

# Linkage Memory in Underivatized Protonated Carbohydrates

Abhigya Mookherjee, Sanjit S. Uppal, Taylor A. Murphree, Miklos Guttman\*

Department of Medicinal Chemistry; University of Washington, Seattle, WA 98195

\*email: [mguttman@uw.edu](mailto:mguttman@uw.edu)

## **ABSTRACT**

Carbohydrates are among the most complex class of biomolecules and even subtle variations in their structures are attributed to diverse biological function. Mass spectrometry has been essential for large scale glycomics and glycoproteomics studies, but the gas-phase structures and sometimes anomalous fragmentation properties of carbohydrates present longstanding challenges. Here we investigate the gas-phase properties of a panel of isomeric protonated disaccharides differing in their linkage configurations. Multiple conformations were evident for most of the structures based on their fragment ion abundances by tandem mass spectrometry, their ion mobilities in several gases, and their deuterium uptake kinetics by gas-phase hydrogen deuterium exchange. Most notably, we find that the properties of the Y-ion fragments are characteristically influenced by the precursor carbohydrate's linkage configuration. This study reveals how protonated carbohydrate fragment ions can retain 'linkage memory' that provides structural insight into their intact precursor.

## INTRODUCTION

Carbohydrates or 'glycans' are abundant biological macromolecules that make up the majority of the glycocalyx on cells in the form of glycoconjugates. Glycosylation is one of the most common and complex post-translational modifications (PTMs) of secreted and cell surface proteins, enabling diverse biological processes like cell adhesion, cell-cell communication, and immune responses <sup>1-3</sup>. Unlike proteins and nucleic acids which are linear polymers built from a template, glycan structures are the sum result of many glycosidase and glycosyltransferase enzymes that can result in polymers with various branching patterns <sup>4</sup>. Additional structural complexity stems from glycans having various positional isomers, linkage stereochemistry ( $\alpha/\beta$  anomers), and different compositions of the various monosaccharide building blocks.

Mass spectrometry (MS) is one of the primary techniques used to investigate the immense diversity of glycans and a great deal of effort has been invested into understanding the gas-phase behavior of carbohydrates. However, MS-based glycan analysis is often confounded by the technique's inability to distinguish among the many potential isomeric structures from a mass measurement. A thorough characterization of the tandem MS behavior of permethylated oligosaccharides paved the way for effective discrimination among potential isomers <sup>5</sup>. Protonated carbohydrates present additional challenges, as they can undergo rearrangement reactions, most notably migration of fucose during collision-induced dissociation (CID) <sup>6</sup>. This migration process is still under investigation <sup>7-9</sup>, and a general structural understanding of the even simple protonated carbohydrate systems is only now starting to emerge <sup>10</sup>. For intact glycoconjugates (i.e. glycopeptides), which are typically analyzed as protonated ions, tandem MS is incapable of providing full information on the branching pattern, linkage information, or stereochemistry of the glycan <sup>11</sup>. Understanding glycobiology at the highest level of detail, including site-specific localization of all glycan structural isomers within glycoproteins <sup>12</sup>, will necessitate a thorough understanding of the properties of carbohydrate fragment ions.

Tools complementary to MS including ion mobility (IM), spectroscopic techniques like cryogenic ion spectroscopy and ultraviolet photodissociation have emerged as powerful gas-

phase techniques with potential applications in glycobiology <sup>13-21</sup>. Ion mobility with mass spectrometry (IM-MS) separates ions by their charge, size, and shape due to interactions with an inert gas. IM-MS has shed light on the structural details of carbohydrates in the gas phase <sup>17,22</sup>, and the emerging next-generation of high-resolution IM-MS show much promise at resolving even subtle isomeric variants within larger glycans <sup>20,23,24</sup>. Gas-phase spectroscopy is emerging to define structural variations in carbohydrates and their structures on its own or in combination with IM-MS <sup>18,20,25-27</sup>. Gas-phase hydrogen/deuterium exchange (gHDX), which tracks the exchangeability of various labile protons in an ion, has also been shown effective at distinguishing carbohydrate isomers <sup>28</sup>.

Thanks to their ability to resolve structural isomers, these new techniques have also shown examples of multiple distinct gas-phase chemical structures obtained for a single distinct carbohydrate, often attributed to the site of protonation or deprotonation <sup>29</sup>. Recent studies have further illustrated the complexity of fragmentation of protonated carbohydrates. Gray et al. using IM-MS found that B-ion fragments of carbohydrates retain information of the stereochemistry at their reducing end even *after* losing their reducing end ('anomeric memory') <sup>17</sup>. This is thought to arise from the generation of different isomeric fragment ion structures, whose ensemble is affected by the structure of the precursor. Similar 'memory effects' were also observed by both IM-MS and MS<sup>n</sup> relative fragmentation propensities even after multi-stage MS <sup>22</sup>. Here we sought to characterize the fragmentation patterns within larger oligosaccharides and the structural disposition of the resulting fragment ions. The study revealed an unprecedented 'linkage memory' effect where carbohydrate Y-ions retain information about linkage patterns at their non-reducing ends. Differences in CID fragmentation patterns, ion mobilities, and gHDX kinetics all corroborate that carbohydrate Y-ions adopt conformational ensembles dependent on their prior linkage configuration. Overall, these studies demonstrate that certain carbohydrate fragments can retain memory of both the linkage configuration and anomeric stereochemistry of their precursor.

## Methods

## **Reagents**

The standards Gal( $\beta$ 1-4)GlcNAc, Gal( $\beta$ 1-6)GlcNAc, Gal( $\beta$ 1-3)GlcNAc, Fuc( $\alpha$ 1-4)GlcNAc, Fuc( $\alpha$ 1-3)GlcNAc and Fuc( $\alpha$ 1-6)GlcNAc were purchased from Carbosynth (Compton, UK). Deuterated ammonia (ND<sub>3</sub>, 99%) and <sup>18</sup>O (98%) water were purchased from Cambridge isotope labs (Tewksbury, MA, USA). All standards from stock solutions were verified by <sup>1</sup>H NMR to ensure purity and conformational identity (Figure S10) as described previously <sup>30</sup>.

## **Ion trap Mass Spectrometry**

Carbohydrates were re-suspended in LC-MS grade optima water and diluted to a working concentration of 10  $\mu$ M in 0.1% formic acid. <sup>18</sup>O labeling of the reducing end hydroxyl of GlcNAc was performed as described previously <sup>31</sup>. Samples were analyzed on a LTQ Orbitrap Mass Spectrometer (Thermo Fisher Scientific) by direct infusion at a rate of  $\sim$ 5  $\mu$ L/min. The protonated precursor ion was subjected to CID using a range of normalized collision energies (CE) of 15 – 20. Full (MS<sub>1</sub>) scans were acquired followed by selection of the most intense fragment ions for further fragmentation with up to four rounds of MS/MS (MS<sup>4</sup>). All the multi-stage tandem MS experiments were performed in triplicate. Spectra were extracted and analyzed using Xcalibur software. MS analysis was coupled to liquid chromatography (Waters Acquity I-class) to resolve  $\alpha/\beta$  anomers. 1  $\mu$ L of a 10-50  $\mu$ M solution of each was injected in pure water and resolved over a 2.1 x 150 mm 300 Å BEH amide column (Waters) using a gradient of 90%B to 55%B over 20 minutes (A: water with 0.1% formic acid; B: acetonitrile with 0.1% formic acid).

