Factors That Influence Fragmentation Behavior of N-Linked Glycopeptide Ions

ARTICLE *in* ANALYTICAL CHEMISTRY · MAY 2008

Impact Factor: 5.64 · DOI: 10.1021/ac800067y · Source: PubMed

6 AUTHORS, INCLUDING:

Eric D. Dodds
University of Nebraska at Lincoln
38 PUBLICATIONS **638** CITATIONS
SEE PROFILE

Bruce German
University of California, Davis
430 PUBLICATIONS **17,879** CITATIONS
SEE PROFILE

Anal. Chem. **2008,** *80,* 3684–3692

Factors That Influence Fragmentation Behavior of N-Linked Glycopeptide Ions

Richard R. Seipert,[†] Eric D. Dodds,[†] Brian H. Clowers,[†] Sean M. Beecroft,[†] J. Bruce German,[‡] and Carlito B. Lebrilla*[,†,§]

Department of Chemistry, Department of Food Science and Technology, and School of Medicine, University of California Davis, One Shields Avenue, Davis, California 95616

The investigation of site-specific glycosylation is essential for further understanding the many biological roles that glycoproteins play; however, existing methods for characterizing site-specific glycosylation either are slow or yield incomplete information. Mass spectrometry (MS) is being applied to investigate site-specific glycosylation with bottom-up proteomic type strategies. When using these approaches, tandem mass spectrometry techniques are often essential to verify glycopeptide composition, minimize false positives, and investigate structure. The fragmentation behavior of glycopeptide ions has previously been investigated with multiple techniques including collision induced dissociation (CID), infrared multiphoton dissociation (IRMPD) and electron capture dissociation (ECD); however, due to the almost exclusive analysis of multiply protonated tryptic glycopeptide ions, some dissociation behaviors of N-linked glycopeptide ions have not been fully elucidated. In this study, IRMPD of N-linked glycopeptides has been investigated with a focus on the effects of charge state, charge carrier, glycan composition, and peptide composition. Each of these parameters was shown to influence the fragmentation behavior of N-linked glycopeptide ions. For example, in contrast to previously reported accounts that IRMPD results only in glycosidic bond cleavage, the fragmentation of singly protonated glycopeptide ions containing a basic amino acid residue almost exclusively resulted in peptide backbone cleavage. The fragmentation of the doubly protonated glycopeptide ion exhibited fragmentation similar to that previously reported; however, when the same glycopeptide was sodium coordinated, a previously inaccessible series of glycan fragments were observed. Molecular modeling calculations suggest that differences in the site of protonation and metal ion coordination may direct glycopeptide ion fragmentation.

Glycosylation is one of the most common post-translational modifications of proteins, but is often the least understood, due to the difficulty of elucidating the structures of these complex glycans.[1] Nonetheless, the diverse nature of protein glycosylation likely controls a variety of biological roles with which these modifications have been associated.[2,3] N-linked glycans, those attached to asparagine residues, are one of the most common forms of protein glycosylation. It is known that N-linked glycosylation occurs at a consensus sequence of NXS, NXT, and less commonly NXC (where X is any amino acid except proline) with high fidelity.[4] However, glycosylation at a given site is not obligatory, and in fact a specific site may be unoccupied in one copy of the protein and occupied with differing glycans in other copies, or glycoforms. Advanced analytical techniques capable of addressing this complexity continue to be a key area of research.

Mass spectrometry can be used to investigate the intricacy of protein glycosylation due to its high sensitivity and ability to help elucidate molecular composition using tandem MS techniques.[5–8] Enzymatic or chemical release of glycans and subsequent analysis with mass spectrometry have proven useful for the identification of protein-bound glycans; however site-specific information is routinely lost with these approaches.[9–14] In order to preserve information on the site-specific glycosylation and corresponding site heterogeneity, it is essential that the glycans remain covalently linked to their respective sites throughout the analysis. As a result, several methods have been developed using both specific and

* Corresponding author. Telephone: 1-530-752-6364. Fax: 1-530-754-5609. E-mail: cblebrilla@ucdavis.edu.

[†] Department of Chemistry.

[‡] Department of Food Science and Technology.

[§] School of Medicine.

(1) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24*, 1241–1252.
(2) Murrell, M. P.; Yarema, K. J.; Levchenko, A. *ChemBioChem* **2004**, *5*, 1334–1347.
(3) Varki, A. *Glycobiology* **1993**, *3*, 97–130.
(4) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R–56R.
(5) Morris, H. R.; Chalabi, S.; Panico, M.; Sutton-Smith, M.; Clark, G. F.; Goldberg, D.; Dell, A. *Int. J. Mass Spectrom.* **2007**, *259*, 16–31.
(6) Harvey, D. J. *Proteomics* **2005**, *5*, 1774–1786.
(7) Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J. C. *Proteomics* **2006**, *6*, 3993–4015.
(8) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. *J. Chromatogr. B* **2007**, *849*, 115–128.
(9) Alvarez-Manilla, G.; Atwood, J.; Guo, Y.; Warren, N. L.; Orlando, R.; Pierce, M. *J. Proteome Res.* **2006**, *5*, 701–708.
(10) Lattova, E.; Chen, V. C.; Varma, S.; Bezabeh, T.; Perreault, H. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1644–1650.
(11) Ramachandran, P.; Boontheung, P.; Xie, Y. M.; Sondej, M.; Wong, D. T.; Loo, J. A. *J. Proteome Res.* **2006**, *5*, 1493–1503.
(12) Faid, V.; Evjen, G.; Tollersrud, O. K.; Michalski, J. C.; Morelle, W. *Glycobiology* **2006**, *16*, 440–461.
(13) Hakansson, K.; Emmett, M. R.; Marshall, A. G.; Davidsson, P.; Nilsson, C. L. *J. Proteome Res.* **2003**, *2*, 581–588.
(14) Wada, Y.; Azadi, P.; Costello, C. E.; Dell, A.; Dwek, R. A.; Geyer, H.; Geyer, R.; Kakehi, K.; Karlsson, N. G.; Kato, K.; Kawasaki, N.; Khoo, K. H.; Kim, S.; Kondo, A.; Lattova, E.; Mechref, Y.; Miyoshi, E.; Nakamura, K.; Narimatsu, H.; Novotny, M. V.; Packer, N. H.; Perreault, H.; Peter-Katalinic, J.; Pohlentz, G.; Reinhold, V. N.; Rudd, P. M.; Suzuki, A.; Taniguchi, N. *Glycobiology* **2007**, *17*, 411–422.

