

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. Bythell^{1*}

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.

Manuscript: CP-ART-04-2018-002620.R1

Submission: 3 October 2018

*Address for correspondence:

Dr. Benjamin J. Bythell, Department of Chemistry and Biochemistry, University of Missouri-St.
Louis, St. Louis, MO 63121. Email: bythellb@umsl.edu

Key Words: Glycans, Mass spectrometry, Collision-induced Dissociation, IRMPD, Ion Structure,
Density functional theory.

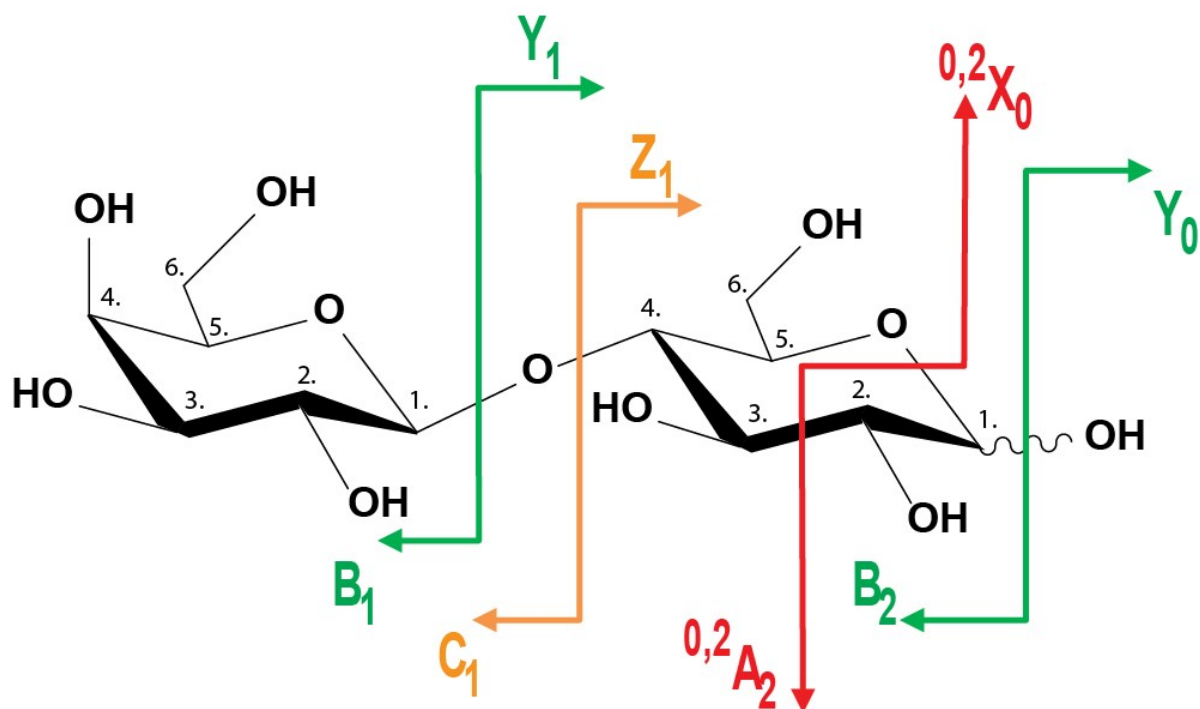
Abstract

We investigate the gas-phase structures and fragmentation chemistry of deprotonated carbohydrate anions using combined tandem mass spectrometry, infrared spectroscopy, regioselective labelling, and theory. Our model system is deprotonated, [lactose-H]⁻. We computationally characterize the rate-determining barriers to glycosidic bond (C_I - Z_I 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_I and Z_I 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_I 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_I anions to ring-open forms were larger, but still modest ($\geq$ 51 kJ mol⁻¹) consistent with 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 offers an explanation for the low abundance of these product anions in our tandem mass spectra.

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} In order 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 ions²⁵ (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 chromatographic²⁸, 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 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 1: Carbohydrate numbering and fragmentation nomenclature of Domon and Costello²⁵ 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 compounds⁴⁴ 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 structure³⁰ 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 MP2⁶⁶ calculations. Regiospecifically ¹³C-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 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.

Experimental Methods

Regioselectively ¹³C-labelled lactose samples (β -D-galactopyranosyl-(1 $\rightarrow$ 4)-D-¹³C₆-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. Breakdown graphs expressing the relative fragment ion signals as a function of collision energy were obtained for labelled and unlabelled [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 was carried out using 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. For unlabelled, deprotonated lactose, the precursor anion was $m/z\ 341.12$ ($\text{C}_{12}\text{H}_{21}\text{O}_{11}^-$) and the fragment peaks were $m/z\ 179.059$ ($\text{C}_6\text{H}_{11}\text{O}_6^-$) and 161.047 ($\text{C}_6\text{H}_9\text{O}_5^-$). 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 labelled (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 in order 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.

