ARTICLE

Evidence of Gas-phase Pyranose-to-furanose Isomerization in Protonated Peptidoglycans

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Shanshan Guan[a,b] and Benjamin J. Bythell[a,b]*

Peptidoglycans are diverse co- and post-translational modifications of key importance in myriad biological processes. Mass spectrometry is employed to infer their biomolecular sequences and stereochemisties, but little is known about the critical gas-phase dissociation processes involved. Here, using tandem mass spectrometry (MS/MS and MS^n), isotopic labelling and high-level simulations, we identify and characterize a facile isomerization reaction that produces furanose N-acetylated ions. This reaction occurs for both O- and N-linked peptidoglycans irrespective of glycosidic linkage stereochemistry (α/β). Dissociation of the glycosidic and other bonds thus occur from the furanose isomer critically altering the reaction feasibility and product ion structures.

1. Introduction

Protein glycosylation covalently attaches oligosaccharides to amino acid side-chains is one the most abundant and diverse co- and post-translational modification classes of proteins in eukaryotes.[1] It is estimated that more than 50% of human proteins are glycosylated.[2] The two most common types of glycosylation are N- and O-, classified according to the heteroatom involved in the glycan-protein (amino acid) linkage. N-linked glycans attach the carbohydrate via an amide group on amino acid sidechain (for example asparagine, Asn) whereas O-linked glycans attach the glycan to a side-chain hydroxyl group (generally serine, Ser, or threonine, Thr).[3] Mammalian N-linked glycans commonly have a core structure that combines mannose and N-acetyl glucose residues, $(Man)_3(GlcNAc)_2$-N-, attached to the Asn-X-Ser/Thr(Cys) motif, where X is any amino acid other than proline, and cysteine, Cys, is rare.[4] The structural diversity of N-glycans is built onto this core structure with addition of further carbohydrate residues. Conversely, O-linked glycosylation shares neither a consensus peptide sequence to indicate the likely potential glycosylation site(s) nor a conserved core structure in the glycan part.[5] In humans, the most common O-linked glycosylation begin with alpha N-acetyl galactose, α-GalNAc[6–8], while β-GlcNAc-linked glycosylation of Ser/Thr is a reversible post-translational modification of nuclear and cytoplasmic proteins.[9] β-GlcNAc-linked glycosylation is unusual in that the β-GlcNAc is not typically glycosylated (extended) further or modified.[10,11] This type of glycosylation plays a significant role in cellular signal transduction and has shown a complex interplay with phosphorylation post-translational modifications.[12–14] Protein glycosylation is involved in diverse metabolic processes such as immune response, protein secretion and transportation, and also affects the properties of the attached protein.[15–23] Protein glycosylation has a significant impact in the pathogenesis of numerous diseases resulting in glycoproteins being identified as disease biomarkers.[24–32] The variations of the glycoform on one glycosylation site among different samples and/or the changes in the glycosylation sites are hypothesized to be further disease biomarkers. Consequently profiling of site-specific glycans has gained increased interest recently.[8,33–35]

Mass spectrometry (MS) is one of the principal techniques for the analysis of glycoproteins. The heterogeneity of glycoforms presents a significant challenge for analyses.[16,36] MS-based analyses of glycoproteins typically are preceded by enzymatic proteolysis.[37–39] There are two general glycoprotein digestion strategies commonly employed: (1) Removal of the oligosaccharide chain(s) from the protein by a glycosidase.[37–40] This arguably reduces the complexity of data interpretation in the following MS-based analysis by independently analyzing the carbohydrate (glycan) and protein components. Regrettably, this is done at the cost of total loss of glycosylation site information;[41–44] (2) The alternative approach uses a proteolytic enzyme to cleave amide bonds within the protein without loss of post translational modifications, resulting in a mixture of peptides and glycopeptides.[40] i.e., retaining peptidoglycan site-information provided the glycopeptide can be sequenced.[38,44] This approach is reliant on being able to successfully obtain the entire glycopeptide sequence from tandem mass spectrometry.

Irrespective of the digestion strategy employed, tandem mass spectrometry (MS/MS) methods are the key to subsequent sequencing of the glycopeptides or glycans. Collision-induced dissociation (CID) is commonly utilized for glycopeptide analysis. CID fragmentation is dominated by glycosidic bond cleavages forming B- and Y-type fragments at

[a] Department of Chemistry and Biochemistry, Ohio University, 307 The Chemistry Building, Athens, OH 45701.
[b] Department of Chemistry and Biochemistry, University of Missouri, 1 University Blvd, St. Louis, MO 63121.
E-mail: bythell@ohio.edu

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

the glycosidic bond.[38–40,44,45] The detection of charged glycan fragments of specific *m/z* values (*m/z* 204.09 for HexNAc; *m/z* 163.06 for hexose, Hex), as diagnostic for the presence of glycopeptides.[5] Identification of the glycans by the currently available algorithms is largely reliant on m/z differences between the fragment ion series.[46–48] The heterogeneity of the glycoforms, the large number of stereo- and linkage isomers coupled with the limited understanding of the dissociation chemistry[32,49] in protonated glycopeptides all limit our ability to extract structural information from the tandem mass spectra. Advances in these areas would substantially improve confidence in structural assignments[32] and thus subsequent experimental design and inference.

In this study, we utilize tandem mass spectrometry, deuterium labelling, and computational chemistry to elucidate the diagnostic fragmentation chemistry of three peptidoglycans: GlcNAc-β-1-Asn, GalNAc-α-1-Ser, and GlcNAc-β-1-Ser. These analytes represent the core structures of N-linked glycosylation, mucin-type glycosylation, and O-GlcNAcylation, respectively.

2. Methods Section

Chemicals.

GlcNAc-β-1-Asn, GalNAc-α-1-Ser, and GlcNAc-β-1-Ser were purchased from Biosynth Carbosynth (San Diego, CA). Deuterium oxide was purchased from Cambridge Isotope Laboratories, Inc (Tewksbury, MA). HPLC-grade Acetonitrile and water were purchased from Sigma-Aldrich (St. Louis, MO).

Experimental methods.

Experimental work was carried out using a Q-Exactive Orbitrap Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA). Tandem mass spectra for $[GlcNAc\text{-}\beta\text{-}1\text{-}Asn+H]^+$, $[GalNAc\text{-}\alpha\text{-}1\text{-}Ser+H]^+$, and $[GlcNAc\text{-}\beta\text{-}1\text{-}Ser]^+$ were obtained by mass selecting the appropriate ion with the quadruple, collision-induced dissociation (CID) in the HCD cell followed by mass-to-charge analysis by the orbitrap mass analyzer. Ionization was by electrospray with the samples infused into the instrument in ~2μM acetonitrile/water/formic acid (50/50/0.1%) solutions at a flow rate of 5 μl min^{-1}. Nitrogen was used as nebulizing, drying, and collision gas. Spectra were collected at mass resolution of 140,000 (FWHM).

Hydrogen–deuterium exchange (HDX) is a chemical reaction in which a covalently bonded hydrogen atom is replaced by a deuterium atom, or vice versa. In D_2O exchangeable protons such as those in hydroxyl or amide group exchanged with deuterons from the solvent. Because the mass of a deuterium is 1 u heavier than a proton, the integer mass shift of the analyte ions provides the number of exchanged protons. By comparing the fragmentation spectra of unlabelled and deuterium labelled samples, the mass shifts of the fragment ions provide insight into both the product ion structures and their mechanism(s) of gas-phase formation. The model analyte systems were dissolved in D_2O for 10 min at room temperature. The resulting solutions were then further diluted in acetonitrile/D_2O (50/50) to a final concentration of ~2μM prior to tandem mass spectral analysis.

Theoretical Methods.

