CONFORMATIONAL AND MECHANISTIC STUDIES OF ASPARAGINE-LINKED GLYCOSYLATION

Thesis by

Karen Lynn Shannon

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1992

(June 1, 1992)

She got here red, all wrinkled, tired, and crying,
And she took her up and held her to breast,
And she sure was glad to get what mama offered,
Then she went to sleep and put her fears to rest.

It didn't seem to matter what she needed,
She could always count on mama to supply,
And regardless of the sleep she might be losing,
She always found a twinkle in her eye.

There ought to be a hall of fame for mamas,
Creation's most unique and precious pearl,
And Heaven help us always to remember,
That the hand that rocks the cradle
Moves the world.

He taught her all the attributes of greatness,
That he knew she wouldn't learn away from home.
And by the time he wore the cover off his Bible
His hair was gray
And his little girl was gone.

There ought to be a hall of fame for papas,
Creation's most unique and precious pearl,
And Heaven help us always to remember,
That the hand that rocks the cradle
Moves the world.

Yes, the hand that rocks the cradle
Moves the world.

to mom and dad

paraphrased from
"The Hand that Rocks the Cradle"
by Ted Harris

Acknowledgements

Many thanks go to my advisor, Barbara Imperiali, for her guidance and encouragement during the past five years. Barbara is solely responsible for my tremendous improvement as a research chemist. She also managed to convince me that traveling three thousand miles away from my mother into an earthquake zone (not to mention recent riot zone) would be fun and adventuresome. I really have appreciated her confidence in me, her friendship, and her advice. I also must congratulate her for her incredible fortitude as an assistant professor! I wish both you and Larry all the best.

I'm indebted to my fellow subgroup members for proofing this manuscript and for moral support this past year: Tamara Hendrickson, Keith Rickert, and Rahul Pathak. I have faith that you will all manage to live without me; however, you are welcome to use 1-900-Shannon whenever!

I can not begin to express my gratitude to the Kleenex crew: Anita Lingenfelter, Suzanne Biggs Fricke, Colleen Costello, Elizabeth Burns, and Sherrie Walsh Campbell. If I only had a copy of all the wacky wisdom and convoluted logic than you've used to cheer me up.....

I especially appreciated the advice and friendship of Caltech's finest sages: Dian Buchness, Alice Dennison, Lois MacBird, Emily Mazurek, and Margaret Collins. I was considering trying to blame you all for the pitiful nature of my finances; but even without the tips on art fairs and neat shops, I probably would have spent the money anyway.

To Louie and Sharon Borbon: I will never forget the 1991 NBA playoff games. I swear that next time we watch a Bulls and Lakers playoff game I will provide tranquilizers for you both and for Milan. I am also eternally indebted to you for generously refinancing Milan with your children's milk money. Seriously, thanks for EVERYTHING!

I have a great deal of respect for George Best; after all, he is one of the few who successfully manage to keep Milan in line. He also has great taste in music. Thanks for listening to all my gripes and for continually pointing out how much worse it could be ("look at it this way..."). May a Republican always be president.

To Milan: I've never had anyone serve me lobster bisque soup as a cold remedy before, nor have I ever seen anyone so fascinated by a cat (nor a cat so spellbound by a person), nor have I ever spent so many blissful evenings suffering through cop movies and sporting events. Thank-you for making me smile and for holding me when life became topsy-turvy.

To those whose lifetime chores will always include picking me up, brushing me off, and sending me back on my way: Mom, Dad, and Laura. I promise that I'll mellow out now (hah!).

It would be impossible for me to specifically thank all the host of other individuals who have provided support, encouragement, and advice over the past five years. I fear without any of you this thesis wouldn't exist.

Abstract

In asparagine-linked glycosylation, a complex carbohydrate is transferred by oligosaccharyl transferase to an asparagine in the consensus triad -Asn-Xaa-Thr/Ser-. The mechanism of this transfer is not fully understood but appears to involve nucleophilic attack by the carboxyamido group of asparagine at the anomeric carbon of the sugar with concomittant displacement of dolichol pyrophosphate. Central to understanding the mechanism of N-glycosylation is ascertaining how the reactivity of a relatively poor nucleophile, the carboxyamido group, can be enhanced.

To increase the amide-nitrogen nucleophilicity, the amide resonance interaction must be disrupted. The necessity for a hydroxy amino acid two residues from the asparagine suggests that this residue may participate in a hydrogen bonding interaction. If the hydroxyl group is hydrogen bonded to the amide oxygen of the asparaginyl side chain, this interaction would increase the pKa of the amide protons through disruption of the amide resonance. Further, a general base could then abstract a proton, thereby increasing the nucleophility of the amide nitrogen.

If such a hydrogen bonding array is crucial to the mechanism of the transfer, the conformation of the peptidyl substrate must allow for the asparagine and threonine side chains to be in close proximity.

A series of tripeptides that satisfy the -Asn-Xaa-Thr/Ser- primary sequence requirement were synthesized and examined as

potential acceptors in an oligosaccharyl transferase assay. It was demonstrated that the dominant solution conformation of substrates in glycosylation were Asx-turns. This conformation is stabilized by an interaction between the asparaginyl carboxyamido oxygen and the backbone amide proton of the hydroxy amino acid residue. The Asx-turn is a common structural motif in peptides and, in this case, brings the carboxyamido and hydroxyl groups within hydrogen bonding distance. Using constrained peptidyl analogs, the Asx-turn was demonstrated to be the bioactive conformation in N-glycosylation.

Further, peptidyl analogs were used to probe the mechanism of glycosylation. A mechanism which is consistent with both the conformational and kinetic data obtained in these studies is proposed.

Table of Contents

Acknowledgements iv

Abstract v

List of Figures ix

List of Tables xii

Abbreviations Used xiv

General Methods 1

Chapter 1: Background Information on Asparagine-Linked Glycosylation 3
Introduction 4
References 24

Chapter 2: Enzyme Assays 27
Introduction 28
Experimental Methods 36
Conclusions 45
Appendices 47
References 52

Chapter 3: Molecular Modelling and Amide Exchange Experiments on Hexapeptides 54
Introduction 55
Experimental Methods-Molecular Modelling 59
Results and Discussion 65
Experimental Methods-Amide Exchange 70
Results and Discussion 72
Conclusion 73
Appendices 75
References 79

Chapter 4: Kinetic and Structural Studies on Tripeptidyl Acceptors 81
Introduction 82
Experimental Methods 89
Results and Discussion 99

Conclusion 111
References 112

Chapter 5: Constrained Analogs 116
Introduction 117
Experimental Methods 125
Results and Discussion 134
Conclusion 148
Appendices 150
References 157

Chapter 6: Mechanistic Implications 159
Introduction 160
Experimental Methods 170
Results and Discussion 175
Conclusion 182
Appendices 183
References 187

NMR Spectra 188

List of Figures

Chapter 1: Background Information on Asparagine-Linked Glycosylation

1-1	Sugar residues found in glycoproteins	5
1-2	Examples of N-linked (a) and O-linked (b) sugar chains of glycoproteins	6
1-3	Cellular location of events in the biosynthesis of glycoproteins	9
1-4	Structure of dolichol phosphate	11
1-5	Dolichol phosphate cycle	12
1-6	Hydrogen bonding array proposed by Marshall	20
1-7	Overall reaction catalyzed by oligosaccharyl transferase	20
1-8	Hydrogen bonding array proposed by Bause	21
1-9	Proposed reaction mechanism for the inhibition of oligosaccharyl transferase by an epoxide containing peptide	22

Chapter 2: Enzyme Assays

2-1	Truncated oligosaccharide donor, Dol-P-P-$GlcNAc_2$	29
2-2	Biosynthesis of the truncated oligosaccharide donor	31
2-3	Synthesis of the truncated oligosaccharide donor	32
2-4	Determination of the oligosaccharide composition of hydrolyzed Dol-P-P-GlcNAc-[^{3}H]GlcNAc	33
2-5	Glycerol dependence of oligosaccharyl transferase	42
2-6	DMSO dependence of oligosaccharyl transferase	43

Chapter 3: Molecular Modelling and Amide Exchange Experiments on hexapeptides

3-1	Hydrogen bonding array proposed by Marshall	57
3-2	Bause and Legler's proposed hydrogen bond array	59
3-3	Timeline of simulated annealing experiment	62
3-4	Dihedral angles measured during simulated dynamics	63
3-5	Ramachandran plots of the lowest energy conformers of Ac-PNGTAV-NH_2, Ac-ANGTAV-NH_2, and Ac-YNPTSV-NH_2	64
3-6	Bond distances measured during simulated dynamics	65
3-7	Percent distribution of bond distances within the lowest energy conformers of Ac-PNGTAV-NH_2	67

3-8	Distribution of bond distances observed for bonds 1 and 2 within the lowest energy conformers of Ac-PNGTAV-NH_2 and Ac-ANGTAV-NH_2	69
3-9	Amide proton exchange rates of Ac-ANGTAV-NH_2	73

Chapter 4: Kinetic and Structural Studies on Tripeptidyl Acceptors

4-1	Representation of a type I β-turn	83
4-2	The two possible β-turn conformations of Ac-Asn-Ala-Thr-NH_2	85
4-3	Representation of the (*i*+4) loop	87
4-4	Representation of an Asx-turn	87
4-5	The possible Asx-turn structure of Ac-Asn-Ala-Thr-NH_2	88
4-6	ROESY spectrum of Ac-Asn-Ala-Thr-NH_2	103
4-7	ROESY spectrum of Ac-Asn-(D)Ala-Thr-NH_2	105
4-8	ROESY spectrum of Ac-Asn-AIB-Thr-NH_2	107
4-9	ROESY spectrum of Ac-Asn-Pro-Thr-NH_2	108
4-10	Graphical representations of connectivity and hydrogen-bonding information derived from spectroscopic studies of six acetyl-protected peptides	104

Chapter 5: Constrained Analogs

5-1	Constrained peptidyl analogs formed via cyclizations	118
5-2	Predominant solution conformation of a cyclic hexapeptides	119
5-3	Predominant solution conformation of cyclo(Pro-D-Tyr(OBzl)-Gly-Ile-Leu-Gln)	120
5-4	The two possible cyclic hexapeptides which incorporate both a Pro-D-Tyr and an Asn-Xaa-Thr sequences	121
5-5	Energy minimized conformation of cyclo(CYNC)TSV	123
5-6	Representations of cyclo(AsnApm)ThrNMe	124
5-7	Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PY_dNGTL)	135
5-8	Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PY_dNGTL)	136
5-9	Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PF_dNNAT)	138

5-10	Glycosyl acceptor abilities of Bz-NLT-NHMe, cyclo(PY_dNGTL), cyclo(PY_dLNGT), and cyclo(PF_dNNAT)	139
5-11	ROESY spectrum of cyclo(CYNCTSV)	141
5-12	Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(CANCTSA)	142
5-13	Macromodel representations of the dimeric pimelic macrocycle	144
5-14	ROESY spectrum of the dimeric pimelic macrocycle	145
5-15	Graphical representation of NOE connectivities and the hydrogen bound amides (*) derived from spectroscopic studies of the pimelic macrocycle	146
5-16	Representation of the linear pimelic dimer, Boc-NX(NXT-OMe)T-OMe	147

Chapter 6: Mechanistic Implications

6-1	Overall reaction catalyzed by oligosaccharyl transferase	160
6-2	Concerted mechanism	161
6-3	Two step mechanism	162
6-4	Proposed reaction pathway catalyzed by lysozyme	163
6-5	Structure of tri-N-acetylglucosaminyl-N-acetylglucosamine lactone (R=GlcNAc)	164
6-6	Bause's proposed activation of the carboxyamido nitrogen	165
6-7	An alternate hydrogen bonding array proposed by Marshall	166
6-8	Reaction mechanism proposed by Clark *et al.*	168
6-9	The hydrogen bonding array which is conformationally consistent with an Asx-turn	169
6-10	Inhibition study of Bz-NLT-NHMe by Bz-NLV-NHMe	177
6-11	Proposed mechanism of N-linked glycosylation consistent both with conformational and mechanistic data	180
6-12	pKa values of a) Asn, b) Tan, and c) Amb	181

List of Tables

Chapter 1: Background Information on Asparagine-Linked Glycosylation

1-1 Glycosyl acceptor properties of unglycosylated acceptor sites in known glycoproteins 17

Chapter 2: Enzyme Assays

2-1 Donor specificity of oligosaccharyl transferase 30

2-2 Efficiency of various detergents in solubilizing Dol-P-P-GlcNAc 39

2-3 Effect of Mg(II) on the activity of enzyme II 39

2-4 Relative acceptor activities for the glycosylation of blocked tripeptides 44

Chapter 3: Molecular Modelling and Amide Exchange Experiments on hexapeptides

3-1 Relative acceptor activities for the glycosylation of peptides by hen oviduct microsomes 56

3-2 Glycosyl acceptor activity of hexapeptides containing modifications of the hydroxyl amino acid 57

3-3 Glycosyl acceptor properties of hexapeptides containing proline within and surrounding the marker sequence 60

Chapter 4: Kinetic and Structural Studies on Tripeptidyl Acceptors

4-1 Enzyme assay results from tripeptide studies 100

4-2 Temperature dependence of amide proton chemical shifts ($-\Delta d/\Delta T$) for Ac-Asn-Xaa-Thr-NH_2 101

4-3 Coupling constants for Ac-Asn-Xaa-Thr-NH_2 102

Chapter 5: Constrained Analogs

5-1 Glycosyl acceptor properties of linear and cyclic peptides 122

5-2 Temperature dependence of the amide proton chemical shifts ($-\Delta d/\Delta T$) for cyclo(PY_dNGTL) and cyclo(PY_dLNGT) in DMSO 135

5-3 Enzyme assay results for the cyclic hexapeptides 139

5-4 Temperature dependence of the amide proton chemical shifts ($-\Delta d/\Delta T$) for cyclo(CANCTSA) and cyclo(CYNCTSV) in DMSO 140

5-5 Enzyme assay results for the cyclic hexapeptides. 142

5-6 Enzyme assay results from dimeric studies 147

Chapter 6: Mechanistic Implications

6-1 Glycosyl acceptor capabilities of hexapeptides in the series Tyr-Asn-Gly-Xaa-Ser-Val 165

6-2 Glycosyl acceptor properties of tripeptides containing sterically modified hyroxyl amino acids within the series Bz-Asn-Leu-Xaa-NHMe 176

6-3 Glycosyl acceptor properties of tripeptides containing sterically modified asparagine residues 178

6-4 Glycosyl acceptor properties of tripeptides containing asparagine analogs with altered proton donor behavior 181

Abbreviations Used

Standardard one and three letter codes are used for naturally occurring amino acid residues.

Ac:	acetyl
Aib, or B:	α-aminoisobutyric acid
Apm:	α-aminopimelic acid
Bn:	benzyl
Boc:	*t*-butoxycarbonyl
Bz:	benzoyl
Dol:	dolichol
Dol-P:	dolichol phosphate
Dol-P-P:	dolichol pyrophosphate
Fmoc:	9-fluorenylmethoxycarbonyl
mRNA:	messenger ribonucleic acid
UDP:	uridine 5'-diphosphate
UMP:	uridine 5'-monophosphate
Xaa:	any naturally occurring amino acid

GENERAL METHODS

Nuclear magnetic resonance spectra were recorded on a Bruker WM-300 (^{1}H, 300 MHz; ^{13}C, 75 MHz) or a Bruker AMX-500 (^{1}H, 500 MHz; ^{13}C, 125 MHz) spectrophotometer. ^{1}H and ^{13}C NMR shifts are reported in parts per million (ppm) downfield from tetrametylsilane as an internal standard on the δ scale. All 2-D ROESY spectra were recorded on a Bruker AMX-500 (500 MHz).

Analytical thin layer chromatography (TLC) was performed with 0.25 mm thick E. Merck pre-coated plates, silica gel 60 F-254. Preparative TLC (PTLC) separations were carried out on 1.0 mm thick E. Merck pre-coated plates, silica gel 60 F-254.

Flash chromatography was carried out using Baker 40 μm silica gel according to the general procedures of Still *et al.* (**1**).

Reagents and solvents were commercial grade and used as supplied. Where anhydrous conditions were employed, solvents were dried before use following the general procedures reported by Perrin and Armarego (**2**).

Liquid scintillation counting was performed on a Beckman LS 5000 TD and reported as disintegrations per minute (dpm). Aqueous samples were counted with Ecolite (ICN) scintillation fluid, organic samples were counted in Betamax (ICN), and proteinaceous pellets were dissolved first in Solvable (DuPont) before counting in Formula 989 (DuPont) scintillation fluid.

High pressure liquid chromatography (HPLC) was performed on a Beckman dual wavelength HPLC.

Centrifugation was performed on a Beckman Superspeed LS-35 equipped with a cooling system.

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

2 Perrin, D.D.; Armarego, W.L.F. *Purification of Laboratory Chemicals*, 3rd ed., Pergamon Press: New York, 1988.

Chapter 1
Background Information on Asparagine-Linked Glycosylation

Glycoproteins are a diverse group of biopolymers which are ubiquitous constituents of cells. They differ from other proteins by having oligosaccharide chains covalently attached to the polypeptide. Glycoproteins range in molecular weight from 15,000 to over 1 million, usually they contain fifteen or fewer sugar units per covalently attached oligosaccharide chain, and their carbohydrate content can range from 1-85% by weight. Glycoproteins are present in most organisms, including: plants, bacteria, fungi, viruses, and animals where they serve a variety of functions. For example, salivary mucins have a lubricating function which is related to their elasticity. In other glycoproteins, the carbohydrate units serve to protect the peptide backbone from proteolysis and denaturation. Of major interest is the apparent role of membrane bound glycoproteins (and glycolipids) located on the outer cell surface to serve as receptors. These membrane-bound proteins may influence cell-cell interactions, and thus be of importance in growth, differentiation, and malignancy. Sugars located at the cell surface also appear to function in cell adhesion, mitogenic stimulation, cytotoxic response, as well as a variety of other biological phenomena.

Nine different sugar residues are found in the oligosaccharide chains attached to the glycoproteins (Figure 1-1).

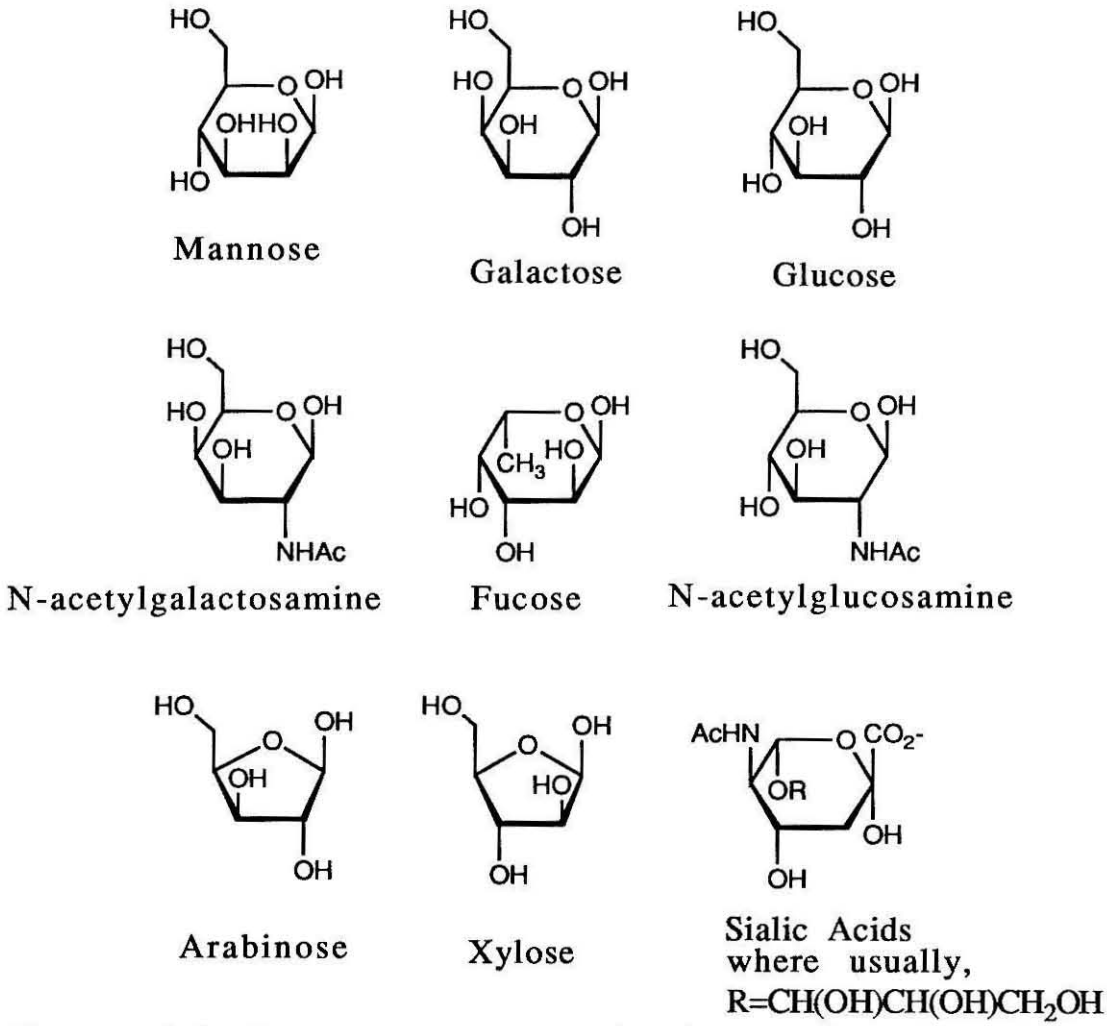


Figure 1-1 Sugar residues found in glycoproteins.

Glucose (Glc) is found mainly in collagen, but galactose (Gal) and mannose (Man) are more common and are widely distributed. The two most frequently found sugars are N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc). Fucose (Fuc), arabinose (Ara), xylose (Xyl), and the sialic acids (Sial) are also found. In general, N-acetylhexosamines are found near the attachment site of the oligosaccharide chain to the protein, whereas sialic acid residues are located more distal in the chain, frequently at terminal sites.

The oligosaccharide chains are attached to the polypeptide backbone of glycoproteins at one of five amino acid residues: asparagine, serine, threonine, hydroxylysine, or hydroxyproline. Depending on the attachment site of the oligosaccharide, the glycosyl-peptide bond can be either of the O- or N-type. Examples of both types of glycoproteins are shown in Figure 1-2 (1).

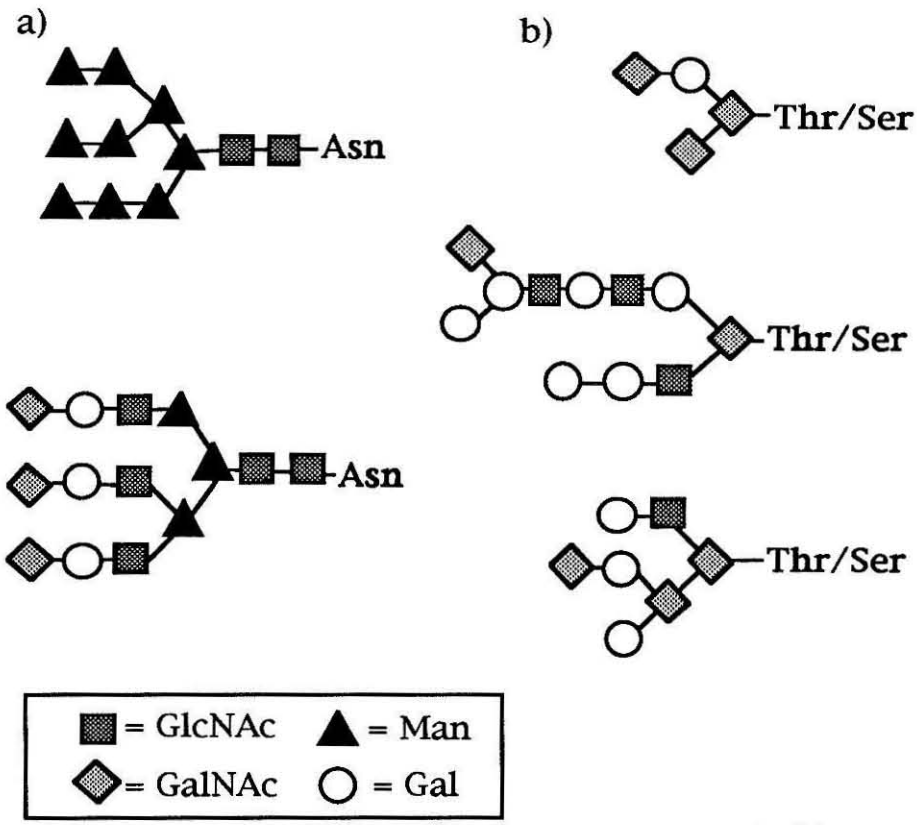


Figure 1-2 Examples of N-linked (a) and O-linked (b) sugar chains of glycoproteins. Structures were selected to contrast the two classes of sugar chains, which show differences in the core sequence, and similarities in the peripheral sequences.

These examples were selected to contrast the two types of glycoproteins. N-linked glycoproteins share a common $Man_3GlcNAc_2$ core sequence (shown in bold) which is attached to an asparagine

residue within a polypeptide. In contrast, there is no single, common carbohydrate core in O-linked glycoproteins. The core sugar residue can be attached to the hydroxyl group of threonine, serine, hydroxyproline, or hydroxylysine. Further, this core sugar can be any of the nine monosaccharides. Sequence diversity is also acquired through extension of the oligosaccharide chains attached to the core. For both types of glycoproteins, this extension commonly occurs from the use of a (Galβ1,4)$_3$GlcNAcβ1,3 sequence repeat (shown in plain text) which is decorated with various terminal sugars (shown in italics). Typically the terminal sugars are α-linked Sial, Fuc, Gal, GlcNAc residues, attached either singly or in combination to one or more hydroxyl groups of the core structure. Branching can also occur by the attachment of extension sequences to galactose residues at both the 3 and 6 positions (**2**).

Glycoproteins are synthesized by the same machinery that produces all other proteins. The biosynthesis event occurs within an intracellular membrane system composed of the endoplasmic reticulum, transfer vesicles, the Golgi apparatus, and secretory vesicles. During translation, glycosylation, processing, and transport, glycoproteins are completely isolated from the cytoplasm.

Our current understanding of N-linked glycoprotein translation is that it begins with the attachment of mRNA to ribosomes in the cytoplasm (Figure 1-3)(**3**). The attachment of free ribosomes to the ER membrane is mediated by a signal recognition particle. A signal sequence (containing about 16 residues) is translated and serves to guide the forming protein through the membrane bilayer into the lumen of the ER. This sequence is subsequently cleaved. Transfer of

preformed oligosaccharide units from a lipid intermediate to asparagine residues, to form N-glycosidic linkages, probably occurs very soon after the polypeptide emerges on the lumen side of the ER.

The glycoprotein then migrates to the Golgi apparatus where further trimming and elongation of the oligosaccharide moiety occurs as well as attachment of oligosaccharide units to hydroxyl residues to form O-glycosidic linkages. Once elongation is completed within the Golgi apparatus, the mature glycoprotein migrates towards the plasma membrane within a transport vesicle. The vesicle membrane fuses with the plasma membrane and secretory glycoproteins are extruded from the cell. Membrane glycoproteins probably become part of the plasma membrane by a lateral diffusion process. These glycoproteins are inserted into the membrane asymmetrically, with their carbohydrate moieties located on the outside surface of the cell.

The biosynthesis of N-glycosidic units proceeds through a different mechanism from that of most of the O-glycosidic carbohydrate units, although the primary sugar donors in the synthesis of both classes of carbohydrate units are sugar nucleotides and the enzymes involved are glycosyltransferases. The simplest route involves the direct transfer of a carbohydrate residue from a sugar nucleotide complex to the polypeptide chain or onto a forming oligosaccharide component. This route predominates in O-glycosylation where it occurs as a post-translational modification in the Golgi apparatus. It is also seen in the later processing of N-glycoproteins.

The second route involves the formation of an intermediate carbohydrate-lipid complex, in which the lipid functions as a carrier

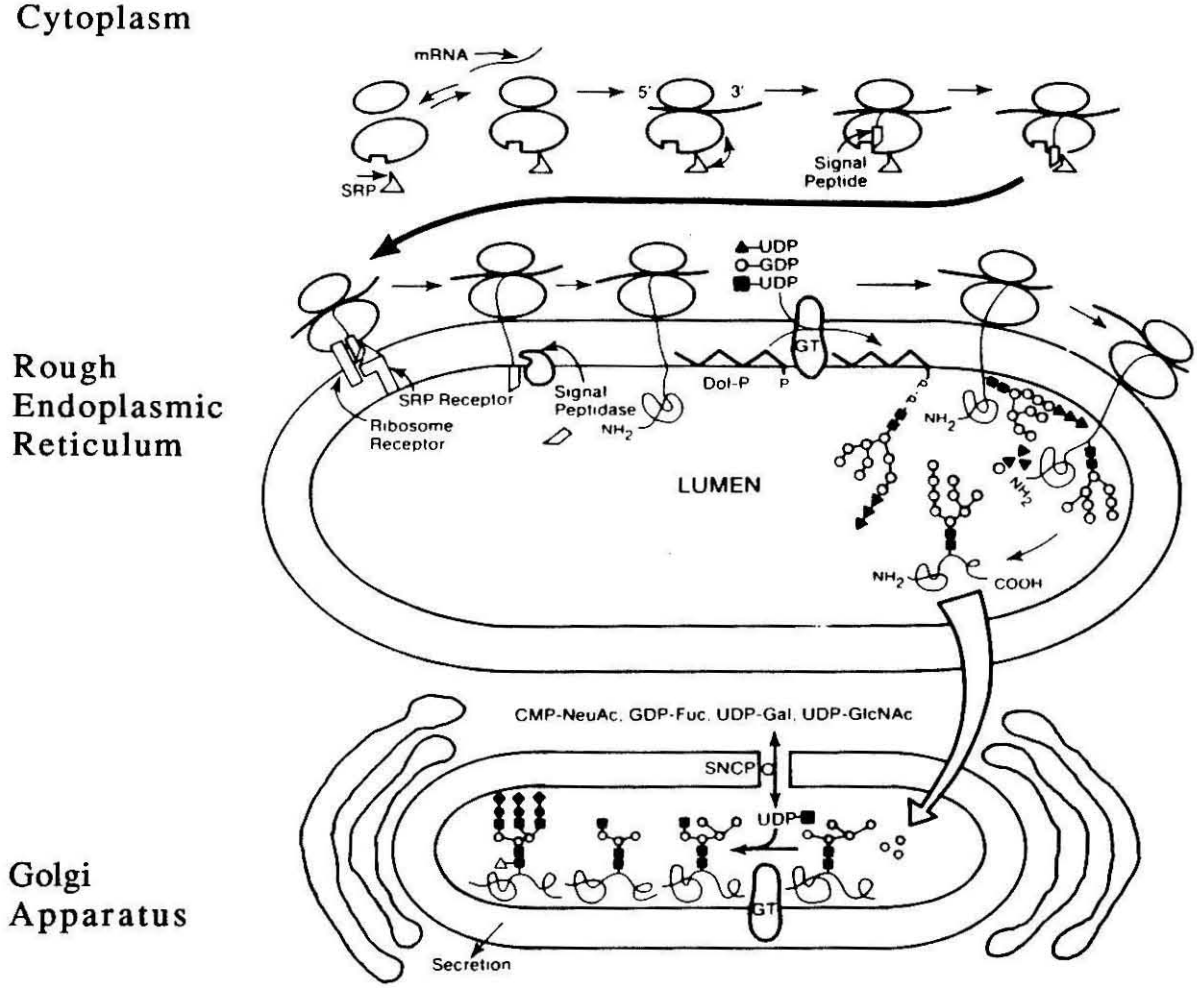


Figure 1-3 Cellular location of events in the biosynthesis of glycoproteins.

for the oligosaccharide chain synthesized on it. This route is used for N-glycosylation. Synthesis of N-glycosidic units can be divided into three distinct steps: 1) assembly of a lipid-linked oligosaccharide intermediate, 2) transfer of the oligosaccharide from the carrier to the growing polypeptide chain, and 3) trimming of the carbohydrate unit and addition of peripheral sugars.

For the synthesis of N-linked glycoproteins, the lipid linked oligosaccharide intermediate is $Glc_3Man_9(GlcNAc)_2$-P-P-Dolichol. This precursor unit is assembled by a series of reactions onto a lipid carrier, dolichol phosphate (**4**). Dolichol phosphate is a polyisoprene derivative located in the cell membrane. The number of isoprene units in the polyprenol is unusually high and varies from 17 to 21 in animals, from 14 to 24 in plants, yeast, and fungi, and from 10 to 12 in bacteria. The length of the fully extended form would be close to 100 Å. For comparison, oleic acid, a common constituent of phospholipids is only 25 Å. Clearly the biological conformation is very unlikely to be the extended form since the width of the ER membrane is generally 40 to 60 Å. The conformation has been shown to be comprised of a central coiled region flanked by two flexible arms (**5**). Dolichol serves to firmly anchor the carbohydrate close to the lumen side of the ER membrane near to the site of the emerging polypeptide. In animals, three of dolichol's isoprenol units have the trans orientation (Figure 1-4).

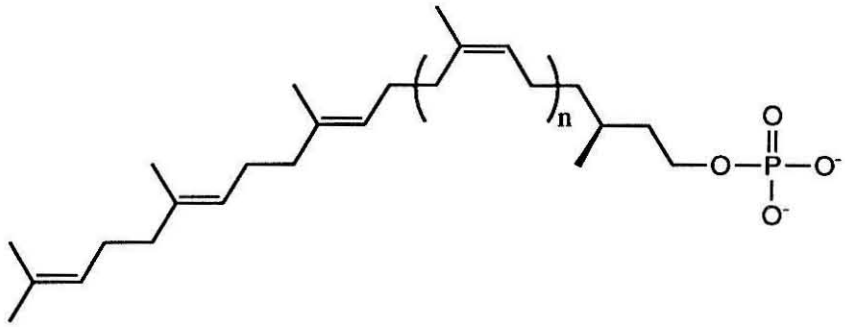


Figure 1-4 Structure of dolichol phosphate (n=9-15).

Also the α-isoprenol unit is saturated and of the two possible enantiomers, only the S-form of dolichol phosphate is involved in glycoprotein biosynthesis (**6**).

The assembly of the oligosaccharide begins with the reversible transfer of N-acetylglucosamine 1-phosphate from UDP-GlcNAc to Dolichol-P (Figure 1-5, **1**), followed by the irreversible transfer of N-acetylglucosamine, also from UDP-GlcNAc, to the GlcNAc-P-P-Dol (Figure 1-5, **2**) to give $(GlcNAc)_2$-P-P-Dol.

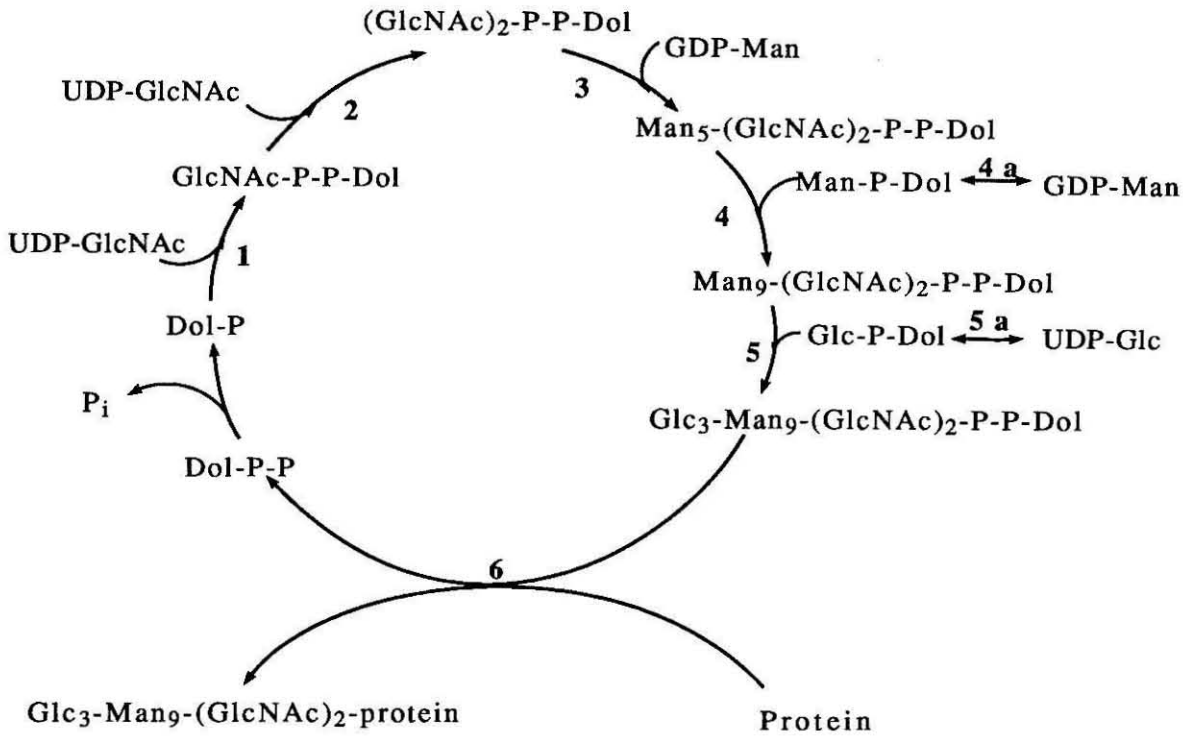


Figure 1-5 Dolichol phosphate cycle

The β-linked mannose residue of the pentasaccharide core is transferred directly from GDP-Man to $(GlcNAc)_2$-P-P-Dol (Figure 1-5, **3**). Attachment of the remaining mannose residues, as well as of the three glucoses, is believed to proceed via the corresponding dolichol phosphate derivatives (Figure 1-5, **4** & **5**), formed respectively from GDP-Man and UDP-Glc (Figure 1-5, **4a** & **5a**). The ultimate product of this cycle is $Glc_3Man_9(GlcNAc)_2$-P-P-Dol. The oligosaccharide is transferred *en bloc* from the lipid carrier to the nascent polypeptide (Figure 1-5, **6**). Thus the final step of the dolichol cycle results in the transfer of a mannose-rich oligosaccharide chain to a polypeptide.

Most of the reactions of the dolichol phosphate cycle proceed with inversion of configuration. However since GlcNAc-P-P-Dol is

formed by a condensation of two phosphates, the sugar-1-phosphate bond is not broken, and thus this sugar remains α-linked. Inversion of configuration of the bond between the oligosaccharide and the pyrophosphate must therefore occur upon transfer of the preformed oligosaccharide to the protein acceptor resulting in the formation of a GlcNAc-β-Asn linkage.

Strong evidence in support of the dolichol phosphate cycle comes from experiments with inhibitors of glycosylation. The most effective of these is tunicamycin, a glucosamine-containing antibiotic isolated from *Streptomyces lysosuperificus*, which inhibits the first step in the dolichol phosphate cycle, the formation of GlcNAc-P-P-Dol (**7**). Other inhibitors include 2-deoxyglucose, 2-fluoroglucose, and 2-fluoromannose (**8**). These inhibitors compete for available dolichol phosphate and lead to formation of altered intermediates. However, these compounds are not recognized by the glycosyl transferase and therefore inhibition of the overall reaction occurs as dolichol phosphate supplies become exhausted.

The transfer of the precursor oligosaccharide to the peptide substrate occurs co-translationally and involves a complex interplay among the nascent polypeptide chain that is being translated through the ER membrane, a polar oligosaccharide chain attached to an extremely hydrophobic lipid carrier, and the enzyme, oligosaccharyl transferase. The latter two components are embedded within the ER membrane.

Although general information about the overall pathway of N-linked glycosylation is available, our understanding of the properties of the oligosaccharyl transferase involved in this process have been

restricted by our inability to purify this enzyme away from its membrane environment. Attempts to purify this enzyme by classical techniques such as detergent solubilization followed by ion-exchange chromatography, gel filtration, etc., have led to complete loss of enzymatic activity (**9**). Presumably this loss of activity reflects that the enzymatic reaction occurs at the interface of the hydrophobic environment of the phospholipid bilayer of the ER. Recently it has been shown that oligosaccharyl transferase activity is associated with a complex composed of two ribophorins and a 48 kd protein (**10**). Thus it is not surprising that conventional manipulations during purification lead to a loss of activity, since these methods would disrupt the multi-subunit complex.

The known properties of oligosaccharyl transferase include a requirement for divalent cations (with Mn^{2+} being the most effective), an optimum pH between 7 and 7.5, and its association with the membrane for maintenance of enzymatic activity (**9**). Until the enzyme can be purified in its active form, specific information regarding the mechanism of the oligosaccharide transfer step must be gleaned from substrate studies of the lipid-linked oligosaccharide and of the polypeptide acceptor.

The affinity of oligosaccharyl transferase for oligosaccharyl-P-P-dolichol seems to diminish as the size of the oligosaccharide decreases. Thus, the "normal" substrate, $Glc_3Man_9GlcNAc_2$-P-P-Dol, is a more effective substrate than shorter versions (**11**). However, shorter lipid-linked oligosaccharides can be transferred to polypeptide acceptors using membrane fractions isolated from a variety of sources (**12-15**). In yeast membrane preparations, the

K_m value for $Glc_3Man_9GlcNAc_2$-P-P-Dol of 0.5 μM does not differ significantly from that for $GlcNAc_2$-P-P-Dol of 1.2 μM (**13**). Thus oligosaccharyl transferase seems to have a rather broad binding specificity with respect to the lipid-linked oligosaccharide.

It is unclear how the presence of the "normal" lipid-linked oligosaccharide substrate influences the recognition of the peptidyl substrate. Studies with oligosaccharyl transferase in yeast membrane preparations indicate that the structure of the lipid linked oligosaccharide affects the apparent K_m of the enzyme for the peptidyl substrate. The K_m of a peptide in the presence of "normal" lipid-linked oligosaccharide substrate, is ten fold lower than the K_m observed when the unnatural substrate, $GlcNAc_2$-P-P-Dol, is present (**13**).

Much more information is available concerning the specificity of oligosaccharyl transferase for the peptidyl substrate. It is now well established that a prerequisite for glycosylation is the occurrence of the acceptor asparagine in the tripeptide sequence, -Asn-Xaa-Thr/Ser-, where Xaa can be any of the 20 naturally occurring amino acids except proline (**16**). Only specific -Asn-Xaa-Thr/Ser- sites in a polypeptide are glycosylated. Most possible acceptor sites that become exposed on the lumenal side of the ER membrane are efficiently glycosylated, but some sites are never utilized. Gavel and von Heijne (**17**) have presented a detailed statistical study of glycosylated and non-glycosylated -Asn-Xaa-Thr/Ser- sites collected from the literature and from the NBRF-PIR database. Their study suggests that the -Asn-Xaa-Thr/Ser- signal leads to glycoslyation approximately thirty percent of the time.

Their study also demonstrated that -Asn-Xaa-Thr/Ser- sites that are located near the C-terminus or which contain proline on either side of the hydroxyl amino acid residue are often not glycosylated.

Experiments with ovalbumin have shown that co-translational glycosylation cannot take place until about 45 residues have been added beyond the glycosylation site (**18**). Therefore the bias of glycosylated -Asn-Xaa-Thr/Ser- sites towards the N-terminus may indicate that glycosylation occurs more easily when the nascent polypeptide is still firmly anchored in the ER membrane. Alternatively, the higher percentage of acceptor sites at the N-terminus may reflect that N-terminal sites spend more time in the lumen of the ER where both the sugar substrate and the transferase are situated. Also, C-terminal sites may be inaccessible to glycosylation if significant folding of the polypeptide has occurred during the translation event (**19**).

Further corroboration that accessibility of the acceptor site is important for glycosylation was provided by studies with polypeptide fragments of known glycoproteins which contained unglycosylated acceptor sites in their native form (Table 1-1) (**20**).

Table 1-1 Glycosyl acceptor properties of unglycosylated acceptor sites in known glycoproteins.

protein	potential site	acceptor upon denaturation?
Ovalbumin	-N_{311}LS-	yes
α-Lactalbumin	-N_{45}QS-	yes
	-N_{74}IS-	yes
RNase A	-N_{34}LT-	yes
DNase	-N_{103}DS-	yes
Prolactin	-N_{31}LS-	yes
Triosephosphate isomerase	-N_{195}VS-	yes
DNase	-N_{103}DS-	no
Catalase	-N_{242}LS-	no
	-N_{437}VT-	no
	-N_{479}FS-	no
Concanavalin A	-N_{118}ST-	no
	-N_{162}GS-	no
Elastase	-N_{66}GT-	no
	-N_{123}NS-	no
	-N_{215}VT-	no
Glyceraldehyde-3P-dehydrogenase	-N_{146}AS-	no
Trypsinogen	-N_{151}SS-	no
Carboxypeptidase A	-N_{90}PS-	no
Alcohol dehydrogenase	-N_{300}LS-	no

In all thirteen cases examined, it was shown that the polypeptides were inactive as acceptors in their native state. In contrast, after denaturation, reduction, and derivatization of the sulfhydryl groups, six of the thirteen fragments were acceptors in oligosaccharyl transferase assays. This suggests that the polypeptide must be

unfolded to serve as an acceptor. However, mere disruption of the tertiary structure is not always sufficient to allow enzymatic glycosylation. The remaining seven peptides were not substrates. No correlation is evident between enzymatic activity and the distribution of amino acids surrounding the -Asn-Xaa-Thr/Ser- site. However, cleavage of concanavalin A and catalase with cyanogen bromide provided mixtures of fragments with acceptor properties (**21**) suggesting that the inability of these polypeptides to be glycosylated may be due to blockage of the acceptor site by folding of the polypeptide and supporting the hypothesis that secondary structure plays a critical role in glycosylation.

Further studies with synthetic peptides have demonstrated that the conformation of the polypeptide is a fairly crucial element of acceptor ability. Synthetic peptides have been shown to serve as acceptors for dolichol-dependent glycosylation catalyzed by microsomal fractions from pig thyroid (**14**), rat liver (**15**), and hen oviduct (**12**). Synthetic peptides containing as few as three amino acid residues can serve as substrates provided that the carboxyl and amino termini are blocked (**15**). The affinity of oligosaccharyl transferase increases with the length of the polypeptide acceptor.

The role of the central amino acid residue, Xaa, in the acceptor sequence was examined by substituting this position with five different residue in the synthetic peptide, -Tyr-Asn-Xaa-Thr-Ser-Val- (**22**). It was found that similar levels of glycosylation occurred when Xaa was Gly, Leu, or Cys; whereas, the level was significantly lower when Xaa was Asp. and the peptide was not glycosylated when Xaa was Pro. Proline differs from other amino acids in two

significant aspects. First, it is an imino acid. Thus if the backbone amide proton of the central residue is important for binding, a peptide containing proline as the central residue could not bind. Second, proline's ring structure severely limits the torsional angles along its backbone. Thus it is also possible that a specific conformation is required for enzymatic activity and proline prevents the polypeptide from attaining this conformation.

The effects of conformational limitations on glycosylation were further studied by examining the enzymatic activity of peptides containing two cysteine residues in different positions within a polypeptide. Disulfide ring formation provided cyclic peptides of various sizes (**23**). This study demonstrated that the linear analogs were glycosylated more actively than their cyclized counterparts. Additionally, as the size of the ring expanded, the levels of glycosylation of the cyclic analogs approached those of their linear analogs. The formation of the disulfide bridge was apparently either blocking access to the acceptor site or preventing the polypeptide from adopting a conformation which could be recognized by the enzyme.

Conformation and accessibility are not the only factors which influence glycosylation. There is also an absolute requirement for a hydrogen-bonding moiety in the side chain of the third residue within the acceptor sequence. *In vitro*, the Asn residue in -Asn-Xaa-Thr- is glycosylated 40-fold more effectively than -Asn-Xaa-Ser-. *In vivo*, additional factors must affect the rate because the frequency of glycoproteins containing glycosylated -Asn-Xaa-Ser- sites is only 3-fold lower than -Asn-Xaa-Thr- sites (**24**). The H-bonding character

of the hydroxyl group is thought to play a crucial role in the interaction of the enzyme with the acceptor site. Marshall (**16**) proposed that the hydroxyl group of the hydroxyl amino acid could be hydrogen bound to the asparagine carboxyamido oxygen as shown in Figure 1-6a. He suggested that the function of this

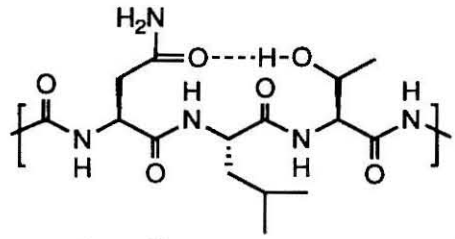


Figure 1-6 Hydrogen bonding array proposed by Marshall.

hydrogen bonding array is to stabilize the conformation which is recognized by the enzyme. However, the role of the hydroxyl group appears to be more extensive and seems to be directly involved in the mechanism of the transfer.

The overall reaction which is catalzyed by oligosaccharyl transferase is shown in Figure 1-7.

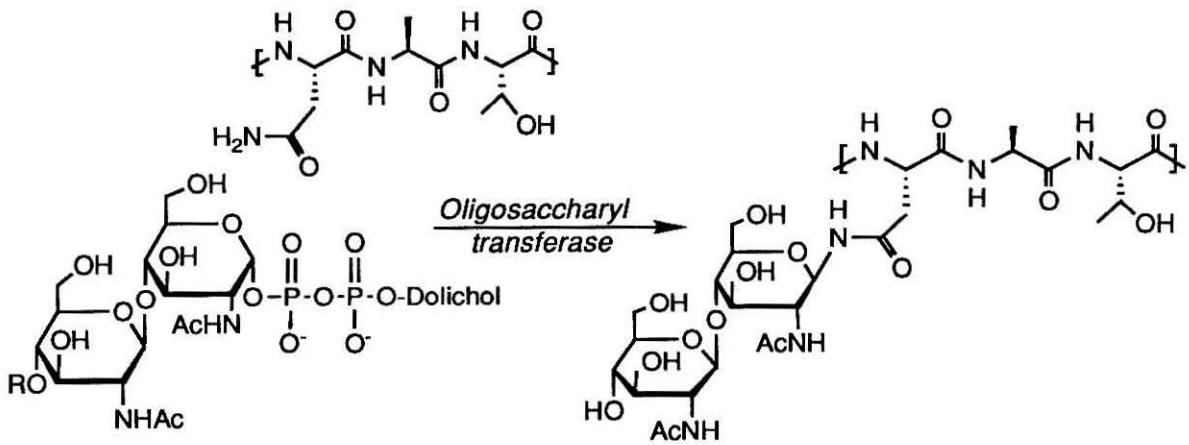


Figure 1-7 Overall reaction catalyzed by oligosaccharyl transferase (R=oligosaccharide).

This reaction is unusual because its mechanism appears to involve nucleophilic attack by the β-carboxyamido group of the Asn side chain at the anomeric carbon of the sugar with concomitant displacement of dolichol pyrophospate. The reaction results in formation of a N-glycosidic bond. Normally a β-carboxyamido nitrogen atom of the type found in an Asn residue would lack significant nucleophilic character. Therefore, it seems that the Asn nucleophilicity must be augmented prior to glycosylation.

Bause proposed an alternative model (Figure 1-8) to explain how the carboxyamido nitrogen could acquire nucleophilic properties.

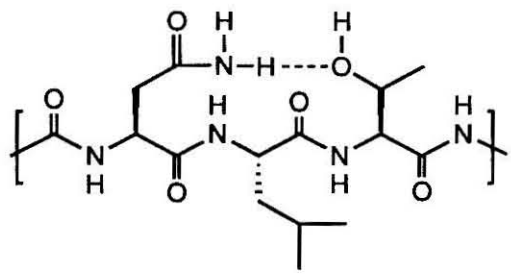


Figure 1-8 Hydrogen bonding array proposed by Bause.

Bause accepted that the presence of the hydroxyl group must be crucial to binding as proposed by Marshall, but he pointed out that in Marshall's model the orientation of the hydrogen bond results in a net decrease in the nucleophilicity of the nitrogen electron pair with a corresponding increase in N-H acidity. Bause proposed that the hydroxy amino acid side chain functions instead as a hydrogen bond acceptor and the amide side chain of the asparagine residue serves as a donor. As a result of this interaction the nucleophilicity of the

amide nitrogen is increased, providing a reactive acceptor for the glycoside.

In support of this model is the finding that a tripeptide containing epoxyethylglycine substituted for the hydroxy amino acid, is an irreversible inhibitor of oligosaccharyl transferase (**25**). In terms of Bause's model, the epoxide mimics the function of the hydrogen bond acceptor, but in this process a basic residue in the enzyme is alkylated, leading to enzyme inactivation (Figure 1-9).

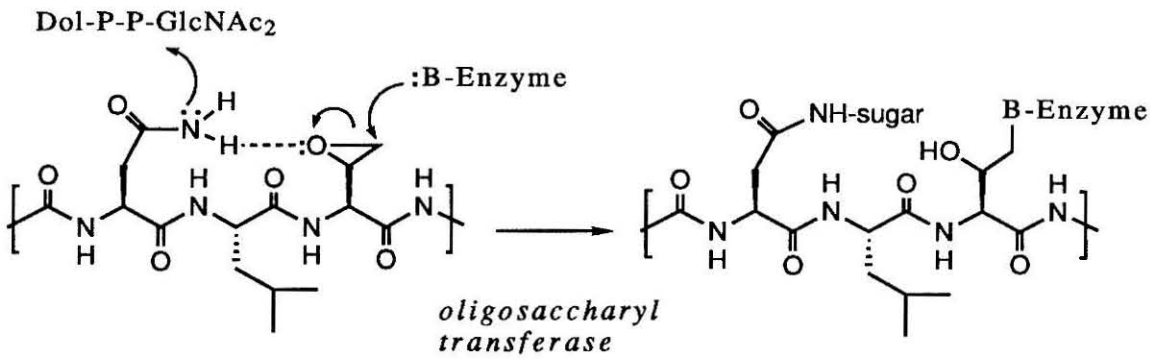


Figure 1-9 Proposed reaction mechanism for the inhibition of oligosaccharyl transferase by an epoxide containing peptide.

With respect to the other essential residue within the acceptor sequence, the asparagine side chain is an absolute requirement for recognition by oligosaccharyl transferase. Substitution of asparagine with glutamine or N-methylasparagine destroys the acceptor activity of the marker sequence (**24**). It has also been shown that synthetic peptides containing *threo*-β-fluroasparagine are extremely poor substrates for in vitro glycosylation, with V_{max}/K_m values only 1% of the analogous Asn-containing peptides (**27**).

From the foregoing account, it is clear that enormous progress has been made in our knowledge of the chemistry and biology of

glycoproteins. Not surprisingly, these findings have raised new questions. Of particular interest are those questions regarding the biosynthesis of the N-glycoproteins. The work described within this thesis was aimed at answering two questions. 1) What is the specific conformation of the polypeptide acceptor which the enzyme recognizes? and 2) What is the mechanism by which the carboxyamido nitrogen of the asparagine side chain becomes activated for attack on the lipid-linked oligosaccharide?

Our approach to answering the first question was to examine a series of flexible peptides which differed by subtle perturbations in their conformation. These peptides were examined for substrate properties with an enzyme assay (described in chapter 2). Computational modelling of hexapeptides with known glycosyl acceptor properties are reported in chapter 3. One- and two-dimensional NMR techniques were employed to ascertain common structural features of the acceptors and the non-binders. The identification of the probable recognition motif, an Asx-turn, is discussed in chapters 4. To further confirm that this motif was indeed the recognition element, constrained peptidyl analogs were synthesized and examined (chapter 5).

In approaching the second question, a mechanism which is consistent with features of this conformational motif is presented in chapter 6. Peptides in which the asparagine residue was replaced with several unnatural residues were used as mechanistic probes.

1. Adapted from, Paulson, J.C. "Glycoproteins: What are the Sugar Chains For?" *Trends in Biol. Sci.* **1989**, *14*, 272-276.

2. Sadler, J.E. In *Biology of Carbohydrates*, 2nd ed.; Ginsburg, V.; Robbins, P.W., Ed.; John Wiley & Sons: New York, 1984; Vol. 2, pp. 87-161.

3. From, Presper, K.A.; Heath, E.C. In *The Enzymology of Post-translational Modification of Proteins*, Freedman, R.B.; Hawkins, H.C., Ed.; Academic Press: New York, 1985; Vol. 2, p. 88.

4. Reviewed by a) Parodi, A.J.; Leloir, L.F. "The Role of Lipid Intermediates in the Glycosylation of Proteins in the Eukaryotic Cell," *Biochim. Biophys. Acta* **1979**, *559*, 1-37. b) Sharon, N.; Lis, H. In *The Proteins*, 3rd ed., Neurath, H.; Hill, R.L., Ed.; Academic Press: London, 1975, Vol. 5; pp. 1-144.

5. Murgolo, N.J.; Patel, A.; Stivala, S.S.; Wong, T.K. "The Conformation of Dolichol," *Biochemistry* **1989**, *28*, 253-260.

6. Low, P.; Peterson, E.; Mizuno, M.; Takigawa, T.; Chojnacki, T.; Dallner, G. "Reaction of Optically Active S- and R-forms of Dolichyl Phosphates with Activated Sugars," *Biosci. Rep.* **1986**, *6*, 677-683.

7. Tkacz, J.S.; Lampen, J.O. "Tunicamycin Inhibition of Polyisoprenyl N-Acetylglucosaminyl Pyrophosphate Formation in Calf-Liver Microsomes," *Biochem. Biophys. Res. Commun.* **1975**, *65*, 248-257.

8. Schwarz, R.T.; Datema, R. "Inhibitors of Protein Glycosylation," *Trends Biochem. Sci.* **1980**, *5*, 65-67.

9. Das, R.; Heath, E.C. "Dolichyldiphosphoryloligosaccharide-Protein Oligosaccharyl-Transferase: Solubilization, Purification, and Properties," *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 3811-3815.

10. Kelleher, D.J.; Kreibich, G.; Gilmore, R. "Oligosaccharyltransferase Activity is Associated with a Protein Complex Composed of Ribophorins I and II and a 48 kd Protein," *Cell* **1992**, *69*, 55-65.

11. Hubbard, S.C.; Ivatt, R.J. "Synthesis and Processing of Asparagine-Linked Oligosaccharides," *Annu. Rev. Biochem.* **1981**, *50*, 555-583.

12. Chen, W.W.; Lennarz, W.J. "Metabolism of Lipid-Linked N-Acetylglucosamine Intermediates," *J. Biol. Chem.* **1977**, *252*, 3473-3479.

13. Sharma, C.B.; Lehle, L.; Tanner, W. "N-Glycosylation of Yeast Proteins," *Eur. J. Biochem.* **1981**, *116*, 101-108.

14. Ronin, C.; Bouchilloux, S., Granier, C.; Van Rietschoten, J. "Enzymic N-Glycosylation of Synthetic Asn-X-Thr Containing Peptides," *FEBS Lett.* **1978**, *96*, 179-182.

15. Behrens, N.H.; Tabora, E. "Dolichol Intermediates in the Glycosylation of Proteins," *Methods Enzymol.* **1977**, *50*, 402-435.

16. Marshall, R.D. "The Nature and Metabolism of the Carbohydrate-Peptide Linkages of Glycoproteins," *Biochem. Soc. Symp.* **1974**, *40*, 17-26.

17. Gavel, Y.; von Heijne, G. "Sequence Differences Between Glycosylated and Non-glycosylated Asn-Xaa-Thr/Ser Acceptor Sites: Implications for Protein Engineering," *Protein Engineering* **1990**, *3*, 433-442.

18. Lennarz, W.J. "Protein Glycosylation in the Endoplasmic Reticulum: Current Topological Issues," *Biochemistry* **1987**, *26*, 7205-7210.

19. Beitema, J.J.; Gaastra, W.; Scheffer, A.J.; Welling, G.W. "Carbohydrate in Pancreatic Ribonucleases," *Eur. J. Biochem.* **1976**, *63*, 441-448.

20. Pless, D.D.; Lennarz, W.J. "Enzymatic Conversion of Proteins to Glycoproteins," *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 134-138.

21. Kronquist, K.E.; Lennarz, W. "Enzymic Conversion of Proteins to Glycoproteins by Lipid-Linked Saccharides: a Study of Potential Exogenous Acceptor Proteins," *J. Supramol. Struct.* **1978**, *8*, 51-56.

22. Bause, E.; Hettkamp, H. "Primary Structural Requirements for the Enzymatic Formation of the N-Glycosidic Bond in Glycoproteins," *FEBS Lett.* **1979**, *108*, 341-344.

23. Bause, E.; Hettkamp, H.; Legler, G. "Conformational Aspects of N-Glycosylation of Proteins," *Biochem. J.* **1982**, *203*, 761-768.

24. Bause, E.; Legler, G. "The Role of the Hydroxy Amino Acid in the Triplet Sequence Asn-Xaa-Thr(Ser) for the N-Glycosylation Step During Glycoprotein Biosynthesis," *Biochem. J.* **1981**, *195*, 639-644.

25. Bause, E. "Active-site Directed Inhibition of Asparagine N-Glycosyltransferases with Epoxy-Peptide Derivatives," *Annu. Rev. Biochem.* **1972**, *41*, 673-702.

26. Welply, J.K.; Shenbagamurthi, P.; Lennarz, W.J.; Naider, F. "Substrate Recognition by Oligosaccharyltransferase," *J. Biol. Chem.* **1983**, *258*, 11856-11863.

27. Rathod, P.K.; Tashjian, A.H.; Abeles, R.H. "Incorporation of Beta-Fluoroasparagine into Peptides Prevents N-Linked Glycosylation. In Vitro Studies with Synthetic Fluoropeptides," *J. Biol. Chem.* **1986**, *261*, 6461-6469.

Chapter 2

Enzyme Assays

The core portion of the carbohydrate sidechain of N-linked glycoproteins is biosynthesized by microsomal membrane-associated enzymes as a complex oligosaccharyldiphosphoryldolichol derivative and is subsequently transfered *en bloc* to specific asparagine residues within a forming polypeptide chain. Selected enzymes of the dolichol pathway have been studied in soluble form after extraction from microsomal membranes of a variety of tissues (**1**). However the enzyme involved in the final step in this series of reactions, the transferase responsible for catalyzing the attachment of the oligosaccharide to the asparaginyl residue of the protein to form the glycosamine linkage, has only been studied in intact membrane preparations. Information concerning the transferase has been restricted to kinetic studies of substrates using enzyme isolated as crude microsomal fractions.

The action of this enzyme was first demonstrated by observing the transfer of an oligosaccharide moiety from a radiolabelled lipid-linked oligosaccharide to denatured forms of three secretory proteins: ovalbumin, α-lactalbumin, and ribonuclease A (**2**). A membrane fraction isolated from hen oviduct was used as a source of the enzyme. The radiolabelled protein product was shown to have increased molecular weight, consistent with attachment of an oligosaccharide side chain.

Further attempts to characterize the transferase have utilized microsomes isolated from thyroid (**4**), yeast (**5**), and liver (**6**) as sources of the enzyme. Generally, transferase activity has been

demonstrated by determining the amount of radiolabel transfer that has occurred in mixtures of ^{3}H- or ^{14}C-labeled oligosaccharyldiphosphoryldolichol and unlabeled peptide acceptors. However these studies have been hampered by a lack of a reliable source of the radiolabeled oligosaccharide donor.

The oligosaccharide donor *in vivo* is $Glc_3Man_9(GlcNAc)_2$-P-P-Dol. This complex donor is biosynthesized by a series of reactions onto a lipid carrier, dolichol phosphate (**7**). However, *in vitro* the enzyme can also transfer a truncated substrate, a dolichyl diphosphoryl chitobiosyl unit (Figure 2-1) (**5**).

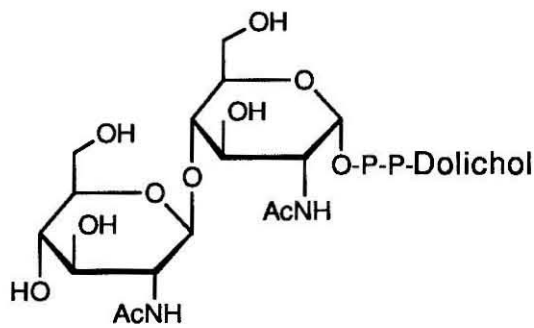


Figure 2-1 Truncated oligosaccharide donor, Dol-P-P-$GlcNAc_2$.

Table 2-1 summarizes the specificity of the transferase towards a variety of oligosaccharyl donors. The "natural substrate" [Dol-P-P-$GlcNAc_2Man_9Glc_3$], the disaccharide [Dol-P-P-$GlcNAc_2$], and the trisaccharide [Dol-P-P-$GlcNAc_2Man$] are similarly good substrates. The nonglucosylated lipid-linked oligosaccharide [Dol-P-P-$GlcNAc_2Man_9$] and the monosaccharide [Dol-P-P-GlcNAc] are not substrates.

Table 2-1 Donor specificity of oligosaccharyl transferase.

substrate	cpm/min	% transfer
Dol-P-P-$GlcNAc_2Man_9Glc_3$	2667	33.3
Dol-P-P-$GlcNAc_2Man_9$	51	0.6
Dol-P-P-$GlcNAc_2Man$	2415	30.1
Dol-P-P-$GlcNAc_2$	2089	26.1
Dol-P-P-GlcNAc	94	1.2

The transfer of radioactivity from the various donors (8000 cpm/min corresponding to 100%) was measured after 15 min incubation.

The observation that the enzyme does not transfer the nonglucosylated large lipid-linked oligosaccharide [Dol-P-P-$GlcNAc_2Man_9$] suggests that the terminal glucoses have an important role in binding. Curiously shorter oligosaccharides regain substrate behavior. It has been suggested that the terminal glucoses serve to induce a conformational change in the enzyme which increases affinity for the peptide substrate. Nonetheless, oligosaccharyl transferase seems to exhibit a rather broad specificity with respect to the lipid-linked oligosaccharide. Since the truncated donor is easier to synthesize, efforts were directed to securing a supply for use in future transferase studies.

The truncated donor can be biosynthesized *in vitro* by the action of two microsomal enzymes which sequentially transfer N-acetylglucosaminyl phosphate (Enzyme I) and GlcNAc (Enzyme II) to dolichyl phosphate (Figure 2-2). However utilization of this method for large scale preparation of radiolabeled oligosaccharide donor is costly as well as problematic due to the lack of a cheap source of

dolichol phosphate and due to product inhibition of Enzyme I by uridine monophosphate (UMP).

$$\text{Dol-P + UDP-GlcNAc} \xrightarrow{\text{Enzyme I}} \text{Dol-P-P-GlcNAc + UMP}$$

$$\text{Dol-P-P-GlcNAc + UDP-GlcNAc} \xrightarrow{\text{Enzyme II}} \text{Dol-P-P-GlcNAc}_2 \text{ + UDP}$$

Figure 2-2 Biosynthesis of the truncated oligosaccharide donor.

Two important developments have provided a more reliable supply of the truncated oligosaccharide donor. The first was the total synthesis of dolichol phosphate from plant polyprenols isolated from the leaves of the *Ginkgo biloba* tree (**8**). Synthetic dolichol phosphate produced in this manner was shown to have identical biological activity to the compound isolated from liver (available from Sigma, 5 mgs costing $217) and was of higher purity as determined by both ^{1}H NMR and TLC.

The second development was the efficient synthesis of the radiolabeled truncated donor as described by Imperiali and Zimmerman (Figure 2-3) (**9**).

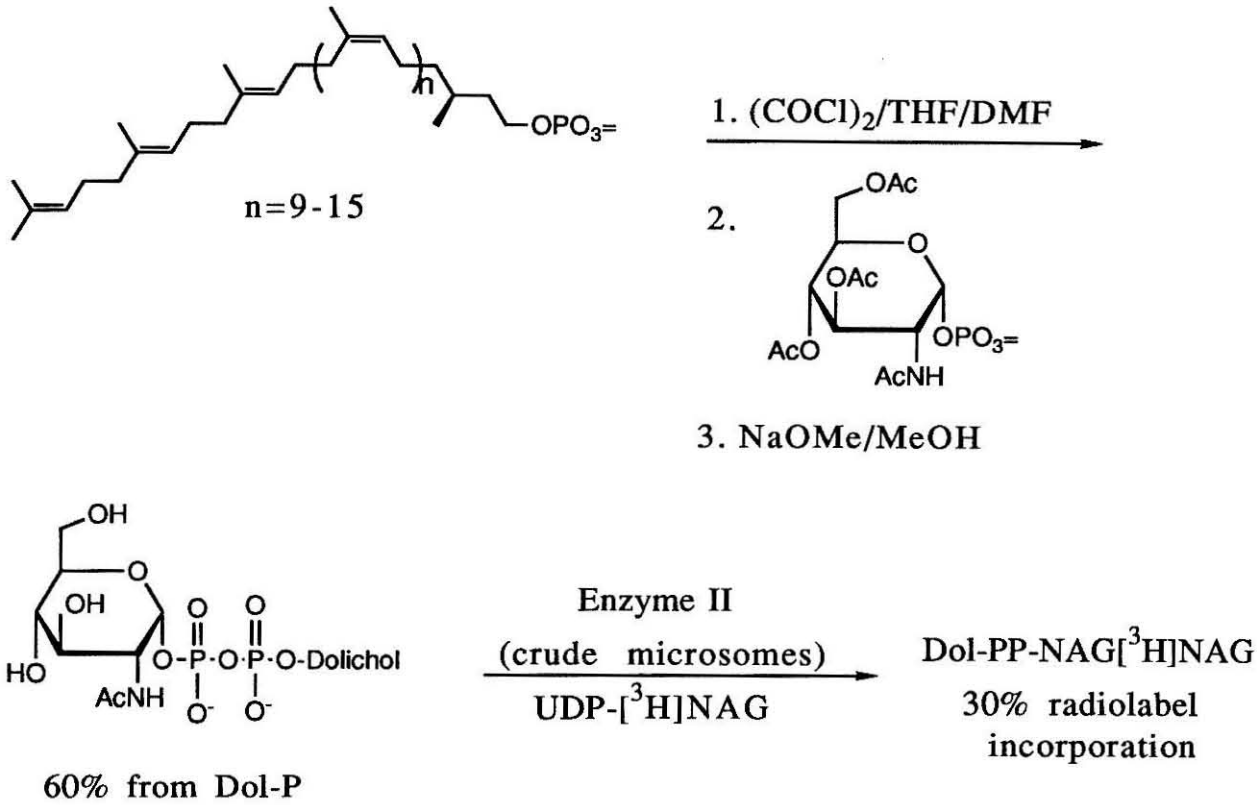


Figure 2-3 Synthesis of the truncated oligosaccharide donor.

To facilitate incorporation of N-acetylglucosaminyl phosphate, Dol-P-P-GlcNAc is chemically synthesized. Dolichol phosphate is first activated as the dichloroderivative. This produces a highly reactive species which can participate in a coupling reaction with per-O-acetyl-N-acetyl glucosaminyl phosphate. The dolichylpyrophosphoryl ester can subsequently be O-deacetylated with 3% anhydrous sodium methoxide. This monosaccharide derivative can be converted into the corresponding disaccharide derivative by reaction with UDP-GlcNAc (radiolabeled at C-6 with ^{3}H) in the presence of fresh porcine liver microsomes (described in detail

in the experimental section). Incorporation of the radiolabel typically proceeds in about 25-30%.

The sugar composition of purified lipid product has been verified by gel filtration. Purification was carried out by eluting diphosphoryl lipid compounds off DEAE-cellulose (acetate form) with 20 mM ammonium acetate. This eluent was concentrated and reapplied to a silica column to separate unreacted Dol-P-P-GlcNAc from Dol-P-P-GlcNAc$_2$. The purified product was characterized by acid catalyzed hydrolysis followed by gel filtration on a Biogel P-2 column calibrated with GlcNAc and chitobiose standards (Figure 2-4) (**10**).

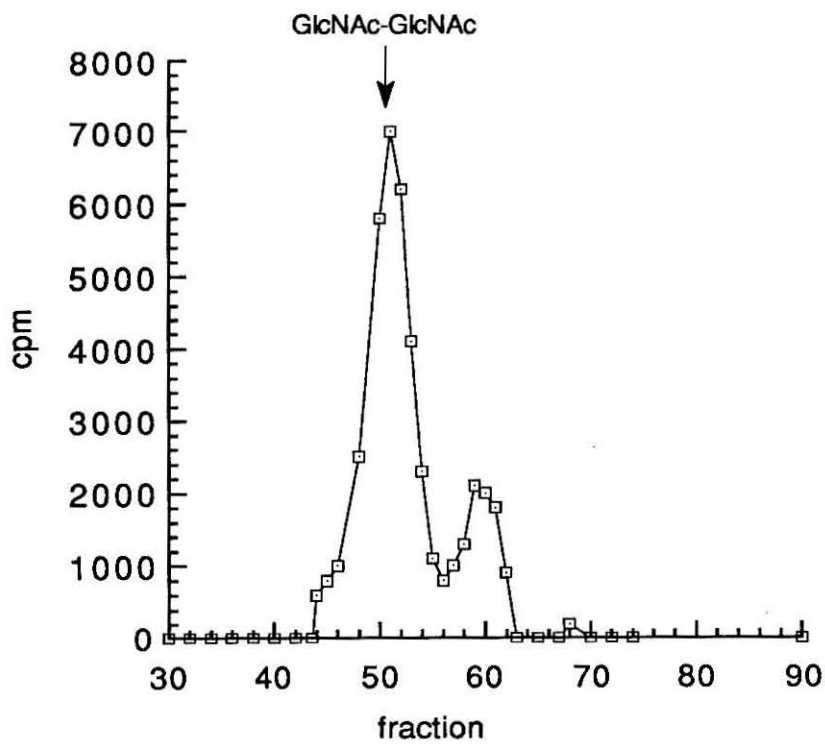


Figure 2-4 Determination of the oligosaccharide composition of hydrolyzed Dol-P-P-GlcNAc-[^{3}H]GlcNAc prepared as outlined above.

This synthetic method can easily produce as much as 20 μCi of Dol-P-P-GlcNAc-[^{3}H]GlcNAc, which represents an ample supply for extensive studies of the transferase.

This facile route to an exogenous supply of the truncated oligosaccharide donor eliminated some of the common problems that were associated with earlier attempts to study oligosaccharyl transferase activity. First, it allowed the concentration of the donor to be varied at constant enzyme concentration. This allowed the determination of the optimal oligosaccharide substrate concentration which maximized the transfer reaction. More importantly, it allowed the assay to be performed at this optimal substrate concentration, thus ensuring that transfer would be observed if it occurred. Additionally, the ability to manipulate the chemical structure of the oligosaccharide donor led to insights into the transferase's specificity towards this substrate (**10**).

This accessible supply of oligosaccharide donor also became the cornerstone of a convenient, inexpensive, and sensitive assay for oligosaccharyl transferase activity. This assay is described in detail later, but briefly it involves transfer of the radiolabeled truncated oligosaccharide donor to synthetic peptides using microsomes isolated from porcine liver as a source of the transferase.

Earlier assay attempts were limited by the amount of radiolabel which could be incorporated into the oligosaccharide donor and subsequently tranferred to the peptidyl acceptor. However, some significant attributes of the transferase were ascertained. Work with synthetic peptides revealed that viable peptidyl acceptors must include the marker sequence, -Asn-Xaa-Thr/Ser-, and that the

amino group of the asparagine and the carboxyl group of the hydroxy amino acid residue must be blocked, either by incorporation of another amino acid or by chemical derivatization (3). The deleterious effect of leaving the termini free might be caused by either an unfavorable charge interaction at the active site of the enzyme or by the disruption of the recognized conformation of the peptide by formation of a salt bridge.

Further it was demonstrated that the effect of the blocking groups is significant (11). The affinity of the peptidyl acceptor increases as the polypeptide backbone extends. Hydrophobic substituents also increase the efficiency of the peptide as a substrate. A role for hydrophobicity is consistent with the biology of the process of glycosylation. *In vivo*, nascent polypeptides attached to membrane bound ribosomes are glycosylated by the membrane-bound transferase. Glycosylation is believed to occur at or near the luminal face of the ER while the protein is passing through the membrane on its way into the lumen. Thus the more a peptide associates with the membrane fraction used in the assay, the better it will interact with the transferase. This is confirmed in assays performed on large polypeptidyl acceptors (41 amino acids). The concentration of nonionic detergent had to be increased suggesting that only peptides which freely associate with the enzyme can be glycosylated, which is consistent with the increase in acceptor properties observed as hydrophobicity was increased.

