994**, *37*, 656-664.

Cort, J., Liu, Z., Lee, G., Harris, S.M., Prickett, K.S., Gaeta, L.S.L. and Andersen, N.H., *Biochem. Biophys. Res. Commun.* **1994**, *204*(3), 1088-1095.

Cort, J. Ph.D. Dissertation, University of Washington, **1997**.

Coste, J., Frérot, E. and Jouin, P. *J. Org. Chem.* **1994**, *59*, 2437-2446.

Cowley, D.J. and Pelton, J.T. *Int. J. Pept. Prot. Res.* **1995**, *46*, 56-64.

Creighton, T.E. *Proteins: Structures and Molecular Properties*, W.H. Freeman & Co., New York, **1984**.

Creighton, T.E. *J. Mol. Biol.* **1975**, *95*, 167-199.

Dalgarno, D.C., Slater, L., Chackalamannil, S., and Senior, M.M. *Int. J. Pept. Prot. Res.* **1992**, *40*, 515-523.

deAlba, E., Angeles Jiménez, M., Rico, M. and Nieto, J.L. *Folding and Design.* **1996**, *1*, 133-144.

Doherty, A.M. *J. Med. Chem.* **1992**, *35*(9), 1493-1508.

Doherty, A.M., Cody, W.L., DePue, P.L., He, J.X., Waite, L.A., Leonard, D.M., Leitz, N.L., Dudley, D.T., Rapundalo, S.T., Hingorani, G.P., Haleen, S.J., LaDouceur, D.M., Hill, K.E., Flynn, M.A. and Reynolds, E.E. *J. Med. Chem.* **1993**, *36*(18), 2585-2594.

Driscoll, P.C., Clore, G.M., Beress, L. and Gronenborn, A.M. *Biochemistry.* **1989**, *28*, 2178-2187.

Endo, S., Inooka, H., Ishibashi, Y., Kitada, C., Mizuta, E. and Fujino, M. *FEBS Lett.* **1989**, *257*, 149-154.

Goto, K., Hama, H. and Kasuya, Y. *Jpn. J. Pharmacol.* **1996**, *72*, 261-290.

Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. *J. Am. Chem. Soc.* **1988**, *110*, 7870-7872.

Ferro, C.J. and Webb, D.J. *Drugs.* **1997**, *53*(Supplement 1), 30-41.

Fields, R. *Meth. Enzymol.* **1972**, *XXV*, 464-469.

Fiori, W.R. and Millhauser, G.L. *Biopolymers (Pept. Sci.).* **1995**, *37*, 243-250.

Freeman, C.M., Catlow, C.R.A., Hemmings, A.M. and Hidler, R.C. *FEBS Lett.* **1986**, *197*, 289-296.

Furchgott, R.F. and Zawadski, J.V. *Nature* **1980**, *288*(5789), 373-376.

Harris, S.M. Ph.D. Dissertation, University of Washington, **1993**.

Hickey, K.A., Rubanyi, G., Paul, R.J. and Highsmith, R.F. *Am. J. Physiol.* **1985**, *248*(5 part 1), C550-C556.

Hirota, N., Mizuno, K. and Goto, Y. *Protein Sci.* **1997**, *6*, 416-421.

Hunt, J.T., Lee, V.G., Liu, E.C.K., Moreland, S., McMullen, D., Webb, M.L. and Bolgar, M. *Int. J. Peptide Protein Res.* **1993**, *8*, 1-10.

Itoh, Y., Yanagisawa, M., Ohkubo, S., Kimura, C., Kosada, T., Inoue, A., Ishida, N., Mitsui, Y., Onda, H., Fujino, M. and Masaki, Y. *FEBS Lett.* **1988**, *231*, 440-444.

Janes, R.W., Peapus, D.H. and Wallace, B.A. *Nature Struct. Biol.* **1994**, *1*, 311-319.

Jasanoff, A. and Fersht, A.R. *Biochemistry.* **1994**, *33*, 2129-2135.

Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. *J. Chem. Phys.* **1979**, *71*, 4546-4553.

Jois, D.S.S., Easwaran, K.R.K., Bednarek, M. and Blout, E.R. *Biopolymers.* **1992**, *32*, 993-1001.

Karne, S., Jayawickreme, C.K. and Lerner, M.R. *J. Biol. Chem.* **1993**, *268*, 19126-19133.

Katahira, R. Umemura, I., Takai, M., Oda, K., Okada, T. and Nosaka, A.Y. *J. Peptide Res.* **1998**, *51*, 155-164.

Kelly, S.M. and Price, N.C. *Biochim. Biophys. Acta.* **1997**, *1338*, 161-165.

Kimura, S., Kasuya, Y., Sawamura, T., Shinmi, O., Sugita, Y., Yanagisawa, M., Goto, K. and Masaki, T. *Biochem. Biophys. Res. Commun.* **1988**, *156*(3), 1182-1186.

Kochva, E., Bdolah, A. and Wollberg, Z. *Toxicon.* **1993**, *31*(5), 541-568.

Koradi, R., Billeter, M. and Wüthrich, K. *J. Mol. Graphics.* **1996**, *14*, 51-55.

Kortemme, T., Ramiréz-Alvarado, M. and Serrano, L. *Science.* **1998**, *281*, 253-256.

Kuprin, S., Gräslund, A., Ehrenberg, A. and Koch, M.H.J. *Biochem. Biophys. Res. Commun.* **1995**, *217*, 1151-1156.

Krystek, S.R., Jr., Bassolino, D.A., Novotny, J., Chen, C., Marschner, T.M. and Andersen, N.H. *FEBS Lett.* **1991**, *281*, 212-218.

Kuroda, M., Yamazaki, K., Taga, T. *FEBS Lett.* **1994**, *355*, 263-266.

Kuwaki, T., Kurihara, H., Cao, W.H., Kurihara, Y., Unewaka, M., Yazaki, Y. and Kumada, M. *Prog. Neurobio.* **1997**, *51*, 545-579.

Lee, G.M., Chen, C., Marschner, T.M. and Andersen, N.H. *FEBS Lett.* **1994**, *355*, 140-146.

Lelj, F., Tamburro, A.M., Villani, V., Grimaldi, P. and Guantieri V. *Biopolymers.* **1992**, *32*, 159-170.

Levitt, M.H., Freeman, R. and Frenkiel, T.A. *J. Magn. Res.* **1982**, *47*(2), 328-330.

Li, N., Kendrick, B.S., Manning, M.C., Carpenter, J.F. and Duman, J.G. *Arch. Biochem. Biophys.* **1998**, *360*(1), 25-32.

Llinás M. and Klein, M.P. *J. Am. Chem. Soc.* **1975**, *97*, 4731-4737.

Masaki, T. and Yanagisawa, M. *Essays in Biochemistry*. **1992**, *27*, 79-89.

Meruka, G., Dyson, H.J. and Wright, P.E. *J. Biomol. NMR*. **1995**, *5*, 14-24.

Millhauser, G.L. *Biochemistry*. **1995**, *34*(12), 3873-3877.

Millhauser, G.L., Stenland, C.J., Hanson, P., Bolin, K.A. and van de Ven, F.J.M. *J. Mol. Biol*. **1997**, *267*, 963-974.

Molday, R.S., Englander, S.W. and Kallen, R.G. *Biochemistry*. **1972**, *11*(2), 150-158.

Muñoz, V. Serrano, L., Angeles Jiménez, M. and Rico, M. *J. Mol. Biol*. **1995**, *247*, 648-669.

Munro, S., Craik, D., McConville, C., Hall, J., Searle, M., Bicknell, W., Scanlon, D. and Chandler, W. *FEBS Lett*. **1991**, *278*, 9-13.

Nakajima, K., Kubo, S., Kumagaye, S.-i., Nishio, H., Tsunemi, M., Inui, T., Kuroda, H., Chino, N., Watanabe, T.X., Kimura, T. and Sakakibara, S. *Biochem. Biophys, Res. Commun*. **1989**, *163*(1), 424-429.

Némethy, G. and Printz, M.P. *Macromolecules*. **1972**, *5*(6), 755-758.

Olsen, U.B., Weis, J.U., Suzdak, P., Lundt, B.F., Madsen, K., Thøgersen, H. and Johansen, N.L. *Pharm. Toxicol*. **1993**, *73*, 60-62.

Palmer, R.M., Ashton, D.S. and Moncada, S. *Nature*. **1988**, *333*(6174), 664-666.

Panek, R.L., Major, T.C., Taylor, D.G., Hingorani, G.P., Dunbar, J.B., Doherty, A.M. and Rapundalo, S.T. *Biochem. Biophys. Res. Commun*. **1992**, *183*, 572-576.

Piotto, M., Saudek, V. and Sklenar, V. *J. Biomol. NMR*. **1992**, *2*, 661-666.

Privalov, P.L. *Protein Folding*, Creighton, T.E., editor, W.H. Freeman & Co., New York, **1992**, 83-126.

Ramachandran, G.N. and Sasisekharan, V. *Adv. Prot. Chem*. **1968**, *23*, 283-437.

Ramiréz-Alvarado, M., Blanco, F.J. and Serrano, L. *Nat. Struct. Biol.* **1996**, *3*(7), 604-612.

Ramiréz-Alvarado, M., Blanco, F.J., Niemann, H. and Serrano, L. *J. Mol. Biol.* **1997**, *273*, 898-912.

Reily, M.D. and Dunbar, J.B., Jr. *Biochem. Biophys. Res. Commun.* **1991**, *178*, 570-577.

Rohl, C.A. and Baldwin, R.L. *Biochemistry.* **1994**, *33*, 7760-7767.

Rubanyi, G.M. and Polokoff, M.A. *Pharm. Rev.* **1994**, *46*(3), 325-415.

Sakurai, T., Yanagisawa, M., Takuwa, Y., Miyazaki, H., Kimura, S., Goto, K. and Masaki, T. *Nature* (London). **1990**, *348*, 732-735.

Sashidar, R.B., Capoor, A.K. and Ramana, D. *J. Immun. Meth.* **1994**, *167*, 121-127.

Satake, K., Okuyama, T., Ohashi, M. and Shinoda, T. *J. Biochem.* **1960**, *47*(5), 656-660.

Saudek, V., Hoflack, J. and Pelton, J.T. *FEBS Lett.* **1989**, *257*, 145-148.

Scheek, R.M., van Gunsteren, W.F. and Kapstein, R. *Methods in Enzymology.* **1989**, *177*, 204-218.

Shaka, A.J., Keeler, J., Frenkiel, T.A. and Freeman, R. *J. Magn. Res.* **1983**, *52*(2), 335-338.

Sibanda, B.L., Blundell, T.L. and Thornton, J.M. *J. Mol. Biol.* **1989**, *206*, 759-777.

Sibanda, B.L. and Thornton, J.M. *Meth. Enzymol.* **1991**, *202*, 59-82.

Simmonson, M.S. and Dunn, M.J. *FASEB J.* **1990**, *4*, 2989-3000.

Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. *J. Magn. Reson. A.* **1993**, *102*, 241-245.

Sokolovsky, M. *J. Neurochem.* **1992**, *59*(3), 809-821.

Spencer, J.R., Antonenko, V.V., Delaet, N.G.J. and Goodman, M. *Int. J. Pept. Prot. Res.* **1992**, *40*, 282-293.

Spokes, R.A., Ghatei, M.A. and Bloom, S.R. *J. Cardiovasc. Pharmacol.* **1989**, *13*(Supp. 5), S191-S192.

Storrs, R.W., Truckses, D. and Wemmer, D.E. *Biopolymers.* **1992**, *32*, 1695-1702.

Takayanagi, R., Kitazumi, K., Takasaki, C., Ohnaka, K., Aimoto, S., Tasaka, K., Ohashi, M. and Nawata, H. *FEBS Lett.* **1991**, *282*, 103-106.

Tam, J.P., Liu, W., Zhang, J.-W., Galantino, M., Bertolero, F., Cristiani, C., Vaghi, F. and deCastiglione, R. *Peptides.* **1994**, *15*(4), 703-708.

Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogoku, Y., Nakajima, K., Kumagaye, S., Kimura, T. and Sakakibara, S. *Prot. Eng.* **1991**, *4*, 509-418.

Thornton, J.M. *Protein Folding*, Creighton, T.E., editor, W.H. Freeman & Co., New York, **1992**, 59-82.

Wallace, B.A. and Janes, R.W. *J. Cardiovasc. Pharm.* **1995**, *26*(Supp. 3) S250-S253.

Walgers, R., Lee, T.C. and Cammers-Goodwin, A. *J. Am. Chem. Soc.* **1998**, *120*, 5073-5079.

Wang, S.-S. *J. Am. Chem. Soc.* **1973**, *95*, 1328-1333.

Webb, M.L. and Meek, T.D. *Med. Res. Rev.* **1997**, *17*(1), 17-67.

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. *J. Biomol. NMR.* **1995**, *5*, 67-81.

Woody, R.W. *Adv. Biophys. Chem.* **1992**, *2*, 37-49.

Woody, R.W. *Methods in Enzymology.* **1995**, *246*, 34-71.

Wüthrich, K. *NMR of Proteins and Nucleic Acids* John Wiley & Sons, New York. **1986.**

Yanagisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K. and Masaki, T. *Nature.* **1988**, *332*, 411-415.

Yanagisawa, M. and Masaki, T. *Trends Pharmacol. Sci.* **1989**, *10*, 374-378.

APPENDIX A: CSD REFERENCE VALUES (298K)

residue	Hα Reference Values aqueous media	DMSO	HN Reference Values aqueous media	DMSO
A (Ala)	4.32	4.32	8.21	8.04
C $(Cys)_2$	4.69	4.64	8.34	8.29
C (Cys^{SH})	4.57		8.23	
D^- (Asp^-)	4.65	4.63	8.29	8.21
E^- (Glu^-)	4.32	4.34	8.37	8.05
F (Phe)	4.61	4.56	8.25	8.14
G (Gly)	3.96	3.76	8.28	8.15
H^+ (His^+)	4.70	4.66	8.35	8.31
I (Ile)	4.15	4.21	8.08	8.26
K (Lys)	4.32	4.29	8.19	8.06
L (Leu)	4.32	4.37	8.17	7.95
M (Met)	4.48	4.41	8.25	8.07
Nle	4.32	4.30	8.22	8.10
N (Asn) / D° (Asp°)	4.72	4.61	8.40	8.14
P (Pro)	4.43	4.36		
Pen $(Pen)_2$	4.65	4.70	8.26	
Q (Gln) / E° (Glu°)	4.34	4.32	8.27	8.06
R (Arg)	4.34	4.38	8.24	8.07
S (Ser)	4.47	4.34	8.26	7.98
T (Thr)	4.37	4.22	8.15	7.83
V (Val)	4.10	4.18	8.08	7.88
W (Trp)	4.65	4.58	8.05	8.09
Y (Tyr)	4.56	4.46	8.13	8.04

Corrections to reference values (from the Andersen lab):

1) Hα effects:

a) NH_3^+ terminal effect: aqueous media, pH ≤ 6.0-6.5, - 0.26ppm
DMSO, ~ -0.46ppm

b) C-terminal effects: CO_2H, $CONH^-$, no correction
CO_2^-, -0.17ppm

c) Xaa-Pro effect: Xaa = Gly, + 0.17ppm; Xaa = others, +0.29ppm

d) Media effects (volume %):

i) Fluoroalcohol effects:
4-6% HFIP (10-20% TFE), + 0.015ppm
10% HFIP (25-30% TFE), + 0.035ppm
12.5-16% HFIP (35-50% TFE), + 0.05ppm
20% HFIP, + 0.065ppm
25-35% HFIP, + 0.075ppm

ii) Acetonitrile effects:
< 15%, no correction
20-45%, -0.03ppm
50-80%, -0.07ppm

2) HN reference values in aqueous media:

a) Standard correction to locate structuring effects: correct reference values to temperatures of experimental values by a gradient of -7.6ppb/°C

b) NH_3^+ terminal effects:
i) "two effect": to second residue, + 0.48ppm
ii) "three effect": to third residue, + 0.20ppm

c) CO_2^- terminal effects: to C-terminal residue, -0.43ppm

3) Current αH reference value for $^{(NMe)}$Ile used in the CSD calculations: 4.75ppm (DMSO media) and 4.73ppm (aqueous media).