## **Ion mobility mass spectrometry (IM-MS)**

Ion mobility measurements were performed on a Waters Synapt G2-Si mass spectrometer. 50  $\mu$ M glycans in 0.1% formic acid were directly infused at 5  $\mu$ L/min. The sample cone voltage was optimized to generate each carbohydrate fragment ion of interest ( $m/z$  366 for Gal-GlcNAcs or 350 for Fuc-GlcNAcs) and mass selected using the quadrupole to avoid any artifacts of ion decay in, during, or after the IM stage. IM measurements in different gasses were

collected using the following settings: N<sub>2</sub> flow rate of 90 mL/min, traveling wave (TW) velocity of 500 m/s and height of 25 V; He flow rate of 120 mL/min, TW velocity and height were 800 m/s and 8 V; and CO<sub>2</sub> flow rate of 42 mL/min, TW velocity and height were 400 m/s and 16 V. The ion current was kept below 10e5 counts/sec to avoid detector saturation, which may lead to broader than expected arrival time distributions (ATDs). These ATDs were fit to Gaussian distributions using custom Excel (Microsoft, Redmond, WA) scripts.

### ***Gas-phase Hydrogen/deuterium exchange (gHDX)***

Instrumental modifications to the Waters Synapt HDMS mass spectrometer to enable infusion of ND<sub>3</sub> and He have been described previously<sup>28</sup>. Analytes were resuspended in Optima LC-MS grade water from Fischer Scientific (Hampton, NH, USA) and diluted to a working concentration of 400 μM in aqueous 0.1% formic acid. An elevated cone voltage was used to generate water loss fragment ions in the source (*m/z* 350 and 366) which were then mass selected with the quadrupole and fragmented in the trap ion guide. A series of acquisitions were collected using a trap collision energy of 5, 10 and 17 for Gal(β1-4)GlcNAc and Gal(β1-3)GlcNAc, 5, 10 and 22 for Gal(β1-6)GlcNAc, and 5, 10 and 15 for Fuc(α1-4)GlcNAc, Fuc(α1-3)GlcNAc, and Fuc(α1-6)GlcNAc. Transfer wave heights of 8V and multiple traveling wave velocities (8 – 512 m/s<sup>-1</sup>) were used to sample a temporal-range of ~0.1 to 10 ms. Unless specified all other MS settings were: source temperature 100 °C, capillary voltage 2.5 kV; sampling cone voltage 30 V; extraction cone voltage 3.3 V; trap CE 15 V; trap bias 8 V; IM TW velocity 100 m/s; IM TW height 5 V; RF confining voltage 350 V; transfer TW height 8 V. The “lock” channel was used to infuse TrisMix, an exchange standard, co-sampled along with the analyte of interest and used to verify that the exchange conditions were identical for all samples<sup>32</sup>. All samples were collected in triplicate. Data was batch converted using scripts in UniDec<sup>33</sup>. Quantification of deuterium uptake, binomial fitting, and bimodal deconvolution were performed using HX-Express v2<sup>34</sup>.

## **RESULTS**

### *Characterization of linkage and anomeric memory by tandem MS*

We first investigated two sets of isomeric disaccharides by MS<sup>n</sup> to examine their fragmentation behavior. Both sets of disaccharides, Gal(β1-x)GlcNAc and Fuc(α1-x)GlcNAc, where x = 3, 4, and 6, are compositionally identical but different in their linkage configurations (Fig. S1B). Upon CID, the protonated precursors (*m/z* 384 for Gal-GlcNAcs and *m/z* 368 for Fuc-GlcNAcs) dissociate easily to lose water from their reducing end to form *m/z* 366 in Gal-GlcNAcs and *m/z* 350 in the case of Fuc-GlcNAcs (Fig. 1A, B). Loss at the reducing end is confirmed by identical dissociation of an <sup>18</sup>O labeled fucosylated disaccharide (*m/z* 370), which dissociates primarily to *m/z* 350, thus confirming that the initial loss of water (-18 Da) is predominantly occurring from the reducing end of the disaccharides (Fig. S1B, C). In the case of the protonated Gal(β1-x)GlcNAcs, all three isomeric precursors fragment to the same product ions but with differing propensities. While *m/z* 222 is the classic Y1 ion formed by cleavage of the glycosidic bond, resulting in protonated GlcNAc; *m/z* 204 is formed by further loss of the reducing end of the protonated GlcNAc. The relative abundances of *m/z* 222 and 204 are visibly distinct for Gal(β1-6)GlcNAc compared to the other two disaccharides (Fig. 1A). Similar to the Gal(β1-x)GlcNAc disaccharides, MS<sup>2</sup> of the protonated Fuc(α1-x)-GlcNAc disaccharides (*m/z* 368) yield *m/z* 350, 222, and 204 (Fig. 1B). The MS<sup>2</sup> spectra of the fucosylated disaccharides is unique only for the (1-4) linkage. Further fragmentation of the water-loss fragment (MS<sup>3</sup>) of all six disaccharides revealed the same product ions namely, *m/z* 204, 186, 168, 144, and 138 with varying relative abundances (Fig. 1C). Though MS<sup>3</sup> (368→350→) of the fucosylated disaccharides yield the same fragment ions as Gal(β1-x)GlcNAcs, there are no unique patterns in the fragmentation spectra among the three linkage isomers (Fig. 1D).

Taking a step further, MS<sup>4</sup> (Figs. 1E, F) was performed on all six disaccharides to fragment the GlcNAc (*m/z* 204) remnant, probing for signature CID fragmentation patterns. In addition to *m/z* 186, 168, 144 and 138, the MS<sup>4</sup> (384→366→204→) spectra of all disaccharides yield *m/z* 126 and 84 which are primary and secondary products of *m/z* 204<sup>22</sup> (Figs. 1E, F). Interestingly, fragmentation of the 204 ion appears similar from Gal(β1-3)GlcNAc and Gal(β1-4)GlcNAc in their

relative abundances, and markedly distinct from Gal( $\beta$ 1-6)GlcNAc. Even more interesting is the MS<sup>4</sup> of the fucosylated disaccharides (368→350→204→) that showed distinct fragmentation spectra for all three isomers (Fig. 1F). The observed differences were quantified by calculating the ratio of the relative abundances of the *m/z* 126 and 186 fragments from triplicate measurements (Figs. 1E, F). For both sets of disaccharides the 126/186 ratio is markedly different for the (1-6) linkage as compared to the (1-4) linkage. For the fucosylated disaccharides, there was a significant difference even between the (1-3) and (1-4) linkages. These varying ratios indicate that the fragmentation propensity, and thereby the structures of the 204 ions differ when generated from precursors of different linkages.

All experiments thus far were performed using direct infusion and thus would probe both the  $\alpha/\beta$  anomers that are present in solution. To examine how the reducing end anomeric configuration influences the structures generated from the precursor, we repeated the analyses using online hydrophilic interaction chromatography (HILIC) to resolve the  $\alpha/\beta$  anomers of each disaccharide (Fig. S2). The MS<sup>4</sup> spectra (384→366→204→) of all six disaccharides showed similar trends to the infusion data (Figs. 1G, H), but in some cases there were large differences between the two anomers. Definitive assignment of the  $\alpha$  and  $\beta$  anomers for the disaccharides cannot be made from this LC-MS data, and we therefore only designate each species as early vs. late eluting. The (1-6) linked Fuc-GlcNAc and Gal-GlcNAc show little dependence on the anomeric state of the precursor, but the (1-4) and (1-3) linked disaccharides generate more *m/z* 126 from the early eluting peak. This is most notable for Fuc( $\alpha$ 1-4)GlcNAc and Gal( $\beta$ 1-3)GlcNAc where the early eluting peaks have more than double the relative intensity for the 126 peak. For this comparison we also included analysis of the GlcNAc monosaccharide, for which the MS<sup>n</sup> relative product ion abundances (222→204→) were completely distinct from those generated from the disaccharides.

