Journal of proteome research

pubs.acs.org/jpr

Article

Hydrophilic Interaction Chromatography Coupled to Ultraviolet Photodissociation Affords Identification, Localization, and Relative Quantitation of Glycans on Intact Glycoproteins

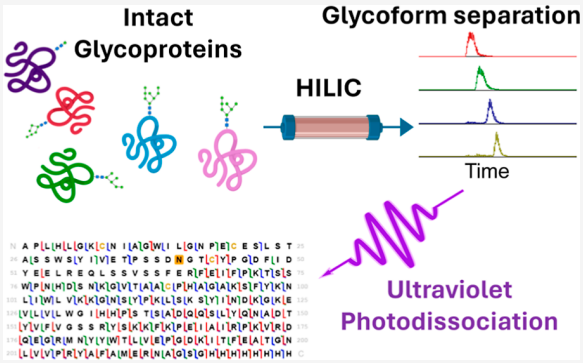
Virginia K. James, Annika A. M. van der Zon, Edwin E. Escobar, Sean D. Dunham, Andrea F. G. Gargano, and Jennifer S. Brodbelt*

Cite This: *J. Proteome Res.* 2024, 23, 4684−4693

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Protein glycosylation is implicated in a wide array of diseases, yet glycoprotein analysis remains elusive owing to the extreme heterogeneity of glycans, including microheterogeneity of some of the glycosites (amino acid residues). Various mass spectrometry (MS) strategies have proven tremendously successful for localizing and identifying glycans, typically utilizing a bottom-up workflow in which glycoproteins are digested to create glycopeptides to facilitate analysis. An emerging alternative is top-down MS that aims to characterize intact glycoproteins to allow precise identification and localization of glycans. The most comprehensive characterization of intact glycoproteins requires integration of a suitable separation method and high performance tandem mass spectrometry to provide both protein sequence information and glycosite localization. Here, we couple ultraviolet photodissociation and hydrophilic interaction chromatography with high resolution mass spectrometry to advance the characterization of intact glycoproteins ranging from 15 to 34 kDa, offering site localization of glycans, providing sequence coverages up to 93%, and affording relative quantitation of individual glycoforms.



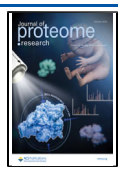
KEYWORDS: *top-down, mass spectrometry, ultraviolet photodissociation, hydrophilic interaction chromatography, glycoprotein, glycan, glycosylation*

INTRODUCTION

Protein glycosylation has emerged as one of the most prevalent, heterogeneous, and dynamic types of post-translational modifications (PTMs), modulating protein structure, function, and interactions.[1] Dysregulation or aberrant glycosylation has been implicated in many diseases, including cancer,[2−4] both neurodegenerative[5−7] and autoimmune diseases,[8,9] among others. Comprehensive analysis of glycoproteins remains challenging owing to their extreme heterogeneity in glycan size and identity.[10] Other PTMs, such as phosphorylation and acetylation, do not consist of repeating, branching or variable units, thus typically resulting in less heterogeneity.[11,12] Conversely, glycans that decorate proteins are composed of multiple saccharides, commonly glucose, galactose, mannose, *N*-acetylglucosamine, and sialic acid, among others, which can be arranged in different orders and branching patterns, each of which has the potential to code for different functions or to modify structures.[13] In addition to the sheer complexity of the glycan compositions, the glycans may be attached at different or multiple sites along the protein sequence, further compounding heterogeneity. Unraveling glycosylation sites and identifying the glycans remain significant analytical challenge.

Mass spectrometry and tandem mass spectrometry have emerged as powerful techniques to characterize glycoproteins. Glycomics, which involves enzymatic cleavage of the glycans from the proteins and analysis of the collection of glycans, provides extensive compositional information but no insights into the specific glycosylation sites of proteins.[14] An alternative approach entails proteolysis of the protein to create peptides followed by analysis of the resulting glycopeptides, often coupled with an enrichment step to enhance the detection of low abundance glycopeptides relative to the enormous population of nonglycosylated peptides.[15−17] This method, known as bottom-up glycoproteomics, provides extensive information about glycan location

Received: July 16, 2024
Revised: September 4, 2024
Accepted: September 10, 2024
Published: September 23, 2024



and site occupancy[18,19] and remains the gold standard for glycoproteomics workflows. Methods that analyze intact proteins offer the potential to map glycoforms, i.e., a depiction of all of the glycosites and their occupancies in proteins. However, the intact glycoprotein strategy requires appropriate separation methods[20] and high-level data processing algorithms[21] for successful translation to high throughput analysis of heterogeneous mixtures. Top-down MS methods can map coexisting glycans, thus revealing proteoform-level information,[22] even in a quantitative fashion in some cases.[23,24] However, relatively few studies have reported top-down analysis of glycoproteins, an endeavor that requires advanced activation techniques[25] and advanced separation techniques,[26] or may utilize partial enzymatic deglycosylation[27−29] to mitigate the analytical challenge presented by glycan heterogeneity which creates an overlap of glycoproteoforms in the m/z domain.[30]

One of the major ongoing challenges in the top-down proteomics workflow is the successful separation of glycoproteins. Traditional reversed phase liquid chromatography (RPLC) methods (C18, C4, PLRP, and others) are suitable for the separation of cleaved permethylated glycans,[31] small glycopeptides,[32−34] and some intact proteins[20] but do not always provide sufficient resolution of larger glycans and glycosylated proteins. Capillary electrophoresis (CE), which involves the separation of analytes based on their electrophoretic mobilities, has shown some success in separating intact glycoproteins,[35,36] along with cleaved glycans.[37] Hydrophilic interaction chromatography (HILIC) has been widely adopted for the separation of cleaved glycans,[38] and glycopeptides,[39−41] and more recently has been adapted for intact glycoproteins.[20,42] Recent advances in HILIC stationary phases have demonstrated especially promising performance metrics for glycoproteins, surpassing the separations achieved by RPLC and CE.[20] For example, the aforementioned HILIC study used a monolith polymer HILIC stationary phase[43,44] to provide sufficient separation of the extremely heterogeneous SARS-CoV-2 spike protein receptor binding domain (RBD) and assigned glycoforms based on intact mass, albeit without top-down MS/MS data to confirm glycoform assignments.[20] Gas-phase charge reduction strategies have also shown some success in dispersing glycoforms in m/z space, thus enabling differentiation of glycoforms.[45,46] Ion mobility, a technique that separates ions based on their shape and charge in the gas phase, has also been implemented to facilitate analysis of glycoproteins and glycopeptides.[26−28]

Sophisticated tandem mass spectrometry methods are also essential for the characterization of glycoproteins.[47] Traditional collisional activated dissociation (CAD) methods are widely used to analyze glycopeptides, affording glycan compositional and peptide sequence information.[19,48,49] Electron-based dissociation methods (ETD, ECD, EThcD) have been successful in localizing glycans on peptides[18,50−52] and small intact glycoproteins[25] and have been adapted for the characterization of the O-linked glycoforms of the RBD.[27] Electron-based and collision-based MS/MS methods may be paired to provide peptide sequence, glycan identification, and glycosite localization of glycopeptides[52−54] and have been used in an integrated manner for countless high throughput bottom-up glycoproteomics studies.[17,18,48,55] For top-down analysis of glycoproteins, the use of high-performance mass spectrometers is essential for the accurate identification of fragment ions, as demonstrated in a recent study that combined CAD and ECD to achieve characterization of glycan structures, glycosite locations, and some of the O-glycoform microheterogeneity of the SARS-CoV-2 spike protein RBD, albeit without chromatographic separation of glycoforms.[27]

An alternative high-energy activation method, ultraviolet photodissociation (UVPD), has shown success in identifying and localizing glycans on peptides.[56−59] UVPD has also been adapted for top-down analysis of proteins in an array of studies, typically yielding higher sequence coverages than other MS/MS methods and moving closer to the goal of complete proteoform analysis.[60−63] The successful inroads of UVPD for top-down analysis motivated our interest in developing a workflow for analyzing intact glycoproteins, among the most challenging targets, owing to the complexity and diversity of the glycan modifications. Herein, we describe a workflow where HILIC is coupled to UVPD-MS to localize glycan sites, identify glycan composition at each site, and provide relative quantitation of glycoforms, as demonstrated for RNase B, partial length hemagglutinin A, and a collection of the SARS-CoV-2 spike protein RBD variants with multiple glycosites.

