Published in final edited form as:

Anal Methods. 2018 ; 10(2): 256–262. doi:10.1039/C7AY02687K.

Dissecting the Dissociation Patterns of Fucosylated Glycopeptides Undergoing CID: A Case Study in Improving Automated Glycopeptide Analysis Scoring Algorithms

Jude C. Lakbub¹, Xiaomeng Su¹, David Hua¹, Eden P. Go¹, and Heather Desaire^{1,*}

¹Ralph N Adams Institute for Bioanalytical Chemistry, Department of Chemistry, University of Kansas, Lawrence, Kansas-66047, United States

Abstract

The need to investigate the fragmentation of fucosylated glycopeptides is driven by recent work showing that at least one, and perhaps many, glycopeptide analysis scoring algorithms are less effective at identifying fucosylated glycopeptides than non-fucosylated glycopeptides. Herein, we study the CID fragmentation characteristics of fucosylated glycopeptides and the scoring rules of the glycopeptide analysis software, GlycoPep Grader, in an effort to improve automated assignments of these important glycopeptides. We identified some prominent product ions from a common fragmentation pathway of fucosylated glycopeptides that were not accounted for in the scoring rules. Based on this finding, we propose new scoring rules for fucosylated glycopeptides that can be incorporated into GlycoPep Grader and other similar analysis software tools to more accurately identify these species. The approach used here, to improve one particular scoring algorithm, could henceforth be used to improve any other algorithm that assigns glycopeptides based on their MS/MS data.

Introduction

Glycopeptide analysis is an important approach for protein glycosylation profiling because it provides glycosylation site-specific information.^{1,2} However, glycopeptide analysis using MS/MS data is challenging, even after the recent emergence of numerous automated tools that facilitate the process. Several bioinformatics tools have been developed to interpret glycopeptide MS/MS data generated by different fragmentation techniques, including electron transfer dissociation (ETD), higher-energy collisional dissociation (HCD), and collision induced dissociation (CID), the most common fragmentation technique used for glycopeptide analysis.^{3–5} Examples of tools that analyze CID glycopeptide data include MAGIC,⁶ GlycoPeptideSearch,⁷ and GlycoPep Grader,⁸ the latter being the focus of this study. Although these tools greatly enhance data analysis speed and reduce turn-around time for protein glycosylation profiling, automated glycopeptide analysis is only going to be as good as the scoring algorithms used to automatically assign the data.

*Corresponding author: Phone: 785-864-3015, Fax: 785-864-5396, hdesaire@ku.edu.

We have recently shown that testing software tools by challenging them with large sets of decoys is one way to probe for weaknesses in these scoring algorithms.⁹ Using this approach, we identified that GlycoPep Grader has some limitation in scoring CID spectra from fucosylated glycopeptides, and this weakness could be detrimental in estimating fucosylation in glycoproteins. Furthermore, it is likely that other glycopeptide analysis tools also have weaknesses in scoring fucosylated glycopeptides, since the scoring rules built into GlycoPep Grader were based on the field's common knowledge of how fucosylated glycopeptides dissociate.

Accurate determination of the fucosylation levels of glycoproteins is vital in understanding some critical biological activities of glycoproteins because the degree of fucosylation can affect the bioactivity of some glycoproteins. For example, an increase in antibody fucosylation can lead to decreased antibody dependent cellular cytotoxicity (ADCC).¹⁰⁻¹² Glycopeptides can be fucosylated at two main sites on the glycan components: (1) core fucosylation, where a fucose residue is appended to a GlcNAc residue on the pentasaccharide core of a glycan moiety; and (2) antenna fucosylation, where the fucose residue is attached to a GlcNAc residue on one of the glycan's antennae.¹³ While antenna fucosylation is rare, core fucosylation is very common, and it is the focus of this study. Although core- and antenna-fucosylated glycopeptide isomers can be distinguished using MS/MS data,¹³ a thorough analysis of the fragmentation patterns of fucosylated glycopeptides is still lacking in the literature, making it difficult to write accurate rules for scoring CID spectra from fucosylated glycopeptides.