EXPERIMENTAL METHODS

2.1 Preparation of crude membrane fraction from porcine liver

Crude microsomal fractions were isolated from porcine liver as described by Zimmerman (**10**). This preparation was based on a procedure described by Behrens and Tabora (**12**).

Materials:

Homogenizing buffer
- 50 mM Tris-HCl, pH 7.2
- 0.25 mM sucrose
- 2.5 mM $MgCl_2$
- 3 mM dithiothreitol

Porcine liver

Glycerol

Procedure:

Fresh pig liver (100 g) is cut into small pieces and blended with 100 mL cold homogenizing buffer. The mixture is pulse-blended at 4°C on high speed until no large lumps are present. The mixture is then centrifuged at 11,000 rpm in a Beckman type 16 rotor for 30 minutes at 4°C in 250 mL centrifugation tubes. The supernatant is decanted into 25 mL tubes and centrifuged for another 30 minutes under the same conditions. The supernatant of this spin is centifuged at 25,000 rpm in a Beckman type 28 rotor for 1.5 hours at 4°C. The pellet of this spin, containing the microsomal fraction, is mixed with glycerol to form a 30% glycerol mixture and stored at -80°C. This procedure typically affords 2.5-5 mL of a microsomal preparation (~100 mg/mL protein). This material can be stored for several months without significant loss of transferase activity.

2.2.1 Enzymatic synthesis of Dol-P-P-GlcNAc-[^{3}H]GlcNAc

Dol-P-P-GlcNAc-[^{3}H]GlcNAc was enzymatically synthesized from synthetic Dol-P-P-GlcNAc as described by Imperiali and Zimmerman (**8**).

Materials:

Dol-P-P-GlcNAc (6 mg/mL in 3/2 chloroform/methanol)
UDP-[^{3}H]GlcNAc
Homogenizing buffer (refer to 2.1)
Nonidet P-40 (10% in Homogenizing buffer)
neat microsomes (prepared in 2.1)
4 mM $MgCl_2$
TUP (no salt) (12/192/186 chloroform/methanol/water)

Procedure:

Dol-P-P-GlcNAc (5 λ) was added to an Eppendorf tube along with 1.5 μCi UDP-[^{3}H]GlcNAc. This mixture was dried under a gentle stream of nitrogen. Nonidet P-40 (10% solution, 10 λ) and 85 λ Homogenizing buffer were added consecutively with vortexing after each addition. Neat microsomes (5 λ) were added to initiate the reaction. The Eppendorf was gently agitated on an orbital shaker for 30 minutes after which time the reaction was quenched with 1.0 mL 3/2 chloroform/methanol and 0.2 mL 4 mM $MgCl_2$. The Eppendorf was vortexed to thoroughly mix the reaction mixture with the quench solution. This mixture was centrifuged at 6,000 rpm for 2 minutes in a microcentrifuge to promote phase separation. The resulting mixture separated into an organic lower phase, an aqueous upper phase, and a proteinaceous interphase. The upper phase was removed. The proteinaceous interphase and the lower phase were

washed with 0.6 mL TUP (no salt). After a second centrifugation, the upper phase was combined with the previously isolated upper phase. The lower phase was removed from the proteinaceous pellet. The pellet was sonicated in 0.5 mL 3/2 chloroform/methanol for 5 minutes. The solution was decanted away from the pellet and combined with the previously isolated lower phase. The lower phase was blown down, resuspended in a known volume of 3/2 chloroform/methanol, and counted for radioactivity. A typical experiment consisting of 20 Eppendorfs yielded around 9,000,000 dpm of crude Dol-P-P-GlcNAc-[^{3}H]GlcNAc (~30% conversion).

The crude product was purified on silica gel G. Unreacted Dol-P-P-GlcNAc was eluted with 72/21/2 chloroform/methanol/water and the product, Dol-P-P-GlcNAc-[^{3}H]GlcNAc, with 60/25/4 chloroform/methanol/water.

2.2.2 Detergent effect (enzyme II)

The efficiency of various detergents used to solubize the Dol-P-P-GlcNAc with the membrane was examined. The microsomal fraction was dissolved in 85 λ Homogenizing buffer and 10 λ of either 10% Nonidet P-40 or 10% Triton X-100 was added. The mixture was gently agitated on an orbital shaker for 30 min. The reaction solution was quenched and worked up as described above. The percent conversion was determined by comparing the fraction of radioactivity present in the organic layer (Dol-P-P-GlcNAc-[^{3}H]GlcNAc) with that present in the aqueous layer (unreacted UDP-[^{3}H]GlcNAc).

Table 2-2 Efficiency of various detergents in solubilizing Dol-P-P-GlcNAc

Detergent	aqueous	organic	total	% conversion
Triton X-100	720000	79000	799000	9.9
Nonidet P-40	692000	135000	827000	16.3

The results are summarized in Table 2-2. Nonidet P-40 shows increased efficiency in solubilizing the lipid-linked oligosaccharide and therefore was used in all further experiments.

2.2.3 Mg(II) requirement (enzyme II)

Incubations with UDP-[^{3}H]GlcNAc were carried out according to standard assay conditions with varying concentrations of $MgCl_2$. The reaction shows an essential requirement for Mg^{2+} ions. Divalent cations have previously been shown to stimulate the activity of enzyme II in microsomes isolated from yeast (**13**).

Table 2-3 Effect of Mg(II) on the activity of enzyme II

conc. $MgCl_2$	aqueous	organic	total	% conversion
0 mM	756000	500	756500	0.07
2.5 mM	600000	139000	739000	18.8
10 mM	720000	79000	799000	9.9

2.3.1 Oligosaccharyl transferase assay- Isolation of glycopeptide

Materials:

Dol-PP-GlcNAc-[^{3}H]GlcNAc, 50,000-200,000 dpm

Assay buffer:

50 mM Tris-HCl, pH 7.5

10 mM $MnCl_2$

1.2% Triton X-100

Micromix:

200 λ crude microsomes

1400 λ assay buffer

peptide in DMSO

TUP with salt (12/192/186/2.69 chloroform/methanol/water/0.25 M $MgCl_2$)

Procedure:

Typically, 200,000 dpm Dol-PP-GlcNAc-[^{3}H]GlcNAc was aliquoted into 1.8 mL Eppendorf tubes and dried under a gentle stream of nitrogen. Peptide solution (10 λ) was added to each Eppendorf and vortexed to dissolve glycolipid. Assay buffer (40 λ) was added and again vortexed. At 0 minutes,150 λ Micromix was added to each Eppendorf and vortexed to ensure thorough mixing. The assay was shaken at room temperature on an orbital shaker during which four 40λ aliquots of the reaction mixture were removed and quenched into 7 mL scintillation vials containing 1.2 mL 3/2/1 chloroform/methanol/4 mM $MgCl_2$.

The quenched reaction aliquots were treated as follows to isolate the glycopeptide product away from the unreacted lipid-linked donor. The upper aqueous phases of the quenched solutions were removed into 7 mL scintillation vials. The lower phase (organic phase and proteinaceous pellet) was reextracted twice with 0.8 mL TUP w/salt. The upper phases were combined

and back extracted with 0.2 mL chloroform. The chloroform was pooled with the lower phase which was left overnight to dry.

The upper aqueous phase was dissolved into about 6 mL Ecolite (ICN) and the amount of tritium was quantitated as disintegrations per minute (dpm). The transfer of tritium into this aqueous phase represents the amount of newly formed glycopeptide.

The organic phase was redissolved in 0.2 mL Solvable (DuPont). After shaking for 2-3 hours, 6 mL Formula 989 (DuPont) was added and the amount of tritium was quantitated. The tritium remaining in this phase represents the amount of unreacted glycolipid. The data was normalized for each Eppendorf.

The observed rates were shown to be due to enzyme-catalyzed glycosylation, and not due to hydrolysis of the lipid-linked substrate to afford a water soluble disaccharide, by the following experiments. First, reactions in the presence of all standard assay constituents initiated with thermally denatured microsomes showed no discernable substrate hydrolysis. Secondly, water soluble radioactive material, obtained from the assay, was prepared on a large scale and subjected to acid catalyzed hydrolysis followed by gel filtration analysis on a Biogel P-2 column (eluent 50 mM acetic acid). The major radioactive material eluted corresponded to a compound with higher molecular weight than di-N-acetylchitobiose which is consistent with formation of glycoprotein.

2.3.2 Glycerol dependence (oligosaccharyl transferase)

Bz-NLT-NMe (0.5 mM) was incubated with Dol-P-P-GlcNAc-[^{3}H]GlcNAc according to standard assay conditions except the amount of assay buffer contained in the micromix was varied. As more assay buffer was added the effective concentration of both the glycerol and the enzyme were decreased.

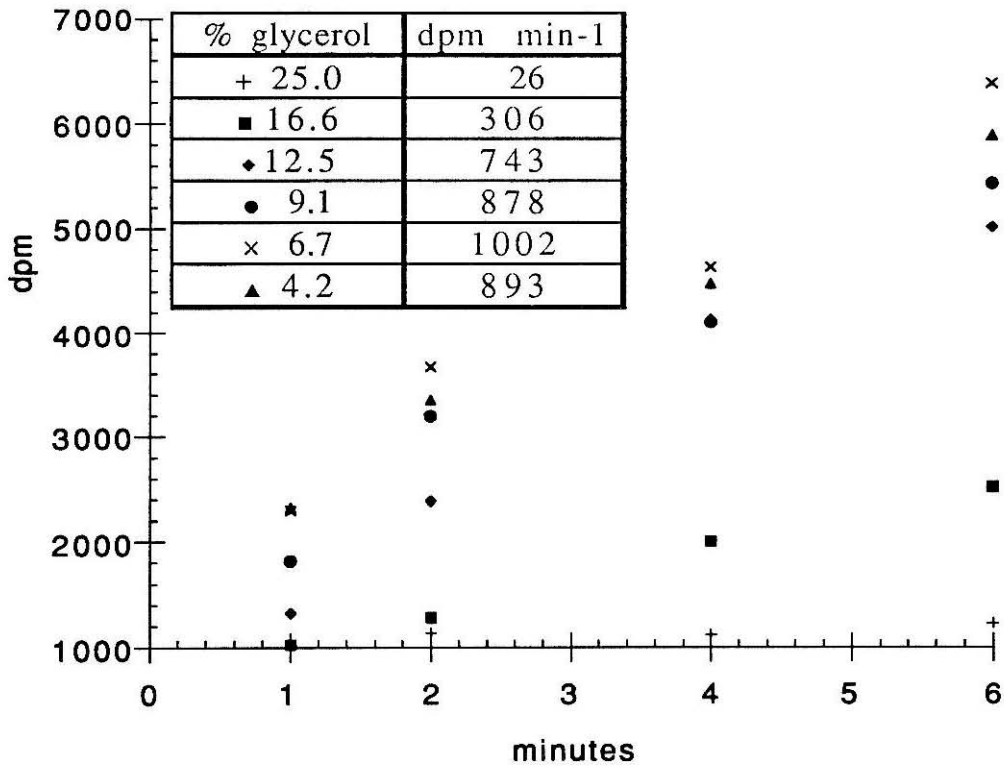
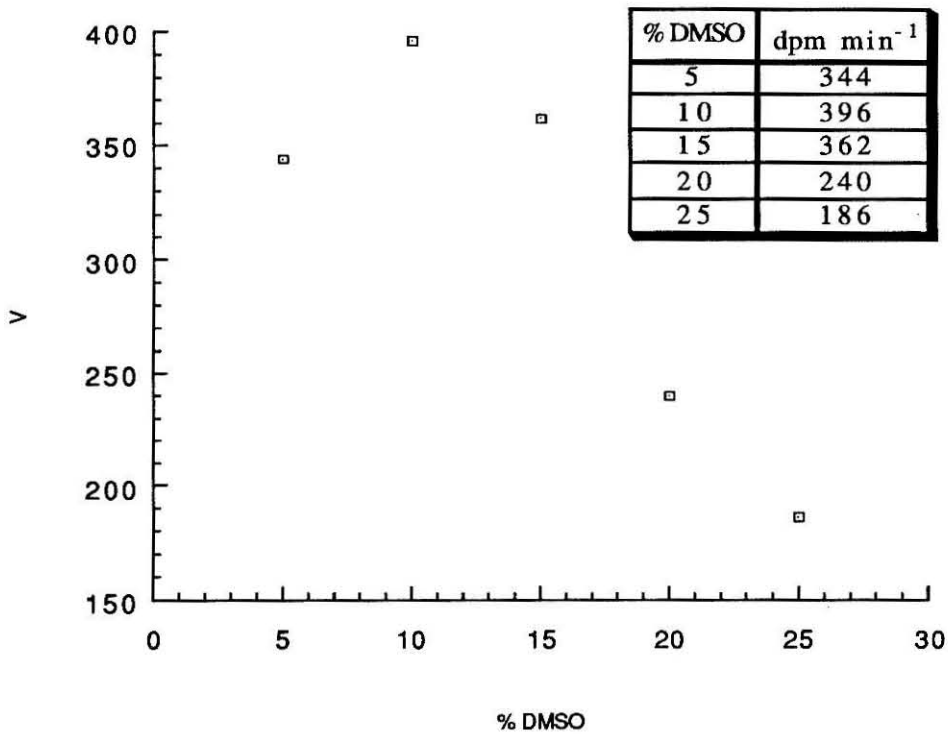


Figure 2-5 Glycerol dependence of oligosaccharyl transferase.

Figure 2-5 demonstrates that as the concentration of glycerol in the assay was decreased the transfer velocity increased. The velocity differences observed are even more striking considering that the effective enzyme concentration also decreased. Glycerol is necessary to maximize the time which the liver can be stored at -85°C (**10**).

2.3.3 DMSO dependence (oligosaccharyl transferase)

Bz-NLT-NMe (0.5 mM) was incubated with Dol-P-P-GlcNAc-[^{3}H]GlcNAc according to standard assay conditions except the percentage of DMSO was varied.



% DMSO	dpm min^{-1}
5	344
10	396
15	362
20	240
25	186

Figure 2-6 DMSO dependence of oligosaccharyl transferase.

Figure 2-6 illustrates that increasing the percentage of DMSO appears to subtly increase the transfer velocity up to 15% after which the velocity drops off. The optimum percentage of DMSO in the transferase assay is around 10%. This finding contradicts an earlier report (**14**) which found that the addition of DMSO had no observable effect on the transferase properties of tripeptides.

The addition of DMSO to the enzyme assay probably has two effects. DMSO has previously been shown by circular dichroism experiments to increase the amount of secondary structure in peptides containing at least five residues (**14**). Thus, DMSO may increase the population of the conformation which is recognized by the enzyme. The second effect might be to solubilize the membrane-bound enzyme and the oligosaccharide donor, or alternatively to mimic the membrane interface where glycosylation occurs.

2.3.3 Effect of Blocking Groups

Table 2-4 summarizes the substrate behavior for five peptides which are composed of the marker sequence blocked with various groups. The glycosyl acceptor properties were determined with Dol-P-P-GlcNAc-[^{3}H]GlcNAc using standard assay conditions. The data were obtained with oligosaccharide donor of identical specific radioactivity and are thus directly comparable. K_m and V_{max} values were determined for each of the peptides as well as a relative acceptor activity value (the plots used to determine these values are included in Appendices 2-1 to 2-5).

Table 2-4 Relative acceptor activities for the glycosylation of blocked tripeptides.

peptide	K_m (mM)	V_{max} (dpm min^{-1})	V_{max}/K_m
Boc-NLT-OMe	1.13	186	164.6
Ac-NLT-OMe	2.68	500	186.6
Ac-NLT-NHMe	0.56	431	769.6
Bz-NLT-OMe	0.54	1550	2870.4
Bz-NLT-NHMe	0.47	4208	8953.2

Bz-NLT-NHMe was the best glycosyl acceptor in this study, as determined by its highest V_{max}/K_m value. Substitution of the N-terminal benzoyl group with an acetyl group decreased the relative rate 10-15 fold. Substitution of the C-terminal methyl ester with a methyl amide increased the relative rate 3-4 fold.

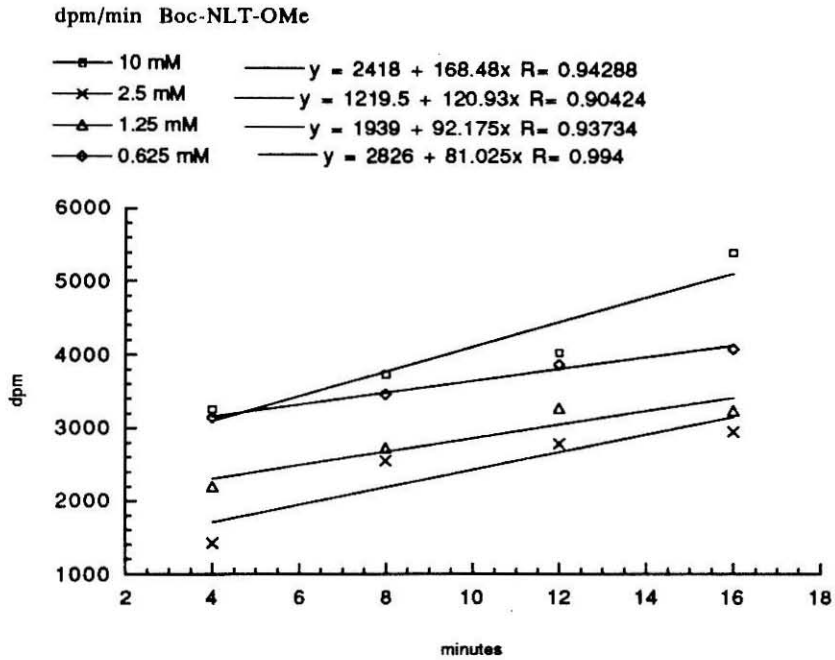
The effect of the various blocking groups on the rate of glycosyl transfer is best understood in terms of the topology of the process. The polypeptide acceptor which interacts with the transferase is still being translated through the ER membrane. Thus, the enzyme active site is likely equipped to recognize a marker sequence which is attached via an amide bond to the remaining portion of the polypeptide. An amide at the C-terminus more accurately mimics the normal substrate than a methyl ester. Furthermore, if this amide proton is involved in a hydrogen bonding interaction with the enzyme, the loss of this proton would lower the binding efficiency of the peptide. Also the lipid linked oligosaccharide donor and the transferase are both attached to the hydrophobic ER membrane. As the hydrophobicity of the terminal blocking groups increases, the ability of the peptide to interact with the membrane increases and glycosylation is promoted.

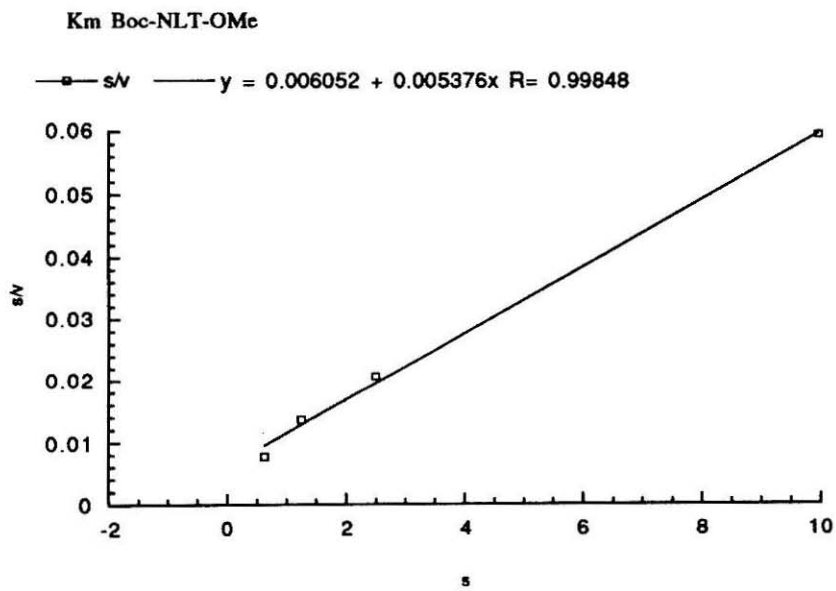
CONCLUSIONS

Small peptides containing the marker sequence, -Asn-Xaa-Thr/Ser-, and blocked at the carboxyl and amino termini can be glycosylated *in vitro* by a microsomal fraction isolated from porcine liver. The glycosyl acceptor abilities of these peptides can

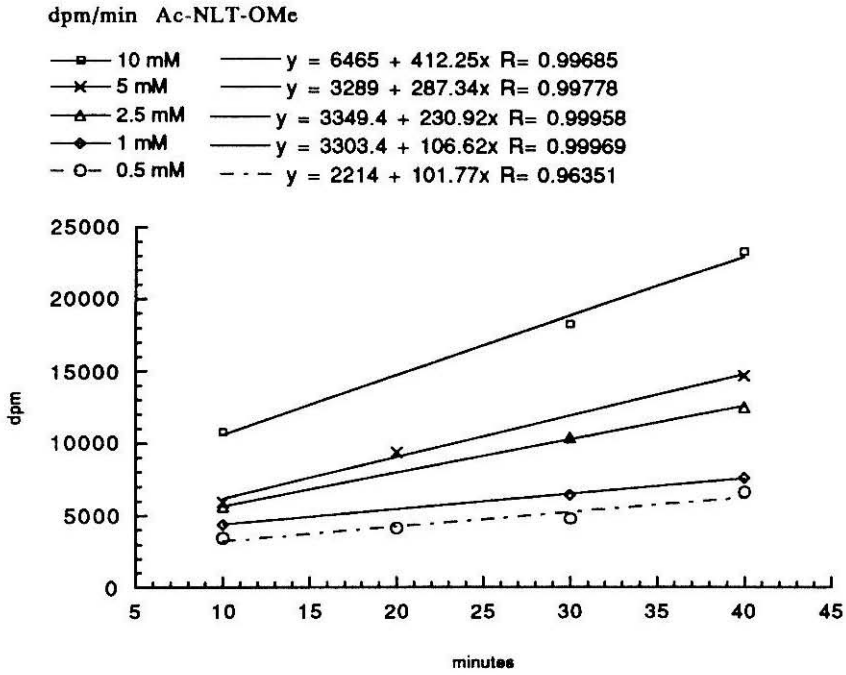
be reliably determined by monitoring the transfer of radiolabelled oliosaccharide from semi-synthetically prepared Dol-P-P-GlcNAc-[^{3}H]GlcNAc to the peptide.

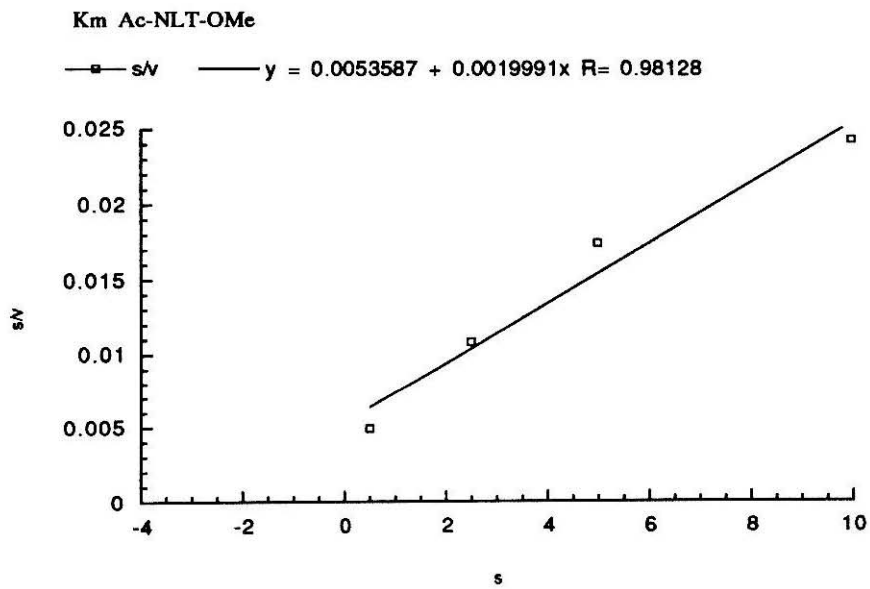
Appendix 2-1 Enzyme kinetic data for Boc-NLT-OMe.



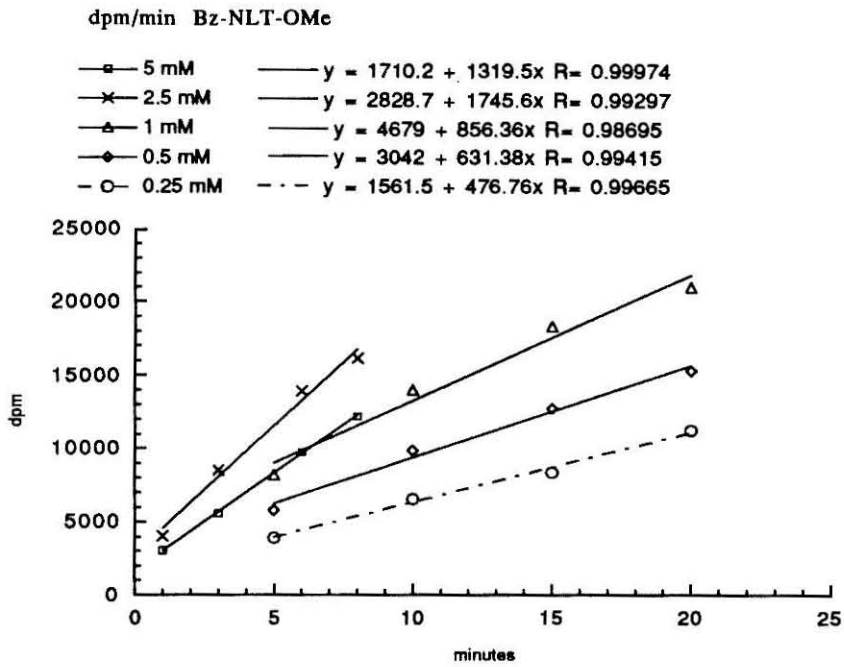


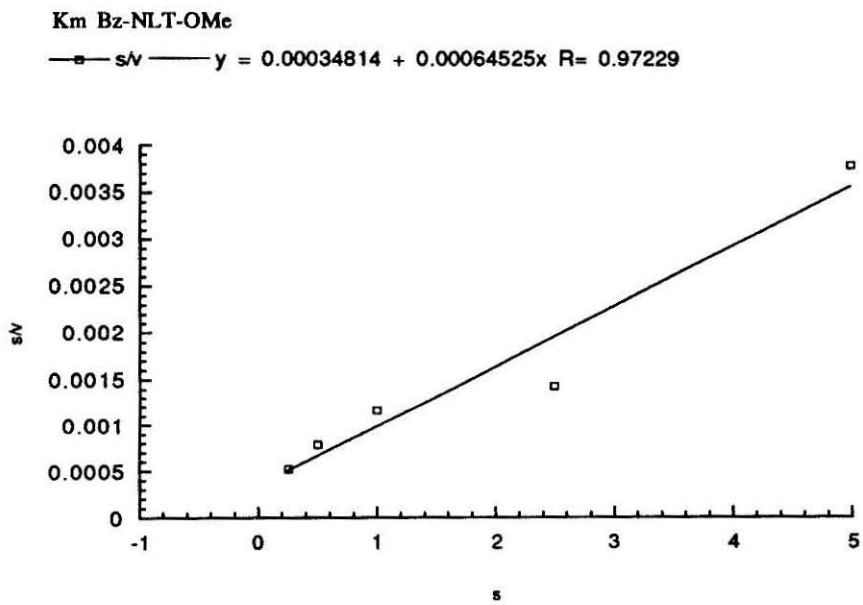
Appendix 2-2 Enzyme kinetic data for Ac-NLT-OMe



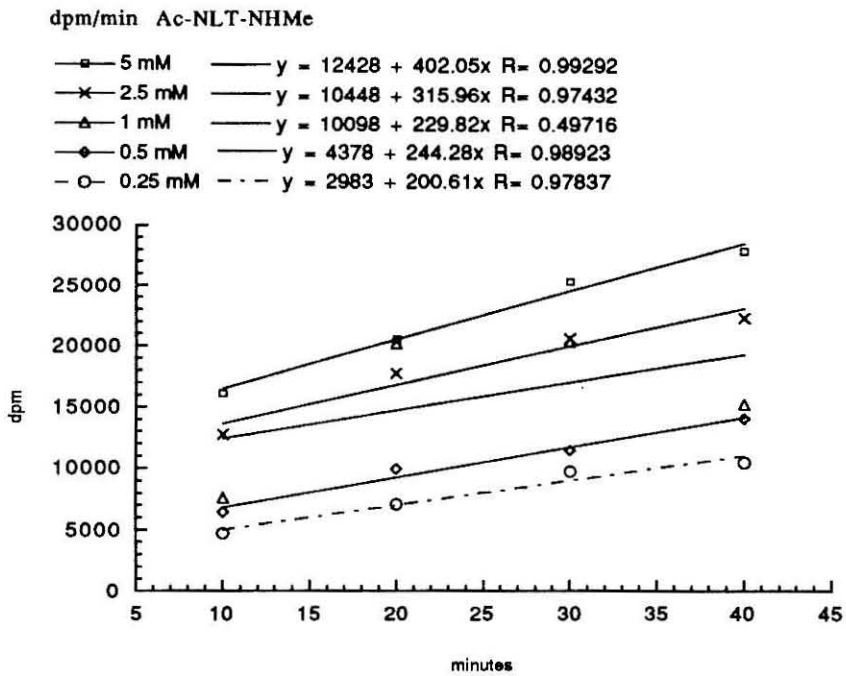


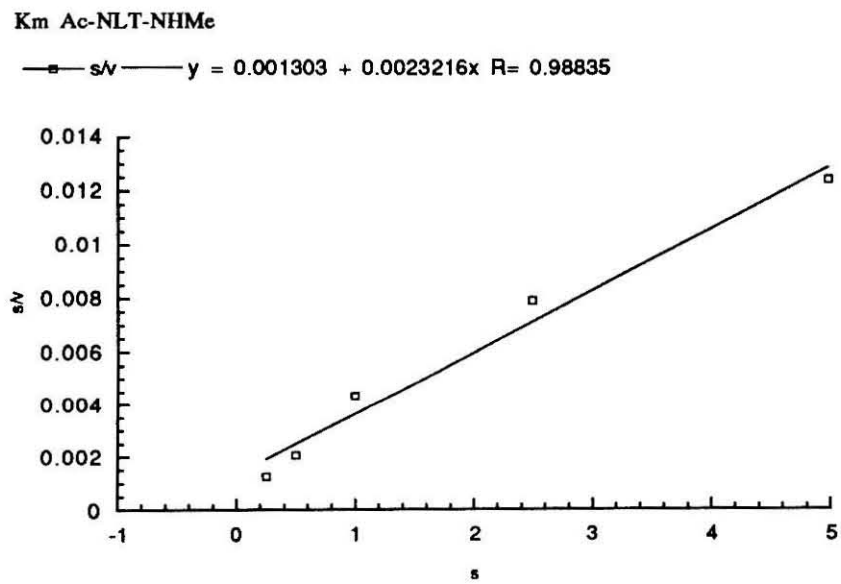
Appendix 2-3 Enzyme kinetic data for Bz-NLT-OMe



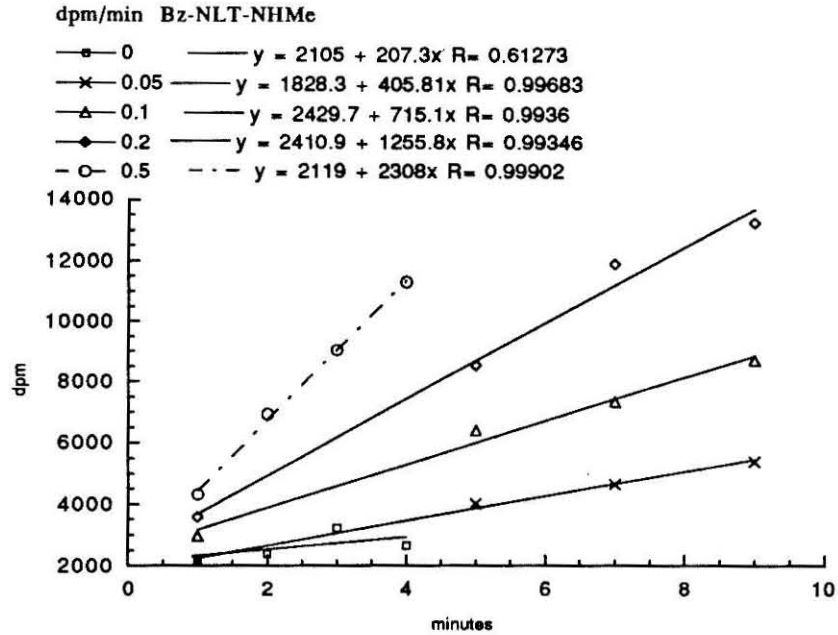


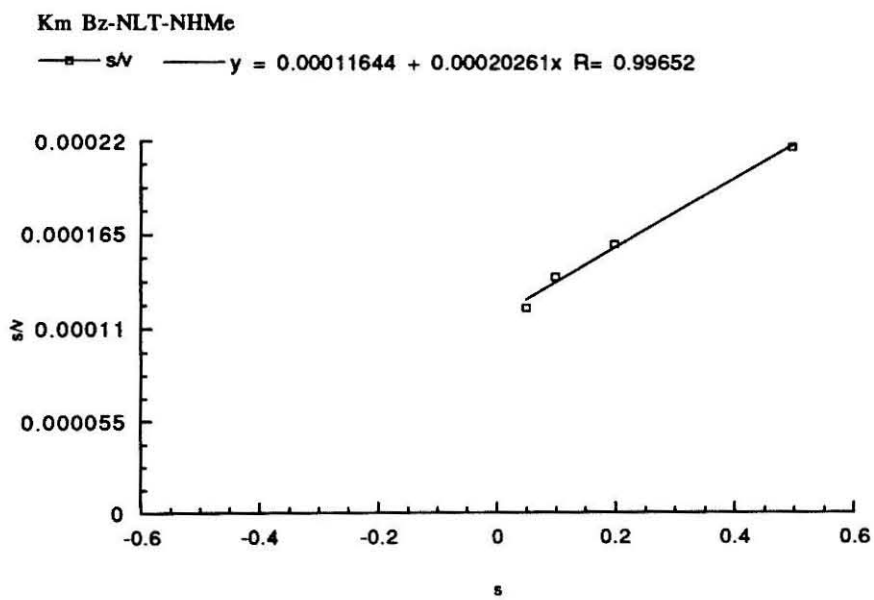
Appendix 2-4 Enzyme kinetic data for Ac-NLT-NHMe





Appendix 2-5 Enzyme kinetic data for Bz-NLT-NHMe





1. a) Heifetz, A.; Elbein, A.D. "Solubilization and Properties of Mannose and N-Acetylglucosamine Transferases Involved in Formation of Polyprenyl-Sugar Intermediates," *J. Biol. Chem.* **1977**, *252*, 3057-3063. b) Keller, R.K.; Boon, D.Y.; Crum, F.C. "N-Acetylglucosamine-1-Phosphate Transferase from Hen Oviduct: Solubilization, Characterization, and Inhibition by Tunicamycin," *Biochemistry* **1979**, *18*, 3946-3952.

2. Pless, D.D.; Lennarz, W.J. "Enzymatic Conversion of Proteins to Glycoproteins," *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 134-138.

3. Hart, G.W.; Brew, K.; Grant, G.A.; Bradshaw, R.A.; Lennarz, W.J. "Primary Structural Requirements for the Enzymic Formation of the N-Glycosidic Bond in Glycoproteins. Studies with Natural and Synthetic Peptides," *J. Biol. Chem.* **1979**, *254*, 9747-9753.

4. Ronin, C.; Bouchilloux, S., Granier, C.; Van Rietschoten, J. "Enzymic N-Glycosylation of Synthetic Asn-X-Thr Containing Peptides," *FEBS Lett.* **1978**, *96*, 179-182.

5. Sharma, C.B.; Lehle, L.; Tanner, W. "N-Glycosylation of Yeast Proteins," *Eur. J. Biochem.* **1981**, *116*, 101-108.

6. Bause, E.; Lehle, L. "Enzymic N-Glycosylation and O-Glycosylation of Synthetic Peptide Acceptors by Dolichol-Linked Sugar Derivatives in Yeast," *Eur. J. Biochem.* **1979**, *101*, 531-540.

7. Reviewed by a) Parodi, A.J.; Leloir, L.F. "The Role of Lipid Intermediates in the Glycosylation of Proteins in the Eukaryotic Cell," *Biochim. Biophys. Acta* **1979**, *559*, 1-37. b) Sharon, N.; Lis, H. In *The Proteins*, 3rd ed. Neurath, H.; Hill, R.L., Ed.; Academic Press: London, 1975; Vol. 5, pp. 1-144.

8. Imperiali, B.; Zimmerman, J.W. "Synthesis of Dolichylpyrophosphate-Linked Oligosaccharides," *Tetrahedron Lett.* **1990**, *31*, 6485-6488.

9. Imperiali, B.; Zimmerman, J.W. "Synthesis of Dolichols via Asymmetric Hydrogenation of Plant Polyprenols," *Tetrahedron Lett.* **1988**, *29*, 5343-5344.

10. Zimmerman, J.W. Ph.D. Thesis, Carnegie Mellon University, June 1990.

11. Welply, J.K.; Shenbagamurthi, P.; Lennarz, W.J.; Naider, F. "Substrate Recognition by Oligosaccharyltransferase," *J. Biol. Chem.* **1983**, *258*, 11856-11863.

12. Behrens, N.H.; Tabora, E. "Dolichol Intermediates in the Glycosylation of Proteins," *Methods Enzymol.* **1977**, *50*, 402-435.

13. Sharma, C.B.; Lehle, L.; Tanner, W. "Solubilization and Characterization of the Initial Enzymes of the Dolichol Pathway from Yeast," *Eur. J. Biochem.* **1982**, *126*, 319-325.

14. Ronin, C.; Aubert, J.-P. "Effect of Dimethylsulfoxide on Two Synthetic Asn-X-Thr Containing Substrates of the Oligosaccharyltransferase," *Biochem. Biophys. Res. Commun.* **1982**, *105*, 909-915.

Chapter 3

Molecular Modelling and Amide Exchange Experiments on Hexapeptides

On the basis of structural analysis of a number of known N-glycosylated proteins, Marshall suggested that the marker sequence, -Asn-Xaa-Thr/Ser-, is a necessary requirement for glycosylation (**1**). However, since not all such sequences in proteins become glycosylated this marker sequence is evidently not the only prerequisite. It is well established that glycosylation is blocked when the central residue of the marker sequence, Xaa, is proline (**2**), but this rule accounts for only a minor portion of all known non-glycosylated marker sequences.

Several researchers have attempted to unravel the other requirements necessary for glycosylation. Gavel and von Heijne have done a thorough statistical analysis of known N-glycosylated proteins (**3**). They looked for correlations between the glycosylation tendency of marker sequences and the position of these sequences within the protein. They found that the frequency of non-glycosylated sites was higher towards the C-terminus of most proteins and also when the site contained proline as either the central residue or the residue flanking the hydroxy amino acid. They postulated that local conformation within and surrounding the marker sequence influences glycosylation. Beyond this, they concluded that there were no local patterns, other than the marker sequence, within the primary amino acid code which correlated with glycosylation.

Their results further substantiated Van Rietschoten's finding that Ac-Asn-Ala-Thr-NH_2 was the smallest peptide to be glycosylated *in vitro* (**4**). The fact that this tripeptide is recognized by the enzyme indicates that all the necessary structural features required for binding are present within the molecule. Thus although

both the overall folding of the protein as well as the local conformation surrounding the marker sequence may affect glycosylation, a tripeptide sequence with blocked termini is sufficient.

The complete tripeptide marker sequence was confirmed to be required for glycosylation by Hart *et al.* (**5**). They showed that a dipeptide containing only the first two residues of the marker sequence could not be glycosylated *in vitro* (Table 3-1). Ac-Asn-Gly-Ser-NHMe was a good acceptor, but Ac-Asn-Gly-NHMe is inactive, clearly indicating that the tripeptidyl marker sequence is the minimal encoded structure necessary for recognition by the enzyme.

Table 3-1 Relative acceptor activities for the glycosylation of peptides by hen oviduct microsomes. Activity is expressed relative to GNLSL.

Peptide	Relative acceptor activity/nmol of peptide
GNLSL	1
Ac-NGS-NHMe	0.7
Ac-NGS-OMe	0.2
Ac-NG-NHMe	<0.01

The significance of the third residue, the hydroxy amino acid, in the marker sequnce has been extensively questioned. Since its presence within the marker sequence is a strict requirement, Marshall postulated that it must be somehow involved in inducing a very specific conformation within the peptide which can be recognized by the enzyme. Marshall proposed that a hydrogen bond

formed between the hydroxyl group of the hydroxy amino acid and the carboxyamido group of the asparagine might form a unique structure within the peptide which the enzyme might then recognize (Figure 3-1) (**1**).

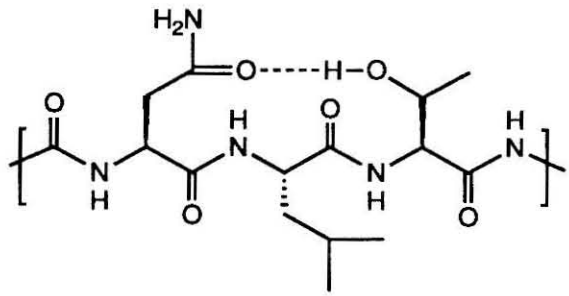


Figure 3-1 Hydrogen bonding array proposed by Marshall.

The role of the hydroxy amino acid in the marker sequence was investigated (**6**) by using a series of hexapeptides derived from Tyr-Asn-Gly-Xaa-Ser-Val by substituting threonine, serine, cysteine, valine and O-methylthreonine for Xaa (Table 3-2).

Table 3-2 Glycosyl-acceptor activity of hexapeptides containing modifications of the hydroxyl amino acid within the marker sequence. Relative transferase activity (V_{max}/K_m) was based on the activity of the serine-containing peptide.

Peptide	K_m (mM)	V_{max} (cpm/10 min)	Relative transferase activity
YNGTSV	0.16	2080	40
YNGSSV	1.6	550	1
YNGCSV	-	-	0.4
YNGVSV	-	-	0
YNGT(OMe)SV	-	-	0

These results show that serine, threonine, and cysteine-containing peptides were glycosylated, although at very different rates, whereas valine and O-methylthreonine containing peptides were not substrates. Replacement of threonine by serine resulted in a four fold decrease in V_{max} and about a ten fold increase in K_m for oligosaccharyl transferase. Replacement of serine by cysteine again decreased acceptor activity two to three fold. These results indicate an absolute requirement for a hydrogen bond donor function in the side chain of the hydroxyl amino acid of the marker sequence and furthermore, point to a considerable influence of this amino acid in binding as well as in turnover. In order to include the observed role of the hydroxyl group as a hydrogen bond donor and at the same time rationalize the substrate behavior exhibited by this series, Bause and Legler proposed an alternate model which contained a hydrogen bond interaction between the side chain amide of asparagine (the proton donor) and the oxygen of the hydroxyl group of the hydroxyl amino acid (the proton acceptor) (Figure 3-2) (**6**). The rationale behind their mechanism will be examined in more detail in chapter 6.

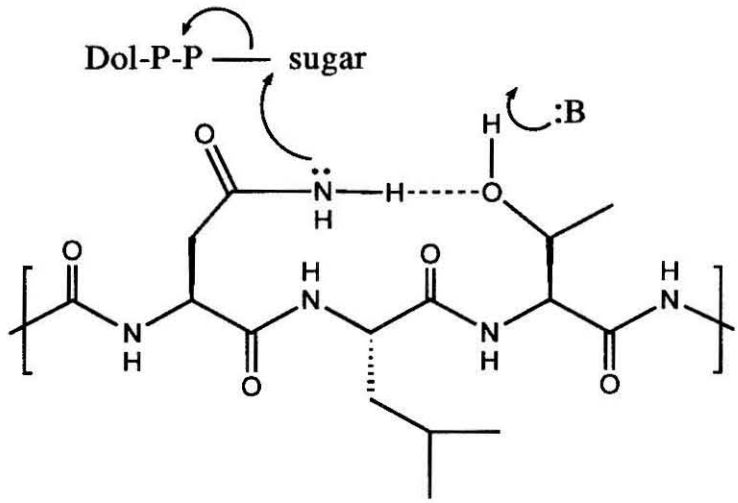


Figure 3-2 Bause and Legler's proposed hydrogen bond array.

The conformations of several tripeptide substrates have previously been examined using energy calculations (7). These studies suggested that despite their diminished size, these peptides exhibit a high probability for secondary structure formation. However these studies were limited by a lack of experimental support.

Our initial approach to exploring the possibility that a specific conformation within the peptide was being recognized by the enzyme included both computational modelling and 1D NMR experiments.

EXPERIMENTAL METHODS

3.1 Computational Modelling

The computational study described below was done in collaboration with C. L. Brooks at Carnegie Mellon University. The modelling program CHARMM (8) was used to perform molecular dynamic studies on a series of hexapeptides in an attempt to

correlate peptide conformation to substrate behavior in glycosylation. The copy of the CHARMM program which was used was under development and had been supplied by Prof. M. Karplus. Computations were performed on a VAX computer and structural data was analyzed with the aid of an Evans and Sutherland Graphics system.

Three hexapeptides were examined: Ac-PNGTAV-NH_2, Ac-ANGTAV-NH_2, and Ac-YNPTSV-NH_2. These peptides were chosen because the substrate behavior of their unblocked counterparts had previously been determined (Table 3-3) (**9**). Of these peptides, PNGTAV and ANGTAV exhibit substrate behavior . The slightly better substrate behavior exhibited by PNGTAV has been attributed to an increased population of the recognized conformer produced by the presence of the constrained imino acid.

Table 3-3 Glycosyl acceptor properties of hexapeptides containing proline within and surrounding the marker sequence.

Peptide	V_{max}/K_m	K_i (mM)
(H_2N^+)PNGTAV(COO^-)	0.51	-
(H_3N^+)ANGTAV(COO^-)	2.6	-
(H_3N^+)YNPTSV(COO^-)	-	>5

CHARMM was used to construct the hexapeptide structures, to energy minimize them using first and second derivative techniques, to perform dynamic simulations, and to analyze the structural and dynamic properties determine by these calculations. The study was aimed at finding a representative sampling of minimum energy structures for each hexapeptide. The compounds were studied as blocked hexapeptides to prevent electrostatic interactions from

dominating the energy studies. In order to obtain the most accurate sampling, the Monte Carlo approach known as simulated annealing (**10**) was used during the dynamics simulation. This approach efficiently finds the global energy minima along the potential energy surface. It differs from conventional iterative minimization because it allows molecules to escape local minima as the temperature is decreased.

3.1.1 Generation and Minimization of Structures

Each structure was treated as an isolated molecule in a vacuum. The structures were generated using CHARMM's protein topology and parameter files. Each structure was initially minimized with an NOE constraint between the side chain amide nitrogen of the asparagine and the hydroxyl oxygen of the threonine. The constraint was used to keep these atoms within hydrogen bonding distance (3.0 Å). This constraint was used since it was proposed that this was the important interaction necessary for binding (**10**). The structures were minimized using adopted-basis set Newton-Raphson minimization (**8**). The NOE constraint was gradually reduced to zero as the energy was minimized. The final minimized structures were then used in molecular dynamic simulations.

An example of a CHARMM Input File used to generate the initial minimized structure for Ac-PNGTAV-NH_2 is included in the Appendix 3-1.

3.1.2 Dynamics simulations

A simulated annealing strategy was adapted into the CHARMM program and employed during the dynamics simulation. First the minimized structure was heated to 598 K over 10 picoseconds, the temperature was maintained at 598 K for 5 picoseconds, the temperature was then lowered back to 298 K over 10 picoseconds, and finally 10 picoseconds of dynamics were performed (Figure 3-3). During the heating and cooling processes, the structure was

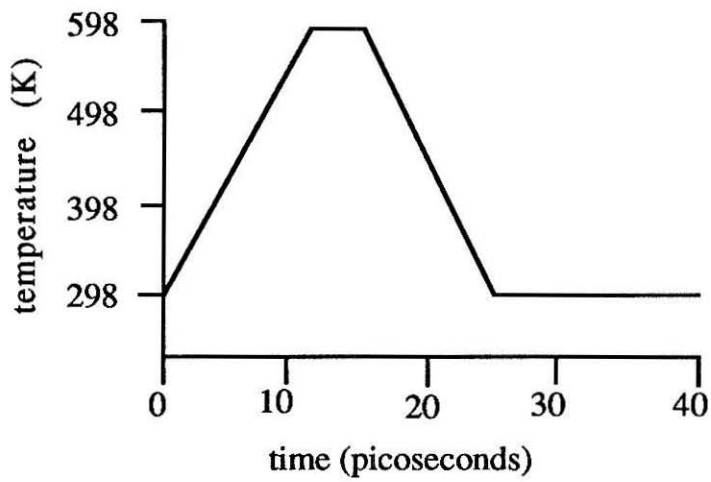


Figure 3-3 Timeline of simulated annealing experiment

continually minimized. During the final 10 picoseconds of the simulation a dynamic trajectory was generated and at 10^{-2} picosecond intervals, snapshots of this trajectories were stored for later analysis.

The dynamics runs were carried out at a dielectric constant of ten. This dielectric value had been previously shown (**11**) to be optimal for observing folding of a peptide during a 20 picosecond simulation. Dielectric values above ten produced structures which did not fold during the simulation, while values below ten produced structures with abnormally high amounts of hydrogen bonding. An example of a CHARMM input file used to generate the dynamic trajectory for Ac-PNGTAV-NH_2 is included in Appendix 3-2.

Six different dihedral angles were followed over the course of each dynamic simulation. These are illustrated in Figure 3-4.

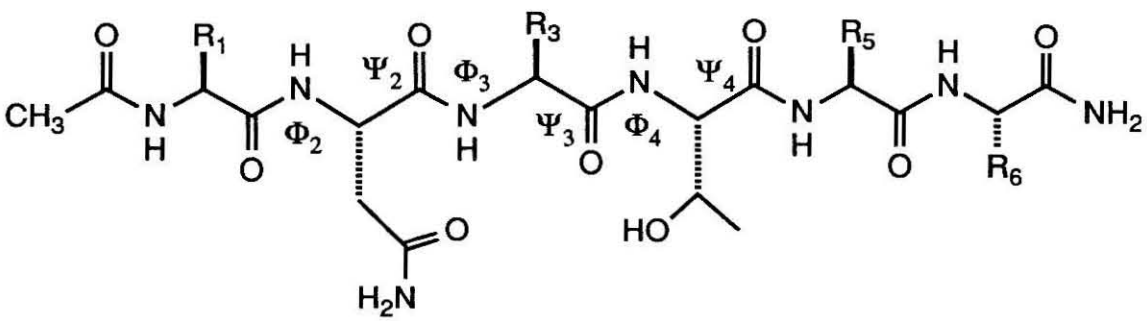


Figure 3-4 Dihedral angles measured during simulated dynamics.

The probability of each dihedral angle, -180° to 180°, was determined and plotted using Ramachandran plots (Figure 3-5).

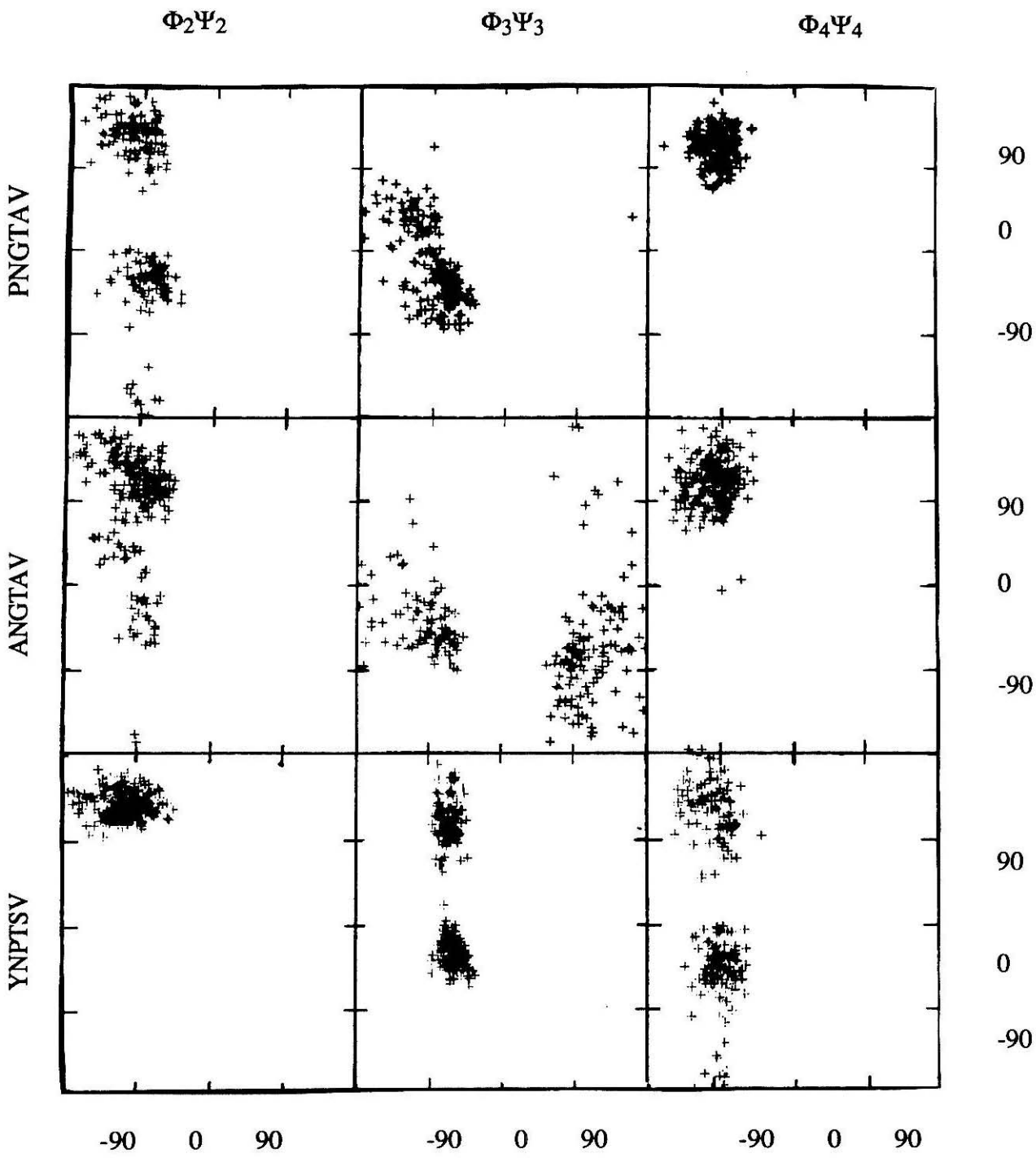


Figure 3-5 Ramachandran plots of the lowest energy conformers of Ac-PNGTAV-NH_2, Ac-ANGTAV-NH_2, and Ac-YNPTSV-NH_2. Only the dihedrals defining the marker sequence are shown.

Four different hydrogen bond distances (Figure 3-6) were also followed during the simulation.

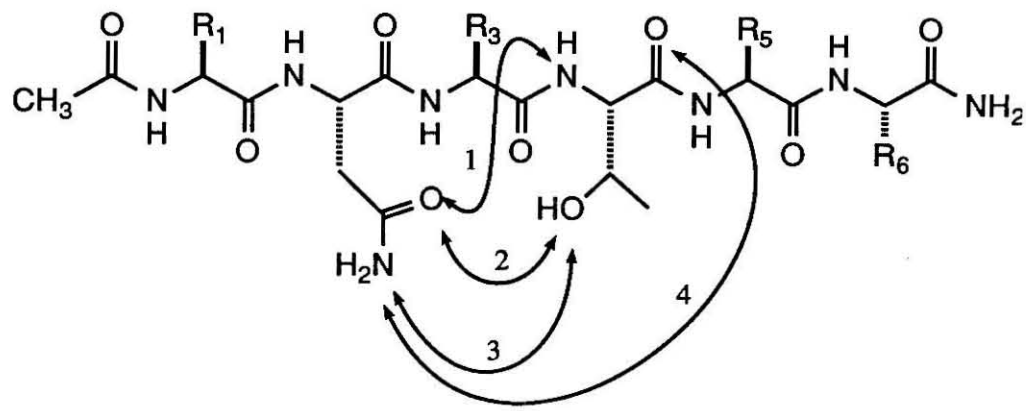


Figure 3-6 Bond distances measured during simulated dynamics.

RESULTS AND DISCUSSION

Ramachandran plots of the Φ,Ψ dihedral angles obtained from the dynamic trajectory are shown in Figure 3-5. The three dihedrals shown define the conformation of the marker sequence within each of three peptides examined. The only correlations between structure and substrate behavior which are evident are that substrates must be able to access the following average dihedral angles: Φ_2, =-90, Ψ_2 = -45; Φ_3 = -80, Ψ_3= -45; and Φ_4 = -90, Ψ_4 = 120. When these values are used to regenerate a structure, a polypeptide with an extended backbone is produced. In comparison, the backbone of the nonsubstrate, Ac-YNPTSV-NH_2, cannot fully extend and is instead kinked due to the restriction around the Φ,Ψ dihedral angles of proline. This data suggests that the enzyme requires a rather extended backbone within the peptide acceptor for binding. Thus, specific backbone dihedral information cannot be derived from these

studies; however, based on bond distance information a dominant conformation does appear to exist.

Significantly, structural data obtained from bond distance analysis correlates with substrate behavior. There seems to be a high probability in the substrate peptides for formation of a hydrogen bond between the side chain carbonyl oxygen of asparagine and the backbone amide of threonine (refer to Figure 3-5, bond 1). This bond distance is routinely 2-4 Å. A second hydrogen bonding interaction between the same carbonyl oxygen and the hydroxyl group occurs to a lesser extent (Figure 3-5, bond 2). In Figure 3-7 the distances of each putative hydrogen bond followed during the simulation of Ac-PNGTAV-NH_2 is plotted as a function of the percentage of time it was observed.

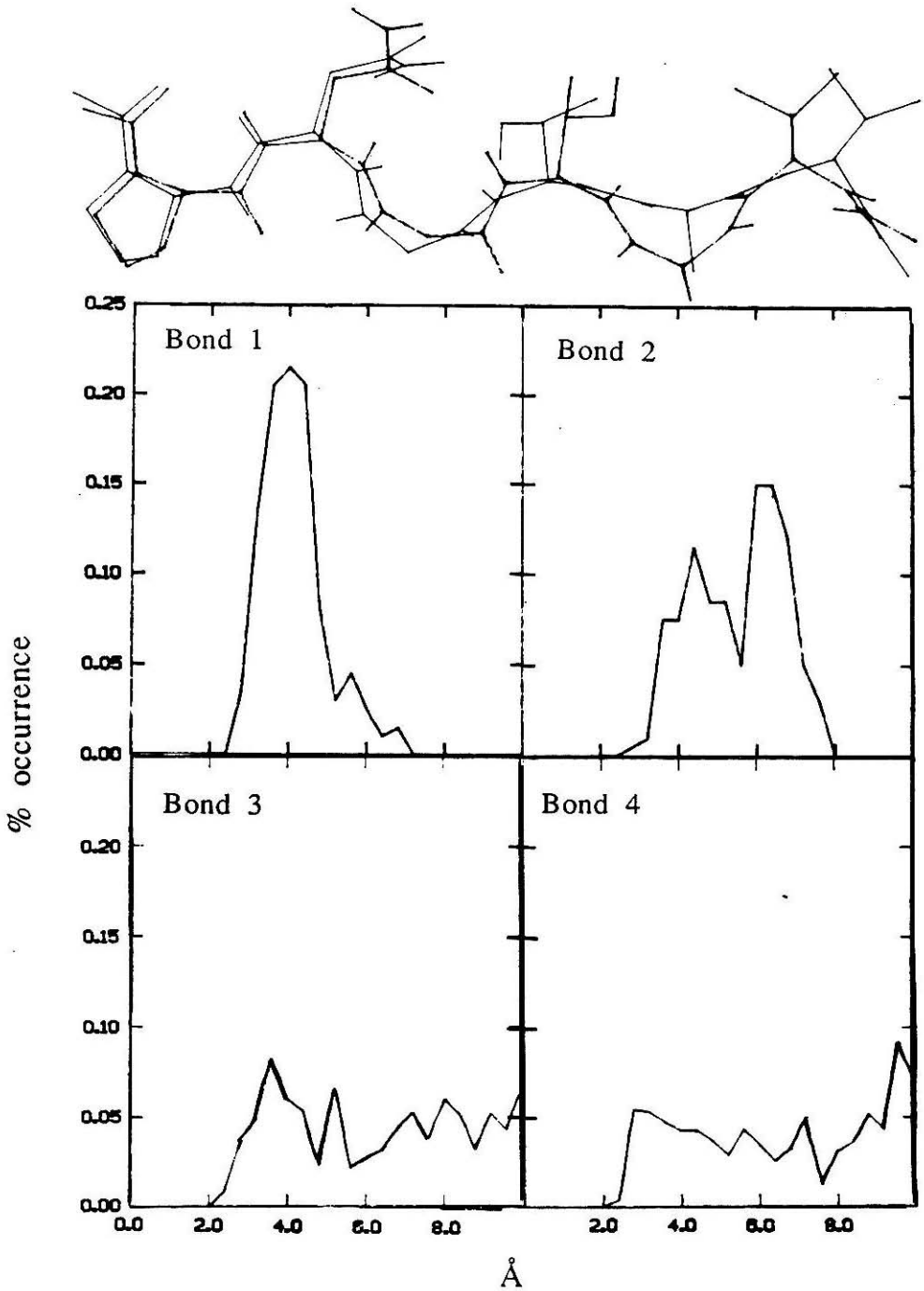


Figure 3-7 Percent distribution of bond distances within the lowest energy conformers of Ac-PNGTAV-NH_2.

The structures superimposed above these plots are representative of the conformations observed during the simulation. In Figure 3-8 bonds 1 and 2 are compared for all three peptides. The average distance for both of these bonds decreases dramatically down the series and correlates well with the observed substrate behavior for these peptides. Thus, hydrogen bond formation between the side chain carbonyl oxygen of asparagine and the backbone amide of threonine appears to be important for substrate activity in glycosylation.

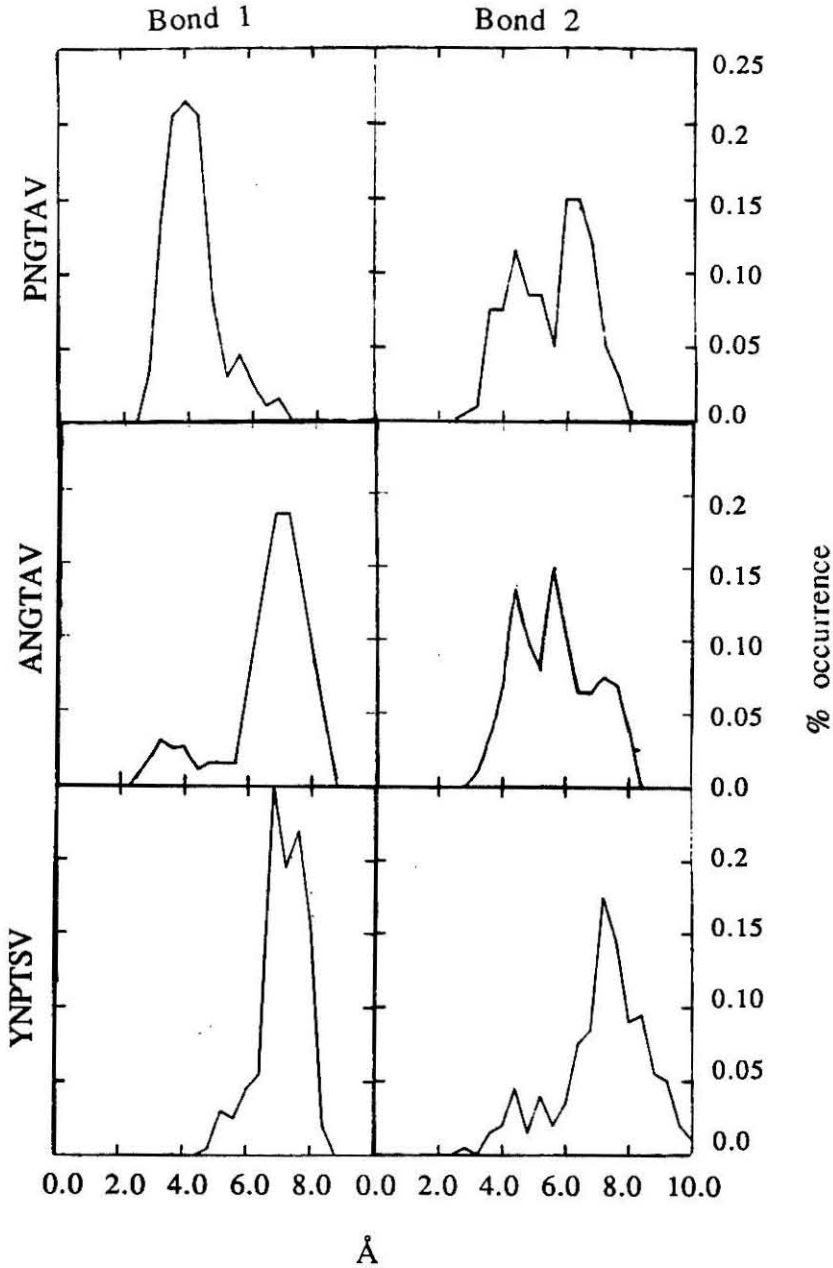


Figure 3-8 Distribution of bond distances observed for bonds 1 and 2 within the lowest energy conformers of Ac-PNGTAV-NH_2 and Ac-ANGTAV-NH_2. Bond distances lengthen as substrate behavior decreases.

EXPERIMENTAL METHODS

3.2 Analysis of Peptide Conformation by ^{1}H NMR

It is generally recognized that the rates of exchange of deuterium for hydrogen in proteins reflects the structure and conformation of these molecules (**12**). However, unraveling the contributions of structure to hydrogen exchange behavior requires knowledge of the micro environment surrounding each proton in the unstructured polypeptide. A variety of factors influence the primary hydrogen exchange rate of a proton within a polypeptide: hydrophobic groups, charge interactions, general acid base catalysis by side chains, and inductive effects which can alter local water chemistry. It has been shown that in polypeptides, the exchange rate of a given backbone amide proton is predominantly influenced by nearest-neighbor substituents and that these influences can be accurately represented by summing the individual effects of each adjacent amino acid. Furthermore, the influence of each of the 20 naturally occurring amino acids has been determined in model studies (**13**).

Using this information, it was possible to predict the relative exchange rate for each of the backbone amides within the two peptidyl substrates, Ac-PNGTAV-NH_2 and Ac-ANGTAV-NH_2. If no secondary structure is present, the relative exchange rates of all of the amide protons are similiar (*i.e.*, within the same order of magnitude). Therefore, any amide proton which had a relative exchange rate which was slower by one order of magnitude or more from the other amide protons was assumed to be solvent shielded.

3.2.1 Synthesis of Hexapeptides

Ac-PNGTAV-NH_2 and Ac-ANGTAV-NH_2 were synthesized via solid phase methods using a DuPont RAMPS semi-manual peptide synthesizer. Fmoc-amino acids were activated as OPfp esters and were coupled to RapidAmide resin (DuPont) in the presence of HOBT. Subsequent amino acids were also activated as Fmoc-AA-OPfp esters and coupled in the presence of HOBT. The coupling steps were monitored for completeness using the Kaiser test. The Fmoc group was deprotected with 50% piperidine/DMF. After chain extension was completed, the N-terminus was deprotected as before and capped with acetic anhydride while the peptide was still attached to the resin. The resin was simultaneously side-chain deprotected and cleaved from the resin using TFA to yield crude product. This crude residue was purified by trituration from methanol with 1/1 Et_2O/Hexane.

PHYSICAL DATA

Ac-PNGTAV-NH_2

^1H NMR (d_6-DMSO) δ: 0.75 (6H, m); 1.0 (3H, d); 1.2 (d, 3H); 1.3 (1H, m); 1.8 (7H, m); 2.5 (2H, m); 3.4 (2H, m); 3.8 (2H, m); 4.0 (2H, m); 4.1 (1H, m); 4.2 (1H, m); 4.4 (1H, q); 4.6 (1H, q); 6.9 (1H, s); 7.0 (1H, s); 7.3 (1H, s); 7.4 (1H, s); 7.6 (1H, t); 7.7 (1H, d); 7.9 (1H, d); 8.2 (1H, d); 8.3 (1H, d).

Ac-ANGTAV-NH_2

^1H NMR (d_6-DMSO) δ: 0.75 (6H, m); 1.0 (3H, d); 1.3 (6H, m); 1.8 (3H, s); 1.9 (1H, m); 2.5 (2H, m); 3.7 (2H, d); 3.9 (1H, m); 4.0 (1H, m); 4.2 (3H, m), 4.5 (1H, m); 4.9 (1H, d); 6.9 (1H, s); 7.0 (1H, s); 7.3 (1H, s); 7.4 (1H, s); 7.6 (1H, d); 7.7 (1H, d); 7.7 (1H, d); 7.9 (2H, m); 8.1 (1H, d); 8.2 (1H, d).

3.2.2 1D Hydrogen Exchange Experiments

Nuclear magnetic resonance spectra (^1H) were recorded on a Bruker 300 MHz spectrophotometer. ^1H NMR shifts were reported in parts per million (ppm) downfield from tetramethylsilane as an internal standard on the δ scale. Peak identification was determined by homonuclear decoupling experiments.

First-order hydrogen-deuterium exchange rates were measured by transfer of deuterium from D_2O to the peptide. The peptide was lyophilized from water (pH 4.0) and redissolved in enough d_6-DMSO to yield a 20 mM solution. The exchange rates were measured at 25°C after the addition of 10% D_20 to the DMSO sample. The NMR signal of the amide protons were recorded over time, integrated, and normalized to the α-proton of the asparagine residue.

RESULTS AND DISCUSSION

The integral values of the amide proton peaks were plotted as a function of time where 0 minutes indicates the addition of D_20 (Figure 3-9). The exchange rate was assumed to be first order, thus the slope of the line represents the observed exchange rate.

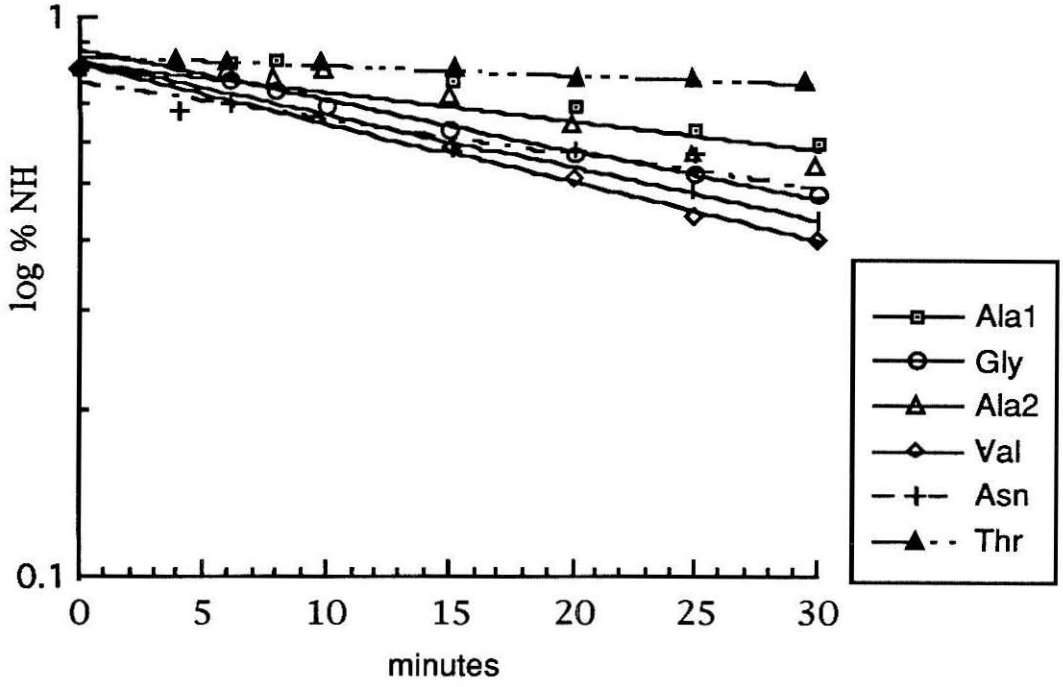


Figure 3-9 Amide proton exchange rates of Ac-ANGTAV-NH_2. The profile for Ac-PNGTAV-NH_2 is similar.

The only proton in both peptides which exhibited a drastically reduced exchange rate was the threonine backbone amide. Thus this amide was concluded to be solvent shielded.

CONCLUSIONS

In order to obtain quantitative information from these experiments it would have been necessary to recalculate the observed exchange rate over a range of pH values. Several factors influenced the decision to seek alternate experiments to obtain structural information. The proton exchange experiment was highly sensitive to temperature, pH, and concentration. Thus, not only was the accumulation of data time-consuming, but the data obtained could only be compared on an individual basis. Since the research

was focused on correlating structure and function, it was necessary to find an alternate approach which would allow us to compare information within a series.

The deciding factor was the upgrade in available NMR equipment which allowed us to aquire 2-D NMR data. Our approach thus switched to studying the secondary structural features present within the peptidyl acceptor using 2-D NOE experiments.

Both the computational study and the amide exchange experiment indicated that the threonine backbone amide was solvent shielded. The modelling study suggested that there was a high probability in substrate peptides for formation of a hydrogen bond between the side chain carbonyl oxygen of asparagine and the backbone amide of threonine. Further experiments were necessary to clearly demonstrate whether this interaction was present.

