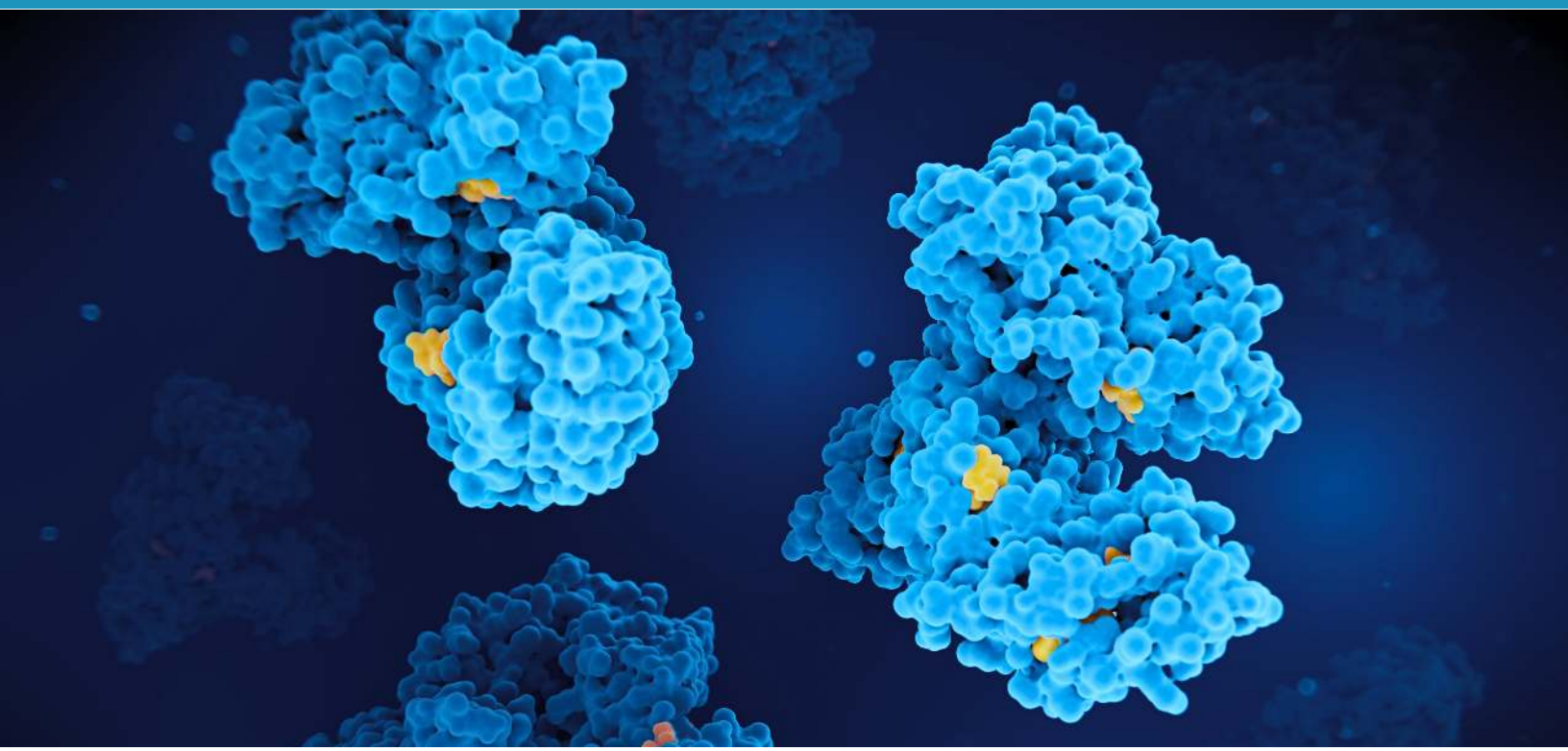
Bioinformatics Solutions Inc.



# N-glycan and O-glycan Profiling of Fetuin by single LC-MS/MS run

Yiming Xiao, PhD., Application Scientist
Kyle Hoffman, PhD., Applications Manager

Bioinformatics Solutions Inc., Waterloo, Canada

## Abstract:

Glycoproteins have gained significant interest due to their pivotal roles in numerous biological processes and their implications for health and disease.[1] Profiling and localization has become essential for unravelling biological processes, paving the way for developing targeted therapies and diagnostic tools in diverse fields ranging from medicine to biotechnology. In this study, we present a straightforward glycosite analysis workflow developed in the BSI Service Lab. Coupled with our innovative software platform, PEAKS GlycanFinder, users can rapidly perform comprehensive glycan profiling.