10.1021/ac800067y CCC: $40.75
Published on Web 03/26/2008

nonspecific proteolytic enzymes to digest the glycoproteins.[15,16] Methods that utilize nonspecific protease enzymes have shown potential for these analyses due to the fact that the overall catalytic activity of these enzymes leads to a peptide footprint surrounding the site of glycosylation while reducing the rest of the protein to individual amino acids or dipeptides. This simple enzymatic process effectively results in enrichment of glycopeptides from protein samples.[17,18] In addition, eliminating nonglycosylated peptides reduces suppression of glycopeptide ionization during mass spectral analysis, which could prevent detection of stoichiometrically minor and less efficiently ionized glycopeptides. Even with relatively pure glycopeptide preparations, it can be difficult to assign composition based solely upon even the most accurate mass measurements due to the vast number of glycopeptide compositional possibilities that can exist within a very narrow mass range.[19,20] Hence, tandem mass spectrometry is often essential to elucidate and verify both composition and structure.

Tandem mass spectrometry has been applied to the structural interrogation of an increasing number of biomolecules including oligosaccharides, peptides, and glycopeptides.[6,8,14,20–24] There are multiple forms of tandem mass spectrometry; however, all of them can be divided into two major groups based upon the method of energy deposition: namely, vibrational and electronic. Vibrational excitation includes infrared multiphoton dissociation (IRMPD) and low-energy collision-induced dissociation (CID), which have both been used to analyze peptides and oligosaccharides. These techniques mainly fragment the molecules along the most labile peptide or glycosidic bonds.[6,23,25–27] Because CID and IRMPD are ergodic processes that allow time for the deposited energy to be distributed throughout the precursor, only the lowest activation barriers for fragmentation are typically overcome. Fragmentation methods based on electronic excitation such as electron capture dissociation (ECD) and electron transfer dissociation (ETD) result in nonergodic dissociation; that is, the dissociation takes place on a rapid time scale, prior to the randomization of energy.[28] This type of dissociation has proven useful for the identification of modified amino acids during peptide sequencing since the product ions retain labile modifications.[22,29] Importantly, ECD and ETD require that precursor ions be multiply charged making these techniques incompatible with ions generated by matrix-assisted laser desorption/ionization (MALDI). Moreover, nonspecific glycopeptides with small peptide moieties may be less prone to multiple charging as compared to larger tryptic glycopeptides.

The application of these fragmentation techniques to the investigation of glycopeptides has greatly enhanced the ability to identify site occupancy in site-specific glycosylation. Glycopeptide ion fragmentation has been found to vary for different instruments and tandem MS methods.[8,24] The use of CID in time-of-flight mass spectrometry (TOF-MS) employing both MALDI and electrospray ionization (ESI) has been demonstrated to induce both glycosidic bond cleavages and peptide bond cleavages. However, MALDI-TOF was more efficient in that both types of fragment ions could be observed in a single MS/MS event.[30–33] Ion trap mass spectrometers (IT-MS) have been used to obtain glycan fragments via CID and more recently peptide backbone fragments via ETD.[20,34–37] The combination of ECD and IRMPD in FT-ICR MS has allowed for the complete fragmentation of glycopeptide ions coupled with the high mass accuracy and resolution inherent to FT-ICR.[38]

ECD allows selective cleavage of the peptide bonds while the intact glycan moiety remains bound to the asparagine side chain.[39] IRMPD and SORI-CID, on the other hand, have been reported to characteristically yield only glycosidic cleavages while leaving the peptide moiety intact.[40] In some cases, IRMPD has been observed to promote peptide backbone cleavage; however, the circumstances that lead to the fragmentation of the peptide backbone cleavage over the attached glycan and vice versa are not well understood.[38]

There have been several studies reported on glycopeptide ion dissociation; however, there is little consensus as to the factors that guide fragmentation. Some have suggested that ionization methods may direct fragmentation of glycopeptide ions due to the observation that glycopeptide ions created by MALDI and ESI fragmented distinctly when subjected to vibrational excitation methods.[8] These differences in fragmentation were attributed vaguely to the ionization techniques; however, the differences were more likely due to the differences in charge states (multiply

(15) An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 5628–5637.
(16) Larsen, M. R.; Hojrup, P.; Roepstorff, P. *Mol. Cell. Proteomics* **2005**, *4*, 107–119.
(17) Clowers, B. H.; Dodds, E. D.; Seipert, R. R.; Lebrilla, C. B. *J. Proteome Res.* **2007**, *6*, 4032–4040.
(18) Temporini, C.; Perani, E.; Calleri, E.; Dolcini, L.; Lubda, D.; Caccialanza, G.; Massolini, G. *Anal. Chem.* **2007**, *79*, 355–363.
(19) Goldberg, D.; Bern, M.; Parry, S.; Sutton-Smith, M.; Panico, M.; Morris, H. R.; Dell, A. *J. Proteome Res.* **2007**, *6*, 3995–4005.
(20) Ren, J. M.; Rejtar, T.; Li, L.; Karger, B. L. *J. Proteome Res.* **2007**, *6*, 3162–3173.
(21) Zhang, J. H.; Schubothe, K.; Li, B. S.; Russell, S.; Lebrilla, C. B. *Anal. Chem.* **2005**, *77*, 208–214.
(22) Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, *15*, 12–16.
(23) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508–548.
(24) Jiang, H.; Desaire, H.; Butnev, V. Y.; Bousfield, G. R. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 750–758.
(25) Lancaster, K. S.; An, H. J.; Li, B. S.; Lebrilla, C. B. *Anal. Chem.* **2006**, *78*, 4990–4997.
(26) Xie, Y. M.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 1590–1598.
(27) Seipert, R. R.; Barboza, M.; Ninonuevo, M. R.; LoCascio, R. G.; Mills, D. A.; Freeman, S. L.; German, J. B.; Lebrilla, C. B. *Anal. Chem.* **2008**, *80*, 159–165.
(28) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.
(29) Kelleher, R. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250–4253.
(30) Bykova, N. V.; Rampitsch, C.; Krokhin, O.; Standing, K. G.; Ens, W. *Anal. Chem.* **2006**, *78*, 1093–1103.
(31) Ito, H.; Takegawa, Y.; Deguchi, K.; Nagai, S.; Nakagawa, H.; Shinohara, Y.; Nishimura, S. I. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 3557–3565.
(32) Yu, Y. Q.; Fournier, J.; Gilar, M.; Gebler, J. C. *Anal. Chem.* **2007**, *79*, 1731–1738.
(33) Harazono, A.; Kawasaki, N.; Itoh, S.; Hashii, N.; Ishii-Watabe, A.; Kawanishi, T.; Hayakawa, T. *Anal. Biochem.* **2006**, *348*, 259–268.
(34) Medzihradszky, K. F. *Methods Enzymol.* **2005**, *405*, 116–138.
(35) Catalina, M. I.; Koeleman, C. A. M.; Deelder, A. M.; Wuhrer, M. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1053–1061.
(36) Peterman, S. M.; Mulholland, J. J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 168–179.
(37) Irungu, J.; Dalpathado, D. S.; Go, E. P.; Jiang, H.; Ha, H. V.; Bousfield, G. R.; Desaire, H. *Anal. Chem.* **2006**, *78*, 1181–1190.
(38) Adamson, J. T.; Hakansson, K. *J. Proteome Res.* **2006**, *5*, 493–501.
(39) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530–4536.
(40) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256–3262.

charged versus singly charged) and possibly the internal energies of the produced glycopeptide ions. Earlier studies have also shown that the nature of the fragmentation of oligosaccharides is dependent upon the coordinating cation, suggesting that glycopeptide ions may yield similar variations.[26]

In order to better understand the fragmentation of glycopeptide ions, we examined the influences of charge state, coordinating ion, glycan composition, and amino acid composition on dissociation. By producing and collisionally cooling the produced ions in the analyzer cell, we separate the energetics of ion formation and vibrational activation. In addition, molecular modeling of selected glycopeptide ions was used as an aid to rationalize the experimentally observed dissociation behaviors.

