

I. INSIGHTS INTO THE DETERMINANTS OF COLLAGEN TRIPLE HELIX
STABILITY

II. INHIBITION OF RNASE A BY ANALOGS OF 3-PRIME-
URIDINEMONOPHOSPHATE

by

Cara Lee Jenkins

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

Doctor of Philosophy
(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2004

A dissertation entitled

I. Insights into the determinants of
collagen triple helix stability.

II. Inhibition of RNase A by
Analogues of 3'-uridinemonophosphate

submitted to the Graduate School of the
University of Wisconsin-Madison
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy

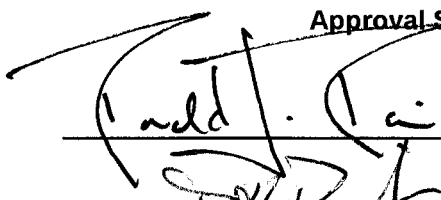
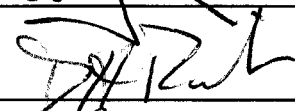

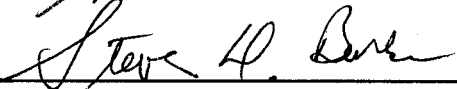

by

Cara Lee Jenkins

Date of Final Oral Examination: June 22, 2004

Month & Year Degree to be awarded: December May August 2004

Approval Signatures of Dissertation Committee

Signature, Dean of Graduate School



I. INSIGHTS INTO THE DETERMINANTS OF COLLAGEN TRIPLE HELIX STABILITY

II. INHIBITION OF RNASE A BY ANALOGS OF 3'-URIDINEMONOPHOSPHATE

Cara Lee Jenkins

Under the supervision of Ronald T. Raines

At the University of Wisconsin–Madison

Collagen, the most abundant protein in animals, is a triple helix composed of Xaa–Yaa–Gly triplet repeats. In natural collagen, the Xaa position is often occupied by proline (Pro), and the Yaa position is often occupied by 4(*R*)-hydroxyproline (Hyp). The three strands are held together by interstrand hydrogen bonds between glycine (Gly) and the Xaa residue of a neighboring strand.

In natural collagen, 3(*S*)-Hydroxyproline (3-Hyp) is found in the rare triplet sequence 3-Hyp–4-Hyp–Gly in natural collagen. Here we show that when it is placed in the Xaa position, 3-Hyp is shown to slightly decrease the thermal stability of triple helical collagen mimics by weakening the interstrand hydrogen bond. Likewise, removing a main-chain–main-chain hydrogen bond from a triple helical collagen mimic drastically reduces its thermal stability.

Replacing the hydroxyl group of Hyp in (Pro-Hyp-Gly)₁₀ with a more electronegative fluoro group had been shown to significantly increase triple helix thermal stability. Here, *O*-acylated Ac-Hyp-OMe residues are shown to have peptide bond cis–trans ratios and backbone dihedral angles are favorable for triple helix formation;

however, the addition of acetyl groups to the Hyp residues in triple helical (Pro-Hyp-Gly)₁₀ slightly decreases triple helix stability, presumably by a steric interference with neighboring strands.

Bicyclic proline analogs, substituted 2-azabicyclo[2.1.1]hexanes (3,5-methanoproline analogs), can help sort out the effects of electronegativity from those of pyrrolidine ring pucker. Little difference was found in the amide cis-trans ratios of H-, OH-, and F-substituted 3,5-methanoproline analogs, in contrast to analogous proline residues. This finding indicates that the role of increasing electronegativity in the 4-position of substituted proline residues is to rigidify the ring and preorganize the peptide backbone into a favorable conformation for triple helix formation.

Ribonuclease A efficiently catalyzes the cleavage of phosphodiester bonds in RNA. 3'-Uridinemonophosphate (3'-UMP) can inhibit this activity. Here, non-natural analogs of 3'-UMP were synthesized and evaluated as inhibitors of wild-type RNase A and its Thr 45→Gly variant. 3'-Nucleotides with the arabino-sugar conformation are more effective than those with ribose sugar moieties at inhibiting wild-type RNase A. In contrast, 3'-nucleotides with ribo-sugars are more effective than those with arabino-sugars at inhibiting T45G RNase A.

Acknowledgements

Any scientific endeavor is an exercise in teamwork. I have been lucky to work with a large number of capable and creative scientists. My advisor, Ron Raines, has provided an environment where creativity and independent thinking are supported and encouraged. I have benefited from his enthusiasm for science and learned a great deal from his prodigious ability in communicating science.

I have also been lucky to work with a variety of collaborators. Grant Krow at Temple University and Eric Eberhardt at Vassar College and their students have figured heavily into two of the projects reported in this dissertation. These collaborations have produced results far better than I could have achieved on my own. In addition, I have collaborated with other students and post-doctoral fellows in the group. Brad Kelemen proposed the nucleotide project, and gave me a great deal of advice and encouragement. Roz Sweeney also worked on the nucleotide project, and helped me learn about enzyme assays. Kim Dickson and Bryan Smith gave me a great deal of help in understanding enzyme kinetics and RNase A behavior.

I have benefited a great deal from innumerable discussions with others working on the collagen project: Drs. Lynn Bretscher, Jonathan Hodges, and Frank Kotch, as well as Eric Benedict and Kim Taylor. Jonathan has been especially gracious with his time and talents in helping me understand some of the finer points of the physical organic approach to understanding collagen stability and structure.

The University of Wisconsin–Madison is a wonderful place to do research, not only because of the excellent instrumentation found here, but also because of the

wonderful people who run the various facilities where the instrumentation is found. Dr. Gary Case, director of the peptide synthesis facility, has taught me a great deal about peptide synthesis, and has been supremely helpful many ways. Dr. Darrell McCaslin, BIFmaster extraordinaire, has taught me in his own unique way about careful experimental design and data analysis. Drs. Ilia Guzei, Martha Vestling and Charlie Fry of the Chemistry Instrument Center have been helpful in gathering and analyzing data from X-ray diffraction, mass spectroscopy, and NMR spectroscopy, respectively.

The members of the Raines lab, past and present, have provided scientific discussions, entertainment, and friendships. I have learned about science and about life from talking and associating with them.

And to save the best for last: I never would have been able to finish the large undertaking of earning a Ph.D degree without the support and encouragement of my friends and family. Caleb has been a wonderful addition to our family, and without the help of a few select babysitters, coming to school every day would have become more than I could handle a long time ago. Rosena Bradford, Miriam Bjork, and Sarah McKay have helped me the past two and a half years to continue my research by taking good care of Caleb during the day. My parents and my parents-in-law have always shown faith in my abilities and interest in my work. They have never questioned my choices or my priorities, for which I am supremely grateful. Too many friends to name here have provided encouragement, support, and much-needed entertainment. David has been there through thick and thin. He has been more supportive than I ever could have dreamed, especially in the difficult endeavor of being a student and a mom.

Table of Contents

Abstract.....	i
Acknowledgements.....	iii
Table of Contents.....	v
List of Figures.....	ix
List of Schemes.....	xii
List of Tables.....	xiii
List of Abbreviations.....	xv

Chapter 1

Insights on the Conformational Stability of Collagen

1.1 Introduction.....	1
1.2 Collagen Mimics.....	3
1.2.1 Structural Studies.....	4
1.2.2 Host-guest Studies.....	9
1.2.3 Tethered Triple-Helical Peptides.....	18
1.2.4 Peptoid Residues.....	24
1.3 4-Substituted Proline Residues.....	28
1.3.1 Hydroxyproline Residues.....	29
1.3.2 Aminoproline Residues.....	32
1.3.3 Fluoroproline Residues.....	33

1.4 Collagen as a Biomaterial.....	39
1.5 Envoi.....	41

Chapter 2

Effect of 3-Hydroxyproline Residues on Collagen Stability.....	43
2.1 Introduction.....	43
2.2 Results and Discussion.....	44
2.2.1 Conformational Stability.....	45
2.2.2 Peptide Bond Isomerization.....	46
2.2.3 Structure of a 3-Hyp Residue.....	47
2.2.4 Steric Effect on Collagen Stability.....	50
2.2.5 Effect of Hydrogen Bonds on Collagen Stability.....	51
2.3 Biological Implications.....	52
2.4 Experimental Section.....	53
NMR Spectra.....	61

Chapter 3

Substituted 2-Azabicyclo[2.1.1]hexanes as Constrained Proline Analogs: Implications for Collagen Stability.....	66
3.1 Introduction.....	66
3.2 Results and Discussion.....	71
3.3 Conclusions.....	75
3.4 Experimental Section.....	76

Chapter 4

O-Acylation of Hydroxyproline Residues: Effect on Peptide Bond Isomerization and

Collagen Stability.....	79
4.1 Introduction.....	79
4.2 Results and Discussion.....	81
4.3 Conclusions.....	90
4.4 Experimental Section.....	91
NMR Spectra.....	97

Chapter 5

Binding of Non-natural 3'-Nucleotides to RNase A.....	113
5.1 Introduction.....	113
5.2 Results.....	114
5.3 Discussion.....	118
5.4 Experimental Section.....	122
NMR Spectra.....	134

Chapter 6

Collagen Mimics Containing Pro-Gly Amide Isosteres: Probing the Importance of Gly-

NH...O=C-Pro Hydrogen Bonds.....	173
6.1 Introduction.....	173
6.2 Results and Discussion.....	175

6.2.1 Synthesis of Fmoc-Tripeptides for Use In Solid Phase Peptide Synthesis.....	175
6.2.2 Solid-Phase Synthesis of Depsipeptides.....	178
6.2.3 Effects of Ester and Alkene Isosteres on Triple Helix Formation and Stability.....	181
6.2.4 Assembly of Heterotrimeric Collagen Mimics.....	187
6.3 Conclusions and Future Directions.....	191
6.4 Experimental Section.....	192
NMR Spectra.....	202
References.....	211

List of Figures

Figure 1.1	Segment of a (Pro-Hyp-Gly) _n triple helix. (a) Ball-and-stick representation indicating 4-hydroxy-L-proline residues and XaaC=O...H-NGly hydrogen bonds. (b) Register of the residues in the three strands of panel a. Atomic coordinates are from (Bella et al., 1994)(PDB entry 1CAG).....	3
Figure 1.2	Distribution of triplets in human collagen types I, II, III, V, XI, IV, VI, VII, VIII, IX, X, and XIII (Ramshaw et al., 1998). Rows and columns refer to the Xaa and Yaa positions of 4040 Gly-Xaa-Yaa triplets. Black boxes indicate triplets found fewer than 3 times (<0.074%), hatched boxes indicate triplets found 3–10 times (0.074–0.25%), and numbers in white boxes represent the occurrence of the most common triplets, rounded to the nearest 0.1%. All Pro residues in the Yaa position are assumed to be hydroxylated. All residues are indicated by their single letter codes, with O indicating Hyp.....	10
Figure 1.3	N- and C-linked collagen mimics.....	20
Figure 1.4	The gauche effect in AcNHCH ₂ CH ₂ F.(O'Hagan et al., 2000) The lobes opposite the C-N and C-F bonds represent σ* antibonding orbitals, which overlap with the σ bonding orbitals of the indicated C-H bonds.....	35
Figure 1.5	Main-chain ω, φ, and ψ torsion angles of a Flp residue. The gauche effect (Fig. 1.4) fixes the pyrrolidine ring pucker (C ^γ -exo). In that ring pucker, the ω, φ, and ψ angles are preorganized at values close to those in the Yaa position of a collagen triple helix (Bretscher et al., 2001). The indicated n→π* interaction contributes to that preorganization.....	36
Figure 2.1	(A) Ortep diagram of crystalline N-(¹³ C ₂ -acetyl)-3(S)-hydroxy-L-proline methyl ester (2.8) drawn with 30% probability ellipsoids. (B) Pyrrolidine ring pucker in crystalline amide 2.8 (left) and crystalline N-acetyl-4(R)-hydroxy-L-proline methyl ester (right) (Panasik et al., 1994).....	48
Figure 2.2	Model of a segment of the triple helix formed by peptide 2.1 (top) and peptide 2.2 (bottom). Each 3-Hyp residue (green), the oxygen of its hydroxyl group (red; labeled), and the steric clash in triple-helical peptide 2.2 (yellow) are indicated.....	51
Figure 2.3	Putative interstrand hydrogen bonds in triple-helical (3-Hyp-4-Hyp-Gly) _n	52

Figure 3.1	Ring puckers in 4-substituted Ac-Pro-OMe. C ^γ -endo pucker is favored when X = H, OH, or F, and Y = H. C ^γ -exo pucker is favored when X = H and Y = OH or F.....	69
Figure 3.2	ORTEP diagrams showing the two crystallographically independent molecules in the unit cell of crystalline Ac-methano-hyp-OMe (3.6), drawn with 30% probability ellipsoids.....	73
Figure 3.3	Superposition of crystalline structures of Ac-methano-hyp-OMe (3.6 , cyan) and Ac-Hyp-OMe (orange) (Panasik et al., 1994).....	73
Figure 3.4	Ball-and-stick diagram showing bond lengths that differ statistically in the two molecules of the X-ray structure of crystalline Ac-methano-hyp-OMe (3.6). The O ₀ ...C ₁ =O ₁ bond lengths and angles are also shown.....	74
Figure 4.1	(A) ORTEP diagrams of crystalline 4.2 (50% probability ellipsoids) and 4.5 (40% probability ellipsoids). (B) Comparison of pyrrolidine ring conformations of crystalline 4.2 and 4.1 (top) (Panasik et al., 1994), and 4.5 and AcProOMe (bottom) (Panasik et al., 1994).....	83
Figure 4.2	van't Hoff plot for the cis-to-trans amide bond isomerization of compounds 4.1 – 4.6 in 1,4-dioxane-d ₈	86
Figure 4.3	MALDI-TOF mass spectrum of peptide 4.9	88
Figure 4.4	(A) Circular dichroism spectra of peptide 9 at 25°C (squares) and 85°C (circles). (B) Thermal denaturation curve for triple helical peptide 9 in 50 mM asetic acid. The squares represent the data points, and the line represents the curve-fit of the data to give T _m = 57.5°C.....	89
Figure 4.5	Segment of a (ProHypGly) _n triple helix. (A) Ball-and-stick model with a Hyp residue highlighted. (B) Space-filling model with an interstrand Hyp)(Pro interaction highlighted. The N-terminus of the triple helical segment is at the top of the figure. Carbon is gray, nitrogen is blue, and oxygen is red. Atomic coordinates are from PDB entry 1CAG (Bella et al., 1994).....	89
Figure 5.1	S- and N-conformation of nucleosides. R = H favors S-conformation; R = OH, F favors N-conformation.....	119
Figure 5.2	3'-UMP bound to wild-type RNase A. The protein backbone is in blue, 3'-UMP is in magenta and the active site residue side chains are in green..	119
Figure 6.1	A segment of a (Pro-Hyp-Gly) triple helix highlighting the interstrand hydrogen bonding pattern. Hydrogen bonds are shown in yellow. Carbon	

	is shown in gray, nitrogen in blue, and oxygen in red. Atomic coordinates are from PDB entry 1CAG (Bella et al., 1994).....	173
Figure 6.2	CD spectra of depsipeptides 6.8 (A) and 6.9 (B) with various concentrations of TMAO.....	183
Figure 6.3	T_m determinations of triple helical 6.8 (A) and 6.9 (B) with various concentrations of TMAO.....	183
Figure 6.4	Extrapolation of T_m values for triple helical (Pro-Pro-Gly) ₁₀ and 6.8 to 0 M TMAO.....	184
Figure 6.5	T_m determinations of triple helical 6.13 with various concentrations of TMAO.....	187

List of Schemes

Scheme 1.1	Synthesis of a collagen mimic with a cystine knot tether (Ottl et al., 1996; Ottl & Moroder, 1999a, b; Ottl et al., 1999; Muller et al., 2000; Ottl et al., 2000). NpysCl is 3-nitropyridyl-2-sulfenyl chloride.....	22
Scheme 2.1	Synthesis of Fmoc-3(S)-Hyp(OtBu) (2.4).....	45
Scheme 2.2	Synthesis of 2.8	46
Scheme 4.1	Synthesis of Compounds 4.1–4.6	82
Scheme 4.2	Acetylation of (Pro-Hyp-Gly) ₁₀ to produce peptide 4.9	87
Scheme 5.1	Synthetic Route to dUMP and dU ^F MP.....	115
Scheme 5.2	Synthetic Route to araUMP.....	116
Scheme 6.1	Synthesis of Fmoc-Pro-Yaa-Gly according to the method of Ottl and Moroder (Yaa = Pro or Hyp) (Ottl et al., 1999).....	176
Scheme 6.2	Synthesis of Fmoc-Pro-Flp-Gly (Holmgren et al., 1999).....	177
Scheme 6.3	Improved synthesis of Fmoc-Pro-Pro-Gly.....	178
Scheme 6.4	Failed synthesis of 6.8 and 6.9	179
Scheme 6.5	Proposed mechanism for two-residue deletion.....	180
Scheme 6.6	Synthesis of 6.10 and 6.11	180
Scheme 6.7	Successful synthesis of 6.8 and 6.9	181
Scheme 6.8	Synthesis of a collagen mimic with a cystine knot (Ottl et al., 1996; Ottl & Moroder, 1999b; Ottl et al., 1999). NpysCl is 3-nitropyridine-2-sulfenyl chloride.....	188
Scheme 6.9	Solid-phase assembly of a model cystine knot.....	190
Scheme 6.10	Proposed on-resin assembly of a cystine-knot-tethered collagen mimic.....	191

List of Tables

Table 1.1	Conformational stability of host–guest triple helices and ranking of frequency of appearance in polyproline II helices (Persikov et al., 2000).....	12
Table 1.2	Effect of Arg on the conformational stability of host–guest triple helices (Yang et al., 1997).....	14
Table 1.3	Effect of non-polar residues on the conformational stability of host–guest triple helices (Shah et al., 1996).....	15
Table 1.4	Effect of charged residues on the conformational stability of host–guest triple helices (Chan et al., 1997).....	17
Table 1.5	Conformational stability of Kemp triacid (KTA)- and related acetyl-terminated triple helices (Goodman et al., 1996).....	23
Table 1.6	Effect of peptoid residues on the conformational stability of Kemp triacid (KTA)- and acetyl-terminated triple helices (Goodman et al., 1998).....	25
Table 1.7	Effect of peptoid residues on the conformational stability of host–guest triple helices (Kwak et al., 1999).....	27
Table 1.8	Effect of 4(R)-hydroxy-L-proline (Hyp) in the Xaa position on the conformational stability of triple helices (Bann & Bächinger, 2000).....	31
Table 1.9	Effect of 4(R)-amino-L-proline (Amp) on the conformational stability of triple helices (Babu & Ganesh, 2001).....	33
Table 1.10	Effect of 4(R)-, 4(S)-fluoro-L-proline (Flp), 4(R)- and 4(S)-hydroxy-L-proline (Hyp) on the conformational stability of triple helices.....	34
Table 1.11	Melting temperatures of collagen mimics containing fluoroproline in the Xaa position.....	38
Table 1.12	Cytotoxicity and cell-binding of triple helices (Johnson et al., 2000).....	40
Table 2.1	Values of T_m for Synthetic (Pro–Hyp–Gly) ₃ –(Xaa–Yaa–Gly)–(Pro–Hyp–Gly) ₃ Triple Helices.....	46
Table 2.2	Main-Chain Angles of Proline and Hydroxyproline residues in Crystalline Amides and Triple Helices.....	49
Table 3.1	Values of $K_{trans/cis}$ for 4-substituted AcXaaOMe.....	70

Table 3.2	Effect of solvent on $K_{\text{trans/cis}}$ of AcXaaOMe.....	72
Table 3.3	Values of φ and ψ dihedral angles for the trans amide isomers of Ac-Xaa-OMe.....	75
Table 4.1	Effect of Hyp and Flp on the conformational stability of collagen triple helices.....	80
Table 4.2	Backbone dihedral angles of crystalline AcProOMe, 4.1 , 4.2 , 4.5 , and 4.7 derived from X-ray diffraction analysis.....	83
Table 4.3	$^{13}\text{C}'$ chemical shift (δ) of AcProOMe and compounds 4.1–4.8	84
Table 4.4	Thermodynamic parameters for amide bond isomerization in compounds 4.1–4.6	85
Table 5.1	Values of pK_a and K_i for 3'-nucleotides.....	117
Table 6.1	MALDI-TOF MS characterization of peptides produced by solid phase peptide synthesis.....	199

List of Abbreviations

3'-UMP	Uridine-3'-phosphate
Ac	Acetyl
Acm	Acetamidomethyl
Ala, A	Alanine
Amp	4(<i>R</i>)-amino-L-proline
araUMP	Arabinouridine-3'-phosphate
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
β -Gal	β -galactosyl
Bn	Benzyl
Boc	<i>t</i> -butoxycarbonyl
Bu	<i>n</i> -Butyl
CAPS	3-(cyclohexylamino)-1-propanesulfonic acid
CD	Circular dichroism
CH ₃ CN	Acetonitrile
CHES	2-(cyclohexylamino)ethanesulfonic acid
Cys, C	Cysteine
DCC	Dicyclohexylcarbodiimide
DIPEA	Diisopropylethylamine
DMF	Dimethyl formamide

dU ^F MP	2'-deoxy-2'-fluorouridine-3'-phosphate
dUMP	2'-deoxyuridine-3'-phosphate
ESI	Electrospray ionization
Et	Ethyl
Et ₂ O	Diethyl ether
EtOAc	Ethyl acetate
Flp	4(<i>R</i>)-fluoro-L-proline
flp	4(<i>S</i>)-fluoro-L-proline
Fmoc	9 <i>H</i> -fluoren-9-ylmethoxycarbonyl
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
HATU	<i>N</i> -(Dimethylamino)-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridine-1-ylmethlene- <i>N</i> -methylmethanaminium hexafluorophosphate- <i>N</i> -oxide
HBTU	2-(1 <i>H</i> -benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
His, H	Histidine
HOAc	Acetic acid
HOBt	<i>N</i> -Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrum
hyp	4(<i>S</i>)-Hydroxy-L-proline
Hyp, O	4(<i>R</i>)-Hydroxy-L-proline

Ile, I	Isoleucine
KTA	Kemp triacid
Leu, L	Leucine
Lys, K	Lysine
MALDI-TOF	Matrix-assisted laser desorption ionization-time of flight
mCPBA	3-chloroperoxybenzoic acid
Me	Methyl
meAla	N-methylalanine
MeOH	Methanol
MES	2-(<i>N</i> -Morpholino)ethanesulfonic acid
Met, M	Methionine
MOPS	3-(<i>N</i> -Morpholino)propanesulfonic acid
MS	Mass spectroscopy
Nleu	<i>N</i> -isobutylglycine
NMR	Nuclear magnetic resonance
Npys	5- or 3-nitro-2-pyridine sulfide
(NPys) ₂	2,2'-dithiobis(5-nitropyridine)
Npys-Cl	3-nitropyridine-2-sulphenyl chloride
OSu	<i>N</i> -oxysuccinimide
PBS	Phosphate-buffered saline
PDB	Brookhaven Protein Databank
Phe, F	Phenylalanine
Pro, P	Proline

PyBOP	Benzotriazole-1-yl-oxy-tris-pyrrolodino-phosphonium hexafluorophosphate
RNase A	Ribonuclease A
Ser, S	Serine
SPPS	Solid-phase peptide synthesis
t-Bu	<i>tert</i> -Butyl
TFA	Trifluoroacetic acid
TFAA	Trifluoroacetic anhydride
Thr, T	Threonine
TIS	Triisopropyl silane
T_m	Melting temperature
TMAO	Trimethylamine oxide
TREN	Tris(2-aminoethyl)amine
TRIS	Tris(carboethoxymethyl)-aminomethane
Trp, W	Tryptophan
Trt	Trityl, triphenylmethyl
TsOH	p-Toluenesulfonic acid
Tyr, Y	Tyrosine
UV	Ultraviolet
Val, V	Valine

Chapter 1[†]

INSIGHTS ON THE CONFORMATIONAL STABILITY OF COLLAGEN

1.1 Introduction

Collagen is the most abundant protein in animals, being a predominant component of connective tissues such as basement membranes, tendons, ligaments, cartilage, bone and skin. Over 19 different types of collagen have been identified, with several new ones being characterized currently. In addition, at least 15 other proteins have been shown to contain collagenous domains (Myllyharju & Kivirikko, 2001).

One of the defining features of collagen is its unique tertiary structure, consisting of three parallel left-handed polyproline II-type strands wound around a common axis to form a triple helix with a shallow right-handed superhelical pitch (Fig. 1.1). The packing of this coiled-coil structure requires that every third residue be glycine (Gly), resulting in a repeating Gly-Xaa-Yaa sequence. The residue in the Xaa position of these triplets is often L-proline (Pro), and the residue in the Yaa position is often 4(*R*)-hydroxy-L-proline (Hyp). Individual triple helices of collagen are organized into fibrils of great tensile strength and flexibility. These fibrils can be arranged and cross-linked so as to support stress efficiently in one, two, or three dimensions in tissues such as tendon, skin, and

[†] Portions of this chapter have been previously published under the same title. Reference: Jenkins, C. L.; Raines, R. T. *Nat. Prod. Rep.* **2002**, *19*, 49-59.

cartilage, respectively. Abnormalities in collagen structure are associated with connective tissue diseases, such as osteogenesis imperfecta, Ehlers–Danlos syndrome, scurvy, and some types of osteoporosis and arthritis (Prockop & Kivirikko, 1995; Prockop, 1998, 1999; Byers, 2000; Myllyharju & Kivirikko, 2001). A complete understanding of the basis for collagen stability (and instability) could lead to new biomaterials and effective therapies for these and other disorders.

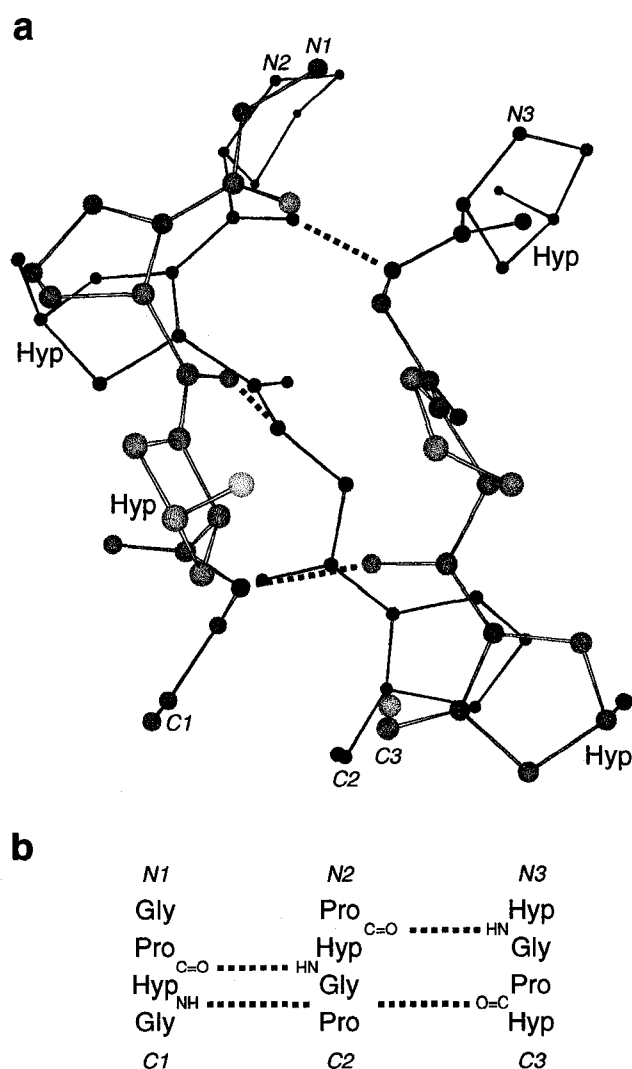


Figure 1.1 Segment of a $(\text{Pro-Hyp-Gly})_n$ triple helix. (a) Ball-and-stick representation indicating 4-hydroxy-L-proline residues and $\text{XaaC=O} \cdots \text{H-NGly}$ hydrogen bonds. (b) Register of the residues in the three strands of panel a. Atomic coordinates are from (Bella et al., 1994) (PDB entry 1CAG).

1.2 Collagen mimics

The most abundant type of collagen is Type I, in which each strand consists of approximately 300 Gly-Xaa-Yaa triplets. Discerning the chemical basis for the conformational stability of such a large molecule is difficult. In 1973, Prockop and coworkers first synthesized small mimics of the collagen triple helix. This seminal work

revealed that Hyp greatly stabilizes a triple helix when its 4(*R*) diastereomer, but not its 4(*S*) diastereomer, is present in the Yaa position (Berg & Prockop, 1973; Inouye et al., 1976). Subsequently, synthetic collagen mimics have been used to reveal the basis for this and other determinants of triple-helix structure and stability (Fields, G.B. & Prockop, 1996).

1.2.1 Structural studies

To understand the stability of collagen, it is important to understand its structure. X-Ray diffraction analyses of biological samples provided the first glimpse into the unique structure of collagen (Rougvie & Bear, 1953; Ramachandran, 1967; Fraser et al., 1979). In the 1950's, a model for the collagen triple helix was proposed by Kartha and Ramachandran (Ramachandran & Kartha, 1954, 1955; Ramakrishnan, 2001) and refined by Rich and Crick (Rich & Crick, 1955, 1961; Riddihough, 1998) to one that is essentially correct.

In 1994, Berman, Brodsky and coworkers used X-ray diffraction analysis to determine the first high-resolution structure of a triple-helical collagen mimic (Bella et al., 1994). The strands in this mimic had the sequence (Pro-Hyp-Gly)₄-Pro-Hyp-Ala-(Pro-Hyp-Gly)₅, and the resulting triple helix is designated here as "Gly→Ala". This landmark structure revealed not only the positions of the atoms of the triple helix, but also a regular network of water molecules surrounding the triple helices in the crystal lattice. The structure appeared to lend support to a hypothesis (Gustavson, 1957; Suzuki et al., 1980) that Hyp stabilizes collagen by forming hydrogen bonds with water

molecules surrounding the triple helix (Bella et al., 1994; Bella et al., 1995; Kramer, R. Z. & Berman, 1998).

Subsequent high-resolution structures of other collagen mimics led to a variety of conclusions about the factors that are most important in collagen stability. For example, Berman, Brodsky, Zagari, Mazzearella, and coworkers obtained crystals of (Pro-Pro-Gly)₁₀ under two different conditions, from which they were able to refine short sections of triple helix (21 residues each) to a resolution of 1.7 and 2.0 Å (Kramer, R.Z. et al., 1998). The two structures have similar molecular structure and hydration patterns, but different crystal packing. In the two structures, as in the Gly→Ala structure, the pyrrolidine ring in each Xaa position has a C^γ-endo (or “down” (Momany et al., 1975)) pucker, and Pro in each Yaa position has a C^γ-exo (or “up”) pucker, with only one exception. The main-chain torsion angles and the first hydration shell around the peptide are similar to those in the Gly→Ala structure, despite the differences in crystal packing. Because the structures of the two peptides (Gly→Ala and (Pro-Pro-Gly)₁₀) are so similar, these workers concluded that Hyp does not affect the triple-helix structure directly and that the contribution of Hyp to triple-helix stability arises only from Hyp–water interactions.

Okuyama and coworkers came to a different conclusion. Independently, they obtained a structure of crystalline (Pro-Pro-Gly)₁₀, from which they were able to refine a triple helix of 21 residues to a resolution of 1.9 Å (Okuyama et al., 1999). They found that the main-chain dihedral angles of their structure had no significant differences from those determined by the Berman group. Using different refinement procedures, the Okuyama group was able to locate only 15 water molecules, rather than the 40 solvent molecules in

the (Pro-Pro-Gly)₁₀ structure of the Berman group. In addition, the Okuyama group found that only 5 of 7 Pro residues in the Xaa position have a C^γ-endo pucker, and only 4 of 7 Pro residues in the Yaa position have a C^γ-exo pucker. This randomness is in conflict with the uniform pattern of Pro puckering observed by the Berman group.

Okuyama and coworkers also determined a structure for (Pro-Hyp-Gly)₁₀ to a resolution of 1.9 Å (Nagarajan et al., 1999). In this structure, the pattern of pyrrolidine ring puckering is similar to that in Gly→Ala. Yet, only 17 water molecules are apparent, and only 3 of 7 Hyp hydroxyl groups participate in hydrogen bonds with water. Okuyama and coworkers concluded that the close similarity of their (Pro-Pro-Gly)₁₀ and (Pro-Hyp-Gly)₁₀ structures indicates that Hyp does not affect directly the molecular structure, and that the nearly equal number of well-defined water molecules in the two structures indicates that the contribution of Hyp to stability is probably not due to an extensive network of water bridges (Nagarajan et al., 1999; Okuyama et al., 1999). This conclusion is consistent with an earlier study in which Engel, Prockop, and coworkers showed that Hyp confers extra stability upon a triple helix even in anhydrous solution (Engel et al., 1977).

A 1.3 Å-resolution structure of crystalline (Pro-Pro-Gly)₁₀ by Zagari, Mazzarella, and coworkers engendered an alternative hypothesis for the contribution of Hyp to collagen thermostability (Berisio et al., 2000; Vitagliano et al., 2001b; Berisio et al., 2002). These workers showed that Pro has a distinct ϕ torsion angle (C'_{i-1}-N_i-C^α_i-C'_i) in the Xaa and Yaa positions: $\phi = (-75 \pm 3)^\circ$ and $\phi = (-60 \pm 2)^\circ$, respectively. They noted that the different angles correlated with the C^γ-endo pucker in the Xaa position and the C^γ-exo pucker in the Yaa position. Because Hyp is more rigid than Pro and favors the C^γ-exo

pucker, Hyp in the Yaa position reduces the number of conformations available to the unfolded state and gives it a higher propensity to fold into a triple helix. Hyp in the Xaa position, however, cannot adopt the C^γ-endo pucker, and is thus a residue with an unfavorable ϕ angle for triple-helix formation. Furthermore, 4(*S*)-Hyp adopts the C^γ-endo pucker, making its ϕ angle inappropriate for the Yaa position. In the Xaa position, the hydroxyl group of 4(*S*)-Hyp is likely to destabilize the triple helix by a steric clash with a Pro residue in the Yaa position of another strand.

The structure of crystalline (Pro-Hyp-Gly)₁₀ refined to 1.4 Å resolution by the same researchers supports this alternative hypothesis for the contribution of Hyp to triple helix stability (Berisio et al., 2001). In this structure, the Pro residues have an average ϕ angle of $(-69.8 \pm 7.9)^\circ$ and the Hyp residues have an average ϕ angle of $(-57.4 \pm 2.0)^\circ$. In addition, the electron density for the pyrrolidine rings of the Hyp residues was well defined, and all had an “up” pucker, whereas the electron density for the Pro residues was not defined well enough to analyze the ring puckers—evidence that Hyp residues are more rigid than Pro.

The structure of a crystalline triple helix containing a sequence from Type III collagen revealed that Pro and Hyp can alter the helical pitch. Each strand of this triple helix has the sequence (Pro-Hyp-Gly)₃-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-(Pro-Hyp-Gly)₄. In the refinement process, no one model was found to fit the electron density map of the entire triple helix (Kramer, R. Z. et al., 1999). Rather, the terminal (Pro/Hyp-rich) regions fit to a model with 7-fold symmetry, as in (Pro-Hyp-Gly)₁₀, but the central (Pro/Hyp-poor) regions fit to a model with 10-fold symmetry, as in natural collagen (Rich & Crick, 1961; Okuyama et al., 1999). This result indicates that the helical pitch is

sequence-dependent, and that the main chain can accommodate small variations in torsion angles and still maintain triple-helicity. This finding forewarns of a potential complication in the interpretation of data from host–guest studies (*vide infra*), as the guest could prefer a different helical pitch than the host.

In addition to X-ray diffraction analysis, nuclear magnetic resonance (NMR) spectroscopy has been used to probe the structure of the collagen triple helix in solution (Long et al., 1992; Li et al., 1993; Melacini & Goodman, 1998). as well as measure its dynamics (Fan et al., 1993; Melacini et al., 2000) and folding kinetics (Liu et al., 1996; Liu et al., 1998; Buevich et al., 2000; Bhate et al., 2002; Xu et al., 2002; Xu et al., 2003). In general, the structural data obtained by NMR spectroscopy and X-ray diffraction analysis are in gratifying agreement.

With the advent of powerful computing techniques a number of interesting computational studies on collagen mimics have also been carried out. Improta and coworkers have carried out computations on [(Pro-Pro-Gly)₁₀]₃ that closely match the high-resolution X-ray structure obtained by Berisio and coworkers (Improta et al., 2002). Mooney and coworkers have calculated the conformational preferences of 4-substituted prolines in collagen mimics of the general sequence [(Pro-Yaa-Gly)₁₀]₃, where Yaa = proline, 4(*R*)-fluoroproline (Flp), 4(*R*)-Hyp, and 4(*R*)-aminoproline (Amp)—both neutral and protonated on the nitrogen at C-4. Their results indicated that the tendency for the pyrrolidine ring in the Yaa position to adopt the C γ -exo conformation increased in the order Amp < Pro < Hyp < Flp < AmpH⁺ (Mooney et al., 2002). They also calculated the energy perturbations upon addition of Gly→Ala substitutions in [(Pro-Hyp-Gly)₁₀]₃ and achieved $\Delta G_{\text{denaturation}}$ values very close to that measured by Brodsky and coworkers for

[(Pro-Hyp-Gly)₄-Pro-Hyp-Ala-(Pro-Hyp-Gly)₅]₃ (Bella et al., 1994; Mooney et al., 2001). They found that the first Gly→Ala substitution gave approximately half of the free energy perturbation, whereas the third Gly→Ala substitution yielded only about 10% of the energy difference. This result indicates that in collagen mutants, the number of Gly mutations, as well as their identity, may be important in predicting the severity of osteogenesis imperfecta cases (Mooney et al., 2001).

1.2.2 Host-guest studies

“Host-guest” peptides and proteins have been used to measure the propensity of individual amino acid residues to form an α -helix (O’Neil & DeGrado, 1990; Blaber et al., 1993; Groebke et al., 1996; Myers et al., 1997; Yang, J. X. et al., 1997) or β -sheet (Kim & Berg, 1993; Minor & Kim, 1994; Otzen & Fersht, 1995; Street & Mayo, 1999). In these studies, a parent sequence is chosen that is known to have the desired secondary structure. A central amino acid residue is then replaced systematically with other residues, and the conformational stability of the resulting structure is measured. This approach has been used to determine the contribution of individual as well as pairs of residues to triple-helix stability (Persikov et al., 2000a). It should be noted here that the term “host-guest” in this context is not intended to imply a non-covalent association; rather, it is a way to refer to a collection of closely-related peptide or protein variants in a concise manner.

The frequency with which a given triplet appears in natural collagen has been discerned by examining 4040 triplets from human fibril- and non-fibril-forming collagens (Ramshaw, J. A. M. et al., 1998). Approximately 49% of all possible triplets

are never or rarely found in fibrillar collagens, 41% are never found in non-fibrillar collagens, and 32% are never found in either type of collagen. The residues Trp and Cys are never found in triple-helical regions of collagen, Tyr is rarely found in the Xaa position, and Phe is never found in the Yaa position. A schematic of the combined triplet distribution of several human collagens is shown in Fig. 1.2. It is interesting to note that no triplet appears with high frequency except for Gly-Pro-Hyp, which is the most stabilizing triplet found in natural collagen (Sakakibara et al., 1973; Engel et al., 1977; Germann & Heidemann, 1988). This survey provides a context for the host-guest studies of collagen mimics, and reduces the number of triplets that need to be studied to obtain a complete picture of the contribution of natural triplets to conformational stability.

Xaa/Yaa	A	C	D	E	F	G	H	I	K	L	M	N	O	Q	R	S	T	V	W	Y
A	0.9		0.5						1.2			0.3	3.4	0.4	1.1	0.4		0.4		
C																				
D	0.4		0.3						1.0				1.5	0.5	1.2	0.3		0.3		
E	1.2		0.4					0.4	2.5		0.3	0.3	2.8	0.8	2.7	0.3	0.6	0.4		
F													2.5	0.3						
G									0.3				0.3							
H													0.5							
I	0.4								0.3				1.5							
K			0.5										1.4							
L	0.5		0.4						1.0				5.5	0.7	0.6		0.3			
M													0.6							
N													0.7							
P	3.4		0.4	0.4				0.9	2.7	0.5	0.4	0.3	10.5	2.5	2.6	1.4	0.8	1.3		
Q	0.3								0.6				1.1	0.3	0.3					
R	0.3		0.5										1.1							
S	0.4								0.4				2.3		0.5			0.3		
T													0.8							
V									0.5				1.1	0.3						
W																				
Y													0.5							

Figure 1.2 Distribution of triplets in human collagen types I, II, III, V, XI, IV, VI, VII, VIII, IX, X, and XIII (Ramshaw, J. A. M. et al., 1998). Rows and columns refer to the Xaa and Yaa positions of 4040 Gly-Xaa-Yaa triplets. Black boxes indicate triplets found fewer than 3 times ($<0.074\%$), hatched boxes indicate triplets found 3–10 times ($0.074\text{--}0.25\%$), and numbers in white boxes represent the occurrence of the most common triplets, rounded to the nearest 0.1%. All Pro residues in the Yaa position are assumed to be hydroxylated. All residues are indicated by their single letter codes, with O indicating Hyp.

A comprehensive host–guest study of individual residues in the Xaa and Yaa positions was reported by Brodsky and coworkers (Persikov et al., 2000b). The guest peptides had the sequence Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Hyp-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂ and Ac-(Gly-Pro-Hyp)₃-Gly-Pro-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂, where Xaa and Yaa were the 19 proteinogenic amino acids other than Pro and Hyp, respectively. They found that the triple helix with Xaa = Pro and Yaa = Hyp was the most stable, with a T_m (which is the temperature at the midpoint of the thermal transition) of 47.3 °C. Surprisingly, replacement of Hyp with Arg in the Yaa position resulted in a triple helix of nearly equal stability ($T_m = 47.2$ °C), whereas all the other amino acid substitutions resulted in triple helices with T_m values that are at least 5 °C lower (Table 1.1). This study revealed that there is a moderate correlation between the contribution of a given residue to triple-helix stability and its propensity to adopt a polyproline II-like conformation. This correlation was shown to be better for the Xaa position than the Yaa position, perhaps because of the greater solvent exposure of the residue in the Xaa position (Jones & Miller, 1991; Simon-Lukasik et al., 2003). Trp, which is not found in natural collagen, was the most destabilizing residue in both the Xaa and Yaa positions; other aromatic amino acid residues (Phe and Tyr) were also destabilizing in both positions.

Table 1.1 Conformational stability of host–guest triple helices and ranking of frequency of appearance in polyproline II helices (Persikov *et al.*, 2000b).

Xaa ^a	<i>T_m</i> (°C)	Yaa ^a	<i>T_m</i> (°C)	polyproline II frequency ^b
Pro	47.3	Hyp	47.3	Pro
Glu	42.9	Arg	47.2	Gln
Ala	41.7	Met	42.6	Arg
Lys	41.5	Ile	41.5	Lys
Arg	40.6	Gln	41.3	Thr
Gln	40.4	Ala	40.9	Leu
Asp	40.1	Val	40.0	Asp
Leu	39.0	Glu	39.7	Met
Val	38.9	Thr	39.7	Ala
Met	38.6	Cys	37.7	Cys
Ile	38.4	Lys	36.8	Val
Asn	38.3	His	35.7	Glu
Ser	38.0	Ser	35.0	Asn
His	36.5	Asp	34.0	Phe
Thr	36.2	Gly	32.7	Ser
Cys	36.1	Leu	31.7	Ile
Tyr	34.3	Asn	30.3	Trp
Phe	33.5	Tyr	30.2	Tyr
Gly	33.2	Phe	28.3	His
Trp	31.9	Trp	26.1	Gly

^a Peptide strands of host–guest triple helices have the sequence Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Hyp-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂ and Ac-(Gly-Pro-Hyp)₃-Gly-Pro-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂.

^b Frequency of occurrence of amino acids in polyproline II regions in globular proteins (Stapley & Creamer, 1999).

The unexpected stability conferred by the guest triplet Gly-Xaa-Arg, which accounts for nearly 10% of the triplets in common human collagens (Fig. 1.2) (Ramshaw, J. A. M.

et al., 1998), was examined further with the peptides listed in Table 1.2 (Yang, W. et al., 1997). The stabilization is not due to cationic charge alone, as triple helices containing Lys are considerably less stable than are those containing Arg (Table 1.1). A triple helix containing the guest triplet Gly-Pro-homoArg, which was synthesized from the guest triplet Gly-Pro-Lys by reaction with a guanidino group transfer reagent, had stability between that of triple helices containing Gly-Pro-Arg and Gly-Pro-Lys. Thus, some property of the guanidino group is important in stabilization, but the longer homoArg side chain decreases stability somewhat. Data on triple helices containing the guest triplets Gly-Ala-Hyp, Gly-Ala-Arg, and Gly-Ala-Lys show that Arg confers stability even when the residue in the Xaa position is not Pro, possibly by forming either an intra- or interchain hydrogen bond with a main-chain carbonyl group. If two Gly-Pro-Arg triplets are adjacent, unfavorable Coulombic interactions destabilize the triple helix somewhat, but separation by one or more intervening triplets abolishes this effect. Still, triple helices with two Gly-Pro-Arg triplets separated by one or two Gly-Pro-Hyp triplets have T_m values slightly lower than does a triple helix containing one Gly-Pro-Arg triplet. The peptide Ac-(Gly-Pro-Arg)₈-Gly-Gly-NH₂ could not be folded into a triple helix until its concentration was 3 mg/mL and the NaCl concentration of the solution was raised to 2 M. Apparently, unfavorable Coulombic interactions between the Arg residues are considerable.

Table 1.2 *Effect of Arg on the conformational stability of host–guest triple helices (Yang, W. et al., 1997).*

Peptide Sequence	T_m (°C) ^a
Ac-(Gly-Pro-Hyp) ₈ -Gly-Gly-NH ₂	45.5
Ac-(Gly-Pro-Hyp) ₃ - Gly-Pro-Arg -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	45.5
Ac-(Gly-Pro-Hyp) ₃ - Gly-Pro-homoArg -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	42.8
Ac-(Gly-Pro-Hyp) ₂ - Gly-Pro-Arg -(Gly-Pro-Hyp) ₂ - Gly-Pro-Arg -(Gly-Pro-Hyp) ₂ -Gly-Gly-NH ₂	42.8
Ac-(Gly-Pro-Hyp) ₃ - Gly-Pro-Arg -Gly-Pro-Hyp- Gly-Pro-Arg -(Gly-Pro-Hyp) ₂ -Gly-Gly-NH ₂	42.2
Ac-(Gly-Pro-Hyp) ₃ - Gly-Arg-Hyp -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	40.6
Ac-(Gly-Pro-Hyp) ₄ - Gly-Pro-Arg-Gly-Pro-Arg -(Gly-Pro-Hyp) ₂ -Gly-Gly-NH ₂	40.4
Ac-(Gly-Pro-Hyp) ₃ - Gly-Ala-Hyp -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	39.9
Ac-(Gly-Pro-Hyp) ₃ - Gly-Ala-Arg -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	38.2
Ac-(Gly-Pro-Hyp) ₃ - Gly-Pro-Lys -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	36.8
Ac-(Gly-Pro-Arg) ₈ -Gly-Gly-NH ₂	32.6 ^b
Ac-(Gly-Pro-Hyp) ₃ - Gly-Ala-Lys -(Gly-Pro-Hyp) ₄ -Gly-Gly-NH ₂	30.8

^a Values of T_m were measured in phosphate-buffered saline (pH 7.0, 150 mM NaCl).

^b Value of T_m was measured in 10 mM sodium phosphate buffer (pH 7.0) containing NaCl (2 M).

The contribution of pairs of common nonpolar residues to triple-helical stability has been examined by comparing a set of host–guest peptides (Shah et al., 1996). The residues Pro and Hyp, along with Ala, Phe, and Leu, which are the most common nonpolar residues found in collagen, were included in various combinations in peptides of the form Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂. The Gly-Xaa-Yaa guest triplets and the T_m values for their triple helices are listed in Table 1.3. Leu and Phe are more destabilizing in the Yaa than in the Xaa position, and Phe is usually more destabilizing than Leu. It is also apparent from the data that hydrophobic residues do not

lend any special stability to collagen triple helices, in contrast to globular proteins, which have hydrophobic cores (Kauzmann, 1959). In the Xaa position of collagen, non-polar residues could participate in intermolecular interactions with other collagen strands or with other proteins that bind to collagen (Shah et al., 1996), as the Xaa position is somewhat more solvent-exposed than is the Yaa position (Jones & Miller, 1991; Simon-Lukasik et al., 2003).

Table 1.3 *Effect of non-polar residues on the conformational stability of host-guest triple helices (Shah et al., 1996).*

Gly-Xaa-Yaa	T_m (°C) ^a	Gly-Xaa-Yaa	T_m (°C) ^a
Gly-Pro-Hyp	44.5	Gly-Leu-Ala	29.5
Gly-Ala-Hyp	39.9	Gly-Ala-Ala	29.3
Gly-Leu-Hyp	39.0	Gly-Pro-Phe	28.3
Gly-Pro-Ala	38.3	Gly-Ala-Leu	27.8
Gly-Phe-Hyp	33.5	Gly-Phe-Ala	23.4
Gly-Pro-Leu	32.7	Gly-Ala-Phe	20.7

^a Values of T_m are for triple helices of Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂ were measured at pH 7.4.

Several pairs of charged residues have been examined in a similar manner (Chan et al., 1997). Glu, Asp, Arg, and Lys were placed individually and in oppositely-charged pairs into peptides with the general sequence Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂. The T_m values of the resulting triple helices are listed in Table 1.4. Comparing the stabilities of triple helices containing pairs of charged residues to those of

triple helices containing individual charged residues at neutral pH indicates that, except for the triplets Gly-Lys-Asp and Gly-Arg-Asp, there is no stabilization from favorable Coulombic interactions. Rather, differences in stability can be accounted for by the effects of individual charged residues. Moreover, charged residues in the Xaa position have little effect on the T_m values of triple helices, and these charged residues decrease conformational stability only marginally compared to Pro and not at all compared to Ala. The difference is much larger when the charged residues are in the Yaa position. An explanation for this phenomenon is that side chains in the Yaa position are less solvent-accessible and closer to side chains in other chains (Jones & Miller, 1991; Simon-Lukasik et al., 2003), leading to unfavorable steric and Coulombic interactions.

Table 1.4 *Effect of charged residues on the conformational stability of host–guest triple helices (Chan et al., 1997).*

Gly-Xaa-Yaa	pH	T_m (°C) ^a	Gly-Xaa-Yaa	pH	T_m (°C) ^a
Gly-Ala-Hyp ^a	7.4	39.9	Gly-Pro-Ala ^b	7.4	38.3
Gly-Pro-Hyp ^a	7.4	44.5	Gly-Ala-Ala ^b	7.4	29.3
Gly-Asp-Hyp	2.7	37.6	Gly-Pro-Asp	2.7	33.1
	7.0	40.1		7.0	30.1
	12.2	38.0		12.2	30.1
Gly-Glu-Hyp	2.7	39.7	Gly-Pro-Glu	2.7	41.9
	7.0	42.9		7.0	39.7
	12.2	40.9		12.2	38.5
Gly-Lys-Hyp	2.7	40.4	Gly-Pro-Lys	2.7	37.1
	7.0	41.5		7.0	36.8
	12.2	38.3		12.2	38.8
Gly-Arg-Hyp	2.7	39.4	Gly-Pro-Arg	2.7	45.5
	7.0	40.6		7.0	44.5
	12.2	38.0		12.2	43.1
Gly-Asp-Lys	2.7	26.5	Gly-Lys-Asp	2.7	30.5
	7.0	30.9		7.0	35.8
	12.2	29.9		12.2	30.2
Gly-Asp-Arg	2.7	33.4	Gly-Arg-Asp	2.7	28.8
	7.0	37.1		7.0	35.0
	12.2	34.4		12.2	31.9
Gly-Glu-Lys	2.7	29.5	Gly-Lys-Glu	2.7	36.5
	7.0	35.0		7.0	35.3
	12.2	33.1		12.2	31.6
Gly-Glu-Arg	2.7	37.3	Gly-Arg-Glu	2.7	35.0
	7.0	40.4		7.0	33.8
	12.2	39.1		12.2	32.2

^a Values of T_m are for triple helices of Ac-(Gly-Pro-Hyp)₃-Gly-Xaa-Yaa-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂. ^b The triplets Gly-Pro-Hyp, Gly-Ala-Hyp, Gly-Pro-Ala, and Gly-Ala-Ala are included for comparison (Shah et al., 1996).

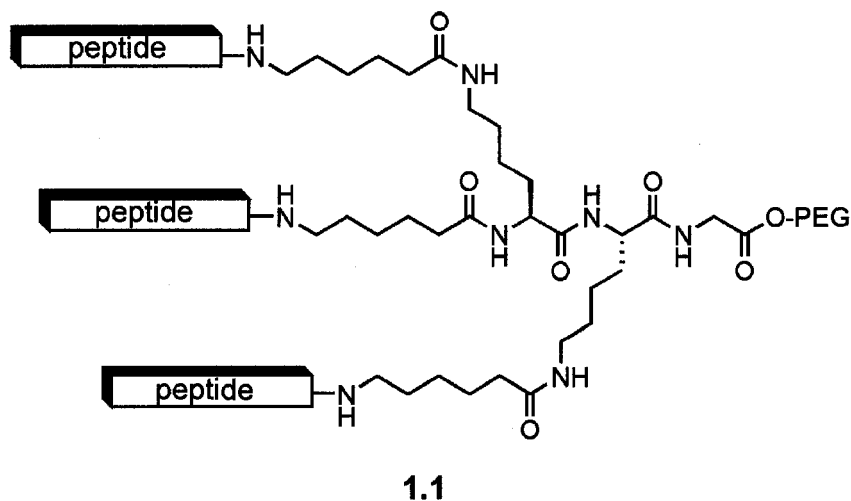
An extensive report of host-guest peptides containing pairs of amino acids excluding Pro and Hyp (including data from the studies cited above) extends the host-guest data to date to include approximately 78% of triplets found in fibrillar collagens (Persikov et al., 2002). The host–guest studies outlined above help to clarify how a particular residue

affects triple-helix stability in a relatively stable triple-helical environment. It is not clear, however, whether this understanding can be applied to regions of collagen lacking Pro and Hyp residues, which can have a different helical pitch (Kramer, R. Z. et al., 1999). The frequency with which a given residue appears in natural collagen (Fig. 1.2) does not necessarily correlate with the stability that that residue imparts to the triple helix in host-guest studies (Table 1.1). Of course, some residues could be present in collagen for other purposes, such as participating in interactions with other biomolecules.

Replacement of Gly with other amino acids has also been studied in host-guest fashion. A review (Persikov et al., 2000a) of these studies leads to the conclusion that the instability caused by these Gly substitutions varies, and that the degree of flexibility in the surrounding peptide can influence the severity of the destabilization. A recent computational study supports these observations (Radmer & Klein, 2004). The identity and location of the Gly substitution also seems to correlate somewhat with the severity of osteogenesis imperfecta, which is caused by Gly substitutions in collagen.

1.2.3 Tethered triple-helical peptides

Several groups have used synthetic methods to tether three collagen-related peptides so as to enhance their triple helicity. For example, Heidemann and coworkers used a di-Lys-based construction with aminohexanoic acid linkers between the three Lys amino groups (one α and two ϵ) and their peptides (**1.1**) (Thakur et al., 1986; Germann & Heidemann, 1988). Fields and coworkers used a similar tether with an orthogonal protecting group strategy to produce triple helices of biological interest (Fields, C.G. et al., 1993; Fields, C. G. et al., 1993).



Tanaka and coworkers used a lysine dimer, similar to those of Heidemann and Fields, but having a β -alanine linker. In addition, the Tanaka group took a different approach to assembling a triple helix. Whereas Fields and Heidemann built their peptides onto the tether one strand at a time, Tanaka and coworkers synthesized a peptide with a Cys residue at its N-terminus. The sulfhydryl group was alkylated with the tether **1.2** (Tanaka, T. et al., 1993).

The Tanaka group also devised a method to cross-link a collagen peptide at both the N- and C-termini (Tanaka, Y. et al., 1998). The N-terminal tether is **1.2**, and the C-terminal linker is **1.3**. The collagen mimic was synthesized to include a Lys residue at the C-terminus of the peptide with a Ser residue attached to N^ε. NaIO₄ oxidation of the peptide generated an aldehyde from this Ser, which then formed an oxime linkage with the aminoxy group on the linker. The N-linked peptides form more stable triple helices than do the unlinked peptides, and the dually N,C-linked peptides **1.4** form even more stable triple helices than do the N-linked ones.

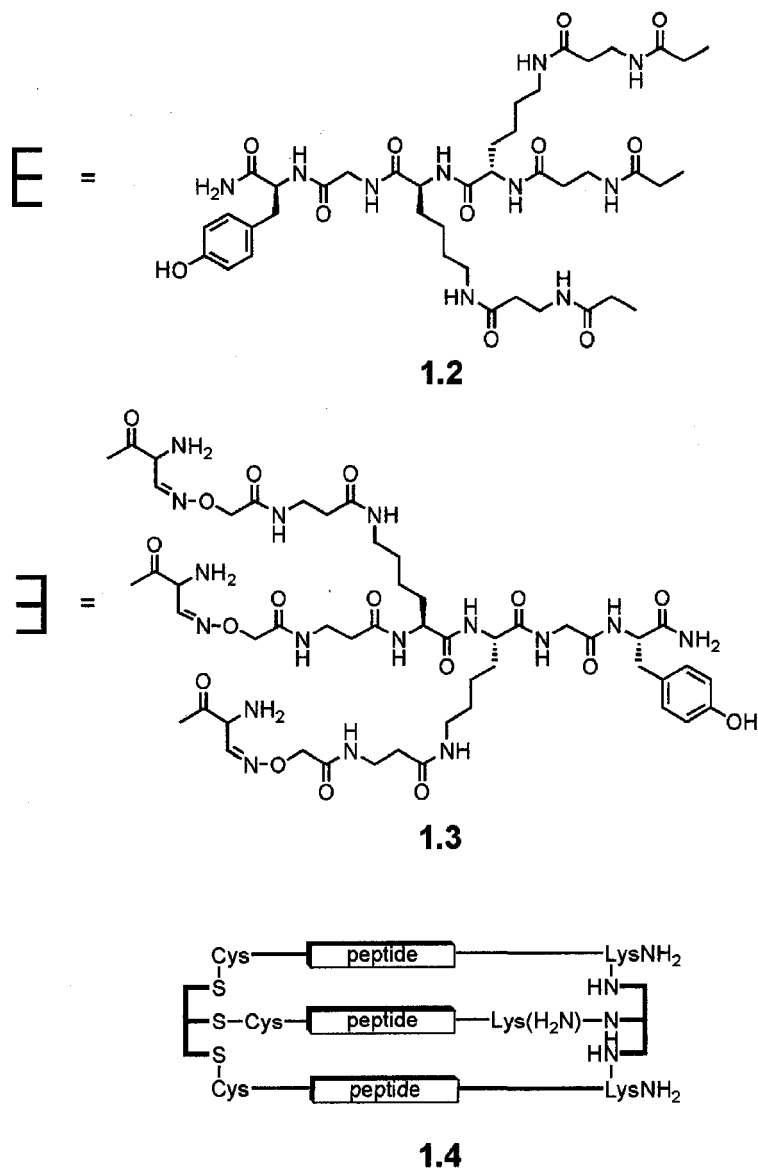
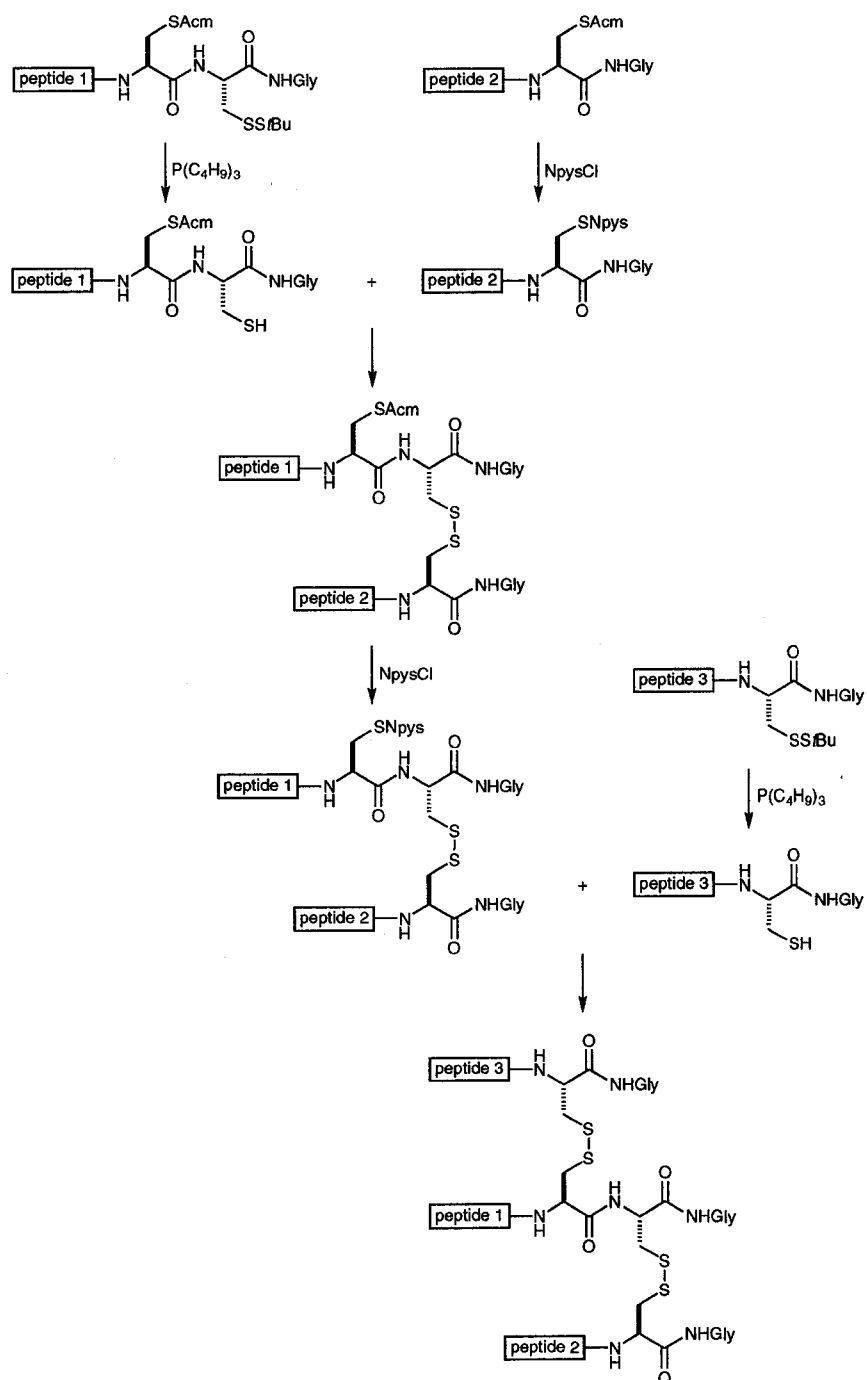


Figure 1.3 *N*- and *C*-linked collagen mimics.

Several groups have designed tethers based on motifs found in nature. Fields and coworkers have attached lipids to the N-termini of collagen mimics and relied on self-assembly processes to drive triple-helix folding and stabilization (Fields, G. B., 1999). Engel, Bächinger, and coworkers used a homotrimeric globular protein in a similar manner (Frank et al., 2001). They created a plasmid that directs *E. coli* to produce a

chimeric protein in which (Gly-Pro-Pro)₁₀ is fused to the 27-residue C-terminal domain of bacteriophage T4 fibritin protein (termed “foldon”). The chimerae formed trimers with a high degree of conformational stability. They subsequently measured the folding kinetics (Boudko et al., 2002) of (Gly-Pro-Pro)₁₀-foldon and its X-ray crystal structure (Stetefeld et al., 2003). Moroder and coworkers used a simplified version of the disulfide bridges found in the C-terminal domain of procollagen to stabilize collagen mimics (Ottl et al., 1996; Ottl & Moroder, 1999a; Ottl et al., 1999) (Ottl & Moroder, 1999b; Ottl et al., 2000) (Muller et al., 2000). They designed a “cystine knot” derived from two pairs of differentially protected cysteine residues such that three strands are tethered in a selective manner, as shown in Scheme 1.1. Many human collagens contain two or three different strands (Myllyharju & Kivirikko, 2001), and the cystine knot provides a facile means to stabilize heterotrimeric triple helices.



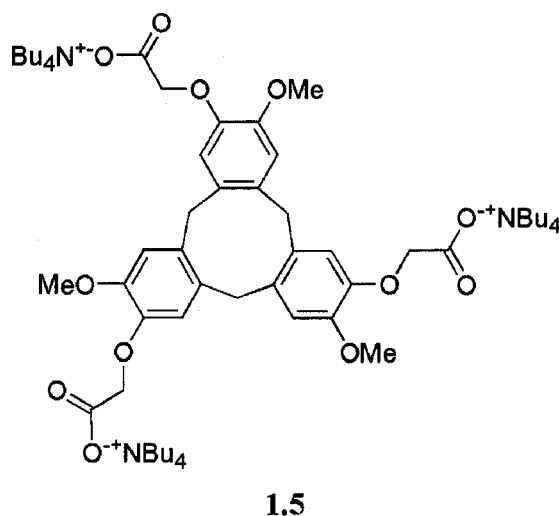
Scheme 1.1 Synthesis of a collagen mimic with a cystine knot tether. (Ottl et al., 1996; Ottl & Moroder, 1999a, b; Ottl et al., 1999; Muller et al., 2000; Ottl et al., 2000) *NpysCl* is 3-nitropyridyl-2-sulfonyl chloride.

Goodman and coworkers used a cyclic tether based on *cis,cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid, also known as the Kemp triacid (Kemp & Petrakis, 1981), to create triple-helical collagen mimics (Goodman et al., 1998). They chose this template because it is rigid and because its three carboxylic acid groups are fixed on the same face of the cyclohexane ring. Condensation of each carbonyl group of the triacid with a Gly residue provides enough flexibility for the collagen-like strands to adopt the requisite register (Fig. 1.1) as well as the larger diameter of the triple helix. In their first use of the Kemp triacid template (Goodman et al., 1996), Goodman and coworkers showed that templated (Gly-Pro-Hyp)_n strands form much more stable triple helices than do acylated strands of equivalent length, as listed in Table 1.5. Indeed, the Kemp triacid template enabled incipient triple-helix formation from strands with only three Gly-Pro-Hyp triplets, which is the shortest triple helix reported to date.

Table 1.5 Conformational stability of Kemp triacid (KTA)- and related acetyl-terminated triple helices (Goodman et al., 1996).

Peptide	<i>T_m</i> (°C)	
	H ₂ O	ethylene glycol/H ₂ O (2:1)
KTA-[Gly-(Gly-Pro-Hyp)-NH ₂] ₃	No transition	No transition
KTA-[Gly-(Gly-Pro-Hyp) ₃ -NH ₂] ₃	30	50
KTA-[Gly-(Gly-Pro-Hyp) ₅ -NH ₂] ₃	70	(Not determined)
KTA-[Gly-(Gly-Pro-Hyp) ₆ -NH ₂] ₃	81	(Not determined)
Ac-Gly-Pro-Hyp-NH ₂	No transition	No transition
Ac-(Gly-Pro-Hyp) ₃ -NH ₂	No transition	No transition
Ac-(Gly-Pro-Hyp) ₅ -NH ₂	18	32
Ac-(Gly-Pro-Hyp) ₆ -NH ₂	26	59

A macrocyclic collagen template molecule has recently been reported to stabilize triple helices as well as Goodman's Kemp triacid template. This template is a glycolate derivative of cyclotrimeratrylene (compound **1.5**), and it was shown that the (+)-enantiomer of the template is better than the (-)-enantiomer at stabilizing triple helices containing either a natural sequence or Pro-Hyp-Gly repeats (Rump et al., 2002).



1.2.4 Peptoid residues

Having shown the stabilizing ability of the Kemp triacid as a template, Goodman and coworkers incorporated N-substituted glycine (peptoid) residues into their collagen mimics. Their initial work focused on a single peptoid residue, N-isobutylglycine (Nleu), which they chose because of its bulky hydrophobic side chain (Goodman et al., 1998). They found that the sequences (Gly-Pro-Nleu)_n and (Gly-Nleu-Pro)_n ($n \geq 9$ and $n \geq 6$, respectively) form stable triple helices, whereas Gly-Nleu-Nleu has to be included in a host-guest fashion within sequences such as (Gly-Pro-Hyp)_n to adopt a triple-helical conformation. In addition, (Gly-Nleu-Pro)_n forms more stable triple helices than does

(Gly-Pro-Nleu)_n, as listed in Table 1.6. The workers reasoned, with the aid of molecular modeling studies, that the isobutyl group of Nleu can form more hydrophobic contacts with Pro in other chains in triple helices composed of (Gly-Nleu-Pro)_n than in those composed of (Gly-Pro-Nleu)_n.

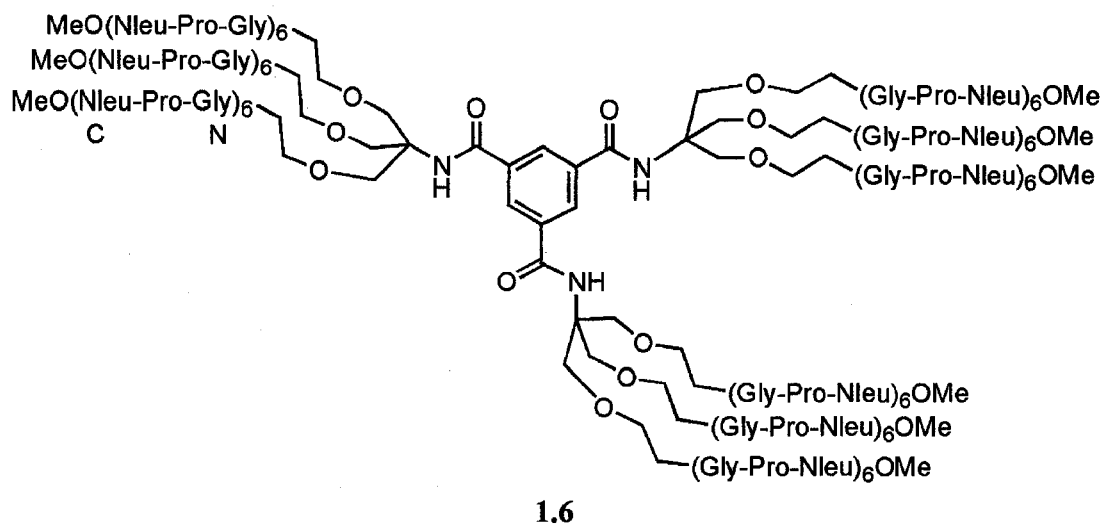
Table 1.6 *Effect of peptoid residues on the conformational stability of Kemp triacid (KTA)- and acetyl-terminated triple helices (Goodman et al., 1998).*

Peptide	<i>T_m</i> (°C)	
	H ₂ O	ethylene glycol/H ₂ O (2:1)
Ac-(Gly-Pro-Nleu) ₆ -NH ₂	No transition	35
Ac-(Gly-Pro-Nleu) ₉ -NH ₂	39	58
(Gly-Pro-Nleu) ₅ -NH ₂	No transition	No transition
(Gly-Pro-Nleu) ₆ -NH ₂	No transition	28
(Gly-Pro-Nleu) ₇ -NH ₂	No transition	39
(Gly-Pro-Nleu) ₉ -NH ₂	28	50
KTA-[Gly-(Gly-Pro-Nleu)-NH ₂] ₃	No transition	No transition
KTA-[Gly-(Gly-Pro-Nleu) ₃ -NH ₂] ₃	No transition	12
KTA-[Gly-(Gly-Pro-Nleu) ₆ -NH ₂] ₃	33	52
KTA-[Gly-(Gly-Pro-Nleu) ₉ -NH ₂] ₃	47	69
Ac-(Gly-Nleu-Pro) ₃ -NH ₂	No transition	No transition
Ac-(Gly-Nleu-Pro) ₆ -NH ₂	26	43
Ac-(Gly-Nleu-Pro) ₉ -NH ₂	^a	^a
KTA-[Gly-(Gly-Nleu-Pro) ₃ -NH ₂] ₃	No transition	22
KTA-[Gly-(Gly-Nleu-Pro) ₃ -NH ₂] ₃	33	52
Ac-(Gly-Pro-Hyp) ₂ -(Gly-Nleu-Nleu) ₂ -(Gly-Pro-Hyp) ₂ -NH ₂	No transition	25
KTA-[Gly-(Gly-Pro-Hyp) ₂ -(Gly-Nleu-Nleu) ₂ -(Gly-Pro-Hyp) ₂ -NH ₂] ₃	20	40

^a Solution became cloudy at 35 °C.

Goodman and coworkers have recently shown that two additional tether molecules are effective at stabilizing triple helices containing the Nleu peptoid residue. Tris(2-aminoethyl)amine (TREN) was found to be more effective than the Kemp triacid at stabilizing triple helices containing five or more (Gly-Nleu-Pro) repeats, but less effective at stabilizing triple helices with three or less (Gly-Nleu-Pro) repeats (Kwak et al., 2002). A dendritic tether molecule consisting of an aromatic tricarboxylic acid coupled with tris(carboethoxymethyl)aminomethane (TRIS) was shown to stabilize triple helices

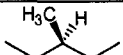
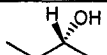

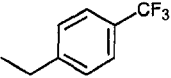
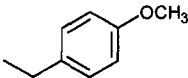
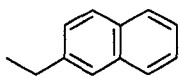
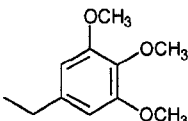

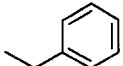

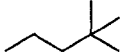
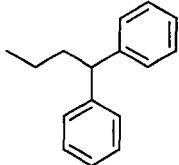
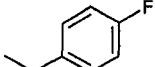
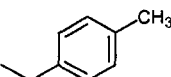

containing (Gly-Nleu-Pro)₆ more effectively than those containing (Gly-Pro-Nleu)₆ sequences (compound **1.6**). It was proposed that the extra triple helix stabilization in the context of the dendrimer arises from intramolecular clustering of the triple helices (Kinberger et al., 2002).



The Goodman group synthesized a series of host–guest peptides with the sequence Ac-(Gly-Nleu-Pro)₃-(Gly-N_x-Pro)₂-(Gly-Nleu-Pro)₃-NH₂, where N_x refers to a peptoid residue (Kwak et al., 1999). The various N_x residues and the *T_m* values of their triple helices are listed in Table 1.7. It is interesting to note that when the side chain of the N_x residue is 2-hydroxyethyl or 2-aminoethyl, triple helices are not formed; whereas when the side chain is 2(*R*)-hydroxypropyl, the resulting triple helix is quite stable. One explanation is that the functionalized linear alkyl chains lacks the steric bulk necessary to stabilize the triple helix. Another explanation is that the solvation of the hydroxyl and amino groups is especially disruptive to triple-helix formation. The workers conclude that Gly-Nleu-Pro is an effective triple-helix promoter, that a variety of peptoid residues can

be incorporated into collagen mimics, and that hydrophobic effects are important for the inter- and intrachain interactions that lead to stable peptoid-containing triple helices in aqueous solution.

Table 1.7 Effect of peptoid residues on the conformational stability of host–guest triple helices (Kwak et al., 1999).

side chain of N _x	<i>T</i> _m (°C) ^a	side chain of N _x	<i>T</i> _m (°C) ^a
	49		42
	47		41
	47		>40 ^b
	46		38
	45		37
	44		24
	44		
	43		not triple helical

^a Values of *T*_m are for triple helices of Ac-(Gly-Nleu-Pro)₃-(Gly-N_x-Pro)₂-(Gly-Nleu-Pro)₃-NH₂ in H₂O.

^b Peptide became insoluble at 40 °C.

Other work (Kersteen & Raines, 2001) examined the contribution of a tertiary amide other than Pro, Hyp, or a peptoid residue to the conformational stability of a triple helix. In their study, Kersteen and Raines determined the conformational stability of triple helices of host–guest peptides with the sequence (Pro-Hyp-Gly)₃-Xaa-Yaa-Gly-(Pro-

Hyp-Gly)₃. In order of decreasing stability, the central triplets are Pro-Hyp-Gly, Pro-Pro-Gly, Ala-Hyp-Gly, Pro-Ala-Gly, Pro-meAla-Gly, and meAla-Pro-Gly, where meAla refers to N-methyl-L-alanine. This residue is identical to Pro and Hyp except for the absence of the $-C^{\gamma}H_2-$ and $-C^{\gamma}HOH-$ groups of the pyrrolidine ring, respectively. These workers concluded that the mere presence of tertiary amides in collagen does not make a major contribution to triple-helix stability. Rather, the conformational restrictions imposed by the pyrrolidine rings of Pro and Hyp are critical.

1.3 4-Substituted proline residues

In common forms of human collagen, Gly-Xaa-Hyp triplets account for nearly 40% of the amino acid sequence (Fig. 1.2) (Ramshaw, J. A. M. et al., 1998). Hyp residues are not incorporated into collagen by ribosomes. Rather, this post-translation modification of Pro residues is mediated by prolyl 4-hydroxylase (Guzman, 1998) after collagen biosynthesis but before the chains form a triple helix. Hydroxylation is critical for the folding of collagen, its secretion to the extracellular matrix, and its further processing and incorporation into fibrils or other structures (Bulleid et al., 1996; Walmsley et al., 1999; Snellman et al., 2000; Byers, 2001). The absence of prolyl 4-hydroxylase is lethal to the nematode *Caenorhabditis elegans* (Friedman et al., 2000; Winter & Page, 2000).

In the 1970's, workers began to notice that the Hyp content of a collagen triple helix correlates with its conformational stability (Burjanadze, 1979, 2000). In 1973, Prockop and coworkers demonstrated that triple-helix stability decreases in the order: (Pro-4(*R*)-Hyp-Gly)₁₀ >> (Pro-Pro-Gly)₁₀ >> (4(*R*)-Hyp-Pro-Gly)₁₀, (Pro-4(*S*)-Hyp-Gly)₁₀, or (4(*S*)-

Hyp-Pro-Gly)₁₀ (Berg & Prockop, 1973; Inouye et al., 1976; Inouye et al., 1982).

Understanding the chemical basis for this finding has motivated much work.

Raines and coworkers showed that attaching an electronegative atom to C^γ has substantial effects on the chemical properties of a proline residue. For example, the nitrogen pK_a of the conjugate acid of 4(*R*)-fluoro-L-proline (FlpOH; 9.23) is lower than that of HypOH (9.68) and ProOH (10.8) (Eberhardt et al., 1996). The nitrogen of AcFlpOMe is more pyramidal than that of AcHypOMe or AcProOMe (Panasik et al., 1994). This result indicates that the nitrogen of AcFlpOMe has greater sp³ character and hence higher electron density. The amide I vibrational mode, which results primarily from the C=O stretching vibration, decreases in the order: AcFlpOMe > AcHypOMe > AcProOMe (Eberhardt et al., 1996). The value of Δ*H*[‡] for amide bond isomerization is smaller for AcFlpOMe than for AcProOMe (Eberhardt et al., 1996). Each of these results is consistent with the traditional picture of amide resonance (Pauling, 1960) coupled with an inductive effect that increases the bond order in the amide C=O bond and decreases the bond order in the amide C–N bond. Raines and coworkers suggested that this inductive effect is the basis for the contribution of Hyp residues to the conformational stability of collagen.

1.3.1 Hydroxyproline residues

New insight on the contribution of Hyp residues to triple-helix stability was inspired by an unusual collagen. Based on the amino acid sequence of cuticle collagen in the hydrothermal vent worm *Riftia pachyptila*, Bann and Bächinger made a pair of peptides with Hyp in the Xaa position: Ac-(Gly-Hyp-Thr)₁₀-NH₂ and Ac-(Gly-Hyp-Thr(β-Gal))₁₀-

NH₂, where Thr(β -Gal) refers to a threonine residue with a galactose moiety on its side chain (Bann & Bächinger, 2000). They compared the ability of these peptides to form triple helices with that of analogous peptides in which Pro replaces Hyp in the Xaa position. The T_m values for these triple helices are listed in Table 1.8. Hyp in the Xaa position is more stabilizing than is Pro. Most dramatically, Ac-(Gly-Pro-Thr)₁₀-NH₂ does not form triple helices to any appreciable extent, whereas Ac-(Gly-Hyp-Thr)₁₀-NH₂ has a T_m value of 19 °C. Glycosylation of the Thr residues adds additional stability. A possible explanation for the stabilization by Hyp in these peptides is that additional hydrogen bonds are forming with water, that Hyp increases the content of *trans* peptide bonds via an inductive effect (Eberhardt et al., 1996), or both. Why then does (Hyp-Pro-Gly)₁₀ not form stable triple helices? (Inouye et al., 1982) Bann and Bächinger suggested that having Hyp and Pro adjacent to each other forces the pyrrolidine rings to adopt unfavorable puckers, which could change the ψ torsion angle ($N_i-C_i^\alpha-C'_{i+1}-N_{i+1}$) to one that is unfavorable for triple-helix formation. They explained the additional triple-helix stabilization achieved upon glycosylation of the Thr residues by invoking a decrease in the activity of water surrounding the peptides, which in turn decreases the ability of water to form hydrogen bonds with main-chain amides and increases the favorability of amide–amide hydrogen bonds. Indeed, ¹H NMR studies of Ac-(Gly-Pro-Thr(β -Gal))_n-NH₂ (n=1, 5, 10) showed that only the peptide where n=10 formed triple helices (T_m =38 °C), and that the Thr-amide protons in the triple helix exchanged with deuterium an order of magnitude more slowly than did the Gly-amide protons (Bann et al., 2003). The authors observed that a triple helix of Ac-(Gly-Hyp-Thr)₁₀-NH₂ has a much larger ΔH_m° than does a triple helix of (Pro-Hyp-Gly)₁₀ or (Pro-Pro-Gly)₁₀. The strong enthalpic interaction

between adjacent Hyp and Thr is presumably due to a hydrogen bond between side-chain hydroxyl groups within a chain, between chains, or with water.