## Introduction:

Fetuin-A (pp63) is a glycoprotein that is abundant in mammalian blood plasma. It is acknowledged as a versatile protein involved in numerous critical biological functions, including the modulation of bone and calcium metabolism, as well as the regulation of the insulin signaling pathway[2]. Additionally, owing to Fetuin-A's distinctive feature of harboring both N-link and O-link glycosylation and its ease of purification [3], it is frequently employed as a standard protein for testing and developing analytical methods aimed at profiling N/O-glycans and pinpointing glycosylation sites. In the study, we utilize our LC-MS/MS developed in our lab coupled with our recently developed software platform, PEAKS GlycanFinder, to analyze the N/O glycan profile of Fetuin from a single LC-MS/MS run (Figure 1) [4].

## Methods:

Bovine fetuin (F3004) was purchased from Sigma-Aldrich. The sample was reduced and alkylated then digested with trypsin at 37ºC overnight. The digested sample was desalted with home-made C18 tip, and the product was separated using a Ultimate3000 UHPLC then analyzed on a Fusion Lumos (Thermo) mass spectrometry with EtHCD mode. All raw data was analyzed using PEAKS GlycanFinder with the following parameters. Precursor, glycan, and peptide fragment ion mass errors were set to 10 PPM, 20 PPM, and 0.02 Da, respectively. Carbamidomethylation was set as a fixed PTM and oxidation (M) was set as a variable PTM. The bovine Fetuin Uniprot sequence entry was used as a database for protein/peptide identification, and N&O glycan database built into PEAKS GlycanFinder were used (N-linked: 1867 entries; O-linked: 265 entries). Only glycopeptides passing a score threshold of 15 were considered.

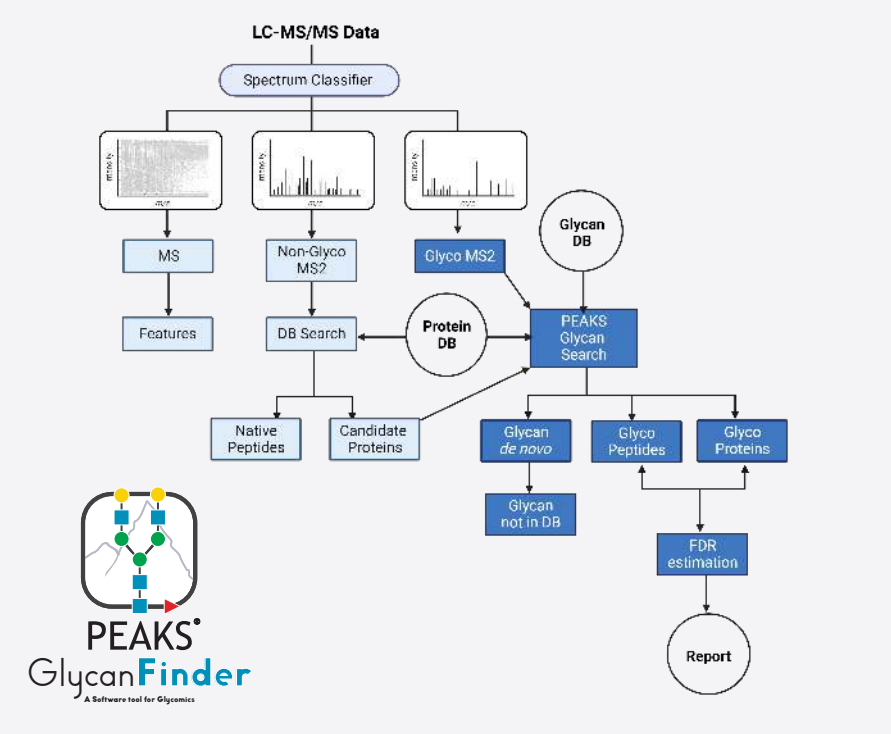



**Fig 1. PEAKS GlycanFinder Workflow**

| | Glycosites | Glycopeptides | Quantifiable Glycopeptides |
|---|---|---|---|
| N-Glycan | N99, N176, N156 | 52 | 33 |
| O-Glycan | 61 | 27 | 18 |

**Table 1. Glycan sites and number and glycopeptides identified in Fetuin.**
**Note: only confident sites are reported here.**

## Results and Discussion:

We identified 79 glycopeptides with 8 high confidence glycosylated sites (Table 1.), and 64% of identified glycopeptides have least one quantifiable feature. The protein coverage by glycopeptides is shown in Figure 2.

The coverage figure generated by GlycanFinder offers a comprehensive overview of glycopeptides and their corresponding glycosylation sites. An ambiguity score, denoted as Ascore, was utilized to assess the confidence level of identified glycosylation sites. In this investigation, a cutoff value of Ascore 20 was established, resulting in the identification of 7 highly confident glycosylation sites. Among these, all three N-glycosylation sites (N 99/156/176) were extensively documented [5][6][7].


**Fig 2. Glycopeptide mapping of Fetuin.** Green bars represent glycopeptides and colored squares on each peptide represent different glycans and carbamidomethylation (red). Confident glycosylation site (localization AScore >= 20) was marked at the top of the protein sequence.