EXPERIMENTAL SECTION

Materials and Reagents. Bovine ribonuclease B (RNase B; SwissProt accession number P61823), chicken egg albumin (CEA; SwissProt accession number P01012), Pronase E, ammonium acetate, ethanolamine, hydrochloric acid, and formic acid were obtained from Sigma-Aldrich (St. Louis, MO). Cyanogen bromide activated Sepharose 4B beads (average particle diameter, 90 μm) were acquired from GE Healthcare (Piscataway, NJ). All other reagents were of analytical grade or higher.

Glycoprotein Digestion. Pronase E was immobilized on Sepharose beads following the procedure described by Clowers et al.[17] In short, 150 mg of cyanogen bromide activated Sepharose beads was hydrated with HCl (1 mM), rinsed with phosphate buffer (100 mM), combined with Pronase (1 mg), and allowed to couple for 24 h. The remaining active sites on the Sepharose beads were then blocked with ethanolamine (1 M) and rinsed with ammonium acetate buffer (100 mM). The Pronase beads were then ready for glycoprotein digestion. Stock glycoprotein solutions were prepared at a concentration of approximately 50 μM in ammonium acetate buffer (100 mM). For nonspecific proteolysis, the glycoprotein stock solution (100 μL) and fresh ammonium acetate buffer (200 μL) were added to the Pronase coupled beads. The digestion was allowed to proceed at 37 °C with gentle mixing, during which the supernatant was sampled at multiple time points ranging from 90 min to 24 h. In order to avoid the transfer of any Pronase coupled beads, the digested solution was centrifuged before sampling the supernatant. Glycopeptide solutions were then analyzed immediately or stored at −20 °C prior to analysis.

No sample cleanup was performed on the digestion supernatants; instead, they were diluted into an electrospray-compatible solvent system for direct MS analysis. All digests were diluted at least 10-fold to a final composition of 50% aqueous acetonitrile buffered with either formic acid (0.1%, pH 3.5) or ammonium acetate (1 mM, pH 7.2), depending on the desired precursor ion. After removal of the supernatant, the Pronase coupled beads were washed and reused for subsequent digestions.

Instrumentation and Analysis. Mass spectral analyses were performed using a 9.4 T Fourier transform ion cyclotron resonance mass spectrometer (IonSpec QFT, Lake Forest, CA) equipped with a Picoview nano-ESI stage (New Objective, Woburn, MA). Samples were introduced via a 1.0 μL full loop injection using a NanoLC-1D microfluidic pump and injector port (Eksigent Technologies, Dublin, CA). The infused solutions were charged via a liquid junction immediately before the nanoelectrospray tip and held at 1500–2100 V with respect to the sample cone interface. Ions entered the system through a 390 μm aperture and were externally accumulated in a storage RF-only hexapole for up to 5 s before injection into the ICR cell.

Individual glycopeptide ions were selected within the ICR cell using stored-waveform inverse Fourier transform (SWIFT) isolation prior to infrared multiphoton dissociation (IRMPD) or SORI-CID. For IRMPD experiments the IR radiation was supplied using a 10.6 μm 20 W CO_2 laser (Parallax Laser Inc., Waltham, MA). The IR radiation was expanded to 0.5 cm with an inline beam expander (Synrad Laser, Mukilteo, WA) to ensure that the isolated ion cloud received maximum exposure during each IR laser pulse. The fragmentation was optimized by varying the IRMPD laser pulse between 500 and 1500 ms. Irradiation time was increased until the majority of the precursor ion was dissociated. For SORI-CID experiments, the isolated precursor ions were subjected to SORI amplitudes of 12.0–14.5 $V_{(b\text{-}p)}$ for 1000 ms with a +1000–2000 Hz frequency offset while the ICR cell pressure was raised to 10^{-6} Torr via pulsed nitrogen gas. Up to 50 mass spectral scans were acquired and averaged to enhance signal-to-noise ratio. Glycopeptide spectra were externally calibrated using the IonSpec Omega data station. The m/z error threshold used for all assignments was 10 ppm.

Molecular Modeling. Molecular modeling of glycopeptide ions was performed with Insight II (Accelrys, San Diego, CA) employing consistent valence force field (CVFF) energy expression coupled with simulated annealing and energy minimization. Simulated annealing permitted a full exploration of the configuration space of a given glycopeptide ion. First, dynamics calculations were performed for each ion at a high temperature to explore all possible configurations. Each ion was then iteratively simulated at incrementally lower temperatures. This technique allowed for full consideration of all ion coordination possibilities. Minimization via steepest descents and conjugate gradients then stepped down the local energy of the ion to obtain the lowest energy structures. Simulations for sodium coordination were performed using simulated annealing starting at 600 K with 100 K steps down to 300 K and corresponding minimization. For protonation, the glycopeptide was first annealed using the same procedure for sodium coordination, then the proton was individually placed at each electron rich site and minimization was performed. In each set of simulations at least 500 lowest energy configurations were explored.

RESULTS AND DISCUSSION

Multiple glycopeptides were produced via the digestion of ribonuclease B (RNase B) and chicken egg albumin (CEA) with immobilized Pronase. RNase B is a well-characterized protein that is known to contain a single glycosylation site at ^{60}N. The site is occupied by high-mannose type oligosaccharides containing between five and nine mannose residues.[41] CEA is also a well-characterized glycoprotein that has two potential N-linked glycosylation sites; however, only ^{292}N has been demonstrated to harbor glycans. This site is highly glycosylated, occupied by both high mannose and hybrid type glycans.[42] Due to the nonspecific nature of Pronase digestion, glycopeptides containing between two and five amino acid residues were formed by modulating the digestion

(41) Fu, D. T.; Chen, L.; Oneill, R. A. *Carbohydr. Res.* **1994**, *261*, 173–186.
(42) Harvey, D. J.; Wing, D. R.; Kuster, B.; Wilson, I. B. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 564–571.

