Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data

William R. Alley, Jr.1,3, Yehia Mechref1,2,3* and Milos V. Novotny1,2,3*

1National Center for Glycomics and Glycoproteomics, Indiana University, Bloomington, IN 47405, USA
2METACyt Biochemical Analysis Center, Indiana University, Bloomington, IN 47405, USA
3Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

Received 12 February 2008; Revised 23 October 2008; Accepted 24 October 2008

Structural characterization of a glycopeptide is not easily attained through collision-induced dissociation (CID), due to the extensive fragmentation of glycan moieties and minimal fragmentation of peptide backbones. In this study, we have exploited the potential of electron-transfer dissociation (ETD) as a complementary approach for peptide fragmentation. Model glycoproteins, including ribonuclease B, fetuin, horseradish peroxidase, and haptoglobin, were used here. In ETD, radical anions transfer an electron to the peptide backbone and induce cleavage of the N–Cα bond. The glycan moiety is retained on the peptide backbone, being largely unaffected by the ETD process. Accordingly, ETD allows not only the identification of the amino acid sequence of a glycopeptide, but also the unambiguous assignment of its glycosylation site. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.

Analysis of post-translational modifications (PTMs) is currently a key research area in proteomic analysis. More than 100 PTMs are known, increasing significantly the overall complexity of proteomic samples. The most common of these, and arguably the most important, is glycosylation. It has been estimated that at least 50% of the proteins in mammalian proteomes are glycosylated; multiple sites of glycosylation are present in a complex sample, resulting in multiple peptide ions per protein. Enzymes can be used to cleave the glycan moiety from the peptide backbone, being largely unaffected by the ETD process. When data acquired from both fragmentation techniques are combined, it is possible to characterize comprehensively the entire glycopeptide. This is being achieved with a mass spectrometer capable of alternating between CID and ETD on-the-fly during an LC/MS/MS analysis. This is demonstrated here with several tryptic glycopeptides.
(CID) tandem mass spectrometry (MS/MS) are often not easily interpretable. During the CID process, glycopeptides undergo preferential glycan fragmentation,15–17 as it represents a lower-energy fragmentation pathway. Only minimal fragmentation of the peptide backbone is observed. While CID is effective in the characterization of the glycan moieties, it does not provide diagnostic information about the peptide backbone. The minimal fragmentation of the peptide backbone has previously provided the justification for the removal of the glycan from the peptide in some experiments.

Recently, electron-based methods such as electron-capture dissociation (ECD),28–31 where a radical anion transfers an electron to a peptide cation, have been introduced as alternative methods of peptide fragmentation. While the exact mechanisms for ECD and ETD fragmentation are not fully understood at present,24–27 both reactions are known to produce c’ ions and z’ radical ions resulting from the scission of the N–Ca bond. Both ECD and ETD induce extensive fragmentation of the peptide backbone with minimal fragmentation of PTMs. This is due to the radical anion induced fragmentation and the absence of vibrational excitation. Moreover, this type of fragmentation results in retaining the PTM on the peptide backbone, thus easily allowing for the site of modification to be identified.

ECD has been successfully applied to investigate glycopeptides.28–31 In these studies, the site of glycosylation could be determined. This has been possible because the glycan remains largely intact and attached to the peptide backbone. However, the potential of ETD for the analysis of glycopeptides has been exploited only recently.17,22,32,38 McLuckey and co-workers have investigated ETD for the tryptic glycopeptide originating from E. cristagalli.22 Electron-transfer dissociation of this glycopeptide produced 12 z’-type ions and no c’ ions when sulfur dioxide radical anions were used as a reactant. When nitrobenzene was used as the radical anion, eleven z’ ions were observed along with three c’ ions. Even more recently, the analysis of three glycopeptides originating from tryptically digested horseradish peroxidase has been reported using a commercially available ion trap mass spectrometer equipped with an ETD source.17,32