#### *Characterization of linkage and anomeric memory by IM-MS*

Ion mobility spectrometry was applied to further characterize and compare the fragment ions from the isomeric disaccharides. The arrival time distributions (ATDs) of the  $m/z$  366 and 350 ions from Gal-GlcNAc and Fuc-GlcNAc are shown in Figs. 2A and B. Not surprisingly, each linkage isomer has a distinct ATD, consistent with a different structure. However, the (1-6) linked disaccharide isomers show two peaks, indicating the presence of two distinct structures. The relative intensity of each peak was dependent on the collision energy used to generate the fragments (Fig. S4). There is also a pronounced shoulder present for the Fuc( $\alpha$ 1-3)GlcNAc, indicating a potential contribution from a second conformer (Fig. 2B, highlighted in yellow). Isolation of the 366/350 peaks with subsequent activation was used to produce and compare the ATDs of the  $m/z$  204 ion produced from each disaccharide. While the ATDs for the  $m/z$  204 fragments from both the (1-4) and (1-3)-linked precursors appear indistinguishable, the ATD from the (1-6) linked precursor is markedly distinct and for the Fuc( $\alpha$ 1-6)GlcNAc it is visibly broader (Figs. 2C, D).

To investigate the gas-phase structures of the fragment ions in greater detail, we applied LC-IM-MS experiments to resolve the  $\alpha/\beta$  anomers and performed the analysis using  $N_2$ , He, and  $CO_2$  as drift gasses to maximize our ability to resolve subtle structural differences<sup>35</sup>. Both anomers of the Gal-GlcNAc ( $m/z$  366) and Fuc-GlcNAcs ( $m/z$  350) yielded the same trends between the linkage isomers in all three drift gases, as the direct infusion data (Figs. 2E, G). The ATDs from the (1-6) linked disaccharides, particularly in  $N_2$  and  $CO_2$ , show clear evidence of multiple conformations that is consistent with multiple Gaussian distributions (Fig. S5). In several instances the  $\alpha/\beta$  anomers gave rise to distinct ATDs. This was most evident for Fuc( $\alpha$ 1-3)GlcNAc where the two anomers were distinct in all three drift gasses (Fig. 2G). The Gal( $\beta$ 1-3)GlcNAc also showed a slight difference between the anomers, most visible in  $N_2$  and  $CO_2$  (Fig. 2E). The ATDs for the (1-4) linked disaccharides were largely invariant among the anomers, with the possible exception of Gal( $\beta$ 1-4)GlcNAc in He, where the 1<sup>st</sup> eluting LC peak has a slightly later peak (Fig. 2E). The (1-6) linked ATDs did not show any obvious difference between the positions of the two

peaks (Fig. 2E), but there was a subtle effect on the relative magnitude of the two peaks (Fig. 2G).

Taking a step further, we examined the mobilities of the  $m/z$  204 ions formed from the  $m/z$  366 and 350 ions from the LC-resolved peaks. While the ATD from the (1-3) and (1-4) look indistinguishable, there is a clear shift in the ATDs from the (1-6) linked disaccharides (Figs. 2F, H), just as seen from the infusion IM-MS data. While most of the ATDs were indistinguishable between the  $\alpha/\beta$  anomers, the  $m/z$  204 ions produced from the (1-3) linked disaccharides show some differences. This is most evident in Gal( $\beta$ 1-3)GlcNAc in  $N_2$ , and for Fuc( $\alpha$ 1-3)GlcNAc in  $CO_2$ , and indicates that the anomeric configuration of a precursor can influence the structures of its fragment ions even after two stages of fragmentation. The  $m/z$  138 ion, which forms one specific structure regardless of the precursor<sup>22</sup>, was invariant in all the samples, confirming that even the subtle differences noted above are not attributed to variations in IM-MS conditions (Fig. S6). Lastly, we compared the  $m/z$  204 ATDs to that produced from monomeric GlcNAc (Fig. 2F, H, bottom panel). While the ATDs for  $m/z$  204 from the (1-4) and (1-3) linked disaccharides appear similar to the free GlcNAc in  $N_2$  and He, they are visibly different from free GlcNAc in  $CO_2$ .

#### *Characterization of linkage memory by gHDX*

gHDX was used as an orthogonal tool for resolving structural variations in the  $m/z$  366, 350, and 204 fragments. The deuterium exchange kinetics for the 350 and 366 ions from each of the disaccharides is different, consistent with their having different structures (Figs. 3A, B). Close inspection of the exchange spectra reveals that the Gal( $\beta$ 1-6)GlcNAc linked 366 ion exists as an ensemble of two dominant structures while the (1-3) and (1-4) linked structures show a predominant isotopic distribution along with a small population of a highly deuterated species (Fig. S7A). For the Fuc-GlcNAc  $m/z$  350 peaks, all of the linkages showed clear evidence of at least two populations, particularly at the longest exchange time points (Fig. S8A).

The  $m/z$  204 fragments generated from all the disaccharides show a maximum exchange of 4 deuterium, consistent with the number of exchangeable protons, but exhibit some notable

differences in their gHDX rates (Figs. 3C, D). For Gal-GlcNAc series, the kinetics for the  $m/z$  204 ion from the (1-3) and (1-4) linked precursors are within experimental error, while it is unique from Gal( $\beta$ 1-6)GlcNAc (Fig. 3C). Interestingly, the gHDX spectra of each  $m/z$  204 shows evidence of a second, highly deuterated population (Fig. S7B, 2.5 ms time point), also suggesting the presence of multiple structures. In the case of the Fuc-GlcNAcs, the gHDX profile for all the  $m/z$  204 were distinct (Fig. 3D).

We also examined the influence of the collision energy (CE) used to generate the fragment ions as this has been observed to influence gHDX kinetics<sup>28</sup>. The profiles of the 366 and 350 ions are indeed offset by altered CE, especially with the 1-6 linkage structures (Figs. S9A, B). In the case of the  $m/z$  204 ions, higher CE results in a drop in the extent of exchange that predominantly affects the shorter time points. The exchange profile at the longer time points still distinguishes the (1-4) linked  $m/z$  204 among Fuc-GlcNAcs and (1-6) linked  $m/z$  204 among Gal-GlcNAcs. To test whether the altered CE was directly affecting the kinetics of deuterium exchange or simply offsetting the distribution of different structures generated, we deconvoluted the isotopic distributions in the spectra that contained well-resolved populations (Figs. 3E, F). Raising the CE used to generate the  $m/z$  366, from 5 to 10, did not change the exchange kinetics of each individual population; rather it decreased the relative abundance of the highly deuterated population from 34 to 20 %. While the gHDX data for  $m/z$  204 did show evidence of multiple populations, these spectra could not be deconvoluted because of their insufficient separation (Fig. S7B, S8B).