Table 1.8 *Effect of 4(R)-hydroxy-L-proline (Hyp) in the Xaa position on the conformational stability of triple helices (Bann & Bächinger, 2000).*

Peptide	T_m (°C) ^a
Ac-(Gly-Pro-Thr) ₁₀ -NH ₂	No transition
Ac-(Gly-Hyp-Thr) ₁₀ -NH ₂	19.2
Ac-(Gly-Pro-Thr(β-Gal)) ₁₀ -NH ₂	39.2
Ac-(Gly-Hyp-Thr(β-Gal)) ₁₀ -NH ₂	54.8

^a Values of T_m were measured in H₂O.

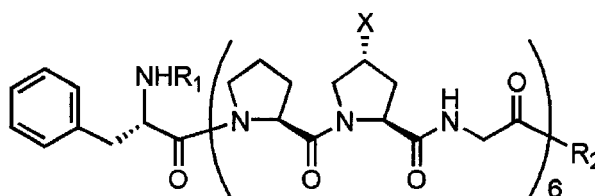
Further studies of peptides with the general structure Ac-(Gly-Hyp-Yaa)_n-NH₂, (where Yaa = Thr, *allo*-Thr, Val, Ser, or Ala) showed that only the peptides containing Thr and Val formed triple helices in water. The Thr- and Val-containing peptides also formed the most stable triple helices in 1,2- and 1,3-propanediol. Molecular modeling indicated that the methyl groups of these two peptides contributed to triple helix stability by providing hydrophobic contacts with the core of the triple helix, and by limiting access of water molecules to the Gly-NH···O=C-Hyp hydrogen bond (Mizuno et al., 2003).

Two groups recently reported studies on non-covalent heterotrimers composed of (Pro-Hyp-Gly)₁₀ and (Pro-Pro-Gly)₁₀. Slatter, Bailey, and coworkers concluded that the non-linear increase in T_m and denaturation enthalpy observed upon replacement of one or two chains of a [(Pro-Pro-Gly)₁₀]₃ triple helix with (Pro-Hyp-Gly)₁₀ could be explained by the introduction of axial asymmetry into the triple helix (this despite the essentially identical triple helical parameters of crystalline [(Pro-Pro-Gly)₁₀]₃ (Berisio et al., 2002)

and $[(\text{Pro-Hyp-Gly})_{10}]_3$ (Berisio et al., 2001)), thus allowing for greater access of water to the main-chain carbonyl groups (Slatter et al., 2003). Berisio, Zagari, and coworkers reported T_m values in agreement with those of Slatter and coworkers, but came to a different conclusion for the increase in triple helix stability with increasing numbers of $(\text{Pro-Hyp-Gly})_{10}$ chains in the triple helix: enthalpic effects play a minor role in the stabilization of collagen by Hyp; instead, Hyp preorganizes the peptide chains to minimize the entropy penalty of triple helix formation (Berisio et al., 2004).

1.3.2 *Aminoproline residues*

Babu and Ganesh reported on the synthesis and stability of collagen mimics containing 4(*R*)-amino-L-proline (Amp) residues in place of Hyp (**1.7-1.10**) (Babu & Ganesh, 2001). They found that triple helices containing Amp residues are more stable than are those containing Hyp, and that the differential stability depends upon pH. In some instances, the difference in T_m values is remarkable—over 30 °C—as is listed in Table 1.9. The effect of Amp residues is, however, a complex function of solution conditions. For example, protonating the Amp amino groups produces both more favorable inductive effects and unfavorable Coulombic interactions. Those unfavorable Coulombic interactions will be more pronounced in a solution of low salt concentration.



1.7: X=NH₂, R₁=H, R₂=OH

1.8: X=OH, R₁=H, R₂=OH

1.9: X=NH₂, R₁=Ac, R₂=NH₂

1.10: X=OH, R₁=Ac, R₂=NH₂

Table 1.9 Effect of 4(*R*)-amino-*L*-proline (*Amp*) on the conformational stability^a of triple helices (Babu & Ganesh, 2001).

peptide	pH ^b			
	3.0	7.0	9.0	12.0
1.7	42.5	31.3	18.5	56.6
1.8	23.0	21.6	15.5	39.6
1.9	60	56.5	26	49
1.10	27	28	27	27

^a Values are for T_m (°C).

^b Buffers: pH 3.0, 20 mM acetate; pH 7.0, 20 mM phosphate; pH 9.0 and 12.0, 20 mM borate. All buffers contained 0.1 M NaCl.

1.3.3 Fluoroproline residues

The 4(*R*) and 4(*S*) diastereomers of Flp were first synthesized by Witkop and coworkers in 1965 for the purpose of studying whether Flp can be incorporated into proteins by biosynthesis, and if so, whether the Flp incorporated into procollagen is defluorinated to yield Hyp (Gottlieb et al., 1965). In vivo studies were carried out by others in 1966 (Bakerman et al., 1966). Apparently, the 4(*S*)-Flp diastereomer inhibits protein synthesis to some extent, but is incorporated into proteins in place of Pro. The

4(*R*)-Flp diastereomer is incorporated into proteins to a larger extent and is converted subsequently to Hyp in collagen strands. Prockop and coworkers later showed that collagen containing 4(*S*)-Flp cannot be exported from cells (Takeuchi & Prockop, 1969; Takeuchi et al., 1969).

The first incorporation of 4(*R*)-Flp into a collagen mimic was reported by Raines and coworkers (Holmgren et al., 1998; Holmgren et al., 1999). The T_m values of triple helices of (Pro-Pro-Gly)₁₀, (Pro-Hyp-Gly)₁₀, and (Pro-Flp-Gly)₁₀ are listed in Table 1.10. Flp imparts remarkable stability to the collagen triple helix. Indeed, three strands of (Pro-Flp-Gly)₁₀ form the most stable collagen mimic of similar size known to date. Moreover, because organic fluorine does not form hydrogen bonds (Howard, J.A.K. et al., 1996; Dunitz & Taylor, 1997), these data confirm that an electron-withdrawing substituent in the 4(*R*) position of Pro can stabilize collagen by a means other than a network of water bridges (Engel & Prockop, 1998).

Table 1.10 *Effect of 4(*R*)-, 4(*S*)-fluoro-L-proline (Flp), 4(*R*)- and 4(*S*)-hydroxy-L-proline (Hyp) on the conformational stability of triple helices.*

Peptide	T_m (°C) ^a
(Pro-Pro-Gly) ₇	6–7 (Shaw & Schurr, 1975)
(Pro-4(<i>R</i>)-Hyp-Gly) ₇	36 (Bretscher et al., 2001)
(Pro-4(<i>R</i>)-Flp-Gly) ₇	45 (Bretscher et al., 2001)
(Pro-4(<i>S</i>)-Flp-Gly) ₇	<2 (Bretscher et al., 2001)
(Pro-Pro-Gly) ₁₀	41 (Holmgren et al., 1998)
(Pro-4(<i>R</i>)-Hyp-Gly) ₁₀	69 (Holmgren et al., 1998)
(Pro-4(<i>R</i>)-Flp-Gly) ₁₀	91 (Holmgren et al., 1998)
(Pro-4(<i>S</i>)-Hyp-Gly) ₁₀	<5 (Inouye et al., 1976)

^a Values of T_m were measured in dilute solutions of aqueous acetic acid.

The remarkable stability imparted by Flp to (Pro-Flp-Gly)₁₀ derives from the interplay of several factors (Bretscher et al., 2001), all of which arise from the inductive effect of the fluorine atom (Eberhardt et al., 1996). First, the *gauche* effect (O'Hagan et al., 2000) dictates the pyrrolidine ring pucker (Bretscher et al., 2001). The *gauche* effect arises when two vicinal carbons bear electronegative substituents. These electronegative substituents prefer to reside *gauche* (60°) to each other so that there is maximum overlap between the σ orbitals of more electropositive substituents, such as hydrogen, and the σ^* orbitals of the electronegative substituents, as shown in Fig. 1.4. As expected from the manifestation of the *gauche* effect, the C ^{γ} -exo ring pucker is predominant in Hyp residues in the Yaa position of collagen-like peptides (Berisio et al., 2001; Vitagliano et al., 2001b), as well as in small-molecule structures of AcHypOMe and AcFlpOMe (Panasik et al., 1994). O'Hagan and colleagues showed that the fluorine–amide *gauche* effect is especially strong (O'Hagan et al., 2000).

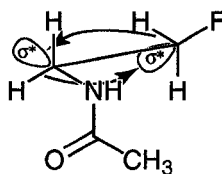


Figure 1.4 The *gauche* effect in AcNHCH₂CH₂F (O'Hagan et al., 2000). The lobes opposite the C-N and C-F bonds represent σ^* antibonding orbitals, which overlap with the σ bonding orbitals of the indicated C-H bonds.

The C ^{γ} -exo ring pucker preorganizes the main-chain torsion angles of Flp residues. The ϕ angle is a function of ring pucker, as described above (Vitagliano et al., 2001b). Likewise, the ψ angle in crystalline AcFlpOMe is 141° (Panasik et al., 1994), which is

close to the $\psi = (150 \pm 9)^\circ$ found for residues in the Yaa position of collagen (Bella et al., 1994). This value of ψ provides a favorable geometry for an interaction between a non-bonding electron pair of an amide oxygen (O'_{i-1}) and the π antibonding orbital of an amide carbon (C'_i), as shown in Fig. 1.5. The $O \cdots C=O$ angle in AcFlpOMe is 98° , which is close to the ideal angle for such an $n \rightarrow \pi^*$ interaction (Bürgi, H. B. et al., 1973; Bürgi, H.B. et al., 1974a; Bürgi, H.B. et al., 1974b). Moreover, the $O \cdots C=O$ distance in AcFlpOMe is only 2.76 Å, which predicates a meaningful interaction. Indeed, the ester carbonyl stretching vibration is lower by 6 cm^{-1} in Ac-4(R)-Flp-OMe than in Ac-4(S)-Flp-OMe, presumably because the $n \rightarrow \pi^*$ interaction decreases the C=O bond order.

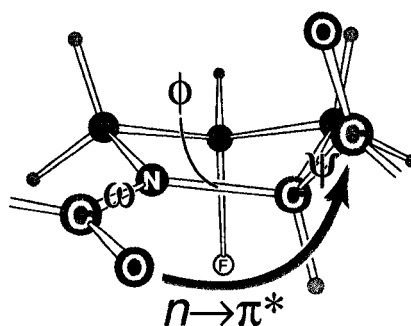


Figure 1.5 Main-chain ω , ϕ , and ψ torsion angles of a Flp residue. The gauche effect (Fig. 1.4) fixes the pyrrolidine ring pucker (C' -exo). In that ring pucker, the ω , ϕ , and ψ angles are preorganized at values close to those in the Yaa position of a collagen triple helix (Bretscher et al., 2001). The indicated $n \rightarrow \pi^*$ interaction contributes to that preorganization.

The $n \rightarrow \pi^*$ interaction stabilizes not only the ideal ψ angle for triple-helix formation, but also the requisite *trans* conformation ($\omega = 180^\circ$) of the Flp peptide bond. In the *cis* conformation ($\omega = 0^\circ$), C'^{α}_{i-1} rather than O'_{i-1} would be proximal to C'_i , and no $n \rightarrow \pi^*$ interaction can occur. Accordingly, as the electronegativity of the substituent in the 4(R)-

position increases, the *trans/cis* ratio of the amide bond also increases (Eberhardt et al., 1996). The reverse trend is true for electronegative 4(*S*) substituents, which impose a C^γ-endo pucker (DeRider et al., 2002) and do not stabilize triple helices in the Yaa position (Inouye et al., 1976; Bretscher et al., 2001). The association of ω angle with pyrrolidine ring pucker provides an explanation for the observation that *cis* prolyl peptide bonds tend to have *endo* ring puckers in crystalline proteins (Milner-White et al., 1992). In summary, 4(*R*)-Flp in the Yaa position stabilizes collagen by a stereoelectronic effect—the *gauche* effect—which fixes the pyrrolidine ring pucker and thus preorganizes all three main-chain torsion angles: ϕ , ψ , and ω (Fig. 5). This same reasoning applies to Hyp and Amp residues.

Ab initio calculations (DeRider et al., 2002) suggested that 4(*S*)-Flp has ϕ and ψ angles that should be favorable for triple helix formation when incorporated into the Xaa position of a collagen mimic. Indeed, (4(*S*)-Flp-Pro-Gly)_n is significantly more stable than (Pro-Pro-Gly)_n and (4(*R*)-Pro-Gly)_n ($n = 7$ (Hodges & Raines, 2003) or 10 (Doi et al., 2003), see Table 1.11), lending further support to the hypothesis that fluoroproline residues stabilize triple helices through stereoelectronic effects.

Table 1.11 *Melting temperatures of collagen mimics containing fluoroproline in the Xaa position.*

Peptide	n	T_m (°C) ^a	reference
(4(<i>S</i>)-Flp-Pro-Gly) _n	7	33	(Hodges & Raines, 2003)
	10	58	(Doi et al., 2003)
(Pro-Pro-Gly) _n	7	6-7	(Shaw & Schurr, 1975)
	10	34	(Kobayashi et al., 1970; Uchiyama et al., 1997)
(4(<i>R</i>)-Flp-Pro-Gly) _n	7	<4	(Hodges & Raines, 2003)
	10	<4	(Doi et al., 2003)

^a Values of T_m were measured in dilute aqueous acetic acid.

An interesting counterpoint to the above studies was reported by Brodsky and coworkers (Persikov et al., 2003). They incorporated the guest triplet (Pro-Flp-Gly) into the central triplet of a (Pro-Hyp-Gly)₈ host and found that the T_m *decreased* by 3.6 °C. In comparison, the guest triplet (Hyp-Pro-Gly) gave a T_m 4.3 °C lower than the parent triple helix, while the guest triplet (Hyp-Hyp-Gly) gave a T_m equal to that of the parent triple helix. Their rationalizations for these observations included invoking a different hydration scheme around Flp than around Hyp residues and suggesting that there may be a different mechanism of stabilization in a homogeneous tripeptide than in a host-guest system.

Moroder and coworkers have used Flp in another context (Renner et al., 2001). They incorporated 4(*R*)-Flp and 4(*S*)-Flp residues in place of a Pro residue with a *cis* peptide bond in barstar. They observed that 4(*S*)-Flp, which favors the *cis* conformation more than does Pro, stabilizes the protein and 4(*R*)-Flp, which favors the *trans* conformation

more than does Pro, destabilizes the protein. Thus, 4(*R*)- and 4(*S*)-Flp residues can be useful tools for protein engineers.

1.4 Collagen as a biomaterial

Collagen is an important biomaterial (Werkmeister & Ramshaw, 1992; Ramshaw, J. A. et al., 1996). For example, collagen is the principal component of biodegradable sutures and artificial heart valves. Some of the advantages of using collagen as a biomaterial are its low immunogenicity and high durability. One disadvantage is that it is difficult to obtain collagen in high purity without degrading its structural integrity. In addition, chemical methods used to cross-link purified collagen can cause cytotoxicity. Furthermore, the collagen that is used most often as a biomaterial is bovine collagen, which can engender allergic and immunological side effects in humans, as well as other health risks. Several different constructs for producing recombinant human fibrillar collagens have been reported. For example, human procollagen II has been isolated from a stably transfected human tumor cell line (Fertala et al., 1994) and from yeast (Myllyharju et al., 2000); human type III collagen has been expressed in baculovirus (Lamberg et al., 1996) and yeast (Myllyharju et al., 2000) systems; and human type I collagen has been produced in transgenic tobacco plants (Ruggiero et al., 2000), mouse mammary glands (Bulleid et al., 2000), and two different yeast strains (Myllyharju et al., 2000; Toman et al., 2000; Olsen et al., 2001).

Despite the numerous studies on collagen mimics, few have been tested as biomaterials. One report by Goodman and coworkers has provided encouraging results (Johnson, G. et al., 2000). In this report, collagen mimics containing Nleu were tested for

the ability to inhibit fibroblast or epithelial cell attachment to polystyrene. The results are listed in Table 1.12. Triple-helical (Gly-Pro-Nleu)_n, but neither (Gly-Nleu-Pro)_n nor (Gly-Pro-Hyp)_n, inhibited cell attachment. It is important to note that none of the peptides tested showed any cytotoxicity. Hence, (Gly-Pro-Nleu)-containing sequences show promise for development into biomaterials.

Table 1.12 *Cytotoxicity and cell-binding of triple helices.* (Johnson, G. et al., 2000)

Peptide	n	Cytotoxicity	Inhibition of Cell Attachment ^c	
			Epithelial cells	Fibroblasts
(Gly-Pro-Nleu) _n -NH ₂	1	—	ND	—
	2	—	ND	ND
	3	—	—	—
	5	—	—	—
	7	—	ND	ND
(Gly-Pro-Nleu) _n -Gly-Pro-NH ₂ ^a	10	—	+	+
(Gly-Pro-Nleu) _n -Gly-Pro-NH ₂ ^b	10	—	+	ND
Ac-(Gly-Pro-Nleu) _n -NH ₂	1	—	—	—
	9	—	+	+
(Gly-Nleu-Pro) _n -NH ₂	10	—	—	ND
Ac-(Gly-Nleu-Pro) _n -NH ₂	3	—	ND	—
	6	—	ND	—
	10	—	insoluble	insoluble
(Gly-Pro-Hyp) _n -NH ₂	9	—	—	—
KTA-[Gly-(Gly-Pro-Nleu) _n -NH ₂] ₃	9	—	+	+
KTA-[Gly-(Gly-Pro-Hyp) _n -NH ₂] ₃	5	—	ND	—
RGES		ND	—	—
GRGDSPK		ND	+	+

^a Peptide was left in solution at 4 °C for a minimum of 7 days before use.

^b Peptide solution was used immediately after preparation.

^c +, inhibition of cell attachment; —, no inhibition; ND, not determined.

Another report by Fields and coworkers (Malkar et al., 2002) showed that the addition of a single 4-fluoroproline residue to a collagen-mimic peptide can have a significant effect on its triple helical stability and thus on its ability to affect cellular responses. They synthesized a series of triple-helical peptides incorporating a sequence from type IV collagen known to promote cell adhesion, spreading, migration, and signal transduction. The natural sequence was flanked with (Gly-Pro-Hyp)₄ repeats at both the N- and C-termini in the parent peptide. In the test peptides, the last Hyp in the (Gly-Pro-Hyp) repeats was replaced by either 4(*R*)- or 4(*S*)-Flp. The peptide containing 4(*R*)-Flp had a higher T_m than the parent peptide (47.0 vs. 42.0 °C), whereas the peptide containing 4(*S*)-Flp had a significantly lower T_m of 35.5 °C. Cell adhesion and spreading assays performed with these peptides showed a positive correlation with triple helix stability; i.e., there was greater cell adhesion and spreading on a surface of the 4(*R*)-Flp-containing peptide than on either of the other two peptides. Fields and coworkers concluded that judicious inclusion of fluoroproline residues could help modulate triple helix stability and help create new collagen ligands, substrates, and biomaterials.

1.5 Envoi

The appearance of high-resolution structures of crystalline collagen triple helices has led to a resurgence of interest in chemical aspects of this ubiquitous protein. With these structures as a guide, chemists, biochemists, and biophysicists have made much progress in delineating the forces responsible for the conformational stability of the triple helix. Still, important questions remain without answers. For example, how much does the ladder of XaaC=O⋯HNGly hydrogen bonds (Fig. 1.1) contribute to stability? In what

contexts does Hyp or Amp in the Xaa position stabilize a triple helix? Do changes in helical pitch affect the results of host–guest studies? Does the rare 3-hydroxy-L-proline residue, which is subject to a *gauche* effect in its pyrrolidine ring, contribute to conformational stability? Which template provides the most accurate collagen mimic? Does greater triple-helix stability translate to greater fibril stability? And, most importantly, how can we use our knowledge of the collagen triple helix to solve real problems in biomedicine?

Chapter 2*

EFFECT OF 3-HYDROXYPROLINE RESIDUES ON COLLAGEN STABILITY

2.1 Introduction

Collagen is the most abundant protein in animals, comprising approximately one-third of the total protein by weight (Prockop & Kivirikko, 1995; Myllyharju & Kivirikko, 2001). Collagen has a unique tertiary structure, which consists of three parallel left-handed polyproline II-type strands wound tightly around a common axis (Ramachandran & Kartha, 1954, 1955; Rich & Crick, 1955, 1961). Maintenance of this triple helix is essential for many biological functions (Fields, G.B., 1995).

Each strand of collagen consists of the repeating sequence: Xaa–Yaa–Gly, where Xaa and Yaa are often proline (Pro) residues. The post-translational hydroxylation of some of these proline residues generates 4(*R*)-hydroxy-L-proline (4-Hyp) residues and, to a lesser extent, 3(*S*)-hydroxy-L-proline (3-Hyp) residues. 3-Hyp, which is found in the natural triplet (3-Hyp–4-Hyp–Gly) (Gryder et al., 1975), arises from the action of the enzyme prolyl-3-hydroxylase rather than by regiochemical ambiguity by the enzyme prolyl-4-hydroxylase (Tryggvason et al., 1976).

* This chapter has been previously published under the same title. Reference: Jenkins, C. L.; Bretscher, L. E.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2003**, *125*, 6422-6427.

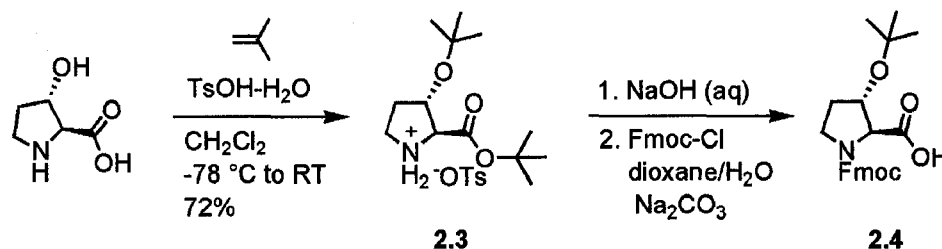
In seminal work, Prockop and coworkers demonstrated that 4-Hyp residues enhance greatly the conformational stability of the collagen triple helix (Berg & Prockop, 1973). This additional stability arises from stereoelectronic effects that preorganize a 4-Hyp residue in a conformation that befits a triple helix (Engel & Prockop, 1998; Holmgren et al., 1998; Bretscher et al., 2001; Jenkins & Raines, 2002). In contrast to 4-Hyp, no information is available on the effect of 3-Hyp residues on collagen stability.

Here, we synthesize collagen-like peptides that contain a 3-Hyp residue. We use these peptides to assess the contribution of 3-Hyp to the conformational stability of the collagen triple helix. We find that the presence of a 3-Hyp residue, in surprising contrast to a 4-Hyp residue, *destabilizes* a collagen triple helix. Like the stability endowed by 4-Hyp residues, the instability imposed by 3-Hyp residues appears to arise largely from inductive effects of its pendant hydroxyl group.

2.2 Results and Discussion

Synthetic collagen mimics have been effective in dissecting the basis for the conformational stability of the collagen triple helix (Fields, G.B. & Prockop, 1996). In particular, host–guest studies have revealed important insights on the contribution of individual amino acid residues, both natural and nonnatural (Shah et al., 1996; Ramshaw, J. A. M. et al., 1998; Kwak et al., 1999; Persikov et al., 2000a; Persikov et al., 2000b; Kersteen & Raines, 2001). To reveal the effect of 3-Hyp residues on triple-helical stability, we synthesized two host–guest peptides in which the central triplet of (Pro–4-Hyp–Gly)₇ was replaced with one containing 3-Hyp. These peptides were (Pro–4-Hyp–Gly)₃–3-Hyp–4-Hyp–Gly–(Pro–4-Hyp–Gly)₃ (**2.1**) and (Pro–4-Hyp–Gly)₃–Pro–3-Hyp–

Gly-(Pro-4-Hyp-Gly)₃ (**2.2**). Peptide **2.1** was designed to elucidate the role of the 3-Hyp-4-Hyp-Gly triplet as is found most often in natural collagen. Peptide **2.2** was designed to reveal the effect of the regiochemistry (3-Hyp *versus* 4-Hyp) of the pendant hydroxyl group. These peptides were synthesized by standard Fmoc/*t*Bu coupling strategies using compound **2.4** (Scheme 2.1) to introduce 3-Hyp residues.



Scheme 2.1 Synthesis of Fmoc-3(S)-Hyp(OtBu) (**2.4**).

2.2.1 Conformational Stability.

The unfolding of triple helices of peptides **2.1** and **2.2** was monitored by CD spectroscopy. The resulting T_m values, along with those of triple helices of (Pro-4-Hyp-Gly)₇ (**2.5**) and (Pro-4-Hyp-Gly)₃-Pro-Pro-Gly-(Pro-4-Hyp-Gly)₃ (**2.6**), are listed in Table 2.1. The values confirm that in its natural (*i.e.*, Yaa) position, 4-Hyp provides more conformational stability to a triple helix than does Pro. In contrast, 3-Hyp provides less conformational stability than does Pro in either its natural (*i.e.*, Xaa) or a nonnatural (*i.e.*, Yaa) position. The additional stability endowed by 4-Hyp is known to arise from stereoelectronic effects (Bretscher et al., 2001; Jenkins & Raines, 2002). What is the origin of the detrimental effect of 3-Hyp on collagen stability?

Table 2.1 Values of T_m for Synthetic (Pro-Hyp-Gly)₃-(Xaa-Yaa-Gly)-(Pro-Hyp-Gly)₃ Triple Helices

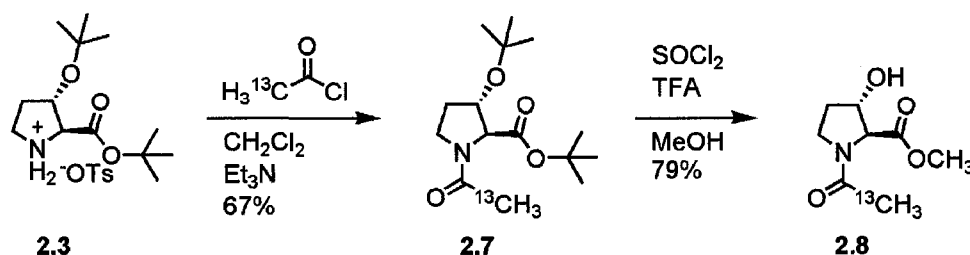
peptide	Xaa-Yaa-Gly	T_m (°C) ^a
2.5	Pro-4-Hyp-Gly	36 ± 2^b
2.1	3-Hyp-4-Hyp-Gly	32.7 ± 0.9
2.6	Pro-Pro-Gly	30.5 ± 2.2^b
2.2	Pro-3-Hyp-Gly	21.0 ± 2.2

^a Values of T_m were determined by CD spectroscopy for peptides (0.2 mM) in 50 mM acetic acid, and are the average (\pm SE) of at least 3 determinations.

^b From (Kersteen & Raines, 2001).

2.2.2 Peptide Bond Isomerization

All of the peptide bonds in triple-helical collagen are in the *trans* (i.e., *Z*) conformation. The 4-hydroxyl group of a 4-Hyp residue increases the *trans/cis* ratio of an Xaa-4-Hyp peptide bond by approximately 50% compared to a Xaa-Pro bond (Eberhardt et al., 1996). We used NMR spectroscopy to determine the effect of the 3-hydroxyl group of a 3-Hyp residue on its *trans/cis* ratio. We found that the *trans/cis* ratio of amide **2.8** (synthesized according to Scheme 2.2) and Ac-Pro-OMe in D₂O were 4.9 and 4.6, respectively. These similar values indicate that peptide bond isomerization makes a negligible contribution to the instability of triple-helical **2.1** and **2.2**.



Scheme 2.2 Synthesis of **2.8**.

2.2.3 Structure of a 3-Hyp Residue.

We used X-ray diffraction analysis to determine the structure of crystalline amide **2.8** (Figure 2.1A). The pyrrolidine ring of crystalline amide **2.8** is puckered such that its N–C^α–C^β–O^γ bond has a dihedral angle of $-81.46 \pm 0.16^\circ$, as would be expected from the manifestation of a *gauche* effect (O'Hagan et al., 2000; Bretscher et al., 2001). The ring pucker in crystalline Ac–3-Hyp–OMe (**2.8**) differs significantly from the ring pucker in crystalline Ac–4-Hyp–OMe (Figure 1B). The conformation of the pyrrolidine ring in amide **2.8** is intermediate between a ¹E envelope and ¹E₂ twisted conformation (Giacovazzo et al., 2002). In the envelope conformation, the flap atom is C_i^β. In the twisted conformation, atoms N_i, C_i^β, and C_i^δ form the basal plane. Atom C_i^β resides 0.456 ± 0.004 Å above that plane, and atom C_i^δ resides 0.153 ± 0.005 Å below that plane. The phase angle $\phi_2 = 9.6 \pm 0.3^\circ$, and the puckering amplitude is $q_2 = 0.378 \pm 0.002$ Å.

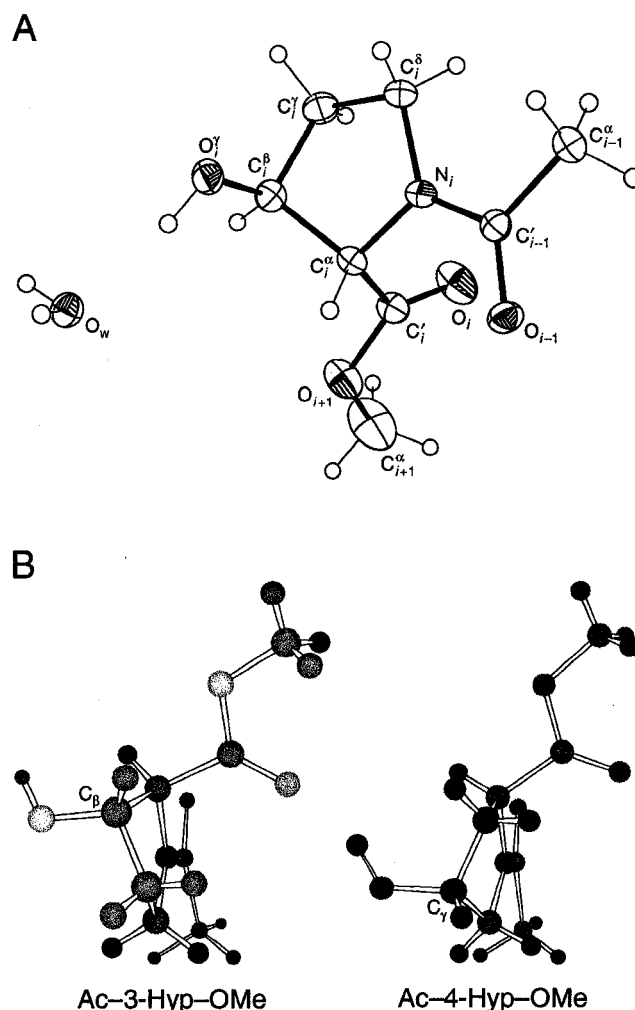


Figure 2.1 (A) ORTEP diagram of crystalline N -($^{13}\text{C}_2$ -acetyl)-3(S)-hydroxy- L -proline methyl ester (**2.8**) drawn with 30% probability ellipsoids. (B) Pyrrolidine ring pucker in crystalline amide **2.8** (left) and crystalline N -acetyl-4(R)-hydroxy- L -proline methyl ester (right) (Panasik et al., 1994).

The structure of crystalline Ac-4-Hyp-OMe contains two symmetry-independent molecules (Panasik et al., 1994). (Thus, there are two numbers herein for each parameter.) In both molecules, the pyrrolidine ring is in the $^5\text{T}_4$ (that is, $^{\text{C}_i^{\beta}}\text{T}_{\text{C}_i^{\gamma}}$) twisted conformation with characteristic phase angle of $\phi_2 = 129.8 \pm 0.7^\circ$ ($127.8 \pm 0.7^\circ$) and puckering amplitude of $q_2 = 0.354 \pm 0.004 \text{ \AA}$ ($0.389 \pm 0.006 \text{ \AA}$) (Giacovazzo et al.,

2002). The basal plane is determined by atoms N_i , C_i^δ , and C_i^α . Atom C_i^β resides 0.28 ± 0.01 Å (0.22 ± 0.01 Å) above that plane, and atom C_i^γ resides 0.35 ± 0.01 Å (0.36 ± 0.01 Å) below the plane.

The ring puckers of 3-Hyp and 4-Hyp (Figure 2.1B) serve to preorganize the $\phi(C'_{i-1}-N_i-C_i-C'_i)$ and $\psi(N_i-C_i-C'_i-N_{i+1})$ dihedral angles in a conformation that is appropriate for the Yaa position of a collagen triple helix (Vitagliano et al., 2001b; Berisio et al., 2002; DeRider et al., 2002). Likewise, a 3-Hyp residue would be stabilizing in the Xaa position relative to Pro because its main-chain dihedral angles are favorable for that position, whereas a 3-Hyp residue would be destabilizing in the Yaa position because its dihedral angles differ significantly from the optimal (Table 2.2). Although 3-Hyp is destabilizing in both positions (Table 2.1), the preorganization that arises from the *gauche* effect does favor a triple helix with 3-Hyp in the Xaa rather than the Yaa position. Unlike 3-Hyp and 4-Hyp residues, a Pro residue is not constrained by a *gauche* effect to adopt a particular pucker, and can therefore accommodate the different ϕ and ψ angles of the Xaa and Yaa positions in a triple helix (Table 2.1).

Table 2.2 Main-Chain Angles of Proline and Hydroxyproline residues in Crystalline Amides and Triple Helices

angle	amide		triple helix			
	Ac-4-Hyp-OMe ^a	[¹³ CH ₃]Ac-3-Hyp-OMe	Pro (Xaa) ^b	Pro (Yaa) ^b	Pro (Xaa) ^c	Hyp (Yaa) ^c
ϕ	-50.9	-79.5	-75	-60	-69.8	-57.4
ψ	145.2	163.7	164	152	162	149.8

^a From (Panasik et al., 1994).

^b From (Berisio et al., 2002).

^c From (Berisio et al., 2001).

To show that hydroxyproline residues have very similar conformations in small molecule as well as peptide contexts we compared the dihedral angles of Ac-4-HypOMe with those of 4-Hyp in the crystal structure 1CAG by pseudorotational analysis (Chacko et al., 1983). We found that the two rings were essentially superimposable, indicating that Hyp ring puckers are not dependent on context, and validating our assumption that 3-Hyp would have the same pucker in a peptide context as we found it to have in amide **2.8**.

2.2.4 Steric Effect on Collagen Stability.

Could steric effects contribute to the instability imparted by 3-Hyp? In a collagen triple helix, residues in the Xaa position are more solvent-exposed than are those in the Yaa position (Jones & Miller, 1991), and are therefore less likely to introduce unfavorable steric interactions. Indeed, residues with bulky side chains are found only rarely in the Yaa position of natural collagen (Ramshaw, J. A. M. et al., 1998). A steric clash between 3-Hyp in the Yaa position with residues in neighboring strands (Figure 2.2, bottom) is likely to diminish the stability of triple-helical **2.2**. This steric clash is absent in triple-helical **2.1** (Figure 2.2, top).

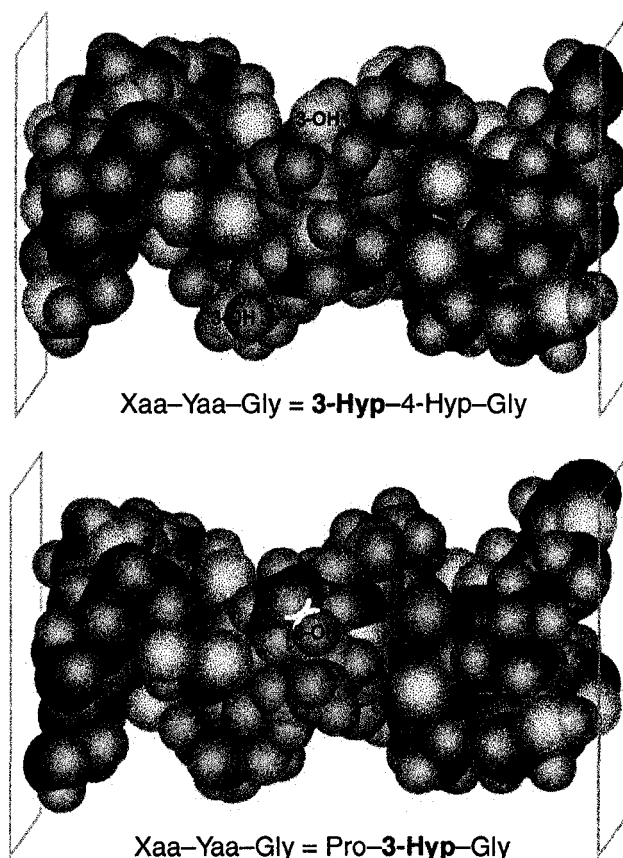


Figure 2.2 Model of a segment of the triple helix formed by peptide 2.1 (top) and peptide 2.2 (bottom). Each 3-Hyp residue (green), the oxygen of its hydroxyl group (red; labeled), and the steric clash in triple-helical peptide 2.2 (yellow) are indicated.

2.2.5 Effect of Hydrogen Bonds on Collagen Stability.

In a collagen triple helix, the C=O of the residue in the Xaa position accepts a hydrogen bond from the glycine N-H of another strand (Figure 3). The carboxyl pK_a of XaaOH is a measure of the ability of the Xaa residue to accept such a hydrogen bond (Holmgren et al., 1999). The pK_a values of the carboxyl groups of 3-HypOH and 4-HypOH were determined by monitoring the effect of pH on the ^1H NMR chemical shift of their α -protons. The carboxyl pK_a of 4-HypOH was found to be 1.80, which is close to the value of 1.82 reported previously (Fasman, 1989), and that of ProOH is 1.95

(Fasman, 1989). In contrast, the pK_a of 3-HypOH is 1.62, making the 3-Hyp residue a weaker hydrogen bond acceptor than Pro. This attribute is likely to contribute to the instability conferred by a 3-Hyp residue in the Xaa position.

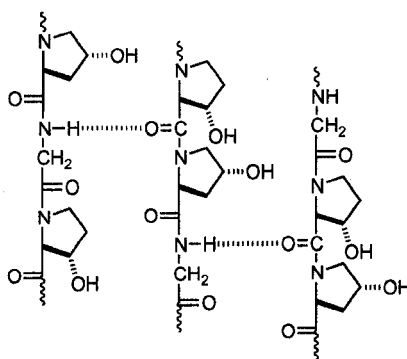


Figure 2.3 Putative interstrand hydrogen bonds in triple-helical (3-Hyp-4-Hyp-Gly)_n.

2.3 Biological Implications.

3-Hyp only decreases the T_m value of triple-helical **2.2** by 3 °C, compared to ≥ 5 °C for other residues in the Xaa position of (Pro-Hyp-Gly)-based host-guest triple helices (Persikov et al., 2000a). Thus, the insertion of 3-Hyp could serve to modulate the local stability of triple helices. This idea is supported by the 10-fold increase of 3-Hyp residues in basement-membrane collagens, as compared to fibrillar collagens (Kefalides, 1973). Additional support for this idea has come from Bachinger and coworkers, who showed that a collagen model peptide with 3-Hyp at every Xaa position does not fold into a triple helix (Mizuno et al., 2004). Apparently, the cumulative destabilizing effect of a large number of 3-Hyp residues is quite large. Collagen triple helices in basement membranes interact with each other, as well as with other biomolecules, in a more varied and complex manner than do those in fibrils. This network of interactions could require

regions of finely tuned conformational stability, and 3-Hyp residues could provide that tuning.

2.4 Experimental Section

General. Reagents were obtained from Aldrich Chemical (Milwaukee, WI) or Fisher Scientific (Hanover Park, IL) and used without further purification. Amino acids and their derivatives were obtained from Fisher Scientific, Bachem Bioscience (King of Prussia, PA), or Novabiochem (San Diego, CA). Dichloromethane was distilled over $\text{CaH}_2(\text{s})$ or drawn from a Baker Cycletainer. Thin-layer chromatography was performed by using aluminum-backed plates coated with silica gel containing F_{254} phosphor and visualized by UV illumination or staining with I_2 , *p*-anisaldehyde stain, or phosphomolybdic acid stain. NMR spectra were obtained with Bruker AC-300 and Varian UNITY-500 spectrometers.

3(*S*)-*tert*-Butoxy-L-Proline *tert*-Butyl Ester Tosylate (2.3). N-(9*H*-Fluoren-9-ylmethoxycarbonyl)-3-*tert*-butoxy-L-proline (2.3) was synthesized by the route in Scheme 2.1. 2-Methyl-propene (approximately 10 mL) was condensed into a pear-shaped flask at $-78\text{ }^\circ\text{C}$ then added to a suspension of 3-hydroxyproline (0.50 g, 3.8 mmol) and *p*-toluenesulfonic acid hydrate (3.00 g, 15.8 mmol) in dichloromethane (25 mL) at $-78\text{ }^\circ\text{C}$. The reaction was stirred for 3 d, allowing the mixture to come to room temperature. The reaction mixture was then cooled to $0\text{ }^\circ\text{C}$, vented carefully, and then poured into a separatory funnel and washed twice with a saturated aqueous solution of NaHCO_3 . The combined aqueous layers were extracted once with dichloromethane, and the combined organic layers were then washed once with water, dried over $\text{MgSO}_4(\text{s})$, filtered, and

concentrated under reduced pressure to yield compound **2.3** as a yellow oil (1.142 g, 72%). ^1H NMR (300 MHz, CDCl_3) δ (ppm) 7.78 (d, $J = 8.1$ Hz, 2H), 7.14 (d, $J = 8.1$ Hz, 2H), 4.29 (m, 1H), 3.99 (d, $J = 2.9$ Hz, 1H), 3.50–3.43 (m, 1H), 3.37–3.27 (m, 1H), 2.33 (s, 3H), 2.00–1.92 (m, 1H), 1.83–1.76 (m, 1H), 1.46 (s, 9H), 1.18 (s, 9H).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-3(*S*)-*tert*-Butoxy-L-Proline (**2.4**).**

Compound **2.3** (412 mg, 0.99 mmol) was suspended in 1.0 M NaOH (10 mL, 10 mmol) and stirred until the reaction mixture became homogeneous, that is, for about 4 h. The reaction mixture was then cooled to 0 °C and acidified with concentrated HCl to pH 0. Na_2CO_3 (1.1 g) was added immediately, followed by dioxane (25 mL). Fmoc-Cl (292 mg, 1.13 mmol) was added in small portions, and the resulting mixture was stirred at room temperature for 4 h. The reaction mixture was then poured into a separatory funnel with 50 mL H_2O and extracted with diethyl ether (3×25 mL). The aqueous layer was then combined with 50 mL ethyl acetate in a beaker, cooled to 0 °C, and acidified with concentrated aqueous HCl to pH 0. The layers were separated, and the aqueous layer was extracted again with EtOAc (25 mL). The combined ethyl acetate layers were dried over $\text{MgSO}_4(\text{s})$ and filtered, and the solvent was removed under reduced pressure to yield compound **2.4** as a white solid (0.148 g, 36.6% over two steps). ^1H NMR (two rotamers, 500 MHz, CDCl_3) δ (ppm) 7.75 and 7.70 (d, $J = 7$ and 7.5 Hz, 2H), 7.60 and 7.55 (dd, $J = 8, 8$ Hz and 7.5, 10 Hz, 2H), 7.41–7.27 (m, 4H), 4.68–4.12 (m, 5H), 3.76–3.61 (m, 2H), 2.12–2.01 (m, 1H), 1.89–1.80 (m, 1H), 1.22 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3 , two rotamers): δ (ppm) = 175.52, 174.12, 155.92, 154.62, 144.00, 143.84, 143.74, 141.26, 127.71, 127.57, 127.05, 126.96, 125.15, 125.06, 124.96, 119.95, 119.86, 75.26, 75.10,

73.67, 67.94, 67.62, 67.34, 66.05, 65.93, 61.15, 61.13, 47.12, 45.46, 45.30, 33.09, 32.23, 28.08. ESI-MS: 408.1805 (observed), 408.1811 (calculated).

To determine whether epimerization of the α -carbon had occurred upon saponification of the *t*-butyl ester **2.3**, the hydroxyl group was deprotected by acidolysis of its *t*-butyl protecting group and the resulting hydroxyl group was converted into its Mosher ester. Briefly, Fmoc-3-HypOH (17 mg, 0.048 mmol) was dissolved in 1 mL dioxane. K_2CO_3 (20 mg, 0.14 mmol) was added to this solution, followed by H_2O (2 drops). A Mosher acid chloride, *R*-(-)- α -methoxy- α -(trifluoromethyl)phenylacetylchloride (0.010 mL, 0.053 mmol), was added, and the resulting mixture was stirred at room temperature for 15 min. EtOAc (1 mL) was then added to the reaction mixture, followed by concentrated HCl (2 drops) and H_2O (1 mL). The layers were mixed thoroughly and then separated. The aqueous layer was extracted twice more with EtOAc, and the combined organic extracts were dried over $MgSO_4(s)$, filtered, and concentrated to a colorless oil. The ^{19}F NMR spectrum of the oil contained only a single peak at -72 ppm, indicating that compound **2.4** had maintained its stereochemical integrity.

N- $^{13}C_2$ -Acetyl-3(*S*)-*tert*-Butoxy-L-Proline *tert*-Butyl Ester (**2.7**). [$^{13}CH_3$]Ac-3-Hyp-OMe (**2.8**) was synthesized by the route shown in Scheme 2.2. The ^{13}C in the acetyl group facilitated the measurement of amide bond *trans/cis* ratios because both the acetyl and ester methyl groups overlap with other resonances in the 1H spectrum. Compound **2.3** (771 mg, 1.86 mmol) was dissolved in 25 mL dry dichloromethane and cooled to 0 °C. $^{13}CH_3COCl$ (200 μ L, 2.78 mmol) was added, followed by triethylamine (800 μ L, 5.74 mmol). The reaction was stirred at 0 °C for 1 h, then at room temperature

overnight. The reaction mixture was transferred to a separatory funnel and washed twice with water. The organic layer was dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure to a brown oil. The product was purified by silica gel chromatography, eluting with CH_2Cl_2 containing MeOH (5% v/v) to yield compound **2.7** as a clear, colorless oil (355 mg, 66.8%). ^1H NMR (300 MHz, CDCl_3) $\delta(\text{ppm})$ = major isomer: 4.28 (bs, 1H), 4.17 (m, 1H), 3.65 (m, 2H), 2.20–1.80 (m, 2H) 2.09 (d, $^1J_{\text{CH}} = 132$ Hz, 3H), 1.46 (s, 9H), 1.22 (s, 9H); minor isomer: $\delta(\text{ppm})$ 4.30 (m, 1H), 4.07 (s, 1H), 3.65 (m, 2H), 2.20–1.80 (m, 2H), 1.96 (d, $^1J_{\text{CH}} = 132$ Hz, 3H), 1.48 (s, 9H), 1.24 (s, 9H). ^{13}C NMR (75 MHz, CDCl_3 , 8 scans) $\delta(\text{ppm})$ = 22.10, 22.25 (^{13}C -labeled carbon).

***N*-($^{13}\text{C}_2$ -acetyl)-3(*S*)-Hydroxy-L-Proline Methyl Ester (2.8).** Compound **2.7** (355 mg, 1.24 mmol) was dissolved in MeOH. Trifluoroacetic acid (0.10 mL) was added, followed by dropwise addition of SOCl_2 (1.00 mL, 13.7 mmol). The reaction mixture was heated at reflux for 1 h, and then cooled to room temperature and concentrated under reduced pressure. The resulting oil was purified by chromatography on silica gel, eluting with CH_2Cl_2 containing MeOH (5% v/v) to yield amide **2.8** as a clear, colorless oil (183 mg, 79%). ^1H NMR (500 MHz, CDCl_3 , two rotamers) $\delta(\text{ppm})$ = major isomer: 4.43 (s, 1H), 4.38 (d, 4 Hz, 1H), 3.71–3.61 (m, 2H), 3.68 (s, 3H), 2.05 (d, $^1J_{\text{CH}} = 128.5$ Hz, 3H), 2.14–2.06 (m, 1H), 2.01–1.94 (m, 1H); minor isomer: 4.50 (d, 4 Hz, 1H), 4.33 (s, 1H), 3.73 (s, 3H), 3.71–3.61 (m, 2H), 1.94 (d, $^1J_{\text{CH}} = 128.5$ Hz, 3H); ^{13}C NMR (125 MHz, CDCl_3 , two rotamers): $\delta(\text{ppm})$ = 174.44, 174.03, 172.98, 75.08, 73.66, 69.46, 67.68, 54.12, 53.83, 46.72, 45.15, 32.64, 30.93, 21.84, 21.60.

$\text{p}K_{\text{a}}$ Determinations. The $\text{p}K_{\text{a}}$ values of the carboxylic acid groups of 3-HypOH and 4-HypOH were determined by ^1H NMR spectroscopy performed in aqueous solutions of

different pH. D₂O stock solutions consisted of 210 mM 3-HypOH (27.6 mg in 1.0 mL), 209 mM 4-HypOH (27.4 mg in 1.0 mL), 104 mM sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, 40.8 mg in 1.0 mL), 100 mM D₃PO₄ (68.5 μ L of 85% w/v D₃PO₄ in 10.0 mL), 100 mM DCl (156 μ L of 20% DCl in 10.0 mL), and 100 mM NaOD (67.6 μ L of fresh 40% w/v NaOD in 10.0 mL). Each sample contained 49 μ L of the 3-Hyp or 4-Hyp stock solution, 1 μ L of the DSS stock solution, and 950 μ L of buffer, according to Tables S1 and S2 in the Supporting Information. The pH of each solution was measured by using a standard pH electrode calibrated in non-deuterated buffers. ¹H NMR spectra were obtained (32 scans, 24 °C), and the chemical shift of the α -proton was determined for each sample. The resulting pH and chemical shift data were fitted to eq 2.1 to yield the pK_a values.

$$\delta = \frac{\delta_{\text{low}} + \delta_{\text{high}} 10^{(\text{pH} - \text{pK}_a)}}{1 + 10^{(\text{pH} - \text{pK}_a)}} \quad (2.1)$$

The acidity of an aqueous solution of a molecule having a titratable group with fraction factor near unity can be measured in D₂O by using a standard pH meter equilibrated against H₂O buffers and adding a correction factor of 0.4 units to the reading (Glasoe & Long, 1960). The difference in pK_a of amino acids in H₂O and D₂O is approximately 0.5 units (Hyman et al., 1960). Thus, a correction of about 0.1 units can be applied to a pK_a measurement obtained by measuring apparent pH values with a standard glass electrode. To determine a more precise correction factor we measured the pK_a of 3-HypOH and 4-HypOH in D₂O in the same manner, and found that the 4-HypOH value

was within 0.02 units of the value reported previously (Fasman, 1989). Thus, we did not apply any correction factor herein to the 3-Hyp pK_a value.

Crystallization of N-($^{13}\text{C}_2$ -Acetyl)-3(*S*)-Hydroxy-L-Proline Methyl Ester (2.8).

Equal portions of a solution of amide **2.8** (67 mg) in 3.0 mL dichloromethane were placed in six half-dram vials. Approximately 15 drops (Pasteur pipette) of a different cosolvent (hexanes, ethyl acetate, diethyl ether, tetrahydrofuran, *t*-butyl methyl ether) or no cosolvent was added to each vial. The vials were then capped loosely and allowed to sit at room temperature undisturbed for several days. The vial with ethyl acetate as a cosolvent contained the crystals most suitable for X-ray diffraction analysis.

X-Ray Diffraction Analysis. A colorless crystal with approximate dimensions $0.50 \times 0.50 \times 0.41 \text{ mm}^3$ was selected under oil under ambient conditions and attached to the tip of a glass capillary. The crystal was mounted in a stream of $\text{N}_2(\text{g})$ at $(173 \pm 2) \text{ K}$ and centered in the X-ray beam by using a video camera. Crystal evaluation and data collection were performed on a Bruker CCD-1000 diffractometer with Mo K_α ($\lambda = 0.71073 \text{ \AA}$) radiation and a diffractometer to crystal distance of 4.9 cm.

Initial cell constants were obtained from three series of ω scans at different starting angles. Each series consisted of 20 frames collected at intervals of 0.3° in a 6° range about ω with an exposure time of 10 s per frame. A total of 102 reflections were obtained. The reflections were indexed successfully by an automated indexing routine built into the SMART program. The final cell constants were calculated from a set of 1508 strong reflections from the actual data collection.

Data were collected by using the hemisphere data collection routine. The reciprocal space was surveyed to the extent of a full sphere to a resolution of 0.80 \AA . A total of 2911

data sets were harvested by collecting one set of frames with 0.3° scans in ω with an exposure time of 10 s per frame. This redundant data set was corrected for Lorentz and polarization effects. The absorption correction was based on fitting a function to the empirical transmission surface as sampled by multiple equivalent measurements (Blessing, 1995). All software and sources of the scattering factors were contained in the SHELXTL (version 5.1) program library (G. Sheldrick, Bruker Analytical X-Ray Systems, Madison, WI).

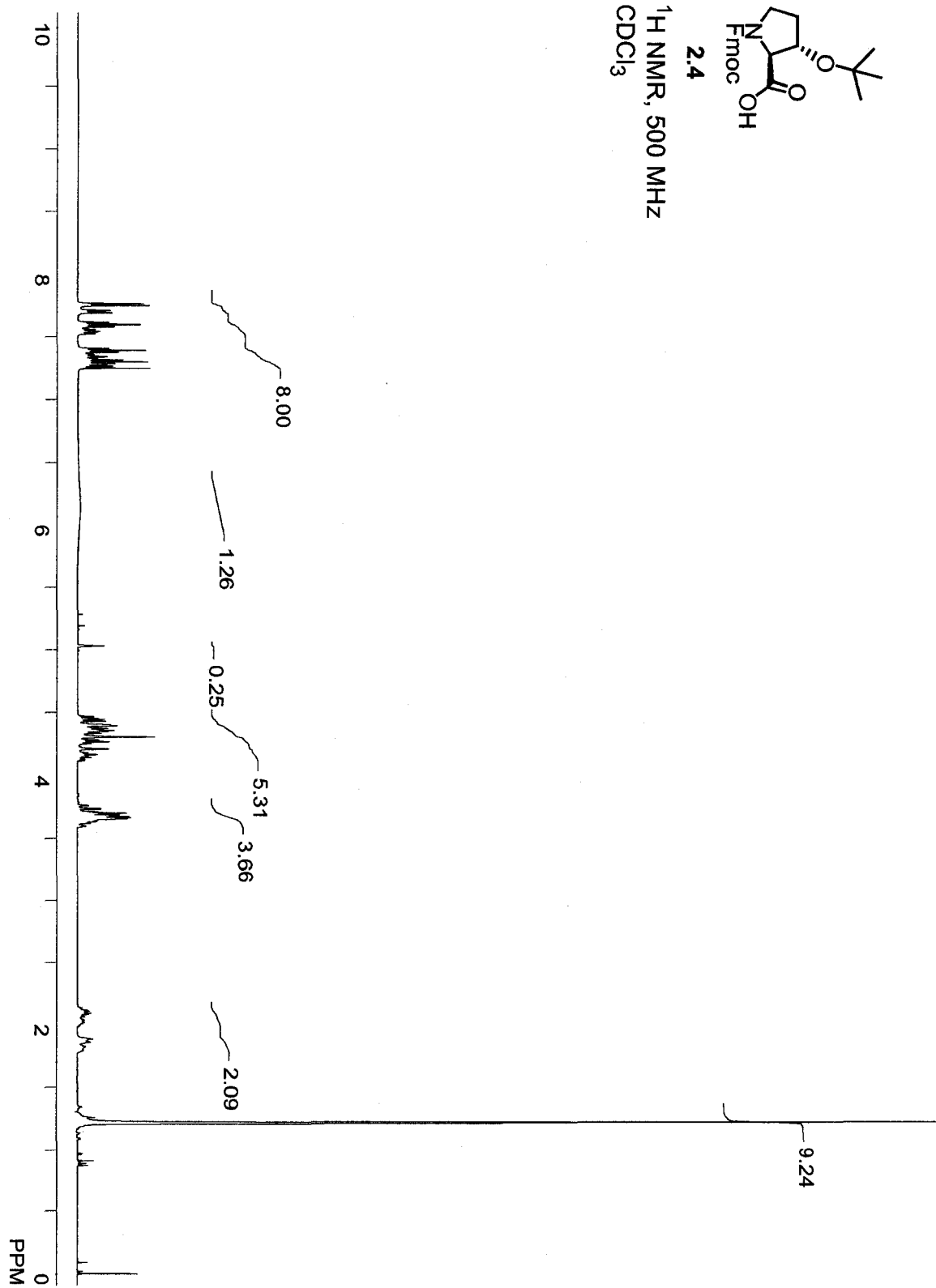
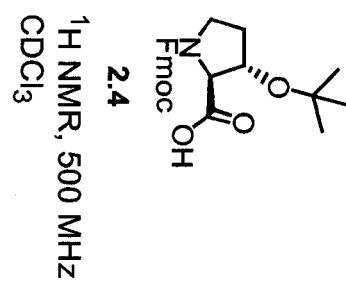
Synthesis of Peptides 2.1 and 2.2. Peptide synthesis was performed on an Applied Biosystems Synergy synthesizer using standard Fmoc/tBu methodology and HBTU as the coupling reagent. 3-Hyp residues were incorporated by using compound 2.4. Mass spectra of the synthetic peptides were obtained with a Perkin–Elmer Voyager MALDI–TOF or Micromass LCT ESI instrument.

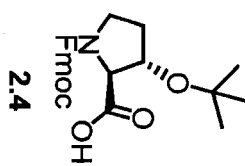
Peptide **2.1**, which was (Pro–4-Hyp–Gly)₃–3-Hyp–4-Hyp–Gly–(Pro–4-Hyp–Gly)₃, was purified by HPLC, eluting with 5–60% v/v solvent B over 35 min (where solvent A was H₂O containing TFA (0.1% v/v), and solvent B was CH₃CN containing TFA (0.1% v/v)). ESI MS: 1905 calculated, 1906 (M + H) observed. Peptide **2.2**, which was (Pro–4-Hyp–Gly)₃–Pro–3-Hyp–Gly–(Pro–4-Hyp–Gly)₃, was purified by HPLC, eluting with 5–60% solvent B over 35 min (where solvent A was H₂O containing TFA (0.1% v/v), and solvent B was CH₃CN containing TFA (0.1% v/v)). ESI MS: 1889 calculated, 1890 (M + H) observed.

Thermal Denaturation Experiments. Values for T_m for each triple helix were determined in triplicate by thermal denaturation experiments monitored by CD spectroscopy on an Aviv 202 SF instrument equipped with an automated temperature

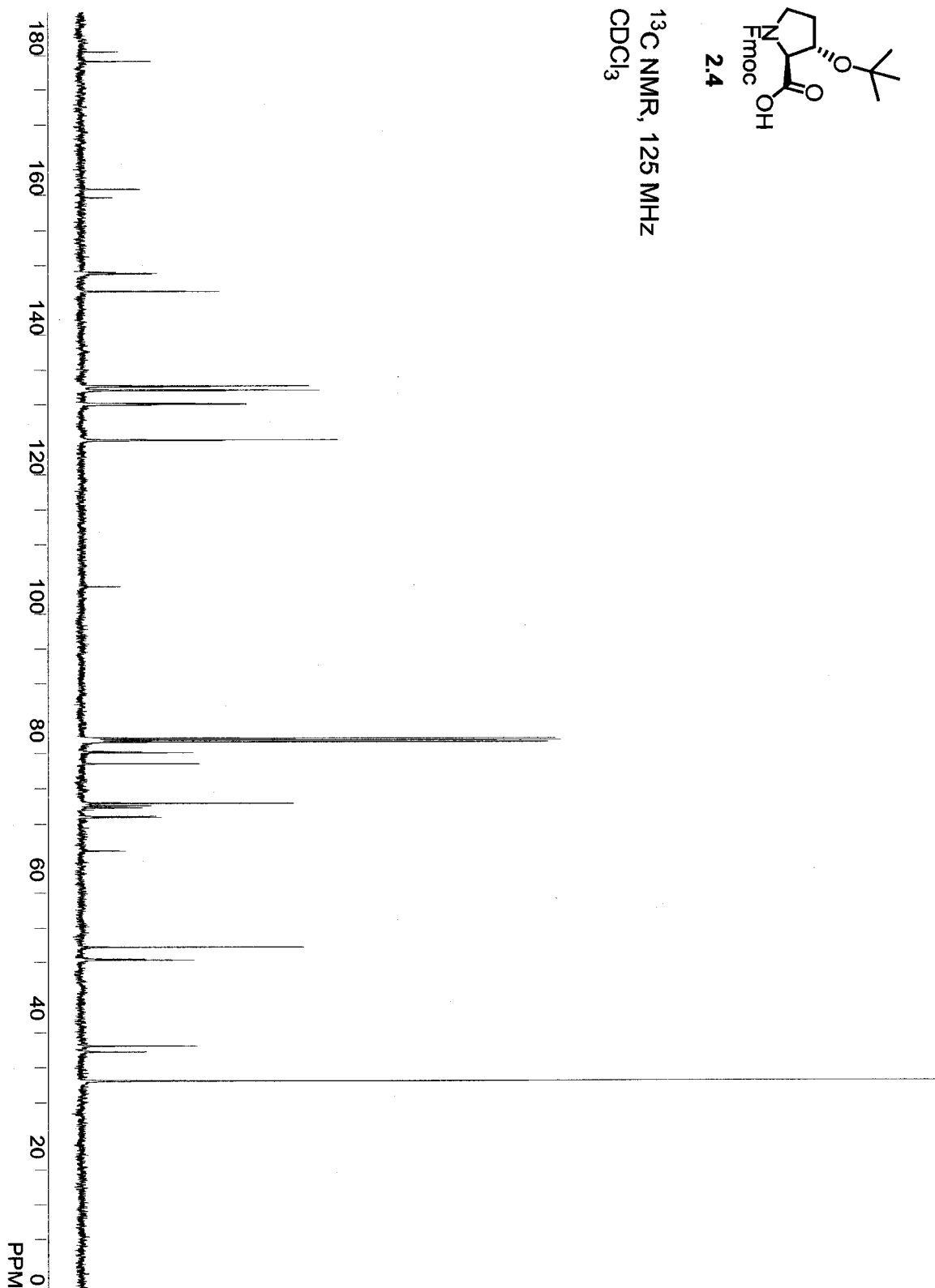
controller. A 0.2 mM solution of peptide **2.1** or **2.2** in 50 mM HOAc was incubated at 4 °C for >24 h. Aliquots of 300 μ L were placed in 0.1-cm pathlength quartz cuvettes that had been equilibrated at 5 °C. Wavelength scans were performed from 200–260 nm at 5 and 50 °C, with a slit width of 1 nm and an averaging time of 3 s. Thermal denaturation experiments were performed by raising the temperature from 5 to 50 °C in 3-°C steps, equilibrating for 5 min at each temperature, and monitoring at 225 nm with a 20-s averaging time. Values of T_m , which is the temperature at the midpoint of the thermal transition, were determined by fitting the data to a two-state model using the software package NLREG v. 4.0 (Philip Sherrod).

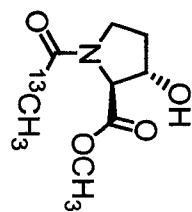
Molecular Modeling. Molecular models of triple-helical **2.1** and **2.2** were created by modifying the structure of crystalline collagen (PDB entry 1CAG(Bella et al., 1994)). The atomic coordinates were imported into the program SYBYL (Tripos, St. Louis, MO). The side-chain methyl group of the three alanine residues were replaced with a hydrogen atom, and a section of the (Pro-Hyp-Gly) triple helix was excised from the C-terminus. This section was subjected to energy minimization with the Tripos force field. A hydroxyl group was added to C^β of the proline residue in the Xaa position of the central triplet in each strand (to mimic peptide **2.1**), or moved from C^γ to C^β in the Yaa position of the central triplet in each strand (to mimic peptide **2.2**). The dihedral angles of the 3-Hyp residues were then altered to match those of the structure of crystalline amide **2.8**. The resulting structures were not subjected to further minimization.





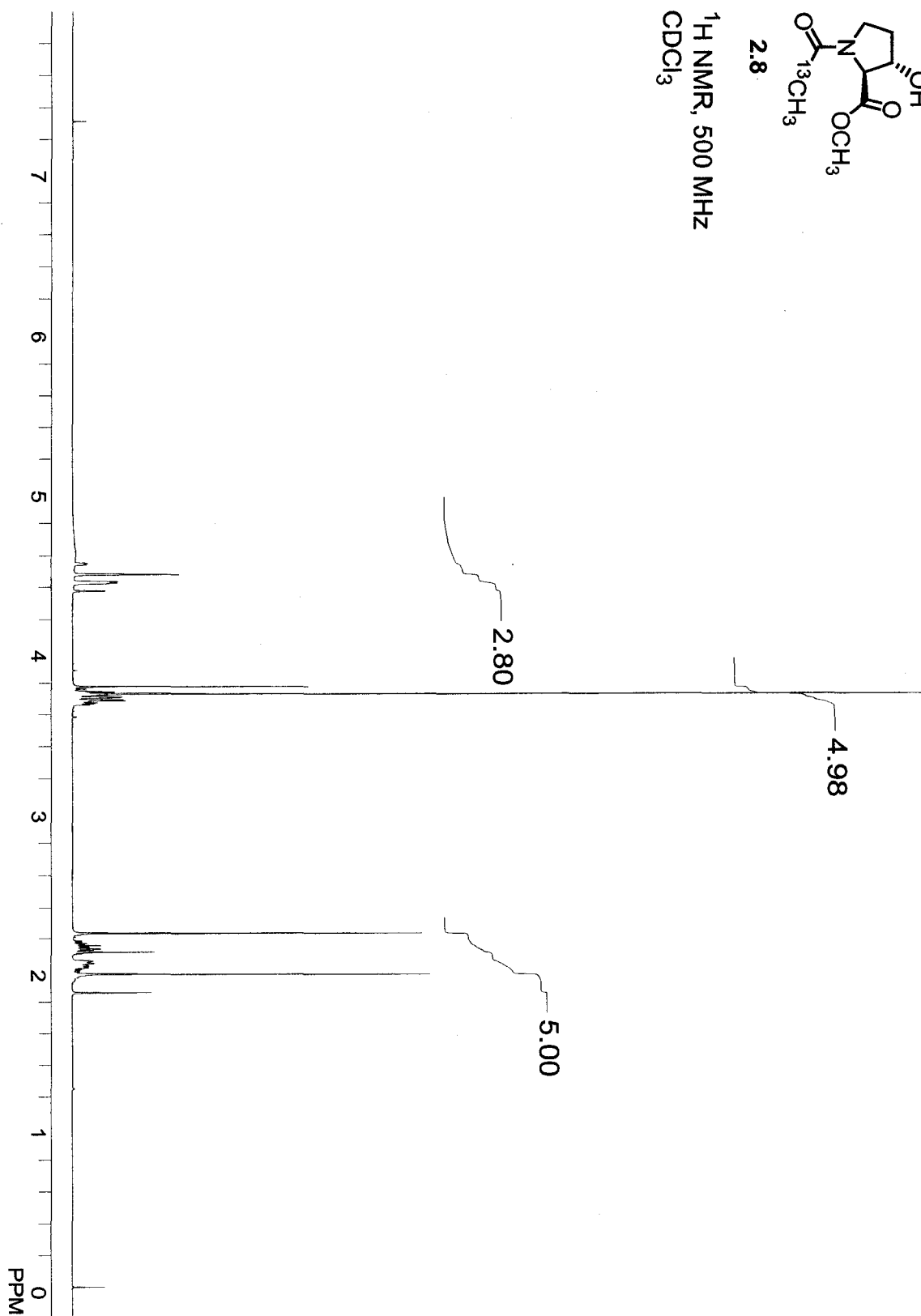
^{13}C NMR, 125 MHz
 CDCl_3

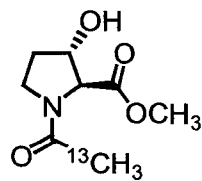




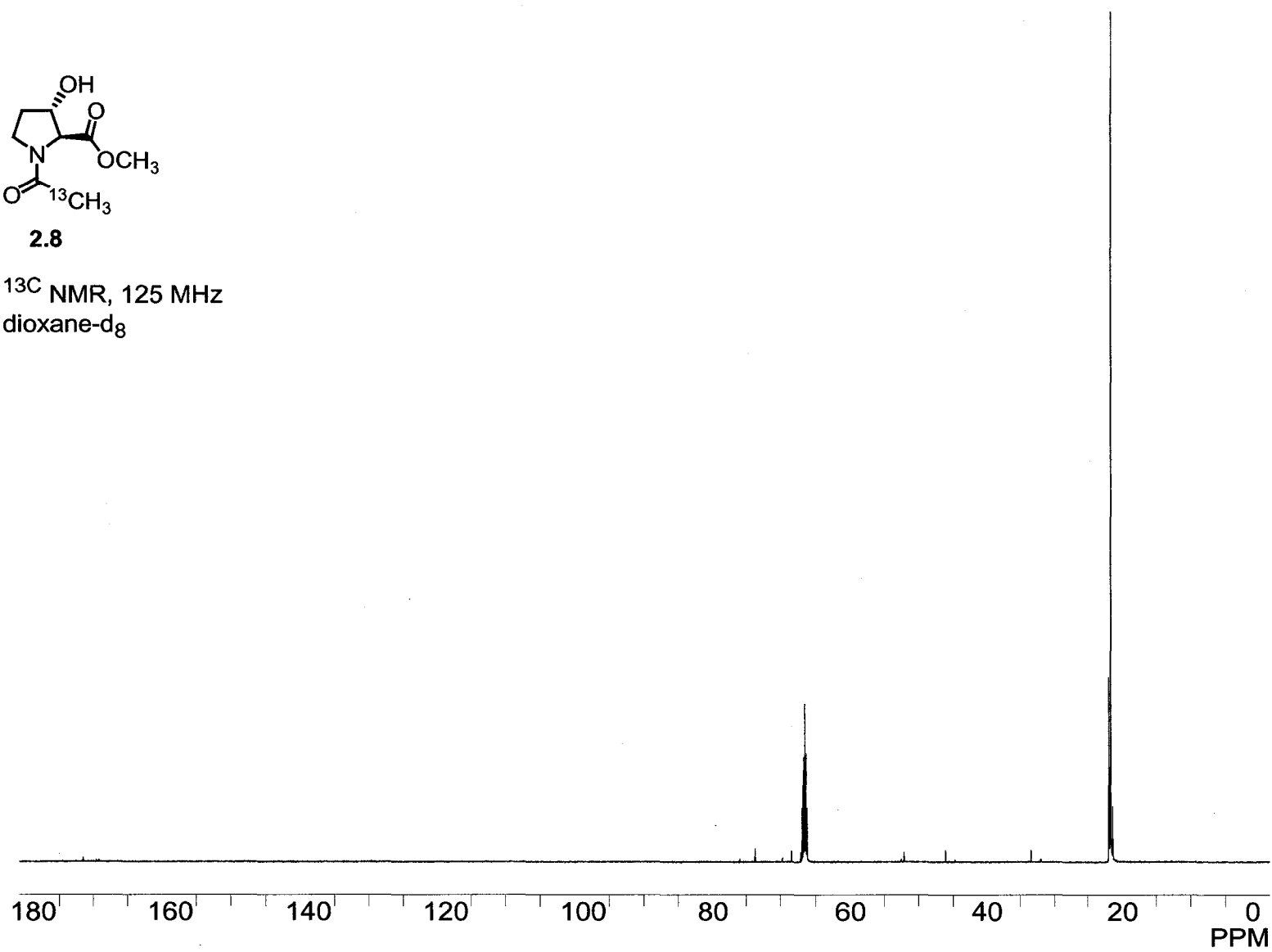
2.8

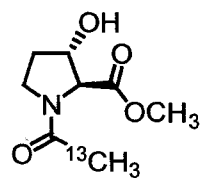
^1H NMR, 500 MHz
 CDCl_3



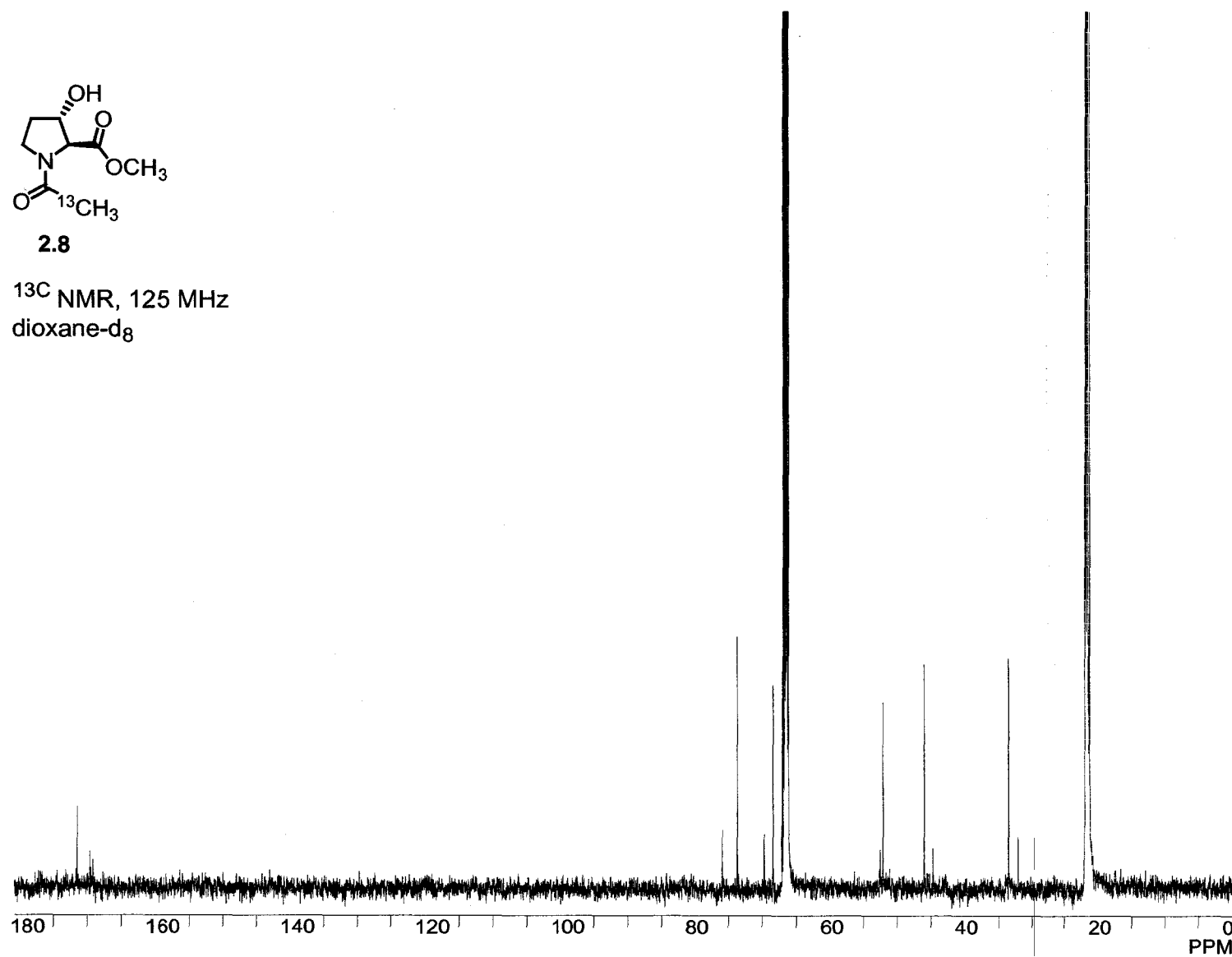
**2.8**

^{13}C NMR, 125 MHz
dioxane- d_8



**2.8**

¹³C NMR, 125 MHz
dioxane-d₈



Chapter 3[‡]

SUBSTITUTED 2-AZABICYCLO[2.1.1]HEXANES AS CONSTRAINED PROLINE ANALOGS: IMPLICATIONS FOR COLLAGEN STABILITY

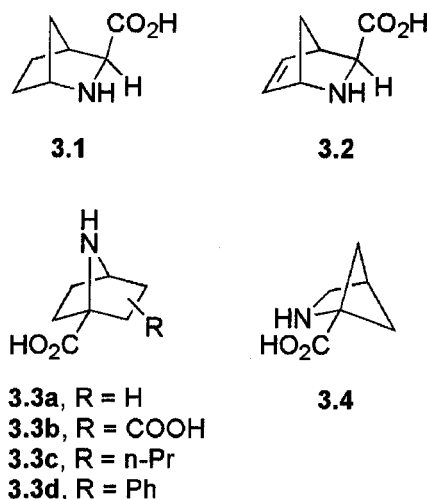
3.1 Introduction

Proline has two prominent attributes that are unique among the proteinogenic amino acids: only proline is a secondary amine, and only proline has a saturated ring (Fisher, 1906). These attributes make proline residues a key determinant of protein structure (MacArthur & Thornton, 1991; Reiersen & Rees, 2001). Accordingly, a deeper understanding of the conformational properties of proline would illuminate challenging problems in protein folding, stability, and design.

As a secondary amine, proline has a much greater propensity than other natural amino acids to form *cis* (that is, *E*) peptide bonds (Stewart et al., 1990; Weiss et al., 1998; Jabs et al., 1999). A variety of methods have been developed to control the *trans/cis* ratio, including buttressing the 2- (Delaney & Madison, 1982), 3- (Delaney & Madison, 1982; Beausoleil et al., 1998), and 5-positions (Magaard et al., 1993; An et al., 1999; Halab et al., 2000; Arnold et al., 2003) with functional groups and including the amide in a ring system that is fused to the pyrrolidine ring (Halab et al., 2000). These approaches endow torsional control of the amide bond but introduce steric bulk that could be undesirable.

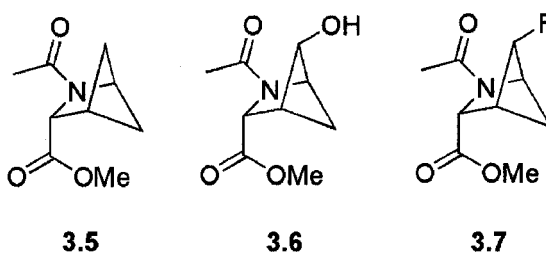
[‡] This chapter has been submitted as a portion of a publication with the same title. Reference: Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. R. *J. Org. Chem.* **2004**, *in press*.

The pyrrolidine ring of proline exists in a variety of puckers, with C^γ being its most aplanar constituent (DeTar & Luthra, 1977; Vitagliano et al., 2001b). A variety of bicyclic proline mimics have been developed to control the conformation of the pyrrolidine ring (3.1–3.4) (Bell et al., 1980; Jung et al., 1981; Gaitanopoulos & Weinstock, 1985; Montelione et al., 1986; Avenoza et al., 1995; Bunuel et al., 2001). All of these proline mimics are rigid enough to fix pyrrolidine ring pucker, but some include elements that make them suboptimal as proline mimics. Of those proline mimics, 3.3a (Han et al., 1999; Avenoza et al., 2002) and 3.4 (Juvvadi et al., 1992) have been employed in the design of peptide-based enzyme inhibitors to reinforce a bioactive peptide conformation, to varying degrees of success.



2-Azabicyclo[2.1.1]hexane (as in 3.5) is a proline analog that displays *both* predominant puckers of the pyrrolidine ring (Krow & Cannon, 2004). This end is achieved by the addition of a single carbon atom to proline, a minimal perturbation. Substitution of a hydrogen at the $C^{\gamma 1}$ or $C^{\gamma 2}$ position of the bicyclic system with a

hydroxyl or fluoro group yields mimics of 4-hydroxyproline (as in 3.6) and 4-fluoroproline (as in 3.7). (2*S*,4*R*)-4-Hydroxyproline (Hyp) residues are prevalent in collagen, which is the most abundant protein in animals. Replacing the Hyp residues with (2*S*,4*R*)-4-fluoroproline (Flp) residues endows synthetic mimics of collagen with extraordinary stability (Holmgren et al., 1998; Holmgren et al., 1999; Jenkins & Raines, 2002).



Substitutions on C-4 (that is, C^γ) of proline residues are known to have a large effect on the *trans*/*cis* ratio (Eberhardt et al., 1996; Bretscher et al., 2000; Bretscher et al., 2001; Renner et al., 2001). For example, electronegative substituents in the 4*R*-position of Pro increase the stability of the *trans* isomer, whereas electronegative substituents in the 4*S*-position decrease that stability (Table 3.1). NMR analyses indicate that Ac-(2*S*,4*R*)-4-fluoroproline-OMe (Ac-Flp-OMe) resides predominantly (86%) in the C^γ-exo pucker in solution, whereas Ac-(2*S*,4*S*)-4-fluoroproline-OMe (Ac-flp-OMe) is found almost exclusively (95%) in the C^γ-endo pucker (Fig. 3.1) (DeRider et al., 2002). This dichotomy can be attributed to the *gauche* effect, which causes the pyrrolidine ring to adopt a pucker that places the nitrogen and fluorine in a *gauche* orientation about the C^{δ2}-C^γ bond (for nomenclature, see Fig. 3.2 and (Nomenclature, 1970)) (O'Hagan et al.,

2000), and has important ramifications for the stability of collagen (Eberhardt et al., 1996; Bretscher et al., 2001; Improta et al., 2001; DeRider et al., 2002; Jenkins & Raines, 2002; Doi et al., 2003; Hodges & Raines, 2003). The relationship between the gauche effect and amide trans/cis isomerization in proline derivatives is thought to arise from both steric and electronic interactions. The stabilization of the trans conformer of Ac-Flp-OMe arises from an $n \rightarrow \pi^*$ interaction between the oxygen of the amide (the electron donor) and the carbon of the methyl ester (the electron acceptor) (Bretscher et al., 2001; DeRider et al., 2002; Hinderaker & Raines, 2003). The relative destabilization of the trans rotamer of Ac-flp-OMe arises from unfavorable steric interactions between the non-bonded electrons on the fluoro and ester groups, which disfavors the $n \rightarrow \pi^*$ interaction (DeRider et al., 2002).