In this investigation, we explore the use of ETD in conjunction with CID in the characterization of tryptic glycopeptides originating from model glycoproteins, such as ribonuclease B, horseradish peroxidase, human haptoglobin, and fetuin. In the general procedure, glycopeptides were subjected to both reversed-phase and activated graphitized carbon nano-liquid chromatography through a commercial microchip attachment, further attached to an ion trap mass spectrometer capable of automatically alternating between CID and ETD modes. The studied glycopeptides thus demonstrate the ability of ETD to efficiently generate c’ and z’ fragments originating from the peptide backbone and provide valuable structural information concerning the glycosylation site and its microheterogeneity. Glycopeptides differing significantly not only in the amino acid sequence, but also in the associated glycan structures were effectively fragmented through ETD. When CID and ETD were used in tandem, both the glycan structure and the amino acid sequence of the glycopeptide under investigation could be easily deduced.

EXPERIMENTAL

Materials

Bovine ribonuclease B, bovine fetuin, horseradish peroxidase, human haptoglobin, TPCK-treated trypsin, formic acid (95–97%), guanidine hydrochloride, and ammonium bicarbonate were purchased from Sigma Chemical Co. (St. Louis, MO, USA). Dithiothreitol (DTT) and iodoacetamide (IAA) were acquired from Bio-Rad Laboratories (Hercules, CA, USA). High-purity water (18 MΩ) and acetonitrile were products of EMD Laboratories (Gibbstown, NJ, USA).

Tryptic digestion

Model glycoproteins (0.2–0.5 mg) were dissolved in 50 mM ammonium bicarbonate buffer containing 6 M guanidine hydrochloride and 10 mM DTT. Glycoprotein denaturation and disulfide bond reduction were performed for 1 h at 50 °C. The reaction mixture was then allowed to cool to room temperature prior to the addition of a 1-μL aliquot of 275 mM IAA, resulting in an IAA final concentration of 25 mM. The alkylation reaction was allowed to proceed for 1 h at room temperature in the dark. Next, samples were diluted to 300 μL with 50 mM ammonium bicarbonate buffer and trypsin was added at an enzyme-to-protein ratio of 1:30. Digestion was allowed to proceed for 18 h at 37 °C. Digestion was quenched by lowering the pH of the peptide solutions to ~3 through the addition of formic acid.

LC/MS/MS

Chip-based nano-liquid chromatographic separations were performed with an Agilent 1200 Series nano-pump system consisting of binary capillary-flow and nano-flow pumps, a vacuum degasser, and an autosampler (Agilent Technologies, Inc., Palo Alto, CA, USA). Chromatographic separations of 500 fmol to 1 pmol of tryptic peptides were performed on either a chip with activated graphitized carbon as the stationary phase (75 μm diameter × 43 mm length with a 40 nL preconcentration trap) or a C18 chip (75 μm diameter × 150 mm length, 40 nL preconcentration trap) packed with 5 μm particle size StableBond Zorbax silica. Peptide samples of 1 μL were injected using the capillary-flow pump operating at 2 μL/min. The total volume for the injection/wash cycle was 4 μL (1 μL of injected peptides, followed by 3 μL of mobile-phase A). The chip’s integrated valve was switched from the ‘enrichment’ position to the ‘analytical’ position to elute peptides from the trapping region into the separation region. Separation was performed using a linear gradient from 3–45% mobile phase B over 39 min with a flow rate of 300 nL/min. Mobile phase A consisted of 97%/3%/0.1% water/acetonitrile/formic acid, while mobile phase B was composed of 97%/3%/0.1% acetonitrile/water/formic acid. The mobile phase compositions were the same for both the capillary- and nano-flow pumps.

The LC system was interfaced to a Bruker Daltonics (Bruker Daltonik GmbH, Bremen, Germany) ultra-high capacity ion trap mass spectrometer via a ChipCube interface (Agilent Technologies, Inc.). The spraying voltage was set to 1825 V (activated graphitized carbon chip) and 1975 V (C18 chip) to sustain a continuous spray. MS scans were performed in the ‘standard enhanced’ mode and MS/MS
scans were performed in ‘ultra’ mode covering the 200–3000 m/z range for the analyses of all glycopeptides. For ETD analyses, a negative chemical ionization source generated fluoranthene radical anions, which were transported to the ion trap and used as the electron-transfer agents. To produce the radical anions, the reactant temperature was set at 65 °C, while the ionization energy and emission current were set at 65 eV and 5.0 μA, respectively. Radical anions were ideally accumulated in the trap for 7.5 ms and allowed to react with the peptide ions for 175 ms. The mass spectrometer was set to alternate between CID and ETD mode for each ion, thus maximizing the MS/MS data collected.

