Hydrophilic interaction chromatography coupled to ultraviolet photodissociation affords identification, localization, and relative quantitation of glycans on intact glycoproteins

Virginia K. James,[1] Annika A.M. van der Zon,[2] Edwin E. Escobar,[1] Andrea F. G. Gargano,[2] Jennifer S. Brodbelt[1]*

AUTHOR ADDRESS:

1. Department of Chemistry, University of Texas at Austin
2. Van 't Hoff Institute for Molecular Science, University of Amsterdam, Science Park 904, Amsterdam, The Netherlands
3. Centre of Analytical Sciences Amsterdam, Science Park 904, Amsterdam, The Netherlands

KEYWORDS : Protein glycosylation, top-down mass spectrometry, ultraviolet photodissociation

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 at the same amino acid residue (glycosite). Top-down mass spectrometry (MS) allows precise identification and localization of glycans on intact proteins, and coupling top-down MS with chromatography allows time-resolved characterization of glycoforms. Here, we couple ultraviolet photodissociation (UVPD) to hydrophilic interaction chromatography (HILIC) to advance the characterization of glycoproteins ranging from 15-34 kDa, offering site localization of glycans, providing sequence coverages up to 93% and relative quantitation of individual glycoforms.

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 of or aberrant glycosylation has been implicated in many diseases, including cancer[2–4] , both neurodegenerative[5–7] and autoimmune diseases,[8,9] among others. Not only have specific glycoproteins been identified as biomarkers of disease state, but disease severity and progression may also be linked to glycosylation patterns.[10–12] Comprehensive analysis of glycoproteins remains elusive owing to extreme heterogeneity in glycan size and identity.[13] Other PTMs, such as phosphorylation and acetylation, do not consist of repeating, branching or variable units, thus resulting in less heterogeneity.[14,15] Conversely, glycans that decorate proteins are composed of multiple monosaccharides, commonly glucose, galactose, mannose, N-acetylglucosamine, and sialic acid, among others, which can be arranged in different orders and branching patterns, each which has the potential to code for different functions or to modify structures.[16,17] In addition to the sheer complexity of the glycan compositions, the glycans may be attached at different and/or multiple sites along the protein sequence, further compounding heterogeneity. Unraveling glycosylation sites and identifying the glycans remains a 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 insight into the specific glycosylation sites of proteins.[18] An alternative strategy 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 non-glycosylated peptides.[17,19] While this method, known as bottom-up glycoproteomics, provides reliable information about glycan location and site occupancy,[17,19] it does not offer comprehensive feedback about glycan branching patterns, linkages, or glycoforms. [20] An overall picture of all the glycosites and their occupancies in a full-length protein is necessary for true combinatorial insight.[20] Only methods that analyze intact proteins offer the potential to map glycoforms, i.e., a depiction of all the glycosites and their occupancies in proteins. However, the intact protein strategy requires appropriate separation methods[21] and high-level data processing algorithms[22] for successful translation to high throughput analysis of heterogeneous mixtures. Top-down methods can map coexisting glycans, thus revealing proteoform-level information,[23] even in a quantitative fashion in some cases.[24,25] However, few studies have reported top-down analysis of glycoproteins, which requires advanced activation techniques,[26] advanced separation techniques[27], or may utilize partial enzymatic deglycosylation[28–30] to mitigate the analytical challenge presented by glycan heterogeneity which creates an overlap of glycoproteoforms in the *m/z* domain.[31]

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,[32] small glycopeptides,[33–35] and some intact proteins,[21] 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,[36,37] along with cleaved glycans.[38] Hydrophilic interaction chromatography (HILIC) has been widely adopted for the separation of cleaved glycans,[39] and glycopeptides,[40–42] and more recently has been adapted for intact glycoproteins.[21,43] Recent advances in HILIC stationary phases have demonstrated especially promising performance metrics for glycoproteins, surpassing the separations achieved by RPLC and CE.[21] For example, the aforementioned HILIC study used a monolith polymer HILIC stationary phase[44,45] to provide sufficient separation of the extremely heterogeneous SARS-CoV-2 spike protein receptor binding domain and assigned glycoforms based on intact mass; albeit without top-down MS/MS data to confirm glycoform assignments.[21] Gas-phase charge reduction strategies have also shown some success in dispersing glycoforms in *m/z* space, thus enabling differentiation of glycoforms;[46,47] these methods commonly suffer from signal suppression so liquid chromatography is still more commonly used for glycoprotein separation. 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.[27–29]

