

CARBOHYDRATES AND AMINO ACIDS: INFRARED MULTIPLE PHOTON
DISSOCIATION SPECTROSCOPY AND DENSITY FUNCTIONAL THEORY
CALCULATIONS

By

CESAR S. CONTRERAS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2008

© 2008 Cesar S. Contreras

To my parents, brothers and sisters, and nieces and nephews

ACKNOWLEDGMENTS

First and foremost, I would like to thank my research advisor, Dr. John R. Eyler. His strong work ethic has encouraged me to work harder throughout the degree. I would like to thank Dr. Eyler for all the great opportunities he has provided, especially for the great privilege to travel and conduct research in The Netherlands. I want to acknowledge my Ph.D. committee members: Dr. Martin T. Vala, Dr. David E. Richardson, Dr. Gar Hoflund and Dr. Rodney J. Bartlett. I would like to thank them for their feedback in the dissertation process and providing clearer explanations of the underlying fundamental knowledge related to this work. Dr. Bartlett has also provided the computational resources needed to complete the work in this thesis and his generosity is greatly appreciated. I would like to thank Dr. David H. Powell for allowing the use of the FTICR-MS in the Mass Spectrometry Services Laboratory to conduct the experiments on the rubidium cation-attached monosaccharides. His willingness to allow the research to continue into the long hours of the night is very much appreciated. Dr. Kathryn R. Williams helped me throughout the writing of the dissertation and her advice and comments were one of the motivations that drove me forward. Discussions with Dr. Nicole A. Horenstein and her course in carbohydrate chemistry provided the basis that I needed and used to understand the results obtained from this work. Although the main project in collaboration with Dr. Horenstein was incomplete by the time of writing this dissertation, her excitement toward carbohydrate chemistry was another driving force toward accomplishing my goals.

I would like to thank Dr. Jos Oomens and Dr. Jeff Steill and the FOM-Institute for Plasma Physics Rijnhuizen, Nieuwegein, The Netherlands for allocating beam time and the use of the FTICR-MS in that institute. I learned a lot from interacting with everyone at FOM and that level of collaboration will be remembered whether I go. I want to also acknowledge Dr. Nicolas C. Polfer for starting the work on the phenylalanine analogs and glycosamine derivatives. His

preparation for life after graduate school and post-doctoral work leading to a faculty position at this major university has shown me what it takes to fulfill my dreams.

I greatly appreciate the friendship of Wright L. Pearson. Many nights we stayed up trying to troubleshoot and take spectra with the OPO laser. The long discussions from anything in politics to the most obscure jazz compositions and of course chemistry will long be remembered, as they helped to stimulate my mind to learn about many more different subjects outside of chemistry. I want to mention my good friend Sabri Alkis, that without his persistent of wanting to have “fun” would have kept me in the lab and I would not have been able to appreciate all that Gainesville and Florida have to offer. I would like to thank all the Eyler and Polfer group members for the many discussions and questions that have kept me on my toes and have pushed me to continue reading and learning about the many mass spectrometric techniques.

I would like to thank my family and parents that instilled in me a hard work ethic and an ability to persist even when the future seems uncertain. My family is always part of everything I do and this dissertation is dedicated to them. There are countless other people that helped throughout the Ph.D. degree and to all of you I am thankful.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	9
LIST OF FIGURES	10
ABSTRACT	15
CHAPTER	
1 INTRODUCTION	17
Saccharide Chemistry	17
Carbohydrate Structure and Terminology	17
Syntheses and Reactions	22
Biological Significance	22
Peptide Chemistry	24
Amino Acids	24
Phenylalanine	24
Protein Structure	25
Objective of Research	26
Overview	28
2 MATERIALS AND METHODS	37
Mass Spectrometric Approaches	37
Fourier Transform Ion Cyclotron Resonance Mass Spectrometry	39
Tandem Mass Spectrometry	52
Infrared Multiple Photon Dissociation	53
Carbon Dioxide Laser	56
Optical Parametric Oscillator	58
Free Electron Laser	60
Implementation of IRMPD with FTICR-MS	62
3 THEORETICAL CALCULATIONS	71
Molecular Modeling	71
Theoretical Background	72
<i>Ab initio</i>	75
Force fields	77
Geometry optimization	79
Vibrational analysis	80
Molecular Dynamics	81
Simulated annealing	81

	Conformational searching	83
4	STRUCTURE DETERMINATION OF PHENYLALANINE ANALOGS	85
	Introduction.....	85
	Experimental.....	87
	Sample Preparation.....	87
	Instrumentation.....	87
	Computational Details.....	88
	Results and Discussion	89
	N-acetylphenylalanine.....	89
	Hydrogen/Deuterium Exchange Experiments.....	91
	AcPhe	91
	AcPheOMe	91
	AcPheGlyOMe	92
	Conclusions.....	93
5	DIFFERENTIATION OF ANOMERS OF D-GLUCOSAMINE AND D- GALACTOSAMINE.....	101
	Introduction.....	101
	Experimental.....	103
	Sample Preparation.....	103
	Instrumentation.....	103
	Computational Details.....	104
	Results and Discussion	105
	Infrared Spectra	105
	Glucosamines	105
	Galactosamines.....	107
	Fragmentation Patterns.....	108
	Conclusion.....	109
6	DIFFERENTIATION OF O-METHYLATED GLYCOSIDE ISOMERS	120
	Introduction.....	120
	Experimental Techniques	121
	Chemicals	121
	Instrumentation.....	122
	Laser Setup.....	122
	Experiment	123
	Computational Details.....	124
	Results and Discussion	125
	FELIX.....	125
	OPO	126
	Glucosides	126
	Galactosides	129
	Conclusions.....	130

7 CONCLUSIONS AND FUTURE WORK.....	144
LIST OF REFERENCES.....	149
BIOGRAPHICAL SKETCH.....	158

LIST OF TABLES

<u>Table</u>	<u>page</u>
4-1. Interaction distance between sodium cation and oxygen atoms on the carbonyls of the carboxyl and on the N-acetyl functional groups for calculated conformers A - D.	94
4-2. Summary of results from H/D exchange experiments.....	94
5-1. Glycosamine isomer structures obtained from the corresponding conformational searches.....	111

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Structure of kojibiose (2-O- α -D-glucopyranosyl-D-glucopyranose).....	29
1-2 Aldoses.....	29
1-3 Fischer projection of D-glucose.....	30
1-4 D-aldoses.....	30
1-5 Common ketoses involved in biological processes	31
1-6 Pyranose and furanose ring formation in equilibrium	31
1-7 Six-membered ring structures of the α -anomers of D-hexoses.....	32
1-8 Ring conformations, for furanoses and pyranoses.....	33
1-9 Disaccharide formation by glycosylation	33
1-10 The polysaccharids amylose and cellulose	34
1-11 Amino acid.....	34
1-12 The 20 amino acids used in proteins.....	35
1-13 General structure of a tripeptide	35
1-14 Process of inverting a β -galactoside to the corresponding α -galactoside by using an inverting glycosidase.	36
2-1 Electron impact ionization and chemical ionization mass spectrum of D-glucose	63
2-2 Initial trajectory of the velocity of the ion	63
2-3 Motion of the ion in a uniform magnetic field.....	64
2-4 Motion of an ion (in a uniform magnetic field) as its cyclotron motion is excited by an alternating electric field.....	64
2-5 Detection of ions in an FTICR-MS.....	65
2-6 Cyclotron, trapping and magnetron motions	65
2-7 Electrospray mist	66
2-8. Ion transfer optics.	66

2-9	Two common Penning traps	66
2-10	The FTICR mass spectrometer at the FOM-Institute	67
2-11	Infrared spectrum of a crystalline sample of β -D-glucopyranose	67
2-12	Initial absorption of photon energy from the ground state to an excited vibrational level and the subsequent intramolecular vibrational relaxation process.....	68
2-13	The CO ₂ laser cavity	68
2-14	Vibrational energy levels of the CO ₂ and N ₂ molecules involved in the lasing mechanism.	69
2-15	The IRMPD spectrum of Mn(CO) ₄ CF ₃ ⁻ , using a tunable CO ₂ laser.....	69
2-16	An OPO laser.....	70
2-17	Free electron laser schematic	70
4-1	Structure of the amino acid phenylalanine.....	94
4-2	Phe analogs studied in this work.....	95
4-3	Dihedral angles defined for the conformational search calculations of AcPhe, AcPheOMe and AcPheGlyOMe	95
4-4	Lowest energy conformers of sodium cationized AcPhe found from conformational search calculations.	96
4-5	Comparison of experimental IRMPD spectrum (red line) of AcPhe with calculated spectra of the theoretically determined lowest energy conformers A-D (blue stick spectra).....	97
4-6	Assignment of spectral bands in the IRMPD spectrum of AcPhe using calculated bands for conformer A.....	98
4-7	IRMPD spectrum of sodium cation-attached AcPhe following solution phase HDX (in red) and calculated spectrum of the doubly deuterated conformer A (in blue).....	98
4-8	Comparison of the singly deuterated sodium cation-attached AcPhe IRMPD spectrum with that of calculated conformer A.....	99
4-9	Overlap of IRMPD spectra for undeuterated and singly-deuterated sodium cation-attached AcPhe.....	99
4-10	Experimental IRMPD spectra and theoretically calculated spectra of phenylalanine analogs	100

5-1	Glycosamine isomers	112
5-2	Side and top views of lithium cationized β -GlcNac	112
5-3	Rotation about C-2 and N, shown with a red arrow.	113
5-4	IRMPD spectra of α -GlcNac (black) and β -GlcNac (red) bound to a lithium cation, showing the 10 cm^{-1} shift of the carbonyl stretch frequency.	113
5-5	Calculated lowest energy structures of lithium cationized α -GlcNac and β -GlcNac, using B3LYP/6-31+G(d).	114
5-6	Comparison of experimental IRMPD and calculated infrared spectra of lithium cationized α -GlcNac	114
5-7	Potential surface scan for rotation about (C-2)-N bond for α -GlcNac.	115
5-8	Potential surface scan for rotation about the (C-2)-N bond for β -GlcNac.	115
5-9	Comparison of experimental IRMPD spectrum (red) to the calculated infrared spectrum (blue) of lithium cationized β -GlcNac.	116
5-10	IRMPD spectra of lithium cationized α -GalNac and β -GalNac, showing the 11 cm^{-1} shift for the carbonyl stretch bands	116
5-11	Calculated infrared spectra of lithium cationized α -GalNac and β -GalNac, indicating a 29 cm^{-1} shift in the C=O stretches.....	117
5-12	Lowest energy structures of lithium cation-attached galactosamine anomers from B3LYP/6-31+G(d)	117
5-13	Potential surface scan for α -GalNac.....	118
5-14	Potential surface scan of β -GalNac.....	118
5-15	Fragmentation ion abundances as a function of laser wavelength for all four isomers... 119	
5-16	Possible fragmentation producing the observed MS/MS spectra of glycosamine isomers	119
6-1	The FELIX-generated IRMPD spectra (600 – 1700 cm^{-1}) of all four glycoside isomers, α Glc, β Glc, α Gal and β Gal.....	131
6-2	The O-methylated glycoside isomers, A) α Glc, B) β Glc, C) α Gal and D) β Gal shown in the ${}^4\text{C}_1$ chair conformation.....	132
6-3	The OPO optics setup	132

6-4	Experimental sequence for obtaining spectra of the O-H stretching region of the rubidium cation-attached glycosides using OPO laser irradiation.....	133
6-5	Experiment sequence for the two laser experiment	133
6-6	Monosaccharide α -D-glucoside.....	134
6-7	Lowest energy conformers of Glycosides.....	134
6-8	Comparison of theoretical and experimental IRMPD spectra for glycoside isomers.....	135
6-9	Calculated spectra for the second lowest energy conformations of D-glucosides anomers from the conformational search.....	135
6-10	IRMPD spectrum of rubidium cation-attached β Glc.....	136
6-11	Comparison of the IRMPD spectrum of β Glc in the C-H stretch region with different irradiation methods	136
6-12	Comparison of calculated spectrum for conformer A (dark blue) and IRMPD spectrum (red) for rubidium cation-attached β Glc.	137
6-13	Experimental spectrum (red) and calculated spectrum for conformer B (green) of rubidium cation-attached β Glc.	137
6-14	Spectral match between experiment (red) and calculated conformer C of rubidium cation-attached β Glc	138
6-15	Spectral comparison of experiment (red) and calculated conformer D (gray) for rubidium cation-attached β Glc	138
6-16	Overlap of all four conformer spectra with IRMPD spectrum for rubidium cation-bound β Glc.....	139
6-17	Calculated conformers for β Glc at the B3LYP/6-31+G(d) level of theory	140
6-18	Comparison of IRMPD spectra for α Glc (blue) and β Glc (red), clearly showing the band at 3637 cm^{-1} is missing for α Glc.....	141
6-19	Comparison of IRMPD and calculated spectra for the lowest energy conformer of α Glc.	141
6-20	Overlap of calculated spectra for conformer A and conformer B with IRMPD spectra of α Glc.....	142
6-21	Conformers of α Glc that correlate with IRMPD spectrum	142

6-22	Comparison of preliminary IRMPD spectra for α Gal (green) and β Gal (yellow) in the O-H stretch region.....	143
6-23	Overlap of IRMPD spectra of all four isomers in the O-H stretch region.....	143

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CARBOHYDRATES AND AMINO ACIDS: INFRARED MULTIPLE PHOTON
DISSOCIATION SPECTROSCOPY AND DENSITY FUNCTIONAL THEORY
CALCULATIONS

By

Cesar S. Contreras

December 2008

Chair: John R. Eyler
Major: Chemistry

In addition to acting as repair agents, stabilizing protein folding or early defense systems in cellular systems, saccharides are especially important in energy storage and enzymatic reactions of proteins. All these areas of research require knowledge of the saccharide structure.

Spectroscopic studies of monosaccharides and amino acids were undertaken to better understand structural conformations in the gas phase. Infrared spectra were obtained by using Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) in conjunction with infrared multiple photon dissociation (IRMPD).

Sodiated phenylalanine analogs were subjected to H/D exchange before their IRMPD spectra were taken. The gas phase H/D exchange experiments of N-acetylphenylalanine indicate that of the two possible locations for exchange to occur, the O-H hydrogen is kinetically favored over the N-H hydrogen. For two larger species, O-methyl N-acetylphenylalanine and N-acetylphenylalanine O-methylglycine, exchange occurred at the N-H site since it was the only one available for exchange, but the D for H substitution only took place in solution and not in the gas phase. Theoretical calculations showed that the phenylalanine analogs, although of different

size, have relatively similar structural features. The sodium cation is predicted to interact with the phenyl ring and also bind to the carbonyl oxygens.

In a second project, IRMPD spectra of N-acetylglycosamines showed that the frequency of the CO stretch was indicative of the particular glycosamine conformation. A band shift of about 10 cm^{-1} was seen between the anomers, α -D-methylglucosamine and β -D-methylglucosamine, while an 11 cm^{-1} shift was seen for the galactosamine anomers. Calculations indicate that the O-methyl group's position (α and β , or axial and equatorial, respectively) and its close proximity to the N-acetyl group cause the orientation of the carbonyl to change in order to minimize steric hindrance, and therefore a band shift is observed for the CO stretch.

A third project involved the setup and use of an optical parametric oscillator (OPO) laser to obtain IRMPD spectra of rubidium cation-bound glycosides. DFT calculations and experimental spectra showed that the anomers of D-glucoside and D-galactoside all have differing hydrogen bonding and locations of rubidium binding, consequently showing distinct spectra in the O-H stretch region.

CHAPTER 1 INTRODUCTION

Saccharide Chemistry

Carbohydrates are everywhere on earth, and they are vital components of our everyday lives. It has been estimated that cellulose comprises 50% of the earth's biomass.¹ In the diet, sugars and starches are metabolized to produce energy, which is stored as adenosine triphosphate, which also contains a sugar as part of its structure. Carbohydrates are essential to the food industry, which uses large amounts of starch, sweet gums and mono- and oligosaccharides. Carbohydrates are also important in textiles, which are largely dependent on cellulose containing materials. The pharmaceutical industry uses carbohydrates for the preparation of antibiotics and intravenous solutions and as components of pills and capsules.

The importance of sugars was recognized by early organic chemists, who studied sugars and their derivatives as early as the 1870s, a decade before Emil Fischer began his 20-year study to classify the configurations of sugars.² But although carbohydrate chemistry is one of the oldest fields in chemistry, interest in the field continues to increase and this has been largely due to the sheer number and complexity of saccharide types.

Carbohydrate Structure and Terminology

To better understand the intricacies and complexities of carbohydrates, the naming conventions used for the isomeric structures of saccharides will be introduced, focusing mainly on the monosaccharide units. The naming convention serves two main purposes: to allow conversion of the IUPAC nomenclature to a more readable format, and to name polymers of monosaccharide units (i.e., polysaccharides) quickly. For example, the disaccharide shown in Figure 1-1 has the cumbersome IUPAC name (3R,4S,5S,6R)-6-(hydroxymethyl)-3-((3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yloxy)tetrahydro-

2H-pyran-2,4,5-triol, but this can be shortened to 2-O- α -D-glucopyranosyl-D-glucopyranose.

Frequently encountered disaccharides also have common names; e.g., 2-O- α -D-glucopyranosyl-D-glucopyranose is called kojibiose.

Aldoses and ketoses. The term *carbohydrate* originated in the sugars known in the 1870s which all had the empirical formulas $C_x(H_2O)_y$, and were thought to be hydrated carbons.³ Today, carbohydrates encompass monomers, oligomers and polymers that are derived from monosaccharides.⁴ Monosaccharide units have the empirical formula CH_2O and are either *aldoses*, which contain both aldehyde and alcohol functionalities, or *ketoses*, with both keto and alcohol groups. The simplest monosaccharide aldose is glyceraldehyde, which contains three carbons (Figure 1-2). Aldoses with 4, 5, and 6 carbons are called aldotetroses, aldopentoses, and aldohexoses, respectively. Monosaccharides can have up to ten carbons, but the most common monosaccharides are pentoses and hexoses.

Fischer projections. Linear chain monosaccharides are usually represented by way of the Fischer projection,⁵ (Figure 1-3). In a Fischer projection, the carbon backbone is oriented vertically with the horizontal groups pointing out the front of the page. The aldehyde group CHO at C-1 is oriented at the top of the vertical chain, and the other carbons are numbered sequentially along the chain from this position. Monosaccharides can have many stereoisomers, since all carbons that are secondary alcohols have four different functional groups and are thus chiral centers. The chain form of D-glucose, whose IUPAC name is (2S,3R,4S,5S)-2,3,4,5,6-pentahydroxyhexanal, has four chiral centers, (Figure 1-3).

D and L sugars. Historically, the designation D or L has referred to the absolute configuration at the highest numbered chiral carbon (C-5 in an aldohexose). In the Fischer projection, if the hydroxyl group on C-5 points to the right, the sugar is designated D. If the OH

points to the left, it is an L sugar. (Note that this should not be confused with d and l , which indicate the direction of rotation of plane-polarized light: d = dextrorotatory = CW rotation; l = levorotatory = CCW rotation.) According to the newer Cahn-Ingold-Prelog convention,⁶ a D sugar has the R configuration at C-5, and S corresponds to L. The D forms of all aldoses through the aldohexoses are shown in Figure 1-4.

Although the majority of monosaccharides are aldoses, a few ketoses are important in biological systems (Figure 1-5). The ketohexose D-fructose links with D-glucose to form the disaccharide sucrose. The L-fucose is a structural component of plant cell walls⁷ and L-rhamnose is a component of bacterial cell membrane, but is also found in a variety of plants.⁸

Cyclic forms. Because aldoses and ketoses contain both carbonyl and hydroxyl functionalities, they readily form cyclic hemiacetals and hemiketals. shows the hemiacetals of D-glucose, which forms the six-membered ring structure by reaction of the C-5 oxygen with C-1, as well as the five-membered ring structure by reaction of the C-4 oxygen with C-1 (Figure 1-6). Each reaction gives rise to a new chiral center at C-1. If the functional group at C-1, designated the anomeric carbon, is in the axial position, the ring structure is designated the α -anomer. Likewise, if the functional group at C-1 is in the equatorial position, the ring structure is designated the β -anomer. The five-membered ring is related to tetrahydrofuran and is called a furanose, while the six-membered ring is called pyranose, after tetrahydropyran. In an aqueous solution, the ring-closed forms are favored over the open-chain form, the abundance of the latter being almost negligible. An equilibrium, called *mutarotation*, is established between the ring structures and the open-chain form, (Figure 1-6). For D-glucose, the equilibrium mixture in an aqueous environment is 38% α -pyranose, 62% β -pyranose, 0.1% α -furanose and less than 0.2% β -furanose, with a negligible concentration of the open-chain form.⁴ This equilibrium ratio is

unique to the D-glucose, and varies for the other monosaccharides, being dependent on the ring structure, functional groups, solvent environment and the conformation of the particular D-aldose.³ For example, an isomer of D-glucose, mannose, exists predominantly as the α -pyranose.⁴ Because of the mutarotation equilibrium, the aldehyde group of the open-chain form can be oxidized (positive Tollens' and Benedict's tests). Thus, all monosaccharides are classified as reducing sugars. Although Fischer projections are useful for visualization of open-chain sugars, ring structures, particularly in the chair conformation, are preferred for hemiacetal and hemiketal forms, (Figure 1-7) for the α -anomers of the D-pyranoses .

Ring conformations. The ring structures of hexoses and pentoses are flexible and exist in many conformations, (Figure 1-8) For furanoses, the ring can be planar and form an equilateral pentagon shape, or envelope (E), or it can be slightly bent and form a twisted pentagon (T). Six-membered rings are more flexible and have many more conformations. For D-glucose, the chair conformation (C) is the most stable in an aqueous solution. It can have two forms. The one shown in Figure 1-8 is denoted 4C_1 , where C-4 is positioned above the plane of the ring and C-1 is positioned below the plane of the ring. The other is 1C_4 and has the opposite orientation, with C-4 below and C-1 above the plane of the ring. The boat conformation (B), can have both C-1 and C-4 above the plane, in which case it is denoted ${}^{1,4}B$, or both C-1 and C-4 below the plane and in this case it is denoted $B_{1,4}$. The chair or boat conformations can be perturbed, forming the half-chair (H) and skewed (S) conformations, respectively.

Polysaccharides. Carbohydrates undergo glycosylation reactions to form molecules that have multiple monosaccharide units linked together. Glycosylation can be understood as the displacement of the hydroxyl group on C-1, the glycosyl donor, by the hydroxyl group, the glycosyl acceptor, of another monosaccharide (Figure 1-9). The disaccharide can be subjected to

further glycosylation to continue the polymerization process. As part of the disaccharide, the donor is called the non-reducing monosaccharide, since it is now bound to another monosaccharide at the anomeric carbon, and cannot undertake ring opening to expose the aldehyde group for oxidation. If the glycosidic bond does not involve the anomeric carbon of the acceptor unit, the disaccharide can still act as a reducing sugar. The different glycosidic linkages are designated (1→1), (1→2), (1→3), (1→4) and (1→6), where the first number in the parentheses indicates the linkage position for the non-reducing end and the second number is the linkage position of the acceptor monosaccharide. The orientation of the linkage depends on the configuration of the anomeric carbon in the donor, further increasing the number of isomers possible when forming polysaccharides through glycosylation.

Oligosaccharides have 2-10 monosaccharides, with the number of monomer units often explicitly stated (i.e., disaccharides, trisaccharides, etc.). There are two classifications of oligosaccharides: true oligosaccharides, which are made up solely of simple monosaccharide units, and conjugate oligosaccharides, which have monosaccharides linked to non-saccharides, for example peptides and lipids. Oligosaccharides can also be grouped as reducing and non-reducing sugars, depending on the linkage to the last donor. The degree of polymerization is small enough that oligosaccharides continue to be soluble in water, similar to monosaccharides.

Polysaccharides have more than 10 monomer units. Similar to oligosaccharides, they are grouped as either true or conjugate polysaccharides. Since the degree of polymerization can be much higher for polysaccharides and there is a higher probability of having more than one type of monosaccharide unit, true polysaccharides are further divided into two subclasses: the homogeneous class, with only one type of monosaccharide as the repeating unit, and the heterogeneous class with more than one type monosaccharide. There are also linear or branched

polysaccharides, the latter being common in glycolipids attached to plasma membrane, or when carbohydrates are attached to surfaces. The solubility decreases as the size of the polymer increases, and many polysaccharides are insoluble in water.

Syntheses and Reactions

One of the major goals of a saccharide synthetic chemist is the production of a select type of saccharide by a “one-pot synthesis,” in which successive chemical reactions are conducted in one reactor, avoiding purification and separation of intermediates. For carbohydrates, one-pot syntheses are not common,⁹⁻¹¹ because the multiple stereocenters provide numerous reaction sites. Efforts to develop new synthetic approaches have also been hindered by the paucity of methods to determine the structures of products which are often formed in very low concentrations.

The molecular complexities of carbohydrates also cause problems in reactions to form derivatives and polymers. Identification of general patterns of reaction has been a difficult task, because product structures vary due to influences of the solvent, temperature, and functional group type, as well as ring size and strain and intra- and intermolecular forces like hydrogen bonding and interactions with metallic species.¹²

While solution phase and wet synthesis research has provided a wealth of information on saccharide structure, it is still unknown what major factors influence the formation and stability of one isomeric form relative to another. It should be apparent that factors controlling the 3-dimensional structure of carbohydrates must be fully studied in order to understand the many processes they mediate.

Biological Significance

Until recently, the role of sugars in biological systems was thought to be limited to energy storage and structural support. However, increased research on glycopeptides and glycoproteins

has indicated that carbohydrates also play a role in biological communication events, in processes such as egg fertilization, microbial infection, inflammation, cancer growth and diabetes.

Glucose is found in proteins synthesized in mammalian cells and is covalently bound to different amino acids in the peptide structure. For example, carbohydrates can be found covalently bonded to the amide nitrogen of asparagines (N-linked glycoproteins) or to the oxygen atom in serine and threonine (O-linked glycoproteins). Carbohydrates bound to proteins affect solubility, protein folding and protect against protein degradation from enzymatic processes. Carbohydrates can also bind to lipids (glycolipids), forming components of the plasma membrane found in all vertebrate cells. More than 400 glycolipids are known, while the most common glycolipids in vertebrates are formed with only 7 distinct monosaccharides. Deoxyribonucleic acid (DNA) is a polymer that has the ribose sugar (2-deoxy-D-erythrose-pentofuranose) as part of its repeating nucleotide units.

Polymers. Further importance of carbohydrate structure in biological function is observed in polysaccharides of glucose. In amylose, α -D-glucoses are connected by $\alpha(1\rightarrow4)$ glycosidic bonds (Figure 1-10). Cellulose, another polysaccharide of D-glucose, contains $\beta(1\rightarrow4)$ glycosidic bonds. Each polysaccharide has a distinct function. Amylose is the soluble component of starch and is a major food source, while cellulose is the structural material in plants. Amylose, because of the $\alpha(1\rightarrow4)$ glycosidic bonds, forms a helical structure, whereas cellulose forms a ribbon-like structure. Cellulose is insoluble in water, and it cannot be digested by humans.

Adding a substituent functional group also changes the properties of the carbohydrate. Chitin is a polysaccharide of D-glucose and has the same $\beta(1-4)$ linkages as cellulose, but has an

N-acetyl group at the C-2 of each glucose unit. Similar to cellulose, chitin provides molecular structural support in cells and forms the hard exoskeleton of insects and shellfish. These and many other discoveries have sparked a renewed interest in carbohydrate chemistry, and chemists having a variety of specialties are trying to solve these glycobiological problems. For all of these research areas, the carbohydrate 3-dimensional structure is important.

Peptide Chemistry

Proteins, along with carbohydrates, nucleic acids, and lipids, are one of the four main groups of molecules that are important in cellular function. Protein chemistry has experienced a growing interest, and mass spectrometric studies have made major contributions in this area.

Amino Acids

The building blocks of proteins are α -amino acids. The designation alpha (α) refers to the point of attachment of the amino group ($-\text{NH}_2$), which is always on the carbon adjacent to the carboxyl group ($-\text{COOH}$) (Figure 1-11). Although often shown as uncharged species (Figure 1-11A), amino acids in biological fluids exist as zwitterions (Figure 1-11B). Twenty amino acids are found in proteins, and each differs by the functional group R (Figure 1-12).

Phenylalanine

Because of its importance in this research, the amino acid phenylalanine will be described in further detail. Phenylalanine (Phe) is one of the essential amino acids, meaning that it is not synthesized by humans and must be included in the diet. Phenylketonuria, a genetic disorder affecting infants and young children, prevents the normal metabolism of phenylalanine. Instead of the normal product, tyrosine, Phe is converted to phenylpyruvate, which interferes with energy release, leading to mental retardation.¹³

Since phenylalanine has a nonpolar side group, it forms part of the hydrophobic region in the interior of a globular protein. Similar to other amino acids, Phe occurs as a zwitterion in

solution. Previous results^{14,15} have indicated that Phe exists in the unionized form in the gas phase. However, recent mass spectrometric studies have shown that phenylalanine may also exist as a zwitterion in the gas phase, if it is bound to specific metals. Gas-phase studies on sodium cation-attached amino acids, including phenylalanine, indicated that rates of deuterium/hydrogen exchange reactions were very fast.¹⁶ For phenylalanine, the exchange rates were $1.12 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1} \text{ molecule}^{-1}$ using CD_3OD and $0.16 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1} \text{ molecule}^{-1}$ using D_2O . These high exchange rates were attributed to the presence of a protonated α -amino group, indicating that the zwitterion form was present.

Protein Structure

In proteins, amino acids are joined by amide linkages (Figure 1-13), often called peptide bonds. Small polymers (up to about 30 amino acid residues) are called polypeptides, which, like polysaccharides, often carry the di-, tri-, etc. prefixes. Proteins are polypeptides with more than 30 amino acids.

The many complexities of carbohydrate structure and conformation were described earlier in this chapter. In proteins, differences arise from the number and sequence of the 20 amino acids (and the presence of $-\text{S}-\text{S}-$ linkages between cysteine residues), referred to as the primary structure. Structural patterns within segments of the protein chain constitute the secondary structure, which depends strongly on the hydrogen bonding between nearby amide groups. The two common types of secondary structure are the α -helix and the β -pleated sheet. The overall shape of the protein, whether it is extended or globular, is called the tertiary structure.

The ability to assemble proteins containing a vast number of amino acid sequences gives an organism the flexibility to tailor proteins for specific functions. Of particular importance in this research is the transport of Na^+ and K^+ ions by specially designed proteins that bind one specific alkali metal and allow its passage through the cell membrane. These ions are essential

to regulatory function within the cell. Another very specialized group of proteins, the Carbohydrate Processing Enzymes are specifically designed for the synthesis of carbohydrates in a single step without the use of protecting groups.⁴ There are two main types of these enzymes; one type makes glycosidic bonds (glycosyltransferases), and the second type breaks glycosidic bonds (glycosidases). Both enzyme types take advantage of binding sites that fit only a specific sugar type. The glycosyltransferase enzymes have two binding sites for the donor and acceptor sugar, respectively. When the enzyme folds together, glycosylation occurs at only one hydroxyl group of the acceptor. The other hydroxyl groups are buried in the binding site and cannot react. The glycosidase enzymes are not as specific, because they only have binding sites for the donor sugar, since the main purpose is to cleave the donor. Figure 1-14 shows a mechanism where a β -galactoside is inverted in the glycosidase binding site to form the corresponding α -galactoside.

Objective of Research

The biological molecules just described have been the subjects of numerous studies in solid and solution phase chemistry. However, elucidation of the 3-dimensional structure in those studies has been hindered by the interactions which occur when the molecule being studied contacts other molecules and ions. Gas phase studies of carbohydrates and amino acids are increasing, driven mainly by the increased sophistication of the gas phase instrumentation developed in recent decades. Gas phase studies are especially useful, because the intermolecular interactions are reduced and are almost negligible in some methods. Therefore the molecule can be studied independent of such interactions, and the intramolecular forces that influence its shape can be studied. Gas phase studies also have the capability of investigating intermolecular interactions one at a time, for example by addition of a metal ion attached to a neutral molecule, so that the direct influence of the interaction on the ion can be studied.

Fourier transform ion cyclotron mass spectrometry (FTICR-MS), in conjunction with ion fragmentation methods, has been used to obtain accurate structural information for gaseous ions. Infrared multiple photon dissociation (IRMPD) has been used by numerous groups to obtain fragmentation patterns of oligosaccharides, proteins and peptides using a fixed-wavelength CO₂ laser, and small ions have been studied using wavelength-tunable CO₂ lasers. A limitation of the latter studies has been the narrow wavelength range of the tunable CO₂ laser (9.1 – 10.9 μm) with numerous gaps where lasing does not occur, resulting in IRMPD spectra obtained with incomplete ir bands. The advent of continuously tunable free electron lasers and the subsequent implementation of them in IRMPD studies has provided the basis for obtaining complete ir spectral bands over a large wavelength range (2 – 250 μm). Recently, another tunable laser, the optical parametric oscillator (OPO), has been used in gas phase studies to probe the conformations of ions. The OPOs in those studies were tunable in the ultraviolet, but newer infrared-tunable OPOs have become available, and the Eyler laboratory, in conjunction with the Mass Spectrometry Services Laboratory at the University of Florida, has implemented an OPO for use with FTICR-MS to obtain IRMPD spectra of carbohydrates. The OPO/FTICR-MS system was used for part of this thesis research, and will be presented in chapter 6.

Even when studied in the gas phase, the complexity of carbohydrates and amino acids is evident in their spectra, and theoretical calculations are crucial to decipher that information. The multiple atoms and flexible nature of amino acids and carbohydrates make them a challenge to model with theoretical methods. Algorithms have been developed to tackle such problems, but often these methods have been used without the understanding of the underlying theoretical framework. It is imperative to understand the theoretical basis in order to obtain accurate and meaningful results.

The objective of the investigations in this thesis was to apply the techniques of both IRMPD-FTICR-MS and theoretical calculations to obtain gas phase structures of amino acid and carbohydrate ions and to differentiate carbohydrate isomer ions. Monosaccharides and mono-peptides were used to study interactions with alkali metal cations. The OPO/FTICR-MS system was implemented to obtain IRMPD spectra in the range from 2700 – 3800 cm^{-1} , the infrared wavelength region containing C-H and O-H stretching bands. The O-H stretches were especially sought after, since the free hydroxyl groups of sugars tend to give strong absorption features in infrared spectra. Theoretical calculations are key to understanding the processes that lead to the differences in the IRMPD of the respective isomers. The calculations are also used to predict how the experimental IRMPD spectral bands shift when analogs of the amino acid phenylalanine are subjected to hydrogen/deuterium exchange.

Overview

The next chapter will describe the experimental methods of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) and Infrared Multiple Photon Dissociation (IRMPD). The principles and mathematical description of FTICR-MS will be presented, as well as the instrumentation used to ionize molecules. The IRMPD mechanism will be summarized and the lasers used as irradiation sources for fragmentation will be described. A general experimental procedure, from ionization to obtaining IRMPD spectra of precursor ions will be introduced. This will be followed by chapter 3, where theoretical calculations involving *ab initio* and classical mechanics methods will be detailed.

Hydrogen/deuterium exchange and subsequent IRMPD experiments of sodium cation-attached phenylalanine analogs will be described in chapter 4. Theoretical calculations were used to predict the band shifts due to H/D exchange and the calculated structures will be discussed. Chapter 5 discusses the differentiation of lithium cation-attached N-

acetylglycosamine isomers through IRMPD spectra in the 600 – 1800 cm^{-1} combined with theoretical calculations. Chapter 6 introduces the earlier work of Valle *et al.*,¹⁷ where IRMPD spectra in the 600 – 1800 cm^{-1} region of rubidium cation-attached glycoside isomers were obtained. Differentiation between the isomers was considered inconclusive. In this work, IRMPD spectra from 2700 – 3800 cm^{-1} of the rubidium cation-attached glycosides are obtained. The properties of the optical parametric oscillator used in these studies are detailed, and the modified experimental sequences to obtain IRMPD spectra are explained. Finally, a conclusion that summarizes the results of the work in chapter 4, 5 and 6 and suggests future work will be presented in chapter 7.

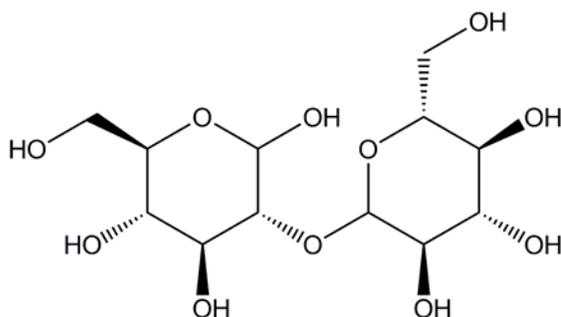


Figure 1-1. Structure of kojibiose (2-O- α -D-glucopyranosyl-D-glucopyranose).

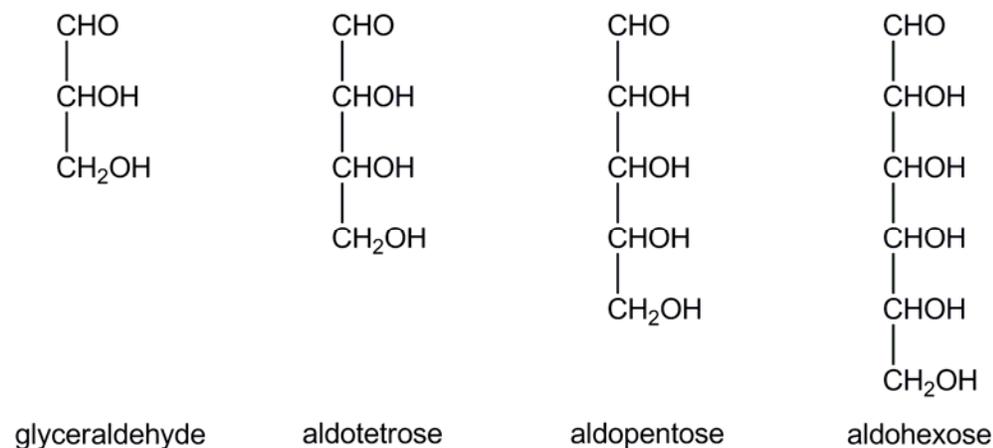


Figure 1-2. Aldoses, illustrating naming as chain size increases.

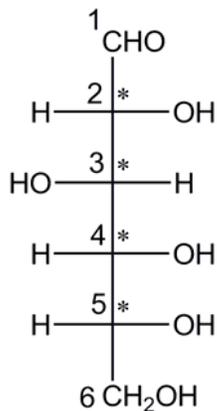


Figure 1-3. Fischer projection of D-glucose with chiral carbons starred.