Simulations were performed to enable characterization of the potential energy surface of each protonated analyte. Initial candidate structures for protonated and multiple potential isomerized ions were generated with Fafoom[50,51] via a genetic algorithm utilizing the MMFF94 force field[52–56]. Fafoom systematically alters dihedral and other angles including sampling ring structure types (chair, boat, and skew) enabling thorough interrogation of the many structural possibilities.[49,57–60] Geometry optimizations of the resulting candidate conformations were performed at the HF/3-21g, B3LYP/6-31G(d), B3LYP/6-31+G(d,p)[61,62] levels and M06-2X/6-31+(d,p)[63,64] levels of theory. Degenerate structures were removed at each stage with the non-degenerate structures utilized as the starting points of subsequent refinement. All density functional calculations of minima, transition structures, product ions and neutrals were performed with the Gaussian 09 suite of programs.[65] Multiple transition structures (TSs) were calculated from multiple precursor ion structures for each potential fragmentation pathway. Vibrational analyses were performed for minima and TSs (all real frequency for minima and 1 imaginary frequency for TSs). The zero-point energy (ZPE) was added to the electronic energy (E_{el}, 0 K) to improve the accuracy of the potential energy surface generated (ΔE_{el+ZPE}, 0 K). In addition, the corresponding standard entropy (ΔH_{298K}), Gibbs free energy (ΔG_{298K}), and entropy (ΔS_{298K}) corrections at 298K were calculated. The reaction pathway through each TS was determined by intrinsic reaction coordinate (IRC) calculations with up to 18 steps in each direction. The terminating points of these calculations (one on product-side, one on reactant-side) were then optimized further to the minima that connect via the TS. Single point calculations of key minima and TSs were performed with 6-311+(2d,p) basis sets for comparison.

Rice–Ramsperger–Kassel–Markus (RRKM) calculations were performed using the energetics, vibrational frequencies, and rotational constants derived from the modeling to approximate the time scale of the fragmentation reactions[66,67]. The Beyer–Swinehart direct count algorithm[68] is used for rotational–vibrational treatment of both the reactant and the transition structure.

3. Results and discussion

Tandem Mass Spectrometry.

The MS/MS spectrum of singly protonated GlcNAc-β-1-Asn (Figure 1a) shows major peaks at *m/z* 204 and *m/z* 133. These are assigned as $C_8H_{14}N_1O_5^+$, the protonated glycan B_1 fragment and $C_4H_9N_2O_3^+$, protonated asparagine, $[Asn+H]^+$. These peaks nominally result from direct dissociation of the glycosidic linkage with the ionizing proton retained either by the glycan fragment to produce the B_1 ion or the asparagine to form the $[Asn+H]^+$ ion. The third abundant peak, *m/z* 126, $C_6H_8N_1O_2^+$, corresponds to neutral loss of $C_2H_6O_3$ (or potentially H_2O and

$C_2H_4O_2$ consecutive losses) from the B_1 ion. The less abundant peak with *m/z* 186, $C_8H_{12}N_1O_4^+$, corresponds to loss of water from the B_1 ion. Similar fragmentation patterns were observed on the MS/MS spectra of singly protonated GlcNAc-β-1-Ser and GalNAc-α-1-Ser (Figure 1b and 1c). Both protonated O-glycosylated-Ser spectra indicate that the major bond cleavage occurs at the glycosidic bond too. Substantial peaks at *m/z* 186, $C_8H_{12}N_1O_4^+$, and *m/z* 126, $C_6H_8N_1O_2^+$, support the occurrence of consecutive dissociation reactions. These peaks differ markedly in abundance to those present in the [GlcNAc-β-1-Asn+H]$^+$ spectrum (Figure 1a).

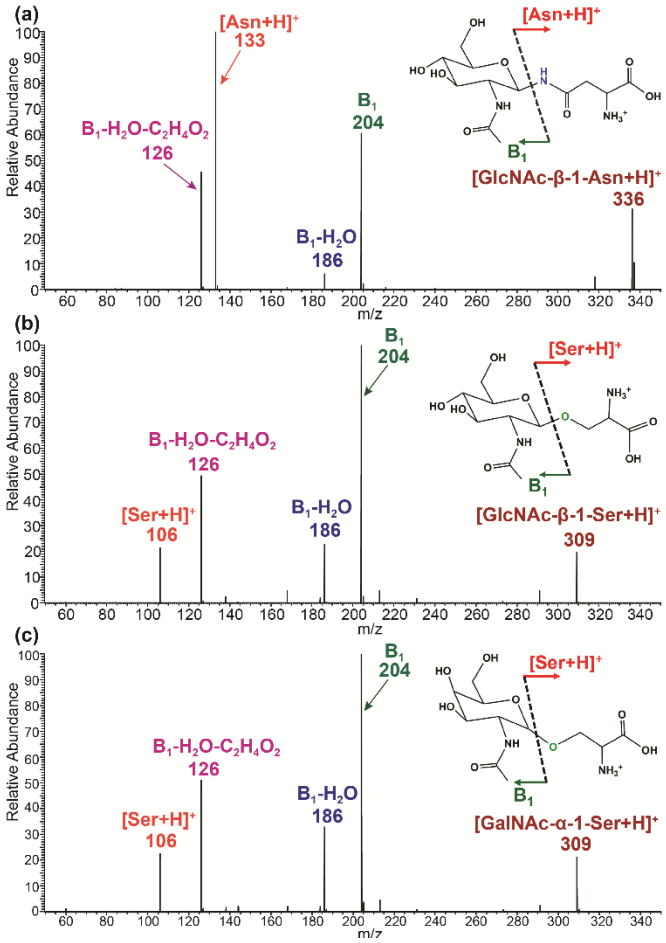


Figure 1. Tandem mass spectra of (a) GlcNac-β-1-Asn, (b) GlcNac-β-1-Ser, and (c) GalNac-α-1-Ser.

Hydrogen-to-deuterium exchange experiments were performed to assess the contributions of C-H versus hydroxyl and N-H protons (Figure S1). Full deuteration of the OH and NH sites in singly charged GlcNAc-β-1-Asn analyte results in a 9u increase in *m/z* corresponding to replacement of 8 OH/NH hydrogens with deuterons and an ionizing D^+. Shifts in *m/z* of the dissociation products were observed as a function of their structures. The B_1 peak shifts 4 u to *m/z* 208, $C_8D_4H_{10}N_1O_5^+$. Two distinct peaks are produced for the [deuterated Asn+D]$^+$ peak; 5u and 6u shifts, corresponding to $C_4D_5H_4N_2O_3^+$ and $C_4D_6H_3N_2O_3^+$. Loss of HDO from the B_1 ion produces a B_1-HDO peak at *m/z* 189. Lastly, a 2u shift (*m/z* 126 to 128) was observed for the B_1-$C_2H_6O_3$ peak (Figure S1a).

For the O-linked analytes we observe exchange of 8 protons to deuterons (7 OH/NH hydrogens and 1 D^+). MS/MS spectra of the resulting [GlcNAc-β-1-Ser-7H+8D]$^+$ ions reveal the mass shifts for the B_1 ion population and its fragments. i.e., Consistent with either forming the same B_1 ion structure(s) derived from GlcNAc (or isomers which dissociate to coincidentally identical elemental compositions). The deuterated serine displays two peaks, [Ser-4H+4D]$^+$ at *m/z* 110 and [Ser-5H+5D]$^+$ at *m/z* 111 (Figure S1b). The GalNAc-α-Ser spectra also show similar mass shift patterns for most peaks. The clear exception is the B_1-water population which splits into two distinct peaks corresponding to B_1-HDO and B_1-D_2O (Figure S1c).