Sophisticated tandem mass spectrometry methods are also essential for the characterization of glycoproteins.[48] Traditional collisional activated dissociation (CAD) methods are widely used to analyze glycopeptides, affording glycan compositional and peptide sequence information.[19,49,50] However, CAD often causes preferential cleavage of the glycans from the peptides, thus making localization of glycans difficult and resulting in low sequence coverage of the middle section of a protein sequence in top-down methods.[51] CAD has been used to characterize O-linked glycans of intact SARS-CoV-2 receptor binding domain (RBD), a protein containing one O-linked glycosite close to the N terminus.[28,29] Electron-based dissociation methods (ETD, ECD, EThcD) have been successful in localizing glycans on peptides[17,52–54] and small intact glycoproteins,[26] and have been adapted for the characterization of the O-linked glycoforms of the RBD.[28] Electron-based and collision-based MS/MS methods may be paired to provide peptide sequence, glycan identification, and glycosite localization of glycopeptides,[54–56] and have been used in an integrated manner for countless high throughput bottom-up glycoproteomics studies.[17,20,49,57] 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.[28]

An alternative high-energy activation method, ultraviolet photodissociation (UVPD), has shown success in identifying and localizing glycans on peptides.[51,58–60] 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.[61–64] 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 the SARS-CoV-2 spike protein receptor binding domain variants.

Materials and methods

Materials and samples

Water, acetonitrile, formic acid, tris(2-carboxyethyl)phosphine (TCEP), trypsin, dithiothreitol (DTT), iodoacetamide (IAM), and trifluoroacetic (TFA) acid (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 A 1 (HA) was a gift from the Georgiou group (UT-Austin) and was produced as previously described.[65] SARS-CoV-2 spike protein receptor binding domain (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 hours at 37 °C. The WT RBD sample was reduced with DTT (5 mM), alkylated with IAM (15 mM) for 30 minutes at room temperature, and digested overnight with trypsin at 37 °C for bottom-up analysis. All protein sequences, masses, and vendors/catalog 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 indicates the number of each unit contained within the glycan.

LC-MS/MS

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[21,44,66] and held at room temperature. Proteins were diluted to 50-500 ng/µL in water with 0.1% formic acid, and 1 µL of solution was injected per experiment. Water with 0.1% formic acid was used for 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 varied by type of sample and are listed in **Tables S2-4**.

Traditional reversed phase chromatography was also used to compare to HILIC (intact proteins) and to 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 the tryptic digest, respectively. The analytical and trap columns had an ID of 75 μm and 100 μm, respectively. The columns were packed to a length 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 the TFA was omitted. The gradient is given in **Table S5** for all 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.[67] A 15 μm ID emitter (New Objective), 35 V of in source collision activation, and 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 ten microscans and a 3E6 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 3E6 AGC target. UVPD spectra were collected at 240,000 resolution, 50-100 microscans, and a 1E6 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 HCD scan events, as previously described.[51]

Data processing

All MS2 data was processed with Prosight Native/TD validator with a mass calibration of 0-10 ppm and default additional settings, including a S/N threshold of 3, cluster tolerance of 0.35, ppm error tolerance of 10, score of 0.5, and 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 NIST glycan mass calculator (https://www.nist.gov/static/glyco-mass-calc/). Quantitation was performed by rendering extracted ion chromatograms (EICs) for the highest abundant charge state of each glycoform in QualBrowser, and peaks of each EIC were integrated using the Genesis algorithm. All fragment ion identifications and glycoform relative abundance used for quantitation are available in the supplemental information. Glycoforms for the multiply glycosylated RBD sample were identified with a custom Matlab script and data available from previous glycoproteomics studies on the RBD.[21] All data is 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, a 15 kDa glycoprotein with a single N-linked glycosite (N34) and eight abundant glycoforms. As previously established,[21,44] the HILIC monolith column provides excellent separation of glycoforms as evidenced by the near baseline chromatographic resolution of four glycoforms in **Figure 1a**. The companion MS1 spectra are shown in **Figure 1b** (and deconvoluted mass spectra in **Figure S1**). We compared the performance of the HILIC separation to a conventional reversed phase separation using PRLP stationary phase (the one most commonly employed for LC-MS analysis of intact proteins), the latter of which did not resolve any of the glycoforms (**Figure S2**). The HILIC separation provided sufficient resolution to isolate individual glycoforms for subsequent MS/MS analysis. The performance of four MS/MS methods, higher-energy collisional dissociation (HCD, a beam-type CAD), ETD, hybrid EThcD, and 193 nm UVPD, was assessed for characterization of RNase B, as summarized in **Figure S3** for the glycoform containing penta-mannose. HCD afforded low sequence coverage (39%) of RNase B and few backbone cleavages close to the glycosylation site (**Figure S3a**), preventing the known N34 glycosite from being confidently differentiated from N44. 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 1c, 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 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. The UVPD runs were not used for quantitative analysis owing to distorted peak shape caused by the averaging needed for MS/MS experiments (**Figure S5**). The relative abundances of four glycoforms are summarized in **Figure 1d**, consistent with those reported based on bottom-up analysis of RNase B glycopeptides (**Table S6**).[69,70] Technical replicates demonstrated low run-to-run variability in elution times and standard deviations of the chromatographic peak areas (**Figure 1d**). Blanks run between samples ensured lack of carryover.

