Published in final edited form as:
Anal Chem. 2010 December 15; 82(24): 10095–10101. doi:10.1021/ac1020722.

Improved mass spectrometric characterization of protein glycosylation reveals unusual glycosylation of maize-derived bovine trypsin

Hao Zhang[a], **Richard Y-C Huang**[a], **Pegah R. Jalili**[b], **Janet W. Irungu**[b], **Gordon R. Nicol**[b], **Kevin B. Ray**[b], **Henry W. Rohrs**[a], and **Michael L. Gross**[a,*]

[a]Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130

[b]Sigma-Aldrich Corporation, 2909 Laclede Ave. St. Louis, Missouri 63103

Abstract

Although bottom-up proteomics using tryptic digests is widely used to locate posttranslational modifications (PTM) in proteins, there are cases where the protein has several potential modification sites within a tryptic fragment, and MS^2 strategies fail to pinpoint the location. We report here a method using two proteolytic enzymes, trypsin and pepsin, in combination followed by tandem mass spectrometric analysis to provide fragments that allow one to locate the modification sites. We used this strategy to find a glycosylation site on bovine trypsin expressed in maize (TrypZean™). Several glycans are present, and all are attached to a nonconsensus *N*-glycosylation site on the protein.

Among the many applications of mass spectrometry (MS) today, one of the most important is the characterization of post-translational modifications (PTMs)1, 2. Glycosylation is an important protein modification and serves a variety of functions in cellular events including signaling and recognition3, 4. Approximately 50% of human proteins are glycosylated5. There are two major glycosylation types, N-linked and O-linked. In the former, oligosaccharides attach at the amide nitrogen of asparagine belonging to a consensus sequence Asn-X-Ser/Thr (where X can be any amino acid other than proline). In O-linked glycosylation, oligosaccharides attach at the hydroxyl oxygen of a serine or threonine 6, 7.

Among the MS-based PTM characterization approaches, the most common is a bottom-up strategy in which proteins are digested enzymatically to create small peptides for MS analysis8. Because this often leads to complex mixtures, liquid chromatography (LC), especially reverse phase liquid chromatography, is used to separate the peptides. For example, Gygi and coworkers9 used bottom-up proteomics to identify a total of 13,720 phosphorylation sites on proteins extracted from *Drosophila* embryos. For determination of glycosylation, the standard bottom up workflow needs to be modified because collisional activation of glycosylated peptides does not usually cause cleavage of the peptide backbone to give sequence information. Several alternative strategies include enzymatic and chemical deglycosylation, chemical derivatization, and non-specific enzyme digestion10, and recent review articles cover the development of MS-based studies of glycosylation11–16. Glycoproteins can be isolated and analyzed intact to gather quantitative information about glycosylation and the number of glycoforms17, 18. Alternatively, the glycan can be removed and isolated, sacrificing information about the location of the PTM. Methods that

*CORRESPONDING AUTHOR FOOTNOTE. Michael L. Gross, Department of Chemistry, Washington University, One Brookings Drive, Campus Box 1134, St. Louis, MO 63130 mgross@wustl.edu.

rely on the nonspecific enzymatic digestion of the glycoprotein to glycopeptides were reported by the Lebrilla19 and the Massolini20 groups. They used immobilized pronase to cut glycoproteins into short glycopeptides, small dipeptides, and amino acids and analyzed the digested product mixtures by LC-MS or MALDI MS to determine the glycan structure and the attachment to the protein prior to deglycosylation. Pepsin also can give better sequence coverage of glycoproteins21, 22.

Here we report the characterization of an unusual N-linked glycosylation on maize-derived bovine trypsin by combining immobilized pepsin digestion, trypsin digestion, and tandem MS (Figure 1). While bovine trypsin is a commonly used enzyme for protein processing, this enzyme was expressed in maize to avoid the presence of unwanted animal-derived proteins when this enzyme is used in proteomics experiments. Previous studies reported that maize-derived bovine trypsin is functionally equivalent to native bovine pancreatic trypsin23. Although there is no glycosylation in the native bovine trypsin, glycosylation of this recombinant protein was detected by SDS/PAGE and glycoprotein staining23. Thus far, the type of glycosylation and its location remain unknown. Given there is no consensus sequence for N-linked glycosylation, it was expected that it would be O-linked.

A variety of approaches were described recently to provide detailed protein and glycosylation information on recombinant protein, but they were not able to detect the site of glycosylation24. To pinpoint the specific amino acid residue that is glycosylated and, thus, determine the type of glycosylation present, we integrated bottom-up proteomics with pepsin digestion as an alternative and fast way to locate glycosylation sites. Because pepsin is effective in acidic conditions, it fits nicely into a bottom-up proteomics workflow where the aqueous phase for gradient chromatography is often aqueous 0.1% formic acid. Pepsin digestion is easily applied to tryptic peptides and can break them into smaller pieces in less than 30 min. Immobilized pepsin is readily available and is easily removed with a centrifuge, thus minimizing interference with the original sample. Many, including the Washington University lab, have worked with pepsin for several years in MS-based hydrogen/deuterium (H/D) exchange experiments25, and have considerable experience using this enzyme.

