Autonomous Dissociation-type Selection for Glycoproteomics using a Real-Time Library Search

Emmajay Sutherland[1†], Tim S. Veth[1†], William D. Barshop[2], Jacob H. Russell[1], Kathryn Kothlow[1], Jesse D. Canterbury[2], Christopher Mullen[2], David Bergen[2], Jingjing Huang[2], Vlad Zabrouskov[2], Romain Huguet[2], Graeme C. McAlister[2], Nicholas M. Riley[1*]

[1]Department of Chemistry, University of Washington, Seattle, WA, USA
[2]Thermo Fisher Scientific, San Jose, California, USA

[†]E.S. and T.S.V contributed equally to this work
*Submit correspondence to nmriley@uw.edu

ABSTRACT

Tandem mass spectrometry (MS/MS) is the gold standard for intact glycopeptide identification, enabling peptide sequence elucidation and site-specific localization of glycan compositions. Beam-type collisional activation is generally sufficient for *N*-glycopeptides, while electron-driven dissociation is crucial for site localization in *O*-glycopeptides. Modern glycoproteomic methods often employ multiple dissociation techniques within a single LC-MS/MS analysis, but this approach frequently sacrifices sensitivity when analyzing multiple glycopeptide classes simultaneously. Here we explore the utility of intelligent data acquisition for glycoproteomics through real-time library searching (RTLS) to match oxonium ion patterns for on-the-fly selection of the appropriate dissociation method. By matching dissociation method with glycopeptide class, this autonomous dissociation-type selection (ADS) generates equivalent numbers of *N*-glycopeptide identifications relative to traditional beam-type collisional activation methods while also yielding comparable numbers of site-localized *O*-glycopeptide identifications relative to conventional electron transfer dissociation-based methods. The ADS approach represents a step forward in glycoproteomics throughput by enabling site-specific characterization of both *N*-and *O*-glycopeptides within the same LC-MS/MS acquisition.

KEYWORDS
Glycoproteomics, *N*-glycopeptides, *O*-glycopeptides, real-time library search, beam-type collisional dissociation, electron transfer dissociation, tandem mass spectrometry, product-dependent triggering, intelligent data acquisition

INTRODUCTION

Protein glycosylation is a highly heterogeneous collection of co- and post-translational modifications.[1,2] Two common classes of glycosylation include *N*-glycosylation, where glycans are attached through nitrogen atoms in asparagine residues within a consensus motif/sequon defined by *N*-X-S/T (where X cannot be proline),[3,4] and mucin-type *O*-glycosylation, where glycans are initiated by an *N*-acetylgalactosamine (GalNAc) linked through oxygen atoms on serine, threonine, and sometimes tyrosine residues.[5,6] Tandem mass spectrometry (MS/MS) remains the leading method for glycoproteomic analysis, enabling site-specific localization of glycan modifications on peptide backbones.[7,8] Higher-energy collisional dissociation (HCD, also called beam-type collisional dissociation) and stepped collision energy HCD (sceHCD) are widely used in *N*-glycopeptide studies, largely due to their fragmentation efficiency, fast acquisition speeds, and ubiquity on most modern mass spectrometers.[9–11] HCD and sceHCD generate numerous glycan-specific fragments useful for glycan compositional assignments, which typically dominate glycopeptide collisional dissociation spectra because of the lability of glycosidic bonds.[11–13] Collision-based dissociation can also generate peptide sequencing ions (e.g., b- and y-type fragment ions), but these ions often lose the entire glycan or retain only a small portion (e.g., a single *N*-acetylhexosamine, a.k.a. HexNAc, residue).[14–17] Peptide sequencing ions that retain the full glycan are ideal for site-specific glycosite localization because they fully explain modifications at each site along the peptide backbone. In the case of *N*-glycosylation, however, most tryptic *N*-glycopeptides contain only one *N*-sequon, and thus only one potential *N*-glycosite, meaning the fragment ion types generated by HCD and sceHCD are considered generally sufficient for standard *N*-glycoproteomic experiments.[18]

Unlike *N*-glycosylation, mucin-type *O*-glycosylation does not have a known consensus motif, and *O*-glycosites often occur in sequences with multiple potential sites of modifications that could harbor numerous combinations of *O*-glycan compositions. These features necessitate alternative forms of fragmentation, such as electron-driven dissociation, that generate peptide sequencing ions that retain glycan modifications after peptide backbone dissociation.[19,20] Electron-transfer dissociation (ETD) is the most widely available alternative dissociation method,[21] and ETD is often conducted using supplemental activation, such as ETD followed by HCD (EThcD), to improve its sequence-informative fragment ion yield.[22–24] EThcD often provides higher quality *N*-glycopeptide spectra than HCD or sceHCD, and it has proven especially useful in *O*-glycopeptide analysis;[25–30] one challenge, however, is the overhead time required to complete the ETD ion-ion reaction. Because both reagent anion accumulation times and ETD reaction times are generally on the order of tens of milliseconds,[31] significantly fewer ETD or EThcD spectra can be collected in a given analysis time compared to the near instantaneous fragmentation achieved in collisional dissociation.

Clearly, collision-based fragmentation and electron-driven dissociation each offer advantages, so current glycoproteomic methods often combine them strategically to capitalize on the strengths of both. A widely implemented data acquisition scheme involves product-dependent (pd) triggering to acquire multiple MS/MS spectra for the same precursor ion using an HCD-pd-EThcD format. These methods rely on 'scout HCD' scans collected for all precursor ions. If glycan-specific oxonium ions are detected in a scout HCD spectrum, it is assumed to be a glycopeptide candidate, and a subsequent EThcD MS/MS spectrum is acquired for that same precursor ion.[32–35] HCD-pd-EThcD are autonomous methods that attempt to maximize scan acquisition time by using EThcD only for potential glycopeptide candidates, and they are the most utilized method for *O*-glycoproteome characterization. That said, they tend to underperform typical (sce)HCD-only methods for *N*-glycopeptide analysis because the time spent performing EThcD for *N*-glycopeptides only adds marginal benefits over the largely sufficient collisional dissociation spectra.[18]

Despite the fact that most complex glycoproteome samples contain *N*- and *O*-glycopeptides, the disparity in performance between sceHCD-only and HCD-pd-EThcD methods leads to *a priori* decisions in methodology to maximize one glycopeptide class or the other. This issue limits sensitivity and throughput, putting a bottleneck on future expansion of glycoproteomics to large study designs and samples starting with low amounts of material (especially considering the necessity of enrichment in glycoproteomics).[36–39] We hypothesized that on-the-fly characterization of glycopeptide classes could significantly address this challenge by introducing a single scan acquisition architecture to analyze *N*- and *O*-glycopeptides without sacrificing data quality for either. Our goal was to classify potential glycopeptide precursor ions in real time to enable autonomous selection of the appropriate dissociation type, i.e., sceHCD only for *N*-glycopeptides or EThcD for *O*-glycopeptides. To accomplish this, we turned to real-time spectral library searching (RTLS), an intelligent data acquisition strategy that matches experimental spectra to library spectra as the instrument acquires data.[40–43] Here, we demonstrate the benefits of our RTLS-enabled strategy, called autonomous dissociation-type selection (ADS), for glycoproteomic analyses of multiple sample types, showing how ADS can maximize *N*- and *O*-glycopeptide characterization in a single LC-MS/MS acquisition.

EXPERIMENTAL PROCEDURES

Digestion of bovine fetuin, human serum, and mucin mixture. Bovine fetuin (Uniprot accession P12763) was purchased from Sigma Aldrich, and normal human serum (31876) was purchased from Invitrogen (Thermo Fisher Scientific). Recombinant MUC16 (Uniprot accession Q8WXI7), CD43 (Uniprot accession P16150), GP1ba (Uniprot accession P07359), Podocalyxin (Uniprot accession O00592), and PSGL-1 (Uniprot accession Q14242) were purchased from R&D Systems. All other reagents were purchased from Sigma Aldrich. Three separate tryptic digests were prepared: 1) bovine fetuin, 2) human serum, 3) a “mucin mix” of MUC16, CD43, GP1ba, Podocalyxin, and PSGL-1. All samples were reconstituted in 100 mM Tris buffer pH 8. For fetuin (3 mg total starting protein amount) and human serum samples (12 mg total starting protein amount), disulfide bonds were reduced and cysteines were carbamidomethylated using a final concentration of 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 40 mM chloroacetamide (CAA), followed by overnight digestion 37 °C using trypsin at a 1:50 (w/w) trypsin:protein ratio. Fetuin and serum peptides were desalted using Strata-X cartridges (Phenomenex) by conditioning the cartridge with 1 mL acetonitrile (ACN) followed by 1 mL 0.2% formic acid (FA) in water. Peptides were acidified with formic acid and then loaded onto the cartridge, followed by a 1 mL wash with 0.2% FA in water. Peptides were eluted with 400 µL of 0.2% FA in 80% ACN and dried via vacuum centrifugation. Glycopeptides were then enriched using a mixed-mode anion exchange protocol adapted from Bermudez and Pitteri.[44] Briefly, Oasis MAX cartridges (Waters) were preconditioned with 3 mL each of ACN, 100 mM triethylammonium acetate in water, 1% trifluoroacetic acid (TFA) in water, and 1% TFA in 95% ACN. Peptides were resuspended in 200 µL 0.1% TFA in 50% ACN and were added to 3 mL of 1% TFA in 95% ACN prior to adding the entire 3.2 mL solution to the cartridge. The entire volume was put over the cartridge, the flow through was collected, and the flow through was reloaded to the cartridge a second time. Then the cartridge was washed with 6 mL 1% TFA in 95% ACN and bound peptides were eluted with 1 mL 0.1% TFA in 50% ACN. Enriched peptides were dried via lyophilization and resuspended in 0.2% FA for LC-MS/MS. For the mucin mix, all proteins (~30 µg each) were resuspended in 100 mM Tris pH 8 and combined prior to a three hour incubation at 37 °C with secreted protease of C1 esterase inhibitor (StcE from *Escherichia coli*, purified as previously reported[45]) at a 1:10 (w/w) enzyme:protein ratio, followed by an overnight digestion at 37 °C in a 1:50 (w/w) trypsin:protein ratio. The resulting peptide mixture was desalted identically to the first desalting step, dried, and

resuspended in 0.2% FA for LC-MS/MS. The same batch of enriched peptides for a given sample was used for all method conditions tested.

