

Mass Spectrometry of Carbohydrates by Experimental and Theoretical Methods

A dissertation presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Doctor of Philosophy

Jordan M. Rabus

August 2021

© 2021 Jordan M. Rabus. All Rights Reserved.

This dissertation titled
Mass Spectrometry of Carbohydrates by Experimental and Theoretical Methods

by

JORDAN M. RABUS

has been approved for
the Department of Chemistry and Biochemistry
and the College of Arts and Sciences by

Benjamin J. Bythell

Associate Professor of Chemistry

Florenz Plassmann

Dean, College of Arts and Sciences

Abstract

RABUS, JORDAN M., Ph.D., August 2021, Analytical Chemistry

Mass Spectrometry of Carbohydrates by Experimental and Theoretical Methods

Director of Dissertation: Benjamin J. Bythell

This dissertation is focused on the study of the fragmentation chemistry of carbohydrates by mass spectrometry and computational methods. By investigating the behavior of model systems, we hypothesize that a general model for carbohydrate fragmentation can be found. The model systems studied are sodiated cellobiose and gentiobiose (Chapter 2), deprotonated lactose (Chapter 3), and β -cyclodextrin (Chapter 4), each representing a different aspect of carbohydrate fragmentation. Experimental techniques, including tandem mass spectrometry, stable isotopic labeling, and infrared multiphoton dissociation were employed. Throughout, we also utilize computational modeling to add a dimension of clarity to the experimental results and provide values usable in subsequent predictive efforts.

Acknowledgments

I have not reached this finish line by working alone. I would like to thank those who have helped me with my goals over the last few years. First, Dr. Benjamin Bythell, for supporting and encouraging me from the very beginning. I would like to thank my committee members for their advice and attention, Dr. Jixin Chen, Dr. Peter Harrington, Dr. Martin Kordesch, as well as Dr. Eike Bauer, Dr. James Bashkin, and Dr. Janet Braddock-Wilking. I would like to thank my collaborators, Dr. Eduardo Carrascosa, Dr. Robert Pellegrinelli, Ali Hassan Abi Khodr, and Dr. Thomas Rizzo. Special thanks to Dr. Philippe Maître and Dr. Estelle Loire for hosting and assisting us while visiting the CLIO facility in Paris, France. Very special thanks to my group members, Dr. Maha Abutokaikah, Dr. Shanshan Guan, Joy Malaba, Ola Bataineh, Jaya Paudel, Ester Mbuna, Dan Simmons, and Howard Mayes: Thank you for all the support and advice you have given me.

Table of Contents

	Page
Abstract.....	3
Acknowledgments	4
List of Tables.....	8
List of Figures	9
List of Schemes	11
Chapter 1 : Introduction.....	12
1.1 Research Goals.....	12
1.2 Mass Spectrometry.....	13
1.2.1 Tandem Mass Spectrometry	14
1.2.2 Quadrupole Mass Filter	15
1.2.3 Time-of-Flight.....	16
1.2.4 Ion Funnel	17
1.2.5 Reflectron.....	18
1.3 Carbohydrates	19
1.3.1 What is a Carbohydrate?.....	19
1.3.2 Carbohydrate Nomenclature	21
1.3.3 Isotopic Labeling.....	23
1.4 Computational Chemistry.....	23
1.4.1 Basics and Primer.....	23
1.4.2 Goals of Calculation	24
1.4.3 Calculation Tools	26
1.4.4 3D Visual Representation	30
1.4.5 Thermochemical Data.....	31
1.4.6 Ion Affinities	32
1.4.7 Rice-Ramsperger-Kassel-Markus (RRKM) Theory	32
1.4.8 Custom Utilities.....	32
1.5 Infrared “Action” Spectroscopy.....	34
1.6 Ion Mobility	36
1.6.1 Structures for Lossless Ion Manipulation (SLIM)	37
1.7 References.....	38

Chapter 2 : Sodium-Cationized Carbohydrate Gas-Phase Fragmentation Chemistry: Influence of Glycosidic Linkage Position		49
2.1	Abstract.....	49
2.2	Introduction.....	50
2.3	Experimental Methods	54
2.4	Theoretical Methods.....	55
2.5	Results and Discussion	56
2.5.1	Experimental Findings.....	56
2.5.2	Energetics of Sodiated Minima.....	60
2.5.3	Glycosidic Bond Cleavage: The B ₁ -Y ₁ Pathways	61
2.5.4	Cross Ring Bond Cleavage Transition Structures: The A _n -X _m Pathways	66
2.5.5	Larger Basis Sets and Alternate Model Chemistries	72
2.6	Conclusions.....	73
2.7	Acknowledgments.....	73
2.8	References.....	74
2.9	Supporting Information	82
Chapter 3 : Deprotonated Carbohydrate Anion Fragmentation Chemistry: Structural Evidence from Tandem Mass Spectrometry, Infra-Red Spectroscopy, and Theory.....		103
3.1	Abstract.....	103
3.2	Introduction.....	104
3.3	Experimental Methods	108
3.4	Theoretical Methods.....	110
3.5	Results and Discussion	112
3.5.1	Tandem Mass Spectrometry of [Lactose-H] ⁻	112
3.5.2	Low Energy Structures of [Lactose-H] ⁻	114
3.5.3	Glycosidic Bond Cleavage Reaction Energetics	115
3.5.4	Z ₁ Structures.....	121
3.5.5	C ₁ Structures	126
3.5.6	Cross-ring Fragmentation Pathways: ^{0,2} A ₂ Anion Formation.....	130
3.6	Conclusions.....	132
3.7	Acknowledgments.....	132
3.8	References.....	133
3.9	Supporting Information	148

Chapter 4 : Unravelling the Structures of Sodiated β -Cyclodextrin and its Fragments ..	162
4.1 Abstract.....	162
4.2 Introduction.....	163
4.3 Methods	167
4.3.1 Experimental Details	167
4.3.2 Computational Approach.....	169
4.4 Results and Discussion	171
4.4.1 CID mass spectrum.....	171
4.4.2 Ion mobility and cryogenic IR action spectra of sodiated α -CD and β -CD 172	
4.4.3 Cryogenic IR Spectra of CID Fragments.....	178
4.4.4 Comparison with Theory	181
4.5 Conclusions.....	189
4.6 Author Contributions.....	191
4.7 Conflicts of Interest.....	191
4.8 Acknowledgements	191
4.9 References.....	191
4.10 Supporting Information	199
Chapter 5 : Conclusion	209

List of Tables

	Page
Table 2-1: Summary of D ₈ -hexose sodiated disaccharide MS/MS spectra. Analytes were generated by exchanging all hydroxyl protons for deuterons	59
Table 2-2: Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- β -D-Glucose) and Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- β -D-Glucose) forms calculated at the M06-2X/6-31+G(d,p) level of theory	65
Table 3-1: Relative energies of the transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- α -D-Glucose anomer) calculated at the M06-2X/6-31++G(d,p) level of theory	120

List of Figures

	Page
Figure 1-1: Example Schematic Diagram of a Time-of-Flight Mass Spectrometer	14
Figure 1-2: Schematic representation of a quadruple mass filter.....	16
Figure 1-3: Polysaccharide components, bonding, stereochemistry and nomenclature of Lactose (β -D-galactopyranosyl-(1-4)-D-Glucose)	19
Figure 1-4: The carbohydrate mass spectrometry fragmentation nomenclature of Domon and Costello illustrated for Lactose	21
Figure 1-5: Example 2D potential energy surface (PES)	25
Figure 1-6: Example of the simplification of a multibody problem to an electron density problem in DFT.....	28
Figure 1-7: Schematic diagram of the calculation filtering process.....	30
Figure 1-8: Example ball and stick model of glucose	31
Figure 1-9: Representative example of a partial SLIM device.	38
Figure 2-1: The carbohydrate fragmentation nomenclature of Domon and Costello	51
Figure 2-2: Example MS/MS spectra ($E_{\text{collisions, lab}} = 38 \text{ eV}$) illustrating the differing degree and nature of fragmentation of the isomeric sodium cationized analytes Cellobiose and Gentiobiose	57
Figure 2-3: Selected minima of the isomeric sodium cationized analytes (a) Cellobiose and (b) Gentiobiose. The β -D-Glucose forms are shown.....	61
Figure 2-4: RRKM unimolecular rate constants, $k \text{ (s}^{-1}\text{)}$ calculated at various internal energies for the lowest energy fragmentation pathways of $[\text{Cellobiose}+\text{Na}]^+$; β -D-Glucose form	71
Figure 3-1: Example Bruker MaXis plus MS/MS spectra of regular and regioselectively labeled, deprotonated lactose ($E_{\text{laboratory}} = 6 \text{ eV}$ in all cases).....	113
Figure 3-2: The lowest-energy C_1 - Z_1 transition structure of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory (a deprotonated 1,4-anhydroglucose-forming TS).....	119
Figure 3-3: Z_1 anion ($m/z 167.066$, $^{13}\text{C}_6\text{H}_9\text{O}_5^-$) infrared action spectroscopy spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}\text{C}_6$ -glucose anions ($m/z 347.129$, $^{13}\text{C}_6\text{C}_6\text{H}_{21}\text{O}_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures.	123
Figure 3-4: C_1 anion ($m/z 179.059$, $\text{C}_6\text{H}_{11}\text{O}_6^-$) infrared action spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}\text{C}_6$ -glucose anions ($m/z 347.129$, $^{13}\text{C}_6\text{C}_6\text{H}_{21}\text{O}_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures.....	129

Figure 4-1: Schematic representation of β -cyclodextrin, with its conical structure represented on the right. The carbon atoms connected to the hydroxyls are labeled C2, C3, and C6 for one glucose unit.....	164
Figure 4-2: Schematic representation of the experimental setup to obtain IR spectra of mass selected CID fragments under cryogenic conditions.	167
Figure 4-3: CID-MS of sodiated β -CD and its fragments, carried out with a potential difference of 100 V and 240 V in our tandem mass spectrometer.	171
Figure 4-4: Arrival times of $[\alpha\text{-CD}+\text{Na}]^+$ and $[\beta\text{-CD}+\text{Na}]^+$ and their corresponding cryogenic IR spectra.	174
Figure 4-5: Cryogenic IR spectra of β -CD complexed with group I cations. The two marked areas denote the transitions associated to the C2 and C3 OH oscillators.....	177
Figure 4-6: Cryogenic IR spectra of the main CID fragments of $[\beta\text{-CD}+\text{Na}]^+$ and the parent ion	179
Figure 4-7: Comparison between the experimental (top) and simulated IR spectra for the m/z 347 fragment.....	183
Figure 4-8: Comparison between the experimental (top) and simulated IR spectra for the m/z 509 fragment.....	186
Figure 4-9: Comparison between the experimental (top) and simulated IR spectra for the m/z 671 fragment.....	189

List of Schemes

	Page
Scheme 2-1: The lowest energy B ₁ -Y ₁ pathway for [Cellobiose+Na] ⁺	63
Scheme 2-2: Ring-opening, then ^{0,2} A ₂ ion formation from [Gentiobiose+Na] ⁺	67
Scheme 2-3: Ring-opening, 2-1 H shift, then ^{0,3} A ₂ ion formation from [Gentiobiose+Na] ⁺	68
Scheme 2-4: Consecutive fragmentation of the ^{0,2} A ₂ ion to form ^{0,4} A ₂ ion formation from [Gentiobiose+Na] ⁺	69
Scheme 2-5: Consecutive fragmentation of the ^{0,2} A ₂ ion to form ^{2,4} A ₂ ion and glycolaldehyde for [Cellobiose+Na] ⁺	72
Scheme 3-1: Carbohydrate numbering and fragmentation nomenclature of Domon and Costello illustrated for lactose (β-D-galactopyranosyl-(1→4)-D-glucose)	106
Scheme 3-2: The lowest-energy glycosidic bond cleavage pathways of deprotonated lactose. 1,4-anhydroglucose-forming pathway initiated from β-D-galactopyranosyl- (1→4)-β-D-glucose and 3,4-anhydroglucose-forming pathway achievable from both anomeric forms.....	117
Scheme 3-3: Ring-opening isomerization of deprotonated 3,4-anhydroglucose Z ₁ anion structures	124
Scheme 3-4: The lowest-energy ^{0,2} A ₂ anion formation pathway of deprotonated lactose	131
Scheme 4-1: Putative ion structure types potentially formed from [β-CD+Na] ⁺ , illustrated for the 2-residue, m/z 347 species: 2-ketone B ₂ ; 1,6-anhydro B ₂ ; 1,2-anhydro B ₂ ; and Z ₂ ion.....	181
Scheme 4-2: Proposed fragmentation mechanism for the formation of the m/z 347 fragment.....	185

Chapter 1 : Introduction

1.1 Research Goals

My research has been the study of carbohydrates by mass spectrometry and computational methods. Using model systems, we aim to build an understanding of the key structures and processes involved in gas-phase carbohydrate fragmentation. To facilitate this goal, a variety of experimental methods will be used, including tandem mass spectrometry, isotopic labeling, and gas-phase infrared spectroscopy.

Computational methods include primarily density functional theory (DFT) calculations of minima and transition states. This also includes the selected use of molecular dynamics calculations as well as the development of a suite of utilities to assist in performing and managing calculations, as well as analysis of the results.

To finish these research goals:

Chapter 2 looks at the fragmentation mechanisms of sodiated β -1,4- and β -1,6-glucose disaccharides cellobiose and gentiobiose. We use isotopic labeling (^{13}C and ^2D) to aid the differentiation of otherwise ambiguous fragments. Although the two carbohydrates only differ in the position of the linkage between them, cellobiose and gentiobiose show quite diverse fragmentation behavior.

Chapter 3 reports on the dissociation chemistry and IR spectroscopy of the deprotonated lactose disaccharide anion. The energetics of glycosidic and cross-ring cleavages are considered computationally, and the potential structure of the fragment ions is probed experimentally using gas-phase IR action spectroscopy.

Chapter 4 looks at the IR spectroscopy of the cyclic carbohydrate cyclodextrin. (poly- α -1,4-glucose) Cyclodextrin is capable of selectively/preferentially binding metal ions similarly to crown ethers and is much larger than the disaccharide systems explored in the preceding chapters. We use a combination of experimental and computational methods to probe the gas-phase structure of the fragment ions and make inferences on the structure of the precursor.

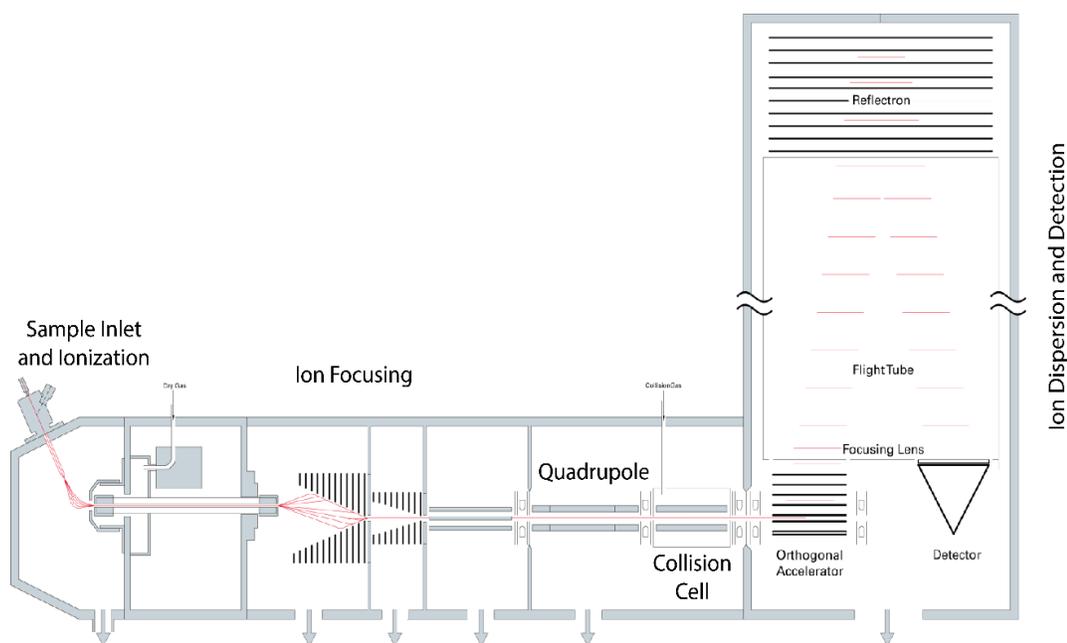
1.2 Mass Spectrometry

Mass spectrometry is an analytical technique allowing for quantitative and qualitative data to be gathered for target molecules. While many instrument designs have been developed and are in use, all mass spectrometers operate on the same fundamental level: (1) gas-phase ions are created from the sample, (2) ions are separated based on the mass-to-charge ratio (m/z) and (3) the ions are detected providing a measure of relative abundance and m/z .

The type of ionization used for the research presented within is electrospray ionization (ESI). In this method, a solution containing the desired analyte is sprayed out of a charged needle towards an oppositely charged inlet plate. When the sample droplets leave the needle, the charge is retained and conserved as the solvent evaporates, ultimately leaving unsolvated, charged analyte ions to be drawn into the mass spectrometer for analysis.

Figure 1-1:

Example Schematic Diagram of a Time-of-Flight Mass Spectrometer



1.2.1 Tandem Mass Spectrometry

A single stage of mass spectrometry is typically insufficient to fully characterize most compounds effectively. We use instead multiple stages of mass spectrometry to not only separate and characterize analyte ions but their fragment ions as well. This tandem mass spectrometry is known as MS/MS or MSⁿ where *n* is the number of tandem stages involved. By building a profile of the fragments produced by a precursor ion, a fuller picture of that precursor can be elucidated. For the work presented below, the primary method by which MS/MS was performed was using collision-induced dissociation (CID). In CID, the analyte ion is subjected to one or more collisions with an inert gas atom or molecule (typically helium, nitrogen, or argon), resulting in energy transfer. The energy

of these collisions can be controlled *via* the use of variable electric fields in the instrument that control the amount of energy that is imparted onto the sample ions as they enter the collision cell, by controlling the relative velocity of the collisions. This energy transfer can be represented symbolically by the center-of-mass collision energy, E_{CM} [1, 2],

$$E_{CM} = E_{Lab} \frac{m_N}{m_i + m_N}$$

where m_N is the mass of the neutral, m_i is the mass of the ion, and E_{Lab} is the ion kinetic energy in the laboratory frame of reference. E_{Lab} represents the amount of energy imparted onto the ion by the accelerating stage that defines the kinetic energy of the ion. If the collision is sufficiently energetic and/or enough collisions occur, enough energy is transferred from the ion/neutral collision to result in the breaking of bonds in the analyte ion. By selection of ions to subject to CID, these fragments can be isolated and detected separately and thus mapped back onto a particular precursor ion.

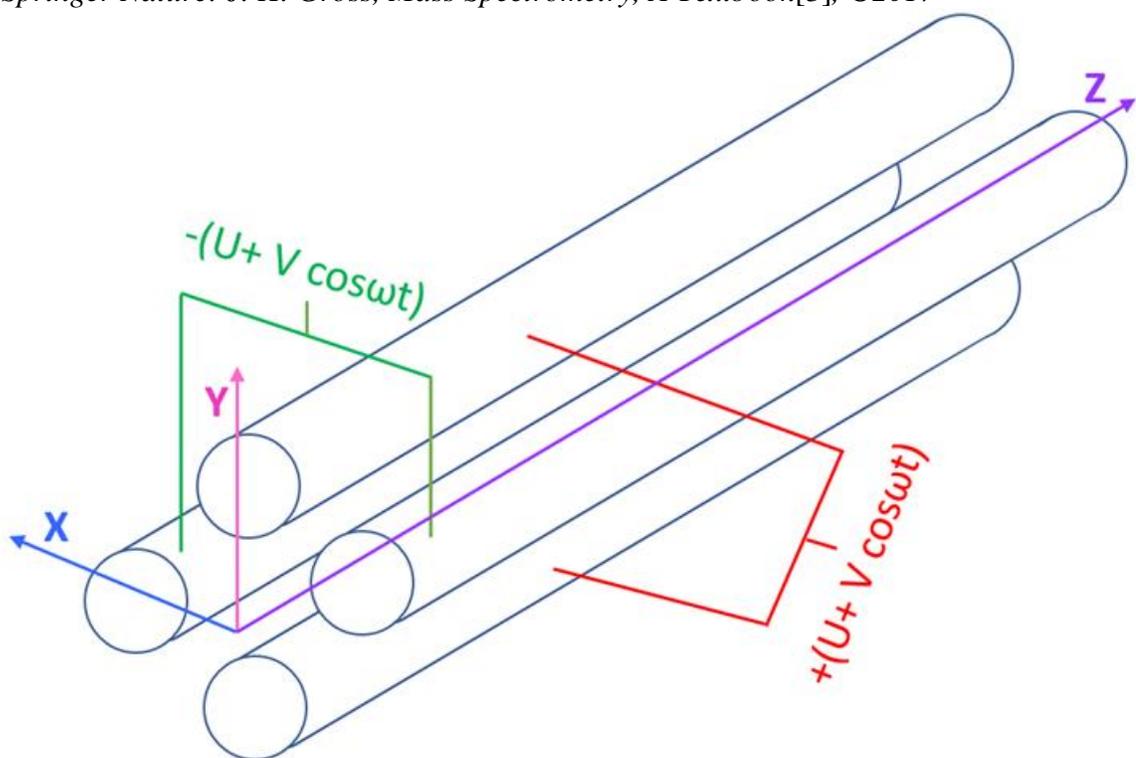
1.2.2 Quadrupole Mass Filter

To accomplish effective MS², accurate selection of a precursor ion is needed prior to fragmentation. In practice, this is accomplished using a quadrupole mass filter. A quadrupole mass filter is a device composed of two pairs of parallel conductive rods, 2 in the XZ plane and two in the YZ plane (Figure 1-2). Each pair rod has a direct current (U) and radio frequency ($V \cos(\omega t)$) component applied to it, with apposing rods being applied opposing charge. Ions travelling through the quadrupole will experience an attractive force corresponding to the appropriate rod(s). As the ions move through the quadrupole, the applied potentials are switched such that an ion does not collide with a

rod but is continually steered back to the center of the quadrupole. By careful selection of the direct current potential and applied frequency, an ion or ions of particular m/z can be preferentially transmitted while other ions are prevented from traversing the quadrupole.

Figure 1-2:

Schematic representation of a quadrupole mass filter. Reprinted by permission from Springer Nature: J. H. Gross, Mass Spectrometry, A Textbook[3], ©2017



1.2.3 Time-of-Flight

The primary type of instrument used in my research was a Time-of-Flight (ToF) mass spectrometer[4, 5]. The principle on which this instrument type operates is as suggested in the name, the physical quantity being measured is the amount of time that an ion takes to traverse a fixed distance. If given the same initial kinetic energy, a heavier

ion will travel more slowly than a lighter ion and therefore take a longer amount of time to cover the same distance. This difference in time, although slight, is sufficient to be distinguished and detected. Symbolically, this can be expressed as follows: Ions are given a known potential energy $E_p = ezU$ where e is the fundamental electric charge, z is the number of charges and U is the magnitude of an applied external electric potential.

Potential energy is converted to kinetic energy in the flight tube, $ezU = \frac{1}{2}mv^2$. Velocity

is length over time, $v = \frac{L}{t}$, so $ezU = \frac{1}{2}m \left(\frac{L}{t}\right)^2$. Rearranging and solving for t yields $t =$

$\frac{L}{\sqrt{2eU}} \sqrt{\frac{m}{z}}$, thus, the time of flight for the ion is directly proportional to the square root of

the mass to charge ratio.

1.2.4 Ion Funnel

In the ion focusing section of the instrument (Figure 1-1) there are a pair of offset dual ion funnels[6]. These ion funnels are designed to radially collect and confine incoming ions after being introduced by ESI. The ion funnels have two primary functions. First, to collect and focus the incoming ions from a diffuse trajectory into a coherent ion beam; and second, to reject any stray neutral molecules which may have been inadvertently introduced. To perform this second function, the two ion funnels are mounted with their axes offset slightly such that ions must follow a curved path to reach the interior of the instrument while any neutrals will continue undeflected and be eventually removed from the system by the vacuum pump.

1.2.5 Reflectron

The flight tube of the instrument depicted in Figure 1-1 features a device commonly known as a reflectron or ion mirror[7–9]. This device is mounted at the top of the flight tube (upper right of Figure 1-1) and is electrostatic in nature. By operating at a slightly higher electrical potential than the kinetic energy of the ions, the reflectron can accept ions, bring them to a halt, and reverse their direction. Ions enter the flight tube from the bottom travelling upwards, enter the reflectron and have their trajectory reversed to travel back down the flight tube to the detector. This serves two purposes: First, the effective flight path of the ions is increased by a factor of ~ 2 , allowing for either increased resolution or a more physically compact instrument depending on design priorities. Second, the reflectron functions to rectify and focus the kinetic energy variations that individual ions of the same m/z may possess. If an individual ion has a higher kinetic energy when entering the flight tube, it will both arrive more quickly and penetrate more deeply into the reflectron, i.e., take a longer flight path before being reoriented towards the detector. Conversely, an individual ion of lower kinetic energy will not penetrate as deep into the reflectron and be reoriented more quickly. The net effect is that although the individual ions of the same m/z may enter the reflectron at slightly different times, they should also remain resident for disparate times and leave the reflectron as a coherent packet to be detected.

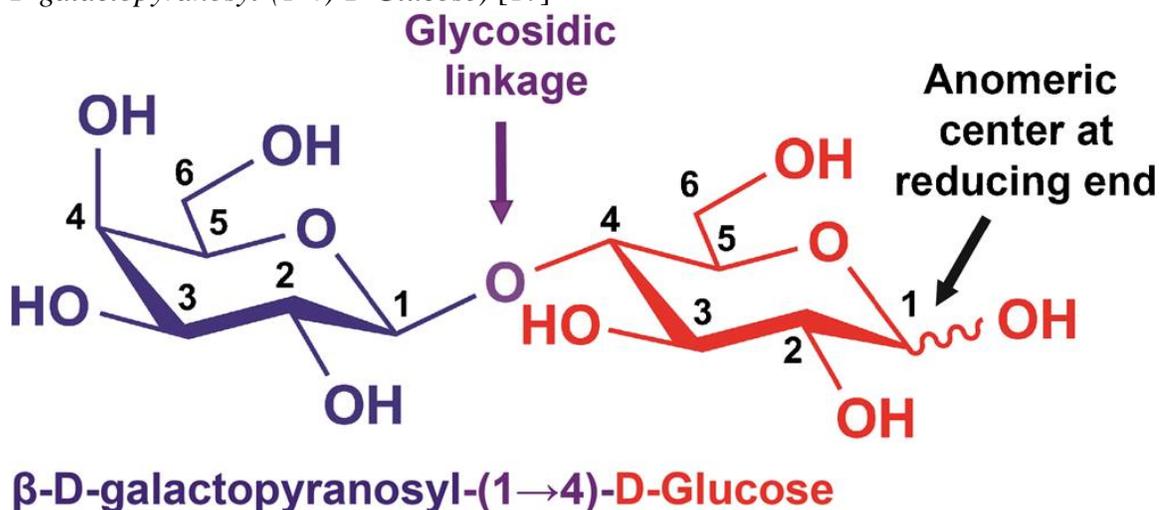
1.3 Carbohydrates

1.3.1 What is a Carbohydrate?

Complex carbohydrates (glycans) are associated with a vast number of major biological functions in humans and other organisms, including but not limited to: joint lubrication, cell growth, and inflammatory and immune responses. [10] Glycans are biomarkers for many cellular processes, as signals of the proper function as well as dysfunction. Changes in level, type, location, and/or structure of glycans inside of a cell, on the cell surface, or within body fluids have been associated with numerous disease processes, with cancers, pneumonia, malaria, diabetes, HIV/AIDS, and hepatitis all being examples of glycan mediated processes. [11–16]. The glycans of interest are also often a low abundance component of a very complex mixture, making their confident identification difficult.

Figure 1-3:

Polysaccharide components, bonding, stereochemistry and nomenclature of Lactose (β -D-galactopyranosyl-(1-4)-D-Glucose) [17]



To understand these processes, a sensitive method to quickly and accurately identify glycan sequence and structure is required. Tandem mass spectrometry is often used for these analyses as it provides an accurate and sensitive measure of the elemental composition of glycans based on their mass-to-charge (m/z) ratios, but certain challenges exist: Glycans are biopolymers composed of monomers which are structural isomers connected by multiple potential linkages (glycosidic bonds), in contrast to other biopolymers such as peptides, in which the linkages between monomer units are constant but the molecular formula of the monomers differs. This isomeric nature of the glycan monomers makes the confident assignment of the structure by mass spectrometry challenging. (Figure 1-3)

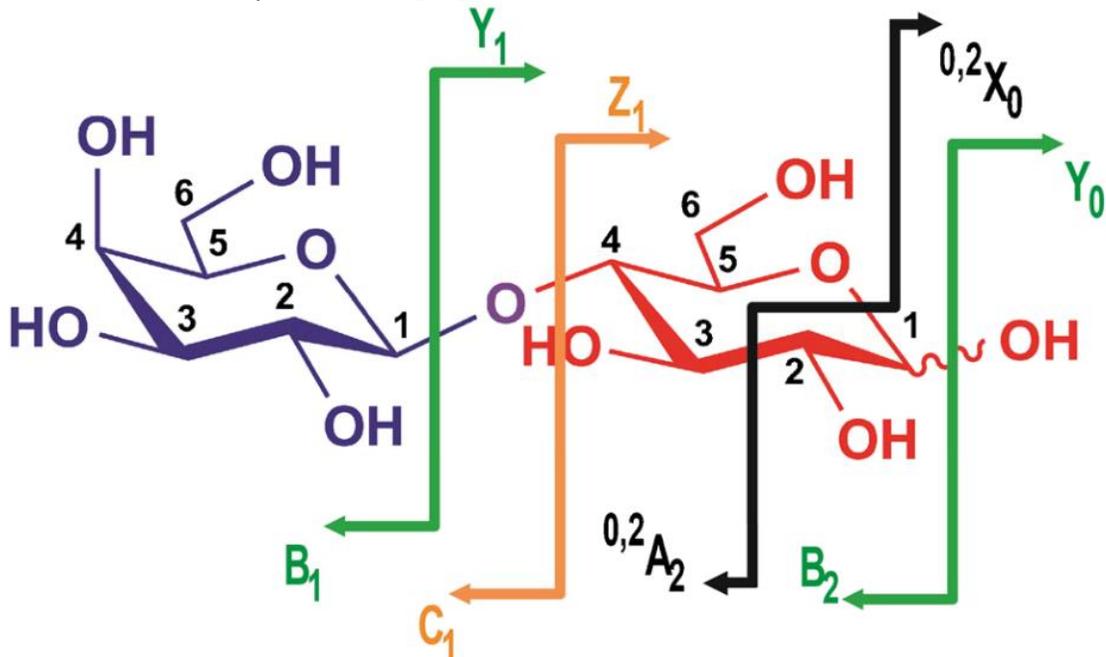
Glycan fragmentation by mass spectrometry results in a characteristic fragment ion pattern [17, 18] which could allow a library of standards to be built and compared against using the fragmentation patterns like fingerprints, but the huge number of isomeric glycans [19], as well as the difficulty in the synthesis of these standards [13], makes this method currently impractical.

Using theoretical methods to study glycan fragmentation *in silico* is an attractive option, but the tools to do so are often not commonly available, extremely resource-intensive, require special expertise to perform the calculations and interpret results, or any/all of the above. We propose the development of a framework by which the computational workflow can be simplified such that practical and useful calculations can be used to help explain complex glycan fragmentation and facilitate effective structural elucidation.

1.3.2 Carbohydrate Nomenclature

Figure 1-4:

*The carbohydrate mass spectrometry fragmentation nomenclature of Domon and Costello illustrated for **Lactose** [20]*



A prerequisite for any mass spectrometric study of glycans is ionization of the analyte: Glycans can be studied as either cations, in which case a proton or a metal cation adduct is present to provide the charge (e.g., $[M+H]^+$, $[M+Li]^+$, $[M+Na]^+$, $[M+Ca]^{+2}$, etc.) [17, 21–26], or as anions, in which the glycan is deprotonated or an anionic adduct is present (e.g., $[M-H]^-$, $[M+Cl]^-$, $[M+NO_3]^-$, etc.) [27–31]. The nomenclature used to describe carbohydrate fragmentation is depicted schematically in Figure 1-4: For a glycan of length $N = n + m$, the glycosidic bond cleavages result in a series of B_n and/or C_n ions if the non-reducing end fragment (left-hand side) retains the charge or Y_m and/or Z_m ions

if the reducing end (right-hand side) fragment keeps the charge. Any cross-ring cleavage results in A_n or X_m ions. These cross-ring cleavages add useful structural information that cannot be determined from glycosidic cleavages alone. Cross-ring cleavages require the breaking of two bonds, which can allow for a more robust structural interpretation.

The type of fragment ion can be determined in many cases using tandem mass spectrometry, but the identification of a particular sequence ion alone is typically insufficient to determine what the position and type of glycosidic linkages are present in a given glycan. For example, the sodiated disaccharides lactose, maltose, and sucrose can all fragment to yield B_1 peaks at m/z 185 and Y_1 peaks at m/z 203.

While the mass-to-charge ratio of these fragment ions may be identical, other distinguishing properties are observed: The presence/absence of given fragments and the relative abundances of the observed fragment ions. These differences are what allows for the identification of a given glycan over another. One approach to identifying carbohydrates by mass spectrometry would be to build a spectral library of possible precursors and their fragments, but such an endeavor is daunting at best due to the sheer variety of potential carbohydrates possible. While custom laboratory synthesis of a given carbohydrate is possible, the skills and facilities needed are not commonly available and/or cost prohibitive. An alternative to this library approach is leveraging computational chemistry to simulate and evaluate novel systems without having to perform costly synthesis, to analyze patterns in fragmentation and make inferences therein to predict and verify the results of an unknown system.

1.3.3 Isotopic Labeling

An additional dimension can be added to the analysis through the use of stable isotopic labeling of carbohydrates. This process involves the substitution of one or more specific atoms in a molecule with a stable isotope. Most commonly this is either swapping $^1\text{H} \rightarrow ^2\text{D}$, $^{12}\text{C} \rightarrow ^{13}\text{C}$ or $^{16}\text{O} \rightarrow ^{18}\text{O}$. For the hydrogen-deuterium exchange, this is often abbreviated to HDX and refers to the swapping of -OH and -NH functional groups to -OD and -ND functional groups. This is easy to perform in the laboratory setting as it can be accomplished by diluting the samples in deuterated solvents. The other forms of stable isotope labeling require the isotope to be incorporated during the synthetic process. Examples of this synthetic isotopic labeling would be enriching a single position in a carbohydrate backbone with all ^{13}C at that position, or a specific $-\text{CH} \rightarrow -\text{CD}$ exchange. The utility of this transformation is found in the mass spectrometer, particularly for mass symmetric carbohydrates which would fragment into otherwise isomeric pieces; without isotopic labeling, it would be impossible to differentiate a C/Y or B/Z ion pair (Figure 1-3). If one half of the carbohydrate would be isotopically labeled, however, the fragments would no longer be isobaric and can be easily identified.

1.4 Computational Chemistry

1.4.1 Basics and Primer

Computational chemistry allows for the estimation of chemical properties for an arbitrary compound without having to perform an experimental measurement. This ability is a useful resource when the experimental conditions are difficult to isolate or replicate, as they are in the mass spectrometer. By use of computational methods in

conjunction with experiments, we can independently evaluate our hypotheses and obtain a more confident result than either method allows for alone.

The fundamental equation in computational chemistry is the Schrödinger equation which takes on the basic form of:

$$\hat{H}\Psi = E\Psi$$

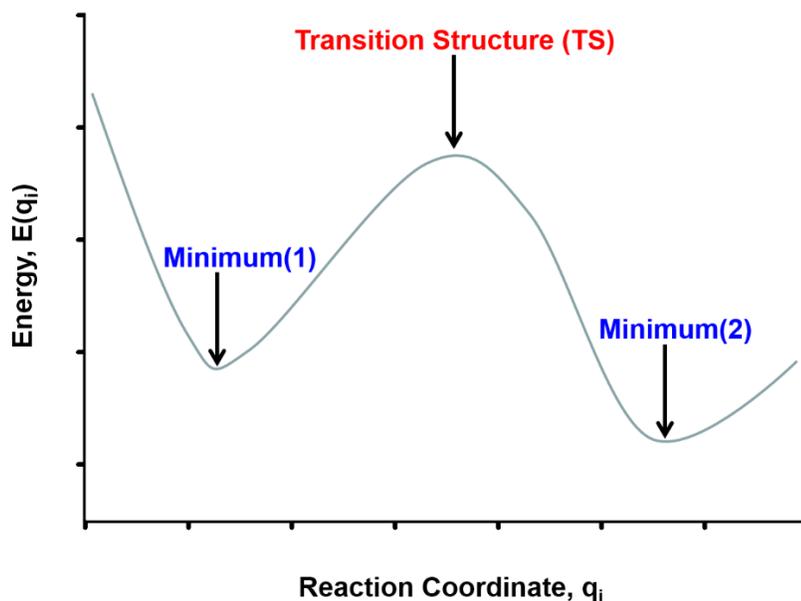
Where \hat{H} is the Hamiltonian operator, corresponding to the total energy of a system, Ψ is the wavefunction for a system, and E is the energy of the system. Although this equation is simple in form, the challenging aspect in practical terms is that the wavefunction for a system is both typically unknown and has no closed analytical form. It must be estimated for each system to be calculated. How this is performed is one of the fundamental differences between the different methods of theoretical calculation.

1.4.2 Goals of Calculation

The goals of the calculations are to calculate the chemical properties of a system. In practical terms, this typically begins with calculating the energy of a molecule/ion. The possible range of energies that a molecule can have form the potential energy surface (Figure 1-5) for the system.

Figure 1-5:

Example 2D potential energy surface (PES)



This surface has many topological features, but it is fundamentally made up of regions of higher and lower energy. Valleys represent stable, local minima on the potential energy surface, while peaks represent transition states linking adjacent minima. This varied energy landscape with multiple minima requires a method by which the assorted minima can be adequately sampled and determined. Specialized tools allow for an automated approach to this sampling. Once an energy minimum has been located, various physiochemical properties can be calculated, including but not limited to fragmentation patterns, rotational constants, vibrational frequencies, NMR shifts, etc.

1.4.3 Calculation Tools

The majority of the computational work presented in this thesis was performed using the Gaussian 09/16 suite of programs[32, 33]. A set of custom utilities was also written in the Python programming language (detailed below). Other software used was NAMD[34] and Fafoom[35, 36] for generating starting structures of carbohydrates, as well as the XTB[37] engine and the CREST[38] conformer generation tool.

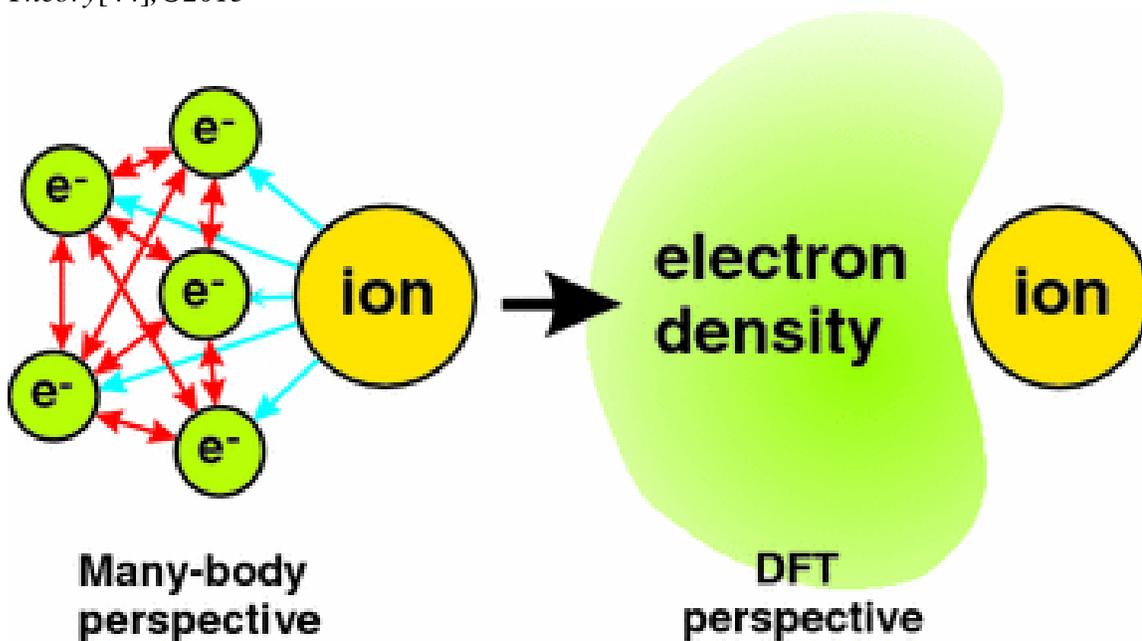
NAMD and Fafoom operate on a similar level of theory, using an empirical force field to optimize structures, but with many different philosophies of operation. NAMD is used to perform temperature replica exchange molecular dynamics on a structure. In basic terms, from a given template molecule, a series of calculations are performed at different simulated temperatures. These differing temperatures allow for the molecule to rotate, vibrate, and generally move enabling the adoption of a variety of conformations. This variety of structures is a function of the energy available, and the structural pool provided. At periodic points in the calculation, certain optimized structures are swapped from one temperature to another to broaden the structural pool and enable additional structures to be located. This gives rise to an alternate name of this technique, simulated annealing. By repeatedly heating and cooling the molecule, a variety of conformations can be accessed. This process is run for a large, fixed, number of cycles. This number is typically set by the user to a high value, trading redundancy for complete coverage. For our calculations typically 1600 results are generated, with varying amounts of degeneracy as a function of the molecule, i.e., for particularly large, flexible molecules a greater number of simulations is likely to be necessary.

Fafoom works similarly in that a template molecule is used to generate conformers, but the method of generation is much more systemic than simulated annealing. Fafoom uses a simple genetic algorithm to take certain topological features from one molecule (specific bond, dihedral angles) and selectively transfer them to other candidates, reoptimizing the structures after each transfer. This process proceeds until further swapping of features fails to produce any new results. Fafoom has a facility to detect and manipulate the six-membered ring conformations as a degree of freedom. As there is a finite number of potential ring conformers (38 distinct conformations of a six-membered ring [39–42]), this method is particularly useful for many carbohydrates of interest. The advantage of this systemic approach over the simulated annealing process is an exhaustive search of the possible conformational space, at the tradeoff of increased time to search and the generation of unlikely/high energy structures.

Gaussian is a much more general-purpose computational chemistry package. In our research, we use it primarily to optimize the geometry of molecules/ions as well as compute their vibrational frequencies. The theory used to calculate these properties can vary, but we use density functional theory (DFT)[43, 44] almost exclusively due to its combination of relatively high accuracy and modest computational requirements. Density functional theory is a method to obtain an approximate wavefunction. In contrast to methods that use atomic or molecular wavefunctions, DFT uses the total electron density of a system to estimate a wavefunction, reducing a multidimensional calculation (with $3 \times N$ variables where N is the number of electrons in the system) to one which only is a function of the three spatial variables. (Figure 1-6)

Figure 1-6:

Example of the simplification of a many-body problem to an electron density problem in DFT. Reprinted by permission from Springer Nature: F. Bechstedt, Density Functional Theory[44], ©2015



For practical calculation, a means of specifying the electronic parameters of a system is required, known as the basis set. The basis set is a mathematical description of the electron distribution surrounding a nucleus to form orbitals. In practical terms, these basis sets take on the form of combinations of Gaussian type functions which work in combination to mimic the true Slater function. By increasing the number of Gaussian functions in the basis set, we can increase the goodness of fit to the true function at the cost of computational time.

The combination of DFT with a modest basis set allows for the practical replication of a much more demanding calculation, (both in terms of basic complexity as

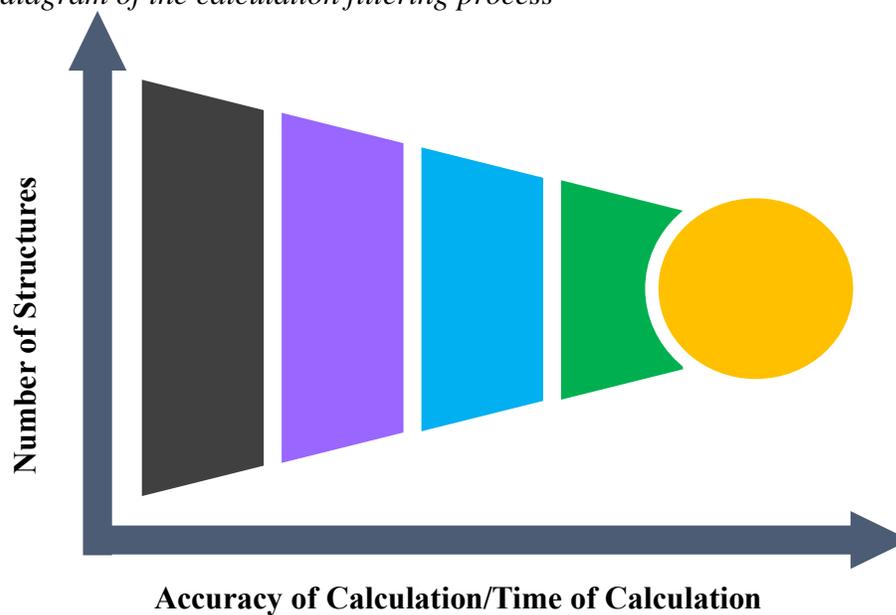
well as requiring a much larger basis set) with a fraction of the time required for the more complex calculation.

The XTB program uses a semi-empirical engine to achieve many of the same functions as Gaussian in a fraction of the time, at a cost to accuracy. The semi-empirical formality replaces computationally intensive steps in the calculation with approximations that are much quicker to calculate. Nevertheless, XTB is useful when faced with a large number of calculations as an initial screening tool. Although the result obtained is not as accurate as that provided by DFT, the qualitative ability to eliminate blatantly unrealistic structures is still very useful. The CREST tool uses XTB as a back-end engine to rapidly generate ensembles of conformers from a given template. It also has a function to generate and rank tautomers from a template, which can be very useful when the site of protonation/deprotonation of a molecule is variable.

Throughout our research, the driving goal for any calculation is practicality; a ‘perfect’ calculation is of no use if it takes infinite time to complete. We try to strike a balance between accuracy and time. This philosophy is illustrated in Figure 1-7 below, with the number of structures being considered decreasing as the complexity (and time) of the calculation increases.

Figure 1-7:

Schematic diagram of the calculation filtering process

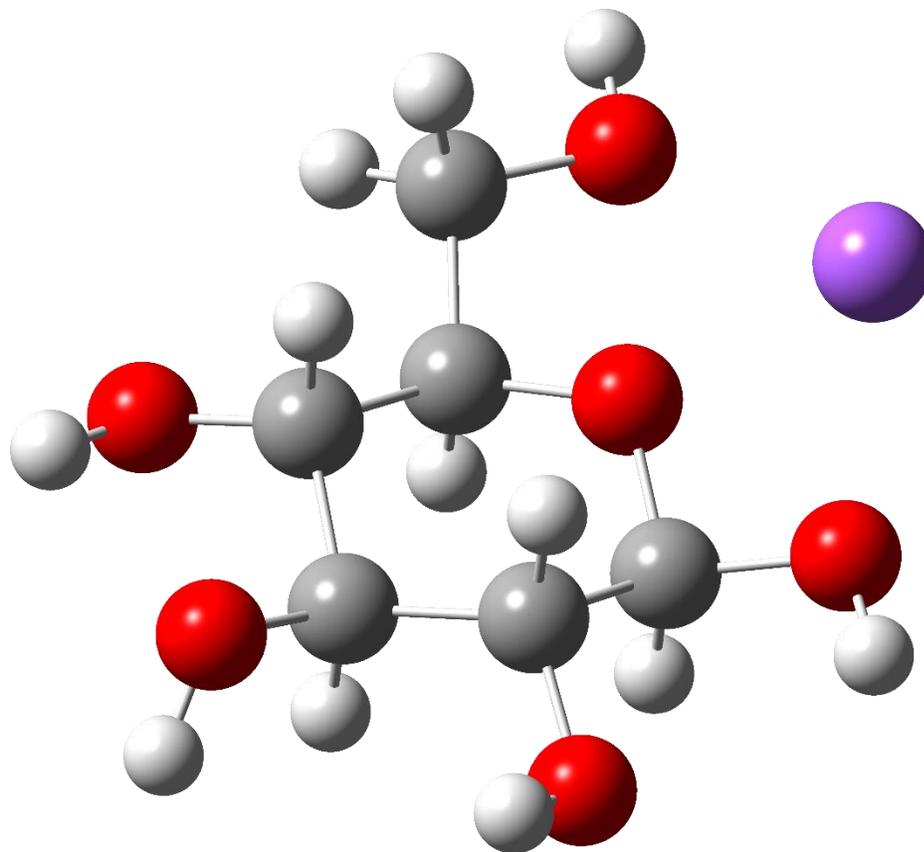


1.4.4 3D Visual Representation

Throughout the following chapters, results of the calculations are visualized as 3D ball and stick models. (Figure 1-8) Atoms are color coded by their type: Carbon in gray, hydrogen in white, oxygen in red, and sodium in purple.

Figure 1-8:

Example ball and stick model of glucose



1.4.5 Thermochemical Data

One of the properties that can be obtained from calculations is the thermochemical contributions for a given system. These parameters allow for the familiar equation for Gibbs free energy to be populated,

$$\Delta G = \Delta H - T\Delta S$$

where ΔG is the change in Gibbs free energy, ΔH is the change in enthalpy, T is the temperature and ΔS is the change in entropy. In addition, the vibrational contributions to

the zero-point energy are also obtainable, allowing for a more realistic energetic description of a system than is offered by the uncorrected total electronic energy.

1.4.6 Ion Affinities

One application of these thermochemical parameters that is used is the calculation of the ion affinity for a particular molecule. This is often referred to specifically as the proton or sodium (lithium, calcium, etc.) affinity of a molecule. The proton affinity of a molecule is defined as the negative ΔH for the reaction $B + H^+ \rightarrow BH^+$ where B is the neutral analyte. Similarly, the sodium affinity of a molecule would be defined as the negative ΔH for the reaction $B + Na^+ \rightarrow BNa^+$. These affinity values can be calculated for different isomers and conformers of a putative analyte and used to help determine the species likely found experimentally. This technique is used in chapter 2.

1.4.7 Rice-Ramsperger-Kassel-Markus (RRKM) Theory

One of the ways which this thermochemical information can be used is in RRKM modeling. RRKM theory[45–47] is a model of unimolecular dissociation reactions and can be used to predict the rate of a reaction/dissociation pathway as a function of internal energy. By modeling reaction rate as a function of energy, we can help explain why some dissociation pathways are observed preferentially over others in terms of experimental conditions and are employed in chapter 2.

1.4.8 Custom Utilities

Several utilities written in the Python programming language were developed over the course of this work to assist with various calculation related tasks.

Mksub.py. Reads a Gaussian input file to create a customized submission script with the specified computational and time resources being allocated automatically, including accounting for resource overhead for the computer scheduler system.

Parselog.py. Reads an arbitrary number of Gaussian output files and automatically extracts the numerical results, sorts, formats, and tabulates the results into a turn-key spreadsheet. Used to deduplicate degenerate structures and quickly visualize relative energies of many structures. Has facilities to automatically detect and adjust the format of the spreadsheet depending on the type of calculation being parsed.

Mkfromlastlog.py. Reads an arbitrary number of Gaussian output files and creates new input files from the final optimized structure. Primarily used when a low-level calculation has been performed as a first wide pass to prepare a set of calculations to run at a more sophisticated level of theory.

Bonds.py. Reads a Gaussian cartesian input file and places a new atom in a specified spatial location relative to the existing structure. This is both useful and needed as ensembles of Gaussian structures are not aligned to a common origin. Originally written to protonate carboxylic acids, but it has since been extended to allow for the addition of any user-specified atom into any chemical environment.

Ffout2gjf.py. Reads a Fafoom output file and creates a set of Gaussian input files with user-specified settings along with a specially formatted submission script to allow for the controlled submission of 10s-100s of calculations with a single command.

1.5 Infrared “Action” Spectroscopy

An additional dimension to the MS/MS or computational results can be obtained from spectroscopic methods. The characteristic vibrational frequencies of an ion may be probed inside a mass spectrometer with a suitably designed experiment. One such experimental method is Infrared Multiphoton Dissociation (IRMPD), whereupon ions are trapped in a confined space and irradiated with an infrared laser[48–51]. If the wavelength of the laser matches a vibrational mode in the ion, a photon can be absorbed. If enough photons are absorbed by the ion, sufficient energy is transferred such that bonds begin to break. This can be observed as a loss of signal for the precursor ion and a gain in signal for a fragment ion. Symbolically this is represented as,

$$IRMPD\ yield = \left(\sum I_f \right) / \left(I_p \sum I_f \right)$$

Where I_p is the precursor ion intensity, I_f is the fragment ion intensity, and the summation is over all fragment ions.