MATERIALS AND METHODS

Materials and Samples

Water, acetonitrile, formic acid, tris(2-carboxyethyl)phosphine (TCEP), trypsin, dithiothreitol (DTT), iodoacetamide (IAM), and trifluoroacetic acid (TFA) (LC−MS grade) were obtained from Thermo Fisher Scientific. RNase B and PNGase F (glycerol-free) were purchased from New England Biolabs. Partial-length hemagglutinin A1 (HA) was a gift from the Georgiou group (UT-Austin) and was produced as previously described.[64] SARS-CoV-2 spike protein RBD variants were obtained from BEI resources, and for most experiments, these samples were partially deglycosylated with 0.75 μL of PNGase F per 20 μg of RBD. All samples were reduced with 150−200 mM TCEP for 2 h at 37 °C. The WT RBD sample was reduced with DTT (5 mM), alkylated with IAM (15 mM) for 30 min at room temperature, and digested overnight with trypsin at 37 °C for bottom-up analysis. All protein sequences, masses, and vendors/catalogue numbers (when applicable) are given in Table S1. All glycans were abbreviated as follows: hexose (H), *N*-acetylhexosamine (N), fucose (F), and sialic acid (S), and the number after each letter which indicates the number of each unit contained within the glycan.

Liquid Chromatography with Tandem Mass Spectrometry

All liquid chromatography experiments were performed on an UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) in trap and elute mode. Trap columns were packed in-house with C4 (5 μm, 120 Å, Dr. Maisch, Inc.) to ∼3 cm length with 100 μM ID and held at 50 °C for the entirety of all experiments. Nano HILIC analytical columns (20 cm length and 200 μm ID) were prepared as previously described[20,43,65] and held at room temperature. Proteins were diluted to 50−500 ng/μL in water with 0.1% formic acid, and 1 μL of the solution was injected per experiment. Water with 0.1% formic acid was used as the loading solvent. Analytical mobile phases consisted of (A) 98% water and 2% acetonitrile, (B) 98% acetonitrile and 2% water, each acidified with 0.1% formic acid and 0.05% TFA. Gradients are varied by type of sample and are listed in Tables S2−S4.

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

Traditional reversed phase chromatography was also used to compare it to HILIC and analyze a standard tryptic digest of the WT RBD sample. PLRP-S packing material (5 μm, 1000 Å, Agilent) and C18 (3 μm, 120 Å, NanoLCMS Solutions) were used for both the trap and analytical columns for HILIC comparison and tryptic digest, respectively. The analytical and trap columns had an ID of 75 and 100 μm, respectively. The columns were packed to lengths of ∼5 cm (trap) and ∼20 cm (analytical). The trap and analytical columns were held at 40 °C and room temperature for all experiments, respectively. The same loading and analytical mobile phases were used for the HILIC method as described above, except that the TFA was omitted. The gradient is given in Table S5 for all experiments. Apart from the fully glycosylated wildtype RBD experiments, all experiments were performed using an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) equipped with a 193 nm Coherent ExciStar XS excimer laser (Santa Clara, CA) to perform UVPD in the low-pressure linear ion trap, as previously described.[66] To characterize the fully glycosylated wild-type RBD, an Orbitrap Fusion Eclipse mass spectrometer (Thermo Fisher Scientific), also equipped with a 193 nm excimer laser, was used. A 15 μm ID emitter (New Objective), 35 V of in-source collision activation, and a spray voltage of 1800 V were used. For the relative quantitation of glycoforms, technical triplicates were performed while collecting only MS1 spectra at 15,000 resolution with 10 microscans and a 3×10^6 AGC target. UVPD analysis was carried out with targeted runs of a single charge state for each glycoform, during which MS1 spectra were collected at 15,000 resolution with two microscans and a 3×10^6 AGC target. UVPD spectra were collected at 240,000 resolution, 50−100 microscans, and a 1×10^6 AGC target or a maximum ion injection time of 500 ms. The most abundant charge state of each glycoform was targeted for UVPD. For the bottom-up glycoproteomics approach, UVPD scan events were triggered by glycan-specific ions in survey higher-energy collisional dissociation (HCD) scan events, as previously described.[59]

Data Processing

All MS2 data were processed with Prosight Native/TD validator[67] with a mass calibration of 0−10 ppm and default additional settings, including a S/N threshold of 3, a cluster tolerance of 0.35, an error tolerance of 10 ppm, a score of 0.5, and a minimum fragment residue length of 2, to confirm isotopic fits of UVPD fragments. Average mass of each glycoform was determined by deconvoluting low resolution MS1 spectra using Unidec.[68] Glycan mass was determined using a NIST glycan mass calculator (https://www.nist.gov/static/glyco-mass-calc/). Quantitation was performed based on creating extracted ion chromatograms (EICs) for the most abundant charge state of each glycoform with a 1 m/z window in QualBrowser, and peaks of each EIC were integrated using the Genesis algorithm. All peaks were smoothed with 7−11 boxcar averages in QualBrowser. All fragment ion identifications and glycoform relative abundance used for quantitation are available in the Supporting Information. Glycoforms for the multiply glycosylated SARS-CoV-2 spike protein RBD were identified with a custom Matlab script and data available from previous glycoproteomics studies on the RBD.[20] All data are available in the public repository jPOST with the accession number JPST002959.

RESULTS AND DISCUSSION

HILIC-UVPD Facilitates Characterization of Protein Sequence and Glycosylation Sites

Development and optimization of the HILIC-UVPD strategy were benchmarked using a well-characterized glycoprotein, RNase B, which is a 15 kDa glycoprotein with a single N-linked glycosite (N34) and eight abundant glycoforms. As previously established,[20,43] the HILIC monolith column provides excellent separation of glycoforms as evidenced by the near baseline chromatographic resolution of four glycoforms in Figure S1a,b. The companion MS1 spectra (Figure S2) also show good separation of each glycoform. We compared the performance of the HILIC separation to a conventional reversed phase separation using the PRLP stationary phase (the one most commonly employed for liquid chromatography−mass spectrometry analysis of intact proteins), the latter of which did not resolve any of the glycoforms (Figure S1c,d). The HILIC separation provided sufficient resolution to isolate individual glycoforms for subsequent MS/MS analysis. The performance of four MS/MS methods, HCD (a beam-type CAD), ETD, hybrid EThcD, and 193 nm UVPD, was assessed for characterization of RNase B, as summarized via the sequence maps in Figure S3 for the glycoform containing penta-mannose. HCD afforded low sequence coverage (39%) of RNase B and yielded few backbone cleavages close to the glycosylation site (Figure S3a), preventing confident localization of the known N34 glycosite. ETD and EThcD offered substantial improvement, yielding 76% and 82% sequence coverage of RNase B, respectively, and both methods produced fragment ions bracketing the glycosylation site (Figure S3b,c). UVPD offered the highest sequence coverage (93%), with numerous fragment ions retaining the glycan and bracketing the N34 site (Figure S3d). A representative UVPD mass spectrum is shown in Figure S4, along with examples of isotopic fits of fragment ions to confirm their assignments. These results motivated the adoption of UVPD for the remainder of the study.

The HILIC method provided sufficient chromatographic resolution to allow the relative quantitation of the RNase B glycoforms. For this strategy, UVPD was used to confidently identify the glycoforms using 50−100 scan averages for the acquisition of MS/MS data, and separate HILIC runs were used for quantitation based on integration of peak areas of the EICs of the most abundant charge state of each intact glycoform (intact masses listed in Table S6). The UVPD runs were not used for quantitative analysis owing to the distorted peak shape caused by the extensive averaging needed for MS/MS analysis (Figure S5, Table S7). The relative abundances of four glycoforms are summarized in Figure S6, consistent with those reported based on bottom-up analysis of RNase B glycopeptides (Table S8).[69,70] Technical replicates demonstrated low run-to-run standard deviations of the chromatographic peak areas (Figure S6). While UVPD can be influenced by the charge state of the precursor ion,[71−73] for this study, the most abundant charge state was selected for UVPD of each glycoform to maximize the signal intensity during the HILIC runs. Blanks run between samples ensured a lack of carryover.