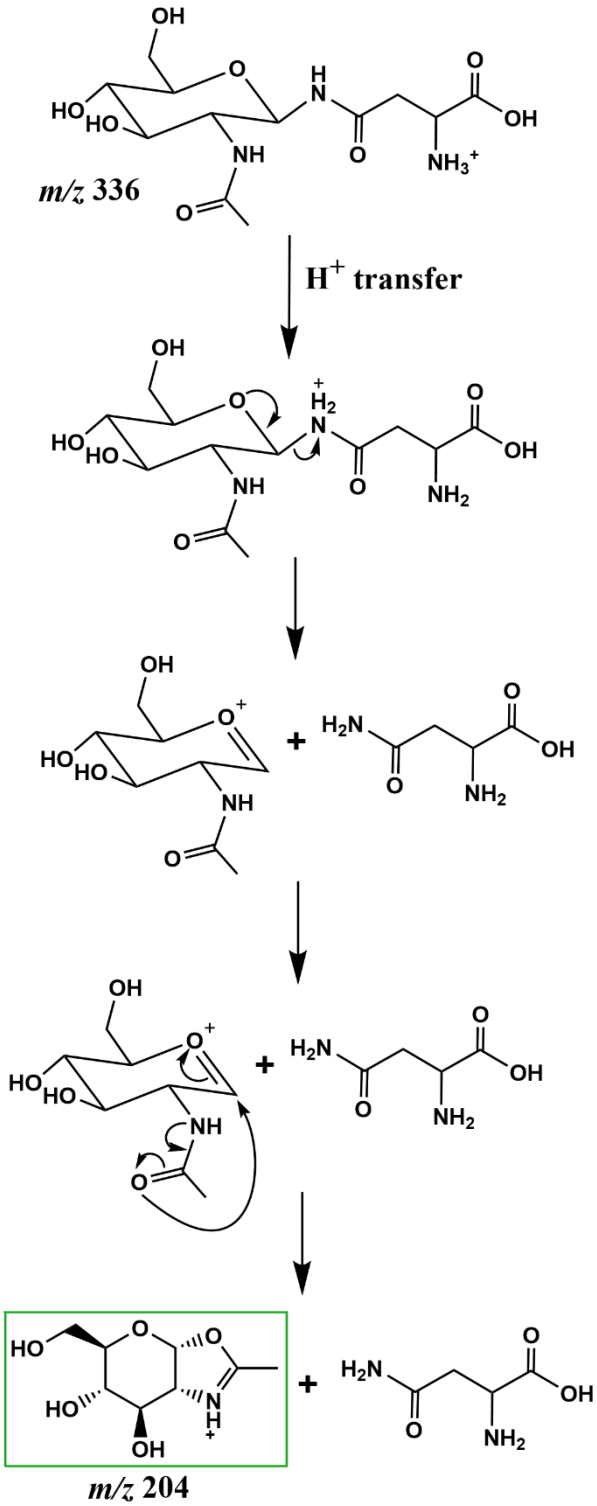
Direct Glycosidic Bond Dissociation Reactions

Consistent with previous reported CID of glycopeptides[38–40,44,45], the three glycosylated amino acids investigated in this study produce dominant bond cleavages at the glycosidic bond. How do these reactions occur?

Our calculations predict that mobilization of a proton to the glycosidic bond nitrogen or oxygen weakens the glycosidic bond[49] enabling its subsequent dissociation. Direct cleavage of the glycosidic bond results in a B_1 ion, *m/z* 204, and neutral asparagine/serine as illustrated in Scheme 1 (Schemes S1 & S2). For the [GlcNAc-β-1-Asn+H]$^+$ precursor our calculations indicate that the rate-determining glycosidic bond cleavage transition structure requires at least 149.3 (B3LYP) or 177.5 (M06-2X) kJ mol^{-1} (Figure S2, Table S1, & Table S2). Glycosidic bond cleavage produces a proton-bound dimer of neutral asparagine and the GlcNAc oxacarbenium ion. The oxacarbenium ion is comparatively unstable and collapses to form a substantially more energetically favorable (by >80 kJ mol^{-1}) bicyclic glucopyranosyl oxazolinium B_1 ion via nucleophilic attack of the carbonyl oxygen into the electropositive carbon 1. This reaction is less energetically demanding than the preceding glycosidic bond cleavage (120.6 or 158.0 5 kJ mol^{-1}, Figure S2b, Table S1 and Table S2). Dimer separation without proton transfer is barrierless and results in bicyclic glucopyranosyl oxazolinium B_1 ions and neutral asparagine. If the neutral asparagine abstracts a proton from the carbohydrate fragment prior to dimer separation, a protonated asparagine ion, [Asn+H]$^+$, will be detected. The [Asn+H]$^+$ is product limited requiring at least 215-267 kJ mol^{-1} (depending on chemical model and site of proton abstraction, Tables S1 & S2). The differing proton (deuteron) abstraction mechanisms prior to dimer separation produce the mass shifts observed in the deuterated analytes.

The same general reaction type was also calculated for the [GlcNAc-β-1-Ser+H]$^+$ and [GalNAc-α-1-Ser+H]$^+$ ions. For [GlcNAc-β-1-Ser+H]$^+$ this direct glycosidic bond cleavage reaction begins with glycosidic oxygen protonation. This is followed by concerted formation of a bicyclic glycopyranosyl oxazolinium B_1 ion and glycosidic bond cleavage (Scheme S1) in essentially a simplified form of the preceding (Asn) mechanism. Glycosidic bond cleavage is rate-limiting for both B_1 and [Ser+H]$^+$ formation (166.2 (B3LYP) or 190.0 (M06-2X) kJ mol^{-1}, Figure S3, Tables S3 & S4). The availability of 2 energetically

similar and feasible proton (deuteron) abstraction pathways is consistent with the doublet of serine ion produced from the deuterated analytes. Lastly, the direct glycosidic bond cleavage reaction of $[GalNAc-\alpha-1-Ser+H]^+$ parallels the other O-glycan in both mechanism and energetics (160.9 (B3LYP) or 184.9 (M06-2X) kJ mol^{-1}, Scheme S2, Figure S3, Tables S3 & S4).

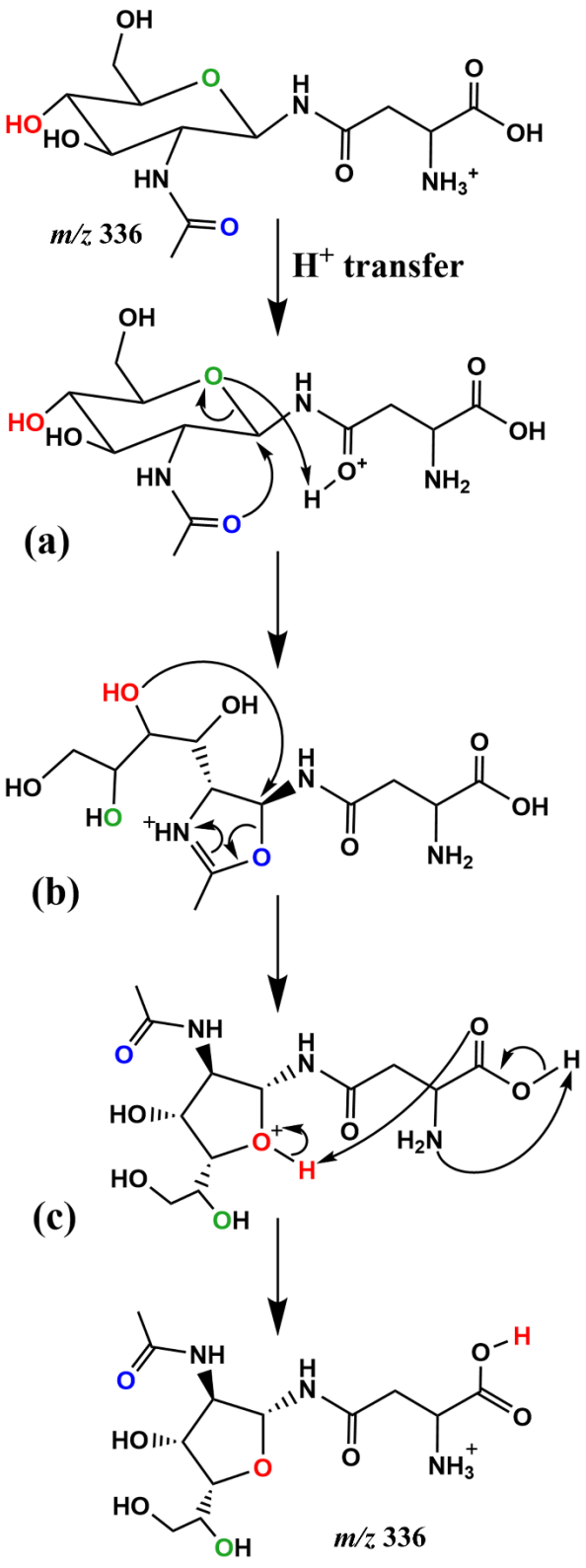


Scheme 1. Conventional (direct) glycosidic bond dissociation illustrated for $[GlcNac-\beta-1-Asn+H]^+$.