## Results and Discussion cont'd:

Regarding high-confidence O-glycosylation sites, in addition to the well-known S296 and S341, one novel site, T37, was discovered and supported by 5 glycopeptides (see example in Figure 3). In addition to these high-confidence glycosylation sites, with manual inspection, S282 site can also be confirmed by internal ion matching.

Using PEAKS GlycanFinder simplifies the process of profiling glycan compositions at specific glycosylation sites. In this instance, we focus on two N-glycosylation sites, N99 and N156. Out of 34 and 9 identified glycan moieties on these sites, 24 and 8 are quantifiable, respectively. The corresponding intensities are depicted in Figure 3, aligning with the sialic acid-rich characteristic of feutin.

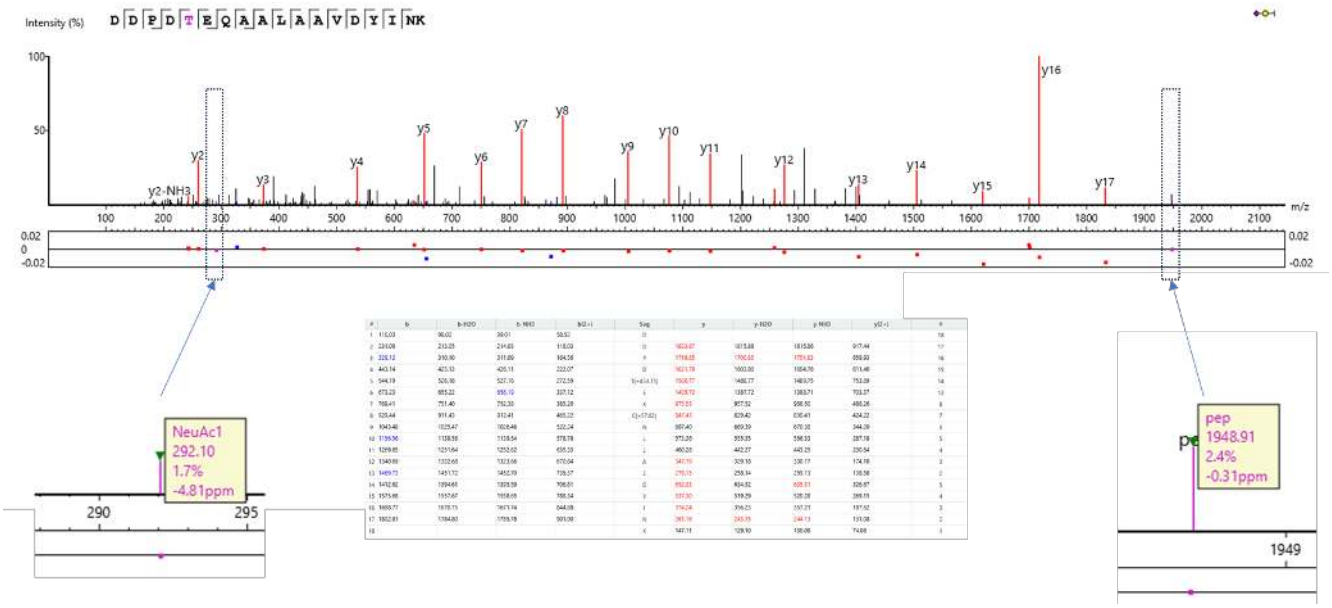



**Fig 3. Spectrum of T37 glycopeptide, DDPDTEQAALAAVDYINK.** O-glycan (Hex)1(NeuAc)1 was identified. Top: Spectrum with fragmentation annotated. Bottom middle: ion matching table, note y14 matched with the glycan fragment ion (Y2). Bottom left: zoomed-out view of NeuAc ion (B1). Down right: zoomed-out view peptide precursor.