Appendix 3-1 CHARMM input file used for minimization

```
*MINIMIZATION
*
OPEN READ UNFORMATTED UNIT 1 NAME TOPH19.MOD
READ RTF UNIT 1
CLOSE UNIT 1

OPEN READ UNFORMATTED UNIT 2 NAME PARAM19.MOD
READ PARA UNIT 2
CLOSE UNIT 2

READ SEQUENCE CARD
* Ac-PNGTAV-NH2
*
8
ACE PRO ASN GLY THR ALA VAL BLK
GENERATE NGT SETUP

IC PARAM
IC SEED 2 CA 2 C 3 N
IC BUILD
IC FILL
IC PURGE

SET 3 3.0

LABEL START

NOE
        RESET
        END

NOE
        ASSIGN SELE ATOM NGT 3 ND2 END SELE ATOM NGT 5 OG1 END 3.0 0.2 0.2
        SCALE @ 3
        TEMP 300
        RMS
        LIST
        END

MINI ABNR NSTEP 50 NPRINT 50
        IC FILL
        INCR 3 BY -0.5
        IF 3 GT 0.0 GOTO START

NOE
        RESET
        END

MINI ABNR NSTEP 250 NPRINT 50
        IC FILL
```

```
OPEN UNIT 7 WRITE FORMATTED NAME PNGTAV.IC
WRITE IC UNIT 7
* INTERNAL COORDINATES
*
OPEN UNIT 8 WRITE FORMATTED NAME PNGTAV.CHR
WRITE COOR CARD UNIT 8
* X, Y, Z COORDINATES
*

OPEN UNIT 9 WRITE FORMATTED NAME PNGTAV.ENG
WRITE ENERGY UNIT 9
* ENERGY
*
STOP
```

Appendix 3-2 CHARMM input file used for dynamics simulation

```
*DYNAMICS-SIMULATED ANNEALING
*
OPEN READ UNFORMATTED UNIT 1 NAME TOPH19.MOD
READ RTF UNIT 1
CLOSE UNIT 1

OPEN READ UNFORMATTED UNIT 2 NAME PARAM19.MOD
READ PARA UNIT 2
CLOSE UNIT 2

READ SEQUENCE CARD
* Ac-PNGTAV-NH2
*
8
ACE PRO ASN GLY THR ALA VAL BLK
GENERATE NGT SETUP

OPEN UNIT 7 READ FORMATTED NAME PNGTAV.CHR
READ COOR CARD UNIT 7
CLOSE UNIT 7

OPEN UNIT 8 WRITE UNFORM NAME DYN.COR
OPEN UNIT 9 WRITE FORM NAME DYN.RES
OPEN UNIT 10 WRITE FORM NAME DYN.CHR
WRITE COOR CARD UNIT 10

SHAKE BONH TOL 1.0E-7

DYNA VERLET STRT NSTEP 10000 TIMESTEP 0.001 IUNCRD 10 NSAVC 10 TWINDH
5.0 TWINDL -5.0 IASORS 0 IASVEL 1 ICHECW 1 IEQFRQ 0 IUNWRITE 11 FIRSTT
298.0 FINALT 598.0 ISEED 2718 INBFRQ 0 CUTNB 11.5 CTTONNB 9.5 CTOFNB 10.5
EPSI 10.0 VDW NBXMOD 5 VSWITCH VDISTANCE

OPEN UNIT 10 WRITE UNFORM NAME DYN.COOR
OPEN UNIT 11 READ FORM NAME DYN.RES
OPEN UNIT 13 WRITE FORM NAME DYN2.RES

DYNA VERLET RESTART NSTEP 5000 TIMESTEP 0.001 IUNCRD 10 NSAVC 10
TWINDH 5.0 TWINDL -5.0 IASORS 0 IASVEL 1 ICHECW 1 IEQFRQ 0 IUNWRITE 13
IUNREAD 11 FIRSTT 598.0 FINALT 598.0 ISEED 230259 INBFRQ 20 CUTNB 11.5
CTTONNB 9.5 CTOFNB 10.5 EPSI 10.0 VDW NBXMOD 5 VSWITCH VDISTANCE

OPEN UNIT 10 WRITE UNFORM NAME DYN.COOR
OPEN UNIT 11 READ FORM NAME DYN2.RES
OPEN UNIT 13 WRITE FORM NAME DYN3.RES

DYNA VERLET RESTART NSTEP 10000 TIMESTEP 0.001 IUNCRD 10 NSAVC 10
TWINDH 5.0 TWINDL -5.0 IASORS 0 IASVEL 1 ICHECW 1 IEQFRQ 0 IUNWRITE 13
IUNREAD 11 FIRSTT 598.0 FINALT 298.0 ISEED 230259 INBFRQ 20 CUTNB 11.5
CTTONNB 9.5 CTOFNB 10.5 EPSI 10.0 VDW NBXMOD 5 VSWITCH VDISTANCE
```

```
OPEN UNIT 10 WRITE UNFORM NAME DYN.COOR
OPEN UNIT 11 READ FORM NAME DYN3.RES
OPEN UNIT 13 WRITE FORM NAME DYN4.RES

DYNA VERLET RESTART NSTEP 10000 TIMESTEP 0.001 IUNCRD 10 NSAVC 10
TWINDH 5.0 TWINDL -5.0 IASORS 0 IASVEL 1 ICHECW   1 IEQFRQ 0 IUNWRITE 11
IUNREAD 13 FIRSTT 298.0 FINALT 298.0 ISEED 230259 INBFRQ 20  CUTNB 11.5
CTTONNB 9.5 CTOFNB 10.5 EPSI 10.0 VDW NBXMOD 5 VSWITCH VDISTANCE

OPEN UNIT 10 WRITE FORM NAME DYN.CRD
WRITE COOR CARD UNIT 10
*DYNAMICS RUN FOR PNGTAV
*

STOP
```

1. Marshall, R.D. "The Nature and Metabolism of the Carbohydrate-Peptide Linkages of Glycoproteins," *Biochem. Soc. Symp.* **1974**, *40*, 17-26.

2. Mononen, I; Karjalainen, E. "Structural comparison of Protein Sequences Around Potential N-Glycosylation Sites," *Biochim. Biophys. Acta* **1984**, *788*, 364-367.

3. Gavel, Y.; von Heijne, G. "Sequence Differences Between Glycosylated and Non-glycosylated Asn-Xaa-Thr/Ser Acceptor Sites: Implications for Protein Engineering," *Protein Engineering* **1990**, *3*, 433-442.

4. Van Rietschoten, J.; Granier, C.; Ronin, C.; Bouchilloux, S. In *Peptides 1978*; Siemon,I.Z.; Kupryszewski, G., Eds.; Proceedings of the 15th European Peptide Symposium; Wroclaw University Press: Wroclaw, 1979; p. 335.

5. Hart, G.W.; Brew, K.; Grant, G.A.; Bradshaw, R.A.; Lennarz, W.J. "Primary Structural Requirements for the Enzymic Formation of the N-Glycosidic Bond in Glycoproteins. Studies with Natural and Synthetic Peptides," *J. Biol. Chem.* **1979**, *254*, 9747-9753.

6. Bause, E.; Legler, G. "The Role of the Hydroxy Amino Acid in the Triplet Sequence Asn-Xaa-Thr(Ser) for the N-Glycosylation Step During Glycoprotein Biosynthesis," *Biochem. J.* **1981**, *195*, 639-644.

7. a) Ricart, J.M.; Perez, J.J.; Pons, M.; Giralt, E. "Conformational Basis of N-Glycosylation of Proteins: Conformational Analysis of Ac-Asn-Ala-Thr-NH2," *Int. J. Biol. Macromol.* **1983**, *5*, 279-282. b) Aubert, J.P.; Biserte, G.; Loucheux-Lefebvre, M.H. "Carbohydrate-peptide Linkage in Glycoproteins," *Arch. Biochem. Biophys.* **1976**, *175*, 410-418.

8. Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminthan, S.; Karplus, M. "CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations," *J. Comp. Chem.* **1983**, *4*, 187-217.

9. Bause, E.; Hettkamp, H.; Legler, G. "Conformational Aspects of N-Glycosylation of Proteins," *Biochem. J.* **1982**, *203*, 761-768.

10. Wilson, S.R.; Cui, W.; Moskowitz, J.W.; Schmidt, K.E. "Conformational Analysis of Flexible Molecules: Location of the Global Minimum Energy Conformation by the Simulated Annealing Method," *Tetrahedron Lett.* **1988**, *29*, 4373-4376.

11. McKelvey, D.R. "Application of CHARMM to the Conformational Aspects of N-Glycosylation of Polypeptides"; summary of postdoctoral work performed in laboratories of Prof. C.L. Brooks; Carnegie Mellon University: Pittsburgh, PA, 1989.

12. Klotz, I.M.; Frank, B.H. "Deuterium-Hydrogen Exchange in Amide N-H Groups," *J. Am. Chem. Soc.* **1965**, *87*, 2721-2728.

13. Molday, R.S.; Englander, S.W.; Kallen, R.G. "Primary Structure Effects on Peptide Group Hydrogen Exchange," *Biochemistry* **1972**, *11*, 150-158.

Chapter 4
Kinetic and Structural Studies on Tripeptidyl Acceptors

The strict requirement of carbohydrate attachment at -Asn-Xaa-Thr/Ser- sites suggests that the polypeptide acceptor should have a very specific conformation, at least temporarily, in order to be recognized at the active site of the transferase. Based upon studies with peptide derivatives which contained modifications of the hydroxy amino acid, it was established that this residue was required for catalytic activity (**1**). These studies further suggested that a hydrogen bond interaction between the side chain of the asparagine and the hydroxy amino acid existed. Consequently, it was proposed that peptidyl acceptors must be able to adopt a conformation that would position these side chains close together in space.

Glycosyl transfer acceptor studies with synthetic peptides have been documented by Aubert *et al.* (**2**), Bause *et al.* (**3**), Aubry *et al.* (**4**) and Abbadi *et al.* (**5**). In addition, structural studies involving circular dichroism (**2, 3a, 4**), infrared, and X-ray analyses (**4, 5**) have been carried out on peptides related to the acceptor sequence. In all cases a clear indication emerges that the presence or absence of acceptor properties could be related to secondary structure forming tendencies in solution. Specifically, β-turn (**2**), "loop" (**3b**) and Asx-turn (**5**) conformations have been proposed.

X-ray crystal structures have been used to analyze the conformations which occur in the polypeptide backbone of several N-glycoproteins (**6**). These studies suggest that there may be a preference for carbohydrate attachment at external bends of protein structure. For example, in virus haemagglutinin, hemocyanin, and α-lactalbumin, the carbohydrates were shown to be attached at -Asn-

Xaa-Thr/Ser- sites located within β-turns. Aubert *et al.* (**2**) and Beeley (**7**) analyzed a number of known glycoproteins with the predictive procedures of Chou and Fasman (**8**), and concluded that there was a high statistical probability for glycosylated asparagine residues to be located within β-turns.

The β-turn is a fundamental structural motif found in proteins which induces the polypeptide chain to reverse its overall direction. β-turns were first described by Geddes (**9**) and later characterized by Venkatachalam (**10**). They involve four amino acid residues (numbered *i* to *i+3*) (Figure 4-1) and are classified according to the backbone dihedral angles of the central residues in the turn (**11**).

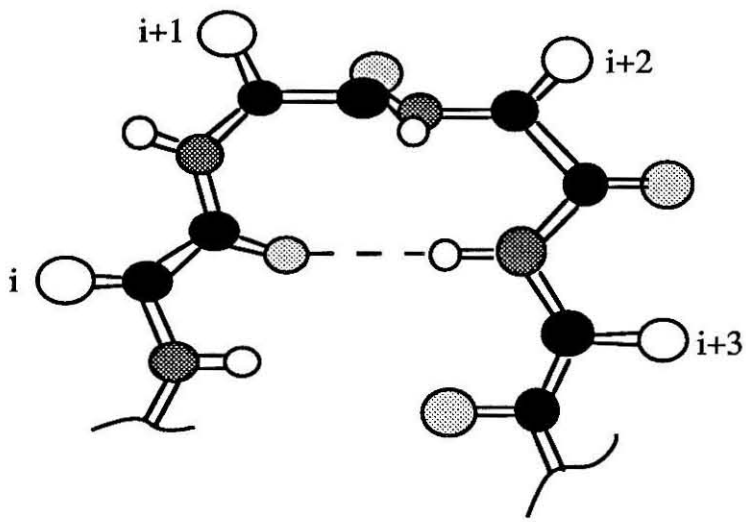


Figure 4-1 Representation of a type I β-turn.

β-turns are stabilized by a hydrogen bond between the amide proton of residue *i+3* and the carbonyl oxygen of residue *i*. Turns are frequently located at the surface of proteins and are usually composed of polar amino acid residues. These polar side chains are generally directed outward due to steric congestion. Biological

activity has often been related to β-turn formation in proteins; for example, turns serve frequently as sites for receptor binding, antibody recognition and post-translational modification. The intrinsic compactness and the exposure of side chains in β-turns makes them ideal recognition elements. However it is unclear whether these biological roles are due strictly to conformational recognition or due instead to propitious surface localization. In regard to glycosylation, it is important to remember that this process is a co-translational event and thus the structure observed in the final globular protein does not necessarily reflect the polypeptide conformation that was present at the time of modification. The amount of secondary structure present in nascent polypeptide chains during glycosylation is not known.

The amount of folding present in exogenous small acceptor peptides is also relevant to the discussion of transferase recognition. Ac-Asn-Ala-Thr-NH_2 is the smallest peptide to be glycosylated *in vitro* (**12**). Accordingly this tripeptide sequence must possess at least the main conformational requirements for the enzymatic process. Small peptides exist in aqueous solution as statistical mixtures of flexible conformers. However, it has been observed that significant populations of β-turn conformations exist in short linear peptides in aqueous solution (**13**). It has also been demonstrated that small peptides assume ordered conformations when they are transferred into hydrophobic environments or into solvents less polar than water (**14**). Since glycosylation most likely occurs at or close to the membrane interface, the peptidyl acceptor is likely to adopt a defined conformation. Thus, it is possible that a turn

conformation could occur and therefore the structure forming potential of the acceptor should be considered.

Statistical analyses of β-turns in proteins of known structure have clearly shown that there are residues which occur more frequently in turns, and also that there are strong positional preferences for some residues (**15**). Polypeptides composed solely of L-amino acids tend to adopt type I β-turn conformations. Interestingly, Asn residues favor the *i+1* position of these turns.

There are two type I β-turns that Ac-Asn-Ala-Thr-NH_2 could adopt (Figure 4-2). In Figure 4-2, structure **a**, the carbonyl oxygen of the acetyl bocking group is hydrogen bonded to the backbone amide of the hydroxy amino acid. In structure **b**, the carbonyl oxygen of the asparagine is hydrogen bonded to the amide proton of the carboxy terminal blocking group.

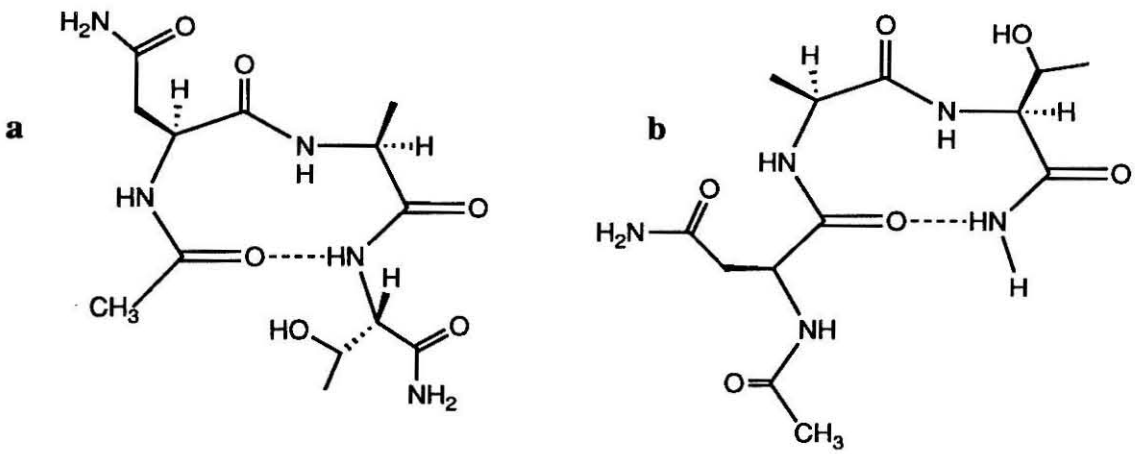


Figure 4-2 The two possible β-turn conformations of Ac-Asn-Ala-Thr-NH_2.

In order to produce the proposed hydrogen bond interaction between the side chains of asparagine and the hydroxy amino acid within the confines of a β-turn, the asparagine residue must occupy

the *i* or *i+2* position within the β-turn. This interaction is not possible when asparagine is located in the *i+1* position. Thus although structure a is predicted to be the most stable β-turn due to positional preferences, structure b is the only conformation which brings the side chains of Asn and the hydroxy amino acid residue within hydrogen bonding distance. This analysis suggests that β-turn formation is unlikely to be important for recognition. The idea that β-turn formation is crucial to glycosylation is further weakened by the realization that although Ac-Asn-Pro-Thr-NH_2 is not a substrate for the transferase (**16**), it exists as a highly stabilized turn in which Asn occupies the *i* position (**17**)

Moreover, Mononen and Karjalainen (**18**) observed that, when they predicted the secondary structure of a large number of known glycoproteins, 70% of the -Asn-Xaa-thr/Ser- sites occurred in β-turns, 20% in β-sheets and 10% in α-helices. Also no difference between glycosylated and non-glycosylated sequences was observed. Their findings were later confirmed by Gavel and von Heijne (**19**) who calculated the β-turn, β-sheet, and α-helix potentials as a function of position using an improved scale. They found no indications that glycosylated -Asn-Xaa-Thr/Ser- sites were more prone to form β-turns than non-glycosylated -Asn-Xaa-Thr/Ser- sites. Thus the involvement of β-turn conformation in glycosylation is not substantiated.

Another conformation which has been suggested to facilitate a hydrogen bond interaction between Asn and Thr is a (*i+4*) "loop structure" (**3b**). This conformation was loosely described by Bause

as a turn composed of five instead of four amino acid residues (Figure 4-3).

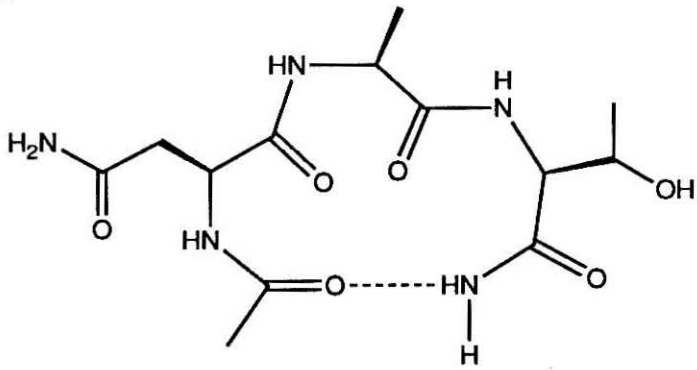


Figure 4-3 Representation of the (*i*+*4*) loop.

No structural data supports this model.

The final conformation which has been implicated in glycosylation is the Asx-turn (**4**). The Asx-turn is topologically equivalent to the β-turn. It is characterized by an interaction closing a 10-membered cycle and involving the side chain carbonyl oxygen of an Asn or Asp residue and the amide proton of the peptide group two residues ahead in the sequence (Figure 4-4) (**20**).

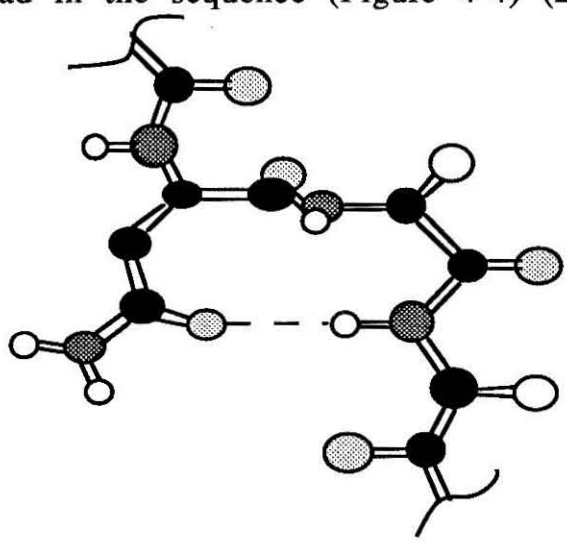


Figure 4-4 Representation of an Asx-turn.

Although the Asx-turn occurs less frequently than the β-turn, about 18% of Asn and Asp residues in proteins occur in this conformation. Asx-turns have been identified in the crystal structures of α-chymotrypsin, prealbumin, and carboxypeptidase A (**21**).

An analysis of the acceptor Ac-Asn-Ala-Thr-NH_2 indicates that not only could it adopt an Asx-turn (Figure 4-5), but that this

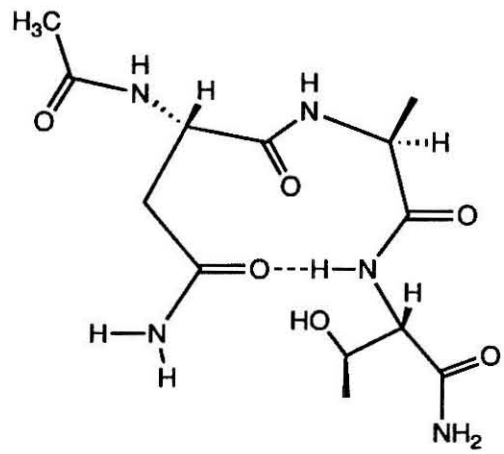


Figure 4-5 The possible Asx-turn structure of Ac-Asn-Ala-Thr-NH_2.

structure would facilitate an interaction between the side chains of Asn and Thr.

More structural information is needed to either consolidate one of the proposed conformational models, or to determine if indeed solution conformational tendencies play any role in the glycosylation process. Thus, a series of tripeptides which contained the -Asn-Xaa-Thr/Ser- marker sequence and differed in the nature of the central residue were investigated. This study included an evaluation of the properties of each peptide as oligosaccharide acceptors in the transferase assay and a parallel examination of the solution

conformational properties using a combination of 1- and 2-D NMR experiments.

EXPERIMENTAL METHODS

The solution conformations of six peptides (Ac-Asn-Ala-Thr-NH_2, Ac-Asn-Leu-Thr-NH_2, Ac-Asn-Asp-Thr-NH_2, Ac-Asn-(D)Ala-Thr-NH_2, Ac-Asn-Pro-Thr-NH_2, Ac-Asn-AIB-Thr-NH_2) were examined through the use of amide proton temperature coefficients, $^3J_{HN\alpha}$ coupling constants (**22**) and NOE studies. The information on through-space interactions was observed through the use of a CAMELSPIN or 2-D ROESY NMR experiments (**23**) in dimethyl sulfoxide. This polar, but aprotic, solvent was chosen since it can be considered to mimic the polar membrane interface where the nascent polypeptide substrate and the membrane associated enzyme oligosaccharyltransferase are localized better than either water or a non-polar organic solvent. Also, it has been shown that oligosaccharyltransferase-catalyzed glycosylation is significantly enhanced in the presence of non-denaturing levels (up to 13.5%) of dimethyl sulfoxide (**24**). This effect might be due to an increase in the population (as observed by CD studies) of a peptide conformer recognized by the oligosaccharyltransferase.

4.1 Peptide Synthesis

The following peptides were synthesized using known solution phase methods (**26**): Ac-Asn-Pro-Thr-NH_2, Bz-Asn-Pro-Thr-NHMe, Ac-Asn-Leu-Thr-NH_2, Bz-Asn-Leu-Thr-NHMe, Ac-Asn-Ala-Thr-NH_2, Ac-

Asn-(D)Ala-Thr-NH_2, Bz-Asn-(D)Ala-Thr-NHMe. A series of generalized reaction procedures is outlined below.

BOC-Xaa-Thr-OMe

To a solution of cold threonine methyl ester (0.85g, 5 mmol) in 5 mL of DMF was added triethylamine (693 μL, 5 mmol). A solution of dicyclohexylcarbodiimide (5.3 mmol), 1-hydroxybenzoltriazole (5.3 mmol), and BOC-Xaa-OH (5.3 mmol) was added slowly with stirring. The reaction was stirred for 12 h. Dicyclohexylurea was filtered off and rinsed with DMF. The reaction mixture was concentrated *in vacuo*. The oily residue was flash chromatographed (Rf 0.2 in 10% methanol-chloroform) to yield the pure dipeptide in about 70% yield.

BOC-Asn-Xaa-Thr-OMe

Trifluoroacetic acid (2.85 mol) was added to a solution of BOC-Xaa-Thr-OMe (1.6 mmol) in 2.85 mL of dichloromethane. The mixture was stirred for 2 h and reduced under a stream of argon. The residue was taken up in DMF and triethylamine (1.7 mmol) was added. The mixture was cooled to 0°C. BOC-Asn-OpNp (1.8 mmol) in 2 mL of DMF was added dropwise. The reaction was stirred for 12 h. The mixture was then concentrated *in vacuo*. The oily yellow residue was flash chromatographed (Rf 0.2 in 10% methanol-chloroform) to yield the tripeptide (about 80%).

BOC-Asn-Xaa-Thr-NH_2

BOC-Asn-Xaa-Thr-OMe (1 mmol) was dissolved in 2 mL of methanol saturated with ammonia. The vessel was sealed tightly and stirred for

18 h. The mixture was reduced under a strearn of argon and quantitiatively yielded the title compound.

Ac-Asn-Xaa-Thr-NH_2

Trifluoroacetic acid (1 mL) was added to a solution of BOC-Asn-Xaa-Thr-NH_2 (0.54 mmol) in 1 mL of dichloromethane. The mixture was stirred for 2 h and then reduced under a stream of argon. The residue was resuspended in 2 mL of 50% aqueous dioxane, cooled to 0°C, and triethylamine (80 µl, 0.54 mmol) was added. Acetic anhydride (102 µL, 0.54 mL) was added dropwise and the mixture was stirred for 18 h. Dioxane was condensed off *in vacuo* and the remaining aqueous solution was purified via Dowex-50 (H^+ form) and Dowex-1 ($0H^-$ form) columns. The pure tripeptide was obtained in about 70%.

BOC-Asn-Xaa-Thr-NHMe

BOC-Asn-Xaa-Thr-OMe (1 mmol) was dissolved in 2 mL of methanol saturated with monomethylamine. The vessel was sealed tightly and stirred for 18 h. The mixture was reduced under a stream of argon and quantitatively yielded the title compound.

Bz-Asn-Xaa-Thr-NHMe

Trifluoroacetic acid (1 mL) was added to a solution of BOC-Asn-Xaa-Thr-NHMe (0.5 mmol) in 1 mL of methylene chloride. The mixture was stirred for 4 h and then reduced under argon. The residue was resuspended in 2 mL of dimethylformamide, cooled to 0°C, and triethylamine (80 µL, 0.54 mmol) was added. Benzoic anhydride (0.116 g, 0.54 mmol) in 1 mL dimethylformamide was added dropwise with

stirring. The mixture was stirred for 10 h. Dimethylformamide was condensed off *in vacuo* and the residue was purified by preparative TLC (Rf 0.5 in 10/1; chloroform/methanol) to yield pure protected tripeptide (~80%).

Ac-Asn-AIB-Thr-NH_2 and Bz-Asn-AIB-Thr-NHMe were synthesized from BOC-Asn-AIB-Thr-OMe as above, however this material was synthesized via the following modified procedure:

BOC-Asn-AIB-OMe

To a solution of cold α-aminoisobutyric methyl ester (0.56 g, 3.6 mmol) in 5 mL of DMF was added 1 equivalent of triethylamine (504 μL, 3.6 mmol). BOC-Asn-OpNp (1.28 g, 3.6 mmol) in 5 mL DMF was added dropwise to the reaction mixture. The reaction was stirred for 6 h and then concentrated *in vacuo*. The oily residue was flash chromatographed (Rf 0.3 in 5/1 $CHCl_3$/MeOH) to yield 0.90 g (81%) of pure dipeptide.

BOC-Asn-Aib-OH

BOC-Asn-Aib-OMe (0.90 g, 2.7 mmol) was saponified in 2.25 mL of 2 N NaOH. After two hours the reaction mixture was neutralized with 1 N HCl and lyophilized. The residue was suspended in MeOH and filtered to remove salts.

BOC-Asn-Aib-Thr-OMe

To a solution of cold threonine methyl ester (0.15 g, 0.88 mmol) in 3 mL of DMF was added 1 equivalent of triethylamine (123 μL, 0.88 mmol). A solution of BOCAsn-Aib-OH (0.3 g, 0.92 mmol), DCC (920 μL of a

1 M solution in DMF), and HOBT (0.124 g, 0.92 mmol) in 1 mL DMF was added dropwise to the reaction mixture with stirring. The reaction was left for 12 hours and condensed *in vacuo*. The residue was flash chromatographed (Rf 0.45 in 10/1 $CHCl_3$/MeOH) to yield pure tripeptide in 71% yield.

Ac-Asn-Asp-Thr-NH_2 and Bz-Asn-Asp-Thr-NH_2 were synthesized from Boc-Asn-Asp(OBn)-Thr-OMe using the following modified procedure.

Ac-Asn-Asp(OBn)-Thr-OMe

Boc-Asn-Asp(OBn)-Thr-OMe (0.5 g, 0.84 mmol) was suspended in 2 mL methylene chloride and 2 mL of trifluoroacetic acid. The solution was sealed tightly and stirred 4 h. The mixture was reduced under a stream of argon and lyophilized from water. The residue was suspended in 3 mL of dimethylformanide and acetic anhydride (95 μL, 1.01 mmol) and triethylamine (140 μL, 1.01 mmol) were added. The solution was stirred for 4 h. The solution was reduced *in vacuo*. Purification of the oily residue via preparative TLC (5/1 chloroform/methanol, Rf 0.4) yielded the pure dipeptide (~75%).

Ac-Asn-Asp-Thr-OMe

Ac-Asn-Asp(OBn)-Thr-OMe (0.3 g, 0.6 mmol) was suspended in 5 mL of methanol. Approximately 60 mgs of 5% Pd-C catalyst was added. The mixture was hydrogenated at atmospheric pressure for 2 h after which time the catalyst was filtered off through a pad of celite. The

filtrate was reduced and afforded the product in essentially quantitative yield.

Ac-Asn-Asp-Thr-NH_2

Ac-Asn-Asp-Thr-OMe (0.24 g) was suspended in 3 mL of methanol saturated with ammonia. The reaction flask was sealed tightly and stirred for 12 h. The solution was reduced under a stream of argon and lyopholized from water to yield pure product (~85%).

PHYSICAL DATA

Ac-Asn-AIB-Thr-NH_2

^{1}H NMR (d_6-DMSO) δ: 8.43 (s, 1H), 8.18 (d, 1H, J = 7 Hz), 7.52 (s, 1H), 7.40 (d, 1H, J = 8.6 Hz), 7.11 (s, 1H), 7.05 (s, 1H), 7.00 (s, 1H), 4.61 (d, 1H, J = 7.9 Hz), 4.46 (q, 1H, J = 7 Hz), 4.07 (m, 1H), 3.93 (m, 1H), 2.50 (m, 2H), 1.82 (s, 3H), 1.33 (s, 6H), 1.01 (d, 3H, J = 6.4 Hz)

^{13}C NMR (d_6-acetone, D_2O) d: 177.4, 175.0, 174.5, 174.3, 172.5, 67.1, 59.2, 57.5, 51.0, 38.6, 24.3, 23.0, 22.0, 19.1

Mass spectroscopy $[MH]^+$ 360
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf = 0.33

Ac-Asn-(D)Ala-Thr-NH_2

^{1}H NMR (d_6-DMSO) δ: 8.33 (d, 1H, J = 6.9 Hz), 8.13 (d, 1H, J = 7.6 Hz) 7.60 (d, 1 H, J = 8.7 Hz), 7.39 (s, 1H, 7.18 (s, 1H) 7.09 (s, 1H), 6.94 (s, 1H), 4.93 (bs, 1H), 4.42 (q, 1H, J = 6.6 Hz), 4.27 (t, 1H, J = 7.3 Hz), 4.03 (m, 2H), 2.42 (m, 2H), 1.84 (s, 3H), 1.22 (d, 3H, J = 7.1 Hz), 0.97 (d, 3H, J = 6.3 Hz)

^{13}C NMR (d_6-acetone, D_2O) δ: 175.5, 174.8, 174.6, 174.5, 172.9, 67.1, 58.9, 50.9, 50.2, 37.0, 22.1, 19.0, 18.8

Mass spectroscopy $[MH]^+$ 346
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf = 0.23

Ac-Asn-Ala-Thr-NH_2

^1H NMR(d_6-DMS0) δ: 8.22 (d, lH, J = 7.0 Hz), 8.13 (d, lH, J = 7.1 Hz), 7.60 (d, 1H, J = 8.3 Hz), 7.43 (s, 1H), 7.06 (s, lH), 7.03 (s, lH), 6.97 (s, lH), 4.79 (d, 1H, J = 5.9 Hz), 4.52 (q, lH, J = 6.7 Hz), 4.22 (t, 1H, J = 7.0 Hz), 4.00 (m, 2H), 2.49 (m, 2H), 1.82(s, 3H), 1.22 (d, 3H, J = 7.1 Hz), 1.01 (d, 3H, J = 6.1 Hz)

^{13}C NMR (D_2O, d_6-acetone) δ: 175.5, 174.6, 174.3, 173.7, 172.9, 67.2, 59.0, 50.5, 43.5,36.7, 22.1, 19.0, 16.6

Mass spectroscopy $[MH]^+$ 346
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf = 0.23

Ac-Asn-Leu-Thr-NH_2

^1H NMR (d_6-DMSO) δ: 8.15 (m, 2H), 7.60 (d, lH, J = 8.4 Hz), 7.44 (s, 1H), 7.04 (s,2H), 6.95 (s, lH), 4.90 (bs, lH), 4.50 (t, lH, J = 7.0 Hz), 4.23 (q, lH, J = 7.4 Hz), 4.02(m, 2H), 2.46 (m, 2H), 1.82 (s, 3H), 1.59 (m, 1H), 1.51 (m, 2H), 1.00 (d, 3H, J = 6.2 Hz),0.86 (d, 3H, J = 6.5 Hz), 0.81 (d, 3H, J = 6.4 Hz)

^{13}C NMR (D_2O, d_6-DMSO) δ:175.0, 174.6, 174.3, 174.2, 173.167.2, 59.0, 53.0, 51. 1 ,50.8, 39.9, 36.5, 29.8, 29.5, 29.3, 24.6

Mass Spectroscopy $[MH]^+$ 388
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf=0.47

Ac-Asn-Pro-Thr-NH_2

^1H NMR (d_6-DMSO) δ: 8.27 (d, 1H, J = 7.8), 7.62 (m, 2H), 7.09 (s, 1H), 6.86 (s, 1H), 4.80 (q,1H, J = 7.1 Hz), 4.31 (q, 1H, J = 5.2 Hz), 4.02 (m, 2H), 3.68 (m, 2H), 2.45 (m, 2H), 2.06(m, 1H), 1.87 (m, 3H), 1.80 (s, 3H), 1.03 (d, 3H, J = 6.3 Hz)

^{13}C NMR (D_20, d_6-acetone) δ: 174.7, 174.2,173.9, 171.7, 67.1, 61.3, 59.0, 48.3, 48.2,36.4, 29.9, 24.9, 21.9, 19.2

Mass Spectroscopy $[MH]^+$372
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf= 0.40

Ac-Asn-Asp-Thr-NH_2

^1H NMR (d_6-DMSO) δ: 8.34 (d, lH, J = 6.4 Hz), 8.19 (d, 1H, J = 7.3 Hz), 7.53 (d, lH, J=8.4 Hz), 7.42 (s, 1H), 7.11 (s, lH), 7.07 (s, lH), 6.97 (s,

1H), 4.78 (bs, 1H), 4.49 (m,2H), 4.01 (m, 2H), 2.50 (m, 4H), 1.85 (s, 3H), 0.97 (d, 3H, J = 6.3 Hz)

^{13}C NMR (d_6-DMSO) δ: 172.1,171.7,171.4, 170.7, 169.7, 66.3, 58.7, 50.2, 49.9, 37.2, 22.5, 20.1

Mass Spectroscopy $[MH]^+$390
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf=0.11

Bz-Asn-Pro-Thr-NHMe

^{1}H NMR (d_6-DMSO) δ: 8.60 (d, 1H, J = 7.2 Hz), 7.74 (d, 2H, J = 7.4 Hz), 7.42 (m, 3H), 7.33 (t, 2H, J = 7.5 Hz), 7.23 (d, 1H, J = 4.2 Hz), 6.94 (s, 1H), 4.89 (q, 1H, J = 7.1 Hz),4.52 (d, 1H, J = 5.4 Hz), 4.23 (m, 1H), 3.94 (m , 2H), 3.69 (m , 2H), 2.49 (m, 2H), 2.45 (d , 3H, J = 4.5 Hz), 1.95 (m , 1H), 1.78 (m, 3H), 0.89 (d, 3H, J = 6.0 Hz)

^{13}C NMR (d_6-DMSO) δ: 172.1, 171.5, 170.9, 170.3, 166.1, 133.6, 13 1.5, 128.3, 127.6, 66.2, 60.4, 58.6, 48.6, 47.1, 36.8, 29.1, 25.8, 24.5, 20.2

Mass Spectroscopy $[MH]^+$448
TLC (5/1; chloroform/methanol) Rf = 0.54

Bz-Asn-AIB-Thr-NHMe

^{1}H NMR (d_6-DMSO) δ: 8.51 (d, 1H, J = 6.7 Hz), 8.42 (s, 1H), 7.72 (d, 2H, J = 7.5 Hz), 7.39 (m, 5 H), 7.18 (d, 1H, J = 8.7 Hz), 6.94 (s, 1 H), 4.58 (q, 1H, J = 6.9 Hz), 4.54 (bs,1H), 3.96 (bs, 1H), 3.87 (m, 2H), 2.44 (d, 3H, J = 4.3 Hz), 1.21 (d, 3H, J = 8.5 Hz), 0.85 (d, 3H, J = 6.4 Hz)

^{13}C NMR (d_6-DMSO) δ: 173.9, 171.9, 171.8, 170.5, 166.7, 133.8, 13 1.6, 128.3, 127.5,
69.2, 66.2, 58.8, 50.4, 51.1, 36.8, 25.8, 25.6, 24.6, 20.2

Mass Spectroscopy [MH] +436
TLC (5/1; chloroform/methanol) Rf = 0.48

Bz-Asn-Asp-Thr-NH_2

^{1}H NMR(d_6-DMSO) δ: 8.47 (d,1H, J = 7.3 Hz), 8.28 (d, 1 H, J = 7.5 H z), 7.71 (d, 2H, J = 7.4 Hz), 7.35 (m, 5H), 6.96 (d, 2H, J = 9.2 Hz), 6.84 (s, 1H), 4.63 (m, 2H), 4.45 (q, 1H, J = 6.9 Hz), 3.90 (m, 2H), 2.49 (m, 4H), 0.86 (d, 3H, J = 6.0 Hz)

^{13}C NMR (d_6-DMSO) δ: 172.1, 172.0, 171.8, 171.4, 170.5, 166.5, 133.9, 131.5, 128.3, 127.5, 66.3, 58.4, 50.8, 49.9, 36.9, 35.7, 20.0

Mass Spectroscopy $[MH]^+$ 452
TLC (4/1/1; *n*-butanol/water/acetic acid) Rf=0.42

Bz Asn-(d)-Ala-Thr-NHMe

^{1}H NMR(d_6-DMSO) δ: 8.59 (d, 1 H, J = 6.9 Hz), 8.02,(d, 1H, J = 6.7 Hz), 7.72 (d, 2H , J = 7.2 Hz), 7.61 (d, lH, J = 7.8 Hz), 7.49 (s, l H), 7.40 (q, lH, J = 7.2 Hz), 7.34 (d, 2H ,J = 7.6 Hz), 7.27 (s, lH), 6.82 (s, lH), 4.68 (d, lH, 5.5Hz), 4.59 (q, lH, J = 6.1 Hz), 4.22 (d, lH, J = 7.0 Hz), 3.93 (m, 2H), 2.49 (m, 5H), 1.10 (d, 3H, J = 6.9 Hz), 0.84 (d, 3H, J = 5.9 Hz)

^{13}C NMR (d_6-DMSO) δ: 172.4, 171.6, 171.0, 166.8, 133.8, 131.6, 128.3, 127.5, 66.3, 58.5, 51.0, 48.6, 36.9, 25.7, 20.0, 18.0

Mass Spectroscopy $[MH]^+$ 422
TLC (5/1; chloroform/methanol) Rf=0.38

Bz-Asn-Leu-Thr-NHMe

^{1}H NMR (d_6-DMSO) δ: 8.64 (d, 1H, J = 7.5 Hz), 8.20 (d, 1H, J = 8.0 Hz), 7.83 (d, 2H, J = 7.4 Hz), 7.60 (d, lH, J = 8.7 Hz), 7.50 (m, 4H), 6.98 (s, lH), 4.83 (d, lH, J = 5.4 Hz), 4.74 (q, lH, J = 6.2 Hz), 4.29 (m, lH), 4.04 (m, 1H), 4.00 (m, 1H), 2.57 (m, 5H), 1.59 (m,1H), 1.52 (m, 2H), 0.98 (d, 3H, J = 6.3 Hz), 0.84 (d, 3H, J = 6.5 Hz), 0.79 (d, 3H, J = 6.5 Hz)

^{13}C NMR (d_6-DMSO) δ: 176.7, 171.9, 171.7, 171.3, 171.2, 131.3, 128.1, 127.3, 107.6, 66.2, 58.3, 51.6, 36.5, 24.0, 23.0, 21.3, 19.9

Mass Spectroscopy $[MH]^+$ 464
TLC (5/1; chloroform/methanol) Rf = 0.42

4.2 NMR Spectroscopy

Ac-Asn-Leu-Thr-NH_2, Ac-Asn-Ala-Thr-NH_2, Ac-Asn-Asp-Thr-NH_2, Ac-Asn-Pro-Thr-NH_2, Ac-Asn-AIB-Thr-NH_2 and Ac-Asn-

(D)Ala-Thr-NH_2 were pretreated by co-evaporation with deuteriochloroform and subsequently dissolved in d_6-dimethyl sulfoxide (concentration approximately 4 mg/mL, 10 mM).

NMR experiments were performed at ambient temperature (298 K) unless otherwise noted. Peak assignments were determined either in 1-D by homonuclear single frequency decoupling experiments or through the use of a 2-D ^1H COSY experiment. In the former case the decoupler power used was maintained at the minimum required to instantaneously saturate the peak of interest. 2-D ROESY spectra were obtained in phase sensitive mode and were multiplied by a squared sine bell apodization function prior to Fourier transformation. The spectral widths in both dimensions were typically 5000 Hz. NOE's were detected using the 2-D spin-locked ROESY experiment, which is best suited to the detection of NOE effects in relatively small systems (MW 400-1500). These spectra were recorded at 299 K. In this experiment the spin lock was achieved by applying a 200 millisecond pulse through the decoupler channel. The relaxation delay was 1.5 s. The transmitter offset was positioned in the center of the spectrum. The ROESY experiment conditions (namely mixing time, radiofrequency power and offset frequency) were varied in initial studies in order to verify that NOE connectivities observed represented proximity and did not in fact result from Hartman-Hahn magnetization transfer (**26**). The data matrix consisted of 256 t_1 increments containing 1K complex points. It should be noted that the ROESY cross peaks, while useful in obtaining qualitative proximity information (<3Å) could not be easily

used to derive quantitative distance information in the same way as NOESY data (27).

The temperature dependence of the assigned amide proton shifts was determined between 299 K and 325 K. In all cases the chemical shift was found to vary linearly with temperature. A minimum of six temperature steps were recorded in each experiment. Temperature calibration of the actual probe temperature was performed using an ethylene glycol standard.

4.3 Oligosaccharyltransferase Assay

The oligosaccharyltransferase assay was performed as described in chapter 2.3.1.

RESULTS AND DISCUSSION

The eleven tripeptides examined in the oligosaccharyl transferase assay belong to two homologous series differing only in the protecting groups used at the carboxyl and amino termini. Six of the tripeptides were blocked at the amino terminus by a benzoyl group and a secondary amide at the carboxyl terminus. The benzoyl-protecting group was used to maximize the detection of any acceptor properties related to the polypeptide (as discussed in chapter 2). In the second series, the tripeptides were blocked at the amino terminus by an acetyl group and by a primary amide at the carboxyl terminus. Spectroscopic studies were performed on the acetyl-protected series since the spectra were devoid of any overlapping proton signals in the diagnostic amide region. By examining both series of peptides, the maximum information concerning both

acceptor properties and solution conformation were derived. Furthermore, substitution of the benzoyl group for the acetyl did not appear to perturb the solution conformation in the 2-D experiment.

The K_m values for the peptides are shown in Table 4-1. The binding affinities were enhanced, as predicted, when the benzoyl group was used as the blocking group. Within the two series, only those peptides with L-amino acids in the central position exhibited acceptor properties. Interestingly, Bz-Asn-Asp-Thr-NH_2 is a substrate, albeit far poorer than the corresponding alanine and leucine containing peptides. This finding is contrary to original reports indicating that peptides containing aspartic acid as the central residue were not tolerated in an oligosaccharide acceptor (**18**). The remaining tripeptides were all either very poor or non-acceptors, regardless of the N-terminal protecting group.

Table 4-1 Enzyme assay results from tripeptide studies.

	K_m /mM[a]	
peptide	X=Ac, Y=NH_2	X=Bz, Y=NHMe
X-Asn-Leu-Thr-Y	1.0	0.25
X-Asn-Ala-Thr-Y	2.0	ND[b]
X-Asn-Asp-Thr-Y	>50	3.3[c]
X-Asn-Pro-Thr-Y	>50	>50
X-Asn-(D)Ala-Thr-Y	>50	>30[d]
X-Asn-AIB-Thr-Y	>50	>30[d]

a Similar V_{max} values were obtained for each peptide; 500 dpm/min ($\sim 5x^{-3}$) nmol min^{-1} mg^{-1} microsomal protein at a DolPPNAGNAG concentration of 8.3 nM

b Not Determined

c Y=NH_2

d Residual activity apparent at high peptide concentrations. K_m lies between 30-50 mM.

The amide proton temperature coefficients for the six acetyl-protected peptides are shown in Table 4-2. These values provide a measure of the amide proton solvent accessibility and/or hydrogen bonding. In polar solvents, coefficients ranging from -4 to -7 ppb/K are typical of solvent exposed protons, whereas those in the 0 to -3 ppb/K range are diagnostic of solvent shielding (values close to 0 ppb/k have been observed for conformationally constrained cyclic peptides in chloroform/dimethyl sulfoxide mixtures (**28**)).

Table 4-2 Temperature dependence of amide proton chemical shifts ($-\Delta\delta/\Delta T$)[a] for Ac-Asn-Xaa-Thr-NH_2[b] in DMSO[c].

	Asn	Asn δNH	Asn δNH	Xaa	Thr	terminus	terminus
Leu	4.5	4.1	5.4	4.8	2.7	3.7	3.7
Ala	5.0	4.5	5.7	5.3	3.1	4.0	4.0
Asp	4.6	3.8	4.8	2.7	1.5	2.9	4.6
D-Ala	5.0	4.5	3.9	7.6	3.0	5.5	5.4
AIB	6.1	5.2	6.4	7.2	3.1	5.0	3.3
Pro	5.5	3.8	2.1	n/a	4.0	6.0	6.0

a In parts per billion Kelvin.
b Xaa = Leu, Ala, Asp, D-Ala, AIB, Pro
c Primary amide proton signals were assigned by comparison with analogous peptides from Bz-Asn-Xaa-Thr-NHMe series and from 2-D NMR data.

With the exception of Ac-Asn-Pro-Thr-NH_2, the threonine amide proton possessed the lowest temperature coefficient. In similar studies with threonine-containing peptides which lack the proximal asparagine residue, the temperature coefficient for the threonine amide proton lies in the "normal" range for an unsequestered proton (**29**). Thus, this effect does not appear to be residue-specific.

The $^3J_{HN\alpha}$ coupling constants for the amide protons within the tripeptide backbones are shown in Table 4-3. The angular dependence of amide proton-Cα coupling constants ($^3J_{HN\alpha}$) has been calibrated from protein structure information (**22**). In general, small coupling constants (2-4 Hz) are observed when the ϕ dihedral angles are in the -30 to -60° range and larger values (7-10 Hz) result from ϕ values of -100 to -140°. The former angles are associated with either regular α-helix secondary structure or the ϕ dihedral of the *n+1* residue within a type I β-turn, while the latter are associated with extended β-sheet structure, random coil or the ϕ dihedral of the *n+2* residue of a β-turn (**11**).

Table 4-3 Coupling constants (in Hertz) for Ac-Asn-Xaa-Thr-NH_2[a].

	Asn	Xaa	Thr
Leu	7.9	8.4	8.4
Ala	7.0	7.7	8.3
Asp	6.4	7.3	8.4
D-Ala	6.9	7.6	8.7
AIB	7.0	n/a	8.6
Pro	7.8	n/a	9.1

a Xaa = Leu, Ala, Asp, D-Ala, AIB, Pro.

Figures 4-6 to Figure 4-9 shows segments of the ROESY spectra for four of the acetyl-protected peptides examined. The 2-D ROESY spectrum of a representative peptidyl substrate, Ac-Asn-Ala-Thr-NH_2, is illustrated in Figure 4-6. (The corresponding spectra of Ac-Asn-Asp-Thr-NH_2 and Ac-Asn-Leu-Thr-NH_2 are not shown, but feature strictly analogous cross peaks.)

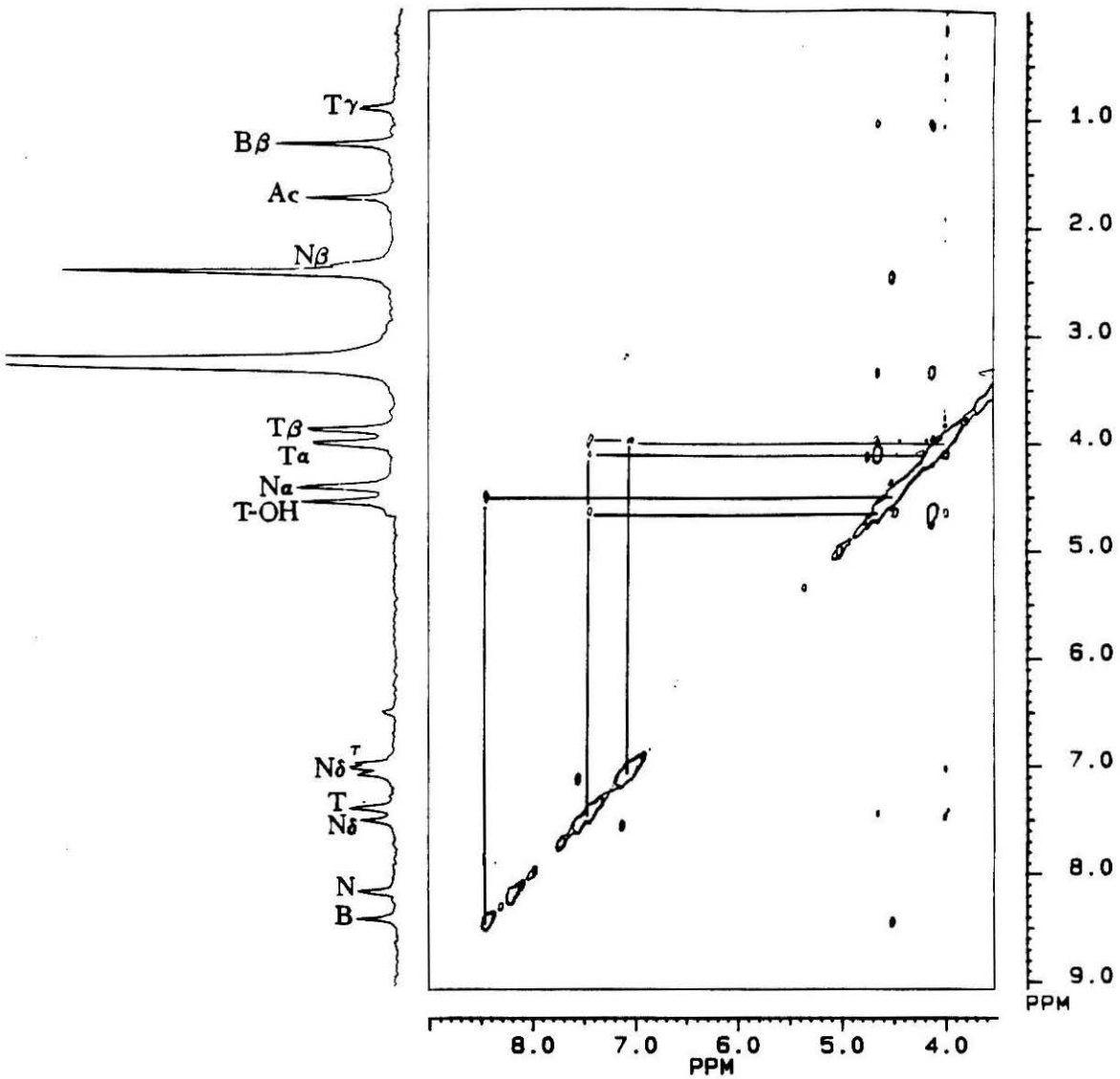


Figure 4-6 ROESY spectrum of Ac-Asn-Ala-Thr-NH_2.

All substrates are characterized by ROESY cross peaks $d_{\alpha N}$(Ac,Asn), $d_{\alpha N}$(Asn,Xaa), $d_{\alpha N}$(Xaa,Thr), $NH_2.d_{\beta N}$(Thr,terminus), $d_{\alpha N}$(Thr,Thr), $d_{\beta N}$(Thr,Thr) and $d_{\alpha N}$(Asn,Asn) which are summarized graphically in Figure 4-10a.

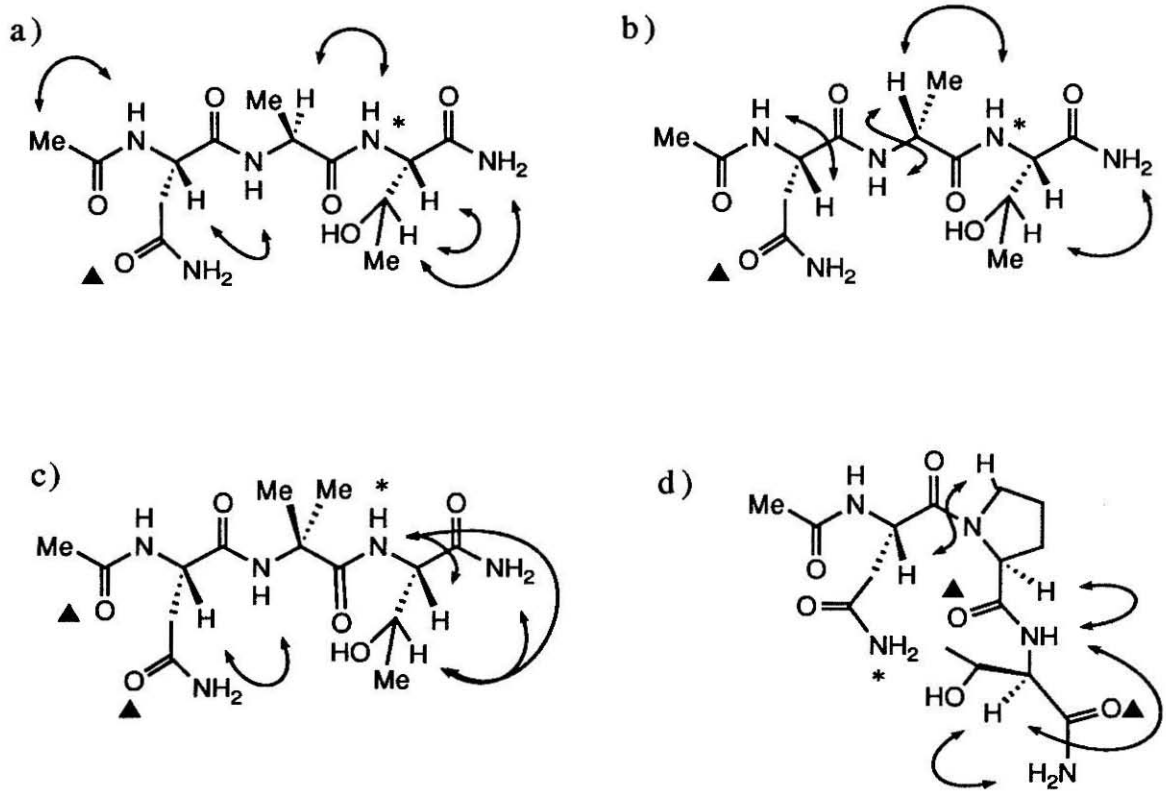


Figure 4-10 Graphical representations of connectivity and hydrogen-bonding information derived from spectroscopic studies of six acetyl-protected peptides. Key: ↔ NOE connectivity; * hydrogen bond donor; ▲ hydrogen bond acceptor. The structures shown are a) Ac-Asn-Ala-Thr-NH_2, b) Ac-Asn-D-Ala-Thr-NH_2, c) Ac-Asn-AIB-Thr-NH_2, and d) Ac-Asn-Pro-Thr-NH_2.

The ROESY data for Ac-Asn-(D)Ala-Thr-NH_2 is shown in Figure 4-7. (D)Alanine substitution results in distinct changes in the crosspeak pattern.

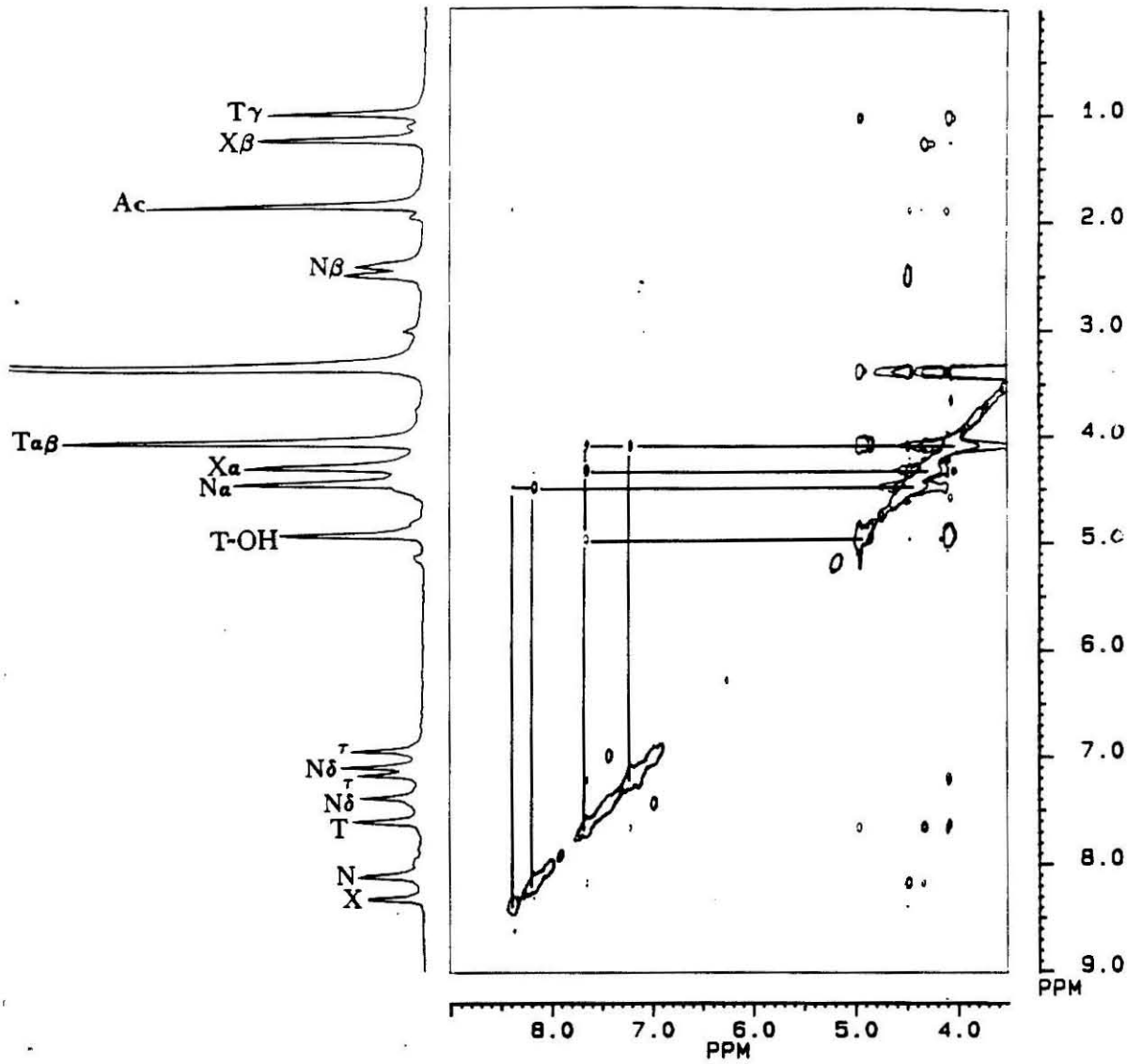


Figure 4-7 ROESY spectrum of Ac-Asn-(D)Ala-Thr-NH_2. X=(D)Alanine.

Firstly, the strong backbone sequential connectivity $d_{\alpha N}$(Asn,Xaa) observed in the substrates is replaced by a very weak (relative to other signals within the same spectrum) crosspeak in Ac-Asn(D)Ala-Thr-NH_2. Instead, a strong intraresidue connectivity $d_{\alpha N}$(Asn, Asn) appeared. Also, despite the change in chirality at the (D)alanine αC, the sequential NOE $d_{\alpha N}$(*d*Ala,Thr) is still present. This information is summarized in Figure 4-10b.

The ROESY spectrum for the peptide Ac-Asn-AIB-Thr-NH_2 is illustrated in Figure 4-8. The first notable feature of the spectrum is that the sequential and intraresidue interactions from the asparagine αCH are of very similar intensity. However since the AIB-αC is disubstituted, no sequential information can be derived for the central residue. The key interproton connectivities are illustrated in Figure 4-10c.

Figure 4-9 illustrates the ROESY spectrum for the non-acceptor peptide Ac-Asn-Pro-Thr-NH_2. The key crosspeaks observed in the ROESY spectrum are $d_{\delta a}$(Pro,Asn), $d_{\alpha N}$(Pro,Thr) and $d_{\alpha N}$(Thr, terminus). These are summarized in Figure 4-10d.

The three tripeptide substrates appear to share common conformational characteristics. First, the threonine NH is shielded from the bulk medium as evidenced by the low temperature coefficients (-1.5 to -3.1 ppb/K) when compared with other values from the same peptide. Thus this amide NH was considered to be shielded from the bulk medium by hydrogen-bonding interaction within the molecule. Due to ring size and amide constraints only two carbonyl groups are likely acceptors in such an interaction. These are the N-terminal acyl carbonyl group or the asparagine side chain

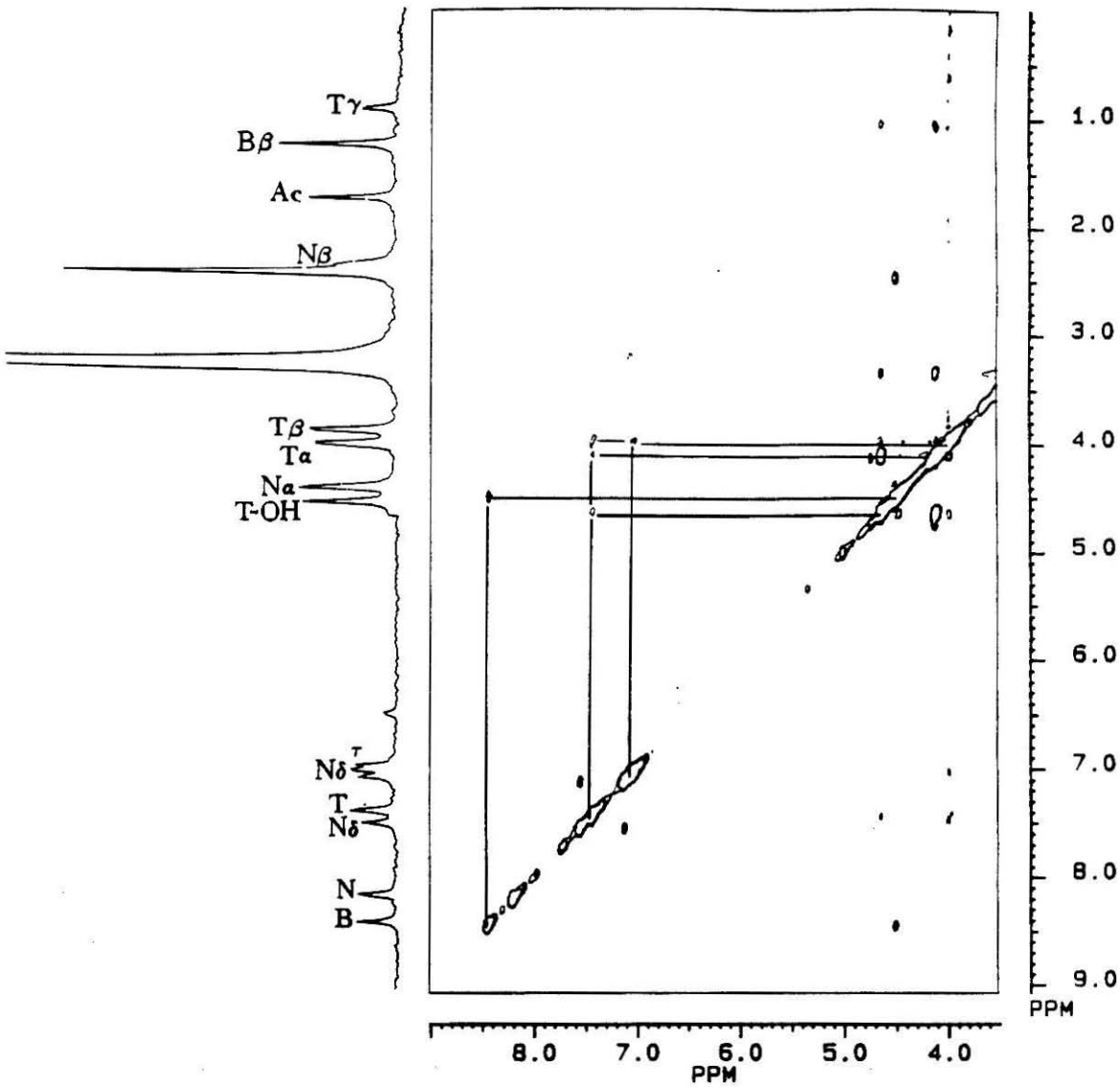


Figure 4-8 ROESY spectrum of Ac-Asn-AIB-Thr-NH_2.

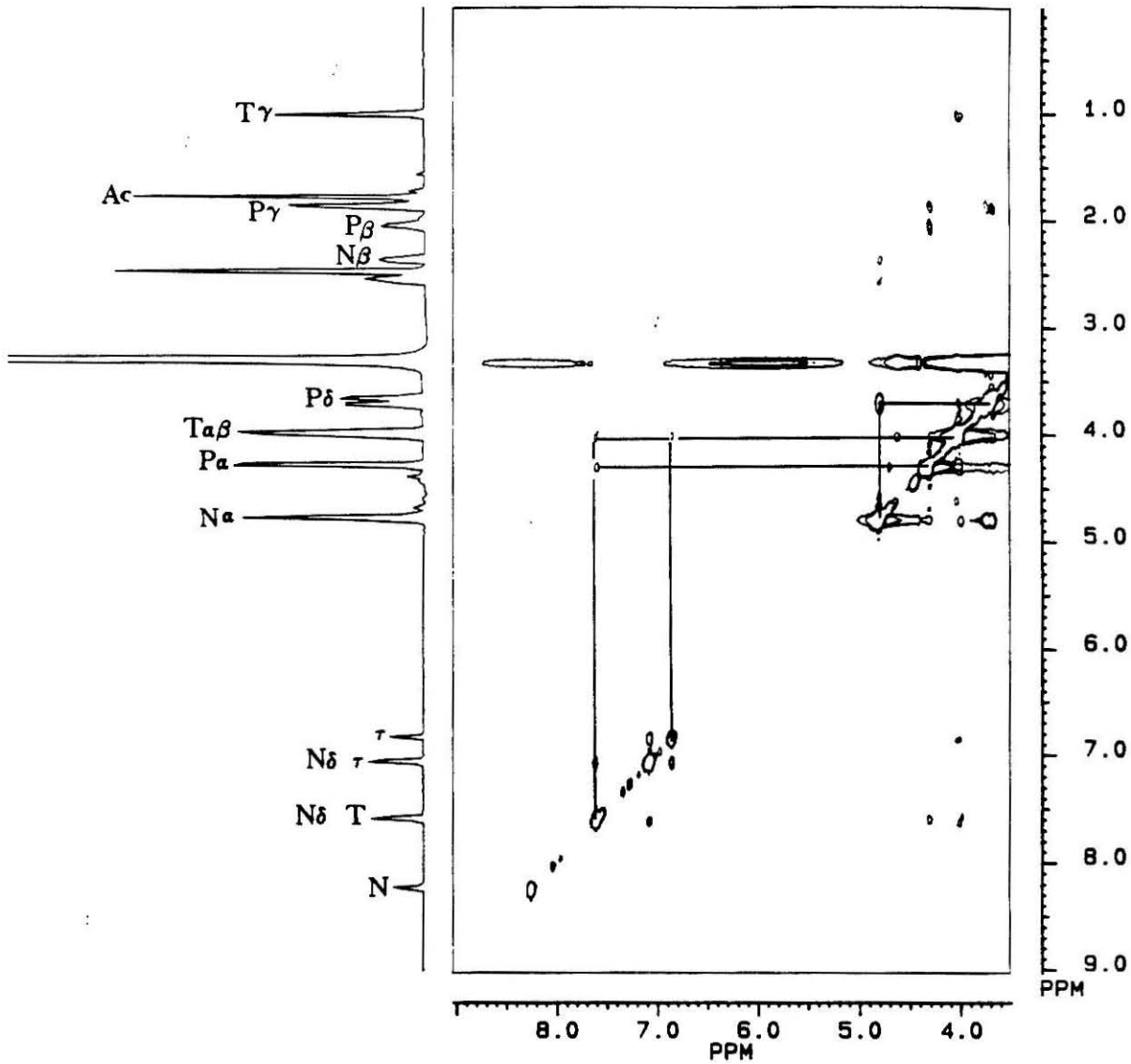


Figure 4-9 ROESY spectrum of Ac-Asn-Pro-Thr-NH_2.

primary amide carbonyl. If the former interaction were present, the $d_{\alpha N}$(Asn,Xaa) connectivity would be absent since this would place these protons on the opposite face of the molecule and hence out of connectivity range. However, with the latter interaction, the $d_{\alpha N}$(Asn,Xaa) would be expected to be strong. The other important connectivities observed, notably the strong intra-residue interactions in the threonine residue are compatible with a backbone-to-side chain H-bonding model. In addition, no other strong NOE connectivities are observed, which is in keeping with the above model.

The ROESY spectra of the acceptor peptides also lack interproton connectivities diagnostic of typical β-turn structures that generally include $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+2)$ interactions (**30**). Instead the substrate ROESY spectra reveal five strong sequential connectivities between residues along the polypeptide backbone indicative of an extended conformation. In addition, the coupling constants for the amide protons within the peptide backbone are also indicative of an extended backbone conformation. These results are consistent with the formation of an Asx-turn.

An examination of the spectroscopic data for the three peptides which are non-binders reveals that these peptides adopt conformations in solution which are distinct from those observed for substrates. In the (D)alanine containing peptide, while its variable temperature data also has a solvent shielded threonine amide, its ROESY crosspeak pattern is quite distinct from the pattern observed for the substrates. A model consistent with these changes would still involve hydrogen-bonding of the threonine NH to the asparagine side

chain amide carbonyl oxygen. However, since the methyl group of the (D)alanine residue crowds the α-carbonyl group of the adjacent asparagine, the amide linking the two residues must rotate by 180° to relieve steric congestion. In this model, the sequential backbone interaction, $d_{\alpha N}$(Asn,Xaa), is broken and a strong intraresidue connectivity, $d_{\alpha N}$(Asn,Asn), is established. The coupling constants observed for this peptide are consistent with this conformation.

The AIB-containing peptide cannot be compared directly with either of the examples described above because of the previously mentioned loss of backbone information due to the disubstituted α-carbon within the peptide backbone. However the variable temperature data reveals that two amide protons (ThrNH and terminus NH) have low temperature coefficients, and one amide (AIB-NH) has a high temperature coefficient analogous to the (D)alanine-containing peptide and distinct from the substrate peptides. The profound effect that α-amino isobutyric acid has on polypeptide conformational behavior has been reviewed by Venkataram and Balaram (**31**).

Finally, it has long been recognized that peptides containing proline as the central amino acid residue seldom act as acceptors in transferase assays. Not surprisingly, both the ROESY and variable temperature data show that proline allows the peptide to adopt a very different solution conformation. The variable temperature data for this peptide is quite distinct. One of the asparagine side chain amide protons has a low temperature coefficient (-2.1 ppb/K) implying that this proton is involved in a H-bonding interaction. From examining conformational models, either the proline or

threonine carbonyl groups could serve as potential H-bond acceptors. In either case, the key interaction between the threonine NH and the carbonyl of the asparagine side chain is broken. This may also explain why for this case the $^{3}J_{HN\alpha}$ value of threonine is somewhat higher (9.1 Hz).

CONCLUSIONS

The three substrates examined share common conformational features in solution. These features are consistent with the presence of an asparagine side chain to peptide backbone (ThrNH) interaction, or an Asx-turn. This interaction could serve to bring the threonine and asparagine side chains (crucial elements for glycosylation) into close proximity with each other, and hence provide a mechanism whereby the functionality of one could influence the reactivity of the other.

Three of the peptides examined (Bz-NPT-NHMe, Bz-N(D)AT-NHMe, and Bz-NBT-NHMe) show very poor or non-acceptor properties. Although it is possible that this diminished binding may be due to steric congestion at the enzyme binding site, our study suggests that in each case alterations in the backbone conformation are deleterious to formation of an Asx-turn as well as to the array of potential hydrogen-bonding interactions associated with this conformation.

1. Bause, E.; Legler, G. "The Role of the Hydroxy Amino Acid in the Triplet Sequence Asn-Xaa-Thr(Ser) for the N-Glycosylation Step During Glycoprotein Biosynthesis," *Biochem. J.* **1981**, *195*, 639-644.

2. Aubert, J.P.; Helbecque, N.; Loucheux-Lefebvre, M.H. "Carbohydrate-Peptide Linkage in Glycoproteins," *Arch. Biochem. Biophys.* **1981**, *208*, 20-29.

3. a) Bause, E. *Biochem. J.* "Structural Requirements of N-Glycosylation of Proteins," **1983**, *209*, 331-336. b) Bause, E., Hettkamp, H., & Legler, G. "Conformational Aspects of N-Glycosylation of Proteins," *Biochem. J.* **1982**, *203*, 761-768.

4. Aubry, A.; Abbadi, A.; Boussard, G.; Marraud, M. "ASX-turn Conformation in the Crystal Structures of Two Tripeptides Related to the Sequence Code for N-Glycosylation," *New Journal of Chemistry* **1987**, *11*, 739-744.

5. Abbadi, A.; Boussard, G.; Marraud, M.; Pichon-Pesme, V.; Aubry, A. In *Second Forum on Peptides* ; Aubry, A., Marraud, M., & Vitoux, B., Eds.; Colloque INSERM, Vol. 74, John Libbey Eurotext Ltd.,1989, pp. 375-378.

6. Beintema, J.J. "Do Asparagine-Linked Carbohydrate Chains in Glycoproteins Have a Preference for β-Bends?" *Bioscience Reports* **1986**, *6*, 709-714.

7. Beeley, J.G. "Peptide Chain Conformation and the Glycosylation of Glycoproteins," *Biochem. Biophys. Res. Commun.* **1977**,*76*, 1051-1055.

8. Chou, P.Y.; Fasman, G.D. "Prediction of Protein Conformation," *Biochemistry* **1974**, *13*, 222-245.

9. Geddes, A.J.; Parker, K.D.; Arkind, E.D.T.; Beighton, E. "Cross-β Conformation in Proteins," *J. Mol. Biol.* **1968**, *32*, 343-358.

10. Venkatachalam, C.M. "Stereochemical Criteria for Polypeptides and Proteins. V. Conformation of a System of Three Linked Peptide Units," *Biopolymers* **1968**,*6*, 1425-1436.

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

12. Van Rietschoten, J.; Granier, C.; Ronin, C.; Bouchilloux, S. "Peptides 1978, Proceedings of the 15th European Peptide Symposium"; Siemon,I.Z.; Kupryszewski, G., Eds.; Wroclaw University Press, Wroclaw 1979, p. 335.

13. Dyson, H.J.; Rance, M.; Houghten, R.A.; Lerner, R.A.; Wright, P.E. "Folding of Immunogenic Peptide Fragments of Proteins in Water Solution. I. Sequence requirements for the formation of a reverse turn," *J. Mol. Biol.* **1988**, *201*, 161-200.

14. Imperiali, B.; Fisher, S.L.; Moats, R.A.; Prins, T.J. "A Conformational Study of Peptides with the General Structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH_2: Spectroscopic Evidence for a Peptide with Significant β-turn Character in Water and in Dimethyl Sulfoxide," *J. Am. Chem. Soc.* **1992**, *118*, 3182-3188.

15. Wilmont, C.M.; Thornton, J.M. "Analysis and Prediction of the Different Types of β-Turn in Proteins," *J. Mol. Biol.* **1988**, *203*, 221-232.

16. Marshall, R.D. "The Nature and Metabolism of the Carbohydrate-Peptide Linkages of Glycoproteins," *Biochem. Soc. Symp.* **1974**, *40*, 17-26.

17. Pichon-Pesme, V.; Aubry, A.; Abbadi, A.; Boussard, G.; Marraud, M. "Crystal Molecular Structures of Two Tripeptides Related to the Sequence Coding for N-Glycosylation," *Int. J. Peptide Protein Res.* **1988**, *32*, 175-182.

18. Mononen, I.; Karjalainen, E. "Structural Comparison of Protein Sequences Around Potential N-Glycosylation Sites," *Biochim. Biophys. Acta* **1984**, *788*, 364-367.

19. Gavel, Y.; von Heijne, G. "Sequence Differences Between Glycosylated and Non-glycosylated Asn-Xaa-Thr/Ser Acceptor Sites: Implications for Protein Engineering," *Protein Engineering* **1990**, *3*, 433-442.

20. Abbadi, A.; Mcharfi, M.; Aubry, S.; Premilat, S.; Boussard, G.; Marraud, M. "Involvement of Side Functions in Peptide

Structures: The Asx Turn. Occurrence and Conformational Aspects," *J. Am. Chem. Soc.* **1991**, *113*, 2729-2735.