By scanning the wavelength of the laser, the fragmentation yield as a function of laser wavelength can be plotted, which can be related to the gas-phase IR absorption spectrum for the ion. These vibrational frequencies can be estimated using computational chemistry for comparison with the experiment. It should be noted that the harmonic oscillator approximation used for calculating the vibrational frequencies can show discrepancies with the experimental IRMPD intensity, but that band positions are typically well preserved[52].

Several experimental considerations need to be noted for this method: The type of instrument used must be capable of trapping ions for a tangible amount of time. In

practical terms, a quadrupole ion trap or Fourier-transform ion cyclotron resonance (FTICR) instrument is required. An additional consideration is the laser wavelength required. A tunable laser is needed to perform a practical IRMPD experiment. While suitable IR lasers have become more widely available, the models typically suited for laboratory deployment are of the optical parametric oscillator (OPO) type and can only access the upper-mid IR region, typically from 3000-4000 cm^{-1} . To investigate the lower frequency region ($\sim 1000\text{-}2000 \text{ cm}^{-1}$), a Free Electron Laser (FEL) is typically used, but such hardware is only available at a handful of facilities around the world currently. For the work presented in chapter 3 on the dissociation and spectroscopy of lactose, the experimental spectroscopy was performed using the FEL located at the Centre Laser Infrarouge d'Orsay (CLIO) in Orsay France[53].

For the work in chapter 4, a modification of the IRMPD scheme described above was used. In this 'tagging' method[54], ions are trapped and cooled to $\sim 60 \text{ K}$. Neutral nitrogen gas is introduced and allowed to mix with the ions. Some of the nitrogen molecules will weakly coordinate to the analyte ions. The analyte ions with the nitrogen tag are then sent from the trap to the detector.

Symbolically this can be represented as $[M] + N_2 \rightarrow [M+N_2]$. As nitrogen has a nominal molecular mass of 28, the ions will be detected at $m/z M+28$.

This tagging scheme can be exploited spectroscopically with a carefully designed experiment. Since the coordination of the nitrogen tag to the analyte is very weak, by irradiating the nitrogen tagged complex with IR photons, any absorption by the analyte ion can be converted into internal vibrational energy. As the energy is redistributed

through the analyte ion, the coordination with the tag can be disrupted, i.e., the tag is lost. Experimentally this is observed as a loss in signal of m/z [M+28] as a function of laser wavelength. This can be extrapolated back to produce action spectra in a similar way to the direct dissociation method described above. One key advantage that is offered by this technique is that due to the low temperature used and the ease of dissociating the nitrogen tag, the recovered line widths tend to be narrower than in conventional IRMPD. (5 cm^{-1} vs $10\text{-}15\text{ cm}^{-1}$ FWHM.)

1.6 Ion Mobility

Ion mobility is an experimental technique that separates ions based on their relative size and shape. For two ions of the same m/z but different conformation, they typically will have different average cross sections. In a typical ion mobility experiment, ions are propelled through a bath of neutral gas molecules using a constant applied electric potential[55]. This is similar to CID but performed at a lower energy than in a CID experiment such that no fragmentation should occur. As the ions travel the neutral gas bath, collisions will occur as a function of the size of the ion. A larger, more extended ion conformation will present a larger cross-sectional area than a more compact conformation and on average experience more neutral gas collisions, slowing its path through the gas bath. The time-of-flight for the ions can be measured and the ion mobility determined. Symbolically this mobility K is given as,

$$K = \frac{v_d}{E} = \frac{l}{t_d E}$$

Where v_d is the velocity of the ion, E the applied external potential, l the length of the mobility cell, and t_d the time taken for the ion to traverse the mobility cell.

If v_d is small compared with the ion thermal velocity, K can be expressed as,

$$K = \frac{3}{16} \sqrt{\frac{2\pi}{\mu k_B T}} \times \frac{z}{N\Omega}$$

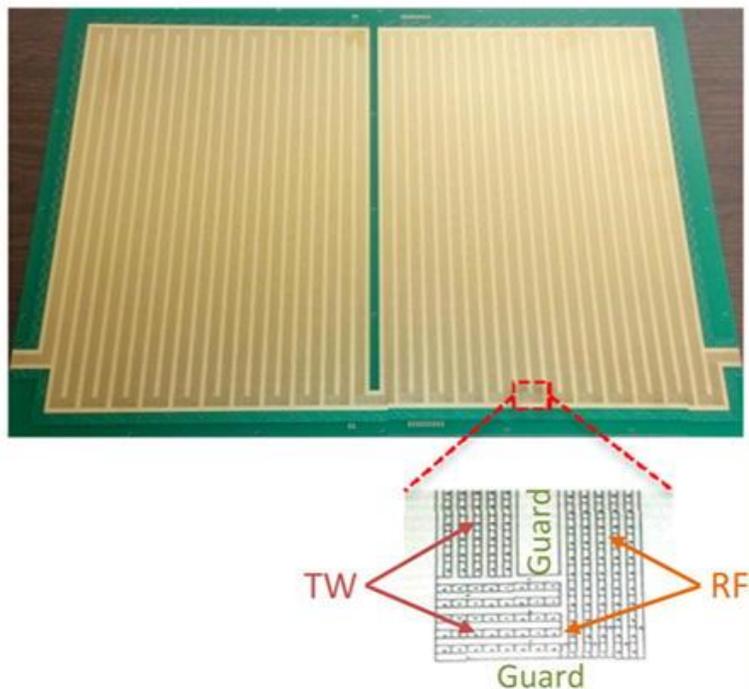
Where μ is the reduced mass of the ion and neutral gas ($\mu = \frac{m_I m_G}{m_I + m_G}$), k_B is the Boltzmann constant, T the temperature, z the analyte charge, N the gas number density, and Ω is the cross-sectional area of the ion.

1.6.1 Structures for Lossless Ion Manipulation (SLIM)

The ion mobility experiments performed in chapter 4 use a device known as a Structures for Lossless Ion Manipulation or SLIM device. (Figure 1-9) This device allows for an extended flight path for ions to be contained in a very compact physical area. The device consists of a pair of printed circuit boards (PCBs) that are mounted a few millimeters apart. The PCBs have a serpentine series of traces affixed to them allowing for ions to be electronically manipulated by varying the applied potentials to different points around the PCB. By assembling the SLIM device in an enclosure pressurized with a suitable collision gas such as nitrogen or helium, ion mobility experiments can be performed with a long mobility cell but compact physical size.

Figure 1-9:

Representative example of a partial SLIM device. Reprinted (adapted) with permission from [56] Copyright 2016 American Chemical Society.



1.7 References

1. Levsen, K., Schwarz, H.: Gas-phase chemistry of collisionally activated ions. *Mass Spectrometry Reviews*. 2, 77–148 (1983). <https://doi.org/10.1002/mas.1280020104>
2. Bordas-Nagy, J., Jennings, K.R.: Collision-induced decomposition of ions. *International Journal of Mass Spectrometry and Ion Processes*. 100, 105–131 (1990). [https://doi.org/10.1016/0168-1176\(90\)85071-9](https://doi.org/10.1016/0168-1176(90)85071-9)
3. Gross, J.H.: Instrumentation. In: Gross, J.H. (ed.) *Mass Spectrometry: A Textbook*. pp. 151–292. Springer International Publishing, Cham (2017)

4. Stephens, W.E.: A Pulsed Mass Spectrometer with Time Dispersion. *Phys. Rev.* 69, 674–674 (1946). <https://doi.org/10.1103/PhysRev.69.674>
5. Wolff, M.M., Stephens, W.E.: A Pulsed Mass Spectrometer with Time Dispersion. *Review of Scientific Instruments.* 24, 616–617 (1953).
<https://doi.org/10.1063/1.1770801>
6. Kelly, R.T., Tolmachev, A.V., Page, J.S., Tang, K., Smith, R.D.: The ion funnel: Theory, implementations, and applications. *Mass Spectrometry Reviews.* 29, 294–312 (2010). <https://doi.org/10.1002/mas.20232>
7. Mamyryn, B.A., Karataev, V.I., Shmikk, D.V., Zagulin, V.A.: The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Journal of Experimental and Theoretical Physics.* 37, 45 (1973)
8. Mamyryn, B.A.: Laser assisted reflectron time-of-flight mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes.* 131, 1–19 (1994).
[https://doi.org/10.1016/0168-1176\(93\)03891-O](https://doi.org/10.1016/0168-1176(93)03891-O)
9. Mamyryn, B.A.: Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry.* 206, 251–266 (2001).
[https://doi.org/10.1016/S1387-3806\(00\)00392-4](https://doi.org/10.1016/S1387-3806(00)00392-4)
10. Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E. eds: *Essentials of Glycobiology.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY) (2009)

11. An, H.J., Kronewitter, S.R., de Leoz, M.L.A., Lebrilla, C.B.: Glycomics and disease markers. *Current Opinion in Chemical Biology*. 13, 601–607 (2009).
<https://doi.org/10.1016/j.cbpa.2009.08.015>
12. Adamczyk, B., Tharmalingam, T., Rudd, P.M.: Glycans as cancer biomarkers. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 1820, 1347–1353 (2012).
<https://doi.org/10.1016/j.bbagen.2011.12.001>
13. Ruhaak, L.R., Miyamoto, S., Lebrilla, C.B.: Developments in the Identification of Glycan Biomarkers for the Detection of Cancer. *Molecular & Cellular Proteomics*. 12, 846–855 (2013). <https://doi.org/10.1074/mcp.R112.026799>
14. Szajda, S.D., Waszkiewicz, N., Chojnowska, S., Zwierz, K.: Carbohydrate markers of pancreatic cancer. *Biochemical Society Transactions*. 39, 340–343 (2011).
<https://doi.org/10.1042/BST0390340>
15. Kirwan, A., Utratna, M., O'Dwyer, M.E., Joshi, L., Kilcoyne, M.: Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *BioMed Research International*. 2015, e490531 (2015). <https://doi.org/10.1155/2015/490531>
16. Panza, M., Pistorio, S.G., Stine, K.J., Demchenko, A.V.: Automated Chemical Oligosaccharide Synthesis: Novel Approach to Traditional Challenges. *Chem. Rev.* 118, 8105–8150 (2018). <https://doi.org/10.1021/acs.chemrev.8b00051>
17. Bythell, B.J., Abutokaikah, M.T., Wagoner, A.R., Guan, S., Rabus, J.M.: Cationized Carbohydrate Gas-Phase Fragmentation Chemistry. *J. Am. Soc. Mass Spectrom.* 28, 688–703 (2017). <https://doi.org/10.1007/s13361-016-1530-x>

18. Chiu, C., Tsai, S.-T., Hsu, P.-J., Huynh, H.T., Chen, J.-L., Phan, H.T., Huang, S.-P., Lin, H.-Y., Kuo, J.-L., Ni, C.-K.: Unexpected Dissociation Mechanism of Sodiated N-Acetylglucosamine and N-Acetylgalactosamine. *J. Phys. Chem. A.* 123, 3441–3453 (2019). <https://doi.org/10.1021/acs.jpca.9b00934>
19. Laine, R.A.: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology.* 4, 759–767 (1994)
20. Domon, B., Costello, C.E.: A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal.* 5, 397–409 (1988)
21. Cancilla, M.T., Penn, S.G., Carroll, J.A., Lebrilla, C.B.: Coordination of Alkali Metals to Oligosaccharides Dictates Fragmentation Behavior in Matrix Assisted Laser Desorption Ionization/Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.* 118, 6736–6745 (1996). <https://doi.org/10.1021/ja9603766>
22. Cancilla, M.T., Penn, S.G., Lebrilla, C.B.: Alkaline Degradation of Oligosaccharides Coupled with Matrix-Assisted Laser Desorption/Ionization Fourier Transform Mass Spectrometry: A Method for Sequencing Oligosaccharides. *Anal. Chem.* 70, 663–672 (1998). <https://doi.org/10.1021/ac9711100>
23. Orlando, R., Bush, C.A., Fenselau, C.: Structural analysis of oligosaccharides by tandem mass spectrometry: Collisional activation of sodium adduct ions.

- Biomedical & Environmental Mass Spectrometry. 19, 747–754 (1990).
<https://doi.org/10.1002/bms.1200191202>
24. Rabus, J.M., Abutokaikah, M.T., Ross, R.T., Bythell, B.J.: Sodium-cationized carbohydrate gas-phase fragmentation chemistry: influence of glycosidic linkage position. *Phys. Chem. Chem. Phys.* 19, 25643–25652 (2017).
<https://doi.org/10.1039/C7CP04738J>
25. Abutokaikah, M.T., Frye, J.W., Tschampel, J., Rabus, J.M., Bythell, B.J.: Fragmentation Pathways of Lithiated Hexose Monosaccharides. *J. Am. Soc. Mass Spectrom.* 1–11 (2018). <https://doi.org/10.1007/s13361-018-1973-3>
26. Kailemia, M.J., Ruhaak, L.R., Lebrilla, C.B., Amster, I.J.: Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments. *Analytical Chemistry.* 86, 196–212 (2014). <https://doi.org/10.1021/ac403969n>
27. Rabus, J.M., Simmons, D.R., Maître, P., Bythell, B.J.: Deprotonated carbohydrate anion fragmentation chemistry: structural evidence from tandem mass spectrometry, infra-red spectroscopy, and theory. *Physical Chemistry Chemical Physics.* 20, 27897–27909 (2018). <https://doi.org/10.1039/C8CP02620C>
28. Bythell, B.J., Rabus, J.M., Wagoner, A.R., Abutokaikah, M.T., Maître, P.: Sequence Ion Structures and Dissociation Chemistry of Deprotonated Sucrose Anions. *J. Am. Soc. Mass Spectrom.* 29, 2380–2393 (2018). <https://doi.org/10.1007/s13361-018-2065-0>

29. Cai, Y., Jiang, Y., Cole, R.B.: Anionic Adducts of Oligosaccharides by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Anal. Chem.* 75, 1638–1644 (2003). <https://doi.org/10.1021/ac0205513>
30. Harvey, D.J.: Fragmentation of negative ions from carbohydrates: Part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. *J Am Soc Mass Spectrom.* 16, 622–630 (2005). <https://doi.org/10.1016/j.jasms.2005.01.004>
31. Cole, R.B., Zhu, J.: Chloride anion attachment in negative ion electrospray ionization mass spectrometry†. *Rapid Communications in Mass Spectrometry.* 13, 607–611 (1999). [https://doi.org/10.1002/\(SICI\)1097-0231\(19990415\)13:7<607::AID-RCM530>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-0231(19990415)13:7<607::AID-RCM530>3.0.CO;2-2)
32. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, J.M., Klene, M., Knox, J.E., 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., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A.,

- Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J.: Gaussian 09, Revision E.01. Gaussian, Inc., Wallingford CT (2009)
33. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A.V., Bloino, J., Janesko, B.G., Gomperts, R., Mennucci, B., Hratchian, H.P., Ortiz, J.V., Izmaylov, A.F., Sonnenberg, J.L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V.G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J.A., Jr., Peralta, J.E., Ogliaro, F., Bearpark, M.J., Heyd, J.J., Brothers, E.N., Kudin, K.N., Staroverov, V.N., Keith, T.A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A.P., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Millam, J.M., Klene, M., Adamo, C., Cammi, R., Ochterski, J.W., Martin, R.L., Morokuma, K., Farkas, O., Foresman, J.B., Fox, D.J.: Gaussian 16 Revision C.01. (2016)
34. Phillips, J.C., Hardy, D.J., Maia, J.D.C., Stone, J.E., Ribeiro, J.V., Bernardi, R.C., Buch, R., Fiorin, G., Héning, J., Jiang, W., McGreevy, R., Melo, M.C.R., Radak, B.K., Skeel, R.D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z., Kalé, L.V., Schulten, K., Chipot, C., Tajkhorshid, E.: Scalable molecular dynamics on CPU and GPU architectures with `NAMD`. *J. Chem. Phys.* 153, 044130 (2020). <https://doi.org/10.1063/5.0014475>

35. Supady, A., Blum, V., Baldauf, C.: First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model.* 55, 2338–2348 (2015).
<https://doi.org/10.1021/acs.jcim.5b00243>
36. Marianski, M., Supady, A., Ingram, T., Schneider, M., Baldauf, C.: Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α -Maltose. *J. Chem. Theory Comput.* 12, 6157–6168 (2016). <https://doi.org/10.1021/acs.jctc.6b00876>
37. Bannwarth, C., Caldeweyher, E., Ehlert, S., Hansen, A., Pracht, P., Seibert, J., Spicher, S., Grimme, S.: Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science.* 11, e1493 (2021).
<https://doi.org/10.1002/wcms.1493>
38. Pracht, P., Bohle, F., Grimme, S.: Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.* 10.1039/C9CP06869D (2020). <https://doi.org/10.1039/C9CP06869D>
39. Cremer, D., Pople, J.A.: General definition of ring puckering coordinates. *J. Am. Chem. Soc.* 97, 1354–1358 (1975). <https://doi.org/10.1021/ja00839a011>
40. Cremer, D., Pople, J.A.: Molecular orbital theory of the electronic structure of organic compounds. XXIII. Pseudorotation in saturated five-membered ring compounds. *J. Am. Chem. Soc.* 97, 1358–1367 (1975).
<https://doi.org/10.1021/ja00839a012>

41. Haasnoot, C.A.G.: The conformation of six-membered rings described by puckering coordinates derived from endocyclic torsion angles. *J. Am. Chem. Soc.* 114, 882–887 (1992). <https://doi.org/10.1021/ja00029a013>
42. Mayes, H.B., Broadbelt, L.J., Beckham, G.T.: How Sugars Pucker: Electronic Structure Calculations Map the Kinetic Landscape of Five Biologically Paramount Monosaccharides and Their Implications for Enzymatic Catalysis. *J. Am. Chem. Soc.* 136, 1008–1022 (2014). <https://doi.org/10.1021/ja410264d>
43. Becke, A.D.: Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* 140, 18A301 (2014). <https://doi.org/10.1063/1.4869598>
44. Bechstedt, F.: Density Functional Theory. In: Bechstedt, F. (ed.) *Many-Body Approach to Electronic Excitations: Concepts and Applications*. pp. 73–88. Springer, Berlin, Heidelberg (2015)
45. Gilbert, R.G., Smith, S.C.: *Theory of unimolecular and recombination reactions*. Blackwell Scientific Publications ; Publishers' Business Services [distributor, Oxford; Boston; Brookline Village, Mass. (1990)
46. Lifshitz, C.: Some recent aspects of unimolecular gas phase ion chemistry. *Chem. Soc. Rev.* 30, 186–192 (2001). <https://doi.org/10.1039/B005243O>
47. Beyer, T., Swinehart, D.F.: Algorithm 448: number of multiply-restricted partitions. *Commun. ACM.* 16, 379 (1973). <https://doi.org/10.1145/362248.362275>
48. Lemaire, J., Boissel, P., Heninger, M., Mauclaire, G., Bellec, G., Mestdagh, H., Simon, A., Caer, S.L., Ortega, J.M., Glotin, F., Maitre, P.: *Gas Phase Infrared*

- Spectroscopy of Selectively Prepared Ions. *Phys. Rev. Lett.* 89, 273002 (2002).
<https://doi.org/10.1103/PhysRevLett.89.273002>
49. MacAleese, L., Maître, P.: Infrared spectroscopy of organometallic ions in the gas phase: From model to real world complexes. *Mass Spectrometry Reviews.* 26, 583–605 (2007). <https://doi.org/10.1002/mas.20138>
50. Oomens, J., Sartakov, B.G., Meijer, G., von Helden, G.: Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions. *International Journal of Mass Spectrometry.* 254, 1–19 (2006).
<https://doi.org/10.1016/j.ijms.2006.05.009>
51. Polfer, N.C., Oomens, J.: Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrometry Reviews.* 28, 468–494 (2009). <https://doi.org/10.1002/mas.20215>
52. Maitre, P., Scuderi, D., Corinti, D., Chiavarino, B., Crestoni, M.E., Fornarini, S.: Applications of Infrared Multiple Photon Dissociation (IRMPD) to the Detection of Posttranslational Modifications. *Chem. Rev.* 120, 3261–3295 (2020).
<https://doi.org/10.1021/acs.chemrev.9b00395>
53. Prazeres, R., Glotin, F., Insa, C., Jaroszynski, D.A., Ortega, J.M.: Two-colour operation of a Free-Electron Laser and applications in the mid-infrared. *Eur. Phys. J. D.* 3, 87–93 (1998). <https://doi.org/10.1007/s100530050151>
54. Masellis, C., Khanal, N., Kamrath, M.Z., Clemmer, D.E., Rizzo, T.R.: Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J.*

- Am. Soc. Mass Spectrom. 28, 2217–2222 (2017). <https://doi.org/10.1007/s13361-017-1728-6>
55. Gabelica, V., Marklund, E.: Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology*. 42, 51–59 (2018).
<https://doi.org/10.1016/j.cbpa.2017.10.022>
56. Deng, L., Ibrahim, Y.M., Hamid, A.M., Garimella, S.V.B., Webb, I.K., Zheng, X., Prost, S.A., Sandoval, J.A., Norheim, R.V., Anderson, G.A., Tolmachev, A.V., Baker, E.S., Smith, R.D.: Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module. *Anal. Chem.* 88, 8957–8964 (2016).
<https://doi.org/10.1021/acs.analchem.6b01915>

Chapter 2 : Sodium-Cationized Carbohydrate Gas-Phase Fragmentation

Chemistry: Influence of Glycosidic Linkage Position

Jordan M. Rabus, Maha T. Abutokaikah, Reginald T. Ross, Benjamin J. Bythell

Department of Chemistry and Biochemistry, University of Missouri, St. Louis, MO

63121, USA

Reproduced from Physical Chemistry Chemical Physics. 2017, 19 (37), 25643–25652

with permission from the PCCP Owner Societies.

<https://doi.org/10.1039/C7CP04738J>

2.1 Abstract

We investigate the gas-phase structures and fragmentation chemistry of two isomeric sodium-cationized carbohydrates using combined tandem mass spectrometry, hydrogen/deuterium exchange experiments, and computational methods. Our model systems are the glucose-based disaccharide analytes, cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)-D-Glucose), and gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)-D-Glucose). These analytes show substantially different tandem mass spectra. We characterize the rate-determining barriers to both the glycosidic and structurally informative cross-ring bond cleavages. Sodiated cellobiose produces abundant Y_1 and B_1 peaks. Our deuterium labeling and computational chemistry approach provides evidence for 1,6-anhydroglucose B_1 ion structures rather than the 1,2-anhydroglucose and oxacarbenium ion structures proposed elsewhere. Unlike those earlier proposals, this finding is consistent with the experimentally observed B_n/Y_m branching ratios. In contrast to cellobiose, sodiated gentiobiose primarily fragments by cross-ring cleavage to form

various A_2 ion types. Fragmentation is facilitated by ring-opening at the reducing end which enables losses of $C_nH_{2n}O_n$ oligomers. Deuterium labeling and theory enables rationalization of these processes. Theory and experiment also support the importance of consecutive fragmentation processes at higher collision energies.

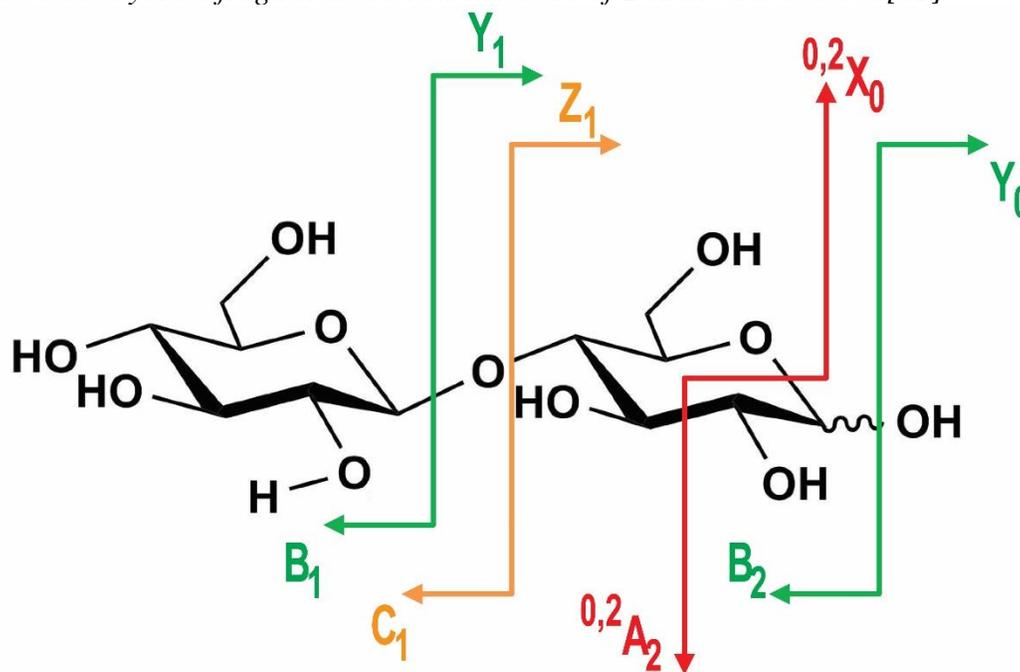
2.2 Introduction

Complex carbohydrates (glycans) play critical roles in a huge number of biological processes. To understand such processes we need to be able to accurately identify which specific sequences and ideally structures are involved.[1-10] The primary approach to determine these sequences for low-abundance species as are found in biological systems is tandem mass spectrometry (MS/MS or MSⁿ).[4, 11-28] Glycans are often analyzed with one or more metal cation adducts present to provide the charge.[4, 9, 10, 14, 16, 17, 29, 30] The mass-to-charge (m/z) ratios of the detected fragments and the corresponding precursor ions provide the elemental composition information from which structure is inferred. Glycosidic bond cleavages produce series of B_n and/or C_n ions if the non-reducing end fragment keeps a charge (Figure 2-1)³¹, Y_m and/or Z_m ions if the reducing end fragment keeps a charge or a mixture of fragments. Cross-ring cleavages to form A_n and/or X_m ions provide additional information which aids sequence assignment and differentiation of isomers. This capability is critical as despite the relatively simple hexose/pentose carbohydrate building blocks there is an enormous number of glycan isomers. To accurately assign glycan structure, knowledge of the carbohydrate monomer identities and their sequence, the position of the glycan linkages and their stereochemistry, as well as the anomericity (α/β) are all necessary.[1-8] Thus, as multiple

isomeric glycans exist, it follows that multiple potential isomeric precursor ions for each fragment ion peak exist as well.

Figure 2-1:

The carbohydrate fragmentation nomenclature of Domon and Costello[31]



For example, sodiated Lactose, Cellobiose, and Gentiobiose can all produce B_1 peaks at m/z 185 and Y_1 peaks at m/z 203. As a result, each sequence assignment currently needs to be checked and confirmed against a standard. In addition to the presence/absence of each sequence ion peak, the relative abundances of the peaks are critical in distinguishing the identity of the analytes.^{28, 32-34} For the many isomeric glycans, this combination of peak prevalence and relative abundance offers the means of identification. This is in marked contrast to peptide systems, where almost all

identification algorithms completely ignore the relative abundance information provided because the fragment ion m/z values typically differ substantially and so readily provide ample ability to discriminate most sequences.[35-37]

Currently, direct comparisons of unknown glycan mass spectra with mass spectral standards cataloged in MS/MS libraries are utilized.[28, 32-34, 38-43] These MS/MS libraries have been highly effective for analyses where particular glycans are known or expected to be prevalent. However, the library-based approach is predicated on the existence of a spectral standard in the library for the glycan under investigation which was collected under similar experimental conditions. A comprehensive spectral library of curated and annotated MS/MS glycan spectra which supports their identification based on the correlation between the experimental and library spectra would be incredibly useful. Identification of unknown glycans from their tandem mass spectra would simply involve comparing the experimental spectrum to the library spectra and looking for a match. Currently, the considerable difficulty in synthesizing standard compounds[44] for many glycans as well as the sheer number of possibilities makes comprehensive glycan library-based approaches impractical.

The field would greatly benefit from any additional means of discrimination between glycans that can be mustered.[45] Recently, ion mobility-based[46-55] separations prior to MS^n have shown some promise, enabling discrimination between isomers for some glycans. Unfortunately, both the nature of the cation and the sugar affect the degree of separation in a non-obvious manner, so separations are far from universally successful. More traditional chromatographic and electrophoretic techniques

are routinely utilized. Separation of unknown structurally isomeric glycans prior to tandem mass spectrometry-based fragmentation is highly desirable. However, this still requires a practical means of accurately interpreting the subsequent MS/MS spectra for each isomer. One area of clear potential is in understanding the factors that govern the gas-phase fragmentation chemistry of charged glycans more accurately.[56] These chemistries ultimately determine our ability to identify glycan sequence and primary structure.[56, 57] The mechanistic understanding of metal-cationized glycan fragmentation chemistry is primarily based on qualitative proposals developed from early mass-spectral analyses of comparatively small systems.[17, 18, 29] In contrast, our initial work in this area indicated that additional gas-phase ion structures need to be considered (for both transition structures and product ions) and that these directly influence the respective product ion branching ratios.[56] If we could augment the size and variety of analytes for which accurate mechanistic information was known, then our predictive capability and confidence in the identifications generated based on this chemical information would improve substantially. A better understanding of the fragmentation chemistry could be leveraged to provide superior confidence in assignment of glycan peak identity, relative abundance, and propensity as a function of conditions.

In the current article, we expand on our initial investigation[56] and present results on the mechanisms and energetics of polysaccharide fragmentation of isomeric disaccharides using a combination of isotopically labeled substrates, tandem mass spectrometry, and computational methods. We investigate two glucose-based disaccharide analytes, Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)-D-Glucose) and

Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)-D-Glucose) cationized with sodium. We compare our findings to the prior literature[18, 23, 29, 56, 58-64] and make structural and energetic predictions based on our data.

2.3 Experimental Methods

Experimental work was carried out using a MaXis plus electrospray-quadrupole time-of-flight mass spectrometer (Bruker, Billerica, MA). Tandem mass spectra for $[M+Na]^+$ ions were obtained by mass-selecting the appropriate ion with the quadrupole, collision-induced dissociation (CID) in the collision cell followed by mass-to-charge dependent product dispersion by the time-of-flight analyzer. Data were collected as a function of collision energy. Experiments were performed using a sample rate of 1 Hz for 1 minute, thus each data point is the average of 60 individual mass spectra. Typical variations in relative abundance were approximately $\pm 0.95\%$ using a confidence interval of 95%. Breakdown graphs expressing the relative fragment ion signals as a function of collision energy were obtained for all species studied.

Ionization was by electrospray with the samples infused into the instrument in $\sim 5 \mu\text{M}$ acetonitrile/water/formic acid (50/50/0.1%) solutions at a flow rate of $3 \mu\text{l min}^{-1}$.

Nitrogen was used as nebulizing, drying, and collision gas. For the deuteration experiments,^{18, 56} analytes were dissolved in D_2O for 10 min at room temperature, and the solution was further diluted in acetonitrile/ D_2O (50/50%) to the final concentration of $\sim 5 \mu\text{M}$. Deuterium oxide was purchased from Cambridge Isotope Laboratories, Inc (Tewksbury, MA). HPLC-grade acetonitrile and water were purchased for Sigma-Aldrich (St. Louis, MO).

2.4 Theoretical Methods

Molecular dynamics simulations were performed to enable effective characterization of the potential energy surface of cationized glycan analytes. We began by sampling the reactant minima with molecular dynamics simulations in NAMD utilizing a CHARMM force field.[65-70] Trial structures were generated via a temperature replica exchange molecular dynamics scheme (T-REMD). Parallel minimization calculations were performed at multiple simulated temperatures, with the periodic exchange of the minimized structures to a different temperature. Snapshots of the structures are saved throughout the run, for further analysis. This process generates too many structures to time-efficiently optimize at a high level of theory. Instead our in-house script groups and sorts these trial structures into families based on hydrogen-bonding and dihedral angles from which only the best candidates are optimized at each successively more realistic level of theory: PM6 \rightarrow HF/3-21g \rightarrow PBE0/6-31G(d) \rightarrow M06-2X/6-31+G(d,p)[71-73] with regrouping and sorting after each level. This technique is a custom modification of prior approaches.[74-77]

The density functional calculations of minima, transition structures, product ions, and neutrals were performed with the Gaussian 09 suite of programs.[78] Targeted single-point calculations at the M06-2X, B3LYP, and MP2(Full) levels of theory were performed on selected, critical precursor ion and transition structures to help assess variability in the description of the chemistry utilizing the 6-311++G(2d,p) basis set. Multiple transition structures (TSs) were calculated for each potential pathway. Minima and TSs were tested by vibrational analysis (all real frequencies and 1 imaginary

frequency, respectively). The potential energy surface generated combined the zero-point energy correction (ZPE) to the electronic energy (E_{el} , 0 K) for improved accuracy ($\Delta E_{\text{el}+\text{ZPE},0\text{K}}$). The related, standard enthalpy ($\Delta H_{298\text{K}}$), Gibbs free energy ($\Delta G_{298\text{K}}$), and entropy ($\Delta S_{298\text{K}}$) corrections to 298 K were also performed. The reaction pathway through each TS was determined by intrinsic reaction coordinate (IRC) calculations with up to 10 steps in each direction. The terminating points of these calculations (one on product-side, one on reactant-side) were then optimized further to determine the minima connected by the TS. Estimates of the metal affinities of the leaving groups were determined as the difference between the zero-point energy-corrected M06-2X/6-31+G(d,p) total electronic energies (E_{el} , 0 K) of the metal-cationized analyte, and the sum of the neutral analyte and the cation at infinite separation.

Rice–Ramsperger–Kassel–Markus (RRKM) calculations were performed using the energetics, vibrational frequencies, and rotational constants derived from the modeling to approximate the time scale of the fragmentation reactions.[79, 80] The Beyer–Swinehart direct count algorithm[81] is used for rotational–vibrational treatment of both the reactant and the transition structure.

2.5 Results and Discussion

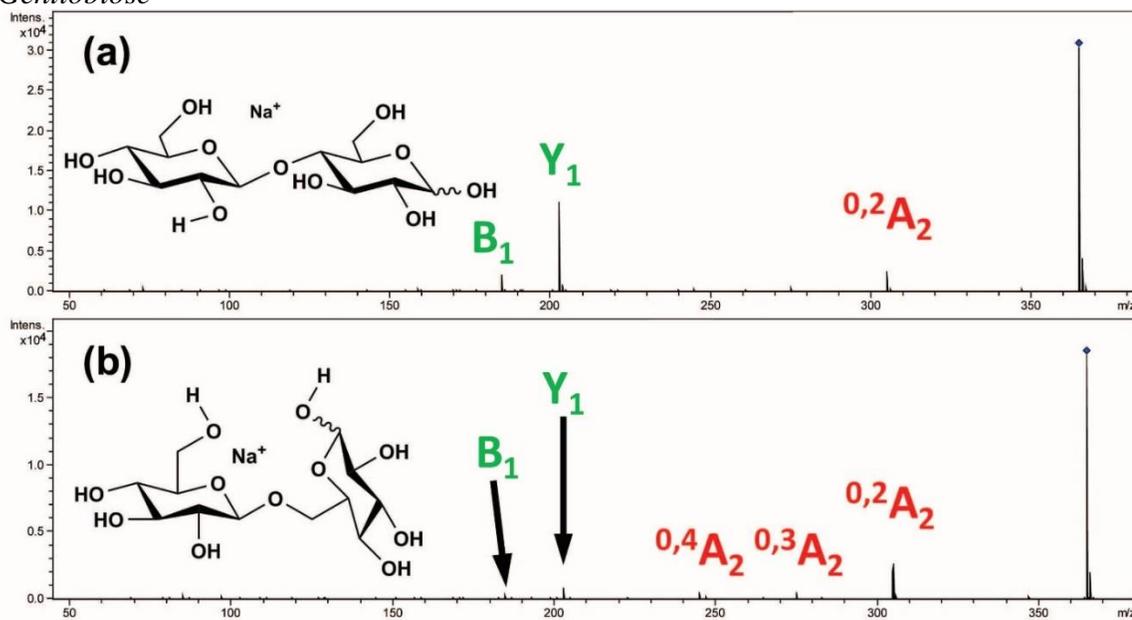
2.5.1 *Experimental Findings*

Consistent with the prevailing literature[17, 18, 29, 46] our experimental data point to clear differences in the fragmentation chemistry of Cellobiose and Gentiobiose when cationized with sodium (Figure 2-2, Figures 2S1-2S4). These differences are both in relative abundance of peaks and order of appearance. $[\text{Cellobiose}+\text{Na}]^+$ predominantly

produces glycosidic bond cleavage product peaks; B_1 and Y_1 . At higher collision energies a $^{0,2}A_2$ peak at m/z 305 is detected and eventually evidence of a very low abundance $^{2,4}A_2$ peak at m/z 245. In contrast, $[\text{Gentiobiose}+\text{Na}]^+$ requires substantially higher collision energies to initiate fragmentation and fragments initially to produce m/z 305, assigned as $^{0,2}A_2$. Formation of the additional cross-ring fragments $^{0,3}A_2$ and $^{0,4}A_2$ (m/z 275 and 245) and glycosidic bond cleavage products requires additional energy to initiate. The abundance of the $^{0,4}A_2$ peak increases substantially at higher collision energies, taking advantage of either consecutive processes or substantial entropic favorability of the relevant pathway(s).

Figure 2-2:

Example MS/MS spectra ($E_{\text{collisions, lab}} = 38 \text{ eV}$) illustrating the differing degree and nature of fragmentation of the isomeric sodium cationized analytes (a) Cellobiose and (b) Gentiobiose



To assess the potential contribution of C-H protons versus hydroxyl protons we performed hydrogen-to-deuterium exchange of all eight hydroxyl protons. This results in an 8 u shift in m/z for the precursor ion and diagnostic shifts for the various fragment ions.[18, 56] These data are summarized in Table 2-1 (Figure 2S5 and 2S6) and will be referenced in the following sections. We shall discuss the structural and mechanistic differences underlying the dissimilar tandem mass spectra presented in Figure 2-2 (and Figures 2S1-2S4), with the aid of our computational data.

Table 2-1:

Summary of D₈-hexose sodiated disaccharide MS/MS spectra. Analytes were generated by exchanging all hydroxyl protons for deuterons

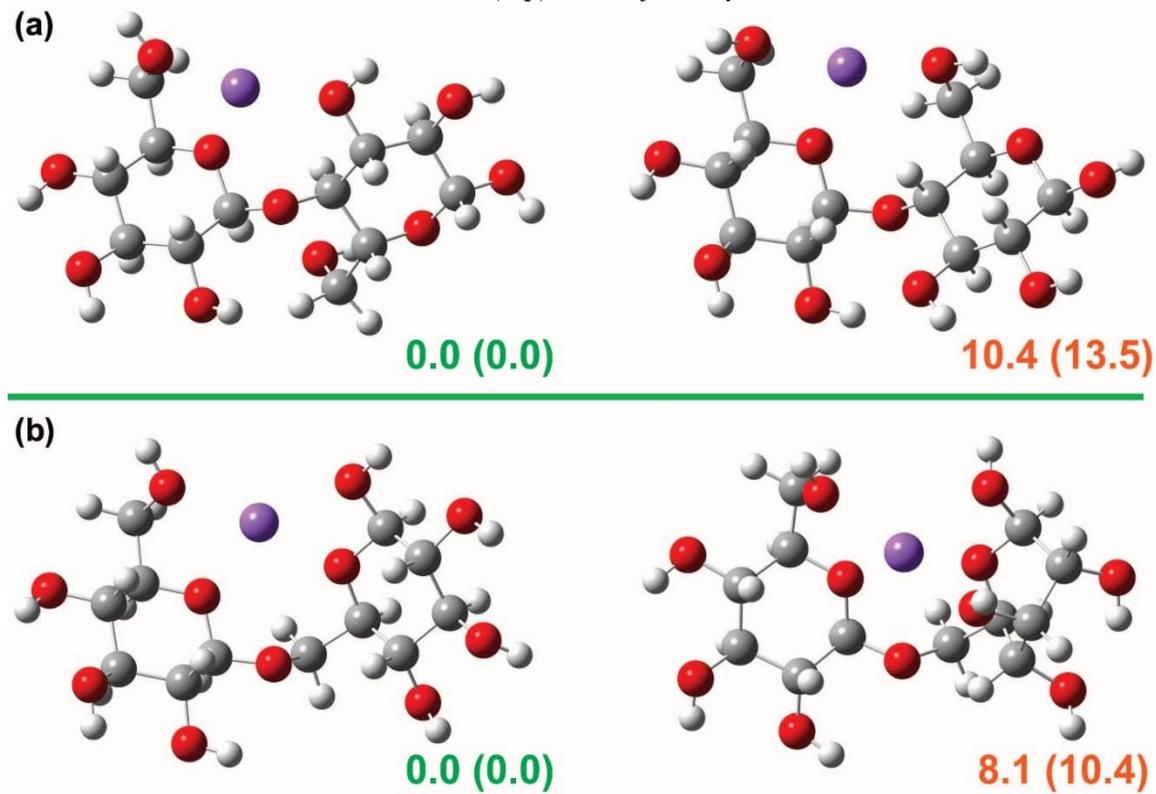
Analyte	<i>m/z</i>	Ion Composition	Assignment	Neutral(s)	$\Delta m/z$	D Retained
Cellobiose	373	[C ₁₂ H ₁₄ D ₈ O ₁₁ +Na] ⁺	[M+Na] ⁺	-	-	8
Cellobiose	353	[C ₁₂ H ₁₄ D ₆ O ₁₀ +Na] ⁺	[M+Na-D ₂ O] ⁺	D ₂ O	-20	6
Cellobiose	311	[C ₁₀ H ₁₂ D ₆ O ₉ +Na] ⁺	^{0,2} A ₂	C ₂ H ₂ D ₂ O ₂	-62	6
Cellobiose	250	[C ₈ H ₉ D ₅ O ₇ +Na] ⁺	^{2,4} A ₂	C ₄ H ₅ D ₃ O ₄	-123	5
Cellobiose	208	[C ₆ H ₇ D ₅ O ₆ +Na] ⁺	<i>Y</i> ₁	C ₆ H ₇ D ₃ O ₅	-168	5
Cellobiose	188	[C ₆ H ₇ D ₃ O ₅ +Na] ⁺	<i>B</i> ₁	C ₆ H ₇ D ₅ O ₆	-185	3
Gentiobiose	373	[C ₁₂ H ₁₄ D ₈ O ₁₁ +Na] ⁺	[M+Na] ⁺	-	-	8
Gentiobiose	353	[C ₁₂ H ₁₄ D ₆ O ₁₀ +Na] ⁺	[M+Na-D ₂ O] ⁺	D ₂ O	-20	6
Gentiobiose	311	[C ₁₀ H ₁₂ D ₆ O ₉ +Na] ⁺	^{0,2} A ₂	C ₂ H ₂ D ₂ O ₂	-62	6
Gentiobiose	280	[C ₉ H ₁₁ D ₅ O ₈ +Na] ⁺	^{0,3} A ₂	C ₃ H ₃ D ₃ O ₃	-93	5
Gentiobiose	249	[C ₈ H ₁₀ D ₄ O ₇ +Na] ⁺	^{0,4} A ₂	C ₄ H ₅ D ₃ O ₄	-124	4
Gentiobiose	208	[C ₆ H ₇ D ₅ O ₆ +Na] ⁺	<i>Y</i> ₁	C ₆ H ₇ D ₃ O ₅	-168	5
Gentiobiose	188	[C ₆ H ₇ D ₃ O ₅ +Na] ⁺	<i>B</i> ₁	C ₆ H ₇ D ₅ O ₆	-185	3

2.5.2 *Energetics of Sodiated Minima*

The lowest energy conformers of sodium cationized Cellobiose are similar to those previously identified for sodium cationized Lactose.[56] There is a substantial number of possibilities[64, 82, 83] making exhaustive discussion impractical. Briefly, hexose rings display chair conformations (Figure 2-3a) with the nearest boat-containing structure (not shown) requiring over 20 kJ mol⁻¹ to populate. In contrast, Gentiobiose while still favoring chair ring conformations in both hexose rings has relatively low energy boat forms too (Figure 2-3b). This is a consequence of the increased flexibility introduced by the 1,6-linkage enabling a greater number of oxygen atoms to coordinate to the sodium cation. The beneficial additional charge-solvation in this cage-like structure is offset by the combination of additional ring strain and a reduction in the number of hydrogen bonds present.

Figure 2-3:

Selected minima of the isomeric sodium cationized analytes (a) Cellobiose and (b) Gentiobiose. The β -D-Glucose forms are shown. Relative energies ($\Delta E_{el+ZPE,0K}$ (ΔG_{298K})) were obtained at the M06-2X/6-31+G(d,p) level of theory



2.5.3 Glycosidic Bond Cleavage: The B_1 - Y_1 Pathways

Consistent with experiment (Figures 2-2, 2S1-2S4), our calculations predict substantial energetic differences between the sodiated Cellobiose and Gentiobiose B_1 - Y_1 pathways. For sodiated Cellobiose, the lowest energy B_1 - Y_1 pathway (Scheme 2-1) begins with rotations of the nonreducing ring and nonreducing terminal carbon-6 hydroxyl group to enable a hydrogen bond between this group and the glycosidic oxygen. This moves the Na^+ away from the glycosidic oxygen forming the reactive configuration. Concerted proton transfer to the glycosidic oxygen, cleavage of the glycosidic bond, then

nucleophilic attack on the electropositive nonreducing end carbon-1 by the newly formed carbon-6 alkoxide forms a 1,6-anhydroglucose structure. This generates an ion-molecule complex (1,6-anhydroglucose \cdots Na $^+$ \cdots D-Glucose) in which competition for the Na $^+$ ensues. Following either proton or sodium ion transfer within the ion-molecule dimer, separation occurs to generate either the sodiated 1,6-anhydrogalactose B_I ion or sodiated Glucose (Y_I ion). Our prior calculations[56] indicated Glucose has a sodium affinity of 183 (α) or 181 (β) kJ mol $^{-1}$ which is consistent with experimental and theoretical work from the Armentrout[84] and Wesdemiotis[85] groups. This value is greater than 1,6-anhydroglucose (167 kJ mol $^{-1}$) and thus consistent with the experimental finding of greater Y_I ion abundance than B_I ($Y_I:B_I$ is \sim 5:1, Figure 2-2).[56, 86] The rate-determining step of the entire process is the B_I - Y_I transition structure (227-231 kJ mol $^{-1}$, $\Delta E_{el+ZPE,0K}$, Table 2-2, Table 2S2). Additionally, this B_I ion structure is consistent with our deuterium labeling data as it retains only 3 hydroxyl deuterons, yielding the expected m/z shift of 3 u (Table 2-1, Scheme 2S1, Figure 2S5). This mechanism differs from that found for [Lactose+Na] $^+$. This is a direct consequence of the presence of Glucose rather than Galactose at the non-reducing end of the molecule which limits the ability to shuttle protons around the non-reducing end and thus favors the present mechanism. We note that both processes generate the 1,6-anhydrohexose product.

Sodiated Gentiobiose glycosidic bond cleavage also proceeds most readily through this type of mechanism (Scheme 2S2), so also produces an ion-molecule complex of 1,6-anhydroGlucose \cdots Na $^+$ \cdots D-Glucose, which then dissociates analogously to the Cellobiose form. The main difference is that the lowest energy Gentiobiose pathway requires a boat conformation at the reducing end of the sugar. Despite this cage-like structure solvating the Na $^+$ cation, a barrier of at least 267 kJ mol $^{-1}$ ($\Delta E_{el+ZPE,OK}$, Table 2-2 (β -anomer), Table 2S2 (α -anomer)) must be overcome to produce B_I or Y_I ions. This value is consistent with the experiments in which the appearance of Y_I and B_I ions required substantially higher collision energies than for [Cellobiose+Na] $^+$.

We considered a substantial variety of alternate mechanisms generating many B_I product structures: oxacarbenium, 1,2-anhydrogalactose, 2-hydroxyglucal, 1,5-anhydro-D-fructose. All of these mechanisms[18, 56, 59, 60] were less energetically competitive. Additionally, the C-H abstraction mechanism[56] is inconsistent with our deuterium labeling experimental data; m/z 208 Y_I , [C $_6$ H $_7$ D $_5$ O $_6$ +Na] $^+$ and m/z 188 B_I , [C $_6$ H $_7$ D $_3$ O $_5$ +Na] $^+$ (Table 2-2, Figure 2S6). Attempts at analogous reactions following ring-opening produced transition structures that were entropically favorable, but considerably less energetically so.

Table 2-2:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- β -D-Glucose) and Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- β -D-Glucose) forms calculated at the M06-2X/6-31+G(d,p) level of theory

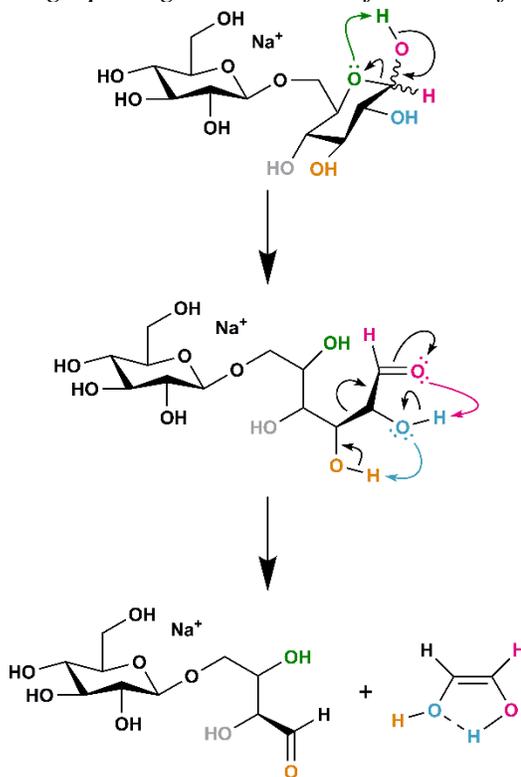
Analyte	Transition Structure	E_{el}/H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298K}/$ kJ mol ⁻¹	$\Delta G_{298K}/$ kJ mol ⁻¹	$\Delta S_{298K}/$ J K ⁻¹ mol ⁻¹
Cellobiose	B_1-Y_1 TS_A	-1459.583018	226.5	224.0	230.1	-20.6
Cellobiose	B_1-Y_1 TS_B	-1459.582210	230.9	229.1	232.0	-9.7
Cellobiose	Ring-opening	-1459.590974	209.2	205.6	216.2	-35.7
Cellobiose	$^{0,2}A_2$	-1459.591155	197.9	197.5	199.3	-6.2
Cellobiose	$^{0,2}A_2 \Rightarrow ^{2,4}A_2$	-1459.535660	335.0	337.7	279.9	194.3
Gentiobiose	B_1-Y_1	-1459.570287	267.5	267.2	265.7	4.9
Gentiobiose	Ring-opening	-1459.582268	188.8	186.1	193.6	-25.3
Gentiobiose	$^{0,2}A_2$	-1459.596969	192.7	193.0	189.3	12.3
Gentiobiose	$^{0,3}A_2$	-1459.571078	259.2	260.6	251.5	30.4
Gentiobiose	$^{0,4}A_2$ (direct)	-1459.544823	316.3	321.5	306.8	49.2
Gentiobiose	$^{0,2}A_2 \Rightarrow ^{0,4}A_2$	-1459.545006	321.6	324.8	262.3	209.7

2.5.4 Cross Ring Bond Cleavage Transition Structures: The A_n-X_m Pathways

For [Gentiobiose+Na]⁺, we observe the $^{0,2}A_2$ peak as the most abundant peak (m/z 305, Figure 2-1b). This reaction is initiated by opening the D-Glucose ring at the anomeric center (Scheme 2-2). Transfer of the hydroxyl proton from the anomeric center (Glucose carbon 1) to the ring oxygen and ring-opening occurs in a concerted manner forming an aldehyde at carbon 1 and a new hydroxyl group at carbon 5 (196 and 189 kJ mol⁻¹ for the α and β anomers ($\Delta E_{el+ZPE,0K}$, Tables 2-2 and 2S2). These pathways are initiated from structures in which the non-reducing end is in a boat conformation. The difference in energy results from whether the proton transfer occurs on the same side of the ring (β anomers). While this interaction is deleterious, it is offset by the additional ability to stabilize the adjacent Na⁺ cation provided by the anomeric oxygen. Our calculations support a retro-aldol mechanism[18, 56, 58] as the most feasible pathway (179.4 or 192.7 kJ mol⁻¹ for the α and β anomers, Scheme 2-2, Scheme 2S4). Consistent with our experimental work, this mechanism produces a loss of 62 u, a *cis*-ethene-1,2-diol (HC(OD)=CH(OD)), from the hydroxyl deuterated form of the analyte, rather than glycolaldehyde (HC(O)-CH₂(OD), 61 u).