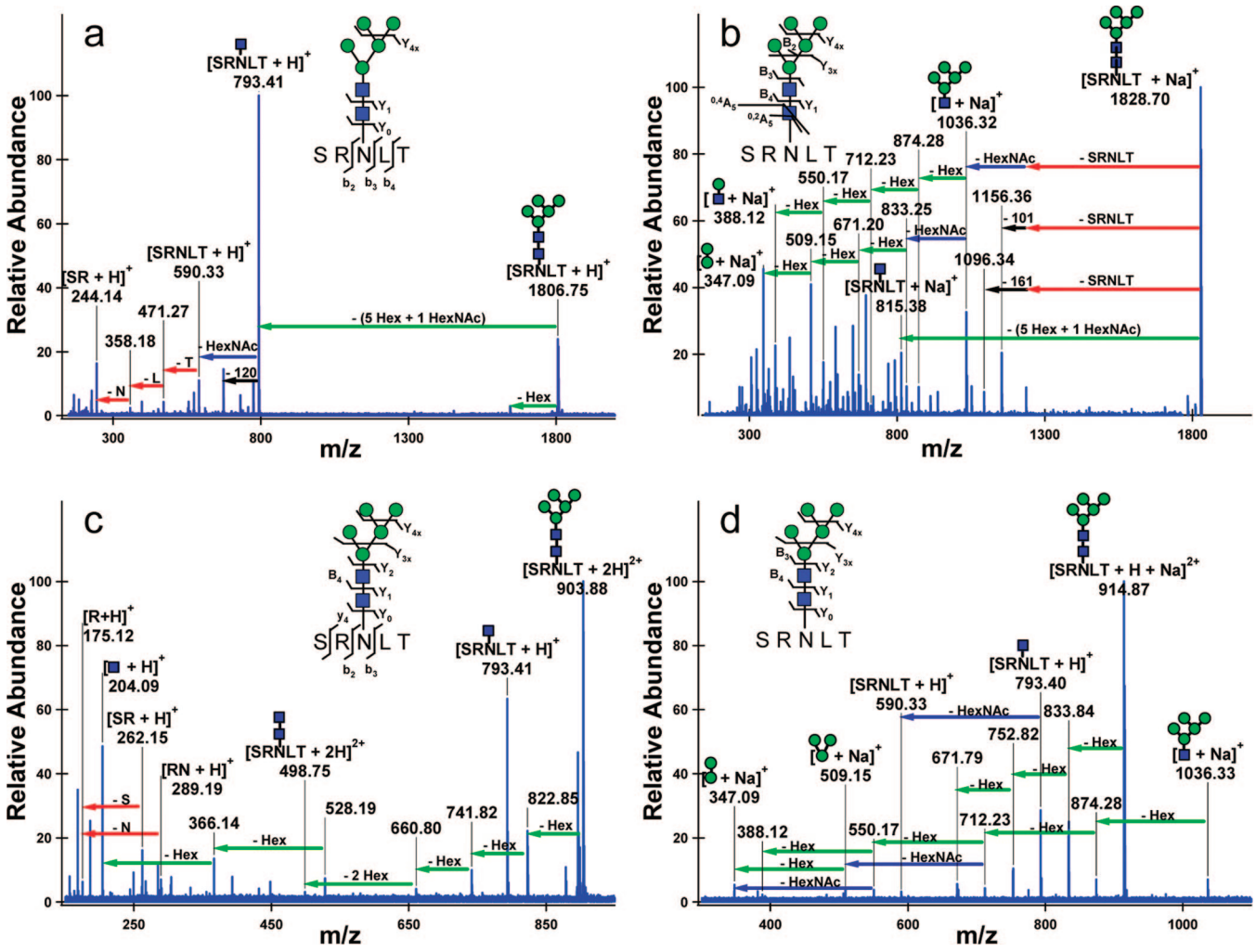

Figure 1. IRMPD nano-ESI FT-ICR mass spectra of the glycopeptide ^{58}SRNLT^{62} + Man5 derived from RNase B with multiple coordinating ions: (a) singly charged proton coordinated glycopeptide *m/z* 1806.75 (5 scans); (b) singly charged sodium coordinated glycopeptide *m/z* 1828.70 (10 scans); (c) doubly charged doubly proton coordinated glycopeptide *m/z* 903.88 (5 scans); (d) doubly charged proton and sodium coordinated glycopeptide *m/z* 914.87 (15 scans).

time. The nomenclature for glycopeptide fragmentation used follows the conventions of Roepstorff and Fohlman for peptide cleavages, and the conventions of Domon and Costello for glycan cleavages.[43,44] To more easily distinguish between peptide and glycan fragments produced from a given precursor, all peptide fragments were marked in lowercase and glycan fragments were marked in uppercase letters. The monosaccharides are represented according to the nomenclature utilized by the Consortium for Functional Glycomics (http://glycomics.scripps.edu/CFGnomenclature.pdf). According to this scheme mannose (Man) and *N*-acetylglucosamine (GlcNAc) are indicated by a green circle and blue square, respectively. In the text of this manuscript, the well-characterized N-linked high mannose glycans will be referred to using the number of mannose residues in the glycan. Thus, the abbreviation Man5 represents the glycan composition $GlcNAc_2Man_5$.

Fragmentation of Protonated vs Sodiated Glycopeptides. Upon IRMPD, it was observed that both charge carrier and the charge state greatly affected the fragmentation behavior of glycopeptide ions. For example, Figure 1 shows the IRMPD spectra of four different ionic forms of the glycopeptide SRNLT + Man5. The singly protonated form of the glycopeptide was observed to have a major fragment ion corresponding to the Y_1 cleavage of the glycan, resulting in the intact peptide attached to a single GlcNAc residue at *m/z* 793.41 (Figure 1a). After this major loss the GlcNAc residue was cleaved followed by b-type peptide cleavages, enabling sequencing of the peptide portion of the glycopeptide. The proton was sequestered by the highly basic arginine, and as a result every fragment ion contained arginine. Very little information regarding the nature of the glycan was obtained. The loss of a single mannose residue via Y_{4x} cleavage was observed from the precursor ion, where the subscript x indicates that the cleavage site was ambiguous due to branching.

In contrast to the simple fragmentation spectrum of the protonated glycopeptide, the sodium coordinated glycopeptide spectrum contained far more fragment peaks (Figure 1b). One of the most intense fragment ions was that resulting from a B_4 cleavage of the glycan (*m/z* 1036.32), yielding the compliment to the most intense peak in the protonated spectrum. After losing the peptide, the glycan was then eroded via glycosidic bond

(43) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
(44) Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397–409.

cleavages, which allowed the connectivity of the glycan to be determined. It was also noted that the asparagine linked GlcNAc yielded multiple types of cross-ring cleavage while all of the other monosaccharide residues yielded only glycosidic bond cleavages. The majority of the unlabeled peaks in Figure 1b correspond to two additional series of glycosidic bond cleavages spaced by a water loss.

