ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Simultaneous N-Deglycosylation and Digestion of Complex Samples on S-Traps Enables Efficient Glycosite Hypothesis Generation

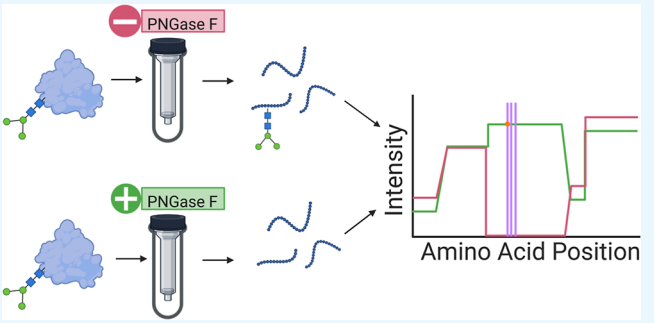
Christine M. DeRosa, Simon D. Weaver, Chien-Wei Wang, Naviya Schuster-Little, and Rebecca J. Whelan*

Cite This: *ACS Omega* 2023, 8, 4410−4418

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: N-linked glycosylation is an important post-translational modification that is difficult to identify and quantify in traditional bottom-up proteomics experiments. Enzymatic deglycosylation of proteins by peptide:*N*-glycosidase F (PNGase F) prior to digestion and subsequent mass spectrometry analysis has been shown to improve coverage of various N-linked glycopeptides, but the inclusion of this step may add up to a day to an already lengthy sample preparation process. An efficient way to integrate deglycosylation with bottom-up proteomics would be a valuable contribution to the glycoproteomics field. Here, we demonstrate a proteomics workflow in which deglycosylation and proteolytic digestion of samples occur simultaneously using suspension trapping (S-Trap). This approach adds no time to standard digestion protocols. Applying this sample preparation strategy to a human serum sample, we demonstrate improved identification of potential N-glycosylated peptides in deglycosylated samples compared with non-deglycosylated samples, identifying 156 unique peptides that contain the N-glycosylation motif (asparagine-X-serine/threonine), the deamidation modification characteristic of PNGase F, and an increase in peptide intensity over a control sample. We expect that this rapid sample preparation strategy will assist in the identification and quantification of both known and potential glycoproteins. Data are available via ProteomeXchange with the identifier PXD037921.



INTRODUCTION

Glycosylation is an important and prevalent post-translational modification that plays a role in diverse biological functions including protein folding, intracellular signaling, and immune response.[1] Glycosylation occurs when glycan chains become covalently attached to side chains of proteins and leads to increased structural heterogeneity.[2] In response to physiological changes, the structure and distribution of glycans at a particular glycosite (micro-heterogeneity) and the presence or absence of a glycan at a glycosite (macro-heterogeneity) can be altered.[3,4] The two most common types of protein glycosylation are O-linked glycosylation, which occurs at the oxygen of serine or threonine residues, and N-linked glycosylation, which occurs at the amide nitrogen of the asparagine residue's side chain in the predictable sequence asparagine-X-serine/threonine (NXS or NXT), where X is any amino acid except proline.[5,6]

Because of their biological importance, glycoproteins are a high-value target for structural characterization and quantitation via mass spectrometry-based proteomics, with bottom-up proteomics being the approach used most frequently. Bottom-up proteomics is a powerful tool that can be used for the identification and quantification of many proteins by digesting proteins into peptides prior to analysis by nano-liquid chromatography−tandem mass spectrometry (LC−MS/MS).[7] However, it is difficult to characterize glycosylation with traditional bottom-up proteomics.[8] While non-modified peptides can be easily matched to their resulting MS spectra, the additional and variable mass of glycans results in the omission of glycopeptides in the results of standard database searches. Many proteomic software packages, such as MSFragger[9,10] and MetaMorpheus,[11,12] are equipped with glycomodules to enable glycoproteomic analysis;[13] however, enrichment steps are almost always necessary to detect glycopeptides[14] whose signals are suppressed by the high number of non-glycosylated peptides in a given sample.[15] Glycopeptide enrichment methods include lectin affinity chromatography,[16] hydrophilic interaction chromatography,[17,18] and titanium dioxide chromatography.[19] These methods can be used individually or in combination to improve glycoprotein coverage.[20] While many enrichment

Received: December 19, 2022
Accepted: January 10, 2023
Published: January 20, 2023

strategies have been developed, most reported techniques achieve only moderate efficiency and display significant bias toward particular subsets of glycopeptides.[21] In addition, inclusion of glycan enrichment increases the number of sample-handling steps required prior to analysis with a concomitant increase in time.

By removing glycans from peptides, either chemically[22] or enzymatically,[23] before or after proteolytic digestion, the base peptides can be identified using traditional bottom-up proteomics without the variable mass addition.[24] A common *N*-glycosidase is peptide:*N*-glycosidase F (PNGase F), which is efficient for the removal of *N*-glycans from proteins.[25] PNGase F cleaves the internal glycoside bond between the asparagine residue and the innermost monosaccharide of the *N*-glycan, *N*-acetylglucosamine (GlcNAc).[26] It is a widely applicable enzyme, able to hydrolyze high-mannose, hybrid, and bi-, tri-, and tetra-antennary oligosaccharides.[25] PNGase deglycosylation results in a free, complete sugar chain and the conversion of asparagine to aspartic acid through a deamidation modification,[27] which can be easily detected using mass spectrometry by a characteristic mass increase of 0.984 Da at the site of deamidation.[28] Although deglycosylating samples prior to traditional digestion can increase proteomic coverage, the required reactions and sample cleanup can add up to a day to already lengthy sample preparation workflows.

Here, we show that PNGase F can be directly coupled with proteolytic digestion using suspension trapping (S-Traps), a powerful tool for proteomics sample processing.[29−32] In S-Trap sample preparation, a sodium dodecyl sulfate (SDS)-solubilized protein solution is reduced and alkylated before acidification and addition of methanol to create a protein suspension. The suspension is then centrifuged through the S-Trap spin column and proteins are retained on a quartz filter, which can then be washed free of contaminants, and a protease of choice can be added to begin digestion. S-Traps are most commonly used with trypsin[33,34] but can also be used with other enzymes such as Lys-C[31] and GluC.[35] S-Trap sample preparation has been used in a variety of applications including protein extraction from mouse microglia cells[36] and proteome comparison of the epiretinal membrane and inner limiting membrane from human retinal cells.[37] The use of S-Traps for simultaneous enzymatic deglycosylation and proteolytic digestion has not yet been reported.[32] We demonstrate that PNGase F deglycosylation and tryptic digestion can occur simultaneously on S-Traps, enabling an optimized and efficient sample preparation workflow that takes the same amount of time as standard proteolytic digestion. We focus on N-glycosylation because of the ease of large-scale complex protein analysis due to the predictable, straightforward sites of N-glycosylation and availability of universal enzymes to specifically remove *N*-glycans from the peptide backbone.[38] By comparing deglycosylated samples with non-deglycosylated (control) samples, we can quickly hypothesize which peptides may contain *N*-glycans based on their amino acid sequences. These potential glycopeptides can then be followed up with additional analysis.