LC-MS/MS data acquisition. Samples were analyzed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to a Thermo EASY-nLC 1200 (Thermo Fisher Scientific). First, samples were reconstituted in 0.1% FA and loaded onto the analytical column at max flow rate with a pressure limit of 700 bar. Using a 300 µL/min flow rate for all steps of the gradient during acquisition, the method started with 1 min of isocratic flow at 3% B, and peptides were linearly eluted using a 70-minute gradient ranging from 3% buffer B (80% ACN with 0.1% FA) to 42% buffer B on a 50 cm PepMap Neo C18 analytical column (Thermo Fisher Scientific). The gradient then increased to 95% B in 1 minute, the column was washed for 8 minutes with 95% B for a total of 80 minutes of acquisition time. After acquisition ended, the column was equilibrated at 3% B. Precursors were ionized using an EasySpray ionization source (Thermo Fisher Scientific) held at +1.6 kV compared to ground, the inlet capillary temperature was held at 275°C, and the RF lens was set to 30%. Varied mass spectrometer settings were tested, but unless explicitly mentioned, the instrument method used a survey MS scan with resolution of 60k at 200 m/z, a normalized AGC target of 100% (400,000 charges), a maximum injection time of 50 ms, and a scan range of 400-1800 m/z. Monoisotopic peptide filtering was enabled, dynamic exclusion was set at 45 seconds after 1 occurrence with a ±10 ppm window, a minimum intensity threshold of 25,000 was used, and charges 2-8 were selected for fragmentation. MS/MS scans were performed in a data-dependent fashion with a cycle time of 2 seconds. Precursors were isolated using a 1.6 m/z quadrupole isolation width and were measured in the Orbitrap with an MS2 resolution of 30k at 200 m/z. For sceHCD-only methods, sceHCD was performed with 20, 30, and 40% normalized collision energies, a scan range of 120-2000 m/z, a normalized AGC target set to 200% (100,000 charges), and a maximum injection time of 150 ms. sceHCD-pd-EThcD methods used the same sceHCD MS/MS settings with a product-dependent trigger filter requiring the presence of four of the following oxonium ions within the top 20 most abundant ions in the sceHCD MS/MS spectrum and with a ±15 ppm tolerance: 126.055, 138.0549, 144.0655, 168.0654, 186.076, 204.0865, 274.0921, 290.087, 292.1027, 308.0976, 366.1395, 657.2349, 673.2298. If those conditions were met, EThcD scans were triggered with calibrated ETD ion-ion reaction parameters[31], a supplemental HCD activation of 25% normalized collision energy, a normalized AGC target set to 300% (150,000 charges), a scan range of 120-4000 m/z, and a maximum injection time of either 200 or 500 ms, as discussed in the text.

Real-time spectral library searching (RTLS). RTLS was conducted in the context of sceHCD-pd-RTLS-EThcD methods, which had identical parameters to the sceHCD-pd-EThcD methods described above, with the exception of RTLS performed after the detection of oxonium ions and before triggering the EThcD scan. The spectral library for RTLS was created using Thermo mzVault and contained two spectra with two peaks each, namely a 138.0549 m/z peak and a 144.0655 m/z peak. The 138.0549 m/z peak was given a relative intensity of 1.0 in both spectra. For the spectrum matching an *O*-glycopeptide, the 144.0655 m/z peak was scaled 1:1 to the 138.0549 m/z peak. For the spectrum matching a *N*-glycopeptide this ratio was scaled 20:1. sceHCD scans were only passed to the RTLS service if scans passed the oxonium ion product-dependent triggering threshold described above. For each experimental spectrum considered by RTLS, cosine similarity scores to the library candidates were calculated.[40] The goal of RTLS in this context is not to match to specific precursor ions, but instead to evaluate oxonium ratios irrespective of precursor ion m/z or mass to "reject" or "promote" the acquisition of additional scans for the desired glycopeptide class.[43] Thus, the precursor search tolerance for RTLS was effectively disabled via the "Similarity Search" option, along with a max search time of 150 ms and an activation energy tolerance of 100. Cosine score thresholds for both *N*- and *O*-glycopeptide spectral matching were empirically evaluated, as discussed in the text. Experimental

spectra scored as a match to the *N*-glycopeptide library spectrum were not permitted to continue through the method sequence, while EThcD scans, with identical parameters as described above, were collected for precursor ions that matched to the *O*-glycopeptide library spectrum.

Data analysis. Data were searched using Byonic (Protein Metrics, v3.11.3)[46] for *N*-glycopeptides and *O*-Pair Search in MetaMorpheus (v0.0.317)[47] for *O*-glycopeptides. Fetuin data was searched with a protein database containing 86 common fetal bovine serum proteins, many of which are commonly co-purified with fetuin. Human serum samples were searched with the canonical human database containing 20,428 entries. Mucin mixture samples were searched against a fasta database containing the five mucin-domain glycoprotein sequences. Byonic was used for *N*-glycopeptide identifications by searching against a database with 72 N-glycans common to bovine plasma (**Table S1**) or 183 common human N-glycans (**Table S2**), and *N*-glycan modifications were denoted as common2 or common1 (meaning they could each occur twice or once) for searches of fetuin and serum data, respectively. The total common and rare max values were both set to 2. Carbamidomethylation (+57,021644) at cysteine was set as a fixed modification and oxidation (+15.994915) at methionine was set as a rare2 variable modification. Cleavage specificity was set as fully specific for C-terminal to R and K residues (fully tryptic) with two missed cleavages allowed. Precursor mass tolerance was set to 10 ppm with fragment mass tolerance(s) set to 20 ppm. Fragmentation was set to HCD or HCD & EThcD for appropriate raw files, and protein FDR was set to 1%. Following Byonic searching, results were filtered further, as previously described[18], with filtering metrics including a Byonic score greater than or equal to 200, a logProb value greater than or equal to 2, a deltaMod score greater than or equal to 10, and peptide length greater than 4 residues. These search parameters and filtering metrics were used for all samples, with only the fasta database varying as appropriate. *O*-Pair Search for *O*-glycopeptides used the same fasta files as appropriate and an *O*-glycan database with 22 entries (**Table S3**).[48] The "Keep top N candidates" feature was set to 50, and Data Type was set as HCD with Child Scan Dissociation set as EThcD. The "Maximum OGlycan Allowed" setting was set to 4, where this number represents both the maximum number of *O*-glycan modifications that could occur on a glycopeptide candidate and the number of times each *O*-glycan could occur per peptide. Under Search Parameters, both "Use Provided Precursor" and "Deconvolute Precursors" were checked. Peak trimming was not enabled. *In silico* digestion parameters were set to generate decoy proteins using reversed sequences, and the initiator methionine feature was set to "Variable". The maximum modification isoforms allowed was 1024, and semi-tryptic digestion was enabled for peptides ranging from 5 to 60 residues. Precursor and product mass tolerances were 10 and 20 ppm, respectively, and the minimum score allowed was 3. Carbamidomethylation on cysteine and oxidation on methionine were fixed and variable modifications, respectively. Results were filtered to keep the highest quality (Level 1) identifications only. The mass spectrometry raw data, fasta files, and search results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD055054.[49]

RESULTS AND DISCUSSION

An oxonium ion ratio to classify *N*- and *O*-glycopeptide spectra. Oxonium ions, also referred to as B-type ions, are low m/z ions (~*m/z* 100-400) that originate from the fragmentation of the overall glycan structure into singly charged mono- and polysaccharide units (or fragments thereof).[50] Halim *et al.* showed that a distinct *N*-acetylglucosamine/*N*-acetylgalactosamine (GlcNAc/GalNAc) ratio arises from the differences observed between the preferred fragmentation pathways of GlcNAc vs GalNAc, highlighting the dependence of oxonium ion intensities on glycan structures.[51] Their data showed that the equatorial C-4 hydrogen of GlcNAc leads to a greater presence of *m/z* 138 and *m/z* 168 ions, while the axial C-4 hydrogen of GalNAc results in the preferential formation of *m/z* 126 and *m/z* 144 ions. Their GlcNAc/GalNAc ratio used a calculation

of all four ions, demonstrating ratios of 2-50 and 0.2-1 for GlcNAc- and GalNAc-containing glycopeptides, respectively.[52] Differences between GalNAc and GlcNAc-containing glycopeptides can also be seen by simply comparing *m/z* 138 and 144,[53–55] and several studies have capitalized on differences in other oxonium ion intensities, as well.[56–59]

To understand if we could use this for our goal of real-time classification of glycopeptide MS/MS spectra, we first looked at the 138/144 ratio differences between *N*- and *O*-glycopeptides from published data.[18] Mammalian *N*-glycans contain GlcNAc residues by default as part of the common chitobiose core, while mucin-type *O*-glycans initiate with a GalNAc residue.[32] **Figure 1A** shows how the 138/144 ratio can potentially be used to distinguish between *N*- and mucin-type *O*-linked glycopeptides, with majority of *N*-glycopeptides generating ratios > 10 and most *O*-glycopeptides generating ratios < 2. Note, most of the *O*-glycopeptide identifications in this data are from core-1 *O*-glycopeptides that contain no GlcNAc residues. Other mucin-type *O*-glycan cores can contain GlcNAc, e.g., core-2, which complicates the 138/144 ratio calculation. Core-1 *O*-glycans are a common and clinically relevant class of mucin-type glycosylation, so we elected to focus on core-1 *O*-glycopeptide for this first implementation of our autonomous dissociation-type selection (ADS) method.