Scheme 2-2:

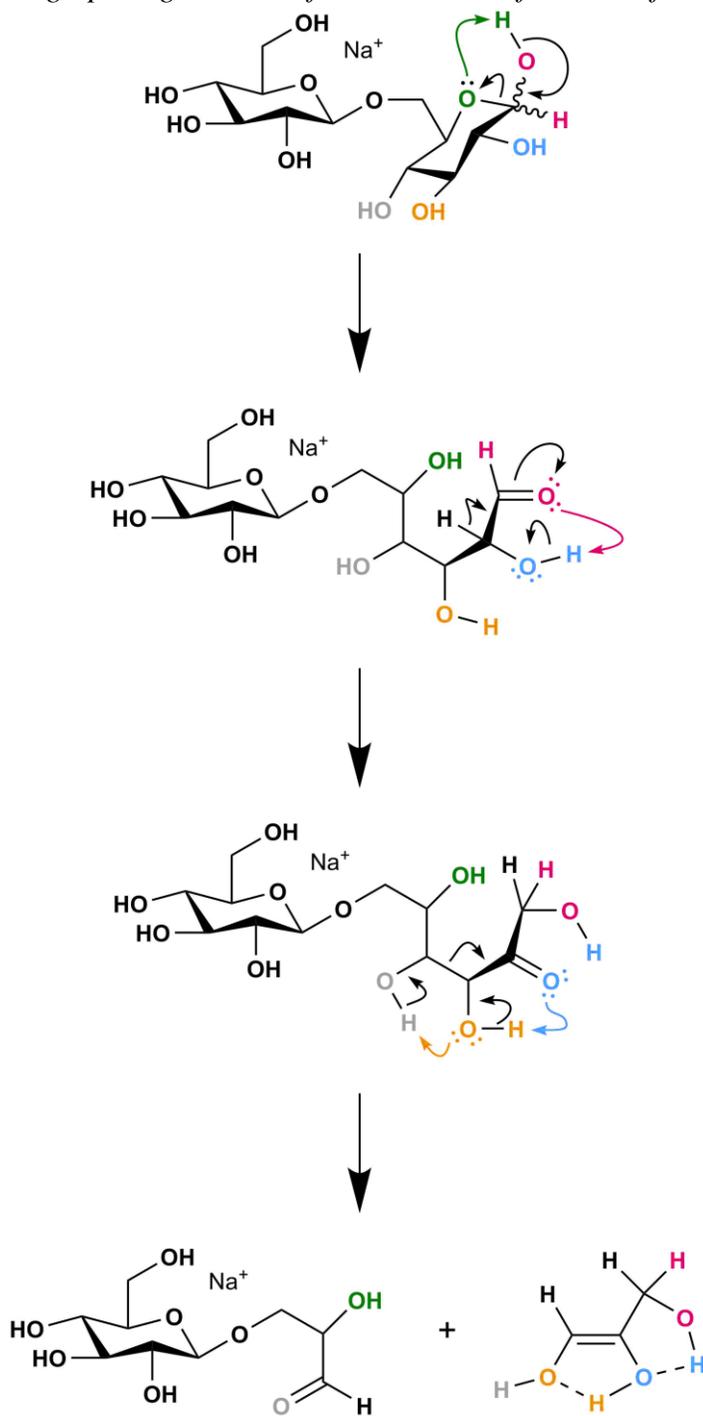
Ring-opening, then $^{0,2}A_2$ ion formation from [Gentiobiose+Na]⁺



At higher collision energies the $^{0,3}A_2$ and $^{0,4}A_2$ peaks are present. For $^{0,3}A_2$ the hydroxyl labeling data indicates a total loss of $C_3H_3D_3O_3$ (Table 2-1, Scheme 2-3, Scheme 2S5, Figure 2S6). This is consistent with a mechanism requiring a C-H shift (not D) from carbon 2 to carbon 1 at the following ring-opening, but prior to a retro-aldol reaction which simultaneously cleaves the carbon 4 to carbon 3 bond and forms $HC(OD)=C(OD)-H_2COD$. The highly strained 2-1 H shift is rate-limiting (245.8 or 259.2 kJ mol^{-1} ; α and β anomers) making the subsequent retro-aldol reaction facile ($<200 \text{ kJ mol}^{-1}$).

Scheme 2-3:

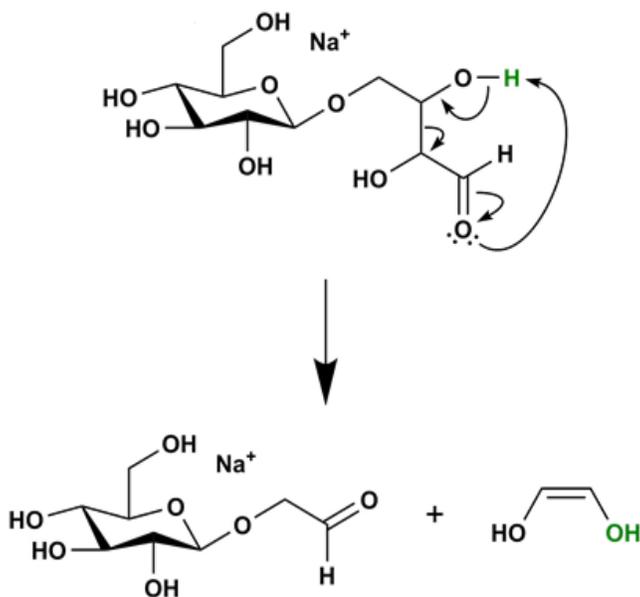
Ring-opening, 2-1 H shift, then $^{0,3}A_2$ ion formation from [Gentiobiose+Na]⁺



Formation of the $^{0,4}A_2$ ion proceeds most effectively through either an energetically demanding, structurally complex direct reaction in which two ethene-1,2-diol ($C_2H_2D_2O_2$) molecules (1 cis, 1 trans) are liberated simultaneously (302.9 or 316.3 kJ mol $^{-1}$; α and β anomers) or as a consecutive retro-aldol reaction from the already formed $^{0,2}A_2$ ion (308.2 or 321.6 kJ mol $^{-1}$; Scheme 4, Scheme S6). The consecutive retro-aldol reaction is dramatically more entropically favorable ($\Delta S_{298K} = 209.7$ J mol $^{-1}$; Table 2-2) meaning that this will overwhelmingly be the route traversed under our experimental conditions. Both pathways are consistent with the hydroxyl labeling data (Table 2-1, Scheme 2S6, Figure 2S6).

Scheme 2-4:

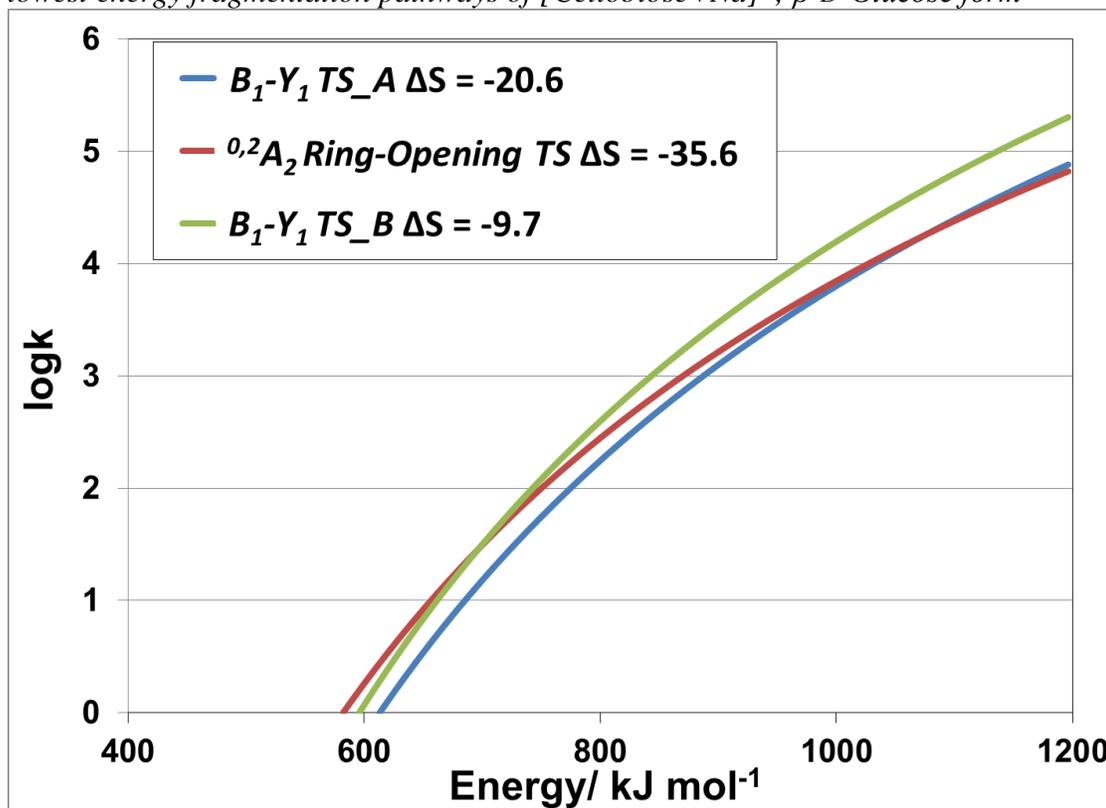
Consecutive fragmentation of the $^{0,2}A_2$ ion to form $^{0,4}A_2$ ion formation from [Gentiobiose+Na] $^+$



Formation of the $^{0,2}A_2$ ions for [Cellobiose+Na]⁺ occurs in an analogous manner (Schemes S7 and S8) to [Gentiobiose+Na]⁺. The $^{0,2}A_2$ -forming reaction is limited by ring-opening of the predominant β -anomer of Cellobiose which is strongly entropically disfavored ($\Delta S_{298K} = -35.7 \text{ J mol}^{-1}$; tight TS, Table 2-2). This enables the enthalpically more demanding, less tight B_I - Y_I transition structures (Table 2-2, B_I - Y_I TS_B , $\Delta S_{298K} = -9.7 \text{ J mol}^{-1}$ and to a lesser extent, B_I - Y_I TS_A , -20.6 J mol^{-1}) to out-compete it under our experimental conditions.⁸⁷⁻⁹⁰ This phenomenon is illustrated in Figure 2-4. The RRKM unimolecular dissociation rate calculations clearly show the B_I - Y_I TS_B pathway (green line) is more competitive than ring-opening reaction under most experimental conditions. However, once ring-opening has occurred subsequent loss of ethane-1,2-diol ($C_2H_2D_2O_2$; Table 2-1) is facile. In contrast, consecutive formation of the $^{2,4}A_2$ ion differs from the Gentiobiose forms. Although nominally the same composition as the [Gentiobiose+Na]⁺ $^{0,4}A_2$ ion, the mechanism of formation varies substantially. Our hydroxyl labeling data clearly shows a peak at m/z 250 corresponding to loss of $C_2H_2D_2O_2$ then $C_2H_3D_1O_2$ (glycolaldehyde) rather than two losses of $C_2H_2D_2O_2$ to generate a peak at m/z 249 (Scheme 2-5, Table 2-1, Figure 2S5). From the $^{0,4}A_2$ ion structure, abstraction of a hydroxyl proton (deuteron) by the aldehyde group with concerted cleavage of the bond connecting carbon 4 to carbon 5 of the original produces the $^{2,4}A_2$ ion (Schemes 2-5 and 2S9). This requires at least 335-341 kJ mol^{-1} to populate, but again is entropically very favorable ($\Delta S_{298K} = 194.3 \text{ J mol}^{-1}$; Table 2-2) due to the consecutive generation of additional gas-phase species (3 overall rather than 2 in the present case). Consistent with experiment, this pathway is likely to become active at higher collisional energies.

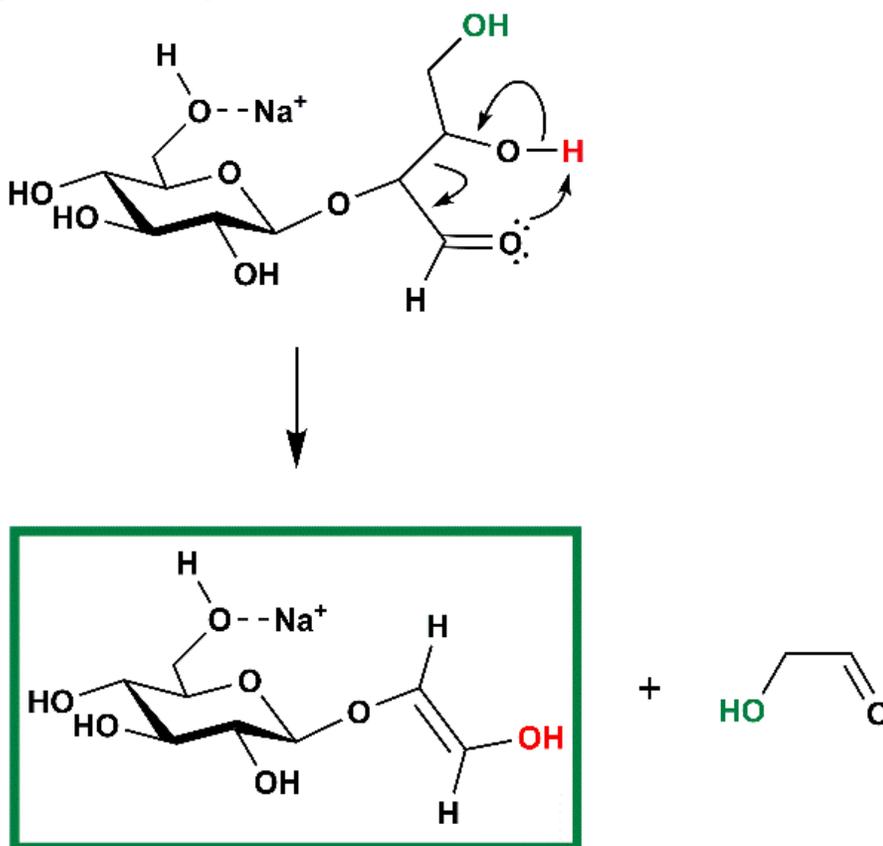
Figure 2-4:

RRKM unimolecular rate constants, k (s^{-1}) calculated at various internal energies for the lowest energy fragmentation pathways of $[\text{Cellobiose}+\text{Na}]^+$; β -D-Glucose form



Scheme 2-5:

Consecutive fragmentation of the $^{0,2}A_2$ ion to form $^{2,4}A_2$ ion and glycolaldehyde for $[\text{Cellobiose}+\text{Na}]^+$



2.5.5 Larger Basis Sets and Alternate Model Chemistries

We tested our findings with higher-level single point calculations at the M06-2X and MP2(full) levels of theory with the 6-311++G(2d,p) basis set. These results are extremely consistent with the M06-2X/6-31+G(d,p) data across both analytes, and between models too: $\Delta\hat{E}_{\text{el}+\text{ZPE},0\text{K}}$ (standard deviation): 1.5 (8.4) kJ mol^{-1} , M06-2X single points; -2.3 (7.9) kJ mol^{-1} MP2(full) single points. The pertinent numbers are available in the supporting information for the interested reader (Tables 2S2-2S5).

2.6 Conclusions

Our combined experiments and calculations indicate that isomeric sodium-cationized glycan analytes can be distinguished based on their fragmentation energetics. The 1,6-glycosidic linkage in Gentiobiose enables additional flexibility and a wider variety of energetically competitive gas-phase ion minima and transition structures (boat conformations at the reducing end). Sodiated Cellobiose strongly prefers chair conformations. Our theoretical predictions of glycosidic and cross-ring transition structure energies are consistent with the branching ratios of products and our hydroxyl labeling data.[86] i.e., the product ion branching ratios can be estimated based on sodium-ion affinities of the ion-molecule dimer's components. Our energetics and labeling data support 1,6-anhydroglucose B_1 ions rather than the previously invoked 1,2-anhydroglucose structures. 1,2-Anhydroglucose structures would lead to substantially different branching ratios (inverted). Additional experimental and theoretical evidence for retro-aldol mechanisms of formation of the $^{0,2}A_2$ peaks (and some others) is provided, consistent with the earlier literature.[18, 56, 58] Subsequent investigations will seek to determine how general the present findings are.

2.7 Acknowledgments

This work was supported by start-up funds from the University of Missouri-St. Louis. Calculations were performed at the Missouri University of Science and Technology Rolla, MO. Maha Abutokaikah thanks the Saudi Arabia Culture Mission for a graduate fellowship.

There are no conflicts of interest to declare.

2.8 References

- 1 D. Vanderschaeghe, N. Festjens, J. Delanghe and N. Callewaert, *Biol Chem.*, 2010, **391**, 149-161.
- 2 G. W. Hart and R. J. Copeland, *Cell*, 2010, **143**, 672–676.
- 3 A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Betozzi, G. W. Hart and M. E. Etzler, *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press: New York, 2008.
- 4 H. J. An, S. R. Kronewitter, M. L. A. De Leoz and C. B. Lebrilla, *Curr. opinion in chemical biology*, 2010, **13**, 601-607.
- 5 A. J. Petrescu, M. R. Wormald and R. A. Dwek, *Curr. Opin. Struct. Biol.*, 2006, **16**, 600-607.
- 6 M. Dalziel, M. Crispin, C. N. Scanlan, N. Zitzmann and R. A. Dwek, *Science*, 2014, **343**, 1235681.
- 7 W. R. Alley Jr. and M. V. Novotny, *Annu. Rev. Anal. Chem.*, 2013, **6**, 237-265.
- 8 C. N. Scanlan, J. Offer, N. Zitzmann and R. A. Dwek, *Nature*, 2007, **446**, 1038-1045.
- 9 M. Thaysen-Andersen and N. H. Packer, *Biochim Biophys Acta.*, 2014, **1844**, 1437-1452.
- 10 M. Thaysen-Andersen, N. H. Packer and B. L. Schulz, *Mol. Cell Proteomics*, 2016, **16**, 1773-1790.
- 11 J. Zaia, *Mass Spectrom. Rev.*, 2004, **23**, 161-227.

- 12 H. J. An, J. W. Froehlich and C. B. Lebrilla, *Current opinion in chemical biology*, 2010, **13**, 421-426.
- 13 Y. Park and C. B. Lebrilla, *Mass Spectrom. Rev.*, 2005, **24** 232-264.
- 14 M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla and I. J. Amster, *Anal. Chem.*, 2013, **86**, 196-212.
- 15 X. Yu, Y. Jiang, Y. Huang, C. E. Costello and C. Lin, *Anal. Chem.*, 2013, **85**, 10017-10021.
- 16 M. T. Cancilla, S. G. Penn and C. B. Lebrilla, *Anal. Chem.*, 1998, **70**, 663-672.
- 17 R. Orlando, C. A. Bush and C. Fenselau, *Biomed. Environ. Mass Spectrom.*, 1990, **19**, 747-754.
- 18 G. E. Hofmeister, Z. Zhou and J. A. Leary, *J. Am. Chem. Soc.*, 1991, **113**, 5964-5970.
- 19 V. N. Reinhold, B. B. Reinhold and C. E. Costello, *Anal. Chem.*, 1995, **67**, 1772-1784.
- 20 Y. Huang, Y. Pu, X. Yu, C. E. Costello and C. Lin, *J. AM. Soc. Mass Spectrom.*, 2016, **27**, 319-328.
- 21 H. Hu, Y. Huang, Y. Mao, X. Yu, Y. Xu, J. Liu, C. Zong, G.-J. Boons, C. Lin, Y. Xia and J. Zaia, *Mol. Cell Prot.*, 2014, **13**, 2490-2502.
- 22 X. Yu, Y. Huang, C. Lin and C. E. Costello, *Anal. Chem.*, 2012, **84**, 7487-7494.
- 23 M. A. Fentabil, R. Daneshfar, E. N. Kitova and J. S. Klassen, *J. Am. Soc. Mass Spectrom*, 2011, **22**, 2171-2178.
- 24 M. R. Asam and G. L. Glish, *J. Am. Soc. Mass Spectrom.*, 1997, **8**, 987-995.

- 25 D. J. Harvey, *Proteomics*, 2001, **1**, 311-328.
- 26 D. J. Harvey, 1999, **18**, 349-451.
- 27 D. J. Harvey, R. H. Bateman and M. R. Green, 1997, **32**, 167-187.
- 28 D. J. Ashline, A. J. Lapadula, Y.-H. Liu, M. Lin, M. Grace, B. Pramanik and V. N. Reinhold, *Anal. Chem.*, 2007, **79**, 3830-3842.
- 29 M. T. Cancilla, S. G. Penn, J. A. Carroll and C. B. Lebrilla, *J. Am. Chem. Soc.*, 1996, **118**, 6736-6745.
- 30 J. Hofmann, A. Stuckmann, M. Crispin, D. J. Harvey, K. Pagel and W. B. Struwe, *Anal. Chem.*, 2017, **89** 2318–2325.
- 31 B. Domon and C. E. Costello, *Glycoconj. J.* , 1988, **5**, 397-409.
- 32 D. J. Ashline, A. J. S. Hanneman, H. Zhang and V. N. Reinhold, *J. Am. Soc. Mass Spectrom*, 2014, **25**, 444–453.
- 33 H. Zhang, S. Singh and V. N. Reinhold, *Anal. Chem.*, 2005, **77**, 6263-6270.
- 34 K. Tseng, J. L. Hedrick and C. B. Lebrilla, *Anal. Chem.*, 1999, **71**, 3747-3754.
- 35 J. K. Eng, A. L. McCormack and J. R. Yates, *J. Am. Soc. Mass Spectrom* 1994, **5**, 976–989.
- 36 D. N. Perkins, D. J. Pappin, ;, D. M. Creasy and J. S. Cottrell, *Electrophoresis*, 1999, **20**, 3551–3567.
- 37 A. Frank and P. Pevzner, *Anal. Chem.*, 2005, **77**, 964–973.
- 38 L. Domokos, D. Hennberg and B. Weimann, *Anal. Chim. Acta*, 1984, **165**, 61-74.
- 39 <https://www.nist.gov/srd/nist-standard-reference-database-1a-v14>.

- 40 B. E. Frewen, G. E. Merrihew, C. C. Wu, W. S. Noble and M. J. MacCoss, *Anal Chem.*, 2006, **78**, 5678-5684.
- 41 A. I. Nesvizhskii, *Methods Mol Biol.*, 2007, **367**, 87-119.
- 42 H. Lam and R. Aebersold, *Methods.*, 2011 **54**, 424-431.
- 43 <http://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:glycan-library>.
- 44 S. S. Nigudkar and A. V. Demchenko, *Chem. Sci.*, 2015, **6**, 2687-2704.
- 45 A. J. Lapadula, P. J. Hatcher, A. J. Hanneman, D. J. Ashline, H. Zhang and V. N. Reinhold, *Anal. Chem.*, 2005, **77** 6271-6279.
- 46 Y. Huang and E. D. Dodds, *Anal. Chem.*, 2013, **85**, 9728-9735.
- 47 Y. Huang and E. D. Dodds, *Anal. Chem.*, 2015, **87**, 5664-5668.
- 48 M. M. Gaye, R. Kurulugama and D. E. Clemmer, *Analyst*, 2015, **140**, 6922-6932.
- 49 W. Hoffmann, J. Hofmann and K. Pagel, *J. Am. Soc. Mass Spectrom*, 2013, **25**, 471-479.
- 50 F. Zhu, M. S. Glover, J. C. Trinidad and D. E. Clemmer, *J. Am. Soc. Mass Spectrom*, 2015, **26**, 25-35.
- 51 M. Zhu, B. Bendiak, B. Clowers and H. H. Hill, *Anal Bioanal Chem.*, 2009, **394**, 1853-1867.
- 52 K. Pagel and D. J. Harvey, *Anal. Chem.*, 2013, **85**, 5138-5145.
- 53 W. B. Struwe, C. Baldauf, J. Hofmann, P. M. Rudd and K. Pagel, *Chem. Commun.*, 2016, **52**, 12353-12356
- 54 C. J. Gray, B. Schindler, L. G. Migas, M. Picmanova, A. R. Allouche, A. P. Green, S. Mandal, M. S. Motawia, R. Sánchez-Pérez, N. Bjarnholt, B. L. Møller, A. M.

- Rijs, P. E. Barran, I. Compagnon, C. E. Eyers and S. L. Flitsch, *Anal. Chem.*, 2017, **89**, 4540-4549.
- 55 C. J. Gray, B. Thomas, R. Upton, L. G. Migas, C. E. Eyers, P. Barran and S. L. Flitsch, *Biochimica et Biophysica Acta*, 2016, **1860**, 1688-1709.
- 56 B. J. Bythell, M. T. Abutokaikah, A. R. Wagoner, S. Guan and J. M. Rabus, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 688-703.
- 57 B. Paizs and S. Suhai, *Mass Spectrom. Rev.*, 2005, **24**, 508-548.
- 58 B. Spengler, J. W. Dolce and R. J. Cotter, *Anal. Chem.*, 1990, **62**, 1721-1737.
- 59 H. Suzuki, A. Kameyama, K. Tachibana, H. Narimatsu and K. Fukui, *Anal. Chem.*, 2009, **81**, 1108–1120.
- 60 T. Yamagaki, K. Fukui and K. Tachibana, *Anal. Chem.*, 2006, **78**, 1015-1022.
- 61 M. T. Rodgers and P. B. Armentrout, *Chem. Rev.*, 2016, **116**, 5642-5687.
- 62 R. R. Wu and M. T. Rodgers, *J. Phys. Chem. B*, 2016, **120**, 4803-4811.
- 63 Y. Zhu, L. A. Hamlow, C. C. He, S. F. Strobehn, J. K. Lee, J. Gao, G. Berden, J. Oomens and M. T. Rodgers, *J. Phys. Chem. B*, 2016, **120**, 8892-8904.
- 64 J.-L. Chen, H. S. Nguan, P.-J. Hsu, S.-T. Tsai, C. Y. Liew, J.-L. Kuo, W.-P. Hu and C.-K. Ni, *Phys.Chem.Chem.Phys.*, 2017, **19**, 15454.
- 65 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comp. Chem.*, 2005, **26**, 1781-1802.
- 66 S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859-1865
- 67 B. R. Brooks, C. L. Brooks, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R.

- Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545-1614
- 68 J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks III, A. D. MacKerell, J. B. Klauda and W. Im, *J. Chem. Theory Comput.*, 2016, **12**, 405-413.
- 69 S. Jo, K. C. Song, H. Desaire, A. D. Mackerell and W. Im, *J. Comput. Chem.*, 2011, **32**, 3135-3141.
- 70 S. Kim, J. Lee, S. Jo, C. L. Brooks, H. S. Lee and W. Im, *J. Comput. Chem.*, 2017, in press.
- 71 J. J. P. Stewart, *J. Mol. Model.*, 2007, **13**, 1173-1213.
- 72 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Phys.*, 2005, **123** (16), 161103.
- 73 Y. Zhao and D. G. Truhlar, *Theor Chem Account*, 2006, **120**, 215.
- 74 C. L. Moss, J. Chamot-Rooke, E. Nicol, J. Brown, I. Campuzano, K. Richardson, J. P. Williams, M. F. Bush, B. Bythell, B. Paizs and F. Turecek, *J. Phys. Chem. B*, 2012, **116**, 3445-3456.
- 75 F. Turecek, T. W. Chung, C. L. Moss, J. A. Wyer, A. Ehlerding, A. I. S. Holm, H. Zettergren, S. B. Nielsen, P. Hvelplund, J. Chamot-Rooke, B. Bythell and B. Paizs, *J. Am. Chem. Soc.*, 2010, **132**, 10728-10740.

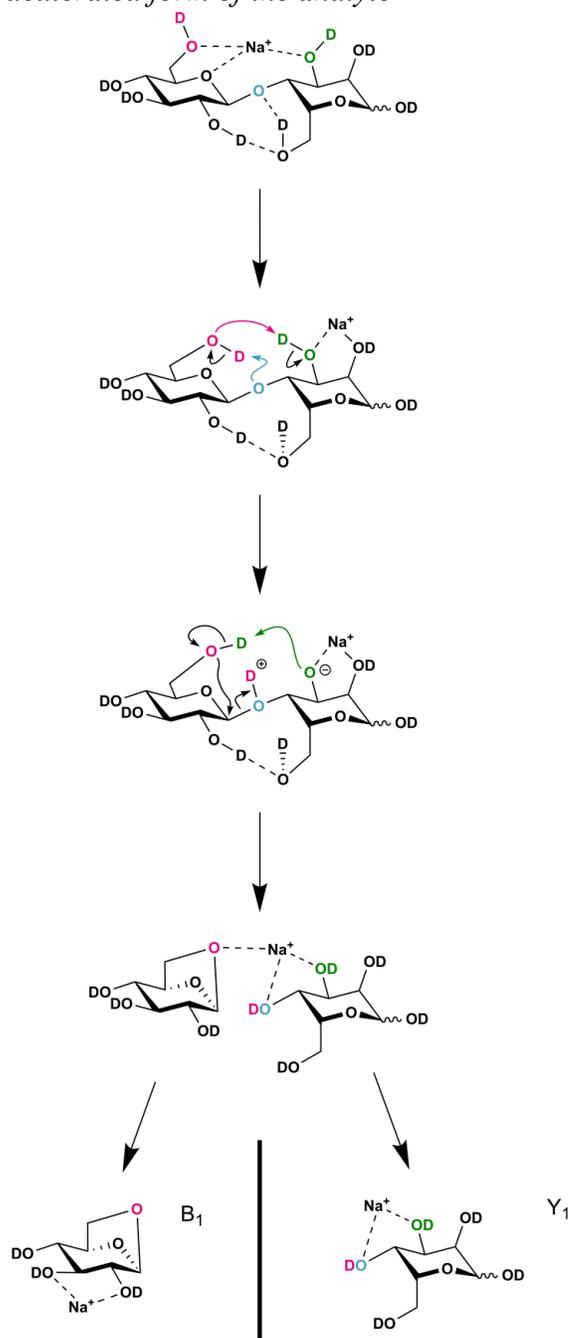
- 76 B. J. Bythell, S. Suhai, A. Somogyi and B. Paizs, *J. Am. Chem. Soc.*, 2009, **131**, 14057-14065.
- 77 T. Wyttenbach, B. Paizs, P. Barran, L. Brecci, D. Liu, S. Suhai, V. H. Wysocki and M. T. Bowers, *J. Am. Chem. Soc.*, 2003, **125**, 13768-13775.
- 78 M. J. T. Frisch, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox *Gaussian 09, Revision C.01*, Gaussian, Inc., Wallingford CT, 2010.
- 79 C. Lifshitz, *Chem. Soc. Rev.*, 2001, **30**, 186–192.
- 80 R. G. Gilbert and S. C. Smith, *Theory of Unimolecular and Recombination Reactions*, Blackwell Scientific Publications. , 1990.
- 81 T. Beyer and D. R. Swinehart, *ACM Commun.* , 1973, **16** 379.

- 82 M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, *J. Chem. Theory Comput.*, 2016, **12**, 6157–6168.
- 83 Y. Tan, N. Zhao, J. Liu, P. Li, C. N. Stedwell, L. Yu and N. C. Polfer, *J. Am. Soc. Mass Spectrom.*, , 2017, DOI:10.1007/s13361-13016-11575-x.
- 84 A. L. Heaton and P. B. Armentrout, *J. Phys. Chem. A*, 2008, **112**, 10156–10167.
- 85 B. A. Cerda and C. Wesdemiotis, *Int.J. Mass Spectrom.*, 1999 **189**, 189–204.
- 86 A. G. Harrison, *J. Mass Spectrom.*, 1999, **34**, 577–589.
- 87 K. Vékey, *J. Mass Spectrom.*, 1996, **31**, 445–463.
- 88 D. Carl, B. K. Chatterjee and P. B. Armentrout, *J. Chem. Phys.*, 2010, **132**, 044303.
- 89 A. L. Heaton and P. B. Armentrout, *J. Am. Chem. Soc.*, 2008, 10227–10232.
- 90 E. Rossich-Molina, A. Eizaguirre, V. Haldys, D. Urban, G. Doisneau, Y. Bourdreux, J.-M. Beau, J.-Y. Salpin and R. Spezia, *ChemPhysChem*, 2017, **18**, 1 – 13.

2.9 Supporting Information

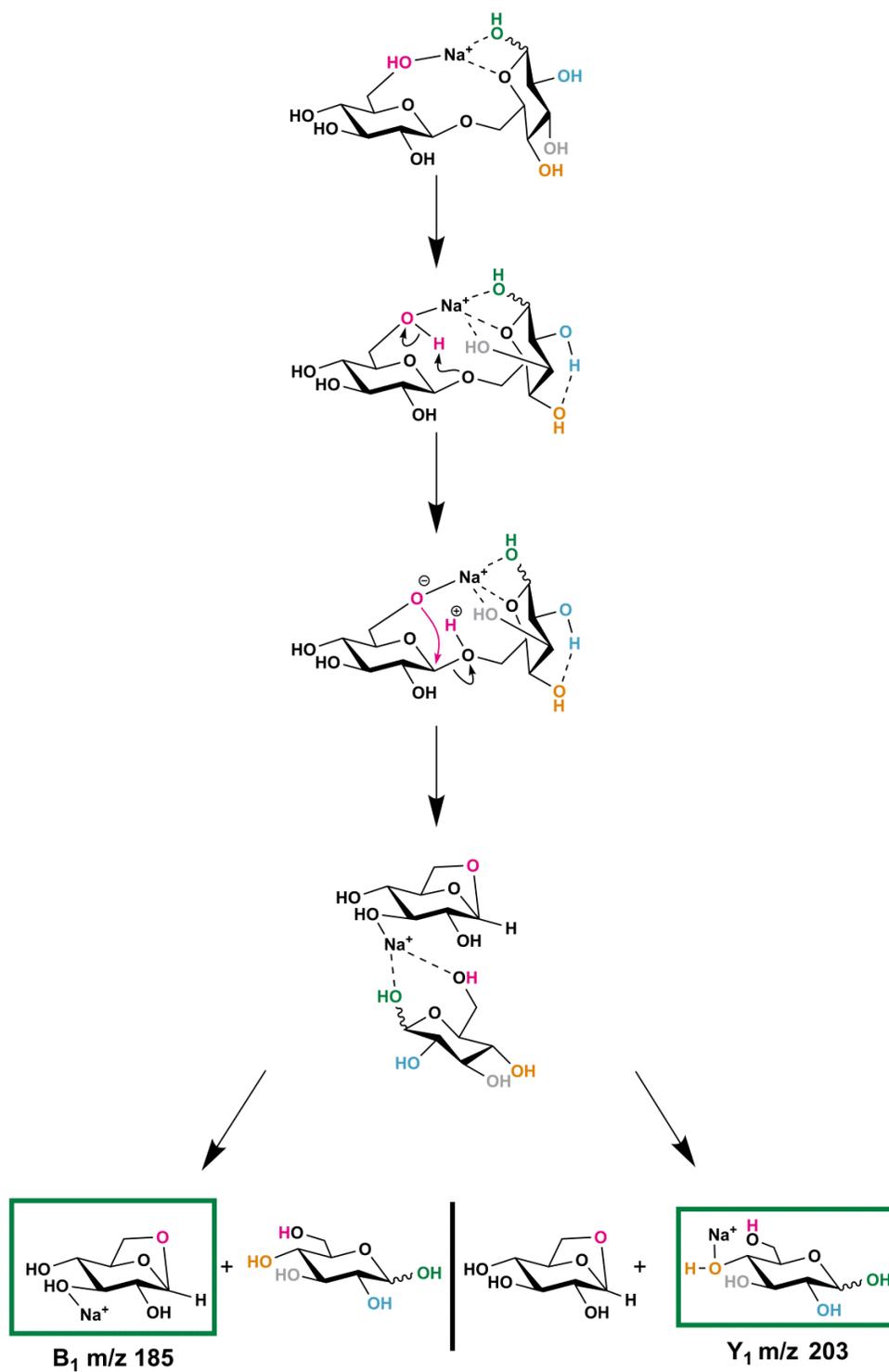
Scheme 2S1:

The lowest energy B₁-Y₁ pathway for [Cellobiose+Na]⁺ illustrated for the hydroxyl deuterated form of the analyte



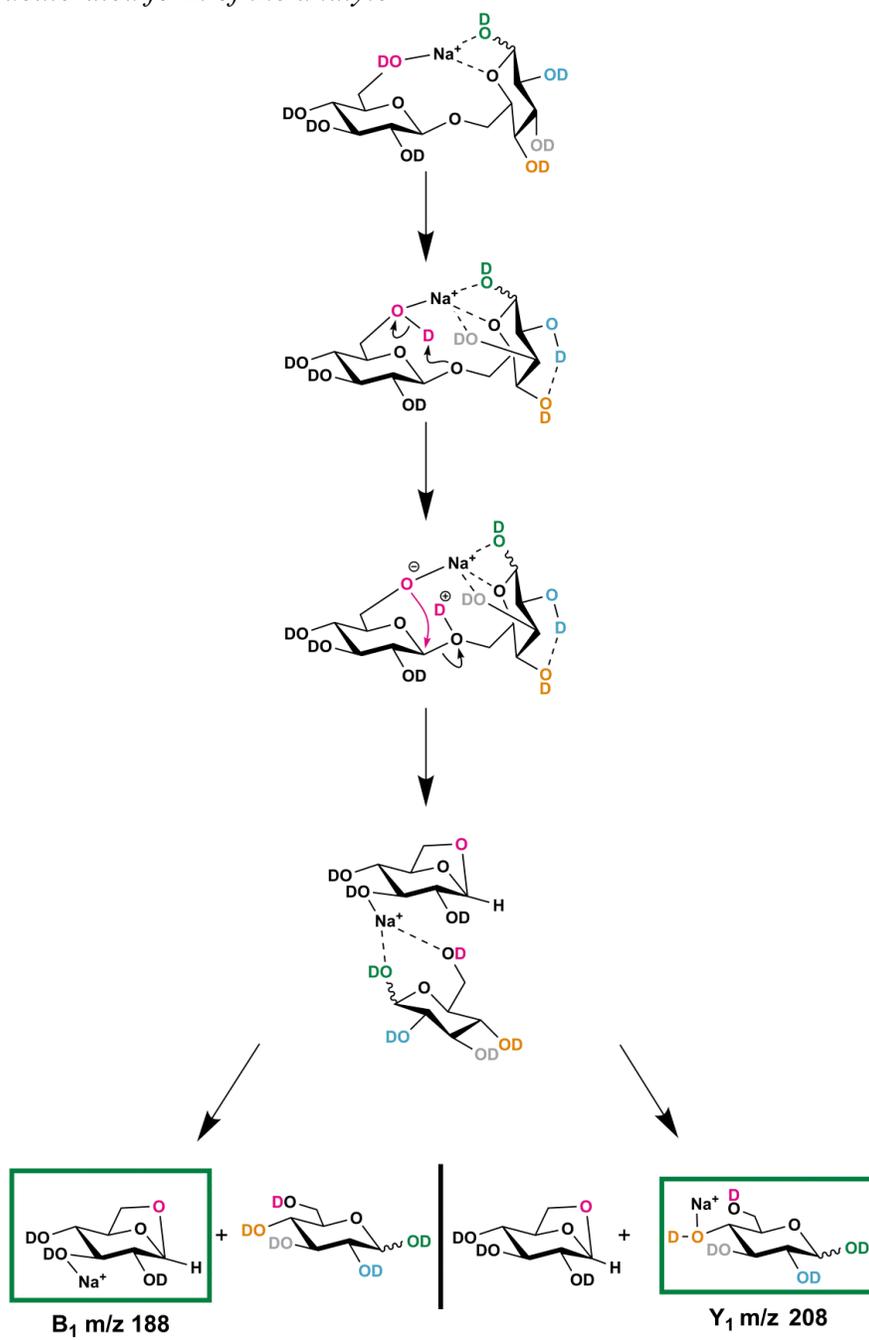
Scheme 2S2:

The lowest energy B₁-Y₁ pathway for [Gentiobiose+Na]⁺



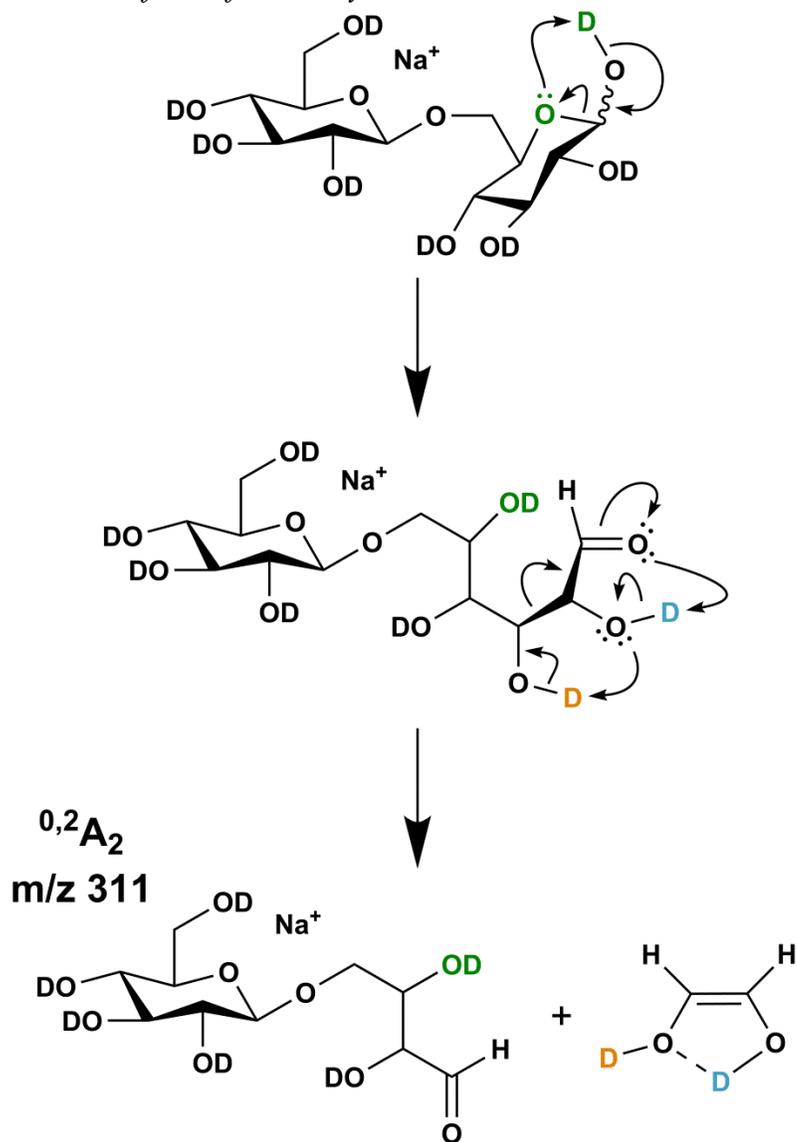
Scheme 2S3:

The lowest energy B₁-Y₁ pathway for [Gentiobiose+Na]⁺ illustrated for the hydroxyl deuterated form of the analyte



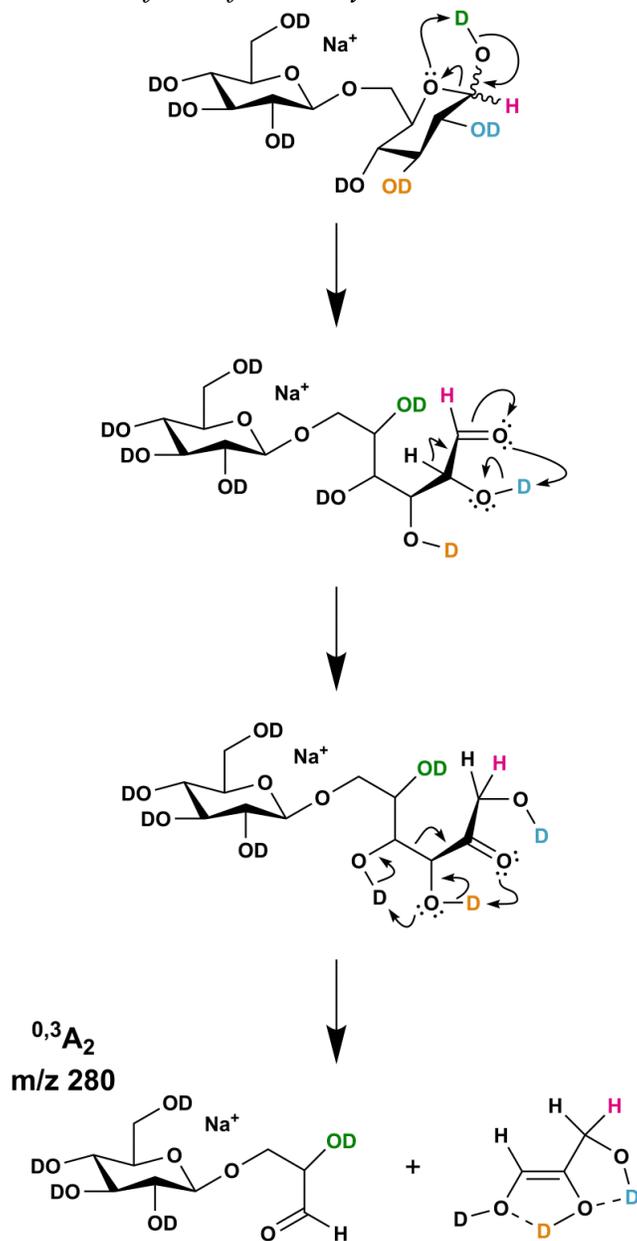
Scheme 2S4:

The lowest energy $^{0,2}A_2-X_0$ pathway for $[Gentiobiose+Na]^+$ illustrated for the hydroxyl deuterated form of the analyte



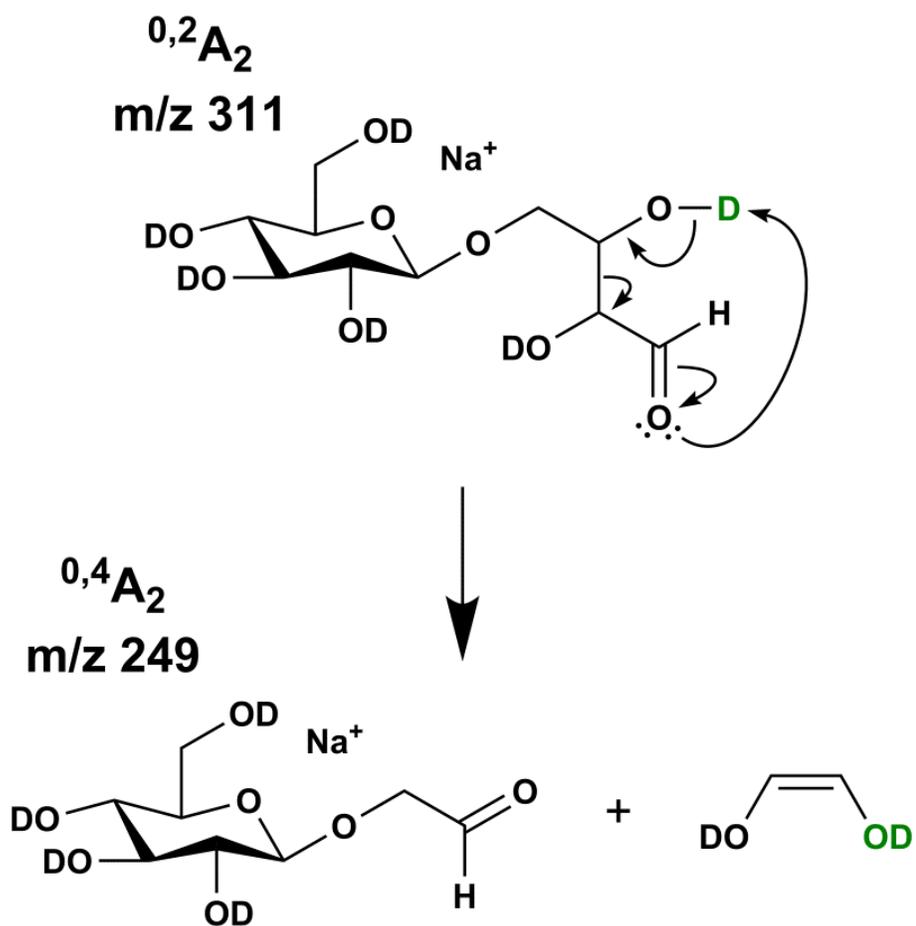
Scheme 2S5:

The lowest energy $^{0,3}A_2$ - X_0 pathway for [Gentiobiose+Na]⁺ illustrated for the hydroxyl deuterated form of the analyte



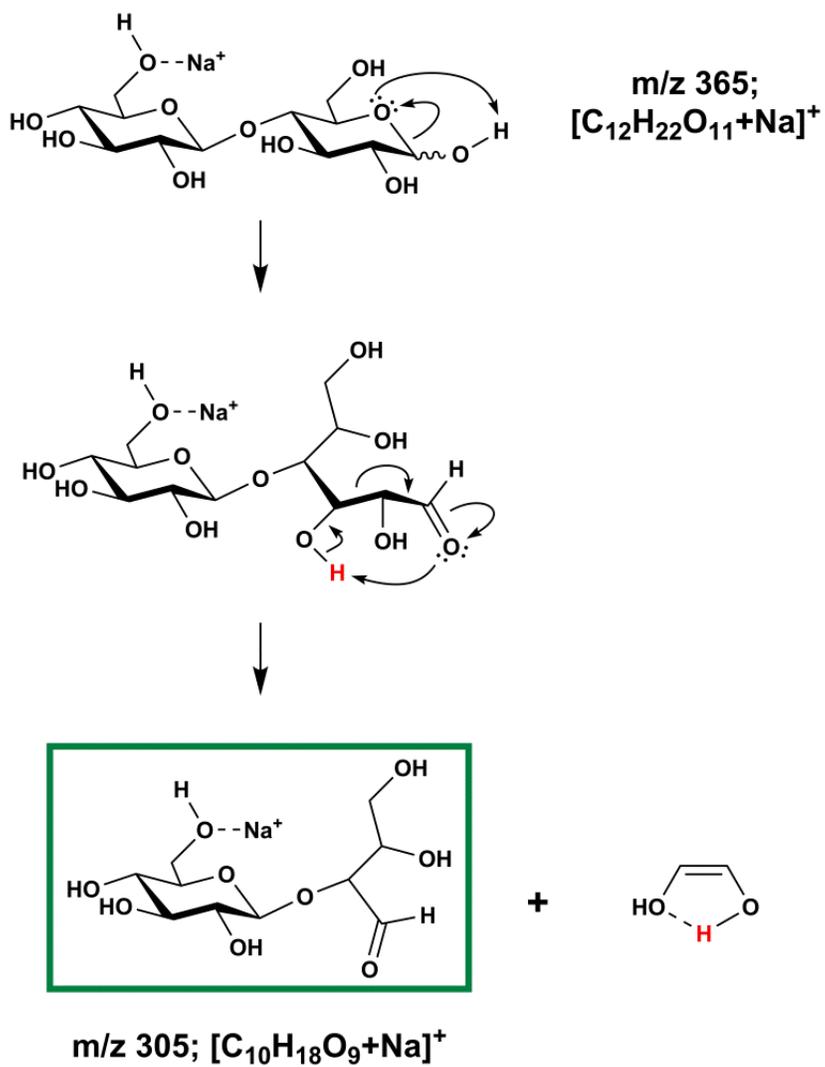
Scheme 2S6:

The $^{0,2}A_2 \rightarrow ^{0,4}A_2$ -formation pathway for $[Gentiobiose+Na]^+$ illustrated for the hydroxyl deuterated form of the analyte