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 ⁴C₁-type structures. Geometry optimizations of the resulting candidate conformations were performed with the Gaussian 09 software package⁸⁰ 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++(2d,2p) level of theory with single point calculations performed at the M06-2X/6-311++(2d,2p), MP2/6-311++(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 product-side, one on reactant-side) were then optimized further to determine which minima were connected to each TS. Calculated B3LYP/6-311++(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.

Results and Discussion

1. Tandem Mass Spectrometry of [Lactose-H]⁻

Our regioselectively labelled mass spectra of deprotonated β-D-galactopyranosyl-(1→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 1, Figure S1, Scheme 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} A_2 , ¹³C₄C₆H₁₇O₉⁻ and m/z 267, ^{0,2} A_2 -H₂O, ¹³C₄C₆H₁₅O₈⁻). When the deprotonated β-D-galactopyranosyl-(1→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 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 1a). This is entirely consistent with the preceding data. For the cross-ring cleavage products (^{0,2} A_2 and ^{0,2} A_2 -H₂O) there are no shifts in m/z relative to the unlabeled (m/z 341, C₁₂O₁₁H₂₁⁻) precursor and the deprotonated β-D-galactopyranosyl-(1→4)-D-¹³C₁-glucose anion

(m/z 347, $^{13}\text{C}_1\text{C}_{11}\text{O}_{11}\text{H}_{21}^-$) spectra. The resulting anion peaks (m/z 281, $^{0,2}\text{A}_2$, $\text{C}_{10}\text{H}_{17}\text{O}_9^-$ and m/z 267, $^{0,2}\text{A}_2-\text{H}_2\text{O}$, $\text{C}_{10}\text{H}_{15}\text{O}_8^-$) indicate (Figure 1b) loss of $^{13}\text{C}_1\text{C}_1\text{H}_4\text{O}_2$ (61 u) and $^{13}\text{C}_1\text{C}_1\text{H}_6\text{O}_3$ (79 u), respectively. These losses must include carbon 1 of the reducing end glucose (Scheme 1, Figure 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}.