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

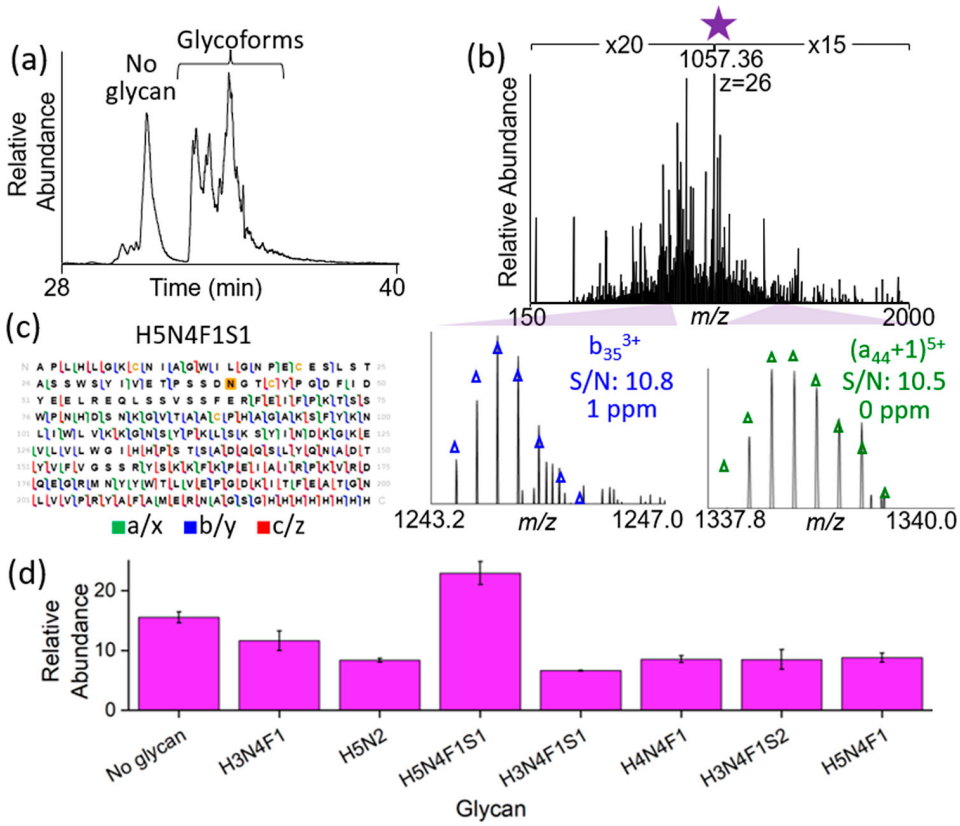


Figure 1. (a) HILIC chromatogram of HA with extracted ion chromatograms and MS1 spectra of glycoforms shown in Figure S8 and average masses of the proteins in Table S9. (b) UVPD mass spectrum (1 pulse, 2 mJ) of the most abundant glycoform of HA (H5N4F1S1, m/z 1057, 26+, t_4 = 34.0 min, marked with a star) with insets showing two glycan-localizing sequence ions. The UVPD fragment ions were used to generate the (c) sequence coverage map. Sequence coverage maps of additional glycoforms are given in Figure S9, and all identified fragment ions are listed in Table S7. (d) Comparison of the relative abundance of glycoforms with error bars representing the standard deviation from three replicates (specific values in Table S10).

Online HILIC UVPD Elucidates Glycan Location and Identity on Complex Viral Glycoproteins

In the previous RNase B example, only a few abundant glycoforms occupied the m/z domain (Figure S2a, top MS1 spectrum). However, many larger glycoproteins contain numerous different glycans located at the same or different glycosites, such as those observed for partial-length HA, an influenza antigen. Partial length HA is a 25 kDa protein with several possible N-linked glycosylations and glycosites. For more massive proteins that are highly charged, each charge state is more closely spaced in the m/z domain, resulting in overlapping charge state distributions of proteoforms, particularly those with similar masses. Without chromatographic separation of proteoforms, spectral congestion is a significant obstacle, as exemplified by the high-resolution mass spectrum of HA shown in Figure S7. The assignment of charge states is confounded, and isolation of individual glycoforms for MS/MS analysis is thwarted. HILIC provided ample separation of HA glycoforms and nonglycosylated HA (Figures 1a, S8a, Table S9), resulting in simplified time-resolved MS1 spectra containing fewer proteoforms (Figure S8b,c) and facilitating the acquisition of MS/MS spectra.

UVPD mass spectra were acquired for all HA glycoforms above 5% relative abundance, with 8 glycoforms in total. The UVPD mass spectrum of the most abundant glycoform (H5N4F1S1) is shown in Figure 1b, displaying numerous fragment ions that localize the glycan, such as $(b_{35})^{3+}$ and $(a_{44} + 1)^{5+}$ ions, with isotopic profiles shown beneath the UVPD mass spectrum. The sequence map is displayed in Figure 1c, resulting in 77% sequence coverage for the H5N4F1S1 glycoform. HILIC-UVPD of 6 other HA glycoforms plus the nonglycosylated protein yielded sequence coverages ranging from 67% to 82% (see sequence maps in Figure S9). The glycan was localized to N40 for all the HA glycoforms (H1N1 A/California/04/2009 strain) based on specific fragment ions that bracketed the glycosite at N40 (corresponding to N97 of full-length HA1). Following the same HILIC quantitation strategy described above, the relative abundances of the HA glycoforms were determined, as summarized in Figure 1d and Table S10. The identification of several abundant glycoforms (10−20% relative abundance) confirmed the high heterogeneity of this glycosite. As HA glycosylation patterns are known to vary with influenza strain,[74] the identification, localization, and relative quantitation of HA glycans provide insights into the evolution of the influenza virus, critical for the development of future vaccines and antiviral therapeutics.

In addition to HA, we investigated the glycosylation of the SARS-CoV-2 spike protein RBD, another viral glycoprotein. The RBD is a 25 kDa immunogenic fragment of the larger spike glycoprotein, typically decorated with approximately 9 kDa of glycans among two N- (N13 and N25) and one O- (full length T323, RBD T5) glycosites.[20,45] We first focused on the characterization of the glycoforms by using PNGase F