The development of efficient automated glycopeptide scoring algorithms depends on the comprehensive analysis and identification of the characteristic fragmentation patterns of glycopeptides containing varying glycan components. Hence, writing accurate glycopeptide scoring algorithms is, perhaps, the most challenging part of software development due to the micro heterogeneity of *N*-linked glycans, since the scoring rules must be specific for each glycan type, and the scoring rules must be based on the predominant product ions of each glycopeptide type.⁸ Currently, very limited information is available on the fragmentation behavior of fucosylated glycopeptides because fucose is generally considered to be a very labile monosaccharide.¹⁴ This limitation in the understanding of the characteristic fragmentation behavior of fucosylated glycopeptides could translate into weaknesses in glycopeptide analysis tools. Hence, a thorough study of the characteristic behavior of fucosylated glycopeptides upon CID fragmentation can uncover prominent product ions that could be used for more accurate assignment of fucosylated glycopeptides to MS/MS data.

Herein, we make use of a recently-published, large dataset containing CID spectra of glycopeptides from well characterized glycoproteins⁹ to study the fragmentation characteristics of fucosylated glycopeptides. The overall goal of this case study is to develop a more precise and accurate understanding of the fragmentation characteristics of fucosylated glycopeptides, so scoring algorithms, such as GlycoPep Grader, can identify these species in CID datasets with higher accuracy. The results of our study show that although the fucose residue is labile, the core fucose does not always dissociate from the glycan during CID fragmentation. In addition to the fragmentation pathway where the fucose residue dissociates from the precursor glycopeptide ion, we identified an additional

fragmentation pathway, where the fucose residue remains attached to the glycan core. A study of the software algorithm, GlycoPep Grader, indicates that this additional fragmentation pathway is not considered when scoring fucosylated glycopeptides. Based on this finding, we propose new scoring rules for fucosylated glycopeptides that could be incorporated into GlycoPep Grader and any other algorithm that scores CID spectra of glycopeptides.

Experimental

The CID data used in this study are from a large dataset of well-characterized spectra that was recently shared online.⁹ The development of this dataset is described in Reference 9. That reference also contains the details regarding the materials and reagents, sample preparation, LC-MS/MS analysis, and generation of the CID dataset.

Results and Discussion

The motivation for this study is driven by a recent publication showing that the scoring algorithm in GlycoPep Grader identifies non-fucosylated glycopeptides with high accuracy, even when tested against 20 decoys per target, but it performs less admirably when identifying fucosylated glycopeptides, when tested against 20 decoys per target.⁹ In order to identify and correct the recently-described weakness in GlycoPep Grader's scoring function for fucosylated glycopeptides, it is important to understand the overall scoring rules of the software. The detailed scoring algorithm has been reported by Woodin et al,⁸ and a brief summary of the scoring algorithm is discussed herein.

GlycoPep Grader scores potential glycopeptide candidate compositions (including decoy glycopeptides) by searching predicted *in-silico* product ions of the glycopeptide compositions against the experimental CID data. The candidate that best matches the experimental spectrum (candidate with highest score) is assigned to the CID spectrum being studied. The software scores candidate glycopeptide compositions using two predominant types of product ions that have been shown to be present in CID spectra of glycopeptides: the peptide plus pentasaccharide core ions, referred to as [peptide + core component] product ions and product ions resulting from step-wise neutral losses of terminal monosaccharide residues from the precursor glycopeptide composition, referred to as [precursor – monosaccharide] product ions. The numbers of predicted peptide and glycan ions matched to the experimental spectra are used to calculate the *PeptideScore* and the *GlycanScore*, respectively, and both scores are used to calculate the overall score (*GlycopeptideScore*) of the glycopeptide candidate. The [peptide + core component] and [precursor – monosaccharide] product ions that are searched by the current version of GlycoPep Grader are shown in Figure 1.