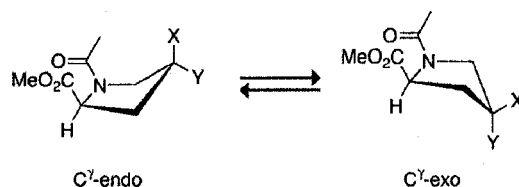
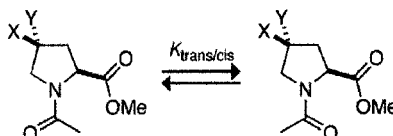


Figure 3.1 Ring puckers in 4-substituted Ac-Pro-OMe. C^γ -endo pucker is favored when $X = H, OH, \text{ or } F$, and $Y = H$. C^γ -exo pucker is favored when $X = H$ and $Y = OH \text{ or } F$.

Table 3.1 Values of $K_{trans/cis}$ for 4-substituted AcXaaOMe.^{a,b}

Xaa	X	Y	$K_{trans/cis}$
Flp	H	F	6.7
Hyp	H	OH	6.1
Pro	H	H	4.6
hyp	OH	H	2.4
flp	F	H	2.5

^a Data are from
(Bretscher et al., 2001)

^b Values were
measured in D₂O at
25 °C by integration
of ¹H NMR spectra

We reasoned that constraining the pucker of proline with a one-carbon (that is, methano) bridge would allow us to dissect the relationship among ring pucker, inductive effects, and peptide backbone conformation in proline derivatives. Accordingly, we synthesized compounds **3.5–3.7** to isolate the influence of ring pucker (which is not variable) and inductive effects (which is variable) on the cis–trans isomerization of proline residues. The results provide insight on the origin of the preference for cis or trans prolyl peptide bonds and have important implications for the conformational stability of collagen.

3.2 Results and Discussion

Compounds **3.5**–**3.7** were synthesized and provided by G. R. Krow and coworkers, Temple University, Philadelphia, PA. The synthetic route to these compounds will be reported in due course. An attractive feature of the methanoproline derivatives is that both pyrrolidine ring puckers are incorporated into a single framework, which allows dissection of the relative contributions of ring pucker and inductive effects on the conformation of substituted prolines. The hydroxyl- and fluoro-substituted methanoproline are analogs of (2*S*,4*S*)-4-hydroxyproline (hyp) and (2*S*,4*S*)-4-fluoroproline (flp), respectively, and the substituent is in an orientation analogous to the disfavored C^γ-exo pucker (O'Hagan et al., 2000; DeRider et al., 2002). In other words, the bicyclic structure fixes the hydroxyl and fluoro groups of compounds **3.6** and **3.7** to be in an antithetical conformation—anti rather than gauche to the pyrrolidine nitrogen about the C^{δ2}–C^{γ1} bond.

We measured the amide trans/cis ratios of **3.5** and **3.6** by using NMR spectroscopy. We used ¹³C NMR spectroscopy because the ¹H resonances of both methyl groups overlapped with those of other protons from the bicyclic ring system. ¹³C NMR spectra were obtained with ¹H-decoupling enabled only during the acquisition phase of the pulse sequence, allowing for no NOE buildup and thus enabling quantitative integration of the relevant peaks. The trans/cis ratios were measured in CDCl₃, dioxane-d₈, and D₂O, and are listed in Table 3.2. The trans/cis ratios of **3.7** were measured by ¹⁹F NMR spectroscopy, as the two ¹⁹F resonances were well-resolved. We found little variation among the three derivatives in a particular solvent. The trans/cis ratios in deuterated dioxane and CDCl₃ were all similar, and the ratios in D₂O were somewhat greater than

those in organic solvent (Eberhardt et al., 1996). These data demonstrate that rigidifying the pyrrolidine ring of proline derivatives by adding a methano bridge abolishes any inductive effect exerted by a fluoro- or a hydroxyl group on the trans/cis ratio of its peptide bond. Apparently, an electronegative substituent on the flexible pyrrolidine ring of proline affects the trans/cis ratio by altering the pucker of the ring.

Table 3.2 *Effect of solvent on $K_{trans/cis}$ of AcXaaOMe.^a*

Xaa	$K_{trans/cis}$		
	CDCl ₃	Dioxane-d ₈	D ₂ O
methano-Pro (3.5)	2.4	2.2	3.5
methano-hyp (3.6)	2.4	2.1	3.6
methano-flp (3.7)	2.7	2.8	3.5

^a Values of $K_{trans/cis}$ were measured in the indicated solvents at 25 °C by integration of ¹³C or ¹⁹F NMR spectra.

Next, we determined the crystalline structure of Ac-methano-hyp-OMe (3.6) by X-ray diffraction analysis. There are two crystallographically independent molecules in the unit cell, as shown in Fig. 3.2. The ring structures of the two molecules are superimposable, while the ester and amide conformations vary slightly relative to one another. The φ angles ($C_{i-1}-N_i-C_i^\alpha-C_i$) differ by 3.1(3)°, the ψ angles (here, $N_i-C_i^\alpha-C_i-O_{i+1}$) differ by 18.0(4)°, and the ω angles ($C_i^\alpha-C_i-N_{i+1}-C_{i+1}^\alpha$) differ by 2.6(2)°. A superposition of the structures of Ac-methano-hyp-OMe and Ac-Hyp-OMe (Panasik et al., 1994) is shown in Fig. 3.3 and clearly depicts the antipodal configuration of the hydroxyl groups on C $^\gamma$.

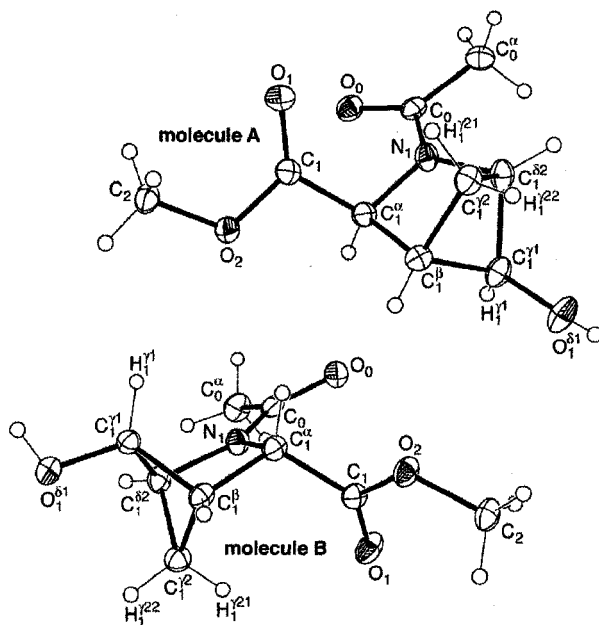


Figure 3.2 ORTEP diagrams showing the two crystallographically independent molecules in the unit cell of crystalline Ac-methano-hyp-OMe (3.6), drawn with 30% probability ellipsoids.

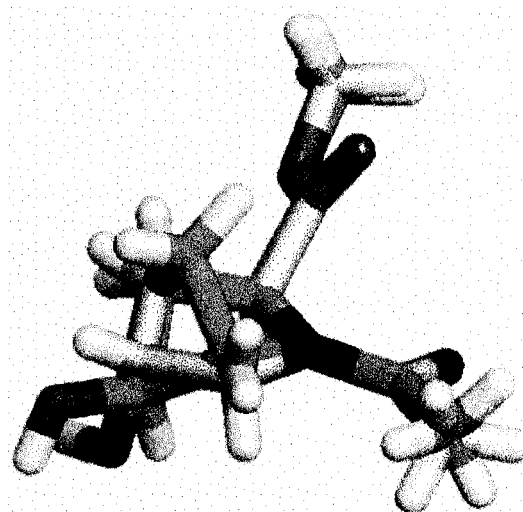


Figure 3.3 Superposition of crystalline structures of Ac-methano-hyp-OMe (3.6, cyan) and Ac-Hyp-OMe (orange) (Panasik et al., 1994).

All relevant structural parameters support the presence of a stronger $n \rightarrow \pi^*$ interaction (Bretscher et al., 2001) in molecule B than in molecule A of crystalline Ac-methano-hyp-OMe, as shown in Fig. 3.4 (DeRider et al., 2002). For example, the $O_0 \cdots C_1$ distance is 2.949(4) Å in molecule B, but 3.092(3) Å in molecule A. In addition, the $C_1=O_1$ bond length is 1.207(3) Å in molecule B, but 1.198(3) Å in molecule A, and the $^+N_1=C_0-O_0^-$ amidic resonance structure appears to be more prevalent in molecule B, which has a shorter N_1-C_0 bond and a longer C_0-O_0 bond than does molecule A. Finally, the $O_0 \cdots C_1=O_1$ angle is closer to the Bürgi-Dunitz optimum of 109° (Bürgi, H. B. et al., 1973; Bürgi, H.B. et al., 1974a; Bürgi, H.B. et al., 1974b; Bürgi, H.B. & Dunitz, 1983; Eliel & Wilen, 1994) in molecule B [$93.2(2)^\circ$] than in molecule A [$83.6(2)^\circ$]. All of these

structural parameters are consistent with greater donation of electron density from the non-bonding electrons of O_0 to the antibonding orbital of the $C_1=O_1$ bond in molecule B, as expected from a stronger $n \rightarrow \pi^*$ interaction (DeRider et al., 2002). Moreover, the congruence of these five structural parameters (three bond lengths, an atom...atom distance, and an atom...atom-atom angle) provides additional support for the existence of a meaningful $n \rightarrow \pi^*$ interaction in Ac-methano-hyp-OMe (**3.6**) as well as a benchmark for detecting $n \rightarrow \pi^*$ interactions in other derivatives of proline.

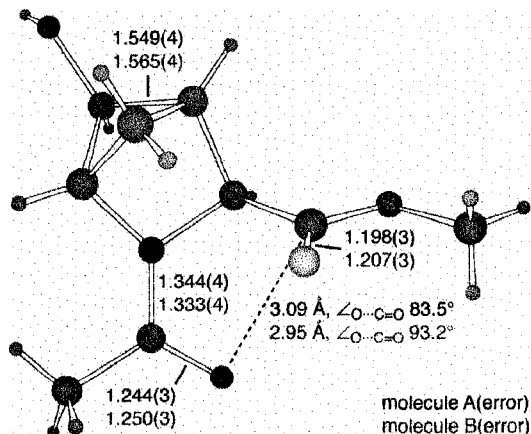


Figure 3.4 Ball-and-stick diagram showing bond lengths that differ statistically in the two molecules of the X-ray structure of crystalline Ac-methano-hyp-OMe (**3.6**). The $O_0 \cdots C_1=O_1$ bond lengths and angles are also shown.

The trans/cis ratios for compounds **3.5**–**3.7** in D_2O are intermediate between those of Ac-Pro-OMe and Ac-hyp-OMe or Ac-flp-OMe (Table 3.1), indicating that the $n \rightarrow \pi^*$ interactions in methanoprolin derivatives are probably weaker than those found in Pro and 4*R*-substituted prolines, but stronger than those in the 4*S*-substituted prolines. Likewise, the ϕ and ψ angles of Ac-methano-hyp-OMe are intermediate between those

of the endo- and exo-puckers of proline derivatives (Table 3.3), which is consistent with its intermediate trans/cis ratio.

Table 3.3 Values of φ and ψ dihedral angles for the trans amide isomers of Ac-Xaa-OMe.^a

Xaa	ring pucker	φ (°)	ψ (°)	reference
Pro	C ^γ -exo	-58.6	143.0	(DeRider et al., 2002)
Pro	C ^γ -endo	-70.0	152.1	(DeRider et al., 2002)
Hyp (1)	C ^γ -exo	-62.0	156.4	(Panasik et al., 1994)
Hyp (2)	C ^γ -exo	-50.9	145.2	(Panasik et al., 1994)
Flp	C ^γ -exo	-59.2	140.8	(DeRider et al., 2002)
flp	C ^γ -endo	-76.4	169.0	(DeRider et al., 2002)
methano-hyp (1)	—	-65.0	169.0	this work
methano-hyp (2)	—	-61.9	153.5	this work

^a Dihedral angles of Xaa = Pro, Flp, and flp are from density functional theory calculations; dihedral angles of Xaa = Hyp and methano-hyp are from X-ray diffraction analysis of crystalline molecules.

3.3 Conclusions

Electronegative substituents in the 4-position of proline residues had been shown to have a substantial effect on the trans/cis ratio of their peptide bonds (Table 3.1) (Eberhardt et al., 1996; Bretscher et al., 2000; Bretscher et al., 2001; Renner et al., 2001). Here, constraining the pucker of the pyrrolidine ring of 4-substituted proline residues with a one-carbon bridge, as in compounds **3.5–3.7**, was shown to abolish the effect of the electronegative substituents on the trans/cis ratio (Table 3.2). Thus, changes in trans/cis ratio arise from changes in ring pucker. This finding suggests that pyrrolidine ring pucker is a key determinant of the stability (or instability) endowed by 4-substituted proline residues on collagen.

3.4 Experimental Section

Measurement of $K_{\text{trans/cis}}$ values of 3.5–3.7. Each compound (10–20 mg) was dissolved in CDCl_3 (approximately 1 mL), and the ^{13}C NMR (3.5 and 3.6) or ^{19}F NMR (3.7) spectrum recorded. The relaxation delay for the measurement of the spectra of 3.5 and 3.6 was 10–18 s to allow for full relaxation of the ^{13}C nuclei. The spectral baselines were corrected and peaks corresponding to the labeled carbon or the fluorine were integrated. The samples were then concentrated under reduced pressure and placed under high vacuum overnight to ensure removal of all residual CDCl_3 . The resulting samples were dissolved in dioxane- d_8 (800 μL), and the spectra were recorded again. The samples were concentrated under reduced pressure and placed under high vacuum overnight. D_2O (800 μL) was then added to each sample followed by enough CD_3OD to effect full dissolution of the sample. The amount of added CD_3OD was less than 20% of the total volume in each case. The samples were filtered, their spectra were recorded, and the trans/cis ratios were determined by integration of the respective resonances.

Crystallization of Ac–methano-hyp–OMe (3.6). Racemic Ac–methano-hyp–OMe (3.6, 20–30 mg) was dissolved in dichloromethane, and the resulting solution was aliquotted into 5 vials. A cosolvent (10–20 drops) was added to each vial with a Pasteur pipette: vial 1—hexanes; vial 2—diethyl ether; vial 3—dioxane; vial 4—no cosolvent; vial 5—ethyl acetate. The vials were capped loosely and allowed to sit at room temperature for approximately 2 days. Vial 5 contained the crystals most suitable for X-ray crystallography, and these crystals were used for X-ray diffraction analysis.

X-Ray Diffraction Data Collection. An air-stable crystal of Ac-methano-hyp-OMe (**3.6**) with approximate dimensions $0.50 \times 0.40 \times 0.40 \text{ mm}^3$ was selected under oil at ambient conditions and attached to the tip of a glass capillary. The crystal was mounted in a stream of cold nitrogen at 173(2) K and centered in the X-ray beam by using a microscope.

Crystal evaluation and data collection were performed on a Bruker P4/CCD-1000 diffractometer with Mo K α ($\lambda = 0.71073 \text{ \AA}$) radiation with a diffractometer-to-crystal distance of 4.999 cm.

Initial cell constants were obtained from three series of ω scans at different starting angles. Each series consisted of 20 frames collected at intervals of 0.3° in a 6° range about ω with an exposure time of 10 s per frame. A total of 69 reflections were obtained. The reflections were indexed successfully by an automated indexing routine built in the SMART program. The final cell constants were calculated from a set of 4952 strong reflections from the actual data collection.

Data were collected by using the multi-run data collection routine. The reciprocal space was surveyed to the extent of a full sphere to a resolution of 0.80 \AA . A total of 12437 data were harvested by collecting one set of 1250 frames with 0.3° scans in ϕ and four sets of 100 frames with 0.3° scans in ω with an exposure time 30 s per frame. This highly redundant dataset was corrected for Lorentz and polarization effects. The absorption correction was based on fitting a function to the empirical transmission surface as sampled by multiple equivalent measurements (Blessing, 1995).

Structure Solution and Refinement. The systematic absences in the diffraction data were consistent for the space groups $P1$ and $P\bar{1}$ (Stowell et al., 1995). The E -statistics

strongly suggested the centrosymmetric space group $P\bar{1}$ that yielded chemically reasonable and computationally stable results of refinement.

A successful solution by the direct methods provided most non-hydrogen atoms from the E -map. The remaining non-hydrogen atoms were located in an alternating series of least-squares cycles and difference Fourier maps. All non-hydrogen atoms were refined with anisotropic displacement coefficients. All hydrogen atoms were included in the structure factor calculation at idealized positions and were allowed to ride on the neighboring atoms with relative isotropic displacement coefficients. There were two chemically equivalent but crystallographically independent molecules of compound **3.6** in the asymmetric unit. Because compound **3.6** crystallized in a centrosymmetric space group, the crystal structure was a racemic mixture of stereoisomers. Several likely intermolecular hydrogen-bonding interactions were observed in the lattice, and formed a series of one-dimensional chains in the ab plane.

The final least-squares refinement of 259 parameters against 3638 data resulted in residuals R (based on F^2 for $I \geq 2\sigma$) and wR (based on F^2 for all data) of 0.0735 and 0.2199, respectively. The final difference Fourier map was featureless.

Chapter 4[§]

O-ACYLATION OF HYDROXYPROLINE RESIDUES: EFFECT ON PEPTIDE BOND ISOMERIZATION AND COLLAGEN STABILITY

4.1 Introduction

Collagen is the most prevalent component of the extracellular matrix in animals (Ramachandran & Reddi, 1976; Nimni, 1988). Collagen has a characteristic tertiary structure consisting of three left-handed polyproline II-type helices wound around a common axis to form a triple helix with a shallow right-handed superhelical pitch. The close packing of the polypeptide chains in the triple helix requires that every third residue be glycine. Each polypeptide chain in fibrillar collagen contains about 300 repeats of the sequence XaaYaaGly, in which Xaa is often (2*S*)-proline (Pro) and Yaa is often (2*S*,4*R*)-4-hydroxyproline (Hyp). The Hyp residues arise from a post-translational modification of Pro residues by the enzyme prolyl 4-hydroxylase. The thermal stability of collagen has a positive correlation with its Hyp content (Burjanadze, 1979, 2000).

The effects of Hyp stereochemistry and sequence on collagen stability have been explored with peptide mimics (Fields, G.B. & Prockop, 1996; Jenkins & Raines, 2002). In seminal work, Prockop and coworkers showed that [(ProHypGly)₁₀]₃ is more stable than [(ProProGly)₁₀]₃, whereas (HypProGly)₁₀ does not exhibit triple helix formation (Berg & Prockop, 1973; Inouye et al., 1982). The diastereomer (2*S*,4*S*)-4-hydroxyproline

[§] The data in this chapter were collected in collaboration with Eric S. Eberhardt and Alexander I. McCloskey, Vassar College, Poughkeepsie, N.Y.

(hyp) in the Xaa or Yaa position precludes triple helix formation (Inouye et al., 1976).

More recent studies have shown that Hyp can enhance triple helical stability in the Xaa position, but not if the Yaa residue is Pro (Bann & Bächinger, 2000; Mizuno et al., 2003).

Previous reports from our laboratory have shown that replacing Hyp in (ProHypGly)_n ($n = 7$ or 10) with (2*S*,4*R*)-4-fluoroproline (Flp) (but not its diastereomer (2*S*,4*S*)-4-fluoroproline (flp)) greatly enhances the stability of its triple helices (Table 4.1) (Holmgren et al., 1998; Holmgren et al., 1999; Bretscher et al., 2001). The increase in stability derives from the enhanced electron-withdrawing ability of the fluoro group. That inductive effect fixes the pucker of the pyrrolidine ring, and thereby preorganizes the peptide backbone dihedral angles of the individual strands into conformations that are favorable for triple helix formation (Improta et al., 2001; DeRider et al., 2002).

Table 4.1 *Effect of Hyp and Flp on the conformational stability of collagen triple helices.*

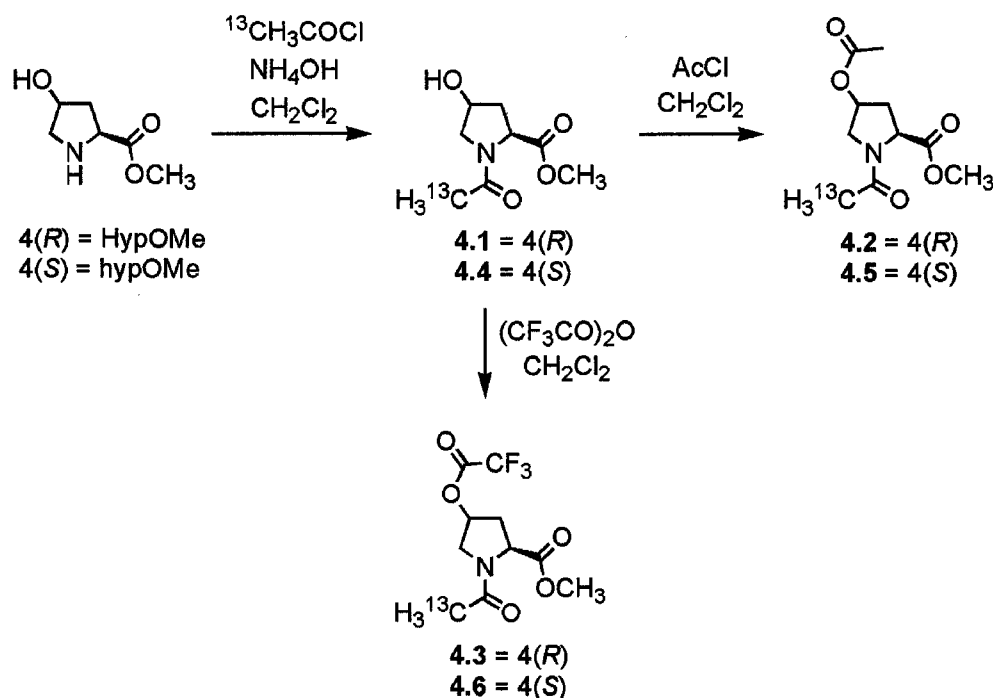
Strand	T_m (°C)	Ref
(ProProGly) ₇	6–7	(Shaw & Schurr, 1975)
(ProHypGly) ₇	36	(Bretscher et al., 2001)
(ProFlpGly) ₇	45	(Bretscher et al., 2001)
(ProProGly) ₁₀	41	(Holmgren et al., 1998; Holmgren et al., 1999)
(ProHypGly) ₁₀	69	(Holmgren et al., 1998; Holmgren et al., 1999)
(ProFlpGly) ₁₀	91	(Holmgren et al., 1998; Holmgren et al., 1999)

What is the best means to put into practice our knowledge about the basis for collagen stability? The chemical synthesis of biopolymers on a large scale is tedious and expensive. In contrast, the chemical modification of natural biopolymers is facile and inexpensive. Can increased inductive effects be used to enhance the conformational stability of natural (as opposed to synthetic) collagen? No known reagent can be used to

replace the hydroxyl group of Hyp residues with a fluoro group in aqueous solution (Ibrahim & Togni, 2004). The electron-withdrawing ability of a hydroxyl group can, however, be increased by its chemical modification. Here, we synthesize and characterize *O*-acyl derivatives of AcHypOMe and AchypOMe. We then determine the effects of *O*-acetylation on the conformational stability of a [(ProHypGly)₁₀]₃ triple helix. The results illuminate intrinsic electronic and steric effects on the stability of the collagen triple helix.

4.2 Results and Discussion

Hydroxyl, acetoxyl, and trifluoroacetoxyl groups confer a wide range of electron-withdrawing ability. These functional groups have conjugate acid pK_a values of 15.6, 4.76, and 0.23, respectively (Dippy et al., 1959), and F values (which report on inductive effects) of 0.33, 0.42, and 0.58, respectively (Hansch et al., 1991). Moreover, acetoxyl and trifluoroacetoxyl groups could be installed into natural collagen by the *O*-acetylation or *O*-trifluoroacetylation of Hyp residues. To reveal the consequences of *O*-acylation of Hyp residues, compounds **4.1**–**4.6** were synthesized by the routes in Scheme 4.1. In **4.1**–**4.6**, hydroxyl, acetoxyl, and trifluoroacetoxyl groups are installed in each configuration at C γ of AcProOMe.



Scheme 4.1 *Synthesis of compounds 4.1–4.6.*

The structures of crystalline **4.2** and **4.5** were determined by X-ray diffraction analysis. These structures are shown in Figure 4.1, along with those of crystalline **4.1** and AcProOMe (Panasik et al., 1994). Compounds **4.1** and **4.2** crystallized with their pyrrolidine rings in the C^γ -exo conformation and a trans amide bond, whereas compound **4.5** and AcProOMe crystallized in the C^γ -endo conformation and a cis amide bond. The φ and ψ angles of **4.2** and **4.5** are similar to those for proline residues with C^γ -exo and C^γ -endo puckers, respectively (Table 4.2) (Vitagliano et al., 2001a). The backbone dihedral angles for **4.2** are intermediate between those of **4.1** and AcFlpOMe (**4.7**), and are favorable for triple helix formation (Berisio et al., 2001, 2002).

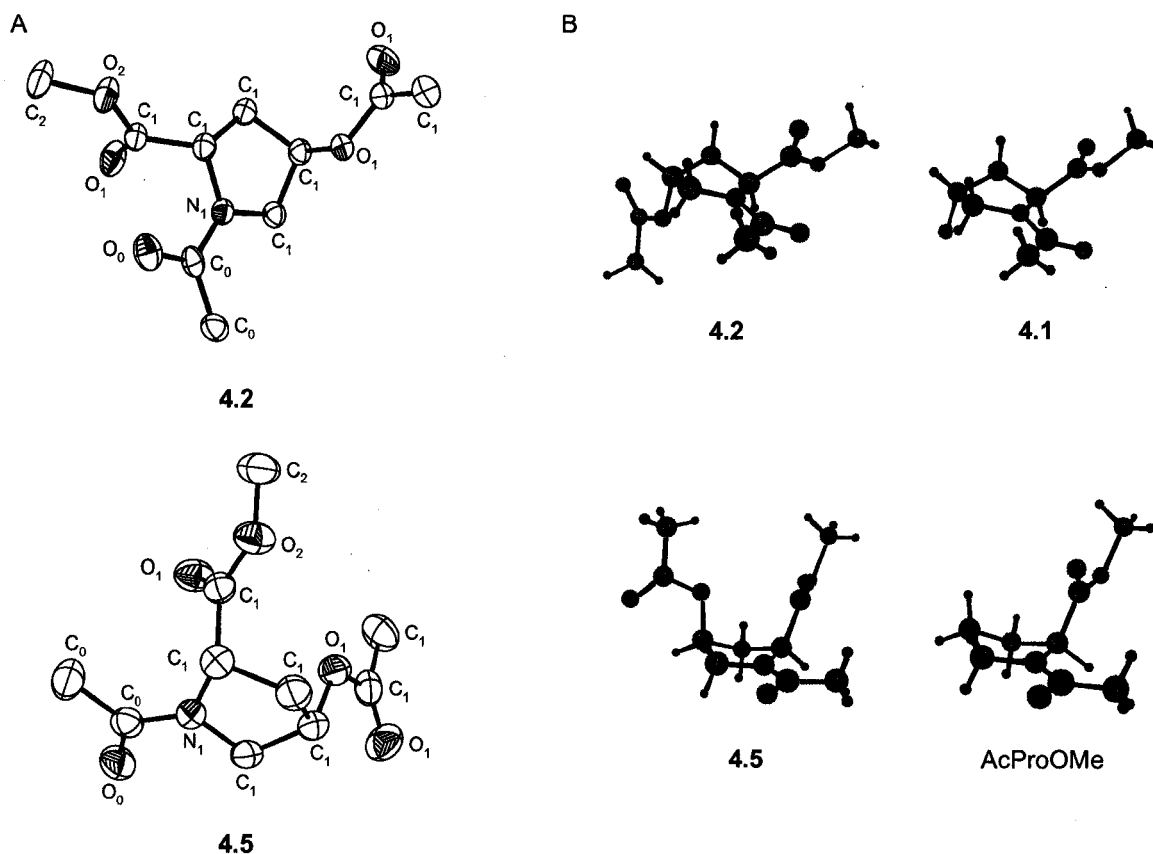


Figure 4.1 (A) ORTEP diagrams of crystalline **4.2** (50% probability ellipsoids) and **4.5** (40% probability ellipsoids). (B) Comparison of pyrrolidine ring conformations of crystalline **4.2** and **4.1** (top), (Panasik et al., 1994) and **4.5** and AcProOMe (bottom) (Panasik et al., 1994).

Table 4.2 Backbone dihedral angles of crystalline AcProOMe, **4.1**, **4.2**, **4.5**, and **4.7** derived from X-ray diffraction analysis^a

Compound	Ring pucker	φ	ψ	ω
AcProOMe ^b	C ^{γ} -endo	-78.9	176.7	-3.1
4.5	C ^{γ} -endo	-73.9	-170.9	-3.4
4.1 ^b	C ^{γ} -exo	-62.0	156.4	-180.0
4.2	C ^{γ} -exo	-58.2	142.0	-179.4
4.7 ^b	C ^{γ} -exo	-55.4	140.6	-177.0

^a φ , C_{i-1}-N_i-C_i ^{α} -C_i; ψ , N_i-C_i ^{α} -C_i-O_{i+1}; ω , O_{i-1}-C_{i-1}-N_i-C_i ^{α}

^b Data are from (Panasik et al., 1994).

^{13}C NMR chemical shifts can report on electron withdrawal by pendant function groups (Friebolin, 1998). The relative electron-withdrawing ability of the hydroxyl, acetoxyl, and trifluoroacetoxyl groups in **4.1–4.6** were assessed by comparing the ^{13}C NMR chemical shifts of their C^γ atoms with those of AcProOMe, AcFlpOMe (**4.7**), and AcflpOMe (**4.8**). These ^{13}C NMR chemical shifts (Table 4.3) indicate that electron withdrawal increases in the order: hydroxyl < acetoxyl < trifluoroacetoxyl, in accord with the conjugate acid $\text{p}K_\text{a}$ and F values of these groups.

Table 4.3 $^{13}\text{C}^\gamma$ chemical shift (δ) of AcProOMe and compounds **4.1–4.8**^a

Compound	δ	Compound	δ
AcProOMe	21.9		
4.1	70.4	4.4	70.7
4.2	73.7	4.5	73.7
4.3	78.5	4.6	78.6
4.7	93.3	4.8	93.5

^a Values were obtained at 25°C in 1,4-dioxane- d_8 .

The cis-to-trans equilibrium constant for prolyl peptide bonds ($K_{\text{Z/E}}$) is close to unity. Yet, all of the peptide bonds in a collagen triple helix are in the trans conformation. Accordingly, substitutions that favor the trans isomer can enhance collagen stability (Eberhardt et al., 1996; Holmgren et al., 1998; Holmgren et al., 1999).

The value of $K_{\text{Z/E}}$ for **4.1–4.6** was assessed by NMR spectroscopy (Forsén & Hoffman, 1963; Grathwohl & Wüthrich, 1981; Led & Gesmar, 1982; Eberhardt et al., 1993; Eberhardt et al., 1996). At 37 °C, the $K_{\text{Z/E}}$ values within a series (4(*R*) versus 4(*S*)) are within error of each other (Table 4.4). The similarity of these values within each

series suggests that any enhancement of preorganization due to increased electron withdrawal at the 4-position is counteracted by other factors. The value of $K_{Z/E}$ is greater for the 4*R* series than the 4*S* series, which is consistent with previous work (Bretscher et al., 2000; Bretscher et al., 2001; Renner et al., 2001; DeRider et al., 2002).

Table 4.4 *Thermodynamic parameters for amide bond isomerization in compounds 4. 1–4.6*^a

Compound	ΔH° (kcal/mol)	ΔS° [cal/(mol·deg)]	$K_{Z/E}$ (37°C)
4.1	-1.65 ± 0.12	-2.59 ± 0.38	4.0 ± 0.4
4.2	-2.24 ± 0.02	-4.68 ± 0.07	3.3 ± 0.3
4.3	-1.75 ± 0.04	-3.15 ± 0.13	3.5 ± 0.4
4.4	-3.82 ± 0.08	-11.19 ± 0.25	1.8 ± 0.2
4.5	-2.78 ± 0.12	-8.01 ± 0.39	1.6 ± 0.4
4.6	-4.08 ± 0.18	-12.18 ± 0.56	1.6 ± 0.6

^a Values (\pm S.E.) are derived from the data in Figure 2.

The effect of temperature on the values of $K_{Z/E}$ for **4.1–4.6** was also assessed by NMR spectroscopy. The resulting van't Hoff plots are shown in Figure 4.2. Values for ΔH° and ΔS° (\pm S.E.) were calculated from linear least-squares fits of these data to eq 4.1:

$$\ln K_{Z/E} = (-\Delta H^\circ / R)(1/T) + \Delta S^\circ / R \quad (4.1)$$

These values are listed in Table 4.4. This analysis assumes that the enthalpic and entropic differences between the cis and trans isomers are independent of temperature, that is, that $\Delta C_p^\circ = 0$ for the isomerization reaction. The linearity of the data in Figure 4.2 indicate that this assumption is likely to be valid.

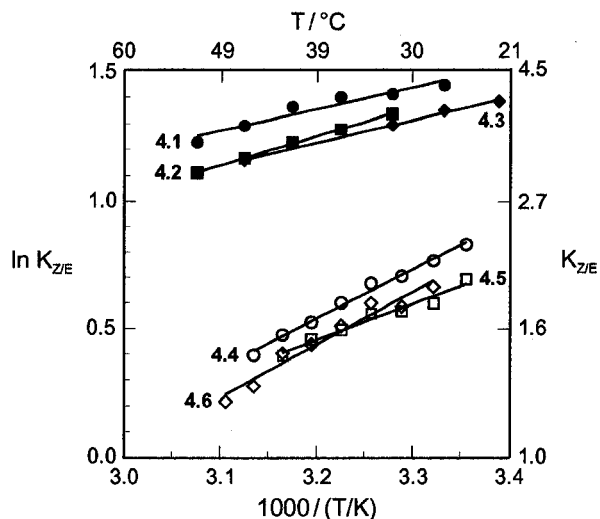


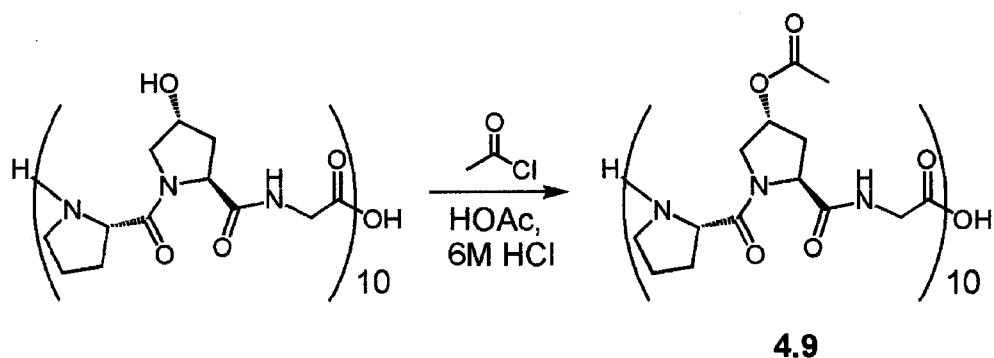
Figure 4.2 *van't Hoff plot for the cis-to-trans amide bond isomerization of compounds 4.1–4.6 in 1,4-dioxane- d_8 .*

In all conditions studied herein, $K_{Z/E} > 1$: the trans isomer of **4.1–4.6** is always more stable than the cis isomer (Figure 4.1, Table 4.4). The values of both ΔH° and ΔS° for compounds **4.4–4.6** are more negative than those for **4.1–4.3**, indicating that isomerization is more enthalpically favorable but entropically unfavorable in the 4*S* series than the 4*R* series (Table 4.4).

Electron withdrawal at C^y in **4.1–4.6** should increase the bond order in the amide C–O bond and decrease the bond order in the amide C–N bond. Indeed, this effect of electron withdrawal was manifested in a faster rate of amide bond isomerization in AcFlpOMe than in AcProOMe (Eberhardt et al., 1996). An analogous difference was not apparent, however, in the rate of cis–trans isomerization of **4.1–4.3** and **4.4–4.6** (data not shown). Apparently, the differential electron withdrawal by hydroxyl, acetoxyl, and trifluoroacetoxyl groups is too small to yield a measurable difference in the rate of cis–

trans isomerization of 4.1–4.3 and 4.4–4.6.

The physicochemical properties of 4.2 suggest that *O*-acetylation of (ProHypGly)₁₀ would not preclude its forming a triple helix and could enhance the stability of that helix. This hypothesis was tested by reacting the hydroxyl groups of (ProHypGly)₁₀ with acetyl chloride in 6 M HCl, as shown in Scheme 4.2. The reaction was done under acidic conditions to preclude acetylation at the *N*-terminus of the peptide (Wilchek & Patchornik, 1964). Mass spectrometry indicated that the acetylation proceeded to completion and without the production of byproducts, producing peptide 4.9 (Figure 4.3). Attempts to acylate (ProHypGly)₁₀ with trifluoroacetyl groups were unsuccessful because the harsher conditions required for selective *O*-trifluoroacetylation led to decomposition of the peptide. Moreover, trifluoroacetoxyl groups are much more labile than acetoxyl groups in water, which would make obtaining *T_m* data of a homogeneous *O*-trifluoroacetylated peptide problematic.



Scheme 4.2 Acetylation of (Pro-Hyp-Gly)₁₀ to produce peptide 4.9.

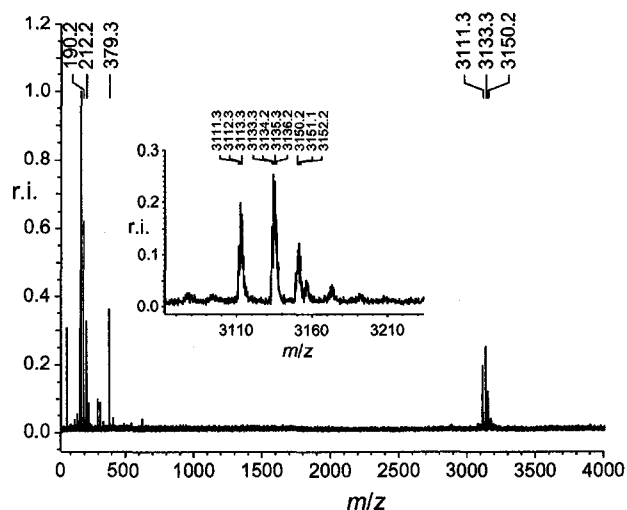


Figure 4.3 MALDI-TOF mass spectrum of peptide 4.9.

After incubation of the peptide 4.9 in 50 mM acetic acid at 4°C for 24 h, CD spectra at 25°C indicated that triple helix formation had occurred (Figure 4.4A). In this solution, the value of T_m of the triple helix (which is the temperature at the midpoint of the thermal transition from native to denatured states) was 57.5°C (Figure 4.4B), compared to 69°C for [(ProHypGly)₁₀]₃ (Holmgren et al., 1998; Holmgren et al., 1999). What is the origin of this decrease in conformational stability? Examination of the structure of the collagen triple helix reveals that addition of an acetyl group to the pseudo-axial hydroxyl group of Hyp could cause significant steric clashes between that acetyl group and a proline residue in a neighboring strand (Figure 4.5).

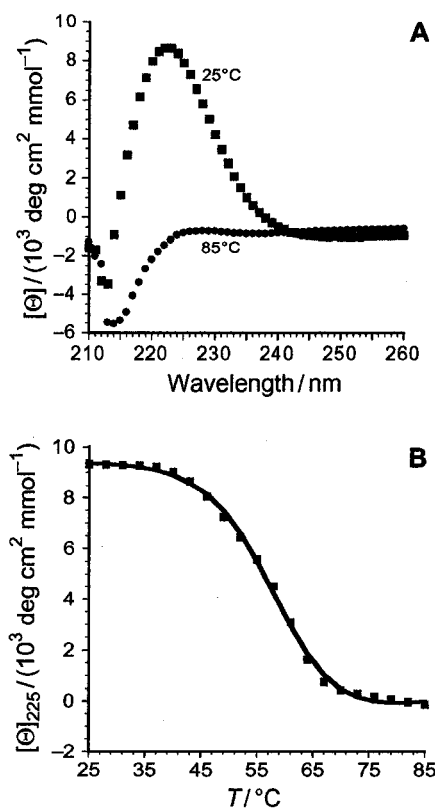


Figure 4.4 (A) Circular dichroism spectra of peptide 9 at 25°C (squares) and 85°C (circles). (B) Thermal denaturation curve for triple helical peptide 9 in 50 mM acetic acid. The squares represent the data points, and the line represents the curve-fit of the data to give $T_m = 57.5^\circ\text{C}$.

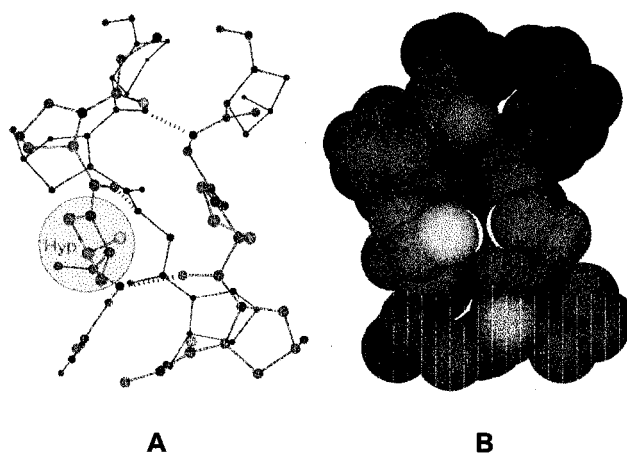


Figure 4.5 Segment of a $(\text{ProHypGly})_n$ triple helix. (A) Ball-and-stick model with a Hyp residue highlighted. (B) Space-filling model with an interstrand Hyp(Pro) interaction

highlighted. The N-terminus of the triple helical segment is at the top of the figure. Carbon is gray, nitrogen is blue, and oxygen is red. Atomic coordinates are from PDB entry 1CAG (Bella et al., 1994).

In an earlier report, *O*-acetylation was found to decrease greatly the stability of a (ProHypGly)₁₀ triple helix ($\Delta T_m = -33^\circ\text{C}$) (Weber & Nitschmann, 1978). In this report, the integrity of the peptide after acetylation and the extent of its acetylation were not apparent. Regardless, the authors used their data to argue that the large decrease in stability was due to the lesser ability of acetoxyl groups (compared to hydroxyl groups) to form water-mediated hydrogen bonds between strands. Instead, we find a small decrease in the stability of [(ProHyp(C(O)CH₃)Gly)₁₀]₃ compared to [(ProHypGly)₁₀]₃ ($\Delta T_m = -11^\circ\text{C}$), which we believe arises from the steric demands of its 30 acetyl groups (Figure 4.5) as well as the slight decrease in the value of K_{ZE} upon *O*-acetylation (Table III). Despite the loss of its hydroxyl groups, (ProHyp(C(O)CH₃)Gly)₁₀ forms a more stable triple helix than does (ProProGly)₁₀ ($\Delta T_m = 17^\circ\text{C}$). Other evidence also argues against the importance of bridging water molecules in mediating collagen stability (Holmgren et al., 1998; Jenkins & Raines, 2002).

4.3 Conclusions

Electron withdrawal by the hydroxyl group is known to have dramatic effects on the physicochemical properties of Hyp and hyp residues, and thereby on the conformational stability of triple helical collagen. We synthesized derivatives of AcHypOMe and AchypOMe containing *O*-acetoxyl and *O*-trifluoroacetoxyl groups, which are more electron-withdrawing than hydroxyl groups. These changes had little impact on the three-dimensional structure of these residues or the thermodynamics of their cis–trans amide

bond isomerization. Chemical modification of the collagen mimic (ProHypGly)₁₀ by *O*-acetylation leads to a modest decrease in the stability of its triple helix. Apparently, *O*-acylation of Hyp residues (in contrast to replacing Hyp with Flp residues) creates a steric conflict in the triple helix. We conclude that *O*-acylation is not a viable means to increase the stability of natural collagen.

4.4 Experimental Section

General. Reagents were obtained from Aldrich Chemical (Milwaukee, WI) or Fisher Scientific (Hanover Park, IL) and used without further purification. Amino acids and their derivatives were obtained from Fisher Scientific, Bachem Bioscience (King of Prussia, PA), or Novabiochem (San Diego, CA). Dichloromethane was drawn from a Baker Cycletainer. Thin-layer chromatography was performed by using aluminum-backed plates coated with silica gel containing F₂₅₄ phosphor and visualized by UV illumination or staining with I₂, *p*-anisaldehyde stain, or phosphomolybdic acid stain. NMR spectra were obtained with Bruker AC-300 and Varian UNITY-500 spectrometers.

***N*-(2-¹³C-Acetyl)-(2*S*,4*R*)-4-hydroxyproline methyl ester (4.1).** (2*S*,4*R*)-4-Hydroxyproline methyl ester hydrochloride (1.821 g, 10.0 mmol) was suspended in CH₂Cl₂ (30 mL), and this suspension was cooled to 0°C. 2-¹³C-Acetyl chloride was added in one portion. A mixture of concentrated NH₄OH (1 mL) and H₂O (10 mL) was then added dropwise. The reaction mixture was stirred at 0°C for 3 h. The layers were separated and then concentrated under reduced pressure. The organic layer contained the product, which was purified by flash chromatography, eluting with MeOH (5–10% v/v) in CH₂Cl₂ to yield **4.1** as 273 mg (14.5%) of a colorless oil. ¹H NMR (300 MHz, CDCl₃,

two rotamers): δ 4.59–4.50 (m, 2H), 3.78, 3.73 (two s, 3H), 3.77–3.71 (m, 1H), 3.55–3.50 (m, 1H), 2.50–2.25 (m, 1H), 2.08, 1.96 (two d, $J = 129.0$, 129.0 Hz, 3H), 2.21–2.00 (m, 1H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 172.8, 170.1 (d, $J = 51.7$ Hz), 69.7, 68.1, 58.6, 57.4, 55.8, 54.3, 52.6, 39.5, 37.7, 22.1 (enriched), 21.5 (enriched). ESI-MS: m/z 211.0849 ($[\text{M}+\text{Na}]^+$); 211.0854 ($[\text{M}+\text{Na}]^+$, calcd).

***N*-(2- ^{13}C -Acetyl)-(2S,4R)-4-acetoxypyrrolidine methyl ester (4.2).** Compound **4.1** (102 mg, 0.54 mmol) was dissolved in dry CH_2Cl_2 (1 mL). Acetyl chloride (0.1 mL, 1.41 mmol) was added to the resulting solution in two portions. The reaction mixture was stirred for approximately 4 h, and then concentrated under reduced pressure to give **4.2** as a colorless oil in quantitative yield. ^1H NMR (300 MHz, CDCl_3 , two rotamers): δ 5.41–5.26 (m, 1H), 4.56–4.52 (m, 1H), 3.95–3.60 (m, 2H), 3.80, 3.75 (two s, 3H), 2.65–2.15 (m, 2H), 2.09, 2.00 (two d, $J = 128.1$, 128.1 Hz, 3H), 2.08 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 172.0, 171.9, 169.6, 169.3 (d, $J = 51.6$ Hz), 72.5, 70.9, 58.4, 57.1, 53.0, 52.6, 52.2, 51.4, 36.9, 34.8, 22.2, 22.1 (enriched), 21.6 (enriched). ESI-MS: m/z 253.0875 ($[\text{M}+\text{Na}]^+$); 253.0881 ($[\text{M}+\text{Na}]^+$, calcd).

***N*-(2- ^{13}C -Acetyl)-(2S,4R)-4-trifluoroacetoxypyrrolidine methyl ester (4.3).**

Compound **4.1** (102 mg, 0.54 mmol) was dissolved in dry CH_2Cl_2 (1 mL). Trifluoroacetic anhydride (0.09 mL, 0.64 mmol) was added to the resulting solution. The reaction mixture was stirred for approximately 2 h, and then concentrated under reduced pressure to give **4.3** as a colorless oil in quantitative yield. ^1H NMR (300 MHz, CDCl_3 , two rotamers): δ 5.62–5.50 (m, 1H), 4.65–4.57 (m, 1H), 4.11–4.03 (m, 1H), 3.82, 3.77 (two s, 3H), 3.81–3.79 (m, 1H), 2.73–2.33 (m, 2H), 2.15, 2.06 (two d, $J = 128.7$, 128.4 Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 171.4, 171.2, 170.4 (d, $J = 51.8$ Hz), 158.5

(app. d, $J = 41.5$ Hz), 156.7 (app. d, $J = 43.1$ Hz), 114.1 (q, $J = 283.8$ Hz), 76.6, 75.2, 58.1, 57.1, 53.0, 52.5, 51.2, 36.4, 34.4, 21.7 (enriched), 20.8 (enriched). ^{19}F NMR (282 MHz, CDCl_3) δ -77.6, -79.2. ESI-MS: m/z 307.0609 ($[\text{M}+\text{Na}]^+$); 307.0599 ($[\text{M}+\text{Na}]^+$, calcd).

***N*-(2- ^{13}C -Acetyl)-(2*S*,4*S*)-4-hydroxyproline methyl ester (4.4) and *N*-(2- ^{13}C -acetyl)-(2*S*,4*S*)-4-(2- ^{13}C)acetoxypoline methyl ester (4.5).** (2*S*,4*S*)-4-Hydroxyproline methyl ester hydrochloride (0.334 g, 1.84 mmol) was dissolved in DMF (4 mL), and the resulting solution was cooled to 0°C. Triethylamine (0.8 mL, 5.74 mmol) was added, upon which a white solid formed. 2- ^{13}C -Acetyl chloride was then added, and the reaction mixture was stirred for 3 h while being allowed to warm to room temperature. The reaction mixture was concentrated under reduced pressure. The residue was suspended in EtOAc, and this suspension was filtered. The filtrate was concentrated to a yellow oil, and the product was purified by flash chromatography, eluting with MeOH (5% v/v) in CH_2Cl_2 to yield **4.4** as a colorless oil (0.235 g, 68%). ^1H NMR (300 MHz, CDCl_3 , two rotamers): δ 4.52 (dd, $J = 9.6, 2.4$ Hz, 1H), 4.49–4.42 (m, 1H), 4.04 (d, $J = 8.1$ Hz, 1H), 3.79, 3.77 (two s, 3H), 3.73–3.57 (m, 2H), 2.40–2.29 (m, 1H), 2.10, 2.02 (two d, $J = 129.0, 129.0$ Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 174.2, 170.3 (d, $J = 51.2$ Hz), 70.7, 68.6, 58.8, 57.3, 56.4, 55.1, 52.5, 39.3, 37.1, 22.1 (enriched), 21.9 (enriched). ESI-HRMS: m/z 211.0772 ($[\text{M}+\text{Na}]^+$); 211.0776 ($[\text{M}+\text{Na}]^+$, calcd).

From the reaction mixture, the bis- ^{13}C -labeled *N,O*-diacetylated compound **4.5** (0.045 g, 10.5%) was also isolated as a colorless oil. ^1H NMR (300 MHz, CDCl_3 , two rotamers): δ 5.32–5.27 (m, 1H), 4.71 (dd, $J = 9.0, 3.0$ Hz, 0.7 Hz), 4.48 (dd, $J = 8.7, 1.8$ Hz, 0.3H), 3.87 (dd, $J = 11.4, 5.4$ Hz, 1H), 3.79, 3.74 (two s, 3H), 3.69–3.64 (m, 1H),

2.59–2.28 (m, 2H), 2.10, 2.05 (two d, $J = 128.1, 128.4$ Hz, 3H), 2.03, 1.98 (two d, $J = 130.2, 129.9$ Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 171.4, 171.3, 170.1, 169.8, 72.7, 71.3, 70.9, 58.7, 57.3, 56.8, 56.5, 53.0, 52.5, 52.2, 50.8, 37.1, 36.9, 34.7, (22.2, 22.1, 20.9, labeled C). ESI-MS (mono- ^{13}C -labeled): m/z 231.1069 ($[\text{M}+\text{H}]^+$); 231.1062 ($[\text{M}+\text{H}]^+$, calcd).

***N*-(2- ^{13}C -Acetyl)-(2S,4S)-4-trifluoroacetoxypoline methyl ester (4.6).**

Compound **4.4** (30 mg, 0.16 mmol) was dissolved in CH_2Cl_2 (1 mL) and trifluoroacetic anhydride (22.5 μL , 0.16 mmol) was added in one portion. The reaction mixture was stirred for 2 h, at which point thin-layer chromatography indicated that starting material was still present. Another portion of trifluoroacetic anhydride (22.5 μL , 0.16 mmol) was added. The reaction mixture was stirred an additional 2 h, concentrated under reduced pressure, and placed under vacuum overnight to give **4.6** as a colorless oil in quantitative yield. ^1H NMR (300 MHz, CDCl_3 , two rotamers): δ 5.57–5.52 (m, 1H), 4.43 (dd, $J = 8.1, 2.7$ Hz, 0.5 H), 4.62 (d, $J = 8.4$ Hz, 0.5H), 4.03–3.82 (m, 2H), 3.78, 3.72 (two s, 3H), 2.76–2.48 (m, 2H), 2.17, 2.13 (two d, $J = 129.0, 129.0$ Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 171.9 (d, $J = 50.1$ Hz), 159.3–156.4 (m), 114.5 (q, $J = 284.0$), 77.4, 59.1, 57.5, 53.3, 53.1, 52.9, 52.5, 36.9, 35.1 (22.1, 22.0, labeled C). ^{19}F NMR (282 MHz, CDCl_3) δ -74.0, -74.1. ESI-MS: m/z 285.0768 ($[\text{M}+\text{H}]^+$); 285.0779 ($[\text{M}+\text{H}]^+$, calcd).

(ProHyp(C(O)CH₃)Gly)₁₀ (4.9). (Pro-Hyp-Gly)₁₀ (Peptides International, Louisville, KY; 3.9 mg, 0.0013 mmol) was dissolved in 6 M HCl (5.6 μL) and glacial acetic acid (11.2 μL) and cooled to 0°C. Acetyl chloride (30 μL , 0.42 mmol) was added, and the resulting mixture was agitated gently for approximately 25 min. The product was

precipitated from diethyl ether and collected by centrifugation. The resulting solid was washed several times with small portions of diethyl ether and then dried under vacuum to yield (ProHyp(C(O)CH₃)Gly)₁₀ (**4.9**) (3.0 mg, 75%) as a white solid. MS (MALDI) m/z 3111.3 (MH⁺ 3110.4).

X-Ray Crystallography. The crystals of **4.2** and **4.5** used for X-ray structure determination grew slowly from the oils obtained after their purification by chromatography. The structure determinations were performed as described previously (Jenkins et al., 2003).

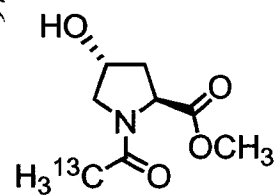
Measurement of K_{ZE} . Values of K_{ZE} for **4.1–4.6** were determined by NMR spectroscopy using a Bruker NMR spectrometer operating at 300.00 MHz for ¹H NMR (**4.1**, **4.3**, and **4.6**) or at 74.43 MHz for ¹³C NMR using a 5-mm broadband probe (**4.1**, **4.2**, **4.4**, and **4.5**). Identical results were obtained for **4.1** using either ¹H or ¹³C NMR spectroscopy. NMR samples of **4.1–4.6** were prepared at concentrations of 10 mM in 1,4-dioxane-*d*₈ (99 atom %). Doubling or halving the concentration of **4.1** and **4.3** in the sample did not alter the values, indicating the absence of effects on K_{ZE} from intermolecular interactions. Experiments were performed using a temperature range of 22–52°C. The temperature settings of the spectrometer were calibrated to within 1°C by reference to a 100% ethylene glycol standard.

Measurement of T_m . The structure and stability of a triple helix of peptide **4.9** was assessed by CD spectroscopy using an Aviv 202 SF instrument equipped with an automated temperature controller. A 0.2 mM solution of peptide **4.9** in 50 mM HOAc was incubated at 4°C for 24 h. Aliquots of 300 μL were placed in 0.1-cm pathlength quartz cuvettes that had been equilibrated at 5°C. Wavelength scans were performed from

200–260 nm at 25 and 85°C, with a slit width of 1 nm and an averaging time of 3 s.

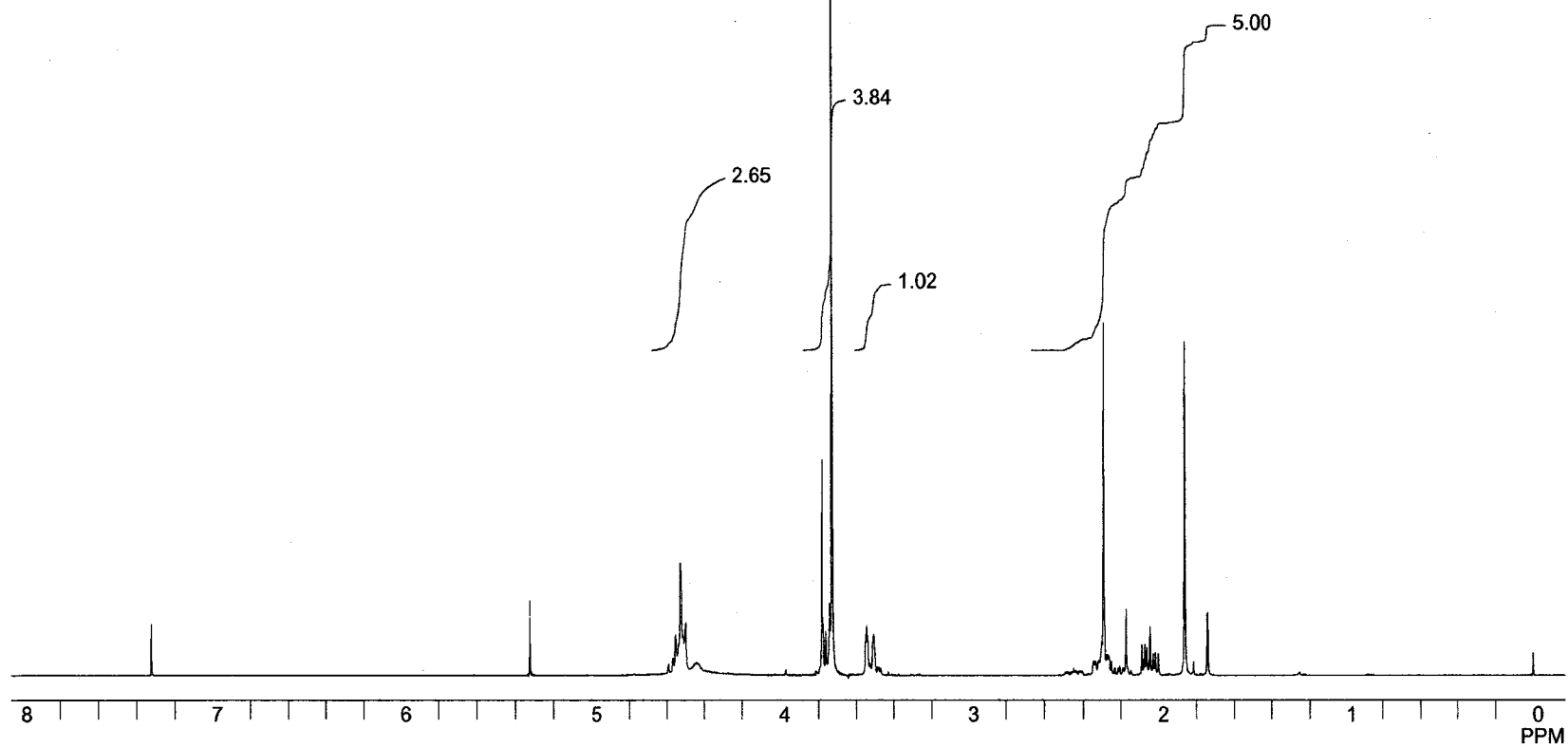
Thermal denaturation experiments were performed by raising the temperature from 5 to 45°C in 3-°C steps, equilibrating for 5 min at each temperature, and monitoring at 225 nm with a 20-s averaging time. Values of T_m , which is the temperature at the midpoint of the thermal transition, were determined in triplicate by fitting the data to a two-state model using the software package NLREG v4.0 (Philip Sherrod).

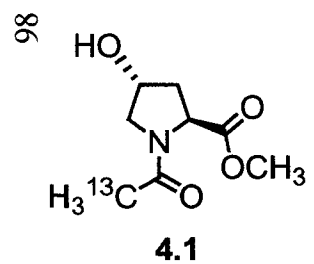
97



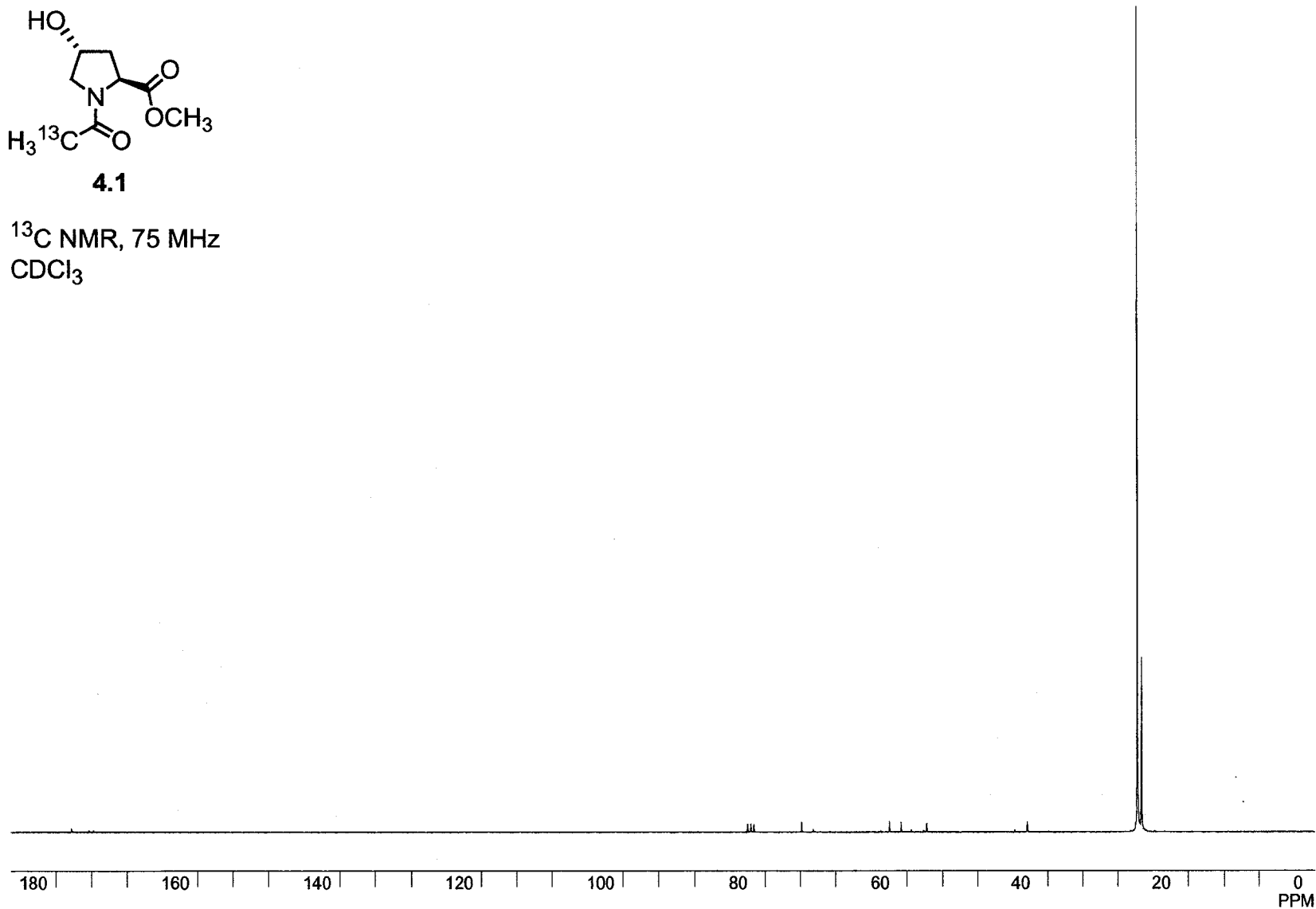
4.1

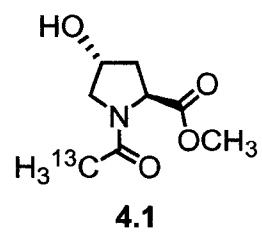
¹H NMR, 300 MHz
CDCl₃



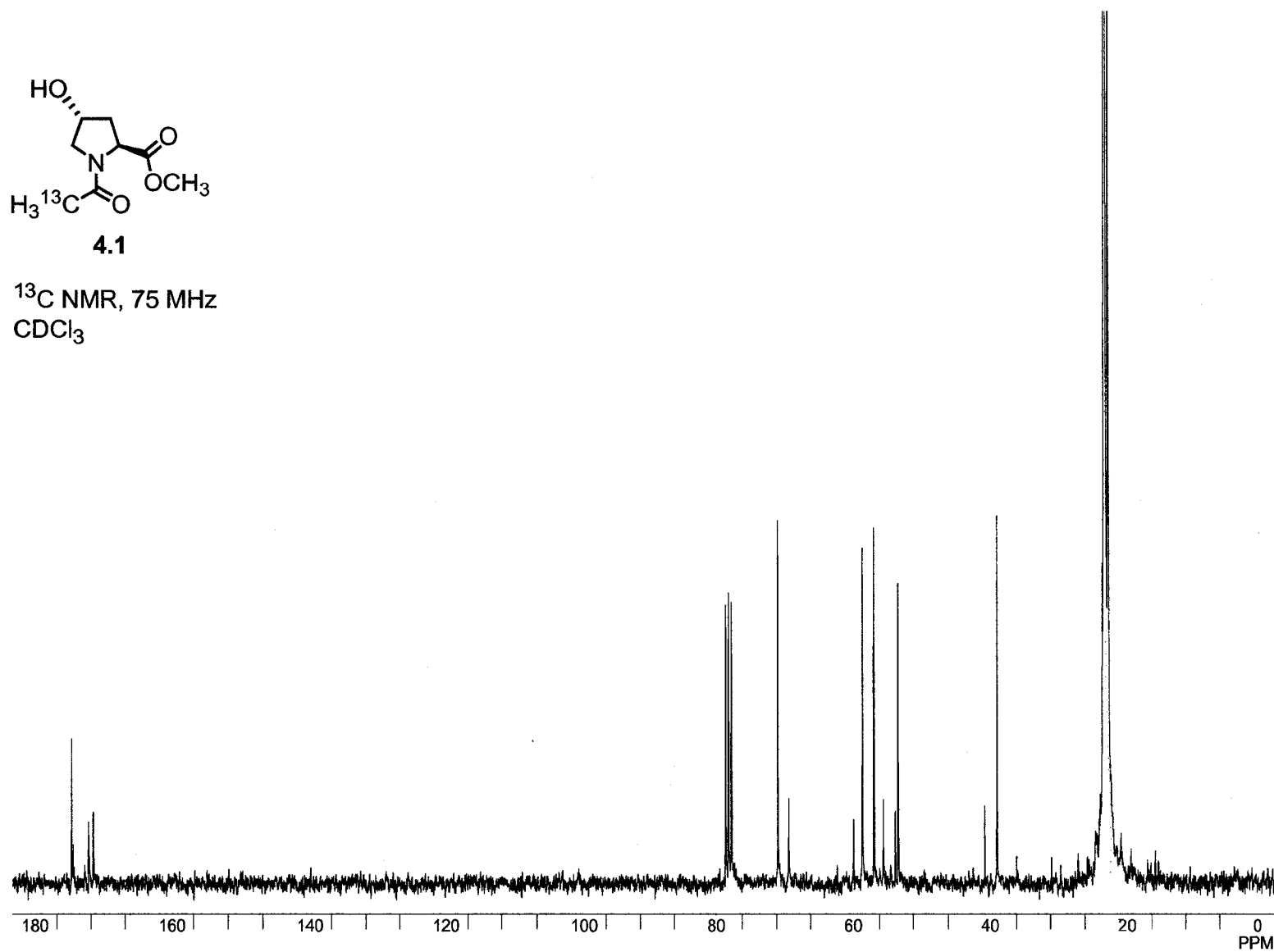


^{13}C NMR, 75 MHz
 CDCl_3

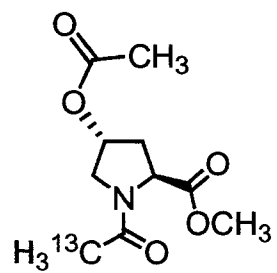




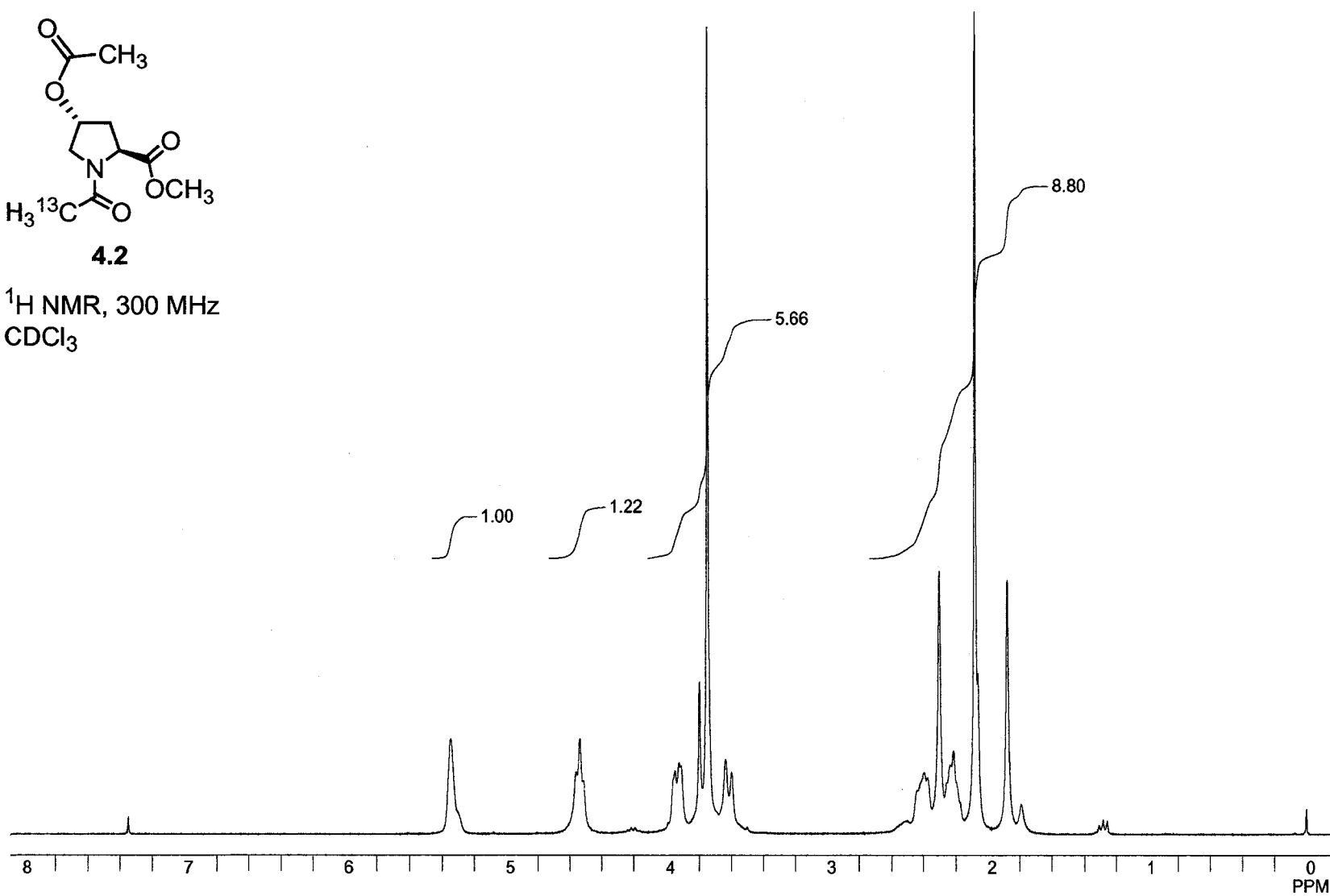
^{13}C NMR, 75 MHz
 CDCl_3

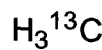
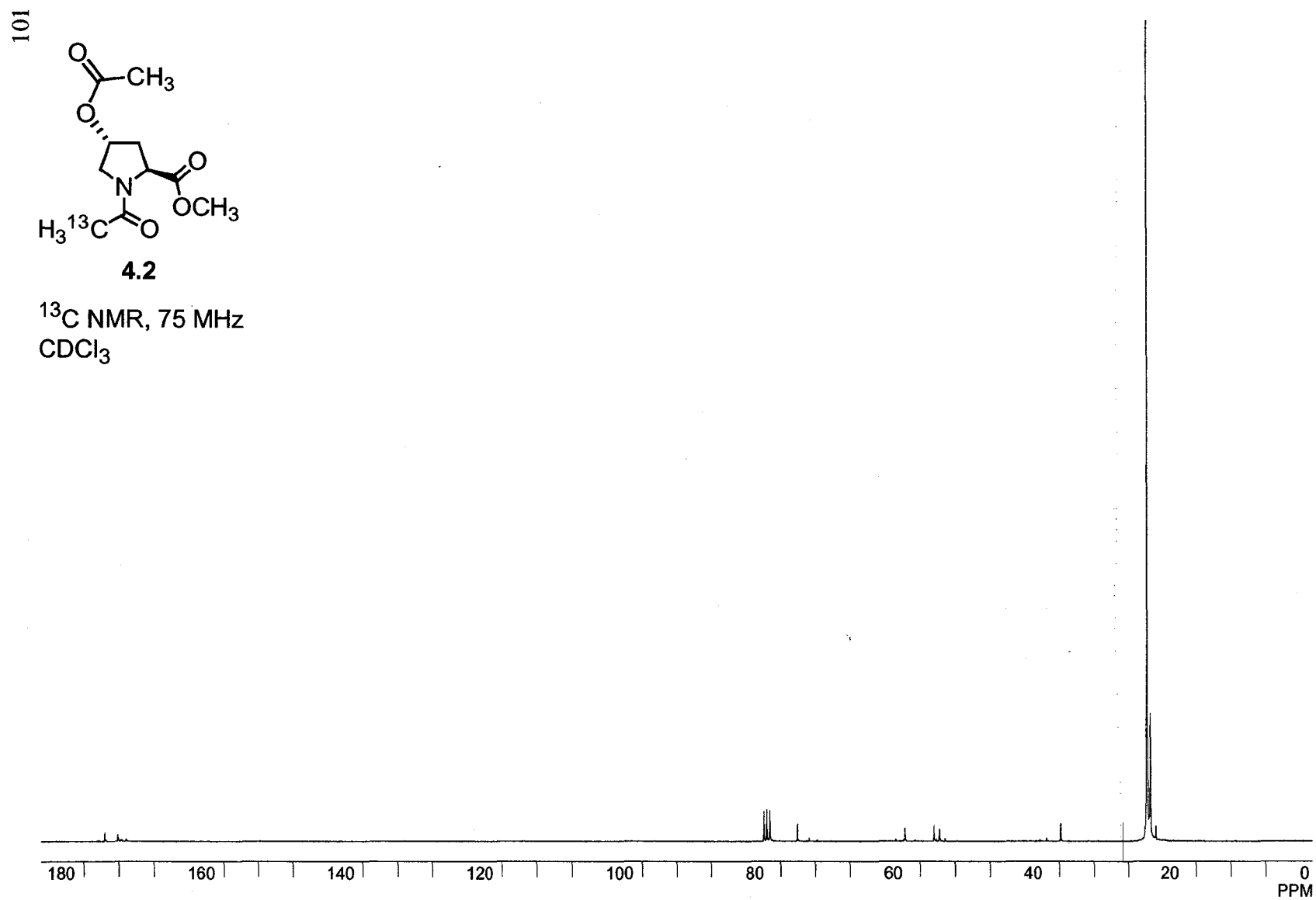