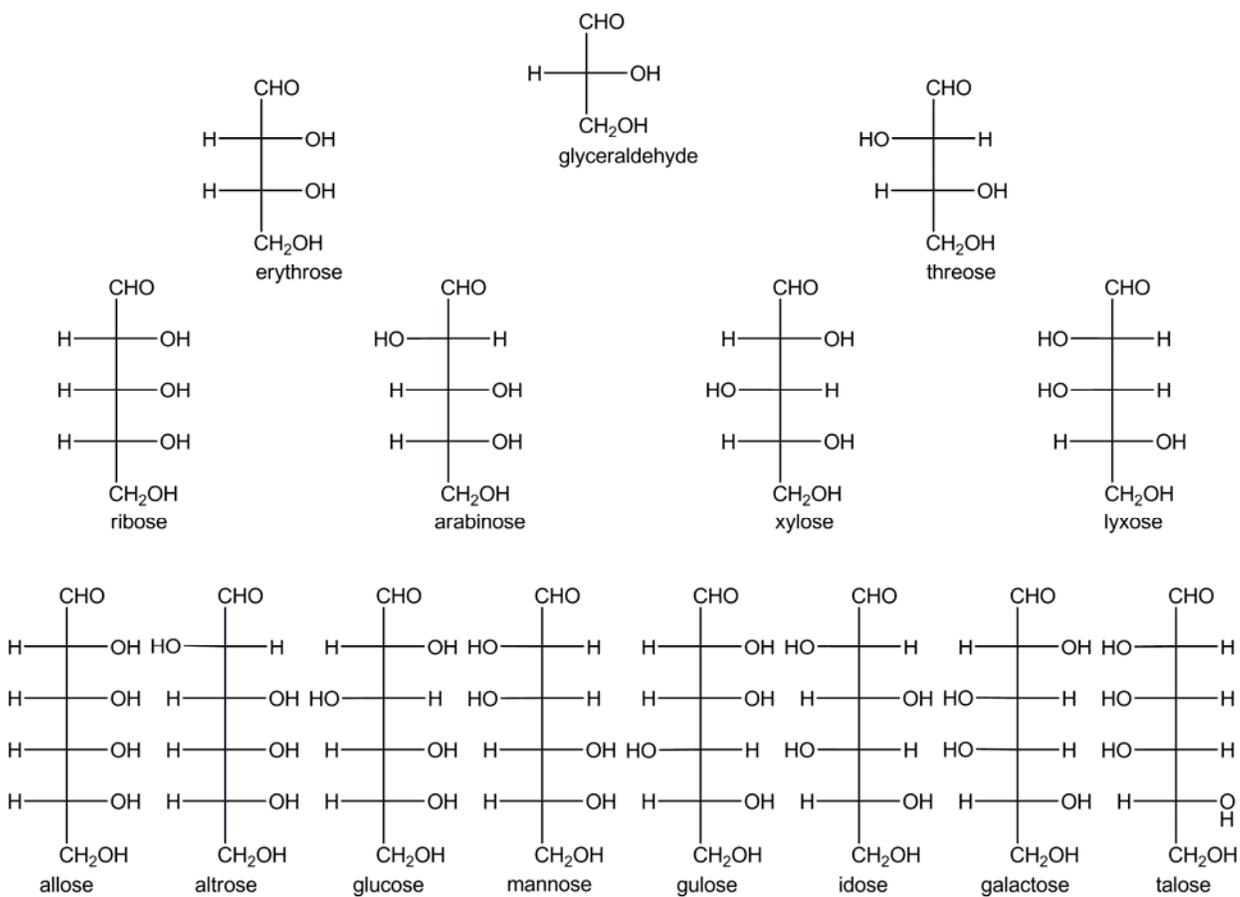


Figure 1-4. D-aldoses

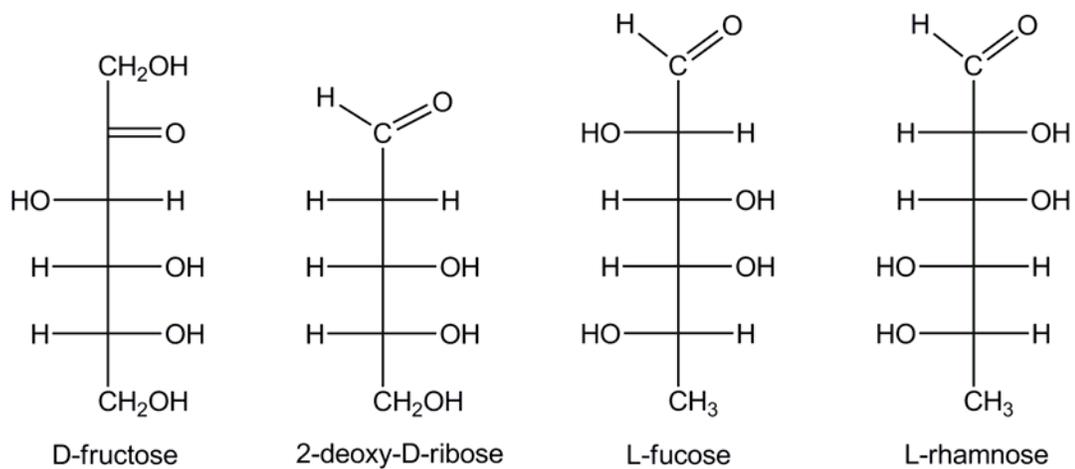


Figure 1-5. Common ketoses involved in biological processes

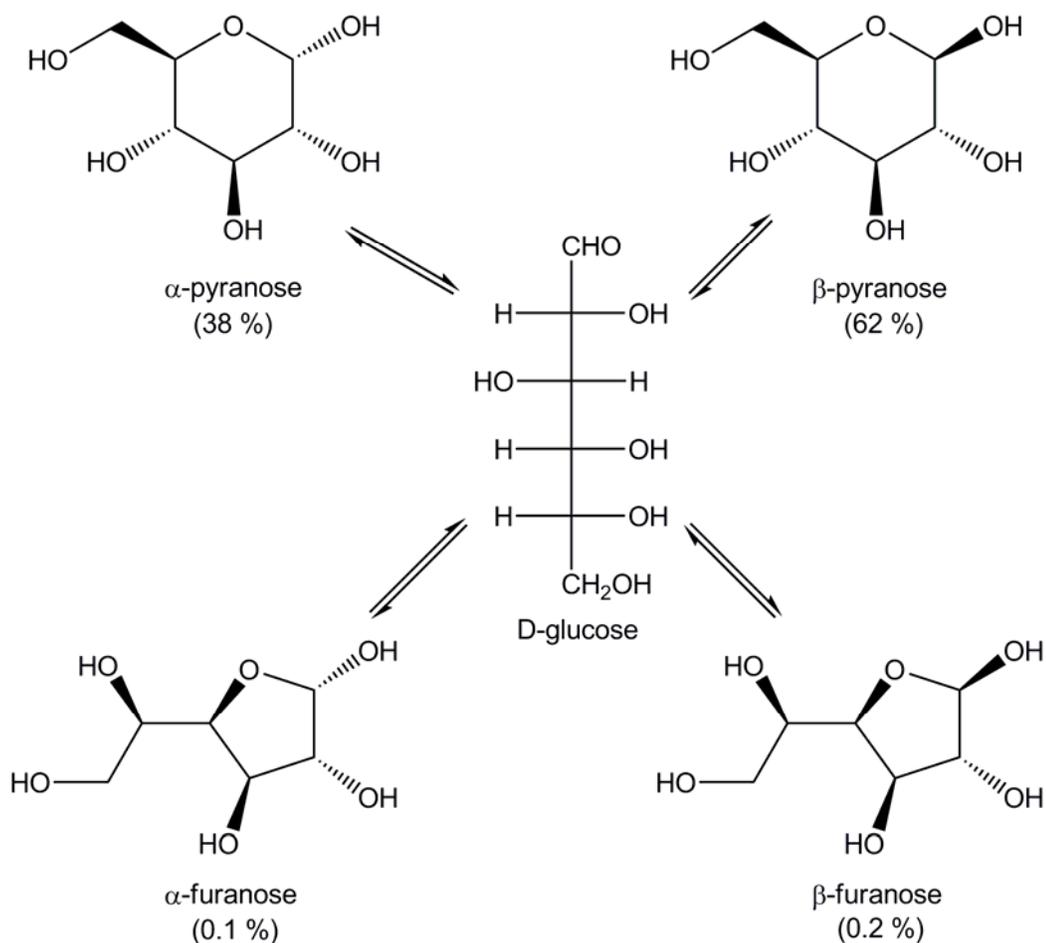


Figure 1-6. Pyranose and furanose ring formation in equilibrium with the linear form of D-glucose. The net abundance of the linear form of D-glucose is negligible in solution.

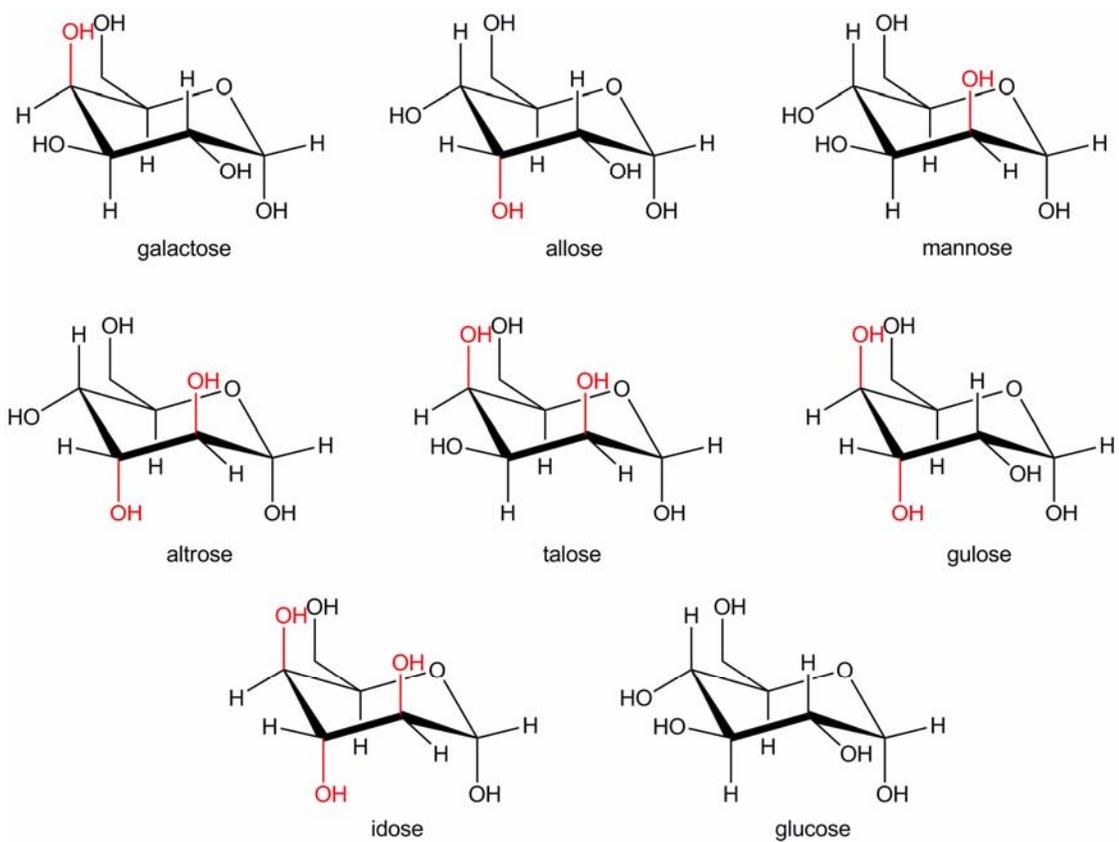


Figure 1-7. Six-membered ring structures of the α -anomers of D-hexoses, showing differences of stereochemistry (highlighted in red) relative to D-glucose (bottom right).

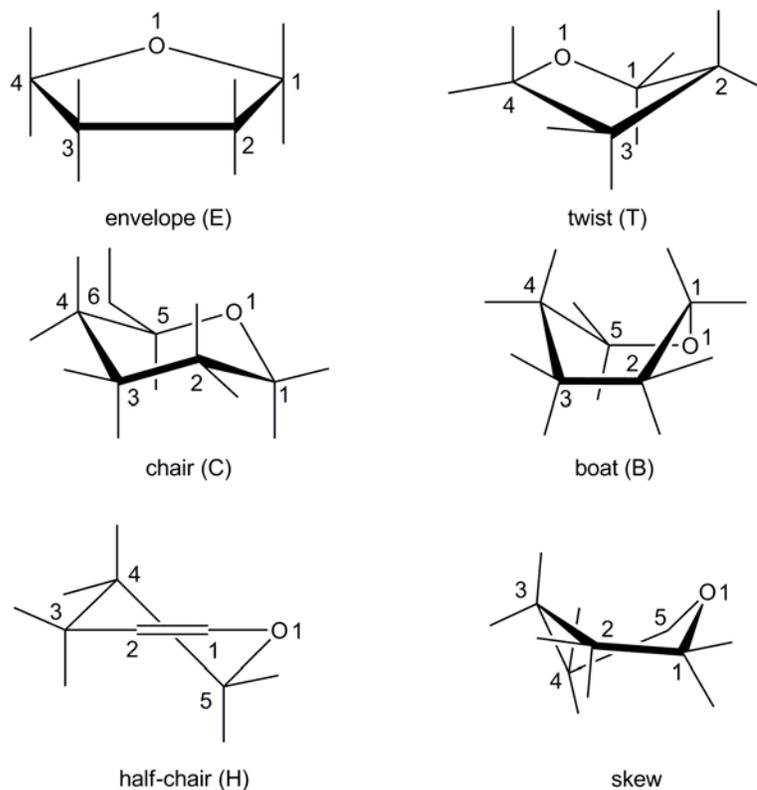


Figure 1-8. Ring conformations, for furanoses and pyranoses. Ring oxygen and carbons are numbered according to IUPAC convention.¹⁸

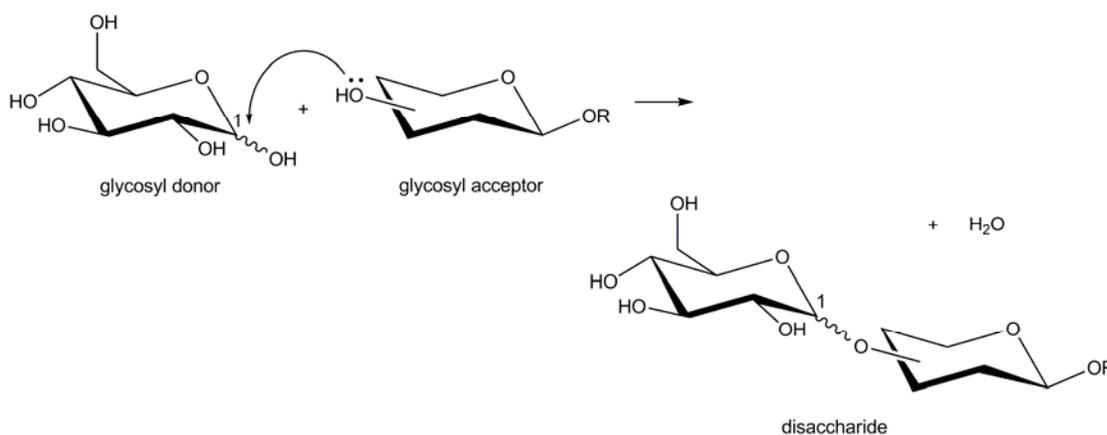


Figure 1-9. Disaccharide formation by glycosylation, showing attack by a monosaccharide acceptor on a monosaccharide donor. For simplicity, the functional groups of the acceptor are symbolized with one hydroxyl group. The linkage can occur at C-1, C-2, C-3, C-4 and C-6 of the acceptor. Stereochemistry at the anomeric carbon of the donor has not been explicitly indicated, and both configurations are possible in the glycosylation reaction

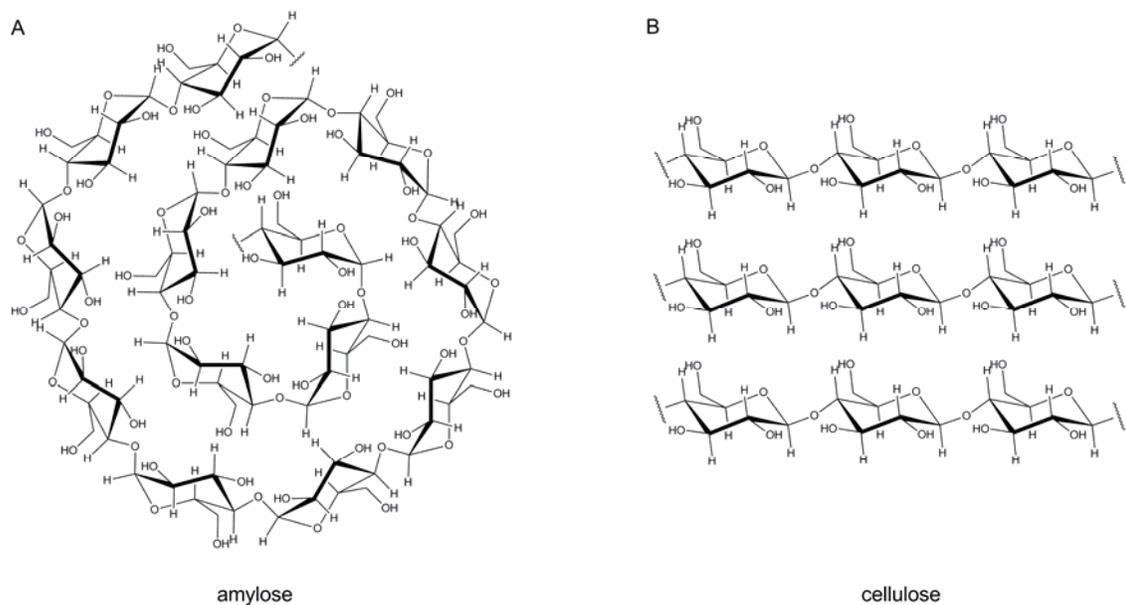


Figure 1-10. A) Representation of the polysaccharide, amylose, showing the helical nature due to the $\alpha(1\rightarrow4)$ glycosidic bonds. B) The $\beta(1\rightarrow4)$ linked cellulose, shown as stacked ribbon-like chains.

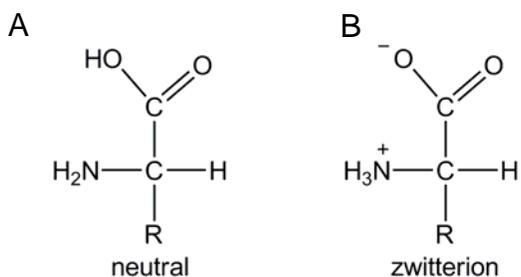


Figure 1-11. Amino acid in the A) unionized and B) zwitterionic forms. The amino group is attached to the alpha carbon, which has the L configuration in all naturally occurring amino acids.

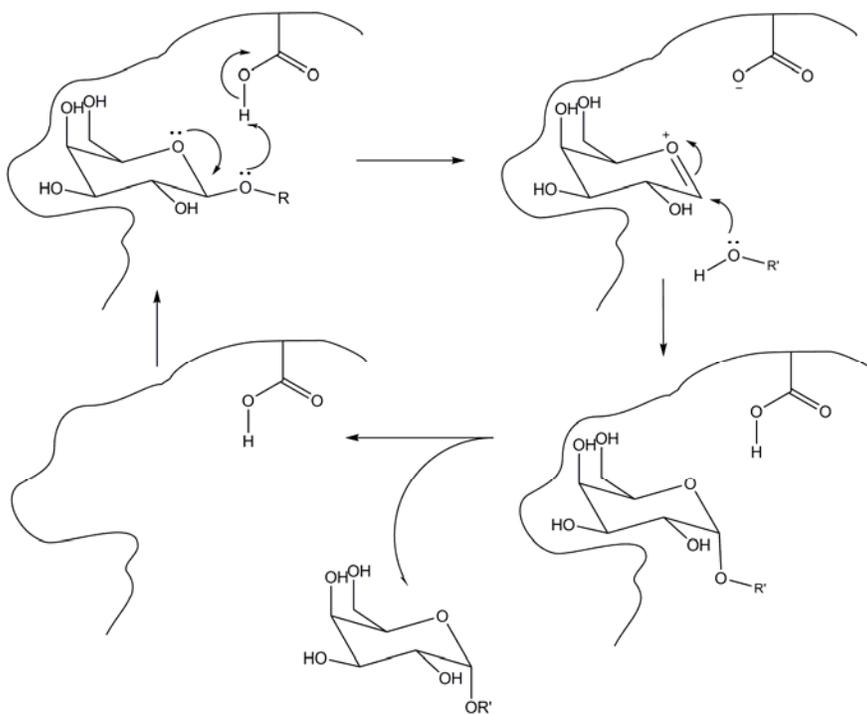


Figure 1-14. Process of inverting a β -galactoside to the corresponding α -galactoside by using an inverting glycosidase. A two step S_N1 process is shown but the inversion is also possible through an S_N2 reaction.

CHAPTER 2 MATERIALS AND METHODS

Mass Spectrometric Approaches

An excellent way to study gas phase ions is through the use of mass spectrometry. Modern mass spectrometry (MS) was pioneered by Dempster and Aston^{19,20} for separation of isotopes. They implemented ion deflecting mass spectrometers that would become the basis of magnetic sector instruments. Other mass spectrometers quickly followed, including the time of flight²¹ and quadrupole mass spectrometers.²² Early mass spectrometric studies of carbohydrates complemented the work of Emil Fischer and his studies of the structure of monosaccharides.²³ Development of mass spectrometry instrumentation to study amino acids, proteins and carbohydrates in the 1980s showed that MS was a robust analytical technique, useful in biochemical research.

Two difficulties arise in obtaining mass spectra of carbohydrates. The low volatility of sugars makes them difficult to vaporize and the sample may require heating to obtain gaseous molecules and ions. Heating sugar samples can cause instabilities, leading to degradation of the sample and fragmentation of parent ions, which can pose a problem when trying to obtain the molecular weight of the sugar. One of the first ionization methods for mass spectrometry, electron impact ionization (EI),²⁴ was used to obtain a mass spectrum of D-glucose in a quadrupole mass spectrometer.³ To obtain the molecular mass of D-glucose, the sample was heated to 130° C under reduced pressure and then bombarded with electrons, generating positive ions of D-glucose. The EI process requires high energy electrons, most often around 70 eV. Because of the high energy electron bombardment, the molecular ion (M^+) of D-glucose rearranges, leading to bond cleavage and formation of fragment ions and neutrals. Therefore, the molecular mass cannot be determined with EI-MS. The mass spectrum of D-glucose from EI-

MS in Figure 2-1 has no corresponding peak for the M^+ ion of D-glucose, which has a molecular weight of 180 Dalton (Da) or atomic mass units. The highest mass is 149 Da, corresponding to loss of CH_2OH from the molecular ion.

Another mass spectrometric study of D-glucose, using the chemical ionization (CI)²⁵ technique, was able to yield the protonated molecular ion and the molecular mass of D-glucose was verified.^{3,26} The chemical ionization process used an ionized gas, in this case methane, for which ion/molecule reactions form CH_5^+ and C_2H_5^+ , which in turn react with the gaseous sugar sample, donate a proton, and form the $[\text{M}+\text{H}]^+$ glucose ion. Figure 2-1 also shows the mass spectrum of D-glucose from the CI-MS experiment, and the highest mass peak, 181 Da, corresponds to D-glucose with the added proton. The abundance of the $[\text{M}+\text{H}]^+$ is low and multiple fragmentation is observed in the CI mass spectrum.

The mass spectra, using fast atom bombardment ionization (FAB),^{27,28} of 21 amino acids were used to obtain their molecular weights.²⁹ The solid sample containing the amino acid was dissolved in a glycerol matrix and transferred to a copper tip. Bombardment with argon atoms (4-6 keV) onto the copper tip produced ions that were analyzed using a Kratos MS 50 magnetic sector mass spectrometer. The molecular ions of the amino acids were evident in the mass spectra, but the amide terminal group fragment ion was obscured by the glycerol matrix and by background peaks from the sample.

These mass spectrometers and their corresponding ionization techniques were able to study small carbohydrates and amino acids, but molecular ions from larger molecules were not produced due to increasing low volatility of the larger species and congestion of peaks in the spectra for those the ions that were formed. Even with small ions, high resolution may be necessary if trying to make a distinction between OH and NH_3 , both having a mass of 17 Da but

differing in mass if more precise measurements are made. Ion cyclotron resonance mass spectrometers made high resolution experiments of biologically relevant ions possible, especially after introduction of the Fourier transform approach by Comisarow and Marshall.³⁰

Fourier transform ion cyclotron mass spectrometry (FTICR-MS),³⁰⁻³⁵ when combined with various dissociation methods, has been established as an important method for analysis of carbohydrates and proteins and was the mass spectrometric technique used in this work. The following section will deal with the theory of ion cyclotron motion, the magnetron motion that is a combination of the cyclotron and trapping motions, and a general description of the detection of ions. The various dissociation methods, especially infrared multiple photon dissociation and the corresponding infrared lasers will be subsequently discussed. The most basic components of the FTICR mass spectrometer will also be described.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

The major difference between FTICR and other mass spectrometric techniques is the method by which ions are trapped and detected. Mathematical description of the trapping and detection of ions has been well documented,³² and an overview of how this occurs is beneficial to understanding the spectra obtained from FTICR.

Cyclotron motion. A charged particle moves in a circular fashion when subjected to a uniform magnetic field.³⁶ The force, F , applied to the charged species is given by Equation 2-1.

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \quad (2-1)$$

where \vec{E} is the electric field, \vec{v} is velocity vector of the ion and \vec{B} is the magnetic field component of the uniform electromagnetic field. The charge of the ion is q , where q and z are used interchangeably. Assuming that the ions are moving in a uniform magnetic field and no electric field is present, the first term on the right side of Equation 2-1 can be neglected. The

vector product of the velocity and the magnetic field can be separated into its individual Cartesian coordinates by evaluating the determinant shown in Equation 2-2:

$$\vec{v} \times \vec{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ v_x & v_y & v_z \\ B_x & B_y & B_z \end{vmatrix} \quad (2-2)$$

$$\det(\vec{v} \times \vec{B}) = \hat{i}(v_y B_z - v_z B_y) - \hat{j}(v_x B_z - v_z B_x) + \hat{k}(v_x B_y - v_y B_x) \quad (2-3)$$

The unit vector components for x, y and z Cartesian coordinates axes are \hat{i} , \hat{j} , and \hat{k} , respectively. If the magnetic field is uniform along the z-axis with no x and y components, then the determinant can be simplified to Equation 2-6.

$$\vec{B} = B_z, \quad B_x = B_y = 0 \quad (2-4)$$

$$\vec{B} = B\hat{k} \quad (2-5)$$

$$\det(\vec{v} \times \vec{B}) = \hat{i}(v_y B_z) - \hat{j}(v_x B_z) \quad (2-6)$$

The acceleration of the ion due to the magnetic field can be described by the derivative of the velocity times the mass (m) of the ion and is equivalent to the force by the uniform magnetic field on the charged species.

$$\vec{F} = m\vec{a} = m \frac{d\vec{v}}{dt} = m \frac{d(\vec{v}_x + \vec{v}_y + \vec{v}_z)}{dt} = q(\vec{v} \times \vec{B}) = q[\hat{i}(v_y B_z) - \hat{j}(v_x B_z)] \quad (2-7)$$

Each acceleration component can be related to a distinct term in the determinant given in Equation 2-6.

$$m \frac{dv_x}{dt} = qv_y B_z \quad (2-8)$$

$$m \frac{dv_y}{dt} = -qv_x B_z \quad (2-9)$$

$$m \frac{dv_z}{dt} = 0 \quad (2-10)$$

Rearranging Equations 2-8 and 2-9 shows that the acceleration is dependent on the mass-to-charge ratio (m/z) of the ion and the magnetic field strength along the z-axis.

$$\frac{dv_x}{dt} = \frac{q}{m} B_z v_y = \omega_c v_y \quad (2-11)$$

$$\frac{dv_y}{dt} = -\frac{q}{m} B_z v_x = -\omega_c v_x, \quad (2-12)$$

where ω_c is the cyclotron frequency of the ion, equal to $(q/m) \cdot B$. There is no influence of the magnetic field on ion motion along the z-axis and cyclotron motion occurs only in the x-y plane.

Equations 2-11 and 2-12 are a set of coupled differential equations and in the general case the velocity as a function of time (t) along each axis can be described by Equations 2-13 and 2-14.

$$v_x(t) = v_y(0)\sin(\omega_c t) + v_x(0)\cos(\omega_c t) \quad (2-13)$$

$$v_y(t) = v_y(0)\cos(\omega_c t) - v_x(0)\sin(\omega_c t) \quad (2-14)$$

The initial velocity vector of an ion can be decomposed into initial x and y velocity vectors, with magnitudes related by

$$v_p(0) = \sqrt{v_x^2(0) + v_y^2(0)} \quad (2-15)$$

The initial angle of motion (γ) is the arctangent of the ratio of the initial y and x velocity vectors (Figure 2-2).

$$v_x(0) = v_p(0)\cos(\gamma) \quad (2-16)$$

$$v_y(0) = v_p(0)\sin(\gamma) \quad (2-17)$$

$$\gamma = \arctan \frac{v_y(0)}{v_x(0)} \quad (2-18)$$

Substituting Equations 2-16 and 2-17 into the general equations for the velocity vectors, and expanding and simplifying Equations 2-13 and 2-14 using the algebraic identities in Equations 2-19 and 2-20 gives Equations 2-21 and 2-22.

$$v_x(t) = v_p(0)\cos(\omega_c t - \gamma) \quad (2-19)$$

$$v_y(t) = -v_p(0)\sin(\omega_c t - \gamma) \quad (2-20)$$

$$v_x(t) = v_p(0)[\sin(\gamma)\sin(\omega_c t) + \cos(\gamma)\cos(\omega_c t)] \quad (2-21)$$

$$v_y(t) = v_p(0)[\sin(\gamma)\cos(\omega_c t) - \cos(\gamma)\sin(\omega_c t)] \quad (2-22)$$

Recognizing that the velocity is a derivative of the position of the ion with respect to time, equations for the x and y coordinates of the ion, with respect to time, are obtained.

$$\frac{dx}{dt} = v_x(t) \quad (2-23)$$

$$\frac{dy}{dt} = v_y(t) \quad (2-24)$$

$$x(t) = \frac{v_p(0)}{\omega_c} \sin(\omega_c t - \gamma) + C_1 \quad (2-25)$$

$$y(t) = \frac{v_p(0)}{\omega_c} \cos(\omega_c t - \gamma) + C_2 \quad (2-26)$$

The constants of integration, C_1 and C_2 , can be obtained in the case when time $t = 0$, and by rearrangement of $x(t)$ and $y(t)$.

$$x(0) = \frac{v_p(0)}{\omega_c} \sin(-\gamma) + C_1, \quad y(0) = \frac{v_p(0)}{\omega_c} \cos(-\gamma) + C_2 \quad (2-27)$$

$$C_1 = x(0) + \frac{v_p(0)}{\omega_c} \sin(\gamma) \quad (2-28)$$

$$C_2 = y(0) - \frac{v_p(0)}{\omega_c} \cos(\gamma) \quad (2-29)$$

Using the identities for $\cos(a-b)$ and $\sin(a-b)$, the equations that describe ion cyclotron motion along the x and y Cartesian coordinates in a uniform magnetic field along the z-axis are given by Equations 2-31 and 2-32.

$$\sin(-\gamma) = -\sin(\gamma), \quad \cos(-\gamma) = \cos(\gamma) \quad (2-30)$$

$$x(t) = x(0) + \frac{v_p(0)}{\omega_c} [\sin(\gamma)(1 - \cos(\omega_c t)) + \cos(\gamma)\sin(\omega_c t)] \quad (2-31)$$

$$y(t) = y(0) - \frac{v_p(0)}{\omega_c} [\cos(\gamma)(1 - \cos(\omega_c t)) - \sin(\gamma)\sin(\omega_c t)] \quad (2-32)$$

The general radial component of a vector provides the description of the ion cyclotron motion near the origin when substituting in Equations 2-31 and 2-32 for $x(t)$ and $y(t)$ in Equation 2-33.

$$|\vec{r}(t)| = \sqrt{x^2(t) + y^2(t)} \quad (2-33)$$

$$|\vec{r}(t)| = \sqrt{2} \frac{v_p(0)}{\omega_c} [1 - \cos(\omega_c t)]^{1/2} \quad (2-34)$$

The angular component of the trajectory near the origin can be described with $\gamma = 0$ and substituting into Equations 2-31 and 2-32 yields Equations 2-35 and 2-36.

$$x(t) = \frac{v_p(0)}{\omega_c} \sin(\omega_c t) \quad (2-35)$$

$$y(t) = -\frac{v_p(0)}{\omega_c} (1 - \cos(\omega_c t)) \quad (2-36)$$

The position of the ion at different time can be calculated, and plotting the position at times $t = \pi/2\omega_c, \pi/\omega_c, 3\pi/2\omega_c, 2\pi/\omega_c$, the ion motion is seen to be circular and periodic (Figure 2-3A).

Equation 2-37 describes the periodicity, or frequency, of the motion.

$$v_c = \frac{\omega_c}{2\pi} \quad (2-37)$$

Motion starting in the opposite direction, $\gamma=\pi$, gives equations for $x(t)$ and $y(t)$ that have opposing trajectories but have the same terms as $x(t)$ and $y(t)$ at $\gamma = 0$.

$$x(t) = -\frac{v_p(0)}{\omega_c} \sin(\omega_c t) \quad (2-38)$$

$$y(t) = \frac{v_p(0)}{\omega_c} (1 - \cos(\omega_c t)) \quad (2-39)$$

Plotting x and y coordinates for times $t = \pi/2\omega_c, \pi/\omega_c, 3\pi/2\omega_c, 2\pi/\omega_c$, the ion motion is again circular and periodic (Figure 2-3B). Since the ICR cell is placed inside the bore of the magnet and parallel to the magnetic field, the ions will develop an internal cyclotron motion perpendicular to the field and therefore ions will move in a circular fashion, dictated by the equations just described. The mass-to-charge ratio of the ion will dictate the cyclotron frequency ω (Equation 2-40). For an ion with mass 221 Da and a charge of +1, in a magnetic field of 4.7 Tesla strength, the ion would have an angular frequency (ω_c) of 324.1 kHz, and a radio frequency of 51.6 kHz would be required to excite the ion into a higher orbit.

$$\omega = \frac{q}{m} B \quad (2-40)$$

Analysis of the units starting with the Lorentz force equations, 2-1 and 2-7, shows the relationship between the frequency and mass-to-charge and magnetic field product.

$$\frac{d\bar{v}}{dt} = \frac{q}{m} (\bar{v} \times \bar{B}) \rightarrow \frac{m}{s^2} = \frac{C \cdot mol}{kg} \left(\frac{m}{s} T \right), \text{ rearranges to give } \frac{1}{s} = \frac{mol}{kg} C \cdot T, \text{ where } C \text{ is the}$$

coulomb charge of the ion and T designates the magnetic field strength in Tesla units.

In order for the ion cyclotron motion to be beneficial for detecting the ions in the ICR cell, the ion cyclotron orbital radius needs to be excited using a uniform electric field oscillating at the cyclotron frequency of an ion with a particular m/z . The presence of an alternating electric field, \bar{E} , can be described by the Equation 2-1.

$$\bar{F} = m \frac{d\bar{v}}{dt} = q(\bar{E} + \bar{v} \times \bar{B}), \quad \bar{B} = B\hat{k} \quad (2-41)$$

The alternating electric field can be described by Equation 2-42.

$$\bar{E} = E_o \sin(\omega_1 t + \phi) \hat{i} \quad (2-42)$$

The magnitude of the electric field is E_o , ω_1 is the frequency of the alternating electric field, and ϕ is the phase of the alternating field. Rearranging Equation 2-41 to obtain the acceleration and substituting in the ion cyclotron frequency, ω_c , gives the coupled partial differentials of the velocity along the x and y axes due to the presence of the electric field.

$$\frac{dv_x}{dt} = \frac{q}{m} E_o \sin(\omega_1 t + \phi) + \omega_c v_y \quad (2-43)$$

$$\frac{dv_y}{dt} = -\omega_c v_x \quad (2-44)$$

Differentiating Equation 2-44, the partial second derivative of the velocity along y is equivalent to the negative partial first derivative of the velocity along x times the ion cyclotron frequency, ω_c .

$$\frac{\partial^2 v_y}{\partial t^2} = -\omega_c \frac{\partial v_x}{\partial t} = -\omega_c \left[\frac{q}{m} E_o \sin(\omega_1 t + \phi) + \omega_c v_y \right] \quad (2-45)$$

Equation 2-45 is a quadratic equation and can be solved using the general solutions of differential equations of the form $(D^2 + \omega_c^2) \cdot v_y$, to obtain velocity components along x and y coordinates.

$$\frac{\partial^2 v_y}{\partial t^2} + \omega_c \frac{q}{m} E_o \sin(\omega_1 t + \phi) + \omega_c^2 v_y = 0 \quad (2-46)$$

$$v_y(t) = c_1 \cos(\omega_c t) + c_2 \sin(\omega_c t) - \omega_c \frac{q}{m} E_o \frac{\sin(\omega_1 t + \phi)}{\omega_c^2 - \omega_1^2} \quad (2-47)$$

$$v_x(t) = c_1 \sin(\omega_c t) - c_2 \cos(\omega_c t) + \omega_1 \frac{q}{m} E_o \frac{\cos(\omega_1 t + \phi)}{\omega_c^2 - \omega_1^2} \quad (2-48)$$

The components $v_y(t)$ and $v_x(t)$ are related by Equation 2-49.

$$\frac{dv_y(t)}{dt} = -\omega_c v_x \quad (2-49)$$

Equation 2-47 reduces to Equation 2-50 when $t = 0$.

$$v_y(t) = c_1 - \omega_c \frac{q}{m} E_o \frac{\sin(\phi)}{\omega_c^2 - \omega_1^2} \quad (2-50)$$

Similarly, the velocity component of along the x-axis can be simplified at $t = 0$.

$$v_x(t) = -c_2 + \omega_1 \frac{q}{m} E_o \frac{\cos(\phi)}{\omega_c^2 - \omega_1^2} \quad (2-51)$$

Let $\phi = x(0) = y(0) = 0$:

$$x(t) = x(0) + \frac{v_p(0) \sin(\gamma)}{\omega_c} (1 - \cos(\omega_c t)) + \frac{v_p(0)}{\omega_c} \cos(\gamma) \sin(\omega_c t) + \frac{q}{m} E_o \frac{\sin(\omega_1 t) - \frac{\omega_1}{\omega_c} \sin(\omega_c t)}{\omega_c^2 - \omega_1^2} \quad (2-52)$$

$$y(t) = y(0) - \frac{v_p(0)\cos(\gamma)}{\omega_c}(1 - \cos(\omega_c t)) - \frac{v_p(0)}{\omega_c}\sin(\gamma)\sin(\omega_c t) - \frac{q}{m} \frac{E_o}{\omega_1 \omega_c} + \frac{q}{m} E_o \frac{\frac{\omega_c}{\omega_1} \cos(\omega_1 t) - \frac{\omega_1}{\omega_c} \cos(\omega_c t)}{\omega_c^2 - \omega_1^2} \quad (2-53)$$

When $\gamma=0$, a situation may arise where $\omega_1 = \omega_c$ and the functions of $x(t)$ and $y(t)$ are undefined. Using L'Hopital's rule, the undefined terms of $x(t)$ and $y(t)$ can be solved. Equation 2-54 gives an example for $x(t)$.

$$\lim_{\omega_1 \rightarrow \omega_c} \frac{\omega_1 \sin(\omega_c t) - \omega_c \sin(\omega_1 t)}{\omega_c^2 - \omega_1^2} \Big|_{\omega_1} = \frac{\sin(\omega_c t) - \omega_c t \cos(\omega_c t)}{-2\omega_1} \quad (2-54)$$

Using the case where $\omega_1 = \omega_c$, the previously undefined terms can now be simplified.

$$\frac{\sin(\omega_c t) - \omega_c t \cos(\omega_c t)}{-2\omega_c} \quad (2-55)$$

$$x(t) = \frac{v_p(0)}{\omega_c} \cos(\gamma) \sin(\omega_c t) + \frac{q}{m} \frac{E_o}{\omega_c} \left[\frac{\sin(\omega_c t) - \omega_c t \cos(\omega_c t)}{-2\omega_c} \right] \quad (2-56)$$

$$y(t) = \frac{v_p(0)}{\omega_c} (\cos(\omega_c t) - 1) + \frac{q}{2m} \frac{E_o}{\omega_c^2} [2\cos(\omega_c t) + \omega_c t \sin(\omega_c t) - 2] \quad (2-57)$$

Equations 2-56 and 2-57 give the motion of the ion as its cyclotron motion is excited by the alternating electric field (Figure 2-4). Analysis of the motion indicates that the important terms that change are those shown in Equations 2-58 and 2-59.³²

$$x(t) = -\frac{q}{2m} \frac{E_o}{\omega_c^2} \omega_c t \cos(\omega_c t) \quad (2-58)$$

$$y(t) = \frac{q}{2m} \frac{E_o}{\omega_c^2} \omega_c t \sin(\omega_c t) \quad (2-59)$$

A schematic description of the FTICR ion detection process follows (Figure 2-5). A broadband RF voltage is applied to two excite plates of the ion trap; if a particular radio frequency is in resonance with the ion cyclotron frequency of a trapped ion with a particular m/z , the resonance condition will excite the ion into a higher cyclotron orbit. The ion cloud is now closer to the detection plates, and as the cloud sweeps by them, an alternating current is induced

on the plates. The current induced by each ion with different m/z in the ion cloud can be plotted as a function of time. The image current can be very convoluted if more than a few ions with different m/z are present (Figure 2-5B). This transient response signal of induced current as a function of time can then be deconvoluted through Fourier transformation. The Fourier analysis transforms the time domain signals into discrete frequency domain values, which correspond to the inherent ion cyclotron resonance frequencies of each mass ion present. Using Equation 2-40, the frequencies are converted into the m/z domain spectrum, or mass spectrum.

Trapping motion. So far, only ideal conditions of cyclotron motion have been discussed, and the kinetic energy of the ions as they enter the cell has been neglected. Ions traveling into the ion trap experience a potential barrier (positive or negative voltage, depending on whether the ion is positive or negative, respectively) at the end plate of the ion trap. The potential is made large enough so that the ions are deflected toward the front of the ion trap. Another potential is raised at the front plate, so that the ions are deflected back and forth within the ion trap, which is called the trapping motion. With the presence of the trapping potential and the magnetic field, a 3-dimensional trap has been created and the resulting potential is described by Equation 2-60.

$$\Phi(r, z) = V_{trap} \left[\gamma + \frac{\alpha}{2a^2} (2z^2 - r^2) \right] \quad (2-60)$$

$$r = \sqrt{x^2 + y^2} \quad (2-61)$$

where r is the radial position of the ion in the xy -plane, V_{trap} is the trapping voltage, a is the measure of trap size and γ and α are constants that depend on the ion trap shape. The electric field is obtained by taking the negative derivative of the potential with respect to the z direction.

$$E(z) = -\frac{d\Phi}{dz} = -\frac{2V_{trap} \alpha}{a^2} z \quad (2-62)$$

The force on the ion applied by the electric field would accelerate the ion along the z axis in a motion parallel to the magnetic field, which is uniform along the z-axis of the laboratory reference frame.

$$F(z) = m \frac{dv_z}{dt} = \frac{2qV_{trap}\alpha}{a^2} z \quad (2-63)$$

The force equation indicates that the trapping motion is a simple harmonic oscillation between the trapping plates. The oscillation frequency, ω_z , is given in Equation 2-64.

$$\omega_z = \sqrt{\frac{2qV_{trap}\alpha}{ma^2}} \quad (2-64)$$

Magnetron motion. The magnetic and electric fields together introduce a third fundamental motion of the trapped ion, called the magnetron motion. The cyclotron and trapping motions are not coupled, although the trapping potential produces an outward-directed electric force that opposes the inward-directed Lorentz force of the magnetic field.

$$F(r) = qE(r) = \frac{qV_{trap}\alpha}{a^2} r \quad (2-65)$$

Combination of both force equations provides the motion of an ion subjected to a 3-dimensional axial trapping quadrupolar electrostatic potential and a uniform magnetic field.

$$F = m\omega^2 r = qB_o \omega r - \frac{qV_{trap}\alpha}{a^2} r \quad (2-66)$$

$$\omega^2 - \frac{qB}{m} \omega + \frac{qV_{trap}\alpha}{ma^2} = 0 \quad (2-67)$$

Equation 2-67 is a quadratic that is independent of the ion position inside the trap (r). Two natural rotational frequencies are obtained from the solution of Equation 2-67.

$$\omega_{\pm} = \frac{\omega_c}{2} \pm \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \quad (2-68)$$

$$\omega_- = \frac{\omega_c}{2} - \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \quad (2-69)$$

Where ω_+ is the “reduced” cyclotron frequency of the ion in the presence of the d.c. trapping potential and ω_- is the magnetron frequency. The magnetron and trapping frequencies are usually less than the cyclotron frequency, so only the cyclotron frequency is detected. The cyclotron, trapping and magnetron motions of the ion (Figure 2-6). As seen from the figure, the magnetron motion is circular and superimposed on the cyclotron motion. Both motions are superimposed on the trapping motion.