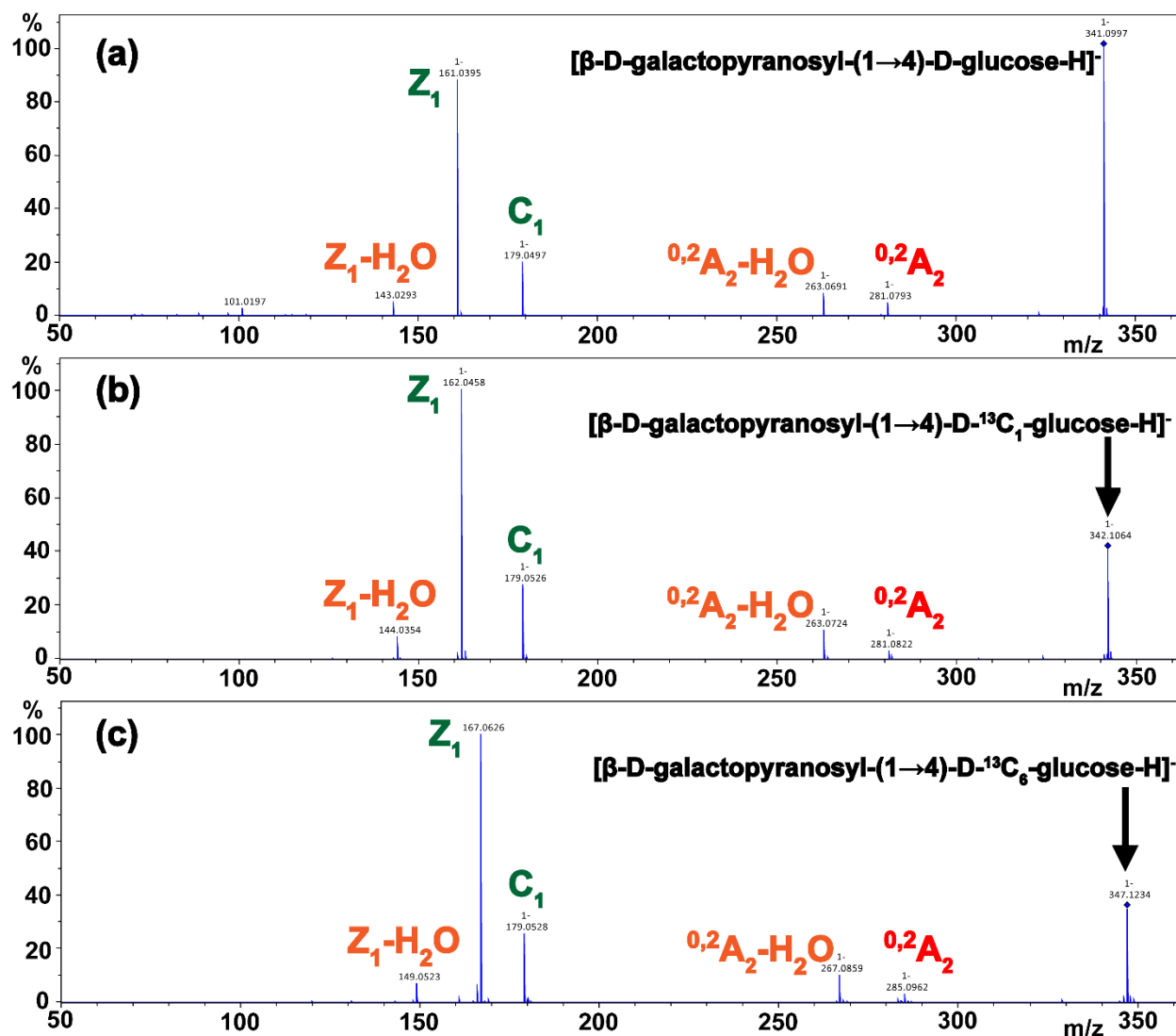


Figure 1: Example Bruker MaXis plus MS/MS spectra of regular and regioselectively labelled, deprotonated lactose ($E_{\text{laboratory}} = 6$ eV in all cases).

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 S2-S4). 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 S4). 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 S5) 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 difficulties⁸²⁻⁸⁶ in examining such systems spectroscopically.

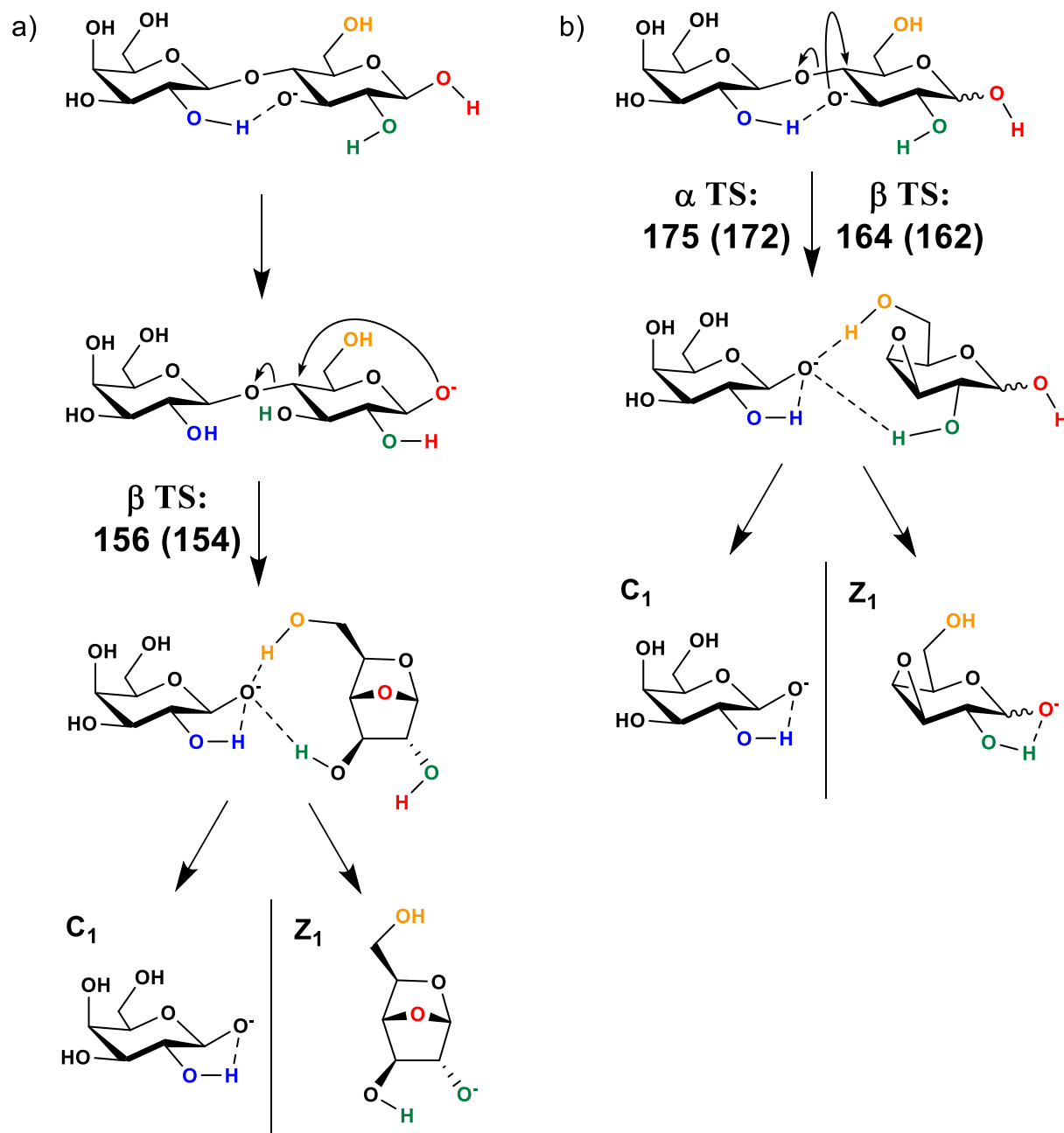
We note however, that while the spectroscopic data is inconclusive, additional structural evidence is available from our MS, MS/MS (Figure 1), and theoretical data. Our MS/MS data do not support a significant population of the aldehyde, ring-open isomer being present. We detect minimal cross-ring fragmentation at low collision energies (Figure 1, $^{0,2}A_2$ and $^{0,2}A_2-H_2O$ anion peaks, Figure S1, Tables 1 & S1, see discussion in Results and Discussion, Section 6). This is consistent with data from multiple earlier solution phase experimental approaches which identify the open-chain aldehyde form in extremely low concentrations (~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 ~0.02%.⁸⁷ Even experiments at high pH where solution phase carbohydrate anions are definitely present indicate 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 3-6) 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. Glycosidic Bond Cleavage Reaction Energetics

