# A Fully Automated Workflow for Glycopeptide Analysis

P1609

Julian Saba[1], Rosa Viner[1],Paul Shan[2], Lei Xin[2], Sergei I. Snovida[3], Edward D. Bodnar[4], Kay-Hooi Khoo[3], and Helene Perreault[4]

[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]Bioinformatics Solutions Inc., Waterloo, ON, Canada; [3]Institute of Biological Chemistry, Academia Sinica, Taiwan; [4]University of Manitoba, Winnipeg, MB, Canada

## Overview

**Purpose:** Development of an automated workflow for comprehensive site-specific glycan/glycopeptide analysis in human serum.

**Methods:** Proteins from human serum were enriched at the glycoprotein and glycopeptide level using different strategies. Enriched glycopeptides were labeled with isobaric mass tags and analyzed on a hybrid ion trap-Oribtrap™ mass spectrometer using combination of higher-energy collisional dissociation (HCD) and electron transfer dissociation (ETD) fragmentation.

**Results:** The described automated workflow significantly improved the rate of success of ETD analysis and simplified the overall glycoproteomics workflow.

## Introduction

Glycosylation is a post-translational modification (PTM) that plays crucial roles in biochemical processes. Structural characterization of glycoproteins and glycopeptides is analytically challenging. Successful application of mass spectrometry (MS) in glycoproteomics depends on adopting a workflows that addresses specific questions relating to a particular sample type. Targeted enrichment of glyco-peptides is one such procedure. In principle, it reduces the complexity of the overall sample matrix, facilitating more sensitive and accurate analysis of the glycopeptides. Others have explored the separate use of $TiO_2$ or graphite as a means to selectively enrich glycopeptides for MS analysis. Here we report on the selective enrichment and characterization of glycopeptides based on combining $TiO_2$ and graphite. This approach was compared to common used zwitterionic Hydrophilic Interaction Liquid Chromato-graphy (ZIC-HILIC)-based strategies. Glycopeptide analysis was performed with a novel acquisition strategy termed high-energy collisional dissociation-accurate mass product ion-dependent electron transfer dissociation (HCD-PD-ETD) (Figure 1). The advantage of this approach is that it streamlined data analysis, improved dynamic range and duty cycle. Additionally, we utilized a novel bioinformatics tool to automate the analysis of the data from the combined fragmentation techniques (Figure 2).

## Methods

### Sample Preparation

Glycoproteins from human serum were isolated using a Thermo Scientific Glycoprotein Isolation Kit Con A per manufacturer's suggestion. Isolated glycoproteins were reduced, alkylated and subjected to enzymatic digestion. Samples were split into multiple fractions and enriched on ZIC®-HILIC (EMD Chemicals Inc.), Thermo Scientific Pierce $TiO_2$ Phosphopeptide Enrichment Kit and/or graphite, or SAX respectively according to the manufacturer's protocol. Upon enrichment, the samples were split into two fractions from each enrichment strategy. For each enrichment strategy, one fraction was labeled with Thermo Fisher Pierce Tandem Mass Tags ($TMT^0$-126™) according to the manufacturer's protocol, while the other fraction was subjected to PNGase F digestion in the presence of $H_2O^{18}$.

### LC/MS

A Thermo Scientific EASY-nLC nano-HPLC system and Michrom Magic™ C18 spray tip 15 cm x 75 µm i.d. column (Michrom BioResources) were used. Gradient elution was performed from 5-45% ACN in 0.1% formic acid over 60 min at a flow rate of 300 nL/min. The samples were analyzed with a Thermo Scientific LTQ Orbitrap Velos hybrid mass spectrometer with ETD. The following MS and MS/MS settings were used: FT: $MS^n$ AGCTarget = 5e4; MS/MS = 1 µscans, 200 ms max ion time; MS = 400-2000 *m/z*, 60000 resolution at *m/z* 400, MS Target = 1e6; MS/MS = Top 10 Data-Dependent™ acquisition HCD Product Dependent acquisition ion trap ETD (Figure 2), Dynamic Exclusion = repeat count 1, Duration 30 sec, Exclusion duration 90 sec; HCD Parameters: Collision Energy = 35%; resolution 7500. $MS^n$ Target Ion Trap = 1e4, 3 µscans, ETD anion AGC target = 2e5, charge dependent ETD reaction time was used. The Thermo Scientific Proteome Discoverer software version 1.2 was used to generate database of glycoproteins and prototype GlycoMaster™ (Bioinformatics Solution) software was used for intact glycopeptides analysis (Figures 2 and 3).