RESULTS AND DISCUSSION

Several glycopeptides from model glycoproteins have been analyzed by both CID and ETD. The CID analyses resulted in extensive fragmentation of the glycan with limited peptide backbone cleavages. Peptide backbone fragmentation was mainly observed when ETD was employed. Minimal glycan fragmentation was observed during the ETD experiment, but some glycan losses were observed as previously reported.

ETD optimization

For efficient ETD, the accumulation and the reaction times were first optimized by injecting 500 fmol of horseradish peroxidase tryptic digest onto the activated graphitized carbon LC chip, while intensities of the product ions were monitored. The glycopeptide NVGLNR with a glycan structure of GlcNAc$_2$(Fuc)Man$_3$(Xyl) was used to evaluate the influence of both radical anion accumulation and reaction times. The ETD-produced fragment ions ($z'$, $z''$, $z'''$, $c'_3$, and $c''_3$) were normalized by dividing the intensity of each fragment by the intensity of the precursor ion observed in the preceding MS scan. This normalization minimized any potential variations caused by performing ETD at different elution times across a chromatographic peak.

The accumulation times of fluoranthene radical anions in the trap were studied at 1, 5, 7.5, and 10 ms (Fig. 1(a)). Each accumulation time experiment was performed three times. Under the conditions previously described for the ETD source, a 1-ms accumulation time resulted in ca. 100 000 fluoranthene radical anions in the trap, while 5 ms and 7.5 ms allowed for the accumulation of ca. 500 000 and 750 000 radical anions, respectively. At 10 ms accumulation, 1 000 000 radical anions were accumulated in the trap. Radical-anion levels were estimated by using the ‘ion-charge control’ feature included with the instrument control software. While ETD fragments were produced at 1 ms, the relative intensities of the different fragments observed were substantially lower than those seen for 5 and 7.5 ms accumulation times. For one ion, $z''$, the relative intensity produced by the 1 ms accumulation time was the greatest for all ions. However, a large standard deviation was noticed and was most likely a statistical anomaly. As a general trend, the relative intensities observed at 7.5 ms accumulation time were optimum. Therefore, an accumulation time of 7.5 ms was used throughout continuation of this work.

The reaction time at which the ions and radical anions were allowed to interact in the ion trap was also investigated.

This was studied by injecting 500 fmol of horseradish peroxidase tryptic digest. Normalization was performed, as described above, using the same ions and at the reaction times of 75, 100, 150, 200, and 250 ms. The reaction time variation and its influence on the efficiency of fragmentation is depicted in Fig. 1(b). As a general trend, the relative intensity of the ions monitored increased as the reaction time was increased from 75 to 200 ms. A significant decrease in the relative intensities of the ions was observed at reaction times greater than 200 ms. This is expected, since at longer reaction times, reactions with fluoranthene radical anions in the trap may have caused further charge-reduction reactions, either through proton transfer/abstraction, or possibly non-dissociative electron transfer. These neutralization events of the produced fragments are expected to prompt the observed lower fragmentation efficiency for extended accumulation times. Therefore, a 175 ms reaction time was employed for the remainder of this work.

Limit of detection

Using the Man$_5$ glycopeptide derived from ribonuclease B and the GlcNAc$_2$(Fuc)Man$_3$(Xyl)-NVGLNR glycopeptide derived from horseradish peroxidase, the minimum amount of material needed for successful ETD fragmentation was determined. All of the ETD fragments for these two glycopeptides were observed at high levels (500 fmol to 1 pmol) as well as much lower levels (data not shown). Glycopeptides were successfully fragmented at 10 fmol in the case of the glycopeptide derived from horseradish peroxidase.
tryptic digest, and 20 fmol in the case of the glycopeptide derived from ribonuclease B tryptic digest. Below these levels, we were unable to detect the glycopeptides.