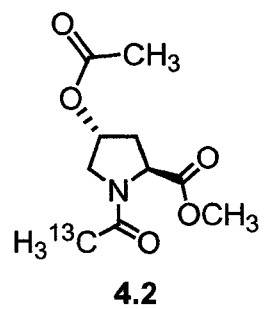
100

**4.2**

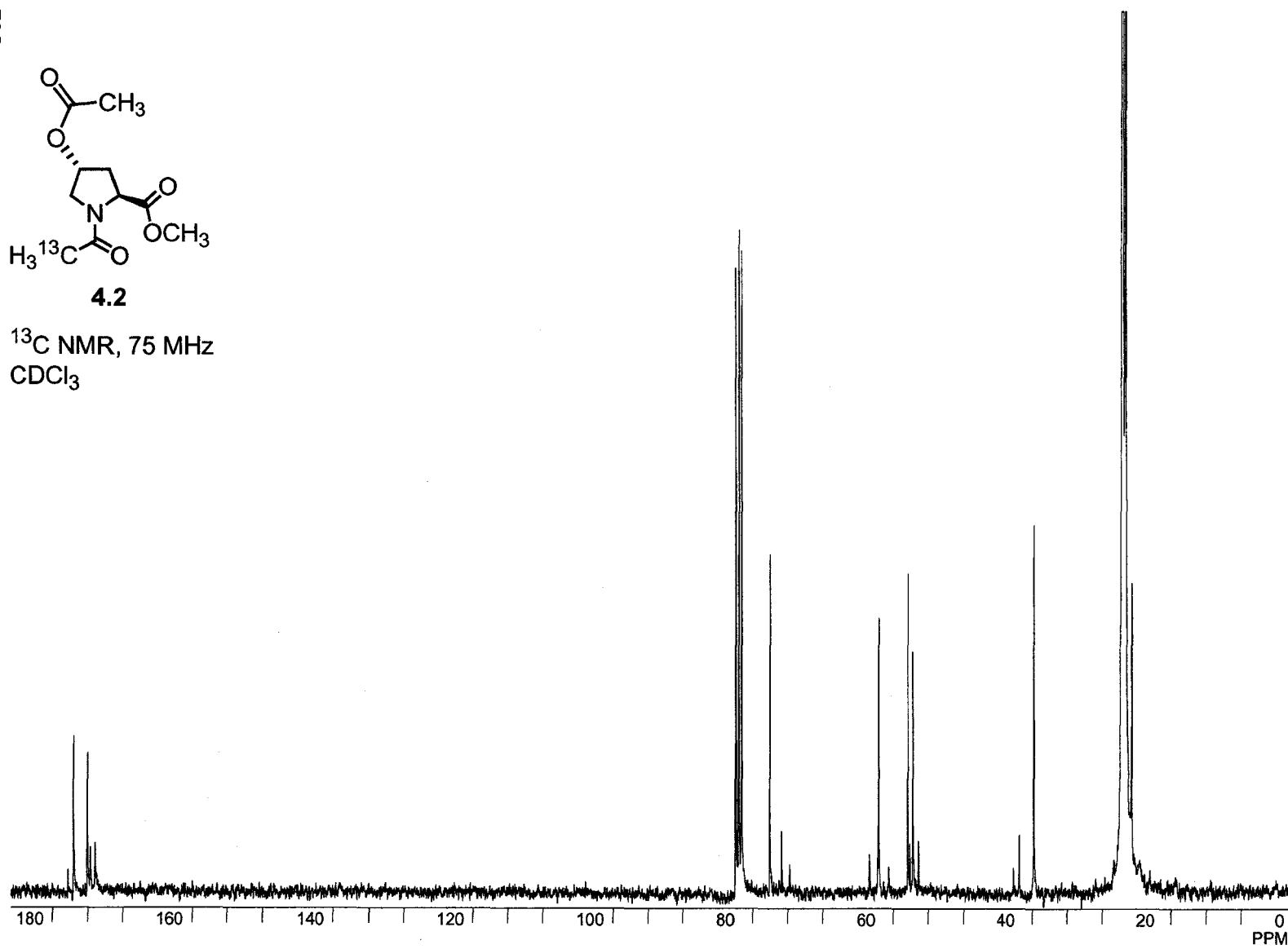
^1H NMR, 300 MHz
 CDCl_3



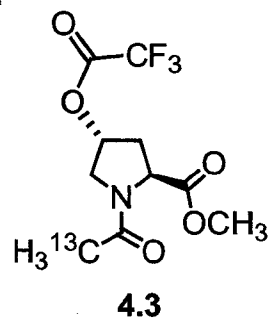
¹³C NMR, 75 MHz
CDCl₃



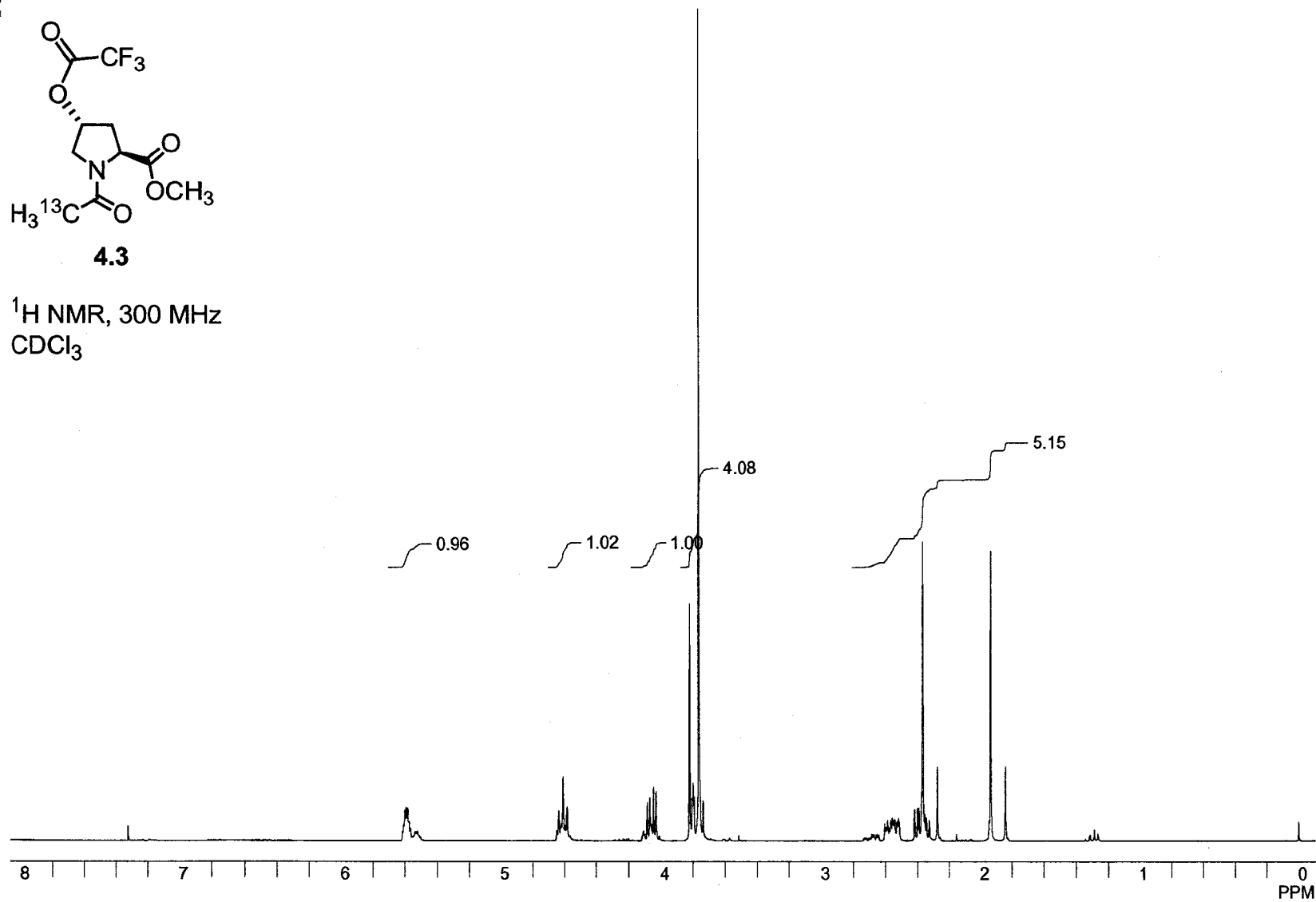
^{13}C NMR, 75 MHz
 CDCl_3

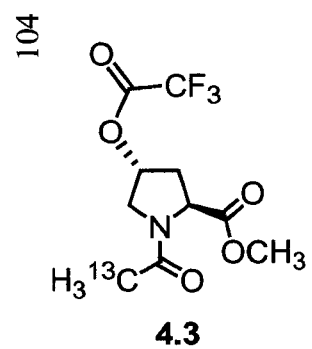


103

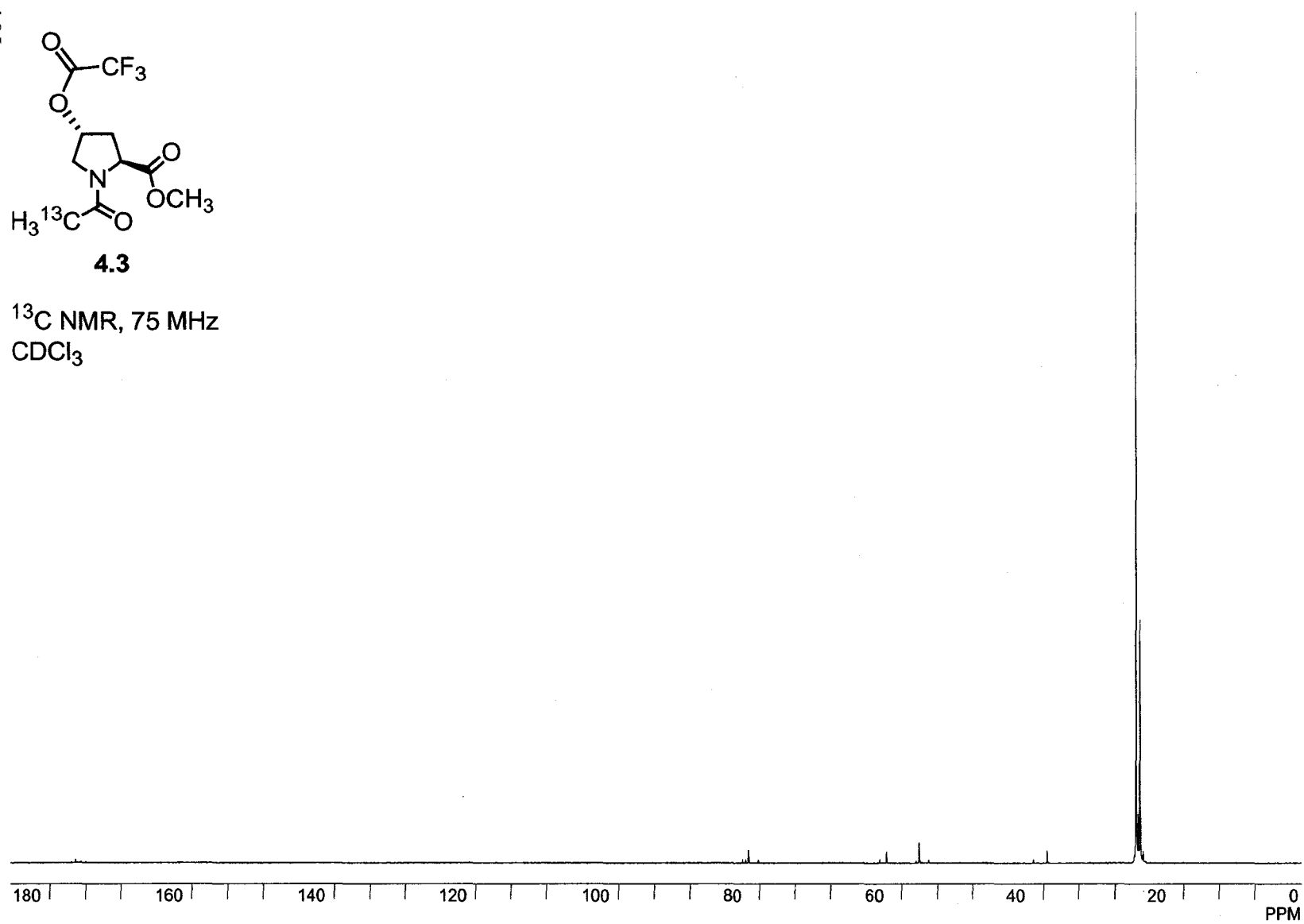


^1H NMR, 300 MHz
 CDCl_3

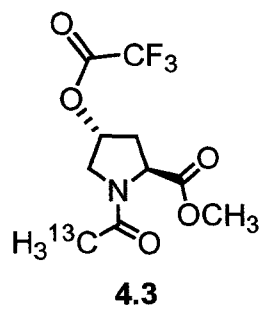




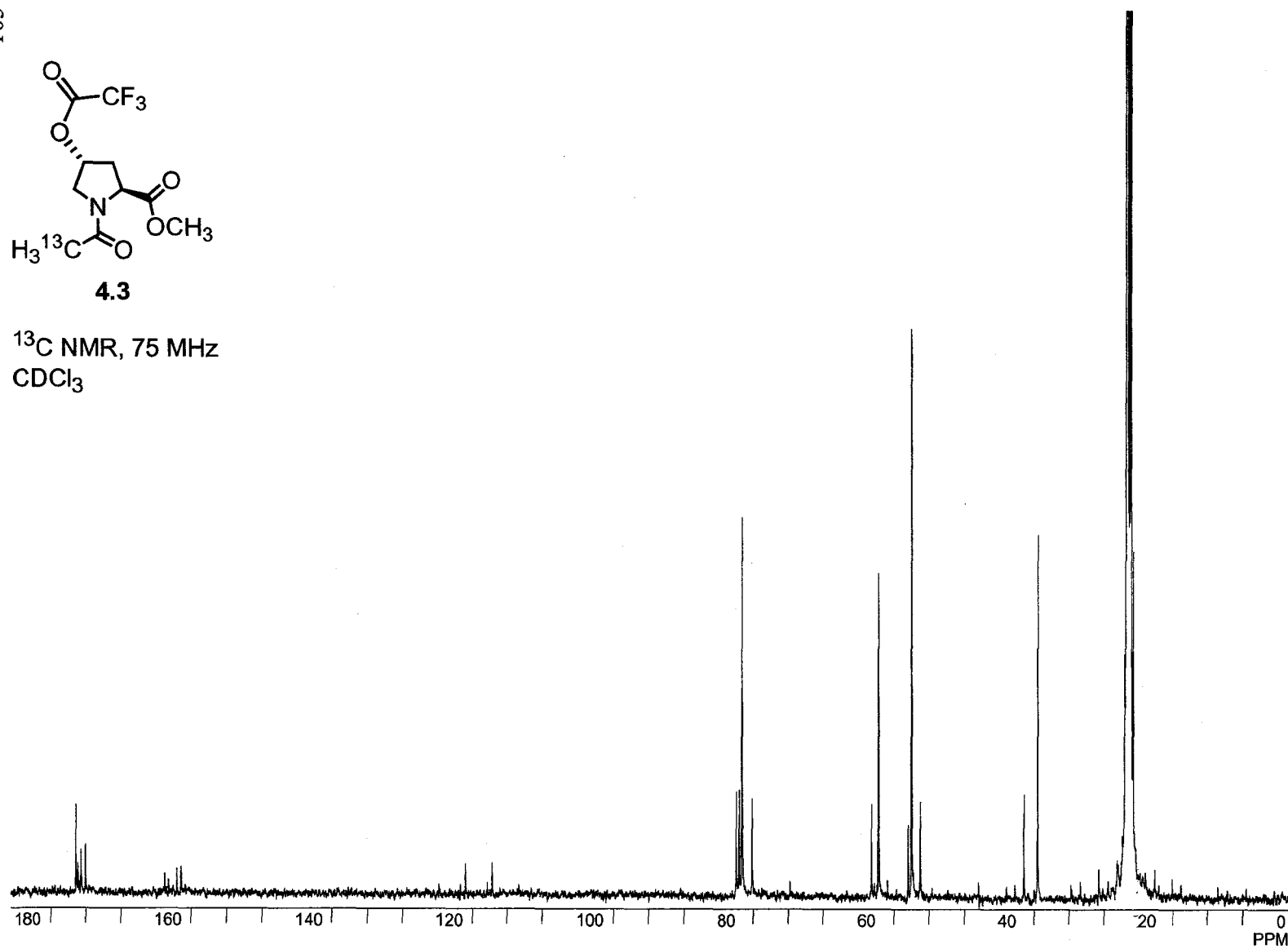
^{13}C NMR, 75 MHz
 CDCl_3



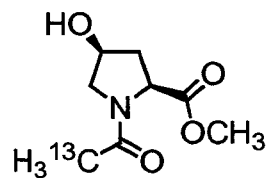
105



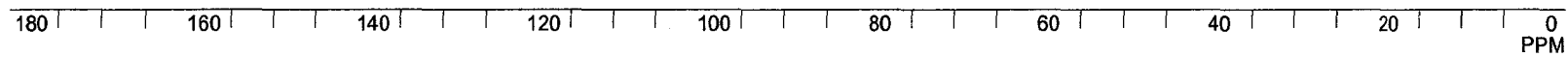
^{13}C NMR, 75 MHz
 CDCl_3

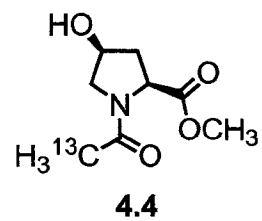


106

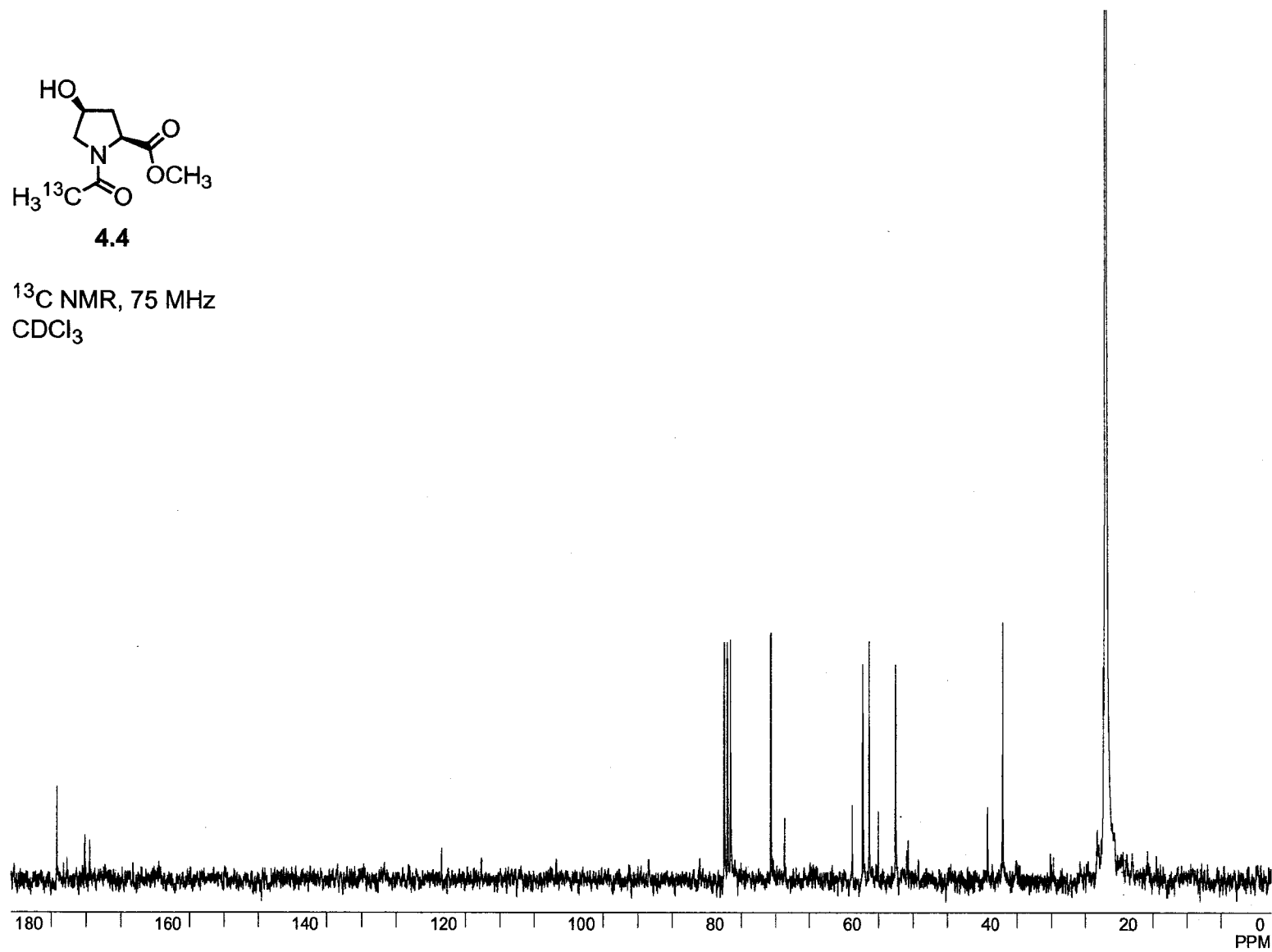
**4.4**

¹³C NMR, 75 MHz
CDCl₃

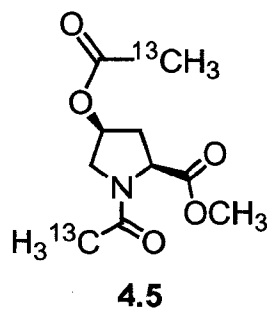




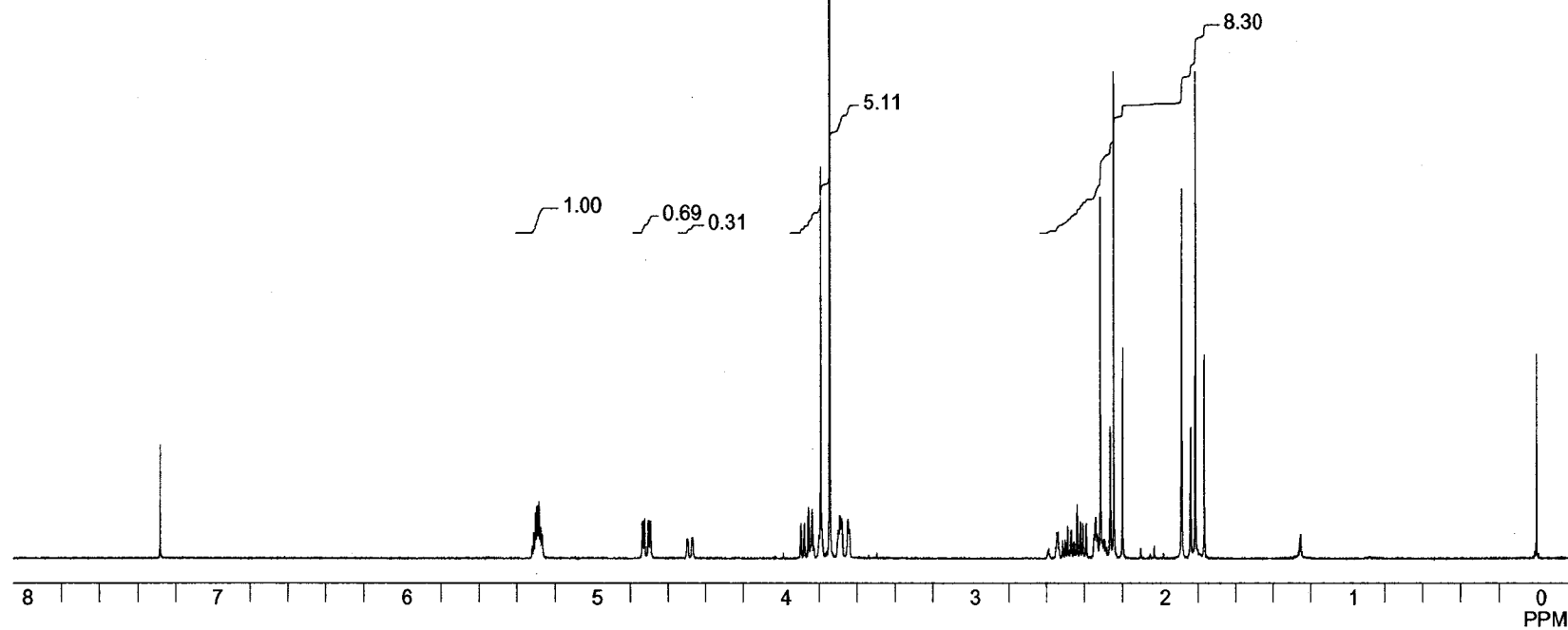
^{13}C NMR, 75 MHz
 CDCl_3

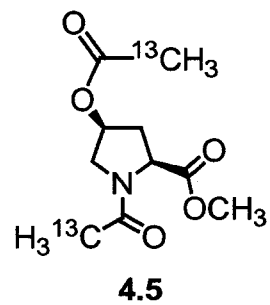


108

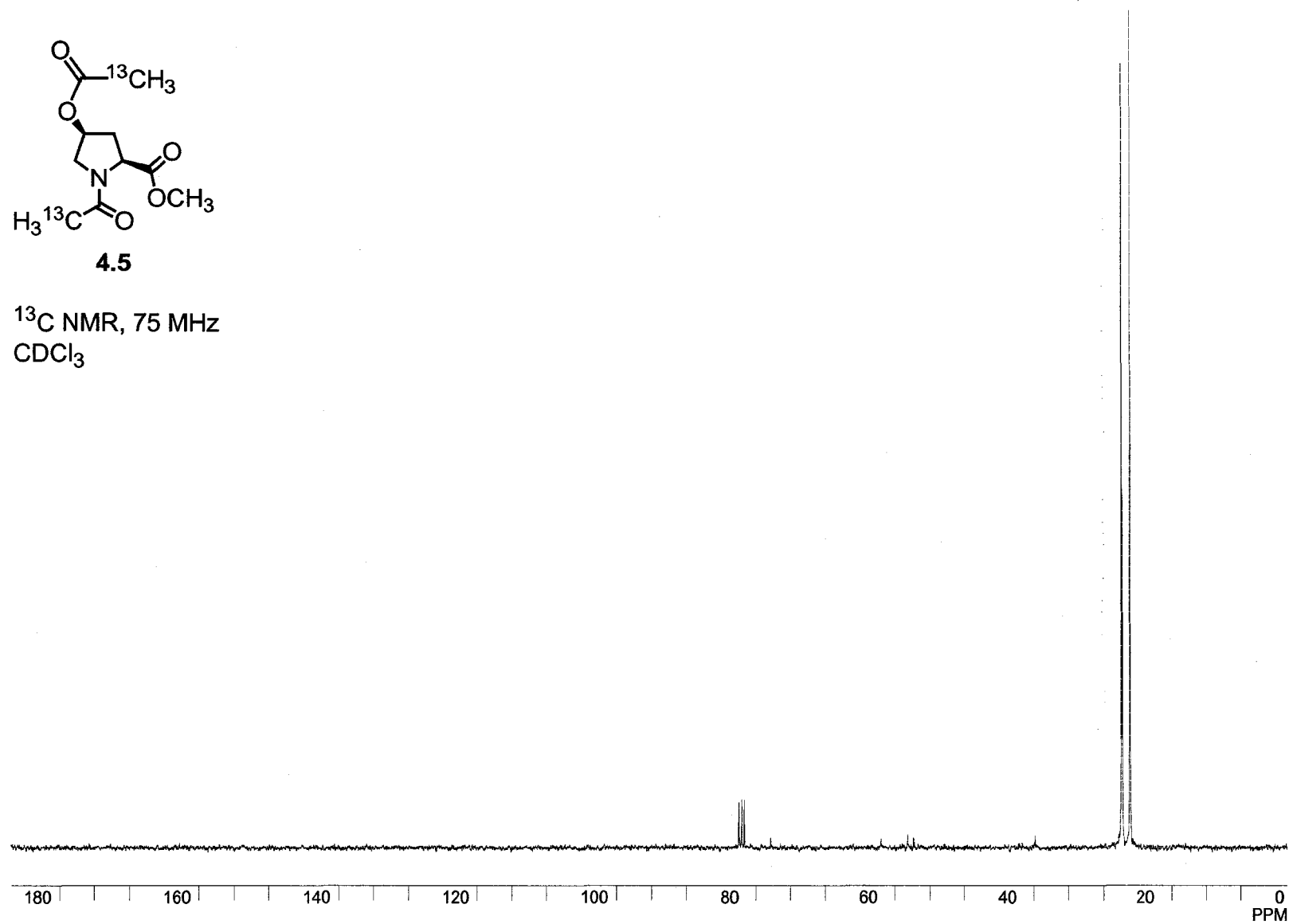


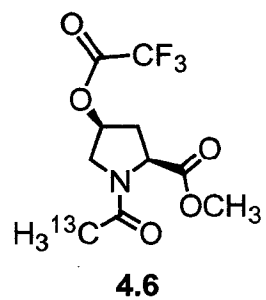
^1H NMR, 300 MHz
 CDCl_3



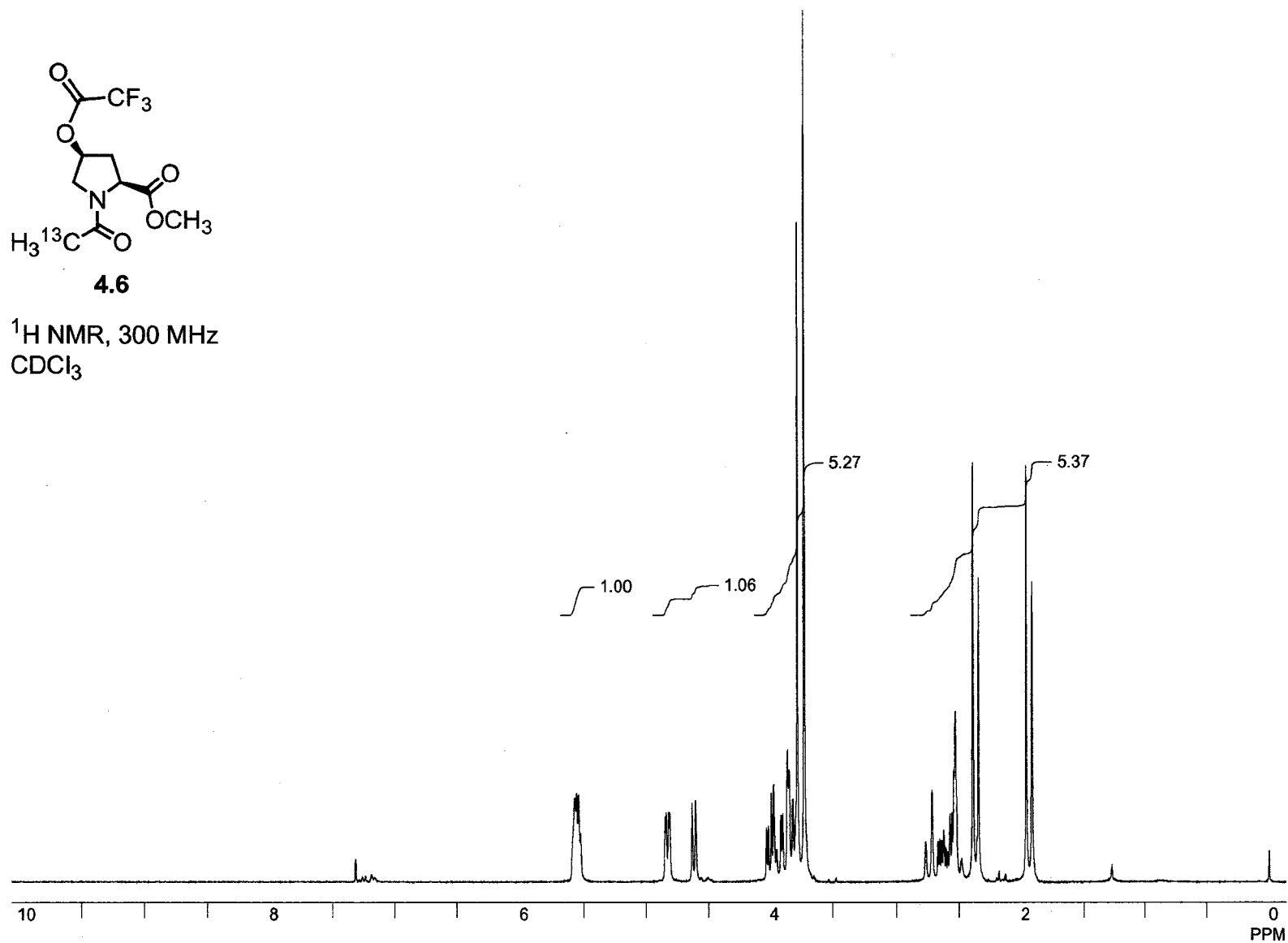


^{13}C NMR, 75 MHz
 CDCl_3

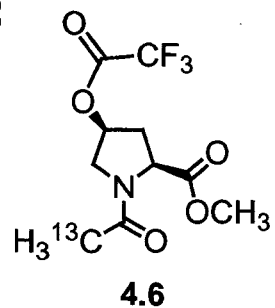




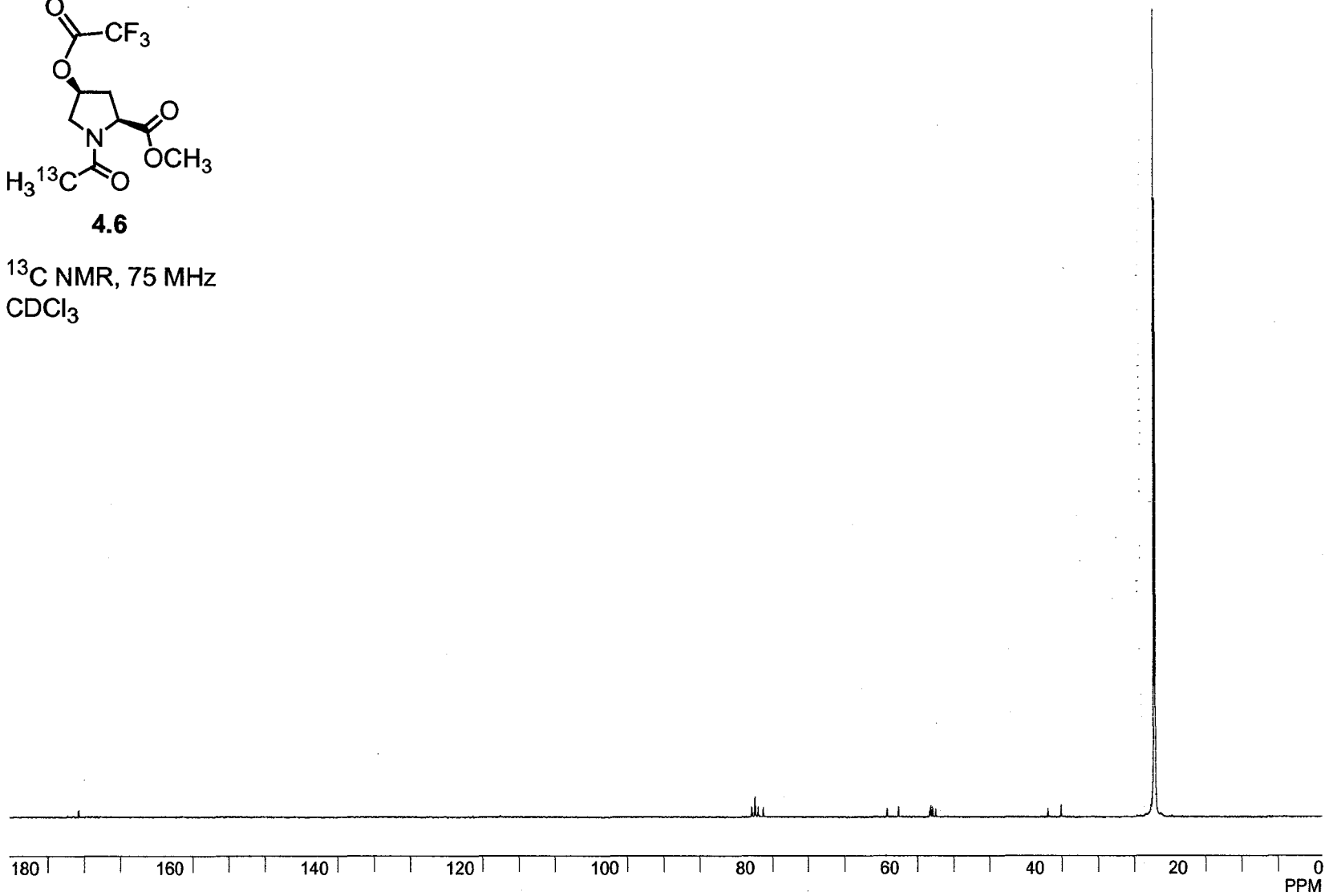
^1H NMR, 300 MHz
 CDCl_3

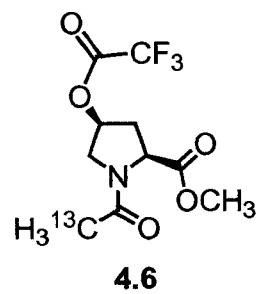


111

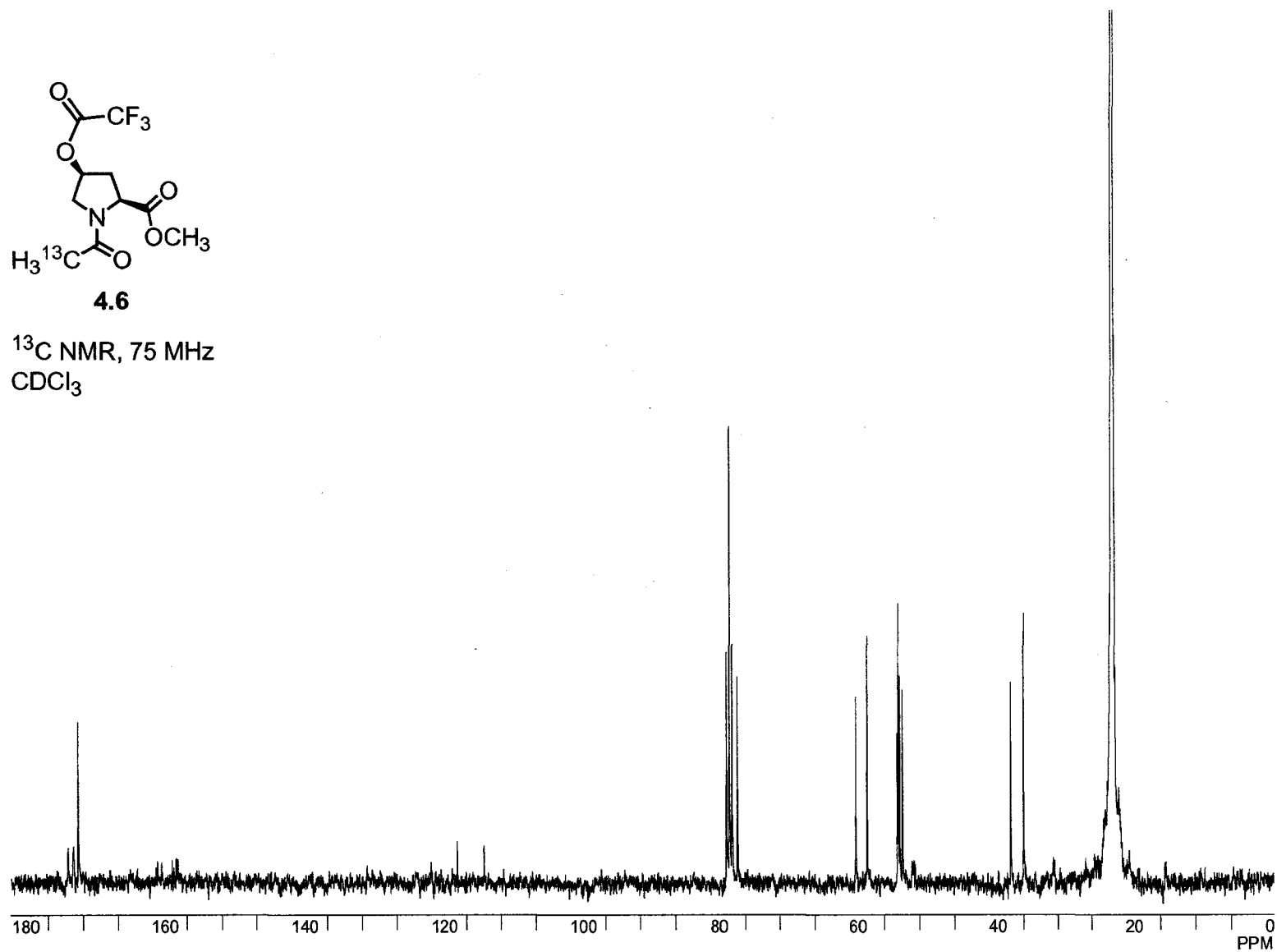


^{13}C NMR, 75 MHz
 CDCl_3





^{13}C NMR, 75 MHz
 CDCl_3



Chapter 5

BINDING OF NONNATURAL 3'-NUCLEOTIDES TO RNASE A

5.1 Introduction

Ribonucleases catalyze the cleavage of RNA. These enzymes are abundant in living systems, where they play a variety of roles systems (D'Alessio & Riordan, 1997; Beintema & Kleineidam, 1998). For example, angiogenin is a homolog of bovine pancreatic ribonuclease (RNase A (Raines, 1998); EC 3.1.27.5) that promotes neovascularization. Angiogenin relies on its ability to cleave RNA for its angiogenic activity (Shapiro et al., 1989; Shapiro & Vallee, 1989). An effective inhibitor of the ribonucleolytic activity of angiogenin could diminish its angiogenic activity, which is an effective means to limit tumor growth (Olson et al., 1995). Selective ribonuclease inhibitors could also be useful tools in studying the roles of various ribonucleases in vitro and in vivo (Leonidas et al., 1999).

Known nucleotide-based inhibitors of ribonucleases have been based on three strategies. Most common are competitive inhibitors that resemble RNA. Shapiro and coworkers have developed especially potent inhibitors of RNase A based on two nucleosides linked by a pyrophosphoryl group (Russo et al., 1997; Russo & Shapiro,

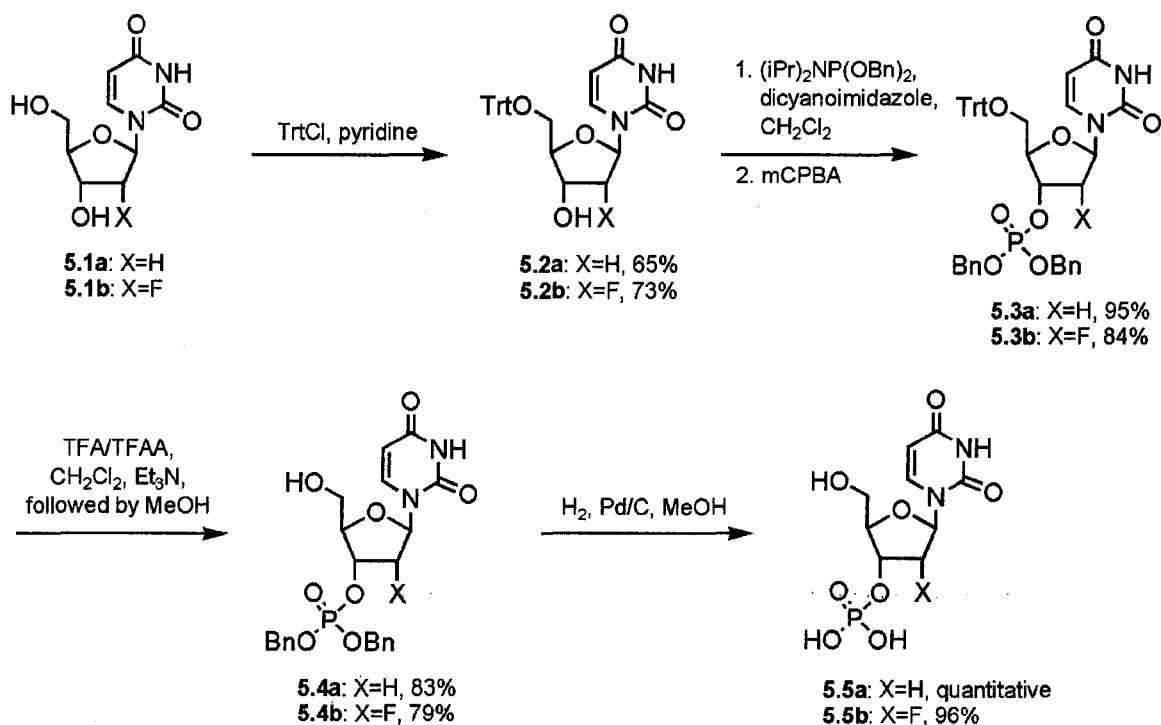
1999; Russo et al., 2001). Using a different approach, Widlanski and coworkers showed that 3'-(4-(fluoromethyl)phenyl phosphate)uridine is a mechanism-based inactivator of RNase A (Stowell et al., 1995). Finally, a new strategy has used an *N*-hydroxyurea nucleotide to recruit zinc(II) and thereby inhibit microbial ribonucleases (Higgin et al., 2003; Makarov et al., 2004). Each of these strategies is based on nucleotides containing a ribose or deoxyribose ring. Is that choice optimal?

Here, we synthesize two 3'-nucleotides containing nonnatural furanose rings: 2'-fluoro-2'-deoxyuridine 3'-phosphate and arabinouridine 3'-phosphate. We measure the inhibition of wild-type RNase A and its T45G variant by these nonnatural 3'-nucleotides and find that both are significantly more potent than deoxyuridine 3'-phosphate and uridine 3'-phosphate, two 3'-nucleotides containing natural furanose rings. These results indicate that the use of nonnatural furanose rings can increase the potency of nucleotide-based inhibitors of RNase A.

5.2 Results

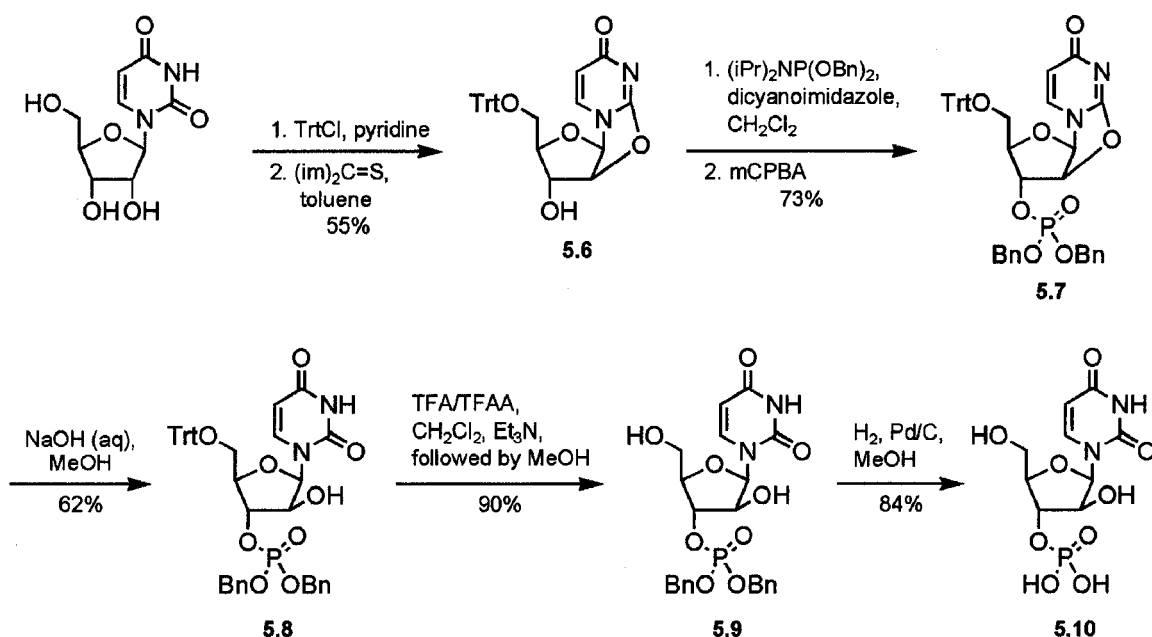
The syntheses of deoxyuridine 3'-phosphate (dUMP, **5.5a**) and 2'-fluoro-2'-deoxyuridine 3'-phosphate (dU^FMP, **5.5b**) were accomplished starting from unprotected nucleosides as shown in Scheme 5.1. Briefly, **5.1** was protected at the 5'-position by treating it with trityl chloride in dry pyridine at reflux to yield **5.2** (Cordington et al., 1964). Subsequent phosphorylation at the 3'-position was achieved by reacting **5.2** with diisopropyl dibenzyl phosphoramidite in the presence of 4,5-dicyanoimidazole followed with oxidation with 3-chloroperoxybenzoic acid to yield **5.3** (Perich & Johns, 1987; Giner & Ferris, 2002). Deprotection was achieved in two steps, removing the trityl group

first in a mixture of dry trifluoroacetic acid and trifluoroacetic anhydride, followed by addition of methanol to complete the initial deprotection (Krainer & Naider, 1993). The resulting dibenzyl phosphate **5.4** was deprotected by hydrogenolysis using Pd/C as the catalyst to give **5.5** in 46–51% overall yield. A synthesis for dUMP has been reported, but involves a phosphorylating agent that is not available commercially (Taktakishvili & Nair, 2000). dU^FMP has also been synthesized previously, but its synthesis involves harsh conditions and an involved purification (Antonov et al., 1976). The advantages of the route in Scheme 5.1 include the mild conditions and high yield of the phosphorylation step, the use of commercially available reagents, the facile deprotection of the trityl group, and the high purity of the final product after debenzylation.



Scheme 5.1 *Synthetic Route to dUMP and dU^FMP*

Arabinouridine 3'-monophosphate (araUMP, **5.10**) was synthesized from uridine according to Scheme 5.2. Uridine was protected at the 5'-position by treatment with trityl chloride in pyridine, then reacted with thiocarbonyldiimidazole to give 5'-trityl-O²,2'-cyclouridine, **5.6**, in moderate yield (Fox & Wempen, 1965). Compound **5.6** was then phosphorylated at the 3'-position to provide **5.7** (Perich & Johns, 1987; Giner & Ferris, 2002). The protected arabinouridine monophosphate **5.8** was generated by treatment of **5.7** with one equivalent of aqueous sodium hydroxide in methanol (Sato et al., 1994), followed by deprotection as described above. This is a novel route to araUMP, starting with inexpensive, commercially available starting materials. Two other routes have been reported: one starts from a 2',3'-epoxyuridine derivative that is not available commercially (Roussev et al., 1997), and results in a mixture of isomers; the other starts from cytidine 2',3'-cyclic phosphate (Nagyvary, 1969; Pollard & Nagyvary, 1973), a relatively expensive starting material.



Scheme 5.2 Synthetic Route to araUMP.

The phosphoryl pK_a values of 3'-UMP, dUMP, dU^FMP, and araUMP were measured by using ³¹P NMR spectroscopy, and are listed in Table 5.1. The pK_a values of 3'-UMP, araUMP, and dU^FMP are within error of each other, whereas the pK_a value of dUMP is, as expected from a previous report (Walz, 1971), greater than that of the other three. The differences in pK_a values likely arise from inductive effects. No stereoelectronic component is apparent, as the phosphoryl groups of UMP and araUMP have the same pK_a values.

Table 5.1 *Values of pK_a and K_i for 3'-nucleotides*

Nucleoside 3'-phosphate	pK_a	K_i (μ M)	
		Wild-type RNase A	T45G RNase A
3'-UMP	5.84 ± 0.05	39 ± 2	89 ± 9
dUMP	6.29 ± 0.07	18 ± 3	≥ 1700
dU ^F MP	5.89 ± 0.10	5.5 ± 0.7	181 ± 15
araUMP	5.85 ± 0.06	6 ± 1	≥ 1000

^a Phosphoryl group pK_a values were determined by ³¹P NMR spectroscopy.

^b Inhibition was determined in MES–NaOH buffer, pH 6.0, containing NaCl (50 mM).

The values of K_i for the four 3'-nucleotides were measured by their ability to inhibit the cleavage of the fluorogenic substrate 6-FAM–dArU(dA)₂–TAMRA by wild-type RNase A and its T45G variant (Kelemen et al., 1999), and are listed in Table 5.1. All four 3'-nucleotides were potent inhibitors of the wild-type enzyme, whereas inhibition of T45G RNase A was less pronounced—by up to three orders-of-magnitude.

5.3 Discussion

3'-Nucleotides with a nonnatural furanose ring can have a greater affinity for wild-type RNase A than do 3'-nucleotides with a ribose or deoxyribose ring. The K_i values increase in the order: $\text{dU}^{\text{F}}\text{MP} \approx \text{araUMP} < \text{dUMP} < 3'\text{-UMP}$ (Table 5.1). 3'-Nucleotides with a dianionic phosphoryl group are known to be more potent inhibitors of RNase A than those with a monoanionic phosphoryl group (Russo et al., 2001). The difference in pK_a values between dUMP and the other three nucleoside 3'-phosphates is approximately 0.4 pK_a units (Table 5.1). Likewise, the K_i value of dUMP with wild-type RNase A is approximately 3-fold higher than those of $\text{dU}^{\text{F}}\text{MP}$ and araUMP. The relative affinity of araUMP for wild-type RNase A measured herein is twofold greater than that reported previously (Pollard & Nagyvary, 1973). 3'-UMP has a higher K_i value with wild-type RNase A than would be expected based on its pK_a . This weaker binding could arise from the 2'-OH group participating in more unfavorable interactions with the enzyme than do the smaller 2'-groups (F or H) of the other 3'-nucleotides. These unfavorable interactions could be reinforced by the tight interaction between the uracil base and Thr45, and by a slight preference for the C3'-endo (N) conformation of 3'-UMP in solution (Davies & Danyluk, 1975) (Figure 5.1). It is interesting to note that in the structure of the crystalline RNase A·3'-UMP complex (Leonidas et al., 2003), the ribose ring is found in the C2'-endo rather than the C3'-endo conformation (Figure 5.2). Like 3'-UMP, $\text{dU}^{\text{F}}\text{MP}$ resides in the C3'-endo conformation (Antonov et al., 1976), but has a smaller fluoro group in its 2'-position. dUMP (Guschlbauer & Jankowski, 1980) and araUMP (Chwang & Sundaralingam, 1973; Venkateswarlu & Ferguson, 1999) are predominantly in the C2'-endo (S) conformation and both have hydrogens where 3'-UMP has a larger hydroxyl

group. The weaker binding of 3'-UMP compared to dUMP and araUMP lends support to the hypothesis that ground-state destabilization contributes to the catalytic prowess of RNase A (Kelemen et al., 2000).

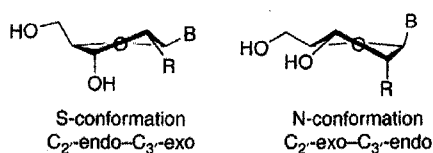


Figure 5.1 *S*- and *N*-conformation of nucleosides. *R* = *H* favors *S*-conformation; *R* = *OH*, *F* favors *N*-conformation.

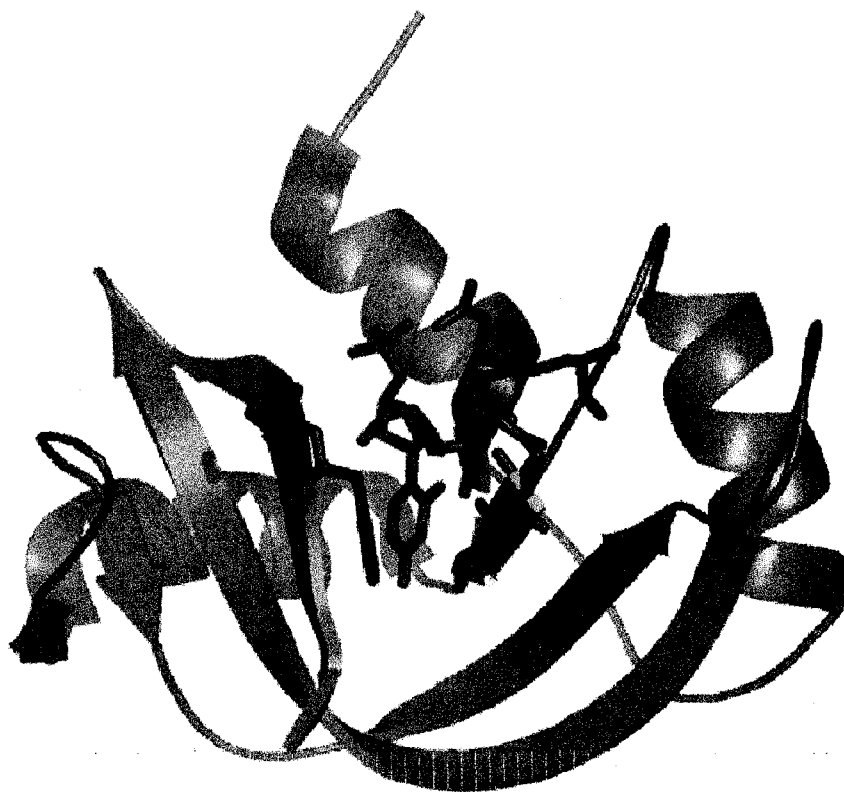


Figure 5.2 3'-UMP bound to wild-type RNase A. The protein backbone is in blue, 3'-UMP is in magenta and the active site residue side chains are in green.

Another possible reason for the K_i value of 3'-UMP being greater than those of the other 3'-nucleotides with wild-type RNase A is that the energetic penalty of desolvating the 2'-OH group is higher than for any of the other 2'-substituents. The 2'-H and -F of dUMP and dU^FMP are more easily desolvated than is a hydroxyl group, due to their lack of ability form hydrogen bonds with water (Howard, Judith A. K. et al., 1996), and the 2'-arabino hydroxyl group of araUMP might not need to be desolvated for that 3'-nucleotide to bind to RNase A because it is oriented away from the active-site residues.

The binding of 3'-nucleotides by T45G RNase A follows a much different trend than does binding by the wild-type enzyme. The K_i values increase in the order: 3'-UMP < dU^FMP << araUMP, dUMP. 3'-UMP has half of the affinity for T45G RNase A than for the wild-type enzyme, whereas dU^FMP has 3% of the affinity. The K_i values for dUMP and araUMP could not be measured, but were estimated to be ≥ 1.7 and ≥ 1 mM, respectively. The general decrease in binding between T45G RNase A and the inhibitors underlines the importance of the interaction between Thr45 and pyrimidine nucleobases in substrate binding (delCardayre & Raines, 1994, 1995; Kelemen et al., 2000). The binding of 3'-nucleotides to T45G RNase A appears to be sensitive to the pucker of the furanose ring: 3'-UMP and dU^FMP, which both reside in the C3'-endo conformation, bind with high micromolar affinities, whereas dUMP and araUMP, which both reside predominantly in the C2'-endo conformation, have at least millimolar affinities. Without the anchoring presence of Thr45, the inhibitor can orient its furanose ring and phosphoryl group so as to optimize favorable contacts with active-site residues. Thus, 3'-UMP could form a hydrogen bond between its 2'-OH and His12 or other active-site residues instead of participating in sterically unfavorable interactions. This putative hydrogen bond could

be the source of the two-fold higher affinity of T45G RNase A for 3'-UMP than dU^FMP. dU^FMP cannot form a hydrogen bond with its 2'-fluoro substituent, but has a favorable conformation and pK_a for binding to the active site at pH 6.

Although araUMP has the same phosphoryl group pK_a value as 3'-UMP and dU^FMP, it has at least an order of magnitude lower affinity for T45G RNase A. Ring pucker could explain the lack of affinity between araUMP and T45G RNase A (compared to 3'-UMP and dU^FMP). When the furanose ring is in the C2'-endo conformation, the nucleobase and phosphoryl group are placed in such a way that neither can interact optimally with active-site residues in the absence of the anchoring site provided by Thr45 in the wild-type enzyme. This lack of favorable interactions should be true for dUMP, as well. dUMP has the additional handicap of having a higher phosphoryl group pK_a than the other 3'-nucleotides, with consequently diminished Coulombic interactions.

Wild-type RNase A appears to be less sensitive to furanose ring pucker in its binding to substrates and 3'-nucleotides than is the T45G variant. The two most effective inhibitors of wild-type RNase A herein are dU^FMP and araUMP, which have different puckers in solution but indistinguishable K_i values. It has already been shown that in the context of a deoxyribose tetranucleotide ligand, a 2'-fluoro-2'-deoxyuridine residue binds more weakly to RNase A than does a 2'-deoxyuridine residue (Kelemen et al., 2000). Hence, we suggest that a most effective residue in creating new inhibitors of RNase A is arabinouridine, which combines a favorable monoester pK_a value with little or no steric hindrance with the RNase A active site.

5.4 Experimental Section

General. Reagents obtained from commercial sources were used without further purification. 2'-Fluoro-2'-deoxyuridine was prepared according to the procedure of Maruyama and coworkers (Sato et al., 1994). Wild-type RNase A and its T45G variant were prepared according to procedures reported previously (delCardayre & Raines, 1994);(delCardayre & Raines, 1995; delCardayre et al., 1995). Dry dichloromethane was drawn from a Baker Cycletainer. Thin-layer chromatography was performed using aluminum-backed plates coated with silica gel containing F₂₅₄ phosphor and visualized by UV illumination or staining with I₂, *p*-anisaldehyde or phosphomolybdic acid. NMR spectra were obtained with a Bruker AC-300 or Varian UNITY-500 spectrometer. Mass spectra were obtained with a Micromass electrospray ionization (ESI) instrument.

5'-Trityl-2'-deoxyuridine (5.2a). 2'-Deoxyuridine (**5.1a**; 0.492 g, 2.16 mmol) was placed in a dry, 50-mL round-bottom flask. Triphenylmethyl chloride (0.714 g, 2.56 mmol) and pyridine (10 mL) were added, and the reaction mixture was stirred for 48 h at room temperature under Ar(g) atmosphere. The reaction mixture was concentrated, and the residue was dissolved in dichloromethane and washed once with 1 M HCl and twice with water. The organic layer was dried over MgSO₄(s), filtered, and concentrated. The residue was crystallized from ethyl acetate/hexanes to yield **5.2a** as a white powder (661 mg, 65.0%). ¹H NMR (300 MHz, DMSO-*d*₆): δ 7.80 (d, *J* = 8.1 Hz, 1H), 7.22–7.41 (m, 15H), 6.29 (t, *J* = 6.3 Hz, 1H), 5.35 (d, *J* = 8.1 Hz, 1H), 4.54 (dt, *J* = 6.1, 3.9 Hz, 1H), 4.06 (dd, *J* = 3.1, 6.8 Hz, 1H), 3.44 (d, *J* = 3.1 Hz, 2H), 2.18–2.49 (ABMX, *J*_{AB} = 13.7 Hz, *J*_{AX} = 6.3 Hz, *J*_{AM} = 4.2 Hz, *J*_{BX} = 6.4 Hz, *J*_{BM} = 0 Hz, 2H).

5'-Trityl-2'-fluoro-2'-deoxyuridine (5.2b). 2'-Fluoro-2'-deoxyuridine (**5.1b**;

3.164 g, 12.85 mmol) was dissolved in dry pyridine (20 mL) and concentrated to an oil under reduced pressure. The resulting oil was dissolved in dry pyridine (50 mL), and trityl chloride (5.485 g, 19.58 mmol) was added, followed by additional dry pyridine (18 mL). The reaction mixture was heated to reflux under Ar(g) for 3 h and 45 min, and then concentrated under reduced pressure to a yellow oil. Residual pyridine was removed by azeotroping with toluene under reduced pressure. The resulting oil was dissolved in dichloromethane and washed once with 1 M HCl, once with saturated NaHCO₃(aq), and once with water. The organic layer was dried over MgSO₄(s), filtered and concentrated. The crude product was purified by silica gel chromatography, eluting with MeOH (2.5–5% v/v) in CH₂Cl₂ to yield **5.2b** as a white solid (4.587 g, 73.1%). ¹H NMR (300 MHz, CDCl₃ + CD₃OD): δ 7.95 (d, J = 8.4 Hz, 1H), 7.42–7.24 (m, 15H), 6.04 (dd, J = 16.5, 1.2 Hz, 1H), 5.28 (d, J = 8.1 Hz, 1H), 4.98 (ddd, $^2J_{F-H}$ = 52.2, J = 4.3, 1.0 Hz, 1H), 4.53 (ddd, J = 22.2, 8.7, 4.2 Hz, 1H), 4.13 (bd, J = 8.4 Hz, 1H), 3.58 (m, 2H). ¹³C NMR (75.4 MHz, CDCl₃ + CD₃OD): δ 163.93, 150.12, 142.91, 139.95, 128.42, 127.73, 127.14, 101.92, 93.58 (d, J = 188.0 Hz), 87.75 (d, J = 34.6 Hz), 87.29, 81.53, 68.01 (d, J = 16.7 Hz), 60.93. ¹⁹F NMR (282.1 MHz, CDCl₃ + CD₃OD): δ –201.44 (m). ESI–MS ($M + Na$): 511.1651 (observed), 511.1645 (calculated).

5'-Trityl-2'-deoxyuridine 3'-dibenzylphosphate (5.3a). Dicyanoimidazole (127 mg, 1.08 mmol) was suspended in dry dichloromethane (25 mL) in an oven-dried 100-mL round-bottom flask equipped with a stir bar. Diisopropyldibenzylphosphoramidite (220 μ L, 0.98 mmol) was added to the suspension at room temperature, and the mixture was allowed to stir for 1.25 h. 5'-Trityl-2'-deoxyuridine (**5.2a**, 175 mg, 0.37 mmol) suspended

in dry CH_2Cl_2 (10 mL) was added, and the reaction mixture was stirred at room temperature for an additional 1.25 h. The reaction mixture was then cooled to 0 °C, and solid *m*-chloroperoxybenzoic acid (351 mg) was added in one portion. The reaction mixture was stirred at 0 °C for approximately 15 min, the ice bath was removed, and the reaction mixture was stirred at room temperature for an additional 1. h. The reaction mixture was poured into a separatory funnel containing ethyl acetate and washed three times with aqueous $\text{Na}_2\text{S}_2\text{O}_5$ (10% w/v), three times with saturated aqueous NaHCO_3 (75 mL total), twice with 1 N HCl (50 mL total), once with water, and once with brine. The organic layer was then dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure. The resulting solid was purified by silica gel chromatography, eluting with CH_2Cl_2 (2.5% v/v) in MeOH to yield **5.3a** as a white solid (256 mg, 94.7%). ^1H NMR (300 MHz, CDCl_3): δ 9.80 (bs, 1H), 7.63 (d, J = 8.4 Hz, 1H), 7.34–7.25 (m, 25H), 6.30 (dd, J = 7.8, 6.3 Hz, 1H), 5.31 (dd, J = 8.4, 1.8 Hz, 1H), 5.09–4.98 (m, 5H), 4.18 (m, 1H), 3.34 (m, 1H), 2.49 (ddd, J = 13.8, 5.7, 2.1 Hz, 1H), 2.24–2.15 (m, 1H). ^{13}C NMR (75.4 MHz, CDCl_3): δ 163.36, 150.22, 142.84, 139.68, 135.27 (d, J = 6.0 Hz), 128.64 (d, J = 2.0 Hz), 128.55 (d, J = 1.8 Hz), 128.51, 128.00 (d, J = 2.6 Hz), 127.38, 87.61, 84.45, 84.32 (d, J = 6.0 Hz), 77.57 (d, J = 5.2 Hz), 69.62 (t, J = 5.6 Hz), 62.94, 39.20. ^{31}P NMR (121.4 MHz, CDCl_3 , ^1H decoupled): δ –1.40. ESI-MS ($M + \text{Na}$): 753.2338 (observed), 753.2342 (calcd).

5'-Trityl-2'-fluoro-2'-deoxyuridine 3'-dibenzylphosphate (5.3b). The preparation of **5.3b** was carried out in a manner similar to that used for the preparation of **5.3a**. The product was purified by silica gel chromatography, eluting with MeOH (2.5 to 5% v/v) in CH_2Cl_2 to yield **5.3b** as a white solid (4.623 g, 84.1%). ^1H NMR (300 MHz, CDCl_3): δ

9.71 (s, 1H), 7.77 (d, $J = 8.1$ Hz, 1H), 7.38–7.21 (m, 25H), 6.07 (d, $J = 16.2$ Hz, 1H), 5.25 (d, $J = 8.4$ Hz, 1H), 5.20–4.87 (m, 6H), 4.23 (bd, $J = 7.2$ Hz, 1H), 3.54 (m, 2H). ^{13}C NMR (75.4 MHz, CDCl_3): δ 163.11, 149.85, 142.76, 139.52, 135.14 (dd, $J = 6.8, 6.6$ Hz), 128.64, 128.53, 128.01, 127.85, 127.45, 102.57, 91.44 (d, $J = 193.6$ Hz), 87.80 (d, $J = 33.6$ Hz), 87.75, 80.42 (d, $J = 8.7$ Hz), 71.96 (dd, $J = 16.1, 4.4$ Hz), 69.86 (dd, $J = 7.0, 6.5$ Hz), 60.61. ^{19}F NMR (282.1 Hz, CDCl_3): δ -200.03 (ddd, $J = 52.2, 16.5, 14.2$ Hz). ^{31}P NMR (121.4 MHz, CDCl_3 , ^1H decoupled): δ -1.50. ESI-MS ($\text{M} + \text{Na}$): 771.2230 (observed), 771.2248 (calculated).

2'-Deoxyuridine 3'-dibenzylphosphate (5.4a). Compound **5.3a** (270 mg, 0.36 mmol) was dissolved in dry CH_2Cl_2 (5 mL) under Ar(g) , and a solution of trifluoroacetic acid (139 μL , 1.8 mmol) and trifluoroacetic anhydride (255 μL , 1.8 mmol) in dry CH_2Cl_2 (0.6 mL) was added by syringe at room temperature. The reaction mixture, which turned bright yellow, was stirred at room temperature for approximately 10 min, cooled to 0 $^\circ\text{C}$, then stirred for an additional 10 min. Upon addition of triethylamine (250 μL , 1.79 mmol), the bright yellow color disappeared. After 5 min, MeOH (10 mL) was added, the reaction mixture was stirred for an additional 5 min, then concentrated under reduced pressure. The residue was dissolved in CH_2Cl_2 , and washed once with 1 M NaCl. The organic layer was dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure. The resulting product was purified by silica gel chromatography, eluting with MeOH (5% v/v) in CH_2Cl_2 to yield **5.4a** as a white solid (151 mg, 82.6%). ^1H NMR (300 MHz, CDCl_3): δ 8.48 (bs, 1H), 7.61 (d, $J = 7.8$, 1H), 7.38–7.34 (m, 10H), 6.08 (dd, $J = 7.5, 6.0$ Hz, 1H), 5.72 (dd, $J = 8.1, 2.1$ Hz, 1H), 5.12–5.02 (m, 4H), 4.99–4.92 (m, 1H), 4.06 (m, 1H), 3.81–3.67 (m, 2H), 2.67 (bt, 1H), 2.40–2.21 (m, 2H). ^{13}C NMR (75.4 MHz,

CDCl₃): δ 163.73, 150.38, 140.70, 135.17 (d, J = 6.7 Hz), 128.76, 128.61, 128.02, 102.54, 85.75, 77.93, 69.78 (dd, J = 3.2, 5.7 Hz), 61.64, 38.72. ³¹P NMR (121.4 MHz, CDCl₃, ¹H decoupled): δ -1.38. ESI-MS (M + Na): 511.1241 (observed), 511.1246 (calculated).

2'-Fluoro-2'-deoxyuridine 3'-dibenzylphosphate (5.4b). The preparation of **5.4b** was carried out in a similar manner to that used for the preparation of **5.4a**. The crude product was purified by silica gel chromatography, eluting with MeOH (2.5–5% v/v) in CH₂Cl₂ to yield **5.4b** as a white solid (1.601 g, 78.6%). ¹H NMR (300 MHz, CDCl₃ + CD₃OD): δ 7.88 (d, J = 8.1 Hz, 1H), 7.38–7.32 (m, 10H), 6.01 (dd, J = 2.7, 15.3 Hz, 1H), 5.73 (d, J = 8.4 Hz, 1H), 5.13–4.85 (m, 6H), 4.14 (m, 1H), 3.78 (m, 2H). ¹³C NMR (75.4 MHz, CDCl₃ + CD₃OD): δ 163.86, 150.19, 140.31, 134.82 (dd, J = 6.5, 5.8 Hz), 128.69, 128.46, 127.89 (d, J = 5.7 Hz), 102.41, 90.94 (d, J = 195.1 Hz), 87.62 (d, J = 33.8 Hz), 82.40 (d, J = 6.3 Hz), 72.43 (dd, J = 14.7, 4.4 Hz), 70.02 (d, J = 5.8 Hz), 59.21. ¹⁹F NMR (282.13 MHz, CDCl₃ + CD₃OD): δ -202.10 (ddd, J = 52.2, 16.5, 14.0). ³¹P NMR (121.4 MHz, CDCl₃ + CD₃OD, ¹H decoupled): δ -1.55. ESI-MS (M + Na): 529.1171 (observed), 529.1152 (calculated).

2'-Deoxyuridine 3'-phosphate (dUMP, 5.5a). 2'-Deoxyuridine 3'-dibenzylphosphate (**5.4a**, 177 mg, 0.36 mmol) was placed in a 100-mL round-bottom flask and flushed with Ar(g) for approximately 5 min. Palladium on carbon (16 mg) was added and the flask flushed again with Ar(g) for approximately 5 min. A 4:1 solution of MeOH and aqueous NH₄HCO₃ (1% w/v) was added slowly, then H₂(g) was introduced via a balloon. The reaction was stirred under H₂(g) for 4 h, then filtered through a Celite plug, concentrated under reduced pressure to dryness, and placed under vacuum overnight to yield **5.5a** as a

colorless solid (133 mg, 100%). ^1H NMR (300 MHz, D_2O): δ 7.69 (d, $J = 8.1$ Hz, 1H), 6.12 (t, $J = 6.7$ Hz, 1H), 5.70 (d, $J = 8.3$ Hz, 1H), 4.54 (septet, $J = 3.6$ Hz, 1H), 4.00 (m, 1H), 3.60 (ABX, $J_{\text{AB}} = 12.4$, $J_{\text{AX}} = 3.4$, $J_{\text{BX}} = 4.7$ Hz, 2H), 2.42–2.34 (m, 2H). ^{13}C NMR (75.4 MHz, D_2O): δ 167.12, 152.53, 143.05, 103.29, 88.10 (d, $J = 5.7$ Hz), 86.56, 75.83 (d, $J = 4.0$ Hz), 62.25, 39.08. ^{31}P NMR (121.4 MHz, D_2O , ^1H decoupled): δ –0.13. ESI–MS ($\text{M} - \text{H}$): 307.0342 (observed), 307.0331 (calculated).

2'-Fluoro-2'-deoxyuridine 3'-phosphate (dU^FMP, **5.5b)** The preparation of **5.5b** was performed in a manner similar to that for the preparation of **5.5a**, to yield **5.5b** as a colorless solid. (1.010 g, 95.7%). ^1H NMR (300 MHz, D_2O): δ 7.66 (d, $J = 8.1$ Hz, 1H), 5.79 (d, $J = 18.9$ Hz, 1H), 5.64 (d, $J = 8.1$ Hz, 1H), 5.04 (dd, $^2J_{\text{F-H}} = 52.2$, $J = 4.3$ Hz, 1H), 4.34 (dtd, $J = 22.5$, 9.0, 4.2 Hz, 1H), 3.94 (m, 1H), 3.67 (ABX, $J_{\text{AB}} = 13.1$, $J_{\text{AX}} = 2.4$, $J_{\text{BX}} = 4.2$ Hz, 2H). ^{13}C NMR (75.4 MHz, D_2O): δ 167.14, 152.15, 143.68, 103.17, 92.78 (d, $J = 190.3$ Hz), 90.71 (d, $J = 35.5$ Hz), 83.21 (d, $J = 7.8$ Hz), 71.82 (d, $J = 14.8$ Hz), 60.64. ^{19}F NMR (282.13 MHz, D_2O): δ –199.02 (ddd, $J = 52.2$, 19.3, 17.4 Hz). ^{31}P NMR (121.4 MHz, D_2O , ^1H decoupled): δ –0.37. ESI–MS ($\text{M} - \text{H}$): 325.0221 (observed), 325.0237 (calculated).

5'-Trityluridine. Uridine (5.040 g, 20.6 mmol) and freshly distilled pyridine (45 mL) were combined in a dry 200-mL round-bottom flask, and cooled to 0 °C under Ar(g). Trityl chloride (5.770 g, 20.7 mmol) in pyridine (25 mL) was added via syringe. The reaction mixture was stirred, allowing it to warm to room temperature, for 4 days. Ethyl acetate was added to the reaction mixture, which was then transferred to a separatory funnel. The organic layer was washed twice with 2 M HCl, once with saturated $\text{NaHCO}_3(\text{aq})$, then once with brine. The organic layer was dried over $\text{MgSO}_4(\text{s})$, filtered,

and concentrated under reduced pressure. The resulting residue was recrystallized from EtOAc/hexanes to yield 5'-trityluridine as a white solid (6.076 g, 61% yield.) ^1H NMR (300 MHz, $\text{CD}_3\text{OD} + \text{CDCl}_3$): δ 7.97 (d, $J = 8.1$ Hz, 1H), 7.23–7.45 (m, 15H), 5.89 (d, $J = 3.3$ Hz, 1H), 5.26 (d, $J = 8.1$ Hz, 1H), 4.42 (dd, $J = 6.1, 5.2$ Hz, 1H), 4.24 (dd, $J = 5.0, 3.3$ Hz, 1H), 4.13 (dt, $J = 5.9, 2.8$ Hz, 1H), 3.49 (ABX, $J_{\text{AB}} = 11.0$ Hz, $J_{\text{AX}} = 3.0$ Hz, $J_{\text{BX}} = 2.5$ Hz, 2H). ^{13}C NMR (75.4 MHz, DMSO): δ 163.04, 150.50, 143.42, 140.62, 128.31, 128.02, 127.20, 101.48, 88.95, 86.42, 82.34, 73.40, 69.53, 63.25. ESI-MS ($\text{M} + \text{Na}$): 509.1677 (observed), 509.1689 (calculated).

5'-Trityl- O^2 ,2'-cycloauridine (5.6). 5'-Trityluridine (10.692 g, 22.1 mmol) and 1,1'-thiocarbonyldiimidazole (5.087 g, 28.5 mmol) were combined in a 250-mL round-bottom flask. Toluene (120 mL) was added, and the reaction mixture was heated to reflux for 1 h. The reaction was then allowed to cool to room temperature. The tan solid product was removed by filtration, washed with MeOH, and recrystallized from MeOH to yield **5.6** as an off-white solid (9.247 g, 89.7%). ^1H NMR (300 MHz, $\text{DMSO}-d_6$): δ 7.94 (d, $J = 7.4$ Hz, 1H), 7.19–7.32 (m, 15H), 6.33 (d, $J = 5.5$ Hz, 1H), 5.99 (d, $J = 4.6$ Hz, 1H), 5.86 (d, $J = 7.5$ Hz, 1H), 5.21 (d, $J = 6.4$ Hz, 1H), 4.24 (m, 1H), 4.06 (m, 1H), 2.89 (m, 2H). ^{13}C NMR (75.4 MHz, DMSO): δ 170.88, 159.25, 143.30, 136.66, 128.02, 127.94, 127.08, 108.88, 89.70, 88.44, 86.64, 85.96, 74.73, 63.01. ESI-MS ($\text{M} + \text{Na}$): 491.1578 (observed), 491.1583 (calculated).

5'-Trityl- O^2 ,2'-cycloauridine 3'-dibenzylphosphate (5.7). The preparation of **5.7** was carried out in a manner similar to that used for the preparation of **5.3a**. The product was purified by silica gel chromatography, eluting with MeOH (5% v/v) in CH_2Cl_2 to yield **5.7** as a white solid (3.417 g, 72.7%). ^1H NMR (300 MHz, CDCl_3): δ 7.32–7.23 (m,

25H), 7.16 (d, $J = 7.2$ Hz, 1H), 6.03 (d, $J = 5.4$ Hz, 1H), 5.87 (d, $J = 7.5$ Hz, 1H), 5.11–4.92 (m, 6H), 4.51 (dd, $J = 7.2, 6.9$ Hz, 1H), 2.85 (m, 2H). ^{13}C NMR (75.4 MHz, CDCl_3): δ 171.07, 158.66, 142.81, 134.90 (d, $J = 5.5$ Hz), 134.10, 128.91 (d, $J = 3.5$ Hz), 128.69 (d, $J = 3.0$ Hz), 128.22, 127.94, 127.35, 110.34, 89.91, 87.11, 86.18 (d, $J = 6.3$ Hz), 85.62 (d, $J = 4.5$ Hz), 79.73 (d, $J = 5.4$ Hz), 70.16 (d, $J = 5.9$ Hz), 62.01. ^{31}P NMR (121.4 MHz, CDCl_3 , ^1H decoupled): δ -1.95. ESI-MS ($\text{M} + \text{Na}$): 751.2151 (observed), 751.2185 (calculated).

5'-Trityl-arabinouridine 3'-dibenzylphosphate (5.8). Methanol (40 mL) was added to **5.6** (2.983 g, 4.09 mmol) in a 100-mL round bottom flask. The solid dissolved at first and then precipitated, so CH_2Cl_2 was added until the solution was clear again. 1 M NaOH (4 mL, 4 mmol) was added dropwise via a Pasteur pipette, and the reaction was stirred at room temperature for 8 h. The bulk of the solvent was evaporated under reduced pressure, and then the residue was dissolved in dichloromethane and washed with water. The pH of the water layer was ~ 13 . Glacial acetic acid (0.5 mL) was added, the layers were separated, and the organic layer was washed again with water. The layers were separated, and the organic layer was washed twice with saturated $\text{NaHCO}_3(\text{aq})$ and twice with brine. The organic layer was dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with 1:1 EtOAc/ CH_2Cl_2 to yield **5.8** as a white solid (1.888 g, 61.8%).

^1H NMR (300 MHz, CDCl_3): δ 10.02 (bs, 1H), 7.59 (d, $J = 8.1$ Hz, 1H), 7.41–7.17 (m, 25H), 6.15 (d, $J = 4.8$ Hz, 1H), 5.35 (d, $J = 8.1$ Hz, 1H), 4.98–4.89 (m, 5H), 4.81 (m, 1H), 4.55 (m, 1H), 4.03 (m, 1H), 3.40 (m, 2H). ^{13}C NMR (75.4 MHz, CDCl_3): δ 164.05, 150.56, 143.19, 141.94, 135.15 (d, $J = 6.1$ Hz), 128.63, 128.53, 128.03 (d, $J = 2.8$ Hz),

127.90, 127.23, 100.99, 87.25, 85.24, 81.05 (d, $J = 5.9$ Hz), 80.94 (d, $J = 9.8$ Hz), 74.50, 69.86 (d, $J = 5.4$ Hz), 62.03. ^{31}P NMR (121.4 MHz, CDCl_3 , ^1H decoupled): δ –1.17. ESI–MS ($\text{M} + \text{Na}$): 769.2303 (observed), 769.2291 (calculated).

Arabinouridine 3'-dibenzylphosphate (5.9). The preparation of **5.9** was carried out in a manner similar to that used for the preparation of **5.4a**. The crude product was purified by silica gel chromatography, eluting with MeOH (5% v/v) in CH_2Cl_2 to yield **5.9** as a light yellow solid (1.084g, 89.6%). ^1H NMR (300 MHz, CDCl_3): δ 10.33 (bs, 1H), 7.70 (d, $J = 8.4$ Hz, 1H), 7.34–7.26 (m, 10H), 6.04 (d, $J = 3.9$ Hz, 1H), 5.60 (d, $J = 8.4$ Hz, 1H), 5.44 (d, $J = 6.3$ Hz, 1H), 5.05–4.98 (m, 4H), 4.82 (m, 1H), 4.52 (m, 1H), 4.46 (m, 1H), 4.03 (m, 1H), 3.76 (m, 2H). ^{13}C NMR (75.4 MHz, CDCl_3): δ 164.59, 150.47, 142.19, 135.02 (d, $J = 6.0$ Hz), 128.74, 128.58, 128.04 (d, $J = 2.8$ Hz), 100.72, 86.02, 83.21, 81.44 (d, $J = 4.1$ Hz), 73.98 (d, $J = 4.0$ Hz), 70.00 (dd, $J = 5.4, 4.8$ Hz), 60.98. ^{31}P NMR (121.4 MHz, CDCl_3 , ^1H decoupled): δ –1.60. ESI–MS ($\text{M} + \text{Na}$): 527.1180 (observed), 527.1195 (calculated).

Arabinouridine 3'-phosphate (araUMP, 5.10). The preparation of **5.10** was carried out in a manner similar to that used for the preparation of **5.5a**. The product was purified by reverse-phase HPLC with elution by the gradient: 0–10 min, 95% A, 5% B; 10–20 min, 95–50% A, 5–50% B; 20–25 min, 50–95% A, 50–5% B. Buffer A was H_2O containing TFA (0.1% v/v); Buffer B was CH_3CN containing TFA (0.1% v/v). The desired product eluted between 6 and 8 min, and the byproduct eluted at 21 min. The fractions were combined and evaporated under reduced pressure to yield **5.10** as a colorless solid (558 mg, 83.8%). ^1H NMR (300 MHz, D_2O): δ 7.67 (d, $J = 8.1$ Hz, 1H), 5.98 (d, $J = 4.2$ Hz, 1H), 5.67 (d, $J = 8.1$ Hz, 1H), 4.41–4.35 (m, 2H), 4.03 (m, 1H), 3.72

(m, 2H). ^{13}C NMR (125.7 MHz, D_2O): δ 167.14, 152.18, 144.08, 101.88, 86.73, 84.16 (d, $J = 4.9$ Hz), 80.51 (broad), 75.18 (d, $J = 4.9$ Hz), 61.55. ^{31}P NMR (121.4 MHz, D_2O , ^1H decoupled): δ -0.37. ESI-MS ($\text{M} - \text{H}$): 323.0272 (observed), 323.0280 (calculated).

pK_a Determination by ^{31}P NMR Spectroscopy. A 3'-nucleotide was dissolved in D_2O (1.0 mL) to make a 100 mM stock solution. An aliquot (100 μL) of the stock solution was added to 0.10 M buffer (900 μL), and the resulting solution was filtered. The buffers used were oxalic acid, citric acid, succinic acid, trisodium citrate, MES, MOPS, TRIS, CHES and CAPS, each adjusted to an appropriate pH with 2 M HCl or 2 M NaOH. Each filtered sample (900 μL) was placed in an NMR tube, and its ^{31}P NMR chemical shift was measured with a Bruker DMX-400 MHz (wide bore) spectrometer equipped with a quattro-nucleus probe or a Bruker DMX-500 MHz spectrometer equipped with a broadband probe, referenced to an external standard of H_3PO_4 , and ^1H -decoupled. The pH of each sample was measured with a Beckman $\Phi 40$ pH meter. Data were fitted to eq 5.1 with the program Deltagraph 4.0.

$$\delta = \frac{\delta_{\text{low}} + \delta_{\text{high}} \times 10^{(\text{pH} - \text{pK}_a)}}{1 + 10^{(\text{pH} - \text{pK}_a)}} \quad (5.1)$$

The reported values are the mean (\pm SE) of two determinations.

K_i Determination for araUMP, dUMP and dU^FMP. The value of K_i for each 3'-nucleotide was measured by determining its ability to inhibit the turnover of the substrate 6-FAM-dArU(dA)₂-6-TAMRA by RNase A (Kelemen et al., 1999).

Fluorescence emission intensity was measured at 515 nm, with excitation at 493 nm.

Each assay was carried out in 2.0 mL of 20 mM MES–NaOH buffer, pH 6.0, containing NaCl (50 mM), RNase A (wild-type, 0.5 pM; T45G, 12.5 pM), and 6-FAM–dArU(dA)₂–6-TAMRA (0.06 μM). The value of $\Delta F/\Delta t$ was measured for 3 min after the addition of RNase A. An aliquot (0.5 μL) of 2 mM inhibitor dissolved in water was added and $\Delta F/\Delta t$ was measured for 3 min in the presence of the inhibitor. Additional aliquots of inhibitor were added at 3-min intervals, doubling the volume of the aliquot with each addition until 8 μL was added, then 4 μL of a 10 mM solution was added, and subsequent additions doubled in volume until 32 μL were added, nine additions in all. In each assay, $\leq 15\%$ of the substrate was cleaved. The loss of fluorescence intensity due to dilution was corrected for by dividing each data point by the corresponding point from an assay in which buffer instead of inhibitor was added to the enzyme/substrate mixture. The values of K_i were determined by fitting the data to eq 5.2 using the program Deltagraph 4.0.

$$\Delta F/\Delta t = (\Delta F/\Delta t)_0 \left(\frac{K_i}{K_i + [I]} \right) \quad (5.2)$$

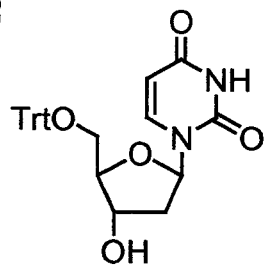
In eq 5.2, $(\Delta F/\Delta t)_0$ is the ribonucleolytic activity prior to inhibitor addition.