EXPERIMENTAL SECTION

Materials

Maize-derived bovine trypsin (trade name “TrypZean™”), water, acetonitrile, and formic acid were obtained from Sigma-Aldrich (St. Louis, MO) at the highest purity available. Immobilized pepsin on agarose was purchased from Thermo Fisher Scientific (Rockford, IL).

Separation of tryptic peptides

Maize-derived bovine trypsin protein (TrypZean) was dissolved in 0.2% formic acid solution to make 200 μL of 500 μM solution. The solution was found to be a mixture of TrypZean and its self-digested peptides, as was confirmed by LC-MS analysis. High molecular weight materials were separated by using a centrifuge at 14000g for 15 min to pass the solution through a 10,000 MWCO filter (Millipore, Kankakee, IL). The flow-through solution containing peptides was taken for LC-MS analysis or for further pepsin digestion.

Pepsin Digestion Protocol

An aliquot of 15 μL of immobilized pepsin on agarose was first washed with water containing 0.2% formic acid. In each washing step, pepsin in water containing 0.2% formic

acid was vortexed for 10 s and briefly centrifuged for 3 s to bring down the pepsin beads. The supernatant was discarded, and new wash solution was added. The washing steps were repeated 5 times. Pepsin digestion of the TrypZean peptides was carried out by adding 20 μL of the filtered peptide solution to the washed pepsin beads and keeping the solution at 37 °C for 30 min. The mixture was resuspended every 5 min. After digestion, the solution was centrifuged briefly (2–3 s) to bring down the agarose beads and the supernatant was taken for LC-MS analysis.

LC-ESI-MS/MS Analysis of Protein Digests

Tryptic or peptic digests (1 μL, 200 fmole digested protein) were diluted in 10 μL water containing 0.1% formic acid. An aliquot (5 μL) was loaded onto a silica capillary column with a PicoFrit tip (New Objective, Inc., Woburn, MA) that was custom packed with C18 reverse phase material (Magic, 0.075 mm × 150 mm, 5 μm, 120 Å, Michrom Bioresources, Inc., Auburn, CA). The gradient was supplied by an Eksigent NanoLC-Ultra 1D (Eksigent Technologies, Inc. Livermore, CA) and was run from 2% solvent B (acetonitrile, 0.1% formic acid) to 60% solvent B over 60 min, then to 80% solvent B for 10 min at 260 nL/min followed by a 12 min re-equilibration step with 100% solvent A (water, 0.1% formic acid). The flow was directed by a PicoView Nanospray Source (PV550, New Objective, Inc., Woburn, MA) into an LTQ Orbitrap (Thermo-Scientific, San Jose, CA) with a spray voltage of 1.8–2.0 kV and a capillary voltage of 27 V. The LTQ Orbitrap was operated in standard data-dependent acquisition mode controlled by Xcalibur 2.0.7 software. Peptide mass spectra (*m/z* range: 350–2000 for tryptic peptides and 300–2000 for peptic peptides) were acquired at high mass resolving power (60,000 for ions of *m/z* 400) with the FT mass spectrometer. The automatic gain control (AGC) value was set at 1×10^6 (Orbitrap) and 3×10^4 (ion trap) with a maximum injection time of 1000 ms (Orbitrap) and 50 ms (ion trap). The LTQ Orbitrap was externally calibrated using a standard calibration mixture of caffeine, MRFA and Ultramark 1621. The mass calibration was checked and repeated before the LC-MS experiments to optimize mass measurement accuracy.

In the data-dependent mode, the six most abundant multiply-charged ions with a minimum intensity of 1000 counts were subjected to collision induced dissociation (CID) or electron transfer dissociation (ETD)26 in the linear ion trap. Precursor activation in CID was performed with an isolation width of 2 Da, an activation time of 30 ms, and normalized collision energy of 35%. For ETD, fragmentation was performed with an isolation width of 3 Da with charge dependent activation time. To improve the quality of the product-ion spectra of the glycopeptides, the mass list of glycopeptides of interest was loaded into the global mass list with their HPLC retention times. The six most abundant ions that matched the input mass list within 10 ppm with a minimum intensity at 500 counts were subjected to CID in the linear ion trap.

For the MS^3 experiment, the three most abundant fragment ions from an MS^2 CID experiment with a minimum intensity of 50 counts were subjected to CID to give detailed sequence information. The MS^3 CID was performed with an isolation width of 5 Da and activation time 30 ms; the normalized collision energy was set at 35%.

Database searching and analysis

Product-ion (MS^2) spectra were processed into mgf files for database searching with DTASuperCharge (V 2.0a6 from the open source proteomics platform MSQuant, version 2.0a81) using default settings27 Resulting mgf files were searched using Mascot (version 2.2.0, Matrix Science, London, UK) against the NCBI nonredundant database with no taxonomic restriction28. The enzyme was set to “trypsin” for tryptic peptides and “none” for peptic peptides. The maximum number of missed cleavages was two for tryptic peptides and

zero for peptic peptides. Methionine oxidation was the only variable modification. The peptide mass tolerance was 15 ppm, allowing for one ^{13}C peak, and the peptide charge was allowed to be +1, +2, or +3. The production mass tolerance was 0.8 Da. Given that the protein of interest was well known, most of the spectral interpretation was manual. Lists of product ions for peptides of interest were generated in silico using MS-Product (Protein Prospector, version 5.6.1, http://prospector.ucsf.edu). Resulting product ion lists were compared with product-ion (MS^2) spectra.