## DISCUSSION

### *Multiple conformations of protonated disaccharides*

Computational and structural studies have provided a wealth of information on carbohydrate fragmentation<sup>10,29,36-40</sup>, however the vast potential structures of protonated carbohydrates and the details of their fragmentation mechanisms are not yet fully understood. Here we have examined a set of disaccharides structures that are common motifs within many

glycoconjugates. An emphasis was placed on the behavior of the water loss (B-type) fragments as these are generated in high abundance during CID of protonated glycoconjugates and are often used as diagnostic ions for identifying glycopeptides<sup>41</sup>. The <sup>18</sup>O labeling experiments (Fig. S1) confirm that the formation of B ions, *m/z* 366 (from Gal-GlcNAc) and 350 (from Fuc-GlcNAc) is by water loss occurring predominantly from the reducing end, consistent with several previous studies<sup>31,42</sup>. In both sets of disaccharides, these isomers were readily distinguishable by both IM-MS and gHDX, further showcasing each technique's ability to resolve carbohydrate isomers (Figs. 2A, B, E, G and 3A, B).

All of the 350 and 366 ions can sample multiple conformations. This is most evident from the IM-MS and gHDX data of the (1-6) linked structures, and to some extent in the (1-3) linked structures. The distribution of the two conformers of the (1-6) linked 366/350 ions is dependent on the collision energy used to generate the fragments (Figs. 3E, 3F, S4). Furthermore, gHDX shows that even the (1-4) and (1-3) linked Fuc-GlcNAc isomers adopt multiple conformers as evidenced by multiple isotopic distributions (Fig. S8A). Therefore, it is likely that all the 350 and 366 ions are comprised of multiple structures and IM-MS is simply unable to resolve some of them. These examples illustrate the highly orthogonal nature of gHDX and IM-MS. While gHDX differentiates ions based on their differences in gas-phase basicity and hydrogen bonding networks<sup>43,44</sup>, separation by IM-MS is based on the shape of the structure and its interactions with the buffer gas<sup>13,14</sup>. The large difference in the subpopulations seen by IM-MS for the (1-6) structures could be attributed to the extra rotational degree of freedom (O6-C6-C5), which bestows additional flexibility allowing a wider range of extended conformations, whereas the (1-3) and (1-4) are more constrained and can only access fewer conformations (details below).

The discovery of 'anomeric memory' has revealed a new paradigm where carbohydrate fragment structures are heavily influenced by the stereochemistry at leaving groups. Despite the loss of the reducing end, the original anomeric configuration influences the structure of B-type ions<sup>17,23</sup>. We directly observe this anomeric memory effect in the water-loss B ions in the anomer-resolved LC-MS data. Interestingly, the stereochemistry at the anomeric end influenced nearly all

of 350/366 ions as seen by IM-MS. In the case of the (1-4) and (1-3) linkages, the effect is a shift of the apparently single ATD, whereas with the (1-6) linkage, it influences the relative intensity of each partially-resolved peak (Figs. 2E, G). It is likely that the distribution of the (1-3) and (1-4) structures is also similarly offset by the anomeric configuration, but the ATDs of the resulting structures are poorly resolved and manifest as an apparent change in the single peak position. Overall, the anomer-resolved data further exemplifies how anomeric memory occurs in a wide range of protonated carbohydrates.

#### *Linkage memory in carbohydrate fragments*

The most intriguing finding from this study was obtained by examining the product ion resulting from a second round of fragmentation of the disaccharides. The  $m/z$  350 and 366 ions fragment primarily to an abundant Y-type ion corresponding to the GlcNAc that has already lost its reducing end hydroxyl ( $m/z$  204). By all methods employed, the structure(s) of the 204 ion generated from each linkage isomer was distinct, revealing that carbohydrate ions not only retain anomeric memory but also a *linkage memory*. Specifically, the 204 ions generated from the (1-6) disaccharides showed a longer ATD, indicating that they have an extended conformation compared to the other two 204 ions (Figs. 2C, D, F, H). The uniqueness of the 204 ion of the (1-6) linked disaccharides is also highlighted in subtle, but significant, differences in the gHDX profiles of the 204 ion (Figs. 3C, D), along with reproducible differences in their fragmentation, particularly with respect to the relative abundance of its fragments,  $m/z$  126 and 186 (Figs. 1E-H). Lastly, we note that both  $MS^n$  and IM-MS show that the behavior of the 204 generated from free GlcNAc monosaccharide is distinct from any of the equivalent fragments produced from the disaccharides (Figs. 1G, 1H, 2F, 2H), further corroborating that prior linkage strongly influences the resulting structure of the 204 ion.

In some cases, the properties of the 204 ion revealed both *linkage and anomeric* memory. Particularly with the (1-4) linked 350 and (1-3) linked 366 disaccharides, the fragmentation propensity of the 204 is distinct for the  $\alpha/\beta$  anomers, despite two rounds of fragmentation since

the initial loss of the reducing end. At this stage, the GlcNAc has lost both its reducing end, which confers the  $\alpha/\beta$  anomericity, and the sugar at the non-reducing end, which defined its linkage connectivity. Nevertheless, the properties of the 204 ion are characteristic of both its former anomericity and linkage connectivity, and thus it is encoded with memory from both ends. In fact, not only does the attachment at a particular linkage site impact the resulting Y-type fragment ions, the linkage stereochemistry and identity of the attached sugar moiety also appear to matter. The CID products of the 204 ion from Gal( $\beta$ 1-3)GlcNAc and Fuc( $\alpha$ 1-3)GlcNAc (Figs. 1E, F) are distinct, implying that the identity of non-reducing end monosaccharide ( $\beta$ -galactose versus  $\alpha$ -fucose) is key along with the linkage type (where the GlcNAc is linked to another monosaccharide). Further support is provided by the gHDX comparisons of the differently linked 204 ions of Gal-GlcNAc and Fuc-GlcNAcs, which show subtle differences for all six  $m/z$  204 ions (Fig. S9E). However, it is likely that some of the differences observed in gHDX profiles are a result of the varying dissociation thresholds of each disaccharide, which may generate a slightly different distribution of fragment ion structures, as seen for the 1-6 linked disaccharides (Figs. 3E, F).

#### *Structural rationalization of linkage memory*

The mechanisms underlying linkage memory tie back to the structure(s) of the  $m/z$  204 ion. Gray et al, proposed the presence of various bridged, open ring, and enol structures in the ensemble where the relative abundances of these structures is influenced by the original anomeric configuration<sup>17</sup>. Indeed the present studies together with previous structural studies of protonated GlcNAc<sup>10,22</sup> can be used to rationalize the linkage memory. With disaccharides, the initial loss of water to generate  $m/z$  350 or 366 may also lead to the formation of bridged structures (Fig. 4). In the case of the (1-4) linkage, there is no hydroxyl at the O4 position and thus no O4-C1 bridged structure at the GlcNAc can form. Similarly, the (1-6) linked structure cannot form the O6-C1 bridged structure. Previous mapping of the fragmentation pathways of GlcNAc revealed that the 204 ion dissociates through two competing pathways to generate either: 1) primarily  $m/z$  186 through water loss; or 2)  $m/z$  126 through a ring rearrangement reaction<sup>22</sup>. The O4-C1 bridge

structure was proposed to initiate the ring rearrangement to form  $m/z$  126. From our current data, the (1-4) linked disaccharides generated the lowest amount of 126 (Figs. 1E, F), which would be consistent with their inability to sample the O4-C1 bridged structure within their structural ensemble. In contrast, the (1-6) linked disaccharide generated the most 126 ion, which may generate relatively more O4-C1 since the formation of the O6-C1 pathway is not accessible.