**CID and ETD of glycopeptides derived from ribonuclease B**

Ribonuclease B is a glycoprotein with a single site of glycosylation at N60, with high-mannose glycan structures of GlcNAc2(Man)5–9. Tryptic digestion of ribonuclease B commonly produces a single missed cleavage resulting in a glycopeptide backbone with an amino acid sequence of SRNTLK to which different glycan structures are attached. Electrospray ionization of this glycopeptide produces the +2 and +3 charge states. The CID analyses of both charge states resulted in extensive glycan fragmentation. As Fig. 2(a) depicts, in the CID fragmentation of the triply charged glycopeptide with the Man5 glycan, the ion at m/z 204 results from a loss of the GlcNAc residue attached to the peptide backbone. The ion at m/z 375.9 is a triply charged ion that results from the loss of five mannose residues, leaving the two core GlcNAc residues attached to the peptide backbone. The ions at m/z 429.2, 483.7, 537.7, and 591.7 are all separated by 54 m/z units, corresponding to a triply charged ion series originating from the loss of 4, 3, 2, or 1 mannose residues, respectively. The ion at m/z 461.4 is a doubly charged ion with only a single GlcNAc residue attached to the peptide backbone. The signals at m/z 724.9, 806.3, and 886.4 are doubly charged ions separated by 81 m/z units originating from the loss of 3, 2, or 1 mannose residues, respectively. A singly charged ion at m/z 921.6 also originates from the loss of 5 mannose residues and a GlcNAc residue.

A vastly different fragmentation pattern was observed with ETD in which significant fragmentation of the peptide backbone was observed (Fig. 2(b)). For the +3 charge state (Fig. 2(b)), c′3–5 ions were observed, along with z′3–5 ions. This spectrum allowed for the easy identification of the site of glycosylation in this peptide, as based on the mass shift of 1330 Da between c′2 and c′3 and between z′3 and z′4. This mass difference corresponds to the addition of N-linked GlcNAc2Man5 glycan, along with asparagine. We have also observed in the spectra ions corresponding to an unreacted precursor (m/z 646.0), the singly charge reduced precursor (m/z 968.5), the doubly charge reduced precursor (m/z 1935.7), along with the precursor ion existing as a +2 charge state possessing a nondissociative radical (m/z 1936.7). Only through the use of ETD was it possible to assign the site of glycosylation as N60.

Interestingly, a loss of 43 Da was observed for the z′5 ion originating from Man5 of ribonuclease B (see labeled peaks in Fig. 2(b)). However, this loss was not observed for the z′4 ion. This loss is most likely due to a radical-induced neutral loss of NH2C.NH2 from the arginine residue or the loss of the N-acetyl moiety originating from a GlcNAc unit, as previously observed in the case of glycopeptides derived from horseradish peroxidase.32 However, the inability to detect this loss in the case of the z4 ion suggests that it originates from the arginine residue. Moreover, no such losses were observed for any c′ ions containing arginine. A loss of 17 Da

![Figure 2. CID (a) and ETD (b) MS/MS of the triply charged Man5 tryptic glycopeptide derived from bovine ribonuclease B. Symbols: blue square: GlcNAc, green circle: mannose. This figure is available in color online at www.interscience.wiley.com/journal/rcm.](image-url)
was observed for the \( z^{-4} \) ion, which may be the result of a loss of ammonia from the N-terminal amine or the side chain of lysine. This was also previously observed and described by Catalina et al. in the case of glycopeptides derived from horseradish peroxidase.\textsuperscript{32} A similar loss was also observed for the +2 ion containing a radical (m/z 1936).

