



THESIS  
3  
2101

This is to certify that the  
dissertation entitled

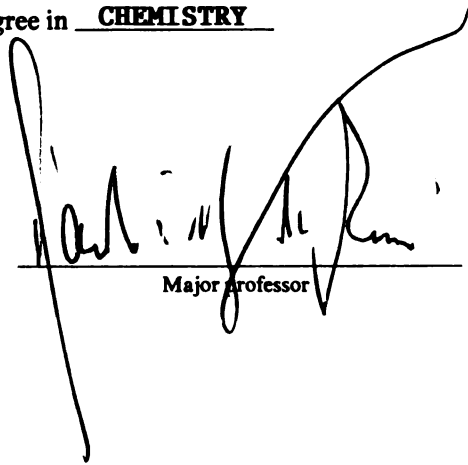
**STATISTICAL, COMPUTATIONAL, AND NMR  
SPECTROSCOPIC METHODS FOR PREDICTING  
CARBOHYDRATE CONFORMATIONS**

presented by

**HUSSEN MOHAMMED**

has been accepted towards fulfillment  
of the requirements for

PhD degree in CHEMISTRY



Major professor

Date 04/16/01

**LIBRARY**  
**Michigan State**  
**University**

**PLACE IN RETURN BOX** to remove this checkout from your record.  
**TO AVOID FINES** return on or before date due.  
**MAY BE RECALLED** with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**STATISTICAL, COMPUTATIONAL, AND NMR SPECTROSCOPIC METHODS  
FOR PREDICTING CARBOHYDRATE CONFORMATIONS**

**By**

**Hussen Mohammed**

**A DISSERTATION**

**Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of**

**DOCTOR OF PHILOSOPHY**

**Department of Chemistry**

**2001**

## **ABSTRACT**

### **STATISTICAL, COMPUTATIONAL, AND NMR SPECTROSCOPIC METHODS FOR PREDICTING CARBOHYDRATE CONFORMATIONS**

By

Hussen Mohammed

In the first part of this work, we report a simple rule based method for predicting the conformations of oligo- and polysaccharides. The method is based on an exhaustive analysis of the entries in the Cambridge Structural Database (CSD) of crystal structures. We analyzed the values of the critical interglycosidic dihedral angles, O5'-C1'-OX-CX ( $\phi$ ) and C1'-OX-CX-CX+1 ( $\psi$ ) from all of the entries which ranged from di-, tri-, tetra- and hexa- to octasaccharides. These results were used to construct a second database of dihedral angles. An almost exclusive preference for specific values of  $\phi$  and  $\psi$  for a specific type of linkage independent of the context was clearly evident. There was symmetry of placement of the points on the  $\phi$ - $\psi$  plots that was determined by the absolute configuration of the anomeric carbon and the carbon to which the linkage is made. These corresponded to the RR, SS, SR, and RS configurations. Three simple and accurate rules are proposed to help predict the conformations of oligo- and polysaccharides. The results support, expand and generalize earlier reports from other groups.

One of the important methods for determining the three-dimensional structures of carbohydrates is using the proton-carbon hetero-nuclear coupling

constant to calculate the critical glycosidic dihedral angle values. In this work we report the synthesis of seven compounds that have rigid glycosidic linkages. These compounds are important to parameterize a new Karplus-type equation. We used these seven compounds and three commercially available molecules to parameterize new Karplus-Type equations relating the vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , to the glycosidic dihedral angles  $\phi$  (H1'-C1'-OX-CX) and  $\psi$  (HX-CX-OX-C1'). The three dimensional structures of the synthetic model compounds **1-7** and the commercially available compounds **8-10** have been refined by Molecular Mechanics and AM1 and PM3 Semi-empirical calculations. The vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , have been measured in solution by two-dimensional excitation-sculptured indirect-detection experiment (EXSIDE) NMR spectroscopy. From the least squares fitting of the measured coupling constants and the glycosidic dihedral angles of the rigid molecules, new Karplus-type equations of the form:  $^3J_{\text{HC}}(\theta) = a \text{Cos}^2(\theta) + b \text{Cos}(\theta) + c$ , were obtained where  $\theta$  stands for the dihedral angles defined by the four atoms H-C-O-C ( $\phi$  and  $\psi$  for carbohydrates).

Finally, we applied the methods developed in this work in combination with existing methods using multi-nuclear, multidimensional NMR spectroscopy, molecular mechanics calculations and molecular dynamics simulations to study the solution conformations and dynamics of the disaccharide maltose and the tetrasaccharide reduced-maltotetraose, and the pseudo-tetrasaccharide acarbose, which is a potent glycohydrolase inhibitor.

**Copyright by  
Hussen Mohammed  
2001**

*To my son, Alula  
my wife, Elizabeth  
and my mother, Nuria*



## **ACKNOWLEDGMENTS**

First I would like to thank my thesis advisor, Dr. Rawle Hollingsworth, for his guidance and patience. I thank Rawle very much for teaching me how to approach problems with a positive attitude. Without his guidance, the completion of my dissertation would have been impossible.

I also want to thank my committee members, Drs. B. Borhan, J. F. Harrison, T. J. Pinnavaia and D. L. Ward and my former committee member Dr. C. K. Chang for their encouragement and for providing me helpful critiques of my dissertation. I deeply thank Drs. Rawle I. Hollingsworth, James F. Harrison, Eugene LeGoff, and Ardeshir Azadnia for writing reference letters.

I am indebted to the Max T. Rogers NMR Facility (Mr. Kermit Johnson and Dr. Long Lee), the Mass Spectrometry Facility and the Computational Chemistry Facilities of the Chemistry and Biochemistry Departments. Especially, I want to thank Kermit for helping me with the multidimensional NMR experiments.

I also want to thank the Hollingsworth group members especially for the fruitful discussions and for making the lab a lively place to work.

I would like to show gratitude to my mother Nuria and my late father Mohammed, to my sisters, Radia, Halima, Sophia and Abrehet and to my brothers, Jemal, Ali, Suleyman, Abdu, Kedir, and Hafiz for their encouragement, love and support. I want to thank my mother for teaching me that education is an important part of life.

Finally, my deepest thanks goes to my wife, Elizabeth Yohannes, and my son, Alula Hussen, for their support, encouragement, sacrifices, and unconditional love.

Thank you all

*Hussen*

## TABLE OF CONTENTS

List of Tables.....	x
List of Figures.....	xiii
List of Schemes.....	xviii
List of Abbreviations.....	xix

### CHAPTER 1

Introduction to Conformational Analysis of Carbohydrates.....	1
Introduction.....	2
The Anomeric and Exo-anomeric Effects.....	9
Crystal Structures from X-ray Crystallography.....	13
Solution Conformations from NMR spectroscopy.....	14
Molecular Mechanics Calculations and Molecular Dynamics Simulations of Carbohydrates.....	21
References.....	26

### CHAPTER 2

A Simple and Accurate Rule-based Method for Predicting the Conformations of Oligo- and Polysaccharides.....	28
Abstract.....	29
Introduction.....	30
Results and Discussion.....	33
Conclusions.....	48
Acknowledgements.....	48
References.....	49

### CHAPTER 3

Design and Synthesis of Molecules with Rigid Glycosidic Linkages.....	58
Abstract.....	59
Introduction.....	60
Design and Synthesis.....	60
Conclusions.....	88
Experimental.....	89
Acknowledgements.....	99
References.....	100

## CHAPTER 4

<b>New Karplus-Type Equations Relating the Glycosidic Dihedral Angles and Vicinal Proton-Carbon Coupling Constants.....</b>	<b>101</b>
<b>Abstract.....</b>	<b>102</b>
<b>Introduction.....</b>	<b>103</b>
<b>Materials and Methods.....</b>	<b>107</b>
<b>Results and Discussion.....</b>	<b>113</b>
<b>Conclusions.....</b>	<b>133</b>
<b>Acknowledgements.....</b>	<b>134</b>
<b>References.....</b>	<b>135</b>

## CHAPTER 5

<b>Studies on the Solution Conformation and Dynamics of Maltose, Reduced Maltotetraose and the Potent Glycohydrolase Inhibitor Acarbose.....</b>	<b>137</b>
<b>Abstract.....</b>	<b>138</b>
<b>Introduction.....</b>	<b>139</b>
<b>Materials and Methods.....</b>	<b>142</b>
<b>Results and Discussion.....</b>	<b>149</b>
<b>Conclusions.....</b>	<b>173</b>
<b>Acknowledgements.....</b>	<b>174</b>
<b>References.....</b>	<b>175</b>

## LIST OF TABLES

<b>Table 1.1.</b>	The solution conformations of the tetrasaccharide $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-1-OPr and its component di- and trisaccharides.....	18
<b>Table 1.2.</b>	The solution conformations of the tetrasaccharide $\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-(1 $\rightarrow$ 3)-[ $\beta$ -D- GlcNAc-(1 $\rightarrow$ 4)-] $\beta$ -D-Man and its component di- and trisaccharides.....	19
<b>Table 1.3.</b>	The solution conformations of the branched tetrasaccharide $\beta$ -D-Xyl-(1 $\rightarrow$ 2) ][ $\alpha$ -D-Man-(1 $\rightarrow$ 3)-][ $\alpha$ -D-Man-( 1 $\rightarrow$ 6)-] $\beta$ -D-Man-OMe and its component di- and trisaccharides.....	20
<b>Table 1.4.</b>	The solution conformations of a family of trisaccharides.....	21
<b>Table 2.1.</b>	The percentages of the total phi-psi space that is populated by oligosaccharide crystals from Cambridge Structural Database .....	35
<b>Table 2.2.</b>	Results of the statistical analysis of the dihedral angles ( $\phi$ and $\psi$ ) of some of the oligosaccharides from the Cambridge Structural Database.....	40
<b>Table 4.1.</b>	$^1\text{H}$ and $^{13}\text{C}$ chemical shift assignments and proton-proton coupling constants of 1,6-anhydroglucopyranose ( <b>8</b> ).....	110
<b>Table 4.2.</b>	$^1\text{H}$ and $^{13}\text{C}$ chemical shift assignments and proton-proton coupling constants of 1,6-anhydrogalactopyranose ( <b>9</b> ).....	111

<b>Table 4.3.</b>	$^1\text{H}$ and $^{13}\text{C}$ chemical shift assignments and proton-proton coupling constants of 1,6-anhydromannopyranose ( <b>10</b> ).....	112
<b>Table 4.4.</b>	Heteronuclear three-bond proton-carbon coupling constants of compounds <b>1-5</b> measured from 2D EXSIDE NMR experiments and the corresponding dihedral angles (torsions). ....	115
<b>Table 4.5.</b>	Heteronuclear three-bond proton-carbon coupling constants of compounds <b>6-10</b> measured from 2D EXSIDE NMR experiments and the corresponding dihedral angles (torsions). ....	116
<b>Table 4.6.</b>	Structures of 1,6-anhydroglucose ( <b>8</b> ) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations.....	119
<b>Table 4.7.</b>	Structures of 1,6-anhydrogalactose ( <b>9</b> ) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations.....	120
<b>Table 4.8.</b>	Structures of 1,6-anhydromannose ( <b>10</b> ) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations.....	121
<b>Table 4.9.</b>	The values of the dihedral angle C2-C3-C4-C5 as a measure of the degree of flattening of the chair conformations of the three anhydro sugars, 1,6-anhydroglucose ( <b>8</b> ), 1,6-anhydrogalactose ( <b>9</b> ), and 1,6-anhydromannose ( <b>10</b> ) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations.....	122

<b>Table 4.10.</b>	The values of the parameters a, b, c and the value of R <sup>2</sup> from a least squares fitting of the heteronuclear coupling constants and the dihedral angles given in Tables 4.4 and 4.5.....	131
<b>Table 5.1.</b>	<sup>1</sup> H and <sup>13</sup> C chemical shift assignments of maltose.....	144
<b>Table 5.2.</b>	<sup>1</sup> H and <sup>13</sup> C chemical shift assignments of reduced maltotetraose.....	145
<b>Table 5.3.</b>	<sup>1</sup> H and <sup>13</sup> C chemical shift assignments of acarbose.....	146
<b>Table 5.4.</b>	Distances calculated from NOESY experiment for reduced maltotetraose.....	147
<b>Table 5.5.</b>	Distances calculated from NOESY experiment for acarbose.....	148
<b>Table 5.6.</b>	Glycosidic dihedral angles from constrained minimization and grid search studies and from crystal structures.....	171
<b>Table 5.7.</b>	Comparison of the predicted vicinal proton-carbon coupling constants using the new Karplus-type equations with that of the three previous equations of Mulloy <i>et. al.</i> , Tvaroska <i>et. al.</i> , and Cloran <i>et. al.</i> with the experimentally measured values. The results are averages of the 2000 data points from Molecular Dynamics simulations.....	172

## LIST OF FIGURES

- Figure 1.1.** The aldotrioses D- and L-glyceraldehyde.....4
- Figure 1.2.** The structure and stereochemical relationships of D-aldoses having three to six carbon skeletons. The highest numbered asymmetric carbon determines the configuration of the aldose in each case.....5
- Figure 1.3.** Cyclization of D-glucose in aqueous solution resulting in the formation of two pyranoses, which show distinct physical, chemical and biological properties.....8
- Figure 1.4.** The  ${}^4C_1$  conformation of  $\beta$ -D-glucose and the  ${}^1C_4$  conformation of  $\beta$ -L-fucose.....8
- Figure 1.5.** The Twist and Envelope conformations of furanoses, with the twist conformation having three of the ring atoms in plane and the envelope conformation with four of the ring atoms in plane.....9
- Figure 1.6.** The anomeric effect that relates to the preference for the axial orientation of the aglycon of glycopyranosides and the exo-anomeric effect that relates to the preference for the aglyconic carbon to be in near syn-clinal orientation to both the ring oxygen and the anomeric hydrogen.....10
- Figure 2.1.** Structure of 1-O-methyl cellobioside and the glycosidic dihedral angles, phi and psi, which determine its conformation. Phi = O5'-C1'-O4-C4 and Psi = C1'-O4-C4-C5.....31



<b>Figure 2.2.</b>	Biograf rigid residue phi-psi maps of (a) cellobiose ( $\beta$ -D-Glc-(1->4)BD-Glc) and (b) melliobiose ( $\alpha$ -D-Gal-(1->6)- $\beta$ -D-Glc).....	36
<b>Figure 2.3.</b>	$\Phi, \Psi$ map of 1-2, 1-3, 1-4 and 1-6 linked oligosaccharides from the Cambridge Structural Database (CSD). Each point on the map corresponds to one glycosidic linkage from the CSD.....	41
<b>Figure 2.5.</b>	Line representation of the structure of dermatan 4-sulfate predicted using the rules in this paper and from X-ray fiber diffraction study. <sup>104</sup> .....	44
<b>Figure 2.6.</b>	Line representation of the structure of hyaluronan (sodium hyaluronate III) predicted using the rules in this paper and from X-ray fiber diffraction study. <sup>104</sup> .....	46
<b>Figure 3.1.</b>	The ten molecules that have rigid glycosidic dihedral angles. While molecules <b>8-10</b> are commercially available, molecules <b>1-7</b> are synthesized in this work.....	62
<b>Figure 3.2.</b>	Proton and carbon nuclear magnetic resonance spectra of 1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\alpha$ -L-fucopyranoside ( <b>1</b> ).....	66
<b>Figure 3.3.</b>	Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzyledene-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\alpha$ -D-glucopyranoside ( <b>2</b> ) .....	69
<b>Figure 3.4.</b>	Proton and carbon nuclear magnetic resonance spectra of 1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -L-fucopyranoside ( <b>3</b> ).....	73
<b>Figure 3.5.</b>	Proton and carbon nuclear magnetic resonance spectra of 3-O-methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-	

	glucopyranoside (4) .....	76
<b>Figure 3.6.</b>	Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzyledene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (5).....	80
<b>Figure 3.7.</b>	Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzyledene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (6) .....	83
<b>Figure 3.8.</b>	Proton and carbon nuclear magnetic resonance spectra of 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (7) .....	86
<b>Figure 4.1.</b>	The ten molecules that have rigid glycosidic dihedral angles. While molecules 8-10 are commercially available, molecules 1-7 are synthesized in this work (Chapter 3).....	105
<b>Figure 4.2.</b>	The new data points plotted against the Karplus-type curves of the previous equations.....	124
<b>Figure 4.3.</b>	The new AM1 data points plotted against the Karplus-type curves of the previous equations and the new equations. ....	126
<b>Figure 4.4.</b>	The new PM3 data points plotted against the Karplus-type curves of the previous equations and the new equations.....	128
<b>Figure 4.5.</b>	The difference in the predicted values of the vicinal coupling constants between the new Karplus-type equations (AM1* and PM3*) and the two previous equations (Tvaroska <i>et. al</i> and Mulloy <i>et. al.</i> ).....	130

**Figure 5.1.** Schematic diagram of maltose, reduced maltotetraose, and acarbose. ....140

**Figure 5.2.** Rigid residue potential energy (PE) surfaces of maltose (top) and acarviosine (the a-b unit of acarbose) (bottom). Isoenergy contours are drawn with interpolation of 1 kcal/mol above the minimum. ....151

**Figure 5.3.** The molecular dynamics population map of maltose, shown above, is consistent with the PE surface. The dihedral angles are defined as  $(O5'-C1'-O4-C4)$  and  $(C1'-O4-C4-C5)$ .....153

**Figure 5.4.** Solution conformations of maltose, reduced maltotetraose and acarbose as determined from constrained minimization using distance constraints from 2D NOESY experiments.....156

**Figure 5.5.** Molecular dynamics population maps of reduced maltotetraose from 200 ps simulation starting from its constraint-minimized structure. The plots are for the glycosidic linkages a-b, b-c, and c-d.....157

**Figure 5.6.** Molecular dynamics trajectories of the phi dihedral angle of reduced maltotetraose from 200 ps simulation starting from its constraint-minimized structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom).....159

**Figure 5.7.** Molecular dynamics trajectories of the psi dihedral angle of reduced maltotetraose from 200 ps simulation starting from its constraint-

minimized structure. The plots are for the glycosidic linkages a-b, b-c, and c-d.....160

**Figure 5.8.** Molecular dynamics population maps of acarbose from 200 ps simulation starting from its constraint-minimized structure. The plots are for the glycosidic linkages a-b, b-c, and c-d.....164

**Figure 5.9.** Molecular dynamics trajectories of the phi dihedral angle of acarbose from 200 ps simulation starting from its constraint-minimized structure. The plots are for the glycosidic linkages a-b, b-c, and c-d.....166

**Figure 5.10.** Molecular dynamics trajectories of the psi dihedral angle of acarbose from 200 ps simulation starting from its constraint-minimized structure. The plots are for the glycosidic linkages a-b, b-c, and c-d.....168

## LIST OF SCHEMES

<b>Scheme 3.1</b> Synthesis of 1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\alpha$ -L-fucopyranoside (1).....	65
<b>Scheme 3.2</b> Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\alpha$ -D-glucopyranoside (2) .....	68
<b>Scheme 3.3</b> Synthesis of 1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -L-fucopyranoside (3) .....	72
<b>Scheme 3.4</b> Synthesis of 3-O-methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (4) .....	75
<b>Scheme 3.5</b> Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (5).....	79
<b>Scheme 3.6</b> Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (6) .....	82
<b>Scheme 3.7</b> Synthesis of 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (7) .....	85

## LIST OF ABBREVIATIONS

Ac	Acetyl
Ac <sub>2</sub> O	Acetic anhydride
d	Doublet
dd	Doublet of doublet
δ	Chemical Shift
DMF	N,N-Dimethyl formamide
2-D NMR	Two-dimensional Nuclear Magnetic Resonance
COSY	Correlated spectroscopy
eq.	Equivalent
EXSIDE	Excitation-Sculptured Indirect-Detection Experiment
FAB	Fast Atom Bombardment
HMQC	Heteronuclear Multi Quantum Coherence spectroscopy
HRMS	High Resolution Mass Spectrometry
Hz	Hertz
LAH	Lithium aluminum hydride
LRMS	Low Resolution Mass Spectrometry
m	Multiplet
MeOH	Methanol
MM	Molecular Mechanics
MD	Molecular Dynamics
m.p.	Melting point

<b>NMR</b>	<b>Nuclear Magnetic Resonance</b>
<b>n.O.e.</b>	<b>Nuclear Overhauser enhancement</b>
<b>NOESY</b>	<b>Nuclear Overhauser enhancement spectroscopy</b>
<b>Py</b>	<b>Pyridine</b>
<b>q</b>	<b>Quartet</b>
<b>r.t.</b>	<b>Room temperature</b>
<b>s</b>	<b>Singlet</b>
<b>TFA</b>	<b>Trifluoroacetic acid</b>
<b>THF</b>	<b>Tetrahydrofuran</b>
<b>TLC</b>	<b>Thin Layer Chromatography</b>
<b>TOCSY</b>	<b>Total Correlated Spectroscopy</b>
<b>pTSA</b>	<b>p-Toluene sulfonic acid</b>

## **Chapter 1**

### **Introduction to Conformational Analysis of Carbohydrates**



## Introduction

Carbohydrates, lipids, DNA, and proteins are four major classes of macromolecules in biology. While proteins are linked by amide bonds and nucleic acids by 3'-5' phosphodiester bonds, carbohydrates can be connected by many different linkage types. In addition, while proteins and nucleic acids are almost exclusively linear, carbohydrates can be highly branched. This complexity allows carbohydrates to provide almost unlimited variations in their structure.

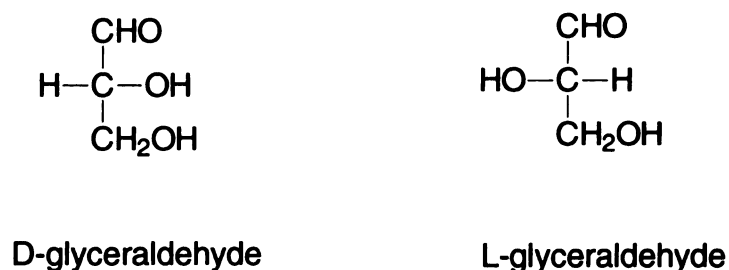
Carbohydrates are commonly classified as monosaccharides, disaccharides, trisaccharides, oligosaccharides and polysaccharides. A monosaccharide is that unit which cannot be further hydrolyzed into smaller carbohydrates. Di-, tri-, tetra-, and pentasaccharides can be hydrolyzed to two, three, four, and five monosaccharide units, respectively. Glycoconjugates are carbohydrates that are covalently attached to other biomolecules like proteins (glycoproteins) and lipids (glycolipids and lipopolysaccharides).

Monosaccharides can be represented by the formula  $C_nH_{2n}O_n$  or  $C_n(H_2O)_n$ . The latter representation led to the name "carbohydrates" or hydrates of carbon. Monosaccharides are further classified as triose, tetrose, pentose, hexose, heptose, and octose depending on the number of atoms in the carbon skeleton. Those monosaccharides with an aldehyde functional group are called aldoses and those with a ketone functional group are called ketoses. The number of carbon atoms and the position of the carbonyl group are both included in the general terms used for identifying the monosaccharides. Aldohexose refers, for

example, to six-carbon sugars with an aldehyde functional group like glucose or galactose. And ketoses are named by adding the suffix “ulose” to the base name, which indicates the number of carbon atoms. Laevulose, commonly known as fructose, is a hexulose containing six carbon atoms and a carbonyl group at position 2. The simplest monosaccharides are water soluble, and most taste sweet.

Glyceraldehyde, an aldotriose, is the smallest monosaccharide. When the central carbon of glyceraldehyde has the R configuration, it is referred as D-glyceraldehyde and when it has the S configuration it is referred as L-glyceraldehyde (**Figure 1.1**). The higher aldoses derived from D-glyceraldehyde belong to the D-series and those derived from L-glyceraldehyde belong to the L-series. D-Sugars are the most common form of sugars in nature, in contrast to the amino acids that commonly exist as L-amino acids in nature. L-Sugars such as L-rhamnose, L-arabinose, and L-fucose also exist in nature. The higher aldoses and ketoses are derived from their respective aldotrioses and ketotetroses by inserting one or more hydroxymethylene (HCOH) groups between the first chiral center and the carbonyl group of the corresponding isomer (see **Figure 1.2** for D-aldoses). A new chiral center is created by each insertion of hydroxymethylene group, thus creating two isomers. In a given monosaccharide with  $n$  chiral centers, there are  $2^n$  isomers or  $2^{n-1}$  enantiomeric pairs. An aldohexose, for example, has  $2^4 = 16$  isomers considering both D and L forms.

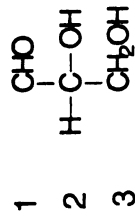
The mnemonic, familiarized by Louis and Mary Fieser of Harvard University, “all altruists gladly make gum in gallon tanks” is a very convenient way to remember the eight aldohexoses. Each word in the mnemonic symbolizes one enantiomer of the hexoses. And the mnemonic “Right AXLe” is a convenient way to remember the aldopentoses, where the four capitalized letters denote the first letter of the four aldopentoses (**Figure 1.2**).



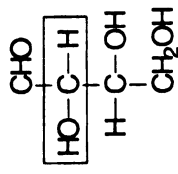
**Figure 1.1** The aldotrioses D- and L-glyceraldehyde.

The dominant forms of monosaccharides in solution are the cyclic hemiacetal forms. Sugar hemiacetals are formed by addition of an alcohol to an aldehyde or a ketone group resulting in anomerization to give two isomers  $\alpha$  and  $\beta$ , where the hydroxyl group at the anomeric carbon is axial or equatorial, respectively (**Figure 1.3**). Aldoses with four or more carbon atoms and ketoses with five or more carbon atoms can form five membered rings, furanose structures. And aldoses with five or more carbon atoms and ketoses with six or more carbon atoms can form six membered rings, pyranose structures. These

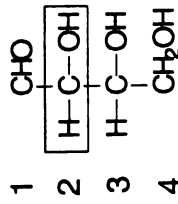
**Figure 1.2.** The structures and stereochemical relationships of D-aldoses that have three to six carbon skeletons. The highest numbered asymmetric carbon determines the configuration of the aldose in each case.



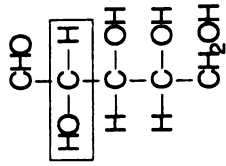
D-glyceraldehyde



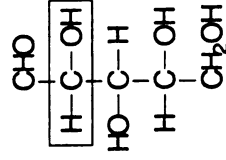
D-Threose



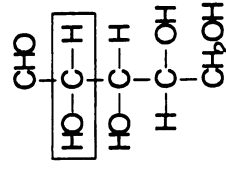
D-Erythrose



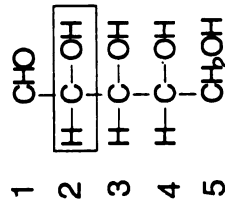
D-Arabinose (Ara)



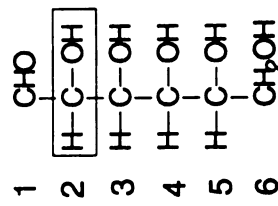
D-Xylose (Xyl)



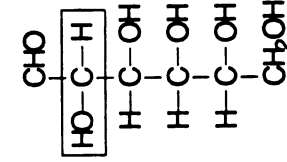
D-Lyxose (Lyx)



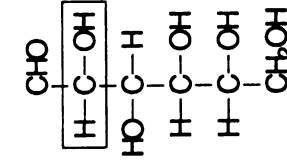
D-Ribose (Rib)



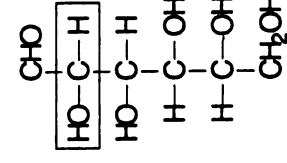
D-Allose (Alo)



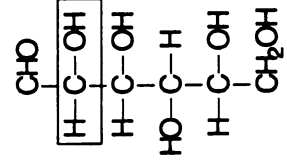
D-Altrose (Alt)



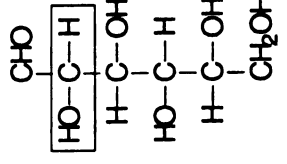
D-Glucose (Glc)



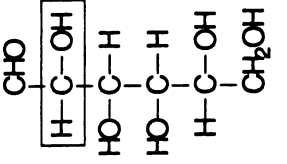
D-Mannose (Man)



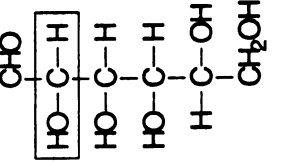
D-Gulose (Gul)



D-Idose (Ido)



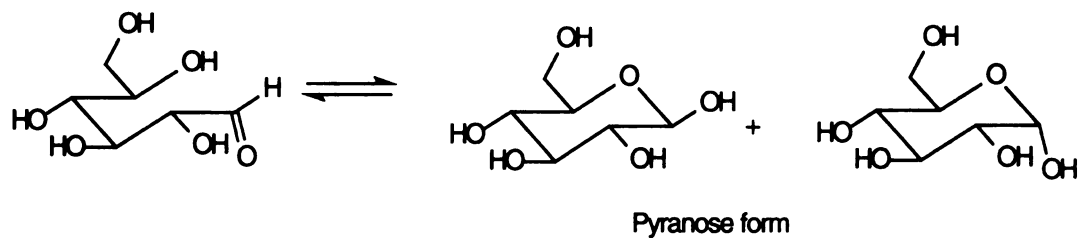
D-Galactose (Gal)



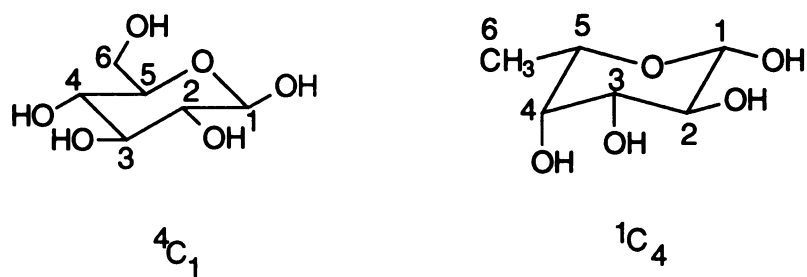
D-Talose (Tal)

tautomers show distinct physical, chemical and biological properties. The composition of sugars in solution, which can be estimated by NMR spectroscopy, depends on the nature of the solvent, the temperature, and the nature of the substituents at the different carbon atoms. Optical rotation measurement can also be used to differentiate between the two anomers. In the D-series, the more dextrorotatory anomer of each sugar is designated  $\alpha$  and the less dextrorotatory is designated  $\beta$ .<sup>1</sup> A freshly prepared solution of  $\alpha$ -D-glucose, for example, has an optical activity,  $[\alpha]_D^{20}$ , of  $+112.2^\circ$  while that of  $\beta$ -D-glucose is  $+18.7^\circ$ . The formation of cyclic sugars creates an additional chiral center at the anomeric carbon, C1 in aldoses and C2 in ketoses. Although Haworth projection formulas of sugars are still used by some, they are somewhat misleading because they suggest that the five and six membered rings are planar which is not the case. For aldohexoses a more clear representation is the chair conformation. There are two possible chair conformations designated  ${}^4C_1$  and  ${}^1C_4$  (**Figure 1.4**) where C stands for chair. In the notation  ${}^4C_1$ , for example, the superscript 4 stands for the atom puckered above the plane of the chair while the subscript 1 stands for the atom puckered below the plane of the chair.

The furanose form is another form, which is found in nature for some of the monosaccharides such as the aldopentoses ribose and deoxyribose in ribonucleic acids (RNA) and deoxyribonucleic acids (DNA). The furanose rings can have two types of conformations, the envelope (E) and the twist (T). In the twist conformation, three of the ring atoms are in plane and one ring atom is



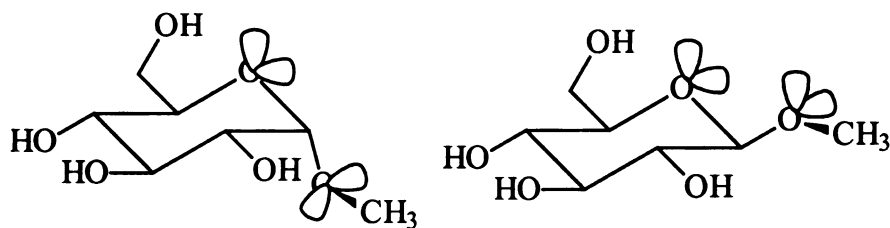
**Figure 1.3** Cyclization of D-glucose in aqueous solution resulting in the formation of two pyranoses, which show distinct physical, chemical and biological properties.



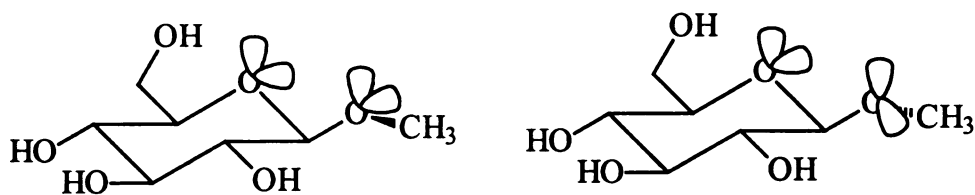
**Figure 1.4** The  ${}^4C_1$  conformation of  $\beta$ -D-glucose (left) and the  ${}^1C_4$  conformation of  $\beta$ -L-fucose.







Axial OMe at the anomeric position preferred over equatorial OMe due to the anomeric effect



Syn-clinal methyl at the anomeric oxygen due to the exo-anomeric effect

**Figure 1.6** The anomeric effect that relates to the preference for the axial orientation of the aglycon of glycopyranosides and the exo-anomeric effect that relates to the preference for the aglyconic carbon to be in near syn-clinal orientation to both the ring oxygen and the anomeric hydrogen.

periplanar to the neighboring C-O bond. Physically there is no difference between the anomeric and exo-anomeric effects.<sup>3</sup> The two different terms were used for the simple practical reason that the anomeric effect relates to the preference for the axial orientation of the aglycon of glycopyranosides while the exo-anomeric effect relates to the preference for the aglyconic carbon to be in near syn-clinal orientation to both the ring oxygen and the anomeric hydrogen (**Figure 1.6**). In the exo-anomeric effect the lone pairs of the exo-oxygen stabilize the ring C-O bond and hence favor syn-clinal methyl at the anomeric oxygen.

As can be seen from the discussions and figures in the preceding sections, carbohydrates can provide almost unlimited variations in their structure and generate a wide variety of complex structures because they can be attached to each other in many different linkage types and they can be highly branched. For example, if we consider a disaccharide of say D-glucose and D-galactose, it can have 32 different structures (each monomer can be  $\alpha$  or  $\beta$ , so the disaccharide can be  $\alpha\alpha$ ,  $\alpha\beta$ ,  $\beta\alpha$ , or  $\beta\beta$  and the two monomers can be linked in four different ways: 1->2, 1->3, 1->4, and 1->6 and the reducing end can be the glucose or the galactose residue). So, carbohydrates provide diverse three-dimensional (3D) structural possibilities and have several applications. For example, recent studies have demonstrated that oligosaccharides are involved in a number of recognition events such as cell adhesion, metastasis, fertilization and embryonic development, amongst others.<sup>4,5</sup> In view of their diverse function in a variety of biological systems, the primary structures and 3D structures of

carbohydrates are of great interest. Due to their diversity and the inability to crystallize such conjugates, the largest pure carbohydrates crystallized so far are tetrasaccharides,<sup>6-10</sup> nuclear magnetic resonance (NMR) spectroscopy in combination with molecular mechanics calculations, is the method of choice for their identification and conformational analysis.

A major problem when using NMR spectroscopy to study carbohydrates is that, except for the anomeric protons, all the protons on oxygenated atoms (e.g., those on carbons 2-6 in aldohexoses) appear in a range of ~3.2 – ~5.3 ppm. Hence, as the number of monosaccharide units increases, the assignment of peaks for each proton (and carbon) becomes increasingly difficult. This means extracting coupling constants and NOESY volumes--which are useful for determining dihedral angles and distances, respectively--becomes very difficult. Doing pure molecular mechanics and grid search studies of higher oligomers and polymers requires a lot of computational space and time. For example, consider a hexasaccharide that has ten glycosidic bonds and thus ten dihedral angles across the glycosidic linkages. To do a simple molecular mechanics grid search study of the hexasaccharide by 10° rotations of each dihedral angle independently involves 36 possible values for each dihedral angle. This gives, for the ten dihedral angles,  $36^{10} \approx 3.7 \times 10^{15}$  starting conformations to be evaluated by molecular mechanics. In addition, these calculations when used alone cannot give reliable or meaningful information on structure. So, *the question is how best to predict the three dimensional structure of carbohydrates.*

The focus of this work is therefore on developing methods that can be used to predict carbohydrate conformations. In the next sections of this chapter we will survey the three major current methods available for the conformational analysis of carbohydrates. These are X-ray crystallography, nuclear magnetic resonance spectroscopy, and computational methods.