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 1b**, top spectrum). However, many larger glycoproteins contain numerous different glycans located at the same or different glycosites, such as observed for partial-length hemagglutinin HA1 (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,

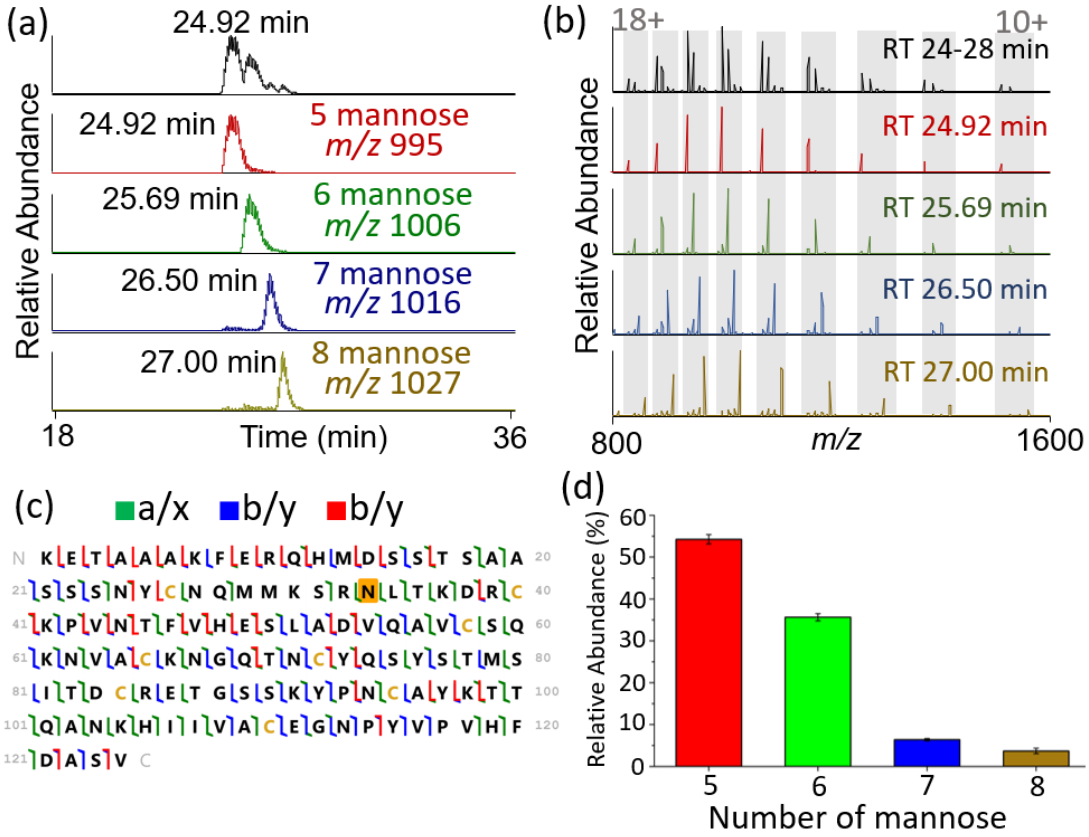


Figure 1. (a) HILIC separation of RNase with base peak chromatogram (top) and extracted ion chromatograms for each glycoform (middle to bottom). (b) MS1 spectra of each glycoform corresponding to each retention time with and charge-assigned spectra shown in **Figure S1** and average intact masses in **Table S7**. (c) The sequence deconvoluted coverage map was derived from UVPD (1 pulse, 2 mJ) of the 15+ charge state (UVPD mass spectrum shown in **Figure S4**). All identified fragment ions are summarized in the supporting information (**Table S8**). The location of the pentamannose glycan is shaded in gold (N34). (d) Relative abundances of glycoforms are compared with error bars representing the standard deviation based on three replicates.

spectral congestion is a significant obstacle, as exemplified by the high-resolution mass spectrum of HA shown in **Figure S6**. The assignment of charge states is not only confounded, but isolation of individual glycoforms for MS/MS analysis is thwarted. HILIC provided ample separation of HA glycoforms and non-glycosylated HA (**Figure 2a, S7a**), resulting in simplified time-resolved MS1 spectra containing fewer proteoforms (**Figure S7b**) and facilitating the acquisition of MS/MS spectra.

