Article

Maximizing Glycoproteomics Results through an Integrated Parallel Accumulation Serial Fragmentation Workflow

Melissa Baerenfaenger,[||] Merel A. Post,[||] Fokje Zijlstra, Alain J. van Gool, Dirk J. Lefeber, and Hans J. C. T. Wessels*

Cite This: *Anal. Chem.* 2024, 96, 8956–8964

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Glycoproteins play important roles in numerous physiological processes and are often implicated in disease. Analysis of site-specific protein glycobiology through glycoproteomics has evolved rapidly in recent years thanks to hardware and software innovations. Particularly, the introduction of parallel accumulation serial fragmentation (PASEF) on hybrid trapped ion mobility time-of-flight mass spectrometry instruments combined deep proteome sequencing with separation of (near-)isobaric precursor ions or converging isotope envelopes through ion mobility separation. However, the reported use of PASEF in integrated glycoproteomics workflows to comprehensively capture the glycoproteome is still limited. To this end, we developed an integrated methodology using timsTOF Pro 2 to enhance N-glycopeptide identifications in complex mixtures. We systematically optimized the ion optics tuning, collision energies, mobility isolation width, and the use of dopant-enriched nitrogen gas (DEN). Thus, we obtained a marked increase in unique glycopeptide identification rates compared to standard proteomics settings, showcasing our results on a large set of glycopeptides. With short liquid chromatography gradients of 30 min, we increased the number of unique N-glycopeptide identifications in human plasma samples from around 100 identifications under standard proteomics conditions to up to 1500 with our optimized glycoproteomics approach, highlighting the need for tailored optimizations to obtain comprehensive data.
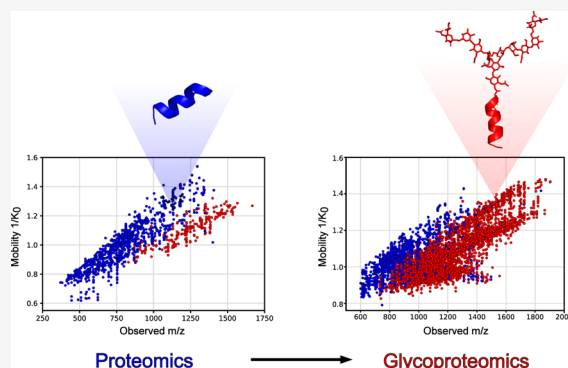
INTRODUCTION

Glycosylation is the most common post-translational modification of human proteins.[1,2] Its implication in all crucial processes of cellular life has resulted in a growing interest to analyze glycosylation to answer biological and clinical questions.[1,3−6] However, studying protein glycosylation is challenging due to the inherent structural complexity caused by micro- and macroheterogeneity and the existence of various glycan isomers.[7−9] As a result, the field of glycoproteomics is lacking behind its bottom-up proteomics counterpart where mass spectrometry (MS) has proven to be a powerful tool for identifying many thousands of proteins in a single LC-MS/MS experiment.[10−12] These advancements were supported by well-established workflows for sample preparation, MS acquisition, and data interpretation for bottom-up proteomics.[13,14] In comparison, the holistic analysis of glycopeptides remains challenging due to the physicochemical properties of glycopeptides and immature technology that lacks standardized workflows. Hence, MS identification rates of glycopeptides from complex mixtures are much lower compared to peptide identifications in bottom-up experiments.[15,16] When compared to bottom-up proteomics, glycoproteomics approaches require a tailored approach that

favors the detection of glycopeptides over peptides. After tryptic digestion of proteins and glycoproteins, the large quantity of peptides in the sample mixture can suppress the detection of less abundant glycopeptides. Therefore, a glycopeptide enrichment can be performed to increase the number of glycopeptides and deplete nonglycosylated peptides.[17,18]

However, further challenges arise in the mass spectrometric detection of glycopeptides. For example, the hydrophilicity of glycopeptides decreases their ionization efficiency in positive ionization mode.[19] To increase ionization efficiency, especially in samples with a large amount of nonglycosylated peptides that could suppress signals of glycopeptides, several strategies have been developed. The use of dopant-enriched nitrogen gas (DEN) has shown to significantly increase desolvation and ionization of hydrophilic analytes such as glycans, glycopep-

tides, or glycoproteins.[20−23] Here, an organic modifier, such as acetonitrile (ACN) or primary alcohols, is used to enrich the nebulizer gas during electrospray ionization, which is not only favorable for the signal intensity of glycopeptides but can also influence their charge state distributions. The use of acetonitrile as DEN has shown to increase the charge of glycopeptides from 2+ to 3+ toward 3+ to 4+ or higher in combination with reversed phase nano-LC-MS.[20] This is a significant advantage, as higher charge states of glycopeptides also address another obstacle in the MS/MS detection of glycopeptides. Similar to peptides, higher charge states of precursor ions provide better MS/MS fragmentation yields at lower collision energies (CE) during collision induced dissociation (CID).[24] This is essential as both the glycan moiety and the peptide moiety of glycopeptides need to yield sufficient fragment ions for an unambiguous characterization. When proteomics-like CID conditions are employed on glycopeptides with low charge states, typically only glycan fragmentation is observed. Fragmentation of the peptide bond rarely yields sufficient b- and y-ions for unambiguous characterization of the peptide moiety. Potential solutions are the use of advanced fragmentation techniques like electron-based dissociation (ExD) that produce c- and z-type peptide fragment ions while retaining the intact glycan moiety.[25,26] However, these techniques are not widely available. More commonly used CID instruments can overcome this obstacle by using higher collision energies to increase the yield of peptide fragment ions or performing CID stepping to fragment the glycan moiety at lower collision energies and the peptide moiety at high collision energies.[24,27]

Another advancement for the holistic detection of glycopeptides in mixtures is the use of ion mobility mass spectrometry.[28] Ion mobility is able to separate ions based on their size and shape in the gas phase and can provide an extra dimension of separation. One ion mobility technique is trapped ion mobility spectrometry (TIMS) in which ions are separated by finding an equilibrium between a constant gas flow and an electric field that allows them to be stored at different spatial positions inside the TIMS funnel. Ions can subsequently be eluted by lowering the electric field potential. In the timsTOF Pro and newer models (Bruker Daltonics), this technology enables parallel accumulation serial fragmentation (PASEF) for fast and sensitive fragmentation of orthogonally isolated precursor ions by mass-to-charge (m/z) and ion mobility. For the field of proteomics, this technology has already demonstrated its use by dramatically increasing the number of peptide identifications in bottom-up proteomics setups.[29] Interestingly, the orthogonal precursor selection allows to focus the MS/MS spectra generation on glycopeptides rather than peptides as size and shape generally differs.[28] Although these advancements in the detection of glycopeptides have been made, a combined methodology leveraging them in one workflow is missing. Hence, we established an integrated glycoproteomics workflow on the timsTOF Pro 2 that utilizes PASEF as well as the increased ionization efficiency of glycopeptides with DEN. We developed our glycoproteomics method starting from default settings for bottom-up proteomics and compared the performance between glycopeptide and peptide identification to identify crucial parameters for glycoproteomic measurements.