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

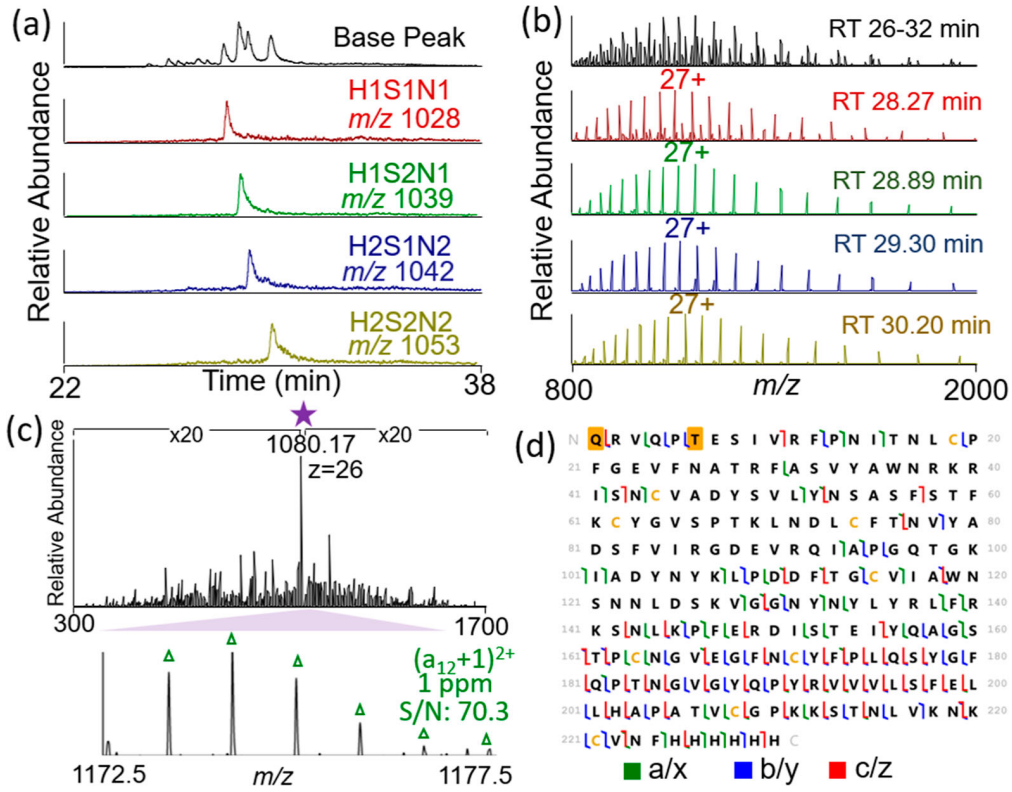


Figure 2. (a) Base peak chromatogram and extracted ion chromatograms of each O-linked glycoform of the SARS-CoV-2 spike protein RBD and (b) corresponding MS1 spectra acquired at specific retention times with full range MS1 spectra and deconvoluted spectra shown in Figure S12. Average masses of the intact proteins are listed in Table S11. (c) UVPD (1 pulse, 2 mJ) mass spectrum of H1S2N1 glycoform (m/z 1080, 26+ charge state) with an expanded view of one site-localizing fragment ion, $(a_{12} + 1)^{2+}$, that contains both the glycan (T6) and the additional modification (addition of pyroGlu (+111 Da), gold-shaded Q) at the N terminus. The precursor is marked with a purple star, and all other ions are sequence ions. (d) Sequence coverage map derived from the UVPD spectrum. The locations of the modifications are shaded in gold. UVPD sequence coverage maps for additional glycoforms are shown in Figure S11, and fragment ion identifications for all glycoforms are provided in Table S7.

to remove the N-linked glycans. O-linked glycans are often the most challenging for identification and localization as they are more labile and may be lost during conventional CAD. However, the fast, high energy activation process of UVPD allows retention of glycans during backbone cleavages of proteins.[59] HILIC provided excellent separation of the O-linked glycoforms of the WT RBD (Figure 2a), as evidenced by the MS1 spectra for each glycoform (Figure 2b). We first localized a nonglycan modification to the N-terminus of the protein, which consists of pyroGlu (+111 Da), a leftover residue from the signal peptide. This modification was confirmed by the analysis of a tryptic digest of the RBD protein (Figure S10) and is consistent with a previous report.[75] UVPD fragment ions of the H1S2N1 glycoform (Figure 2c) confirmed the location of the O-linked glycan as T6 when considering the addition of one amino acid pyroGlu to the N terminus, T5 without Q1 (Figure 2d). The detection of fewer sequence ions that bracket the glycan (i.e., short N-terminal or very long C-terminal fragment ions) is attributed to partial glycan dissociation, making it more challenging to search for and assign the glycan-modified fragment ions with confidence. A plethora of fragment ions characterize the C-terminal region of the protein, and the overall sequence coverage is 51% as reflected in the sequence map in Figure 2d. UVPD fragment ion identifications for all glycoforms (H1S1N1, H2S1N2, H1S2N1, and H2S2N2) are provided in the Supporting Information (Table S7), and the corresponding sequence coverage maps for the H1S1N1, H2S1N2, and H2S2N2 glycoforms are shown in Figure S11. UVPD resulted in sequence coverages ranging from 28% to 51% for each glycoform.

Comparing Glycosylation across Glycoprotein Variants

As the COVID-19 pandemic has evolved, the emergence of new viral variants remains a primary concern, and elucidating changes in glycosylation patterns in the SARS-CoV-2 proteome, such as the mutational hot spot spike protein, may help correlate the impact of mutations on transmissibility and virulence.[76] Here, we compare five RBD variants containing different point mutations (Table S1). The N-linked glycans were removed using PNGase treatment to allow the O-glycosylation sites to be targeted. HILIC-MS provided good separation and MS1 analysis of the O-linked glycoforms of each variant (see base peak ion chromatograms, MS1 spectra in Figures S13−S17, and intact masses in Table S11). UVPD spectra were collected for each O-glycoform, and the corresponding fragment ion identifications are provided in Table S7, along with the resulting sequence maps in Figures S18−S22. The UVPD data confirmed that all RBD samples have the same O-linked glycosite, T6 or T5, for the WT and variant RBD samples, respectively. The distributions of O-glycoforms for each variant and the WT protein are summarized in Figure 3. The identities of the RBD glycans are similar to those reported previously,[27,28] and

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

the specific glycosylation patterns of each variant are generally similar except for the E484 K variant for which the H1S1N1 glycan was not detected. Interestingly, none of the RBD variants contained the H2S1N2 glycan, which was identified for only the wild-type RBD (Figure 3). All variants and the WT RBD share the same most abundant glycan (H1S2N2); however, the WT RBD exhibited a much lower abundance of this glycoform (39%) than the variant samples (average 74%) and literature values (65%).[27,28] This trend continues with the lower abundance glycoforms, H1S1N1 and H2S2N2, as the variants agree more closely with literature values than the WT which shows higher abundances of H1S1N1 and H2S2N2 (Table S12). Given that the RBDs examined here and elsewhere[27,28] are products of overexpression, it must be considered that the expression cell lines (HEK293 for WT, HEK293T for variants) or other expression conditions may impact glycosylation patterns.

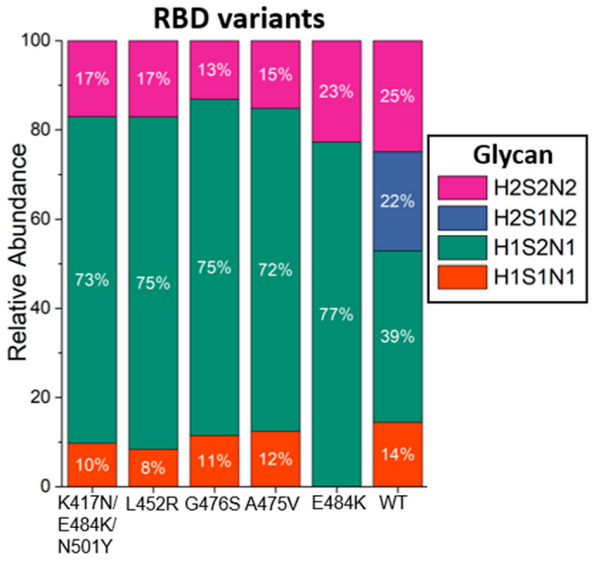


Figure 3. Relative abundances of each glycoform. Chromatograms and MS1 spectra for the RBD variant samples are shown in Figures S13 and S17, average masses of the intact proteins are shown in Table S11, and sequence coverage maps generated for each glycoform are shown in Figures S18−S22.

Online HILIC UVPD Enables Characterization of Multiply Glycosylated Proteins

The most compelling attribute of top-down analysis of proteins is the ability to map multiple modification sites in context, revealing combinatorial patterns of post-translational modifications. Glycosylated proteins often have multiple N- and O- glycosylation sites, creating extremely heterogeneous sets of proteoforms even for relatively small proteins like the 30−34 kDa RBD. HILIC-MS was used to separate fully glycosylated RBD (Figure 4a), a protein exhibiting high heterogeneity, which confounds the identification of specific glycoforms (Figure 4b). Although HILIC did not provide baseline separation of each glycoform, it offered sufficient chromatographic separation when implemented along with the isolation of narrow m/z fractions of the MS1 spectra to winnow the number of glycoforms selected, enabling the confident selection of individual glycoforms.