The IRMPD spectrum of the doubly protonated glycopeptide yielded the same major fragment ion as the singly protonated spectrum, i.e. *m/z* 793.41 (Figure 1c). However, the doubly protonated glycopeptide yielded additional glycan fragments. Interestingly, the majority of the glycan fragmentation reactions produced ions that retained the doubly charged state. These products are due to Y-type fragments such as *m/z* 822.85, 741.82, 660.80, and 498.75. This behavior was consistent with previous observations reported in the literature.[8,20] The presence of this additional proton appears to aid in the fragmentation of the glycan, yet it was apparently retained by the peptide-containing fragment upon dissociation. The first proton was sequestered by the highly basic arginine, and as a result the majority of the fragment ions contain arginine. The second proton was not fixed on a specific basic site and could participate in additional dissociation reactions either in the peptide or glycan portion of the ion. This behavior was expected according to the mobile proton model of peptide fragmentation.[23,45]

The doubly charged glycopeptide coordinated with a sodium ion and a proton yields a mixture of all of the fragments previously discussed (Figure 1d). Two of the major fragment ions were the result of B_4 and Y_1 cleavage of the glycan, with the sodium coordinated to the glycan portion (*m/z* 1036.33) and the proton affixed to the peptide (*m/z* 793.40). Each of these fragments then dissociated in a similar manner to their singly charged counterparts; however, fewer fragments were observed as compared to the singly coordinated cases due to the lower abundance of the doubly charged precursor ion.

The difference between singly and doubly protonated glycopeptides has been seldom noticed. This is possibly due to the large masses of tryptic glycopeptides, which are most often multiply charged under ESI conditions. For this reason, the majority of fragmentation studies involving glycopeptides have focused on multiply protonated ions. In contrast, the smaller glycopeptides resulting from Pronase digestion are primarily observed as singly charged ions. It should be noted that the observed fragmentation behavior of singly protonated glycopeptides is not unique to Pronase digestion products. As recently noted by Yu and co-workers, singly protonated tryptic glycopeptides ionized by MALDI yielded CID fragments consistent with protonation of the peptide moiety.[32]

As the singly protonated glycopeptide spectrum was unique in its fragmentation behavior, other glycopeptides were analyzed to determine the generality of this type of fragmentation (Figure 2). All analyses resulted in similar types of dissociation with the Y_1 cleavage dominating the MS/MS spectra, followed by the loss of the remaining GlcNAc and peptide cleavages. For example, a

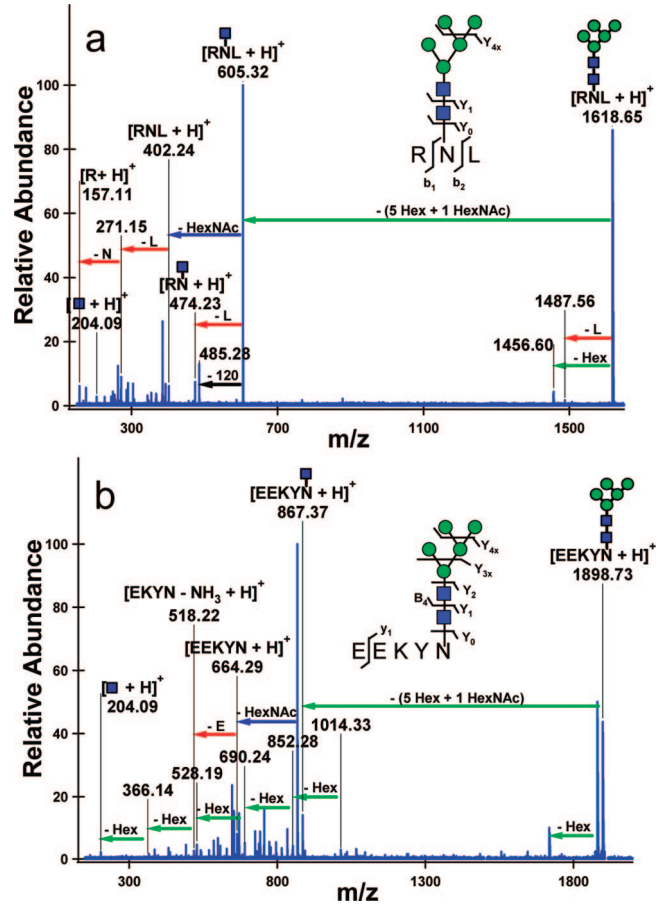

Figure 2. IRMPD nano-ESI FT-ICR mass spectra of proton coordinated glycopeptides yield similar fragmentation behavior: (a) RNase B glycopeptide ^{59}RNL^{61} + Man5 *m/z* 1618.65 (10 scans); (b) CEA glycopeptide ^{288}EEKYN^{292} + Man5 *m/z* 1898.73 (10 scans).

shorter glycopeptide derived from RNase B, RNL + Man5 (*m/z* 1618.65), was analyzed by IRMPD with the fragmentation behavior mirroring that of the previously analyzed glycopeptide (Figure 2a). The most abundant fragment was once again the Y_1 cleavage of the glycan *m/z* 605.32, resulting in the loss of $GlcNAcMan_5$ followed by the loss of the remaining GlcNAc and then subsequent peptide cleavages. The entire peptide was also sequentially fragmented from that point, allowing the sequence of the peptide to be determined unequivocally. A glycopeptide from a CEA digest, EEKYN + Man5 (*m/z* 1898.73), yielded a similar fragmentation pattern (Figure 2b). The Y_1 cleavage was the major fragment ion; however, the neutral loss of water from the side chain of the glutamic acid shifted the mass of the major peak from *m/z* 867.37 to *m/z* 849.36. The loss of a single glutamic acid residue was observed via y_1 cleavage after the loss of the remaining GlcNAc residue. In this case some fragmentation of the glycan was also observed; however, the intensity of these glycan fragments was an order of magnitude less than the peptide-containing fragments, presumably due to sequestration of the proton by the lysine residue. The observation of more glycan fragments was attributed to the significantly lower gas phase basicity of lysine when compared to arginine, allowing for protonation of the B_4 fragment (*m/z* 1014.33) of the glycopeptide and its respective fragments.[46]

To further understand the dissociation behavior of protonated N-linked glycopeptide ions, SORI-CID was performed on the

(45) Cox, K. A.; Gaskell, S. J.; Morris, M.; Whiting, A. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 522–531.