RESULTS AND DISCUSSION

A mass spectrometric analysis from previous work of intact TrypZean showed a set of peaks larger than the molecular weight of bovine trypsin corresponding to glycosylated form23, 24. The distinct peaks were approximately 203 Da apart, consistent with glycosylation23. Previous LC-MS analysis of the tryptic fragments of TrypZean revealed that the difference between the glycosylated and nonglycosylated forms of the protein resided on the tryptic fragment ^{70}SIVHPSYNSNTLNNDIMLIK^{89} but the specific site of glycosylation was not identified 24.

There are eight possible glycosylation sites on this sequence: four potential O-linked and four potential non consensus N-linked sites. Glycoforms of any of these sites would have the same intact protein masses. Therefore, sequencing is required to identify the glycosylation site on TrypZean.

Although LC-MS is a powerful tool for characterizing many protein PTMs, characterization of glycosylation with tandem MS remains difficult. CID of glycopeptides often yields fragment ions that correspond to cleavages of the glycan, thus providing little information about the site of glycosylation on the peptide sequence. This proved to be the case for the glycopeptide Ser_{70}-Lys_{89} (Figure 2). New electron-driven dissociation technology29, 30 affords peptide backbone fragmentation without the loss of labile PTMs such as phosphorylyation and glycosylation29. Several studies showed that the sites of glycan attachment can be determined by using ETD31 or ECD32. We employed directed33 and data-dependent ETD LC-MS, including the newly available charge-dependent ETD on the Thermo Orbitrap. Previous work provided the *m/z* values of various glycopeptides from TrypZean24. To locate the site of glycosylation, these glycopeptides were subjected to data dependent or directed CID and ETD analysis. The glycopeptide at *m/z* 1154 corresponding to Ser_{70}-Lys_{89} + $Hex_3HexNAc_2Fuc_1Xyl_1$ was selected as an example to illustrate how the analysis was performed. We observed several peptide backbone cleavages in the resulting series of c and z ions, as illustrated in the ETD spectrum (Figure 3). Owing to the large additional mass of the glycan moieties (1000–1800 Da), some glycosylated peptide fragment ions had *m/z* values that were out of the ion trap mass range (*m/z* 350–2000), and the observed c and z ions only covered 60% of the peptide sequence. Furthermore, oxidation (+16) at Met_{87} caused mass shifts in both the c and z ion series for peptides including Met_{87}.

Although a Mascot search found no glycosylation, a manual search using ions predicted by Protein Prospector did; both c and z ion series suggested glycosylation at Asn_{77}. The c and z ions showing the glycan attachments originated from high-mass precursor ions of high-charge states. Owing to the relative low efficiency of ETD and the low glycopeptide ion abundance, only five of the most abundant glycopeptide ions could be dissociated by ETD during the LC-MS run (Table S1). Even with only 60% sequence coverage, c and z fragment ions from different forms of glycopeptides suggested that Asn_{77} is the glycosylation site. The ion series containing the modified site was not complete. Therefore, the fragment ions without modification, while highly suggestive, do not provide a definitive location for the glycosylation. Because glycosylation reduces the proton affinity of the glycopeptides, the

resulting lower charge states move the *m/z* of large glycopeptides like this one outside the optimal range of the ion trap. These lower charge states are also problematic for ETD, which is more effective with higher charge states. Evidence from ETD and CID spectra showed that all seven forms of glycopeptide Ser70-Lys89 have similar glycosylation fragmentation patterns and pointed to Asn_{77} as the site of attachment.

To improve the sequence coverage of TrypZean glycopeptide Ser_{70}-Lys_{89} and to locate more precisely the sites of glycosylation, we employed two strategies. First, MS^3 CID was conducted on product ions that exhibited neutral losses of oligosaccharide fragments. Given that the most abundant ions formed in the MS^2 experiment are those from which oligosaccharides were lost, it seemed reasonable to pursue subsequent steps of collisional activation. Moreover, previous studies showed a promising outcome with two sets of b and y ions in product-ion spectrum from MS^3 of glycopeptides34. With most of the glycan removed, the available energy from additional steps of collisional activation likely would lead to cleavage of the desired peptide bonds. The most abundant ion was a peptide ion with glycosidically linked HexNAc (peptide-HexNAc). CID of its product ions was conducted in the data-dependent mode by setting up the mass spectrometer to fragment the three most abundant ions formed in the MS_2 of the glycopeptides. We observed two sets of b and y ions, one glycosylated and the other nonglycosylated in the MS^3 experiment (resulting product-ion spectrum of peptide-HexNAc of *m/z* 1246.88 is in Figure 4). The presence of two sets of ions increased the complexity of spectrum. Owing to overlap of fragment ions from both glycosylated and nonglycosylated peptides, the fragment ions due to water or ammonia losses, and the low signal-to-noise ratios in several regions of the spectrum, the confidence in assigning the glycan site was dramatically reduced with increasing peptide length. As in the ETD spectrum, the MS^3-generated product-ion spectrum of tryptic peptide-HexNAc did not provide a complete unambiguous ion series that could locate the glycosylation site.