The (1-6) linked disaccharides not being able to generate a O6-C1 structure is also consistent with the IM-MS data. The 204 ion from the (1-6) linkage, which will have free rotation at the C5-C6-O6, appears less compact, while the (1-3) and (1-4) species, which can form the O6-C1 structure to keep the O6 more constrained, appear more compact with shorter ATDs. We would like to emphasize that the reaction mechanisms of the formation of the 350/366 ions and its tentative structures proposed here are likely oversimplified, yet they provide a plausible explanation for the observed linkage memory in the subsequent 204 ions. More detailed and extensive work is required to characterize and identify the full ensemble of ion structures, which will also help verify the proposed reaction mechanisms.

The direct observation of multiple gas-phase structures from several disaccharides further strengthens the existing argument that carbohydrate fragment ions exist as ensembles of structures<sup>10,17,38,40</sup>. More importantly, the structural ensemble of both B- and Y-type fragments are influenced by the identity of their precursors and both the reducing and non-reducing ends. Even an apparently simple chemically homogenous structure can form several conformers that are poorly, partially, or well-resolved in the gas-phase. Furthermore, the presence of multiple conformers of both parent and fragment ions will have to be considered when benchmarking properties of carbohydrate ions, such as IM databases seeking to standardize collision cross section values. Though protonated carbohydrates have proven more complex than metal adducted or derivatized carbohydrates, a fundamental understanding of carbohydrate fragments and the principles that govern their formation will be necessary not only for structural studies, but for the development of much needed technologies to characterize the complex biological functions of the diverse glycan and glycoconjugate repertoire. Overall, the diagnostic CID

patterns, ion mobilities, and gHDX behavior of the differently linked Y fragments, illustrate the potential of fragment ions in determining intact precursor carbohydrate structures.

## ACKNOWLEDGEMENTS

The authors wish to thank Dale Whittington, and J. Scott Edgar for assistance with data collection and Benjamin Bythell and Isabel Compagnon for insightful discussions. This work was funded by NIH NIGMS grant R01GM127579 and University of Washington Royalty Research Fund award (#A118776).

## SUPPORTING INFORMATION

Figures S1-S10, showing the structures of disaccharides and fragmentation of  $^{18}\text{O}$  labeled Fuc-GlcNAc; HILIC separation of the anomers of Gal-GlcNAc and Fuc-GlcNAc; LC-MS<sup>3</sup> of the  $\alpha/\beta$  anomers of Gal-GlcNAc  $m/z$  366, and Fuc-GlcNAc  $m/z$  350 ions; collision energy dependence on ATDs of (1-6) linked disaccharide  $m/z$  350 and 366 ions; deconvolution of the ATDs of Gal( $\beta$ 1-6)GlcNAc  $m/z$  366, and Fuc( $\alpha$ 1-6)GlcNAc  $m/z$  350; ATDs of the  $m/z$  138 ion of Gal(1-x)GlcNAc; gHDX spectra of Gal( $\beta$ 1-x)GlcNAc  $m/z$  366 and 204; gHDX spectra of Gal( $\beta$ 1-x)GlcNAc  $m/z$  350 and 204; collision energy effects on the gHDX kinetics of the  $m/z$  366, 350 and 204 fragments, CE effects on ATDs of (1-6) linked  $m/z$  366/350, and NMR spectra of six disaccharides.

## REFERENCES

- (1) Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3-49.
- (2) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in health and disease. *Nature Reviews Nephrology* **2019**, *15*, 346-366.

- (3) Gahmberg, C. G.; Tolvanen, M. Why mammalian cell surface proteins are glycoproteins. *Trends in Biochemical Sciences* **1996**, *21*, 308-311.
- (4) Varki, A. Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. *Cold Spring Harb Perspect Biol* **2011**, *3*.
- (5) Sheeley, D. M.; Reinhold, V. N. Structural Characterization of Carbohydrate Sequence, Linkage, and Branching in a Quadrupole Ion Trap Mass Spectrometer: Neutral Oligosaccharides and N-Linked Glycans. *Analytical Chemistry* **1998**, *70*, 3053-3059.
- (6) Wuhrer, M.; Koeleman, C. A. M.; Hokke, C. H.; Deelder, A. M. Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Communications in Mass Spectrometry* **2006**, *20*, 1747-1754.
- (7) Mucha, E.; Lettow, M.; Marianski, M.; Thomas, D. A.; Struwe, W. B.; Harvey, D. J.; Meijer, G.; Seeberger, P. H.; von Helden, G.; Pagel, K. Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry. *Angewandte Chemie International Edition* **2018**, *57*, 7440-7443.
- (8) Sastre Toraño, J.; Gagarinov, I. A.; Vos, G. M.; Broszeit, F.; Srivastava, A. D.; Palmer, M.; Langridge, J. I.; Aizpurua-Olaizola, O.; Somovilla, V. J.; Boons, G.-J. Ion-Mobility Spectrometry Can Assign Exact Fucosyl Positions in Glycans and Prevent Misinterpretation of Mass-Spectrometry Data After Gas-Phase Rearrangement. *Angewandte Chemie International Edition* **2019**, *58*, 17616-17620.
- (9) Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K. The role of the mobile proton in fucose migration. *Analytical and Bioanalytical Chemistry* **2019**, *411*, 4637-4645.
- (10) Bythell, B. J.; Abutokaikah, M. T.; Wagoner, A. R.; Guan, S.; Rabus, J. M. Cationized Carbohydrate Gas-Phase Fragmentation Chemistry. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 688-703.

- (11) Pap, A.; Klement, E.; Hunyadi-Gulyas, E.; Darula, Z.; Medzihradszky, K. F. Status Report on the High-Throughput Characterization of Complex Intact O-Glycopeptide Mixtures. *Journal of the American Society for Mass Spectrometry* **2018**, *29*, 1210-1220.
- (12) Schumacher, K. N.; Dodds, E. D. A case for protein-level and site-level specificity in glycoproteomic studies of disease. *Glycoconjugate Journal* **2016**, *33*, 377-385.
- (13) Clowers, B. H.; Dwivedi, P.; Steiner, W. E.; Hill, H. H.; Bendiak, B. Separation of Sodiated Isobaric Disaccharides and Trisaccharides Using Electrospray Ionization-Atmospheric Pressure Ion Mobility-Time of Flight Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 660-669.
- (14) Dwivedi, P.; Bendiak, B.; Clowers, B. H.; Hill, H. H. Rapid Resolution of Carbohydrate Isomers by Electrospray Ionization Ambient Pressure Ion Mobility Spectrometry-Time-of-Flight Mass Spectrometry (ESI-APIMS-TOFMS). *Journal of the American Society for Mass Spectrometry* **2007**, *18*, 1163-1175.
- (15) Brodbelt, J. S. Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chemical Society reviews* **2014**, *43*, 2757-2783.
- (16) Tan, Y.; Polfer, N. C. Linkage and Anomeric Differentiation in Trisaccharides by Sequential Fragmentation and Variable-Wavelength Infrared Photodissociation. *Journal of the American Society for Mass Spectrometry* **2015**, *26*, 359-368.
- (17) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. Bottom-Up Elucidation of Glycosidic Bond Stereochemistry. *Analytical Chemistry* **2017**, *89*, 4540-4549.
- (18) Mucha, E.; González Flórez, A. I.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schöllkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K. Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angewandte Chemie (International ed. in English)* **2017**, *56*, 11248-11251.