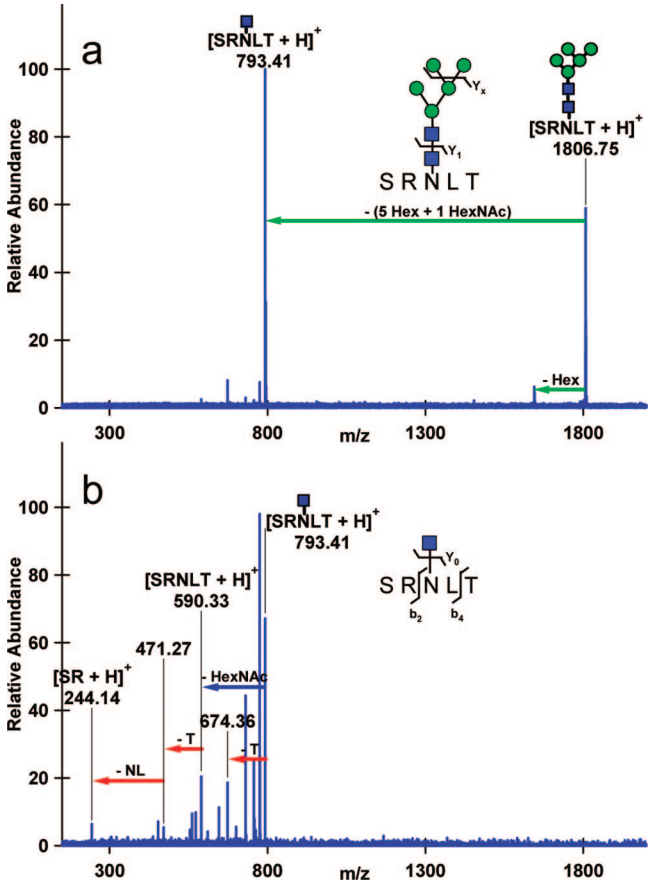

Figure 3. SORI-CID nano-ESI FT-ICR mass spectra: (a) MS/MS of protonated RNase B glycopeptide ^{58}SRNLT^{62} + Man5 *m/z* 1806.75 (5 scans); (b) MS/MS/MS of the major product from (a), *m/z* 793.41 (50 scans).

glycopeptide SRNLT + Man5 (Figure 3). The fragment ions produced after one stage of MS/MS were dominated by the Y_1 cleavage of the glycan resulting in the peptide bound to a single GlcNAc, *m/z* 793.41 (Figure 3a). This fragment was then isolated and subjected to another round of SORI-CID (MS^3) resulting in the loss of the GlcNAc and multiple peptide fragments (Figure 3b). The Y_1 ion arising from the cleavage of the chitobiose core has previously been shown to predominate in N-linked glycopeptide MS/MS in IT-MS with CID, while MS^3 experiments result in fragmentation of peptide bonds.[8] These results have been demonstrated for multiple types of glycans including complex-type sialylated structures, and correlate with the SORI-CID data in Figure 3. Since IRMPD is capable of producing fragment ions in a single event which would require multiple stages of SORI-CID, the appearance of peptide backbone fragments in the IRMPD of protonated N-linked glycopeptide ions was not unexpected.[21] Since peptide backbone fragmentation is evidently a sequential process, that is, the peptide fragments are themselves derived from the Y_1 fragment, the absence of peptide fragments during previous efforts may have been due to insufficient IR irradiation time. Interestingly, previously reported results obtained for high mannose type glycopeptides demonstrate this point, as those irradiated for shorter periods of time only yielded glycan fragments, while those that were irradiated for twice as long yielded primarily peptide fragments or a mixture of both fragmentation types.[38]

Glycopeptides without Basic Residues. Since the dramatic difference in fragmentation of singly and doubly protonated glycopeptide ions appeared to be a result of charge sequestration and gas phase basicity, additional glycopeptides were investigated to determine whether prominent peptide bond cleavages were conserved in the absence of basic residues. In order to obtain glycopeptide ions without basic residues, digestion of RNase B and CEA was performed for 24 h to reduce the size of the peptide portion. Two distinct glycopeptide ions with the same peptide moiety (NL) but different glycans were compared as the protonated and sodium coordinated ions (Figure 4). The protonated RNase B glycopeptide NL + Man5 (*m/z* 1462.56) yielded a mixture of peptide and glycan fragments upon IRMPD (Figure 4a). In the absence of basic amino acid residues, the cleavage of the GlcNAc−GlcNAc bond was observed as both Y_1 (*m/z* 449.22) and B_4 (*m/z* 1014.36) fragments of comparable but low abundances. These results were in sharp contrast to singly protonated glycopeptides with basic residues, in which the GlcNAc−GlcNAc glycosidic cleavage was the major product ion and was observed exclusively as the Y_1 fragment (Figures 1a and 2a,b). Evidently, the lack of a basic amino acid residue permits greater proton mobility and thus allows more extensive glycan fragmentation and formation of product ions that do not include the peptide portion. Each glycosidic bond was cleaved, and the most abundant fragment ions were protonated mono- and disaccharide ions at *m/z* 204.09 and 366.14. The protonated glycopeptide NL + $HexNAc_4Hex_5$ (*m/z* 1868.74) derived from CEA (Figure 4c) exhibited fragmentation characteristics similar to the previously discussed example (Figure 4a). The loss of a terminating HexNAc residue from the intact glycopeptide ion was among the most abundant fragments. Again, the GlcNAc−GlcNAc bond cleavage was observed as both Y_1 and B_4 fragments, and these ions were low in abundance compared to the Y_1 fragment characteristic of singly protonated glycopeptides with basic residues. The loss of leucine was also observed, indicating that complex type glycopeptides also can yield peptide fragments.

The sodium adducts of these two glycopeptides were also subjected to IRMPD (Figure 4b and 4d). The dissociation patterns of these ions were very similar to that observed for the sodiated form of SRNLT + Man5 (Figure 1b). This result was not unexpected, because the presence or absence of basic amino acid residues should exert less influence on the fragmentation of sodiated ions as compared to protonated ions. This result also suggests that, for some glycopeptides, fragmentation of sodiated ions may be preferred over protonated ions since the dissociation behavior of sodiated glycopeptides is less sensitive to amino acid composition. The major product ions in these spectra corresponded to the B_4 cleavage of the GlcNAc−GlcNAc bond (corresponding to *m/z* 1036.33 in Figure 4b and *m/z* 1442.50 in Figure 4d), and $^{0,2}A_5$ cross-ring cleavage of the peptide linked GlcNAc (corresponding to *m/z* 1156.38 in Figure 4b and *m/z* 1562.55 in Figure 4d). The composition and connectivity of the glycan portions of the glycopeptides were confirmed based upon the multiple series of fragment ions that were found. The structure of the glycan portion of the CEA derived glycopeptide in Figure

(46) Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201–217.

