Novel Antibodies for the Simple and Efficient Enrichment of Native O-GlcNAc Modified Peptides

Authors
Rajan A. Burt, Borislav Dejanovic, Hayley J. Peckham, Kimberly A. Lee, Xiang Li, Johain R. Ounadjela, Anjana Rao, Stacy A. Malaker, Steven A. Carr, and Samuel A. Myers

Correspondence
scarr@broad.mit.edu; sam@lji.org

In Brief
O-GlcNAc, the single N-acetylglucosamine coupled to serine and threonines of nucleocytoplasmic proteins, is an enigmatic posttranslational modification implicated in most cellular functions. Methods to study the O-GlcNAc modified proteome have been challenging, in part, due to limitations in its enrichment from native tissue. Here, we characterize novel antibodies for the enrichment of native O-GlcNAc modified peptides. Our enrichment strategy shows strong sensitivity and excellent specificity for O-GlcNAc. We expect these reagents to be a significant advancement for the field.

Graphical Abstract
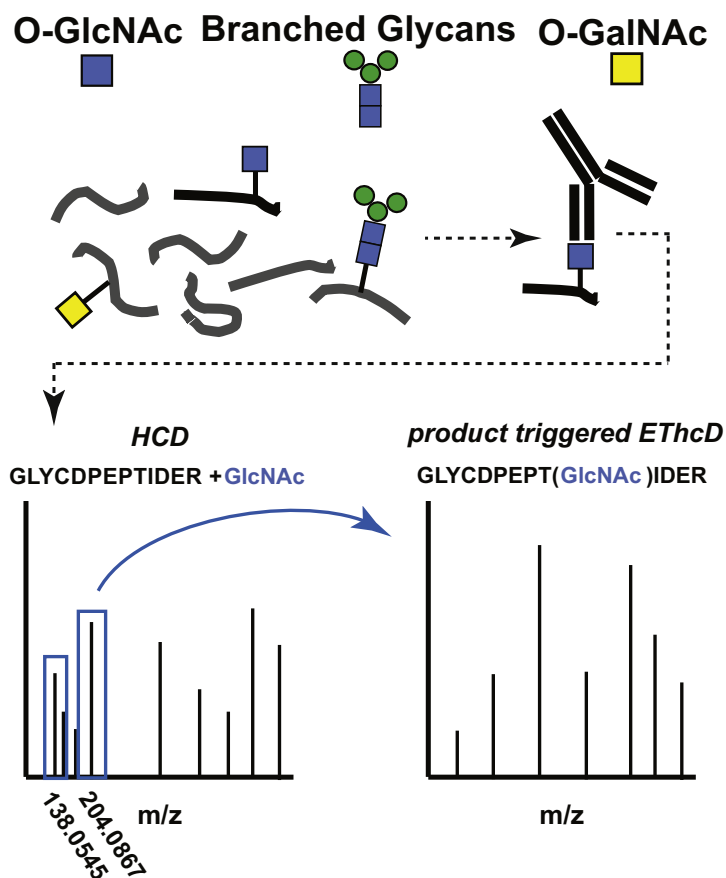
Highlights

- Anti-O-GlcNAc antibodies are fast and simple enrichment reagents.
- Anti-O-GlcNAc antibodies are sensitive and achieve significant depth of coverage.
- Anti-O-GlcNAc antibodies are specific for singular O-GlcNAc modifications.
- Anti-O-GlcNAc antibody enrichment techniques can be applied to cells and tissues.
- HCD product-triggered EThcD data acquisition improves depth of coverage.

Novel Antibodies for the Simple and Efficient Enrichment of Native O-GlcNAc Modified Peptides

Rajan A. Burt[1] , Borislav Dejanovic[2], Hayley J. Peckham[3], Kimberly A. Lee[3], Xiang Li[4], Johain R. Ounadjela[1] , Anjana Rao[4,5,6] , Stacy A. Malaker[7], Steven A. Carr[1,*], and Samuel A. Myers[1,4,*]

Antibodies against posttranslational modifications (PTMs) such as lysine acetylation, ubiquitin remnants, or phosphotyrosine have resulted in significant advances in our understanding of the fundamental roles of these PTMs in biology. However, the roles of a number of PTMs remain largely unexplored due to the lack of robust enrichment reagents. The addition of N-acetylglucosamine to serine and threonine residues (O-GlcNAc) by the O-GlcNAc transferase (OGT) is a PTM implicated in numerous biological processes and disease states but with limited techniques for its study. Here, we evaluate a new mixture of anti-O-GlcNAc monoclonal antibodies for the immunoprecipitation of native O-GlcNAcylated peptides from cells and tissues. The anti-O-GlcNAc antibodies display good sensitivity and high specificity toward O-GlcNAc-modified peptides and do not recognize O-GalNAc or GlcNAc in extended glycans. Applying this antibody-based enrichment strategy to synaptosomes from mouse brain tissue samples, we identified over 1300 unique O-GlcNAc-modified peptides and over 1000 sites using just a fraction of sample preparation and instrument time required in other landmark investigations of O-GlcNAcylation. Our rapid and robust method greatly simplifies the analysis of O-GlcNAc signaling and will help to elucidate the role of this challenging PTM in health and disease.