APPENDIX B: MISCELLANEOUS NMR TABLES – AQUEOUS MEDIA

Table B.1: HLDIIW in 10% ethylene glycol, 90% acidic water at 290K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
His 16	exchanged	4.304	3.389, 3.389	δ 7.399, ε 8.697
Leu 17	8.739 [-7.453]	4.410	1.588, ?	γ 1.574, δ,δ′ 0.917, 0.880
Asp 18	8.720 [-6.263]	bleached	2.856, 2.728	
Ile 19	8.054 [-8.830]	4.053	1.677	γ1,γ1′ 1.358, 1.035
				γ2 0.557, δ 0.766
Ile 20	8.066 [-8.204]	4.160	1.770	γ1,γ1′ 1.386, 1.093
				γ2 0.817, δ 0.803
Trp 21	8.097 [-9.624]	4.716	3.357, 3.212	δ 7.233, ε1 10.102, ε3 7.661
				ζ2 7.473, ζ3 7.148, η2 7.221

Table B.2: HLDIIW in 30% ethylene glycol, 10% acetic acid, 40% water at 290K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
His 16	exchanged	4.333	3.435, 3.373	δ 7.432, ε 8.743
Leu 17	8.746 [-5.270]	4.419	1.605, 1.543	γ 1.606, δ,δ′ 0.920, 0.886
Asp 18	8.759 [-6.539]	4.806	2.909, 2.776	
Ile 19	7.955 [-7.034]	4.092	1.683	γ1,γ1′ 1.359, 1.023
				γ2 0.576, δ 0.754
Ile 20	8.022 [-6.953]	4.185	1.767	γ1,γ1′ 1.392, 1.087
				γ2 0.819, δ 0.796
Trp 21	8.099 [-8.288]	4.715	3.341, 3.217	δ 7.222, ε1 10.094, ε3 7.604
				ζ2 7.432, ζ3 7.103, η2 7.179

Table B.3: VYFAHLDIIW in 45% ethylene glycol, 8% acetic acid, 47% water at 290K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Val 12	exchanged	3.716	2.051	γ,γ′ 0.869, 0.840
Tyr 13	8.502 [-4.457]	4.585	2.869	δ,δ′ 7.032, ε,ε′ 6.743
Phe 14	8.127 [-6.821]	4.545	3.008, 2.819	δ,δ′ 7.159, ε,ε′ 7.245, ζ 7.212?
Ala 15	8.184 [-5.843]	4.169	1.285	
His 16	8.367 [-4.829]	4.625	3.250, 3.158	δ 7.290, ε 8.648
Leu 17	8.277 [-4.714]	4.335	1.609, 1.410	γ 1.534, δ,δ′ 0.882, 0.840
Asp 18	8.534 [-4.736]	4.718	2.891, 2.742	
Ile 19	7.897 [-4.543]	4.092	1.681	γ1,γ1′ 1.339, 0.999
				γ2 0.563, δ 0.744
Ile 20	8.031 [-5.529]	4.186	1.747	γ1,γ1′ 1.380, 1.072
				γ2 0.808, δ 0.786
Trp 21	8.146 [-4.886]	4.679	3.318, 3.320	δ 7.192, ε1 10.138, ε3 7.573
				ζ2 7.405, ζ3 7.071, η2 7.155

Table B.4: GSHLDIIW in aqueous 40% ethylene glycol, 5% acetic acid at 300K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.868, 3.868?		
Ser 15	8.590 [-4.498]	4.471	3.815, 3.815	
His 16	8.626 [-5.953]	4.720	3.274, 3.162	δ 7.310, ε 8.641
Leu 17	8.294 [-5.234]	4.321	1.576, 1.401	γ 1.546, δ,δ′ 0.905, 0.857
Asp 18	8.478 [-5.602]	4.712	2.883, 2.744	
Ile 19	7.841 [-5.267]	4.077	1.691	γ1,γ1′ 1.342, 1.012
				γ2 0.586, δ 0.752
Ile 20	7.962 [-6.522]	4.165	1.760	γ1,γ1′ 1.377, 1.080
				γ2 0.806, δ 0.794
Trp 21	8.065 [-7.475]	4.684	3.330, 3.214	δ 7.198, ε1 10.124, ε3 7.593
				ζ2 7.425, ζ3 7.091, η2 7.169

Table B.5: GSHLDIIW in aqueous 15% HFIP, 34% ethylene glycol, 2.5% acetic acid at 285K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.938, 3.938?		
Ser 15	8.653 [-4.933]	4.552	3.891, 3.821?	
His 16	8.637 [-5.987]	4.788	3.346, 3.218	δ 7.358, ε 8.673
Leu 17	8.328 [-5.585]	4.385	1.640, 1.640?	γ 1.609, δ,δ′ 0.953, 0.910
Asp 18	8.466 [-5.870]	4.798	2.924, 2.802	
Ile 19	7.835 [-6.254]	4.142	1.772	γ1,γ1′ 1.417, 1.084
				γ2 0.686, δ 0.819
Ile 20	7.826 [-4.632]	4.224	1.821	γ1,γ1′ 1.432, 1.122
				γ2 0.858, δ 0.841
Trp 21	7.865 [-6.120]	4.757	3.377, 3.276	δ 7.277, ε1 9.986, ε3 7.629
				ζ2 7.456, ζ3 7.127, η2 7.199

Table B.6: GSHLDI$^{(NMe)}$IW in aqueous 5% acetic acid at 285K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.887, 3.887?		
Ser 15	8.727 [-5.996]	4.482	3.832, 3.832?	
His 16	8.813 [-7.334]	4.753	3.290, 3.184	δ 7.300, ε 8.616
Leu 17	8.427 [-6.378]	4.343	1.597, 1.597?	γ 1.546, δ,δ′ 0.918, 0.863
Asp 18	8.624 [-6.826]	4.705	2.853, 2.735	
Ile 19	8.139 [-9.436]	4.442	1.599	γ1,γ1′ 1.227, 1.027
				γ2 0.261, δ 0.792
Ile 20	n/a	4.732	2.000	γ1,γ1′ 1.411, 1.049
				γ2 0.883, δ 0.711, NMe 2.982
Trp 21	8.352 [-10.362]	4.605	3.397, 3.171	δ 7.247, ε1 10.173, ε3 7.673
				ζ2 7.478, ζ3 7.147, η2 7.230

Table B.7: GSHLDI$^{(NMe)}$IW in aqueous 30% ethylene glycol, 3.5% acetic acid at 285K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.876, 3.876?		
Ser 15	8.703 [-5.435]	4.466	3.830, 3.809	
His 16	8.777 [-7.135]	4.730	3.284, 3.172	δ 7.315, ε 8.650
Leu 17	8.406 [-5.998]	4.317	1.588, 1.588?	γ 1.526, δ,δ′ 0.909. 0.858
Asp 18	8.598 [-6.386]	4.691	2.840, 2.713	
Ile 19	8.064 [-8.094]	4.420	1.580	γ1,γ1′ 1.402, 1.002
				γ2 0.231, δ 0.746
Ile 20	n/a	4.720	2.003	γ1,γ1′ 1.218, 0.919
				γ2 0.876, δ 0.779, NMe 2.976
Trp 21	8.243 [-8.875]	4.576	3.365, 3.148	δ 7.223, ε1 10.203, ε3 7.630
				ζ2 7.443, ζ3 7.114, η2 7.195

Table B.8: GSHLDI$^{(NMe)}$IW in aqueous 15% HFIP, 25.5% ethylene glycol, 3% acetic acid at 285K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.948, 3.948?		
Ser 15	8.741 [-5.909]	4.544	3.888, 3.888?	
His 16	8.698 [-6.504]	4.808	3.372, 3.207	δ 7.361, ε 8.684
Leu 17	8.397 [-6.167]	4.388	1.643, 1.643?	γ 1.601, δ,δ′ 0.951, 0.908
Asp 18	8.560 [-6.860]	4.828	2.886, 2.769	
Ile 19	7.882 [-6.146]	4.487	1.583	γ1,γ1′ 1.425, 1.025
				γ2 0.288, δ 0.775
Ile 20	n/a	4.698	2.063	γ1,γ1′ 1.287, 0.966
				γ2 0.909, δ 0.828 NMe 3.014
Trp 21	7.805 [-4.913]	4.621	3.352, 3.227	δ 7.248, ε1 9.951, ε3 7.630
				ζ2 7.450, ζ3 7.129, η2 7.211

Table B.9: Ac-GSHLDI$^{(NMe)}$IW in aqueous 5% acetic acid at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	8.409 [-7.418]	3.950, 3.950?		**AcMe**
Ser 15	8.392 [-7.099]	4.428	3.832, 3.811	
His 16	8.706 [-6.944]	4.724	3.306, 3.180	δ 7.281, ε 8.596
Leu 17	8.291 [-5.877]	4.322	1.605, 1.605?	γ 1.533, δ,δ′ 0.915, 0.879
Asp 18	8.579 [-6.570]	4.697	2.867, 2.746	
Ile 19	8.113 [-9.102]	4.444	1.596	γ1,γ1′ 1.411, 1.029
				γ2 0.255, δ 0.765
Ile 20	n/a	4.731	1.992	γ1,γ1′ 1.231, 0.925
				γ2 0.860, δ 0.785, NMe 2.982
Trp 21	8.360 [-10.268]	4.594	3.394, 3.164	δ 7.246, ε1 10.175, ε3 7.668
				ζ2 7.474, ζ3 7.144, η2 7.226

Table B.10: Ac-GSHLDI$^{(NMe)}$IW in aqueous 30% ethylene glycol, 3.5% acetic acid at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	8.394 [-6.751]	3.937, 3.937?		AcMe 2.058
Ser 15	8.364 [-6.814]	4.413	3.822, 3.797	
His 16	8.664 [-6.697]	4.699	3.359, 3.144	δ 7.294, ε 8.626
Leu 17	8.251 [-5.286]	4.314	1.604, 1.604?	γ 1.531, δ,δ′ 0.912, 0.853
Asp 18	8.539 [-6.052]	4.684	2.855, 2.730	
Ile 19	8.026 [-7.828]	4.427	1.579	γ1,γ1′ 1.398, 1.000
				γ2 0.228, δ 0.743
Ile 20	n/a	4.716	1.994	γ1,γ1′ 1.223, 0.919
				γ2 0.868, δ 0.779, NMe 2.973
Trp 21	8.232 [-8.722]	4.565	3.302, 3.178	δ 7.222, ε1 10.197, ε3 7.113
				ζ2 7.442, ζ3 7.627, η2 7.194

Table B.11: Ac-GSHLDI$^{(NMe)}$IW in aqueous 15% HFIP, 25.5% ethylene glycol, 3% acetic acid at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	8.314 [-6.454]	3.984, 3.984?		AcMe 2.109
Ser 15	8.352 [-6.703]	4.474	3.902, 3.862	
His 16	8.577 [-5.863]	4.753	3.231, 3.356	δ 7.336, ε 8.650
Leu 17	8.211 [-4.822]	4.388	1.666, 1.666?	γ 1.592, δ,δ′ 0.955, 0.904
Asp 18	8.467 [-6.036]	4.800	2.901, 2.782	
Ile 19	7.847 [-5.690]	4.483	1.580	γ1,γ1′ 1.427, 1.014
				γ2 0.275, δ 0.771
Ile 20	n/a	4.697	2.053	γ1,γ1′ 1.295, 0.964
				γ2 0.916, δ 0.826, NMe 3.011
Trp 21	7.792 [-4.668]	4.616	3.373, 3.201	δ 7.248, ε1 9.936, ε3 7.629
				ζ2 7.444, ζ3 7.127, η2 7.207

Table B.12: EAVYFAHLD in aqueous 30% ethylene glycol, 3 equivalents TFA at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	4.054	2.131, 2.131	γ,γ′ 2.542, 2.542
Ala 11	8.766 [-5.982]	4.361	1.314	
Val 12	8.158 [-7.645]	3.999	1.901	γ,γ′ 0.867, 0.749
Tyr 13	8.206 [-7.905]	4.512	2.885, 2.885	δ,δ′ 6.992, ε,ε′ 6.730
Phe 14	8.008 [-5.686]	4.499	3.054, 2.885	δ,δ′ 7.182, ε,ε′ 7.296, ζ 7.253
Ala 15	8.189 [-7.179]	4.157	1.292	
His 16	8.375 [-6.079]	4.645	3.256, 3.160	δ 7.282, ε 8.611
Leu 17	8.355 [-8.471]	4.361	1.596, 1.596	γ 1.596, δ,δ′ 0.908, 0.853
Asp 18	8.532 [-7.210]	4.718	2.936, 2.936	

Table B.13: EAVYFAHLD in aqueous 25% HFIP, 17.5% ethylene glycol, 3 equivalents TFA at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	4.143	2.215, 2.215	γ,γ′ 2.606, 2.606
Ala 11	8.746 [-6.497]	4.460	1.388	
Val 12	8.008 [-7.425]	4.060	2.005	γ,γ′ 0.926, 0.826
Tyr 13	7.841 [-5.323]	4.501	2.914, 2.914	δ,δ′ 6.994, ε,ε′ 6.801
Phe 14	7.588 [-2.964]	4.542	3.160, 2.996	δ,δ′ 7.215, ε,ε′ 7.356, ζ 7.303
Ala 15	7.887 [-6.066]	4.226	1.361	
His 16	8.129 [-3.641]	4.720	3.352, 3.201	δ 7.303, ε 8.581
Leu 17	8.088 [-3.671]	4.433	1.710, 1.628	γ 1.655, δ,δ′ 0.963, 0.906
Asp 18	8.258 [-5.168]	4.788	2.969, 2.969	

Table B.14: EAVYFAHLDIIW in aqueous 20% HFIP, 32% acetic acid at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	4.180	2.198, 2.198	γ,γ′ 2.598, 2.598
Ala 11	8.670 [-5.962]	4.483	1.368	
Val 12	8.055 [-7.368]	4.127	1.944	γ,γ′ 0.894, 0.777
Tyr 13	8.058 [-7.410]	4.624	2.921, 2.860	δ,δ′ 6.984, ε,ε′ 6.764
Phe 14	7.894 [-6.436]	4.638	3.112, 2.931	δ,δ′ 7.201, ε,ε′ 7.291, ζ 7.244
Ala 15	8.055 [-7.368]	4.276	1.359	
His 16	8.258 [-5.540]	4.757	3.337, 3.224	δ 7.308, ε 8.624
Leu 17	8.164 [-5.387]	4.429	1.676, 1.560	γ 1.592, δ,δ′ 0.921, 0.880
Asp 18	8.452 [-6.140]	4.833	2.950, 2.822	
Ile 19	7.943 [-6.185]	4.159	1.753	γ1,γ1′ 1.756, 1.067
				γ2 0.624, δ 0.804
Ile 20	7.902 [-6.968]	4.280	1.814	γ1,γ1′ 1.817, 1.112
				γ2 0.860, δ 0.820
Trp 21	7.917 [-6.643]	4.834	3.386, 3.283	δ 7.198, ε1 9.731, ε3 7.605
				ζ2 7.416, ζ3 7.108, η2 7.181

Table B.15: EAVYFAHLDIIW in aqueous 40% HFIP, 24% acetic acid at 285K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	4.176	2.205, 2.191	γ,γ′ 2.603, 2.603
Ala 11	8.587 [-5.355]	4.499	1.380	
Val 12	7.951 [-6.804]	4.110	1.961	γ,γ′ 0.904, 0.795
Tyr 13	7.904 [-6.548]	4.589	2.896, 2.896	δ,δ′ 6.977, ε,ε′ 6.777
Phe 14	7.711 [-5.144]	4.630	3.127, 2.952	δ,δ′ 7.189, ε,ε′ 7.307, ζ 7.254
Ala 15	7.890 [-6.148]	4.259	1.364	
His 16	8.113 [-4.338]	4.740	3.337, 3.226	δ 7.276, ε 8.534
Leu 17	8.040 [-4.214]	4.414	1.691, 1.590	γ 1.604, δ,δ′ 0.939, 0.887
Asp 18	8.345 [-5.309]	4.864	2.938, 2.831	
Ile 19	7.862 [-5.234]	4.163	1.783	γ1,γ1′ 1.438, 1.081
				γ2 0.679, δ 0.832
Ile 20	7.753 [-5.978]	4.259	1.824	γ1,γ1′ 1.438, 1.122
				γ2 0.863, δ 0.832
Trp 21	7.725 [-6.150]	4.837	3.393, 3.310	δ 7.185, ε1 9.514, ε3 7.618
				ζ2 7.424, ζ3 7.132, η2 7.203