Instrumentation. To introduce ions into the ion trap, various sources are used for internal or external ion production. As was mentioned earlier, electron impact ionization (EI) was one of the first internal sources, producing ions by bombardment of a gaseous sample with electrons. In FTICR instruments the EI source is inside the vacuum chamber, near the magnetic field and the ion trap. A positively charged repeller plate in the source moves the cations, formed from the bombardment of the gaseous sample with high energy electrons (70 eV), into the ion trap to be detected. External sources tend to be farther from the magnet, and therefore ions produced in them need to be guided with electrostatic optics toward the ion trap. The longer path requires a vacuum pumping system that will keep the vacuum low enough so the ions do not suffer too many collisions as they travel toward the ion trap. This places stringent pumping requirements on FTICR-MS, since it requires very low pressures ($< 10^{-9}$ torr) to keep ions trapped for long periods of time. A number of different ionization techniques can be employed to produce saccharide ions in mass spectrometers, both within the ion source and externally. Currently, the most popular ones are matrix assisted laser desorption ionization (MALDI),³⁷ electrospray ionization (ESI),³⁸⁻⁴³ and, in the recent past, fast atom bombardment (FAB). The external source developed by John Fenn and co-workers, electrospray ionization (ESI), has allowed FTICR-MS

to develop as a valuable tool for studying molecules of many sizes, including carbohydrates and peptides, and was the ionization method used for this work.

Electrospray ionization is able to introduce liquid samples that have a low vapor pressure, into the gas phase. ESI creates a fine mist by pushing the liquid sample through a small diameter tubing of silica. As the mist moves farther away from the tip, the mist spray expands and the individual droplets become smaller and smaller due to evaporation, until the solvent and the ions are separated from each other and the bare ions enter the capillary,⁴⁴ although ions surrounded by solvent continue to be present inside the capillary (Figure 2-7). A positive or negative electrostatic potential of approximately 3000 V is placed on the capillary to attract ions of opposite charge, thus taking the ions from an atmospheric pressure environment to a vacuum region of 10^{-3} torr. As the ions travel through the capillary, further desolvation is facilitated by heating the capillary in the range of $100^{\circ} - 200^{\circ}$ C. The ions pass through a set of skimmers to improve ion abundance by removing neutral molecules. The ions at this point have differing kinetic energies, and therefore are trapped and accumulated in a hexapole by an RF voltage. The ions are held for 500 - 1000 ms and are sent as a bunch toward the ion trap. The ion cloud needs to be kept from diverging due to the electrostatic repulsion between cations and the increasing low pressure gradient in the vacuum chamber. A set of electrostatic plates, or einzel lenses, are used to focus and slightly deflect the ions, although some FTICR-MS instruments use octopole ion guides to steer the ions into the ion trap (Figure 2-8).

Because of the gentle nature of the ESI process, molecular ions as well as multiply-charged ions are observed in the mass spectrum. The multiply-charged ion formation is an important feature of ESI, since FTICR mass spectrometers measure m/z , making it possible to observe large molecules in an instrument with a restricted mass range.

An example of the use of mass spectrometry to study sugars is found in the work of Cole *et al.*, where the oligosaccharide ions of maltose (G_2) through maltoheptaose (G_7) were formed by electrospray ionization and analyzed with a quadrupole mass spectrometer.⁴⁵ Using LiCl as the ionization agent, the oligosaccharide ions of maltose attached a lithium cation, but also formed complexes where multiple LiCl molecules were attached to the α -glucose monosaccharide units of the oligosaccharide ions. The number of LiCl attached increased as the oligomers became larger. Theoretical calculations indicated that the lithium cation was triply coordinated to hydroxyl groups and the glucose ring oxygen atom.

In another study involving proteins, the mass spectrum of myoglobin was observed using ESI, with myoglobin seen to form ions having multiple charge states.⁴⁶ The charge states from $z = 8^+$ to 23^+ were observed in the mass spectrum. Using the m/z of one of the observed charge states, the molecular weight of horse myoglobin was determined (MW = 16,951 Da).

The ion traps used with FTICR-MS are Penning traps.⁴⁷ Named after F. M. Penning by Hans G. Dehmelt, the most basic Penning trap is the cubical ion trap or cell,⁴⁸ which has six electrostatic plates (Figure 2-9). The front and back plates have an aperture at the center of each plate to allow the introduction of ions or radiation into the cell. These plates serve to trap ions by placing electrostatic voltages on the plates, creating a potential barrier that will deflect the ions back and forth within the ion trap. Two excite plates, placed left and right, and two detect plates, placed top and bottom, form the body of the trap. The excite plates have an RF voltage applied to them to increase the radial motion of the ions in the cell (as can be derived from Equations 2-58 and 2-59), while the detect plates are connected through a resistor to capture the alternating current induced by the ion charge as the ions sweep by the plates.

Similar functionality can be obtained by a closed cylindrical cell, where the trapping plates are circular instead of square or rectangular, and the detection and excite plates form a cylinder of four separate plates. A closed cylindrical cell was used in the FTICR-MS experiments conducted at the University of Florida. Another common configuration is the open cylindrical cell, where the trapping plates are replaced by two sets of cylindrical plates, similar in design to the cylindrical detect and excite plates. The open cylindrical cells have the advantage of accumulating the maximum amount of ions within the restrictions of the diameter of the cell. While almost any shape of cell can be constructed to trap ions, the open cylindrical cell has lately been the preferred choice to include in commercial instruments due to its high S/N, ion accessibility, simplicity in design and low cost construction. An open cylindrical cell was used in FTICR-MS experiments conducted at the FOM-Institute (Figure 2-10).

Tandem Mass Spectrometry

The FTICR-MS has the highest mass resolving power of any mass spectrometry technique. Separation of ions within 1 Da can quite easily be accomplished and accuracies of better than 300 ppb have been demonstrated.⁴⁹ One of the drawbacks of FTICR-MS and, indeed, any mass spectrometric technique is that isomer differentiation is not possible, since isomers have the same m/z . Tandem MS methods (MS^n) have been developed to overcome this limitation. As was previously discussed, trapping of the ions is essential for the detection of molecular ions. But trapping ions in the ICR cell can also serve to induce fragmentation of precursor ions. In MS^n experiments, a precursor ion is usually fragmented and may show a distinct fragmentation pattern relative to the other isomers. Further fragmentation of a fragment ion that was produced by dissociation of the precursor ion can be done to obtain a third level of identification (MS^3). The tandem MS process is limited by the S/N and abundance of the ion being fragmented, but MS^4 and higher have been conducted with FTICR-MS.⁴⁶

It has been shown for carbohydrate isomers that depending on the conditions of tandem MS experiments, the fragmentation patterns obtained for different isomers may or may not be distinct. Leary *et al.* obtained low-energy collision induced dissociation (CID) tandem mass spectra of monolithiated disaccharides that showed many cleavages of the reducing end monosaccharide, but there was little differentiation between the different linkages.⁵⁰ In the same study, each dilithiated disaccharide did produce distinct fragmentation ion patterns and it was possible to assign the (1→6) glycosidic bond linkages instead of the (1→4) linkages to four glucose oligosaccharide isomers.

The binding strength of alkali metal ions to carbohydrates was obtained using infrared multiple photon dissociation (IRMPD) and collision induced dissociation tandem mass spectrometric techniques. Cancilla and co-workers determined that the binding energy of alkali metal ions to oligosaccharides decreased going down group IA of the periodic table.^{51,52} Lithium is bound strongly enough to the carbohydrates that during irradiation and subsequent fragmentation its direct loss is much less relative to that of the other alkali metal cations. For ions containing K^+ , Rb^+ and Cs^+ only loss of the alkali metal cation was seen when the complexes were subjected to IRMPD. Using CID, other fragmentation channels were observed, even for precursor ions containing the larger alkali metal cations. Infrared multiple photon dissociation can also be used to obtain infrared spectra of ions in the gas phase, and therefore will be discussed next.

Infrared Multiple Photon Dissociation

Detailed knowledge of the primary structure of sugars and amino acids is the major goal of this work and tandem MS techniques that are complementary with gas phase infrared spectroscopy were utilized. But typical infrared spectroscopy bands of solid and solution phase samples are too broad to distinguish characteristics for sugar or amino acid isomers. The FTIR

spectrum of β -D-glucopyranose obtained from a crystalline sample in a KBr pellet illustrates the broad peaks in the infrared spectrum (Figure 2-11).

For gas phase photon absorbance spectroscopy the abundance of the ions is too low to detect a change in the incident radiation intensity. Action or consequence spectroscopy must be used and is distinct from absorbance spectroscopy in that the absorption of radiation is not explicitly detected, but instead, the action or consequence of absorption of radiation is monitored. One particular type of action spectroscopy, infrared multiple photon dissociation (IRMPD), is used to obtain infrared spectra of gas phase molecules or ions. By monitoring the fragmentation of the precursor ion and its signal loss in the mass spectrum at a particular irradiation wavelength, a spectrum of the loss as a function of different wavelengths can be constructed.

Mechanism. Infrared multiple photon dissociation (IRMPD) occurs because the fluence, or density over time, of the laser intensity is large enough that multiple photons of the same wavelength are absorbed by an ion, causing the ion to fragment. This phenomenon was first observed in the 1960s as the CO₂ laser became more powerful and had higher photon intensities. The collisionless environment that the FTICR-MS technique provides was used by Beauchamp and co-workers, who initially demonstrated the use of low intensity CO₂ lasers to induce fragmentation of ions in low pressure ICR traps.^{53,54} Absorption of photons whose frequency is in resonance with the frequency of a vibrational mode (ν_i), excites the ion into a higher vibrational level, (ν_i+1) (Figure 2-12). Intramolecular vibrational redistribution (IVR) distributes the energy throughout the other normal modes, and the original vibrational mode returns to its ground state. Another photon with the same wavelength can again be absorbed into the ground state of ν_i , and intramolecular redistribution of the photon energy occurs again. The sequential

absorption process is slow, and the FTICR-MS provides a collisionless environment so that collisional relaxation does not occur. Photon energy absorption and IVR continue so that in essence the process reaches a quasi-continuum of vibrational energy levels. Some vibrational modes may have shallow potential curvature, and as IVR occurs one of the modes may be shallow enough that the energy being redistributed into this mode will be large enough to cause dissociation. The quasicontinuum condition allows absorption of photon energy at any wavelength into the vibrational mode that has the lowest potential energy barrier toward dissociation, facilitating bond cleavage. Fragmentation can be observed from this weakest bond or other modes with relatively low potential energy barriers, and in FTICR-MS these fragmentations are observed in the tandem mass spectrum. It is important to note that if initially the laser frequency is not in resonance with at least one vibrational mode of the molecule, then the IVR process will be inefficient, and therefore subsequent dissociation cannot occur and fragmentation will be minimal.

The IRMPD experiment uses a high fluence, frequency tunable infrared laser and the overall fragmentation as a function of laser frequency, or IRMPD spectrum, can be obtained. The most common tunable infrared lasers for obtaining IRMPD spectra will be discussed in the next section. These include the CO₂, optical parametric oscillator and free electron lasers. For the most part all lasers have the following features:

- A lasing medium
- An optical resonant cavity
- A “pumping” mechanism that provides excitation to the medium to achieve population inversion

There are numerous configurations for each laser, so only the basic properties will be described and the features used in this work will be highlighted.

Carbon Dioxide Laser

One of the first lasers to be used for spectroscopy, the CO₂ laser^{55,56} is used throughout science, the manufacturing industry and medical fields in multiple applications. The lasing medium is a gas mixture that is mainly composed of the following gases: CO₂, N₂, He, Xe and H₂O. The gas mixture fills a chamber enclosure where a discharge is applied. The optical resonator includes one totally reflective surface and one semi-reflective mirror that form the chamber enclosure ends. For wavelength tunable CO₂ lasers, the totally reflective surface is replaced by a Brewster window and a diffraction grating.⁵⁷ The grating is rotated so that a different lasing wavelength will be in resonance in the optical cavity at a particular grating setting.

The basic mechanism of lasing follows. The gas discharge excites the CO₂ molecules from the ground vibrational level to higher vibrational levels. The discharge also excites the N₂ gas, and when N₂ collides with the CO₂ gas, energy is transferred causing the CO₂ molecule to go into a higher vibrational state (Figure 2-14). The N₂ vibrations are in close resonance with the ν_3 vibrational mode of CO₂ (asymmetric stretch), and therefore energy is transferred into this state. Relaxation occurs into lower lying excited vibrational levels, and photons are released within the resonant cavity. The photons with wavelengths that are resonant with the cavity length will dominate (constructive interference), and will induce the CO₂ molecules to absorb more radiation and stimulate emission of radiation with the same wavelength. This process of amplification continues to the point where the population of molecules occupying the higher energy ν_3 state is greater than the molecules occupying the other ν_1 and ν_2 states, or what is termed population inversion. The gain saturation is exceeded if the coherent beam of photons is intense enough to overcome the threshold of the resonant cavity, and the laser beam exits through the semi-reflective mirror. For a tunable CO₂, the length of the cavity can be changed so

that any of the rotational-vibrational modes of the CO₂ molecule can be in resonance within the cavity, and lasing occurs at the wavelength equal to the energy level of the rotational-vibrational mode.

The gas discharge increases the temperature of the lasing medium and degrades the lasing efficiency over time if the laser is not cooled. Most CO₂ lasers have a cooling system that runs water or a coolant through an isolated outer chamber of the resonance cavity, but some low power CO₂ lasers are air-cooled. Cooling is also obtained with the introduction of He in the gas mixture. He gas cools the gas discharge by increasing the thermal conductivity of the gas mixture, so that heat is quickly transferred to the walls of the resonance cavity, which is surrounded by the coolant chamber. The ionization potential of He (20 eV) is too high to influence the discharge (1 – 3 eV), but the discharge current can now be increased since the temperature of the gas mixture is lower due to the addition of He. The resulting effect is higher lasing efficiency and power. The addition of Xe to the gas mixture (0.5 torr) also increases laser efficiency and power by affecting the discharge conditions. The low ionization potential of Xe (12 eV) facilitates new electron production and maintains a constant discharge with a constant current. Because the current is constant, the electric field in the medium decreases and the mean energy of the electrons in the discharge is reduced. This is favorable for the CO₂ and N₂ gases, since their collision cross section increases when the electron energy is less than 4 eV. The resonance cavity can be sealed or the gas mixture can be allowed to flow through the chamber. In a sealed chamber, the electric discharge induces the breakdown of CO₂ to form CO, which degrades the lasing capabilities over time. Addition of 0.2 torr of H₂O is crucial, since it virtually eliminates the formation of CO and increases the lifetime of the laser⁵⁶ through the chemical reaction of $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{H}_2$. The constant concentration of CO₂ obtained by

the addition of H₂O maintains the pumping rate and therefore the lasing efficiency is not reduced. When the gas mixture flows through the chamber, a gas pumping system is used to carry away the degraded mixture. This requires a constant supply of the gas mixture, which can increase the cost of daily operation of the laser. The fast flow setup has the potential of obtaining higher powers, but advances in both types have been made and either setup can produce laser beams with high intensities adequate for IRMPD experiments.

Earlier work of Eyler and co-workers to obtain IRMPD spectra was done with CO₂ lasers.^{58,59} In a study by Shin and Beauchamp, the IRMPD spectrum of Mn(CO)₄CF₃⁻ was obtained by monitoring the loss of two carbon monoxide molecules as a function of CO₂ laser irradiation at different wavelengths.⁶⁰ One spectrum from that study (Figure 2-15) shows that CO₂ lasers do not provide continuous coverage of infrared bands. Nevertheless, the CO₂ laser has many uses, one of which is to obtain IRMPD fragmentation patterns of carbohydrates and amino acids. In chapter 5, the CO₂ laser will also be used as an off resonance photon source for IRMPD experiments using an optical parametric oscillator (OPO) laser, which will be discussed next.

Optical Parametric Oscillator

An optical parametric oscillator (OPO)⁶¹ consists of an optical resonator cavity and a non-linear crystal. The resonance cavity can be a two-mirror cavity or one end of the crystal can serve as the second reflective surface of the resonator (Figure 2-16). The OPO uses a pump laser and takes advantage of the properties of a non-linear crystal to obtain coherent radiation at wavelengths longer than the pump laser's wavelength. The pump photon is converted into two photons within the non-linear crystal, and their resultant laser beams are called the signal and idler beams. Due to conservation of energy, the idler and signal beam frequencies are equal to the pump laser's frequency.

$$\nu_p = \nu_i + \nu_s \quad (2-70)$$

Only the frequencies that satisfy the conservation of momentum will be generated efficiently.

$$\kappa_p = \kappa_i + \kappa_s \quad (2-71)$$

$$\kappa_j = 2\pi n_j h \nu_j \quad (2-72)$$

$$n_p \nu_p = n_i \nu_i + n_s \nu_s \quad (2-73)$$

The term κ is the wave vector of the frequency ν , n is the index of refraction at frequency ν_j , and h is Planck's constant. This phase matching condition allows for lasing at almost any wavelength, limited by the transparency region of the non-linear crystal. Because the OPO itself is not involved in a stimulated emission process to obtain a population inversion, it requires a pump laser with good beam quality and high power density. For a continuous-wave (cw) OPO, the threshold is met when the gain exceeds losses in the cavity resonator. The OPO can be constructed so that a single idler beam is output, or both idler and signal beams are present, termed a double resonant OPO. In the double resonant OPO, both beams depend on the configuration of the resonator so each beam radius is decreased to maximize the gain. The size of the beam radii is limited by the laser induced damage threshold and birefringence of the crystal. The interacting wavelengths in the resonance cavity and the crystal deposit heat throughout the crystal, establishing a gradient of hot and cold regions with the surface. Phase matching is dependent on a constant temperature and therefore cannot be maintained throughout the crystal, limiting the power output. The temperature gradient can be minimized if the crystal is placed within an oven and kept at a constant temperature. Fabricated non-linear crystals, such as a periodically poled lithium niobate (ppLiNbO₃) crystal, have high damage thresholds, so that Nd/YAG lasers of a few watts can be used as pump lasers. The high damage threshold of ppLiNbO₃ allows the crystal to be heated to high temperatures (150° C), increasing the efficiency and power of the output beams. Tuning of an OPO can be accomplished by changing

the temperature of the crystal, changing the phase matching conditions of the idler and signal beams. The ppLiNbO₃ crystal has multiple poling periods corresponding to different wavelength ranges that allow for lasing from 1.38-2.0 and 2.28-4.67 microns at a set temperature between 50 – 150° C. An etalon, placed within the resonance cavity, is used to tune a particular set of idler and signal wavelengths within a particular poling period and temperature range. To enhance one resonance mode within the cavity, a piezo-electric mirror can be used as one end of the resonance cavity and combined with lock-in electronics to keep the laser wavelength in resonance.

Optical parametric oscillators and amplifiers are relatively new technology and IRMPD spectra of carbohydrates are rare. Ultraviolet pulsed-OPOs have been used in ion dip spectroscopy studies,⁶² and Simons and co-workers have studied the glycosidic linkages of lactosides,⁶³ and hydrogen bonding of the monosaccharides mannose, galactose, glucose and lactose.⁶⁴ In the chapter 5, IRMPD spectra of the glycoside anomers of glucose and galactose, obtained with an infrared cw-OPO, will be presented.

Free Electron Laser

Ever since an FTICR-MS was coupled to a free electron laser (FEL)⁶⁵ at the FOM-Institute for Plasma Physics in The Netherlands (FOM-Institute) by Valle et al.,⁶⁶ other national laboratories have followed. There are a total of 13 locations that house FEL's for scientific research⁶⁷ and 3 of these are able to conduct IRMPD studies using the FEL with an FTICR-MS^{68,69} or other ion trap mass spectrometers.^{70,71} In the following section, the general lasing and tuning properties of the free electron laser will be summarized. Detailed properties of the FOM-Institute's FEL, the Free Electron Laser for eXperiments (FELIX), will be discussed in chapters 3 and 4 and therefore only relevant information regarding specific properties will be included in this summary.

An accelerating electron gun produces a beam of electrons that are guided into a resonant cavity and passed through an undulating magnetic field (Figure 2-17). In the resonance cavity, the electrons are accelerated with the magnetic field to relativistic speeds, and release photons. The electrons tend to bunch and thus radiate in phase. The bunching characteristic of the electrons makes all FELs pulsed-wave lasers. Gain amplification is required to overcome cavity losses and slight detuning of the resonant wavelength is done to increase the net gain of the optical wave. Because of the nature of the free electron beam, FELs are continuously tunable and have large output wavelength ranges (FELIX is tunable between 2.0 – 250 μm). A high energy electron beam (1 – 800 MeV, 2 – 500 Amp) and optimizing for maximum gain can obtain femtosecond pulses with 4 MW – 36 GW, producing laser power of 0.1 – 10 MW at 1% efficiency. An FEL produces short laser pulses of 500 fs – 10 ps that are separated by a few hundred ps – ns. Tuning can be accomplished in a number of different ways:

- The resonant wavelength can be changed by changing the electron beam energy, and this is necessary when the undulator magnets are fixed.
- Changing the magnetic field strength will alter the resonant wavelength, and the power output will be independent of the laser wavelength for a large range. Tuning by changing the magnetic field requires a high quality B field for fine wavelength selection.
- The laser can be operated at an odd harmonic of the fundamental wavelength, where the gain can be slightly higher.
- The FEL can be continuously tuned by moving the undulator magnets or wigglers a certain distance apart so that the free electrons and their released photons undulate at a specific wavelength.

A study by vonHelden et al. looked at the mid-ir spectra of six gas-phase conformers of phenylalanine.⁷² A neutral molecular beam was formed by laser desorption (1064 nm irradiation) of a graphite-phenylalanine solid mixture. A UV resonant 2-photon ionization (R2PI) experiment was conducted that scanned the wavelength of a Nd/YAG pumped dye laser (37500 – 37700 cm^{-1}) and indicated that six conformers of phenylalanine were present. The IR-

UV ion dip spectrum of each conformer was obtained by monitoring the ion signal loss as FELIX was scanned from 300 – 1900 cm^{-1} followed by irradiation with the dye laser set to a specific conformer's ionization wavelength. The ir spectrum of each conformer was distinctly different and density functional theory calculations indicated that the phenylalanine conformers have a “puckered” structure.

Implementation of IRMPD with FTICR-MS

General experimental procedure. An external ESI source coupled to an FTICR with a superconducting 4.7 tesla magnet was employed. To reduce collisions, pressures of approximately 10^{-9} torr were obtained near the ICR cell in the vacuum chamber, allowing for long trapping times without loss of ion signal due to dephasing collisions.³² Depending on the nature of the experiment, the solvent (mainly a methanol and water mixture) included either a dilute protic acid or an alkali metal salt to ionize the sample. The solution was electrosprayed into the FTICR and through the aid of electrostatic optics and differential pumping stages the ions were transferred to the ICR cell in the high homogeneity field of the superconducting magnet. Once the precursor ion was isolated, by ejection of all other ions using stored waveform inverse Fourier transform excitation (SWIFT),⁷³ the laser beam was introduced into the ICR cell. After laser fragmentation occurred, the fragment ions were excited by an RF broadband pulse and an image current was induced at the detection plates of the ICR cell so that the transient response could be obtained, Fourier transformed, and the mass spectrum processed and stored digitally. The laser was then tuned to a different lasing mode so that the fragmentation process could be repeated at this new lasing frequency. The lasers used provided infrared radiation between 600 – 2000 cm^{-1} and 2600 – 3900 cm^{-1} depending on the laser type used.

The IRMPD spectrum, while not directly indicating the amount of photon energy absorbed, does indicate what vibrational modes are infrared active. In other words, if a

theoretical vibrational spectrum of the molecular ion were obtained, those modes with the highest infrared intensities should be the most active in the absorption process of IRMPD, and would be predicted to produce the largest abundance of the fragment ions. Before obtaining a theoretical infrared spectrum of the ion of interest, the most stable conformations of the ions need to be predicted by theoretical calculations.

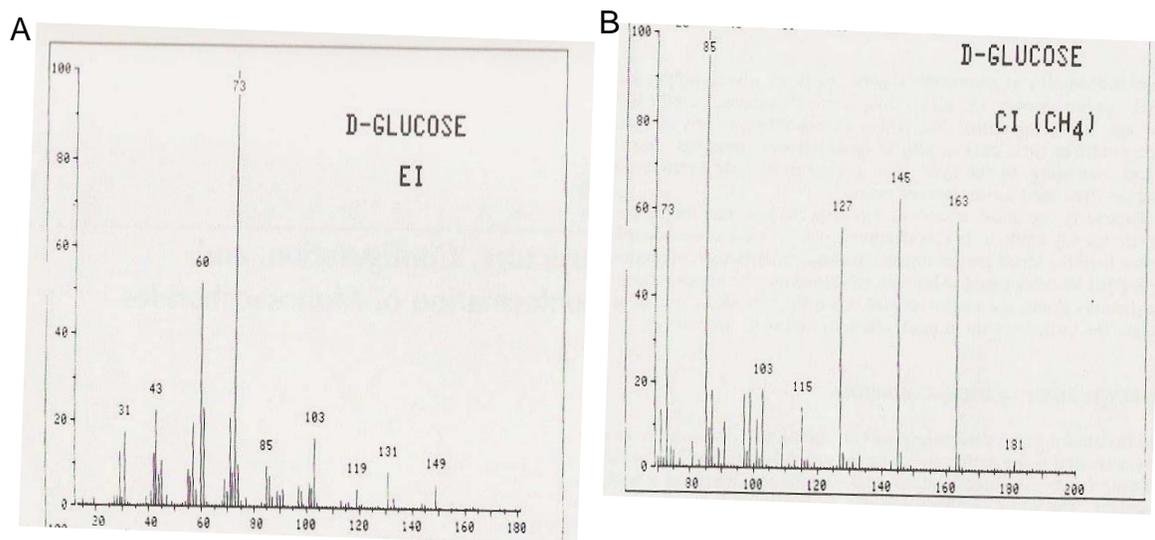


Figure 2-1. Mass spectrum of D-glucose by A) Electron impact ionization and B) chemical ionization, clearly showing the $[M+H]^+$ ion at 181 m/z . Images from *Carbohydrate Chemistry: Monosaccharides and their Oligomers*, by El Khadem *et al.*, Copyright © 1997.³ Reprinted with permission of Academic Press, Inc.

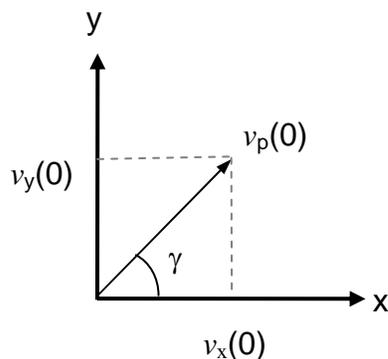


Figure 2-2. The initial trajectory of the ion, showing the initial velocity vector and angle of motion in the x-y plane.

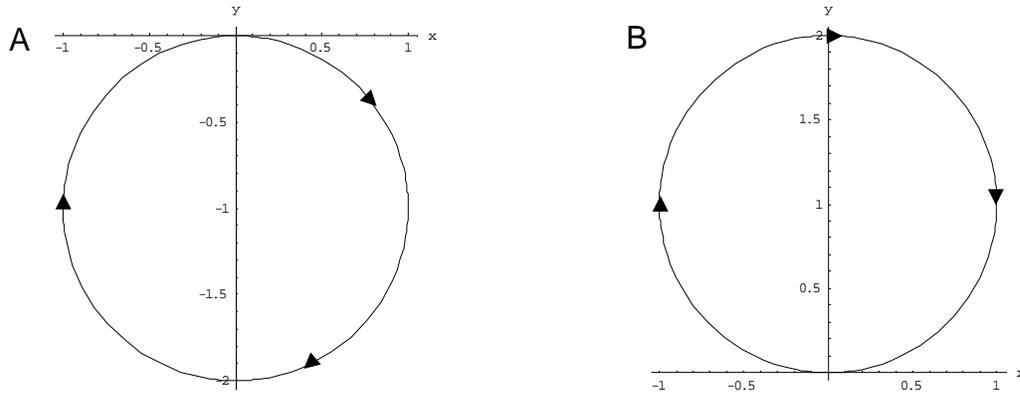


Figure 2-3. Motion of the ion in a uniform magnetic field, for the conditions when A) $\gamma = 0$ and B) $\gamma = \pi$. For simplicity the values for v_p and ω_c were taken as unity.

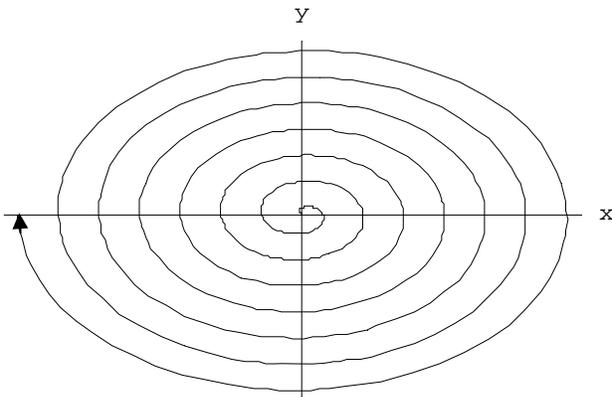


Figure 2-4. The motion of an ion (in a uniform magnetic field) as its cyclotron motion is excited by an alternating electric field.

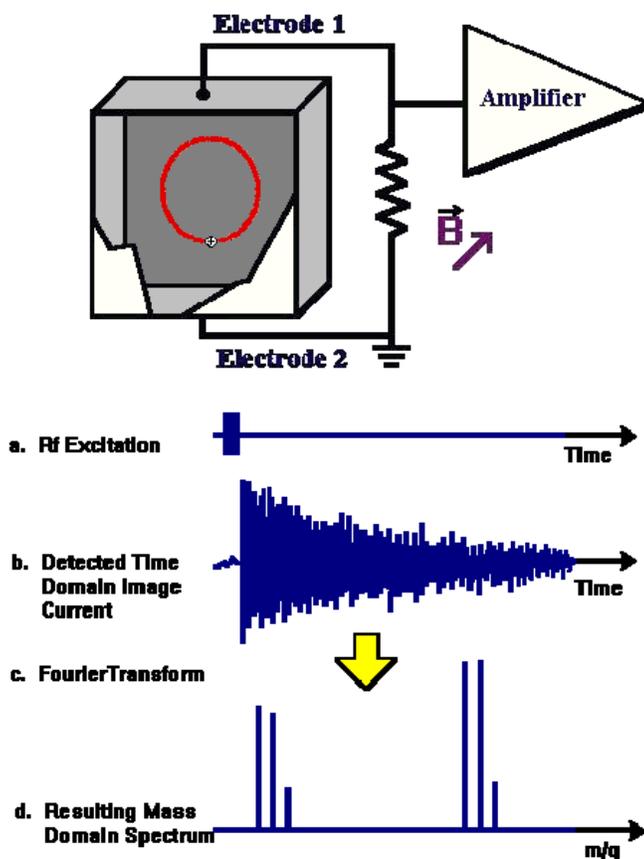


Figure 2-5. Schematic for detection of ions in an FTICR-MS. Image courtesy of Environmental Molecular Sciences Laboratory website.⁷⁴

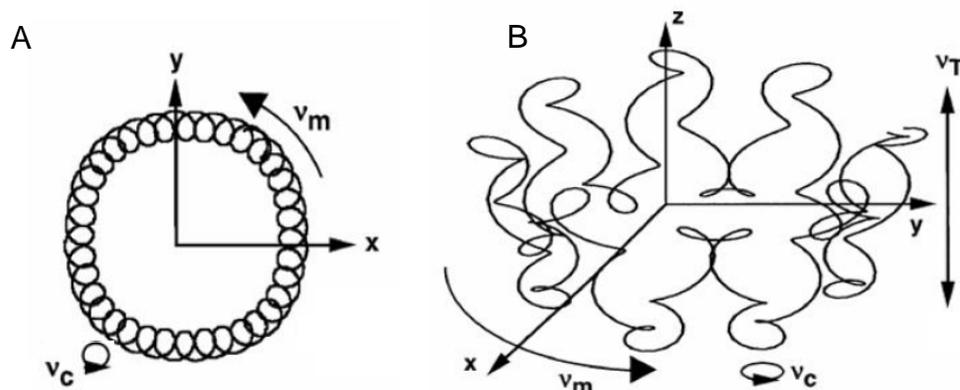


Figure 2-6. A) The superimposed cyclotron and magnetron motions, and B) the superimposed cyclotron, trapping and magnetron motions from “Fourier transform ion cyclotron resonance mass spectrometry: A primer,” by Marshall *et al.*, *Mass Spectrometry Reviews*, Copyright © 1998.³² Reprinted with permission of John Wiley & Sons, Inc.

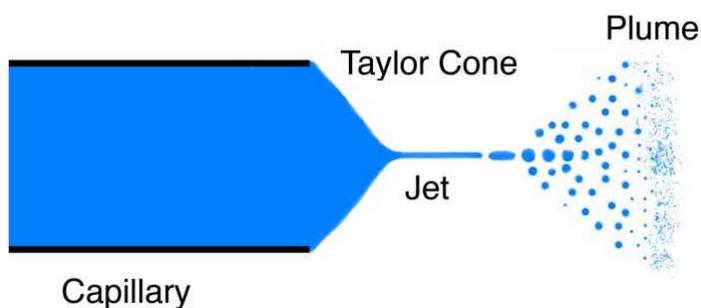


Figure 2-7. Electrospray mist. Reproduced by permission from Creative Commons.⁷⁵

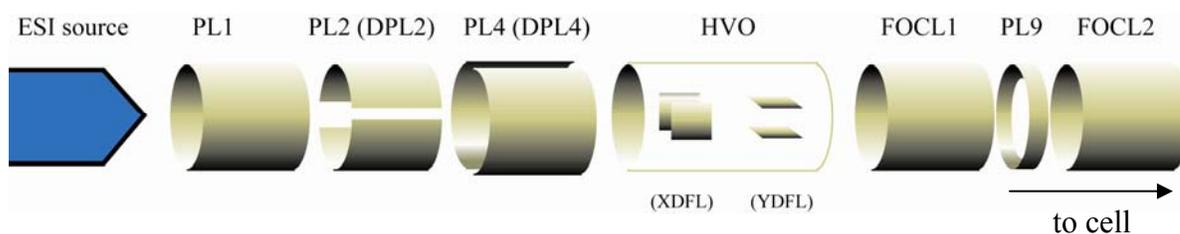


Figure 2-8. Drawing of ion transfer optics and names of each optic component.

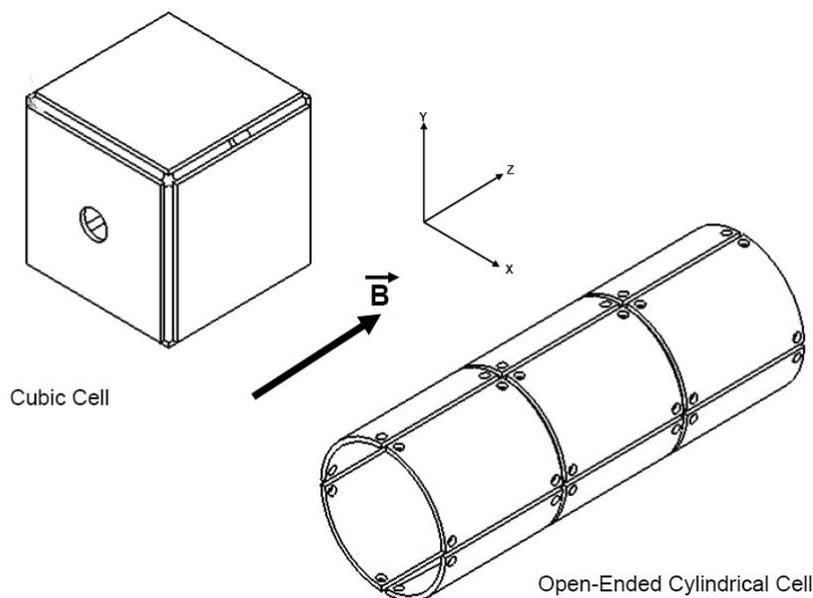


Figure 2-9. Two common Penning traps, a cubic cell and an open-ended cylindrical cell are shown orientated relative to the uniform magnetic field. Reproduced by permission of Valle *et al.*¹⁷

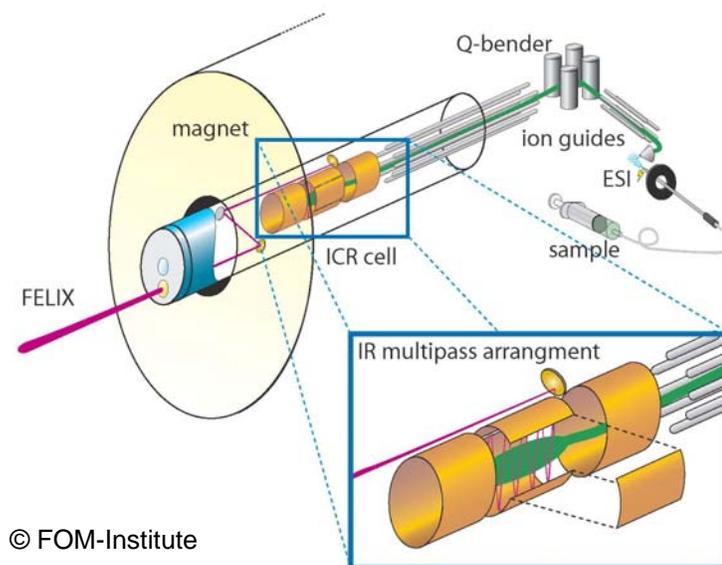


Figure 2-10. FTICR mass spectrometer at the FOM-Institute. The open-ended cylindrical cell is shown in the inset. Drawing courtesy of Dr. Jos Oomens.

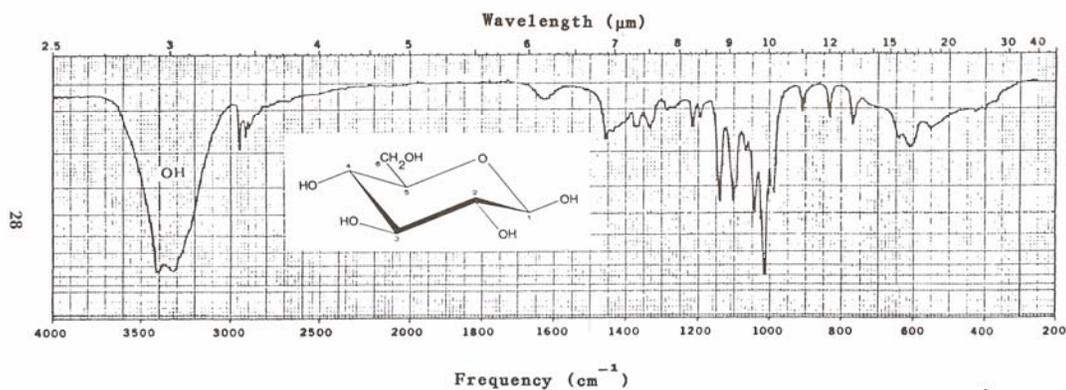


Figure 2-11. Infrared spectrum of a crystalline sample of β -D-glucopyranose in a KBr pellet. Images from *Carbohydrate Chemistry: Monosaccharides and their Oligomers*, by El Khadem *et al.*, Copyright © 1997.³ Reprinted with permission of Academic Press, Inc.

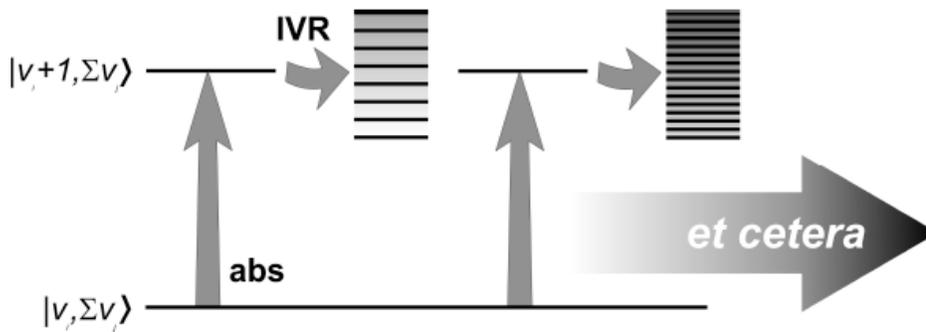


Figure 2-12. The initial absorption of photon energy from the ground state to an excited vibrational level and the subsequent intramolecular vibrational relaxation process.