To compute the *PeptideScore*, the software searches for six possible product ions at different charge states. These include: (1) the [naked peptide], (2) [peptide + HexNAc], (3) [peptide + 2HexNAc], (4) [peptide + 2HexNAc + Hex], (5) [peptide + 2HexNAc + 2Hex], and (6) [peptide + 2HexNAc + 3Hex]. For the *GlycanScore*, the [precursor – monosaccharide] ions searched by GlycoPep Grader depend on the type of glycan component of the glycopeptide

candidate. The glycan types are divided into eight categories: (1) Non-fucosylated high-mannose type glycans; (2) fucosylated high-mannose type glycans; (3) Non-fucosylated complex or hybrid type glycans containing sialic acid; (4) fucosylated complex or hybrid type glycans containing sialic acid; (5) fucosylated complex or hybrid type glycans that do not contain sialic acid and have more HexNAc than Hex residues; (6) fucosylated complex or hybrid type glycans that do not contain sialic acid and have more Hex than HexNAc residues; (7) complex or hybrid type glycan structures with a greater number of HexNAc than Hex residues and do not contain fucose or sialic acid residues; and (8) complex or hybrid type glycans that have more Hex than HexNAc residues and do not contain fucose or sialic acid residues.

We studied the CID fragmentation patterns and the scoring rules for high-mannose and non-fucosylated hybrid/complex glycopeptide compositions, which are accurately scored by GlycoPep Grader, and compared them to those of fucosylated glycopeptide compositions. The aim was to better understand why the scoring rules for the nonfucosylated glycopeptides seemed to be more accurate than the rules for the fucosylated glycopeptides.

Figure 2 shows example CID spectra that exhibit the characteristic fragmentation patterns of high-mannose and non-fucosylated hybrid/complex type glycopeptides. CID fragmentation of high-mannose type glycopeptides were characterized by neutral losses of Hex residues from the precursor glycopeptide ion (Figure 2A), while the fragmentation of non-fucosylated hybrid/complex type glycopeptides were characterized by neutral losses of Hex and HexNAc residues from the precursor ion (Figure 2B). No alternative fragmentation pathways were identified. Both spectra contain the [peptide + core component] (peaks with black star) and the [precursor – monosaccharide] product ions (peaks with red star), which are used by GlycoPep Grader to score glycopeptide compositions. The [peptide + core component] ions are present in both the charge state of the precursor glycopeptide and at lower charge states, while the [precursor – monosaccharide] product ions are mainly present at the charge state of the precursor ion. The CID spectra in Figure 2 are representative of the fragmentation patterns observed across a diverse group of high-mannose and non-fucosylated hybrid/complex type glycopeptides.

After studying the CID spectra and confirming the characteristic product ions, we investigated whether GlycoPep Grader's scoring algorithm is designed to search for the characteristic product ions in the CID spectra of high-mannose and non-fucosylated hybrid/complex type glycopeptides. When calculating the *PeptideScores* for both types of glycopeptide compositions, GlycoPep Grader searches for the six possible [peptide + core component] product ions at both the charge state of the precursor ion and at decremented charge states. For the *GlycanScores*, the software searches for the predominant product ions resulting from the neutral losses of Hex residues (for high-mannose glycopeptides) and Hex and HexNAc residues (for non-fucosylated glycopeptides) from the glycopeptide precursor ions. Overall, our study of GlycoPep Grader's scoring algorithm for high-mannose and non-fucosylated glycopeptides indicates that the rules are well written, since the software searches for the product ions that are typically present in the CID spectra from these types of glycopeptides. The well-written rules correlate with the high numbers of high-mannose and

non-fucosylated glycopeptides that out-scored decoy glycopeptides when GlycoPep Grader's scoring algorithm was assessed recently.⁹

We next studied the CID spectra and GlycoPep Grader's scoring rules for fucosylated glycopeptides to investigate if the same correlation exists between the characteristic product ions in spectra of fucosylated glycopeptides and their scoring rules. For a thorough study of the fragmentation patterns of core-fucosylated glycopeptides undergoing CID, two sets of model fucosylated glycopeptides from well-characterized proteins were used. These include fucosylated glycopeptides that contain sialic acid and fucosylated glycopeptides that do not contain sialic acid. After extensive analysis of CID data from these model glycopeptides, we identified two characteristic fragmentation pathways. One pathway results in product ions without appended fucose, while the other generates product ions containing the core fucose residue.