Fluorescence Data Analysis. During the course of K_i assays the fluorescence intensity was found to be quenched by the inhibitors at high inhibitor concentration. To correct for this quenching, the following assay was conducted. To 2.0 mL of 20 mM MES–NaOH buffer, pH 6.0, containing NaCl (50 mM) was added 1 μL (60 μM) of the substrate 6-FAM–dArU(dA)₂–6-TAMRA, followed 3 min later by 2 μL of a concentrated solution of wild-type RNase A (1.5 mM). At 3-min intervals thereafter, aliquots of

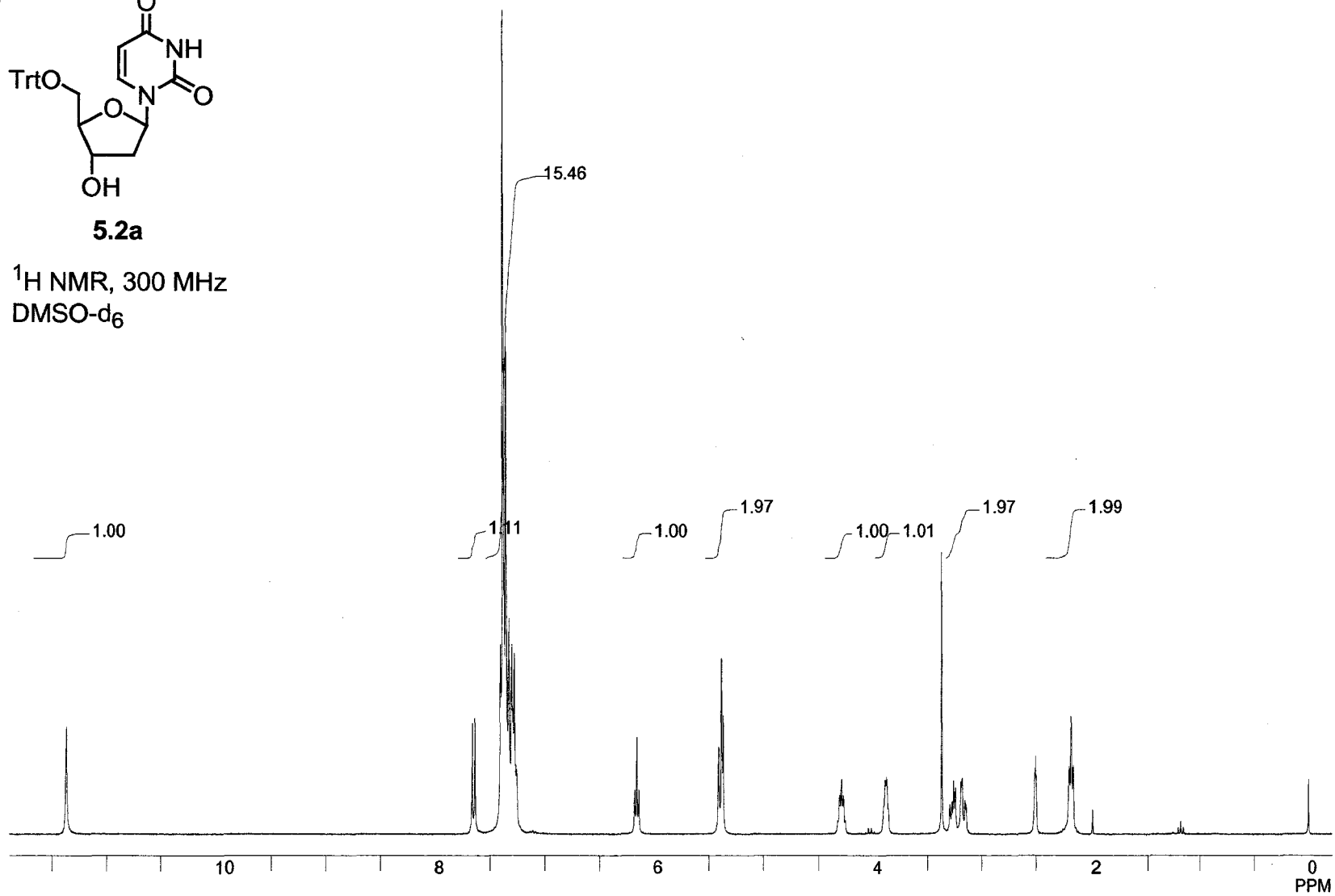
araUMP (1.72 mM) were added, beginning with 5 μ L and doubling in volume until 40 μ L had been added, 4 additions in all, and the fluorescence intensity was measured. In a separate assay, aliquots (5, 10, 20, and 40 μ L) of 25.7 mM araUMP were added, and the fluorescence intensity was measured. The two data sets were corrected for loss of fluorescence intensity due to dilution, then combined and fitted to eq 5.3 using the program Deltagraph 4.0. A quenching correction factor for each point was calculated using eq 5.3 (where F_{∞} is the value of the final fluorescence intensity measurement and $k = -30.77$) and the inhibitor concentration in the cuvette. Each value of $\Delta F/\Delta t$ was divided by the correction factor to give the corrected value. The correction factor was the same for all the inhibitors, assuming that the fluorescence quenching arises from the uracil moiety of the inhibitors.

$$y = (1 - F_{\infty})e^{kx} + F_{\infty} \quad (5.3)$$

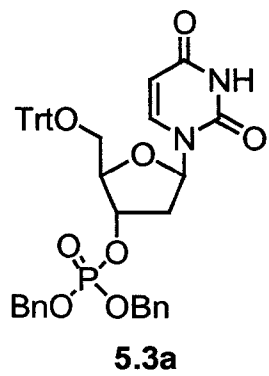
134

**5.2a**

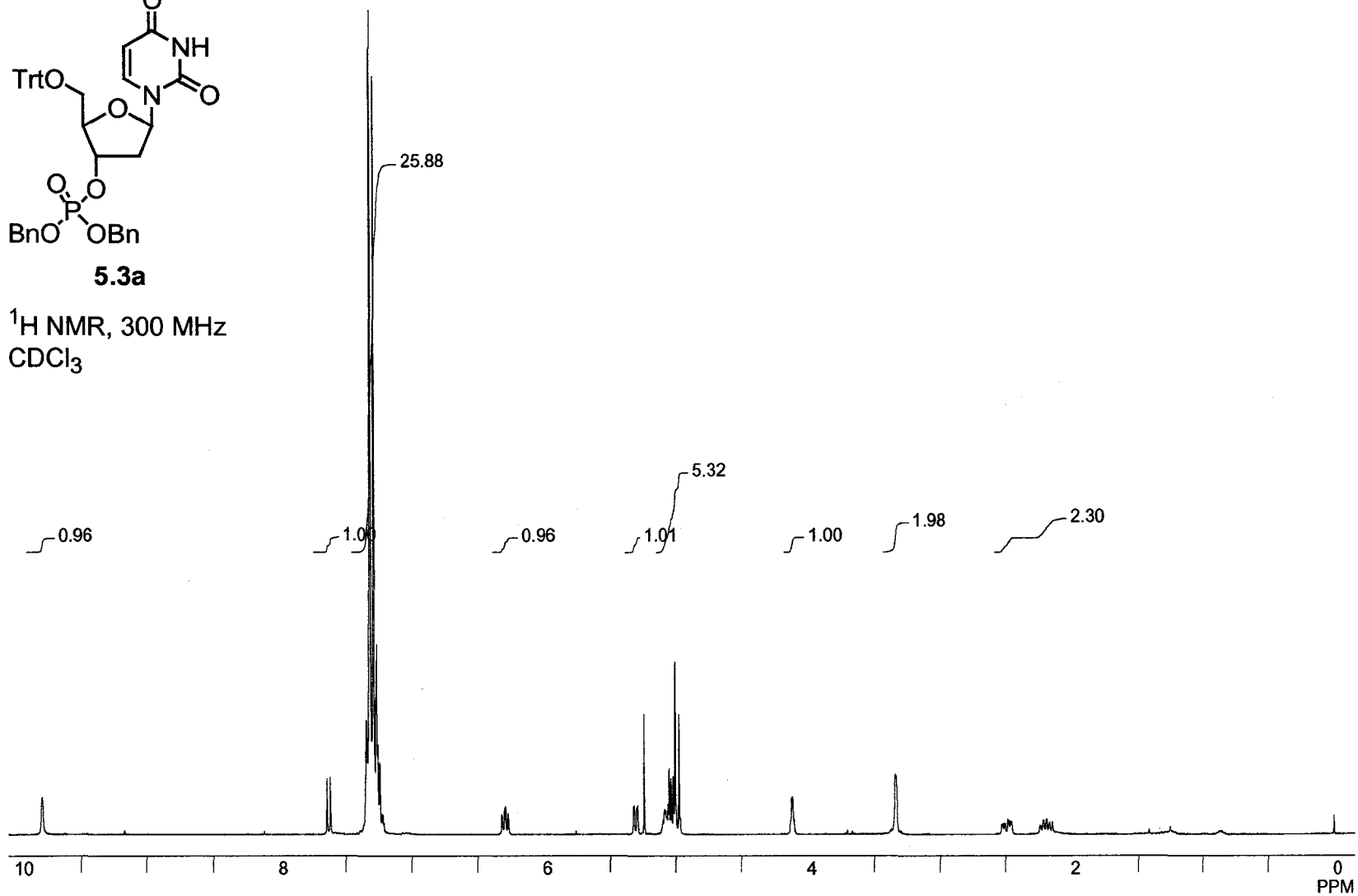
^1H NMR, 300 MHz
DMSO- d_6



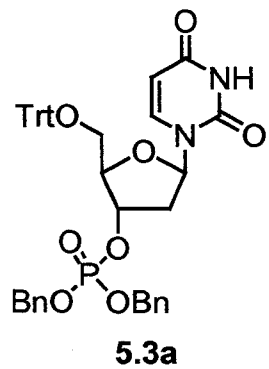
135



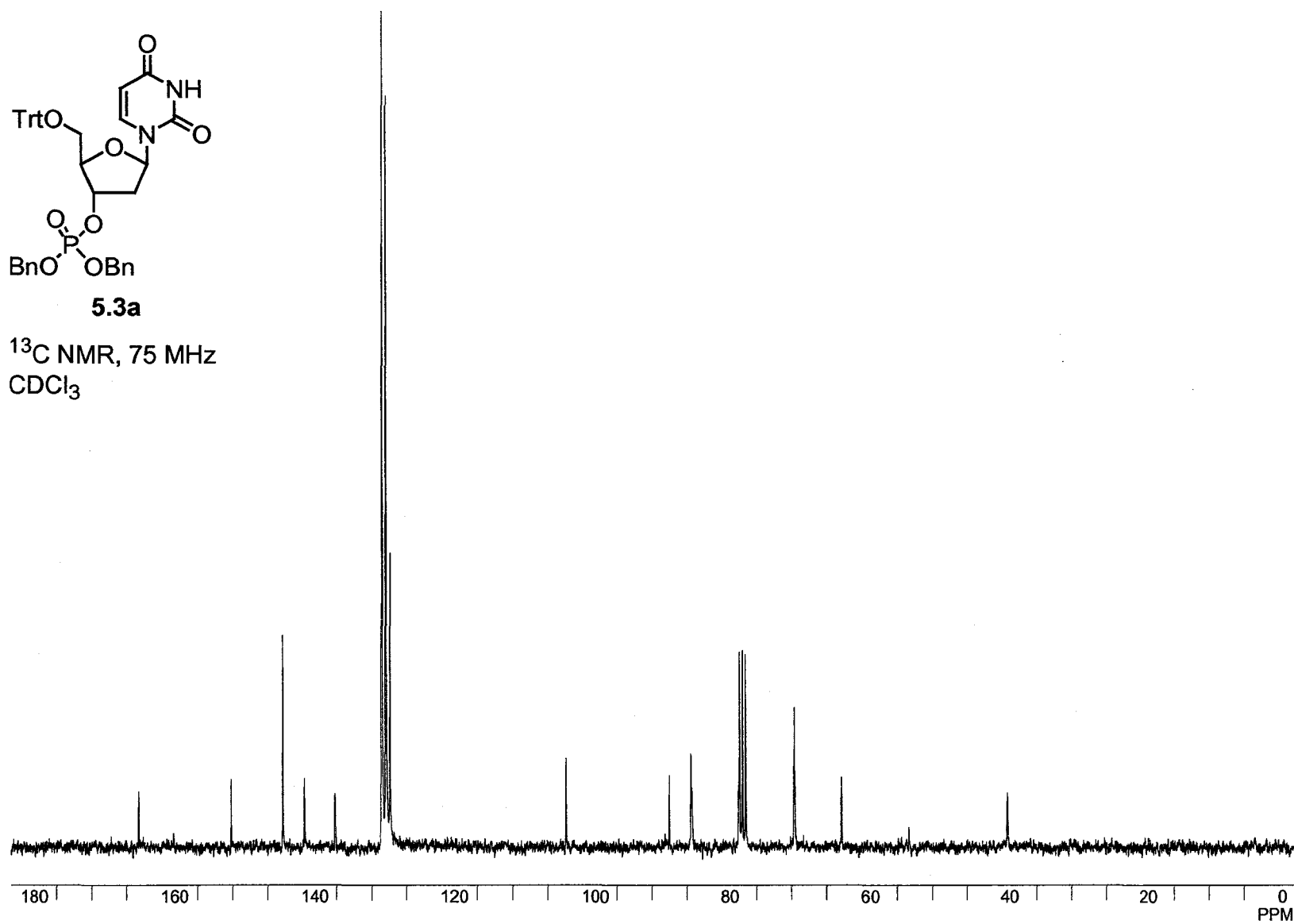
^1H NMR, 300 MHz
 CDCl_3

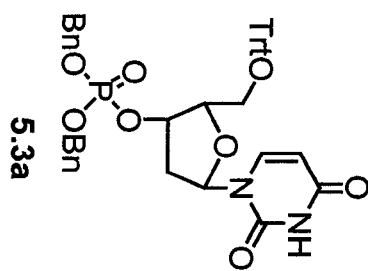


136

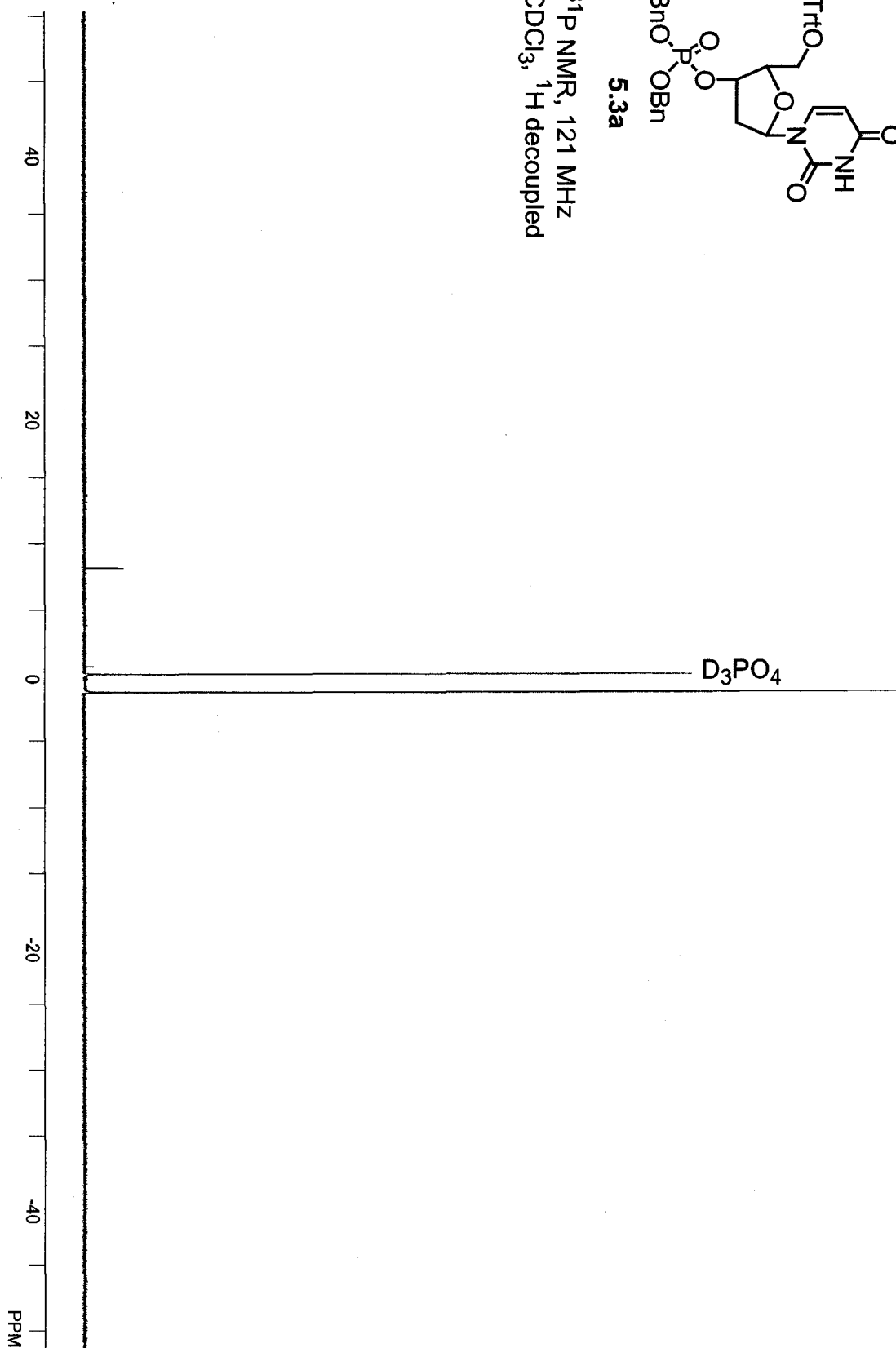


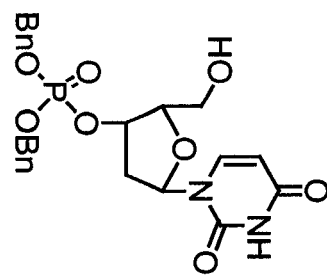
^{13}C NMR, 75 MHz
 CDCl_3



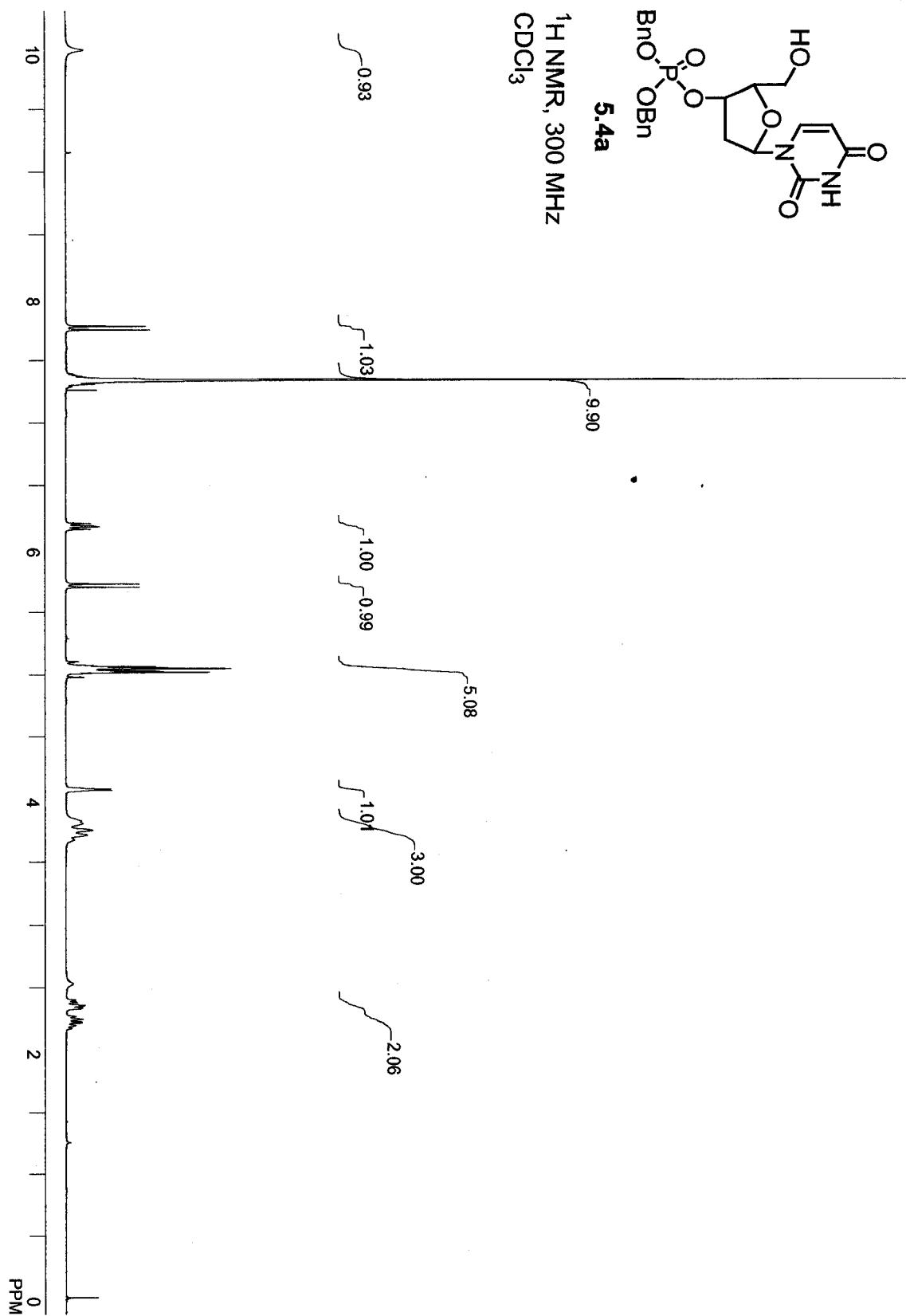


^{31}P NMR, 121 MHz
 CDCl_3 , ^1H decoupled

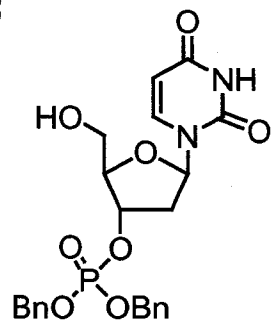


**5.4a**

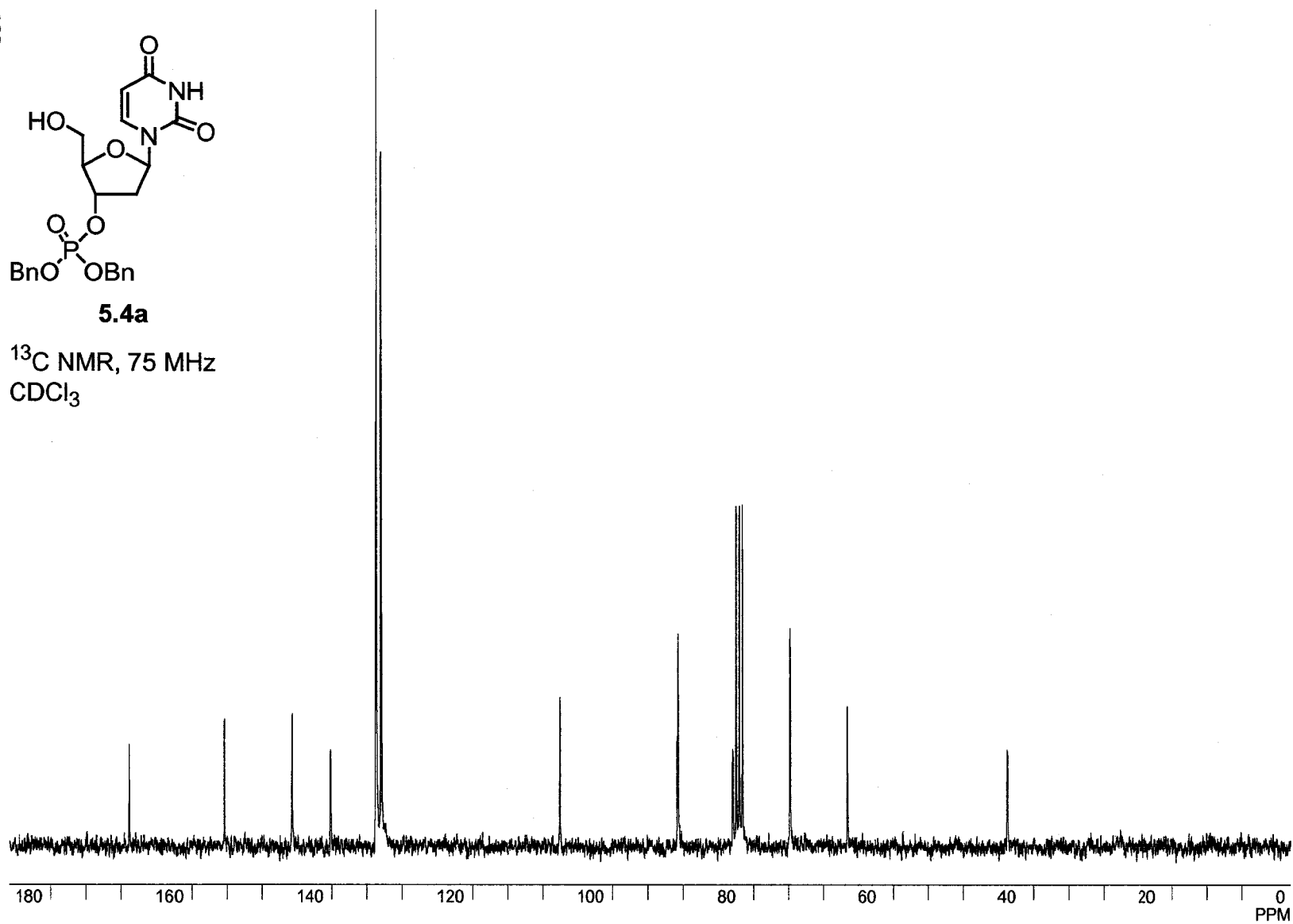
^1H NMR, 300 MHz
 CDCl_3

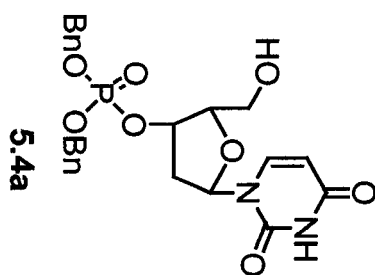


139

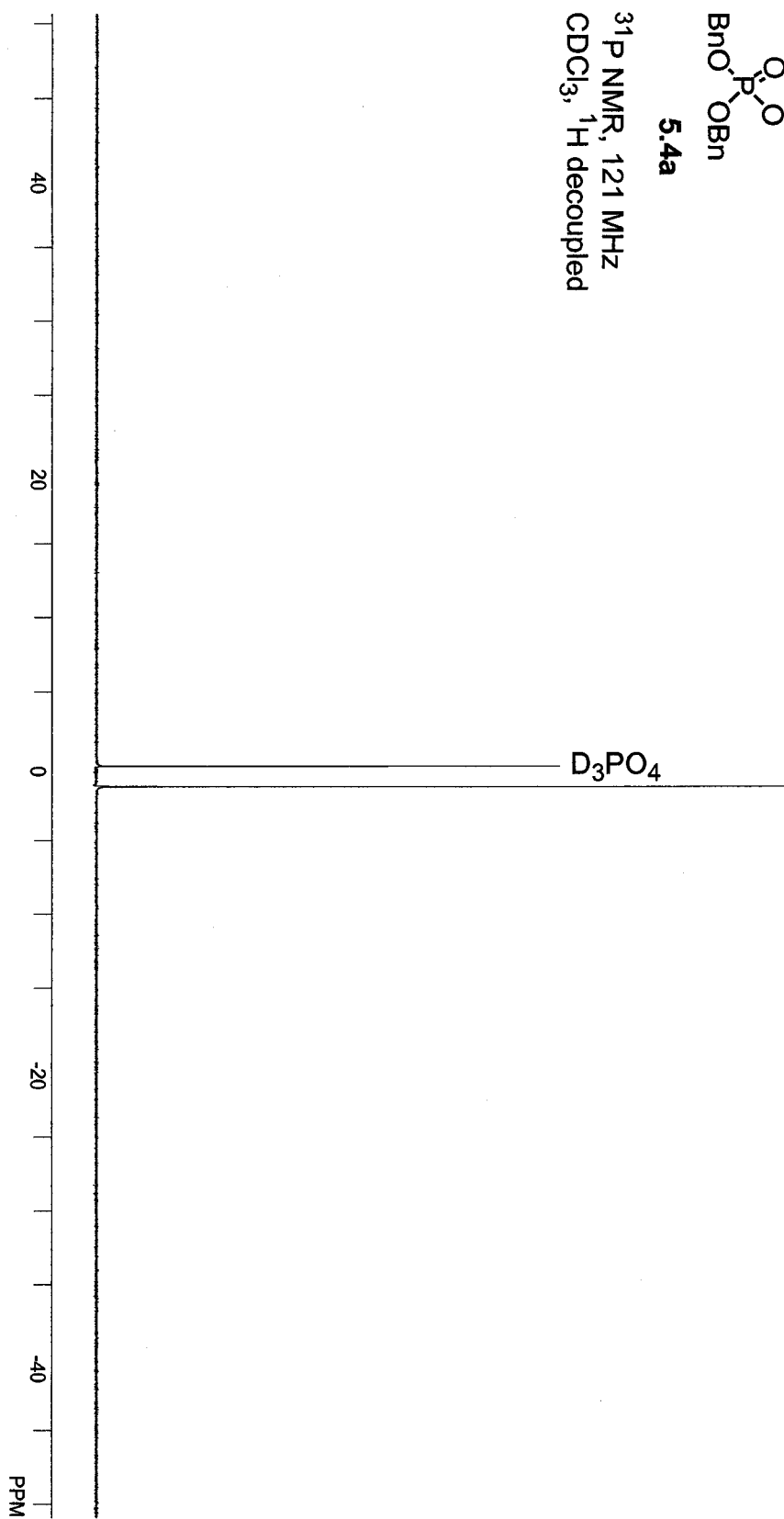
**5.4a**

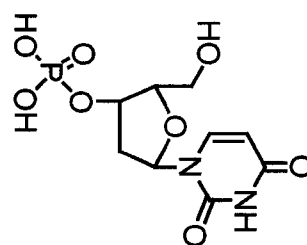
^{13}C NMR, 75 MHz
 CDCl_3



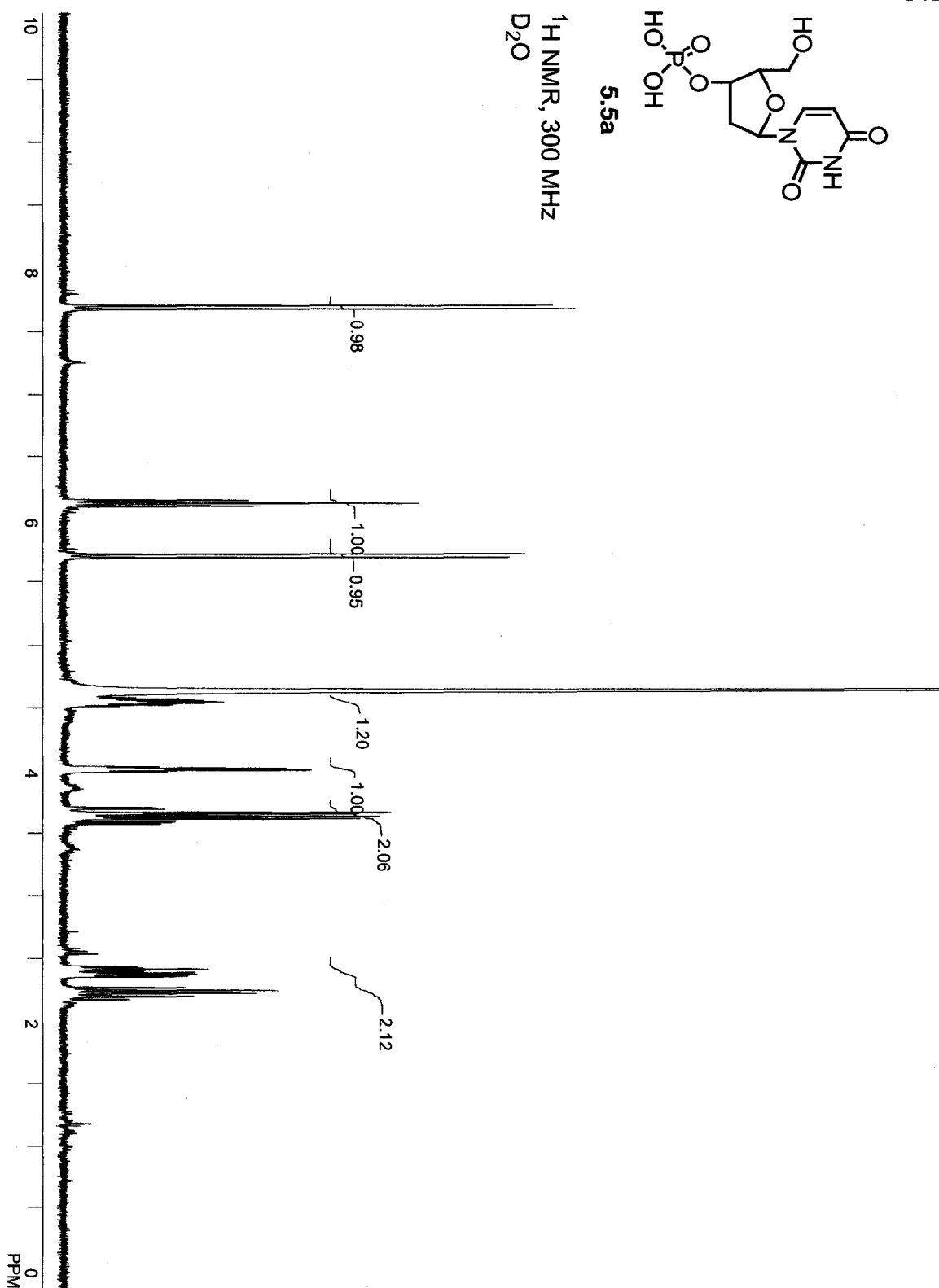


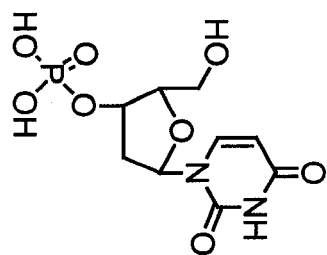
^{31}P NMR, 121 MHz
 CDCl_3 , ^1H decoupled



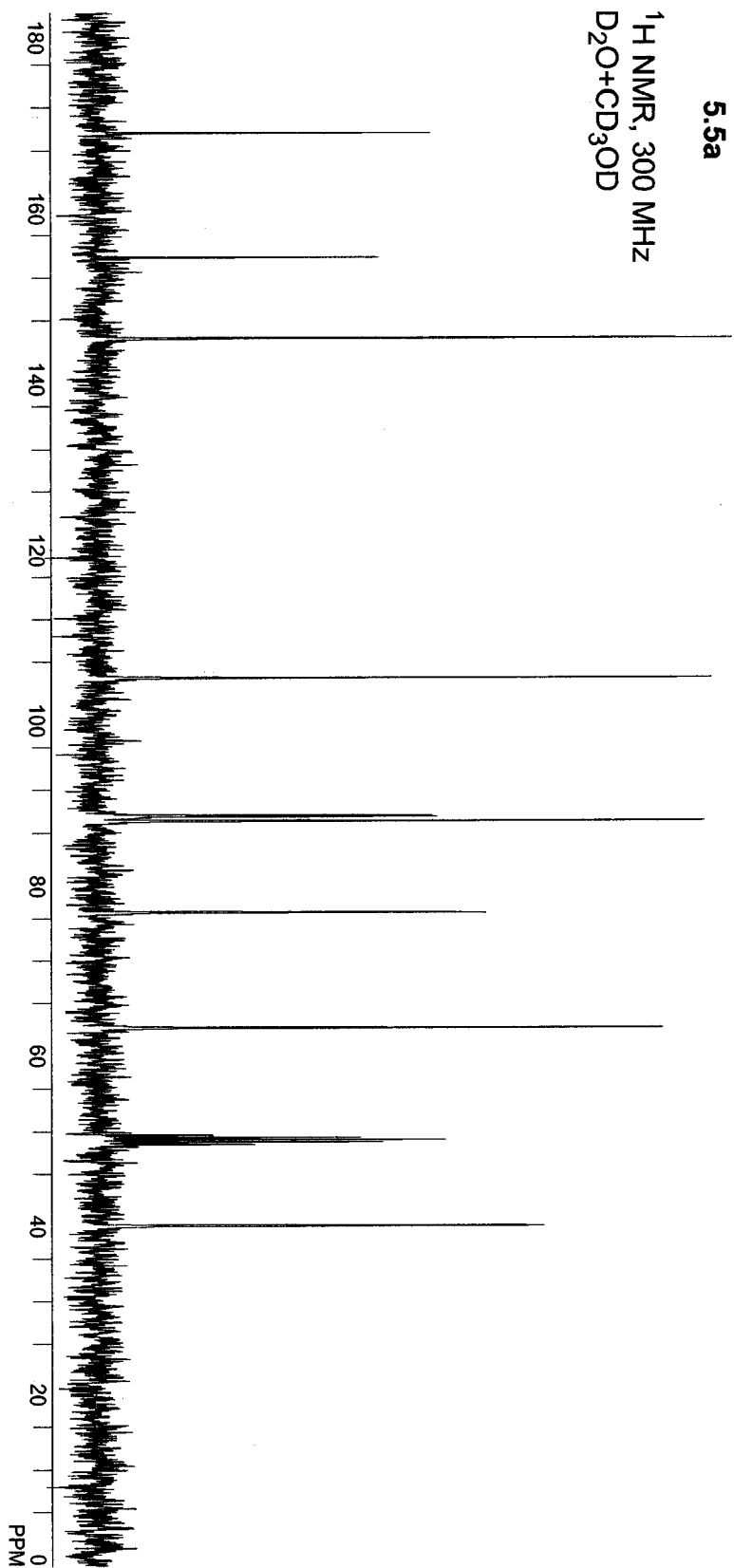
**5.5a**

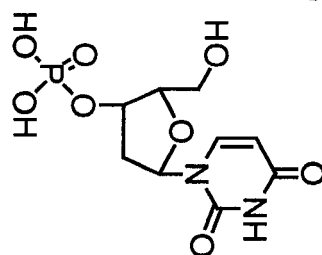
^1H NMR, 300 MHz
 D_2O



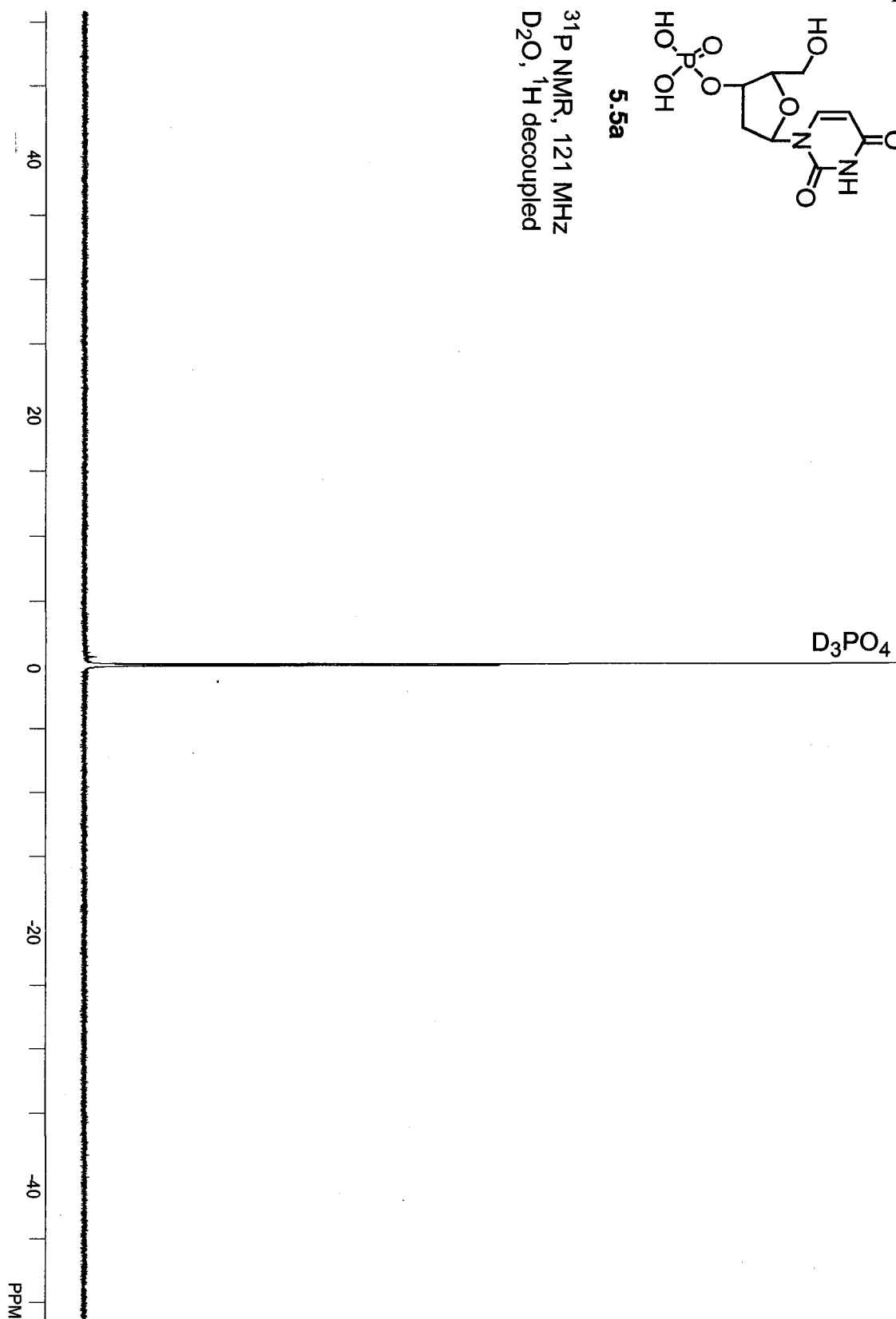
**5.5a**

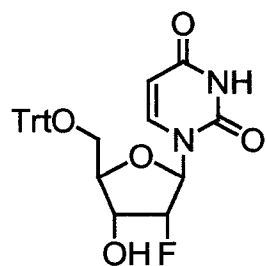
^1H NMR, 300 MHz
 $\text{D}_2\text{O} + \text{CD}_3\text{OD}$



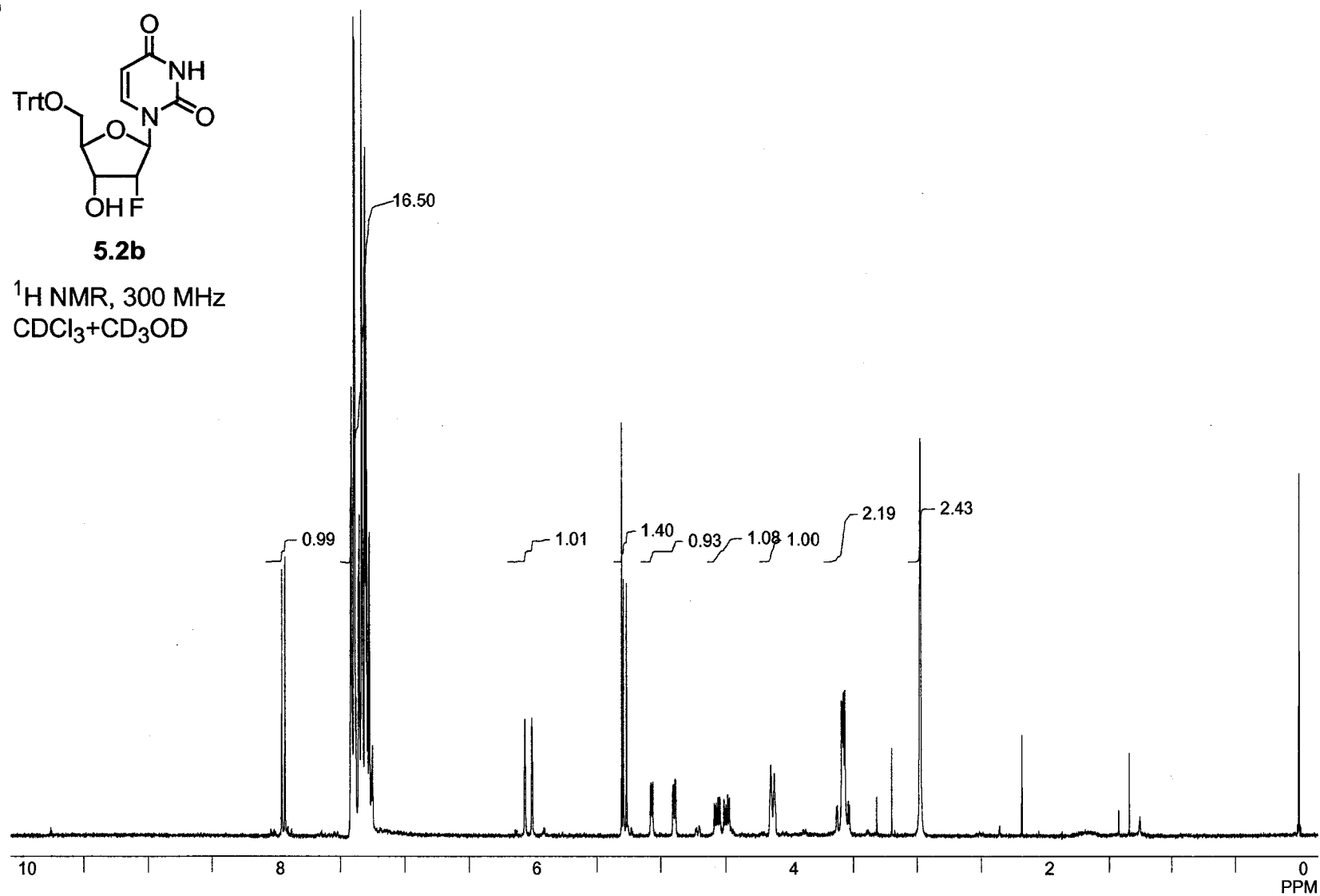
**5.5a**

^{31}P NMR, 121 MHz
 D_2O , ^1H decoupled

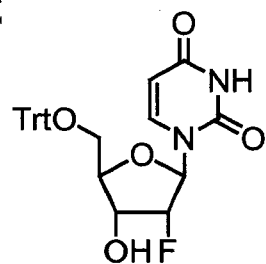


**5.2b**

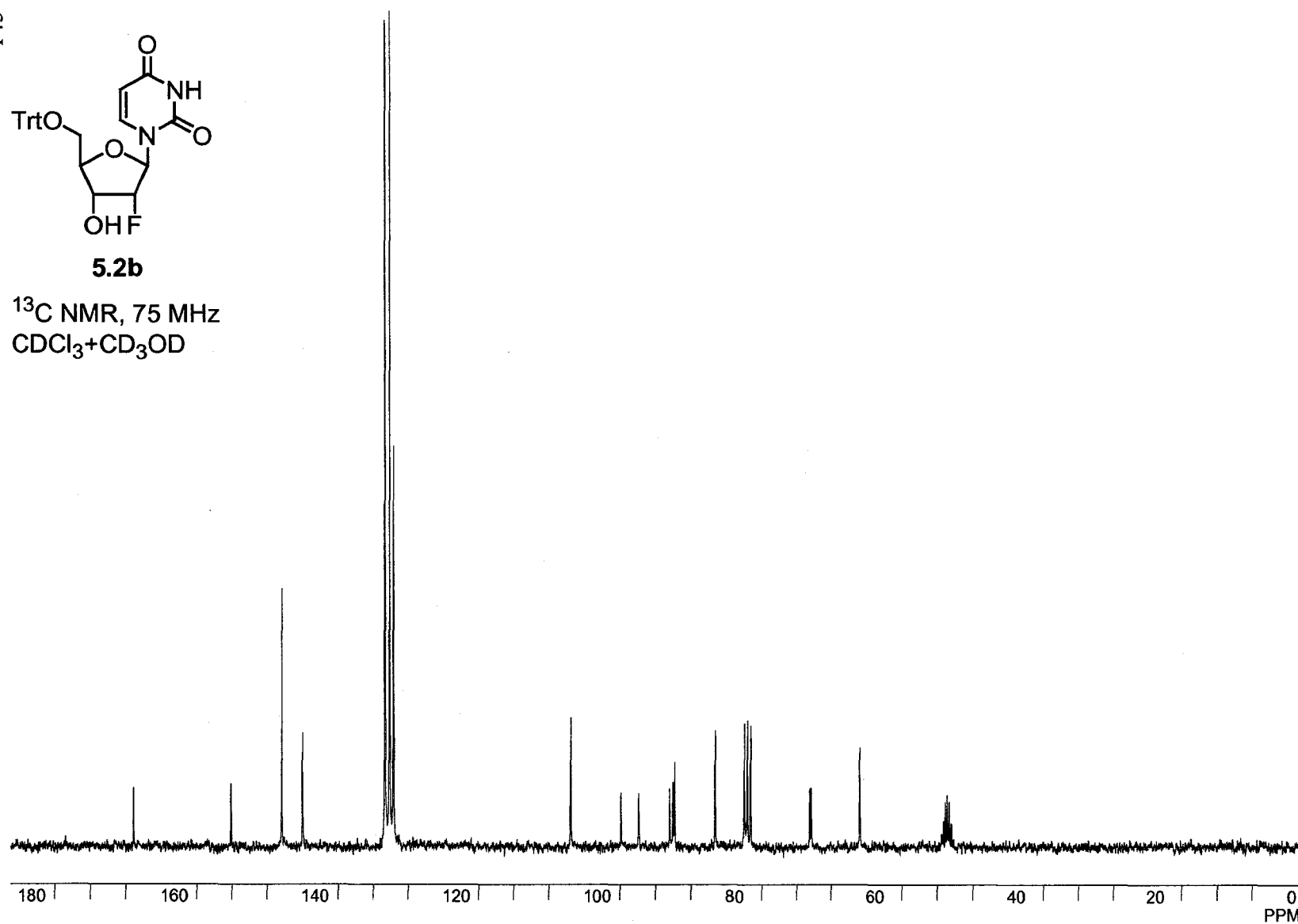
^1H NMR, 300 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$



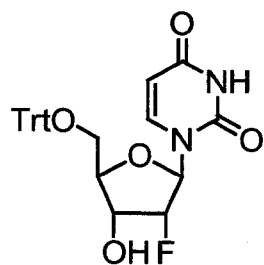
145

**5.2b**

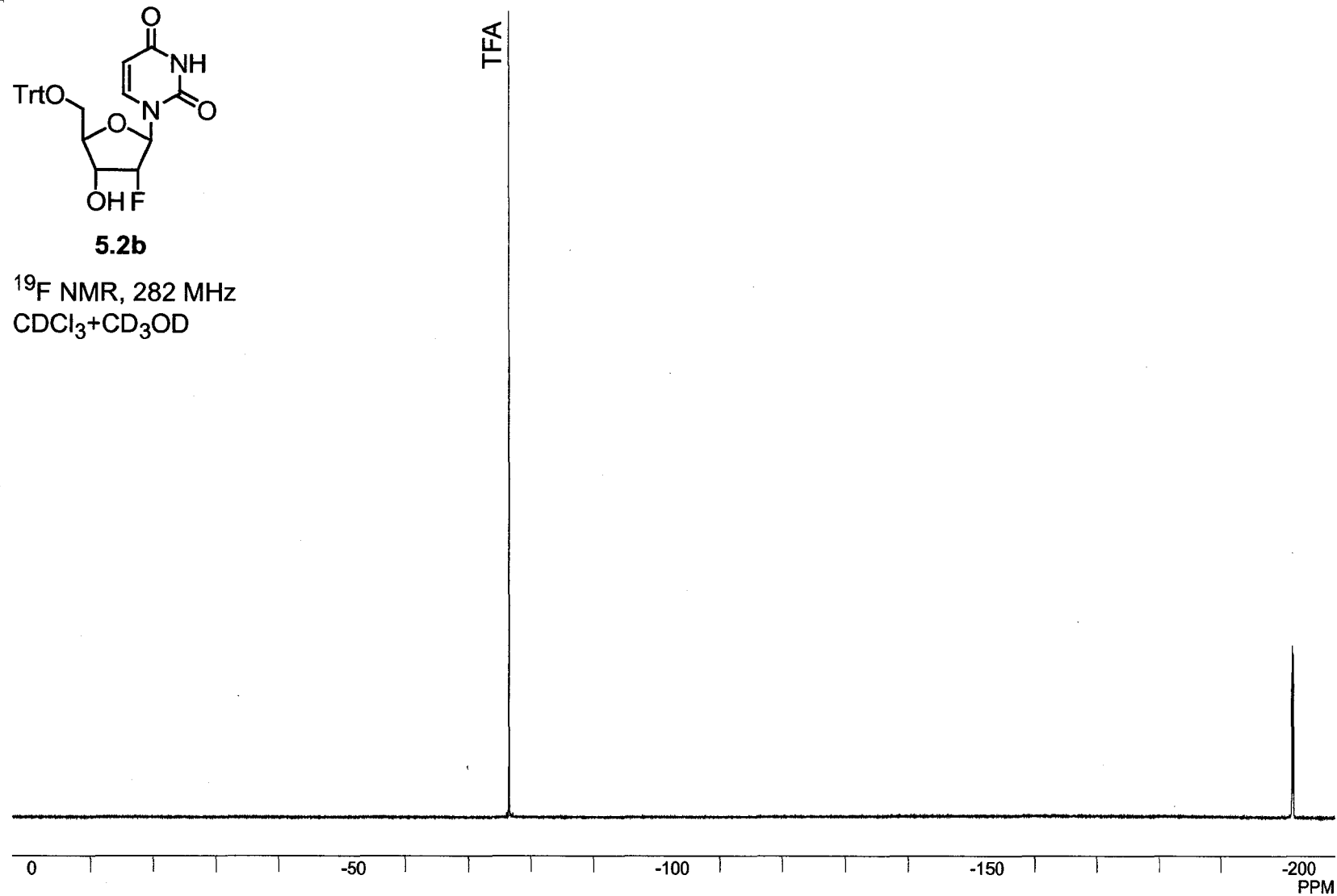
^{13}C NMR, 75 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$



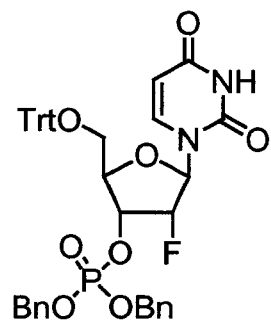
146

**5.2b**

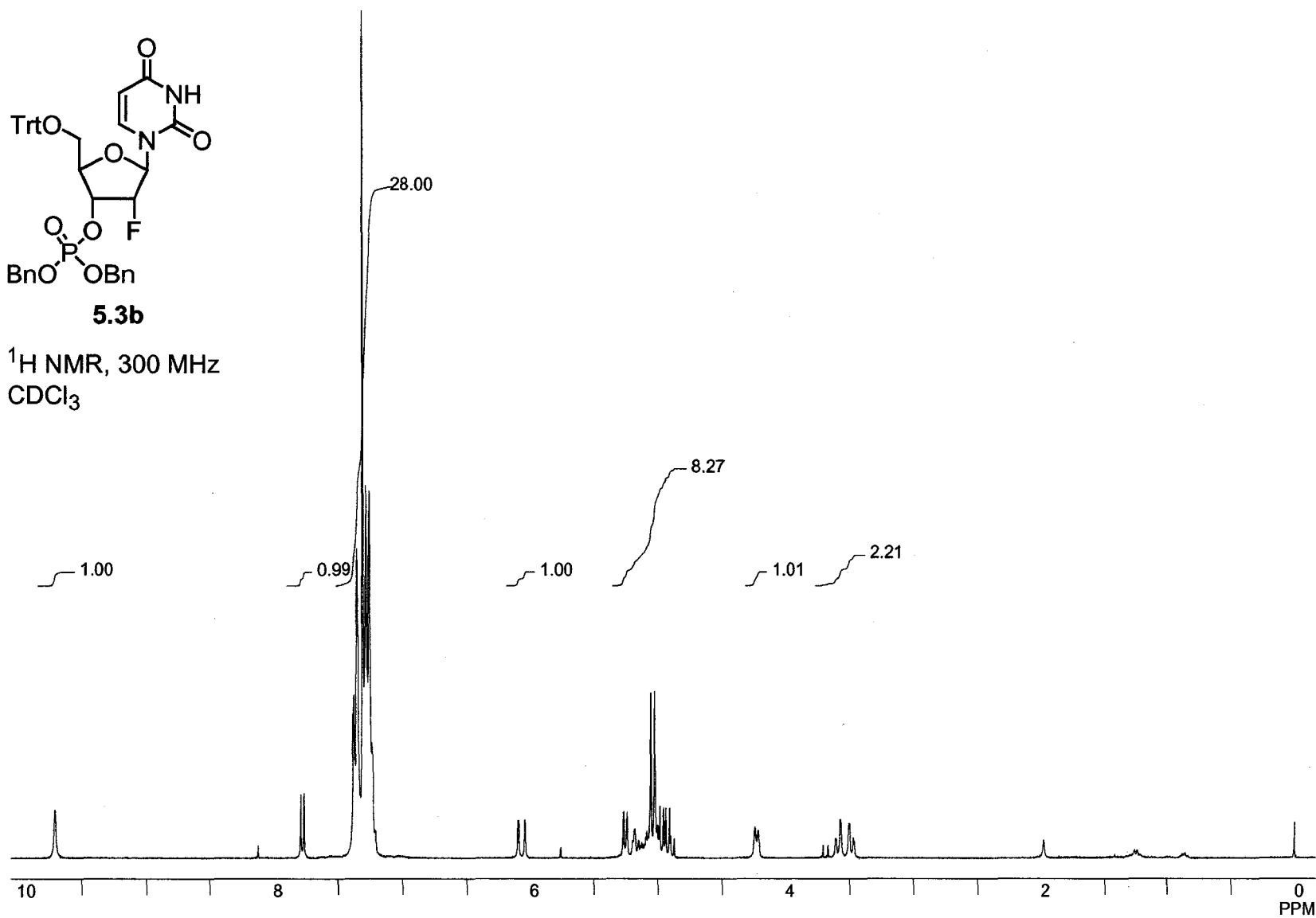
^{19}F NMR, 282 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$



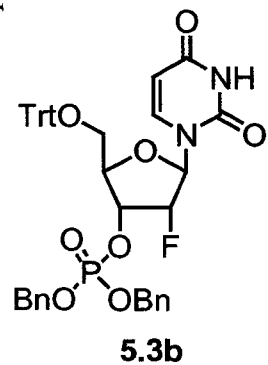
147

**5.3b**

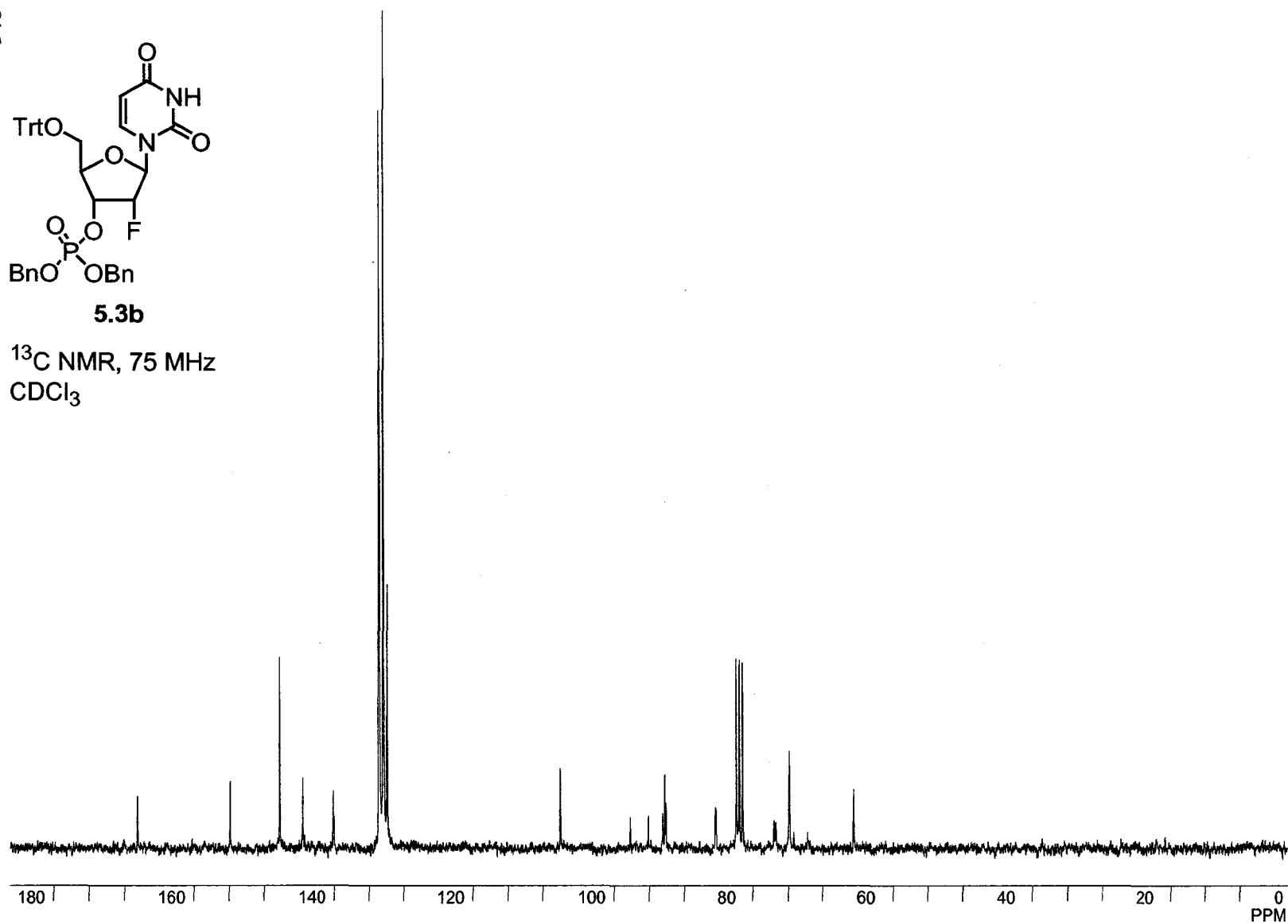
^1H NMR, 300 MHz
 CDCl_3

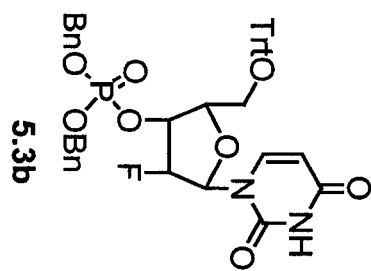


148

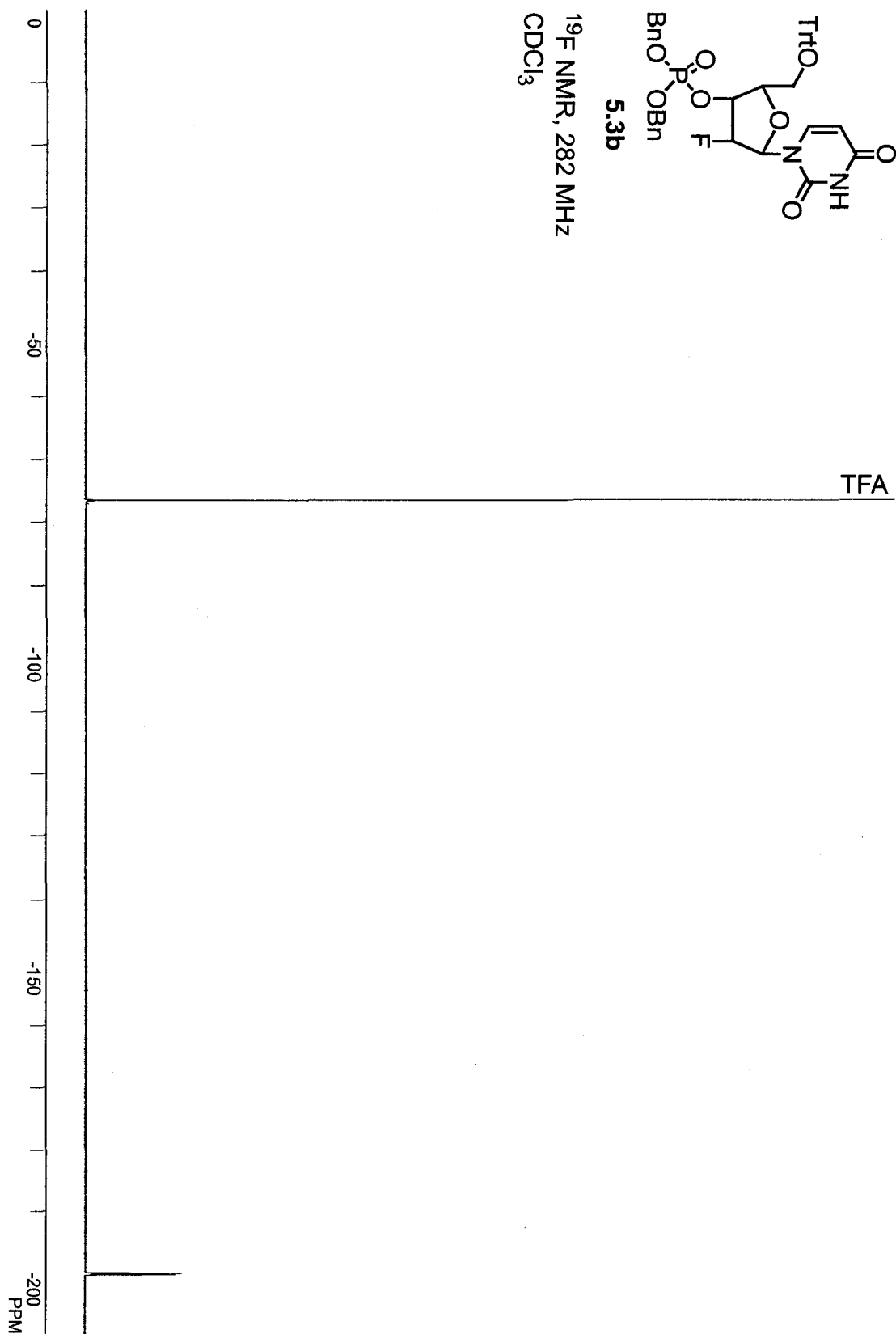


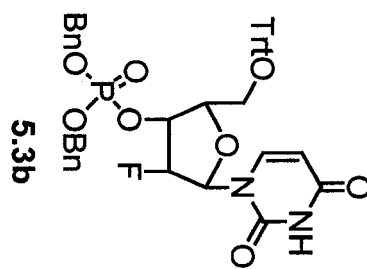
^{13}C NMR, 75 MHz
 CDCl_3



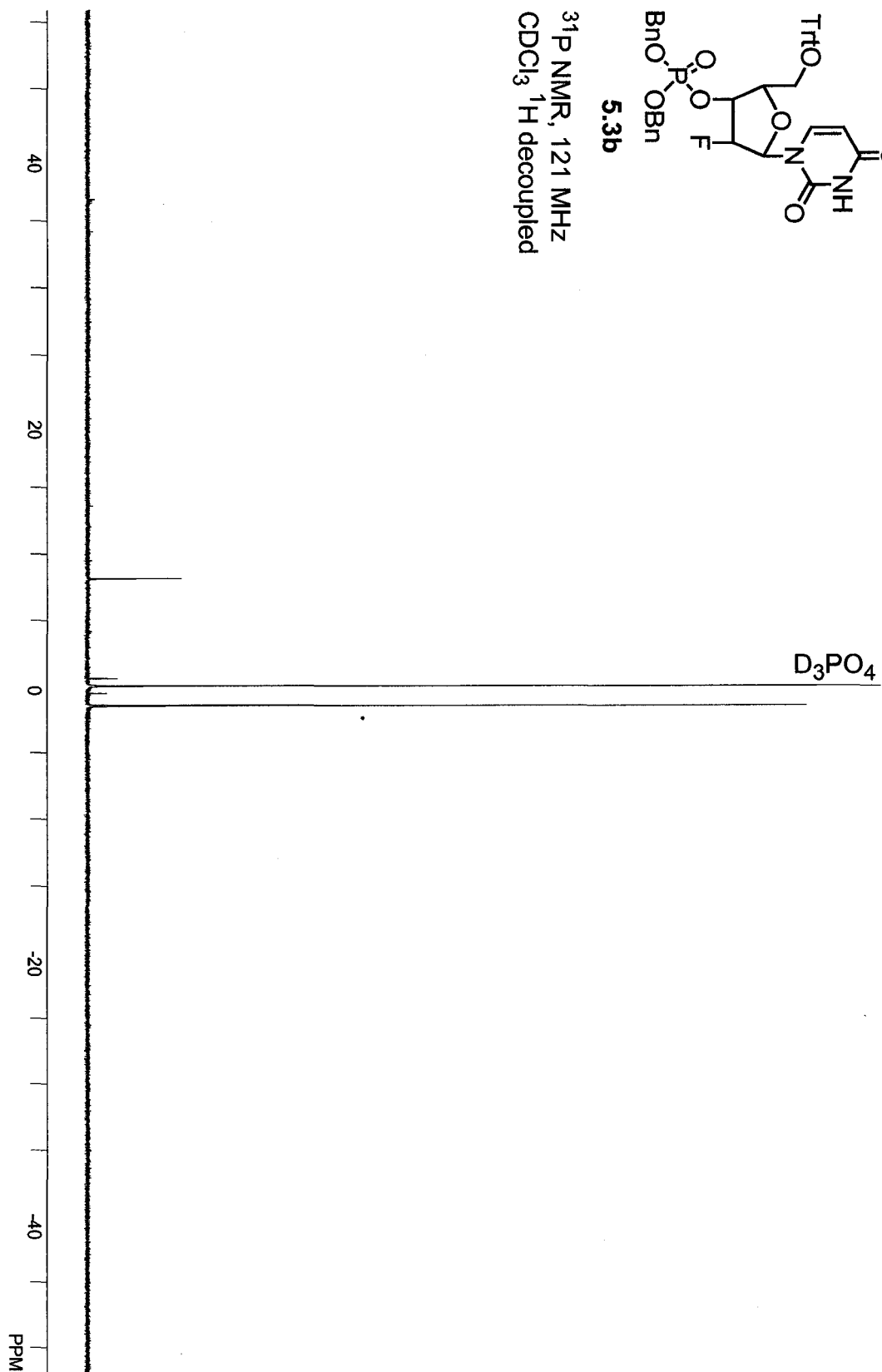


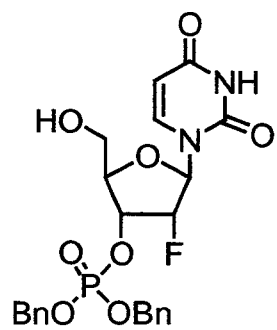
^{19}F NMR, 282 MHz
 CDCl_3



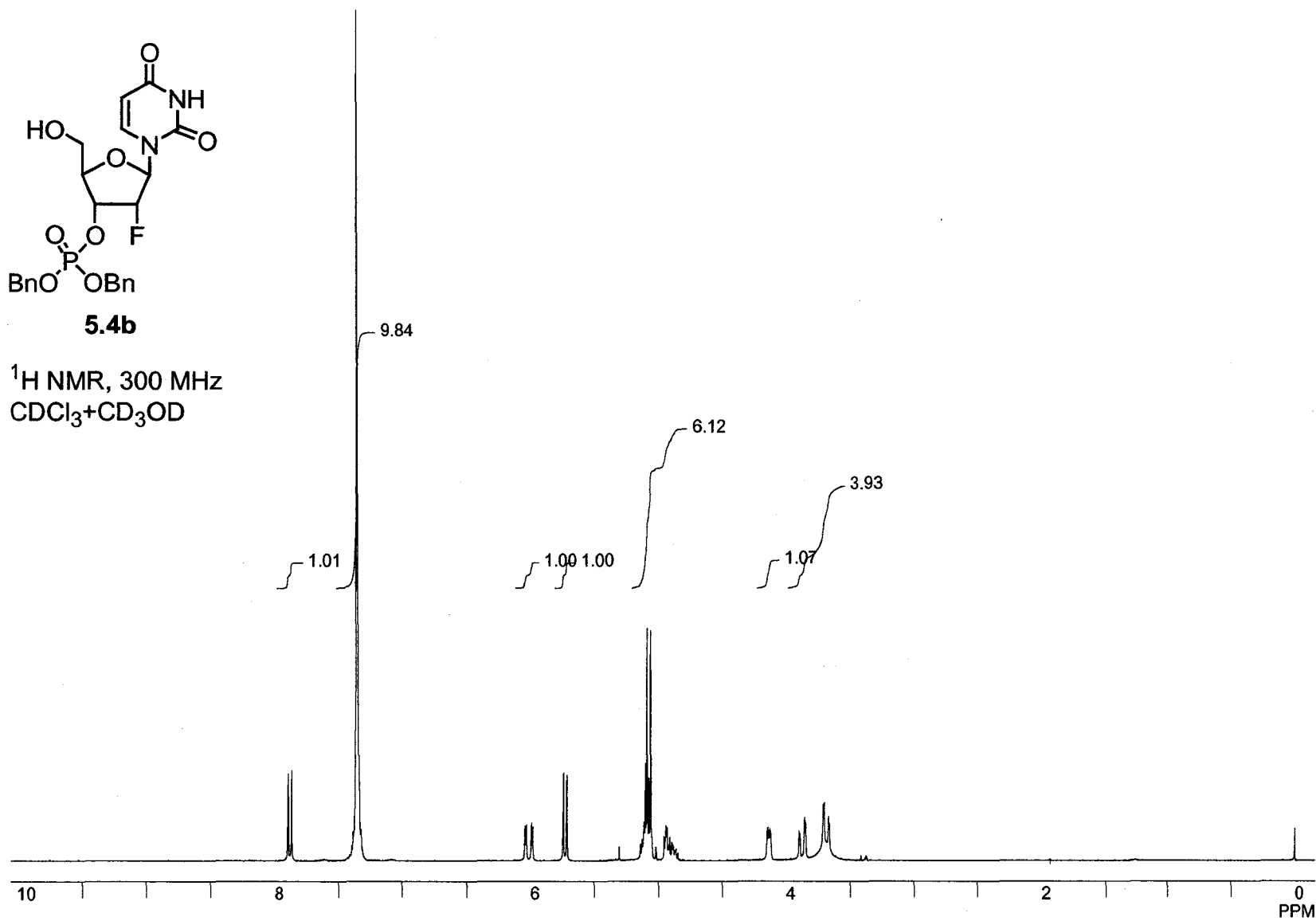


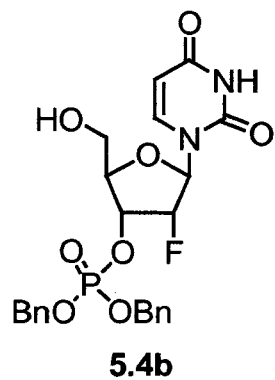
^{31}P NMR, 121 MHz
 CDCl_3 ^1H decoupled



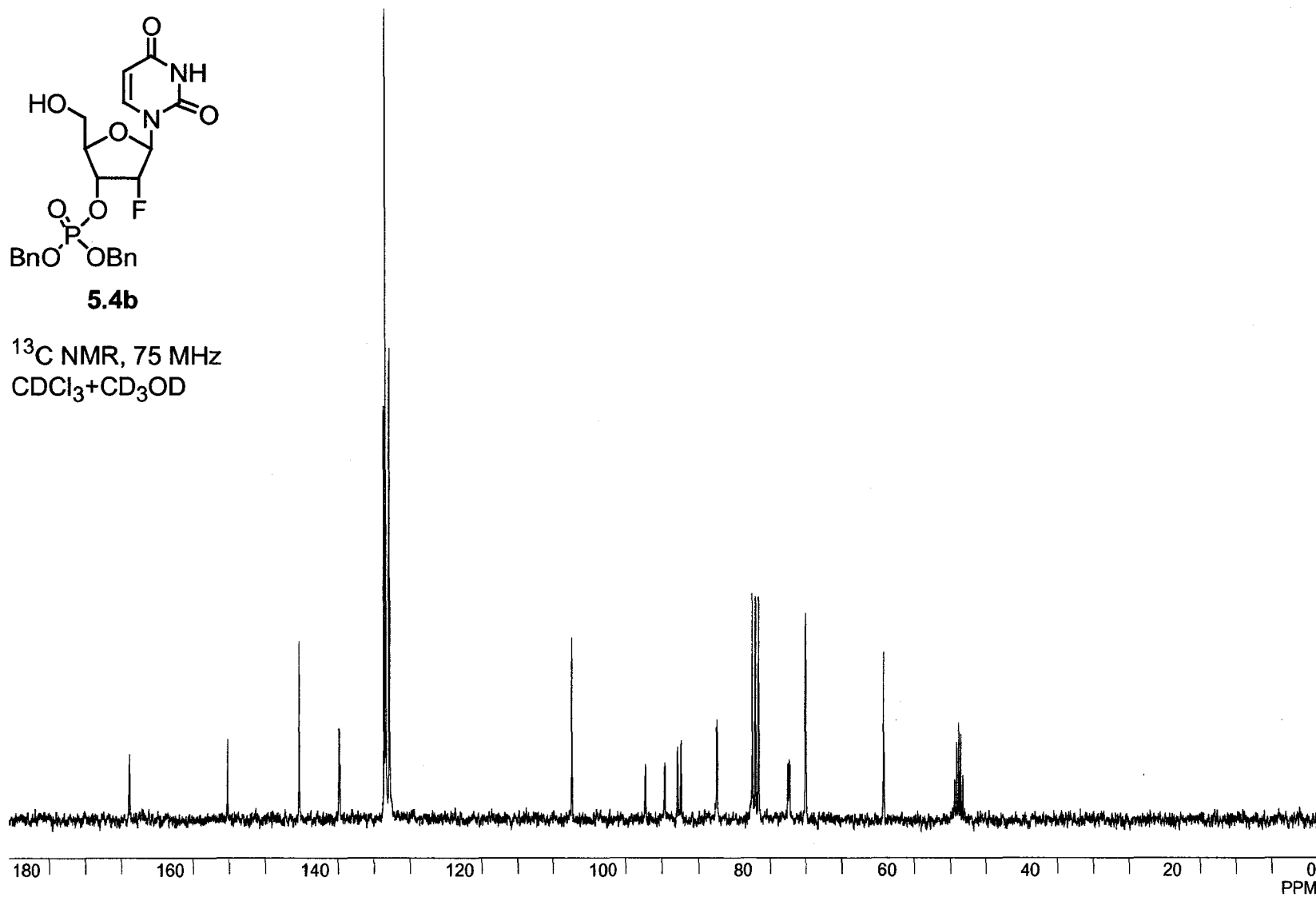
**5.4b**

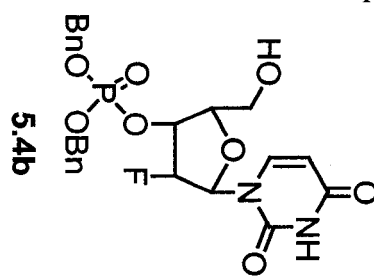
^1H NMR, 300 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$