We optimized electrospray conditions, ion optics, MS/MS precursor selection, and collision energies to increase glycopeptide identification for enriched glycopeptides from tryptic human plasma digests. Our optimized method enabled identification of ∼1100−1500 glycopeptides using LC gradient times as short as 15 to 30 min, respectively, which enables comprehensive glycoproteome coverage at considerable sample throughput for (pre)clinical applications.

■ EXPERIMENTAL SECTION

Tryptic Digestion and Glycopeptide Enrichment from Human Plasma Samples. Plasma samples of healthy donors were obtained from the Sanquin blood bank (Nijmegen, Netherlands) according to their protocols of informed consent. Samples from 5 individuals were pooled and preparation was performed as described previously.[4] 10 μL of human plasma was denatured with 10 μL urea (8 M urea, 10 mM Tris-HCl, pH 8.0) and reduced with 15 μL of 10 mM dithiothreitol for 30 min at room temperature. Alkylation was performed by the addition of 15 μL of 50 mM 2-chloroacetamide and incubation in the dark for 20 min at room temperature. Proteins were first digested with LysC (1 μg LysC/50 μg protein) for 3 h at room temperature. Subsequently, samples were diluted with three volumes of a 50 mM ammonium bicarbonate buffer. A tryptic digest was performed overnight at 37 °C by the addition of 1 μg of trypsin per 50 μg of protein. Glycopeptides were enriched using Sepharose CL-4B beads (Merck). 100 μL of slurry was added to a well of a 0.20 μm pore size 96 multiwell filter plate (AcroPrep Advance, VWR). The beads were washed three times with 20% ethanol and 83% acetonitrile. After the digested sample was applied, the plate was incubated for 20 min at room temperature while shaking. The filter plate was centrifuged to remove the supernatant, and beads were washed three times with 83% ACN and three times with 83% ACN with 0.1% trifluoroacetic acid (TFA). Glycopeptides were eluted with 50 μL of water.

LC-MS/MS Analysis. The samples were analyzed by employing a nanoElute liquid chromatography system (Bruker Daltonics) connected to a timsTOF Pro 2 instrument (Bruker Daltonics). A CaptiveSprayer nanoflow electrospray ionization source was used either with dopant-enriched nitrogen gas via the nanobooster or without DEN. Separation of peptides and glycopeptides was achieved on an Elute Fifteen C18 reversed-phase column (0.075 mm ID × 150 mm length, 1.9 μm particles, 120 Å pore size) operated at 45 °C. The elution gradient for most optimization steps consisted of a linear increase from 7% to 45% acetonitrile in 0.1% formic acid (FA) and 0.02% TFA over 15 min, with a flow rate of 500 nL/min. More information on the used gradients can be found in Table S1. Mass spectrometry measurements were conducted in positive ionization mode with a capillary voltage of 1500 V and, if used, a nanobooster gas pressure of 0.2 bar of N$_2$. MS conditions were optimized starting from default setting for proteomics measurements provided by Bruker Daltonics and adjusted to facilitate glycopeptide identification. This proteomics method is a data-dependent acquisition with 1.1 s duty cycle time and is used for peptide identifications under standard proteomics conditions. The MS conditions of the proteomics method and optimized glycoproteomics methods can be found in Table S2. All measurements were performed in duplicate.

Optimization Strategy. We optimized several acquisition parameters to establish an integrated glycoproteomics method on timsTOF Pro 2. First, we optimized the collision energies to achieve ideal MS/MS fragmentation. Subsequently, we optimized the ion optics setting to enhance transmission of

glycopeptides. PASEF conditions, namely, the target intensity, TIMS isolation width (also referred to as measuring time), and the polygon region for PASEF precursor selection were adjusted. Finally, we evaluated the effect of different LC gradients and benchmarked our optimized method with and without the use of DEN. Also, overview of all used MS acquisition parameters for each optimization step can be found in Table S3. All MS data is deposited in ProteomeXchange repository under the identifier PXD047898.[30]

Data Interpretation. Glycopeptides were identified using MSFragger Glyco. MSFragger searches were performed using fragpipe 15.0, MSFragger 3.4, and philosopher 4.1.1. The glyco-N-HCD search parameters were: 20 ppm mass tolerance with an isotope error of 0−2, semitryptic enzyme specificity, peptide length 5−50 amino acids, and 600−20,000 m/z range. Carbamidomethylation at cysteine residues was set as fixed modification, whereas methionine oxidation and N-terminal ammonia loss were set as variable modifications. Human secreted protein reference sequence database was downloaded from Uniprot (containing 4029 entries, downloaded on 2021.11.22) and glycan mass offsets were extracted for unique compositions in the GlyGen glycan reference database (containing 475 glycan compositions, downloaded on 2022.22.04).[31,32] For the final (glyco)peptide sequence matches (PSM), identified peptides, glycans, and proteins were filtered to 1% FDR using a sequential filtering step to remove PSMs, glycans, and peptides from proteins that did not pass FDR criteria. This FDR filtering has been proposed previously to robustly determine the glycan composition of N-glycopeptides from tandem MS data.[33] DataAnalysis version 5.3 (Bruker Daltonics) was used for raw data analysis. The software MSFragger Glyco identifies glycan compositions based on mass offsets. With this approach, a high fragmentation of the peptide moiety yields optimal results. Other software tools that take glycan fragmentation into account can benefit from collision energy stepping.[24,25,34] We therefore provide the optimized instrument method both with and without TIMS stepping in the ProteomeXchange repository under identifier PXD047898.

Data was visualized using Python 3.10.10 with the packages Matplotlib 3.7.1, Scipy 1.10.1, Numpy 1.24.3, and Pandas 1.5.3. All data points represent the average of two replicate measurements. Peptide and glycopeptide images were created using PyMOL version 2.5.5.