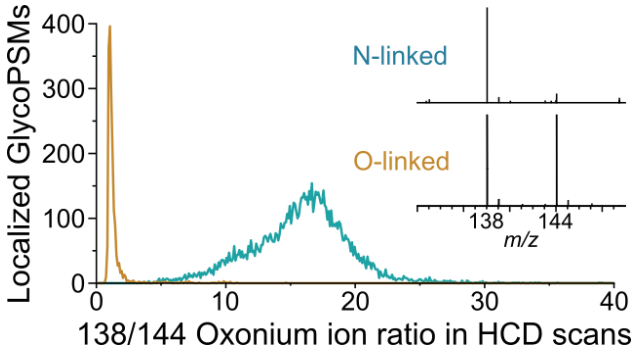

Figure 1. Oxonium ion ratios can indicate the glycopeptide type and may help select a specific dissociation method. A simple ratio between *m/z* 138/144 oxonium ions can indicate whether a glycopeptide likely belongs to the *N*-linked (containing high degrees of GlcNAc) or mucin-type *O*-linked (containing little GlcNAc, but mostly GalNAc) class. This data represents *O*-glycopeptides with predominantly core-1 mucin-type *O*-glycans, the focus of the *O*-glycopeptides in this study.

To exploit this *N*- vs. *O*-glycopeptide difference in 138/144 ratios, we turned to real-time library searching (RTLS) to match oxonium ion intensities during LC-MS/MS analysis. We constructed two library spectra, relating to either *N*- or *O*-glycopeptides, containing only m/z 138.0549 and m/z 144.0655 with intensities at either a 20:1 (*N*-glycopeptide) or 1:1 (*O*-glycopeptide) ratio. The RLTS service uses an mzVault spectral library search on the instrumental data system to assign a cosine similarity score of experimental spectra to library entries. This cosine score is returned to the instrument firmware by the RTLS service, allowing spectral matches to be used for further decision making. The goal of RTLS in ADS is to evaluate oxonium ion patterns irrespective of precursor ion m/z or mass to reject or promote further acquisition of EThcD scans depending on glycopeptide class. Because of this, large precursor search and activation energy tolerances atypical of RTLS parameters for molecule-specific libraries were a key part of RTLS settings for this small oxonium ion library. We started our ADS method design using the traditional sceHCD-only and sceHCD-pd-EThcD architectures commonly used in glycoproteomics (**Figure 2**). We elected to use sceHCD for all collision-based spectra because of previously described benefits.[18] We then designed RTLS-enabled ADS methods to use the same oxonium ion-based product-

dependent triggering to filter which sceHCD scans get sent to the RTLS service. If sceHCD MS/MS spectra had sufficient oxonium ion signal to pass the product-dependent filter, they were compared to the two library spectra to determine if their 138/144 ratios more closely matched *N*- or *O*-glycopeptides. If cosine similarity scores indicated that a spectrum was classified as an *O*-glycopeptide, an EThcD MS/MS spectrum was acquired for that same precursor ion. If RTLS classified the spectrum as an *N*-glycopeptide, no further MS/MS scans were acquired for that precursor ion (as governed by typical dynamic exclusion rules). **Figure 2** shows how ADS methods can be described as a minor addition to the typical pd-triggered workflow, i.e., sceHCD-pd-RTLS-EThcD.

To understand how RTLS cosine scoring would affect categorization of glycopeptides, we examined the distribution of cosine scores as a function of 138/144 ratio (**Figure S1**). Spectra classified by RTLS as *N*-glycopeptides typically had higher cosine scores (~20-80) relative to spectra classified as *O*-glycopeptides (~5-50), which is likely a function of the dominant signal from m/z 138 in *N*-glycopeptide spectra. Indeed, m/z 138 can often be present even when m/z 144 is absent in *N*-glycopeptide spectra with relatively low signal, so we set an infinitely high 138/144 ratio from spectra missing the m/z 144 signal to be a ratio of 35. Next, we tested cosine score thresholds using sceHCD-pd-RTLS-EThcD methods relative to a standard sceHCD-pd-EThcD method (**Figure S1**). Using a cosine score above 30 for the RTLS filter translated to fewer *N*-glycopeptide identifications that the standard method, while a score above 15 as the filter for *O*-glycopeptides meant fewer identifications. With this, we set the RTLS cosine score threshold to be 5 for *O*-glycopeptides in our ADS methods, since 1) the main filtering step only relied on quality of *O*-glycopeptides and 2) we wanted to trigger EThcD scans for as many *O*-glycopeptides as we could, even if they were lower quality sceHCD spectra.

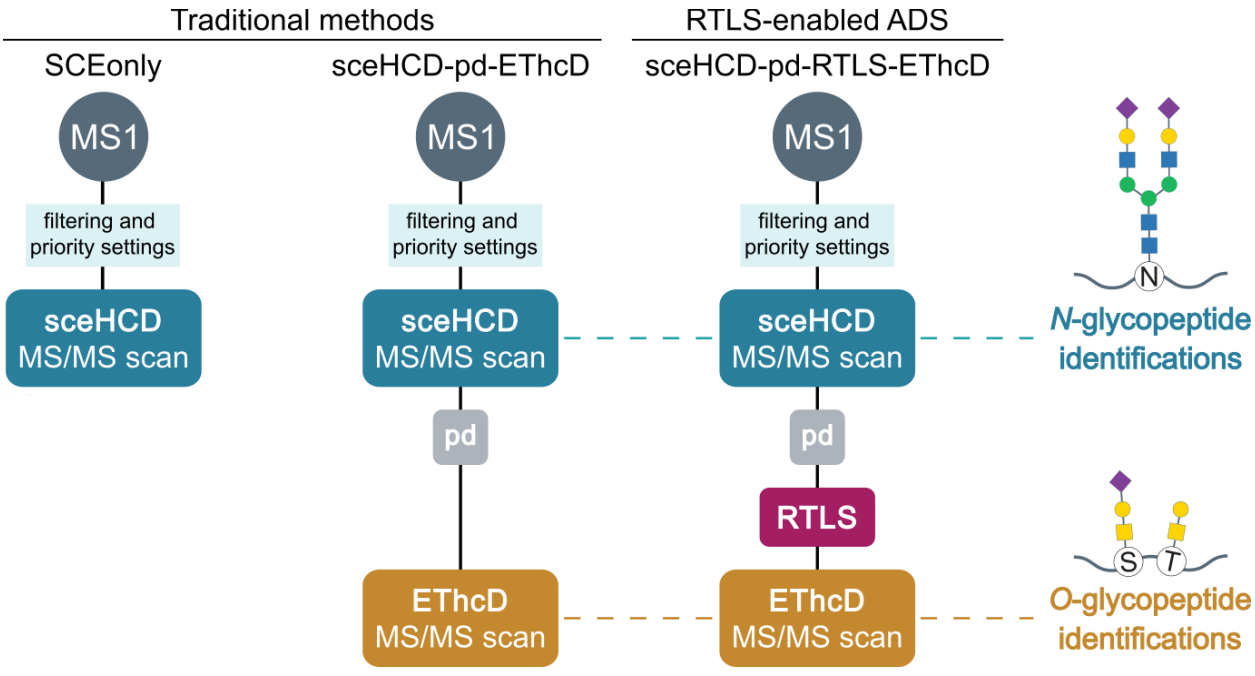

Figure 2. Conventional MS methods for glycoproteomics compared to the real-time library search (RTLS)-enabled autonomous dissociation-type selection (ADS) methods introduced in this study. Glycoproteomic studies often rely on purely collision-based dissociation (secHCD-only) methods or product-dependent triggering methods that combine collision- and electron-based dissociation (sceHCD-pd-EThcD). ADS methods insert an RTLS function as an additional decision point between the product-

dependent filter that purely evaluates the presence of oxonium ions. The RTLS service returns a match to *N*- or *O*-glycopeptide spectra, allowing the instrument to collect a subsequent MS/MS scan based on user-defined parameters. sceHCD scans are generally useful for *N*-glycopeptide identifications while electron-based dissociation is necessary for site-specific analysis of *O*-glycopeptides. Conventional sceHCD-pd-EThcD triggers EThcD for all glycopeptides, regardlesss of class. In this study, we used RTLS to classify likely glycopeptide precursor ions as *N*- or *O*-linked to acquire EThcD scans only for likely *O*-glycopeptides. Permitting EThcD scans to only be acquired when needed is the central premise of ADS, with the hope of collecting as many sceHCD spectra as possible to maximize glycoproteome sampling depth for both *N*- and *O*-glycopeptides. To inspect how our ADS method affected scan count, we compared sceHCD and EThcD scan acquisition for three methods: a sceHCD-only method, a sceHCD-pd-EThcD method, and an ADS (i.e., sceHCD-pd-RTLS-EThcD) method. While the sceHCD-only method generated the most scans (24,678 sceHCD spectra), the ADS method significantly improved the number of sceHCD scans taken over conventional sceHCD-pd-EThcD method (21,494 vs 8,679) while still acquiring 383 EThcD scans (**Figure S2**). The 4,242 EThcD scans in the traditional sceHCD-pd-EThcD method significantly slowed down scan acquisition, limiting the number of total spectra that can be acquired. Similarly, EThcD spectra in ADS methods were likely the reason for slightly fewer overall MS/MS scans. To verify that the RTLS service itself was not significantly slowing down data acquisition, we plotted the distribution of round-trip times each RTLS instance required (**Figure S2**). When compared to the typical ion accumulation times needed for sceHCD MS/MS spectra (**Figure S2**), it is clear that the several milliseconds needed per RTLS instance fits well within the scan overhead time available from the >100 ms needed per scan for ion accumulation. Therefore, RTLS did not contribute any significant effect to the overall scan acquisition rate in ADS methods.