APPENDIX C: MISCELLANEOUS NMR TABLES – DMSO MEDIA

Table C.1: DAEAVYFAHLDIIW in DMSO, 2% water, 3 equivalents TFA at 295K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	exch.	4.115	2.890, 2.694	
Ala 9	8.368 [-2.563]	4.344	1.266	
Glu 10	8.155 [-3.620]	4.258	1.902, 1.761	γ,γ′ 2.278, 2.278
Ala 11	8.038 [-3.702]	4.299	1.166	
Val 12	7.778 [-3.624]	4.075	1.874	γ,γ′ 0.726, 0.726
Tyr 13	7.851 [-3.456]	4.425	2.827, 2.627	δ,δ′ 6.969, ε,ε′ 6.613
Phe 14	8.037 [-4.478]	4.547	3.029, 2.805	δ,δ′ 7.236, ε,ε′ 7.254, ζ 7.206
Ala 15	8.216 [-5.684]	4.266	1.223	
His 16	8.201 [-3.721]	4.632	3.119, 3.010	δ 7.362, ε 8.927
Leu 17	8.071 [-4.014]	4.342	1.472	γ 1.597, δ,δ′ 0.872, 0.845
Asp 18	8.527 [-3.963]	4.630	2.730, 2.515	
Ile 19	7.584 [-3.015]	4.237	1.686	γ1,γ1′ 1.369, 0.996
				γ2 0.742, δ 0.754
Ile 20	7.889 [-3.909]	4.233	1.723	γ1,γ1′ 1.405, 1.048
				γ2 0.810, δ 0.797
Trp 21	8.208 [-5.684]	4.521	3.152, 3.055	δ 7.146, ε1 10.833, ε3 7.534
				ζ2 7.355, ζ3 7.005, η2 7.092

Table C.2: DAEAVYFAHLDI$^{(NMe)}$IW in DMSO, 2% water, 3 equivalents TFA at 300K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	exch.	4.120	2.893, 2.701	
Ala 9	8.632 [-3.364]	4.348	1.271	
Glu 10	8.142 [-4.754]	4.257	1.903, 1.774	γ,γ′ 2.282, 2.282
Ala 11	8.018 [-4.744]	4.298	1.171	
Val 12	7.763 [-4.988]	4.073	1.875	γ,γ′ 0.731, 0.731
Tyr 13	7.847 [-4.572]	4.430	2.830, 2.631	δ,δ′ 6.966, ε,ε′ 6.613
Phe 14	8.016 [-5.843]	4.542	3.034, 2.807	δ,δ′ 7.245, ε,ε′ ?, ζ ?
Ala 15	8.192 [-7.052]	4.261	1.225	
His 16	8.181 [-4.749]	4.628	3.132, 3.014	δ 7.361, ε 8.917
Leu 17	8.043 [-4.869]	4.344	1.457, 1.457?	γ 1.605, δ,δ′ 0.880, 0.855
Asp 18	8.454 [-4.878]	4.593	2.622, 2.476	
Ile 19	7.836 [-7.258]	4.478	1.669	γ1,γ1′ 1.438, 0.979
				γ2 0.567, δ 0.755
$^{(NMe)}$Ile 20	n/a	4.762	1.934	γ1,γ1′ 1.197, 0.874
				γ2 0.855, δ 0.779, NMe 2.968
Trp 21	8.222 [-9.857]	4.457	3.154, 3.052	δ 7.092, ε1 10.809, ε3 7.511
				ζ2 7.352, ζ3 7.008, η2 7.092

Table C.3: Ac-DAEAVYFAHLDI$^{(NMe)}$IW in DMSO, 2% water, 3 equivalents TFA at 295K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Asp 8	8.276 [-4.635]	4.560	2.699, 2.499	Ac-Me 1.864
Ala 9	8.101 [-4.529]	4.213	1.237	
Glu 10	7.971 [-3.292]	4.223	1.917, 1.777	γ,γ′ 2.281, 2.281?
Ala 11	7.942 [-4.740]	4.269	1.179	
Val 12	7.737 [-5.280]	4.044	1.878	γ,γ′ 0.738, 0.718
Tyr 13	7.833 [-4.969]	4.397	2.822, 2.622	δ,δ′ 6.961, ε,ε′ 6.608
Phe 14	8.007 [-6.345]	4.533	3.032, 2.804	δ,δ′ 7.251, ε,ε′ 7.251, ζ 7.203
Ala 15	8.201 [-7.516]	4.255	1.218	
His 16	8.185 [-5.389]	4.626	3.121, 3.002	δ 7.381, ε 8.952
Leu 17	8.046 [-5.385]	4.342	1.458, 1.458?	γ 1.597, δ,δ′ 0.876, 0.850
Asp 18	8.463 [-5.208]	4.584	2.601, 2.457	
Ile 19	7.891 [-8.452]	4.453	1.667	γ1,γ1′ 1.444, 0.978
				γ2 0.564, δ 0.745
$^{(NMe)}$Ile 20	n/a	4.776	1.929	γ1,γ1′ 1.199, 0.879
				γ2 0.850, δ 0.779, NMe 2.971
Trp 21	8.296 [-10.401]	4.441	3.145, 3.042	δ 7.088, ε1 10.837, ε3 7.511
				ζ2 7.349, ζ3 7.004, η2 7.088

Table C.4: GSHLDIIW in DMSO, 5% water, 6% acetic acid, 0.5% TFA at 310K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.207?, 3.207?		
Ser 15	8.519 [-3.464]	4.429	3.650, 3.594	
His 16	8.370 [-3.897]	4.638	3.187, 3.027	δ 7.339, ε 8.895
Leu 17	8.030 [-3.690]	4.331	1.522, 1.470	γ 1.608, δ,δ′ 0.888, 0.858
Asp 18	8.460 [-4.529]	4.638	2.754, 2.538	
Ile 19	7.549 [-3.347]	4.230	1.707	γ1,γ1′ 1.385, 1.011
				γ2 0.755, δ 0.778
Ile 20	7.829 [-4.240]	4.233	1.738	γ1,γ1′ 1.414, 1.057
				γ2 0.809, δ 0.798
Trp 21	8.097 [-6.198]	4.542	3.158, 3.071	δ 7.144, ε1 10.749, ε3 7.537
				ζ2 7.359, ζ3 7.002, η2 7.087

Table C.5: Ac-GSHLDIIW in DMSO, 2% water, 3 equivalents TFA at 315K

		Chemical Shift (ppm)		
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	8.144 [-3.780]	3.743, 3.743?		Ac-Me 1.903
Ser 15	7.956 [-4.151]	4.288	3.647, 3.600	
His 16	8.193 [-3.398]	4.590	3.185, 3.032	δ 7.326, ε 8.877
Leu 17	7.896 [-3.086]	4.308	1.476, 1.476?	γ 1.590, δ,δ′ 0.885, 0.851
Asp 18	8.356 [-4.197]	4.619	2.746, 2.525	
Ile 19	7.493 [-3.019]	4.219	1.700	γ1,γ1′ 1.379, 1.005
				γ2 0.747, δ 0.756
Ile 20	7.791 [-3.905]	4.219	1.732	γ1,γ1′ 1.405, 1.052
				γ2 0.795, δ 0.791
Trp 21	8.071 [-5.715]	4.529	3.156, 3.048	δ 7.137, ε1 10.751, ε3 7.529
				ζ2 7.357, ζ3 7.000, η2 7.086

Table C.6: GSHLDI$^{(NMe)}$IW in DMSO, 2% water, 3 equivalents TFA at 315K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Gly 14	exch.	3.592?, 3.592?		
Ser 15	8.476 [-3.923]	4.424	3.636, 3.605	
His 16	8.320 [-4.405]	4.634	3.183, 3.022	δ 7.334, ε 8.888
Leu 17	7.971 [-3.937]	4.328	1.675, 1.675?	γ 1.609, δ,δ′ 0.897, 0.867
Asp 18	8.368 [-4.692]	4.597	2.653, 2.490	
Ile 19	7.690 [-7.726]	4.499	1.678	γ1,γ1′ 1.436, 0.998
$^{(NMe)}$Ile 20	n/a	4.752	1.951	γ1,γ1′ 1.215, 0.901
Trp 21	8.042 [-9.685]	4.485	3.154, 3.062	δ 7.090, ε1 10.734, ε3 7.510

Table C.7: EAVYFAHLD in DMSO, 2% water, 3 equivalents TFA at 300K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	3.787	1.904, 1.889	γ,γ′ 2.367, 2.321
Ala 11	8.613 [-2.928]	4.380	1.167	
Val 12	7.912 [-3.886]	4.092	1.851	γ,γ′ 0.725, 0.708
Tyr 13	7.871 [-3.458]	4.428	2.801, 2.606	δ,δ′ 6.956, ε,ε′ 6.592
Phe 14	8.051 [-4.588]	4.542	3.002, 2.781	δ,δ′ 7.225, ε,ε′ 7.225, ζ 7.225?
Ala 15	8.219 [-5.516]	4.242	1.200	
His 16	8.195 [-3.844]	4.603	3.101, 2.989	δ 7.343, ε 8.900
Leu 17	8.056 [-4.039]	4.332	1.461, 1.461?	γ 1.601, δ,δ′ 0.871, 0.834
Asp 18	8.452 [-4.873]	4.556	2.708, 2.614	

Table C.8: Ac-EAVYFAHLD in DMSO, 2% water, 3 equivalents TFA at 305K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	8.088 [-3.148]	4.252	1.879, 1.754	γ,γ′ 2.282, 2.266, Ac-Me 1.883
Ala 11	8.137 [-3.355]	4.274	1.185	
Val 12	7.695 [-3.214]	4.043	1.871	γ,γ′ 0.739, 0.720
Tyr 13	7.803 [-3.292]	4.404	2.816, 2.632	δ,δ′ 6.961, ε,ε′ 6.614
Phe 14	7.959 [-4.466]	4.532	3.037, 2.813	δ,δ′ 7.249, ε,ε′ 7.249?, ζ 7.249?
Ala 15	8.139 [-5.453]	4.243	1.226	
His 16	8.166 [-3.920]	4.622	3.133, 3.007	δ 7.363, ε 8.909
Leu 17	8.025 [-3.905]	4.358	1.484, 1.484?	γ 1.631, δ,δ′ 0.891, 0.857
Asp 18	8.425 [-4.689]	4.575	2.729, 2.621	

Table C.9: EAVYFAHLDIIW in DMSO at 315K

	Chemical Shift (ppm)			
Residue	HN [Δδ/ΔT]	α	β,β′	Others
Glu 10	exch.	3.827	1.943, 1.943?	γ,γ′ 2.369, 2.369
Ala 11	8.533 [-3.574]	4.442	1.203	
Val 12	7.840 [-4.869]	4.136	1.895	γ,γ′ 0.762, 0.762
Tyr 13	7.798 [-4.286]	4.463	2.840, 2.640	δ,δ′ 6.964, ε,ε′ 6.609
Phe 14	7.955 [-5.227]	4.581	3.038, 2.822	δ,δ′ 7.229, ε,ε′ 7.229?, ζ 7.229?
Ala 15	8.101 [-6.656]	4.296	1.232	
His 16	8.012 [-3.736]	4.550	3.034, 2.959	δ 7.041, ε ?
Leu 17	7.920 [-5.284]	4.329	1.490	γ 1.561, δ,δ′ 0.870, 0.832
Asp 18	8.572 [-6.761]	4.638	2.773, 2.516	
Ile 19	7.441 [-2.991]	4.246	1.711	γ1,γ1′ 1.383, 1.010
				γ2 0.764, δ 0.764
Ile 20	7.748 [-5.122]	4.247	1.738	γ1,γ1′ 1.424, 1.070
				γ2 0.817, δ 0.807
Trp 21	8.032 [-6.961]	4.541	3.155, 3.065	δ 7.138, ε1 10.762, ε3 7.534
				ζ2 7.344, ζ3 6.991, η2 7.075

APPENDIX D: PEN-1 DISTANCE AND DIHEDRAL CONSTRAINTS

Final [$Pen^{3,15}Nle^{7}$]-ET-1 Distance Constraints, from Version 6.0
Pen-1 ET analog in 50% aqueous glycol at pH 3.5

1st value is equilibrium value (d), 2nd value is negative correction (d-), 3rd value is positive correction (d+)