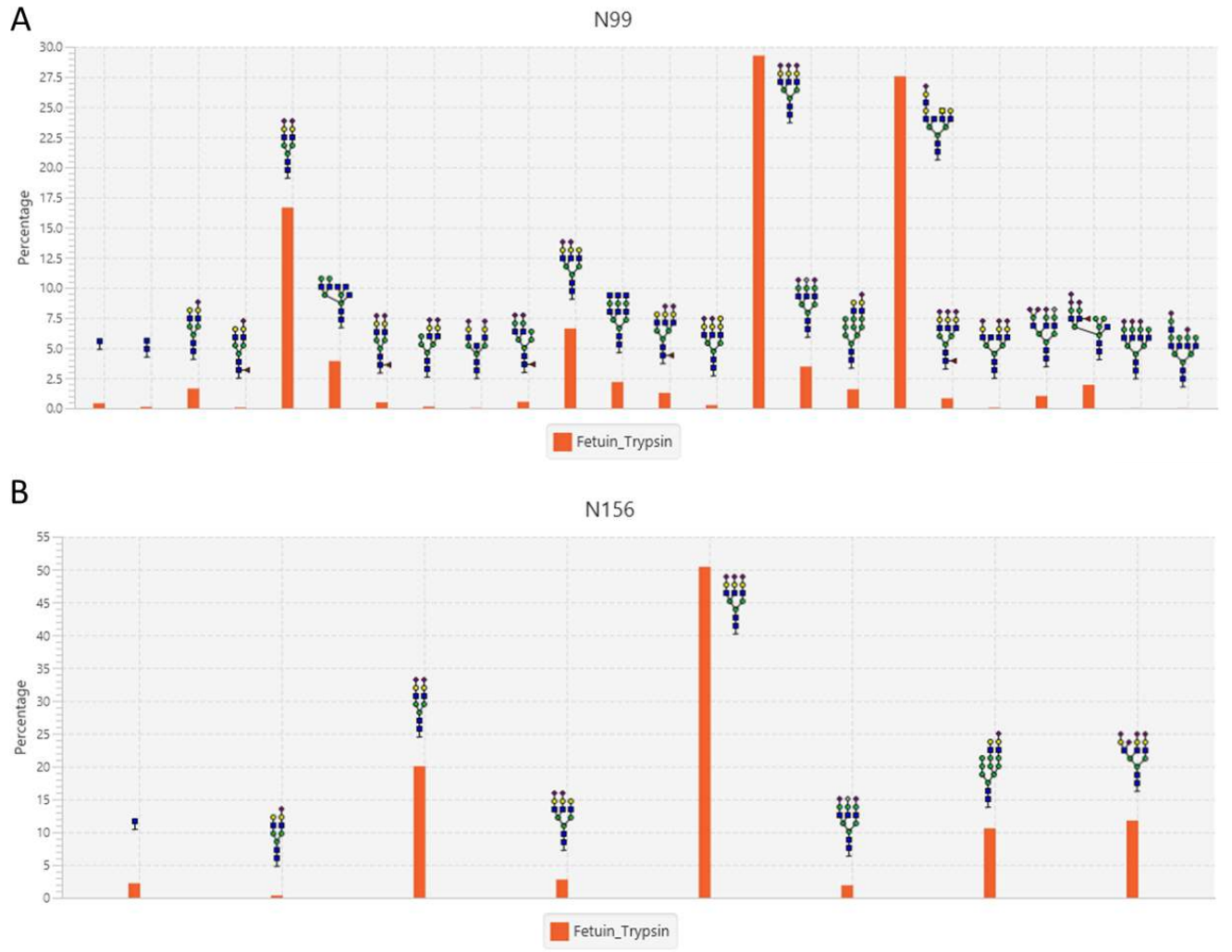



**Fig 4. Glycan profiles of Fetuin at N99 (A) and N156 (B). Relative abundance of each glycoform is shown as bar chart.**

## Results and Discussion cont'd:

Comparison of the relative abundance of 6 glycoforms at N99 and N156, as reported in literature [5][7][8], is presented in Figure 5, revealing a consistent outcome. Noteworthy is the slightly lower relative abundance of (HexNAc)5(Hex)6(NeuAc)3 at N99, quantified by PEAKS GlycanFinder. The difference is ascribed to PEAKS GlycanFinder's ability to identify a greater number of measurable glycoforms in comparison to existing literature. This expanded pool of quantifiable glycopeptides results in a lower relative abundance of the specific glycoform. The pie chart in Figure 6 provides a direct visualization of the phenomenon, showcasing the distribution of glycoforms on the specific residue. Analysis reveals that among the top 5 most prevalent glycoforms identified by Glycanfinder at site N99, two have limited representation in existing literature.