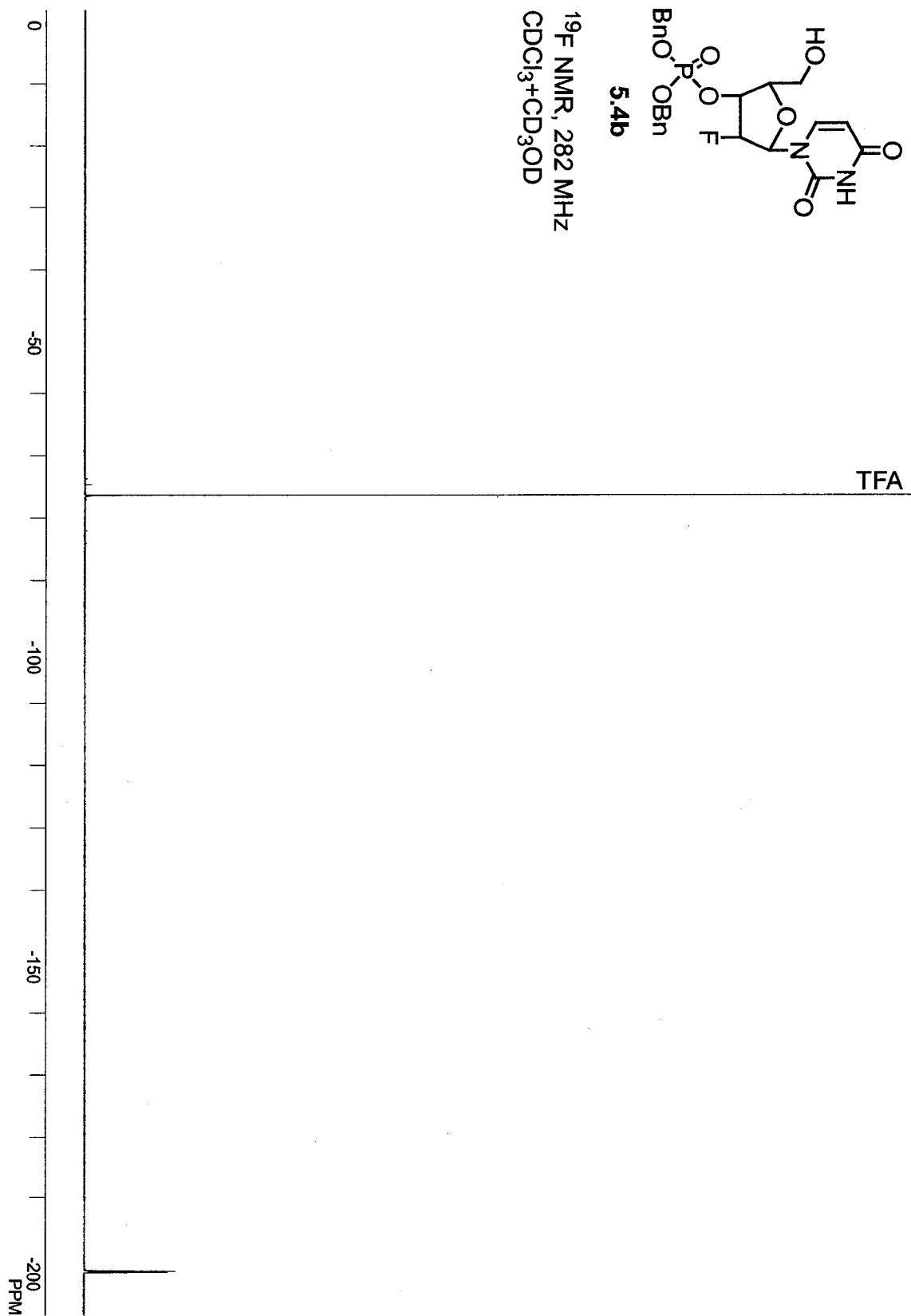


^{13}C NMR, 75 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$

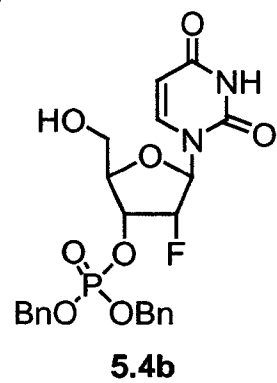




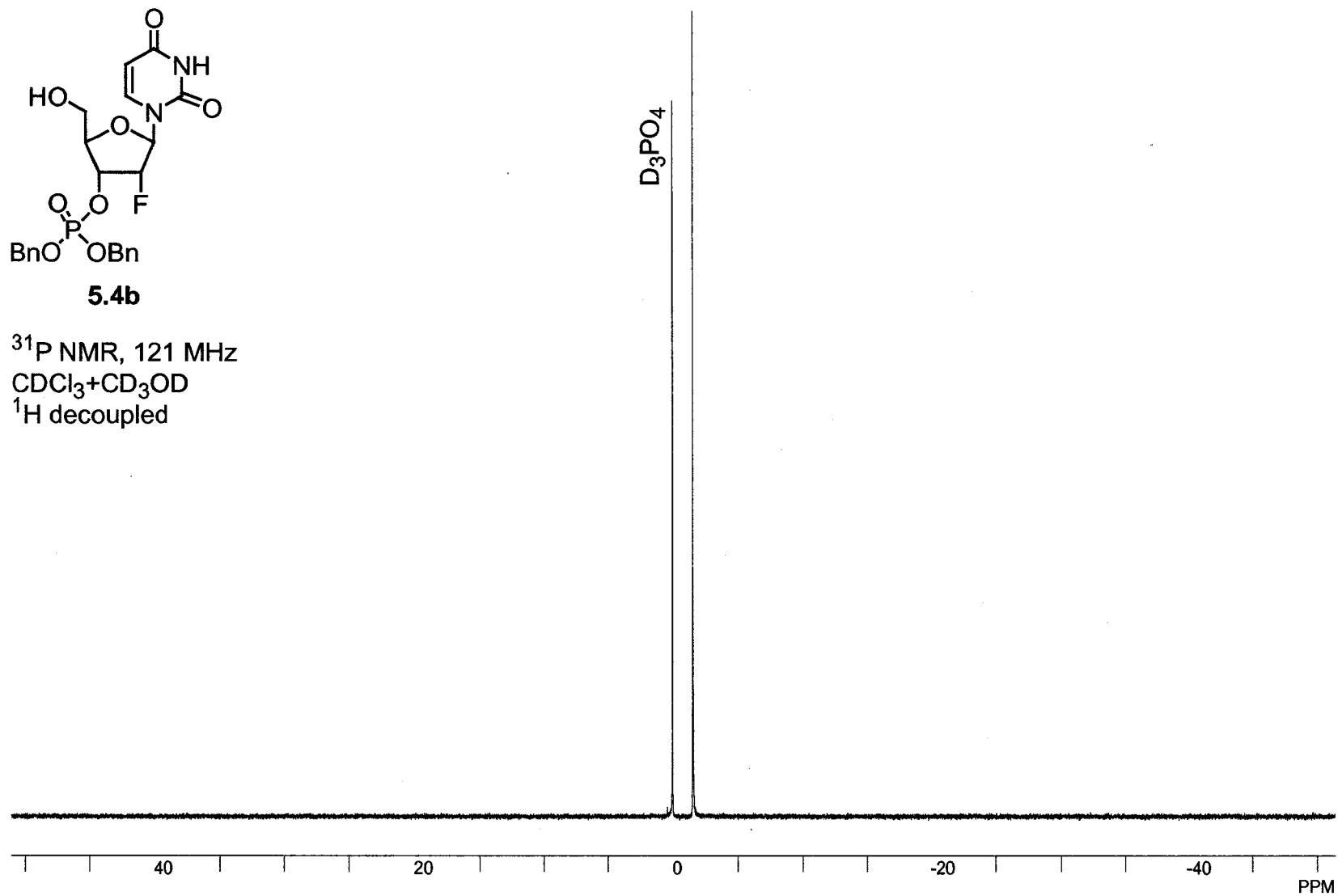
^{19}F NMR, 282 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$

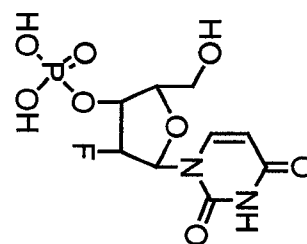


154

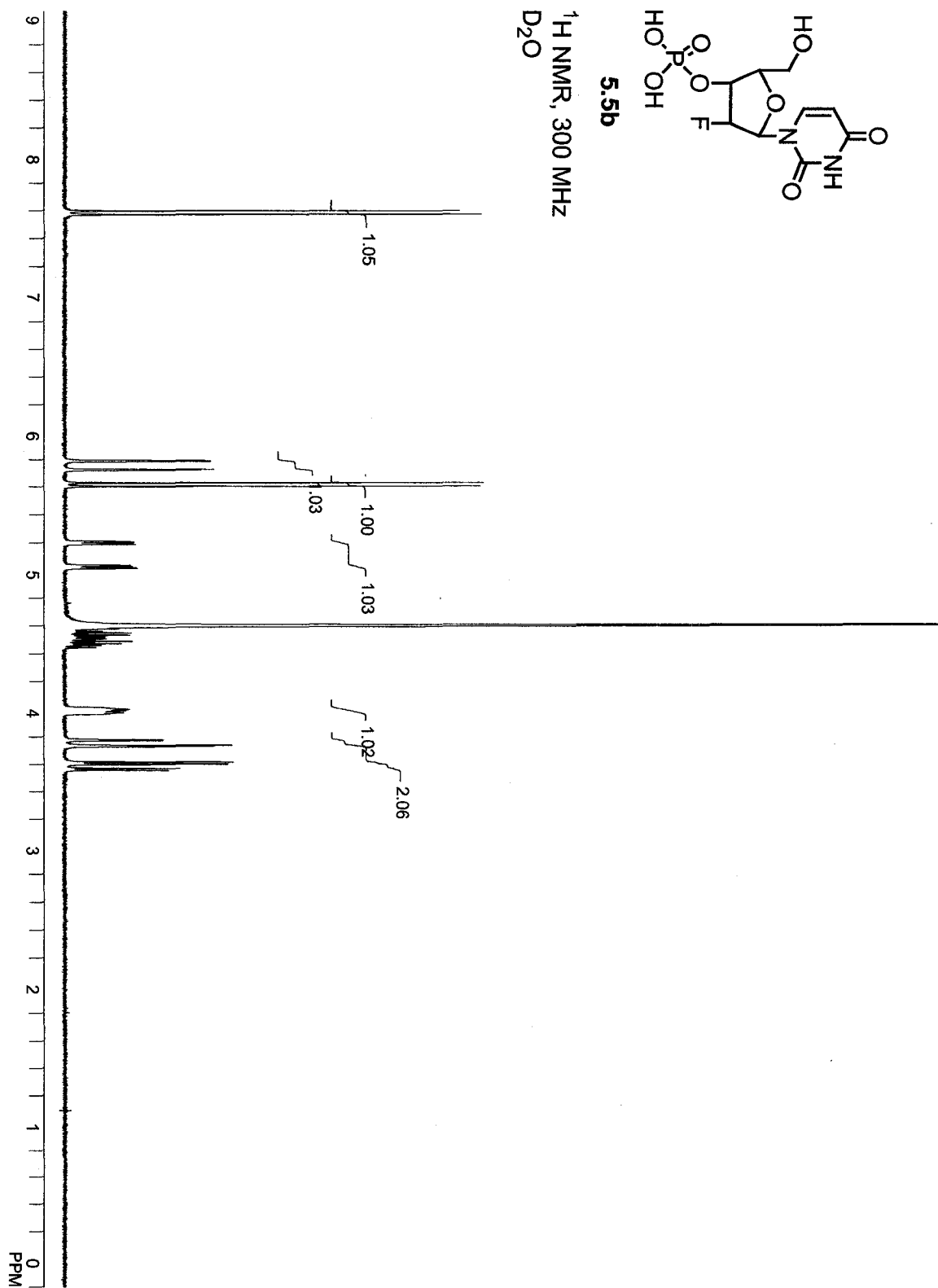


^{31}P NMR, 121 MHz
 $\text{CDCl}_3 + \text{CD}_3\text{OD}$
 ^1H decoupled

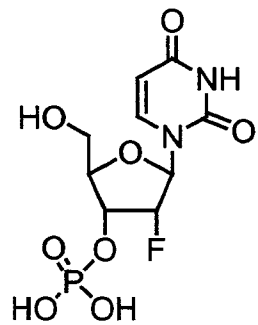


**5.5b**

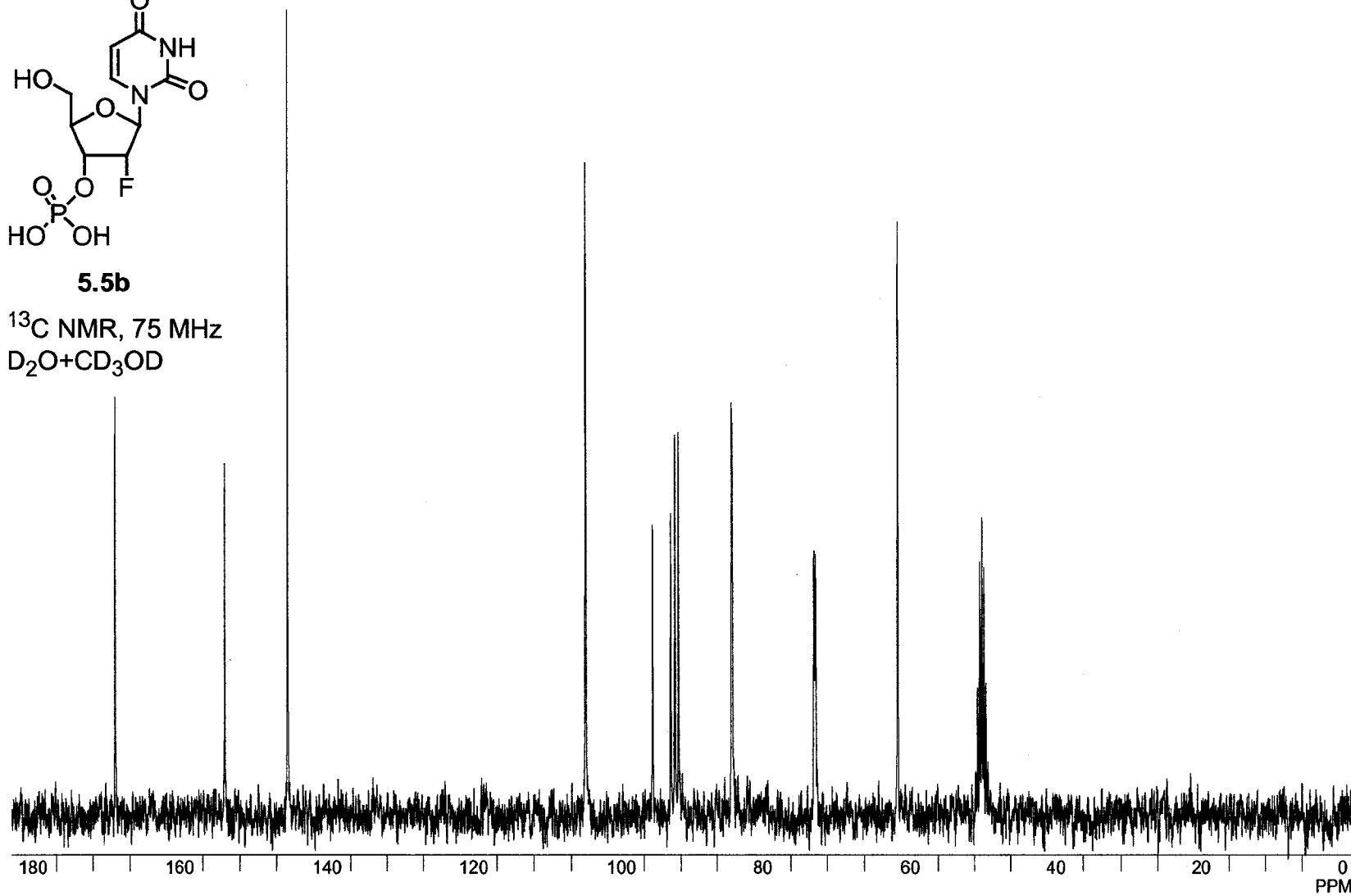
^1H NMR, 300 MHz
 D_2O

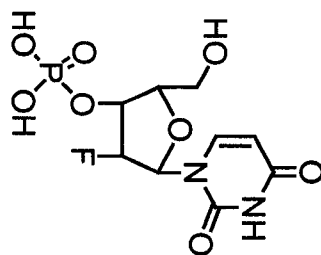


156

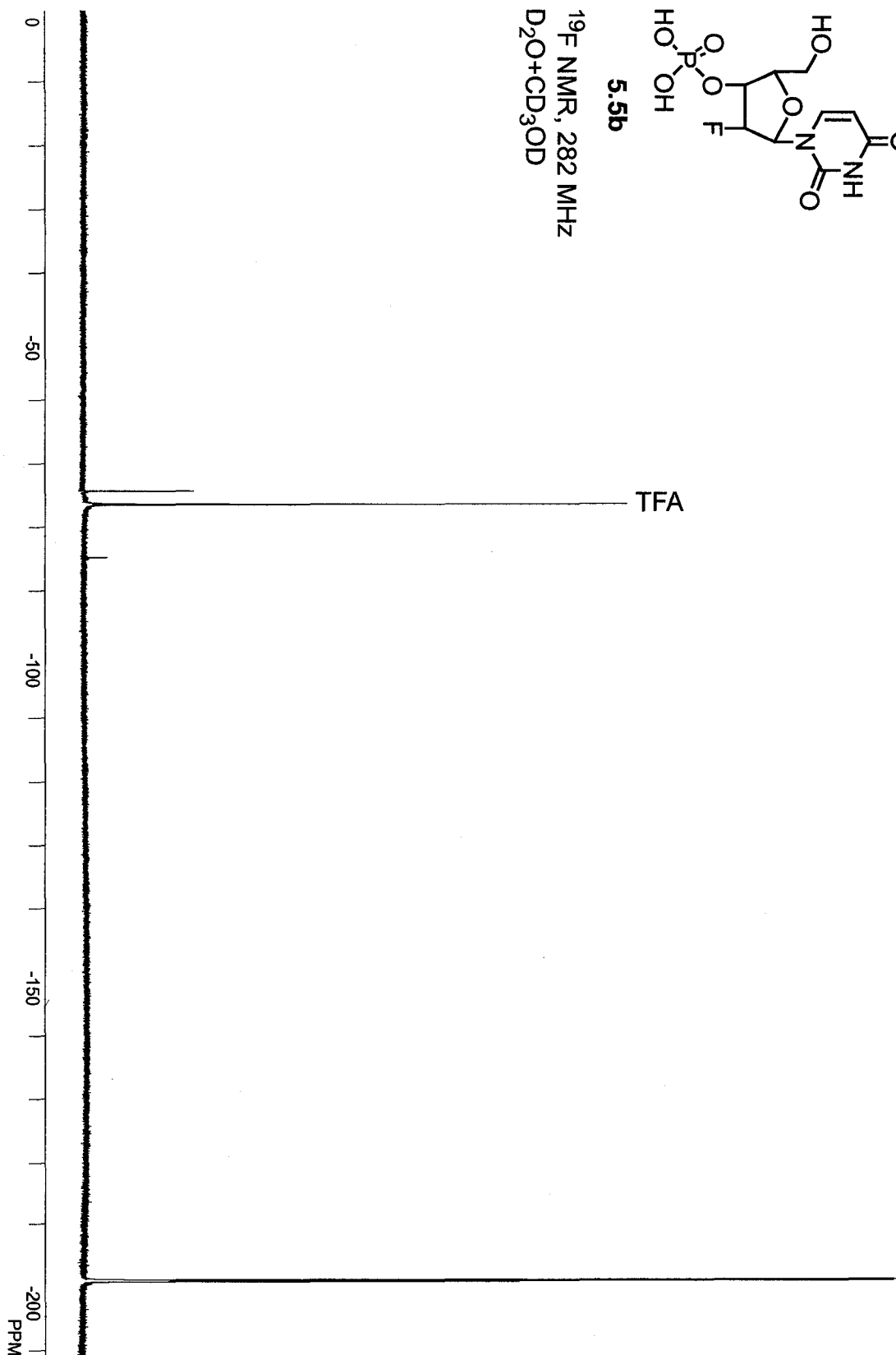
**5.5b**

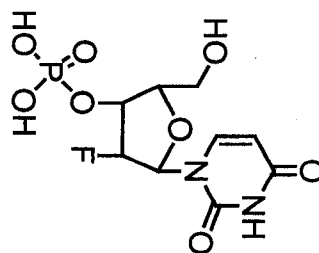
^{13}C NMR, 75 MHz
 $\text{D}_2\text{O} + \text{CD}_3\text{OD}$



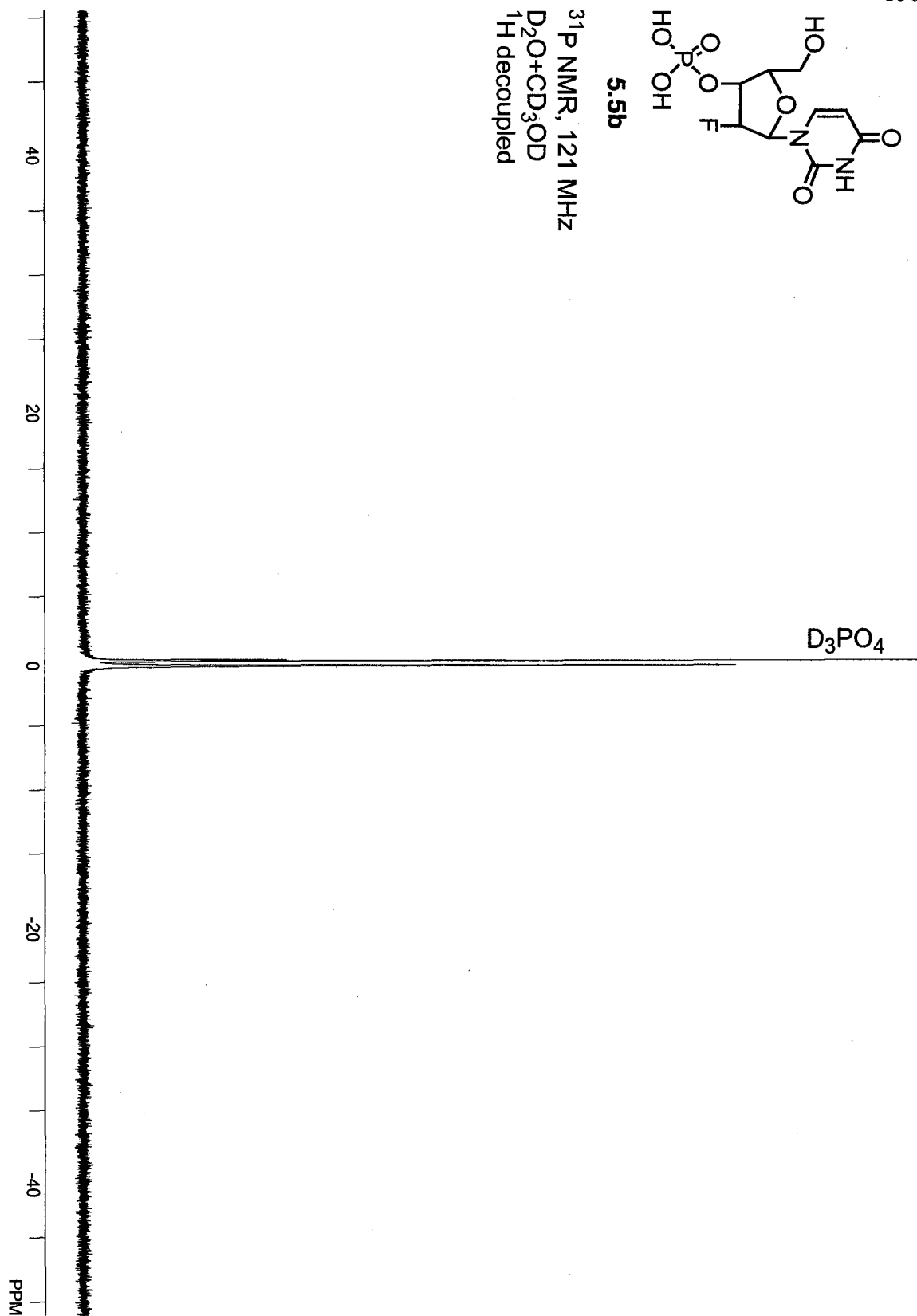
**5.5b**

^{19}F NMR, 282 MHz
 $\text{D}_2\text{O} + \text{CD}_3\text{OD}$

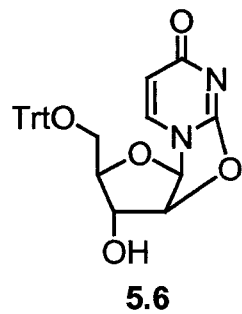


**5.5b**

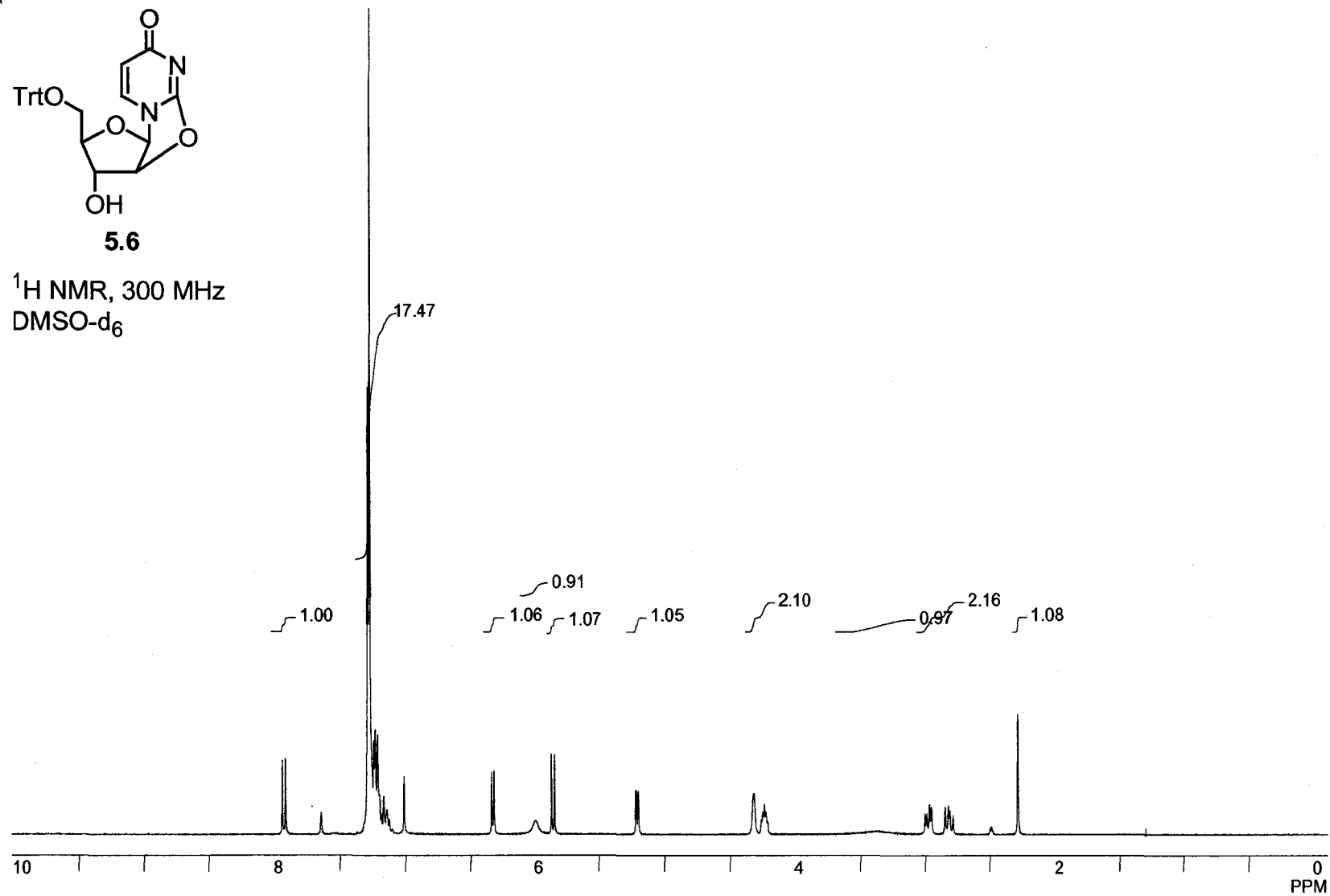
^{31}P NMR, 121 MHz
 $\text{D}_2\text{O} + \text{CD}_3\text{OD}$
 ^1H decoupled

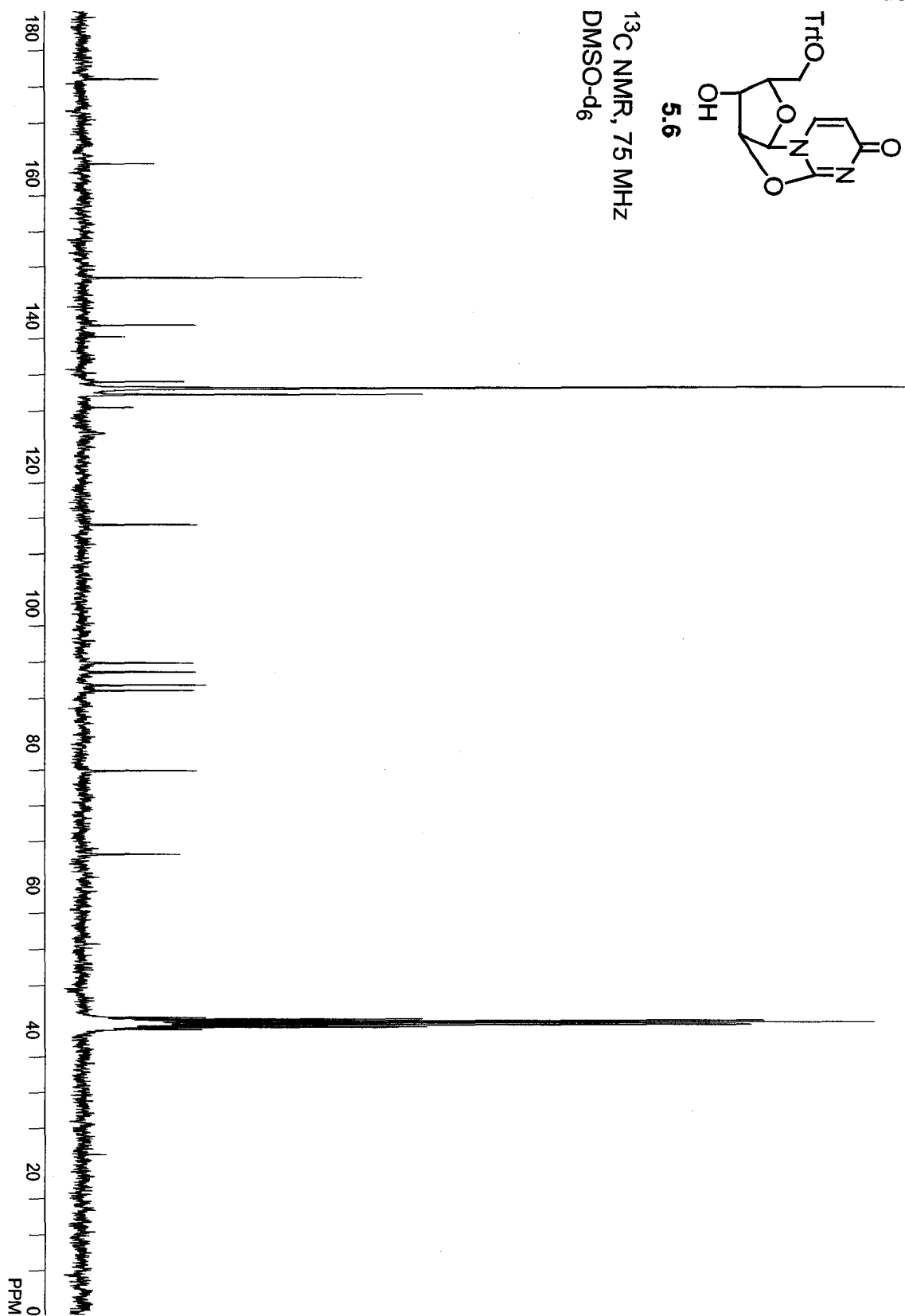
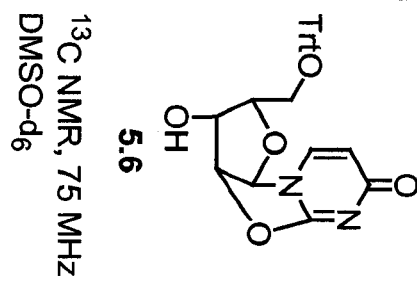


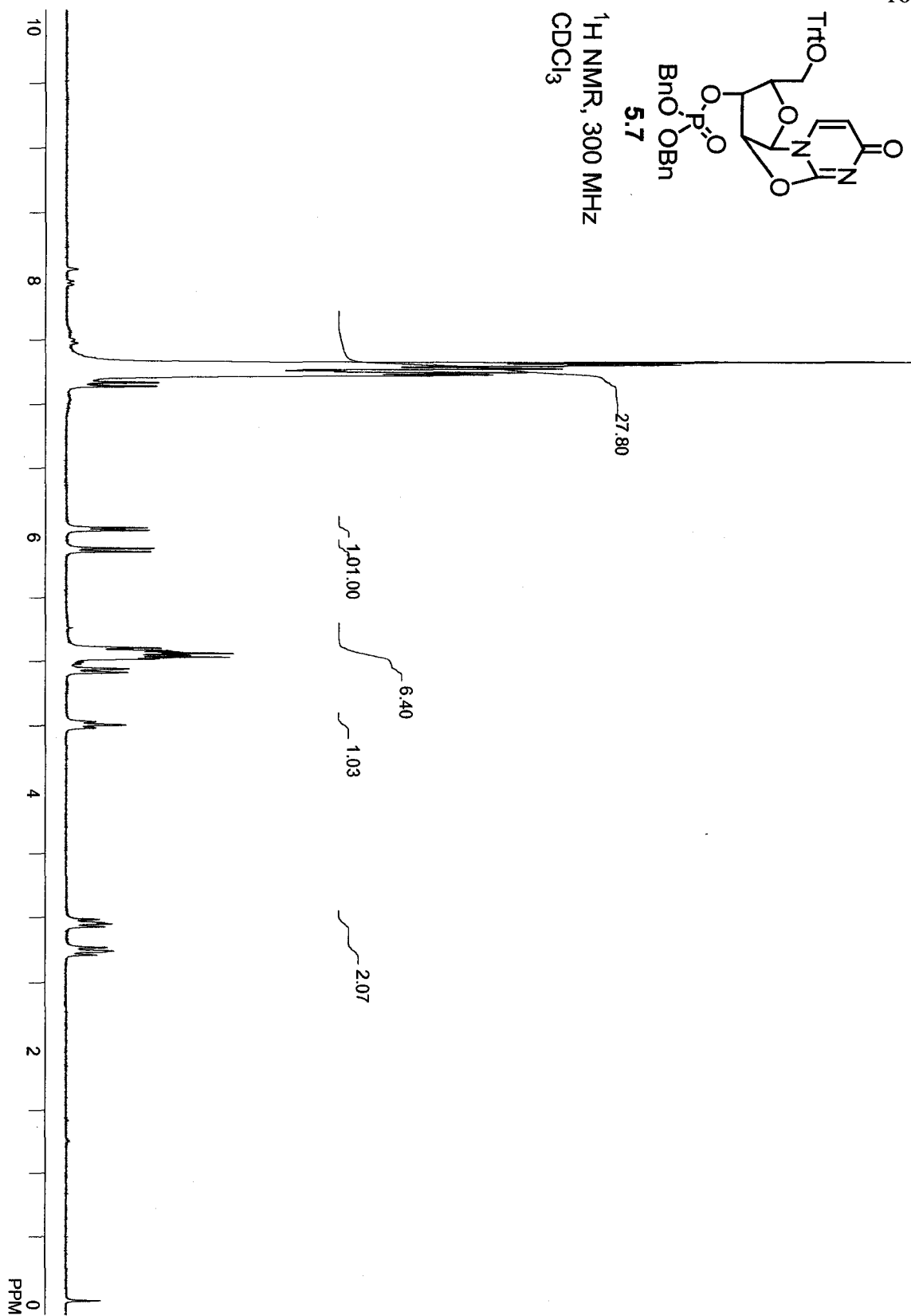
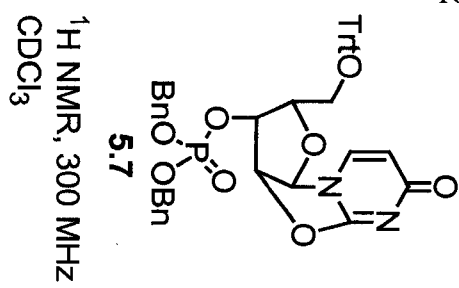
159

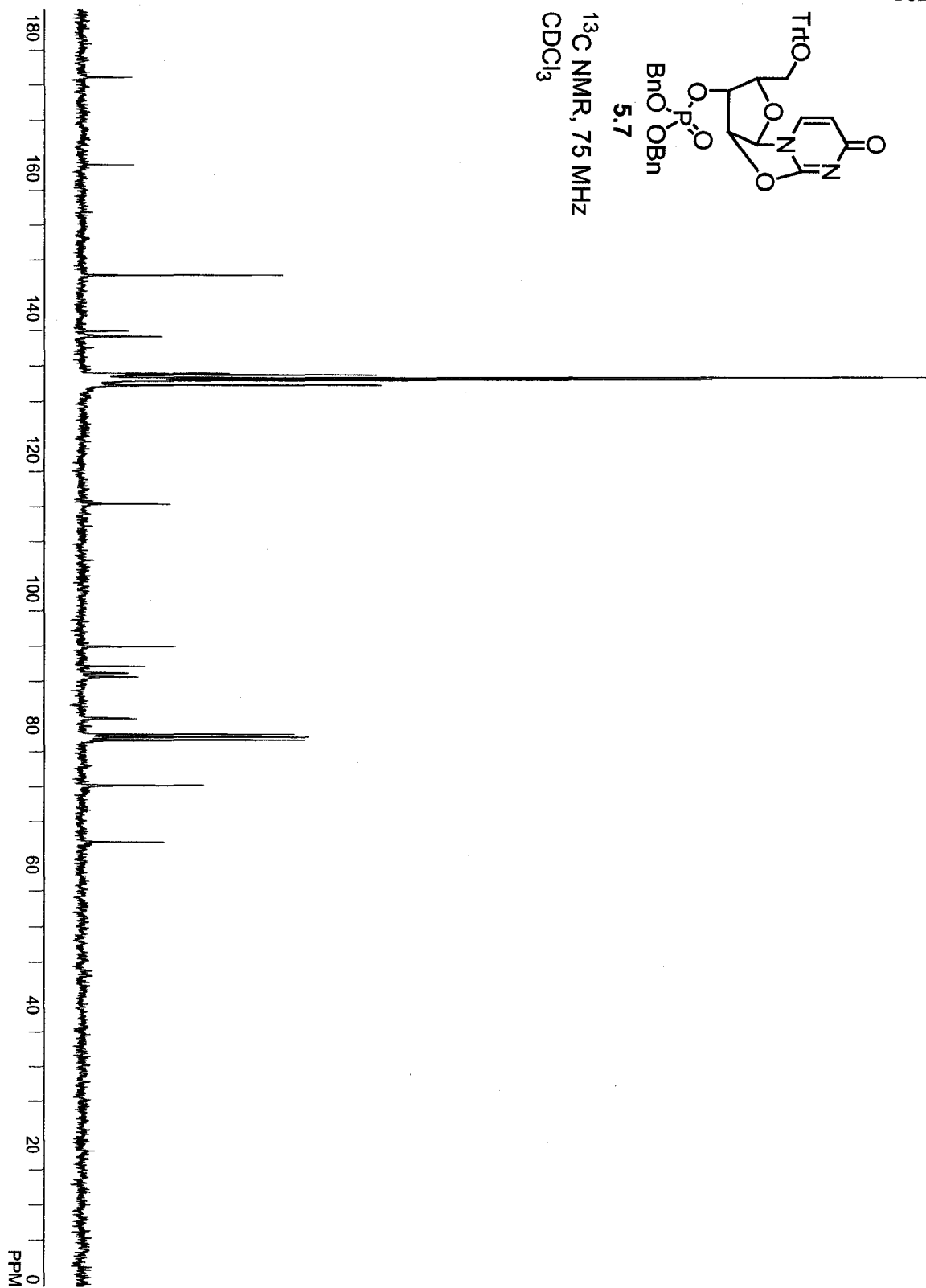
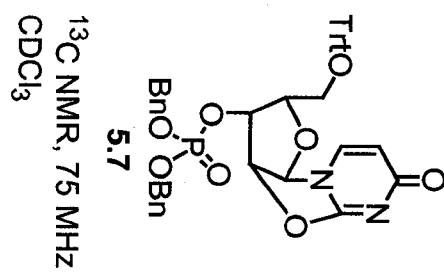


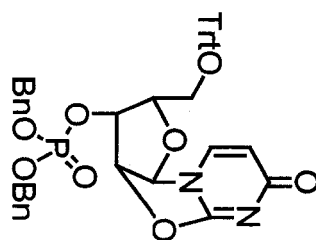
^1H NMR, 300 MHz
DMSO- d_6



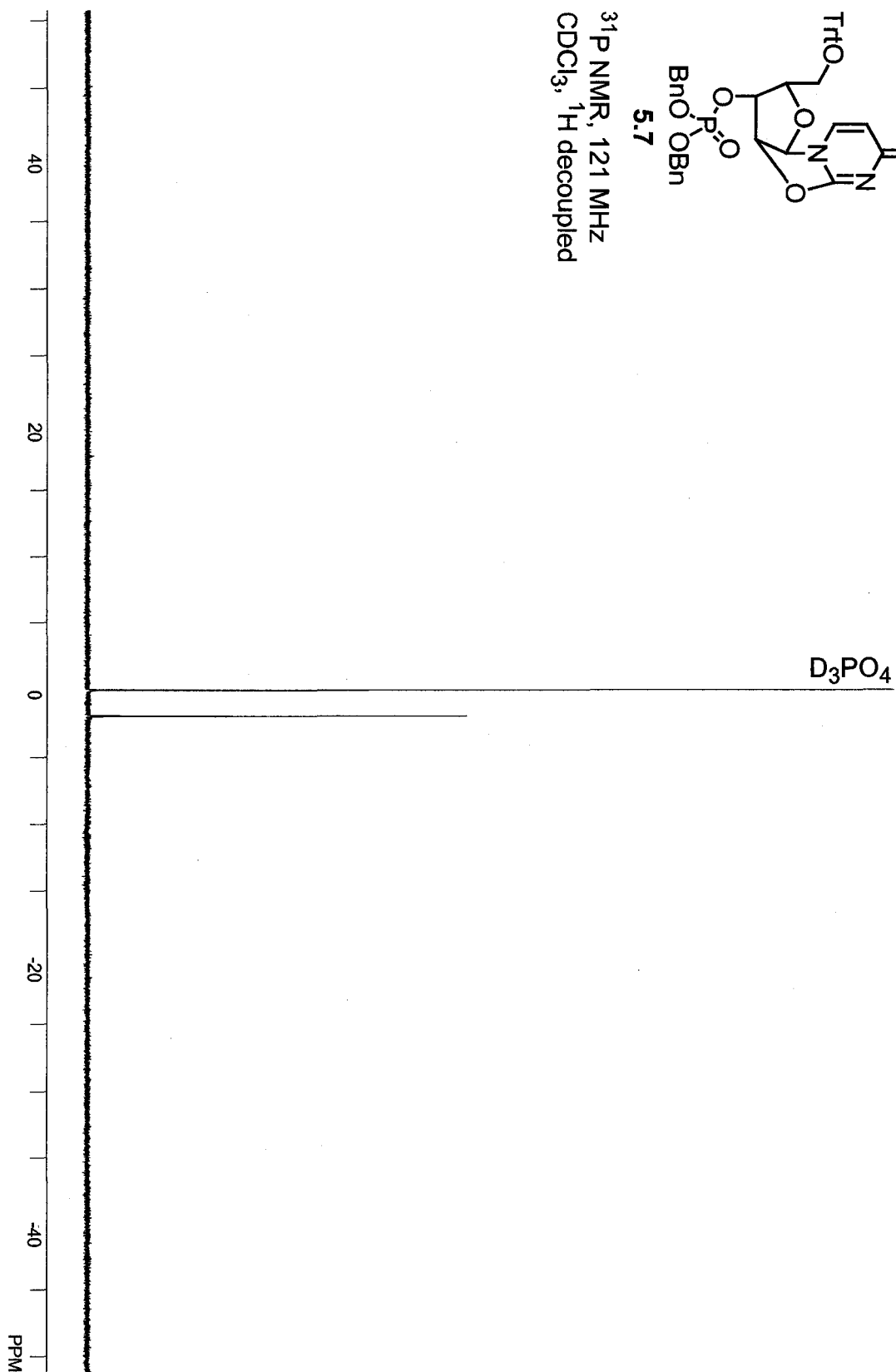


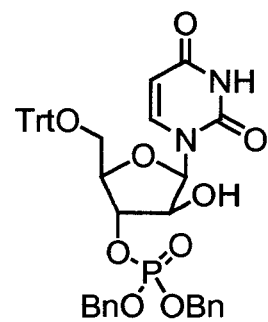




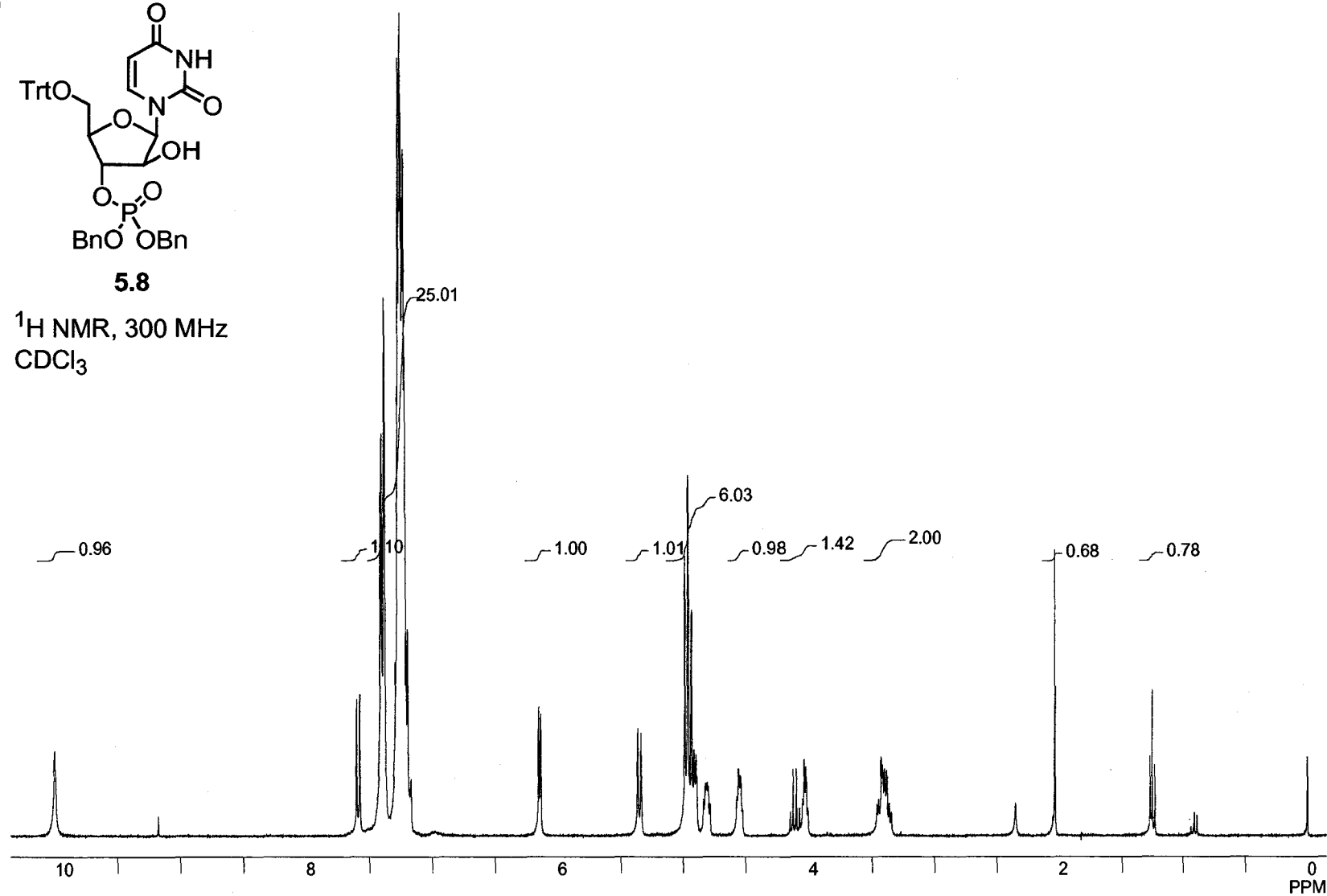


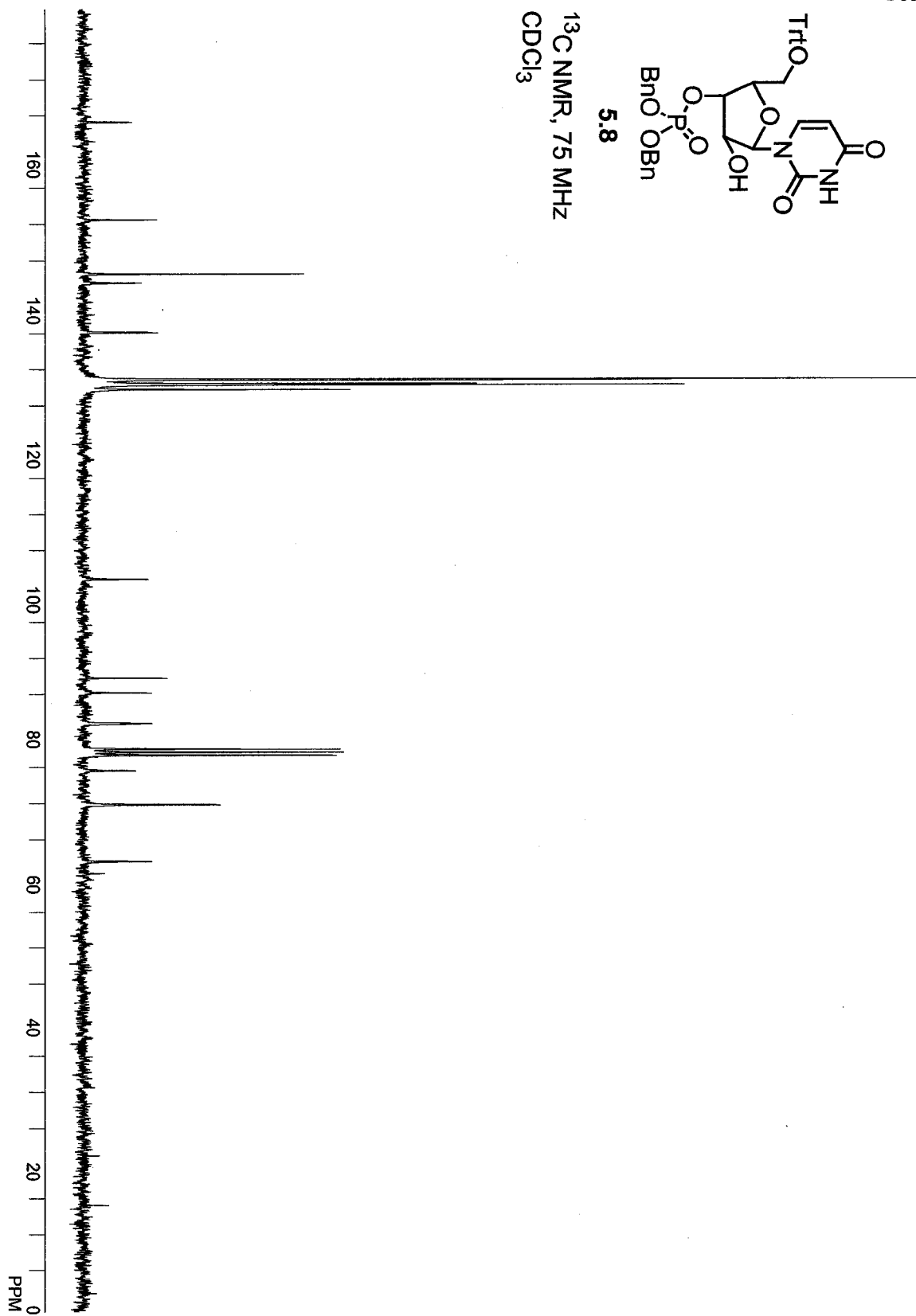
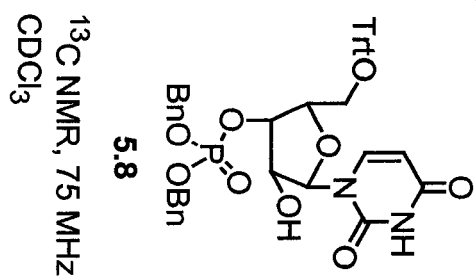
^{31}P NMR, 121 MHz
 CDCl_3 , ^1H decoupled

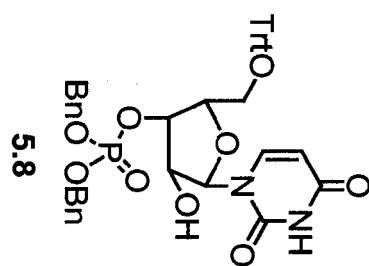


**5.8**

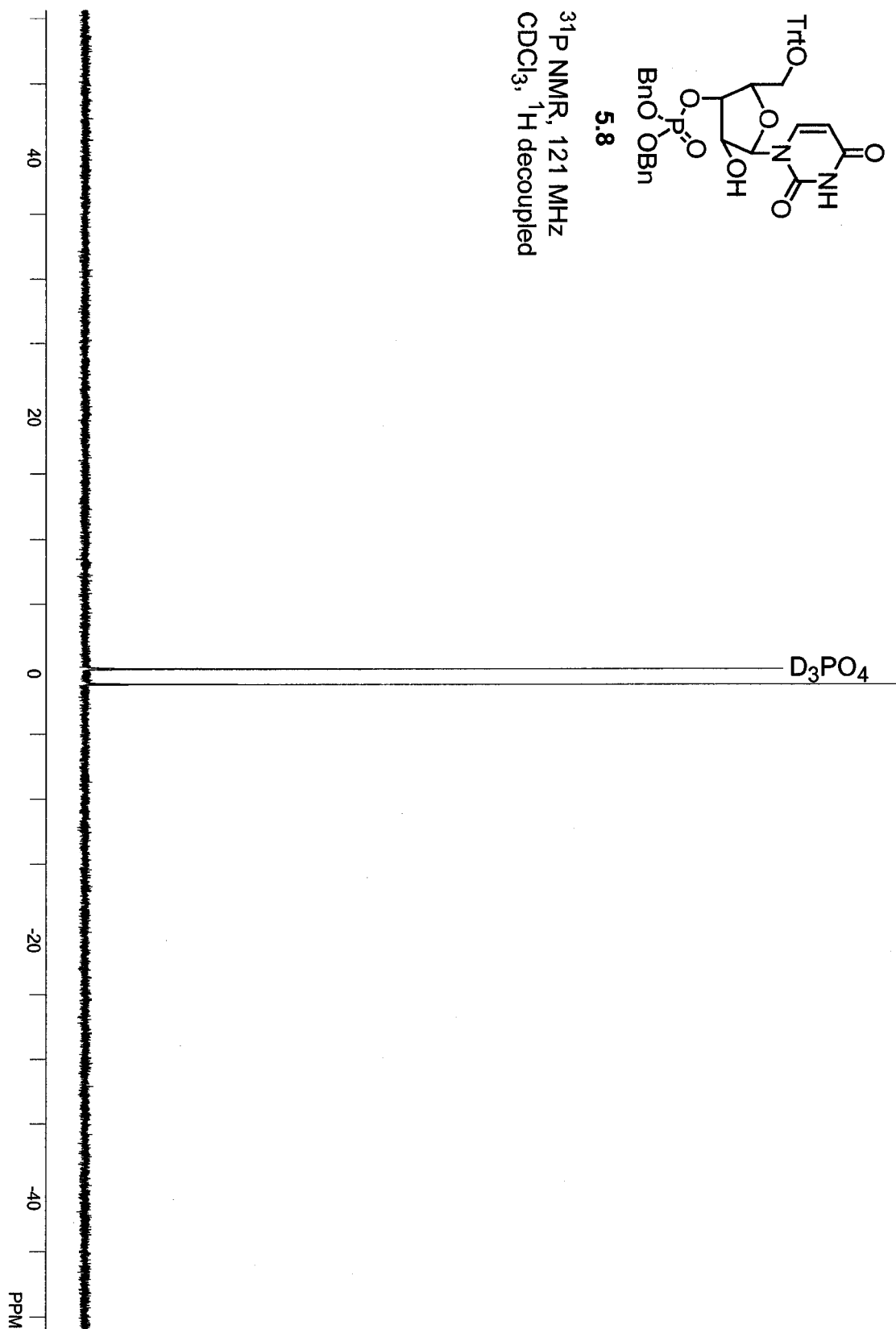
^1H NMR, 300 MHz
 CDCl_3

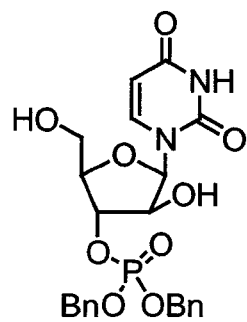




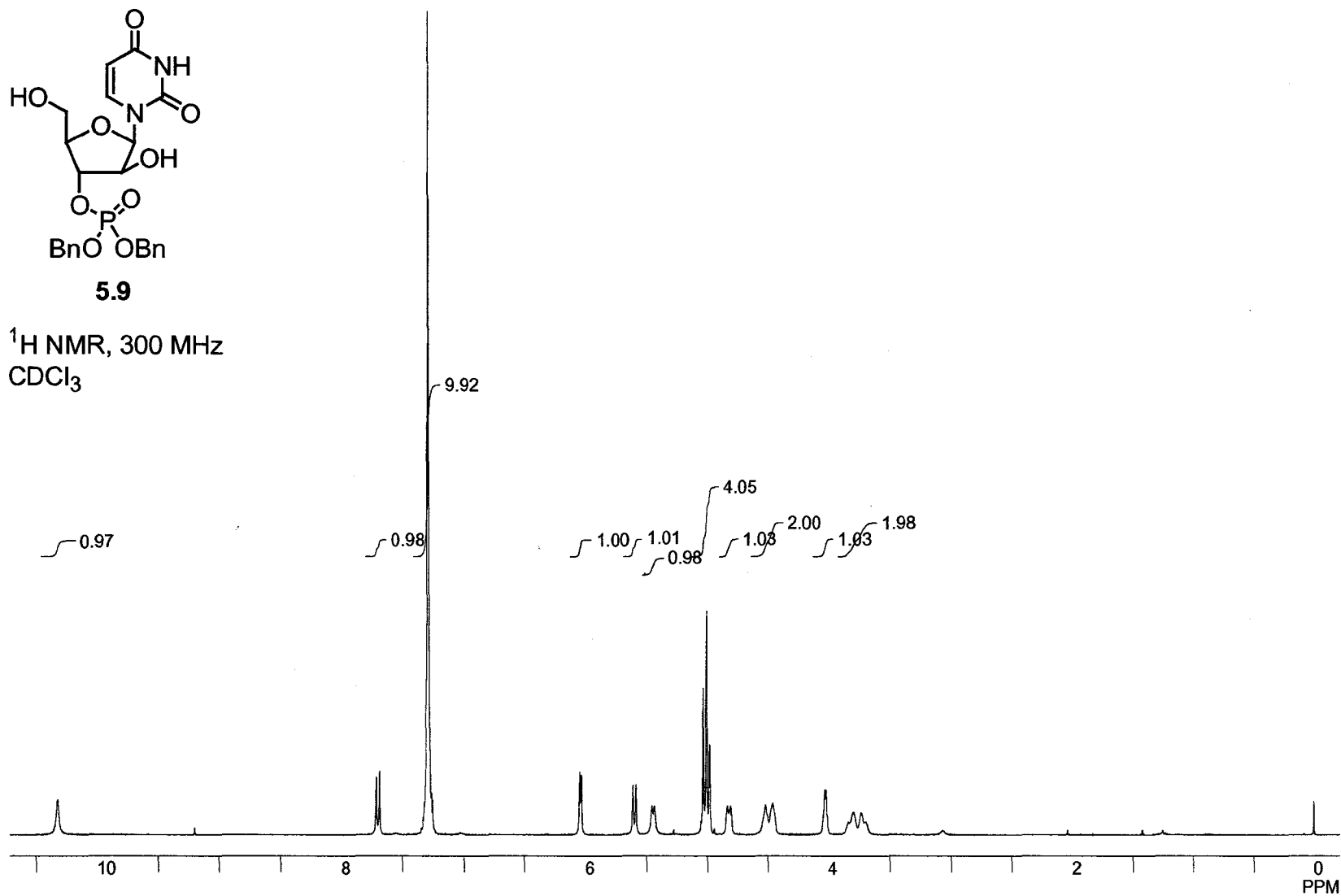


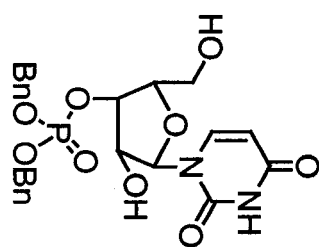
^{31}P NMR, 121 MHz
 CDCl_3 , ^1H decoupled



**5.9**

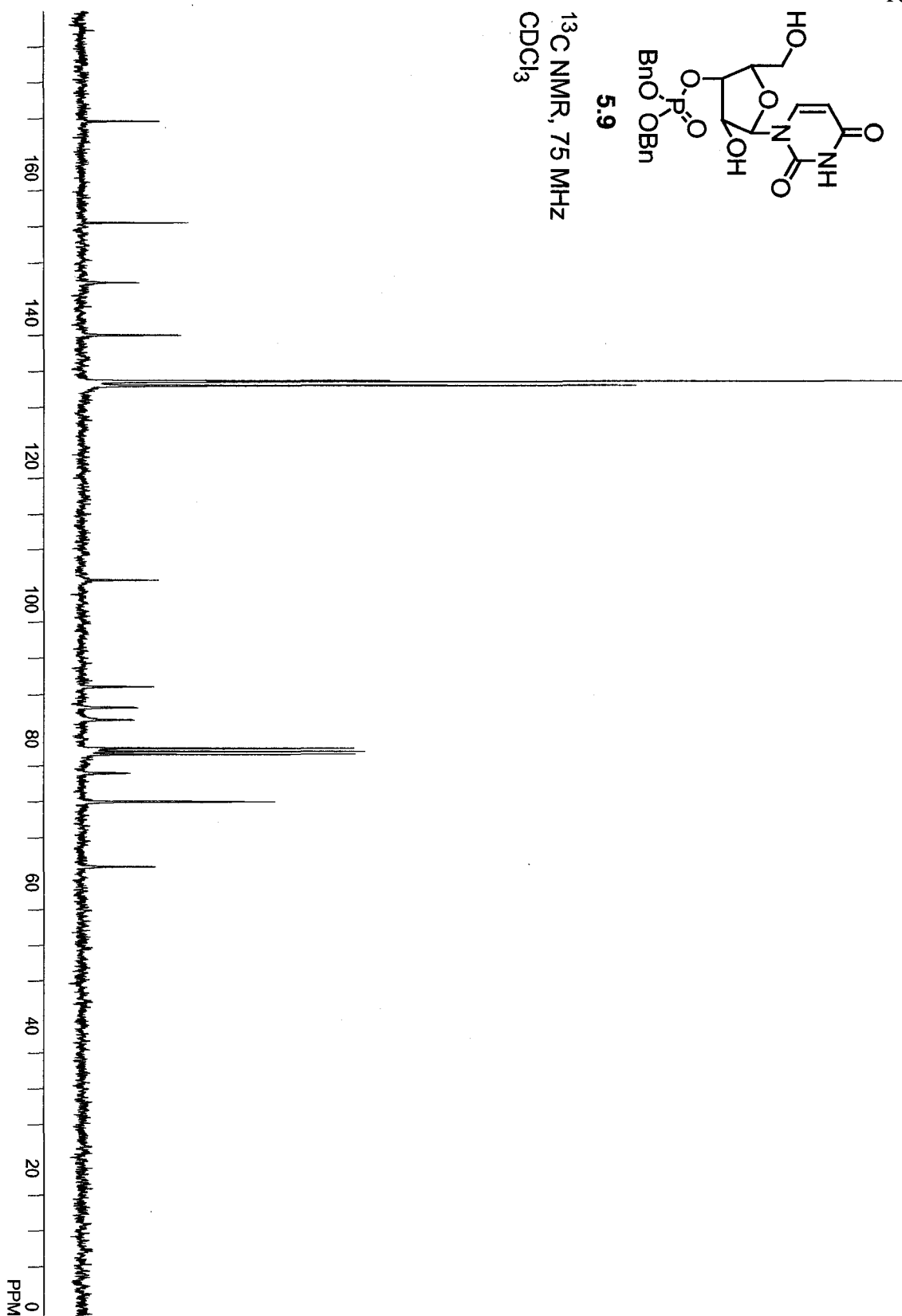
^1H NMR, 300 MHz
 CDCl_3

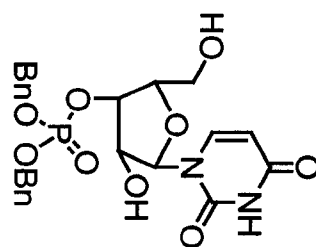




5.9

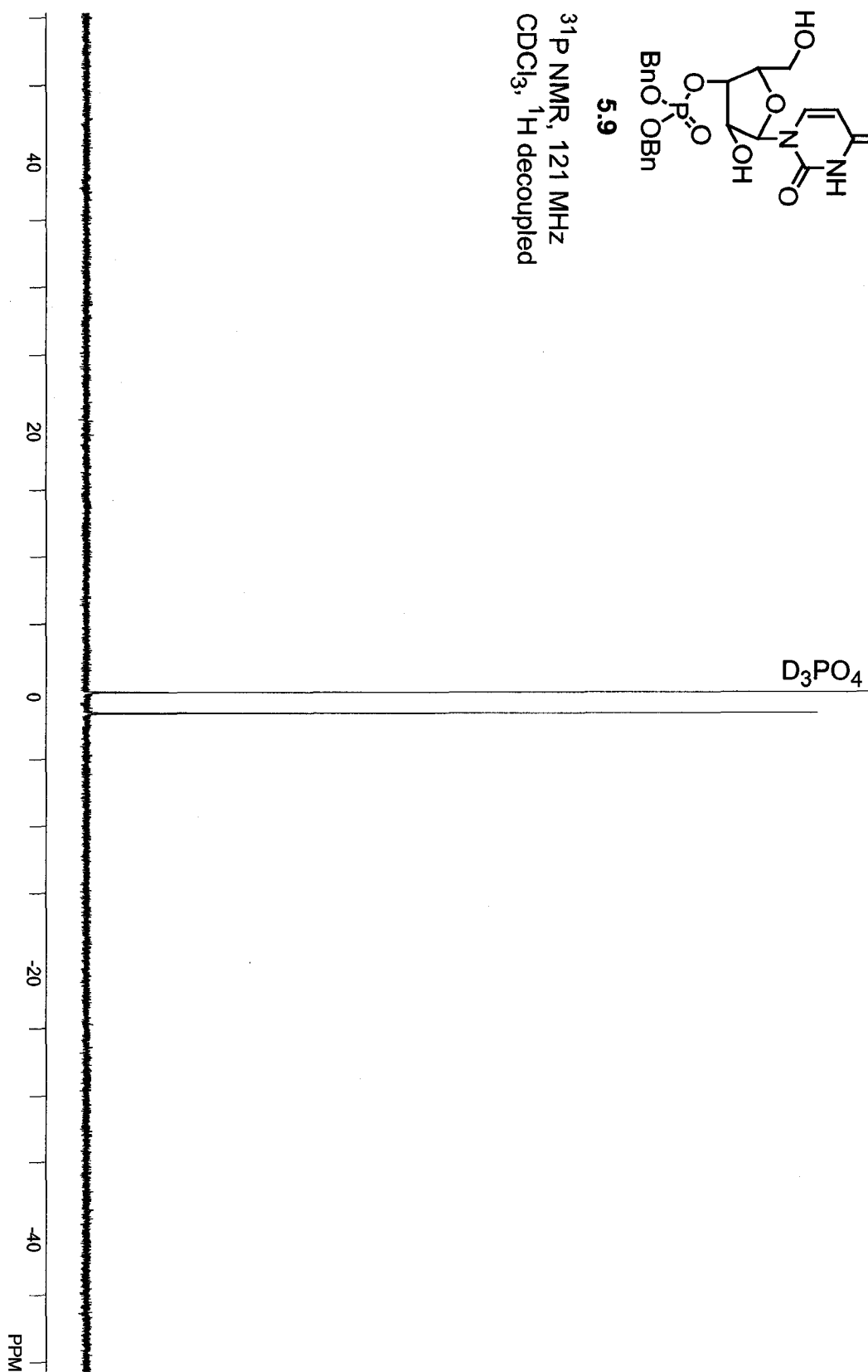
^{13}C NMR, 75 MHz
 CDCl_3

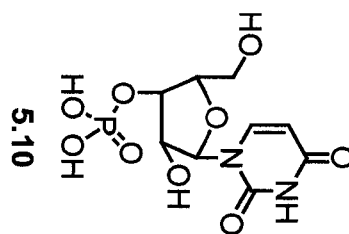




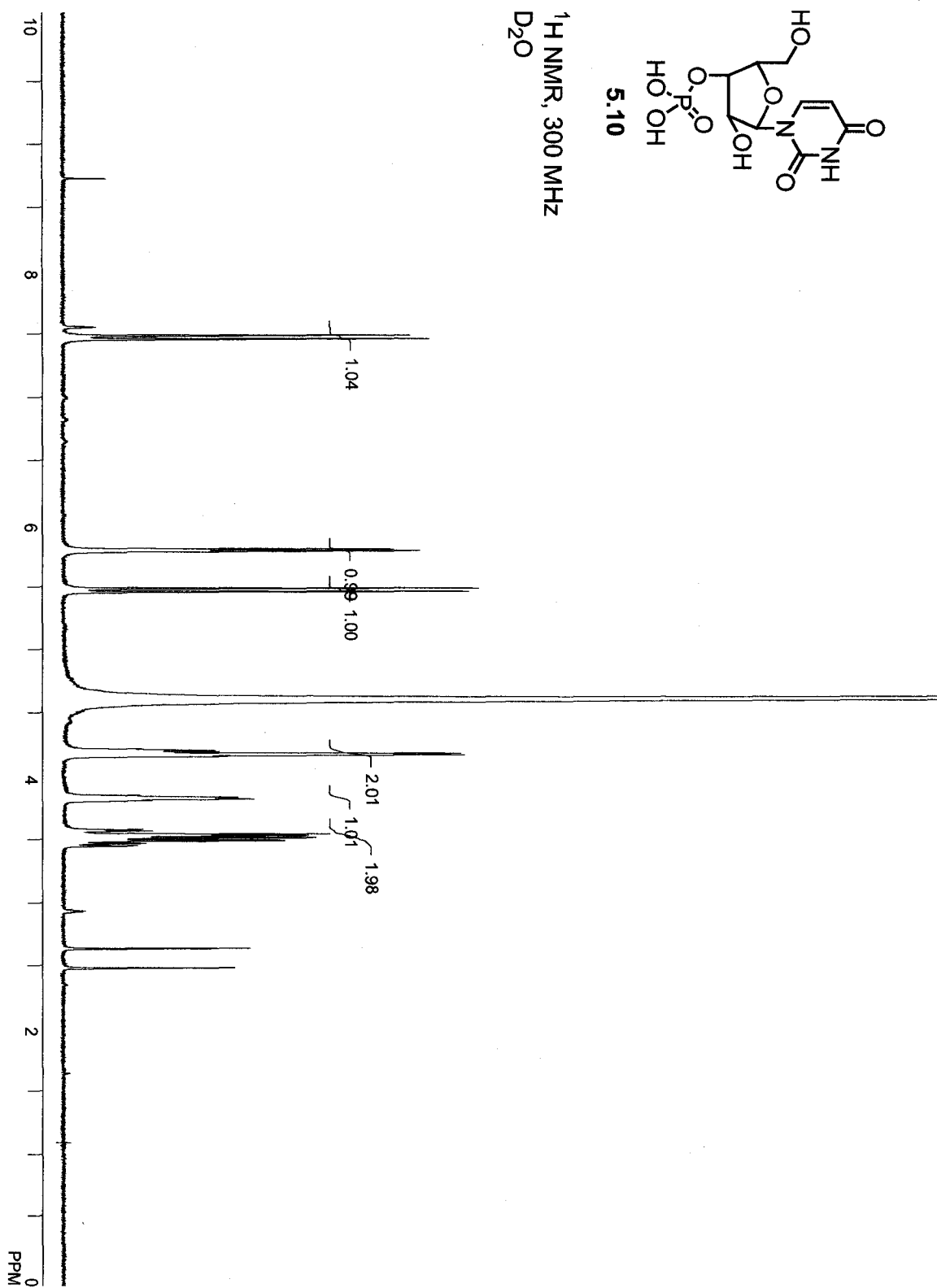
5.9

^{31}P NMR, 121 MHz
 CDCl_3 , ^1H decoupled

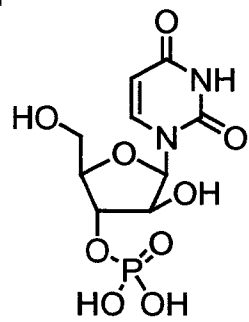




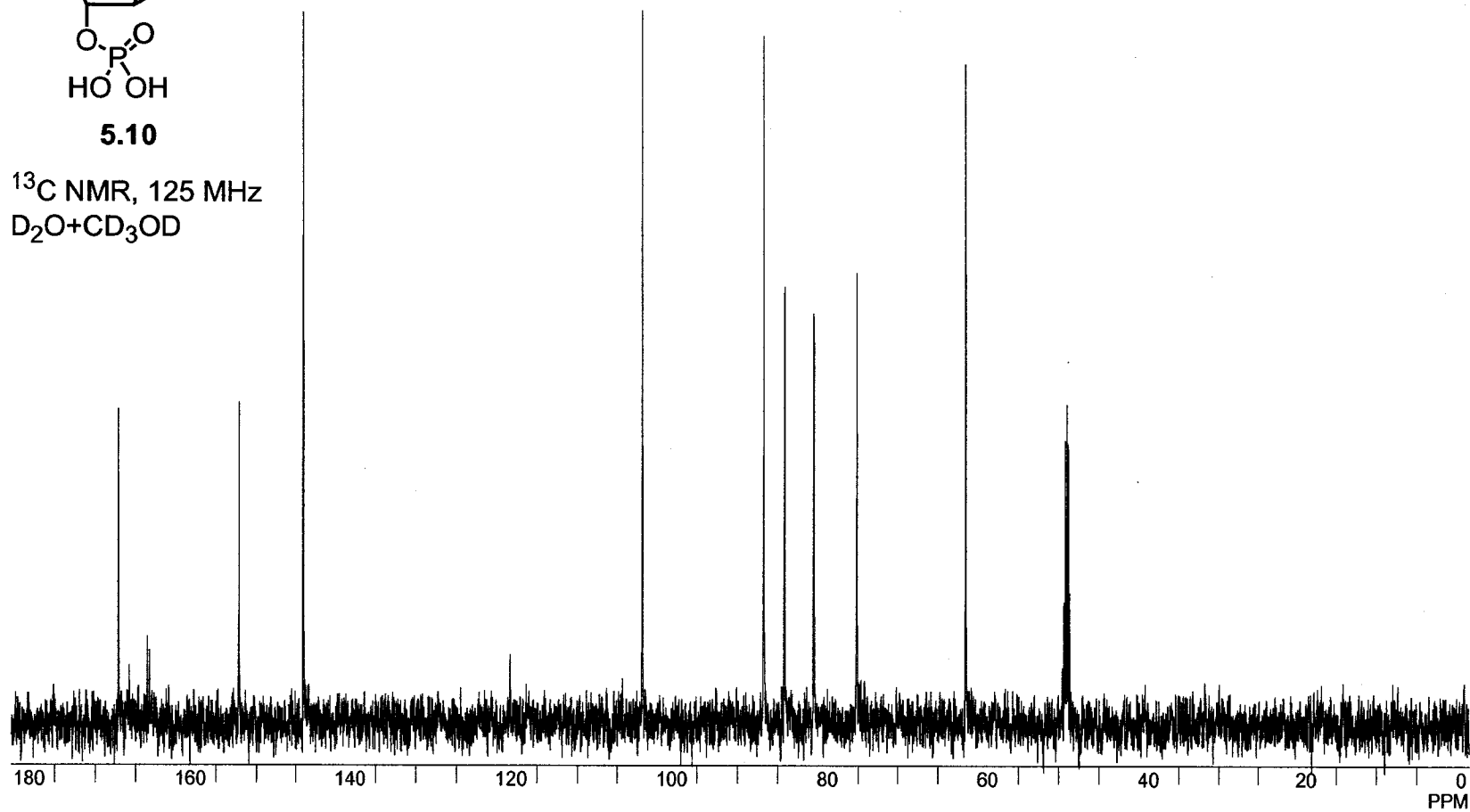
^1H NMR, 300 MHz
 D_2O

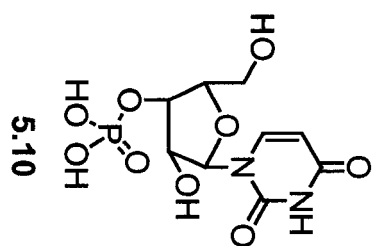


171

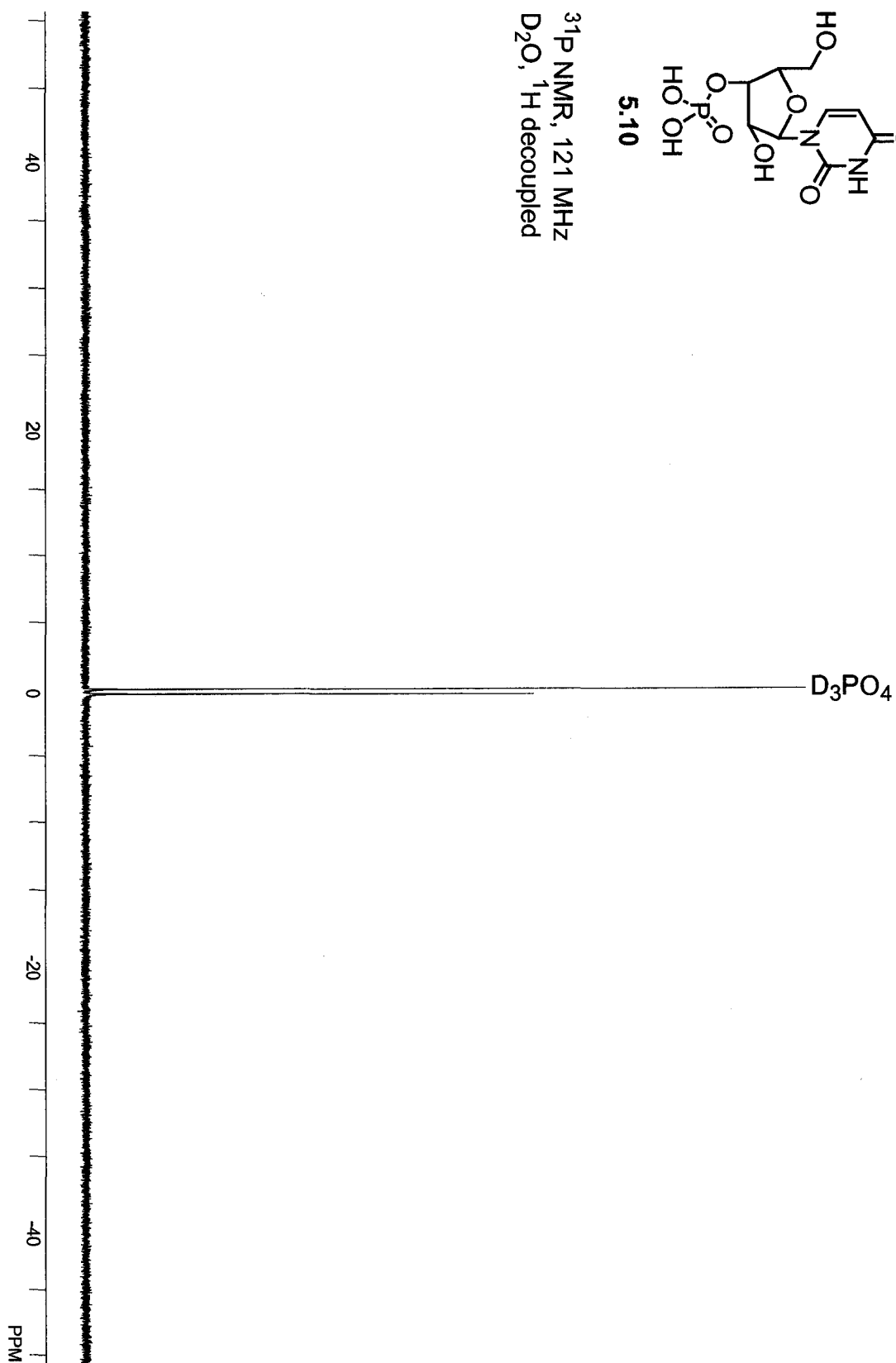
**5.10**

^{13}C NMR, 125 MHz
 $\text{D}_2\text{O} + \text{CD}_3\text{OD}$





^{31}P NMR, 121 MHz
 D_2O , ^1H decoupled



Chapter 6

COLLAGEN MIMICS CONTAINING PRO-GLY AMIDE ISOSTERES: PROBING THE IMPORTANCE OF $\text{GLY-NH}\cdots\text{O}=\text{C-PRO}$ HYDROGEN BONDS

6.1 Introduction

The collagen triple helix is a unique tertiary structure. The triple helix consists of three left-handed polyproline type II helices wound around a common axis to form a right-handed helix with a shallow super-helical pitch. In order for the protein to fold correctly, every third residue must be glycine (Persikov et al., 2000a), creating a triplet repeat of the sequence Xaa-Yaa-Gly. The glycine residues form hydrogen bonds with the residues in the Xaa position, with the Gly-NH as the proton donor and the carbonyl of the Xaa residue acting as the acceptor. The resulting hydrogen bond ladder (Fig 6.1) is thought to be a major contributor to the thermal stability of the triple helix (Ramachandran, 1967), but the strength of the hydrogen bonds has not yet been measured.

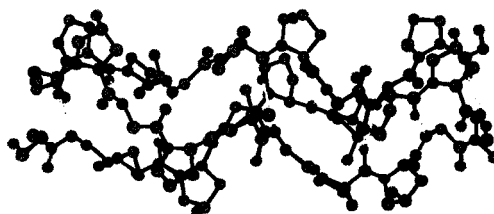


Figure 6.1 A segment of a (Pro-Hyp-Gly) triple helix highlighting the interstrand hydrogen bonding pattern. Hydrogen bonds are shown in yellow. Carbon is shown in gray, nitrogen in blue, and oxygen in red. Atomic coordinates are from PDB entry 1CAG (Bella et al., 1994).

One attempt to measure the strength of the interstrand hydrogen bond in collagen mimics was to measure the H/D fractionation factors of (Pro-Pro-Gly)₁₀ and (Pro-Hyp-Gly)₁₀ triple helices. The fractionation factor is the extent to which a particular hydrogen-bonding site becomes enriched in deuterium over hydrogen relative to solvent, and a smaller fractionation factor indicates a stronger hydrogen bond (Loh & Markley, 1994). This study led to the conclusion that the hydrogen bond strengths were similar in both peptides, but did not provide an estimate of those strengths (Danielson & Raines, 2000). We desired a method to more accurately measure the interstrand hydrogen bond strength in collagen mimics.