CATEGORY 1:

```
remarks  PEN1_CAT1.tbl
remarks  NOE Classification: Key (category 1), version 6.0
remarks  NOTE:  All 3,15 gamma constraints in PEN_CAT1W.tbl
remarks  NOTE:  all beta/gamma H's wildcarded except 1,3,12,11,15
remarks  created by GML 6-11-93
remarks  edited 7-13-94

set echo=false end
set wrnlev=0 end

remarks  Group 1:  HNi/HNi+1
assign (resid   2 and  name HN  )(resid   3 and  name HN  ) 4.00 0.50 3.10
assign (resid   3 and  name HN  )(resid   4 and  name HN  ) 4.00 0.50 3.10
assign (resid   8 and  name HN  )(resid   9 and  name HN  ) 3.80 0.30 3.10
assign (resid  10 and  name HN  )(resid  11 and  name HN  ) 2.80 0.30 0.30
assign (resid  11 and  name HN  )(resid  12 and  name HN  ) 2.80 0.30 0.30
assign (resid  14 and  name HN  )(resid  15 and  name HN  ) 3.10 0.40 0.40
assign (resid  15 and  name HN  )(resid  16 and  name HN  ) 3.00 0.40 0.25
assign (resid  16 and  name HN  )(resid  17 and  name HN  ) 3.50 0.30 3.10

remarks  Group 2:  HNi/HNi+2
assign (resid   9 and  name HN  )(resid  11 and  name HN  ) 4.00 0.70 5.10
assign (resid  10 and  name HN  )(resid  12 and  name HN  ) 4.00 0.70 5.10
assign (resid  11 and  name HN  )(resid  13 and  name HN  ) 4.00 0.70 5.10

remarks  Group 4:  HAi/HNi+1
assign (resid   1 and  name HA  )(resid   2 and  name HN  ) 2.20 0.20 0.35
assign (resid   2 and  name HA  )(resid   3 and  name HN  ) 2.20 0.20 0.35
assign (resid   8 and  name HA  )(resid   9 and  name HN  ) 2.30 0.20 0.45
```

```
assign (resid   9 and name HA  )(resid  10 and name HN  ) 3.40 0.20 0.30
assign (resid  10 and name HA  )(resid  11 and name HN  ) 3.60 0.30 0.40
assign (resid  11 and name HA  )(resid  12 and name HN  ) 3.40 0.30 0.30
assign (resid  12 and name HA  )(resid  13 and name HN  ) 3.60 0.30 0.30
assign (resid  13 and name HA  )(resid  14 and name HN  ) 3.20 0.20 0.30
assign (resid  14 and name HA  )(resid  15 and name HN  ) 3.60 0.20 0.30
assign (resid  16 and name HA  )(resid  17 and name HN  ) 3.20 0.25 0.50

remarks Group 5: HAi/HNi+2
assign (resid   8 and name HA  )(resid  10 and name HN  ) 4.00 0.80 0.60

remarks Group 6: HAi/HNi+3
assign (resid   9 and name HA  )(resid  12 and name HN  ) 3.10 0.25 0.30
assign (resid  10 and name HA  )(resid  13 and name HN  ) 3.50 0.50 1.20
assign (resid  12 and name HA  )(resid  15 and name HN  ) 3.50 0.25 0.55
assign (resid  13 and name HA  )(resid  16 and name HN  ) 4.00 0.60 1.40
assign (resid  14 and name HA  )(resid  17 and name HN  ) 3.80 0.30 0.70
assign (resid  15 and name HA  )(resid  18 and name HN  ) 3.50 0.30 0.50

remarks Group 9: HBi/HNi+1
assign (resid   1 and name HB* )(resid   2 and name HN  ) 3.30 0.30 0.40
assign (resid   2 and name HB* )(resid   3 and name HN  ) 3.00 0.40 0.40
assign (resid   8 and name HB* )(resid   9 and name HN  ) 4.00 0.30 3.10
assign (resid   9 and name HB* )(resid  10 and name HN  ) 3.00 0.40 0.40
assign (resid  10 and name HB* )(resid  11 and name HN  ) 2.80 0.30 0.60
assign (resid  11 and name HB* )(resid  12 and name HN  ) 3.00 0.30 0.40
assign (resid  12 and name HB  )(resid  13 and name HN  ) 2.70 0.50 0.30
assign (resid  13 and name HB* )(resid  14 and name HN  ) 2.70 0.55 0.50
assign (resid  14 and name HB* )(resid  15 and name HN  ) 3.30 0.20 0.90
assign (resid  16 and name HB* )(resid  17 and name HN  ) 3.50 0.30 0.80

remarks Group 13: HAi/HBi+3
assign (resid   8 and name HA  )(resid  11 and name HB* ) 3.80 0.60 5.10
assign (resid   9 and name HA  )(resid  12 and name HB  ) 2.55 0.40 0.35
assign (resid  10 and name HA  )(resid  13 and name HB* ) 3.15 0.45 0.40
assign (resid  11 and name HA  )(resid  14 and name HB* ) 3.50 0.50 1.40
assign (resid  13 and name HA  )(resid  16 and name HB* ) 3.40 0.55 0.50
assign (resid  15 and name HA  )(resid  18 and name HB* ) 3.30 0.30 0.50
assign (resid  14 and name HA  )(resid  17 and name HB* ) 3.50 0.50 0.80
assign (resid  16 and name HA  )(resid  19 and name HB  ) 4.00 0.40 9.90
```

```
remarks  Group 15: HGi/HNi
assign (resid  12 and  name HG2* )(resid  12 and  name HN  ) 2.60 0.40 0.40
assign (resid  12 and  name HG1* )(resid  12 and  name HN  ) 4.00 0.40 0.40

remarks  Group 16:  HGi/HNi+1
assign (resid   7 and  name HG*  )(resid   8 and  name HN  ) 4.00 0.40 9.90
assign (resid   9 and  name HG*  )(resid  10 and  name HN  ) 3.80 0.40 9.90
assign (resid  12 and  name HG2* )(resid  13 and  name HN  ) 3.80 0.30 0.70
assign (resid  12 and  name HG1* )(resid  13 and  name HN  ) 3.80 0.40 0.70

remarks  Group 18:  Miscellaneous Interresidue constraints
assign (resid   1 and  name HB*  )(resid  15 and  name HN   ) 3.80 0.35 1.20
assign (resid   9 and  name HA   )(resid  12 and  name HG1* ) 4.50 0.50 3.10
assign (resid   9 and  name HA   )(resid  12 and  name HG2* ) 3.30 0.40 0.40
assign (resid  10 and  name HA   )(resid  13 and  name HD*  ) 3.30 0.30 0.60


remarks  PEN1_CAT1W.tbl
remarks  Model 1:  3GU=1 3GD=2  15GU=1 15GD=2
remarks  created by GML 6-11-93
remarks  edited 7-13-94

remarks  Group 15: HGi/HNi
assign (resid   3 and  name HG1* )(resid   3 and  name HN  ) 2.60 0.40 0.40
assign (resid   3 and  name HG2* )(resid   3 and  name HN  ) 4.20 0.20 0.40
assign (resid  15 and  name HG*  )(resid  15 and  name HN  ) 3.00 0.40 0.60

remarks          Group 16: HGi/HNi+1
assign (resid   3 and  name HG** )(resid   4 and  name HN  ) 3.40 0.60 0.50
assign (resid  15 and  name HG** )(resid  16 and  name HN  ) 3.30 0.30 0.60

remarks  Group 17: Pen Methyls
assign (resid   1 and  name HB*  )(resid  15 and  name HG2* ) 3.50 0.30 1.20
assign (resid   1 and  name HB*  )(resid  15 and  name HG1* ) 3.70 0.20 1.20
assign (resid   3 and  name HG** )(resid   5 and  name HN   ) 2.90 0.60 0.80
assign (resid   3 and  name HG2* )(resid  11 and  name HB*  ) 3.30 0.40 0.80
assign (resid  12 and  name HA   )(resid  15 and  name HG2* ) 3.40 0.80 0.80
assign (resid  12 and  name HA   )(resid  15 and  name HG1* ) 3.20 0.50 0.50
```

CATEGORY 2:

```
remarks  PEN1_CAT2.tbl
remarks  NOE Classification: (category 2), version 6.0
remarks  NOTE:  All 3,15 gamma constraints in PEN_CAT2X.tbl
remarks  NOTE:  all beta/gamma H's wildcarded except 1,3,12,11,15
remarks  created by GML 6-11-93
remarks  edited 7-13-94

remarks  Group 1:  HNi/HNi+1
assign (resid   4 and  name HN  )(resid   5 and  name HN  ) 3.00 0.60 0.40
assign (resid   5 and  name HN  )(resid   6 and  name HN  ) 4.00 1.00 3.10
assign (resid   6 and  name HN  )(resid   7 and  name HN  ) 3.00 0.50 0.50
assign (resid   9 and  name HN  )(resid  10 and  name HN  ) 3.00 0.30 0.40
assign (resid  12 and  name HN  )(resid  13 and  name HN  ) 3.00 0.40 0.70
assign (resid  13 and  name HN  )(resid  14 and  name HN  ) 2.80 0.20 0.40

remarks  Group 2:  HNi/HNi+2
assign (resid   4 and  name HN  )(resid   6 and  name HN  ) 4.00 0.70 5.10
assign (resid   8 and  name HN  )(resid  10 and  name HN  ) 4.00 0.40 5.10
assign (resid  12 and  name HN  )(resid  14 and  name HN  ) 4.00 0.40 5.10
assign (resid  13 and  name HN  )(resid  15 and  name HN  ) 4.00 0.40 5.10
assign (resid  14 and  name HN  )(resid  16 and  name HN  ) 4.00 0.40 5.10
assign (resid  15 and  name HN  )(resid  17 and  name HN  ) 4.00 0.40 5.10
assign (resid  16 and  name HN  )(resid  18 and  name HN  ) 4.00 0.40 5.10
assign (resid  17 and  name HN  )(resid  19 and  name HN  ) 4.00 0.70 5.10
assign (resid  18 and  name HN  )(resid  20 and  name HN  ) 4.00 0.70 5.10

remarks  Group 3:  HAi/HNi
assign (resid  10 and  name HA  )(resid  10 and  name HN  ) 2.30 0.25 0.25
assign (resid  11 and  name HA  )(resid  11 and  name HN  ) 2.80 0.30 0.25
assign (resid  18 and  name HA  )(resid  18 and  name HN  ) 2.30 0.25 0.25

remarks  Group 4:  HAi/HNi+1
assign (resid  15 and  name HA  )(resid  16 and  name HN  ) 3.20 0.40 0.20
assign (resid  19 and  name HA  )(resid  20 and  name HN  ) 2.35 0.20 0.30

remarks  Group 5:  HAi/HNi+2
assign (resid   6 and  name HA  )(resid   8 and  name HN  ) 3.30 0.40 0.20
assign (resid   9 and  name HA  )(resid  11 and  name HN  ) 4.00 0.80 5.10
assign (resid  12 and  name HA  )(resid  14 and  name HN  ) 4.00 0.40 5.10
assign (resid  13 and  name HA  )(resid  15 and  name HN  ) 4.00 0.40 1.40
```

```
assign (resid  14 and name HA  )(resid  16 and name HN  ) 4.00 0.40 5.10
assign (resid  15 and name HA  )(resid  17 and name HN  ) 4.00 0.40 1.40

remarks Group 6: HAi/HNi+3
assign (resid   8 and name HA  )(resid  11 and name HN  ) 4.00 0.60 9.90
assign (resid  11 and name HA  )(resid  14 and name HN  ) 4.00 0.70 5.10

remarks Group 8: HBi/HNi
assign (resid   5 and name HB* )(resid   5 and name HN  ) 2.70 0.20 0.25
assign (resid   7 and name HB* )(resid   7 and name HN  ) 2.70 0.20 0.25
assign (resid   8 and name HB* )(resid   8 and name HN  ) 2.70 0.20 0.35
assign (resid  11 and name HB* )(resid  11 and name HN  ) 2.70 0.45 0.40
assign (resid  12 and name HB  )(resid  12 and name HN  ) 2.50 0.20 0.20
assign (resid  13 and name HB* )(resid  13 and name HN  ) 2.55 0.45 0.40
assign (resid  14 and name HB* )(resid  14 and name HN  ) 2.60 0.45 0.50

remarks Group 10: HBi/HNi+2
assign (resid  12 and name HB  )(resid  14 and name HN  ) 4.00 0.40 5.10

remarks Group 11: HAi/HBi
assign (resid   5 and name HA  )(resid   5 and name HB* ) 2.50 0.20 0.25
assign (resid   7 and name HA  )(resid   7 and name HB* ) 2.50 0.20 0.25
assign (resid   8 and name HA  )(resid   8 and name HB* ) 2.50 0.20 0.25
assign (resid  11 and name HA  )(resid  11 and name HB* ) 2.50 0.20 0.25
assign (resid  12 and name HA  )(resid  12 and name HB  ) 3.00 0.20 0.15

remarks Group 14: HAi/HGi
assign (resid  12 and name HA  )(resid  12 and name HG2* ) 2.80 0.30 0.40
assign (resid  12 and name HA  )(resid  12 and name HG1* ) 3.10 0.40 0.50

remarks Group 16: HGi/HNi+1
assign (resid  10 and name HG* )(resid  11 and name HN  ) 3.80 0.70 0.50

remarks Group 20: Disulfides
assign (resid   1 and name CB  )(resid  15 and name CB  ) 3.94 0.50 0.40
assign (resid   1 and name CB  )(resid  15 and name SG  ) 3.05 0.15 0.32
assign (resid   1 and name SG  )(resid  15 and name CB  ) 3.05 0.15 0.32
assign (resid   1 and name SG  )(resid  15 and name SG  ) 2.02 0.06 0.05
assign (resid   3 and name CB  )(resid  11 and name CB  ) 3.94 0.50 0.40
assign (resid   3 and name CB  )(resid  11 and name SG  ) 3.05 0.15 0.32
assign (resid   3 and name SG  )(resid  11 and name CB  ) 3.05 0.15 0.32
```

assign (resid 3 and name SG)(resid 11 and name SG) 2.02 0.06 0.05

remarks **PEN1_CAT2A.tbl**
remarks Model 1,2: 15GU=2 15GD=1
remarks Model *: 1BU=* 1BD=*
remarks created by GML 8-19-93
remarks edited 7-13-94

remarks Group 9: HBi/HNi+1
assign (resid 1 and name HB*)(resid 2 and name HN) 3.30 0.30 0.40

remarks Group 10: HBi/HNi+2
assign (resid 1 and name HB*)(resid 3 and name HN) 4.00 0.60 5.10

remarks Group 17: Pen Methyls
assign (resid 1 and name HB*)(resid 15 and name HG2*) 3.40 0.40 0.50
assign (resid 1 and name HB*)(resid 15 and name HG2*) 3.50 0.10 1.40

remarks Group 18: Miscellaneous Interresidue contraints
assign (resid 1 and name HB*)(resid 15 and name HN) 3.80 0.60 1.40

remarks **PEN1_CAT2B.tbl**
remarks Model 1,3: 3GU=1 3GD=2
remarks Model D: 11BU=2 11BD=1
remarks created by GML 8-19-93
remarks edited 7-13-94

remarks Group 8: HBi/HNi
assign (resid 11 and name HB2)(resid 11 and name HN) 2.75 0.45 0.40
assign (resid 11 and name HB1)(resid 11 and name HN) 2.65 0.35 0.40

remarks Group 9: HBi/HNi+1
assign (resid 11 and name HB2)(resid 12 and name HN) 3.20 0.30 0.40
assign (resid 11 and name HB1)(resid 12 and name HN) 3.00 0.30 0.30

remarks Group 11: HAi/HBi
assign (resid 11 and name HA)(resid 11 and name HB2) 2.50 0.20 0.25
assign (resid 11 and name HA)(resid 11 and name HB1) 3.00 0.20 0.15

```
remarks Group 13: HAi/HBi+3
assign (resid  8 and name HA  )(resid  11 and name HB1 ) 3.80 0.60 5.10
assign (resid  8 and name HA  )(resid  11 and name HB2 ) 3.80 0.40 5.10

remarks Group 17: Pen Methyls
assign (resid  3 and name HG2* )(resid  11 and name HB2 ) 3.50 0.20 1.20
assign (resid  3 and name HG2* )(resid  11 and name HB1 ) 3.20 0.40 0.70
```

```
remarks PEN1_CAT2X.tbl
remarks Model 1: 3GU=1 3GD=2 15GU=1 15GD=2
remarks created by GML 6-11-93
remarks edited 7-13-94

remarks Group 14: HAi/HGi
assign (resid  3 and name HA  )(resid  3 and name HG2* ) 2.80 0.30 0.40
assign (resid  3 and name HA  )(resid  3 and name HG1* ) 3.80 0.30 0.40
assign (resid 15 and name HA  )(resid 15 and name HG2* ) 3.40 0.40 0.70
assign (resid 15 and name HA  )(resid 15 and name HG1* ) 3.20 0.40 0.70

remarks Group 15: HGi/HNi
assign (resid 15 and name HG2* )(resid 15 and name HN  ) 2.80 0.40 0.40
assign (resid 15 and name HG1* )(resid 15 and name HN  ) 3.25 0.20 0.55

remarks Group 16: HGi/HNi+1
assign (resid  3 and name HG2* )(resid  4 and name HN  ) 3.40 0.60 0.50
assign (resid  3 and name HG1* )(resid  4 and name HN  ) 3.80 0.40 0.50
assign (resid 15 and name HG1* )(resid 16 and name HN  ) 3.50 0.30 0.70
assign (resid 15 and name HG2* )(resid 16 and name HN  ) 3.30 0.30 0.60

remarks Group 17: Pen methyls
assign (resid  2 and name HA   )(resid 15 and name HG** ) 3.70 0.50 0.80
assign (resid  3 and name HG1* )(resid  5 and name HN   ) 3.00 0.60 0.40
assign (resid  3 and name HG2* )(resid  5 and name HN   ) 4.40 0.60 9.10
assign (resid  3 and name HG1* )(resid 11 and name HN   ) 4.40 0.30 9.10
assign (resid  3 and name HG1* )(resid 11 and name HB*  ) 4.20 0.30 5.10
```