21. Richardson, J.S. "The anatomy and taxonomy of protein structure," *Adv. Protein Chem.* **1981**, *34*, 167-339.

22. Pardi, A.; Billeter, M.; Wuthrich, K. "Calibration of the Angular Dependence of the Amide Proton-Cα Proton Coupling Constants, $^3J_{HN\alpha}$, in a Globular Protein. Use of $^3J_{HN\alpha}$ for Identification of Helical Secondary Structure," *J. Mol. Biol.* **1984**, *180*, 741-751.

23. a) Bax, A. "Correction of Cross-peak Intensities in Two-Dimensional Spin-Locked NOE Spectroscopy for Offset and Hartmann-Hahn Effects," *J. Mag. Res.* **1988**,*77*, 134-147. b) Bax, A.; Davis, D.G. "Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy," *J. Mag. Res.* **1985**, *63*, 207-213. c) Bothner-By, A.A.; Stephens, R.L.; Lee, J.; Warren, C.D.; Jeanloz, R.W. "Structure Determination of a Tetrasaccharide: Transient Nuclear Overhauser Effects in the Rotating Frame," *J. Am. Chem. Soc.* **1984**,*106*, 811-813.

24. Ronin, C.; Aubert, J.-P. "Effect of Dimethylsulfoxide on Two Synthetic Asn-X-Thr Containing Substrates of the Oligosaccharyltransferase," *Biochem. Biophys. Res. Commun.* **1982**,*105*, 909-915.

25. Bodansky, M. Principles of Peptide Synthesis, 1984, Springer-Verlag, Berlin.

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

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

28. Giersch, L.M.; Thompson, K.F.; Rockwell, A.L.; Briggs, M.S. "Conformation-Function Relationships in Hydrophobic Peptides: Interior Turns and Signal Sequences," *Biopolymers* **1985**, *24*, 117-135.

29. Imperiali, B.; Fisher, S.L. unpublished results.

30. Wuthrich, K.; Billeter, M.; Braun, W. "Polypeptide Secondary Structure Determination by Nuclear Magnetic Resonance Observation of Short Proton-Proton Distances," *J. Mol. Biol.* **1984**, *180*, 715-740.

31. Venkataram, P.B.; Balaram, P. "The Stereochemistry of Peptides Containing α-Aminoisobutyric Acid," *CRC Crit. Rev. Biochem.* **1984**, *16*, 307-348.

Chapter 5
Constrained Analogs

With few exceptions short peptides (fewer than 10 residues) tend to be unordered in dilute aqueous solutions; that is, they adopt a large number of different conformations that are in dynamic equilibrium. However, small peptides can adopt defined conformations when they bind to receptors. Pinpointing the active conformation is relatively difficult because of the large number of possible conformations and due to the changing demographics of this conformational population as the solvent environment changes. Since the major conformations which exist in solution may not be relevant to those recognized by the enzyme at the active site, the solution conformations of a peptide deduced by experimental methods, although interesting in establishing the dynamic range of conformations available to a given peptide, are not necessarily relevant to the enzyme-bound conformation (**1**).

The design of conformationally constrained peptides is a powerful method for elucidating the biologically active conformation of a peptide which is recognized by an enzyme (**2**). Conformationally constrained peptides can be used to determine which (if any) of the many conformations observed in solution are responsible for binding. In addition, conformationally constrained peptides bind more tightly to their receptors, because their binding occurs with a smaller decrease in conformational entropy than the binding of an unconstrained analog (**3**).

Various types of conformationally restricted peptides can be produced by chemically modifying either the polypeptide backbone or the side chain functionalities of various residues (**4**). It is also possible to restrict conformational freedom through

short-range cyclizations (5). Cyclizations can be achieved by covalent bonding of backbone to backbone groups, side chain to side chain groups, or backbone to side chain groups (Figure 5-1).

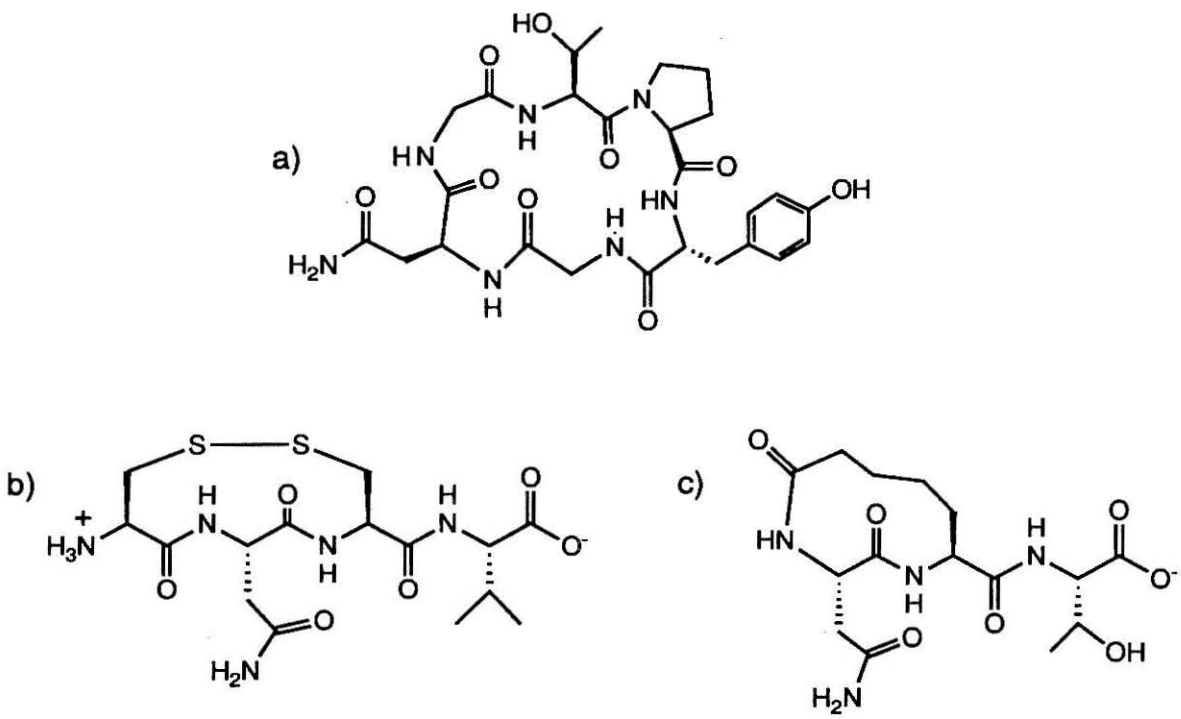


Figure 5-1 Constrained peptidyl analogs formed via cyclization of a) backbone to backbone groups, b) side chain to side chain groups, and c) backbone to side chain groups.

In chapter 4 it was demonstrated that substrates for glycosylation share common conformational features in solution. These features were consistent with the presence of an Asx-turn. In an effort to determine if the Asx-turn is the bioactive conformation of the polypeptide acceptor, three types of constrained analogs were designed and synthesized. These

analogs were examined for: greater conformational integrity, increased specificity, and increased reactivity.

The first series of constrained peptides examined were constructed via amide bond formation between the peptide amino and carboxy termini (cyclizations via backbone to backbone groups). Cyclic hexapeptides have proven useful in structural studies because they have fewer degrees of freedom than their linear analogs, yet they retain sufficient flexibility so that their conformations are not strictly dominated by steric interactions. Studies on cyclic hexapeptides which contain repeating tripeptide sequences have been shown to prefer a conformation which contains two *i+3* to *i* intramolecular hydrogen bonds such that the structure is essentially two fused β-turns (**6**) (Figure 5-2).

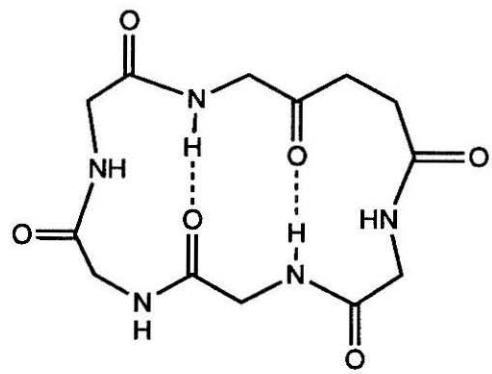


Figure 5-2 Predominant solution conformation of a cyclic hexapeptide.

Incorporation of proline residues into these cyclic peptides further reduces their flexibility. The presence of proline also reduces the number of possible β-turn conformers, since proline can occupy only the *i+1* or *i+2* positions. Flexibility is further restricted by incorporation of a D-aromatic residue following the proline.

Kopple extensively studied cyclo(L-Xaa-L-Pro-D-Phe)$_2$ peptides and concluded that these peptides consisted of two type II β-turns with the amide protons of the Xaa residue solvent shielded and transannularly hydrogen bonded. Crystal structure analysis of these peptides demonstrated that the 4→1 hydrogen bonds were bent out of the plane of the β-turn and were thus relatively weak. Gierasch *et al.* (7) improved on the two turn cyclic hexpeptide model by retaining only one Pro-D-Phe sequence and thereby increasing the flexibility at the turn formed opposite to it. Their peptide, cyclo(Pro-D-Tyr(OBzl)-Gly-Ile-Leu-Gln), was determined to exist in one preferred conformation which was composed of two β-turns with two strong transannular hydrogen bonds (Figure 5-3).

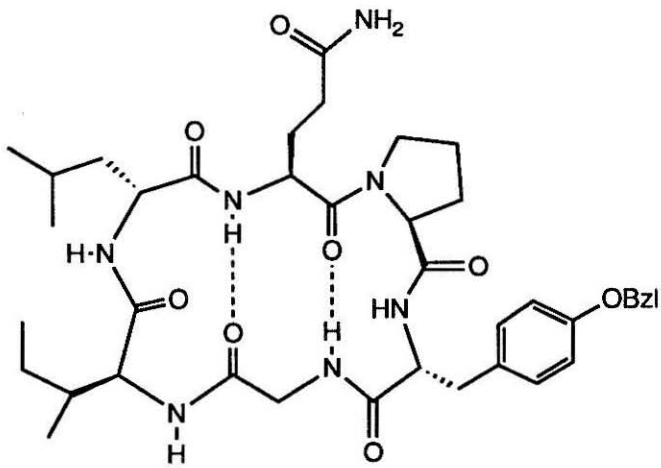


Figure 5-3 Predominant solution conformation of cyclo(Pro-D-Tyr(OBzl)-Gly-Ile-Leu-Gln).

Incorporation of an Asn-Xaa-Thr sequence into a cyclic hexapeptide which also contains a Pro-D-Phe sequence will fix the marker sequence within a type II β-turn. The marker sequence

can be positioned within the β-turn by frameshifting the asparagine into either the *i* or the *i+1* position (Figure 5-4).

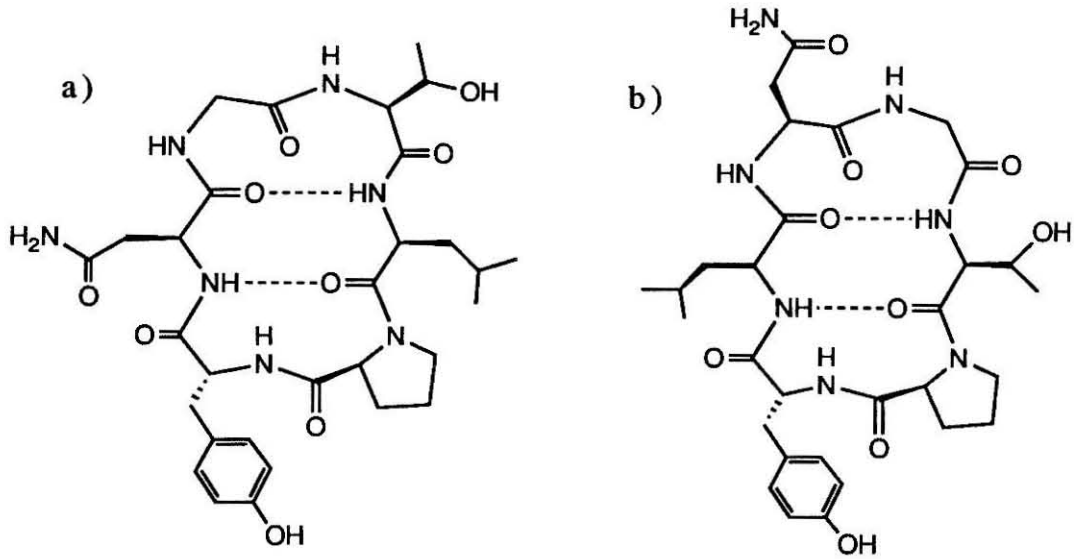


Figure 5-4 The two possible cyclic hexapeptides which incorporate both a Pro-D-Tyr and an Asn-Xaa-Thr sequences.

By examining the substrate behavior of these cyclic hexapeptides, the importance of β-turn formation in glycosylation can be conclusively established. Additionally, the positional preference of the asparagine within the marker sequence can be probed.

The second series of constrained analogs examined consists of peptides containing disulfide bonds (cyclization via side chain to side chain groups). Bause *et al.* (**8**) examined a series of peptides which contained two cysteine residues in various positions relative to the marker sequence (Table 5-1).

Table 5-1 Glycosyl-acceptor properties of linear and cyclic peptides. The essential asparagine and threonine residues are shown in bold.

	V_{max}/K_m	
peptide	linear	cyclic
CYNGTSVCG	9.6	0.7
CYNGTTCG	1.5	0.1
CYNCTSV	3.3	3.5
CNGTSVCG	1.5	0.6
CNGTSCG	1.2	0.07
CNGTCG	0.2	0.0
TNGTSV	1.0	n/a

In general the cyclic peptides show, with the exception of the CYNCTSV, a lower ability to function as glycosyl acceptors than their linear counterparts. Cyclization of the linear peptides by disulfide bond formation was accompanied by drastic increases in K_m as well as decreases in V_{max}. The exception, CYNCTSV, differs from the other peptides in that one cysteine residue is occupying the central position of the marker sequence. Disulfide formation, in this case, constrains only the asparagine residue; the C-terminal peptide segment including the essential threonine remains fully flexible (Figure 5-5).

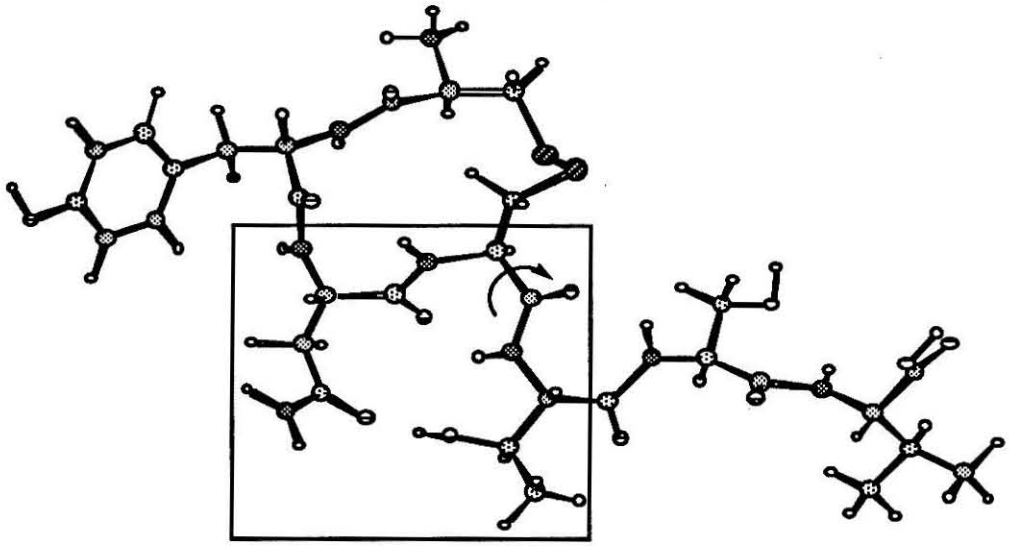


Figure 5-5 Energy minimized conformation of cyclo(CYNC)TSV. The marker sequence is boxed and the flexibility of the C-terminus is represented with an arrow.

Bause *et al.* (**8**) attempted to correlate their results with the ability to form β-turn structure. However, they did not attempt to examine the conformational properties of this series with any structural studies. Molecular modelling of the linear compounds shown in Table 5-1 does not indicate any particular predispostion for β-turn formation. Nor do their cyclic counterparts lead to any major disruptions of the distance separating the asparagine and threonine side chain functionalities. The modelling studies do indicate that disulfide formation is deleterious to Asx-turn formation in all cases except CYNCTSV. For CYNCTSV, an Asx-turn is still viable; however, other conformations which are characterized by intra-ring hydrogen bonds are preferred.

Examination of the solution conformations of these disulfide containing peptides should help to evaluate the importance of

Asx-turn formation in glycosylation. Unfortunately, the Asx-turn is not strictly constrained in these analogs.

To complement the disulfide analogs, a highly constrained Asx-turn containing analog was designed. This analog was to be cyclized via amide bond formation between the side chain carboxyl group of aminopimelic acid (Apm) and the backbone amide of asparagine (Figure 5-6).

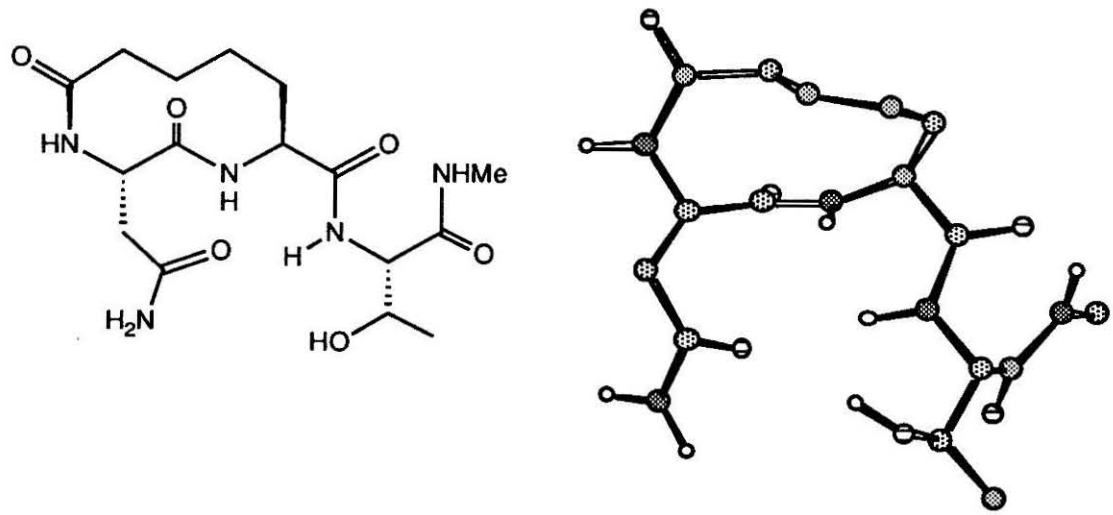


Figure 5-6 Representations of cyclo(AsnApm)ThrNMe.

The 10-membered macrocycle thus formed is much less flexible than the 14-membered disulfide ring in cyclo(CYNC)TSV. This macrocycle is also composed mainly of methylene units thus eliminating the intra-ring hydrogen bond interactions which destabilized the Asx-turn in cyclo(CYNC)TSV. Modelling studies of this compound predict that it will form a fairly rigid Asx-turn with a *cis* amide bond at the site of ring closure. In the minimized structure of cyclo(AsnApm)ThrNHMe, the side chain carbonyl oxygen of asparagine and the backbone amide of the threonine

are 2.6 Å apart. This conformation is further stabilized by hydrogen bonding interactions between the hydroxyl group of threonine and the side chain carbonyl oxygen of asparagine.

The three types of analogs designed were examined both for substrate behavior in glycosylation and for Asx-turn formation. The presence of the latter structure was evaluated through the use of both nuclear Overhauser effects and amide proton temperature coefficients.

EXPERIMENTAL METHODS

5.1 Synthesis of cyclic hexapeptides

Three linear peptides were synthesized on a Milligen 9050 automated peptide synthesizer on KA-resin: PYNGTL, PYLNGT, and TPFNNA. Peptide couplings were accomplished utilizing Fmoc-AA-OPfp esters as the active species and HOBT to prevent racemization. Side chain deprotection as well as cleavage of the peptide from the resin was accomplished with 9/0.5/0.3/0.2 TFA/thioanisole/ethanedithiol/anisole. Crude product was triturated from TFA with 2/1 Et_2O/Hexane. Pure product was obtained by preparative reverse phase HPLC (Ultrasphere ODS, 5 μm, 10 mm x 25 cm column).

Cyclic peptides were obtained by dissolving the linear peptides in DMF (0.02 M), neutralizing the N-termini with NMM, and activating the carboxyl terminus with 1 equivalent each of BOP and HOBT. Cyclization usually proceeded in 24 hrs. Cyclized product was readily distinquished from linear precursor by TLC in 1/1/1/1 HOAc/H_2O/EtOAc/n-BuOH. The reactions were worked

up by condensing the DMF, resuspending the residue in methanol, and triturating crude product with Et_2O. Pure product was obtained by preparative reverse phase HPLC (Ultrasphere ODS, 5 μm, 10 mm x 25 cm column).

PHYSICAL DATA- NMR are included at the end of the chapter.

cyclo(PY$_d$NGTL)

^{1}H NMR (d_6-DMSO) δ: 0.9 (m, 6H), 1.1 (d, 3H), 1.3 (m, 1H), 1.6 (m, 1H), 1.8 (m, 1H), 2.0 (m, 1H), 2.65 (m, 2H), 2.6 (m, 2H), 3.2 (m, 1H), 3.4 (m, 2H)4.0 (m, 1H), 4.25 (m, 3H) 4.6 (m, 2H), 6.65 (d, 2H), 6.7 (s, 1H), 7.05 (d, 2H), 7.3 (s, 1H), 7.6 (d, 1H), 8.2 (d, 1H), 8.9 (d, 1H)

^{13}C NMR (d_6-DMSO) δ: 172.01, 171.98, 171.04, 170.80, 170.75, 169.97, 169.08, 155.56, 129.74, 128.84, 114.81, 65.49, 60.77, 59.85, 55.33, 50.02, 48.30, 46.47, 44.67, 40.98, 36.24, 34.91, 27.88, 25.26, 24.02, 23.35, 20.85, 20.55

Mass Spectroscopy $[MH]^+$ 646

cyclo(PY$_d$LNGT)

^{1}H NMR (d_6-DMSO) δ: 0.9 (m, 6H), 1.1 (d, 3H), 1.45 (m, 1H), 1.55 (m, 1H), 1.6 (m, 1H), 2.0 (m, 1H), 2.65 (m, 2H), 2.7 (m, 2H), 3.4 (m, 1H), 3.7 (m, 1H), 3.8 (m, 1H0, 3.9 (m, 1H), 4.3 (m, 1H), 4.4 (m, 1H), 4.5 (m, 1H), 4.6 (m, 1H), 4.8 (m, 1H), 4.9 (m, 1H), 6.65 (d, 2H), 6.9 (s, 1H), 7.0 (d, 1H), 7.2 (t, 1H), 7.4 (s, 1H), 8.0 (d, 1H), 8.3 (m, 2H), 8.4 (d, 1H)

^{13}C NMR (d_6-DMSO) δ: 171.91, 171.68, 171.00, 170.35, 170.25, 168.68, 167.36, 129.9, 127.22, 114.7, 67.26, 61.19, 59.31, 55.39, 53.32, 51.60, 47.28, 41.47, 40.90, 38.74, 35.72, 31.72, 28.32, 25.02, 24.00, 22.82, 21.51, 19.03

Mass Spectroscopy $[MH]^+$ 646

cyclo(PF$_d$NNAT)

^{1}H NMR (d_6-DMSO/$CDCl_3$) δ: 1.04 (d, 3H, J=6.25 Hz), 1.3 (d, 3H, J=7.8 Hz), 1.53 (m, 1H), 1.86 (m, 2H), 1.94 (m, 1H), 2.54 (m, 1H0, 2.8 (m, 3H), 3.41 (d, 1H, J=11.1 Hz), 3.50 (m, 1H), 3.63 (m, 1H), 4.07 (m, 1H), 4.2 (m, 1H), 4.38 (m, 2H), 4.51 (m, 2H), 4.63 (bs, 1H), 4.67 (m, 1H), 6.73 (s, 1H), 6.79 (s, 1H), 7.11 (m, 2H), 7.17 (m, 3H), 7.36 (s, 1H), 7.90 (D, 1H, J=8.7 Hz), 8.39 (d, 1H, J=6.9 Hz), 8.61 (d, 1H, J=8.2), 8.64 (d, 1H, J=9.5 Hz), 8.72 (d, 1H, J=9.4 Hz)

^{13}C NMR (d_6-DMSO/$CDCl_3$) δ: 173.1, 171.5, 171.4, 170.2, 170.1, 168.8, 168.7, 138.7, 129.0, 128.8, 127.7, 125.8, 66.46, 61.78, 55.6, 53.8, 52.5, 50.1, 49.7, 47.0, 40.1, 36.4, 36.2, 35.7, 29.0, 28.6, 28.7, 27.8, 25.3, 19.5, 19.0, 18.7

Mass Spectroscopy $[MH]^+$ 645

5.2 Synthesis of disulfide containing peptides

CYNCTSV and CANCTSA were synthesized on the Milligen 9050 using FMOC-AA-KA resin and BOP/HOBT couplings. Thiols were protected with trityl groups. The linear peptides were triturated from MeOH with Et_2O until pure by HPLC. Disulfide formation was promoted by suspending the peptides in water (0.01 M) and adjusting the pH to 8.5 with NH_4OH. The solutions were stirred overnight with aeration. Disulfide formation was monitored using Ellman's test for free thiol (**9**). Cyclized products were isolated by preparative reverse phase HPLC (Ultrasphere ODS, 5 μm, 10 mm x 25 cm column) using a linear gradient (0-40% acetonitrile/water over 20 minutes). Disulfide containing compounds were retained 13 minutes compared to their reduced forms which were retained 22 minutes.

cyclo(CYNC)TSV

^{1}H NMR (d_6-DMSO) δ: 0.81 (q, 6H, J=4.9 Hz), 1.06 (d, 3H, J=6.4 Hz), 2.06 (m, 1H), 2.39 (m, 1H), 2.73 (m, 1H), 2.9 (m, 3H), 3.05

(m, 1H), 3.10 (m, 1H), 3.2 (m, 1H), 3.72 (m, 2H), 4.08 (m, 1H), 4.17 (m, 2H), 4.27 (m, 2H), 4.39 (t, 1H, J=5.4 Hz), 4.55 (m, 1H), 4.6 (m, 1H), 4.83 (m, 1H), 5.05 (m, 1H), 6.66 (d, 2H, J=8.9 Hz), 6.86 (s, 1H), 6.98 (d, 2H, J=8.4 Hz), 7.42 (s, 1H), 7.57 (d, 1H, J=8.1 Hz), 7.83 (d, 2H, J=8.4 Hz), 7.87 (d, 2H, J=8.0 Hz), 7.95 (d, 2H, J=8.3 Hz), 8.7 (d, 2H, J=7.9 Hz), 9.2 (bs, 1H)

^{13}C NMR (d_6-DMSO) δ: 171.9, 170.5, 170.2, 169.8, 169.2, 169.0, 156.0, 129.9, 126.8, 115.1, 66.5, 61.5, 57.9, 57.3, 54.8, 52.1, 50.7, 49.8, 36.8, 35.4, 29.9, 18.9, 17.9

Mass Spectroscopy $[MH]^+$ 787 (exact mass)

cyclo(CANC)TSA

^{1}H NMR (d_6-DMSO) δ: 1.03 (d, 3H), 1.24 (m, 6H), 2.65 (m, 2H), 2.71 (m, 1H), 3.07 (m, 3H), 3.60 (d, 2H), 3.97 (s, 2H), 4.17 (t, 1H), 4.29 (m, 3H), 4.42 (m, 1H), 4.63 (m, 1H), 5.1 (d, 1H), 6.85 (s, 1H), 7.39 (s, 1H), 7.44 (d, 1H), 7.85 (d, 1H), 7.96 (d, 1H), 8.05 (d, 1H), 8.23 (d, 1H), 8.76 (d, 1H)

^{13}C NMR (d_6-DMSO) δ:173.76, 171.94, 171.59, 170.15, 169.50, 169.19, 169.13, 66.55, 61.45, 57.81, 54.81, 52.79, 50.96, 49.69, 49.50, 47.57, 37.81, 37.28, 35.32, 18.91, 17.56, 17.14

Mass Spectroscopy $[MH]^+$ 667 (exact mass)

5.3 Synthesis of the Pimelic macrocycle

Aminopimelic acid is available from Sigma as the D,L racemate. The L-enantiomer can be obtained by acylation followed by enantioselective hydrolysis with porcine kidney acylase (**10**).

L-Apm-OBn

Concentrated H_2SO_4 (0.8 mL) was mixed with 18 ml Et_2O. BnOH (12 ml, 15 mmol) was added dropwise to this solution and the Et_2O was blown off with a gentle stream of N_2. L-Pimelic Acid (2 g, 12.4 mmol) was added to the solution and stirred at 25°C for 6 hrs. The solution was diluted with 10 mL EtOH and neutralized with pyridine. Crude precipitate was collected, redissolved in boiling water, and allowed to cool overnight in the refrigerator. 1.8 g pure product was obtained (Yield 65%).
TLC (4/1/1 HAc/H_2O/*n*-BnOH) Rf=0.2

Boc-Apm-OBn

Apm-OBn (1.4 g, 5.58 mmol) was dissolved in 10 mL 50% aq. dioxane. Et_3N (1.2 mL, 8.37 mmol) was added and stirred, followed by the addition of 1.5 g BOC-ON (6.14 mmol). The solution was stirred for 3 hrs at RT after which time the dioxane was condensed off. The aqueous layer was acidified to pH 7.0 and extracted with EtOAc. The organic phase was dried over $NaSO_4$, filtered, and condensed (Yield 3.5 g). Crude product was flash chromatographed (10/1 $CHCl_3$/MeOH Rf=0.25) and yielded 1.63 g pure product (Yield 81%).

^1H NMR ($CDCl_3$) δ: 1.35 (s, 9H), 1.37 (m, 2H), 1.66 (m, 3H), 1.87 (m, 1H), 2.40 (t, 2H), 4.38 (m, 1H), 5.12 (s, 2H), 7.42 (m, 5H)

Boc-Apm(OBn)-Thr-OMe

Et_3N (180 mL, 0.86 mmol) was added to 0.179 g HCl.Thr-OMe (0.66 mmol) in 2 mL THF and the solution was stirred. HOBT

(0.098 g, 0.73 mmol) and 0.194 g DCC (0.73 mmol) were added to 0.266 g Boc-Apm(OBn) in 3 mL THF. This solution was added dropwise to the Thr-OMe solution. Reaction was stirred overnight. DCU was filtered off and rinsed with 2 mL THF. THF was concentrated and the residue was purified via flash chromatography (8/1 Et_2O/Hexane). Pure product was obtained (0.22 g, 62%).

^1H NMR ($CDCl_3$) δ: 1.2 (d, 3H), 1.43 (s, 9H), 1.65 (m, 3H), 1.86 (m, 1H), 2.37 (t, 2H), 3.75 (s, 3H), 4.11 (m, 1H), 4.34 (m, 1H), 4.56 (m, 1H), 5.05 (d, 1H), 5.11 (s, 2H), 6.80 (d, 1H), 6.88 (d, 1H), 7.34 (m, 5H)

Boc-Asn-Apm(OBn)-Thr-OMe

Boc-Apm-Thr-OMe (0.383 g, 0.798 mmol) was dissolved in 4 mL 50% TFA/Me_2Cl_2. The reaction was stirred for 20 minutes after which the TFA was blown off with a gentle stream of N_2. Me_2Cl_2 was condensed and the residue was dried with 5% toluene/MeOH until traces of TFA were gone. Residue was redissolved in 2 mL DMF and 221 μL Et_3N (1.59 mmol). Boc-Asn (0.177 g, 0.76 mmol), 0.103 g HOBT , and 0.338 g BOP were dissolved in 2 mL DMF and added dropwise to the TFA salt solution. Reaction was stirred for 10 hrs after which DMF was condensed off. Residue was flash chromatographed (20/1 $CHCl_3$/MeOH Rf = 0.45) and yielded 0.43 g pure product (yield 92%).

The synthesis was also performed with D,L-pimelic acid. The diastereomers (DS) could be separated at this stage by flash chromatography (20/1 $CHCl_3$/MeOH; Rf DS-1=0.5, Rf DS-2=0.45).

DS-2 was determined to be the L,L,L tripeptide by comparison against the L,L,L tripeptide obtained above.

^{1}H NMR (d_6-DMSO) δ: 1.05 (d, 3H), 1.17 (t, 2H), 1.37 (s, 10H), 1.54 (m, 3H), 1.68 (m, 1H), 2.3 (t, 2H), 2.45 (m, 2H), 3.6 (s, 3H), 4.1 (m, 1H), 4.24 (m, 2H), 4.38 (m, 1H), 5.07 (s, 2H), 6.9 (s, 1H), 7.0 (d, 2H), 7.30 (s, 1H), 7.32 (m, 5H), 7.74 (d, 1H), 8.03 (d, 1H).10 (s, 2H), 6.13 (s, 1H), 7.10 (s, 1H), 7.34 (m, 5H), 7.48 (d, 1H), 7.71 (d, 1H)

^{13}C NMR (d_6-DMSO) δ: 172.6, 171.9, 171.6, 171.1, 170.8, 155.1, 136.2, 128.4, 127.8, 126.8, 124.2, 118.9, 78.2, 66.2, 66.1, 57.8, 51.8, 51.6, 51.2, 45.7, 37.0, 36.4, 33.4, 31.2, 28.0, 24.2, 24.0, 19.9, 8.5

Boc-Asn-Apm-Thr-OMe

Boc-Asn-Apm(OBn)-Thr-OMe (0.0795 g, 0.135 mmol) was dissolved in 2 mL of MeOH and flushed with N_2. Palladium (10 mg, 5% absorbed on carbon) was added and solution was again flushed with N_2. Solution was hydrogenated at atmospheric pressure with vigorous stirring for 35 min. Palladium was filtered off through celite and rinsed with 50% aq. MeOH. 52 mgs pure product was obtained (Yield 78%). No further product was obtained after rigorously rinsing both flask and celite. Subsequent reactions ranged in yields from 70-92%.

^{1}H NMR (d_6-DMSO) δ: 1.04 (d, 3H), 1.27 (m, 2H), 1.46 (s, 9H), 1.48 (m, 3H), 1.65 (m, 1H), 2.17 (t, 2H), 2.34 (m, 2H), 3.6 (s, 3H), 4.10 (m, 1H), 4.23 (m, 2H), 4.38 (m, 1H), 4.9 (s, 1H), 6.87 (s, 1H), 6.99 (d, 1H), 7.27 (s, 1H), 7.72 (d, 1H), 8.04 (d, 1H)

^{13}C NMR (d_6-DMSO) δ:174.3, 172.0, 171.6, 171.1, 170.8, 155.1, 78.2, 66.2, 57.8, 51.9, 51.7, 51.2, 36.9, 33.5, 31.9, 28.1, 24.4, 24.2, 19.9

Boc-Asn-Apm-(Asn-Apm(OBn)-Thr-OMe)-Thr-OMe

Boc-Asn-Apm(OBn)-Thr-OMe (0.033 g, 0.05 mmol) dissolved in 1 mL 50% TFA/Me_2Cl_2. The reaction was stirred for 20 minutes after which the TFA was blown off with a gentle stream of N_2. Me_2Cl_2 was condensed and the residue was dried with 5% toluene/MeOH until traces of TFA were gone. Residue was then dissolved in 0.25 mL DMF and 5.6 μL Et_3N (0.055 mmol). Boc-Asn-Apm(OH)-Thr-OMe (0.025 g, 0.05 mmol) and 0.024 g BOP were dissolved in 0.25 mL DMF and added dropwise to the TFA salt solution. Reaction was stirred for 10 hrs after which DMF was condensed off. Residue was flash chromatographed (20/1 $CHCl_3$/MeOH Rf 0.23) (yield 52%).

^1H NMR (d_6-DMSO) δ: 0.8 (d, 6H), 1.0 (t, 5H), 1.2 (s, 11H), 1.4 (m, 5H), 1.9 (t, 2H), 2.1 (t, 2H), 2.4 (m, 4H), 3.4 (d, 6H), 4.0 (m, 2H), 4.1 (m, 1H), 4.25 (m, 2H), 4.8 (s, 2H), 7.1 (m, 5H).

Pimelic dimer Boc-Asn-Apm-(Asn-Apm(OBn)-Thr-OMe)-Thr-OMe (0.040 g, 0.055 mmol) was dissolved in 1 mL of MeOH and flushed with N_2. Palladium (10 mg, 5% absorbed on carbon) was added and solution was again flushed with N_2. Solution was hydrogenated at atmospheric pressure with vigorous stirring for 35 min. The palladium was filtered off through celite and rinsed with 50% aq. MeOH. The residue was dissolved in 1 mL 50% TFA/Me_2Cl_2 and stirred for 20 minutes. TFA was blown off with a gentle stream of N_2. Me_2Cl_2 was condensed and the residue was dried with 5% toluene/MeOH until traces of TFA were gone.

^{1}H NMR (d_6-DMSO) δ: 1.01 (m, 6H), 1.3 (m, 4H), 1.5 (m , 6H), 1.6 (m, 2H). 2.1 (t, 2H), 2.2 (t, 2H), 2.4 (m, 4H), 3.6 (d, 6H), 4.0 (m, 1H), 4.1 (m, 2H), 4.2 (m, 2H), 4.3 (m, 1H), 4.4 (m, 1H), 4.5 (m, 1H), 4.6 (m, 1H), 5.0 (s, 1H), 6.7 (s, 1H), 7.2 (s, 1H), 7.4 (s, 1H), 7.6 (s, 1H), 7.8 (d, 1H), 8.0 (d, 1H), 8.1 (d, 1H), 8.2 (d, 1H), 8.4 (d, 1H)

^{13}C NMR (d_6-DMSO) δ: 174.3, 172.1, 171.9, 171.8, 171.6, 170.9, 170.8, 66.1, 57.9, 52.3, 52.0, 51.7, 49.3, 41.4, 41.0, 40.8, 40.0, 37.0, 36.1, 35.0, 33.6, 31.7, 31.5, 24.9, 24.7, 24.5, 21.6, 20.0, 19.9

Pimelic macrocycle

0.011 g H_3N(+)-Asn-Apm-(Asn-Apm-Thr-OMe)-Thr-OMe was dissolved in 450 μL DMF (0.03M) and neutralized with 2 μL Et_3N. 12.4 mg BOP was added and stirred for 1 hr after which time 1.5 eq Et_3N (3 μL, 0.021 mmol) were added. 200 μL DMF was added after 24 hrs to resolubilize the reaction mixture. The reaction was allowed to stir a further 48 hrs. DMF was condensed off and the residue was triturated from methanol with Et_2O. The crude product was purified by preparative reverse phase HPLC (Ultrasphere ODS, 5 μm, 10 mm x 25 cm column) and yielded 6 mgs.

^{1}H NMR (d_6-DMSO) δ:1.0 (m, 3H), 1.1 (m, 1H), 1.2 (m , 2H), 1.4 (m, 3H). 2.1 (t, 2H), 2.4 (m, 2H), 3.4 (d, 3H), 4.1 (m, 1H), 4.2 (m, 1H), 4.3 (m, 1H), 4.6 (m, 1H), 5.0 (s, 1H), 6.7 (s, 1H), 7.4 (s, 1H), 7.6 (d, 1H), 7.8 (d, 1H), 8.0 (d, 1H)

^{13}C NMR (d_6-DMSO) δ: 172.1, 171.8, 171.7, 170.7, 170.6, 66.1, 57.9, 51.8, 51.5, 49.3, 40.0, 36.1, 34.7, 31.6, 28.9, 28.7, 28.4, 24.4, 24.2, 19.8

5.4 NMR Spectroscopy on Constrained Analogs

NMR experiments were performed at ambient temperature (298 K) in d_6-dimethylsulfoxide with the exception of cyclo(PF$_d$NNAT) which were obtained in 50% $CDCl_3/d_6$-DMSO. 2-D ROESY spectra were obtained as described in chapter 4.2.

5.5 Oligosaccharyltransferase Assay

The oligosaccharyltransferase assay was performed as described in chapter 2.3.1. The percentage of dimethylsulfoxide was raised to 15% in the assays of the pimelic macrocycle in order to solubilize this analog.

RESULTS AND DISCUSSION

The cyclic hexapeptides, cyclo(PY$_d$NGTL) and cyclo(PY$_d$LNGT), were designed to fix the glycosylation marker sequence within the *i* and *i+1* positions of a type I β-turn. The Pro-(D)-Tyr sequence, which strongly favors a type II β-turn, was used to encourage the adoption of a β-turn by the remaining four residues. The solution conformations of these peptides were analyzed by ^{1}H-NMR. The proton resonances were assigned via a 2-D COSY experiment. Table 5-2 illustrates the temperature coefficents for each proton in the amide region of cyclo(PY$_d$NGTL) and cyclo(PY$_d$LNGT). As discussed in chapter 4, coefficients ranging from -4 to -7 ppb/K are typical of solvent exposed protons, whereas those in the 0 to -3 ppb/K range are diagnostic of solvent shielding.

Table 5-2 Temperature dependence of the amide proton chemical shifts ($-\Delta\delta/\Delta T$) for cyclo(PY_dNGTL) and cyclo(PY_dLNGT) in DMSO.

peptide	Tyr	Asn	Gly	Thr	Leu
PYNGTL	6.94	2.15	8.29	6.94	2.64
PYLNGT	5.55	5.35	7.96	6.82	3.75

The ROESY spectra of each peptide are shown in Figures 5-7 and 5-8 respectively.

The following observations were made with repect to cyclo(PY_dNGTL): 1) small temperature dependencies are seen for the Asn and Leu NH resonances, 2) the Gly Hα-NH coupling constants are extremely different (9.3 vs. 4.6 Hz), 3) a set of NH(*i*) to NH(*i*+*1*) NOE connectivities are seen in the ROESY experiment (Leu-NH to Thr-NH and Asn-NH to Tyr-NH), one at each end of the molecule, and 4) Pro-αH interacts with Tyr-NH. Figure 5-7 summarizes this information.

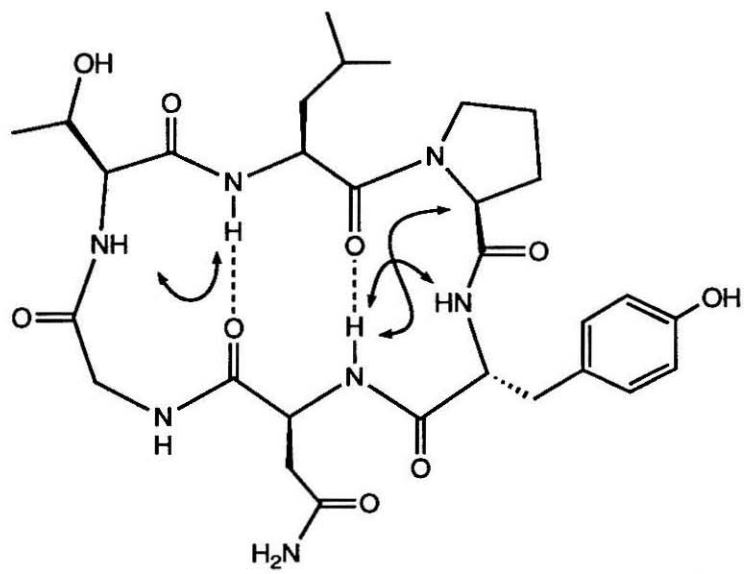


Figure 5-7 Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PY_dNGTL).

Cyclo(PY_dNGTL) thus appeared to adopt both turns as predicted. Asparagine was fixed in the i position of the turn.

Curiously enough, cyclo(PY_dLNGT) did not form the same well defined turns. The only amide that exhibited a low temperature dependency was Leu-NH. The Gly Hα-NH coupling constants were equal indicating that these α protons were magnetically equivalent and therefore not constrained. Only one NH(i) to NH(i+1) NOE connectivity was seen (Leu-NH to Tyr-NH) indicating that only the turn containing the Pro-(D)-Tyr pair was fixed. The Pro αH to Tyr-NH interaction was also visible. The lack of formation of the second β-turn was further corroborated by the existence of a minor conformation in both the ^{13}C and ^{1}H NMR spectra (about 1/3 in DMSO). Figure 5-8 summarizes the data for cyclo(PY_dLNGT).

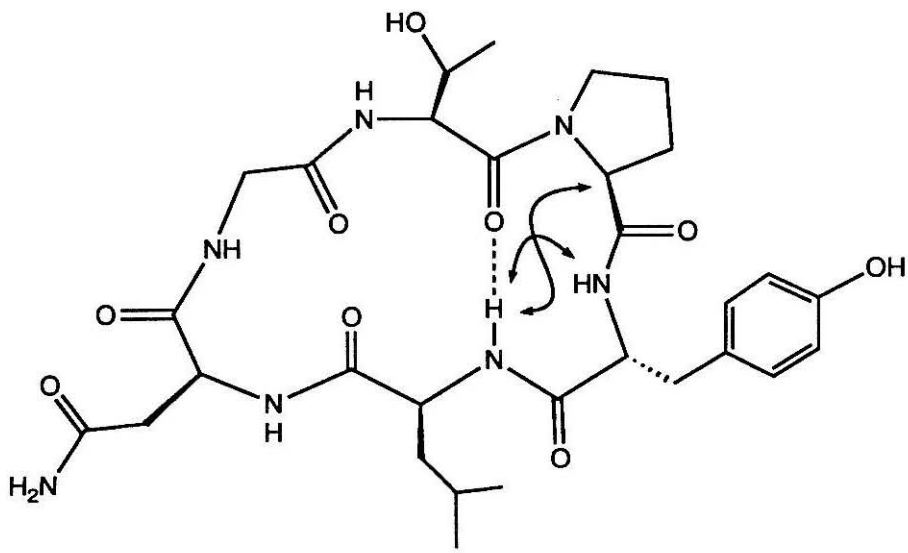


Figure 5-8 Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PY_dNGTL).

There are a number of possibile explanations as to why cyclo(PY$_d$NGTL) doesn't form the second turn. CPK modelling of this peptide indicates that most of its functionality is on the same face of the macrocycle, unlike cyclo(PY$_d$LNGT). Steric congestion may cause the backbone to twist into a more accomodating conformation. Alternatively, asparagine's high positional preference for the *i* position within a β-turn may cause the backbone to contort into an alternate loop structure.

In order to study the effect of asparagine within the *i+1* position of the turn, a third cyclic hexapeptide, cyclo(PF$_d$NNAT) was designed and synthesized. This peptide substituted a second asparagine residue for the leucine. Asparagine is smaller than leucine and thus it would cause less steric crowding, and it should also encourage a β-turn in which the asparagine of the marker sequence is fixed in the *i+1* position.

The ^{1}H- NMR analyses of this compound indicated that the second β-turn formed. The temperature coefficients and diagnostic ROESY connectivities are shown in Figure 5-9.

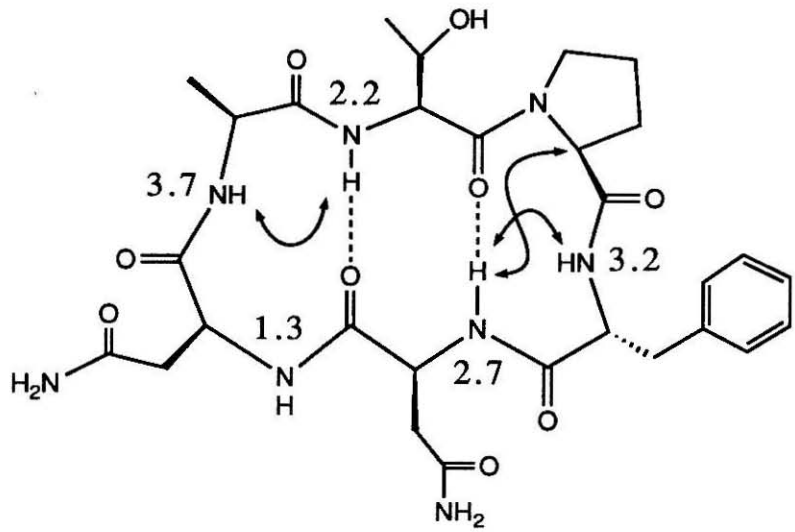


Figure 5-9 Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(PF_dNNAT). Temperature coefficients for the amide protons are shown.

The glycosyl acceptor properties for these three peptides were determined using standard assay conditions. The peptides were examined using 50,000 dpm of Dol-P-P-GlcNAc-[^{3}H]GlcNAc. The results are shown in Table 5-3. The only peptide which exhibited any activity in the transferase assay was cyclo(PY_dNGTL). The relative activity of this peptide was 100 fold lower than the known substrate Bz-NLT-NHMe. Cyclo(PY_dLNGT) and cyclo(PF_dNNAT) had no detectable activity at concentrations up to 10 mM.

Table 5-3 Enzyme assay results for the cyclic hexapeptides. Bz-NLT-NHMe is included as a standard.

peptide	K_m (mM)	V_{max} (dpm min^{-1})	V_{max}/K_m
cyclo(PY_dNGTL)	5-10	~250	~40
cyclo(PY_dNGTL)	-	-	-
cyclo(PY_dNGTL)	-	-	-
Bz-NLT-NHMe	0.55	2702	4914

The amount of Dol-P-P-GlcNAc-[^{3}H]GlcNAc was increased to 200,000 dpm to boost the detection limits of the assay. Figure 5-10 illustrates the results of this experiment.

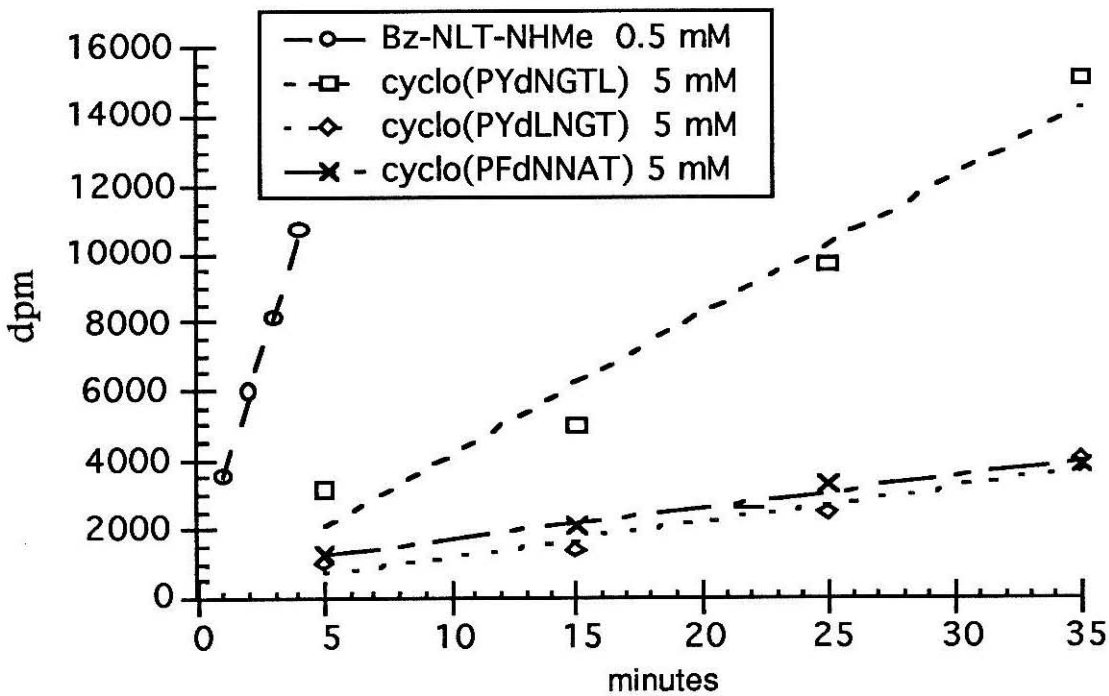


Figure 5-10 Glycosyl acceptor abilities of 0.5 mM Bz-NLT-NHMe, 5 mM cyclo(PY_dNGTL), 5 mM cyclo(PY_dLNGT), and 5 mM cyclo(PF_dNNAT).

It is clear that all three cyclic hexapeptides were extremely poor acceptors in the glycosylation assay. Since these peptides were

fixed into β-turns, it is highly unlikely that this conformation is the active recognition element for glycosylation.

Cyclo(CYNCTSV) has previously been reported to have similar acceptor properties to its reduced form, linear(CYNCTSV) (**8**). Thus, although disulfide formation decreases the conformational space accessible to this peptide, the active conformation must still be populated. The solution conformations of cyclo(CANCTSA) and cyclo(CYNCTSV) were examined by ^{1}H-NMR to determine if a major conformation was evident.

Table 5-4 illustrates the temperature coefficents for each proton in the amide region of cyclo(CANCTSA) and cyclo(CYNCTSV).

Table 5-4 Temperature dependence of the amide proton chemical shifts ($-\Delta\delta/\Delta T$) for cyclo(CANCTSA) and cyclo(CYNCTSV) in DMSO.

peptide	Xaa-NH	Asn-NH	Cys-NH	Thr-NH	Ser-NH	Xaa-NH
cyclo(CANCTSA)	2.4	5.2	1.6	4.6	4.6	5.2
cyclo(CYNCTSV)	2.1	5.1	1.6	4.0	4.3	5.3

The portion of the ROESY spectrum of cyclo(CYNCTSV) is shown in Figure 5-11. The ROESY spectra for cyclo(CANCTSA) and cyclo(CYNCTSV) are equivalent. There are two diagnostic connectivities which can seen in the spectra: Ala-NH to Asn-NH and Ala-NH to Cys-NH. These connectivities suggest that the macrocycle formed by the disulfide bond is extremely rigid and is characterized by intra-ring hydrogen bonding interactions. The portion of the polypeptide outside of this ring is extremely flexible and lacks significant structure.

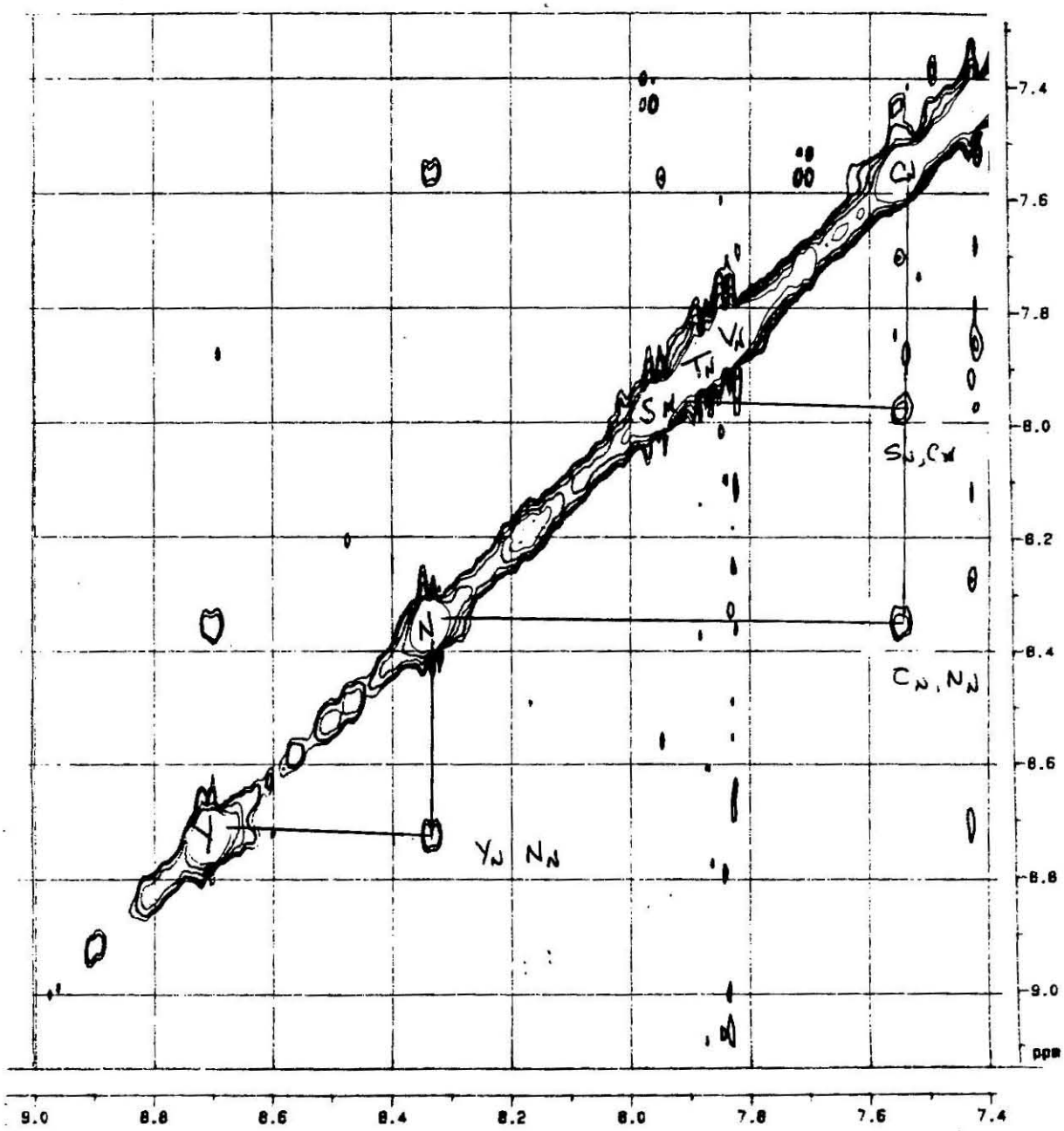


Figure 5-11 ROESY spectrum of cyclo(CYNCTSV)

Figure 5-12 illustrates the NOE connectivities observed and the hydrogen bonding interactions suggested by this data.

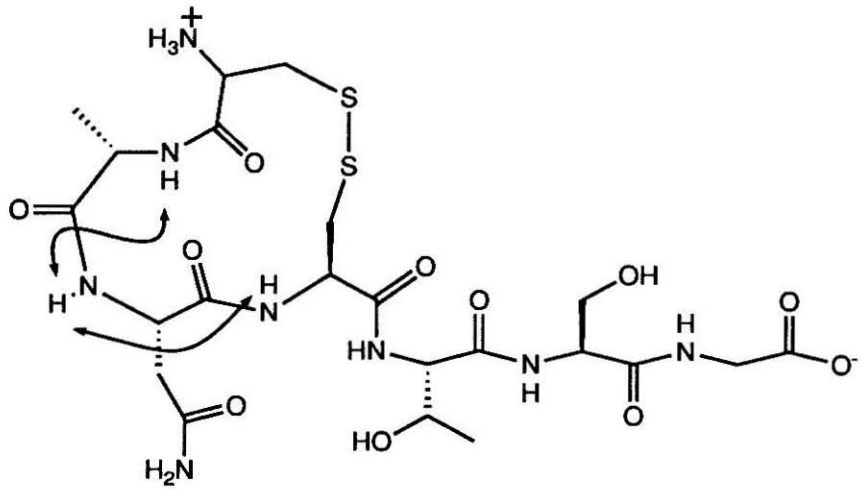


Figure 5-12 Graphical representation of connectivity and hydrogen bonding information derived from spectroscopic studies of cyclo(CANCTSA).

The glycosyl acceptor properties for cyclo(CANCTSA) and cyclo(CYNCTSV) were determined using standard assay conditions (Table 5-5).

Table 5-5 Enzyme assay results for the cyclic hexapeptides.

peptide	K_m (mM)	V_{max} (dpm min^{-1})	V_{max}/K_m
Bz-NLT-NHMe	0.47	4208	8953
cyclo(CANCTSA)	2.06	6369	3091
cyclo(CYNCTSV)	0.33	6402	19400

The enhanced binding of cyclo(CYNCTSV) compared to cyclo(CANCTSA) demonstrates the effect of hydrophobicity on binding (refer to chapter 2). Since these peptides have identical solution conformations as determined by ^{1}H-NMR, the binding efficiency must solely be due to hydrophobic effects.

It was previously demonstrated that the linear and cyclized forms of CYNCTSV have identical transferase activity (**8**). Although the cyclized form has less degrees of conformational freedom than the linear form, the marker sequence is not rigidly fixed into any particular conformation. Thus the marker sequence is neither fixed into the active conformation nor is it prevented from adopting it. This is consistent with the enzymatic data which suggests that the active conformation is equally attainable to both the cyclized and linear forms.

Since the tripeptide data in chapter 4 suggested that the active conformation was an Asx-turn, a peptidyl acceptor that was locked into this conformation was designed. The disulfide peptides discussed above served as a template. Cyclo(Asn-Apm)-Thr-NHMe is also cyclized from the side chain of its central residue, Apm; however, the macrocycle is closed by an amide bond between the asparagyl amino terminus and the side chain carboxylate of Apm. The macrocycle contains only 10 heteroatoms and thus should be less flexible than the 14 member disulfide ring.

The final ring closure step in the synthesis of cyclo(Asn-Apm)-Thr-NHMe proved challenging. This was not unexpected due to the highly constrained nature of the macrocycle. Cyclization of less sterically hindered ring systems have been reported to proceed with low yields (**11**). The sole product isolated from several attempted ring closures was the symmetrical dimer shown in Figure 5-13.

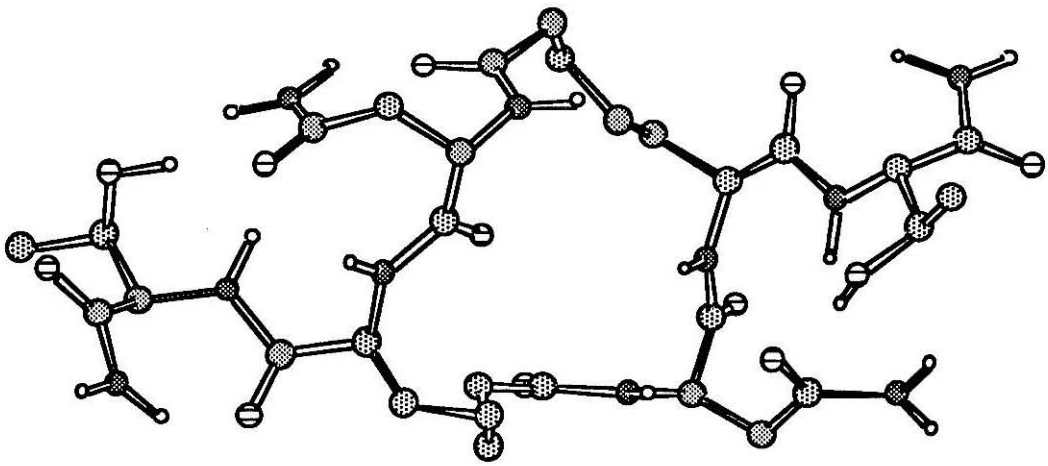


Figure 5-13 Macromodel representations of the dimeric pimelic macrocycle.

The ^{1}H-NMR spectra of this compound contains only one resonance for each pair of protons. The dimer appears to be rigidly constrained as indicated by: 1) the relative insolubility of the compound in most solvents, 2) the stability of the 1-D solution conformation at elevated temperatures (at least to 330 K), and the observation of only one population of NOE connectivities (at both 273 K and 330 K). The isolated compound was determined to be dimeric by mass spectroscopy.

The ROESY spectrum of this dimeric material is shown in Figure 5-14.

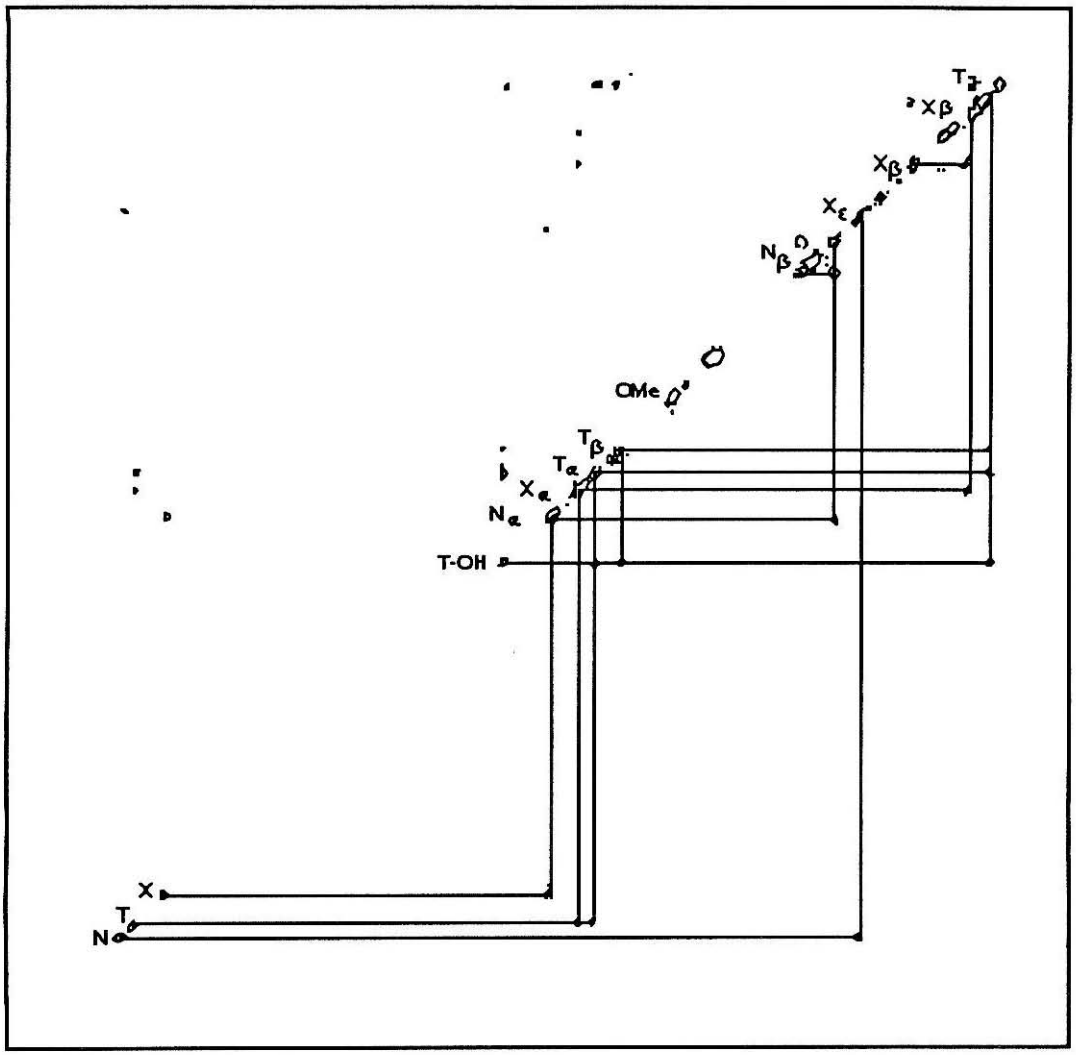


Figure 5-14 ROESY spectrum of the dimeric pimelic macrocycle in DMSO.

The NOE connectivities observed indicated that the macrocycle is fixed into Asx-turn conformation. A strong backbone sequential connectivity $d_{\alpha N}$(Asn, Xaa) is observed. Also present is a strong $d_{\varepsilon N}$(Apm, Asn) connectivity. The temperature coefficients for the amide protons are: Asn-NH = 4.6, Apm-NH = 4.9, and Thr-NH = 2.4. Thr-NH is the only solvent shielded amide proton (Figure 5-15).

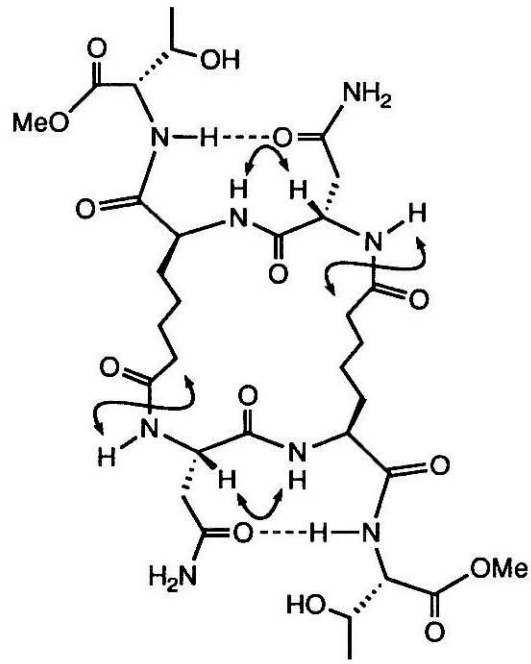


Figure 5-15 Graphical representation of NOE connectivities and the hydrogen bound amides derived from spectroscopic studies of the dimeric pimelic macrocycle.

The glycosyl acceptor properties for the dimeric pimelic macrocycle were determined using standard assay conditions (Table 5-6). The uncyclized dimer was also examined to determine the effect of cyclization on enzymatic activity.

Since the carboxyl termini of both the linear and cyclized dimeric material are methyl esters, the data was also compared against the closest versions of the standard substrate which were Boc-NLT-OMe and Ac-NLT-OMe.

Table 5-6 Enzyme assay results from dimeric studies.

peptide	K_m (mM)	V_{max} (dpm min^{-1})	V_{max}/K_m
Boc-NLT-OMe	1.13	186	165
Boc-NX(NXT-OMe)T-OMe[1]	>30[2]	-	-
Ac-NLT-OMe	3.2	524	164
pimelic macrocycle	0.55	263	478

1 X = Apm(OBn).
2 Residual activity was apparent at 15 mM.

The enzymatic data was surprising for the following reason. The linear pimelic dimer, Boc-NX(NXT-OMe)T-OMe, was a very poor acceptor. This was unexpected since it is has two potential glycosylation sites which are not constrained (Figure 5-16).

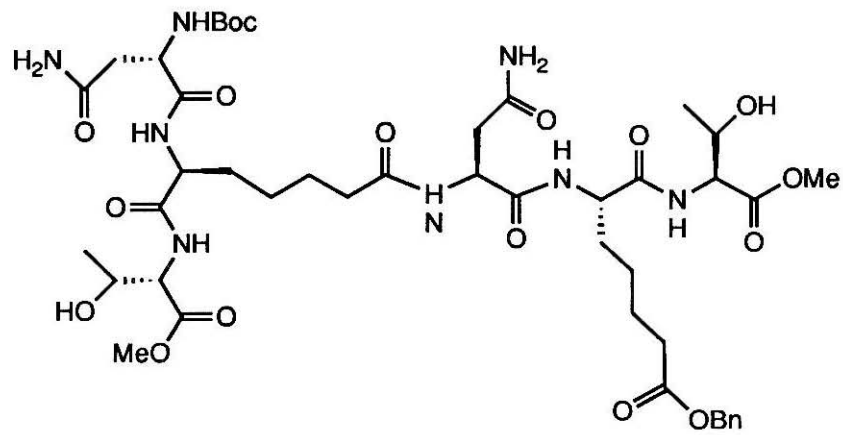


Figure 5-16 Representation of the linear pimelic dimer, Boc-NX(NXT-OMe)T-OMe. X = Apm(OBn)

Once the macrocycle formed, the compound acquired substrate behavior. The relative activity of this macrocycle compared to Ac-NLT-OMe is slightly deceiving due to the presence of two putative glycosylation sites. The data is also complicated by the insolubility of the macrocycle; at concentrations higher than 20 mM, the macrocycle is only partially soluble. The macrocycle is poorly soluble in water which makes performing the assay more complicated. It is likely that the actual concentration of the compound in the assay is lower than recorded. Thus the velocities observed for the macrocycle relate to lower concentrations. It is likely that if the macrocycle could be better solubilized, the V_{max} would increase.

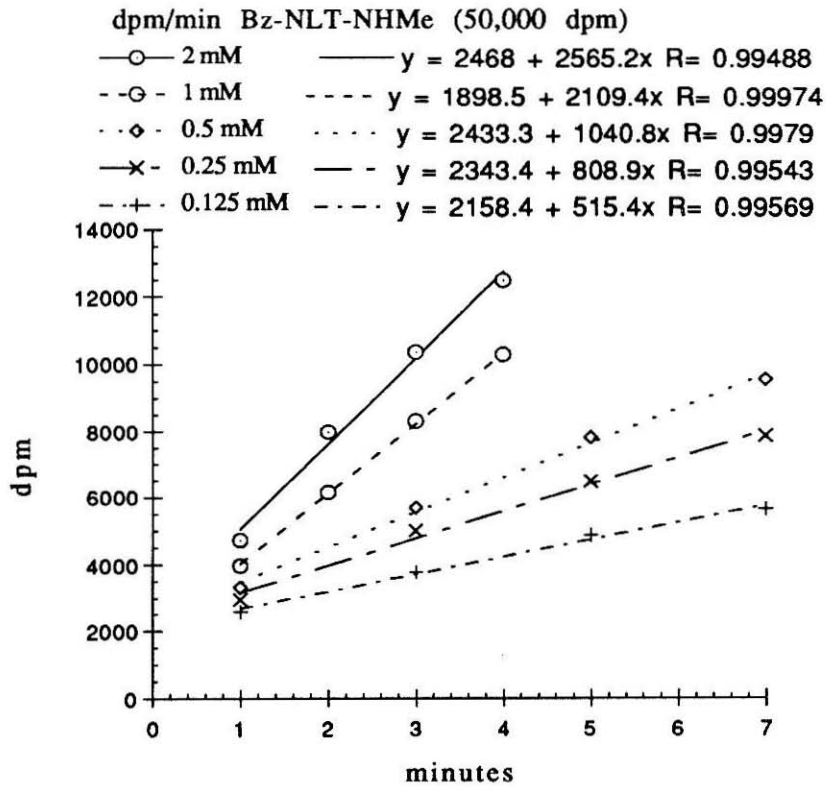
Nonetheless, an examination of the kinetic data reveals that the macrocycle is 4 fold more active relative to Ac-NLT-OMe. If both marker sequences are serving as active acceptors, then the activity is still twice as high as the unconstrained standard. More importantly, the pimelic containing peptide only acquires activity when it is cyclized. Since the solution conformation of this compound indicated that it is rigidly fixed into Asx-turn conformation, it is highly likely that this is indeed the bioactive conformation.

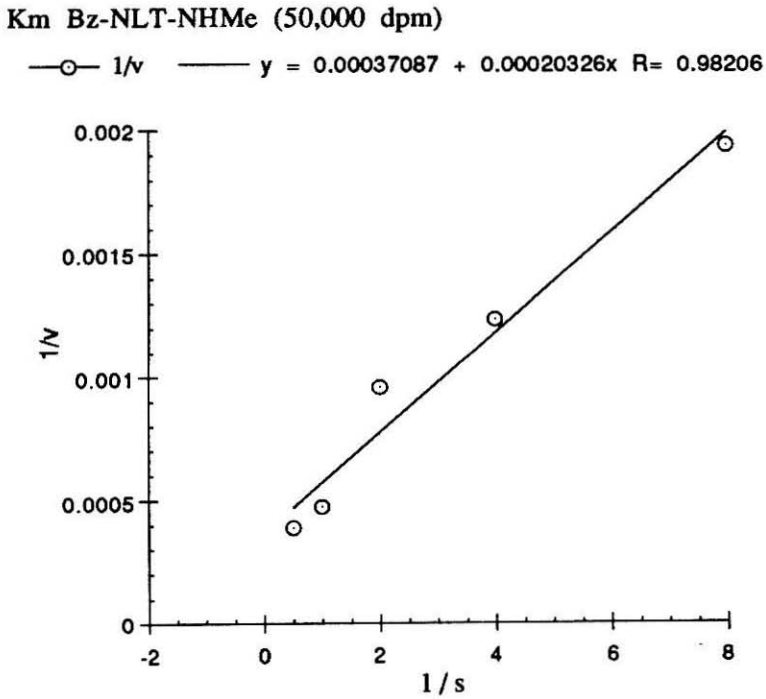
CONCLUSIONS

Hexapeptides fixed into β-turn conformations were shown to be relatively poor acceptors for glycosylation. Instead, the bioactive conformation was demonstrated to be an Asx-turn. Further analogs which fix the Asx-turn while containing only one

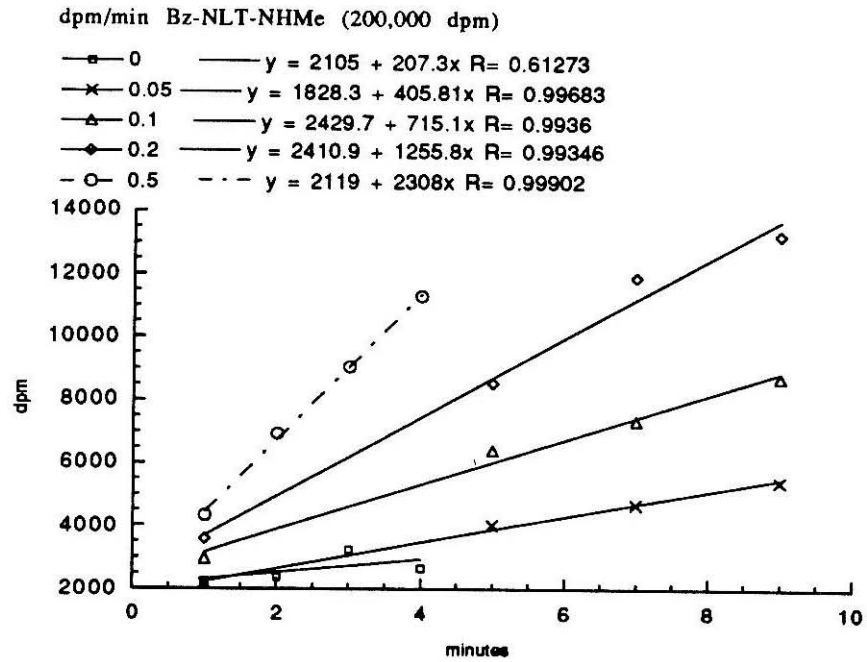
putative glycosylation site are required before the enzymatic properties can be evaluated definitively.