An approach to estimating hydrogen bond strength in proteins such as T4 lysozyme (Koh et al., 1997), staphylococcal nuclease (Chapman et al., 1997; Shin et al., 1997), chymotrypsin inhibitor 2 (Beligere & Dawson, 2000), turkey ovomucoid third domain (Lu et al., 1997), the GCN4 coiled coil domain (Blankenship et al., 2002), serine protease–protein inhibitor complexes (Lu et al., 1999), 4-oxalocrotonate tautomerase (Silinski & Fitzgerald, 2003), and amyloid fibrils (Gordon & Meredith, 2003), has been to replace the amide N-H linkage in question with an ester, eliminating the hydrogen bond donor in one (or more (Beligere & Dawson, 2000)) site. The substitution of ester for amide is considered a conservative substitution because both moieties have similar conformational preferences (Wiberg & Laidig, 1987), and inclusion of lactic acid unit in place of an alanine residue in an α -helical model peptide has been shown to induce minimal structural perturbations (Karle et al., 2001; Aravinda et al., 2002). On the other hand, it has been shown that there is more backbone flexibility around an ester linkage than around an amide (Mammi & Goodman, 1986). While esters and amides are

approximately isosteric, they have somewhat different electronic properties, which introduces differences in their Coulombic interactions and hydrogen bond accepting capabilities. As a result, the hydrogen bond strengths obtained in this way are estimates, as it is difficult to correct for these differences, but these estimates have been found quite informative.

Another approach to eliminating hydrogen-bonding capability with minimal structural perturbation is to replace an amide bond with an alkene, creating a dipeptide alkene isostere (Hann et al., 1980; Wipf & Fritch, 1994). A variety of approaches to the synthesis of dipeptide alkene isosteres has been reported in the literature (Gardner et al., 1995; Masse et al., 1997; Wipf et al., 1998; Gardner et al., 1999; Oishi et al., 2002; Tamamura et al., 2002; Vasbinder & Miller, 2002; Tamamura et al., 2003; Wang et al., 2003), but their use in biological systems has been largely restricted to replacing amide groups in peptide-derived enzyme inhibitors to decrease their susceptibility to proteolysis (Johnson, R. L., 1984; Tamamura et al., 2003).

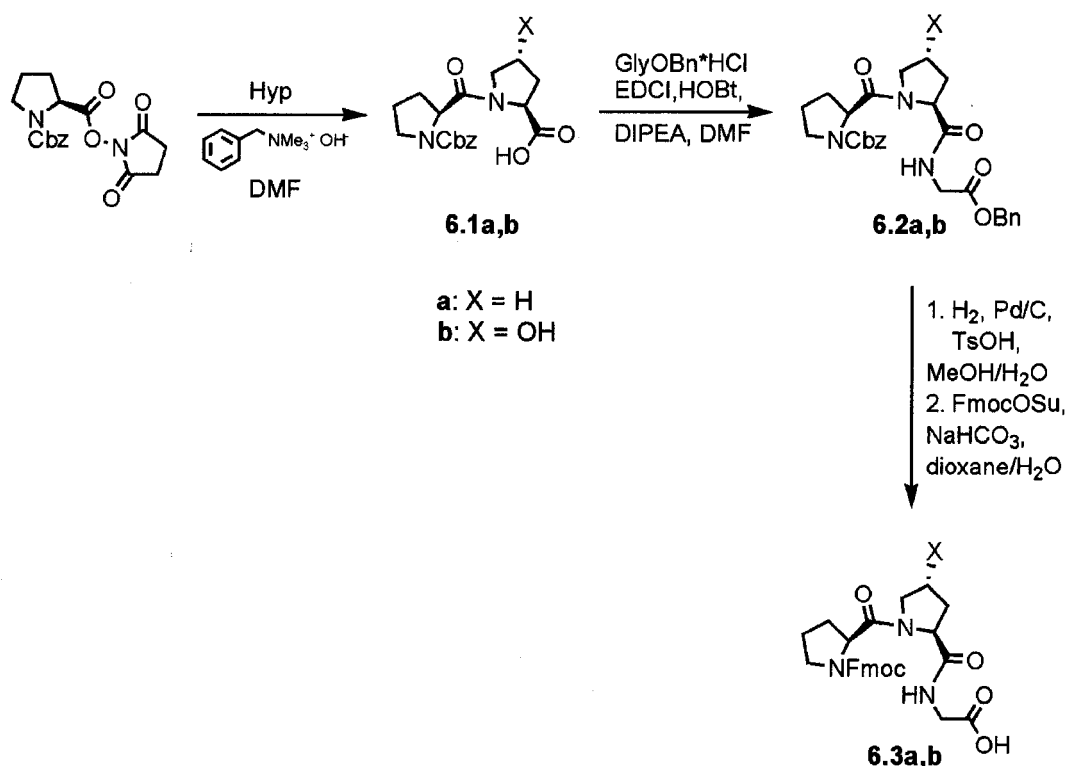
We hypothesized that replacing one glycine N-H linkage in a collagen mimic peptide with an ester or alkene linkage would allow us to estimate the contribution of interstrand hydrogen bonds to triple helix stability.

6.2 Results and Discussion

6.2.1 *Synthesis of Fmoc-tripeptides for Use in Solid Phase Peptide Synthesis*

To facilitate the assembly of collagen mimic peptides with the repeating sequence (Pro-Yaa-Gly)₁₀ (Yaa = Pro or 4(*R*)-hydroxy-L-proline, Hyp) we synthesized the tripeptides Fmoc-Pro-Pro-Gly and Fmoc-Pro-Hyp-Gly in solution according to the

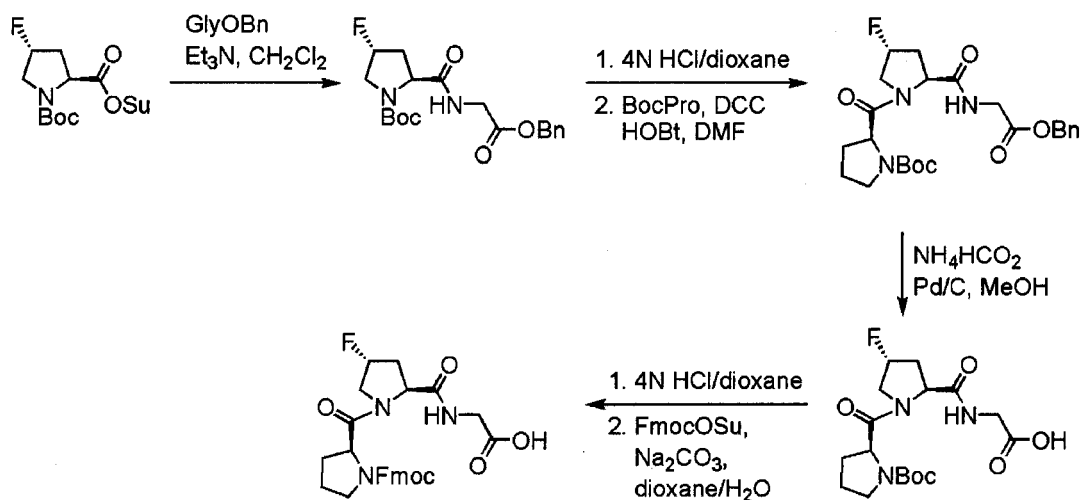
procedure of Ottl and Moroder (Scheme 6.1) (Ottl et al., 1999). Their reported overall yield for the synthesis of Fmoc-Pro-Hyp-Gly (75% over 4 steps) was much higher than we could achieve, as our average yields over 4 steps were 16.2% for Fmoc-Pro-Hyp-Gly and 19.2% for Fmoc-Pro-Pro-Gly. In addition, we found that the yield of the final step was poor because of difficulty in purifying the final product. As a result, we desired a more efficient route to Fmoc-tripeptides starting from readily available starting materials.



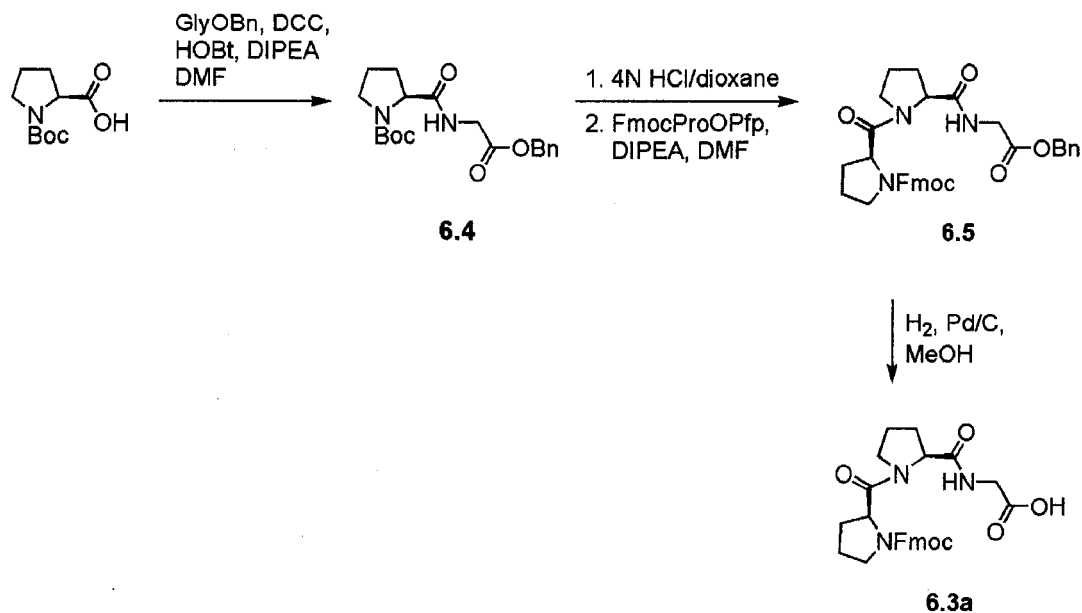
Scheme 6.1 Synthesis of Fmoc-Pro-Yaa-Gly according to the method of Ottl and Moroder (Yaa = Pro or Hyp) (Ottl et al., 1999).

Holmgren, et al. reported a 5-step synthesis for Fmoc-Pro-Flp-Gly (where Flp = 4(*R*)-fluoro-L-proline) based on standard Boc chemistry with an average overall yield of 16.7% (Scheme 6.2) (Holmgren et al., 1999). We thought that this general route could be shortened and made more efficient by eliminating the need to switch from an N-terminal

Boc protecting group to Fmoc in the final step. Following the steps outlined in Scheme 6.3, with careful monitoring the final hydrogenolysis reaction, we found that we could achieve an overall yield of 56.6% over three steps for Fmoc-Pro-Pro-Gly starting from Boc-Pro. It has also been shown that this synthetic scheme works well for the synthesis of Fmoc-Pro-Hyp-Gly (F.W. Kotch, personal communication) and Fmoc-Pro-Flp-Gly (C.L.J., data not shown.) One of the primary advantages of this route is the high yield and facile purification of the final product from the byproducts of the benzyl deprotection.



Scheme 6.2 Synthesis of Fmoc-Pro-Flp-Gly (Holmgren et al., 1999).

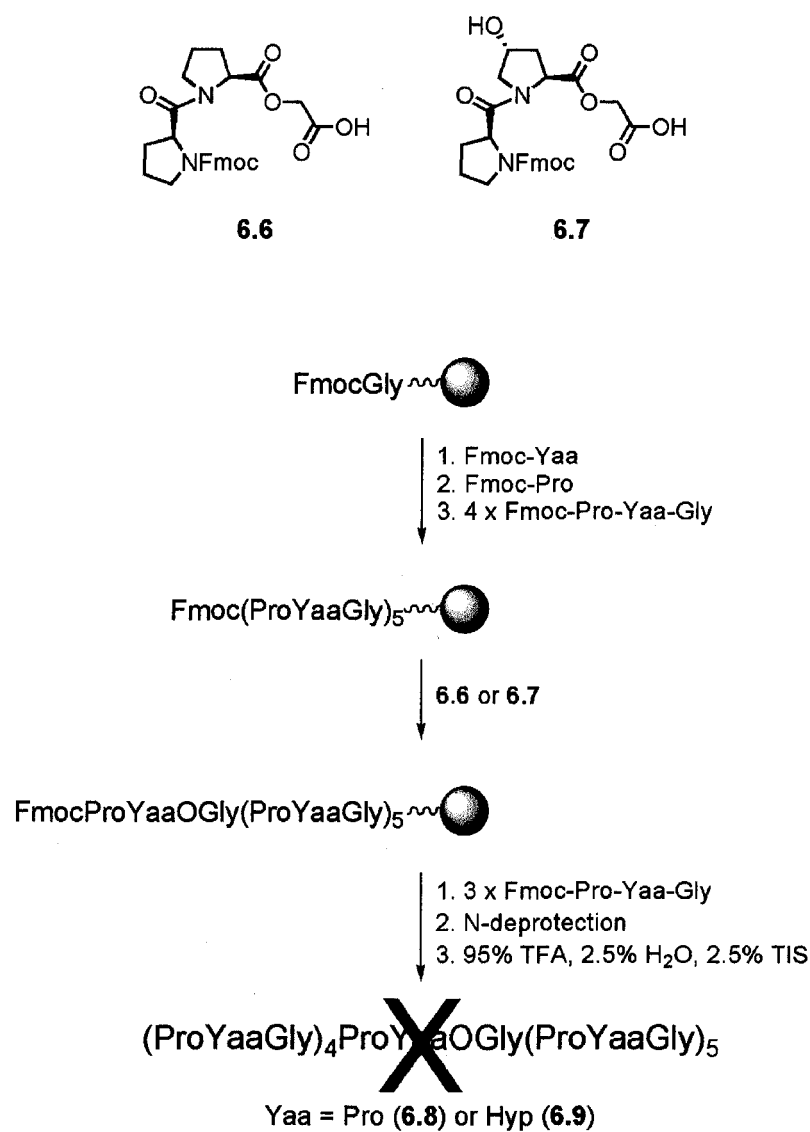


Scheme 6.3 Improved synthesis of Fmoc-Pro-Pro-Gly.

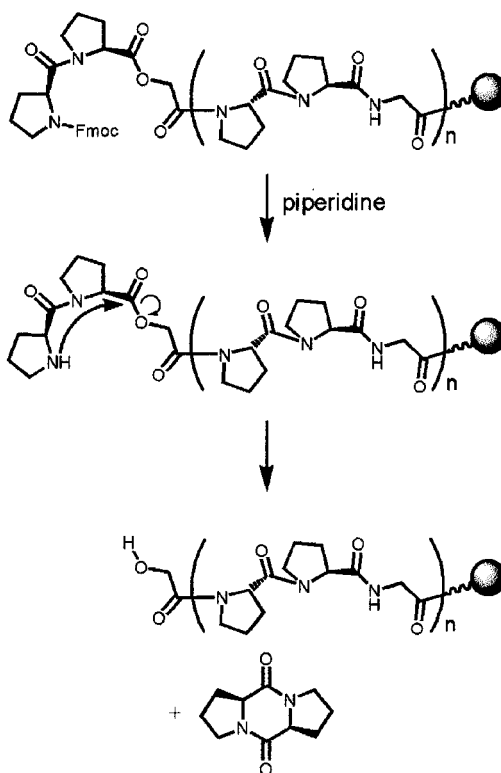
6.2.2 Solid Phase Synthesis of Depsipeptides

We also synthesized the Fmoc-trimers **6.6** and **6.7** using the method of Ottil and Moroder (Ottil et al., 1999) to facilitate assembly of depsipeptides **6.8** and **6.9**, which were designed to eliminate one hydrogen bond donor from each strand of the triple helix. We found that the solid phase peptide synthesis (SPPS) of **6.8** and **6.9** according to Scheme 6.4 resulted in the isolation of peptides lacking a Pro-Pro segment in the case of **6.8**, or a Pro-Hyp segment in the case of **6.9**. The loss of these dipeptide segments was presumably due to diketopiperazine formation (Gisin & Merrifield, 1972), which was facilitated by the enhanced leaving group ability of glycolate compared to glycine (Scheme 6.5). To minimize diketopiperazine formation, we built up the ester-containing sections of **6.8** and **6.9** by incorporating two Fmoc-dipeptide segments (compound **6.10** or **6.11**, Scheme 6.6) and two Fmoc-protected single amino acids (Scheme 6.7). This route avoids the presence of a free N-terminal nucleophile two amino acids away from the labile ester during the

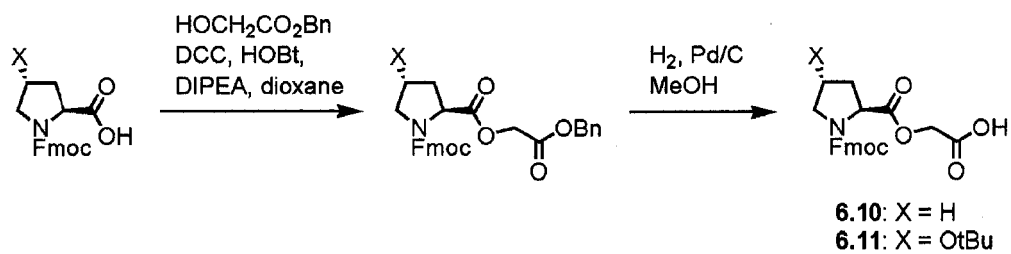
synthesis, eliminating the diketopiperazine formation. Cleavage of the peptides from the resin with TFA and subsequent HPLC purification yielded full-length **6.8** and **6.9**, as determined by MALDI-TOF MS.



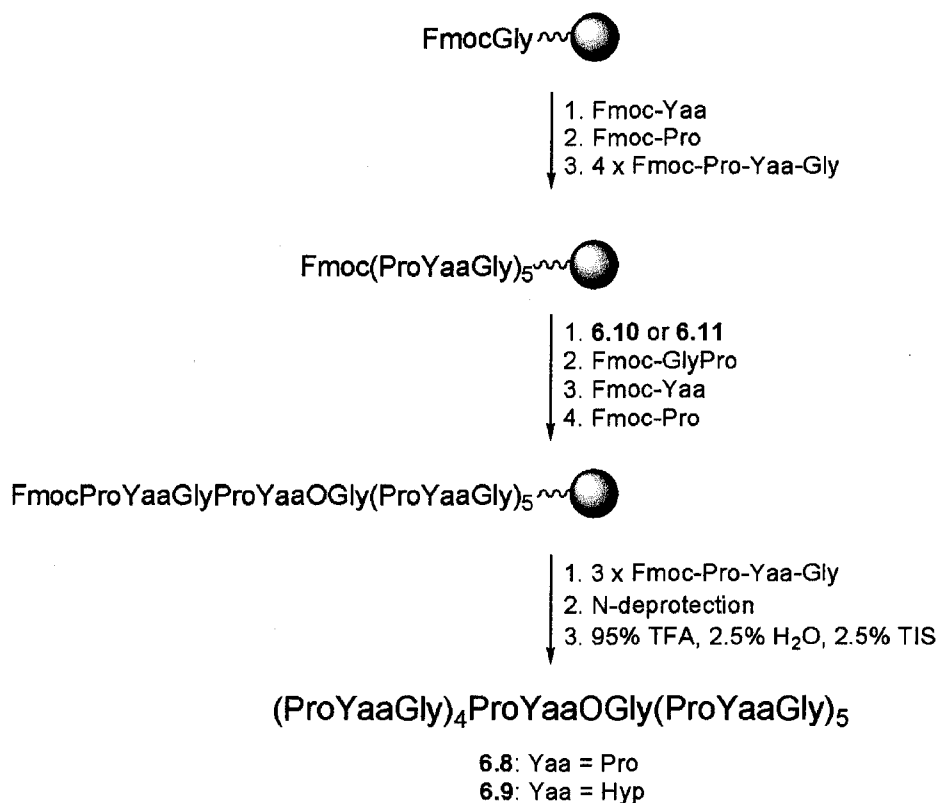
Scheme 6.4 Failed synthesis of **6.8** and **6.9**.



Scheme 6.5 *Proposed mechanism for two-residue deletion.*



Scheme 6.6 *Synthesis of 6.10 and 6.11.*



Scheme 6.7 *Successful synthesis of 6.8 and 6.9*

6.2.3 Effects of Ester and Alkene Isosteres on Triple Helix Formation and Stability

Depsipeptides **6.8** and **6.9** were incubated at 4 °C at a concentration of 0.2 mM in 50 mM HOAc for approximately 24 h, and then studied by CD spectroscopy. Wavelength scans from 200–260 nm indicated that neither peptide had assembled into a triple helix. The absence of the triple helix was confirmed by the linear decrease in ellipticity with increasing temperature. In contrast, (Pro-Pro-Gly)₁₀ had a melting temperature (T_m , which is the temperature at the midpoint of the thermal transition between native and unfolded states) of 41 °C and (Pro-Hyp-Gly)₁₀ had a T_m of 69 °C under the same conditions. The failure of **6.8** and **6.9** to assemble into triple helices indicates that each interstrand

hydrogen bond contributes significantly to triple helix stability, but we could not form a good estimate of the magnitude of that contribution.

In studying host-guest collagen mimics with the general structure Ac-(Gly-Pro-Hyp)₃-Zaa-Pro-Hyp-(Gly-Pro-Hyp)₄-Gly-Gly-NH₂, Brodsky and coworkers used the osmolyte trimethylamine oxide (TMAO) to enhance triple helix thermal stability so that they could compare the melting temperatures of a variety of Gly→Zaa variants (Beck et al., 2000). TMAO is thought to enhance protein stability by an unfavorable interaction with backbone amide groups. This unfavorable interaction is probably an indirect interaction arising from enhancement of water structure by TMAO (Zou et al., 2002). We hypothesized that TMAO might stabilize triple helices of **6.8** and **6.9** enough for us to measure T_m values and compare them to those of (Pro-Hyp-Gly)₁₀ and (Pro-Pro-Gly)₁₀ under the same conditions, thus enabling us to obtain an estimate of hydrogen bond strengths. Indeed, incubating **6.8** and **6.9** with various concentrations of TMAO in phosphate-buffered saline (PBS, pH 7.4, 150 mM NaCl) for 24 h led to increasing ellipticity at 225 nm with increasing TMAO concentration at 4 °C, characteristic of increasing triple helical content (Fig. 6.2). Triple helices of depsipeptide **6.8** began to show a thermal transition in solutions containing 2 M TMAO or higher, while those of depsipeptide **6.9** did not exhibit clear transitions at any TMAO concentration, as shown in Fig. 6.3. Mass spectra of the samples of **6.9** after thermal melt experiments indicated that the depsipeptide had decomposed. In comparison, **6.8** showed minimal decomposition after thermal melt experiments. No further experiments were carried out with depsipeptide **6.9** due to its propensity for decomposition; however, according to the fractionation factor experiments cited above (Danielson & Raines, 2000), the interstrand

hydrogen bond strengths for (Pro-Pro-Gly)₁₀ and (Pro-Hyp-Gly)₁₀ are similar. Hence, dipeptide isosteres based on Pro-Gly should be a good approximation for those based on Hyp-Gly.

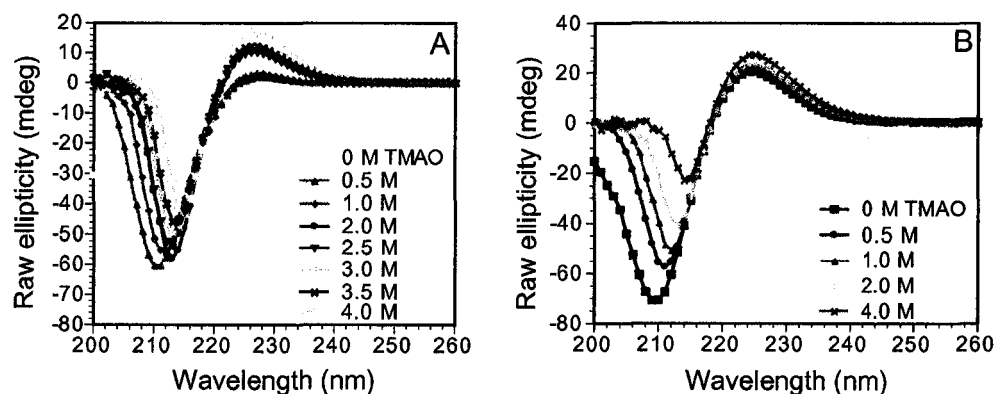


Figure 6.2 CD spectra of depsipeptides **6.8** (A) and **6.9** (B) with various concentrations of TMAO.

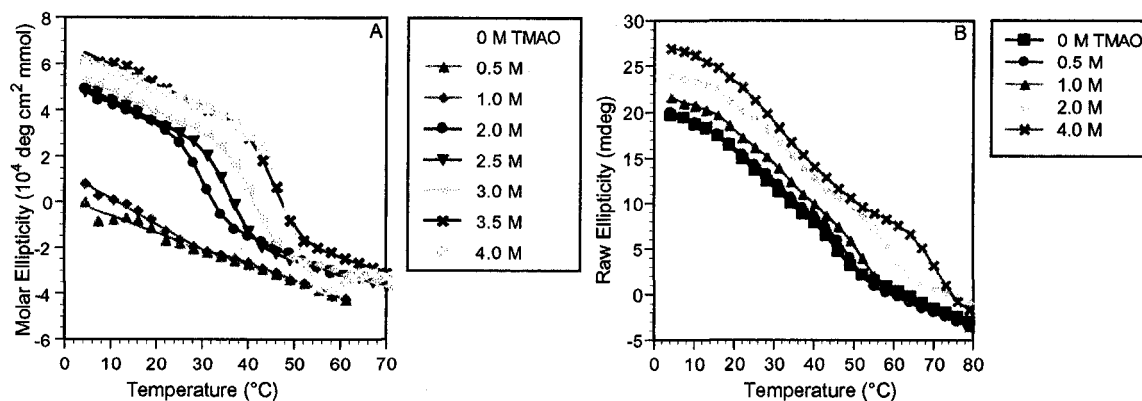


Figure 6.3 T_m determinations of triple helical **6.8** (A) and **6.9** (B) with various concentrations of TMAO.

The T_m values of triple helical **6.8** (above a threshold concentration of 2 M TMAO) and (Pro-Pro-Gly)₁₀ increased linearly with TMAO concentration, allowing us to extrapolate to obtain T_m values at 0 M TMAO for both peptides (Fig. 6.4). The extrapolated T_m value for (Pro-Pro-Gly)₁₀ was 32.8 °C, identical to the measured value of

33 °C in the absence of TMAO. The extrapolated T_m value for **6.8** was 10.7 °C, giving a ΔT_m of 22 °C. This value closely matches the ΔT_m at 2 M TMAO, which was observed to be 20 °C. Using eq. 6.1 with a ΔS value of 0.21 kcal/mol (Holmgren et al., 1999), this difference in melting temperature translates to a $\Delta\Delta G^\circ_{\text{obs}}$ value of 4.2 kcal/mol. This value, however, does not take into account the differences in solvation between amides and esters.

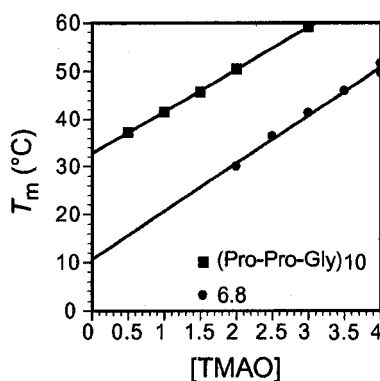


Figure 6.4 Extrapolation of T_m values for triple helical (Pro-Pro-Gly)₁₀ and **6.8** to 0 M TMAO.

Differences in solvation between amides and esters can be estimated by the $\Delta\Delta G^\circ$ of octanol-water partitioning for analogous small molecule amides and esters. For example, the $\Delta\Delta G^\circ(\text{octanol/water})$ for methyl acetate vs. *N*-methlyacetamide is -1.7 kcal/mol (Leo et al., 1971). Taking solvation effects into account is one way to correct for differences in electrostatics between amides and esters. Eq. 6.2 (F = folded, U = unfolded) (Koh et al., 1997) gives a method for incorporating this correction factor into the estimate of hydrogen bond strength. Thus, the $\Delta\Delta G^\circ$ for removal of one hydrogen bond donor from each strand of a triple helix is estimated to be ≥ -5.9 kcal/mol, therefore

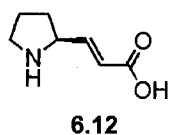
each hydrogen bond contributes approximately 2 kcal/mol to the stability of the triple helix.

$$\Delta\Delta G^\circ = \Delta T_m \Delta S \quad (6.1)$$

$$\Delta G^\circ(\text{F})_{\text{backbone H-bond}} - \Delta\Delta G^\circ(\text{F} \rightarrow \text{U})_{\text{electrostatics}} =$$

$$\Delta\Delta G^\circ_{\text{octanol} \rightarrow \text{water}} - \Delta\Delta G^\circ_{\text{obs}} \geq -1.7 \text{ kcal/mol} - 4.2 \text{ kcal/mol} = -5.9 \text{ kcal/mol} \quad (6.2)$$

The replacement of a Pro-Gly unit in (Pro-Pro-Gly)₁₀ with the dipeptide isostere **6.12** (also denoted Pro=Gly) yields peptide **6.13**, which lacks both a hydrogen bond donor and an acceptor. The Fmoc derivative of **6.12** was a generous gift from Prof. Scott J. Miller, and the synthesis of **6.13** by standard SPPS protocols went smoothly. Incubation of **6.13** for 24 h in 50 mM HOAc yielded no triple helix formation, as determined by CD wavelength scan and thermal melt experiments. The CD spectrum from 200-260 nm exhibited only a small maximum near 225 nm, similar to the signature spectra of a polyproline type II helix (Feng et al., 1996). The CD signal at 227 nm decreased linearly with temperature, confirming the observation that no significant amount of triple helix had formed.



6.13

Repeating the thermal melt experiments on the peptide **6.13** in increasing concentrations of TMAO revealed that no transition could be observed at less than 3.5 M TMAO (Fig 6.5). The T_m value of triple helical **6.13** in 3.5 M TMAO was 14.0 °C, and that in 4.0 M TMAO was 19.5 °C. Extrapolation of the T_m data to 0 M TMAO would indicate a melting temperature of -24.7 °C; however, with only two points on which to base the extrapolation, the error in the extrapolated T_m is likely to be very large. Furthermore, it is difficult to interpret this set of data in any quantitative manner, since the effects of removing an additional hydrogen bond acceptor and the conformational effects of Pro=Gly versus Pro-Gly on peptide backbone torsion angles φ and ψ (especially in the context of a collagen triple helix) are unknown. The conclusion that may be drawn from the replacement of a Pro-Gly segment by alkene isostere **6.12** is that the resulting triple helix is much less stable than when the amide linkage of a Pro-Gly unit is replaced by an ester. The most likely reason for the comparative decrease in the stability of triple helices containing the alkene isostere vs. the ester linkage may be more rigid enforcement of φ and ψ angles in the alkene isostere that are unfavorable for triple helix formation; however, a larger disruption in solvation at the alkene moiety than the ester may also play a role.

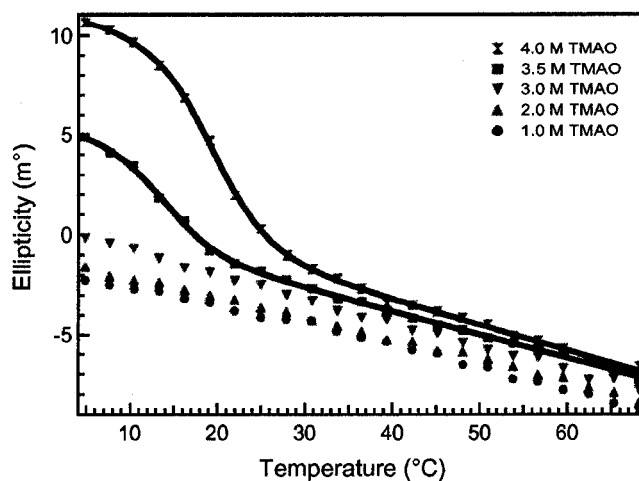
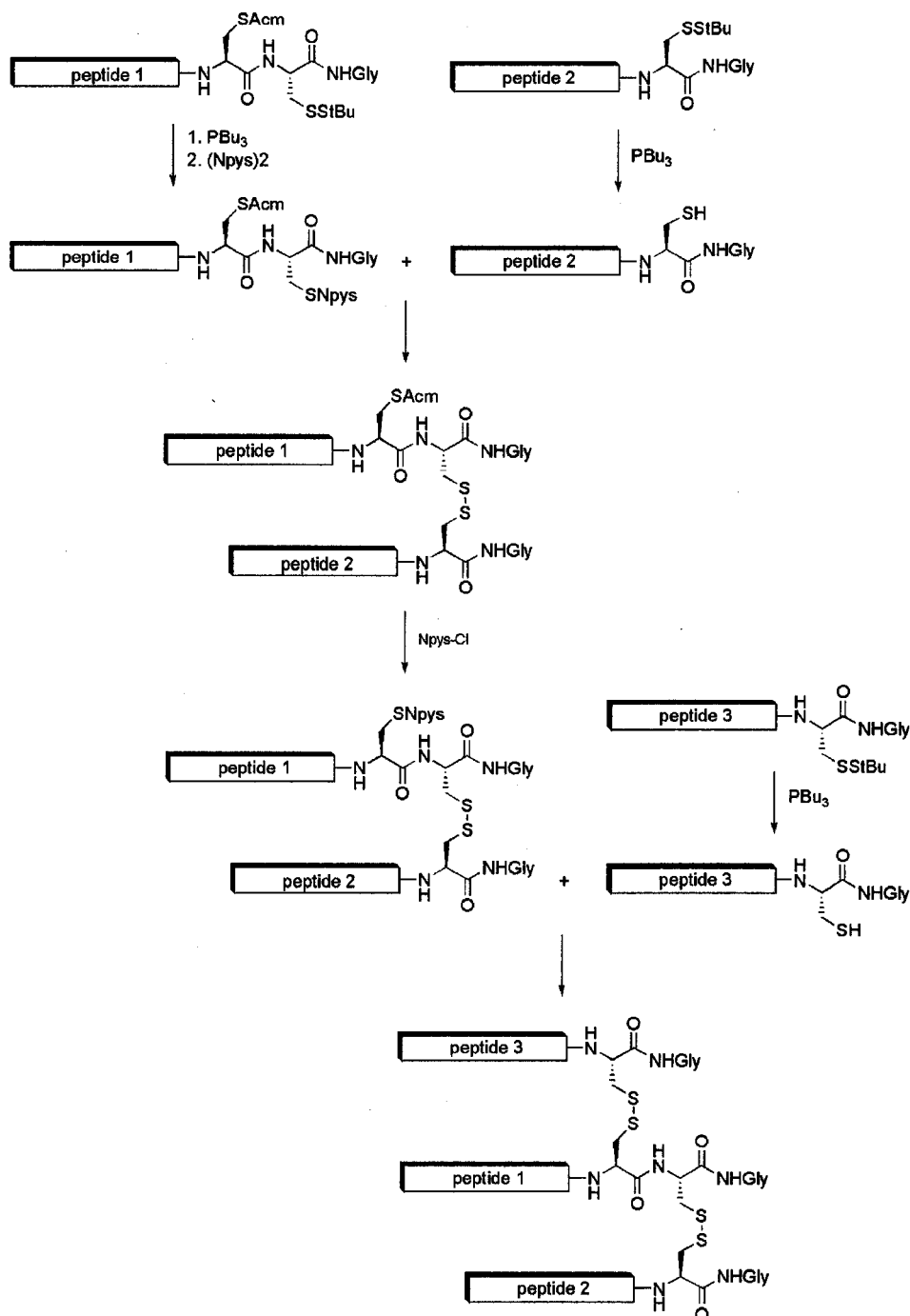


Figure 6.5 T_m determinations of triple helical 6.13 with various concentrations of TMAO.

6.2.4. Assembly of Heterotrimeric Collagen Mimics

To obtain a more direct measurement of the contribution of one interstrand hydrogen bond to triple helix stability, formation of a triple helix lacking only one hydrogen bond (rather than three) is necessary. The formation of such a heterotrimer requires the ability to link the strands together covalently and selectively. Moroder and coworkers have devised a method to do this by the selective formation of two disulfide bonds, creating a minimal “cystine knot” (Scheme 6.8) (Ottl et al., 1996; Ottl & Moroder, 1999a; Ottl et al., 1999). They have also shown that this process is difficult when peptides are used which have a high propensity to form stable triple helices at room temperature (e.g., peptides containing predominantly Pro-Hyp-Gly repeats) (Sacca et al., 2002).

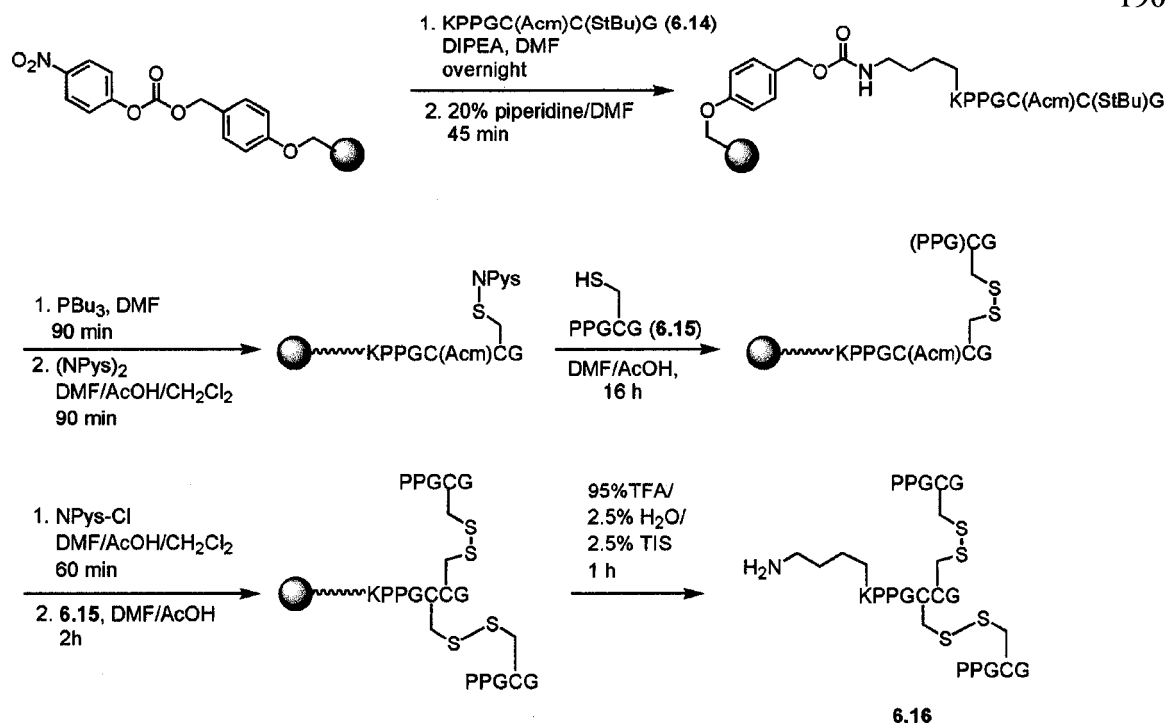


Scheme 6.8 Synthesis of a collagen mimic with a cystine knot (Ottl et al., 1996; Ottl & Moroder, 1999b; Ottl et al., 1999). NpysCl is 3-nitropyridine-2-sulfonyl chloride.

We desired a more efficient synthesis of heterotrimers using the cystine knot technology, and hypothesized that this might be accomplished by doing the assembly on

a solid support, as outlined in Scheme 6.9. The advantages of doing the assembly on a solid support are twofold: (1) the deprotection and activation reactions can be carried out using large excesses of reagents without requiring subsequent HPLC purification of the activated peptide, and (2) the coupling reactions can be carried out with larger excesses of peptides without the difficult problem of removing any homodimer formed during the reactions, because it can simply be washed away from the heterodimer bound to the resin. Furthermore, the excess peptide can easily be recovered from the reaction mixture.

We tested this idea by forming a heterotrimer consisting of one Pro-Pro-Gly repeat per strand, as shown in Scheme 6.9. The protected constituent peptides **6.14** and **6.15** were synthesized by standard SPPS methods in excellent yields. Peptide **6.14** was attached via the N-terminal lysine residue to commercially available *p*-nitrophenylcarbonate-derivatized Wang resin. The Cys(StBu) residue was deprotected by reaction with PBU_3 and activated by reaction with 2,2'-dithiobis(5-nitropyridine) ($(\text{Npys})_2$) forming a labile mixed disulfide.(Rabanal et al., 1996) Deprotection of 1.5 eq. of peptide **6.15** with PBU_3 and subsequent reaction with activated **6.14** on-resin yielded the resin-bound dimer. Deprotection and concomitant activation of the Cys(Acm) residue on-resin with 3-nitro-2-pyridine sulfonyl chloride (Npys-Cl) was followed by reaction with another 1.5 equivalents of **6.15**. Subsequent cleavage of the product from the resin with TFA afforded the assembled trimer **6.16** in 55% crude yield, assuming 100% loading of **6.14** onto the resin.

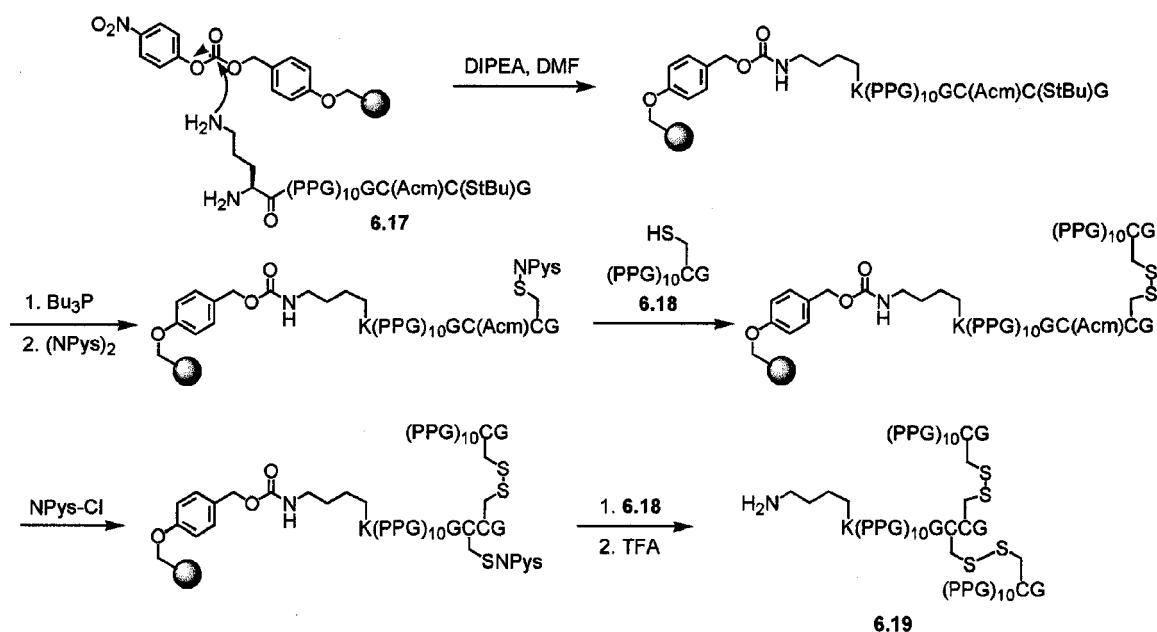


Scheme 6.9 Solid-phase assembly of a model cystine knot.

With the model trimer **6.16** in hand, we attempted to repeat the sequence with full-length peptides **6.17** and **6.18** to make trimer **6.19** (Scheme 6.10). Unfortunately, we found that it was difficult to achieve the coupling between the N-terminal Lys of **6.17** and the activated Wang resin. Repeated attempts at reliably effecting the loading of **6.17** via the Lys residue onto a variety of other *p*-nitrophenylcarbonate-activated resins (Novasyn TGATM, PEGATM, and CLEARTM resins) also failed.

The reason for the failure of peptide **6.17** to attach to a variety of resins is unclear. It is clear that the chemistry of forming a carbamate linkage between the N-terminal Lys residue and the activated carbonate moiety on the resin is not the main problem, as peptide **6.14** was able to be attached efficiently to Wang resin in this manner. It is likely that the size or the conformation of **6.17** is the barrier to loading. Xaa-Pro bonds have a

high propensity to adopt the cis-amide conformation (Stewart et al., 1990; Weiss et al., 1998; Jabs et al., 1999), which would increase the effective size of the peptide, decreasing its accessibility to the active sites in the pores of the resin. In addition, it is also possible that the peptides are assembling into triple helices in solution, which would also create a barrier for access of the Lys residue to the resin.



Scheme 6.10 Proposed on-resin assembly of a cystine-knot-tethered collagen mimic.

6.3. Conclusions and Future Directions

Assembly of heterotrimeric collagen mimics on solid phase remains an important goal. Encouraging preliminary results have been obtained by synthesizing the peptide (Pro-Pro-Gly)₁₀-Gly-Cys(Acm)-Cys(StBu)G (**6.20**) on PEGA resin and performing the deprotection, activation, and coupling reactions shown in Scheme 6.10 without cleaving and purifying **6.20** and reattaching it to the resin (F. W. Kotch, unpublished results).

Purification of the final full-length trimer from byproducts still remains as a significant challenge.

Once an effective synthesis of disulfide-linked heterotrimers is achieved, measurement of melting temperatures by CD spectroscopy will provide an avenue to measuring the contribution of interstrand hydrogen bonds to collagen triple helix stability. Another important approach to this determination will be to measure the melting temperatures of the tethered triple helices by differential scanning calorimetry (DSC), which will provide more direct access to thermodynamic values associated with triple helix denaturation than CD determination of melting temperatures can provide.

The analysis of homotrimeric collagen mimics has provided some insight into the contribution of interstrand hydrogen bonds to triple helix stability. One estimate of the strength of this hydrogen bond is ~ 2 kcal/mol, well within the accepted range for hydrogen bonds in proteins. The study of heterotrimeric collagen mimics containing amide bond isosteres will provide further insight into the contribution of interstrand hydrogen bonds to collagen stability because it will provide a way to minimize the structural perturbations that may contribute to the observed decrease in stability of the depsipeptide- and alkene isostere-containing collagen mimics we have studied.

6.4. Experimental Section

General. Reagents were obtained from Aldrich Chemical (Milwaukee, WI) or Fisher Scientific (Hanover Park, IL) and used without further purification. Amino acids and their derivatives were obtained from Fisher Scientific, Bachem Bioscience (King of Prussia, PA), or Novabiochem (San Diego, CA). Dichloromethane was distilled over

CaH₂(s) or drawn from a Baker Cycletainer. Thin-layer chromatography was performed by using aluminum-backed plates coated with silica gel containing F₂₅₄ phosphor and visualized by UV illumination or staining with I₂, *p*-anisaldehyde stain, or phosphomolybdic acid stain. NMR spectra were obtained with Bruker AC-300 and Varian UNITY-500 spectrometers.

***N*-Benzyloxycarbonyl-prolylproline (6.1a).** Benzyltrimethylammonium hydroxide (Triton B, 20.2 mL, 40% w/v in MeOH, 48.4 mmol) was added to proline (5.485 g, 47.6 mmol), and additional MeOH was added to effect complete dissolution of the proline. The mixture was concentrated under reduced pressure to remove the methanol and then placed under high vacuum for 1 h to obtain a waxy solid. The solid was dissolved in dry DMF (80 mL) and then added to a solution of *N*-benzyloxycarbonyl prolinoxy succinimide (Z-Pro-OSu, 15.053 g, 43.5 mmol) in DMF (100 mL). The reaction mixture was stirred overnight at room temperature. The solvent was removed under reduced pressure, and the residue was dissolved in 5% (w/v) KHCO₃ (aq) and washed twice with diethyl ether. The aqueous layer was placed under aspirator vacuum for 1 h to remove residual ether, then acidified with 6 N HCl to pH 2. The precipitate was collected by filtration, washed with water, and air-dried overnight, giving **6.1a** (12.962 g, 89%), which was carried on without further purification. ¹H NMR (300 MHz, CDCl₃+CD₃OD, two rotamers): δ 7.40-7.27 (m, 5H), 5.12, 5.03 (two ABq, *J* = 12.3, 11.7 Hz, 2H), 4.58-4.30 (m, 3H), 3.84-3.33 (m, 4H), 2.30-1.67 (m, 8H). ¹³C NMR (75 MHz, CDCl₃+CD₃OD, two rotamers): δ 173.6, 173.4, 171.7, 171.4, 155.0, 154.2, 136.3, 136.1, 128.2, 128.0, 127.9, 127.8, 127.5, 67.2, 66.9, 58.8, 58.7, 58.0, 57.4, 47.0, 46.6, 46.5, 29.8, 28.8, 24.7, 24.0, 23.4. ESI-HRMS: *m/z* 345.1434 ([*M*-H]⁺); 345.1451 ([*M*-H]⁺, calcd).

***N*-Benzyloxycarbonyl-prolylprolylglycine benzyl ester (6.2a). *N*-**

Benzyloxycarbonyl-prolylproline (**6.1a**, 8.003 g, 23.1 mmol), glycine benzyl ester tosylate salt (10.158g, 30.1 mmol), PyBOP (15.226 g, 29.3 mmol), and hydroxybenzotriazole hydrate (4.401 g, 28.7 mmol) were dissolved in dry DMF (250 mL). Diisopropylethylamine (15.0 mL, 86.1 mmol) was added and the mixture was stirred at room temperature for 24 h. The solvent was removed under reduced pressure, the resulting residue was dissolved in EtOAc, and this solution was washed twice each with 5% (w/v) KHSO₄ and 5% (w/v) KHCO₃, and once each with water and brine. The organic layer was dried over MgSO₄(s), filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with 5% (v/v) MeOH/CH₂Cl₂. Some of the fractions contained residual impurities, and these fractions were pooled and repurified by chromatography, eluting with EtOAc. The fractions (from both purifications) containing pure product were pooled and concentrated to yield **6.2a** (9.135 g, 80%) as a white solid. ¹H NMR (400 MHz, CDCl₃, two rotamers): δ 7.39-7.24 (m, 10H), 5.18-4.94 (m, 4H), 4.66-3.29 (m, 8H), 2.56-1.46 (m, 8H).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolylprolylglycine (6.3a). (Method A)**

To **6.2a** (9.135 g, 18.5 mmol) under an Ar(g) atmosphere was added tosic acid monohydrate (4.405 g, 23.2 mmol) and Pd/C (10 wt % Pd, 0.946 g), followed by a solution of MeOH (140 mL) and water (130 mL). The flask was flushed with Ar(g), H₂(g) was introduced via a balloon, and the reaction was stirred overnight. The white solid did not all dissolve initially, so MeOH (60 mL) was added, followed by more H₂(g) via balloon. The reaction mixture was again stirred overnight, filtered through Celite, and concentrated under reduced pressure. To a solution of the crude product in dioxane (100

mL) and water (100 mL) was added 9-fluorenylmethyloxycarbonyloxy succinimide (FmocOSu, 7.508 g, 22.3 mmol). Additional dioxane was added until the FmocOSu dissolved, followed by $\text{NaHCO}_3(\text{s})$ (3.125 g, 37.2 mmol). The reaction mixture was stirred for 3 d, and concentrated under reduced pressure to remove the bulk of the dioxane. Aqueous KHCO_3 (2.5% w/v, 150 mL) was added and the resulting solution was washed twice with Et_2O and twice with EtOAc. The aqueous layer was placed under aspirator vacuum for 1 h to remove residual organic solvents, acidified with 6 N HCl to pH 3, and allowed to stir for 1 h. The oil that initially formed became solid, and was collected by filtration and rinsed with water. A second crop of product was obtained by re-acidifying the aqueous layer. The combined solids were air-dried overnight to give **6.3a** (3.417 g, 38 %) as a white solid. ^1H NMR (300 MHz, CDCl_3): δ 7.74-2.8 (m, 8H), 4.62-2.95 (m, 11H), 2.20-1.72 (m, 8H). ESI-MS: m/z 492.4 ($[\text{M}+\text{H}]^+$); 492.5 ($[\text{M}+\text{H}]^+$, calcd).

t-Butyloxycarbonyl-prolylglycine benzyl ester (6.4). t-Butyloxycarbonyl-proline (5.268 g, 24.5 mmol), glycine benzyl ester tosylate salt (9.088 g, 26.9 mmol), dicyclohexylcarbodiimide (5.075 g, 24.5 mmol), and hydroxybenzotriazole hydrate (3.756 g, 24.5 mmol) were dissolved in dry DMF (200 mL) and diisopropylethylamine (12.8 mL, 73.5 mmol) was added. The reaction mixture was stirred overnight, filtered to remove DCU, and concentrated under reduced pressure. The residue was dissolved in EtOAc and washed twice each with 5% (w/v) KHCO_3 and 5% (w/v) KHSO_4 , and once with brine. The organic layer was dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with 1:1 hexane/EtOAc to yield **6.4** (6.987 g, 79%) as a white solid. ^1H NMR (300 MHz,

CDCl₃): δ 7.39-7.28 (m, 5H), 5.17 (ABq, J = 12.3 Hz, 2H), 4.36-3.99 (m, 3H), 3.50-3.30 (m, 2H), 2.36-1.82 (m, 4H), 1.46 (s, 9H).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolylprolylglycine benzyl ester (6.5).**

BocProGlyOBn (**6.4**, 4.259 g, 11.8 mmol) was dissolved in dioxane (50 mL) and placed under Ar(g) atmosphere. 4 N HCl in dioxane (29 mL, 116 mmol) was added via syringe. The reaction mixture was stirred for 1.5 h, and then concentrated to dryness. The residue was dissolved in dry DMF (80 mL), and diisopropylethylamine (4.2 mL, 24.1 mmol) was added, followed by FmocPro pentafluorophenyl ester (5.934 g, 11.8 mmol). Additional DMF (20 mL) was added, then the reaction mixture was stirred overnight and then concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with EtOAc, to yield **6.5** (5.392 g, 79%) as a white solid. ¹H NMR (300 MHz, CDCl₃): δ 7.77-7.17 (m, 13H), 5.19-4.94 (m, 2H), 4.73-3.02 (m, 11H), 2.57-1.61 (m, 8H). ¹³C NMR (125 MHz, DMSO-d₆, two rotamers): δ 171.85, 171.75, 171.2, 170.6, 170.1, 170.0, 153.74, 153.67, 144.0, 143.8, 140.7, 140.6, 127.65, 127.57, 127.2, 127.0, 125.1, 124.81, 124.76, 120.1, 120.0, 66.5, 66.1, 59.1, 59.0, 57.8, 57.3, 46.9, 46.6, 46.3, 46.1, 29.5, 28.91, 28.87, 28.6, 24.33, 24.26, 23.7, 22.6.

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolylprolylglycine (6.3a). (Method B)**

Pd/C (10 wt % Pd, 0.372 g) was added to solid **6.5** (3.692 g, 6.35 mmol) under an Ar(g) atmosphere. MeOH (100 mL) was added and the flask was flushed well with Ar(g). H₂(g) was added via a balloon and the reaction mixture was stirred for 4.75 h, filtered through Celite, and concentrated under reduced pressure. The resulting solid was triturated overnight with EtOAc/Et₂O and then filtered and air-dried to yield **6.3a** (2.7 g, 86%) as a white solid. ¹H NMR (300 MHz, CD₃OD): δ 7.82-7.28 (m, 8H), 4.69-3.35 (m, 11H),

2.29-1.74 (m, 8H). ^{13}C NMR (125 MHz, DMSO- d_6 , two rotamers): δ 171.84, 171.76, 170.1, 170.0, 153.8, 153.7, 144.0, 143.8, 143.7, 140.7, 140.6, 127.64, 127.57, 127.1, 127.0, 125.1, 124.81, 124.76, 120.1, 120.0, 66.5, 66.0, 59.2, 59.0, 57.8, 57.3, 46.9, 46.6, 46.3, 46.1, 29.5, 28.91, 28.87, 28.6, 24.33, 24.26, 23.7, 22.6. ESI-HRMS: m/z 490.1965 ($[\text{M}-\text{H}]^-$); 490.1978 ($[\text{M}-\text{H}]^-$, calcd).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolylprolylglycolate (6.6)** Compound 6.6 was prepared according to the method of Moroder and Ottil (Method A) (Ottil et al., 1999). ^1H NMR (300 MHz, DMSO- d_6): δ 7.87-7.27 (m, 8H), 4.68-4.10 (m, 8H), 3.73-3.09 (m, 4H), 2.25-1.73 (m, 8H). ESI-MS: m/z 493.4 ($[\text{M}+\text{H}]^+$); 493.5 ($[\text{M}+\text{H}]^+$, calcd).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolyl-4(*R*)-hydroxyprolylglycolate (6.7)** Compound 6.7 was prepared according to the method of Moroder and Ottil (Method A) (Ottil et al., 1999). ^1H NMR (300 MHz, DMSO- d_6): δ 7.85-3.0 (m, 8H), 4.62-4.20 (m, 8H), 3.70-3.35 (m, 4H), 2.30-1.75 (m, 8H). ESI-MS: m/z 509.2 ($[\text{M}+\text{H}]^+$); 509.2 ($[\text{M}+\text{H}]^+$, calcd).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-glycylproline.** Glycylproline (1.022 g, 5.94 mmol) was dissolved in $\text{Na}_2\text{CO}_3(\text{aq})$ (10% (w/v), 20 mL), and a solution of Fmoc-Cl (1.696 g, 6.56 mmol) in dioxane (15 mL) was added. Additional dioxane (30 mL) was added and the reaction mixture was stirred for 3 h. Water (100 mL) was added and the reaction mixture was washed twice with Et_2O . The aqueous layer was acidified with 6 N HCl to pH 2 and extracted four times with EtOAc. The combined EtOAc extracts were dried over $\text{MgSO}_4(\text{s})$, filtered, concentrated under reduced pressure, and placed under high vacuum overnight to yield Fmoc-Gly-Pro (2.255 g, 96%) as a white solid. ^1H NMR (300 MHz, CDCl_3): δ 9.84 (bs, 1H), 7.76-7.26 (m, 8H), 6.19, 5.93 (two broad m, 1H),

4.60-3.41 (m, 8H), 2.31-1.87 (m, 4H). ^{13}C NMR (75 MHz, CDCl_3): δ 174.0, 168.5, 156.4, 145.5, 143.8, 141.2, 127.7, 127.0, 125.2, 119.9, 67.6, 67.3, 61.4, 59.3, 47.0, 46.3, 43.3, 28.5, 24.6. ESI-MS: m/z 393.1 ($[\text{M}-\text{H}]^-$); 393.2 ($[\text{M}-\text{H}]^-$, calcd).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-prolylglycolate (6.10). 9-**

Fluorenylmethoxycarbonylproline (FmocPro, 0.515 g, 1.53 mmol) and hydroxybenzotriazole (0.275 g, 1.80 mmol) were dissolved in dry CH_2Cl_2 (25 mL). Dicyclohexylcarbodiimide (0.325 g, 1.58 mmol) was added, followed by benzyl glycolate (0.25 mL, 1.76 mmol) and diisopropylethylamine (0.55 mL, 3.16 mmol). The reaction was stirred overnight, filtered to remove DCU, and concentrated under reduced pressure. The resulting yellow oil was dissolved in EtOAc, washed twice each with 5% (w/v) KHCO_3 and 5% (w/v) KHSO_4 , then dried over $\text{MgSO}_4(\text{s})$, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with 25% (v/v) EtOAc/hexanes to yield the benzyl ester (0.281 g, 38%) as a white foam. The benzyl ester (507 mg, 1.04 mmol) was combined with Pd/C (10 wt % Pd, 49 mg) under Ar(g) and MeOH (40 mL) was added. The flask was flushed with Ar(g) and $\text{H}_2(\text{g})$ was introduced via a balloon. The reaction mixture was stirred for 3 h, then filtered through Celite and concentrated under reduced pressure. The crude product was purified by silica gel chromatography, eluting with 5–10 % (v/v) MeOH/ CH_2Cl_2 to yield **6.10** (0.172 g, 42%) as a white solid. ^1H NMR (300 MHz, CDCl_3): δ 9.12 (bs, 1H), 7.75-7.25 (m, 8H), 4.73-4.14 (m, 5H), 3.64-3.41 (m, 2H), 2.27-1.81 (m, 4H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers): δ 171.7, 170.7, 170.6, 155.2, 154.7, 143.9, 143.8, 143.6, 143.5, 141.2, 127.6, 127.0, 125.0, 124.9, 124.8, 124.7, 119.9, 67.7, 67.4, 60.7, 60.5, 58.9, 58.4,

47.0, 46.9, 46.5, 30.7, 29.7, 24.1, 23.1. ESI-HRMS: m/z 394.1273 ($[M-H]^+$); 394.1291 ($[M-H]^+$, calcd).

***N*-(9*H*-Fluoren-9-ylmethoxycarbonyl)-4(*R*)-*t*-butoxypropylglycolate (6.11)**

Compound **6.11** was synthesized by a method similar to that used for **6.10** to yield 215 mg (55%) of a white foam. ^1H NMR (300 MHz, CDCl_3) δ 7.78-7.28 (m, 8H), 4.73 (ABq, $J = 16.3$, 2H), 4.61-4.15 (m, 5H), 3.76 (app. dt, $J = 10.3$, 6.0 Hz, 1H), 3.34 (app dt, $J = 11.2$, 5.1 Hz, 1H), 2.34-2.20 (m, 2H), 1.21, 1.16 (two s, 9H). ^{13}C NMR (75 MHz, CDCl_3 , two rotamers) δ 171.7, 170.9, 155.3, 154.4, 143.8, 143.7, 141.3, 127.7, 127.0, 125.1, 124.9, 124.8, 119.5, 74.4, 69.1, 68.2, 67.9, 67.3, 60.6, 60.2, 57.8, 57.3, 53.4, 53.3, 47.1, 37.5, 28.2. ESI-HRMS: m/z 466.1871 ($[M-H]^+$), 466.1866 ($[M-H]^+$, calcd).

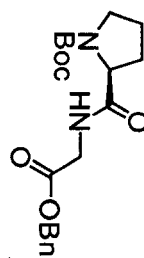
Solid phase peptide synthesis. Peptides were assembled on solid phase using an Applied Biosystems Model 432A (Synergy) peptide synthesizer using HBTU with HOBT and DIPEA as the coupling reagent, or an Applied Biosystems Pioneer instrument using HATU/DIPEA as the coupling reagent. Both instruments employ the Fmoc/*t*-Bu coupling strategy. It was found that the best resin for collagen mimic peptide synthesis using the Synergy instrument was 2-chlorotrityl resin, and the best resin for the Pioneer instrument was polyethyleneglycol-grafted polystyrene (PEG-PS) resin. Syntheses were carried out using 1.5–3 eq. of each segment to be coupled. Cleavage of the peptide from the resin was achieved using a mixture 95% TFA/2.5% H_2O /2.5% triisopropyl silane (TIS) for 1.5–3 h. Peptides were isolated by precipitation from the cleavage cocktail with *t*-butyl methyl ether, purified if necessary by preparative reverse-phase HPLC, and characterized by MALDI-TOF MS (Table 6.1).

Table 6.1 MALDI-TOF MS characterization of peptides produced by solid phase peptide synthesis.

Peptide	MALDI-TOF MS	
	Observed (M+H) ⁺	Calculated (M+H) ⁺
6.8	2530.9	2531.3
6.9	2691.2	2691.2
6.13	2513.8	2513.3
6.14	820.8	820.5
6.15 (StBu-protected)	518.9	518.4
6.16	1516.0	1515.6
6.17	3138.2	3137.7
6.18 (StBu-protected)	2778.7	2778.3

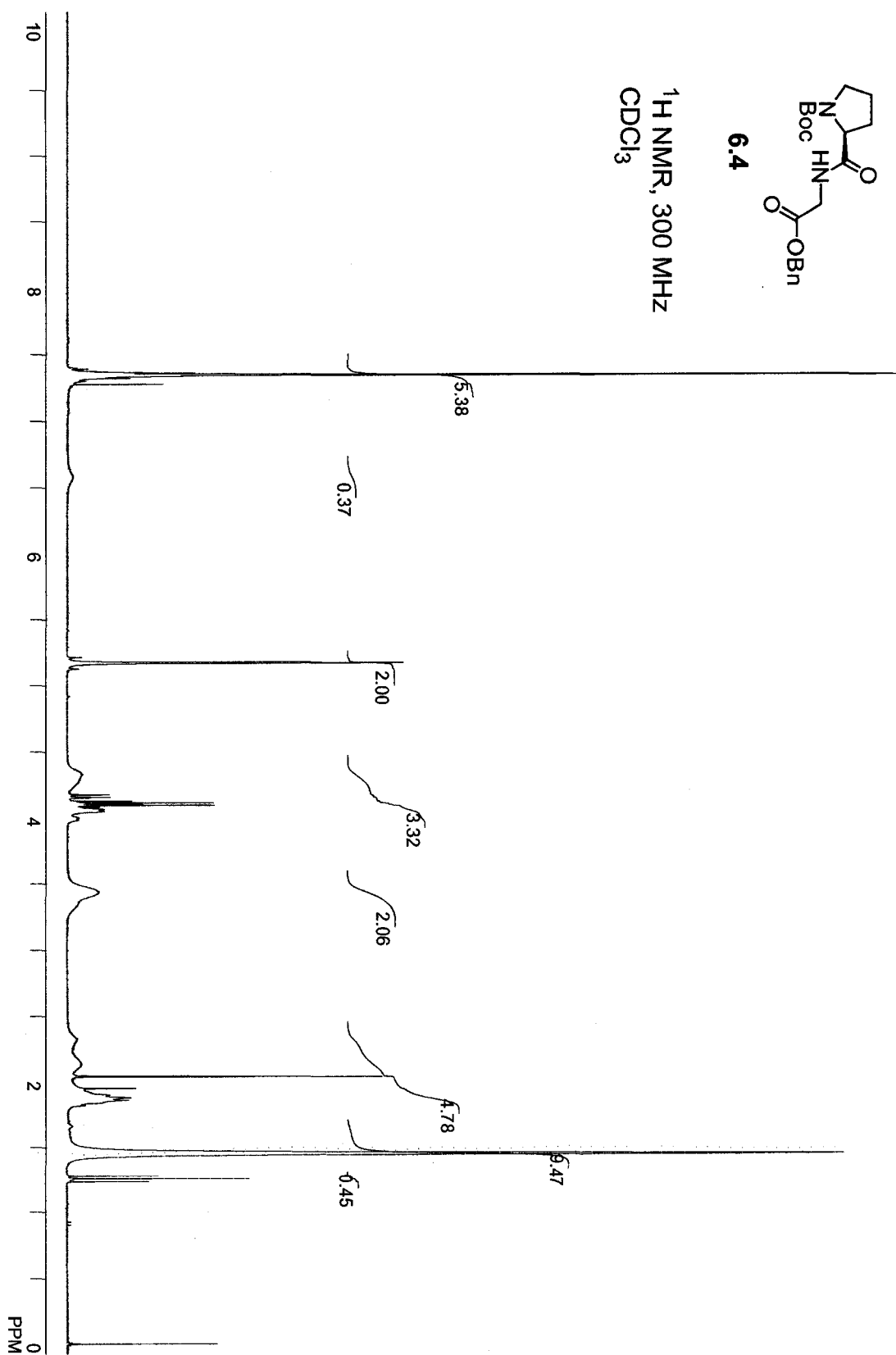
Circular Dichroism (CD) spectroscopy of depsipeptides 6.8 and 6.9. The depsipeptides were dissolved in buffer (either 50 mM HOAc, pH 3, or 0–4 M TMAO in phosphate-buffered saline, pH 7.4, 150 mM NaCl) to achieve a final concentration of ~0.2 mM, then were incubated at 4 °C for at least 24 h prior to study. Solutions for study were degassed on ice under vacuum for approximately 20 min immediately before placement into pre-cooled 1-mm-pathlength quartz cuvettes. Wavelength spectra were recorded at 4 °C from 260–200 nm, sampling every 1 nm and averaging for 3 s at each wavelength. Temperature scans were conducted from 4 °C to 60–80 °C (depending on the expected melting temperatures of the triple helices being studied), raising the temperature in 3-°C steps, equilibrating for 5 min at each temperature, and monitoring at 225 nm with a 20-s averaging time for each data point. Wavelength spectra were recorded again at the highest temperature used in the temperature scans using the same parameters as at 4 °C. Melting temperatures were determined by fitting the raw CD data using the software package NLREG 4.0 (Phillip Sherrod, 1998.)

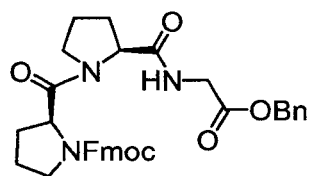
Circular Dichroism (CD) spectroscopy of peptide 6.13. Peptide 6.13 was dissolved in 50 mM HOAc (0.36 mM) or phosphate-buffered saline (PBS, pH 7.4, 0.22 mM) with various concentrations of TMAO (0-4 M) and the solutions was placed at 4 °C for >24 h. Samples were degassed under vacuum for 15 min at 0°C before performing CD measurements. Wavelength spectra were recorded from 200-260 nm at 4°C, sampling every 1 nm with a 3-s averaging time. Temperature scans were run by raising the temperature of the samples in 3 °C increments from 4–73 °C and equilibrating for 10 min at each temperature before taking a wavelength spectrum from 210-260 nm, sampling for 3 s every 3 nm. Wavelength spectra were also recorded at 73 °C, using the same parameters as for the wavelength scan at 4 °C.