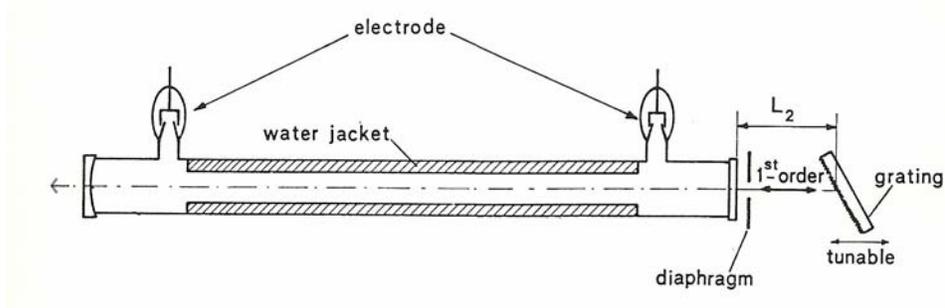


Figure 2-13. The CO₂ laser cavity. Shown are electrodes for gas discharge, the water jacket part of the cooling system, diffraction grating and a diaphragm to select one harmonic of a lasing mode. Brewster window is not shown. Adapted from *The CO₂ Laser*, by Witteman *et al.*, Copyright © 1987.⁵⁶ Reprinted with permission of Springer-Verlag, Inc.

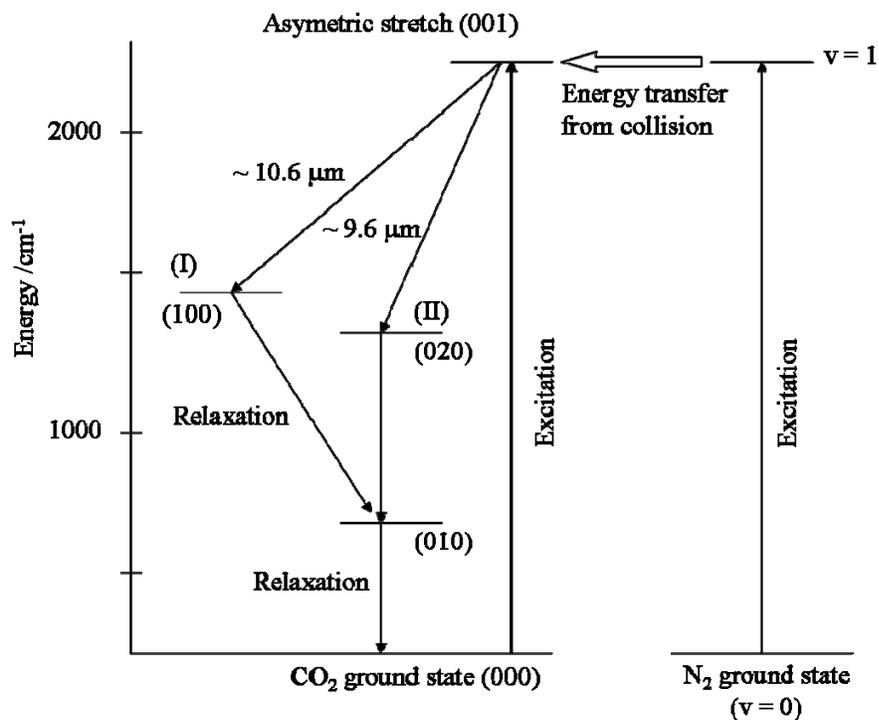


Figure 2-14. Vibrational energy levels of the CO₂ and N₂ molecules involved in the lasing mechanism. Excitation of CO₂, by collision with N₂, activates the asymmetric stretch (001). Photons are released when the molecule relaxes down to (I) symmetric stretch or (II) one of the angle-bending modes. Only the fundamental vibrations are shown, but each mode is coupled with rotations, and relaxation can occur to a rotational-vibrational mode with a slightly different wavelength than the fundamental.

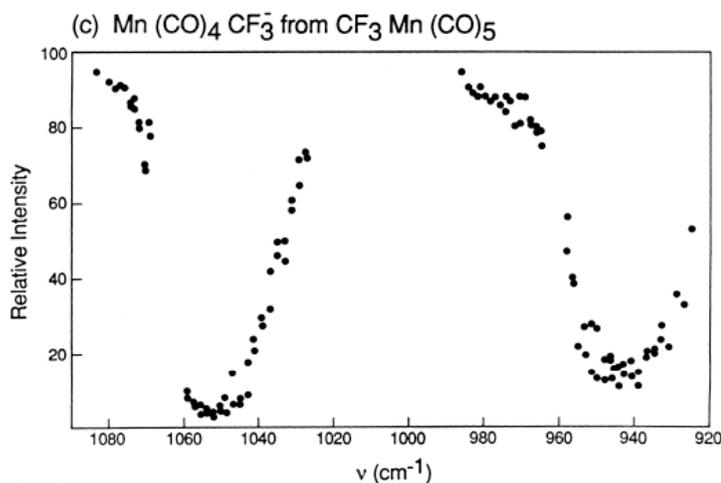


Figure 2-15. IRMPD spectrum of Mn(CO)₄CF₃⁻, using a tunable CO₂ laser. Image reproduced from "Infrared multiphoton dissociation spectrum of CF₃Mn(CO)₃(NO)⁻," by S. K. Shin and J. L. Beauchamp, *Journal of American Chemical Society*, Copyright © 1990. Reprinted with permission of American Chemical Society.

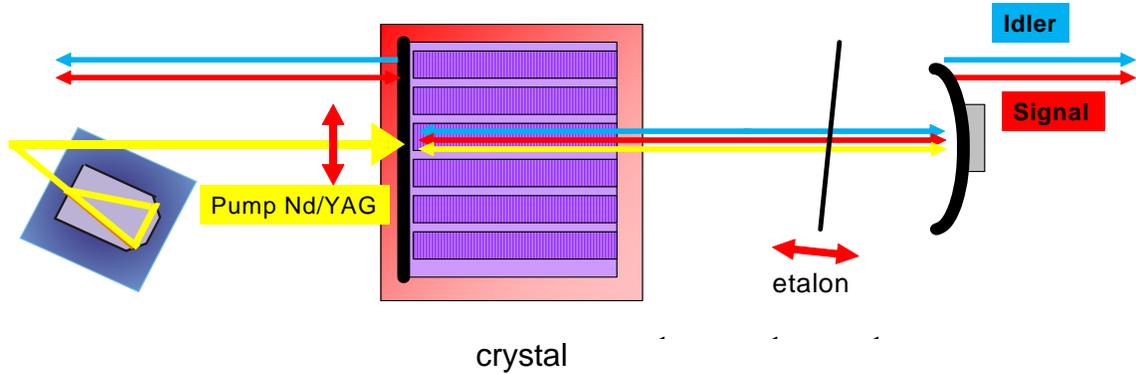
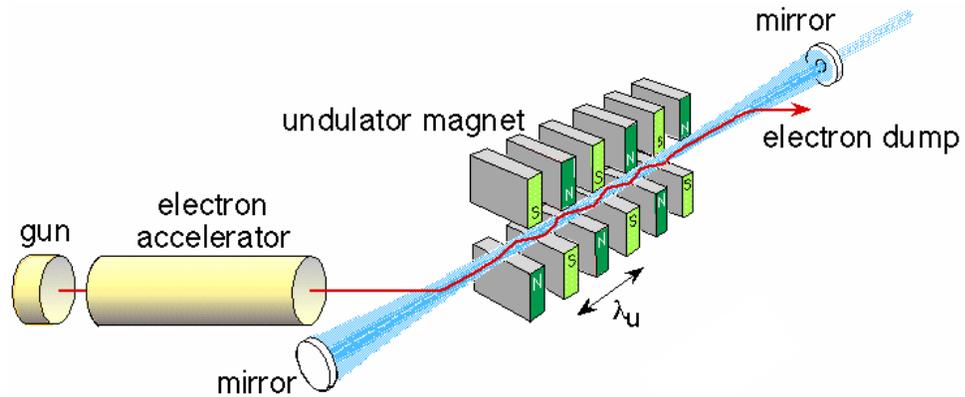


Figure 2-16. Diagram of an OPO laser, showing the pump beam of an Nd/YAG laser (yellow) being converted into the idler (blue) and signal beams (red) by a periodically poled LiNbO₃ crystal. The resonance cavity is defined by one semi-reflective mirror (right end) and a reflective coating on the surface of the crystal (left end). An intracavity etalon is shown, that helps tune a particular set of idler and signal wavelengths. Image courtesy of Wright L. Pearson and LINOS Photonics, Inc.



© FOM-Institute

Figure 2-17. Free electron laser schematic, courtesy of the FOM-Institute.

CHAPTER 3 THEORETICAL CALCULATIONS

Molecular Modeling

Theoretical chemistry is crucial in obtaining structural information from an IRMPD experiment and can also serve to probe the dynamics of the molecular species. In order to obtain the theoretical vibrational spectrum of a species, a number of points need to be considered. First, the correct structure needs to be found, or at least one that will reproduce the spectrum within reasonable error.⁷⁶ The existence of multiple minima on the potential energy surface of a molecule makes finding the most stable structures the first priority when trying to obtain theoretical vibrational frequencies of carbohydrates and amino acids. This requires exploring the potential energy surface of the dissociating ion to find a global minimum structure.

Assignment of vibrational modes in the IRMPD spectrum of a molecular ion requires that either 1) the bands are well resolved and/or can be deconvoluted to obtain all $3N-6$ vibrational modes, or 2) theoretical calculations reproduce the experimental infrared spectrum. Option 1 is difficult to obtain with action spectroscopy for a number of reasons.

- Typical IRMPD experiments are conducted at room temperature, where a range of excited rotational energy levels couple with vibrations of the molecule so each band is slightly broadened.
- The conditions of FTICR trapping lead to magnetron and trapping motions of the ions which produce a spatial distribution of the tightly-packed ion cloud, so that non-uniform irradiation of the ions results.
- Fluctuation in ion signals due to inherent electronic noise will introduce an uncertainty that will be reflected in the IRMPD spectrum.
- Some modes may be degenerate, and cannot be resolved with the current resolution of the FTICR-MS/IRMPD experiment.

An example would be C-H stretches that have similar molecular environments such that they have the similar stretching frequencies. This is evident in sugars, where the carbons 2 – 4 of

a six-membered D-glucoside can have similar electron density throughout the region, and the C-H stretches form one band around 3000 cm^{-1} with a full width at half maximum of 250 cm^{-1} .

If correction of these conditions were possible, assignment of bands would continue to be more time consuming without the use of theoretical calculations. The following section will introduce methods to obtain theoretical structures and their vibrational spectra. The general Hartree-Fock (HF) and density functional theory (DFT) methods in terms of the Hamiltonian will be described. Gaussian or Pople type basis sets used to describe the wavefunctions of the molecule will be introduced, focusing on those that are used to describe carbohydrates and amino acids. The empirical and semi-empirical force fields parameters that were used for the random torsion angle search algorithm will be discussed as an efficient way to obtain multiple conformations of the molecule. The conformational algorithm itself will be subsequently summarized. Finally, obtaining the theoretical infrared spectra will be discussed.

Theoretical Background

Both quantum and molecular mechanics methods examine potential energy surfaces with single point calculations, geometry optimizations and molecular dynamics simulations. In quantum mechanics, the Schrödinger equation gives the wavefunctions and energies of a molecule.

$$H\Psi = E\Psi \quad (3-1)$$

The term H is the Hamiltonian, Ψ is the wavefunction and E is the energy. The Schrödinger equation is an eigenvalue equation of the form in Equation 3-2.

$$Of = cf \quad (3-2)$$

The operator, O , acts on the function (f) producing an eigenvalue and the function itself. The molecular Hamiltonian is composed of the following operators:

$$H = K_N + K_e + U_{NN} + U_{ee} + U_{Ne} \quad (3-3)$$

K_N is the kinetic energy of the nuclei (N), K_e is the kinetic energy of the electrons (e), U_{NN} is the nuclear-nuclear repulsion term, U_{ee} is the electron-electron repulsion term and U_{Ne} is the attractive potential between nuclei and electrons. Nuclei are many times more massive than electrons, and the motion of the nuclei can be considered fixed relative to the motion of the electrons. This approximation, termed the Born-Oppenheimer approximation, indicates that the electron distribution is only dependent on the fixed positions of the nuclei and not their velocities. The molecular Hamiltonian then is simplified to only electronic terms, and is called H_{elec} .

$$H = H_{elec} = K_e + U_{ee} + U_{Ne} \quad (3-4)$$

This allows the electronic portion and the nuclear portion of the Schrödinger equation to be solved independently. Solving for the electronic portion produces an effective nuclear potential, E^{eff} that depends on nuclear coordinates and describes the potential energy surface for the system.

$$H_{elec}\psi_{elec} = E^{eff}\psi_{elec} \quad (3-5)$$

The effective nuclear potential is also used as part of the nuclear Hamiltonian (H_{nucl}) in the Schrödinger equation of nuclear motion to describe vibrations, translations and rotations of the nuclei.

Many different wavefunctions are solutions to the Schrödinger equation, corresponding to the different stationary states of the molecule. The wavefunction must meet the following two requirements for the eigenvalue equation to be solvable: 1) The ψ should be normalizable, so that integration over all space gives the probability density of all electrons.

$$\int_{-\infty}^{+\infty} |\psi_i^2(r)| dr = 1 \quad (3-6)$$

The probability density for all electrons within a system is given by ψ^2 and r is an arbitrary coordinate. 2) The wavefunction needs to be *antisymmetric*, since all electrons are indistinguishable but cannot occupy the same space.

$$\psi(i, j) = -\psi(j, i) \quad (3-7)$$

The wavefunction of the electronic Schrödinger equation is termed a molecular orbital of the molecule and each molecular orbital is approximated with a linear combination of atomic orbitals.

$$\psi_i = \sum_v c_{vc} \phi_v \quad (3-8)$$

This wavefunction is not *antisymmetric* and would not constitute a solution to the Schrödinger equation. The electron spin has yet to be included in the description of the electron. For simplicity, electrons can have spin up (+1/2) or spin down (-1/2) and are designated α and β , respectively. For each molecular orbital, two electrons of opposite spin can occupy one orbital. A determinant of the spin-orbit functions will form an *antisymmetric* wavefunction. For a closed shell molecule, the determinant is represented in Equation 3-9 where each row represents all possible assignments of an electron i to all orbital-spin combinations.

$$\psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\bar{r}_1)\alpha(1) & \phi_1(\bar{r}_1)\beta(1) & \dots & \phi_{\frac{n}{2}}(\bar{r}_1)\alpha(1) & \phi_{\frac{n}{2}}(\bar{r}_1)\beta(1) \\ \phi_1(\bar{r}_2)\alpha(2) & \phi_1(\bar{r}_2)\beta(2) & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \phi_1(\bar{r}_n)\alpha(n) & \phi_1(\bar{r}_n)\beta(n) & \dots & \phi_{\frac{n}{2}}(\bar{r}_n)\alpha(n) & \phi_{\frac{n}{2}}(\bar{r}_n)\beta(n) \end{vmatrix} \quad (3-9)$$

Basis sets. The linear combination of atomic orbitals that make up the molecular orbital can themselves be approximated with a finite set of one electron basis functions.

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} \quad (3-10)$$

N is the number of finite basis functions and the $c_{\mu i}$ are molecular orbital expansion coefficients.

Electronic structure programs most often use Gaussian type basis functions, where the s type function is given in Equation 3-11.

$$g_s = \left(\frac{2\alpha}{\pi} \right)^{3/4} e^{-\alpha r^2} \quad (3-11)$$

The atomic orbital exponent α is different for each atom in the molecule and r is the spherical coordinate that defines the radial extension of the function. Linear combinations of primitive Gaussians (Equation 3-11) are used to form the entire basis function, and are called contracted Gaussians.

$$\chi_\mu = \sum_p d_{\mu p} g_p \quad (3-12)$$

The molecular orbital with contracted Gaussians is given in Equation 3-13.

$$\phi_i = \sum_\mu c_{\mu i} \left(\sum_p d_{\mu p} g_p \right) \quad (3-13)$$

Ab initio

Hartree-Fock (HF) is the most common quantum mechanical method used to obtain properties of molecules. Although not an *ab initio* method, density functional theory (DFT) is included in this section due to its similarity to the Hartree-Fock method.

Hartree-Fock. The Hamiltonian for the one electron wavefunctions just described is designated the Fock operator (F). The Fock operator is a sum of the kinetic energy of a single electron, the potential on the electron by the fixed nuclei and the average potential from the other N-1 electrons. To solve the equation, electronic structure algorithms proceed as follow:

- Evaluate the integrals and account for the atomic positions
- Form an initial guess for the molecular expansion coefficients

- Form the Fock matrix
- Solve the Fock equation
- Repeat the procedure until the wavefunction of one electron is consistent with that of the other electrons in an average field produced by the N-1 electrons

The last step indicates that the wavefunction has converged. The above discussion has dealt with closed shell systems or restricted HF. For open shell molecules, the unrestricted method is used to treat unpaired electrons. The electron can have different spin states and the molecular orbital includes the extra factors, given by Equations 3-14 and 3-15.

$$\phi_i^\alpha = \sum_{\mu} c_{\mu i}^{\alpha} \chi_{\mu} \quad (3-14)$$

$$\phi_i^\beta = \sum_{\mu} c_{\mu i}^{\beta} \chi_{\mu} \quad (3-15)$$

The different spin states produce two sets of orbitals that permit proper dissociation to separate atoms and correct delocalized orbitals. One drawback is that the eigenfunctions are not pure spin states and suffer from spin contamination from higher spin states.

Density functional theory. The wavefunction probability density ψ^2 describes the probability of all electrons in a system. Density functional theory uses the probability density to model electron correlation by using functionals to describe the molecule. From the work of Kohn and Sham, the functionals currently used partition the electronic energy into several terms in Equation 3-16.

$$E = E^T + E^V + E^J + E^{XC} \quad (3-16)$$

$$\rho = E^T + E^V + E^J \quad (3-17)$$

T is the kinetic energy, V is the nucleus-electron attraction and nucleus-nucleus repulsion, J is the electron-electron repulsion and XC represents the exchange-correlation. The terms

$E^T + E^V + E^J$ correspond to the charge distribution of ρ and are similar to the potentials that

were described for Hartree-Fock. The term E^{XC} is the exchange-correlation term and includes a

portion of the electron-electron interactions. The exchange-correlation term is usually divided into two parts, referred to as exchange and correlation, respectively.

$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho) \quad (3-18)$$

All three terms are functionals of the electron density, and many exchange and many correlation functionals have been parametrized to be used for DFT calculations. One of the more popular functionals to use for molecular systems and for carbohydrates and amino acids is the hybrid functional, B3LYP. It is called a hybrid functional since it includes the Hartree-Fock exchange term.

$$E^{XC} = c_{HF} E_{HF}^X + c_{DFT} E_{DFT}^{XC} \quad (3-19)$$

In general DFT calculations proceed similar to HF calculations, except with the added E^{XC} term.

Force fields

In molecular mechanics methods the electron energy is represented explicitly through force field parameters. Molecular mechanics force fields use the equations of classical mechanics to describe the potential energy surface. Simple analytical functions are used to describe the molecule. For example, the compression and stretching of a bond is described as a simple harmonic oscillator, and is analogous to a mass on a spring.

$$V = \frac{1}{2}k(r - r_o)^2 \quad (3-20)$$

The potential energy is V , k is the force constant and r_o and r are the equilibrium and displacement distances of the atoms in the bond, respectively. Both k and r_o are known quantities for a pair of atoms and are called force field parameters. The known quantities have been obtained by fitting the force field to experimental data for specific types of atoms. Some recent modifications to force fields have included the use of *ab initio* methods to obtain the fitting parameters. The potential energy of the molecular system is the sum of all the individual

components of the potential. Using the AMBER (Assisted Model Building and Energy Refinement) force field, the components of the potential describe bonds, angles, torsions, improper torsions, van derWaals interactions, electrostatics and hydrogen bonding (lone pair interactions are included for sulfur).

$$E_{AMBER} = E_{bond} + E_{angle} + E_{dihedral} + E_{vdW} + E_{static} + E_{Hbond} + E_{lonepair} \quad (3-21)$$

$$E_{bond} = \sum_{bonds} k_r (r - r_0)^2 \quad (3-22)$$

$$E_{angle} = \sum_{angles} k_\theta (\theta - \theta_0)^2 \quad (3-23)$$

$$E_{dihedral} = \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \phi_o)] \quad (3-24)$$

$$E_{vdW} = \sum_{ij \in vdW} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right] \quad (3-25)$$

$$E_{static} = \sum_{ij \in static} \left[\frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (3-26)$$

$$E_{bHbond} = \sum \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right] \quad (3-27)$$

Other force fields have similar energy components and some may add an extra term in a component. All force fields have a different set of parameter values for all the constants.

The force fields that have been modified to study carbohydrates are molecular mechanics (MM),⁷⁷ AMBER,⁷⁸ OPLS^{79,80} and CHARMM.⁸¹ MM and AMBER force fields have been used extensively to model large proteins, peptides and carbohydrates, since these molecular systems are too large to be studied with *ab initio* methods. The classical mechanic force fields cannot describe interactions that are purely quantum mechanical. Nevertheless, these force fields can be used to obtain an initial set of structural isomers that can be analyzed by *ab initio* methods.

Geometry optimization

The geometry optimization is a minimization of the energy in the Schrödinger equation used to describe the molecule. It is implemented in robust computer molecular modeling software, and at the core it has the following steps:

- Define an adequate basis set to describe the wavefunction of the molecule
- Solve the eigenvalue equation by minimizing the eigenvalue (energy) that is described by the wavefunction through the variational principle
- Use an algorithm that steps the atoms' positions in the direction that lowers the total potential energy. At each step the eigenvalue is minimized and a new optimized wavefunction is obtained
- Vary the atoms' position toward the minimum of the potential energy surface, until the energy difference from one step compared to that of the next step is zero, or more practically, a set threshold limit

The algorithm implemented in software uses any number of trajectory finding methods to numerically minimize the energy.⁸²⁻⁸⁴ This is distinct from the energy minimization to obtain the normalized wavefunction. The algorithm can be thought of as, although not strictly, following the potential energy surface of the molecule to find the minimum of the potential well. The algorithm is designed so that if the $E_{\min}(n+1) > E_{\min}(n)$, where n is the optimization step, then the step will be discarded and the calculation will return to the previous atomic positions of $E_{\min}(n)$ and a new step will be taken until $E_{\min}(n+1) \sim E_{\min}(n)$, and the energy is said to have converged to a stable point. The derivative, $dE/dq \sim 0$ is a stationary point on the potential energy surface. A stable point can be defined to be a minimum, a maximum or an n th order saddle point. The second derivative of the equation is obtained to indicate that the stationary point is a minimum on the potential energy surface. The second derivative also describes the motion of the nuclei along the potential energy surface and is used to calculate molecular vibrations.

Vibrational analysis

The equilibrium positions of the atoms are determined by solving the electronic Schrödinger equation. The kinetic energy (T) of nuclei about their equilibrium positions is given in Equation 3-28.

$$T = \frac{1}{2} \sum_{i=1}^{3N} \left(\frac{\partial q_i}{\partial t} \right)^2 \quad (3-28)$$

$$q_i = m_i^{1/2} (x_i - x_i^{eq}) \quad (3-29)$$

Here q_i are the mass-weighted Cartesian displacement coordinates (shown only for the x coordinates in Equation 3-29) and N is the number of nuclei in the molecule. The potential energy (U) can be expanded as a Taylor series about the equilibrium positions, and to a first approximation only the quadratic terms are used. Although the energies can be evaluated quantum mechanically, the vibrational motions can be solved using Newton's second law (the first derivative term vanishes at the equilibrium position).

$$U = U_{eq} + \sum_{i=1}^{3N} \left(\frac{\partial U}{\partial q_i} \right)_{eq} q_i + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left(\frac{\partial^2 U}{\partial q_i \partial q_j} \right)_{eq} q_i q_j + \dots \quad (3-30)$$

$$U = U_{eq} + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left(\frac{\partial^2 U}{\partial q_i \partial q_j} \right)_{eq} q_i q_j \quad (3-31)$$

By approximating to the second order, one can obtain harmonic approximations of the vibrations, and until recently 2nd order approximations were the major method implemented in quantum chemistry codes to solve for the normal modes of molecular systems. The eigenvalue equation uses the wavefunction and potential energy surface obtained from the geometry optimization, starting at the converged point and perturbs the system by moving the atoms along

the potential energy surface. The normal modes will be used to obtain calculated spectra to compare to experimental IRMPD data.

Molecular Dynamics

Carbohydrates are complex molecules that can have more than one local minimum. For example, a hexose monosaccharide can exist in a linear form, but can also be in a five or six-membered ring structure. A monosaccharide like glucose can have up to five chiral centers, whose functional groups have rotatable bonds. The six-membered ring structure can have different configurations, with chair, boat, envelope, twist, and skewed the most stable configurations, but others are possible. A study of L-ideronate had to include the 4C_1 , 1C_4 and 2S_0 ring structures to obtain the average structure from the experiment.⁸⁵ Although not as complicated to model, amino acids are also linear chain polyatomics and have multiple conformations. There is a set of robust methods to deal with global minimum optimization or the multiple minima problem, most of which involve molecular dynamics theory implementation in chemistry computer algorithms and codes. The most common are simulated annealing,⁸⁶ replica exchange molecular dynamics(REMD)⁸⁷ and dihedral angle searching.⁸⁸ Although not reported in this work, simulated annealing will be summarized to contrast to the dihedral angle search method.

Simulated annealing

Simulated annealing requires that a molecule start at a local minimum so that its energy can be increased to extract the molecule from its potential well, in a state well above the local minimum. Exploring the conformation space to find the global minimum is done by taking the geometry optimized starting structure and increasing the system temperature above 300 K and even higher than 1500 K, depending on the molecule size and flexibility. Heating occurs for only a few ps (in the time scale of the calculation and not real time). By virtual “heating” of the

glycoside, local minima barriers will be overcome, and thus the molecule will sample many states on the potential energy surface. Controlled cooling allows the molecule to find lower energy configurations than the initial configuration, and given sufficient simulation time the global minimum can potentially be reached.⁸⁹ To find the global minimum, the system is slowly cooled to 0 K for the rest of the simulated annealing calculation. Adequate cooling time is 100 – 1000 ps depending on molecular size. The time step of the simulation is crucial and should be on the order of a few femtoseconds since the integration step needs to be smaller than the vibrational frequencies or the integration will be inadequate. The major drawback to using this method is the short time step in the simulation. For a 30 atom monosaccharide, using an empirical force field, a simulated annealing run for 1 ns of simulated time will take approximately 15 days on a Pentium IV 1.3 GHz processor with 500 MB of RAM. While codes exist that will allow for multiple cpu processing, this usually requires access to a supercomputer or a network cluster of computers. If using *ab initio* molecular dynamics, large amount of computing time will be required even while using high performance multiple processor systems. Similar to simulated annealing, REMD requires a small time step for proper integration. REMD also requires a large amount of computational power since multiple simulations need to be processed simultaneously, and it is mainly implemented in algorithms for use on high performance multiple processor systems.

Since this is the first step in obtaining the theoretical vibrational frequencies of the sugars and amino acids in this work, the bulk of the time should not be spent finding different conformations. Therefore the computational method needs to proceed quickly in obtaining different conformations, so that they can be further analyzed with *ab initio* methods to describe

the amino acids and carbohydrates adequately. A method that will search the conformational space adequately while keeping computational time manageable is needed.

Conformational searching

Torsion angle searching is a global minimum search method that provides multiple conformations so that they can be compared with each other to find the lowest energy structure. A strict conformational analysis (CA) searches all the dihedral angles of the molecule, incrementing each by a few degrees at a time and at each point a geometry optimization is performed. For an ion with a few torsion angles, because of high symmetry or because it contains only a few atoms, the conformational analysis (CA) will be relatively simple. But for molecules with many degrees of freedom (e.g., glycosides), a quantum mechanical conformational analysis of all torsion angles can be extremely computationally expensive, even at low levels of *ab initio* theory.

Random conformational searching is a method where the dihedral angles can be altered by a software algorithm that emphasizes finding low energy structures. For faster results while using a single processor, the empirical or classical mechanics force fields are employed. *Ab initio* or density functional geometry optimization is done only in the last step of the overall process, to overcome the limitations of the empirical and semi-empirical force fields. The algorithm as implemented in the Hyperchem suite of programs is based on the usage directed approach⁹⁰ and has the following general scheme:

- Definition of dihedral angles by the user
- Randomization of the dihedral angles to create a new structure
- Geometry optimization for the new structure
- Comparison of the energy converged structure to the set of other structures that have been made by the randomization step, to check for duplicate structures and high energy species

The algorithm relies on comparing the structures created in the search to verify that all of the structures are distinct from one another. The spatial orientation of torsions angles and bonds are compared to a specified precision to check for the duplicates. The total energy obtained from the geometry optimization is compared to that of other structures and if they are within a specified limit, the structure may be considered the same if the other criteria are also met. The usage-directed method proceeds depending on the following scenarios:

- If the structure is found to be the same as one of the others in the search, then it is tabulated as such and the search continues with this structure being used for the randomization of the dihedral angles
- If the structure is distinct from any of the others, and is within the total energy limit, then it is tabulated as a new structure and the search continues with this structure being used for the randomization of the dihedral angles
- If the structure is distinct but its total energy is higher than a user specified limit, the structure is discarded and the search continues by applying randomization of the dihedral angles to the previous structure that was accepted by the search

The number of created structures can be modified, to keep the computation time reasonable (default value is 1000 structures). A table is created with all the distinct conformers arranged from lowest to highest energy with the final values for the dihedral angles of each structure are listed. Individual conformers are extracted for further analysis with higher levels of theory.

CHAPTER 4 STRUCTURE DETERMINATION OF PHENYLALANINE ANALOGS

Introduction

Hydrogen deuterium exchange (H/D) is the method in which particularly labile hydrogens atoms are substituted with deuterium atoms in the molecule or ion of interest.⁹¹ H/D exchange has been used in solution phase studies, mainly NMR work, to obtain structural information.⁹² H/D exchange studies have also been conducted in the gas phase and early work from Beauchamp, *et al.* probed the proton affinity and exchange rates of labile hydrogen atoms.⁹³ Ion-molecule reactions involving H/D exchange have also been followed in mass spectrometry experiments, for example to investigate the dynamics of unsaturated platinum complexes.⁹⁴ Utilizing the technique of infrared multiple photon dissociation (IRMPD),⁵⁸ infrared spectra of gas phase deuterated ions can be obtained, and these have recently been used to obtain structural information for small ions^{59,69,95-98} and biologically relevant ions^{66,99-103} by observing the spectral shifts that occur in the infrared spectra due to deuteration.

H/D exchange experiments can provide a great deal of structural information, especially when coupled with mass spectrometric and/or infrared spectral techniques. Substitution of hydrogen by deuterium has primarily been used to obtain information on the tertiary structure of proteins,¹⁰⁴ but can also provide primary structural information for smaller species such as the phenylalanine analogs studied in this work. In the case of amino acids, exchange is seen for hydrogen atoms bound to the Lewis base oxygen and nitrogen atoms.¹⁰⁵

Phenylalanine (Phe), one of the essential amino acids, has been studied extensively both in solution and in the gas phase.^{14,101,106-111} Recent work has shown that Phe behaves differently in each phase, in part due to structural differences of the molecule in the two phases.¹⁵ The binding of alkali metals to Phe in the gas phase has been especially well characterized.^{14,106,110}

However, much less is known of the effect that protecting groups or modifications to the basic amino acid composition have on the phenylalanine structure. Recent work by Dunbar and co-workers¹⁰¹ showed that by extending the amino acid chain, by forming a dipeptide, an attached alkali cation has a higher possibility of forming a chelating complex. The sodium cation tends to interact with electronegative sites on the extended chain, subject to steric hindrance constraints. For sodiated PheAla and AlaPhe, the most energetically favorable structures involved a sodium cation - π interaction with the phenyl ring. The phenylalanine moiety tends to pucker to allow the alkali metal to interact with both the phenyl ring and the lone pair electrons on the oxygen or nitrogen atoms. To further investigate the effect that modification of phenylalanine has on the chelating complexes, in the work reported here phenylalanine analogs bound to sodium cations were studied in the gas phase after either solution- or gas-phase hydrogen/deuterium exchange.

Mass spectrometric techniques use low quantities of material and are ideally suited for the low abundance yield of products obtained from organic synthesis of the amino acid analogs. Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS)^{31-35,112-114} has the advantage of trapping ions for extended periods of time so that laser irradiation of the sample can be performed without deactivation of excited ions by collisions with background gas. In this work sodium cations were attached to the phenylalanine analogs in solution, and the resulting charged complexes were desolvated during the electrospray ionization process^{39,42,43} to produce isolated ions. Infrared multiple photon dissociation spectra of the Phe analogs cationized by sodium then provided structural information, when combined with theoretical calculations. Density functional theory (DFT)^{76,115,116} calculations were used to predict the lowest energy structures of each analog. The hybrid functional B3LYP, which has been shown to be adequate for predicting infrared spectra of Phe,^{72,101} was used for all calculations.

Neutral Phe has three potential hydrogen atoms which can undergo H/D exchange (Figure 4-1). Limiting the deuteration sites to one or two locations, by the addition of protecting groups, allows observed spectral shifts to be clearly defined and simplifies band assignment and subsequent structure elucidation. The analogs of Phe studied in this work had protecting groups on all but one or two of the possible exchange sites. The analogs, N-acetylphenylalanine (AcPhe), O-methyl N-acetylphenylalanine (AcPheOMe) and N-acetylphenylalanine O-methylglycine (AcPheGlyOMe) (Figure 4-2). AcPhe has two possible exchangeable hydrogens (in the O-H and N-H groups), AcPheOMe has one, and AcPheGlyOMe has two.

Experimental

Sample Preparation

Phe analogs were synthesized from a commercial sample of L-phenylalanine (Sigma-Aldrich Co.) by Mr. Alfred Chung at the Proteomics Division of the Interdisciplinary Center for Biotechnology Research at the University of Florida. Stock solutions were made by dissolving 0.01 mg/mL of the analogs in an 80:20 CH₃OD:D₂O solution.

For solution phase H/D exchange, a solution of 80:20 CH₃OD:D₂O was used to dilute the stock sample to 1 mmol, and an equimolar amount of NaCl was added. Samples were allowed to undergo exchange for a minimum of 20 minutes followed by electrospray introduction into the FTICR-MS. For gas phase exchange, ND₃ was introduced into the ICR vacuum chamber at a background pressure of 5E-7 torr, and allowed to react with the cationized Phe analogs for 1-5 seconds before laser irradiation occurred. For species where no gas phase exchange was detected, reaction time was increased up to 35 seconds to verify that no reaction had taken place.

Instrumentation

To deduce sites of sodium interaction and to assist in ionizing the peptide molecules, the phenylalanine analogs were electrosprayed with sodium, producing the amino acid - sodium

cation complexes. Electrospray ionization was conducted with a Z-spray source (Micromass/Waters Corporation, Milford, MA) at a solution flow rate of 20 $\mu\text{L}/\text{min}$, with the flow rates of the nebulizing gas and the desolvation gas (N_2) set to 32 and 150 l/hr, respectively, and the electrospray needle-skimmer voltage difference set at 3 kV. A 4.7T superconducting magnet (Cryomagnetics Inc., Oak Ridge, TN) was used with a laboratory-constructed pumping system, ion trap and electronics console. The FTICR-MS system has been described previously⁶⁶. Precursor ions were isolated using the stored waveform inverse Fourier transform (SWIFT)^{73,117-120} technique to eject all unwanted ions. Ions were detected using the broadband detection mode covering a mass range from 20 to 2000 Da. For the IRMPD activation, the Free Electron Laser for Infrared eXperiments (FELIX)^{65,121,122} was used with the wavelength scanned over the range of 5.5 – 12.5 μm (800 – 1800 cm^{-1}). Mass-selected precursor ions were irradiated in the open cylindrical FTICR ion trap for a period of 4 seconds at each wavelength with an average power of 50 – 60 mJ per macropulse and a laser repetition rate of 5 macropulses per second. Typically, four individual transients were accumulated at each wavelength to improve the signal-to-noise ratio. The depletion signal of the precursor ion was monitored throughout the experiment.

Computational Details

The Hyperchem suite of programs¹²³ was used to assign trial structures to the phenylalanine analogs. Depending on the analog, methyl or acetyl groups were added to a core Phe structure to produce trial AcPhe, AcPheOMe and AcPheGlyOMe structures. Sodium was initially placed near the carboxyl or carbonyl group of the Phe analogs, and the complex was given an overall charge of +1. It should be noted that although the sodium was placed near the carboxyl or carbonyl group, no bond was defined between sodium and any other atom, and

therefore sodium was free to move during the geometry optimization steps of the calculation. Explicit definition of a bond was considered unnecessary since the quantum mechanical AM1 force field was used initially to model the interactions of the analog complexes. The dihedral angles which were varied in the conformational search were explicitly defined (Figure 4-3).

Using the usage directed approach as described in chapter 3,⁹⁰ 1000 structures were created within the Hyperchem software to adequately probe the conformational space of each analog. During the conformational search, the algorithm eliminated duplicate structures by comparing energies, torsion angles, and RMS fit residual errors between corresponding atoms. The comparison threshold values were set to 0.05 kcal/mol, 10° and 0.25 Å, respectively. All conformers within 15 kcal/mol of the lowest energy conformer were further refined by geometry optimization using Gaussian03¹²⁴ at the B3LYP/6-31G(d) level of theory.^{125,126} The conformations were compared and duplicate conformations having the same energy (within 0.0001 hartree), and the same vibrational spectrum as the comparison conformer, were discarded. Using the B3LYP/6-311++G(d,p) level of theory,^{127,128} an average of 45 conformers were further geometry optimized, and a vibrational frequency calculation, at the same level of theory, was performed for each of these conformations. The calculated frequencies were scaled by 0.965, followed by comparison to experimental spectra.

Results and Discussion

N-acetylphenylalanine

Calculated infrared spectra for the four lowest energy conformers (Figure 4-4) of sodium-cationized AcPhe are compared to the experimental IRMPD spectrum obtained for this complex in Figure 4-5. Conformer A has a cation- π interaction (Figure 4-4) while conformer B does not, although the conformers have comparable energies (0.04 kcal/mol). The calculated spectrum of conformer B gives a poorer match to the experimental data relative to that of conformer A.

When compared to the experimental C=O stretching bands (1660 and 1720 cm^{-1}), the calculated C=O stretches for conformer B are red shifted. The interaction between the sodium cation and the carbonyl groups is stronger for this conformer since the cation- π interaction does not occur in conformer B. The C=O stretch frequencies of conformers C and D also do not fit the experimental data as well as those for conformer A. For both conformers C and D, one calculated C=O stretch band is red shifted and one is blue shifted when compared to the experimental spectrum. The calculated structures for conformers C and D have only one carbonyl interacting with the sodium cation and the carboxyl carbonyl group does not interact with either the sodium cation or the phenyl group. The free carboxyl carbonyl stretch shows a large blue shift, while the carbonyl interacting with the sodium cation is red shifted due to a stronger interaction between the oxygen and the nitrogen atoms. Table 4-1 gives the calculated distances between the sodium cation and the oxygens of the n-acetyl and carboxyl groups, indicative of the strength of electrostatic interaction.

With the scaling factor correction, the rms differences between the experimental data and the calculated conformer C=O stretch frequencies are 5 cm^{-1} for A, 20 cm^{-1} for B, 29 cm^{-1} for C and 40 cm^{-1} for D, with frequency values for the experimental spectra taken from the peak maxima of the experimental bands. All other calculated bands for conformers C and D do not agree well with experimental data (except that conformer D shows reasonable agreement with the bands at approximately 1150 and 1525 cm^{-1}). The observed IRMPD spectral bands can be assigned using the calculated results for conformer A (Figure 4-6) although small contributions to the IRMPD spectrum from the other conformers is possible, due to the similar relative energies of the structures. Bands I, II and V would be expected to shift following substitution of

a deuterium atom for a hydrogen atom and are used extensively in the discussion of IRMPD spectra of H/D-exchanged AcPhe analogs below.

Hydrogen/Deuterium Exchange Experiments

AcPhe

For the solution phase H/D exchange of AcPhe, the mass spectrum indicated that, as would be expected, two exchanges had taken place, one involving the amine hydrogen and one the carboxylate hydrogen. The IRMPD spectrum of sodium cation-attached AcPhe following solution phase HDX (Figure 4-7).

In experiments involving gas-phase H/D exchange of sodium cation-attached AcPhe, the resultant mass spectra showed only a 1 Da mass increase, indicating that only one D for H substitution had occurred. An IRMPD spectrum of the singly deuterated complex ion was obtained and the calculated spectra for sodium cation-attached AcPhe with one D substituted for either the amide or carboxylate hydrogen were compared (Figure 4-8). Agreement with the experimental spectrum is much better for the spectrum calculated for D exchange with the carboxylate hydrogen. This conclusion is further supported by the presentation in Figure 4-9, where the spectra for undeuterated and singly-deuterated sodium cation-attached AcPhe are overlaid, with C=O stretching and O-H, O-D, N-H, and N-D bending mode regions indicated.

AcPheOMe

For sodium cation-attached AcPheOMe, no gas-phase H/D exchange was observed. AcPheOMe does not have a carboxylate hydrogen (it has been replaced by the methyl group), and exchange of the amide hydrogen is apparently quite slow in the gas phase (as was seen for sodium cation-attached AcPhe). One H/D exchange took place in solution-phase experiments, presumably at the amide hydrogen, as confirmed by the agreement between the IRMPD spectrum of mono-deuterated sodium cation-attached AcPheOMe and the calculated spectrum

for the amide-deuterated species (Figure 4-10A). The lowest energy sodium cation-attached AcPheOMe structure used for DFT calculations of the mono-deuterated spectrum (Figure 4-10C).