CATEGORY 3:

```
remarks PEN1_CAT3.tbl
remarks NOE Classification: category 3, version 6.0
```

```
remarks  NOTE:  All 3,15 gamma constraints in PEN_CAT3Y.tbl
remarks  NOTE:  all beta/gamma H's wildcarded except 1,3,12,11,15
remarks  created by GML 6-11-93
remarks  edited 7-13-94

remarks  Group 1:  HNi/HNi+1
assign (resid   7 and  name HN  )(resid   8 and  name HN  ) 3.10 0.70 3.10
assign (resid  18 and  name HN  )(resid  19 and  name HN  ) 3.00 0.40 0.40
assign (resid  19 and  name HN  )(resid  20 and  name HN  ) 3.35 0.40 0.90
assign (resid  20 and  name HN  )(resid  21 and  name HN  ) 3.10 0.30 0.40

remarks  Group 2:  HNi/HNi+2
assign (resid   4 and  name HN  )(resid   6 and  name HN  ) 4.00 0.70 5.10
assign (resid  17 and  name HN  )(resid  19 and  name HN  ) 4.00 0.70 5.10
assign (resid  18 and  name HN  )(resid  20 and  name HN  ) 4.00 0.70 5.10

remarks  Group 3:  HAi/HNi
assign (resid   2 and  name HA  )(resid   2 and  name HN  ) 2.80 0.30 0.25
assign (resid   4 and  name HA  )(resid   4 and  name HN  ) 2.70 0.35 0.20
assign (resid   6 and  name HA  )(resid   6 and  name HN  ) 2.80 0.30 0.25
assign (resid   7 and  name HA  )(resid   7 and  name HN  ) 2.40 0.30 0.20
assign (resid   8 and  name HA  )(resid   8 and  name HN  ) 2.70 0.35 0.20
assign (resid   9 and  name HA  )(resid   9 and  name HN  ) 2.60 0.25 0.25
assign (resid  17 and  name HA  )(resid  17 and  name HN  ) 2.40 0.30 0.20

remarks  Group 4:  HAi/HNi+1
assign (resid  17 and  name HA  )(resid  18 and  name HN  ) 2.35 0.40 0.30
assign (resid  18 and  name HA  )(resid  19 and  name HN  ) 2.40 0.30 0.30
assign (resid  20 and  name HA  )(resid  21 and  name HN  ) 2.25 0.20 0.35

remarks  Group 5:  HAi/HNi+2
assign (resid   4 and  name HA  )(resid   6 and  name HN  ) 4.00 0.30 5.10
assign (resid   5 and  name HA  )(resid   7 and  name HN  ) 3.70 0.60 0.90

remarks  Group 7:  HAi/HNi4
assign (resid   9 and  name HA  )(resid  13 and  name HN  ) 4.00 0.60 9.90
assign (resid  10 and  name HA  )(resid  14 and  name HN  ) 4.00 0.70 9.90
assign (resid  11 and  name HA  )(resid  15 and  name HN  ) 4.00 0.60 9.90
assign (resid  12 and  name HA  )(resid  16 and  name HN  ) 4.00 0.70 9.90
assign (resid  13 and  name HA  )(resid  17 and  name HN  ) 4.00 0.70 9.90
assign (resid  14 and  name HA  )(resid  18 and  name HN  ) 4.00 0.60 9.90
```

```
remarks Group 8: HBi/HNi
assign (resid   2 and name HB* )(resid   2 and name HN  ) 3.45 0.45 0.60
assign (resid   4 and name HB* )(resid   4 and name HN  ) 3.35 0.45 0.50
assign (resid   6 and name HB* )(resid   6 and name HN  ) 3.50 0.70 0.80
assign (resid   9 and name HB* )(resid   9 and name HN  ) 3.15 0.45 0.50
assign (resid  10 and name HB* )(resid  10 and name HN  ) 2.60 0.50 0.50
assign (resid  16 and name HB* )(resid  16 and name HN  ) 2.55 0.25 0.70
assign (resid  17 and name HB* )(resid  17 and name HN  ) 2.45 0.45 0.35
assign (resid  18 and name HB* )(resid  18 and name HN  ) 2.90 0.40 0.40
assign (resid  19 and name HB  )(resid  19 and name HN  ) 2.60 0.40 0.40
assign (resid  20 and name HB  )(resid  20 and name HN  ) 2.60 0.40 0.40
assign (resid  21 and name HB* )(resid  21 and name HN  ) 2.65 0.20 0.25

remarks Group 9: HBi/HNi+1
assign (resid   6 and name HB* )(resid   7 and name HN  ) 3.80 0.40 0.60
assign (resid   7 and name HB* )(resid   8 and name HN  ) 3.30 0.60 1.00

remarks Group 10: HBi/HNi+2
assign (resid   5 and name HB* )(resid   7 and name HN  ) 3.80 0.50 5.10
assign (resid   8 and name HA  )(resid  10 and name HB* ) 4.00 0.40 9.90
assign (resid   9 and name HA  )(resid  11 and name HB* ) 4.00 0.40 9.90
assign (resid  10 and name HA  )(resid  12 and name HB  ) 4.00 0.80 5.00
assign (resid  11 and name HA  )(resid  13 and name HB* ) 4.00 0.40 9.90
assign (resid  14 and name HA  )(resid  16 and name HB* ) 4.00 0.40 9.90

remarks Group 11: HAi/HBi
assign (resid  14 and name HA  )(resid  14 and name HB* ) 2.50 0.20 0.25

remarks Group 13: HAi/HBi+3
assign (resid   1 and name HA  )(resid   4 and name HB* ) 4.00 0.40 9.90
assign (resid   2 and name HA  )(resid   5 and name HB* ) 4.00 0.40 9.90
assign (resid   3 and name HA  )(resid   6 and name HB* ) 4.00 0.40 9.90
assign (resid   4 and name HA  )(resid   7 and name HB* ) 4.00 0.70 9.90
assign (resid   5 and name HA  )(resid   8 and name HB* ) 4.00 0.40 9.90
assign (resid   6 and name HA  )(resid   9 and name HB* ) 4.00 0.40 9.90
assign (resid   7 and name HA  )(resid  10 and name HB* ) 3.80 0.60 5.10
assign (resid  17 and name HA  )(resid  20 and name HB  ) 3.90 0.80 9.90
assign (resid  18 and name HA  )(resid  21 and name HB* ) 4.00 0.40 9.90

remarks Group 14: HAi/HGi
```

```
assign (resid  19 and  name HA   )(resid  19 and  name HG2* ) 2.80 0.30 0.50

remarks  Group 15:  HGi/HNi
assign (resid   7 and  name HG*  )(resid   7 and  name HN   ) 3.00 0.40 0.40
assign (resid  10 and  name HG*  )(resid  10 and  name HN   ) 3.30 0.40 0.40
assign (resid  19 and  name HG2* )(resid  19 and  name HN   ) 3.30 0.60 0.20

remarks  Group 16:  HGi/HNi+1
assign (resid  19 and  name HG2* )(resid  20 and  name HN   ) 3.30 0.50 0.40

remarks  Group 18:  Miscellaneous Interresidue contraints
assign (resid  10 and  name HA   )(resid  13 and  name HE*  ) 4.50 0.70 3.10
assign (resid  11 and  name HA   )(resid  14 and  name HD*  ) 3.40 0.50 1.20
assign (resid  12 and  name HG2* )(resid  13 and  name HD*  ) 4.50 0.40 9.90
assign (resid  12 and  name HG1* )(resid  13 and  name HD*  ) 4.50 0.60 9.90
assign (resid  12 and  name HG2* )(resid  13 and  name HE*  ) 4.00 0.35 0.80
assign (resid  12 and  name HG1* )(resid  13 and  name HE*  ) 4.00 0.35 0.80
assign (resid  12 and  name HG2* )(resid  16 and  name HD2  ) 4.20 0.40 9.00
assign (resid  12 and  name HG1* )(resid  16 and  name HD2  ) 3.70 0.80 1.00
assign (resid  12 and  name HG2* )(resid  16 and  name HE1  ) 4.20 0.40 9.90
assign (resid  12 and  name HG1* )(resid  16 and  name HE1  ) 4.20 0.60 9.90
assign (resid  13 and  name HE*  )(resid  17 and  name HD*  ) 3.40 0.50 1.00
assign (resid  14 and  name HA   )(resid  17 and  name HD*  ) 4.00 0.80 1.40
assign (resid  17 and  name HD*  )(resid  18 and  name HB*  ) 4.00 0.30 9.00
assign (resid  18 and  name HB*  )(resid  21 and  name HN   ) 4.00 0.50 9.10

remarks  Group 19:  Miscellaneous Intraresidue contraints
assign (resid  13 and  name HA   )(resid  13 and  name HD*  ) 3.00 0.35 0.70
assign (resid  14 and  name HA   )(resid  14 and  name HD*  ) 3.20 0.40 0.80


remarks  PEN1_CAT3Y.tbl
remarks  Model 1:  3GU=1 3GD=2  15GU=1 15GD=2
remarks  created by GML 6-11-93
remarks  edited 7-13-94

remarks  Group 17:  Pen methyls
assign (resid   2 and  name HN   )(resid  15 and  name HG2* ) 3.80 0.60 0.80
assign (resid   2 and  name HN   )(resid  15 and  name HG1* ) 4.50 0.40 4.50
assign (resid   3 and  name HG2* )(resid   6 and  name HN   ) 4.00 0.70 5.00
assign (resid   3 and  name HG1* )(resid   6 and  name HN   ) 4.00 0.50 9.90
```

```
assign (resid   3 and name HG2* )(resid   8 and name HB*  ) 3.40 0.40 1.40
assign (resid   3 and name HG1* )(resid   8 and name HB*  ) 4.00 0.40 3.10
assign (resid   3 and name HG1* )(resid  12 and name HA   ) 4.20 0.60 1.20
assign (resid   3 and name HG2* )(resid  12 and name HG1* ) 3.50 0.30 3.20
assign (resid   3 and name HG2* )(resid  12 and name HG2* ) 3.30 0.30 1.00
assign (resid   3 and name HG1* )(resid  12 and name HG1* ) 3.40 0.30 3.20
assign (resid   3 and name HG1* )(resid  12 and name HG2* ) 3.30 0.60 1.50
assign (resid  14 and name HD*  )(resid  15 and name HG1* ) 3.80 0.60 0.40
assign (resid  14 and name HD*  )(resid  15 and name HG2* ) 4.00 0.30 3.10
assign (resid  15 and name HG2* )(resid  16 and name HD2  ) 3.60 0.30 1.00
```

CATEGORY 4:

```
remarks  PEN1_CAT4.tbl
remarks  NOE Classification: category 4, version 6.0
remarks  NOTE:  All 3,15 gamma constraints in PEN_CAT4Z.tbl
remarks  NOTE:  all beta/gamma H's wildcarded except 1,3,12,11,15
remarks  created by GML 6-11-93
remarks  edited 7-13-94

remarks  Group 1:  HNi/HNi+1
assign (resid  17 and name HN   )(resid  18 and name HN   ) 3.10 0.40 0.55

remarks  Group 2:  HNi/HNi+2
assign (resid   2 and name HN   )(resid   4 and name HN   ) 4.00 0.40 5.10
assign (resid   3 and name HN   )(resid   5 and name HN   ) 4.00 0.40 5.10
assign (resid   6 and name HN   )(resid   8 and name HN   ) 4.00 0.40 5.10
assign (resid   7 and name HN   )(resid   9 and name HN   ) 4.00 0.40 5.10
assign (resid  19 and name HN   )(resid  21 and name HN   ) 4.00 0.40 5.10

remarks  Group 3:  HAi/HNi
assign (resid  12 and name HA   )(resid  12 and name HN   ) 2.50 0.35 0.25
assign (resid  13 and name HA   )(resid  13 and name HN   ) 2.50 0.35 0.25
assign (resid  14 and name HA   )(resid  14 and name HN   ) 2.60 0.35 0.25
assign (resid  15 and name HA   )(resid  15 and name HN   ) 2.80 0.20 0.25
assign (resid  16 and name HA   )(resid  16 and name HN   ) 2.80 0.20 0.25
assign (resid  19 and name HA   )(resid  19 and name HN   ) 2.45 0.30 0.25
assign (resid  21 and name HA   )(resid  21 and name HN   ) 2.80 0.30 0.25

remarks  Group 4:  HAi/HNi+1
assign (resid   5 and name HA   )(resid   6 and name HN   ) 2.65 0.30 0.40
```

```
assign (resid   6 and  name HA  )(resid   7 and  name HN  ) 2.65 0.30 0.60
assign (resid   7 and  name HA  )(resid   8 and  name HN  ) 3.20 0.30 0.80

remarks Group 5: HAi/HNi+2
assign (resid   1 and  name HA  )(resid   3 and  name HN  ) 4.00 0.50 5.10
assign (resid   2 and  name HA  )(resid   4 and  name HN  ) 4.00 0.50 5.10
assign (resid   7 and  name HA  )(resid   9 and  name HN  ) 4.20 0.50 4.10

remarks Group 6: HAi/HNi+3
assign (resid   4 and  name HA  )(resid   7 and  name HN  ) 4.00 0.60 5.10
assign (resid   5 and  name HA  )(resid   8 and  name HN  ) 4.00 0.60 5.10
assign (resid   6 and  name HA  )(resid   9 and  name HN  ) 4.00 0.60 9.90
assign (resid   7 and  name HA  )(resid  10 and  name HN  ) 4.00 1.00 9.90
assign (resid  16 and  name HA  )(resid  19 and  name HN  ) 4.00 0.40 9.90
assign (resid  17 and  name HA  )(resid  20 and  name HN  ) 3.50 0.40 5.10
assign (resid  18 and  name HA  )(resid  21 and  name HN  ) 4.00 0.40 9.90

remarks Group 7: HAi/Ni+4
assign (resid   1 and  name HA  )(resid   5 and  name HN  ) 4.00 0.60 9.90
assign (resid   2 and  name HA  )(resid   6 and  name HN  ) 4.00 0.60 9.90
assign (resid   4 and  name HA  )(resid   8 and  name HN  ) 4.00 0.60 9.90
assign (resid   5 and  name HA  )(resid   9 and  name HN  ) 4.00 0.80 9.90
assign (resid   6 and  name HA  )(resid  10 and  name HN  ) 4.00 0.60 9.90
assign (resid   7 and  name HA  )(resid  11 and  name HN  ) 4.00 0.80 9.90
assign (resid  15 and  name HA  )(resid  19 and  name HN  ) 4.00 0.60 9.90
assign (resid  16 and  name HA  )(resid  20 and  name HN  ) 4.00 0.60 9.90
assign (resid  17 and  name HA  )(resid  21 and  name HN  ) 4.00 0.80 9.90

remarks Group 9: HBi/HNi+1
assign (resid   4 and  name HB* )(resid   5 and  name HN  ) 3.50 0.80 0.65
assign (resid   5 and  name HB* )(resid   6 and  name HN  ) 4.00 0.80 9.90
assign (resid  17 and  name HB* )(resid  18 and  name HN  ) 3.60 0.80 0.50
assign (resid  18 and  name HB* )(resid  19 and  name HN  ) 3.50 0.30 0.50
assign (resid  19 and  name HB  )(resid  20 and  name HN  ) 3.30 0.60 0.50
assign (resid  20 and  name HB  )(resid  21 and  name HN  ) 3.30 0.60 0.50

remarks Group 10: HBi/HNi+2
assign (resid   2 and  name HB* )(resid   4 and  name HN  ) 4.00 0.60 5.10
assign (resid   4 and  name HB* )(resid   6 and  name HN  ) 4.00 0.60 5.10
assign (resid   6 and  name HB* )(resid   8 and  name HN  ) 4.00 0.70 5.10
assign (resid   7 and  name HB* )(resid   9 and  name HN  ) 4.00 0.70 5.10
```

```
assign (resid   8 and name HB* )(resid  10 and name HN  ) 4.00 0.60 5.10
assign (resid   9 and name HB* )(resid  11 and name HN  ) 4.00 0.60 5.10
assign (resid  10 and name HB* )(resid  12 and name HN  ) 4.00 0.80 5.10
assign (resid  11 and name HB* )(resid  13 and name HN  ) 4.00 0.60 5.10
assign (resid  13 and name HB* )(resid  15 and name HN  ) 4.00 0.60 5.10
assign (resid  14 and name HB* )(resid  16 and name HN  ) 4.00 0.60 5.10
assign (resid  16 and name HB* )(resid  18 and name HN  ) 4.00 0.60 5.10
assign (resid  17 and name HB* )(resid  19 and name HN  ) 3.60 0.40 5.10

remarks Group 12: HAi/HBi+2
assign (resid   2 and name HA  )(resid   4 and name HB* ) 4.00 0.40 9.90
assign (resid   3 and name HA  )(resid   5 and name HB* ) 4.00 0.40 9.90
assign (resid   4 and name HA  )(resid   6 and name HB* ) 4.00 0.50 9.90
assign (resid   5 and name HA  )(resid   7 and name HB* ) 4.00 0.40 9.90
assign (resid   6 and name HA  )(resid   8 and name HB* ) 4.00 0.40 9.90
assign (resid   7 and name HA  )(resid   9 and name HB* ) 4.00 0.70 9.00
assign (resid  15 and name HA  )(resid  17 and name HB* ) 4.00 0.40 9.90
assign (resid  16 and name HA  )(resid  18 and name HB* ) 4.00 0.40 9.90
assign (resid  17 and name HA  )(resid  19 and name HB  ) 4.00 0.50 9.90
assign (resid  18 and name HA  )(resid  20 and name HB  ) 4.00 0.40 9.90
assign (resid  19 and name HA  )(resid  21 and name HB* ) 4.00 0.40 9.90

remarks Group 15: HGi/HNi
assign (resid   6 and name HG  )(resid   6 and name HN  ) 3.10 0.30 0.40
assign (resid   9 and name HG* )(resid   9 and name HN  ) 3.30 0.30 3.10
assign (resid  17 and name HG  )(resid  17 and name HN  ) 3.50 0.50 0.60
assign (resid  19 and name HG1* )(resid  19 and name HN  ) 3.30 0.30 0.50
assign (resid  20 and name HG1* )(resid  20 and name HN  ) 3.30 0.30 0.50

remarks Group 16: HGi/HNi+1
assign (resid  17 and name HG  )(resid  18 and name HN  ) 3.80 0.40 1.80
assign (resid  19 and name HG1* )(resid  20 and name HN  ) 3.80 0.40 0.80
assign (resid  20 and name HG1* )(resid  21 and name HN  ) 4.00 0.60 9.90
assign (resid  20 and name HG2* )(resid  21 and name HN  ) 4.00 0.80 1.00

remarks Group 18: Miscellaneous Interresidue contraints
assign (resid   4 and name HB* )(resid   7 and name HN  ) 4.00 0.80 5.10
assign (resid   4 and name HB* )(resid   7 and name HB* ) 4.00 0.50 9.00
assign (resid   4 and name HB* )(resid   7 and name HG* ) 4.00 0.50 9.00
assign (resid   5 and name HN  )(resid   7 and name HG* ) 4.00 0.70 9.00
assign (resid   5 and name HA  )(resid   6 and name HG  ) 4.00 0.60 9.00
```

```
assign (resid   7 and name HG*  )(resid   8 and name HB*  ) 4.00 0.30 9.00
assign (resid   6 and name HA   )(resid  12 and name HG1* ) 4.00 1.20 9.90
assign (resid   6 and name HA   )(resid  12 and name HG2* ) 3.70 0.50 1.00
assign (resid   8 and name HA   )(resid   9 and name HB*  ) 4.00 0.60 9.00
assign (resid   8 and name HB*  )(resid  10 and name HB*  ) 4.00 0.50 9.00
assign (resid   8 and name HB*  )(resid  11 and name HN   ) 4.00 0.30 9.00
assign (resid   9 and name HA   )(resid  18 and name HB*  ) 4.00 0.30 9.00
assign (resid   9 and name HB*  )(resid  12 and name HN   ) 4.50 0.50 5.10
assign (resid   9 and name HB*  )(resid  13 and name HB*  ) 4.00 0.40 9.00
assign (resid   9 and name HG*  )(resid  13 and name HD*  ) 4.00 0.60 9.00
assign (resid   9 and name HD*  )(resid  13 and name HD*  ) 4.00 0.60 9.00
assign (resid   9 and name HE*  )(resid  12 and name HB   ) 4.00 0.30 9.90
assign (resid   9 and name HE*  )(resid  12 and name HG*  ) 4.00 0.30 9.90
assign (resid  10 and name HG*  )(resid  14 and name HD*  ) 4.00 0.70 0.90
assign (resid  10 and name HG*  )(resid  14 and name HE*  ) 3.90 0.50 1.00
assign (resid  12 and name HB   )(resid  14 and name HD*  ) 4.00 0.60 0.50
assign (resid  13 and name HD*  )(resid  14 and name HN   ) 3.80 0.50 0.80
assign (resid  13 and name HD*  )(resid  14 and name HA   ) 3.80 0.40 0.85
assign (resid  13 and name HD*  )(resid  17 and name HD*  ) 4.00 0.50 1.20
assign (resid  14 and name HD*  )(resid  16 and name HE1  ) 4.00 0.50 9.00
assign (resid  16 and name HA   )(resid  19 and name HG1* ) 4.00 0.50 9.00
assign (resid  16 and name HA   )(resid  19 and name HD*  ) 4.00 0.50 9.00
assign (resid  16 and name HD2  )(resid  17 and name HD*  ) 3.80 0.50 1.20
assign (resid  17 and name HD*  )(resid  18 and name HN   ) 4.00 0.60 5.10
assign (resid  18 and name HN   )(resid  19 and name HG1* ) 4.00 0.45 9.00
assign (resid  18 and name HN   )(resid  19 and name HG2* ) 4.00 0.30 9.00
assign (resid  19 and name HG2* )(resid  21 and name HN   ) 3.50 0.50 1.50
assign (resid  19 and name HD*  )(resid  20 and name HN   ) 4.00 0.50 5.10


remarks PEN1_CAT4Z.tbl
remarks Model 1: 3GU=1 3GD=2 15GU=1 15GD=2
remarks created by GML 6-11-93
remarks edited 7-13-94

remarks Group 17: Pen methyls
assign (resid   1 and name HA   )(resid   3 and name HG2* ) 4.00 0.50 9.00
assign (resid   1 and name HA   )(resid   3 and name HG1* ) 4.00 0.50 9.00
assign (resid   2 and name HN   )(resid   3 and name HG1* ) 4.00 0.20 5.10
assign (resid   2 and name HA   )(resid   3 and name HG1* ) 4.00 0.50 9.00
assign (resid   3 and name HG1* )(resid  12 and name HN   ) 4.00 0.20 1.40
```

```
assign (resid   3 and  name HG2* )(resid  12 and  name HN    ) 3.60 0.50 1.10
assign (resid   3 and  name HG2* )(resid  12 and  name HA    ) 4.00 0.80 5.10
assign (resid   3 and  name HG** )(resid  12 and  name HB    ) 4.40 0.40 9.00
assign (resid  12 and  name HN    )(resid  15 and  name HG2* ) 5.00 0.50 9.10
assign (resid  12 and  name HN    )(resid  15 and  name HG1* ) 4.50 0.40 9.10
assign (resid  12 and  name HB    )(resid  15 and  name HG1* ) 4.00 0.40 9.00
assign (resid  12 and  name HG** )(resid  15 and  name HG** ) 3.70 0.40 1.20
assign (resid  12 and  name HG1* )(resid  15 and  name HG1* ) 3.50 0.40 5.50
assign (resid  12 and  name HG1* )(resid  15 and  name HG2* ) 3.70 0.40 1.20
assign (resid  12 and  name HG2* )(resid  15 and  name HG1* ) 4.00 0.50 5.50
assign (resid  12 and  name HG2* )(resid  15 and  name HG2* ) 4.00 0.80 5.50
assign (resid  15 and  name HG1* )(resid  16 and  name HD2   ) 4.00 0.20 3.10
assign (resid  15 and  name HG** )(resid  16 and  name HE1   ) 4.00 0.40 5.10
assign (resid  15 and  name HG2* )(resid  18 and  name HN    ) 4.00 0.45 9.10

set echo=true end
set wrnlev=5 end
```