The lowest energy transition structures for the C_I - Z_I glycosidic bond cleavage reactions of [lactose-H]⁻ are predicted to vary with anomeric configuration. Our calculations support predomination of two, related intra-glucose S_N2-like mechanisms (Scheme 2, **Figure 2**, Table 1) producing either deprotonated 1,4-anhydroglucose or 3,4-anhydroglucose structures (Figure S6). The reaction forming 1,4-anhydroglucose (Scheme 2a, Figure 2) is the lowest energy pathway for the β-D-glucose anomeric form, requiring at least 156 (154) kJ mol⁻¹, ΔE_{el+ZPE,0K} (ΔG_{298K}). 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_I 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, formation of [1,4-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 2a).



Scheme 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 $\Delta E_{\text{el+ZPE,0K}}$ (ΔG_{298}), calculated at the M06-2X/6-31++G(d,p) level of theory.

The main alternative C_I - Z_I reaction type follows a similar, though more strained S_N2 -like mechanisms to glycosidic bond cleavage (Figure S7 & S8) resulting in different ion-molecule complexes which comprise a deprotonated β -D-galactose anion and 3,4-anhydroglucose (Scheme 2b). The relevant glycosidic bond cleavage transition structures require at least 175 (172) kJ mol^{-1} , $\Delta E_{\text{el+ZPE,0K}}$ ($\Delta G_{298\text{K}}$) for the α -glucose anomer and 164 (162) kJ mol^{-1} for the β -glucose anomer to populate (Table 1). This reflects the additional strain involved in formation of the 3,4-epoxides (Figures S7 and S8). From the ion-molecule complex formed, production of the $[3,4\text{-anhydroglucose-H}]^-$, Z_I ion structure necessitates transfer of a proton to the deprotonated β -D-galactose anion (Scheme 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 S1).