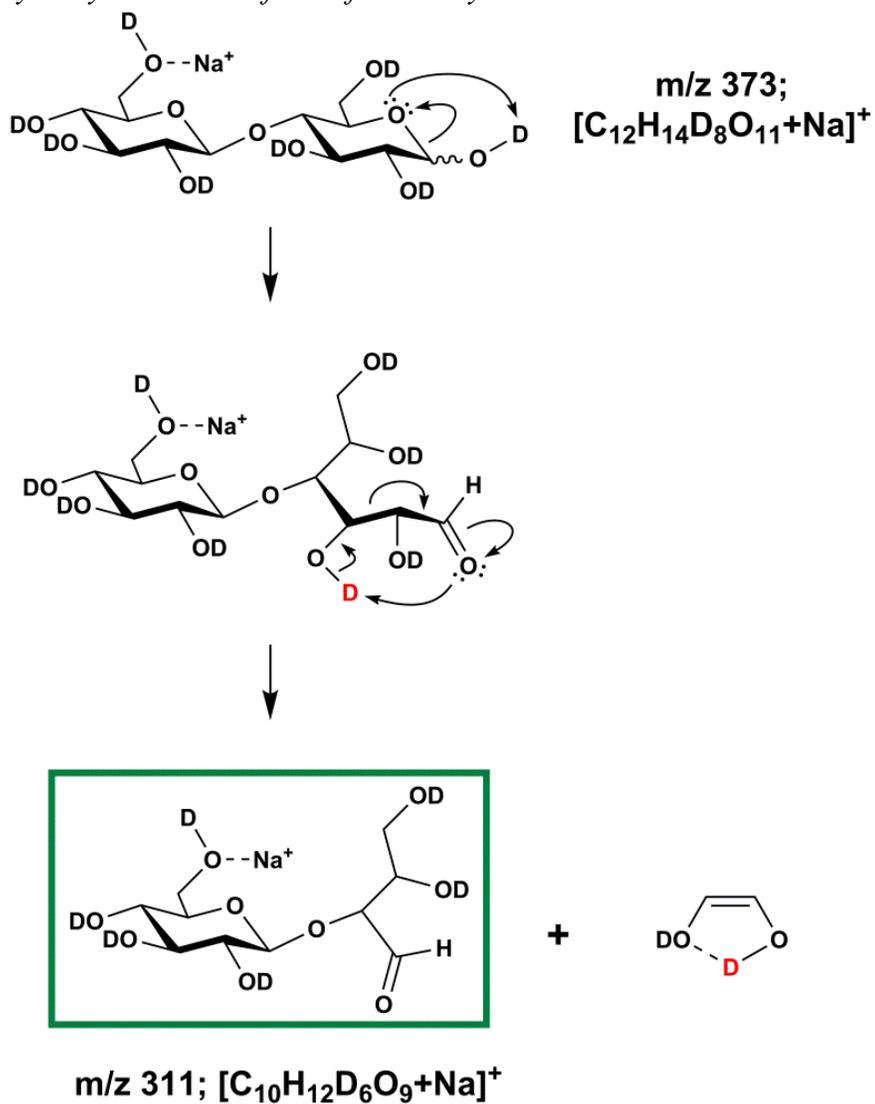
Scheme 2S7:

The lowest energy $^{0,2}A_2$ -forming pathway for [Cellobiose+Na]⁺



Scheme 2S8:

The lowest energy $^{0,2}A_2$ -forming pathway for $[\text{Cellobiose}+\text{Na}]^+$ illustrated for the hydroxyl deuterated form of the analyte



Scheme 2S9:

The lowest energy $^{0,2}A_2 \rightarrow ^{2,4}A_2$ -formation for $[Cellobiose+Na]^+$ illustrated for the hydroxyl deuterated form of the analyte. Loss of glycolaldehyde includes a single deuterium corresponding to a loss of 61 u. This is in contrast with the deuterated Gentiobiose $^{0,2}A_2$ ion (Scheme 2-6, Table 2-1) which loses 62 u to form the $^{2,4}A_2$ ion (m/z 249) and *cis*-ethene-1,2-diol; $HC(OD)=CH(OD)$

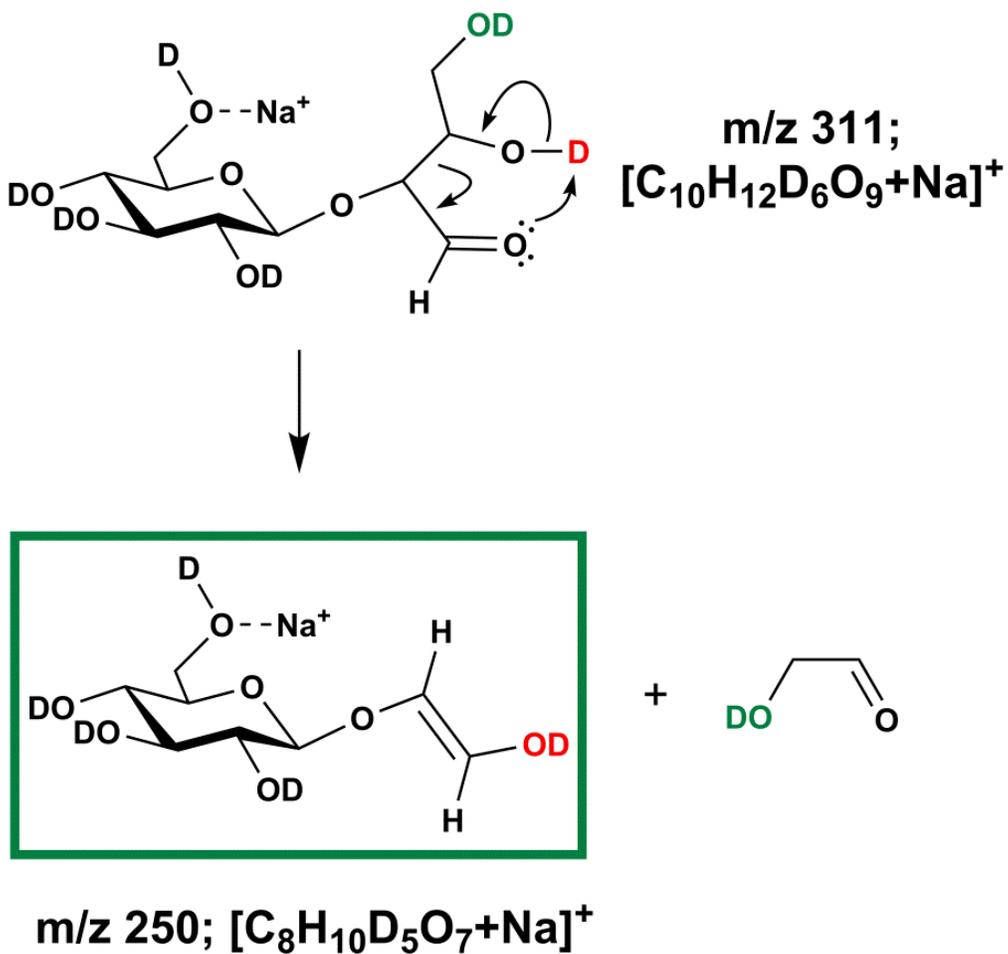


Table 2S1:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- α -D-Glucose) & Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- α -D-Glucose) forms calculated at the M06-2X/6-31+G(d,p) level of theory

Analyte	Transition Structure	E _{el} /H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298K}/$ kJ mol ⁻¹	$\Delta G_{298K}/$ kJ mol ⁻¹	$\Delta S_{298K}/$ J mol ⁻¹
Cellobiose	<i>B₁-Y₁ TS_A</i>	-1459.585487	229.3	228.4	227.9	1.4
Cellobiose	<i>B₁-Y₁ TS_B</i>	-1459.584806	229.5	227.7	231.9	-14.0
Cellobiose	Ring-opening	-1459.590974	214.9	215.7	211.7	13.6
Cellobiose	^{0,2} A ₂	-1459.591155	204.1	204.7	203.4	4.4
Cellobiose	^{0,2} A ₂ => ^{2,4} A ₂	-1459.535660	341.3	344.9	283.9	204.8
Gentiobiose	<i>B₁-Y₁</i>	-1459.565898	267.4	267.3	263.3	13.3
Gentiobiose	Ring-opening	-1459.591961	196.4	195.9	195.7	0.8
Gentiobiose	^{0,2} A ₂	-1459.596969	179.4	180.5	174.1	21.5
Gentiobiose	^{0,3} A ₂	-1459.571078	245.8	248.1	236.3	39.6
Gentiobiose	^{0,4} A ₂ (direct)	-1459.544823	302.9	309.0	291.6	58.4
Gentiobiose	^{0,2} A ₂ => ^{0,4} A ₂	-1459.545006	308.2	312.3	247.1	218.9

Table 2S2:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- β -D-Glucose) & Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- β -D-Glucose) forms calculated at the M06-2X/6-311++G(2d,p)//M06-2X/6-31+G(d,p) level of theory

Analyte	Transition Structure	E _{el} /H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298K}/$ kJ mol ⁻¹	$\Delta G_{298K}/$ kJ mol ⁻¹	$\Delta S_{298K}/$ J K ⁻¹ mol ⁻¹
Cellobiose	<i>B₁-Y₁ TS_A</i>	-1459.962526	225.4	222.9	229.0	-20.6
Cellobiose	<i>B₁-Y₁ TS_B</i>	-1459.962334	228.2	226.4	229.3	-9.7
Cellobiose	Ring-opening <i>^{0,2}A₂</i>	-1459.969094	211.8	208.2	218.8	-35.7
Cellobiose	<i>^{0,2}A₂ => ^{2,4}A₂</i>	-1459.919256	323.2	325.9	268.0	194.3
<hr/>						
Gentiobiose	<i>B₁-Y₁</i>	-1459.951330	262.7	262.4	261.0	4.9
Gentiobiose	Ring-opening <i>^{0,2}A₂</i>	-1459.978358	193.5	190.8	198.3	-25.3
Gentiobiose	<i>^{0,3}A₂</i>	-1459.953202	251.2	252.9	243.9	30.4
Gentiobiose	<i>^{0,4}A₂ (direct)</i>	-1459.926336	310.3	315.5	300.8	49.2
Gentiobiose	<i>^{0,2}A₂ => ^{0,4}A₂</i>	-1459.92843	310.6	311.8	251.2	209.7

Table 2S3:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- α -D-Glucose) and Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- α -D-Glucose) forms calculated at the M06-2X/6-311++G(2d,p)//M06-2X/6-31+G(d,p) level of theory

Analyte	Transition Structure	E _{el} /H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298K}/$ kJ mol ⁻¹	$\Delta G_{298K}/$ kJ mol ⁻¹	$\Delta S_{298K}/$ J K ⁻¹ mol ⁻¹
Cellobiose	<i>B₁-Y₁ TS_A</i>	-1459.965384	226.5	225.6	225.1	1.4
Cellobiose	<i>B₁-Y₁ TS_B</i>	-1459.962334	233.6	232.4	232.5	-14.0
Cellobiose	Ring-opening <i>^{0,2}A₂</i>	-1459.969094	228.0	227.3	223.9	13.6
Cellobiose	<i>^{0,2}A₂ => ^{2,4}A₂</i>	-1459.919256	352.9	351.3	352.5	204.8
Gentiobiose	<i>B₁-Y₁</i>	-1459.947344	269.6	268.7	265.6	13.3
Gentiobiose	Ring-opening	-1459.970425	209.0	208.1	205.1	0.8
Gentiobiose	<i>^{0,3}A₂</i>	-1459.953202	236.4	238.7	226.9	39.6
Gentiobiose	<i>^{0,4}A₂ (direct)</i>	-1459.926336	295.1	301.2	283.8	58.4
Gentiobiose	<i>^{0,2}A₂ => ^{0,4}A₂</i>	-1459.92843	295.4	299.5	234.2	218.9

Table 2S4:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- β -D-Glucose) & Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- β -D-Glucose) forms calculated at the MP2(Full)/6-311++G(2d,p)//M06-2X/6-31+G(d,p) level of theory

Analyte	Transition Structure	E_{el}/H	$\Delta E_{el+ZPE,0K}/$ $kJ\ mol^{-1}$	$\Delta H_{298K}/$ $kJ\ mol^{-1}$	$\Delta G_{298K}/$ $kJ\ mol^{-1}$	$\Delta S_{298K}/$ $J\ K^{-1}\ mol^{-1}$
Cellobiose	$B_1-Y_1\ TS_A$	-1457.576976	227.4	224.9	231.1	-20.6
Cellobiose	$B_1-Y_1\ TS_B$	-1457.576802	230.2	228.4	231.4	-9.7
Cellobiose	Ring-opening $^{0,2}A_2$	-1457.58682	205.2	201.6	212.3	-35.7
Cellobiose	$^{0,2}A_2 = >^{2,4}A_2$	-1457.528782	338.2	340.8	282.9	194.3
<hr/>						
Gentiobiose	B_1-Y_1	-1457.565286	264.5	264.1	262.7	4.9
Gentiobiose	Ring-opening $^{0,2}A_2$	-1457.594017	190.8	188.1	195.6	-25.3
Gentiobiose	$^{0,3}A_2$	-1457.567628	252.1	253.4	244.4	30.4
Gentiobiose	$^{0,4}A_2$ (direct)	-1457.534519	327.1	332.4	317.7	49.2
Gentiobiose	$^{0,2}A_2 = >^{0,4}A_2$	-1457.537467	325.2	328.4	265.9	209.7

Table 2S5:

Relative Energies of the Transition Structures of sodiated Cellobiose (β -D-glucopyranosyl-(1 \rightarrow 4)- α -D-Glucose) & Gentiobiose (β -D-glucopyranosyl-(1 \rightarrow 6)- α -D-Glucose) forms calculated at the MP2(Full)/6-311++G(2d,p)//M06-2X/6-31+G(d,p) level of theory

Analyte	Transition Structure	E _{el} /H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298K}/$ kJ mol ⁻¹	$\Delta G_{298K}/$ kJ mol ⁻¹	$\Delta S_{298K}/$ J K ⁻¹ mol ⁻¹
Cellobiose	<i>B₁-Y₁ TS_A</i>	-1457.579963	227.6	226.7	226.2	1.4
Cellobiose	<i>B₁-Y₁ TS_B</i>	-1457.578078	236.0	2334.7	234.8	-14.0
Cellobiose	Ring-opening <i>^{0,2}A₂</i>	-1457.585463	219.6	218.9	215.5	13.6
Cellobiose	<i>^{0,2}A₂ => ^{2,4}A₂</i>	-1457.528782	367.3	365.7	366.8	204.8
Gentiobiose	<i>B₁-Y₁</i>	-1457.56123	269.3	268.3	265.3	13.3
Gentiobiose	Ring-opening <i>^{0,2}A₂</i>	-1457.588121	198.7	197.7	194.7	0.8
Gentiobiose	<i>^{0,3}A₂</i>	-1457.567628	234.6	236.9	225.1	39.6
Gentiobiose	<i>^{0,4}A₂ (direct)</i>	-1457.534519	309.7	315.8	298.4	58.4
Gentiobiose	<i>^{0,2}A₂ => ^{0,4}A₂</i>	-1457.537467	307.8	309.4	311.9	246.6

Figure 2S1:

Relative proportions of peak current as a function of collision energy for [Cellobiose+Na]⁺. Ion current reduction due to direct sodium loss (and/or scattering) beyond laboratory collision energies of 46 eV was too large to provide meaningful additional data points. See Chen et al., (Phys.Chem.Chem.Phys., 2017, 19, 15454) for specifics on the energetics of desodiation using [glucose+Na]⁺

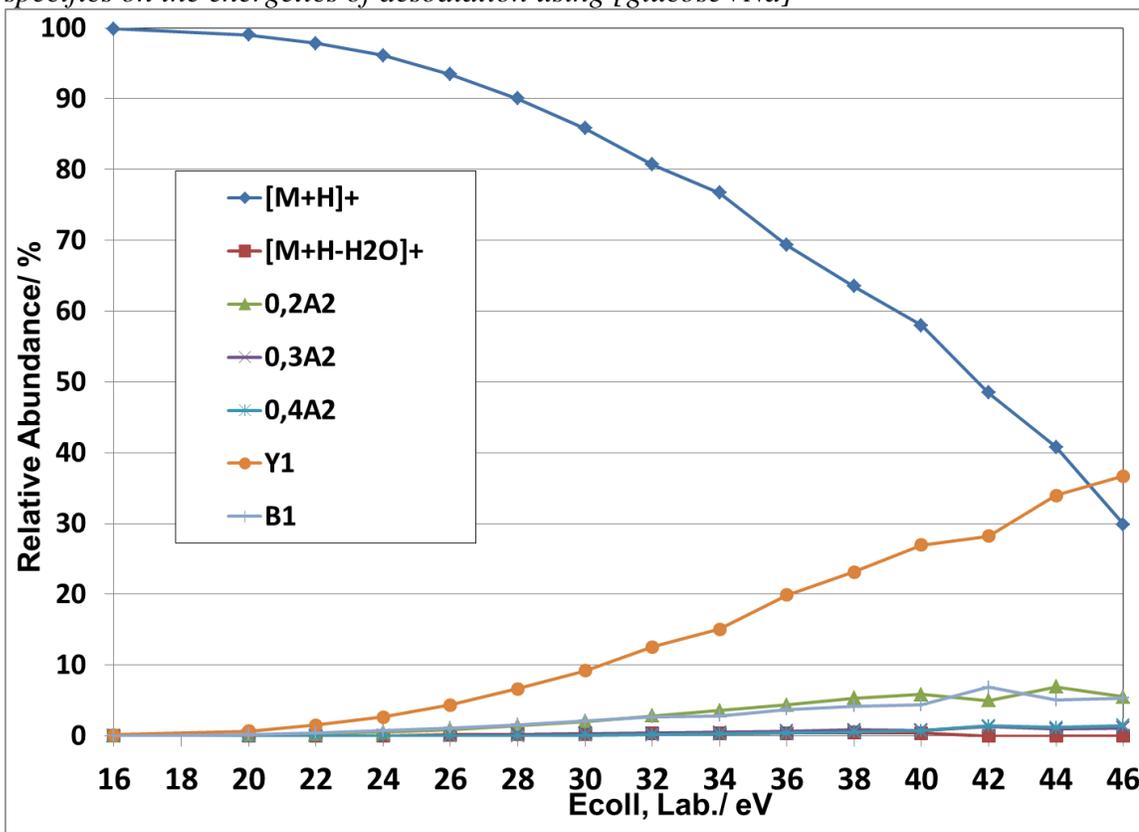


Figure 2S2:

Relative proportions (with adjusted scale) of peak current as a function of collision energy for $[\text{Cellobiose}+\text{Na}]^+$. Ion current reduction due to direct sodium loss (and/or scattering) beyond laboratory collision energies of 46 eV was too large to provide meaningful additional data points. See Chen et al., (*Phys.Chem.Chem.Phys.*, 2017, **19**, 15454) for specifics on the energetics of desodiation using $[\text{glucose}+\text{Na}]^+$

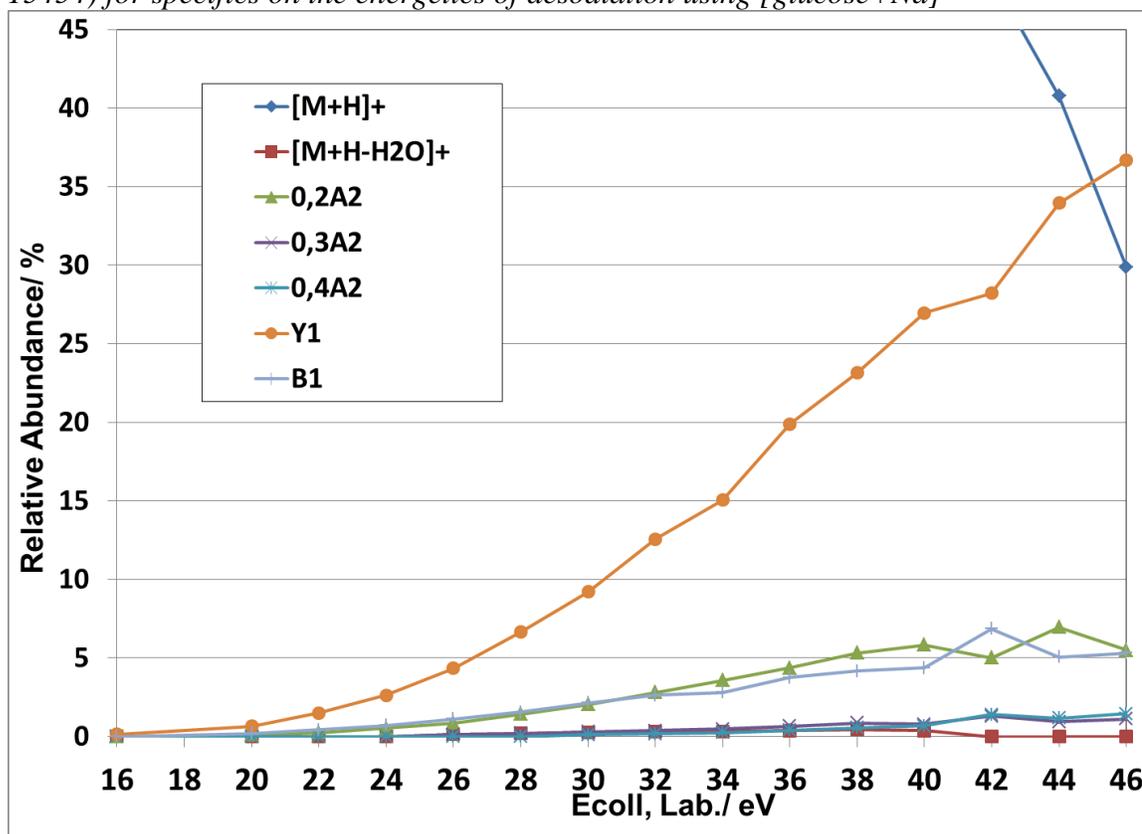


Figure 2S3:

Relative proportions of peak current as a function of collision energy for $[\text{Gentiobiose}+\text{Na}]^+$. Ion current reduction due to direct sodium loss (and/or scattering) beyond laboratory collision energies of 46 eV was too large to provide meaningful additional data points. See Chen et al., (*Phys.Chem.Chem.Phys.*, 2017, **19**, 15454) for specifics on the energetics of desodiation using $[\text{glucose}+\text{Na}]^+$

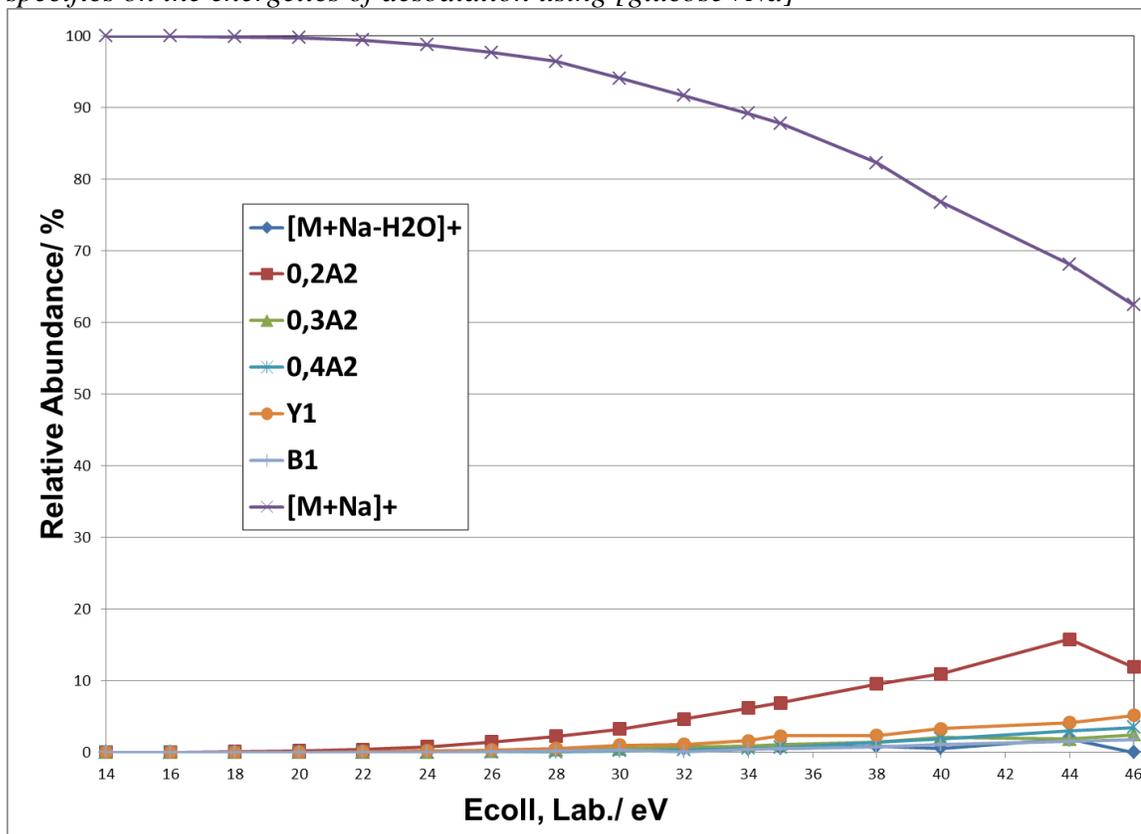


Figure 2S4:

Relative proportions of peak current (excluding precursor, see Figure 2S3) as a function of collision energy for $[Gentiobiose+Na]^+$. Ion current reduction due to direct sodium loss (and/or scattering) beyond laboratory collision energies of 46 eV was too large to provide meaningful additional data points. See Chen et al., (*Phys.Chem.Chem.Phys.*, 2017, **19**, 15454) for specifics on the energetics of desodiation using $[glucose+Na]^+$

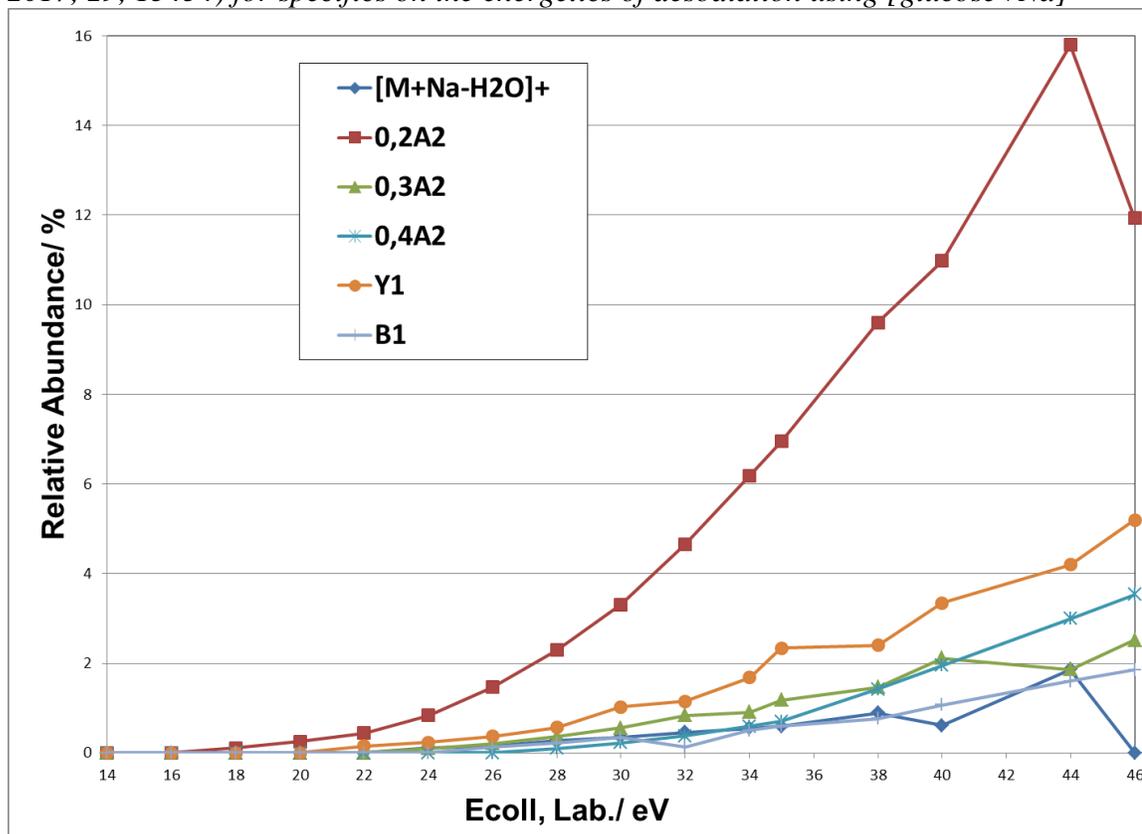


Figure 2S5:

Example MS/MS spectra of sodiated deuterated hydroxyl Cellobiose, $[D_8\text{-Cellobiose}+\text{Na}]^+$

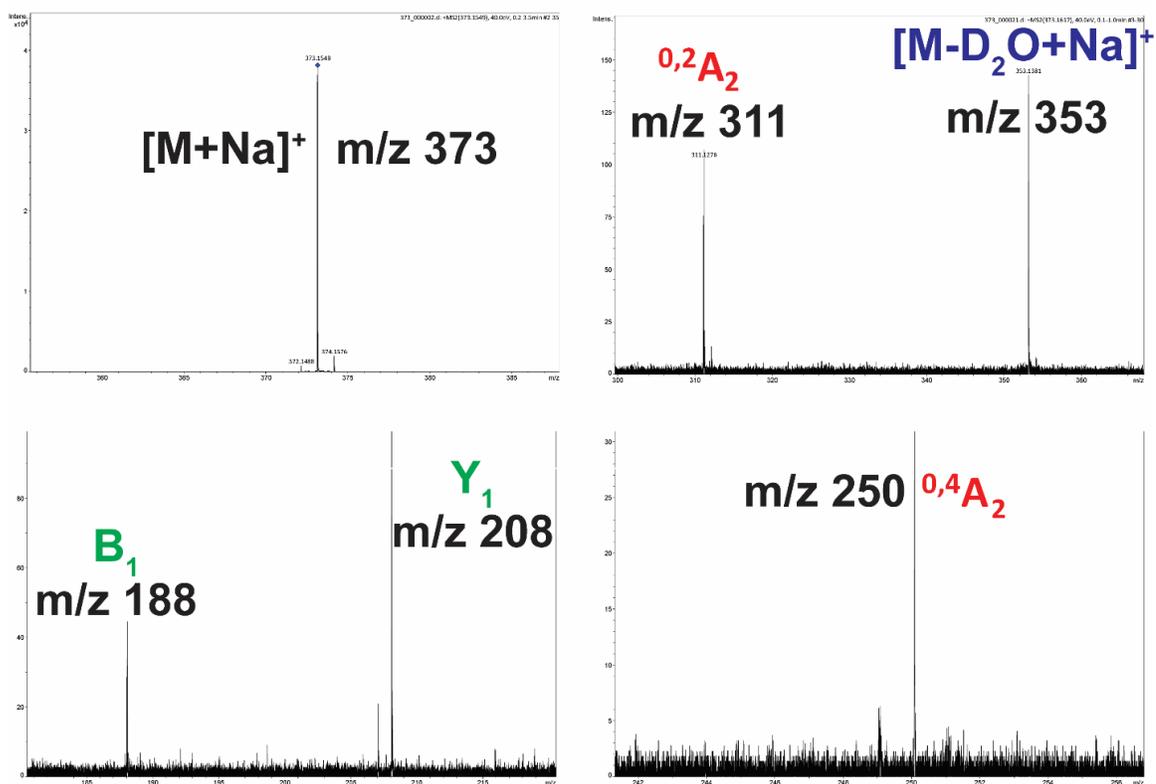
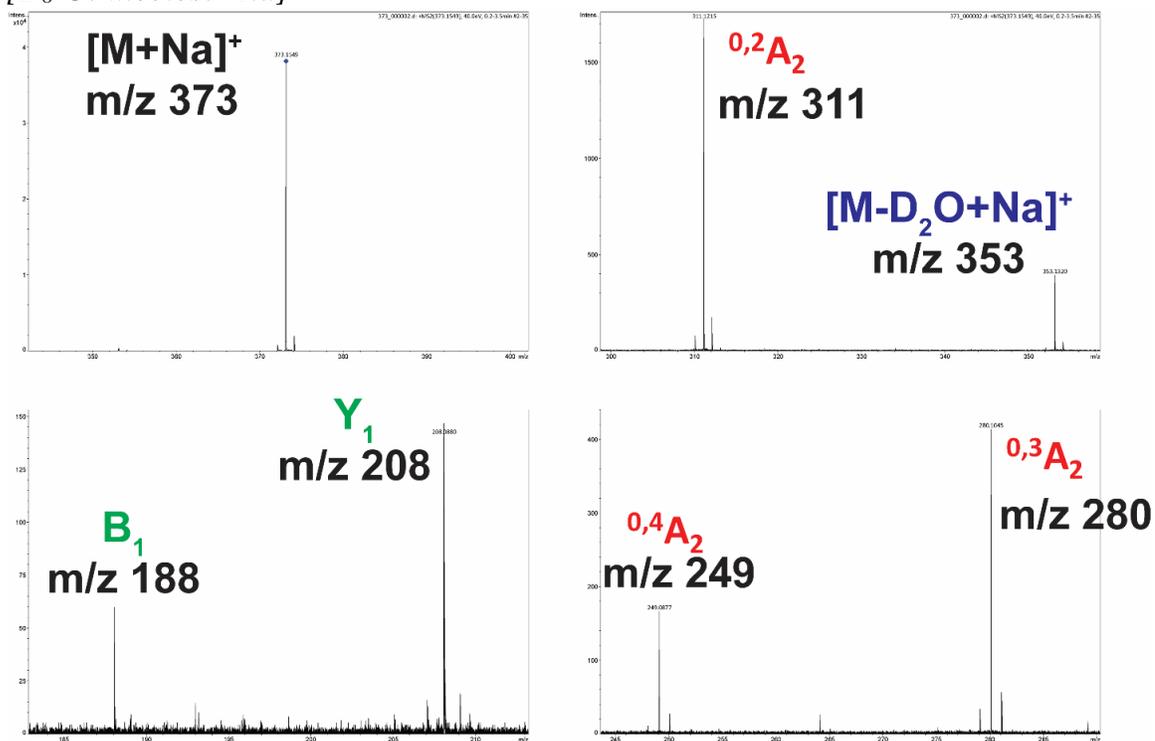


Figure 2S6:

Example MS/MS spectra of hydroxyl deuterated, sodiated Gentiobiose, $[D_8\text{-Gentiobiose}+\text{Na}]^+$



Additional Instrumentation specifics for sodiated Cellobiose and Gentiobiose MS

collection:

Collision energy: 2-50 eV in 2 eV steps

Collision cell RF: 300 Vpp

Transfer Time 80 μ s for both.

PrePulseStorage 5 μ s.

Source: ESI.

Capillary: 3400 V

End Plate Offset: -500 V

Dry Gas (N₂): 0.4 Bar, 4 L/min, 200 °C

Transfer:

Funnel/Multipole RF: 250 Vpp

**Chapter 3 : Deprotonated Carbohydrate Anion Fragmentation Chemistry:
Structural Evidence from Tandem Mass Spectrometry, Infra-Red Spectroscopy,
and Theory**

Jordan M. Rabus¹, Daniel R. Simmons¹, Philippe Maître², Benjamin J. Bythell¹

¹Department of Chemistry and Biochemistry, University of Missouri, St. Louis, MO
63121, USA.

²Laboratoire de Chimie Physique (UMR8000), CNRS, Univ. Paris-Sud, Université Paris-
Saclay, 91405, Orsay, France.

Reproduced from Physical Chemistry Chemical Physics **2018**, 20 (44), 27897–27909

with permission from the PCCP Owner Societies.

<https://doi.org/10.1039/C8CP02620C>

3.1 Abstract

We investigate the gas-phase structures and fragmentation chemistry of deprotonated carbohydrate anions using combined tandem mass spectrometry, infrared spectroscopy, regioselective labeling, and theory. Our model system is deprotonated, [lactose-H]⁻. We computationally characterize the rate-determining barriers to glycosidic bond (C_1 - Z_1 reactions) and cross-ring cleavages and compare these predictions to our tandem mass spectrometric and infrared spectroscopy data. The glycosidic bond cleavage product data support complex mixtures of anion structures in both the C_1 and Z_1 anion populations. The specific nature of these distributions is predicted to be directly affected by the nature of the anomeric configuration of the precursor anion and the distribution of energies imparted. i.e., Z_1 anions produced from the β -glucose anomeric form have a

differing distribution of product ion structures than do those from the α -glucose anomeric form. The most readily formed Z_1 anions ([1,4-anhydroglucose-H]⁻ structures) are produced from the β -glucose anomers, and do not ring-open and isomerize as the hemiacetal group is no longer present. In contrast, the [3,4-anhydroglucose-H]⁻, Z_1 anion structures which are most readily produced from α -glucose forms, can ring-open through very low barriers (<25 kJ mol⁻¹) to form energetically and entropically favorable aldehyde isomers assigned with a carbonyl stretch at ~1640 cm⁻¹. Barriers to interconversion of the pyranose [β -galactose-H]⁻, C_1 anions to ring-open forms were larger but still modestly (≥ 51 kJ mol⁻¹) consistent with the evidence of both forms presence in the infrared spectrum. For the cross-ring cleavage $^{0,2}A_2$ anions, ring-opening at the glucose hemiacetal of [lactose-H]⁻ is rate-limiting (>180 (α -), >197 kJ mol⁻¹ (β -anomers)). This finding explains the low abundance of these product anions in our tandem mass spectra.

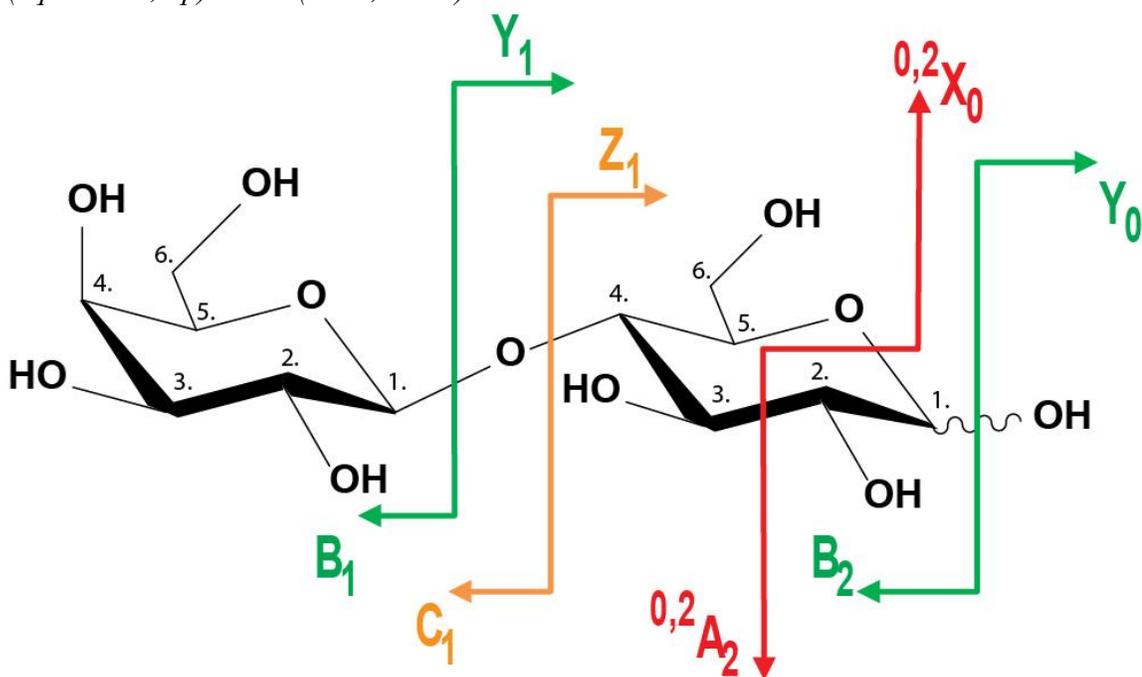
3.2 Introduction

Complex carbohydrates (glycans) are associated with the vast majority of major biological processes.[1] Most proteins and therefore enzymes are glycosylated.[1,2] Glycans are biomarkers for many normally functioning cellular processes. Conversely, changes in the level, location, nature, and/or structure of glycans in body fluids, on cell surfaces, within cells, are increasingly associated with a multitude of disease states.[3–10] To effectively study such processes, we require fast, accurate, and highly sensitive identification (and eventually quantitation) methods for the many different classes of glycan potentially involved.[4,10–17]

Tandem mass spectrometry (MS/MS) utilizing collisional (or other) ion activation methods[1,11,18–24] coupled with prior liquid and/or gas-phase separations has become a major tool in these endeavors. The resulting fragment ions[25] (Scheme 1) then provide the means of sequence and sometimes structural identification.[1,11,12,26–29] The success of this and other tandem mass spectrometry-based approaches is predicated on being able to differentiate the enormous number of isomeric glycan possibilities based on the resulting fragmentation pattern, known chemical information about the analyte source material, as well as any chromatographic[28], ion mobility,[12,18,26,30–33] and/or other diagnostic information[32,34–39] collected for the specific analyte. Unfortunately, all of these pieces of information are not routinely available together, nor immediately assignable when dealing with unknown glycans from a biological source material. Worse still is the issue of isomers. Each glycan and glycan fragment has a huge number of compositional isomers meaning that mass-to-charge (m/z) information alone is seldom sufficient to get the make a confident identification of an unknown (unlike for protonated peptide analytes, for example).

Scheme 3-1:

Carbohydrate numbering and fragmentation nomenclature of Domon and Costello[25] illustrated for lactose (β -D-galactopyranosyl-(1 \rightarrow 4)-D-glucose). The hydroxyl bond at carbon 1 of the glucose is wave-like indicating that the precursor is a mixture of the β (equatorial, up) and α (axial, down) anomers



Irrespective of the approach to MS/MS of charged glycans, we need a means of interpreting the data after it has been generated. If a standard compound or tandem mass spectral library[40–43] is available for the particular glycan analyte analyzed on the same instrument type under very similar experimental conditions, then identification of the analyte is often immediately possible. Unfortunately, this situation is far from routine. The considerable difficulty in synthesizing standard compounds[44] for many glycans and the sheer number of possibilities for library components make comprehensive glycan library-based approaches impractical. An alternative strategy would be to harness an

improved understanding of the gas-phase fragmentation chemistry of the glycan ions[45–48] implemented as rules in an algorithm to generate and then compare theoretical to experimental spectra. Assignments of glycan sequence and structure could then be made and tested.

While most of the early glycan analyses were performed with ionization of glycans by metal cationization[13,29,49–52], many more recent analyses have been performed in the negative mode following experimental generation of anionic, deprotonated glycans.[1,11,12,18,19,26,30,53–58] Some of these approaches have shown isomeric discrimination for precursor anions based on gas-phase structure[30] or MS/MS spectra.[54,59–61] Despite the seeming potential of such approaches, our theoretical understanding of the gas-phase structures and dissociation chemistries involved is relatively underdeveloped.[22,58,59,62]

In the present article, we present some of our initial attempts to characterize deprotonated glycan ions using our combined experimental and computational approach. We characterize the potential energy surface of deprotonated lactose, a simple disaccharide present in milk with density functional[63–65] and MP2[66] calculations. Regiospecifically ^{13}C -labelled lactose analytes enable facile separation of otherwise isomeric fragments (C_n vs. Y_m , B_n vs. Z_m , A_n vs. X_m ions, Scheme 3-1) which we then interrogate with infrared action spectroscopy[62,67–69] and theory.[45,46,48,62] We discuss how the key structures are generated by MS/MS, the critical bond cleavage mechanisms, their relative energies, and the resulting mix of structures.

3.3 Experimental Methods

Regioselectively ^{13}C -labelled lactose samples (β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}\text{C}_6$ -glucose) were purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA). Lactose (β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}\text{C}_1$ -glucose) was purchased from Omicron Biochemicals, Inc. (South Bend, IN, USA).

Tandem mass spectrometric work was carried out using a MaXis plus electrospray-quadrupole time-of-flight mass spectrometer (Bruker, Billerica, MA). MS/MS spectra were obtained by quadrupole isolation of the precursor ion followed by collision-induced-dissociation (CID) in the collision cell, then product ion dispersion by the time-of-flight analyzer. Data were collected as a function of collision energy. Experiments were performed using a sample rate of 1 Hz for 1 minute, thus each data point is the average of 60 individual mass spectra. Typical variations in relative abundance were approximately $\pm 0.95\%$ using a confidence interval of 95%. Breakdown graphs expressing the relative fragment ion signals as a function of collision energy were obtained for labeled and unlabeled [lactose-H] $^+$ ions. Ionization was by electrospray with the samples infused into the instrument in $\sim 1 \mu\text{M}$ acetonitrile/water (50/50) solutions at a flow rate of $3 \mu\text{L min}^{-1}$. Nitrogen was used as nebulizing, drying, and collision gas.

Experimental spectroscopic work was carried out in a Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker Apex IV Qe, Bremen, Germany).[70] Ionization was by electrospray with the various lactose samples infused into the instrument in $\sim 1 \mu\text{M}$ acetonitrile/water (50/50) solutions at a flow rate of $3 \mu\text{L min}^{-1}$. Low-energy collisional activation and subsequent thermalization of the generated

fragment ions can also be achieved in a linear hexapole pressurized with Ar at $\sim 10^{-3}$ mbar. Fragments were produced by CID of the deprotonated precursors, isolated in the quadrupole, prior to transfer to the ICR cell and infrared multiple photon dissociation (IRMPD) analysis.[67,71] IRMPD spectroscopy used the free-electron laser (FEL) at the Centre Laser Infrarouge d'Orsay (CLIO).[71] For each analyte, we follow all the precursor ion and fragment ion peaks. A typical duration of experiment was approximately 80 minutes with a sample rate of 1 data point every 15 seconds, giving approximately 320 data points total. Typical variations in relative abundance over this duration were approximately $\pm 1.14\%$ using a confidence interval of 95%. For unlabeled, deprotonated lactose, the precursor anion was m/z 341.12 ($C_{12}H_{21}O_{11}^-$) and the fragment peaks were m/z 179.059 ($C_6H_{11}O_6^-$) and 161.047 ($C_6H_9O_5^-$). For spectroscopic investigation of the fragments, utilization of labelled samples was necessary as the two primary fragment anions have ambiguous origin and potentially differing structures (m/z 179.059, $C_6H_{11}O_6^-$, is either a C_I anion, a Y_I anion or a mixture of the 2 possibilities; m/z 161.047, $C_6H_9O_5^-$, is either a B_I anion, a Z_I anion or a mixture of the 2 possibilities, Scheme 1). Consequently, we collisionally fragmented the deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}C_6$ -glucose anions (m/z 347.129, $^{13}C_6C_6H_{21}O_{11}^-$) to produce the subsequent analytes for spectroscopic analysis. The predominant fragments were the C_I , m/z 179.059 ($C_6H_{11}O_6^-$) and the Z_I , m/z 167.066 ($^{13}C_6H_9O_5^-$) peaks. For the Z_I ion, m/z 167.066 ($^{13}C_6H_9O_5^-$), the fragment peaks were m/z 105.038 ($^{13}C_4H_5O_3^-$), 87.0277 ($^{13}C_4H_3O_2^-$), and 76.039 ($^{13}C_3H_5O_2^-$); for the m/z 179.059 ($C_6H_{11}O_6^-$), C_I ion, the fragment peaks were m/z 161.047 ($C_6H_9O_5^-$), 143.036 ($C_6H_7O_4^-$), 131.035 ($C_5H_7O_4^-$), and

113.025 ($C_5H_5O_3^-$). We utilized the deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}C_1$ -glucose analyte in which carbon 1 of the reducing end (glucose) was ^{13}C labeled (Scheme 1) to identify the position from which the carbon atoms were lost in $^{0,2}A_2$ anion-forming reactions. An analogous approach was utilized in the preceding tandem mass spectrometry experiments on the MaXis plus. The abundances of the photofragments and their corresponding precursors were recorded as a function of the IR wavelength to derive the IR action spectra where the IRMPD efficiency is plotted against the photon energy. Both the IR FEL power variation and wavelength are monitored online while recording the IRMPD spectrum. For this purpose, a small fraction of the IR beam is directed towards a power meter. A second power meter is used to record the IR absorption spectrum of a polystyrene film using a second fraction of the IR FEL beam. As a result, at each wavelength during the scan, relative power, polystyrene absorption, and a mass spectrum are simultaneously recorded. Wavelength and power corrections can thus be made during the data treatment.

3.4 Theoretical Methods

Simulations were performed to enable effective characterization of the potential energy surface of glycan analytes. Initial candidate structures for lactose, as well as multiple potential fragment ions, were systemically generated via the tool Fafoom[72–74], a genetic algorithm. The structures were optimized using the MMFF94 Force Field.[75–79] This approach samples a wide range of ring structures incorporating multiple chair, boat, and skew forms enabling thorough interrogation of the potential energy surface. i.e., not just 4C_1 -type structures. Geometry optimizations of the resulting

candidate conformations were performed with the Gaussian 09 software package[80] at the B3LYP/6-31G(d)[81], then M06-2X/6-31++G(d,p)[46] levels of theory. Degenerate structures were removed at each stage, and the non-degenerate structures were further optimized at the B3LYP/6-311++G(2d,2p) level of theory with single point calculations performed at the M06-2X/6-311++G(2d,2p), MP2/6-311++G(2d,2p), and M06-2X/aug-cc-pVDZ levels of theory to assess energetic variability as a function of model and basis set. Additional, targeted manual adjustment and supplementation of the structural pool analyzed were performed as necessary, to ensure chemically relevant species are not being neglected. Density functional theory calculations of minima, transition structures, product ions and neutrals present on each potential reaction pathway were performed at the M06-2X/6-31++G(d,p) level of theory. Multiple transition structures (TSs) were calculated for each potential pathway. Minima and TSs were tested by vibrational analysis (all real frequencies or 1 imaginary frequency, respectively). The potential energy surface generated combined the zero-point energy correction (ZPE) to the electronic energy (E_{el} , 0 K) for improved accuracy ($\Delta E_{el+ZPE,0K}$). The related, standard enthalpy (ΔH_{298}), Gibbs free energy (ΔG_{298}), and entropy (ΔS_{298}) corrections to 298 K were also determined. The reaction pathway through each TS was determined by intrinsic reaction coordinate (IRC) calculations with up to 10 steps in each direction. The terminating points of these calculations (one on the product-side, one on the reactant-side) were then optimized further to determine which minima were connected to each TS. Calculated B3LYP/6-311++G(2d,2p) vibrational frequencies were utilized for comparisons with the experimental “action” IR spectroscopy spectra. A scaling factor of

0.967 was utilized for the vibrational frequencies. A 12 cm^{-1} full width at half maximum Lorentzian line shape was employed for comparison to the experimental spectra.

3.5 Results and Discussion

3.5.1 Tandem Mass Spectrometry of [Lactose-H]⁻

Our regioselectively labelled mass spectra of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D-¹³C₆-glucose anions (nominal m/z 347, ¹³C₆H₂₁O₁₁⁻) support primary cleavage of the glycosidic bond adjacent to the reducing end (glucose) of deprotonated lactose (Figure 3-1, Figure 3S1, Scheme 3-1); the C_I - Z_I reaction. The Z_I ion (¹³C₆H₉O₅⁻, nominal m/z 167) is the most prevalent fragment followed by the C_I ion (C₆H₁₁O₆⁻, m/z 179).

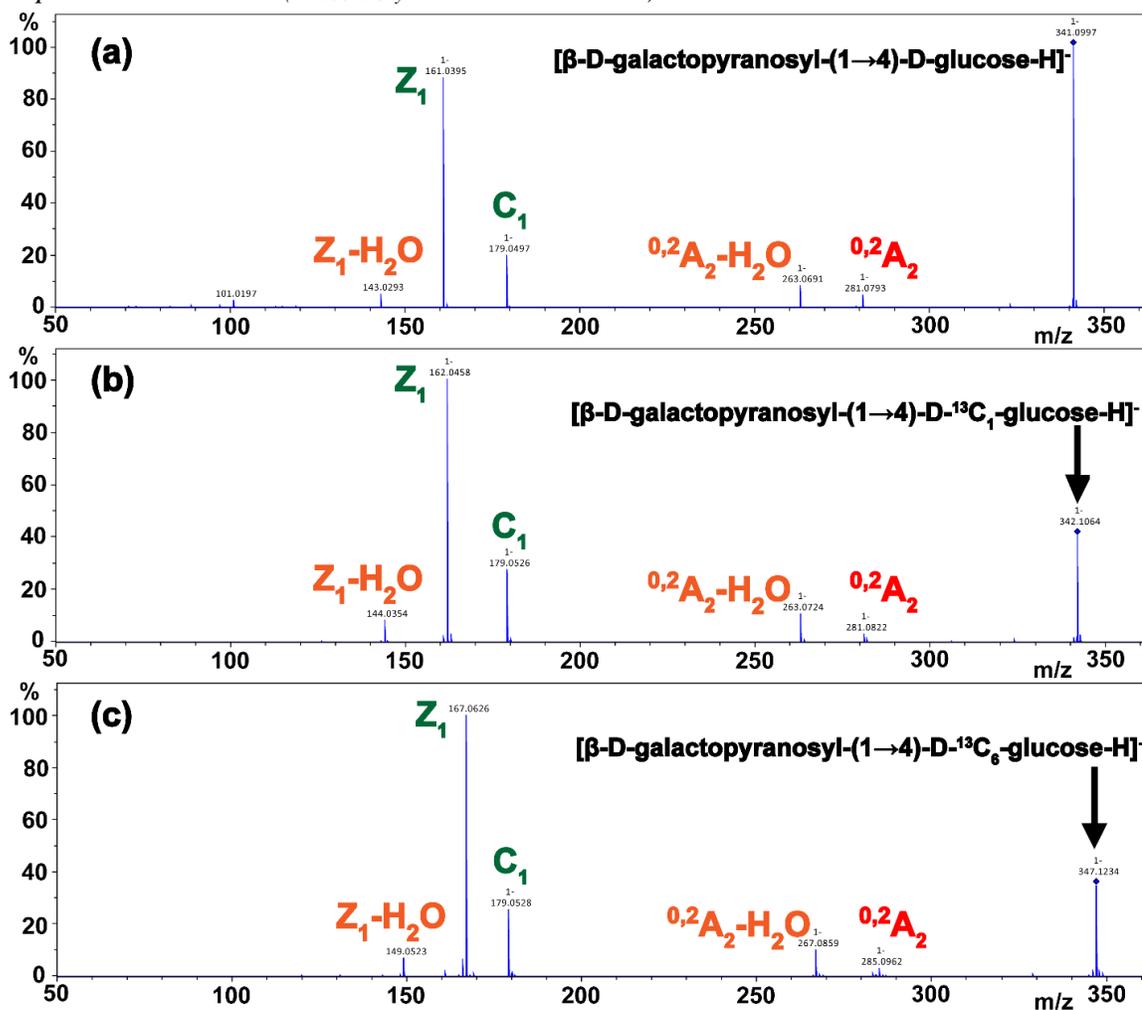
Additional peaks corresponding to reactions involving ring-opening and C-C bond cleavage follow (m/z 285, ^{0,2}A₂, ¹³C₄C₆H₁₇O₉⁻ and m/z 267, ^{0,2}A₂-H₂O, ¹³C₄C₆H₁₅O₈⁻).