To evaluate ADS methods for glycoproteomics, we prepared three samples with differing compositions and complexities and evaluated the number of localized unique glycopeptides identified from one sceHCD-only, two sceHCD-pd-EThcD, and two ADS methods (**Figure 3**). We first tested conventional and ADS method performance on semi-complex mixture of glycopeptides derived from a tryptic bovine fetuin digestion (which also includes common glycoproteins co-purified with fetuin) enriched for glycopeptides using a mixed-mode electrostatic repulsion hydrophilic interaction approach (**Figure 3A**). This sample, a common standard used in glycoproteomics, has both *N*- and *O*-glycopeptides with moderate complexities in their peptide and glycan compositions. The sceHCD-only method generated significantly more unique *N*-glycopeptide identifications than the traditional sceHCD-pd-EThcD, matching expected trends from previous.[18] Also expected was the inability of the sceHCD-only method to return reliable localized *O*-glycopeptide identifications. The traditional sceHCD-pd-EThcD methods did enable site-specific analysis of *O*-glycopeptides, generating a combination of *N*- and *O*-glycopeptide identifications that totaled more than the *N*-glycopeptides seen with sceHCD alone. That said, this gain in *O*-glycopeptide identification came to the detriment of *N*-glycopeptide identifications, ultimately meaning the pool of *N*-glycopeptides went under sampled with sceHCD-pd-EThcD. We used methods with 200 or 500 ms injection times for EThcD MS/MS scans to see if improving the quality of EThcD spectra would aid in *O*-glycopeptide identifications. A small increase in *O*-glycopeptide identifications without a decrease in *N*-glycopeptides was seen in the method using longer injection times, indicating that sceHCD-pd-EThcD methods may benefit from ensuring that EThcD scans are high quality, even if each EThcD scan takes longer to acquire. Regardless, our new ADS approach was able to provide site-specific *O*-glycopeptide identifications only obtainable with the EThcD MS/MS scans while maintaining *N*-glycopeptide identifications consistent with the sceHCD-only method. A similar benefit in EThcD-based *O*-glycopeptide identifications was also observed for the ADS method with a 500 ms maximum injection time (relative to 200 ms) for EThcD MS/MS scans. We also briefly explored an alternative ADS architecture with this sample that used a branched scan acquisition tree to collect a second

sceHCD MS/MS scan if the RTLS service returned a match to *N*-glycopeptides (**Figure S3**), but we ultimately elected to stick with the linear method architecture (**Figure 2**) for this remainder of this study.

We repeated these method comparisons for a tryptic digest of human serum, also enriched for glycopeptides (**Figure 3B**). This sample allowed us to evaluate ADS methods in a sample where *N*-glycosylation is dominant, due to numerous abundant and heterogeneous *N*-glycoproteins in serum. A similar drop in *N*-glycopeptide identifications was seen with conventional sceHCD-pd-EThcD methods when compared to the sceHCD-only method. Unlike the fetuin sample, however, there were not enough *O*-glycopeptides identifiable in the sample to increase the total number of identifications in sceHCD-pd-EThcD methods above sceHCD-only identifications, nor did the longer injection time for EThcD scans generate a significant difference in *O*-glycopeptide identifications. Overall, this drop in performance for sceHCD-pd-EThcD represents a clear example of why some glycoscientists choose sceHCD-only methods for their work. Conversely, our new ADS methods maintained the *N*-glycopeptide identifications seen with the sceHCD-only method while also generating *O*-glycopeptide identifications. The 500 ms injection time for EThcD MS/MS scans in our ADS method did offer a slight boost in *O*-glycopeptide identifications, supporting the idea that ensuring high quality EThcD scans is worth the extra overhead time. ADS minimized the detriment of this overhead time, providing the benefits of EThcD fragmentation only when it is required by the analyte.

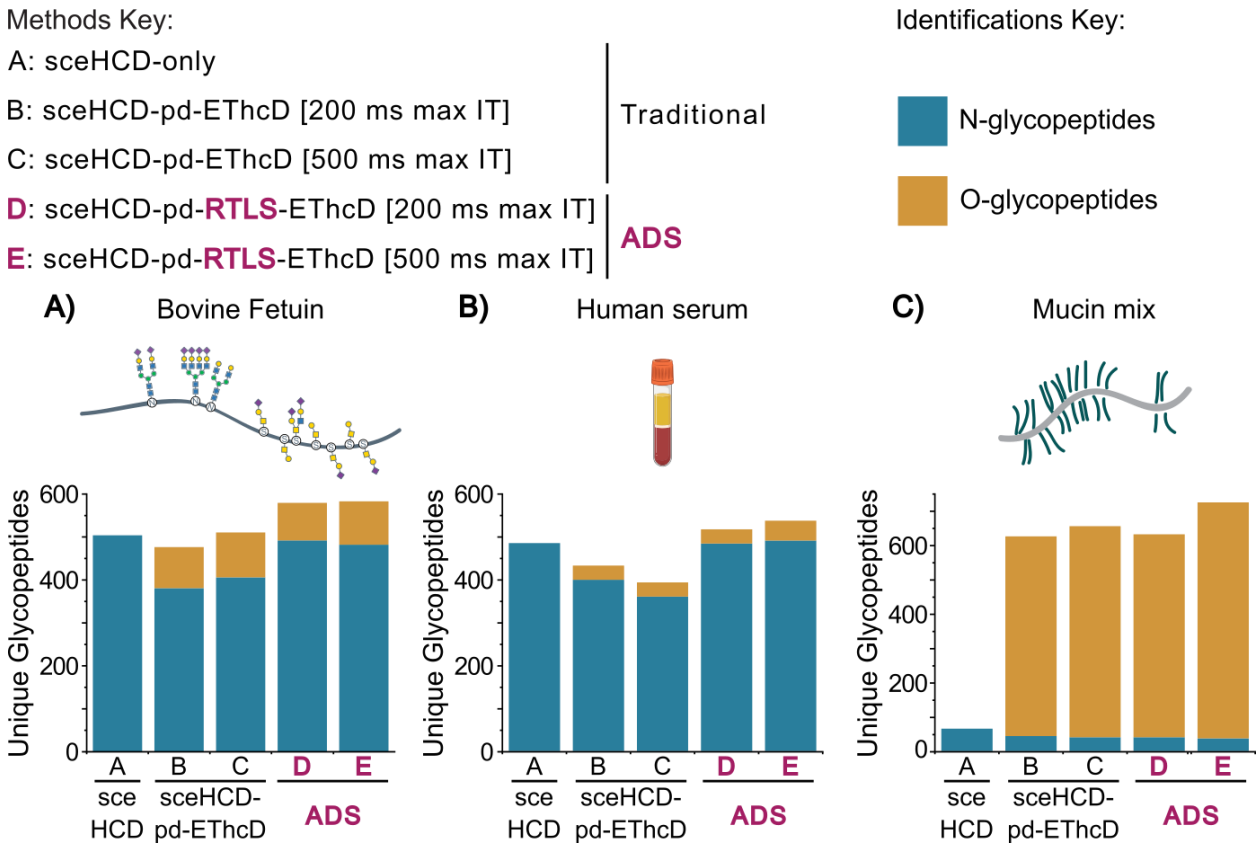

Figure 3. Performance of ADS methods relative to conventional glycoproteomic methods for glycopeptides from three different sources. Five methods (A-E) were compared, including one sceHCD-only (A), two sceHCD-pd-EThcD (B and C), and two ADS (i.e., sceHCD-pd-RTLS-EThcD) methods (D and

E). Methods with EThcD scans included a maximum injection time (max IT) set to either 200 ms or 500 ms for each precursor selected for EThcD MS/MS. These methods were compared across glycopeptides from **A)** bovine fetuin, **B)** human serum, and **C)** a mixture of mucin-domain glycoproteins. Blue and orange bars indicate the number of unique localized *N*- and *O*-glycopeptides identified, respectively, with ADS methods highlighted in pink text. All data reported are averages from two technical injection replicates.

To assess the converse of plasma, where *N*-glycopeptides are dominant, we also wanted to evaluate ADS in a sample where *O*-glycopeptides are dominant. We created a mixture of recombinant human mucin-domain glycoproteins sequentially digested with a mucin-specific *O*-glycoprotease, StcE, followed by trypsin.[45] These mucin domains contain dense regions of *O*-glycosites that can be cleaved by *O*-glycoproteases like StcE into MS-amenable *O*-glycopeptides. These mucin-domain glycoproteins also contain some *N*-glycosites, making them a good test case for ADS methods. As shown in **Figure 3C**, conventional sceHCD-pd-EThcD methods offer a clear advantage over sceHCD-only methods because of the high number of *O*-glycopeptides present. A small drop in *N*-glycopeptide identifications is seen, as before, but the gain in localized *O*-glycopeptides far outweighs this minor loss in *N*-glycopeptide characterization. Here, the longer, 500 ms injection time for EThcD scans also shows a benefit, which is logical given the reliance on high quality EThcD spectra for characterizing this sample. ADS methods with a 200 ms injection time for EThcD scans performed equivalently as traditional sceHCD-pd-EThcD methods, which is an important benchmark to hit. The ADS method's guiding principle, to only collect EThcD scans when needed, could hurt *O*-glycopeptide identifications if not implemented correctly, but this data supports that the ADS discrimination of *O*-glycopeptide spectra was not detrimental to characterizing a sample that is dominated by *O*-glycosylation. Furthermore, the ADS method with a 500 ms maximum injection time appeared to offer a slight advantage over all other methods tested here.

