PUTTING THE PIECES TOGETHER AGAIN: CHARACTERIZING TRISACCHARIDES BY THE ENERGETICS OF THEIR PRIMARY FRAGMENTATION PATHWAYS AND THEIR ION MOBILITY

SEAN OVERTON

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Master's degree in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

Identification of polysaccharides is not a straightforward task due to the high degree of stereochemistry present in their isobaric monomers. Their isobaric nature causes traditional mass spectrometry to fall short when trying to differentiate not only the conformation of the monomers but the position of the glycosidic bonds that bind them. This structural information is important for biochemists as they study the role of different glycans in biological processes.

Tandem mass spectrometry (MS/MS) allows the study of the fragment ions formed during collision induced dissociation (CID), the fragments formed depend on the structure and stability of the precursor molecule and can be used to identify the compounds. These fragmentation pathways will be as complex as the species that form them. To date, typical saccharide fragments are separated into three groups that represent the major fragments: Cross-ring cleavages (A/X), and those resulting from cleaving different sides of the glycosidic bond (B/Y) and (C/Z).

Ion mobility separation (IMS) has shown to have some success at discerning polysaccharide conformers and those of other biopolymers such as proteins and polynucleotides. Ion mobility separates gas-phase ions by colliding them with non-reactive gases and relating respective increase in flight time to their collision cross-section (CCS).

In this study, the relative energetics of the first steps of the cross-ring cleavage and both glycosidic bond cleavage channels for *isomaltotriose* [glc(α1-6)glc(α1-6)glc] as well as a minor water loss channel were explored using density functional theory (DFT) calculations

at the B3LYP/6-31+g(d) level of theory. It was demonstrated that charge-remote mechanisms are a viable alternative to charge-directed mechanisms when under the high energy short time scale conditions present during an ESI-MS/MS experiment.

To verify the efficiency of ion mobility for isomeric separation, the relative experimental CCS of isomaltotriose [glc(α1-6)glc(α1-6)glc], maltotriose [glc(α1-4)glc(α1-4)glc], panose [glc(α1-6)glc(α1-4)glc] and raffinose [gal(α1-6)glc(α1-2)fru] were determined by comparison with literature CCS values for dextran, a variable-length oligomer of α1-6 linked glucose was used as an external calibrant. The experimental CCS of the precursor ions were compared to literature values when available as well as the calculated effective values of the optimized DFT geometries using the trajectory method of the MOBCAL computational suite.

As phosphate is often used as an adducting agent to increase the intensity of the precursor ion when running an IMS experiment, the effect of its presence on the fragmentation of isomaltotriose and large isomaltooligosaccharides was studied. It was seen that depending on the location of the phosphate ion, it will preferentially dissociate leaving behind a neutral glycan. This explains the low abundance of fragment ions observed when selecting a phosphate-adducted precursor ion during an MS/MS experiment.

IMS and MS-MS are complementary methods that can be used to identify monomers within a polysaccharide and how they are bound.

Acknowledgments

I would like to express my gratitude to Professor Paul Mayer and Dr. Justin Renaud who have guided me throughout my academic career. As a young struggling student who thought he had no place in a laboratory, they saw potential in me and have allowed me to work on some amazing projects and grow into a scientist with a passion for analytical chemistry.

I also want to thank everyone from the Mayer laboratory past and present. Over almost a decade you have been there to provide valuable insight and encouragement.

I would like to thank Petra Chea who during her B.Sc. Honours thesis was an exceptional student whom I had the privilege to guide as she joined my research project.

Finally, but in no way the least, I would like to thank my lovely wife, Valérie. I had the honor of marrying you at the beginning of not only my master's degree but also a global pandemic. You were always there to provide encouragement.

Table of contents

List of figures

List of reaction schemes

List of tables

List of abbreviations

In alphabetical order,

B3LYP – Becke, 3-parameter, Lee-Yang-Parr hybrid functionals.

CCS – Collisional cross section, in square angstroms

CRC – Cross-ring cleavage

DC – Direct current

DFT – Density functional theory

E_{COM} – Center of mass frame kinetic energy, in electron volts.

E_{LAB} – Lab frame kinetic energy, in electron volts.

ESI – Electrospray ionization

eV – Electron volts

Fru - Fructose

GBC – Glycosidic bond cleavage

Glc – Glucose

Gal – Galactose

IMS – Ion mobility separation

IMO – Isomaltooligosaccharide

ISO – Isomaltotriose

MAL – Maltotriose

MS – Mass spectrometry

MS/MS – Tandem mass spectrometry

NIST – National institute of standards and technology

PAN – Panose

PES – Potential energy surface

PM3 – Parametric model 3

p.t. – Proton transfer

QM/MD – Quantum mechanics/ Molecular dynamics

RAF – Raffinose

RF – Radiofrequency

RRKM – Rice-Ramsperger-Kassel-Markus theory

STQN – Synchronous transit-guided quasi-newton

TM – Trajectory method

TWIMS – Travelling-wave ion mobility separation

Statement of contribution

Petra Chea did the experimental ion mobility experiments and dextran-derived collisional cross section calculations for isomaltotriose, maltotriose, raffinose and panose during her B.Sc. honours project. Her work was expanded on in chapter 7.

The tandem mass spectrometry experiments for isomaltotriose and the phosphate adducted isomaltooligosaccharides as well as all DFT and MOBCAL calculations were done by me.

Chapter 1. Introduction

Biological polymers are the workhorses of living organisms and are responsible for all biological activities. The study of their structures allows researchers to understand their functions and is the origin of many fields of study or omics. These biopolymers are generally separated into four families: the polynucleotides, the proteins, the lipids, and the carbohydrates. However, this classification is rather simplistic since the complex macromolecules that allow life are often a mixture of these families.

Carbohydrates are normally thought of as simple short-term energy reserves procured through nourishment, but these molecules include themselves into a lot of biological mechanisms. This is due to the formation of glycoconjugates, a glycan or polysaccharide incorporated into another biopolymer. When bound to a protein they are called glycoproteins and they are called glycolipids when bound to a lipid. These then adopt the name of glycoprotein when part of a protein and glycolipid when part of a lipid. The three-dimensional structure of the glycan is of interest due to its ability to affect the structure of the glycoconjugate as well as its function. Glycoproteins are responsible for an array of roles. For instance, the broad mucin family are the main component in the secreted mucus in our mouth, gastrointestinal tract, and female genital tract. They allow lubrification and prevent unwanted bacteria from attaching themselves to the intestinal tract [1]. Another example of an important glycoprotein is immunoglobulin G. It is the dominant antibody in humans and contains multiple glycosylation sites. Gradual variations in the glycan composition in an individual can be due to aging however faster variations have been observed in multiple diseases such as Parkinson's, cancers, autoimmune diseases, and diabetes [2].

The study of glycan structures via mass spectrometry is not trivial in nature. The inherent difficulty is due to the high degree of stereochemistry and isomerisation in the different functional groups, which all lead to different isobaric species that cannot be differentiated via mass spectrometry alone. As shown in Figure 1, carbohydrates can adopt a cyclic structure as either a five-membered ring called a furanose such as fructose (Fru) or a six-membered ring called a pyranose such as glucose (Glc) or it can be linear. In addition, they are differentiated by the orientation of the different hydroxyl groups. The orientation of the hydroxyl group attached to the anomeric carbon differentiates the alpha and beta carbohydrates and the other hydroxyls dictate the name of the carbohydrate such as glucose and galactose (Gal). In addition to the differences found in the individual monomers, the glycosidic bond resulting from a condensation reaction forms a bridge between the anomeric carbon of a carbohydrate to any of the other carbons on second sugar. This variability leads to different three-dimensional configurations as well as branching structures that are all isobaric and indistinguishable to a mass spectrometer also shown in Figure 1.

Carbohydrates can undergo an array of biochemical modifications as shown in Figure 2. These modifications vary its functionality and amplifies the diversity of carbohydrates. These modifications change the mass of the molecule, which allows the identification of the types of modifications present on the glycan but not the location of the modified hydroxyl group. Since carbohydrates are often isobaric, complementary methods to mass spectrometry are needed to study them. The addition of tandem mass spectrometry, which breaks the carbohydrate ions into structure dependant fragments, can give some information on the location of the modification site [3]. Another method is the use of ion

mobility separation, which separates ions according to their collisional cross sections. How these methods work will be explained in further detail in Chapter 2 and supplemental theory on previous results obtained with them are detailed in Chapters 5 and 7.

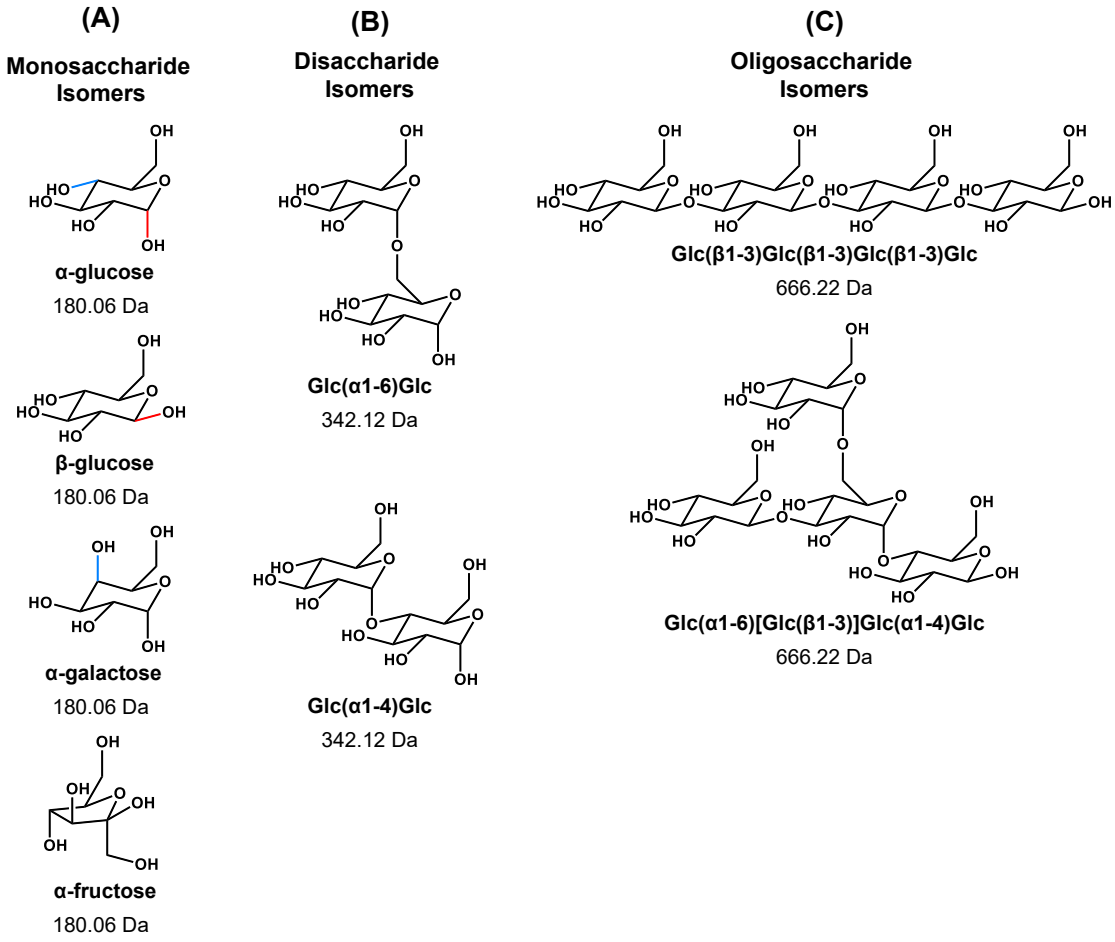


Figure 1. Isobaric nature of unmodified carbohydrate isomers. (A) The unmodified monosaccharides are isobaric and are differentiated by the orientation of the hydroxyl groups as highlighted by the red C-O bond demonstrating the difference between α- or β-carbohydrates or the blue C-O bond demonstrates the differing bond for glucose and galactose, or they can also have a five or six membered ring conformation as shown for fructose. (B) The disaccharides have the increased isomerization from the positioning of the glycosidic bond formed during the dehydration reaction that connects the two sugars. (C)This variability in the orientation of the glycosidic bond can lead to the formation of complex isobaric bridged species.

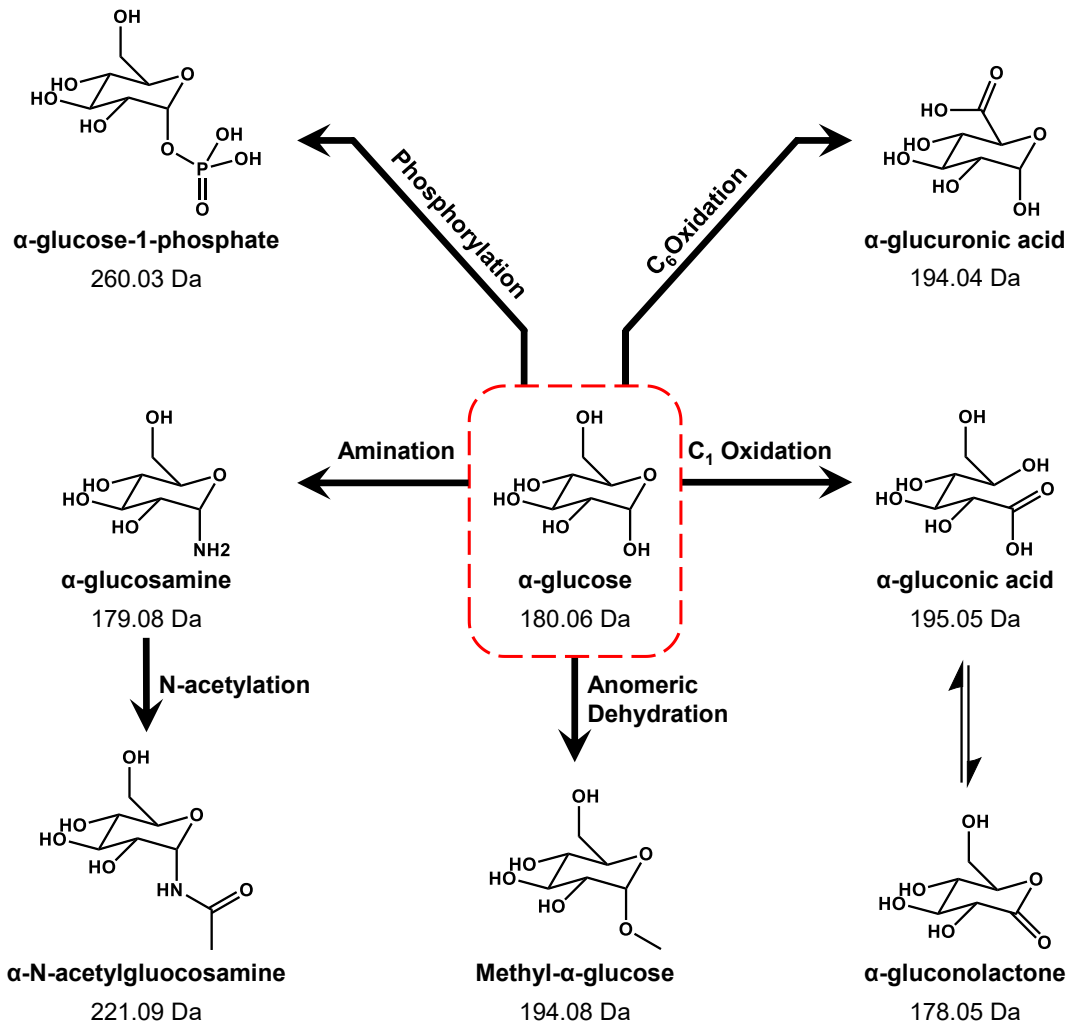


Figure 2. Post-glycosylation modifications of glucose. Enzymatic reactions give more functionality to carbohydrates by chemically altering the structure. These modifications shift the total mass of the glycan allowing for identification of the type of modification, but not which hydroxyl group was changed.

Goal

To better understand how tandem mass spectrometry and ion mobility separation can be used to characterize glycans, our study explores the energetics of the primary fragmentation pathways. It also investigates the binding strength of phosphate as it is commonly used as an adduct to increase ion sensitivity. Finally, we probe the measured collisional cross sections of select trisaccharides.

Chapter 2. Instrumentation

2.1 Electrospray Ionization

Electrospray ionization (ESI) is a commonly used technique for producing the glycan ions needed for their analysis via mass spectrometry. It is considered a soft technique for liquid sample ionization as it primarily produces quasi-molecular gas ions through addition of a charged species (adduction) or addition/abstraction of a proton instead of producing an abundance of fragments in the source [4]. As biopolymers tend to be fragile compounds, a soft ionization technique is essential for their analysis. As shown in Figure 3, the liquid sample is pumped through a charged capillary that is surrounded by a sheath containing the nebulizing gas, in this case nitrogen. The gas flow and the applied voltage cause microdroplets to form at the end the capillary. For negative mode ESI, as is the case of this study, the negative charges on the surface will be neutralized by removing a proton from the proximal solvent molecules. These newly formed ions will want to return to the neutral state by removing a proton from nearby molecules. The source chamber is heated to allow the solvents to evaporate. As the sample evaporates, the more volatile solvent will become less prevalent and eventually the analyte can be deprotonated. As the droplets get smaller, the remaining charges on the surface get closer to each other causing coulombic repulsion. Eventually, this repulsion becomes too great and the droplet explodes leaving a mixture of molecules and complexes in the gas phase. Another way of generating charged analyte is by forming a quasi-molecular ion through adduction. This consist of forming a gas phase complex between the neutral analyte and a small anionic atom like a halogen or a small molecule generally a base such as a phosphate [5]. These gas phase ions can then be electronically guided through the instrument.

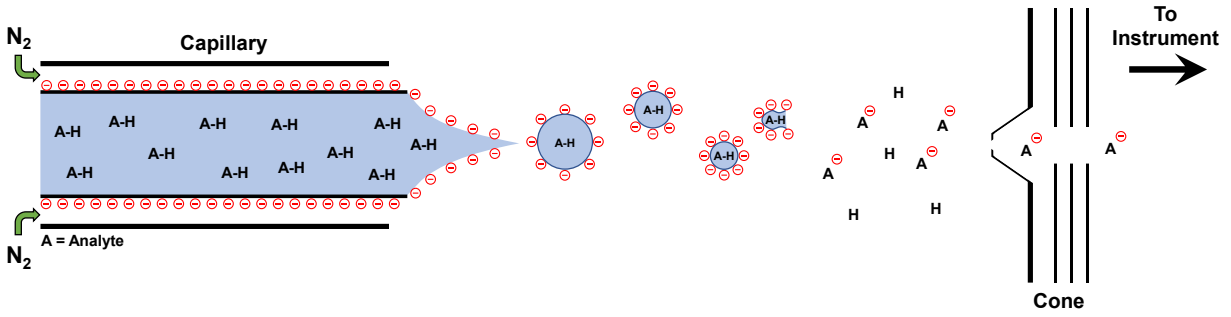


Figure 3. Schematic for the electrospray process in negative ionization mode. The sample solution containing the analyte (A) is extruded through a negatively charged capillary. The combination of the electrical potential and the nebulizer gas (N_2) forms charged microdroplets, which are then evaporated using heat. The charge is transferred to the analyte when the negative charges on the droplets explosively repel each other and break the droplet.

2.2 Quadrupole mass selection

The ESI source is indiscriminate in selecting what molecules get ionized in the sample solution. The number of charges per droplet is finite and the more complex the solution matrix is, the less probable it becomes for the desired analyte ion to form. That probability stems from the equilibrium constants of the protonation-deprotonation reaction. If an impurity in the sample has a more favorable rate of deprotonation than the analyte, it will be more prominent in the resulting mass spectrum. As the type of ion formed in the source is difficult to control, it can cause problems when trying to study ion fragmentation as it then becomes difficult to determine which fragments come from which ion. For this purpose, a quadrupole mass filter is used.

The quadrupole consists of four metal poles, two of them positively charged and the others negatively charged. The polarity of the poles is rapidly switched back and forth using a radiofrequency (RF) alternating current. An offset direct current (DC) voltage is applied to maintain forward momentum of the ions [6]. As shown in Figure 4, The RF and

DC parameters that allow for a stable trajectory through the quadrupole overlap over several mass-to-charge ratios. Since there are small ranges of RF/DC values that allow for the isolation of specific mass-to-charge ratios, the resolution capabilities of a quadrupole are limit ed.

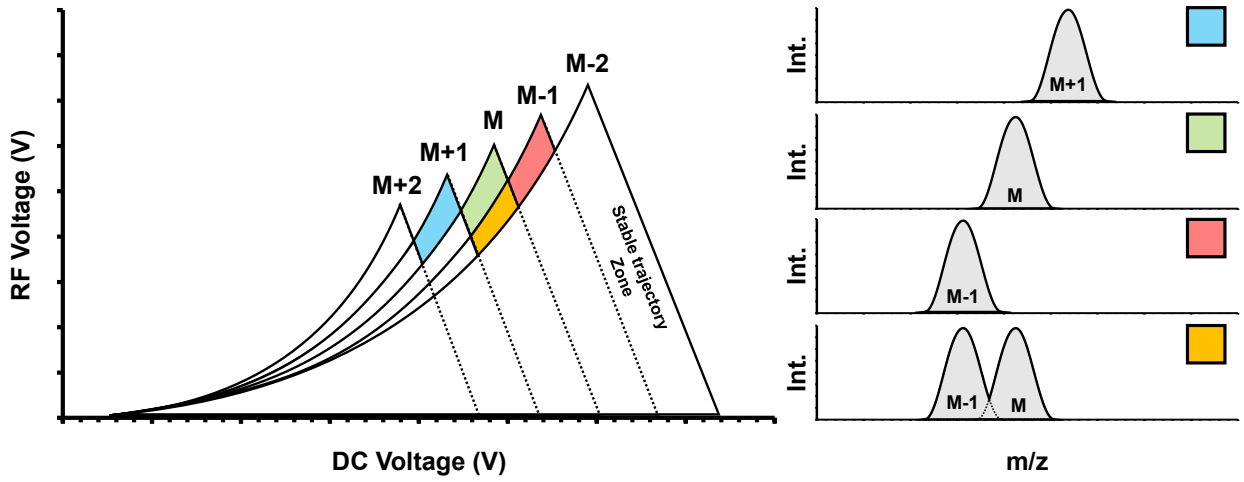


Figure 4. Quadrupole mass selection diagram. The areas formed by each curved triangle shows the DC and RF values that can produce a stable trajectory for five imaginary ions of mass M±2. The red, green and blue areas show the values that would isolate M-1, M and M+2 respectively. The orange area shows the values that would lead to improper separation of M and M-1.

2.3 Collision-Induced Dissociation

Once the precursor ion is selected by the quadrupole, it can proceed with collision-induced dissociation (CID) in which an ion is accelerated by an electric field before it collides with an inert gas like argon. The kinetic energy given to the molecule can be tuned to control the amount of fragmentation that occurs. This energy is labeled as the lab-frame kinetic energy (E_{LAB}) and the electron volt (eV) is used. During the inelastic collisions with the argon atoms, a portion of its total kinetic energy is internalized and dispersed through different vibrational and rotational modes. When the vibrational energy is great enough, the bonds will irreversibly stretch and break as shown in Figure 5. One

product of the reaction, referred to as the fragment ion, keeps the charge and can be electronically guided to the detector. The other product is neutral meaning it cannot be guided towards the detector and is lost to waste through the pump systems. This fragment is generally referred to as the neutral loss. The mass of the neutral loss can be inferred through subtraction of the fragment ions mass from that of the precursor ion.

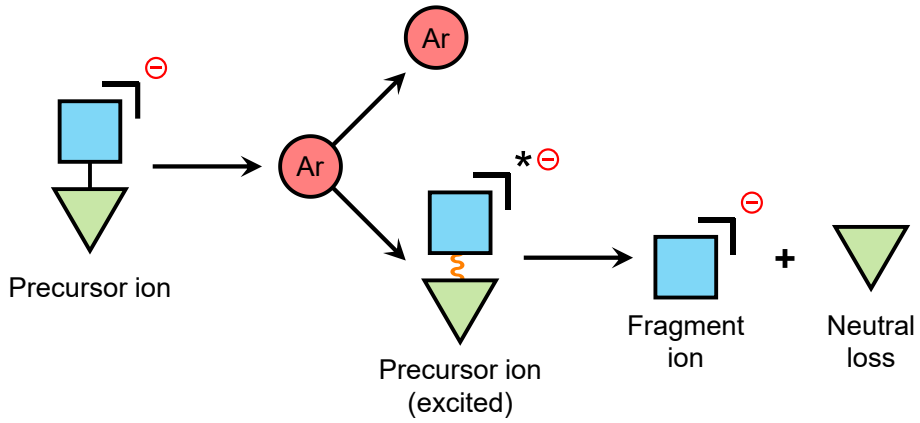


Figure 5. Collision-induced dissociation process. As the accelerated precursor ion collides with an argon atom, some of the kinetic energy is internalized which rotationally and vibrationally excites it. The ion will fragment if the internal energy is sufficient to break bonds or undergo unimolecular reactions.

To study the relation between the lab-frame energy and the fragments formed, the center of mass frame energy (E_{COM}) is used as an approximation and is calculated using equation 1. This is the total amount of the lab-frame energy that can be internalized by considering the relative mass difference between the precursor ion (m_{PI}) and that of the collision gas (m_{CG})[7].

$$E_{COM} = E_{LAB}\left(\frac{m_{CG}}{m_{CG} + m_{PI}}\right) \qquad \text{(eq. 1)}$$

2.4 Ion mobility Separation

2.4.1 Drift tube Ion mobility

The second orthogonal method often used to differentiate glycans is ion mobility separation. This technique accelerates ions through a counter current of an unreactive gas like helium or nitrogen. The ion collides multiple times with the drift gas causing it to decelerate depending on the number collision events that occur. This is directly proportional to the surface area available to collide also called the collisional cross section (CCS).

The earliest and simplest form of ion mobility was done using an ion drift tube [8]. It consists of a series of concentric metal rings that have a uniform electric field across them as shown in Figure 6. The ions are released in distinct packets giving the start time (t_0) for the kinetic experiment. The stop time (t_f) is determined when the ions are detected at the other end of the drift tube. From these two times, the drift time (t_d) is calculated and presented to the operator. As shown in Equation 2, the drift time is proportional to the collisional cross section (Ω), the length of the ions trajectory (L), the drift gas pressure (P) and density number (N_0) and the relative mass of the ion and the gas atoms (μ). These values influence the frequency and momentum transfer efficiency of the collision events. The drift time is inversely proportional to the amplitude of the accelerating electric field (E) and its effect relative to the ions charge (Z). This influences the initial velocity of the ion entering the tube and at high values, the less time is available for the ion to collide with the drift gas.

$$t_d = \Omega \frac{16}{\sqrt{18\pi}} \frac{\sqrt{k_B T}}{Z\, e} \frac{L}{E} \frac{P}{720} \frac{273}{T} \sqrt{\mu} N_0 \qquad \text{(eq. 2)}$$

The advantage of the drift tube is that the ion's trajectory is known to be equal to length of the drift tube. This allows the measurement of the absolute collisional cross section, which can be compared to computational values. The main disadvantage of the ion drift tube is the relation of pressure and tube length. In order to attain high resolution separation either a long tube is needed which increases the overall loss of signal or higher pressures are needed which do not go well with mass spectrometers as they use a decreasing pressure differential to aid in drawing the gaseous molecules into the instrument. The addition of a high-pressure cell in the middle of the ion's path towards the detector would destabilize the ion beam leading to a loss of signal.

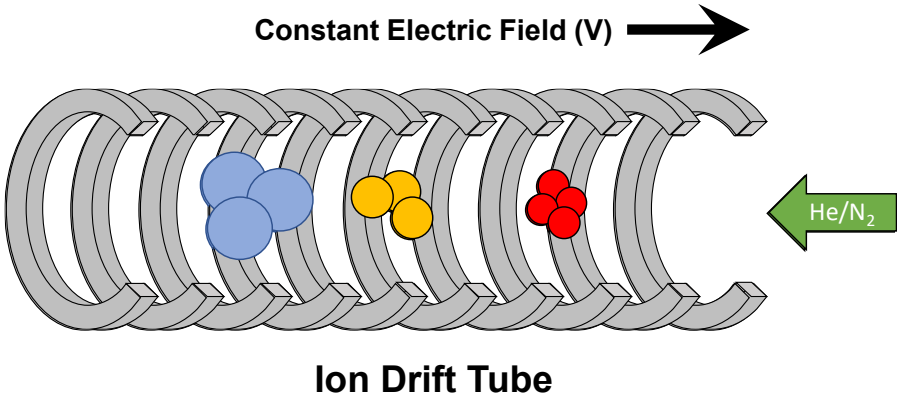


Figure 6. Ion Drift Tube Schematic. The ion drift tube has a constant electric field that accelerates the ions. As they pass through a counter current of neutral gas (He/N_2), the ions are slowed proportionally to their collisional cross section which allows for their separation.

2.4.2 Travelling wave Ion mobility.

Different designs exist for incorporating an ion mobility cell within the flow of a mass spectrometer. The one designed by Waters Corp. for the Synapt series mass spectrometers uses a travelling wave ion mobility separation (TWIMS) system. It is built on a similar design as the drift tube, except it uses an alternating electric field instead of a uniform one as shown in Figure 7. The current is applied to sets of five adjacent rings with no electric field in between sets. The active sections move down the cell to form wave-like burst of acceleration. The rate at which the active sections move through the cell and the amplitude of the applied electric field are called the wave velocity and height respectively, both can be set by the operator [9]. This process effectively increases the length of the ion's total trajectory while maintaining low drift gas pressures. The disadvantage to this system is that the trajectory length is unknown. This uncertainty means that the effective CCS values cannot be calculated from the drift times, but CCS values can be calculated relative to a drift time calibration curve. In general, the calibrant of choice should be chemically similar to the analytes. For this reason, dextran 1000 was used as the calibrant using the method and theoretical CCS values developed by Hofmann et al [10]. It is a mixture of low-mass glucose polymer ranging between 3 to 25 units joined by (α1,6) glycosidic bonds that is produced biochemically by the *Leuconostoc mesenteroides* bacteria.

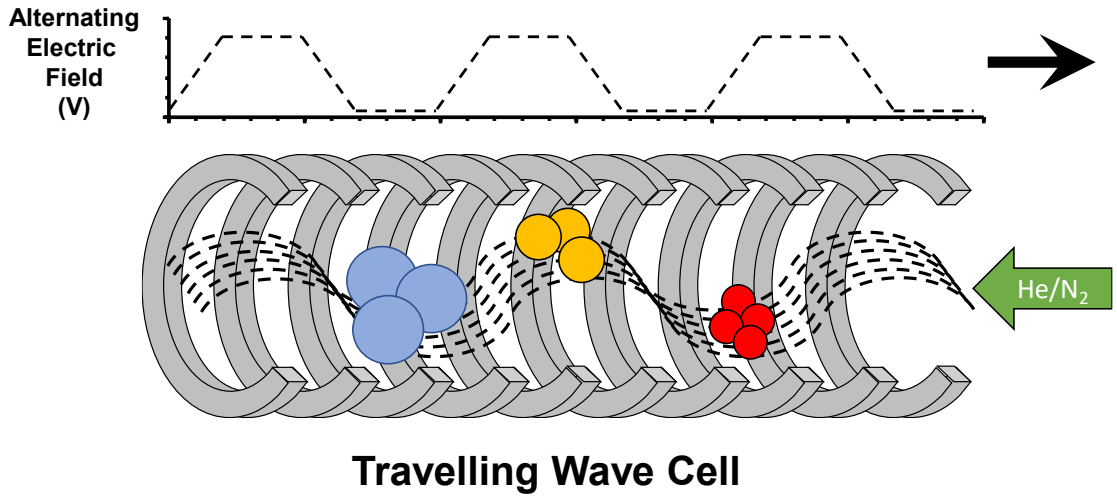


Figure 7. Schematic of a travelling wave ion mobility (TWIM) cell. Instead of a constant electric field, a TWIM cell accelerates the ions in waves by charging small sections of the cell at time.

Chapter 3. Computational Theory

3.1 Structure and transition state optimization

In order to understand the experimental results, the energetic and geometric properties of the different precursor ions and the unimolecular reactions that produce select fragments had to be calculated. This is done by using the quantum mechanical models available within the Gaussian 16 suite, specifically those for density functional theory (DFT). Unlike *ab initio* which calculates the electron distribution by solving the Schrödinger equations for both the energy and wave functions, DFT uses the less computationally expense approach of calculating the total electron density with Kohn Sham single electron equations [11]. To approximate the electron-electron interactions of polyelectron molecules, the Becke, 3-parameter, Lee-Yang-Parr (B3LYP) hybrid functionals are used [12, 13]. This method works well for calculating the geometry of the lowest energy structures. This is done by optimizing the structures and calculating the internal energy at each step which forms a potential energy surface (PES). The program then looks for a minimum which can either be the ground state geometry which we can relate to gas phase observations or a saddle-point between two different minima. The difference is determined by looking at the calculated vibrational frequencies. A saddle-point will have an imaginary frequency (read as a negative value) associated to the transitional vibrational mode. The ground structure will have no imaginary frequencies attributed to it. More than one imaginary mode means that the structure most likely failed to optimize fully and could be located on a shoulder of the PES.

The time and computational resources needed by these calculations depends not only on the functionals sets used but also the basis set. The choice of which must be a balance

between accuracy and speed. Higher basis sets with more functions to compute per atom will give more accurate results but at the cost of exponentially longer calculation times. With the calculations often run-on shared servers, there is a limit and economical cost attributed to these calculations which should be thought about when choosing a basis set. Since an unmodified trisaccharide contains 66 atoms (65 for the deprotonated species and 73 for the phosphate adducted species), and the atoms of interest are carbon, hydrogen, oxygen and phosphorus, it was decided to use the 6-31+G(d) basis set to calculate the optimised geometries of the precursor ions, the intermediates, and the products. Since basis sets tend to be very complex, only a brief summary will be given into the meaning of the basis set [14].

The 6-31G means that the inner shell or core is represented by an s orbital approximation calculated by a linear combination of six Gaussian primitive functions, or Gaussians for short. The valence shell is represented by an inner and outer orbital (hence why it is often referred to as a two-function or double zeta set). The inner orbital is calculated by a linear combination of three Gaussians and the outer orbital is calculated by a single Gaussian. The (d) label represents the addition of d-type polarization functions for all non-hydrogen atoms. Finally, the plus symbol (+) represents the addition of diffuse s- and p-functions for all non-hydrogen atoms which are used to better approximate the anion's effect on molecular orbitals.

The transition states for select reactions were completed using a mixture of calculation techniques that increase in both time demand but also success rates in finding the transition state. The first method used was the synchronous transit-guided quasi-newton (STQN) method which uses a mixture of linear and quadratic synchronous transit vectors

to correlate the changes between a reactant and product structure in the case of the two-structure approach (QST2) [15]. A three-structure (QST3) approach which includes a best-guess transition state structure can be used with greater success rates but is highly dependant on the quality of the guess work with relations to the transition state structure. A mod-redundant scan can be used to calculate the internal energy as a set transformation is scanned such as the change in angle during a [1,2] proton transfer or the stretching of a bond as it gets ready to cleave. In general, the highest energy structure along this scanned parameter is a valid guess for a transition state structure. The third method used which is also the most demanding is to take the putative transition state structure and perform a Berny optimisation in which we state that we are looking for a local energy maximum instead of a minima. Then during every optimisation step, the force constants are calculated (this is generally only done at the beginning of the calculation). This last method has had the most success at finding a transition state although with a large complex molecule such as a trisaccharide, it is still more of an art than anything.

Once a transition state structure is found, it is verified using the intrinsic reaction coordinate (IRC) method. This observes the movement of the transition states imaginary mode and optimizes the structure in steps in both the reverse and forwards reaction direction. A true transition state will converge to the same structures as the reactants and the products respectively.

For each structure, the reported zero-point was scaled to correct for the basis-set dependant bias. The scaling factor was taken from the computational chemistry comparison and benchmark database provided by the national institute of standards and

technology (NIST) [16]. The corrected zero-point energy was added to the reported electronic energy to determine the internal energy of the optimised structure.

3.2 Rice-Ramsperger-Kassel-Markus (RRKM) theory

As tandem mass spectrometry involves unimolecular reactions, the Rice-Ramsperger-Kassel-Markus (RRKM) theory can be used to study them. Equation 3 describes the formula behind the Fortran executable program developed by Professor Mayer based off his work with Tomas Baer [17]. It describes the function of the microcanonical rate constant (k) which is proportionate to the degeneracy of the reaction (σ) and the sum of the states for a transition state of interest ($N^{\ddagger}(E\text{-}E_0)$) between zero energy and the maximum energy of ($E\text{-}E_0$) where E is the given energy and E_0 is the 0 K activation energy of the reaction. The calculated function is inversely proportional to Planck's constant (*h*) and the energy dependant density of states function for the reactant (ρ).

$$k(E) = \frac{\sigma N^{\ddagger}(E - E_0)}{h\rho(E)} \qquad \text{(eq. 3)}$$

This function was used in cases when there are multiple isomeric structures of similar ground state energy as is the case for the precursor ions formed in the source. When this occurs, it becomes increasingly difficult to assess if there is a single preferred structure among them or if there is competition. For this purpose, the density of states, ρ(E), was calculated from the vibration frequencies of the optimized structures using the direct-count method of Beyer-Swinehart [18]. The calculated density of states for each structure were then normalized to the total population of the isomers at a given energy allowing a statistical determination of the dominant isomer at low energies.

Equation 3 was also used to determine the determine the competition between reaction pathways by looking at the function k(E). If the reaction rate trendlines were close to each other, there could be competition. They could sometimes intersect meaning one pathway is more favorable than the other only over a certain internal energy range [19].

3.3 Collisional Cross Section Calculations

Using the optimized structures, their collisional cross sections (CCS) can be calculated using different computational simulations. For this study, the trajectory method (TM) available with M.F. Jarrold's MOBCAL software was used [9, 20]. The program simulates the structure and orientates it randomly in a three-dimensional space. It then bombards the space with helium atoms and measures their changes in trajectory as they interact with the analyte. This is then computed into the collisional cross section of that orientation. The simulation is repeated as many times as programed with different analyte orientations. The average CCS is then calculated with its standard deviation.

Chapter 4. Procedures

4.1 Reagents and Materials

The pure trisaccharide standards: maltotriose, raffinose and panose as well as reagent grade ammonium formate and ammonium phosphate were purchased from Sigma-Aldrich (Oakville, ON, Canada). Isomaltotriose was mass selected as the trimeric species in the dextran 1000 ion mobility standard isolated from *Leuconostoc mesenteroides* was also purchased from Sigma-Aldrich. UHPLC-MS grade methanol and water were purchased from Thermo Fisher Scientific (Nepean, ON, Canada). Millex® PTFE syringe filter tips (1mLx0.45µm) were also purchased from Thermo Fisher Scientific.

4.2 Sample Preparation

The trisaccharide and ion mobility calibration samples were prepared at a concentration of 10 mg/mL in methanol and water (1:1) that was spiked with either 0.1 M ammonium formate to promote the deprotonation species or 0.1 M ammonium phosphate to form the adducted species. The samples were then filtered through 1mLx0.45µm PTFE syringe filters tips.

4.3 Electrospray Tandem Mass Spectrometry

Electrospray mass spectrometry was performed on a Waters Synapt G1 quadrupole-time-of-flight instrument equipped with a T-Wave ion mobility cell (Waters Ltd., Manchester, UK). The sample was directly infused into the electrospray source with a syringe pump at a rate of 50 µL/min. The capillary voltage was 3.0 kV and the cone voltage was set at 50 V. The desolvation and cone gas were both nitrogen and set to a flowrate of 175 and 40 L/h respectively. The desolvation and source temperatures were set to 250 and 200°C respectively.

When studying the deprotonated species, m/z 503 was mass-selected by the quadrupole mass filter. M/z 601 was selected for the phosphate adducted species. The ions then passed to the collision cell (transfer cell) containing argon at a flow rate of 1 mL/min. Collision-induced dissociation experiments were run with an acceleration voltage or lab frame collision energy ranging between 2 to 50 eV.

4.5 Travelling Wave Ion mobility

During the ion mobility experiments, helium was used as the drift gas at a flow of 90 L/min. The travelling wave had an amplitude (height) of 11 V and a velocity of 400 m/s. Dextran 1000 was used as an external calibrant for the relative CCS calculations.

4.6 Energy calculations via DFT and Statistical mechanics

Density functional theory calculations were performed with the Gaussian 16 suite of programs [21] available through Compute Canada (Toronto, Canada) using the B3LYP/6-31+g(d) level of theory.

These calculations were used to determine the lowest energy deprotonation site of isomaltotriose as well as the relative energy of the product ions and neutrals of the different fragmentation pathways. Transition state calculations were done with a mixture of transit-guided quasi-newton (STQN), two-structure (QST2) and three-structure (QST3), methods. Transition states were confirmed by the intrinsic reaction coordinate (IRC) method in Gaussian 16.

4.7 Collisional cross section calculations.

The collisional cross sections for isomaltotriose, maltotriose, raffinose and panose were calculated using the trajectory method in the MOBCAL program. Each analysis was done with a single structure at a time by inputting the three-dimensional geometric coordinates of the atoms and Mulliken charges provided by the DFT calculations. Each analysis involved two hundred random orientations of the input geometry.

Chapter 5. Study of the energetics of deprotonated Isomaltotriose fragmentation.

Carbohydrate fragmentation is highly complex, and the ions formed can be separated into three different families. Glycosidic bond cleavages (GBC) which consist of cleaving the C-O bond on either side of the glycosidic bond, cross-ring cleavages (CRC) when the ion results from breaking apart one of the carbohydrate monomers and satellite cleavages when either one of the hydroxyl groups is lost in the form of water or a modified group is removed.

In 1988, Domon & Costello put forward a method for naming these fragments based on the nomenclature used for protein fragmentation studies and is shown in Figure 8 [22]. They used the letters $^{k,l}A_i$, B_i and C_i to denote the ions that contain the non-reducing end of the polysaccharide and are numbered from values of i from 1 to n (n being the total number of units) to identify the number of carbohydrate rings contained in the ion. They then used $^{k,l}X_j$, Y_j and Z_j to demonstrate if the ion contains the reducing end. These are numbered from values of j from 0 to n-1. It is important to note that the sum of i and j for ions resulting from the same cleavage always equals n. The CRC ions use two superscripts (k,l) before the A and X label to demonstrate which two bonds are cleaved. Starting with the number 0 corresponding to the bond between the anomeric carbon and the oxygen of the ether group forming the carbohydrate ring, then moving clockwise around the ring. The labels B and Y correspond to breaking the glycosidic C-O bond on the side of the non-reducing end and the labels C and Z correspond to the breaking of the glycosidic C-O bond on the side closer to the reducing end.