The preceding discussion offers a reasonable mechanism for the glycosidic bond cleavage reactions. However, there are difficulties. We found no feasible direct means of formation of the *m/z* 126 peak. Our and literature[45,69,70] MS^3 data support the hypothesis that this ion is formed from the *m/z* 204, B_1 ion population, by loss of $C_2H_6O_3$. We were unable to locate any energetically feasible means of losing $C_2H_6O_3$ from this glycopyranosyl oxazolinium B_1 ion as either one neutral molecule or as water then $C_2H_4O_2$. Consequently, the possibility of alternate means of dissociation leading to other B_1 ion structures were investigated.

Pyranose to Furanose Isomerization followed by Glycosidic Bond Cleavage

In solution the pyranose ring form of hexose sugars overwhelmingly predominates. As tandem MS occurs in the gas-phase, we wondered if isomerization were feasible under CID conditions? Our calculations indicate that it is. The isomerization process begins with opening of the pyranose ring. For $[GlcNAc-\beta-1-Asn+H]^+$ this reaction occurs by initially protonating the glycan ring oxygen (Scheme 2). The carbonyl oxygen of the N-acetyl group then nucleophilically attacks into carbon 1 in an S_N2-like TS which cleavages the ring-ether bond to carbon 1 forming an oxazoline-derivative (Scheme 2; Schemes S3 & S4 for the O-Ser congeners). The new oxazoline-derivative enables free rotation of the glycan carbon chain (carbons 3 to 6). Ergo, the hydroxyl oxygen on carbon 4 can now nucleophilically attack into carbon 1 eliminating the oxazoline-derivative in a further S_N2 reaction that generates a furanose ring structure (Figure 2, Figure S4). In direct contrast to all preceding literature proposals, our calculations indicate these isomerization reactions are energetically less demanding than the previously calculated direct glycosidic bond cleavage reactions in all cases. (Table S1-S6). i.e., for *both* the N- and O-linked glycans. The obvious next question is whether this enables formation of the glycosidic bond cleavage products?



Scheme 2. Pyranose to furanose isomerization pathway for $[GlcNac-\beta-1-Asn+H]^+$. (a) Concerted S_N2 pyranose ring-opening and oxazoline derivative formation; (b) oxazoline derivative ring-opening and formation of furanose structure; (c) proton transfer from the protonated furanose ring oxygen to the low-energy N-terminal amine of the Asn.

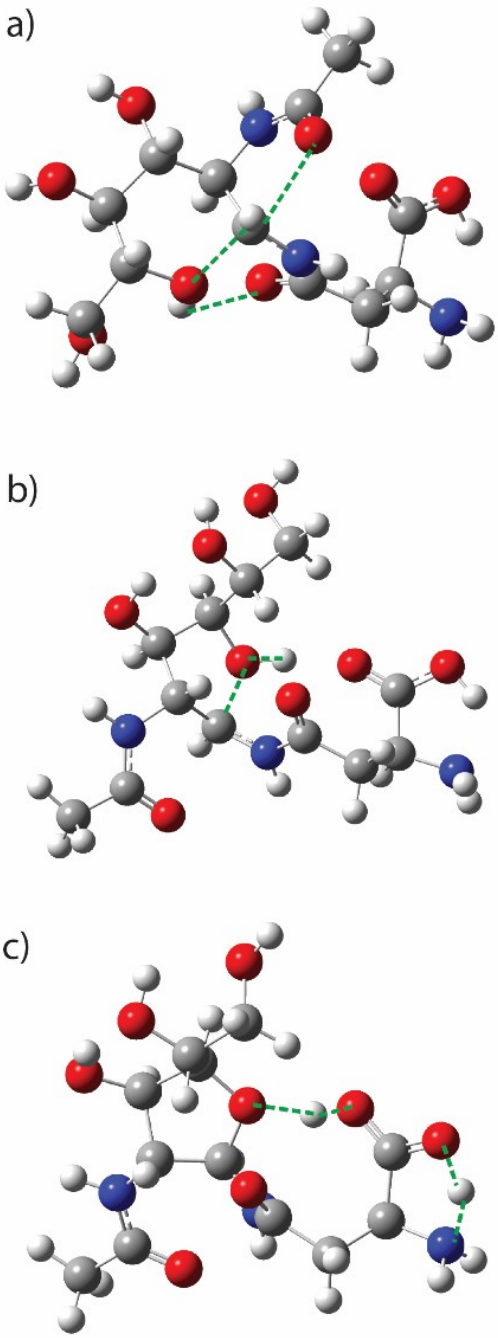
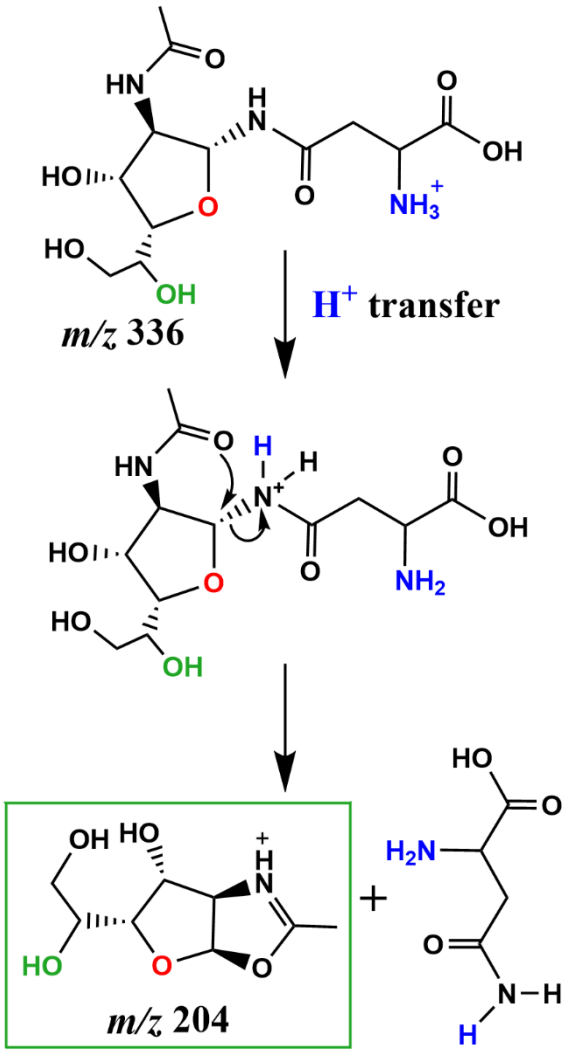


Figure 2. Transition structures of pyranose to furanose isomerization reaction in $[GlcNac-\beta-1-Asn+H]^+$. (a) Concerted S_N2 pyranose ring-opening and oxazoline derivative formation; (b) oxazoline derivative ring-opening and formation of furanose structure; (c) proton transfer from the protonated furanose ring oxygen to the low-energy N-terminal amine of the Asn.