UVPD mass spectra were acquired for all HA glycoforms above 5% relative abundance, eight glycoforms in total. The UVPD mass spectrum of the most abundant glycoform (H5N4F1S1) is shown in **Figure 2b**, displaying numerous fragment ions that contain the glycan, such as $(y_{190}\text{-}1)^{23+}$ and $(a_{44}+1)^{5+}$ ions, with isotopic profiles shown beneath the UVPD mass spectrum. The sequence map is displayed in **Figure 2c**, resulting in 77% sequence coverage for the H5N4F1S1 glycoform. HILIC-UVPD of six other HA glycoforms plus the non-glycosylated protein yielded sequence coverages ranging from 67% to 82% (**Figure S8**). 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 2d** and **Table S8**. The identification of several abundant glycoforms (10-20% relative abundance) confirmed the high heterogeneity at this glycosite. As HA glycosylation patterns are known to vary with influenza strain,[71] the identification, localization, and relative quantitation of HA glycans provide insight into the development of future vaccines and antiviral therapeutics.

In addition to HA, we investigated the glycosylation of SARS-CoV-2 spike protein receptor binding domain (RBD), another viral glycoprotein. The RBD is a 25 kDa immunogenic fragment of the large 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.[21,46] We first focused on the characterization of the O-linked glycoforms by using PNGase F to remove the N-linked glycans. O-linked glycans are often the most chall enging 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.[51]

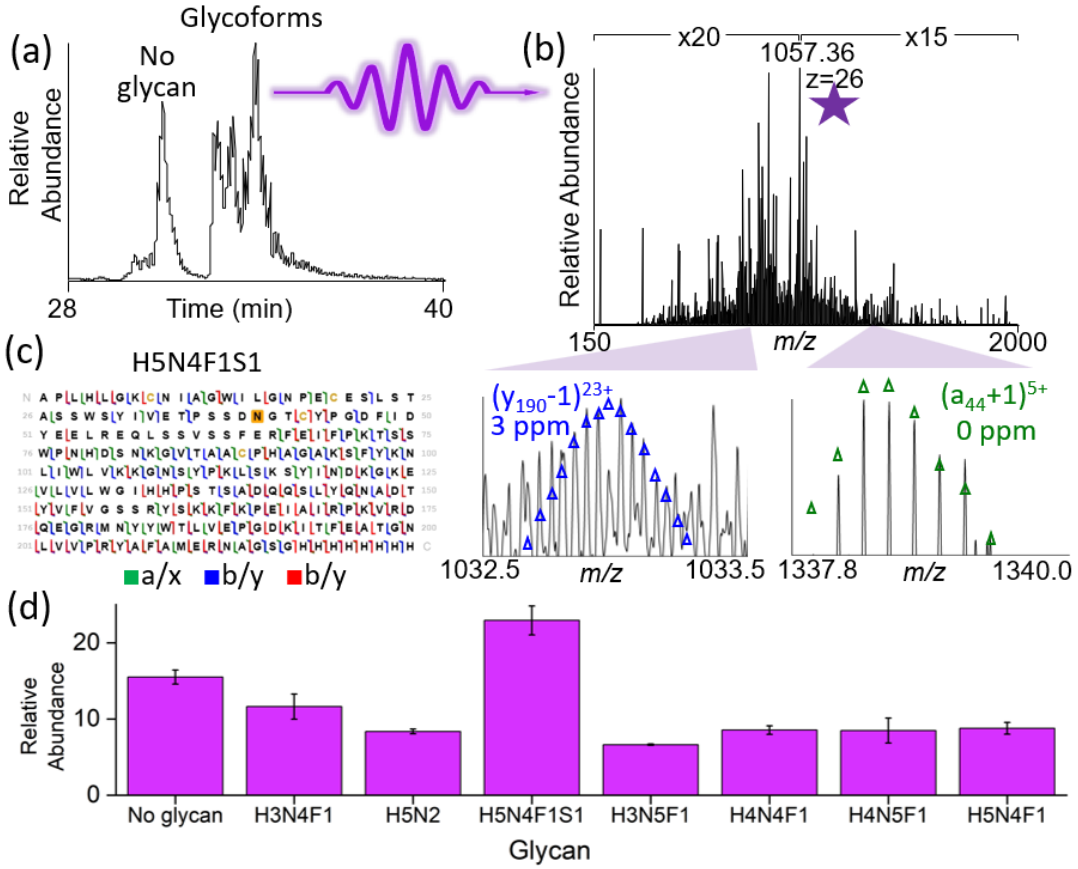


Figure 2. (a) HILIC chromatogram of HA with extracted ion chromatograms and MS1 spectra of glycoforms shown in **Figure S7** and intact mass average masses in **Table S9**. (b) UVPD mass spectrum (1 pulse, 2 mJ) of the most abundant glycoform of HA (H5N4F1S1, *m/z* 1057, t_4 = 34.0 min, highlighted with a star) with inserts showing two glycan-containing 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 S8,** and all identified fragment ions are provided in **Table S8**. (d) Comparison of the relative abundance of glycoforms with error bars representing the standard deviation from three replicates (see specific values in **Table S10**).