A representative CID spectrum of a fucosylated glycopeptide containing sialic acid that exhibits both fragmentation pathways is shown in Figure 3. Figure 3A shows the fragmentation pathway that occurs when the labile sialic acid and fucose residues dissociate from the glycan component. The predominant [precursor-monosaccharide] product ion (e.g. peak at m/z 1486.4) results from the neutral loss of sialic acid from the glycopeptide precursor ion (black oval) and subsequent neutral loss of the fucose residue, followed by neutral losses of Hex and HexNAc residues from the glycan's antennae that led to several [precursor – monosaccharide] product ions (peaks indicated with red star) and [peptide + core component] product ions (peaks indicated with black star) that do not contain the fucose residue. The most abundant ion that does not contain fucose residue is the Y_1 ion (e.g. peak at m/z 1364.8) at a charge state lower than that of the precursor ion.

Figure 3B shows the second fragmentation pathway, during which the sialic acid residue dissociates from the glycan component, but the core-fucose residue does not. The predominant [precursor-monosaccharide] product ion (peak at m/z 1486.4) also results from the loss of sialic acid from the glycopeptide precursor ion (black oval), and the subsequent step-wise neutral losses of monosaccharide residues from the glycan's antennae, without loss of fucose, leads to product ions that have the appended core fucose residue. The most abundant ion that contains the appended fucose residue is the Y_1 +Fuc ion (e.g. peak at m/z 1510.9) at a charge state lower than that of the precursor ion.

For both fragmentation pathways, the [precursor – monosaccharide] product ions were observed only at the charge state of the precursor glycopeptide while the [peptide + core component] product were observed at both the charge state of the precursor and the next lower charge state. The characteristic oxonium ion for sialylated glycopeptides at m/z 657^{15,16} was also observed, as well as other common oxonium ions (at m/z 528 and 690) of *N*-linked glycopeptides.

Analysis of CID data from fucosylated glycopeptides that do not contain sialic acid also reveals the same two fragmentation pathways described above. A representative example is shown in Figure 4. The characteristic [peptide + core component] and [precursor – monosaccharide] product ions from the two pathways were identified irrespective of the

charge state of the precursor glycopeptide and the relative numbers of Hex and HexNAc residues of the glycan component.

After identifying the two common CID fragmentation pathways for fucosylated glycopeptides, we studied GlycoPep Grader's scoring rules to determine whether the software searches for the characteristic product ions (ions with and without appended fucose) generated by the two pathways when scoring CID data from fucosylated glycopeptides. We noticed that GlycoPep Grader searches for product ions resulting from only one of the two possible fragmentation pathways, the product ions that do not contain appended fucose. Based on this finding, we recommend that the algorithm be redesigned to search for additional product ions (product ions that contain the appended core fucose residue) to increase the scores of fucosylated glycopeptides. The proposed ions are shown in Figure 5; the new product ions containing core-fucose are indicated by red asterisks. By comparison, the product ions that are searched for by the current version of GlycoPep Grader are shown in Figure 1. These ions do not include species that have the fucose remaining on the glycopeptide.

Finally, we note that in addition to the ions discussed above, all types of glycopeptides generate oxonium ions while undergoing CID. These ions are shown in orange in Figures 2–4 and include m/z 366, 528, 657, and 690. The presence of these ions have been described previously.^{15–18} Aside from ion, m/z 657, which is useful for identifying glycopeptides containing sialic acid,^{15,16} oxonium ions generally are not very useful in *discriminating* between one glycoform type and another; rather, they are mainly useful at determining whether or not a particular CID spectrum is from a glycopeptide. Hence, it is highly unlikely that the scores of fucosylated glycopeptides could be improved significantly by including additional oxonium ions into the scoring algorithm, particularly because none of the common oxonium ions contain fucose.