Glycosidic Bond Cleavage Reactions from the Furanose Isomers

Glycosidic bond cleavage reactions are again initiated by proton mobilization to the glycosidic nitrogen/oxygen which weakens the glycosidic bond.[49] Nucleophilic attack by the carbonyl oxygen of the N-Acetyl group into carbon 1 with concerted cleavage of the glycosidic bond results in a dimer comprised of a glycofuranosyl oxazolinium B_1 ion and neutral

amino acid (Schemes 3, S5, S6, and Figure S5). Both levels of theory predict these dissociations to be more energetically favorable than the corresponding glycosidic bond cleavage reactions from the initial pyranose forms (Table S1-S6). All other aspects of this pathway are consistent with the previous discussion (D/H and charged amino acid ion formation mechanisms).



Scheme 3. Glycosidic bond cleavage following pyranose-furanose isomerization illustrated for [GlcNAc-β-1-Asn+H]$^+$.

As a check of the viability of these reactions, we performed a series of RRKM unimolecular dissociation rate calculations to investigate whether the isomerized furanose form glycosidic bond cleavage pathway is kinetically competitive (Figures 3 and S6-S9). Figure 3 summarizes these data by comparing the rate limiting steps of the two pathways for each analyte ion. In all cases relevant to our CID conditions (logk ≈ 2-6), the furanose pathways are favored. The GlcNAc reaction rates (Figure 3a & 3b) are more similar at high internal energies. Practically most of the analyte ions will have dissociated well before these internal energies can be accessed in our instrument. The reaction rate of protonated GalNAc O-glycans favors the isomerization followed by dissociation to a greater degree. Consequently, we would expect furanose oxazolinium B_1 ions to be the major component of our m/z 204 ion population in all cases. This finding, if general, directly contradicts many of the prevailing assumptions about glycan sequence and glycan ion dissociation chemistry.

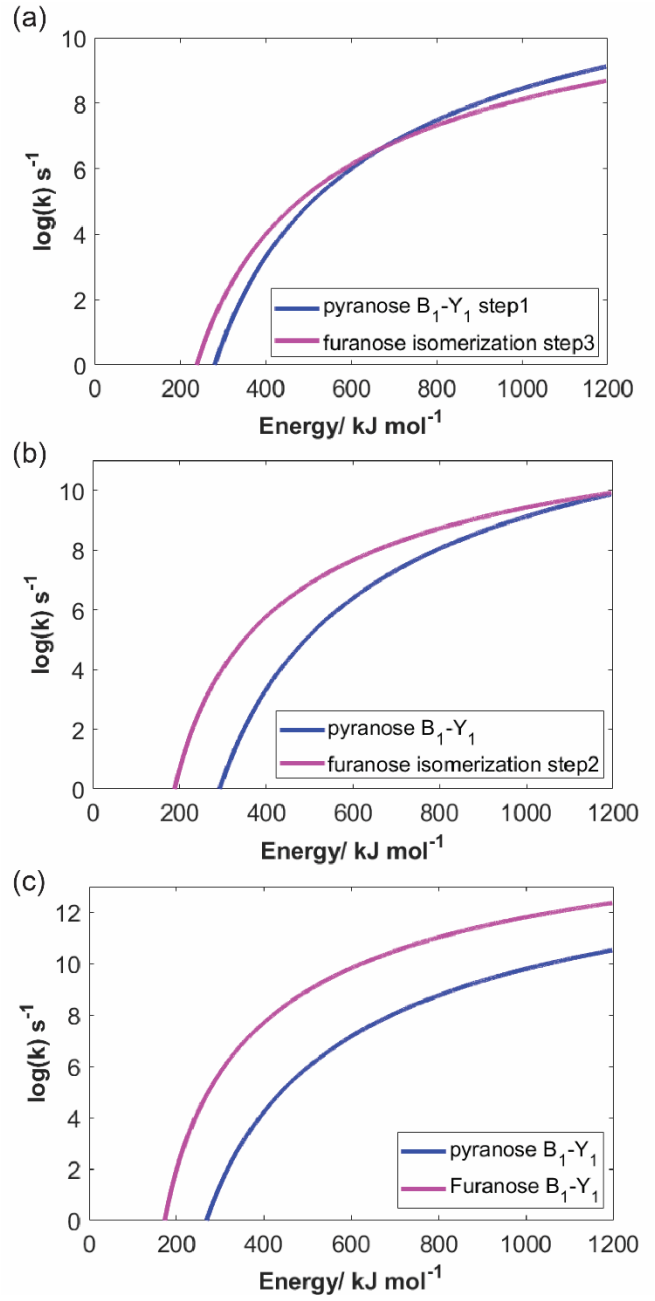


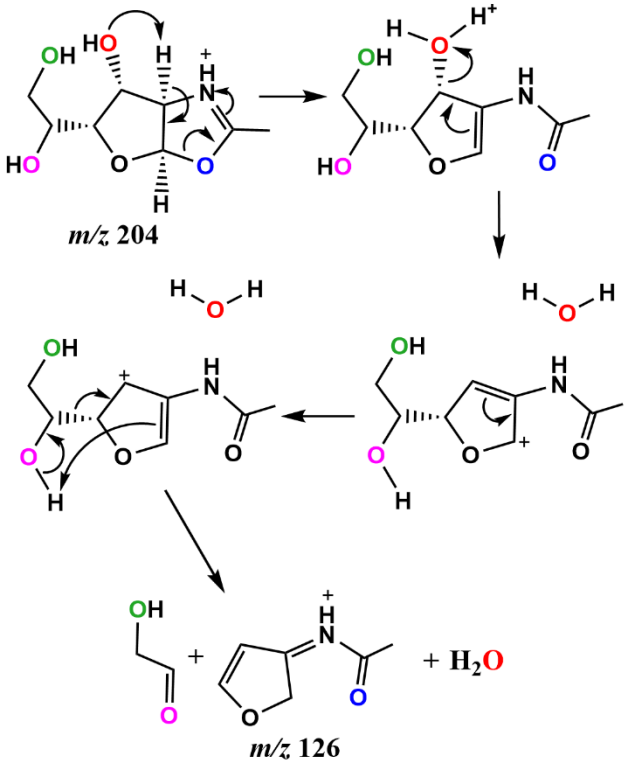
Figure 3. Plots of log(unimolecular rate constants, k) vs. internal energy for the rate-determining steps of peptidoglycan dissociation: (a) [GlcNAc-β-1-Asn+H]$^+$, (b) [GlcNAc-β-1-Ser+H]$^+$ (c) [GalNAc-α-1-Ser+H]$^+$.

Consecutive dissociation of the furanose oxazolinium B_1 ion: Formation of m/z 126

The bicyclic furanose oxazolinium B_1 ion can dissociate further to produce the *m/z* 126 ion. For the GlcNAc B_1 ions, our calculations predict that H_2O loss is initiated by C-alpha proton

abstraction from carbon 2 by the adjacent hydroxyl oxygen on carbon 3 resulting in a newly formed double bond between carbons 1 and 2 and concerted opening of the oxazoline ring (Scheme 4; Figure 4 a & b). This proton extraction step is rate-limiting and requires at least 296.9 or 360.0 kJ mol^{-1} for [GlcNAc-β-1-Asn+H]$^+$, or 284.3 or 351.0 kJ mol^{-1} for [GlcNAc-β-1-Ser+H]$^+$ (B3LYP then M06-2X).

For our deuterated GlcNAc spectra, this water loss occurred as HDO consistent with the present mechanism. The resultant carbocation is then able to abstract a proton from the hydroxyl group on carbon 5. Concerted bond cleavage between carbons 4 and 5 enables loss of glycolaldehyde, $C_2H_4O_2$, to generate the *m/z* 126 ion (Scheme 4; Figure 4c). This $C_2H_4O_2$ loss (203.4 or 257.9 kJ mol^{-1}, B3LYP then M06-2X) for [GlcNAc-β-1-Asn+H]$^+$, 190.7 or 248.8 kJ mol^{-1} for [GlcNAc-β-1-Ser+H]$^+$) is less energetically demanding than the preceding H_2O loss reaction so will occur spontaneously after the H_2O loss. Furthermore, consecutive reactions are very entropically favorable (Table S1-S4). Consequently, the small amount of B_1-H_2O ion, m/z 186, detected relative to the abundant *m/z* 126 peak in Figure 1 is rationalized.