O-GlcNAc is the monosaccharide addition of N-acetylglucosamine to serine and threonine residues of nuclear, cytosolic, and mitochondrial proteins. Distinct from N- and O-linked ER and Golgi glycosylation pathways, O-GlcNAc is considered to be a dynamic, regulatory, intracellular posttranslational modification (PTM). Unlike other types of glycosylation, O-GlcNAc is not elaborated beyond the initiating modification. Understanding the functional relevance of specific O-GlcNAc sites has proven challenging for several reasons. The sole enzymes responsible for the addition and removal of O-GlcNAc, OGT, and OGA/MGEA5, respectively, are essential for embryonic development (1, 2). Tissue-specific deletions of either gene present different and often deleterious effects depending on the context (3–7). The mechanisms underlying these phenotypic changes are likely pleiotropic as O-GlcNAc signaling affects nearly every cellular process (Reviewed in Zachara *et al.* (8)). As such, having an experimental strategy to confidently map O-GlcNAc sites across the proteome would be a powerful tool to dissect the functions of this diverse PTM.

Mapping O-GlcNAcylation affords its own challenges including enriching O-GlcNAc-modified peptides from complex biological mixtures and unambiguously identifying sites of modification. Regarding the latter, the glycosidic bond between the GlcNAc and the hydroxyl-containing residue is labile. Typical collisional activation conditions (CID, HCD, etc.) result in near stoichiometric loss of the GlcNAc moiety from the precursor glycopeptide. This removes the information required for site assignment (9, 10). Fragmentation using electron capture dissociation (11) and its more recent replacement electron transfer dissociation (ETD) have been well documented to preserve the O-glycosidic bond while breaking the peptide backbone to enable unambiguous modification site assignment (12–14). ETD, however, is inherently slower than HCD and is not compatible with widely used isobaric peptide labeling quantification strategies such as iTRAQ and TMT. Consequently, HCD has been used in combination with ETD by using an HCD-product-dependent ETD acquisition method (HCD-pd-ETD), where sugar oxonium ions and fragments thereof are detected in the HCD scan triggering ETD on the same, reaccumulated precursor (9, 15).

https://doi.org/10.1016/j.mcpro.2021.100167

Several strategies have been developed to enrich O-GlcNAc-modified peptides, though most have notable drawbacks (Reviewed in Maynard and Chalkley (16)). Metabolic labeling with monosaccharides bearing bioorthogonal chemical moieties has been used successfully in multiple iterations (17–19). By culturing cells with chemically modified monosaccharides, these sugars are incorporated into metabolic pathways that eventually modify proteins. Modified peptides or proteins can then be coupled to an affinity handle for subsequent enrichment and analysis by immunoblotting or mass spectrometry. What metabolic labeling gains in sensitivity, it loses in applicability, being incompatible with sample sources that cannot be cultured in the laboratory. Chemo-enzymatic approaches have also been successful, where azide or alkyne-modified N-acetylgalactosamine (GalNAc) is enzymatically transferred by a mutant GalT1 to terminal GlcNAc moieties before being coupled to an affinity handle (14, 20). The widespread adoption of this approach, however, has been hindered by its complexity.

Enrichment of native O-GlcNAc-modified peptides is important to expand our knowledge of this PTM's function, allowing cells and/or tissues from mouse models or human patients to be compared across biological states. Phenyl-boronic acid can coordinate the cis-diol moiety of glucose for enrichment of native O-GlcNAc-modified peptides, though its sensitivity and specificity have not been fully explored (21, 22). Wheat germ agglutinin (WGA), a lectin that has a weak affinity for GlcNAc, has been used successfully for O-GlcNAc-modified peptide enrichment across a wide range of cells and tissues (23–25). Although WGA-based lectin weak affinity chromatography (LWAC) has provided some of the largest datasets of O-GlcNAc-modified peptides (24, 26, 27) its lack of specificity is well documented, and the amount of input peptide needed can be prohibitive (24, 28, 29).

A simple and sensitive strategy to enrich for native O-GlcNAc-modified peptides in cells and tissues would greatly facilitate the study and understanding of O-GlcNAc signaling on a global and local (site-specific) scale. Here, we characterize a newly developed mixture of anti-O-GlcNAc monoclonal antibodies capable of enriching native O-GlcNAc-modified peptides from cells and tissues using peptide input amounts compatible with many sample types. The antibody mixture shows excellent specificity toward O-GlcNAc-modified peptides, reducing the number of coenriched glycan-containing peptides that are common with other enrichment strategies. We show that these anti-O-GlcNAc antibodies can be coupled with offline fractionation to enable deep characterization of the O-GlcNAc proteome from mouse brain tissue. We believe this commercially available enrichment reagent will be a breakthrough for our ability to map global, site-specific O-GlcNAcylation patterns across a diverse range of biological samples.