Dihedral Constraints:

```
remarks: dih_v10
remarks: Dihedral Constraints Table for Pen3,15 Nle7 Et-1
remarks: edited by GML 7-17-93
remarks: Version 1.0
restraints dihedral

 assign ( resid 1 and name n )                                  {psi 1}
        ( resid 1 and name ca)
        ( resid 1 and name c )
        ( resid 2 and name n ) 80.0 150.0 40.0 2

 assign ( resid 1 and name c )                                  {phi 2}
        ( resid 2 and name n )
        ( resid 2 and name ca)
        ( resid 2 and name c ) 20.0 -140.0 60.0 2

 assign ( resid 2 and name n )                                  {psi 2}
        ( resid 2 and name ca)
        ( resid 2 and name c )
        ( resid 3 and name n ) 40.0 150.0 40.0 2
```

```
assign ( resid 2 and name c )                    {phi 3}
       ( resid 3 and name n )
       ( resid 3 and name ca)
       ( resid 3 and name c ) 30.0 -160.0 40.0 2

assign ( resid 4 and name c )                    {phi 4}
       ( resid 5 and name n )
       ( resid 5 and name ca)
       ( resid 5 and name c ) 40.0 -160.0 40.0 2

assign ( resid 5 and name n )                    {psi 5}
       ( resid 5 and name ca)
       ( resid 5 and name c )
       ( resid 6 and name n ) 50.0 150.0 40.0 2

assign ( resid 5 and name c )                    {phi 6}
       ( resid 6 and name n )
       ( resid 6 and name ca)
       ( resid 6 and name c ) 100.0 -20.0 80.0 2

assign ( resid 6 and name n )                    {psi 6}
       ( resid 6 and name ca)
       ( resid 6 and name c )
       ( resid 7 and name n ) 100.0 20.0 60.0 2

assign ( resid 6 and name c )                    {phi 7}
       ( resid 7 and name n )
       ( resid 7 and name ca)
       ( resid 7 and name c ) 40.0 -20.0 80.0 2

assign ( resid 7 and name n )                    {psi 7}
       ( resid 7 and name ca)
       ( resid 7 and name c )
       ( resid 8 and name n ) 30.0 20.0 60.0 2

assign ( resid 8 and name n )                    {psi 8}
       ( resid 8 and name ca)
       ( resid 8 and name c )
       ( resid 9 and name n ) 80.0 150.0 60.0 2

assign ( resid 8 and name c )                    {phi 9}
```

```
        ( resid 9 and name n )
        ( resid 9 and name ca)
        ( resid 9 and name c ) 150.0 -20.0 80.0 2

assign ( resid 9 and name n )                    {psi 9}
        ( resid 9 and name ca)
        ( resid 9 and name c )
        ( resid 10 and name n) 150.0 10.0 60.0 2

assign ( resid 9 and name c )                    {phi 10}
        ( resid 10 and name n )
        ( resid 10 and name ca)
        ( resid 10 and name c ) 60.0 -20.0 80.0 2

assign ( resid 10 and name n )                   {psi 10}
        ( resid 10 and name ca)
        ( resid 10 and name c )
        ( resid 11 and name n ) 100.0 20.0 60.0 2

assign ( resid 10 and name c )                   {phi 11}
        ( resid 11 and name n )
        ( resid 11 and name ca)
        ( resid 11 and name c ) 150.0 -20.0 80.0 2

assign ( resid 11 and name n )                   {psi 11}
        ( resid 11 and name ca)
        ( resid 11 and name c )
        ( resid 12 and name n ) 100.0 20.0 60.0 2

assign ( resid 11 and name c )                   {phi 12}
        ( resid 12 and name n )
        ( resid 12 and name ca)
        ( resid 12 and name c ) 150.0 -10.0 60.0 2

assign ( resid 12 and name n )                   {psi 12}
        ( resid 12 and name ca)
        ( resid 12 and name c )
        ( resid 13 and name n ) 150.0 10.0 50.0 2

assign ( resid 12 and name c )                   {phi 13}
        ( resid 13 and name n )
```

```
        ( resid 13 and name ca)
        ( resid 13 and name c ) 150.0 -20.0 80.0 2

assign ( resid 13 and name n )                          {psi 13}
        ( resid 13 and name ca)
        ( resid 13 and name c )
        ( resid 14 and name n ) 80.0 20.0 60.0 2

assign ( resid 13 and name c )                          {phi 14}
        ( resid 14 and name n )
        ( resid 14 and name ca)
        ( resid 14 and name c ) 100.0 -20.0 80.0 2

assign ( resid 14 and name n )                          {psi 14}
        ( resid 14 and name ca)
        ( resid 14 and name c )
        ( resid 15 and name n ) 150.0 20.0 60.0 2

assign ( resid 15 and name n )                          {psi 15}
        ( resid 15 and name ca)
        ( resid 15 and name c )
        ( resid 16 and name n ) 60.0 20.0 60.0 2

assign ( resid 15 and name c )                          {phi 16}
        ( resid 16 and name n )
        ( resid 16 and name ca)
        ( resid 16 and name c ) 40.0 -20.0 80.0 2

assign ( resid 16 and name n )                          {psi 16}
        ( resid 16 and name ca)
        ( resid 16 and name c )
        ( resid 17 and name n ) 40.0 0.0 60.0 2

assign ( resid 16 and name c )                          {phi 17}
        ( resid 17 and name n )
        ( resid 17 and name ca)
        ( resid 17 and name c ) 40.0 0.0 80.0 2

assign ( resid 17 and name n )                          {psi 17}
        ( resid 17 and name ca)
        ( resid 17 and name c )
```

```
        ( resid 18 and name n ) 10.0 20.0 60.0 2

assign  ( resid 17 and name c )                          {phi 18}
        ( resid 18 and name n )
        ( resid 18 and name ca)
        ( resid 18 and name c ) 40.0 -20.0 80.0 2

assign  ( resid 20 and name n )                          {psi 20}
        ( resid 20 and name ca)
        ( resid 20 and name c )
        ( resid 21 and name n ) 40.0 150.0 60.0 2


scale=$kcdi
end
```

APPENDIX E: X-PLOR PROTOCOLS

NOTE: Part 1 (sequence builder) and Part 2 (patching protocols/coordinate builder) corresponds to AcNMe-14mer models. Parts 3-5 (1st-3rd dynamics runs) correspond to Pen-1 models, but modified to include loops and remove redundant distance constraint classes. Trial categories have also been removed from the macro files. 1st stage protocol patches the disulfide bonds. 2nd stage protocol begins with the patched disulfide bonds. 3rd stage protocol (refinement stages) employs a longer high temp dynamics simulation (16ps at 1000K). Also applies Lennard-Jones potentials during the final energy minimization.

PART 1: SEQUENCE BUILDER

```
remarks        file generate/NMeET.inp
remarks        Generate structure file and hydrogens for NMeET

topology @topallhdg_a1.pro  end          {read topology file}
                                         {defines residues/atoms}

parameter                                {read parameter file}
  @parallhdg_a1.pro                      {defines dihedrals/bonds}
  improper ct ct ha ct 250.0 0 0.0       {  impropers/angles/charges}
end

segment                                  {sequence builder}
  name="NMeET"
  chain
              LINK PEPT   HEAD - *  TAIL + *   END
                          first prop tail + pro end
remark                    first nter tail + * end      {N-term capped}
            last cter head - * end
  sequence ACE ASP ALA GLU ALA VAL TYR PHE ALA HIS LEU ASP ILE ILE
TRP end
 end
end

flags exclude vdw elec end
```

```
write structure output=NMeET.psf end                    {writes protein structure file}
stop
```

PART 2: PATCHING PROTOCOL/COORDINATE BUILDER

```
remarks  file  temp_nme.inp
remarks  Generates a "template" coordinate set for the NMeET analog.
remarks  This produces an arbitrary extended conformation with
remarks  ideal geometry.
remarks
remarks  Author: Axel T. Brunger
remarks   modified by GML 1-1-99

evaluate ($numgen=5)                                    {number of structures created}
evaluate ($count=0)                                     {initialize count}
evaluate ($rootname="NMeET")                            {initialize rootname}
evaluate ($psfin=$rootname +".psf")

parameter @parallhdg_a1.pro  end                        {*Read parameters.*}
topology @topallhdg_a1.pro  end                         {Read topology files}

while ($count < $numgen) loop big                       {define main loop}
                structure reset end                     {reset structure files}
                coordinates initialize end              {reset coordinate files}
                evaluate ($count=$count+1)              {counter}

                evaluate ($pdbout=$rootname +"_"+ encode($count) +".pdb")

structure @@$psfin end                                  {read psf}

vector ident (x) ( all )                                {initialize vectors}
vector do (x=x/10.) ( all )
vector do (y=random(0.5) ) ( all )
vector do (z=random(0.5) ) ( all )

vector do (fbeta=50) (all)              {*Friction coefficient, in 1/ps.*}
vector do (mass=100) (all)              {*Heavy masses, in amus.*}
```

```
parameter
  nbonds
    cutnb=5.5 rcon=20. nbxmod=-2 repel=0.9  wmin=0.1 tolerance=1.
    rexp=2 irexp=2 inhibit=0.25
  end
end

flags exclude * include bond angle vdw end
minimize powell nstep=50  nprint=10 end                    {first energy minimization}
flags include impr end
minimize powell nstep=50 nprint=10 end                     {second energy min.}

dynamics  verlet                                           {initial dynamics run at 300K}
  nstep=50  timestep=0.001 iasvel=maxwell  firsttemp= 300.
  tcoupling = true  tbath = 300.   nprint=50  iprfrq=0
end

parameter
  nbonds
    rcon=2. nbxmod=-3 repel=0.75
  end
end

minimize powell nstep=100 nprint=25 end                    {third energy min.}

dynamics  verlet                                           {second dynamics run at 300K}
  nstep=500  timestep=0.005 iasvel=maxwell  firsttemp= 300.
  tcoupling = true  tbath = 300.   nprint=100  iprfrq=0
end

flags exclude vdw elec end
vector do (mass=1.) ( name h* )
hbuild selection=( name h* ) phistep=360 end
flags include vdw elec end

minimize powell nstep=200 nprint=50 end                    {final energy min.}

topology

 presidue NME                                              {N-methyl patching statement}
              modify ATOM N              TYPE=NH1  CHARGE=-0.360 END
```

```
                add atom CN   charge=-0.3    type=CT end
                add atom HN1 charge=0.1     type=HA end
                add atom HN2 charge=0.1     type=HA end
                add atom HN3 charge=0.1     type=HA end
                delete atom HN  end

                bond N CN     bond CN HN1 bond CN HN2 bond CN HN3
                angles N CN HN1       angles N CN HN2     angles N CN HN3
                angles HN1 CN HN2   angles HN1 CN HN3  angles HN2 CN HN3
 end
end

patch NME                                              {apply patching statement}
  reference=nil=( resid 13 )                           { define patched residue}
end

write coordinates output=$pdbout end                   {*Write coordinates.*}
end loop big                                           {end main loop}
stop
```