## **Crystal Structure from X-ray Crystallography**

The most direct and powerful experimental tool to determine molecular structures and intermolecular interactions at atomic resolution is X-Ray diffraction from single crystals. Under crystallizing conditions helical structures invariably grow much faster along the helix axis than in the other directions. This means, it is seldom possible to achieve the isotropic growth rate necessary to produce single crystals in the case of helix forming polymers such as polysaccharides.

This is a major experimental setback with non-crystalline materials. However, it is often possible to prepare polycrystalline and/or oriented specimens in the form of fibers or films that are suitable for X-ray diffraction analysis. Because of this X-ray structures of several polysaccharides composed of simple to complex repeating units have been determined. These include the structural polysaccharides; cellulose, mannan, chitin, and xylan; the storage polysaccharide amylose and some of its derivatives: hyaluronan, chondroitin, keratan, and other polysaccharides in the glycosaminoglycan family; agarose, alginate, carrageenans, curdlan, gellan, and pectates, which are gel-formers; and branched polysaccharides such as galactomannan, welan, and xanthan.<sup>11</sup>

A library of the solid-state conformations of carbohydrates from X-ray crystallography studies will be built and statistically analyzed to come out with a simple and accurate rule-based method to predict carbohydrate conformations (Chapter two).

## **Solution Conformation from NMR Spectroscopy**

Nuclear magnetic resonance spectroscopy (NMR) in combination with molecular mechanics and molecular dynamics studies is one of the best known and widely used experimental methods for the conformational analysis of carbohydrates in solution. The first task here is to perform combinations of homonuclear- and heteronuclear-(one- and) two-dimensional NMR experiments in order to allow the assignments of the signals for the oligosaccharide. This include: double quantum filtered J-correlated spectroscopy (DQF-COSY) spectra,<sup>12</sup> total correlated spectroscopy (TOCSY) spectra,<sup>13</sup> heteronuclear multi-quantum coherence (HMQC) experiment,<sup>14</sup> heteronuclear multi-bond coherence (HMBC) experiment,<sup>15</sup> and nuclear Overhauser effect (n.O.e.) experiment .<sup>16</sup>

Analysis of the spin systems for the glycoside ring systems with 1-D traces from DQF-COSY and TOCSY spectra gives the connectivities. Once the assignments are made, n.O.e. experiments are done to determine the various inter-nuclear distances. Spectra of the n.O.e. experiments are acquired at different mixing times to address the potential problem of spin diffusion. Assuming isotropic motion then, the cross-relaxation rates and thus inter-nuclear distances between nuclei  $i$  and  $j$ ,  $r_{ij}$ , for each pair wise interaction will be

extracted from the volume of the NOESY peak,  $V_{ij}$ , measured at different mixing times using the relation:

$$r_{ij} = r_{\text{ref}} (V_{\text{ref}} / V_{ij})^{1/6} \quad (1)$$

where,  $r$  = distance,  $V$  = volume, ref = reference pair of nuclei the distance between which is fixed and known. This is followed by constrained minimization using the distances obtained at the previous step to determine the average conformation about glycosidic linkages in solution.

The second important structural information comes from the calculation of dihedral angles from coupling constants measured from NMR experiments. To calculate the dihedral angles between the glycoside residues from the vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , we use a 'Karplus type' equation<sup>17-19</sup> of the form:

$$^3J_{\text{CH}} = a \cos^2(\phi) + b \cos(\phi) + c \quad (2)$$

where  $^3J_{\text{CH}}$  is the three bond proton-carbon hetero-nuclear coupling constant and  $\phi$  is dihedral angle. The Haasnoot - De Leeuw - Altona relationship<sup>20</sup> can be used to calculate the  $\omega$  dihedral angle defined as H5-C5-C6-H6 between the glycoside residues from the coupling constants  $^3J_{\text{HH}}$  for 1→6 glycosidic linkages.

The final step is to use the distances and the angles obtained from the different NMR spectroscopy experiments as constraints in the force fields of

molecular mechanics calculations are performed using, for example, the Drieding force fields<sup>21</sup> implemented in the BIOGRAF program. The Drieding force field includes a harmonic term for the bond distortions, a simple harmonic function for the valence angle term, an improper torsion term to evaluate inversions about atomic centers, a cosine expansion torsional term, a Coulombic potential for evaluating electrostatic energy, a Leonard-Jones 12-6 potential for van der Waals contributions and a Leonard-Jones 12-10 potential for evaluating hydrogen bonding contributions.

NMR spectroscopy data represent an average conformational space when used for the determination of 3D structure. Often it is not possible to deduce the magnitude of internal motions from NMR data alone. To further complicate the matters, the average is not a linear average in the case of n.O.e. measurements, but weighed by the  $r^{-6}$  dependence of the n.O.e. up on the inter-nuclear distance.<sup>22</sup> This means that a single n.O.e. measurement is not a very precise tool. The measurement of an n.O.e. of magnitude  $n$  % could be interpreted in terms of a semi-rigid conformer where the n.O.e. corresponds to exactly to an inter-nuclear distance  $r$ , or equally, it could correspond to a situation where the two nuclei are in proximity for 10 % of the time, and a considerable distance apart for 90 % of the time, thus leading to a completely erroneous conclusions regarding the dominant conformer if interpreted as a semi-rigid body. In oligosaccharides, the number of constraints is small, and long-range constraints are almost never observed. Therefore, n.O.e. data should be used in combination with theoretical calculations. The combination of n.O.e. data and

theoretical predictions achieves several desirable purposes. First, the energy minimized structure obtained using the n.O.e. constraints is more likely to represent the global minimum energy structure than the use of purely theoretical energies, since the available conformational space is restricted by the additional constraints.<sup>23</sup> Second, molecular dynamics simulations<sup>24,25</sup> at ambient temperature (300K) can be run on the energy minimized structure in the presence and absence of n.O.e. constraints. The former allows one to determine whether an alternative, lower energy conformer exists by searching over the conformational space near the obtained minimum. The latter gives some measure of the extent of molecular motion in the sense that the extent of torsional oscillations about  $\phi$  and  $\psi$  for each linkage can be measured. This is important because the energy minimized structures representing the conformation at 0 K are not biologically important, and we are more interested in the 'conformation' at normal temperatures. Since this is a dynamic structure, the 'solution conformation' is defined as an average in Cartesian space together with root mean square deviations per atom with respect to the center of mass to gain some measure of mobility.

Another approach in combining n.O.e. data with theoretical studies involves calculation of the populations of conformers on the potential energy surface followed by computation of theoretical n.O.e.s from these conformers, and comparison with those generated experimentally.<sup>26,27</sup>

In the next few paragraphs we will discuss some examples of the solution conformations of oligosaccharides from the literature. These examples



demonstrate that the solution conformation of disaccharides seems to be preserved in structures containing three or more monosaccharide units.

The preferred solution conformations of  $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-1-OPr and its component di- and trisaccharides were determined by combination of n.O.e. and MM and MD studies<sup>28,29</sup> The results are summarized in **Table 1.1** where a, b, c and d denote the first, second, third and fourth monosaccharide units, respectively. The consistency of the values of the dihedral angles in the different oligosaccharides show that the conformations of the disaccharides a-b, b-c, and c-d is preserved in the trisaccharides a-b-c and b-c-d, and the tetrasaccharide a-b-c-d.

**Table 1.1.** The solution conformations of the tetrasaccharide  $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-1-OPr and its component di- and trisaccharides.

Molecule	$\phi/\psi$ of a-b	$\phi/\psi$ of b-c	$\phi/\psi$ of c-d
$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal a b	36/23		
$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc b c		39/20	
$\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal c d			28/33
$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc a b c	36/27	40/27	
$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal b c d		38/23	32/34
$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc – (1 $\rightarrow$ 3)- $\alpha$ -D-Gal a b c d	34/26	40/23	25/38

A similar study was done on  $\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-(1 $\rightarrow$ 3)-[ $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-] $\beta$ -D-Man and related di- and trisaccharides.<sup>30</sup> The results are summarized in **Table 1.2** where a, b, c and d denote the first, second, third and fourth monosaccharide units, respectively. Here again, the conformation of the individual disaccharides is consistent the same disaccharide in the trisaccharide and tetrasaccharide units. Note that, while the previous example contains only linear carbohydrate chains, the second example contains branched oligosaccharides (a-b-[c]-d).

**Table 1.2.** The solution conformations of the tetrasaccharide  $\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-(1 $\rightarrow$ 3)-[ $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-] $\beta$ -D-Man and its component di- and trisaccharides.

Molecule	$\phi/\psi$ of a-b	$\phi/\psi$ of b-d	$\phi/\psi$ of c-d
$\alpha$ -D-Man-(1 $\rightarrow$ 3)- $\beta$ -D-Man b d		-50/-10	
$\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-(1 $\rightarrow$ 3)- $\beta$ -D-Man a b d	40/30	-50/-20	
$\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-(1 $\rightarrow$ 3)-[ $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-] $\beta$ -D-Man a b c d	40/30	-50/-20	60/10

In the third example the solution conformation of three oligosaccharides have been carried out using rotating frame n.O.e. experiments in combination

with Hard Sphere Exo-Anomeric effect calculations<sup>31</sup> and the results are shown in **Table 1.3** where  $\phi'$  and  $\psi'$  are the dihedral angles defined by O5'-C1'-OX-CX and C1'-OX-CX-CX+1, respectively (where X= 2, 3 or 6). Except for the 1→6 linkage, which is more flexible than other linkage types due to its three bonds across the glycosidic linkage, the geometry of the other linkages has only one value and do not change much from one compound to another.

**Table 1.3.** The solution conformations of the branched tetrasaccharide  $\beta$ -D-Xyl-(1→2) ) $[\alpha$ -D-Man-(1→3)-] $[\alpha$ -D-Man-( 1→6)-] $\beta$ -D-Man-OMe and its component and trisaccharides.

Molecule	$\phi'/\psi'$ of a-c	$\phi'/\psi'$ of b-c	$\phi'/\psi'$ of d-c	$P_{\omega=60}$ : $P_{\omega=180}$
$\beta$ -D-Xyl-(1→2) $[\alpha$ -D-Man-(1→6)-] $\beta$ -D-Man-OMe a b c	-70/120	17/180		4:6
$\beta$ -D-Xyl-(1→2) $[\alpha$ -D-Man-(1→3)-] $\beta$ -D-Man-OMe a d c	-80/125		80/-140	
$\beta$ -D-Xyl-(1→2) ) $[\alpha$ -D-Man-(1→3)-] $[\alpha$ -D-Man-( a d b 1→6)-] $\beta$ -D -Man-OMe c	-80/125	80/170	80/-140	4:6

In the last example, on the basis of the n.O.e. data and theoretical calculations, it was found out that only one predominant conformer was found to be selected for each of the compounds shown in **Table 1.4**.<sup>32</sup> Here also, even though the oligosaccharides are branched, the change in the dihedral angles for the same type of linkage from compound to compound is very small.

**Table 1.4.** The solution conformations of a family of trisaccharides.

Molecule	$\phi/\psi$ of a-c	$\phi/\psi$ of b-c	$\phi/\psi$ of d-c	$\phi/\psi$ of e-c
$\alpha$ -D-Glc-(1→4)[ $\alpha$ -D-Man-(1→3)-]- $\beta$ -D-Gal-OMe a b c	-68.3/ -40.0	-65.5/ -51.9		
$\beta$ -D-Glc-(1→4)[ $\alpha$ -D-Man-(1→3)-]- $\beta$ -D-Gal-OMe d b c		-68.4/ -55.2	57.7/ 15.3	
$\alpha$ -D-Glc-(1→4)[ $\alpha$ -L-Rha-(1→3)-]- $\beta$ -D-Gal-OMe a e c	-71.8/ -39.7			51.7/ -10.0
$\beta$ -D-Glc-(1→4)[ $\alpha$ -L-Rha-(1→3)-]- $\beta$ -D-Gal-OMe d e c			50.7/ 17.5	51.7/ -23.4

From these and other examples we have learned that the solution conformation of disaccharides is reasonably consistent in the disaccharides and when those same disaccharide units are in the middle of a bigger oligo- or polysaccharide. This is further discussed in more detail in Chapter Two.

## **Molecular Mechanics Calculations and Molecular Dynamics**

### **Simulations of Carbohydrates**

Molecular mechanics grid search studies of the conformations about the glycosidic linkages and molecular dynamics simulations will be done to determine the possible conformational states with a weighting factor which tells the probability of being at a certain conformational state. Due to the smaller number of available constraints especially for long-range connectivities in oligosaccharides, it is necessary to rely upon potential energy surfaces as an aid to the interpretation of NMR data. In principle, the best estimates of the potential

surface for isolated molecules are obtained from ab *initio* calculations since a minimal number of approximations are inherent in the method. However, it is impractical to compute internal energies with full optimization for other than small organic molecules, due to the huge number of calculations required. In fact, the computational time increases approximately to the fifth power of the number of degrees of freedom of the system.<sup>23</sup> Semi-empirical calculations are also quantum mechanical in nature, but in order to reduce the computational complexity, a large number of integrals in the calculation are replaced by empirical data.<sup>33</sup> But, this is still dependent upon the cube of the number of degrees of freedom of the system, and is therefore still impractical for oligosaccharides.<sup>23</sup> Molecular mechanics (MM) methods use an entirely classical treatment. In molecular mechanics, a molecule is viewed as a collection of points (atoms) connected by springs (bonds) with different elasticities (force constants). The forces holding the atoms together are described by potential energy functions of structural features like bond lengths, bond angles, non-bonded interactions, and so on. In comparison with quantum mechanical methods, molecular mechanics methods are very much faster, and it is possible to obtain a minimized geometry for a large macromolecule in a useable time. The precise form of the force fields used in this calculations varies according to its source, but takes the general form of the equation below,<sup>23</sup> where  $r$  refers to bond lengths,  $\theta$  to valence angles,  $\phi$  to dihedral angles, with equilibrium values denoted by subscript  $e$ . The torsional term contains Fourier terms of the order  $n$  and phase  $\gamma$  to simulate various torsional functions. The fourth term groups the non-bonded

(van der Waals) interactions, together with a coulombic term containing the partial atomic charges  $q_i$  and  $q_j$ , and the fifth term represents the hydrogen-bonding potential. The accuracy of the calculations depends upon the choice of  $K_r$ ,  $K_\theta$ ,  $K_n$ ,  $A_{ij}$ ,  $B_{ij}$ ,  $q_i$ ,  $q_j$ ,  $C_{ij}$  and  $D_{ij}$ . These parameters are derived from a combination of *ab initio* calculations on small molecules and from experimental data.

$$\begin{aligned}
 E_{\text{tot}} = & \sum K_r(r-r_e)^2 + \sum K_\theta(\theta-\theta_e)^2 + \sum K_n(1+\cos(n\phi-\gamma)) + \sum (A_{ij}/R_{ij}^{12}-B_{ij}/R_{ij}^6) + \\
 & \text{bond} \qquad \qquad \text{angle} \qquad \qquad \text{torsion angle} \qquad \qquad \text{van der Waals} \\
 & \sum (q_i q_j / \epsilon R_{ij}) + \sum (C_{ij}/R_{ij}^{12}-D_{ij}/R_{ij}^{10}) + \sum K_n(n-n_e)^2 \qquad \qquad (3) \\
 & \text{electrostatic} \qquad \text{hydrogen bond} \qquad \qquad \text{n.O.e.}
 \end{aligned}$$

Now in Chapter Two, we have collected and statistically analyzed crystal structures of carbohydrates and came out with a simple rule-based method for predicting the conformations of oligo- and polysaccharides. In Chapter Three, we describe the design and synthesis of seven molecules with rigid glycosidic linkages. These molecules are important to parameterize a new Karplus-type equation that will be useful to calculate the glycosidic dihedral angles of carbohydrates from vicinal proton-carbon coupling constants and hence help predict carbohydrate conformations.

In Chapter four, we optimized the geometries, of the seven molecules whose synthesis is described in Chapter Three and three other commercially available molecules, using molecular mechanics and AM1 and PM3 semi-

empirical calculations. We then measured the vicinal proton-carbon heteronuclear coupling constants using a two-dimensional NMR spectroscopy, excitation-sculptured indirect-detection experiment (EXSIDE). From the least squares fitting of the coupling constants and the dihedral angles, we came out with a new Karplus-type equation that will be used to predict the glycosidic dihedral angles from NMR coupling constants or vice versa.

Finally, in Chapter Five we describe our study of the solution conformation and dynamics of the disaccharide maltose, the tetrasaccharide reduced maltotetraose, and the pseudotetrasaccharide acarbose using existing methods and the new methods that we have developed and described in chapters two to four.

## References

1. Hudson, C. S. *J. Am. Chem. Soc.* **1909**, *31*, 66-80.
2. Lemieux, R.; Koto, S. *Tetrahedron*, **1974**, *30*, 1933.
3. Lemieux, R.; Koto, S.; Voisin, D. *The origin and Consequences of the Anomeric Effect*, **1979**, p2.
4. Rademacher, T. W.; Parekh, R. B. & Dweck, R. A. *Ann. Rev. Biochem.* **1988**, *57*, 785.
5. Phillips, M. L.; Nudelman, E.; Gaeta, F. C. A.; Perez, M.; Shingel, A. K.; Hakamori, S.; Paulson, J. C. *Science* **1990**, *250*, 1130.
6. Gilardi, R.; Flippen-Andersen, J. L. *Acta Crystallogr.* **1986** *C43*, 806.
7. Jeffrey, G. A.; Huang, D. B. *Carbohydr. Res.* **1993**, *247*, 37.
8. Okuyama, K.; Noguci, K.; Saitoh, M.; Ohno, S.; Fujii, S.; Tsukada, M.; Takeda, H.; Hidano, T. *Bull. Chem. Soc. Jap.* **1993**, *66*, 374.
9. Gessler, K.; Krauss, N.; Steiner, T.; Betzel, C.; Sandmann, C.; Saenger, W. *Science* **1994**, *266*, 1027.
10. Raymond, S.; Heyraud, A.; Tran Qui, D.; Kuick, Å.; Chanzy, H. *Macromolecules* **1995**, *28*, 2096.
11. Chandrasekaran, R. *Adv. Carbohydr. Chem. Biochem.* **1997**, *52*, 311.
12. Ernst, R., Bodenhausen, G., & Wokahun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimension*, **1987** pp 431-440, Oxford University Press, Oxford.
13. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355.
14. Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565.



15. Titman, J.; Neuhaus, D.; Keeler, J. *J. Magn. Reson.* **1989**, *85*, 111.
16. Sangers, J. K.; Hunter, B. K. *Modern NMR Spectroscopy*, **1993**, pp 193-197, Oxford University Press, Oxford.
17. Tvaroska, I; Hricovini, M; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359.
18. Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988**, *184*, 39.
19. Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1999**, *121*, 9843.
20. Haasnoot, C. A.; De Leeuw, F. A.; Altona, C. *Tetrahedron* **1980**, *36*, 2783.
21. Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897.
22. Noggle, J. H.; Schirmer, R. E. *The Nuclear Overhauser Effect-Chemical Applications*, Academic Press, London, New York, **1971**.
23. Homans, S. W. *Progress in NMR Spectroscopy* **1990**, *22*, 55; Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, T.C.; Ferguson, D.M.; Seibel, G.L.; Singh, U.S.; Weiner, P.K.; Kollman, P.A. AMBER 4.1, University of California, San Francisco, **1995**.
24. Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179.
25. Karplus, M.; McCammon *Ann. Rev. Biochemistry* **1983**, *52*, 263.
26. Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664.
27. Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6676.
28. Toma, L.; Ciuffreda, P.; Colombo, D.; Ronchetti, F.; Lay, L.; Panza, L. *Helv. Chim. Acta* **1994**, *77*, 668.

29. Toma, L.; Colombo, D.; Roncheti, F.; Panza, L.; Russo, G. *Helv. Chim. Acta* **1995**, *78*, 636.
30. Brisson, J. R.; Carver, J. P. *Biochemistry* **1983**, *22*, 3671.
31. Leeflang, B. R.; Bouwstra, J. B.; Kerékgyarto, J.; Kamerling, J. P.; Vliegthant, J. F. B. *Carbohydr. Res.* **1990**, *208*, 117.
32. Lipkind, G. M.; Shashkov, A. S.; Nechaev, O. A.; Torger, V. I.; Shibaev, N. N.; Kochetkov, N. K. *Carbohydr. Res.* **1989** *195*, 27.
33. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

## **Chapter 2**

### **A Simple and Accurate Rule-based Method for Predicting the Conformation of Oligo- and Polysaccharides**

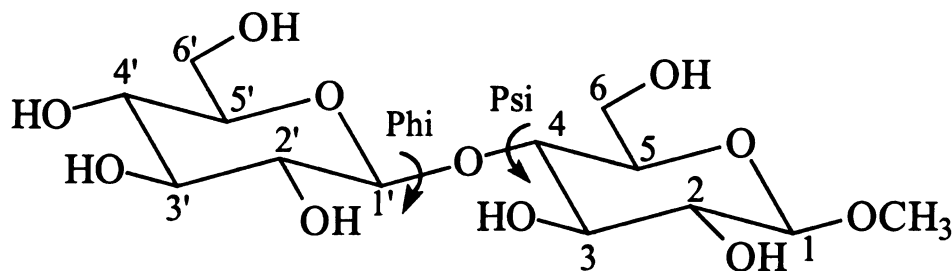
(Adapted from the paper entitled 'A Simple and Accurate Rule-based Method for Predicting the Conformation of Oligo- and Polysaccharides', Hussen Mohammed and Rawle I. Hollingsworth, *J. Org. Chem.* **2001**, accepted for publication.)

## **Abstract**

In this chapter, we report a simple rule based method for predicting the conformations of oligo- and polysaccharides. The method is based on an exhaustive analysis of the entries in the Cambridge Structural Database (CSD) of crystal structures. We analyzed the values of the critical interglycosidic dihedral angles, O5'-C1'-OX-CX ( $\phi$ ) and C1'-OX-CX-CX+1 ( $\psi$ ) from all of the entries which ranged from di-, tri-, tetra- and hexa- to octasaccharides. These results were used to construct a second database of dihedral angles. The maximum percentage of the total possible  $\phi$ - $\psi$  space that is populated by a given linkage was found to be 4.0 % and that was for  $\alpha$ -1->2 linkages. An almost exclusive preference for specific values of  $\phi$  and  $\psi$  for a specific type of linkage independently of the context was clearly evident. These values coincided with those indicated in a molecular mechanics potential energy search. There was symmetry of placement of the points on the  $\phi$ - $\psi$  plots that was determined by the absolute configuration of the anomeric carbon and the carbon to which the linkage is made. These corresponded to the RR, SS, SR, and RS configurations. Three simple and accurate rules are proposed to help predict the conformations of oligo- and polysaccharides where the glycosyl donor is a pyranose residue. The results support, expand and generalize earlier reports from other groups.

## **Introduction**

Carbohydrates are among the most abundant naturally occurring organic substances on earth. They are found in large storage polymers such as starch, in cell wall components such as cellulose and hemi-cellulose, in insect exoskeletons as chitin, as bacterial cell surface antigens, on viruses as the immunodominant species that govern infection, as the predominant substances on blood cell surfaces that determine blood type and in a large variety of other roles and functions. Nucleic acids contain a large proportion of carbohydrates in the forms of ribose and deoxy-ribose. Despite their importance, methods for determining the 3-dimensional structure and properties of even mildly complex carbohydrates have lagged behind those of other biomolecules. There are several reasons for this. Firstly, complex carbohydrates are generally difficult to crystallize and there are therefore only a handful of X-ray crystal structures and these are limited to fiber analyses. The nuclear magnetic resonance (NMR) spectra of complex carbohydrates are very complicated and characterized by a high degree of overlap and second-order coupling effects. Because of this, it is usually not possible to get anywhere near a complete assignment. This often precludes solution 3-dimensional structure determination by this method.<sup>1</sup> The third reason is that carbohydrates, for example those found in glycoproteins, are often very heterogeneous in nature. Their biological function is based on this heterogeneity. Current structural and analytical tools, especially those based on diffraction methods, are extremely limited in utility because they require homogeneity and uniformity of structure.



**Figure 2.1.** Structure of 1-O-methyl cellobioside and the glycosidic dihedral angles, phi and psi, which determine its conformation. Phi = O5'-C1'-O4-C4 and Psi = C1'-O4-C4-C5.

There is, however, a fairly significant body of information on the 3-dimensional structure of very small carbohydrate molecules such as monosaccharides and di- and trisaccharides. It might be possible to make projections from these structures to more complex ones. There are three important questions to be asked. Firstly, how dependent is the conformation about the glycosidic linkage on the nature of the residues attached at that linkage? Secondly, how do substitutions in the glycosyl residue affect the conformation of the glycosidic linkage of that same residue? Thirdly, how sensitive is the conformation of a particular type of linkage (e.g.  $\alpha$ -1- $\rightarrow$ 4) between two residues to the context in which that linkage appears? Is this conformation the same in different molecules? If favorable and reliable answers to these

questions could be obtained, a way of predicting the conformation of complex carbohydrate molecules could be developed.

There is more than just an indication that the conformation of the bonds involved in a specific type of inter-glycosidic linkage between two carbohydrate rings might be independent of the context in which the linkage appears. This is implied in the Hudson rules of iso-rotation.<sup>2</sup> These rules allow the estimation of the net optical rotation of carbohydrate molecules containing more than one residue by summing fractional contributions from each residue. The anomeric carbon makes the largest contribution because the ring oxygen, the anomeric carbon and the anomeric oxygen atom form the system with the most intense U.V. chromophore in typical carbohydrates without  $\pi$ -systems. Because the intensity of this chromophore (and therefore the optical rotatory dispersion) is a function of the geometry of the interglycosidic bond, it is reasonable to conclude that significant changes in geometry of a given type of linkage do not occur in different molecular contexts. In an earlier study with a limited number of glycosides, it was observed that the dihedral angles of a particular type of linkages generally fall between the limits of  $\pm 30^\circ$ , irrespective of the monosaccharides involved.<sup>3</sup> The idea that the geometry about interglycosidic linkages is relatively independent of the molecular context in which they occur is exploited in the molecular builder computer program POLYS by Engelsen et al. This program produces three-dimensional structures of carbohydrates from preoptimized monosaccharide structures and from computer analyses describing the linkage properties of disaccharide fragments.<sup>4</sup>

## Results and Discussion

A case for supporting the conjecture that linkages should be relatively insensitive to context can also be made by analyzing the potential energy surface obtained by systematically varying the dihedral ( $\phi$ ) about the anomeric linkage relative to the one about the linkage that connects the glycosyl residue to the aglycon or second carbohydrate residue ( $\psi$ ) (**Figure 2.1**) and evaluating the energy at each step using molecular mechanics methods. **Figure 2.2** shows  $\phi, \psi$  maps of (a) cellobiose ( $\beta$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc) and (b) mellobiose ( $\alpha$ -D-Gal-(1 $\rightarrow$ 6)-D-Glc). Dih-1 stands for  $\phi$ , which is defined as H1'-C1'-OX-CX and Dih-2 stands for  $\psi$ , which is defined as C1'-OX-CX-HX. For cellobiose X is 4 and for mellobiose it is 6. The contour lines are drawn with a gap of 1 kcal/mol for each line. Such potential energy surfaces in hexopyranosides are characterized by a single deep potential energy well (**Figure 2.2a**) except for 1-6 linkages where two relatively broad but connected minima are obtained (**Figure 2.2b**). The depth of the potential energy well is typically 6 kcal/mol or higher. This is much higher than  $kT$  (the characteristic energy) at room temperature. Because of this, residues should be essentially locked in this conformation unless the penalty can be offset by some strong interaction such as an electrostatic one if charged groups are present.

A definitive proof of the conjecture was sought by examining the entire Cambridge database of crystal structures and analyzing the phi and psi values for all of the carbohydrate entries. These include simple glycosides,<sup>5-8</sup> disaccharides,<sup>9-82</sup> trisaccharides,<sup>83-97</sup> tetrasaccharides,<sup>98-100</sup> hexasaccharides,<sup>101</sup>



and octasaccharides<sup>102</sup>. The database contained a total of 164 phi and 153 psi angles. No database entries were ignored. The results of the search were completely supportive of our conjecture. An almost exclusive preference for a specific value of phi and psi for a specific type of linkage independently of context was clearly evident. These values coincided with those indicated in the molecular mechanics potential energy surface search. **Table 2.1** shows the total area of the  $\phi, \psi$  space populated by the different linkages of oligosaccharides where both the donor and acceptor residues have the D configuration. Only a very small area of this space is populated. Of the available 360 by 360 degrees of the  $\phi, \psi$  space, the maximum area populated by a given linkage of oligosaccharides is 4.0% and that is for the  $\alpha$ -1->2 linkage (**Table 2.1**). The percentages in **Table 2.1** were obtained by using the relation  $(\phi \text{ max} - \phi \text{ min})(\psi \text{ max} - \psi \text{ min}) / (360 \times 360)$ . The actual area populated by the oligosaccharides within the rectangle is much smaller. This indicates that the actual percentage of  $\phi, \psi$ -space that is populated is very limited. The value corresponding to the maximum probability that the angle of a specific linkage lay within a given percentage of the total configurational space of 360 degrees and the mean value of the dihedral angle is given for some of the oligosaccharides in **Table 2.2**. Also included in **Table 2.2** are the standard deviation and the percentage of  $\phi$  or  $\psi$  space populated by a given linkage. From the linkages where both the donor and acceptor residues are pyranose rings, the highest percentage of  $\phi$  space populated is 14 % and that is for the  $\alpha$ -Glc-1->4-Glc linkage. The corresponding value for the  $\psi$  space is obtained for  $\beta$ -Glc-1->3-Glc linkage and it is 16 %.

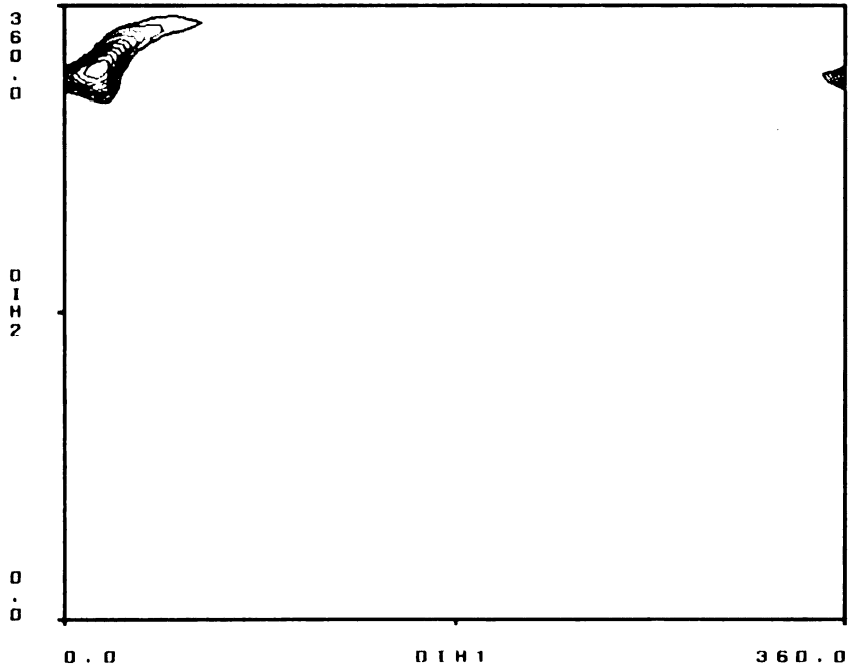
**Table 2.1.** The percentages of the total phi-psi space that is populated by oligosaccharide crystals from Cambridge Structural Database.

Linkage	% Area	Linkage	% Area
$\beta$ -1 $\rightarrow$ 4	1.66	$\alpha$ -1 $\rightarrow$ 4	3.35
$\beta$ -1 $\rightarrow$ 2	1.87	$\alpha$ -1 $\rightarrow$ 2	3.98
$\beta$ -1 $\rightarrow$ 6	2.55	$\alpha$ -1 $\rightarrow$ 6	2.68
$\beta$ -1 $\rightarrow$ 3	2.99	$\alpha$ -1 $\rightarrow$ 3	1.75

Because of geometric arguments, there is a symmetry of placement of the points on the potential energy surface that is determined by the absolute configuration of the anomeric carbon and the carbon to which the linkage is made if that carbon is a carbohydrate ring carbon (**Figure 2.3**). These correspond to the configurations RR, SS, SR and RS. The first character in each pair corresponds to the configuration of the anomeric carbon. In the case of alpha D-glucosides and the entire alpha D-aldohexopyranosides then this first character is R. If the residue is attached to a primary carbon in the aglycon (e.g. the 6-position of a hexose) the second character is not relevant. For the corresponding beta glycosides the configuration at C1 is S. The configurations for the acceptor carbons of a D-glucose are R, S, and R for the acceptor carbons 2, 3, and 4, respectively. The reverse holds for the L-sugars. As noted earlier, there are two minima on the potential energy surface if the linkage is to the primary carbon at the 6-position of an aldohexopyranoside.

**Figure 2.2** Biograf rigid residue  $\phi, \psi$  maps of (a) cellobiose ( $\beta$ -D-Glc-(1->4)BD-Glc) and (b) melliobiose ( $\alpha$ -D-Gal-(1->6)- $\beta$ -D-Glc). Dih-1 stands for  $\phi$ , which is H1'-C1'-OX-CX where X is 4 for cellobiose and 6 for melliobiose and Dih-2 stands for  $\psi$ , which is C1'-O4-C4-H4 for cellobiose and C1'-O6-C6-C5 for melliobiose. The contour lines are drawn with a gap of 1 kcal/mol for each line. This figure is a color presentation for clarity.

a.



b.



Structures corresponding to the occupation of one or the other of these minima were observed. This study is limited to the anomeric linkage of pyranosides and the results of the statistical analysis of the built database of carbohydrate dihedral angles from CSD can be summarized in the following rules:

(1) For a disaccharide linkage between D-pyranose sugars in which the bond from the bridging oxygen to both residues is equatorial (e.g. a residue that is  $\beta$ -linked to an equatorial site on the other residue) the average value  $\pm$  standard deviation of  $\phi$  and  $\psi$  (and  $\omega$  for 1-6 linkages) are given below in degrees as  $\phi, \psi$  pairs:

(a)  $-88 \pm 11$ ,  $-128 \pm 15$  for 1-4 linkages,

(b)  $-79 \pm 10$ ,  $109 \pm 34$  for 1-3 linkages,

(c)  $-75 \pm 11$ ,  $-129 \pm 23$  for 1-2 linkages, and

(d) For 1-6 linkages  $\phi$  is  $-71 \pm 10$  with  $(\psi, \omega)$  of  $(-170 \pm 3, 83 \pm 7)$  when the acceptor is galactose and  $(164 \pm 4, 71 \pm 7)$  with a 30% probability and  $(-166 \pm 10, -63 \pm 1)$  with a 70% probability when the acceptor is glucose.

(2) When the oxygen is axial on the donor carbon the absolute value of  $\phi$  is similar but the sign is reversed. The absolute value of  $\psi$  is similar but its sign is reversed if the bridging oxygen is axial to the acceptor carbon (e.g. the 2-position of mannose or the 4-position of galactose). The sign of  $\phi$  is reversed if the donor residue is of the L configuration and the sign of  $\psi$  is reversed when the acceptor

residue is of the L configuration. Thus for a D residue that is  $\alpha$ -linked to an equatorial site of another D residue, the average value  $\pm$  standard deviation of  $\phi$  and  $\psi$  (and  $\omega$  for 1-6 linkages) are given below in degrees as  $\phi, \psi$  pairs.