CONCLUSIONS

Glycoproteomic methods often differ in their tandem MS dissociation method based on the glycopeptide class they are targeting, such as collision-based dissociation for *N*-glycopeptides and electron-based dissociation for *O*-glycopeptides. A common trend in data acquisition has been to combine the collision- and electron-based dissociation in a single method through product-dependent triggering that uses oxonium ion detection in real time to initiate subsequent MS/MS scans with alternative fragmentation methods. Recently, however, this combination of methods, such as HCD-pd-EThcD, has been abandoned by some in favor of sceHCD-only methods that benefit from high scan acquisition speeds. This comes at the detriment of glycopeptides that require alternative fragmentation methods for glycosite localization, e.g., *O*-glycopeptides. Here, we addressed this challenge with a proof-of-principle study describing a new intelligent data acquisition strategy called autonomous dissociation-type selection (ADS). ADS uses a real-time library search to categorize glycopeptide precursor ions as likely *N*- or *O*-glycopeptide species, allowing on-the-fly selection of the appropriate fragmentation method without sacrificing sampling depth for either glycopeptide class.

ADS methods are straightforward to implement on current quadrupole-Orbitrap-linear ion trap Tribrid mass spectrometers that have the RTLS service enabled, e.g., Orbitrap Eclipse and Orbitrap Ascend systems. This first demonstration of ADS focused on differentiating *N*-glycopeptides from *O*-glycopeptides that were largely core-1 species. More detailed oxonium ion signatures have been explored in the literature, which are active areas of investigation for future iterations of ADS methods.[11,52,60–62] It is also possible that ADS methods could benefit glycoproteomic workflows that incorporate real-time searching[63] or other electron-driven and photodissociation methods,[64–68] and it could improve characterization of glycopeptides with multiple *N*-glycosites or with co-occupied *N*- and *O*-glycosites.[69,70] Incorporating more robust

libraries will enable better discrimination of glycopeptide classes and will increase the sensitivity of ADS methods to further improve glycoproteome characterization.

NOTES, COMPETING FINANCIAL INTERESTS, AND AUTHOR CONTRIBUTIONS

W.D.B., J.D.C., C.M., D.B., J.H., V.Z., R.H., and G.C.M. are employees of Thermo Fisher Scientific, the manufacturer of the instrumentation used in this research. N.M.R. receives support from Thermo Fisher Scientific under a nondisclosure agreement and is a consultant for Tegmine Therapeutics, Cartography Biosciences, and Augment Biologics. Experiments described in this manuscript were conducted on instruments housed in a Thermo Fisher Scientific facility. Conceptualization, N.M.R., W.D.B., and G.C.M; methodology, N.M.R., W.D.B., J.D.C., C.M., D.B., J.H., V.Z., R.H., and G.C.M; investigation, N.M.R., E.S., T.S.V., J.H.R., and K.K.; writing – original draft, E.S., T.S.V., J.H.R., K.K., and N.M.R.; writing – review & editing, E.S., T.S.V., and N.M.R. with input from all authors.

SUPPORTING INFORMATION

Supporting Information, including Figures S1-3 and Tables S1-3, is available free of charge online.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institutes of Health under Award Number R00GM147304 (N.M.R.), by an ASMS Research Award, and by a Washington Research Foundation Postdoctoral Fellowship (E.S). We thank John E. P. Syka for helpful conversations throughout this study. The mass spectrometry raw data, fasta files, and search results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD055054.[49]

REFERENCES

(1) *Essentials of Glycobiology*, 4th ed.; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.

(2) Schjoldager, K. T.; Narimatsu, Y.; Joshi, H. J.; Clausen, H. Global View of Human Protein Glycosylation Pathways and Functions. *Nat. Rev. Mol. Cell Biol. 2020 2112* **2020**, *21* (12), 729–749. https://doi.org/10.1038/s41580-020-00294-x.

(3) Hirata, T.; Kizuka, Y. N-Glycosylation. In *The Role of Glycosylation in Health and Disease*; Lauc, G., Trbojević-Akmačić, I., Eds.; Springer International Publishing: Cham, 2021; pp 3–24. https://doi.org/10.1007/978-3-030-70115-4_1.

(4) Schwarz, F.; Aebi, M. Mechanisms and Principles of N-Linked Protein Glycosylation. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 576–582. https://doi.org/10.1016/j.sbi.2011.08.005.

(5) Wandall, H. H.; Nielsen, M. A. I.; King-Smith, S.; de Haan, N.; Bagdonaite, I. Global Functions of O-Glycosylation: Promises and Challenges in O-Glycobiology. *FEBS J.* **2021**, *288* (24), 7183–7212. https://doi.org/10.1111/febs.16148.

(6) Li, J.; Guo, B.; Zhang, W.; Yue, S.; Huang, S.; Gao, S.; Ma, J.; Cipollo, J. F.; Yang, S. Recent Advances in Demystifying O-Glycosylation in Health and Disease. *PROTEOMICS* **2022**, *22* (23–24), 2200156. https://doi.org/10.1002/pmic.202200156.

(7) Bagdonaite, I.; Malaker, S. A.; Polasky, D. A.; Riley, N. M.; Schjoldager, K.; Vakhrushev, S. Y.; Halim, A.; Aoki-Kinoshita, K. F.; Nesvizhskii, A. I.; Bertozzi, C. R.; Wandall, H. H.; Parker, B. L.; Thaysen-Andersen, M.; Scott, N. E. Glycoproteomics. *Nat. Rev. Methods Primer* **2022**, *2* (1), 1–29. https://doi.org/10.1038/s43586-022-00128-4.
(8) Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilleke, E.; Lebrilla, C. B. Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem. Rev.* **2018**, *118* (17), 7886–7930. https://doi.org/10.1021/acs.chemrev.7b00732.
(9) Chau, T. H.; Chernykh, A.; Kawahara, R.; Thaysen-Andersen, M. Critical Considerations in *N*-Glycoproteomics. *Curr. Opin. Chem. Biol.* **2023**, *73*, 102272. https://doi.org/10.1016/j.cbpa.2023.102272.
(10) Peters-Clarke, T. M.; Coon, J. J.; Riley, N. M. Instrumentation at the Leading Edge of Proteomics. *Anal. Chem.* **2024**, *96* (20), 7976–8010. https://doi.org/10.1021/acs.analchem.3c04497.
(11) Hoffmann, M.; Pioch, M.; Pralow, A.; Hennig, R.; Kottler, R.; Reichl, U.; Rapp, E. The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses—Exploiting the Diagnostic Potential of Fragment Ions. *PROTEOMICS* **2018**, *18* (24), 1800282. https://doi.org/10.1002/pmic.201800282.
(12) 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. pGlyco 2.0 Enables Precision N-Glycoproteomics with Comprehensive Quality Control and One-Step Mass Spectrometry for Intact Glycopeptide Identification. *Nat. Commun.* **2017**, *8* (1), 438. https://doi.org/10.1038/s41467-017-00535-2.
(13) Shen, J.; Jia, L.; Dang, L.; Su, Y.; Zhang, J.; Xu, Y.; Zhu, B.; Chen, Z.; Wu, J.; Lan, R.; Hao, Z.; Ma, C.; Zhao, T.; Gao, N.; Bai, J.; Zhi, Y.; Li, J.; Zhang, J.; Sun, S. StrucGP: De Novo Structural Sequencing of Site-Specific N-Glycan on Glycoproteins Using a Modularization Strategy. *Nat. Methods* **2021**, *18* (8), 921–929. https://doi.org/10.1038/s41592-021-01209-0.
(14) Hevér, H.; Xue, A.; Nagy, K.; Komka, K.; Vékey, K.; Drahos, L.; Révész, Á. Can We Boost N-Glycopeptide Identification Confidence? Smart Collision Energy Choice Taking into Account Structure and Search Engine. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (2), 333–343. https://doi.org/10.1021/jasms.3c00375.
(15) Hevér, H.; Nagy, K.; Xue, A.; Sugár, S.; Komka, K.; Vékey, K.; Drahos, L.; Révész, Á. Diversity Matters: Optimal Collision Energies for Tandem Mass Spectrometric Analysis of a Large Set of N-Glycopeptides. *J. Proteome Res.* **2022**, *21* (11), 2743–2753. https://doi.org/10.1021/acs.jproteome.2c00519.
(16) Yang, H.; Yang, C.; Sun, T. Characterization of Glycopeptides Using a Stepped Higher-Energy C-Trap Dissociation Approach on a Hybrid Quadrupole Orbitrap. *Rapid Commun. Mass Spectrom.* **2018**, *32* (16), 1353–1362. https://doi.org/10.1002/rcm.8191.
(17) Wang, Y.; Tian, Z. New Energy Setup Strategy for Intact N-Glycopeptides Characterization Using Higher-Energy Collisional Dissociation. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (3), 651–657. https://doi.org/10.1021/jasms.9b00089.
(18) Riley, N. M.; Malaker, S. A.; Driessen, M. D.; Bertozzi, C. R. Optimal Dissociation Methods Differ for N- and O-Glycopeptides. *J. Proteome Res.* **2020**, *19* (8), 3286–3301. https://doi.org/10.1021/acs.jproteome.0c00218.
(19) Darula, Z.; Medzihradszky, K. F. Analysis of Mammalian O-Glycopeptides - We Have Made a Good Start, but There Is a Long Way to Go. *Mol. Cell. Proteomics* **2018**, *17* (1), 2–17. https://doi.org/10.1074/mcp.MR117.000126.
(20) Khoo, K. H. Advances toward Mapping the Full Extent of Protein Site-Specific O-GalNAc Glycosylation That Better Reflects Underlying Glycomic Complexity. *Curr. Opin. Struct. Biol.* **2019**, *56*, 146–154. https://doi.org/10.1016/j.sbi.2019.02.007.