**CID and ETD of horseradish peroxidase**

Horseradish peroxidase is a glycoprotein with eight sites of glycosylation. Several types of glycans are present; however, the major glycan structure has been reported to be GlcNAc\(_2\)(Fuc)Man\(_3\)(Xyl).\textsuperscript{33} A glycopeptide derived (with the amino acid sequence of: LYNFSNTGLPDPTLNTTYLQTLR) from tryptically digested horseradish peroxidase contains two sites of glycosylation, N\(_{216}\) and N\(_{228}\). CID fragmentation of this glycopeptide ion (Fig. 3(a)) resulted in a spectrum which was difficult to interpret, as the multiple sites of glycosylation of this peptide can lead to many different fragmentation pathways. However, several diagnostic ions were present in this spectrum that did allow for a partial glycan sequence determination. In the CID spectrum, an ion is present at m/z 366 which is most likely Man-GlcNAc. A second ion, 162 m/z units higher at m/z 528.4, is GlcNAc-Man\(_2\). The addition of the third mannose unit is evidenced by the ion occurring at m/z 690. Thus, only a partial glycan sequence can be obtained from this spectrum. Moreover, the extensive fragmentation of the glycan, with partial glycan structures being retained, was assigned as illustrated in Fig. 3(a). In this figure, the asterisk represents the peptide backbone. More complex fragmentation patterns were observed due to the double glycosylation sites within this peptide. Unfortunately, the product ions originating from this glycopeptide are a mix of cleavages simultaneously occurring at both sites, thus making it impossible to assign definitive glycan structures to each of the glycosylation sites. For example, the ion observed at m/z 1599 indicates that the peptide has retained two GlcNAc residues along with a fucose residue. Similarly, the ions at m/z values of 1626 and 1525 represent the product ions retaining two and three GlcNAc units, respectively. The retention of the GlcNAc\(_5\) Man residue on the peptide backbone is represented by the ion at m/z 1606. The ion with GlcNAc\(_5\)Man\(_2\) has been observed at m/z 1686. At m/z 1700, three GlcNAc residues are retained on the peptide backbone in addition to a fucose residue. The ions at m/z values of 1707, 1781, and 1788 correspond to the peptide backbone with three GlcNAc residues and a mannose residue; three GlcNAc residues, a mannose residue, and a fucose residue; and three GlcNAc residues and two mannose residues, respectively. Glycopeptide fragments containing three GlcNAc residues, a fucose residue and a mannose residue are observed at m/z 1662, while the loss of one entire glycan residue is suggested by the fragment ion observed at m/z 1907. The fragment ion at m/z 2008 is indicative of an ion retaining one complete structure plus an additional GlcNAc residue. Although the CID spectrum is rich in fragment ions, the majority of which result from the fragmentation of the glycan moieties, the assignment of the glycosylation sites and their glycan attachment is practically impossible due to the doubly glycosylated nature of this particular glycopeptide.

![CID spectrum of glycopeptide](image1)

![ETD spectrum of glycopeptide](image2)

**Figure 3.** CID (a) and ETD (b) MS/MS of the doubly glycosylated glycopeptide LYNFSNTGLPDPTLNTTYLQTLR derived from tryptically digested horseradish peroxidase. The asterisk represents the peptide backbone. Symbols: blue square: GlcNAc, green circle: mannose, red triangle: fucose, orange star: xylose. This figure is available in color online at www.interscience.wiley.com/journal/rcm.
In comparison, 7 of the possible 19 c’ ions and 5 of the possible 20 z’ ions are observed in the ETD spectrum of the same glycopeptide (Fig. 3(b)). We were able to assign c’ ions from c’3–10 except for the c’9 ion which was not present due to the presence of a proline residue at this position. We were also able to assign the z’4 ions. Although the c’5 ion was not observed in the spectrum, the wealth of information in the ETD spectrum is sufficient to allow unequivocal identification of the amino acid sequence of this peptide, and thus assign both glycosylation sites. Observing the c’3 ion at m/z 1579.5 and the remaining c’ ions permits the assignment of the site of glycosylation as N216. Unfortunately, ETD fragmentation of this glycopeptide did not yield a z’10 fragment; thus we are not able to determine, from this data alone, the second site of glycosylation. However, combining the c’ and z’ series of ions permits the correct determination of the amino acid sequence for this glycopeptide. The fragments observed in the ETD spectrum allowed the amino acid sequence assignment, which in conjunction with the precursor ion m/z value and the charge state assignment suggests the presence of two glycosylation sites. If only a single glycan was present, the quadruply charged ion would have been observed at m/z 955. However, this ion was not seen, suggesting the presence of two glycosylation sites. The glycan structure attached to both glycosylation sites is GlcNAc2(Fuc)Man3(Xyl), which must be associated with N228.