HILIC provided excellent separation of the O-linked glycoforms of the WT RBD (**Figure 3a**), as evidenced by the MS1 spectra for each glycoform (**Figure 3b**). We first localized a non-glycan 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 analysis of a tryptic digest of the RBD protein (**Figure S9**) and is consistent with a previous report.[72] UVPD of the H1S2N1 glycoform (**Figure 3c**) yielded fragment ions that confirmed the location of the O-linked glycan as T6 when considering the addition of one amino acid pyroGlu (Q) to the N terminus, T5 without Q1 (**Figure 3d**). 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%. UVPD fragment ion identifications for all glycoforms (H1S1N1, H2S1N2, H1S2N1, and H2S2N2) are provided in the supporting information (**Table S8**), and the corresponding sequence coverage maps for H1S1N1, H2S1N2, and H2S2N2 glycoforms are shown in **Figure S10**. 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 immunogenic fragment of the larger spike glycoprotein, on transmissibility and virulence.[73] Here, we compare five RBD variants containing different point mutations (**Table S1**). The N-linked glycans were removed using PNGase F 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 and extracted ion chromatograms in **Figures S12-S16**). UVPD spectra were collected for each O-glycoform, and the correspondingfragment ion identifications are provided in the supplemental information (**Table S8**), along with the resulting sequence maps in **Figures S17-S21**. The UVPD data confirmed that all RBD samples have the same O-linked glycosite, T6 and T5, for the WT and variant RBD samples, respectively. The distribution of O-glycoforms for each variant and the WT protein are summarized in **Figure 4**. The identities of the RBD glycans are similar to those reported previously,[28,29] and the specific glycosylation patterns of each variant are generally similar except for the E484K variant for which the H1S1N1

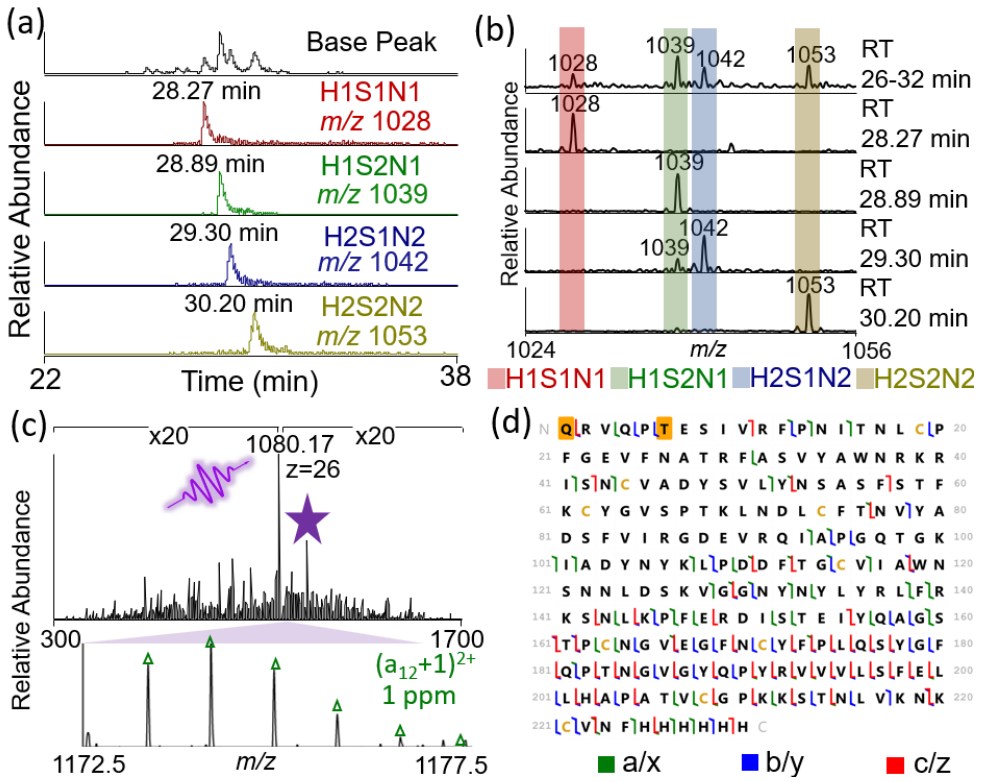


Figure 3. (a) Base peak chromatogram and extracted ion chromatograms of each O-linked glycoform of SARS-CoV-2 spike protein receptor binding domain (RBD) and (b) corresponding MS1 spectra of a single charge state (26+) at each retention time with full range MS1 spectra and deconvoluted spectra shown in **Figure S10**. Average intact masses are given in **Table S10**. (c) UVPD (1 pulse, 2 mJ) mass spectrum of the 26+ charge state of the H1S2N1 glycoform (*m/z* 1080) 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) at the N terminus). (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 S8**.

glycan was not detected. Interestingly, none of the RBD variants contained the H2S1N2 glycan, which was only identified for the wild-type RBD (**Figure 4**). 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%).[28,29] This trend continues with the variants agree more closely with literature values than the lower abundance glycoforms, H1S1N1 and H2S2N2, as the WT which shows higher abundances of H1S1N1 and H2S2N2 (**Table S11).** Given that the RBDs examined here and elsewhere[28,29] 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.