By combining the set of putative glycoforms reported in a previous study[20] and the UVPD data acquired here, we

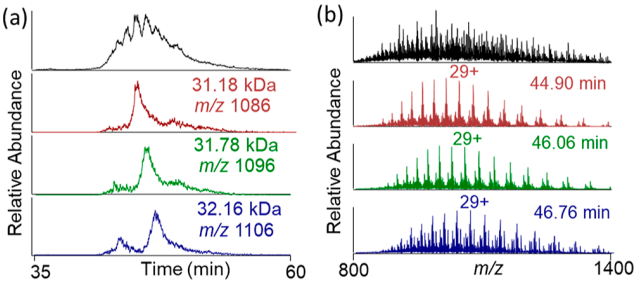


Figure 4. (a) Chromatograms including a base peak (top trace) and EICs of the most abundant glycoforms of the WT RBD. (b) MS1 spectra acquired at the specified retention time.

identify the three most abundant glycoforms found for the fully glycosylated wildtype RBD (Figure 5). For each selected glycoform, fragment ions generated by UVPD enabled the distinction of the glycans from each other. In the case of the T6 (H1S2N1), N14 (N5H6F3S1), and N26 (N5H4F1) glycoforms (Figure 5a), the accurate mass assignment of the a_{12}^{2+} ion identifies the T6 glycosite and the dual glycosylation of the N14 and N26 sites. For the T6 (H1S2N1), N14 (H6N9F1), and N26 (H5N5) (Figure 5b) and T6 (H1S1N1), N14 (H5N4F3), and N26 (H6N5F2S2) (Figure 5c) glycoforms, the $(y_{226} - 1)^{28+}$ and y_{229}^{29+} ions retain all three glycosites. While these data show the potential of HILIC-UVPD to distinguish glycan composition at different glycosites, sequence information was relatively low for the glycan-containing N-terminus, preventing confident glycan localization. As the fully glycosylated protein contains approximately 6 kDa of glycans, the likelihood of glycan fragmentation concurrent with protein fragmentation increases, thus reducing the number of protein fragment ions with glycans intact.

■ CONCLUSIONS

Glycoproteins remain one of the most challenging analytical targets, and their complexity is reflected in their myriad biological functional and structural roles. Here, we investigated the glycan profiles of viral proteins by using an integrated HILIC-UVPD strategy. After separation of glycoforms by HILIC, UVPD was used to characterize individual glycoforms, and the relative abundances of glycoforms could be compared using their chromatographic peak areas. Sequence coverages up to 93% were obtained by UVPD, allowing the determination of glycosites. For partial length HA1 from influenza, the glycosite was localized to N40, and the glycan composition at this site was found to be heterogeneous with eight different glycans for the H1N1 strain analyzed here. O-glycoforms of the SARS-CoV-2 RBD were identified and found to be consistent with prior reports regarding glycan identification and abundance. We also investigated the O-glycoforms of point mutants of the RBD and found that those mutants retained the same glycosite but exhibited variations in the distribution and compositions of the glycan profiles. As spike protein mutations are common in emerging SARS-CoV-2 variants,[77−79] this array of RBD point mutants showcases the ability of HILIC-UVPD to characterize glycosylation patterns across the evolution of spike proteins. Finally, we advanced the characterization of multiply glycosylated species, in this case, the fully glycosylated RBD, allowing insights into glycosylations that may occur in tandem. While several heterogeneous glyco-

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

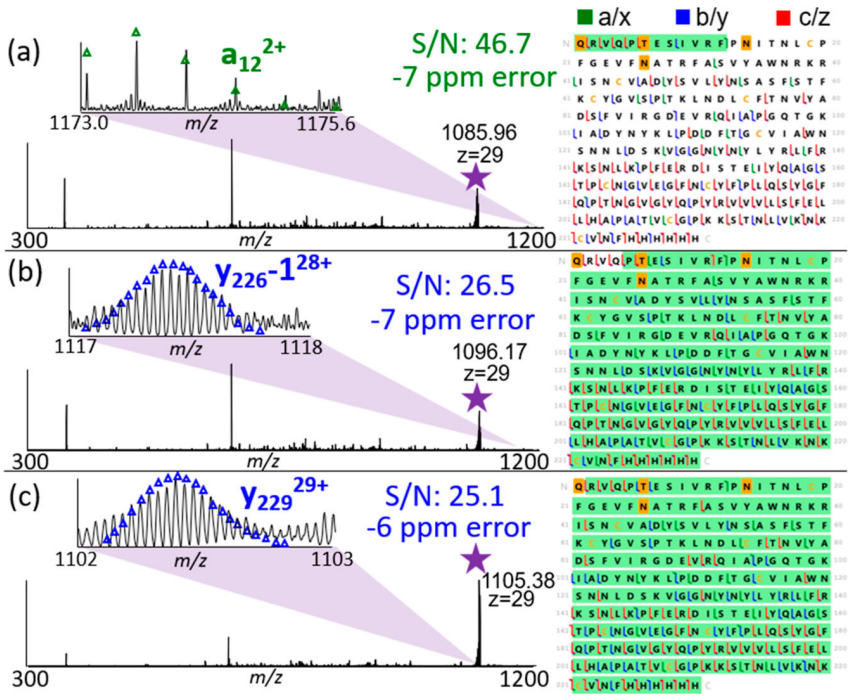


Figure 5. Localization of glycans in fully glycosylated wildtype RBD using UVPD (1 pulse, 0.5 mJ) for (a) T6 (H1S2N1), N14 (N5H6F3S1) and N26 (N5H4F1) glycoform (29+), (b) T6 (H1S2N1), N14 (H6N9F1), and N26 (H5N5) glycoform (29+), and (c) T6 (H1S1N1), N14 (H5N4F3), and N26 (H6N5F2S2) glycoform (29+). Each glycoform also contains a pyroGlu (+111 Da) modification at the N-terminus. The corresponding sequence maps are shown on the right. The locations of the modifications are shaded in gold, and the green shading corresponds to the section of the protein encompassed by the fragment ion expanded in each inset on the left. Lists of identified fragment ions are summarized in Table S7.

forms were studied here, protein glycosylation in biological systems can be exponentially more complex and the glycoproteins themselves may be far larger in size.[80] Further improvements in instrumentation to enhance the sensitivity of top-down methods and the development of more sophisticated data processing methods to decipher specific glycan compositions based on the assignment of glycan fragment ions are key future goals that will be critical for comprehensive top-down glycoproteoform analysis.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.4c00600.

Protein sequences, masses and catalogue numbers; 40 min HILIC gradient; 60 min HILIC gradient; 90 min HLIC gradient; PLRP gradient; average intact mass of RNase B glycoforms; relative abundances of RNase B glycoforms compared to reported values; average intact mass of HA glycoforms; relative abundances of HA glycoforms; average intact masses of RBD; relative abundances of RBD O-linked glycoforms compared to reported values; PLRP vs HILIC separations for RNase B; ESI mass spectra and deconvoluted mass spectra obtained for RNase B; sequence maps based on HCD, ETD, EThcD, and UVPD of RNase B; UVPD spectrum of RNase B and a few fragment identifications; comparison of chromatographic peak shapes for MS1 only vs UVPD run; relative abundance of RNase B glycoforms; ESI mass spectrum of HA collected via direct infusion; EICs and corresponding ESI mass spectra for each glycoform of HA; Sequence coverage maps of additional HA glycoforms derived from UVPD; UVPD spectrum of a tryptic peptide that confirms the N terminal modification of the WT RBD; sequence maps of each O-glycoform of WT RBD based on UVPD; full range and deconvoluted MS1 spectra of WT RBD; EICs for the most abundant charge states of each O-glycoform of K417N/E484 K/N501Y RBD and MS1 spectra corresponding to the retention time at the peak maximum; EICs for the most abundant charge states of each L452R RBD O-glycoform and MS1 spectra corresponding to the retention time at the peak maximum.; EICs for the most abundant charge states of each G476S RBD O-glycoform and MS1 spectra corresponding to the retention time at the peak maximum; EICs for the most abundant charge states of each A475 V RBD O-glycoform and MS1 spectra corresponding to the retention time at the peak maximum; EICs for the most abundant charge states of each E484 K RBD O-glycoform and MS1 spectra corresponding to the retention time at the peak maximum; sequence maps of each O-glycoform of K417N/E484 K/N501Y RBD based on UVPD; sequence maps of each O-glycoform of L452R RBD based on UVPD; sequence maps of each O-glycoform of G476S RBD based on UVPD; sequence maps of each O-glycoform of A475VRBD based on UVPD; and sequence maps of each O-glycoform of E484 K RBD based on UVPD (PDF)