Another glycopeptide originating from horseradish peroxidase possesses a tryptic peptide backbone of GLIQSDQELFSSPNATDTIPLVR, which produced a triply charged ion with an m/z value of 1274. This glycopeptide again produced a CID spectrum that was difficult to interpret (Fig. 4(a)). The ions at m/z values of 366, 528, and 690 are representative of the fragments composed of one, two or three mannose residues attached to a GlcNAc residue, respectively. Interestingly, the CID spectrum of this glycopeptide contained both the glycan fragmentation and extensive fragmentation of the peptide backbone indicated by a continuous series of y-type ions from y10–14. The fragment ions from y10–14 retained a single GlcNAc unit. These fragments observed in the spectrum are consistent with the recently published results.17

The ETD spectrum of the same glycopeptide is shown in Fig. 4(b). The spectrum is rich in valuable information pertaining to the peptide backbone. The ETD MS/MS spectrum included a series of z’ ions from z’5–12, representing 8 of the 20 possible z’ fragments. Interestingly, no c’ ions were observed for this glycopeptide. The reason for this remains unclear, yet it may be possible that these fragments were generated but remained attached due to intermolecular hydrogen bonding. The mass difference between the z’9 and z’10 ions is 1290, which is consistent with a glycan structure of GlcNAc2(Fuc)Man3(Xyl) attached to an asparagine residue, suggesting N285 as the glycosylation site.

**CID and ETD of human haptoglobin**

Human haptoglobin is a glycoprotein containing four sites of glycosylation. In healthy subjects, the major glycan associated with haptoglobin has been characterized as a disialylated biantennary complex glycan, while a trisialyated...
triantennary complex glycan is the most abundant of the minor glycans. Additionally, there are monosialylated biantennary and disialylated triantennary glycans present in trace amounts. Low levels of fucosylated trisialylated triantennary and disialylated biantennary complex glycans have also been observed by another group. Fucosylated haptoglobin glycans have been identified as a biomarker for pancreatic cancer and ovarian cancer. Moreover, increased levels of fucosylation have also been observed in patients with inflammatory joint disease, such as rheumatoid arthritis. Hence, monitoring changes in the glycosylation of this glycoprotein has significant diagnostic value and can be accomplished by combining CID and ETD data. This allows the glycan structure to be characterized by the determination of the sequence of the associated peptide, and perhaps, most importantly, the site of glycosylation and its microheterogeneity. This could be achieved with a single LC/MS/MS analysis using sub-microgram amounts of glycoproteins, which appears feasible.

The CID MS/MS spectrum of a glycopeptide derived from tryptically digested human haptoglobin is shown in Fig. 5(a). An ion observed at m/z 657 in this spectrum indicates the presence of sialic acid. This ion consists of the GlcNAc-Gal-NeuNAc monosaccharide sequence originating from the glycan structure. Additionally observed in the spectrum is an ion at m/z 366, which corresponds to the GlcNAc-Gal moiety.

The signal at m/z 1237.8 is a triply charged ion originating from the loss of a sialic acid (mass: 291 Da). The signal at m/z 1115.7 is a triply charged ion that results from the loss of one antennary structure composed of GlcNAc-Gal-NeuNAc (mass 656 Da), while the ion at m/z 1672.8 is the doubly charged version resulting from the same mass loss. The triply charged fragment ion present at m/z 1139.6 results from the loss of two sialic acid residues, while that at m/z 1345.1 is a doubly charged fragment ion that results from the loss of both antennary structures. A doubly charged fragment ion is present at 1527.6, resulting from the loss of a complete antennary structure and the loss of an additional sialic acid residue. An ion present at m/z 1447.7 is doubly charged and is the product ion originating from the loss of an antenna with its attached mannose residue, in addition to the loss of sialic acid from the second antenna. A doubly charged fragment ion at m/z 1246.1 is also observed, which originates from the loss of both antennas and a mannose residue. Another doubly charged fragment ion is present at m/z 1181.2, resulting from the loss of both antennas and two mannose residues. At m/z 1019.1, a doubly charged ion is present that corresponds to the peptide backbone with only the two core GlcNAc residues. Glycan structural assignment for this glycopeptide appears possible; however, no information pertaining to the peptide backbone sequence was observed in the CID spectrum.