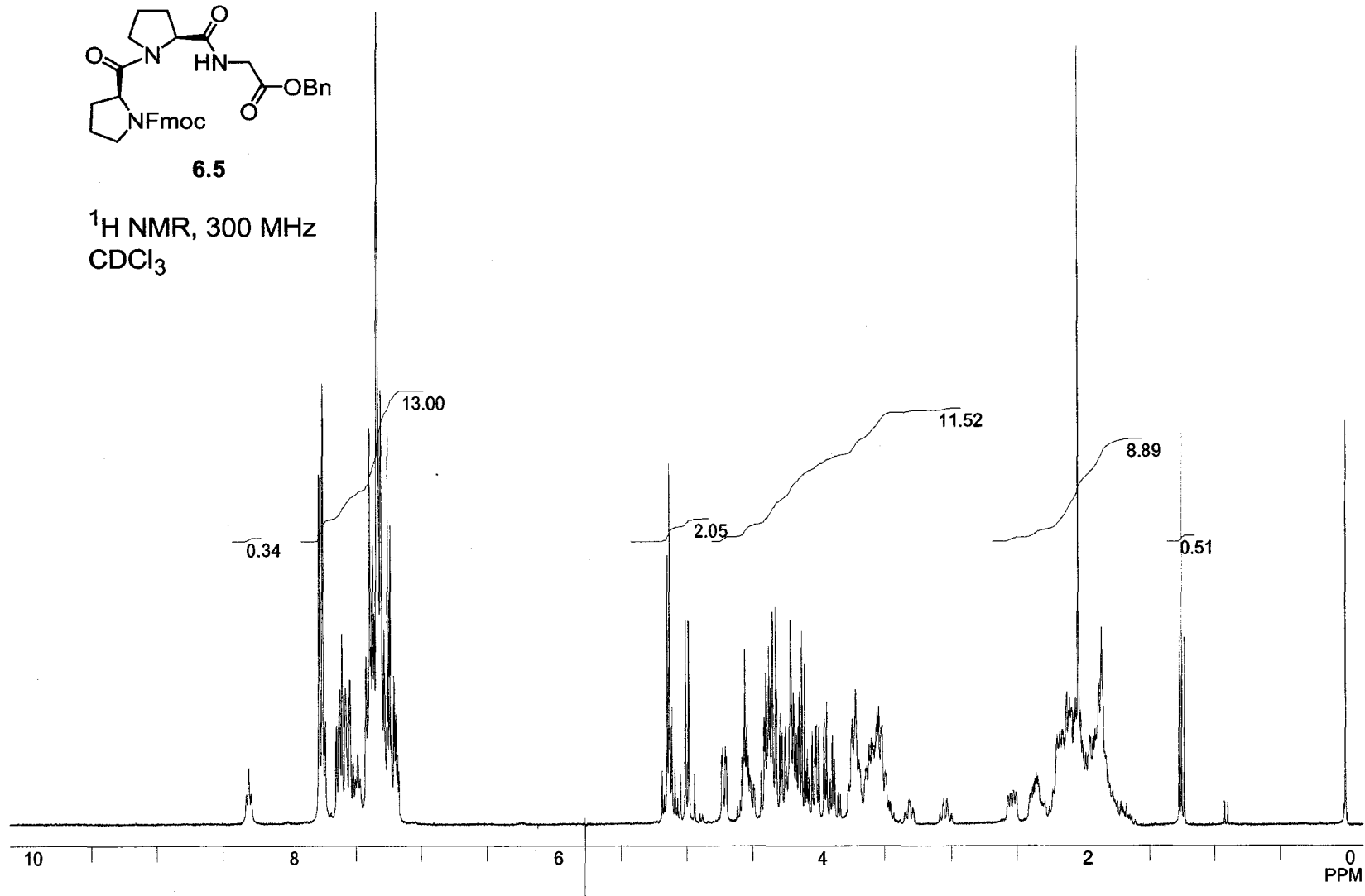
6.4

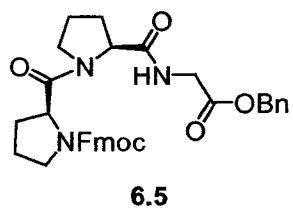
^1H NMR, 300 MHz
 CDCl_3



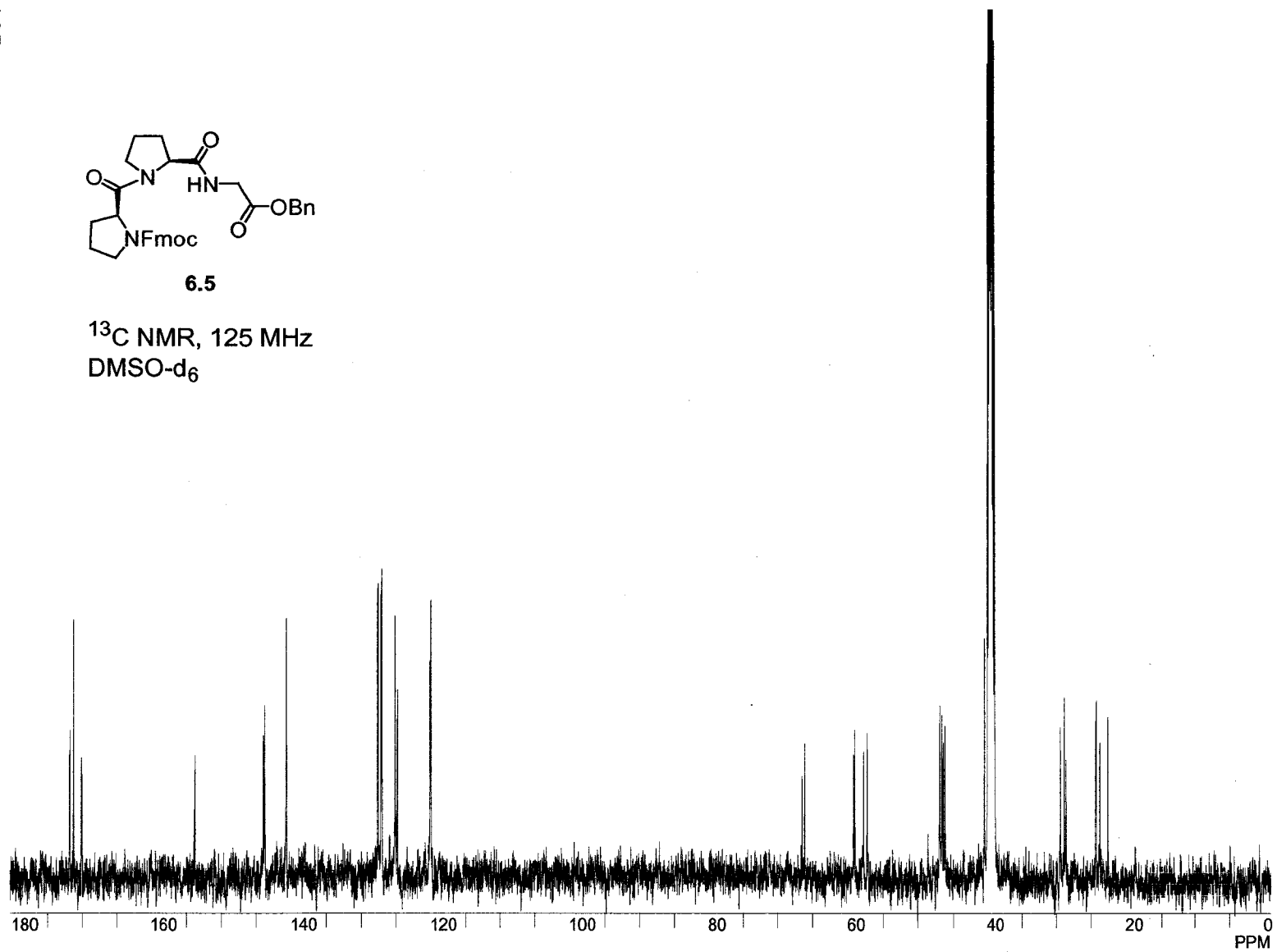
**6.5**

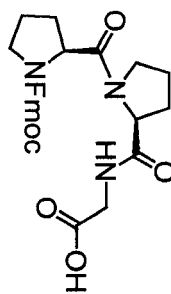
^1H NMR, 300 MHz
 CDCl_3



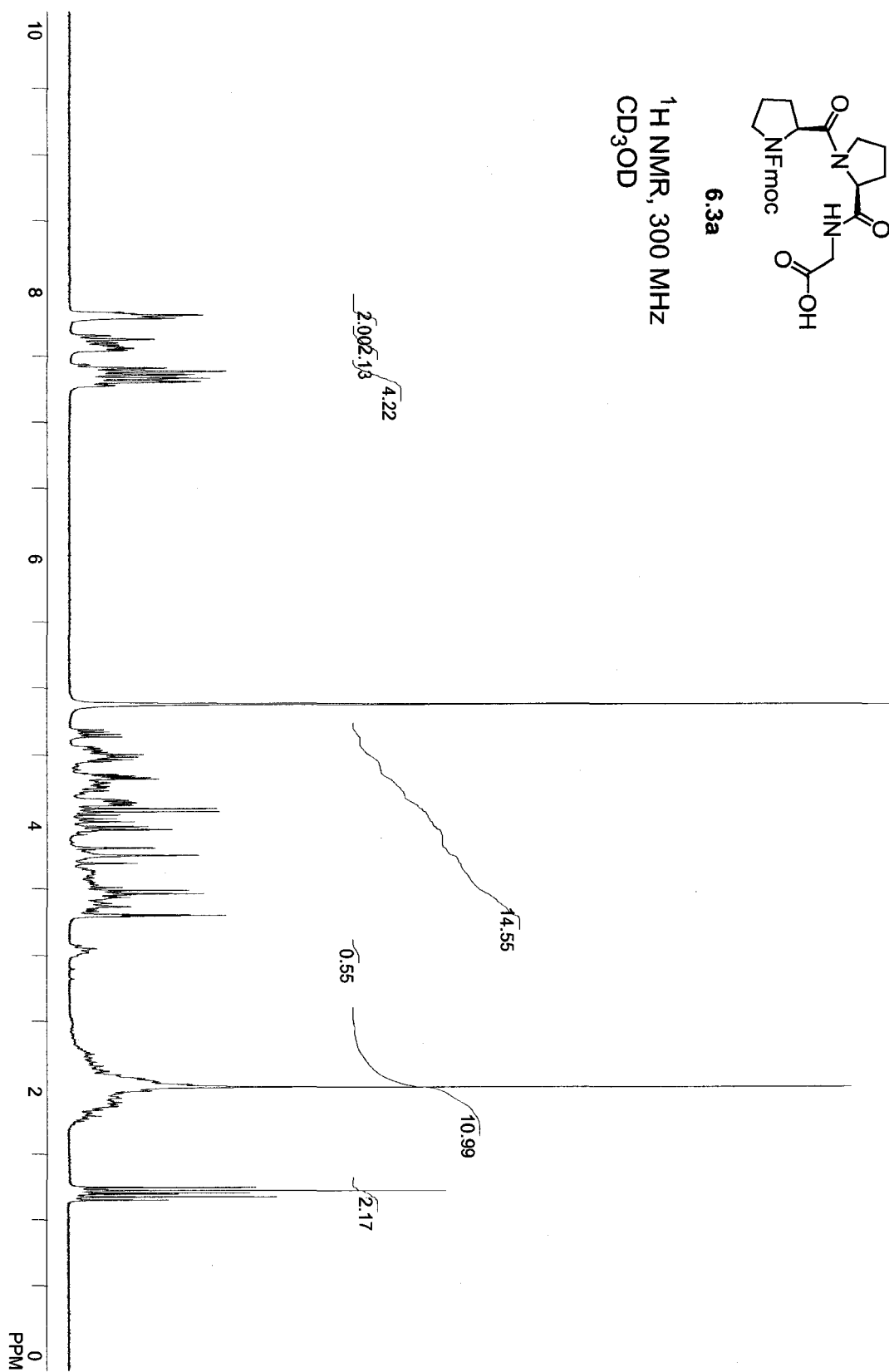


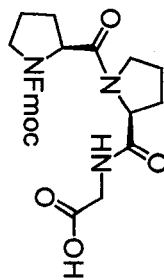
^{13}C NMR, 125 MHz
DMSO- d_6





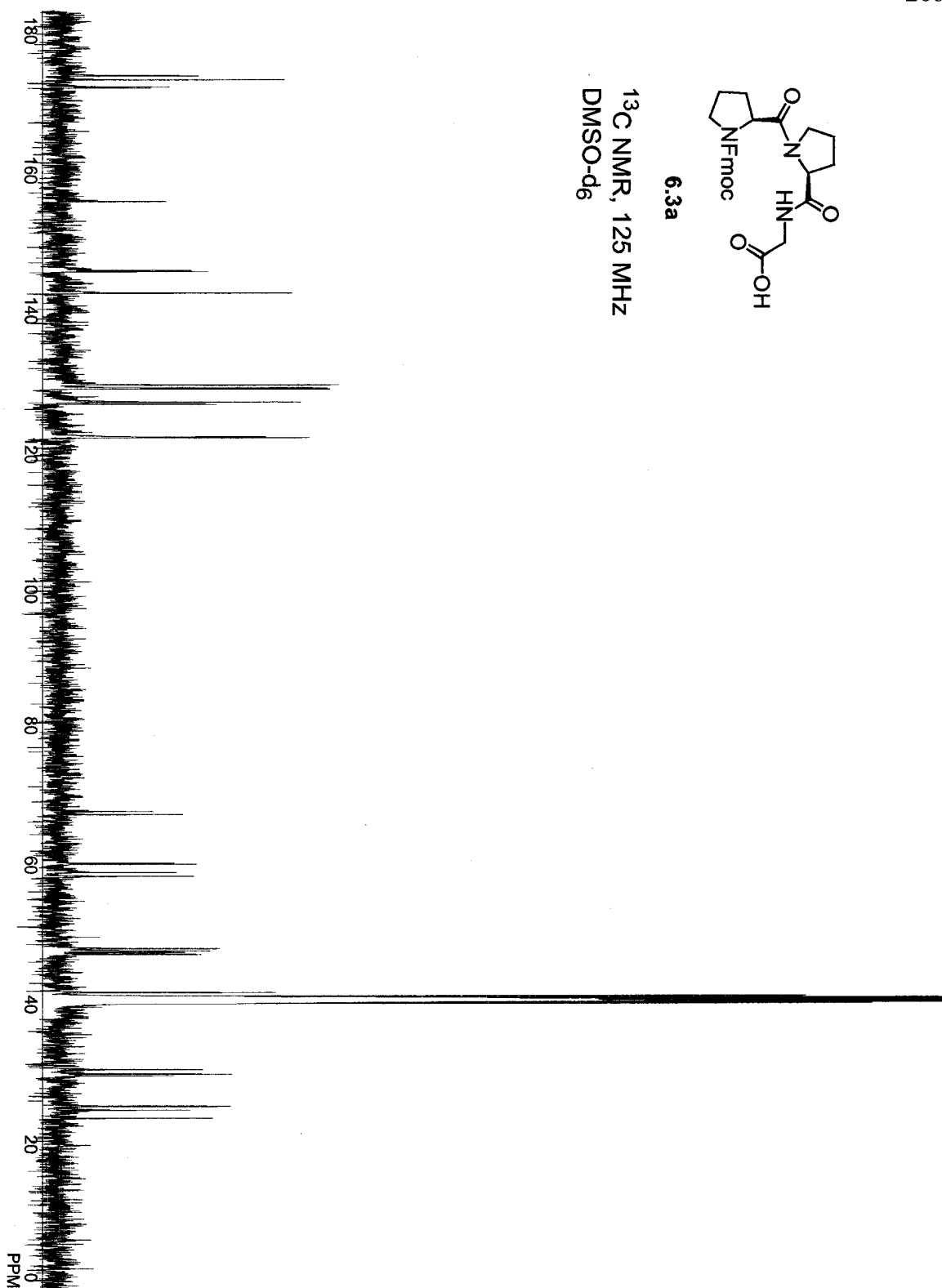
6.3a

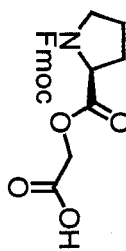
¹H NMR, 300 MHz
CD₃OD



6.3a

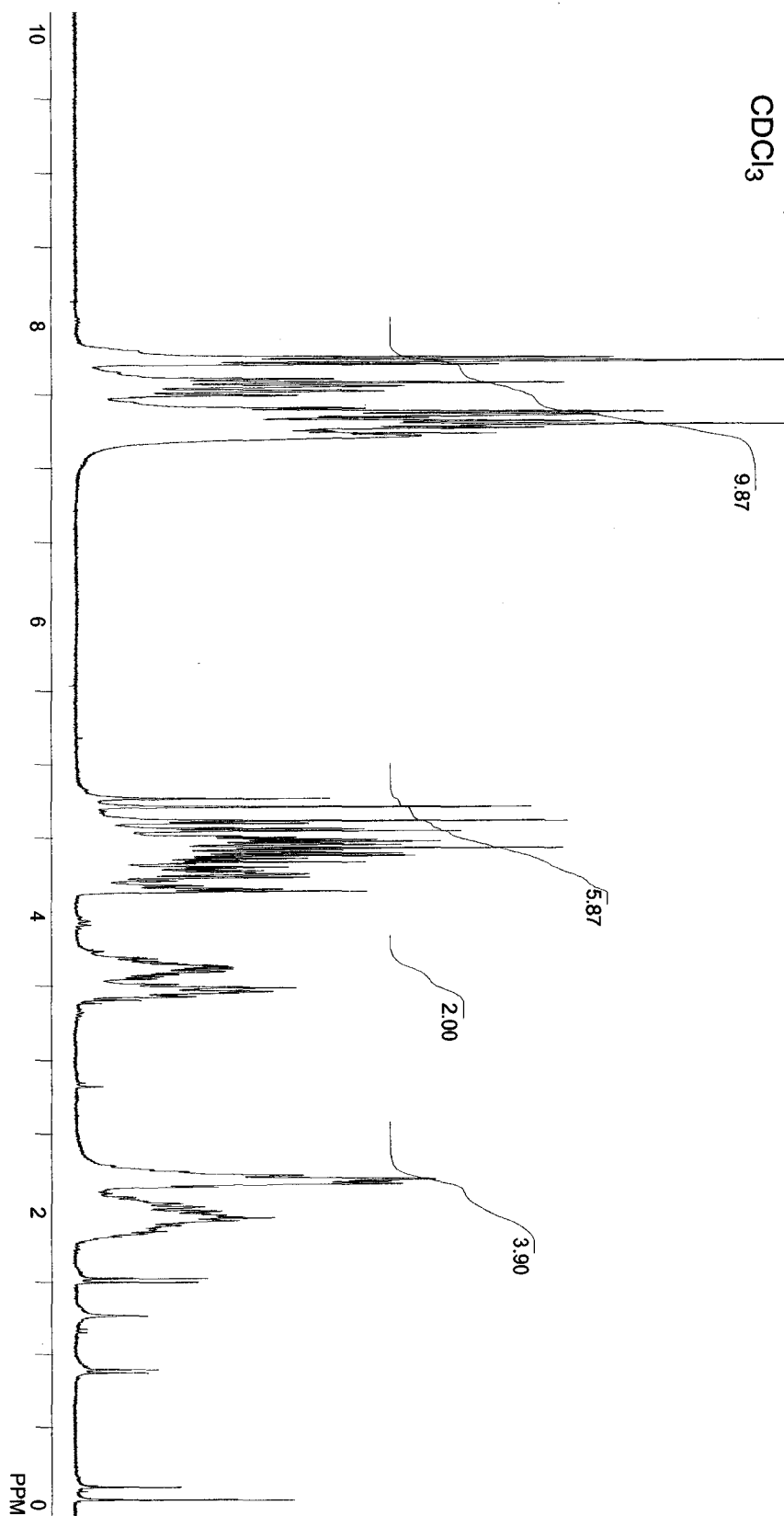
^{13}C NMR, 125 MHz
DMSO- d_6

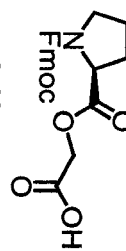




6.10

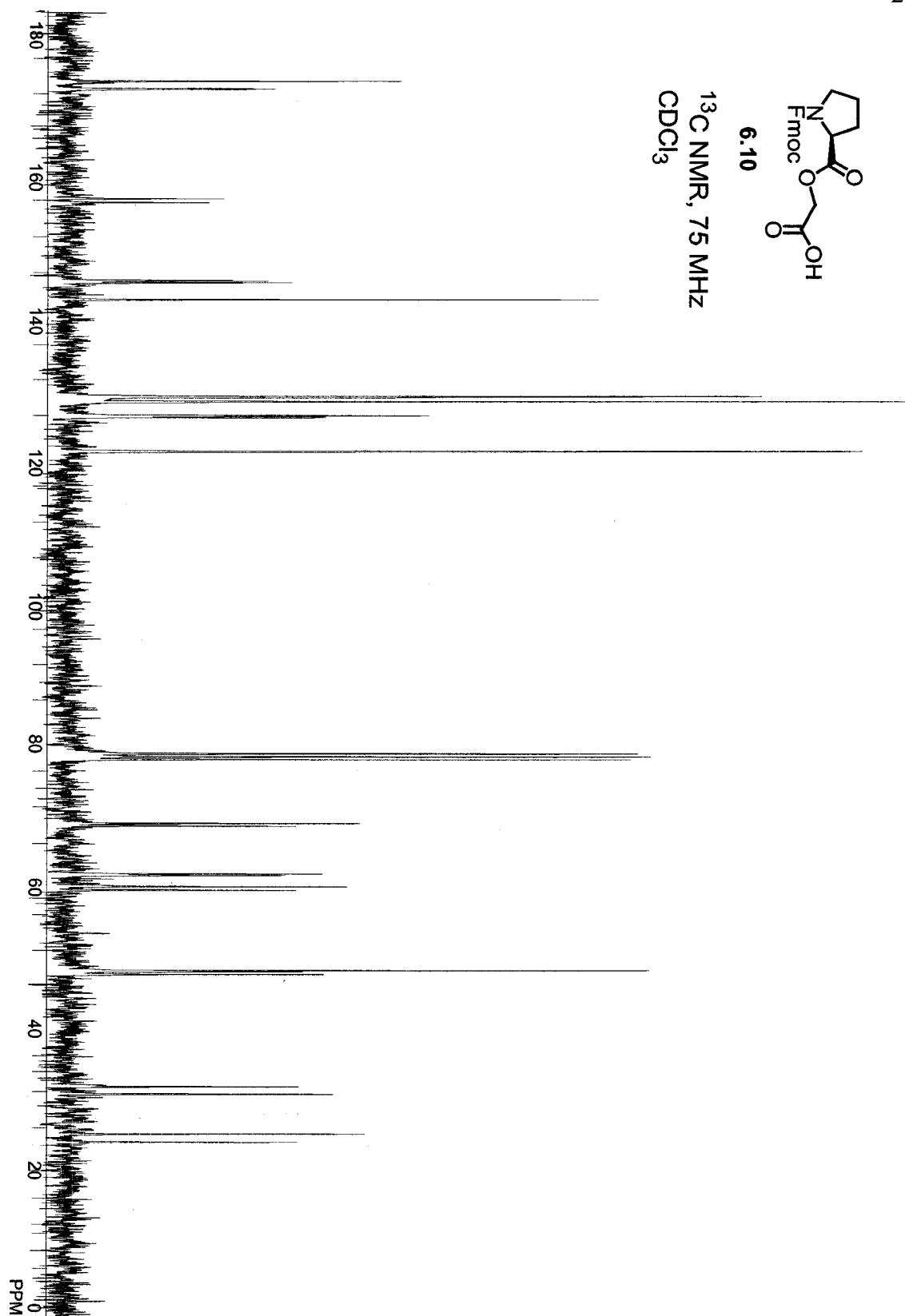
^1H NMR, 300 MHz
 CDCl_3

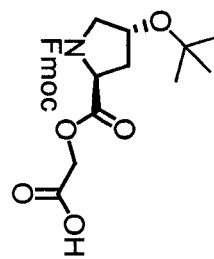




6.10

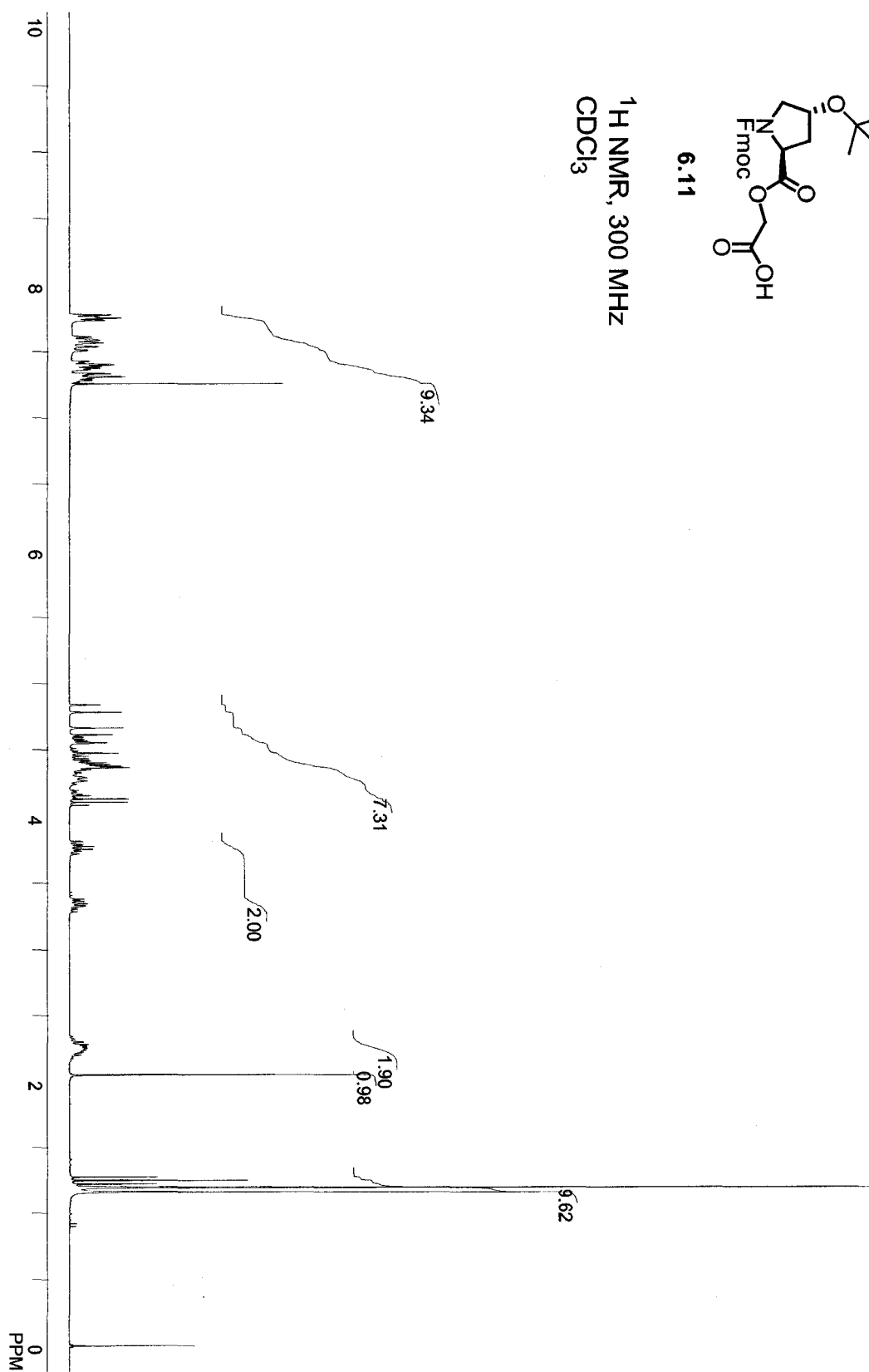
^{13}C NMR, 75 MHz
 CDCl_3

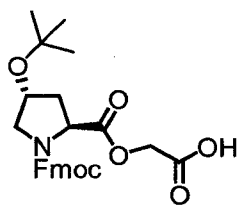




6.11

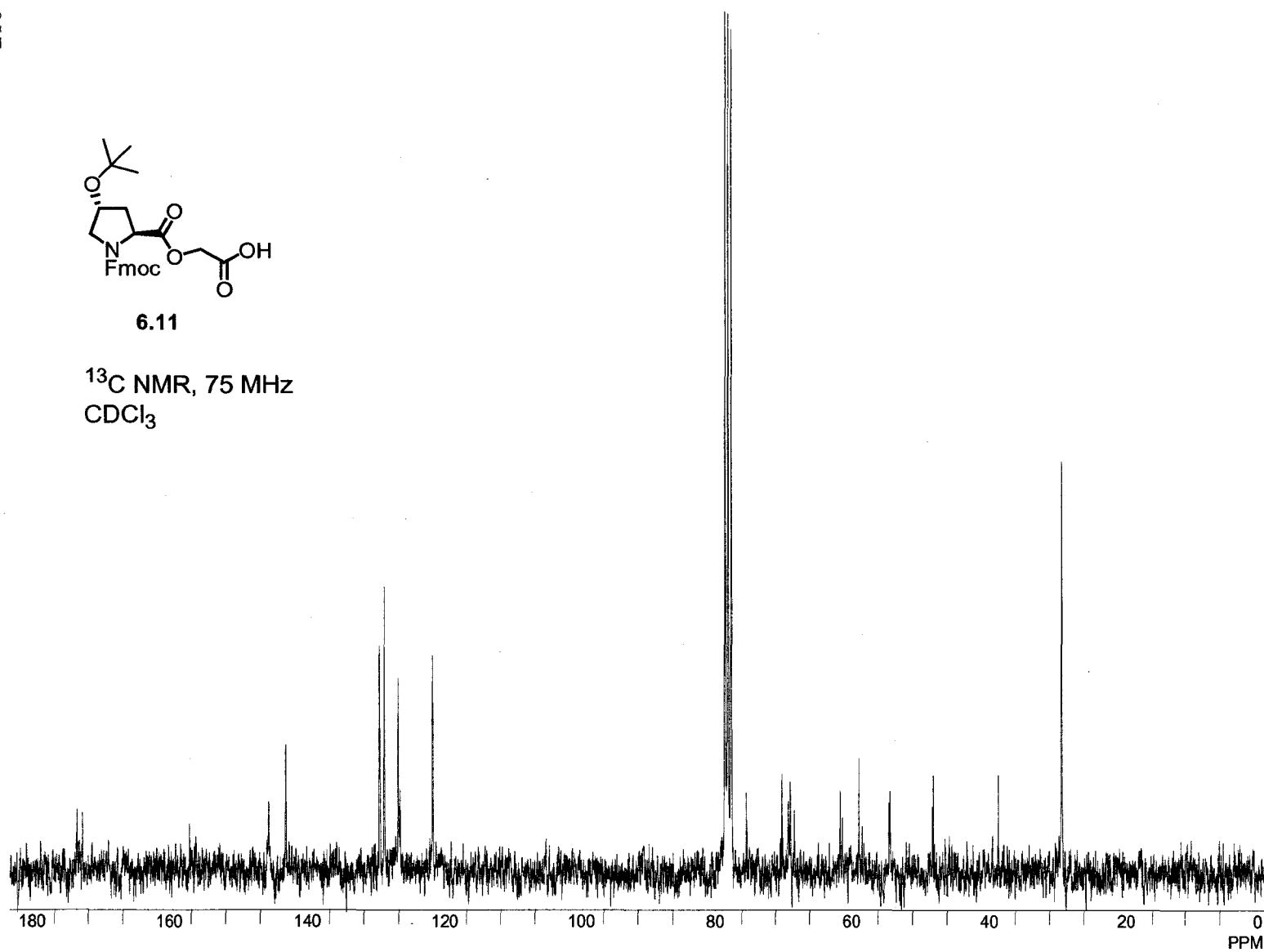
^1H NMR, 300 MHz
 CDCl_3





6.11

¹³C NMR, 75 MHz
CDCl₃



References

- An, S.-S. A., Lester, C. C., Peng, J.-L., Li, Y.-J., Rothwarf, D. M., Welker, E., Thannhauser, T. W., Zhang, L. S., Tam, J. P. & Sheraga, H. A. (1999). Retention of the Cis Proline Conformation in Tripeptide Fragments of Bovine Pancreatic Ribonuclease a Containing a Non-Natural Proline Analogue, 5,5-Dimethylproline. *J. Am. Chem. Soc.* **121**, 11558-11566.
- Antonov, I. V., Dudkin, S. M., Karpeiskii, M. Y. & Yakovlev, G. I. (1976). The Conformations of Phosphorylating Derivatives of 2'-Fluoro-2'-Deoxyuridine in Solution. *Sov. J. Bioorg. Chem.* **2**, 863-872.
- Aravinda, S., Shamala, N., Das, C. & Balaram, P. (2002). Structural Analysis of Peptide Helices Containing Centrally Positioned Lactic Acid Residues. *Biopolymers* **64**, 255-267.
- Arnold, U., Hinderaker, M. P., Koeditz, J., Golbik, R., Ulbrich-Hoffmann, R. & Raines, R. T. (2003). Protein Prosthesis: A Nonnatural Residue Accelerates Folding and Increases Stability. *J. Am. Chem. Soc.* **125**, 7500-7501.
- Avenoza, A., Busto, J. H., Peregrina, J. M. & Rodriguez, F. (2002). Incorporation of Aha into Model Dipeptides as an Inducer of a Beta-Turn with a Distorted Amide Bond. Conformational Analysis. *J. Org. Chem.* **67**, 4241-4349.
- Avenoza, A., Cativiela, C., Busto, J. H. & Peregrina, J. M. (1995). Exo-2-Phenyl-7-Azabicyclo[2.2.1]Heptane-1-Carboxylic Acid: A New Constrained Proline Analog. *Tetrahedron Lett.* **36**, 7123-7126.
- Babu, I. R. & Ganesh, K. N. (2001). Enhanced Triple Helix Stability of Collagen Peptides with 4R-Aminopropyl (Amp) Residues: Relative Roles of Electrostatic and Hydrogen Bonding Effects. *J. Am. Chem. Soc.* **123**, 2079-2080.
- Bakerman, S., Martin, R. L., Burgstahler, A. W. & Hayden, J. W. (1966). In Vivo Studies with Fluoroprolines. *Nature* **212**, 849-850.
- Bann, J. G. & Bächinger, H. P. (2000). Glycosylation/Hydroxylation-Induced Stabilization of the Collagen Triple Helix: 4-Trans-Hydroxyproline in the Xaa Position Can Stabilize the Triple Helix. *J. Biol. Chem.* **275**, 24466-24469.
- Bann, J. G., Bächinger, H. P. & Peyton, D. H. (2003). Role of Carbohydrate in Stabilizing the Triple Helix in a Model for Deep-Sea Hydrothermal Vent Worm Collagen. *Biochemistry* **42**, 4042-4048.

- Beausoleil, E., Sharma, R., Michnick, S. W. & Lubell, W. D. (1998). Alkyl-3-Position Substituents Retard the Isomerization of Prolyl and Hydroxyprolyl Amides in Water. *J. Org. Chem.* **63**, 6572-6578.
- Beck, K., Chan, V. C., Shenoy, N., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. (2000). Destabilization of Osteogenesis Imperfecta Collagen-Like Model Peptides Correlates with the Identity of the Residue Replacing Glycine. *Proc. Nat. Acad. Sci. USA* **97**, 4273-4278.
- Beintema, J. J. & Kleinedam, R. G. (1998). The Ribonuclease a Superfamily: General Discussion. *Cell. Mol. Life Sci.* **54**, 825-842.
- Beligere, G. S. & Dawson, P. E. (2000). Design, Synthesis, and Characterization of 4-Ester Cl₂, a Model for Backbone Hydrogen Bonding in Protein Alpha-Helices. *J. Am. Chem. Soc.* **122**, 12079-12082.
- Bell, E. A., Qureshi, M. Y., Pryce, R. J., Janzen, D. H., Lemke, P. & Clardy, J. (1980). 2,4-Methanoproline (2-Carboxy-2,4-Methanopyrrolidine) and 2,4-Methanoglutamic Acid (1-Amino-1,3-Dicarboxycyclobutane) in Seeds of *Ateleia Herbert Smithii* Pittier (Leguminosae). *J. Am. Chem. Soc.* **102**, 1409-1412.
- Bella, J., Brodsky, B. & Berman, H. M. (1995). Hydration Structure of a Collagen Peptide. *Structure* **3**, 893-906.
- Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. (1994). Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 Å Resolution. *Science* **266**, 75-81.
- Berg, R. A. & Prockop, D. J. (1973). The Thermal Transition of a Non-Hydroxylated Form of Collagen. Evidence for a Role for Hydroxyproline in Stabilizing the Triple Helix of Collagen. *Biochem. Biophys. Res. Comm.* **52**, 115-120.
- Berisio, R., Granata, V., Vitagliano, L. & Zagari, A. (2004). Characterization of Collagen-Like Heterotrimers: Implications for Collagen Stability. *Biopolymers* **73**, 682-688.
- Berisio, R., Vitagliano, L., Mazzarella, L. & Zagari, A. (2001). Crystal Structure of a Collagen-Like Polypeptide with Repeating Sequence Pro-Hyp-Gly at 1.4 Å Resolution: Implications for Collagen Hydration. *Biopolymers* **56**, 8-13.
- Berisio, R., Vitagliano, L., Mazzarella, L. & Zagari, A. (2002). Crystal Structure of the Collagen Triple Helix Model [(Pro-Pro-Gly)₁₀]₃. *Protein Sci.* **11**, 262-270.
- Berisio, R., Vitagliano, L., Sorrentino, G., Carotenuto, L., Piccolo, C., Mazzarella, L. & Zagari, A. (2000). Effects of Microgravity on the Crystal Quality of a Collagen-Like Polypeptide. *Acta Crystallogr. D* **56**, 55-61.

- Bhate, M., Wang, X., Baum, J. & Brodsky, B. (2002). Folding and Conformational Consequence of Glycine to Alanine Replacements at Different Positions in a Collagen Model Peptide. *Biochemistry* **41**, 6539-6547.
- Blaber, M., Zhang, X. & Matthews, B. W. (1993). Structural Basis of Amino Acid Alpha Helix Propensity. *Science* **260**, 1637-1640.
- Blankenship, J. W., Balambika, R. & Dawson, P. E. (2002). Probing Backbone Hydrogen Bonds in the Hydrophobic Core of GCN4. *Biochemistry* **41**, 15676-15684.
- Blessing, R. H. (1995). An Empirical Correction for Absorption Anisotropy. *Acta Cryst* **A51**, 33-38.
- Boudko, S., Frank, S., Kammerer, R. A., Stetefeld, J., Schulthess, T., Landwehr, R., Lustig, A., Bachinger, H. P. & Engel, J. (2002). Nucleation and Propagation of the Collagen Triple Helix in Single-Chain and Trimerized Peptides: Transition from Third- to First-Order Kinetics. *J. Mol. Biol.* **317**, 459-470.
- Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L. & Raines, R. T. (2001). Conformational Stability of Collagen Relies on a Stereoelectronic Effect. *J. Am. Chem. Soc.* **123**, 777-778.
- Bretscher, L. E., Taylor, K. M. & Raines, R. T. (2000). Effect of Fluoro-Substituted Proline Residues on the Conformational Stability of Triple-Helical Collagen Mimics. Peptides for the New Millennium: Proceedings of the Sixteenth American Peptide Symposium. Fields, G. B., Tam, J. P. & Barany, G. Dordrecht, The Netherlands, Kluwer Academic, 355-356.
- Buevich, A. B., Dai, Q. H., Liu, X., Brodsky, B. & Baum, J. (2000). Site-Specific Nmr Monitoring of Cis-Trans Isomerization in the Folding of the Proline-Rich Collagen Triple Helix. *Biochemistry* **39**, 4299-4308.
- Bulleid, N. J., John, D. C. A. & Kadler, K. E. (2000). Recombinant Expression Systems for the Production of Collagen. *Biochem. Soc. Trans.* **28**, 350-353.
- Bulleid, N. J., Wilson, R. & Lees, J. F. (1996). Type-III Procollagen Assembly in Semi-Intact Cells: Chain Association, Nucleation and Triple-Helix Folding Do Not Require Formation of Inter-Chain Disulphide Bonds but Triple-Helix Nucleation Does Require Hydroxylation. *Biochem. J.* **317**, 195-202.
- Bunuel, E., Gil, A. M., Diaz de Vallegas, M. D. & Cativiela, C. (2001). Synthesis of Constrained Prolines by Diels-Alder Reaction Using a Chiral Unsaturated Oxazolone Derived from (R)-Glyceraldehyde as a Starting Material. *Tetrahedron* **57**, 6417-6427.

- Bürgi, H. B. & Dunitz, J. D. (1983). From Crystal Statics to Chemical Dynamics. *Acc. Chem. Res.* **16**, 153-161.
- Bürgi, H. B., Dunitz, J. D., Lehn, J. M. & Wipff, G. (1974a). Stereochemistry of Reaction Paths at Carbonyl Centres. *Tetrahedron* **30**, 1563-1572.
- Bürgi, H. B., Dunitz, J. D. & Shefter, E. (1973). Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **95**(15), 5065-5067.
- Bürgi, H. B., Lehn, J. M. & Wipff, G. (1974b). An *Ab Initio* Study of Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **96**, 1965-1966.
- Burjanadze, T. V. (1979). Hydroxyproline Content and Location in Relation to Collagen Thermal Stability. *Biopolymers* **18**, 931-938.
- Burjanadze, T. V. (2000). New Analysis of the Phylogenetic Change of Collagen Thermostability. *Biopolymers* **53**, 523-528.
- Byers, P. H. (2000). Collagens: Building Blocks at the End of the Development Line. *Clin. Genet.* **58**, 270-279.
- Byers, P. H. (2001). Folding Defects in Fibrillar Collagens. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 151-158.
- Chacko, K. K., Swaminathan, S. & Veena, K. R. (1983). The Conformation of Proline Using the Concept of Pseudorotation. *Curr. Sci. India* **52**, 660-663.
- Chan, V. C., Ramshaw, J. A. M., Kirkpatrick, A., Beck, K. & Brodsky, B. (1997). Positional Preferences of Ionizable Residues in Gly-X-Y Triplets of the Collagen Triple Helix. *J. Biol. Chem.* **272**, 31441-31446.
- Chapman, E., Thorson, J. S. & Schultz, P. G. (1997). Mutational Analysis of Backbone Hydrogen Bonds in Staphylococcal Nuclease. *J. Am. Chem. Soc.* **119**, 7151-7152.
- Chwang, A. K. & Sundaralingam, M. (1973). Intramolecular Hydrogen Bonding in 1-Beta-D-Arabinofuranosylcytosine (Ara-C). *Nature New Biology* **243**, 78-79.
- Codington, J. F., Doerr, I. L. & Fox, J. J. (1964). Nucleosides. XVIII. Synthesis of 2'-Fluorothymidine, 2'-Fluorodeoxyuridine, and Other 2'-Halogeno-2'-Deoxynucleosides. *J. Org. Chem.* **29**, 558-564.
- D'Alessio, G. & Riordan, J. F. (1997). Ribonucleases: Structures and Functions. New York, Academic Press.

- Danielson, M. A. & Raines, R. T. (2000). Contribution of Mainchain-Mainchain Hydrogen Bonds to the Conformational Stability of Triple-Helical Collagen. Peptides for the New Millennium: Proceedings of the Sixteenth American Peptide Symposium. Fields, G. B., Tam, J. P. & Barany, G. Dordrecht, The Netherlands, Kluwer Academic, 347-348.
- Davies, D. B. & Danyluk, S. S. (1975). Nuclear Magnetic Resonance Studies of 2'- and 3'-Ribonucleotide Structures in Solution. *Biochemistry* **14**(3), 543-554.
- Delaney, N. G. & Madison, V. (1982). Novel Conformational Distributions of Methylproline Peptides. *J. Am. Chem. Soc.* **104**, 6635-6641.
- delCardayre, S. B. & Raines, R. T. (1994). Structural Determinants of Enzyme Processivity. *Biochemistry* **33**, 6031-6037.
- delCardayre, S. B. & Raines, R. T. (1995). A Residue to Residue Hydrogen Bond Mediates the Nucleotide Specificity of Ribonuclease A. *J. Mol. Biol.* **252**, 328-336.
- delCardayre, S. B., Ribo, M., Yokel, E. M., Quirk, D. J., Rutter, W. J. & Raines, R. T. (1995). Engineering Ribonuclease A: Production, Purification, and Characterization of Wild-Type Enzyme and Mutants at Gln11. *Protein Eng.* **8**, 261-173.
- DeRider, M. L., Wilkens, S. J., Waddell, M. J., Bretscher, L. E., Weinhold, F., Raines, R. T. & Markley, J. L. (2002). Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **124**, 2497-2505.
- DeTar, D. F. & Luthra, N. P. (1977). Conformations of Proline. *J. Am. Chem. Soc.* **99**(4), 1232-1244.
- Dippy, J. F. J., Hughes, S. R. C. & Rozanski, A. (1959). The Dissociation Constants of Some Symmetrically Disubstituted Succinic Acids. *J. Chem. Soc.*, 2492-2498.
- Doi, T., Nishi, Y., Uchiyama, S., Nishiuchi, Y., Nakazawa, T., Ohkubo, T. & Kobayashi, Y. (2003). Characterization of Collagen Model Peptides Containing 4-Fluoroproline; (4(S)-Flpprogly)10 Forms a Triple Helix, but (4(R)-Flpprogly)10 Does Not. *J. Am. Chem. Soc.* **125**, 9922-9923.
- Dunitz, J. D. & Taylor, R. (1997). Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem. Eur. J.* **3**, 89-98.

- Eberhardt, E. S., Loh, S. N. & Raines, R. T. (1993). Thermodynamic Origin of Prolyl Peptide Bond Isomers. *Tetrahedron Lett.* **34**, 3055-3056.
- Eberhardt, E. S., Panasik, N., Jr. & Raines, R. T. (1996). Inductive Effects on the Energetics of Prolyl Peptide Bond Isomerization: Implications for Collagen Folding and Stability. *J. Am. Chem. Soc.* **118**, 12261-12266.
- Eliel, E. L. & Wilen, S. H. (1994). Stereochemistry of Organic Compounds. New York, Wiley Interscience.
- Engel, J., Chen, H.-T., Prockop, D. J. & Klump, H. (1977). The Triple Helix-Coil Conversion of Collagen-Like Polytripeptides in Aqueous and Nonaqueous Solvents. Comparison of the Thermodynamic Parameters and the Binding of Water to (L-Pro-L-Pro-Gly)_n and (L-Pro-L-Hyp-Gly)_n. *Biopolymers* **16**, 601-622.
- Engel, J. & Prockop, D. J. (1998). Does Bound Water Contribute to the Stability of Collagen? *Matrix Biol.* **17**, 679-680.
- Fan, P., Li, M. H., Brodsky, B. & Baum, J. (1993). Backbone Dynamics of (Pro-Hyp-Gly)₁₀ and a Designed Collagen-Like Triple-Helical Peptide by ¹⁵N Nmr Relaxation and Hydrogen-Exchange Measurements. *Biochemistry* **32**, 13299-13309.
- Fasman, G. D., Ed. (1989). Practical Handbook of Biochemistry and Molecular Biology. Boca Raton, FL, CRC Press.
- Feng, Y., Melacini, G., Taulane, J. P. & Goodman, M. (1996). Acetyl-Terminated and Template-Assembled Collagen-Based Polypeptides Composed of Gly-Pro-Hyp Sequences. 2. Synthesis and Conformational Analysis by Circular Dichroism, Ultraviolet Absorbance, and Optical Rotation. *J. Am. Chem. Soc.* **118**, 10351-10358.
- Fertala, A., Sieron, A. L., Ganguly, A., Li, S.-H., Ala-Kokko, L., Anumula, K. R. & Prockop, D. J. (1994). Synthesis of Recombinant Human Procollagen II in a Stably Transfected Tumour Cell Line (Ht1080). *Biochem. J.* **298**, 31-37.
- Fields, C. G., Lovdahl, C. M., Miles, A. J., Matthias Hagen, V. L. & Fields, G. B. (1993). Solid-Phase Synthesis and Stability of Triple-Helical Peptides Incorporating Native Collagen Sequences. *Biopolymers* **33**, 1695-1707.
- Fields, C. G., Mickelson, D. J., Drake, S. L., McCarthy, J. B. & Fields, G. B. (1993). Melanoma Cell Adhesion and Spreading Activities of a Synthetic 124-Residue Triple-Helical "Mini-Collagen". *J. Biol. Chem.* **268**, 14153-14160.

- Fields, G. B. (1995). The Collagen Triple-Helix: Correlation of Conformation with Biological Activities. *Connect. Tissue Res.* **31**, 235-243.
- Fields, G. B. (1999). Induction of Protein-Like Molecular Architecture by Self-Assembly Processes. *Bioorg. Med. Chem.* **7**, 75-81.
- Fields, G. B. & Prockop, D. J. (1996). Perspectives on the Synthesis and Application of Triple-Helical, Collagen-Model Peptides. *Biopolymers* **40**, 345-357.
- Fisher, E. (1906). *Berichte der deutschen chemischen Gesellschaft* **39**, 530-610.
- Forsén, S. & Hoffman, R. A. (1963). Study of Moderately Rapid Chemical Exchange Reactions by Means of Nuclear Magnetic Double Resonance. *J. Chem. Phys.* **39**, 2892-2901.
- Fox, J. J. & Wempen, I. (1965). Nucleosides XXVI. A Facile Synthesis of 2, 2'-Anhydro-Arabino Pyrimidine Nucleosides. *Tetrahedron Lett.*, 643-646.
- Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J. G., Guo, Y., Lustig, A., Bachinger, H. P. & Engel, J. (2001). Stabilization of Short Collagen-Like Triple Helices by Protein Engineering. *J. Mol. Biol.* **308**, 1081-1089.
- Fraser, R. D. B., MacRae, T. P. & Suzuki, E. (1979). Chain Conformation in the Collagen Molecule. *J. Mol. Biol.* **129**, 463-481.
- Friebolin, H. (1998). Basic One- and Two-Dimensional Nmr Spectroscopy, 3rd Ed. New York, VCH.
- Friedman, L., Higgin, J. J., Moulder, G., Barstead, R., Raines, R. T. & Kimble, J. (2000). Prolyl 4-Hydroxylase Is Required for Viability and Morphogenesis in *Caenorhabditis Elegans*. *Proc. Natl. Acad. Sci. USA* **97**, 4736-4741.
- Gaitanopoulos, D. E. & Weinstock, J. (1985). 2-Azabicyclo[2.2.1]Heptane-3-Carboxylic Acid--a Bicyclic Proline. *J. Het. Chem.* **22**, 957-959.
- Gardner, R. R., Liang, G.-B. & Gellman, S. H. (1995). An Achiral Dipeptide Mimetic That Promotes Beta-Hairpin Formation. *J. Am. Chem. Soc.* **117**, 3280-3281.
- Gardner, R. R., Liang, G.-B. & Gellman, S. H. (1999). Beta-Turn and Beta-Hairpin Mimicry with Tetrasubstituted Alkenes. *J. Am. Chem. Soc.* **121**, 1806-1816.
- Germann, H.-P. & Heidemann, E. (1988). A Synthetic Model of Collagen: An Experimental Investigation of the Triple-Helix Stability. *Biopolymers* **27**, 157-163.

- Giacovazzo, C., Monaco, H. L., Artioli, G., Viterbo, D., Ferraris, G., Gilli, G., Zanotti, G. & Catti, M. (2002). Fundamentals of Crystallography, 2nd Edn. Oxford, UK, Oxford University Press.
- Giner, J.-L. & Ferris, W. V., Jr. (2002). Synthesis of 2-C-Methyl-D-Erythritol-2,4-Cyclopyrophosphate. *Org. Lett.* **4**(7), 1225-1226.
- Gisin, B. F. & Merrifield, R. B. (1972). Carboxyl-Catalyzed Intramolecular Aminolysis. A Side Reaction in Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* **94**, 3102-3106.
- Glasoe, P. K. & Long, F. A. (1960). Use of Glass Electrodes to Measure Acidities in Deuterium Oxide. *J. Phys. Chem.* **64**, 188-190.
- Goodman, M., Bhumralkar, M., Jefferson, E. A., Kwak, J. & Locardi, E. (1998). Collagen Mimetics. *Biopolymers* **47**, 127-142.
- Goodman, M., Feng, Y., Melacini, G. & Taulane, J. P. (1996). A Template-Induced Incipient Collagen-Like Triple-Helical Structure. *J. Am. Chem. Soc.* **118**, 5156-5157.
- Gordon, D. J. & Meredith, S. C. (2003). Probing the Role of Backbone Hydrogen Bonding in Beta-Amyloid Fibrils with Inhibitor Peptides Containing Ester Bonds at Alternate Positions. *Biochemistry* **42**, 475-485.
- Gottlieb, A. A., Fujita, Y., Udenfriend, S. & Witkop, B. (1965). Incorporation of Cis- and Trans-4-Fluoro-L-Prolines into Proteins and Hydroxylation of the Trans Isomer During Collagen Biosynthesis. *Biochemistry* **4**, 2507-2513.
- Grathwohl, C. & Wüthrich, K. (1981). NMR Studies of the Rates of Proline *Cis* – *Trans* Isomerization in Oligopeptides. *Biopolymers* **20**, 2623-2633.
- Groebke, K., Renold, P., Tsang, K. Y., Allen, T. J., McClure, K. F. & Kemp, D. S. (1996). Template-Nucleated Alanine-Lysine Helices Are Stabilized by Position-Dependent Interactions between the Lysine Side Chain and the Helix Barrel. *Proc. Nat. Acad. Sci. USA* **93**, 4025-4029.
- Gryder, R. M., Lamon, M. & Adams, E. (1975). Sequence Position of 3-Hydroxyproline in Basement Membrane Collagen. *J. Biol. Chem.* **250**, 2470-2474.
- Guschlbauer, W. & Jankowski, K. (1980). Nucleoside Conformation Is Determined by the Electronegativity of the Sugar Substituent. *Nuc. Acids Res.* **8**(6), 1421-1433.
- Gustavson, K. H. (1957). Oxford, UK, Blackwell.

- Guzman, N. A., Ed. (1998). Prolyl Hydroxylase, Protein Disulfide Isomerase, and Other Structurally Related Proteins. New York, Marcel Dekker.
- Halab, L., Gosselin, F. & Lubell, W. D. (2000). Design, Synthesis, and Conformational Analysis of Azacycloalkane Amino Acids as Conformationally Constrained Probes for Mimicry of Peptide Secondary Structures. *Biopolymers (Pept. Sci.)* **55**, 101-122.
- Han, W., Pelletier, J. C., Mersinger, L. J., Kettner, C. A. & Hodge, C. N. (1999). 7-Azabicycloheptane Carboxylic Acid: A Proline Replacement in a Boroarginine Thrombin Inhibitor. *Org. Lett.* **1**, 1875-1877.
- Hann, M. M., Sammes, P. G., Kennewell, P. D. & Taylor, J. B. (1980). On Double Bond Isosteres of the Peptide Bond; an Enkephalin Analogue. *J. C. S. Chem. Commun.*, 234-235.
- Hansch, C., Leo, A. & Taft, R. W. (1991). A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* **91**, 165-195.
- Higgin, J. J., Yakovlev, G. I., Mitkevich, V. A., Makarov, A. A. & Raines, R. T. (2003). Zinc-Mediated Inhibition of Ribonuclease Sa by an N-Hydroxyurea Nucleotide and Its Basis. *Bioorg. Med. Chem. Lett.* **13**, 409-412.
- Hinderaker, M. P. & Raines, R. T. (2003). An Electronic Effect on Protein Structure. *Protein Sci.* **12**, 1188-1194.
- Hodges, J. A. & Raines, R. T. (2003). Stereoelectronic Effects on Collagen Stability: The Dichotomy of 4-Fluoroproline Diastereomers. *J. Am. Chem. Soc.* **125**, 9262-9263.
- Holmgren, S. K., Bretscher, L. E., Taylor, K. M. & Raines, R. T. (1999). A Hyperstable Collagen Mimic. *Chem. Biol.* **6**, 63-70.
- Holmgren, S. K., Taylor, K. M., Bretscher, L. E. & Raines, R. T. (1998). Code for Collagen's Stability Deciphered. *Nature* **392**, 666-667.
- Howard, J. A. K., Hoy, V. J., O'Hagan, D. & Smith, G. T. (1996). How Good Is Fluorine as a Hydrogen Bond Acceptor? *Tetrahedron* **52**, 12613-12622.
- Hyman, H. H., Kaganove, A. & Katz, J. J. (1960). The Basicity of Amino Acids in D₂O. *J. Phys. Chem.* **64**, 1653-1655.
- Ibrahim, H. & Togni, A. (2004). Enantioselective Halogenation Reactions. *Chem. Commun.*, 1147-1155.

- Improta, R., Benzi, C. & Barone, V. (2001). Understanding the Role of Stereoelectronic Effects in Determining Collagen Stability. 1. A Quantum Mechanical Study of Proline, Hydroxyproline, and Fluoroproline Dipeptide Analogues in Aqueous Solution. *J. Am. Chem. Soc.* **123**, 12568-12577.
- Improta, R., Mele, F., Crescenzi, O., Benzi, C. & Barone, V. (2002). Understanding the Role of Stereoelectronic Effects in Determining Collagen Stability. 2. A Quantum Mechanical/Molecular Mechanical Study of (Proline-Proline-Glycine)_n Polypeptides. *J. Am. Chem. Soc.* **124**, 7857-7865.
- Inouye, K., Kobayashi, Y., Kyogoku, Y., Kishida, Y., Sakakibara, S. & Prockop, D. J. (1982). Synthesis and Physical Properties of (Hydroxyproline-Proline-Glycine)₁₀. Hydroxyproline in the X-Position Decreases the Melting Temperature of the Collagen Triple Helix. *Arch. Biochem. Biophys.* **219**, 198-203.
- Inouye, K., Sakakibara, S. & Prockop, D. J. (1976). Effects of the Stereo-Configuration of the Hydroxyl Group in 4-Hydroxyproline on the Triple-Helical Structures Formed by Homogenous Peptides Resembling Collagen. *Biochim. Biophys. Acta* **420**, 133-141.
- Jabs, A., Weiss, M. S. & Hilgenfeld, R. (1999). Non-Proline Cis Peptide Bonds in Proteins. *J. Mol. Biol.* **286**, 291-304.
- Jenkins, C. L., Bretscher, L. E., Guzei, I. A. & Raines, R. T. (2003). Effect of 3-Hydroxyproline Residues on Collagen Stability. *J. Am. Chem. Soc.* **125**, 6422-6427.
- Jenkins, C. L. & Raines, R. T. (2002). Insights on the Conformational Stability of Collagen. *Nat. Prod. Rep.* **19**, 49-59.
- Johnson, G., Jenkins, M., McClean, K. M., Griesser, H. J., Kwak, J., Goodman, M. & Steele, J. G. (2000). Peptoid-Containing Collagen Mimetics with Cell Binding Activity. *J. Biomed. Mater. Res.* **51**, 612-624.
- Johnson, R. L. (1984). Inhibition of Renin by Substrate Analogue Inhibitors Containing the Olefinic Amino Acid 5(S)-Amino-7-Methyl-3(E)-Octenoic Acid. *J. Med. Chem.* **27**, 1351-1354.
- Jones, E. Y. & Miller, A. (1991). Analysis of Structural Design Features in Collagen. *J. Mol. Biol.* **218**, 209-219.
- Jung, M. E., Shishido, K., Light, L. & Davis, L. (1981). Preparation of Di- and Triacylimines and Their Use in the Synthesis of Nitrogen Heterocycles. *Tetrahedron Lett.* **22**, 4607-4610.

- Juvvadi, P., Dooley, D. J., Humblet, C. C., Lu, G. H., Lunney, E. A., Panek, R. L., Skeeane, R. & Marshall, G. R. (1992). Bradykinin and Angiotensin II Analogs Containing a Conformationally Constrained Proline Analog. *Int. J. Peptide Protein Res.* **40**, 163-170.
- Karle, I. L., Das, C. & Balaram, P. (2001). Effects of Hydrogen-Bond Deletion on Peptide Helices: Structural Characterization of Depsipeptides Containing Lactic Acid. *Biopolymers* **59**, 276-289.
- Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* **14**, 1-63.
- Kefalides, N. A. (1973). Structure and Biosynthesis of Basement Membranes. *Int. Rev. Connect Tissue Res.* **6**, 63-104.
- Kelemen, B. R., Klink, T. A., Behlke, M. A., Eubanks, S. R., Leland, P. A. & Raines, R. T. (1999). Hypersensitive Substrate for Ribonucleases. *Nuc. Acids Res.* **27**(18), 3696-3701.
- Kelemen, B. R., Schultz, L. W., Sweeney, R. Y. & Raines, R. T. (2000). Excavating an Active Site: The Nucleobase Specificity of Ribonuclease A. *Biochemistry* **39**, 14487-14494.
- Kemp, D. S. & Petrakis, K. S. (1981). Synthesis and Conformational Analysis of Cis,Cis-1,3,5-Trimethylcyclohexane-1,3,5-Tricarboxylic Acid. *J. Org. Chem.* **46**, 5140-5143.
- Kersteen, E. A. & Raines, R. T. (2001). Contribution of Tertiary Amides to the Conformational Stability of Collagen Triple Helices. *Biopolymers* **59**, 24-28.
- Kim, C. W. A. & Berg, J. M. (1993). Thermodynamic Beta-Sheet Propensities Measured Using a Zinc-Finger Host Peptide. *Nature* **362**, 267-270.
- Kinberger, G. A., Cai, W. & Goodman, M. (2002). Collagen Mimetic Dendrimers. *J. Am. Chem. Soc.* **124**, 15162-15163.
- Kobayashi, Y., Sakai, R., Kakiuchi, K. & Isemura, T. (1970). Physicochemical Analysis of (Pro-Pro-Gly)_n with Defined Molecular Weight--Temperature Dependence of Molecular Weight in Aqueous Solution. *Biopolymers* **9**, 415-425.
- Koh, J. T., Cornish, V. W. & Schultz, P. G. (1997). An Experimental Approach to Evaluating the Role of Backbone Interactions in Proteins Using Unnatural Amino Acid Mutagenesis. *Biochemistry* **36**, 11314-11322.

- Krainer, E. & Naider, F. (1993). A New Method for the Detritylation of Alcohols Bearing Other Reducible and Acid-Hydrolyzable Functionalities. *Tetrahedron Lett.* **34**(11), 1713-1716.
- Kramer, R. Z., Bella, J., Mayville, P., Brodsky, B. & Berman, H. M. (1999). Sequence Dependent Conformational Variations of Collagen Triple-Helical Structure. *Nat. Struct. Biol.* **6**, 454-457.
- Kramer, R. Z. & Berman, H. M. (1998). Patterns of Hydration in Crystalline Collagen Peptides. *J. Biomol. Struct. Dyn.* **16**, 367-380.
- Kramer, R. Z., Vitagliano, L., Bella, J., Berisio, R., Mazzarella, L., Brodsky, B., Zagari, A. & Berman, H. M. (1998). X-Ray Crystallographic Determination of a Collagen-Like Peptide with the Repeating Sequence (Pro-Pro-Gly). *J. Mol. Biol.* **280**, 623-638.
- Krow, G. R. & Cannon, K. C. (2004). Azabicyclo[2.1.1]Hexanes. A Review. *Heterocycles* **62**, 877-898.
- Kwak, J., De Capua, A., Locardi, E. & Goodman, M. (2002). TREN (Tris(2-Aminoethyl)Amine): An Effective Scaffold for the Assembly of Triple Helical Collagen Mimetic Structures. *J. Am. Chem. Soc.* **124**, 14085-14091.
- Kwak, J., Jefferson, E. A., Bhumralkar, M. & Goodman, M. (1999). Triple Helical Stabilities of Guest-Host Collagen Mimetic Structures. *Bioorg. Med. Chem.* **7**, 153-160.
- Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T. & Kivirikko, K. I. (1996). Characterization of Human Type Iii Collagen Expressed in a Baculovirus System. *J. Biol. Chem.* **271**, 11988-11995.
- Led, J. J. & Gesmar, H. (1982). The Applicability of the Magnetization-Transfer Technique to Determine Chemical Exchange Rates in Extreme Cases. The Importance of Complementary Experiments. *J. Mag. Res.* **49**, 444-463.
- Leo, A., Hansch, C. & Elkins, D. (1971). Partition Coefficients and Their Uses. *Chem. Rev.* **71**, 525-616.
- Leonidas, D. D., Chavali, G. B., Oikonomakos, N. G., Chrysina, E. D., Kosmopoulou, M. N., Vlassi, M., Frankling, C. & Acharya, K. R. (2003). High Resolution Crystal Structures of Ribonuclease a Complexed with Adenylic and Uridylic Nucleotide Inhibitors. Implications for Structure-Based Design of Ribonucleolytic Inhibitors. *Prot. Sci.* **12**, 2559-2574.

- Leonidas, D. D., Shapiro, R., Irons, L. I., Russo, N. & Acharya, K. R. (1999). Toward Rational Design of Ribonuclease Inhibitors: High Resolution Crystal Structure of a Ribonuclease a Complex with a Potent 3',5'-Pyrophosphate-Linked Dinucleotide Inhibitor. *Biochemistry* **38**, 10287-10287.
- Li, M.-H., Fan, P., Brodsky, B. & Baum, J. (1993). Two-Dimensional NMR Assignments and Conformation of (Pro-Hyp-Gly)₁₀ and a Designed Collagen Triple-Helical Peptide. *Biochemistry* **32**, 7377-7387.
- Liu, X., Kim, S., Dai, Q. H., Brodsky, B. & Baum, J. (1998). Nuclear Magnetic Resonance Shows Asymmetric Loss of Triple Helix in Peptides Modeling a Collagen Mutation in Brittle Bone Disease. *Biochemistry* **37**, 15528-15533.
- Liu, X., Siegel, D. L., Fan, P., Brodsky, B. & Baum, J. (1996). Direct NMR Measurement of the Folding Kinetics of a Trimeric Peptide. *Biochemistry* **35**, 4306-4313.
- Loh, S. N. & Markley, J. L. (1994). Hydrogen Bonding in Proteins as Studied by Amide Hydrogen D/H Fractionation Factors: Application to Staphylococcal Nuclease. *Biochemistry* **33**, 1029-1036.
- Long, C. G., Li, M. H., Baum, J. & Brodsky, B. (1992). Nuclear Magnetic Resonance and Circular Dichroism Studies of a Triple-Helical Peptide with a Glycine Substitution. *J. Mol. Biol.* **225**, 1-4.
- Lu, W., Qasim, M. A., Laskowski, M., Jr. & Kent, S. B. H. (1997). Probing Intermolecular Main Chain Hydrogen Bonding in Serine Proteinase-Protein Inhibitor Complexes: Chemical Synthesis of Backbone-Engineered Turkey Ovomucoid Third Domain. *Biochemistry* **36**, 673-679.
- Lu, W., Randal, M., Kossiakoff, A. A. & Kent, S. B. H. (1999). Probing Intermolecular Backbone H-Bonding in Serine Proteinase-Protein Inhibitor Complexes. *Chem. Biol.* **6**, 419-427.
- MacArthur, M. W. & Thornton, J. M. (1991). Influence of Proline Residues on Protein Conformation. *J. Mol. Biol.* **218**, 397-412.
- Magaard, V. W., Sanchez, R. M., Bean, J. W. & Moore, M. L. (1993). A Convenient Synthesis of the Conformationally Constrained Amino Acid 5,5-Dimethylproline. *Tetrahedron Lett.* **34**, 381-384.
- Makarov, A. A., Yakovlev, G. I., Mitkevich, V. A., Higgin, J. J. & Raines, R. T. (2004). Zinc(II)-Mediated Inhibition of Ribonuclease Sa by an N-Hydroxyurea Nucleotide and Its Basis. *Biochem. Biophys. Res. Commun.*, in press.

- Malkar, N. B., Lauer-Fields, J. L., Borgia, J. A. & Fields, G. B. (2002). Modulation of Triple Helix Stability and Subsequent Melanoma Cellular Responses by Single-Site Substitution of Fluoroproline Derivatives. *Biochemistry* **41**, 6054-6064.
- Mammi, S. & Goodman, M. (1986). Polydepsipeptides. 13. Synthesis and ^1H NMR Analysis of Collagen Model Structures. *Int. J. Peptide Protein Res.* **28**, 29-44.
- Masse, C. E., Knight, B. S., Stavropoulos, P. & Panek, J. S. (1997). Asymmetric C-N Bond Constructions Via Crotlylsilane Addition Reactions: A Stereocontrolled Route to Dipeptide Isosteres. *J. Am. Chem. Soc.* **119**, 6040-6047.
- Melacini, G., Bonvin, A. M. J. J., Goodman, M., Boelens, R. & Kaptein, R. (2000). Hydration Dynamics of the Collagen Triple Helix by NMR. *J. Mol. Biol.* **300**, 1041-1048.
- Melacini, G. & Goodman, M. (1998). Improved Method for the Stereospecific ^1H -NMR Assignments in Collagen-Like Triple Helices. *Chirality* **10**, 28-34.
- Milner-White, J. E., Bell, L. H. & Maccallum, P. H. (1992). Pyrrolidine Ring Puckering in *Cis* and *Trans*-Proline Residues in Proteins and Polypeptides. *J. Mol. Biol.* **228**, 725-734.
- Minor, D. L. & Kim, P. S. (1994). Context Is a Major Determinant of Beta Sheet Propensity. *Nature* **371**, 264-267.
- Mizuno, K., Hayashi, T. & Bachinger, H. P. (2003). Hydroxylation-Induced Stabilization of the Collagen Triple Helix. Further Characterization of Peptides with 4(R)-Hydroxyproline in the Xaa Position. *J. Biol. Chem.* **278**, 32373-32379.
- Mizuno, K., Hayashi, T., Peyton, D. H. & Bachinger, H. P. (2004). The Peptides Acetyl-(Gly-3(S)Hyp-4(R)Hyp) $_{10}$ -NH $_2$ and Acetyl-(Gly-Pro-3(S)Hyp) $_{10}$ -NH $_2$ Do Not Form a Collagen Triple Helix. *J. Biol. Chem.* **279**, 282-287.
- Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. (1975). Energy Parameters in Polypeptides. VII. Geometric Parameters, Partial Atomic Charges, Nonbonded Interactions, Hydrogen Bond Interactions, and Intrinsic Torsional Potentials for the Naturally Occurring Amino Acids. *J. Phys. Chem.* **79**, 2361-2381.
- Montelione, G. T., Hughes, P., Clardy, J. & Scheraga, H. A. (1986). Conformational Properties of 2,4-Methanoproline (2-Carboxy-2,4-Methanopyrrolidine) in Peptides: Determination of Preferred Peptide Bond Conformation in Aqueous Solution by Proton Overhauser Measurements. *J. Am. Chem. Soc.* **108**, 6765-6773.

- Mooney, S. D., Huang, C. C., Kollman, P. A. & Klein, T. E. (2001). Computed Free Energy Differences between Point Mutations in a Collagen-Like Peptide. *Biopolymers* **58**, 347-353.
- Mooney, S. D., Kollman, P. A. & Klein, T. E. (2002). Conformational Preferences of Substituted Prolines in the Collagen Triple Helix. *Biopolymers* **64**, 63-71.
- Muller, J. C., Ottil, J. & Moroder, L. (2000). Heterotrimeric Collagen Peptides as Fluorogenic Collagenase Substrates: Synthesis, Conformational Properties, and Enzymatic Digestion. *Biochemistry* **39**, 5111-5116.
- Myers, J. K., Pace, C. N. & Scholtz, J. M. (1997). A Direct Comparison of Helix Propensity in Proteins and Peptides. *Proc. Nat. Acad. Sci. USA* **94**, 2833-2837.
- Myllyharju, J. & Kivirikko, K. I. (2001). Collagens and Collagen-Related Diseases. *Ann. Med.* **33**, 7-21.
- Myllyharju, J., Nokelainen, M., Vuorela, A. & Kivirikko, K. I. (2000). Expression of Recombinant Human Type I-III Collagens in the Yeast *Pichia Pastoris*. *Biochem. Soc. Trans.* **28**, 353-357.
- Nagarajan, V., Kamitori, S. & Okuyama, K. (1999). Structure Analysis of a Collagen-Model Peptide with a (Pro-Hyp-Gly) Sequence Repeat. *J. Biochem. (Tokyo)* **125**, 310-318.
- Nagyvary, J. (1969). Arabinonucleotides. II. The Synthesis of O²,2'-Anhydrocytidine 3'-Phosphate, a Precursor of 1-Beta-D-Arabinosylcytosine. *J. Am. Chem. Soc.* **91**, 5409-5410.
- Nimni, M. E., Ed. (1988). Collagen 1-4. Boca Raton, FL, CRC Press.
- Nomenclature (1970). IUPAC-IUB Commission on Biochemical Nomenclature. *Biochemistry* **9**, 3471-3479.
- O'Hagan, D., Bilton, C., Howard, J. A. K., Knight, L. & Tozer, D. J. (2000). The Preferred Conformation of N- β -Fluoroethylamides. Observation of the Fluorine Amide Gauche Effect. *J. Chem. Soc., Perkin Trans. 2*, 605-607.
- Oishi, S., Kamano, T., Niida, A., Odagaki, Y., Hamanaka, N., Yamamoto, M., Ajito, K., Tamamura, H., Otaka, A. & Fujii, N. (2002). Diastereoselective Synthesis of New Psi[(E)-CH=CMe]- and Psi[(Z)-CH=CMe]-Type Alkene Dipeptide Isosteres by Organocopper Reagents and Application to Conformationally Restricted Cyclic RGD Peptidomimetics. *J. Org. Chem.* **67**, 6162-6173.

- Okuyama, K., Nagarajan, V. & Kamitori, S. (1999). 7/2-Helical Model for Collagen - Evidence from Model Peptides. *Proc. Indian Acad. Sci.* **111**, 19-34.
- Olsen, D. R., Leigh, S. D., Chang, R., McMullin, H., Ong, W., Tai, E., Chisholm, G., Birk, D. E., Berg, R. A., Hitzeman, R. A. & Toman, P. D. (2001). Production of Human Type I Collagen in Yeast Reveals Unexpected New Insights into the Molecular Assembly of Collagen Trimers. *J. Biol. Chem.* **276**, 24038-24043.
- Olson, K. A., Fett, J. W., French, T. C., Key, M. E. & Vallee, B. L. (1995). Angiogenin Antagonists Prevent Tumor Growth in Vivo. *Proc. Natl. Acad. Sci. USA* **92**, 442-446.
- O'Neil, K. T. & DeGrado, W. F. (1990). A Thermodynamic Scale of the Helix-Forming Tendencies of the Commonly Occurring Amino Acids. *Science* **250**, 646-651.
- Ottl, J., Battistuta, R., Pieper, M., Tschesche, H., Bode, W., Kuhn, K. & Moroder, L. (1996). Design and Synthesis of Heterotrimeric Collagen Peptides with a Built-in Cystine Knot. Models for Collagen Catabolism by Matrix-Metalloproteases. *FEBS Lett.* **398**, 31-36.
- Ottl, J., Gabriel, D., Murphy, G., Knauper, V., Tominaga, Y., Nagase, H., Kroger, M., Tschesche, H., Bode, W. & Moroder, L. (2000). Recognition and Catabolism of Synthetic Heterotrimeric Collagen Peptides by Matrix Metalloproteinases. *Chem. Biol.* **7**, 119-132.
- Ottl, J. & Moroder, L. (1999a). Disulfide-Bridged Heterotrimeric Collagen Peptides Containing the Collagenase Cleavage Site of Collagen Type I. Synthesis and Conformational Properties. *J. Am. Chem. Soc.* **121**, 653-661.
- Ottl, J. & Moroder, L. (1999b). A New Strategy for Regioselective Interstrand Disulfide Bridging of Multiple Cysteine Peptides. *Tetrahedron Lett.* **40**, 1487-1490.
- Ottl, J., Musiol, H. J. & Moroder, L. (1999). Heterotrimeric Collagen Peptides Containing Functional Epitopes. Synthesis of Single-Stranded Collagen Peptides Related to the Collagenase Cleavage Site. *J. Peptide Sci.* **5**, 103-110.
- Otzen, D. E. & Fersht, A. R. (1995). Side-Chain Determinants of Beta-Sheet Stability. *Biochemistry* **34**, 5718-5724.
- Panasik, N., Jr., Eberhardt, E. S., Edison, A. S., Powell, D. R. & Raines, R. T. (1994). Inductive Effects on the Structure of Proline Residues. *Int. J. Pept. Protein Res.* **44**, 262-269.
- Pauling, L. (1960). The Nature of the Chemical Bond, 3rd Ed. Ithaca, NY, Cornell University Press.

- Perich, J. W. & Johns, R. B. (1987). A New, Convenient and Efficient General Procedure for the Conversion of Alcohols into Their Dibenzyl Phosphorotriesters Using N,N-Diethyl Dibenzyl Phosphoramidite. *Tetrahedron Lett.* **28**, 101-102.
- Persikov, A. V., Ramshaw, J. A. M. & Brodsky, B. (2000a). Collagen Model Peptides: Sequence Dependence of Triple-Helix Stability. *Biopolymers* **55**, 436-450.
- Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. (2000b). Amino Acid Propensities for the Collagen Triple Helix. *Biochemistry* **39**, 14960-14967.
- Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. (2002). Peptide Investigations of Pairwise Interactions in the Collagen Triple Helix. *J. Mol. Biol.* **316**, 385-394.
- Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. (2003). Triple Helix Propensity of Hydroxyproline and Fluoroproline: Comparison of Host-Guest and Repeating Tripeptide Collagen Models. *J. Am. Chem. Soc.* **125**, 11500-11501.
- Pollard, D. R. & Nagyvary, J. (1973). Inhibition of Pancreatic Ribonuclease a by Arabinonucleotides. *Biochemistry* **12**, 1063-1066.
- Prockop, D. J. (1998). What Holds Us Together? Why Do Some of Us Fall Apart? What Can We Do About It? *Matrix Biol.* **16**, 519-528.
- Prockop, D. J. (1999). Pleasant Surprises *En Route* from the Biochemistry of Collagen to Attempts at Gene Therapy. *Biochem. Soc. Trans.* **27**, 15-31.
- Prockop, D. J. & Kivirikko, K. I. (1995). Collagens: Molecular Biology, Diseases, and Potentials for Therapy. *Annu. Rev. Biochem.* **64**, 403-434.
- Rabanal, F., DeGrado, W. F. & Dutton, P. L. (1996). Use of 2,2'-Dithiobis(5-Nitropyridine) for the Heterodimerization of Cysteine Containing Peptides. Introduction of the 5-Nitro-2-Pyridinesulfonyl Group. *Tetrahedron Lett.* **37**, 1347-1350.
- Radmer, R. J. & Klein, T. E. (2004). Severity of Osteogenesis Imperfecta and Structure of a Collagen-Like Peptide Model a Lethal Mutation Site. *Biochemistry* **43**, 5314-5323.
- Raines, R. T. (1998). Ribonuclease A. *Chem. Rev.* **98**, 1045-1066.
- Ramachandran, G. N., Ed. (1967). Chemistry of Collagen. London, Academic Press.
- Ramachandran, G. N. & Kartha, G. (1954). Structure of Collagen. *Nature* **174**, 269-270.

- Ramachandran, G. N. & Kartha, G. (1955). Structure of Collagen. *Nature* **176**, 593-595.
- Ramachandran, G. N. & Reddi, A. H., Eds. (1976). Biochemistry of Collagen. New York, Plenum Press.
- Ramakrishnan, C. (2001). In Memorium: Professor G. N. Ramachandran (1922–2001). *Protein Sci.* **10**, 1689-1691.
- Ramshaw, J. A., Werkmeister, J. A. & Glattauer, V. (1996). Collagen-Based Biomaterials. *Biotechnol. Genet. Eng. Rev.* **13**, 335-382.
- Ramshaw, J. A. M., Shah, N. K. & Brodsky, B. (1998). Gly-X-Y Tripeptide Frequencies in Collagen: A Context for Host-Guest Triple-Helical Peptides. *J. Struct. Biol.* **122**, 86-91.
- Reiersen, H. & Rees, A. R. (2001). The Hunchback and Its Neighbours: Proline as an Environmental Modulator. *Trends Biochem. Sci.* **26**, 679-684.
- Renner, C., Alefelder, S., Bae, J. H., Budisa, N., Huber, R. & Moroder, L. (2001). Fluoroprolines as Tools for Protein Design and Engineering. *Angew. Chem. Int. Ed. Engl.* **40**, 923-925.
- Rich, A. & Crick, F. H. C. (1955). The Structure of Collagen. *Nature* **176**, 915-916.
- Rich, A. & Crick, F. H. C. (1961). The Molecular Structure of Collagen. *J. Mol. Biol.* **3**, 483-506.
- Riddihough, G. (1998). Structure of Collagen. *Nat. Struct. Biol.* **5**, 858-859.
- Rougvie, M. A. & Bear, R. S. (1953). An X-Ray Diffraction Investigation of Swelling by Collagen. *J. Amer. Leather Chem. Ass.* **48**, 735-751.
- Roussev, C. D., Simeonov, M. F. & Petkov, D. D. (1997). Arabinonucleotide Synthesis by the Epoxide Route. *J. Org. Chem.* **62**, 5238-5240.
- Ruggiero, F., Exposito, J.-Y., Bournat, P., Gruber, V., Perret, S., Comte, J., Olganier, B., Garrone, R. & Thiesen, M. (2000). Triple Helix Assembly and Processing of Human Collagen Produced in Transgenic Tobacco Plants. *FEBS Lett.* **469**, 132-136.
- Rump, E. T., Rijkers, D. T. S., Hilbers, H. W., de Groot, P. G. & Liskamp, R. M. J. (2002). Cyclotrimeratrylene (CTV) as a New Chiral Triacid Scaffold Capable of Inducing Triple Helix Formation of Collagen Peptides Containing Either a Native Sequence or Pro-Hyp-Gly Repeats. *Chem. Eur. J.* **8**, 4613-4621.

- Russo, A., Acharya, K. R. & Shapiro, R. (2001). Small Molecule Inhibitors of Rnase a and Related Enzymes. *Methods Enzym.* **341**, 629-648.
- Russo, A. & Shapiro, R. (1999). Potent Inhibition of Mammalian Ribonucleases by 3',5'-Pyrophosphate-Linked Nucleotides. *J. Biol. Chem.* **274**, 14902-14908.
- Russo, A., Shapiro, R. & Vallee, B. L. (1997). 5'-Diphosphoadenosine 3'-Phosphate Is a Potent Inhibitor of Bovine Pancreatic Ribonuclease A. *Biochem. Biophys. Res. Commun.* **231**, 671-674.
- Sacca, B., Barth, D., Musiol, H. J. & Moroder, L. (2002). Conformation-Dependent Side Reactions in Interstrand-Disulfide Bridging of Trimeric Collagenous Peptides by Regioselective Cysteine Chemistry. *J. Peptide Sci.* **8**, 205-210.
- Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y. & Prockop, D. J. (1973). Synthesis of (Pro-Hyp-Gly)_n of Defined Molecular Weights. Evidence for the Stabilization of Collagen Triple Helix by Hydroxyproline. *Biochim. Biophys. Acta.* **303**, 198-202.
- Sato, Y., Utsumi, K., Maruyama, T., Kimura, T., Yamamoto, I. & Richman, D. D. (1994). Synthesis and Hypnotic and Anti-Human Immunodeficiency Virus-1 Activities of N³-Substituted 2'-Deoxy-2'-Fluorouridines. *Chem. Pharm. Bull.* **42**, 595-598.
- Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C. & Brodsky, B. (1996). A Host – Guest Set of Triple-Helical Peptides: Stability of Gly – X – Y Triplets Containing Common Nonpolar Residues. *Biochemistry* **35**, 10262-10268.
- Shapiro, R., Fox, E. A. & Riordan, J. F. (1989). Role of Lysines in Human Angiogenin: Chemical Modification and Site-Directed Mutagenesis. *Biochemistry* **28**, 1726-1732.
- Shapiro, R. & Vallee, B. L. (1989). Site-Directed Mutagenesis of Histidine-13 and Histidine-114 of Human Angiogenin. Alanine Derivatives Inhibit Angiogenin-Induced Angiogenesis. *Biochemistry* **28**, 7401-7408.
- Shaw, B. R. & Schurr, J. M. (1975). The Association Reaction of Collagen Model Polypeptides (Pro-Pro-Gly)_n. *Biopolymers* **14**, 1951-1985.
- Shin, I., Ting, A. Y. & Schultz, P. G. (1997). Analysis of Backbone Hydrogen Bonding in a Beta Turn of Staphylococcal Nuclease. *J. Am. Chem. Soc.* **119**, 12667-12668.

- Silinski, P. & Fitzgerald, M. C. (2003). Comparative Analysis of Two Different Amide-to-Ester Bond Mutations in the Beta-Sheet of 4-Oxalocrotonate Tautomerase. *Biochemistry* **42**, 6620-6630.
- Simon-Lukasik, K. V., Persikov, A. V., Brodsky, B., Ramshaw, J. A. M., Laws, W. R., Ross, J. B. A. & Ludescher, R. D. (2003). Fluorescence Determination of Tryptophan Side-Chain Accessibility and Dynamics in Triple-Helical Collagen-Like Peptides. *Biophys. J.* **84**, 501-508.
- Slatter, D. A., Miles, C. A. & Bailey, A. J. (2003). Asymmetry in the Triple Helix of Collagen-Like Heterotrimers Confirms That External Bonds Stabilize Collagen Structure. *J. Mol. Biol.* **329**, 175-183.
- Snellman, A., Keranen, M.-R., Hagg, P. O., Lamberg, A., Hiltunen, J. K., Kivirikko, K. I. & Pihlajaniemi, T. (2000). Type XIII Collagen Forms Homotrimers with Three Triple Helical Collagenous Domains and Its Association into Disulfide-Bonded Trimers Is Enhanced by Prolyl-4-Hydroxylase. *J. Biol. Chem.* **275**, 8936-8944.
- Stapley, B. J. & Creamer, T. P. (1999). A Survey of Left-Handed Polyproline II Helices. *Protein Sci.* **8**, 587-595.
- Stetefeld, J., Frank, S., Jenny, M., Schulthess, T., Kammerer, R. A., Boudko, S., Landwehr, R., Okuyama, K. & Engel, J. (2003). Collagen Stabilization at the Atomic Level: Crystal Structure of Designed (Gly-Pro-Pro)₁₀-Foldon. *Structure* **11**, 339-346.
- Stewart, D. E., Sarkar, A. & Wampler, J. E. (1990). Occurrence and Role of Cis Peptide Bonds in Protein Structures. *J. Mol. Biol.* **214**, 253-260.
- Stowell, J. K., Widlanski, T. S., Kutateladze, T. G. & Raines, R. T. (1995). Mechanism-Based Inactivation of Ribonuclease A. *J. Org. Chem.* **60**, 6930-6936.
- Street, A. G. & Mayo, S. L. (1999). Intrinsic Beta-Sheet Propensities Result from Van Der Waals Interactions between Side Chains and the Local Backbone. *Proc. Nat. Acad. Sci. USA* **96**, 9074-9076.
- Suzuki, E., Fraser, R. D. B. & MacRae, T. P. (1980). Role of Hydroxyproline in the Stabilization of the Collagen Molecule Via Water Molecules. *Int. J. Biol. Macromol.* **2**, 54-56.
- Takeuchi, T. & Prockop, D. J. (1969). Biosynthesis of Abnormal Collagens with Amino Acid Analogues. I. Incorporation of L-Azetidine-2-Carboxylic Acid and *Cis*-4-Fluoro-L-Proline into Protocollagen and Collagen. *Biochim. Biophys. Acta* **175**, 142-155.

- Takeuchi, T., Rosenbloom, J. & Prockop, D. J. (1969). Inability of Cartilage Cells to Extrude Collagen Polypeptides Containing L-Azetidine-2-Carboxylic Acid or Cis-4-Fluoro-L-Proline. *Biochim. Biophys. Acta* **175**, 156-164.
- Taktakishvili, M. & Nair, V. (2000). A New Pethod for the Phosphorylation of Nucleosides. *Tetrahedron Lett.* **41**, 7173-7176.
- Tamamura, H., Hiramatsu, K., Miyamoto, K., Omagari, A., Oishi, S., Nakashima, H., Yamamoto, N., Kuroda, Y., Nakagawa, T., Otaka, A. & Fujii, N. (2002). Synthesis and Evaluation of Pseudopeptide Analogues of a Specific CXCR4 Inhibitor, T140: The Insertion of an (E)-Alkene Dipeptide Isostere into the Beta-II'-Turn Moiety. *Bioorg. Med. Chem. Lett.* **12**, 923-928.
- Tamamura, H., Koh, Y., Ueda, S., Sasaki, Y., Yamasaki, T., Aoiki, M., Maeda, K., Watai, Y., Arikuni, H., Otaka, A., Mitsuya, H. & Fujii, N. (2003). Reduction of Peptide Character of HIV Protease Inhibitors That Exhibit Nanomolar Potency against Multidrug Resistant HIV-1 Strains. *J. Med. Chem.* **46**, 1764-1768.
- Tanaka, T., Wada, Y., Nakamura, H., Doi, T., Imanishi, T. & Kodama, T. (1993). A Synthetic Model of Collagen Structure Taken from Bovine Macrophage Scavenger Receptor. *FEBS Lett.* **334**, 272-276.
- Tanaka, Y., Suzuki, K. & Tanaka, T. (1998). Synthesis and Stabilization of Amino and Carboxy Terminal Constrained Collagenous Peptides. *J. Peptide Res.* **51**, 413-419.
- Thakur, S., Vadolas, D., Germann, H. P. & Heidemann, E. (1986). Influence of Different Tripeptides on the Stability of the Collagen Triple Helix. II. An Experimental Approach with Appropriate Variations of a Trimer Model Oligotripeptide. *Biopolymers* **25**, 1081-1086.
- Toman, P. D., Chisholm, G., McMullin, H., Giere, L. M., Olsen, D. R., Kovach, R. J., Leigh, S. D., Fong, B. E., Chang, R., Daniels, G. A., Berg, R. A. & Hitzeman, R. A. (2000). Production of Recombinant Human Type I Procollagen Trimers Using a Four-Gene Expression System in the Yeast *Saccharomyces Cerevisiae*. *J. Biol. Chem.* **275**, 23303-23309.
- Tryggvason, K., Risteli, J. & Kivirikko, K. (1976). Separation of Prolyl 3-Hydroxylase and 4-Hydroxylase Activities and the 4-Hydroxyproline Requirement for Synthesis of 3-Hydroxyproline. *Biochem. Biophys. Res. Commun.* **76**, 275-281.
- Uchiyama, S., Kai, T., Kajiyama, K., Kobayashi, Y. & Tomiyama, T. (1997). Measurement of Thermodynamic Quantities in the Heating-Rate Dependent Thermal Transitions of Sequenced Polytripeptides. *Chem. Phys. Lett.* **281**, 92-96.

- Vasbinder, M. M. & Miller, S. J. (2002). Synthesis of the Pro-Gly Dipeptide Alkene Isostere Using Olefin Cross-Metathesis. *J. Org. Chem.* **67**, 6240-6242.
- Venkateswarlu, D. & Ferguson, D. M. (1999). Effects of C2'-Substitution on Arabinonucleic Acid Structure and Conformation. *J. Am. Chem. Soc.* **121**, 5609-5610.
- Vitagliano, L., Berisio, R., Mastrangelo, A., Mazzarella, L. & Zagari, A. (2001a). Preferred Proline Puckerings in Cis and Trans Peptide Groups: Implications for Collagen Stability. *Protein Sci.* **10**, 2627-2632.
- Vitagliano, L., Berisio, R., Mazzarella, L. & Zagari, A. (2001b). Structural Bases of Collagen Stabilization Induced by Proline Hydroxylation. *Biopolymers* **58**, 459-464.
- Walmsley, A. R., Batten, M. R., Lad, U. & Bulleid, N. J. (1999). Intracellular Retention of Procollagen with the Endoplasmic Reticulum Is Mediated by Prolyl-4-Hydroxylase. *J. Biol. Chem.* **274**, 14884-14892.
- Walz, F. G., Jr. (1971). Kinetic and Equilibrium Studies on the Interaction of Ribonuclease a and 2'-Deoxyuridine-3'-Phosphate. *Biochemistry* **10**, 2156-2162.
- Wang, X. J., Hart, S. A., Xu, B., Mason, M. D., Goodell, J. R. & Etzkorn, F. A. (2003). Serine-Cis-Proline and Serine-Trans-Proline Isosteres: Stereoselective Synthesis of (Z)- and (E)-Alkene Mimics by Still-Wittig and Ireland-Claisen Rearrangements. *J. Org. Chem.* **68**, 2343-2349.
- Weber, R. W. & Nitschmann, H. (1978). Der Einfluss Der O-Acetylierung Auf Das Knoformative Verhalten Des Kollagen-Modell Peptides (L-Pro-L-Hyp-Gly)₁₀ Und Von Gelatine. *Helv. Chim. Acta* **61**, 701-708.
- Weiss, M. S., Jabs, A. & Hilgenfeld, R. (1998). Peptide Bonds Revisited. *Nat. Struct. Biol.* **5**, 676.
- Werkmeister, J. A. & Ramshaw, J. A. M., Eds. (1992). Collagen Biomaterials. Barking, Essex, England, Elsevier Science.
- Wiberg, K. B. & Laidig, K. E. (1987). Barriers to Rotation Adjacent to Double Bonds. 3. The C-O Barrier in Formic Acid, Methyl Formate, Acetic Acid, and Methyl Acetate. The Origin of Ester and Amide "Resonance". *J. Am. Chem. Soc.* **109**, 5935-5943.
- Wilchek, M. & Patchornik, A. (1964). The Synthesis of O-Acetylhydroxy- α -Amino Acids. *J. Org. Chem.* **29**, 1629-1630.

- Winter, A. D. & Page, A. P. (2000). Prolyl 4-Hydroxylase Is an Essential Procollagen-Modifying Enzyme Required for Exoskeleton Formation and the Maintenance of Body Shape in the Nematode *Caenorhabditis Elegans*. *Mol. Cell. Biol.* **20**, 4084-4093.
- Wipf, P. & Fritch, P. C. (1994). S_N2' -Reactions of Peptide Aziridines. A Cuprate-Based Approach to (E)-Alkene Isosteres. *J. Org. Chem.* **59**, 4875-4886.
- Wipf, P., Henninger, T. C. & Geib, S. J. (1998). Methyl- and (Trifluoromethyl)Alkene Peptide Isosteres: Synthesis and Evaluation of Their Potential as Beta-Turn Promoters and Peptide Mimetics. *J. Org. Chem.* **63**, 6088-6089.
- Xu, Y., Bhate, M. & Brodsky, B. (2002). Characterization of the Nucleation and Folding of a Collagen Triple Helix Peptide. *Biochemistry* **41**, 8143-8151.
- Xu, Y., Hyde, T., Wang, X., Bhate, M., Brodsky, B. & Baum, J. (2003). NMR and CD Spectroscopy Show That Imino Acid Restriction of the Unfolded State Leads to Efficient Folding. *Biochemistry* **42**, 8696-8703.
- Yang, J. X., Spek, E. J., Gong, Y. X., Zhou, H. X. & Kallenbach, N. R. (1997). The Role of Contest on Alpha-Helix Stabilization: Host-Guest Analysis in a Mixed Background Peptide Model. *Protein Sci.* **6**, 1264-1272.
- Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. (1997). Glyproarg Confers Stability Similar to Glyprohyp in the Collagen Triple Helix of Host-Guest Peptides. *J. Biol. Chem.* **272**, 28837-28840.
- Zou, Q., Bennion, B. J., Daggett, V. & Murphy, K. P. (2002). The Molecular Mechanism of Stabilization of Proteins by Tmao and Its Ability to Counteract the Effects of Urea. *J. Am. Chem. Soc.* **124**, 1192-1202.