- (19) Riggs, D. L.; Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K.; Julian, R. R. Glycan Isomer Identification Using Ultraviolet Photodissociation Initiated Radical Chemistry. *Analytical Chemistry* **2018**, *90*, 11581-11588.
- (20) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Analytical Chemistry* **2019**, *91*, 4876-4882.
- (21) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L. Advancing Solutions to the Carbohydrate Sequencing Challenge. *Journal of the American Chemical Society* **2019**, *141*, 14463-14479.
- (22) Mookherjee, A.; Uppal, S. S.; Guttman, M. Dissection of Fragmentation Pathways in Protonated N-Acetylhexosamines. *Analytical Chemistry* **2018**, *90*, 11883-11891.
- (23) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *Journal of the American Society for Mass Spectrometry* **2019**, *30*, 1028-1037.
- (24) Wojcik, R.; Webb, I. K.; Deng, L.; Garimella, S. V.; Prost, S. A.; Ibrahim, Y. M.; Baker, E. S.; Smith, R. D. Lipid and Glycolipid Isomer Analyses Using Ultra-High Resolution Ion Mobility Spectrometry Separations. *International journal of molecular sciences* **2017**, *18*.
- (25) Scutelnic, V.; Rizzo, T. R. Cryogenic Ion Spectroscopy for Identification of Monosaccharide Anomers. *The Journal of Physical Chemistry A* **2019**, *123*, 2815-2819.
- (26) Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. Cryogenic IR spectroscopy combined with ion mobility spectrometry for the analysis of human milk oligosaccharides. *Analyst* **2018**, *143*, 1846-1852.

- (27) Warnke, S.; Ben Faleh, A.; Pellegrinelli, R. P.; Yalovenko, N.; Rizzo, T. R. Combining ultra-high resolution ion mobility spectrometry with cryogenic IR spectroscopy for the study of biomolecular ions. *Faraday discussions* **2019**, *217*, 114-125.
- (28) Uppal, S. S.; Beasley, S. E.; Scian, M.; Guttman, M. Gas-Phase Hydrogen/Deuterium Exchange for Distinguishing Isomeric Carbohydrate Ions. *Analytical Chemistry* **2017**, *89*, 4737-4742.
- (29) Struwe, W. B.; Baldauf, C.; Hofmann, J.; Rudd, P. M.; Pagel, K. Ion mobility separation of deprotonated oligosaccharide isomers – evidence for gas-phase charge migration. *Chemical Communications* **2016**, *52*, 12353-12356.
- (30) Alonge, K. M.; Logsdon, A. F.; Murphree, T. A.; Banks, W. A.; Keene, C. D.; Edgar, J. S.; Whittington, D.; Schwartz, M. W.; Guttman, M. Quantitative analysis of chondroitin sulfate disaccharides from human and rodent fixed brain tissue by electrospray ionization-tandem mass spectrometry. *Glycobiology* **2019**, *29*, 847-860.
- (31) Viseux, N.; de Hoffmann, E.; Domon, B. Structural Assignment of Permethylated Oligosaccharide Subunits Using Sequential Tandem Mass Spectrometry. *Analytical Chemistry* **1998**, *70*, 4951-4959.
- (32) Uppal, S. S.; Mookherjee, A.; Harkewicz, R.; Beasley, S. E.; Bush, M. F.; Guttman, M. High-Precision, Gas-Phase Hydrogen/Deuterium-Exchange Kinetics by Mass Spectrometry Enabled by Exchange Standards. *Analytical chemistry* **2020**, *92*, 7725-7732.
- (33) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical Chemistry* **2015**, *87*, 4370-4376.
- (34) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of Overlapped and Noisy Hydrogen/Deuterium Exchange Mass Spectra. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 1906-1912.

- (35) Davidson, K. L.; Bush, M. F. Effects of Drift Gas Selection on the Ambient-Temperature, Ion Mobility Mass Spectrometry Analysis of Amino Acids. *Anal Chem* **2017**, *89*, 2017-2023.
- (36) Rabus, J. M.; Abutokaikah, M. T.; Ross, R. T.; Bythell, B. J. Sodium-cationized carbohydrate gas-phase fragmentation chemistry: influence of glycosidic linkage position. *Physical Chemistry Chemical Physics* **2017**, *19*, 25643-25652.
- (37) Mucha, E.; Stuckmann, A.; Marianski, M.; Struwe, W. B.; Meijer, G.; Pagel, K. In-depth structural analysis of glycans in the gas phase. *Chemical Science* **2019**, *10*, 1272-1284.
- (38) Rossich Molina, E.; Eizaguirre, A.; Haldys, V.; Urban, D.; Doisneau, G.; Bourdreux, Y.; Beau, J.-M.; Salpin, J.-Y.; Spezia, R. Characterization of Protonated Model Disaccharides from Tandem Mass Spectrometry and Chemical Dynamics Simulations. *ChemPhysChem* **2017**, *18*, 2812-2823.
- (39) Yamagaki, T.; Fukui, K.; Tachibana, K. Analysis of Glycosyl Bond Cleavage and Related Isotope Effects in Collision-Induced Dissociation Quadrupole-Time-of-Flight Mass Spectrometry of Isomeric Trehaloses. *Analytical Chemistry* **2006**, *78*, 1015-1022.
- (40) Fentabil, M. A.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Blackbody Infrared Radiative Dissociation of Protonated Oligosaccharides. *Journal of the American Society for Mass Spectrometry* **2011**, *22*, 2171-2178.
- (41) Nwosu, C. C.; Seipert, R. R.; Strum, J. S.; Hua, S. S.; An, H. J.; Zivkovic, A. M.; German, B. J.; Lebrilla, C. B. Simultaneous and Extensive Site-specific N- and O-Glycosylation Analysis in Protein Mixtures. *Journal of Proteome Research* **2011**, *10*, 2612-2624.
- (42) Hofmeister, G. E.; Zhou, Z.; Leary, J. A. Linkage position determination in lithium-cationized disaccharides: tandem mass spectrometry and semiempirical calculations. *Journal of the American Chemical Society* **1991**, *113*, 5964-5970.
- (43) Reyzer, M. L.; Brodbelt, J. S. Gas-phase H/D exchange reactions of polyamine complexes: (M + H)<sup>+</sup>, (M + alkali metal)<sup>+</sup>, and (M + 2H)<sup>2+</sup>. *Journal of the American Society for Mass Spectrometry* **2000**, *11*, 711-721.

(44) Wyttenbach, T.; Bowers, M. T. Gas phase conformations of biological molecules: The hydrogen/deuterium exchange mechanism. *Journal of the American Society for Mass Spectrometry* **1999**, *10*, 9-14.