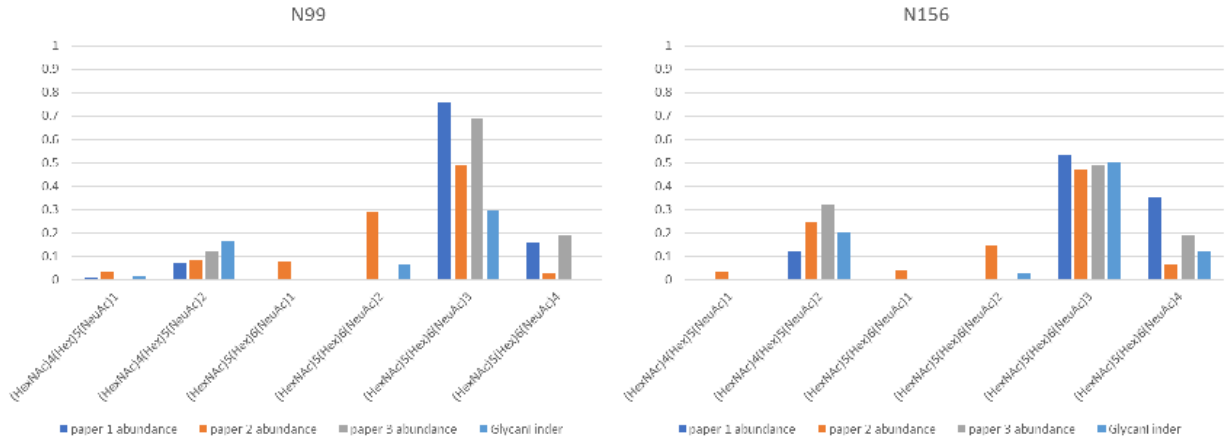
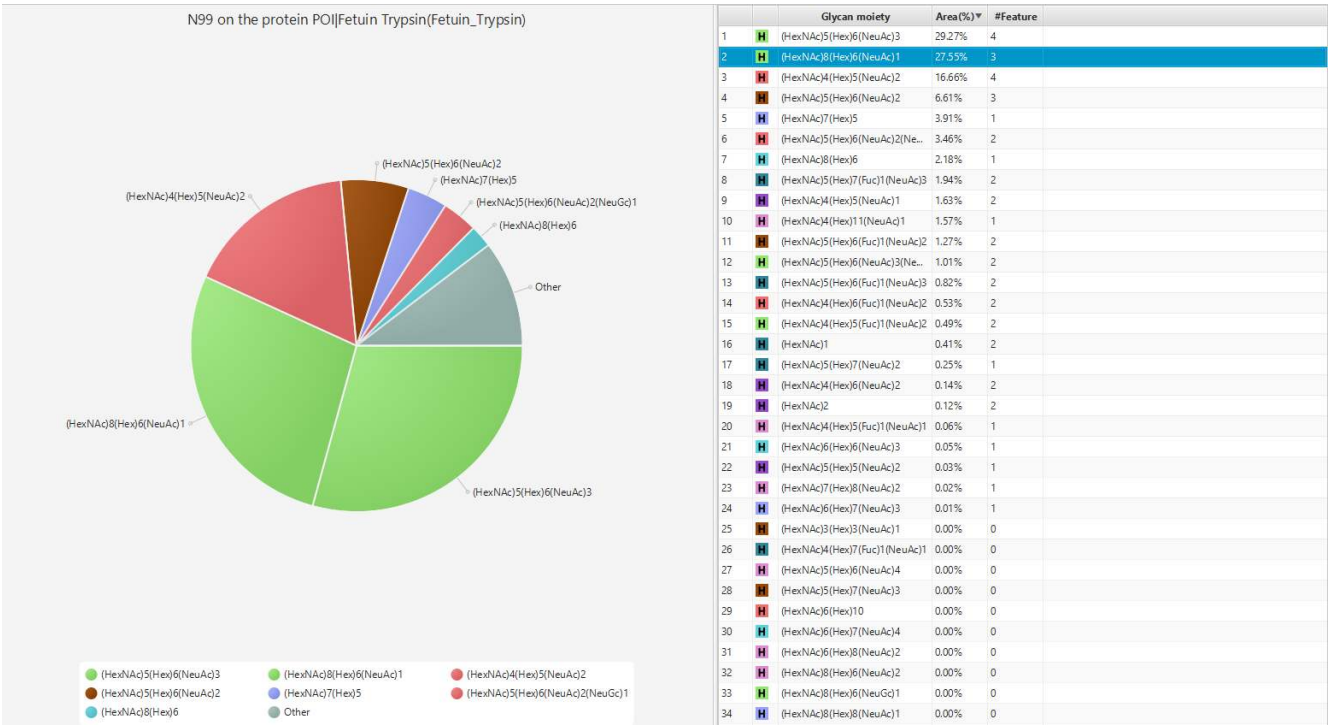



**Fig 5. Comparison of glycoform relative abundance between literature and PEAKS GlycanFinder.** Data was sourced from [5][8][7], identified as papers 1, 2, and 3, respectively.