AcPheGlyOMe

Two exchanges of D for H were observed for AcPheGlyOMe in solution, and none in gas-phase experiments. This protected dipeptide has two amide hydrogens, and DFT calculations predict the most stable structure of the sodium cation-attached complex (Figure 4-10D). Agreement between the IRMPD spectrum obtained for the sodium cation-attached doubly-deuterated ion from solution-phase HDX experiments and the calculated spectrum (Figure 4-10B) for the structure in Fig. 4-10D with both amide hydrogens replaced with deuterium is quite good.

The most stable structures calculated for both sodium cation-attached AcPheOMe and AcPheGlyOMe involve interaction of the sodium cation with the phenyl ring of the phenylalanine side chain and lone pairs on all carbonyl oxygens. Figures 4-10C and 4-10D predict that these complexes will have a puckered conformation. The structures in Figure 4-10 correspond to the undeuterated species of AcPheOMe and AcPheGlyOMe, and the corresponding spectra are for the deuterated ions. The optimized structures for the undeuterated ions were used to calculate the spectral shift of the deuterated ions.

Results for H/D exchange experiments involving the Phe analogs are summarized in Table 4-2. The results found in this study and others of phenylalanine indicate that the species behaves quite differently when interacting with a chelating metal, which seems to fix the structure into one conformation. Gas phase H/D exchange has been observed for the amide hydrogen on other amino acids.^{105,129} Direct substitution of the amide hydrogen with deuterium is slow, and earlier studies indicate that if a hydroxyl group is available, the exchange will occur at the hydroxyl

hydrogen followed by deuterium migration to the amide position. Formation of a chelated sodium cation complex apparently impedes migration of deuterium to the amide hydrogen, as seen for sodium cation-attached AcPhe, where exchange occurred at the carboxylate hydrogen but not at the amide hydrogen. In preliminary gas-phase H/D exchange experiments involving protonated (as opposed to sodium cation-attached) AcPhe, three hydrogen atoms were exchanged with deuterium. Deuterium migration from the carboxylate group to the amide nitrogen was not impeded for this species.

Conclusions

Shifts of vibrational spectral bands upon H/D exchange with sodium cation-attached phenylalanine analogs have been predicted theoretically and compared to those observed in gas-phase infrared multiple photon dissociation spectra. The H/D exchange experiments not only provided band assignment confirmation, but simultaneously helped identify structural features of the Phe analogs being studied. Structural characteristics common to all phenylalanine analogs in this study include the following.

- Chelating complexes form when the sodium cation binds to Phe analogs
- The lowest energy structures for all three analogs include cation - π interactions
- Torsion of carbonyls toward the phenyl ring, in order to interact with the sodium cation, causes the conformation to stay relatively rigid and in a puckered state
- Complexation with the sodium cation hinders D for H exchange at the amide group hydrogen(s).

These results differ from those found for other amino acids in the gas phase, where gas phase exchange can be observed at the amine groups.

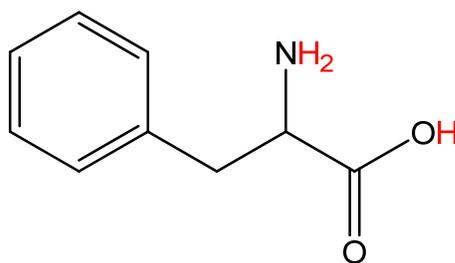
Table 4-1. Interaction distance between sodium cation and oxygen atoms on the carbonyls of the carboxyl and on the N-acetyl functional groups for calculated conformers A - D.

Calculated Distances between Sodium and Oxygen Atoms		
AcPhe Conformer	Carboxyl Carbonyl O	N-acetyl Carbonyl O
A	2.22	2.35
B	2.26	2.17
C	n/a	2.14
D	n/a	2.14

Distance is reported in Angstroms.

Table 4-2. Summary of results from H/D exchange experiments, where no D for H substitution was observed for sodium cation-attached AcPheOMe and AcPheGlyOMe in the gas phase

N-acetylphenylalanine species	Deuteration method	Deuteration sites
Sodiated (AcPhe)	Gas phase	1 – OH
	Solution	2 – OH, NH
Sodiated (AcPheOMe)	Gas phase	None
	Solution	1 – NH
Sodiated (AcPheGlyOMe)	Gas phase	None
	Solution	2 – NH



Phenylalanine (Phe)

Figure 4-1. Phenylalanine with exchangeable hydrogens highlighted in red

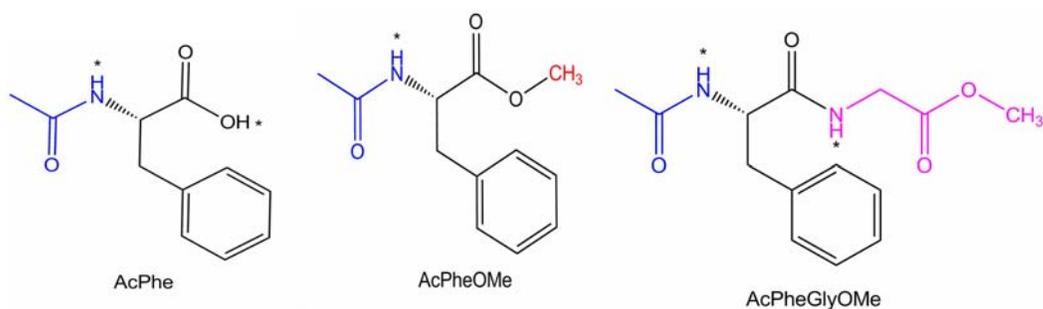


Figure 4-2. Phe analogs studied in this work. The portion highlighted in blue represents the N-acetyl adduct, highlighted in red is the methyl group, and the portion highlighted in purple indicates the O-methylated glycine. Possible exchangeable hydrogen atoms are indicated by asterisks.

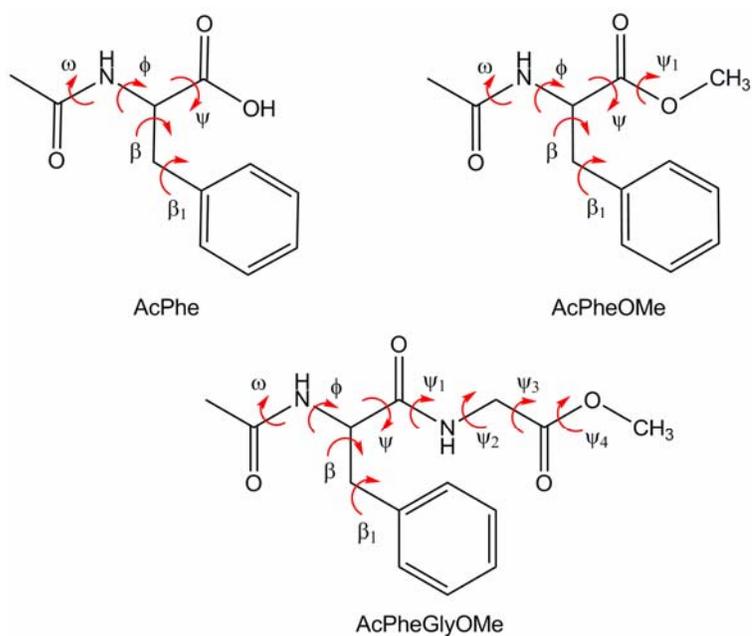


Figure 4-3. Dihedral angles defined for the conformational search calculations of AcPhe, AcPheOMe and AcPheGlyOMe (shown in red arrows). The labelling scheme for the dihedral angles is similar to that of References 133 and 134.^{130,131}

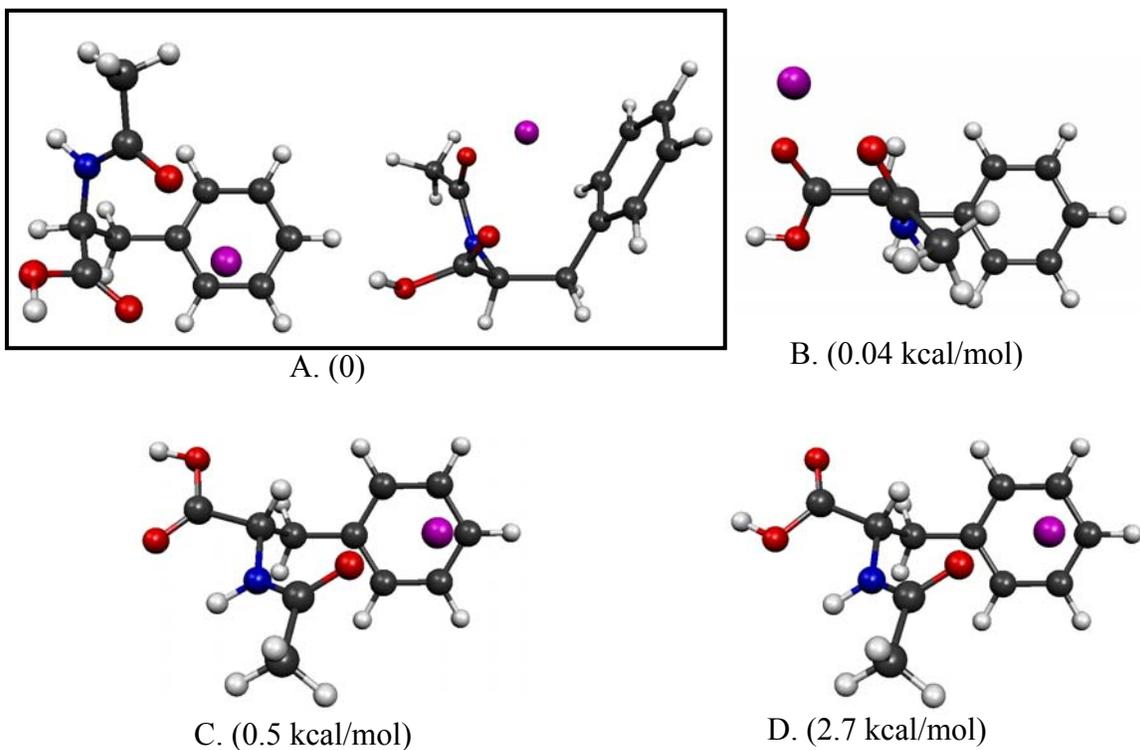


Figure 4-4. Lowest energy conformers of sodium cationized AcPhe found from conformational search calculations. Relative energies (given in parentheses) were obtained at the B3LYP/6-31++G(d,p) level of theory and are zero-point corrected. An alternate view is shown for conformer A to show the puckering of the structure.

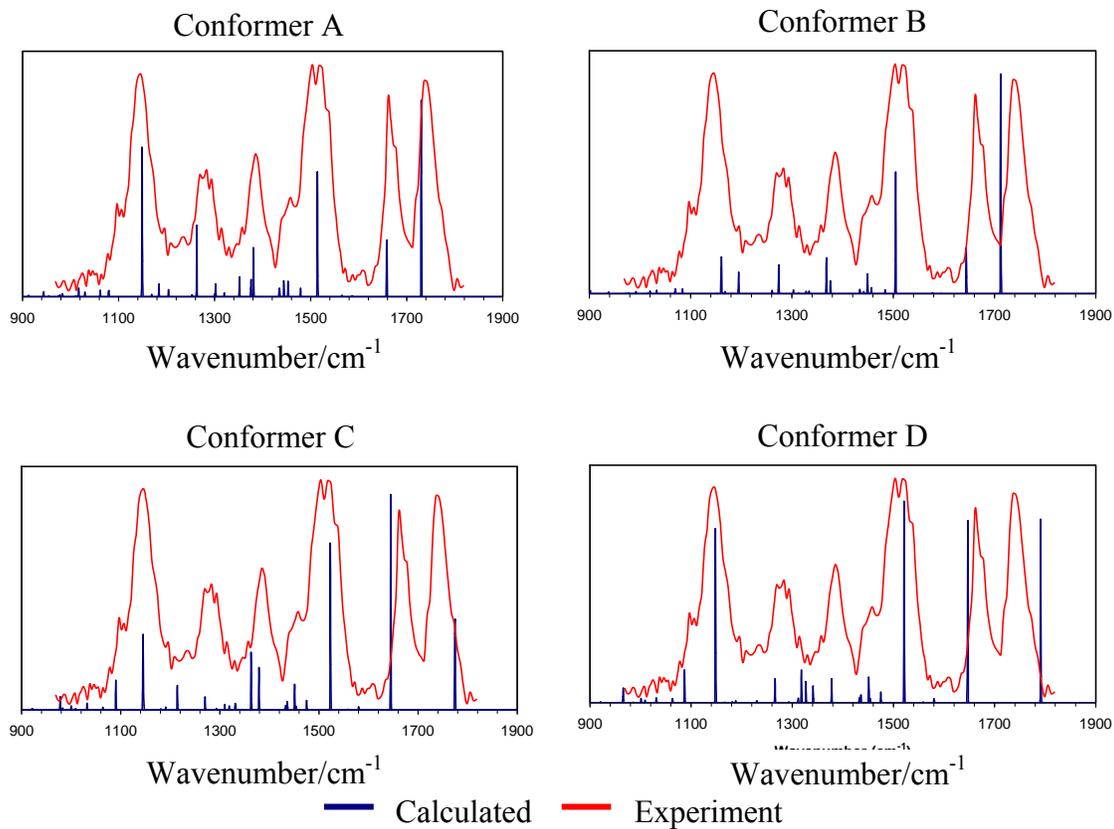


Figure 4-5. Comparison of experimental IRMPD spectrum (red line) of AcPhe with calculated spectra of the theoretically determined lowest energy conformers A-D (blue stick spectra). Calculated vibrational frequencies are scaled by a factor of 0.965.

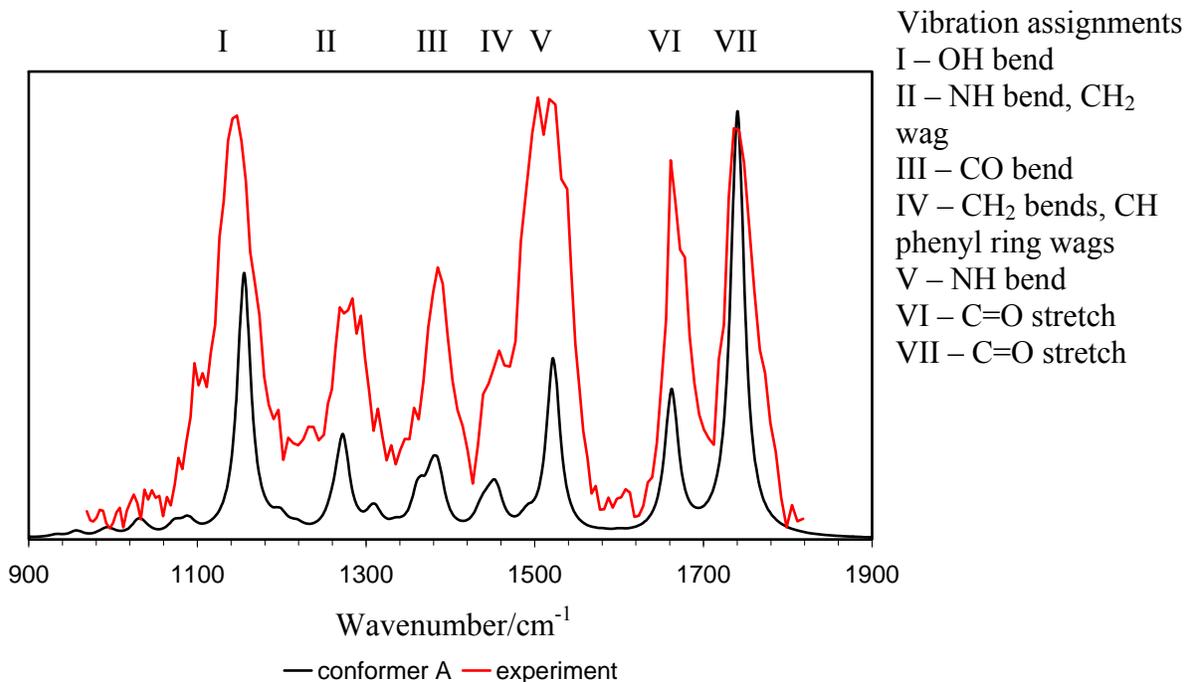


Figure 4-6. Assignment of spectral bands in the IRMPD spectrum of AcPhe using calculated bands for conformer A. The calculated spectrum has been convoluted with a 20 cm⁻¹ Gaussian profile.

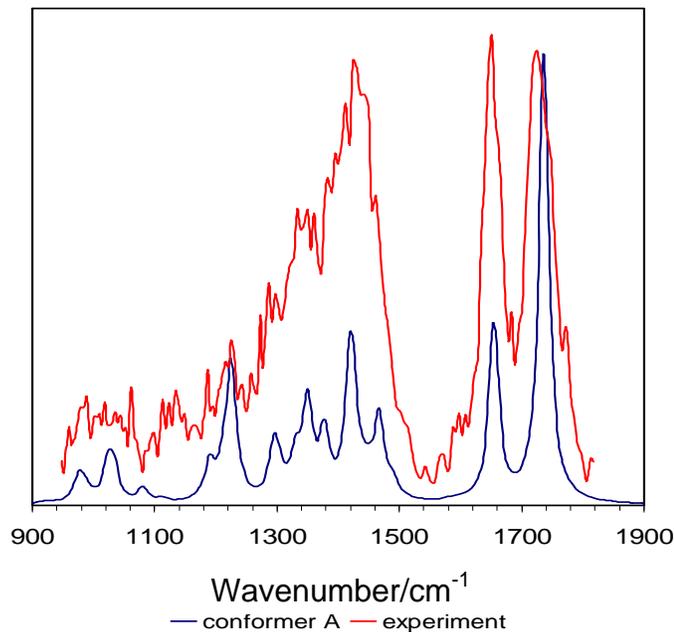


Figure 4-7. IRMPD spectrum of sodium cation-attached AcPhe following solution phase HDX (in red) and calculated spectrum of the doubly deuterated conformer A (in blue). Deuterium substitution occurred at both the N-H and COOH hydrogens.

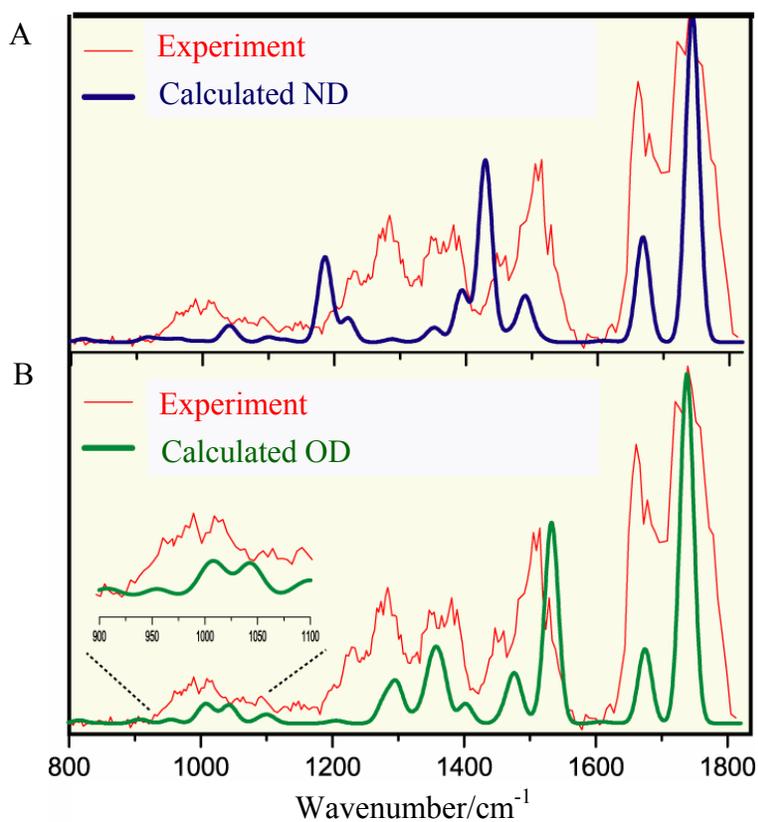


Figure 4-8. Comparison of the singly deuterated sodium cation-attached AcPhe IRMPD spectrum with that of calculated conformer A if exchange had occurred at A) the N-H hydrogen or at B) the O-H hydrogen.

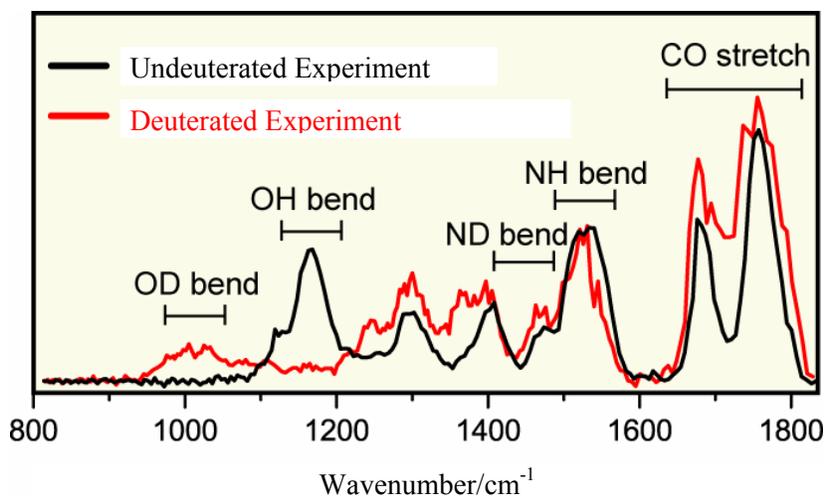


Figure 4-9. Experimental spectra for undeuterated and singly-deuterated sodium cation-attached AcPhe, indicating that deuterium was exchanged with the carboxylate hydrogen.

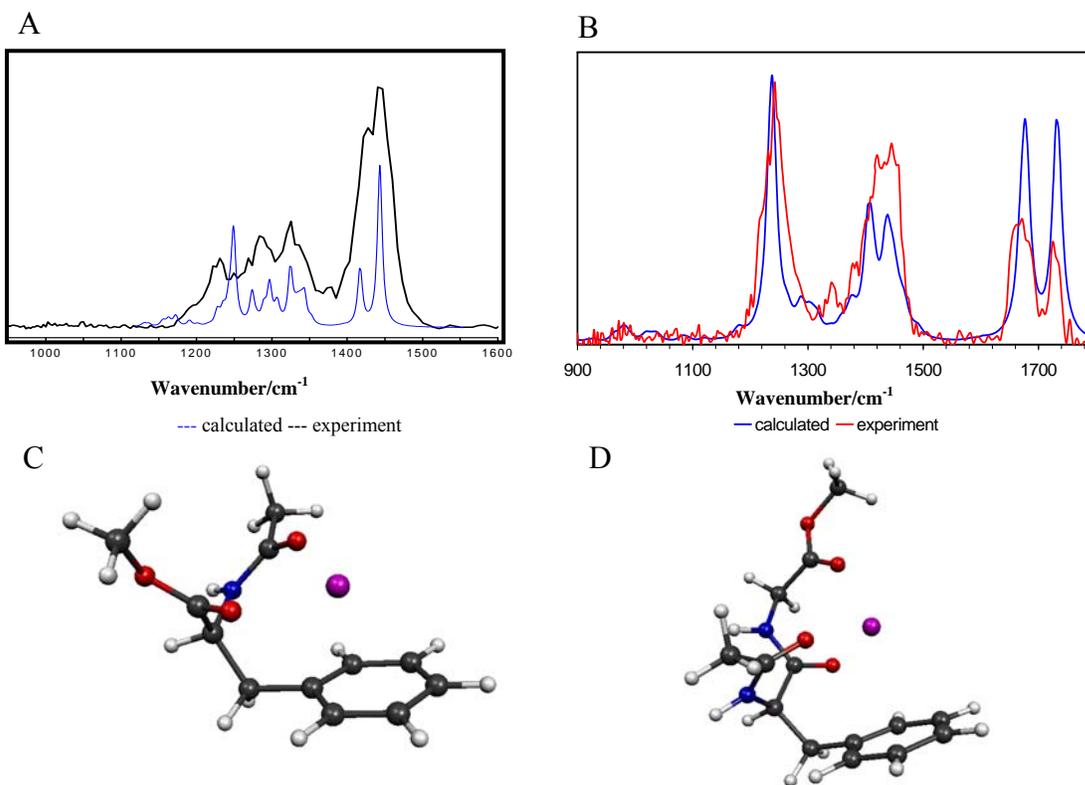


Figure 4-10. Experimental IRMPD spectra and theoretically calculated spectra for mono-deuterated sodium cation-attached AcPheOMe (A) and doubly-deuterated sodium cation-attached AcPheGlyOMe (B). The lowest energy structures found for sodium cation-attached AcPheOMe and AcPheGlyOMe are shown as (C) and (D), respectively.

CHAPTER 5 DIFFERENTIATION OF ANOMERS OF D-GLUCOSAMINE AND D-GALACTOSAMINE

Introduction

Gas-phase experiments continue to gain prominence in chemistry and biology since mass spectrometric techniques have evolved to allow routine study of proteins and other large molecules.^{132,133} Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) experiments,^{30,31,33,113,114} when combined with sophisticated ionization techniques such as electrospray ionization (ESI),³⁸⁻⁴¹ permit the study of sugar complexes without requiring sample heating for volatilization, which can often fragment or otherwise destabilize precursor ions of interest. As stated in chapter 2, sugars can be studied at low pressures (below 10^{-9} torr) in FTICR-MS, providing a collisionless environment where fragmentation processes due to laser irradiation can be studied. Collision induced dissociation and single-frequency infrared multiple photon dissociation (IRMPD) methods have been used to study carbohydrate isomers in tandem MS experiments (MS/MS).^{51,52,134} These two methods frequently produce similar fragmentation patterns for isomeric ions, making differentiation difficult. Tandem mass spectrometry at varying laser irradiation wavelengths has been shown to provide an extra degree of analysis that can help in differentiation of sugars, and recent studies using FTICR-MS in conjunction with IRMPD have shed new light on the structures of carbohydrates.^{17,66,135}

Carbohydrates, even simple monosaccharides, are flexible and can assume many configurations, sometimes favoring one particular conformation over another depending on environmental influences such as interaction with different solvents, metals, and intramolecular hydrogen bonding. Glycan conformation can strongly influence the functions of proteins to which they are bound, as glycans are important in cell-cell signaling and have been found to aid in protein folding, thus stabilizing the protein structure.

To obtain a complete picture of the sugar's role in protein-based systems, the interactions between alkali metals and sugars must be understood. Sodium and potassium play a crucial role in regulating cell functions and their blood concentration levels are indicators of well-functioning biological systems. Lithium ions are used as strong prescription drugs to regulate mood.

Practically, alkali metal-attached sugars are easily ionizable with the electrospray process and yield high abundances of the precursor ions which allow MS/MS experiments to be carried out reliably. If the alkali metal-sugar interaction is strong, fragmentation pathways other than simple alkali ion loss may become more prevalent when ions are submitted to irradiation with an infrared laser. Intramolecular fragmentation is usually observed for lithium-bound saccharides.⁵²

¹³⁶ Identification and differentiation of conformational isomers, specifically methylated α - and β -D-glucosamines (α -GlcNac and β -GlcNac) and methylated α and β D-galactosamines (α -GalNac and β -GalNac) bound to the lithium cation (Figure 5-1) has been made possible in this work through observing their infrared spectra and fragmentation pattern differences as a function of irradiation wavelength.

Plotting the fragmentation of the glycosamines as a function of laser irradiation wavelength yields IRMPD action spectra. These gas phase infrared spectra can be compared with theoretical calculations of the vibrational frequencies to obtain band assignments and structural information for each glycosamine isomer. Molecular mechanics conformational searches using the AMBER force field¹³⁷ were conducted and density functional theory (DFT)^{115,116,138} geometry optimizations and frequency calculations were used to identify spectral features of each glycosamine isomer. The results obtained are expected to provide the basis for future experimental work involving di- and trisaccharides.

Experimental

Sample Preparation

Glycosamines were obtained from Prof. Brad Bendiak, Department of Cellular and Structural Biology, University of Colorado Health Sciences Center. Stock solutions were made by dissolving 0.01 mg/mL of each glycosamine in 80:20 CH₃OH:H₂O solution. Stock samples were diluted by a factor of 10 and were mixed with an equimolar amount of LiCl before introduction in the electrospray apparatus.

Instrumentation

All ESI-FTICR-MS experiments were carried out using a laboratory built FTICR spectrometer equipped with a 4.7 T superconducting magnet (Cryomagnetics Inc., Oak Ridge, TN), which has been described previously.⁶⁶ An external Z-spray source (Micromass/Waters Corporation, Milford, MA) was used for sample injection at a flow rate of 10 μ L/ min. Electrospray ionization efficiency was increased through the use of both nebulizer and desolvation gas (N₂), with a continuous flow of 35 and 155 l/hr, respectively. Source temperature was set to 52° C and the desolvation gas temperature was 125° C. The electrospray needle-skimmer voltage difference was set to 3 kV. Precursor ions were isolated using stored waveform inverse Fourier transform (SWIFT) waveforms¹¹⁸⁻¹²⁰ to eject all other ions. Ions were detected using a broadband detection mode covering a mass range from 20 to 2000 Da. For the IRMPD activation, a free electron laser was used to scan the infrared fingerprint region of 900 – 1800 cm⁻¹ of each glucosamine anomer. Four individual transients were accumulated at each wavelength to improve the signal-to-noise ratio. Three IRMPD spectra were averaged for each anomer and the resultant spectrum used to determine the peak position of the C=O stretch band. The appearance of fragment ions with m/z 210, 132, 127, and 122 and were monitored as a function of irradiation wavelength.

Computational Details

Using the structure database provided in the Hyperchem software,¹²³ each anomer of D-GlcNac and D-GalNac was created with a clockwise hydrogen bonding network, which has been shown in literature to be the most stable solution phase structure for D-glucopyranosides.¹³⁹⁻¹⁴¹ Lithium was initially centered above the ring system of the glycan approximately 2.5 Å from the ring oxygen (O1) (Figure 5-2). Each glycosamine was given an overall charge of +1.

A conformational search of torsion angles was done to probe the rearrangement of each glycosamine isomer as it interacted with the lithium cation. All calculations were performed in an isolated environment. All H-C-O-H torsion angles were defined so that only the O-H bonds rotated about the CO bonds, thus keeping the structural definitions for each glycosamine intact. The ring torsion angles were defined and were varied in the conformational search, while ensuring that the chiral center orientation did not change. Using the usage-directed approach,⁹⁰ 1000 different structures were created with the algorithm. For each step, the dihedral angles were randomly changed to create a new structure, and a geometry optimization using the AMBER force field was performed to verify existence of a stable structure. This optimized structure was then compared to the other stable structures previously obtained from the conformational search, and duplicate structures were discarded. On average, the 100 most stable structures within 15 kcal/mol of the lowest energy structure found were then geometry optimized using Gaussian 03¹²⁴ with the B3LYP/6-31+G(d) level of theory.¹²⁶ Similar to the procedure described in Chapter 3, duplicate structures were discarded and a frequency calculation, at this same level of theory, was carried out for each structure to verify minima and to obtain theoretical infrared spectra for comparison with experimentally obtained spectra. This level of theory has been used previously to study sugars in the gas phase. A scaling factor of 0.98 was used for the calculated frequencies, and is within the reported average scaling factor using this basis set.^{142,143}

For further verification that the lowest energy conformer had been found for each glycosamine isomer, a rigid potential surface scan of the H-(C-2)-N-H dihedral angle was done (Figure 5-3) using Gaussian 03 with B3LYP/6-31+G(d). Single point calculations were performed as the amide group was rotated in 10 degree increments about the (C-2)-N bond, while keeping the rest of the molecule fixed.

Results and Discussion

Infrared Spectra

Glucosamines

The IRMPD spectra of both lithium cationized α - and β -GlcNac indicate that the spectral bands in the range from 800 to 1500 cm^{-1} are broad and there is no clear distinction between the two anomers (Figure 5-4). For the band near 1652 cm^{-1} , however, a shift in position in this band between the two anomers can be clearly seen. Using the average of 3 independently determined IRMPD spectra an average shift of 10 cm^{-1} was calculated.

Calculations indicate that the band near 1652 cm^{-1} is the C=O stretch of the carbonyl on the acetamido (-NHCO(CH₃)) group. There is a corresponding shift of 20 cm^{-1} between the theoretically calculated frequencies of the two anomers of D-GlcNac, but the experimental and theoretical shifts agree given the rms error of 34 cm^{-1} reported for B3LYP/6-31G(d) vibrational frequencies.¹⁴⁴ The lowest energy structures for lithium cationized α -GlcNac and β -GlcNac were calculated and both anomers have the chair 4C_1 conformation, where carbon 4 is above the plane of the sugar ring and carbon 1 is below the plane of the ring (Figure 5-5). Hydrogen bonding is evident. For α -GlcNac, two hydrogen bonds are formed between the hydroxyl groups from C-3 through C-6, while for β -GlcNac only one hydrogen bond forms between C-3 and C-4.

Analysis of hydrogen bonding and the most probable interaction locations for the lithium for each anomer follows.

α -GlcNac. The calculated spectrum of the lowest energy structure found for lithium cationized α -GlcNac is similar to the experimental spectrum, having the same broad features from 800 – 1500 cm^{-1} (Figure 5-6). The calculated C=O stretch frequency after scaling is 1648 cm^{-1} . With the O-methyl group at C-1 in the axial position, the amide hydrogen in α -GlcNac is located nearest to the oxygen of the O-methyl. To minimize steric hindrance, the carbonyl is oriented away from the O-methyl group. A rigid potential surface scan of the dihedral angle rotation about the (C-2)-N bond led to one stable minimum, with three other possible stationary points along the potential energy surface (Figure 5-7). The point of highest energy corresponds to the structure in which the carbonyl oxygen is in close contact with the o-methyl group. Structures with the cation at other locations of the sugar were higher in energy by more than 10 kcal/mol.

β -GlcNac. For β -GlcNac, the theoretical calculations show that the O-methyl group is in an equatorial position, that the acetamido group has reoriented so that the amide hydrogen is located nearest to the oxygen of the O-methyl group, and steric hindrance is minimized between the carbonyl and O-methyl groups. Results of the potential surface scan (Figure 5-8) indicate that a torsion angle of 6° for the H-(C-2)-N-H dihedral angle is found for the lowest energy structure in the complete range of angles from 0 to 360° , corresponding to the lowest energy structure from the conformational search. Another minimum is found in the dihedral scan (4 kcal/mol higher in energy), with large potential energy barriers between the two minima (54 and 242 kcal/mol, respectively). The highest energy structure occurs for a dihedral angle which places the N-acetyl methyl group closest to the hydrogen of the C-3 hydroxyl group. In the

lowest energy structure, the lithium cation interacts with the O1 oxygen as well as the oxygens of the O-methyl and C-6 hydroxyl groups. The Mulliken calculated charge¹⁴⁵ on the lithium cation (0.62) is close to the calculated charge of lithium on α -GlcNac (0.66), although the interaction in this case is with three oxygens instead of two oxygens. The small difference in lithium charge between the two anomers can arise from the strong interaction of the lithium cation with the carbonyl oxygen lone pair electrons of α -GlcNac, reducing the total charge by a similar magnitude to that of the three β -GlcNac oxygens interacting with the lithium cation. Other structures where the lithium cation is located at different positions relative to β -GlcNac were calculated to be higher in energy by more than 9 kcal/mol and the calculated and experimental spectra were compared (Figure 5-9).

Galactosamines

Based on analysis of the glucosamines discussed above, where the major distinguishing feature in the IRMPD spectra was the carbonyl stretching band, spectra for the anomers of galactosamine were only obtained in the range 1400 – 1700 cm^{-1} (Figure 5-10). An 11 cm^{-1} shift in band position was seen between the carbonyl stretching peaks of lithium cationized α -GalNac and β -GalNac. The calculated spectra indicate a shift of 29 cm^{-1} for the C=O stretch frequencies (Figure 5-11).

The lowest energy structures from the DFT calculations indicate a ${}^4\text{C}_1$ ring conformation (Figure 4-12). B3LYP/6-31+G(d) calculations of lithium cation-attached α -GalNac indicate that hydrogen bonding occurs between the C-3 and C-4 hydroxyl groups, and with the carbonyl oxygen and the hydroxyl group at C-3. The other potential hydrogen bonds are interrupted due to the lithium cation located between the C-4 and C-5 functional groups. There is also a hydrogen bond between the amide and the O-methyl oxygen. β -GalNac has one hydrogen bond

between the C-3 and C-4 O-H groups. The position of the acetamido group relative to the O-methyl group is similar to that found for the glucosamines. Interaction of the amide hydrogen with the O-methyl oxygen and the lessening of steric hindrance cause the carbonyl to change position relative to the O-methyl group; these are the two major features that influence the shift of the C=O stretch frequencies.

Potential surface scans of α -GalNac and β -GalNac (torsion about the (C-2)-N bond), similar to those carried out for the glucosamines, indicated that the conformers found in the conformational search and used throughout this analysis were the lowest points on the 1-dimensional surface (Figures 5-13 and 5-14, respectively). The highest energy points on the surface correspond to structures in which the carbonyl oxygen is in close proximity to the O-methyl oxygen at C-1 for both anomers of D-GalNac. These are analogous to the highest energy points on the potential surfaces of the D-GlcNac anomers. The lithium cation in the complex with α -GalNac is bound to the O-1, C-6 and C-4 hydroxyl oxygens, and has a charge of 0.57. The interactions of the lithium cation in β -GalNac are the same as in α -GalNac (O-1, C-6 and C-4 hydroxyl oxygens) corresponding to a similar charge of 0.58 on the lithium cation. Other structures with a different lithium position in the α -GalNac complex have a calculated energy of 10 kcal/mol or higher, while for β -GalNac the more energetic locations are 5 kcal/mol or higher in energy. The charge on the lithium cation is less for the galactosamines than for the glucosamines, as is evident from the interactions just reported. Table of the relative energies using DFT for the structures from the conformational search (Table 5-1) are given below.

Fragmentation Patterns

Four fragments were seen for each of the sugars in the tandem mass spectra produced by IRMPD fragmentation (Figure 5-15). The fragmentation channels involved the loss of CH₄O

(210 m/z), fragments resulting from cross ring cleavage at C-1 - O1 and C-2 - C-3(127 m/z) and (132 m/z) and fragments resulting from cross ring cleavage at C-1 - O1 and C-2 - C-3 (122 m/z) (Figure 5-16). The proposed fragment ions include the lithium cation, except m/z 132. From Figure 5-15, the abundance of the 122 m/z fragment ion for α -GlcNac is higher than for the other three isomers, and the DFT calculated structures suggest that the increase in abundance may be due to the lithium cation being located between the acetamido and C-3 functional groups. As cross ring cleavage occurs to produce the 122 m/z fragment (which includes the acetamido group), attachment of lithium to the leaving group would be facile. For β -GlcNac and α/β -GalNac, fragmentation ions and their relative abundances are similar, showing no discernable differences in the MS/MS spectra. DFT calculations support these results, since calculated structures show the lithium cation in a similar position for each isomer, thus predicting that the lithium cation-bound fragments would have similar abundance ratios.

Conclusion

Glycosamine isomer differentiation in the gas phase is possible with IRMPD spectra if each isomer has distinct infrared spectral features or fragmentation patterns in the tandem mass spectrum. In this study, for the lithium cationized anomers of glucosamine and galactosamine, C=O stretch band shifts were observed (10 and 11 cm^{-1} , respectively) and can be used to differentiate the anomers. The shifts between the alpha and beta anomers were explained using theoretically calculated spectra as being due to the difference in location of the acetamido carbonyl, which is influenced by the repulsion of the O-methyl oxygen on carbon atom 1 and the carbonyl. The favorable interaction of the amide hydrogen with the O-methyl oxygen also influences the location of the carbonyl, since both are part of the acetamido group. Calculated structures indicate that the lithium cation is located between C-3 and the acetamido group for α -

GlcNac. This is corroborated by the tandem mass spectrum fragmentation pattern that indicates a larger abundance of the acetamido-lithium fragment ion for α -GlcNac than for the other isomers. As discussed in the results section, the other 3 isomers have a similar location of the lithium cation. In the calculated structures of β -GlcNac and α/β -GalNac, the lithium cation is found to interact with 3 oxygens, near the ring oxygen. Fragmentation patterns in the MS/MS spectrum of the three isomers are similar, possibly due to the similar position of the lithium cation for all three. Other defining spectral features, which might be seen in different ir wavelength ranges, could be used to further differentiate the isomers. A study of glycosides and their isomer differentiation using bands in the mid-ir and O-H stretch regions will be presented in chapter 6.

Table 5-1. Glycosamine isomer structures obtained from the corresponding conformational searches. Relative energies were obtained at the B3LYP/6-31+G(d) level of theory. Conformer 1 of each isomer appears in Figures 5-5 and 5-12.