MATERIALS AND METHODS

Materials. Male human serum (HS), SDS, iodoacetamide (IAA), triethylammonium bicarbonate (TEAB), albumin from chicken egg white (ovalbumin), trypsin inhibitor from chicken egg white (ovomucoid), and conalbumin from chicken egg white (ovotransferrin) were purchased from Millipore Sigma (St. Louis, MO). Tris(2-carboxyethyl)phosphine (TCEP), deoxycholic acid (DCA), phosphoric acid, and methanol (Burdick & Jackson) were obtained through VWR. Pure formic acid (99% purity) (FA), acetonitrile (ACN), and C_{18} ZipTips were purchased from Fisher Scientific (Hanover Park, IL). S-Traps were purchased from ProtiFi (Huntington, NY), and trypsin gold was purchased from Promega (Madison, WI). Peptide:*N*-glycosidase F (PNGase F), deglycosylation mix buffer 1, and deglycosylation mix buffer 2 were purchased from New England BioLabs (Ipswich, MA). Tris-glycine sample buffer and Novex 16% Tris-glycine mini protein gels were purchased from Thermo (Waltham, MA). RNase B was a gift from Matthew Champion in the Department of Chemistry and Biochemistry at the University of Notre Dame.

Gels. For RNase B gels, 20 μg of RNase B was mixed with 1 μg of PNGase F and one of four reaction buffers (water, 100 mM TEAB, deglycosylation mix buffer 1, or deglycosylation mix buffer 2). For chicken protein gels, 20 μg of chicken protein ovalbumin, ovomucoid, ovotransferrin, or a cocktail of all three proteins (20 μg each, 60 μg total protein) was mixed with 1 μg of PNGase F in 100 mM TEAB buffer. All reactions were incubated at 37 °C for 3 h, quenched by mixing with equal volume of Tris-glycine sample buffer, and denatured at 95 °C for 10 min. 3 μg of total protein from each reaction was separated by 16% SDS−polyacrylamide gel electrophoresis (SDS−PAGE) at 100 V for 3 h and stained with the colloidal Coomassie blue stain (Bio-rad).

Deglycosylation and Digestion. 20 μg of total protein was digested following previously described methods.[39] Briefly, proteins were denatured and reduced with 10% SDS and 10 mM TCEP at 95 °C for 10 min. 0.2% DCA was included as a passivating agent and 100 mM TEAB was included for buffering. Proteins were alkylated using 10 mM IAA for 30 min at room temperature in the dark. The alkylation reaction was quenched with 1.2% phosphoric acid. A protein suspension was formed by the addition of 100 mM TEAB in 90% methanol, and the suspension was spun onto an S-Trap device and washed following manufacturer's instructions. 1 μL PNGase F in 100 mM TEAB was added to the S-Trap to deglycosylate the proteins retained on the S-Trap and incubated at 37 °C for 0, 2, or 4 h. Proteins were then digested using 750 ng trypsin gold in 100 mM TEAB. Following digestion, peptides were eluted using 100 mM TEAB followed by 0.1% FA in water. The reaction was quenched with 10% FA and a third elution was performed with 50% ACN and 0.1% FA. All eluates were combined, and peptides were desalted using ZipTips and reconstituted at 200 ng/μL in 0.5% FA and 4% ACN for mass spectrometry analysis.

Mass Spectrometry. Peptides were analyzed with a Waters NanoAcquity liquid chromatograph (LC) paired with a Q-Exactive mass spectrometer (Thermo Scientific). 200 ng of peptides was injected into the LC system equipped with a peptide BEH C18 column (Waters, 100 μm × 100 mm, 1.7 μm particle size). Peptides were separated over a 48 min gradient with a flow rate of 0.9 μL/min with a two-solvent system, where solvent A was water containing 0.1% FA and solvent B was ACN containing 0.1% FA. The following linear gradient was used for all samples: 4% B from 0 to 8 min, 4−7% B from 8 to 10 min, 7−33% B from 10 to 30 min, 33−90% B from 30 to 33 min, 90% B until 36 min, 90−4% B for 1 min, and reequilibration at 4% B from 37 to 48 min. The mass

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418

spectrometer settings were identical to those described previously.[40]

Database Searching (PEAKS). Raw data files were searched using PEAKS Online X build 1.4.2020-10-21 (Bioinformatics Solutions, Waterloo, ON, Canada)[41] using the current *Homo sapiens* Uniprot database (downloaded June 27, 2022),[42] with the MUC16 entry (accession ID: Q8WXI7, 14, 152 amino acids) replaced with the version from the 2016 SwissProt database (22, 152 amino acids) for the HS samples, and the current *Gallus gallus* Uniprot database (downloaded June 24, 2022) for the chicken protein samples. The digestion enzyme was set to trypsin with a maximum of two missed cleavages. Precursor mass tolerance was set to 10 ppm, and fragment mass error tolerance was set to 0.02 Da. Carbamidomethylation of C was added as a global modification, and deamidation of N and Q, oxidation of M, pyro-glu conversion from E and Q, and sodium adduction were set as variable modifications. A peptide FDR was set to 1% and protein −10 log *P* was set to ≥20. Peptides of length 6 to 45 amino acids were considered, and common contaminants (including keratin) were filtered out. Two or more unique peptides were required for protein identification.

Database Searching (MSFragger). Control HS injections (no PNGase F) from the 0 h time points were searched using the glyco-*N*-open-HCD workflow[9] in MSFragger (v3.5) with FragPipe (v18.0) and Philosopher (v4.4.0) to detect glycopeptides for comparative purposes. Default settings for this workflow were used, along with the *H. sapiens* database as described above.

Data Analysis. Data analysis was performed using Python (version 3.9.12) in Spyder (version 5.1.5).[43,44] Plots were made with seaborn,[45] PrIntMap-R (https://championlab.shinyapps.io/printmap-r/),[46] matplotlib-venn,[47] and matplotlib,[48] and the following packages were used for data analysis: pandas,[49] re,[43] and numpy.[50]

RESULTS AND DISCUSSION

PNGase F Activity Is Compatible with S-Trap Sample Preparation. The direct compatibility of the S-Trap platform with PNGase F and buffers suitable for PNGase F activity has not previously been demonstrated. We investigated whether PNGase F is compatible with S-Trap sample preparation using RNase B as a model glycoprotein containing one known *N*-glycan that results in a characteristic mass shift upon cleavage.[51] PNGase F is known to be effective at removing glycans in protein deglycosylation mix buffers 1 and 2,[52] whereas TEAB is known to be compatible with S-Trap sample preparation and mass spectrometry analysis.[53] To test whether a single buffer system would be compatible with both PNGase F and S-Traps, RNase B was deglycosylated using PNGase F in four different buffer conditions: the two deglycosylation buffers provided with commercially available PNGase F, TEAB, and water. As shown in Figure 1A, PNGase F is found to be effective in all four buffer environments. The ability of PNGase F to function effectively in TEAB eliminates the need for any buffer exchange steps during sample preparation.