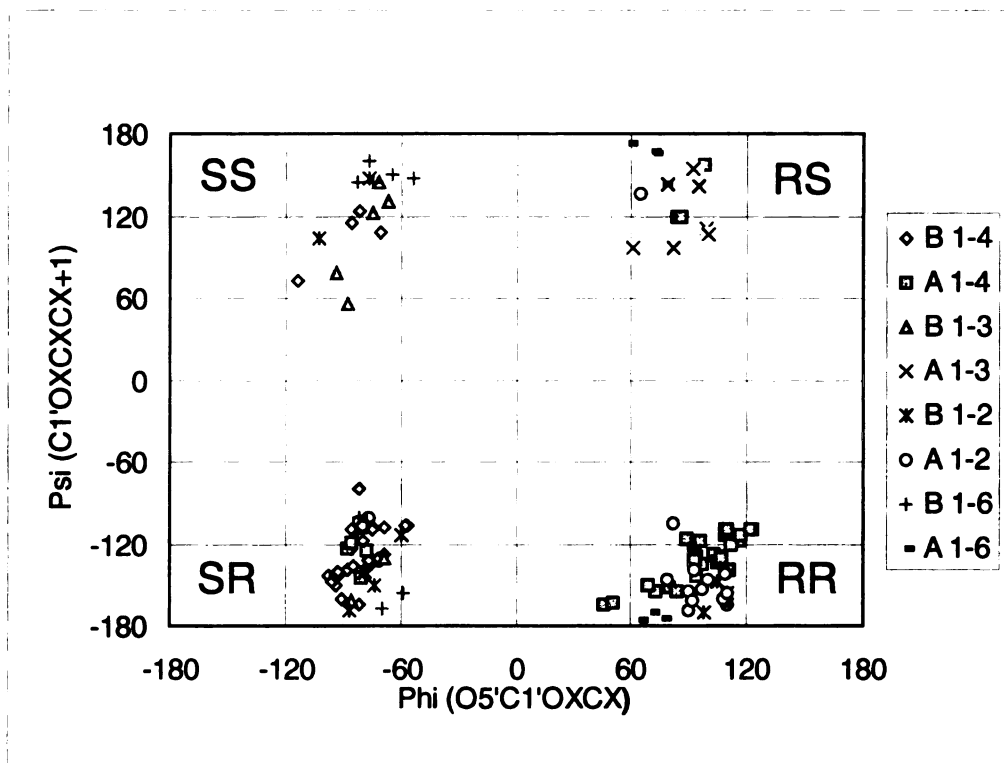
- (a)  $98 \pm 19, -128 \pm 22$  for 1-4 linkages
- (b)  $87 \pm 14, 121 \pm 24$  for 1-3 linkages,
- (c)  $99 \pm 11, -146 \pm 28$  for 1-2 linkages, and
- (d) For 1-6 linkages  $\phi$  is  $72 \pm 8$  with the same ( $\psi, \omega$ ) values as in Rule 1(d).

(3) Rules 1 and 2 also apply for acyl glycosides or for glycosides bearing simple substituents at other sites in the ring.

It is noteworthy that there is only a very small difference in the average values of  $\phi$  and  $\psi$  between the 1- $\rightarrow$ 4 and 1- $\rightarrow$ 2 linkages as shown in Rules 1 (a) and (c) and 2 (a) and (c). The reason for this is that both 1- $\rightarrow$ 4 and 1- $\rightarrow$ 2 linkages belong to the same region of  $\phi, \psi$  space, which is the SR region for the  $\beta$ -1- $\rightarrow$ 4 and  $\beta$ -1- $\rightarrow$ 2 linkages and the RR region for the  $\alpha$ -1- $\rightarrow$ 4 and  $\alpha$ -1- $\rightarrow$ 2 linkages (**Figure 2.3**). In the case of  $\beta$ -1- $\rightarrow$ 3 linkages (Rule 1(b)), the average value of  $\phi$  (magnitude and sign) is practically identical to the corresponding values of the  $\beta$ -1- $\rightarrow$ 4 and  $\beta$ -1- $\rightarrow$ 2 linkages (Rule 1(a) and (c)) but the average value of  $\psi$  is of similar magnitude and opposite in sign to the corresponding values of the  $\beta$ -1- $\rightarrow$ 4 and  $\beta$ -1- $\rightarrow$ 2 linkages. Here again the reason is that while  $\beta$ -1- $\rightarrow$ 3 linkages belong to the SS region of the  $\phi, \psi$  space, the  $\beta$ -1- $\rightarrow$ 4 and  $\beta$ -1- $\rightarrow$ 2 linkages belong to the SR region (**Figure 2.3**). Similarly, the  $\alpha$ -1- $\rightarrow$ 3 linkage belongs to the RS region

**Table 2.2.** Results of the statistical analysis of the dihedral angles ( $\phi$  and  $\psi$ ) of some of the oligosaccharides from the Cambridge Structural Database.

Linkages	$\alpha$ -Glc-(1 $\rightarrow$ 4)- Glc	$\beta$ -Glc-(1 $\rightarrow$ 4)- Glc	$\beta$ -Glc-(1 $\rightarrow$ 3)- Glc	$\beta$ -Gal-(1 $\rightarrow$ 4)- Glc	$\beta$ -Man-(1 $\rightarrow$ 4)- Man	$\alpha$ -Glc-(1 $\rightarrow$ 2)- Fru	
$\Phi$	No. Of records	28	13	6	13	4	16
	Mean	96	-82	-80	-78	-84	101
	Standard Deviation	20	9	10	13	12	10
	% $\Phi$ space	14	7	8	10	7	8
$\Psi$	No. Of records	28	13	6	13	4	16
	Mean	-129	-127	111	-127	-140	-143
	Standard Deviation	22	16	36	15	11	31
	% $\Psi$ space	16	15	16	10	6	18



**Figure 2.3**  $\Phi, \Psi$  map of 1-2, 1-3, 1-4 and 1-6 linked oligosaccharides from the Cambridge Structural Database (CSD). Each point on the map corresponds to one glycosidic linkage from the CSD.  $\Phi = O5'-C1'-OX-CX$  and  $\Psi = C1'-OX-CX-CX+1$ , where  $X = 2, 3, 4,$  and  $6$  for 1-2, 1-3, 1-4, and 1-6 linkages, respectively.



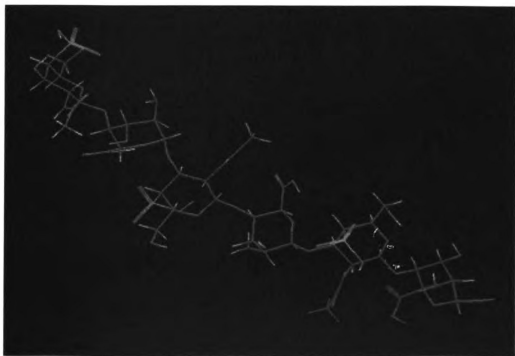
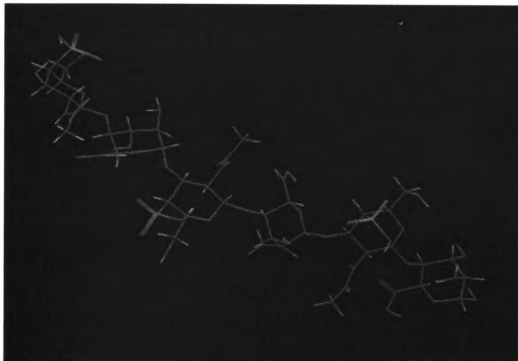
while the  $\alpha$ -1- $\rightarrow$ 4 and  $\alpha$ -1- $\rightarrow$ 2 linkages reside in the RR region. Comparing rules 1(a-d) with the corresponding rules 2(a-d), we can see that while the average values of  $\psi$  have the same sign and similar magnitudes, the average values of  $\phi$  have similar magnitude but opposite sign. This is because the anomeric carbon of the donor residues in the two rules have opposite configurations, S in rules 1(a-d) and R in rules 2(a-d), while the configuration of the acceptor carbons stay the same in the two rules.

It should be mentioned that care should be taken when using the average values of  $\phi$  and  $\psi$  in Rules 1(d) and 2(d) for 1- $\rightarrow$ 6 linkages because of the relatively greater flexibility of 1- $\rightarrow$ 6 linkages as compared to 1- $\rightarrow$ 4, 1- $\rightarrow$ 3 and 1- $\rightarrow$ 2 linkages. The higher flexibility of 1- $\rightarrow$ 6 linkages arise from the lower barrier of a methylene group to rotation. We have observed that when a glucose residue is attached at the anomeric carbon (C1) of the acceptor glucose of 1- $\rightarrow$ 6 linkages, the signs of both  $\psi$  and  $\omega$  are positive but when C1 has a free hydroxyl or a fructose residue attached to it, the signs of both  $\psi$  and  $\omega$  are negative (Rule 1(d) and 2(d)). It seems that the sign of  $\omega$  dictates the sign of  $\psi$  for 1- $\rightarrow$ 6 linkages where the acceptor is glucose. A larger database of 1- $\rightarrow$ 6 linkages is needed to confirm this. The  $(\psi, \omega)$  average values in Rules 1(d) and 2(d) are the same because of the achiral nature of C6. Even though there are no 1- $\rightarrow$ 6 linkages with the acceptor residue having the L configuration in the Cambridge Structural Database, we expect that the sign of  $\omega$  and hence the sign of  $\psi$  might be opposite to the case when the acceptor residue has the D configuration. In this

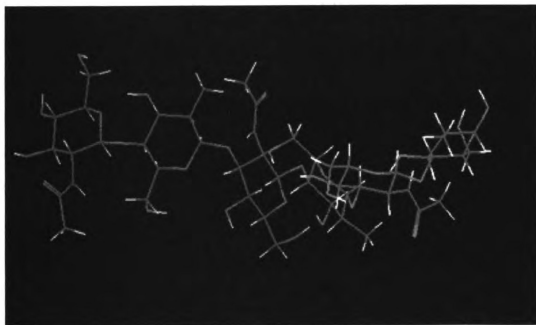
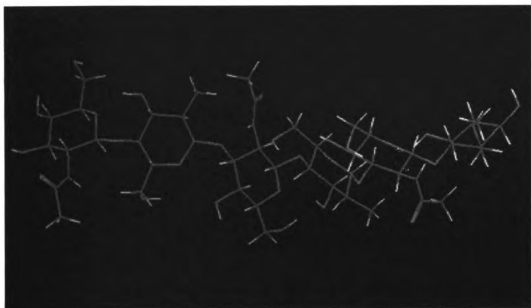
work, we observed that the average values of  $\phi$  shift from  $71^\circ$  for 1- $\rightarrow$ 6 linked oligosaccharides to  $80^\circ$  for  $\beta$ -1- $\rightarrow$ X (X = 2, 3, 4) to  $94^\circ$  degrees for  $\alpha$ -1- $\rightarrow$ X (X = 2, 3, 4) linked oligosaccharides as the steric demand increases.

To test the prediction power of the rules proposed in this paper we plotted the structure of two polysaccharides, dermatan 4-sulfate (**Figure 2.5**) and hyaluronan (sodium hyaluronate III) (**Figure 2.6**), using the dihedral angles from X-ray fiber diffraction studies<sup>104</sup> with the BIOGRAF molecular mechanics program<sup>105</sup> and compared them with those predicted using the rules. Dermatan 4-sulfate is a polysaccharide with the repeating unit [- $\rightarrow$ 3)- $\beta$ -D-GalNAc4SO<sub>3</sub><sup>-</sup>-(1- $\rightarrow$ 4)- $\beta$ -L-IdoA-(1- $\rightarrow$ )] and interacts with collagen, elastin, and some glycoproteins to maintain structural integrity of the tissues involved. Hyaluronan is an unsulfated glucosaminoglycan found in mammalian connective tissues, where it forms the central core of the proteoglycan aggregate.<sup>104</sup> It is a polymer composed of alternating 1- $\rightarrow$ 4 and 1- $\rightarrow$ 3 linkages [- $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1- $\rightarrow$ 4)- $\beta$ -D-GlcA-(1- $\rightarrow$ )]. As shown in **Figure 2.5** and **Figure 2.6**, the predicted structures are comparable to the structures from X-ray fiber diffraction studies. These rules will be useful for predicting the three dimensional structures of pure carbohydrates and the carbohydrate parts of glycoproteins and glycolipids. These structures will be useful in understanding molecular recognition in biological processes for example.

**Figure 2.5** Line representation of the structure of dermatan 4-sulfate predicted using the rules in this paper (bottom) and from X-ray fiber diffraction study (top).<sup>104</sup> Dermatan 4-sulfate is a polysaccharide with the repeating unit [->3)- $\beta$ -D-GalNAc4SO<sub>3</sub><sup>-</sup>-(1->4)- $\beta$ -L-IdoA-(1->] and interacts with collagen, elastin, and some glycoproteins to maintain structural integrity of the tissues involved. This figure is a color presentation for clarity.



**Figure 2.6** Line representation of the structure of hyaluronan (sodium hyaluronate III) predicted using the rules in this paper (bottom) and from X-ray fiber diffraction study (top).<sup>104</sup> Hyaluronan is an unsulfated glucosaminoglycan found in mammalian connective tissues, where it forms the central core of the proteoglycan aggregate. It is a polymer composed of alternating 1->4 and 1->3 linkages [->3)- $\beta$ -D-GlcNAc-(1->4)- $\beta$ -D-GlcA-(1->]. This figure is a color presentation for clarity.



## Conclusions

We describe herein a reliable method for predicting oligo- and polysaccharide conformations. The consistency of prediction especially when mixed linkages are present is high. This indicates that further substitution adjacent to the site of the linkage does not have a dominant impact at this level of refinement. The exo-anomeric effect<sup>1</sup> is a phenomenon that specially relates to the conformation about glycosidic linkages and deserves some special comment. It relates to the preference for a gauche conformation about the O-R bond of the aglycones of sugars. The effect should lead to values of +60 and -65 degrees for the  $\phi$  angles of  $\alpha$ - and  $\beta$ -glycosides, respectively.<sup>1a</sup> Values of  $\phi$  that range from 61 – 74° for  $\alpha$ - and from 68 - 87 for  $\beta$ -methyl glycosides with corresponding ranges of 75 – 121° and 71 – 105°, respectively for  $\alpha$ - and  $\beta$ -linked oligosaccharides were reported in an earlier work.<sup>103</sup> These and the values reported here are within a range consistent with an exo-anomeric effect. It is still not clear, however, to what extent this effect plays a role especially since a term defining it was not implicitly assigned in the force fields but results that agree with the databases were obtained anyway. There is yet another level of refinement to the predictions that one can make. This would include an assessment of steric, electronic and volume effects of substituents proximal to the linkage sites. The biasing due to the anomeric effect will doubtless play a key role in this next level of prediction.

## Acknowledgements

This work was supported by the Michigan State University Research Excellence Fund.

## References

1. For some of the early works on the prediction of the conformations of oligosaccharides and modeling of their solution conformations by a combination of NMR and calculation, see (a) Lemieux, R. U.; Koto, S. *Tetrahedron* **1974**, *30*, 1933. (b) Lemieux, R. U.; Koto, S.; Voisin, D. *Origin and Consequences of the anomeric effect, ACS Symposium series, No. 87, Chapter 2* **1979**, p17. (c) Lemieux, R. U.; Bock, K. *Jap. J. Antibiot.* **1979**, *32*, 163. (d) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. *Can. J. chem.* **1980**, *58*, 631. (e) Thogersen, H.; Lemieux, R. U.; Bock, K.; Meyer, B. *Can. J. chem.* **1982**, *60*, 44. (f) Bock, K.; Lemieux, R. U. *Carbohydr. Res.* **1982**, *100*, 63. (g) Lemieux, R. U.; Bock, K. *Arch. Biochem. Biophys.* **1983**, *221*, 125. (h) Sabesan, S.; Bock, K.; Lemieux, R. U. *Can. J. Chem.* **1984**, *62*, 1034. (i) Vandonselaar, M.; Delbaere, L. T. J.; Spohr, U.; Lemieux, R. U. *J. Biol. Chem.* **1987**, *262*, 10848. (j) Delbaere, L. T. J.; Vandonselaar, M.; Prasad, L.; Quail, J. W.; Nikrad, P. V.; Pearlstone, J. R.; Carpenter, L. B.; smillie, L. B.; Spohr, U.; Lemieux, R. U. *Trans. Amer. Crystallogr. Assoc.* **1989**, *25*, 65. (k) Otter, A.; Lemieux, R. U.; Ball, R. G.; Venot, A. P.; Hidsgaul, O.; Bundle, D. R. *Eur. J. Biochem.* **1999**, *259*, 295. (l) Bock, K.; Arnap, J.; Loenngren, J. *Eur. J. Biochem.* **1982**, *129*, 171. (m) Bock, K. *Pure & Appl. Chem.* **1983**, *55*, 605. (n) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664. (o) Meyer, B. *Topics in current chemistry: Carbohydrate chemistry*, ed. Thiem, J., Berlin, **1990**, *154*, 141.



2. Hudson, C. S. *J. Amer. Chem. Soc.* **1930**, *52*, 1680; Hudson, C. S. *J. Amer. Chem. Soc.* **1930**, *52*, 1707.
3. Jeffrey, G. A. *Acta Cryst. Section B* **1990**, *46*, 89.
4. Engelsen, S. B.; Cross, S.; Mackie, W.; Perez, S. *Biopolymers*, **1996**, *39*, 417.
5. Jeffrey, G. A.; Takagi, S. *Acta Crystallogr., Section B* **1977**, *33*, 738.
6. Gatehouse, B. M.; Poppleton, B. J. *Acta Crystallogr., Section B* **1970**, *26*, 1761.
7. Jeffrey, G. A.; McMullen, R. K.; Takagi, S. *Acta Crystallogr., Section B* **1977**, *33*, 728.
8. Shalaby, M. A.; Fronczek, F. R.; Younathan, E. S. *Carbohydr. Res.* **1994**, *258*, 267.
9. Taga, T.; Sumiya, S.; Osaki, K.; Utamura, T.; Koizumi, K. *Acta Crystallogr., Section B* **1981**, *37*, 963.
10. Mo, H.; Jensen, L. H. *Acta Crystallogr., Section B* **1978**, *34*, 1562.
11. Longchambon, F.; Ohanessian, J.; Gillier-Pandraud, H.; Duchet, D.; Jacquinet, J. C.; Sinay, P. *Acta Crystallogr., Section B* **1981**, *37*, 601.
12. Foces-Foces, C.; Cano, F. H.; Garcia-Blanco, S. *Acta Crystallogr., Section B* **1980**, *36*, 377.
13. Ollis, J.; James, V. J.; Angual, S. J.; Pojer, P. M. *Carbohydr. Res.* **1978**, *60*, 219.
14. Smith, J. L.; Sundaralingam, M. *Acta Crystallogr., Section B* **1981**, *37*, 1095.

15. Rihs, G.; Traxler, P. *Helv. Chim. Acta* **1981**, *64*, 1533.
16. Neuman, A.; Becquart, J.; Avenel, D.; Gillier-Pandraud, H.; Sinay, P. *Carbohydr. Res.* **1985**, *139*, 23.
17. Hirotsu, K.; Simada, A. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1872.
18. Lugar, P.; Vangehr, K.; Bock, K.; Pauleen, H. *Carbohydr. Res.* **1983**, *117*, 23.
19. Nassimbeni, L. R.; Nirer, M. L.; Cragg, G. M.; Pettit, G. R. *Acta Crystallogr. Section C (Cr. Str. Comm.)* **1985**, *41*, 728.
20. Jacobson, R. A.; Wunderlich, J. A.; Lipscomb, W. N. *Acta Crystallogr.* **1961**, *14*, 598.
21. Brown, C. J. *J. Chem. Soc* **1966**, A 927.
22. Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr., Section B* **1968**, *24*, 830.
23. Senma, M.; Taga, T.; Osak, K. *Chem. Lett.* **1974**, 1415.
24. Perez, S.; Vergelati, C. *Acta Crystallogr., Section B* **1984**, *40*, 294.
25. Svensson, G.; Albertsson, J.; Svensson, C.; Mangusson, G.; Dahmen, J. *Carbohydr. Res.* **1986**, *146*, 29.
26. Perez, S.; Vergelati, C.; Tran, V. H. *Acta Crystallogr., Section B* **1985**, *41*, 262.
27. Jeffrey, G. A.; Nanni, R. *Carbohydr. Res.* **1985**, *137*, 21.
28. Wong, R. Y.; Horowitz, R. M. *J. Chem. Soc., Perkin Trans. 1*, **1986**, 843.
29. Brown, J. M.; Cook, S. J.; Jones, R. H.; Khan, R. *Tetrahedron* **1986**, *42*, 5089.

30. Srikrishnan, T.; Chowdhary, M. S.; Matta, K.L. *Carbohydr. Res.* **1989**, *186*, 167.
31. Nikitin, A. V.; Andranov, V. I.; Myasnikova, R. M.; Firgang, S. I.; Usov, A. I.; Sopin, V. F.; Pertsin, A. I. *Kristallografiya* **1986**, *31*, 676.
32. Lamba, D.; Burden, C.; Mackie, W.; Sheldrick, B. *Carbohydr. Res.* **1986**, *153*, 205.
33. Lamba, D.; Burden, C.; Mackie, W.; Sheldrick, B. *Carbohydr. Res.* **1986**, *155*, 11.
34. Ezure, Y.; Yoshikuni, Y.; Ojima, N.; Sugiyama, M.; Hirotsu, K.; Higuchi, T. *Acta Crystallogr. Section C (Cr. Str. Comm.)* **1987**, *43*, 1809.
35. Nikitin, A. V.; Shibanov, T. A.; Firgang, S. I.; Usov, A. I.; Sopin, V. F.; Myasnikova, R. M. *Kristallografiya* **1987**, *32*, 896.
36. Takagi, S.; Jeffrey, G. A. *Acta Crystallogr., Section B* **1977**, *33*, 2377.
37. Arene, F.; Neuman, A.; Longchambon, F. *C. R. Acad. Sci. C Chim.* **1979**, *288*, 331.
38. Rohrer, D. C.; Sarko, A.; Bluhm, T. L.; Lee, Y. C. *Acta Crystallogr., Section B* **1980**, *36*, 650.
39. Ilin, S. G.; Lindeman, S. V.; Struchkov, Yo. T.; Levina, E. V.; Storik, V. A.; Yelyakov, G. B. *Dokl. Akad. SSSR* **1992**, *323*, 290.
40. Taga, T.; Inagaki, E.; Fujmori, Y.; Fujita, K.; Hara, K. *Carbohydr. Res.* **1993**, *241*, 63.
41. Duee, E.; Grand, A.; Tran, V. H. *Acta Crystallogr., Section B* **1981**, *37*, 850.

42. Kanters, J. A.; Schouten, A.; van Bommel, M. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1990**, *46*, 2408.
43. Yamada, H.; Nishizawa, M.; Katayama, C. *Tetrahedron Lett.* **1992**, *33*, 4009.
44. Kanters, J. A.; Schrenberg, R. L.; Leeflang, B. R.; Kroon, J.; Mathlouthi, M. *Carbohydr. Res.* **1988**, *180*, 175.
45. Jin, G.; Yamagata, Y.; Tomita, K. *Chem. Pharm. Bull.* **1990**, *38*, 297.
46. Bugg, C. E. *J. Amer. Chem. Soc.* **1973**, *95*, 908.
47. Cook, W. J.; Bugg, C. E. *Acta Crystallogr., Section B* **1973**, *29*, 907.
48. Beerers, C. A.; Hansen, H. N. *Acta Crystallogr., Section B* **1971**, *27*, 1323.
49. Noordik, J. H.; Beurskens, P. T.; Bennema, P.; Visser, R. A.; Gould, R. O. *Z. Kristallogr.* **1984**, *168*, 59.
50. Fries, D. C.; Rao, S. T.; Sundaralingam, M. *Acta Crystallogr., Section B* **1971**, *27*, 994.
51. Ohanessian, J.; Longchambon, F.; Arene, F. *Acta Crystallogr., Section B* **1978**, *34*, 3666.
52. Oliver, J. D.; Strickland, L. C. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1984**, *40*, 820.
53. Hanson, J. C.; Sieker, L. C.; Jensen, L. H. *Acta Crystallogr., Section B* **1973**, *29*, 797.
54. Brown, G. M.; Levy, H. A. *Acta Crystallogr., Section B* **1973**, *29*, 790.
55. Hynes, R. C.; Le Page, Y. *J. Appl. Crystallogr.* **1991**, *24*, 352.

56. Thiem, J.; Kleeberg, M.; Klaska, K. H. *Carbohydr. Res.* **1989**, *189*, 65.
57. Quingley, G. J.; Sarko, A.; Merchessault, R. H. *J. Amer. Chem. Soc.* **1970**, *92*, 5834.
58. Gress, M. E.; Jeffrey, G. A. *Acta Crystallogr., Section B* **1977**, *33*, 2490.
59. Takusagawa, F.; Jacobson, R. A. *Acta Crystallogr., Section B* **1978**, *34*, 213.
60. Ham, J. T.; Williams, D. G. *Acta Crystallogr., Section B* **1970**, *26*, 1373.
61. Cannon, J. R.; Raston, C. L.; Toia, R. F.; White, A. H. *Aust. J. Chem.* **1980**, *33* 2229.
62. Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1967**, *23*, 1038.
63. Foces-Foces, C.; Cano, F. H.; Garcia-Blanco, S. *Acta Crystallogr., Section B* **1981**, *37*, 1270.
64. Tanase, t.; Nouchi, r.; Oka, y.; Kato, M.; Nakamura, N.; Yamamura, T.; Yamamoto, Y.; Yano, S. *J. Chem. Soc., Dalton Trans.* **1993**, 2645.
65. Van der Sluis, P.; Kroon, J. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1990**, *46*, 2171.
66. Kivikoski, J.; Pitkanen, I.; Nurmi, J. *Carbohydr. Res.* **1992**, *232*, 189.
67. Jin, G; Fujii, S.; Tomita, K. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1990**, *46*, 1866.
68. Kanters, J. A.; Shouter, A.; van Bommel, M. *J. Mol. Struct.* **1990**, *238*, 403.
69. Kivikoski, J.; Valkomen, J.; Nurmi, J. *Carbohydr. Res.* **1992**, *232*, 53.
70. Lamba, D.; Segre, A. L.; Glover, S.; Mackie, W.; Sheldrick, B.; Perez, S. *Carbohydr. Res.* **1990**, *208*, 215.

71. Lee, C. K.; Koh, L.L. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1993**, *49*, 621.
72. Noguchi, K.; Okuyama, K.; Kitamura, S.; Takeo, K. *Carbohydr. Res.* **1992**, *237*, 33.
73. Koll, P.; Petrusova, M.; Petrus, I.; Zimmer, B.; Morf, M.; Kopf, J. *Carbohydr. Res.* **1993**, *248*, 37.
74. Lindon, A.; Lee, C. K. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1995**, *51*, 1007.
75. Lee, C. K.; Linden, A. *Carbohydr. Res.* **1994**, *264*, 319.
76. Lindon, A.; Lee, C. K. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1995**, *51*, 1012.
77. Prawit, H.; Mahidol, C.; Ruchirawat, S.; Prawat, U.; Tuntiwachwuttikul, P.; Tooptakong, U.; Tayler, W. C.; Pakawachai, C.; Skelton, B. W.; White, A. H. *Phytochemistry* **1995**, *40*, 1167.
78. Kanie, O.; Takeda, T.; Hatano, K. *Carbohydr. Res.* **1995**, *276*, 409.
79. Wieczorek, M. W.; Blaszczyk, J.; Krol, B. W. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1996**, *52*, 1193.
80. Brisse, F.; Marchessault, R. H.; Perez, S.; Zugenmaier, P. *J. Amer. Chem. Soc.* **1982**, *104*, 7470.
81. Kivikoski, J.; Pitkaren, I.; Valkomen, J.; Heikkila, H. *Carbohydr. Res.* **1992**, *223*, 45.
82. Connor, S. O.; Lam, L. K. T.; Jones, N. D.; Chasey, M. O. *J. Org. Chem.* **1976**, *41*, 2087.

83. Perez, S.; Brisse, F. *Acta Crystallogr., Section B* **1977**, *33*, 2578.
84. Hoogendorp, J. O.; de Kok, A. J.; Romers, C. *Carbohydr. Res.* **1983**, *114*, 169.
85. Moran, R. A.; Richards, G. F. *Acta Crystallogr., Section B* **1973**, *29*, 2770.
86. Taga, T.; Inagaki, E.; Fujmori, Y.; Nakamura, S. *Carbohydr. Res.* **1993**, *240*, 39.
87. Taga, T.; Inagaki, E.; Fujmori, Y.; Nakamura, S. *Carbohydr. Res.* **1994**, *251*, 203.
88. Jeffrey, G. A.; Huang, D. *Carbohydr. Res.* **1991**, *222*, 47.
89. Bequart, J.; Neuman, A.; Gillier-Pandraud, H. *Carbohydr. Res.* **1982**, *111*, 9.
90. Jeffrey, G. A.; Park, Y. J. *Acta Crystallogr., Section B* **1972**, *28*, 257.
91. Rohrer, D. C. *Acta Crystallogr., Section B* **1972**, *28*, 425.
92. Ferretti, V.; Bertolasi, V.; Gilli, G.; Accorsi, C. A. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1984**, *40*, 531.
93. Berman, H. M. *Acta Crystallogr., Section B* **1970**, *26*, 290.
94. Jeffrey, G. A.; Huang, D. *Carbohydr. Res.* **1990**, *206*, 173.
95. Avenel, D.; Neuman, A.; Gillier-Pandraud, H. *Acta Crystallogr., Section B* **1976**, *32*, 2598.
96. Imberty, A.; Perez, S. *Carbohydr. Res.* **1988**, *181*, 41.
97. Carter, G. T.; Phillipson, D. W.; West, R. R.; Borders, D. B. *J. Org. Chem.* **1993**, *58*, 6588.
98. Jeffrey, G. A.; Huang, D. *Carbohydr. Res.* **1991**, *210*, 89.

99. Nardi, R. G.; Flippen-Anderson, J. L. *Acta Crystallogr., Section C (Cr. Str. Comm.)* **1987**, *43*, 806.
100. Noguchi, K.; Kobayashi, E.; Oguyama, K.; Kitamura, S.; Takeo, K.; Ohno, S. *Carbohydr. Res.* **1994**, *258*, 35.
101. Hinrichs, W.; Saenger, W. *J. Amer. Chem. Soc.* **1990**, *112*, 2789.
102. Ashton, P. R.; Brown, C. L.; Menzer, S.; Nepogodiev, S. A.; Stoddart, J. F.; Williams, D. J. *Chem. Eur. J.* **1996**, *2*, 580.
103. Perez, S.; Marchessault, R.H. *Carbohydr. Res.* **1978**, *65*, 114.
104. Chandrasekaran, R. *Adv. Carbohydr. Chem. Biochem.* **1997**, *52*, 311.
105. Molecular Simulations, Inc. Waltham, MA 02154 U. S. A.



## **Chapter 3**

### **Design and Synthesis of Compounds with Rigid Glycosidic Linkages**

## Abstract

One of the important methods for determining the three-dimensional structures of carbohydrates is using the proton-carbon coupling constants to calculate the critical glycosidic dihedral angle values. In this work we report the synthesis of seven compounds that have rigid glycosidic linkages. These are: 1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\alpha$ -L-fucopyranoside (1), 4,6-O-benzylidene-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\alpha$ -D-glucopyranoside (2), 1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -L-fucopyranoside (3), 3-O-methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (4), 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (5), 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (6), and 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (7). These compounds are important to parameterize a new Karplus-type equation relating the glycosidic dihedral angles to the vicinal proton-carbon coupling constant of carbohydrates.

## Introduction

Previous equations by Tvaroska and coworkers<sup>1</sup> and Mulloy and coworkers<sup>2</sup> for determining glycosidic dihedral angles of carbohydrates from vicinal proton-carbon coupling constants were parameterized by using vicinal proton-carbon heteronuclear coupling constants measured in solution and dihedral angles taken from crystal structures. While the solution structure of carbohydrates is an average of several conformations, the crystal structure represents one of the solution structures where crystal-packing forces play a role. In addition, while our study (Chapter Two) revealed that glycosidic dihedral angles defined by the four atoms H-C-O-C are almost always between 0 and 90 degrees, the equations by Tvaroska and coworkers<sup>1</sup> and Mulloy and coworkers<sup>2</sup> are parameterized with a very small number of data points in the range between 0 and 90 degrees..

It would then be very valuable to design and synthesize compounds with rigid glycosidic linkages with predetermined dihedral angles and measure the interglycosidic heteronuclear coupling constants in solution. By relating the values of the dihedral angles to the three-bond heteronuclear coupling constant,  $^3J_{CH}$ , we can then come up with a new and more accurate Karplus type equation for carbohydrates.

## Design and Synthesis

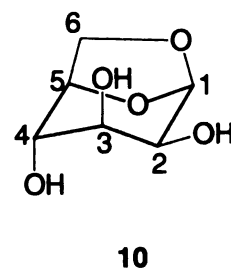
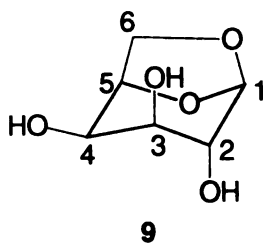
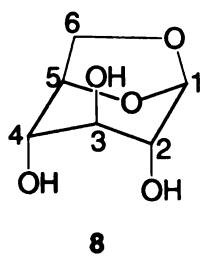
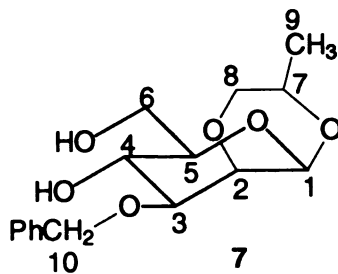
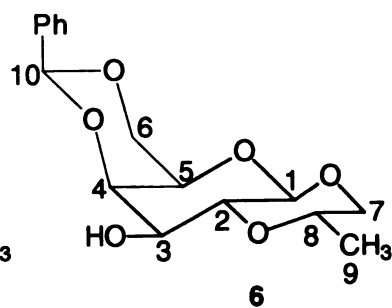
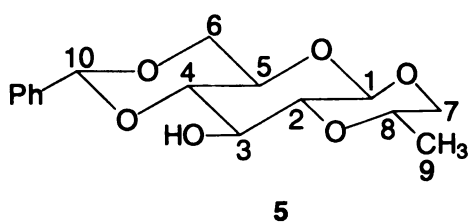
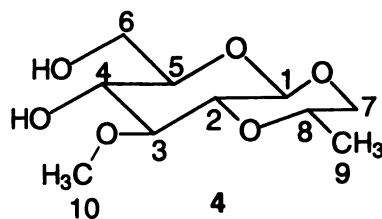
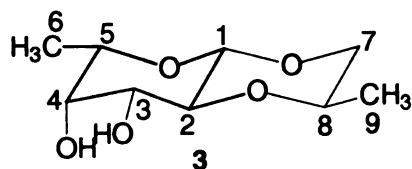
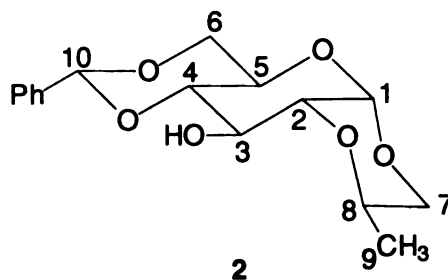
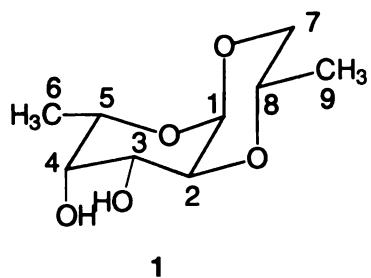
We designed and synthesized seven compounds with rigid glycosidic linkages for the purpose of parameterizing a new Karplus-type equation relating

carbohydrate glycosidic dihedral angles and vicinal proton-carbon coupling constants. The method we chose to rigidize the glycosidic linkage is forming a six-membered ring around and including the O2-C2-C1-O1 fragment of monosaccharides (**Figure 3.1**). We prepared the seven compounds starting from different monosaccharides, which commonly occur in nature. These compounds have different electronic environments around the anomeric carbon (C1) because of the different substitutions and/or configurations at the different sites of the monosaccharides. This is important because, the electronic environment around the glycosidic linkages may affect the heteronuclear spin – spin coupling constant, which is transmitted through bonding electrons.

The synthesis is designed in such away that the 2-hydroxy group of the monosaccharides acts as a nucleophile for asymmetric oxidation of the prochiral C=C double bond of the allyl group attached to the anomeric position (**Scheme 3.1 – 3.6**).

The set of compounds we plan to prepare include  $\alpha$ - and  $\beta$ -linked glycosides with varying glycosidic dihedral angles. In addition, the compounds represent the most common monosaccharide units in carbohydrates, namely, glucose, galactose, mannose and fucose. As mentioned earlier we plan to prepare compounds with rigid dihedral angles by forming a six-membered ring around the O2-C2-C1-O1 portion of the monosaccharides. The problem with making the ring involving the glycosidic bonds is that the values of the glycosidic dihedral angles we will be able to make are limited to certain values. But we believe that the Karplus type equation parameterized for these dihedral angles

**Figure 3.1** The ten molecules that have rigid glycosidic dihedral angles. While molecules **8-10** are commercially available, molecules **1-7** are synthesized in this work.