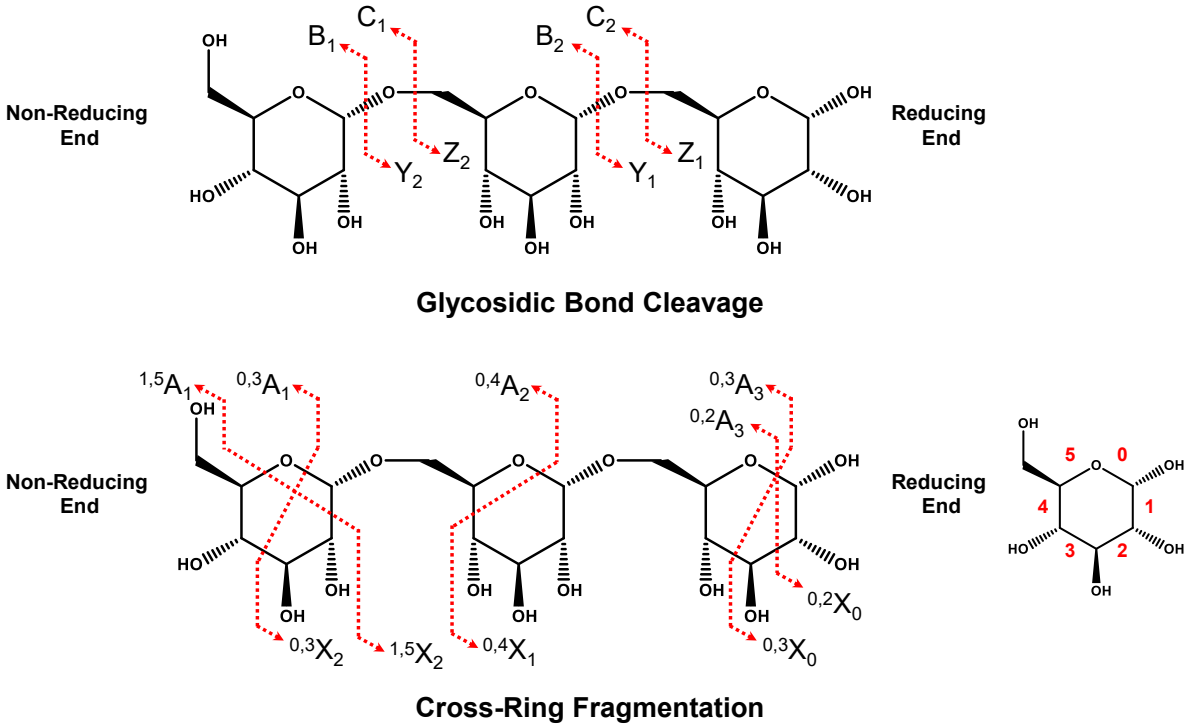


Figure 8. Domon & Costello's nomenclature for the identification of carbohydrate fragment ions. (Top) the labels C_i and Z_j represent the possible ions formed when the C-O bond on the reducing end of the glycosidic bond is cleaved. If the charge is found on the non-reducing end that gained the glycosidic oxygen, it is a C ion. If the charge is on dehydrated reducing end, it is a Z ion. The labels B_i and Y_j represent the possible ions formed when the C-O bond on the non-reducing end of the glycosidic bond is cleaved. If the charge is found on the dehydrated non-reducing end, it is a B ion. If the charge is on reducing end that gained the oxygen, it is a Y ion. (Bottom) CRC cleavage ions are represented by the labels ${}^{k,l}A_i$ if the charge is on the non-reducing end or ${}^{k,l}X_j$ if the charge is on the reducing end. The subscripts i and j represent the number of carbohydrate moieties remaining in the ion. At all times, the sum of i and j for ions resulting from the cleavage of the same bonds equal the total number of carbohydrate rings. The superscript k and l represent the two bonds cleaved during CRC.

To simplify future discussion, an addendum is going to be used to demonstrate which atom of which ring is being discussed for the purposes of deprotonation and reaction schemes as shown in Figure 9. It should be noted that this addendum applies solely to linear polysaccharide chains and is not meant for branched species. The atoms will be

numbered based on the carbon that they are attached to starting always from the anomeric carbon and proceeding clockwise around the carbon chain that forms the ring. The rings will be identified by lowercase roman numerals starting from the reducing end.

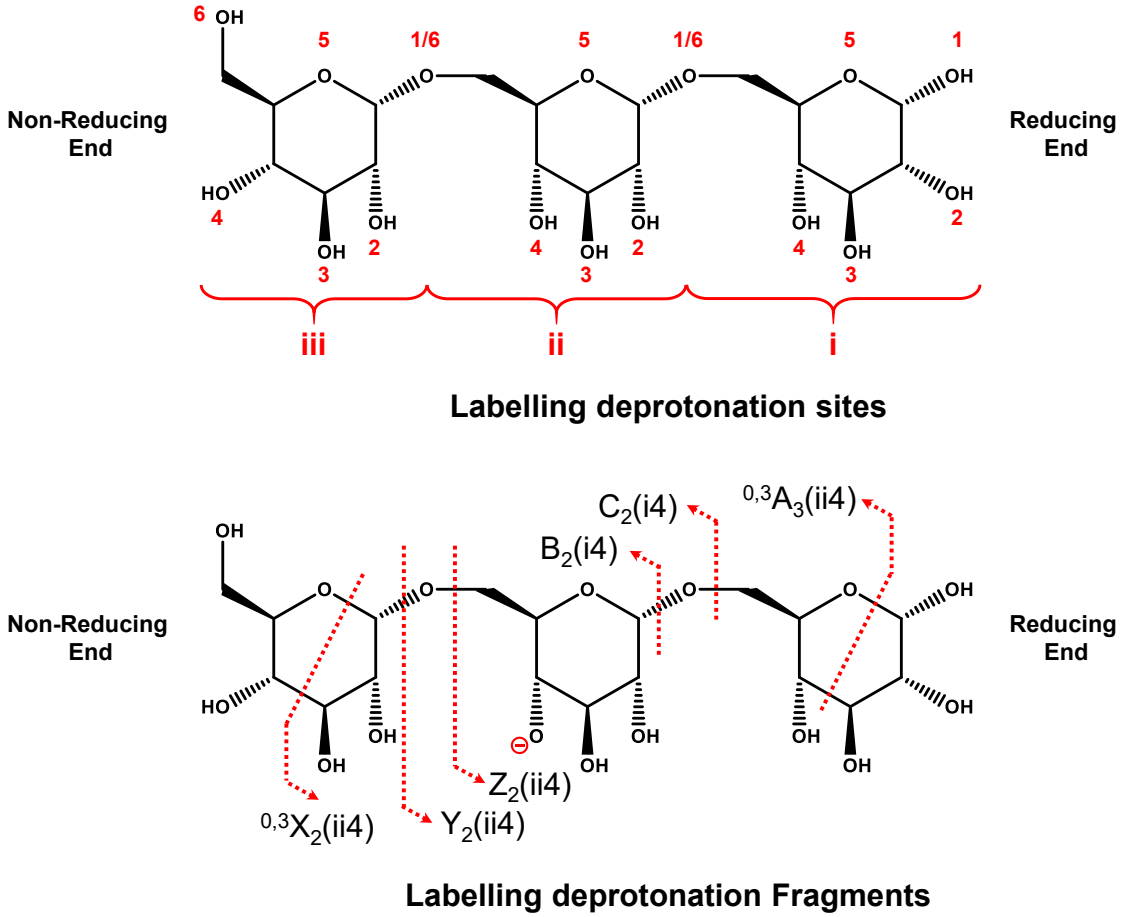


Figure 9. Nomenclature used to label the deprotonation sites of the different ions. (Top) the oxygen atoms are numbered according to the carbon they are attached to. The rings are labeled with a lowercase roman numeral starting from the reducing end. (Bottom) examples are shown for the application of this nomenclature in conjunction with Domon and Costello's nomenclature.

Mechanisms for the fragmentation of simple monosaccharides and disaccharides have previously been studied using different computation methods. Using AM-1 (Austin model) semi-empirical calculations, Carrol et al.[23] used glucose deprotonated at the anomeric carbon as a basis for determining the energetics of a charge-directed mechanism for

cross-ring cleavage by opening the carbohydrate ring followed by the loss of $C_2H_4O_2$ ion after multiple charge transfers along the opened ring. In addition, they proposed a possible charge-remote GBC mechanism for the formation of C_2 and Z_1 ions from isomaltotriose (α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→6)-α-D-glucose). It involved a [1,3]-proton transfer from the β-carbon to the ether group of the glycosidic bond, this was however not studied computationally.

Using the same computational method, Mulroney et al. [24] used disaccharide models such as sophorose (β-D-glucopyranosyl-(1→2)-D-glucose) deprotonated at the ii4 position. They proposed that the loss of water starts with the [1,3]-proton transfer from a β-carbon to the leaving hydroxyl group resulting in the formation of an enol.

Using density functional theory (DFT) and Minnesota functionals (M06-2X/6-31++G(d,p)), Rabus et al.[25] used the disaccharide lactose (β-D-galactopyranosyl-(1→4)-D-glucose) as their basis. They calculated that the deprotonation site should be the i4 position forming a hydrogen bond with the ii2 hydroxyl group. This hydrogen bonding between the two rings forms a stable 9-membered pseudo-ring structure. Using calculations and isotopic labelling, they determined that the GBC fragments at m/z 179 and 161 resulted from the same bond breaking leading to the formation of the C_1 and Z_1 ions respectively.

5.1 The fragmentation roadmap for isomaltotriose.

In order to limit the annotations on each image, the mass spectra for the CID of isomaltotriose (m/z 503) at three different lab frame acceleration energies (2, 10 and 20 eV) was separated over the three following figures. Figure 10 shows the fragment ions in the mass to charge range of 350 to 550. This corresponds to ions containing more than

two sugar units. As seen, the molecular ion decomposes by a multitude of reaction pathways to form the high abundance of fragment ions observed. These fragment ions are similar to those previously observed with deprotonated mono- and disaccharides. It can lose water (m/z 485), it can lose varying amounts of formaldehyde-like moieties $(CH_2O)_n$ (m/z 473, 443, 413 and 383) often referred to as cross-ring cleavage (CRC) or A/X ions. As the energy is increased, the precursor ion and most ions that result from cleaving across the reducing ring are gone except for the loss of $C_4H_8O_4$ which can no longer lose formaldehyde-like units.

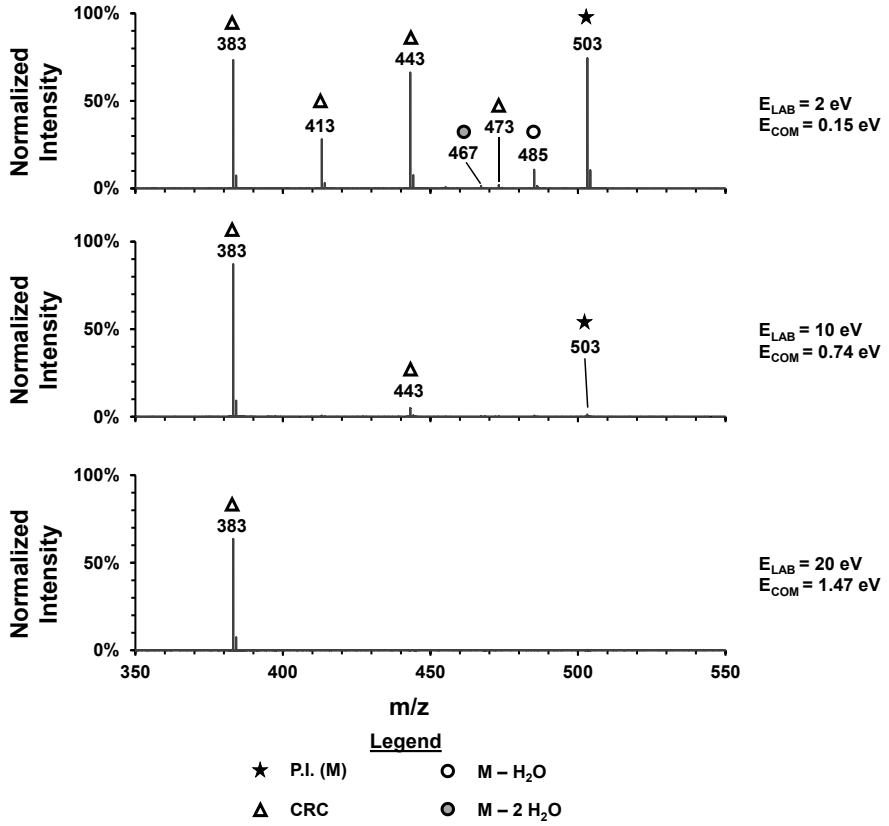


Figure 10. CID mass spectra for deprotonated isomaltriose in the 350 to 550 m/z range. The CID spectra for the lab frame collision energies of 2 (top), 10 (middle) and 20 eV (bottom) over the 350 to 550 m/z range show the precursor ion (m/z 503) the fragment ions containing more than two sugar units. The precursor ion either loses water or undergoes CRC. As the energy increases only ion from the equivalent loss of $C_4H_8O_4$ remains in this mass range the rest of the ions contain two or less sugar units.

Afterwards, the glycosidic bond between either ring can be cleaved on either side of the bridging oxygen (m/z 341,323, 179, 161). Figure 11 shows the region between 200 and 350 m/z which corresponds to the disaccharide fragments ions. It shows a similar trend as the precursor ion. The 341-ion corresponds to either a C_2 or Y_2 ion resulting from the loss of a dehydrated carbohydrate moiety and is the most abundant ion at low energies. The other disaccharide GBC ions are significantly lower in abundance. They are the 323-ion corresponding to a B_2 or Z_2 type ion of a disaccharide containing a singly dehydrated carbohydrate and the 305-ion corresponding to a doubly dehydrated disaccharide that could result from a B/Z type fragmentation from the 485-ion. It is also possible that the 323 and 305-ions are also sequential loses of water from the 341-ion. After cleaving a carbohydrate unit, the newly formed reducing end can open and undergo CRC reactions. The 341-ion would produce the 311, 281, 251 and 221 ions. The 323-ion would produce the 263 and 203 ions (the 293 and 233-ions were seen in very low abundance). The 305-ion would produce the 245 and 185 ions (shown in Figure 12).

At the low mass region of the spectra, the trend continues after the loss of the second carbohydrate to form the monosaccharide fragments. Since there are multiple ways to produce the monosaccharides, the mass range between 50 and 200 m/z is more complex to analyse as shown in Figure 12. The main GBC fragment is the 179-ion corresponding to either a C_1 or Y_1 ion or deprotonated glucose. This can come directly from the precursor ion or from the C_2/Y_2 (341) ion undergoing a second C/Y type fragmentation. The 161-ion which corresponds to a B_1 or Z_1 ion can come from three possible sources: as a direct fragment from the precursor ion, from a 341-ion that then undergoes a B/Z type fragmentation or a 323-ion undergoing a C/Y type fragmentation. The 143 ion has the

lowest abundance of the GBC ions since it results from consecutive B/Z type fragmentations to produce the doubly dehydrated species.

The monosaccharide GBC ions will then open the last carbohydrate ring and form CRC ions. The 149, 119, 89 and 59-ions come from the C_1/Y_1 (179) ion and are more abundant than the other monosaccharide CRC ions. The 131, 101, 71 ions are produced by the B_1/Z_1 (161) ion. Lastly, the 113-ion comes from the doubly dehydrated 143-ion.

All the observed ions can be used to determine the complete fragmentation roadmap as shown in Figure 13 that ties the fragment ions with their sources. This demonstrates the high degree of complexity in the fragmentation of a simple trisaccharide. The purpose of this map is to relate any mass shifts to the possible type and location of any biochemical modifications to the glycan. To better understand these processes, this study looks into the energetics of the primary fragmentation pathways stemming solely from the precursor ion in order to determine the possible competition between them.

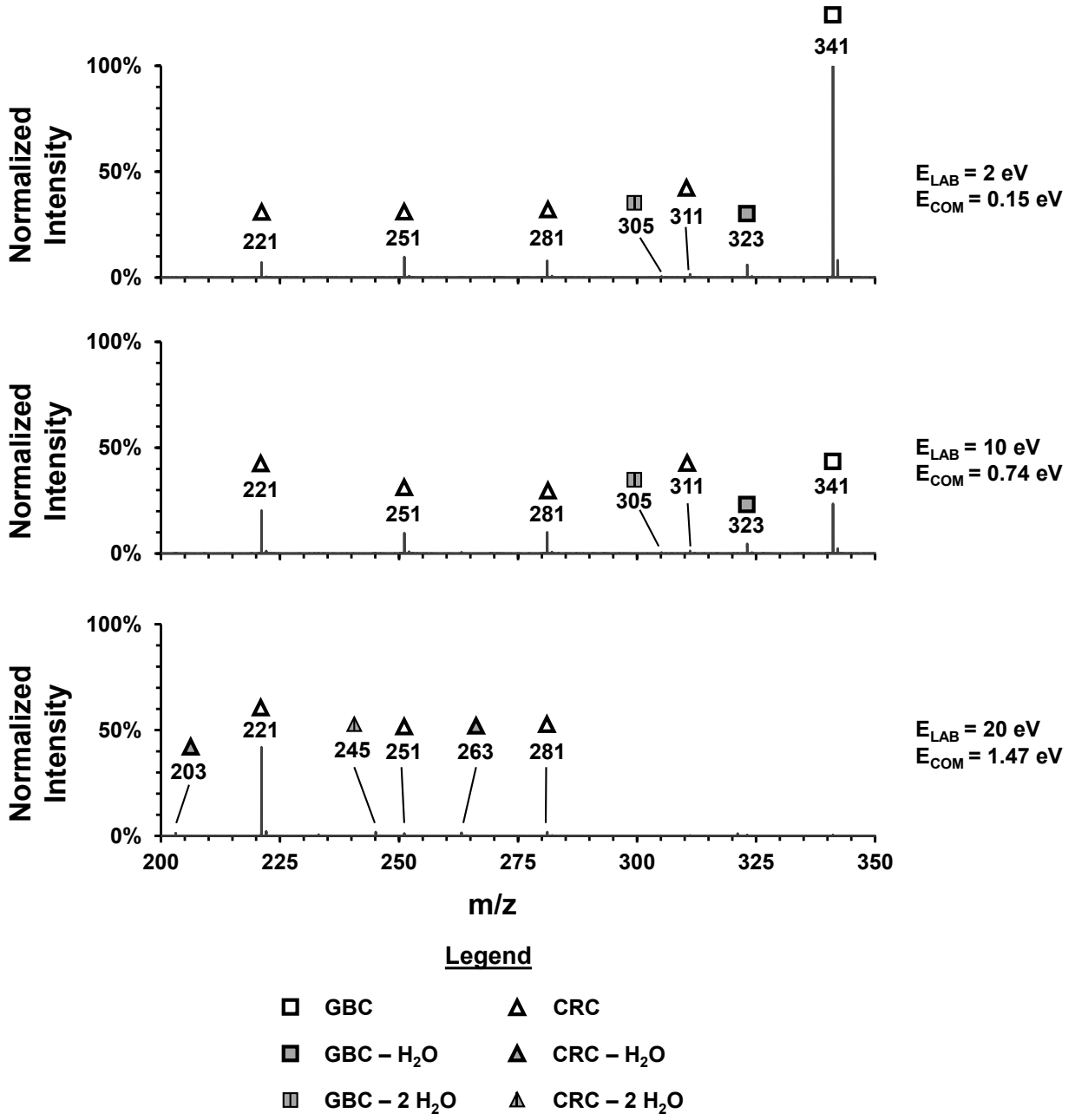


Figure 11. CID mass spectra for deprotonated isomaltriose in the 200 to 350 m/z range. The CID spectra for the lab frame collision energies of 2 (top), 10 (middle) and 20 eV (bottom) over the 200 to 350 m/z range show the fragment ions of the dimeric species. The 341 and 323 ions correspond respectively to the disaccharide and dehydrate disaccharide ions that result from GBC. These then start to disappear as they undergo similar water-loss or cross-ring cleavage. At high energy, the main remaining ion in this mass range corresponds to the equivalent loss of $C_4H_8O_4$ from the 341 ion.

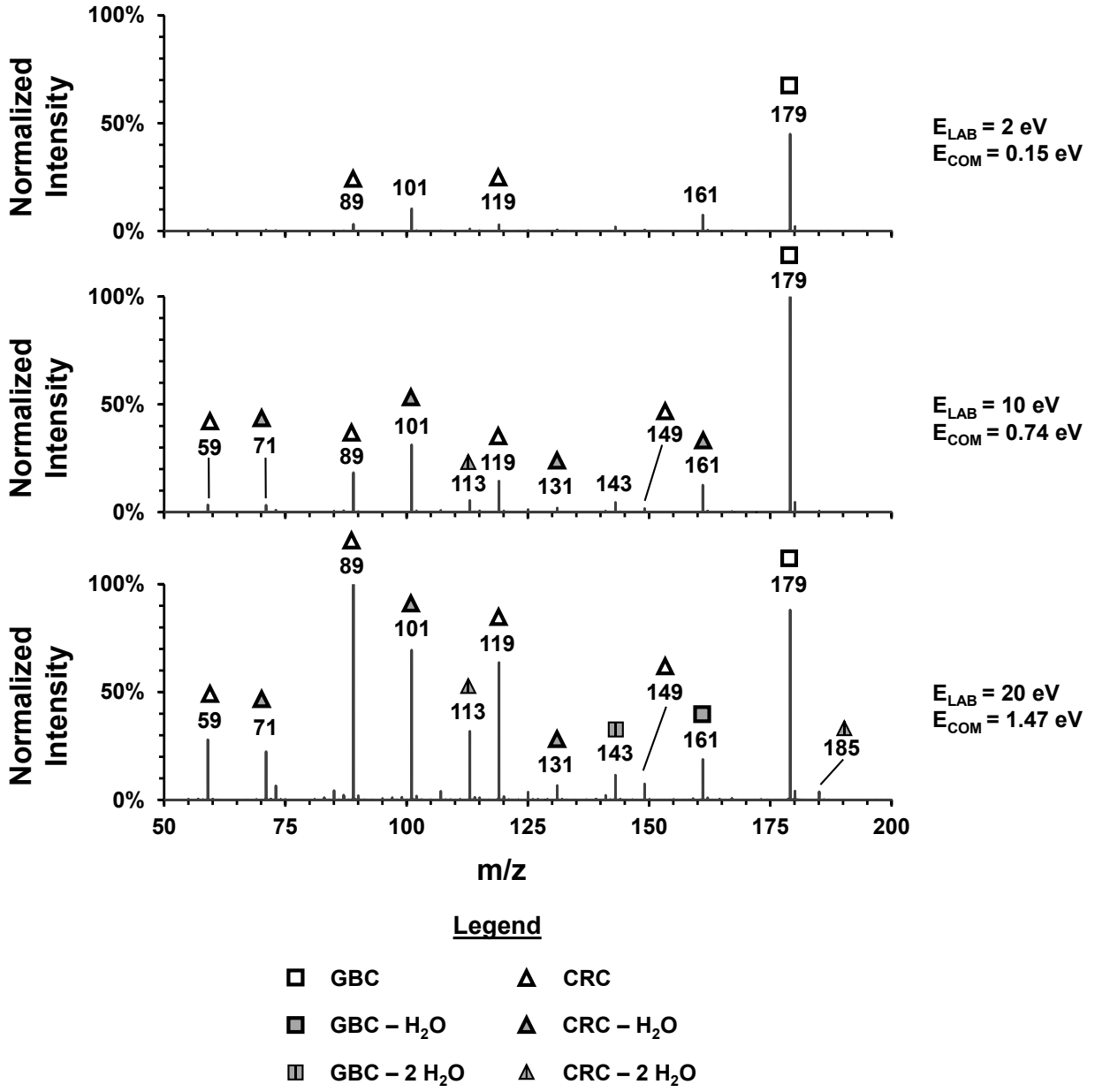


Figure 12. CID mass spectra for deprotonated isomaltriose in the 50 to 200 m/z range. The CID spectra for the lab frame collision energies of 2 (top), 10 (middle) and 20 eV (bottom) over the 50 to 200 m/z range show the fragment ions of the monomeric species. The 179 and 161 ions correspond respectively to the monosaccharide and dehydrate monosaccharide ions that result from GBC. These then start to disappear as they undergo similar water-loss or cross-ring cleavage. At high energy, the main remaining ion in this mass range corresponds to the equivalent loss of $C_3H_6O_3$ from the 179 ion.

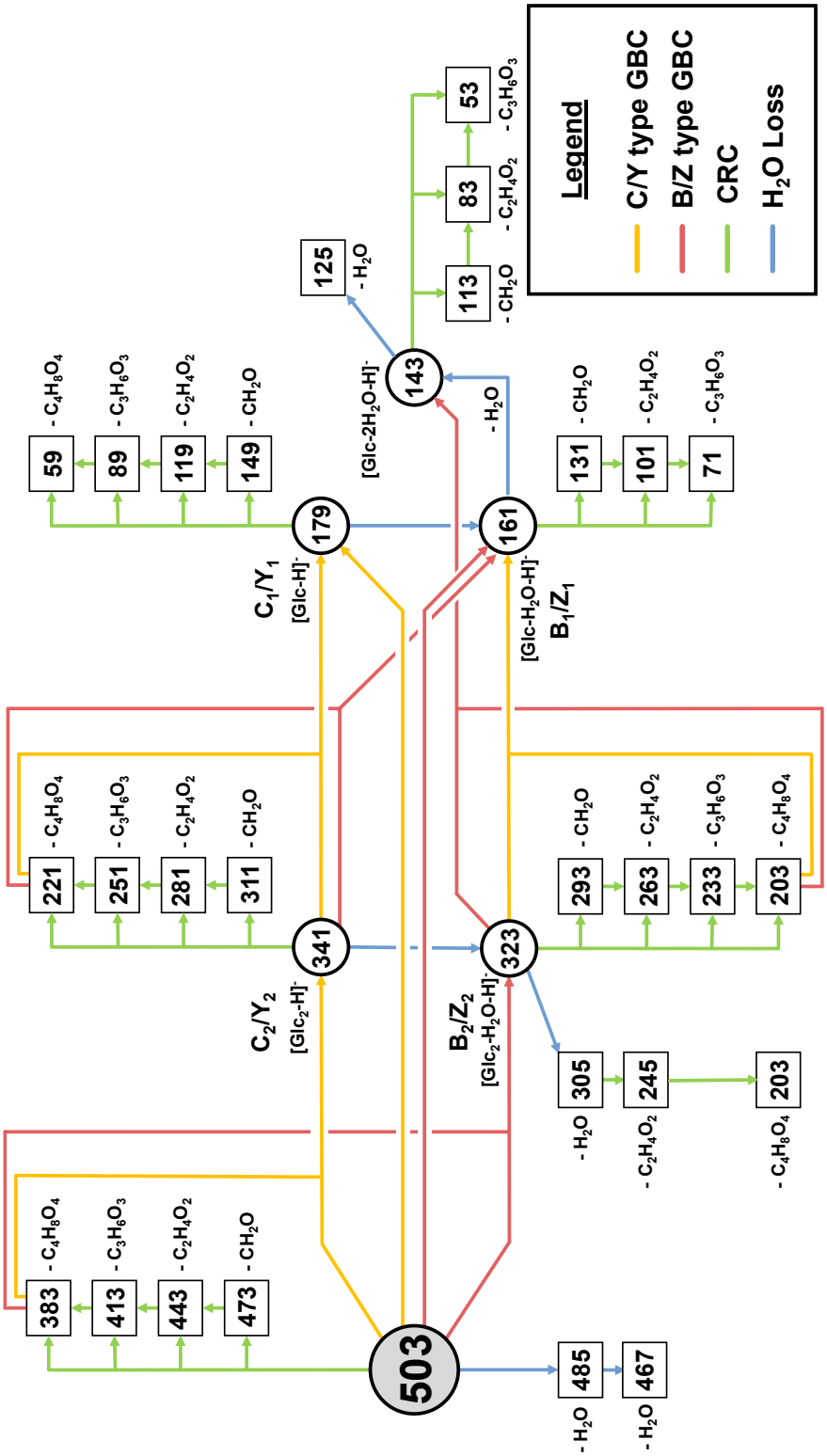


Figure 13. Fragmentation roadmap for deprotonated isomaltotriose (m/z 503).

5.2 Determining the structure of deprotonated isomaltotriose

The first step of studying the energetics of carbohydrate fragmentation is to determine the reactant of the reaction, the precursor ion coming out of the source. Any one of the eleven hydroxyl groups in isomaltotriose is a possible ionization site for the deprotonation reaction.

The most stable forms of gas-phase isomaltotriose leaving the source contain hydrogen bonding between the rings formed between the fourth hydroxyl and the second hydroxyl of the adjacent ring as seen by Rabus et al. The mobility of the proton along this hydrogen bond makes it an easier proton to remove and the hydrogen bridge is strengthened by the negative charge. Trisaccharides like isomaltotriose have two different intercyclic hydrogen bonding sites. According to the calculated internal energies shown in Table 1, the system is most stable when the charge is shared between the fourth hydroxyl group on the middle ring (ii4) and second hydroxyl group of the non-reducing end (iii2). Optimized structures for all species in Table 1 can be found together in Figure 14 or individually in annex 2.

The activation energy for the charge transfer between these two sites was determined to be negligeable at the level of theory used and more energy is needed to transfer the charge out of this bridging position for charge-directed reactions to proceed elsewhere in the molecule.

Table 1. Relative Internal energies of isomaltotriose deprotonated at the different hydroxyl groups. The energies are shown relative to the lowest energy isomer (ii4).

Precursor ion deprotonation site	Relative internal Energy (eV)
i1	0.90
i2	1.14
i3	0.66
i4	0.57
ii2	0.61
ii3	0.41
ii4	0.00
iii2	0.06
iii3	0.32
iii4	0.43
iii6	0.71

To determine the energetic bias for either the ii4 or iii2 position as a starting structure, the density of states for the different precursor ions was calculated from the vibration frequencies of the optimized structure. At the lowest energy used in the experiment (0.15 eV), the ii4 position is the dominant species with the iii2 position gaining in importance as the energy is increased, to a maximum of 30%, as shown in Figure 15. All the other positions never went beyond 0.5% of the total composition.

Although the ii4 isomer is the most likely starting point for the reaction mechanism, the relatively barrierless transition between it and the iii2 isomer does not exclude either from both being a valid precursor ion moving forwards.

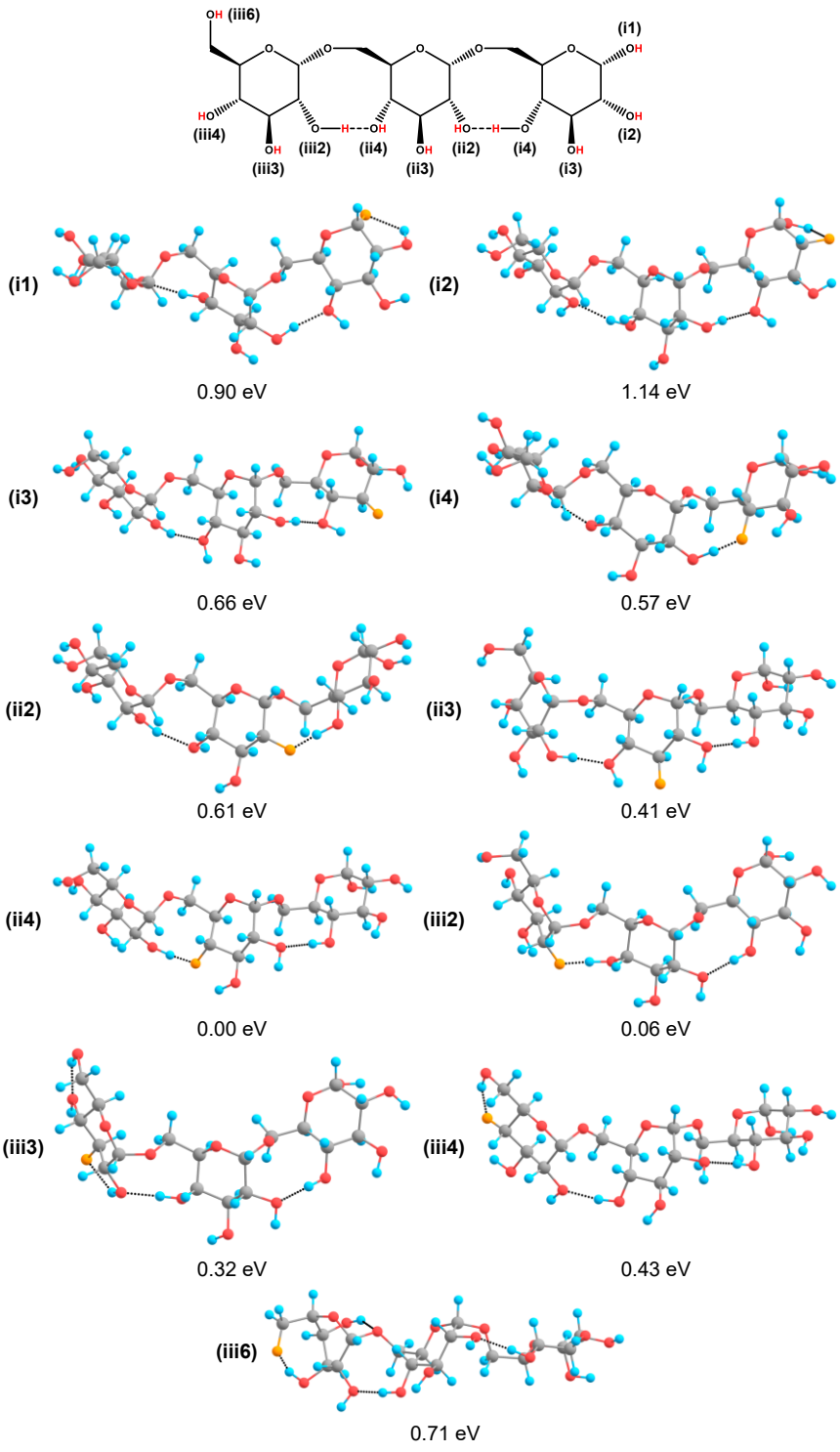


Figure 14. Optimized structures and energies of isomaltotriose deprotonated at different hydroxyl groups. The deprotonated oxygen is coloured orange instead of red to highlight its location.

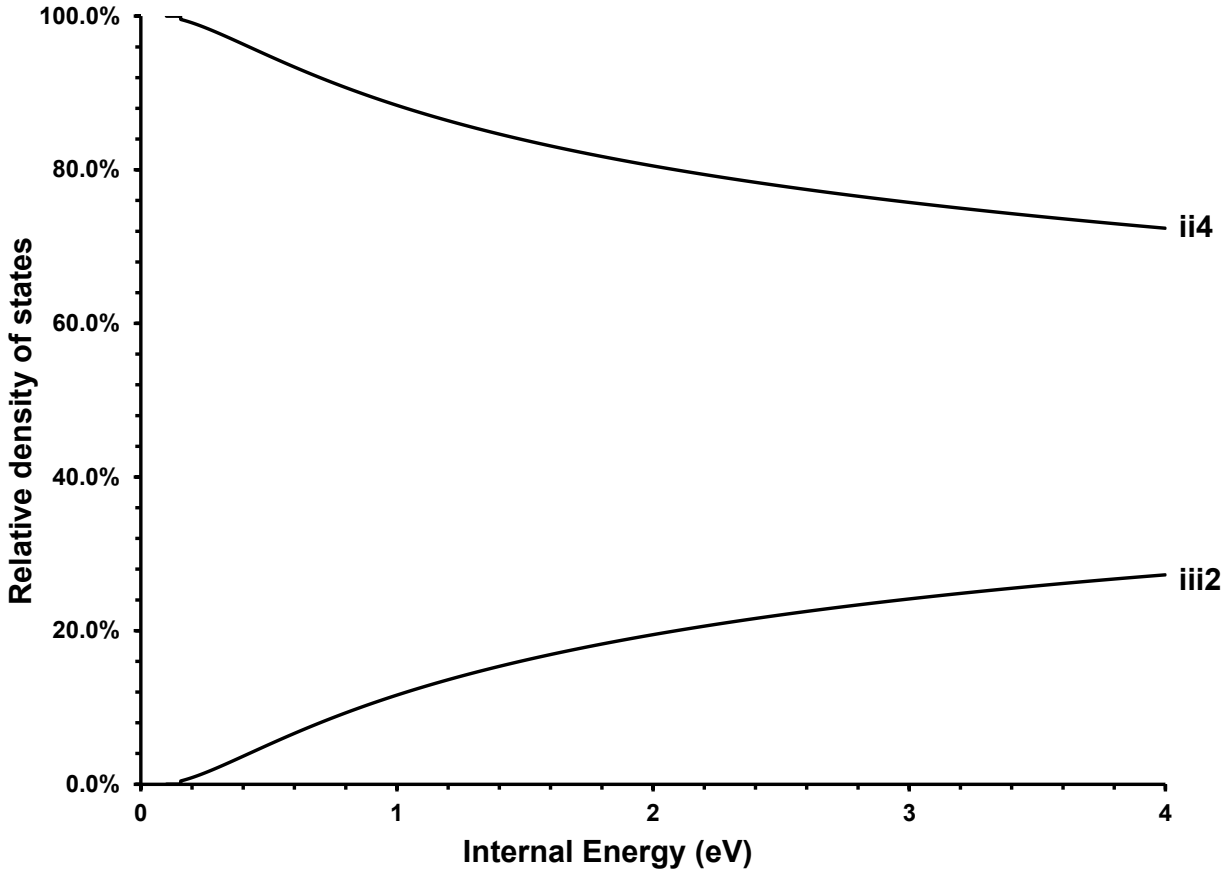


Figure 15. Relative density of states for the isomers of deprotonated isomaltotriose. The ii4 species is the primary precursor ion as it occupies more states than the other conformers. The iii2 species presents some competition at higher energies and is considered a secondary precursor ion. All of the other species never occupied more than 0.5% of the possible states.

5.3. Glycosidic bond Cleavage

Isomaltotriose has two glycosidic bonds meaning there are four different possible cleavage sites as shown in Figure 16. Figure 17 shows that the relative abundance of the intact sugar cleavages (C/Y type) are the dominant fragmentation types when compared to the dehydrated sugar cleavages (B/Z type). The maxima for the abundancies demonstrate that the C_2/Y_2 (m/z 341) fragment ions are the first glycosidic bond cleavage ions formed then the B_2/Z_2 (m/z 323) ions hit their maxima at a slightly elevated internal

energy. The disaccharides then dissociate in a similar manner to the trisaccharide and can form the C_1/Y_1 (m/z 179) or B_1/Z_1 monosaccharide ions. The monosaccharide ions could also be formed in competition with the disaccharide formation pathway as well as sequentially.

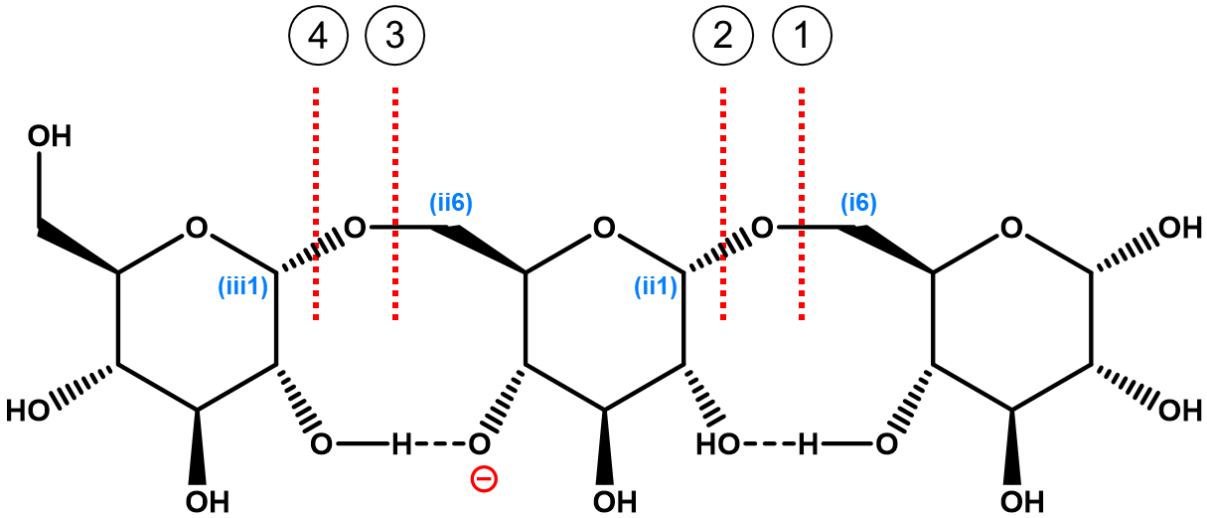


Figure 16. Illustration of the four different bonds broken by glycosidic bond cleavage. The order presented from the reducing end to the non-reducing end is the same as they will be presented.

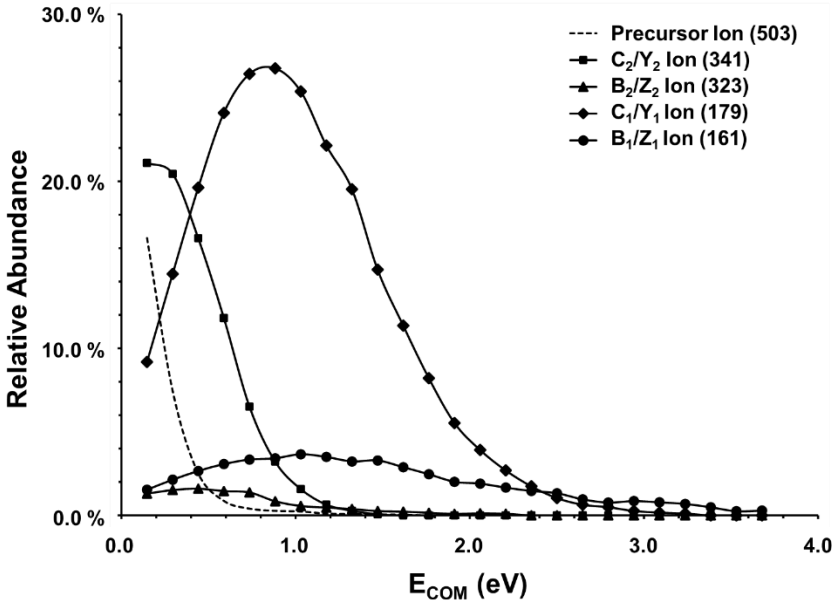
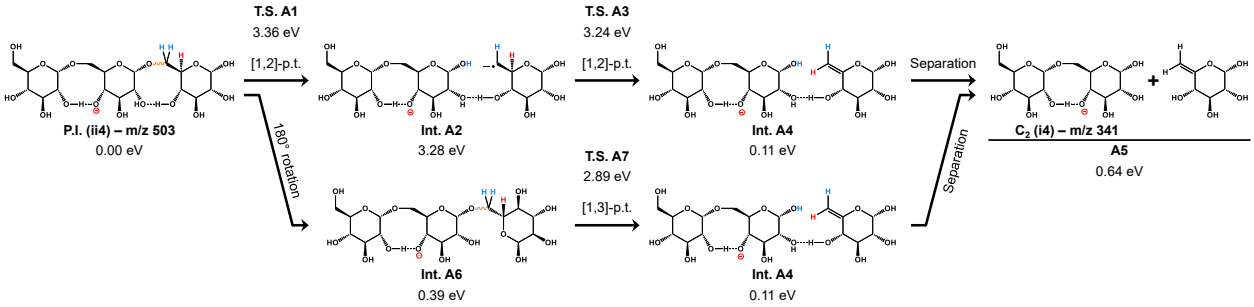


Figure 17. ESI-MS/MS breakdown diagram of the different glycosidic bond cleavage ions. As the precursor ion dissociates, the C/Y type ions are preferentially formed over the B/Z type ions. When comparing the disaccharide lines (m/z 323, 341) to the monosaccharide lines (m/z 161, 179) it can be seen that although there may be some competition between the neutral loss of a monosaccharide or a disaccharide, these pathways are most likely sequential in nature.