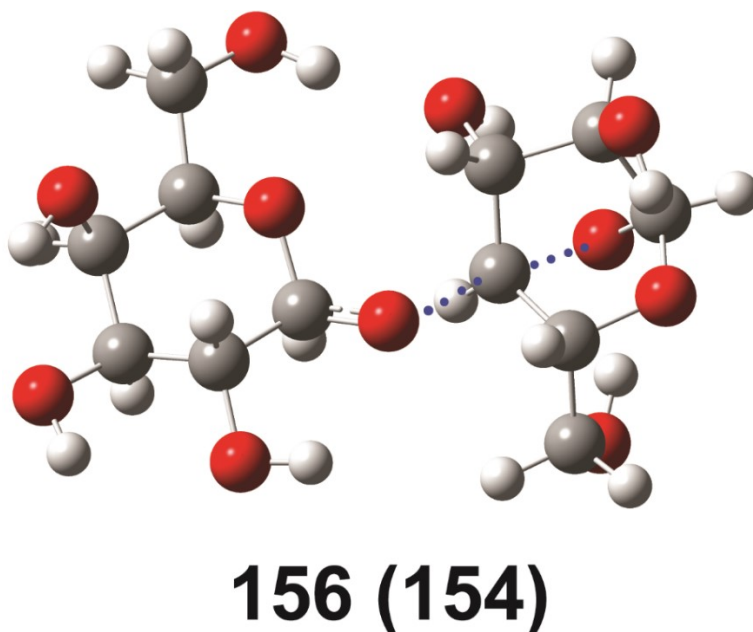


Figure 2: The lowest-energy C_I - Z_I transition structure of deprotonated lactose (β -D-galactopyranosyl-(1 $\rightarrow$ 4)- β -D-Glucose) calculated at the M06-2X/6-31++G(d,p) level of theory (a deprotonated 1,4-anhydroglucose-forming TS). Values in kJ mol^{-1} : $\Delta E_{\text{el+ZPE,0K}}$ (ΔG_{298}).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

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	ΔS_{298} /J 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	ΔS_{298} /J 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	-1296.94447	-1296.591076	116.1	120.7	108.1	42.3

Table 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.

4. Z₁ Structures

Our Z₁ anion infrared action spectrum is shown in Figure 3. The lower energy features are potentially consistent with the [1,4-anhydroglucose-H]⁻, Z₁ anion structures shown in panels a, b and c (secondary alkoxides) providing some of the ion population (Figure 3). We note that these 3 rigid, bicyclic structures (Figure S6a) 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, Figure S9). These barriers are very low at only 23 (23) kJ mol^{-1} , $\Delta E_{\text{el+ZPE,0K}}$ (ΔG_{298}). This barrier indicates that these structures should be easily accessible under our experimental conditions (Scheme 3). Furthermore, the ring-open Z_I anions are substantially more entropically favorable ($\Delta S_{298\text{K}} \geq 23 \text{ J mol}^{-1}$) indicating that this transformation will produce highly favorable product anions as activation level increases (Figure S6). 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}

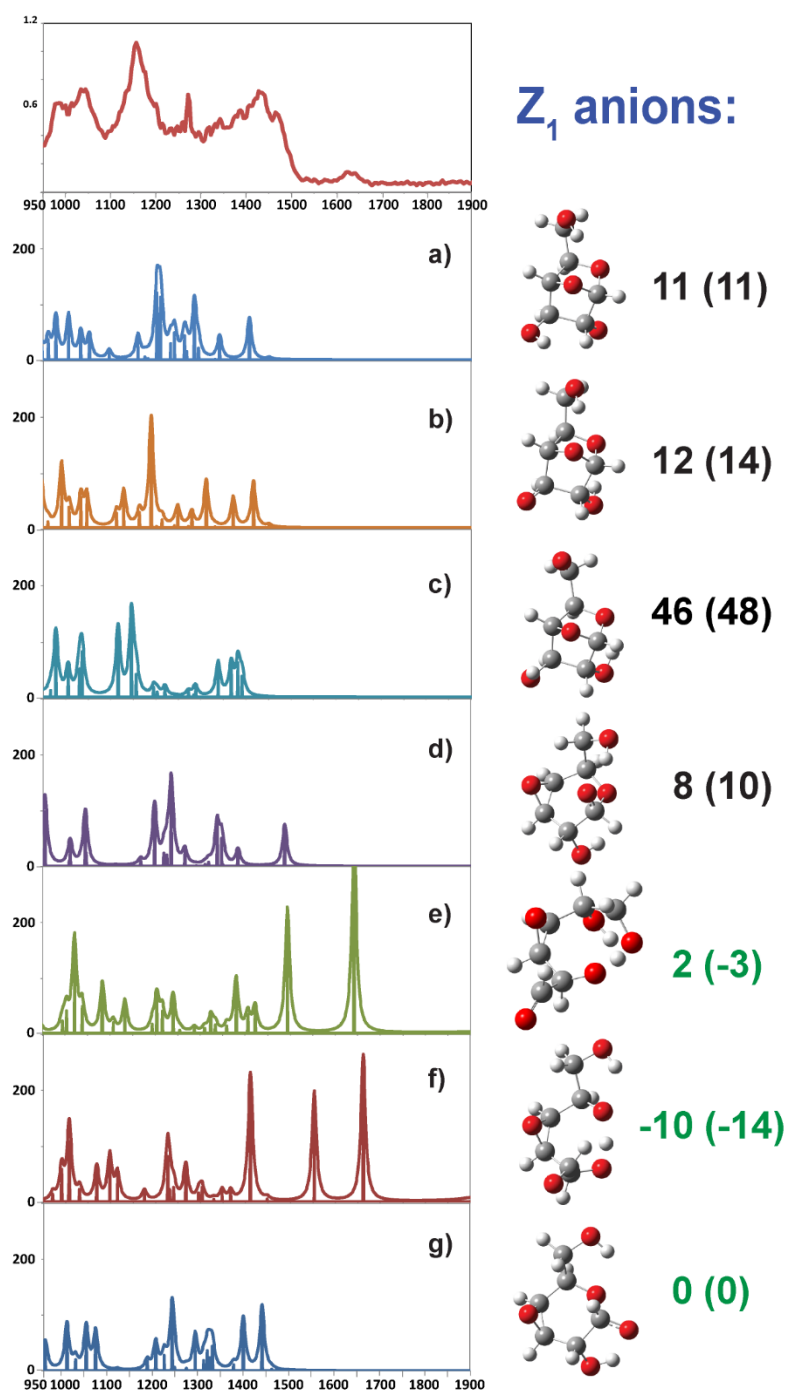
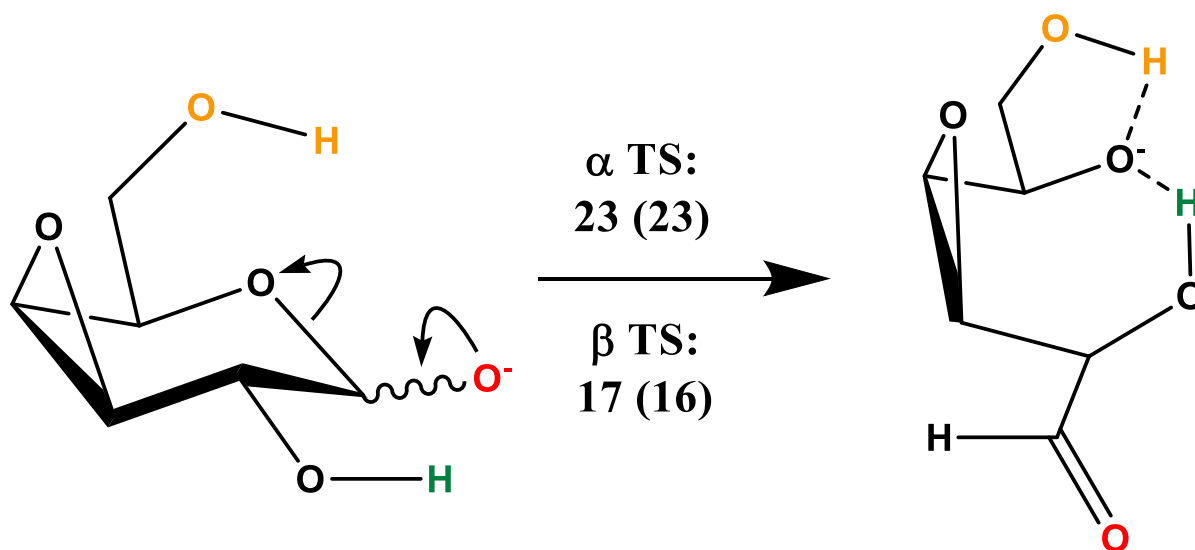