To further investigate the compatibility of PNGase F with S-Traps, the performance time of PNGase F was examined. Five samples of RNase B were treated with PNGase F and incubated for 1, 5, 15, 60, or 120 min, followed by heat shock to stop the reaction. A similar gel shift was observed for all samples, indicating that the deglycosylation reaction reaches completion by the first time point (Figure 1B). We

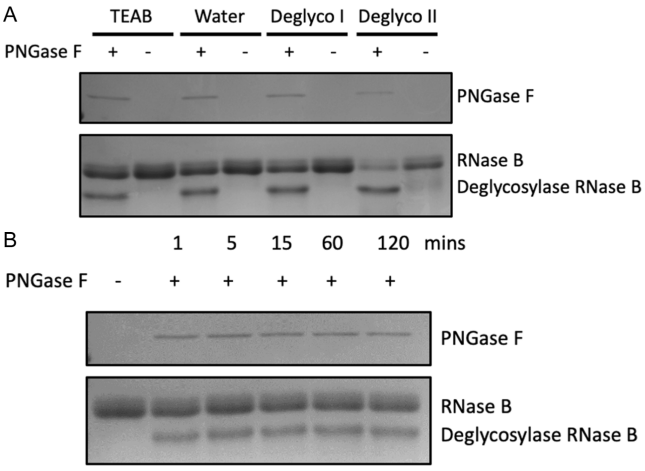


Figure 1. PNGase F works in buffers compatible with S-Traps and mass spectrometry and deglycosylates RNAse B rapidly. (A) Gel shift assay measuring PNGase F activity in different buffer conditions. RNase B was treated with (+) or without (−) PNGase F, and 3 μg protein was loaded in each lane. (B) Gel shift assay measuring PNGase F activity over the course of 2 h. RNase B was treated with (+) or without (−) PNGase F, and 3 μg protein was loaded in each lane. Cropped gel images are shown for clarity; uncropped gel images are shown in the Supporting Information, Figure S1.

hypothesize that because PNGase F works quickly and performs the same at the beginning and end of a two-hour period, it is not necessary to delay the addition of trypsin after the addition of PNGase F.

Specific Known Glycosites Can Be Identified in Chicken Proteins When Deglycosylated and Digested on S-Trap. Having confirmed the compatibility of PNGase F with S-Trap buffers, we examined three chicken proteins, which all contain at least one previously annotated glycosite in Uniprot. Each protein—albumin from chicken egg white (ovalbumin, accession ID: P01012), trypsin inhibitor from chicken egg white (ovomucoid, accession ID: P01005), and conalbumin from chicken egg white (ovotransferrin, accession ID: P02789)—was successfully deglycosylated using PNGase F individually (Supporting Information, Figure S2, lanes 2−7) and combined and deglycosylated together (Supporting Information, Figure S2, lanes 8 and 9).

The cocktail of chicken proteins was deglycosylated and digested simultaneously. For each amino acid in the protein sequence of interest, the area (Figure 2) or number of peptide-spectrum matches (PSMs) (Supporting Information, Figure S3) was summed, and the ratio of deglycosylated sample to control sample, or fold change, was calculated. Ovalbumin has one annotated N-glycosylation site at N293 with heterogeneous neutral glycan attachments.[54] We observed a 129× increase in intensity for the peptides containing that residue, whereas we did not see an increase in peptide intensity at N312, which is part of an N-glycosylation motif but is not an annotated glycosylation site (Figure 2A). This evidence supports the annotation of N293 but not N312 as N-glycosites. Ovomucoid has five known N-glycosylation sites at N34, N77, N93, N99, and N199[17] and one unannotated site containing the N-glycosylation motif at N182. Comparing the observed peptide intensity of the deglycosylated sample to the control sample across the amino acid sequence of ovomucoid, we observed a more than 1000× increase at N93 and N99, a 60× increase at N34, and a 50× increase at N77 (Figure 2B).