| | | Glycan moiety | Area(%) | #Feature |
|---|---|---|---|---|
| 1 | H | (HexNAc)5(Hex)6(NeuAc)3 | 29.27% | 4 |
| 2 | H | (HexNAc)8(Hex)6(NeuAc)1 | 27.55% | 3 |
| 3 | H | (HexNAc)4(Hex)5(NeuAc)2 | 16.66% | 4 |
| 4 | H | (HexNAc)5(Hex)6(NeuAc)2 | 6.61% | 3 |
| 5 | H | (HexNAc)7(Hex)5 | 3.91% | 1 |
| 6 | H | (HexNAc)5(Hex)6(NeuAc)2(Ne... | 3.46% | 2 |
| 7 | H | (HexNAc)8(Hex)6 | 2.18% | 1 |
| 8 | H | (HexNAc)5(Hex)7(Fuc)1(NeuAc)3 | 1.94% | 2 |
| 9 | H | (HexNAc)4(Hex)5(NeuAc)1 | 1.63% | 2 |
| 10 | H | (HexNAc)4(Hex)11(NeuAc)1 | 1.57% | 1 |
| 11 | H | (HexNAc)5(Hex)6(Fuc)1(NeuAc)2 | 1.27% | 2 |
| 12 | H | (HexNAc)5(Hex)6(NeuAc)3(Ne... | 1.01% | 2 |
| 13 | H | (HexNAc)5(Hex)6(Fuc)1(NeuAc)3 | 0.82% | 2 |
| 14 | H | (HexNAc)4(Hex)6(Fuc)1(NeuAc)2 | 0.53% | 2 |
| 15 | H | (HexNAc)4(Hex)5(Fuc)1(NeuAc)2 | 0.49% | 2 |
| 16 | H | (HexNAc)1 | 0.41% | 2 |
| 17 | H | (HexNAc)5(Hex)7(NeuAc)2 | 0.25% | 1 |
| 18 | H | (HexNAc)4(Hex)6(NeuAc)2 | 0.14% | 2 |
| 19 | H | (HexNAc)2 | 0.12% | 2 |
| 20 | H | (HexNAc)4(Hex)5(Fuc)1(NeuAc)1 | 0.06% | 1 |
| 21 | H | (HexNAc)6(Hex)6(NeuAc)3 | 0.05% | 1 |
| 22 | H | (HexNAc)5(Hex)5(NeuAc)2 | 0.03% | 1 |
| 23 | H | (HexNAc)7(Hex)8(NeuAc)2 | 0.02% | 1 |
| 24 | H | (HexNAc)6(Hex)7(NeuAc)3 | 0.01% | 1 |
| 25 | H | (HexNAc)3(Hex)3(NeuAc)1 | 0.00% | 0 |
| 26 | H | (HexNAc)4(Hex)7(Fuc)1(NeuAc)1 | 0.00% | 0 |
| 27 | H | (HexNAc)5(Hex)6(NeuAc)4 | 0.00% | 0 |
| 28 | H | (HexNAc)5(Hex)7(NeuAc)3 | 0.00% | 0 |
| 29 | H | (HexNAc)6(Hex)10 | 0.00% | 0 |
| 30 | H | (HexNAc)6(Hex)7(NeuAc)4 | 0.00% | 0 |
| 31 | H | (HexNAc)6(Hex)8(NeuAc)2 | 0.00% | 0 |
| 32 | H | (HexNAc)8(Hex)6(NeuAc)2 | 0.00% | 0 |
| 33 | H | (HexNAc)8(Hex)6(NeuGc)1 | 0.00% | 0 |
| 34 | H | (HexNAc)8(Hex)8(NeuAc)1 | 0.00% | 0 |

**Fig 6. Pie Chart for glycoform distribution on N99.** Left: Pie chart of relative glycan abundances. Glycoforms with less than 2% area are included in "other". Right: All identified glycan moieties.