In the second approach, we included an additional digestion step using pepsin, which is effective in protein digestion for H/D amide exchange experiments. The digestion made use of immobilized pepsin to cleave non-specifically the tryptic peptide of interest into smaller pieces. At low pH, pepsin cleaves the C-terminal of most amino acids with an aromatic side chain. While low pH can lead to hydrolysis of glycans, we consistently found that the conditions employed here (0.1% formic acid, pH 2, digestion time from 2 to 30 mins) did not. Systematic studies of pepsin for H/D exchange show that immobilized pepsin has an average cleavage probability of ~14% under acidic conditions35. It also retains its cleavage efficiency in the presence of nearby glycan, unlike trypsin21. We reasoned that digestion would reduce the length of the TrypZean peptides and improve our sequence coverage, specifically for TrypZean glycopeptide Ser_{70}-Lys_{89}. We also could control the peptide length by varying the time of pepsin digestion. Based on the product-ion spectra obtained for the peptic peptides, the glycosylated site was definitively identified as Asn_{77}, in agreement with the ETD results. The data ruled out O-linked glycosylation. A targeted MS^3 experiment on the peptic peptides that retained glycosylation confirmed this interpretation. Given that the lengths of peptic peptides are smaller than that of original tryptic peptide, the quality of product-ion spectra (MS^3) is improved, and these spectra show the location of the glycan attachment (Figure 5). Due to the low charge states (1+ or 2+) of small peptic peptides, ETD produced low quality spectra that did not locate the glycosylation site.

We observed a total 12 peptic peptides from the TrypZean glycopeptide Ser_{70}-Lys_{89} in the LC-MS experiment (Figure 6). Using extracted ion chromatograms based on accurate mass measurement and MS^2, we detected seven glycosylated peptic fragments above the noise level. Although it is possible there are other glycan sites in peptides hiding below the noise

threshold, the most abundant glycans are attached at Asn77. This use of immobilized pepsin on the tryptic fragments of TrypZean provided the most definitive results.

CONCLUSIONS

Proteomics-based LC-MS experiments that include pepsin digestion have potential for determining glycosylation or other modification sites. Pepsin digestion can be easily incorporated into a standard proteomics workflow with minor changes to the digestion protocol and the database search. Immobilized pepsin digests rapidly and can be removed easily by centrifugation. Its reactivity can be controlled with time, temperature, and pH, and this allows optimizing desired peptide backbone cleavages while minimizing loss of PTMs. The nonspecific cleavages by pepsin give more sequence coverage for glycopeptides in tandem MS experiments. The reduced size of the peptides leads to more confident identification of the modified sites than for some tryptic peptides that have multiple possible modification sites that are N-linked, O-linked, or both. The reduction in size of the peptides helps bypass current experimental limitations when large glycosylated peptide ions appear with low charge states. This method should work well for identifying multiple glycosylation sites, either N or O-linked, in single peptides produced with pepsin alone or in combination with other enzymatic digestion.

Although application of this method was successful in resolving the modification site of Trypzean, it does not supplant the standard bottom-up strategy using trypsin. Many modifications are easily identified by using standard methods, especially if there is only one modification and one site or if the possible sites are widely separated in a tryptic fragment. The addition of pepsin digestion will be most useful in locating complicated PTMs where there are multiple modifications and/or multiple modification sites. Because pepsin is much less specific than trypsin, the signal for each peptic fragment will be reduced, and more sample will be needed for this approach. This problem can be alleviated by including an enrichment step (e.g., TiO_2 enrichment column for phosphopeptides, lectin enrichment column for glycopeptides). An alternative approach, top-down proteomics, may also provide PTM information, but the required instrumentation is expensive and not yet widely available.

Using this approach, we identified an unusual glycosylation site by implementing these relatively simple changes to a commonly employed trypsin-based proteomics work flow. The outcome provides, to our knowledge, the first definitive experimental evidence proving the existence of the nonconsensus N-linked glycosyation in this transgenic product. Although we cannot rule out low-level glycosylation at other sites, that at Asn77 is the most abundant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to Drs. Jim Walters and Manolo Plasencia for advice on the analysis of glycosylation. The principal funding for this research was from the National Centers for Research Resource of the NIH (2P41RR000954). Some support was from Merck, and MLG is a consultant for Merck.