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418

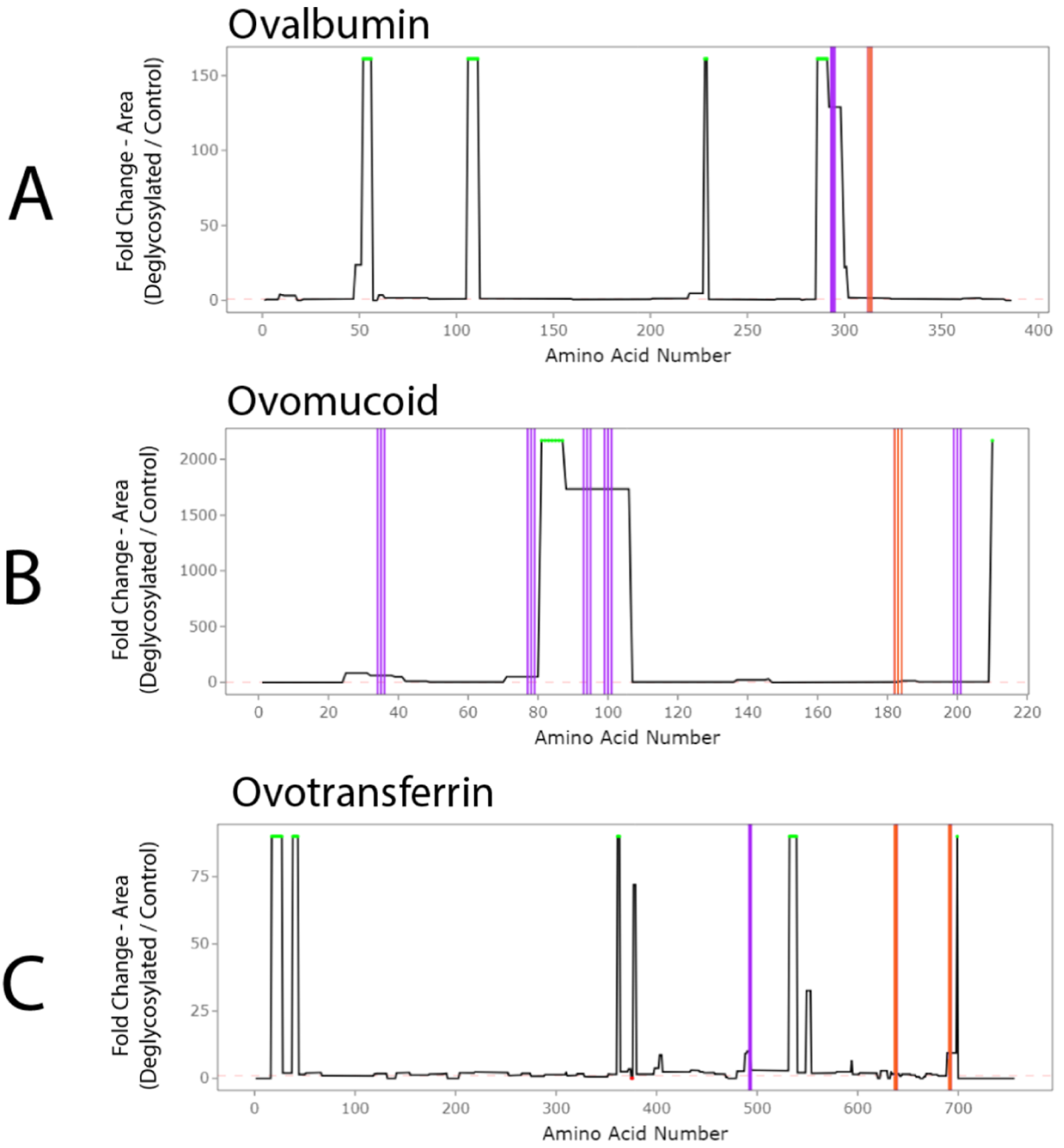


Figure 2. Chicken proteins with known glycosites can be deglycosylated using PNGase F on S-Trap. Specific known glycosites can be identified in (A) albumin from chicken egg white (ovalbumin), (B) trypsin inhibitor from chicken egg white (ovomucoid), and (C) conalbumin from chicken egg white (ovotransferrin). Locations of annotated *N*-glycan motifs are indicated by vertical purple lines, and those of unannotated *N*-glycan motifs are indicated by vertical orange lines. Fold change of the deglycosylated sample relative to the control sample is represented on the *y*-axis, with green representing infinite fold change and red representing negative infinite fold change. Each trace is an average area fold change resulting from $n = 9$ injections.

Ovotransferrin has one known N-glycosylation site at N492[55] and two unannotated sites at N637 and N691. We observed a 10-fold increase in intensity for peptides containing the N492 site in the deglycosylated sample compared to the control (Figure 2C). These trends are highly reproducible across technical replicates (Supporting Information, Figure S4).

Coupling Deglycosylation and S-Trap Sample Preparation Increases the Number of Peptides Identified Containing the *N*-Glycan Motif and Deamidation Modification. We next sought to characterize the suitability of the S-Trap platform to perform simultaneous N-deglycosylation and trypsin digestion of more complex samples. Using HS as a model sample, we investigated three different time intervals (0, 2, and 4 h) between the addition of PNGase F and trypsin (Figure 3A). HS is a well-characterized sample of moderate-to-high complexity containing hundreds of proteins, and it has great clinical relevance as a readily accessible biofluid for diagnosis and prognosis.[56] Across all replicates and time points, an average of 72 ± 5% peptides were identified in both the control and deglycosylated samples ($n = 9$). Over time, we did not observe any trend in the number of peptides identified in the deglycosylated samples,