EXPERIMENTAL PROCEDURES

Anti-O-GlcNAc Antibody Development

Polyclonal antibodies were produced by immunizing New Zealand White rabbits with randomized peptide libraries containing serine and threonine residues modified with O-linked GlcNAc coupled to KLH. Rabbits were selected for monoclonal antibody development based on reactivity in ELISA and Western blot assays. Four clones were selected for inclusion in the immunoenrichment kit (PTMScan O-GlcNAc [GlcNAc-S/T] Motif Kit #95220, Cell Signaling Technology, Inc) to provide the broadest coverage of O-GlcNAc sites.

Sample Generation

Mouse mESCs were routinely passaged by standard methods using 0.5% trypsin and ESC media (KO-DMEM, 10% FBS, 2 mM glutamine, 1× nonessential amino acids, 0.1 mM β-mercaptoethanol, and recombinant leukemia inhibitory factor). Synaptosomal preps were prepared as previously described (30). Briefly, ten mouse brains were homogenized in homogenization buffer (5 mM HEPES (pH 7.4), 1 mM MgCl2, 0.5 mM CaCl2) supplemented with phosphatase inhibitors (PhosStop, Roche), protease inhibitors (cOmplete mini, Roche) and 10 μM PUGNAc using a Teflon homogenizer. After 10 min of centrifugation at 1400g, the supernatant was centrifuged at 13,800g for 10 min. The pellet was resuspended in 0.32 M Tris-buffered sucrose and ultracentrifuged into 1.2, 1.0, 0.85 M sucrose gradient at 82,500g, 2 h. Synaptosome fractions between the 1 and 1.2 M sucrose layer were collected, pelletized, and stored at −80 °C for downstream processing. All steps were performed with ice-cold buffers and centrifugation steps were performed at 4 °C.

Sample Preparation

Samples were prepared by two different methods. For urea-based digestion, cell pellets were resuspended in a 8M urea lysis buffer solution containing 75 mM NaCl, 50 mM Tris HCl (pH 8.0), 1 mM EDTA, Aprotinin (2 μg/μl) (Sigma-Aldrich), Leupeptin (10 μg/μl) (Roche), PMSF (1 mM) (Sigma-Aldrich), Phosphatase Inhibitor Cocktail 2 (1:100 dilution of commercial stock, Sigma-Aldrich), Phosphatase Inhibitor Cocktail 3 (1:100 dilution of commercial stock, Sigma-Aldrich), and NaF (10 mM). Protein concentrations were determined with a bicinchoninic acid (BCA) protein assay (Pierce). Lysis buffer solution was used to dilute protein concentrations to 8 mg/ml. Proteins were then reduced with 5 mM dithiothreitol (DTT) while shaking for 1 h at room temperature. Proteins were subsequently alkylated with 10 mM of iodoacetamide for 45 min at room temperature shaking in the dark. Before digestion, proteins were diluted with 50 mM Tris HCl (pH 8.0) until the urea concentration was below 2 M. Proteins were then digested for 2 h with LysC (Wako Laboratories) for 2 h at room temperature. Following this, proteins were digested with sequencing-grade trypsin (Promega) at room temperature while shaking overnight. Enzymatic digestion for both LysC and trypsin was performed at a 1:50 enzyme–substrate ratio. After digestion, peptides were acidified through the addition of formic acid (FA) to 1% and insoluble material was pelleted by centrifugation at 4000g for 10 min. Peptides were desalted with 500 mg reversed-phase tC18 Sep-Pak cartridge (Waters). Cartridges were conditioned with 5 ml of ACN, 5 ml of 50% ACN 0.1% FA and equilibrated four times with 5 ml of 0.1% trifluoroacetic acid (TFA). Peptides were loaded onto Sep-Pak cartridges, washed three times with 0.1% TFA, once with 1% FA, and then eluted with 50% ACN plus 0.1% FA. Peptide solutions were frozen, concentrated by vacuum centrifugation, and then resuspended in 30% ACN 0.1% FA. Peptide concentration was assessed *via* a BCA assay. Peptides were then aliquoted, frozen, dried to near completion, and stored at −80 °C.

For SDS-based lysis, cell pellets or synaptosome preps were resuspended in 1% or 3% SDS, respectively, in a lysis buffer solution containing 50 mM TEAB, 1 mM EDTA, Aprotinin (2 µg/µl), Leupeptin (10 µg/µl), PMSF (1 mM), Phosphatase Inhibitor Cocktail 2 (Sigma, 1:100 dilution of commercial stock), Phosphatase Inhibitor Cocktail 3 (Sigma, 1:100 dilution of commercial stock), and NaF (10 mM). Samples were sonicated using a Misonix Ultrasonic Liquid Processor (Fisher Scientific) probe sonicator pulsed twice for 4 s at an amplitude of 30 before a 10 min, 4 °C centrifugation at 20,000g. Protein concentrations were assessed by BCA, reduced with 5 mM DTT for 1 h, and subsequently alkylated with 10 mM iodoacetamide for 45 min in the dark while shaking at room temperature for both reactions. Proteins were precipitated using suspension trapping (31, 32) using S-Trap midi columns (Protifi) per manufacturer's recommendations. After digestion with trypsin, peptides were eluted, dried by vacuum centrifugation, and resuspended in 30% ACN 0.1% FA aqueous solutions. After quantification by BCA, peptides were aliquoted in their desired amounts, frozen, dried by vacuum centrifugation, and stored at –80 °C.