REFERENCES

1. Mann M, Jensen ON. Nat Biotechnol. 2003; 21:255–261. [PubMed: 12610572]
2. Lebrilla CB, Mahal LK. Curr Opin Chem Biol. 2009; 13:373–374. [PubMed: 19733502]

3. Dennis JW, Granovsky M, Warren CE. Bioessays. 1999; 21:412–421. [PubMed: 10376012]
4. Lebrilla CB, An HJ. Mol Biosyst. 2009; 5:17–20. [PubMed: 19081926]
5. Wong CH. J Org Chem. 2005; 70:4219–4225. [PubMed: 15903293]
6. Kornfeld R, Kornfeld S. Annu Rev Biochem. 1985; 54:631–664. [PubMed: 3896128]
7. Schachter H. Biochem Cell Biol. 1986; 64:163–181. [PubMed: 3521675]
8. Hoffman MD, Sniatynski MJ, Kast J. Anal Chim Acta. 2008; 627:50–61. [PubMed: 18790127]
9. Zhai B, Villen J, Beausoleil SA, Mintseris J, Gygi SP. J Proteome Res. 2008; 7:1675–1682. [PubMed: 18327897]
10. An HJ, Froehlich JW, Lebrilla CB. Curr Opin Chem Biol. 2009; 13:421–426. [PubMed: 19700364]
11. Dell A, Morris HR. Science. 2001; 291:2351–2356. [PubMed: 11269315]
12. Zaia J. Chem Biol. 2008; 15:881–892. [PubMed: 18804025]
13. Zaia J. Mass Spectrom Rev. 2004; 23:161–227. [PubMed: 14966796]
14. Bond MR, Kohler JJ. Curr Opin Chem Biol. 2007; 11:52–58. [PubMed: 17174139]
15. Wuhrer M, Catalina MI, Deelder AM, Hokke CH. J Chromatogr B Analyt Technol Biomed Life Sci. 2007; 849:115–128.
16. Harvey DJ. Expert Rev Proteomics. 2005; 2:87–101. [PubMed: 15966855]
17. Fridriksson EK, Beavil A, Holowka D, Gould HJ, Baird B, McLafferty FW. Biochemistry. 2000; 39:3369–3376. [PubMed: 10727230]
18. Reid GE, Stephenson JL Jr, McLuckey SA. Anal Chem. 2002; 74:577–583. [PubMed: 11838679]
19. Clowers BH, Dodds ED, Seipert RR, Lebrilla CB. J Proteome Res. 2007; 6:4032–4040. [PubMed: 17824634]
20. Temporini C, Perani E, Calleri E, Dolcini L, Lubda D, Caccialanza G, Massolini G. Anal Chem. 2007; 79:355–363. [PubMed: 17194161]
21. Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L, Zou H. J Proteome Res. 2009; 8:651–661. [PubMed: 19159218]
22. Carlsson SR, Fukuda M. J Biol Chem. 1990; 265:20488–20495. [PubMed: 2243102]
23. Woodard SL, Mayor JM, Bailey MR, Barker DK, Love RT, Lane JR, Delaney DE, McComas-Wagner JM, Mallubhotla HD, Hood EE, Dangott LJ, Tichy SE, Howard JA. Biotechnol Appl Biochem. 2003; 38:123–130. [PubMed: 12749769]
24. Jalili, P.; Nicol, G.; Irungu, J.; Ray, K. 56th ASMS Conference on Mass Spectrometry and Allied Topics; 2008.
25. Sperry JB, Shi X, Rempel DL, Nishimura Y, Akashi S, Gross ML. Biochemistry. 2008; 47:1797–1807. [PubMed: 18197706]
26. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Proc Natl Acad Sci U S A. 2004; 101:9528–9533. [PubMed: 15210983]
27. Mortensen P, Gouw JW, Olsen JV, Ong SE, Rigbolt KT, Bunkenborg J, Cox J, Foster LJ, Heck AJ, Blagoev B, Andersen JS, Mann M. J Proteome Res. 9:393–403. [PubMed: 19888749]
28. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551–3567. [PubMed: 10612281]
29. Mikesh LM, Ueberheide B, Chi A, Coon JJ, Syka JE, Shabanowitz J, Hunt DF. Biochim Biophys Acta. 2006; 1764:1811–1822. [PubMed: 17118725]
30. Wiesner J, Premsler T, Sickmann A. Proteomics. 2008; 8:4466–4483. [PubMed: 18972526]
31. Catalina MI, Koeleman CA, Deelder AM, Wuhrer M. Rapid Commun Mass Spectrom. 2007; 21:1053–1061. [PubMed: 17311219]
32. Kjeldsen F, Haselmann KF, Budnik BA, Sorensen ES, Zubarev RA. Anal Chem. 2003; 75:2355–2361. [PubMed: 12918977]
33. Schmidt A, Claassen M, Aebersold R. Curr Opin Chem Biol. 2009; 13:510–517. [PubMed: 19775930]
34. Nilsson J, Ruetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm G, Larson G. Nat Methods. 2009; 6:809–811. [PubMed: 19838169]
35. Hamuro Y, Coales SJ, Molnar KS, Tuske SJ, Morrow JA. Rapid Commun Mass Spectrom. 2008; 22:1041–1046. [PubMed: 18327892]

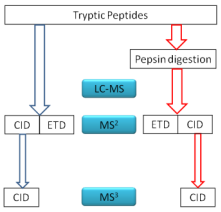

Figure 1.