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418

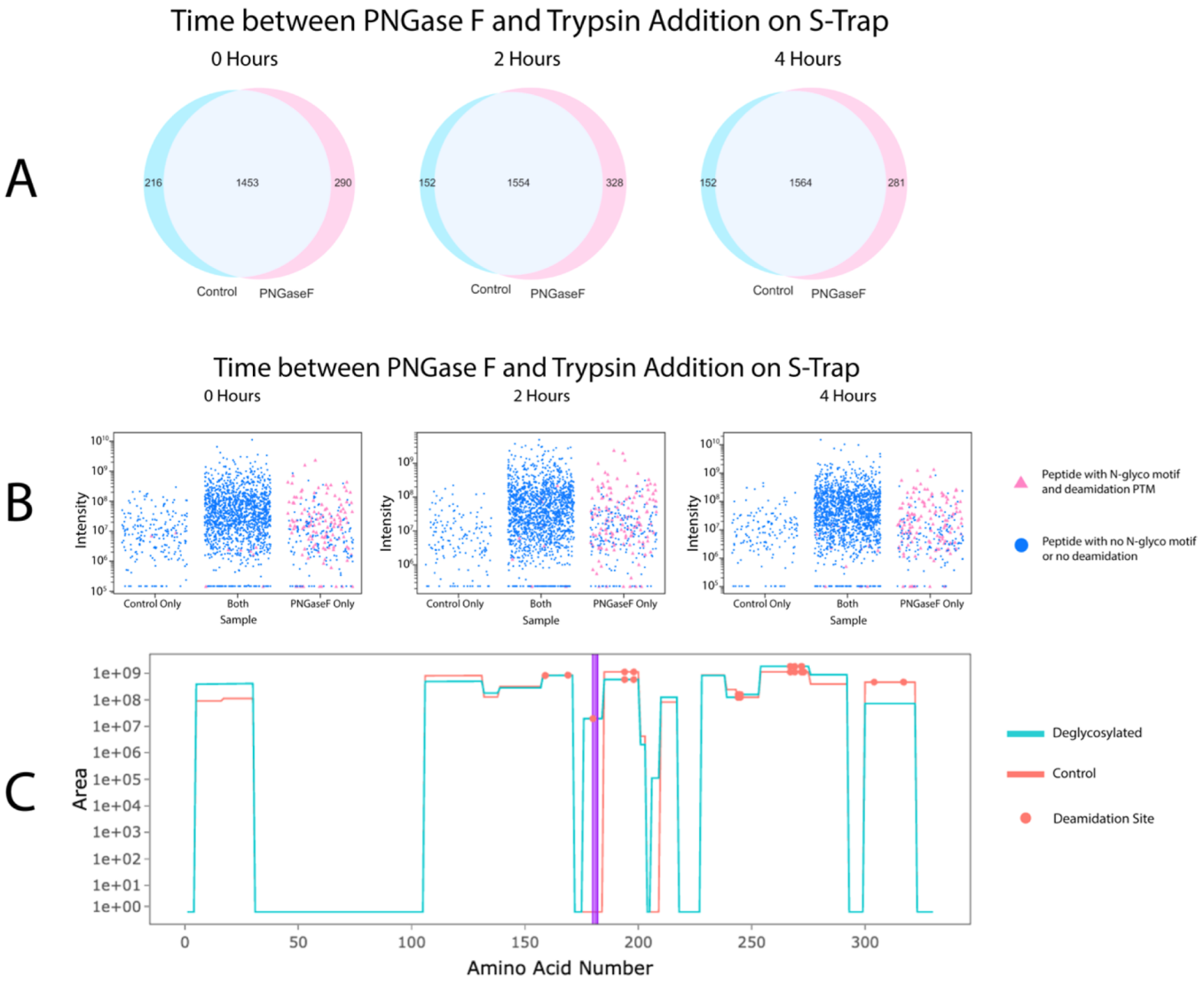


Figure 3. Deglycosylated and digested HS; peptide data analysis. (A) PNGase F does not interfere with trypsin's activity or the number of peptides identified. Overlap in peptide identification between control sample (blue) and deglycosylated sample (pink) in one biological replicate (replicate A) over three time points of trypsin addition. Other biological replicates show the same trend and are found in the Supporting Information (Figure S5). (B) Many more peptides containing the *N*-glycan motif and deamidation modification are found. Identified peptides by sample are plotted against their intensity across three time points of trypsin addition. Peptides that were detected but did not have an intensity value were assigned the minimum intensity of the plot, the approximate limit of quantitation of the instrument. Peptides containing both the *N*-glycan motif (N-X-S/T) and deamidation modification are plotted as pink triangles and those without appear as blue dots. Data shown are from one biological replicate (replicate A) with triplicate injections. Other biological replicates show the same trend and can be found in the Supporting Information (Figure S6). (C) The specific known N-glycosylation site can be identified in the heavy chain of IgG. The area is plotted for each amino acid position of the protein in the control (orange) and deglycosylated (blue) samples. Locations of annotated N-glycosites are indicated by vertical purple lines, and locations of deamidation PTMs are represented by orange dots. Data are shown from one biological replicate (replicate A, 0 h). Other biological replicates and time points display the same trend.

and on average, there was less than 15% change in the number of peptides identified across three biological replicates. Given that there are a comparable number of total peptides identified in both the deglycosylated and control samples, we conclude that PNGase F does not interfere with the digestive activity of trypsin.

Additionally, at each time point, many more peptides containing the *N*-glycan motif (NXT/S) and a deamidation modification are identified in the PNGase F sample compared to the control sample (Figure 3B). 156 unique potential N-glycosylated peptides were identified across all replicates and time points, with an average of 97 ± 25 potential N-glycosylated peptides identified per sample (n = 9). A list of these peptides for each biological replicate can be found in the Supporting Information (Tables S1−S10). Because PNGase F is used to remove *N*-glycans from proteins and results in the deamidation of the asparagine residue (a mass increase in 0.98 Da), we hypothesize that peptides containing both the motif and the deamidation shift were successfully deglycosylated with PNGase F.[57] Furthermore, given that the number of total peptides and the number of peptides containing the *N*-glycan motif and deamidation modification do not improve with increased deglycosylation time, we conclude that trypsin and PNGase F may be added simultaneously to the S-Trap. This allows for the deglycosylation assay to be run in the same amount of time as a standard digestion.

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418

Many proteins identified in the HS sample exhibit an increase in intensity in peptides containing *N*-glycan motifs and deamidation modifications in the deglycosylated sample trace compared to the control trace. Follow-up analysis of these peptides yields evidence of previously quantified glycoproteins with known and annotated glycosites.[58] IgG is an abundant protein in HS that contains an annotated N-glycosite in the Fc domain of the heavy chain[54] with importance in the function of immunoglobulin both in immunity and as drugs.[59] In the deglycosylated samples, we identify peptides containing a deamidated asparagine residue at the known site of N-glycosylation (N180) that were not identified in the control samples (Figure 3C). The area fold-change plot for this protein can be seen in the Supporting Information, Figure S7. Many other proteins with probable glycosylation sites were identified, including apolipoprotein (Supporting Information, Figure S8), α-2-HS-glycoprotein (Supporting Information, Figure S9), and haptoglobin (Supporting Information, Figure S10). We conclude that this method of simultaneous N-deglycosylation and proteolytic digestion generates valuable information about potential N-glycosites that can be validated for the presence of confirmed glycans.

The generation of PNGase F-derived peptides was considered as a possible source of contamination. To investigate whether detectable PNGase F peptides are generated via trypsin digestion, we searched raw data files from the 0 h deglycosylated and control serum samples in PEAKS, using the HS database with the addition of the PNGase F sequence from the Protein Data Bank (https://www.rcsb.org/structure/1pgs). All other search parameters were as described above. In the control sample, 0 PNGase F peptides were identified. In the deglycosylated sample, 29 unique PNGase F peptides were identified, corresponding to 84% PNGase F sequence coverage. The most intense PNGase F peptide had an area of 4.9×10^8, while the most intense non-PNGase F peptide had an area of 5.4×10^9. We interpret these data to mean that trypsin is digesting PNGase F to produce detectable peptides but that these peptides will not be identified in database searches of human samples. A similar number of total peptides identified between the control and PNGase F samples across all time points (Figure 3) supports the claim that PNGase F peptide generation does not interfere with the detection of target peptides.

Deglycosylation on S-Trap Complements N-Glyco Peptide Database Search. MSFragger is a database search software that contains a N-Glyco search module that analyzes tandem mass spectra of intact glycopeptides and then identifies and validates their glycan sites.[9] In comparing three HS biological replicates from the 0 h time point, 53 unique glycopeptides were identified by MSFragger's N-Glyco search. Using the deglycosylation method described in this report, 138 unique potential *N*-glycopeptides were identified in the same three biological replicates, 160% more than were identified in the MSFragger N-Glyco search for intact glycopeptides (Figure 4A). The total number of peptides identified between each identification method differed by no more than 3% (Figure 4B). This observation suggests that the increase in potential *N*-glycopeptides is not simply due to an overall increase in peptides from the different search strategies, although this analysis was not intended to directly compare search techniques but rather to enable a comparison between intact glycopeptide analysis and deglycopeptide analysis. While we see an increase in the number of *N*-glycopeptides using S-Trap deglycosylation, it is important to note that no information about the glycan composition or attached mass is revealed by this approach, unlike MSFragger, which identifies glycans from their mass difference. Deglycosylation on S-Trap can be used to generate hypotheses about which proteins contain N-linked glycans and can identify potential glycopeptides missed in *N*-glycopeptide database searches, but other methods are necessary for validation. Deglycosylation and digestion on S-Trap combined with the use of N-Glyco search modules can provide a robust assay for discovery and validation of glycans on proteins of interest.