Figure 3: Z₁ anion (m/z 167.066, $^{13}\text{C}_6\text{H}_9\text{O}_5^-$) infrared action spectroscopy spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 $\rightarrow$ 4)-D- $^{13}\text{C}_6$ -glucose anions (m/z 347.129, $^{13}\text{C}_6\text{C}_6\text{H}_{21}\text{O}_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures. 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: $\Delta E_{\text{el}} + \text{ZPE}_{0\text{K}}$ (ΔG_{298}).



Scheme 3. Ring-opening isomerization of deprotonated 3,4-anhydroglucose Z_1 anion structures. Values in kJ mol^{-1} , $\Delta E_{\text{el+ZPE,0K}}$ (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 $\sim 1180\text{--}1220\text{ cm}^{-1}$ 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 $\sim 979, 1008, 1024, 1030, 1035, 1037,$ and 1053 cm^{-1} ; Coupled C-CH₂-OH bending bands $\sim 1030\text{--}1060\text{ cm}^{-1}$ are also predicted; The region from $\sim 1090\text{--}1170\text{ cm}^{-1}$ 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 $1145\text{--}1165$ and $\sim 1280\text{ cm}^{-1}$ ranges. Ring hydroxyl ($1343, 1367, 1373, 1383\text{ cm}^{-1}$) and primary hydroxyl ($1409, 1418\text{ cm}^{-1}$) bends provide the remainder of potentially bands assignable for these structures (panels a-c). The deprotonated 3,4-anhydroglucopyranose Z_1 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 2) is in error, the fragmentation distribution should be substantially biased in favor of the [1,4-anhydroglucose-H]⁻, Z_I anion structures (Scheme 2a, Figure 2, panels a-c, Table 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.