Enrichment of O-GlcNAc-Modified Peptides by Immunoprecipitation

Peptides were resuspended in chilled immunoaffinity purification (IAP) buffer at 1 mg/ml. Anti-O-GlcNAc noncovalently linked to protein-A coated agarose bead aliquots (PTMScan O-GlcNAc [GlcNAc-S/T] Motif Kit #95220Cell Signaling Technology, Inc) were washed twice and resuspended in 160 µl of chilled IAP buffer. A range of reagent amounts were used (0.125× to 1× the recommended amount) for the antibody titration experiment. For all subsequent experiments, 0.25× (40 µl) of the resuspended antibody aliquot was used. Anti-O-GlcNAc Ab-beads were incubated with peptide samples for 2 h at 4 °C with end-over-end tumbling. After incubation, beads were pelleted by centrifugation for 1 min. Flow-throughs were transferred to separate tubes and frozen. Beads were first washed once with chilled IAP buffer and then twice with chilled PBS. Samples were eluted through resuspension with 150 µl of a 0.15% TFA aqueous solution with light manual agitation for 5 min. Samples were centrifuged for 1 min, and eluents were transferred to equilibrated Stage tips. The elution procedure was repeated twice before continuing with Stage tip sample clean up.

For all glycopeptide enrichment sample cleanup, Stage tips (33) were prepared by stacking two punches of Empore 2215-C18 Octadecyl 47 mm disks (Bioanalytical Technologies) at the bottom of a 200 µl pipette tip. Stage tips were washed with 100 µl of methanol, 100 µl of 50% ACN 0.1% FA, and were equilibrated with 100 µl of 0.1% FA twice. Samples were loaded onto Stage tips after equilibration, washed twice with 100 µl of 0.1% FA, and eluted with 50 µl of 50% ACN with 0.1% FA. Samples were then transferred to HPLC tubes, frozen, and dried by vacuum centrifugation. Samples were then resuspended in 9 µl of solvent A (3% ACN 0.1% FA) before LC-MS/MS analysis.

Enrichment of Glycopeptides With WGA

WGA-based glycopeptide enrichment was performed on 1 mg and 125 µg of urea digested mESC peptide aliquots in duplicate. Peptide samples were first resuspended in chilled IAP at 1 mg/ml. Next, 600 µl of WGA Agarose slurry (Vector Labs) stock solution was washed twice and resuspended in 1.2 ml of chilled IAP buffer. In total, 400 µl of the washed WGA slurry was transferred to their appropriate labeled 1.5 ml tubes, and their respective peptide samples were then transferred. Samples were incubated at 4 °C with end-over-end tumbling. After incubation, samples were pelleted by centrifugation for 1 min. Flow-throughs were transferred to separate tubes and preserved. Beads were first washed once with chilled IAP buffer and then twice with

chilled PBS. Peptides were then eluted by resuspending samples in 300 µl of 0.15% TFA. The resuspended samples were then immediately transferred to the top of equilibrated Stage-tips and allowed to trans-elute onto C18 disks for 5 min. Stage tips were washed once with 75 µl of 0.1% FA and eluted with 50 µl of 50% ACN with 0.1% FA. Samples were transferred to HPLC vials, frozen, and dried by vacuum centrifugation. Samples were resuspended in 9 µl of solvent A (3% ACN 0.1% FA).

Basic Reversed-Phase (bRP) Chromatography

Four milligrams of S-trap digested synaptosome peptides were fractionated by basic reversed-phase (bRP) chromatography. bRP solvent A was made in an aqueous solution of 5 mM of Ammonium Formate and 2% ACN (pH 10). bRP Solvent B was made in a solution of 5 mM Ammonium Formate and 90% ACN (pH 10). A 4.6 mm inner diameter 250 mm long Agilent Zorbax 300 Extend-C18 (Agilent) was used as the bRP column. Chromatography occurred over a 96-min gradient with fractions collected in a Whatman 2 ml 96 well plate (GE Healthcare). A 1260 Infinity II high-performance liquid chromatography (HPLC) instrument was used (Agilent). The chromatography gradient proceeded with a solvent B composition scheme of: 0% to 5 min, 5% at 7.66 min, 28.5% at 64.98 min, 34% at 70.48 min, 60% at 83.48 min, and 60% solvent B continuing for the remainder of the gradient. Peptides were concatenated into a 12-fraction scheme as previously described (34). Peptides were then frozen and dried by vacuum centrifugation in preparation for O-GlcNAc enrichment.