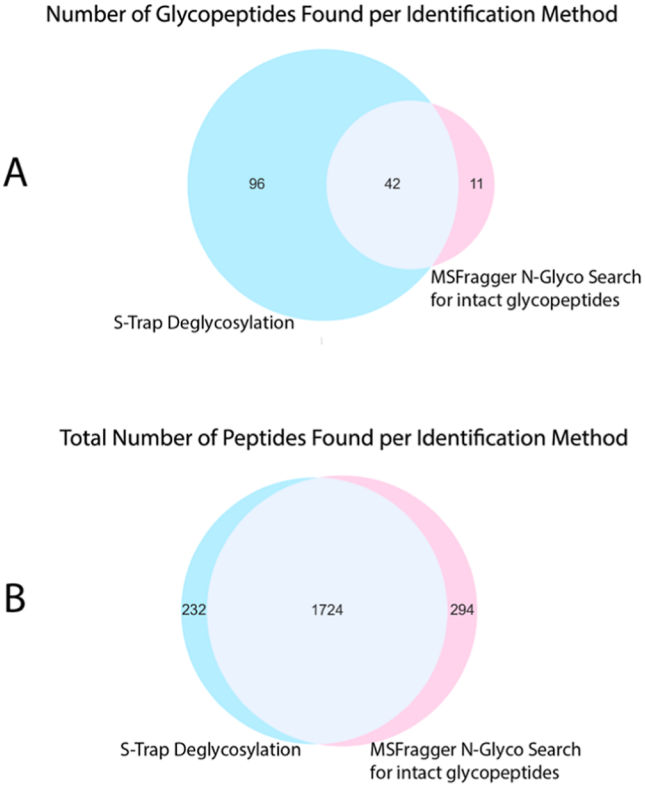


Figure 4. Comparison of peptides identified using S-Trap deglycosylation and MSFragger N-Glyco search of intact glycopeptides. (A) Overlap in *N*-glycopeptide identification between S-Trap deglycosylation (blue) and MSFragger N-Glyco search of intact glycopeptides (pink) in three biological replicates of 0 h time point of trypsin addition. (B) Overlap in total peptide identification between S-Trap deglycosylation (blue) and MSFragger N-Glyco search of intact glycopeptides (pink) in three biological replicates of 0 h time point of trypsin addition.

CONCLUSIONS

We demonstrate that simultaneously deglycosylating and digesting protein samples on an S-Trap improves the identification of potential N-glycosylated peptides. By co-incubating PNGase F and trypsin on S-Traps, we perform sample preparation in the same amount of time as a standard protein digestion. Identifying peptides that appeared in only deglycosylated samples, had a deamidation modification, and contained an *N*-glycan motif was a fast and effective way to generate a hypothesized list of glycopeptides. Future experiments could use this method to discover potential biomarkers in diseases such as cancer, where aberrant glycosylation is known to play a role. We anticipate using this technique in our

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418

future work to study glycosylation in samples from ovarian cancer patients to determine whether differential glycosylation is present throughout the proteome. Other proteins in a complex proteome with N-glycosylation such as monoclonal antibodies and the Rh blood group proteins could be studied using this method as well. Additionally, with the use of suitable enzymes, this technique could be extended to the more difficult, but important, question of O-linked glycosylation on proteins of interest.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c08071.

Gel images, glycosite maps, Venn diagrams, scatter plots of identified peptides and glycopeptides as a function of time interval between enzyme additions, and tables of identified peptides and corresponding proteins (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Rebecca J. Whelan − *Department of Chemistry, University of Kansas, Lawrence, Kansas 66047, United States;* orcid.org/0000-0002-9293-1528; Phone: 785-864-1980; Email: rwhelan1@ku.edu

Authors

Christine M. DeRosa − *Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States;* orcid.org/0000-0003-3685-9098

Simon D. Weaver − *Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; Integrated Biomedical Sciences Graduate Program, University of Notre Dame, Notre Dame, Indiana 46656, United States;* orcid.org/0000-0003-4354-0309

Chien-Wei Wang − *Department of Chemistry, University of Kansas, Lawrence, Kansas 66047, United States;* orcid.org/0000-0002-0449-8366

Naviya Schuster-Little − *Department of Chemistry, University of Kansas, Lawrence, Kansas 66047, United States;* orcid.org/0000-0001-7336-9042

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c08071

Notes

The authors declare no competing financial interest.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[60] partner repository with the data set identifier PXD037921.

■ ACKNOWLEDGMENTS

This work was supported by the University of Notre Dame Advancing Our Vision Fund in Analytical Science and Engineering (to R.J.W.). S.D.W. and N.S.L. received fellowship support from the Chemistry−Biochemistry−Biology Interface (CBBI) Program at the University of Notre Dame, supported by training grant T32GM075762 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The authors thank Bill Boggess and the University of Notre Dame Mass Spectrometry and Proteomics Facility for expert technical assistance. The authors thank Matthew Champion (University of Notre Dame) for providing RNaseB, Manish Patankar (University of Wisconsin) for helpful discussions, and two anonymous reviewers for constructive and insightful comments.

■ REFERENCES

(1) Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3−49.

(2) Pirillo, A.; Svecla, M.; Catapano, A. L.; Holleboom, A. G.; Norata, G. D. Impact of protein glycosylation on lipoprotein metabolism and atherosclerosis. *Cardiovasc. Res.* **2021**, *117*, 1033−1045.

(3) Stavenhagen, K.; Hinneburg, H.; Thaysen-Andersen, M.; Hartmann, L.; Silva, D. V.; Fuchser, J.; Kaspar, S.; Rapp, E.; Seeberger, P. H.; Kolarich, D. Quantitative mapping of glycoprotein micro-heterogeneity and macro-heterogeneity: an evaluation of mass spectrometry signal strengths using synthetic peptides and glycopeptides. *J. Mass Spectrom.* **2013**, *48*, 627−639.

(4) Zacchi, L. F.; Schulz, B. L. N-glycoprotein macroheterogeneity: biological implications and proteomic characterization. *Glycoconjugate J.* **2015**, *33*, 359−376.

(5) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126*, 855−867.

(6) Mohanty, S.; P Chaudhary, B. P.; Zoetewey, D. Structural Insight into the Mechanism of N -Linked Glycosylation by Oligosaccharyltransferase. *Biomolecules* **2020**, *10*, 624.

(7) Dupree, E. J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B. A.; Darie, C. C. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of This Field. *Proteomes* **2020**, *8*, 14.

(8) Angel, P. M.; Lim, J.; Wells, L.; Bergmann, C.; Orlando, R. A potential pitfall in ^{18}O-based N-linked glycosylation site mapping. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 674−682.

(9) Polasky, D. A.; Yu, F.; Teo, G. C.; Nesvizhskii, A. I. Fast and comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco. *Nat. Methods* **2020**, *17*, 1125−1132.

(10) Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry−based proteomics. *Nat. Methods* **2017**, *14*, 513.

(11) Lu, L.; Riley, N. M.; Shortreed, M. R.; Bertozzi, C. R.; Smith, L. M. O-Pair Search with MetaMorpheus for O-glycopeptide characterization. *Nat. Methods* **2020**, *17*, 1133.

(12) Solntsev, S. K.; Shortreed, M. R.; Frey, B. L.; Smith, L. M. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J. Proteome Res.* **2018**, *17*, 1844−1851.

(13) Rangel-Angarita, V.; Malaker, S. A. Mucinomics as the Next Frontier of Mass Spectrometry. *ACS Chem. Biol.* **2021**, *16*, 1866−1883.

(14) Mariño, J.; Bones, J. J.; Kattla, P. M.; Rudd, K. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* **2010**, *6*, 713−723.

(15) Dalpathado, D. S.; Desaire, H. Glycopeptide analysis by mass spectrometry. *Analyst* **2008**, *133*, 731−738.

(16) Yang, Z.; Hancock, W. S. Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. *J. Chromatogr. A* **2005**, *1070*, 57−64.