Conformer #	α -GlcNac	β -GlcNac	α -GalNac	β -GalNac
1	0.00	0.00	0.00	0.00
2	4.67	0.01	2.70	3.42
3	5.60	1.79	15.66	5.00
4	15.56	1.80	17.21	5.01
5	17.45	2.60	17.30	6.65
6	18.66	2.61	17.41	11.03
7	24.55	2.94	17.42	12.41
8	30.06	5.19	19.09	12.56
9	30.37	5.27	19.10	13.21
10	30.62	5.31	25.56	13.37
11	34.25	5.53	25.72	13.92
12	35.65	6.78	27.71	15.41
13	36.04	7.72	32.75	16.38
14	36.34	8.86	34.95	18.30
15	38.24	15.16	-	18.33
16	38.48	-	-	21.34
17	39.37	-	-	22.58
18	40.82	-	-	-
19	47.42	-	-	-
20	50.99	-	-	-

- indicates that the 15 kcal/mol cutoff was met with the specified number of structures

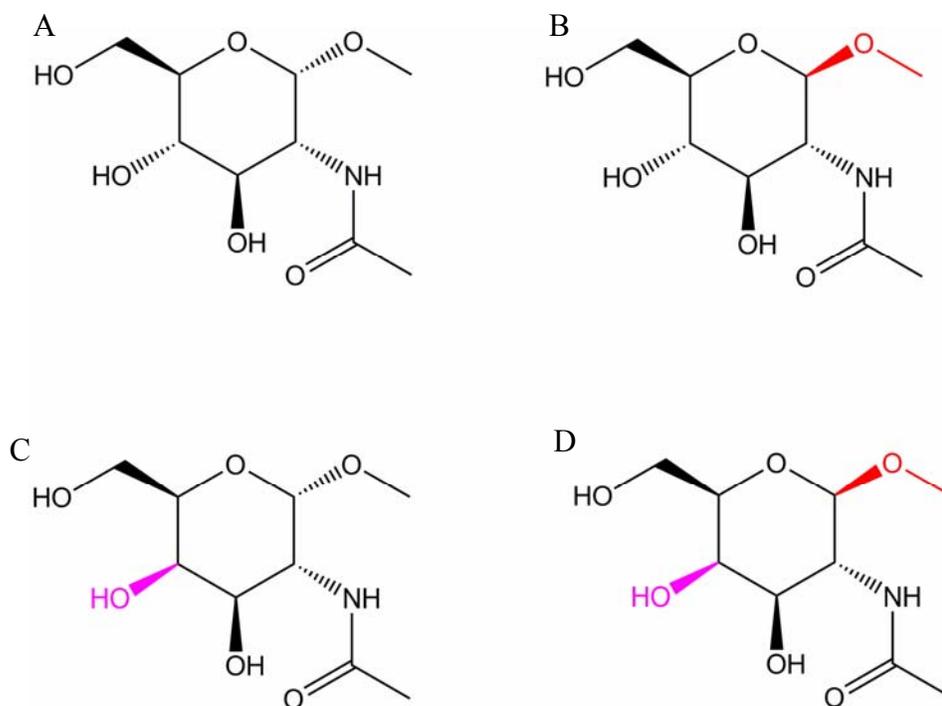


Figure 5-1. Isomers A) α -GlcNac, B) β -GlcNac, C) α -GalNac and D) β -GalNac. The differences between α - and β - structures at carbon 1 and those of glucose and galactose on carbon 4 are shown in red and purple, respectively.

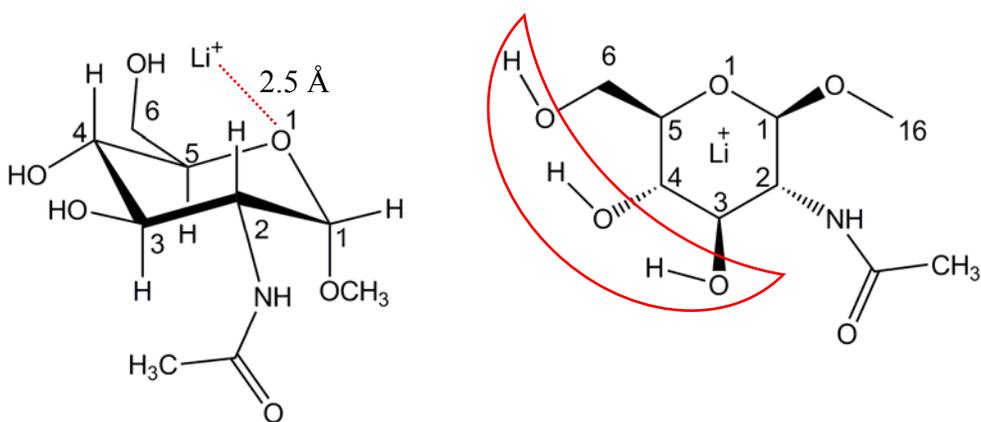


Figure 5-2. Side and top views of lithium cationized β -GlcNac. The lithium cation was initially placed above the ring, approximately 2.5 Å from O1. The clockwise hydrogen bonding network is visible in the top view and is encircled in red. The carbons and ring oxygen are numbered.

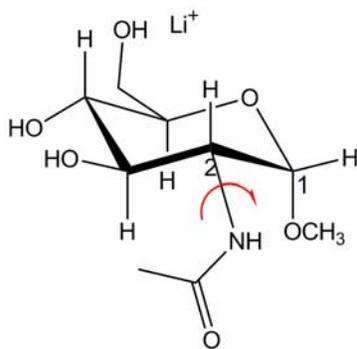


Figure 5-3. Rotation about C-2 and N, shown with a red arrow.

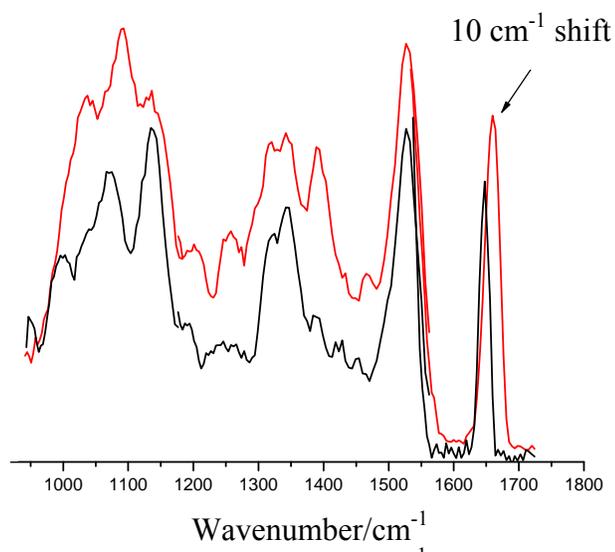


Figure 5-4. IRMPD spectra of lithium cation-bound α -GlcNAc (black) and β -GlcNAc (red), showing the 10 cm^{-1} shift of the carbonyl stretch frequency.

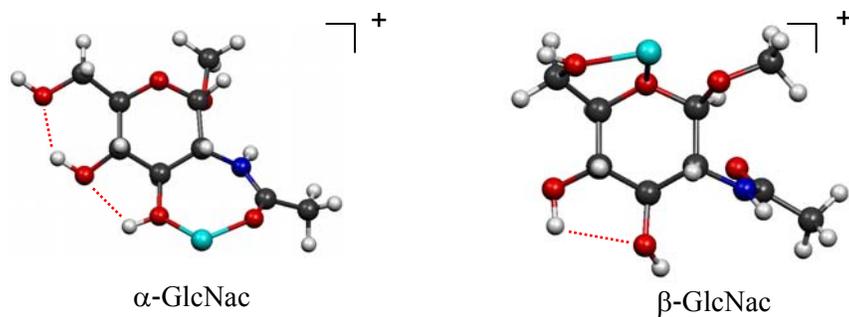


Figure 5-5. Calculated lowest energy structures of lithium cationized α -GlcNac and β -GlcNac, using B3LYP/6-31+G(d). Lithium is shown in blue and hydrogen bonding is indicated with dashed dots.

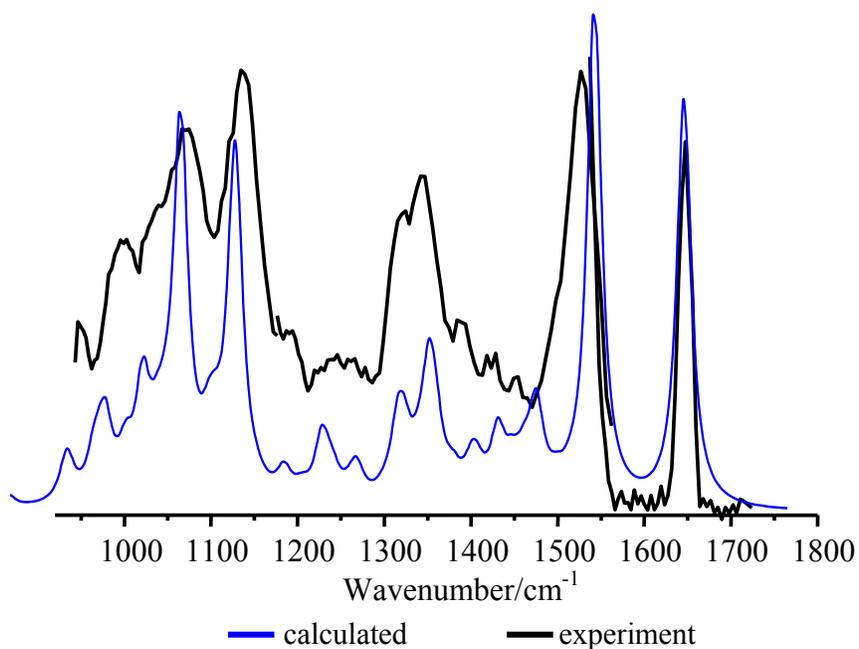


Figure 5-6. Comparison of experimental IRMPD and calculated infrared spectra of lithium cationized α -GlcNac. The calculated spectrum has been scaled by 0.98, and a 20 cm^{-1} Gaussian band profile has been used.

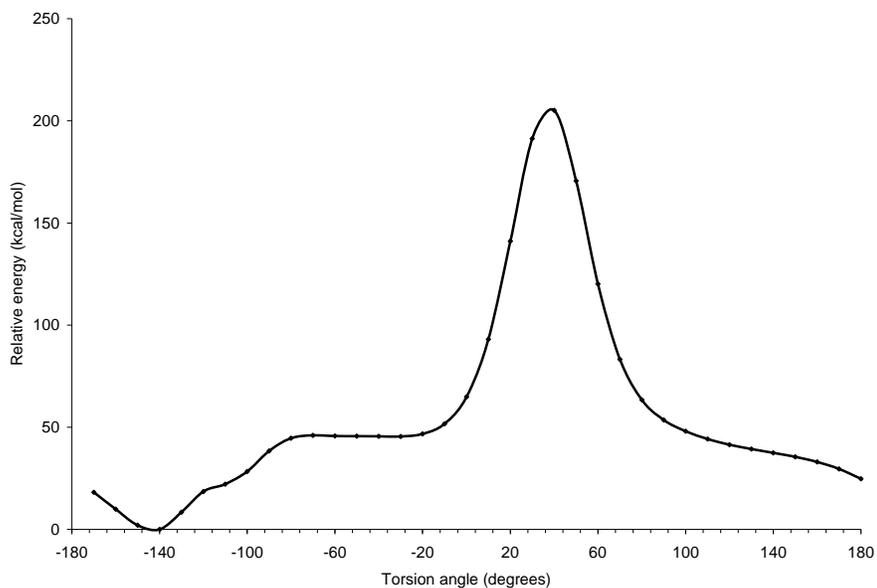


Figure 5-7. Potential surface scan for rotation about (C-2)-N bond for α -GlcNac.

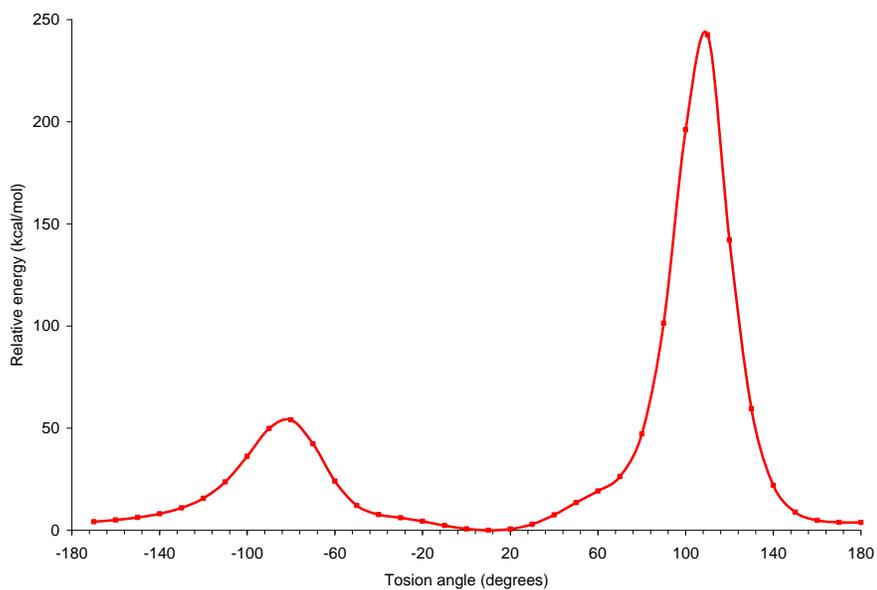


Figure 5-8. Potential surface scan for rotation about the (C-2)-N bond for β -GlcNac.

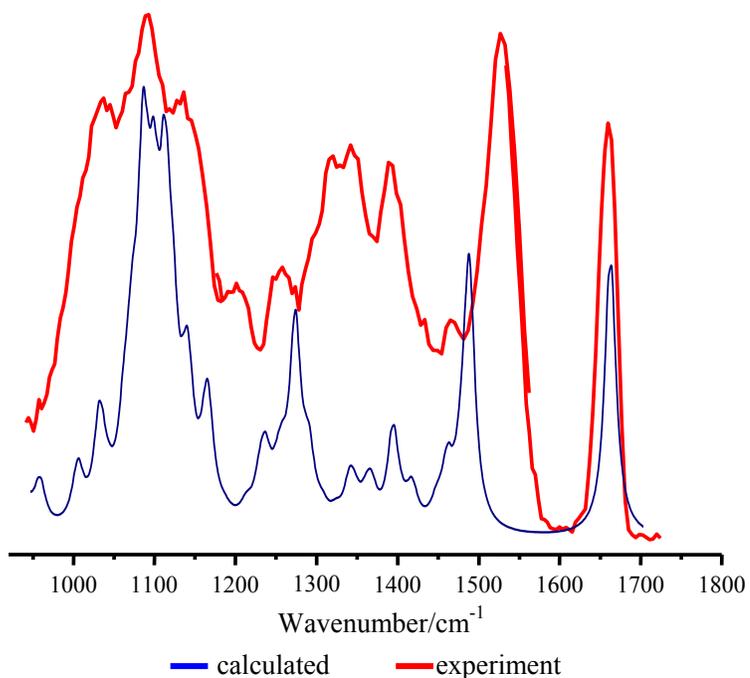


Figure 5-9. Comparison of experimental IRMPD spectrum (red) to the calculated infrared spectrum (blue) of lithium cationized β -GlcNac.

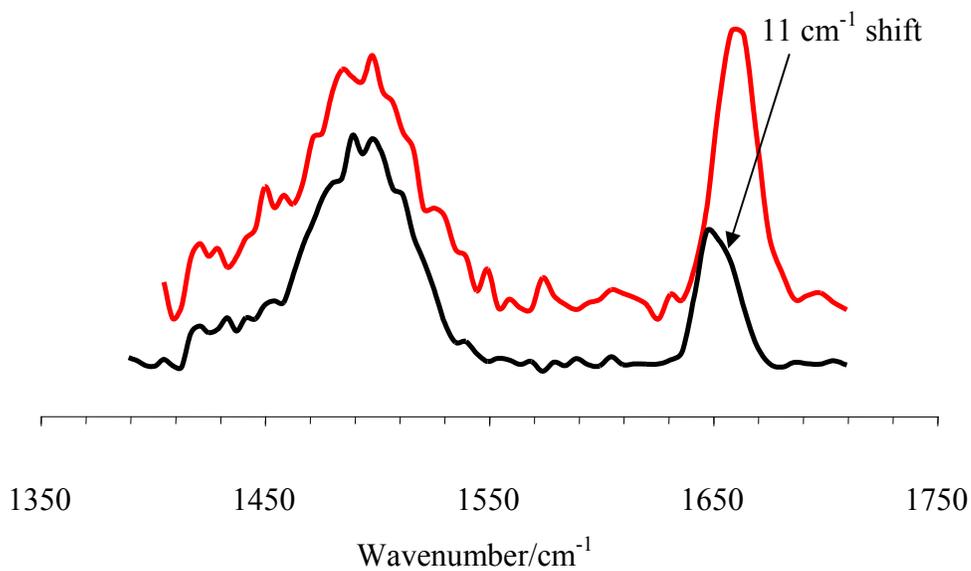


Figure 5-10. IRMPD spectra of lithium cationized α -GalNac and β -GalNac, showing the 11 cm^{-1} shift for the carbonyl stretch bands.

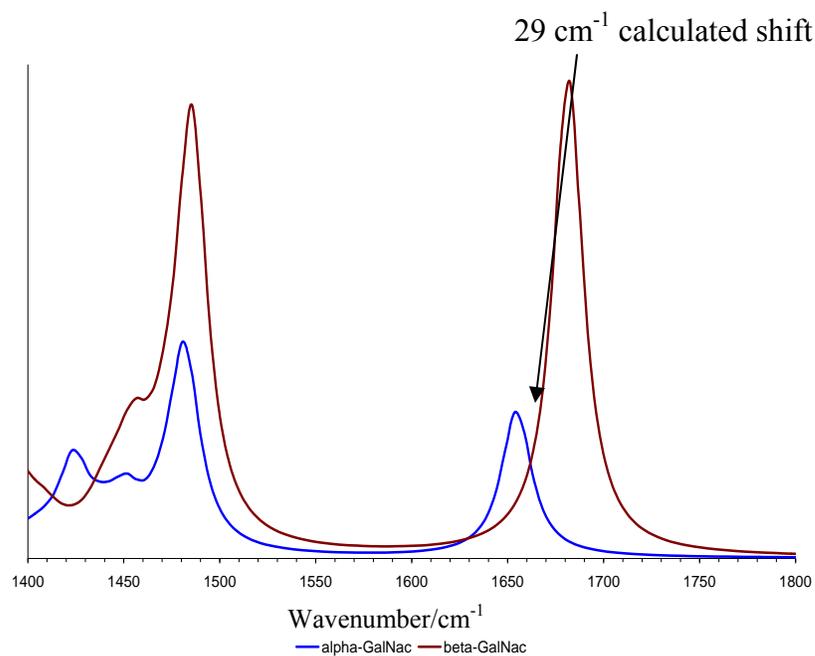


Figure 5-11. Calculated infrared spectra of lithium cationized α -GalNac and β -GalNac, indicating a 29 cm^{-1} shift in the C=O stretches.

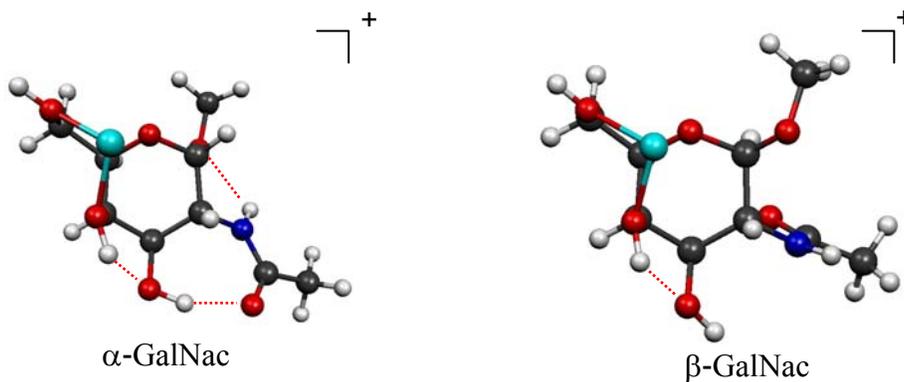


Figure 5-12. Lowest energy structures of lithium cation-attached galactosamine anomers from B3LYP/6-31+G(d), showing hydrogen bonding. Lithium is shown in blue.

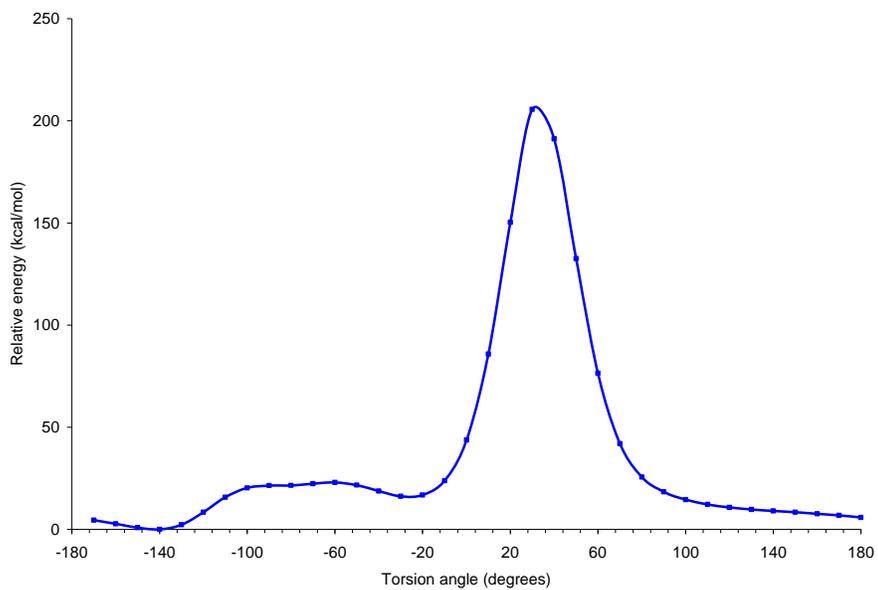


Figure 5-13. Potential surface scan for α -GalNac.

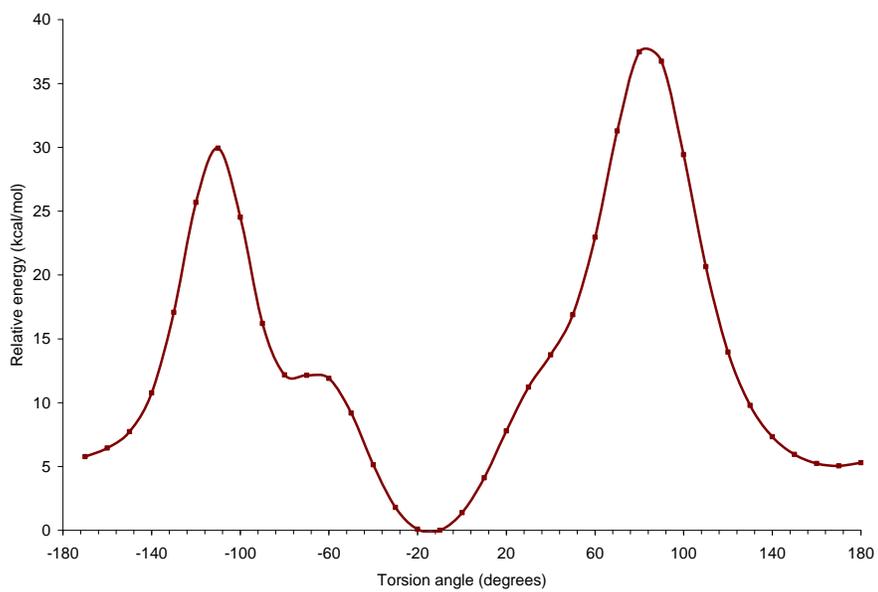


Figure 5-14. Potential surface scan of β -GalNac.

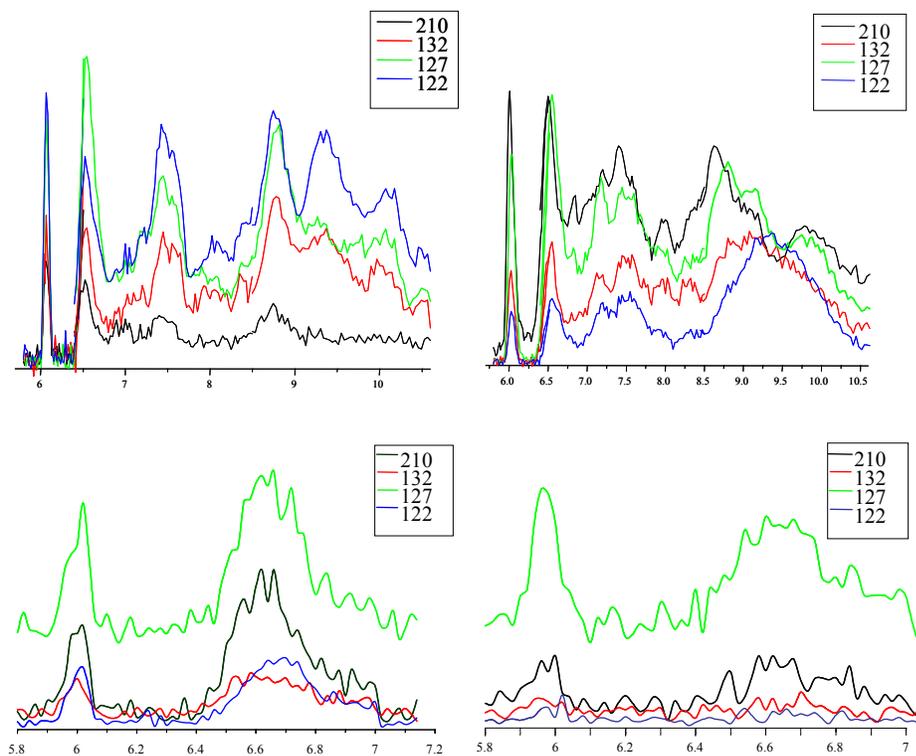


Figure 5-15. Fragment ion abundances as a function of laser wavelength for all four isomers. The abscissa has units of wavelength (μm).

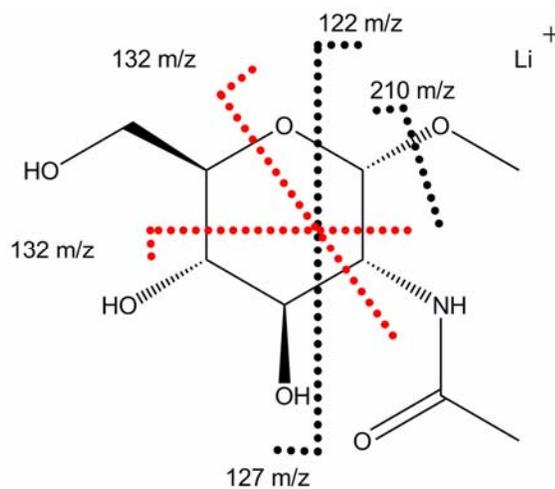


Figure 5-16. Possible fragmentation producing the observed MS/MS spectra of glycosamine isomers. Black dashed dots indicate lithium cation-attached fragment ions and red dashed dots indicate fragment ions without the lithium cation.

CHAPTER 6 DIFFERENTIATION OF O-METHYLATED GLYCOSIDE ISOMERS

Introduction

Mass spectrometry has been used successfully to obtain sequence and linkage information about oligosaccharides.¹³⁵ However, as many oligosaccharides are structural isomers with identical mass, differentiation of their isomeric structures can be challenging, if not impossible, with this approach. Chapters 1 and 2 mention that one of the recently developed methods for obtaining structural information about sugars combines infrared spectroscopy with theoretical calculations.^{17,146-150} With solid and solution phase samples, infrared bands of sugars and many other large molecules are broad, and therefore it is difficult to observe multiple identifying features in their spectra.¹⁵¹ A gas phase experiment may have less spectral broadening since there is no solvent present and therefore provide more identifying spectral features that allow for differentiation of isomers. In the gas phase, a typical infrared absorption experiment is difficult to conduct, since concentration of the species of interest is too low to detect absorbance of radiation of the incident laser beam by the gaseous sample. Nevertheless, as has been discussed throughout this work, a gas phase spectrum is obtainable using action spectroscopy methods.

Infrared multiple photon dissociation (IRMPD)¹⁵² spectra were obtained by Valle *et al.* for the rubidium cation-attached glycoside isomers O-methyl- α -D-glucoside (α Glc), O-methyl- β -D-glucoside (β Glc), O-methyl- α -D-galactoside (α Gal) and O-methyl- β -D-galactoside (β Gal) at the FOM-Institute for Plasma Physics Rijnhuizen, The Netherlands using the Free Electron Laser for Infrared eXperiments (FELIX).^{17,66,122} One major band for each isomer is observed in the 600 – 1700 cm^{-1} range, with few distinguishing features (Figure 6-1).

However, the O-methylated glycosides have four hydroxyl groups (Figure 6-2), which are predicted to give rise to intense stretching bands in the 3500 cm^{-1} wavenumber region of the infrared spectrum. Therefore, IRMPD spectra of the O-H stretching region were obtained to positively differentiate between all four isomers. Extensive calculations were carried out with density functional theory using the B3LYP hybrid functional and 6-31+G(d) basis set.^{126,128} This chapter presents IRMPD spectra of the same four O-methyl D-glycoside isomers studied by Valle *et al.*, but the spectra were taken using an FTICR-MS coupled with a continuous wave optical parametric oscillator (cw-OPO) laser.^{57,61,153} The OPO optics setup and the experiment sequence are introduced in the experimental techniques section. Results and spectral characteristics of each isomer in the mid-ir and the O-H stretching region are discussed, as well as hydrogen bonding and the rubidium cation interactions that influence the O-H stretch band positions in the IRMPD spectra. Setup of the OPO optics and acquisition of IRMPD spectra using the OPO laser was done in conjunction with Wright L. Pearson. All figures of experimental spectra were plotted by Mr. Pearson. In this chapter, the appropriate abbreviation, for example βGlc for Rb^+ -O-methyl- β -D-glucopyranoside, will be used to indicate the rubidium cation-attached monosaccharide as opposed to the neutral molecule.

Experimental Techniques

Chemicals

O-methyl D-glycoside samples were provided by Prof. Brad Bendiak, Department of Cellular and Structural Biology, University of Colorado Health Sciences Center. Solvents and inorganic salts were obtained from Sigma-Aldrich, and were used without further purification. Solid 10 mg samples of D-glycosides were dissolved in a 10 mL solvent mixture of 80:20 MeOH:H₂O. Stock solutions were diluted to $1 \times 10^{-4}\text{M}$ with the same solvent mixture, and an equimolar amount of RbCl was added as the ionization agent.

Instrumentation

An FTICR 4.7T Apex II mass spectrometer (Bruker, Billerica, MA) coupled with an ESI source (Analytica of Branford, Inc., Branford, CT), which is located in the Mass Spectrometry Services laboratory at the University of Florida was used for experiments with the OPO laser. The ESI source had a flow rate of 2 $\mu\text{L}/\text{min}$, and ionization efficiency was aided with a nebulizer and a desolvation gas, both N_2 gas, with continuous flows of 35 and 155 l/hr, respectively. The electrospray needle-capillary voltage difference was set to 3.6 kV. Precursor ions were isolated using swept frequency ejection pulses to remove all other ions. Ions were detected using the broadband detection mode covering a mass range from 70 to 500 Da. Data was collected with the Bruker X-MASS™ data acquisition system.

For IRMPD spectra obtained with FELIX, a 4.7T superconducting magnet (Cryomagnetics Inc., Oak Ridge, TN) was used with a laboratory-constructed pumping system, ion trap and electronics console and with a Z-spray electrospray ionization source. The FELIX setup was mentioned in chapters 3 and 4, and is described in the article by Valle *et al.*⁶⁶

Laser Setup

The tunable, continuous-wave OS 4000 optical parametric oscillator (OPO) laser (LINOS Photonics, München, Germany) uses a Nd/YAG pump laser (2W, 1064 nm) to produce two output beams from the nonlinear OPO crystal, the idler and signal beams. The two beams are dependent on the pump radiation and are correlated by Equation 6-1.

$$\frac{1}{\lambda_p} = \frac{1}{\lambda_s} + \frac{1}{\lambda_i} \quad (6-1)$$

The λ_p , λ_s and λ_i are the wavelengths of the pump laser and the signal and idler beams, respectively. Tuning the OPO is accomplished by changing the poling period of the OPO crystal. A total of 18 poling periods, with different wavelength ranges, allow for lasing in the

wavelength ranges from 1.38-2.0 and 2.28-4.67 microns. Within a poling period, finer tuning is accomplished by changing the temperature of the OPO crystal from 50 – 150° C. An etalon is used to minimize and discriminate between the competing lasing wavelengths, allowing for fine tuning of a specified wavelength. A piezo-electric mirror makes up one end of the resonance cavity and uses lock-in electronics to keep the laser wavelength in resonance. The current filtering lenses allow the use of the two idler beams created at both ends of the resonance cavity. Figure 6-3 shows idler beam 1 exiting at the front of the OPO enclosure box and idler beam 2 exiting at the back of the OPO enclosure box. Both idler beams are oriented by a set of mirrors to enter the ICR cell collinearly. Two Uniblitz CS25 shutters (Vincent Associates, Rochester, NY) block the idler beams from entering the cell when closed and divert the beams to a detector to record the power output. The signal beam exits the OPO enclosure box adjacent to idler beam 1 and is directed into a wavemeter (WA-1500, EXFO Electro-Optical Engineering Inc., Plano, TX). A 10.6 micron continuous wave CO₂ laser (Synrad, Inc., Mukilteo, WA) is aligned to enter the ICR cell and is collinear with the OPO idler beams. The CO₂ laser can be used independently or as an off-resonance laser in a two-laser experiment.^{154,155} The entire optics setup, with the exception of a short path of the signal beam before it enters the wavemeter, is enclosed in a nitrogen gas purge box designed by Mr. Pearson to eliminate absorbance of laser radiation from water vapor in the air.

Experiment

OH Stretches.

After electrospray ionization, trapping and isolation in the ICR cell, rubidium cation-attached glycoside ions were irradiated for 10 s by the OPO idler beams, followed by excitation and detection of fragment ions. A typical experimental sequence at each wavelength, the appearance of Rb cation signal as well as the loss of precursor ion signal were monitored *via* the

mass spectrum (Figure 6-4). The peaks due to each isotope of rubidium, ^{85}Rb and ^{87}Rb , and their corresponding complex glycoside ions, were monitored and used to form the composite IRMPD spectrum.

CH Stretches. Weak absorption by the C-H stretching modes of the rubidium cation-attached glycosides produced little or no loss of the rubidium cations even after 10 s of irradiation with the OPO lasers. Using a fixed wavelength (10.6 μm) CO_2 laser, an on-resonance, off-resonance two laser experiment was conducted, where the CO_2 laser was the off-resonance laser used to provide more photon fluence for dissociation. The mid-ir spectra in Figure 6-1 show little or no absorption at or near the 10.6 μm ($\sim 940\text{ cm}^{-1}$) wavelength, and fragmentation is not expected to occur at this wavelength. The CO_2 irradiation time and power were adjusted such that none of the glycoside-rubidium cation complexes were observed to fragment due to the off-resonance radiation alone. The experimental sequence included irradiation for 10 s with OPO idler laser beams, and irradiation with the CO_2 laser during the last 3 s of the irradiation period (Figure 6-5). In regions where it was suspected that this two laser sequence did not provide enough photons for absorbance by the C-H stretch bands to induce dissociation of the rubidium cation, total irradiation time was increased to 15 s, with the last 7 s including the CO_2 laser. At each wavelength, ten individual transients were accumulated to improve the signal-to-noise ratio. For D-glucosides, three IRMPD spectra (conducted on different days) were used to calculate the average spectrum for each isomer. Preliminary IRMPD spectra of D-galactosides will also be discussed.

Computational Details

Using the Hyperchem¹²³ drawing database, O-methylated-D-glucosides were drawn with a clockwise hydrogen bonding network, which has been shown in literature to stabilize

monosaccharides.^{139,140,156} For each isomer, the rubidium cation was placed centered above the glycoside ring system, approximately 3 Å from O1 (Figure 6-6). The dihedral angles within the ring atoms were defined and were varied in the conformational search by using the built-in ring flex and torsion algorithm of Hyperchem.⁹⁰ The H-C-O-H torsion angles were defined so that the O-H functional group would rotate about the CO bonds, keeping the torsion angle changes separate from the ring conformation changes. All torsion angles were randomly varied for 1000 searches. For each of the 1000 structures created, a geometry optimization using the AMBER^{77,137} force field was performed. Comparison of conformers was done to find duplicate structures. The parameters used to test for duplicate structures were the atomic positions, dihedral angles and total energy, whose threshold values were set to 0.25 Å, 10°, and 0.05 kcal/mol, respectively. All conformers within 15 kcal/mol of the lowest energy conformer were further geometry optimized using B3LYP/6-31+G(d), plus LANL2DZ basis^{157,158} for Rb, in Gaussian03.¹²⁴ For the glucoside anomers, further ring conformations were explored whenever the conformational search did not provide adequate structures to account for IRMPD spectra bands of each anomer. Chair (⁴C₁) conformations with differing rubidium cation position were only explored for D-galactosides. Vibrational analysis was then done for all the conformations at the B3LYP/6-31+G(d) level of theory. Frequencies were scaled by 0.97 for O-H stretches and mid-ir region, and 0.96 for C-H stretches. These scaling factor values are well within the error for factors used for similar sugars.^{85,140,141,143,146,156,159-162}

Results and Discussion

FELIX

The lowest energy conformations for each isomer from the conformational search (Figure 6-7) and the calculated infrared spectra of the conformers were used to compare to FELIX IRMPD spectra (Figure 6-8). Using a 20 cm⁻¹ Gaussian band profile to mimic the width of the

experimental spectral bands, the calculated spectra of the lowest energy structures seem to agree with the experimental spectra (Figure 6-8). Using the next set of higher energy structures calculated for each isomer and comparing them to the corresponding IRMPD spectrum, it can be seen that these also seem to agree with the experimental spectra (Figure 6-9). Although the experimental band that appears for each isomer in the 900 – 1400 cm^{-1} range has a few small features that can potentially be used for differentiation between isomers, all the calculated spectra include those same features and therefore a distinction between isomers is unambiguous. The experimental spectra band is too broad to be used for identification and differentiation of each isomer.

OPO

Glucosides

β -D-glucoside. The IRMPD spectrum of rubidium cation-attached O-methyl- β Glc from 2750 – 3750 cm^{-1} (Figure 6-10) with the general region for C-H and O-H stretches are indicated. The most common region for free O-H stretches,¹⁶³ namely for those hydroxyl groups that are not interacting with another atom, molecule, or ion, is expected to be in the range of 3500 – 3700 cm^{-1} . Hydroxyl groups that are involved in hydrogen bonding are expected to have O-H stretches in the lower energy range of the O-H stretch region (3200 – 3600 cm^{-1}). For β Glc in Figure 6-10, the O-H stretch region shows four major bands, 3450, 3560, 3637 and 3677 cm^{-1} , labeled I, II, III and IV, respectively. The methylated sugar has four O-H groups, accounting for the number of bands observed in the spectrum. The bandwidths are 32, 20, 9, and 21 cm^{-1} full width at half maximum, respectively. The side band near band II has a bandwidth of 17 cm^{-1} . The C-H stretch region has two bands from 2800 – 3050 cm^{-1} . Figure 6-11 shows a comparison of the C-H stretch region spectrum when using only the OPO laser and the spectrum from the

two laser experiment. All that is observed in the latter is enhancement of the bands with no new bands or shifting of the bands seen in the one laser experiment.

The simulated spectra were compared to the OPO-IRMPD spectrum. For the lowest energy conformer (A) (Figure 6-7) there is minimal agreement with the experimental spectrum (Figure 6-12). Comparing the simulated spectrum of the next lowest energy conformer (B) (0.8 kcal/mol higher in energy) with that of the experiment, the agreement is similar to that of conformer (A) (Figure 6-13). Two other conformations that are within 7 kcal/mol of the lowest energy structure are compared to the experimental data (Figures 6-14 and 6-15). An overlap of spectra for all four calculated conformers with the IRMPD spectrum of β Glc shows that these four conformers account for all the spectral features in the IRMPD spectrum, and all are within the margin of error for the DFT calculations at this level of theory (Figure 6-16).

The the first three conformers of β Glc (Figure 6-17) have a 4C_1 ring conformation, and a fourth conformer D has a 1C_4 ring conformation. The carbohydrate nomenclature was discussed in chapter 1. To account for bands I and II in the IRMPD spectrum, the large red shift of the O-H stretches must be explained. Conformer D has hydrogen bonds between C-3 and C-6 hydroxyl groups and C-6 and the oxygen of the O-methyl group of carbon atom 1. This results in the shifting of the O-H stretching frequencies for the respective hydroxyl groups to band I (3450 cm^{-1}) and II (3560 cm^{-1}) for conformer D. The location of the rubidium cation is similar for conformers A and B, but changes for conformers C and D. For conformers A and B, the rubidium cation is above the plane of the ring. For conformer C, the rubidium cation is between C-4 and C-5 functional groups, and for conformer D, the rubidium cation is below the plane of ring. In all cases, the rubidium cation is interacting with the maximum possible number of oxygen atoms. For conformer C, the position of the rubidium cation and hydrogen bonding

network cause the O-H stretch of the C-4 hydroxyl to red shift toward band II. For all four conformations of β Glc, the free O-H stretches were of highest energy, corresponding to the band at 3677 cm^{-1} . Band IV, although indicating that the O-H stretch has not been perturbed, cannot be generally assigned to a particular hydroxyl group since each conformation has unperturbed hydroxyl groups at different carbon atoms. It should be noted that conformers A – D have relative energies up to 6.49 kcal/mol, which is within the error for DFT calculations at this level of theory,^{144,164-166} and potentially all conformations can be present in the gas phase experiment.