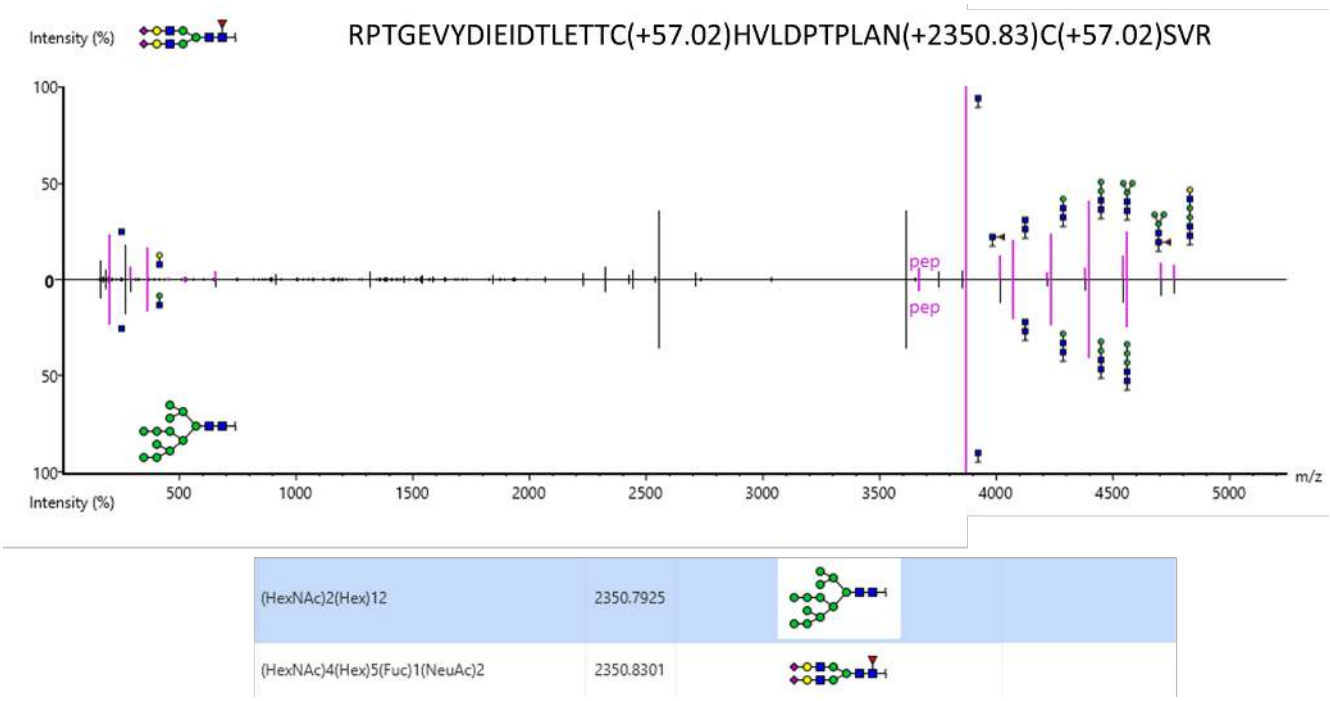



**Fig 7. Mirror plot of glycopeptide RPTGEVYDIEIDTLETTCHVLDPTPLAN(+2350.83)CSVR with two potential glycoform annotations.** The structures and mass of the two glycoforms are listed at the bottom.

## Results and Discussion cont'd:

A significant challenge in identifying glycoforms arises from the close mass proximity of two glycoforms, particularly with large glycan moieties. To address this issue, PEAKS GlycanFinder introduces an S-score to distinguish the best glycan from the second-best. Additionally, a mirror plot is provided for a direct comparison of fragment annotations, as illustrated in Figure 7. In the example presented, glycopeptide RPTGEVYDIEIDTLETTCHVLDPTPLAN(+2350.83)CSVR can potentially be assigned two glycoforms: (HexNAc)4(Hex)5(Fuc)1(NeuAc)2 and (HexNAc)2(Hex)12. These glycoforms exhibit a mere 0.04 Da mass difference, making it nearly impossible to differentiate them based solely on precursor m/z values. However, the mirror plot in Figure 6 reveals a greater number of Y ions associated with (HexNAc)4(Hex)5(Fuc)1(NeuAc)2, indicating its higher confidence as the preferred glycan.

Besides high confidence glycopeptide identifications, PEAKS GlycanFinder also provides additional information for glycan moieties not found in the glycan database that was searched against. This information is summarized into the Denovo tab in three categories: Glycan search, Glycan denovo and Open glycan. In short, Glycan search entails the identification of glycans through a standard peptide-first search followed by a glycan database search. When a spectrum does not meet the criteria for glycopeptide identification, an Open glycan search is performed. This refers to peptide-spectrum matches (PSMs) in which the glycan is searched using a glycan-first search strategy against the glycan database. Conversely, Glycan denovo is performed when glycopeptide spectra cannot be identified from the provided databases. Then a search for the Y1 ion is initiated, followed by glycan de novo sequencing on the spectrum. Within this project, 596 PSMs are categorized under the OpenGlycan section in the denovo tab. An example is presented in Figure 8, where a high-quality glycan spectrum is discernible in the top panel. However, despite peptide fragmentation in the low mass range (lower panel, zoomed in at the low mass range), no peptide match can be identified. This outcome aligns with expectations, given that only one provided sequence (feutin) was used in the search, and it is possible that other low-abundance glycoproteins coexist in the sample. This underscores the capability of the denovo tab in PEAKS GlycanFinder to facilitate glycopeptide/glycan identification even when an insufficient protein/glycan database is used.