Modification of standard bottom-up MS to investigate protein glycosylation sites by integrating pepsin digestion and MS^3 into the original work flow. Tryptic peptides were separated by HPLC then analyzed by CID and ETD in MS^2 to provide protein sequence and PTM information. MS^3 increased the sequence coverage of glycopeptides. Alternatively, tryptic peptides could be digested further with pepsin to give smaller peptides for CID analysis and improve the identification of PTM sites, especially in peptides that have multiple modification sites.

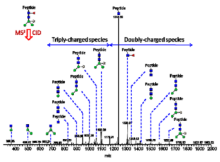

Figure 2.

Product-ion (CID) spectrum of TrypZean glycopeptide Ser_{70}-Lys_{89} (peptide + $Hex_3HexNAc_2Xyl_1Fuc_1$). The triply charged ion of *m/z* 1154.5253 was the precursor ion. [N-acetylglucosamine, or HexNAc = blue square; Fucose, or Fuc = red triangle; Hexose, or Hex = green circle; Xylose, or Xyl = pink star]

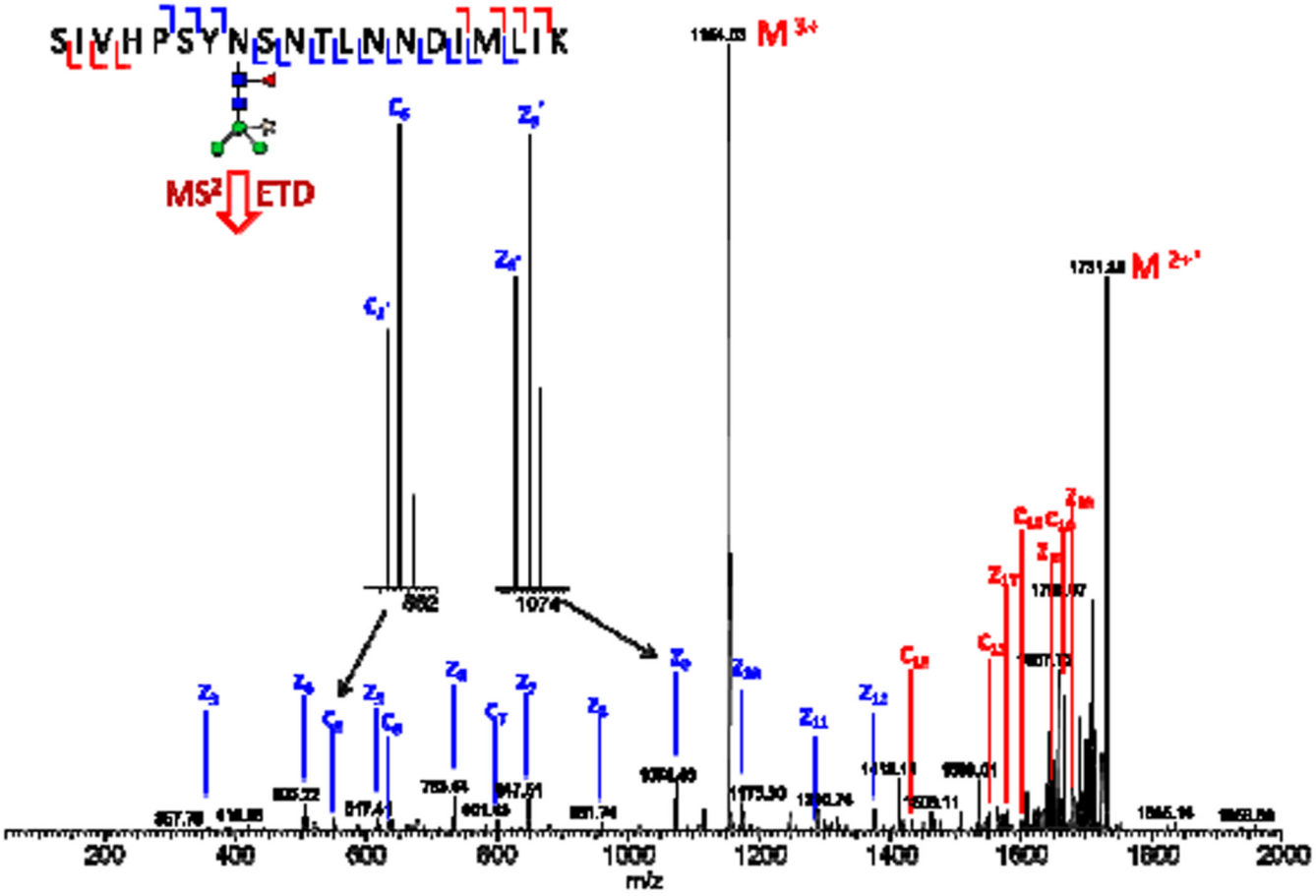

Figure 3.

Product-ion (ETD) spectra of TrypZean glycopeptide Ser_{70}-Lys_{89} (peptide + $Hex_3HexNAc_2Xyl_1Fuc_1$). The triply charged ion of *m/z* 1154.5253 is the precursor ion. The peptide fragment ions without glycan attached are labeled in blue, whereas glycan-attached peptide fragment ions are labeled in red. [N-acetylglucosamine, or HexNAc = blue square; fucose, or Fuc = red triangle; hexose, or Hex = green circle; xylose, or Xyl = pink star].