5. C_I Structures

Our C_I anion infrared action spectroscopy spectrum is shown in Figure 4. Consistent with the prior literature^{58,62} we observe bands above $\sim 1600\text{ cm}^{-1}$ 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 population rather than a single environment. From multiple deprotonated β -D-galactopyranose C_I anion forms we calculated barriers to formation of ring-open conformers. Our calculations indicate that this process requires more energy than for the [3,4-anhydroglucose-H]⁻, Z_I 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 S10). 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 $\geq 180\text{ kJ mol}^{-1}$ (Table 1), indicating that ring-opening should not necessarily be expected to immediately occur for larger C_n ions. The β -D-galactopyranose C_I anion forms are enthalpically preferred over the ring-open forms (at 298 K). However, the ring-open product ions are hugely entropically more favorable ($>40\text{ J K}^{-1}\text{ mol}^{-1}$) resulting in Gibbs free energies similar to those of the lowest energy deprotonated β -D-galactopyranose forms at 298 K (Figure S11). In summary, 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, 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 near identical environments and such environments 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.

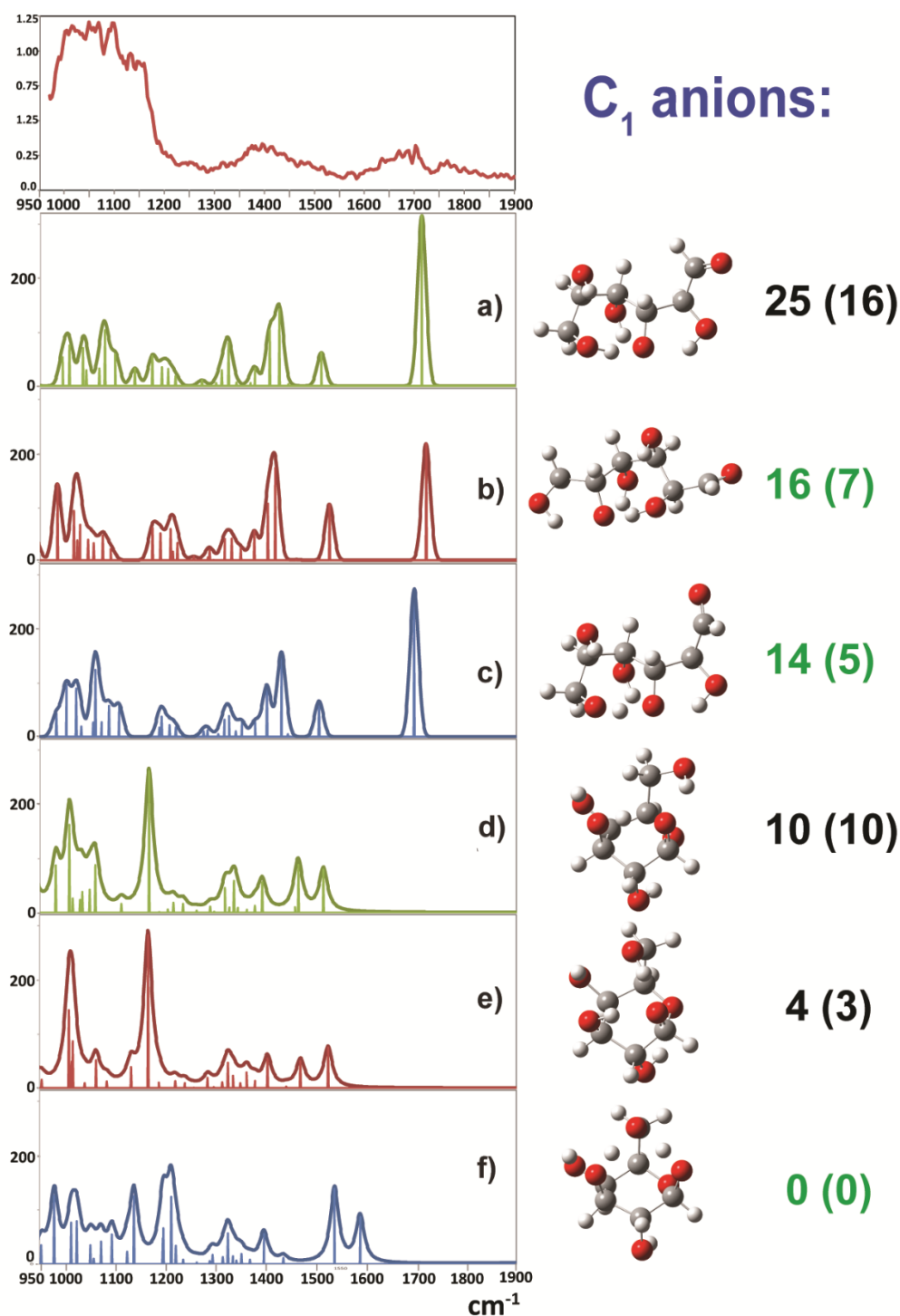
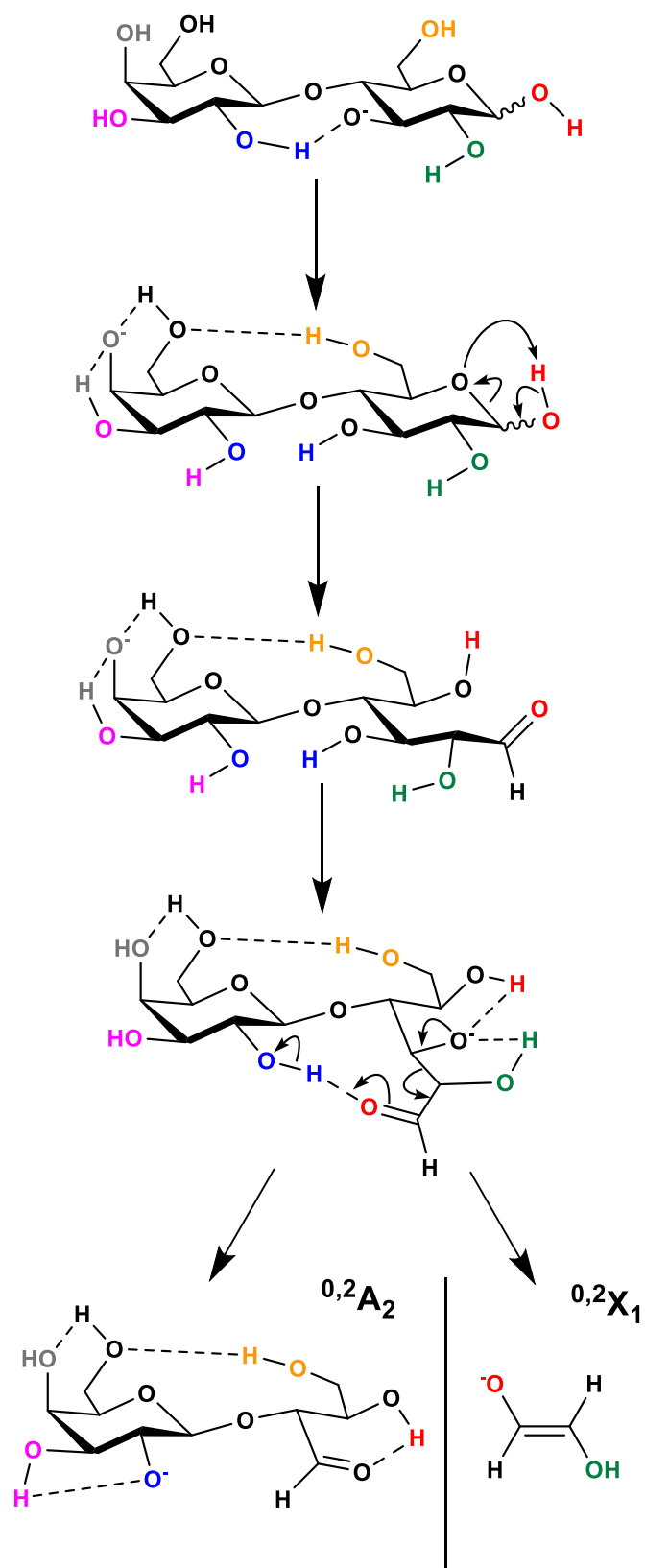