When the deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D-¹³C₁-glucose analyte (m/z 342, ¹³C₁C₁₁H₂₁O₁₁⁻) in which carbon 1 of the reducing end (glucose) is a ¹³C (Scheme 1) was instead isolated then fragmented (Figure 3-1b), the Z_I peak shifted by 1 u (¹³C₁C₅H₉O₅⁻, m/z 162) while the C_I ion (C₆H₁₁O₆⁻, m/z 179) peak did not shift relative to the unlabeled analyte spectrum (Figure 3-1a). This is entirely consistent with the preceding data. For the cross-ring cleavage products (^{0,2}A₂ and ^{0,2}A₂-H₂O) there are no shifts in m/z relative to the unlabeled (m/z 341, C₁₂O₁₁H₂₁⁻) precursor and the deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D-¹³C₁-glucose anion (m/z 347, ¹³C₁C₁₁O₁₁H₂₁⁻) spectra. The resulting anion peaks (m/z 281, ^{0,2}A₂, C₁₀H₁₇O₉⁻ and m/z 267, ^{0,2}A₂-H₂O, C₁₀H₁₅O₈⁻) indicate (Figure 3-1b) loss of ¹³C₁C₁H₄O₂ (61 u) and ¹³C₁C₁H₆O₃ (79 u), respectively. These losses must include carbon 1 of the reducing end glucose (Scheme 3-1, Figure 3-

1b). Lastly, we should note that no substantial B_1 or Y_1 anion peaks were detected providing clear evidence of the comparative lack of competitiveness of cleavage adjacent to the galactose residue (B_1 - Y_1 reaction) in comparison to earlier cationized analytes.[29,45,46,49]

Figure 3-1:

Example Bruker MaXis plus MS/MS spectra of regular and regioselectively labeled, deprotonated lactose ($E_{\text{laboratory}} = 6 \text{ eV}$ in all cases)



3.5.2 Low Energy Structures of [Lactose-H]⁻

The lowest energy conformers of [lactose-H]⁻ are predicted to have a proton bridged between the galactose C2 oxygen and glucose C3 oxygen (Figure 3S2-3S4). Similar to solution phase data, our calculations support both the α - and β -glucose pyranose forms being substantially more energetically favorable than the lowest energy ring-open forms (Figure 3S4). The relative energies shift with hydrogen-bonding pattern, angle of each glycosidic bond, ring form, and anomericity. Our calculations also sampled each of the other sites of potential hydroxyl deprotonation. While not all of these are predicted to be energetically competitive, several of the reducing-end hydroxyls produce relatively low energy deprotonated forms.

Comparison between calculated frequencies of the α - and β -glucose pyranose and the ring-open isomers for these structures and our experimental data (Figure 3S5) provide essentially no ability to discern which structures are likely populated. An extremely broad and intense experimental feature was recorded between 900 and 1400 cm⁻¹. Beyond 1400 cm⁻¹ the spectral complexity is greatly reduced, but so is the apparent strength of the bands present underlining some of the inherent difficulties[82–86] in examining such systems spectroscopically.

We note however, that while the spectroscopic data is inconclusive, additional structural evidence is available from our MS, MS/MS (Figure 3-1), and theoretical data. Our MS/MS data do not support a significant population of the aldehyde, ring-open isomer being present. We detect minimal cross-ring fragmentation at low collision energies (Figure 3-1, ^{0,2}A₂ and ^{0,2}A₂-H₂O anion peaks, Figure 3S1, Tables 3-1 & 3S1, see

discussion in Results and Discussion, Section 3.5.6). This is consistent with data from multiple earlier solution-phase experimental approaches which identify the open-chain aldehyde form in extremely low concentrations ($\sim 0.002\%$).^[87–90] Heating solutions to 82 °C did result in increased detection of the open-chain aldehyde, but only to a concentration of $\sim 0.02\%$.^[87] Even experiments at high pH where solution-phase carbohydrate anions are present indicate the overwhelming dominance of the pyranose forms.^[91] Consequently, if the aldehyde is being formed at all prior to CID, it would have to be in the electrospray process. Based on our calculations (Results and Discussion, Sections 3.5.3–3.5.6) this should result in substantial fragmentation of the precursor ions too. We see no evidence to support this, i.e., no substantial low m/z ion current or cross-ring fragment peaks present in the initial MS stage.

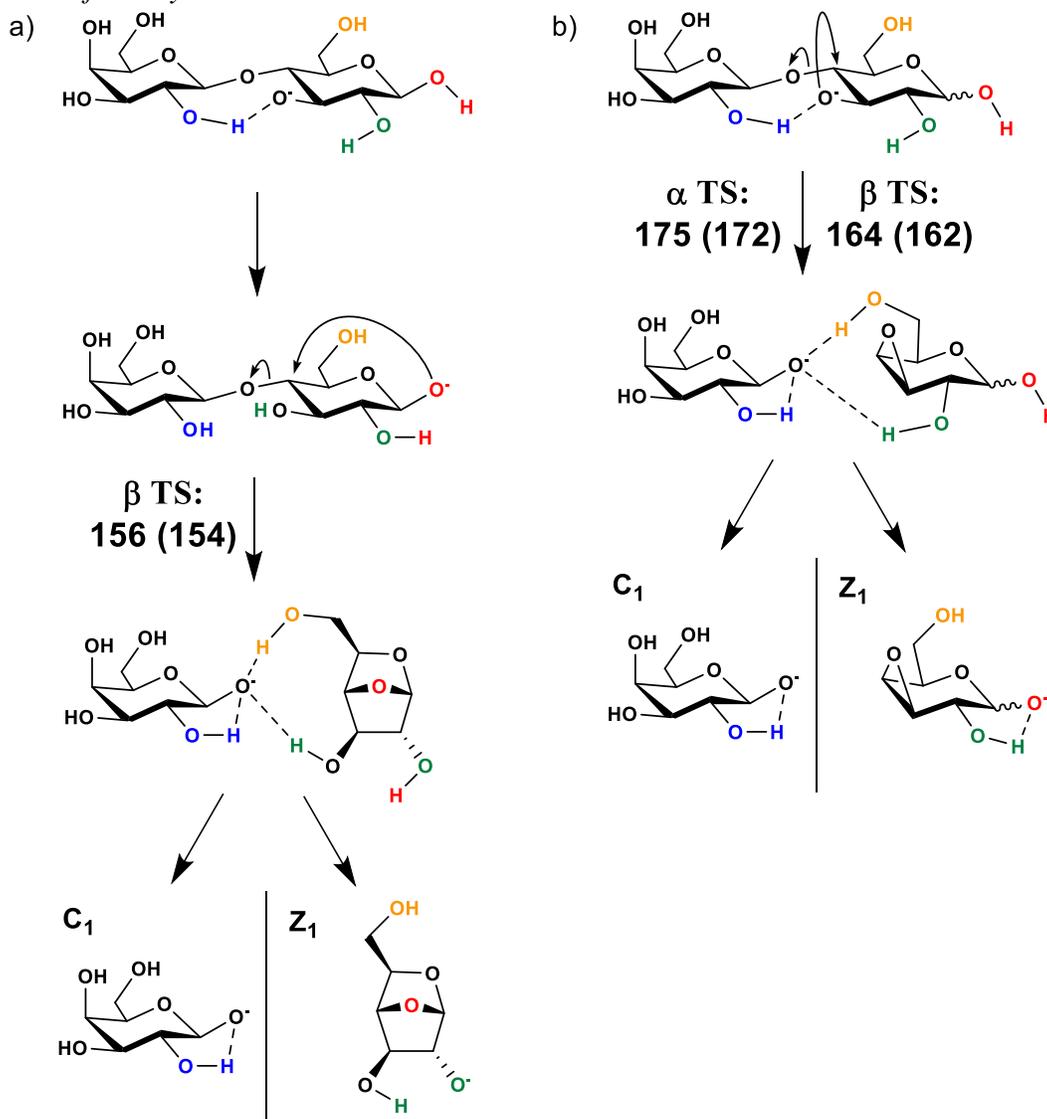
3.5.3 Glycosidic Bond Cleavage Reaction Energetics

The lowest energy transition structures for the C_1 - Z_1 glycosidic bond cleavage reactions of [lactose-H]⁻ are predicted to vary with anomeric configuration. Our calculations support the predomination of two, related intra-glucose S_N2-like mechanisms (Scheme 3-2, Figure 3-2, Table 3-1) producing either deprotonated 1,4-anhydroglucose or 3,4-anhydroglucose structures (Figure 3S6). The reaction forming 1,4-anhydroglucose (Scheme 3-2a, Figure 3-2) is the lowest energy pathway for the β -D-glucose anomeric form, requiring at least 156 (154) kJ mol⁻¹, $\Delta E_{\text{el+ZPE,0K}}$ ($\Delta G_{298\text{K}}$). Following glycosidic bond cleavage, an ion-molecule complex of deprotonated β -D-galactose and 1,4-anhydroglucose is formed. If the complex separates directly, a C_1 anion structure can be detected. Alternatively, one or more proton transfers can occur within the ion-

molecule complex prior to separation. The subsequent dissociation will then produce either a C_I or Z_I anion depending on which fragment gets neutralized at the end of the final proton transfer prior to complex dissociation. Thus, the formation of $[1,4\text{-anhydroglucose-H}]^-$, Z_I anion structures necessitates abstraction of a proton from the neutral 1,4-anhydroglucose prior to complex separation and product detection (Scheme 3-2a).

Scheme 3-2:

The lowest-energy glycosidic bond cleavage pathways of deprotonated lactose. a) 1,4-anhydroglucose-forming pathway initiated from β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-glucose and b) 3,4-anhydroglucose-forming pathway achievable from both anomeric forms. Values in kJ mol^{-1} $\Delta E_{el+ZPE,0K}$ (ΔG_{298}), calculated at the M06-2X/6-31++G(d,p) level of theory



The main alternative C_1 - Z_1 reaction type follows a similar, though more strained S_N2 -like mechanisms to glycosidic bond cleavage (Figure 3S7 & 3S8) resulting in

different ion-molecule complexes which comprise a deprotonated β -D-galactose anion and 3,4-anhydroglucose (Scheme 3-2b). The relevant glycosidic bond cleavage transition structures require at least 175 (172) kJ mol^{-1} , $\Delta E_{\text{el}+\text{ZPE},0\text{K}}$ ($\Delta G_{298\text{K}}$) for the α -glucose anomer and 164 (162) kJ mol^{-1} for the β -glucose anomer to populate (Table 3-1). This reflects the additional strain involved in the formation of the 3,4-epoxides (Figures 3S7 and 3S8). From the ion-molecule complex formed, production of the $[\text{3,4-anhydroglucose-H}]^-$, Z_I ion structure necessitates the transfer of a proton to the deprotonated β -D-galactose anion (Scheme 3-2b) in a manner directly analogous to that discussed previously for 1,4-anhydroglucose type Z_I anions. Production of a third family of Z_I ion *directly* from potential ring-open conformers is not supported by our calculations as the barriers (221 (215) kJ mol^{-1}) to these reactions are substantially higher than either glucopyranose form (Table 3S1).

Figure 3-2:

The lowest-energy C_1-Z_1 transition structure of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory (a deprotonated 1,4-anhydroglucose-forming TS). Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298}). The reaction coordinate is illustrated with blue dots

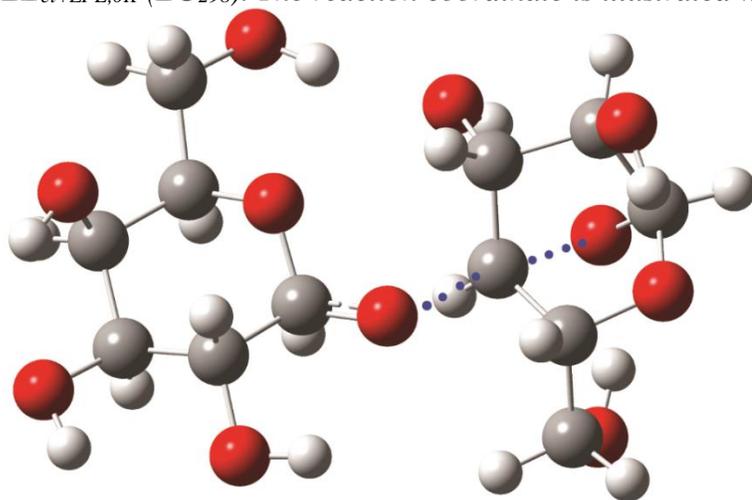
**156 (154)**

Table 3-1:

Relative energies of the transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- α -D-Glucose anomer) calculated at the M06-2X/6-31++G(d,p) level of theory. The α and β forms are both present but are incapable of interconverting to any significant extent prior to fragmentation occurring. The β -anomer Global Minimum is 5.0 (3.6) kJ mol⁻¹, $\Delta E_{el+ZPE,0K}$ (ΔG_{298}) above the α -anomer

α -D-Glucose anomer	E_{el}/H	E_{el+ZPE}/H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298}/$ kJ mol ⁻¹	$\Delta G_{298}/$ kJ mol ⁻¹	$\Delta S_{298}/$ J K ⁻¹ mol ⁻¹
Global Minimum (α)	-1297.000619	-1296.637181	0	0	0	0
C ₁ -Z ₁ TS (3,4-anhydroglucose)	-1296.932761	-1296.570500	175.1	175.7	172.3	11.4
C ₁ -Z ₁ TS (3,4-anhydro ring open)	-1296.909679	-1296.551085	226.0	231.2	218.1	43.7
Ring opening TS	-1296.920038	-1296.562144	197.0	195.8	197.3	-4.9
^{0,2} A ₂ formation TS_A	-1296.949413	-1296.591725	119.3	124.2	111.4	42.7
^{0,2} A ₂ formation TS_B	-1296.944470	-1296.591076	121.0	126.6	111.7	53.0
β -D-Glucose anomer	E_{el}/H	E_{el+ZPE}/H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298}/$ kJ mol ⁻¹	$\Delta G_{298}/$ kJ mol ⁻¹	$\Delta S_{298}/$ J K ⁻¹ mol ⁻¹
Global Minimum (β)	-1296.996960	-1296.635278	0	0	0	0
C ₁ -Z ₁ TS (1,4-anhydroglucose)	-1296.938463	-1296.575823	156.1	155.8	154.4	4.7
C ₁ -Z ₁ TS (3,4-anhydroglucose)	-1296.935448	-1296.572877	163.8	164.5	161.9	8.6
C ₁ -Z ₁ TS (3,4-anhydro ring open)	-1296.909679	-1296.551085	221.0	225.2	214.5	35.8
Ring opening TS	-1296.924404	-1296.566837	179.7	177.5	181.7	-14.2
^{0,2} A ₂ formation TS_A	-1296.949413	-1296.591725	114.3	118.2	107.8	34.7
^{0,2} A ₂ formation TS_B	-1296.944447	-1296.591076	116.1	120.7	108.1	42.3

3.5.4 Z_I Structures

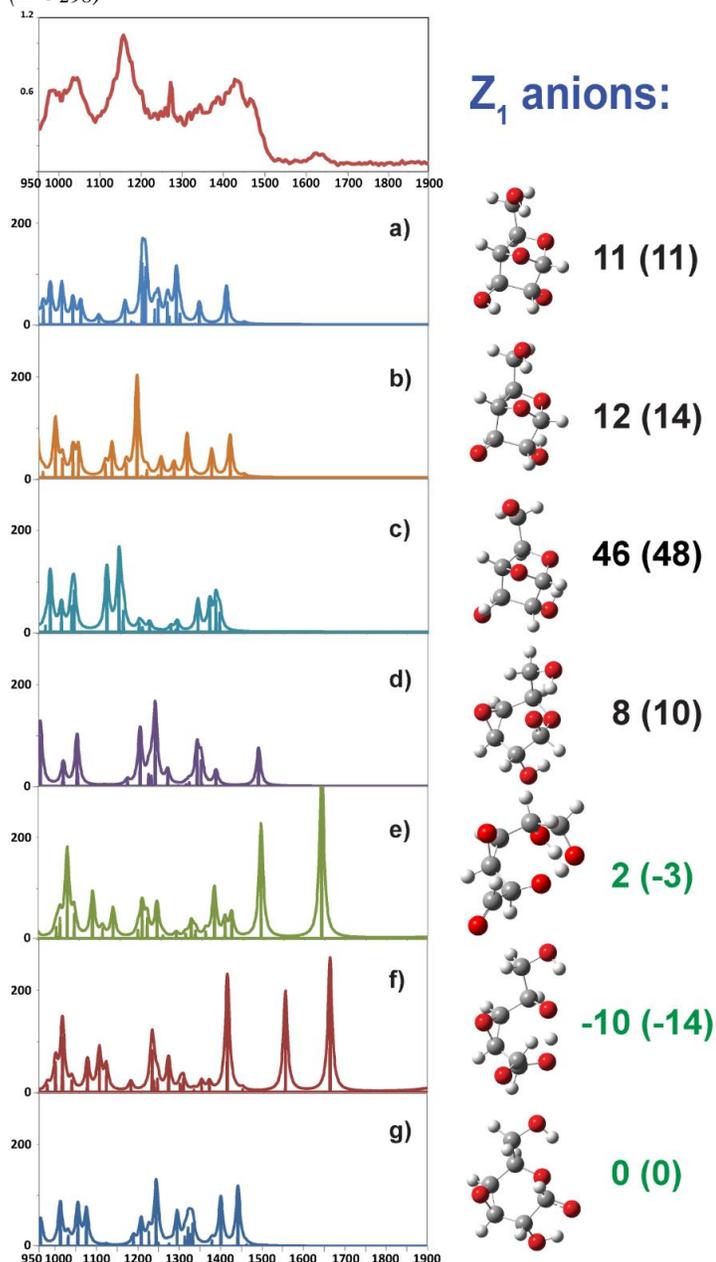
Our Z_I anion infrared action spectrum is shown in Figure 3-3. The lower energy features are potentially consistent with the [1,4-anhydroglucose-H]⁻, Z_I anion structures shown in panels a, b, and c (secondary alkoxides) providing some of the ion population (Figure 3-3). We note that these 3 rigid, bicyclic structures (Figure 3S6a) cannot interconvert or readily ring-open due to the lack of a hemiacetal group. This potential assignment is made more based on mechanistic evidence from the TS calculations than confidence in the spectroscopic evidence though. Why? Carbohydrate ions typically have many bands in this spectral region.[35,58] In the present case, a broad but structured absorbance with maxima at 980, 1040, 1150, 1270, and 1430 cm⁻¹ is detected. However, potentially the most structurally diagnostic band in the entire spectrum is the one centered near 1640 cm⁻¹. i.e., our IRMPD spectrum suggests the presence of a carbonyl stretch within the Z_I ion population, so one or more ring-open structure(s) is present (panels e and f). Given our earlier statements that argued against *direct* formation of Z_I anions from ring-open precursors and that the predicted [1,4-anhydroglucose-H]⁻ do not ring-open, how is this ion being generated?

Formation of [3,4-anhydroglucose-H]⁻, Z_I ion structures is predicted to predominate from the α -glucose anomer of [lactose-H]⁻. From this precursor anion, we calculated the barrier to ring-opening at the hemiacetal functional group (Scheme 3-3, Figure 3S9). These barriers are very low at only 23 (23) kJ mol⁻¹, $\Delta E_{el+ZPE,0K}$ (ΔG_{298}). This barrier indicates that these structures should be easily accessible under our experimental conditions (Scheme 3-3). Furthermore, the ring-open Z_I anions are substantially more entropically favorable ($\Delta S_{298K} \geq 23$ J mol⁻¹) indicating that this

transformation will produce highly favorable product anions as activation level increases (Figure 3S6). Together, this interpretation points to distinct populations of Z_I anion structures from each of the two anomeric configurations. Analogous proposals have recently been advocated for some cationized monosaccharides.⁴⁷ These findings, if general, will have a marked effect in helping explain the earlier tandem mass spectrometric findings from the Bendiak and Xia groups.[20,21]

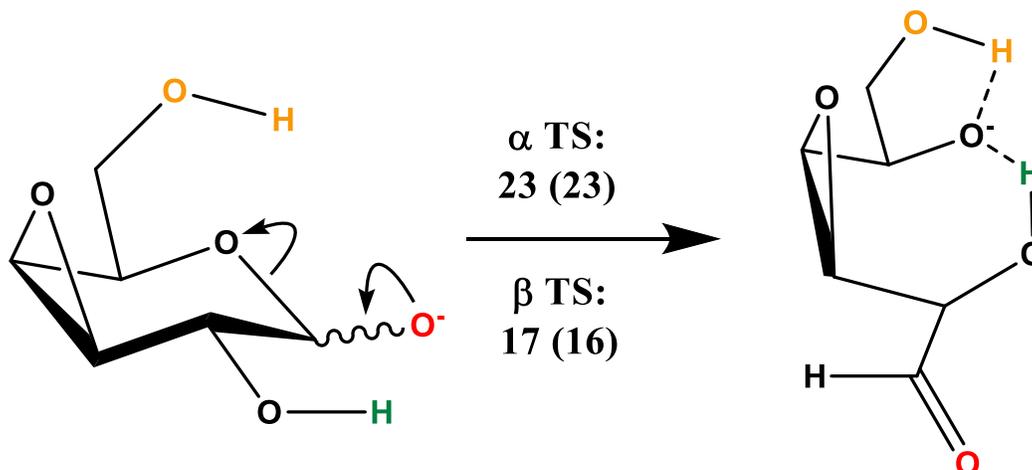
Figure 3-3:

Z_1 anion (m/z 167.066, $^{13}\text{C}_6\text{H}_9\text{O}_5^-$) infrared action spectroscopy spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}\text{C}_6$ -glucose anions (m/z 347.129, $^{13}\text{C}_6\text{H}_{21}\text{O}_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures. Panels a-c: [1,4-anhydroglucose-H] $^-$ structures; panel d: [3,4-anhydro- β -glucose-H] $^-$; panels e and f: ring-open 3,4-anhydroglucose; panel g: [3,4-anhydro- α -glucose-H] $^-$. Relative energies calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,OK}$ (ΔG_{298})



Scheme 3-3:

Ring-opening isomerization of deprotonated 3,4-anhydroglucose Z₁ anion structures. Values in kJ mol⁻¹, $\Delta E_{el+ZPE,0K}$ (ΔG_{298}) calculated at the M06-2X/6-31++G(d,p) level of theory



Consequently, we are observing the population that emerged from that ion-molecule complexes. i.e., at least some kinetic rather than thermodynamic product anion control. In terms of which modes are *potentially* providing the observed IRMPD spectrum, we have: The strong features in the ~1180-1220 cm⁻¹ region are consistent with C-O- stretches coupled to the C-H and C-O-H bends (symmetric and/or asymmetric); Analogous, lower energy, coupled motions between the ring ether and adjacent hydroxyl groups are consistent with theoretical bands at ~979, 1008, 1024, 1030, 1035, 1037, and 1053 cm⁻¹; Coupled C-CH₂-OH bending bands ~1030-1060 cm⁻¹ are also predicted; The region from ~1090-1170 cm⁻¹ is consistent with primary alkoxide C-O⁻ stretches coupled to C-H bends and vibrations; These combined C-H motions from C-(H)C-O⁻ species produce strong theoretical absorptions in the 1145-1165 and ~1280 cm⁻¹ ranges. Ring hydroxyl (1343, 1367, 1373, 1383 cm⁻¹) and primary hydroxyl (1409, 1418 cm⁻¹) bends provide the

remainder of potentially assignable bands for these structures (panels a-c). The deprotonated 3,4-anhydroglucopyranose Z_I anions, if present, are predicted to have intense hydrogen-bonded hydroxyl bending motions ($\text{O-H}\cdots\text{O-C}$) at 1438 (α , panel g) and 1459 (β , panel d) cm^{-1} . Based on our calculations, the experimental band at approximately 1640 cm^{-1} is consistent with one or more carbonyl stretches in the ring-open aldehyde isomers (panels e and f).

An obvious potential problem with the hypothesis that some of the anion population is ring-open aldehyde is shown in panels e and f. A strong band is predicted at 1494 cm^{-1} in panel e and at 1553 cm^{-1} in panel f without clear corresponding experimental features. Our explanation for this is a combination of the fact that: (1) The majority of the Z_I anion population formed is actually represented by panels a-c, due to the low barrier for this process and the inability for these product anions to ring-open, and (2) that the highlighted bands (1494 cm^{-1} and 1553 cm^{-1}) correspond to motions ($\text{C-O}\cdots\text{H}^+\cdots\text{O-C}$) up and down rather than back and forth between the two oxygens. These vibrations have been shown to result in much broader experimental bands than predicted by all but the most sophisticated theoretical approaches[84,86] and for the bands themselves to be sensitive to small adjustments in structure.[82–86] Additional support for this hypothesis comes from the fact that lactose typically has approximately double the population of pyranose β -glucose than α -glucose anomer in solution. Consequently, unless the combined MS and theoretical data (Results and Discussion, section 3.5.2) is in error, the fragmentation distribution should be substantially biased in favor of the [1,4-anhydroglucose-H] $^-$, Z_I anion structures (Scheme 3-2a, Figure 3-2, panels a-c, Table

3-1) due to the pyranose β -glucose's combined relatively large precursor ion population, low barrier to fragmentation (156 (154) kJ mol⁻¹), and energetic inability to ring-open (180 (182) kJ mol⁻¹) and interconvert (192 (194) kJ mol⁻¹) to the α -glucose form.

3.5.5 C_1 Structures

Our C_1 anion infrared action spectroscopy spectrum is shown in Figure 3-4. Consistent with the prior literature^{58,62} we observe bands above ~ 1600 cm⁻¹ that are unexplainable without consideration of the ring-open isomers. This feature is broad and asymmetric, so is more consistent with multiple carbonyl stretches, of differing populations rather than a single environment. From multiple deprotonated β -D-galactopyranose C_1 anion forms we calculated barriers to the formation of ring-open conformers. Our calculations indicate that this process requires more energy than for the [3,4-anhydroglucose-H]⁻, Z_1 anion structures. However, this is still only 51 (47) kJ mol⁻¹, $\Delta E_{\text{el+ZPE,0K}}$ (ΔG_{298}) kJ mol⁻¹ and is an entropically favorable process (Figure 3S10). Major rearrangements in peptide sequence ions have been detected spectroscopically despite similar isomerization barriers.[92–98] This barrier is substantially lower than that calculated for ring-opening of [lactose-H]⁻ which required ≥ 180 kJ mol⁻¹ (Table 3-1), indicating that ring-opening should not necessarily be expected to immediately occur for larger C_n ions. The β -D-galactopyranose C_1 anion forms are enthalpically preferred over the ring-open forms (at 298 K). However, the ring-open product ions are hugely entropically more favorable (>40 J K⁻¹ mol⁻¹) resulting in Gibbs free energies similar to those of the lowest energy deprotonated β -D-galactopyranose forms at 298 K (Figure

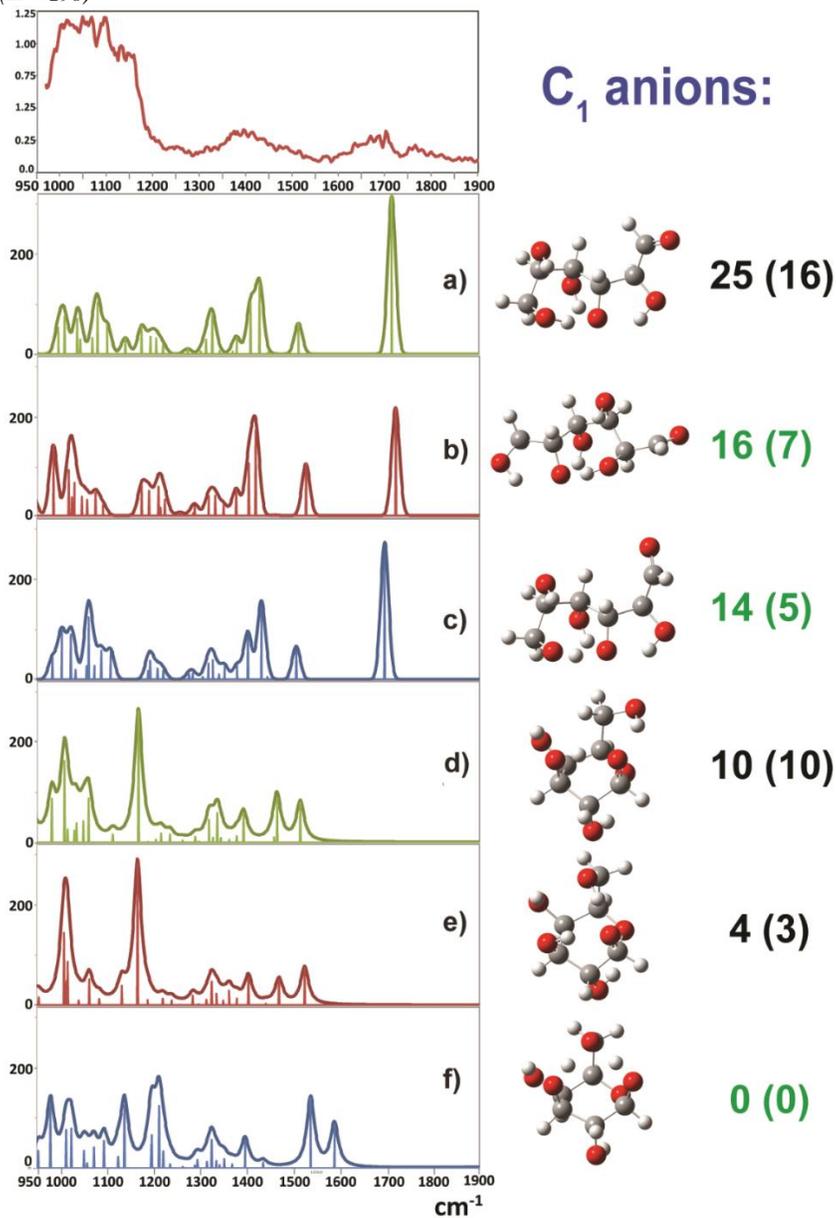
3S11). In summary, the formation of at least some population of these ring-open aldehyde C_I anion structures is likely.

The ring-open aldehyde C_I anion structures have intense predicted bands between ~ 1300 and 1550 cm^{-1} . Our experimental spectrum features a broad band, an unresolved band in this energy range centered at $\sim 1400\text{ cm}^{-1}$. The low energy ($<1200\text{ cm}^{-1}$) part of the spectrum is intense, but not diagnostic as theoretical bands are predicted for both pyranose and ring-open aldehyde structures (various hydrogen-bonded, primary alkoxide C-O⁻ stretches coupled to adjacent O-H bends). Thus, we cannot entirely discard the possibility that some of the population is pyranose isomers. Contrary to this assertion is the fact that the least convincing pyranose theoretical spectrum is produced by the lowest energy β -D-galactopyranose structure (panel f) in which two, coupled, hydroxyl proton oscillations equidistant from the deprotonated primary alkoxide anion are predicted at 1538 (asymmetric) and 1590 (symmetric) cm^{-1} for this structure (whereas no substantial experimental bands are detected between ~ 1540 - 1600 cm^{-1}). However, these protons are in nearly identical environments and such environments have been shown to result in much broader experimental bands than predicted by all but the most sophisticated theoretical approaches[84,86] and for the bands themselves to be sensitive to small adjustments in structure.[82–86] In summary the combined mass spectrometric, spectroscopic, and theoretical data provides evidence that ring-open, deprotonated anhydrogalactose structures can be formed and remain populated on a timescale at least that of our spectroscopic measurements. The deprotonated β -D-galactopyranose anion structures are readily formable under our experimental conditions. The ambiguity in the

lower energy part of our IR spectrum makes it impossible to rule out populations of these isomers though, despite the lack of a band solely diagnostic for the pyranose ring form in our experimental range.

Figure 3-4:

C_1 anion (m/z 179.059, $C_6H_{11}O_6^-$) infrared action spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D- $^{13}C_6$ -glucose anions (m/z 347.129, $^{13}C_6C_6H_{21}O_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures. Panels a-c: Ring-open deprotonated galactose structures; Panels d-f: [β -D-galactose-H] $^-$. Relative energies calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298})

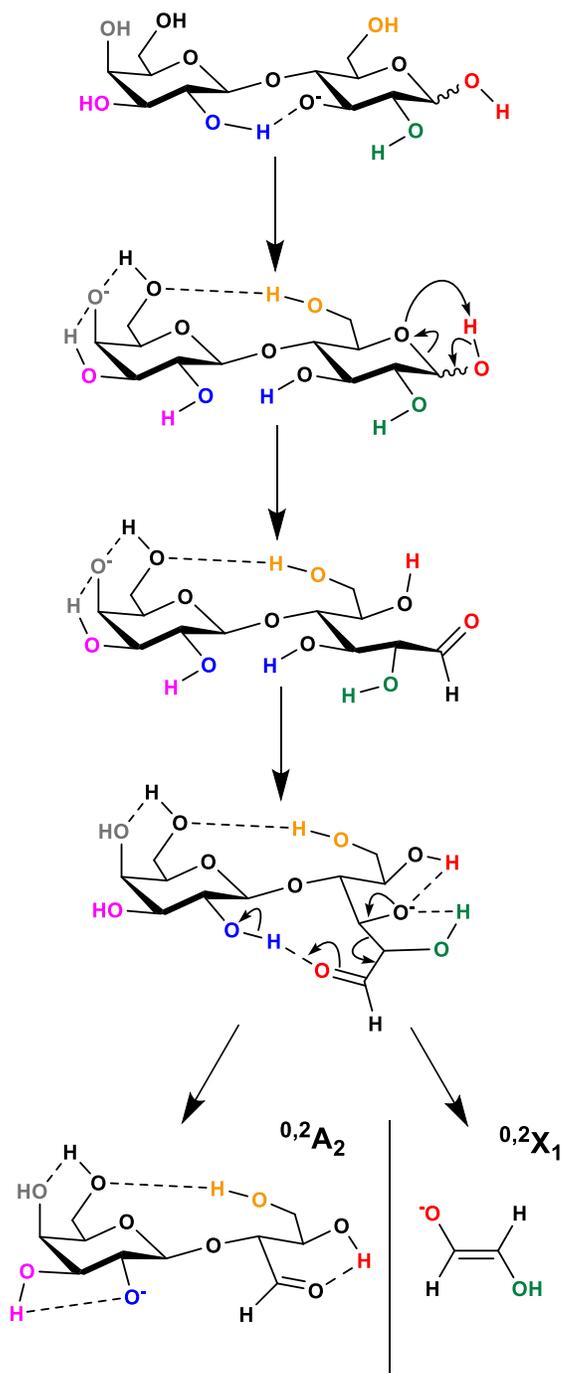


3.5.6 Cross-ring Fragmentation Pathways: $^{0,2}A_2$ Anion Formation

Lastly, we briefly discuss the cross-ring cleavage reactions of [lactose-H]⁻. The abundance of these fragments was low (Figure 3-1, Figure 3S1) precluding spectroscopic investigation. Our calculations predict that this entire process is governed by the ring-opening reaction with the subsequent cleavage of the C2-C3 bond being substantially less energetically demanding (Table 3-1, Scheme 3-4, Figures 3S12 & 3S13). This result is consistent with both the lack of $^{0,2}A_2$ anion peaks at low collision energies experimentally and the population of ring-open isomers prior to collisional activation being very small (Table 3S1). Were this not the case, $^{0,2}A_2$ anion peaks would be the most readily formed (Table 3-1, Table 3S1). The barrier to ring-opening (≥ 180 kJ mol⁻¹) is affected by anomeric configuration with our calculations indicating the α -anomer required ~ 17 kJ mol⁻¹ more energy to initiate this process (Table 3-1, Figures 3S12 & 3S13). Following ring-opening, facile rotations and proton transfers enable cleavage of the C2-C3 bond to most readily furnish a dimer of $^{0,2}A_2$ and trans-1,2-ethendiol (Scheme 3-4, Figures 3S12 & 3S13). The dimer then separates to yield the detected $^{0,2}A_2$ anion. Similar to recent cationized data[51,52], the ring-opening barrier is found to be entropically disfavored (Table 3-1) which likely contributes to this pathway only contributing a minor portion of the ion current detected.

Scheme 3-4:

The lowest-energy ${}^{0,2}A_2$ anion formation pathway of deprotonated lactose



3.6 Conclusions

Our combined experiments and calculations support a mixture of gas-phase ion structures comprising the Z_I anion population generated from [lactose-H]⁻. Evidence for ring-open aldehyde C_I anion structures is provided, although some pyranose populations could not be discounted. Isomerization barriers between *some* though not all of the product anion forms are extremely low enabling facile interconversion. In contrast to the other ring-product structures (both C_I and Z_I types), the Z_I ion structures most readily formed from the β -anomers ([1,4-anhydroglucose-H]⁻) are substantially more stable. This is due to the elimination of the hemiacetal functional group during the glycosidic bond cleavage process. The ring-opening barriers for [lactose-H]⁻ are predicted to be much more substantial than for β -D-galactopyranose anions thereby rate-limiting cross-ring cleavage reactions (0,2A_2). If true in general, this likely prevents larger C_n (or Y_m) anions from facily opening/fragmenting most readily via cross-ring cleavage. Subsequent investigations will seek to broaden the size, complexity, and range of functional groups present within deprotonated carbohydrate anion analytes to test how general the present findings are.

3.7 Acknowledgments

This work was supported by start-up funds from the University of Missouri-St. Louis Department of Chemistry and Biochemistry and a 2018 University of Missouri-St. Louis Research Award. Calculations were performed at the Missouri University of Science and Technology Rolla, MO. This material is based (in part) upon work supported by the National Science Foundation under 1808394 (to BJB). Calculations were

performed at the Missouri University of Science and Technology Rolla, MO. The research leading to this result has been supported by the project CALIPSOplus under the Grant Agreement 730872 from the EU Framework Programme for Research and Innovation HORIZON 2020. The authors are grateful to E. Loire, J. M. Ortega, F. Gobert, and N. Jestin for technical support.

3.8 References

- 1 A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Betozzi, G. W. Hart and M. E. Etzler, *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press: New York, 2008.
- 2 D. S. Alonzi, D. C. A. Neville, R. H. Lachmann, R. A. Dwek and T. D. Butters, Glucosylated free oligosaccharides are biomarkers of endoplasmic- reticulum alpha-glucosidase inhibition, *Biochem. J.*, 2008, **409**, 571–580.
- 3 B. Adamczyk, T. Tharmalingam and P. M. Rudd, Glycans as cancer biomarkers, *Biochim. Biophys. Acta*, 2012, **1820**, 1347–1353.
- 4 L. R. Ruhaak, S. Miyamoto and C. B. Lebrilla, Developments in the identification of glycan biomarkers for the detection of cancer, *Mol. Cell. Proteomics MCP*, 2013, **12**, 846–855.
- 5 Z. Yin and X. Huang, Recent Development in Carbohydrate Based Anti-cancer Vaccines, *J. Carbohydr. Chem.*, 2012, **31**, 143–186.
- 6 C. A. Reis, H. Osorio, L. Silva, C. Gomes and L. David, Alterations in glycosylation as biomarkers for cancer detection, *J. Clin. Pathol.*, 2010, **63**, 322–329.

- 7 Y.-L. Huang, J.-T. Hung, S. K. C. Cheung, H.-Y. Lee, K.-C. Chu, S.-T. Li, Y.-C. Lin, C.-T. Ren, T.-J. R. Cheng, T.-L. Hsu, A. L. Yu, C.-Y. Wu and C.-H. Wong, Carbohydrate-based vaccines with a glycolipid adjuvant for breast cancer, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 2517–2522.
- 8 N. Fujitani, J. Furukawa, K. Araki, T. Fujioka, Y. Takegawa, J. Piao, T. Nishioka, T. Tamura, T. Nikaido, M. Ito, Y. Nakamura and Y. Shinohara, Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers, *Proc. Natl. Acad. Sci.*, 2013, **110**, 2105–2110.
- 9 D. L. Meany, Z. Zhang, L. J. Sokoll, H. Zhang and D. W. Chan, Glycoproteomics for prostate cancer detection: changes in serum PSA glycosylation patterns, *J. Proteome Res.*, 2009, **8**, 613–619.
- 10 C. B. Lebrilla and H. J. An, The prospects of glycan biomarkers for the diagnosis of diseases, *Mol. Biosyst.*, 2009, **5**, 17–20.
- 11 M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla and I. J. Amster, Oligosaccharide analysis by mass spectrometry: a review of recent developments, *Anal Chem*, 2013, **86**, 196–212.
- 12 X. Zheng, X. Zhang, N. S. Schocker, R. S. Renslow, D. J. Orton, J. Khamsi, R. A. Ashmus, I. C. Almeida, K. Tang, C. E. Costello, R. D. Smith, K. Michael and E. S. Baker, Enhancing glycan isomer separations with metal ions and positive and negative polarity ion mobility spectrometry-mass spectrometry analyses, *Anal. Bioanal. Chem.*, 2017, **409**, 467–476.

- 13 V. N. Reinhold, B. B. Reinhold and C. E. Costello, Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID, *Anal Chem*, 1995, **67**, 1772–1784.
- 14 E. Mirgorodskaya, N. G. Karlsson, C. Sihlbom, G. Larson and C. L. Nilsson, Cracking the Sugar Code by Mass Spectrometry: An Invited Perspective in Honor of Dr. Catherine E. Costello, Recipient of the 2017 ASMS Distinguished Contribution Award, *J. Am. Soc. Mass Spectrom.*, , DOI:10.1007/s13361-018-1912-3.
- 15 X. Yu, Y. Jiang, Y. Huang, C. E. Costello and C. Lin, Detailed Glycan Structural Characterization by Electronic Excitation Dissociation, *Anal Chem*, 2013, **85**, 10017–10021.
- 16 D. J. Ashline, A. J. S. Hanneman, H. Zhang and V. N. Reinhold, Structural Documentation of Glycan Epitopes: Sequential Mass Spectrometry and Spectral Matching, *J Am Soc Mass Spectrom*, 2014, **25**, 444–453.
- 17 A. J. Lapadula, P. J. Hatcher, A. J. Hanneman, D. J. Ashline, H. Zhang and V. N. Reinhold, Congruent Strategies for Carbohydrate Sequencing. 3. OSCAR: An Algorithm for Assigning Oligosaccharide Topology from MS_n Data, *Anal Chem*, 2005, **77**, 6271–6279.
- 18 D. J. Harvey, C. A. Scarff, M. Edgeworth, W. B. Struwe, K. Pagel, K. Thalassinou, M. Crispin and J. Scrivens, Travelling-wave ion mobility and negative ion fragmentation of high-mannose N-glycans, *J Mass Spectrom*, 2016, **51**, 219–235.

- 19 W. R. Alley Jr. and M. V. Novotny, Structural glycomic analyses at high sensitivity: a decade of progress, *Annu Rev Anal Chem*, 2013, **6**, 237–265.
- 20 C. Konda, F. Londry, B. Bendiak and Y. Xia, Assignment of the Stereochemistry and Anomeric Configuration of Sugars within Oligosaccharides via Overlapping Disaccharide Ladders using MSn, *J Am Soc Mass Spectrom*, 2014, **25**, 1441–1450.
- 21 C. Konda, B. Bendiak and Y. Xia, Differentiation of the Stereochemistry and Anomeric Configuration for 1-3 Linked Disaccharides Via Tandem Mass Spectrometry and 18O-labeling, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 347–358.
- 22 B. Bendiak and T. T. Fang, Assignment of the stereochemistry and anomeric configuration of structurally informative product ions derived from disaccharides: infrared photodissociation of glycosyl-glycolaldehydes in the negative ion mode, *Carbohydr. Res.*, 2010, **345**, 2390–2400.
- 23 T. T. Fang and B. Bendiak, The Stereochemical Dependence of Unimolecular Dissociation of Monosaccharide-Glycolaldehyde Anions in the Gas Phase: A Basis for Assignment of the Stereochemistry and Anomeric Configuration of Monosaccharides in Oligosaccharides by Mass Spectrometry via a Key Discriminatory Product Ion of Disaccharide Fragmentation, *m / z 221*, *J. Am. Chem. Soc.*, 2007, **129**, 9721–9736.
- 24 T. T. Fang, J. Zirrolli and B. Bendiak, Differentiation of the anomeric configuration and ring form of glucosyl-glycolaldehyde anions in the gas phase by mass spectrometry: isomeric discrimination between *m/z* 221 anions derived from

- disaccharides and chemical synthesis of m/z 221 standards, *Carbohydr. Res.*, 2007, **342**, 217–235.
- 25 B. Domon and C. E. Costello, A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates, *Glycoconj J*, 1988, **5**, 397–409.
- 26 K. A. Morrison and B. H. Clowers, Contemporary glycomic approaches using ion mobility–mass spectrometry, *Curr. Opin. Chem. Biol.*, 2018, **42**, 119–129.
- 27 Y. Park and C. B. Lebrilla, Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides, *Mass Spectrom Rev*, 2005, **24**, 232–264.
- 28 L. Veillon, Y. Huang, W. Peng, X. Dong, B. G. Cho and Y. Mechref, Characterization of isomeric glycan structures by LC-MS/MS, *ELECTROPHORESIS*, 2017, **38**, 2100–2114.
- 29 G. E. . Hofmeister, Z. . Zhou and J. A. Leary, Linkage Position Determination in Lithium-Cationized Disaccharides: Tandem Mass Spectrometry and Semiempirical Calculations, *J Am Chem Soc*, 1991, **113**, 5964–5970.
- 30 W. B. Struwe, C. Baldauf, J. Hofmann, P. M. Rudd and K. Pagel, Ion mobility of deprotonated oligosaccharide isomers - evidence for gas-phase charge migration, *Chem Commun*, 2016, **52**, 12353–12356.
- 31 M. M. Gaye, G. Nagy, D. E. Clemmer and N. L. B. Pohl, Multidimensional Analysis of 16 Glucose Isomers by Ion Mobility Spectrometry, *Anal. Chem.*, , DOI:10.1021/acs.analchem.5b04280.

- 32 N. Khanal, C. Masellis, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS, *Anal Chem*, 2017, **89**, 7601–7606.
- 33 C. J. Gray, B. Thomas, R. Upton, L. G. Migas, C. E. Eyers, P. Barran and S. L. Flitsch, Applications of ionmobility mass spectrometry for high throughput, high resolution glycan analysis, *Biochim. Biophys. Acta*, 2016, **1860**, 1688–1709.
- 34 C. Masellis, N. Khanal, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification, *J Am Soc Mass Spectrom*, 2017, DOI: 10.1007/s13361-017-1728-6.
- 35 E. Mucha, A. I. González Flórez, M. Marianski, D. A. Thomas, W. Hoffmann, W. B. Struwe, H. S. Hahm, S. Gewinner, W. Schöllkopf, P. H. Seeberger, G. von Helden and K. Pagel, Glycan fingerprinting using cold-ion infrared spectroscopy, *Angew Chem Int Ed*, 2017, **56**, 11248–11251.
- 36 B. Schindler, L. Barnes, G. Renois, C. Gray, S. Chambert, S. Fort, S. Flitsch, C. Loison, A.-R. Allouche and I. Compagnon, Anomeric memory of the glycosidic bond upon fragmentation and its consequences for carbohydrate sequencing, *Nat. Commun.*, 2017, **8**, 973.
- 37 C. J. Gray, B. Schindler, L. G. Migas, M. Picmanova, A. R. Allouche, A. P. Green, S. Mandal, M. S. Motawia, R. Sánchez-Pérez, N. Bjarnholt, B. L. Møller, A. M. Rijs, P. E. Barran, I. Compagnon, C. E. Eyers and S. L. Flitsch, Bottom-Up Elucidation of Glycosidic Bond Stereochemistry, *Anal Chem*, 2017, **89**, 4540–4549.

- 38 M. T. Campbell, D. Chen and G. L. Glish, Identifying the D-Pentoses Using Water Adduction to Lithium Cationized Molecule, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 1420–1424.
- 39 M. T. Campbell, D. Chen and G. L. Glish, Distinguishing Linkage Position and Anomeric Configuration of Glucose–Glucose Disaccharides by Water Adduction to Lithiated Molecules, *Anal. Chem.*, 2018, **90**, 2048–2054.
- 40 M. P. Campbell, C. A. Hayes, W. B. Struwe, M. R. Wilkins, K. F. D. J. Harvey, P. M. Rudd, D. Kolarich, F. Lisacek, N. G. Karlsson and N. H. Packer, *UniCarbKB: Putting the pieces together for glycomics research*, 2011.
- 41 M. P. Campbell, T. Nguyen-Khuong, C. A. Hayes, S. A. Flowers, K. Alagesan, D. Kolarich, N. H. Packer and N. G. Karlsson, Validation of the curation pipeline of UniCarb-DB: building a global glycan reference MS/MS repository, *Biochim. Biophys. Acta*, 2014, **1844**, 108–116.
- 42 C. A. Hayes, N. G. Karlsson, W. B. Struwe, F. Lisacek, P. M. Rudd, N. H. Packer and M. P. Campbell, UniCarb-DB: a database resource for glycomic discovery, *Bioinforma. Oxf. Engl.*, 2011, **27**, 1343–1344.
- 43 NIST, chemdata:glycan-library [], <http://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:glycan-library>., (accessed March 10, 2018).
- 44 S. S. Nigudkar and A. V. Demchenko, Stereocontrolled 1,2-cis glycosylation as the driving force of progress in synthetic carbohydrate chemistry, *Chem Sci*, 2015, **6**, 2687–2704.