Scheme 4. Mechanism of m/z 126 formation from the bicyclic furanose oxazolinium structured B_1 ion (β-GlcNAc-form).

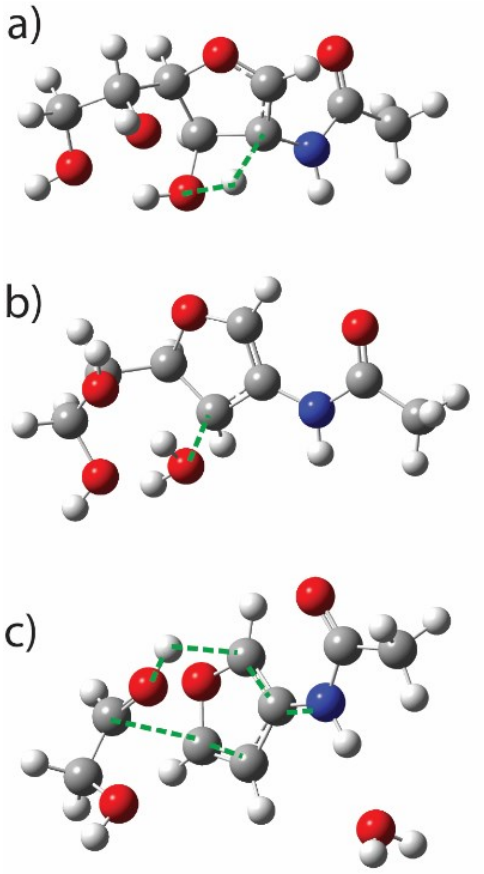


Figure 4. Transition structures that generate the m/z 126 ion from the β-GlcNAc-bicyclic furanose oxazolinium B_1 structure: (a) proton extraction, (b) H_2O loss, and (c) loss $C_2H_4O_2$.

The furanose oxazolinium B_1 ion structure produced from [GalNAc-α-1-Ser+H]$^+$ also produced a substantial peak at m/z 126. Our calculations predict that the same type of consecutive reaction mechanism is primarily responsible for this peak too (Scheme S7, Table S5-S6, Figure S10). Our combined data offer a mechanistic explanation of the dissociation chemistry consistent with Nilsson and co-workers[69,71] conclusion that the m/z 126 and 144 peaks "correspond to entirely different structures, which have followed completely different decomposition pathways"[71]. Consistent with the present dataset, Mookherjee et al.[72] recently provided evidence that cast doubt on the mechanism and product ion structure originally proposed by Yu et al.[71] for the key m/z 126 peak using model protonated HexNAc analytes rather than peptidoglycans.

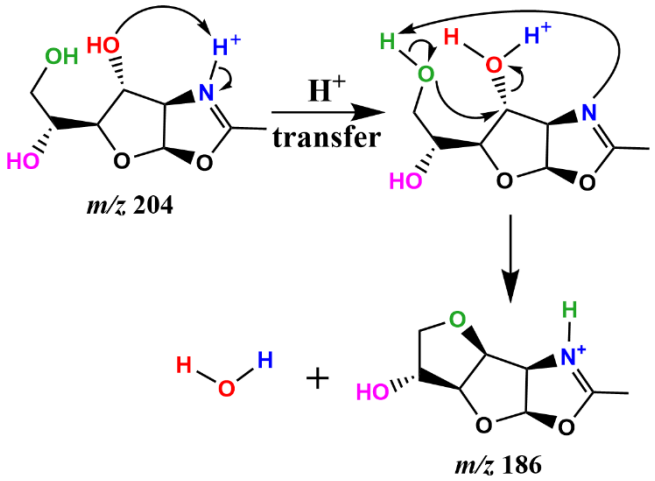
Mechanistic Support from ^{2}D Labelling: Differentiation of GalNAc and GlcNAc O-Glycans by Deuteration of Hydroxyl and NH Groups

The MS/MS spectra of the isomers, protonated O-linked GlcNAc-β-1-Ser and GalNAc-α-1-Ser analytes are quite similar, making confident differentiation nontrivial. In contrast, the deuterium labelling data shows a clear difference between these analytes; the B_1-H_2O peaks (originally *m/z* 186) enable differentiation between these two isomeric ions: (1) The $[B_1\text{-}H_2O]^+$ ions produced by the [GlcNAc-β-1-Ser/Asn+H]$^+$ predominantly have a 3 u shift to *m/z*=189, $[B_1\text{-}4H\text{+}4D\text{-}HDO]^+$;

(2) The $[B_1\text{-}H_2O]^+$ ions from [GalNAc-α-1-Ser+H]$^+$ split into 2 peaks at *m/z* 188 and *m/z* 189, shifts of 2 or 3 u respectively. These correspond to $[B_1\text{-}4H\text{+}4D\text{-}D_2O]^+$ and $[B_1\text{-}4H\text{+}4D\text{-}HDO]^+$ ions.

For the GlcNAc furanose oxazolinium B_1 ion the H_2O loss reaction involves abstraction of a C_{alpha}-H proton by the hydroxyl oxygen (Scheme 4). i.e., Consistent with loss of HDO in the deuterated experiments to yield $[B_1\text{-}4H\text{+}4D\text{-}HDO]^+$ at *m/z*=189. Calculated barriers support this pathway.

In contrast, the GalNAc furanose oxazolinium B_1 ion places C5 and C6 and their hydroxyl groups on the same side of the furanose ring (Scheme 5). In addition to enabling differing hydrogen bonding and charge solvation patterns, this allows a second competitive water loss pathway (Scheme 5). This water loss pathway is initiated by proton (deuteron) mobilization the from oxazolinium nitrogen to the C3 hydroxyl oxygen (Scheme 5; Figure 5a). The proximity of the C5 and C6 backbone and hydroxyl groups now permits nucleophilic attack by the C6 primary alcohol group into the electropositive C3 position with concerted expulsion of a water molecule (Scheme 5). This ring-forming reaction is once again S_N2-like (Figure 5b) and is followed by immediate abstraction of the C6 hydroxyl proton by the newly adjacent amide nitrogen. Loss of D_2O from the deuterated form is consistent with this mechanism and the calculations indicate this pathway is energetically comparable with the HDO loss pathway (Table S5 and S6), consistent with the deuterium labelling experimental results.



Scheme 5. H_2O (D_2O) loss pathway of the bicyclic α-GalNac-furanose oxazolinium B_1 ion.

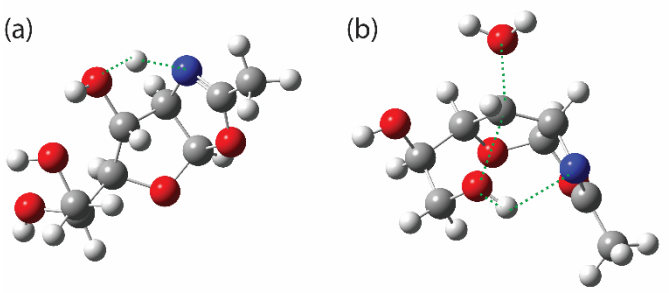


Figure 5. Transition state structures of H_2O (D_2O) loss reactions of the GalNAc bicyclic furanose oxazolinium B_1 ion: (a) proton (deuteron) transfer from oxazolinium nitrogen to Carbon 3 hydroxyl group, (b) loss of H_2O (D_2O).