(21) Riley, N. M.; Coon, J. J. The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal. Chem.* **2018**, *90* (1), 40–64. https://doi.org/10.1021/acs.analchem.7b04810.
(22) Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L. Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-Enabled Intact Glycopeptide/Glycoproteome Characterization. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1751–1764. https://doi.org/10.1007/s13361-017-1701-4.
(23) Riley, N. M.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Capturing Site-Specific Heterogeneity with Large-Scale N-Glycoproteome Analysis. *Nat. Commun.* **2019**, *10* (1), 1311. https://doi.org/10.1038/s41467-019-09222-w.
(24) Glover, M. S.; Yu, Q.; Chen, Z.; Shi, X.; Kent, K. C.; Li, L. Characterization of Intact Sialylated Glycopeptides and Phosphorylated Glycopeptides from IMAC Enriched Samples by EThcD Fragmentation: Toward Combining Phosphoproteomics and Glycoproteomics. *Int. J. Mass Spectrom.* **2017**. https://doi.org/10.1016/J.IJMS.2017.09.002.
(25) Riley, N. M.; Malaker, S. A.; Bertozzi, C. R. Electron-Based Dissociation Is Needed for O-Glycopeptides Derived from OpeRATOR Proteolysis. *Anal. Chem.* **2020**, *92* (22), 14878–14884. https://doi.org/10.1021/acs.analchem.0c02950.
(26) Pap, A.; Klement, E.; Hunyadi-Gulyas, E.; Darula, Z.; Medzihradszky, K. F. Status Report on the High-Throughput Characterization of Complex Intact O-Glycopeptide Mixtures. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1210–1220. https://doi.org/10.1007/s13361-018-1945-7.
(27) Darula, Z.; Pap, Á.; Medzihradszky, K. F. Extended Sialylated O-Glycan Repertoire of Human Urinary Glycoproteins Discovered and Characterized Using Electron-Transfer/Higher-Energy Collision Dissociation. *J. Proteome Res.* **2019**, *18* (1), 280–291. https://doi.org/10.1021/ACS.JPROTEOME.8B00587/SUPPL_FILE/PR8B00587_SI_002.PDF.
(28) Pap, A.; Kiraly, I. E.; Medzihradszky, K. F.; Darula, Z. Multiple Layers of Complexity in O-Glycosylation Illustrated With the Urinary Glycoproteome. *Mol. Cell. Proteomics* **2022**, *21* (12). https://doi.org/10.1016/j.mcpro.2022.100439.
(29) Chen, Z.; Wang, D.; Yu, Q.; Johnson, J.; Shipman, R.; Zhong, X.; Huang, J.; Yu, Q.; Zetterberg, H.; Asthana, S.; Carlsson, C.; Okonkwo, O.; Li, L. In-Depth Site-Specific O-Glycosylation Analysis of Glycoproteins and Endogenous Peptides in Cerebrospinal Fluid (CSF) from Healthy Individuals, Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD) Patients. *ACS Chem. Biol.* **2022**, *17* (11), 3059–3068. https://doi.org/10.1021/acschembio.1c00932.
(30) Burt, R. A.; Dejanovic, B.; Peckham, H. J.; Lee, K. A.; Li, X.; Ounadjela, J. R.; Rao, A.; Malaker, S. A.; Carr, S. A.; Myers, S. A. Novel Antibodies for the Simple and Efficient Enrichment of Native O-GlcNAc Modified Peptides. *Mol. Cell. Proteomics MCP* **2021**, *20*, 100167. https://doi.org/10.1016/j.mcpro.2021.100167.
(31) Rose, C. M.; Rush, M. J. P.; Riley, N. M.; Merrill, A. E.; Kwiecien, N. W.; Holden, D. D.; Mullen, C.; Westphall, M. S.; Coon, J. J. A Calibration Routine for Efficient ETD in Large-Scale Proteomics. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (11), 1848–1857. https://doi.org/10.1007/s13361-015-1183-1.
(32) Singh, C.; Zampronio, C. G.; Creese, A. J.; Cooper, H. J. Higher Energy Collision Dissociation (HCD) Product Ion-Triggered Electron Transfer Dissociation (ETD) Mass Spectrometry for the Analysis of N-Linked Glycoproteins. *J. Proteome Res.* **2012**, *11* (9), 4517–4525. https://doi.org/10.1021/pr300257c.
(33) Wu, S.-W.; Pu, T.-H.; Viner, R.; Khoo, K.-H. Novel LC-MS2 Product Dependent Parallel Data Acquisition Function and Data Analysis Workflow for Sequencing and Identification of Intact Glycopeptides. *Anal. Chem.* **2014**, *86* (11), 5478–5486. https://doi.org/10.1021/ac500945m.

(34) Zhao, P.; Viner, R.; Teo, C. F.; Boons, G.-J.; Horn, D.; Wells, L. Combining High-Energy C-Trap Dissociation and Electron Transfer Dissociation for Protein O-GlcNAc Modification Site Assignment. *J. Proteome Res.* **2011**, *10* (9), 4088–4104. https://doi.org/10.1021/pr2002726.
(35) Saba, J.; Dutta, S.; Hemenway, E.; Viner, R. Increasing the Productivity of Glycopeptides Analysis by Using Higher-Energy Collision Dissociation-Accurate Mass-Product-Dependent Electron Transfer Dissociation. *Int. J. Proteomics* **2012**, *2012* (1), 560391. https://doi.org/10.1155/2012/560391.
(36) Chrysinas, P.; Venkatesan, S.; Ang, I.; Ghosh, V.; Chen, C.; Neelamegham, S.; Gunawan, R. Cell and Tissue-Specific Glycosylation Pathways Informed by Single-Cell Transcriptomics. bioRxiv January 3, 2024, p 2023.09.26.559616. https://doi.org/10.1101/2023.09.26.559616.
(37) de Haan, N.; Pučić-Baković, M.; Novokmet, M.; Falck, D.; Lageveen-Kammeijer, G.; Razdorov, G.; Vučković, F.; Trbojević-Akmačić, I.; Gornik, O.; Hanić, M.; Wuhrer, M.; Lauc, G.; the The Human Glycome Project. Developments and Perspectives in High-Throughput Protein Glycomics: Enabling the Analysis of Thousands of Samples. *Glycobiology* **2022**, *32* (8), 651–663. https://doi.org/10.1093/glycob/cwac026.
(38) Marie, A.-L.; Gao, Y.; Ivanov, A. R. Native N-Glycome Profiling of Single Cells and Ng-Level Blood Isolates Using Label-Free Capillary Electrophoresis-Mass Spectrometry. *Nat. Commun.* **2024**, *15* (1), 3847. https://doi.org/10.1038/s41467-024-47772-w.
(39) Riley, N. M.; Bertozzi, C. R.; Pitteri, S. J. A Pragmatic Guide to Enrichment Strategies for Mass Spectrometry-Based Glycoproteomics. *Mol. Cell. Proteomics* **2020**, mcp.R120.002277. https://doi.org/10.1074/mcp.r120.002277.
(40) Bills, B.; Barshop, W. D.; Sharma, S.; Canterbury, J.; Robitaille, A. M.; Goodwin, M.; Senko, M. W.; Zabrouskov, V. Novel Real-Time Library Search Driven Data Acquisition Strategy for Identification and Characterization of Metabolites. *Anal. Chem.* **2022**, *94* (9), 3749–3755. https://doi.org/10.1021/acs.analchem.1c04336.
(41) Brademan, D. R.; Overmyer, K. A.; He, Y.; Barshop, W. D.; Canterbury, J. D.; Bills, B. J.; Anderson, B. J.; Hutchins, P. D.; Sharma, S.; Zabrouskov, V.; McAlister, G. C.; Coon, J. J. Improved Structural Characterization of Glycerophospholipids and Sphingomyelins with Real-Time Library Searching. *Anal. Chem.* **2023**, *95* (20), 7813–7821. https://doi.org/10.1021/acs.analchem.2c04633.
(42) McGann, C. D.; Barshop, W. D.; Canterbury, J. D.; Lin, C.; Gabriel, W.; Huang, J.; Bergen, D.; Zabrouskov, V.; Melani, R. D.; Wilhelm, M.; McAlister, G. C.; Schweppe, D. K. Real-Time Spectral Library Matching for Sample Multiplexed Quantitative Proteomics. *J. Proteome Res.* **2023**, *22* (9), 2836–2846. https://doi.org/10.1021/acs.jproteome.3c00085.
(43) Ruwolt, M.; He, Y.; Borges Lima, D.; Barshop, W.; Broichhagen, J.; Huguet, R.; Viner, R.; Liu, F. Real-Time Library Search Increases Cross-Link Identification Depth across All Levels of Sample Complexity. *Anal. Chem.* **2023**, *95* (12), 5248–5255. https://doi.org/10.1021/acs.analchem.2c05141.
(44) Bermudez, A.; Pitteri, S. J. Enrichment of Intact Glycopeptides Using Strong Anion Exchange and Electrostatic Repulsion Hydrophilic Interaction Chromatography. *Methods Mol. Biol.* **2021**, *2271*, 107–120. https://doi.org/10.1007/978-1-0716-1241-5_8.
(45) Shon, D. J.; Malaker, S. A.; Pedram, K.; Yang, E.; Krishnan, V.; Dorigo, O.; Bertozzi, C. R. An Enzymatic Toolkit for Selective Proteolysis, Detection, and Visualization of Mucin-Domain Glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (35), 21299–21307. https://doi.org/10.1073/pnas.2012196117.
(46) Roushan, A.; Wilson, G. M.; Kletter, D.; Sen, K. I.; Tang, W.; Kil, Y. J.; Carlson, E.; Bern, M. Peak Filtering, Peak Annotation, and Wildcard Search for Glycoproteomics. *Mol. Cell. Proteomics* **2021**, *20*, 100011. https://doi.org/10.1074/MCP.RA120.002260.