- 45 B. J. Bythell, M. T. Abutokaikah, A. R. Wagoner, S. Guan and J. M. Rabus, Cationized Carbohydrate Gas-phase Fragmentation Chemistry, *J Am Soc Mass Spectrom*, 2017, **28**, 688–703.
- 46 J. M. Rabus, M. T. Abutokaikah, R. T. Ross and B. J. Bythell, Sodium-cationized carbohydrate gas-phase fragmentation chemistry: influence of glycosidic linkage position, *Phys Chem Chem Phys*, 2017, **19**, 25643–25652.
- 47 J.-L. Chen, H. S. Nguan, P.-J. Hsu, S.-T. Tsai, C. Y. Liew, J.-L. Kuo, W.-P. Hu and C.-K. Ni, Collision-induced dissociation of sodiated glucose and identification of anomeric configuration, *Phys Chem Chem Phys*, 2017, **19**, 15454–15462.
- 48 M. T. Abutokaikah, J. W. Frye, J. Tschampel, J. R. Rabus and B. J. Bythell, Fragmentation Pathways of Lithiated Hexose Monosaccharides, *J Am Soc Mass Spectrom*, 2018, DOI: 10.1007/s13361-018-1973-3.
- 49 Z. Zhou, S. Ogden and J. A. Leary, Linkage position determination in oligosaccharides: mass spectrometry (MS/MS) study of lithium-cationized carbohydrates, *J. Org. Chem.*, 1990, **55**, 5444–5446.
- 50 M. T. ; Cancilla, S. G. ; Penn, J. A. ; Carroll and C. B. ; Lebrilla, Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization/Fourier transform mass spectrometry, *J Am Chem Soc*, 1996, **118**, 6736–6745.
- 51 M. T. ; Cancilla, S. G. ; Penn and C. B. ; Lebrilla, Alkaline Degradation of Oligosaccharides Coupled with Matrix-Assisted Laser Desorption/Ionization

- Fourier Transform Mass Spectrometry: A Method for Sequencing Oligosaccharides, *Anal Chem*, 1998, **70**, 663–672.
- 52 B. ; Spengler, J. W. ; Dolce and R. J. Cotter, Infrared Laser Desorption Mass Spectrometry of Oligosaccharides: Fragmentation Mechanisms and Isomer Analysis, *Anal Chem*, 1990, **62**, 1721–1737.
- 53 D. J. Harvey, Fragmentation of negative ions from carbohydrates: Part 2. Fragmentation of high-mannose N-linked glycans, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 631–646.
- 54 D. J. Harvey, Fragmentation of negative ions from carbohydrates: Part 3. Fragmentation of hybrid and complex N-linked glycans, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 647–659.
- 55 D. J. Harvey and P. M. Rudd, Fragmentation of negative ions from N-linked carbohydrates. Part 5: Anionic N-linked glycans, *Int. J. Mass Spectrom.*, 2011, **305**, 120–130.
- 56 Pfenninger, A., Karas, M., B. Finke and B. Stahl, Structural analysis of underivatized neutral human milk oligosaccharides in the negative ion mode by nano-electrospray MS(n) (part 1: methodology)., 2002, **13**, 1331–40.
- 57 P. Domann, D. I. R. Spencer and D. J. Harvey, Production and fragmentation of negative ions from neutral N-linked carbohydrates ionized by matrix-assisted laser desorption/ionization, *Rapid Commun. Mass Spectrom.*, 2012, **26**, 469–479.
- 58 D. J. Brown, S. E. Stefan, G. Berden, J. D. Steill, J. Oomens, J. R. Eyler and B. Bendiak, Direct evidence for the ring opening of monosaccharide anions in the gas

- phase: photodissociation of aldohexoses and aldohexoses derived from disaccharides using variable-wavelength infrared irradiation in the carbonyl stretch region, *Carbohydr. Res.*, 2011, **346**, 2469–2481.
- 59 R. A. Doohan, C. A. Hayes, B. Harhen and N. G. Karlsson, Negative Ion CID Fragmentation of O-linked Oligosaccharide Aldoses—Charge Induced and Charge Remote Fragmentation, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 1052–1062.
- 60 C. Ashwood, C.-H. Lin, M. Thaysen-Andersen and N. H. Packer, Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS, *J. Am. Soc. Mass Spectrom.*, 2018, 1–16.
- 61 D. J. Harvey, Fragmentation of negative ions from carbohydrates: Part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 622–630.
- 62 B. J. Bythell, J. M. Rabus, A. R. Wagoner, M. T. Abutokaikah and P. Maitre, Sequence Ion Structures and Dissociation Chemistry of Deprotonated Sucrose Anions, *J Am Soc Mass Spectrom*, 2018, DOI: 10.1007/s13361-018-2065-0.
- 63 P. J. Stephens, J. F. Devlin, C. F. Chabalowski and M. J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, *J Phys Chem*, 1994, **98**, 11623–11627.
- 64 Y. Zhao and D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states,

- and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
- 65 Y. Zhao, N. E. Schultz and D. G. Truhlar, Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions, *J Chem Phys*, 2005, **123** (16), 161103.
- 66 C. Møller and M. S. Plesset, A note on an approximation treatment for many-electron systems, *Phys Rev E*, 1934, **46**, 618–622.
- 67 J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. L. Caer, J. M. Ortega, F. Glotin and P. Maitre, Gas Phase Infrared Spectroscopy of Selectively Prepared Ions, *Phys. Rev. Lett.*, 2002, **89**, 273002.
- 68 MacAleese, Luke and Maître, Philippe, Infrared spectroscopy of organometallic ions in the gas phase: From model to real world complexes, *Mass Spectrom. Rev.*, 2007, **26**, 583–605.
- 69 N. C. Polfer and J. Oomens, Reaction products in mass spectrometry elucidated with infrared spectroscopy, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3804–3817.
- 70 J. M. Bakker, T. Besson, J. Lemaire, D. Scuderi and P. Maître, Gas-Phase Structure of a π -Allyl–Palladium Complex: Efficient Infrared Spectroscopy in a 7 T Fourier Transform Mass Spectrometer, *J. Phys. Chem. A*, 2007, **111**, 13415–13424.
- 71 R. Prazeres, F. Glotin, C. Insa, D. A. Jaroszynski and J. M. Ortega, Two-colour operation of a Free-Electron Laser and applications in the mid-infrared, *Eur. Phys. J. - At. Mol. Opt. Plasma Phys.*, 1998, **3**, 87–93.

- 72 A. Supady, adrianasupady/fafoom, <https://github.com/adrianasupady/fafoom>, (accessed August 25, 2017).
- 73 M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α -Maltose, *J Chem Theory Comput*, 2016, **12**, 6157–6168.
- 74 A. Supady, V. Blum and C. Baldauf, First-Principles Molecular Structure Search with a Genetic Algorithm, *J. Chem. Inf. Model.*, 2015, **55**, 2338–2348.
- 75 T. A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *J. Comput. Chem.*, 1996, **17**, 490–519.
- 76 T. A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comput. Chem.*, 1996, **17**, 520–552.
- 77 T. A. Halgren, Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94, *J. Comput. Chem.*, 1996, **17**, 553–586.
- 78 T. A. Halgren and R. B. Nachbar, Merck molecular force field. IV. conformational energies and geometries for MMFF94, *J. Comput. Chem.*, 1996, **17**, 587–615.
- 79 T. A. Halgren, Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules, *J. Comput. Chem.*, 1996, **17**, 616–641.
- 80 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji,

- M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision E.01*, Gaussian, Inc., Wallingford CT, 2009.
- 81 A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 82 C. M. Leavitt, A. F. DeBlase, C. J. Johnson, M. van Stipdonk, A. B. McCoy and M. A. Johnson, Hiding in Plain Sight: Unmasking the Diffuse Spectral Signatures of the Protonated N-Terminus in Isolated Dipeptides Cooled in a Cryogenic Ion Trap, *J. Phys. Chem. Lett.*, 2013, **4**, 3450–3457.
- 83 N. Yang, C. H. Duong, P. J. Kelleher, M. A. Johnson and A. B. McCoy, Isolation of site-specific anharmonicities of individual water molecules in the $I^-\cdot(H_2O)_2$ complex using tag-free, isotopomer selective IR-IR double resonance, *Chem. Phys. Lett.*, 2017, **690**, 159–171.

- 84 J. J. Talbot, X. Cheng, J. D. Herr and R. P. Steele, Vibrational Signatures of Electronic Properties in Oxidized Water: Unraveling the Anomalous Spectrum of the Water Dimer Cation, *J. Am. Chem. Soc.*, 2016, **138**, 11936–11945.
- 85 J. A. Fournier, C. T. Wolke, M. A. Johnson, T. T. Odbadrakh, K. D. Jordan, S. M. Kathmann and S. S. Xantheas, Snapshots of Proton Accommodation at a Microscopic Water Surface: Understanding the Vibrational Spectral Signatures of the Charge Defect in Cryogenically Cooled $\text{H}^+ (\text{H}_2\text{O})_{n=2-28}$ Clusters, *J. Phys. Chem. A*, 2015, **119**, 9425–9440.
- 86 J. D. Herr, J. Talbot and R. P. Steele, Structural Progression in Clusters of Ionized Water, $(\text{H}_2\text{O})_{n=1-5}^+$, *J. Phys. Chem. A*, 2015, **119**, 752–766.
- 87 S.R. Maple and A. Allerhand, Detailed tautomeric equilibrium of aqueous D-glucose. Observation of six tautomers by ultrahigh resolution carbon-13 NMR, *J Am Chem Soc*, 1987, **109**, 3168–3169.
- 88 J. M. Los, L. B. Simpson and K. Wiesner, The Kinetics of Mutarotation of D-Glucose with Consideration of an Intermediate Free-aldehyde Form, *J Am Chem Soc*, 1956, **78**, 1564–1568.
- 89 J. P. Dworkin and S. L. Miller, *Carbohydr Res*, 2000, **329**, 359–365.
- 90 L. D. Hayward and S. J. Angyal, *Carbohydr Res*, 1977, **53**, 13–20.
- 91 B. E. Lewis and V. L. Schramm, *J Am Chem Soc*, 2003, **125**, 7872–7877.
- 92 N. C. Polfer, J. Oomens, S. Suhai and B. Paizs, Infrared spectroscopy and theoretical studies on gas-phase protonated leu-enkephalin and its fragments: direct

- experimental evidence for the mobile proton, *J. Am. Chem. Soc.*, 2007, **129**, 5887–5897.
- 93 X. Chen, L. Yu, J. D. Steill, J. Oomens and N. C. Polfer, Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: oligoglycine b(2)-b(8), *J. Am. Chem. Soc.*, 2009, **131**, 18272–18282.
- 94 C. Bleiholder, S. Osburn, T. D. Williams, S. Suhai, M. Van Stipdonk, A. G. Harrison and B. Paizs, Sequence-scrambling fragmentation pathways of protonated peptides., *J. Am. Chem. Soc.*, 2008, **130**, 17774–89.
- 95 U. Erlekam, B. J. Bythell, D. Scuderi, M. Van Stipdonk, B. Paizs and P. Maître, Infrared spectroscopy of fragments of protonated peptides: direct evidence for macrocyclic structures of b5 ions, *J. Am. Chem. Soc.*, 2009, **131**, 11503–11508.
- 96 B. J. Bythell, P. Maitre and B. Paizs, Cyclization and Rearrangement Reactions of a_n Fragment Ions of Protonated Peptides, *J Am Chem Soc*, 2010, **132**, 14766–14779.
- 97 B. J. Bythell, O. Hernandez, V. Steinmetz, B. Paizs and P. Maître, Tyrosine side-chain catalyzed proton transfer in the YG a2 ion revealed by theory and IR spectroscopy in the ‘fingerprint’ and XH (X=C, N, O) stretching regions, *Int. J. Mass Spectrom.*, 2012, **316–318**, 227–234.
- 98 J. Zhao, J. K.-C. Lau, J. Grzetic, U. H. Verkerk, J. Oomens, K. W. M. Siu and A. C. Hopkinson, Structures of a_n^* Ions Derived from Protonated Pentaglycine and Pentaalanine: Results from IRMPD Spectroscopy and DFT Calculations, *J. Am. Soc. Mass Spectrom.*, 2013, **24**, 1957–1968.

3.9 Supporting Information

Figure 3S1:

Energy-resolved tandem mass spectra of deprotonated lactose-(glucose- $^{13}\text{C}_6$). KEY: m/z 347, [lactose-(glucose- $^{13}\text{C}_6$)-H] $^-$; m/z 167, Z_1 ($^{13}\text{C}_6\text{H}_9\text{O}_5^-$); m/z 179, C_1 ($\text{C}_6\text{H}_{11}\text{O}_6^-$); m/z 149, $Z_1\text{-H}_2\text{O}$ ($^{13}\text{C}_6\text{H}_7\text{O}_4^-$); m/z 161, B_1 or $C_1\text{-H}_2\text{O}$ ($\text{C}_6\text{H}_9\text{O}_5^-$); m/z 285, $^{0,2}A_2$ ($^{13}\text{C}_4\text{C}_6\text{H}_{17}\text{O}_9^-$); m/z 267, $^{0,2}A_2\text{-H}_2\text{O}$ ($^{13}\text{C}_4\text{C}_6\text{H}_{15}\text{O}_8^-$)

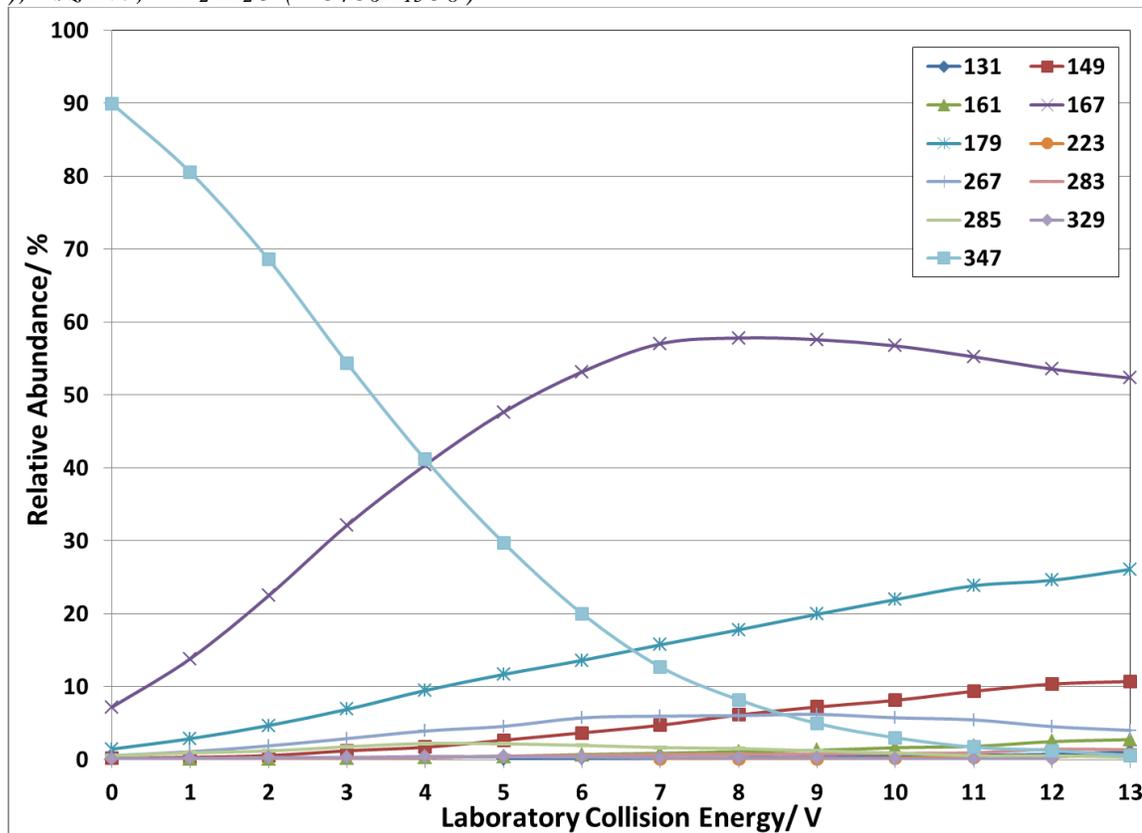


Figure 3S2:

Relative energies of selected low-energy conformers of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- α -D-Glucose anomer) calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol⁻¹: $\Delta E_{el+ZPE,0K}$ (ΔG_{298K})

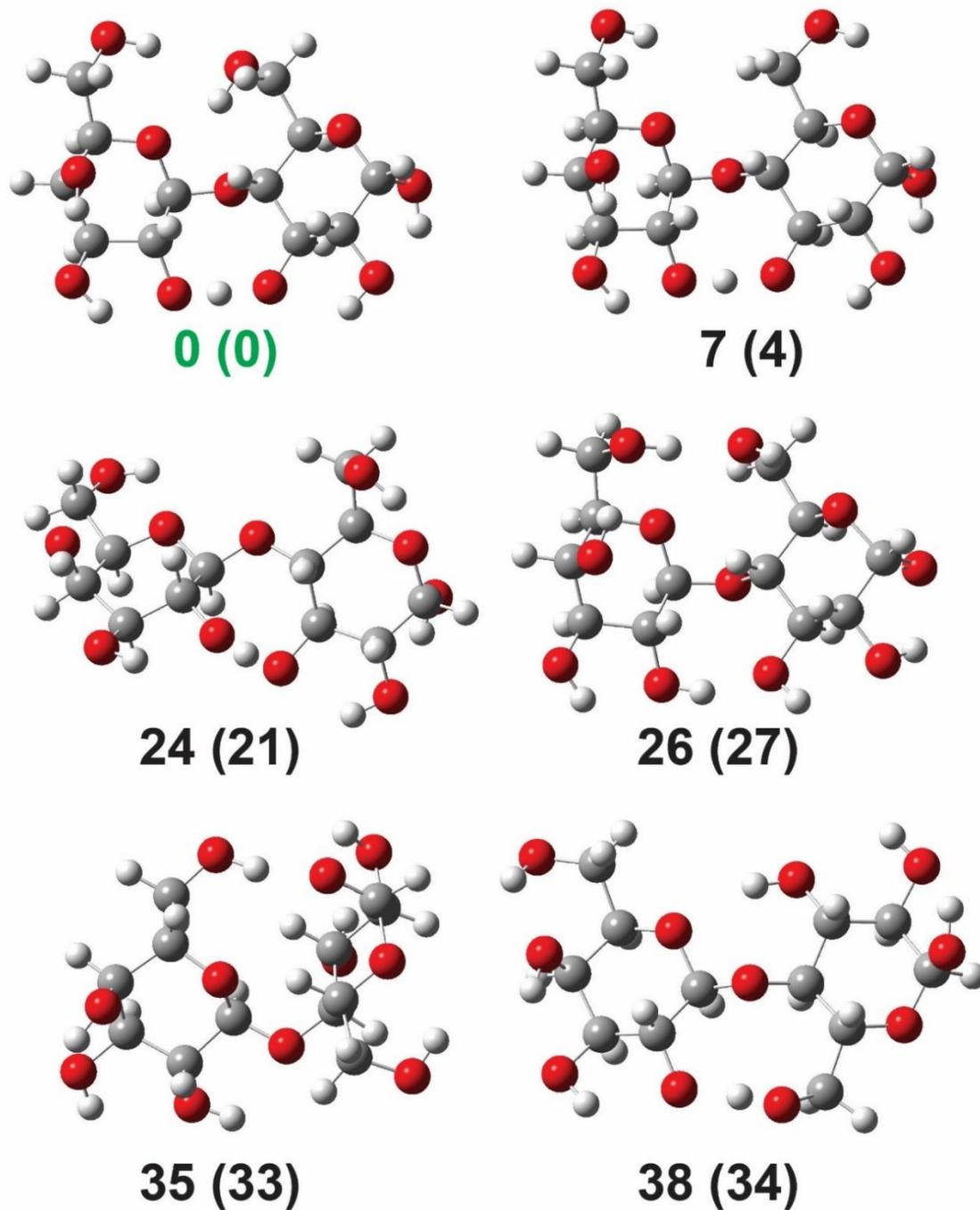


Figure 3S3:

Relative energies of selected low-energy conformers of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose anomer) calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The β -anomer GM is 5.0 (3.6) kJ mol^{-1} , $\Delta E_{el+ZPE,0K}$ (ΔG_{298}) above the α -anomer (Table 3-1)

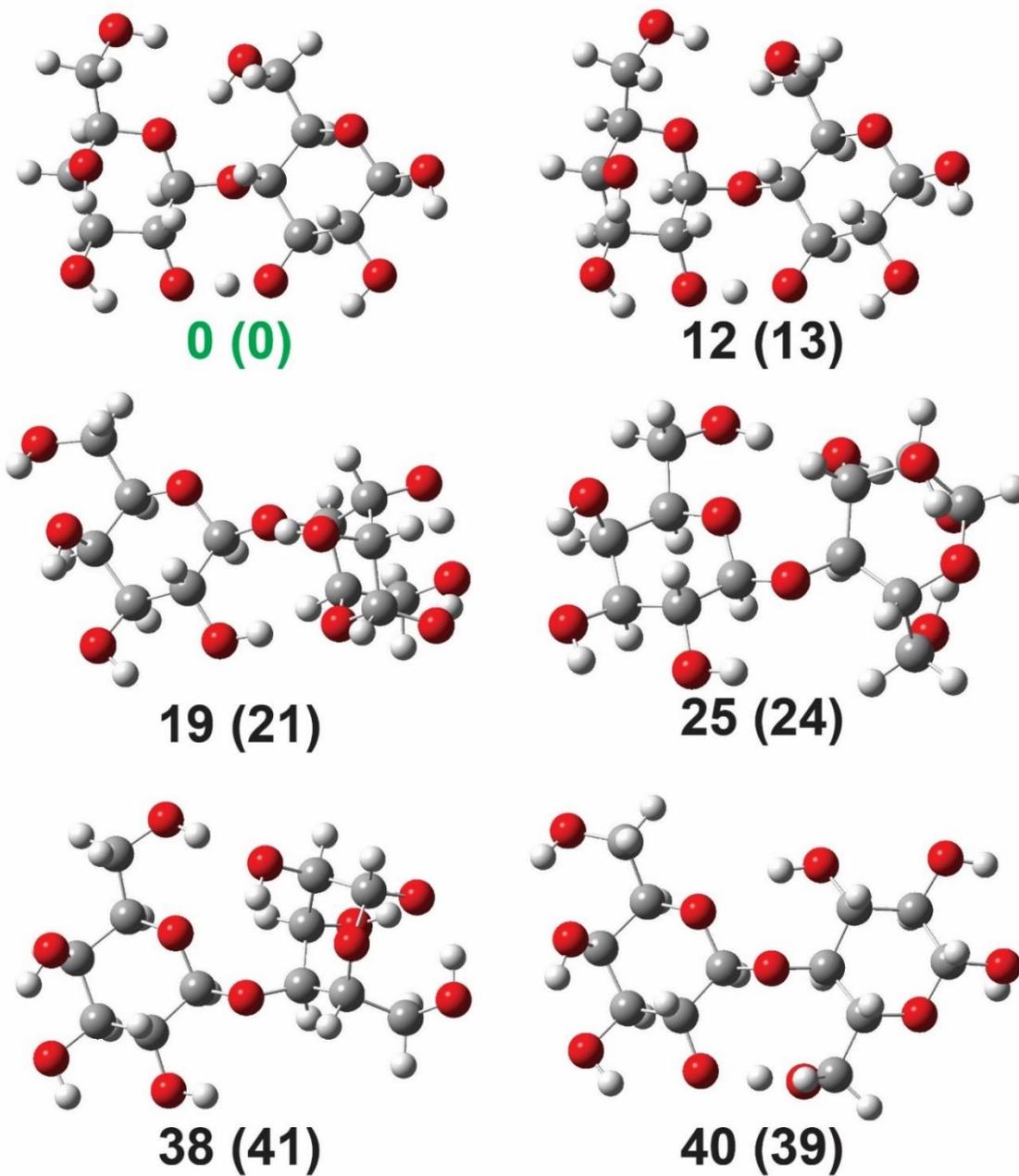


Figure 3S4:

Relative energies of selected low-energy conformers of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose ring-open) calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The ring-open GM is 32.1 (28.1) kJ mol^{-1} , $\Delta E_{el+ZPE,0K}$ (ΔG_{298}) above the α -anomer (Table 3-1) and 27.1 (24.5) kJ mol^{-1} above the β -anomer (Table 3-2)

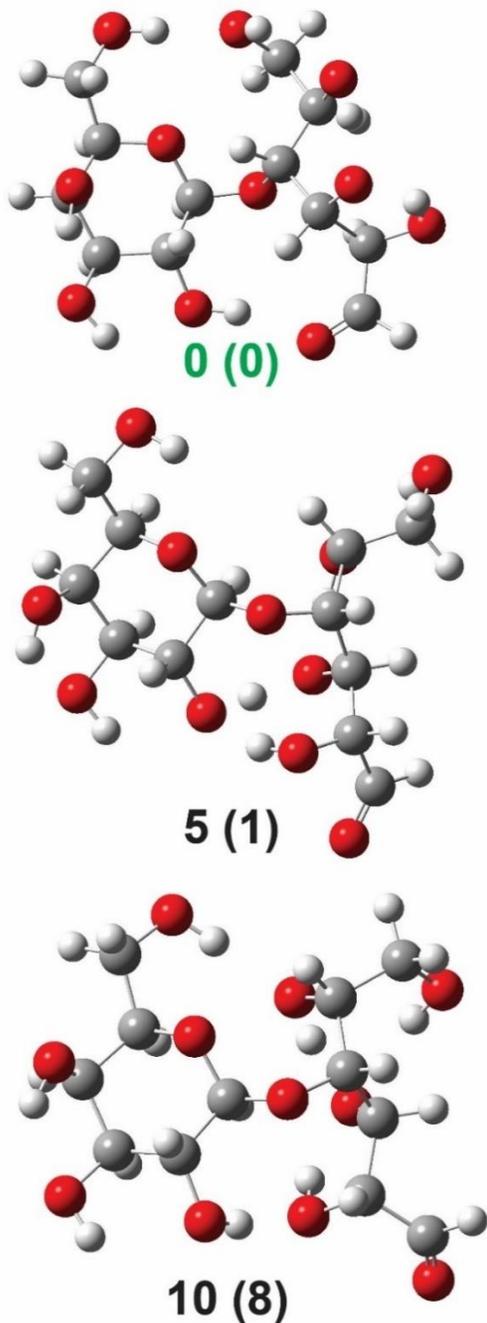


Figure 3S5:

Comparison to the 2 lowest energy structures of deprotonated β -D-galactopyranosyl-(1 \rightarrow 4)-D-glucose (m/z 341.12 ($C_{12}H_{21}O_{11}^-$)) with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures. Panels a and b are ring-open forms, panels c and d β -glucose anomers, while panels e and f are α -glucose forms. Relative energies calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}) and are normalized to the α -glucose forms. Structural changes between the 2 levels of theory are negligible

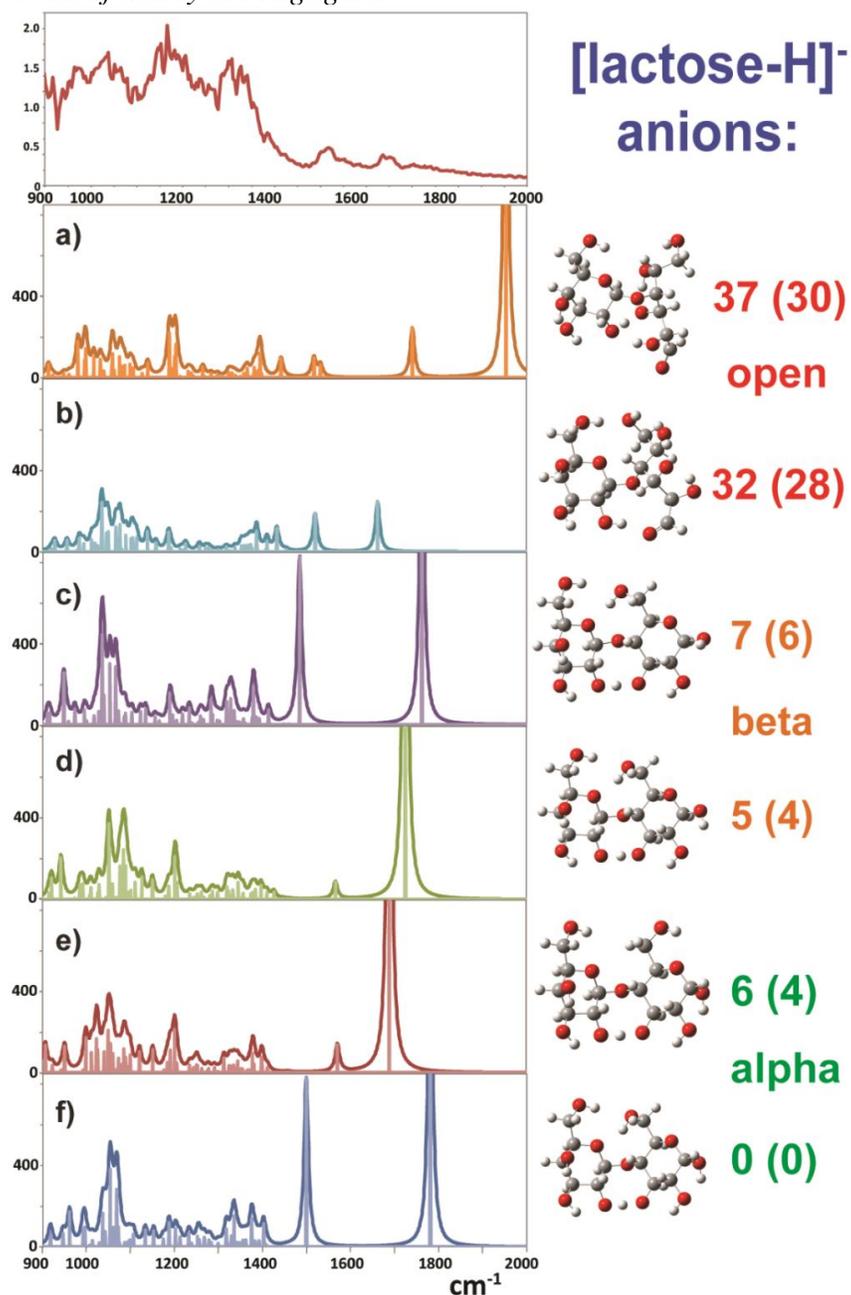


Figure 3S6:

Selected Z_1 anion low-energy conformers calculated at the M06-2X/6-31++G(d,p) level of theory. a) deprotonated 1,4-anhydroglucose structures, [1,4-anhydroglucose-H]⁻; b) deprotonated 3,4-anhydroglucose, [3,4-anhydro- β -glucose-H]⁻; c) deprotonated 3,4-anhydroglucose [3,4-anhydro- α -glucose-H]⁻; d) deprotonated, ring-open aldehyde isomers. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K})

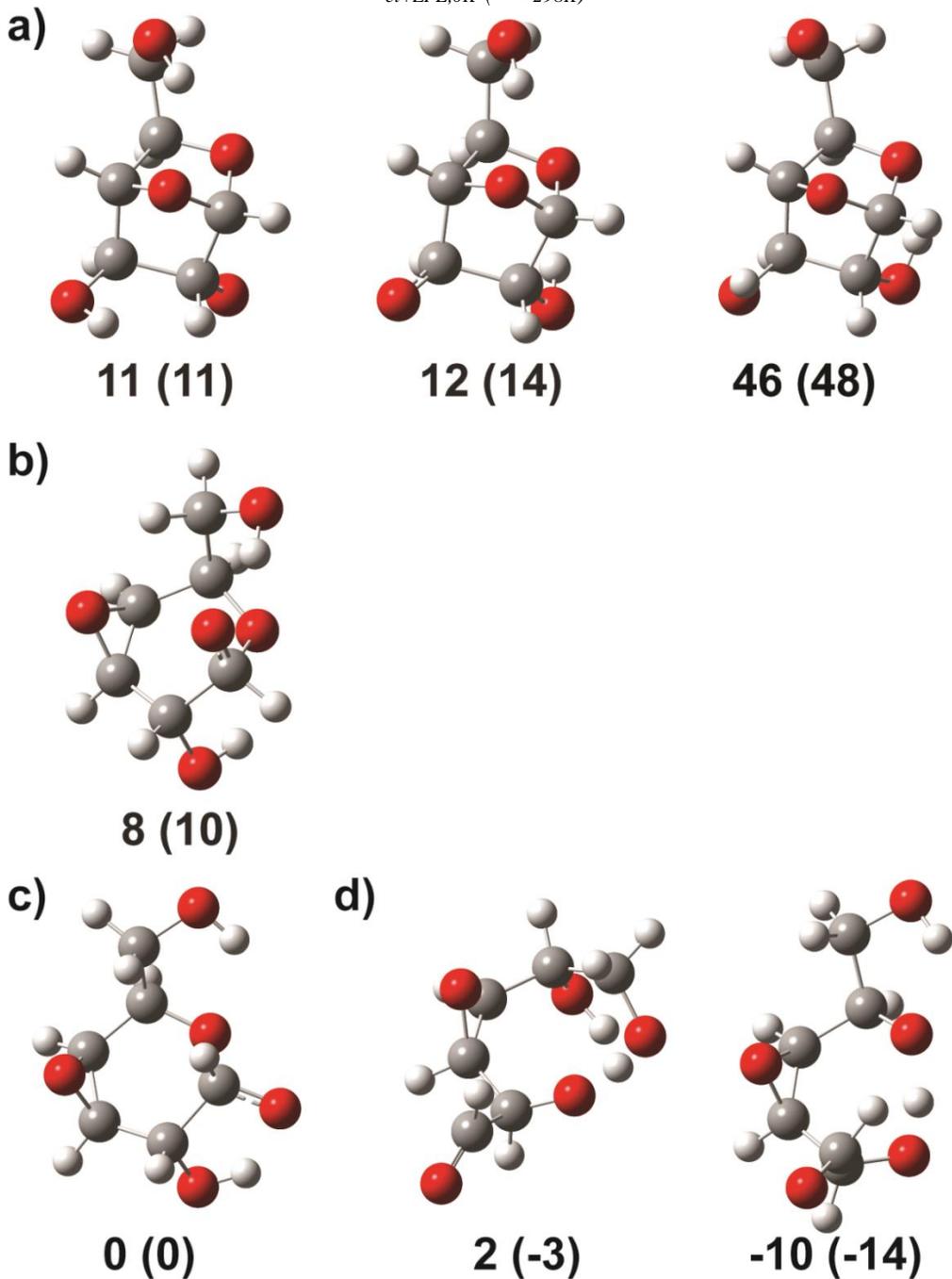


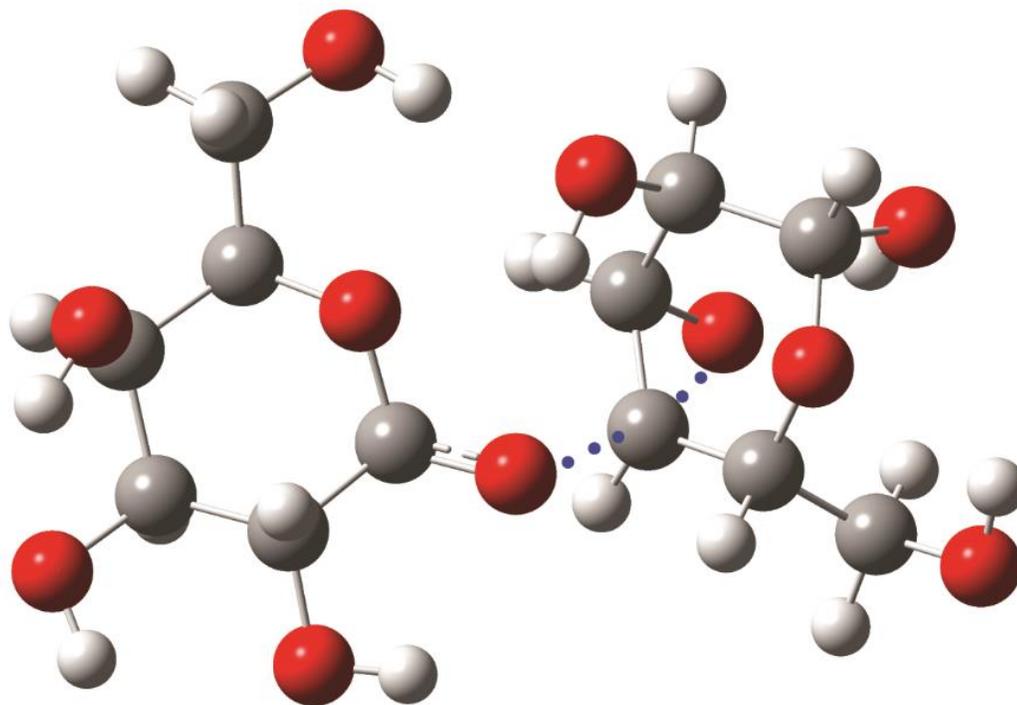
Table 3S1:

Relative energies of the transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)-D-Glucose ring-open) calculated at the M06-2X/6-31++G(d,p) level of theory. The ring-open GM is 32.1 kJ mol⁻¹ ($\Delta E_{el+ZPE,0K}$) above the α -anomer (Table 3-1)

Minima and TSs	E_{el}/H	E_{el+ZPE}/H	$\Delta E_{el+ZPE,0K}/$ kJ mol ⁻¹	$\Delta H_{298}/$ kJ mol ⁻¹	$\Delta G_{298}/$ kJ mol ⁻¹	$\Delta S_{298}/$ J K ⁻¹ mol ⁻¹
GM	-1296.986125	-1296.624955	0	0	0	0
C ₁ -Z ₁ TS_1	-1296.909679	-1296.551085	193.9	196.5	190.0	21.8
C ₁ -Z ₁ TS_2	-1296.908922	-1296.550436	195.6	198.7	190.1	28.8
^{0,2} A ₂ formation TS_1	-1296.949413	-1296.591725	87.2	89.5	83.3	20.7
^{0,2} A ₂ formation TS_2	-1296.94447	-1296.591076	88.9	92.0	83.5	28.3
Ring-forming TS (α)	-1296.920038	-1296.562144	164.9	161.1	169.1	-26.8
Ring-forming TS (β)	-1296.924404	-1296.566837	152.6	148.7	157.2	-28.2

Figure 3S7:

The higher-energy 3,4-anhydroglucose-forming C_1-Z_1 transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The reaction coordinate is illustrated with blue dots



164 (162)

Figure 3S8:

The lowest-energy 3,4-anhydroglucose-forming C_1-Z_1 transition structure of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- α -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The reaction coordinate is illustrated with blue dots

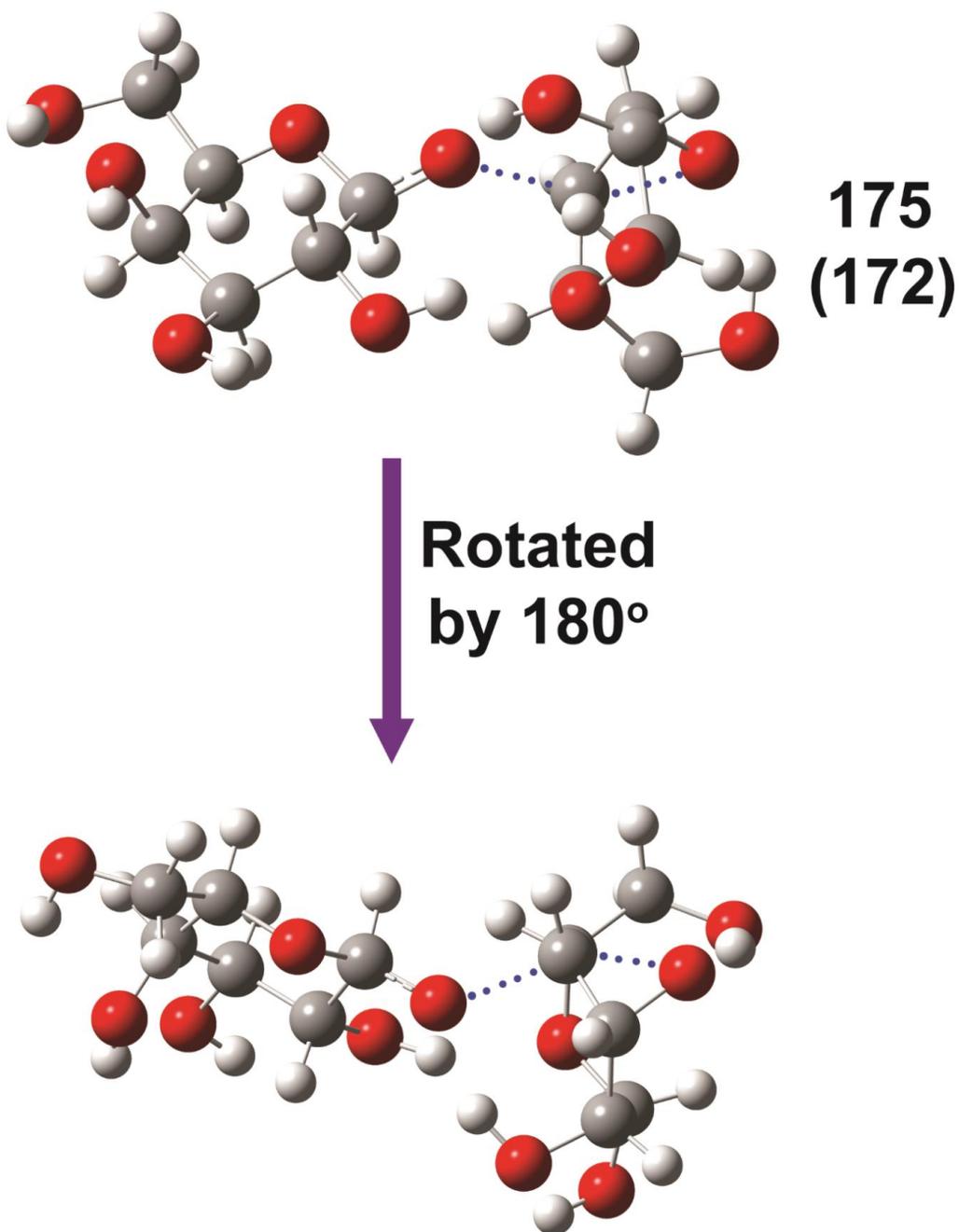


Figure 3S9:

The low-energy Z_1 -ring-opening isomerization transition structures generated from of the a) deprotonated 3,4-anhydro- α -glucose anion and b) deprotonated 3,4-anhydro- β -glucose anion structures. Relative energies calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The reaction coordinate is illustrated with blue dots

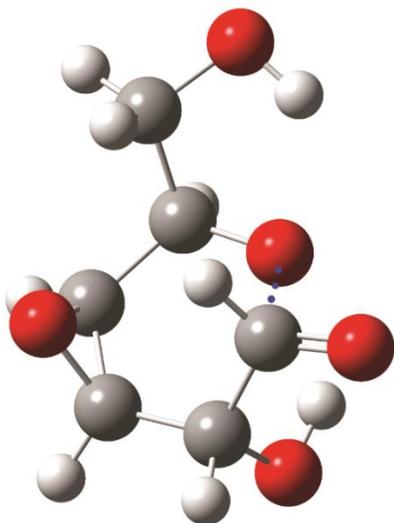
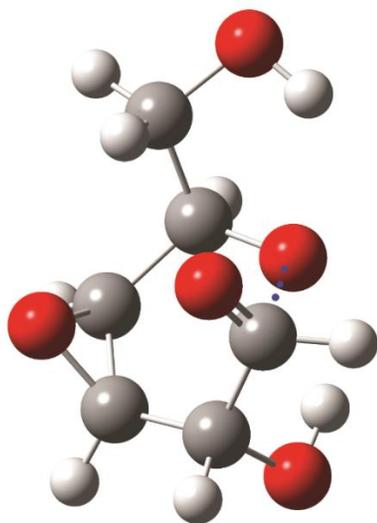
a)**23 (23)****b)****17 (16)**

Figure 3S10:

The low-energy C₁-ring-opening isomerization transition structures generated from the deprotonated β-D-galactopyranose anion structures. Relative energies calculated at the M06-2X/6-31++G(d,p) level of theory. Values in kJ mol⁻¹: ΔE_{el+ZPE,0K} (ΔG_{298K}). The reaction coordinate is illustrated with blue dots

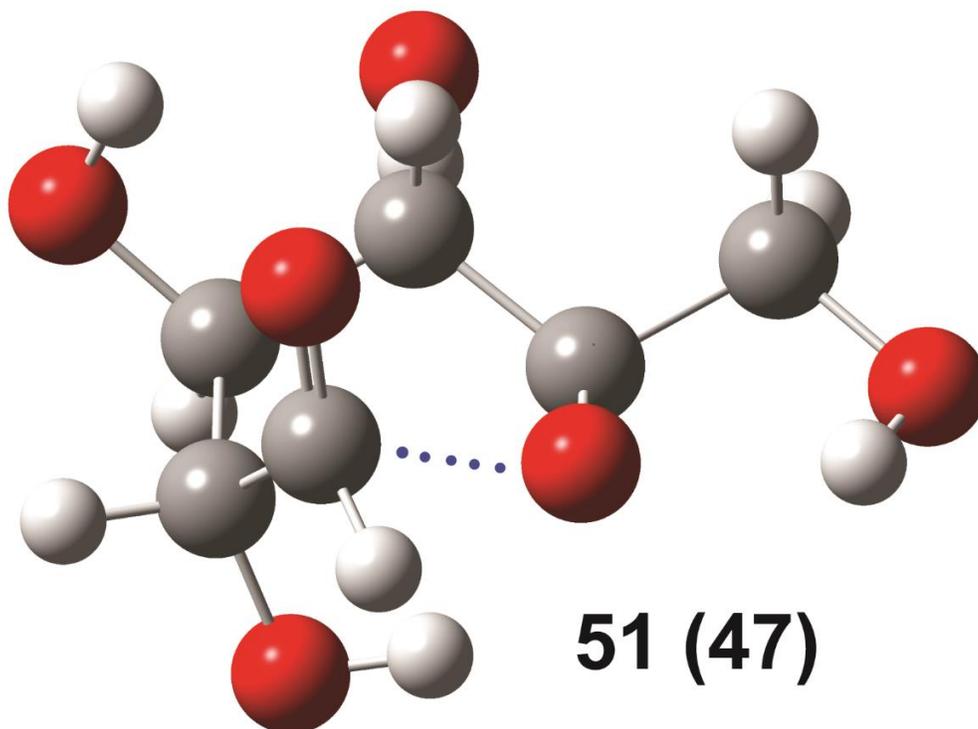


Figure 3S11:

Selected C_1 anion low-energy conformers calculated at the M06-2X/6-31++G(d,p) level of theory. a) deprotonated β -D-galactose C_1 structures, [β -D-galactose-H] $^-$; b) deprotonated, ring-open aldehyde C_1 structures. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K})

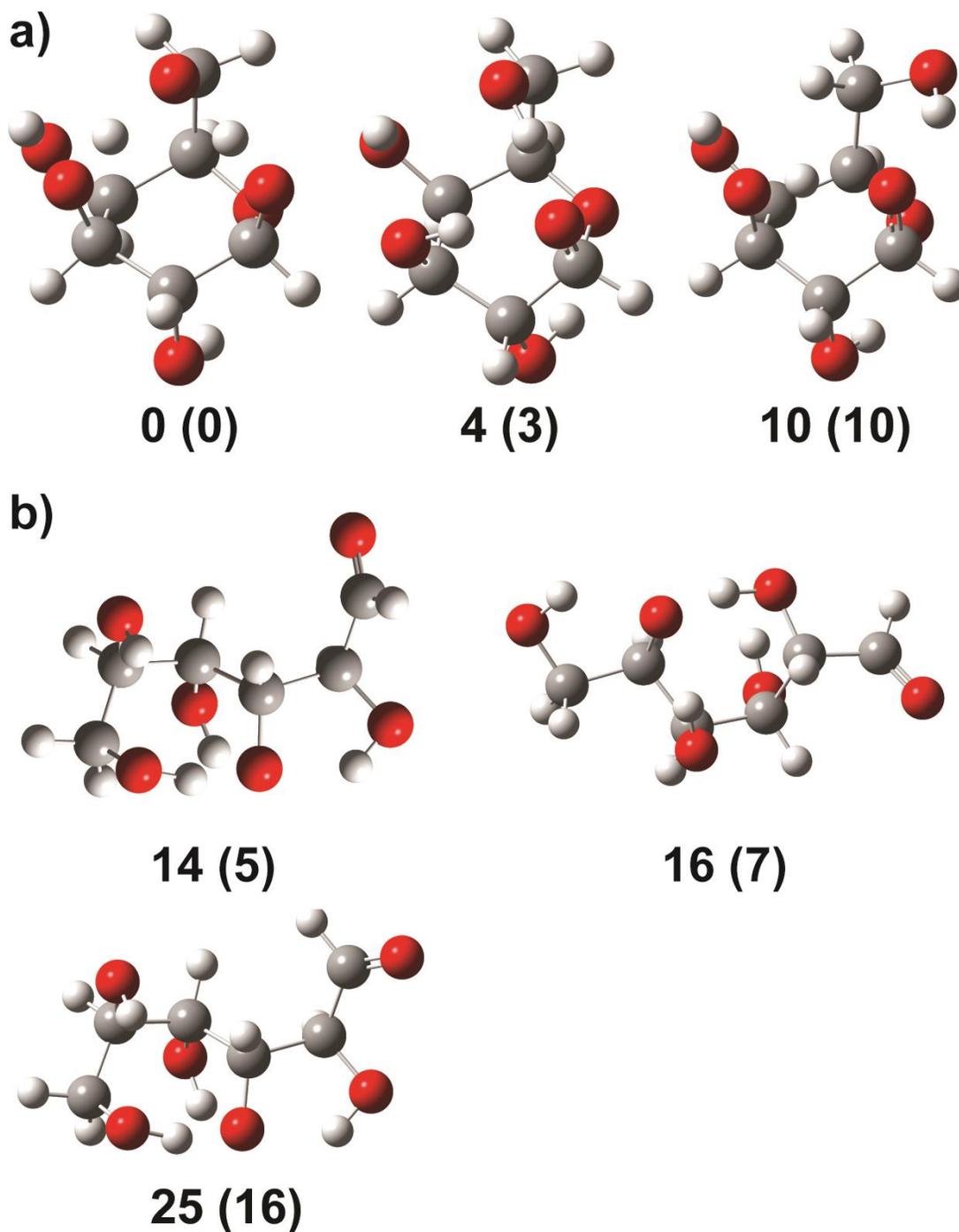


Figure 3S12:

The lowest-energy $^{0,2}A_2$ -forming transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- α -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory. a) ring-opening TS; b) $^{0,2}A_2$ formation TS_A. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The reaction coordinate is illustrated with blue dots

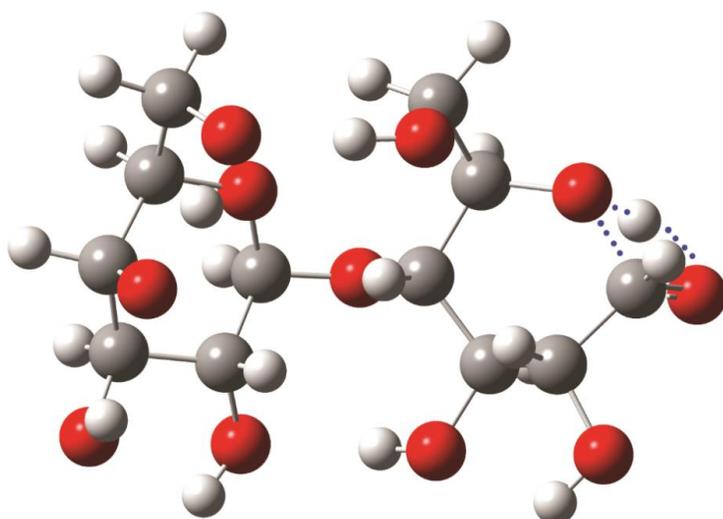
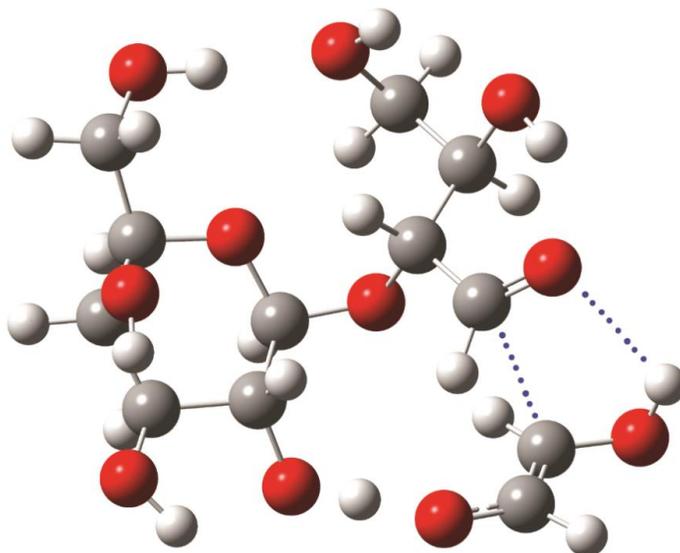
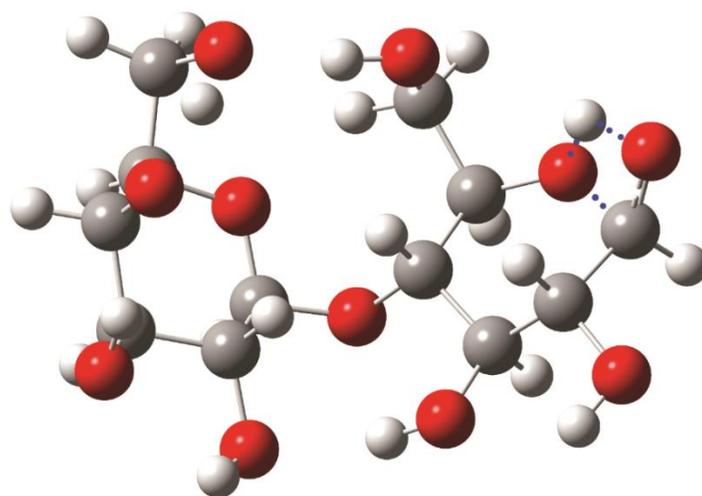
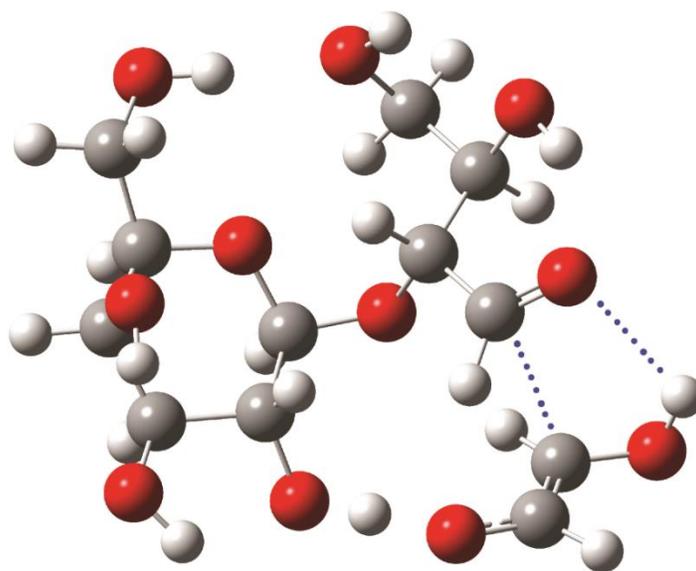
a)**197 (197)****b)****119 (111)**

Figure 3S13:

The lowest-energy 0,2A_2 -forming transition structures of deprotonated lactose (β -D-galactopyranosyl-(1 \rightarrow 4)- β -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory. a) ring-opening TS; b) 0,2A_2 formation TS_A. Values in kJ mol^{-1} : $\Delta E_{el+ZPE,0K}$ (ΔG_{298K}). The reaction coordinate is illustrated with blue dots

a)**180 (182)****b)****114 (108)**

Chapter 4 : Unravelling the Structures of Sodiated β -Cyclodextrin and its

Fragments

Jordan M. Rabus,^b Robert P. Pellegrinelli,^a Ali Hassan Abi Khodr,^a Benjamin J. Bythell,^b

Thomas R. Rizzo,^a and Eduardo Carrascosa^a

- a. Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, EPFL SB ISIC LCPM, Station 6, CH-1015 Lausanne, Switzerland
- b. Department of Chemistry and Biochemistry, Ohio University, 391 Clippinger Laboratories, Athens, Ohio 45701, United States

Reproduced from Physical Chemistry Chemical Physics **2021**, **23**, 13714-13723 with permission from the PCCP Owner Societies.