(17) Hägglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P. A New Strategy for Identification of N-Glycosylated Proteins and Unambiguous Assignment of Their Glycosylation Sites Using HILIC Enrichment and Partial Deglycosylation. *J. Proteome Res.* **2004**, *3*, 556−566.

(18) Chen, R.; Seebun, D.; Ye, M.; Zou, H.; Figeys, D. Site-specific characterization of cell membrane N-glycosylation with integrated hydrophilic interaction chromatography solid phase extraction and LC−MS/MS. *J. Proteomics* **2014**, *103*, 194−203.

(19) Larsen, M. R.; Jensen, S. S.; Jakobsen, L. A.; Heegaard, N. H. H. Exploring the Sialiome Using Titanium Dioxide Chromatography and Mass Spectrometry. *Mol. Cell. Proteomics* **2007**, *6*, 1778−1787.
(20) Huang, J.; Wan, H.; Yao, Y.; Li, J.; Cheng, K.; Mao, J.; Chen, J.; Wang, Y.; Qin, H.; Zhang, W.; Ye, M.; Zou, H. Highly Efficient Release of Glycopeptides from Hydrazide Beads by Hydroxylamine Assisted PNGase F Deglycosylation for N-Glycoproteome Analysis. *Anal. Chem.* **2015**, *87*, 10199−10204.
(21) Morgenstern, D.; Wolf-Levy, H.; Tickotsky-Moskovitz, N.; Cooper, I.; Buchman, A. S.; Bennett, D. A.; Beeri, M. S.; Levin, Y. Optimized Glycopeptide Enrichment Method−It Is All About the Sauce. *Anal. Chem.* **2022**, *94*, 10308.
(22) Chen, W.; Smeekens, J. M.; Wu, R. Comprehensive Analysis of Protein N-Glycosylation Sites by Combining Chemical Deglycosylation with LC−MS. *J. Proteome Res.* **2014**, *13*, 1466−1473.
(23) Segu, Z. M.; Hussein, A.; Novotny, M. V.; Mechref, Y. Assigning N-Glycosylation Sites of Glycoproteins Using LC/MSMS in Conjunction with Endo-M/Exoglycosidase Mixture. *J. Proteome Res.* **2010**, *9*, 3598−3607.
(24) Mills, K.; Johnson, A. W.; Diettrich, O.; Clayton, P. T.; Winchester, B. G. A strategy for the identification of site-specific glycosylation in glycoproteins using MALDI TOF MS. *Tetrahedron: Asymmetry* **2000**, *11*, 75.
(25) Tarentino, A. L.; Gomez, C. M.; Plummer, T. H. Deglycosylation of Asparagine-Linked Glycans by Peptide: N-Glycosidase F. *Biochemistry* **1985**, *24*, 4665−4671.
(26) Huang, Y.; Orlando, R. Kinetics of N-Glycan Release from Human Immunoglobulin G (IgG) by PNGase F: All Glycans Are Not Created Equal. *J. Biomol. Tech.* **2017**, *28*, 150−157.
(27) Palmisano, G.; Melo-Braga, M. N.; Engholm-Keller, K.; Parker, B. L.; Larsen, M. R. Chemical Deamidation: A Common Pitfall in Large-Scale N-Linked Glycoproteomic Mass Spectrometry-Based Analyses. *J. Proteome Res.* **2012**, *11*, 1949−1957.
(28) Hao, P.; Adav, S. S.; Gallart-Palau, X.; Sze, S. K. Recent advances in mass spectrometric analysis of protein deamidation. *Mass Spectrom. Rev.* **2016**, *36*, 677−692.
(29) Ludwig, K. R.; Schroll, M. M.; Hummon, A. B. Comparison of in-solution, FASP, and S-trap based digestion methods for bottom-up proteomic studies. *J. Proteome Res.* **2018**, *17*, 2480−2490.
(30) HaileMariam, M.; Eguez, R. V.; Singh, H.; Bekele, S.; Ameni, G.; Pieper, R.; Yu, Y. S. Trap, an ultrafast sample-preparation approach for shotgun proteomics. *J. Proteome Res.* **2018**, *17*, 2917−2924.
(31) Zougman, A.; Selby, P. J.; Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **2014**, *14*, 1006.
(32) Duong, V.-A.; Park, J.-M.; Lee, H. A review of suspension trapping digestion method in bottom-up proteomics. *J. Sep. Sci.* **2022**, *45*, 3150.
(33) Elinger, D.; Gabashvili, A.; Levin, Y. Suspension Trapping (S-Trap) Is Compatible with Typical Protein Extraction Buffers and Detergents for Bottom-Up Proteomics. *J. Proteome Res.* **2019**, *18*, 1441−1445.
(34) Yang, Y.; Anderson, E.; Zhang, S. Evaluation of six sample preparation procedures for qualitative and quantitative proteomics analysis of milk fat globule membrane. *Electrophoresis* **2018**, *39*, 2332.
(35) Bouwman, K. M.; Tomris, I.; Turner, H. L.; van der Woude, R.; Shamorkina, T. M.; Bosman, G. P.; Rockx, B.; Herfst, S.; Snijder, J.; Haagmans, B.; Ward, A. B.; Boons, G. J.; de Vries, R. P. Multimerization- and glycosylation-dependent receptor binding of SARS-CoV-2 spike proteins. *PLoS Pathog.* **2021**, *17*, No. e1009282.
(36) Guergues, J.; Zhang, P.; Liu, B.; Stevens, S. M. Improved Methodology for Sensitive and Rapid Quantitative Proteomic Analysis of Adult-Derived Mouse Microglia: Application to a Novel In Vitro Mouse Microglial Cell Model. *Proteomics* **2019**, *19*, 1800469.
(37) Christakopoulos, C.; Cehofski, L. J.; Christensen, S. R.; Vorum, H.; Honoré, B. Proteomics reveals a set of highly enriched proteins in epiretinal membrane compared with inner limiting membrane. *Exp. Eye Res.* **2019**, *186*, 107722.
(38) Pan, S.; Chen, R.; Aebersold, R.; Brentnall, T. A. Mass Spectrometry Based Glycoproteomics—From a Proteomics Perspective. *Mol. Cell. Proteomics* **2011**, *10*, R110.
(39) Schuster-Little, N.; Madera, S.; Whelan, R. Developing a mass spectrometry−based assay for the ovarian cancer biomarker CA125 (MUC16) using suspension trapping (STrap). *Anal. Bioanal. Chem.* **2020**, *412*, 6361−6370.
(40) Weaver, S. D.; Schuster-Little, N.; Whelan, R. J. Preparative capillary electrophoresis (CE) fractionation of protein digests improves protein and peptide identification in bottom-up proteomics. *Anal. Methods* **2022**, *14*, 1103.
(41) Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G. A.; Ma, B. PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification. *Mol. Cell. Proteomics* **2012**, *11*, M111.010587.
(42) Bateman, A.; Martin, M.; Orchard, S.; Magrane, M.; Agivetova, R.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E. H.; Britto, R.; Bursteinas, B.; Bye-A-Jee, H.; Coetzee, R.; Cukura, A.; Da Silva, A.; Denny, P.; Dogan, T.; Ebenezer, T.; Fan, J.; Castro, L. G.; Garmiri, P.; Georghiou, G.; Gonzales, L.; Hatton-Ellis, E.; Hussein, A.; Ignatchenko, A.; Insana, G.; Ishtiaq, R.; Jokinen, P.; Joshi, V.; Jyothi, D.; Lock, A.; Lopez, R.; Luciani, A.; Luo, J.; Lussi, Y.; Macdougall, A.; Madeira, F.; Mahmoudy, M.; Menchi, M.; Mishra, A.; Moulang, K.; Nightingale, A.; Oliveira, C. S.; Pundir, S.; Qi, G.; Raj, S.; Rice, D.; Lopez, M. R.; Saidi, R.; Sampson, J.; Sawford, T.; Speretta, E.; Turner, E.; Tyagi, N.; Vasudev, P.; Volynkin, V.; Warner, K.; Watkins, X.; Zaru, R.; Zellner, H.; Bridge, A.; Poux, S.; Redaschi, N.; Aimo, L.; Argoud-Puy, G.; Auchincloss, A.; Axelsen, K.; Bansal, P.; Baratin, D.; Blatter, M.; Bolleman, J.; Boutet, E.; Breuza, L.; Casals-Casas, C.; De Castro, E.; Echioukh, K. C.; Coudert, E.; Cuche, B.; Doche, M.; Dornevil, D.; Estreicher, A.; Famiglietti, M. L.; Feuermann, M.; Gasteiger, E.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz-Gumowski, N.; Hinz, U.; Hulo, C.; Hyka-Nouspikel, N.; Jungo, F.; Keller, G.; Kerhornou, A.; Lara, V.; Le Mercier, P.; Lieberherr, D.; Lombardot, T.; Martin, X.; Masson, P.; Morgat, A.; Neto, T. B.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pourcel, L.; Pozzato, M.; Pruess, M.; Rivoire, C.; Sigrist, C.; Sonesson, K.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C.; Chen, Y.; Garavelli, J. S.; Huang, H.; Laiho, K.; Mcgarvey, P.; Natale, D. A.; Ross, K.; Vinayaka, C. R.; Wang, Q.; Wang, Y.; Yeh, L.; Zhang, J.; Ruch, P.; Teodoro, D. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480.
(43) Van Rossom, G. *The Python Library Reference*, 2022; *313*.
(44) Raybaut, P. *Spyder-documentation*, 2021; *5.1*, 5.
(45) Waskom, M. L. Seaborn: statistical data visualization. *JOSS* **2021**, *6*, 3021.
(46) Weaver, S. D.; DeRosa, C. M.; Schultz, S. R.; Champion, M. M. PrIntMap-R: An Online Application for Intraprotein Intensity and Peptide Visualization from Bottom-Up Proteomics. *J. Proteome Res.* **2023**, DOI: 10.1021/acs.jproteome.2c00606.
(47) Tretyakov, K. *Matplotlib-venn Documentation*, 2022; *011.7*.
(48) Hunter, J. D. Matplotlib: a 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.
(49) Mckinney, W. Data Structures for Statistical Computing in Python. *SciPy* **2010**, *445*, 51−56.
(50) Harris, C. R.; Millman, K. J.; Van Der Walt, S. J.; Gommers, J.; Virtanen, R.; Cournapeau, P.; Wieser, D.; Taylor, E.; Berg, J.; Smith, S.; Kern, N. J.; Picus, R.; Hoyer, M.; van Kerkwijk, S.; Brett, M. H.; Haldane, M.; del Río, A.; Wiebe, J. F.; Peterson, M.; Gérard-Marchant, P.; Sheppard, P.; Reddy, K.; Weckesser, T.; Abbasi, W.; Gohlke, H.; Oliphant, C.; Oliphant, T. E. Array programming with NumPy. *Nature* **2020**, *585*, 357.
(51) Hargett, A. A.; Marcella, A. M.; Yu, H.; Li, C.; Orwenyo, J.; Battistel, M. D.; Wang, L.; Freedberg, D. I. Glycosylation States on Intact Proteins Determined by NMR Spectroscopy. *Molecules* **2021**, *26*, 4308.
(52) Dong, L.; Xie, J.; Wang, Y.; Jiang, H.; Chen, K.; Li, D.; Wang, J.; Liu, Y.; He, J.; Zhou, J.; Zhang, L.; Lu, X.; Zou, X.; Wang, X.; Wang, Q.; Chen, Z.; Zuo, D. Mannose ameliorates experimental