## Results

Studies have shown efficient enrichment of neutral and sialylated *N*-linked glycopeptides by ZIC-HILIC[1]; while $TiO_2$ shows specific binding to sialylated glycopeptides[2]. Graphite columns have shown affinity for glycopeptides with a smaller peptide backbone.[3] In our experiments, $TiO_2$ and graphite were both used alone and combined for a two-step enrichment approach to ensure all sizes of glycopeptides would be enriched. These were compared against the most commonly used ZIC-HILIC based strategy. To test the enrichment methods, a tryptic digest of human serum was examined. First, enriched digests were incubated with the enzyme PNGase F in the presence of $H_2O^{18}$ to obtain de-glycosylated peptides. Samples were then analyzed using conventional shotgun proteomics to generate a database of glycoproteins in Proteome Discoverer™ software (Figure 3, right column). This database was then brought into GlycoMaster to analyze HCD-PD-ETD data of intact glycopeptides (Figure 2). GlycoMaster software was used to extract the information of glycan fragmentation from the HCD spectrum and peptide fragmentation of the peptide backbone from the ETD spectrum to produce results on the overall glycopeptide structure (Figure 3, left column). Figure 4 summarizes the results from the two different enrichment strategies for human serum. Both approaches show high numbers of unique glycopeptides and glycoproteins with an overlap of only 35%. Though ZIC-HILIC has shown to provide efficient enrichment of sialylated glycopeptides, it also binds with neutral glycopeptides. $TiO_2$ ,on the other hand, selectively binds to sialylated glycopeptides over

**FIGURE 1. Schematic representation of HCD accurate-mass product-dependent ETD (HCD-PD-ETD) acquisition method.**

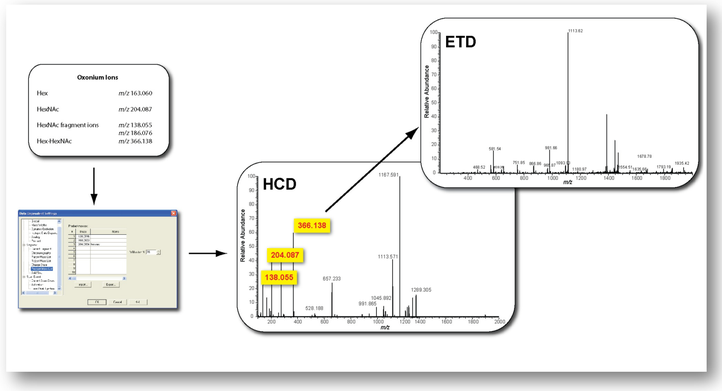



**FIGURE 2. User interface of GlycoMaster software.**

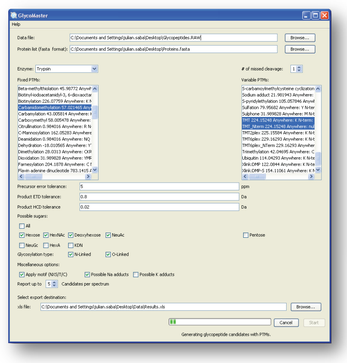



**FIGURE 3. Glycopeptide identification workflow.**

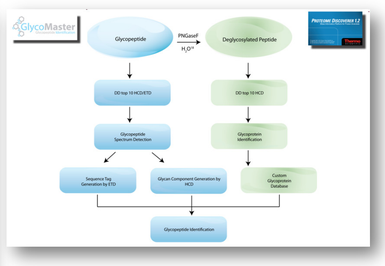
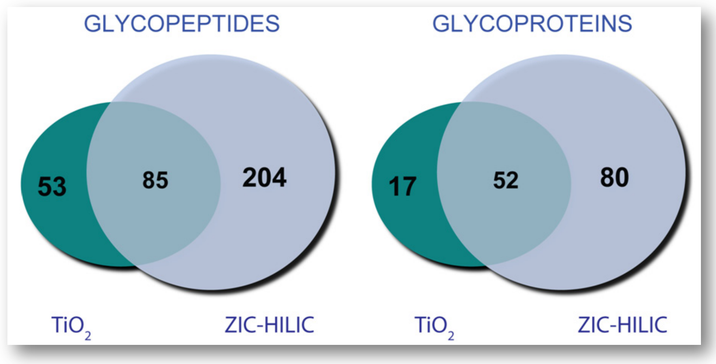