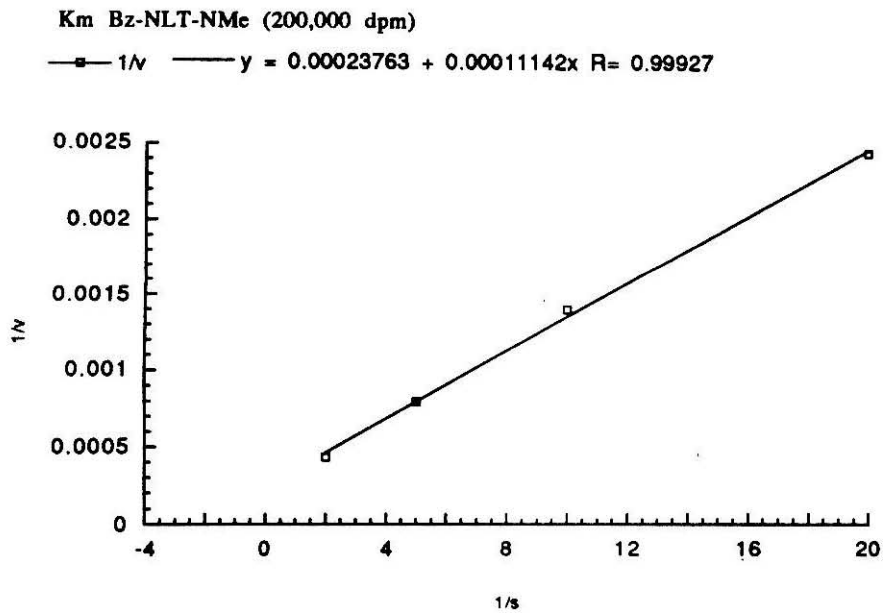
Appendix 5-1 Enzyme kinetic data for Bz-NLT-NHMe (50,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



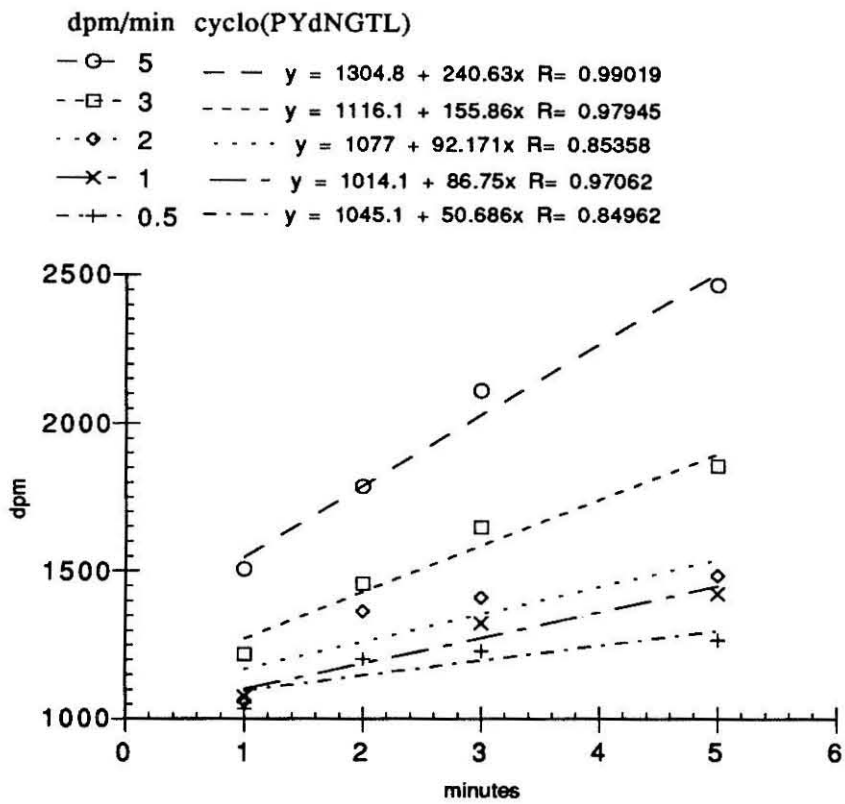


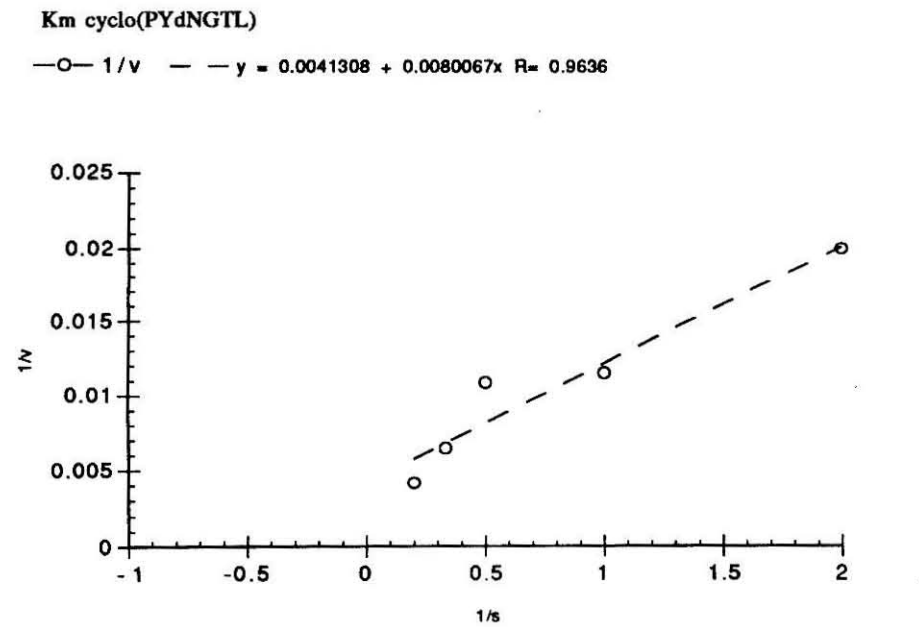
Appendix 5-2 Enzyme kinetic data for Bz-NLT-NHMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



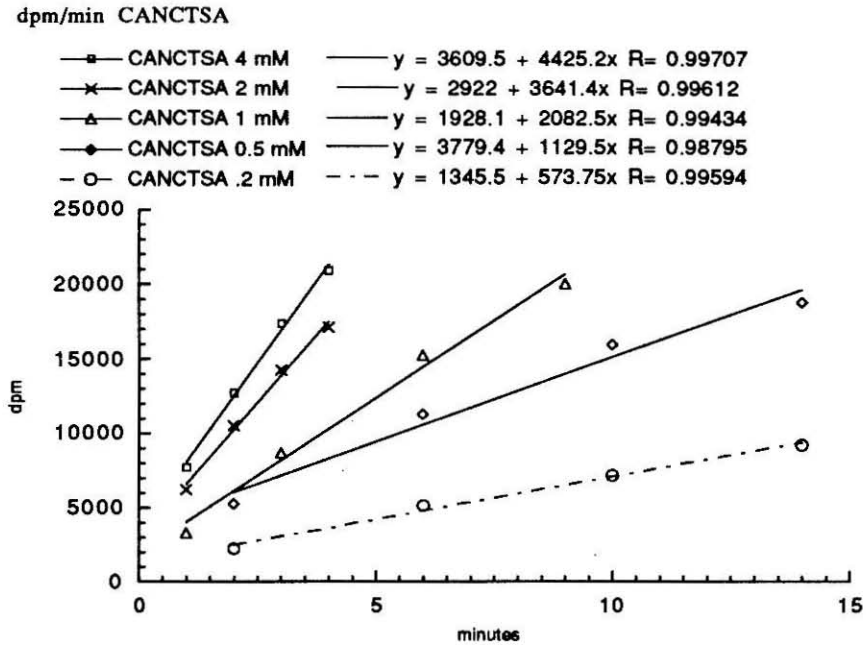


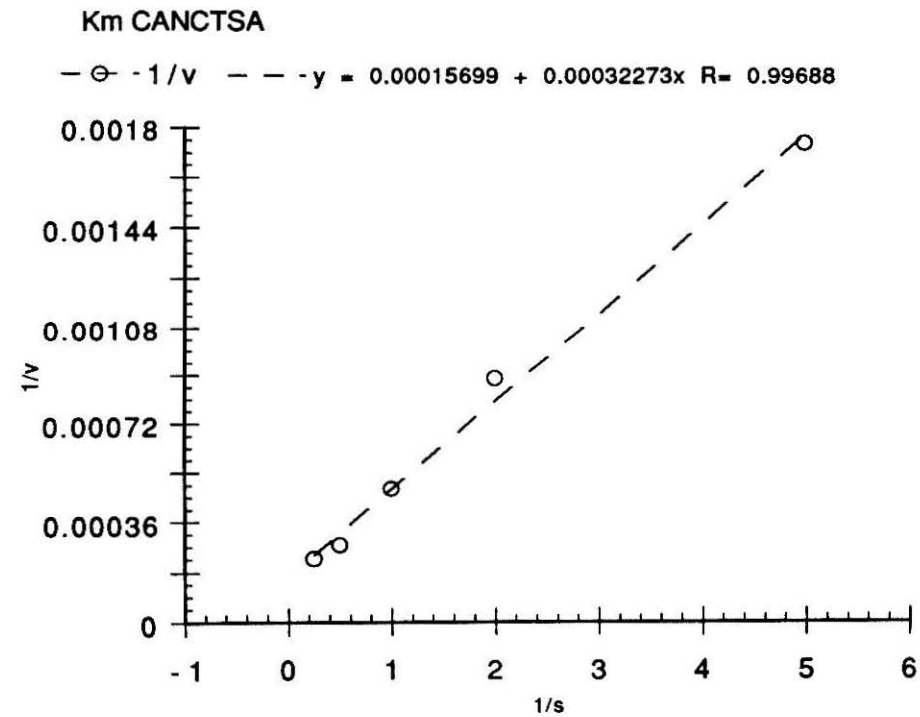
Appendix 5-3 Enzyme kinetic data for cyclo(PY_dNGTL) (50,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



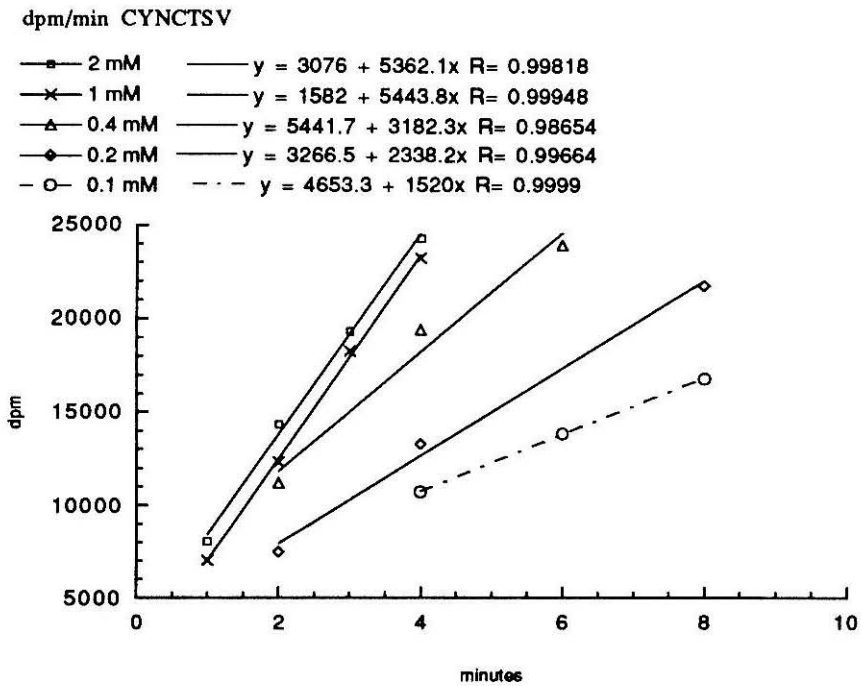


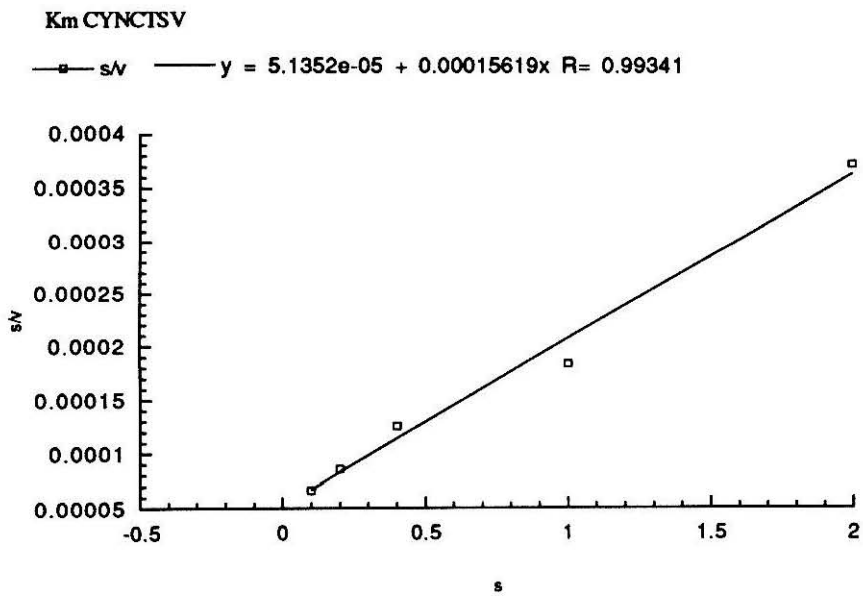
Appendix 5-4 Enzyme kinetic data for cyclo(CANCTSA) (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



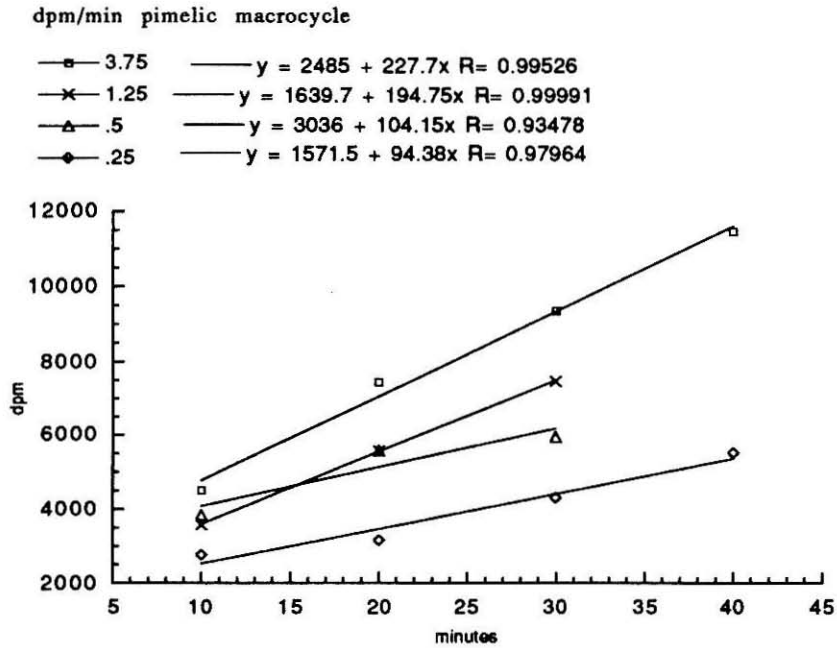


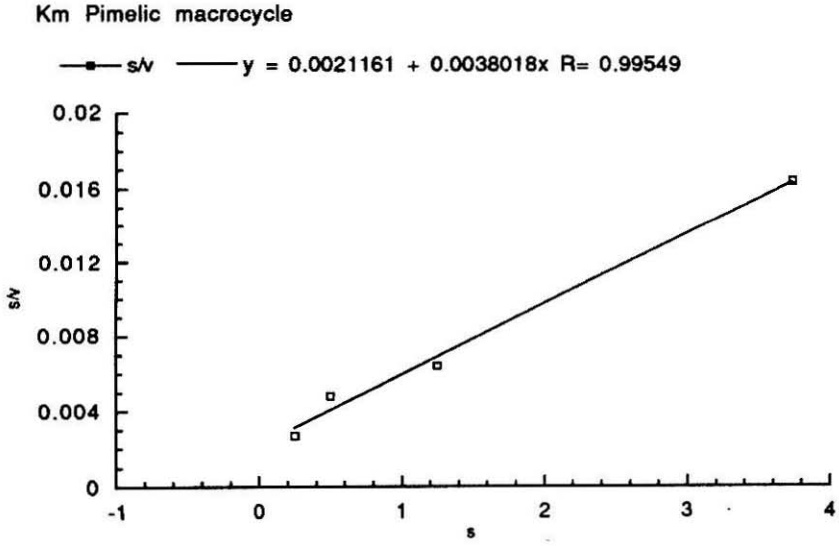
Appendix 5-5 Enzyme kinetic data for cyclo(CYNCTSV) (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



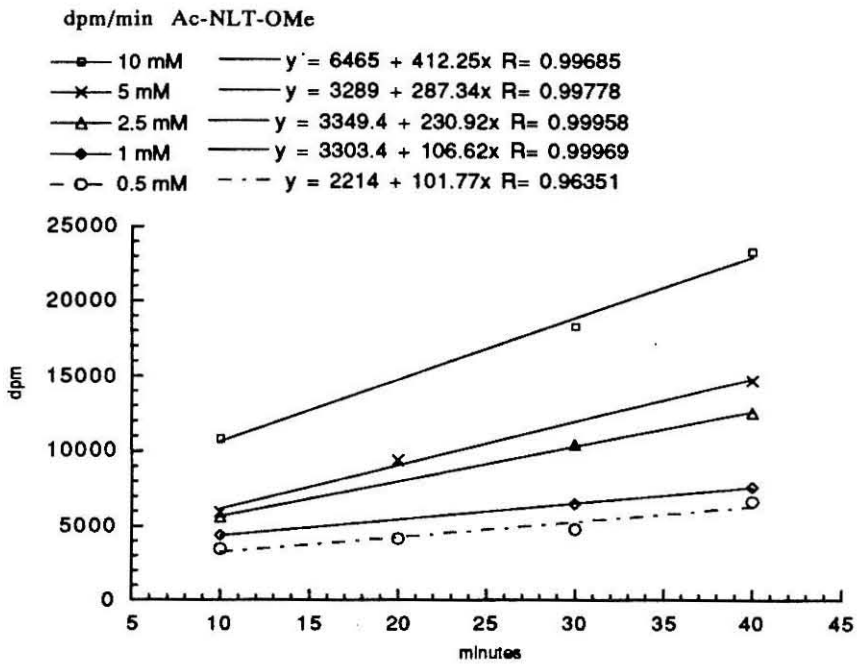


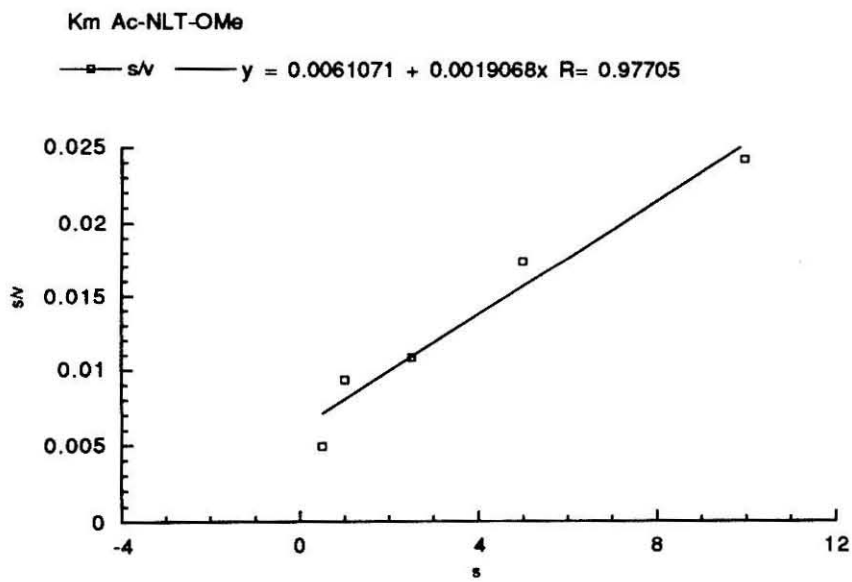
Appendix 5-6 Enzyme kinetic data for pimelic macrocycle (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)





Appendix 5-7 Enzyme kinetic data for Ac-NLT-OMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)





1. Hruby, V.J. "A Perspective on the Application of Physical Methods to Peptide Conformational-Biological Activity Studies," *Peptides* **1985**, 7, 1-14.

2. a) Hruby, V.J. in "Perspectives in Peptide Chemistry," (Wieland, T.; Geiger, R.; Eberle, A., eds.), Karger, Basel, 1981, p. 207. b) Hruby, V.J. "Conformational Restrictions of Biologically Active Peptides via Amino Acid Side Chain Groups," *Life Sci.* **1982**, *31*, 189-199.

3. DeGrado, W.F. "Design of Peptides and Proteins," *Adv. Protein. Chem.* **1988**, *39*, 51-124.

4. Hruby, V.J.; Kao, L.-F.; Hirning, L.D.; Burks, T.F. In "Peptides: Structure and Function," Deber,C.M.; Hruby, V.J.; Kopple, K.D., Eds.; Pierce Chemical Co., Rockford, Illinois, 1985; p. 487.

5. reviewed by Toniolo, C. "Conformationally restricted peptides through short-range cyclizations", *Int. J. Peptide Protein Res.* **1990**, *35*, 287-300.

6. a) Kopple, K.D.; Ohnishi, M.; Go, A. "Conformations of Cyclic Peptides. VI. Factors Influencing Mono-,1,4- Di-, and 1,2,4-Trisubstituted Cyclic Hexapeptide Backbones," *J. Am. Chem. Soc.* **1972**, *94*, 973-981. b) Kopple, K.D.; Ohnishi, M.; Go, A. "Conformations of Cyclic Peptides. IV. Nuclear Magnetic Resonance Studies of cyclo-Pentaglycyl-L-leucyl and cyclo-Diglycyl-L-histidyldiglycl-L-tyrosyl," *Biochemistry* **1969**, *8*, 4087-4095.

7. Gierasch, L.M.; Thompson, K.F.; Rockwell, A.L.; Briggs, M.S. "Conformation-Function Relationships in Hydrophobic Peptides: Interior Turns and Signal Sequences," *Biopolymers* **1985**, *24*, 117-135.

8. Bause, E.; Hettkamp, H.; Legler, G. "Conformational Aspects of N-Glycosylation of Proteins," *Biochem. J.* **1982**, *203*, 761-768.

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

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

11. Evans, D.A.; Ellman, J.A. "The Total Synthesis of the Isodityrosine-derived Cyclic Tripeptides OF4949-111 and K-13. Determination of the Absolute Configuration of K-13," *J. Am. Chem. Soc.* **1989**, *111*, 1063-1067.

Chapter 6
Mechanistic Implications

As discussed in the preceeding chapters, the central reaction in N-linked glycosylation is catalyzed by oligosaccharyl transferase and culminates in the formation of the primary bond between carbohydrate and protein. Overall, a complex carbohydrate moiety is transferred to an asparagine residue which is located within the consensus triad -Asn-Xaa-Thr/Ser- (Figure 6-1). The mechanism of this transfer is not fully understood but appears to involve nucleophilic attack by the carboxyamido group of the asparagine at the sp^3 anomeric carbon of the sugar.

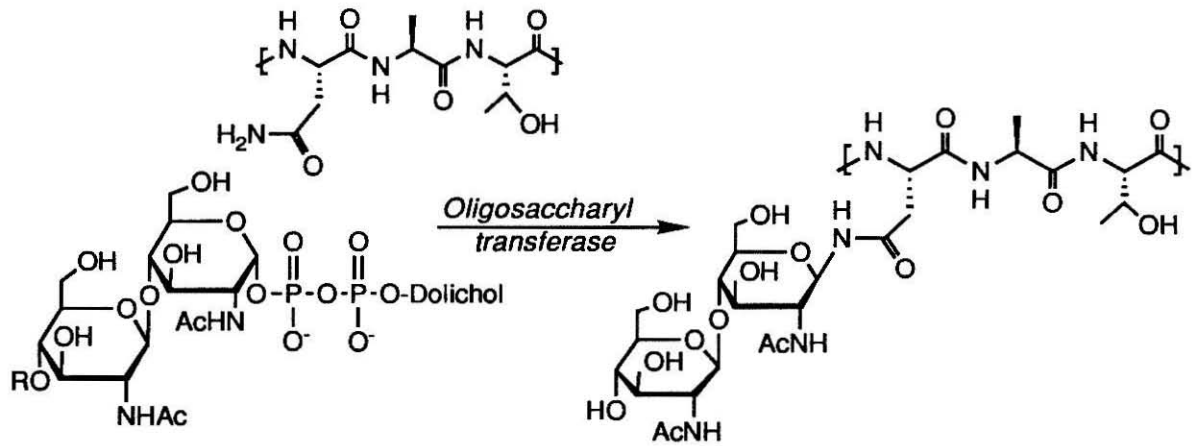


Figure 6-1 Overall reaction catalyzed by oligosaccharyl transferase.

Little mechanistic information regarding this transfer step is available for two reasons: 1) the purification of this membrane bound enzyme is treacherous and 2) once removed from its lipid environment the enzyme is extremely unstable.

Until the enzyme can be purified in its active form, specific information regarding the mechanism of the oligosaccharide

transfer step must be gleaned from substrate studies of the lipid-linked oligosaccharide donor and of the polypeptidyl acceptor.

The transfer of carbohydrate from the lipid donor to the peptidyl acceptor occurs with inversion of configuration at the anomeric center. This process could occur in a one-step displacement reaction with the carboxyamido nitrogen attacking C_1 of N-acetylglucosamine with concomitant displacement of dolichol pyrophosphate (Figure 6-2).

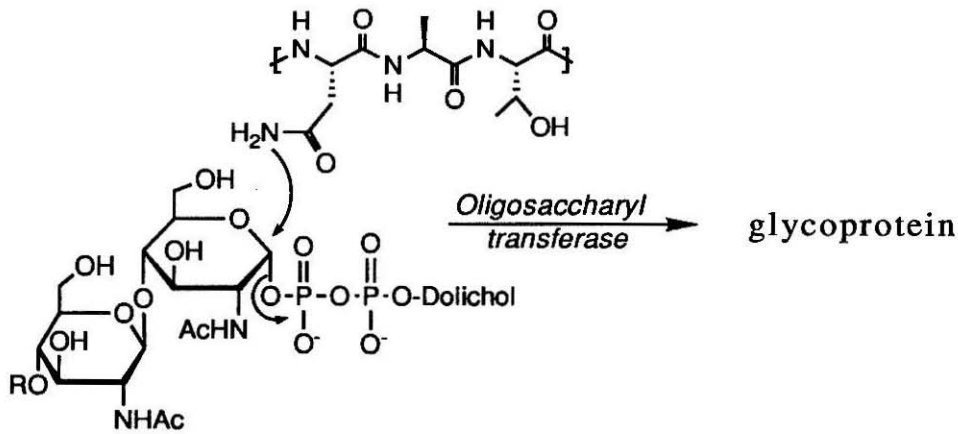


Figure 6-2 Concerted mechanism

Alternatively, the reaction could proceed via an oxocarbonium sugar intermediate, generated by cleavage of the carbon-oxygen bond, with subsequent attack by the amide nitrogen (Figure 6-3).

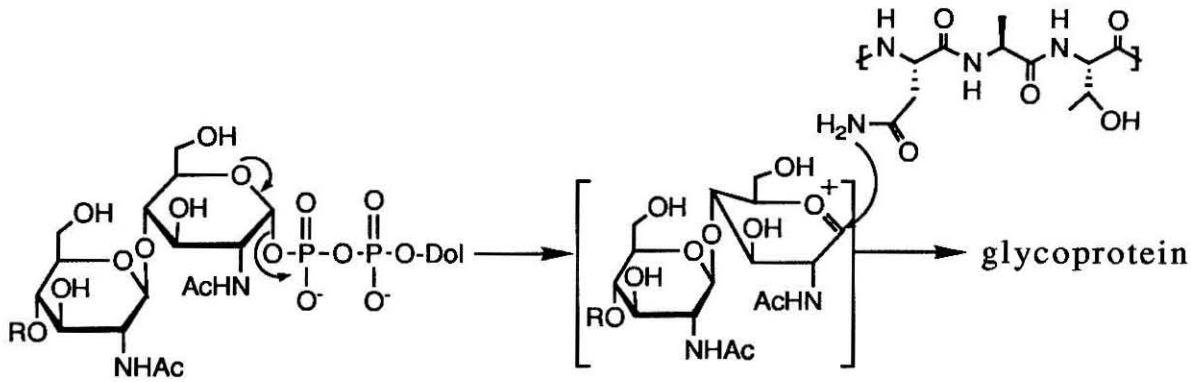


Figure 6-3 Two step mechanism

The oxocarbonium species may be stabilized by the neighboring N-acetyl group. A similar two step mechanism has been proposed for the reaction catalyzed lysozyme, a β-glucosidase (**1**). Lysozyme catalyzes the hydrolysis of a polysaccharide composed of alternating β(1,4) linked units of N-acetylglucosamine and N-acetylmuramic acid. The reaction proceeds with retention of configuration (Figure 6-4).

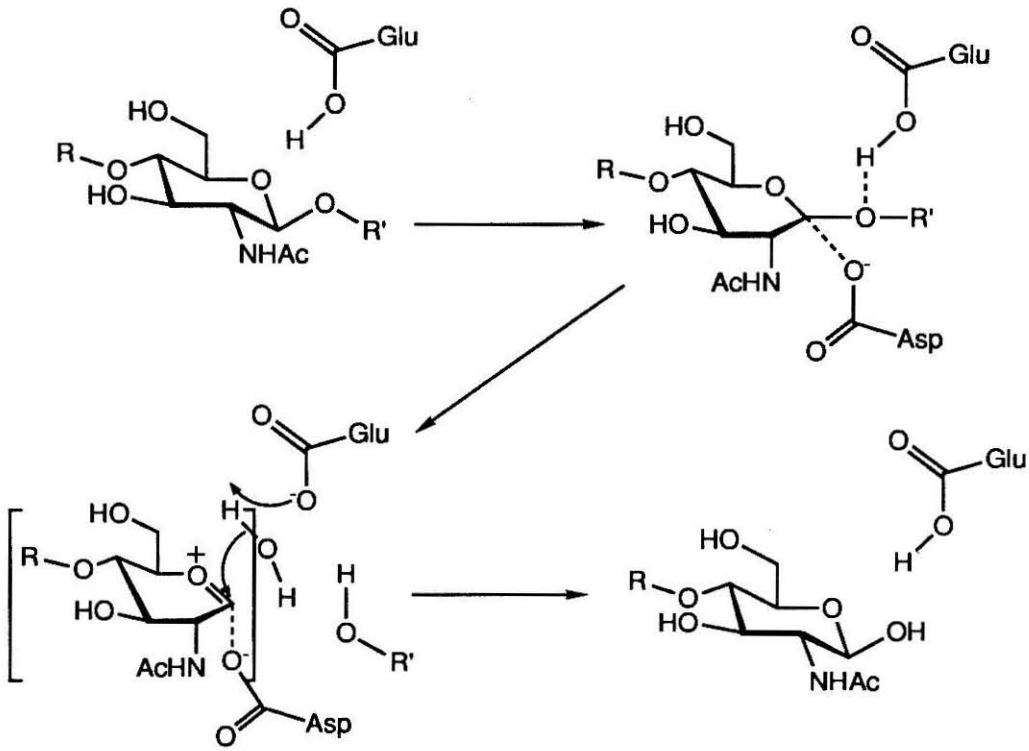


Figure 6-4 Proposed reaction pathway catalyzed by lysozyme.

A glutamate residue located at the active site weakens the β(1,4) bond by withdrawing electron density into a hydrogen bonding interaction. A second residue, aspartate, electrostatically stablizes the oxocarbonium ion intermediate which forms in the bond breaking step. This intermediate is highly electrophilic and is readily attacked by water. The intermediate is characterized by a half-chair conformation and a sp^2 center within the ring. Evidence supporting this mechanism was provided by the inhibition of lysozyme with a transition state analog, tri-N-acetylglucosaminyl-N-acetylglucosamine lactone (Figure 6-5).

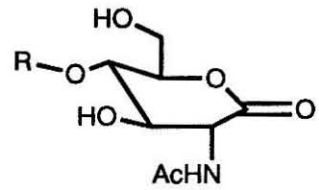


Figure 6-5 Structure of tri-N-acetylglucosaminyl-N-acetylglucosamine lactone (R=GlcNAc).

The two-step lysozyme-like mechanism proposed for glycosylation has never been demonstrated. Stereochemically, the mechanism does not require formation of the oxocarbonium species since the product, containing an $\alpha(1,4)$ linkage, would be generated by direct attack with inversion of configuration; however, a two step mechanism cannot be ruled out.

Central to understanding the mechanism of N-glycosylation is ascertaining how the reactivity of a relatively poor nucleophile, the carboxyamido group, can be enhanced. Amides are planar and resonance stabilized. Since nitrogen's lone pair of electrons are involved in a π-bonding interaction with carbon, the nucleophilicity of nitrogen is reduced. To increase the amide-nitrogen nucleophilicity, the amide resonance interaction must be disrupted.

One way of disrupting the amide resonance interaction is through a hydrogen bonding array. The necessity of a hydroxyl amino acid two residues away from the asparagine has prompted several researchers to postulate that this residue functions as either a hydrogen bond donor or acceptor. Bause and Legler proposed that a general base near the active site could remove the proton from the β-hydroxyl oxygen (**2**). The resulting oxyanion

would be strongly nucleophilic and could subsequently abstact a proton from the β-carboxyamido nitrogen of the asparagine (Figure 6-6). This would render the asparagine sufficiently nucleophilic at nitrogen to attack the lipid-linked oligosaccharide donor.

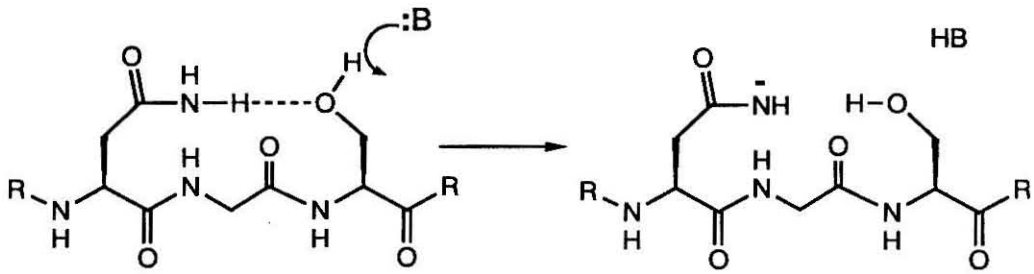


Figure 6-6 Bause's proposed activation of the carboxyamido nitrogen.

This proton relay system would explain the observed differences in the glycosylation rates of cysteine, serine, and threonine (Table 6-1) (**2**).

Table 6-1 Glycosyl acceptor capabilities of hexapeptides in the series Tyr-Asn-Gly-Xaa-Ser-Val. Relative transferase activity (V_{max}/K_m) was based on the serine-containing peptide.

Xaa	K_m (mM)	V_{max} (cpm/10 min)	relative transferase activity
Thr	0.16	2080	40
Ser	1.6	550	1
Cys	-	-	0.4[a]
Val	-	-	-
Thr(OMe)	-	-	-

[a] residual activity was evident at high concentrations of peptide.

These results showed that serine-, threonine-, and cysteine-containing peptides could be glycosylated, although at very different rates, whereas valine- and O-methylthreonine-containing peptides were not substrates. The observed differences in the transfer rates (cys < ser < thr) are compatible with the decreasing efficiency of these residues to serve as hydrogen bond acceptors. Thus there appears to be an absolute requirement connected with the hydrogen bond donor ability of the hydroxyl amino acid.

In an earlier attempt to interpret the acid-base titration behavior of glycopeptides, Marshall constructed a different hydrogen bonding array (**3**). In this model, the hydroxyl group of the hydroxyl amino acid is involved in a hydrogen bond with the amide oxygen of the asparaginyl side chain (Figure 6-7).

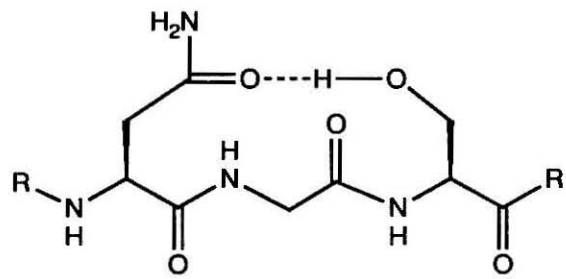


Figure 6-7 An alternate hydrogen bonding array proposed by Marshall.

Under these conditions a nucleophilic attack on the lipid-linked oligosaccharide would only be expected when it is preceeded by abstraction of a proton from the amide group. However, initial deprotonation is unlikely due to the relatively low base strengths

of the likely candidates (thiolate, carboxylate, imidazole) compared to the pK_a (14-15) of the carboxyamido nitrogen donor.

Another possible mode of disrupting the amide resonance and thereby increasing the carboxyamido nucleophilicity is through the addition of a nucleophile to the carbonyl carbon to generate a tetrahedral intermediate. Breakdown of this intermediate with expulsion of the amide nitrogen would result in production of nucleophilic ammonia and an acyl derivitive of the nucleophile. This scheme is analogous to the well-established mechanism for glutamine-dependent amino transfer reactions (**4**) and has been recently suggested for the glycosyl transfer reaction (**5**).

In this proposal, asparagine is activated by an enzyme bound base and deaminates to form an enzyme bound acyl intermediate with release of ammonia. Subsequent nucleophilic displacement of pyrophosphyldolichol by ammonia generates an intermediate 1-amino oligosaccharide. This intermediate then recombines with the acyl enzyme intermediate to produce the final glycopeptide.

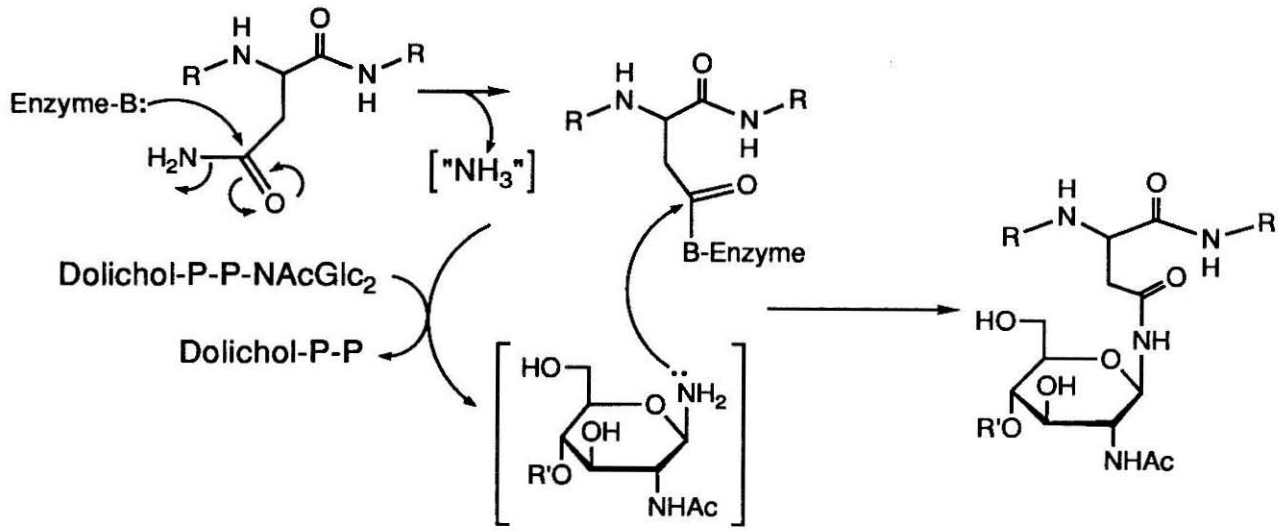


Figure 6-8 Reaction mechanism proposed by Clark *et al.*

The downfall of this mechanism is that it neither profers a role for the essential hydroxy amino acid residue in the transfer reaction, nor does it attempt to explain the specificity observed for the reaction.

The former mechanistic proposals which incorporate hydrogen bonding arrays are more likely because they not only require the hydroxy residue but also have stabilized secondary structures. In the preceeding chapters it was demonstrated that the conformation of the polypeptidyl acceptor is a crucial element in glycosylation. The conformation of the polypeptidyl acceptor establishes a microenvironment surrounding the carboxyamido group which may subtly govern the specificity observed in glycosylation.

In chapters 3 to 5, the dominant solution conformation of the peptidyl acceptors examined was shown to be an Asx-turn. This structure is formed via a hydrogen bonding interaction

between the carbonyl oxygen of the asparagine side chain and the amide proton of the hydroxy residue. The Asx-turn is a common structural motif found in proteins. When the glycosylation triad exists in this conformation, the carboxyamido oxygen and the hydroxyl proton are within hydrogen bonding distance. Thus if hydrogen bonding interactions provide a means of inducing specificity into the mechanism, they are probably best depicted by a modified version of the array first described by Marshall (Figure 6-9).

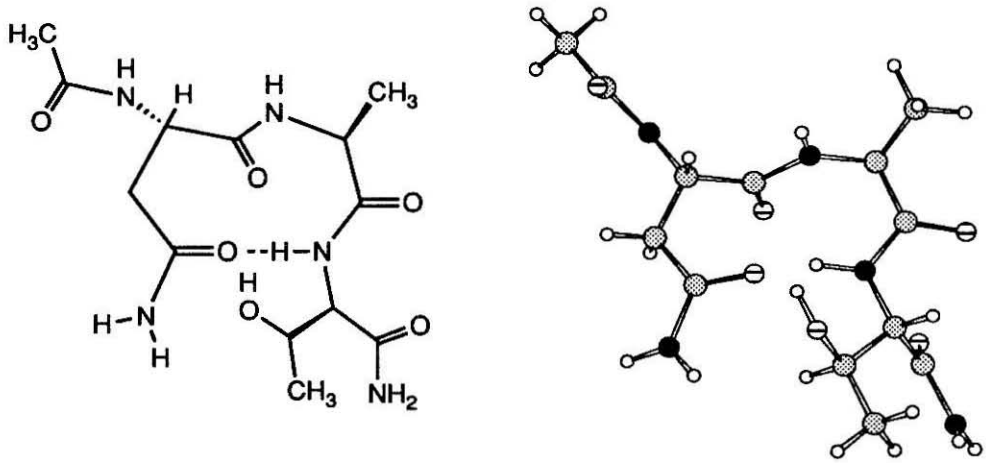


Figure 6-9 The hydrogen bonding array which is conformationally consistent with an Asx-turn.

This array contains a second hydrogen bond between the side chain carbonyl oxygen of asparagine and the hydroxyl proton of the hydroxy amino acid. This added hydrogen bonding interaction may serve both to further stabilize the Asx-turn and to align the hydroxyl proton with the more basic lone pair of electrons of the carbonyl oxygen.

This chapter summarizes the use of peptidyl analogs to probe the catalytic mechanism of N-glycosylation. We have

synthesized a number of tripeptides that differ from one another by replacement of the essential residues, Asn and Thr/Ser, with various unnatural amino acid residues.

EXPERIMENTAL METHODS

Peptide Synthesis

All peptides were synthesized using standard solution phase methods (6). Coupling reactions were performed with DCC/HOBT/triethylamine in THF or DMF at a minimum reagent concentration of 0.1 M.

Bz-Ala(β-cyano)-Leu-Thr-NHMe

BOC-Ala(β-CN) was prepared in 72% yield from Ala(β-CN) with BOC-ON and triethylamine in dioxane-water. BOC-Leu.H_2O was reacted with Thr-OMe.HCl using DCC/HOBT/triethylamine in DMF to afford BOC-Leu-Thr-OMe in 90%. BOC-Leu-Thr-OMe was deprotected with trifluoroacetic acid and dichloromethane followed by the reaction with BOC-Ala(β-CN) and DCC/HOBT/triethylamine and BOC-Ala(β-CN)-Leu-Thr-OMe was obtained in 88%. Then the reaction with methylamine in methanol gave 45% of BOC-Ala(β-CN)-Leu-Thr-NHMe. BOC group was removed as described previously. Subsequent reaction with benzoic anhydride and triethylamine in DMF yielded 60% of Bz-Ala(β-CN)-Leu-Thr-NHMe after purification (PTLC eluent 5:1 chloroform/methanol).

^{1}H NMR (d_6-DMSO) δ: 0.81 (3H, d, J=6.5 Hz); 0.84 (3H, d, J=6.5 Hz); 0.98 (3H, d, J=6.3 Hz); 1.48 (2H, m); 1.59 (1H, m); 2.58 (3H, d,

J=4.6 Hz); 2.94 (1H, dd, J=10.3, 16.8 Hz); 3.02 (1H, dd, J=4.7, 16.8 Hz); 3.96 (1H, m); 4.08 (1H, dd, J=3.9, 8.7 Hz); 4.32 (1H, m); 4.82 (1H, m); 7.50 (2H, t, J=8.1 Hz); 7.56 (1H, d, J=7.3 Hz); 7.61 (1H, d, 8.1 Hz); 7.59 (1H, d, J=8.7 Hz); 7.88 (2H, d, J=7.4 Hz); 8.28 (1H, d, J=7.9 Hz); 8.96 (1H, d, J=8.1 Hz).

^{13}C NMR (d_6-DMSO) δ: 19.89, 21.47, 22.99, 24.05, 25.55, 40.21, 49.57, 51.76, 58.10, 66.36, 118.18, 127.43, 128.28, 131.64, 133.49, 166.66, 168.79, 170.19, 171.64.

High resolution MS [M^+] Calcd. for $C_{22}H_{32}N_5O_5$: 446.2403, obsd. 446.2415.

Bz-Ala-[β-(CSNH2)]-Leu-Thr-NHMe

Bz-Ala(β-cyano)-Leu-Thr-NHMe (0.208 mmol, 92.8 mg) was placed in a flask and 4 mL of saturated ammonia solution in absolute ethanol (4.3 M) was added. Then hydrogen sulfide was introduced to the reaction mixture at 0°C. The solution was stirred at room temperature for one day and the solvent was removed. Dry methanol was added to the mixture and sulfur was filtered out. Then the mixture was purified with PTLC (silica, chloroform/methanol 8/1) and 53.6 mg (53%) of thioamide was obtained. The compound was further purified by recrystallization from methanol/chloroform/hexane (1/1/10) which afforded analytically pure thioamide as a white solid (33 mg, 33%).

^1H NMR (d_6-DMSO) δ: 0.81 (3H, d, J=6.4 Hz); 0.85 (3H, d, J=6.5 Hz); 0.98 (3H, d, J=6.3 Hz); 1.49 (1H, m); 1.52 (1H, m); 1.58 (1H, m); 2.59 (3H, d, J=4.5 Hz); 2.97 (2H, br s); 3.98 (1H, m); 4.05 (1H, m); 4.33 (1H, m); 4.84 (1H, d, J=5.1 Hz); 4.96 (1H, m); 7.49 (2H, t, J=7.5 Hz); 7.54 (1H, t, J=7.4 Hz); 7.59 (1H, d, J=8.7 Hz); 7.84 (2H, d, J=7.5 Hz); 8.11 (1H, d, 8.0 Hz); 8.59 (1H, d, J=7.7 Hz); 9.15 (1H, s); 9.54 (1H, s).

^{13}C NMR (d_6-DMSO) δ: 19.91, 21.42, 23.15, 24.05, 25.58, 39.00, 40.20, 45.06, 53.28, 58.22, 66.29, 127.39, 128.18, 131.19, 133.96, 166.50, 170.22, 170.70, 171.88, 203.98.

High resolution MS [M^+]. Calcd. for $C_{22}H_{34}N_5O_5S$: 480.2281, obsd. 480.2269.

Bz-Ala(β-[C(NH_2)=NH.HCl])-Leu-Thr-NHMe (Bz-Tan-Leu-Thr-NHMe)

Under nitrogen atmosphere, A saturated HCl solution in absolute ethanol (8.6 M, 2 mL) was added to Bz-Ala(β-cyano)-Leu-Thr-NHMe (58.2 mg, 0.131 mmol) at 0°C. The solution was stirred at that temperature for 1 h until no starting material remained by TLC (silica, chloroform:methanol=5:1). The solvent was removed with nitrogen flow at room temperature. Care was taken to avoid moisture. The mixture was dried *in vacuo* after which a saturated ammonia solution in absolute ethanol (4.3 M, 2 mL) was added at RT. After stirring 1 h at 0°C, the mixture was concentrated and dissolved in distilled water (5 mL) and washed with chloroform (3 x 5 mL). Water was removed and the resulting solid was purified via preparative HPLC.

^1H NMR (d_6-DMSO) δ: 0.84 (d, 3H, J=6.3Hz), 0.87 (d, 3H, J=6.3Hz), 0.96 (d, 3H, J=6.4Hz), 1.51 (m, 2H), 1.55 (m, 1H), 2.54 (d, 3H, J=4.5Hz), 2.80 (dd, 1H), 2.98 (dd, 1H), 3.99 (m, 1H), 4.09 (m, 1H), 4.40 (m, 1H), 4.85 (m, 1H), 7.48 (t, 2H, J=7.5Hz), 7.56 (t, 1H, J=7.4Hz), 7.56 (d, 1H, J=4.6Hz), 7.69 (d, 1H, J=8.6Hz), 7.87 (d, 2H, J=-7.5Hz), 8.24 (d, 1H, J=8.4Hz), 8.75 (br s, 2H), 8.76 (d, 1H, J=7.2Hz), 8.86 (br s, 2H).

^{13}C NMR (d_6-DMSO) δ: 19.97, 21.41, 23.09, 24.15, 25.59, 28.35, 34.39, 40.43, 51.44, 51.68, 58.39, 66.24, 127.56, 128.27, 131.73, 133.46, 166.95, 167.86, 169.47, 170.30, 171.79.

High resolution MS [M^+]. Calcd. for $C_{22}H_{35}N_6O_5$: 463.2669, obsd. 463.2663.

Bz-γ-Aminobutyrine-Leu-Thr-NHMe (Bz-Amb-Leu-Thr-NHMe)

Bz-Ala(CN)-Leu-Thr-NHMe (20 mg, 0.046 mmol) was dissolved in 3.6 mL methanol/acetic acid/chloroform (3/0.1/0.5). Under nitrogen atmosphere, 50 mg of 5% Pd/carbon was added and the solution was then stirred under atmospheric hydrogen pressure for 8 h. At this time the reaction mixture was filtered and the product purified by PTLC (chloroform:methanol 5:1). The reaction afforded 13 mg (59% yield) of the target amine as a white solid.

^{1}H NMR (d_6-DMSO) δ: 0.84 (d, 3H, J=6.3 Hz); 0.88 (d, 3H, J=6.4 Hz); 0.98 (d, 3H, J=6.1 Hz); 1.50 (m, 2H); 1.63 (m, 1H); 1.99 (m, 1H); 2.07 (m, 1H); 2.59 (d, 3H, J=4.2 Hz); 2.90 (m, 2H); 3.96 (m, 1H); 4.09 (m, 1H); 4.35 (m, 1H); 4.54 (m, 1H); 4.83 (d, 1H, J=4.7 Hz); 7.47 (t, 2H, J=7.4 Hz); 7.55 (t, 1H, J=7.1 Hz); 7.65 (d, 1H, J=4.5 Hz); 7.66 (d, 1H, J=8.9 Hz); 7.76 (br s, 2H); 7.88 (d, 2H, J=7.5 Hz); 8.15 (d, 1H, J=7.6 Hz); 8.63 (d, 1H, J=7.5 Hz).

^{13}C NMR (d_6-DMSO) δ: 19.88, 21.41, 23.05, 24.09, 25.51, 29.63, 36.22, 40.59, 51.20, 51.60, 58.33, 66.40, 127.55, 128.12, 131.37, 133.81, 166.63, 170.31, 170.84, 171.90.

High resolution MS [M^+]. Calcd. for $C_{22}H_{36}N_5O_5$: 450.2716, obsd. 450.2715.

Bz-Asp-Leu-Thr-NHMe

^{1}H NMR (d_6-DMSO) δ: 0.801 (d, 3H, J= 6.4 Hz), 0.839 (d, 3H, J= 6.6 Hz), 0.972 (d, 3H, J= 6.4 Hz), 1.483 (m, 2H), 1.586 (m, 1H), 2.580 (d, 3H, J= 4.7 Hz), 2.686 (dd, 1H, J= 16.8 Hz, 9.1 Hz), 2.802 (dd, 1H, J = 16.7 Hz, 5.0 Hz), 3.978 (m, 1H), 4.061 (dd, 1H, J = 8.6 Hz, 3.7 Hz), 4.301(m, 1H), 4.782 (m, 1H), 7.483

(m, 2H), 7.530 (m, 3H), 7.840 (d, 2H, J= 8.5 Hz), 8.104 (d, 1H, J= 8.0 Hz), 8.696 (d, 1H, J= 7.5 Hz).

^{13}C NMR (d_6-DMSO) δ: 19.93, 21.48, 23.07, 24.10, 25.62, 35.62, 35.64, 50.29, 51.64, 58.20, 66.27, 127.46, 128.17, 131.37, 133.93, 166.55, 170.24, 170.85, 171.85, 171.91

Mass Spectroscopy $[MH]^+$ 464

Bz-Asp(γ-OMe)-Leu-Thr-NHMe (7)

^{1}H NMR (d_6-DMSO) δ: 0.804 (d, 3H, J=6.6 Hz), 0.843 (d, 3H, J=6.6 Hz), 0.976 (d, 3H, J= 6.3 Hz), 1.486 (m, 2H), 1.600 (m, 1H), 2.583 (d, 3H, J=4.5 Hz), 2.763 (dd, 1H, J=16.5 Hz, 8.9 Hz), 2.887 (dd, 1 H, J=16.5 Hz, 5.5 Hz), 3.587 (s, 3H), 3.981 (m, 1H), 4.069 (dd, 1H, J=8.6 Hz, 3.7 Hz), 4.301 (m, 1H), 4.831 (m, 2H), 7.468 (t, 2H, J=7.6 Hz), 7.539 (m, 3H), 7.837 (d, 2H, J=8.5 Hz), 8.152 (d, 1H, J=8.0 Hz), 8.722 (d, 1H, J=7.7 Hz).

^{13}C NMR (d_6-DMSO) δ: 19.91, 21.48, 23.04, 24.08, 25.51, 35.36, 35.40, 50.06, 51.49, 51.68, 58.13, 66.27, 127.43, 128.18, 131.40, 133.86, 166.55, 170.19, 170.50, 170.76, 171.83

High resolution MS $[M^+]$ = Calculated for $C_{23}H_{34}N_4O_7$: 479.2506; observed: 479.2504.

Bz-Asn(β-NHMe)-Leu-Thr-NHMe (7)

^{1}H NMR (d_6-DMSO) δ: 0.797 (d, 3H, J=6.5 Hz), 0.848 (d, 3H, J=6.5 Hz), 0.991 (d, 3H, J=6.5 Hz), 1.504 (m, 2H), 1.601 (m, 1H), 2.567 (d, 3H, J=4.5 Hz), 2.570 (m, 1H), 2.586 (d, 3H, J=5.0 Hz), 2.689 (dd, 1H, J=6.3 Hz, J=15.3 Hz), 4.001 (m, 1H), 4.068 (m, 1H), 4.293 (m, 1H), 4.750 (m, 1H), 4.804 (d, 1H, J=5.5 Hz), 7.452 (m, 2H), 7.542 (m, 2H), 7.591 (d, 1H, J=8.5 Hz), 7.826 (d, 2H, J=7.5 Hz), 7.897 (d, 1H, J=4.5 Hz), 8.187 (d, 1H, J=8.0 Hz), 8.628 (d, 1H, J=8.0 Hz)

^{13}C NMR (d_6-DMSO) δ: 19.97, 21.34, 23.08, 24.07, 25.52, 25.62, 36.76, 50.71, 51.67, 58.40, 66.30, 127.36, 128.21, 131.35, 133.98, 166.37, 170.09, 170.29, 171.10, 171.98

Mass Spectroscopy $[MH]^+$ 478

Bz-Asn-Leu-allo(Thr)-NHMe

^{1}H NMR (d_6-DMSO) δ: 0.81 (dd, 6H, J=6.6 Hz), 1.00 (d, 3H, 6.3 Hz), 1.4 (m, 2H), 1.6 (m, 2H), 2.5 (d, 3H, J=4.4 Hz), 2.55 (m, 2H), 3.8 (m, 1H), 4.06 (m, 2H), 4.25 (m, 1H), 4.85 (m, 1H), 4.86 (bs, 1H), 7.0 (s, 1H), 7.3 (s, 1H), 7.46 (m, 3H), 7.52 (m, 1H), 7.76 (m, 1H), 8.1 (d, 1H, J=7.6 Hz), 8.62 (d, 1H, J=7.6 Hz)

Mass Spectroscopy $[MH]^+$ 465

Bz-Asn-Leu-Ser-NHMe

^{1}H NMR (d_6-DMSO) δ: 0.8 (dd, 6H, J=6.4 Hz), 1.6 (m, 3H), 2.65 (m, 2H), 3.3 (s, 1H), 4.2 (m, 1H), 4.8 (m, 1H), 7.4 (m, 3H), 7.6 (m, 3H), 7.8 (1H, d, J=7.4 Hz), 7.9 (1H, d, J=7.6 Hz)

Mass Spectroscopy $[MH]^+$ 451

Bz-Asn-Leu-Val-NHMe

^{1}H NMR (d_6-DMSO) δ: 0.80 (d, 9H, J=6.6 Hz), 0.84 (d, 3H, J=6.5 Hz), 1.46 (m, 2H), 1.59 (m, 1H), 1.94 (m, 1H), 2.56 (d, 3H, J=4.4 Hz), 2.57 (dd, 1H), 2.63 (dd, 1H, J=5.9, 15.6 Hz), 4.02 (m, 1H), 4.34 (m, 1H), 4.74 (m, 1H), 6.95 (s, 1H), 7.37 (s, 1H), 7.46 (t, 2H, J=7.4 Hz), 7.53 (t, 1H, J=7.0 Hz), 7.63 (d, 1H, J=9.1 Hz), 7.75 (d, 1H), 7.83 (d, 2H, J=7.7.4 Hz), 8.08 (d, 1H, J=8.1 Hz), 8.59 (d, 1H, J=7.8 Hz) ppm.

Mass Spectroscopy $[MH]^+$ 463

RESULTS AND DISCUSSION

The transfer kinetics of eight peptides (Bz-Asn-Leu-Ser-NHMe, Bz-Asn-Leu-allo(Thr)-NHMe, Bz-Asn-Leu-Thr-NHMe, Bz-Asp-Leu-Thr-NHMe, Bz-Asp(γ-NMe)-Leu-Thr-NHMe, Bz-Asp(γ-OMe)-Leu-Thr-NHMe, Bz-Tam-Leu-Thr-NHMe, and Bz-Amb-Leu-Thr-NHMe) were determined using standard assay conditions and were compared against the standard substrate, Bz-Asn-Leu-Thr-

NHMe, to probe for relative enzymatic activity. These peptides were categorized as containing either a modification of the hydroxyl amino acid or of the asparagine residue.

The first three peptides were designed to probe the importance of the hydroxyl group as a binding determinant. The importance of the proton acceptor ability of the hydroxy residue was discussed early. The observed differences in the glycosylation rates of peptides containing modified hydroxy residues within the marker sequence was related to the decrease in base strength of the hydrogen bond acceptor functions (cysteine < serine < threonine). However the particular efficiency of the threonine-containing peptide compared to the serine-containing peptide (40/1) is not fully explainable with this rationale. To probe for binding interactions between the hydroxyl group and the enzyme, the stereochemistry of the hydroxy residue was varied. The kinetic data for these peptides is shown in Table 6-2.

Table 6-2 Glycosyl acceptor properties of tripeptides containing sterically modified hyroxyl amino acids within the series Bz-Asn-Leu-Xaa-NHMe.

peptide	K_m (mM)	V_{max} dpm min^{-1}	V_{max}/K_m	K_i (mM)
Bz-NLT-NHMe	0.47	4208	8953	-
Bz-NLS-NHMe	3.3	507	154	-
Bz-NL_{allo}T-NHMe	-	-	-	>10
Bz-NLV-NHMe	-	-	-	>10

Although the peptide formed by substitution of the hydroxyl group with a methyl group (Thr→Val) is sterically similar to the standard substrate, the glycosyl acceptor ability is

abolished. Furthermore, the peptide no longer binds to the enzyme as evidenced by inhibition studies with Bz-NLT-NHMe. No inhibition is evident at either low or high concentrations of substrate (Figure 6-10).

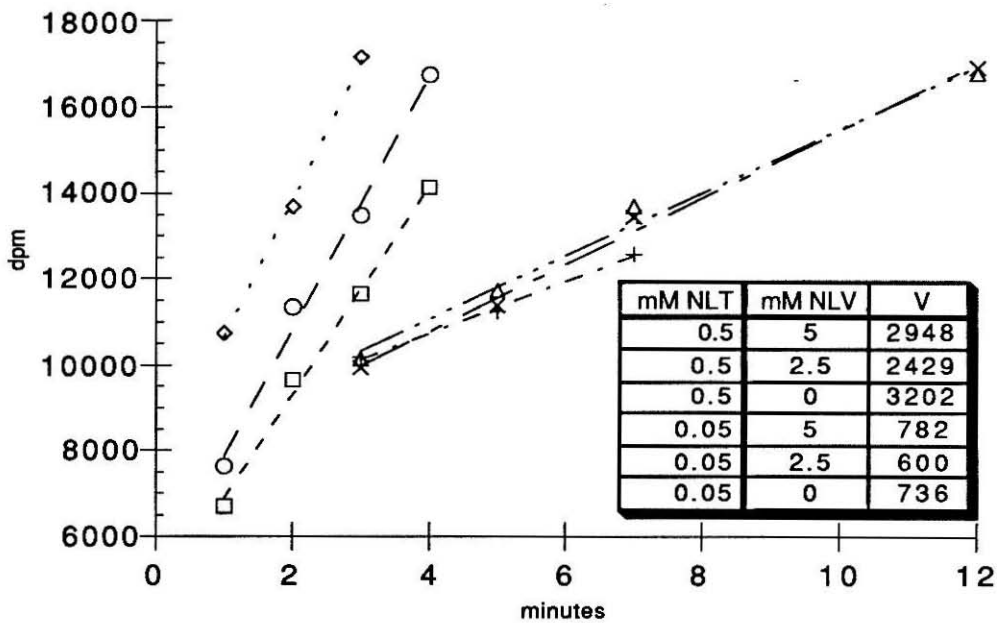


Figure 6-10 Inhibition study of Bz-NLT-NHMe by Bz-NLV-NHMe.

Examination of the 2-D ROESY spectra of Bz-NLV-NHMe indicated that the peptide exists predominantly in solution as an Asx-turn. Thus the hydroxyl group must make important binding interactions with the enzyme.

The position of the hydroxyl group also appears to affect binding. The sole modification in Bz-Asn-Leu-(allo)Thr-NHMe occurs at the Thr β-carbon. The stereochemistry of the α,β positions in naturally occurring threonine is (S), (R); in (allo)Thr it is (S), (S). This modification also resulted in a compound which

was a non-binder (non-acceptor, no inhibition). This suggests that the hydroxy residue may bind within a chiral pocket located on the surface of the transferase. The existence of such a pocket would explain the dramatic increase in efficiency of threonine-containing sites as compared to serine-containing sites. If a chiral pocket exists, the additional methyl group could serve to restrict rotation of the hydroxyl group, favorably positioning the hydroxyl proton for a hydrogen bonding interaction.

The balance of the peptides examined were used to probe for chemical features which are essential to the acceptor ability of the asparagine residue. Modifications which were designed to subtly alter the environment surrounding asparagine are shown in Table 6-3.

Table 6-3 Glycosyl acceptor properties of tripeptides containing sterically modified asparagine residues.

peptide	K_m (mM)	V_{max} dpm min^{-1}	V_{max}/K_m	K_i (mM)
Bz-NLT-NHMe	0.47	4208	8953	-
Bz-DLT-NHMe	-	-	-	>10
Bz-D(OγMe)LT-NHMe	-	-	-	>10
Bz-N(NγMe)LT-NHMe	-	-	-	>10

It has long been known that -Gln-Xaa-Thr- sites are not recognized as acceptor sites despite the fact that glutamine differs from asparagine by only one methylene unit. This may be because glutamine does not form a stable short-range side chain to main chain interaction similar to the Asx-turn. Consequently although -Gln-Xaa-Thr- sequences possess identical functionality to the marker sequence they cannot adopt a similar conformation.

Moreover, -Asp-Xaa-Thr- sites also are not recognized as acceptor sites. Bz-DLT-NHMe was neither an acceptor nor an inhibitor in the transferase assay although it contains the "active" hydroxy amino acid, Thr, and is expected to adopt an Asx-turn conformation. This suggests that a full negative charge is not tolerated by the active site. Accordingly, the mechanism suggested by Bause and Legler (Figure 6-6) in which a full negative charge is induced at nitrogen by the abstraction of a β-carboxyamido proton is unlikely since Bz-DLT-NHMe mimics the charge distribution expected for this transition state and thus should inhibit the transferase.

Bz-D(OγMe)LT-NHMe and Bz-N(NγMe)LT-NHMe likewise are nonbinders in the transferase assay although they also contain threonine and also are expected to adopt an Asx-turn conformation. This was surprising because neither modication was expected to deleteriously affect either the conformation of the peptide or the hydrogen bonding interaction between the asparaginyl carbonyl oxygen and the hydroxyl proton. However, both modifications did drastically alter the proton donor ability of the carboxyamido group. Thus it also appears that the proton donor ability of asparagine is essential for binding.

A mechanism consistent with these findings is shown in Figure 6-11.

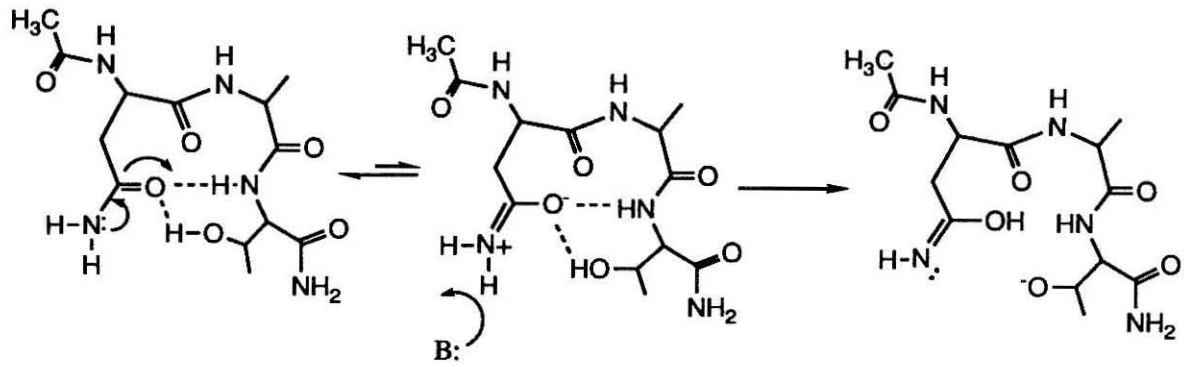


Figure 6-11 Proposed mechanism of N-linked glycosylation consistent both with conformational and mechanistic data.

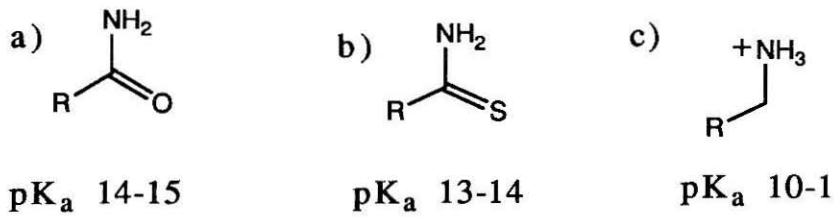
The side chains of asparagine and the hydroxy amino acid are brought into proximity by formation of a stabilized Asx-turn. Both the hydroxyl group and the backbone amide of the hydroxy amino acid interact with the carbonyl oxygen and increase the acid dissociation constant of the amide protons by stabilizing the alternate tautomer, the imidate. A general base at the active site can then abstract a proton from the carboxyamido nitrogen, thereby generating electron density at nitrogen and providing a means for nucleophilic attack at the C_1 sugar of the lipid-linked oligosaccharide. This mechanism relies on the peptide backbone either 1) to catalyze formation of the alternate tautomer or 2) to stabilize the alternate tautomer, thus any modifications which promote the formation of the imidate should result in more efficient acceptors.

The following modifications affected the proton donor behavior of the carboxyamido nitrogen (Table 6-4).

Table 6-4 Glycosyl acceptor properties of tripeptides containing asparagine analogs with altered proton donor behavior.

peptide	K_m (mM)	V_{max} dpm min^{-1}	V_{max}/K_m	K_i (mM)
Bz-NLT-NHMe	0.47	4208	8953	-
Bz-Amb-LT-NHMe	-	-	-	~1
Bz-Tan-LT-NHMe	0.26	344	1323	-

Bz-Amb-LT-NHMe competitively inhibits the transferase and has a K_i comparable to the K_m of the standard substrate, Bz-NLT-NHMe. The proton donor in Bz-Amb-LT-NHMe has a lower pK_a than the donor in the standard substrate, so it more readily donates a proton (Figure 6-12).



pK_a 14-15 pK_a 13-14 pK_a 10-12

Figure 6-12 pKa values of a) Asn, b) Tan, and c) Amb.

Amb lacks a carbonyl oxygen so it cannot form the necessary hydrogen bonding interactions to stabilize the recognized conformation, an Asx-turn; however it can achieve an analogous structure. Since the K_i of this inhibitor is similar to the K_m, the increase in proton donor efficiency apparently compensates for the lack of stablized Asx-turn structure.

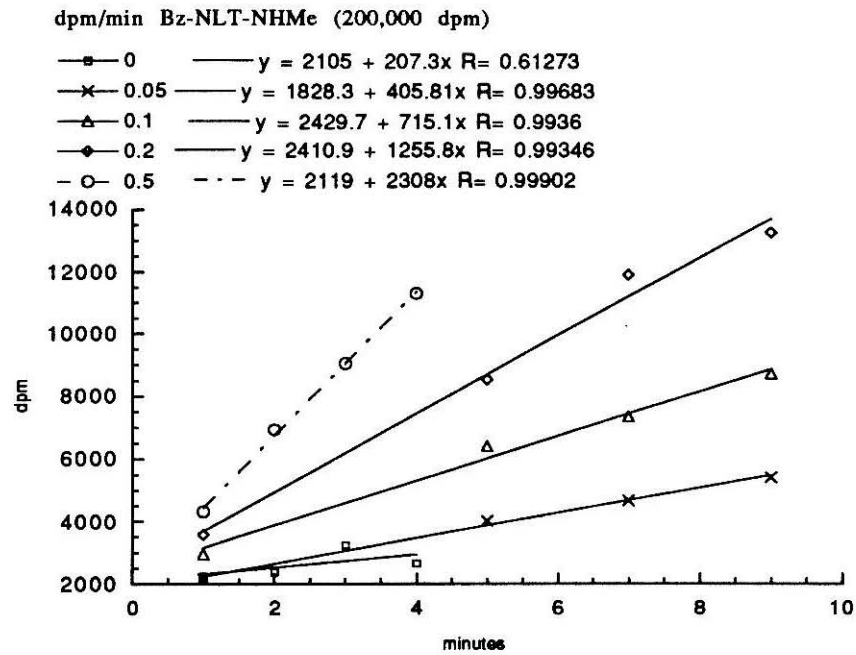
Tan is the first example of a modified asparagine analog which functions as a substrate in glycosylation. The K_m of Bz-Tan-LT-NHMe was consistently shown to be slightly lower than Bz-

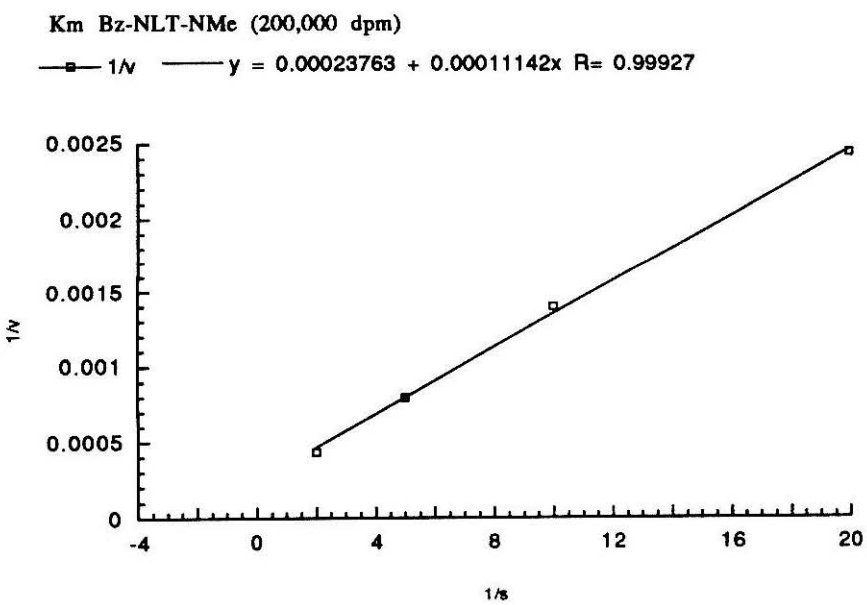
NLT-NHMe indicating that this analog actually binds more efficiently to the enzyme. Sulfur is a weaker base than oxygen which causes the thioimidate tautomer to be less populated than the imidate tautomer. Consequently thioamides are stronger acids than carboxyamides and for this reason may bind more efficiently. The lower Vmax of Bz-Tan-LT-NHMe can also be related to the weaker basicity of sulfur. Since the thioimidate tautomer is less populated, the interconversion happens less frequently and thus the nucleophilic species is not generated as often.

CONCLUSIONS

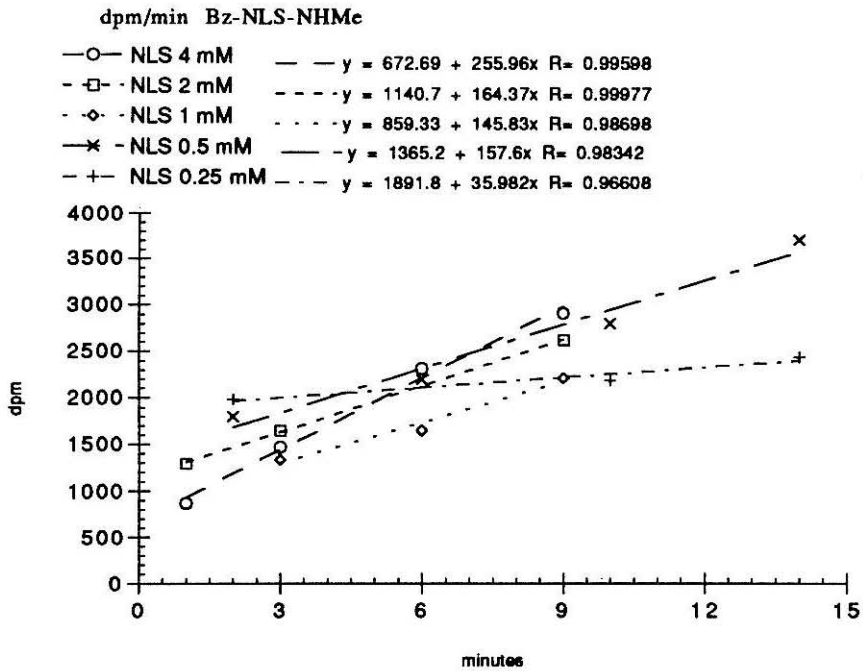
The mechanism described above presents an ideal way for specificity to be generated in the glycosylation reaction despite the presence of seemingly more reactive groups within the protein. The backbone of the polypeptide acceptor not only forms a recognized conformation but it also catalyzes the formation of a nucleophilic species. The marker sequence must adopt an Asx-turn conformation in order to function as an acceptor for oligosaccharyl transferase. This conformation enhances the reactivity of the asparagyl side chain carboxyamido group by promoting the formation of an imidate intermediate. Thus the specificity observed in N-glycosylation is generated by the conformation of the peptidyl acceptor.