(47) 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* (11), 1133–1138. https://doi.org/10.1038/s41592-020-00985-5.
(48) Riley, N.; Bertozzi, C. Deciphering O-Glycoprotease Substrate Preferences with O-Pair Search. *Mol. Omics* **2022**, 10.1039/D2MO00244B. https://doi.org/10.1039/D2MO00244B.
(49) Perez-Riverol, Y.; Bai, J.; Bandla, C.; García-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.; Kundu, D. J.; 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* (D1), D543–D552. https://doi.org/10.1093/nar/gkab1038.
(50) Domon, B.; Costello, C. E. A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconj. J.* **1988**, *5* (4), 397–409. https://doi.org/10.1007/BF01049915.
(51) Halim, A.; Westerlind, U.; Pett, C.; Schorlemer, M.; Rüetschi, U.; Brinkmalm, G.; Sihlbom, C.; Lengqvist, J.; Larson, G.; Nilsson, J. Assignment of Saccharide Identities through Analysis of Oxonium Ion Fragmentation Profiles in LC-MS/MS of Glycopeptides. *J. Proteome Res.* **2014**, *13* (12), 6024–6032. https://doi.org/10.1021/PR500898R/SUPPL_FILE/PR500898R_SI_001.PDF.
(52) Yu, J.; Schorlemer, M.; Gomez Toledo, A.; Pett, C.; Sihlbom, C.; Larson, G.; Westerlind, U.; Nilsson, J. Distinctive MS/MS Fragmentation Pathways of Glycopeptide-Generated Oxonium Ions Provide Evidence of the Glycan Structure. *Chem. – Eur. J.* **2016**, *22* (3), 1114–1124. https://doi.org/10.1002/chem.201503659.
(53) Pirro, M.; Mohammed, Y.; de Ru, A. H.; Janssen, G. M. C.; Tjokrodirijo, R. T. N.; Madunić, K.; Wuhrer, M.; van Veelen, P. A.; Hensbergen, P. J. Oxonium Ion Guided Analysis of Quantitative Proteomics Data Reveals Site-Specific O-Glycosylation of Anterior Gradient Protein 2 (AGR2). *Int. J. Mol. Sci.* **2021**, *22* (10), 5369. https://doi.org/10.3390/ijms22105369.
(54) Stavenhagen, K.; Kayili, H. M.; Holst, S.; Koeleman, C. A. M.; Engel, R.; Wouters, D.; Zeerleder, S.; Salih, B.; Wuhrer, M. N- and O-Glycosylation Analysis of Human C1-Inhibitor Reveals Extensive Mucin-Type O-Glycosylation. *Mol. Cell. Proteomics* **2018**, *17* (6), 1225–1238. https://doi.org/10.1074/mcp.RA117.000240.
(55) Malaker, S. A.; Riley, N. M.; Shon, D. J.; Pedram, K.; Krishnan, V.; Dorigo, O.; Bertozzi, C. R. Revealing the Human Mucinome. *Nat. Commun.* **2022**, *13* (1), 3542. https://doi.org/10.1038/s41467-022-31062-4.
(56) White, M. E. H.; Sinn, L. R.; Jones, D. M.; de Folter, J.; Aulakh, S. K.; Wang, Z.; Flynn, H. R.; Krüger, L.; Tober-Lau, P.; Demichev, V.; Kurth, F.; Mülleder, M.; Blanchard, V.; Messner, C. B.; Ralser, M. Oxonium Ion Scanning Mass Spectrometry for Large-Scale Plasma Glycoproteomics. *Nat. Biomed. Eng.* **2024**, *8* (3), 233–247. https://doi.org/10.1038/s41551-023-01067-5.
(57) Madsen, J. A.; Farutin, V.; Lin, Y. Y.; Smith, S.; Capila, I. Data-Independent Oxonium Ion Profiling of Multi-Glycosylated Biotherapeutics. *mAbs* **2018**, *10* (7), 968. https://doi.org/10.1080/19420862.2018.1494106.
(58) Mukherjee, S.; Jankevics, A.; Busch, F.; Lubeck, M.; Zou, Y.; Kruppa, G.; Heck, A. J. R.; Scheltema, R. A.; Reiding, K. R. Oxonium Ion–Guided Optimization of Ion Mobility–Assisted Glycoproteomics on the timsTOF Pro. *Mol. Cell. Proteomics* **2023**, *22* (2). https://doi.org/10.1016/j.mcpro.2022.100486.
(59) Toghi Eshghi, S.; Yang, W.; Hu, Y.; Shah, P.; Sun, S.; Li, X.; Zhang, H. Classification of Tandem Mass Spectra for Identification of N- and O-Linked Glycopeptides. *Sci. Rep.* **2016**, *6*, 37189.

(60) Vékey, K.; Ozohanics, O.; Tóth, E.; Jekő, A.; Révész, Á.; Krenyácz, J.; Drahos, L. Fragmentation Characteristics of Glycopeptides. *Int. J. Mass Spectrom.* **2013**, *345–347*, 71–79. https://doi.org/10.1016/j.ijms.2012.08.031.
(61) Park, G. W.; Lee, J. W.; Lee, H. K.; Shin, J. H.; Kim, J. Y.; Yoo, J. S. Classification of Mucin-Type O-Glycopeptides Using Higher-Energy Collisional Dissociation in Mass Spectrometry. *Anal. Chem.* **2020**, *92* (14), 9772–9781. https://doi.org/10.1021/ACS.ANALCHEM.0C01218/SUPPL_FILE/AC0C01218_SI_004.XLSX.
(62) DeBono, N. J.; Moh, E. S. X.; Packer, N. H. Experimentally Determined Diagnostic Ions for Identification of Peptide Glycotopes. *J. Proteome Res.* **2024**, *23* (7), 2661–2673. https://doi.org/10.1021/acs.jproteome.3c00858.
(63) Armony, G.; Brehmer, S.; Srikumar, T.; Pfennig, L.; Zijlstra, F.; Trede, D.; Kruppa, G.; Lefeber, D. J.; van Gool, A. J.; Wessels, H. J. C. T. The GlycoPaSER Prototype as a Real-Time N-Glycopeptide Identification Tool Based on the PaSER Parallel Computing Platform. *Int. J. Mol. Sci.* **2023**, *24* (9), 7869. https://doi.org/10.3390/ijms24097869.
(64) Beckman, J. S.; Voinov, V. G.; Hare, M.; Sturgeon, D.; Vasil'ev, Y.; Oppenheimer, D.; Shaw, J. B.; Wu, S.; Glaskin, R.; Klein, C.; Schwarzer, C.; Stafford, G. Improved Protein and PTM Characterization with a Practical Electron-Based Fragmentation on Q-TOF Instruments. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (8), 2081–2091. https://doi.org/10.1021/JASMS.0C00482.
(65) Baba, T.; Ryumin, P.; Duchoslav, E.; Chen, K.; Chelur, A.; Loyd, B.; Chernushevich, I. Dissociation of Biomolecules by an Intense Low-Energy Electron Beam in a High Sensitivity Time-of-Flight Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (8), 1964–1975. https://doi.org/10.1021/JASMS.0C00425.
(66) Li, R.; Xia, C.; Wu, S.; Downs, M. J.; Tong, H.; Tursumamat, N.; Zaia, J.; Costello, C. E.; Lin, C.; Wei, J. Direct and Detailed Site-Specific Glycopeptide Characterization by Higher-Energy Electron-Activated Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2024**, *96* (3), 1251–1258. https://doi.org/10.1021/acs.analchem.3c04484.
(67) Macauslane, K. L.; Pegg, C. L.; Nouwens, A. S.; Kerr, E. D.; Seitanidou, J.; Schulz, B. L. Electron-Activated Dissociation and Collision-Induced Dissociation Glycopeptide Fragmentation for Improved Glycoproteomics. *Anal. Chem.* **2024**, *96* (27), 10986–10994. https://doi.org/10.1021/acs.analchem.4c01450.
(68) Helms, A.; Escobar, E. E.; Vainauskas, S.; Taron, C. H.; Brodbelt, J. S. Ultraviolet Photodissociation Permits Comprehensive Characterization of O-Glycopeptides Cleaved with O-Glycoprotease IMPa. *Anal. Chem.* **2023**, *95* (24), 9280–9287. https://doi.org/10.1021/acs.analchem.3c01111.
(69) Chongsaritsinsuk, J.; Rangel-Angarita, V.; Mahoney, K. E.; Lucas, T. M.; Enny, O. M.; Katemauswa, M.; Malaker, S. A. Quantification and Site-Specific Analysis of Co-Occupied N- and O-Glycopeptides. bioRxiv July 7, 2024, p 2024.07.06.602348. https://doi.org/10.1101/2024.07.06.602348.
(70) Khatri, K.; Pu, Y.; Klein, J. A.; Wei, J.; Costello, C. E.; Lin, C.; Zaia, J. Comparison of Collisional and Electron-Based Dissociation Modes for Middle-Down Analysis of Multiply Glycosylated Peptides. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1075–1085. https://doi.org/10.1007/s13361-018-1909-y.