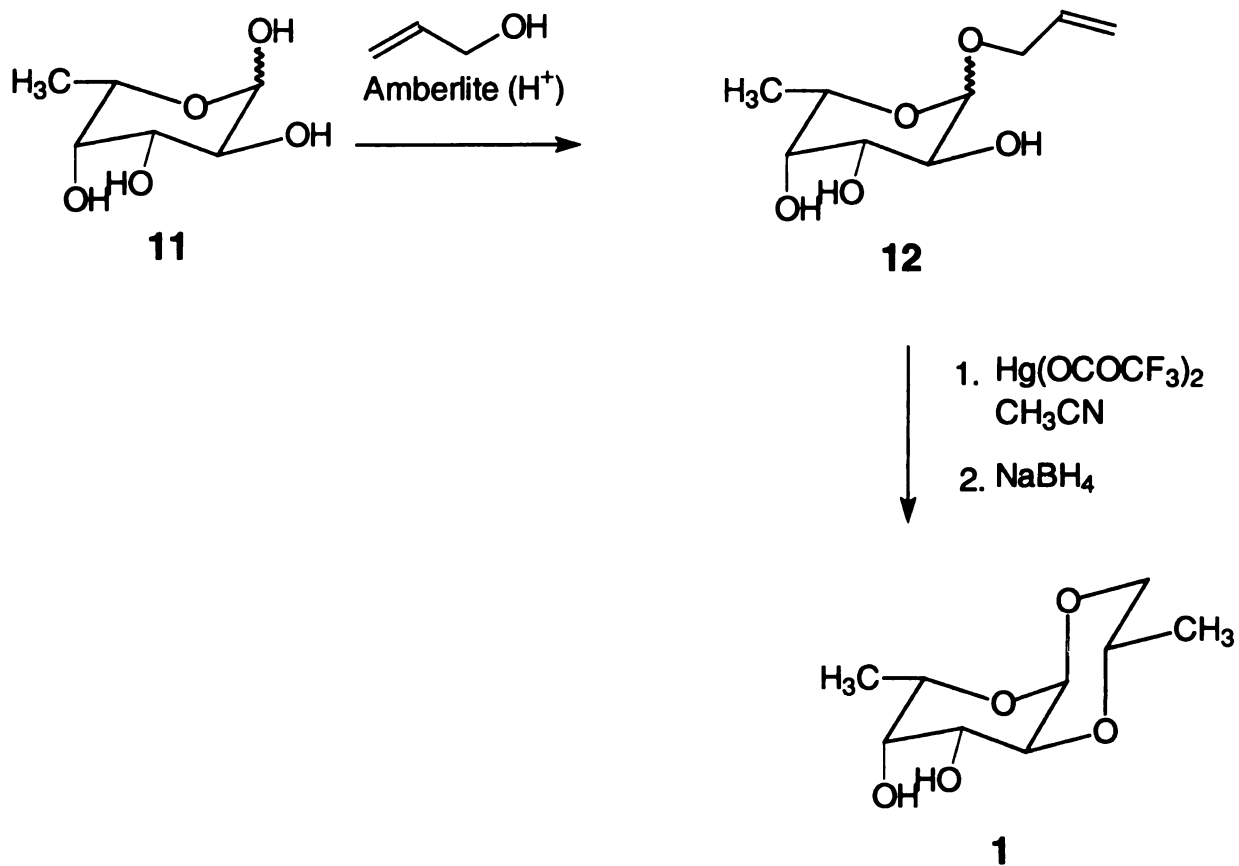


should reasonably predict the coupling constant or dihedral angle values for the remaining range. The compounds we prepare will have around 60 or 180 degrees for the glycosidic dihedral angles. For the dihedral angles between 60 and 180 degrees, we will use the commercially available 1,6-anhydroglucopyranose (8), 1,6-anhydrogalactopyranose (9), and 1,6-anhydromannopyranose (10) (Figure 3.1). These three compounds cover the range of about 110 to 165 degrees.

### Preparation of $\alpha$ -glycosides

The  $\alpha$ -glycosides with rigid glycosidic linkages were prepared by first allylating the anomeric carbon by allyl alcohol using the traditional Fischer glycosidation. This was followed by oxymercuration and intramolecular nucleophilic attack by the 2-hydroxyl group of the monosaccharide<sup>3</sup> to form the six-membered ring with high or exclusive preference of the CH<sub>2</sub>HgOAc group for the equatorial position. Finally, sodium borohydride reduction gave the final products (Scheme 3.1 and 3.2).

The Fischer glycosidation of L-fucose gave almost exclusively the  $\alpha$ -L-fucoside (12) while that of D-glucose gave a mixture of  $\alpha$ - and  $\beta$ -glucosides with a quantitative yield. To separate the  $\alpha$ - and  $\beta$ -glucosides and to improve their solubility in non-aqueous solvents, they were benzylated<sup>4</sup> using benzylidene dimethyl acetal in dimethyl formamide in the presence of a catalytic amount of p-toluene sulfonic acid (pTSA) to give allyl 4,6-O-benzylidene-D-

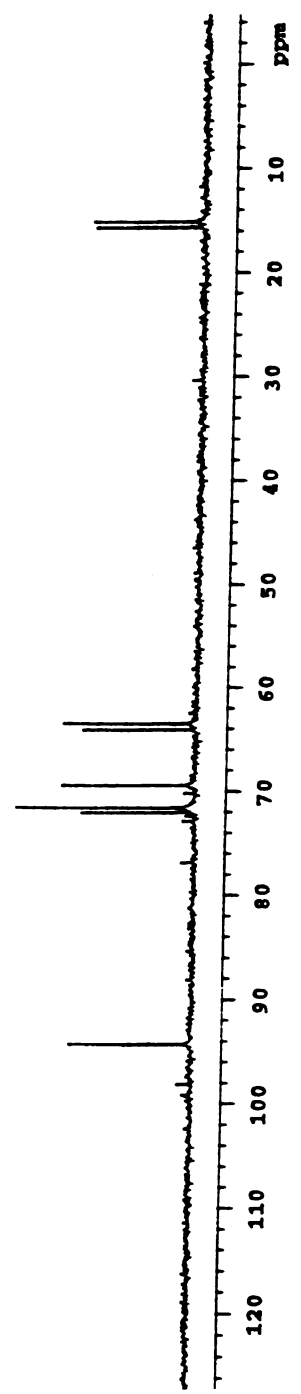
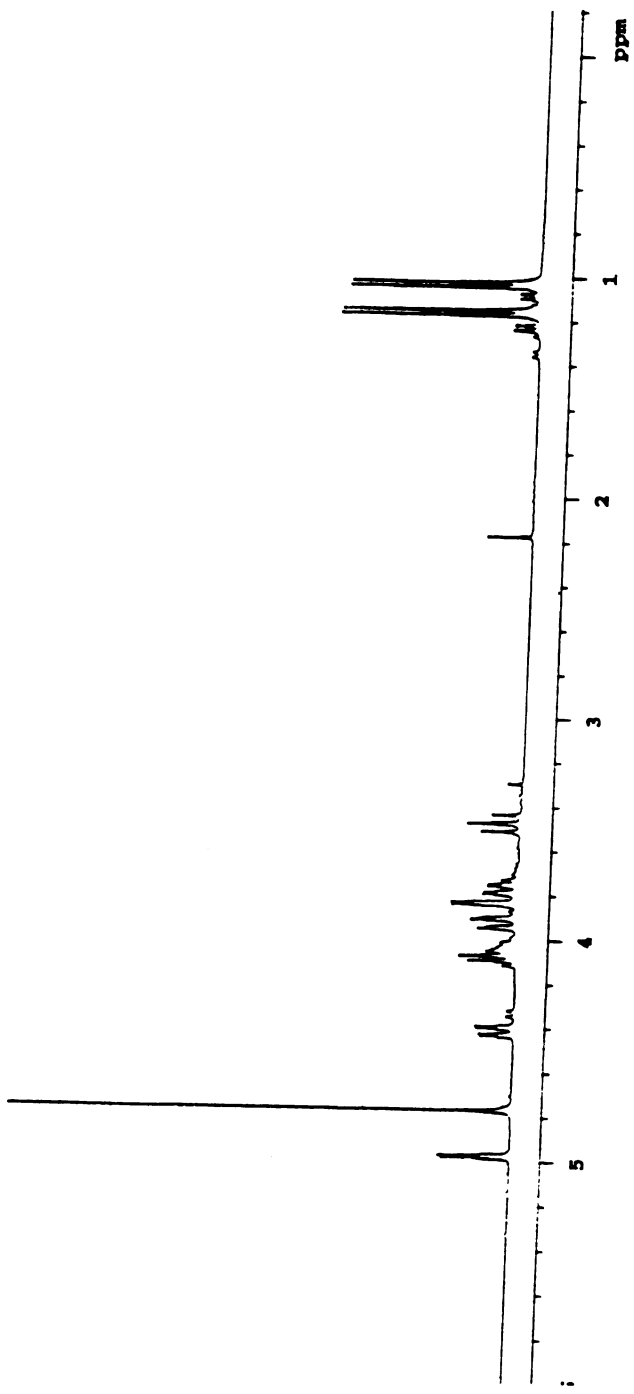


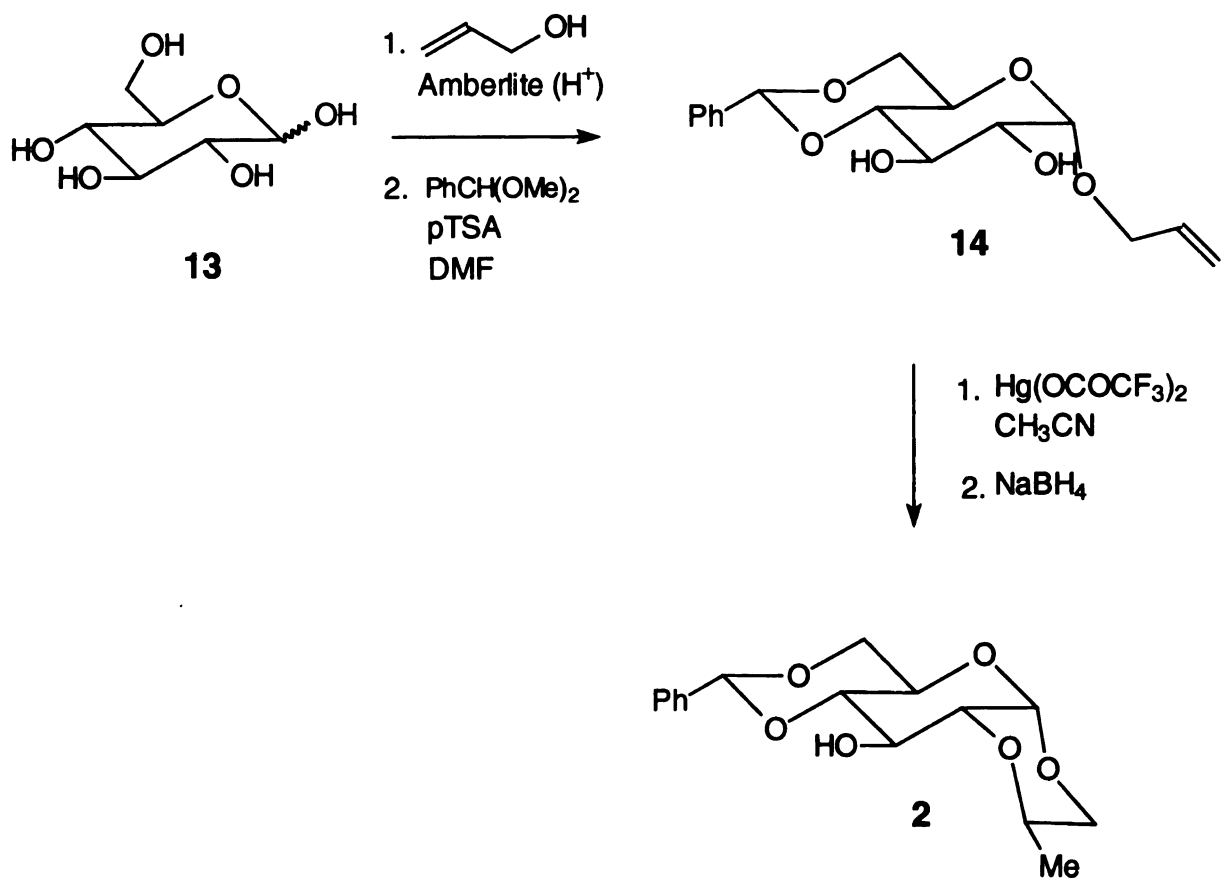
**Scheme 3.1** Synthesis of 1,2-O-[2-methyl-(S)-1,2-ethanediyl]-α-L-fucopyranoside

(1).



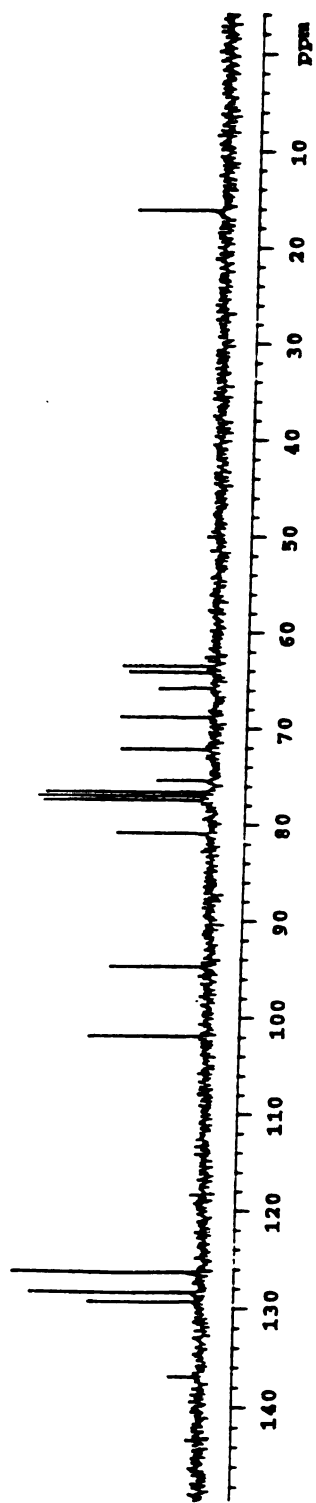
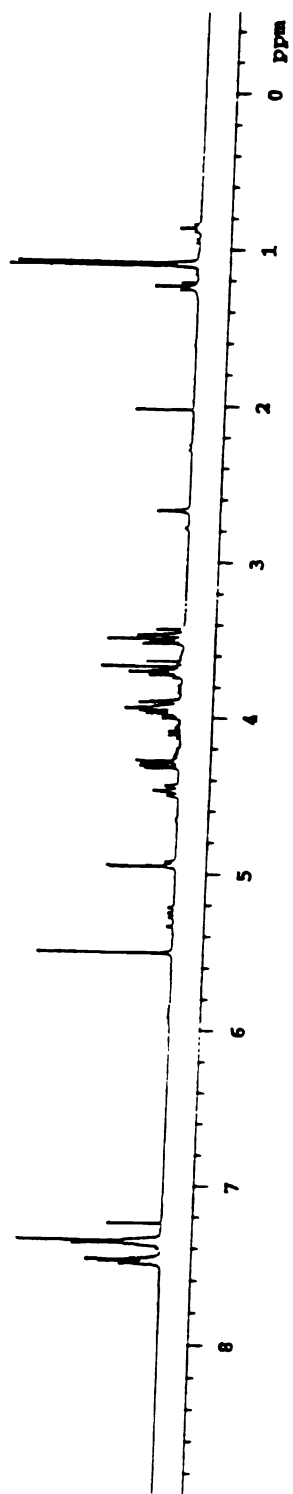
**Figure 3.2** Proton and carbon nuclear magnetic resonance spectra of 1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\alpha$ -L-fucopyranoside (1).





**Scheme 3.2** Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\alpha$ -D-glucopyranoside (**2**).

**Figure 3.3** Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzylidene-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\alpha$ -D-glucopyranoside (**2**).

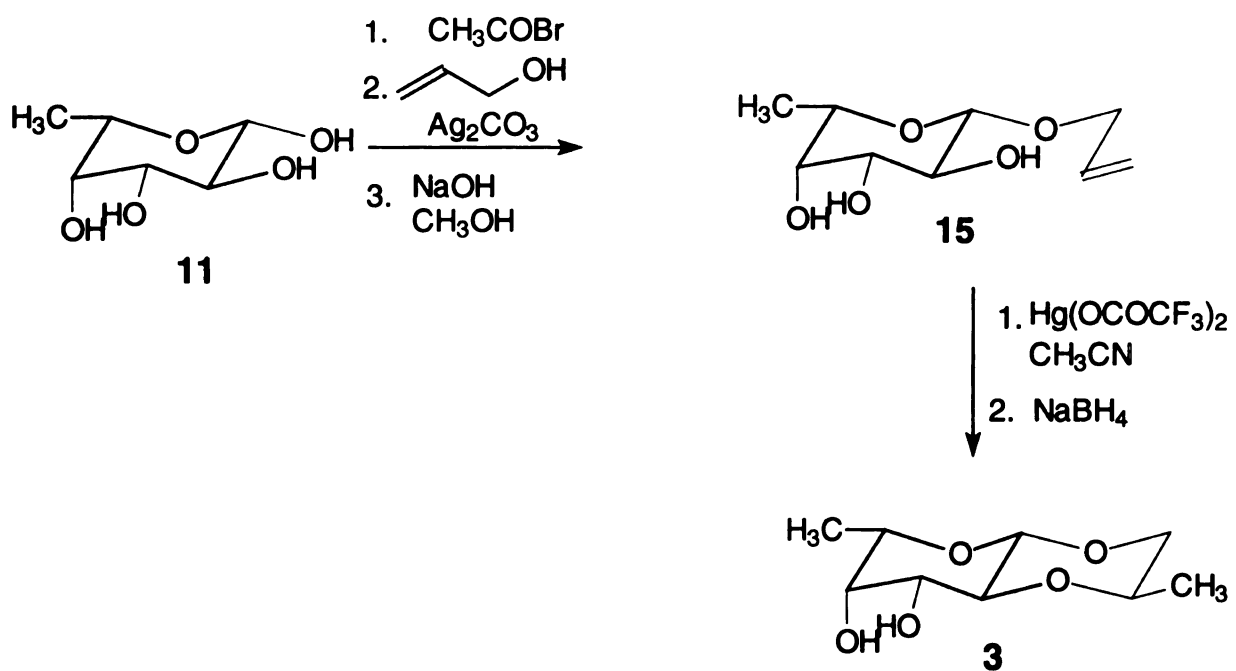


glucosides. Crystallization in hexane-ethyl acetate (2:1) gave crystals of the allyl 4,6-O-benzylidene- $\alpha$ -D-glucoside (**14**). Oxymercuration of compounds **12** and **14** was then performed using mercuric trifluoroacetate in acetonitrile. Attack of the oxymercured double bond by the hydroxyl group at C2 position formed a six-membered ring, creating a new chiral center, C8. The methylene mercuric trifluoroacetate group is almost exclusively at the equatorial position because of the 1,3-diaxial interactions between C3 and the CH<sub>2</sub>HgOAc group if it was to be at the axial position. Finally reduction of the cyclized products with sodium borohydride in methanol converted the CH<sub>2</sub>HgOAc group into a methyl group affording the final products (**1** and **2**).

### **Preparation of $\beta$ -glycosides**

The  $\beta$ -glycosides (**3-6**) were prepared by allylation of the monosaccharides using the Koenigs-Knorr reaction followed by oxymercuration and intramolecular nucleophilic attack by the 2-hydroxyl group of the monosaccharides to form a six-membered ring and sodium borohydride reduction to give the final cyclic  $\beta$ -glycosides (**Schemes 3.3 – 3.6**).

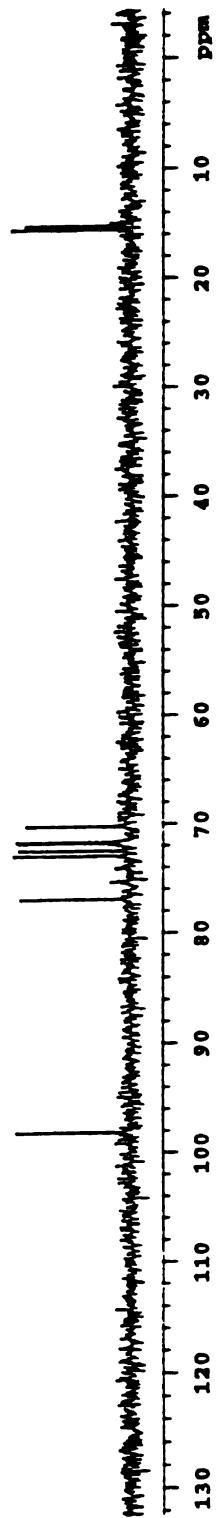
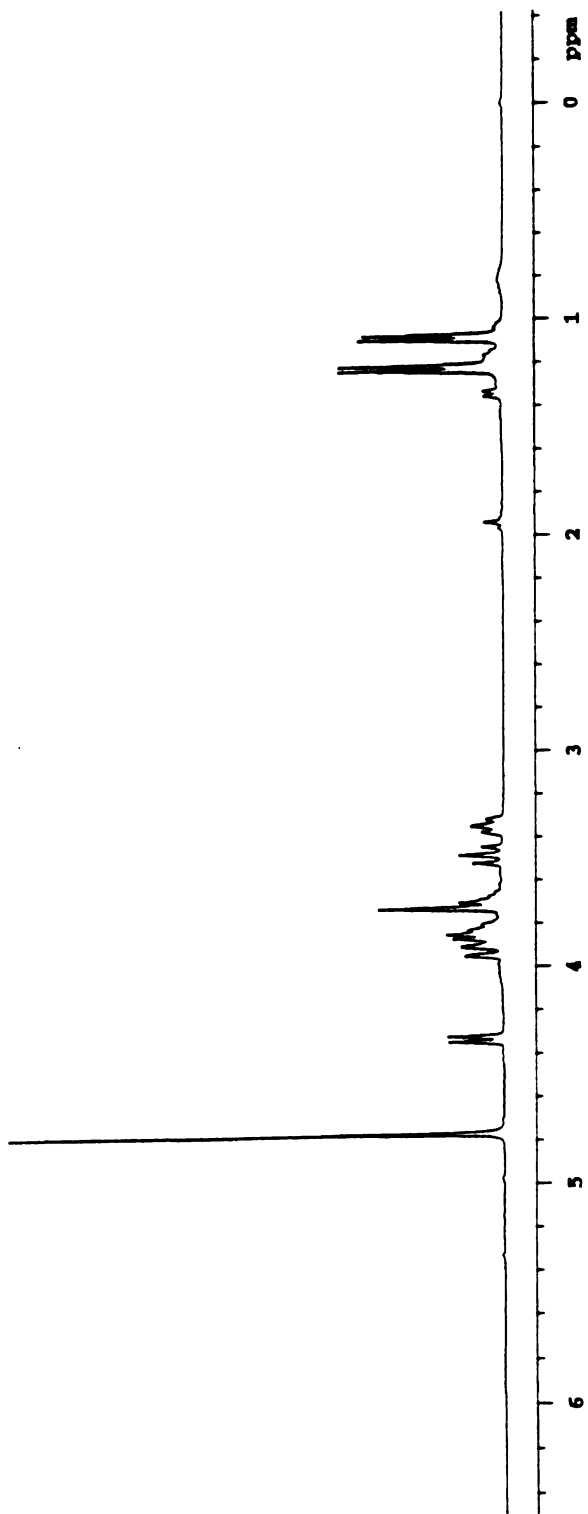
Starting from the commercially available D-glucose, D-galactose, L-fucose and 3-O-methyl D-glucose, acetobromination<sup>5</sup> was accomplished using acetyl bromide in acetic acid. This activates the anomeric carbon by putting a bromine group at it and protects the 2-, 3-, 4- and 6-hydroxyl groups with acetyl groups. After rotaevaporation of the solvent, the remaining residue was allylated<sup>6</sup> by allyl

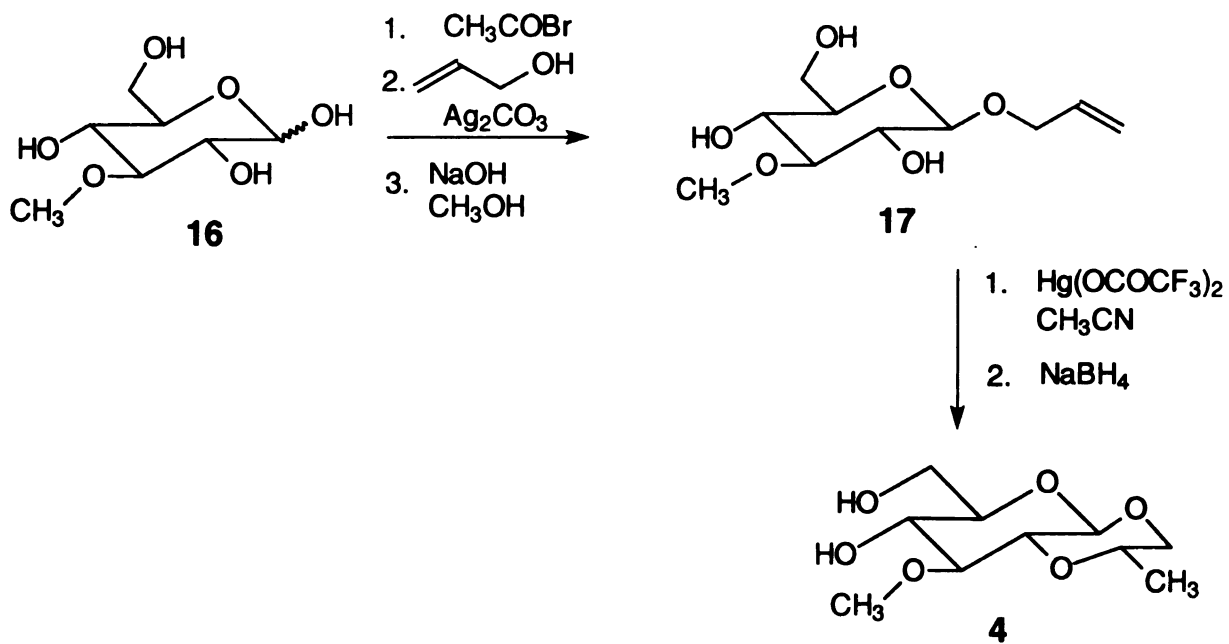


**Scheme 3.3** Synthesis of 1,2-O-[2-methyl-(R)-1,2-ethanediyl]-β-L-fucopyranoside (3).

**Figure 3.4** Proton and carbon nuclear magnetic resonance spectra of 1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -L-fucopyranoside (**3**).

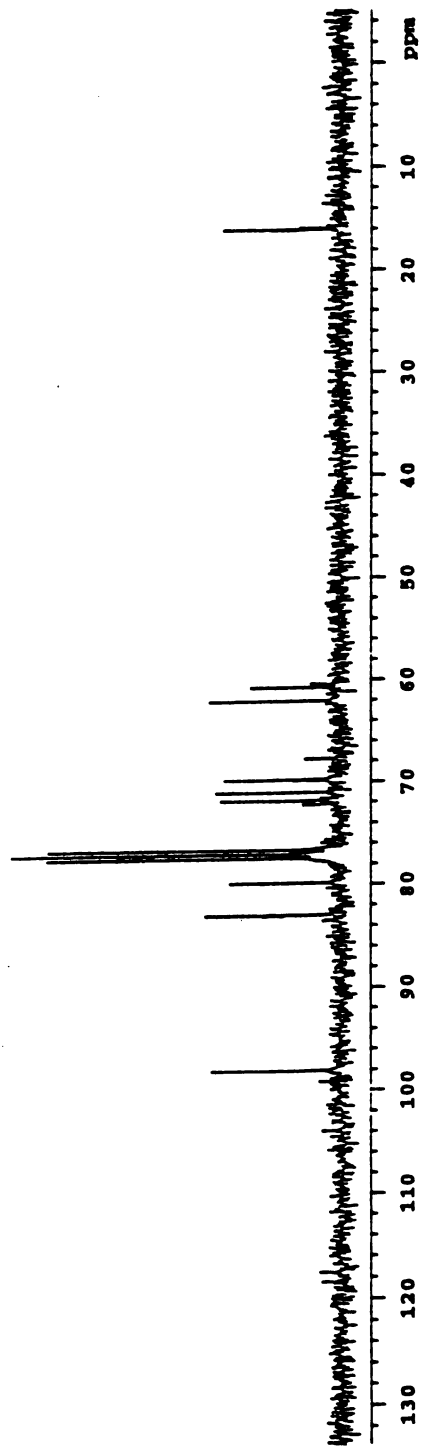
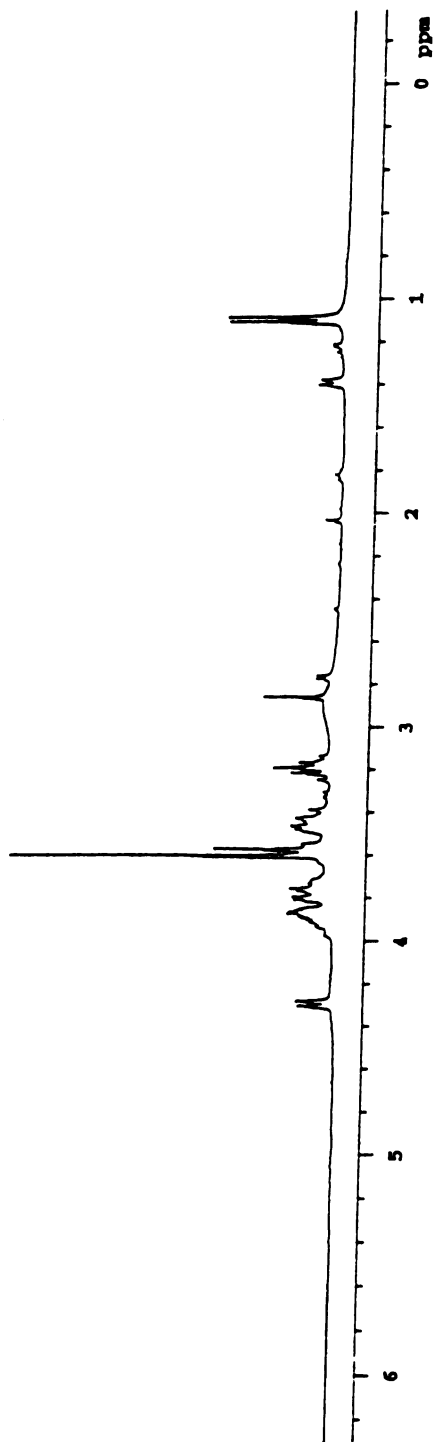






**Scheme 3.4** Synthesis of 3-O-methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]-β-D-glucopyranoside (**4**).

**Figure 3.5** Proton and carbon nuclear magnetic resonance spectra of 3-O-methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (**4**).



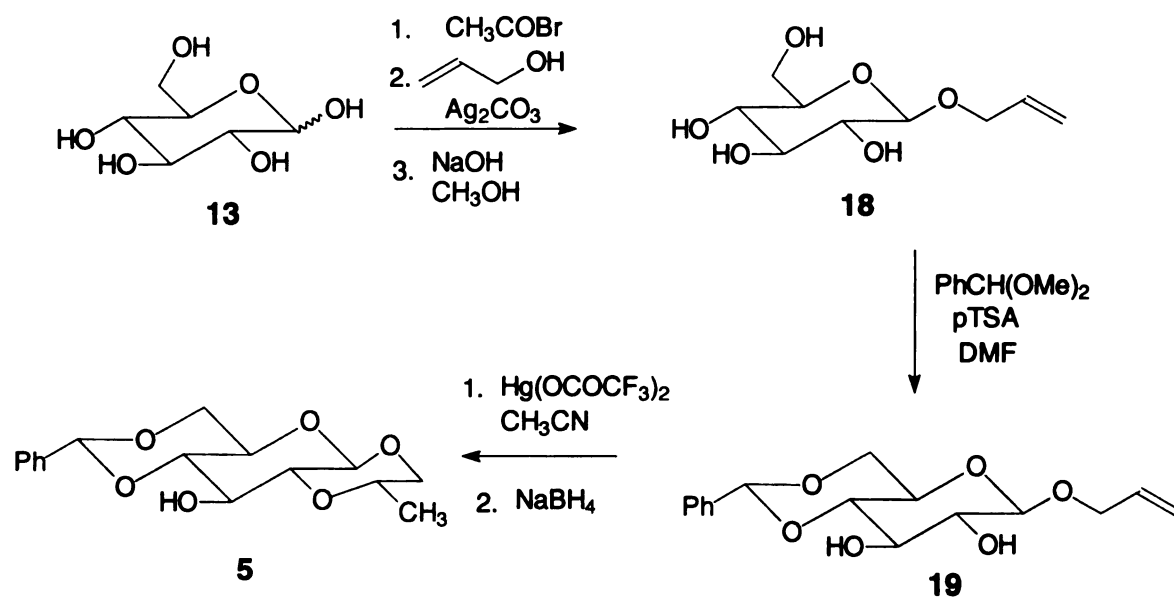
alcohol in the presence of silver carbonate to give exclusively the allyl  $\beta$ -glycosides. The acetyl groups at the 2-, 3-, 4- and 6-positions were then deacetylated using sodium methoxide in methanol to give the allyl  $\beta$ -L-fucoside (15), allyl 3-O-methyl- $\beta$ -D-glucoside (17), allyl  $\beta$ -D-glucoside (18), allyl  $\beta$ -D-galactoside (21).

Allyl  $\beta$ -D-glucoside (18) and allyl  $\beta$ -D-galactoside (21) were benzylated by benzylidene dimethyl acetal in dimethyl formamide in the presence of a catalytic amount of *p*-toluene sulfonic acid to give allyl 4,6-O-benzylidene- $\beta$ -D-glucoside (19) and allyl 4,6-O-benzylidene- $\beta$ -D-galactoside (22), respectively. The allyl  $\beta$ -glycosides 15, 17, 19 and 22 were then oxymercured using mercuric trifluoroacetate in acetonitrile. Attack of the oxymercured double bond by the hydroxyl group at C2 position formed a six-membered ring, creating a new chiral center, C8. These compounds were finally reduced with sodium borohydride similar to the allyl  $\alpha$ -glycosides, the  $\text{CH}_2\text{HgOAc}$  group being converted into a methyl group to give the final products (3-6).