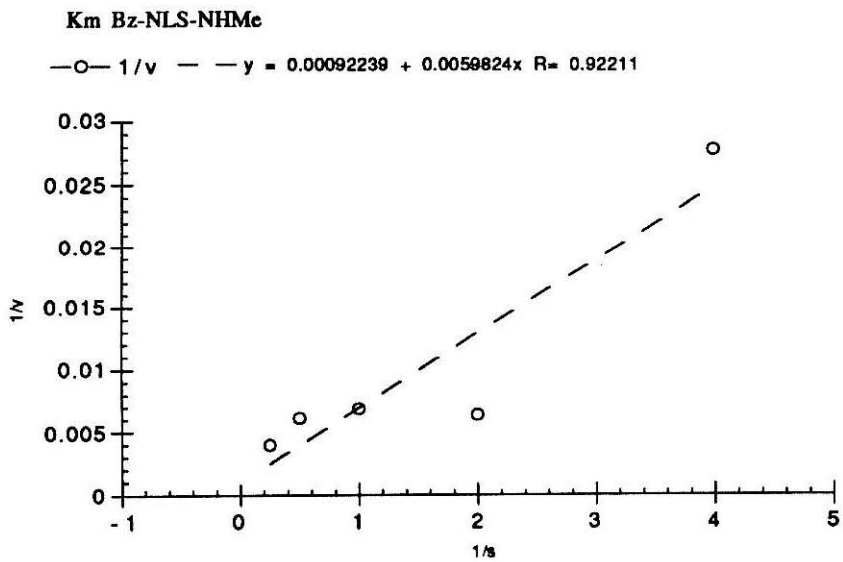
Appendix 6-1 Enzyme kinetic data for Bz-NLT-NHMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



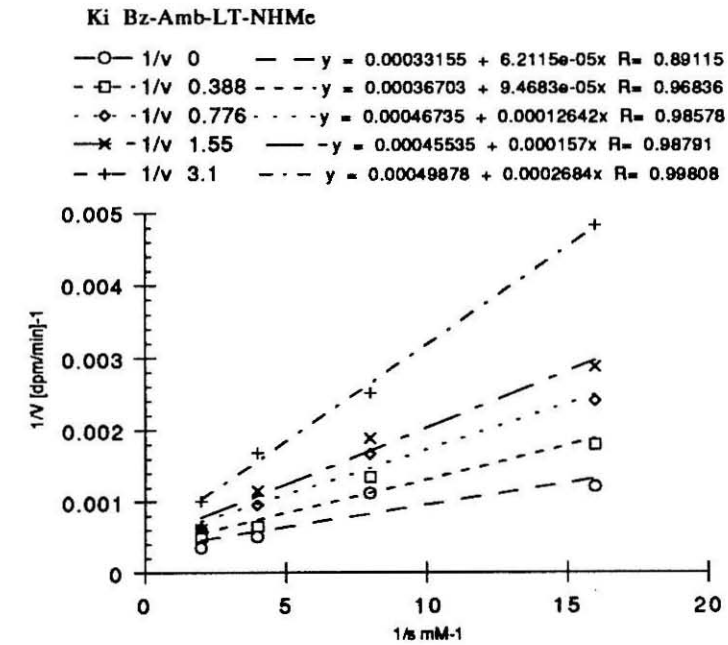


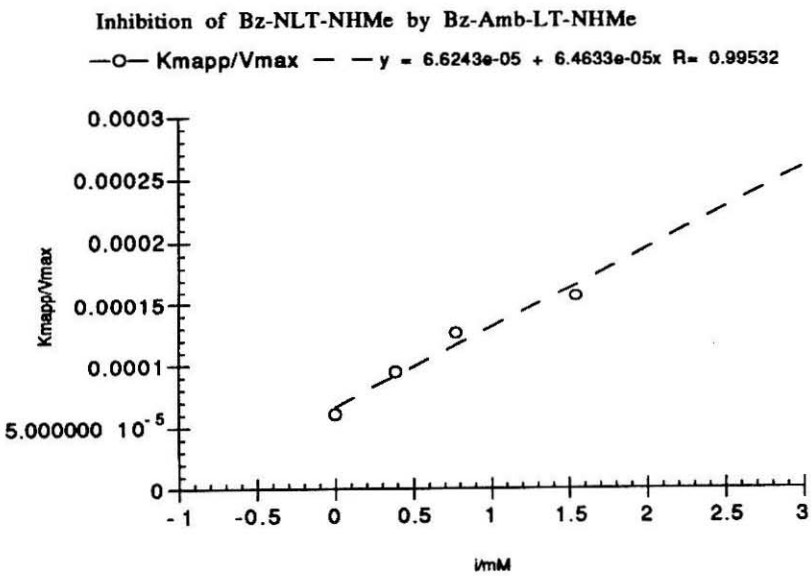
Appendix 6-2 Enzyme kinetic data for Bz-NLS-NHMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)



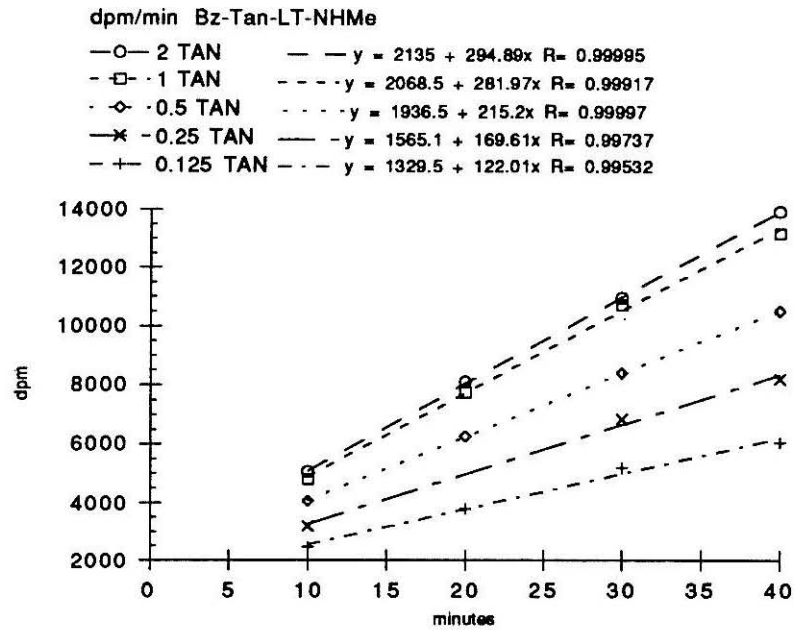


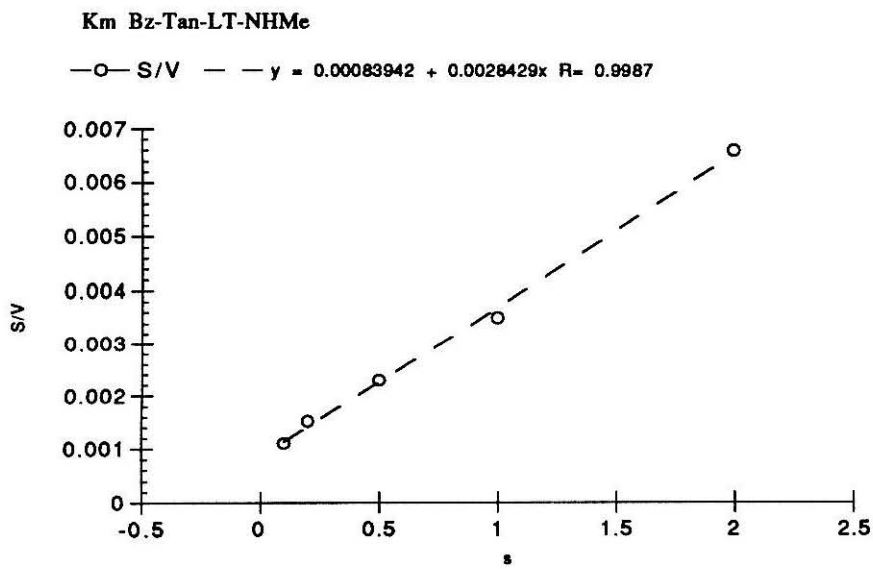
Appendix 6-3 Enzyme kinetic data for Bz-Amb-LT-NHMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)





Appendix 6-4 Enzyme kinetic data for Bz-Tan-LT-NHMe (200,000 dpm Dol-P-P-GlcNAc-[^{3}H]GlcNAc)





1. Blake, C.C.F.; Koenig, D.F.; Mair, G.A.; North, A.C.T.; Phillips, D.C.; Sharma, V.R. "Stucture of hen egg-white lysozyme," *Nature* **1965**, *206*, 757.

2. Bause, E.; Legler, G. "The Role of the Hydroxy Amino Acid in the Triplet Sequence Asn-Xaa-Thr(Ser) for the N-Glycosylation Step During Glycoprotein Biosynthesis," *Biochem. J.* **1981**, *195*, 639-644.

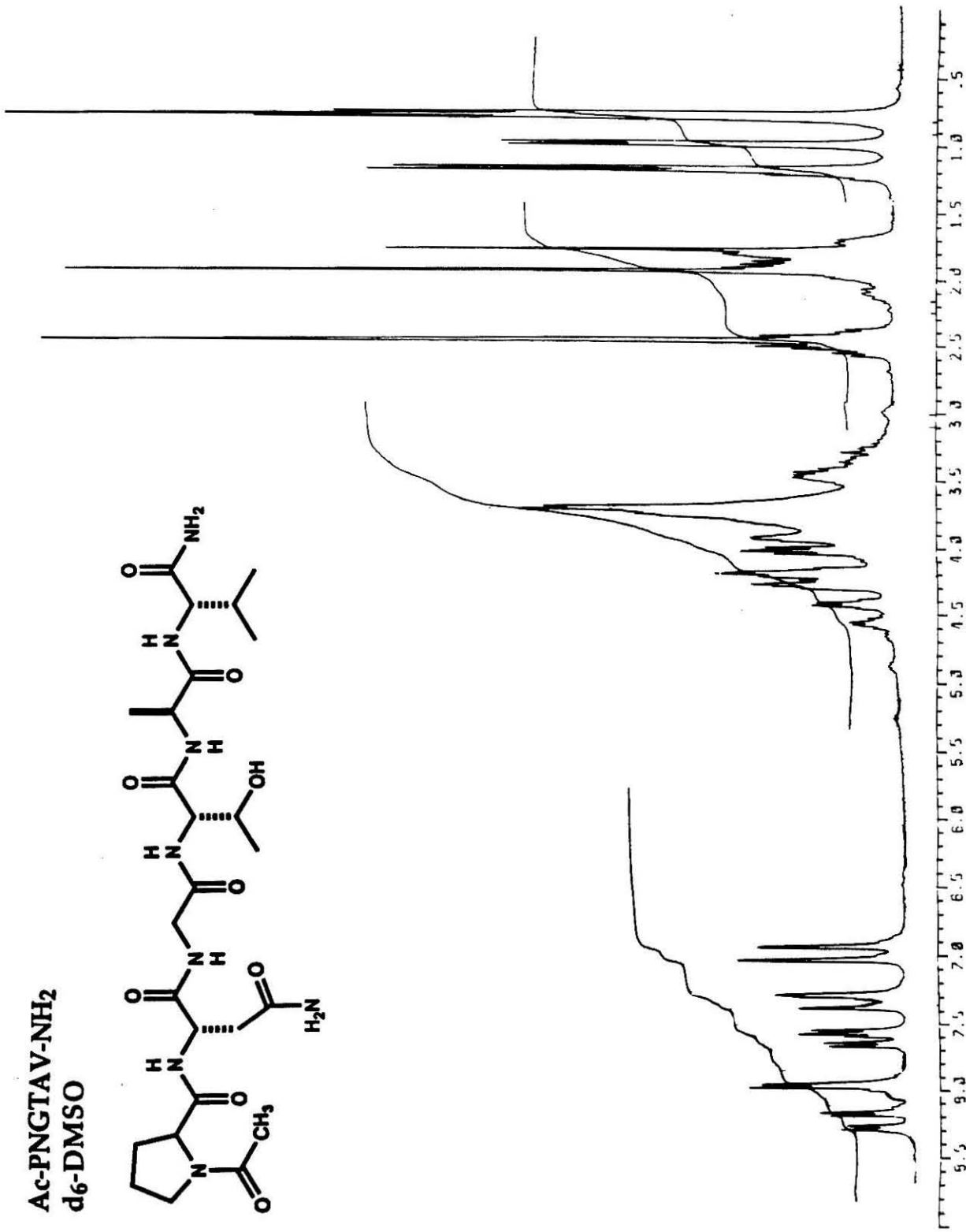
3. Marshall, R.D. "The Nature and Metabolism of the Carbohydrate-Peptide Linkages of Glycoproteins," *Biochem. Soc. Symp.* **1974**, *40*, 17-26.

4. Walsh, C. *Enzymatic Reaction Mechanisms*; W.H. Freeman & Co.: San Fransisco, 1979; pp. 138-156.

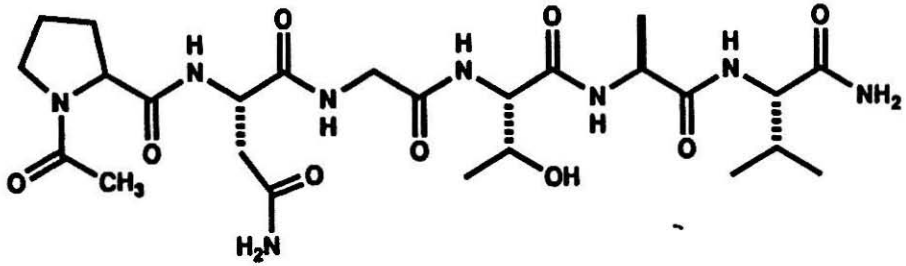
5. Clark, R.S.; Banerjee, S.; Coward, J.K. "Yeast Oligosaccharyltransferase: Glycosylation of Peptide Substrates and Chemical Characterization of the Glycopeptide Product," *J. Org. Chem.* **1990**, *55*, 6275-6285.

6. Bodansky, M. *Principles of Peptide Synthesis*, Springer-Verlag: Berlin, 1984.

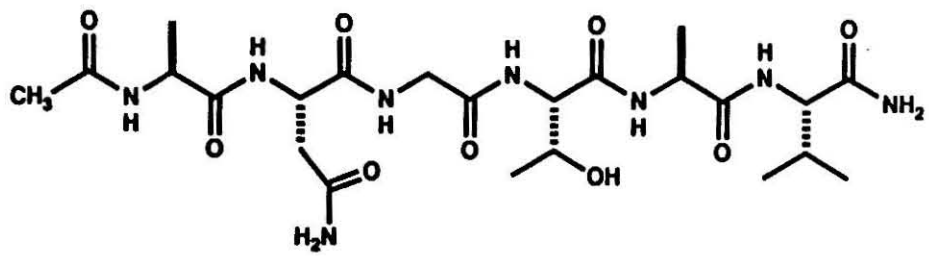
7. Prepared by Rickert, K.W., California Institute of Technology, 1991.