Conclusion

Recently, we showed that GlycoPep Grader has unprecedented strengths in scoring CID data from high-mannose and non-fucosylated hybrid/complex glycopeptides, but the software does not score spectra from fucosylated glycopeptides as effectively. After extensive studies of GlycoPep Grader's scoring algorithm and the fragmentation characteristics of CID data from high-mannose, non-fucosylated hybrid/complex, and fucosylated hybrid/complex glycopeptides, we identified an alternative fragmentation pathway for fucosylated glycopeptides that is not incorporated in GlycoPep Grader's scoring rules. The alternative fragmentation pathway leads to [Peptide + core component] and [precursor – monosaccharide] product ions that contain an appended fucose residue. The incorporation of scoring rules for product ions that contain appended fucose residues should significantly improve the success of correctly assigning fucosylated glycopeptides. This case study, in evaluating and improving the scoring rules for one particular software product, also could be replicated for other similar tools, leading to better software and, ultimately, better analytical capabilities in the glycoproteomics field.

Acknowledgments

The authors gratefully acknowledge the support of NIH grant R01GM103547 to HD.

References

1. Go EP, Herschhorn A, Gu C, Castillo-Menendez L, Zhang S, Mao Y, Chen H, Ding H, Wakefield JK, Hua D, Liao HX, Kappes JC, Sodroski J, Desaire HJ. *Viol.* 2015; 89:8245–8257.
2. Desaire H. *Molecular & Cellular Proteomics*. 2013; 12:893–901. [PubMed: 23389047]
3. Woodin CL, Maxon M, Desaire H. *Analyst*. 2013; 138:2793–2803. [PubMed: 23293784]
4. Hu H, Khatri K, Klein J, Leymarie N, Zaia J. *Glycoconj J*. 2016; 33:285–296. [PubMed: 26612686]
5. Hu H, Khatri K, Zaia J. *Mass Spectrom Rev*. 2017; 36:475–498. [PubMed: 26728195]
6. Lynn K-S, Chen C-C, Lih TM, Cheng C-W, Su W-C, Chang C-H, Cheng C-Y, Hsu W-L, Chen Y-J, Sung T-Y. *Anal Chem*. 2015; 87:2466–2473. [PubMed: 25629585]
7. Chandler KB, Pompach P, Goldman R, Edwards N. *J Proteome Res*. 2013; 12:3652–3666. [PubMed: 23829323]
8. Woodin CL, Hua D, Maxon M, Rebecchi KR, Go EP, Desaire H. *Anal Chem*. 2012; 84:4821–4829. [PubMed: 22540370]
9. Lakbub JC, Su XM, Zhu ZK, Patabandige MW, Hua D, Go EP, Desaire H. *J Proteome Res*. 2017; 16(8):3002–3008. [PubMed: 28691494]
10. Hodoniczky J, Zheng YZ, James DC. *Biotechnol Prog*. 2005; 21:1644–1652. [PubMed: 16321047]
11. Iida S, Misaka H, Inoue M, Shibata M, Nakano R, Yamane-Ohnuki N, Wakitani M, Yano K, Shitara K, Satoh M. *Clin Cancer Res*. 2006; 12:2879–2887. [PubMed: 16675584]
12. Jefferis R. *Nat Rev Drug Discov*. 2009; 8:226–234. [PubMed: 19247305]
13. Vékey K, Ozohanics O, Tóth E, Jek A, Révész Á, Krenyácz J, Drahos L. *Int J Mass Spectrom*. 2013; 345-347:71–79.
14. Tajiri M, Kadoya M, Wada Y. *J Proteome Res*. 2009; 8:688–693. [PubMed: 19099505]
15. Nwosu CC, Seipert RR, Strum JS, Hua SS, An HJ, Zivkovic AM, German BJ, Lebrilla CB. *J Proteome Res*. 2011; 10:2612–2624. [PubMed: 21469647]
16. Song E, Pyreddy S, Mechref Y. *Rapid Comm Mass Spectrom*. 2012; 26:1941–1954.
17. Conboy JJ, Henion JD. *J Am Soc Mass Spectrom*. 1992; 3:804–814. [PubMed: 24234703]
18. Itoh, S., Takakura, D., Kawasaki, N., Yamaguchi, T. *The Protein Protocols Handbook*. Walker, JM., editor. Humana Press; Totowa, NJ: 2009. p. 1419-1430.