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, in essence, dissecting combinatorial post-translational modification patterns. 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 for fully glycosylated RBD HILIC (**Figures 5a, S22**). Although HILIC did not provide baseline separation

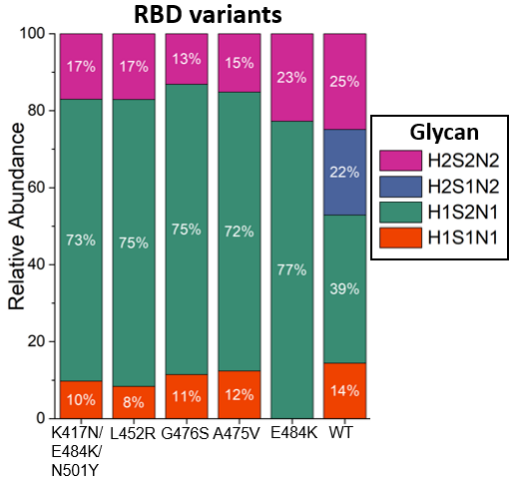


Figure 4. Relative abundances of each glycoform. Chromatograms and MS1 spectra for the RBD variant samples are shown in **Figures S12-S16,** average intact masses are shown in **Table S10,** and sequence coverage maps generated for each glycoform are shown in **Figures S17-S21**.

of individual glycoforms, it offered sufficient chromatographic separation when implemented along with isolation of narrow *m/z* bands to winnow the number of glycoforms selected, enabling UVPD characterization of the more abundant glycoforms (**Figure 5b**). Combining the set of putative glycoforms reported in a previous study[21] and the UVPD data acquired here, we showcase an example of aglycoform identified for the fully glycosylated RBD (**Figure 5c**). Accurate mass assignments of the $(a_{14}+1)^{4+}$ and $(x_{224}+1)^{26+}$ fragment ions support the localization of all three glycans. 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 advances that are critical for realizing the goal of comprehensive top-down glycoproteoform analysis.

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 using an integrated HILIC-UVPD strategy. After separation of glycoforms by HILIC, UVPD was used to characterize individual glycoforms and the relative abundance 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 heterogenous 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[74–76], 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 insight into glycosylations that may occur in tandem.

ASSOCIATED CONTENT

Supporting Information. PDF: Protein sequences, masses, and catalog numbers, LC gradients, comparison of glycoform abundances to literature values, PLRP vs HILIC separations for RNase B, deconvoluted RNase B MS1 spectra, comparison of chromatographic peak shape for RNase B MS1 only and UVPD runs, ESI mass spectra of HA collected without LC separation, sequence coverage maps and extracted ion chromatograms for RNase B (HCD, ETD, EThcD, and UVPD), HA (UVPD only) and RBDs (UVPD only). Spreadsheet: Identified fragment ions in MS2 spectra for RNase B (HCD, ETD, EThcD, and UVPD), HA (UVPD only) and RBDs (UVPD only). Relative quantitation results for RNase B, HA, and the RBDs. This material is available