PART 3: 1ST STAGE DYNAMICS SIMULATION

```
remarks:   file XPLOR3.1 com-file
remarks:   1st stage dynamics annealing starting from a non-cyclic
remarks:   structure, ends with a patched cyclic structure

evaluate ($numgen=10)                          {number of structures generated}
evaluate ($count=0)                                {initialize count}
evaluate ($rootname="rhelix")                      {initialize rootname}
evaluate ($psfin=$rootname + "xplor.psf")          {define input psf}
evaluate ($psfout=$rootname + "_1.psf")            {define output psf}

parameter  @parallhdg_a1.pro  end                  {initialize parameter files}
topology   @topallhdg_a1.pro  end                  {initialize topology files}

while ($count < $numgen) loop main                 {start MAIN LOOP}
                structure reset end
                {reset structure file}
                coordinates initialize end         {reset coordinate files}
```

```
            evaluate ($count = $count+1)                {counter}

            evaluate ($pdbin=$rootname + "_" + encode($count) + ".pdb")
            evaluate ($pdbout=$rootname+ "_" + encode($count) + "a.pdb")

structure @@psfin end                       {read input psf}
coordinates @@pdbin.pdb                     {read input pdb}

evaluate ($knoe  = 10.0)                    {define initial NOE constant}
evaluate ($asym  =  0.1)                    {define initial asymptote constant}
evaluate ($asyt  =  0.1)
evaluate ($kcdi  =  0.1)                    {define initial dihedral constant}
flags
 exclude elec dihe
end

noe                                         {define distance constraints weights}
 reset                                      { set below/ switch parameters from}
 nrestraints = 1000                         {  soft to square energy wells}
 ceiling = 100                              {  different for each class}

class cat1 @pen1_cat1.tbl
 averaging  cat1 r-6
 potential  cat1 soft
 scale      cat1 $knoe
 sqoffset   cat1 0.0
 sqconstant cat1 1.0
 sqexponent cat1 2
 soexponent cat1 1
 asymptote  cat1 $asym
 rswitch    cat1 0.8

class cat2 @pen1_cat2.tbl
 averaging  cat2 r-6
 potential  cat2 soft
 scale      cat2 $knoe
 sqoffset   cat2 0.0
 sqconstant cat2 1.0
 sqexponent cat2 2
 soexponent cat2 1
 asymptote  cat2 $asym
```

```
  rswitch    cat2 0.7

class cat3 @pen1_cat3.tbl
  averaging  cat3 r-6
  potential  cat3 soft
  scale      cat3 $knoe
  sqoffset   cat3 0.0
  sqconstant cat3 1.0
  sqexponent cat3 2
  soexponent cat3 1
  asymptote  cat3 $asym
  rswitch    cat3 0.6

class cat4 @pen1_cat4.tbl
  averaging  cat4 r-6
  potential  cat4 soft
  scale      cat4 $knoe
  sqoffset   cat4 0.0
  sqconstant cat4 1.0
  sqexponent cat4 2
  soexponent cat4 1
  asymptote  cat4 $asyt
  rswitch    cat4 0.5
end

@pen1.dih                                   {define dihedral constraints}

evaluate ($rcon  = 0.0002)
parameters
  nbonds
    wmin =  0.01
    cutnb =  4.5
    tolerance 0.3
    repel=   1.0
    rexp  = 2
    irex  = 2
    rcon=$rcon
  end
end

flags exclude noe cdih end
```

```
mini powell nstep=200 drop= 10 end                {initial energy minimization}
flags include noe cdih end                        { $rcon set to 0.0002}

evaluate ($rcon  = 0.002)
parameters
 nbonds
   wmin =  0.01
   cutnb =  4.5
   tolerance 0.3
   repel=   1.0
   rexp  = 2
   irex  = 2
   rcon=$rcon
 end
end

flags exclude noe cdih end
mini powell nstep=200 drop= 10 end                {2nd energy minimization}
flags include noe cdih end                        { $rcon set to 0.002}

vector do (mass  =  10.0) (all)                   {initialize vectors}
vector do (fbeta = 100.0) (all)
evaluate  ($bath  = 1000.01)                      {define high temperature}
constraints interaction (all) (all)
 weights
   angles   0.4
   impropers 0.1
 end
end

shake                                             {shake algorithm}
 reference = parameter bonds (all) (all)
 tolerance = 1.0e-2
end

remarks:  In loop stg1 dihedral weights ($kcdi) are increased while knoe values remain
remarks:          low to favor known local geometry in the starting structure.

while ( $kcdi < 0.6 ) loop stg1                   {Dynamics LOOP 1}
evaluate ( $kcdi = min(0.6, $kcdi + 0.1))
dynamics verlet                                   { 0.05ps dynamics}
```

```
    nstep=100
    timestep=0.0005
    iasvel=current
    tcoupling=true
    tbath=$bath
    nprint=50
    iprfrq=100
    ntrfr=0
end
end loop stg1                                          {end LOOP STG1}

remarks:  the $kcdi values are then reduced to 15% of their maximal
remarks:  values for 10 ps of 1000  dynamics with fully scaled knoe values

evaluate ( $kcdi=0.15 )
noe                                                    {re-define NOE weights}
scale cat1 100
scale cat2 60
scale cat3 40
scale cat4 25
end

dynamics  verlet                                       {Second dynamics run}
    nstep=100                                          {  0.2ps dynamics}
    timestep=0.002
    iasvel=maxwell
    firsttemperature = $bath
    tcoupling = true
    tbath = $bath
    ntrfr = 999999999
    nprint=50
    iprfrq=100
end

evaluate ($kcdi=0.05)
while ($rcon  < 0.1) loop stg2                         {Dynamics LOOP STG2}
  evaluate ($asym  = min(1.0, $asym  + 0.10))
  noe asymp all $asym  end
  evaluate ($kcdi  = min(0.6, $kcdi  + 0.10))
  evaluate ($rcon  = min(0.10001, $rcon  * 1.3))
  parameters
```

```
    nbonds repel 1.0 rcon $rcon  end
  end
  dynamics  verlet                                    { 0.4ps dynamics}
    nstep=200
    timestep=0.002
    iasvel=current
    tcoupling = true
    tbath = $bath
    nprint=50
    iprfrq=100
    ntrfr = 0
  end
end loop stg2                                         {end LOOP STG2}

noe                                    {switch from soft to square energy well}
 potential cat1 square
 potential cat2 square
 potential cat3 square
end

evaluate ($kcdi=0.05)
while ($rcon  < 0.1) loop stg3                        {Dynamics LOOP STG3}
  evaluate ($asyt  = min(1.0, $asyt  + 0.10))         { asym/asyt slopes increased}
  noe asymp all $asyt  end
  evaluate ($kcdi  = min(0.6, $kcdi  + 0.10))         {dihedral weights increased}
  evaluate ($rcon  = min(0.10001, $rcon  * 1.3))      {rcon weights increased}
  parameters
    nbonds repel 1.0 rcon $rcon  end
  end
  dynamics  verlet                                    { 0.4ps dynamics run}
    nstep=200
    timestep=0.002
    iasvel=current
    tcoupling = true
    tbath = $ bath
    nprint=50
    iprfrq=100
    ntrfr = 0
  end
end loop stg3                                         {end LOOP STG3}
```

```
noe
potential cat4 square          {switch from soft to square energy well}
end

parameters                     { vanderWaal radii are reduced to 80% when the }
 nbonds                        { kvdw value is increased}
  repel 0.8
  rconst 4
  wmin 1.5
 end
end

energy end
if ($grad > 100) then
 mini powell nstep=25 drop=50 end                    {energy minimization}
end if

while ($bath  gt 300) loop stage4                    {Dynamics LOOP STG4}
 dynamics  verlet                          { slow cooling with full value square well }
   nstep=60                                { noe and torsion constraints}
   timestep=0.002                          {  0.12ps dynamics each round}
   iasvel=current
   tcoupling = true
   tbath = $bath
   nprint=30 iprfrq = 30
   ntrfrq = 0
 end
 evaluate ($bath  = $bath  - 50)
end loop stage4                                      {end LOOP STG4}

shake reset end
constraints interaction (all) (all)
 weights
  angles   1.0
  impropers 1.0
 end
end

noe                                                  {redefine NOE weights for }
scale cat1 50                                        { last energy min, half value}
scale cat2 30                                        { wrt to dynamics run}
```

```
scale cat3 20
scale cat4 10
end

mini powell nstep=200 drop = 10.0 end                    {final energy minimization}
topology

  presidue disu                                          {disulfide patching statement}
              group
                 delete     atom 1HG1   end
                 modify     atom 1CB    charge= 0.20 end
                 modify     atom 1SG    charge=-0.20 end
              group
                 delete     atom 2HG1   end
                 modify     atom 2CB    charge= 0.20 end
                 modify     atom 2SG    charge=-0.20 end

  add bond 1SG 2SG
  add angle  1CB 1SG 2SG
  add angle  1SG 2SG 2CB
  add ic 1CA 1CB 1SG 2SG  0.0 0.0   180.0  0.0 0.0
  add ic 1CB 1SG 2SG 2CB  0.0 0.0   180.0  0.0 0.0
  add ic 1SG 2SG 2CB 2CA  0.0 0.0   180.0  0.0 0.0
end

patch disu                                               {apply patching statement}
   reference=1=( resid 1 ) reference=2=( resid 15 )
end
patch disu
   reference=1=( resid 3 ) reference=2=( resid 11 )
end

write coordinates output=$pdbout end                     {write coordinate file}
print threshold=0.3 noe                                  {print NOE violations}
end loop main                                            {end MAIN LOOP}

write structure output=$psfout end                       {Write new psf file}
stop
```

PART 4: 2ND STAGE DYNAMICS SIMULATION

```
remarks:    continuation of XPLOR3.1 file(=stage 2), also used as is for
remarks:    constrainted dynamics beginning with previously cyclized
remarks:    structures.
remarks:    this comfile is also used for the final refinement which includes
remarks:    additional "low bounds only" constraints that represent nOes
remarks:    predicted in noesy simulations but unobserved in any experimental
remarks:    NOESY spectrum.  This is called stage 2.

evaluate ($numgen=10)                          {number of structures generated}
evaluate ($count=0)                                {initialize count}
evaluate ($rootname="rhelix")                      {initialize rootname}
evaluate ($psfin=$rootname + "_1.psf")             {define input psf}
evaluate ($psfout=$rootname + "_2.psf")            {define output psf}

parameter  @parallhdg_a1.pro  end                  {initialize parameter files}
topology   @topallhdg_a1.pro  end                  {initialize topology files}

while ($count < $numgen) loop main                 {start MAIN LOOP}
              structure reset end                  {reset structure file}
              coordinates initialize end           {reset coordinate files}
              evaluate ($count = $count+1)         {counter}

              evaluate ($pdbin=$rootname + "_" + encode($count) + "a.pdb")
              evaluate ($pdbout=$rootname+ "_" + encode($count) + "b.pdb")

nbfix s s 462 13.6 462 13.6  end
structure @@psfin end                          {read input psf}
coordinates @@pdbin.pdb                        {read input pdb}

evaluate ($knoe  = 10.0)                       {define initial NOE weight}
evaluate ($asym  = 0.35)                       {define initial well slopes}
evaluate ($kcdi  =  0.1)                       {define initial dihedral weight}
flags
 exclude elec dihe
end

noe
 reset
 nrestraints = 1000
```

```
  ceiling = 100

class cat1 @pen1_cat1.tbl                    {define distance constraints}
  averaging  cat1 r-6
  potential  cat1 soft
  scale      cat1 $knoe
  sqoffset   cat1 0.0
  sqconstant cat1 1.0
  sqexponent cat1 2
  soexponent cat1 1
  asymptote  cat1 $asym
  rswitch    cat1 0.8

class cat2 @pen1_cat2.tbl
  averaging  cat2 r-6
  potential  cat2 soft
  scale      cat2 $knoe
  sqoffset   cat2 0.0
  sqconstant cat2 1.0
  sqexponent cat2 2
  soexponent cat2 1
  asymptote  cat2 $asym
  rswitch    cat2 0.7

class cat3 @pen1_cat3.tbl
  averaging  cat3 r-6
  potential  cat3 soft
  scale      cat3 $knoe
  sqoffset   cat3 0.0
  sqconstant cat3 1.0
  sqexponent cat3 2
  soexponent cat3 1
  asymptote  cat3 $asym
  rswitch    cat3 0.6

class cat4 @pen1_cat4.tbl
  averaging  cat4 r-6
  potential  cat4 soft
  scale      cat4 $knoe
  sqoffset   cat4 0.0
  sqconstant cat4 1.0
```

```
 sqexponent cat4 2
 soexponent cat4 1
 asymptote  cat4 $asym
 rswitch    cat4 0.5
end

@pen1.dih                                   {define dihedral constraints}

evaluate ($rcon  = 0.002)
parameters
 nbonds
  wmin =  0.01
  cutnb =  4.5
  tolerance 0.3
  repel=   1.0
  rexp  = 2
  irex  = 2
  rcon=$rcon
 end
end

flags exclude noe cdih end
mini powell nstep=150 drop= 10 end        {initial energy minimization}
flags include noe cdih end

vector do (mass  =  10.0) (all)
vector do (fbeta = 100.0) (all)
evaluate  ($bath  = 1000.01)
constraints interaction (all) (all)
 weights
  angles   0.4
  impropers 0.1
 end
end
shake
 reference = parameter bonds (all) (all)
 tolerance = 1.0e-2
end

set seed=72494620 end                          {randomizer seed number}
vector do (vx=maxwell(600)) (all)       {in some structures, noe deviations }
```

```
vector do (vy=maxwell(600)) (all)       {are lower and Enoe is insufficient }
vector do (vz=maxwell(600)) (all)       {to warm system                     }

while ( $kcdi < 0.6 ) loop stg1                 {Dynamics LOOP STG1}
evaluate ( $kcdi = min(0.6, $kcdi + 0.1))
dynamics verlet
    nstep=100                                   {as in previous non-cyclic example this    }
    timestep=0.001                              {dynamics course enforces the torsion      }
    iasvel=current                              {constraints while knoe is at a modest level}
    tcoupling=true                              { 0.1ps dynamics run}
    tbath=$bath
    nprint=50
    iprfrq=100
    ntrfr=0
end
end loop stg1                                   {end LOOP STG1}

evaluate ( $kcdi=0.30 )
noe                                             {reset NOE weights}
scale cat1 100
scale cat2 60
scale cat3 40
scale cat4 25
end

dynamics  verlet                                {2nd dynamics run}
    nstep=2000                                  { 4ps dynamics run at high temp}
    timestep=0.002
    iasvel=maxwell
    firsttemperature = $bath
    tcoupling = true
    tbath = $bath
    ntrfr = 999999999
    nprint=50
    iprfrq=100
end

evaluate ($kcdi=0.05)
while ($rcon  < 0.2) loop stg2                  {Dynamics LOOP STG2}
  evaluate ($asym  = min(0.6, $asym  + 0.05))
  noe asymp all $asym  end
```

```
  evaluate ($kcdi  = min(0.6, $kcdi  + 0.05))
  evaluate ($rcon  = min(0.20001, $rcon  * 1.4))
  parameters
   nbonds repel 0.9 rcon $rcon  end
  end
  dynamics  verlet                          { 0.2ps dynamics run}
    nstep=100
    timestep=0.002
    iasvel=current
    tcoupling = true
    tbath = $bath
    nprint=50
    iprfrq=100
    ntrfr = 0
  end
end loop stg2                               {end LOOP STG2}

noe                                         {reset energy potentials}
  potential cat1 square                     { from soft to square wells}
  potential cat2 square
  potential cat3 square
potential cat4 square
end

parameters
  nbonds
    repel 0.8
    rconst 4
    wmin 1.5
  end
end
energy end
if ($grad > 100) then
  mini powell nstep=25 drop=50 end          {energy minimization}
end if
evaluate ($kcdi=0.30)

while ($bath  gt 300) loop stage3           {Dynamics LOOP STG3}
  dynamics  verlet                          {simulated annealing steps}
    nstep=60                                {  0.12ps dynamics/each step}
    timestep=0.002 {ps}
```

```
    iasvel=current
    tcoupling = true
    tbath = $bath
    nprint=30 iprfrq = 30
    ntrfrq = 0
  end
  evaluate ($bath  = $bath  - 50)
end loop stage3                          {end LOOP STG3}

shake reset end
constraints interaction (all) (all)
  weights
    angles   1.0
    impropers 1.0
  end
end
evaluate($kcdi=0.15)

noe                                      {reset NOE weights for final}
scale cat1 50                            { energy minimization}
scale cat2 30
scale cat3 20
scale cat4 10
end

mini powell nstep=200 drop = 10.0 end    {final energy minimization}
write coordinates output=$pdbout  end    {write coordinates}
print threshold=0.3 noe                  {print NOE violations}
end loop main                            {end MAIN LOOP}

write structure output=$psfout  end      {write structure file: optional}
stop
```