5.3.1 Glycosidic bond cleavage 1: O(ii1)–C(i6)

Starting at the reducing end, the first glycosidic bond is between the sixth carbon of the reducing ring and the oxygen attached to the first carbon of the middle ring. As shown in Scheme A, the charge-remote cleavage of this bond requires the transfer of a proton to the oxygen to form a hydroxyl group and an alkene between the fifth and sixth carbons of the reducing ring. The proton donated for the formation of the hydroxyl group can come from the alpha carbon in the form of a [1,2] proton transfer (p.t.); however, this leads to a high energy primary radical anion as the glycosidic bond is broken (A2). If formed, it would quickly undergo a second [1,2] proton transfer between the beta and alpha carbon to form the stable ion-neutral complex (A4). A second, and more probable, option mentioned by

Carroll et al. involves a [1,3] proton transfer directly from the beta carbon to immediately form the ion-neutral complex without forming a radical anion. This mechanism is not immediately available to the precursor ion since the beta-proton is in the staggered position. To make the beta-transfer available, the reducing ring must rotate 180° around the C(i5)–C(i6) bond to put the proton in a favorable eclipsed position as demonstrated in Figure 18. The availability of the [1,3] pathway significantly lowers the energy needed to form the C_2 ion.



Scheme A. Charge-remote glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of the C_2(i4) ion (m/z 341). The two α-hydrogens of the glycosidic oxygen are coloured in blue and its β-hydrogen is labeled in red to help distinguish the proton movement in the [1,2]-p.t. pathway (top) and the [1,3]-p.t. pathway (bottom). A larger version of the scheme can be found in annex 1.

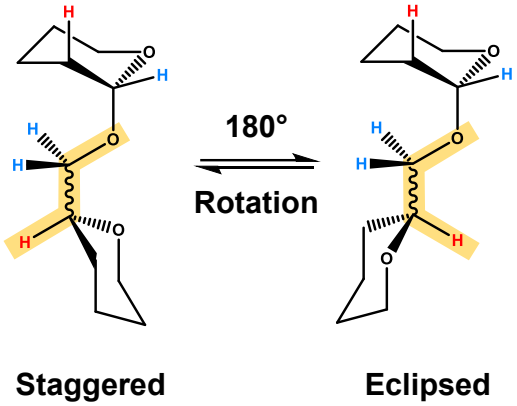


Figure 18. Staggered and eclipsed conformations of the reducing end glycosidic β-hydrogen. As the ring on the reducing end is rotated around the C_5-C_6, the β-hydrogen is moved into an eclipsed conformation allowing the [1,3]-p.t. to proceed.

The competition between the sequential [1,2]-p.t. and the [1,3]-p.t. pathways was probed using RRKM theory. As shown in Figure 19, the rate constants for the direct [1,3]-pathway is significantly higher than the sequential mechanism. The trends show parallel trends meaning that no matter what the internal energy of the molecule is, it will take the [1,3]-p.t. pathway if available.

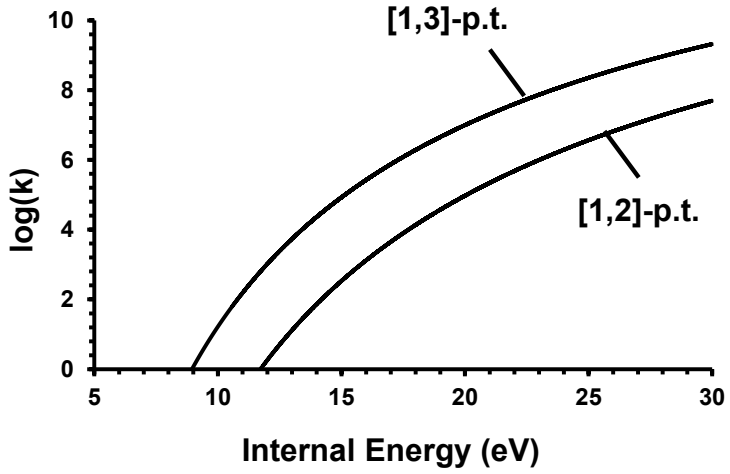
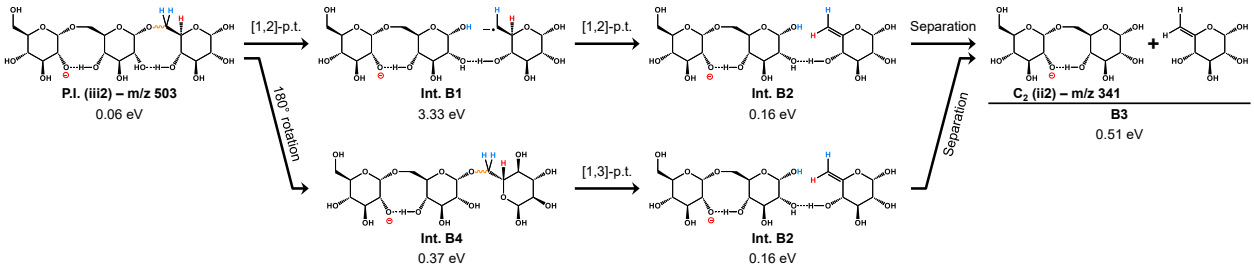


Figure 19. Internal energy dependant rate constant curves for the [1,2] and [1,3]-p.t. pathways.

Previous studies have shown the charge-directed formation of the Z_1 ion by breaking the same bond followed by charge transfer to the neutral during the ion-neutral phase (A4); however, this requires the charge to migrate outside of its favored position between the non-reducing ring and the middle ring. Transferring the charge to the iii2 position at any point of the mechanism leads to an isobaric ion, C_2(ii2) with negligeable differences in energy for the intermediates and only a minor drop in energy of 0.13 eV for the product as demonstrated in Scheme B.

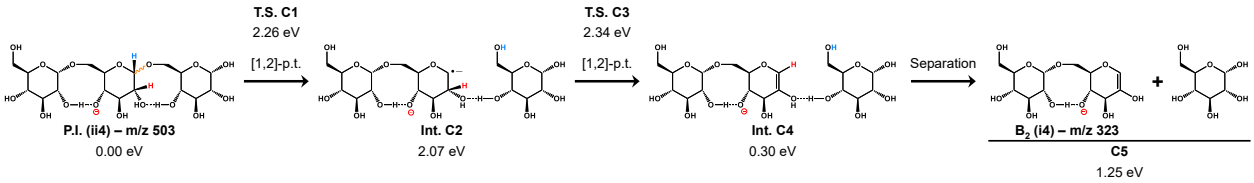


Scheme B. Charge-remote glycosidic bond cleavage mechanisms for the secondary precursor ion leading to the formation of the C_2(ii2) ion (m/z 341). The two α-hydrogens of the glycosidic oxygen are coloured in blue and its β-hydrogen is labeled in red to help distinguish the proton movement in the [1,2]-p.t. pathway (top) and the [1,3]-p.t. pathway (bottom). A larger version of the scheme can be found in annex 1.

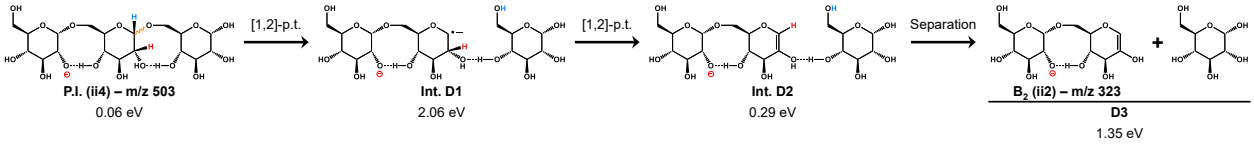
5.3.2 Glycosidic bond cleavage 2: C(ii1)–O(ii1)

The second possible bond cleavage is between the first carbon of the middle ring and the oxygen attached to it. As demonstrated in Scheme C, a proton is transferred from the alpha carbon C(ii1) to the attached oxygen and the bond is broken (C2). This [1,2] proton transfer forms a radical anion on a secondary carbon instead of a primary as seen in structure A2, the latter being a significantly more stable intermediate with a 1.21 eV drop in internal energy. The second [1,2] proton-transfer then forms the π-bond between C(ii1) and C(ii2) of the B_2(i4) ion complexed with a neutral glucose moiety. Contrary to the previous mechanism, this cleavage can only go through the sequential [1,2] proton transfer pathway since there is not enough rotational freedom to reposition the beta-proton for the direct [1,3] proton transfer. As the alpha proton is transferred and the bond is broken, the radical anion is formed on a secondary carbon instead of a primary carbon seen in structure A2. This significantly stabilizes the structure resulting in a 1.21 eV drop in relative internal energy. Like the previous mechanism, if the charge is shifted to the iii2 position at any point of the reaction, the relative energies of the intermediates are

indistinguishable, and the final product, B_2 (ii2), which is isobaric B_2(i4) has a slightly lowered relative energy by 0.11 eV as shown in Scheme D. Assuming these reactions are rapid, the charge would not migrate outside of the favored position enough before the separation of the ion-neutral complex to transfer to the cleaved glucose unit to form the Y_1 ion (m/z 179).



Scheme C. Charge-remote glycosidic bond cleavage mechanisms of the primary precursor ion leading to the formation of the B2(i4) ion (m/z 323). The α-hydrogen of the glycosidic oxygen is coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the scheme can be found in annex 1.

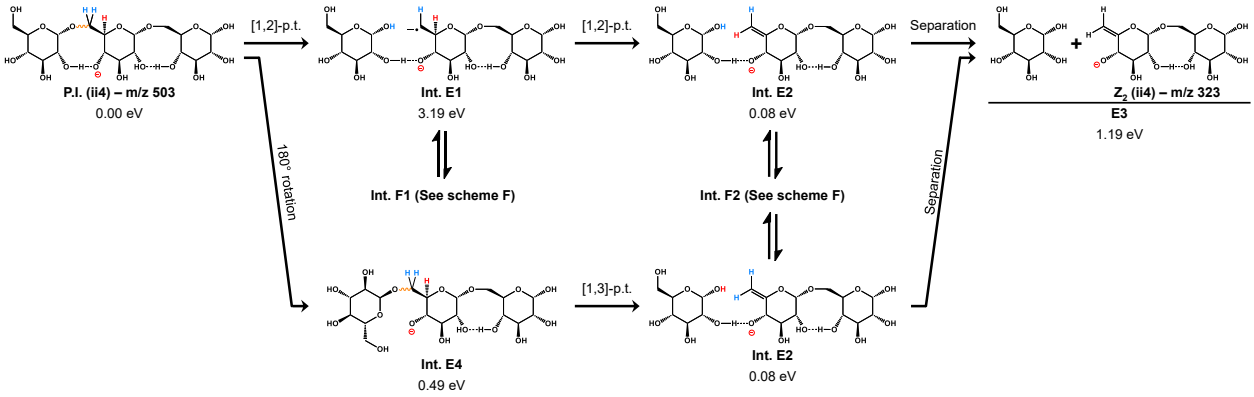


Scheme D. Charge-remote glycosidic bond cleavage mechanisms of the secondary precursor ion leading to the formation of the B_2(ii2) ion (m/z 323). The α-hydrogen of the glycosidic oxygen is coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the scheme can be found in annex 1.

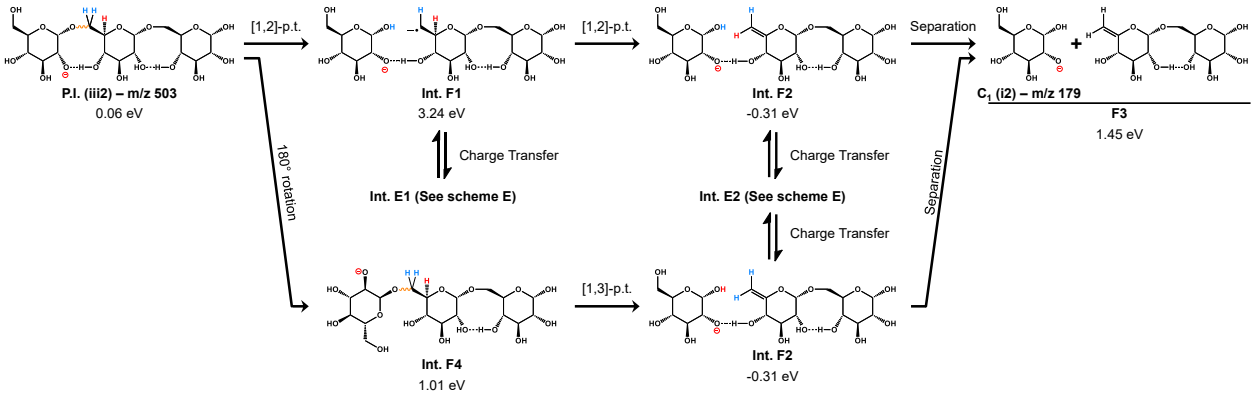
5.3.3 Glycosidic bond cleavage 3: C(ii6)–O(iii1)

Moving to the glycosidic bond that connects the central ring to the non-reducing end ring, cleaving this third bond is done through a similar mechanism as seen for the cleavage of the O(ii1)–C(i6) bond, and is shown in Scheme E. The reaction can proceed either through sequential [1,2] proton transfers or rotation around the C(ii6)–O(iii1) bond before undergoing a [1,3] proton transfer. The energies of the intermediates compared to those

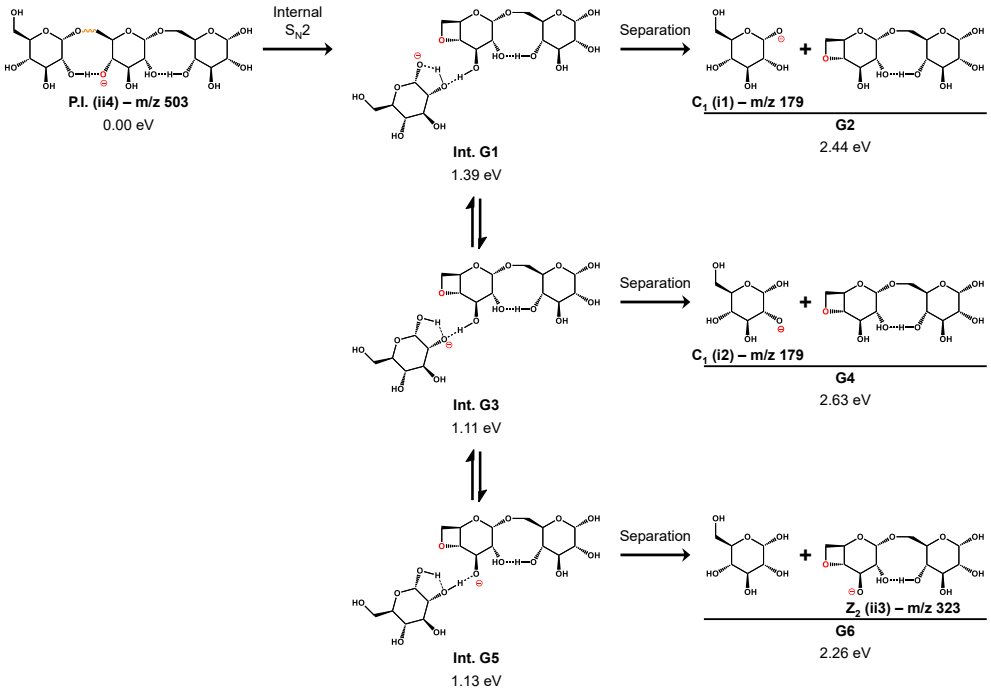
in Scheme A are relatively similar however the product, Z_2(ii4) (m/z 323), sits at a higher relative energy than the C_2(i4) ion formed. Regardless of the process, the charge can possibly be transferred from either the [1,2] intermediate (E1) or the ion-neutral complex (E2) to the neutral glucose unit to form the C_1(i2) ion (m/z 179) which is one way to form the monosaccharide ion directly from the precursor ion as shown in Scheme F. The proximity of the charge to the bond cleaved opens the possibility of a charge-directed mechanism as demonstrated for a disaccharide by Rabus et al. and was calculated and shown in Scheme G. The oxygen containing the charge, O(ii4) can bend and create a bond with C(ii6) in an internal S_N2 reaction to form α-4,6-anhydroglucopyrosyl (1→6)glucose (E2) containing an oxetane-like structure as the glycosidic bond is cleaved. The charge remains on the glucose unit but can shift charge over to the other molecule to form the Z_2(ii3) ion (m/z 323) with a slightly lower total energy.



Scheme E. Charge-remote glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of the Z_2(ii4) ion (m/z 323). Where indicated, it is possible to charge transfer into the intermediates of Scheme F. The α-hydrogens of the glycosidic oxygen are coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the schemes can be found in annex 1.



Scheme F. Charge-remote glycosidic bond cleavage mechanisms for the secondary precursor ion leading to the formation of the C_1(i2) ion (m/z 179). Where indicated, it is possible to charge transfer into the intermediates of Scheme E. The α-hydrogens of the glycosidic oxygen are coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the schemes can be found in annex 1.

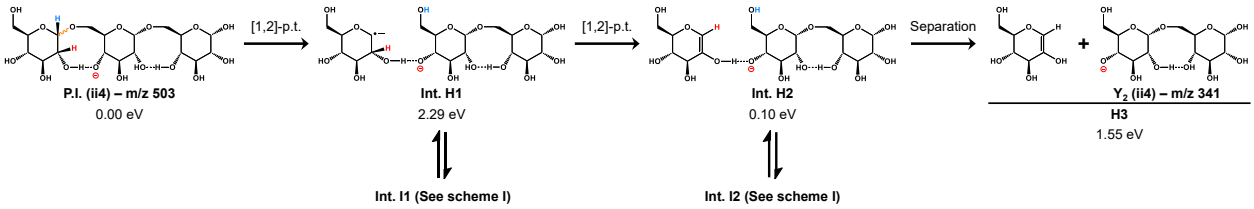


Scheme G. Charge-direct glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of either the C_1(i1) ion (m/z 179) or the Z_2(ii3) ion (m/z 323). The deprotonated oxygen (red) forms on oxetane-like structure. The ring-strain increases the product energy by approximately 1.2 eV when compared to the product energies of F3. The charge can be transferred twice to either form the less stable C_1(i2) ion or the more stable cyclic Z_2(ii3) ion. A larger version of the scheme can be found in annex 1.

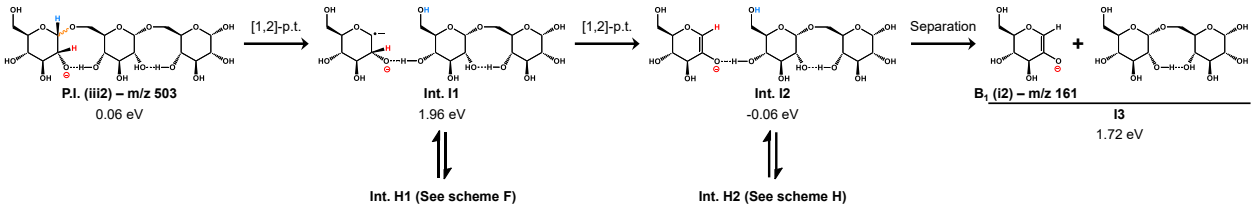
5.3.4 Glycosidic bond cleavage 4: C(iii1)–O(iii1)

Similar to the cleavage of the C(ii1)-O(ii1) bond, the fourth glycosidic bond cleavage reaction can only go through the sequential [1,2] pathway as shown in Scheme H to form the Y_2(ii4) ion (m/z 341) due to the inability to rotate the β-proton an to undergo a [1,3] proton transfer. As mentioned previously, the transition state and intermediate energies of the sequential pathway (H1 and H2) are diminished enough relative to the cleaving of the other side of the glycosidic bond since the radical anion is in a secondary position instead of the less favorable primary position.

As the charge is part of the hydrogen bridge that connects ionic and neutral moieties in the intermediate (H2) and the ion-neutral complex (H4), the charge can be transferred to form analogs as if the secondary precursor ion (iii2) was the reactant as shown in Scheme I.

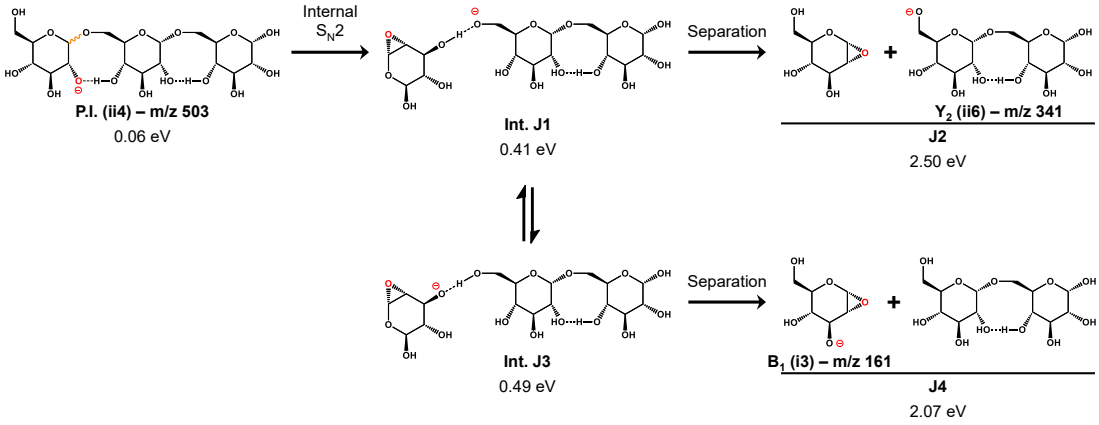


Scheme H. Charge-remote glycosidic bond cleavage mechanisms of the primary precursor ion leading to the formation of the Y_2(i4) ion (m/z 341). Where indicated, it is possible to charge transfer into the intermediates of Scheme I. The α-hydrogen of the glycosidic oxygen is coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the scheme can be found in annex 1.



Scheme I. Charge-remote glycosidic bond cleavage mechanisms of the secondary precursor ion leading to the formation of the B_1(i2) ion (m/z 161). Where indicated, it is possible to charge transfer into the intermediates of Scheme H. The α-hydrogen of the glycosidic oxygen is coloured in blue, and its β-hydrogen is labeled in red to help distinguish the proton movement during the reaction. A larger version of the scheme can be found in annex 1.

If the charge is transferred to form the iii2 precursor ion, it can undergo the charge-directed mechanism shown in Scheme J. The deprotonated oxygen, O(iii2), can for a epoxide with C(iii1) to form 1,2-anhydro-glucose as the glycosidic bond is broken (J1). The charge is transferred to the O(iii1) forming the Y_2(ii6) ion (m/z 341). It is slightly more energetically favorable to then transfer the charge before the ion-neutral complex separates (J32) to form the B_1(i3) ion (H6).



Scheme J Charge-direct glycosidic bond cleavage mechanisms for the secondary precursor ion leading to the formation of either the B_1(i3) ion (m/z 161) or the Y_2(ii6) ion (m/z 323). The deprotonated oxygen (red) forms on oxetane-like structure. The ring-strain increases the product energy by approximately 1 eV when compared to the product energies of H3. The charge can be transferred to form the more stable cyclic B_1(i3) ion. A larger version of the scheme can be found in annex 1.

5.4. Cross-ring fragmentation.

The second type of fragmentation for carbohydrates is cross-ring cleavage which consists of the loss of $(CH_2O)_n$. This is observed experimentally for isomaltotriose for n values of 1, 2, 3, and 4 as the loss of 30, 60, 90 and 120 Da respectively from either the precursor ion, or any of the glycosidic bond cleavage ions (B/C/Y/Z type ions). The diagram for the relative abundancies of the cross-ring fragment ions lost from the precursor ion, shown in Figure 20, demonstrates experimentally that the different ions are not solely formed through the sequential loss of CH_2O but more in competition with each other. After losing the first CH_2O unit (m/z 473), the fragment ion can quickly lose either 1, 2 or 3 units. This happens rapidly since the m/z 473 ion disappears shortly after the energy is increased. Regardless of the initial number of formaldehyde units lost, the ions will eventually fragment to form the $C_4H_8O_4$ loss ion (m/z 383) before it fragments further.

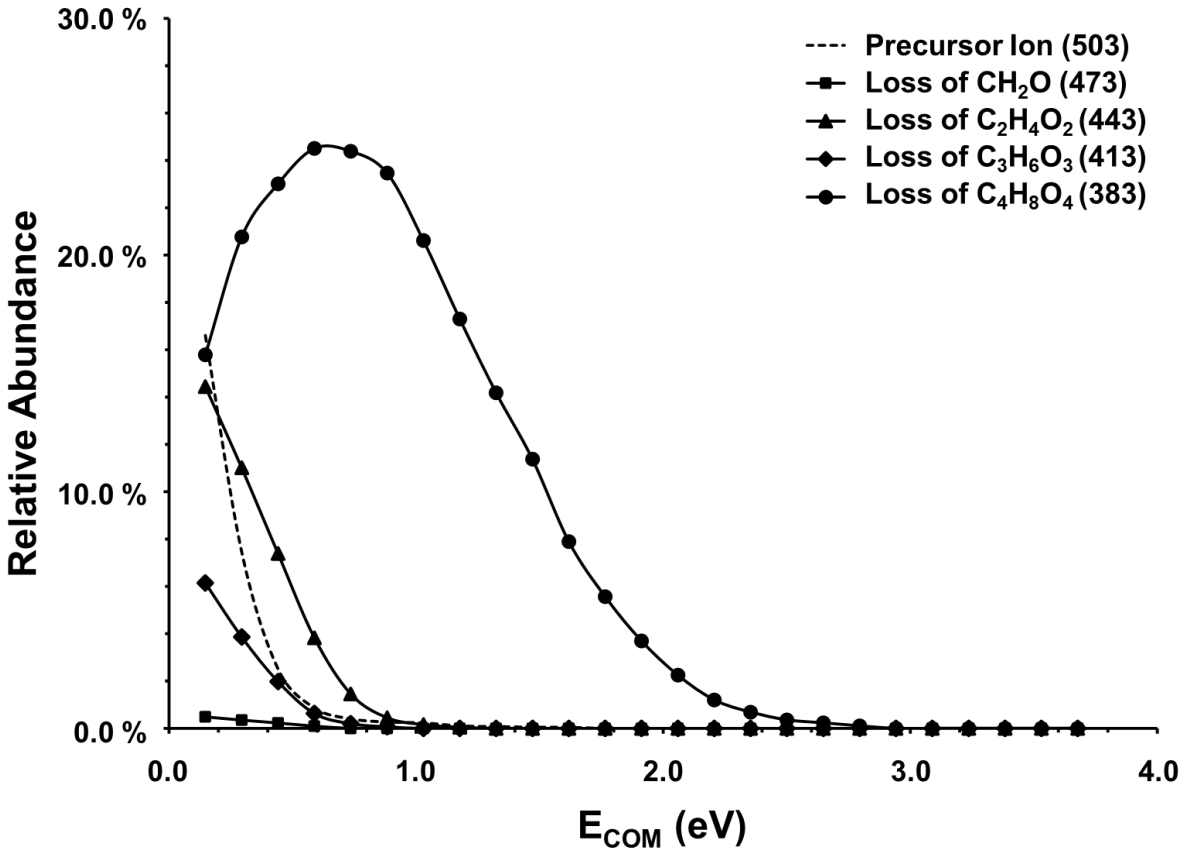
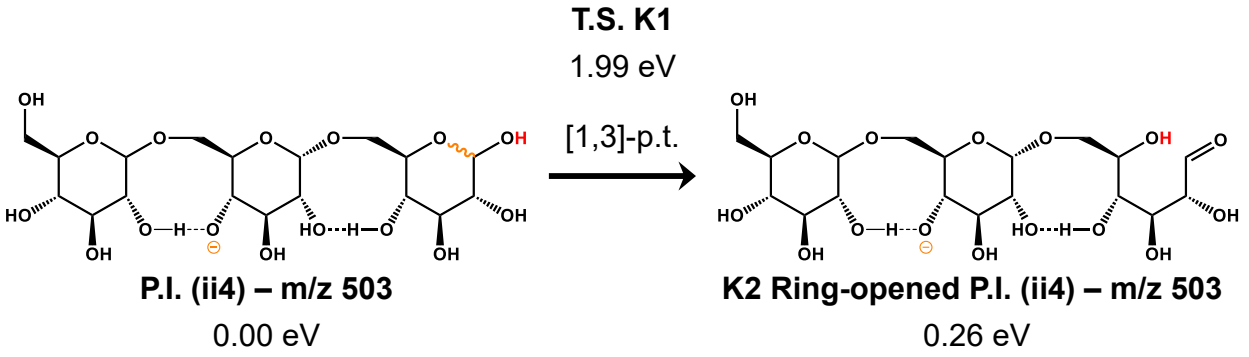


Figure 20. ESI-MS/MS breakdown diagram of the different cross-ring fragmentation ions coming from the precursor ion. As it dissociates, it can loss between one and four units of CH_2O. The loss of an even number of units is more prominent than the loss of an odd number. The ions resulting from the loss of one to three units fragment further to form the have all disappeared around 1.0 eV as they form the $C_4H_8O_4$ loss ion. Afterwards this ion losses the remains of the reducing-end sugar to form either the B_2 or C_2 ion depending on which moiety keeps the bonding oxygen.

5.4.1. Opening of the reducing carbohydrate ring.

As proposed by Carroll et al.[23], the first step of the cross-ring cleavage mechanisms is the opening of the reducing ring. Previous studies have used either charge-remote or charge-direct ring-opening reactions or a mixture of the two. Due to the distance of the charge from the reducing ring, a charge-remote method was adopted as shown in Scheme K. The proton on the anomeric oxygen undergoes a [1,3] proton-transfer to the oxygen inside the ring. The terminal aldehyde is then formed as the bond is broken. This does slightly destabilize the precursor ion as the cyclic structure is more

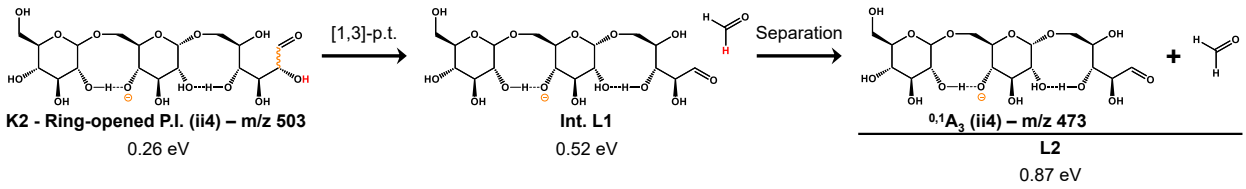
thermodynamically favorable. The energies of the open-ringed ii4 and iii2 structures both increase relative to each other. That and their distance from the reaction site mean that the mechanisms for cross-ring cleavage are effectively the same regardless of which of the two reactants are used. For simplicity and because it was determined to be the primary precursor ion, the ii4 structure will be used.



Scheme K. Charge-remote ring-opening mechanism for the primary precursor ion (ii4). The precursor ion can undergo a [1,3] proton transfer between the anomeric hydroxyl group and the oxygen forming the cyclic carbohydrate structure. A larger version of the scheme can be found in annex 1.

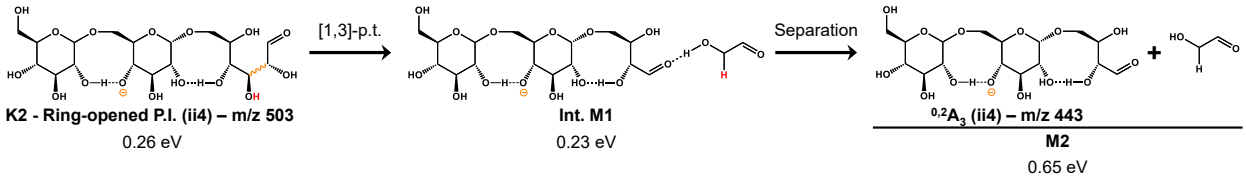
5.4.2. Cross-ring cleavage mechanisms

After the reducing ring is opened, the ion can go through four competing reaction channels involving a [1,3] proton transfer from the hydroxyl group on one side of the cleaving bond to the carbon on the other side. A new aldehyde group is formed as the bond is broken allowing for similar sequential reactions. The first and highest energy pathway is the transfer from O(i2) to C(i1) to lose formaldehyde (CH_2O) and form the $^{0,1}A_3$(ii4) ion (m/z 473 Da) as shown in Scheme L. The higher internal energy of the intermediate (L1) could be explained by the lack of hydrogen bonding needed to form the ion-neutral complex, without that added stability it quickly separates.



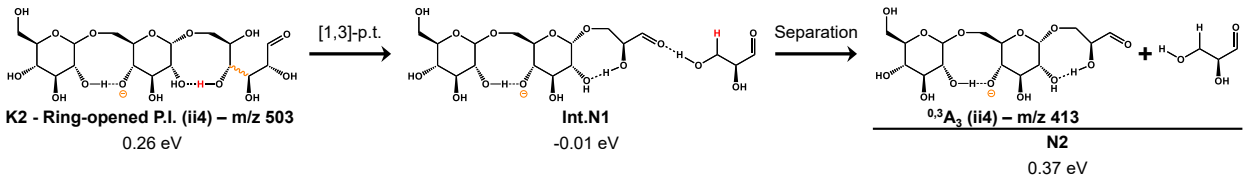
Scheme L. Charge-remote mechanism for the formation of the $^{0,1}A_3$(ii4) ion (m/z 473). Following a [1,3] proton transfer between the proton of the second hydroxyl group and the anomeric carbon on the reducing ring, a formaldehyde molecule is lost. A larger version of the scheme can be found in annex 1.

The second involves the transfer from O(i3) to C(i2) to lose hydroxyethanal or glycolaldehyde ($C_2H_4O_2$) and form the $^{0,2}A_3$(ii4) ion (m/z 443) as shown in Scheme M. The presence of the alcohol group on the neutral allows for hydrogen bonding in the ion-neutral complex intermediate (M1) which gives more stability when compared to the previous mechanism.



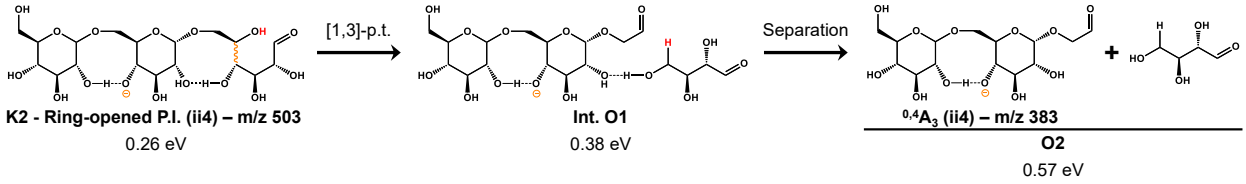
Scheme M. Charge-remote mechanism for the formation of the $^{0,2}A_3$(ii4) ion (m/z 443). Following a [1,3] proton transfer between the proton of the third hydroxyl group and the second carbon on the reducing ring, a glycolaldehyde molecule is lost. A larger version of the scheme can be found in annex 1.

The most energetically favorable cross-ring cleavage is the transfer from O(i4) to C(i3) to lose 2,3-dihydroxylpropanal or glyceraldehyde ($C_3H_6O_3$) and form the $^{0,3}A_3$(ii4) ion (m/z 413), Scheme N. The stability of the increased chain length for the neutral aldehyde helps lower the energy requirements for the product. Another factor that stabilizes these structures is that the intercyclic hydrogen bridge between the reducing and middle rings at O(i4) and O(ii2) shifts to O(i5). This conversion from a nine-membered pseudo-ring structure to an eight-membered structure could lower the ring strain of the system[26].



Scheme N. Charge-remote mechanism for the formation of the $^{0,3}A_3$(ii4) ion (m/z 413). Following a [1,3] proton transfer between the proton of the fourth hydroxyl group and the third carbon on the reducing ring, a glyceraldehyde molecule is lost. A larger version of the scheme can be found in annex 1.

The fourth possible cleavage involves the transfer from O(i5) to C(i4) to lose 2,3,4-trihydroxybutanal ($C_4H_8O_4$) or threose and form the $^{0,4}A_3$(ii4) ion (m/z 383) as shown in Scheme O. This cleavage leads to a product that is most comparable in energy to the second cross-ring cleavage.



Scheme O. Charge-remote mechanism for the formation of the $^{0,3}A_3$(ii4) ion (m/z 413). Following a [1,3] proton transfer between the proton of the fifth hydroxyl group and the fourth carbon on the reducing ring, a threose molecule is lost. A larger version of the scheme can be found in Annexe 1.

5.4.3. Moving on from cross-ring cleavage.

Since the cross-ring cleavage mechanism do not directly involve the glycosidic bond, any of the ions formed can undergo a similar mechanism to the glycosidic bond cleavage forming either a B_2 ion (m/z 323) or C_2 ion (m/z 341) as presented in Schemes A and C. The predominance of the m/z 383 ion relative to the other cross-ring fragments could be because the $^{0,4}A_3$(ii4) ion (M2) makes a more favorable precursor for the glycosidic bond

cleavage since its β-hydrogen is attached to an aldehyde group which greatly increases its acidity and reactivity.

5.5. Peripheral fragmentation or water-loss fragmentation

The water-loss pathway for isomaltotriose represented by the ion at m/z485 is a minor pathway only present at low energies. As shown in Figure 21, it is a very short-lived ion before proceeding through the previously mentioned pathways. Ultimately, it will lead to an ion that is isobaric with either of the B or Z type ions.

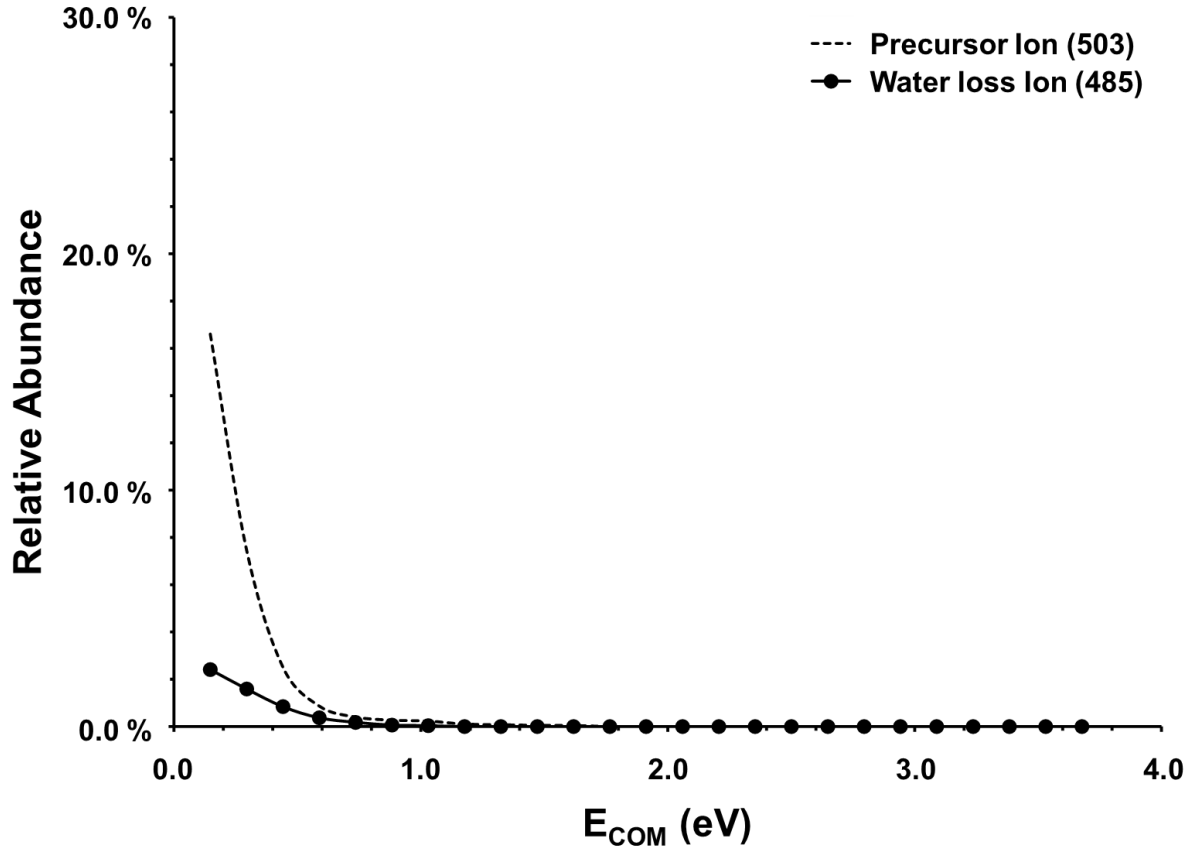
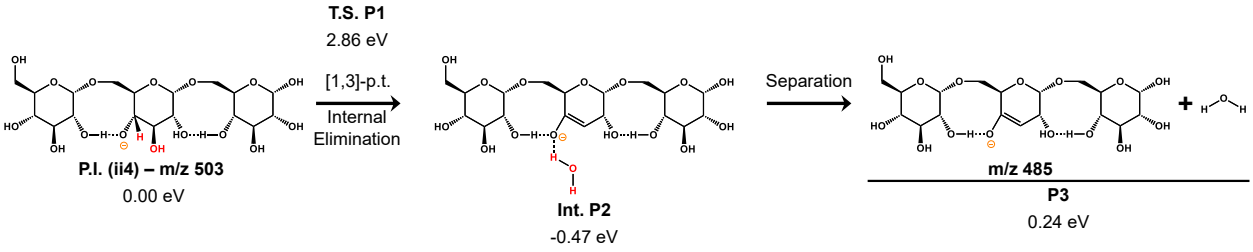


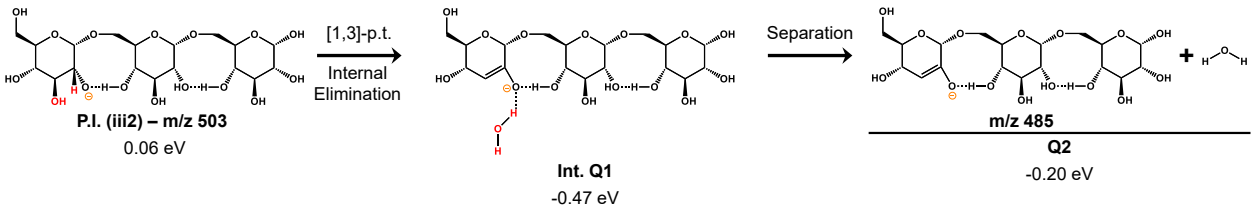
Figure 21. Partial ESI-MS/MS breakdown diagram of water loss ion (m/z 485). As the precursor ion dissociates, water is quickly lost before proceeding with further fragmentation.

The dehydration reaction can occur on any of the hydroxyl groups on the trisaccharide; however, one of them leads to a more favorable final energy as shown in Scheme P. The carbohydrates prefer to lose the hydroxyl group adjacent to the oxygen holding the charge. For the main precursor ion, the hydroxyl group attached to C(ii3) takes the proton bound to C(ii4) containing the alkoxide through a [1,3] proton transfer to form a transitory water bound intermediate that quickly removes the water molecule to form the ion-neutral complex (P1). The enol structure formed allows for the charge to be dispersed and mitigated through the newly formed aromaticity.



Scheme P. Charge-remote mechanism for water-loss from the primary precursor ion. The proton (red) that is in the alpha position to the deprotonated oxygen undergoes a [1,3] proton transfer to the adjacent hydroxyl group (ii3) to form a water molecule. As it is removed an enol is formed allowing for more stability as the charge on the oxygen can be delocalized through the pi bond. A larger version of the schemes can be found in Annexe 1.