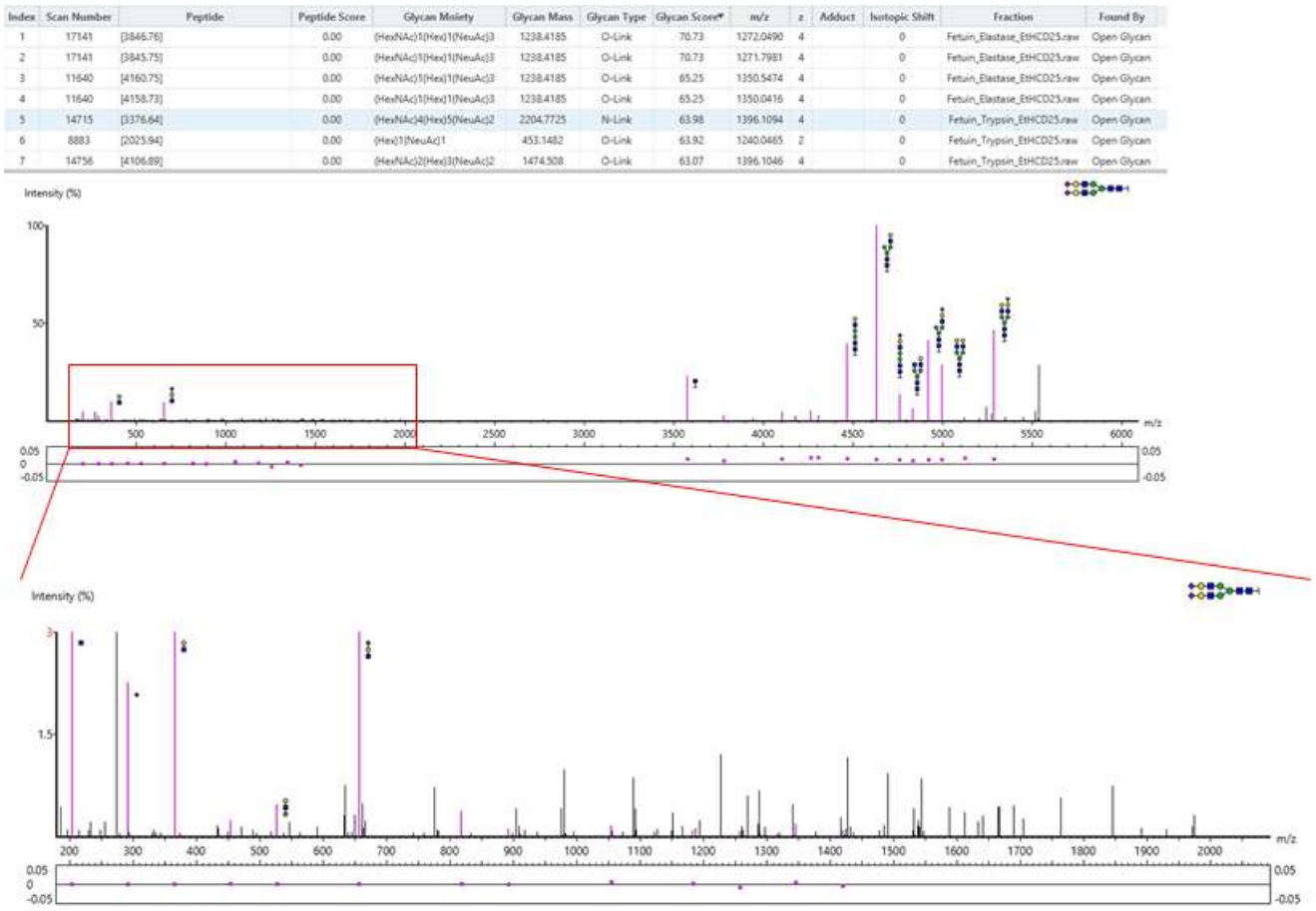



**Fig 8. Glycopeptide identification under de novo tab.** Top: Glycan ion match annotation; bottom: zoom in on the low mass range.

## Summary

PEAKS GlycanFinder offers a consolidated work platform that significantly decreases workload and enhances the efficiency of data analysis in glycoproteomics.

## References:


1. Bagdonaite, I. et al. Glycoproteomics. Nat Rev Methods Primers 2, 48 (2022).
2. Chekol Abebe, E. et al. The structure, biosynthesis, and biological roles of fetuin-A: A review. Frontiers in Cell and Developmental Biology 10, (2022).
3. Cartellieri, S., Hamer, O., Helmholz, H. & Niemeyer, B. One-step affinity purification of fetuin from fetal bovine serum. Biotechnology and Applied Biochemistry 35, 83–89 (2002).
4. Sun, W. et al. Glycopeptide database search and de novo sequencing with PEAKS GlycanFinder enable highly sensitive glycoproteomics. Nat Commun 14, 4046 (2023).
5. Windwarder, M. & Altmann, F. Site-specific analysis of the O-glycosylation of bovine fetuin by electron-transfer dissociation mass spectrometry. J Proteomics 108, 258–268 (2014).
6. Nwosu, C. C. et al. Simultaneous and Extensive Site-specific N- and O-Glycosylation Analysis in Protein Mixtures. J. Proteome Res. 10, 2612–2624 (2011).
7. Lin, Y.-H., Franc, V. & Heck, A. J. R. Similar Albeit Not the Same: In-Depth Analysis of Proteoforms of Human Serum, Bovine Serum, and Recombinant Human Fetuin. J Proteome Res 17, 2861–2869 (2018).
8. Thaysen-Andersen, M., Mysling, S. & Højrup, P. Site-Specific Glycoprofiling of N-Linked Glycopeptides Using MALDI-TOF MS: Strong Correlation between Signal Strength and Glycoform Quantities. Anal. Chem. 81, 3933–3943 (2009).

**Bioinformatics Solutions, Inc.**

140 Columbia St, Suite 202
Waterloo, Ontario N2L 3K8
Canada

Tel: (519) 885-8288
Fax: (519) 885-9075

sales@bioinfor.com
www.bioinfor.com