PART 5: 3RD STAGE DYNAMICS SIMULATION

```
remarks:    continuation of XPLOR3.1 file(=stage 3), also used as is for
remarks:    constrainted dynamics beginning with previously cyclized
remarks:    structures.
remarks:    this comfile is also used for the final refinement which includes
```

```
remarks:    additional "low bounds only" constraints that represent nOes
remarks:    predicted in noesy simulations but unobserved in any experimental
remarks:    NOESY spectrum. This is called stage 3.
remarks:    NOTE: contains Lennard-Jones Potentials

evaluate ($numgen=10)                          {number of structures generated}
evaluate ($count=0)                                {initialize count}
evaluate ($rootname="rhelix")                      {initialize rootname}
evaluate ($psfin=$rootname + "_2.psf")             {define input psf}
evaluate ($psfout=$rootname + "_3.psf")            {define output psf}

parameter  @parallhdg_a1.pro  end
topology   @topallhdg_a1.pro  end

while ($count < $numgen) loop main                 {start MAIN LOOP}
              structure reset end                  {reset structure file}
              coordinates initialize end           {reset coordinate files}
              evaluate ($count = $count+1)         {counter}

              evaluate ($pdbin=$rootname + "_" + encode($count) + "b.pdb")
              evaluate ($pdbout=$rootname+ "_" + encode($count) + "c.pdb")


nbfix s s 462 13.6 462 13.6  end
structure @@psfin end                      {read input psf}
coordinates @@pdbin.pdb                    {read input pdb}

evaluate ($knoe  = 10.0)                   {define initial NOE weight}
evaluate ($asym  = 0.35)                   {define initial well slopes}
evaluate ($kcdi  =  0.1)                   {define initial dihedral weight}
flags
 exclude elec dihe
end

noe                                {define distance constraints}
 reset
 nrestraints = 1000
 ceiling = 100

class cat1 @pen1_cat1.tbl
 averaging  cat1 r-6
```

```
  potential  cat1 soft
  scale      cat1 $knoe
  sqoffset   cat1 0.0
  sqconstant cat1 1.0
  sqexponent cat1 2
  soexponent cat1 1
  asymptote  cat1 $asym
  rswitch    cat1 0.8

class cat2 @pen1_cat2.tbl
  averaging  cat2 r-6
  potential  cat2 soft
  scale      cat2 $knoe
  sqoffset   cat2 0.0
  sqconstant cat2 1.0
  sqexponent cat2 2
  soexponent cat2 1
  asymptote  cat2 $asym
  rswitch    cat2 0.7

class cat3 @pen1_cat3.tbl
  averaging  cat3 r-6
  potential  cat3 soft
  scale      cat3 $knoe
  sqoffset   cat3 0.0
  sqconstant cat3 1.0
  sqexponent cat3 2
  soexponent cat3 1
  asymptote  cat3 $asym
  rswitch    cat3 0.6

class cat4 @pen1_cat4.tbl
  averaging  cat4 r-6
  potential  cat4 soft
  scale      cat4 $knoe
  sqoffset   cat4 0.0
  sqconstant cat4 1.0
  sqexponent cat4 2
  soexponent cat4 1
  asymptote  cat4 $asym
  rswitch    cat4 0.5
```

```
end

@pen1.dih                                   {define dihedral constraints}

evaluate ($rcon  = 0.002)
parameters
 nbonds
   wmin =  0.01
   cutnb =  4.5
   tolerance 0.3
   repel=   1.0
   rexp  = 2
   irex  = 2
   rcon=$rcon
 end
end

flags exclude noe cdih end
mini powell nstep=150 drop= 10 end          {initial energy minimization}
flags include noe cdih end

vector do (mass  =  10.0) (all)
vector do (fbeta = 100.0) (all)
evaluate  ($bath  = 1000.01)
constraints interaction (all) (all)
 weights
   angles   0.4
   impropers 0.1
 end
end
shake
 reference = parameter bonds (all) (all)
 tolerance = 1.0e-2
end

set seed=72494620 end                       {randomizer seed number}
vector do (vx=maxwell(600)) (all)
vector do (vy=maxwell(600)) (all)
vector do (vz=maxwell(600)) (all)

while ( $kcdi < 0.6 ) loop stg1             {Dynamics LOOP STG1}
```

```
evaluate ( $kcdi = min(0.6, $kcdi + 0.1))
dynamics verlet                          { 0.1ps dynamics run}
    nstep=100
    timestep=0.001
    iasvel=current
    tcoupling=true
    tbath=$bath
    nprint=50
    iprfrq=100
    ntrfr=0
end
end loop stg1                            {end LOOP STG1}

evaluate ( $kcdi=0.30 )

noe                                      {reset NOE weights}
scale cat1 100
scale cat2 60
scale cat3 40
scale cat4 25
end

dynamics  verlet                         {2nd dynamics protocol}
    nstep=8000                           { 16ps dynamics run at high temp}
    timestep=0.002
    iasvel=maxwell
    firsttemperature = $bath
    tcoupling = true
    tbath = $bath
    ntrfr = 999999999
    nprint=50
    iprfrq=100
end
evaluate ($kcdi=0.05)

while ($rcon  < 0.2) loop stg2           {Dynamics LOOP STG2}
  evaluate ($asym  = min(0.8, $asym  + 0.08))
  noe asymp all $asym  end
  evaluate ($kcdi  = min(0.4, $kcdi  + 0.05))
  evaluate ($rcon  = min(0.20001, $rcon  * 1.4))
  parameters
```

```
    nbonds repel 0.9 rcon $rcon  end
   end
   dynamics  verlet                          {0.2ps dynamics run}
     nstep=100
     timestep=0.002
     iasvel=current
     tcoupling = true
     tbath = $bath
     nprint=50
     iprfrq=100
     ntrfr = 0
   end
end loop stg2                                {end LOOP STG2}

noe                                          {reset energy potentials from}
  potential cat1 square                      { soft to square wells}
  potential cat2 square
  potential cat3 square
  potential cat4 square
end

parameters
  nbonds
    repel 0.8
    rconst 4
    wmin 1.5
  end
end
energy end

if ($grad > 100) then
  mini powell nstep=25 drop=50 end
end if
noe
scale cat5 5
scale cat6 0.1
end
evaluate ($kcdi=0.20)

while ($bath  gt 300) loop stage3            {Dynamics LOOP STG3}
  dynamics  verlet                           { simulated annealing}
```

```
    nstep=60                                { 0.12ps dynamics/each step}
    timestep=0.002 {ps}
    iasvel=current
    tcoupling = true
    tbath = $bath
    nprint=30 iprfrq = 30
    ntrfrq = 0
  end
  evaluate ($bath  = $bath  - 50)
end loop stage3                             {end LOOP STG3}

shake reset end
constraints interaction (all) (all)
  weights
    angles   1.0
    impropers 1.0
  end
end
parameters                                  {Lennard-Jones Potential Parameters}
  nbonds                                    { applied during last energy min.}
    wmin   =  1.50
    cutnb  =  8.50
    tolerance 0.50
    ctofnb =  7.50
    ctonnb =  6.50
    repel  =  0.00
    vswitch
  end
end
evaluate($kcdi=0.10)

noe                                         {reset NOE weights for final}
scale cat1 50                               { energy minimization}
scale cat2 30
scale cat3 20
scale cat4 10
end

mini powell nstep=500 drop = 10.0 end       {long energy min. protocol}
write coordinates output=$pdbout  end       {write coordinates}
print threshold=0.3 noe                     {print NOE violations}
```

```
end loop main                                   {end MAIN LOOP}

write structure output=$psfout  end             {write structure file: optional}
stop
```

PART 6: ANALYSIS PROTOCOL (PHI/PSI REPORT)

```
remarks:      file phipsi.com -- used to report phi and psi dihedral angles
remarks:      for pdb coordinate files. For use with X-PLOR ver 3.1
remarks:      Created 3/8/94 by GM Lee
remarks:      NOTE:  user needs to customize the program by changing
remarks:                  values (for numgen, totres, rootname, and psf/pdb
remarks:                  suffixes).  If the peptide contains an N-terminal
remarks:                  acyl group (or C-term amide), user needs to remark
remarks:                  the evaluate phi (or evaluate psi) statement at the
remarks:                  beginning (or end) of the program.

set display=phsi3ac2.lst end                              {phi/psi display file}
set precision=5 end                                       {number of sig. figs.}
evaluate ($numgen=60)                                     {number of structures}
evaluate ($totres=21)                                     {number of residues}
evaluate ($rootname="5p")                                 {rootname of pdb/psf}
evaluate ($pdbsuffix="_3ac.pdb")                          {pdb suffix}
evaluate ($psfsuffix="1_3ac.psf")                         {psf suffix}
evaluate ($count=0)                                       {initialize count}

evaluate ($remarks= "   model     residue     phi                              psi")
evaluate ($dash= "-----------------------------------------------")
display $remarks
display $dash
evaluate ($restotal=$totres - 1)
evaluate ($psfname=$rootname + $psfsuffix)
structure @@$psfname end                                  {load psf file (only one)}

while ($count < $numgen) loop big                         {loop for multiple pdbs}
coordinates initialize end                                {initialize coordinates}
 evaluate ($count=$count+1)                               {reassign count}
 evaluate ($pdbfile=$rootname + encode($count) + $pdbsuffix)
```

```
coordinates @@$pdbfile                          {load pdb files}

vector ident (store9)  (name ca)                {loads atoms}

set echo=off end                        {reduces lines in output}
set message=off end                     { file. This is optional}

evaluate ($pre_atom=0)                  {initialize n-1 residue}
evaluate ($pre_pre_atom=0)              {initialize n-2 residue}
evaluate ($resnumber=0)                 {initialize residue number}

for $atom in id (name ca) loop small    {loop for measuring dihedrals}
            if ($resnumber = 1 ) then
              pick dihedral
                (byresidue (id $pre_atom) and name N)
                (byresidue (id $pre_atom) and name CA) {psi for residue 1}
                (byresidue (id $pre_atom) and name C)
                (byresidue (id $atom) and name N)
              geometry
            evaluate ($psi=$result)
            vector show element (resname) (byresidue (id $pre_atom) and name ca)
            evaluate ($resname=$result)
            evaluate ($phi="--START--")
            display  $pdbfile    $resname              $resnumber    $phi
            $psi
            end if

            if ($pre_pre_atom > 0) then     pick dihedral
               (byresidue (id $pre_pre_atom) and name C)
               (byresidue (id $pre_atom) and name N)                {phi}
               (byresidue (id $pre_atom) and name CA)
               (byresidue (id $pre_atom) and name C)
              geometry
              evaluate ($phi=$result)

              pick dihedral
               (byresidue (id $pre_atom) and name N)
               (byresidue (id $pre_atom) and name CA)               {psi}
               (byresidue (id $pre_atom) and name C)
               (byresidue (id $atom) and name N)
               geometry
```

```
                evaluate ($psi=$result)
                vector show element (resname) (byresidue (id $pre_atom) and name ca)
                evaluate ($resname=$result)
                display $pdbfile   $resname $resnumber   $phi          $psi
                end if

                if ($resnumber = $restotal ) then exit small end if
                evaluate ($pre_pre_atom=$pre_atom)      {reassign residue}
                evaluate ($pre_atom=$atom)              {numbers and names}
                evaluate ($resnumber=$resnumber+1)
end loop small                                          {end dihedral loop}

evaluate ($pre_pre_atom=$pre_atom)                      {reassigns residue numbers}
evaluate ($pre_atom=$atom)                              {and names for C-terminal}
evaluate ($resnumber=$resnumber+1)                      {residue}

if ($resnumber = $totres) then
                pick dihedral
                 (byresidue (id $pre_pre_atom) and name C)
                 (byresidue (id $pre_atom) and name N)
                 (byresidue (id $pre_atom) and name CA)  {phi c-term residue}
                 (byresidue (id $pre_atom) and name C)
                 geometry
                evaluate ($phi=$result)
                vector show element (resname) (byresidue (id $pre_atom) and name ca)
                evaluate ($resname=$result)
                evaluate ($psi= "--END--")
                display  $pdbfile          $resname $resnumber          $phi
                $psi
                display $dash
end if

end loop big                                {end structure loop}
stop                                        {terminates program}
```

VITA

Gregory Mitchell Lee

Birthdate: December 16, 1969

Birthplace: San Francisco, California

Education:

1999	Doctor of Philosophy	University of Washington, Seattle, Washington
1991	Bachelor of Science	University of California, Davis, California

Publications and Presentations:

1992 "Herbaceamide, A Chlorinated, N-Acyl Amino Ester from the Marine Sponge, *Dysidea Herbacea*." *Tetrahedron Letters*, **33**(50), 7671-7674.
(with T.F. Molinski.)

1994 "Configurational Analysis of New Furanosesquiterpenes from *Dysidea Herbacea*. Assignment of Absolute Stereochemistry." *Tetrahedron*, **50**(13), 3879-3888.
(with P.A. Searle, N.M. Jamal and T.F. Molinski.)

1994 "β-Structure in Human Amylin and Two Designer β-Peptides: CD and NMR Spectroscopic Comparisons Suggest Soluble β-Oligomers and the Absence of Significant Populations of β-Strand Dimers." *Biochemical and Biophysical Research Communications*, **204**(3), 1088-1095.
(with J. Cort, Z. Liu, S.M. Harris, K.S. Prickett, L.S.L. Gaeta and N.H Andersen.)

1994 "Does the Solid-State Structure of Endothelin-1 Provide Insights Concerning the Solution-State Conformational Equilibrium?" *FEBS Letters*, **355**, 140-146.
(with C. Chen, T.M. Marschner and N.H. Andersen.)

1995 "The Endothelin C-Terminal Signal Fragment: Determinants of the Conformational Equilibrium *IN SITU* and Detached." ***Protein and Peptide Letters***, **1**(4), 215-222.
(with N.H. Andersen and C. Chen.)

1997 "Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging." *J. Am. Chem. Soc.*, **119**, 8547-8561.
(with N.H. Andersen, J.W. Neidigh, S.M. Harris, Z. Liu and H. Tong.)

1998 "Conformational Features of C-Terminal Endothelin-1 Antagonist Analogs." Poster presented at the 39th ENC, Pacific Grove, CA.
(with C. Chen, S.M. Harris and N.H. Andersen.)