Tables of identified fragment ions from MS2 experiments and relative quantitation of RNase B, HA, and the RBDs (XLSX)

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

AUTHOR INFORMATION

Corresponding Author

Jennifer S. Brodbelt − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0003-3207-0217; Email: jbrodbelt@cm.utexas.edu

Authors

Virginia K. James − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States*

Annika A. M. van der Zon − *van 't Hoff Institute for Molecular Science, University of Amsterdam, Amsterdam 1098 XH, The Netherlands; Centre of Analytical Sciences Amsterdam, Amsterdam 1098 XH, The Netherlands*

Edwin E. Escobar − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States*

Sean D. Dunham − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-6214-7468

Andrea F. G. Gargano − *van 't Hoff Institute for Molecular Science, University of Amsterdam, Amsterdam 1098 XH, The Netherlands; Centre of Analytical Sciences Amsterdam, Amsterdam 1098 XH, The Netherlands;* orcid.org/0000-0003-3361-7341

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.4c00600

Funding

Funding from the National Science Foundation (CHE-2203602) and the Robert A. Welch Foundation (F-1155) is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Mowei Zhou for insights about implementation of HILIC-MS. The gift of HA from the Georgiou lab (UT-Austin) is gratefully acknowledged. BEI resources: The following reagent was produced under HHSN272201400008C and obtained through BEI Resources, NIAID, NIH: Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, Wuhan-Hu-1 with C-Terminal Histidine Tag, Recombinant from HEK293T Cells, NR-52946. The following reagents were obtained through BEI Resources, NIAID, NIH: Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, L452R Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55403, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, G476S Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55401, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, K417N/E484 K/N501Y Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55414, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, A475 V Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55402, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, E484 K Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55400.

REFERENCES

(1) Conibear, A. C. Deciphering Protein Post-Translational Modifications Using Chemical Biology Tools. *Nat. Rev. Chem* **2020**, *4* (12), 674−695.

(2) Kailemia, M. J.; Park, D.; Lebrilla, C. B. Glycans and Glycoproteins as Specific Biomarkers for Cancer. *Anal. Bioanal. Chem.* **2017**, *409* (2), 395−410.

(3) Scott, D. A.; Drake, R. R. Glycosylation and Its Implications in Breast Cancer. *Expert Rev. of Proteomics* **2019**, *16* (8), 665−680.

(4) Silsirivanit, A. Glycosylation markers in cancer; Makowski, G. S., Ed.; Elsevier, 2019; Vol. *89*, pp 189−213..*Glycosylation Markers in Cancer*

(5) Chai, A. B.; Leung, G. K. F.; Callaghan, R.; Gelissen, I. C. P-Glycoprotein: A Role in the Export of Amyloid-β in Alzheimer's Disease? *FEBS J.* **2020**, *287* (4), 612−625.

(6) Diaz-Ortiz, M. E.; Seo, Y.; Posavi, M.; Carceles Cordon, M.; Clark, E.; Jain, N.; Charan, R.; Gallagher, M. D.; Unger, T. L.; Amari, N.; Skrinak, R. T.; Davila-Rivera, R.; Brody, E. M.; Han, N.; Zack, R.; Van Deerlin, V. M.; Tropea, T. F.; Luk, K. C.; Lee, E. B.; Weintraub, D.; Chen-Plotkin, A. S. GPNMB Confers Risk for Parkinson's Disease through Interaction with α-Synuclein. *Science* **2022**, *377* (6608), No. eabk0637.

(7) Xu, M.; Zhou, M.; Li, S.; Zhen, X.; Yang, S. Glycoproteins as Diagnostic and Prognostic Biomarkers for Neurodegenerative Diseases: A Glycoproteomic Approach. *J. Neurosci. Res.* **2021**, *99* (5), 1308−1324.

(8) Edavalath, S.; Rai, M. K.; Gupta, V.; Mishra, R.; Misra, D. P.; Gupta, L.; Agarwal, V. Tacrolimus Induces Remission in Refractory and Relapsing Lupus Nephritis by Decreasing P-Glycoprotein Expression and Function on Peripheral Blood Lymphocytes. *Rheumatol. Int.* **2022**, *42* (8), 1347−1354.

(9) Rudman, N.; Gornik, O.; Lauc, G. Altered N-Glycosylation Profiles as Potential Biomarkers and Drug Targets in Diabetes. *FEBS Lett.* **2019**, *593* (13), 1598−1615.

(10) Oliveira, T.; Thaysen-Andersen, M.; Packer, N. H.; Kolarich, D. The Hitchhiker's Guide to Glycoproteomics. *Biochem. Soc. Trans.* **2021**, *49* (4), 1643−1662.

(11) Ramazi, S.; Zahiri, J. Post-Translational Modifications in Proteins: Resources, Tools and Prediction Methods. *Database* **2021**, *2021*, baab012.

(12) Leutert, M.; Entwisle, S. W.; Villén, J. Decoding Post-Translational Modification Crosstalk With Proteomics. *Mol. Cell. Proteomics* **2021**, *20*, 100129.

(13) Čaval, T.; Heck, A. J. R.; Reiding, K. R. Meta-Heterogeneity: Evaluating and Describing the Diversity in Glycosylation Between Sites on the Same Glycoprotein. *Mol. Cell. Proteomics* **2021**, *20*, 100010.

(14) de Haan, N.; Yang, S.; Cipollo, J.; Wuhrer, M. Glycomics Studies Using Sialic Acid Derivatization and Mass Spectrometry. *Nat. Rev. Chem* **2020**, *4* (5), 229−242.

(15) Piovesana, S.; Cavaliere, C.; Cerrato, A.; Laganà, A.; Montone, C. M.; Capriotti, A. L. Recent Trends in Glycoproteomics by Characterization of Intact Glycopeptides. *Anal. Bioanal. Chem.* **2023**, *415* (18), 3727−3738.

(16) Chang, D.; Zaia, J. Methods to Improve Quantitative Glycoprotein Coverage from Bottom-up LC-MS Data. *Mass Spectrom. Rev.* **2022**, *41* (6), 922−937.

(17) 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 Primers* **2022**, *2* (1), 48.

(18) 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.

(19) Hwang, H.; Jeong, H. K.; Lee, H. K.; Park, G. W.; Lee, J. Y.; Lee, S. Y.; Kang, Y.-M.; An, H. J.; Kang, J. G.; Ko, J.-H.; Kim, J. Y.; Yoo, J. S. Machine Learning Classifies Core and Outer Fucosylation