The region from $2800 - 3050\text{ cm}^{-1}$ has multiple C-H stretching modes, and similar to the mid-ir spectra, the peaks are too broad to contribute in making a distinction between the isomers, and will not be discussed. An extensive conformational search for all rubidium cation-attached β Glc, where all torsion angles and rings conformations, continues to be explored by the author and Wright L. Pearson. The spectrum for β Glc will be used as a basis of comparison for spectra of the other three isomers.

α -D-glucoside. The O-H stretch region of rubidium cation-attached α Glc shows three bands, with the band at 3637 cm^{-1} that appears for β Glc is clearly missing. The three bands have similar positions to those of β Glc (Figure 6-18). The bandwidths of the α Glc O-H stretch bands are 45, 54 and 21 cm^{-1} , respectively. The C-H stretch region shows two bands in the range of $2900 - 3050\text{ cm}^{-1}$ and the intensities of the two bands have changed with respect to those of β Glc. Two O-H stretch bands of α Glc overlap bands I and II of β Glc and are indicative that the hydroxyl groups contributing to these bands are involved in strong hydrogen bonding and are perturbed by the rubidium cation. From the results of β Glc, the third band of α Glc at 3677 cm^{-1} should correspond to an O-H stretch of a hydroxyl group not involved in hydrogen bonding. Similar to β Glc, the calculated spectrum of the lowest energy structure found for α Glc did not

account for all bands present in the IRMPD spectrum (Figure 6-19). The theoretical spectrum for two other conformations of α Glc, one having a 1C_4 ring conformation and the other with a 0_3S ring conformation, have better correlation to the experimental spectrum of α Glc (Figure 6-20), and the conformations are shown in Figure 6-21. The calculated bands of these conformers are significantly red shifted, but do not completely overlap all O-H stretch bands, and further conformational calculations are required for better correlation between calculated and experimental spectra.

Galactosides

Spectra for the rubidium cation-attached O-methyl-D-galactosides was also taken, although only a single pass in the range of $3500 - 3800 \text{ cm}^{-1}$ has been completed, and therefore will not be compared to calculated structures. For α Gal, the range from 3400 to 3580 cm^{-1} has been scanned twice and the average of the two spectra was calculated and the IRMPD spectra of α Gal and β Gal was compared (Figure 6-22). A quick summary of the IRMPD spectral features for the galactosides is discussed, relative to the findings of the glucoside isomers.

β -D-Galactoside. Two bands appear for β Gal, centered at 3580 and 3650 cm^{-1} , with bandwidths of 17 and 30 cm^{-1} , respectively. The two bands fall within the region for bands II and III of β Glc, and it is expected that the gas phase conformations for β Gal have hydroxyl groups involved in hydrogen bonding and rubidium interactions.

α -D-galactoside. The O-H region has 3 distinguishable bands that are broad, appearing at 3550 , 3600 and 3660 cm^{-1} , with respective bandwidths of 38 , 27 and 30 cm^{-1} . In Figures 22 and 23, α Gal is shown to have distinct spectrum relative to the other three isomers. The band near 3600 cm^{-1} differentiates α Gal and β Gal, since β Gal has a nearby band that is shifted to 3580 cm^{-1} . Similar to β Gal, the O-H stretches also fall within bands II and III of β Glc, with a slight

overlap of band IV. The gas phase conformations for α Gal are expected to have hydroxyl groups whose stretches are perturbed by the rubidium cation and hydrogen bonding, but the perturbations are not expected to be as strong as those for the glucoside isomers.

A spectrum showing IRMPD spectra for all four isomers indicates that each isomer has peaks that overlap band II and band III of β Glc, with a slight overlap at band IV (Figure 6-23). Comparison of all four isomer spectra, there is a clear indication that differentiation is possible using the infrared spectra of the O-H stretch region. Differentiation between the glucoside anomers, α Glc and β Glc, can be easily accomplished if the region of band III is scanned, since it is in this wavelength range that no O-H stretch is expected for α Glc. Differentiation between the galactoside anomers, α Gal and β Gal, is possible due to the band shift of β Gal near 3580 cm^{-1} to 3600 cm^{-1} of α Gal. Finally, differentiation of the glucosides and galactosides is possible using their respective IRMPD spectra, since both galactosides do not have band peaks in the region of band I of β Glc,. A peak appearing in the region of band I of β Glc can be used as an analytical marker to indicate the presence of rubidium cation-attached glucosides, and not galactosides, in an unidentified sample.

Conclusions

To differentiate the rubidium cation-attached isomers of O-methyl- α -D-glucoside, O-methyl- β -D-glucoside, O-methyl- α -D-galactoside and O-methyl- β -D-galactoside, Fourier transform ion cyclotron resonance mass spectrometry in conjunction with infrared multiple photon dissociation was used to obtain infrared spectra of the rubidium-bound isomers in the gas phase. Previous work of Valle *et al.* at the FOM-Institute for Plasma Physics Rijnhuizen, The Netherlands, showed few discernable differences in the spectra of all four isomers in the spectral range of $600 - 1700\text{ cm}^{-1}$, so that differentiation of conformational isomers using DFT

calculations was dubious. Work was done to couple an optical parametric oscillator laser system with an FTICR-MS instrument to obtain IRMPD spectra at 2600 – 3900 cm^{-1} . Theoretical calculations and experimental spectra indicate that the spatial orientation which distinguishes each isomer (α or β , glucose or galactose) influences the strength of the hydrogen bonding network. The rubidium cation location also perturbs the O-H stretches and the hydrogen bonding network, making differentiation of all four glycoside isomers possible.

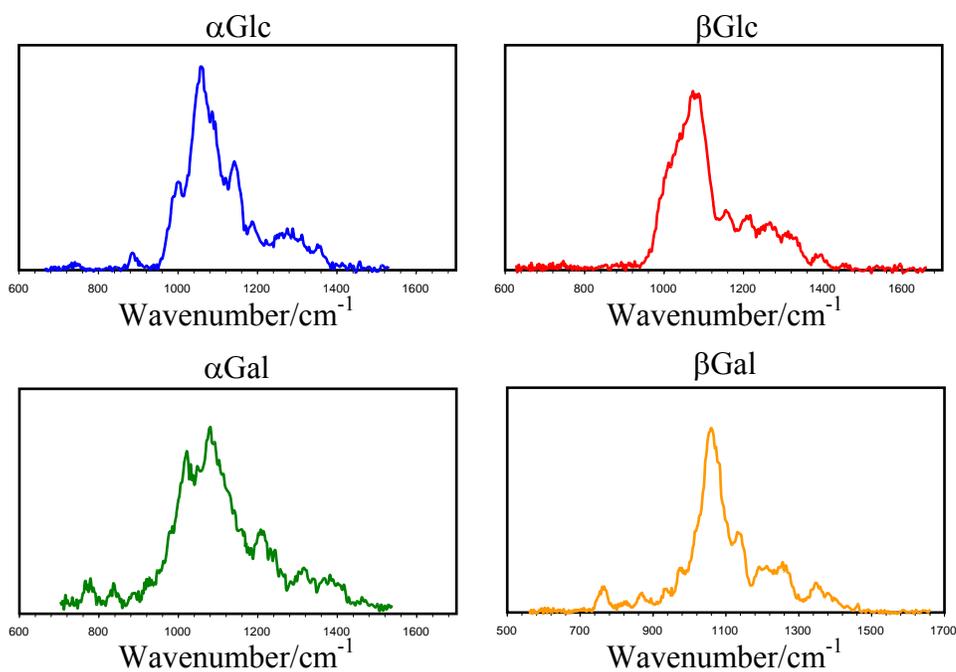


Figure 6-1. FELIX-generated IRMPD spectra (600 – 1700 cm^{-1}) of all four glycoside isomers, αGlc , βGlc , αGal and βGal . Spectra were obtained by Valle *et al.* and included in Jose J. Valle's dissertation.¹⁷

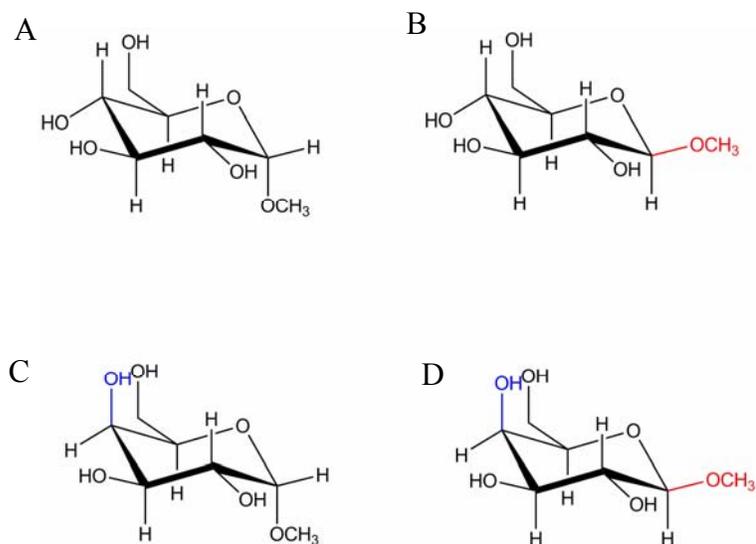


Figure 6-2. O-methylated glycoside isomers, A) α Glc, B) β Glc, C) α Gal and D) β Gal shown in the 4C_1 chair conformation. A difference in structure for beta isomers is shown in red, and differences between glucose and galactose are highlighted in blue.

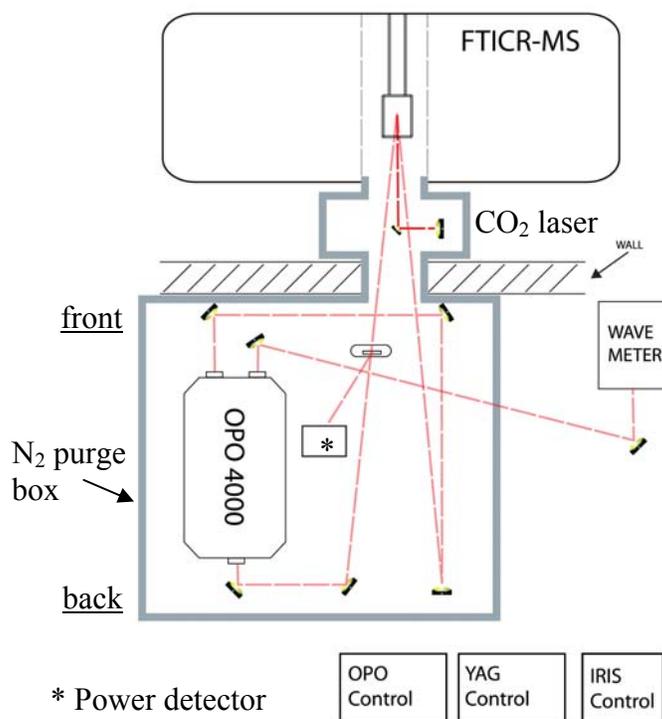


Figure 6-3. OPO optics setup, showing idler beams (as well as CO_2 laser) directed into the ICR cell and signal beam directed to the wavemeter. The setup includes two iris shutters that divert the idler beams into the power detector when they are closed (for clarity, only one iris shutter is shown).

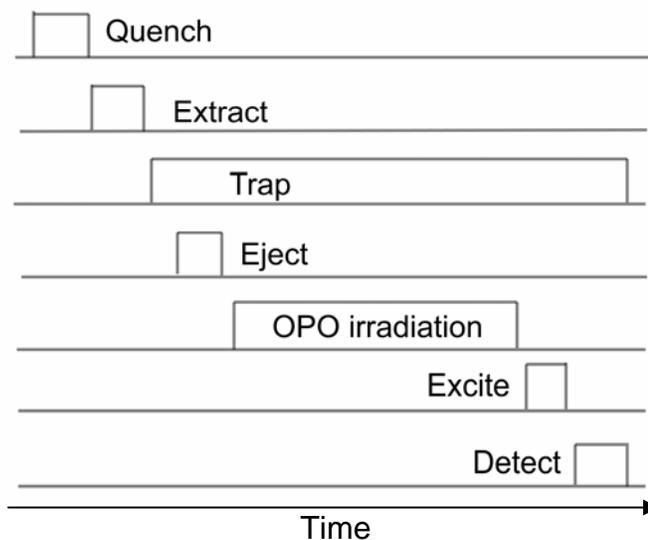


Figure 6-4. Experimental sequence for obtaining spectra of the O-H stretching region of the rubidium cation-attached glycosides using OPO laser irradiation.

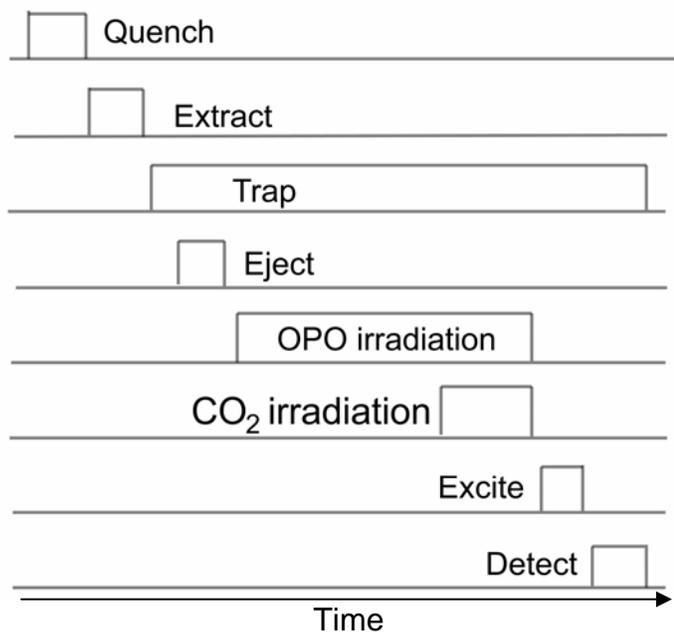


Figure 6-5. Experiment sequence for the two laser experiment. The CO₂ laser output was tuned to be off-resonance with any vibrational modes, and did not promote dissociation when used alone.

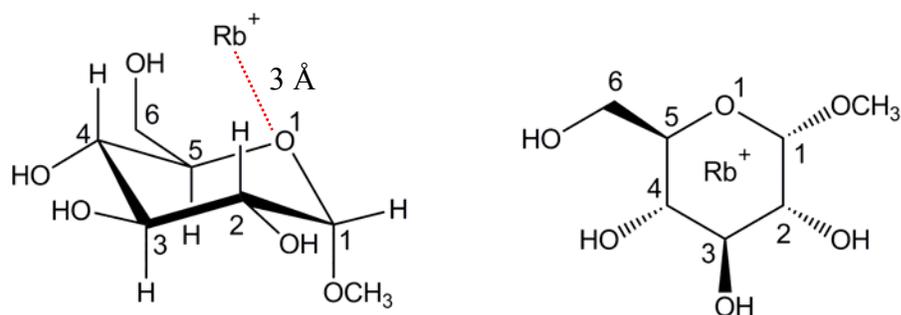


Figure 6-6. The monosaccharide α -D-glucoside (side and top views) with the rubidium cation centered above ring and $\sim 3 \text{ \AA}$ from O1. Ring oxygen (O1) and the carbons of the monosaccharide are numbered.

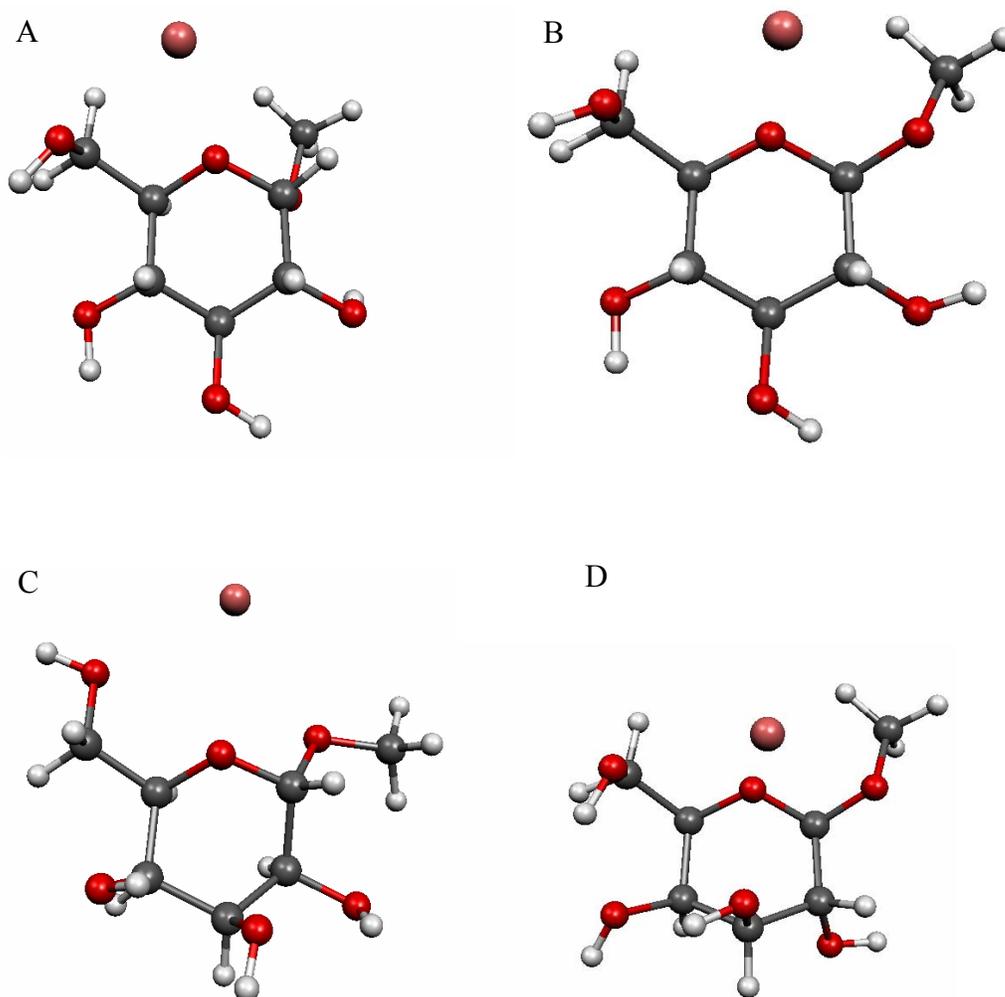


Figure 6-7. Lowest energy conformers of A) α Glc B) β Glc C) α Gal and D) β Gal obtained from each isomers conformational search. Rubidium is shown in purple.

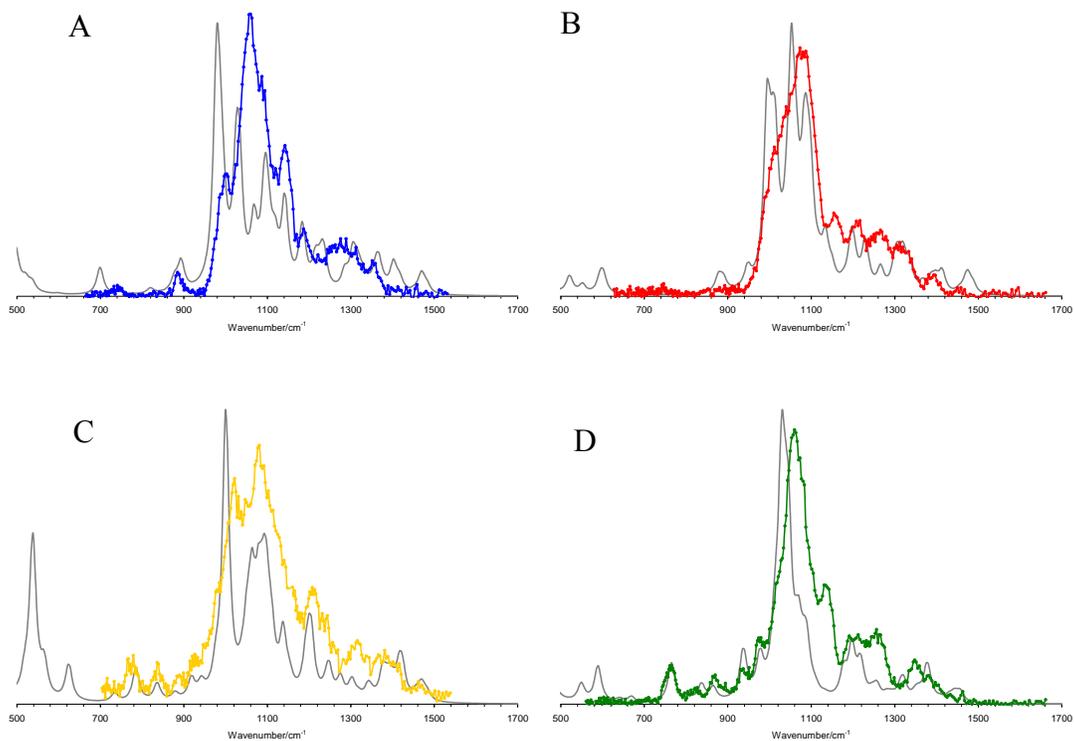


Figure 6-8. Comparison of theoretical (gray) and experimental IRMPD spectra for the isomers A) α Glc, B) β Glc, C) α Gal and D) β Gal. Theoretical spectral frequencies have been scaled by 0.97 and a 20 cm^{-1} Gaussian band profile has been applied.

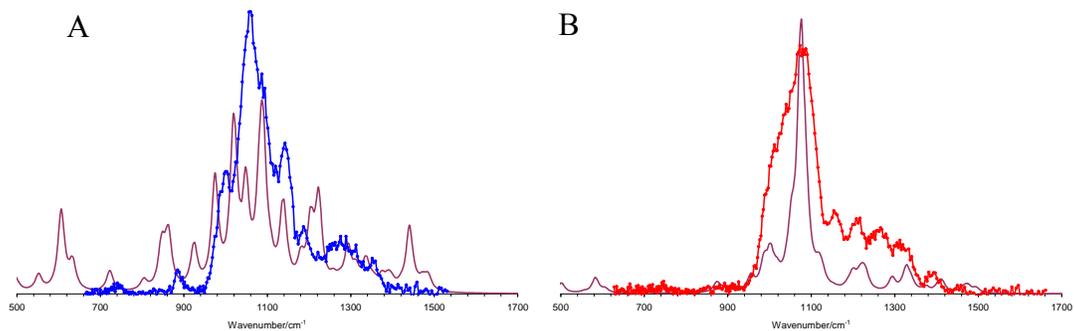


Figure 6-9. Calculated spectra for the second lowest energy conformations of D-glucosides anomers from the conformational search. Labeling is the same as in Figure 6-8. Theoretical spectra of each isomer are shown in purple.

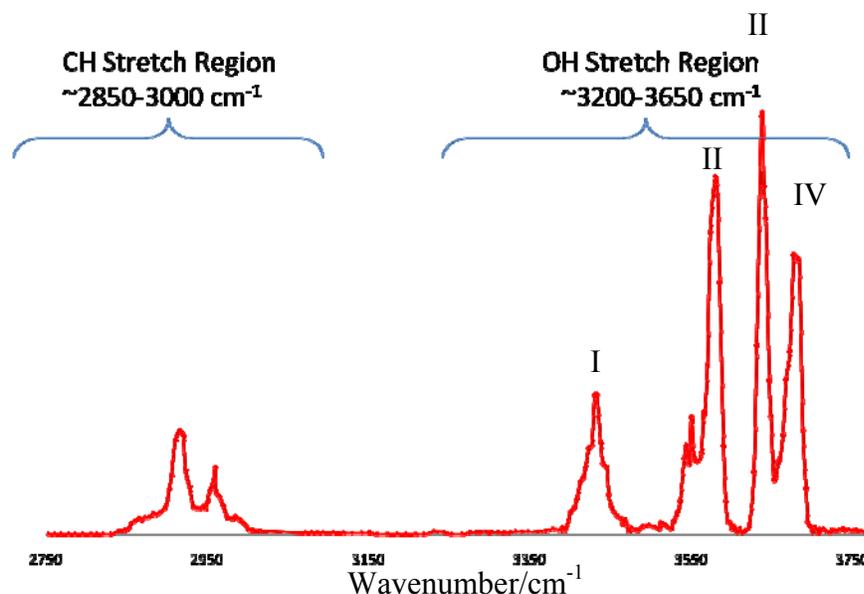


Figure 6-10. IRMPD spectrum of rubidium cation-attached β Glc, with the typical ranges for C-H and O-H stretches indicated. The four major bands in the O-H stretch region are labeled. Image courtesy of Wright L. Pearson.

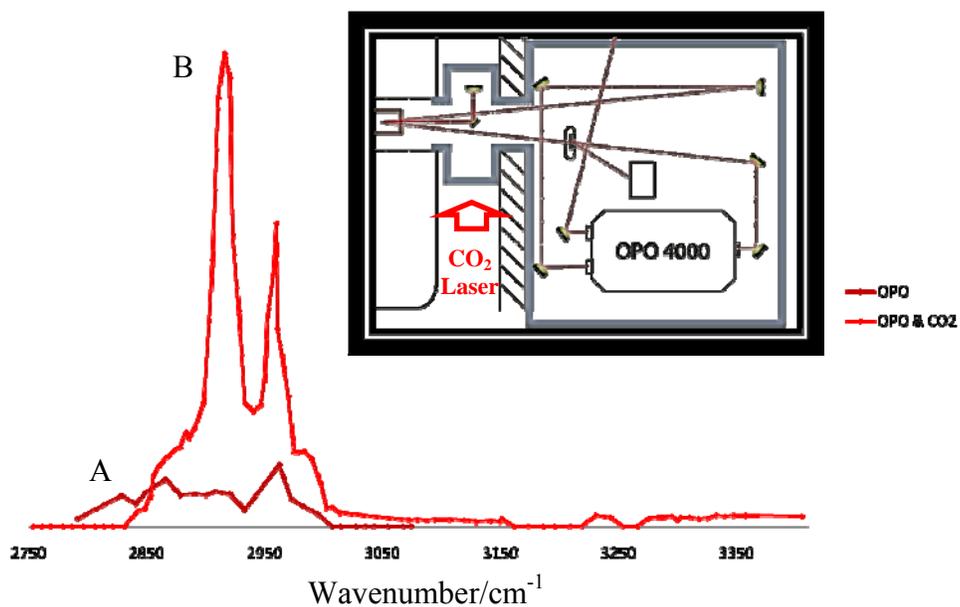


Figure 6-11. Comparison of the IRMPD spectrum of β Glc in the C-H stretch region with A) irradiation with the on-resonance OPO laser and B) irradiation with both the OPO and CO_2 laser (as described in the experimental section) showing the enhancement of the peaks. Image courtesy of Wright L. Pearson.

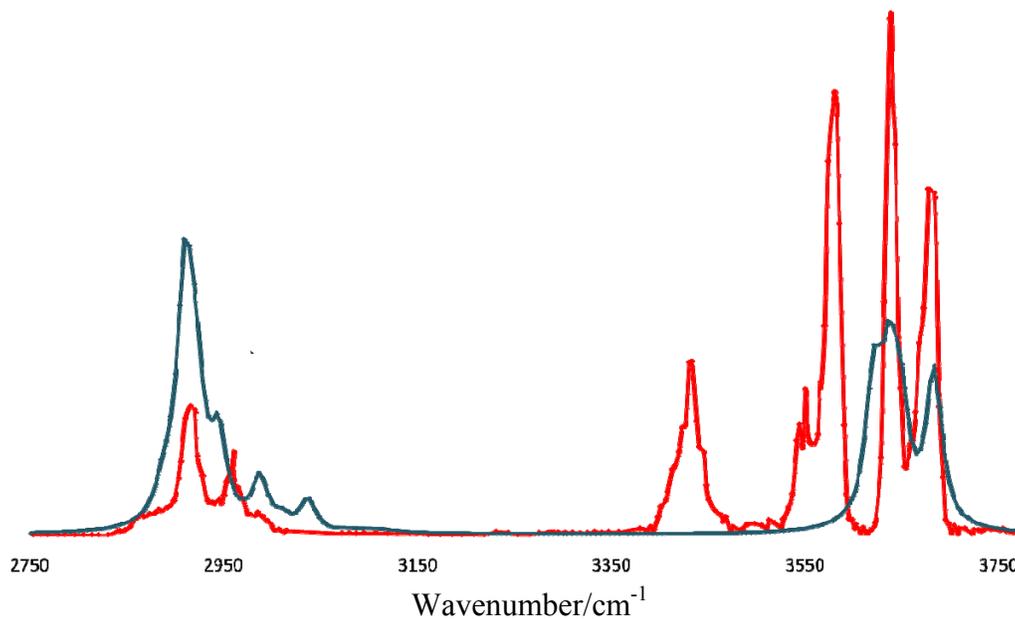


Figure 6-12. Comparison of calculated spectrum for conformer A (dark blue) and IRMPD spectrum (red) for rubidium cation-attached β Glc. Image courtesy of Wright L. Pearson.

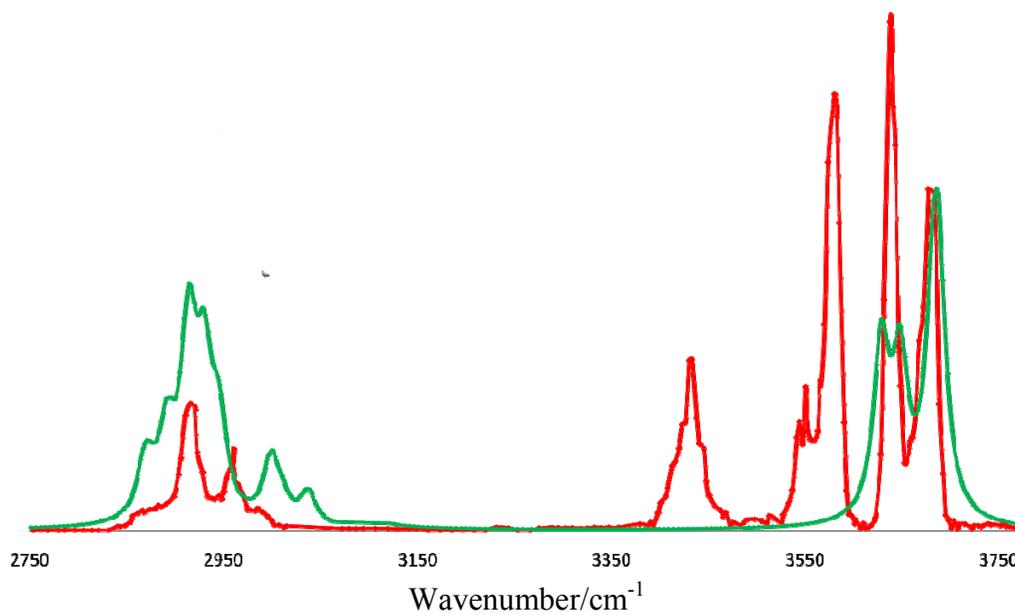


Figure 6-13. Experimental spectrum (red) and calculated spectrum for conformer B (green) of rubidium cation-attached β Glc. Image courtesy of Wright L. Pearson.

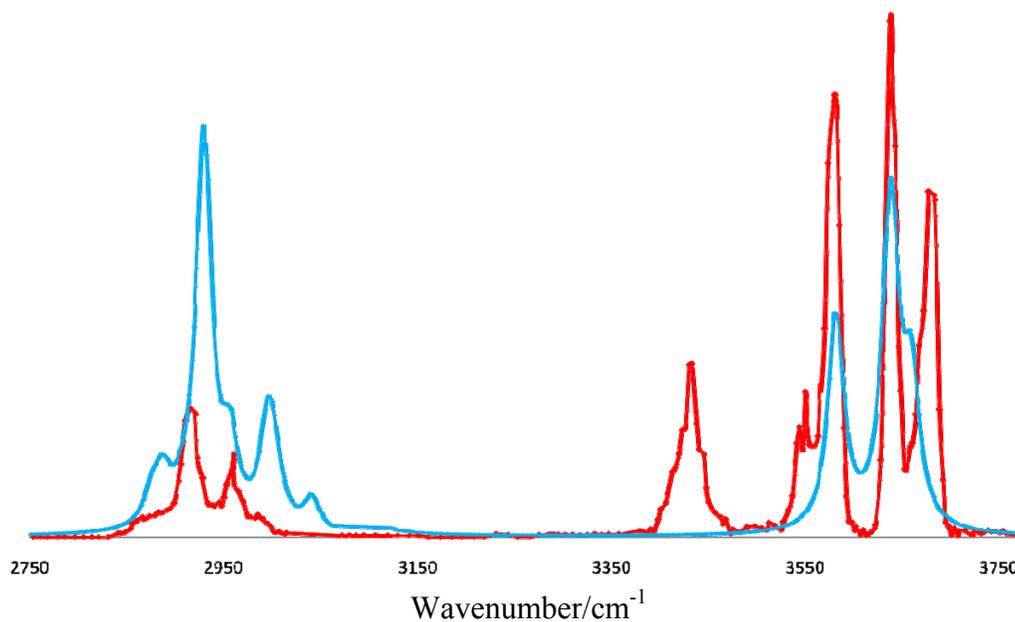


Figure 6-14. Spectral match between experiment (red) and calculated conformer C of rubidium cation-attached β Glc. The red shift toward the band near 3560 cm^{-1} is clearly evident. Image courtesy of Wright L. Pearson.

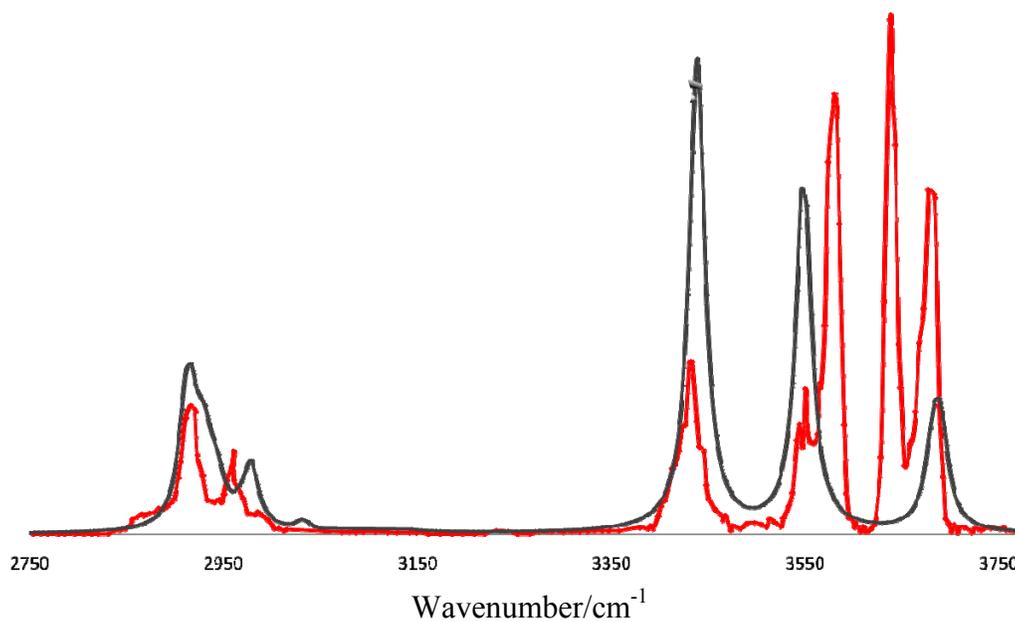


Figure 6-15. Spectral comparison of experiment (red) and calculated conformer D (gray) for rubidium cation-attached β Glc, showing a red shift for O-H stretches of C-3 (3450 cm^{-1}) and C-6 (3560 cm^{-1}). Image courtesy of Wright L. Pearson.

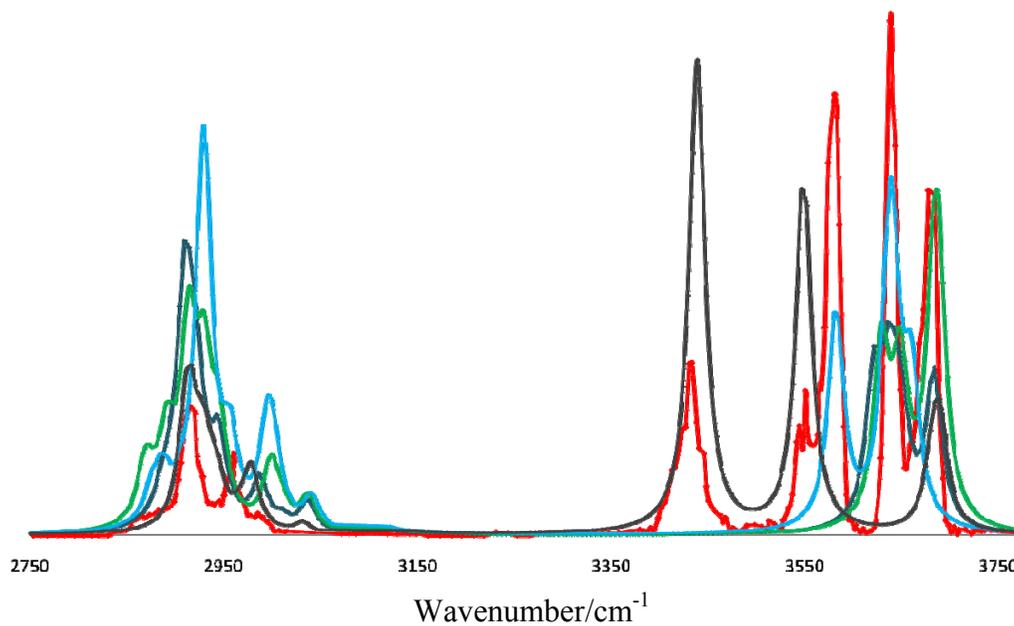


Figure 6-16. Overlap of all four conformer spectra with IRMPD spectrum for rubidium cation-bound β Glc, accounting for all bands present in experimental spectrum. Spectrum for conformer A is in dark blue, conformer B is the green line, conformer C is the blue line, and the gray line is the spectrum of conformer D. Image courtesy of Wright L. Pearson.

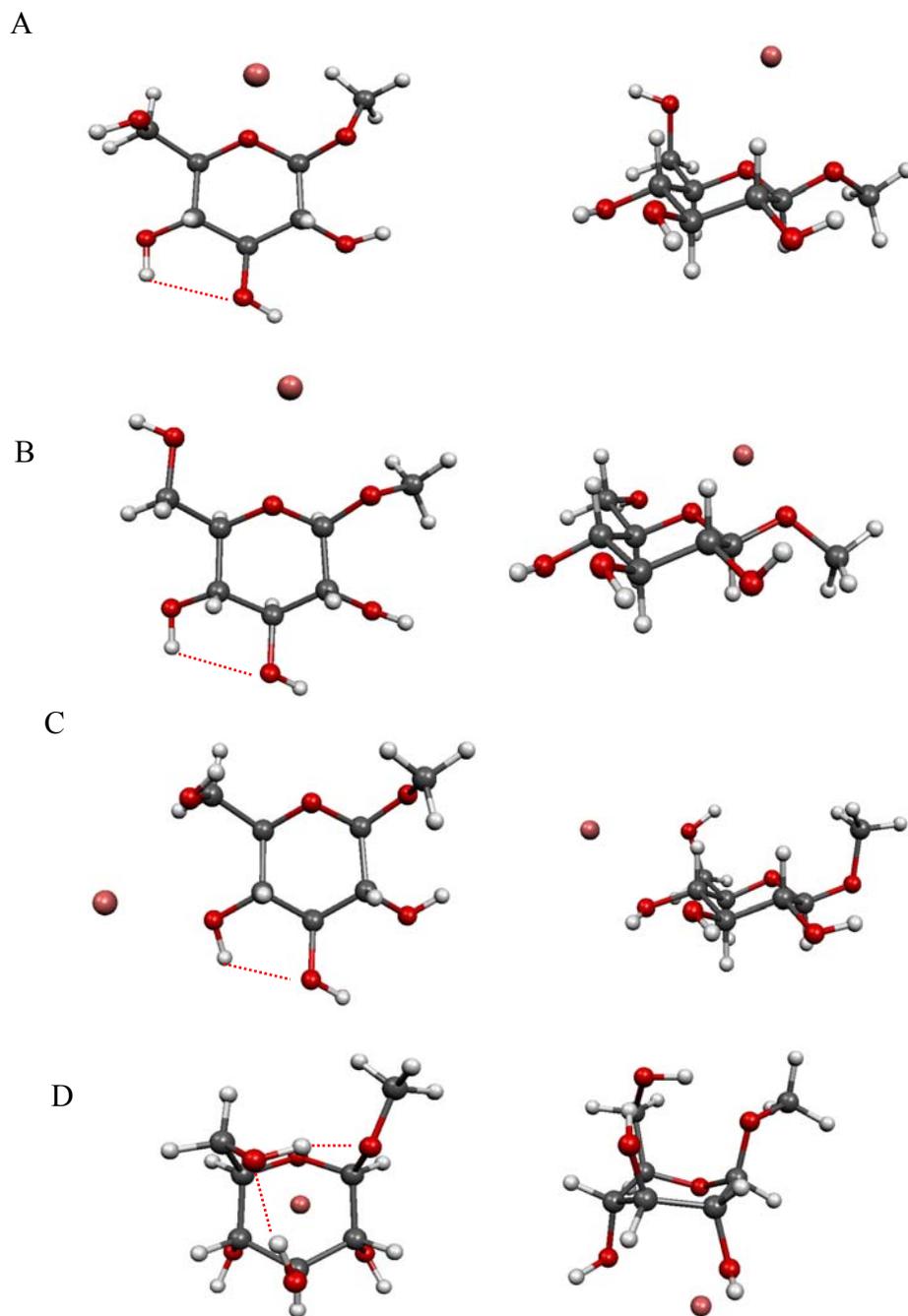


Figure 6-17. Calculated conformers for β Glc at the B3LYP/6-31+G(d) level of theory that account for experimental spectral bands, shown in top and side views. Hydrogen bonding is shown with red dashed dots (top view only) and the rubidium cation is colored tan.