■ RESULTS AND DISCUSSION

To develop a generic analytical PASEF workflow for holistic glycoproteomics, we used enriched glycopeptides from human plasma samples. Although our methodology can be applied to various sample types, enriched glycopeptides from human plasma not only represent a complex mixture with heterogeneous glycopeptides but are also biologically and clinically relevant. The enriched glycopeptides were subjected to LC-MS/MS measurements, starting with the default parameters for proteomics on the timsTOF platforms (for proteomics parameters, see Table S2). We subsequently optimized collision energies for glycopeptide fragmentation, electrospray source conditions and ion optics, PASEF settings, and gradient length. To verify the outcome of our optimization steps, we evaluated results based on the number of identified glycopeptides in relation to the number of peptide identifications. MSFragger Glyco software was used for (glyco)peptide identification, which identifies the peptide

moiety of glycopeptides using a modified open mass search strategy whereas glycan moiety compositions are identified based on mass offset and limited glycan fragmentation data in combination with reference glycan compositions.[35] Finally, we benchmarked our optimized method against the standard proteomics method using acetonitrile-enriched nitrogen gas to enhance the ionization efficiency of the glycopeptides. For all of our optimization steps, we focused on increasing the number of unique glycopeptide identifications. Additionally, we also show the effect of our method on peptide identifications. A schematic overview of the instrument configuration and conceptual optimization steps is shown in Figure 1.
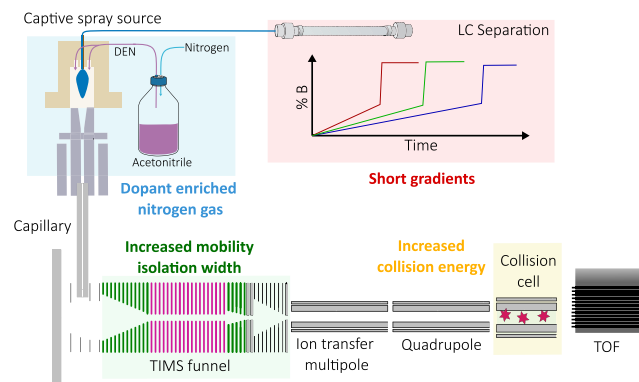

Figure 1. Optimization scheme. Our methodology integrates the advantages of DEN, optimized PASEF settings with ideal mobility isolation width and collision energies to increase glycopeptide identification with short LC run times of 15 to 30 min.

Optimizing CID Energies. The identification of glycopeptides by tandem mass spectrometry poses a significant challenge due to the need for optimal fragmentation of both the peptide moiety and the intrinsically different glycan moiety. Conventional shotgun proteomics CID energies typically yield predominant glycan fragmentation with insufficient peptide fragment ions to confidently identify the glycopeptide.[36] To address this limitation, we systematically adjusted collision energies to enhance the yield of the peptide backbone fragments in our approach.

The timsTOF Pro 2 instruments enable collision energy scaling based on precursor ion mobility values. At lower ion mobility values, higher charge states and smaller glycopeptides are expected, requiring less energy to yield the peptide backbone fragments. Conversely, higher mobility values indicated lower charge states and larger glycopeptides. Starting from the default proteomics collision energy slope, we incrementally increased the slope by steps of +10% up to +60% as seen in Figure 2A. An almost linear increase in the number of unique glycopeptide identifications with increasing collision energies was observed (Figure 2B). In contrast, the number of peptide identifications decreased at higher collision energies. Furthermore, the optimal collision energy values were found to be dependent on the precursor charge state. For example, glycopeptides with a charge state of +5 did not benefit from increasing the collision energies above +30% as illustrated in Figure 2C. As our final optimized glycoproteomics method employs acetonitrile-enriched nitrogen gas, resulting in an overall charge state increase of glycopeptides, we did not increase collision energies any further and
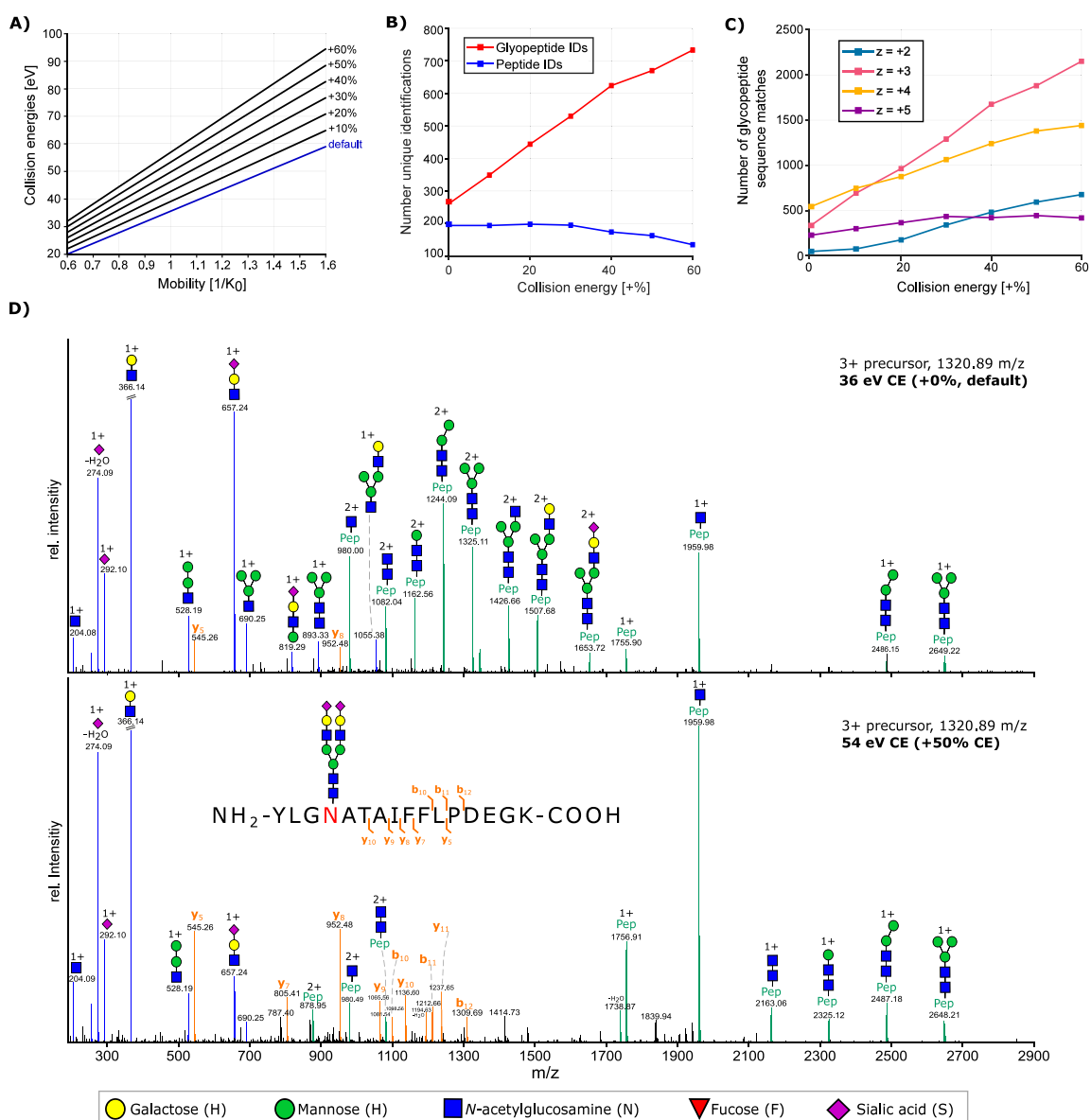
Figure 2. Collision energy optimization. (A) Collision energies are set as a function of ion mobility and increase linearly with higher mobility values. The slope and offset were increased in steps of 10% compared to the proteomics default settings to optimize collision energies for glycopeptide fragmentation. (B) Increasing the collision energies led to an increase in unique glycopeptide identifications from less than 300 to over 700 at higher CE values while peptide identifications decreased. (C) Ideal CE values depended on precursor charge state. Precursors with a lower charge state (+2, +3) benefited more from increased CE values compared to highly charge precursors (+4, +5). (D) Exemplary MS/MS spectrum of a +3 glycopeptide with standard proteomics CE values shows mainly the fragmentation of the glycan moiety as oxonium ions (blue) or glycan fragments with intact peptide residue (green). At higher CE values (+50%), more b- and y-ion fragments of the peptide backbone (orange) are observed to support peptide moiety identification by protein sequence database search algorithms. The annotation of glycan fragments represents only one possibility. Other isomers are plausible.