Figure 5. CID (a) and ETD (b) MS/MS of the glycopeptide VVLHPNYSQVDIGLIK derived from human haptoglobin. The asterisks indicate the loss of water from one of the sialic acid structures. Symbols: blue square: GlcNAc, green circle: mannose, yellow circle: galactose, purple diamond: sialic acid. This figure is available in color online at www.interscience.wiley.com/journal/rcm.
The ETD spectrum of the same peptide demonstrated extensive fragmentation of the peptide backbone, and minimal fragmentation of the glycan structure (Fig. 5(b)). Below ca. 1000 \( m/z \), \( z' \) ions dominated the spectrum, with \( z'_{2–8} \) being present. In this region of the spectrum, only a single \( c' \) ion, \( c'_{5–8} \) was observed. However, at the \( m/z \) values above ca. 1000, the opposite trend was observed; the spectrum is dominated by \( c' \) ions; \( c'_{5–15} \) are present as doubly charged ions with no \( z' \) ions being present. In the ETD spectrum of this glycopeptide, 11 \( c' \) ions (73.3% of the 15 possible \( c' \) fragments) and 7 \( z' \) ions (46.7% of the 15 possible \( z' \) fragments) were observed. Overall, 60% of the 30 possible ions were observed in the ETD spectrum, thus facilitating assignment of the glycosylation site. Based on the detection of \( c'_{5–15} \), the glycosylation site of this glycopeptide was deduced to be at N241 and could only be assigned using ETD. Interestingly, ETD fragments containing the glycan structure were present with a loss of water and are designated with an asterisk in Fig. 5(b). This loss most likely originates from one of the sialic acid residues.

**Limitations of ETD**

ETD suffers from several drawbacks, limiting its applicability to a wide range of bioanalytical problems. First, the fragmentation efficiency of ETD is low. For the glycopeptides studied in this work, the fragmentation efficiency was approximately 20%. The major reaction pathways observed in this work were charge reduction reactions, as opposed to the needed fragmentation-reaction pathways. This has potentially adverse effects for important trace level analytes in the samples derived from biological sources.

The second major limitation encountered by ETD is its limited useful \( m/z \) range. ETD appears to be limited to an \( m/z \) range of less than about 1400. This was discussed previously for the peptide analysis and is demonstrated here for the glycopeptides derived from fetuin. The CID fragmentation of a tryptic glycopeptide derived from the digestion of bovine fetuin at \( m/z \) 1634 is shown in Fig. 6(a). This peptide has an amino acid sequence of RPTGEVYIDTLETCHVLDPPLANCSVR with a trisialyated triantennary complex glycan. The CID spectrum of this peptide demonstrates extensive glycosidic fragmentation and the lack of any peptide backbone fragmentation. However, when this same glycopeptide was subjected to ETD fragmentation, no diagnostic glycopeptide backbone cleavage products were observed. This trend was also observed for other glycopeptides above \( m/z \) 1400 (data not shown).