A similar reaction was found for the secondary precursor ion as shown in Scheme Q. The hydroxyl group in the iii3 position steals the adjacent alkoxide's alpha proton to form a water molecule, which is rapidly eliminated to form the ion-neutral complex (Q1). The intermediates for both precursor ions are almost identical in energy however, the final product for the secondary precursor ion (Q2) is much lower in internal energy and is exothermic in nature.



Scheme Q. Charge-remote mechanism for water-loss from the secondary precursor ion. The proton (red) that is in the alpha position to the deprotonated oxygen undergoes a [1,3] proton transfer to the adjacent hydroxyl group (iii3) to form a water molecule. As it is removed an enol is formed allowing for more stability as the charge on the oxygen can be delocalized through the pi bond. A larger version of the schemes can be found in annex 1.

5.6. Assessing the competition between the pathways.

As previously mentioned, polysaccharide fragments fall into three different groups: glycosidic bond cleavage (GBC), cross-ring cleavage (CRC) and peripheral cleavage (PC) generally seen as the water-loss fragmentation. The purpose of this chapter was to test the feasibility of charge-remote reactions that could produce these fragments and see how they compare to previously proposed charge-directed mechanisms that are driven by multiple charge migrations along the hydroxyl groups of the glycan.

The activation energies (E_a) for each scheme were calculated as the energy difference between the precursor ion and the highest transition state of the mechanism. Due to the high degree of difficulty encountered in trying to calculate the transition state structures at the adopted level of theory. These values are used as approximations for the similar reactions allowing for a semi-empirical comparison. For example, the activation energy determined for Scheme A involving the cleaving of the C/Z bond between the middle and reducing ring of the primary precursor ion can be used to approximate the energy required to undergo the similar reaction for the secondary precursor ion in Scheme B as well as the comparable [1,3]-p.t. involved in cleaving the C/Z bond between the non-reducing ring

and the middle ring in Schemes E and F. This assumption is based off the observation that the intermediates and ion-neutral complexes involved have similar internal energies. A similar thought process is used for the sequential [1,2]-p.t. pathway used to cleave the B/Y bonds in Schemes D, E and F by using the activation energy calculated from Scheme C. This assumption was also used to estimate the activation energy of the water-loss mechanism for the secondary precursor ion in Scheme Q by adopting the value calculated for the primary precursor ion in Scheme P.

In Rabus et al.'s work on lactose, they calculated a charge-remote ring-opening activation energy of 2.04 eV which is comparable to our calculated value of 1.99 eV for the transition state of Scheme K. They then calculated an activation energy of 1.23 eV for the charge-directed mechanism for the loss of $C_2H_4O_2$. This showed that the overall cross-ring cleavage mechanisms were dependant on the energy of the ring-opening reaction. For that reason, the activation energy of 1.99 eV was rounded to 2.0 eV and used as the assumed activation energy for cross-ring cleavage Schemes K – O. Their calculated activation energy for the charged-directed cleaving of the C/Z glycosidic bond was used as a ball-park figure for the charge-directed GBC Schemes G and J.

The total reaction energies ($\Delta_r H$) for each scheme were calculated as the energy difference between the precursor ion and the summed product energies of the ion and the neutral loss. For these calculations, the CRC schemes used the closed ring precursor ion as the initial reactant instead of the ring-opened product formed in Scheme K. The values for E_a and $\Delta_r H$ for each scheme are shown in Table 2.

Table 2. Comparative reaction energies for the unimolecular fragmentation reactions of isomaltotriose.

Reaction Scheme	Cleavage Type	Precursor Ion	Ion Formed	Approx. E_a (eV)	$\Delta_r H$ (eV)
A	CR - GBC (C_2/Z_1)	Primary	C_2(i4) m/z 341	2.9[a]	0.64
B		Secondary	C_2(ii2) m/z 341		0.45
C	CR - GBC (B_2/Y_1) [C(ii1) – O]	Primary	B_2(i4) m/z 323	2.3[b]	1.25
D		Secondary	B_2(ii2) m/z 323		1.29
E	CR - GBC (C_1/Z_2) [O – C(ii6)]	Primary	Z_2(ii4) m/z 323	2.9[a]	1.19
F		Secondary	C_1(i2) m/z 179		1.39
G	CD - GBC (C_1/Z_2) [O – C(ii6)]	Primary	C_1(i1) m/z 179	1.7[c]	2.44
		Secondary	Z_2-cyclic(ii3) m/z 323		2.26
H	CR - GBC (B_1/Y_2) [C(iii1) – O]	Primary	Y_2(ii4) m/z 341	2.3[b]	1.55
I		Secondary	B_1(i2) m/z 161		1.66
J	CD - GBC (B_1/Y_2) [C(iii1) – O]	Primary	Y_2(ii6) m/z 341	1.7[c]	2.57
		Secondary	B_1-cyclic(i3) m/z 161		2.01
K	CRC – ring-opening	Primary	Ring-opened precursor ion	2.0[d]	-
L	CRC – CH_2O loss	Primary	$^{0,1}A_3$(ii4) m/z 473		0.87
M	CRC – $C_2H_4O_2$ loss	Primary	$^{0,2}A_3$(ii4) m/z 443		0.65
N	CRC – $C_3H_6O_3$ loss	Primary	$^{0,3}A_3$(ii4) m/z 413		0.37
O	CRC – $C_4H_8O_4$ loss	Primary	$^{0,4}A_3$(ii4) m/z 383		0.57
P	Water-loss (ii3)	Primary	M – H_2O	2.9[e]	0.24
Q	Water-loss (iii3)	Secondary	M – H_2O		-0.20

CR: Charge-remote, **CD**: charge-directed, **GBC**: Glycosidic bond cleavage, **CRC**: Cross-ring cleavage.

[a] The activation energy calculated for Scheme A is used to approximate the values for the similar reactions for Schemes B,E and F.

[b] The activation energy calculated for Scheme C is used to approximate the values for the similar reactions for Schemes D, H and I.

[c] The activation energy for Schemes G and J are speculative values based on the charge-directed activation energies observed by Rabus et al.

[d] As observed by Carrol et al. for glucose, the ring-opening step was the rate-determining step for the cross-ring cleavage energies.

[e] The activation energy calculated for Scheme P is used to approximate the value for the similar reaction in Scheme Q.

By observing the activation energies, we get a kinetic point view of how isomaltotriose fragments. To start, the charge-directed mechanism would be more favorable since it has the lowest activation energy at 1.7 eV. However, this would lead to either the Z_2(m/z 341) or the B_1(m/z 179) being the dominant ions which was not observed experimentally. The reason for this could be because it is not enthalpically favorable to form either an oxirane (Z_2) or an oxetane(B_1) cyclic structure as they have much higher product energies of 2.26 and 2.01 eV respectively when compared to their charge-remote equivalents that had a much higher activation energy but lower product energies of 1.19 and 1.66 eV respectively. Looking at the other mechanisms, by order of their activation energies, the next priority would go to the CRC ions which reflects our observations that the cross-ring cleavage has to occur before the GBC happens or they would not be observed. After that, the cleaving of the B/Y bond would occur at a greater rate than the C/Z cleavage. Looking at the enthalpies of these reactions, it would form in majority the B_2 (m/z 323) ion followed by the Y_2 (m/z 341) ion and lastly the B_1 (161 m/z) ion. Although our experimental results do agree with the priority being placed on the direct formation of a dimer over that of a monomer, the 341 ion dominates over the 323 ion. The highest activation energies go to the C/Z bond cleavage which produces the dominant C_2 (m/z 341) ion and the water-loss cleavage. What is most likely happening in the experiment is that the internal energies of the precursor ion coming from the source are much higher than the acceleration energy given to it. As it collides with the argon it has enough energy to access any of the mechanisms. The ions observed are going to be the most enthalpically stable ions. Any excess energy given by further acceleration causes them to undergo further fragmentation to reach a more stable ion.

Going with this hypothesis, the fragment ions with the lowest enthalpies are the CRC ions coming form the precursor ion ranging between 0.37 and 0.87 eV. Afterwards there are the C_2 isomers at 0.45 and 0.64 eV formed by cleaving the C/Z bond between the reducing end and the middle ring. This corroborates our observations as they are the dominant ions at low energy as well as Carrol et al.'s results of the isotopically labeled lactose disaccharide that demonstrated it was the C/Z bond that was cleaved. It should be noted that the water-loss fragment ions (m/z 485) have even lower enthalpies at -0.20 and 0.24 eV. Since these ions quickly disappear it is possible that they quickly undergo the same cleaving of the C/Z bond which would form an isomer of the B_2 ion.

Afterwards at nearly double the enthalpy, the cleaving of the B/Y bond between the reducing ring and the middle ring leads to the formation of the B_2 isomers (m/z 323) at 1.25 and 1.29 eV. This could explain why the 341 ion dominates over the 323 ion.

The cleaving of the C/Z bond between the non-reducing end and the middle ring has similar energies to the last as it forms either the Z_2 ion which has an enthalpy of 1.19 eV and is isobaric to the B_2 ions or it can form the C_1 ion(m/z 179) which has an enthalpy of 1.39 eV. As this energy is higher than that of the C_2 ion, it fits with the idea that the 179 ion could come directly from the precursor ion but the majority of it would be the sequential loss form the C_2 ion.

Finally, cleaving across the B/Y bond between the non-reducing ring and the middle ring can produce the Y_2 ion (m/z 341) which would supplement the intensity of the C_2 ion or it could form the B_1 ion (m/z 161). The have respective enthalpies of 1.55 and 1.66 eV.

5.7. How tandem MS helps in glycan sequencing.

Knowing which bonds are more prominently cleaved can help scientists sequence glycans. As they fragment sequentially starting from the reducing end, the location and type of modified sites can be determined by comparing with the unmodified version. During glycosidic bond cleavage, the glycan will lose a mass of 162 as the C-type ion is formed. A shift in that mass would suggest that one of the hydroxyl groups have been modified and could determine which type of modification by the magnitude of the mass shift. Cross-ring cleavage can give even more information as it has previously been observed by Carrol et al. that the CRC fragment ions formed depend on the type of glycosidic bond cleavage. These are generally observed as mass loss of 30, 60, 90 or 120 as units of CH_2O are lost. A loss of 78 can be observed only in the case of (1,2) and (1,3) linked carbohydrates. Shifts in these masses could help indicate which hydroxyl group was modified.

Chapter 6. Study of the thermodynamics of phosphate adduction

6.1. ESI-MS/MS of phosphate adducted glycans

When studying glycans via ESI-MS, sensitivity and in-source stability is always an issue. It is important that the ions are produced efficiently enough to be seen at low concentrations without increasing the ionization energy that could cause the newly formed ion to be excited and fragment before leaving the source. As mentioned previously, deprotonation is one way to form an anionic species and the efficiency of that ionization depends primarily on the ability of the solvent to remove a proton from the analyte as the microdroplet is ruptured. The hydroxyl groups on carbohydrates have very low acidities and are not easily removed. Another issue with the deprotonated species is that they tend to have low stability leading to fragmentation in the source. Another method of ion formation is achieved by forming a gas-phase quasi-molecular ion complex comprised of the analyte and a small easily ionizable species called the adduct. In negative ionization mode, the adducts are generally halogens (Cl^-, Br^- or I^-) or small bases ($H_2PO_4^-$, HCO_3^-, NO_3^-, SO_3^-, etc…). The complex is held together by various dipole interactions with the addition of hydrogen bonding in the case of the small bases [27]. In a study by D.J. Harvey, it was demonstrated that phosphate and nitrate buffers help produce a high abundance of quasi-molecular ions of glycans [28]. As a part of my honours thesis, the fragmentation of different lengths of phosphate-adducted isomaltooligosaccharides (IMOs) that make up the dextran 1000 calibrant was studied [29]. Phosphate was chosen as the adducting agent as the theoretical collisional cross sections were available.

As shown in the mass spectra of Figure 22, during a CID experiment run on phosphate adducted isomaltotriose (m/z 601), the internalized energy incurred on collision with the argon causes the excited complex to choose one of three pathways. The complex can separate, leaving the charge on the dihydrogen phosphate ion (m/z 97). Any additional energy will cause the dihydrogen phosphate ion to lose water to form a relatively stable PO_3^- ion (m/z 79) [30]. The second option is for the dihydrogen phosphate ion to form the neutral phosphoric acid by abstracting a proton from the analyte. This forms the deprotonated isomaltotriose, which follows the same reaction pathways discussed in Chapter 5. The third pathway is to undergo similar fragmentation of the trisaccharide before losing the phosphate adduct. However, these ions have very low abundancies and are of little relevance. These reactions allowed us to plot the fragmentation roadmap in for phosphate adducted species in Figure 23. This is in relation to the fragmentation of the deprotonated species previously discussed.

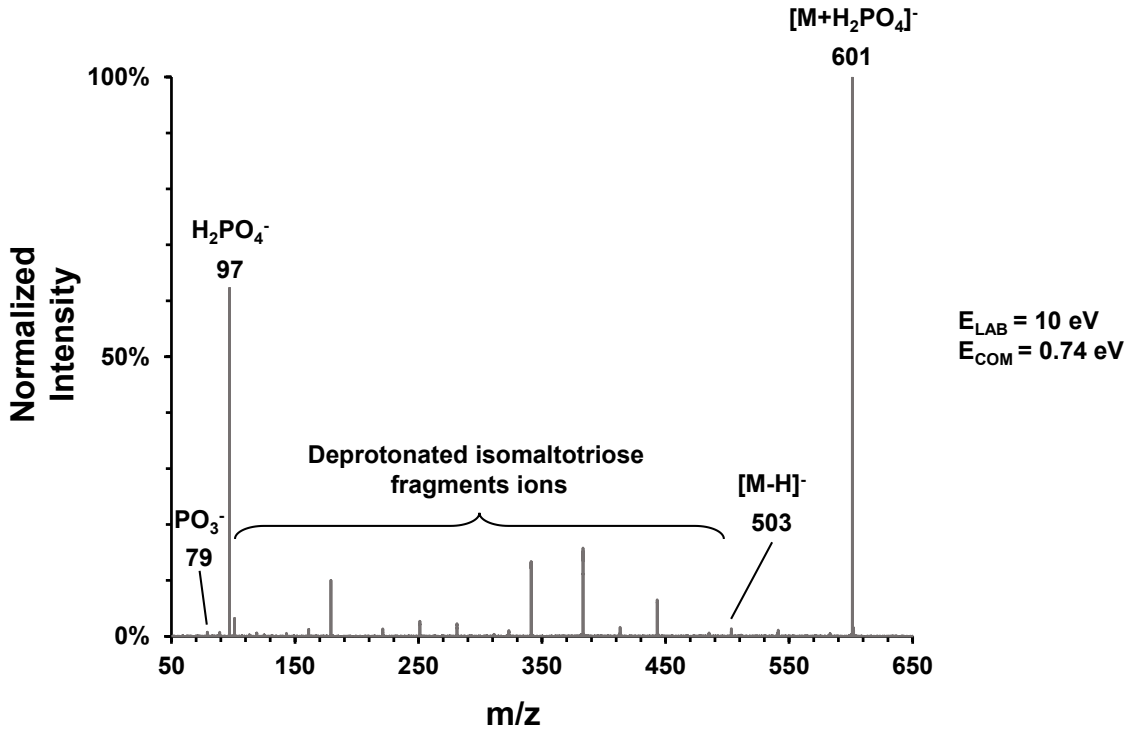


Figure 22. CID Mass spectra of phosphate adducted isomaltotriose (m/z 601). As the adduct-glycan complex is excited it can either dissociate to form a phosphate ion (m/z 97) or the phosphate ion can abstract a proton before separation to form the deprotonated glycan (m/z 503) which then fragments further.

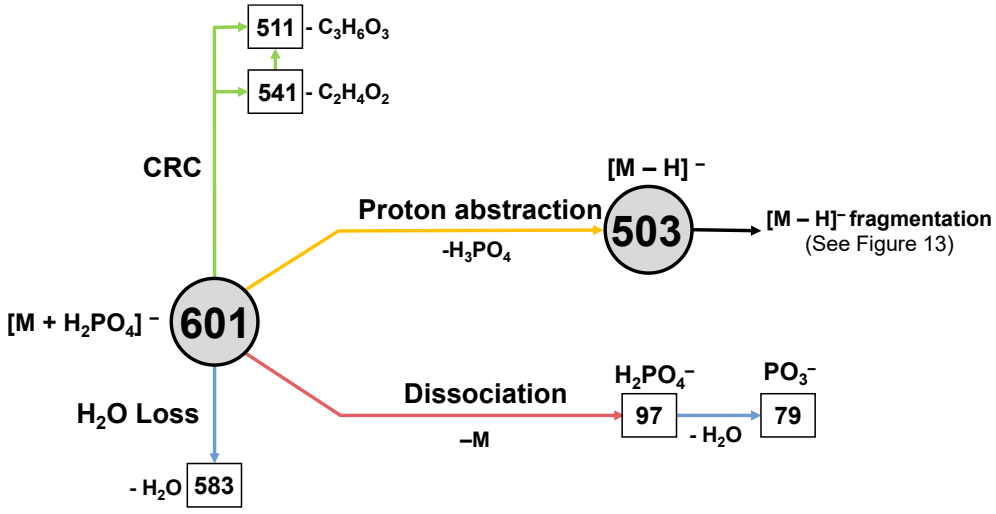


Figure 23. Abbreviated fragmentation roadmap for phosphate adducted isomaltotriose. The phosphate adduct species can undergo complex dissociation to form the phosphate ion or it can abstract a proton before the dissociation occurs to form the deprotonated species and fragment through the previously described roadmap. Although very low in abundance, some fragmentation of the glycan can occur before the complex dissociates.

The breakdown diagram in Figure 24 demonstrates the competition between the complex separation and the proton abstraction processes. At lower energies, the adducted ions adopt either pathway in relatively equal amounts. As the energy is increased, the loss of the phosphate ion begins to out compete the proton abstraction process. This has the drawback of producing low amounts of the deprotonated fragment ions that could give needed structural information. Phosphate is best used when interested in the accurate mass or ion mobility of the precursor ion. In order to understand why complex separation becomes more favorable at high energies, the reaction energies of the precursor ion and the product ions for both processes were studied.

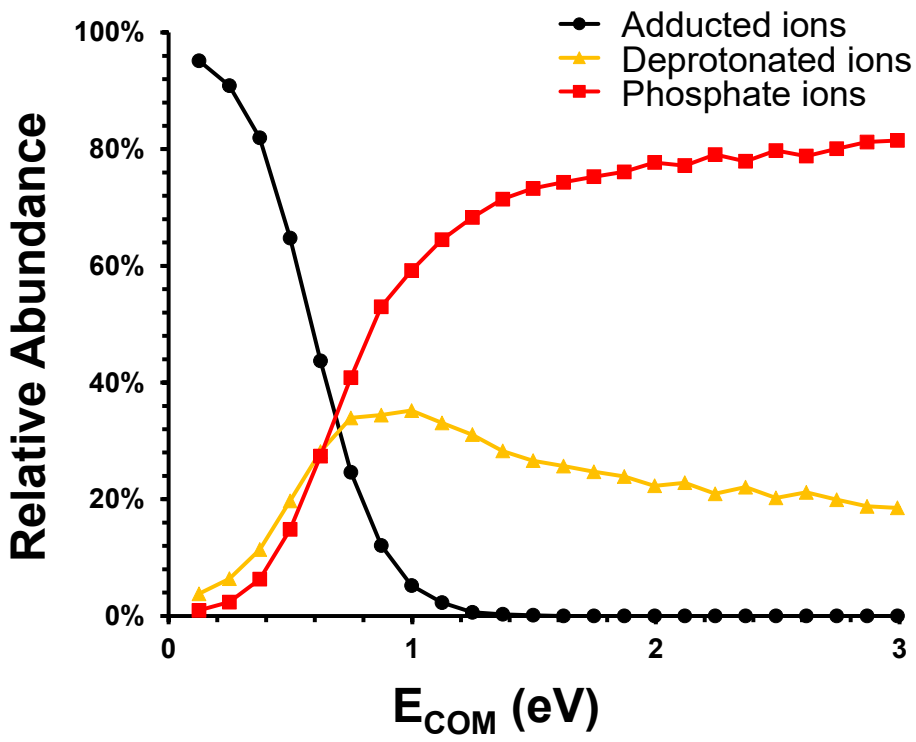


Figure 24. Breakdown diagram of phosphate adducted isomaltotriose. As the phosphate adducted species (black) dissociates, it can dissociate, and the phosphate ions are observed (red). If the phosphate ion can abstract a proton before dissociation, the deprotonated glycan ions are observed (yellow).

6.2. Determining the structure of phosphate adducted isomaltotriose.

In order to probe the possible structures for phosphate adduction, the optimised neutral structure for isomaltotriose was used as a starting point for the neutral analyte component of the gas-phase complex. Afterwards, in a position near two adjacent hydroxyl groups on the trisaccharide structure the dihydrogen phosphate ion was place in four different orientations as shown in Figure 25. The first orientation allows for the ketone and alkoxide groups to act as hydrogen acceptors from the hydroxyl groups. The second has the two hydroxyl groups attached to the phosphorus acting as hydrogen donors to the oxygens of the two hydroxyl groups of the glycan. The last two has the ion acting as a hydrogen donor and acceptor with the roles switch between the two orientations.

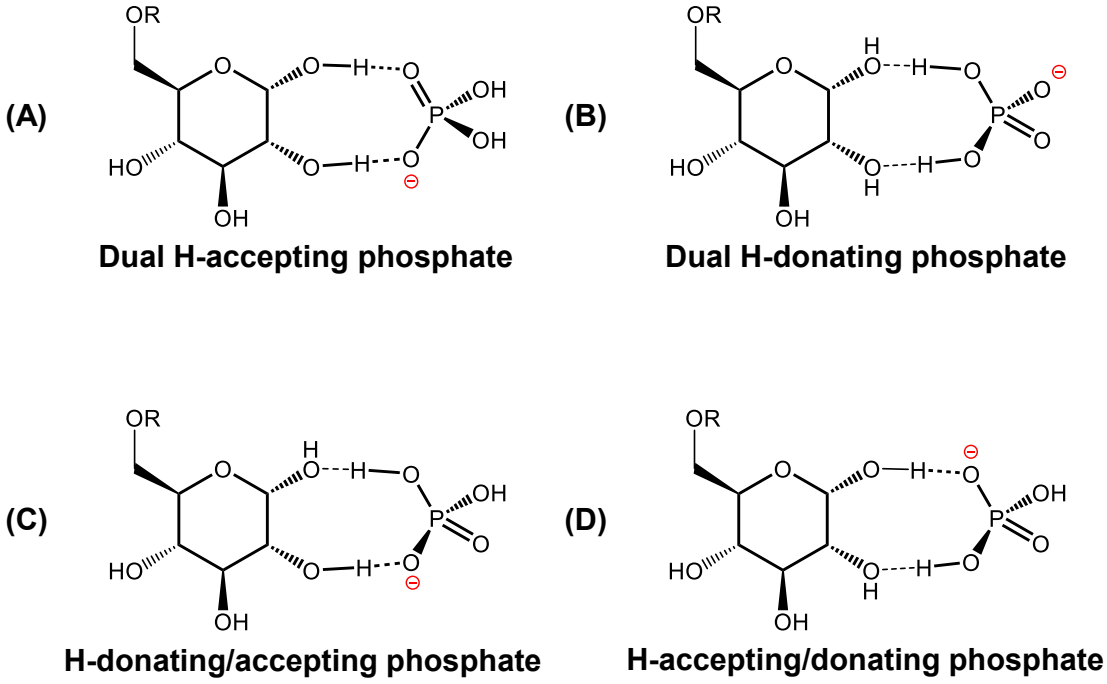


Figure 25. Four possible binding conformations used to probe the possible structures of the phosphate adducted species. The phosphate ion was position between two adjacent hydroxyl groups so that it can either accept two protons (A), donate two protons (B) or a mixture of the two (C and D).

After probing the different sites and orientations, any duplicate geometries were set aside leading to a total of thirty-three distinct structures. Due to the high number of possible structures, the relative density of states was calculated and plotted for each of them as shown in Figure 26. It demonstrated that only four species occupied the majority of the states. Their structures are shown in Figure 27. The most prominent structure (PA1) has the three rings wrapped around the phosphate ion which is accepting three hydrogens and donating one back to the glycan. As the energy is increased, the phosphate tends to migrate out from between the three rings towards the reducing end.

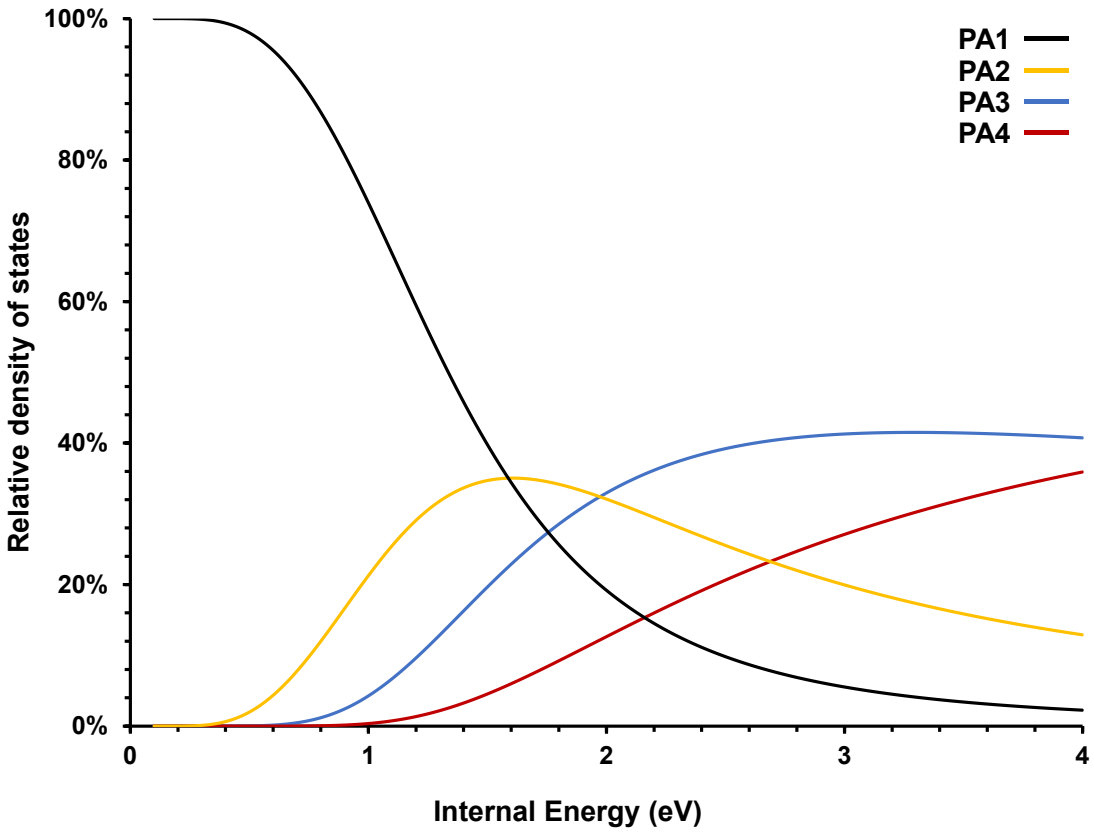


Figure 26. Relative density of states for the isomers of phosphate adducted isomaltotriose. The geometry for the four isomers that occupy the majority of the states are presented in Figure 27. The 29 other isomers never achieve more than 5% of the total states.

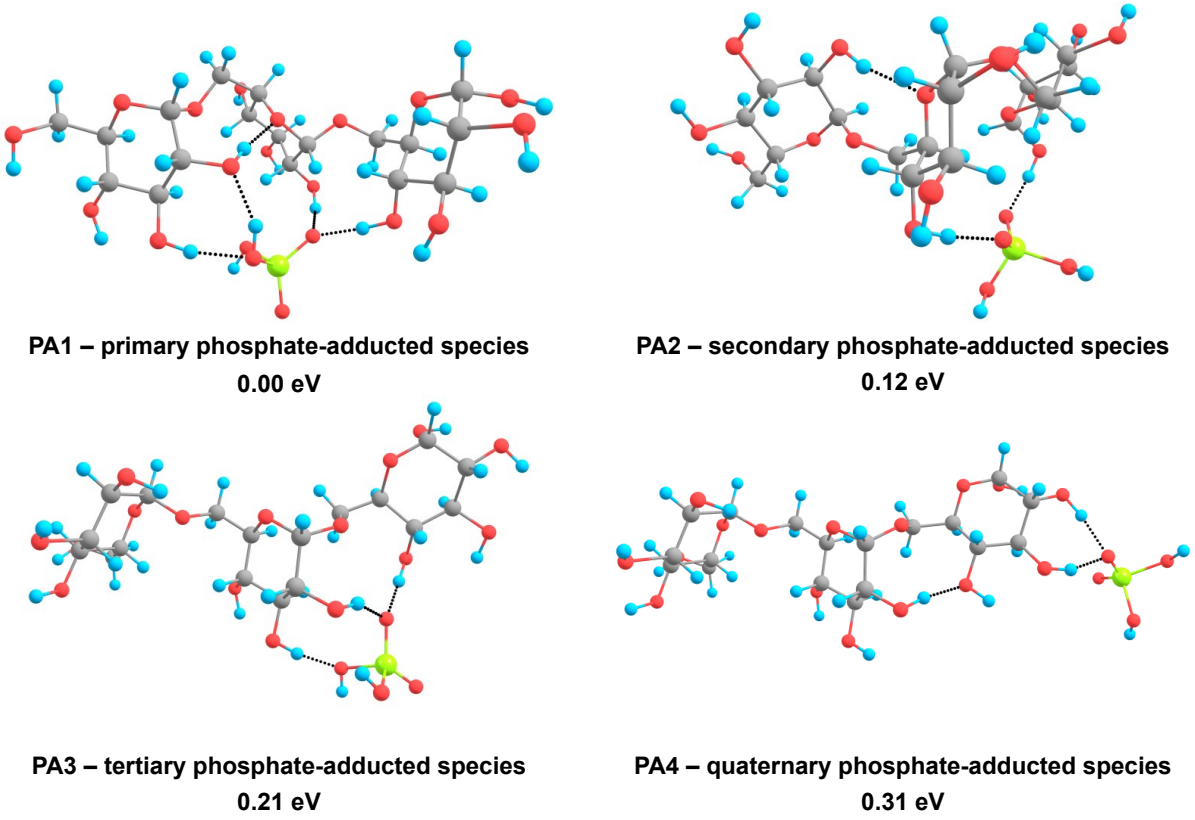


Figure 27. Optimised DFT structures of the four prominent isomers of phosphate adducted isomaltotriose. The isomers are numbered according to their importance. The internal energies are relative to the lowest energy isomer (PA1).

6.3. Probing the competition between abstraction and separation

Once the precursor ion structures were determined, the product energies for both pathways were calculated. As shown in Figure 28, there is only a 0.2 eV difference between the Gibbs reaction energy for both pathways, with the proton abstraction reaction being slightly more favorable. Since these energies are so close, the resulting ion depends on the competition between the rate of separation and the rate of proton abstraction before the complex dissociates. The dissociation rate depends on the binding energy between the phosphate and the analyte. As the proton that is transferred is most likely part of one of the hydrogen bridges, the positioning of the charge along that bridge should not affect the other intermolecular interactions meaning the binding energies

should remain relatively the same regardless of with component bears the charge. This means that the resulting ion is primarily dependant on the acidity of the abstracted proton.

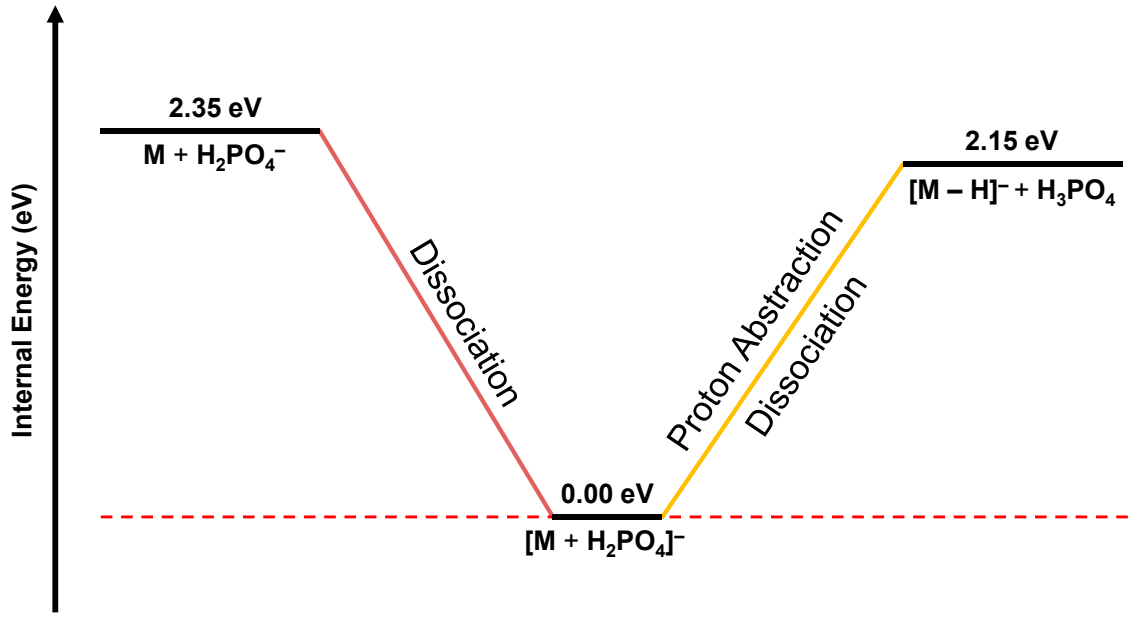


Figure 28. Energy diagram for the complex dissociation and the proton abstraction pathways. The presented energies are relative to the internal energy of the primary phosphate adducted precursor ion (PA1). There is only small difference in energy between the dissociation products and the proton abstraction products.

As seen in Chapter 5 and in Houdier and Perez's work on sucrose, the most acidic protons in a trisaccharide are those bound to the oxygens that form the intercyclic hydrogen bridges [31]. Looking back at the structures in Figure 27, the primary adducted precursor ion (PA1) is accepting the more acidic protons from the two hydroxyl groups between the reducing and middle ring. At this position, the abstraction rate would be at its greatest and to a degree comparable to the rate of complex dissociation as seen by the almost 1:1 nature of the breakdown diagram trends for both reactions at low energies. Afterwards, the phosphate starts to migrate towards the reducing end where it no longer has access to the more acidic protons causing the abstraction rate to drop and the abundance of the deprotonated species goes with it.

6.4. Observations with longer isomaltooligosaccharides (IMOs).

The IMOs present in the dextran 1000 ion mobility calibrant allowed for the study of phosphate adducting on species containing three to seven glucose residues. The CID experiments were completed during my honours research project and their breakdown curves presented in Figure 29. First, looking at the adducted species, the added chain length increases the stability of the complex. This was observed for the deprotonated species as well. This is most likely due to the added rotational and vibrational degrees of freedom available through the general flexibility of the glycosidic bonds. As the oligosaccharide increases in size, it takes more internal energy to start cleaving bonds as more energy can be dissipated through rotational and vibrational modes. For the phosphate adducted species, the energy is transferred along the structure before it can overcome the binding energy of the complex.

Now looking at the product ions formed, there seems to be a greater discrepancy between the rate of abstraction and the rate of complex dissociation. Assuming that it is still the intercyclic hydroxyl groups that are favored for deprotonation, the acidity should not be greatly affected by the chain length just the number of possible abstraction points. As the number of residues increases, the statistical probability of the phosphate abstracting a proton before it is released rises. If the larger glycans behave similarly to the trisaccharide, the phosphate would migrate out from a favored abstraction position to the reducing end of the sugar.

For the heptameric IMO (E), the onset energies seem to become identical again similar to the trisaccharide. This could possibly be a sign of a shift in either the binding energy or the accessibility of the intercyclic hydroxyl groups although which is the primary cause is

difficult to determine as both of which could be explained if the heptamer starts to adopt a helical structure as proposed by Fittolani et al. [32]. They proposed that polysaccharides can form helical structures at a size of seven residues per turn. The phosphate could sit in the middle of the ring giving heightened access to the necessary abstraction groups however the binding energy would increase from the added interactions. If the phosphate ion sits on the outside of the ring, it would have less access to the hydroxyl groups and would have a lowered binding energy. Further studies would be needed to determine where the phosphate ion would sit. Due to the size of the molecule, DFT calculations would be too demanding. For that reason, a semi-empirical method which takes approximations from empirical data to simplify the calculations should be used. The parametric method 3 (PM3) elaborated by J.J.P. Stewart would work for the atoms in question [33]. Another possibility is to use molecular mechanics/ molecular dynamics (MM/MD) which treats the molecular bonds as springs and using force constants to approximate the interatomic interactions [34]. Both methods shorten the calculation time (and cost) but are less precise due to the approximations used. They are generally reserved for large biomolecular interactions such as protein-protein or enzyme-substrate binding where the precise behavior of individual atoms is less relevant.

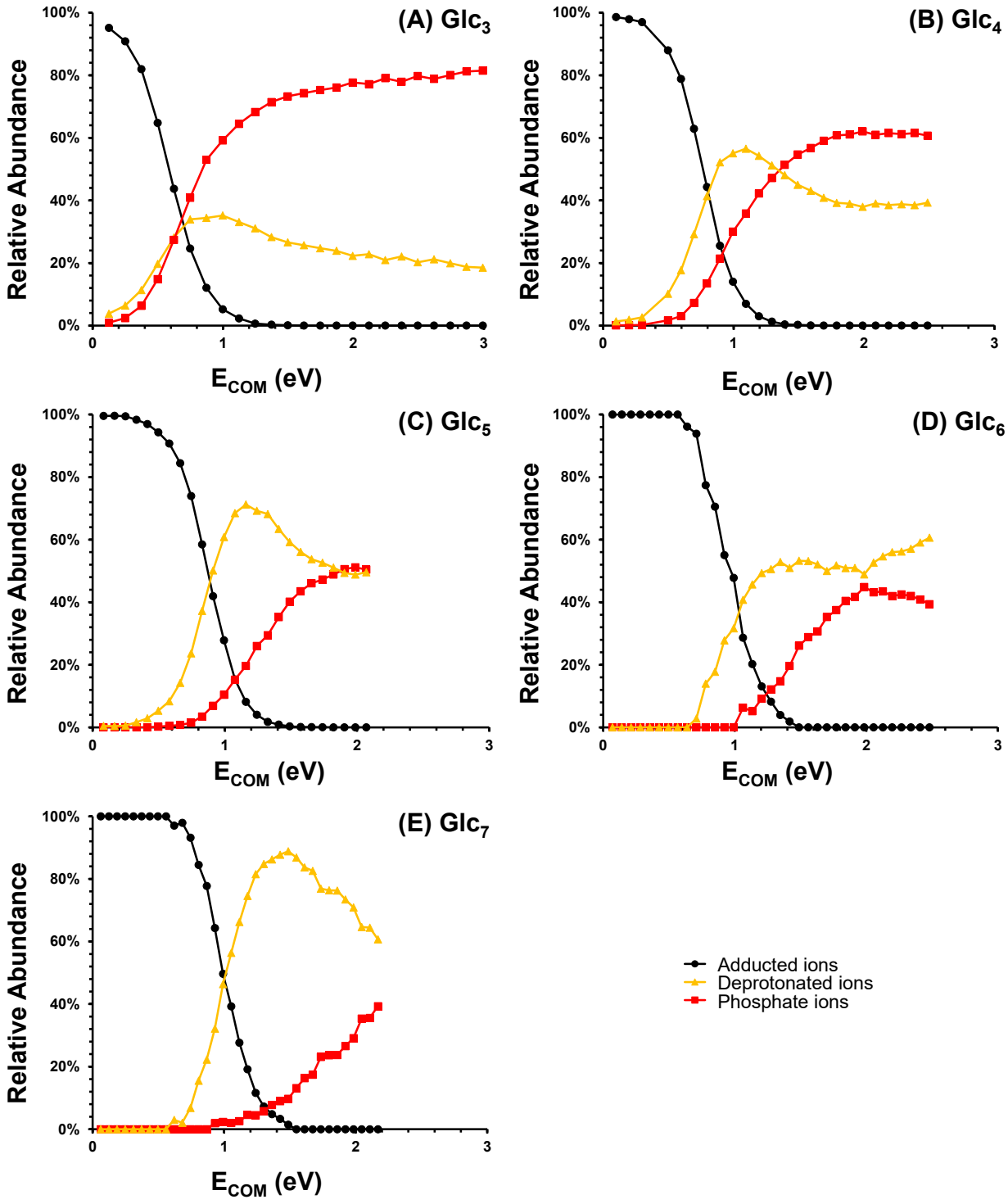


Figure 29. Breakdown diagrams for the dissociation and proton abstraction pathways of isomaltooligosaccharides. The curves for trimer (A), the tetramer (B), the pentamer (C), the hexamer (D) and the heptamer (E) demonstrate the effect of the carbohydrate chain length on the competition between dissociation and the proton abstraction pathways.

Chapter 7. Ion mobility characterization of the trisaccharide isomers.

As previously mentioned in Chapter 2.4, ion mobility separation (IMS) is another orthogonal method for characterizing glycans. These experiments measure the relative deceleration of the ions as they undergo a multitude of low energy collisions with the drift gas. The travel time of each ion is reported as the drift time in the millisecond time scale which is then used to calculate either the absolute or relative collisional cross sections (CCS) depending on the type of mobility cell used. This method has already shown that it can effectively differentiate between isobaric glycan ions in several works. Jin et al. demonstrated its usefulness in characterizing 77 different O-linked glycans released from the porcine gastric mucin glycoprotein and the glycoproteins present in human saliva samples which include branched and linear isomers as well as sequence isomers [35]. Since its increase in popularity, CCS databases for different biopolymers can now be found like the one produced by Zheng et al. [36]. Rashid et al. studied an array of deprotonated glucose polymers and they observed that the primary component shifts in CCS values [37]. These analyses can also be done using positive electrospray ionization as demonstrated by Przybylski and Bonnet who studied the mobility of different trisaccharides as they formed different quasi-molecular ions either through protonation or adduction of sodium, ammonium, etc. They used this information to characterize the short glycans present in honey samples. An interesting observation of theirs was that the protonation had very little effect on the conformation of the species but when larger adducts were added, the glycans would adopt measurably different conformations in order to stabilize the structure. This signified that although deprotonation or protonation

are better for fragmentation studies, the adducted species provide better separation during ion mobility experiments [38].

As a part of her fourth-year B.Sc. honours project in the Department of Chemistry and Biomolecular Sciences, Petra Chea measured the drift times and calculated the relative CCS values for four different isobaric trisaccharide ions by comparing them to a dextran calibration curve, their structures are shown in Figure 30. Isomaltotriose(ISO) which is found in dextran, maltotriose(MAL) and panose(PAN) are composed solely of glucose residues that are linked together with different glycosidic bonds conformations. ISO has two α1,6 glycosidic bonds and MAL has two α1,4 glycosidic bonds while PAN contains a mix of the two types. These three species were chosen for this study as they can demonstrate the effect that the glycosidic bond orientation has on both their fragmentation (qualitatively studied in Petra's work) and their relative collisional cross sections. Raffinose, which was readily available allowed for a glimpse at the effect of changing the carbohydrate moieties as it is composed of a galactose residue connected through an α1,6 bond to a central glucose residue which is then connected by an α1,2 bond to a fructose residue. In order to corroborate her observations, the optimised DFT structures were calculated for each deprotonation site and can be found in Annexes A2.2 to A2.4. Originally our thought was that isomaltotriose, having two (1,6)-linked glycosidic bonds with an added methylene group in them, would have the largest CCS when compared to maltotriose which would be more compact having two (1,4)-linked glycosidic bonds. Panose having a mix of the two types of glycosidic bonds would sit in between. Raffinose, having a (1,6)-linked glycosidic bond and a (1,2)-linked glycosidic bond, would behave similar to panose as the 1,2 and 1,4 linkages are similar in geometry. Raffinose could

have a lower CCS due to the presence of a five-membered furanose ring instead of a six-membered pyranose ring.