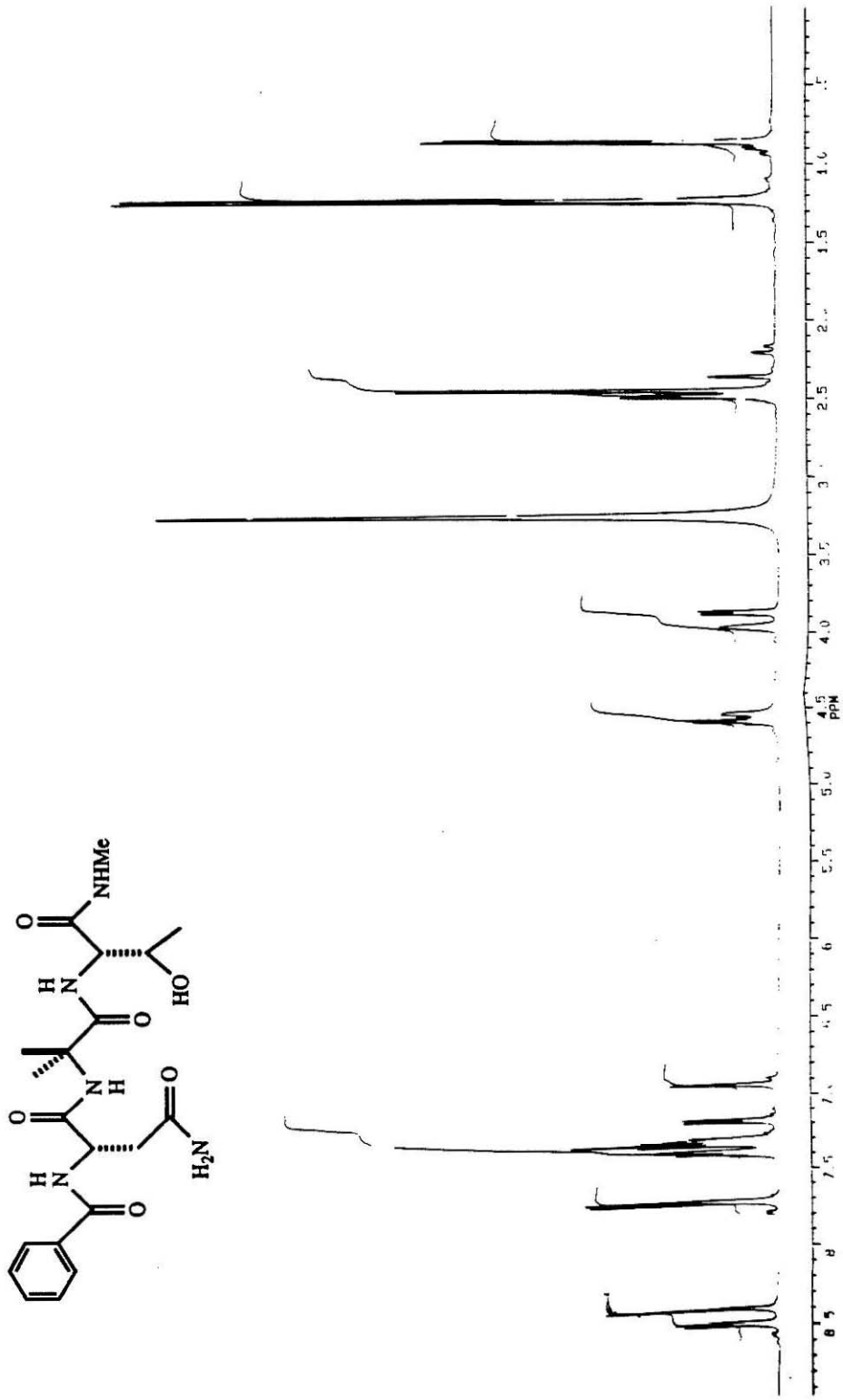

Ac-PNGTAV-NH2
d6-DMSO
PPM

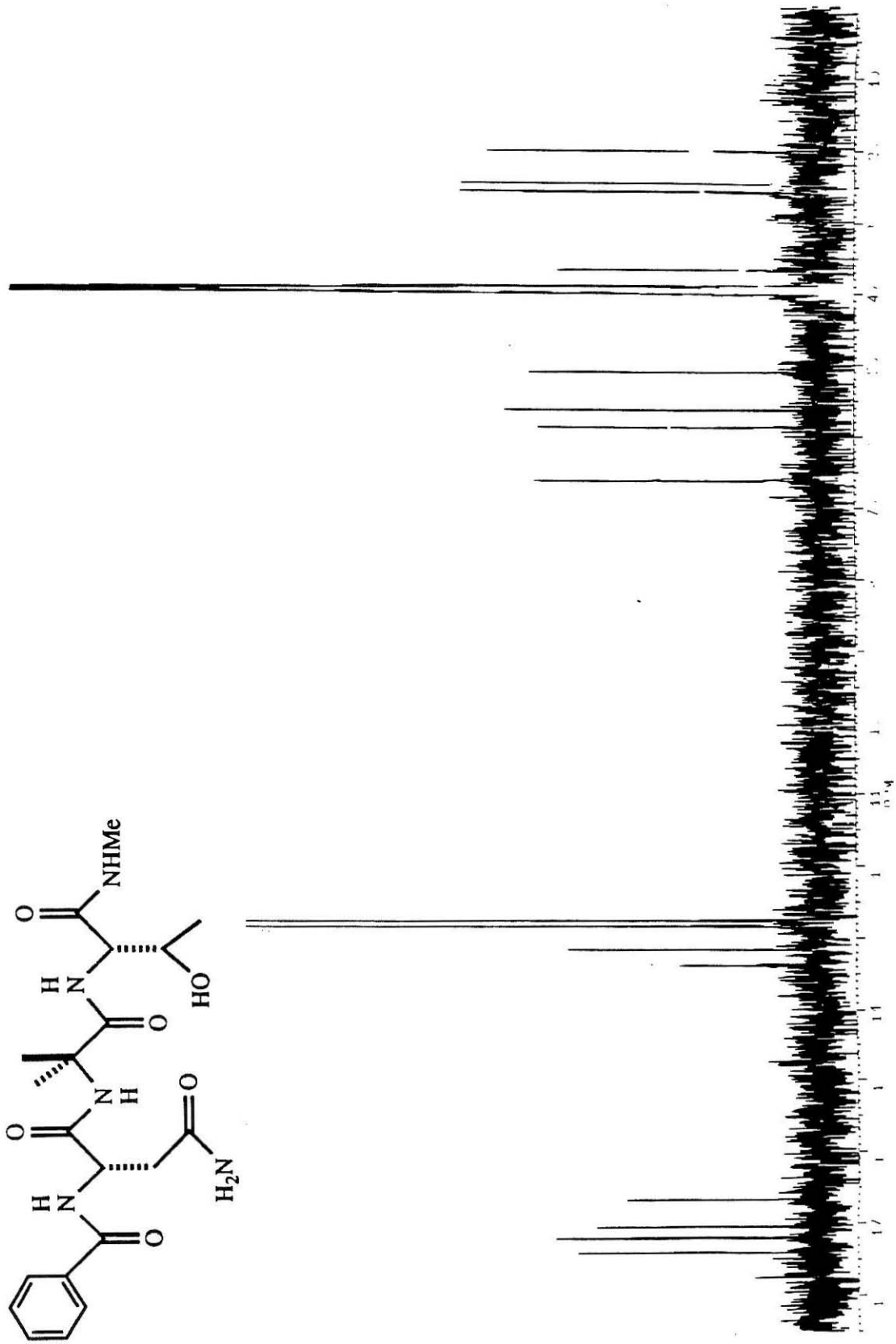
Ac-PNGTAV-NH_2
d_6-DMSO

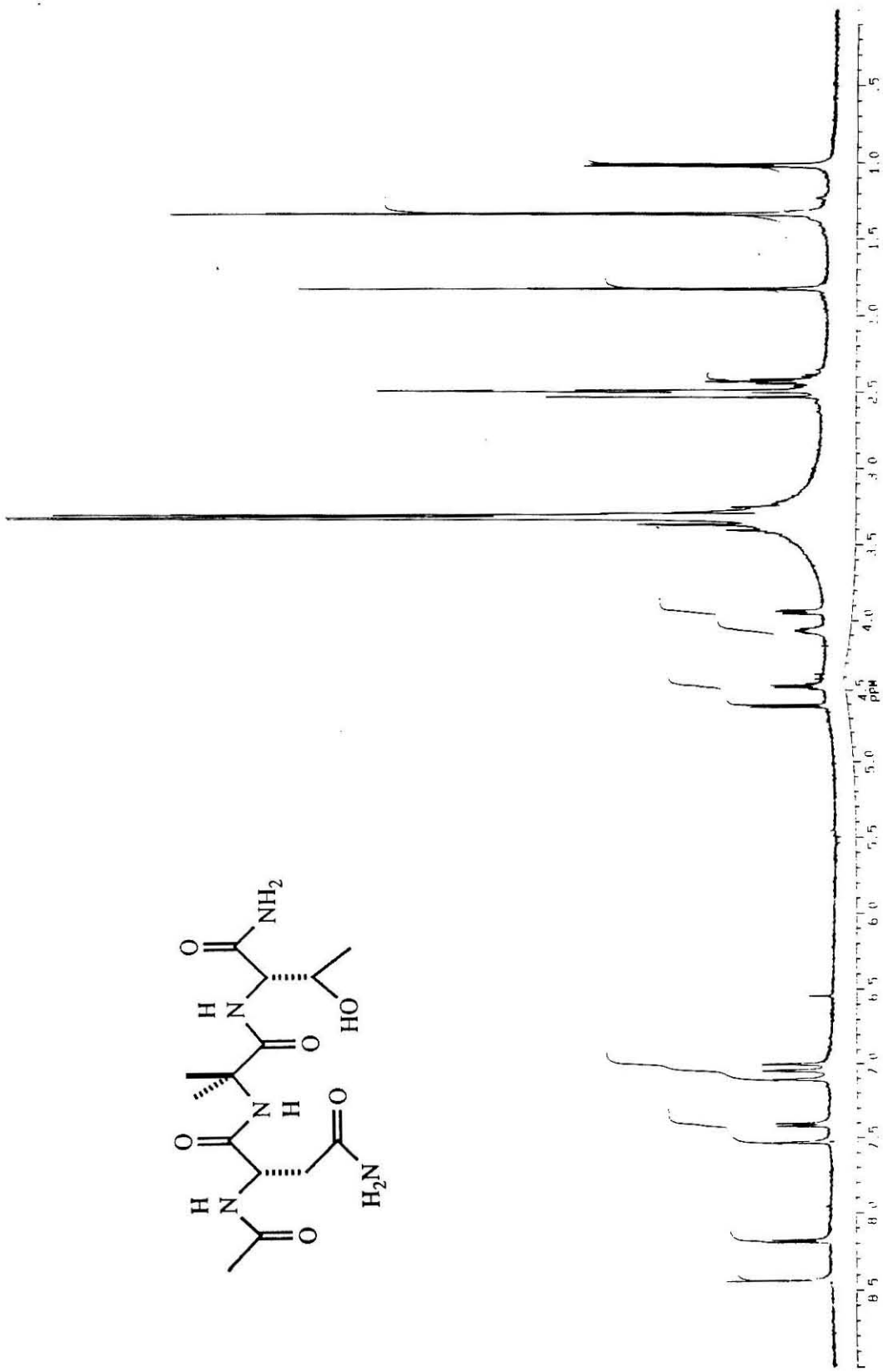


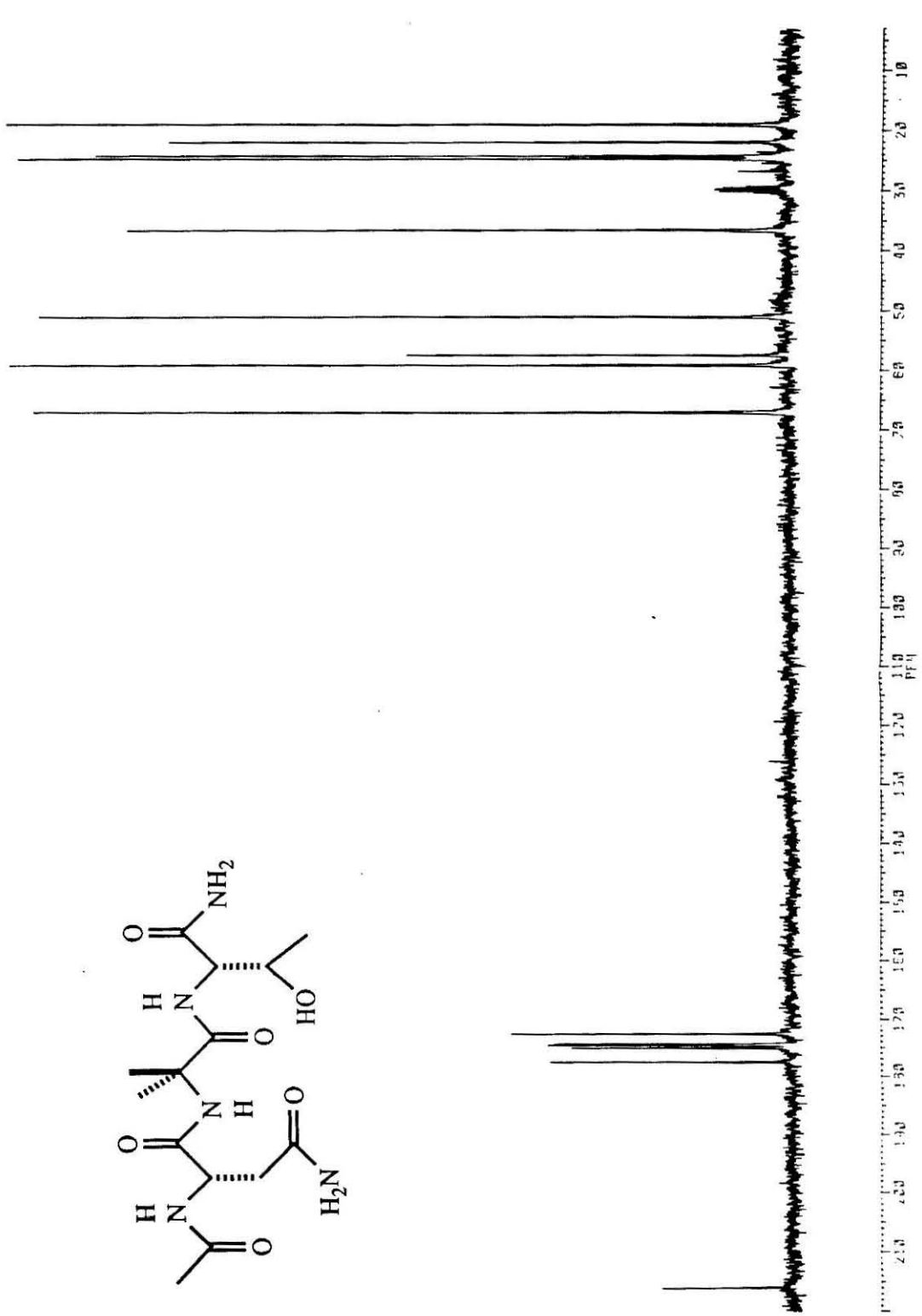
Ac-ANGTAV-NH_2
d_6-DMSO

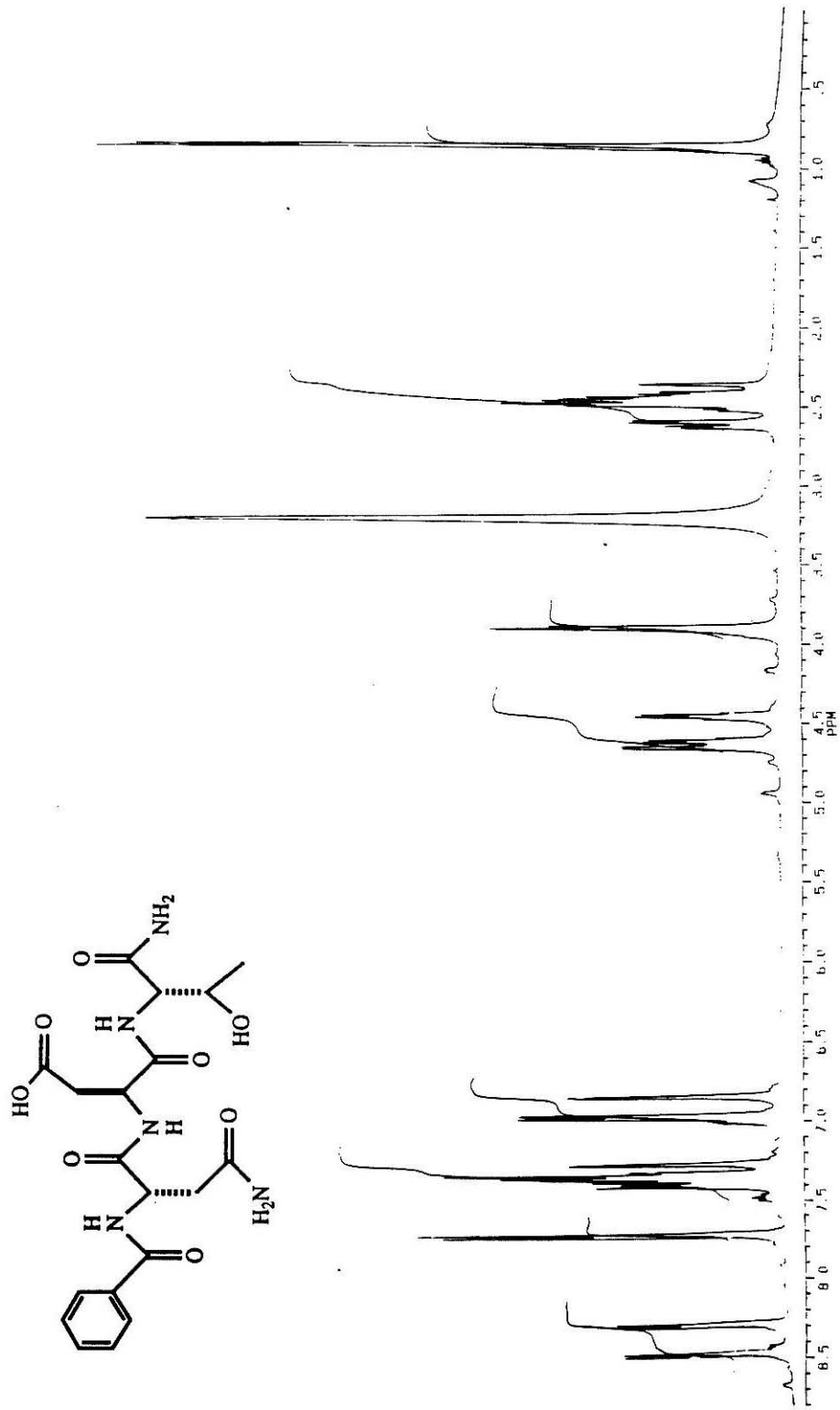


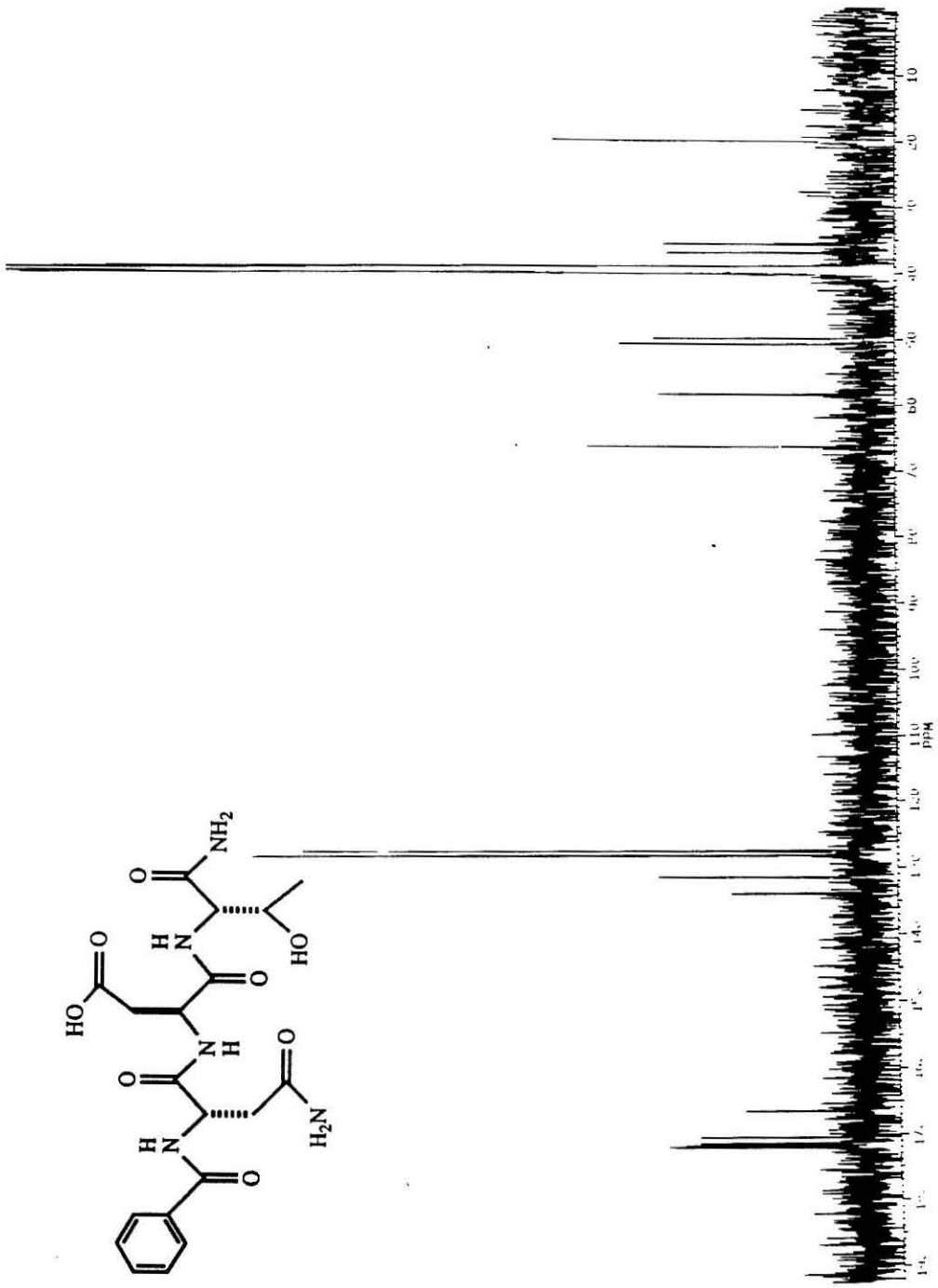

NHMe
HO
H_2N
PPM

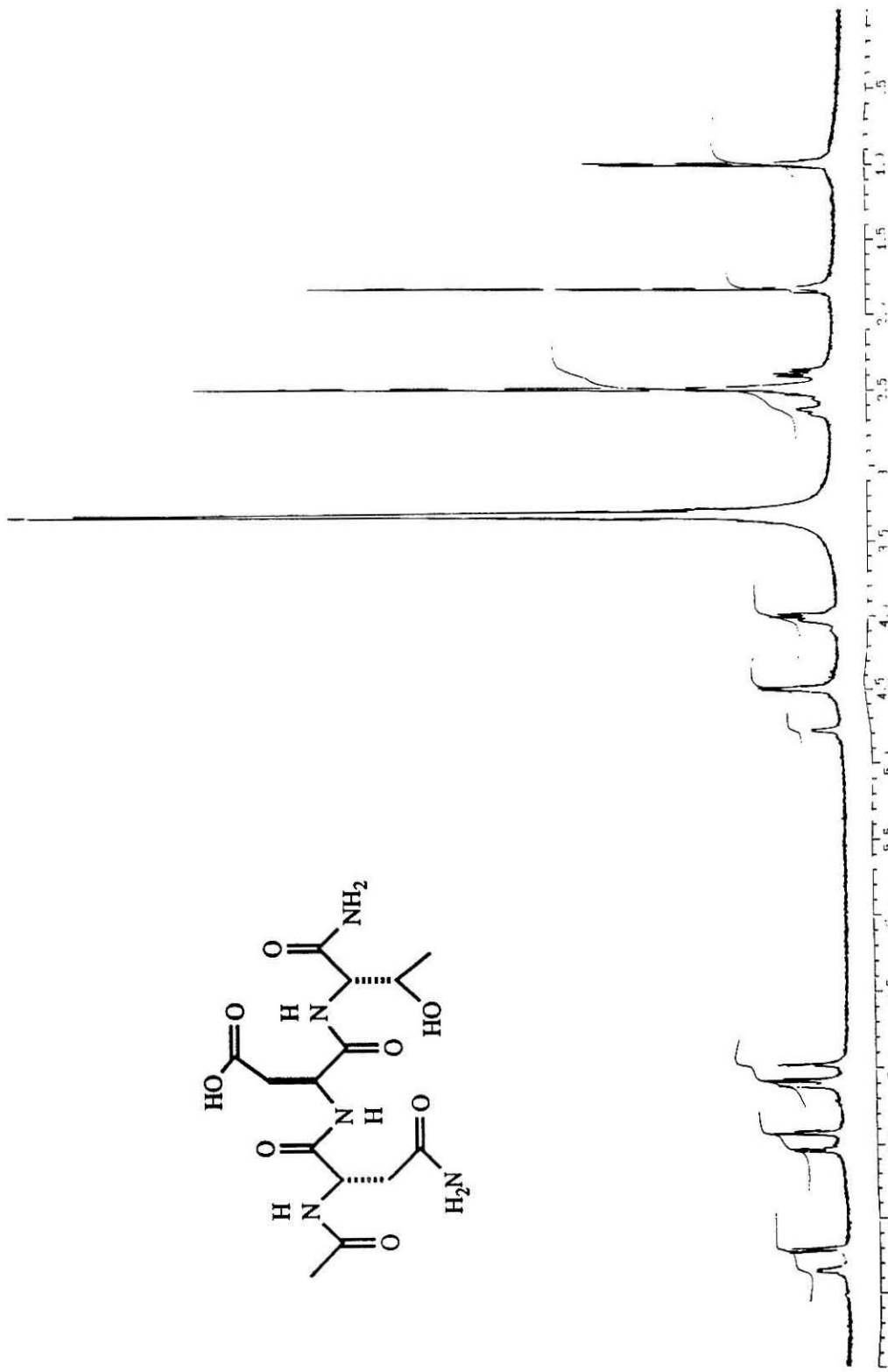

NHMe
O
H
N
HO
O
N
H
O
O
H
N
H_2N
O

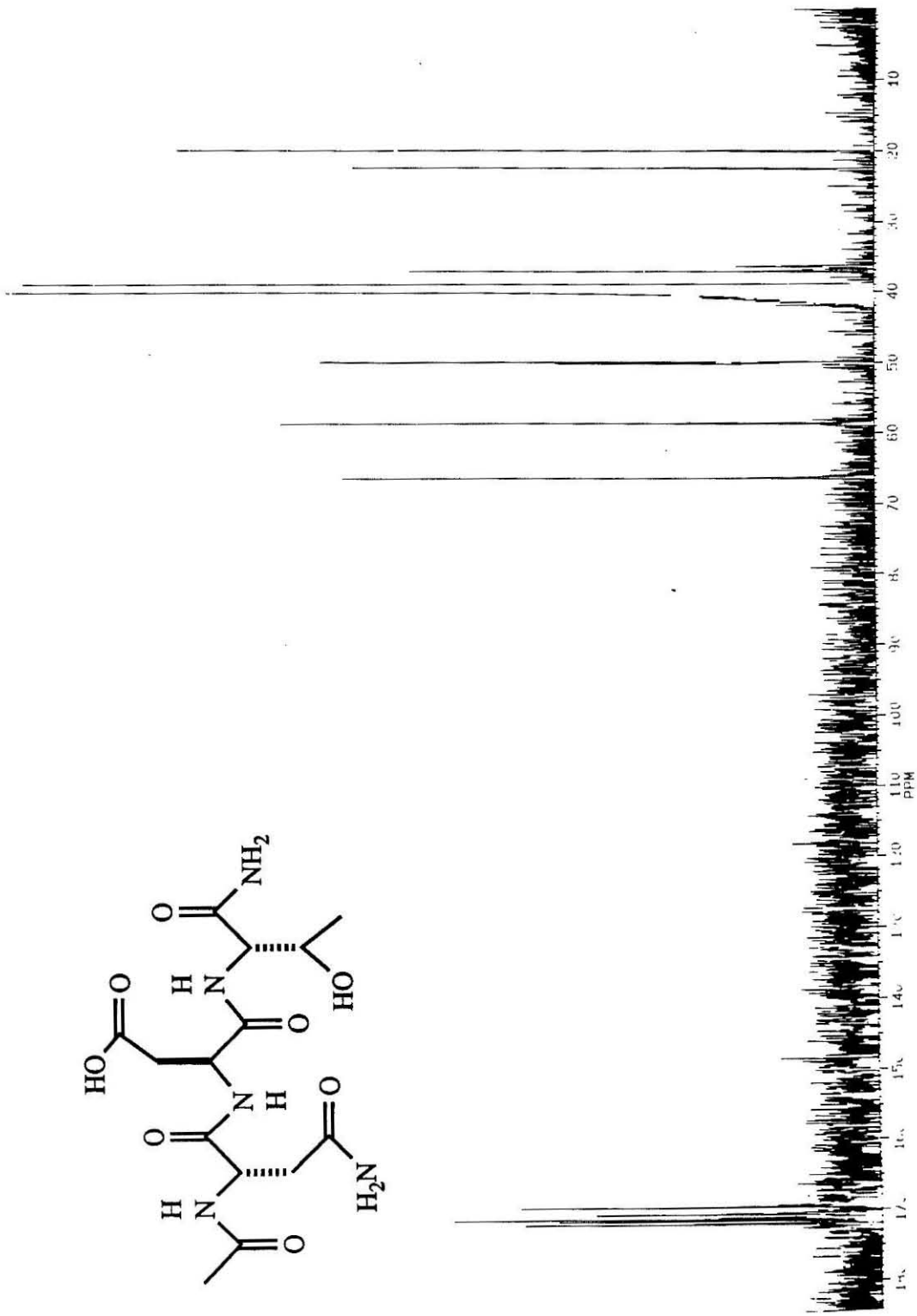

NH2
O
H
N
HO
O
N
H
O
H
N
O
H2N
O
ppm

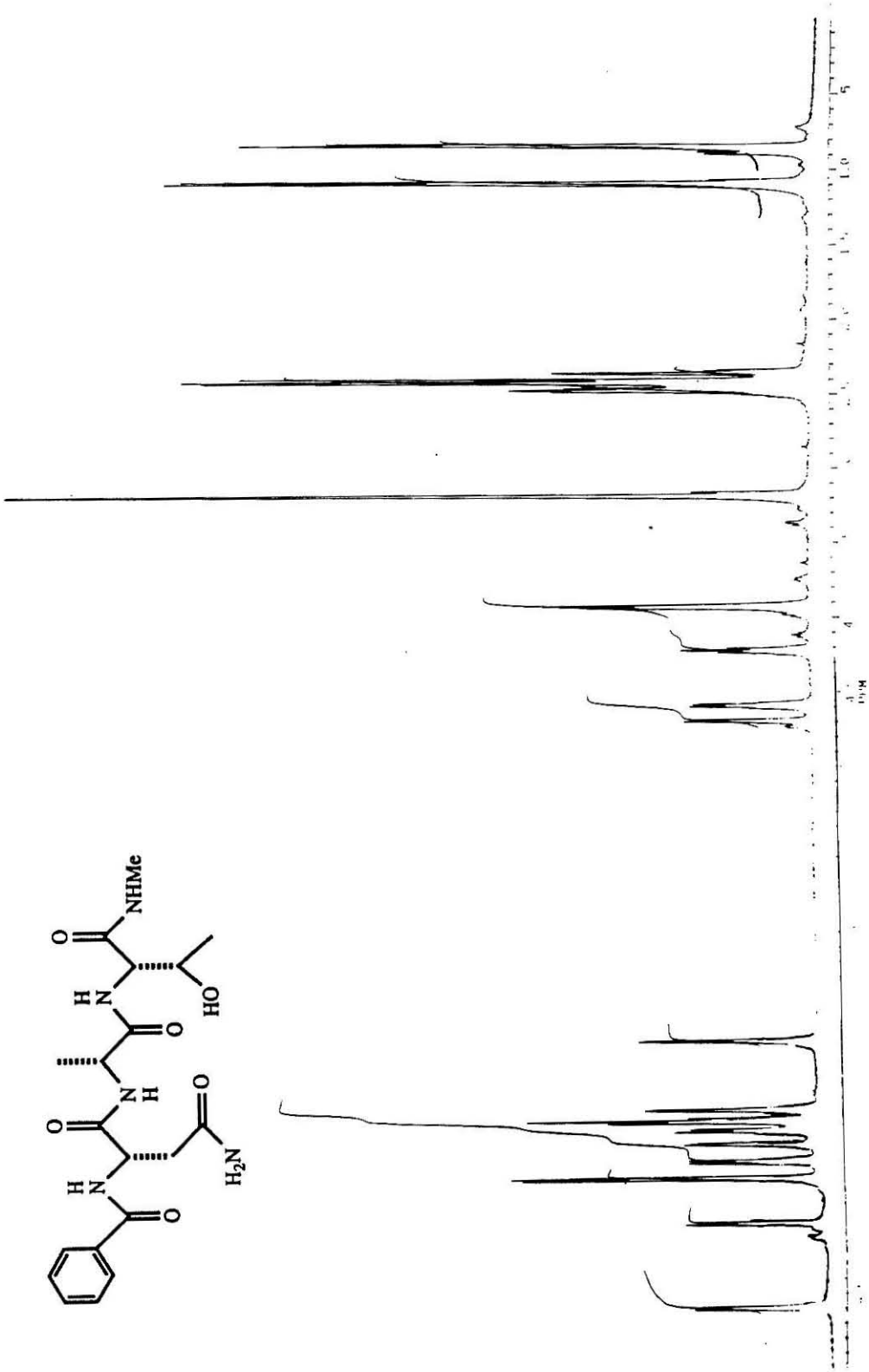

NH2
O
H
N
HO
O
N
H
O
O
H
N
H2N
O
PPM

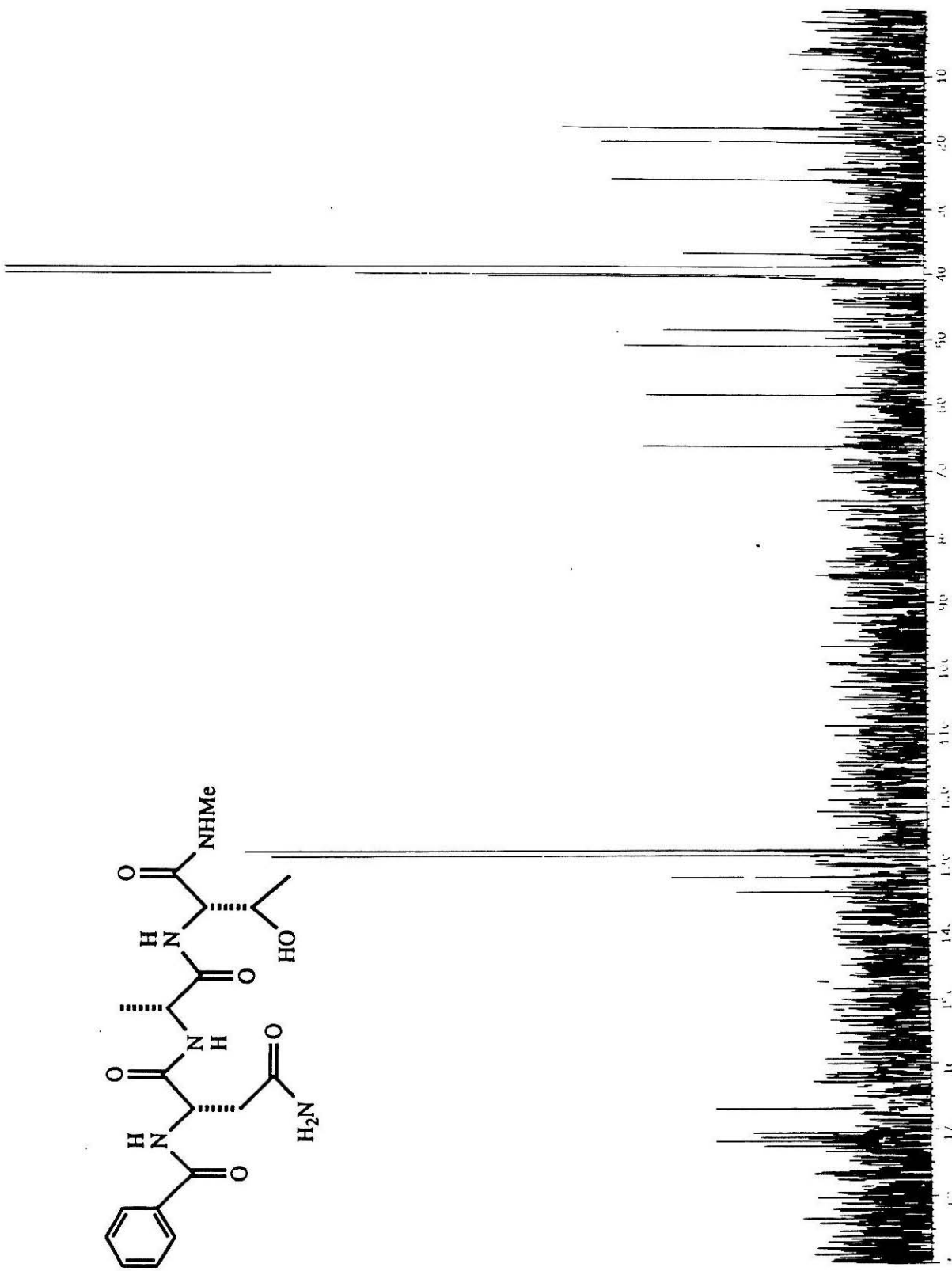

HO
O
O
H
N
O
NH2
O
HO
N
H
O
H
N
O
O
H2N
ppm

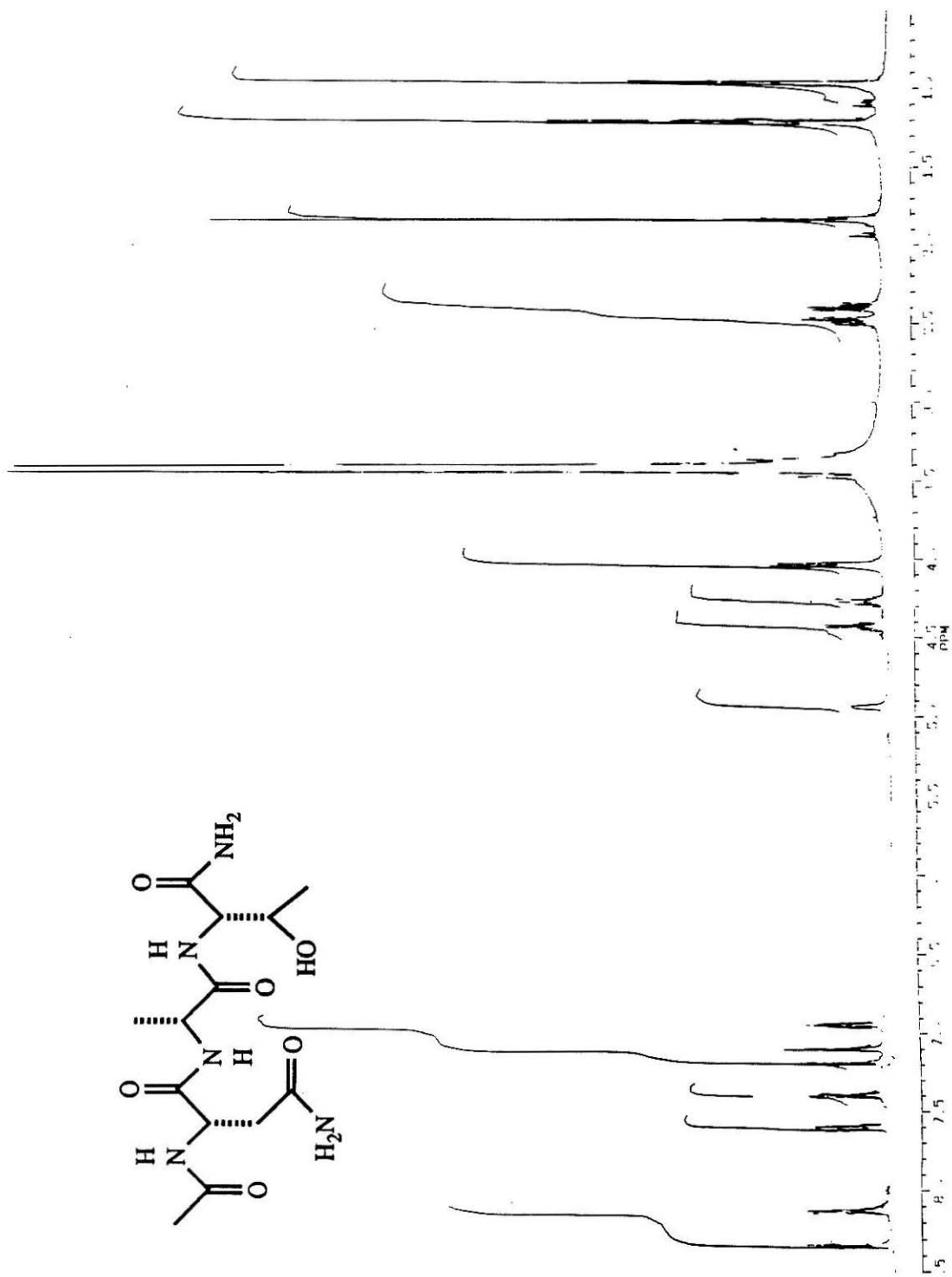

HO
O
H
N
O
O
NH2
HO
N
H
O
O
H
N
O
H2N
ppm

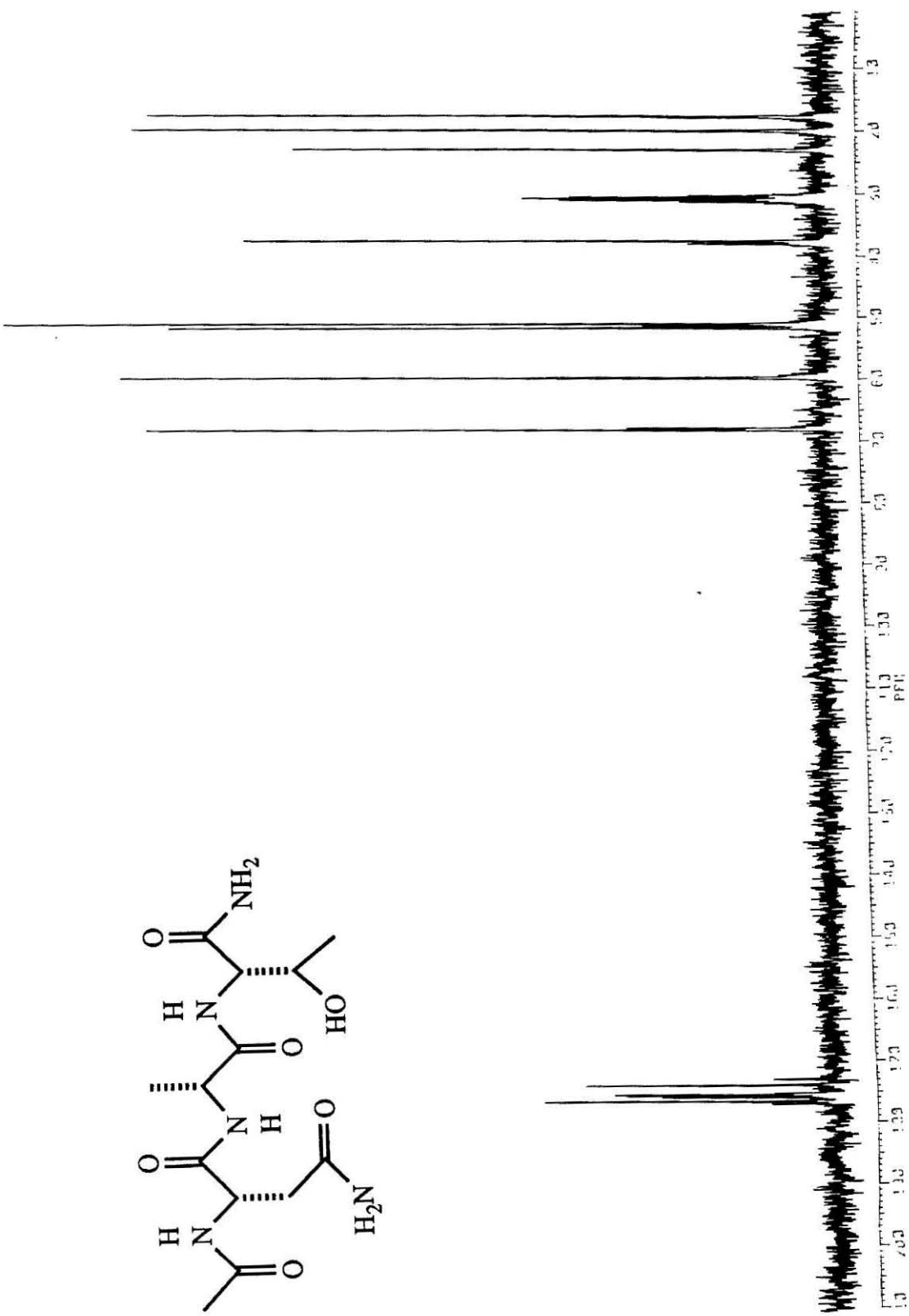

HO
O
O
H
N
O
NH2
O
HO
N
H
O
H
N
O
O
H2N
O
8.5
7.5
5.5
4.5
PPM
2.5
1.5

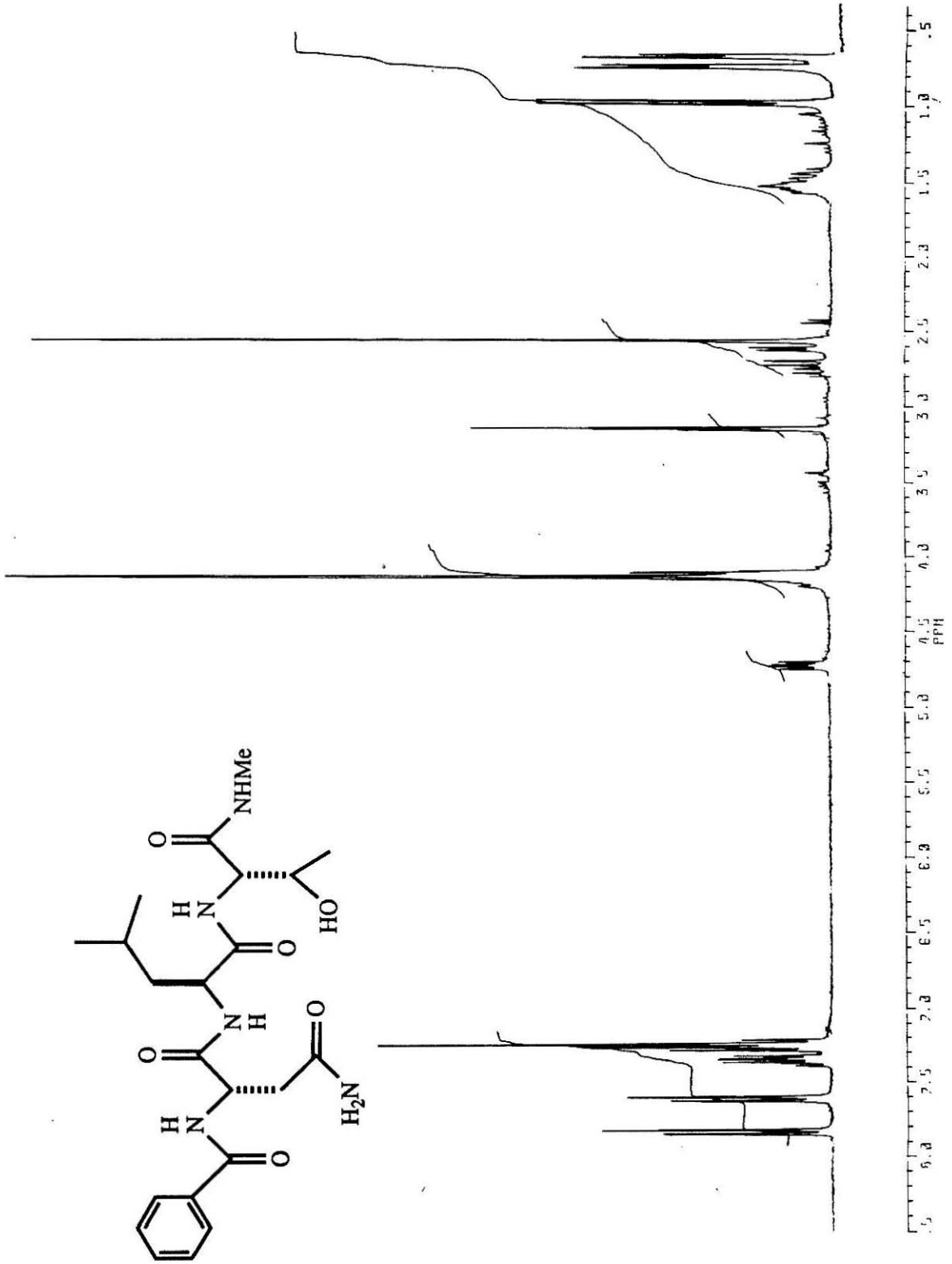

HO
O
O
H
N
O
NH2
O
HO
N
H
O
O
H
N
O
H2N
O
PPM

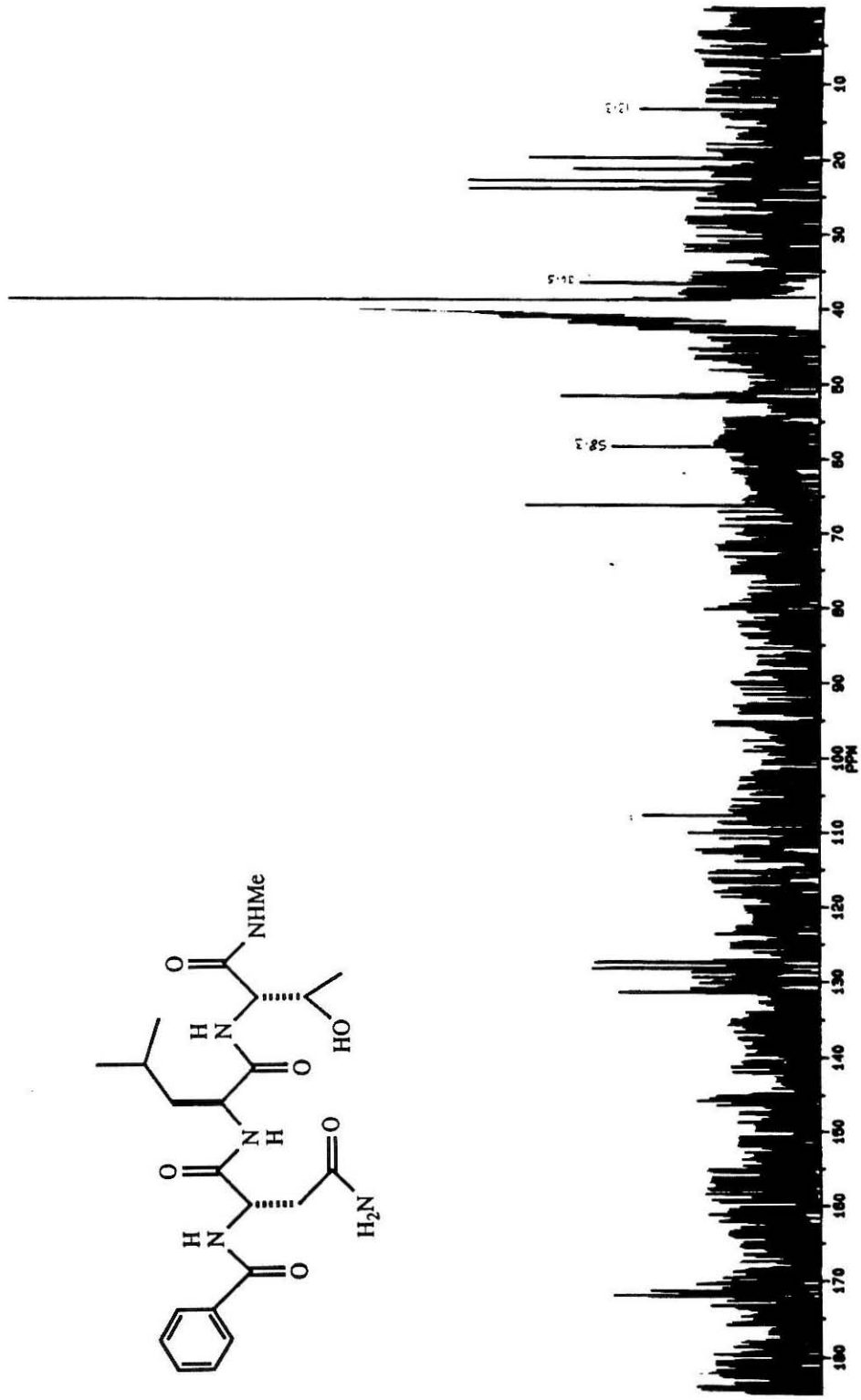

NHMe
O
H
N
HO
O
N
H
O
O
H
N
O
H2N
O
PPM

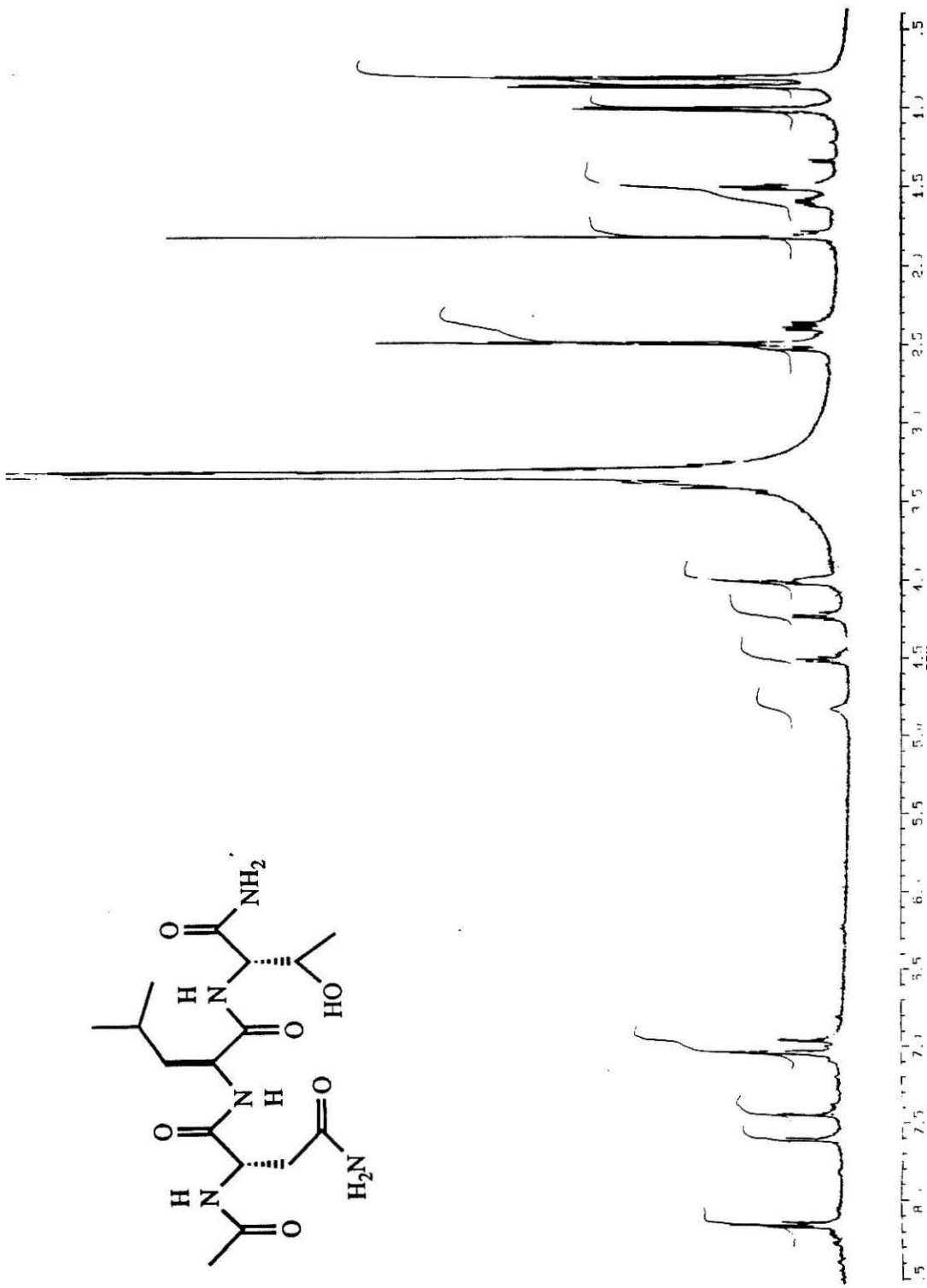

NHMe
O
H
N
HO
O
N
H
O
O
H
N
O
H_2N
O

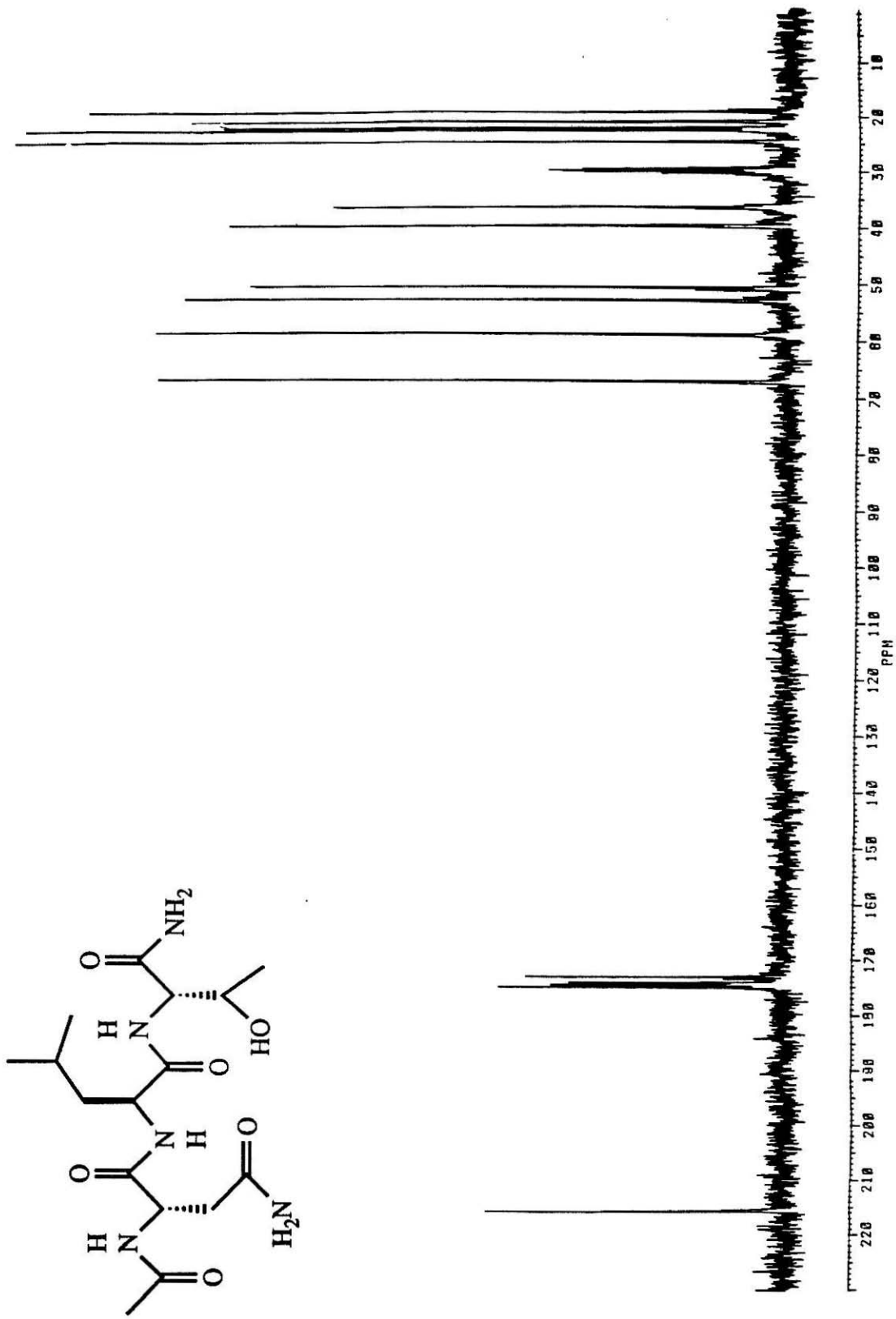

O
NH2
H
N
HO
O
N
H
O
H2N
H
N
O

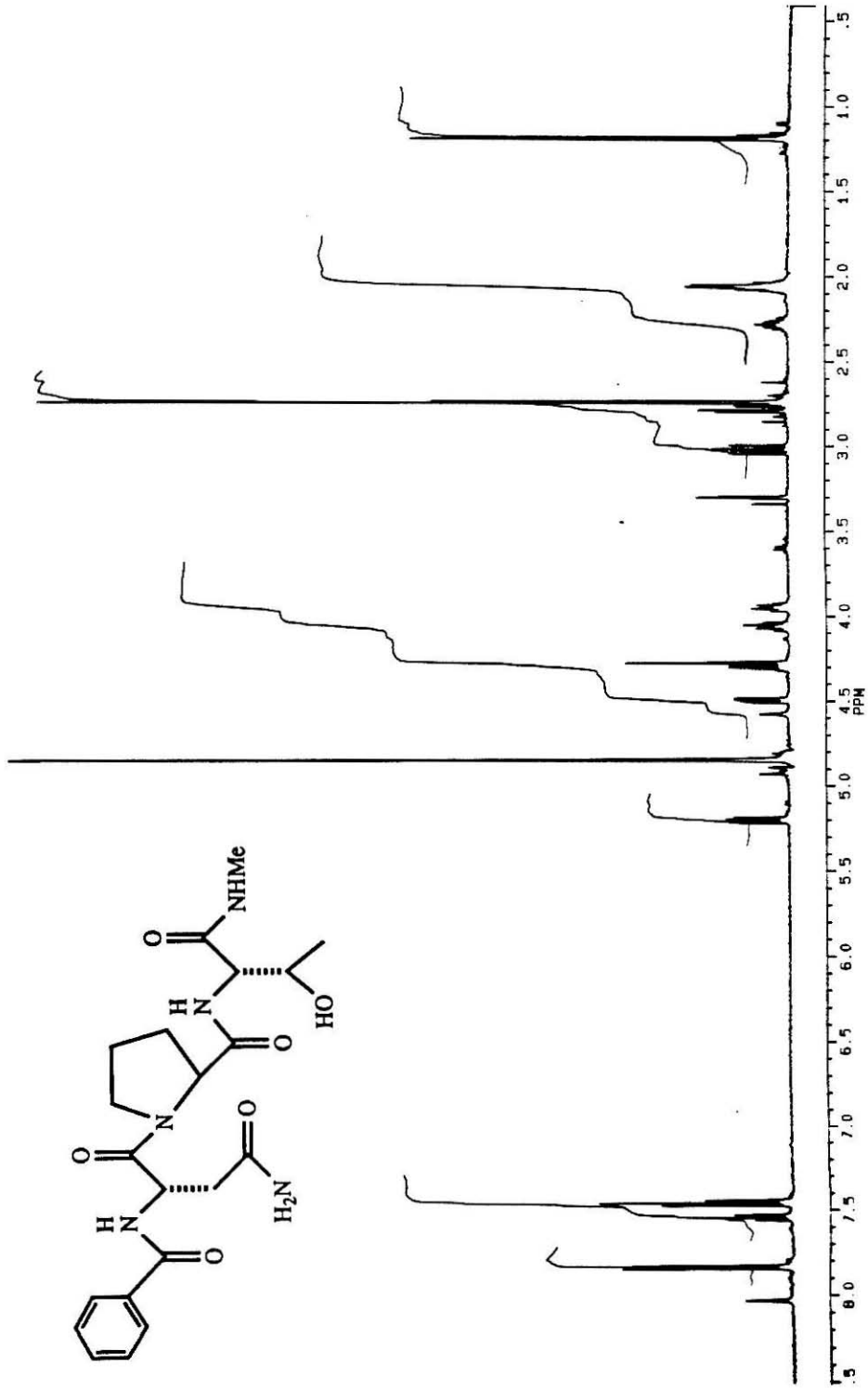


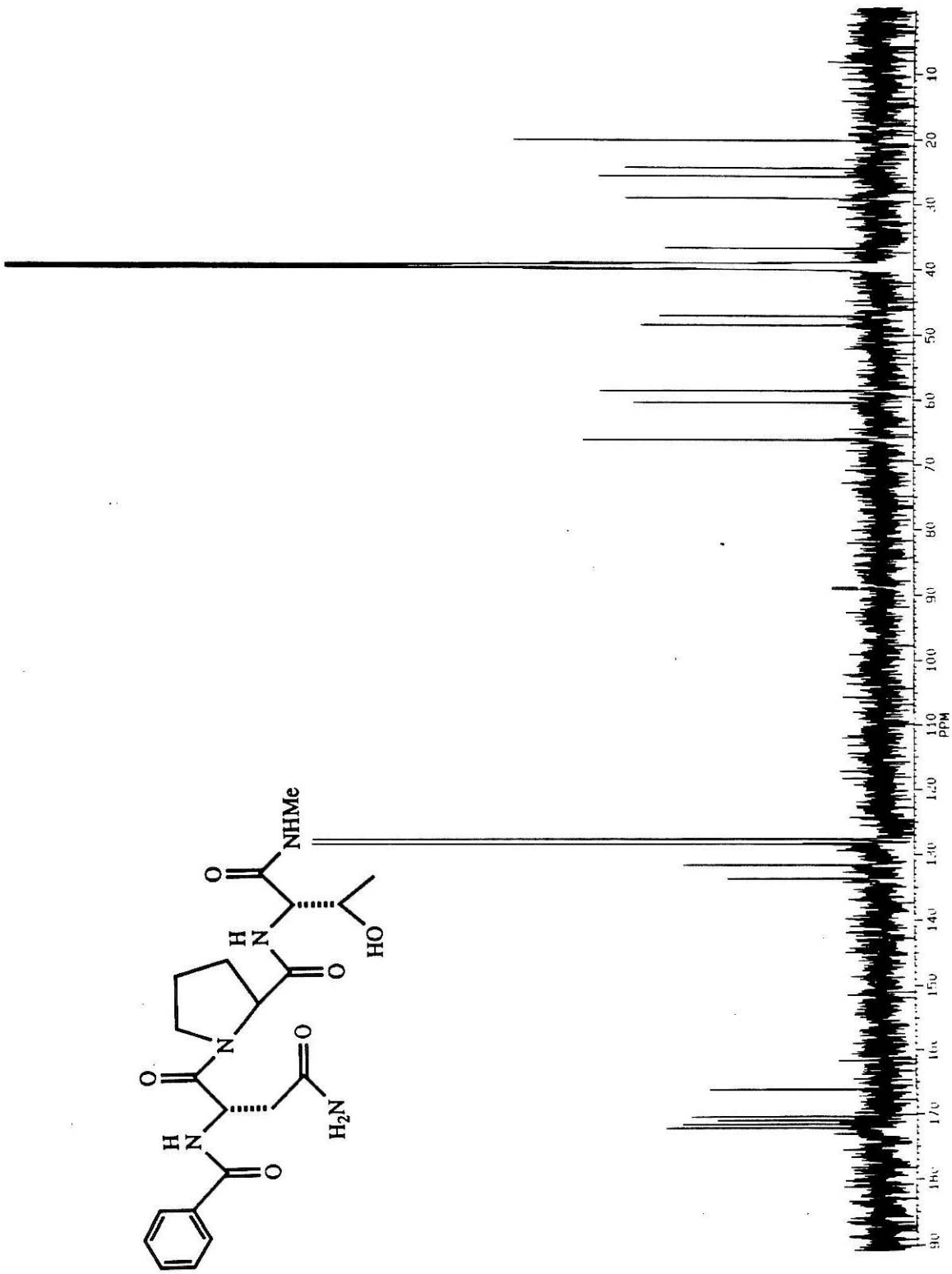

NHMe
O
H
N
HO
O
N
H
O
O
H
N
O
H2N
PPM

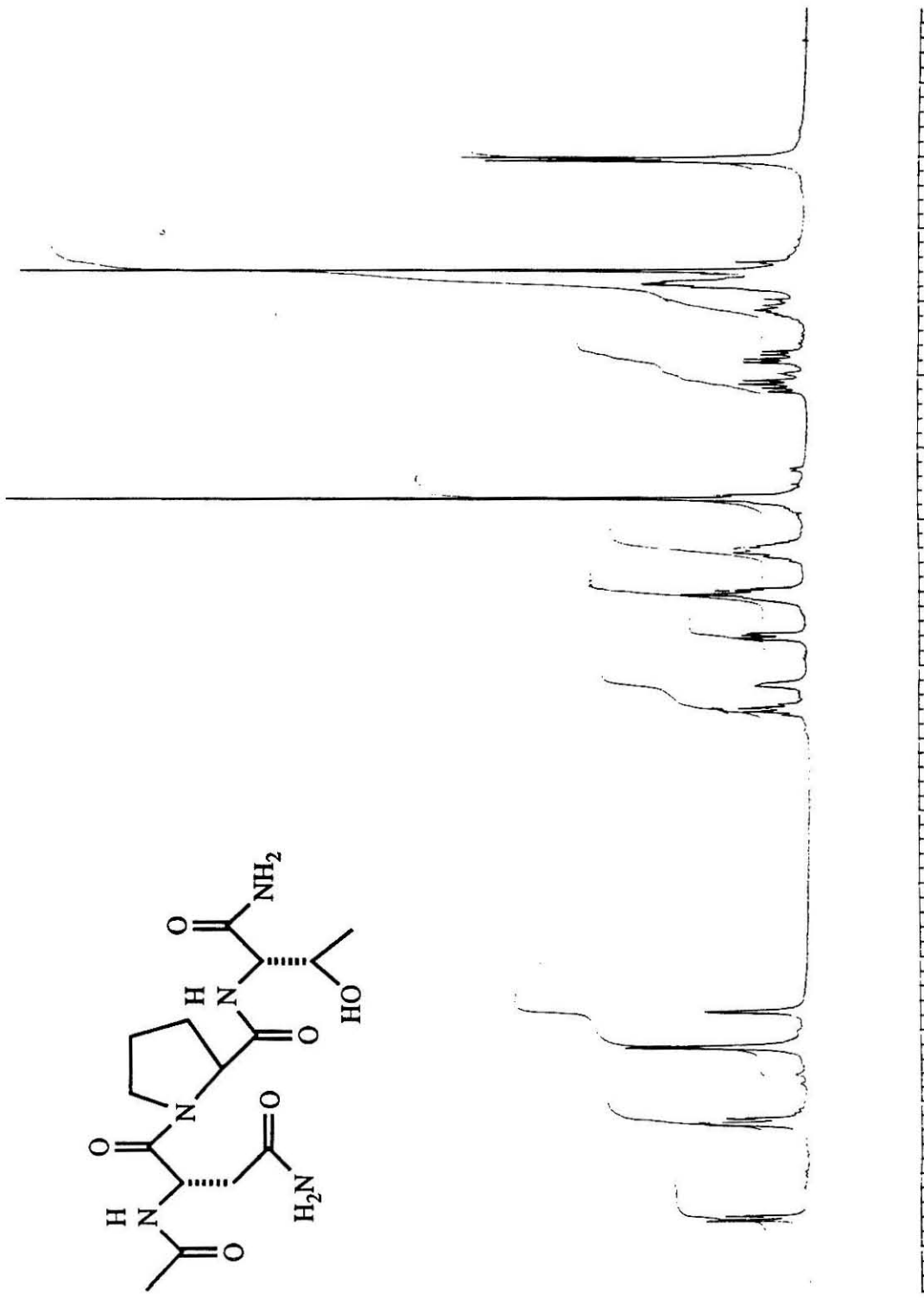

NHMe
O
H
N
HO
O
N
H
O
O
H
N
O
H_2N
58.3
36.5
12.3
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
PPM

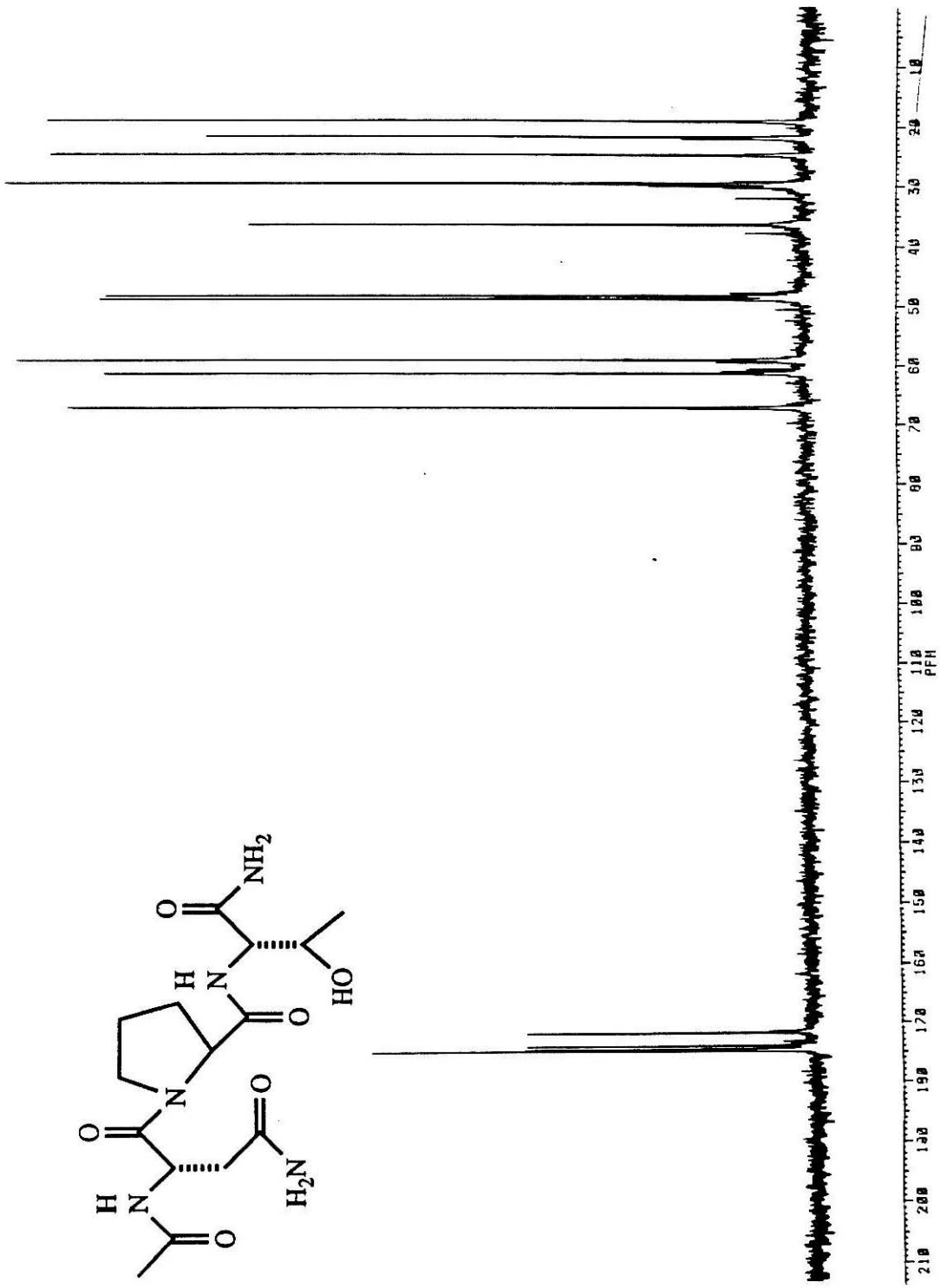

NH2
O
H
N
HO
O
N
H
O
H
N
O
H2N
O
5.5
1.5
PPM
1.5

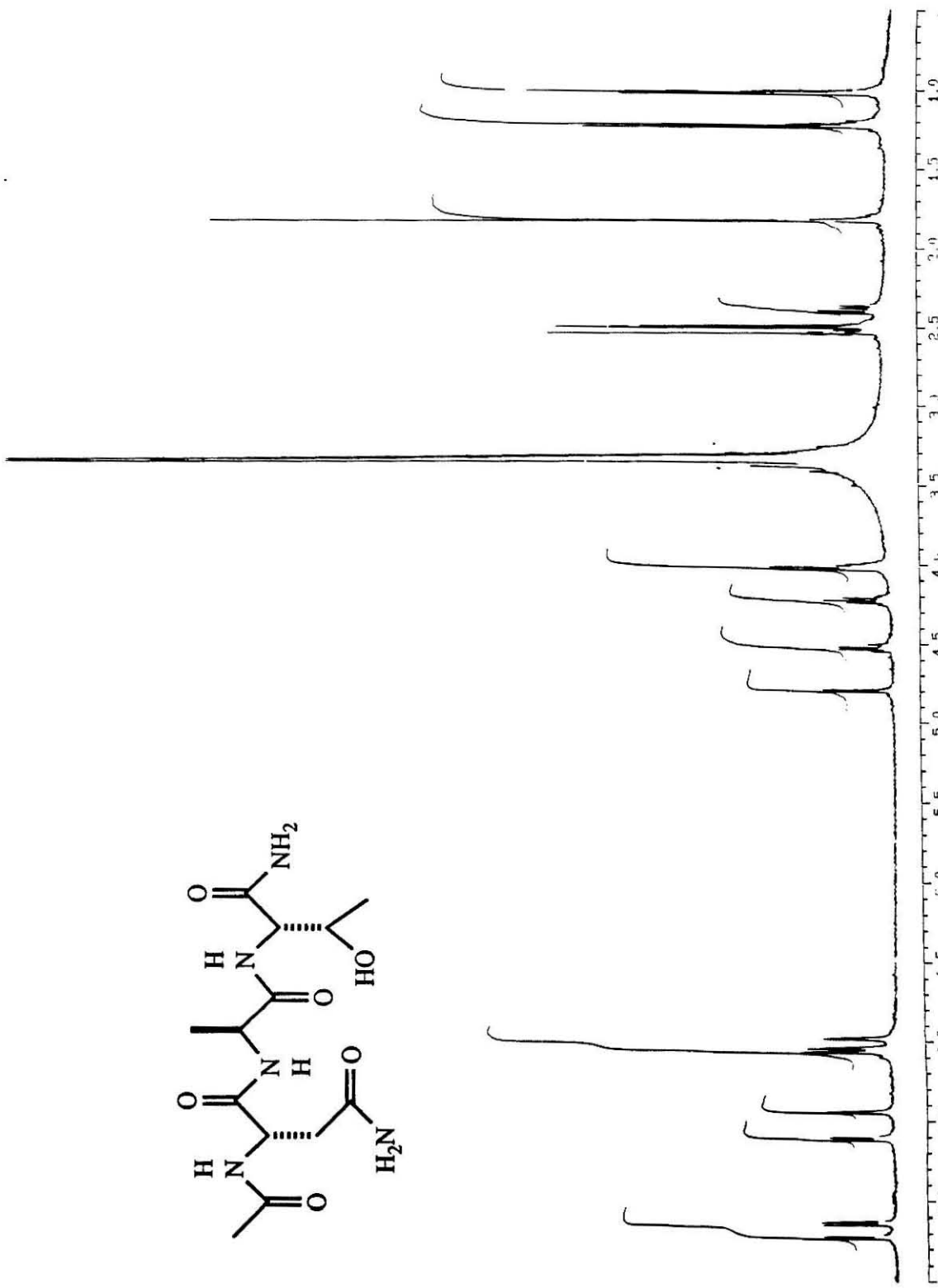

NH2
O
H
N
HO
O
N
H
O
O
H2N
H
N
O
PPM
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