Mass Spectrometric Data Acquisition

Peptides were separated online using an Easy-nLC 1200 HPLC (Thermo Fisher Scientific). Samples were loaded onto a capillary column with an integrated emitter tip and heated to 52 °C. The capillary column was packed in-house with ReproSil-Pur C18-AQ 1.9 µm beads (Dr Maisch GmbH) to an approximate length of 27 cm. Once loaded onto the capillary column, peptides were separated through acidic reversed-phase chromatography with solvent A composed of 3% ACN and 0.1% FA and solvent B composed of 90% ACN and 0.1% FA. Chromatography proceeded for 110 min with the following gradient of solvent B composition: 2% solvent B at the start, 6% at 1 min, 30% at 85 min, 60% at 94 min, 90% at 95 min, 90% at 100 min, 50% at 101 min, and 50% solvent B continuing for the remainder of the gradient. Chromatography had a flow rate of 200 nl/min until 101 min, at which the flow rate increased to 500 nl/min for the remainder of the gradient.

Data was acquired on a Thermo Scientific Orbitrap Fusion Lumos mass spectrometer. 4/9 µl of sample was injected for each LC-MS/MS run, with the exception of the HCD collision energy (CE) ramp experiment where 3 µl was injected per run. All MS1 scans were performed in the Orbitrap at 60K resolution with a scan range of 350 to 1800 m/z with a maximum injection time of 50 ms. Unless otherwise specified, all samples were assessed using the HCD-pd-ETD method for MS2 data acquisition. HCD scans containing either a 204.0867 m/z or 138.0545 m/z peak within 15 ppm of the top 20 most intense peaks would trigger a separate ETD scan on the same precursor. All ETD in this study was performed with 25% HCD supplemental activation (EThcD). MS2 scans were analyzed in the Orbitrap for both HCD and EThcD at 15,000 resolution and 1.7 m/z isolation window, with 105 ms injection time and normalized AGC of 200% for HCD and a 120 ms injection time with a normalized AGC of 800% for EThcD. HCD was set at 30 CE for antibody and peptide titration experiments. HCD CE was set to 40 for the 12 fraction synaptosome O-GlcNAc enriched samples. A two second cycle duty time was set for MS/MS events, with monoisotopic precursor determination set to "Peptide," a filter intensity threshold set to 1.0e4, a charge state filter between 2 and 8,

and a 30 s dynamic exclusion time. Precursor priorities were set to highest charge state and lowest m/z.

Data for EThcD alone was acquired by reinjection of the 12 fraction O-GlcNAc enriched synaptosome samples. MS1 parameters remained the same as previously described. EThcD was conducted with 25% supplemental HCD activation, analyzed in the Orbitrap at a resolution of 15,000 with an isolation window of 1.7, a normalized AGC target of 60%, and a 50-ms injection time. The duty cycle was set to 20, with monoisotopic precursor determination set to "Peptide," a filter intensity threshold set to 1.0e4, a charge state filter between 2 and 6, and a 45 s dynamic exclusion time.

Data Analysis

Raw files were processed and searched using Byonic v3.9.6 (Protein Metrics Inc). Cleavage specificity was set to RK with C-terminal cleavage and "Fully Specific" specificity. Only two missed cleavages were allowed. Precursor and product ion mass tolerances were set to 10 and 20 ppm, respectively. Cysteine carbamidomethylation was set as a fixed modification, where methionine oxidation was set to variable and "common2." Protein N-term acetylation was set to "rare1." HexNAc (N-acetylhexosamine) at S and T was set to "common3" and phospho at S, T, and Y was set to "common2." A total of four common and one rare modification was allowed. The Uniprot Mouse database (December 28, 2017, 47,069 entries), containing common laboratory contaminants, was the sequence database used for all subsequent searches. Peptide identifications were extracted using Byos (Protein Metrics) for each run individually and combined into a single report for further processing. Subcellular localization annotation was added to all identifications through Uniprot Retrieve/ID mapping (https://www.uniprot.org/uploadlists/). Unique protein accession number lists were generated by collapsing redundant accession numbers in Excel and submitted to Uniprot Retrieve/ID mapping with the "From UniProtKB AC/ID" and "To UniProtKB" options selected. A subcellular localization column was added to submitted lists through the column editing tab, and reports were downloaded as excel files. Subcellular localization annotations were added to Byos outputs by merging on the protein accession number in Python using the pandas library. PSMs were flagged for the presence of NXS or NXT motifs, where X cannot be P, in the "sequence (unformatted)" column through a pandas regex expression.

PSMs were excluded if they did not include one HexNAc mass addition or fell below a two-dimensional posterior error probability (PEP2D) cutoff of 0.05. "Localized HexNAc PSMs" were putatively identified by filtering out all HexNAc PSMs with a Delta Mod Score greater than or equal to ten (35). Distinct glycopeptides were determined differently for HCD than for EthcD in the synaptosome analysis. For HCD-based identifications, site assignments forced by Byonic's output were removed from the glycopeptide sequence and concatenated with the number of HexNAcs determined by the precursor mass. Sequence_#HexNAcs were then collapsed, retaining the PSM with the highest Delta Mod Score, and are referred to as "Distinct glycopeptides." For EThcD -based identifications, the glycosylation site within the peptide was retained, and PSMs were collapsed retaining the PSM with the highest Delta Mod Score. Distinct glycopeptides for both fragmentation methods were determined in the same way for the reproducibility analysis, where site assignments forced by Byonic's output were removed from the glycopeptide sequence, concatenated with the number of HexNAcs determined by the precursor mass, and collapsed down to unique entries.