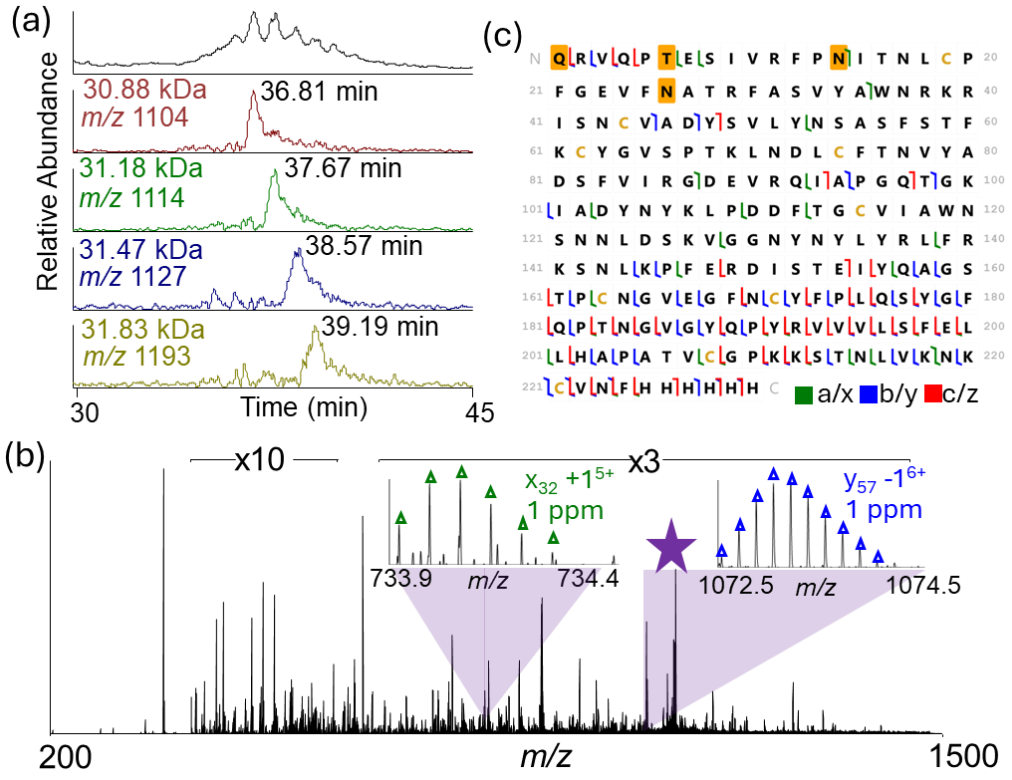


Figure 5. (a) Chromatograms include a base peak (top trace) and EICs of the most abundant species of the WT RBD. (b) UVPD mass spectrum of the 28+ charge state of the 31.18 kDa glycoform and the selected precursor is labeled with a star. (c) The resulting sequence coverage map displays the identified modified residues (T6 (H1S2N1), N14 (N6H4F1), and N26 (N5H3F2S1)). The glycoform also contains a pyroGlu (+111 Da) modification at the N-terminus. The locations of the modifications are shaded in gold. MS1 spectra of other glycoforms are shown in **Figure S22**.

free of charge via the Internet at http://pubs.acs.org.xx

AUTHOR INFORMATION

Corresponding Author

* Correspondence to: jbrodbelt@cm.utexas.edu

Funding Sources

Funding from the National Institutes of Health (R35GM13965) and the Robert A. Welch Foundation (F-1155) to J.S.B. is gratefully acknowledged.

ACKNOWLEDGMENTS

The gift of HA from the Georgiou lab (UT-Austin) is gratefully acknowledged. The authors acknowledge Sean Dunham for helpful conversations on top-down data acquisition and Dr. Mowei Zhou for insight about implementation of HILIC-MS. 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/E484K/N501Y Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55414, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, A475V Variant with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-55402, Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, E484K 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 Review of Proteomics* **2019**, *16* (8), 665–680. https://doi.org/10.1080/14789450.2019.1645604.

(4) Silsirivanit, A. Chapter Five - Glycosylation Markers in Cancer. In *Advances in Clinical Chemistry*; Makowski, G. S., Ed.; Elsevier, 2019; Vol. 89, pp 189–213.

(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? *The FEBS Journal* **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), eabk0637.

(7) Xu, M.; Zhou, M.; Li, S.; Zhen, X.; Yang, S. Glycoproteins as Diagnostic and Prognostic Biomarkers for Neurodegenerative Diseases: A Glycoproteomic Approach. *Journal of Neuroscience Research* **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 Letters* **2019**, *593* (13), 1598–1615.

(10) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. *Nat Rev Nephrol* **2019**, *15* (6), 346–366.

(11) Gudelj, I.; Lauc, G.; Pezer, M. Immunoglobulin G Glycosylation in Aging and Diseases. *Cellular Immunology* **2018**, *333*, 65–79.

(12) Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat Rev Cancer* **2015**, *15* (9), 540–555.

(13) Oliveira, T.; Thaysen-Andersen, M.; Packer, N. H.; Kolarich, D. The Hitchhiker's Guide to Glycoproteomics. *Biochemical Society Transactions* **2021**, *49* (4), 1643–1662.

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

(15) Leutert, M.; Entwisle, S. W.; Villén, J. Decoding Post-Translational Modification Crosstalk With Proteomics. *Molecular & Cellular Proteomics* **2021**, *20*. h

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

(17) 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.

(18) 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.

(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) 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), 1–29.

(21) 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.

(22) 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.

(23) 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.

(24) Cupp-Sutton, K. A.; Wu, S. High-Throughput Quantitative Top-down Proteomics. *Mol. Omics* **2020**, *16* (2), 91–99.

(25) 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*[S]. *Molecular & Cellular Proteomics* **2019**, *18* (3), 594–605.