continued our optimization steps with a collision energy increase of +50% compared to default proteomics settings. Additionally, it is worth noting that collision energies in most timsTOF Pro instruments are limited to 100 eV to avoid arcing. This technical limitation prevented further optimization of glycopeptide fragmentation at even higher collision energies.

Improving Ion Transmission. Optimizing the collision energies dramatically increased the number of unique glycopeptide identifications from under 300 to over 700 identifications. To further increase identifications, we continued with optimizing ion optics, namely, the prepulse storage and transfer time, which are responsible for the transmission of ions from the collision cell to the TOF orthogonal accelerator.

Furthermore, we optimized the collision cell RF. To determine a plausible range for glycopeptide optimization, we first optimized transmission of tunemix signals (Agilent technologies), covering a wide m/z range via direct infusion experiments (Figure S1). For glycopeptide identification, the transmission of low m/z ions as well as higher m/z ions is necessary to allow the detection of high-mass glycopeptide precursor ions as well as low-mass oxonium ions after fragmentation. Hence, we chose to optimize glycopeptide identification with prepulse storage times, transfer times, and collision cell RF values that allow the detection of all these ions. Ramping of the prepulse storage and transfer time did not result in a strong increase of glycopeptide identifications over
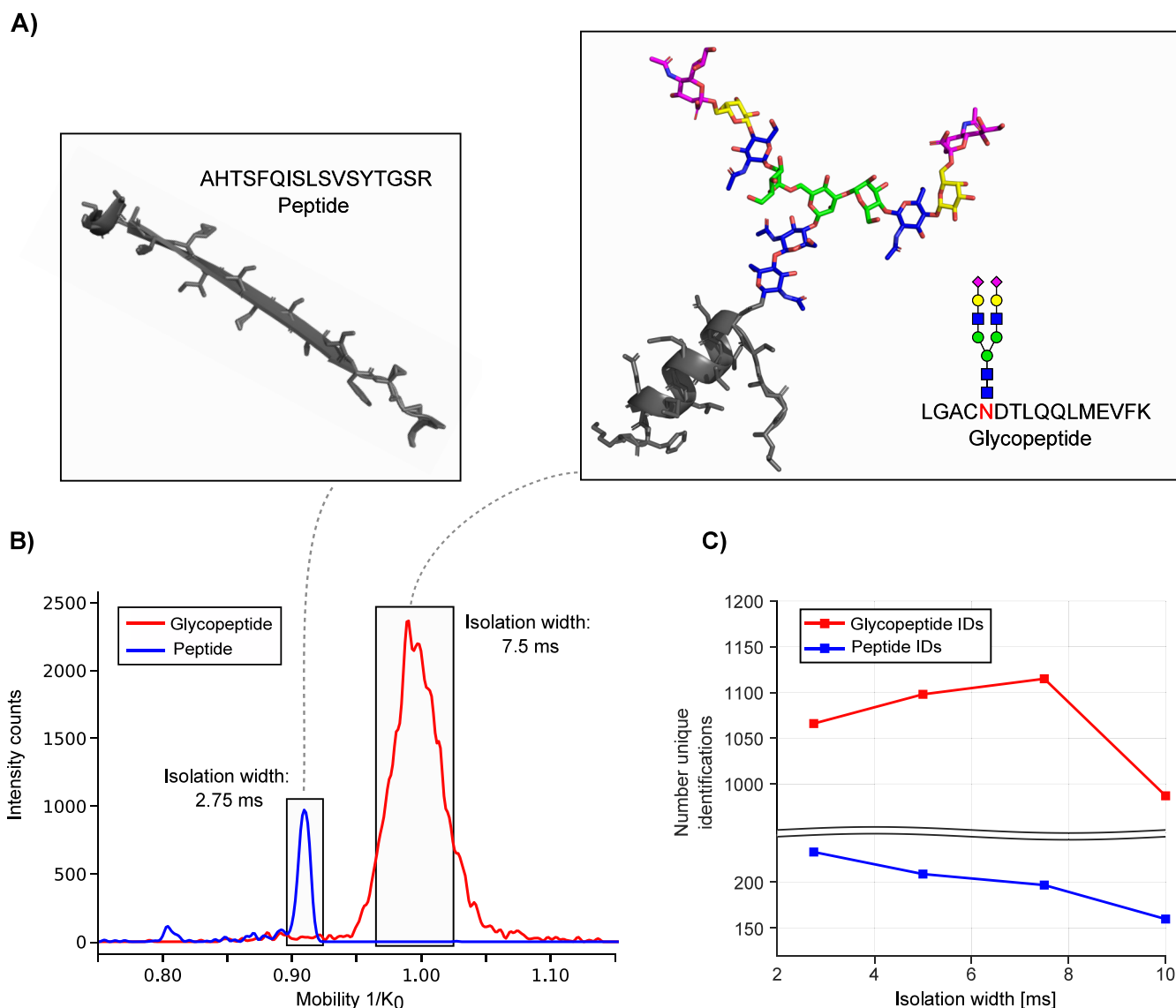
Figure 3. TIMS isolation width optimization. (A) Structure of one selected peptide (pdb entry 7VON, alpha-2-macroglobulin) and one selected glycopeptide (pdb entry 1ATH with an attached glycan, human antithrombin III). (B) Mobilogram of the selected peptide (blue, sequence AHTSFQISLSVSYTGSR) and glycopeptide (red, sequence LGAC**N**DTLQQLMEVFK decorated with glycan composition H5N4S2). The standard TIMS isolation width of 2.75 ms is ideal for the isolation of the peptide; however, a much larger isolation with of 7.5 ms is needed for the glycopeptide due to a broader mobility distribution. (C) Number of peptide and glycopeptide identifications with increasing TIMS isolation width.