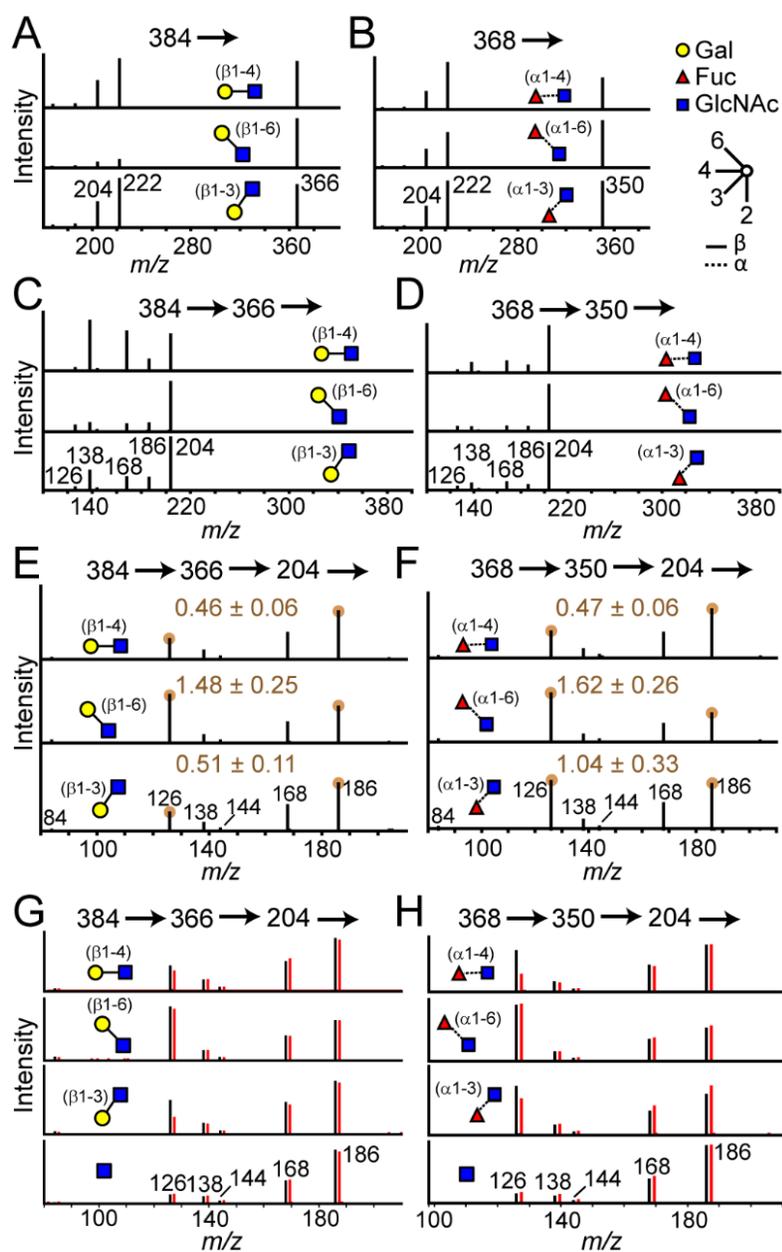


Figure 1. MS<sup>n</sup> spectra of Gal(β1-4)GlcNAc (top), Gal(β1-6)GlcNAc (middle), and Gal(β1-3)GlcNAc (bottom) are shown in (A, C, E). MS<sup>n</sup> spectra of Fuc(α1-4)GlcNAc (top), Fuc(α1-6)GlcNAc (middle), and Fuc(α1-3)GlcNAc (bottom) are shown in (B, D, F). The masses in each panel: *m/z* 384, 368, 366, 350, and 204, have been selected for each stage of MS<sup>n</sup> shown by the arrows. Structures of each disaccharide are shown in the panel. The symbols show the three monosaccharides GlcNAc (in blue), galactose (in yellow) and fucose (in red) that form the six disaccharides, while the solid/dashed lines denote the glycosidic bond between the monosaccharides with the angle of the line specifying a linkage configuration (see legend). Solid lines denote a beta (β) linkage while the dashed lines denote the alpha (α) linkage along with the originating carbon atom and ending carbon atom numbered involved in the linkage. (G, H) Anomers of the disaccharides were resolved by HILIC and analyzed online by MS<sup>4</sup>. The corresponding MS<sup>3</sup> data are shown in Figure S3. The spectra correspond to the intensities of the same ions from the early (black) and late (red) eluting LC peaks. GlcNAc monomer analyzed by the same HILIC-MS method is shown on the bottom spectra (222 → 204 →).

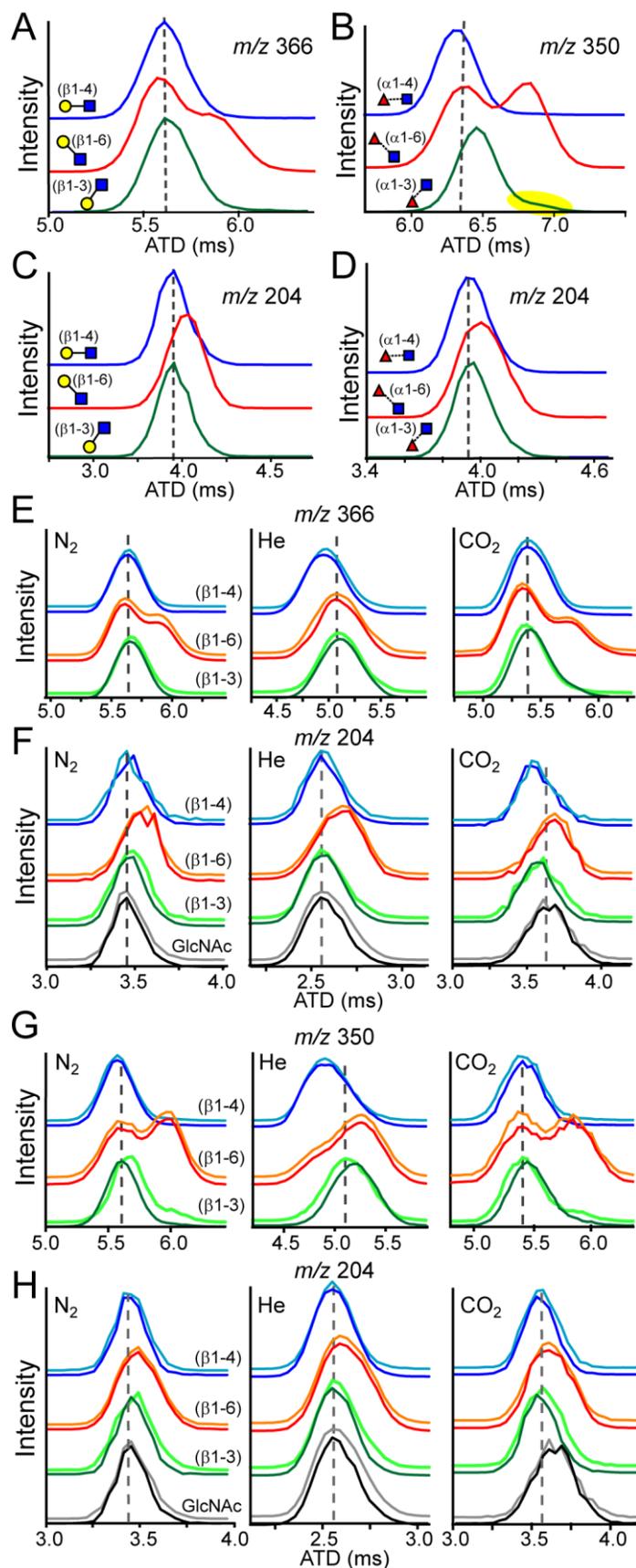


Figure 2. The arrival time distributions (ATDs) in  $N_2$  of (A)  $m/z$  366, (B)  $m/z$  350, (C)  $m/z$  204 from 366, and (D)  $m/z$  204 from 350. IM- $MS^n$  analysis of LC resolved anomers of Gal( $\beta 1-x$ )GlcNAc and Fuc( $\alpha 1-x$ )GlcNAc are shown in panel (E-H). ATDs of (E)  $m/z$  366 and (F)  $m/z$  204 of Gal-GlcNAc, (G)  $m/z$  350 and (H)  $m/z$  204 of the Fuc-GlcNAc disaccharides are shown in using  $N_2$ , He, and  $CO_2$  as the drift gas. The early eluting peak is shown in light blue (1-3), orange (1-6), light green (1-3), and gray (GlcNAc) while the later eluting peak is dark blue, red, dark green, and black. Dashed lines are included to aid in visualizing minor differences in peak positions.