(26) 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.

(27) 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.

(28) 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.

(29) S. Roberts, D.; 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. *Chemical Science* **2022**, *13* (36), 10944–10949.

(30) Wu, D.; Robinson, C. V. Native Top-Down Mass Spectrometry Reveals a Role for Interfacial Glycans on Therapeutic Cytokine and Hormone Assemblies. *Angewandte Chemie International Edition n/a* (n/a).

(31) Reid, D. J.; Thibert, S.; Zhou, M. Dissecting the Structural Heterogeneity of Proteins by Native Mass Spectrometry. *Protein Science* **2023**, *32* (4), e4612. https://doi.org/10.1002/pro.4612.

(32) Kurz, S.; Sheikh, M. O.; Lu, S.; Wells, L.; Tiemeyer, M. Separation and Identification of Permethylated Glycan Isomers by Reversed Phase NanoLC-NSI-MSn. *Molecular & Cellular Proteomics* **2021**, *20*. https://doi.org/10.1074/mcp.RA120.002266.

(33) 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. *Journal of Chromatography B* **2019**, *1110–1111*, 101–107.

(34) 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.

(35) Yin, H.; Zhu, J. Methods for Quantification of Glycopeptides by Liquid Separation and Mass Spectrometry. *Mass Spectrometry Reviews* **2023**, *42* (2), e21771. https://doi.org/10.1002/mas.21771.

(36) 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. https://doi.org/10.1002/elps.201900333.

(37) 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. *Journal of Chromatography B* **2021**, *1179*, 122855.

(38) Tomnikova, A.; Kozlík, P.; Křížek, T. Monosaccharide Profiling of Glycoproteins by Capillary Electrophoresis with Contactless Conductivity Detection. *ELECTROPHORESIS* **2022**, *43* (20), 1963–1970.

(39) 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. *Analytical Biochemistry* **2021**, *635*, 114447.

(40) 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. https://doi.org/10.7171/jbt.16-2703-003.

(41) 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. *Journal of Chromatography B* **2023**, *1228*, 123852.

(42) 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 Proteom* **2022**, *13* (2), 95–108.

(43) 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. In *Carbohydrate Analysis by Modern Liquid Phase Separation Techniques (Second Edition)*; El Rassi, Z., Ed.; Elsevier: Amsterdam, 2021; pp 209–278.

(44) Passamonti, M.; de Roos, C.; Schoenmakers, P. J.; Gargano, A. F. G. 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.

(45) 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. *Journal of Chromatography A* **2023**, *1688*, 463721.

(46) 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.

(47) 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.; Syka, J. E. P.; Huguet, R.; Marty, M. T.; Sandoval, W. Exposing the Molecular Heterogeneity of Glycosylated Biotherapeutics. bioRxiv May 11, 2023, p 2023.05.10.540271.

(48) 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.

(49) 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.

(50) 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. *Clinical Proteomics* **2019**, *16* (1), 35.

(51) 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.

(52) 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.

(53) 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.

(54) 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.

(55) 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.

(56) 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.

(57) 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**, 1–15.

(58) 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.

(59) Ko, B. J.; Brodbelt, J. S. Comparison of Glycopeptide Fragmentation by Collision Induced Dissociation and Ultraviolet Photodissociation. *International Journal of Mass Spectrometry* **2015**, *377*, 385–392.

(60) 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.

(61) 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.

(62) 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.

(63) 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**.

(64) 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.

(65) 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.

(66) 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.

(67) 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.

(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. https://doi.org/10.1007/s00216-011-5109-x.

(71) 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.

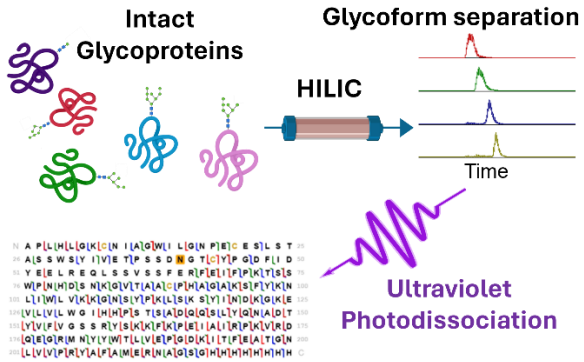
(72) D'Ippolito, R. A.; Drew, M. R.; Mehalko, J.; Snead, K.; Wall, V.; Putman, Z.; Esposito, D.; DeHart, C. J. Refining the N-Termini of the SARS-CoV-2 Spike Protein and Its Discrete Receptor-Binding Domain. *J Proteome Res* **2021**, *20* (9), 4427–4434.

(73) 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 (Lausanne)* **2022**, *9*, 815389.

(74) 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..

(75) 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.

(76) 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.



Insert Table of Contents artwork here