### **Preparation of the $\beta$ -mannoside**

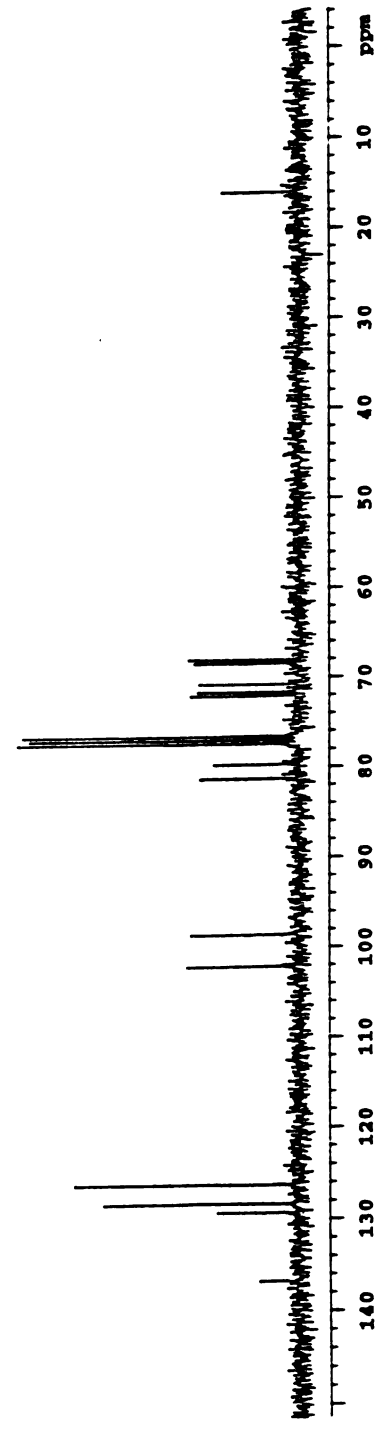
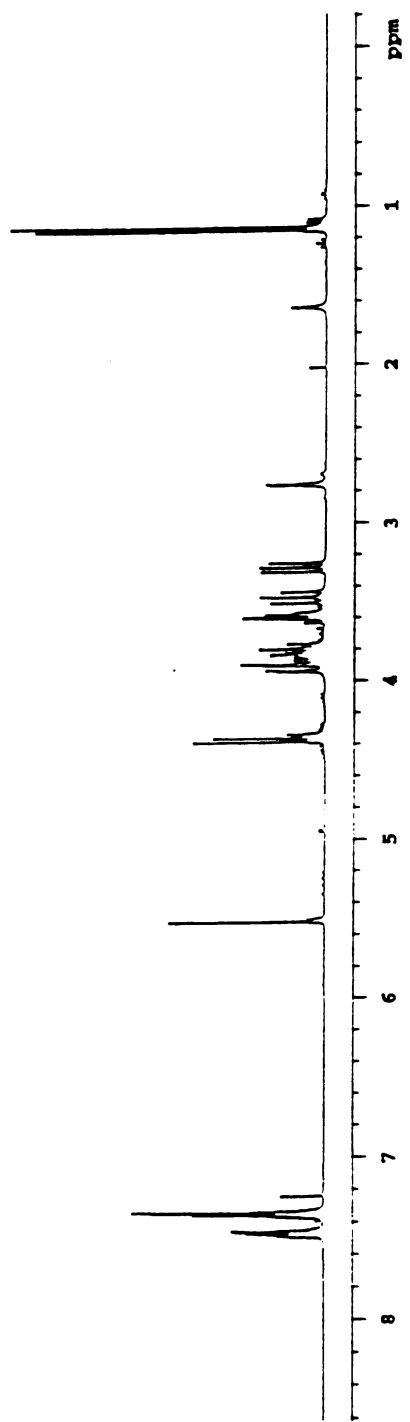
Our attempt to prepare the  $\beta$ -mannoside using a similar procedure to that of the other  $\beta$ -glycosides was not successful. As a result we designed a different method to prepare the cyclic  $\beta$ -mannoside (7) (Scheme 3.7).

Starting from the commercially available methyl  $\alpha$ -mannoside (23), benzylidene acetal was performed by using benzylidene dimethyl acetal in

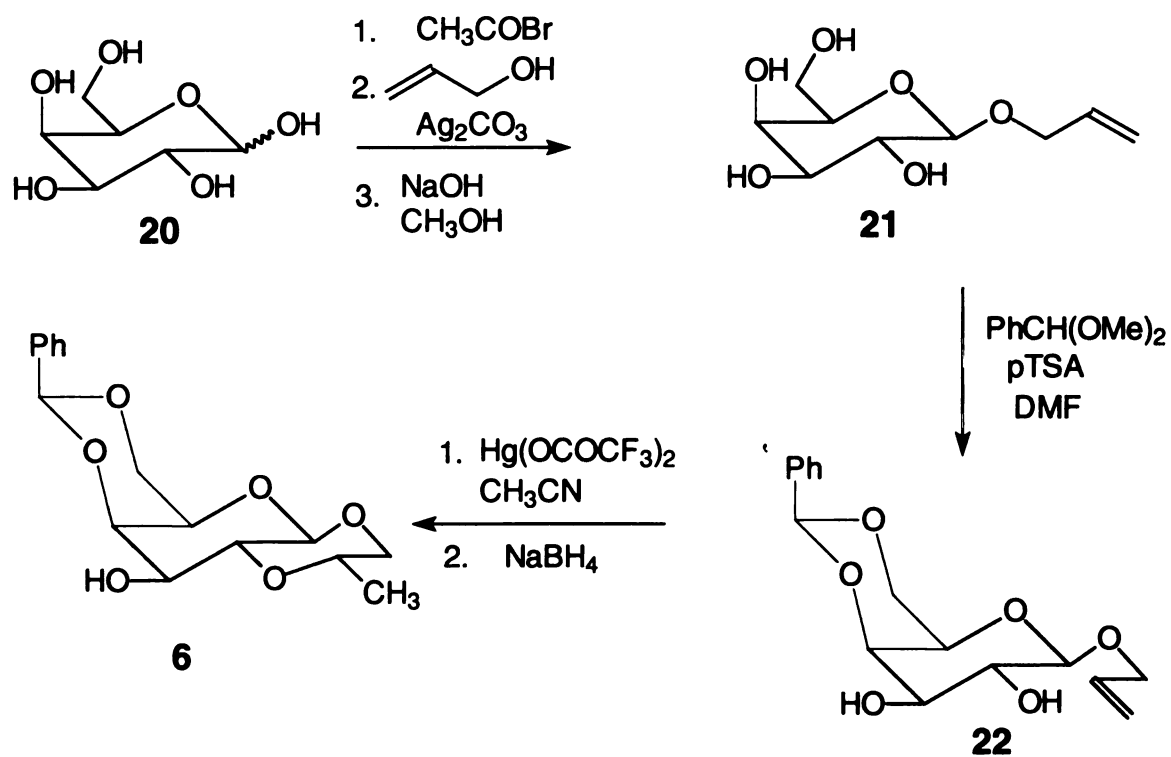


**Scheme 3.5** Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]-β-D-glucopyranoside (**5**).

**Figure 3.6** Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (**5**).

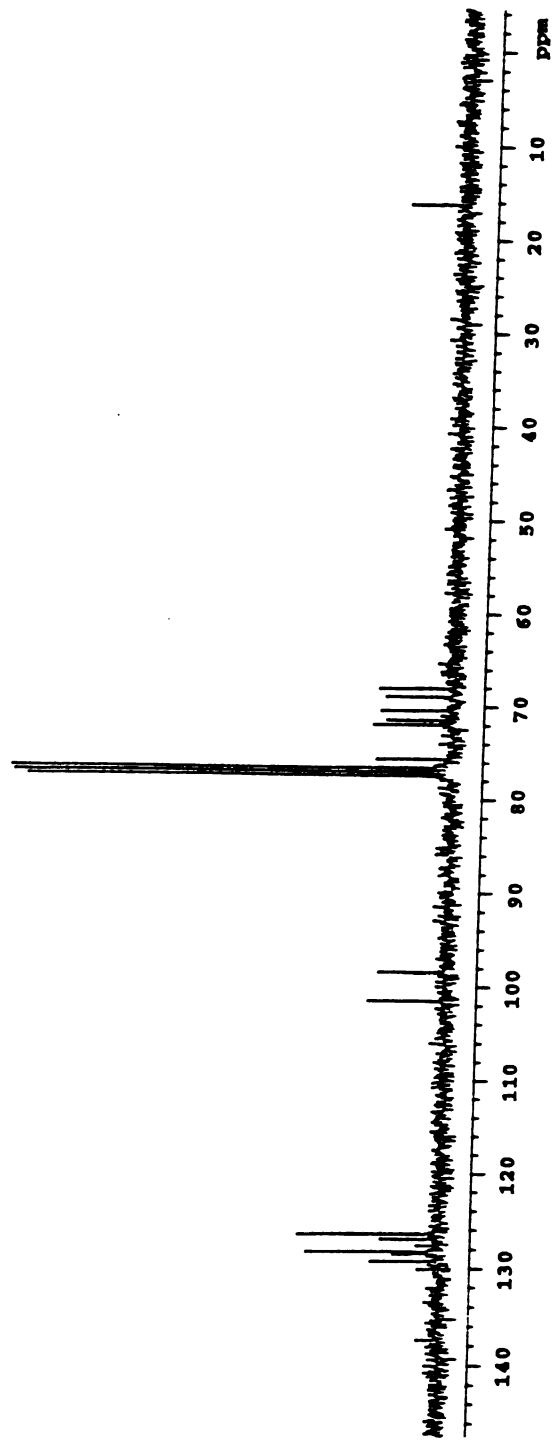
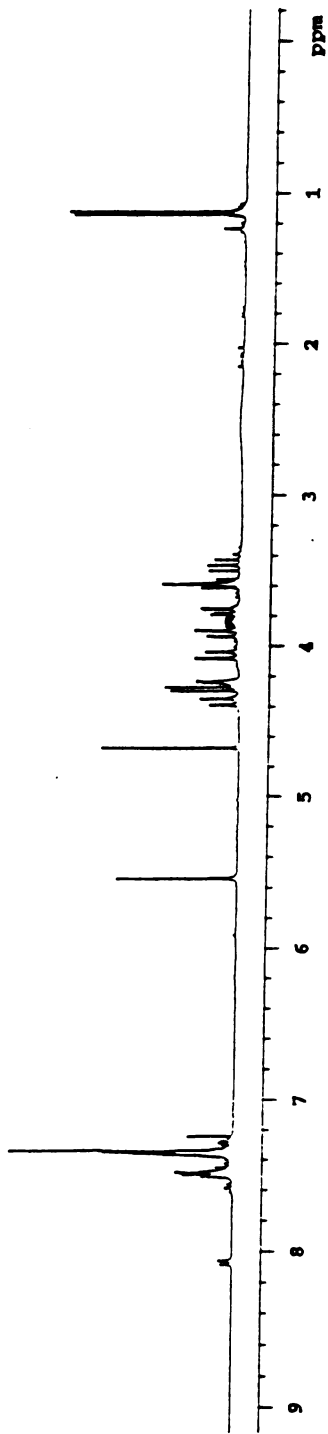


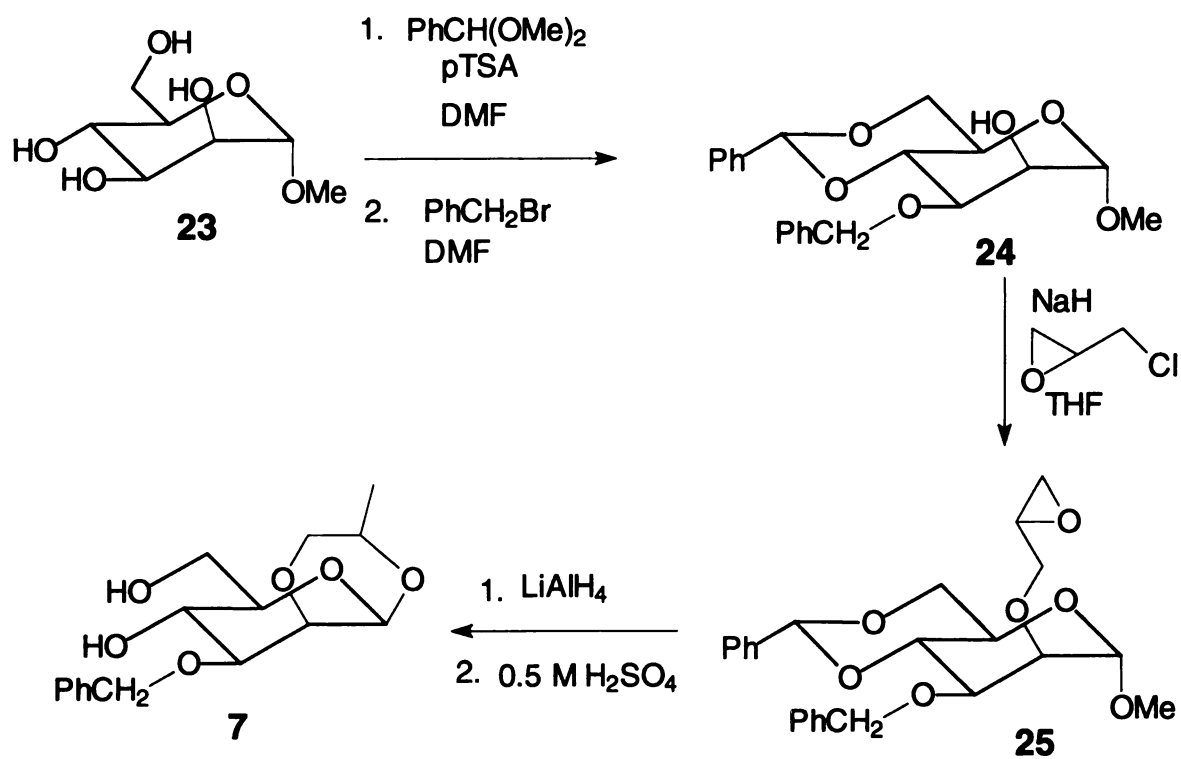




**Scheme 3.6** Synthesis of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]-β-D-galactopyranoside (**6**).

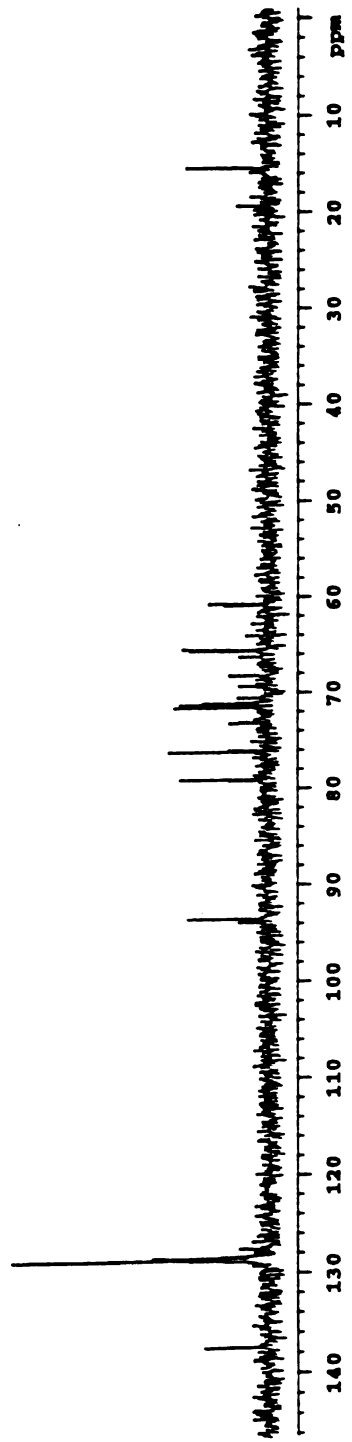
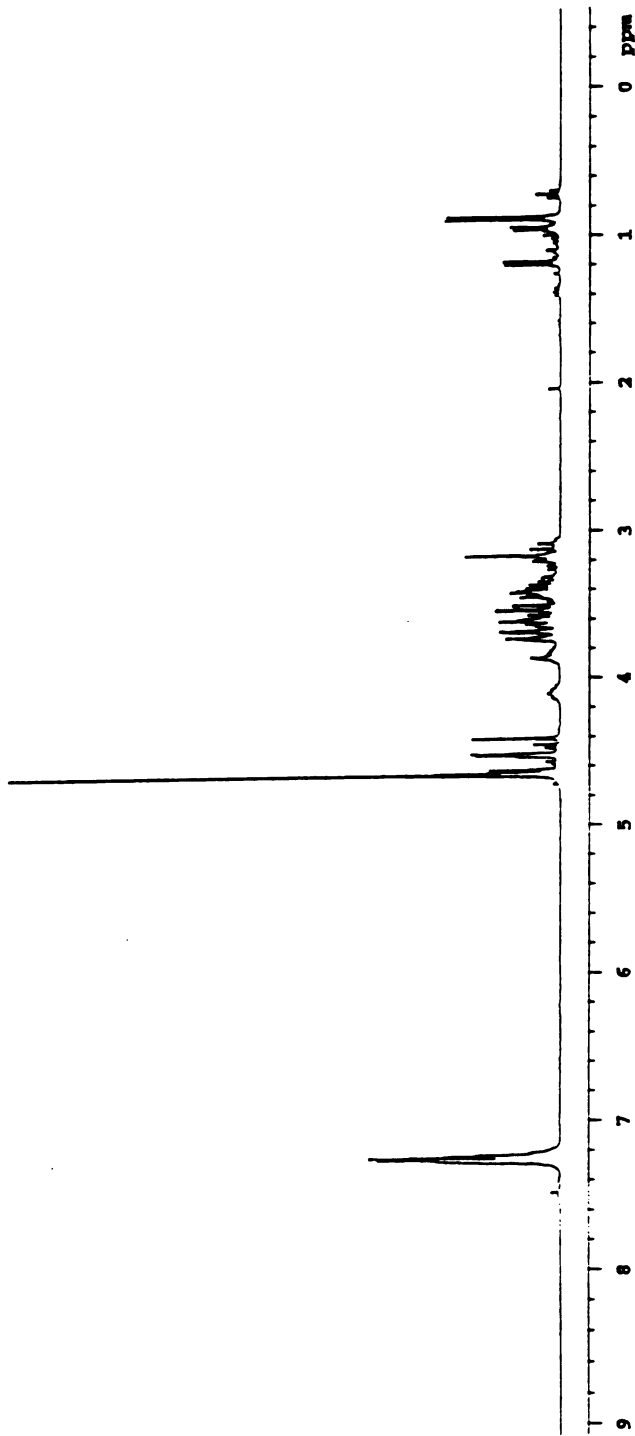
**Figure 3.7** Proton and carbon nuclear magnetic resonance spectra of 4,6-O-benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (**6**).





**Scheme 3.7** Synthesis of 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]-β-D-mannopyranoside (**7**).

**Figure 3.8** Proton and carbon nuclear magnetic resonance spectra of 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (**7**).



dimethyl formamide in the presence of a catalytic amount of p-toluene sulfonic acid to give methyl 4,6-O-benzylidene- $\alpha$ -D-mannoside. This resulted in protection of the 4- and 6-hydroxyl groups of the methyl  $\alpha$ -D-mannoside. We then protected the 3- hydroxyl group using benzyl bromide<sup>7</sup> in tetrahydrofuran to get methyl 3-O-benzyl-4,6-O-benzylidene- $\alpha$ -D-mannoside (**24**). Reaction of **24** with racemic epichlorohydrin gave the diastereomeric **25**. This was followed by lithium aluminum hydride reduction<sup>8</sup> of the epoxide and acid hydrolysis<sup>9</sup> to give the final product 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (**7**) after purification by flash chromatography using methanol-chloroform (1:6).

## Conclusions

Seven compounds (**1-7**) having rigid and predetermined glycosidic dihedral angles were designed and synthesized starting from commercially available monosaccharide units. Their synthesis was accomplished by allylation (Fischer glycosidation for  $\alpha$ -glycosides and Koenigs-Knorr reaction for  $\beta$ -glycosides), oxymercureic activation of the allylic double bond, and intramolecular attack by the 2-hydroxy group to form a six-membered ring and finally reduction by sodium borohydride. Synthesis of **7** required a different method as shown in **Scheme 3.7**. These compounds (**1-7**) are important for parameterization of a new Karplus-type equation relating the vicinal proton-carbon heteronuclear coupling constant,  $^3J_{CH}$ , with dihedral angles defined by H-C-O-C (see Chapter Four).

## Experimental

### General Techniques

All reagents used were reagent grade. Reaction temperatures were measured externally. NMR spectra were recorded at 300 MHz with a Varian spectrometer at the Max T. Rogers NMR Facility, Michigan State University. They were obtained at 20 °C and chemical shifts are reported relative to the residue solvent peak unless otherwise specified. Melting points were measured using Fischer-Johns Melting Point Apparatus. Optical rotations were measured using a Perkin Elmer polarimeter at 589 nm. Infra Red spectra were recorded using Nicolet FTIR spectrometer. High-resolution mass spectra (HRMS) were recorded on a JEOL HX-110-HF spectrometer using Fast Atom Bombardment (FAB) conditions and an N-benzyl alcohol matrix. TLC was performed on pre-coated plates of silica gel 60F, 0.20 mm thick (Merck). For detection, the plates were sprayed with 50% H<sub>2</sub>SO<sub>4</sub> (Orcinol) and heated for 3-5 minutes at 130 °C. Flash chromatography was performed on Aldrich silica gel 60 (Merck), 200-400 mesh. Yields refer to chromatographically and spectroscopically (<sup>1</sup>H NMR) homogeneous materials.

### Synthesis

#### **1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\alpha$ -L-fucoside (1).**

A mixture of L-Fucose 1.0 g and dry Dowex-50wx-8 (H<sup>+</sup>) resin (0.5 g) in 15 mL allyl alcohol was stirred and heated under reflux (bath temperature 110 °C) for 100 min. After cooling, the resin was filtered off, washed with anhydrous



ethanol (2 x 10 mL) and the combined filtrate was concentrated. The residue was co-evaporated with a 1:1 mixture of benzene and anhydrous ethanol (2 x 15 mL) to yield 1.2 g of syrup. Then, mercuric trifluoroacetate (3.32 g) in 50 mL acetonitrile was added drop wise in to allyl  $\alpha$ -D-fucoside (**12**) (0.81 g) in 50 mL acetonitrile. The reaction mixture was stirred overnight at room temperature after which 10 mL of 3M sodium hydroxide solution was added and stirring was continued for a further 10 minutes. Sodium borohydride (0.60 g) dissolved in 10 mL of sodium hydroxide solution was added and the reaction mixture was stirred for 3 h at room temperature. The suspension was then filtered through celite and the clear filtrate was concentrated under vacuum to remove the solvent. The residue thus obtained was purified by column chromatography (chloroform/methanol = 6:1). Yield: 0.41 g (50 %). M.p. 141-144 °C.  $^1\text{H}$  NMR (300 MHz,  $\text{D}_2\text{O}$ ):  $\delta$  0.92 (3H, d,  $J = 6.45$  Hz, Me), 1.05 (1H, d,  $J = 6.45$  Hz, Me), 3.36 (1H, dd,  $J = 12.30, 10.55$  Hz, H7ax), 3.66 (1H, dd,  $J = 3.52, 10.55$  Hz, H2), 3.73 (1H, d,  $J = 2.93$  Hz, H4), 3.81 (1H, dd,  $J = 2.93, 12.30$  Hz, H7e), 3.93 (1H, ddd,  $J = 6.45, 2.93, 10.54$  Hz, H8), 3.97 (1H, q,  $J = 6.45$  Hz, H5), 4.30 (1H, dd,  $J = 3.52, 10.55$  Hz, H3), 4.87 (1H, d,  $J = 3.52$  Hz, H1);  $^{13}\text{C}$  NMR (75 MHz,  $\text{D}_2\text{O}$ ):  $\delta$  94.5(C1), 72.73(C4), 71.74(C2), 71.69(C7), 69.55(C5), 64.30(C3), 63.67(C8), 15.77(C6), 15.30(C9); FTIR ( $\text{CHCl}_3$  cast) 3418, 2976, 2901, 1449, 1338, 1173, 1143, 1109, 1079, 999, 960, 839, 758  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{20} = -83.93^{\circ}$  (c 3.45,  $\text{CH}_3\text{OH}$ ); FAB-HRMS (NBA): calcd  $\text{C}_9\text{H}_{17}\text{O}_5$   $[\text{M}+\text{H}]^+$ , 205.1076, found 205.1086.

**4,6-O-Benzylidene-1,2-O-[2-methyl-(R)-1,2-ethanediy]- $\alpha$ -D-glucopyranoside (2).**

A mixture of D-glucose 6.0 g, allyl alcohol 25 mL) and dry Dowex-50wx-8 (H<sup>+</sup>) resin (2 g) was stirred and heated under reflux (bath temperature 110 °C) for 100 min. After cooling, the resin was filtered off, washed with anhydrous ethanol (2 x 10 mL) and the combined filtrate was concentrated. The residue was co-evaporated with a 1:1 mixture of benzene and anhydrous ethanol (2 x 15 mL). The residual allyl D-glucopyranoside was directly benzylidenated with 2.5 mL of benzaldehyde dimethyl acetal in 50 mL dimethyl formamide, by shaking for 15-20 h in the presence of catalytic amount of p-toluenesulfonic acid monohydrate. After addition of NaHCO<sub>3</sub> (0.31 g), the solvent was then evaporated and the residue was treated with ice water (15 mL) and extracted with ethyl acetate (3 X 15 mL), and the organic phase washed with water, dried over anhydrous sodium sulfate, and evaporated. Crystallization from hexane / ethyl acetate gave Allyl 4,6-O-benzylidene- $\alpha$ -D-glucopyranoside (2.11 g, 37% yield). Then mercuric trifluoroacetate (2.2 g) in 100 mL acetonitrile was added drop wise in to allyl 4,6-O-benzylidene- $\alpha$ -D-glucopyranoside (**14**) (0.8 g) in 150 mL acetonitrile. The reaction mixture was stirred at room temperature for 6 h after which 10 mL of 3M sodium hydroxide solution was added and stirring was continued for a further 10 minutes. Sodium borohydride (0.4 g) dissolved in 10 mL of sodium hydroxide solution was added and the reaction mixture was stirred for 3 h at room temperature. The suspension was then filtered through celite and the clear filtrate was concentrated under vacuum at 30 °C to remove the acetonitrile. The

residue thus obtained was purified by column chromatography (hexane/ethyl acetate = 1:2). Yield: 0.22 g (25 %). M.p. 217 - 220 °C. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 4.96 (H1, d, J = 3.6 Hz), 3.71 (H2, dd, J = 3.46, 9.22 Hz), 4.49 (H3, dd, J = 9.22, 9.57 Hz), 3.50 (H4, dd, J = 9.57, 9.44 Hz), 3.96 (H5, ddd, J = 4.72, 9.44, 10.30 Hz), 4.30 (H6eq, dd, J = 4.72, 10.43 Hz), 3.67 (H6ax, dd, J = 10.43, 10.30 Hz), 3.92 (H7eq, dd, J = 2.85, 11.73 Hz), 3.47 (H7ax, dd, J = 11.73, 10.51 Hz), 3.97 (H8, ddd, J = 2.85, 6.10, 10.51 Hz), 1.08 (H9, d, J = 6.10 Hz), 5.51 (H10, s), 7.41 (Ph); <sup>13</sup>CNMR (75 MHz, CDCl<sub>3</sub>): δ 128.1 and 125.4 (Ph), 101.2 (C10), 94.0 (C1), 80.2 (C4), 74.8 (C2), 71.4 (C7), 68.1 (C6), 65.1 (C3), 63.4 (C5), 62.7 (C8), 15.2 (C9); FTIR (CHCl<sub>3</sub> cast) 3411, 2969, 2928, 2863, 1453, 1372, 1173, 1164, 1148, 1130, 1117, 1087, 1076, 1028, 1017, 969, 943, 885, 822, 750, 698 cm<sup>-1</sup>; [α]<sub>D</sub><sup>20</sup> = +47.7° (c 0.595, CHCl<sub>3</sub>).

### **1,2-O-[2-methyl-(R)-1,2-ethanediyl]-β-L-fucoside (3).**

L-Fucose 2.0 g was added at once into acetyl bromide under stirring at 0°C. Soon after this, cold acetic acid (8 mL) was added to the suspension. During the next 10 min, all the sugar dissolved, with a fairly violent evolution of HBr. After the evolution of gas stopped, the ice-bath was removed and the reaction mixture was stirred for 1.5 hr at ambient temperature. The excess reagents were evaporated under reduced pressure at room temperature and then co-evaporated with toluene three times. The syrup thus obtained was treated with allyl alcohol (15 mL) and Ag<sub>2</sub>CO<sub>3</sub> (2.8 g). After passing through celite, it was then rotaevaporated. The residue was then deacetylated with 0.4 %

NaOMe in 100 mL methanol. It was then neutralized with Amberlite (H<sup>+</sup>) ion exchange resin. The resin was filtered off and the solution was evaporated under low pressure. The crude residue obtained was directly cyclized as follows. Mercuric trifluoroacetate (1.66 g) in 100 mL acetonitrile was added drop wise in to allyl  $\beta$ -L-fuccopyranoside (0.4 g) in 150 mL acetonitrile. The reaction mixture was stirred at room temperature for 6 h after which 10 mL of 3M sodium hydroxide solution was added and stirring was continued for a further 10 minutes. Sodium borohydride (0.3 g) dissolved in 10 mL of sodium hydroxide solution was added and the reaction mixture was stirred for 3 h at room temperature. The suspension was then filtered through celite and the clear filtrate was concentrated under vacuum at 30 °C to remove the acetonitrile. The residue thus obtained was purified by column chromatography (chloroform : methanol = 6:1). Yield: 0.16 g (40 %). M.p. 164 - 167 °C. <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O):  $\delta$  0.98 (3H, d, J = 6.45 Hz, Me-9), 1.12 (3H, d, J = 6.45 Hz, Me-6), 3.38 (1H, dd, J = 11.58, 10.80 Hz, H7ax), 3.25 (1H, dd, J = 7.62, 9.96 Hz, H2), 3.62 (1H, d, J = 3.51 Hz, H4), 3.83 (1H, dd, J = 2.35, 11.58 Hz, H7e), 3.73 (1H, ddd, J = 6.45, 2.35, 10.80 Hz, H8), 3.76 (1H, q, J = 6.45 Hz, H5), 3.61 (1H, dd, J = 3.51, 9.96 Hz, H3), 4.23 (1H, d, J = 7.62 Hz, H1); <sup>13</sup>CNMR (75 MHz, D<sub>2</sub>O):  $\delta$  98.2 (C1), 76.9 (C2), 72.8 (C5), 72.8 (C8), 71.8 (C3), 71.8 (C4), 71.6 (C7), 15.4(C6), 15.2(C9); FTIR (CHCl<sub>3</sub> cast) 3460, 2983, 2916, 1447, 1370, 1277, 1156, 1131, 1100, 1075, 1002 cm<sup>-1</sup>;  $[\alpha]_D^{20}$  = -39.3° (c 0.82, CH<sub>3</sub>OH); FAB-HRMS (NBA): calcd C<sub>9</sub>H<sub>17</sub>O<sub>5</sub> [M+H]<sup>+</sup>, 205.1076, found 205.1077.

### **3-O-Methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucoside (4).**

3-O-Methyl-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucoside (4) was prepared using a procedure similar to the one used to prepare 1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -L-fucoside (3) starting from 3-O-methyl-D-glucose. Yield: 0.15 g (26 %).  $^1\text{H NMR}$  (300 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  1.00 (3H, d,  $J = 6.41$  Hz, Me-9), 3.45 (3H, s, OMe-10), 3.07 (1H, dd,  $J = 7.83, 9.96$  Hz, H2), 3.26 (1H, dd,  $J = 8.00, 9.96$  Hz, H3), 3.34 (1H, d,  $J = 8.00, 9.61$  Hz, H4), 3.40 (1H, dd,  $J = 12.27, 10.85$  Hz, H7a), 3.60 (1H, dd,  $J = 4.98, 12.45$  Hz, H6), 3.70 (1H, ddd,  $J = 2.49, 6.41, 10.85$  Hz, H8), 3.71 (1H, dd,  $J = 9.61, 4.98$  Hz, H5), 3.75 (1H, d,  $J = 12.45, 9.96$  Hz, H6'), 3.85 (1H, dd, 2.49, 12.27 Hz, H7a), 4.34 (1H, d,  $J = 7.83$  Hz, H1);  $^{13}\text{CNMR}$  (75 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  97.5 (C1), 78.3 (C2), 82.3 (C3), 71.4 (C4), 77.7 (C5), 60.6 (C6), 71.4 (C7), 72.2 (C8), 15.2 (C9), 79.7 (OMe-10); FTIR ( $\text{CHCl}_3$  cast) 3421, 2916, 1456, 1376, 1278, 1151, 1121, 1068, 1034  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{20} = +33.7^\circ$  (c 0.72,  $\text{CH}_3\text{OH}$ ); FAB-HRMS (NBA): calcd  $\text{C}_{10}\text{H}_{19}\text{O}_6$   $[\text{M}+\text{H}]^+$ , 235.1182, found 235.1174.

### **4,6-O-Benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (5).**

D-Glucose (2.0 g) was added at once into acetyl bromide under stirring at  $0^\circ\text{C}$ . Soon after this, cold acetic acid (8 mL) was added to the suspension. During the next 10 min, all the sugar dissolved, with a fairly violent evolution of HBr. After the evolution of gas stopped, the ice-bath was removed and the reaction mixture was stirred for 1.5 hr at ambient temperature. The excess

reagents were evaporated under reduced pressure at room temperature and then co-evaporated with toluene three times. The syrup thus obtained was treated with allyl alcohol (15 mL) and  $\text{Ag}_2\text{CO}_3$  (2.8 g). After passing through celite, it was then rotaevaporated. The residue was then deacetylated with 0.4 % NaOMe in 100 mL methanol. It was then neutralized with Amberlite ( $\text{H}^+$ ) ion exchange resin. The resin was filtered off and the solution was evaporated under low pressure. The residue obtained was directly benzylidenated as follows. The residual allyl  $\beta$ -D-glucopyranoside was directly benzylidenated with 0.83 mL of benzaldehyde dimethyl acetal in 50 mL dimethyl formamide, by shaking for 15-20 h in the presence of catalytic amount of p-toluenesulfonic acid monohydrate. After addition of  $\text{NaHCO}_3$  (0.3 g), the solvent was evaporated, the residue treated with ice-water (15 mL) and extracted with ethyl acetate (3 X 15 mL), and the organic phase washed with water, dried over anhydrous sodium sulfate, and evaporated. Then, mercuric trifluoroacetate (2.2 g) in 100 mL acetonitrile was added drop wise in to allyl 4,6-O-benzylidene- $\beta$ -D-glucopyranoside (**19**) (0.8 g) in 150 mL acetonitrile. The reaction mixture was stirred at room temperature for 6 h after which 10 mL of 3M sodium hydroxide solution was added and stirring was continued for a further 10 minutes. Sodium borohydride (0.4 g) dissolved in 10 mL of sodium hydroxide solution was added and the reaction mixture was stirred for 3 h at room temperature. The suspension was then filtered through celite and the clear filtrate was concentrated under vacuum at 30 °C to remove the acetonitrile. The residue was then purified by column chromatography (hexane/ethyl acetate = 1:2). Yield:

0.25 g (31 %). M.p. 188 - 190 °C.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.16 (3H, d,  $J = 6.23$  Hz, Me-9), 3.30 (1H, dd,  $J = 7.8, 9.52$  Hz, H2), 3.49 (1H, dd,  $J = 10.72, 11.72$  Hz, H7ax), 3.85 (1H, dd,  $J = 6.22, 2.93, 12.8$  Hz, H8), 3.82 (1H, dd,  $J = 9.89, 10.25$  Hz, H3), 3.89 (1H, 8.42, 8.42, H5), 3.94 (1H, dd,  $J = 11.72, 2.93$  Hz, H7e), 4.20 (1H, d 11.7 Hz, H6), 3.61 (1H, dd,  $J = 10.25, 8.42$  Hz, H4), 4.39 (1H, d,  $J = 7.69$  Hz, H1), 5.54 (1H, s, H10), 7.40 (5H, m, Ph);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  129.4, 128.4 (Ph), 102.38 (C10), 98.87 (C1), 81.54 (C2), 71.03 (C5), 68.68 (C3), 68.31 (C4), 72.27 (C7), 71.87 (C8), 68.65 (C6), 15.2 (C9); FTIR ( $\text{CHCl}_3$  cast) 3459, 2975, 2871, 1452, 1370, 1314, 1277, 1146, 1120, 1094, 1066, 1055, 1029, 1006, 971, 844, 749, 699  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{20} = +10.0^\circ$  (c 1.22,  $\text{CHCl}_3$ ).

**4,6-O-Benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (6).**

A procedure similar to the one we used to prepare 4,6-O-Benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-glucopyranoside (5) was used to prepare 4,6-O-Benzylidene-1,2-O-[2-methyl-(S)-1,2-ethanediyl]- $\beta$ -D-galactopyranoside (6) starting from D-galactose. Yield: 0.24 g (30 %).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  4.30 (H1, d,  $J = 7.69$  Hz), 3.60 (H2, dd,  $J = 7.69, 9.89$  Hz), 3.79 (H3, dd,  $J = 9.89, 3.84$  Hz), 4.25 (H4, dd,  $J = 3.84, 0.91$  Hz), 3.61 (H5, d,  $J = 1.46$  Hz), 4.39 (H6, dd,  $J = 1.46, 12.67$  Hz), 4.08 (H6', d,  $J = 12.67$  Hz), 3.92 (H7eq, dd,  $J = 2.75, 11.89$  Hz), 3.47 (H7ax, dd,  $J = 11.89, 10.25$  Hz), 3.85 (H8, ddd,  $J = 2.75, 6.41, 10.25$  Hz), 1.15 (H9, d,  $J = 6.41$  Hz), 5.56 (H10, s), 7.37 (Ph);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  129.4, 127.3, 126.6 (Ph), 101.7 (C10), 98.7 (C1), 77.3 (C2),

70.6 (C3), 75.9 (C4), 68.3 (C5), 69.1 (C6), 72.1 (C7), 71.7 (C8), 16.3 (C9); FTIR (CHCl<sub>3</sub> cast) 3463, 2973, 2871, 1713, 1452, 1401, 1370, 1331, 1315, 1288, 1250, 1218, 1173, 1159, 1142, 1124, 1101, 1079, 1060, 1024, 1007, 973, 919, 889 cm<sup>-1</sup>; [ $\alpha$ ]<sub>D</sub><sup>20</sup> = +52.4° (c 0.55, CHCl<sub>3</sub>); FAB-HRMS (NBA): calcd C<sub>16</sub>H<sub>21</sub>O<sub>6</sub> [M+H]<sup>+</sup>, 309.1338, found 309.1349.