colitis by protecting intestinal barrier integrity. *Nat. Commun.* **2022**, *13*, 4804.
(53) Baghdady, Y. Z.; Schug, K. A. Qualitative evaluation of high pH mass spectrometry-compatible reversed phase liquid chromatography for altered selectivity in separations of intact proteins. *J. Chromatogr. A* **2019**, *1599*, 108−114.
(54) 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.* **2009**, *81*, 3933−3943.
(55) Williams, J.; Elleman, T. C.; Kingston, I. B.; Wilkins, A. G.; Kuhn, K. A. The Primary Structure of Hen Ovotransferrin. *Eur. J. Biochem.* **1982**, *122*, 297−303.
(56) Zhang, A.; Sun, H.; Yan, G.; Han, Y.; Wang, X. Serum Proteomics in Biomedical Research: A Systematic Review. *Appl. Biochem. Biotechnol.* **2013**, *170*, 774−786.
(57) Dang, L.; Jia, L.; Zhi, Y.; Li, P.; Zhao, T.; Zhu, B.; Lan, R.; Hu, Y.; Zhang, H.; Sun, S. Mapping human N-linked glycoproteins and glycosylation sites using mass spectrometry. *TrAC, Trends Anal. Chem.* **2019**, *114*, 143−150.
(58) Zhang, H.; Li, X.; Martin, D. B.; Aebersold, R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003**, *21*, 660−666.
(59) Bi, M.; Bai, B.; Tian, Z. Structure-Specific N-Glycoproteomics Characterization of NIST Monoclonal Antibody Reference Material 8671. *J. Proteome Res.* **2022**, *21*, 1276−1284.
(60) Perez-Riverol, Y.; Bai, J.; Bandla, C.; García-Seisdedos, S.; Hewapathirana, D.; Kamatchinathan, S.; Kundu, D.; Prakash, A.; Frericks-Zipper, A.; Eisenacher, M.; Walzer, M.; Wang, S.; Brazma, A.; Vizcaíno, J. A. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **2022**, *50*, D543−D552.

https://doi.org/10.1021/acsomega.2c08071
ACS Omega 2023, 8, 4410−4418