peptide identifications (Figure S2). In contrast, we found that the collision cell RF has a strong influence on glycopeptide identification and that increased Vpp values promote the identification of glycopeptides while decreasing the identification of peptides (Figure S2). With increased RF Vpp, the kinetic energy of ions is increased, enabling them to maintain their trajectory through the collision cell, thus improving the transmission of ions with higher m/z values.[37,38] This is crucial as glycopeptides show y-ion fragments at high m/z values due to fragment ions containing the whole peptide backbone and different glycan fragments at reduced charge states. For glycopeptide identification, determining the correct peptide mass based on the Y0 "peptide-only" or Y1 "peptide+HexNAc" fragment is essential (see annotation of MS/MS spectra in Figure 2D).[24,35] These Y-ions are observed at higher m/z values compared to that of peptide fragment ions. Hence, we increased the collision cell RF from 1500 to 1700 eV. We chose not to increase the collision cell RF any further as it

would result in the loss of diagnostic oxonium ions from glycan fragmentation at lower m/z values (see Figure S1), which can be relevant for certain annotation tools.

Optimizing PASEF Settings. One of the key features of the timsTOF platform, which provides high sensitivity and fast MS/MS acquisition with simultaneous ion mobility separation, is parallel accumulation and serial fragmentation (PASEF) of ions.[39,40] After drastically increasing the glycopeptide identification by adjusting collision energies, we evaluated PASEF parameters for optimal MS/MS data generation. First, we tested the effect of the target intensity for MS/MS acquisition, as we expected that higher target intensities would lead to an increased intensity and signal-to-noise ratio of peptide backbone ions. However, increasing the target intensity in a range from 20,000 counts to 150,000 counts did not result in an increase of glycopeptide identifications (see Figure S3). Next, we focused on the TIMS isolation width as PASEF precursor ion selection is based on m/z values and mobility. As

ions are sequentially released from the TIMS funnel, the quadrupole is isolating precursor ions by quickly switching between different m/z positions while achieving MS/MS acquisition rates of up to 110 Hz in timsTOF Pro 2. Considering typical ion mobility peak widths of peptides, the quadrupole isolation time is set between 2.5 and 4.0 ms, which can theoretically result in the selection of 240 to 400 precursor ions per second.[40] For our measurements, we chose a ramp time of 100 ms with a 100% duty cycle. In the default proteomics method, the mobility range is set from 0.6 to 1.6 1/K_0 [V·s/cm^2] and the TIMS isolation width is set to 2.75 ms. This results in a mobility isolation window of ± 0.028 1/K_0 [V·s/cm^2]. Isolation widths of 2.75 ms are ideal for peptides, where typical half-widths of 1 ms are observed under similar conditions.[40] However, this isolation width might not be ideal for glycopeptides as the glycan moiety not only consist of complex isomers but also multiple conformations for each isomer.[41] Hence, we tested TIMS isolation widths of 2.75, 5.00, 7.50, and 10.0 ms to determine the effect on glycopeptide identification rates. Figure 3A shows the structure of one selected peptide and one selected glycopeptide to demonstrate typical differences in size and also rotatable bonds in peptides and glycopeptides. Mobilograms for both the peptide and glycopeptide are shown in Figure 3B. The TIMS isolation width of 2.75 ms is sufficient to capture the entire mobility peak for the peptide; however, a broader isolation width is needed for glycopeptides to avoid ion loss during MS/MS isolation. When the isolation width was ramped from 2.75 to 10 ms, the number of unique glycopeptide identifications increased with isolation widths up to 7.5 ms before decreasing at 10 ms (Figure 3C). For peptides, an increase in isolation width decreases the number of identifications. This could be due to a reduction in MS/MS detection of unique precursor ions, as fewer precursor ions are selected per TIMS ramp.

In addition to increasing the TIMS isolation width, we specifically selected a mass and mobility region that favored the selection of glycopeptides over peptides as the MS/MS precursor. As demonstrated before, glycopeptides show a distinct difference in mass and ion mobility and selecting this region of interest by defining a polygon for MS/MS precursor selection can increase the number of MS/MS spectra of glycopeptides.[28] The polygon used in the proteomics default setting and our revised polygon can be found in Figure S4.

Influence of LC-Gradient Length. High sample throughput is not only a crucial parameter in bottom-up proteomics but also essential for glycoproteomics to enable experiments with large sample cohorts or routine testing in clinical setups. To achieve higher throughput, liquid chromatography gradients can be reduced if the peak capacity is still sufficient and MS/MS sampling is fast.[42] The timsTOF Pro 2 offers additional separation power in the mobility dimension as well as fast MS/MS acquisition rates to sample LC peaks with small peak widths, therefore facilitating high identification rates with short LC gradients. To determine how gradient length affects glycopeptide identification on our 15 cm reversed-phase nanoflow column, we tested different gradient lengths from 5 to 60 min (Figure 4). Information on the gradient conditions can be found in Table S1.

Unique glycopeptide identifications drastically increased when increasing between gradient times of 5 to 15 min from over 300 to almost 1200 identifications. Longer gradient times of 30 min still showed an improvement in glycopeptide identifications; however, the number of identifications
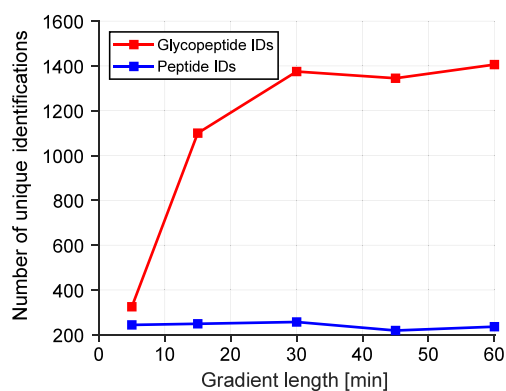

Figure 4. Influence of gradient length on the number of unique peptides (blue) and glycopeptides (red).

stagnated with gradient times over 30 min. Hence, for the following measurements, a gradient length of 30 min was selected.

Benchmarking of Optimized Conditions under Use of Dopant-Enriched Nitrogen Gas. After combining optimized parameters, we benchmarked our glycoproteomics method against the bottom-up proteomics default method with and without the use of DEN. As reported previously, the use of DEN can drastically increase ionization efficiencies of polar compounds such as glycans, glycoproteins, and glycopeptides.[20−23] Hence, we evaluated how much our use of acetonitrile as DEN influences glycopeptide identification rates under proteomics settings, as well as with optimized glycoproteomics settings. As expected, the use of DEN strongly influences the number of glycopeptide identifications for both the default proteomics method and the developed glycoproteomics method as seen in Figure 5.
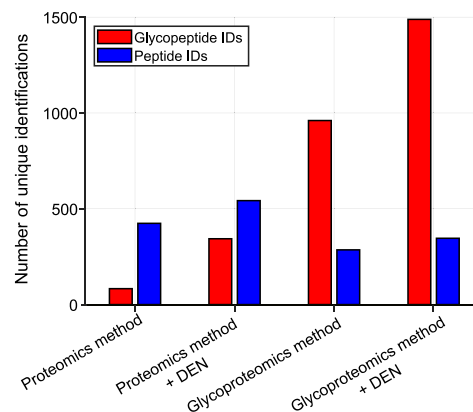

Figure 5. Benchmarking of glycoproteomics method against default proteomics settings employing a 30 min gradient with and without the use of DEN.