### **3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (7).**

Methyl  $\alpha$ -D-mannopyranoside glucopyranoside 4.0 g was directly benzylidened with 2.5 mL of benzaldehyde dimethyl acetal in 50 mL dimethyl formamide, by shaking for 15-20 h in the presence of catalytic amount of *p*-toluene sulfonic acid monohydrate. After addition of NaHCO<sub>3</sub> (0.5 g), the solvent was evaporated, the residue treated with ice-water (17 mL) and extracted with ethyl acetate (3 X 20 mL), and the organic phase washed with water, dried over anhydrous sodium sulfate, and evaporated. Then, a mixture of Methyl 4,6-O-benzylidene- $\alpha$ -D-mannopyranoside (2.5 g) and dibutyl tin oxide (2.2 g) in 100 mL methanol was refluxed for 1 h at 60 °C. The solvent was then rotaevaporated and the residue dissolved in 125 mL dimethyl formamide. To this solution 1.8 mL of benzyl bromide was added and was stirred at 100 °C for 30 min. It was then concentrated and purified by flash chromatography with ethyl acetate-hexane (3:2) to give methyl 3-O-benzyl-4,6-O-benzylidene- $\alpha$ -D-mannopyranoside (**24**). Then, to a solution of 0.64 g of methyl 3-O-benzyl-4,6-O-benzylidene- $\alpha$ -D-mannopyranoside (**24**) in 30.0 mL of tetrahydrofuran was added 0.41 g of sodium hydride, 22.9 mg of imidazole and 22.6 mg of tetrabutyl ammonium



bromide. It was then refluxed for 0.5 h and 1.6 g of epichlorohydrin was added and refluxed overnight. To this mixture 20.0 mL of methanol was added drop wise followed by 42.0 mL of dilute aqueous HCl (0.12 % v/v). It was then extracted with methylene chloride (3 x 60 mL), the organic layer washed with dilute solutions of sodium carbonate and brine and dried over anhydrous sodium sulfate and the solvent evaporated under reduced pressure. And the residue was purified by flash chromatography (hexane-ethyl acetate 6:1) to give 0.54 g product (57% yield). To a solution of this epoxide (0.54 g) in 75 mL of tetrahydrofuran at 0°C under N<sub>2</sub> was added 0.46 mg of Lithium Aluminum Hydride. The mixture was stirred at 0°C for 1 h and at room temperature for 5 h. The reaction was quenched by the addition of saturated aqueous solution of sodium sulfate until a white, granular precipitate was observed. The THF solution was then decanted, and the salts were washed with additional THF. The combined organic solution was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give 0.54 g (99% for the reduction step) of 1-O-methyl-2-(2-hydroxypropyl)-3-O-benzyl-4,6-O-benzylidene- $\alpha$ -D-mannopyranoside. Hydrolysis of methyl 2-(2-hydroxy propyl)-3-O-benzyl-4,6-O-benzylidene- $\alpha$ -D-mannopyranoside was performed by refluxing overnight in 0.5 M H<sub>2</sub>SO<sub>4</sub> and then purification by flash chromatography (methanol/chloroform 1/6) gave 0.1 g (22% yield) of 3-O-benzyl-1,2-O-[2-methyl-(R)-1,2-ethanediyl]- $\beta$ -D-mannopyranoside (7). <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O):  $\delta$  1.03 (3H, d, J = 6.59 Hz, Me-9), 3.34 (1H, ddd, J = 2.20,5.37,9.77 Hz, H5), 3.27 (1H, dd, J = 10.74,11.96 Hz, H8ax), 3.70 (1H, dd, J = 9.76,9.77 Hz, H4), 3.87 (1H, dd, J = 2.20,12.20 Hz,

H6), 3.81 (1H, dd, J = 2.92,11.96 Hz, H8e), 3.57 (1H, dd, J = 3.66,9.76 Hz, H3), 3.74 (1H, dd, J = 5.37,12.20 Hz, H6'), 4.02 (1H, dd, J = 3.66, 1.46 Hz, H2), 4.27 (1H, ddd, J = 10.74, 2.93, 6.59 Hz, H7), 4.78 (1H, d, J = 1.46 Hz, H1), 4.54 (2H, s, CH<sub>2</sub>-10), 7.34 (5H, m, Ph); <sup>13</sup>CNMR (75 MHz, D<sub>2</sub>O): δ 128.9, 137.0 (Ph), 93.5 (C1), 71.1 (C2), 76.3 (C5), 79.2 (C3), 71.2 (C7), 66.0 (C4), 66.0 (C8), 73.0 (C10), 62.0 (C6), 15.4 (C9); FTIR (CHCl<sub>3</sub> cast) 3419, 2871, 1496, 1454, 1367, 1274, 1208, 1173, 1071, 860, 741, 699 cm<sup>-1</sup>; [α]<sub>D</sub><sup>20</sup> = -18.4° (c 1.23, CH<sub>3</sub>OH); FAB-HRMS (NBA): calcd C<sub>16</sub>H<sub>23</sub>O<sub>6</sub> [M+H]<sup>+</sup>, 311.1495, found 311.1496.

## Acknowledgements

This work was supported by the Michigan State University Research Excellence Fund.

## References

1. Tvaroska, I; Hricovini, M; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359.
2. Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988**, *184*, 39.
3. Huang, G.; Hollingsworth, R. I. *Tetrahedron* **1998**, *54*, 1355.
4. Pelyvás, I.; Lindhorst, T.; Thiem, J. *Liebigs Ann. Chem.* **1990**, 761.
5. Koto, S.; Yoshida, T.; Takenaka, K.; Zen, S. *Bull. Chem. Soc. Jpn.* **1982**, *55*, 3667.
6. Koto, S.; Morishima, N.; Takenaka, K.; Kanemitsu, K.; Shimoura, N.; Kase, M.; Kojiro, S.; Nakamura, T.; Kawase, T.; Zen, S. *Bull. Chem. Soc. Jpn.* **1989**, *62*, 3549.
7. Nashed, M. A.; Anderson, L. *Tetrahedron Lett.* **1976**, *39*, 3503.
8. Loughhead, D. G. *J. Org. Chem.* **1985**, *50*, 3931.
9. Tomaszewski, M. J.; Holme, K. R.; Perlin, A. S. *Carbohydr. Res.* **1991**, *217*, 237.

## **Chapter 4**

### **New Karplus-type Equations Relating the Glycosidic Dihedral Angles with Vicinal Proton-Carbon Coupling Constants.**

## Abstract

We report here new Karplus-Type equations relating the vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , to the glycosidic dihedral angles  $\phi$  (H1'-C1'-OX-CX) and  $\psi$  (HX-CX-OX-C1'). The three dimensional structures of the synthetic model compounds 1-7 and the commercially available compounds 8-10 have been studied by Molecular Mechanics and AM1 and PM3 Semi-empirical calculations. The vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , have been measured in solution by two-dimensional excitation-sculptured indirect-detection experiment (EXSIDE) NMR spectroscopy. From the least squares fitting of the measured coupling constants and the glycosidic dihedral angles of the rigid molecules, equations of the form:  $^3J_{\text{HC}}(\theta) = a \text{Cos}^2(\theta) + b \text{Cos}(\theta) + c$ , were obtained where  $\theta$  stands for the dihedral angles defined by the four atoms H-C-O-C ( $\phi$  and  $\psi$  for carbohydrates).

## Introduction.

The equation  ${}^3J_{\text{HC}}(\theta) = 5.7 \cos^2(\theta) - 0.6 \cos(\theta) + 0.5$  (**equation 1**), where  $\theta$  stands for the dihedral angle defined by the atoms H-C-O-C, has been proposed by Tvaroska *et. al.*<sup>1</sup> and a similar equation  ${}^3J_{\text{HC}}(\theta) = 5.5 \cos^2(\theta) - 0.7 \cos(\theta) + 0.6$  (**equation 2**) has also been proposed by Mulloy *et. al.*<sup>2</sup> The above equations have been used to relate the vicinal proton-carbon heteronuclear coupling constants across the glycosidic bonds of carbohydrates to the glycosidic dihedral angles. However, more often than not, the structure of the carbohydrates predicted by using the above equations is different from the structures predicted by other methods.<sup>3</sup> We believe that the reason for the discrepancy is the following:

i) The glycosidic dihedral angles of carbohydrates defined by the atoms H-C-O-C is, to our knowledge, almost always in the range 0 to 80 degrees. This is based on analysis of the crystal structure of carbohydrates from Cambridge Structural Database. This is discussed in detail in chapter two.

ii) Only 3 data points are used to parameterize the above **equation 1** for the region 0 to 85 degrees. The remaining 14 data points are in the range of 86 to 180 degrees. Even those 3 data points, as pointed out by authors, are “only estimated values and therefore the precision of these data is lower”.<sup>1</sup> The accuracy of the values of the three through-bond coupling constants is also not given. And in the case of **equation 2** only 6 of the 18 data points are in the range of 0 to 90 degrees and the remaining 12 are in the range of 90 to 180 degrees. In addition, equations 1 and 2 are parameterized using coupling constants

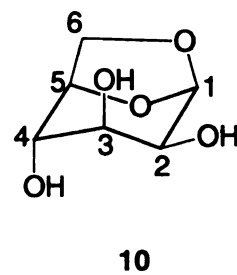
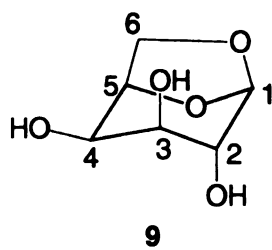
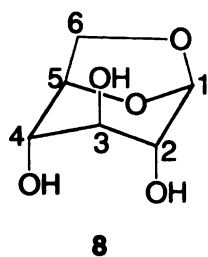
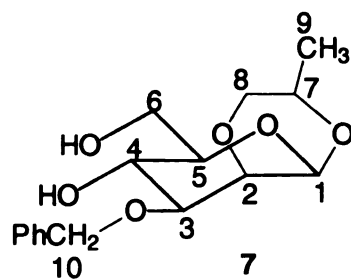
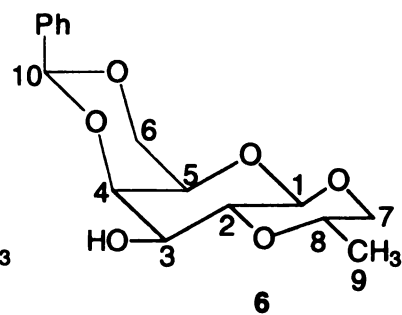
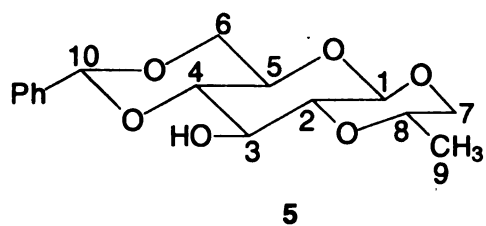
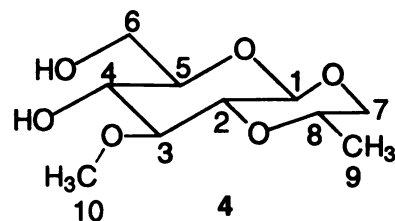
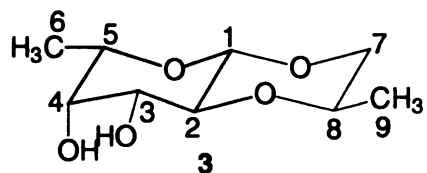
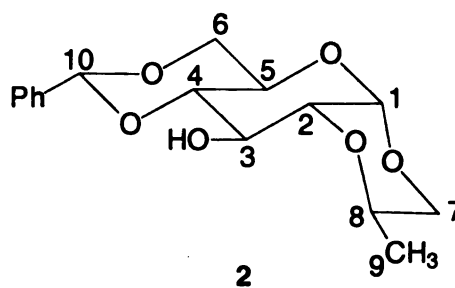
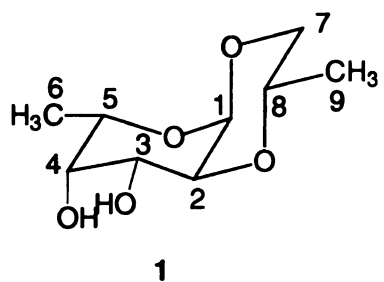
measured in solution and dihedral angles from crystal structures. But the solution conformations of their model compounds and carbohydrates, in general, may not be identical with the crystal structures.

iii) The electronic environments (electronegativity and anomeric effect, among others) are different for the two glycosidic dihedral angles,  $\phi$  (H1'-C1'-OX-CX) and  $\psi$  (HX-CX-OX-C1') and hence the magnitudes of the through-bond coupling constants for a given dihedral angle may be different. And this difference might require two equations to describe the dependence of the three bond proton-carbon heteronuclear coupling constants of carbohydrates across the glycosidic bonds to the glycosidic dihedral angles, one for each angle  $\phi$  (H1'-C1'-OX-CX) and  $\psi$  (HX-CX-OX-C1').

More recently Cloran *et al.* reported a new Karplus type equation,  $^3J_{HC}(\theta) = 7.49 \cos^2(\theta) - 0.96 \cos(\theta) + 0.15$  (**equation 3**), where  $\theta$  stands for the dihedral angle defined by the atoms H-C-O-C, relating the vicinal proton-carbon coupling constants to the glycosidic dihedral angles from pure theoretical calculations of the coupling constants and geometries of their model compounds.<sup>4,5</sup> This equation shows a big difference with the previous two equations of Tvaroska *et al.*<sup>1</sup> and Mulloy *et al.*<sup>2</sup> at lower and higher dihedral angles and factors such as solvation, basis set limitations, the small set of geometries and structures studied are suggested as possible factors for the deviations from the above mentioned equations. The equation of Haasnoot *et al.* is a well-parameterized Karplus-type equation for  $\omega$  dihedral angle of

**Figure 4.1** The ten molecules that have rigid glycosidic dihedral angles. While molecules **8-10** are commercially available, molecules **1-7** are synthesized in this work (Chapter 3).





carbohydrates relating the dihedral angles defined by H-C-C-H with the proton-proton coupling constants for 1->6 linked sugars.<sup>6</sup>

Because of the reasons discussed above we decided to synthesize molecules with rigid glycosidic linkages, measure their vicinal proton-carbon coupling constants in solution, refine their dihedral angles by molecular mechanics and semi-empirical (AM1 and PM3) calculations, and least squares fit the dihedral angles and coupling constants to come out with new Karplus-type equations. The synthesis of the model compounds is presented in Chapter Three. The structural studies and the parameterization of the new Karplus-type equation are reported in this Chapter.

## **Material and Methods.**

### **Model compounds.**

The model compounds 1-7 were synthesized in this work and their synthesis and characterization is described in Chapter Three. And compounds 8-10 were commercially available and were used as is. Seven of the model compounds are bicyclic and the remaining three are tricyclic. The model compounds contained  $\alpha$  and  $\beta$  sugars, D and L sugars, substituted and unsubstituted sugars and the most common monosaccharide units in nature, namely, glucose, galactose, mannose, and fucose.

## **Nuclear Magnetic Resonance (NMR) Spectroscopy.**

Nuclear Magnetic Resonance spectra of compounds **1, 3, 4** and **7-10** were obtained in deuterated water, while that of **2, 5, and 6** were obtained in deuterated chloroform. Spectra are recorded at 30 °C. The residual solvent peaks were used as reference. Combinations of homonuclear- and heteronuclear- (one- and) two-dimensional NMR experiments were performed in order to allow the assignments of the signals for the model compounds and to measure the coupling constants. This included: double quantum filtered J-correlated spectroscopy (DQF-COSY) spectra (Ernst *et al.*, 1987),<sup>7</sup> total correlated spectroscopy (TOCSY) spectra (Bax & Davis, 1985),<sup>8</sup> heteronuclear multi-quantum coherence (HMQC) experiment (Bax & Subramanian, 1986),<sup>9</sup> heteronuclear multi-bond coherence (HMBC) experiment (Titman *et al.*, 1989),<sup>10</sup> and nuclear Overhauser effect (NOESY) experiment (Sangers *et al.*, 1993).<sup>11</sup>

Analysis of the spin systems for the model compounds from the 1D NMR spectra and from DQF-COSY and TOCSY spectra gave the connectivities. Once the assignments are made NOESY experiment was used to determine the various inter-nuclear distances. Spectra of the NOESY experiments were acquired at different mixing times to address the potential problem of spin diffusion.

The vicinal proton-carbon coupling constants were obtained from excitation-sculptured indirect-detection (EXSIDE) experiments (Krishnamurthy *et al.*, 1996)<sup>12</sup>. Excitation bandwidth selection, also known as excitation-sculpting, was carried out with the Varian software-internal package using a Gaussian

cascade Q3 pulse of 14 ms duration for selective inversion of the selected protons in compounds 1-10. A J-scaling factor of 15 was used and 16 scans of 2048 complex data points were collected for  $512t_1$  increments. EXSIDE provides sensitivity comparable to a nonselective gradient HMBC experiment, but with pure absorptive line shapes. Long-range proton-carbon coupling constants were J-scaled and measured along the  $F_1$  (carbon) dimension. In the EXSIDE spectrum, the active coupling constant is measured with no interference from any passive couplings, in contrast to other methods where the proton-carbon coupling constants are measured along the  $F_2$  (proton) dimension with overlapping homonuclear couplings. But the resonances of interest should be well resolved from any homonuclear-coupled spin partner. The EXSIDE experiment is especially very useful for carbohydrates since the anomeric protons are further away from the rest of the bulk proton signals and hence can be selectively excited. The error in measured coupling constants was  $\pm 0.1$  Hz. All of the two dimensional NMR spectra were measured on Varian Inova 600 MHz NMR spectrometer operating at 600 MHz for protons and 150 MHz for carbons. The tables of assignment of the proton and carbon chemical shifts and coupling constants of compounds 8 – 10 is given in **Tables 4.1 – 4.3**, while compounds 1 – 7 were fully characterized and their proton and carbon chemical shifts and coupling constants assigned in Chapter Three.

**Table 4.1.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments and proton-proton coupling constants of 1,6-anhydroglucopyranose (**8**).

1,6-anhydro glucose ( <b>8</b> )	$^1\text{H}$ ( $\delta$ , ppm)	$^3J_{\text{HH}}$ (Hz)	$^{13}\text{C}$ ( $\delta$ , ppm)
1	5.40	1.17 ( $^3J_{\text{H1H2}}$ )	101.7
2	3.47	---	70.5
3	3.63	---	72.7
4	3.62	---	71.1
5	4.58	5.86 ( $^3J_{\text{H5H6'}}$ )	76.5
6	4.03	7.62 ( $^3J_{\text{H6H6'}}$ )	65.4
6'	3.71	5.86, 7.62 ( $^3J_{\text{H5H6'}}$ , $^3J_{\text{H6H6'}}$ )	65.4

**Table 4.2.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments and proton-proton coupling constants of 1,6-anhydrogalactopyranose (**9**).

1,6-anhydro galactose ( <b>9</b> )	$^1\text{H}$ ( $\delta$ , ppm)	$^3\text{J}_{\text{HH}}$ (Hz)	$^{13}\text{C}$ ( $\delta$ , ppm)
1	5.37	1.46 ( $^3\text{J}_{\text{H1H2}}$ )	100.9
2	3.90	4.89 ( $^3\text{J}_{\text{H2H3}}$ )	73.4
3	3.76	1.71 ( $^3\text{J}_{\text{H3H4}}$ )	71.5
4	4.01	4.89 ( $^3\text{J}_{\text{H4H5}}$ )	64.5
5	4.46	4.64, 4.88 ( $^3\text{J}_{\text{H4H5}}$ , $^3\text{J}_{\text{H5H6'}}$ )	74.5
6	4.28	7.81 ( $^3\text{J}_{\text{H6H6'}}$ )	63.7
6'	3.61	5.12, 7.81 ( $^3\text{J}_{\text{H5H6'}}$ , $^3\text{J}_{\text{H6H6'}}$ )	63.7

**Table 4.3.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments and proton-proton coupling constants of 1,6-anhydromannopyranose (10).

1,6-anhydro mannose (10)	$^1\text{H}$ ( $\delta$ , ppm)	$^3J_{\text{HH}}$ (Hz)	$^{13}\text{C}$ ( $\delta$ , ppm)
1	5.28	---	101.3
2	3.60	4.88 ( $^3J_{\text{H2H3}}$ )	66.0
3	3.88	1.83 ( $^3J_{\text{H3H4}}$ )	71.6
4	3.84	1.46 ( $^3J_{\text{H4H5}}$ )	70.3
5	4.48	5.86 ( $^3J_{\text{H5H6'}}$ )	75.8
6	4.10	7.57 ( $^3J_{\text{H6H6'}}$ )	64.8
6'	3.62	5.86, 7.57 ( $^3J_{\text{H5H6'}}$ , $^3J_{\text{H6H6'}}$ )	64.8

## **Molecular Mechanics and Semi-empirical Calculations**

Molecular mechanics calculations were performed using the Alchemy 2000 program (Tripos, Inc. St. Louis, MO, USA) using the MM3 force fields. Minimizations are performed using the conjugate gradient method.

Semi-empirical calculations were performed using the Alchemy 2000 program (Tripos, Inc. St. Louis, MO, USA). The geometries of the molecules were optimized using the AM1<sup>13</sup> and PM3<sup>14</sup> Hamiltonians in the Alchemy 2000 program and the geometries were re-optimized after they were perturbed. **Figures 4.2-4.5** are color presentations for clarity.

## **Results and Discussion**

We observed that sucrose has a dihedral angle of 8 degrees in the crystal structure in the Cambridge Structural Database. Also its 2-hydroxyl group and the C1 hydroxymethyl group are perfectly situated to form an eight membered ring by acetal formation. Based on this information we prepared the sucrose acetal hoping that it will serve as a model for a dihedral angle value of about zero degrees. But the measured three-bond proton-carbon coupling constant, 2.07 Hz, is actually consistent with a W structure of the cyclooctane ring with a dihedral angle value of around 60 degrees. As a result we parameterized the Karplus-type equation based on our studies of compounds 1-10 which does not include dihedral angles around 0 degrees.



Tvaroska *et. al.* used 17 data points to parameterize the Karplus-type equation  ${}^3J_{\text{HC}}(\theta) = 5.7 \text{ Cos}^2(\theta) - 0.6 \text{ Cos}(\theta) + 0.5$ , where  $\theta$  stands for the dihedral angle defined by the atoms H-C-O-C. Of the 17 data points, 14 correspond to magnitudes of  $\theta$  between 86 and 180 degrees and of the remaining 3, 2 are for values of  $\theta$  being 60 degrees with  ${}^3J_{\text{HC}}$  of 2.2 and 2.6 Hz and the remaining 1 is for a value of  $\theta$  being 10 degrees with  ${}^3J_{\text{HC}}$  of 5.2 Hz. In addition, the 3 data points that were used to parameterize their Karplus-type equation for the range of 0 to 86 degrees are 'only estimated values and therefore the precision these is lower'.<sup>1</sup> The studies of Mulloy *et. al.* has 18 data points, 12 of which are above 105 degrees. Furthermore, both equations 1 and 2 were obtained by fitting crystal geometries with experimentally measured coupling constants in solution. But the solution conformations of their model compounds and carbohydrates, in general, may not be identical with the crystal structures.

Our studies (Chapter two) revealed that the values of the critical dihedral angles of carbohydrates that determine their three dimensional structure are almost always in the range of 0 to 85 degrees. It became apparent then that the equations of Tvaroska *et. al.* and Mulloy *et. al.* are poorly parameterized for carbohydrates.

We used 28 data points out of which 13 are between 0 and 90 degrees and 15 are between 90 and 180 degrees. For the range of  $\theta$  between 100 and 180 degrees our data points (Tables 4.4 and 4.5) lay close to the Karplus-type curves of Tvaroska *et. al.* and Mulloy *et. al.* (Figure 4.2) This is not surprising

**Table 4.4.** Heteronuclear three-bond proton-carbon coupling constants of compounds **1-5** measured from 2D EXSIDE NMR experiments and the corresponding dihedral angles (torsions). The compound number is given in bracket.

Dihedral	$^3J_{CH}$	Torsion (MM calculation)	Torsion (AM1 calculation)	Torsion (PM3 calculation)
H1C1O5C5 (1)	6.837	-175.8	173.1	168.3
H1C1O7C7 (1)	1.211	-59.7	-72.1	-77.7
H7aC7O7C1 (1)	1.469	59.3	68.1	71.7
H5C5O5C1 (1)	1.95	55.5	58.2	63.1
H1C1O1C7 (2)	2.273	59.3	62.5	71.1
H7axC7O1C1 (2)	1.667	59.5	-59.8	-75.5
H1C1O5C5 (2)	6.671	176.7	-174.1	-163.1
H7aC7O7C1 (3)	1.225	-61.0	-68.6	-73.8
H7eC7O7C1 (3)	6.813	179.4	173.7	171.4
H5C5O5C1 (3)	1.711	58.8	58.6	61.4
H1C1O7C7 (4)	2.197	-60.2	-63.1	-66.1
H7aC7O7C1 (4)	1.211	60.6	67.5	74.0
H7eC7O7C1 (4)	7.015	-179.6	-174.8	-171.2
H1C1O5C5 (5)	1.947	60.1	62.4	63.6
H7eC7O7C1 (5)	7.203	-179.1	-173.8	-171.2
H5C5O5C1 (5)	1.953	-58.1	-59.6	-60.5

**Table 4.5.** Heteronuclear three-bond proton-carbon coupling constants of compounds **6-10** measured from 2D EXSIDE NMR experiments and the corresponding dihedral angles (torsions).

Dihedral	$^3J_{CH}$	Torsion (MM calculation)	Torsion (AM1 calculation)	Torsion (PM3 calculation)
H7aC7O7C1 ( <b>6</b> )	1.225	61.3	69.9	74.0
H5C5O5C1 ( <b>6</b> )	1.095	-60.6	-70.1	-70.2
H1C1O7C7 ( <b>7</b> )	6.587	178.8	-174.7	-166.2
H1C1O6C6 ( <b>8</b> )	4.883	138.3	149.7	148.5
H6C6O6C1 ( <b>8</b> )	3.175	-107.5	-130.3	-130.0
H5C5O5C1 ( <b>8</b> )	5.878	164.0	159.0	160.1
H1C1O6C6 ( <b>9</b> )	4.871	138.9	147.6	147.1
H6C6O6C1 ( <b>9</b> )	2.692	-107.5	-126.5	-127.0
H5C5O5C1 ( <b>9</b> )	5.867	164.4	159.6	160.6
H1C1O6C6 ( <b>10</b> )	5.398	137.6	148.6	147.2
H6C6O6C1 ( <b>10</b> )	3.170	-106.2	-127.9	-127.1
H5C5O5C1 ( <b>10</b> )	5.605	164.3	159.2	160.7

because their equations are actually reasonably well optimized in this range of dihedral angles because the higher percentage (14 of 17 in Tvaroska *et. al.* and 12 of 18 in Mulloy *et. al.*) of their data points belong to this range of dihedral angles. The fact that our data points (AM1 and PM3 optimized geometries with the experimentally measured heteronuclear coupling constants) are in good agreement with the predictions by the equations of Tvaroska *et. al.* and Mulloy *et. al.* for the range of dihedral values between 100 and 180 degrees gave us confidence that the geometries obtained by the semi-empirical calculations are reasonably reliable.

In the important range of dihedral angles for carbohydrates, 0 to 85 degrees, our data points are significantly different from the Karplus-type curves of Tvaroska *et. al.* and Mulloy *et. al.* (**Figure 4. 2**). And indeed, the least square fit of our data points (e.g. **equation 4**:  $^3J_{\text{HC}}(\theta) = 5.597 \text{ Cos}^2(\theta) - 0.421 \text{ Cos}(\theta) + 0.722$ ) gives a different prediction with better  $R^2$  values (**Table 4.10**).

Molecular mechanics and PM3 and AM1 semi-empirical calculations were used to refine the structures of the already constrained bi- and tri-cyclic model compounds (**1 – 10**). Molecular mechanics optimization of the structures of compounds **1 – 10** revealed that the dihedral angles obtained from this studies are far way from the Karplus-type curves of Tvaroska *et. al.* and Mulloy *et. al.* (**Figure 4.2**). Especially for the dihedral angles around 60 degrees, the dihedral angles from the MM calculations are between 59 and 61 degrees for all molecules even though the values of the vicinal proton-carbon heteronuclear

coupling constants are distributed between 2.27 and 1.09 Hz (**Table 4.4** and **4.5** and **Figure 4.2**). In contrast the corresponding dihedral angles of the molecules obtained from geometry optimizations by PM3 and AM1 semi-empirical calculations are distributed between 58 and 78 degrees, which is consistent with the distribution of the vicinal proton-carbon heteronuclear coupling constants between 2.27 and 1.09 Hz (**Tables 4.4** and **4.5** and **Figures 4.3** and **4.4**). Indeed, the least squares fitting of the dihedral angles refined by the molecular mechanics calculations to the experimentally measured vicinal proton-carbon heteronuclear coupling constants gave a much lower  $R^2$  value than that of the dihedral angles refined by semi-empirical calculations and the measured coupling constants (**Table 4.10**). The least squares fitting of the dihedral angles refined by semi-empirical geometry optimizations with the experimentally measured vicinal proton-carbon heteronuclear coupling constants gave the highest  $R^2$  value. **Table 4.10** gives the coefficients of the Karplus-type equation from least squares fitting. For each case (e.g. AM1), three equations were obtained. These were for the  $\phi$  dihedral angle alone (AM1- $\phi$ ), for the  $\psi$  dihedral angle alone (AM1- $\psi$ ), and for both  $\phi$  and  $\psi$  angles (AM1- $\phi\psi$ ). It was clear that the equation for the  $\phi$  dihedral angle alone (e.g. AM1- $\phi$ ) predicts a higher value of the coupling constant than that for the  $\psi$  dihedral angle alone (AM1- $\psi$ ). This confirms our expectation that different equations are required for the  $\phi$  and  $\psi$  dihedral angles. On the other hand, dividing our database into two reduces the

**Table 4.6.** Structures of 1,6-anhydroglucose (**8**) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations. The relative energies are from PM3 single point energy calculation of the crystal structures and the optimized structures using MM, AM1 or PM3 geometry optimizations. Ref code stands for the reference codes given in the Cambridge Structural Database.

Compound	H1C1O6C6	H5C5O5C1	H6pC6O6C1	Ref code	Relative E (kcal/mol)
1,6-Glcanh	140.90	164.03	-115.90	Ahglpy01	9.61
<b>(8)</b>	144.42	161.14	-106.67	Ahglpy10	2.39
MM	147.2	162.9	-118.8	Ahglpy10	0.78
AM1	144.7	162.9	-122.2	Ahglpy10	0.17
PM3	146.8	160.9	-127.2	Ahglpy10	0.0
MM	146.6	163.2	-117.5	Ahglpy01	0.81
AM1	146.5	161.8	-123.5	Ahglpy01	0.18
PM3	146.8	160.9	-127.1	Ahglpy01	0.0

**Table 4.7.** Structures of 1,6-anhydrogalactose (**9**) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations. The relative energies are from PM3 single point energy calculation of the crystal structures and the optimized structures using MM, AM1 or PM3 geometry optimizations. Ref code stands for the reference codes given in the Cambridge Structural Database.

Compound	H1C1O6C6	H5C5O5C1	H6pC6O6C1	Refcode	Relative E (kcal/mol)
1,6- Galanh ( <b>9</b> )	120.13	-173.36	-88.29	Ahgalp3	4.26
	133.68	156.92	-110.09	Ahgalp2	7.75
	122.86	163.13	-108.84	Ahgalp1	5.39
MM	149.2	162.1	-122.5	Ahgalp1	0.84
AM1	143.9	161.7	-120.7	Ahgalp1	0.16
PM3	148.4	160.5	-129.4	Ahgalp1	0.0
MM	148.3	162.4	-120.8	Ahgalp2	0.83
AM1	150.6	158.7	-130.4	Ahgalp2	0.27
PM3	147.2	161.3	-127.2	Ahgalp2	0.18
MM	148.3	162.6	-120.7	Ahgalp3	0.86
AM1	143.9	161.8	-120.6	Ahgalp3	0.15
PM3	148.1	160.7	-129.3	Ahgalp3	0.0

**Table 4.8.** Structures of 1,6-anhydromannose (**10**) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations. The relative energies are from PM3 single point energy calculation of the crystal structures and the optimized structures using MM, AM1 or PM3 geometry optimizations. Ref code stands for the reference codes given in the Cambridge Structural Database.

Compound	H1C1O6C6	H5C5O5C1	H6pC6O6C1	Ref code	Relative E (kcal/mol)
1,6-Mananh ( <b>10</b> )	142.51	165.68	-118.35	Biftuf1	6.91
	144.68	161.17	-121.26	Biftuf2	6.65
MM	147.0	162.9	-117.7	Biftuf1	0.86
AM1	146.7	160.5	-124.6	Biftuf1	0.18
PM3	146.9	161.0	-126.6	Biftuf1	0.0
MM	147.2	162.3	-118.6	Biftuf2	0.71
AM1	146.4	161.0	-124.2	Biftuf2	0.19
PM3	145.9	161.8	-124.8	Biftuf2	0.1



**Table 4.9.** The values of the dihedral angle C2-C3-C4-C5 as a measure of the degree of flattening of the chair conformations of the three 1,6-anhydro sugars, 1,6-anhydroglucose (8), 1,6-anhydrogalactose (9), and 1,6-anhydromannose (10) from Cambridge Structural Database and from MM, AM1 and PM3 geometry optimizations. The relative energies are from PM3 single point energy calculation of the crystal structures and the optimized structures using MM, AM1 or PM3 geometry optimizations. Refcode stands for the reference codes given in the Cambridge Structural Database.