Motif analysis was performed by centering each distinct glycosylation site with a Delta Mod Score greater than or equal to 10, with five flanking amino acids. Sequence logos were created by Weblogo (https://weblogo.berkeley.edu/logo.cgi) using default parameters and by PhosphositePlus (https://www.phosphosite.org/staticMotifAnalysis#) using the Automated Selection as background. Identification overlap analysis was performed using distinct glycopeptides or distinct Uniprot accession numbers/proteins from each run.

Quantification of HexNAc Fragment Ion-Containing Spectra

MS2 peak lists were generated for 1 mg mESC peptide WGA & Ab-based enrichments. Spectral.raw files were first processed in Spectrum Mill version BI.07.04.210 (Broad Institute of MIT & Harvard) using the Extractor feature to obtain.mzXML files. Spectral filtering was set between 600 and 6000 Da mass range for precursors and a sequence tag length of greater than 0. The instrument setting was set to ESI Q-Exactive HCD, and carbamidomethylation on cysteines was set as a fixed modification. Spectra from the same precursor or within a retention time window of ±60 s were merged.

Extracted files were uploaded to the MS-Filter feature in Protein Prospector v 6.2.2 (UCSF Mass Spectrometry Facility). Precursor mass range was set to 600 to 6000 Da, charge filter was set to 2 to 8, masses were set to be monoisotopic, parental tolerances were set to 200 ppm, fragment tolerances were set to 20 ppm, an ESI-Q-high-res instrument setting was selected, and Max MSMS Pks was selected.

Peak lists were exported from Protein Prospector as.txt files. Files were uploaded to a Python script using the pandas library. The Python script was developed to convert peak lists into dataframes where scan numbers were row identifiers, predetermined peaks were column identifiers (366.140, 204.087, 186.076, 168.066, 144.065, 138.055, 126.055, and 274.092 within a tolerance of 20 ppm), and dataframe values were peak intensities. A value.count() function was applied to these data frames to obtain the number of spectra (MS2 events) with 204, 366, and 274 peaks.

Distribution of Ratios for HexNAc Fragment Ion Fragments

Applying the Python script described earlier, dataframes for glyco-diagnostic ion containing MS2 spectra were created for CD43, MUC16, MUC16 treated with sialidase, MUC16 treated with sialidase and PNGaseF, and a 1 mg mESC peptide Ab-based enrichment. A GlcNAc/GalNAc ratio column was created by taking the sum of the 138 and 168 peaks divided by the sum of the 126 and 144 peaks for MS2 spectra with those four peaks present. The distribution of this ratio was plotted via a kernel estimate density using distplot() function in the seaborn library. The same strategy was also applied for the 138/144 peak ratios. Outlier values, as defined by values three z-scores away from the mean, were removed for the plotted ratio columns. Computational removal of HexHexNAc and sialic acid was performed by subsetting data frames to contain row indices without values in either the 366.140 or 274.092 columns.

Annotating HexNAc Fragment Ion Intensities and Distributions

MS2 peak lists for the 1 mg mouse embryonic stem cell (supplemental Table S1) and synaptosome enrichments (supplemental Table S2) were generated in the manner described previously, with the additional setting of Ignore EThcD scans applied to obtain.mzXML files from Spectrum Mill. The MS2 peak lists were processed into dataframes as described previously. An additional Python script was developed to match Byos PSMs to MS2 events. MS2 event dataframes were iterated through to find the MS2 event with the minimal difference in retention time and precursor mass for each PSM. The MS2 event with the minimal difference in retention time and precursor mass was considered a match if the charge states between the PSM and the MS2 event were the exact same. The HexNAc fragment ion intensities and 138/144 ratio were annotated for HCD PSMs that found matching MS2 events. These annotations are accessible via the

supplemental tables for review. The Python scripts developed are publicly available (https://github.com/rBurt-96/Burt_MCP_2021).