of N-Glycoproteins Using Mass Spectrometry. *Sci. Rep.* **2020**, *10* (1), 318.
(20) Wilson, J. W.; Bilbao, A.; Wang, J.; Liao, Y.-C.; Velickovic, D.; Wojcik, R.; Passamonti, M.; Zhao, R.; Gargano, A. F. G.; Gerbasi, V. R.; Paša-Tolić, L.; Baker, S. E.; Zhou, M. Online Hydrophilic Interaction Chromatography (HILIC) Enhanced Top-Down Mass Spectrometry Characterization of the SARS-CoV-2 Spike Receptor-Binding Domain. *Anal. Chem.* **2022**, *94* (15), 5909−5917.
(21) Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; Heck, A. J. R. Hybrid Mass Spectrometry Approaches in Glycoprotein Analysis and Their Usage in Scoring Biosimilarity. *Nat. Commun.* **2016**, *7* (1), 13397.
(22) Rommelfanger, S. R.; Zhou, M.; Shaghasi, H.; Tzeng, S.-C.; Evans, B. S.; Paša-Tolić, L.; Umen, J. G.; Pesavento, J. J. An Improved Top-Down Mass Spectrometry Characterization of Chlamydomonas Reinhardtii Histones and Their Post-Translational Modifications. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (7), 1671−1688.
(23) Cupp-Sutton, K. A.; Wu, S. High-Throughput Quantitative Top-down Proteomics. *Mol. Omics* **2020**, *16* (2), 91−99.
(24) Lin, Z.; Wei, L.; Cai, W.; Zhu, Y.; Tucholski, T.; Mitchell, S. D.; Guo, W.; Ford, S. P.; Diffee, G. M.; Ge, Y. Simultaneous Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome. *Mol. Cell. Proteomics* **2019**, *18* (3), 594−605.
(25) Bourgoin-Voillard, S.; Leymarie, N.; Costello, C. E. Top-down Tandem Mass Spectrometry on RNase A and B Using a Qh/FT-ICR Hybrid Mass Spectrometer. *Proteomics* **2014**, *14* (10), 1174−1184.
(26) Miller, S. A.; Jeanne Dit Fouque, K.; Hard, E. R.; Balana, A. T.; Kaplan, D.; Voinov, V. G.; Ridgeway, M. E.; Park, M. A.; Anderson, G. A.; Pratt, M. R.; Fernandez-Lima, F. Top/Middle-Down Characterization of α-Synuclein Glycoforms. *Anal. Chem.* **2023**, *95* (49), 18039−18045.
(27) Roberts, D. S.; Mann, M.; Melby, J. A.; Larson, E. J.; Zhu, Y.; Brasier, A. R.; Jin, S.; Ge, Y. Structural O-Glycoform Heterogeneity of the SARS-CoV-2 Spike Protein Receptor-Binding Domain Revealed by Top-Down Mass Spectrometry. *J. Am. Chem. Soc.* **2021**, *143* (31), 12014−12024.
(28) Roberts, D. S.; Mann, M. H.; Li, B.; Kim, D. R.; Braiser, A.; Jin, S.; Ge, Y. Distinct Core Glycan and O -Glycoform Utilization of SARS-CoV-2 Omicron Variant Spike Protein RBD Revealed by Top-down Mass Spectrometry. *Chem. Sci.* **2022**, *13* (36), 10944−10949.
(29) Wu, D.; Robinson, C. V. Native Top-Down Mass Spectrometry Reveals a Role for Interfacial Glycans on Therapeutic Cytokine and Hormone Assemblies. *Angew. Chem.* **2022**, *134* (49), No. e202213170.
(30) Reid, D. J.; Thibert, S.; Zhou, M. Dissecting the Structural Heterogeneity of Proteins by Native Mass Spectrometry. *Protein Sci.* **2023**, *32* (4), No. e4612.
(31) Kurz, S.; Sheikh, M. O.; Lu, S.; Wells, L.; Tiemeyer, M. Separation and Identification of Permethylated Glycan Isomers by Reversed Phase NanoLC-NSI-MSn. *Mol. Cell. Proteomics* **2021**, *20*, 100045.
(32) Ji, E. S.; Lee, H. K.; Park, G. W.; Kim, K. H.; Kim, J. Y.; Yoo, J. S. Isomer Separation of Sialylated O- and N-Linked Glycopeptides Using Reversed-Phase LC-MS/MS at High Temperature. *J. Chromatogr. B* **2019**, *1110−1111*, 101−107.
(33) Chen, S.-Y.; Dong, M.; Yang, G.; Zhou, Y.; Clark, D. J.; Lih, T. M.; Schnaubelt, M.; Liu, Z.; Zhang, H. Glycans, Glycosite, and Intact Glycopeptide Analysis of N-Linked Glycoproteins Using Liquid Handling Systems. *Anal. Chem.* **2020**, *92* (2), 1680−1686.
(34) Yin, H.; Zhu, J. Methods for Quantification of Glycopeptides by Liquid Separation and Mass Spectrometry. *Mass Spectrom. Rev.* **2023**, *42* (2), 887−917.
(35) Jia, Y.; Cao, J.; Zhou, J.; Zhou, P. Methyl Chitosan Coating for Glycoform Analysis of Glycoproteins by Capillary Electrophoresis. *Electrophoresis* **2020**, *41* (9), 729−734.
(36) Zhang, C.; Schumacher, K. N.; Dodds, E. D.; Hage, D. S. Glycoprotein Analysis Using Lectin Microcolumns and Capillary Electrophoresis: Characterization of Alpha1-Acid Glycoprotein by Combined Separation Methods. *J. Chromatogr. B* **2021**, *1179*, 122855.
(37) Tomnikova, A.; Kozlík, P.; Křížek, T. Monosaccharide Profiling of Glycoproteins by Capillary Electrophoresis with Contactless Conductivity Detection. *Electrophoresis* **2022**, *43* (20), 1963−1970.
(38) Guo, L.; Nayak, S.; Mao, Y.; Li, N. Glycine Additive Enhances Sensitivity for N- and O-Glycan Analysis with Hydrophilic Interaction Chromatography-Electrospray Ionization-Mass Spectrometry. *Anal. Biochem.* **2021**, *635*, 114447.
(39) Huang, Y.; Nie, Y.; Boyes, B.; Orlando, R. Resolving Isomeric Glycopeptide Glycoforms with Hydrophilic Interaction Chromatography (HILIC). *J. Biomol. Tech.* **2016**, *27* (3), 98−104.
(40) Li, Y.; Guo, W.; Zhang, Q.; Yang, B.; Zhang, Y.; Yang, Y.; Liu, G.; Pan, L.; Zhang, W.; Kong, D. Improved Analysis ZIC-HILIC-HCD-Orbitrap Method for Mapping the Glycopeptide by Mass Spectrometry. *J. Chromatogr. B* **2023**, *1228*, 123852.
(41) Sran, K. S.; Sharma, Y.; Kaur, T.; Rao, A. Post-Translational Modifications and Glycoprofiling of Palivizumab by UHPLC-RPLC/HILIC and Mass Spectrometry. *J. Proteins Proteomics* **2022**, *13* (2), 95−108.
(42) Gargano, A. F. G.; Haselberg, R.; Somsen, G. W. Chapter 5 - Hydrophilic Interaction Liquid Chromatography-Mass Spectrometry for the Characterization of Glycoproteins at the Glycan, Peptide, Subunit, and Intact Level; El Rassi, Z., Ed.; Elsevier: Amsterdam, 2021; pp 209−278.*Carbohydrate Analysis by Modern Liquid Phase Separation Techniques (Second ed.)*
(43) Passamonti, M.; de Roos, C.; Schoenmakers, P. J.; Gargano, A. F. Poly(acrylamide-*co*-*N*,*N*′-methylenebisacrylamide) Monoliths for High-Peak-Capacity Hydrophilic-Interaction Chromatography−High-Resolution Mass Spectrometry of Intact Proteins at Low Trifluoroacetic Acid Content. *Anal. Chem.* **2021**, *93* (48), 16000−16007.
(44) Passamonti, M.; Zhai, Z.; Moreschini, M.; Wilson, J. W.; Zhou, M.; Schoenmakers, P. J.; Gargano, A. F. G. Influence of Ion-Pairing Reagents on the Separation of Intact Glycoproteins Using Hydrophilic-Interaction Liquid Chromatography - High-Resolution Mass Spectrometry. *J. Chromatogr. A* **2023**, *1688*, 463721.
(45) Yang, Y.; Du, Y.; Kaltashov, I. A. The Utility of Native MS for Understanding the Mechanism of Action of Repurposed Therapeutics in COVID-19: Heparin as a Disruptor of the SARS-CoV-2 Interaction with Its Host Cell Receptor. *Anal. Chem.* **2020**, *92* (16), 10930−10934.
(46) Schachner, L. F.; Mullen, C.; Phung, W.; Hinkle, J. D.; Beardsley, M. I.; Bentley, T.; Day, P.; Tsai, C.; Sukumaran, S.; Baginski, T.; DiCara, D.; Agard, N.; Masureel, M.; Gober, J.; ElSohly, A.; Melani, R.; Syka, J. E. P.; Huguet, R.; Marty, M. T.; et al. Exposing the Molecular Heterogeneity of Glycosylated Biotherapeutics. *Nat. Commun.* **2024**, *15*, 3259.
(47) 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.
(48) 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.
(49) Sun, S.; Hu, Y.; Ao, M.; Shah, P.; Chen, J.; Yang, W.; Jia, X.; Tian, Y.; Thomas, S.; Zhang, H. N-GlycositeAtlas: A Database Resource for Mass Spectrometry-Based Human N-Linked Glycoprotein and Glycosylation Site Mapping. *Clin. Proteomics* **2019**, *16* (1), 35.
(50) Halim, A.; Rüetschi, U.; Larson, G.; Nilsson, J. LC-MS/MS Characterization of O-Glycosylation Sites and Glycan Structures of Human Cerebrospinal Fluid Glycoproteins. *J. Proteome Res.* **2013**, *12* (2), 573−584.