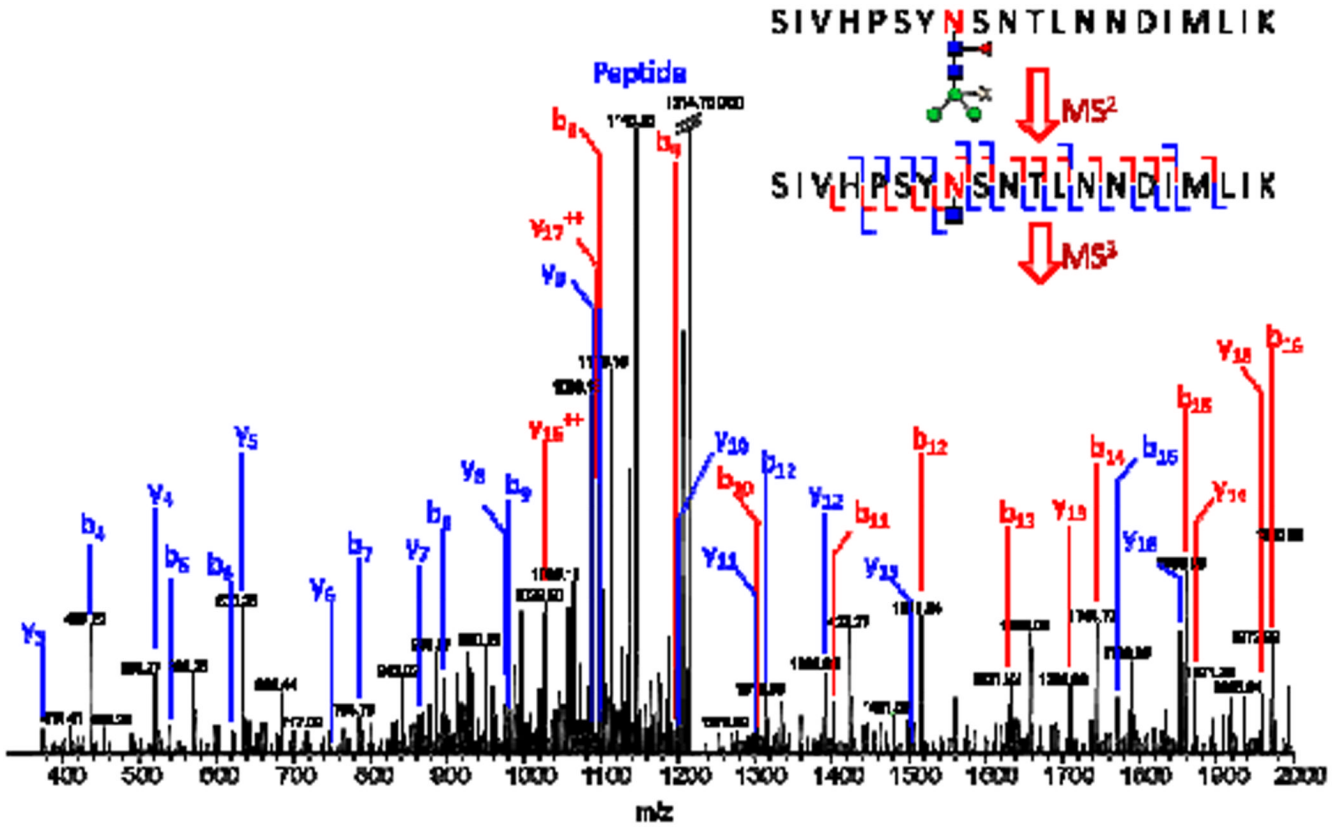

Figure 4.

Product-ion (MS^3) spectrum of TrypZean glycopeptide Ser_{70}-Lys_{89} (peptide + $Hex_3HexNAc_2Xyl_1Fuc_1$). The triply-charged precursor ion of *m/z* 1154.5253 was isolated and dissociated by CID. The most abundant product ion of *m/z* 1246.9, the glycosidically linked HexNAc (peptide-HexNAc), was isolated and activated by CID. The peptide fragment ions without glycan attached are labeled in blue, whereas glycan-attached peptide fragment ions are labeled in red. [*N*-acetylglucosamine, or HexNAc = blue square; fucose, or Fuc = red triangle; hexose, or Hex = green circle; xylose, or Xyl = pink star].