Preparation of CD43 and MUC16 for HexNAc Fragment Ion Comparison

Samples were prepared as previously described (36, 37). Briefly, recombinant glycoproteins (R&D Systems 1658-PD and 3345-PS) were digested overnight with mucinase SmEnhancin (38) at 37 °C in a total reaction volume of 12 μl in 50 mM ammonium bicarbonate. In sialidase-treated samples, 1 μl of dilute (1:40) sialoEXO (Genovis) was also added to the mucinase digestion. The volume was then increased to 19 μl with 50 mM ammonium bicarbonate. For PNGaseF-treated samples, 1 μl of enzyme (Promega) was added to 99 μl of 50 mM ammonium bicarbonate, and 1 μl of this reaction was added to each mucinase reaction vial. De-N-glycosylation reactions were incubated for 8 to 12 h at 37 °C. For reactions without PNGaseF, 1 μl of ammonium bicarbonate was added. Reduction and alkylation were performed according to ProteaseMax (Promega) protocols. A total of 1 μl of 0.5 M DTT was added, and the samples were incubated at 56 °C for 20 min, followed by the addition of 2.7 μl of 0.55 M iodoacetamide at room temperature for 15 min in the dark. Digestion was completed by adding sequencing-grade trypsin (Promega) at a 1:20 E:S ratio overnight at 37 °C and quenched by adding 0.3 μl of glacial acetic acid. C18 cleanup was performed using 1 ml strataX columns (Phenomenex). Each column was incubated with 1 ml of acetonitrile once, followed by one 1 ml rinse of buffer A (0.1% formic acid in water). The samples were diluted to 1 ml in buffer A and loaded through the column, then rinsed with buffer A. Finally, the samples were eluted with three rinses of 100 μl of buffer B (0.5% formic acid, 80% acetonitrile) and dried by vacuum centrifugation. The samples were reconstituted in 10 μl of buffer A for MS analysis. Samples were analyzed on an Orbitrap Fusion Tribrid (Thermo) using an HCD-pd-ETD method, as described above, with no supplemental activation for ETD.

RESULTS

Anti-O-GlcNAc Antibodies Provide a Simple, Sensitive Approach to Enrich Native O-GlcNAc-Modified Peptides

Four new anti-O-GlcNAc rabbit monoclonal antibodies (Abs) were recently developed by Cell Signaling Technology and are provided as a proprietary mixture noncovalently coupled to Protein A-coated agarose beads (PTMScan O-GlcNAc [GlcNAc-S/T] Motif kit #9522). To initially characterize the performance of this Ab-based enrichment reagent, we generated tryptic digests of whole cell lysates (WCL) from mouse embryonic stem cells (mESCs). mESCs were chosen due to their relatively high levels of O-GlcNAc and because OGT is essential for embryonic development (1, 2, 39). Data were acquired on a Fusion Lumos Tribrid mass spectrometer using an HCD product-triggering EThcD (HCD-pd-EThcD) acquisition method. HCD was performed on all multiply charged precursor ions. Upon detection of the glyco-specific HexNAc fragment ion masses of 204.087 or 138.055 (+/− 15 ppm) in the HCD scan, the precursor was reaccumulated for a separate ETD scan with supplemental HCD (25%) activation (EThcD) (9, 15). As HexNAc fragment ion detection in the HCD scan is crucial to trigger the subsequent EThcD scan necessary for confident O-HexNAc (O-linked

N-acetylhexosamine) site assignment, we first tested various HCD collision energies (CEs) using WCL digests of mESC enriched using the anti-O-GlcNAc Ab mixture. We found that higher HCD CEs caused a slight decrease in the number of HexNAc-containing HCD peptide spectral matches (PSMs) but increased the number of HexNAc-containing EThcD-derived PSMs (Fig. 1A). Higher HCD CEs also resulted in higher spectral quality for both HCD and the product-triggered EThcD (pd-EThcD), measured by the improved confidence (PEP2D score) in glycopeptide identifications (Fig. 1B). This coincided with more PSMs with higher charge states (Fig. 1C). We therefore chose these data acquisition conditions for analysis of O-GlcNAcylated peptides.

We next optimized the amount of Ab-conjugated beads needed to observe the highest number of HexNAc-containing PSMs and the lowest number of unmodified PSMs using 1 mg of mESC WCL digest. We found that use of one-fourth the amount of Ab recommended for O-GlcNAc-peptide yielded equivalent numbers of HexNAc-containing PSMs as the full amount while providing the best enrichment specificity (Fig. 1, D and E), consistent with previous anti-PTM Ab work (40, 41). Reproducibility was evaluated across three parallel enrichments of the same mESC sample. We found the overlap of glycoprotein identifications to be 48% across all three replicates or an average of 61% across any two (Fig. 1F). At the distinct glycopeptide level, the overlap across all three enrichments was 34%, or an average of 49% for across any two (Fig. 1G). This level of reproducibility across replicated is similar to that obtained using antiphosphotyrosine antibodies (42).

To evaluate the sensitivity of the anti-O-GlcNAc Abs, we enriched from decreasing amounts of input peptides, keeping the amount of Ab-beads constant. We observed a nearly linear decrease in HexNAc-containing PSMs as input decreased, where the greatest drop occurred between 125 and 250 μg peptides (Fig. 2A). The site, peptide, and protein identification results of the 1 mg enrichment are available in supplemental Table S1. We compared these results to an alternative, commonly used native O-GlcNAc-modified peptide enrichment reagent WGA (23, 24, 26, 43), at our high and low input loads. In batch mode enrichments, the anti-O-GlcNAc Abs greatly improved the number of HexNAc-containing PSMs across all input levels compared with WGA (Fig. 2A). It should be noted WGA was used in batch enrichment format rather than in its more commonly used column format for a more direct interpretation of the results (24, 26, 27). Separating the HexNAc-containing PSMs by unimolecular dissociation mode showed similar trends, illustrating that although HCD provides little site assignment information, it is responsible for roughly two-thirds of PSM identifications (Fig. 2B).