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693

(51) Zhu, H.; Qiu, C.; Gryniewicz-Ruzicka, C. M.; Keire, D. A.; Ye, H. Multiplexed Comparative Analysis of Intact Glycopeptides Using Electron-Transfer Dissociation and Synchronous Precursor Selection Based Triple-Stage Mass Spectrometry. *Anal. Chem.* **2020**, *92* (11), 7547−7555.
(52) 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.
(53) 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.
(54) Mao, Y.; Wang, S.; Zhao, Y.; Konstantinidi, A.; Sun, L.; Ye, Z.; Vakhrushev, S. Y. Systematic Evaluation of Fragmentation Methods for Unlabeled and Isobaric Mass Tag-Labeled O-Glycopeptides. *Anal. Chem.* **2021**, *93* (32), 11167−11175.
(55) 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.* **2023**, *8*, 233−247.
(56) Escobar, E. E.; Wang, S.; Goswami, R.; Lanzillotti, M. B.; Li, L.; McLellan, J. S.; Brodbelt, J. S. Analysis of Viral Spike Protein N-Glycosylation Using Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2022**, *94* (15), 5776−5784.
(57) Ko, B. J.; Brodbelt, J. S. Comparison of Glycopeptide Fragmentation by Collision Induced Dissociation and Ultraviolet Photodissociation. *Int. J. Mass Spectrom.* **2015**, *377*, 385−392.
(58) 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.
(59) Escobar, E. E.; King, D. T.; Serrano-Negrón, J. E.; Alteen, M. G.; Vocadlo, D. J.; Brodbelt, J. S. Precision Mapping of O-Linked N-Acetylglucosamine Sites in Proteins Using Ultraviolet Photodissociation Mass Spectrometry. *J. Am. Chem. Soc.* **2020**, *142* (26), 11569−11577.
(60) Blevins, M. S.; Juetten, K. J.; James, V. K.; Butalewicz, J. P.; Escobar, E. E.; Lanzillotti, M. B.; Sanders, J. D.; Fort, K. L.; Brodbelt, J. S. Nanohydrophobic Interaction Chromatography Coupled to Ultraviolet Photodissociation Mass Spectrometry for the Analysis of Intact Proteins in Low Charge States. *J. Proteome Res.* **2022**, *21* (10), 2493−2503.
(61) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646−12651.
(62) Liu, F. C.; Ridgeway, M. E.; Wootton, C. A.; Theisen, A.; Panczyk, E. M.; Meier, F.; Park, M. A.; Bleiholder, C. Top-Down Protein Analysis by Tandem-Trapped Ion Mobility Spectrometry/Mass Spectrometry (Tandem-TIMS/MS) Coupled with Ultraviolet Photodissociation (UVPD) and Parallel Accumulation/Serial Fragmentation (PASEF) MS/MS Analysis. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 2232−2246.
(63) Shaw, J. B.; Liu, W.; Vasil'ev, Y. V.; Bracken, C. C.; Malhan, N.; Guthals, A.; Beckman, J. S.; Voinov, V. G. Direct Determination of Antibody Chain Pairing by Top-down and Middle-down Mass Spectrometry Using Electron Capture Dissociation and Ultraviolet Photodissociation. *Anal. Chem.* **2020**, *92* (1), 766−773.
(64) Mehaffey, M. R.; Lee, J.; Jung, J.; Lanzillotti, M. B.; Escobar, E. E.; Morgenstern, K. R.; Georgiou, G.; Brodbelt, J. S. Mapping a Conformational Epitope of Hemagglutinin A Using Native Mass Spectrometry and Ultraviolet Photodissociation. *Anal. Chem.* **2020**, *92* (17), 11869−11878.
(65) Gargano, A. F. G.; Roca, L. S.; Fellers, R. T.; Bocxe, M.; Domínguez-Vega, E.; Somsen, G. W. Capillary HILIC-MS: A New Tool for Sensitive Top-Down Proteomics. *Anal. Chem.* **2018**, *90* (11), 6601−6609.
(66) Klein, D. R.; Holden, D. D.; Brodbelt, J. S. Shotgun Analysis of Rough-Type Lipopolysaccharides Using Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2016**, *88* (1), 1044−1051.
(67) Durbin, K. R.; Robey, M. T.; Voong, L. N.; Fellers, R. T.; Lutomski, C. A.; El-Baba, T. J.; Robinson, C. V.; Kelleher, N. L. ProSight Native: Defining Protein Complex Composition from Native Top-Down Mass Spectrometry Data. *J. Proteome Res.* **2023**, *22* (8), 2660−2668.
(68) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370−4376.
(69) An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. Determination of N-Glycosylation Sites and Site Heterogeneity in Glycoproteins. *Anal. Chem.* **2003**, *75* (20), 5628−5637.
(70) Hua, S.; Nwosu, C. C.; Strum, J. S.; Seipert, R. R.; An, H. J.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. Site-Specific Protein Glycosylation Analysis with Glycan Isomer Differentiation. *Anal. Bioanal. Chem.* **2012**, *403* (5), 1291−1302.
(71) Becher, S.; Wang, H.; Leeming, M. G.; Donald, W. A.; Heiles, S. Influence of Protein Ion Charge State on 213 Nm Top-down UVPD. *Analyst* **2021**, *146* (12), 3977−3987.
(72) Juetten, K. J.; Brodbelt, J. S. Top-Down Analysis of Supercharged Proteins Using Collision-Electron-and Photon-Based Activation Methods. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 1467−1476.
(73) Hellinger, J.; Brodbelt, J. S. Impact of Charge State on Characterization of Large Middle-Down Sized Peptides by Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (8), 1647−1656.
(74) Tate, M. D.; Job, E. R.; Deng, Y.-M.; Gunalan, V.; Maurer-Stroh, S.; Reading, P. C. Playing Hide and Seek: How Glycosylation of the Influenza Virus Hemagglutinin Can Modulate the Immune Response to Infection. *Viruses* **2014**, *6* (3), 1294−1316.
(75) D'Ippolito, R. A.; Drew, M. R.; Mehalko, J.; Snead, K.; Wall, V.; Putman, Z.; Esposito, D.; DeHart, C. J. Refining the N-Termini of the SARS-CoV-2 Spike Protein and Its Discrete Receptor-Binding Domain. *J. Proteome Res.* **2021**, *20* (9), 4427−4434.
(76) Thakur, S.; Sasi, S.; Pillai, S. G.; Nag, A.; Shukla, D.; Singhal, R.; Phalke, S.; Velu, G. S. K. SARS-CoV-2 Mutations and Their Impact on Diagnostics, Therapeutics and Vaccines. *Front. Med.* **2022**, *9*, 815389.
(77) Gupta, D.; Sharma, P.; Singh, M.; Kumar, M.; Ethayathulla, A. S.; Kaur, P. Structural and Functional Insights into the Spike Protein Mutations of Emerging SARS-CoV-2 Variants. *Cell. Mol. Life Sci.* **2021**, *78* (24), 7967−7989.
(78) Zhang, L.; Jackson, C. B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B. D.; Rangarajan, E. S.; Pan, A.; Vanderheiden, A.; Suthar, M. S.; Li, W.; Izard, T.; Rader, C.; Farzan, M.; Choe, H. SARS-CoV-2 Spike-Protein D614G Mutation Increases Virion Spike Density and Infectivity. *Nat. Commun.* **2020**, *11* (1), 6013.
(79) Magazine, N.; Zhang, T.; Wu, Y.; McGee, M. C.; Veggiani, G.; Huang, W. Mutations and Evolution of the SARS-CoV-2 Spike Protein. *Viruses* **2022**, *14* (3), 640.
(80) 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**, *21* (12), 729−749.

https://doi.org/10.1021/acs.jproteome.4c00600
J. Proteome Res. 2024, 23, 4684−4693