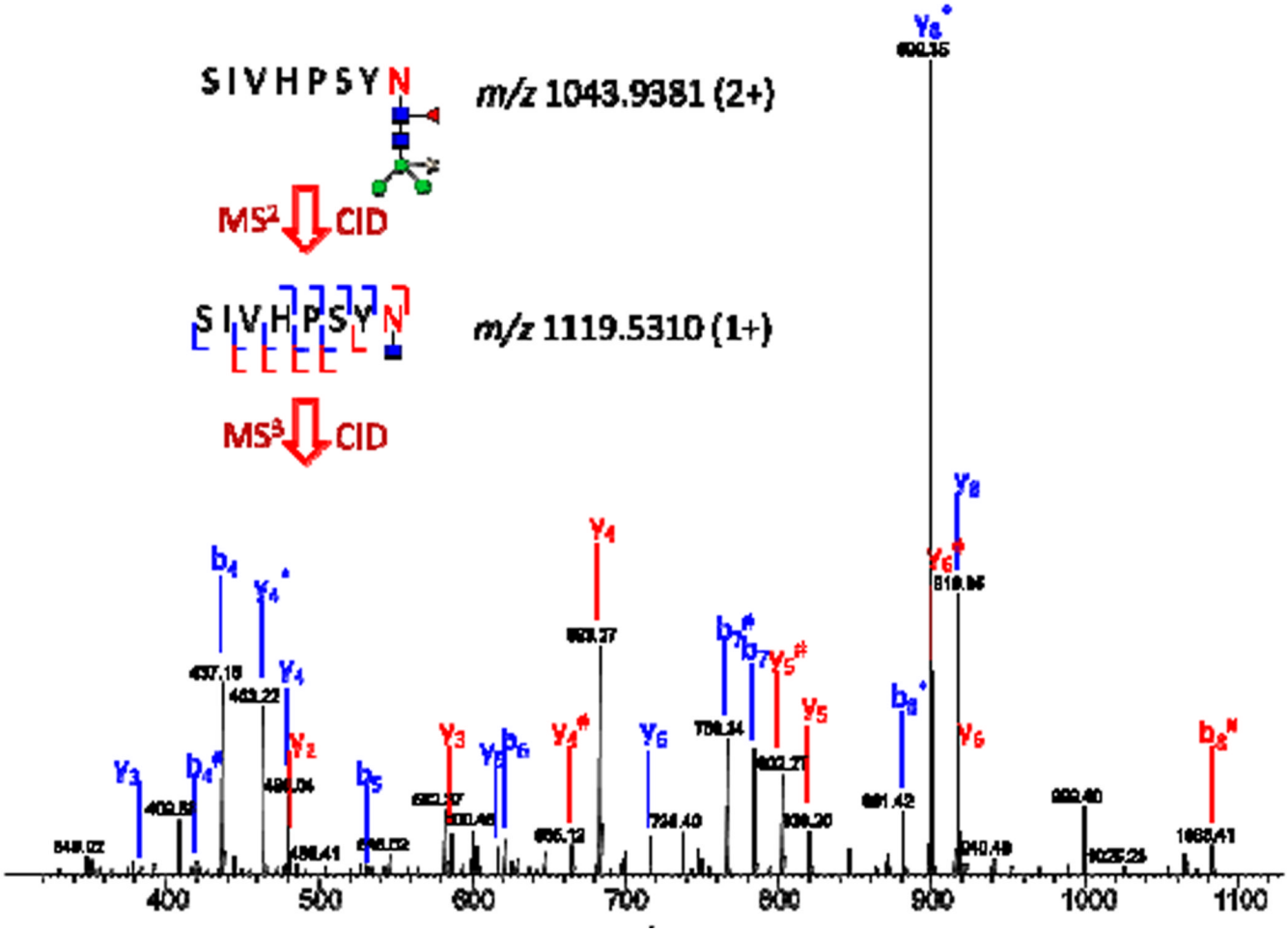

Figure 5.

Product-ion (MS^3) spectrum of TrypZean glycopeptide Ser_{70}-Asn_{77} (peptide + $Hex_3HexNAc_2Xyl_1Fuc_1$). The doubly charged precursor ion of *m/z* 1043.9381 was isolated and dissociated by CID. The most abundant product ion of *m/z* 1119.53, the glycosidically linked HexNAc (peptide-HexNAc), was isolated and activated by CID. The peptide fragment ions without glycan attached are labeled in blue, whereas glycan-attached peptide fragment ions are labeled in red. Peaks corresponding to ammonia loss are labeled with asterisks(*), and those corresponding to water loss are labeled with octothorpes(#). [N-acetylglucosamine, or HexNAc = blue square; fucose, or Fuc = red triangle; hexose, or Hex = green circle; xylose, or Xyl = pink star].

Peptide induced by Pepsin Digestion	Glycosylation	Oxidation
SIVHPS	No	No
SIVHPSY	No	No
SIVHPSYN	Yes	No
SIVHPSYNS	Yes	No
SIVHPSYNSN	Yes	No
SIVHPSYNSNTLN	Yes	No
SIVHPSYNSNTLNN	Yes	No
SIVHPSYNSNTLNNDI	Yes	No
SIVHPSYNSNTLNNDIM	Yes	Yes
NDIMLIK	No	Yes
DIMLIK	No	Yes
IMLIK	No	Yes

Figure 6.

Pepsin cleavage chart of TrypZean glycopeptide Ser_{70}-Lys_{89}. The distribution of modifications including glycosylation and oxidation on various peptic peptides are listed. All peptic peptides were identified by using both accurate mass measurements and sequence information from collisional activation.