Conclusions

We investigated the fragmentation chemistry of three peptidoglycans, GlcNAc-β-1-Asn, GalNAc-α-1-Ser, and GlcNAc-β- 1-Ser which represent the core structures of N-linked glycosylation, mucin-type glycosylation, and O-GlcNAcylation. We find that the glycosidic bond cleavage reactions following pyranose to furanose isomerization are *more favorable* than the direct glycosidic bond cleavage from the pyranose form. Ergo, furanose oxazolinium B_1 ions are the predominant m/z 204 ion structure within the gas-phase populations. The furanose oxazolinium B_1 ion structures enable subsequent formation of the abundant *m/z* 126 ions a major impediment to earlier hypotheses. Our mechanisms are consistent with the present deuterium labelled MS/MS data, unimolecular rate constants calculations and the literature.

Future work will involve testing the generality of these findings and their potential application in sequencing of larger protonated peptidoglycans. Whether similar processes are accessible to protonated oligosaccharides/glycans which lack the charge solvating amino acid/peptide functional groups adjacent to the reducing end of the carbohydrate will be addressed.

Conflicts of interest

There are no conflicts to declare.

Author Contributions

Both authors contributed substantially to all front-end aspects of this project and writing of the manuscript. SG discovered the key isomerization mechanism. BJB was responsible for administration, funding, and supervision.

Acknowledgements

The work was supported by the National Science Foundation (CHE-1948611). The Q-Exactive Plus Orbitrap was obtained with

support from the NSF MRI program (CHE-1428787). Calculations were performed at the Missouri University of Science and Technology, Rolla, MO with support from the NSF (ACI-1919789), at Ohio University, and at the Ohio Supercomputer Center, http://osc.edu/ark:/19495/f5s1ph73.

References

1 K. Furykawa and A. Kobata, Protein glycosylation, *Current Opinion in Biotechnology*, 1992, **3**, 554–559.
2 R. Apweiler, H. Hermjakob and N. Sharon, On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database, *Biochimica et Biophysica Acta - General Subjects*, 1999, **1473**, 4–8.
3 A. Helenius and M. Aebi, Intracellular Functions of N-Linked Glycans, *Science (New York, N.Y.)*, 2001, **291**, 2364–2369.
4 Y. Satomi, Y. Shimonishi and T. Takao, N-glycosylation at Asn491 in the Asn-Xaa-Cys motif of human transferrin, *FEBS Letters*, 2004, **576**, 51–56.
5 Z. Zhu and H. Desaire, Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry, *Annual Review of Analytical Chemistry*, 2015, **8**, 463–483.
6 P. H. Jensen, D. Kolarich and N. H. Packer, Mucin-type O-glycosylation - Putting the pieces together, *FEBS Journal*, 2010, **277**, 81–94.
7 W. R. Alley and M. V. Novotny, Structural Glycomic Analyses at High Sensitivity: A Decade of Progress, *Annual Review of Analytical Chemistry*, 2013, **6**, 237–265.
8 C. Reily, T. J. Stewart, M. B. Renfrow and J. Novak, Glycosylation in health and disease, *Nature Reviews Nephrology*, 2019, **15**, 346–366.
9 C. R. Torres and G. W. Hart, Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc., *The Journal of biological chemistry*, 1984, **259**, 3308–3317.
10 J. A. Hanover, C. K. Cohen, M. C. Willingham and M. K. Park, O-linked N-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins., *Journal of Biological Chemistry*, 1987, **262**, 9887–9894.
11 G. D. Holt and G. W. Hart, The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharidie linkage, O-linked GlcNAc, *Journal of Biological Chemistry*, 1986, **261**, 8049–8057.
12 K. P. Kearse and G. W. Hart, Lymphocyte activation induces rapid changes in nuclear and cytoplasmic glycoproteins, *Proceedings of the National Academy of Sciences of the United States of America*, 1991, **88**, 1701–1705.
13 R. J. Copeland, J. W. Bullen and G. W. Hart, Cross-talk between GlcNAcylation and phosphorylation: Roles in insulin resistance and glucose toxicity, *American Journal of Physiology - Endocrinology and Metabolism*, 2008, **295**, 17–28.
14 Q. Zeidan and G. W. Hart, The intersections between O-GlcNAcylation and phosphorylation: Implications for multiple signaling pathways, *Journal of Cell Science*, 2010, **123**, 13–22.
15 O. Gornik, L. Royle, D. Harvey, C. Radcliffe, R. Saldova, R. Dwek, P. Rudd and G. Lauc, Changes of serum glycans during sepsis and acute pancreatitis, *Glycobiology*, 2007, **17**, 1321–1332.
16 B. Adamczyk, T. Tharmalingam and P. M. Rudd, Glycans as cancer biomarkers, *Biochimica et Biophysica Acta - General Subjects*, 2012, **1820**, 1347–1353.
17 M. Sperandio, C. A. Gleissner and K. Ley, Glycosylation in immune cell trafficking, *Immunological Reviews*, 2009, **230**, 97–113.
18 N. Tanichi and H. Korekane, Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics, *BMB Reports*, 2011, **44**, 772–781.
19 K. Ohtsubo and J. D. Marth, Glycosylation in Cellular Mechanisms of Health and Disease, *Cell*, 2006, **126**, 855–867.
20 K. J. Brown, A. Vanderver, E. P. Hoffman, R. Schiffmann and Y. Hathout, Characterization of transferrin glycopeptide structures in human cerebrospinal fluid, *International Journal of Mass Spectrometry*, 2012, **312**, 97–106.
21 Y. Chen, S. Hojo, N. Matsumoto and K. Yamamoto, Regulation of Mac-2BP secretion is mediated by its N-glycan binding to ERGIC-53, *Glycobiology*, 2013, **23**, 904–916.
22 M. Nagae, K. Morita-Matsumoto, S. Arai, I. Wada, Y. Matsumoto, K. Saito, Y. Hashimoto and Y. Yamaguchi, Structural change of N-glycan exposes hydrophobic surface of human transferrin, *Glycobiology*, 2014, **24**, 693–702.
23 R. G. Spiro, Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds, *Glycobiology*, 2002, **12**, 43R-56R.
24 J. Bones, S. Mittermayr, N. O'Donoghue, A. Guttman and P. M. Rudd, Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation, *Analytical Chemistry*, 2010, **82**, 10208–10215.
25 R. Saldova, L. Royle, C. Radcliffe, U. Abd hamid, R. Evans, J. Arnold, R. Banks, R. Huston, D. Harvey, R. Antrobus, S. Petrescu, R. Dwek and P. Rudd, Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG, *Glycobiology*, 2007, **17**, 1344–1356.
26 E. Miyoshi and M. Nakano, Fucosylated haptoglobin is a novel marker for pancreatic cancer: Detailed analyses of oligosaccharide structures, *Proteomics*, 2008, **8**, 3257–3262.
27 N. Okuyama, Y. Ide, M. Nakano, T. Nakagawa, K. Yamanaka, K. Moriwaki, K. Murata, H. Ohigashi, S. Yokoyama, H. Eguchi, O. Ishikawa, T. Ito, M. Kato, A. Kasahara, S. Kawano, J. Gu, N. Taniguchi and E. Miyoshi, Fucosylated haptoglobin is a novel marker for pancreatic cancer: A detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation, *International Journal of Cancer*, 2006, **118**, 2803–2808.
28 D. H. Dube and C. R. Bertozzi, Glycans in cancer and inflammation - Potential for therapeutics and diagnostics, *Nature Reviews Drug Discovery*, 2005, **4**, 477–488.
29 S. Rachagani, M. P. Torres, N. Moniaux and S. K. Batra, Current status of mucins in the diagnosis and therapy of cancer, *BioFactors*, 2009, **35**, 509–527.
30 A. Cazet, S. Julien, M. Bobowski, M. A. Krzewinski-Recchi, A. Harduin-Lepers, S. Groux-Degroote and P. Delannoy, Consequences of the expression of sialylated antigens in breast cancer, *Carbohydrate Research*, 2010, **345**, 1377–1383.
31 Y. Tian and H. Zhang, Characterization of disease-associated N-linked glycoproteins, *PROTEOMICS*, 2012, **13**, 504–511.
32 A. Chernykh, R. Kawahara and M. Thaysen-Andersen, Towards structure-focused glycoproteomics, *Biochemical Society Transactions*, 2021, **49**, 161–186.
33 Z. Darula and K. F. Medzihradszky, Analysis of mammalian O-glycopeptides - We have made a good start, but there is a long way to go, *Molecular and Cellular Proteomics*, 2018, **17**, 2–17.