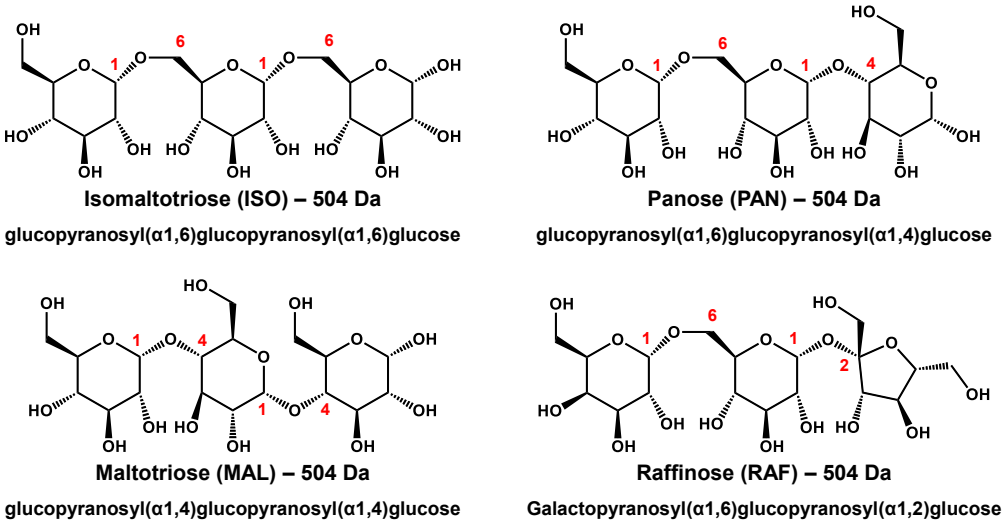


Figure 30. Chemical structures of the four trisaccharides studied by ion mobility separation. Isomaltotriose, panose and maltotriose are glucose trisaccharides containing different glycosidic linkages. Raffinose contains comparable glycosidic linkages to panose but contains a non-reducing galactose residue and a reducing fructose residue.

The intensities of the relative drift time distributions for the different trisaccharide species varied depending on the exact concentration and ionization efficiency of the glycan. To simplify the interpretation of Figure 31, the intensities of each were normalised to their highest value. The data in that figure came from two experimental runs on different days marked as A and B. Although the ion mobility parameters were set the same, they gave slightly different drift times. As isomaltotriose was part of the dextran calibrant, the drift times were then centered relative to mean drift time of isomaltotriose in their respective runs. The calibration curve derived CCS values are shown in Table 3.

The CCS values for the DFT optimised structures were calculated using MOBCAL. Instead of only calculating the MOBCAL values for each lowest energy structure, the

values were calculated for each structure derived from a different deprotonation site and then averaged. The individual CCS values can be found in Annexe 4 and the averaged values are presented in Table 3 with their standard deviation. This is done with a different assumption than that used in the MS/MS study. Instead of the lowest energy structure colliding in a single high energy event and then fragmenting over a short time scale, the molecules are encountering many low-energy collision events causing the molecules to gradually increase in thermal energy over a longer time scale. This means there are less limitations on what deprotonated conformation is passing through the mobility cell at a given time.

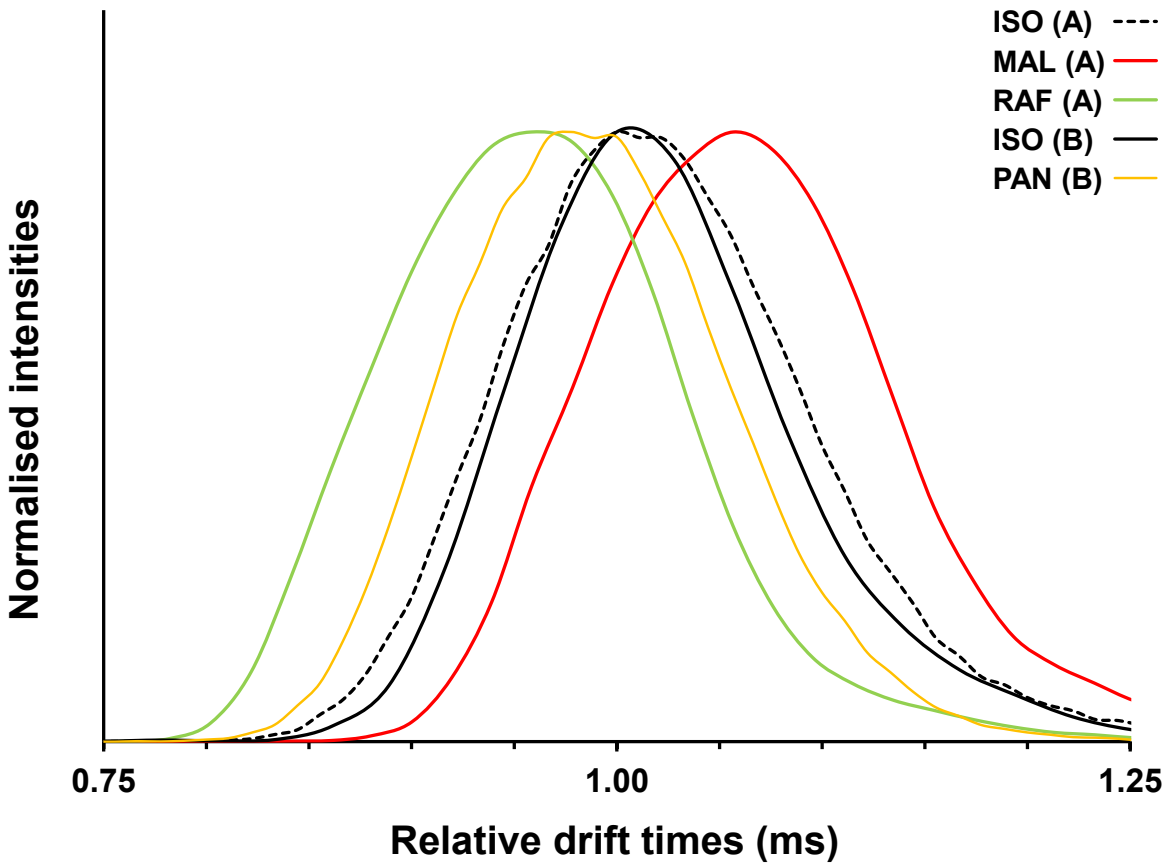


Figure 31. Relative drift time distributions for the studied trisaccharides. The samples were run on different dates marked as A and B. The drift times are normalized to the drift times of isomaltotriose (black continuous and dotted lines) for their respective

runs. The intensities are normalized to their respective maxima in order to correct for sample concentration or ionization efficiency. Panose and raffinose have shorter drift times and maltotriose has a longer drift time than the isomaltotriose standard present in the dextran calibrant.

Table 3. Experimental, calculated, and literature CCS values for the protonated and deprotonated trisaccharides under helium and nitrogen.

Glycan	Experimental $[M-H]^-$ CCS in helium[a] ($\pm$ 1 Å^2)	Averaged MOBCAL $[M-H]^-$ CCS in helium[b] ($\pm$ 3 Å^2)	Literature $[M-H]^-$ CCS in helium (Å^2)	Literature $[M-H]^-$ CCS in nitrogen (Å^2)	Literature $[M+H]^+$ CCS in nitrogen[f] (Å^2)
ISO	(136.62) [c]	152	136.62[c]	202.9 ± 0.5[e]	205.0 ± 4.8
PAN	134	151	139[d]	-	204.6 ± 4.4
MAL	144	148	146[d]	213 ± 1[e]	205.9 ± 3.7
RAF	131	153	-	197.6 ± 0.2[e]	205.3 ± 2.7

[a] The experimental values were calculated by Petra Chea using dextran (including isomaltotriose) as a the calibrant.
[b] The calculated CCS values from MOBCAL were averaged over the available deprotonation sites. The individual optimised structures are presented in Annexe 2. The individual CCS values are presented in Annexe 4.
[c] CCS values of deprotonated isomaltotriose from the dextran calibrant in both helium and nitrogen were experimentally determined by DTIM as published by Hofmann et al. [10].
[d] CCS values of deprotonated panose and maltotriose in helium were determined by Rashid et al. using α,ω-Dicarboxy-terminated polystyrene as a calibrant [37].
[e] CCS values of deprotonated isomaltotriose, maltotriose and raffinose in nitrogen were experimentally determined by DTIM and published by Zheng et al. [36].
[f] CCS values of all four protonated trisaccharides in nitrogen as published by Przybylski and Bonnet[38].

The experimental CCS values do not agree with our original hypothesis as maltotriose has a noticeably increased CCS compared to isomaltotriose. However, the MOBCAL calculated results seem to agree with that hypothesis meaning there is a disconnection between the experimental processes and those simulated using the trajectory method. It is possible that two different phenomena are competing: The effective geometry and the collisional efficiency. One of the MS/MS observations made by Petra Chea in regards to these trisaccharides was that the presence of one or more (1,4) glycosidic bonds when

compared to isomaltotriose significantly lowered the stability of the precursor ion when it undergoes fragmentation. We proposed that the (1,4) glycosidic bond lowered the flexibility of the molecule which meant less energy could be dissipated through rotational and vibrational modes before the excess energy caused fragmentation. This same process could explain what is observed under these lower energy conditions. Maltotriose is less flexible meaning that as it collides with the helium atoms, less energy is converted into the internal potential energy, and more is converted into kinetic energy causing it to drift more. This behaves more like the elastic collisions assumed in the ion mobility process. Although it is geometrically smaller and more compact, it is easier to slow down. Isomaltotriose is much more flexible meaning more energy can be converted into potential energy instead of kinetic energy leading to a loss in collisional efficiency and less drift. Isomaltotriose would undergo a more inelastic-like collision process. This discrepancy in CCS values were observed experimentally in both helium and nitrogen for the deprotonated and protonated species.

Panose would have a flexibility and a geometry somewhere between that of isomaltotriose and maltotriose. If collision efficiency was the dominant factor, it would have an experimental CCS value between that of isomaltotriose and maltotriose which is not the case. It could be that the addition of more than one inflexible (1,4) linkage increases the elasticity of the collisions almost exponentially. A single (1,4) linkage having less effect on the drift time causes the geometry factor to be more important in lowering the measured value. Raffinose behaves like panose according to its linkages, but its smaller experimental CCS value would suggest a smaller geometry from the presence of a five membered ring instead of the six membered ring. Zheng et al. also observed this

drop in CCS although slightly diminished. The MOBCAL values showed an increase in the CCS of raffinose although they should be taken with a grain of salt as demonstrated by the high standard errors and the observed over-evaluation of all MOBCAL values. It is still uncertain what component of its code causes MOBCAL to inflated CCS values by approximately 10%. Other software could be used in the future such as Larriba's ion mobility spectrometry suite (IMoS) which Przybylski and Bonnet used and got comparable results to their experimental data [38, 39].

However, the observation that not only geometry but also flexibility are both factors that influence the measured experimental values does mean that unless factored in, the computationally derived values should not be the determining factor when using CCS values to identify compounds. It is preferred to use an experimentally derived CCS databases, specifically those obtained using a DTIM instrument for carbohydrate sequencing. Ion mobility can be used to distinguish isobaric species in a mixture with a greater emphasis on differentiating the types of glycosidic bonds present as well as the difference between pyranoses and furanoses. It still is uncertain about the ability to distinguish similar carbohydrates such as galactose and glucose.

Chapter 8. Conclusion and Future work

The art of sequencing glycans is an important part in understanding their functionality either on their own or as a part of a larger biopolymer. As tandem mass spectrometry and ion mobility separation are the most commonly used tools today, their uses and limitations were probed as part of this study.

Sequencing deprotonated polysaccharides by tandem mass spectrometry relies on the understanding of their fragmentation roadmaps. These were generally considered to be charge-directed unimolecular mechanisms reliant on the migration of the charge along the different hydroxyl groups which works within a long time-scale experiment. This study looked at the possibility of charge-remote mechanisms that could produce the observed ions that would occur in the more energetic short time scales prevalent during an ESI-MS/MS experiment. It is important to note that at under these conditions, both types of mechanism would be possible, and the ions would result from their competition. We proposed that glycosidic bond cleavage could occur using a [1,3]-p.t. mechanism if the β-proton could adopt an eclipsed conformation through rotation of one or more rings. This explained the prevalence of the C_2 ion at m/z 341 over that of the B_2 ion at m/z 323. A future experiment could involve looking at the fragmentation of a glucose trisaccharide containing β(1,6) glycosidic linkages. It would have the β-proton for the B/Y bond cleavage in the eclipsed position, which is not possible for with the α-glycan. It might allow for more competition between the B_2 and C_2 ions. The cross-ring cleavage mechanisms can occur after a ring opening step through a similar [1,3]-p.t. between the non-leaving hydroxyl group adjacent to the cleaving bond and the adjacent leaving carbon. In the

future, the presence of a modified hydroxyl group such as the aminated glucosamine could be studied. The modification could either behave in a similar fashion allowing the location of the modified site to be probed the shift in either the neutral or ion mass. As the hydroxyl group is the proton donor, depending on the modification it could hinder the cross-ring cleavage entirely.

Ion mobility separation is the second orthogonal method studied. It has shown promise when trying to differentiate between glycosidic bond linkages and between a furanose and a pyranose. However, it is still uncertain about the possibility of distinguishing between the difference in orientation of the individual hydroxyl groups. Perhaps with increased capabilities of the later generation instruments this will be improved. As protonation and deprotonation sites have little effect on the overall structure, it is best to use a large adducting species such as phosphate or nitrate which allow for more drastic and measurable conformation changes with the added bonus of increased precursor ion intensity. It does come with the drawback of hindering the MS/MS fragmentation as there is distinct competition between the proton abstraction reactions needed to kick-start the fragmentation and the loss ion-neutral complex separation. This leads to very low fragment ion intensities. To maximise results, the MS/MS and IMS experiments should be done with two different solutions (with only one containing an adducting agent). If sample buffers must be used, it could be possible to add an inline desalting column that would be used only when running the tandem MS experiment.

When it comes to identifying glycans using their CCS, it is recommended to compare the experimental results with those found in databases especially those populated with DTIM data. The theoretical values calculated in MOBCAL do not always consider the possibility

of inelastic collisions and their effect on drift times as is the case when comparing different glycosidic linkages. Other software suites do exist that have shown some promise.

References

1. Bergstrom, K.S.B., Xia, L.: Mucin-type O-glycans and their roles in intestinal homeostasis, /pmc/articles/PMC3858029/, (2013)

2. Gudelj, I., Lauc, G., Pezer, M.: Immunoglobulin G glycosylation in aging and diseases. Cellular Immunology. 333, 65–79 (2018). https://doi.org/10.1016/j.cellimm.2018.07.009

3. Zaia, J.: Mass spectrometry of oligosaccharides. Mass Spectrometry Reviews. 23, 161–227 (2004). https://doi.org/10.1002/mas.10073

4. Ashcroft, A.E.: Ionization Methods in Organic Mass Spectrometry. (1997). https://doi.org/10.1039/9781847551191

5. Zhu, J., Cole, R.B.: Formation and decompositions of chloride adduct ions, [M + Cl] − , in negative ion electrospray ionization mass spectrometry. Journal of the American Society for Mass Spectrometry 2000 11:11. 11, 932–941 (2000). https://doi.org/10.1016/S1044-0305(00)00164-1

6. Miller, P.E., Denton, M.B.: The quadrupole mass filter: Basic operating concepts. Journal of Chemical Education. 63, 617–622 (1986). https://doi.org/10.1021/ED063P617

7. Solano, E.A., Mohamed, S., Mayer, P.M.: Modeling collision energy transfer in APCI/CID mass spectra of PAHs using thermal-like post-collision internal energy distributions. The Journal of Chemical Physics. 145, 164311 (2016). https://doi.org/10.1063/1.4966186

8. Wyttenbach, T., Bowers, M.T.: Gas-Phase Conformations: The Ion Mobility/Ion Chromatography Method. Top Curr Chem. 225, 207–232 (2003). https://doi.org/10.1007/3-540-36113-8_6

9. Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Structural information from ion mobility measurements: Effects of the long-range potential. Journal of Physical Chemistry. 100, (1996). https://doi.org/10.1021/jp961623v

10. Hofmann, J., Struwe, W.B., Scarff, C.A., Scrivens, J.H., Harvey, D.J., Pagel, K.: Estimating Collision Cross Sections of Negatively Charged N-Glycans using Traveling Wave Ion Mobility-Mass Spectrometry. Analytical Chemistry. 86, 10789–10795 (2014). https://doi.org/10.1021/AC5028353

11. Parr, R.G.: Density Functional Theory of Atoms and Molecules. Horizons of Quantum Chemistry. 5–15 (1980). https://doi.org/10.1007/978-94-009-9027-2_2

12. Becke, A.D.: Density-functional thermochemistry. III. The role of exact exchange. The Journal of Chemical Physics. 98, 5648 (1998). https://doi.org/10.1063/1.464913

13. C, L., W, Y., RG, P.: Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical review. B, Condensed matter. 37, 785–789 (1988). https://doi.org/10.1103/PHYSREVB.37.785

14. Ditchfield, R., Hehre, W.J., Pople, J.A.: Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. The Journal of Chemical Physics. 54, 724 (2003). https://doi.org/10.1063/1.1674902

15. Peng, C., Ayala, P.Y., Schlegel, H.B., Frisch, M.J.: Using Redundant Internal Coordinates to Optimize Equilibrium Geometries and Transition States. https://doi.org/10.1002/(SICI)1096-987X(19960115)17:1

16. NIST Computational Chemistry Comparison and Benchmark Database, http://cccbdb.nist.gov/

17. Baer, T., Mayer, P.M.: Statistical Rice-Ramsperger-Kassel-Marcus quasiequilibrium theory calculations in mass spectrometry. Journal of the American Society for Mass Spectrometry 1997 8:2. 8, 103–115 (1997). https://doi.org/10.1016/S1044-0305(96)00212-7

18. Beyer, T., Swinehart, D.F.: Algorithm 448: Number of multiply-restricted partitions. Communications of the ACM. 16, 379 (1973). https://doi.org/10.1145/362248.362275

19. Baer, T., Hase, W.L.: Unimolecular Reaction Dynamics Theory and Experiments. (1996)

20. Jarrold M.F.: https://nano.lab.indiana.edu/software/

21. 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 Jr., J.A., 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)

22. Domon, B., Costello, C.E.: A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate Journal. 5, 397–409 (1988). https://doi.org/10.1007/BF01049915

23. Carroll, J.A., Willard, D., Lebrilla, C.B.: Energetics of cross-ring cleavages and their relevance to the linkage determination of oligosaccharides. Analytica Chimica Acta. 307, 431–447 (1995). https://doi.org/10.1016/0003-2670(94)00514-M

24. Mulroney, B., Peel, J.B., Traeger, J.C.: Theoretical study of deprotonated glucopyranosyl disaccharide fragmentation. Journal of Mass Spectrometry. 34, 856–871 (1999). https://doi.org/10.1002/(SICI)1096-9888(199908)34:8<856::AID-JMS841>3.0.CO;2-8

25. 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

26. Engler, E.M., Andose, J.D., von Schleyer, P.R.: Critical Evaluation of Molecular Mechanics. Journal of the American Chemical Society. 95, 8005–8025 (1973). https://doi.org/10.1021/ja00805a012

27. Zhu, J., Cole, R.B.: Formation and decompositions of chloride adduct ions, [M + Cl]−, in negative ion electrospray ionization mass spectrometry. Journal of the American Society for Mass Spectrometry. 11, 932–941 (2000). https://doi.org/10.1016/S1044-0305(00)00164-1

28. 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. Journal of the American Society for Mass Spectrometry. 16, 622–630 (2005). https://doi.org/10.1016/J.JASMS.2005.01.004

29. Sean M. Overton: Surf's Up: Using travelling-wave ion mobility tandem mass spectrometry to characterize glycans, (2020)

30. Morris, R.A., Viggiano, A.A.: Chemistry of PO−, PO2−, and PO3− in the gas phase. The Journal of Chemical Physics. 109, 4126 (1998). https://doi.org/10.1063/1.477015

31. Houdier, S., Pérez, S.: Assessing Sucrose Hydroxyl Acidities Through Semiempirical Calculations. http://dx.doi.org.proxy.bib.uottawa.ca/10.1080/07328309508005399. 14, 1117–1132 (2006). https://doi.org/10.1080/07328309508005399

32. Fittolani, G., Seeberger, P.H., Delbianco, M.: Helical polysaccharides. Peptide Science. 112, e24124 (2020). https://doi.org/10.1002/PEP2.24124

33. Stewart, J.J.P.: Optimization of parameters for semiempirical methods I. Method. Journal of Computational Chemistry. 10, 209–220 (1989). https://doi.org/10.1002/JCC.540100208

34. Warshel, A., Levitt, M.: Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. Journal of Molecular Biology. 103, 227–249 (1976). https://doi.org/10.1016/0022-2836(76)90311-9

35. Jin, C., Harvey, D.J., Struwe, W.B., Karlsson, N.G.: Separation of Isomeric O-Glycans by Ion Mobility and Liquid Chromatography–Mass Spectrometry. Analytical Chemistry. 91, 10604–10613 (2019). https://doi.org/10.1021/ACS.ANALCHEM.9B01772

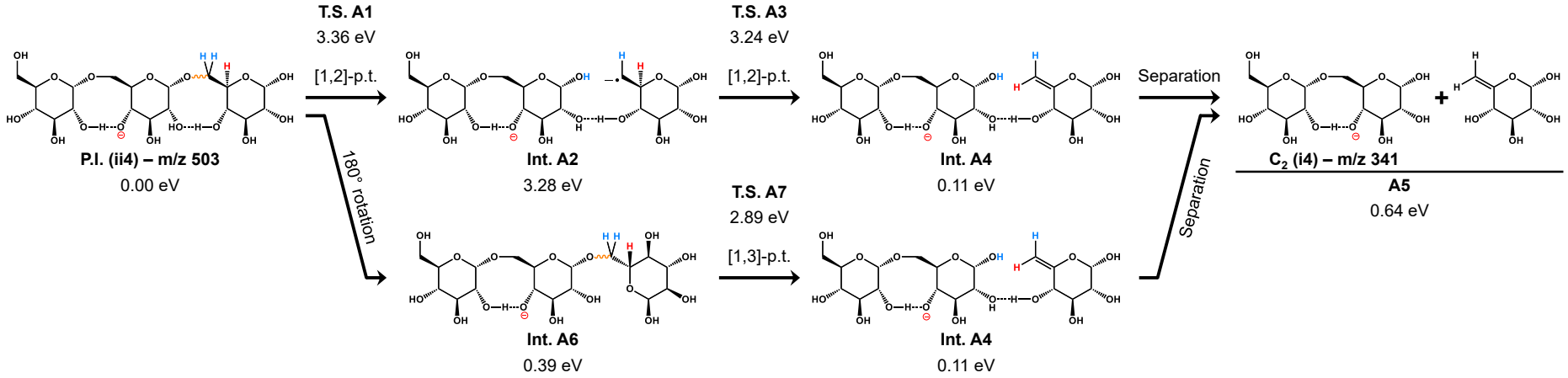
36. Zheng, X., Aly, N.A., Zhou, Y., Dupuis, K.T., Bilbao, A., Paurus, V.L., Orton, D.J., Wilson, R., Payne, S.H., Smith, R.D., Baker, E.S.: A structural examination and collision cross section database for over 500 metabolites and xenobiotics using drift tube ion mobility spectrometry. Chemical Science. 8, 7724–7736 (2017). https://doi.org/10.1039/C7SC03464D

37. Rashid, A.M., Saalbach, G., Bornemann, S.: Discrimination of large maltooligosaccharides from isobaric dextran and pullulan using ion mobility mass spectrometry. Rapid Communications in Mass Spectrometry. 28, 191 (2013). https://doi.org/10.1002/RCM.6771

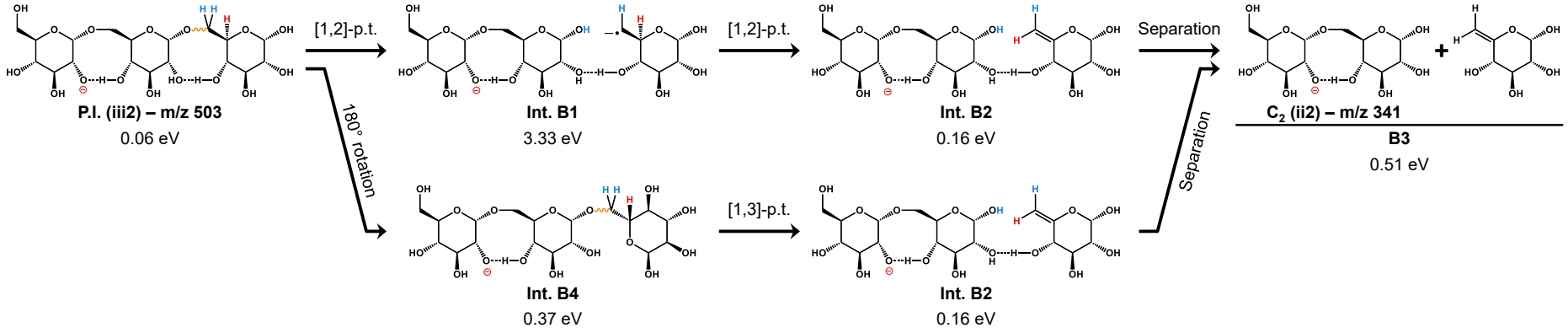
38. Przybylski, C., Bonnet, V.: Discrimination of isomeric trisaccharides and their relative quantification in honeys using trapped ion mobility spectrometry. Food chemistry. 341, (2021). https://doi.org/10.1016/J.FOODCHEM.2020.128182

39. Larriba, C., Christopher J. Hogan, Jr.: Ion Mobilities in Diatomic Gases: Measurement versus Prediction with Non-Specular Scattering Models. Journal of Physical Chemistry A. 117, 3887–3901 (2013). https://doi.org/10.1021/JP312432Z

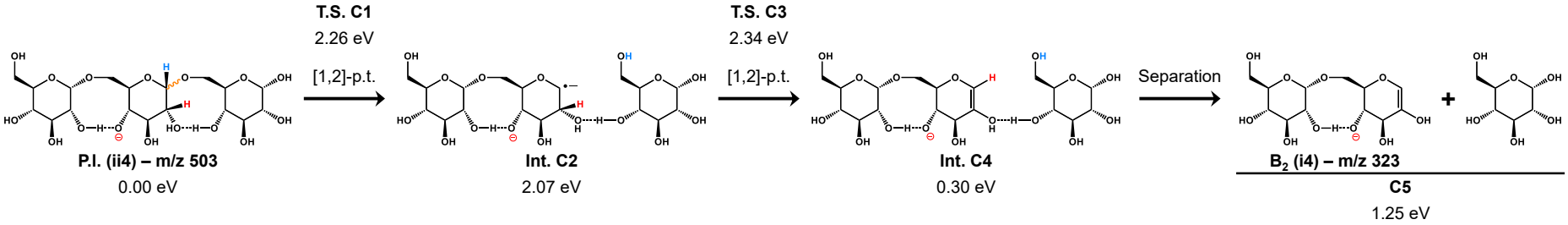
Annexe 1 – Full size reaction schemes



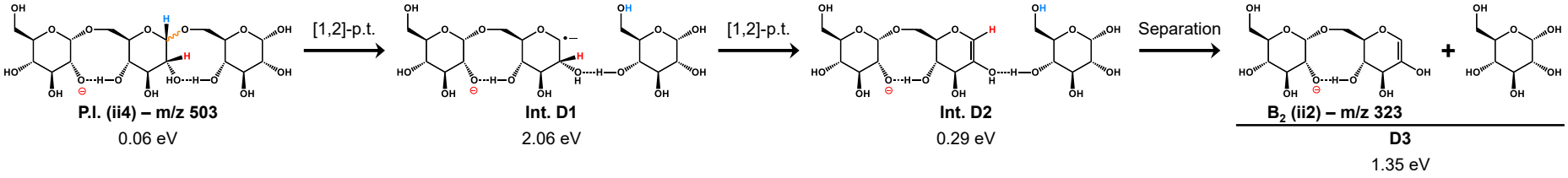
Scheme A. Charge-remote glycosidic bond cleavage mechanisms of the primary precursor ion leading to the formation of the C_2(i4) ion (341 Da).



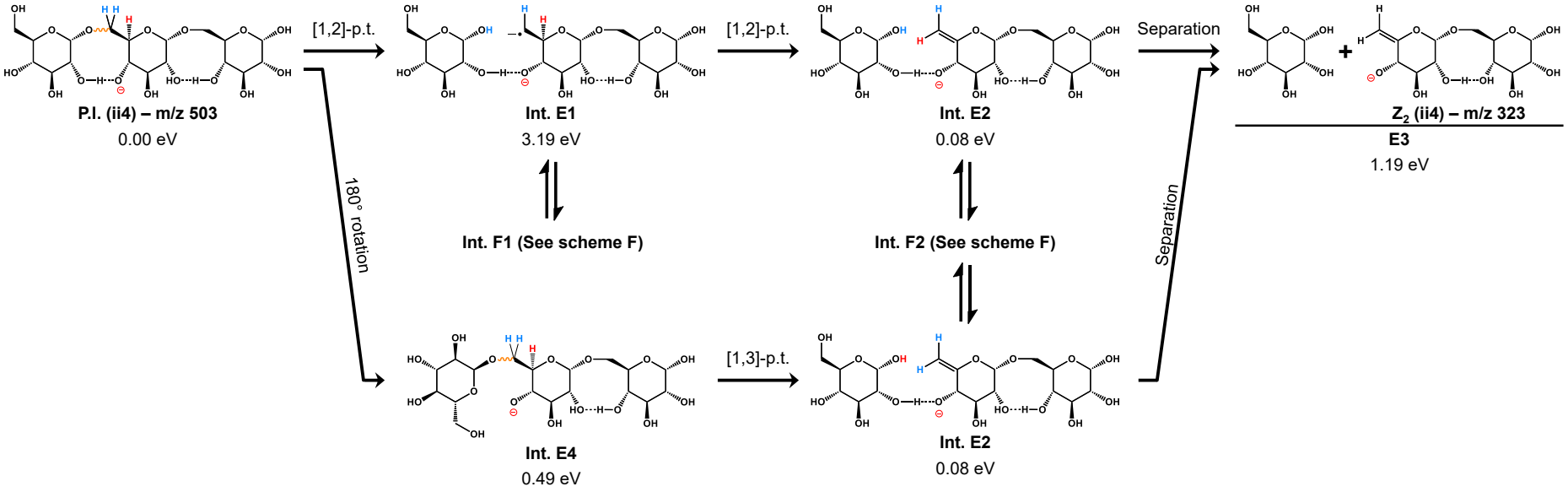
Scheme B. Charge-remote glycosidic bond cleavage mechanisms of the secondary precursor ion leading to the formation of the C_2(ii2) ion (341 Da).



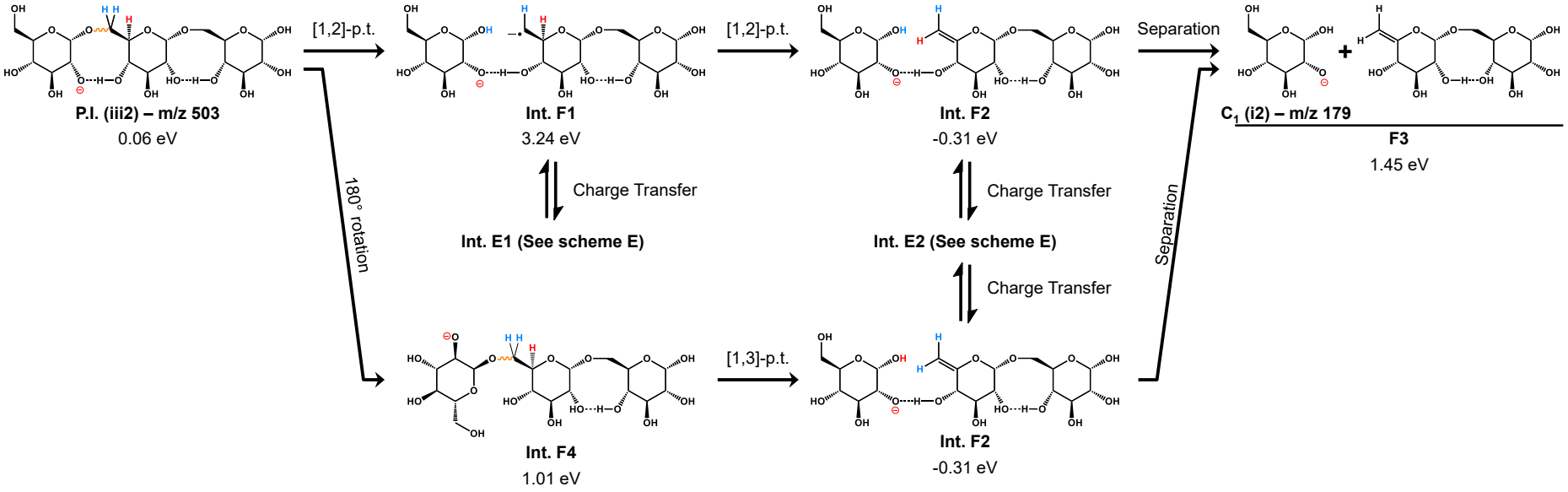
Scheme C. Charge-remote glycosidic bond cleavage mechanisms of the primary precursor ion leading to the formation of the B_2(i4) ion (m/z 323).



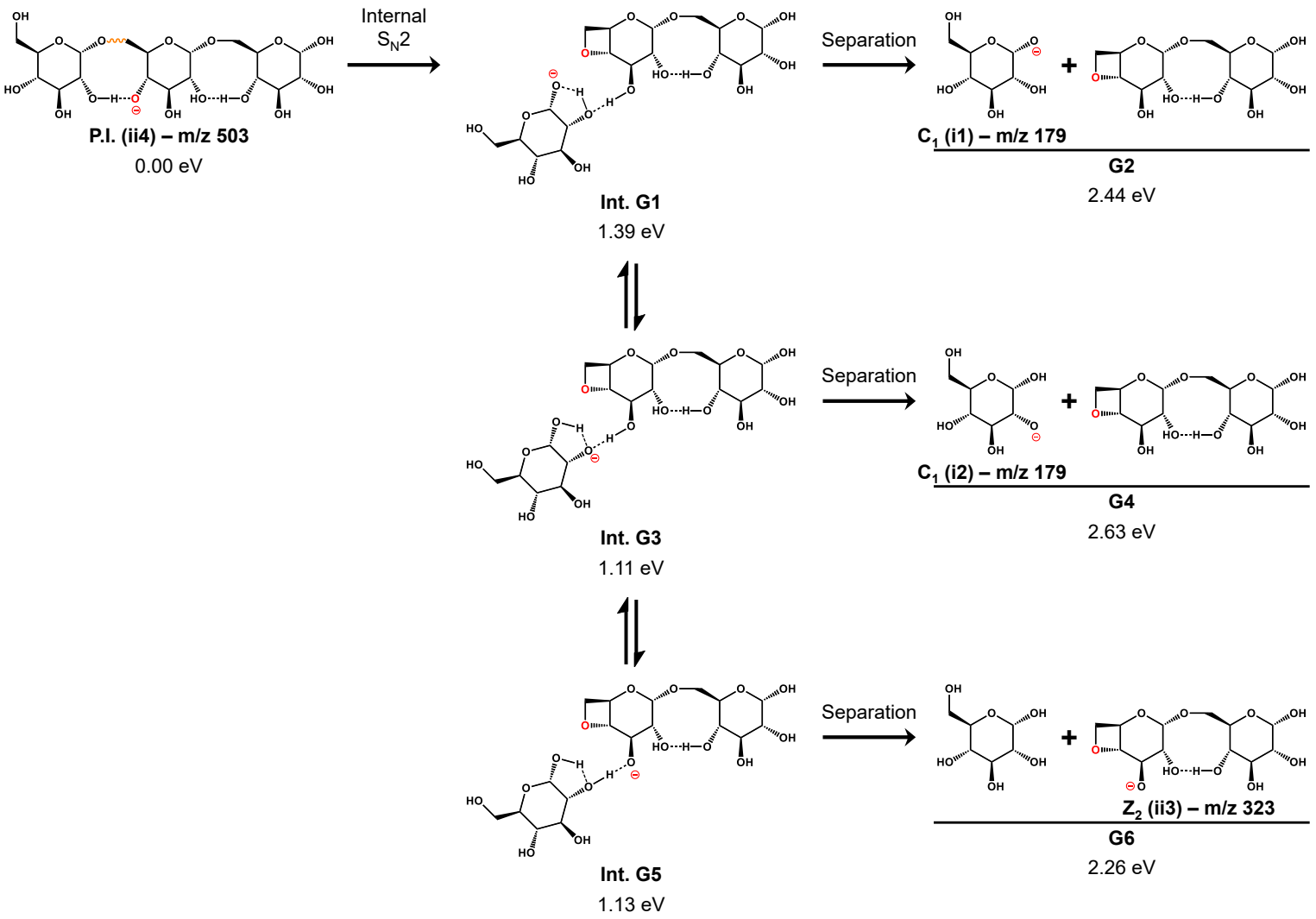
Scheme D. Charge-remote glycosidic bond cleavage mechanisms of the secondary precursor ion leading to the formation of the B_2(ii2) ion (m/z 323).



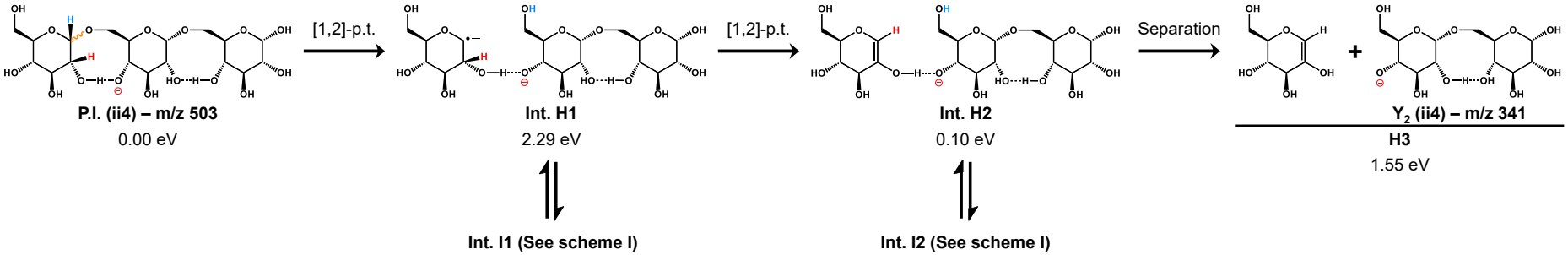
Scheme E. Charge-remote glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of the Z_2(ii4) ion (m/z 323) with the possibility of charge transferring into Scheme F where indicated.



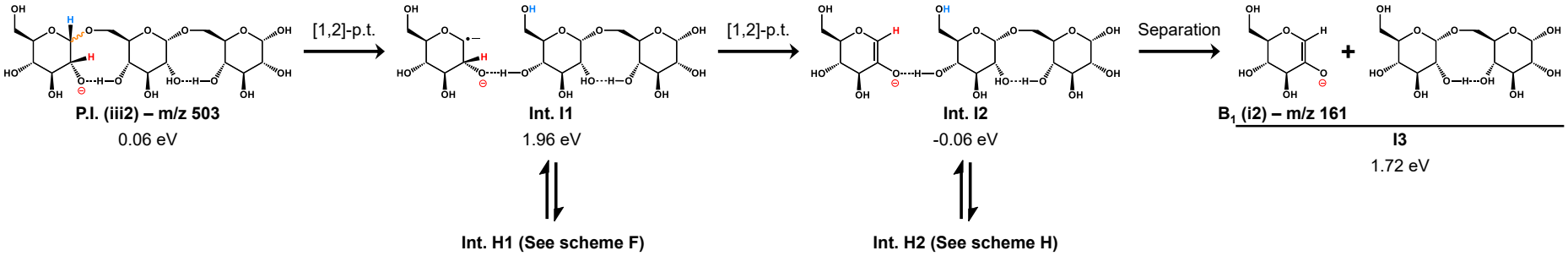
Scheme F. Charge-remote glycosidic bond cleavage mechanisms for the secondary precursor ion leading to the formation of the C_1(i2) ion (m/z 179) with the possibility of charge transferring into scheme E where indicated.



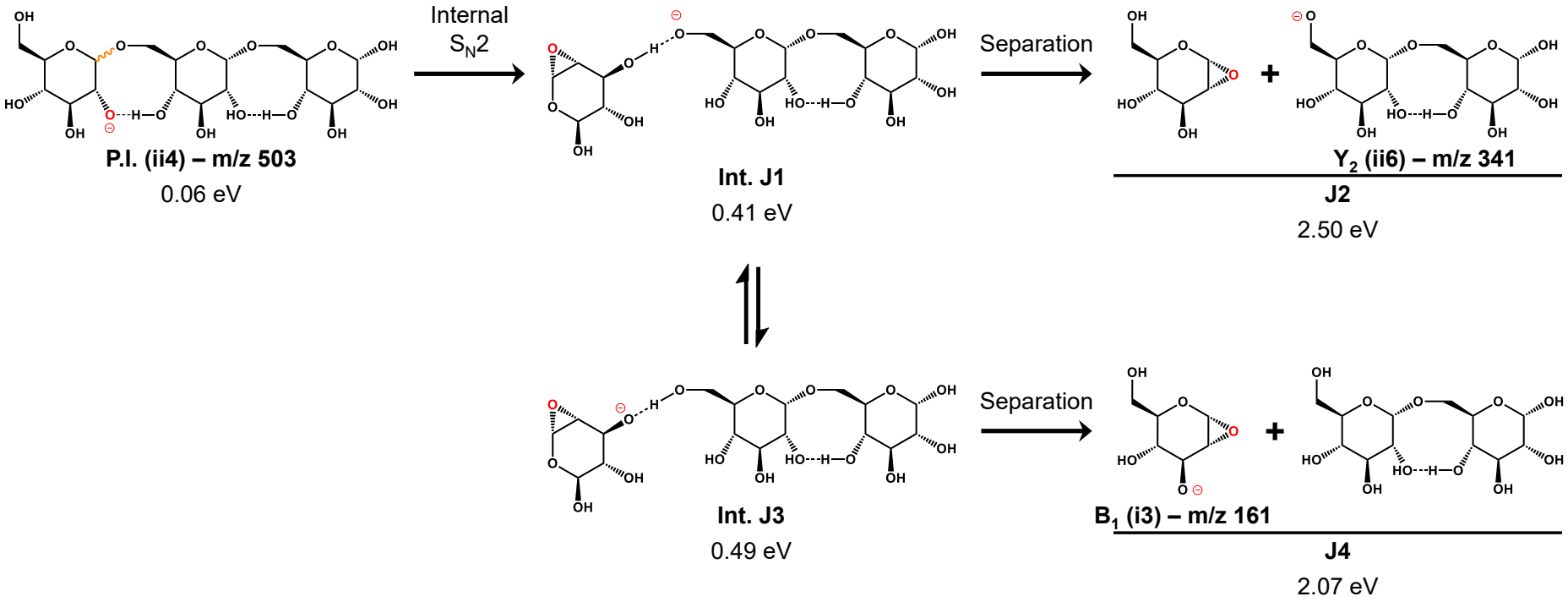
Scheme G. Charge-direct glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of the C_1(i1) ion (m/z 179) with the possibility of charge transferring to form the Z_2(ii3) ion (m/z 323).