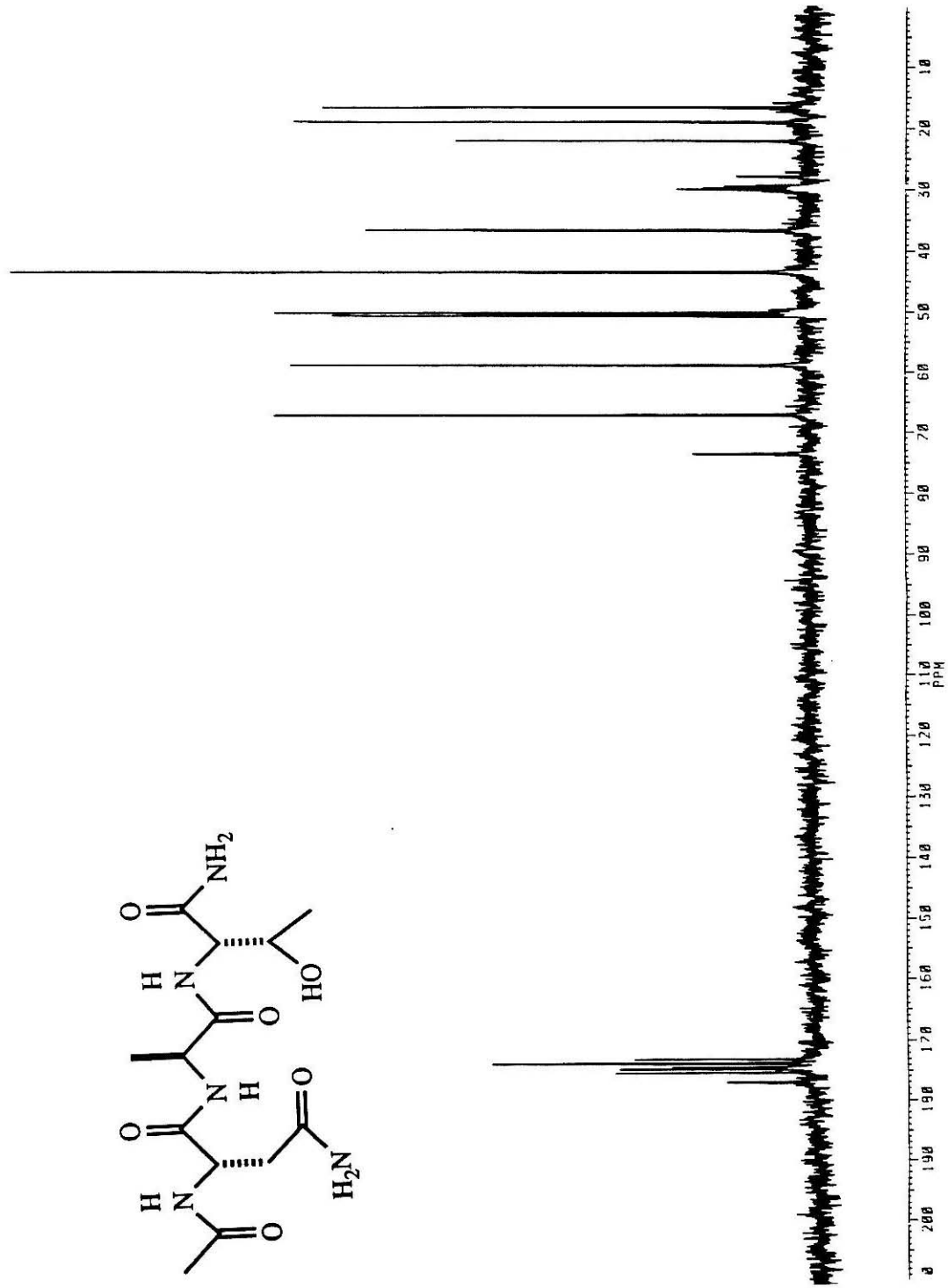

NHMe
HO
H_2N
PPM

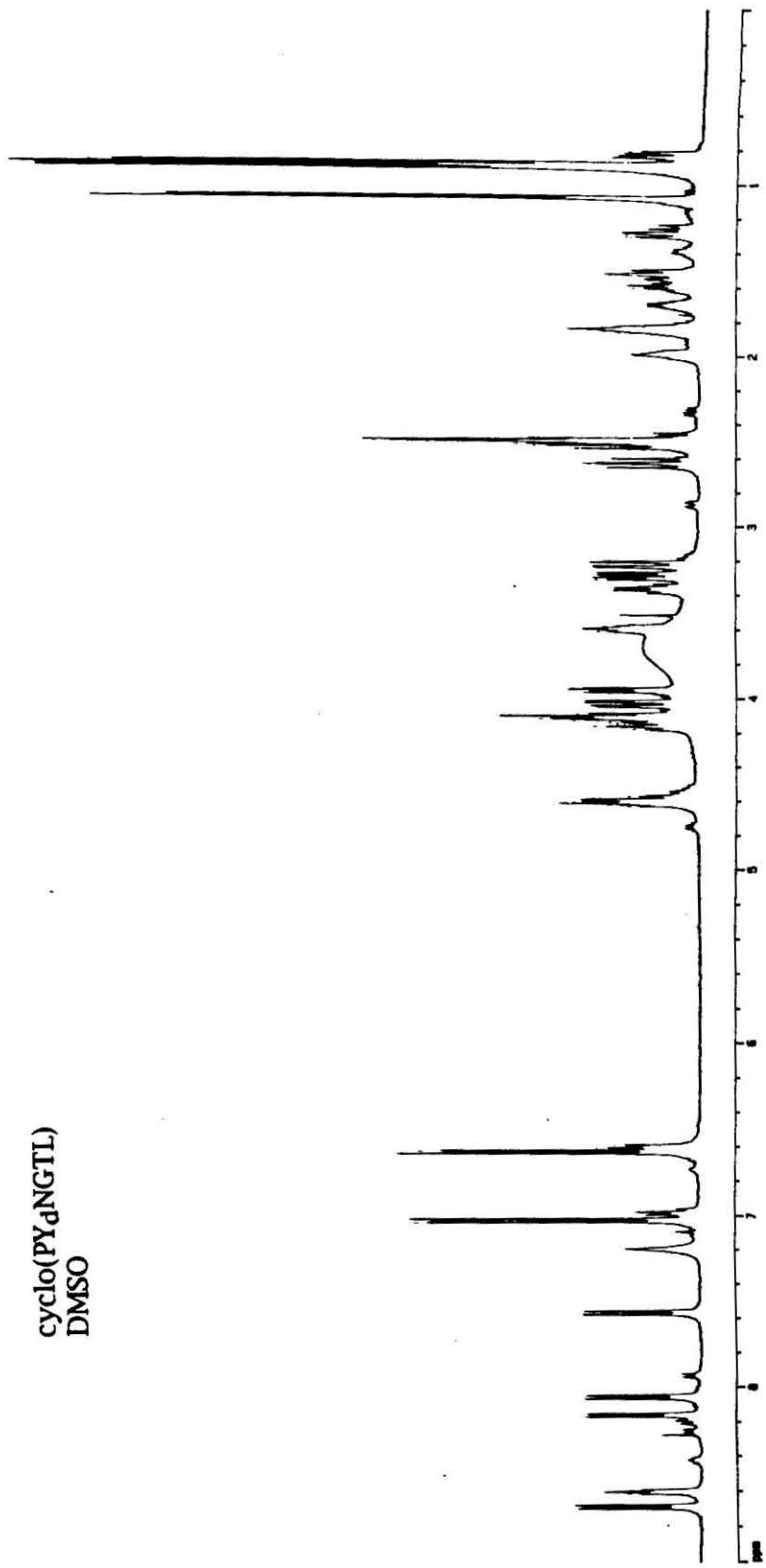

NHMe
O
H
N
HO
O
O
N
O
O
H
N
H_2N
O
PPM

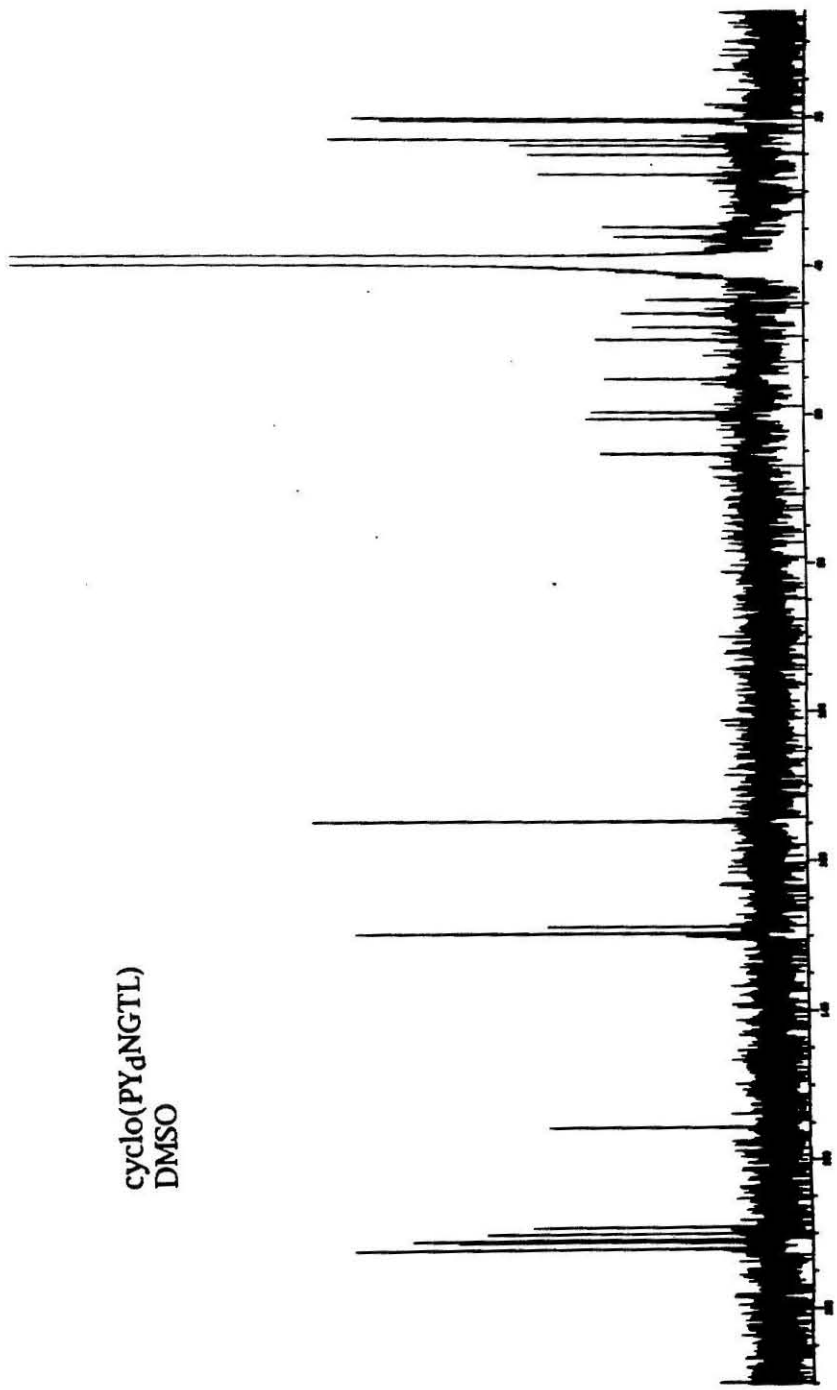

O
NH2
H
N
HO
O
N
O
O
H2N
H
N
O
4.5
PPM

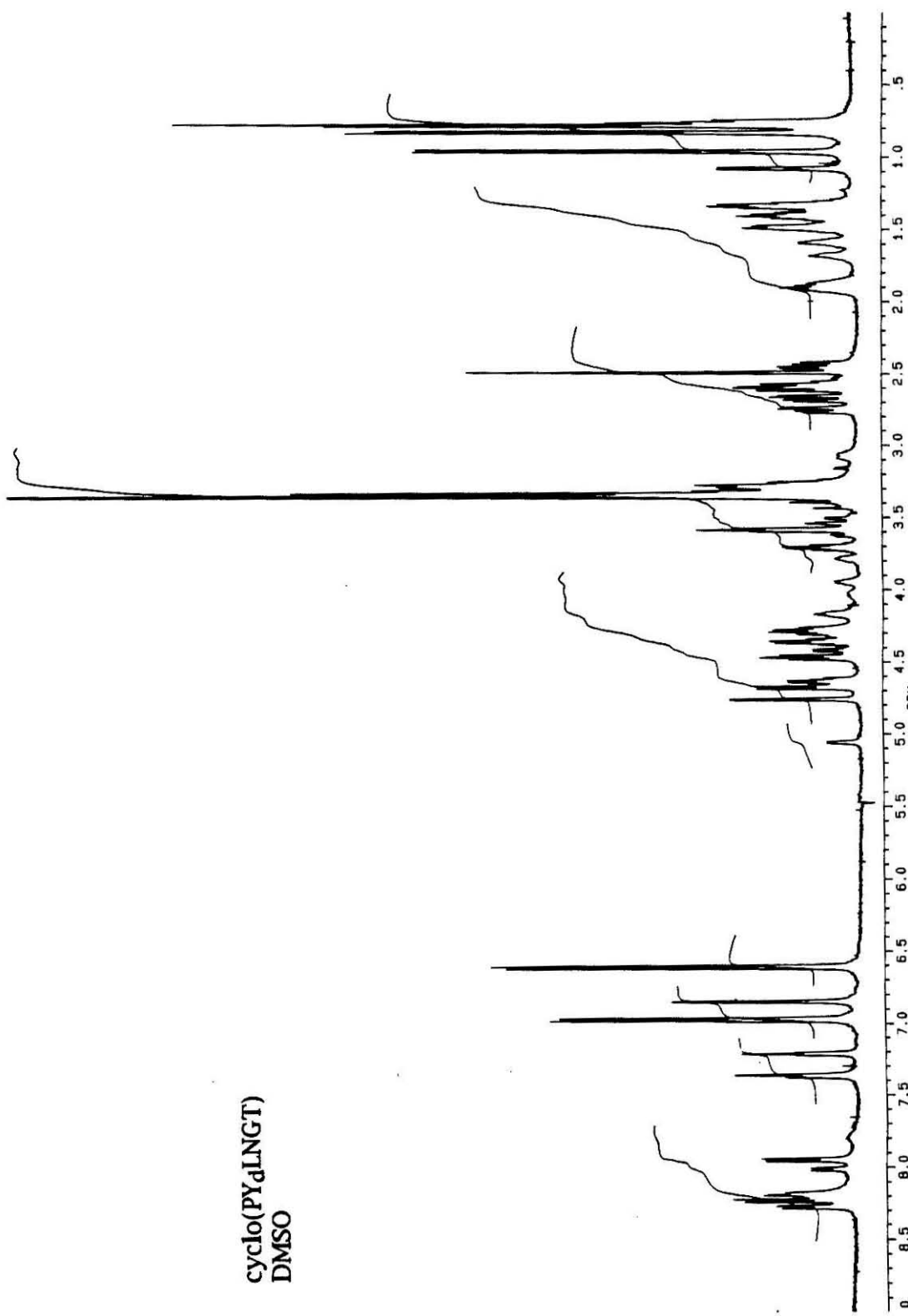

O
NH2
H
N
HO
O
N
O
O
H
N
H2N
O
PPM

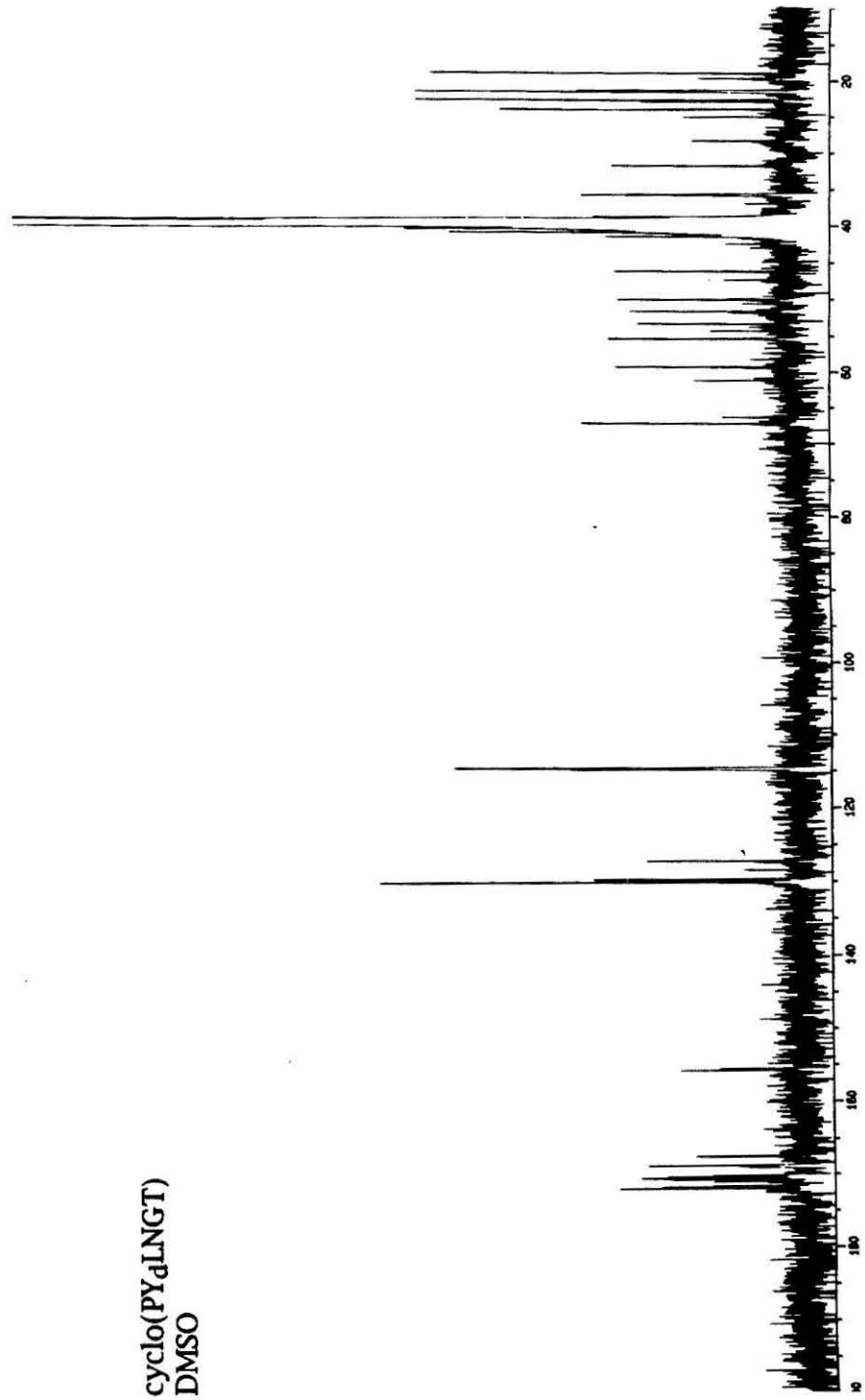

NH2
O
H
N
HO
O
N
H
O
O
H
N
H2N
O
PPM

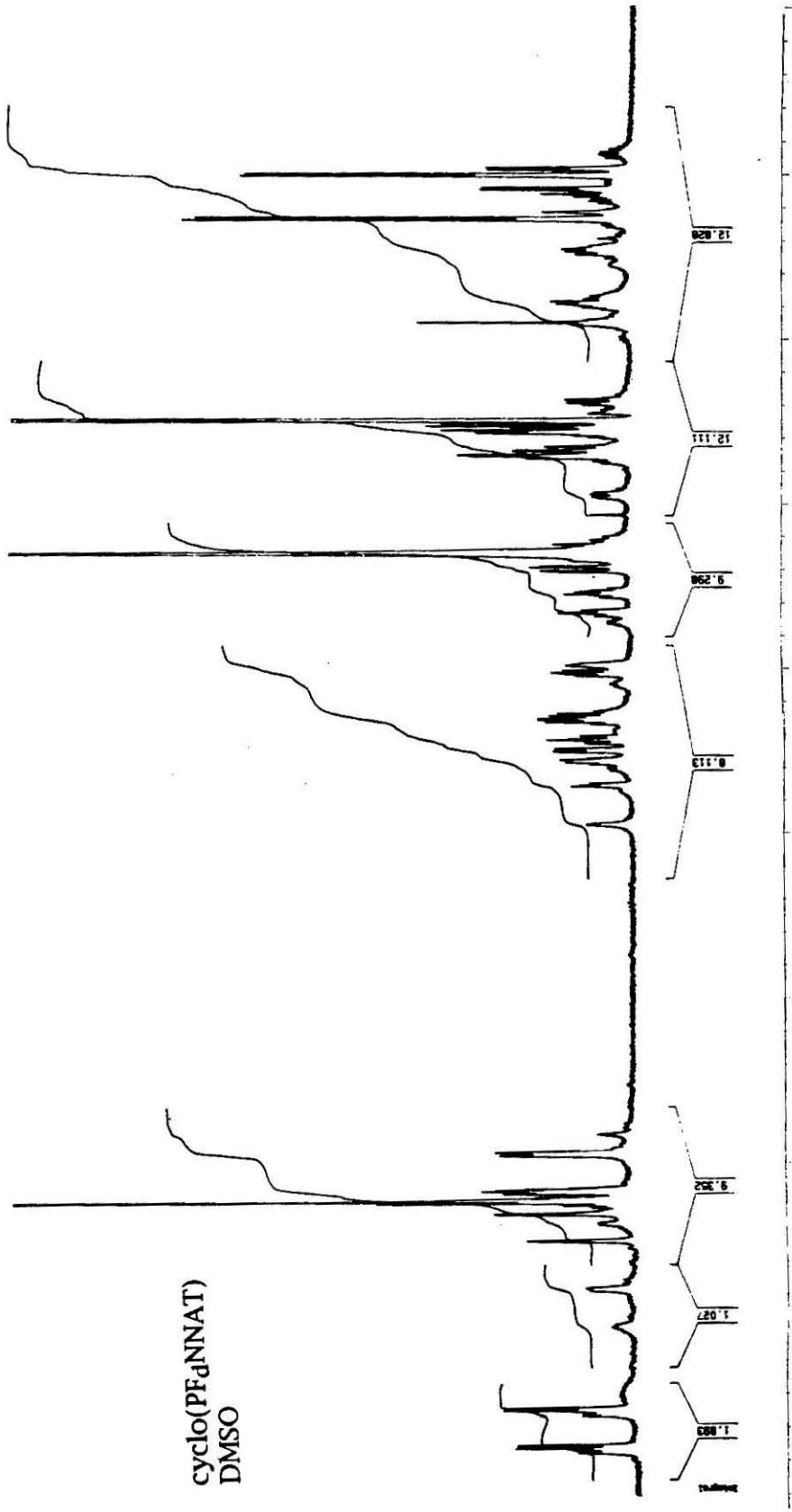

NH2
O
H
N
HO
O
N
H
O
O
H2N
H
N
O
PPM

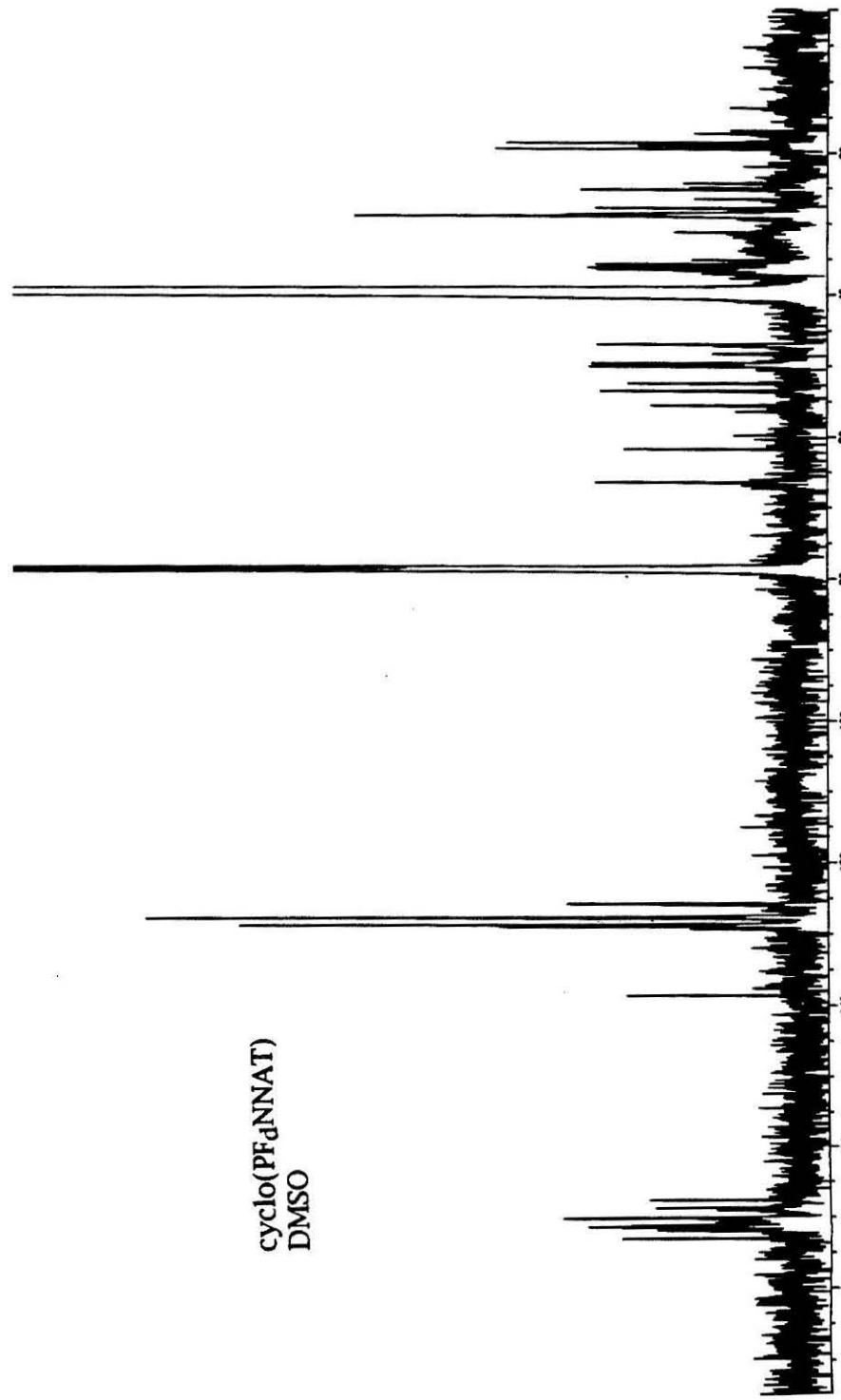

cyclo(PYdNGTL)
DMSO

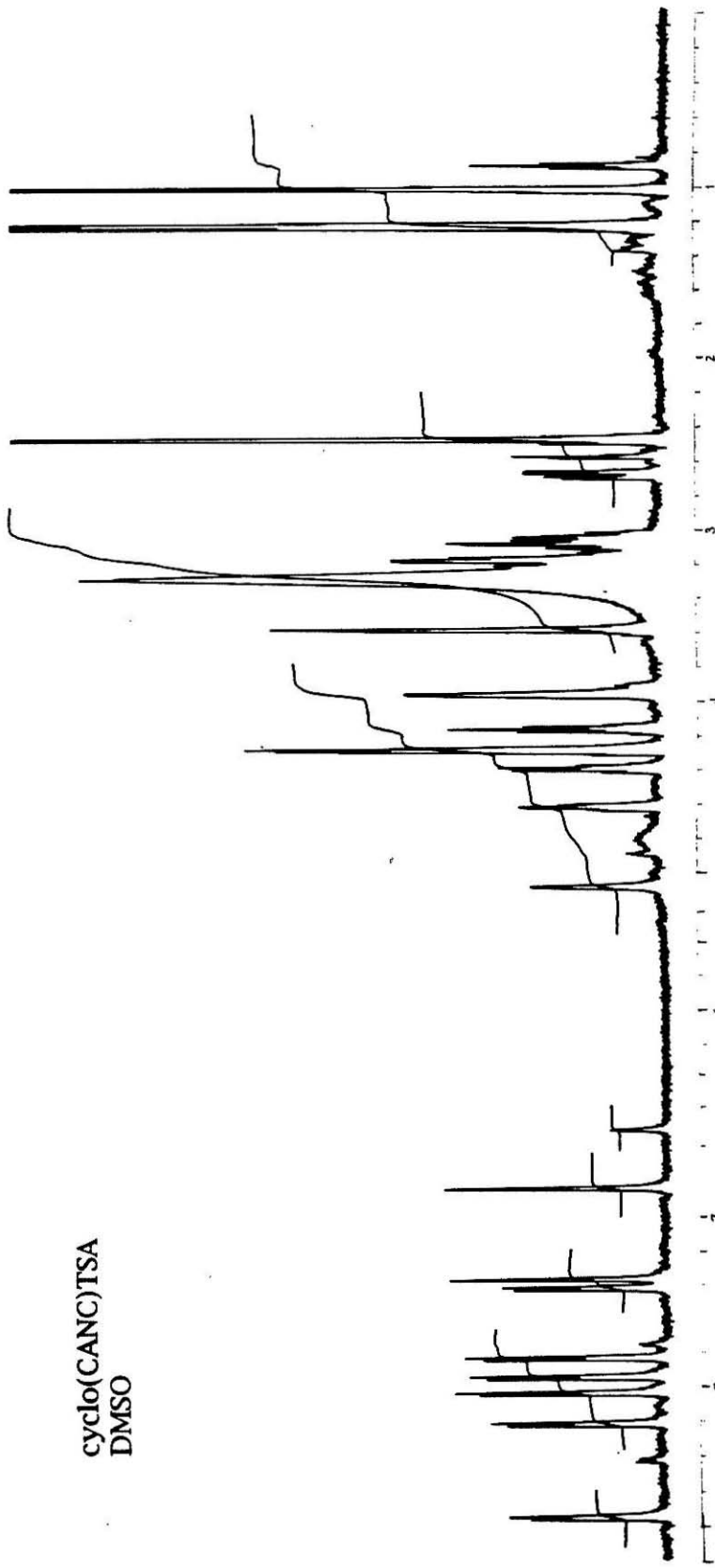

cyclo(PYdNGTL)
DMSO

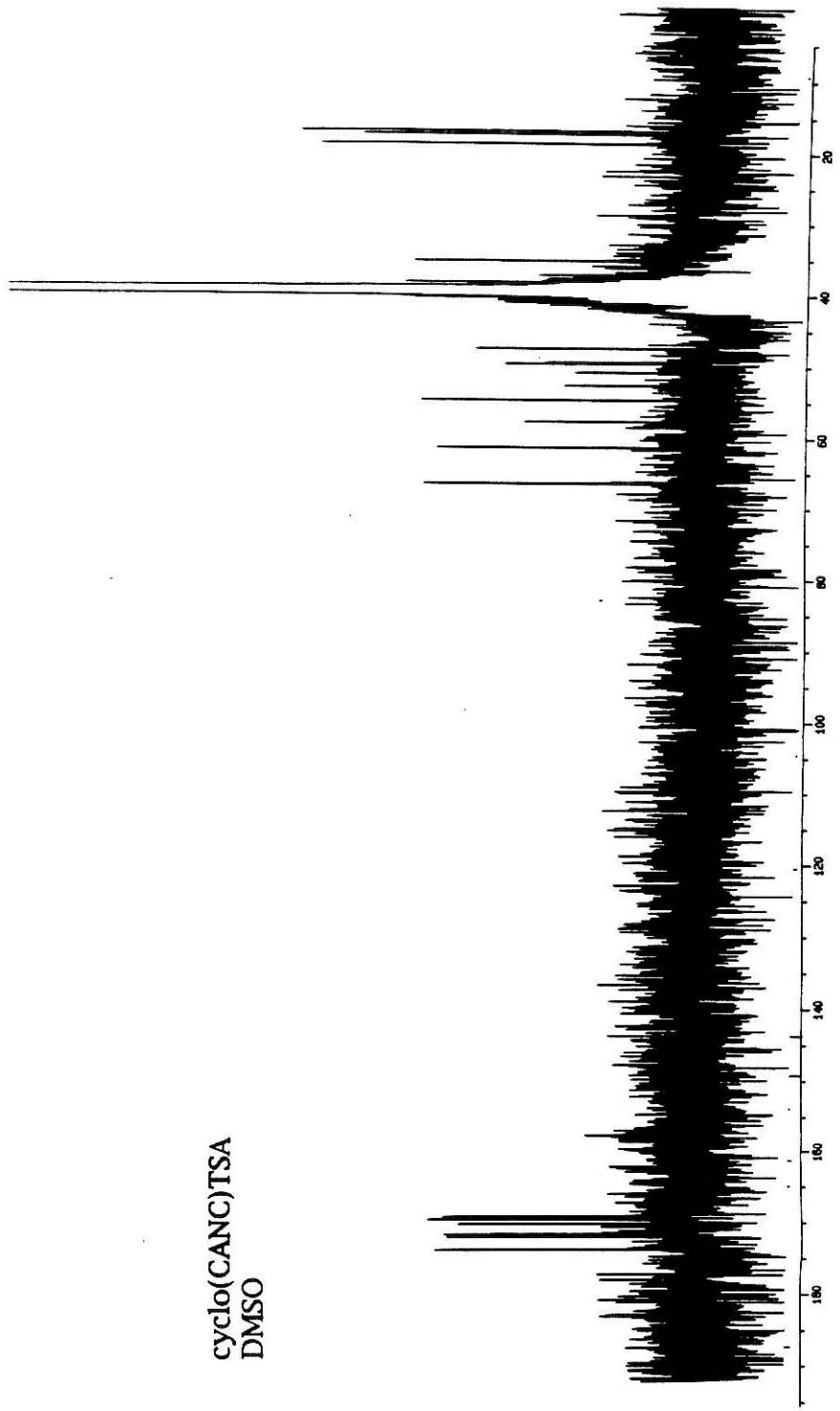

cyclo(PYdLNGT)
DMSO
0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
PPM

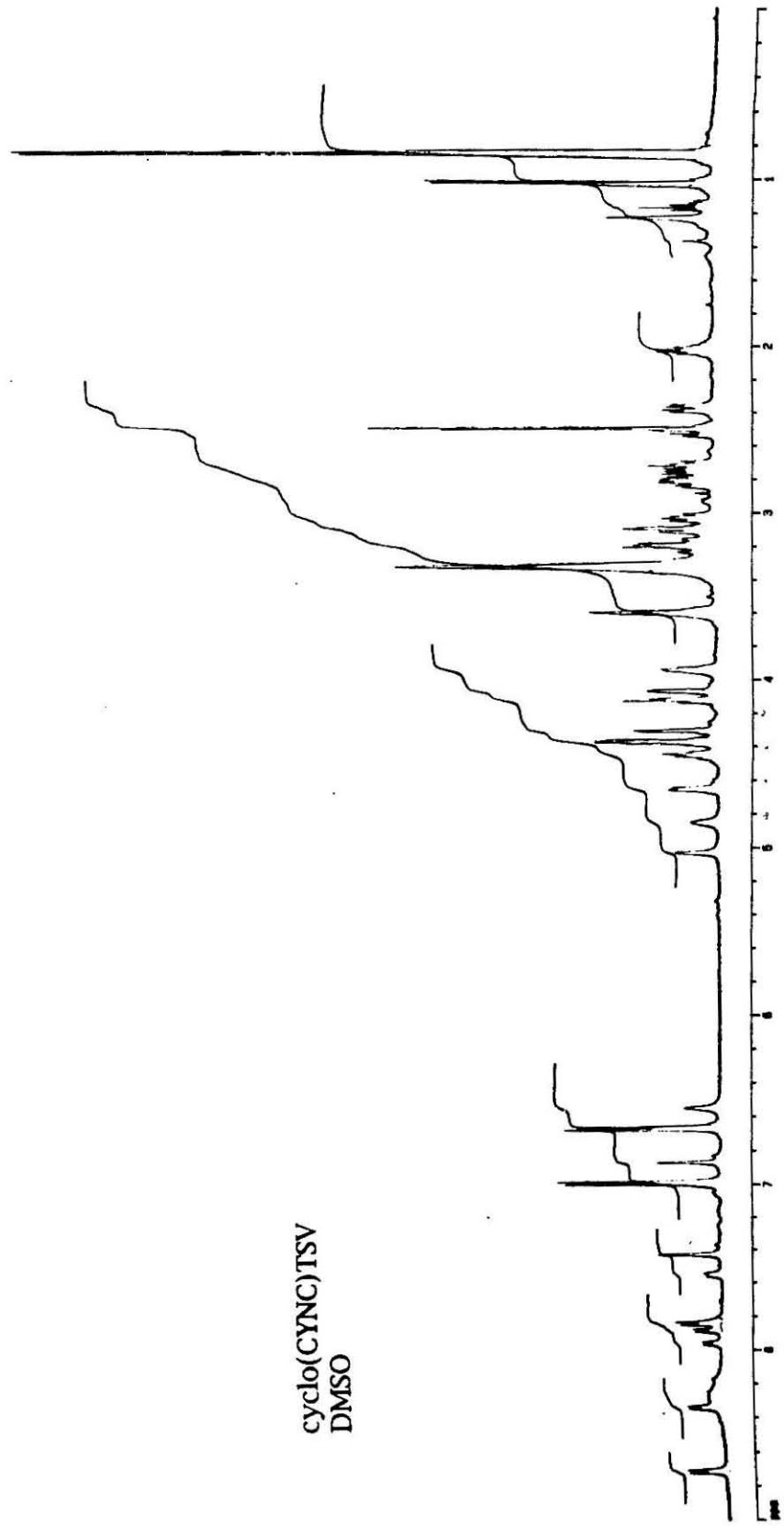

cyclo(PY$_d$LNGT)
DMSO

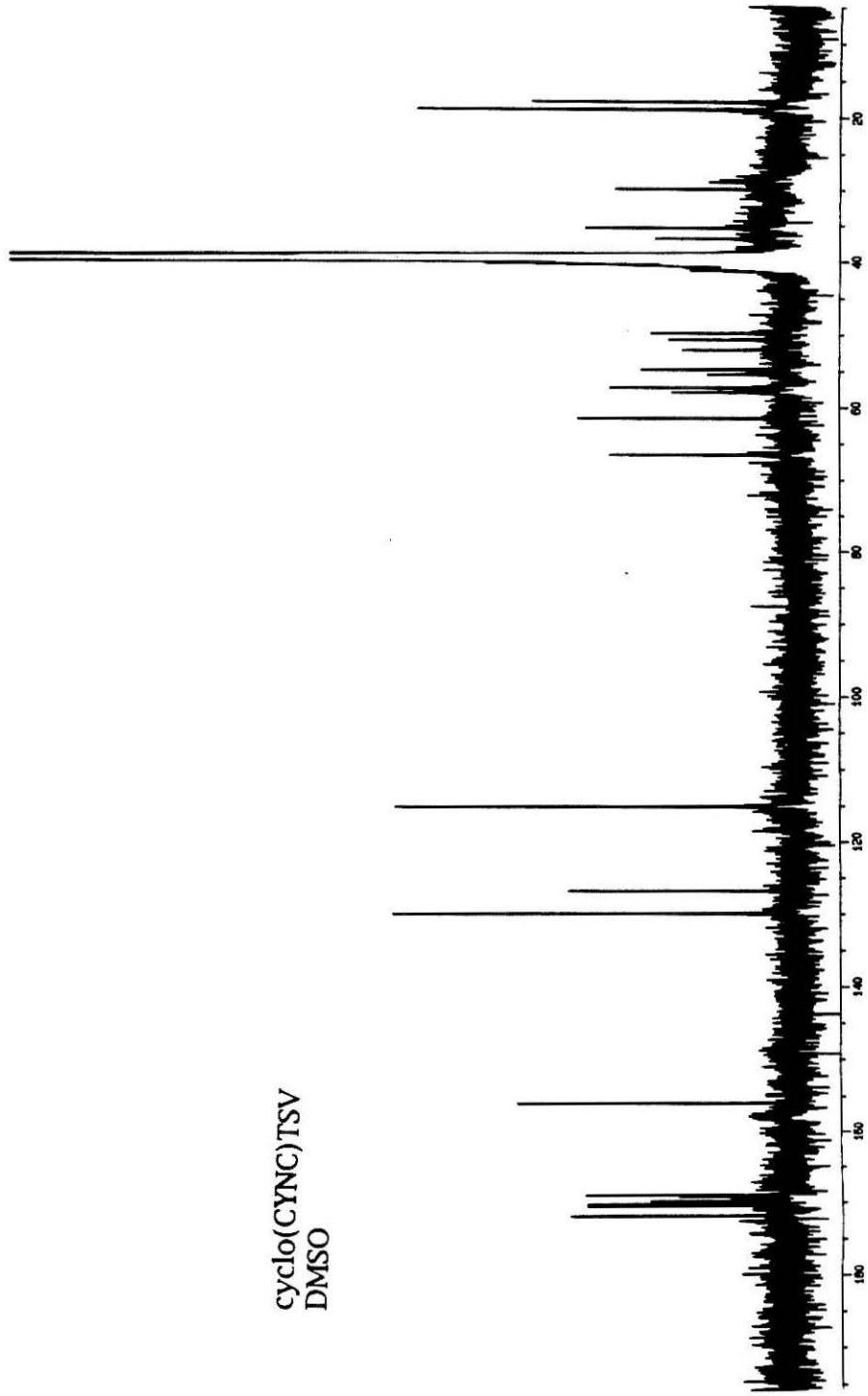

cyclo(PFdNNAT)
DMSO
1.883
1.027
9.352
8.113
9.298
12.111
12.626
5
6
7

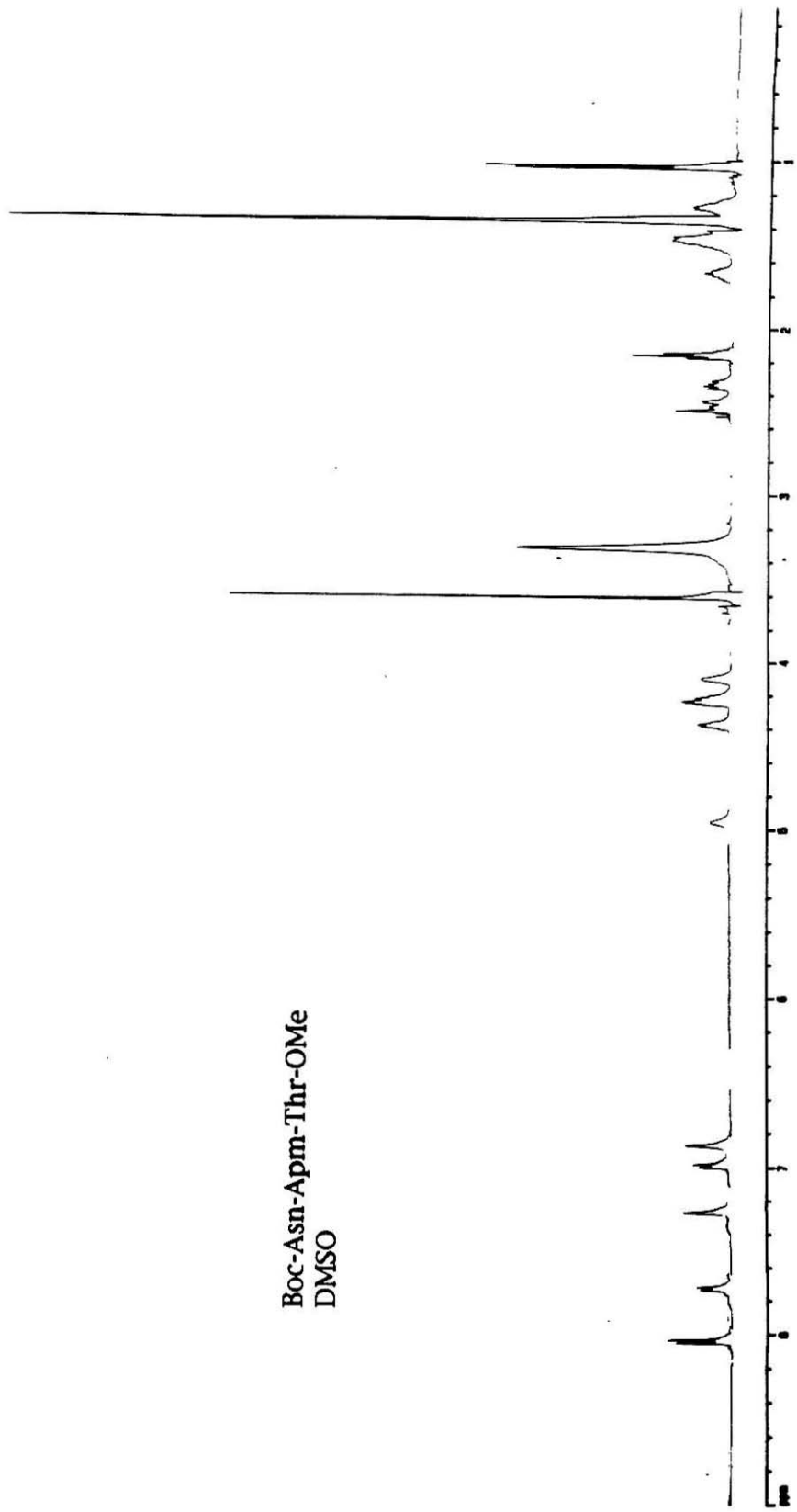

cyclo(PFdNNAT)
DMSO
20
40
60
80
100
120
140
160
180

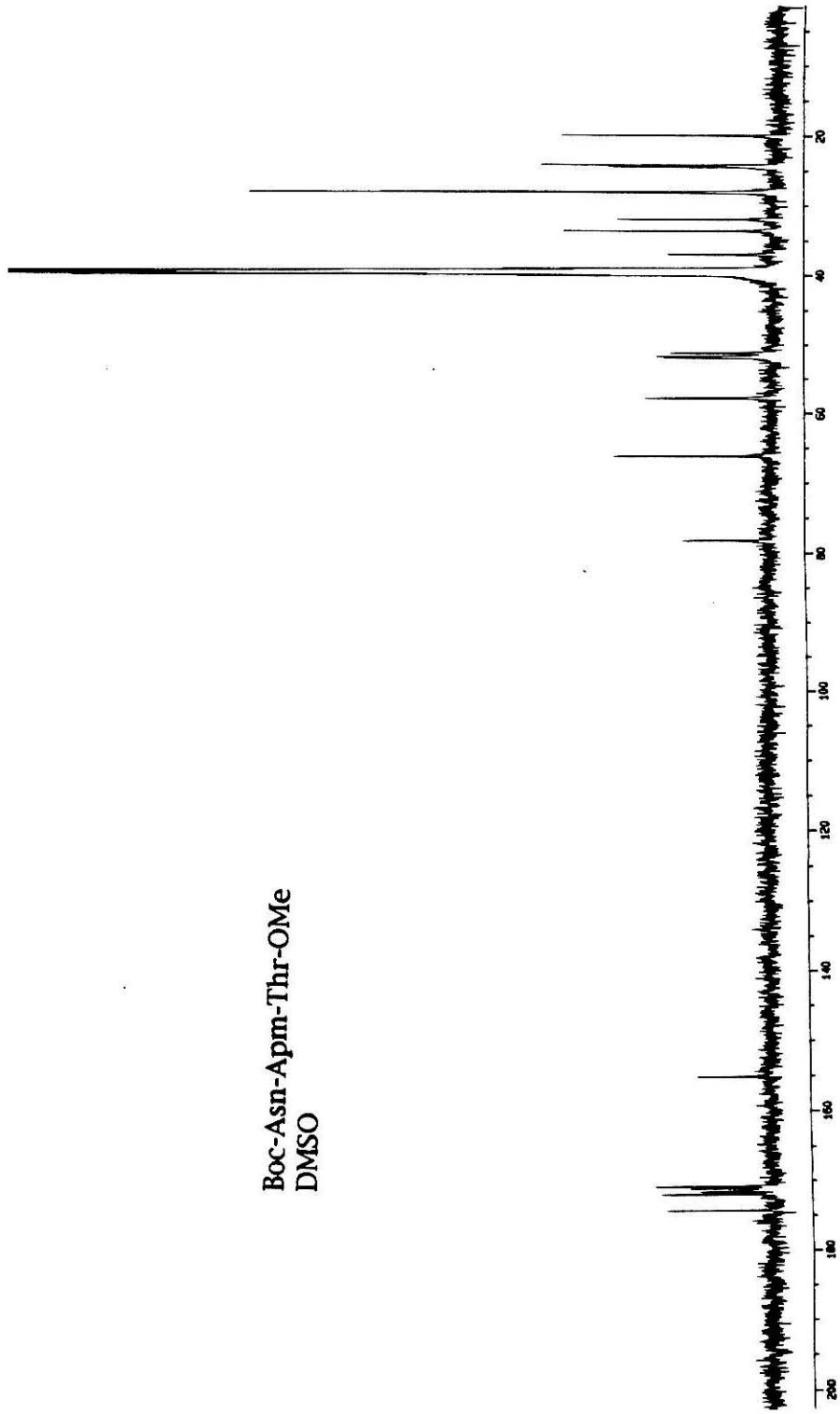

cyclo(CANC)TSA
DMSO

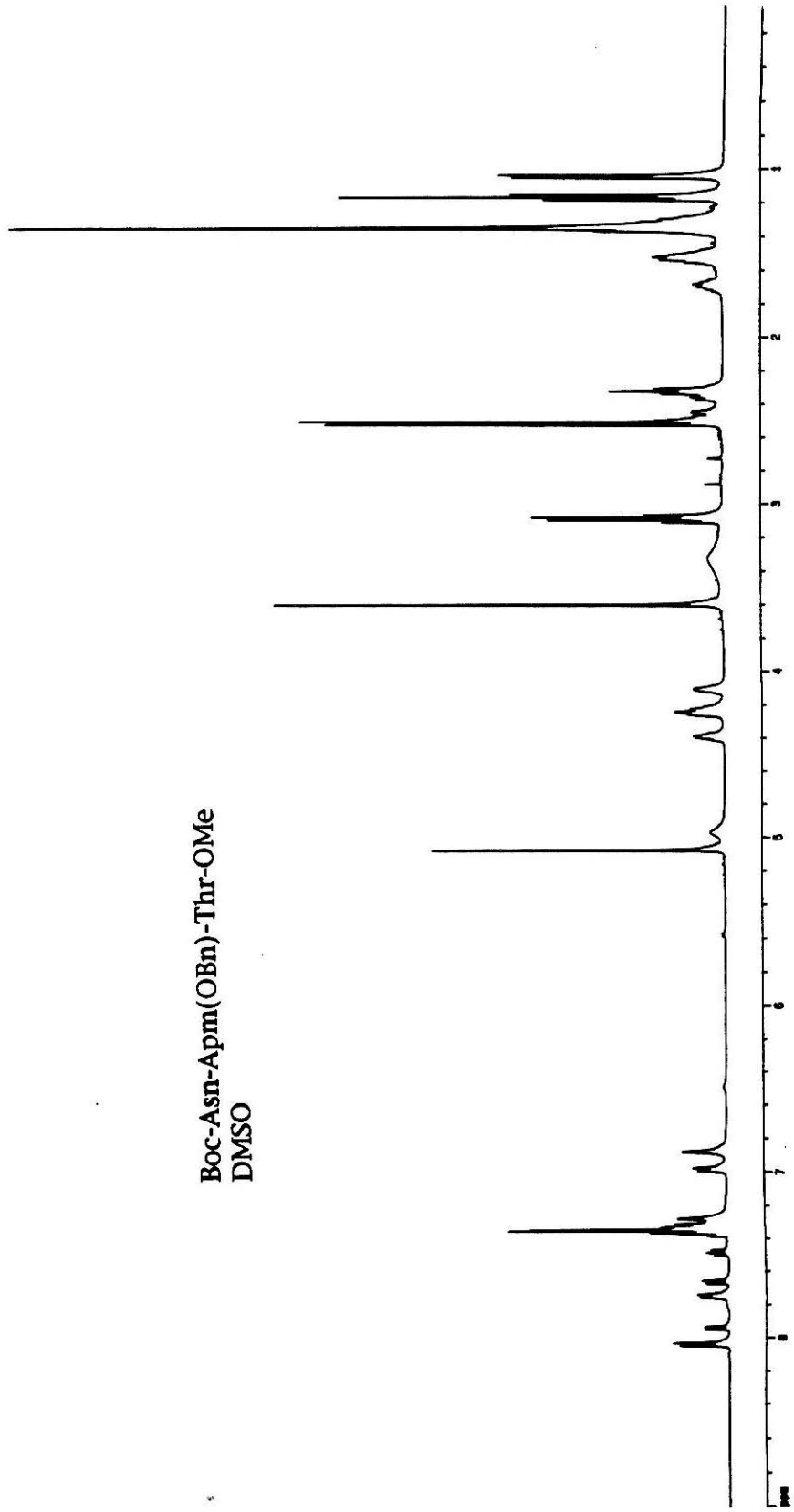

cyclo(CANC)TSA
DMSO
20
40
60
80
100
120
140
160
180

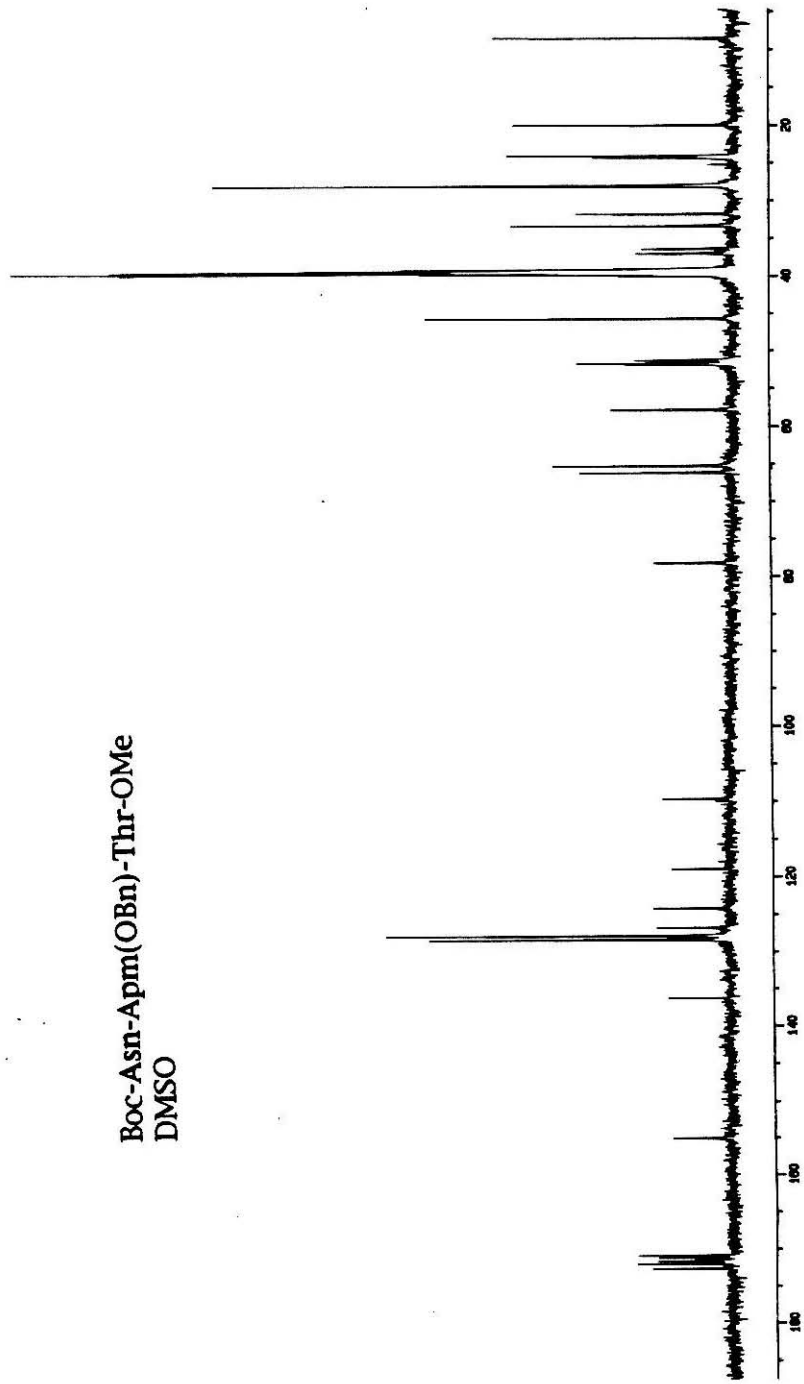

cyclo(CYNC)TSV
DMSO


cyclo(CYNC)TSV
DMSO
20
40
60
80
100
120
140
160
180

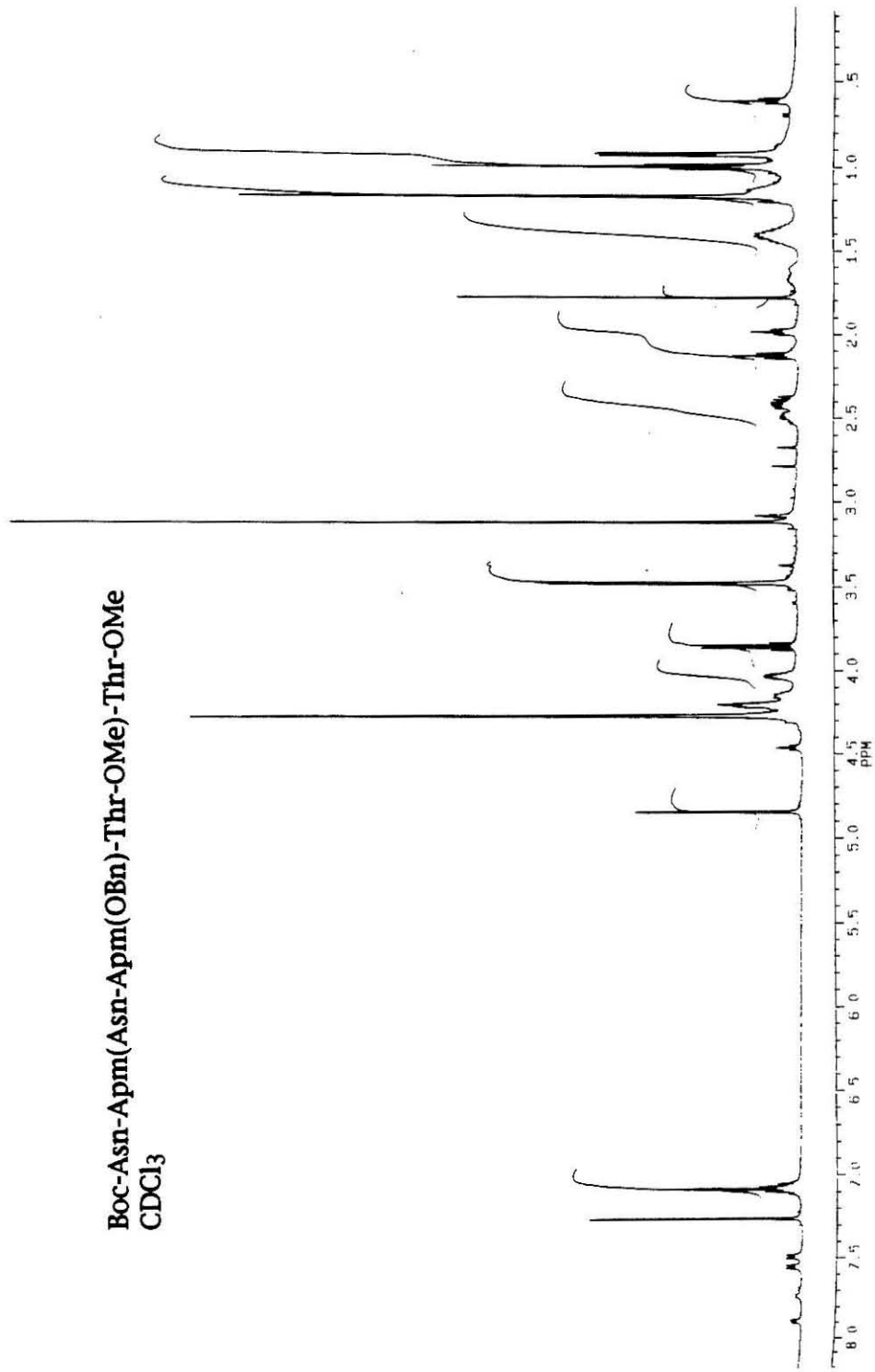

Boc-Asn-Apm-Thr-OMe
DMSO

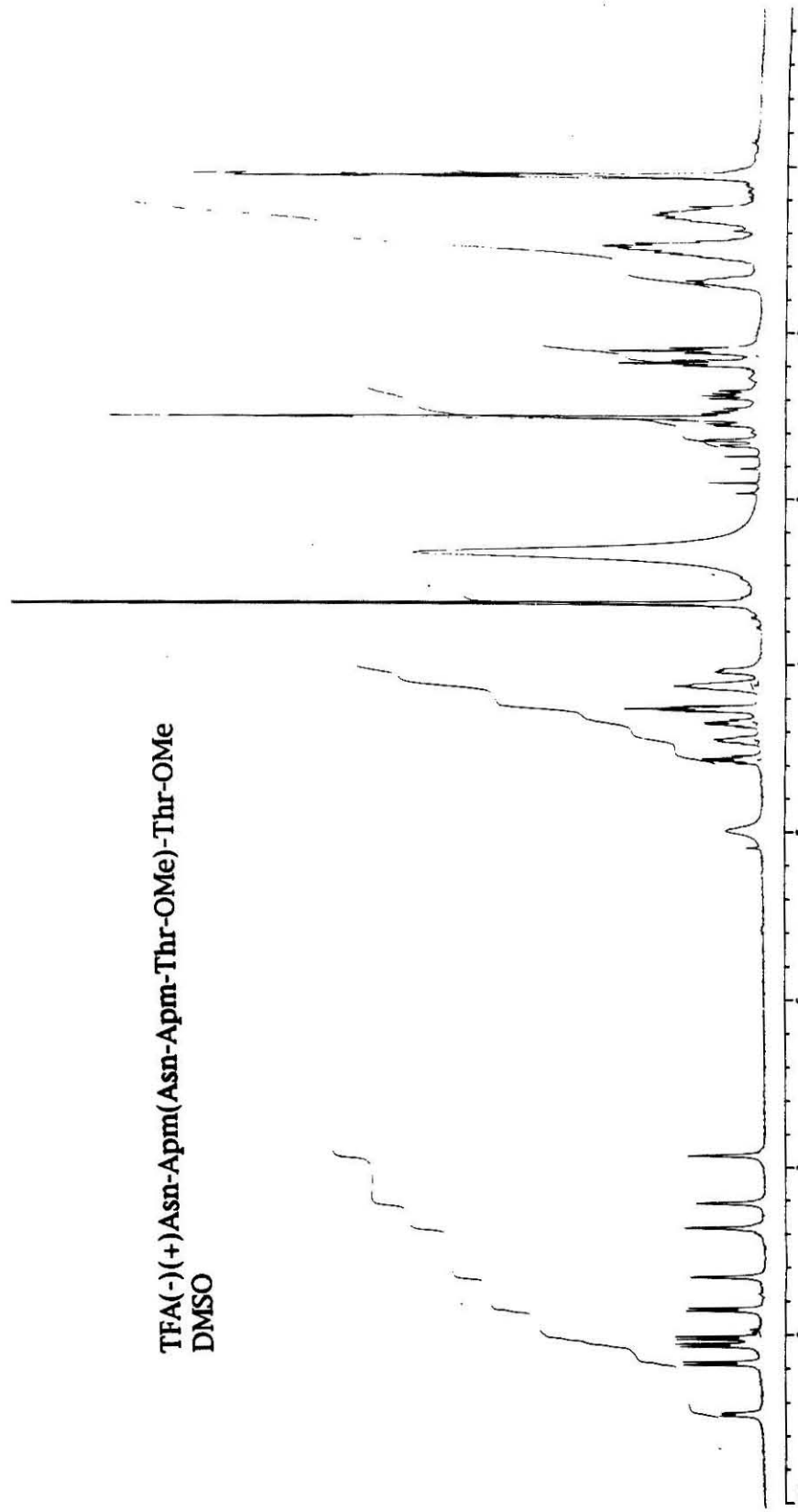

Boc-Asn-Apm-Thr-OMe
DMSO
200
180
160
140
120
100
80
60
40
20


Boc-Asn-Apm(OBn)-Thr-OMe
DMSO

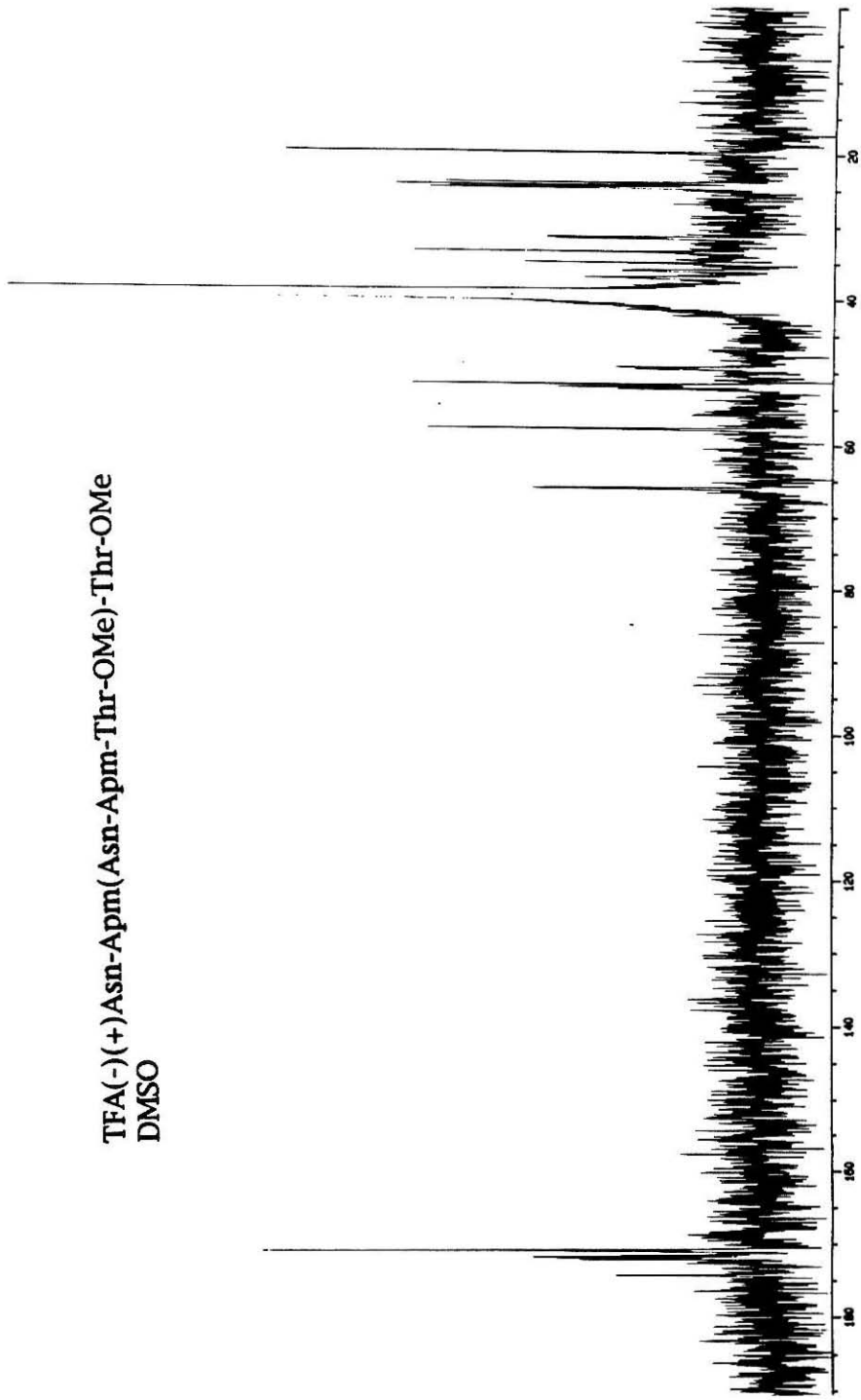

Boc-Asn-Apm(OBn)-Thr-OMe
DMSO

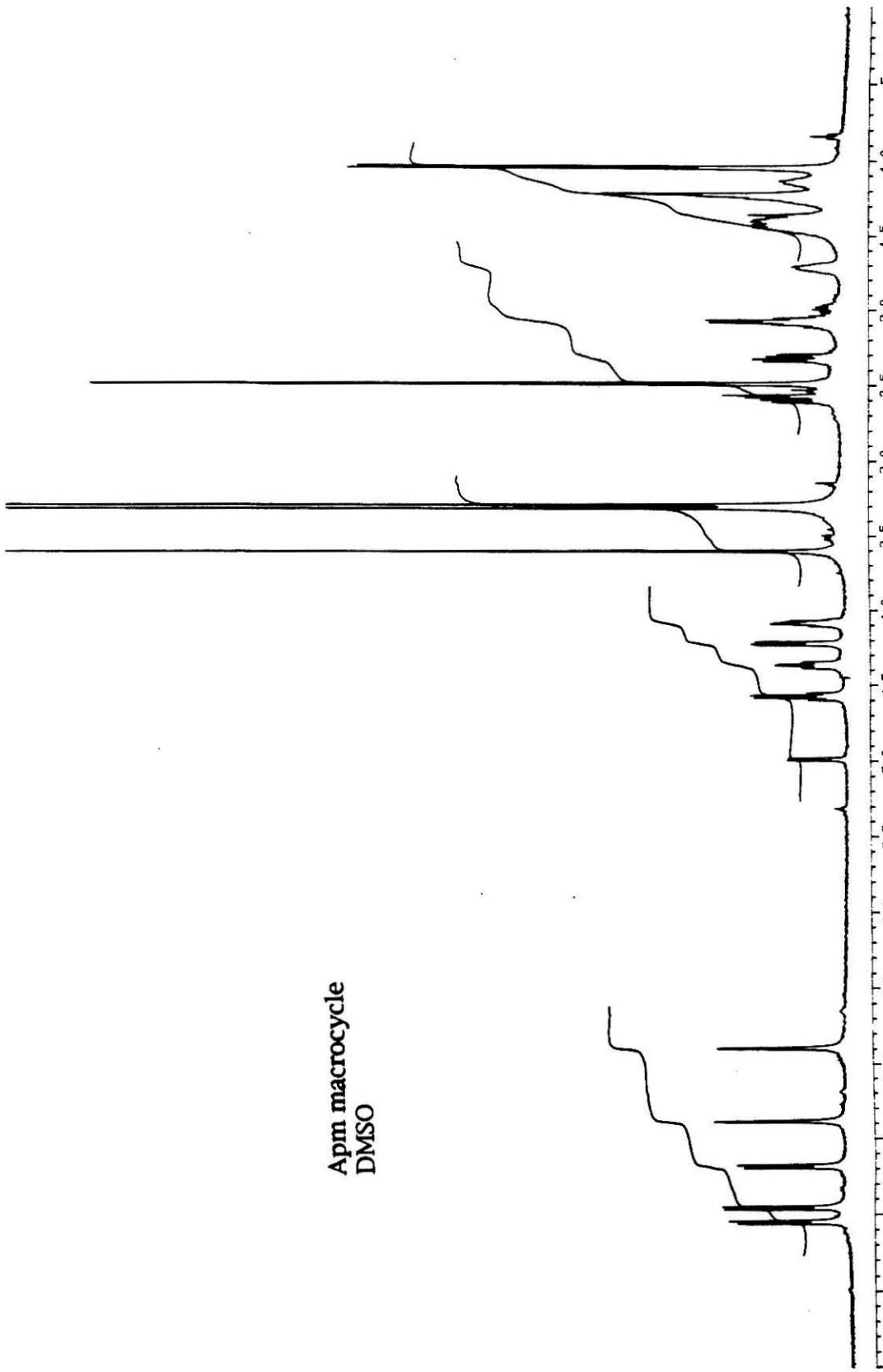
Boc-Asn-Apm(Asn-Apm(OBn)-Thr-OMe)-Thr-OMe
$CDCl_3$

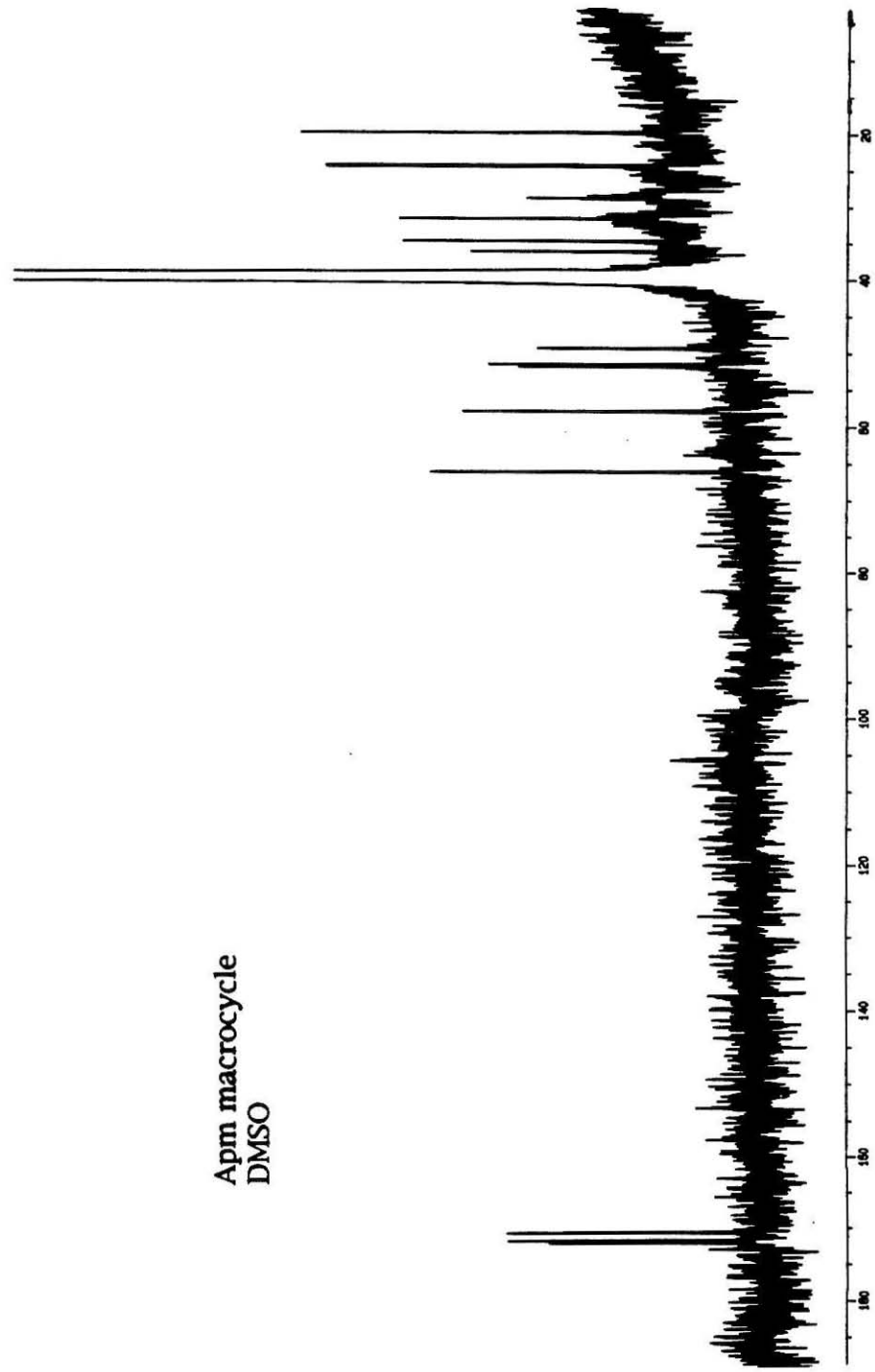


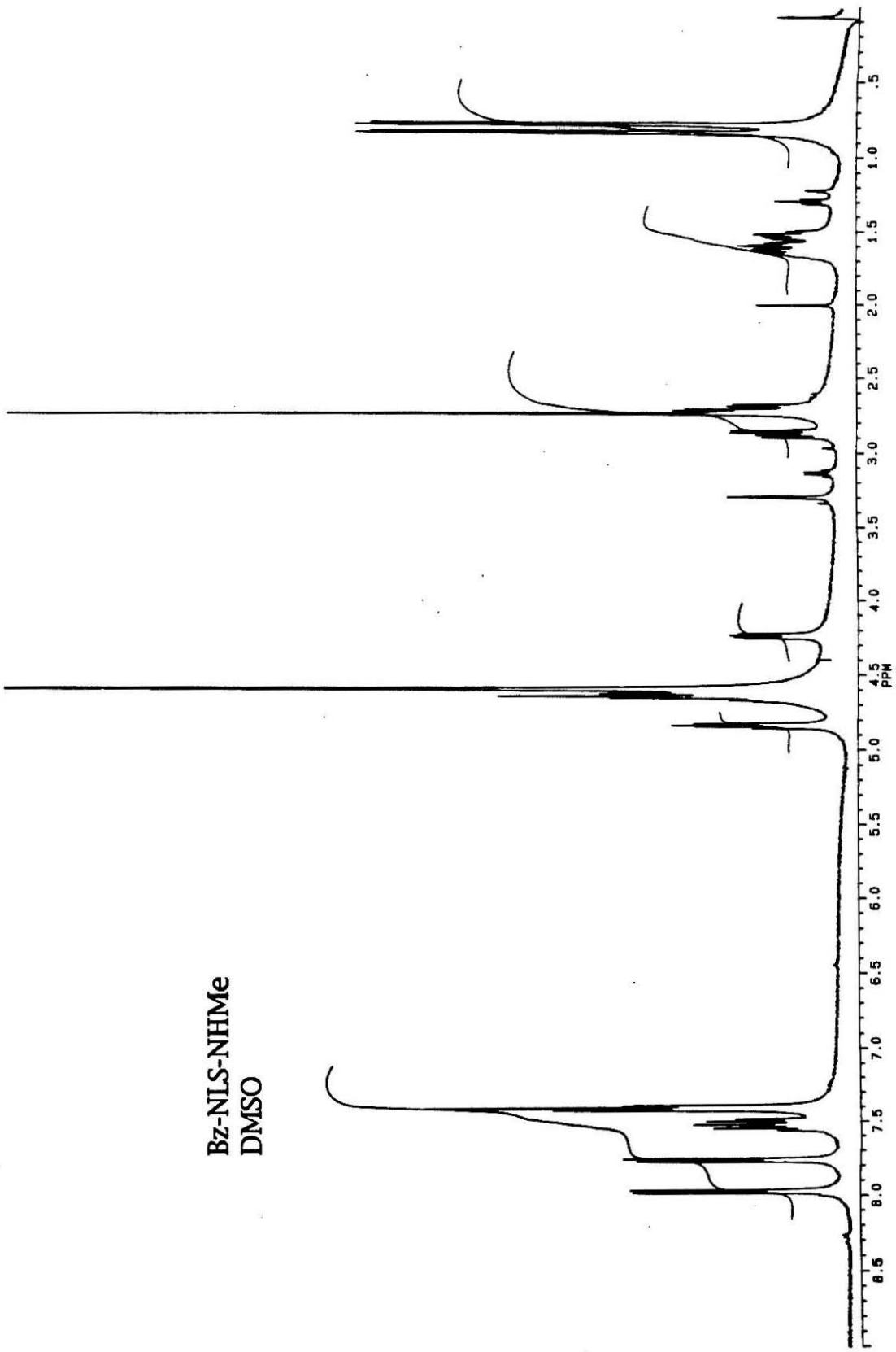

TFA(-)(+)Asn-Apm(Asn-Apm-Thr-OMe)-Thr-OMe
DMSO
ppm

TFA(-)(+)Asn-Apm(Asn-Apm-Thr-OMe)-Thr-OMe
DMSO

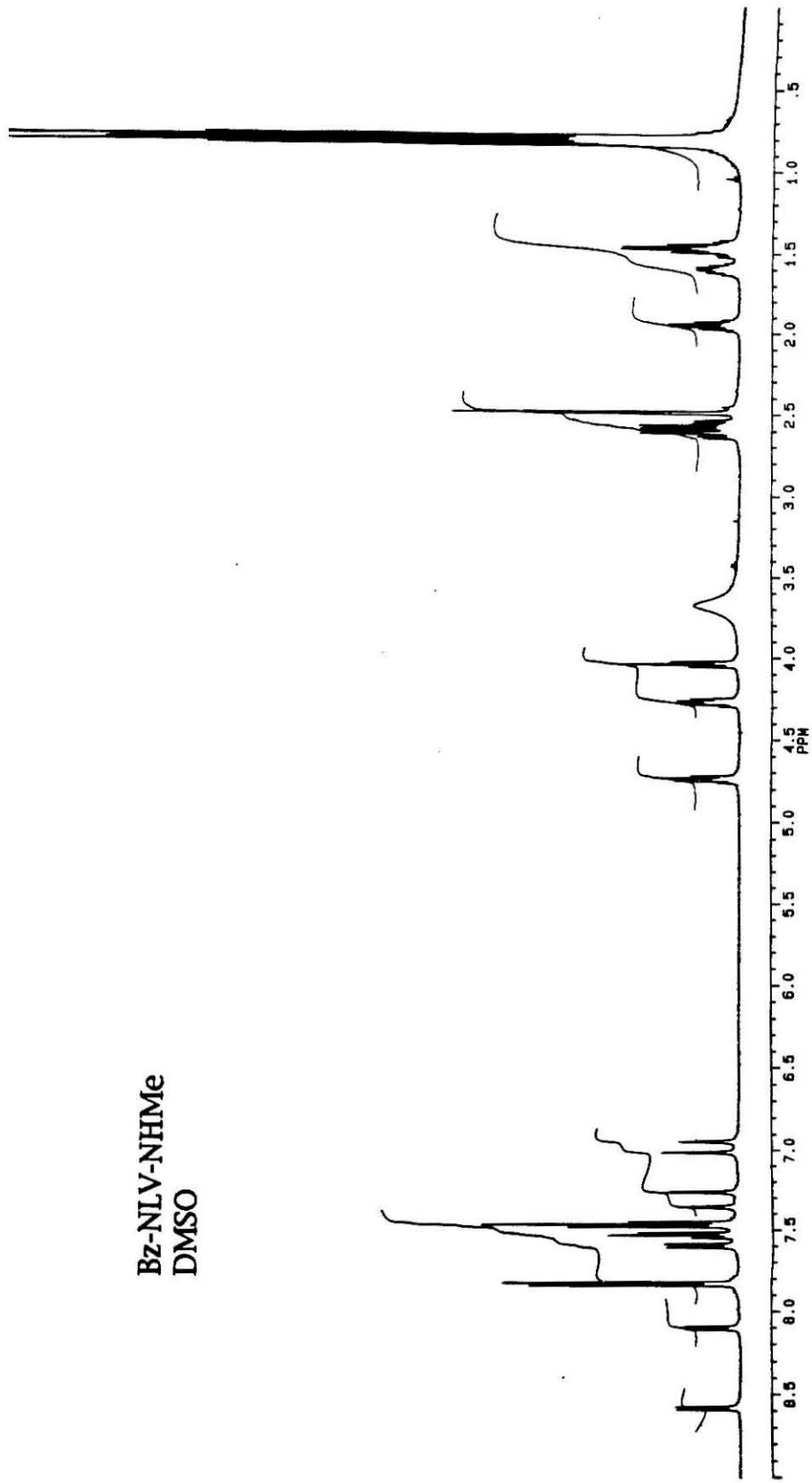


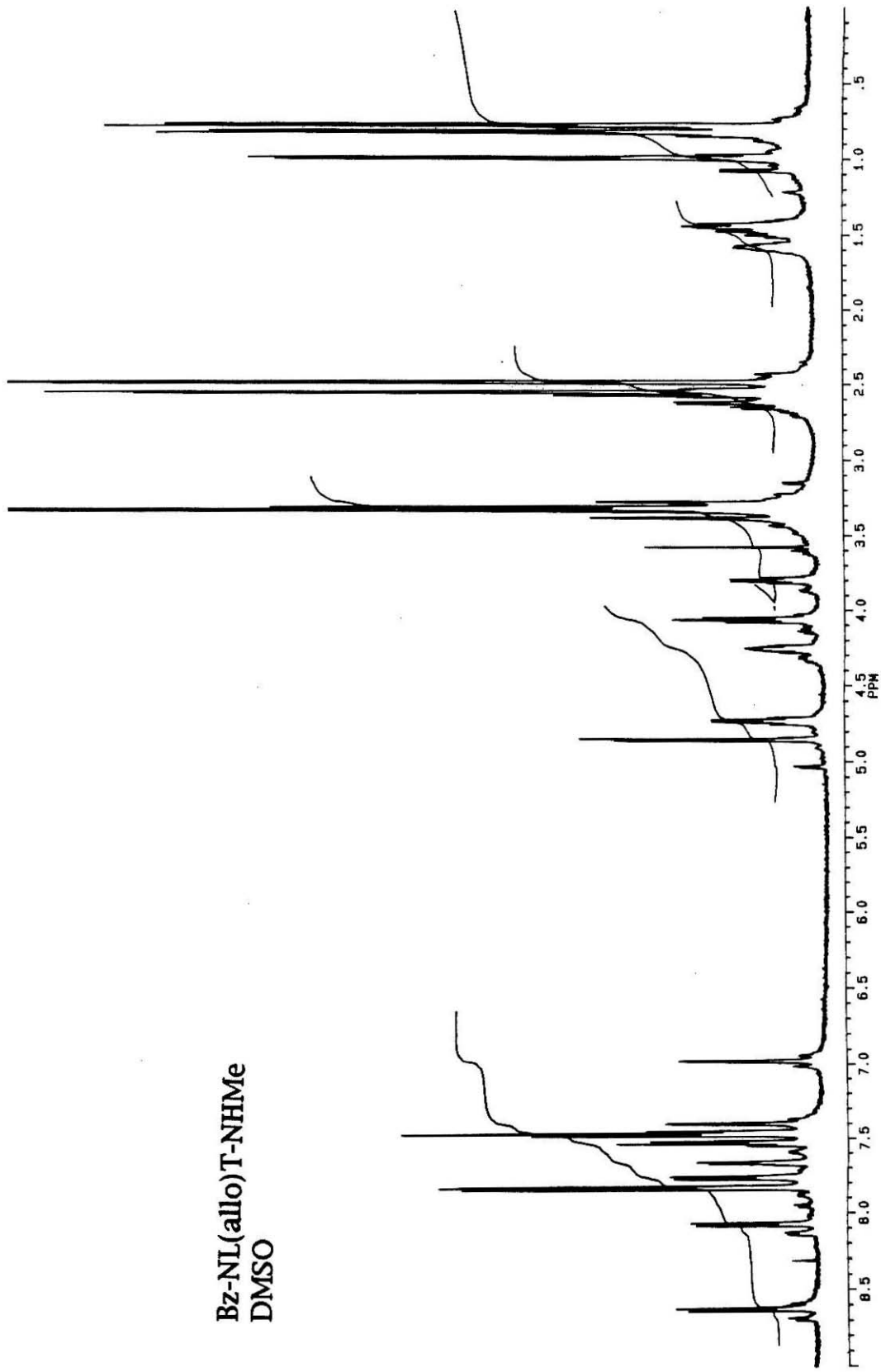

Apm macrocycle
DMSO
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
PPM
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5

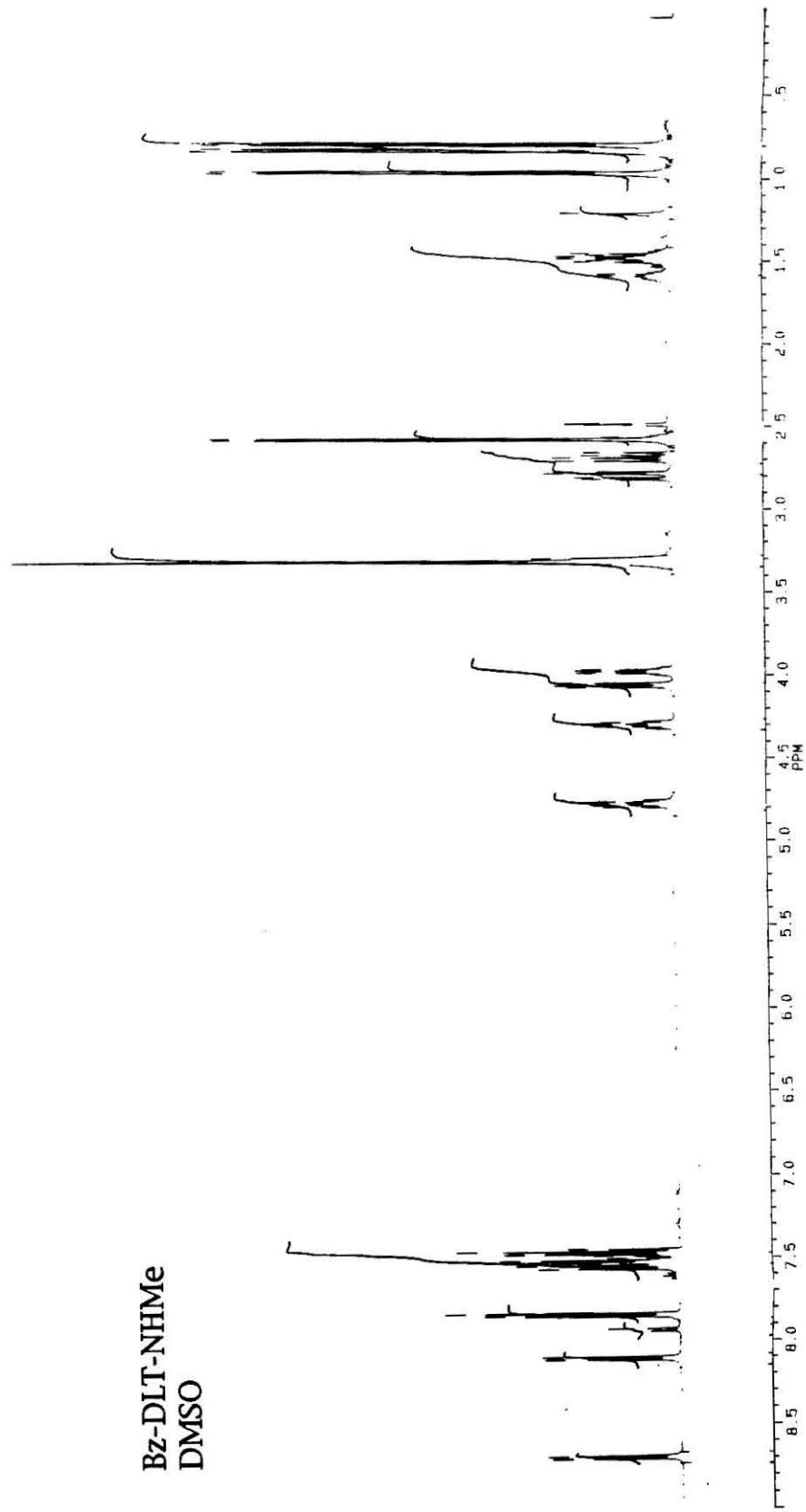

Apm macrocycle
DMSO
20
40
60
80
100
120
140
160
180

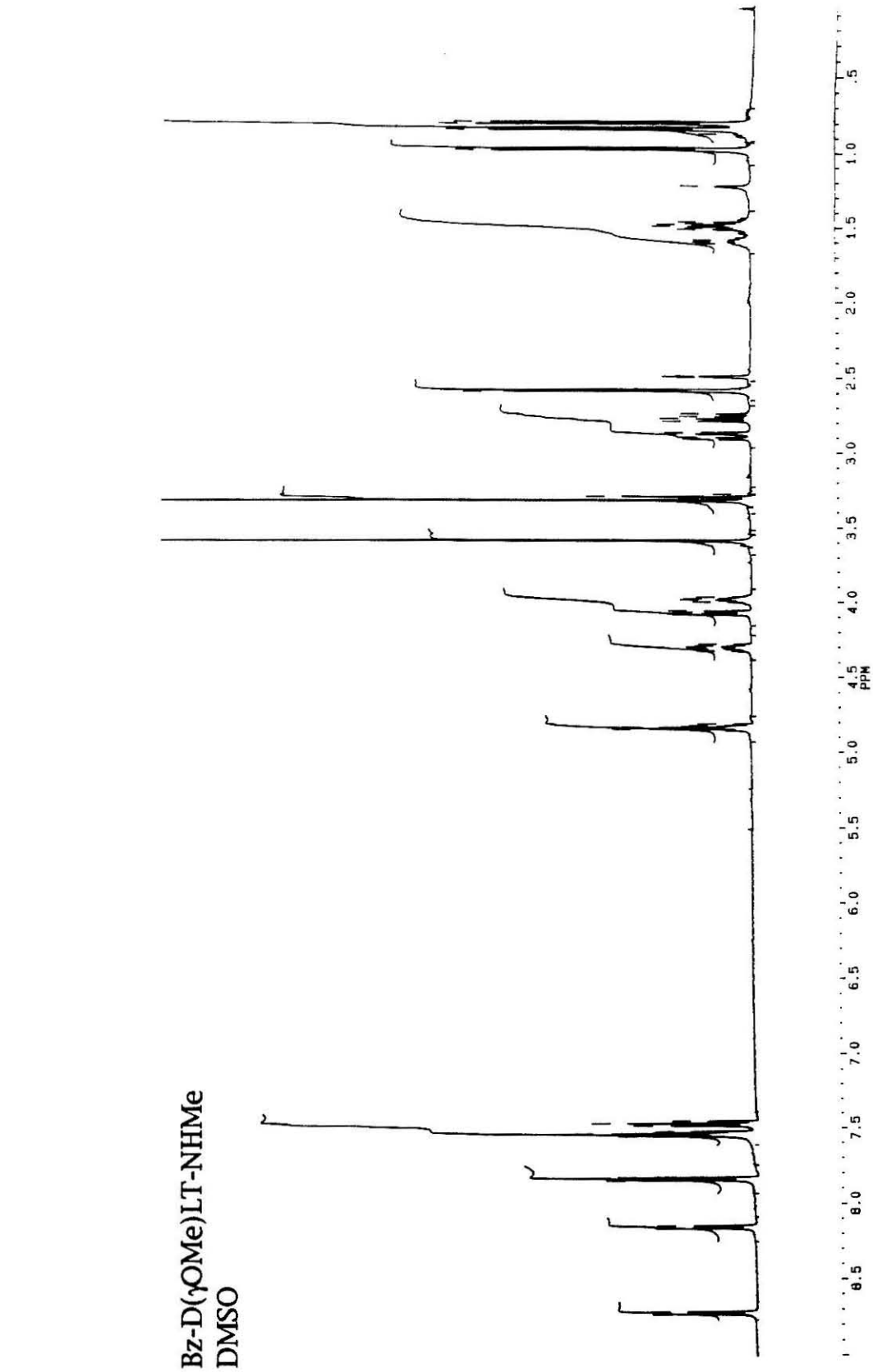

Bz-NLS-NHMe
DMSO
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
PPM
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5

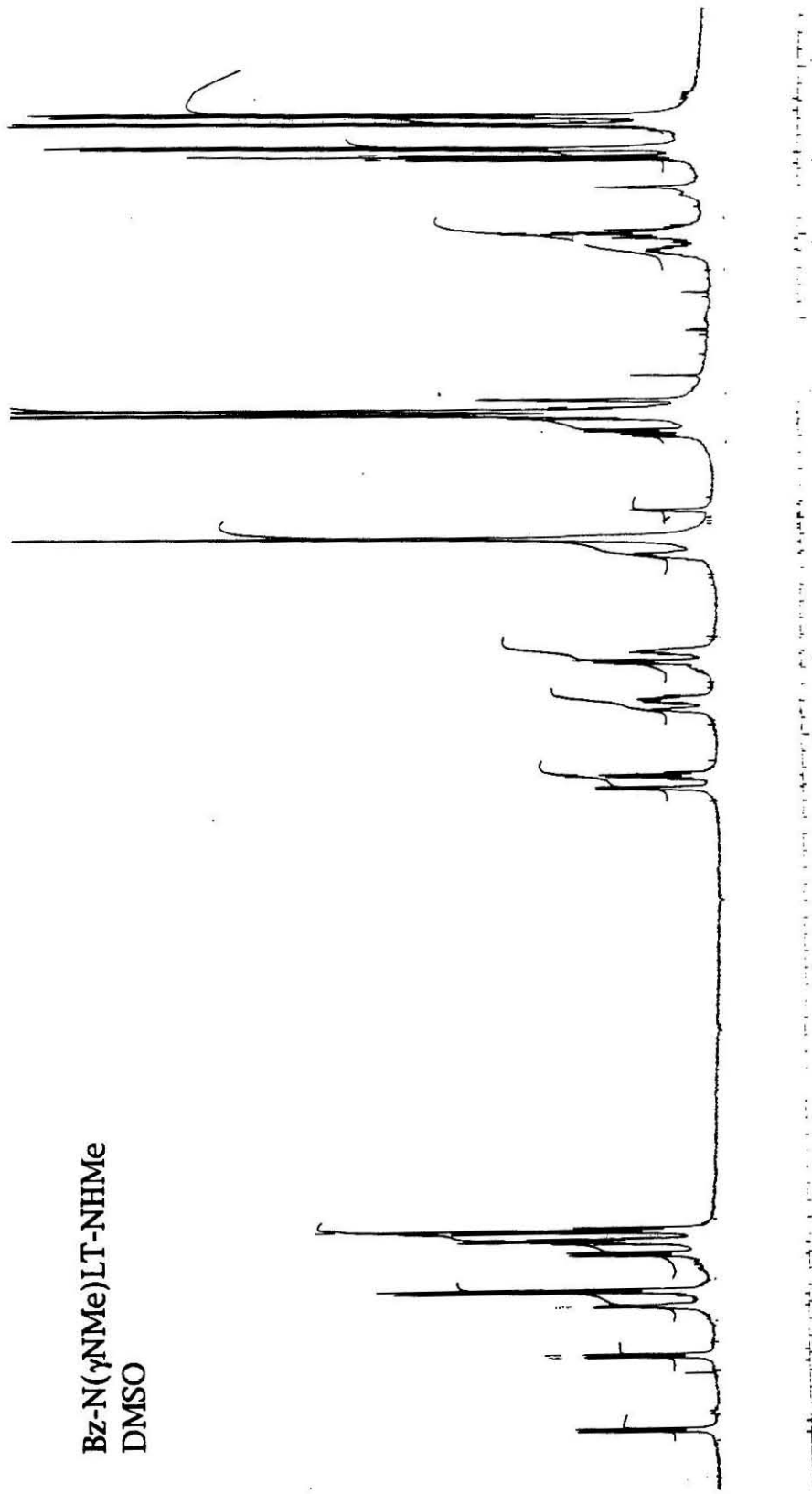
Bz-NLV-NHMe
DMSO

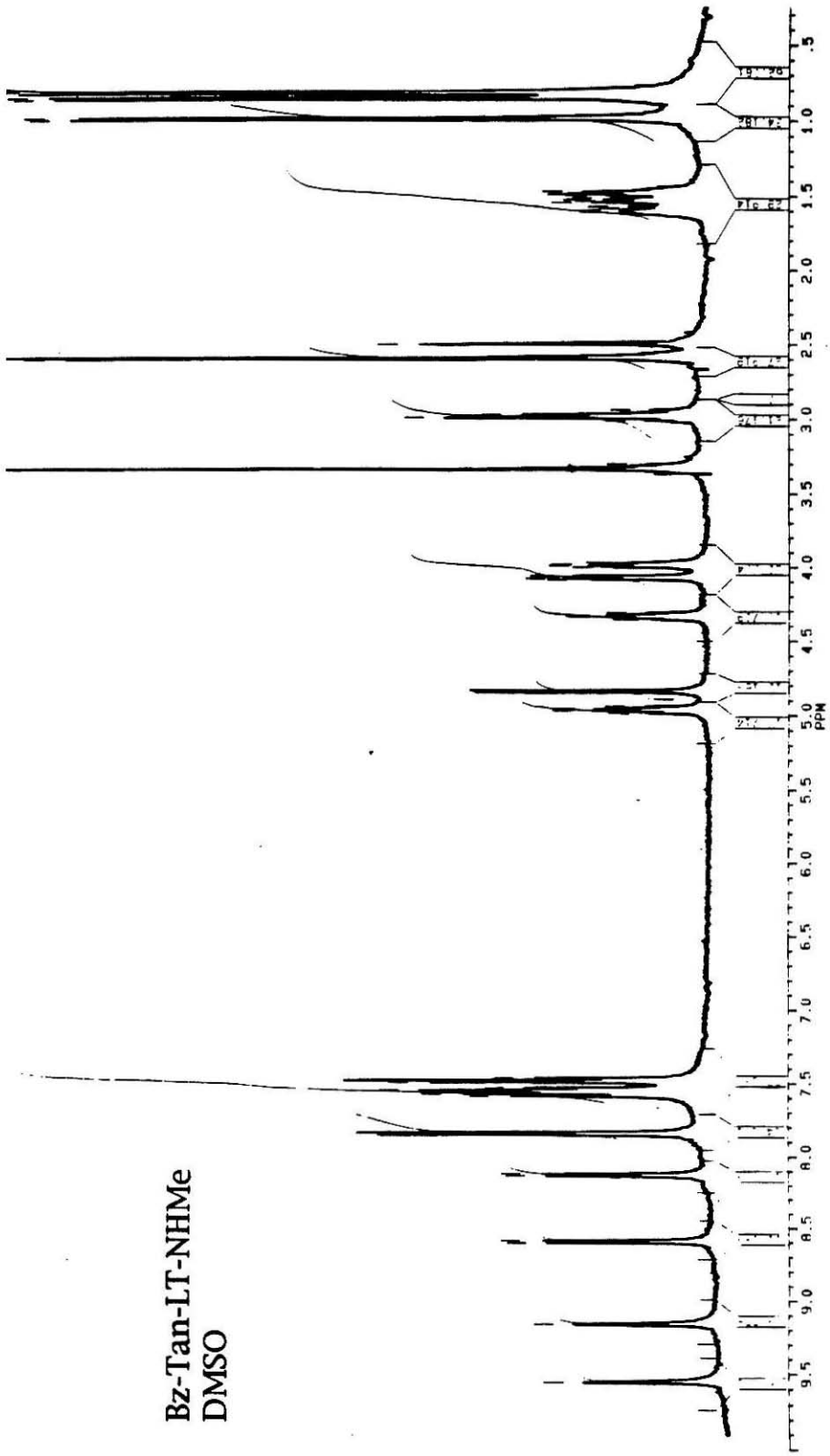



Bz-NL(allo)T-NHMe
DMSO
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
PPM

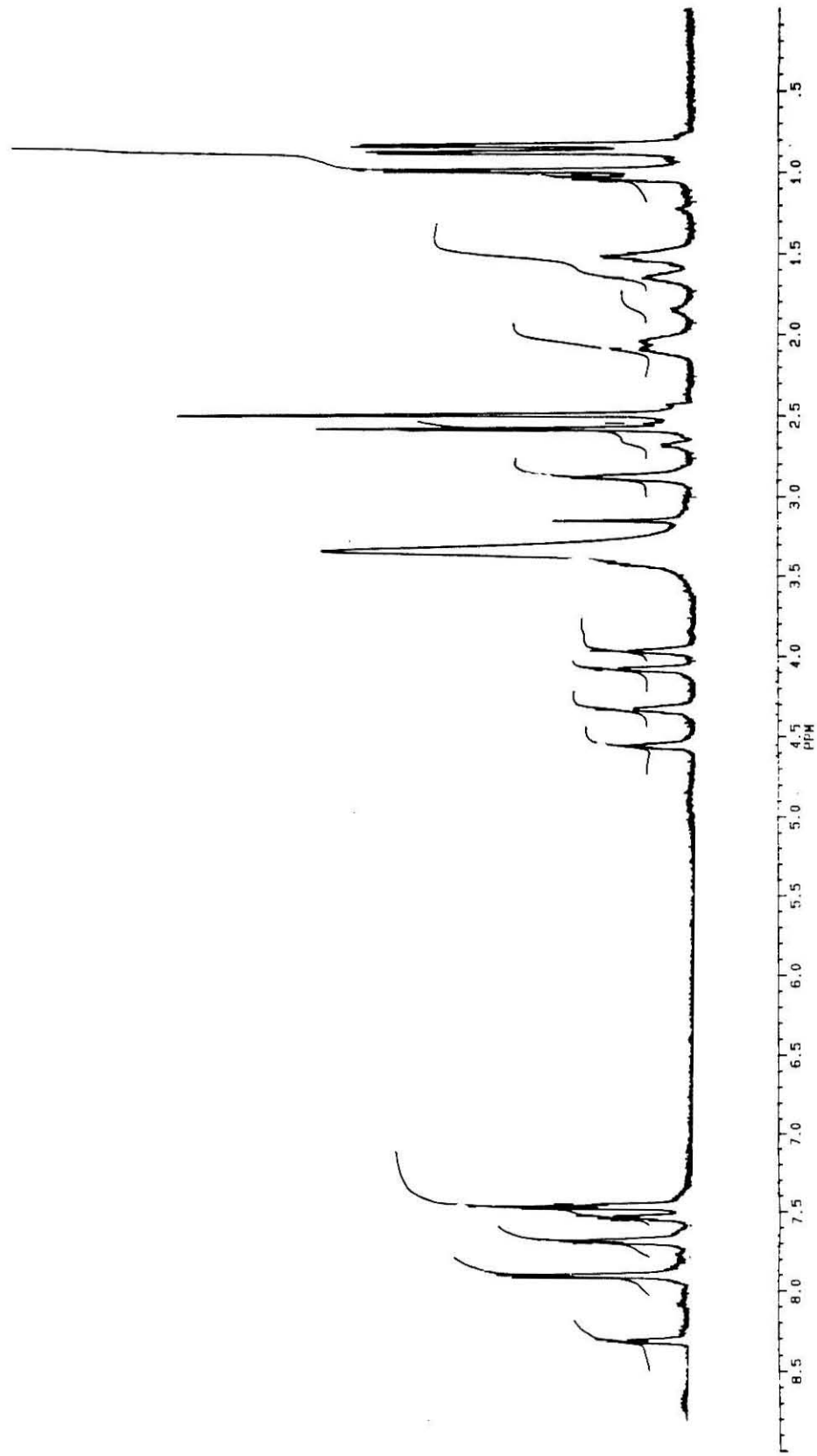

Bz-DLT-NHMe
DMSO
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
PPM
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5


Bz-D(γOMe)LT-NHMe
DMSO
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
PPM

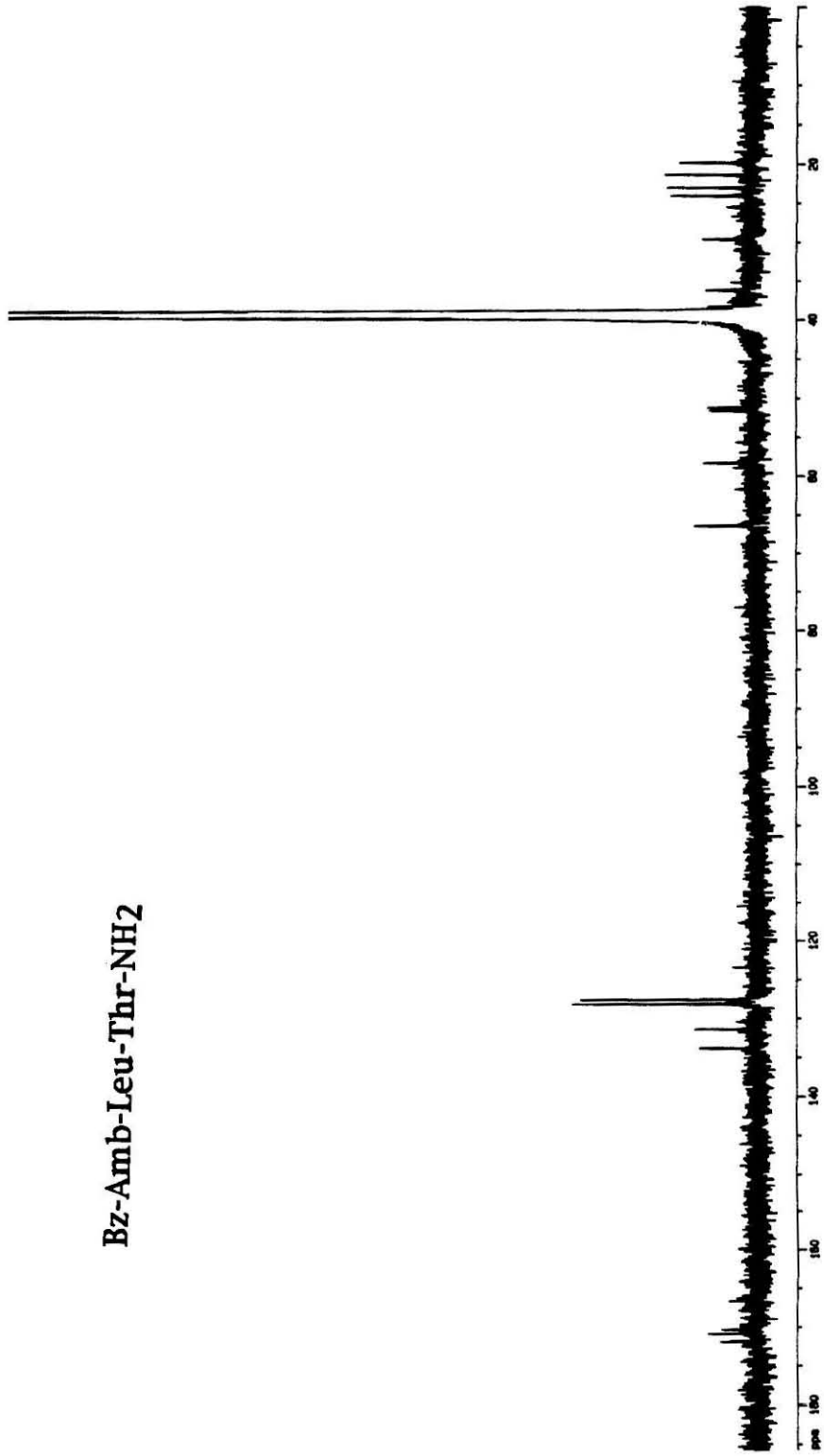

Bz-N(γNMe)LT-NHMe
DMSO

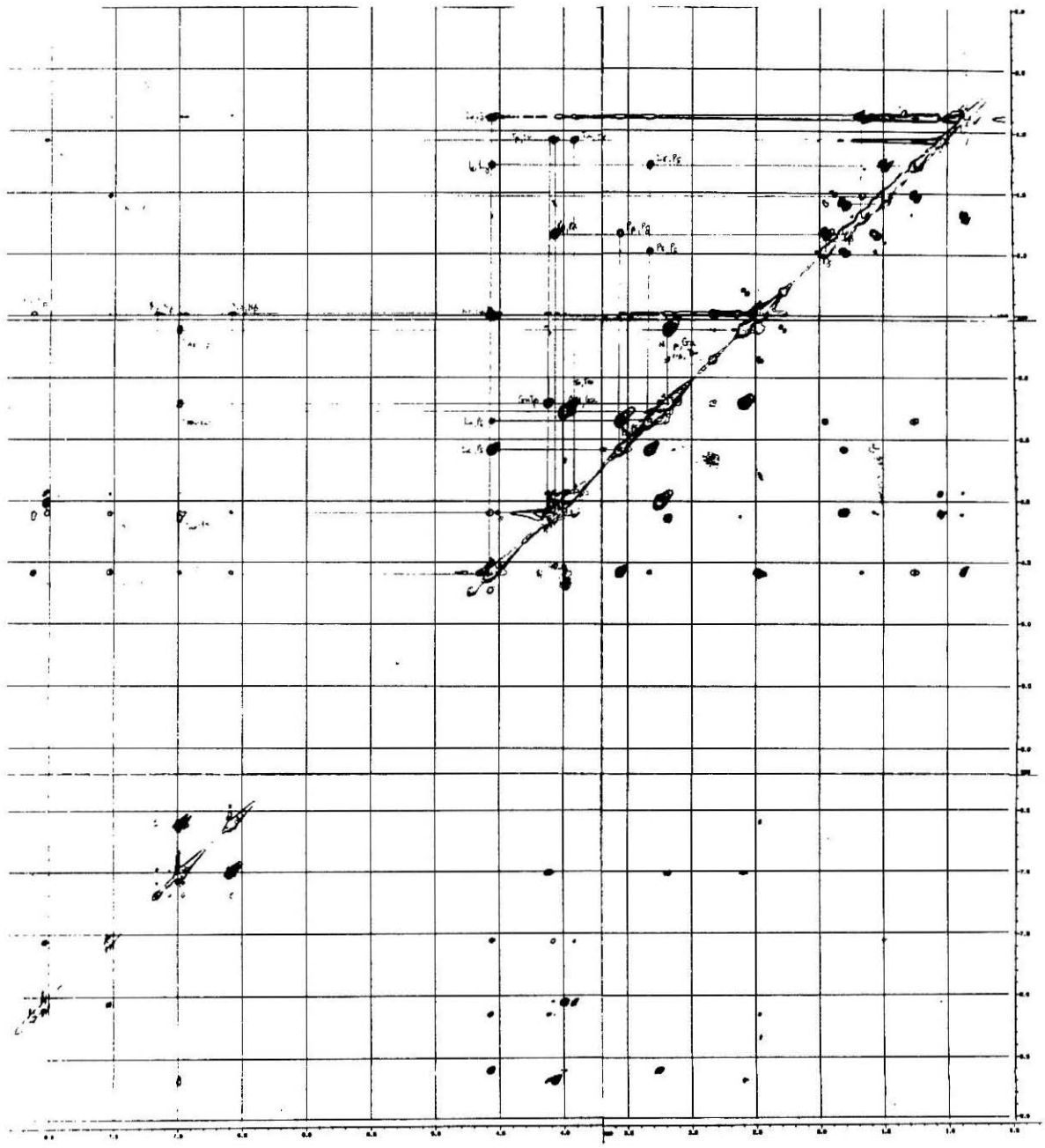

Bz-Tan-LT-NHMe
DMSO
PPM

Bz-Tan-LT-NHMe
DMSO

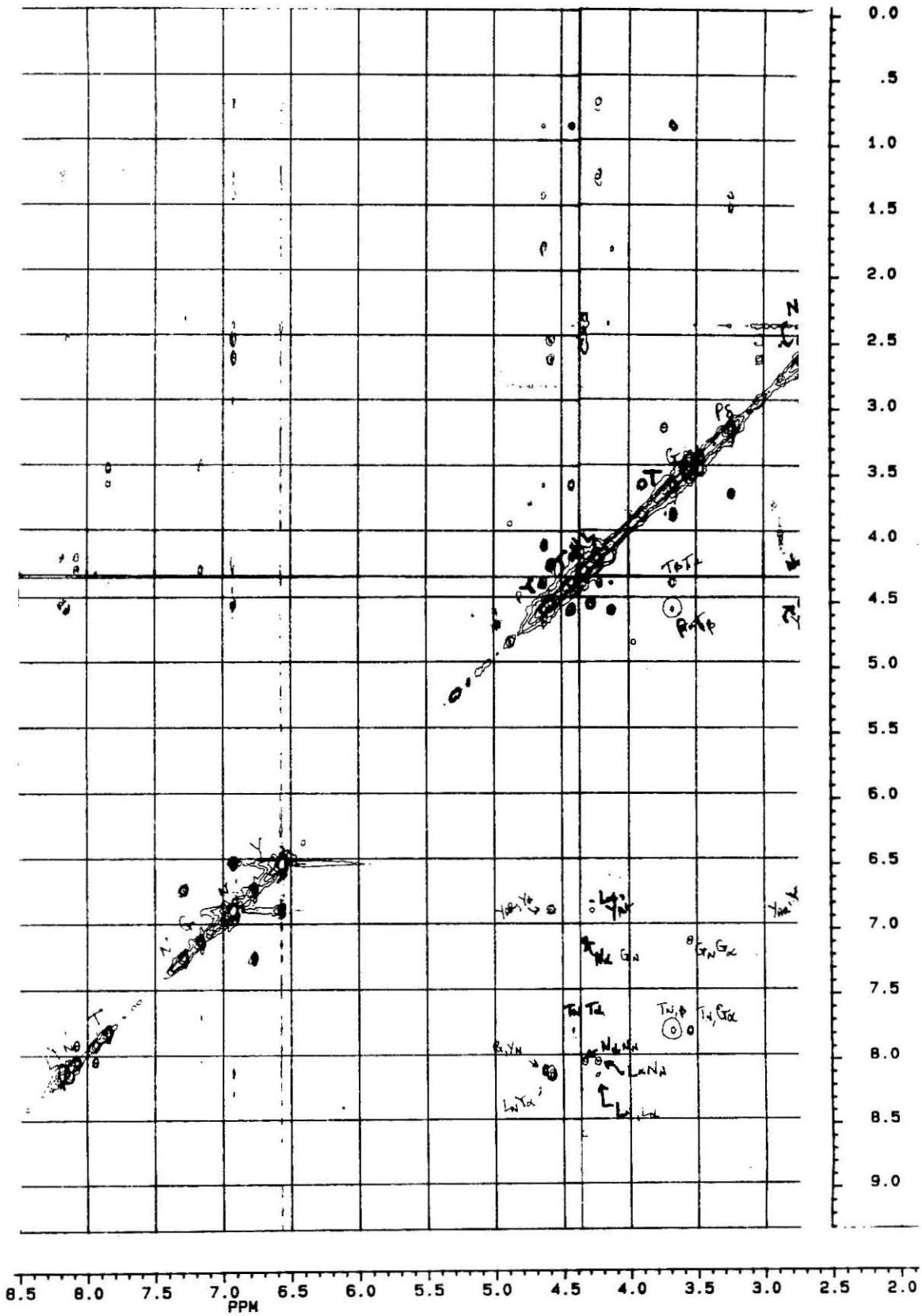

8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
PPM
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5

Bz-Amb-Leu-Thr-NH_2

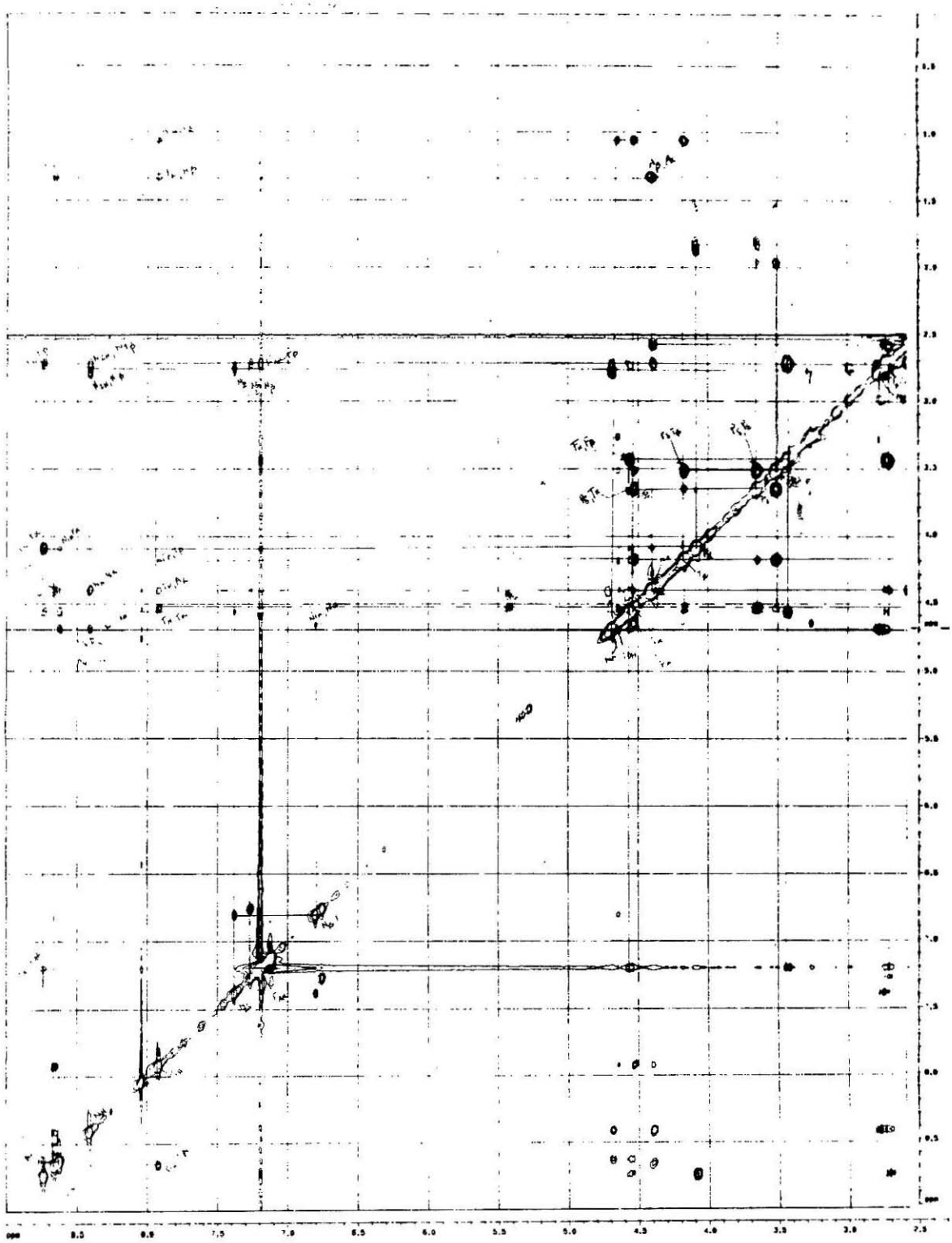




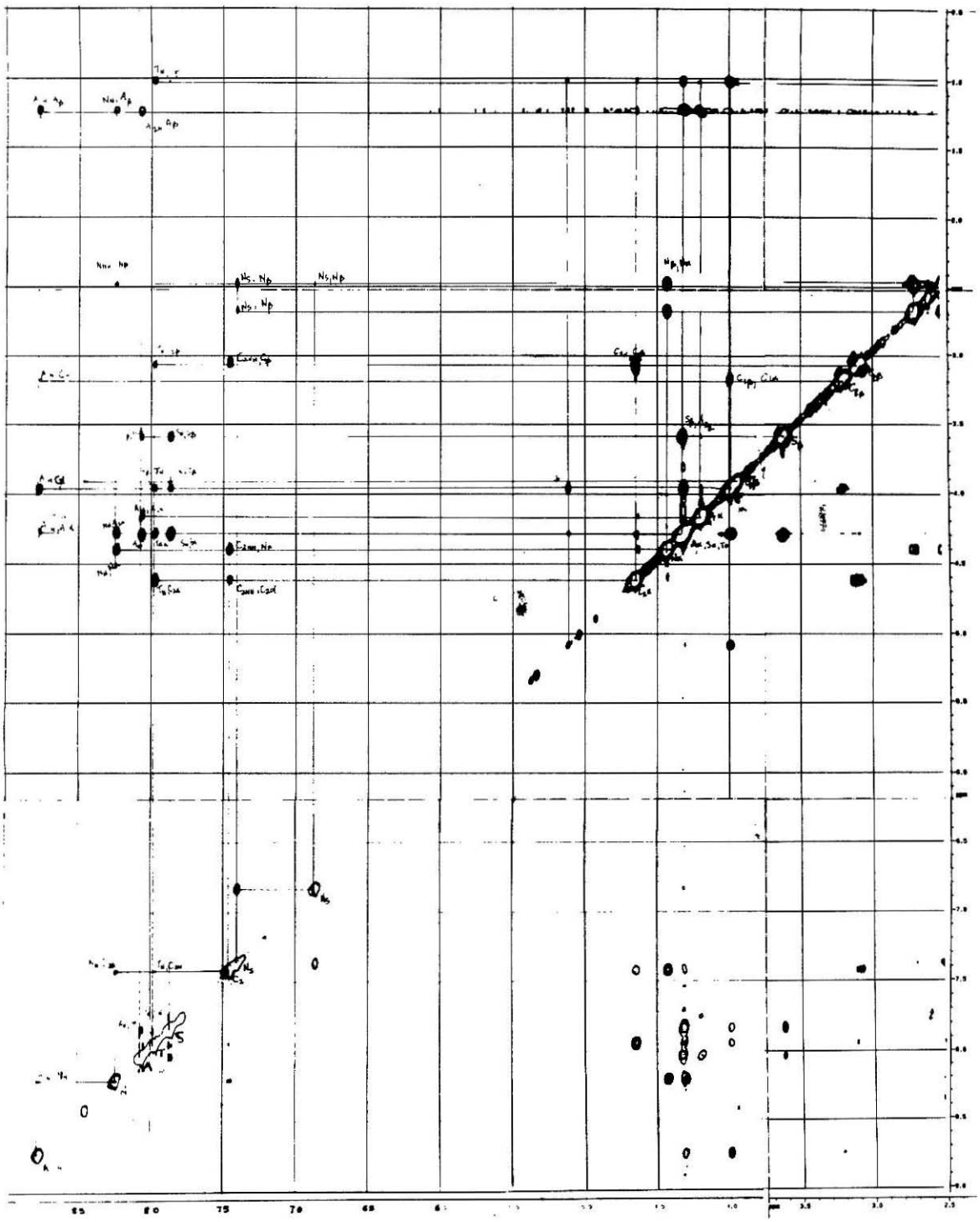
ROESY spectra
cyclo(PY$_d$NGTL)



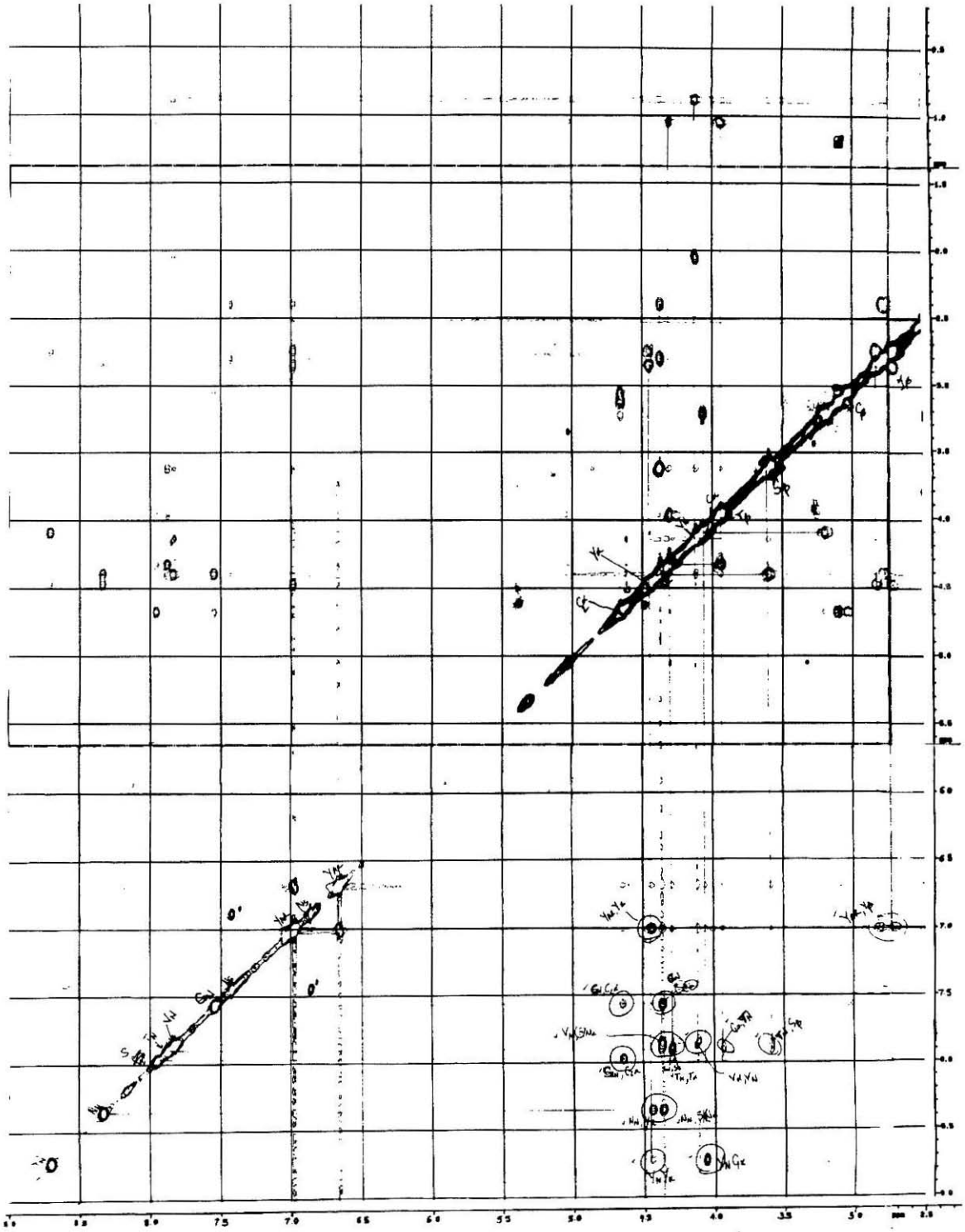
ROESY spectra
cyclo(PY$_d$LNGT)



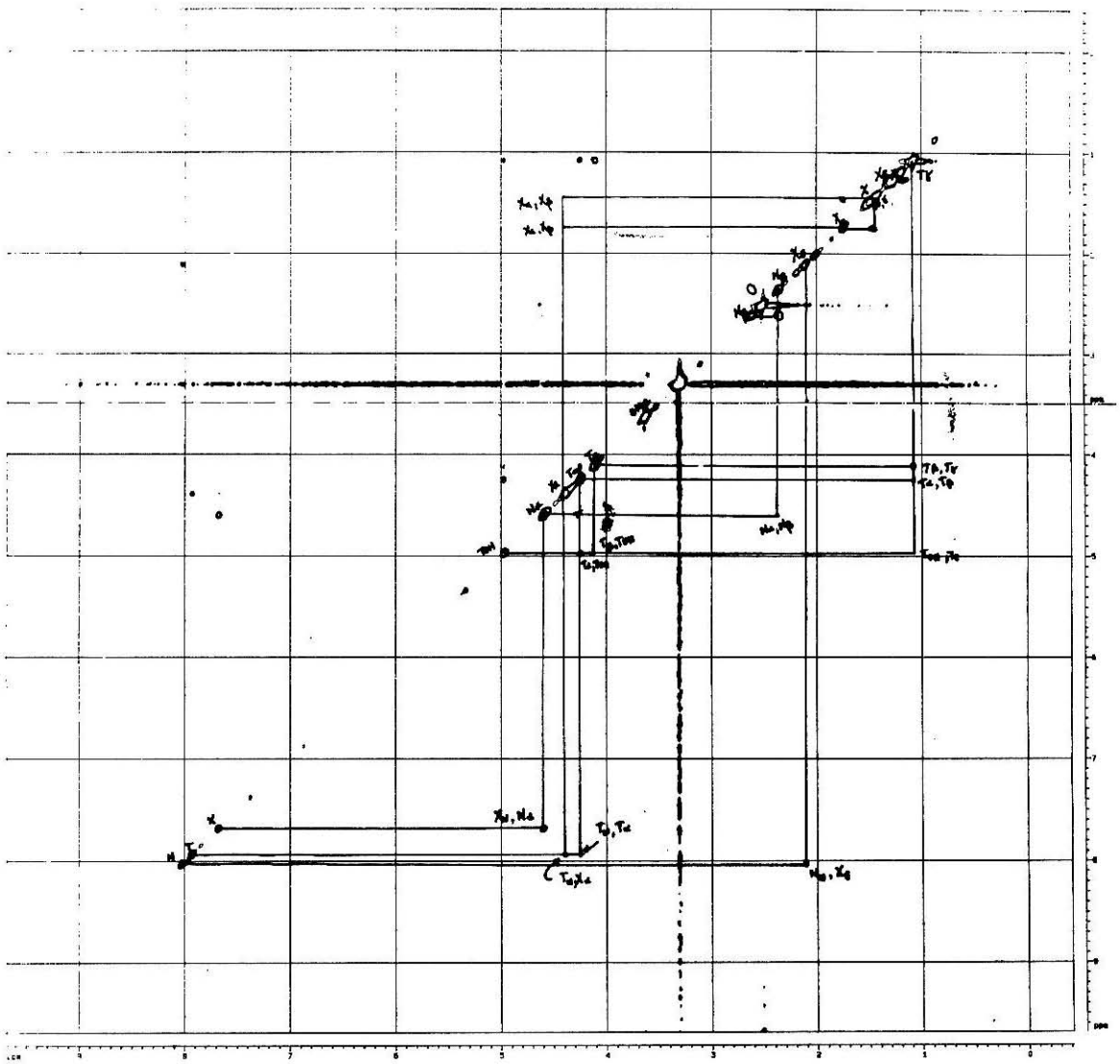
ROESY spectra
cyclo(PF$_d$NNAT)



ROESY spectra
cyclo(CANCTSA)



ROESY spectra
cyclo(CYNCTSV)



ROESY spectra
Pimelic macrocycle