34 X. You, H. Qin and M. Ye, Recent advances in methods for the analysis of protein o-glycosylation at proteome level, *Journal of Separation Science*, 2018, **41**, 248–261.
35 K. Ohtsubo and J. D. Marth, Glycosylation in Cellular Mechanisms of Health and Disease, *Cell*, 2006, **126**, 855–867.
36 H. J. An and C. B. Lebrilla, Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial), *Mass Spectrometry Reviews*, 2011, **30**, 560–578.
37 S. Pan, R. Chen, R. Aebersold and T. A. Brentnall, Mass spectrometry based glycoproteomics - From a proteomics perspective, *Molecular and Cellular Proteomics*, 2011, **10**, 1–14.
38 M. Wuhrer, M. I. Catalina, A. M. Deelder and C. H. Hokke, Glycoproteomics based on tandem mass spectrometry of glycopeptides, *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 2007, **849**, 115–128.
39 Morelle, W and J. Michalski, Analysis of protein glycosylation by mass spectrometry, *Nature protocols*, 2007, **2**, 1585–1602.
40 K. Vékey, O. Ozohanics, E. Tóth, A. Jek, Á. Révész, J. Krenyácz and L. Drahos, Fragmentation characteristics of glycopeptides, *International Journal of Mass Spectrometry*, 2013, **345–347**, 71–79.
41 G. Zauner, C. A. M. Koeleman, A. M. Deelder and M. Wuhrer, Protein glycosylation analysis by HILIC-LCMS of Proteinase K-generated N- and Oglycopeptides, *Journal of Separation Science*, 2010, **33**, 903–910.
42 E. D. Dodds, R. R. Seipert, B. H. Clowers, J. B. German and C. B. Lebrilla, Analytical performance of immobilized pronase for glycopeptide footprinting and implications for surpassing reductionist glycoproteomics, *Journal of Proteome Research*, 2009, **8**, 502–512.
43 P. H. Jensen, D. Kolarich and N. H. Packer, Mucin-type O-glycosylation - Putting the pieces together, *FEBS Journal*, 2010, **277**, 81–94.
44 E. D. Dodds, Gas-phase dissociation of glycosylated peptide ions, *Mass Spectrometry Reviews*, 2012, **31**, 666–682.
45 J. Nilsson, Liquid chromatography-tandem mass spectrometry-based fragmentation analysis of glycopeptides, *Glycoconjugate Journal*, 2016, **33**, 261–272.
46 H. Peltoniemi, S. Joenväärä and R. Renkonen, De novo glycan structure search with the CID MS/MS spectra of native N-glycopeptides, *Glycobiology*, 2009, **19**, 707–714.
47 O. Ozohanics, J. Krenyacz, K. Ludányi, F. Pollreisz, K. Vékey and L. Drahos, GlycoMiner: a new software tool to elucidate glycopeptide composition, *Rapid Communications in Mass Spectrometry*, 2008, **22**, 3245–3254.
48 D. Goldberg, M. Bern, S. Parry, M. Sutton-Smith, M. Panico, H. R. Morris and A. Dell, Automated N-glycopeptide identification using a combination of single- and tandem-MS, *Journal of Proteome Research*, 2007, **6**, 3995–4005.
49 B. J. Bythell, M. T. Abutokaikah, A. R. Wagoner, S. Guan and J. M. Rabus, Cationized carbohydrate gas-phase fragmentation chemistry, *Journal of The American Society for Mass Spectrometry*, 2017, **28**, 688–703.
50 M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α-Maltose, *Journal of Chemical Theory and Computation*, 2016, **12**, 6157–6168.
51 A. Supady, V. Blum and C. Baldauf, First-Principles Molecular Structure Search with a Genetic Algorithm, *Journal of Chemical Information and Modeling*, 2015, **55**, 2338–2348.
52 T. A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *Journal of Computational Chemistry*, 1996, **17**, 490–519.
53 T. A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *Journal of Computational Chemistry*, 1996, **17**, 520–552.
54 T. A. Halgren, Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94, *Journal of Computational Chemistry*, 1996, **17**, 553–586.
55 T. A. Halgren and R. B. Nachbar, Merck molecular force field. IV. conformational energies and geometries for MMFF94, *Journal of Computational Chemistry*, 1996, **17**, 587–615.
56 T. A. Halgren, Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules, *Journal of Computational Chemistry*, 1996, **17**, 616–641.
57 J. M. Rabus, M. T. Abutokaikah, R. T. Ross and B. J. Bythell, Sodium-cationized carbohydrate gas-phase fragmentation chemistry: influence of glycosidic linkage position, *Phys. Chem. Chem. Phys.*, 2017, **19**, 25643–25652.
58 J. M. Rabus, R. P. Pellegrinelli, A. H. A. Khodr, B. J. Bythell, T. R. Rizzo and E. Carrascosa, Unravelling the structures of sodiated β-cyclodextrin and its fragments, *Phys. Chem. Chem. Phys.*, 2021, **23**, 13714–13723.
59 M. T. Abutokaikah, J. W. Frye, J. Tschampel, J. R. Rabus and B. J. Bythell, Fragmentation Pathways of Lithiated Hexose Monosaccharides, *J. Am. Soc. Mass Spectrom.*, 2018, **29**, 1627–1637.
60 J. M. Rabus, D. R. Simmons, P. Maître and B. J. Bythell, Deprotonated carbohydrate anion fragmentation chemistry: structural evidence from tandem mass spectrometry, infra-red spectroscopy, and theory, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27897–27909.
61 A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics*, 1993, **98**, 5648–5652.
62 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Physical Review B*, 1988, **37**, 785–789.
63 Y. Zhao and D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function, *Theoretical Chemistry Accounts*, 2008, **120**, 215–241.
64 Y. Zhao, N. E. Schultz and D. G. Truhlar, Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions, *The Journal of Chemical Physics*, 2005, **123**, 161103.
65 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L.

Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Farkas, O., J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision E.01*, Gaussian, Inc., Wallingford CT, 2009.

66 R. G. Gilbert and S. C. Smith, *Theory of unimolecular and recombination reactions*, 1990.

67 C. Lifshitz, Some recent aspects of unimolecular gas phase ion chemistry, *Chemical Society Reviews*, 2001, **30**, 186–192.

68 T. Beyer and D. F. Swinehart, Algorithm 448: Number of multiply-restricted partitions, *Communications of the ACM*, 1973, **16**, 379.

69 A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Rüetschi, G. Brinkmalm, C. Sihlbom, J. Lengqvist, G. Larson and J. Nilsson, Assignment of Saccharide Identities through Analysis of Oxonium Ion Fragmentation Profiles in LC–MS/MS of Glycopeptides, *J. Proteome Res.*, 2014, **13**, 6024–6032.

70 C. Ashwood, C.-H. Lin, M. Thaysen-Andersen and N. H. Packer, Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS, *J. Am. Soc. Mass Spectrom.*, 2018, 1–16.

71 J. Yu, M. Schorlemer, A. Gomez Toledo, C. Pett, C. Sihlbom, G. Larson, U. Westerlind and J. Nilsson, Distinctive MS/MS Fragmentation Pathways of Glycopeptide-Generated Oxonium Ions Provide Evidence of the Glycan Structure, *Chemistry – A European Journal*, 2016, **22**, 1114–1124.

72 A. Mookherjee, S. S. Uppal and M. Guttman, Dissection of Fragmentation Pathways in Protonated *N* -Acetylhexosamines, *Analytical Chemistry*, 2018, **90**, 11883–11891.