Recently, the abovementioned limitations have been partially overcome by performing low-energy CID on the charge-reduced precursor ions generated by ETD. Using an LTQ mass spectrometer (Thermo-Fisher, Santa Clara, CA, USA), the Coon group significantly improved the fragmentation efficiency for doubly charged ions by implementing a procedure termed EtCaD (electron-transfer/collisionally-
activated dissociation). To generate \( c' \) and \( z' \) ions, the Q-activation value was lowered from 0.25 to 0.18 and the normalized collision energy was reduced from 35% to 20%. However, when this method was applied to the analysis of phosphopeptides, the loss of phosphate was observed, indicating the need to further optimize this methodology for PTM analysis.\(^{37}\)

More recently, this method has been refined by Karger and coworkers.\(^{38}\) A more gentle CID method was developed and was termed CR-CID (charge-reduced collision-induced dissociation). In this technique, the Q-activation value was reduced to 0.15 (0.25 is commonly utilized), while the normalized collision energy was lowered to 10% (35% is commonly employed). When applying this method to the analysis of phosphorylation and glycosylation, the PTMs were retained on the peptide backbone. This method was successfully applied to a 37-amino acid glycopeptide derived from EGFR with an \( m/z \) value of 1900. By using ETD only, this

![Figure 7. CID MS\(^3\) of charge-reduced species generated by ETD of glycopeptide CANLVVPVIPNATLDQITGK with a triantennary, trisialylated glycan (a) and glycopeptide QDOQICYNTTYLNVQR with a triantennary, disialylated glycan (b). Both glycopeptides are derived from \( \alpha\)-acidic glycoprotein. Symbols: blue square: GlcNAc, green circle: mannose, yellow circle: galactose, purple diamond: sialic acid. This figure is available in color online at www.interscience.wiley.com/journal/rcm.](https://www.interscience.wiley.com/journal/rcm)
glycopeptide resulted in eight z’ ions, while the CR-CID produced 19 z’ ions and a single y ion located on an aspartic acid/proline pair.

However, when we attempted to reproduce these results on our instrument using glycopeptides derived from tryptically digested α-1-acid glycoprotein, we observed typical CID fragmentation. The CID spectra acquired from the charge-reduced precursor resulted only in glycosidic bond cleavage; no ETD-type fragments were observed, as shown in Figs. 7(a) and 7(b). This is due to the inability to independently vary the CID energy parameters, thus hindering the possibility of duplicating the ETCaD and the CR-CID results. Therefore, the limited m/z range of ETD associated with the mass spectrometer used here results in a reduced application to characterize N-linked glycopeptides with complex glycan structures. ETD, in general, appears to be very effective for the characterization of some N-linked glycopeptides (depending on the instrument being used), phosphopeptides, O-linked glycans, and other PTMs.

CONCLUSIONS

In this study, several glycopeptides derived from model glycoproteins were investigated using a mass spectrometer capable of performing both CID and ETD experiments. As demonstrated previously by two different groups, the CID spectra showed extensive fragmentation of the glycan moiety, with very limited fragmentation of the associated peptide backbone. However, the CID MS/MS spectra permitted the determination of glycan structures. The CID information permitted the characterization of the microheterogeneity of a glycosylation site, but little information pertaining to the amino acid sequence of a peptide backbone can be deduced, making the site of modification difficult to assess. ETD offers a fragmentation complementary to CID, allowing the determination of the amino acid sequence and the glycosylation site(s) of a glycopeptide. In ETD fragmentation, the peptide backbone is fragmented, yielding a series of c’ and z’ ions, while the attached glycans remained largely intact and are retained on the peptide backbone throughout the process. This allowed not only the determination of the peptide amino acid sequence, but the fact that the glycan remained attached to the backbone facilitated the unambiguous determination of the site of glycosylation and its microheterogeneity.

When the data acquired from both modes of fragmentation are combined, it is possible to characterize fully the glycopeptides in many cases. Combination of the two complementary fragmentation techniques should allow investigators to characterize rapidly and unequivocally the glycan composition and its attachment to a peptide backbone in a single LC/MS/MS experiment utilizing sub-microgram quantities of glycoproteins.

Acknowledgements

This work was supported by the METACyt Biochemical Analysis Center, a component of the Indiana Metabolomics and Cytomics Initiative (METACyt), funded by a grant from Eli Lilly Endowment, Grant No. GM24349 from the National Institute of General Medical Sciences, U.S. Department of Health and Human Services, and Grant No. RR018942 from the National Center for Glycomics and Glycoproteomics (NCGG) at Indiana University.

REFERENCES