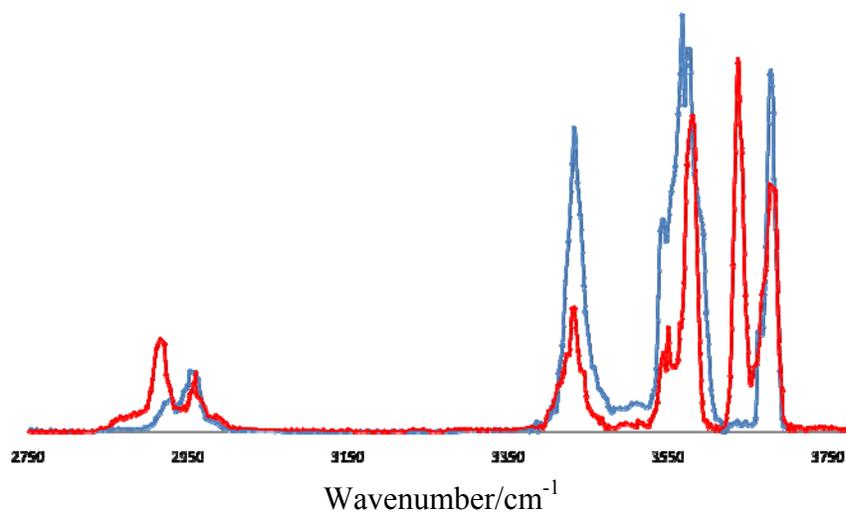


Figure 6-18. Comparison of IRMPD spectra for α Glc (blue) and β Glc (red), clearly showing the band at 3637 cm^{-1} is missing for α Glc. Image courtesy of Wright L. Pearson.

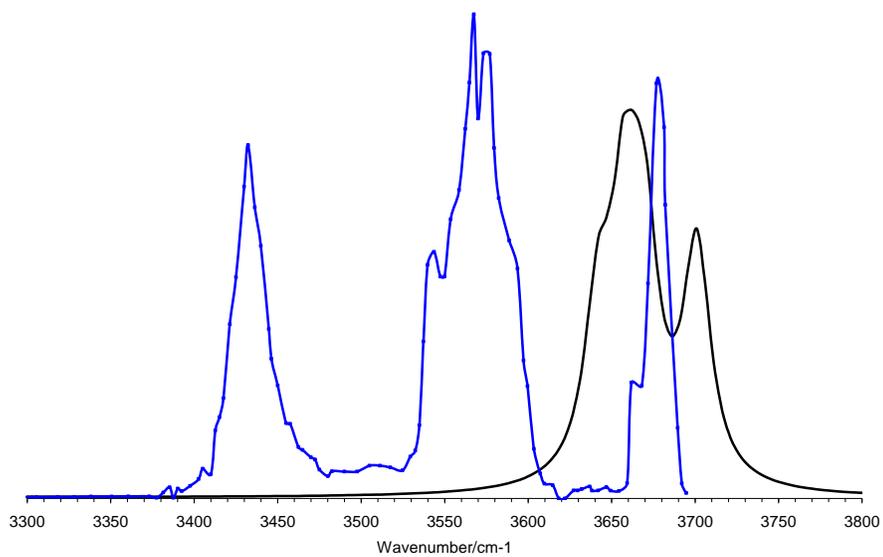


Figure 6-19. Comparison of IRMPD and calculated spectra for the lowest energy conformer of α Glc.

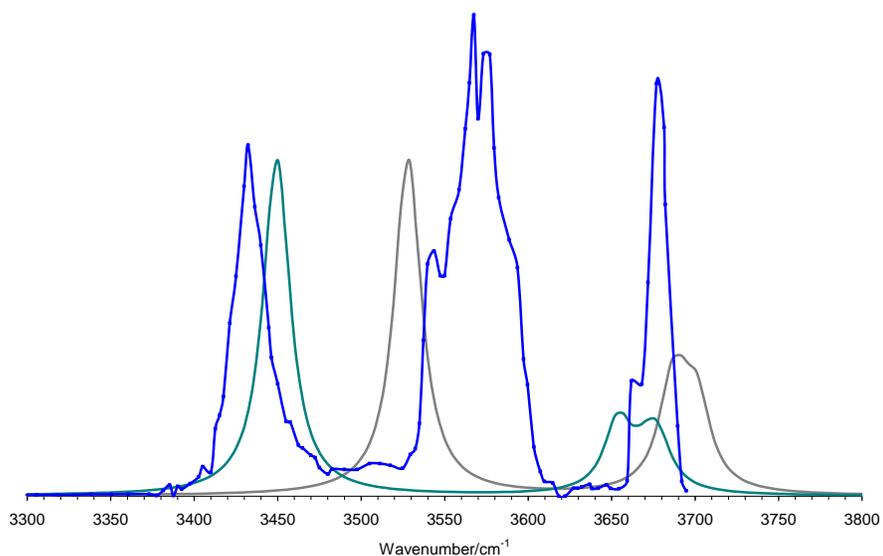


Figure 6-20. Overlap of calculated spectra for conformer A and conformer B with IRMPD spectra of α Glc.

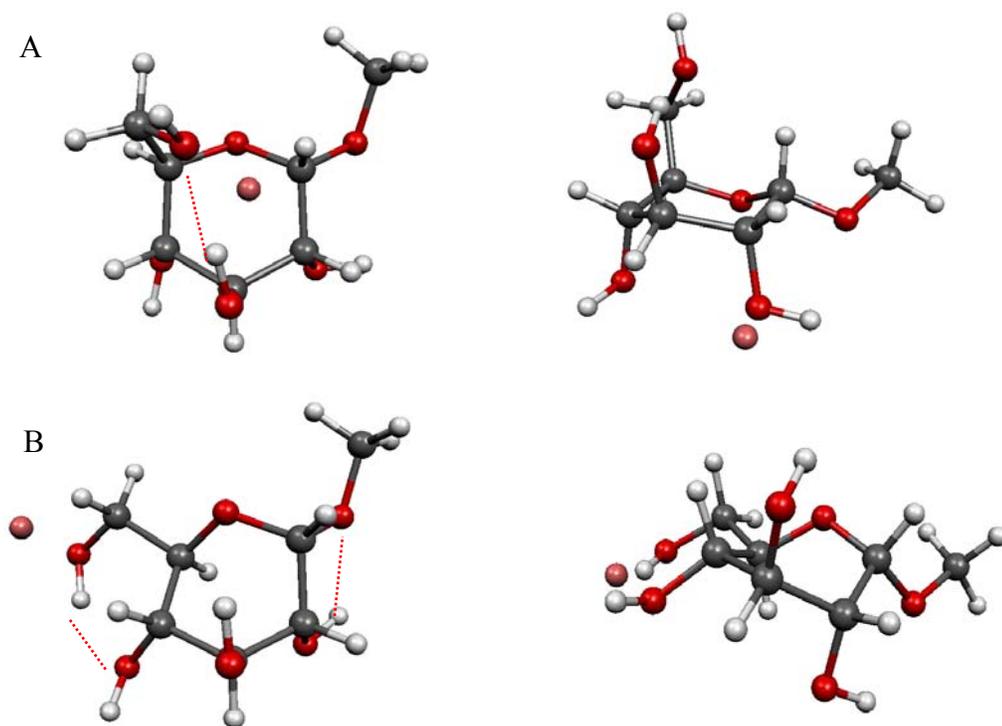


Figure 6-21. Conformers of α Glc that correlate with IRMPD spectrum, shown in top and side views. Hydrogen bonding is marked with red dashed dots (top view only) and the rubidium cation is colored tan. Relative energies are in parenthesis and were calculated at the B3LYP/6-31+G(d) level of theory.

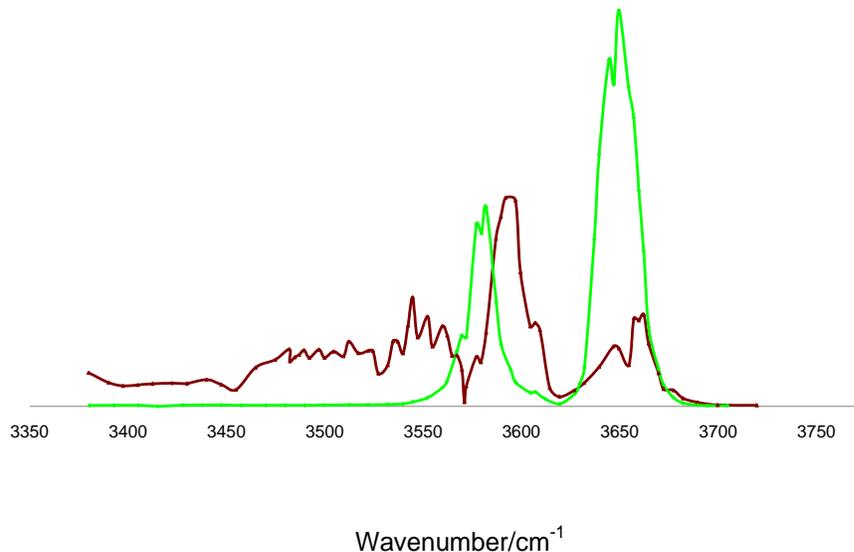


Figure 6-22. Comparison of preliminary IRMPD spectra for α Gal (dark red) and β Gal (green) in the O-H stretch region. A shift occurs for the band near 3580 cm^{-1} , and more spectral features are evident for the α Gal anomer.

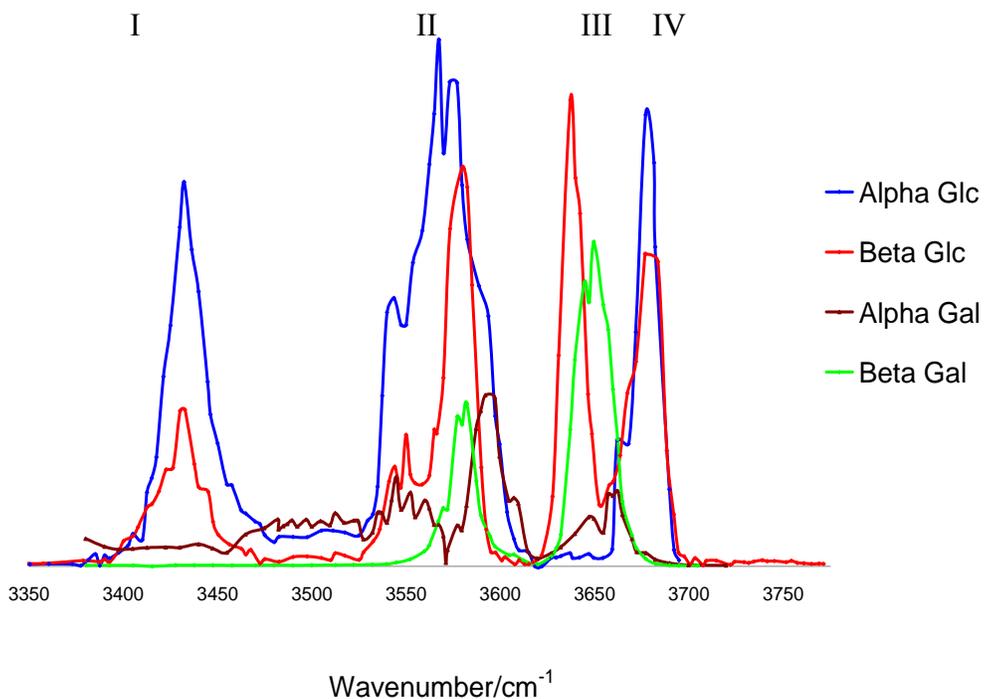


Figure 6-23. Overlap of IRMPD spectra of all four isomers in the O-H stretch region. The major spectral bands are labeled.

CHAPTER 7 CONCLUSIONS AND FUTURE WORK

Fourier transform ion cyclotron resonance mass spectrometry and infrared multiple photon dissociation have been used to obtain gas phase infrared spectra of alkali cation-attached glycosides, N-acetylglycosamines and phenylalanine derivatives. The experiments were combined with theoretical calculations to obtain the best predictions of the structures of the respective ions.

The IRMPD-FTICR-MS experiments proved invaluable in differentiating carbohydrate isomers and in providing well-resolved spectra of phenylalanine analogs. The IRMPD spectra obtained in the 800 – 1800 cm^{-1} were used in conjunction with theoretical calculations to predict the band shifts due to deuterium/hydrogen exchange of sodium cation-attached phenylalanine analogs. It was also possible to differentiate the α - and β -anomers of N-acetylglycosamines by observing the shift in the C=O stretch. IRMPD spectra in the 800- 1800 cm^{-1} were also obtained for rubidium cation-attached glycosides, but the band features were quite broad and differentiation between the respective isomers was inconclusive. Instead, using an optical parametric oscillator coupled to an FTICR-MS, IRMPD spectra in the 2750 – 3800 cm^{-1} range were obtained for the rubidium cation-attached glycosides. Differentiation was possible because the bands were narrow (12 – 45 cm^{-1}), and each isomer displayed distinct peaks in the O-H stretch region (3400 – 3800 cm^{-1}).

Fragmentation of the precursor ion as a function of irradiation wavelength was also useful in obtaining structural information and alkali metal cation binding locations. Binding of the lithium cation to carbohydrates induces fragmentation of the sugar when subjected to irradiation with an infrared laser. For the lithium cation-attached N-acetylglycosamines, the fragment ions with m/z 210, 132, 127 and 122 were observed after irradiation and subsequent fragmentation of

each precursor ion. Fragmentation ion abundances for α -GalNac, β -GalNac and β -GlcNac were similar, where the order of abundance of the fragment ions was $127 > 210 > 132 \sim 122$. However, α -GlyNac displayed a distinct abundance order: $122 \sim 127 > 132 > 210$. The 122 m/z fragment ion arises from the lithium cation attached to a cross ring cleavage fragment containing the N-acetyl group. The calculated structures support this result, indicating that the lithium cation is located between the C-3 hydroxyl group and the N-acetyl group for α -GlcNac. For the other three isomers, the Li^+ is located near the ring oxygen.

Theoretical calculations were an integral part of this work. The Conformational searches, in conjunction with the DFT calculations, provided good spectral matching with experimental spectra for the phenylalanine analogs. The lowest energy structures obtained in the conformational searches gave the best matches to the IRMPD spectra for all three phenylalanine analogs, and band assignments were based on the results of the DFT calculations. For deuterium/hydrogen exchange experiments, the bands that involved R-NH and R-OH vibrations in the undeuterated analog were expected to shift after each analog was subjected to H/D exchange and subsequent IRMPD spectra were taken. The bands shifts were correctly predicted by theory. After confirmation with experiment, the calculated structures of the phenylalanine analogs were found to have a “puckered” structure, in which all the carbonyl oxygens interact with Na^+ and the phenyl ring. A cation- π interaction was found for all of the lowest energy structures.

The conformational search and DFT calculations approach was also adequate to obtain structures for N-acetylglycosamines. The calculations indicated that the spectral band shifts are due to the location of the carbonyl of the N-acetyl group. The orientation of the N-acetyl group and the relative position of the carbonyl are influenced by the repulsion between the O-methyl

oxygen and the carbonyl oxygen. The C=O band shift seems to be a localized effect so that differentiation of anomers is possible. However, IRMPD could not differentiate between D-GalNac and D-GlcNac. For example, in the IRMPD spectra of β -GlcNac and β -GalNac, the C=O stretch bands overlap almost completely. A similar situation was observed with α -GlcNac and α -GalNac, whose C=O stretch bands overlap, thus preventing differentiation of the two isomers using only their IRMPD spectra. However, theoretical calculations were able to predict the band shifts by finding the orientation of the carbonyl in the N-acetyl group of each glycosamine isomer.

Conformational searches and DFT calculations did not provide an adequate match for the IRMPD spectra of the rubidium cation-attached glycosides. Although similar to the N-acetylglycosamines, the free glycoside sugars can form extensive hydrogen bonded networks that may not be adequately modeled with the levels of theory used. The rubidium cation requires that the empirical AMBER force field be used in the conformational search, instead of the semi-empirical AM1 force field used for other two projects. Probably more important is the need to use an effective core potential basis set in the DFT calculations involving rubidium. This most likely increases the discrepancies between calculated and experimental spectra of the rubidium cation-attached glycosides. In contrast to the N-acetylglycosamines, all the hydroxyl groups were probed in the IRMPD spectra of the Rb^+ glycosides and therefore theoretical calculations were required to find the correct orientations. The hydroxyl groups are distinct from the carbonyl group, in that they may be involved in extensive hydrogen bonded networks and are easily influenced by external perturbations, such as the rubidium cation. The ring conformation also affects the hydroxyl's orientation and the ring flexing is more cumbersome to model. Aside from these drawbacks, the calculations for the rubidium cation-attached glycosides did seem to

indicate that the major influences in O-H band shifts are due to the spatial orientation particular to each isomer (α or β , glucose or galactose) and the resulting effects on the strength of the hydrogen bonded network. Theoretical calculations also indicate that the rubidium cation location perturbs the O-H stretches and the hydrogen bonded network, thus making differentiation of all four glycoside isomers possible

As with any project, future studies need to be undertaken to gain deeper insight into the chemistry of the molecules being studied. For phenylalanine analogs, IRMPD spectra of deuterated and undeuterated proton-bound Phe analogs would provide structural information about the ions. The experiments and calculations would help explain why gas phase H/D exchange at the acetamido group is observed for proton-bound AcPhe but not for the corresponding sodium cation-attached AcPhe. Calculated and IRMPD spectra of sodium cation-attached di- and tripeptides containing phenylalanine are needed to observe if lack of exchange at the amide group occurs for the larger species.

For the N-acetylglycosamines, it would be useful to obtain IRMPD spectra in the range of $2600 - 3900 \text{ cm}^{-1}$ using the OPO/FTICR-MS setup, in order to confirm the lowest energy calculated structures used as a comparison to IRMPD spectra in the $800 - 1800 \text{ cm}^{-1}$. IRMPD spectra of rubidium cation-bound GlcNac and GalNac's isomers would be useful to determine if the C=O stretch undergoes a similar shifting pattern observed for the Li^+ attached glycosamines. This would be especially helpful for α -GlcNac, whose lowest energy structure has lithium bound to the carbonyl group itself, and is probably affecting the shift. A gas phase H/D exchange experiment and IRMPD spectra are also needed for the deuterated GlcNac and GalNac ions.

Although all four isomers of the D-glycosides can be differentiated by IRMPD, theoretical structures are needed for α -D-glucoside and β -D-glucoside encompassing all the IRMPD

spectral features. Theoretical calculations of D-galactosides are also needed and are underway to explore more of the conformational space than the previous conformational searches. Molecular dynamic simulations that mimic the experimental conditions may prove useful, so that a more representative set of conformations can be obtained. Experiments similar to those of lithium cation-attached N-acetylglycosamines, but using lithium cation-attached glycosides, are needed to obtain fragmentation patterns as a function of irradiation wavelengths to provide structural information. As a more extensive and long term project, the OPO/FTICR-MS setup could be used to obtain IRMPD spectra for the other D-hexoses to test if differentiation can be established for the entire set of isomeric monosaccharides. Theoretical calculations, H/D exchange experiments and IRMPD spectra of lithium-bound species have proven useful for all the work herein presented, and they would also be used in the analysis of the D-hexoses to obtain structural information.

LIST OF REFERENCES

- (1) Dickerson, R.; Geis, I. *Chemistry, Matter, and the Universe: An Integrated Approach to General Chemistry*; Benjamin-Cummings: New York, 1976.
- (2) Ramberg, P. *Chemical Structure, Spatial Arrangement: The Early History of Stereochemistry, 1874-1914*; Ashgate Publishing: Vermont, 2003.
- (3) El Khadem, H. *Carbohydrate chemistry: monosaccharides and their oligomers*; Academic Press: New York, 1997.
- (4) Davis, B.; Fairbanks, A. *Carbohydrate chemistry*; Oxford University: New York, 2002.
- (5) McMurry, J. *Organic Chemistry*; Brooks.Cole: Pacific Grove, California, **1992**.
- (6) Cahn, R.; Ingold, C.; Prelog, V. *Angew. Chem. Int. Ed. Eng.* **1966**, 5, 385.
- (7) Becker, D.; Lowe, J. *Glycobiology* **2003**, 13, 41.
- (8) Hauser, G.; Karnovsky, M. *J. Biol. Chem* **1958**, 233, 287.
- (9) Wang, Y.; Zhang, L.; Ye, X. *Comb. Chem. High Throughput Screening* **2006**, 9, 63.
- (10) Wang, C.; Lee, J.; Luo, S.; Kulkarni, S.; Huang, Y.; Lee, C.; Chang, K.; Hung, S. *Nature* **2007**, 446, 896.
- (11) Stick, R. *Carbohydrates: The Sweet Molecules of Life*; Academic Press: New York, 2001.
- (12) Smith, D.; Woerpel, K. *Org. Biomol. Chem.* **2006**, 4, 1195.
- (13) Centerwall, S.; Centerwall, W. The Discovery of Phenylketonuria: The Story of a Young Couple, Two Retarded Children, and a Scientist. In *Pediatrics*; *Am. Acad. Pediatrics*, 2000; Vol. 105; pp 89.
- (14) Gapeev, A.; Dunbar, R. C. *J. Am. Chem. Soc.* **2001**, 123, 8360.
- (15) Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, 3, 1.
- (16) Rožman, M. *J. Am. Soc. Mass Spectrom.* **2005**, 16, 1846.
- (17) Valle, J. J. Differentiation of Carbohydrate Stereoisomers by Infrared Multiple Photon Dissociation using a Free Electron Laser and a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. Ph.D. Thesis, University of Florida, 2005.
- (18) McNaught, A. *Carbohydr. Res.* **1997**, 297, 1.
- (19) Dempster, A. *Proc. Am. Phil. Soc* **1935**, 75, 755.

- (20) Aston, F. *Mass spectra and isotopes*; Edward Arnold: London, 1942.
- (21) Stephens, W. *Phys. Rev.* **1946**, *69*, 46.
- (22) Dawson, P. *Quadrupole Mass Spectrometry and Its Applications*; American Institute of Physics: New York, 1995.
- (23) *Nobel Lectures, Chemistry 1901-1921*; Elsevier Publishing Company: Amsterdam, 1966.
- (24) Lawrence, E. *Phys. Rev.* **1926**, *28*, 947.
- (25) Munson, M.; Field, F. *J. Am. Chem. Soc.* **1966**, *88*, 2621.
- (26) Suming, H.; Yaozu, C.; Longfei, J.; Shuman, X. *Org. Mass Spectrom.* **1985**, *20*, 719.
- (27) Surman, D.; Vickerman, J. *J. Chem. Soc., Chem Commun* **1981**, *1981*, 324.
- (28) Barber, M.; Bordoli, R.; Sedgwick, R.; Tyler, A. *J. Chem. Soc., Chem. Commun.* **1981**, *1981*, 325.
- (29) Williams, D.; Bradley, C.; Santikarn, S.; Bojesen, G. *Biochem. J.* **1982**, *201*, 105.
- (30) Comisarow, M.; Marshall, A. Fourier transform ion cyclotron resonance spectroscopy method and apparatus; Google Patents, 1976.
- (31) Marshall, A. *Acc. Chem. Res.* **1985**, *18*, 316.
- (32) Marshall, A.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1.
- (33) Marshall, A. *Int. J. Mass Spectrom.* **2000**, *200*, 331.
- (34) Marshall, A.; Hendrickson, C. *Int. J. Mass Spectrom.* **2002**, *215*, 59.
- (35) Marshall, A. G.; Hendrickson, C. L.; Emmett, M. R.; Rodgers, R. P.; Blakney, G. T.; Nilsson, C. L. *Eur. J. Mass Spectrom.* **2007**, *13*, 57.
- (36) Brobeck, W.; Lawrence, E.; MacKenzie, K.; McMillan, E.; Serber, R.; Sewell, D.; Simpson, K.; Thornton, R. *Phys. Rev.* **1947**, *71*, 449.
- (37) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53.
- (38) Daniel, R.; Chevolot, L.; Carrascal, M.; Tissot, B.; Mourão, P. A. S.; Abian, J. *Carbohydr. Res.* **2007**, *342*, 826.

- (39) Boyd, R. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 1191.
- (40) Mann, M.; Wilm, M. *Trends Biochem. Sci.* **1995**, *20*, 219.
- (41) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882.
- (42) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Mass Spectrom. Rev.* **1990**, *9*, 37.
- (43) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246*, 64.
- (44) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.
- (45) Striegel, A.; Timpa, J.; Piotrowiak, P.; Cole, R. *Int. J. Mass Spectrom. Ion Processes* **1997**, *162*, 45.
- (46) Siuzdak, G. *Mass Spectrometry for Biotechnology*; Academic Press: New York, 1996.
- (47) Caravatti, P.; Allemann, M. *Org. Mass. Spectrom* **1991**, *26*, 514.
- (48) Comisarow, M. *Int. J. Mass Spectrom. Ion Phys.* **1981**, *37*, 251.
- (49) Schaub, T.; Hendrickson, C.; Horning, S.; Quinn, J.; Senko, M.; Marshall, A. *Anal. Chem.* **2008**.
- (50) Zhou, Z.; Ogden, S.; Leary, J. *J. Org. Chem.* **1990**, *55*, 5444.
- (51) Cancilla, M.; Penn, S.; Carroll, J.; Lebrilla, C. *J. Am. Chem. Soc* **1996**, *118*, 6736.
- (52) Cancilla, M.; Wong, A.; Voss, L.; Lebrilla, C. *Anal. Chem* **1999**, *71*, 3206.
- (53) Bomse, D.; Woodin, R.; Beauchamp, J. *J. Am. Chem. Soc.* **1979**, *101*, 5503.
- (54) Woodin, R.; Bomse, D.; Beauchamp, J. *J. Amer. Chem. Soc* **1978**, *100*, 3248.
- (55) Patel, C. *Phys. Rev.* **1964**, *136*, 1187.
- (56) Witteman, W. *The CO₂ laser*; Springer-Verlag: Berlin; New York, 1987.
- (57) Duarte, F. *Tunable Laser Applications*; CRC Press: Boca Raton, Florida, 1995.
- (58) Peiris, D. M.; Cheeseman, M. A.; Ramanathan, R.; Eyler, J. R. *J. Phys. Chem.* **1993**, *97*, 7839.

- (59) Peiris, D. M.; Riveros, J. M.; Eyler, J. R. *Int. J. Mass Spectrom. Ion Processes* **1996**, *159*, 169.
- (60) Shin, S.; Beauchamp, J. *J. Am. Chem. Soc.* **1990**, *112*, 2066.
- (61) Bosenberg, W. R.; Drobshoff, A.; Alexander, J. I.; Myers, L. E.; Byer, R. L. *Optics Lett.* **1996**, *21*, 713.
- (62) Snoek, L.; Robertson, E.; Kroemer, R.; Simons, J. *Chem. Phys. Lett.* **2000**, *321*, 49.
- (63) Jockusch, R.; Kroemer, R.; Talbot, F.; Snoek, L.; Carcabal, P.; Simons, J.; Havenith, M.; Bakker, J.; Compagnon, I.; Meijer, G. *J. Am. Chem. Soc.* **2004**, *126*, 5709.
- (64) Carcabal, P.; Jockusch, R. A.; Hunig, I.; Snoek, L. C.; Kroemer, R. T.; Davis, B. G.; Gamblin, D. P.; Compagnon, I.; Oomens, J.; Simons, J. P. *J. Am. Chem. Soc.* **2005**, *127*, 11414.
- (65) Verhoeven, T.; Bongers, W.; Bratman, V.; Caplan, M.; Denisov, G.; van der Geer, C.; Manintveld, P.; Poelman, A.; Plomp, J.; Savilov, A. *Plasma Science, IEEE Trans.* **1999**, *27*, 1084.
- (66) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Rev. Sci. Instr.* **2005**, *76*, 023103/1.
- (67) The World Wide Web Virtual Library: Free Electron Laser, 2008.
- (68) Fukui, K.; Takada, Y.; Sumiyoshi, T.; Imai, T.; Takahashi, K. *J. Phys. Chem. B* **2006**, *110*, 16111.
- (69) Moore, D. T.; Oomens, J.; Van Der Meer, L.; Von Helden, G.; Meijer, G.; Valle, J.; Marshall, A. G.; Eyler, J. R. *Chem. Phys. Chem.* **2004**, *5*, 740.
- (70) Lucas, B.; Grégoire, G.; Lemaire, J.; Maître, P.; Glotin, F.; Schermann, J.; Desfrancois, C. *Int. J. Mass Spectrom.* **2005**, *243*, 105.
- (71) Lucas, B.; Grégoire, G.; Lemaire, J.; Maître, P.; Ortega, J.; Rupenyan, A.; Reimann, B.; Schermann, J.; Desfrancois, C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2659.
- (72) von Helden, G.; Compagnon, I.; Blom, M. N.; Frankowski, M.; Erlekam, U.; Oomens, J.; Brauer, B.; Gerber, R. B.; Meijer, G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1248.
- (73) Guan, S.; Marshall, A. *Int. J. Mass Spectrom. Ion Processes* **1996**, *157*, 5.
- (74) Hixson, K. K. FTICR Theory; Environmental Molecular Sciences Laboratory. http://www.emsl.pnl.gov/docs/msd/mass_spec/home/ftictut.html, 2000; Vol. 2008.

- (75) Kkmurray. Electropray diagram. GNU Free Documentation License, Creative Commons, 2007.
- (76) Koch, W.; Holthausen, M. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: New York, 2001.
- (77) Hemmingsen, L.; Madsen, D. E.; Esbensen, A. L.; Olsen, L.; Engelsen, S. B. *Carbohydr. Res., Molecular Modeling of Carbohydrates* **2004**, 339, 937.
- (78) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T. *J. Comp. Chem.* **2003**, 24, 1999.
- (79) Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *Q. Rev. Biophys* **1993**, 26, 49.
- (80) Jorgensen, W.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, 110, 1657.
- (81) MacKerell, A.; Bashford, D.; Bellott, M.; Dunbrack, R.; Evanseck, J.; Field, M.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. *J. Phys. Chem. B* **1998**, 102, 3586.
- (82) Ypma, T. *SIAM Review* **1995**, 37, 531.
- (83) Hestenes, M.; Stiefel, E. *J. Res. Nat. Bur. Stand* **1952**, 49, 409.
- (84) Hildebrand, F. *Introduction to Numerical Analysis*; McGraw-Hill: Columbus, Ohio 1974.
- (85) French, A.; Brady, J. *Computer modeling of carbohydrate molecules*; American Chemical Society: Washington, DC, 1990.
- (86) Kirkpatrick, S.; Gelatt, C.; Vecchi, M. *Science* **1983**, 220, 671.
- (87) Swendsen, R.; Wang, J. *Phys. Rev. Lett.* **1986**, 57, 2607.
- (88) *HyperChem: Computational Chemistry.*; Hypercube: Gainesville, Florida, 1996.
- (89) Naidoo, K.; Brady, J. *Chem. Phys.* **1997**, 224, 263.
- (90) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, 111, 4379.
- (91) Atzrodt, J.; Fey, T.; Zimmermann, J. *Angew. Chem. Int. Ed.* **2007**, 46, 7744.
- (92) Kohen, A.; Limbach, H. *Isotope Effects in Chemistry and Biology*; CRC Press: Boca Raton, Florida, 2006.
- (93) Campbell, S.; Rodgers, M.; Marzluff, E.; Beauchamp, J. *J. Am. Chem. Soc.* **1995**, 117, 12840.

- (94) Reid, G.; O'Hair, R.; Styles, M.; McFadyen, W.; Simpson, R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1701.
- (95) Moore, D. T.; Oomens, J.; Eyler, J. R.; von Helden, G.; Meijer, G.; Dunbar, R. C. *J. Am. Chem. Soc.* **2005**, *127*, 7243.
- (96) Peiris, D. M.; Yang, Y.; Ramanathan, R.; Williams, K. R.; Watson, C. H.; Eyler, J. R. *Int. J. Mass Spectrom. Ion Processes* **1996**, *157/158*, 365.
- (97) Dunbar, R. C.; Moore, D. T.; Oomens, J. *J. Phys. Chem. A* **2006**, *110*, 8316.
- (98) Szczepanski, J.; Wang, H.; Vala, M.; Tielens, A. G. G. M.; Eyler, J. R.; Oomens, J. *Astrophys. J.* **2006**, *646*, 666.
- (99) *Analytical Chemistry*. Washington, DC, United States, **2008**, *80*, 1.
- (100) Polfer, N. C.; Oomens, J.; Dunbar, R. C. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2744.
- (101) Polfer, N. C.; Oomens, J.; Dunbar, R. C. *Chem. Phys. Chem.* **2008**, *9*, 579.
- (102) Wu, G.; Stace, A. J. *Chem. Phys. Lett.* **2005**, *412*, 1.
- (103) Ho, Y.-P.; Yang, Y.-C.; Klippenstein, S. J.; Dunbar, R. C. *J. Phys. Chem.* **1995**, *99*, 12115.
- (104) Thomas E. Wales, J. R. E. *Mass Spectrom. Rev.* **2006**, *25*, 158.
- (105) Tian, Z.; Pawlow, A.; Poutsma, J. C.; Kass, S. R. *J. Am. Chem. Soc.* **2007**, *129*, 5403.
- (106) Ruan, C.; Rodgers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 14600.
- (107) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321*, 49.
- (108) Polfer, N. C.; Oomens, J.; Moore, D. T.; Von Helden, G.; Meijer, G.; Dunbar, R. C. *J. Am. Chem. Soc.* **2006**, *128*, 517.
- (109) Dunbar, R. C. *J. Phys. Chem. A* **2000**, *104*, 8067.
- (110) Ryzhov, V.; Dunbar, R. C.; Cerda, B.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1037.
- (111) Bubb, W. *Concepts in Magnetic Resonance, Part A* **2003**, *19*, 1.
- (112) Dunbar, R. C. *Techniques of Chemistry (New York)* **1986**, *6*, 903.

- (113) Baykut, G.; Eyler, J. R. *TrAC, Trends in Anal. Chem.* **1986**, *5*, 44.
- (114) Dunbar, R. C. *Ft-Icr/Ms* **1991**, 1.
- (115) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133
- (116) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864
- (117) Marshall, A.; Wang, T.; Ricca, T. *J. Am. Chem. Soc.* **1985**, *107*, 7893.
- (118) Chen, L.; Wang, T.; Ricca, T.; Marshall, A. *Anal. Chem.* **1987**, *59*, 449.
- (119) Chen, L.; Marshall, A. *Rapid Commun. Mass Spectrom.* **1987**, *1*, 39.
- (120) Guan, S.; Marshall, A. *Anal. Chem.* **1993**, *65*, 1288.
- (121) Eyler, J. R. *Abstracts of Papers, 223rd ACS National Meeting, Orlando, FL, United States, April 7-11, 2002* **2002**, CHED.
- (122) Oepts, D.; van der Meer, A.; van Amersfoort, P. *Infrared Phys. Tech.* **1995**, *36*, 297.
- (123) HyperChem(TM) Professional; 7.51 ed.; Hypercube, Inc.
- (124) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03; D.02 ed.*; Gaussian, Inc.: Wallingford CT, 2004.
- (125) Hariharan, P. *J. Chem. Phys.* **1982**, *77*, 3654.
- (126) Rassolov, V.; Pople, J.; Ratner, M.; Windus, T. *J. Chem. Phys.* **1998**, *109*, 1223.
- (127) Hehre, W. J. D., R.; Pople, J.A. *J. Chem. Phys* **1972**, *56*, 2257.
- (128) Clark, T.; Chandrasekhar, J.; Spitznagel, G.; Schleyer, P. *J. Comp. Chem.* **1983**, *4*, 294.

- (129) Dookeran, N.; Harrison, A. *J. Mass. Spectrom* **1996**, *30*, 666.
- (130) Dunfield, L.; Burgess, A.; Scheraga, H. *J. Phys. Chem.* **1978**, *82*, 2609.
- (131) Momany, F.; McGuire, R.; Burgess, A.; Scheraga, H. *J. Phys. Chem.* **1975**, *79*, 2361.
- (132) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246*, 64.
- (133) Stephenson, J. L., Jr.; Booth, M. M.; Boue, S. M.; Eyler, J. R.; Yost, R. A. *ACS Symposium Series* **1996**, *619*, 512.
- (134) Park, Y.; Lebrilla, C. *Mass Spectrom. Rev.* **2005**, *24*, 232.
- (135) Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B. *Anal. Chem.* **2006**, *78*, 670.
- (136) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161.
- (137) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (138) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785
- (139) Zuccarello, F.; Buemi, G. *Carbohydr. Res.* **1995**, *273*, 129.
- (140) Momany, F.; Appell, M.; Willett, J.; Schnupf, U.; Bosma, W. *Carbohydr. Res.* **2006**, *341*, 525.
- (141) Momany, F.; Appell, M.; Willett, J.; Bosma, W. *Carbohydr. Res.* **2005**, *340*, 1638
- (142) Irikura, K.; Johnson III, R.; Kacker, R. *J. Phys. Chem. A* **2005**, *109*, 8430.
- (143) Russell, D.; Johnson, I. NIST Computational Chemistry Comparison and Benchmark Database; NIST Standard Reference Database Number 101, 2006.
- (144) Scott, A.; Radom, L. *J. Phys. Chem.* **1996**.
- (145) Mulliken, R. *J. Chem. Phys* **1955**, *23*, 1833.
- (146) Momany, F. A.; Appell, M.; Strati, G.; Willett, J. L. *Carbohydr. Res.* **2004**, *339*, 553.
- (147) Silva, C. *Theo. Chim. Acta* **2006**, *116*, 137.
- (148) Vasko, P. D.; Blackwell, J.; Koenig, J. L. *Carbohydr. Res.* **1971**, *19*, 297.

- (149) Vasko, P. D.; Blackwell, J.; Koenig, J. L. *Carbohydr. Res.* **1972**, *23*, 407.
- (150) Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209.
- (151) Capman, J. *Practical Organic Mass Spectrometry: A Guide for Chemical and Biochemical Analysis*, 2nd ed.; Wiley: Chichester, New York, 1993.
- (152) Lyman, J. L.; Galbraith, H. W.; Ackerhalt, J. R. *Los Alamos Sci.* **1982**, *3*, 66.
- (153) Giordmaine, J. A.; Miller, R. C. *Phys. Rev. Lett.* **1965**, *14*, 973
- (154) Watson, C. H.; Zimmerman, J. A.; Bruce, J. E.; Eyler, J. R. **1991**, *95*, 6081.
- (155) Zimmerman, J. A.; Watson, C. H.; Eyler, J. R. *Anal. Chem.* **1991**, *63*, 361.
- (156) Schnupf, U.; Willett, J. L.; Bosma, W. B.; Momany, F. A. *Carbohydr. Res.* **2007**, *342*, 196.
- (157) Russo, T.; Martin, R.; Hay, P. *J. Phys. Chem.* **1995**, *99*, 17085.
- (158) Wadt, W.; Hay, P. *J. Chem. Phys.* **1985**, *82*, 270.
- (159) Appell, M.; Strati, G.; Willett, J. L.; Momany, F. A. *Carbohydr. Res.* **2004**, *339*, 537.
- (160) Appell, M.; Willett, J. L.; Momany, F. A. *Carbohydr. Res.* **2005**, *340*, 459.
- (161) Gerbst, A.; Grachev, A.; Shashkov, A.; Nifantiev, N. *Russ. J. Bioorg. Chem.* **2007**, *33*, 24.
- (162) Hineno, M. *Carbohydr. Res.* **1977**, *56*, 219.
- (163) Silverstein, R.; Bassler, G.; Morrill, T. *Spectrometric Identification of Organic Compounds*; Wiley: New York, 1991.
- (164) Boese, A.; Martin, J.; Handy, N. *J. Chem. Phys.* **2003**, *119*, 3005.
- (165) Manuel Alcamí, O. M., Manuel Yáñez, . *Mass Spectrom. Rev.* **2001**, *20*, 195.
- (166) Zheng, Y.; Ornstein, R.; Leary, J. *J. Mol. Struct.* **1997**, *389*, 233.

BIOGRAPHICAL SKETCH

Born in Tijuana, Baja California, Mexico, to Roberto G. Contreras and Maria Elia Contreras Sanchez, Cesar S. Contreras spent the first five years of childhood living in Tijuana with his siblings: Roberto, Martha, Diana, Maria, Elizabeth, Sandra and Daniel. Cesar and his family migrated to United States and settled in San Diego, California, which he called home for the next 23 years. Cesar graduated from Clairemont High with honors in Spanish language and literature and political science. After attending San Diego Mesa Community College, and earning an associate degree in liberal arts and sciences, Cesar briefly attended CSU San Marcos, before enrolling at San Diego State University (SDSU). At SDSU he conducted research with Dr. Andrew L. Cooksy, and was chosen as a MARC scholar, two events that would help fulfill his dreams of attending graduate school in Chemistry. He obtained his Bachelor of Science from SDSU in 2002, and after a year off, decided to continue his studies at the University of Florida. Throughout his educational career, Cesar held several jobs that contributed to his growth as a person and a scientist. Cesar was employed as a chemical research technician for Schumacher Industries, a division of Air Products and Chemical, Inc. Cesar was also employed at the San Diego Research Laboratory, before heading to graduate school at UF.