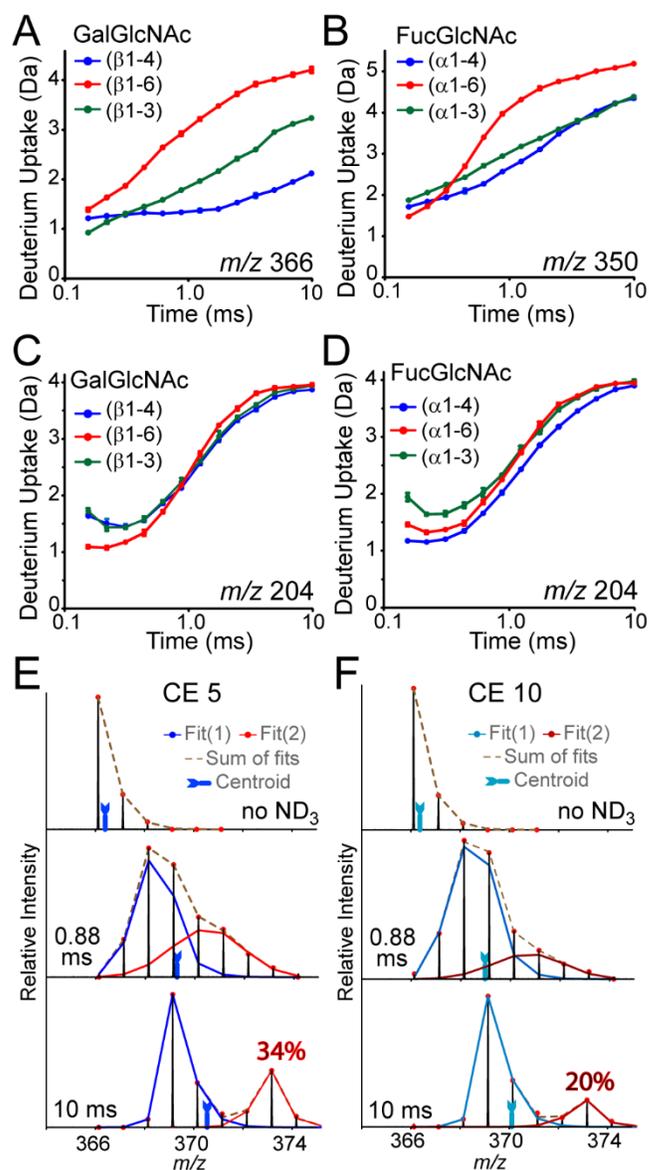


Figure 3. gHDX analysis of  $m/z$  366 (A),  $m/z$  204 (C) of Gal( $\beta$ 1-x)GlcNAc disaccharides and  $m/z$  350 (B),  $m/z$  204 (D) of Fuc( $\alpha$ 1-x)GlcNAc disaccharides. Deuterium uptake for the various fragment ions is shown as a function of incubation time with ND<sub>3</sub>. The points and error bars represent the average and standard deviations from triplicate measurements. (E/F) Example mass gHDX spectra from the  $m/z$  366 peak of Gal( $\beta$ 1-6)GlcNAc generated from the precursor ion ( $m/z$  386) using a CE of 5 (E) or 10 (F). The blue and red fits indicate the two populations that are resolved, and the relative abundance of the red population is indicated in the inset.

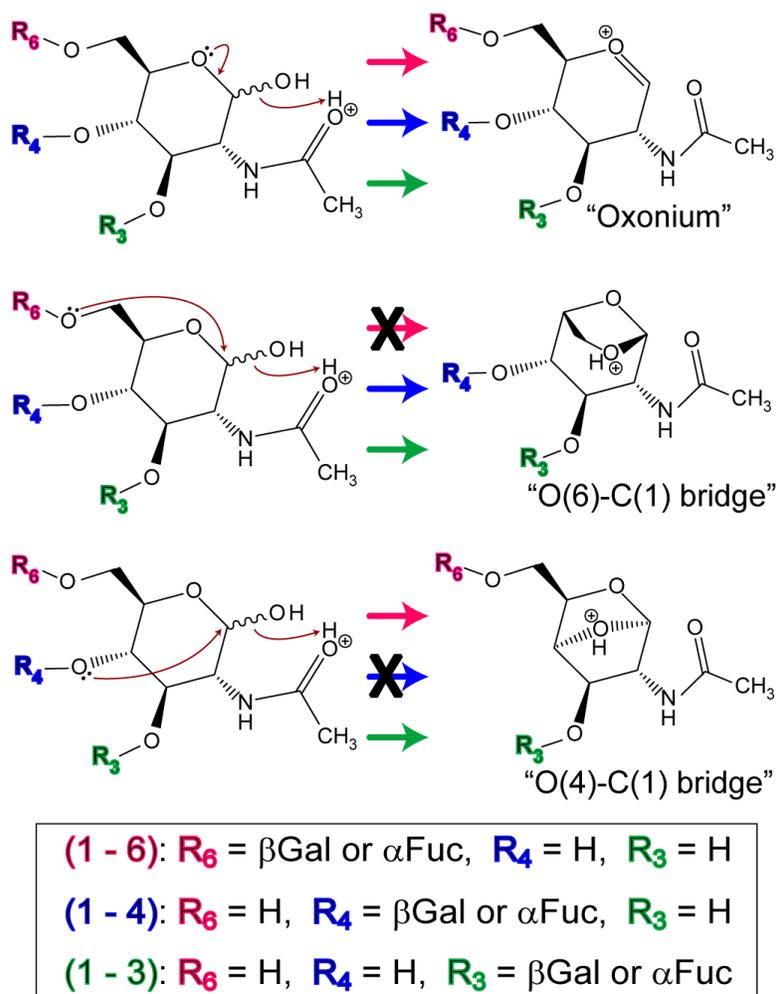


Figure 4. Proposed structural rationalization of linkage memory. GlcNAc linked to another monosaccharide at either the 6 (magenta), 4 (blue), or 3 (green) position undergoes a loss of the reducing end hydroxyl upon fragmentation. The legend at the bottom indicates the positions of the  $\beta$ -Gal or  $\alpha$ -Fuc in the disaccharides used in this study. Red arrows indicate potential mechanisms that drive formation of either the classic oxonium structure (top), a O6-C1 linked bridge structure (middle), or a O4-C1 linked bridge structure (bottom). Due to the linkage configuration, the (1-6) linked disaccharides cannot form the O6-C1 bridge structure, and likewise the (1-4) linked disaccharide cannot form the O4-C1 bridge structure (pathways denoted by Xs).