**FIGURE 4. Comparison of ZIC-HILIC *vs* $TiO_2$ for enrichment of human serum glycopeptides.**



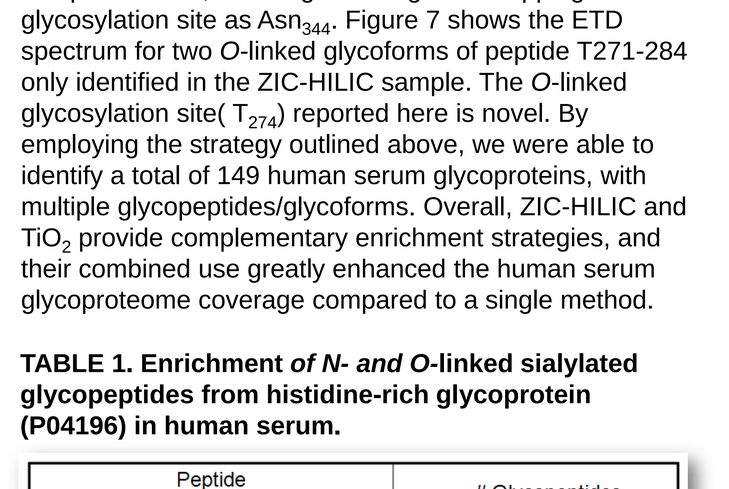

neutral glycopeptides, thus providing much more efficient enrichment of these acidic species. Our results confirm this observation as the unique glycopeptides present in the Venn diagram for $TiO_2$ were *N*-linked and acidic in nature, while the unique glycopeptides in ZIC-HILIC were mainly *O*-linked and acidic or *N*-linked and neutral.

Table 1 summarizes results for a histidine-rich glycoprotein identified from human serum. This protein contains both *N*- and *O*-linked glycosylation sites. Table 1 and Figures 5-7 show examples of the protein glycopeptides unique to each enrichment. Figure 6 displays the HCD spectrum for a bi-antennary histidine-rich glycoprotein *N*-linked glycopeptide T340-353 corresponding to *m/z* 958.383 in Figure 5. The corresponding ion trap ETD spectrum is shown in Figure 6. The ETD spectrum at precursor charge +4 generates complete z-ions, enabling unambiguous mapping of the *N*-glycosylation site as $Asn_{344}$. Figure 7 shows the ETD spectrum for two *O*-linked glycoforms of peptide T271-284 only identified in the ZIC-HILIC sample. The *O*-linked glycosylation site( $T_{274}$) reported here is novel. By employing the strategy outlined above, we were able to identify a total of 149 human serum glycoproteins, with multiple glycopeptides/glycoforms. Overall, ZIC-HILIC and $TiO_2$ provide complementary enrichment strategies, and their combined use greatly enhanced the human serum glycoproteome coverage compared to a single method.

**TABLE 1. Enrichment *of N- and O*-linked sialylated glycopeptides from histidine-rich glycoprotein (P04196) in human serum.**

| Peptide | # Glycopeptides | |
|---|---|---|
| | ZIC-HILIC | $TiO_2$ |
| HSHN**N**NSDLHPHK | 4 | 10 |
| SST**T**KPPFKPHGSR | 3 | 3 |
| VE**N**TTVYYLVLDVQESDCSVLSRK | 2 | 0 |
| VIDF**N**CTTSSVSSALANTK | 2 | 0 |

**FIGURE 5. Enrichment efficiency of ZIC-HILIC *vs* $TiO_2$ for sialylated *N*-linked glycopeptides T340-353 from histidine-rich glycoprotein in human serum.**




A novel acquisition strategy called HCD-PD-ETD (Figure 1) has also been implemented and compared to traditional HCD/ETD methods. In this approach, an LTQ Orbitrap Velos™ mass spectrometer equipped with ETD acquires HCD spectra in a data-dependent fashion. The instrument identifies diagnostic glycan oxonium (product) ions at ppm mass accuracy on the fly in the HCD spectra and only

**FIGURE 6. Orbitrap HCD spectrum and ion trap ETD spectrum of histidine-rich glycoprotein *N*-linked glycopeptide T340-353 precursor at *m/z* 958.383 (4+) from human serum.**