We observed that not only glycopeptide identifications increase with the use of DEN, but also the number of peptide identifications. Still, the use of DEN seems to have a stronger effect on glycopeptides compared to peptides, as shown by the 4-fold increase in glycopeptide identifications under proteomics conditions and the increase in identifications from 960 to almost 1500 under glycoproteomics conditions (Figure 5).

It is worth mentioning that when DEN was used, the mobility values for peptides and glycopeptides experienced a small shift toward higher mobility, which can be seen in Figure
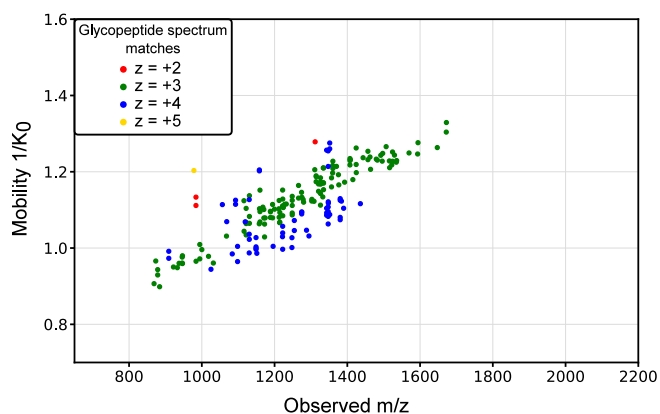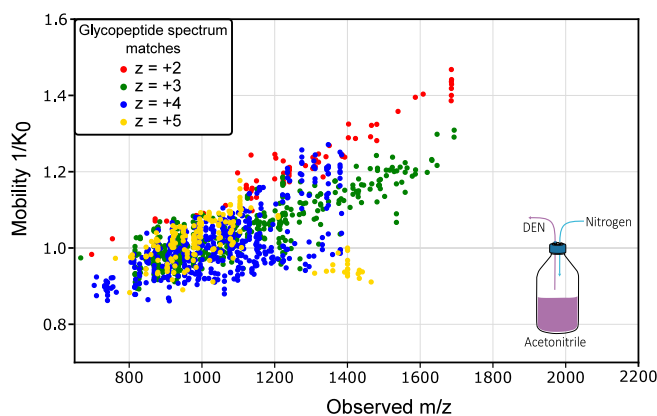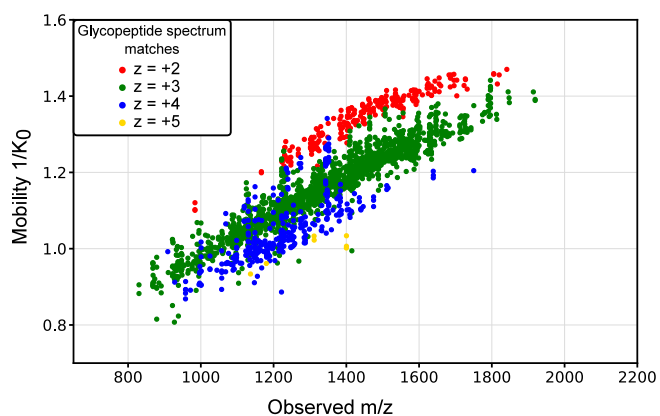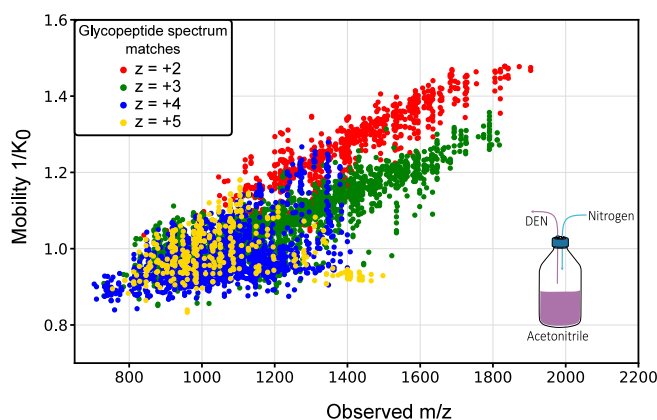
Figure 6. Charge state distribution of the glycopeptide sequence matches with and without the use of DEN. (A) Using standard proteomics condition without DEN results in a low number of glycopeptide spectrum matches. (C) Optimized glycoproteomics method drastically increases the number of glycopeptide spectrum matches. For both conditions, the dominant charge state is +3. (B, D) When using DEN, the number of glycopeptide spectrum matches drastically increases for both methods. In addition, more ions with a charge state of +4 and +5 are observed.

S4. To evaluate the significance of shifting mobility values with the use of DEN, we measured the shift in mobility for one calibrant ion with different nanobooster fill levels as seen in Figure S5. A slight shift in mobility was observed; however, a regular refilling of the nanobooster with acetonitrile enabled us to obtain stable results during glycoproteomics measurements.

In addition, we were interested to see how charge states are influenced by using DEN. Hence, we plotted the glycopeptide spectrum matches per charge state in the m/z and mobility dimension for all four conditions (Figure 6). Without the use of DEN, we dominantly observe that the glycopeptide spectrum matches with a charge state of +3. This distribution shifts toward an almost equal occurrence of charge states of +3 and +4 with DEN both under default proteomics conditions and when using the optimized glycoproteomics method (Figure 6 and Figure S6). Remarkably, the number of glycopeptide spectrum matches with a charge state of +2 is low when using the default proteomics method with and without DEN (Figure 6). We conclude that this is caused mainly by the low collision energies of the proteomics method, which do not allow sufficient fragmentation of glycopeptide ions with low charge states (see also Figure 2C).

Furthermore, we were interested to see if our methodology is biased against certain glycan compositions and plotted the number of glycopeptide sequence matches per glycan class, number of neuraminic acids, and number of fucoses for the

proteomics method and the optimized glycoproteomics method with and without the use of DEN (Figure S7). In addition, we plotted the number of identified glycopeptide sequence matches for the most abundant glycan compositions for the default proteomics and optimized glycoproteomics method with the use of DEN (Figure S8). We found that the standard proteomics method without DEN shows a slight bias toward high-mannose type glycans, whereas the optimized glycoproteomics method with DEN yields an increased number of sialylated complex-type glycans. Although these results are more in line with previously published data on the human plasma glycoproteome, it must be kept in mind that we did not perform a relative quantification of glycopeptides, which makes a direct comparison with existing data on the human glycoproteome challanging.[43] Full tables with all glycopeptide identifications for all conditions can be found in the ProteomeXchange repository under the identifier PXD047898.