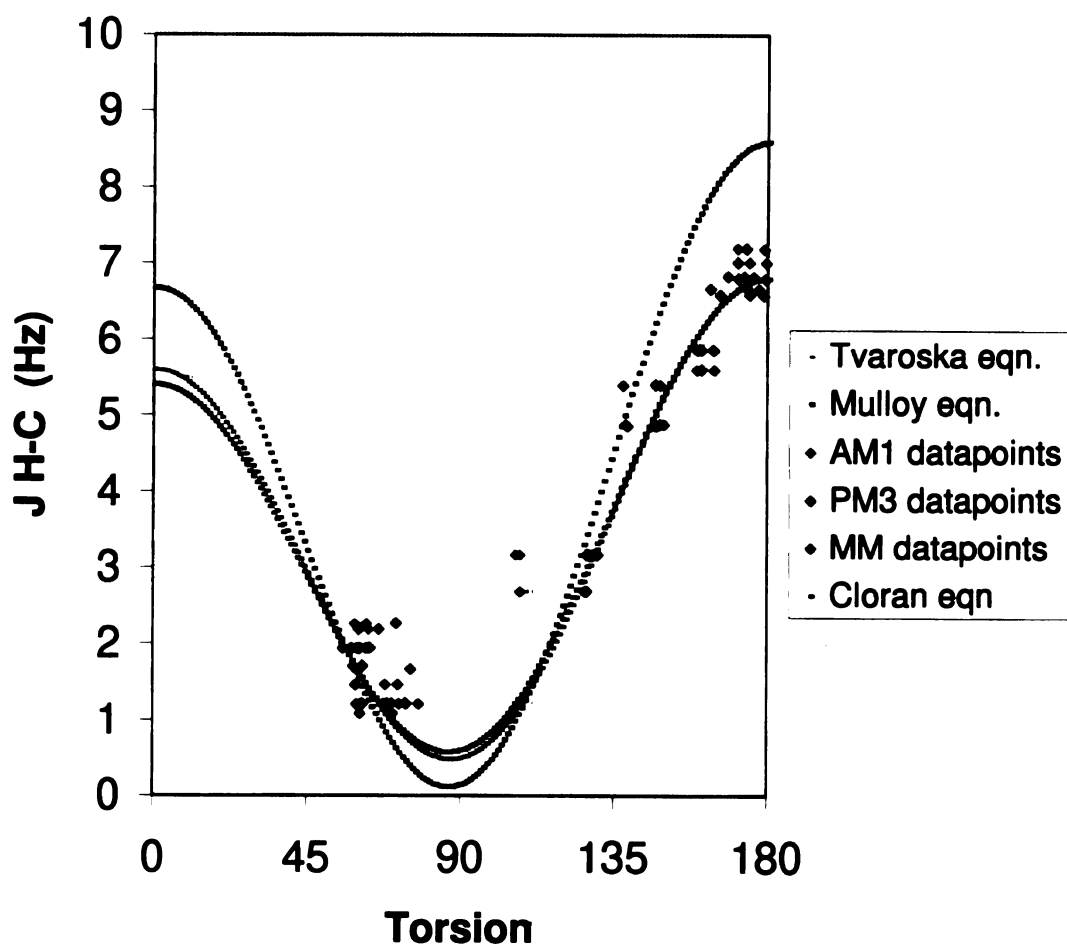
REFCODE	MM	AM1	PM3	Crystal structure	Relative energy (kcal/mol)
AHGLPY10 (8)	-47.2	-34.7	-21.9	-35.1	0.0
AHGLPY01 (8)	-47.9	-36.0	-22.7	-35.2	7.22
AHGALP1 (9)	-46.3	-35.2	-31.9	-39.8	1.13
AHGALP2 (9)	-47.0	-35.3	-37.0	-38.9	3.49
AHGALP3 (9)	-47.4	-35.1	-32.3	-39.1	0.0
BIFTUF1 (10)	-47.7	-35.5	-28.4	-41.0	0.26
BIFTUF2 (10)	-46.0	-37.7	-31.5	-37.9	0.0

accuracy. So, at this stage we recommend that a single equation (**equation 6**, AM1- $\phi\psi$  or **equation 4**, AM1\*) be used until the accuracy of the individual equations (e.g. AM1- $\phi$  and AM1- $\psi$ ) is confirmed with more data points. Since we don't have data points around 0 degrees, to check how different the equations would be if a data point around 0 degrees was present, we added one more data point to our 28 data points. That data point is 5.6 Hz for 11 degrees.<sup>2</sup> The equations thus obtained are termed MM\*, AM1\*, and PM3\* in **Table 4.10**.

There were two crystal structures of 1,6-anhydroglucose (**8**) and two for that of 1,6-anhydromannose (**10**) and three for 1,6-anhydrogalactose (**9**) in the Cambridge Structural Database (**Tables 4.6 – 4.9**). The structures were similar except for the dihedral angles involving C6 that showed some variations. A single point energy calculation of these crystal structures revealed that some of the structures differ by as much as 7.2 kcal/mol. So in solution these molecules may exist as an ensemble average of different conformations. The two crystal structures of 1,6-anhydroglucose (**8**) differ by 7.2 kcal/mol and this corresponded to compound **8** spending 99.9995 % of the time in the lower energy conformer. And the two crystal structures of 1,6-anhydrogalactose (**9**) differ by energies of 3.49 and 1.13 kcal/mol from the third and lowest energy conformer and this corresponded to populations of 0.23 %, 12.81 %, and 86.96 % for the three crystal structures. The two crystal structures of 1,6-anhydromannose (**10**) differ by energy of 0.26 kcal/mol and this means compound **8** would spend 39.16 % of the time in one conformer and 60.842 % in the other conformer.

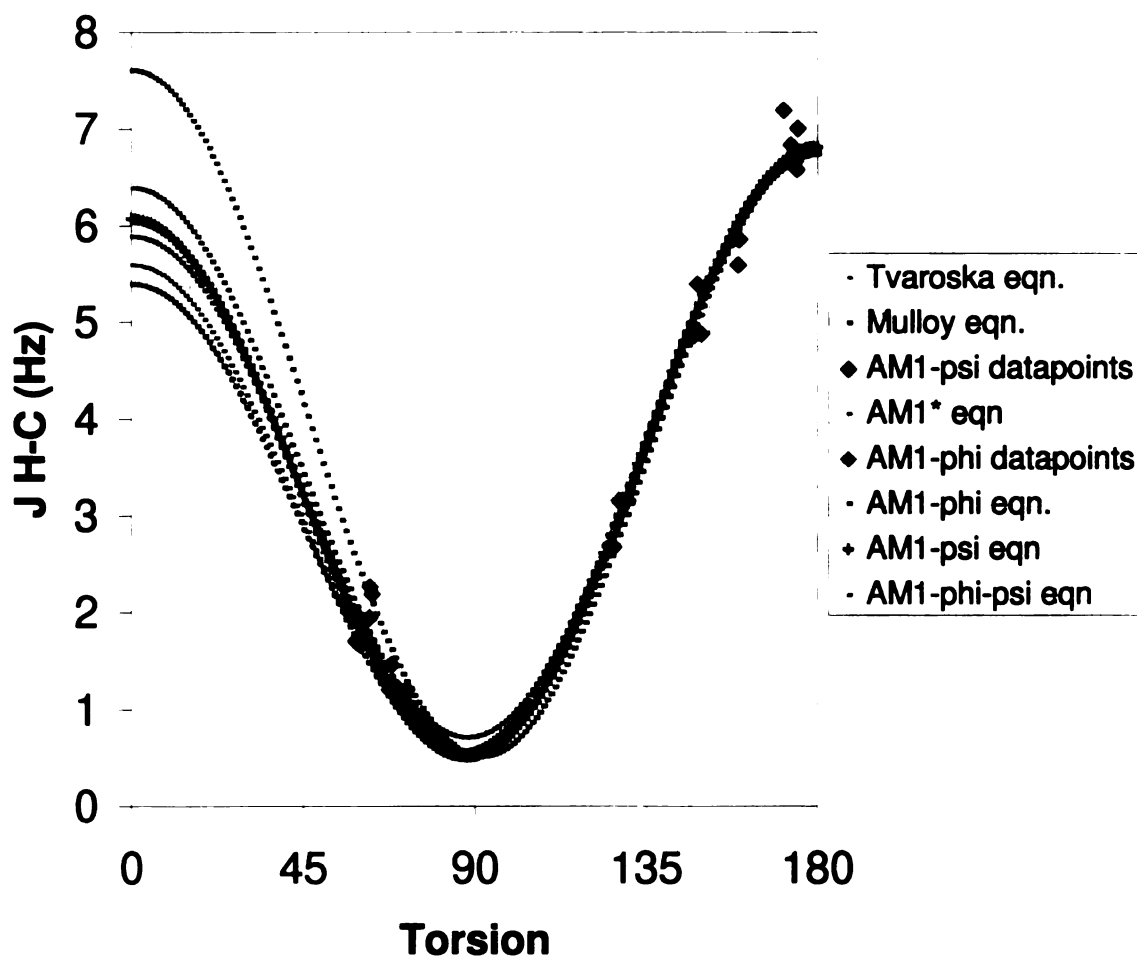
**Figure 4.2.** The new data points plotted against the Karplus-type curves of the previous equations. Our data points are from the experimental coupling constants and from AM1, PM3, and MM geometry optimizations, respectively. This figure is a color presentation for clarity.

## Curves of the three old equations with the new data points



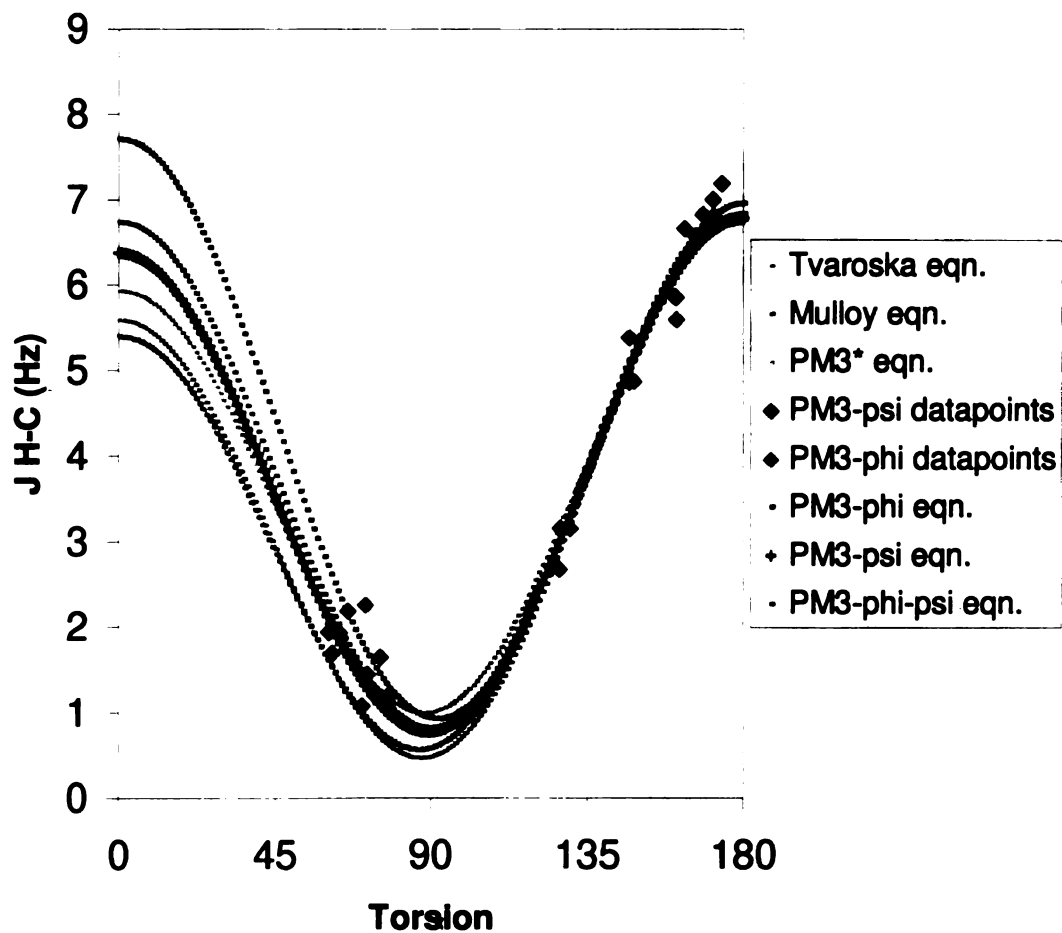
**Figure 4.3.** The new AM1 data points plotted against the Karplus-type curves of the previous equations and the new equations. This figure is a color presentation for clarity.

**Curves of new equations (AM1) and the two old equations with the new data points**

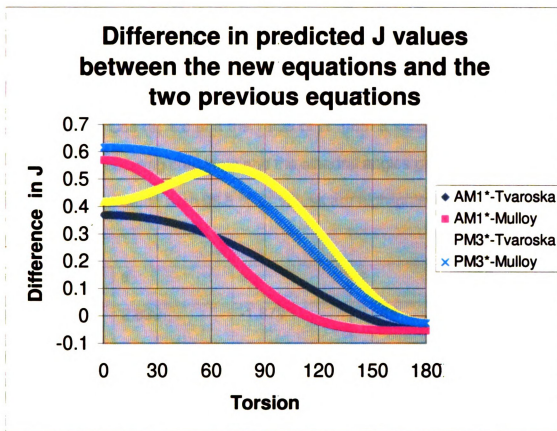


**Figure 4.4.** The new PM3 data points plotted against the Karplus-type curves of the previous equations and the new equations. This figure is a color presentation for clarity.

### Curves of new equations (PM3) and the two old equations with the new data points







**Figure 4.5.** The difference in the predicted values of the vicinal coupling constants between the new Karplus-type equations (AM1\* and PM3\*) and the two previous equations (Tvaroska *et. al.* and Mulloy *et. al.*). This figure is a color presentation for clarity.

**Table 4.10.** The values of the parameters a, b, c of the equation  ${}^3J_{\text{HC}}(\theta) = a \cos^2(\theta) + b \cos(\theta) + c$  and the value of  $R^2$  from a least squares fitting of the heteronuclear coupling constants and the dihedral angles given in Tables 4.4 and 4.5. Those with better than 98.5 %  $R^2$  values are highlighted.

	a	b	c	$R^2$
MM- $\Phi$	2.639	-1.879	2.184	97.7
MM- $\Psi$	2.270	-2.221	2.059	97.1
MM- $\Phi\Psi$	2.380	-2.128	2.102	97.0
PM3- $\Phi$	6.395	0.375	0.943	98.0
PM3- $\Psi$	5.793	-0.120	0.787	98.2
PM3- $\Phi\Psi$	5.953	-0.051	0.846	97.8
<b>AM1-<math>\Phi</math></b>	<b>6.672</b>	<b>0.413</b>	<b>0.524</b>	<b>99.2</b>
<b>AM1-<math>\Psi</math></b>	<b>5.926</b>	<b>-0.371</b>	<b>0.504</b>	<b>99.1</b>
<b>AM1-<math>\Phi\Psi</math></b>	<b>6.020</b>	<b>-0.193</b>	<b>0.573</b>	<b>98.7</b>
MM *	4.004	-1.280	1.513	93.7
<b>AM1*</b>	<b>5.597</b>	<b>-0.421</b>	<b>0.722</b>	<b>98.6</b>
PM3*	5.335	-0.414	1.020	97.7

Geometry optimizations starting from each of the crystal structures of 1,6-anhydroglucose (8), 1,6-anhydrogalactose (9), and 1,6-anhydromannose (10), using molecular mechanics and AM1 and PM3 semi-empirical calculations, gave structures which are lower in energy than any of the crystal structures and which are consistent between the different calculation methods. These structures are also consistent with structures obtained from AM1 and PM3 geometry optimizations starting from built structures (as opposed to starting from crystal geometries). But the MM geometry optimization starting from built structures gave structures relatively not consistent with any of the structures. The main difference is on the dihedral angles involving C6 and the extent of flattening of their chair conformation. While AM1 geometry optimization consistently gave the same chair conformation as that of the crystal structures, PM3 geometry optimizations consistently flatten the chair by about 9 degrees and the MM optimizations do the opposite of PM3 by about 9 degrees (Table 4.9). These results suggest that the refinement of the dihedral angles of the model compounds using the AM1 semi-empirical methods give reliable results.

Previous studies on AM1 and PM3 calculations have shown that the PM3 semi-empirical calculations can lead to physically incorrect results, for example in carbohydrates, because of distortions of the potential energy surface (relative energy versus H---O distance in the intra- or intermolecular interactions).<sup>15,16</sup> The authors suggest that the PM3 method be used with care. Our own studies on the model compounds 1 – 10 using the PM3 and AM1 semi-empirical calculations revealed that the values of the vicinal heteronuclear coupling

constants for similar dihedral angles are more spread for the PM3 results than the results from AM1 calculations. Furthermore, PM3 optimizations tend to flatten the chair conformation of the 1,6-anhydro sugars (8-10) (Table 4.9). This was also apparent in the higher  $R^2$  value for the AM1 results (equation 4:  ${}^3J_{\text{HC}}(\theta) = 5.597 \text{ Cos}^2(\theta) - 0.421 \text{ Cos}(\theta) + 0.722$ ) than the PM3 results (equation 5:  ${}^3J_{\text{HC}}(\theta) = 5.335 \text{ Cos}^2(\theta) - 0.414 \text{ Cos}(\theta) + 1.020$ ) (Table 4.10). As a result, we suggest that the equation (equation 4 i.e.  ${}^3J_{\text{HC}}(\theta) = 5.597 \text{ Cos}^2(\theta) - 0.421 \text{ Cos}(\theta) + 0.722$ ) obtained from the least squares fitting of the dihedral angles refined by AM1 semi-empirical geometry optimizations to the experimentally measured vicinal proton-carbon coupling constants,  ${}^3J_{\text{HC}}$ , be used for predicting the dihedral angles defined by the four atoms H-C-O-C ( $\phi$  and  $\psi$  for carbohydrates) from vicinal proton-carbon coupling constants,  ${}^3J_{\text{HC}}$ , or vice versa.

## Conclusions

The three dimensional structures of the synthetic model compounds 1-7 and the commercially available compounds 8-10 have been refined by Molecular Mechanics and AM1 and PM3 Semi-empirical calculations. The vicinal proton-carbon coupling constants,  ${}^3J_{\text{HC}}$ , have been measured in solution by two-dimensional excitation-sculptured indirect-detection experiment (EXSIDE) NMR spectroscopy. From the least squares fitting of the measured coupling constants and the glycosidic dihedral angles of the rigid molecules, equation of the form:  ${}^3J_{\text{HC}}(\theta) = a \text{ Cos}^2(\theta) + b \text{ Cos}(\theta) + c$ , are proposed where  $\theta$  stands for the dihedral angles defined by the four atoms H-C-O-C ( $\phi$  and  $\psi$  for carbohydrates).

This new Karplus-Type equations relating the vicinal proton-carbon coupling constants,  $^3J_{\text{HC}}$ , to the glycosidic dihedral angles  $\phi$  (H1'-C1'-OX-CX) and  $\psi$  (HX-CX-OX-C1'), respectively have been compared to the previous equations and shown to give a higher  $R^2$  value.

### **Acknowledgements**

This work was supported by the Michigan State University Research Excellence Fund.

## References

1. Tvaroska, I; Hricovini, M; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359.
2. Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988**, *184*, 39.
3. Rundolf, T.; Kjellberg, A.; Damberg, C.; Nishida, T.; Widmalm, G. *Magn. Reson. Chem.* **1998**, *36*, 839.
4. Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1999**, *121*, 9843.
5. Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, P. B.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, S. *J. Am. Chem. Soc.* **1998**, *120*, 11158.
6. Haasnoot, C. A., De Leeuw, F. A., & Altona, C. *Tetrahedron* **1980**, *36*, 2783.
7. Ernst, R., Bodenhausen, G., & Wokahun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimension*, **1987**, pp 431-440, Oxford University Press, Oxford.
8. Bax, A. & Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355.
9. Bax, A. & Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565.
10. Titman, J., Neuhaus, D., & Keeler, J. *J. Magn. Reson.* **1989**, *85*, 111.
11. Sangers, J. K., & Hunter, B.K. *Modern NMR Spectroscopy*, **1993**, pp 193-197, Oxford University Press, Oxford.
12. Krishnamurthy, V. V. *J. Magn. Reson.* **1996**, *121*, 33.
13. Dewar, M. J. S.; Zoebish, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

14. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
15. Csonka, G. I.; Angyan, J. G. *J. Mol. Struct. (Theochem)* **1997**, *393*, 31.
16. Kallies, B; Mitzner, R. *J. Mol. Model.* **1995**, *1*, 68.

## **Chapter 5**

# **Studies on the Solution Conformations and Dynamics of Maltose, Reduced Maltotetraose, and the Potent Glycohydrolase Inhibitor Acarbose**



## **Abstract.**

We applied the rule-based method (Chapter Two) and the new Karplus-type equations (Chapters Three and Four) that we have developed in this work in combination with existing methods using multi-nuclear, multidimensional NMR spectroscopy, molecular mechanics calculations and molecular dynamics simulations to study the solution conformations and dynamics of a family of  $\alpha$ -1->4 linked oligosaccharides namely, the disaccharide maltose and the tetrasaccharide reduced-maltotetraose, and the pseudo-tetrasaccharide acarbose, which is a potent glycohydrolase inhibitor. The study revealed that all three oligosaccharides occupy the RR region of the  $\phi$ - $\psi$  space as predicted by the rule-based method. Also we found out that the pseudotetrasaccharide acarbose is more flexible than maltose and reduced maltotetraose as revealed from the NMR spectroscopy and molecular dynamics studies.

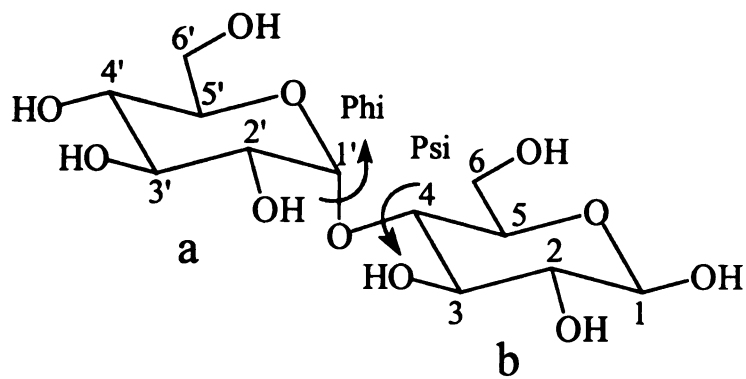
## **Introduction.**

Carbohydrates provide diverse three-dimensional (3D) structural possibilities and have several applications. For example, recent studies have demonstrated that oligosaccharides are involved in a number of recognition events such as cell adhesion, metastasis, fertilization and embryonic development, amongst others.<sup>1,2</sup> In view of their diverse function in a variety of biological systems, the primary structures and 3D structures of carbohydrates are of great interest. Due to their diversity and the inability to crystallize such conjugates, nuclear magnetic resonance (NMR) spectroscopy, in combination with molecular mechanics calculations and molecular dynamics studies, is a method of choice for their identification and conformational analysis.

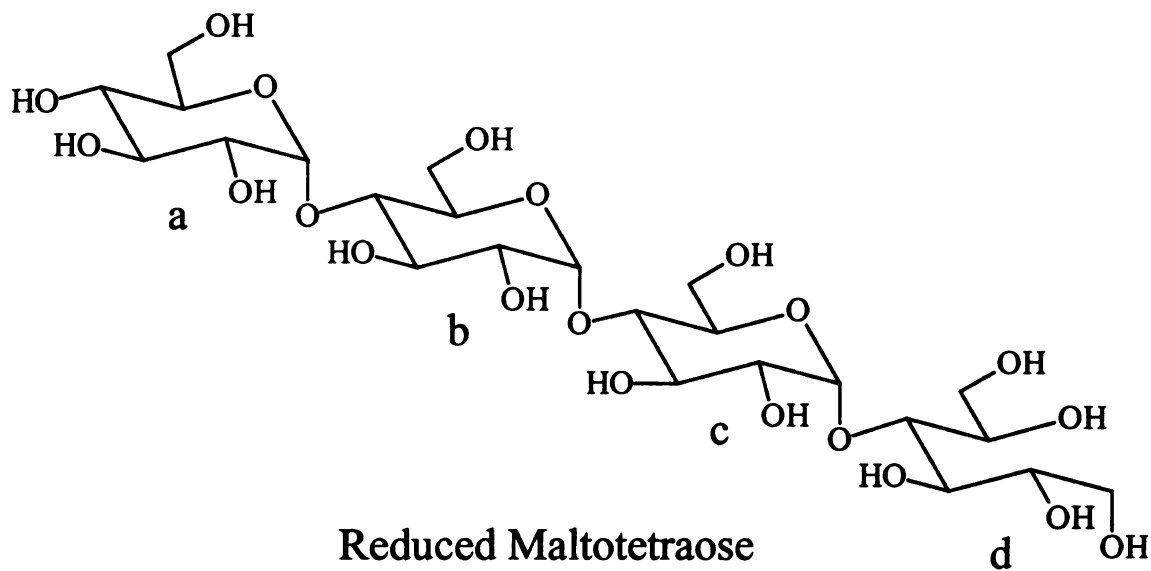
A major problem when using NMR spectroscopy to study carbohydrates is that, except for the anomeric protons, all the protons on oxygenated atoms (e.g., those on carbons 2-6 in aldohexoses) appear in a range of ~3.3 ppm to ~5.3 ppm. Hence, as the number of monosaccharide units increases, the assignment of peaks for each proton (and carbon) becomes increasingly difficult. This means extracting coupling constants and NOESY volumes--which are useful for determining dihedral angles and distances, respectively--becomes very difficult.

In chapter two we have collected and analyzed the crystal structures of oligosaccharides and proposed a simple and accurate rule based method for predicting the conformations of oligo and polysaccharides.<sup>3</sup> Because solution conformations of carbohydrates may not be identical to their crystal structure, it would be very useful to come out with a rule-based method for predicting

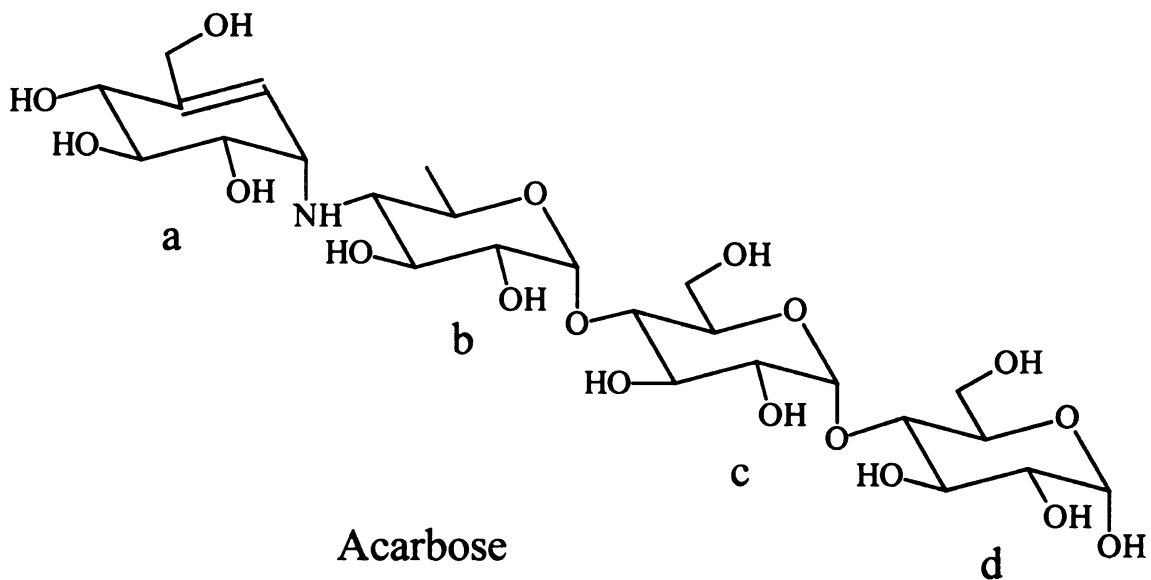
**Figure 5.1** Schematic diagrams of maltose, reduced maltotetraose, and acarbose.



**Maltose**



**Reduced Maltotetraose**



**Acarbose**

carbohydrate conformations in solution by analyzing a library of solution conformations of carbohydrates. In an effort to build a database of solution conformations of carbohydrates and to see how the rule-based method and the new Karplus-type equations compare to existing methods and equations we decided to study the solution conformation and dynamics of maltose, reduced maltotetraose, and the glycohydrolase inhibitor acarbose by a combination of nuclear magnetic resonance (NMR) spectroscopy, molecular mechanics calculations and molecular dynamics simulations.

## **Material and Methods**

### **Oligosaccharides**

The commercially available maltose is used as it is. The reduced maltotetraose was obtained by reduction of the commercially available maltotetraose with sodium borohydride in methanol. Acarbose was generously supplied by Professor Jack Preiss, Department of Biochemistry, Michigan State University.

### **Nuclear Magnetic Resonance (NMR) Spectroscopy**

Nuclear Magnetic Resonance spectra of the three oligosaccharides were obtained in deuterated water. Spectra were recorded at 30 °C. The residual solvent peaks were used as reference. Combinations of homonuclear- and heteronuclear- (one- and) two-dimensional NMR experiments were performed in

order to allow the assignments of the signals for the three oligosaccharides and to measure the coupling constants. This included: double quantum filtered J-correlated spectroscopy (DQF-COSY) spectra (Ernst *et al.*, 1987),<sup>4</sup> total correlated spectroscopy (TOCSY) spectra (Bax & Davis, 1985),<sup>5</sup> heteronuclear multi-quantum coherence (HMQC) experiment (Bax & Subramanian, 1986),<sup>6</sup> and nuclear Overhauser effect (NOESY) experiment (Sangers *et al.*, 1993).<sup>7</sup>

Analysis of the spin systems for the three oligosaccharides from the 1D NMR spectra and from DQF-COSY and TOCSY spectra gave the connectivities. The tables of assignment of the proton and carbon chemical shifts of the three oligosaccharides are given in **Tables 5.1 – 5.3**. Once the assignments were made NOESY experiments were performed to determine the various inter-nuclear distances (**Table 5.4 and 5.5**). Spectra of the NOESY experiments were acquired at different mixing times (300, 400, and 500 ms) to address the potential problem of spin diffusion. The vicinal proton-carbon coupling constants were obtained from excitation-sculptured indirect-detection (EXSIDE) experiments (Krishnamurthy *et al.*, 1996).<sup>8</sup> The EXSIDE spectra were measured on Varian Inova 600 MHz NMR spectrometer operating at 600 MHz for protons and 150 MHz for carbons. And the rest of the one- and two-dimensional NMR spectra were measured on Varian VXR 500 MHz NMR spectrometer operating at 500 MHz for protons and 125 MHz for carbons.

**Table 5.1.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments (in ppm) of maltose.

H/C	Residue a		Residue b	
1	5.32	99.1	4.56	95.3
2	3.49	72.3	3.16	73.7
3	3.61	72.6	3.67	75.8
4	3.33	69.1	3.51	76.5
5	3.63	71.4	3.54	74.3
6	3.76	60.1	3.82	60.4
6	3.65	60.1	3.66	60.4

**Table 5.2.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments (in ppm) of reduced maltotetraose.

H/C	Residue a		Residue b		Residue c		Residue d	
1	5.38	101.6	5.38	101.6	5.12	102.2	3.64	62.2
2	3.60	73.3	3.60	73.3	3.60	73.3	3.65	64.3
3	3.97	74.7	3.97	74.7	3.97	74.7	3.67	64.5
4	3.41	71.2	3.65	78.8	3.65	78.8	3.88	83.7
5	3.71	74.6	3.69	74.6	3.66	74.6	3.71	63.9
6	3.84	62.2	3.84	62.1	3.84	62.2	3.64	62.2
6	3.65	62.2	3.64	62.1	3.64	62.2	3.64	62.2



**Table 5.3.**  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift assignments (in ppm) of acarbose.

H/C	Residue a		Residue b		Residue c		Residue d			
1	3.52	57.1	5.27	101	5.35	100	4.61	96.6	5.19	92.7
2	3.77	73.8	3.56	71.6	3.53	72.2	3.23	75.0	3.51	71.6
3	3.72	73.9	3.55	73.7	3.93	74.3	3.73	77.3	3.89	74.3
4	4.00	72.1	2.46	65.7	3.60	78.0	3.59	77.9	3.62	78.1
5	--	140	3.74	70.4	3.80	72.1	3.53	75.5	3.90	71.0
6	4.18	62.8	1.30	47.3	3.92	61.6	3.86	61.6	3.88	61.6
6	4.08	62.8	--	47.3	3.83	61.6	3.78	61.6	3.75	61.6
7	5.86	124								

**Table 5.4.** Distances calculated from NOESY experiment for reduced maltotetraose. The letters A, B, C, and D in column 1 refer to the different residues and the numbers refer to the protons (e.g. A1-B4 refer to proton 1 of residue A with proton 4 of residue B).

Atom pairs	NOESY Volume	Distances in Å
A1-B4	79.2	2.37
B1-C4	81.2	2.35
C1-D4	100	2.28
B1-C6	31.2	2.76
A1-B6	22.9	2.91
C1-D6	18.7	3.00
A2-A4	37.5	2.67
A4-A5	16.7	3.06
A4-A6	10.4	3.31
A3-A5	16.7	2.70

**Table 5.5.** Distances calculated from NOESY experiment for acarbose. The letters A, B, C, and D in column 1 refer to the different residues and the numbers refer to the protons (e.g. A1-B4 refer to proton 1 of residue A with proton 4 of residue B).

Atom pairs	NOESY Volume	Distances in Å
A6-A7	3	3.54
A1-A7	18	2.63
A6'-A7	5	3.25
A4-A7	3	3.54
C1-D4	61	2.15
B1-C4	49	2.22
A7-B6	6	3.16
A7-B4	1	4.25
B2-B4	16	2.68
A1-B4	17	2.65

## **Molecular Mechanics Calculations and Molecular Dynamics Simulations**

Molecular mechanics calculations and molecular dynamics simulations were performed using the DREIDING force fields (Mayo et al., 1990)<sup>9</sup> implemented in the BIOGRAF<sup>10</sup> (Molecular Simulations Inc., Waltham, MA 02154) program. The Drieding force field includes a harmonic term for the bond distortions, a simple harmonic function for the valence angle term, an improper torsion term to evaluate inversions about atomic centers, a cosine expansion torsional term, a Coulombic potential for evaluating electrostatic energy, a Leonard-Jones 12-6 potential for van der Waals contributions and a Leonard-Jones 12-10 potential for evaluating hydrogen bonding contributions. Minimizations were performed using the conjugate gradient method. The default parameters given in this program for carbohydrate rings were used without modification since they have been validated earlier (Wang *et. al.* 1996).<sup>11</sup> The calculated distance constraints, from NOESY experiments, were included as harmonic restraints with a force constant of 10,000 kcal per mol per Å<sup>2</sup> that was gradually relaxed to 1000 and 10 kcal per mol per Å<sup>2</sup>. Grid search studies were performed using sequential search mode in the Biograf program, with a 5 degrees step growth search from 0 to 360 degrees, for each defined dihedral angle. **Figures 5.2 and 5.3** are color presentations for clarity

## **Results and Discussion**

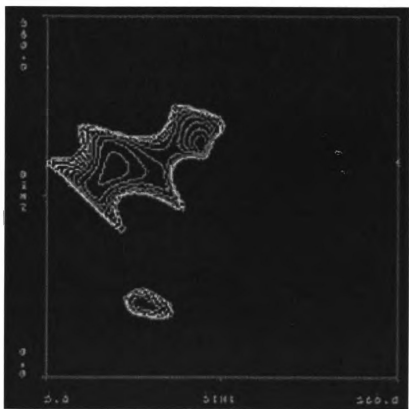
The three oligosaccharides namely, maltose, reduced maltotetraose and acarbose, all have  $\alpha$ -1->4 linkages which is the linkage type of the

polysaccharide amylose. While the disaccharide maltose has one  $\alpha$ -1->4 linkage, the tetrasaccharides maltotetraose and acarbose each have three  $\alpha$ -1->4 linkages. Maltose and maltotetraose are the building blocks of the polysaccharide amylose. Acarbose is a potent inhibitor of  $\alpha$ -glucosidases. It is used for treatment of diabetes, hyperlipidaemia and obesity and is marketed as Glucobay.

The potential energy (PE) surfaces of the three oligosaccharides in this study were calculated using the Dreiding<sup>9</sup> force field in the Biograf<sup>10</sup> molecular mechanics (MM) program. The PE surfaces provide a first estimation of the conformational regions that are energetically accessible. While two minima (**Figure 5.2**) were found for maltose and maltotetraose (the second minimum 9 kcal/mol above the first minimum), the acarviosine (a-b) unit of acarbose had two minima with 1 kcal/mol energy difference and a third minimum that was at 6 kcal/mol above the lowest minimum. This indicated that while maltose and maltotetraose should have a reasonably rigid conformation, the a-b unit of acarbose should be more flexible because of the fact that its residue-a at the non-reducing end is not actually a monosaccharide. The isoenergy contours in **Figure 5.2** were drawn with interpolation of 1 kcal/mol above the minimum.

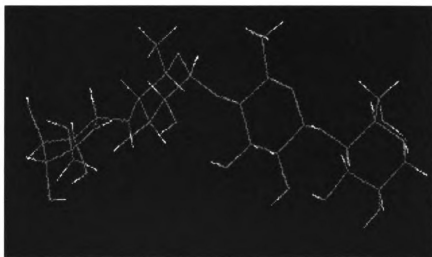
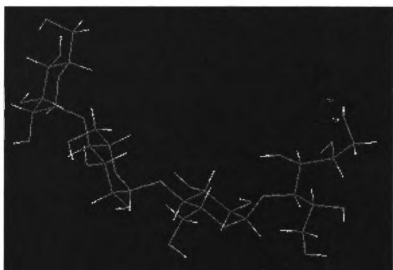
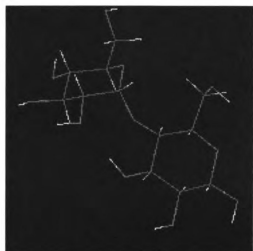
The next task was to perform combinations of homonuclear and heteronuclear, one- and two-dimensional NMR experiments in order to allow the assignments of the signals for the three oligosaccharides and to extract internuclear distances. This included performing double quantum filtered J-correlated spectroscopy (DQF-COSY) spectra,<sup>4</sup> total correlated spectroscopy

**Figure 5.2** Rigid residue potential energy (PE) surfaces of maltose (top) and acarviosine (the a-b unit of acarbose) (bottom). Isoenergy contours are drawn with interpolation of 1 kcal/mol above the minimum. Notice that while maltose has practically a single minimum, acarviosine has two minima with 1 kcal/mol PE difference and a third minimum at 6 kcal/mol higher. The dihedral angles of maltose are defined as (O5'-C1'-O4-C4) for  $\phi$  and (C1'-O4-C4-C5) for  $\psi$ . The dihedral angles acarviosine are defined as (C7'-C1'-N4-C4) for  $\phi$  and (C1'-N4-C4-C5) for  $\psi$ . This figure is a color presentation for clarity.