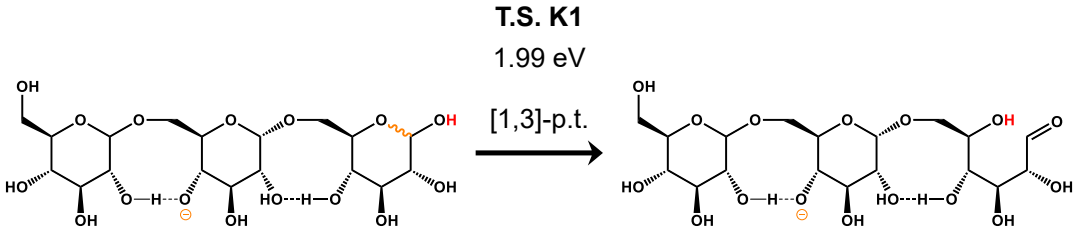
Scheme H. Charge-remote glycosidic bond cleavage mechanisms for the primary precursor ion leading to the formation of the Y_2(ii4) ion (m/z 341). The proximity of the charge within the hydrogen bridge that connects the intermediates, and the ion-neutral complex allows for charge transferring into scheme I.



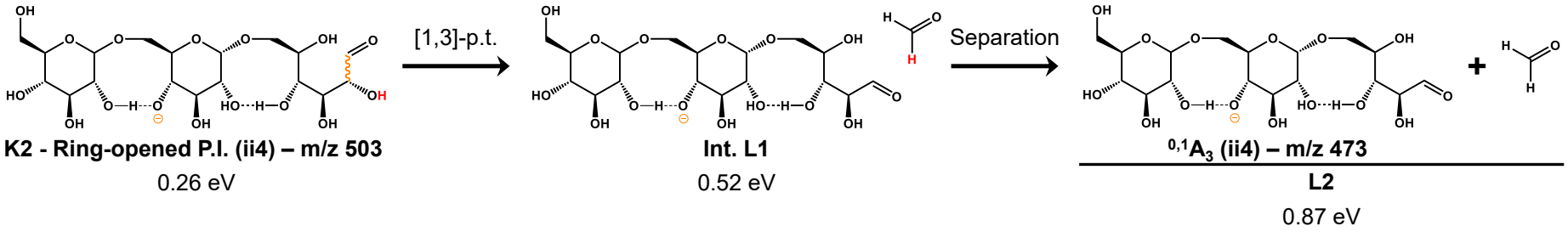
Scheme I. Charge-remote glycosidic bond cleavage mechanisms for the secondary precursor ion leading to the formation of the B_1(i2) ion (m/z 161). The proximity of the charge within the hydrogen bridge that connects the intermediates, and the ion-neutral complex allows for charge transferring into scheme H.



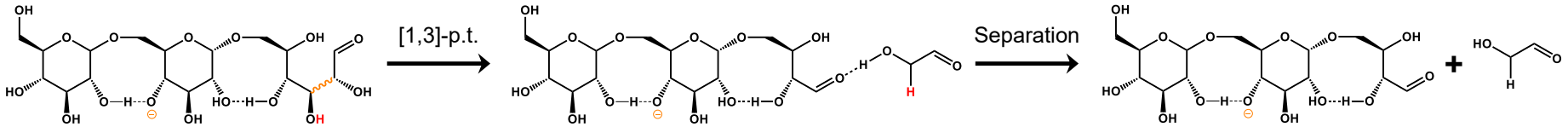
Scheme J. Charge-directed glycosidic bond cleavage mechanism for the secondary precursor ion leading to the formation of the Y_2(ii6) ion (m/z 341). The proximity of the charge within the hydrogen bridge that connects the ion-neutral complex allows for charge to be transferred to the 1,2-anhydroglucose neutral in order to form a B_1(i3) ion (m/z 161).



Scheme K. Charge-remote ring-opening mechanism for the primary precursor ion (ii4) involving a [1,3] proton transfer between the anomeric hydroxyl group and the oxygen forming the cyclic carbohydrate structure



Scheme L. Charge-remote mechanism for the formation of the $^{0,1}A_3$(ii4) ion (m/z 473) via the loss of formaldehyde through a [1,3] proton transfer between the proton of the second hydroxyl group and the anomeric carbon on the reducing ring.



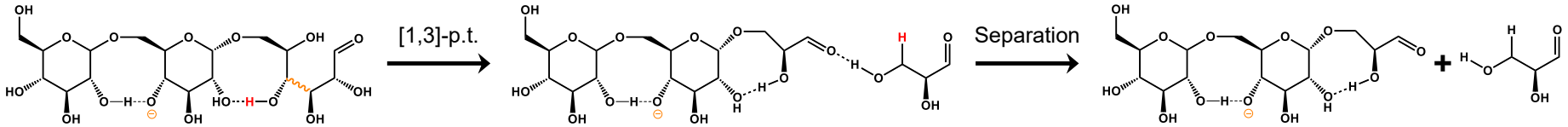
K2 - Ring-opened P.I. (ii4) – m/z 503
0.26 eV

Int. M1
0.23 eV

$^{0,2}A_3$ (ii4) – m/z 443

M2
0.65 eV

Scheme M. Charge-remote mechanism for the formation of the $^{0,2}A_3$(ii4) ion (m/z 443) via the loss of glycolaldehyde through a [1,3] proton transfer between the proton of the third hydroxyl group and the second carbon on the reducing ring.



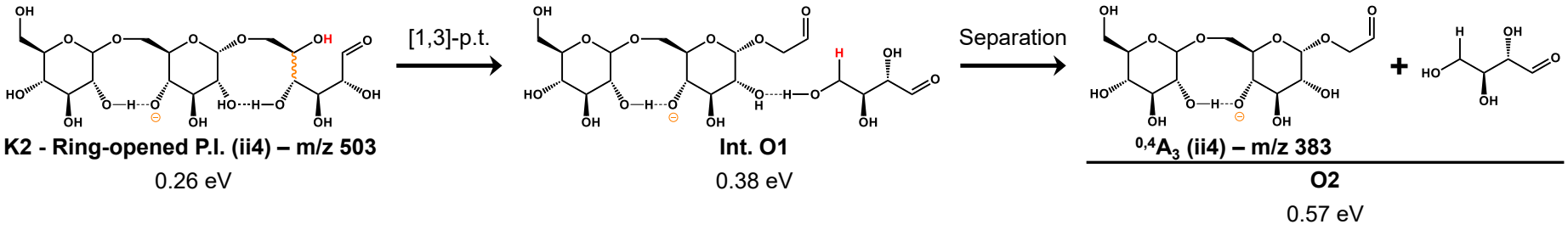
K2 - Ring-opened P.I. (ii4) – m/z 503
0.26 eV

Int.N1
-0.01 eV

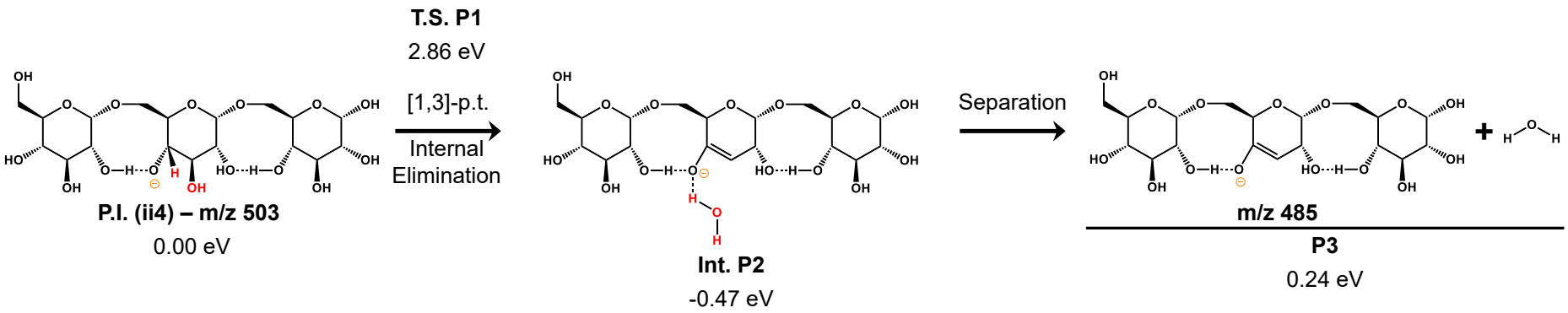
$^{0,3}A_3$ (ii4) – m/z 413

N2
0.37 eV

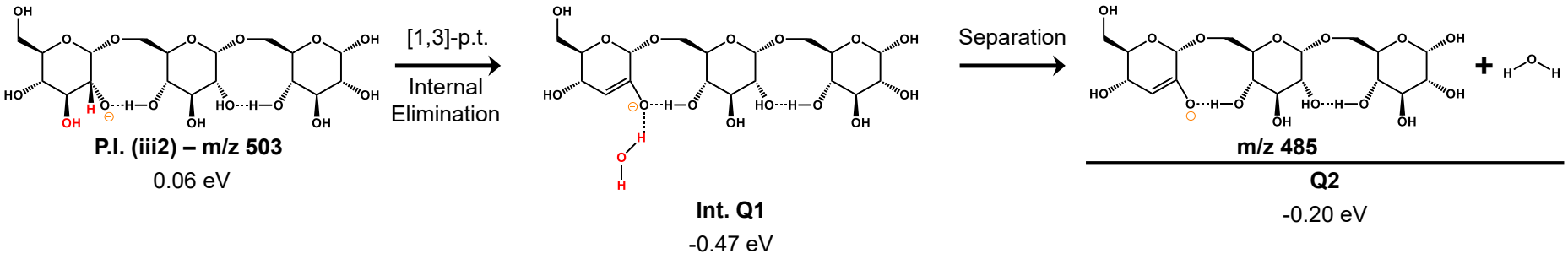
Scheme N. Charge-remote mechanism for the formation of the $^{0,3}A_3$(ii4) ion (413 Da) via the loss of glyceraldehyde through a [1,3] proton transfer between the proton of the fourth hydroxyl group and the third carbon on the reducing ring. A larger version of the schemes can be found in annex 1.



Scheme O. Charge-remote mechanism for the formation of the $^{0,4}A_3$(ii4) ion (m/z 383) via the loss of threose through a [1,3] proton transfer between the proton of the fifth hydroxyl group and the fourth carbon on the reducing ring. A larger version of the schemes can be found in annex 1.



Scheme P. Charge-remote mechanism for the formation loss of water from the ii3 position of the primary precursor ion leading to the production of the m/z 485 ion.

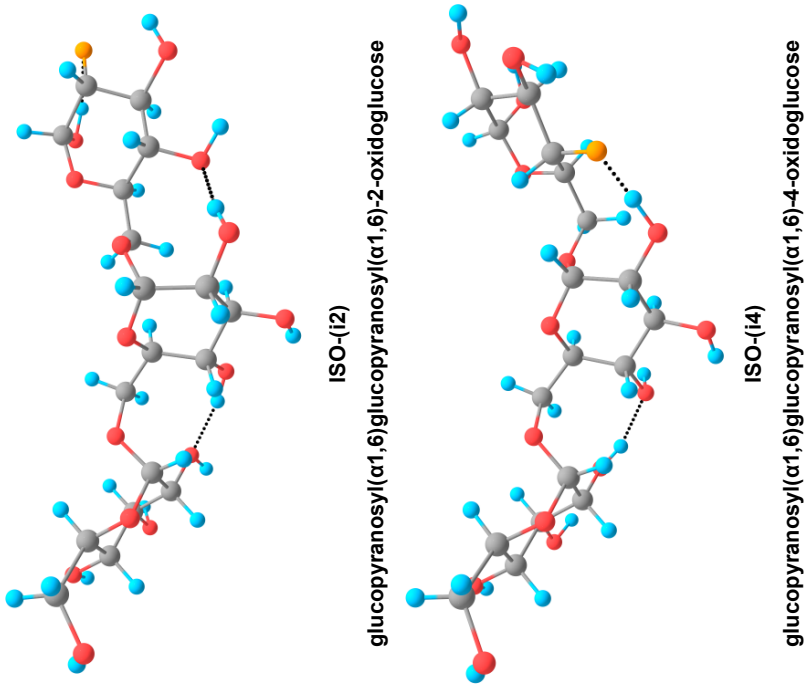


Scheme Q. Charge-remote mechanism for the formation loss of water from the iii3 position of the secondary precursor ion leading to the production of the m/z 485 ion.

Annexe 2 – Optimised precursor ion structures of the trisaccharides .

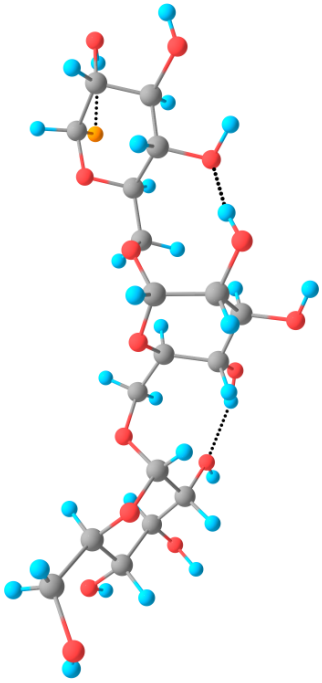
A2.1 Isomaltriose

The deprotonated oxygen is highlighted in orange.

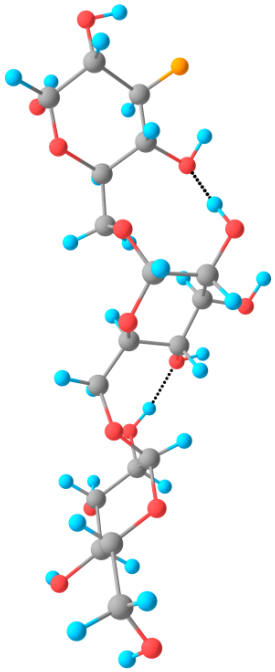


ISO-(i2) glucopyranosyl(α1,6)glucopyranosyl(α1,6)-2-oxidoglucose

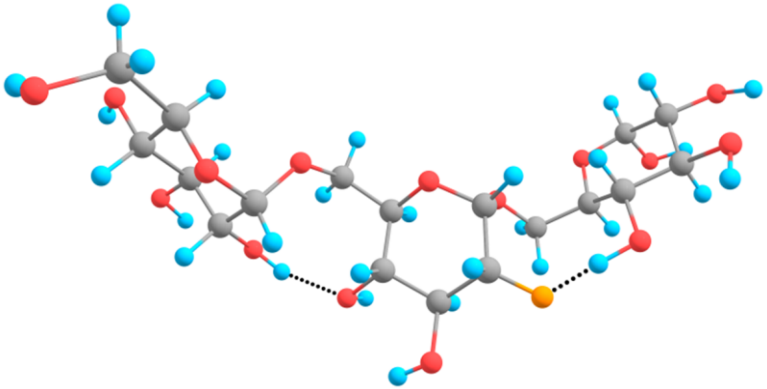
ISO-(i4) glucopyranosyl(α1,6)glucopyranosyl(α1,6)-4-oxidoglucose



ISO-(i1) glucopyranosyl(α1,6)glucopyranosyl(α1,6)-1-oxidoglucose

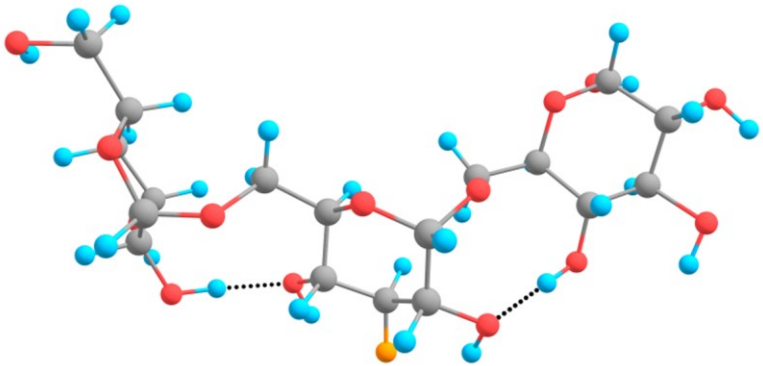


ISO-(i3) glucopyranosyl(α1,6)glucopyranosyl(α1,6)-3-oxidoglucose



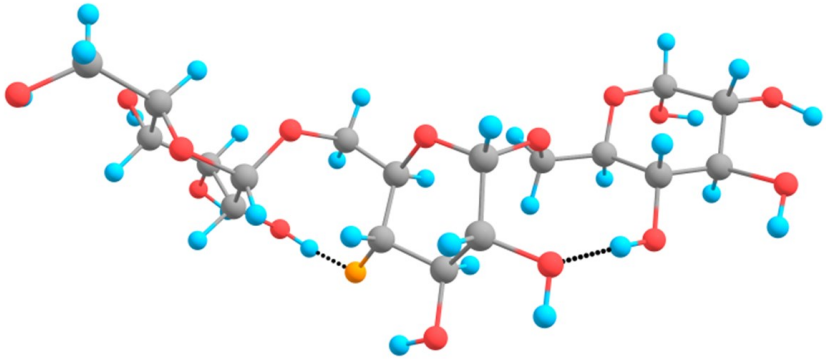
ISO-(ii2)

glucopyranosyl(α1,6)-2-oxido-glucopyranosyl(α1,6)glucose



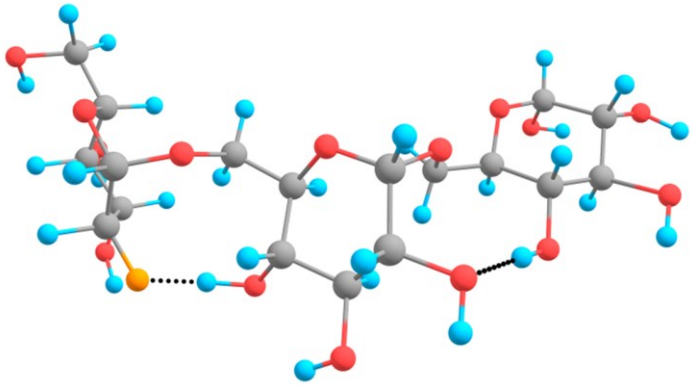
ISO-(ii3)

glucopyranosyl(α1,6)-3-oxido-glucopyranosyl(α1,6)glucose



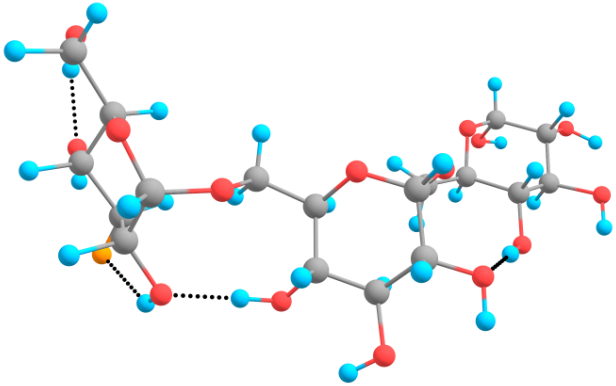
ISO-(ii4)

glucopyranosyl(α1,6)-4-oxido-glucopyranosyl(α1,6)glucose



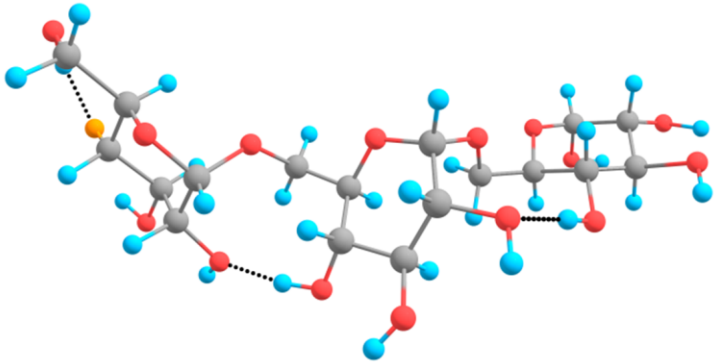
ISO-(iii2)

2-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,6)glucose



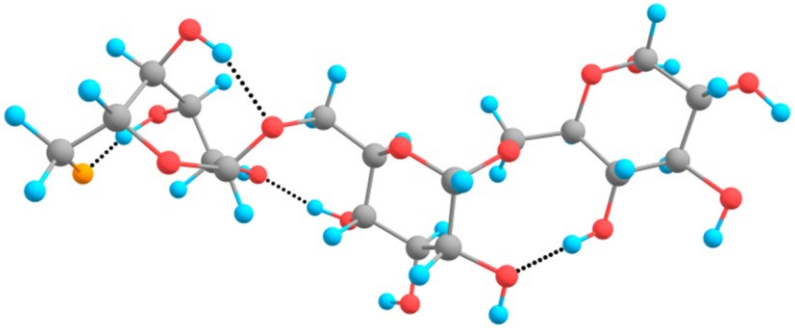
ISO-(iii3)

2-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,6)glucose



ISO-(iii4)

2-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,6)glucose

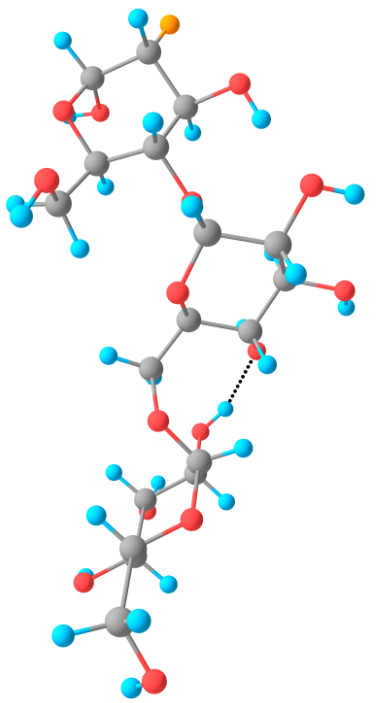


ISO-(iii6)

2-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,6)glucose

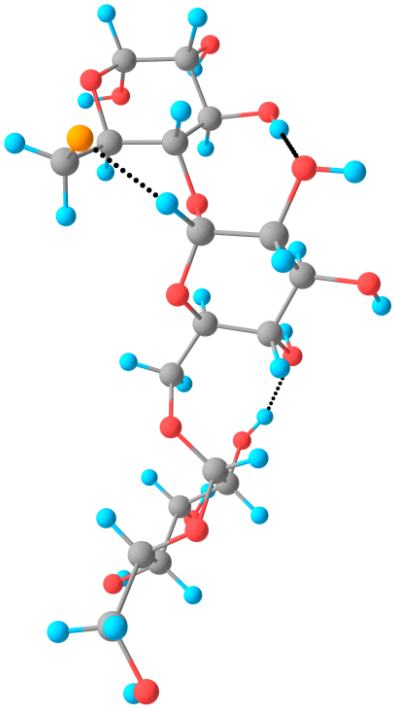
A2.2. Panose (PAN)

The deprotonated oxygen is highlighted in orange.



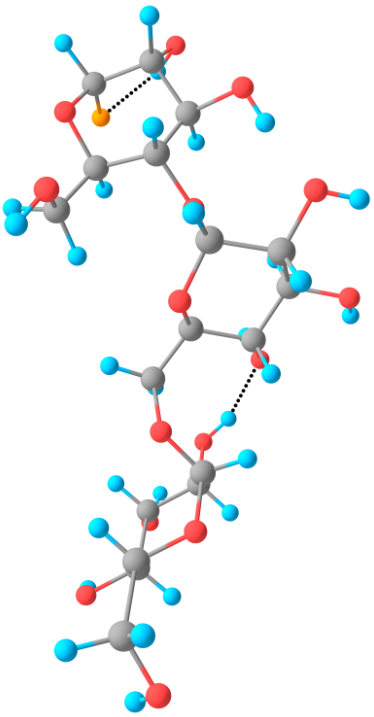
PAN-(i2)

glucopyranosyl(α1,6)glucopyranosyl(α1,4)-2-oxidoglucose



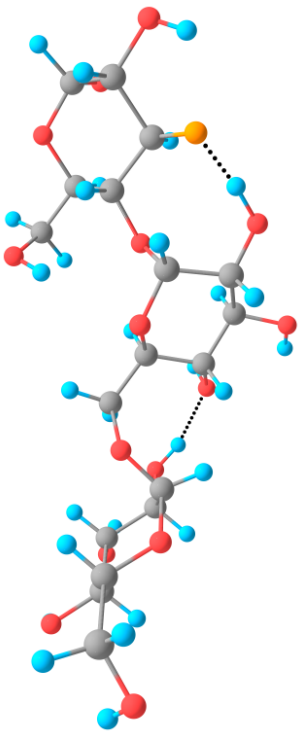
PAN-(i6)

glucopyranosyl(α1,6)glucopyranosyl(α1,4)-6-oxidoglucose



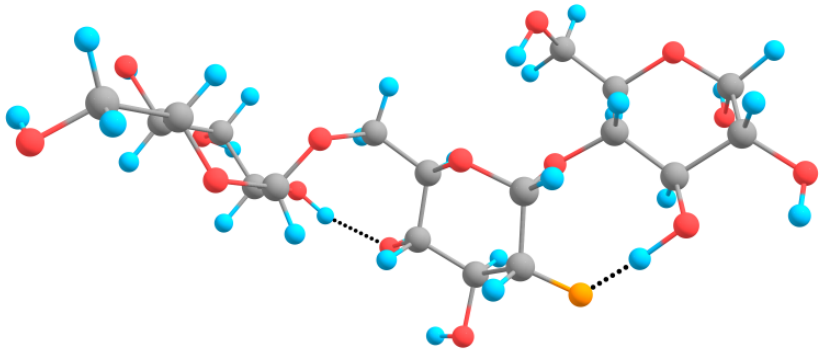
PAN-(i1)

glucopyranosyl(α1,6)glucopyranosyl(α1,4)-1-oxidoglucose



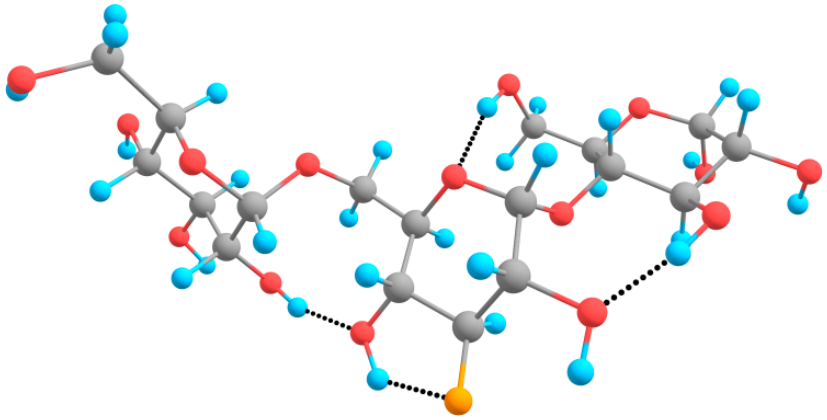
PAN-(i3)

glucopyranosyl(α1,6)glucopyranosyl(α1,4)-3-oxidoglucose



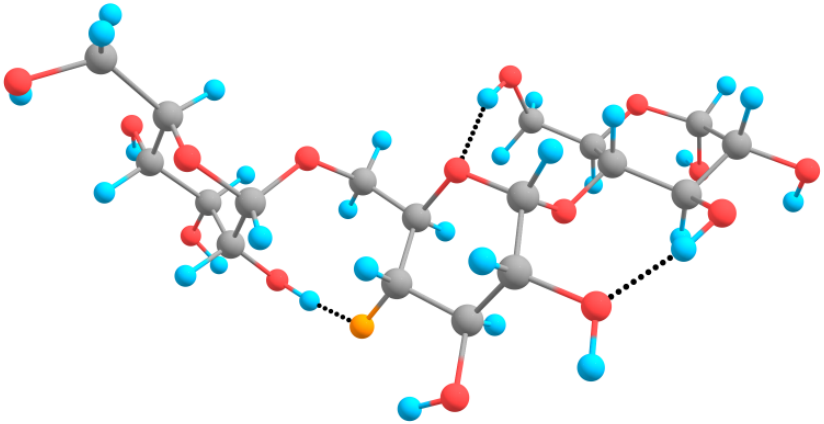
PAN-(ii2)

glucopyranosyl(α1,6)-2-oxidoglucopyranosyl(α1,4)glucose



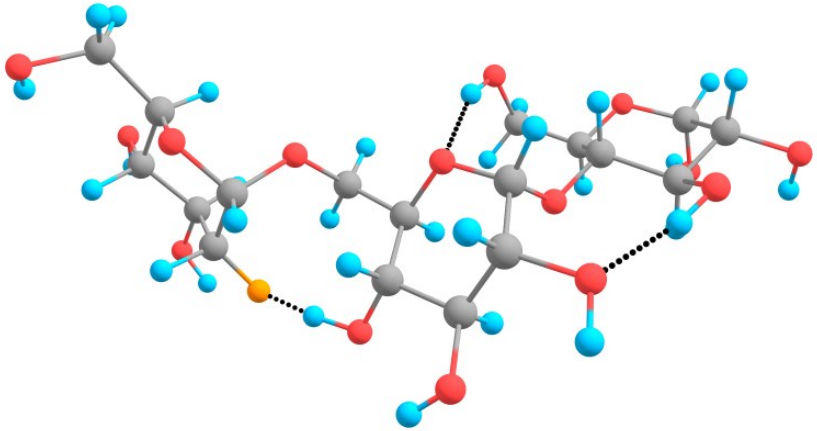
PAN-(ii3)

glucopyranosyl(α1,6)-3-oxidoglucopyranosyl(α1,4)glucose



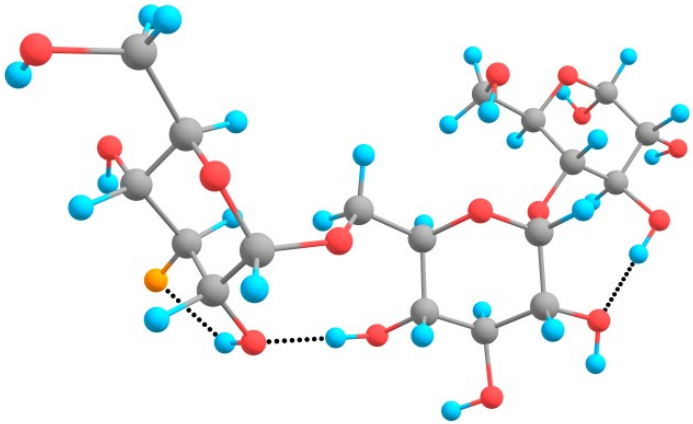
PAN-(ii4)

glucopyranosyl(α1,6)-4-oxidoglucopyranosyl(α1,4)glucose



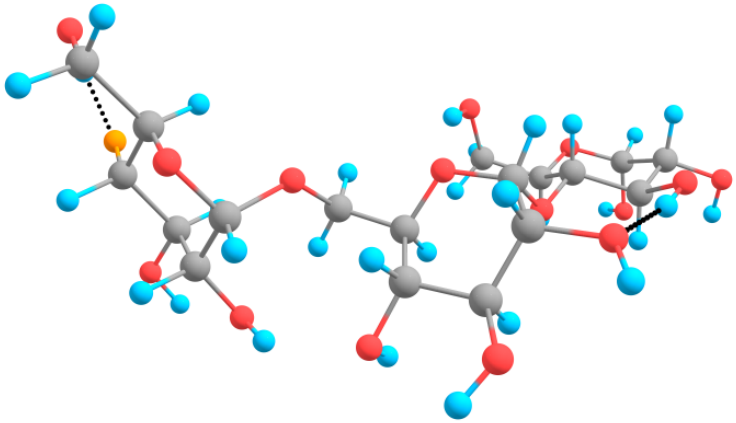
PAN-(iii2)

2-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,4)glucose



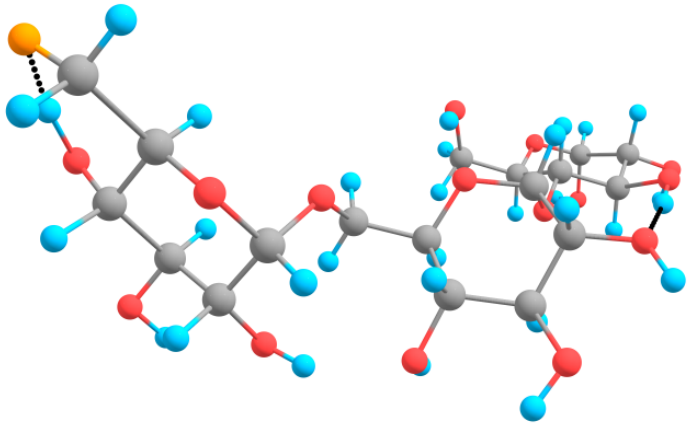
PAN-(iii3)

3-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,4)glucose



PAN-(iii4)

4-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,4)glucose

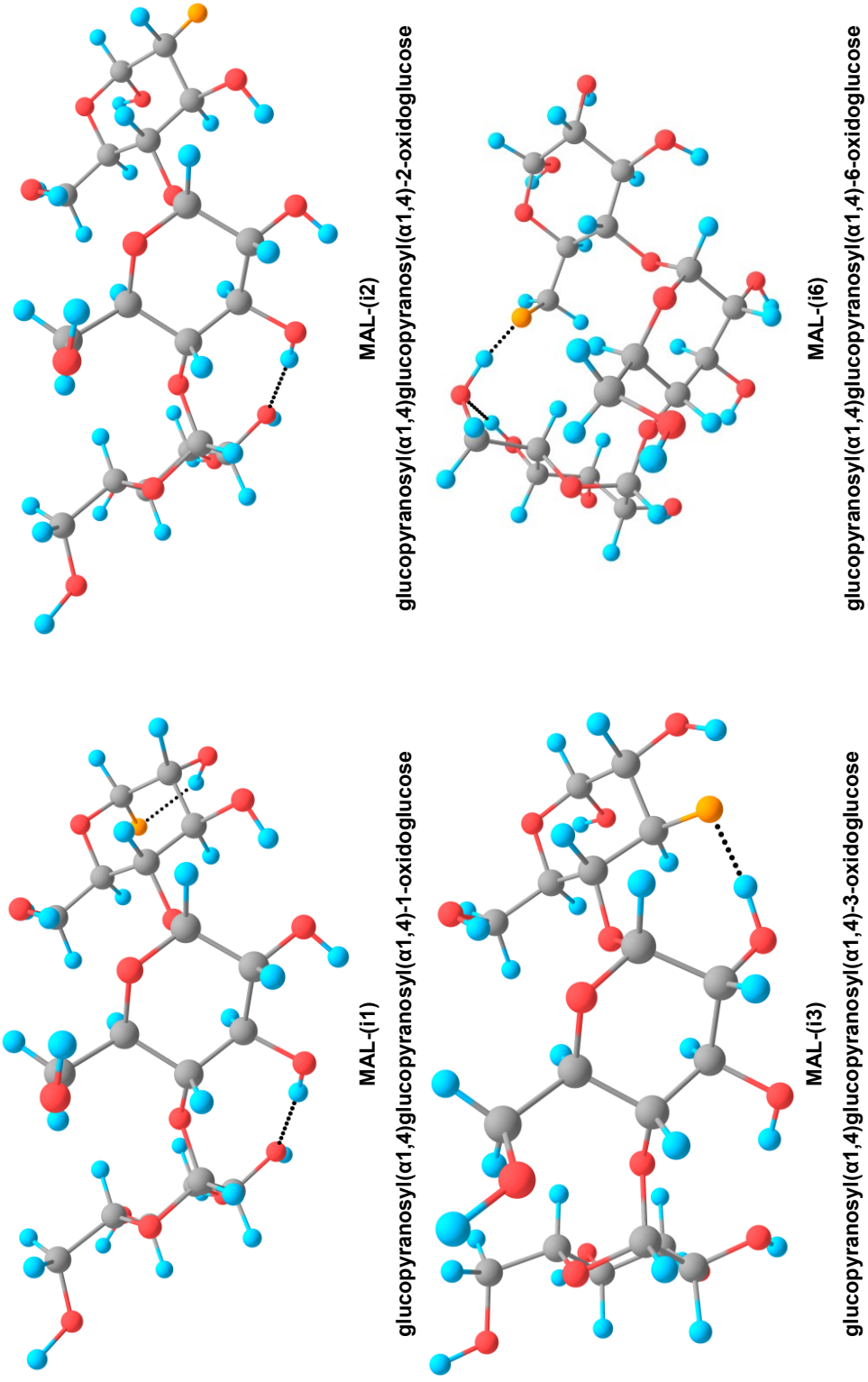


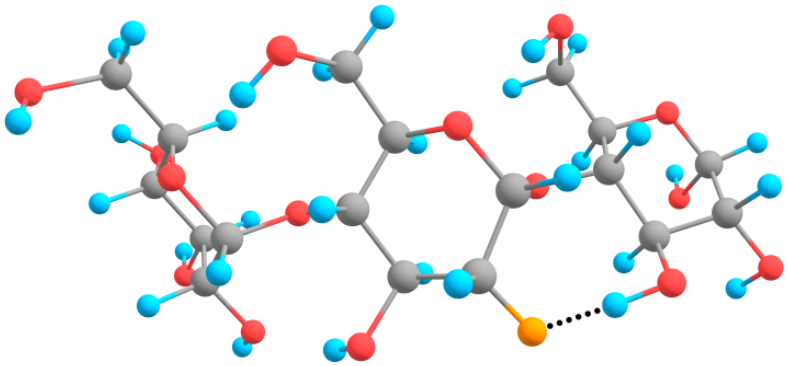
PAN-(iii6)

6-oxidoglucopyranosyl(α1,6)glucopyranosyl(α1,4)glucose

A2.3. Maltotriose (MAL)

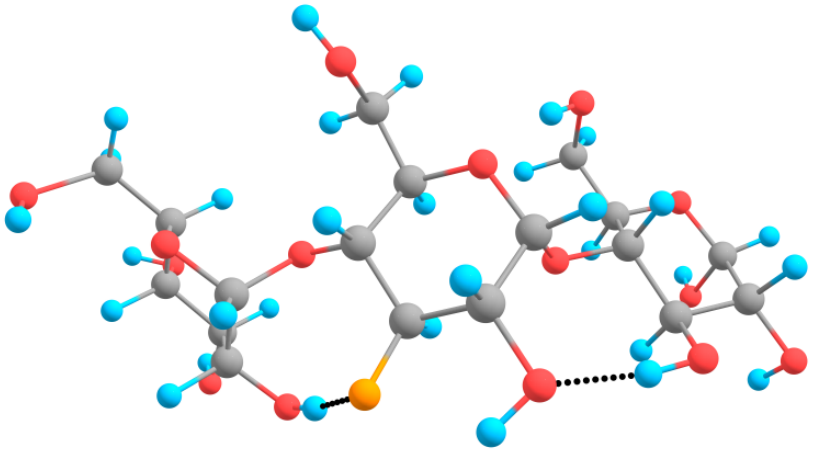
The deprotonated oxygen is highlighted in orange.





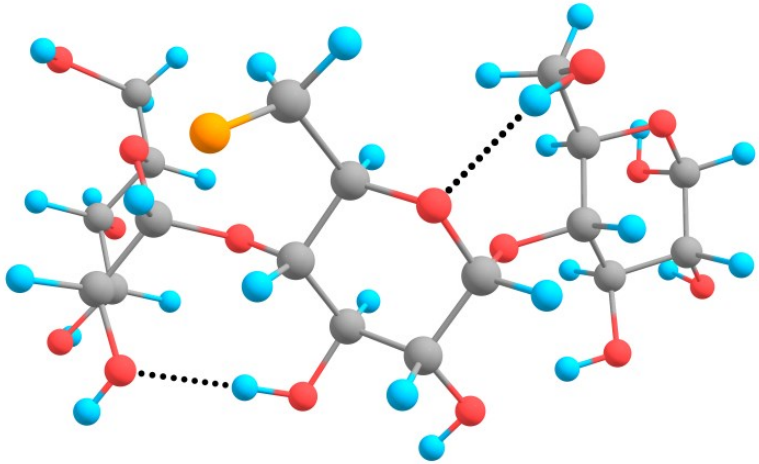
MAL-(ii2)

glucopyranosyl(α1,4)-2-oxidoglucopyranosyl(α1,4)glucose



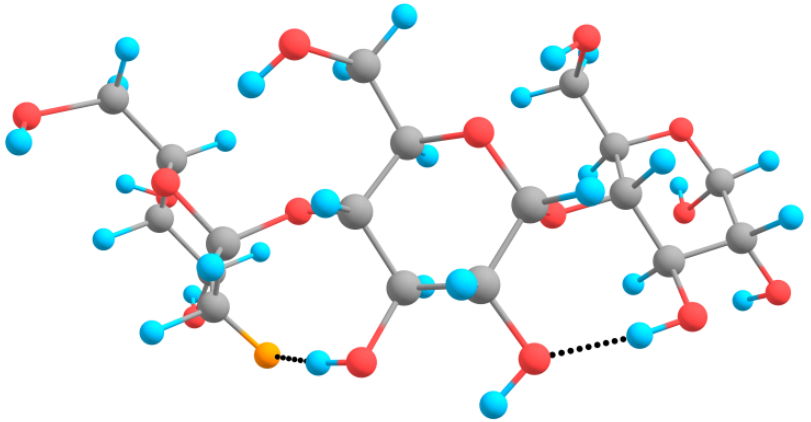
MAL-(ii3)

glucopyranosyl(α1,4)-3-oxidoglucopyranosyl(α1,4)glucose



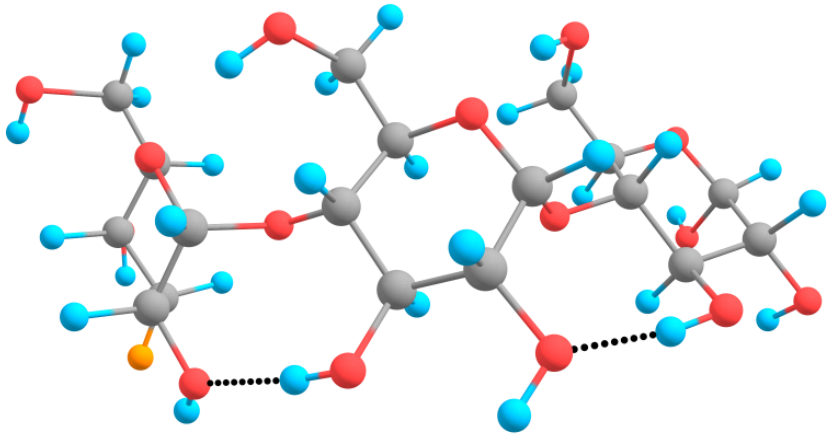
MAL-(ii6)

glucopyranosyl(α1,4)-6-oxidoglucopyranosyl(α1,4)glucose



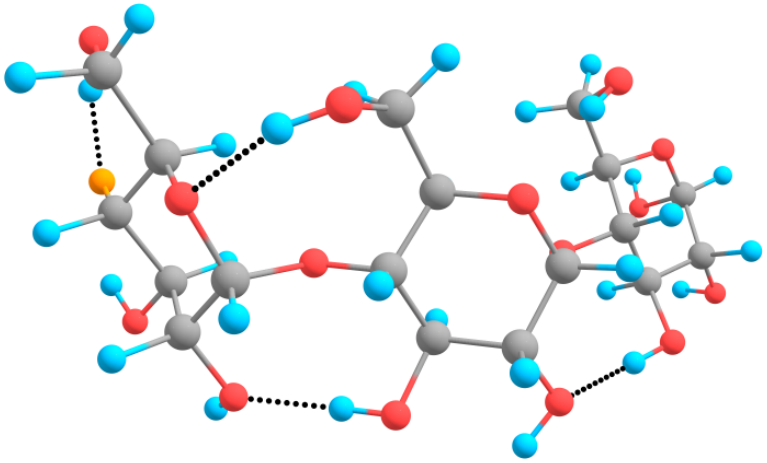
MAL-(iii2)

2-oxidoglucopyranosyl(α1,4)glucopyranosyl(α1,4)glucose



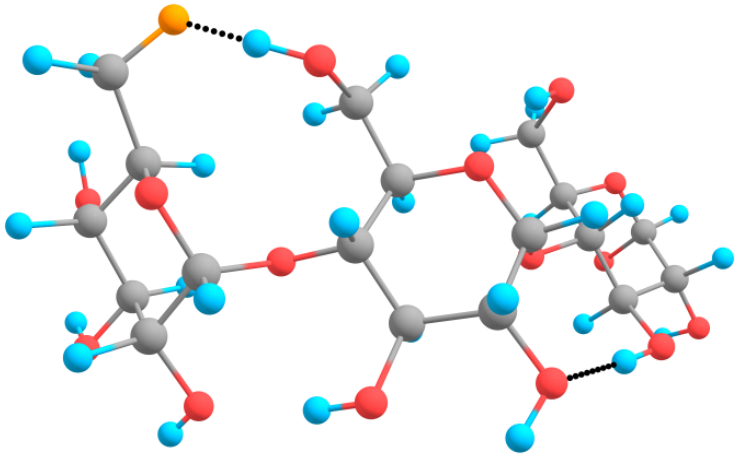
MAL-(iii3)

3-oxidoglucopyranosyl(α1,4)glucopyranosyl(α1,4)glucose



MAL-(iii4)

4-oxidoglucopyranosyl(α1,4)glucopyranosyl(α1,4)glucose

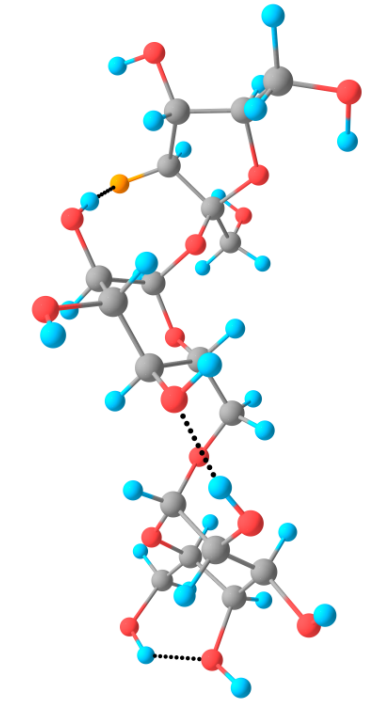


MAL-(iii6)

6-oxidoglucopyranosyl(α1,4)glucopyranosyl(α1,4)glucose

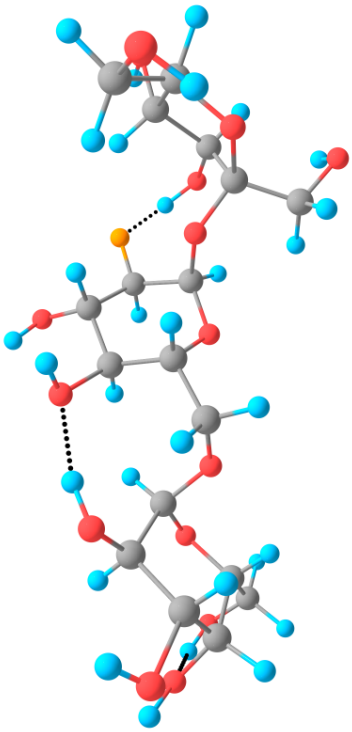
A2.4 Raffinose (RAF)

The deprotonated oxygen is highlighted in orange.