Figure 1.

Representation of the predominant product ions searched by GlycoPep Grader when scoring *N*-linked glycopeptide compositions against CID spectra. Six possible [peptide + core component] ions (A) are searched for as well as several [precursor-monomer] ions (B) for each of the eight groups (see text) of glycopeptides. The monosaccharide neutral losses evaluated for group 1 (see text) are shown in the purple oval. For group 2, the relevant losses are shown in both the purple oval and the yellow circle. For group 3, the relevant losses are shown in the blue oval. For group 4, they are shown in the blue oval and yellow circle. For group 5, they are shown in the orange oval and yellow circle. For group 6, they are shown in the green oval and yellow circle. For groups 7 and 8, neutral losses are in the orange oval at bottom left, and by the green oval at bottom right, respectively. This figure is reprinted with permission from Reference 8.

Figure 2.

CID spectra of a high-mannose glycopeptide from RNase B (**A**) and a non-fucosylated hybrid type glycopeptide from HIV-1 envelope glycoprotein, C.97ZA012 gp140 (**B**). The spectra show the characteristic fragmentation patterns of high-mannose and non-fucosylated H/C type glycopeptides. High-mannose spectra (**A**) are characterized by neutral losses of Hex residues from the precursor ion (black oval) and spectra of non-fucosylated H/C type glycopeptides (**B**) are characterized by neutral losses of Hex and HexNAc residues from the precursor ion (black oval). Both spectra contain [peptide + core component] ions (black stars) and [precursor – monosaccharide] product ions (red stars), which are used by GlycoPep Grader to score *N*-linked glycopeptides. Different oxonium ions (orange stars) are also observed, and no alternative fragmentation pathway was identified.

(A) Fragmentation pathway 1: Fucose dissociates from glycan core

(B) Fragmentation pathway 2: Fucose remains attached to glycan core

Figure 3.

Representative CID spectra of a fucosylated complex type glycopeptide containing sialic acid showing two fragmentation pathways observed for fucosylated glycopeptides. **(A)** shows the fragmentation path that leads to [precursor – monosaccharide] product ions (red star) and [peptide + core component] product ions (black star) that do not contain the core fucose. **(B)** shows the fragmentation pathway that generates [peptide + core component] product ions (black star) and [precursor – monosaccharide] product ions (red star) that contain the appended core fucose. Oxonium ions (orange star) are also observed. Ions generated by each fragmentation pathway are annotated and shown in black while non-related peaks are shown in gray. Precursor ions are in black ovals.

(A) Fragmentation pathway 1: Fucose dissociates from glycan core

(B) Fragmentation pathway 2: Fucose remains attached to glycan core

Figure 4.

Representative CID spectra of a fucosylated complex type glycopeptide that does not contain sialic acid showing two fragmentation pathways observed for fucosylated glycopeptides. **(A)** The fragmentation path that leads to [precursor – monosaccharide] product ions (red star) and [peptide + core component] product ions (black star) that do not contain the core fucose. **(B)** The fragmentation pathway that generates [peptide + core component] product ions (black star) and [precursor – monosaccharide] product ions (red star) that contain the appended core fucose. Oxonium ions (orange star) are also observed. Ions generated by each fragmentation pathway are annotated and shown in black while non-related peaks are shown in gray. Precursor ions are in black ovals.

Product ions proposed for future GlycoPep Grader scoring

Figure 5.

Schematic representation of the proposed product ions for future GlycoPep Grader scoring. The proposed product ions include ions containing appended fucose (with red asterisks) for the scoring of fucosylated glycopeptides.