Figure 4: C_1 anion (m/z 179.059, $C_6H_{11}O_6^-$) infrared action spectrum with analytes generated from CID of deprotonated β -D-galactopyranosyl-(1 $\rightarrow$ 4)-D- $^{13}C_6$ -glucose anions (m/z 347.129, $^{13}C_6C_6H_{21}O_{11}^-$), compared to the lowest energy structural possibilities with frequencies calculated from B3LYP/6-311++G(2d,2p) optimized structures. Panels a-c: Ring-open deprotonated galactose structures; Panels d-f: $[\beta$ -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}).

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 1, Figure S1) 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 1, Scheme 4, Figures S12 & S13). 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 S1). Were this not the case, $^{0,2}A_2$ anion peaks would be the most readily formed (Table 1, Table S1). 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 1, Figures S12 & S13). 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 4, Figures S12 & S13). 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 1) which likely contributes to this pathway only contributing a minor portion of the ion current detected.



Scheme 4. The lowest-energy $^{0,2}A_2$ anion formation pathway of deprotonated lactose.

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 population 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 ($^{0,2}A_2$). If true in general, this likely prevents larger C_n (or Y_m) anions from facilely 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.

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.

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 MSn 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. Mestdag, 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 $\text{I}^+\cdots(\text{H}_2\text{O})_2$ 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 an 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.