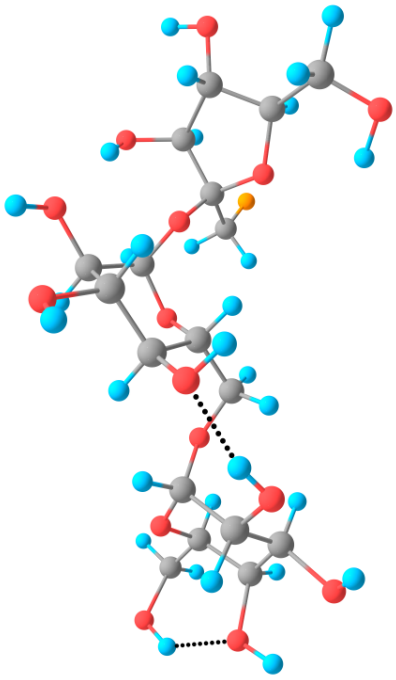
RAF-(i3)

galactopyranosyl(α1,6)glucopyranosyl(α1,2)-3-oxidofructose



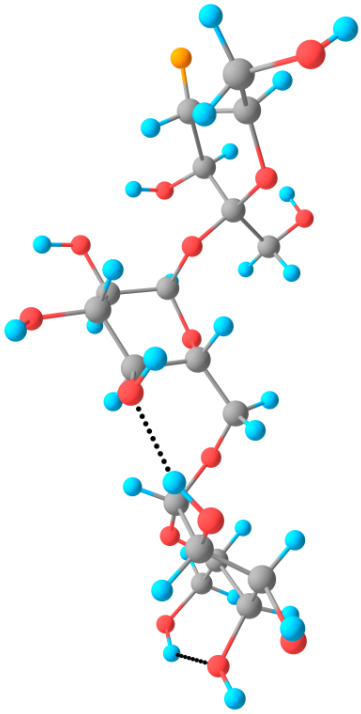
RAF-(ii2)

galactopyranosyl(α1,6)-2-oxidoglucopyranosyl(α1,2)fructose



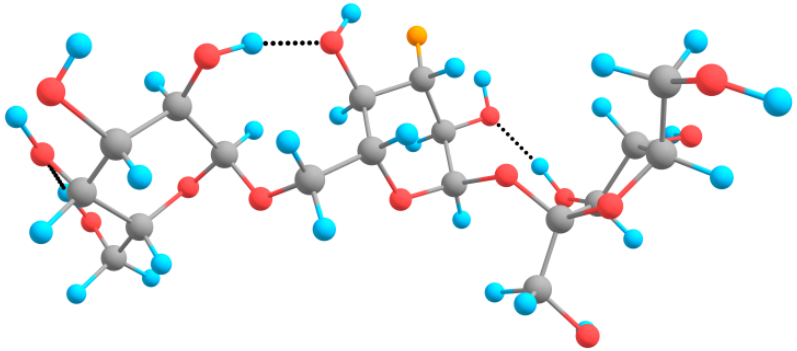
RAF-(i1)

galactopyranosyl(α1,6)glucopyranosyl(α1,2)-1-oxidofructose



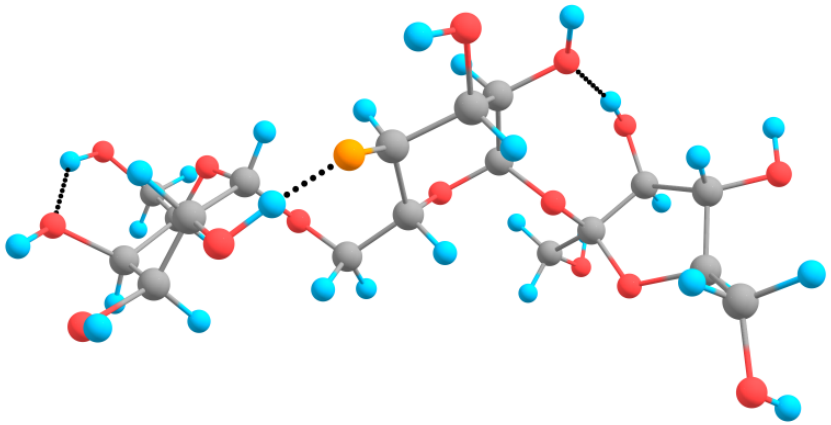
RAF-(i4)

galactopyranosyl(α1,6)glucopyranosyl(α1,2)-4-oxidofructose



RAF-(ii3)

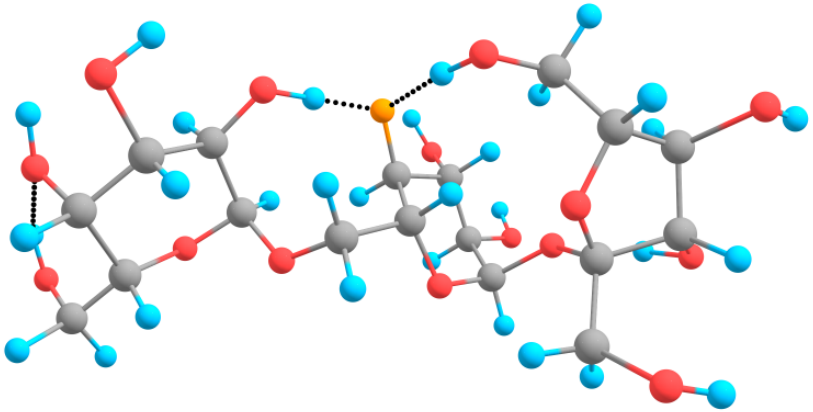
galactopyranosyl(α1,6)-2-oxidoglucopyranosyl(α1,2)fructose



RAF-(ii4a)

galactopyranosyl(α1,6)-4-oxidoglucopyranosyl(α1,2)fructose

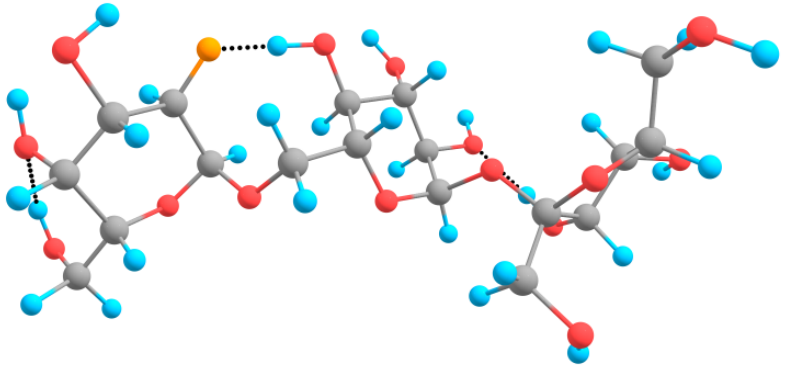
unfolded isomer



RAF-(ii4b)

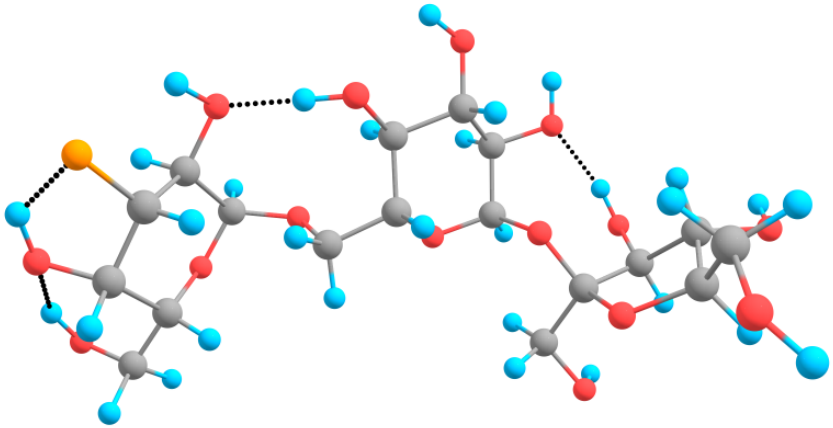
galactopyranosyl(α1,6)-4-oxidoglucopyranosyl(α1,2)fructose

Folded isomer



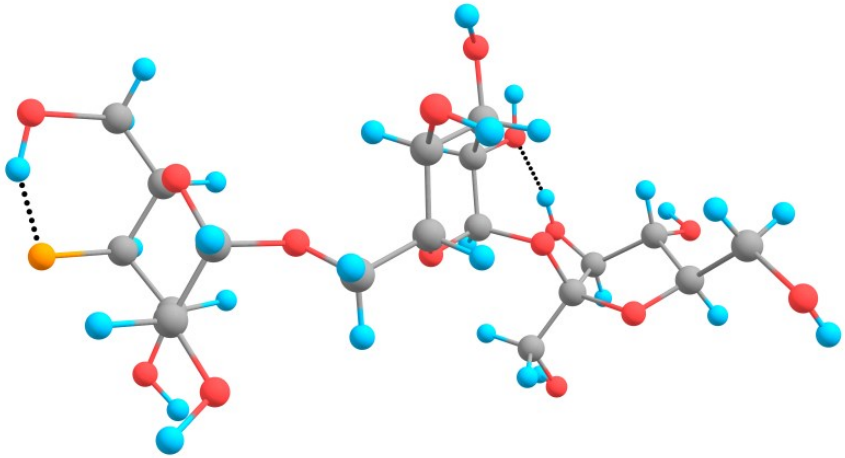
RAF-(iii2)

2-oxidogalactopyranosyl(α1,6)glucopyranosyl(α1,2)fructose



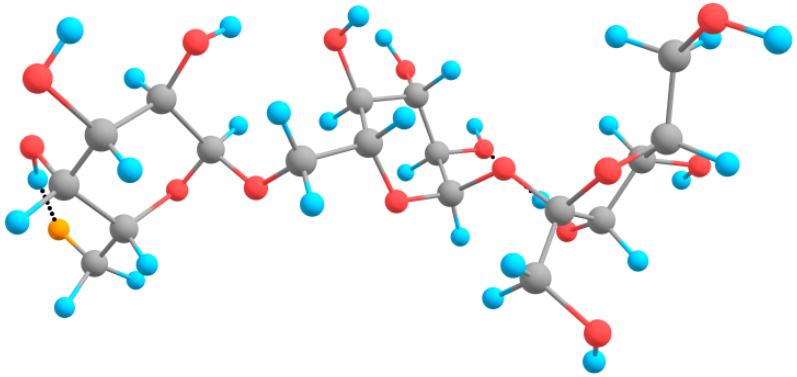
RAF-(iii3)

3-oxidogalactopyranosyl(α1,6)glucopyranosyl(α1,2)fructose



RAF-(iii4)

4-oxidogalactopyranosyl(α1,6)glucopyranosyl(α1,2)fructose

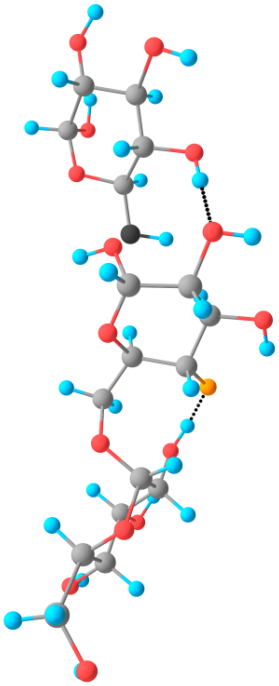


RAF-(iii6)

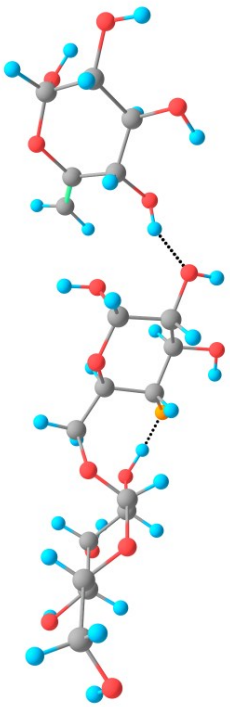
6-oxidogalactopyranosyl(α1,6)glucopyranosyl(α1,2)fructose

Annexe 3 – Optimised structures from the reaction schemes.

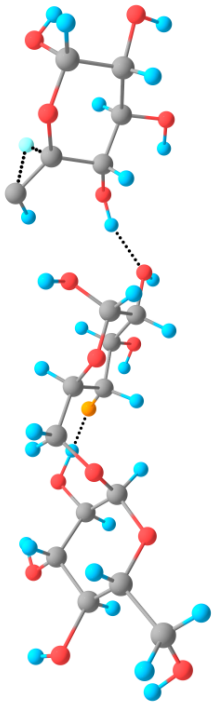
The deprotonated oxygen is highlighted in orange.
The migrating protons in transition states are highlighted in light blue.
The carbons holding a radical anion are highlighted in black



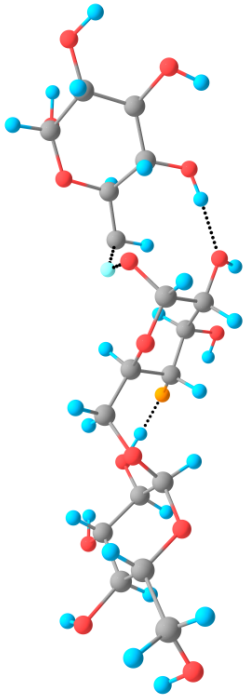
A1 – initial [1,2]-p.t. transition state for the C_2 glycosidic bond cleavage of the primary precursor ion.



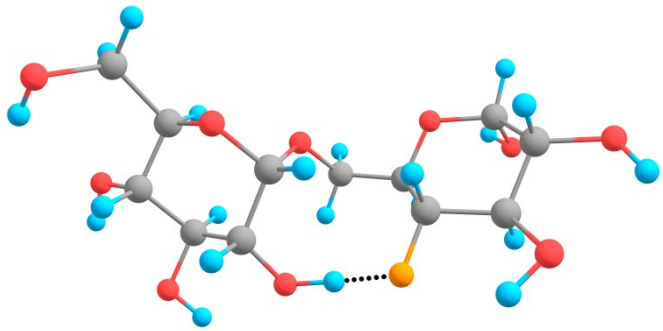
A2 – initial [1,2]-p.t. radical anion intermediate C_2 glycosidic bond cleavage of the primary precursor ion.



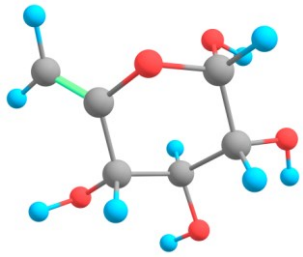
A3 – second [1,2]-p.t. transition state C_2 glycosidic bond cleavage of the primary precursor ion.



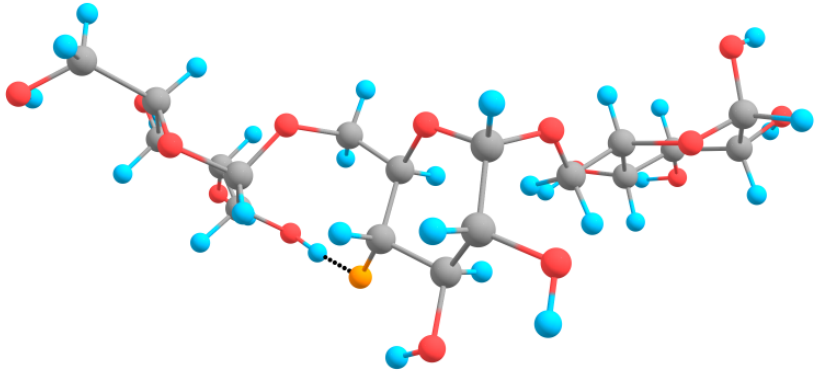
A4 – Ion-Neutral complex for the C_2 glycosidic bond cleavage of the primary precursor ion.



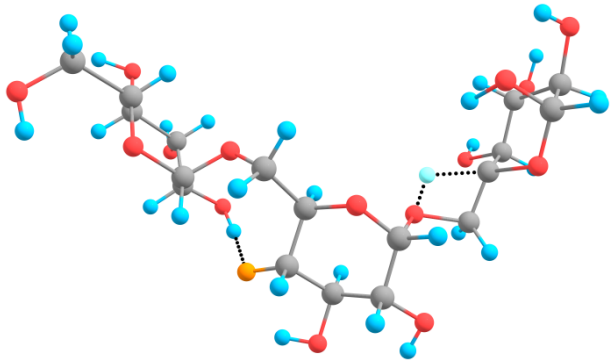
A5(a) – C_2(i4) product ion



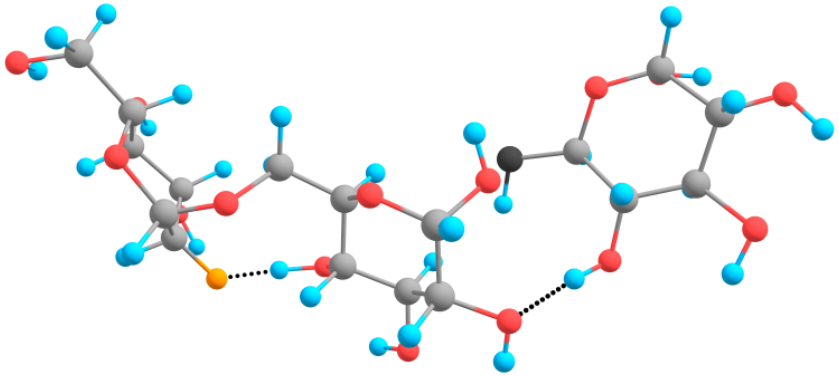
A5(b) – C_2(i4) neutral loss



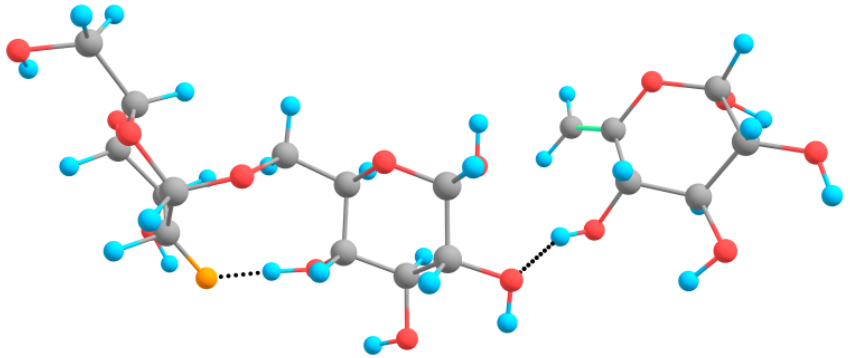
A6 – Flipped-ring intermediate for the C_2 glycosidic bond cleavage of the primary precursor ion.



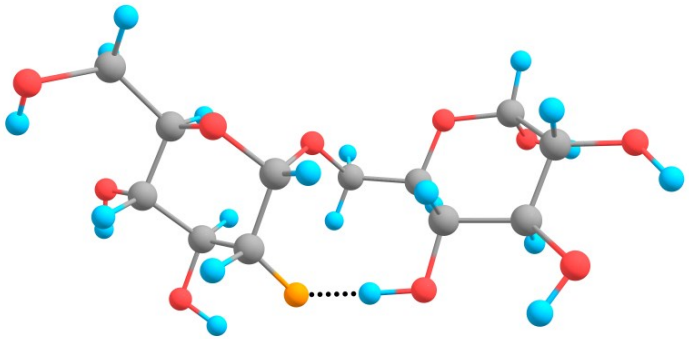
A7 – [1,3]-p.t. transition state for the C_2 glycosidic bond cleavage of the primary precursor ion.



B1 – Initial [1,2]-p.t. radical anion intermediate for the C_2 glycosidic bond cleavage of the secondary precursor ion.



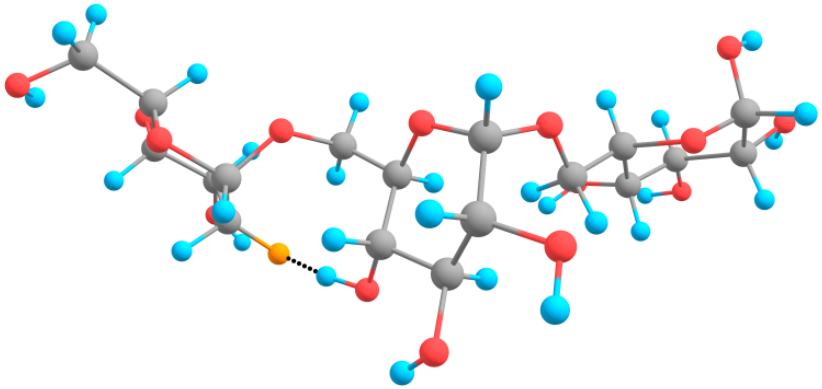
B2 – Ion-Neutral complex for the C_2 glycosidic bond cleavage of the secondary precursor ion.



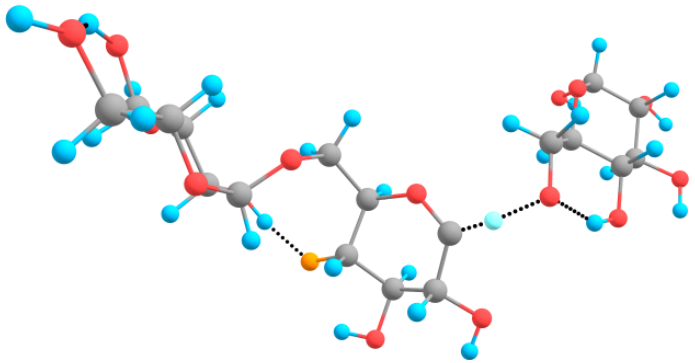
B3(a) - C_2(ii2) product ion

See A5(b)

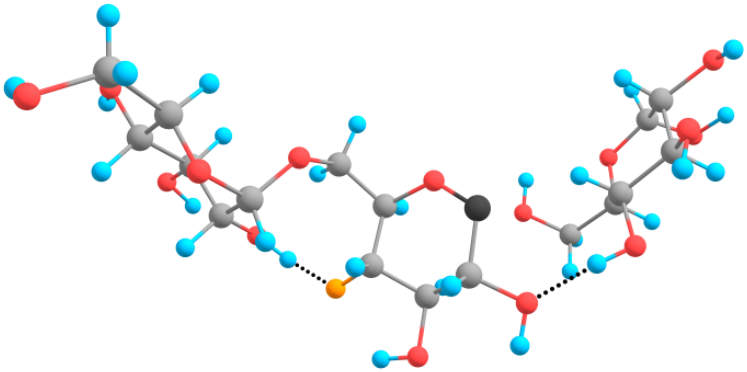
B3(b) - C_2(ii2) neutral loss



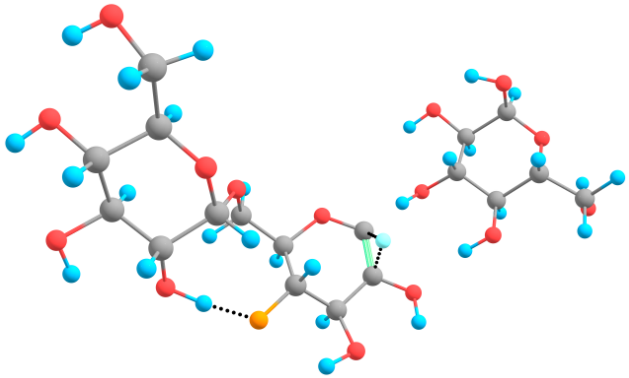
B4 – Flipped-ring intermediate for the C_2 glycosidic bond cleavage of the secondary precursor ion.



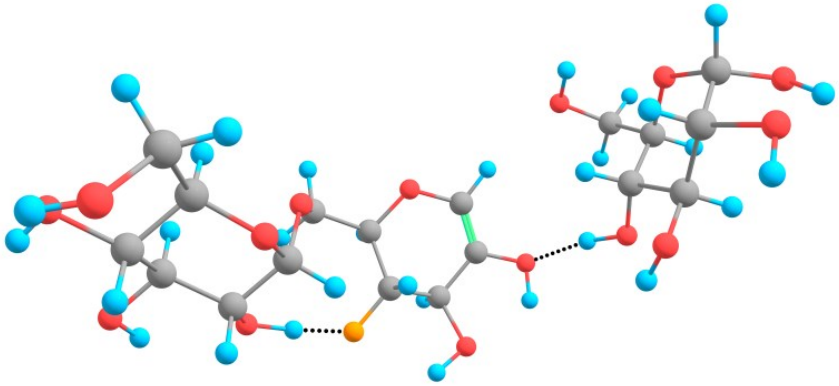
C1 – Initial [1,2]-p.t. transition state for the B_2 glycosidic bond cleavage of the primary precursor ion.



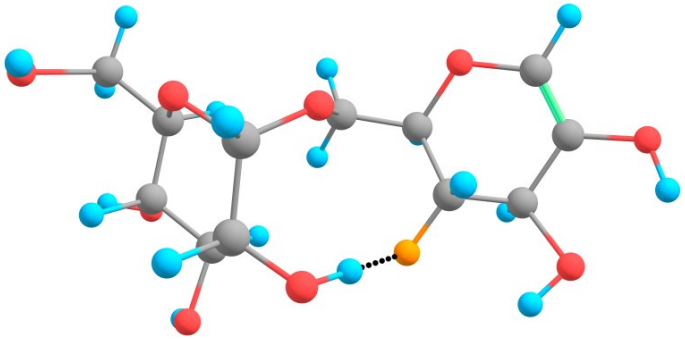
C2 – Initial [1,2]-p.t. radical anion intermediate for the B_2 glycosidic bond cleavage of the primary precursor ion.



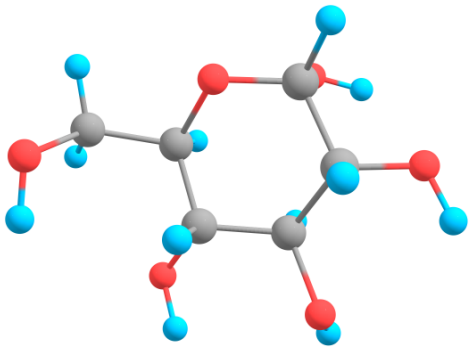
C3 – Second [1,2]-p.t. transition state for the B_2 glycosidic bond cleavage of the primary precursor ion.



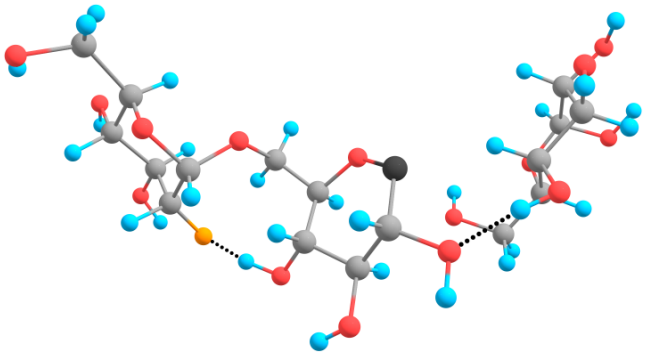
C4 – Ion-Neutral complex for the B_2 glycosidic bond cleavage of the primary precursor ion.



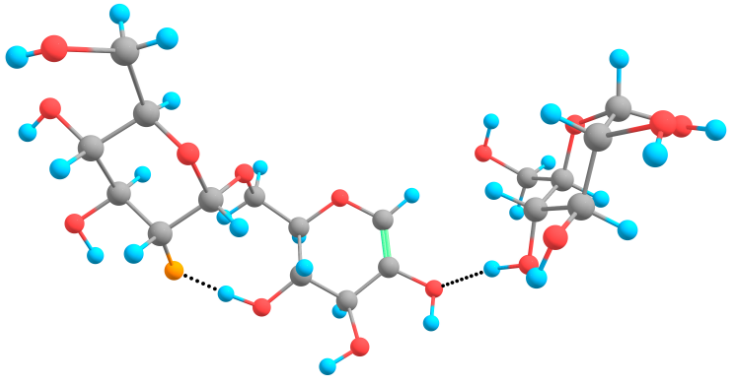
C5(a) – B_2(i4) product ion



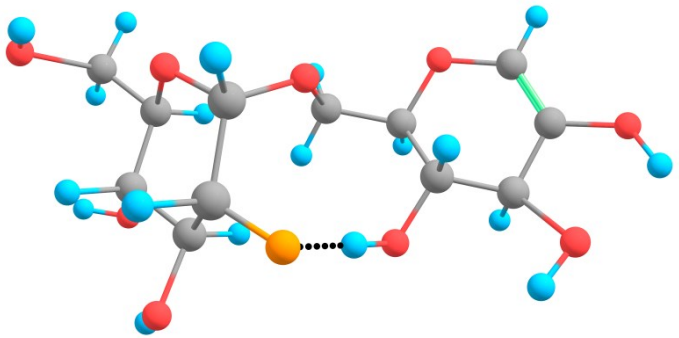
C5(b) – B_2(i4) neutral loss



D1 – Initial [1,2]-p.t. radical anion intermediate for the B_2 glycosidic bond cleavage of the secondary precursor ion.



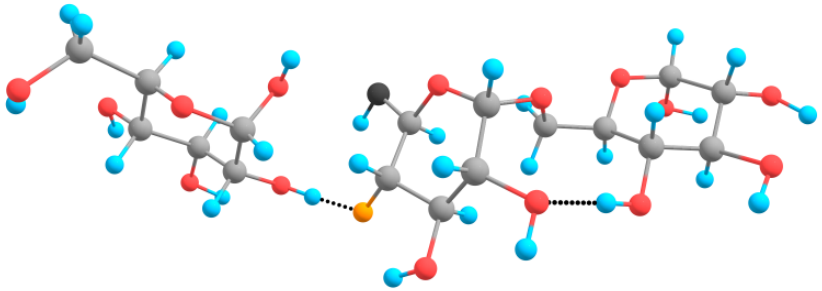
D2 – Ion-Neutral complex for the B_2 glycosidic bond cleavage of the secondary precursor ion.



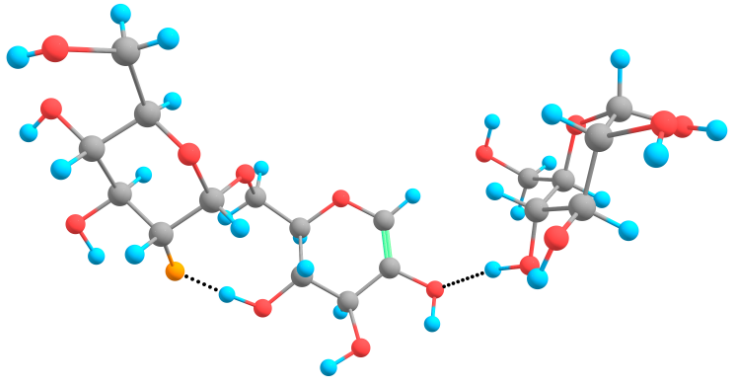
D3(a) – B_2(ii2) product ion

See C5(b)

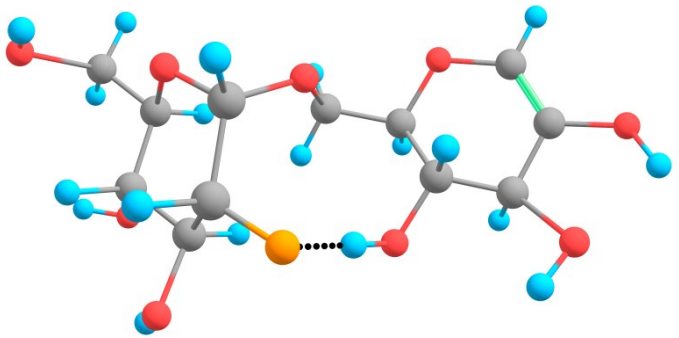
D3(b) – B_2(ii2) neutral loss



E1 – Initial [1,2]-p.t. radical anion intermediate for the Z_2 glycosidic bond cleavage of the primary precursor ion.



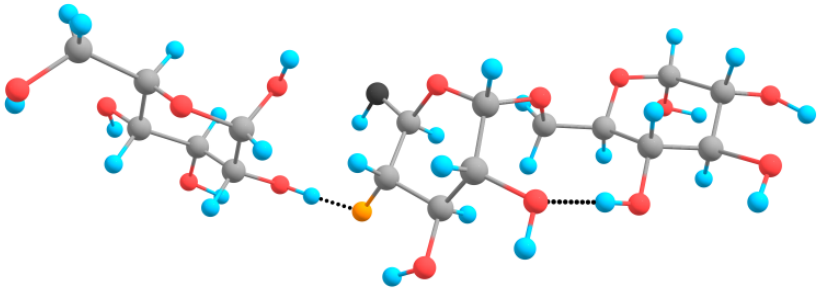
D2 – Ion-Neutral complex for the B_2 glycosidic bond cleavage of the secondary precursor ion.



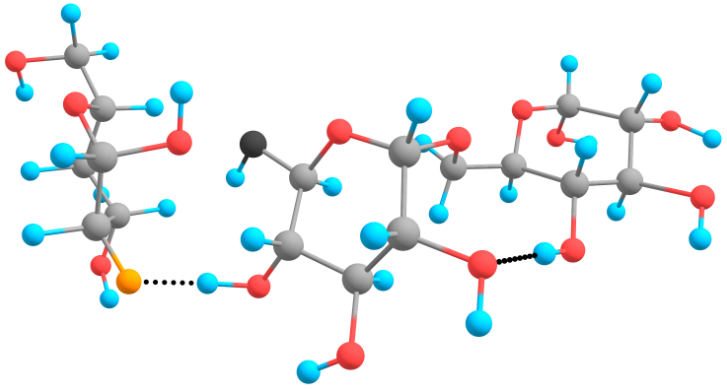
D3(a) – B_2(ii2) product ion

See C5(b)

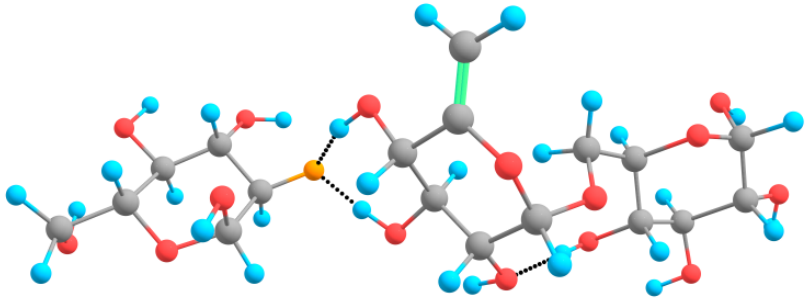
D3(b) – B_2(ii2) neutral loss



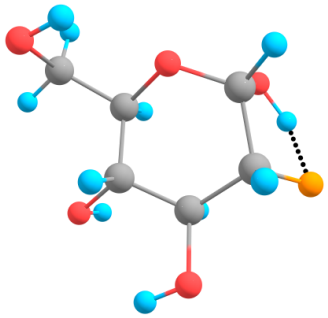
E1 – Initial [1,2]-p.t. radical anion intermediate for the Z_2 glycosidic bond cleavage of the primary precursor ion.