**Figure 5.3** Solution conformations of maltose (top), reduced maltotetraose (middle) and acarbose (bottom) as determined from constrained minimization using distance constraints from 2D NOESY experiments. This figure is a color presentation for clarity.



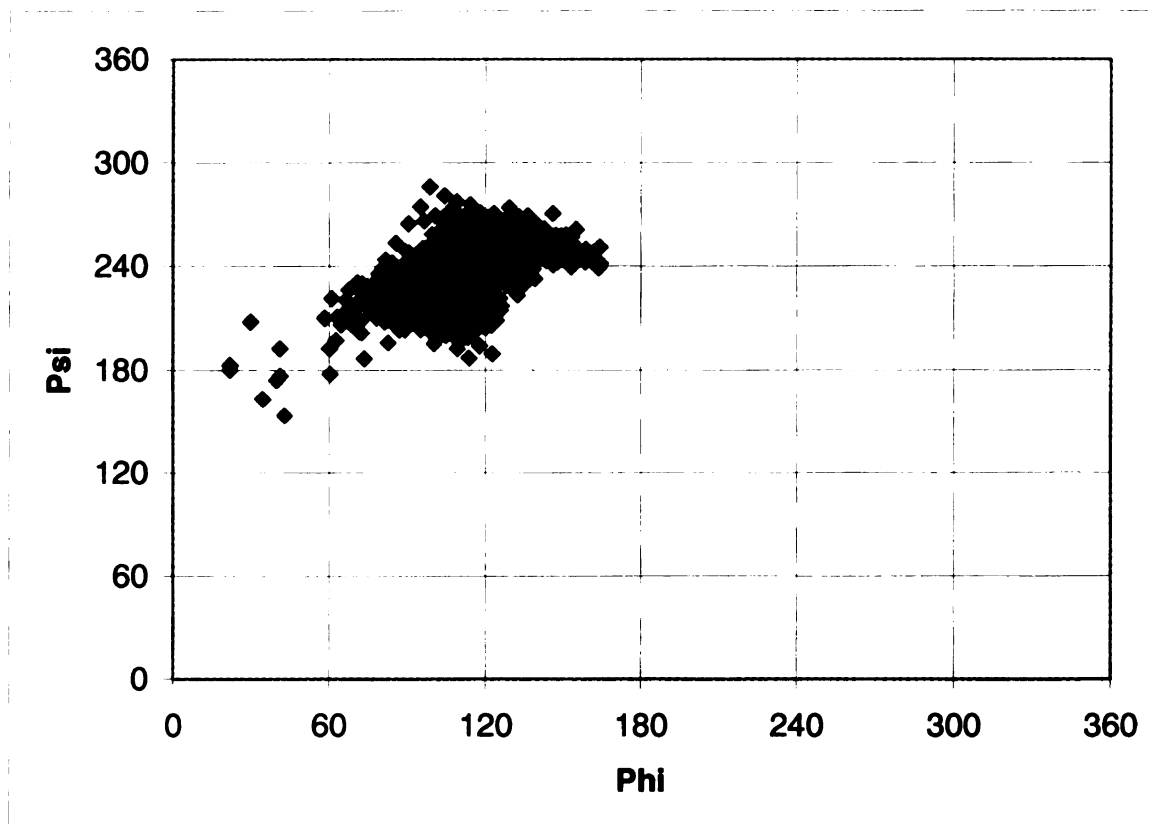


trace  
zation  
color

(TOCSY) spectra,<sup>5</sup> heteronuclear multi-quantum coherence (HMQC) experiment,<sup>6</sup> and two-dimensional nuclear Overhauser effect (n.O.e.) experiment (NOESY).<sup>7</sup>

Analysis of the spin systems for the glycoside ring systems of the three oligosaccharides with 1-D traces from DQF-COSY and TOCSY spectra gave the connectivities. Once the assignments were made (**Table 5.1-5.3**), n.O.e. experiments were performed to determine the various inter-nuclear distances. Spectra of the n.O.e. experiments were acquired at different mixing times to address the potential problem of spin diffusion. Assuming isotropic motion then, the cross-relaxation rates and thus inter-nuclear distances between nuclei *i* and *j*,  $r_{ij}$ , for each pair wise interaction were extracted (**Tables 5.4-5.5**) from the volume of the NOESY peak,  $V_{ij}$ , using the relation  $r_{ij} = r_{ref} (V_{ref} / V_{ij})^{1/6}$  where,  $r$  = distance,  $V$  = volume,  $ref$  = reference pair of nuclei the distance between which is fixed and known.

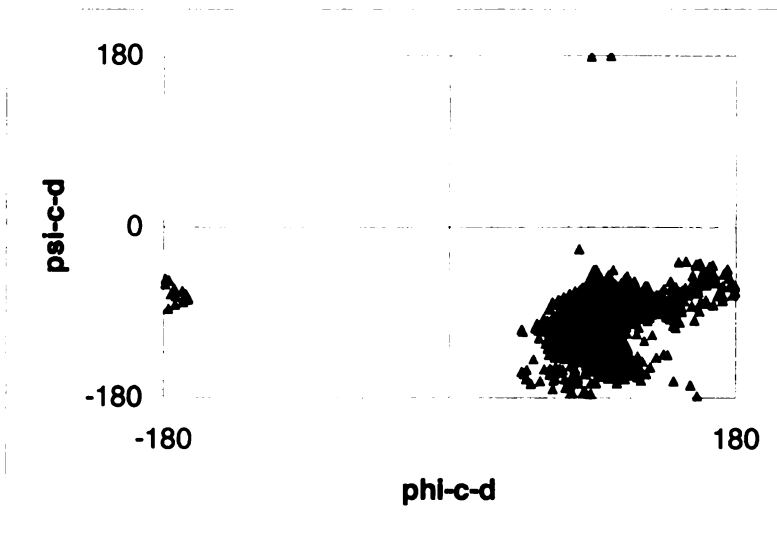
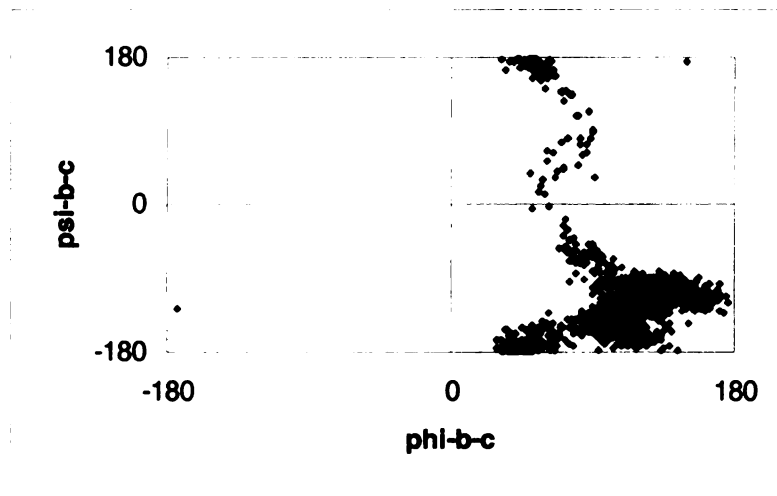
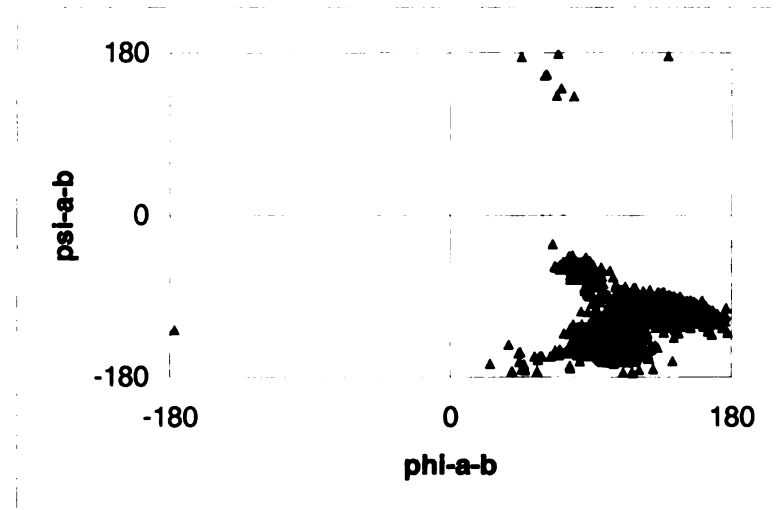
This was followed by constrained minimization using the inter-nuclear distances obtained at the previous step as constraints in the force fields of molecular mechanics calculations (the Drieding force fields<sup>9</sup> implemented in the BIOGRAF program) to determine the average conformation about glycosidic linkages of these oligosaccharides in solution. For maltose the distance between H1' and H4 was 2.35 Å. The constraint-minimized solution conformations of maltose, reduced maltotetraose, and acarbose though similar were not identical as shown in **Figure 5.3** and **Table 5.6**. Notice that the two tetrasaccharides,



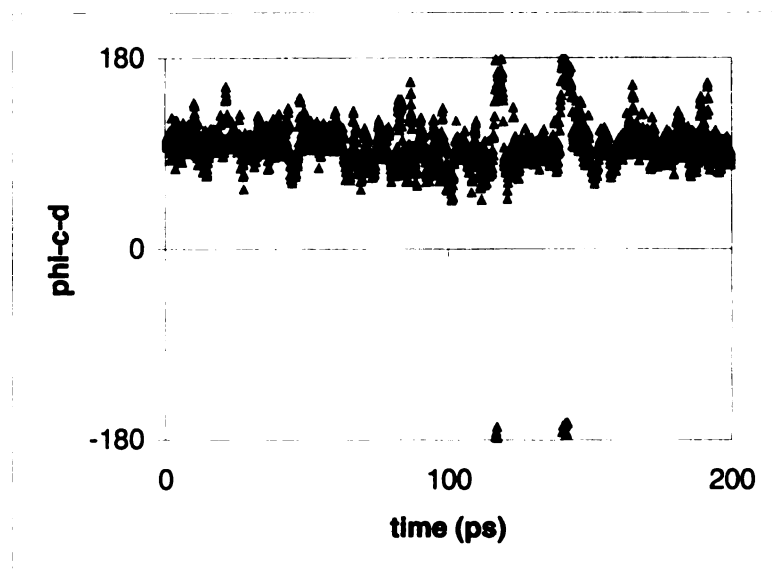
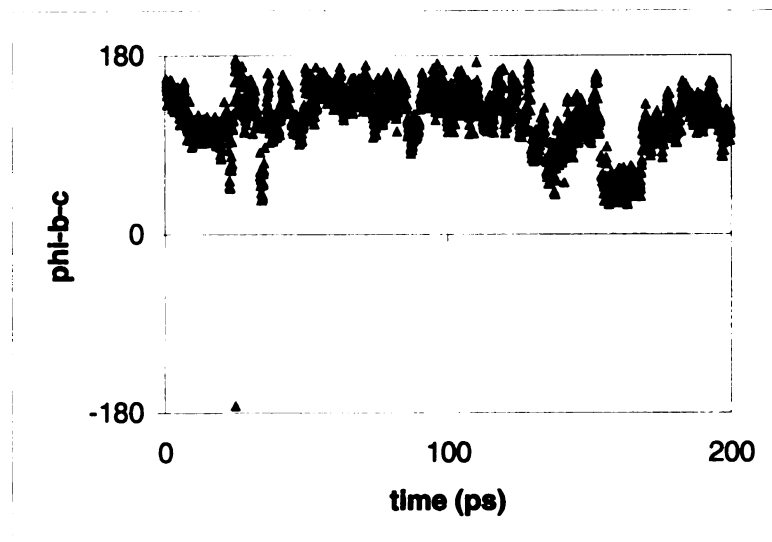
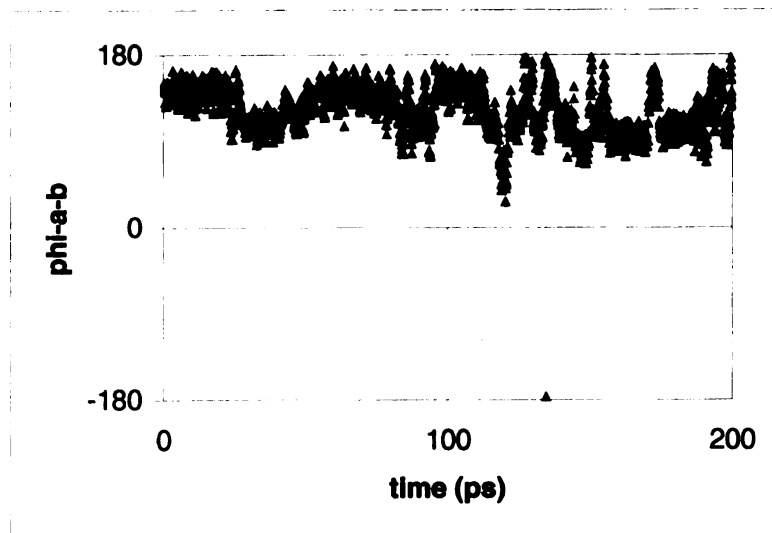
**Figure 5.4** The molecular dynamics population map of maltose, shown above, is consistent with the PE surface. The dihedral angles are defined as (O5'-C1'-O4-C4) for  $\phi$  and (C1'-O4-C4-C5) for  $\psi$ .

**Figure 5.5** Molecular dynamics population maps of reduced maltotetraose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is the  $\phi$  dihedral angle (O5'-C1'-O4-C4) and the y-axis is the  $\psi$  dihedral angle (C1'-O4-C4-C5), both plotted from -180 to 180°.

atrace  
ucture  
nd c-d  
axis is

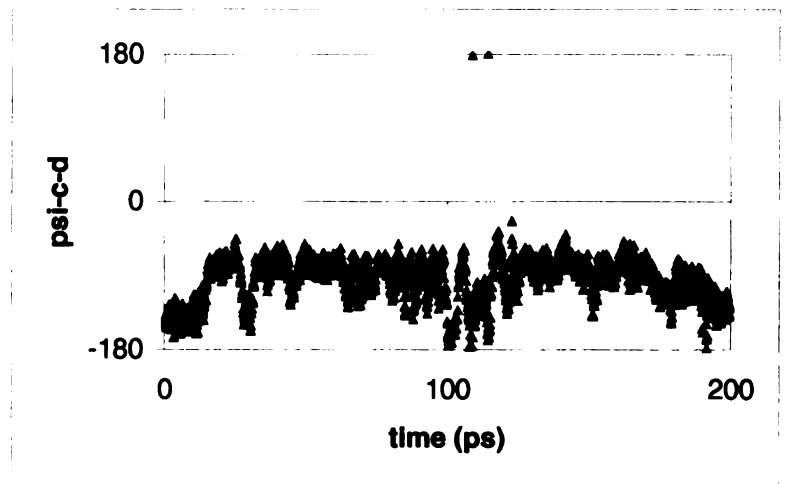
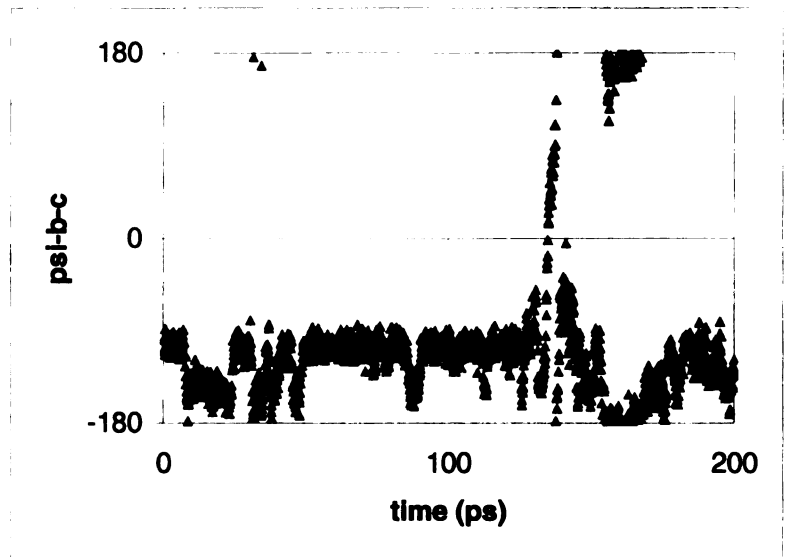
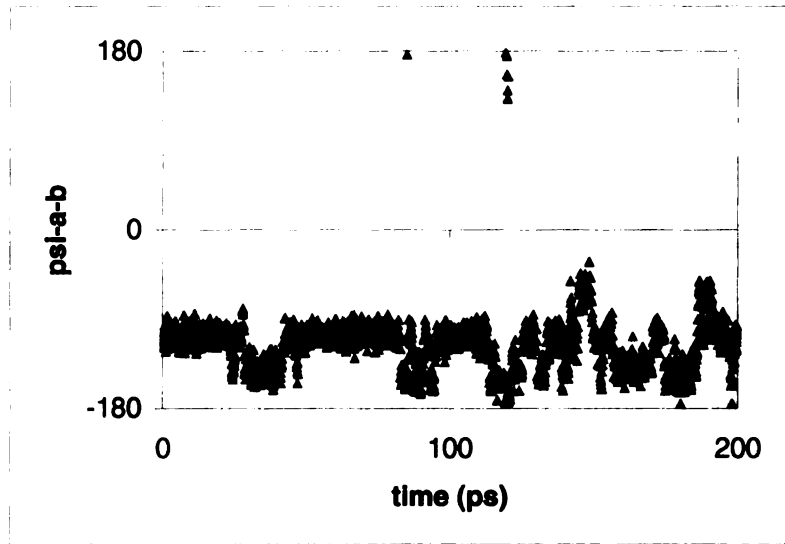


**Figure 5.6** Molecular dynamics Trajectories of the  $\phi$  angle of reduced maltotetraose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is time in ps and the y-axis is the  $\phi$  dihedral angle (O5'-C1'-O4-C4) plotted from -180 to 180°.



**Figure 5.7** Molecular dynamics Trajectories of the  $\psi$  angle of reduced maltotetraose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is time in ps from 0 to 200 ps and the y-axis is the  $\psi$  dihedral angle (C1'-O4-C4-C5) plotted from  $-180$  to  $180^\circ$ .



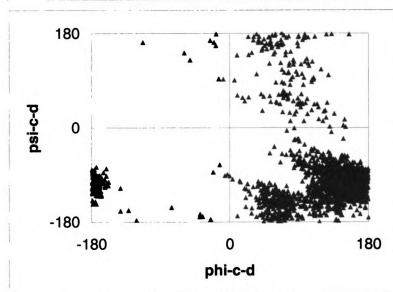
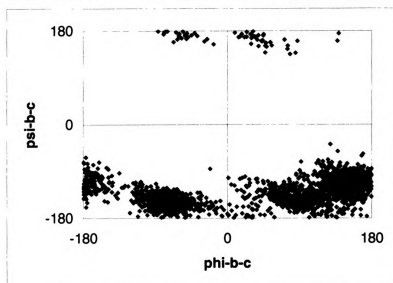
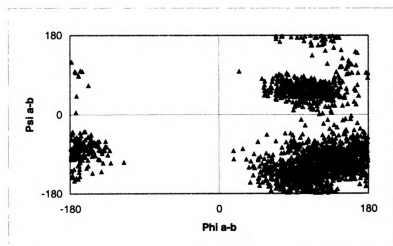


reduced maltotetraose and acarbose, showed the beginning of the helical type conformation that is known for amylose.

Because solution conformations are ensemble averages of many conformations, molecular dynamics simulations were performed in order to gain insight into the dynamics of the structures of the oligosaccharides in solution. Starting from the constraint-minimized solution conformations of the three oligosaccharides, 200 ps molecular dynamics simulations (canonical NVT) were performed using the Biograf molecular mechanics program.<sup>10</sup> The results of the molecular dynamics simulations are shown in **Figures 5.4-5.10**. The molecular dynamics population maps of all the three oligosaccharides (**Figures 5.4, 5.5, and 5.8**) were consistent with the potential energy surfaces (**Figure 5.2**). Notice that in **Figures 5.2 and 5.4** the dihedral angles are plotted from 0 to 360 degrees and not -180 to 180 degrees and as a result the areas populated by the oligosaccharides may look different from the other population maps that are drawn from -180 to 180 degrees, but it is just a shift of 180 degrees in both dihedral angles.

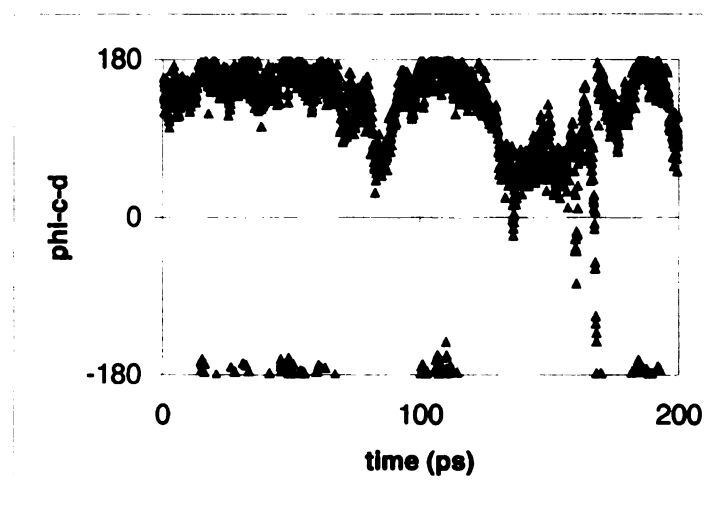
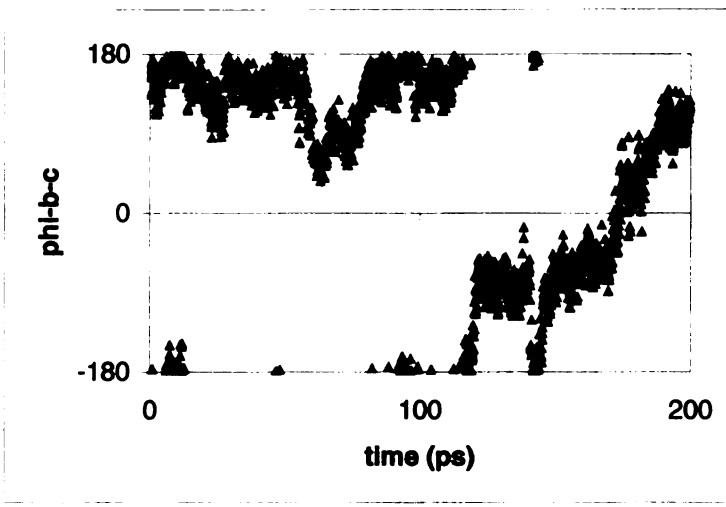
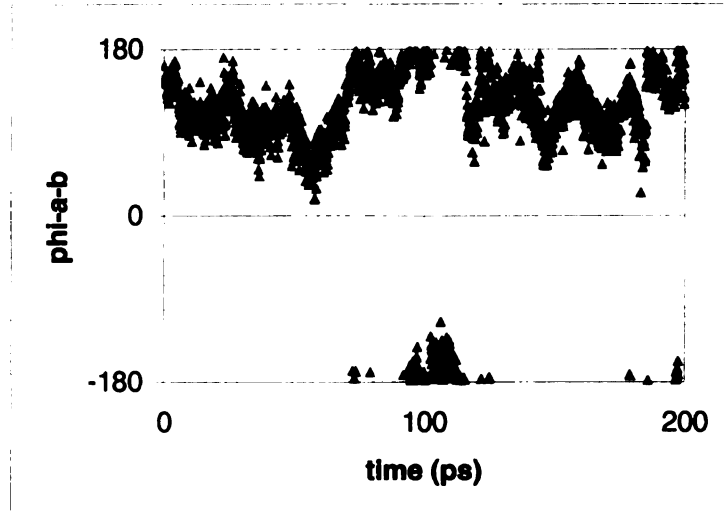
The molecular dynamics simulation studies revealed that, as the study of the potential energy surface indicated, the acarviosine (a-b) unit of acarbose showed significantly greater flexibility than all the other linkages (**Figure 5.4-5.10**). In fact, the b-c and c-d units of acarbose were also slightly more flexible than the linkages of maltose and reduced maltotetraose and this may be associated with the greater flexibility of the a-b unit of acarbose. The molecular dynamics trajectories of the  $\phi$  and  $\psi$  angles of the three oligosaccharides

**Figure 5.8** Molecular dynamics population maps of acarbose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is the  $\phi$  dihedral angle (O5,C7'-C1'-O4,N4-C4) and the y-axis is the  $\psi$  dihedral angle (C1'-O4,N4-C4-C5), both plotted from -180 to 180°.



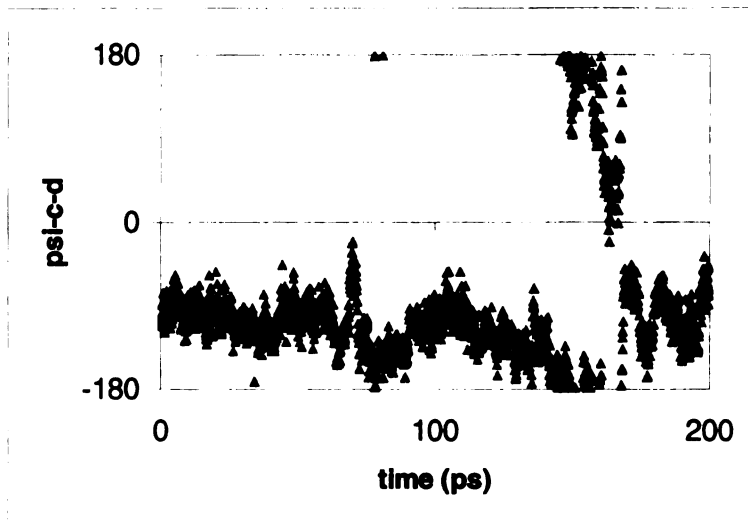
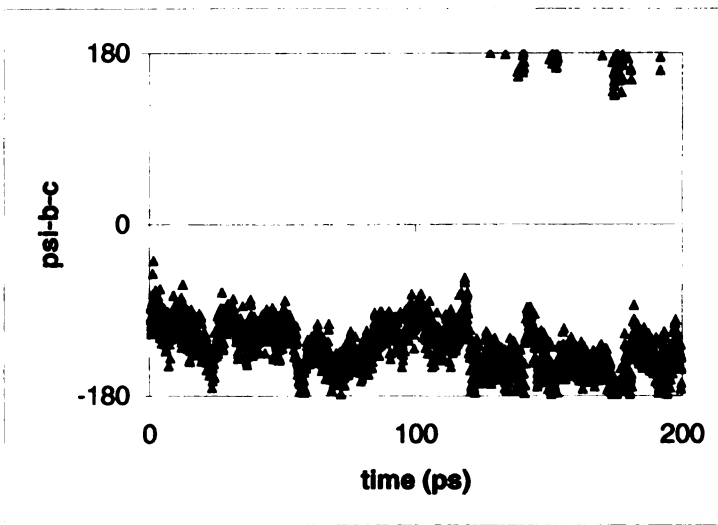
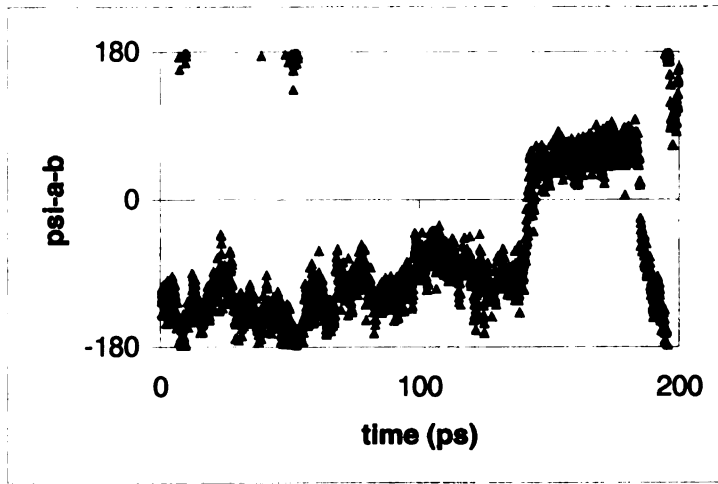
**Figure 5.9** Molecular dynamics Trajectories of the  $\phi$  angle of acarbose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is time in ps from 0 to 200 ps and the y-axis is the  $\phi$  dihedral angle (C7',O5'-C1'-N4,O4-C4) plotted from  $-180$  to  $180^\circ$ .

use from  
ure. The  
bottom).  
ral angle



**Figure 5.10** Molecular dynamics Trajectories of the  $\psi$  angle of acarbose from 200 ps simulation starting from its constraint-minimized solution structure. The plots are for the glycosidic linkages a-b (top), b-c (middle) and c-d (bottom). The x-axis is time in ps from 0 to 200 ps and the y-axis is the  $\psi$  dihedral angle (C1'-N4,O4-C4-C5) plotted from  $-180$  to  $180^\circ$ .

se from  
re. The  
ottomi  
l angle





(**Figures 5.6, 5.7, 5.9, and 5.10**) showed the variation of the dihedral angles against time as the simulation progresses. It was evident that acarbose showed frequent transitions between different conformations as compared to reduced maltotetraose. Comparison of the solution conformations, the grid-search minimum energy conformations, and the crystal structures of the three oligosaccharides (**Table 5.6**) revealed that the structures are reasonably consistent with each other. Two conformations of the a-b unit of acarbose have been inferred to exist at neutral, acidic and alkaline pHs.<sup>12,13</sup> This is consistent with our studies as shown in **Figures 5.2 and 5.9**. The rule based prediction of the dihedral angles of an  $\alpha$ -1- $\rightarrow$ 4 linkage in oligosaccharides (Chapter Two) gave 98, -128 degrees for  $\phi$  and  $\psi$ .<sup>3</sup> This is consistent with conformations determined by NMR spectroscopy and molecular mechanics calculations and molecular dynamics simulations and structures from X-ray crystallography (**Table 5.6**).

An important piece of structural information comes from the calculation of dihedral angles from vicinal proton-carbon coupling constants measured from NMR spectroscopy, or vice versa. To calculate the dihedral angles ( $\theta$ ) between the glycoside residues from the vicinal proton-carbon coupling constants,  $^3J_{HC}$ , we used the three previous<sup>14-16</sup> and the new (Chapter Four) Karplus-type equations which have the general form  $^3J_{CH} = a \cos^2(\phi) + b \cos(\phi) + c$ . The  $^3J_{CH}$  values measured from 2D EXSIDE experiments for the dihedral angles H1b-C1b-O4c-C4c and H1c-C1c-O4d-C4d of acarbose were 2.60 and 3.60 Hz, respectively. The corresponding value for maltose was 4.88 Hz. These values

**Table 5.6.** Values of the glycosidic dihedral angles of the three oligosaccharides from constrained minimization, grid search studies and from crystal structure.

Molecule	Linkage	Constrained Minimization	Grid Search	Crystal Structure
Maltose	a-b	120.56, -105.83	100, -145	124, -108 <sup>a</sup>
Reduced maltotetraose	a-b	124.34, -104.20	100, -145	
	b-c	123.27, -103.06	100, -145	
	c-d	148.49, -114.70	100, -145	
Acarbose	a-b	156.71, -120.61	160, -130	33,-151 <sup>b</sup> ( 104,-113) <sup>c</sup>
	b-c	148.86, -103.48	100, -145	110,-120 ( 91,-180)
	c-d	137.38, -101.33	100, -145	121,-107 (111,-117 )

<sup>a</sup> Ref. 17. <sup>b</sup> Ref. 18. <sup>c</sup> PDB database 1GAH.

**Table 5.7.** Comparison of the predictions of the new Karplus-type equations (Chapter Four) with that of the three previous equations of Mulloy *et. al.*<sup>15</sup> and Tvaroska *et. al.*<sup>14</sup> and Cloran *et. al.*<sup>16</sup> with the experimentally measured values of the vicinal coupling constant,  $^3J_{\text{HC}}$ , in Hz. The results are averages of the 2000 data points from molecular dynamics simulations that were performed starting from constraint-minimized solution structures.

	Maltose	Acarbose	
	$\phi$ a-b	$\phi$ b-c	$\phi$ c-d
Mulloy <i>et. al.</i> <sup>15</sup>	4.39	2.33	3.21
Tvaroska <i>et. al.</i> <sup>14</sup>	4.54	2.38	3.30
Cloran <i>ET. AL.</i> <sup>16</sup>	5.31	2.50	3.70
<b>Equation 6</b> (AM1- $\phi\psi$ ) (This work)	5.24	2.83	3.86
<b>Equation 4</b> (AM1*) (This work)	4.84	2.65	3.59
Experiment (This work)	4.88	2.60	3.60

were not consistent with a single constraint minimized solution conformation, but were found to be more consistent with average of conformations in solution from molecular dynamics simulations starting from the constraint minimized solution conformation. In all the cases the vicinal proton-carbon coupling constants predicted by the new Karplus-type equations showed better agreement with the measured coupling constants as compared to the three previous Karplus-type equations (**Table 5.7**).

## **Conclusions**

The solution conformations and dynamics of maltose, reduced maltotetraose, and acarbose were studied by a combination of NMR spectroscopy, molecular mechanics calculations and molecular dynamics simulations. The results of the rule-based method (Chapter two) and the new Karplus-type equations (Chapters Three and Four) were compared with previous methods and equations. The results indicate that the rule based prediction method can give a reasonable prediction of the conformations of oligo- and polysaccharides. The new Karplus-type equations gave results in very good and better agreements with experiment than the previous equations. The greater flexibility of acarbose as compared to maltose and reduced maltotetraose was evident. This flexibility might be one of the important factors for the ability of acarbose to bind to many enzymes.

## **Acknowledgements**

This work was supported by the Michigan State University Research Excellence Fund. We thank Professor Jack Preiss, Department of Biochemistry, Michigan State University for generously supplying acarbose.

## References

1. Rademacher, T. W.; Parekh, R. B. & Dweck, R. A. *Ann. Rev. Biochem.* **1988**, *57*, 785.
2. Phillips, M. L.; Nudelman, E.; Gaeta, F. C. A.; Perez, M.; Shingel, A. K.; Hakamori, S. ; Paulson, J. C. *Science* **1990**, *250*, 1130.
3. Mohammed, H.; Hollingsworth, R. I. *J. Org. Chem.* **2001** (accepted for publication).
4. Ernst, R., Bodenhausen, G., & Wokahun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimension*, **1987** pp 431-440, Oxford University Press, Oxford.
5. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355.
6. Bax, A.; Subramanian, S. *J. Magn, Reson.* **1986**, *67*, 565.
7. Sangers, J. K., & Hunter, B.K. *Modern NMR Spectroscopy*, **1993** pp 193-197, Oxford University Press, Oxford.
8. Krishnamurthy, V. V. *J. Magn. Reson.* **1996**, *121*, 33.
9. Mayo, S. L.; Olafson, B. D.;Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897.
10. Molecular Simulations, Inc. Waltham, MA 02154.
11. Wang, Y.; Hollingsworth, R. I. *Biochemistry* **1996**, *35*, 5647.
12. Bock, K.; Pederson, H. *Carbohydr. Res.* **1984**, *132*, 142.
13. Bock, K.; Meldal, M.; Refn, S. *Carbohydr. Res.* **1991**, *221*, 1.
14. Tvaroska, I; Hricovini, M; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359.
15. Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988**, *184*, 39.

16. Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1999**, *121*, 9843.
17. Quingley, B. I.; Sarko, A.; Marchessault, R. H. *J. Am. Chem. Soc.* **1970**, *92*, 5834.
18. Brzozowski, A. M.; Davis, G. J. *Biochemistry* **1997**, *36*, 10837.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02112 6135