■ CONCLUSIONS

In this work, we developed an optimal integrated PASEF workflow for glycoproteomics measurements. By optimizing multiple parameters, including collision energies (CE), ion optics, and ion mobility settings, we established a robust workflow for the holistic identification of glycopeptides. Our results show that the use of dopant-enriched nitrogen (DEN)

gas and optimization of collision energies have a significant impact on glycopeptide identification rates. Unlike most approaches that optimize collision energies or report the benefits of DEN on a limited number of glycopeptides, our study provides a broader perspective by evaluating a larger number of glycopeptides.[20,28] However, it must be considered that the ideal collision energy values reported here are advantageous for the used glycan offset search strategy and that other data interpretation platforms may require slightly different settings to target the glycan moiety of glycopeptides.

By comparing peptide and glycopeptide identification rates, we demonstrate the need to adjust parameters based on the molecular characteristics of the analytes. This emphasizes the inherent complexity of glycoproteomics and underlines that a one-size-fits-all approach is not applicable to obtaining optimal results. Our adjusted methodology, which integrates optimized collision energies, PASEF settings, and DEN, provides a high-throughput approach suitable for glycopeptide analysis in complex biological samples on the widely used timsTOF platform. Thus, this work contributes to the growing field of glycoproteomics and paves the way for improved characterization of post-translational modifications in complex biological samples, enabling broad applications in both research and (pre-)clinical settings.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.3c05874.

> Ion optics tuning; tuning of ion optics for ideal transmission and identification of glycopeptides; optimization of target intensity; glycopeptide and peptide IDs plotted by m/z vs mobility value for proteomics method; change in mobility of one agilent tunemix; charge state distribution of glycopeptides; grouped glycopeptide sequence matches for the default proteomics method; identified glycopeptide sequence matches for proteomics method and optimized glycoproteomics method; Table S1: tested nanoLC gradients; Table S2: settings of the default proteomics method and optimized parameters for the developed glycoproteomics method; Table S3: optimization scheme (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Hans J. C. T. Wessels − *Translational Metabolic Laboratory, Department of Human Genetics, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands*; Email: Hans.Wessels@radboudumc.nl

Authors

Melissa Baerenfaenger − *Department of Neurology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands; Division of BioAnalytical Chemistry, AIMMS Amsterdam Institute of Molecular and Life Sciences, Vrije Universiteit Amsterdam, Amsterdam 1081 HZ, Netherlands;* ⓞ orcid.org/0000-0002-2855-924X

Merel A. Post − *Department of Neurology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands*

Fokje Zijlstra − *Translational Metabolic Laboratory, Department of Human Genetics, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands*

Alain J. van Gool − *Translational Metabolic Laboratory, Department of Human Genetics, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands*

Dirk J. Lefeber − *Department of Neurology, Donders Institute for Brain, Cognition, and Behavior and Translational Metabolic Laboratory, Department of Human Genetics, Radboud University Medical Center, Nijmegen 6525 GA, Netherlands*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.analchem.3c05874

Author Contributions

‖M.B. and M.A.P. contributed equally to this work and share first authorship

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Schjoldager, K. T.; Narimatsu, Y.; Joshi, H. J.; Clausen, H. *Nat. Rev. Mol. Cell Biol.* **2020**, *21* (12), 729−749.

(2) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochimica et biophysica acta* **1999**, *1473* (1), 4−8.

(3) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. *Nature Reviews Nephrology* **2019**, *15* (6), 346−366.

(4) Wessels, H. J. C. T.; Kulkarni, P.; van Dael, M.; Suppers, A.; Willems, E.; Zijlstra, F.; Kragt, E.; Gloerich, J.; Schmit, P.-O.; Pengelley, S.; Marx, K.; van Gool, A. J.; Lefeber, D. J., Plasma glycoproteomics delivers high-specificity disease biomarkers by detecting site-specific glycosylation abnormalities. *bioRxiv* **2022**, DOI: 10.1101/2022.05.31.494121.

(5) Baerenfaenger, M.; Post, M. A.; Langerhorst, P.; Huijben, K.; Zijlstra, F.; Jacobs, J. F. M.; Verbeek, M. M.; Wessels, H. J. C. T.; Lefeber, D. J. *International Journal of Molecular Sciences* **2023**, *24* (3), 1937.

(6) Suttapitugsakul, S.; Sun, F.; Wu, R. *Anal. Chem.* **2020**, *92* (1), 267−291.

(7) Čaval, T.; Heck, A. J. R.; Reiding, K. R. *Molecular & Cellular Proteomics* **2021**, *20*, 100010.

(8) Liew, C. Y.; Yen, C.-C.; Chen, J.-L.; Tsai, S.-T.; Pawar, S.; Wu, C.-Y.; Ni, C.-K. *Commun. Chem.* **2021**, *4* (1), 92.

(9) Chen, Z.; Glover, M. S.; Li, L. *Curr. Opin. Chem. Biol.* **2018**, *42*, 1−8.

(10) Deutsch, E. W.; Omenn, G. S.; Sun, Z.; Maes, M.; Pernemalm, M.; Palaniappan, K. K.; Letunica, N.; Vandenbrouck, Y.; Brun, V.; Tao, S.-c.; Yu, X.; Geyer, P. E.; Ignjatovic, V.; Moritz, R. L.; Schwenk, J. M. *J. Proteome Res.* **2021**, *20* (12), 5241−5263.

(11) Müller, J. B.; Geyer, P. E.; Colaço, A. R.; Treit, P. V.; Strauss, M. T.; Oroshi, M.; Doll, S.; Virreira Winter, S.; Bader, J. M.; Köhler, N.; Theis, F.; Santos, A.; Mann, M. *Nature* **2020**, *582* (7813), 592−596.

(12) Thomas, D. R.; Scott, N. E. *Curr. Opin. Struct. Biol.* **2021**, *68*, 18−25.

(13) Duong, V.-A.; Lee, H. *International Journal of Molecular Sciences* **2023**, *24* (6), 5350.

(14) O'Rourke, M. B.; Town, S. E. L.; Dalla, P. V.; Bicknell, F.; Koh Belic, N.; Violi, J. P.; Steele, J. R.; Padula, M. P. *Proteomes* **2019**, *7* (3), 29.

(15) Yang, Y.; Franc, V.; Heck, A. J. R. *Trends Biotechnol.* **2017**, *35* (7), 598−609.

(16) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. *Journal of Chromatography B* **2007**, *849* (1), 115−128.