FIGURES

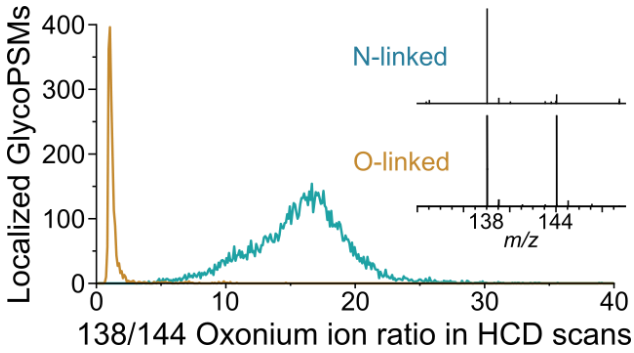

Figure 3. Oxonium ion ratios can indicate the glycopeptide type and may help select a specific dissociation method. A simple ratio between *m/z* 138/144 oxonium ions can indicate whether a glycopeptide likely belongs to the *N*-linked (containing high degrees of GlcNAc) or mucin-type *O*-linked (containing little GlcNAc, but mostly GalNAc) class. This data represents *O*-glycopeptides with predominantly core-1 mucin-type *O*-glycans, the focus of the *O*-glycopeptides in this study.

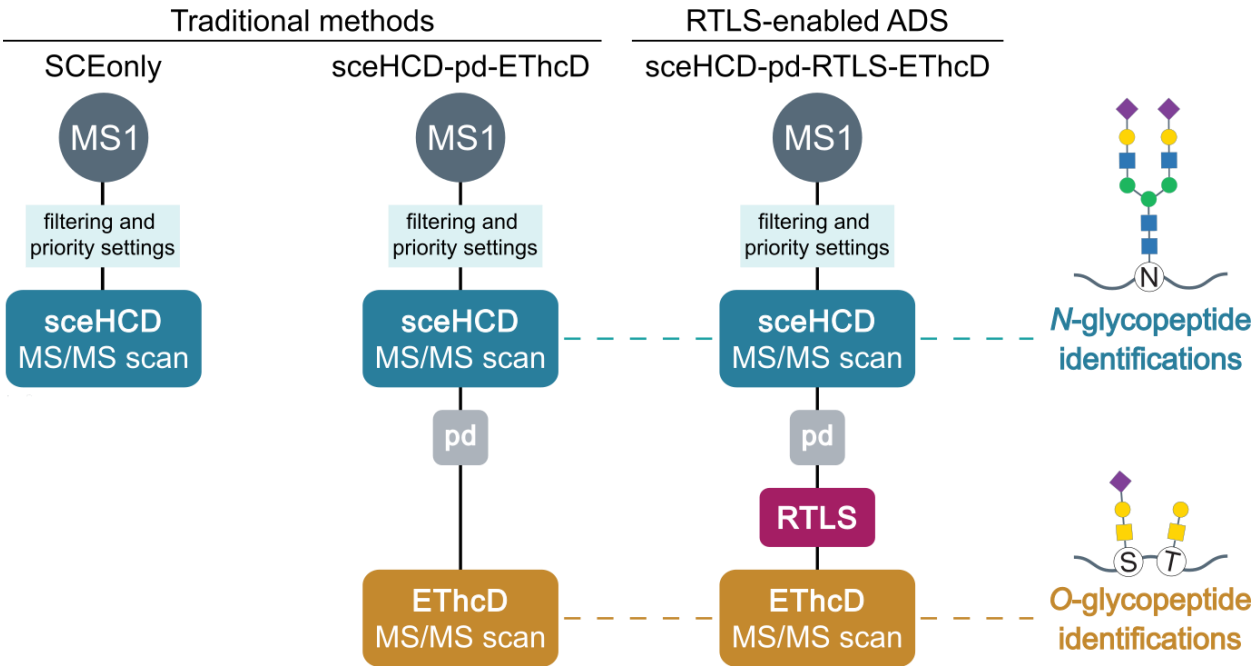

Figure 4. Conventional MS methods for glcyoproteomics compared to the real-time libray serach (RTLS)-enabled autonomous dissociation-type selection (ADS) methods introduced in this study. Glycoproteomic studies often rely on purely collision-based dissociation (secHCD-only) methods or product-dependent triggering methods that combine collision- and electron-based dissociation (sceHCD-pd-EThcD). ADS methods insert a RTLS function as an additional decision point between the product-dependent filter that purely evaluates the presence of oxonium ions. The RTLS service returns a match to *N*- or *O*-glycopeptide spectra, allowing the instrument to collect a subsequent MS/MS scan based on user-defined parameters. sceHCD scans are generally useful for *N*-glycopeptide identifications while electron-based dissociation is necessary for site-specific analysis of *O*-glycopeptides. Conventional sceHCD-pd-EThcD triggers EThcD for all glycopeptides, regardlesss of class. In this study, we used RTLS to classify likely glycopeptide precursor ions as *N*- or *O*-linked to acquire EThcD scans only for likely *O*-glycopeptides.

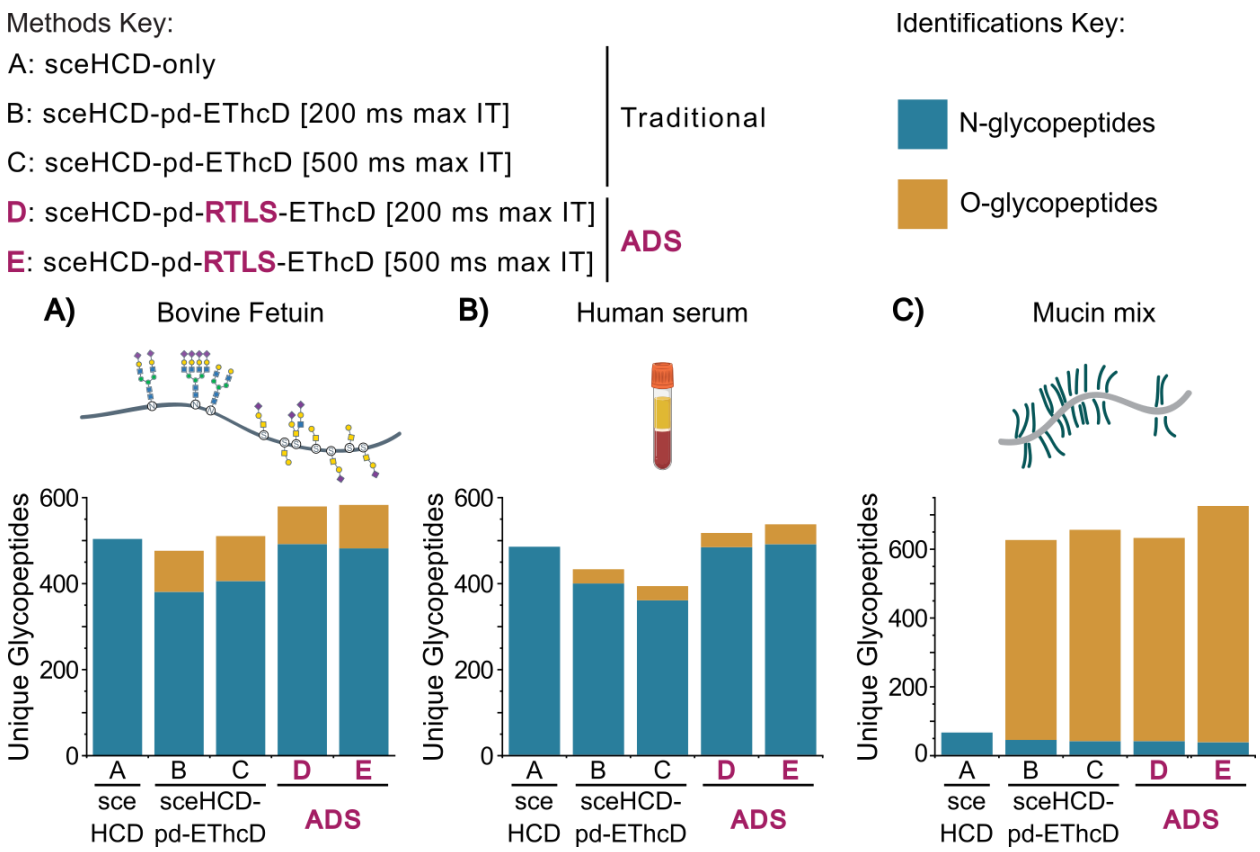

Figure 3. Performance of ADS methods relative to conventional glycorpteomic methods for glycopeptides from three different sources. Five methods (A-E) were compared, including one sceHCD-only (A), two sceHCD-pd-EThcD (B and C), and two ADS (i.e., sceHCD-pd-RTLS-EThcD) methods (D and E). Methods with EThcD scans included a maximum injection time (max IT) set to either 200 ms or 500 ms for each precursor selected for EThcD MS/MS. These methods were compared across glycopeptides from **A)** bovine fetuin, **B)** herum serum, and **C)** a mixture of mucin-domain glycoproteins. Navy and teal bars indicate the number of unique localized *N*- and *O*-glycopeptides identified, respectively, with ADS methods highlighted in orange text. All data reported are averages from two technical injection replicates.

For TOC only