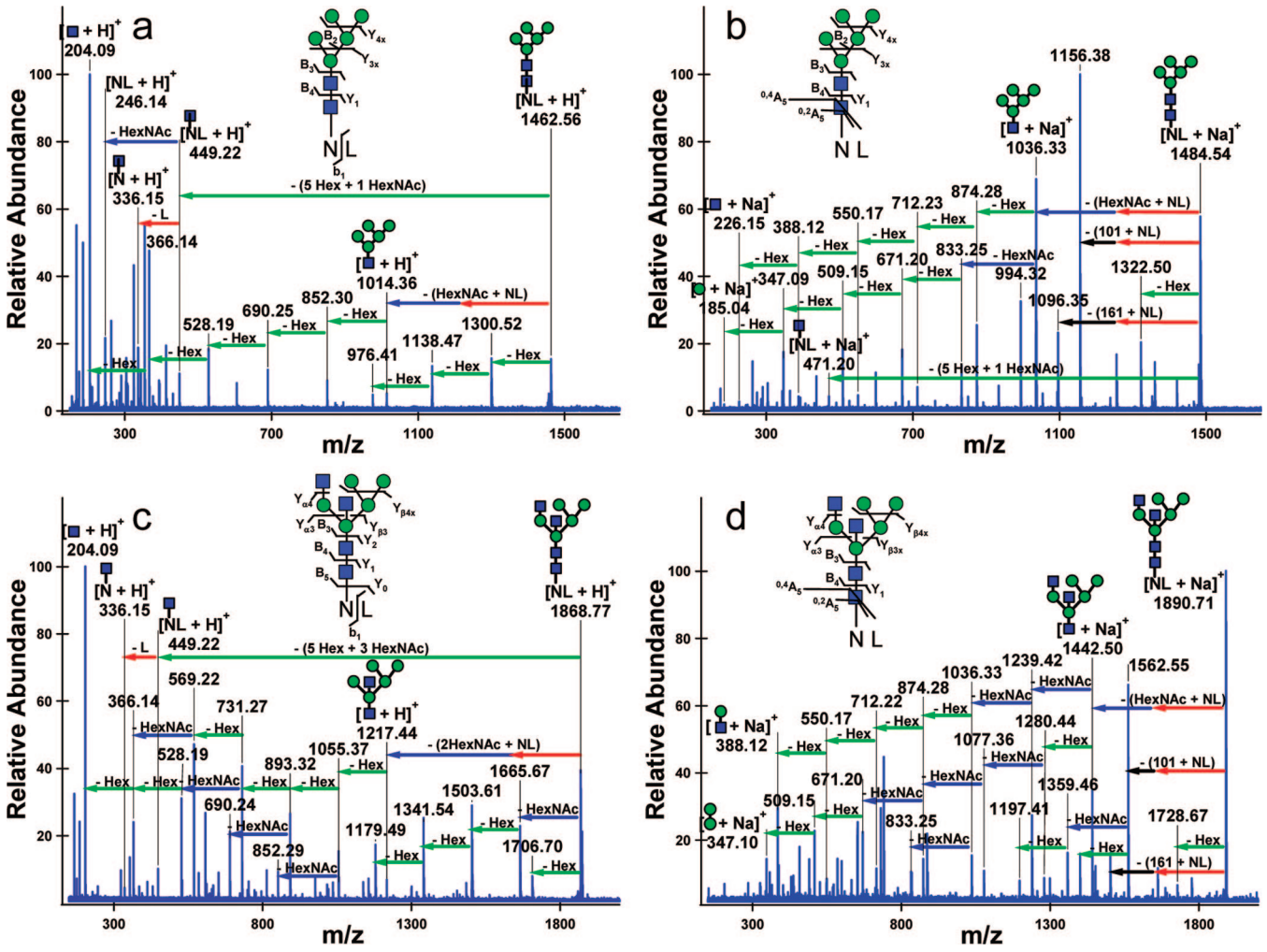

Figure 4. IRMPD nano-ESI FT-ICR mass spectra of glycopeptides without basic residues. The RNase B glycopeptide ^{60}NL^{61} + Man5: (a) proton coordinated *m*/*z* 1462.56 (10 scans); (b) sodium coordinated *m*/*z* 1484.54 (10 scans). The CEA glycopeptide ^{292}NL^{293} + $HexNAc_4Hex_5$ (hybrid type glycan); (c) proton coordinated *m*/*z* 1868.74 (25 scans); (d) sodium coordinated *m*/*z* 1890.71 (10 scans).

4d was consistent with work performed by Harvey and coworkers.[42] On the basis of Figure 4a−d, the dissociation behavior of high mannose and complex type glycopeptide ions appeared to have similar dependencies on charge carrier, possibly extending to charge state and peptide composition.

Molecular Modeling. In order to more fully understand the structural basis for the fragmentation behavior of N-linked glycopeptide ions, multiple gas phase simulations were performed to investigate the nature of the cation coordination to glycopeptides. Since IRMPD and CID are slow fragmentation processes, the lowest energy fragmentation pathways predominantly lead to the majority of the product ions.[47] The intact glycopeptides were initially modeled in order to investigate sites of sodium coordination and protonation and to investigate whether these specific sites could direct fragmentation. It was observed experimentally that the GlcNAc−GlcNAc bond was one of the most labile bonds in both the protonated and sodiated forms of the molecules; however, in the protonated form, the peptide-containing fragment was observed while the glycan portion of the glycopeptide was detected for sodiated precursor. Molecular modeling of this species indicated that the proton was localized at the most basic amino acid residue. Data in Table 1 illustrate that the difference in energy for protonation of the arginine in SRNLT + Man5 is favored energetically over the next lowest energy site by 62.02 kcal/mol. The same was true for the lysine-containing EEKYN + Man5, although the energy difference was not quite as large due to the lower gas phase basicity of lysine as compared to arginine.[46] On the other hand, protonation of glycopeptides without strongly basic residues, (i.e., NL + Man5) predicted multiple sites of protonation due to small energy differences between the lowest energy structure (in which the proton was coordinated to the carboxyl group on the asparagine bound GlcNAc) and other relatively low energy protonation sites. Hence, computationally, the proton was predicted to be more mobile in glycopeptides without strongly basic residues and as a result glycan fragments which were not possible in the presence of strongly basic residues became accessible (Figure 1). The *p* values recorded in Table 1 illustrate the statistical significance of the coordination site specificity. The lower *p* values correspond to more significant coordinations, with *p* values greater than 0.05 having no significance.

The molecular simulations of sodium adduction to the glycopeptides indicated a multidentate coordination, with the sodium ion coordinated mainly to the hydroxyl oxygens on the glycan.

(47) Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* **2003**, *22*, 158–181.

Table 1. Total Energies from Molecular Modeling for Each Glycopeptide and Their Respective Coordinating Ions[a]

glycopeptide	cation	lowest energy site	energy of lowest site (kcal/mol)	Δ from next lowest site (kcal/mol)	p value of lowest site
SRNLT + Man5	H^+	R: C=N	−233.39 ± 0.18	−62.02	1.39E−05
EEKYN + Man5	H^+	K: C−N	−213.48 ± 0.02	−40.15	5.36E−04
NL + Man5	H^+	GlcNAc 1: =O	−162.78 ± 0.001	−0.53	7.73E−02
SRNLT + Man5	Na^+	glycan/peptide	−91.53 ± 0.17	−0.09	2.30E−02
EEKYN + Man5	Na^+	glycan/peptide	−88.94 ± 0.14	−1.68	6.39E−03

[a] The lowest energy structures and sites of coordination are indicated for each data set. Sodium coordination was multidentate in each case, and hence the site of coordination was a pocket between the glycan and peptide portion of the molecule. The difference in energy from the next lowest site indicates the specificity of the coordination. Additionally, the statistical p values relate the lowest energy site to the entire data set (smaller values indicate more specificity).