(17) Li, J.; Zhang, J.; Xu, M.; Yang, Z.; Yue, S.; Zhou, W.; Gui, C.; Zhang, H.; Li, S.; Wang, P. G.; Yang, S. *J. Sep. Sci.* **2022**, *45* (16), 3169−3186.

(18) Riley, N. M.; Bertozzi, C. R.; Pitteri, S. J. *Mol. Cell. Proteomics* **2021**, *20*, 100029.

(19) Stavenhagen, K.; Hinneburg, H.; Thaysen-Andersen, M.; Hartmann, L.; Silva, D. V.; Fuchser, J.; Kaspar, S.; Rapp, E.; Seeberger, P. H.; Kolarich, D. *J. Mass Spectrom.* **2013**, *48* (6), 627.

(20) Alagesan, K.; Kolarich, D. *bioRxiv* **2019**, No. 597922.

(21) Kammeijer, G. S.; Kohler, I.; Jansen, B. C.; Hensbergen, P. J.; Mayboroda, O. A.; Falck, D.; Wuhrer, M. *Anal. Chem.* **2016**, *88* (11), 5849−56.

(22) Madunić, K.; Wagt, S.; Zhang, T.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M. *Anal. Chem.* **2021**, *93* (18), 6919−6923.

(23) van Schaick, G.; Wuhrer, M.; Domínguez-Vega, E. *Anal. Chim. Acta* **2023**, *1265*, No. 341271.

(24) Hinneburg, H.; Stavenhagen, K.; Schweiger-Hufnagel, U.; Pengelley, S.; Jabs, W.; Seeberger, P. H.; Silva, D. V.; Wuhrer, M.; Kolarich, D. *Journal of The American Society for Mass Spectrometry* **2016**, *27* (3), 507−519.

(25) Riley, N. M.; Malaker, S. A.; Driessen, M. D.; Bertozzi, C. R. *J. Proteome Res.* **2020**, *19* (8), 3286−3301.

(26) Alagesan, K.; Hinneburg, H.; Seeberger, P. H.; Silva, D. V.; Kolarich, D. *Glycoconjugate Journal* **2019**, *36* (6), 487−493.

(27) Hoffmann, M.; Pioch, M.; Pralow, A.; Hennig, R.; Kottler, R.; Reichl, U.; Rapp, E. *Proteomics* **2018**, *18* (24), e1800282.

(28) Mukherjee, S.; Jankevics, A.; Busch, F.; Lubeck, M.; Zou, Y.; Kruppa, G.; Heck, A. J. R.; Scheltema, R. A.; Reiding, K. R. *Molecular & cellular proteomics: MCP* **2023**, *22* (2), No. 100486.

(29) Meier, F.; Park, M. A.; Mann, M. *Mol. Cell. Proteomics* **2021**, *20*, No. 100138.

(30) Deutsch, E. W.; Bandeira, N.; Perez-Riverol, Y.; Sharma, V.; Carver, J. J.; Mendoza, L.; Kundu, D. J.; Wang, S.; Bandla, C.; Kamatchinathan, S.; Hewapathirana, S.; Pullman, B. S.; Wertz, J.; Sun, Z.; Kawano, S.; Okuda, S.; Watanabe, Y.; MacLean, B.; MacCoss, M. J.; Zhu, Y.; Ishihama, Y.; Vizcaíno, J. A. *Nucleic Acids Res.* **2022**, *51* (D1), D1539−D1548.

(31) The UniProt Consotium. *Nucleic Acids Res.* **2023**, *51* (D1), D523−D531.

(32) York, W. S.; Mazumder, R.; Ranzinger, R.; Edwards, N.; Kahsay, R.; Aoki-Kinoshita, K. F.; Campbell, M. P.; Cummings, R. D.; Feizi, T.; Martin, M.; Natale, D. A.; Packer, N. H.; Woods, R. J.; Agarwal, G.; Arpinar, S.; Bhat, S.; Blake, J.; Castro, L. J. G.; Fochtman, B.; Gildersleeve, J.; Goldman, R.; Holmes, X.; Jain, V.; Kulkarni, S.; Mahadik, R.; Mehta, A.; Mousavi, R.; Nakarakommula, S.; Navelkar, R.; Pattabiraman, N.; Pierce, M. J.; Ross, K.; Vasudev, P.; Vora, J.; Williamson, T.; Zhang, W. *Glycobiology* **2020**, *30* (2), 72−73.

(33) Polasky, D. A.; Geiszler, D. J.; Yu, F.; Nesvizhskii, A. I. *Molecular & Cellular Proteomics* **2022**, *21* (3), No. 100205.

(34) Liu, M.-Q.; Zeng, W.-F.; Fang, P.; Cao, W.-Q.; Liu, C.; Yan, G.-Q.; Zhang, Y.; Peng, C.; Wu, J.-Q.; Zhang, X.-J.; Tu, H.-J.; Chi, H.; Sun, R.-X.; Cao, Y.; Dong, M.-Q.; Jiang, B.-Y.; Huang, J.-M.; Shen, H.-L.; Wong, C. C. L.; He, S.-M.; Yang, P.-Y. *Nat. Commun.* **2017**, *8* (1), 438.

(35) Polasky, D. A.; Yu, F.; Teo, G. C.; Nesvizhskii, A. I. *Nat. Methods* **2020**, *17* (11), 1125−1132.

(36) Janiszewska, E.; Kmieciak, A.; Kacperczyk, M.; Witkowska, A.; Kratz, E. M.; Rauf, A. *Oxid. Med. Cell. Longevity* **2022**, *2022*, 7657876.

(37) Hill, D. W.; Baveghems, C. L.; Albaugh, D. R.; Kormos, T. M.; Lai, S.; Ng, H. K.; Grant, D. F. *Rapid communications in mass spectrometry: RCM* **2012**, *26* (19), 2303−10.

(38) Dubey, R.; Hill, D. W.; Lai, S.; Ming-Hui, C.; Grant, D. F. *Metabolomics* **2015**, *11* (3), 753−763.

(39) Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Räther, O.; Mann, M. *Molecular & cellular proteomics: MCP* **2018**, *17* (12), 2534−2545.

(40) Meier, F.; Beck, S.; Grassl, N.; Lubeck, M.; Park, M. A.; Raether, O.; Mann, M. *J. Proteome Res.* **2015**, *14* (12), 5378−5387.

(41) Wei, J.; Tang, Y.; Ridgeway, M. E.; Park, M. A.; Costello, C. E.; Lin, C. *Anal. Chem.* **2020**, *92* (19), 13211−13220.

(42) Hsieh, E. J.; Bereman, M. S.; Durand, S.; Valaskovic, G. A.; MacCoss, M. J. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (1), 148−53.

(43) Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S. M.; Bondt, A.; Wuhrer, M. *Glycoconjugate Journal* **2016**, *33* (3), 309−343.