<https://doi.org/10.1039/D1CP01058A>.

4.1 Abstract

We present cryogenic infrared spectra of sodiated β -cyclodextrin $[\beta\text{-CD+Na}]^+$, a common cyclic oligosaccharide, and its main dissociation products upon collision-induced dissociation (CID). We characterize the parent ions using high-resolution ion mobility spectrometry and cryogenic infrared action spectroscopy, while the fragments are characterized by their mass and cryogenic infrared spectra. We observe sodium-cationized fragments that differ in mass by 162 u, corresponding to B_n/Z_m ions. For the m/z 347 product ion, electronic structure calculations are consistent with the formation of the lowest energy 2-ketone B_2 ion structure. For the m/z 509 product ion, both the calculated 2-ketone B_3 and the Z_3 structures show similarities with the experimental spectrum. The theoretical structures most consistent with the spectrum of the m/z 671

ions is a slightly higher energy 2-ketone B₄ structure. Overall, the data suggest a consistent formation mechanism for all the observed fragments.

4.2 Introduction

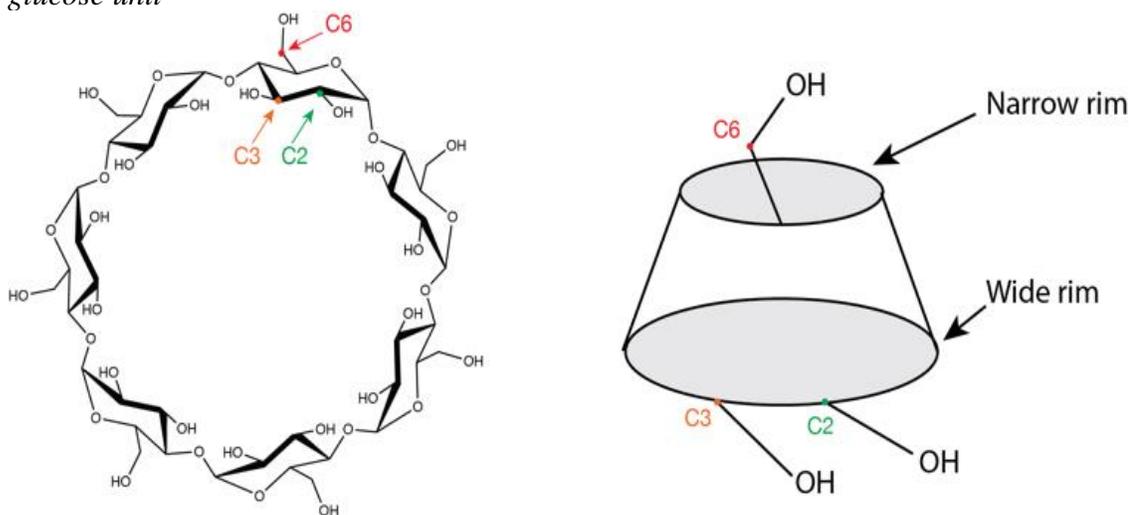
The synthesis and biomedical applications of large cyclic molecular systems have received growing attention, mainly due to their ability to bind selectively to substrates based on non-covalent interactions.[1] Crown ethers are arguably the most popular and best-investigated host molecules for complexing metal ions and are widely used in organic and supramolecular chemistry.[2, 3] Cyclic oligosaccharides occur both naturally and synthetically, and they share similar properties with crown ethers in being able to act as a host for metal cations, organometallic complexes, and biologically relevant molecules.[4] Among the advantages of these compounds is their relative ease of derivatization, allowing for a suitable functional group to be attached to the cyclic oligosaccharide, which can alter their affinity to a certain enantiomer and thus make them good chiral selectors. This is of great interest in separation techniques such as liquid chromatography and capillary electrophoreses.[5, 6]

The best-known cyclic oligosaccharides are cyclodextrins (CDs), which are produced enzymatically from starch and occur in three different configurations: α , β , and γ , made up of 6, 7, and 8 D-glucopyranose units, respectively, and linked together by α -(1-4) glycosidic bonds.[7] The non-toxicity and inexpensive synthesis of cyclodextrins has made them the molecules of choice for many applications including drug delivery, cosmetics, and food processing.[8-11] Very recently, a modified β -CD has been shown to have broad-spectrum anti-viral properties, including against herpes and Zika.[12]

The properties of cyclodextrins arise from their unique conical cylindrical structure, with the hydroxyls pointing outwards while the glycosidic oxygens and the hydrogens line the interior,[13] leading to a hydrophobic zone inside the cylinder, and a hydrophilic zone on the exterior. This allows CDs to form complexes with a wide variety of compounds, ranging from nonpolar molecules that are captured in the hydrophobic cavity, to polar molecules and ions which preferentially interact with the outer polar surface. Because of this feature, cyclodextrins are water-soluble and can be used to bring poorly soluble compounds into aqueous solution by enclosing them in their hydrophobic cavity. For β -CD, the C2 and C3 hydroxyls form a complete hydrogen-bonded network on the wide rim, while the C6 hydroxyls form a separate network on the narrow rim, resulting in a rigid structure (Figure 4-1).[14]

Figure 4-1:

Schematic representation of β -cyclodextrin, with its conical structure represented on the right. The carbon atoms connected to the hydroxyls are labeled C2, C3, and C6 for one glucose unit



Complexation of CDs with guest molecules and ions can distort these structures. It is therefore important to identify these subtle structural details and determine how they affect the physicochemical and biological properties of these species. For this purpose, techniques are needed that can isolate, control and probe molecules. Tandem mass spectrometry (MSⁿ) coupled with soft-ionization techniques, such as electrospray ionization (ESI) and matrix-assisted laser desorption (MALDI), has emerged as a key tool to investigate large molecules and complexes.[15] Mass spectrometry has thus been used to study many properties of cyclodextrins, including their host-guest interaction with various organic molecules, their binding selectivity with different metal cations, and their fragmentation patterns in the gas phase.[16-21] MS coupled to collision-induced dissociation (CID) is often used in biomedical and pharmaceutical contexts to reconstruct the primary structure and binding characteristics of complexed and functionalized CDs. For instance, fragmentation spectra can help determine the chirality of target analytes present in CD-based host-guest complexes,[22,23] as well as identify regioisomers of functionalized CDs.[20] Furthermore, studies that use CDs to promote or catalyze the synthesis of biodegradable products rely on CID-MS studies to determine the number and structural arrangement of derivative host units attached to the CD guest and consequently optimize the synthetic strategy.[24,25] The homomolecular nature of cyclodextrins makes them interesting models for CID studies because they are limited to three types of fragmentation: B/Y, C/Z, and cross-ring fragmentation.[26] Fragmentation of β -CD-metal cation complexes by CID has primarily shown a series of product cations separated by 162 u, which corresponds to a dehydrated glucose unit that requires fragmentation of

two glycosidic linkages.[26-28] In addition, other cross-ring fragmentation channels have been observed, especially for anions.[29]

Despite being a powerful tool for fragmentation analysis, MS alone is unable to provide all desirable isomer-specific information, such as preferential isomer formation upon reaction or fragmentation. In the case of cationized cyclodextrins, MS studies have been unable to address whether CID preferentially forms B_n or Z_m ions. However, the addition of a vibrational spectroscopy dimension to MS can help identify structural details of CID fragments. While vibrational spectroscopy of ions in mass spectrometers dates back to the early '80's from the work of Beauchamp and coworkers,[30, 31] the development of user-friendly infrared lasers has resulted in the reemergence of infrared (IR) action spectroscopy as a structural probe for gas-phase ions, including biomolecular ions.[32-35] Our group has recently demonstrated that cryogenic IR spectroscopy coupled with ion mobility spectrometry (IMS) can identify various isomeric glycans ranging from monosaccharides to nine-membered N-glycans.[36-39] In a very recent study we have been able to identify anomeric CID fragments from single glycans by their IR spectra.[40] The advantage of measuring spectra of cryogenically cooled ions is that it greatly reduces thermal inhomogeneous broadening, resulting in significantly enhanced spectral resolution, which is particularly important for the analysis of large molecules possessing inherently congested spectra.[33]

Here we investigate the CID products of sodiated β -CD using cryogenic IR spectroscopy. Understanding which fragmentation pathway(s) is (are) responsible for

product ions will provide insight into potential dissociation mechanisms of oligosaccharides which, until now, remain only partially understood.

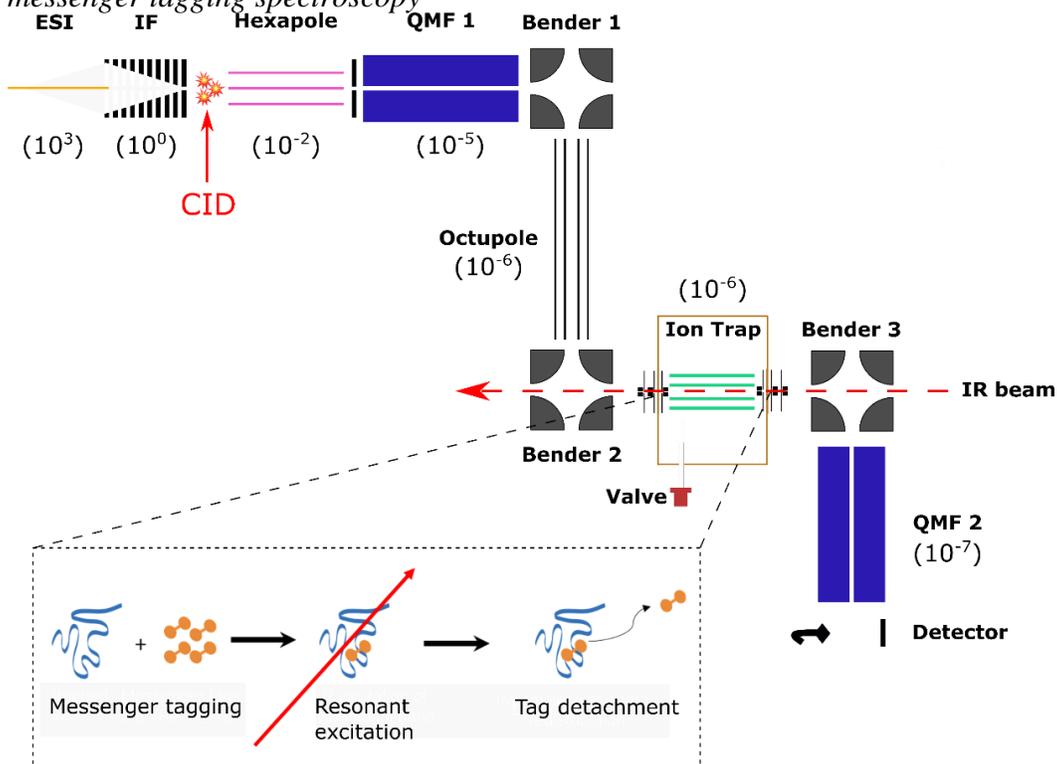
4.3 Methods

4.3.1 Experimental Details

To explore the fragmentation characteristics and corresponding structural signatures of $[\beta\text{-CD}+\text{Na}]^+$, tandem mass spectrometry, collision-induced dissociation and laser spectroscopy are combined in a homebuilt instrument, schematically depicted in Figure 4-2, the details of which have been previously reported.[41]

Figure 4-2:

Schematic representation of the experimental setup to obtain IR spectra of mass selected CID fragments under cryogenic conditions. Pressures, in mbar, are indicated in brackets for each differentially pumped section. The bottom left inset illustrates the principle of messenger tagging spectroscopy



β -Cyclodextrin (purchased from Sigma Aldrich and used with no further purification) was dissolved in a water-methanol (1:1) mixture at 0.1 mM concentration, complexed to the group I cations, and introduced into the gas phase *via* nano-electrospray (nESI). In the case of sodiated β -CD cations, the electrosprayed ions were radially confined in an RF ion funnel at ~ 1 mbar before being accelerated through a potential difference of ~ 240 V into a hexapole ion trap at a pressure of 10^{-2} mbar, creating fragments *via* CID. The trapped precursor ions and their fragments were then extracted as 100 μ s pulses at 10 Hz and passed through a quadrupole mass filter, where ions of a designated m/z range are selected. The transmitted ions were subsequently guided through two electrostatic benders and an octupole ion guide before being focused into a linear octupole ion trap that is cooled by a closed-cycle cryocooler (SHI Cryogenics). Upon entering the trap, the ions are thermalized through collisions with a He/N₂ buffer gas mixture (90:10) that was previously introduced as 250-350 μ s pulses and pre-cooled to the temperature of the copper trap housing. The trap temperature was held at 50-60K using a resistive heater attached to the copper housing to avoid condensation of N₂. To obtain single-photon infrared spectra at cryogenic conditions, the target ions were “tagged” by forming weakly bound complexes with N₂. The trapped messenger-tagged species were then irradiated with IR light from a pulsed tunable optical parametric oscillator (OPO, LaserVision). In the event of resonant absorption, the vibrational energy rapidly redistributes within the molecule, thereby detaching the weakly-bound N₂ tag. All the ions were then extracted from the trap and passed through a second quadrupole mass filter where they were detected using a Channeltron. By selecting the m/z associated with

the messenger-tagged target ion and monitoring its depletion as a function of laser wavenumber in a *laser on-off* experiment, cryogenic infrared action spectra were obtained. These spectra were normalized to the OPO power, which slightly varies over the wavelength range of the scan.

To investigate whether multiple isomers coexist in the gas phase for the studied species, analogous solutions of sodiated β -CD and α -CD were electrosprayed into a recently developed ultrahigh-resolution ion mobility spectrometer attached to a cryogenic trap and a time-of-flight mass spectrometer.[42] The ion mobility device allows for separation path lengths of over 10 meters and has demonstrated the ability to separate structurally analogous saccharide anomers. Thus, structural differences arising from differing hydrogen bonding arrangements, ring deformations (chair, boat, skew), and sites of sodiation should be well resolved.

4.3.2 Computational Approach

To understand the formation mechanism and structure of the observed fragments, theoretical simulations were carried out on the three smallest fragment ions. As previously,[43-46] simulations were performed using the genetic algorithm tool Fafoom[47, 48] to enable effective characterization of the potential energy surface. The structures were optimized using the MMFF94 force field.[49-53] This approach samples a wide range of ring structures incorporating multiple chair, boat, and skew forms, enabling a thorough interrogation of the potential energy surface. All oxygen sites had sodium cations added using custom scripts to generate starting points for the calculations. Geometry optimizations of the resulting candidate conformations were performed with

the Gaussian 09 software package[54] at the HF/3-21G, B3LYP/6-31G(d), and B3LYP/6-31+G(d,p) levels of theory.[55-57] For the putative charged fragments of the largest species with m/z 671, a combination of a preliminary PM6 optimization with a HF/3-21G single point energy calculation was performed prior to the aforementioned series of calculations. Degenerate structures were removed at each stage, and only the most competitive non-degenerate structures were optimized at each successively realistic level of theory. Additional, targeted manual adjustment and supplementation of the structural pool analyzed were performed to reduce the chance that chemically relevant species had been neglected. Subsequent simulations utilizing the CREST package,[58-60] which systemically places a sodium ion at each lone pair of the molecule, with subsequent molecular dynamics using a semiempirical engine followed by B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) calculations broadened the pool of low (and high) energy candidate structures but failed to generate new global minima. All minima were tested by vibrational analysis (all real frequencies). The potential energy surface generated combined the zero-point energy correction (ZPE) to the electronic energy (E_{el} , 0 K) for improved accuracy ($\Delta E_{el+ZPE,0K}$). The related, standard enthalpy (ΔH_{298K}), Gibbs free energy (ΔG_{298K}), and entropy (ΔS_{298K}) corrections to 298 K were also determined. Calculated B3LYP/6-31+G(d,p) vibrational frequencies were utilized for comparisons with the experimental spectra. A scaling factor of 0.955 was used for all vibrational frequencies and a 5 cm^{-1} full width at half maximum Lorentzian line shape was employed for comparison to the experimental spectra. Targeted single-point calculations at the M06-2X, and ω B97X-D levels of theory were also performed on the selected minimum

energy structures to help address variability. Subsequent targeted M06-2X and ω B97X-D optimizations and frequency calculations with 6-31+G(d,p) basis sets on m/z 671 candidates were consistent with the B3LYP data.

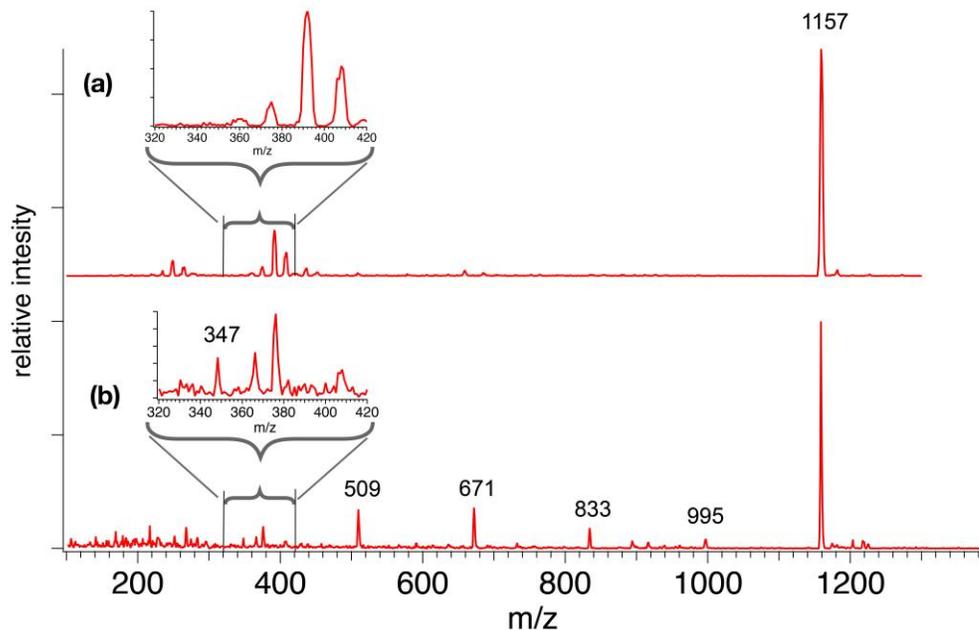
4.4 Results and Discussion

4.4.1 CID mass spectrum

Figure 4-3 presents mass spectra of sodiated β -CD under largely non-dissociative conditions (a) and CID conditions (b). The spectrum in Figure 3(b) exhibits the $[\beta\text{-CD}+\text{Na}]^+$ precursor ion mass at m/z 1157, as well as a sequence of lower mass peaks at m/z 995, 833, 671, 509 and 347.

Figure 4-3:

CID-MS of sodiated β -CD and its fragments, carried out with a potential difference of 100 V (a) and 240 V (b) in our tandem mass spectrometer. Consecutive loss of 162 m/z is observed from the parent at 240 V, while no fragments are observed at 100 V



This series indicates spacings of 162 u, a fragment moiety that has been previously observed[20, 27] and corresponds to a singly charged, dehydrated glucose monosaccharide. Detection and analysis of the monomeric dehydrated glucose fragment at m/z 185 was not attempted, mainly due to the increased number of overlapping ions in the low mass range. A minor peak at m/z 893 corresponds to the loss of 264 mass units and is ascribed to a minor cross-ring fragmentation process also observed in previous investigations.[27, 29] The masses corresponding to the three- and four-glucose unit fragments (m/z 509 and m/z 671) appear as the most intense CID products, whereas the formation of two- and six-residue fragments appear with much lower intensity.

It is difficult to determine whether the fragmentation occurs in a sequential or simultaneous manner. However, similar peaks and relative intensities are produced in a commercial time-of-flight (ToF) spectrometer (Waters Q-ToF Premier, Figure 4S1) and our home-built instrument. To further investigate this, a series of CID experiments at increasing collision energies were performed on the Q-ToF, and all the main fragments appeared at the same energy threshold, providing no evidence for a sequential process (see Figure 4S2).

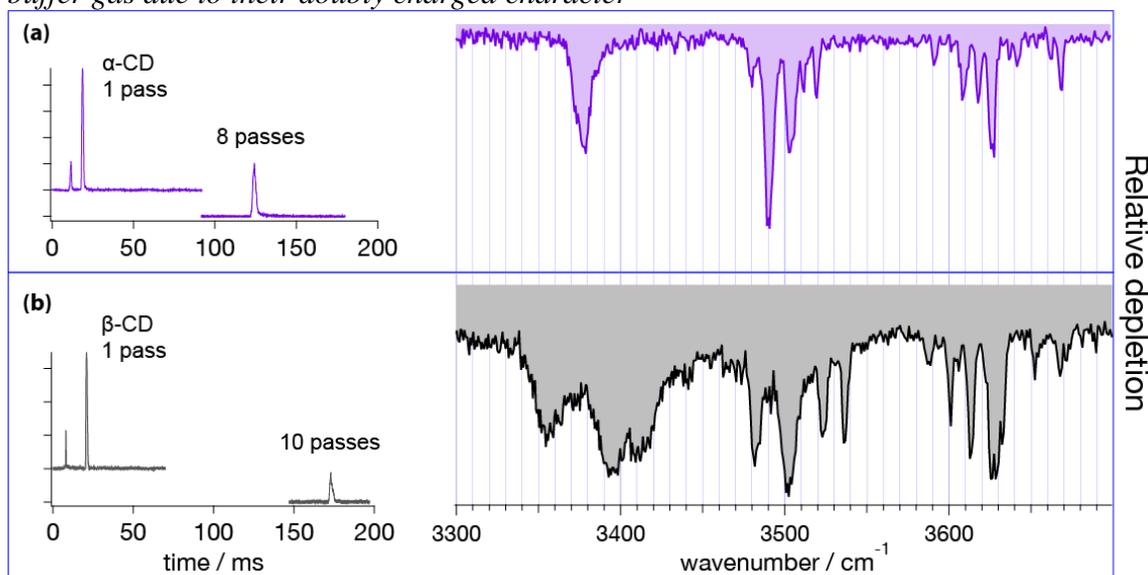
4.4.2 Ion mobility and cryogenic IR action spectra of sodiated α -CD and β -CD

Precursor [β -CD+Na]⁺ ions can adopt different conformations depending on the relative orientation of the OH groups at both the narrow and wide rims of the conical cylindrical structure (Figure 4-1). The relative orientation of these donor-acceptor hydrogen bonding networks at both edges of the molecule can potentially lead to four different combinations, depending on whether the hydrogen bonding pattern is clockwise

(cw) or counter-clockwise (cc): cc-cc, cc-cw, cw-cc, and cw-cw.[61] Distortion of these symmetric conformations through the inclusion of Na^+ might significantly increase the number of conformers. To evaluate whether electrosprayed $[\beta\text{-CD}+\text{Na}]^+$ ions form a distribution of conformers, additional experiments were carried out using a recently designed instrument that combines ultrahigh-resolution traveling-wave ion mobility spectrometry (IMS) using structures for lossless ion manipulation (SLIM),[62, 63] with cryogenic infrared spectroscopy to provide isomer-specific structural identification of complex molecules.[42] A schematic representation of this instrument is shown in Figure 4S3. In this work, sodiated $\beta\text{-CD}$ was cycled through the IMS section of this apparatus over a total length of ~ 17 meters, which is sufficient to separate small structural differences.[39] However as shown in Figure 4-4(b), only one primary, a well-resolved peak was observed in the arrival time distribution (ATD) under these conditions. A similar experiment on the six-member cyclic structure $[\alpha\text{-CD}+\text{Na}]^+$ also reveals a single peak in the ATD (Figure 4-4(a)). These results can be potentially interpreted in three ways: (1) the monomodal ATDs demonstrate the presence of a single, stable gas-phase conformer for both sodiated $\alpha\text{-}$ and $\beta\text{-CD}$; (2) both cyclodextrins consist of multiple stable conformers (with different relative orientations of the hydrogen bonding networks, or different coordination number between the Na^+ and CD, for example) that are so structurally similar that they cannot be separated using our SLIM-IMS device; or (3) the monomodal ATD results from multiple conformers that rapidly interconvert.

Figure 4-4:

Arrival times of $[\alpha\text{-CD}+\text{Na}]^+$ (a) and $[\beta\text{-CD}+\text{Na}]^+$ (b) and their corresponding cryogenic IR spectra. Note that the fast ATD peak in the left panels of (a) and (b) corresponds to doubly charged dimers of the corresponding cyclodextrins that cannot be discriminated by our quadrupole mass filter and are propelled faster through the IMS buffer gas due to their doubly charged character



Option (3) appears highly unlikely, given the rigid geometry of both sodiated β -CD and α -CD and the consequent significant strain associated with the concerted rearrangement of the hydrogen-bonded hydroxyl groups. We cannot completely rule out the possibility of (2), that structurally similar $[\alpha\text{-CD}+\text{Na}]^+$ or $[\beta\text{-CD}+\text{Na}]^+$ conformers are present within the same ATD peak. Structural changes resulting from different coordination numbers of the sodium cation will likely lead to structural distortions easily resolved by our ion mobility device. Even if two or more cyclodextrin conformers would have such similar CCS that our SLIM-IMS is unable to resolve them, corresponding calculations would also not be able to reproduce the slight difference in CCS, as the

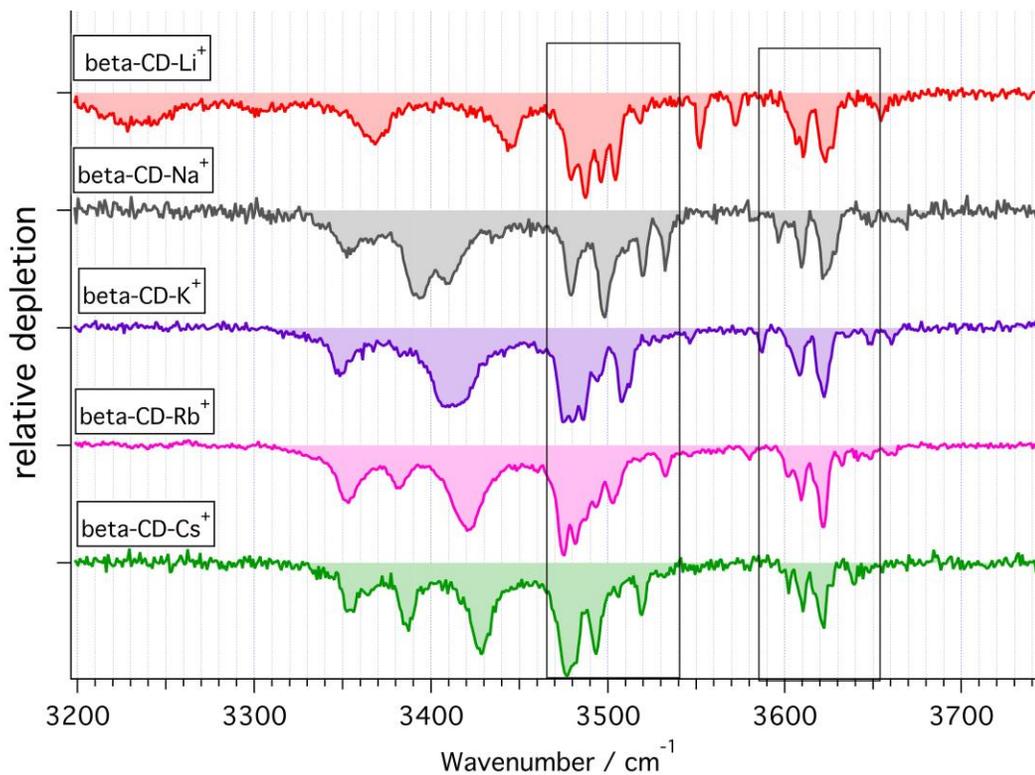
resolution of our IMS instrument exceeds the accuracy of such calculations. As discussed more fully below, cryogenic IR spectroscopy is highly sensitive to structural differences, and the number of bands in each of the spectra of Figure 4-4 is consistent with the presence of single conformers. Thus, the most straightforward conclusion from the data of Figure 4-4 is that both sodiated α - and β -CD exist as single conformers in the gas phase.

The cryogenic IR spectra corresponding to the monomodal ATD distributions of sodiated α - and β -CD is shown in Figure 4-4 were obtained by gating the monomer ATD peak, storing and tagging the corresponding ions in a cryogenic ion trap, and monitoring the depletion of tagged species as a function of IR wavenumber using a ToF spectrometer. Both IR spectra show similarities, with three distinct regions of IR transitions representing three types of OH oscillators. The lowest energy range of the spectra (3330-3450 cm^{-1}) exhibits broad bands for both $[\alpha\text{-CD+Na}]^+$ and $[\beta\text{-CD+Na}]^+$. The $[\alpha\text{-CD+Na}]^+$ spectrum presents a slightly asymmetric feature at 3375 cm^{-1} , while the spectrum of $[\beta\text{-CD+Na}]^+$ shows three distinct bands with maxima at 3355, 3395, and 3410 cm^{-1} . Previous theoretical studies on neutral and metal-complexed CDs predict a tight hydrogen-bonded network between the C6 hydroxyls on the narrow rim and find that Na^+ and other metal cations preferentially coordinate with these hydroxyls.[64-67] Additional cryogenic IR spectra of β -CD complexed to other group I metal cations (Figure 4-5) demonstrate that the ionic radius strongly affects the bands below 3450 cm^{-1} . The IR spectra in this wavenumber region show three major bands for all complexed β -CD cations, whereas only the complexes with Cs^+ and Rb^+ show significant spectral

similarity. These experimental findings support the notion that the C6 hydroxyls are binding sites for the metal cations, but also hint at the formation of additional hydrogen bonding networks between these C6 OHs, which vary in number and strength depending on the complexed metal cation. For instance, the small Li^+ will be complexed by a reduced amount of C6 O-atoms, thus likely promoting one or more strong hydrogen bonds between the hydroxyls undergoing metal complexation and other neighboring free C6 OHs. Such a strong hydrogen bond could be the reason for the broad low energy band centered at 3225 cm^{-1} , which is only present in the IR spectrum of $[\beta\text{-CD-Li}]^+$. In contrast to the lithium case, the large radii Cs^+ and Rb^+ cations will be complexed by many of the seven C6 hydroxyl groups of $\beta\text{-CD}$, preventing the strong hydrogen bonds predicted for the Li^+ complex. These symmetry considerations agree with the theoretical predictions of Gamez and coworkers, which find a trivalent coordination of the C6 hydroxyls for the $[\alpha\text{-CD-Li}]^+$ complex, whereas the lowest energy $[\alpha\text{-CD-Cs}]^+$ complex consists of five C6 Oxygens binding to the metal cation.[67] Based on these arguments, an even stronger symmetric complexation could be inferred from the $[\alpha\text{-CD-Na}]^+$ spectrum (Figure 4(a)), where the presence of one single transition centered at 3370 cm^{-1} plausibly arises from a structural arrangement with the C6 hydroxyls binding to the metal cation, leading to a network of concerted hydrogen bonds. In contrast, the low energy spectral region in sodiated $\beta\text{-CD}$ hints at a more asymmetric binding pattern to the metal cation.

Figure 4-5:

Cryogenic IR spectra of β -CD complexed with group I cations. The two marked areas denote the transitions associated to the C2 and C3 OH oscillators



The differences between sodiated α - and β -CD in this spectral region can be thus attributed to different sodium binding patterns and consequent symmetry distortions of the two molecules. In summary, we assign these low energy transitions to the C6 hydroxyl stretch vibrations and associate the differences in this region to the number of metal-binding C6 hydroxyls as well as the resulting strength of additional intramolecular hydrogen bonds between these OHs.

The higher energy part of the spectra shows two groups of narrow transitions in the 3450-3550 cm^{-1} and 3560-3700 cm^{-1} ranges. Based on the above assignment, these

oscillators correspond to the C2 and C3 hydroxyls located on the wider rim. Theoretical studies have repeatedly shown that these hydroxyls form a hydrogen-bonded network that can adopt two different configurations around the rim (cw and ccw), either with the C3 hydroxyls hydrogen-bonded to the partially free C2 hydroxyls of the adjacent glucose moieties, or the reversed configuration. While these two possible configurations prevent us from assigning the groups of IR transitions to specific hydroxyls, the 3450-3550 cm^{-1} and 3560-3700 cm^{-1} transitions correspond to the hydrogen-bonded and partially free hydroxyls on the wide rim, respectively. The similarity between sodiated α - and β -CD in these spectral regions further supports that the metal cation does not bind here, but rather at the narrow rim.

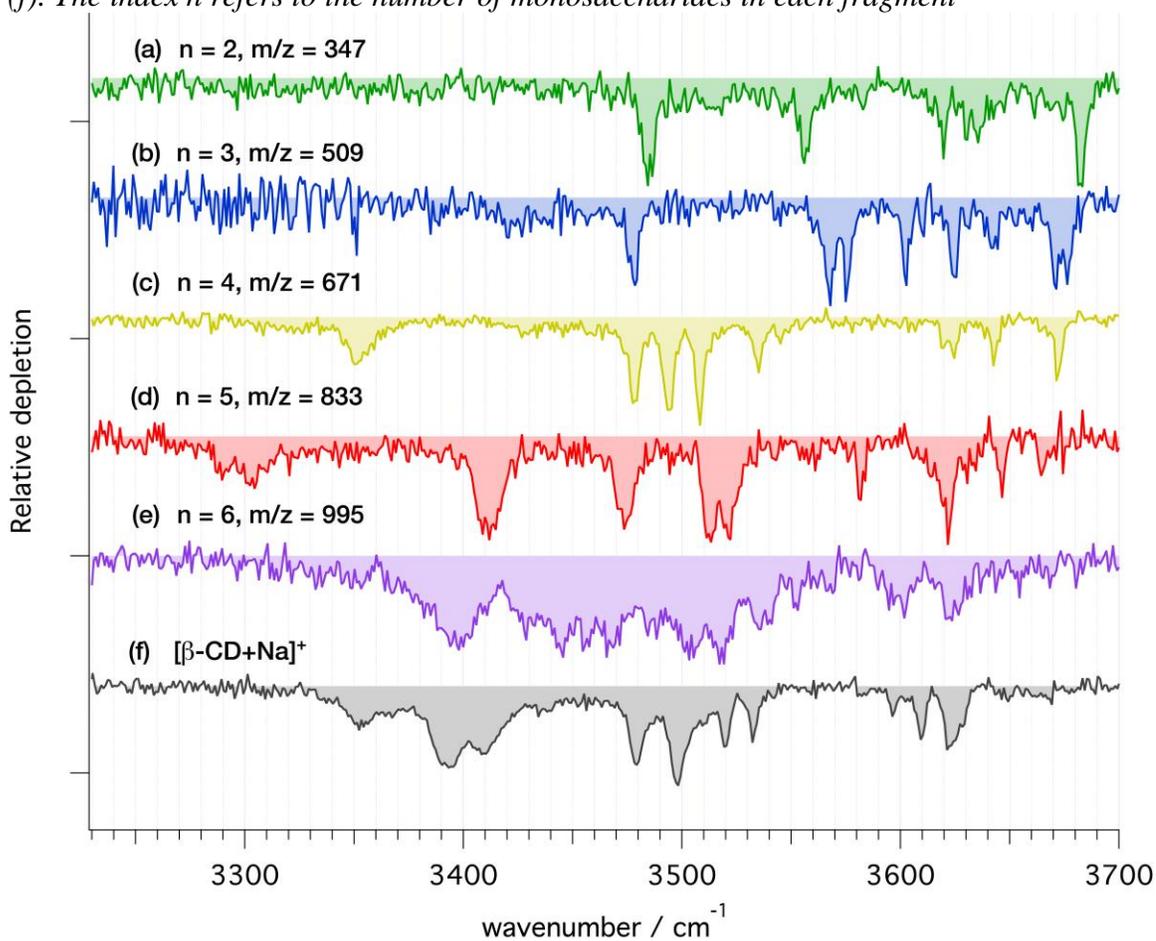
4.4.3 Cryogenic IR Spectra of CID Fragments

Sodiated fragments of $[\beta\text{-CD+Na}]^+$ were produced by collision-induced dissociation between the ion funnel (IF, Figure 4-2) and the hexapole followed by m/z selection in the first quadrupole and storage in the cryogenic trap, where their IR spectra were recorded (Figure 4-6(a-f)). These spectra provide well-resolved transitions that can be used to help determine the structure(s) of the fragment ions. Note that the IR spectrum of the precursor sodiated β -CD (Figure 4-6(f)) obtained on the mass spectrometer shown in Figure 4-2 is highly consistent with the $[\beta\text{-CD+Na}]^+$ spectrum (Figure 4-4(b)) obtained using the IMS-MS device (see Figure 4S4). This spectral comparison demonstrates that the conformational landscape of sodiated β -CD is similar in both instruments. Following all the above arguments, we expect that the CID fragments of Figure 4-3(a) are produced from a single $[\beta\text{-CD+Na}]^+$ conformer.

To associate the observed spectroscopic features with a specific structure, compositional sampling simulations followed by electronic structure and vibrational frequency calculations were performed on the sodiated fragments containing 2, 3, and 4 residues (m/z 347, 509, and 671, respectively).

Figure 4-6:

Cryogenic IR spectra of the main CID fragments of $[\beta\text{-CD}+\text{Na}]^+$ (a-e) and the parent ion (f). The index n refers to the number of monosaccharides in each fragment



The spectra of the 2- and 3-residue sodiated fragments (Figure 4-6(a, b)) exhibit a series of well-resolved bands above 3530 cm^{-1} as well as one band in the $3480\text{-}3490\text{ cm}^{-1}$ range. Neither spectrum show bands below 3470 cm^{-1} . In both cases, the observed number of transitions match the expected number of OH stretch vibrations if one assumes that there is only a single conformer. The most distinct IR signatures are observed for the 4-residue fragment (Figure 4-6(c)). The three clumps of well-resolved bands resemble the spectral features observed for $[\alpha\text{-CD-Na}]^+$ (see Figure 4-4(a)). Synthesis of similar cyclic tetrasaccharides has been previously carried out,[68, 69] and very recently the first synthesis of a 4-member cyclodextrin was reported.[70] However, both the significant strain imposed on the glucose units[71] together with the predicted high barrier to re-cyclization make the formation of a 4-member cyclodextrin unlikely. Despite this, we tested this hypothesis computationally, as discussed below.

In contrast to the spectra of the smaller fragments, the 5- and 6-residue ions (Figure 4-6(d-e)) exhibit a significant number of broad bands at lower wavenumber, which are characteristic of hydrogen-bonded OH stretch vibrations. The spectrum of the 5-residue fragment contains a few weak bands in the $3280\text{-}3320\text{ cm}^{-1}$ range, four strong broad features in the $3400\text{-}3550\text{ cm}^{-1}$ range, as well as several weaker transitions at higher wavenumber. There are fewer bands for this fragment than the number of OH oscillators, indicating that some bands overlap. The 6-residue fragment (Figure 4-6(e)) exhibits a less resolved IR spectrum, with overlapping bands across the entire range between $3360\text{-}3630\text{ cm}^{-1}$. Comparing the $n=6$ fragment to the isomeric $[\alpha\text{-CD-Na}]^+$ (Figure 4-4(a)) reveals a difference in the number, position, and breadth of the bands,

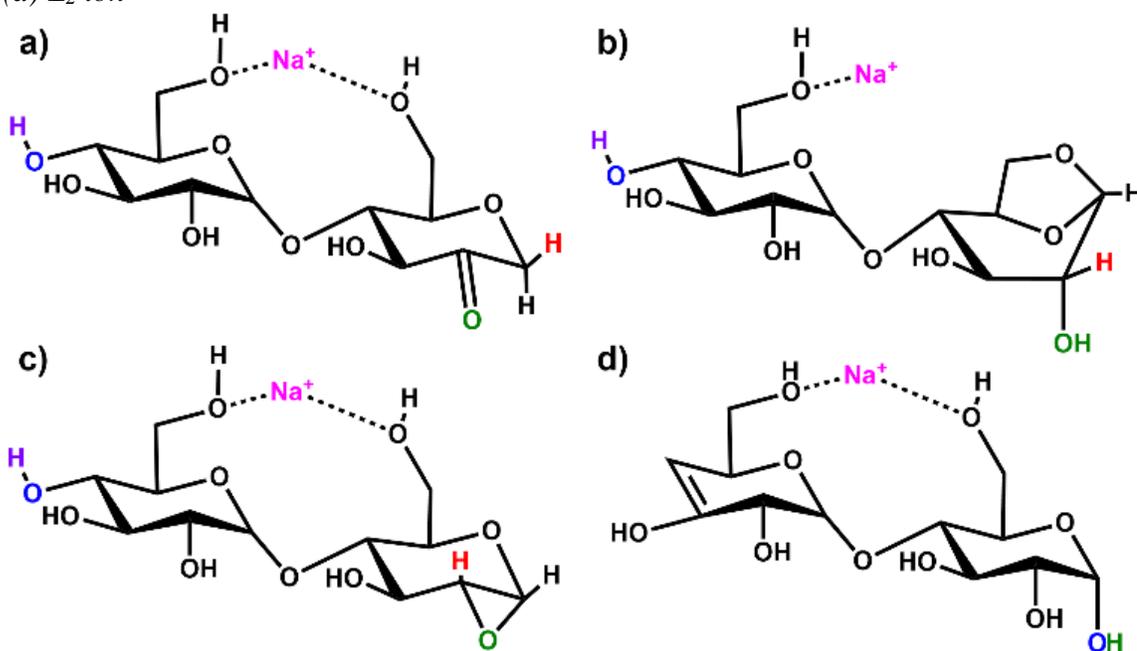
which may reflect the difference between a more symmetric, rigid, cyclic structure and an open, asymmetric one. The spectroscopic differences between the $n=6$ fragment and $[\alpha\text{-CD}+\text{Na}]^+$ thus imply the presence of a substantial energy barrier to re-cyclization of the CID fragment. Rather, one (or several) open fragment structure(s) must be stabilized and kinetically trapped in its (their) potential minimum (minima) through collisional cooling.

4.4.4 Comparison with Theory

Simulations of the 2-, 3-, and 4-residue fragments resulted in a significant number of low-energy conformations. For the 2- and 3-residue fragments, simulations predict four distinct types of energetically stable structures (Scheme 4-1), with an additional cyclic form found for the 4-residue dissociation product.

Scheme 4-1:

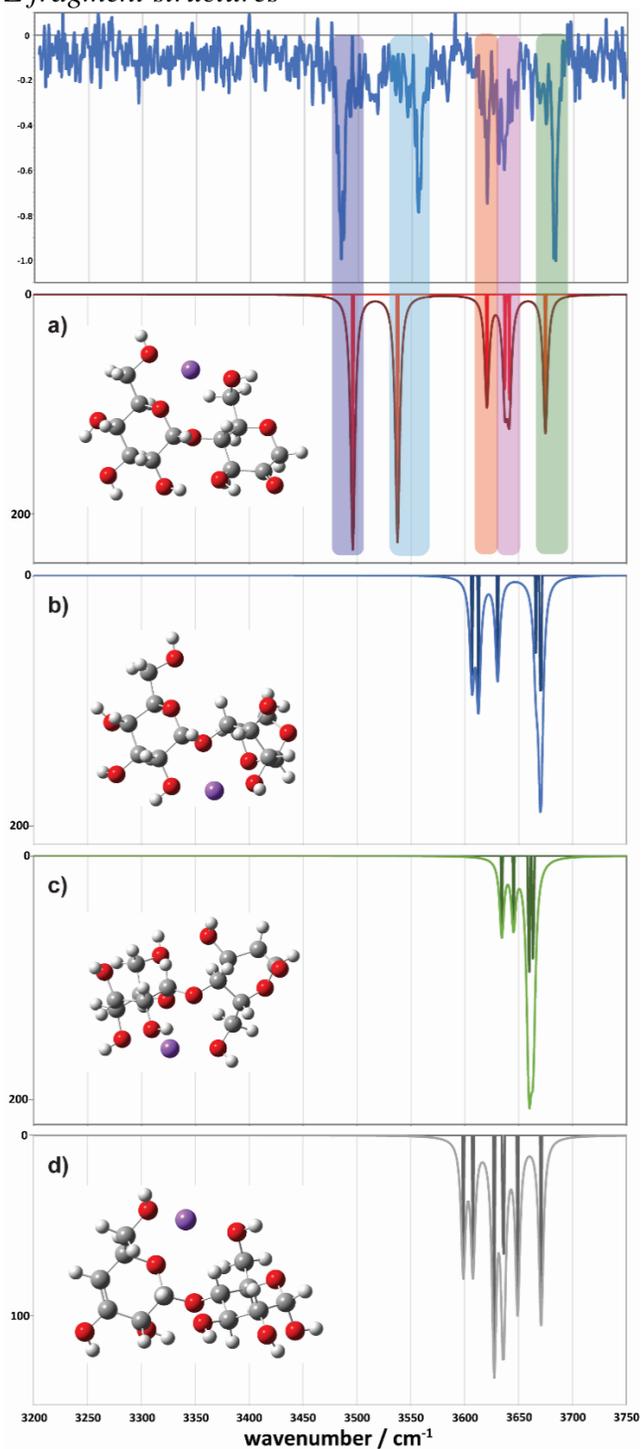
Putative ion structure types potentially formed from $[\beta\text{-CD}+\text{Na}]^+$, illustrated for the 2-residue, m/z 347 species: (a) 2-ketone B_2 ; (b) 1,6-anhydro B_2 ; (c) 1,2-anhydro B_2 ; and (d) Z_2 ion



Three of the four fragments are B ions, while the remaining fragment species is a product of a Z-type fragmentation. Our B3LYP data support 2-ketone B-type fragments[43, 44, 72, 73] as the lowest energy structures. The mechanistically less likely to be formed 1,6-anhydro B ion structure[45, 74] is next followed by the epoxide 1,2-anhydro structures, which are consistently least energetically favorable (≥ 85 kJ mol⁻¹). The Z fragments are predicted to be 10-20 kJ mol⁻¹ higher in energy than the 2-ketone B ion structures. Large basis set M06-2X and ω b97X-D single-point calculations predict the 2-ketone and 1,6-anhydro terminated B ion structures and the Z ion structures to be of similar relative energy. All levels of theory substantially disfavor the strained 1,2-anhydro terminated B ion structures. The lowest energy structures of each type of fragment are shown in Figures 4S5, 4S6, and 4S7 of the Supporting Information, and the corresponding relative energies are indicated in Table 4S1.

Figure 4-7:

Comparison between the experimental (top) and simulated IR spectra for the m/z 347 fragment. (a) represents the ketone, (b) the 1-6 anhydro, (c) the 1-2 anhydro and (d) the Z fragment structures



For the 2-residue fragment (m/z 347), Figure 4-7 depicts the comparison between the experimental spectrum and the IR transitions calculated for the lowest energy B_2 and Z_2 ion structures. This result strongly suggests the formation of a 2-ketone B_2 ion structure (Figure 4-6a). Consequently, we assign: (1) the $\sim 3485\text{ cm}^{-1}$ band to the symmetric O-H stretch of the carbon 3 hydroxyl of residue 2, H-bonded to the adjacent ketone oxygen and the O-H of the carbon 2 hydroxyl of residue 1 (also H-bonded); (2) the 3556 cm^{-1} band to the asymmetric stretch of the same hydroxyl groups; (3) the 3620 cm^{-1} band to the carbon 4 hydroxyl of residue 1, O-H stretch; (4) the 3630 cm^{-1} band to the carbon 3 hydroxyl of residue 1, O-H stretch; (5) the 3636 cm^{-1} band to the carbon 6 primary hydroxyl of residue 2, O-H stretch; and (6) the 3683 cm^{-1} band to the carbon 6 primary hydroxyl of residue 1, free O-H stretch. Following these assignments, and in contrast to the cyclic structures proposed for the metal complexes (Figure 4-5), the low energy bands observed for the sodiated ring-opened β -CD fragments do not necessarily relate to the C6 hydroxyls.

The proposed mechanism for this fragmentation process is presented in Scheme 2 and is adapted from earlier proposals for linear systems.[43, 44, 72, 73] Other candidate fragmentation pathways are detailed in Figure 4S8. The mechanism involves abstraction of a carbon 2 hydroxyl proton, formation of a ketone, and a 1,2-hydride shift with concerted glycosidic bond cleavage.[43, 44, 72, 73] This ring-opening simultaneously creates the first 2-ketone and a new C4 hydroxyl group from the cleaved glycosidic linkage. i.e., nominally a linear B_7 2-ketone terminated structure. Further dissociation of a second glycosidic bond in an identical manner results in the generation of 2-residue and

5-residue structures with the sodium cation in an ion-molecule complex.[74] Provided the complex has sufficient lifetime, a competition for the sodium cation ensues in which the fragment with the larger sodium affinity should most frequently predominate.[45, 74] If the 2-residue fragment separates with the sodium attached, an m/z 347 ion is detected, and if the 5-residue fragment separates with the sodium attached, an m/z 833 ion is detected.