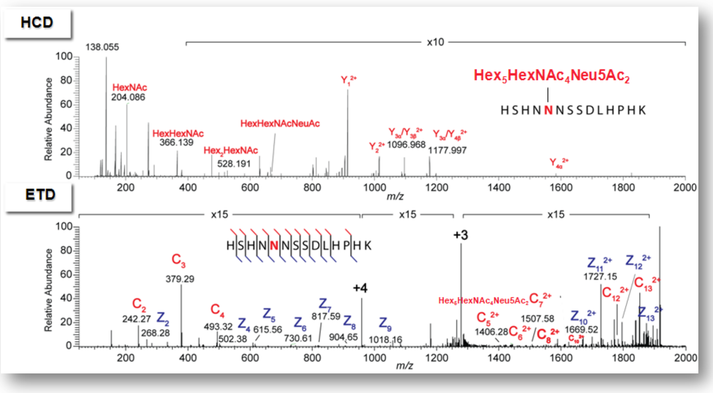
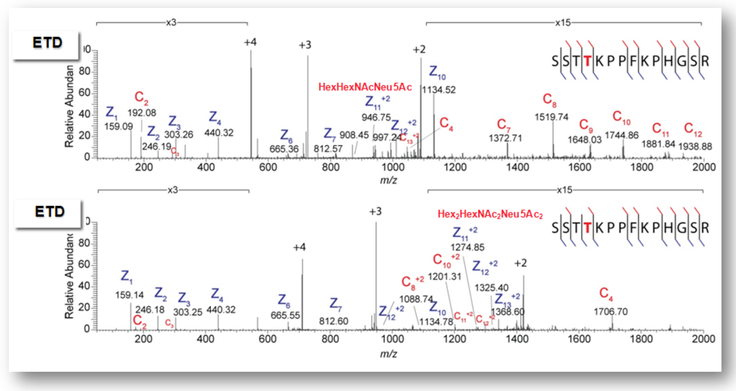
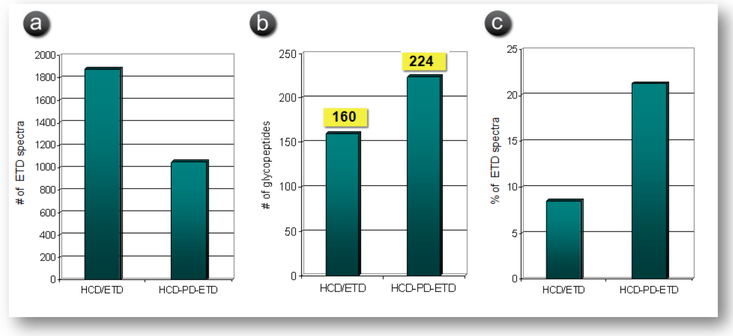



**FIGURE 7. LC-MS ion trap ETD spectrum of histidine-rich glycoprotein *O*-linked glycopeptide T271-284 precursor at *m/z* 546.515 (4+) and *m/z* 710.571 (4+) from human serum.**




triggers ETD spectra for the glycopeptide precursors. This approach increases overall productivity for MS analysis of glycopeptides by acquiring ETD spectra only when a glycopeptide is detected. Additionally, this approach minimizes overall file size and the number of ETD spectra that are extrapolated to characterize glycopeptides. This automated workflow was applied to analysis of glycopeptides enriched (ZIC-HILIC, $TiO_2$ and SAX) from human serum. Overall, 706 intact *N*- and *O*-linked glycopeptides were identified in human serum. This novel HCD-PD-ETD acquisition strategy significantly outperformed the standard HCD/ETD alternating acquisition method. Figure 8 shows the comparison of the HCD-PD-ETD approach to the traditional alternating HCD/ETD strategy for the SAX-enriched sample. Overall, 800 less ETD spectra are acquired in the HCD-PD-ETD approach (Figure 8a), but far more glycopeptides are identified in comparison to alternating HCD/ETD (224 vs 160, Figure 8b). The HCD-PD-ETD approach resulted in more than twice as many ETD spectra identified than the alternating HCD/ETD strategy (Figure 8c).

**FIGURE 8. Comparison of HCD/ETD *vs* HCD-PD-ETD for SAX enriched sample. (a.) Number of ETD spectra acquired, (b.) number of glycopeptides identified and (c.) percentage of ETD spectra identified.**




## Conclusion

- A complementary approach of different enrichment strategies for glycopeptide analysis was demonstrated.
- The novel HCD-PD-ETD approach on an LTQ Orbitrap Velos mass spectrometer increased the overall productivity for MS analysis of glycopeptides by streamlining data analysis and improving dynamic range and duty cycle.
- GlycoMaster software automated the analysis of the data acquired from the combined fragmentation techniques.
- The fully automated workflow enabled identification of 149 glycoproteins, 706 *N*- and *O*-linked glycopeptides, and 195 glycosylation sites.

## References

1. Takegawa, Y.; Deguchi, K.; Ito, H.; Keira, T.; Nakagawa, H.; Nishimura, S.; Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography, *J Sep Sci.* **2006**, 16, 2533-2540.
2. Larsen, M.R.; Jensen, S.S.; Jakobsen, L.A.; Heegaard, N.H. Exploring the sialiome using titanium dioxide chromatography and mass spectrometry, *Mol. Cell Proteomics,* **2007**, 6, 1778-1787.
3. Larsen, M.R.; Hojrup, P.; Roepstorff, P. Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol Cell Proteomics*, **2005**, 4, 107-119.



Thermo SCIENTIFIC