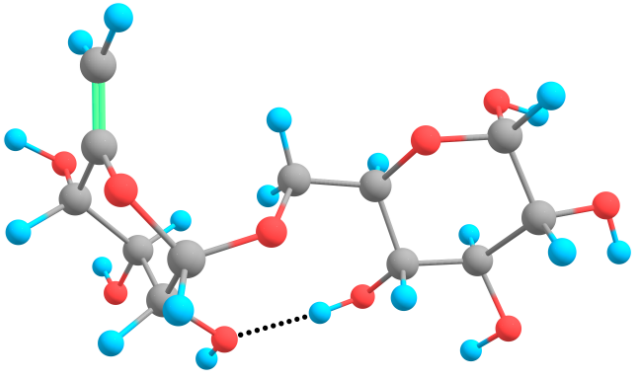
F1 – Initial [1,2]-p.t. radical anion intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



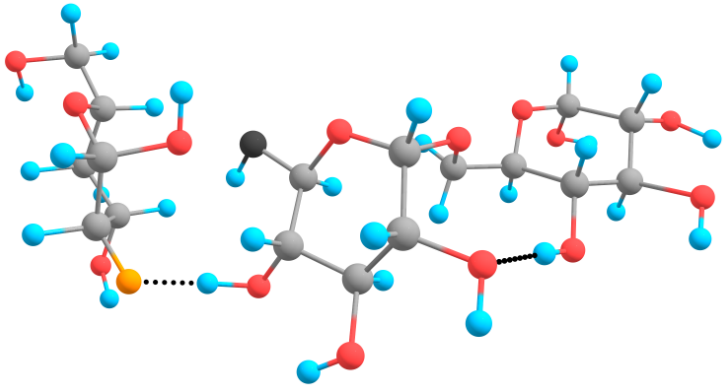
F2 – Ion-Neutral complex intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



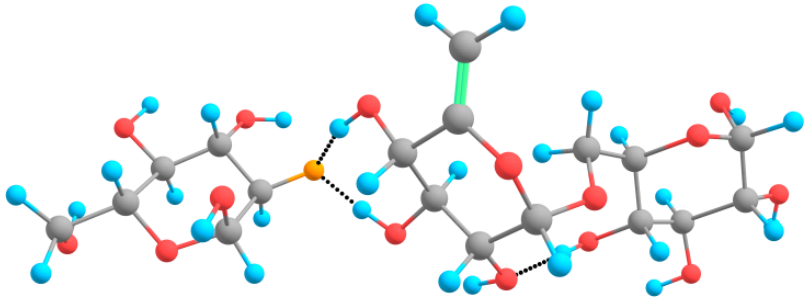
F3(a) – C_1(i2) product ion



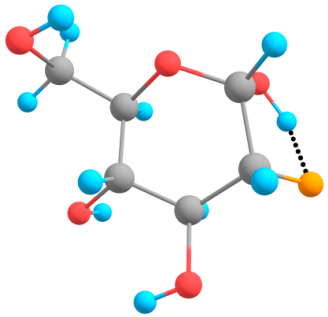
F3(b) – C_1(i2) neutral loss



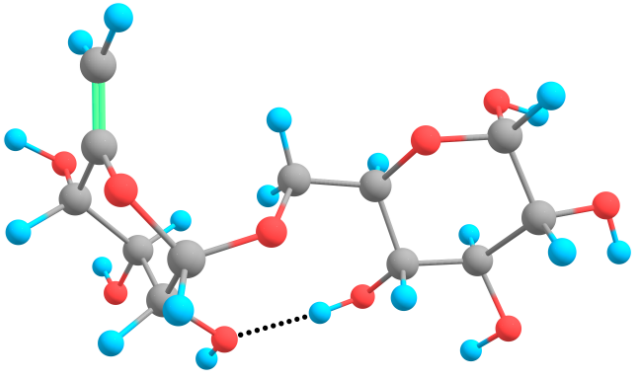
F1 – Initial [1,2]-p.t. radical anion intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



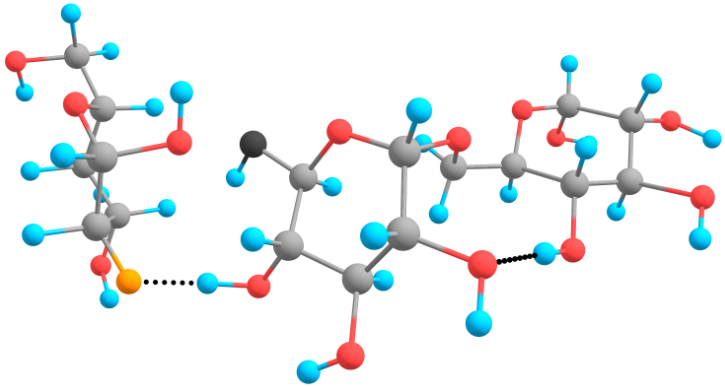
F2 – Ion-Neutral complex intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



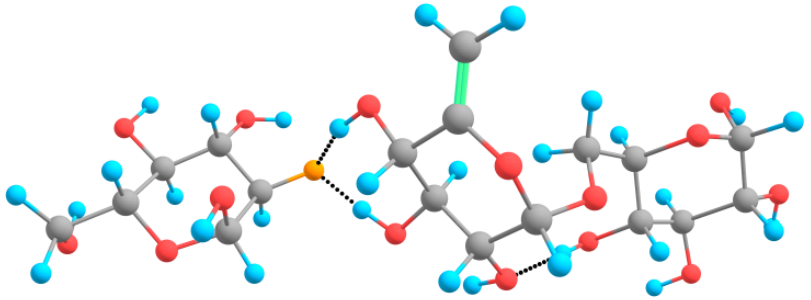
F3(a) – C_1(i2) product ion



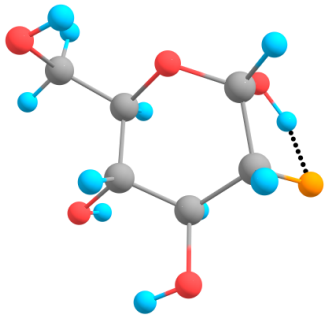
F3(b) – C_1(i2) neutral loss



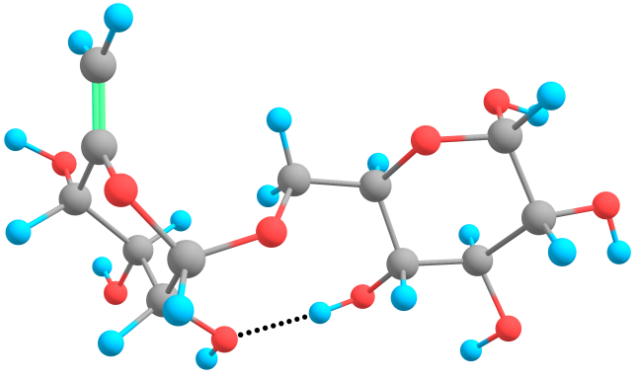
F1 – Initial [1,2]-p.t. radical anion intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



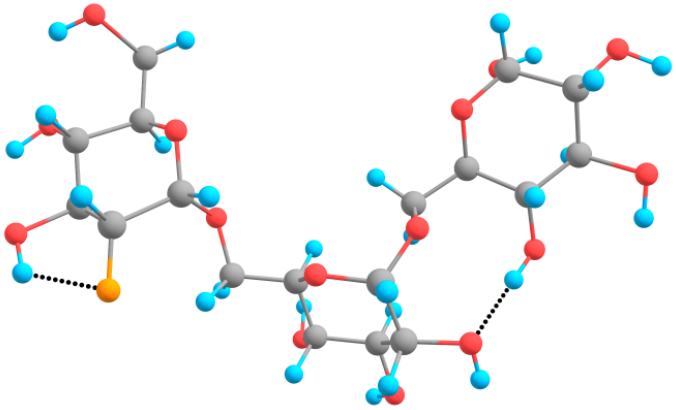
F2 – Ion-Neutral complex intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



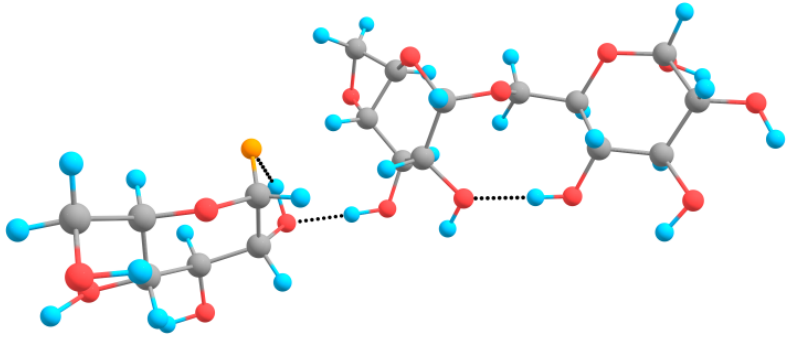
F3(a) – C_1(i2) product ion



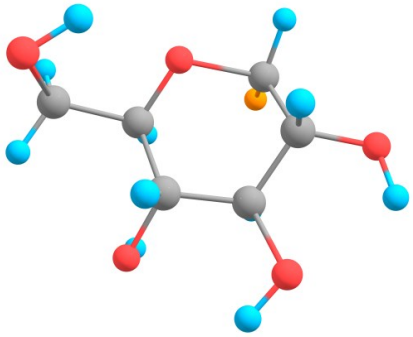
F3(b) – C_1(i2) neutral loss



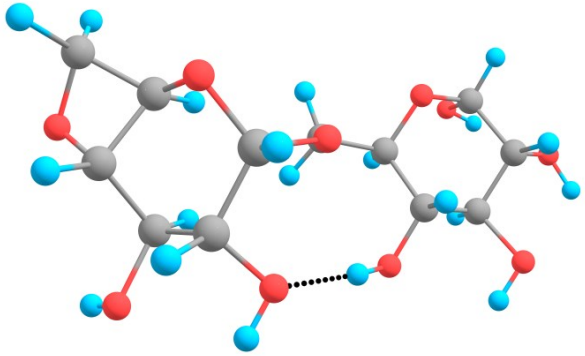
F4 – Flipped-ring intermediate for the C_1 glycosidic bond cleavage of the secondary precursor ion.



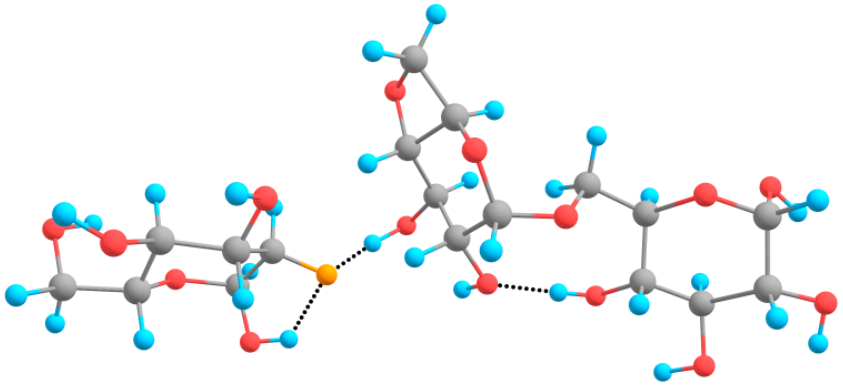
G1 – Charge-direct Ion-Neutral complex for the C_1 glycosidic bond cleavage of the primary precursor ion.



G2(a) – C_1(i1) product ion



G2(b) – C_1(i1) Charge-directed neutral loss



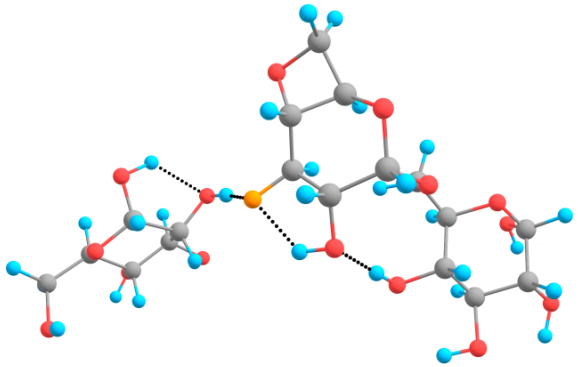
G3 – Charge-direct Ion-Neutral complex for the C_1 glycosidic bond cleavage of the primary precursor ion after the first charge-transfer.

See F3(a)

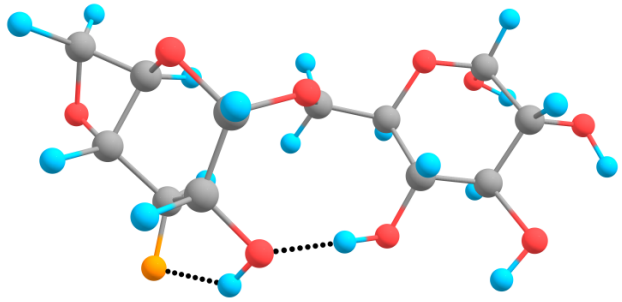
G4(a) – C_1(i2) product ion

See G2(b)

G4(b) – C_1(i2) Charge-directed neutral loss



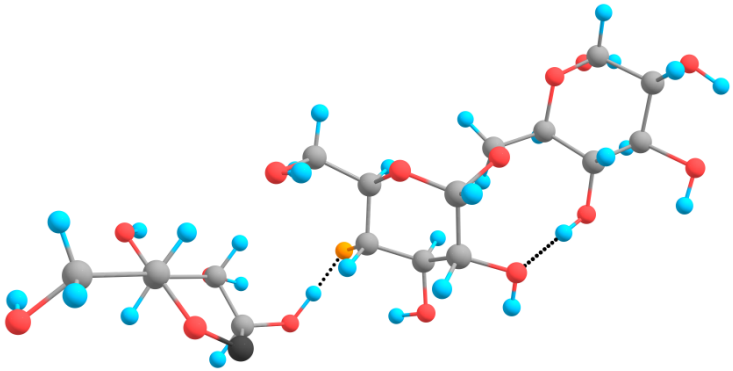
G5 – Charge-direct Ion-Neutral complex for the Z_2 glycosidic bond cleavage of the primary precursor ion after the second charge-transfer.



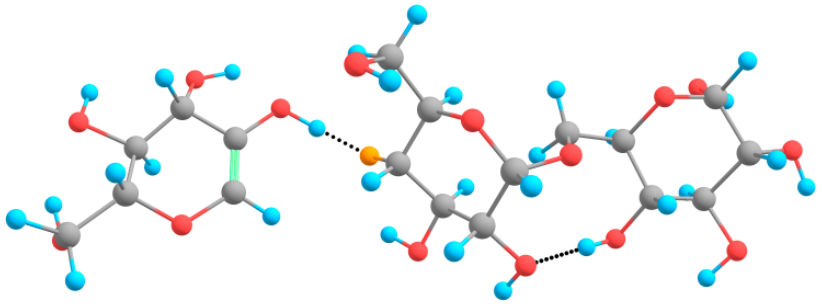
G6(a) – Z_2(ii3) Charge-directed product ion

See C5(b)

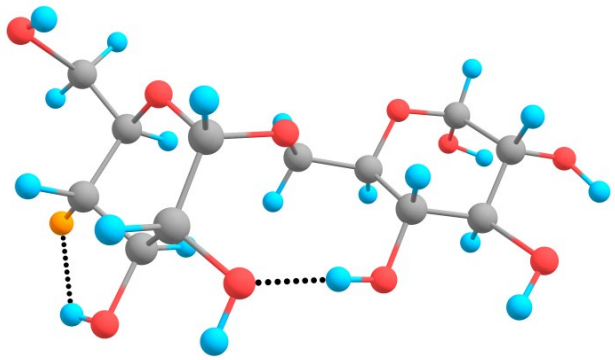
G6(b) – Z_2(ii3) neutral loss



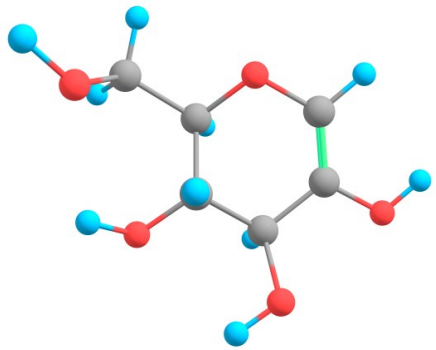
H1 – Initial [1,2]-p.t. radical anion intermediate for the Y_2 glycosidic bond cleavage of the primary precursor ion



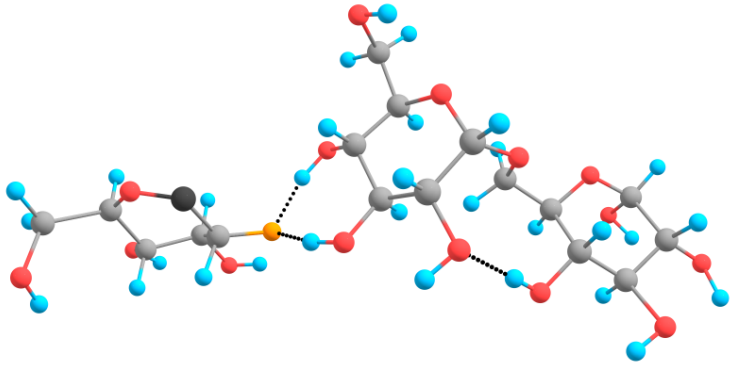
H2 – Ion-neutral complex for the Y_2 glycosidic bond cleavage of the primary precursor ion



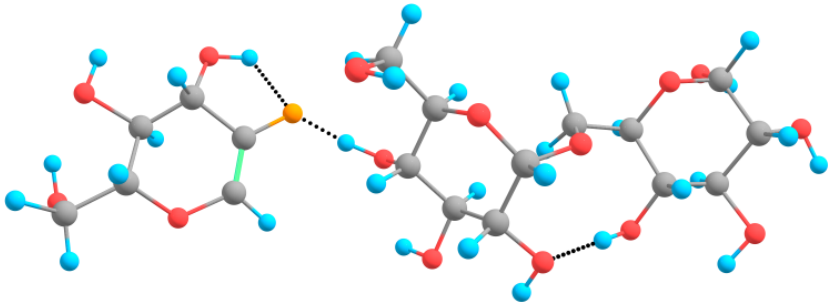
H3(a) – Y_2(ii4) product ion



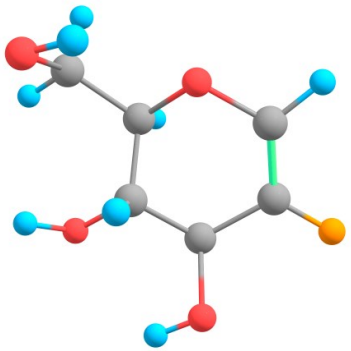
H3(b) – Y_2(ii4) neutral loss



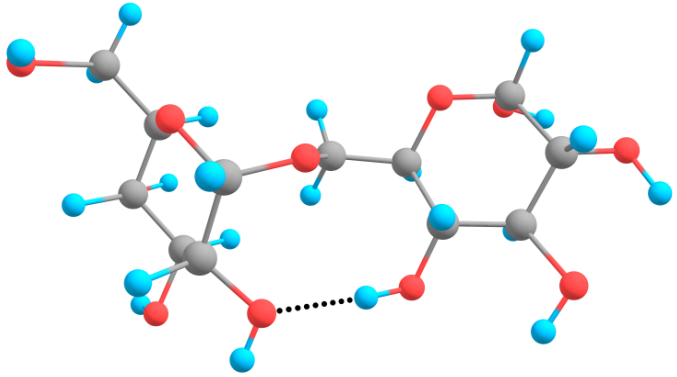
I1 – Initial [1,2]-p.t. radical anion intermediate for the B_1 glycosidic bond cleavage of the secondary precursor ion



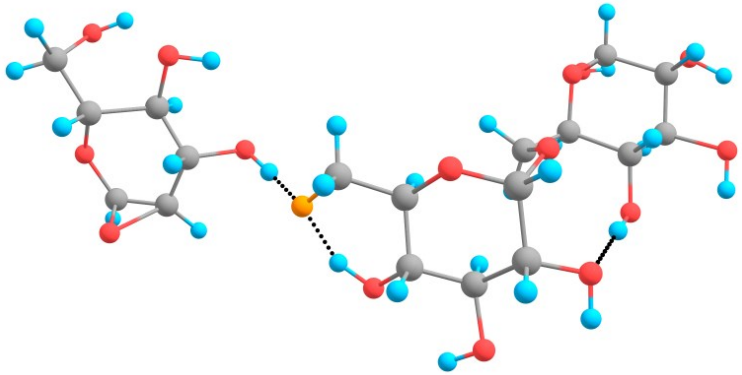
I2 – Ion-Neutral complex for the B_1 glycosidic bond cleavage of the secondary precursor ion



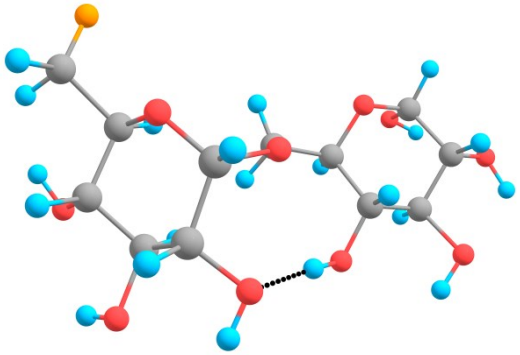
I3(a) – B_1(i2) product ion



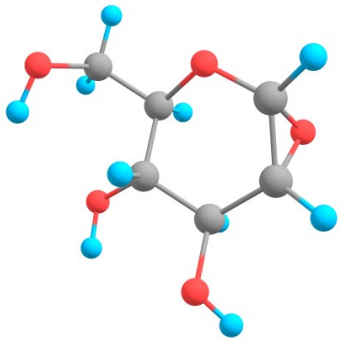
I3(b) – B_1(i2) neutral loss



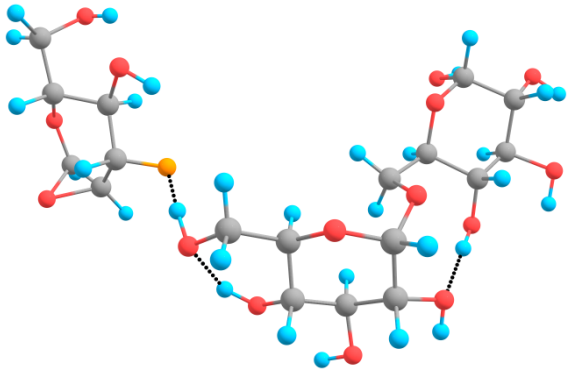
J1 – Charge-direct Ion-Neutral complex for the C_2 glycosidic bond cleavage of the secondary precursor ion



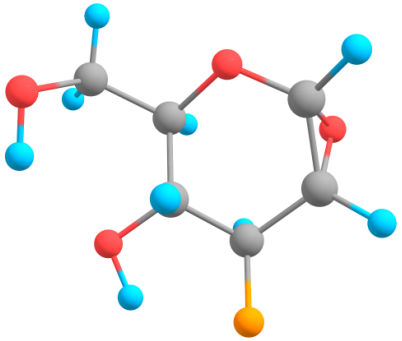
J2(a) – C_2(ii6) product ion



J2(b) – C_2(ii6) Charge-direct neutral loss



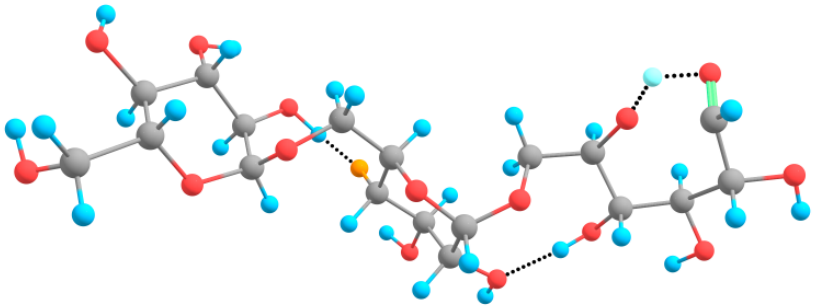
J3 – Charge-direct Ion-Neutral complex for the B_1 glycosidic bond cleavage of the secondary precursor ion after charge transfer.



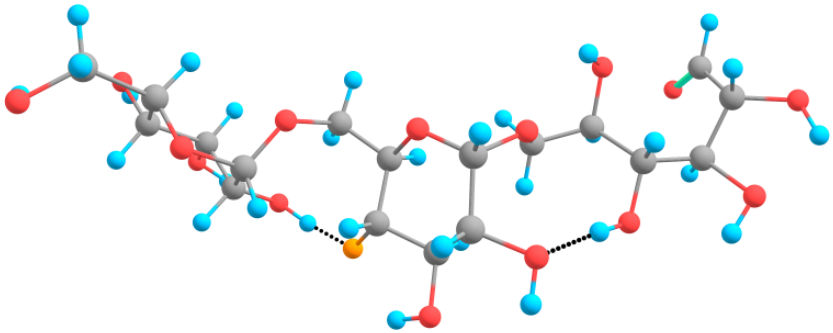
J4(a) – B_1(i3) Charge-directed product ion

See I3(b)

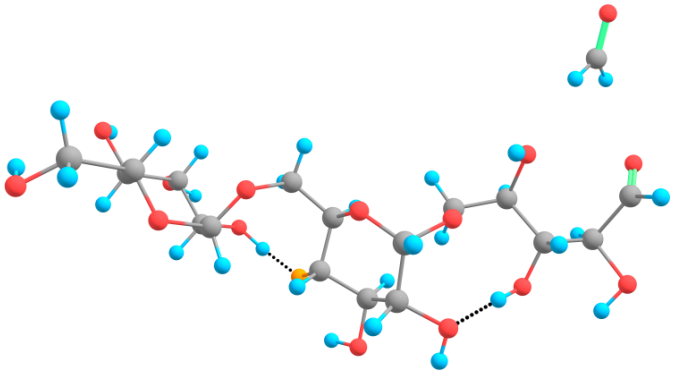
J4(b) – B_1(i3) neutral loss



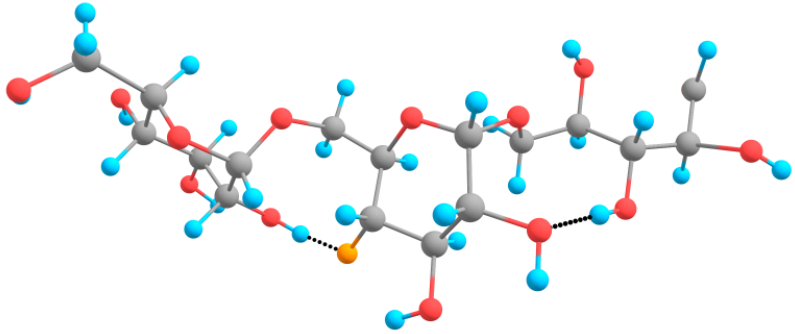
K1 – Ring-opening transition state



K2 – Ring-opened primary precursor ion

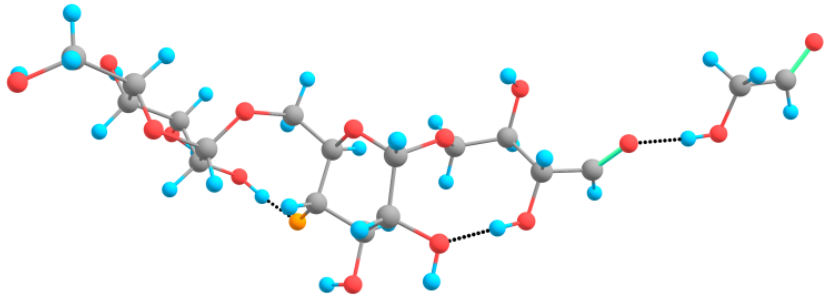


L1 – Ion-Neutral complex for CH_2O from the ring-opened primary precursor ion.

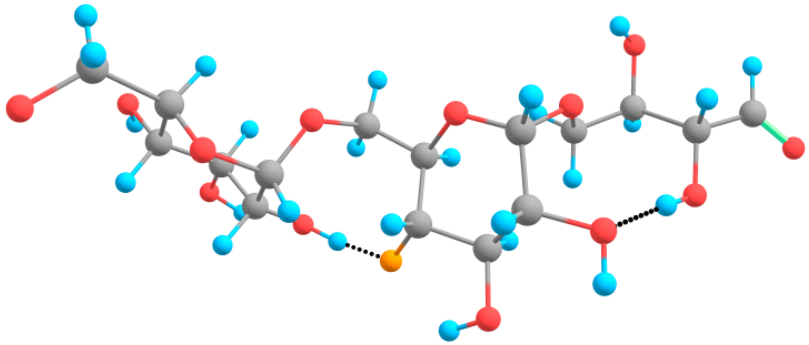


L2(a) – $^{0,1}A_3$(ii4) product ion

L2(b) – $^{0,1}A_3$(ii4) neutral loss

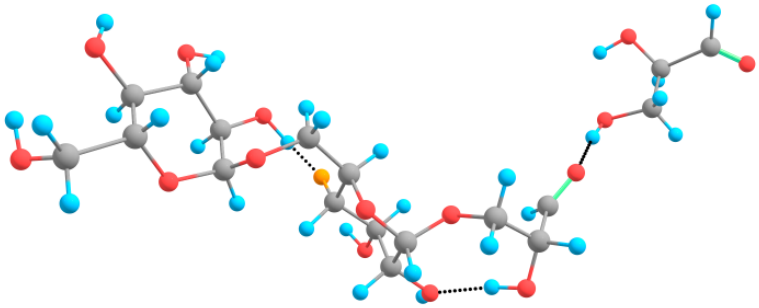


M1 – Ion-Neutral complex for $C_2H_4O_2$ from the ring-opened primary precursor ion.

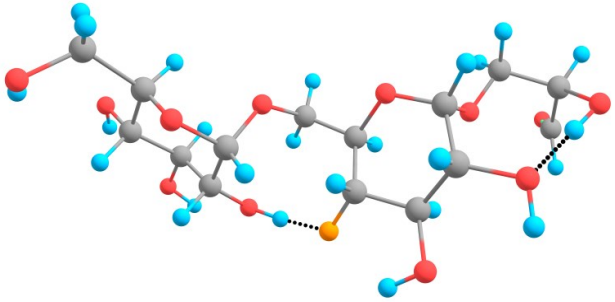


M2(a) – $^{0,2}A_3$(ii4) product ion

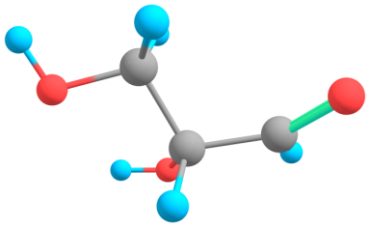
M2(b) – $^{0,2}A_3$(ii4) neutral loss



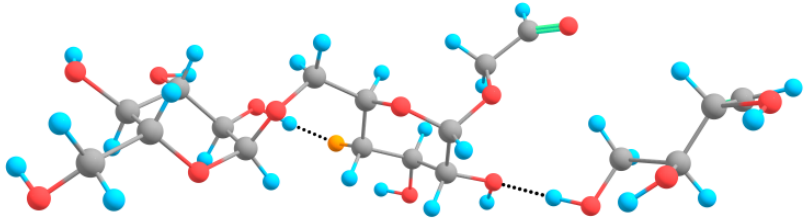
N1 – Ion-Neutral complex for $C_3H_6O_3$ from the ring-opened primary precursor ion.



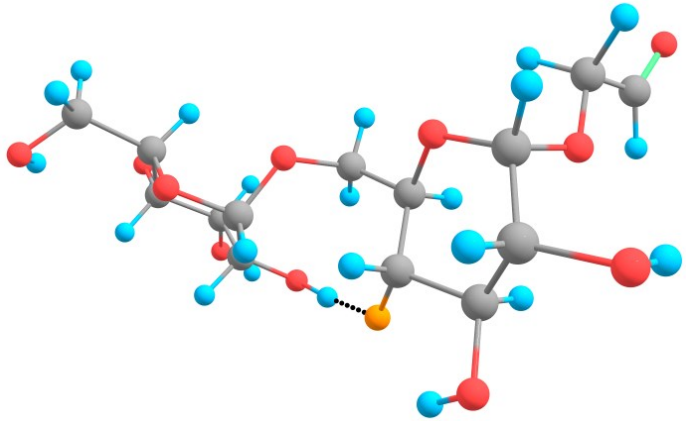
N2(a) – $^{0,3}A_3$(ii4) product ion



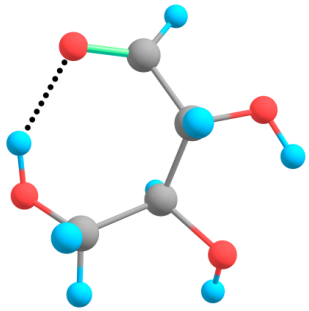
N2(b) – $^{0,3}A_3$(ii4) neutral loss



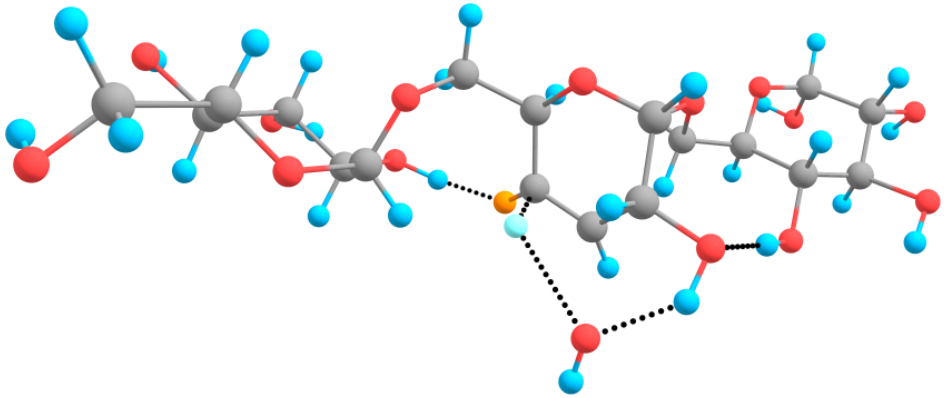
O1 – Ion-Neutral complex for $C_4H_8O_4$ from the ring-opened primary precursor ion.



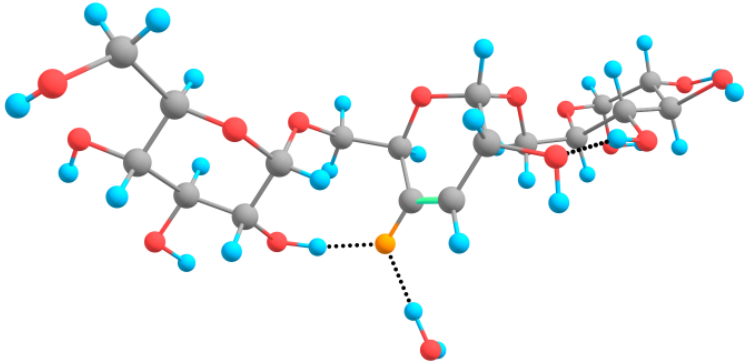
O2(a) – $^{0,4}A_3$(ii4) product ion



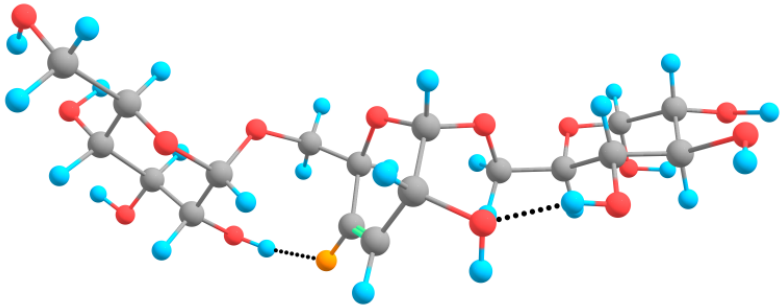
O2(b) – $^{0,4}A_3$(ii4) neutral loss



P1 – Transition state for the concerted [1,3]-p.t. and elminiation of water from the primary precursor ion.

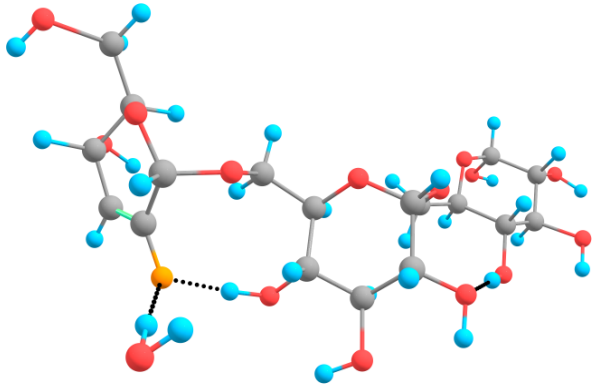


P2 – Ion-Neutral complex for water loss from the primary precursor ion.

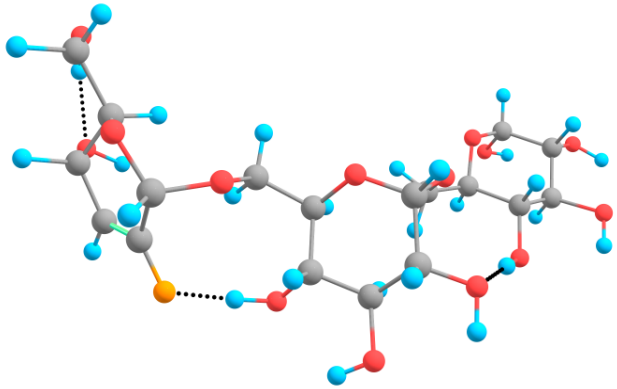


P2(a) – Water loss product ion from the primary precursor ion

P2(b) – Water neutral loss



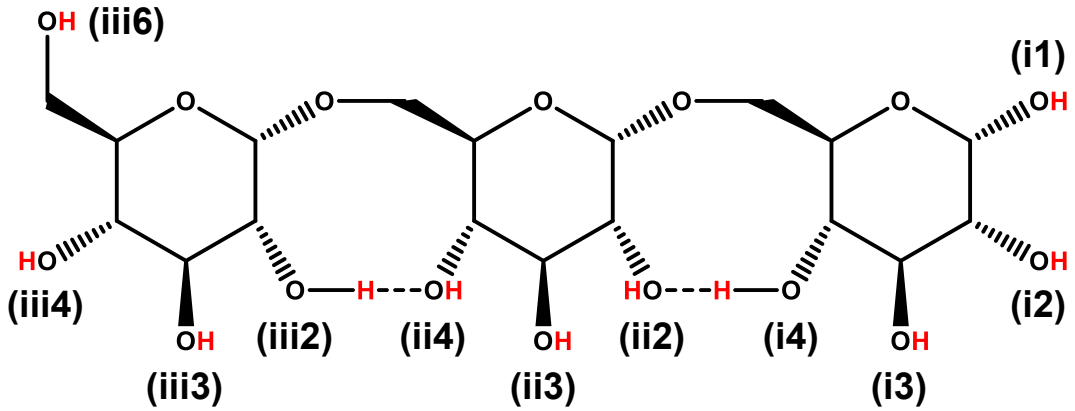
Q1 – Ion-Neutral complex for water loss from the secondary precursor ion.



Q2(a) – Water loss product ion from the secondary precursor ion

Annexe 4. Calculated CCS values for the deprotonated trisaccharides.

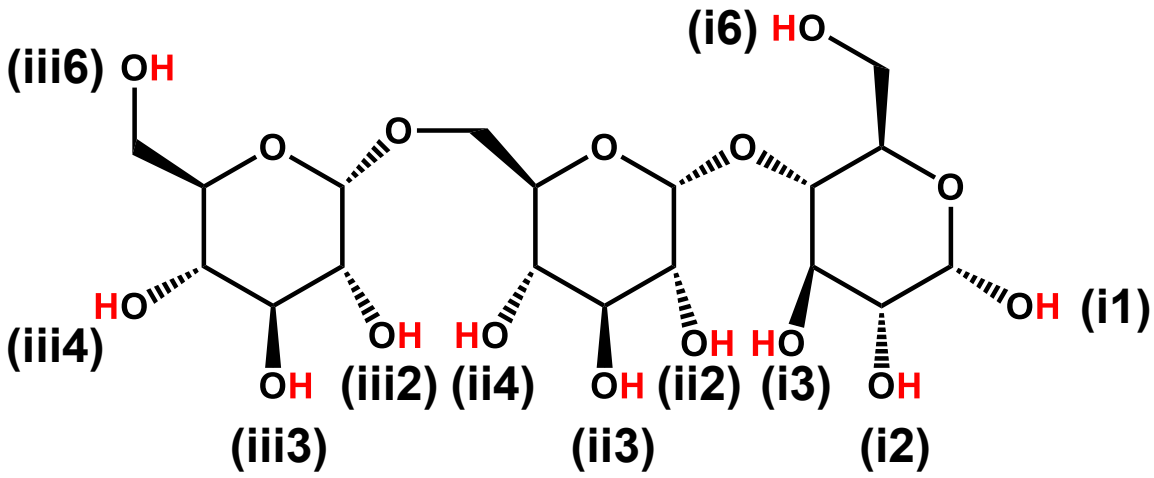
A4.1 Collisional cross sections for deprotonated isomaltotriose (ISO).



Structure	Relative Internal Energy† (eV)	Calculated CCS (± 2 Å²)
ISO-(i1)	0.90	153.(6)
ISO-(i2)	0.14	154.(4)
ISO-(i3)	0.66	153.(6)
ISO-(i4)	0.57	152.(0)
ISO-(ii2)	0.61	152.(7)
ISO-(ii3)	0.41	151.(7)
ISO-(ii4)	**0.00**	**152.(1)**
ISO-(iii2)	0.06	151.(3)
ISO-(iii3)	0.32	151.(8)
ISO-(iii4)	0.43	152.(7)
ISO-(iii6)	0.71	149.(6)
	Average CCS (Å²)	**152 ± 3**

†The energy values were normalized to the lowest energy structure for isomaltotriose (red)

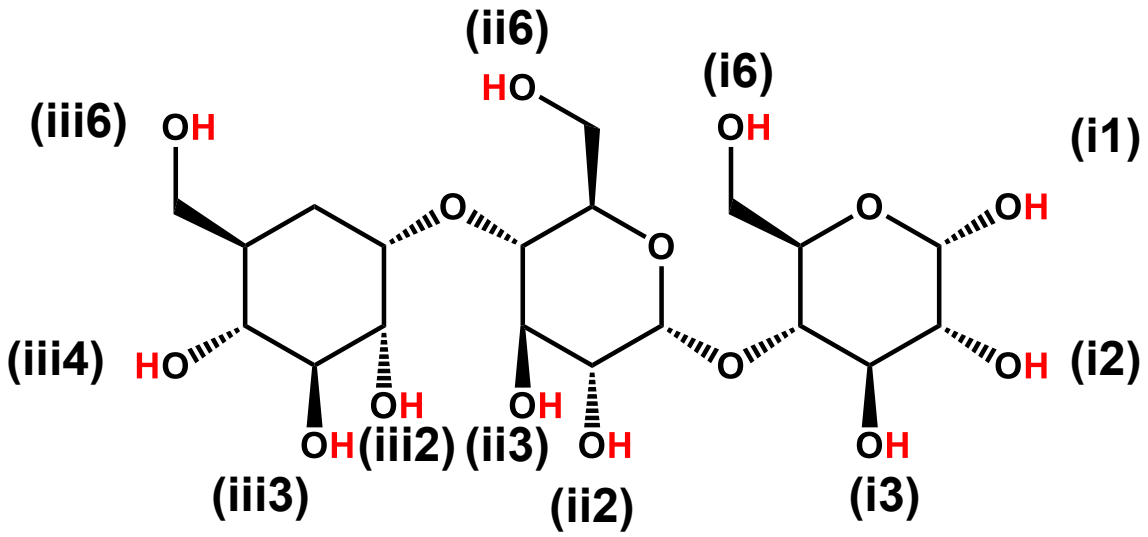
A4.2 Collisional cross sections for deprotonated panose (PAN)



Structure	Relative Internal Energy† (eV)	Calculated CCS (± 2 Å^2)
PAN-(i1)	1.20	152.(8)
PAN-(i2)	2.28	154.(2)
PAN-(i3)	0.76	151.(4)
PAN-(i6)	1.72	153.(1)
PAN-(ii2)	0.78	151.(0)
PAN-(ii3)	0.46	149.(4)
PAN-(ii4)	0.01	149.(1)
PAN-(iii2)	**0.00**	**149.(3)**
PAN-(iii3)	0.46	149.(5)
PAN-(iii4)	1.39	151.(7)
PAN-(iii6)	1.47	151.(7)
	Average CCS (Å^2)	**151 ± 3**

†The energy values were normalized to the lowest energy structure for panose (red)

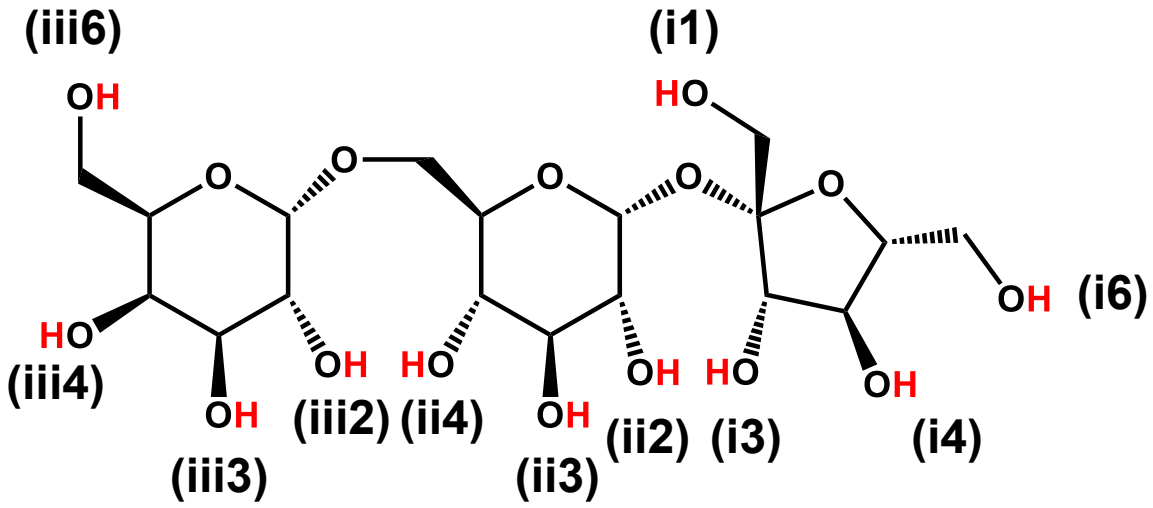
A4.3 Collisional cross sections for deprotonated maltotriose (MAL)



Structure	Relative Internal Energy† (eV)	Calculated CCS (± 2 Å^2)
MAL-(i1)	0.71	149.(6)
MAL-(i2)	1.73	151.(2)
MAL-(i3)	0.77	150.(0)
MAL-(i6)	0.48	140.(5)
MAL-(ii2)	0.71	148.(5)
MAL-(ii3)	0.22	150.(5)
MAL-(ii6)	0.98	150.(3)
MAL-(iii2)	0.25	149.(1)
MAL-(iii3)	0.26	148.(4)
MAL-(iii4)	**0.00**	**147.(4)**
MAL-(iii6)	0.45	145.(7)
	Average CCS (Å^2)	**148 ± 3**

†The energy values were normalized to the lowest energy structure for maltotriose (red).

A4.4 Collisional cross sections for deprotonated raffinose (RAF)



Structure	Relative Internal Energy† (eV)	Calculated CCS (± 2 Å^2)
RAF-(i1)	1.37	155.(4)
RAF-(i3)	1.42	152.(2)
RAF-(i4)	0.35	153.(4)
RAF-(ii2)	1.00	151.(5)
RAF-(ii3)	0.46	152.(9)
RAF-(ii4a)	0.07	153.(2)
RAF-(ii4b)	**0.00**	**147.(4)**
RAF-(iii2)	0.05	153.(6)
RAF-(iii3)	0.41	155.(9)
RAF-(iii4)	1.22	158.(1)
RAF-(iii6)	1.42	155.(8)
	Average CCS (Å^2)	**154 ± 3**

†The energy values were normalized to the lowest energy structure for raffinose (red)