Scheme 4-2:

Proposed fragmentation mechanism for the formation of the m/z 347 fragment

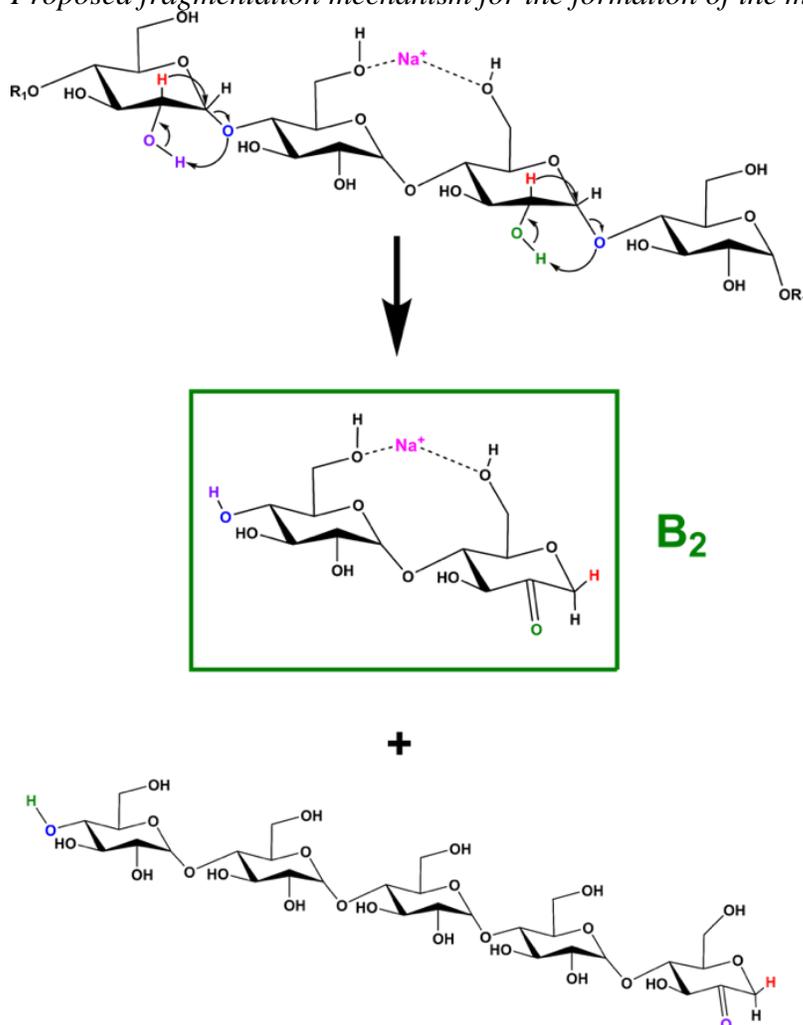
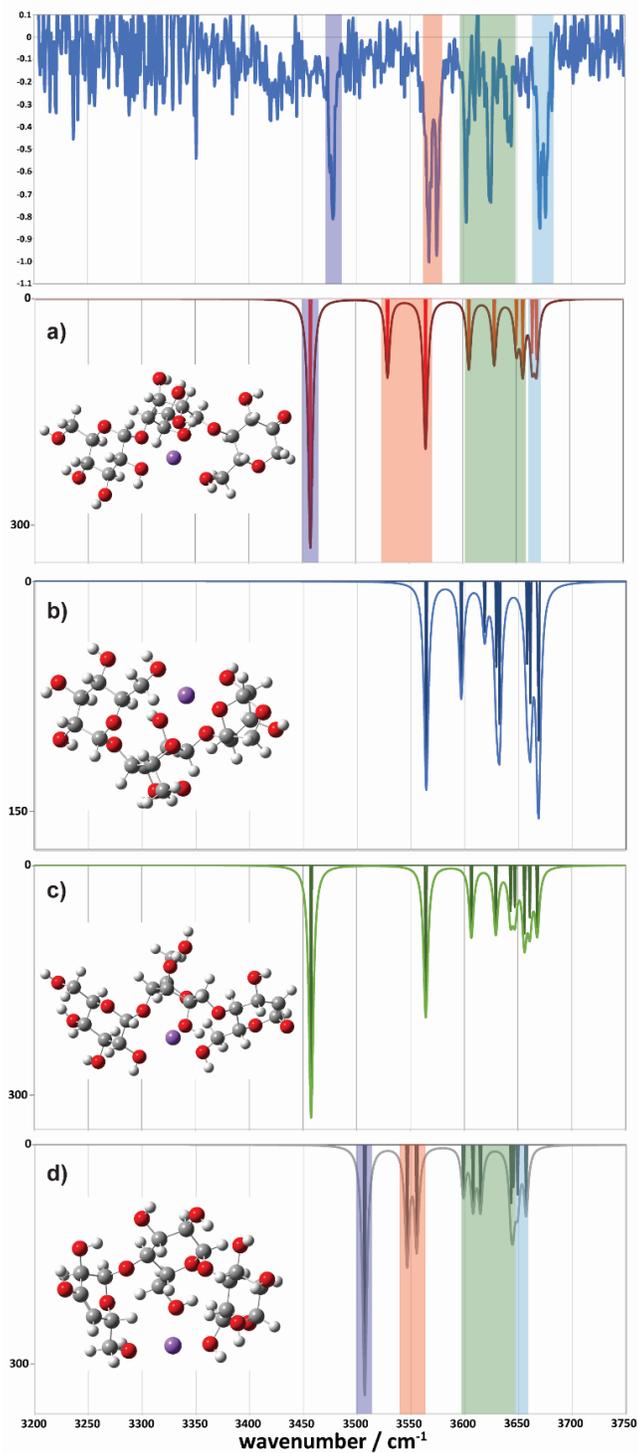


Figure 4-8:

Comparison between the experimental (top) and simulated IR spectra for the m/z 509 fragment. (a) represents the ketone, (b) the 1-6 anhydro, (c) the 1-2 anhydro B_3 , and (d) the Z_3 fragment structures



Comparison between experimental and theoretical IR spectra of the 3-residue fragment is less conclusive (Figure 4-8). None of the theoretical spectra of the lowest energy structures are perfectly consistent with the experiment. The lowest energy 2-ketone B₃-type structure is reasonably consistent with experiment (Figure 4-8a, mean deviation = 5.9 cm⁻¹, R² = 0.97). However, the Z₃-type structure shows a similar quality of fit (Figure 4-8d, mean deviation = 12.0 cm⁻¹, R² = 0.94). In contrast, the 1,2-anhydro B₃ prediction lacks sufficient intermediate energy bands and there is no evidence to support any population of the 1,6-anhydro B₃ structure. The latter is consistent with structural arguments against the feasibility of the cross-ring proton transfers necessary to generate these structures from alpha-linked sugars (Figure 4S8c).

If our tentative assignment of the 2-ketone B₃ or the Z₃ ion structure for the *m/z* 509 peak is correct, this offers a potential explanation for the origin of the *m/z* 671 spectrum—that is, analogous 3-residue and 4-residue structures together with a sodium cation comprise the ion-molecule complex formed following dissociation of the 2 glycosidic bonds. As both neutrals are of similar size and comprise the same functional groups, similar sodium affinities should be expected, enabling both ions to be formed and detected, provided the ion-molecule complex has sufficient lifetime to enable sodium cation transfer between the fragments.[45, 74]

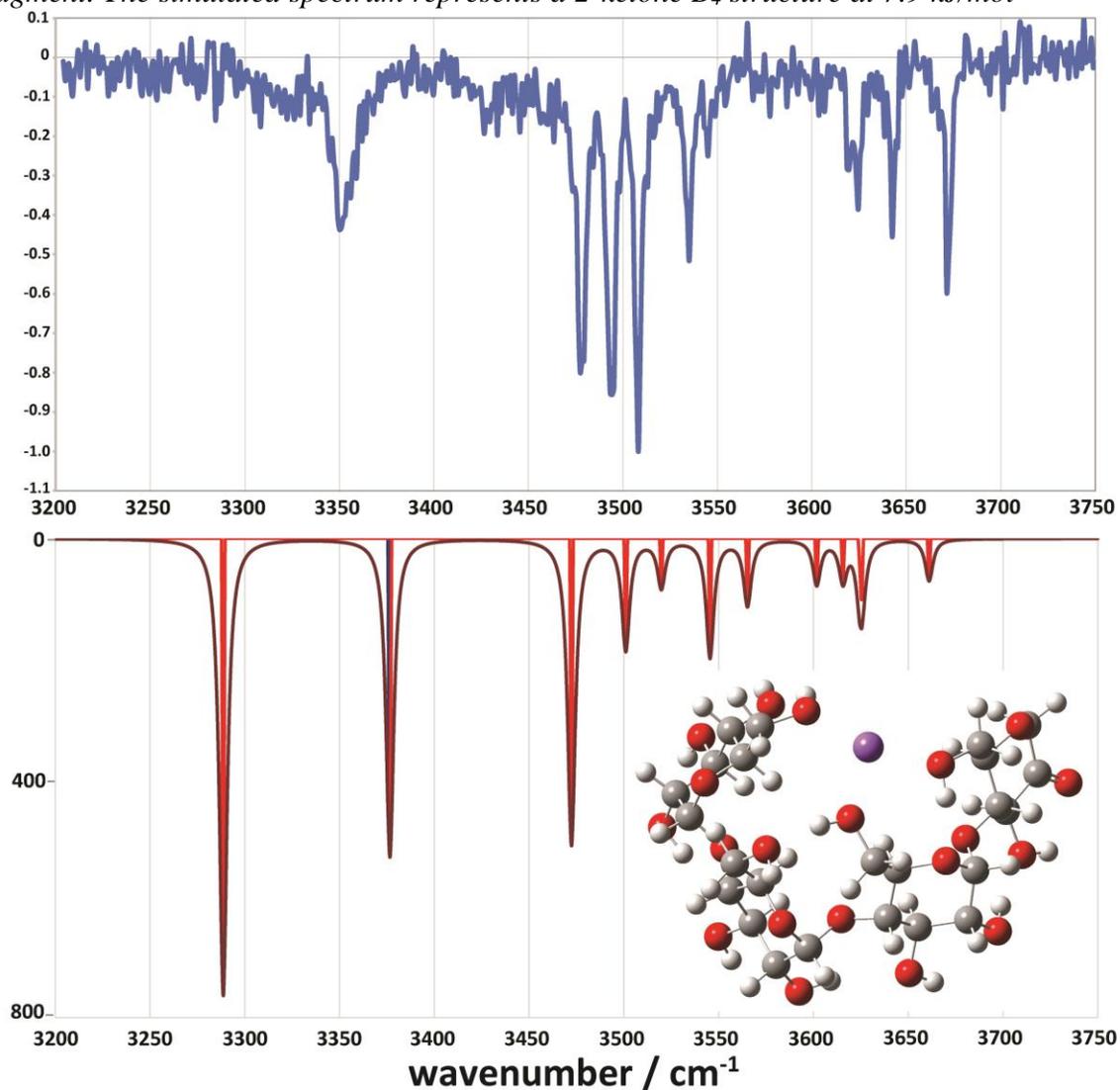
Following this argument, either a 2-ketone B₄ or a Z₄ structure would be expected for the 4-residue fragment, but none of the lowest energy 2-ketone B₄ nor Z₄ theoretical spectra are particularly consistent with the experimental spectrum for *m/z* 671 (Figure

4S7). This is similarly true for the other fragment structure types including potential cyclic candidate structures.

However, we did locate a low-energy 2-ketone B₄ structure that was far more consistent with the experimental spectrum (Figure 4-9). The broad bands at low wavenumber correspond to strongly H-bonded C6-OH stretches (predicted at 3290 and 3377 cm⁻¹), consistent with the broad feature at ~3310-3380 cm⁻¹. The bands at intermediate wavenumber correspond to a mixture of H-bonded environments (C3-OH, C6-OH, C2-OH, C2-OH, C2-OH) at 3473, 3502, 3520, 3546, and 3566 cm⁻¹, respectively. The bands at high wavenumber (3602, 3616, 3626, and 3662 cm⁻¹) comprise three C3-OH stretches and a free C6-OH stretch.

Figure 4-9:

Comparison between the experimental (top) and simulated IR spectra for the m/z 671 fragment. The simulated spectrum represents a 2-ketone B_4 structure at 7.9 kJ/mol



4.5 Conclusions

This work reports fingerprint vibrational spectra for a series of metal complexed β -CD cations, focusing on sodiated β -CD and its CID products. We obtain distinct, well-resolved infrared spectra for each of the main fragments. Simulations followed by

electronic structure and vibrational frequency calculations were performed for the 2-, 3-, and 4-residue ions corresponding to B or Z ion structures and compared to the experimental results. The main findings of this work are:

- (1) Collisional activation of $[\beta\text{-CD}+\text{Na}]^+$ results in a series of fragment ions differing by 162 u. Cryogenic IR spectroscopy experiments on the major fragment ions provide well-resolved vibrational fingerprints in the 3200-3750 cm^{-1} spectral region.
- (2) Our combined experiments and calculations support the formation of a single gas-phase conformer in each of the m/z 347, 509, and 671 species. The experimental spectrum for $n=5$ suggests the formation of a single fragment conformer, while that for the 6-residue fragment suggests the presence of more than one stable conformer.
- (3) There are four main types of potential fragmentation products for the 2-, 3-, and 4-residue fragments. For the 2-residue fragment, the predicted lowest energy 2-ketone-type fragment is clearly the most similar to the experimental spectrum. For the 3-residue fragment, both the predicted lowest energy 2-ketone-type and Z_3 -type fragments show significant similarities to the experimental spectrum. The 4-residue fragment is also most consistent with a 2-ketone type structure, which in turn would support the same for the 3-residue fragment (Scheme 4-2). Thus, all the fragments seem to form 2-ketone structures.
- (4) The results presented here also demonstrate that cryogenic IR spectroscopy and theory can be used to directly relate spectroscopic patterns in CD-metal cationic complexes to the location and degree of coordination between the host CD and the metal ion.

In a broader context, this work demonstrates the potential of our method to unravel atomic-level structural features of complex biomolecular systems that can help address key macromolecular functions.

4.6 Author Contributions

E.C. conceived the experiment; J.M.R. and B.J.B. performed the theoretical simulations; J.M.R. and B.J.B. analysed the simulation results; R.P.P. and E.C. carried out the experiments. R.P.P. analysed the experimental data; A.B.K took the data shown in Figure 4; J.M.R., A.B.K, B.J.B., and T.R.R contributed to data discussion and interpretation; R.P.P., J.M.R., B.J.B., T.R.R. and E.C. wrote the manuscript.

4.7 Conflicts of Interest

There are no conflicts to declare.

4.8 Acknowledgements

E.C., R.P.P., A.B.K., and T.R.R. thank the Swiss National Science Foundation (Grant 200020_184838), and the EPFL for the financial support of this work.

J.M.R. and B.J.B. were supported by the National Science Foundation under CHE-1948611 and ACI-1919789, Ohio University, and performed calculations at the Ohio Supercomputer Center (<http://osc.edu/ark:/19495/f5s1ph73>).

4.9 References

1. G. Gattuso, S. A. Nepogodiev and J. F. Stoddart, *Chem. Rev.*, 1998, **98**, 1919-1958.
2. C. J. Pedersen, *J. Am. Chem. Soc.*, 1967, **89**, 7017-7036.

3. C. J. Pedersen, *J. Am. Chem. Soc.*, 1967, **89**, 2495-2496.
4. F. Hapiot, S. Tilloy and E. Monflier, *Chem. Rev.*, 2006, **106**, 767-781.
5. A. F. Casy and A. D. Mercer, *Magnetic Resonance in Chemistry*, 1988, **26**, 765-774.
6. P. Řezanka, K. Navrátilová, M. Řezanka, V. Král and D. Sýkora, *Electrophoresis*, 2014, **35**, 2701-2721.
7. A. Biwer, G. Antranikian and E. Heinzle, *Appl. Microbiol. Biotechnol.*, 2002, **59**, 609-617.
8. G. Crini, *Chem. Rev.*, 2014, **114**, 10940-10975.
9. M. E. Davis and M. E. Brewster, *Nat. Rev. Drug Discov.*, 2004, **3**, 1023-1035.
10. E. M. M. Del Valle, *Process Biochem.*, 2004, **39**, 1033-1046.
11. H. M. A. M. Dias, F. Berbicz, F. Pedrochi, M. L. Baesso and G. Matioli, *Food Res. Int.*, 2010, **43**, 1104-1110.
12. S. T. Jones, V. Cagno, M. Janeček, D. Ortiz, N. Gasilova, J. Piret, M. Gasbarri, D. A. Constant, Y. Han, L. Vuković, P. Král, L. Kaiser, S. Huang, S. Constant, K. Kirkegaard, G. Boivin, F. Stellacci and C. Tapparel, *Sci. Adv.*, 2020, **6**, eaax9318.
13. K. Lindner and W. Saenger, *Angew. Chem. Int. Ed.*, 1978, **17**, 694-695.
14. J. Szejtli, *Chem. Rev.*, 1998, **98**, 1743-1754.
15. J. Griffiths, *Anal. Chem.*, 2008, **80**, 5678-5683.
16. V. Gabelica, N. Galic and E. De Pauw, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 946-953.
17. C. B. Lebrilla, *Acc. Chem. Res.*, 2001, **34**, 653-661.

18. N. Marangoci, M. Mares, M. Sillion, A. Fifere, C. Varganici, A. Nicolescu, C. Deleanu, A. Coroaba, M. Pinteala and B. C. Simionescu, *Results Pharma Sci.*, 2011, **1**, 27-37.
19. S. Reale, E. Teixidò and F. d. Angelis, *Ann. Chim.*, 2005, **95**, 375-381.
20. S. Sforza, G. Galaverna, R. Corradini, A. Dossena and R. Marchelli, *J. Am. Soc. Mass Spectrom.*, 2003, **14**, 124-135.
21. P. Su, A. J. Smith, J. Warneke and J. Laskin, *J. Am. Soc. Mass Spectrom.*, 2019, **30**, 1934-1945.
22. F. He, J. Ramirez and C. B. Lebrilla, *J. Am. Chem. Soc.*, 1999, **121**, 4726-4727.
23. C. Sonnendecker, S. Thürmann, C. Przybylski, F. D. Zitzmann, N. Heinke, Y. Krauke, K. Monks, A. A. Robitzki, D. Belder and W. Zimmermann, *Angew. Chem. Int. Ed.*, 2019, **58**, 6411-6414.
24. C. Peptu, M. Danchenko, Ľ. Škultéty and J. Mosnáček, *Molecules*, 2018, **23**.
25. Q. Zhang, G.-Z. Li, C. R. Becer and D. M. Haddleton, *Chem. Commun.*, 2012, **48**, 8063-8065.
26. B. Domon and C. E. Costello, *Glycoconjugate J*, 1988, **5**, 397-409.
27. R. Frański, B. Gierczyk, G. Schroeder, S. Beck, A. Springer and M. Linscheid, *Carbohydr. Res.*, 2005, **340**, 1567-1572.
28. K. P. Madhusudanan, *J. Mass Spectrom.*, 2003, **38**, 409-416.
29. Y. Liu and D. E. Clemmer, *Anal. Chem.*, 1997, **69**, 2504-2509.
30. M. F. Jarrold, A. J. Illes, N. J. Kirchner, W. Wagnerredeker, M. T. Bowers, M. L. Mandich and J. L. Beauchamp, *J. Phys. Chem.*, 1983, **87**, 2213-2221.

31. C. A. Wight and J. L. Beauchamp, *J. Am. Chem. Soc.*, 1981, **103**, 6499-6501.
32. J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. L. Caer, J. M. Ortega, F. Glotin and P. Maitre, *Phys. Rev. Lett.*, 2002, **89**, 273002.
33. T. Rizzo and O. Boyarkin, in *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, eds. A. M. Rijs and J. Oomens, Springer International Publishing, Switzerland, 2015, vol. 364, ch. 579, pp. 43-97.
34. T. R. Rizzo, J. A. Stearns and O. V. Boyarkin, *Int. Rev. Phys. Chem.*, 2009, **28**, 481-515.
35. A. M. Rijs and J. Oomens, eds., *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Springer International Publishing, 2015.
36. I. Dyukova, E. Carrascosa, R. P. Pellegrinelli and T. R. Rizzo, *Anal. Chem.*, 2020, **92**, 1658-1662.
37. N. Khanal, C. Masellis, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, *Analyst*, 2018, **143**, 1846-1852.
38. C. Masellis, N. Khanal, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 2217-2222.
39. S. Warnke, A. Ben Faleh, V. Scutelnic and T. R. Rizzo, *J. Am. Soc. Mass Spectrom.*, 2019, **30**, 2204-2211.
40. R. P. Pellegrinelli, L. Yue, E. Carrascosa, S. Warnke, A. Ben Faleh and T. R. Rizzo, *J. Am. Chem. Soc.*, 2020, **142**, 5948-5951.

41. A. Svendsen, U. J. Lorenz, O. V. Boyarkin and T. R. Rizzo, *Rev. Sci. Instrum.*, 2010, **81**, 073107.
42. A. Ben Faleh, S. Warnke and T. R. Rizzo, *Anal. Chem.*, 2019, **91**, 4876-4882.
43. M. T. Abutokaikah, J. W. Frye, J. Tschampel, J. M. Rabus and B. J. Bythell, *J. Am. Soc. Mass Spectrom.*, 2018, **29**, 1627-1637.
44. B. J. Bythell, J. M. Rabus, A. R. Wagoner, M. T. Abutokaikah and P. Maître, *J. Am. Soc. Mass Spectrom.*, 2018, **29**, 2380-2393.
45. J. M. Rabus, M. T. Abutokaikah, R. T. Ross and B. J. Bythell, *Phys. Chem. Chem. Phys.*, 2017, **19**, 25643-25652.
46. J. M. Rabus, D. R. Simmons, P. Maître and B. J. Bythell, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27897-27909.
47. M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, *J. Chem. Theory Comput.*, 2016, **12**, 6157-6168.
48. A. Supady, V. Blum and C. Baldauf, *J. Chem. Inf. Model.*, 2015, **55**, 2338-2348.
49. T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 490-519.
50. T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 520-552.
51. T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 553-586.
52. T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 616-641.
53. T. A. Halgren and R. B. Nachbar, *J. Comput. Chem.*, 1996, **17**, 587-615.
54. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg,

- M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision A.02*, 2009.
55. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
56. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
57. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623-11627.
58. S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 2847-2862.
59. P. Pracht, F. Bohle and S. Grimme, *Phys. Chem. Chem. Phys.*, 2020, **22**, 7169-7192.
60. P. Pracht and S. Grimme, *Chemical Science*, 2021, DOI: 10.1039/d1sc00621e.
61. W. Snor, E. Liedl, P. Weiss-Greiler, A. Karpfen, H. Viernstein and P. Wolschann, *Chem. Phys. Lett.*, 2007, **441**, 159-162.
62. L. Deng, Y. M. Ibrahim, A. M. Hamid, S. V. B. Garimella, I. K. Webb, X. Zheng, S. A. Prost, J. A. Sandoval, R. V. Norheim, G. A. Anderson, A. V. Tolmachev, E. S. Baker and R. D. Smith, *Anal. Chem.*, 2016, **88**, 8957-8964.

63. A. M. Hamid, S. V. B. Garimella, Y. M. Ibrahim, L. Deng, X. Zheng, I. K. Webb, G. A. Anderson, S. A. Prost, R. V. Norheim, A. V. Tolmachev, E. S. Baker and R. D. Smith, *Anal. Chem.*, 2016, **88**, 8949-8956.
64. C. Przybylski, V. Bonnet and C. Cézard, *Phys. Chem. Chem. Phys.*, 2015, **17**, 19288-19305.
65. A. Stachowicz, A. Styrz, J. Korchowiec, A. Modaressi and M. Rogalski, *Theor. Chem. Acc.*, 2011, **130**, 939-953.
66. S. E. Angelova, V. K. Nikolova and T. M. Dudev, *Phys. Chem. Chem. Phys.*, 2017, **19**, 15129-15136.
67. F. Gámez, P. Hurtado, A. R. Hortal, B. Martínez-Haya, G. Berden and J. Oomens, *ChemPhysChem*, 2013, **14**, 400-407.
68. G. C. Daskhan and N. Jayaraman, *Chem. Commun.*, 2014, **50**, 8554-8557.
69. K. Mukai, H. Watanabe, K. Oku, T. Nishimoto, M. Kubota, H. Chaen, S. Fukuda and M. Kurimoto, *Carbohydr. Res.*, 2005, **340**, 1469-1474.
70. D. Ikuta, Y. Hirata, S. Wakamori, H. Shimada, Y. Tomabechi, Y. Kawasaki, K. Ikeuchi, T. Hagimori, S. Matsumoto and H. Yamada, *Science*, 2019, **364**, 674-677.
71. S. Immel, J. Brickmann and F. W. Lichtenthaler, *Liebigs Ann.*, 1995, **1995**, 929-942.
72. J.-L. Chen, H. S. Nguan, P.-J. Hsu, S.-T. Tsai, C. Y. Liew, J.-L. Kuo, W.-P. Hu and C.-K. Ni, *Phys. Chem. Chem. Phys.*, 2017, **19**, 15454-15462.

73. H. T. Huynh, H. T. Phan, P.-J. Hsu, J.-L. Chen, H. S. Nguan, S.-T. Tsai, T. Roongcharoen, C. Y. Liew, C.-K. Ni and J.-L. Kuo, *Phys. Chem. Chem. Phys.*, 2018, **20**, 19614-19624.
74. B. J. Bythell, M. T. Abutokaikah, A. R. Wagoner, S. Guan and J. M. Rabus, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 688-703.

4.10 Supporting Information

Figure 4S1:

MS and MS² spectra of sodiated β -CD at different fragmentation energies in the collision cell of a commercial QToF spectrometer (Waters Permier). The MS spectrum shows the presence of protonated fragments either already present in solution or produced in-source. Formation of the main sodiated fragments starts at voltage difference of around 80V between the quadrupole bias and the collision cell bias. The main fragment masses are marked by green shaded areas

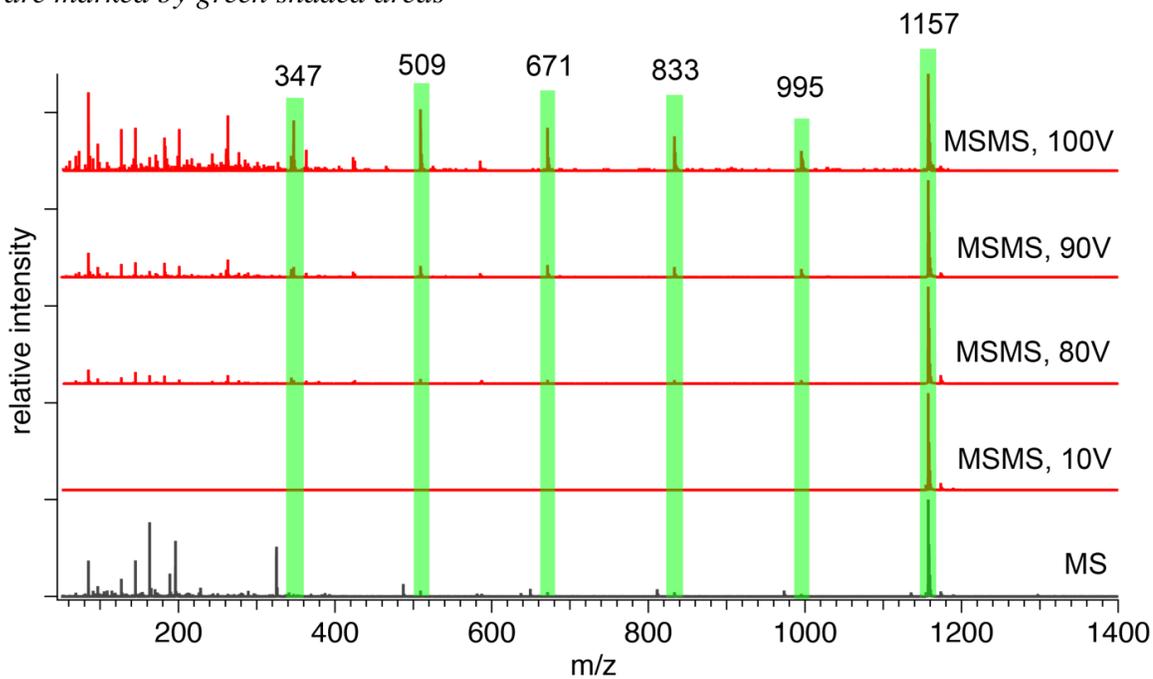


Figure 4S2:

Relative intensity of CID fragments of sodiated β -CD normalized to the precursor intensity as a function of collision cell energies on a commercial QToF spectrometer (Waters Permier)

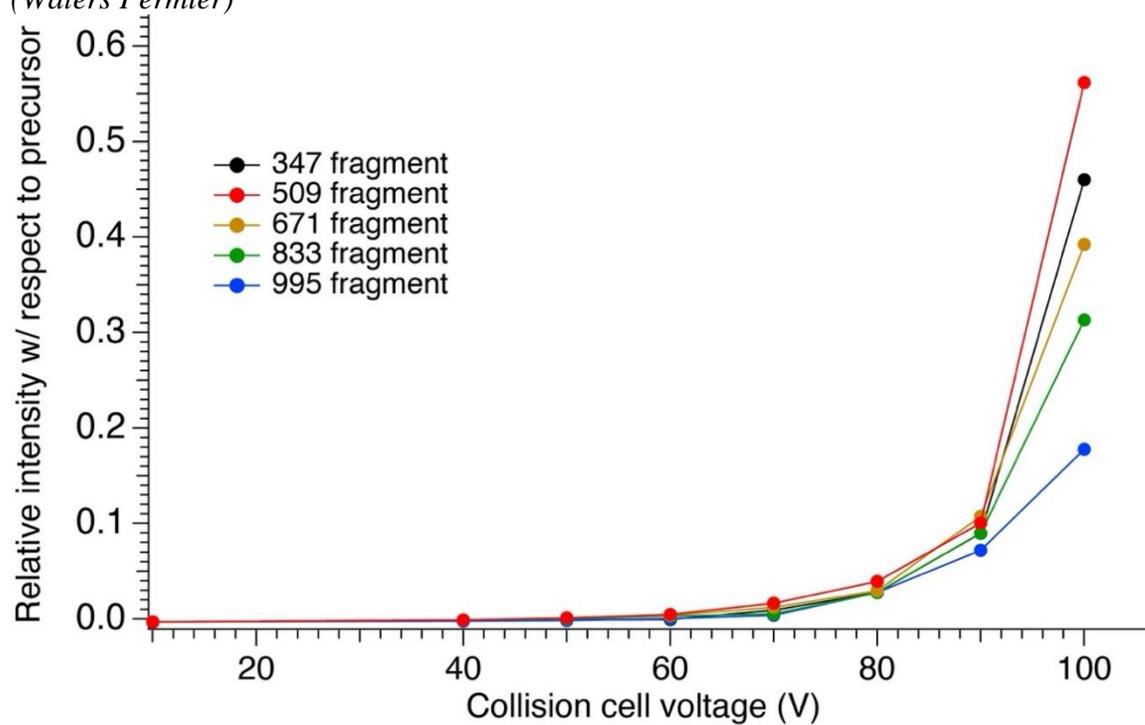


Figure 4S3:

Schematic representation of the homebuilt ion mobility mass spectrometer used to obtain the ATDs and cryogenic spectra shown in Figure 4-4 of the manuscript, adapted with permission from reference 38. The apparatus consists of the same electrospray source, an ion funnel trap, a homebuilt SLIM type ion mobility spectrometer, a quadrupole mass filter and a cryogenically cooled ion trap coupled to an orthogonal time-of-flight spectrometer

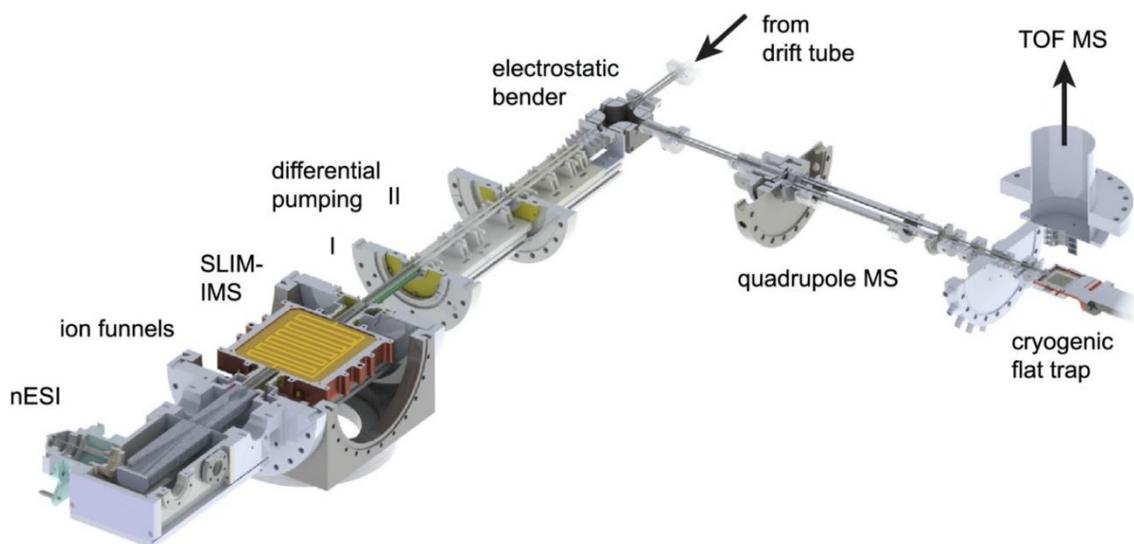


Figure 4S4:

Cryogenic IR spectra of sodiated β -CD measured on the tandem mass spectrometer shown in Figure 2 of the manuscript (upper spectrum) and on the IMS-MS instrument shown in Figure S3 (lower spectrum). Both spectra were taken at the same temperature conditions and using N_2 as tagging molecules. The small shift observed in the bands come from a slightly different laser calibration

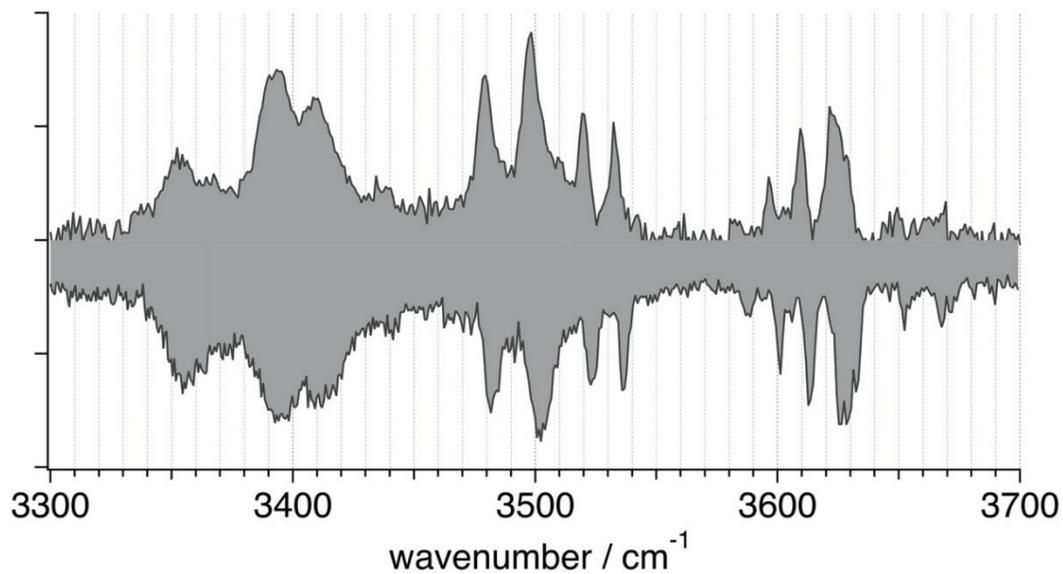


Figure 4S5:

Comparison of m/z 347 lowest energy theoretical spectra with the experimental IR spectrum: lowest energy (a) 2-ketone B_2 ion structure; (b) 1,6-anhydro B_2 ion structure; (c) 1,2-anhydro B_2 ion structure; (d) Z_2 ion structure calculated at the B3LYP/6-31+G(d,p) level of theory. This Figure is reproduced from Figure 4-7 of the manuscript as single-column, full size figure for better visibility of the results

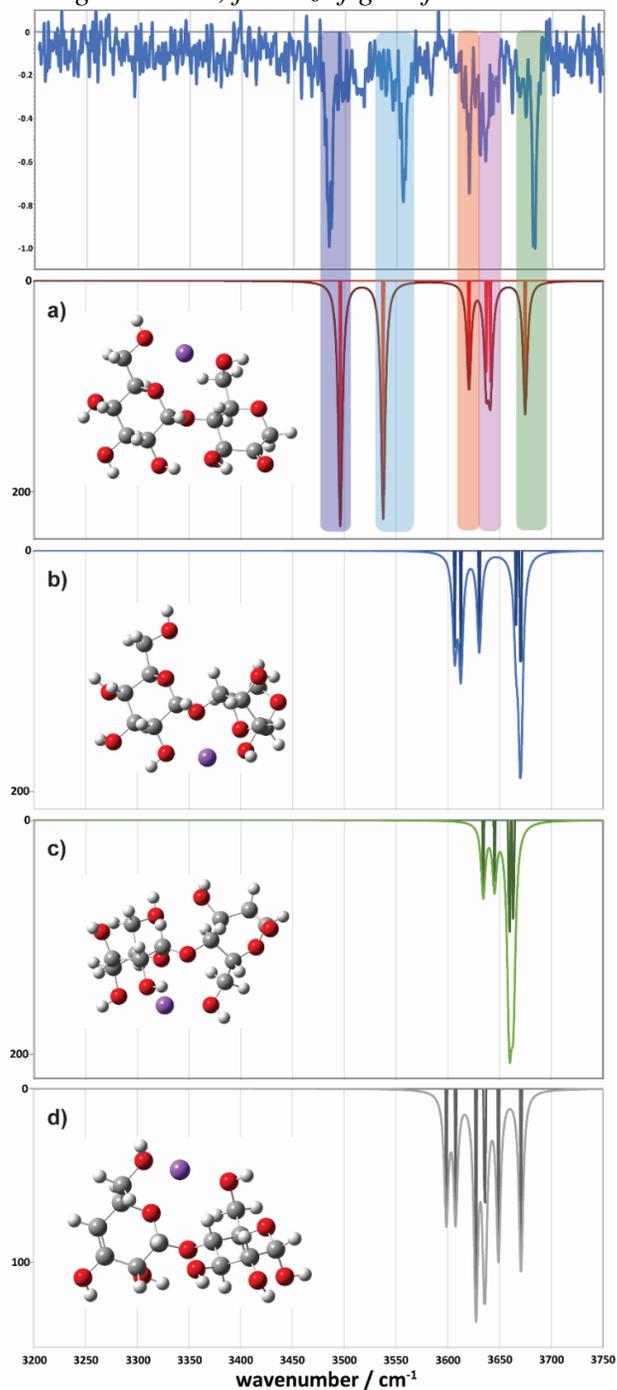


Figure 4S6:

Comparison of m/z 509 lowest energy theoretical spectra with the experimental IR spectrum: lowest energy (a) 2-ketone B_3 ion structure; (b) 1,6-anhydro B_3 ion structure; (c) 1,2-anhydro B_3 ion structure; (d) Z_3 ion structure calculated at the B3LYP/6-31+G(d,p) level of theory. This Figure is reproduced from Figure 4-8 of the manuscript as single-column, full size figure for better visibility of the results

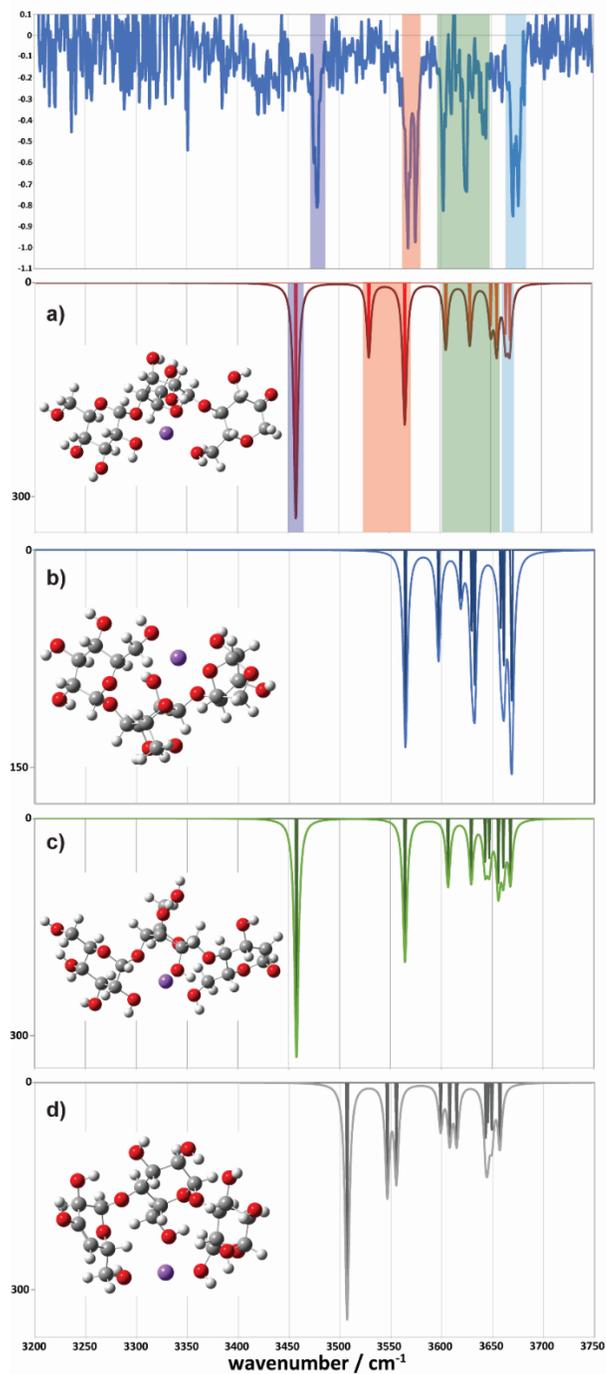
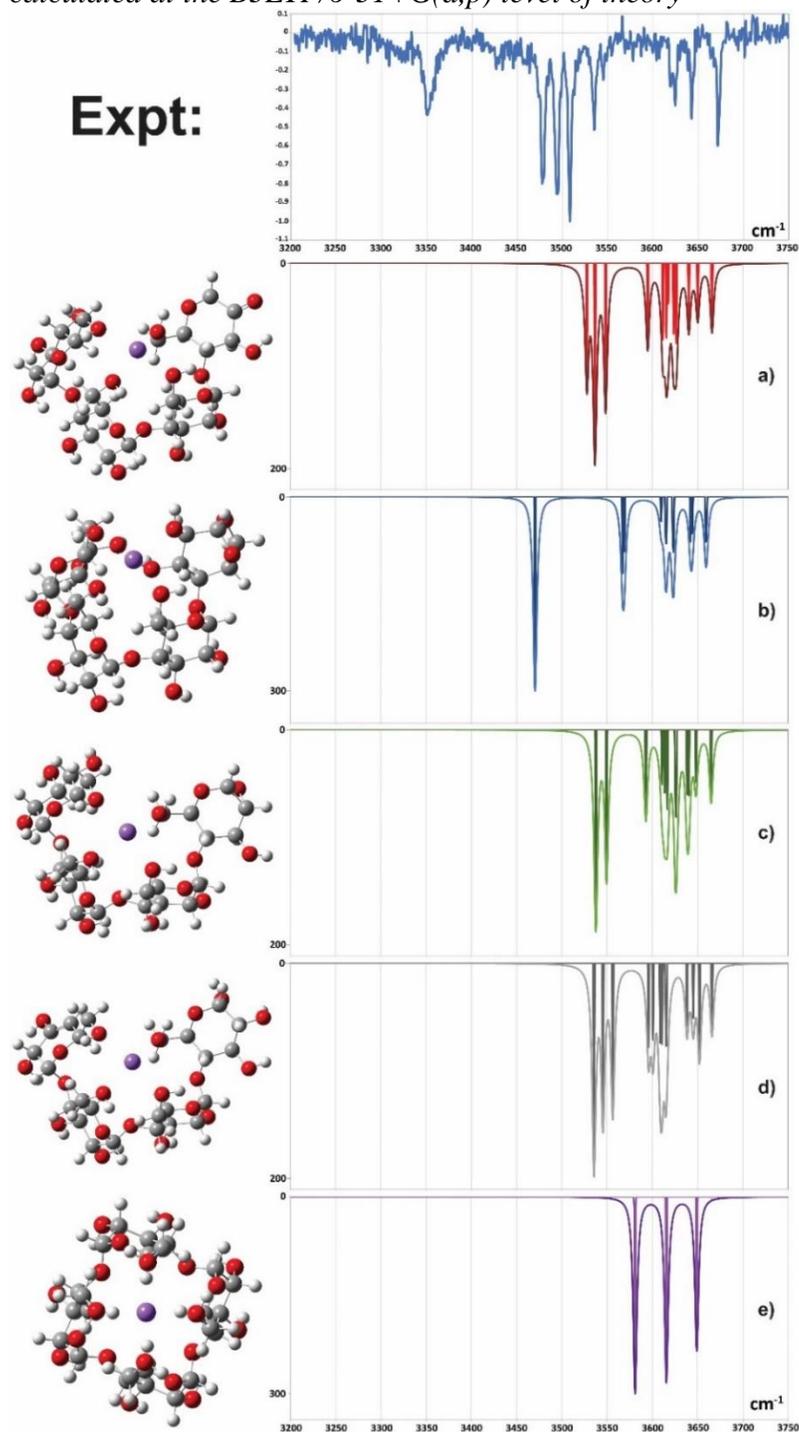


Figure 4S7:

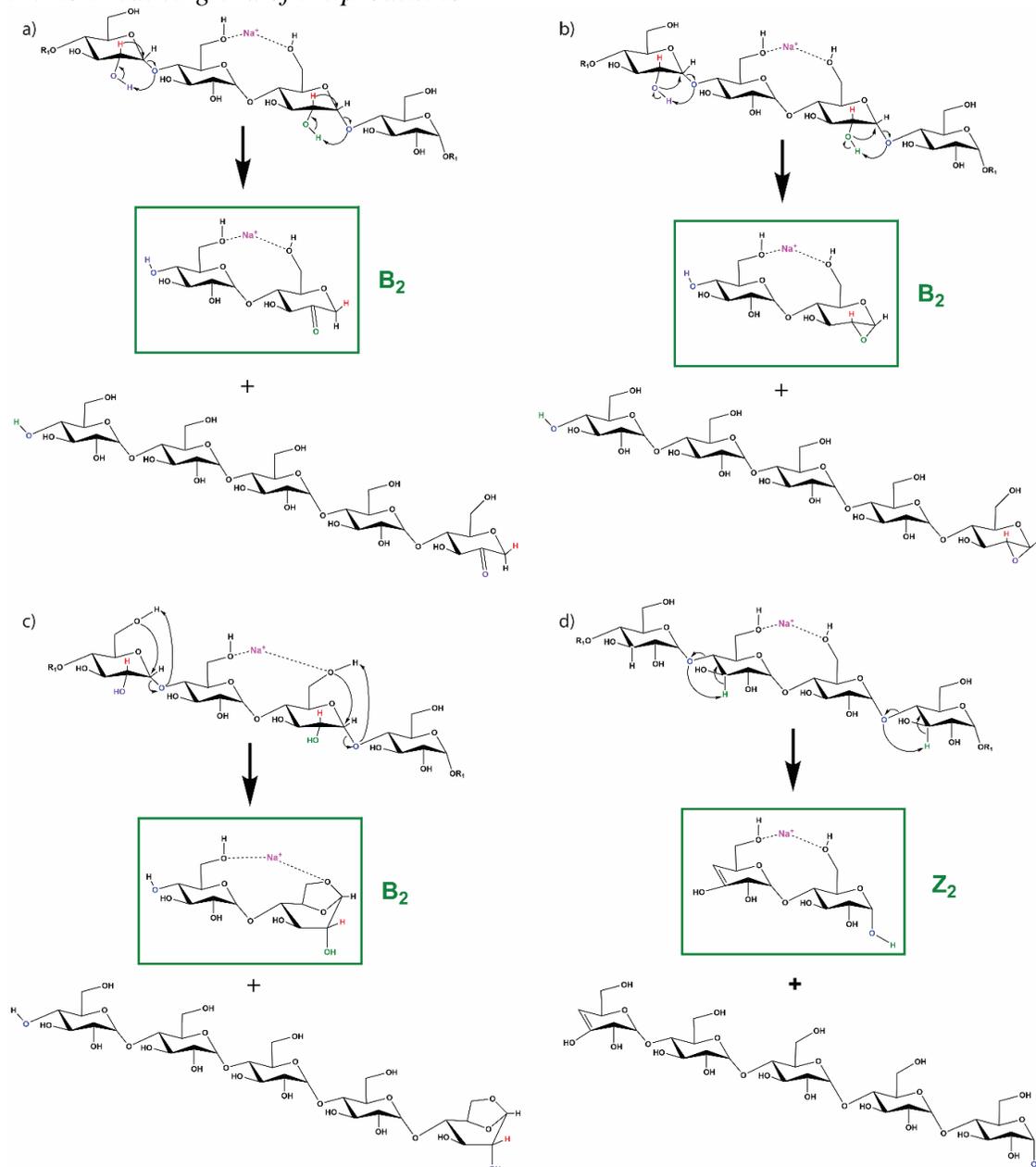
Comparison of m/z 671 lowest energy theoretical spectra with the experimental IR spectrum: lowest energy (a) 2-ketone B_4 ion structure; (b) 1,6-anhydro B_4 ion structure; (c) 1,2-anhydro B_4 ion structure; (d) Z_4 ion structure; (e) Cyclic ion (symmetrical) calculated at the B3LYP/6-31+G(d,p) level of theory



Putative fragmentation mechanisms for CID products

Figure 4S8:

Simplified mechanisms for glycosidic bond dissociation in $[\beta\text{-CD}+\text{Na}]^+$ leading to the (a) 2-ketone B_2 ion, (b) the 1,2-anhydro B_2 ion, (c) the 1,6-anhydro B_2 ion products and the (d) Z_2 ion product. The latter involves formation of an alkene bond between C3 and C4 of the non-reducing end of the product ion



Relative energies of the candidate ion structures calculated at different levels of theory

Table 4S1:

Relative energies of m/z 347 minima. Superscript “a” indicates single point calculations performed using the structure of each type calculated at the B3LYP/6-31+G(d,p) level of theory. Energies are given in kJ mol⁻¹

m/z 347 Structures	$\Delta E_{e1,0K+ZPE}$ (ΔG_{298K}) B3LYP/6-31+G(d,p)	$\Delta E_{e1,0K}$ M06-2X/6-311+G(2d,p) ^a	$\Delta E_{e1,0K}$ wB97XD/6-311+G(2d,p) ^a
Ketone B ₂	0 (0)	7.0	0
Z ₂	14.9 (12.8)	14.4	15.8
1,6-andydro B ₂	22.2 (24.2)	0	3.4
1,2-andydro B ₂	106.4 (106.2)	84.8	90.0

Table 4S2:

Relative energies of m/z 509 minima. Superscript “a” indicates single point calculations performed using the structure calculated at the B3LYP/6-31+G(d,p) level of theory. Energies are given in kJ mol⁻¹

m/z 509 Structures	$\Delta E_{e1,0K+ZPE}$ (ΔG_{298K}) B3LYP/6-31+G(d,p)	$\Delta E_{e1,0K}$ M06-2X/6-311+G(2d,p) ^a	$\Delta E_{e1,0K}$ wB97XD/6-311+G(2d,p) ^a
Ketone B ₃	0 (0)	9.9	0
Z ₃	21.8 (32.0)	21.3	19.5
1,6-andydro B ₃	29.2 (44.8)	0	0.3
1,2-andydro B ₃	86.5 (92.0)	83.7	85.5

Table 4S3:

Relative energies of *m/z* 671 minima. Superscript “a” indicates single point calculations performed using the structure calculated at the B3LYP/6-31+G(d,p) level of theory. Energies are given in kJ mol^{-1}

<i>m/z</i> 671 Structures	$\Delta E_{\text{el,0K}} + \text{ZPE}$ ($\Delta G_{298\text{K}}$) B3LYP/6-31+G(d,p)	$\Delta E_{\text{el,0K}}$ M06-2X/6-311+G(2d,p) ^a	$\Delta E_{\text{el,0K}}$ wB97XD/6-311+G(2d,p) ^a
Ketone B ₄	0 (0)	0	0
Z ₄	11.5 (14.9)	2.0	11.0
1,6-andydro B ₄	41.6 (37.5)	1.4	4.2
Cyclic	56.3 (73.4)	58.6	10.2
1,2-andydro B ₄	89.0 (90.2)	74.4	83.4

Chapter 5 : Conclusion

The study of carbohydrates by mass spectrometry is an active area of research. Although narrowly composed, the fragmentation chemistry of carbohydrates is quite diverse, allowing for the distinction of otherwise ambiguous spectra. Experimental methods such as isotopic labeling allow for the disambiguation of otherwise isomeric fragments but are not always feasible. With the use of computational methods, an additional dimension of analysis is afforded to further disambiguate the experimental results.

By examining model systems, we have built up an initial understanding of the key structures and dissociation chemistries relevant to some gas-phase carbohydrate ions. Based on these data we hypothesize that the fragmentation behavior of carbohydrates can be projected onto larger, more complex/biologically relevant systems, allowing for their analysis where previously this was not possible. The model systems studied were the disaccharides cellobiose and gentiobiose, the disaccharide lactose, and a series of cyclic carbohydrates known as cyclodextrins. Systems were studied by MS/MS using CID then computationally modeled using a series of DFT calculations, systemically investigating possible structural outcomes and their energetic feasibility. For the lactose and cyclodextrin, gas-phase “action” spectroscopy was also used to interrogate the systems to gain additional insight into the structural possibilities of the systems and their fragments. To accomplish this, theoretical modeling was used to calculate the vibrational spectra for a series of potential fragment ion structures and compared them with the experiment. In so doing structural hypotheses were rigorously tested. The results from these model

systems will be used as a foundation from which the mass spectrometric study of larger and/or more sophisticated carbohydrates can be launched.



OHIO
UNIVERSITY

Thesis and Dissertation Services