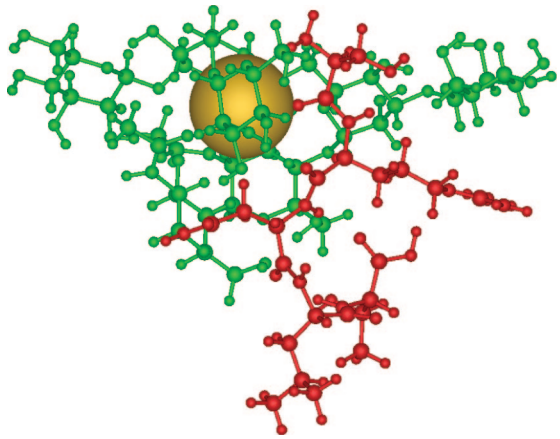

Figure 5. Simulated coordination of sodium cation to the RNase B glycopeptide ^{58}SRNLT^{62} + Man5. The sodium (yellow) is represented as space filling, while the glycan (green) and peptide (red) portions are represented as ball and stick. The sodium is coordinated in a multidentate fashion with the majority of the coordination stemming from the hydroxyl oxygens and little interaction with the peptide.

Figure 5 shows the sodium ion enveloped by the glycopeptide SRNLT + Man5; the sodium (yellow) is represented as space filling, and the glycan portion (green) and peptide portion (red) are both represented as ball and stick. Due to the complex coordination and the flexibility of the molecule, multiple sodium coordination possibilities exist with similar internal energies; complimentary results have been obtained for glycans.[48,49] Even though there was not a single site of coordination, it was determined that the glycan portion coordinates more completely with the sodium in the majority of low energy structures.

Molecular modeling of fragment ions was performed to further investigate the effects of cation coordination on the observed fragment ions. One of the most labile bonds of N-linked glycopeptides, the GlcNAc−GlcNAc bond, was broken, and each fragment was set at a distance such that any interaction of the fragments was prohibited. In separate modeling experiments one fragment was coordinated to a proton or sodium ion while the other was left uncoordinated, and then the complementary experiments were performed. These experiments illustrated the energy preference for the observation of the sodiated glycan moiety by 10 kcal/mol over the peptide-containing portion (Table 2). However, protonation was shown to greatly favor the peptide-containing moiety due to the presence of strongly basic residues.

Table 2. Total Energy for Molecular Modeling of the Major Fragment Ions Resulting from GlcNAc−GlcNAc Bond Cleavage[a]

glycopeptide	cation	glycan coordinated − peptide coordinated energy (kcal/mol)	p value of lowest site
SRNLT + Man5	H^+	49.58 ± 3.65	4.89E−04
EEKYN + Man5	H^+	26.53 ± 1.51	2.38E−04
SRNLT + Man5	Na^+	−10.61 ± 5.69	1.92E−02
EEKYN + Man5	Na^+	−8.05 ± 5.95	1.06E−02

[a] Sodium coordination to the glycan portion of the molecule results in overall lower energy of the system, while protonation of the peptide-containing moiety results in lower energy as demonstrated by the energy difference.

These results further illustrate that singly protonated glycopeptide ions subjected to fragmentation will preferentially produce ions that contain the peptide portion, while singly sodiated glycopeptides will produce glycan-containing fragments.

CONCLUSIONS

Fragmentation of N-linked glycopeptides via IRMPD has been shown to be dependent upon the charge carrier, charge state, and peptide composition of the glycopeptide ion. Singly protonated glycopeptide ions that contain basic amino acid residues yielded mainly peptide-containing fragments after cleavage of the GlcNAc−GlcNAc bond. In a complementary manner, multiply protonated glycopeptide ions with basic amino acid residues yielded mainly multiply protonated products of glycosidic bond cleavage, with some relatively low abundance, singly charged products of peptide backbone cleavage. The fragmentation of protonated glycopeptide ions without basic amino acid residues yielded mainly glycosidic cleavage fragments, a significant contrast to the predominant peptide backbone cleavage observed in the presence of basic residues. Sodium coordinated glycopeptides were observed to yield multiple series of glycosidic bond cleavages, regardless of peptide or glycan composition. The major fragments of sodium coordinated glycopeptides were the result of either GlcNAc−GlcNAc bond cleavage or cross-ring cleavage of the peptide linking GlcNAc, with the glycan portion retaining the charge in both cases. In general, the GlcNAc−GlcNAc bond was observed to be one of the most labile bonds in singly charged N-linked glycopeptides, with the glycan fragment retaining the charge from sodiated precursor ions and the peptide portion retaining the charge from protonated precursor ions. The overall fragmentation behavior of N-linked glycopeptide ions was found to be dependent upon the chemistry of each particular ion, and

(48) Fukui, K.; Kameyama, A.; Mukai, Y.; Takahashi, K.; Ikeda, N.; Akiyama, Y.; Narimatsu, H. *Carbohydr. Res.* **2006**, *341*, 624–633.
(49) Fukui, K.; Takada, Y.; Sumiyoshi, T.; Imai, T.; Takahashi, K. *J. Phys. Chem. B* **2006**, *110*, 16111–16116.

as a result ions with different peptide composition, charge carrier, and charge state have exhibited distinct fragmentation behaviors. The differential fragmentation behavior of N-linked glycopeptide ions can be used to aid in elucidation of site specific glycosylation by allowing both peptide and glycan dissociation products to be obtained from the same glycopeptide through alteration of the charge state or the coordinating ion.

While IRMPD and ECD have previously been shown to provide complimentary dissociation information for glycopeptide ions, the present work does not support the previous notion that IRMPD exclusively provides glycan fragmentation of glycopeptide ions. On the contrary, we have shown that IRMPD can provide both types of cleavage depending on the intrinsic chemistry of the precursor ion. These considerations are particularly relevant to the analysis of glycopeptides produced by nonspecific proteolysis, since the location and presence of basic amino acid residues is not fixed, as is the case for tryptic glycopeptides. This work improves our understanding of N-linked glycopeptide fragmentation and promises to improve an important analytical tool for glycopeptide analysis and their applications in the field of glycoproteomics.

ACKNOWLEDGMENT

The authors would like to acknowledge the following funding sources: California Dairy Research Foundation (06 LEC-01-NH), Dairy Management Incorporated, University of California Discovery Grant (05GEB01NHB), and the National Institutes of Health (GM 49077).

Received for review January 10, 2008. Accepted February 25, 2008.

AC800067Y