Anti-O-GlcNAc Antibodies Are Highly Specific for O-GlcNAc-Modified Peptides

To evaluate the specificity of the anti-O-GlcNAc Abs for O-GlcNAc-modified peptides, we compared the results
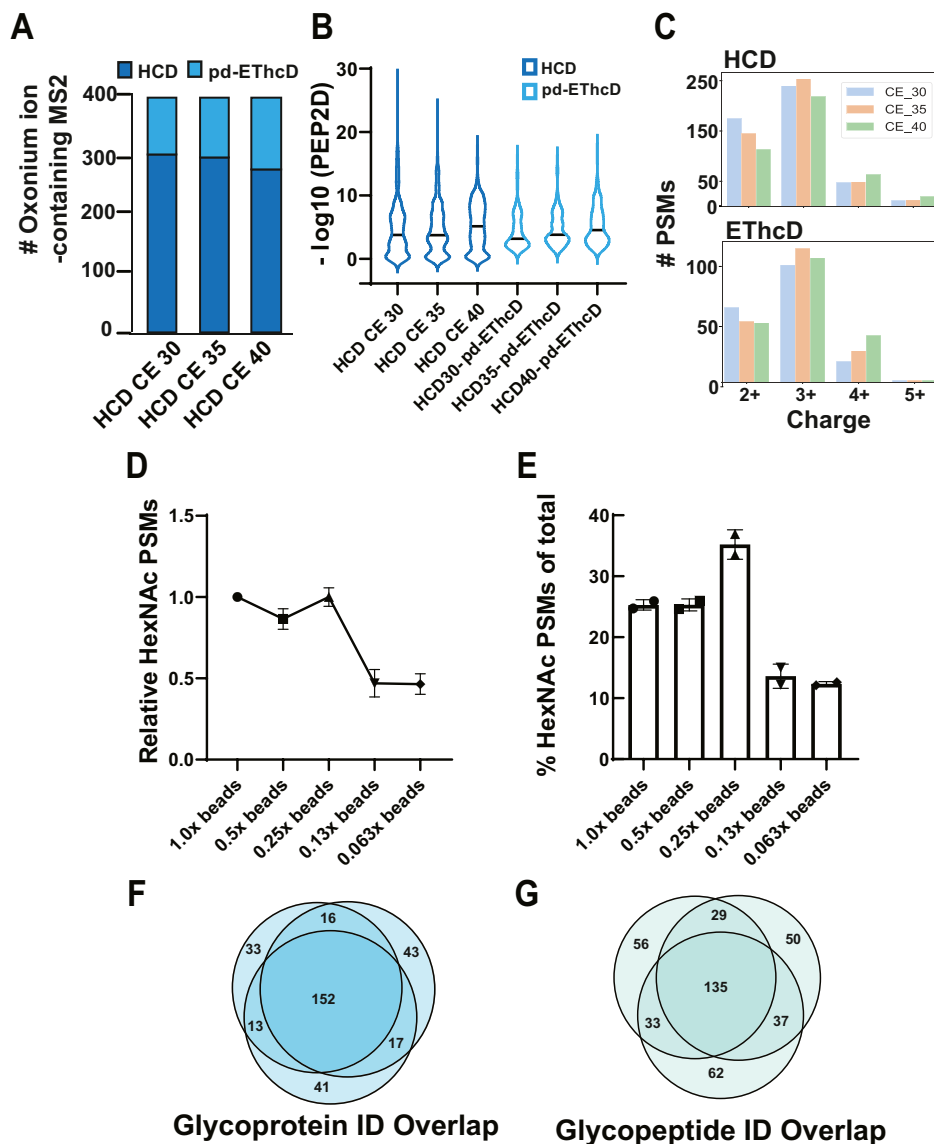
FIG. 1. **Optimization of data acquisition and antibody enrichment conditions.** *A*, total HexNAc-containing PSMs across different HCD collision energies (CEs) (300 μg of mESC WCL). Both HCD and product triggered EThcD (pd-EthcD) PSMs are shown. Only HCD CEs were altered. *B*, distribution of PEP2D scores for HexNAc-containing PSMs separated by dissociation method. Only HCD CEs were altered. EThcD supplemental activation energy was held constant at 25%. *C*, number of PSMs per charge state across different HCD CEs. *D*, relative HexNAc-containing PSMs titrating down the amount of antibody from the recommended levels. Values are relative to the 1× experiment. *E*, O-HexNAc glycopeptide specificity for decreasing amounts of antibody. Specificity was calculated by dividing the number of HexNAc-containing PSMs by the total number of PSMs. *F*, overlap in distinct glycoprotein identifications from triplicate 1 mg mESC WCL enrichments. *G*, overlap in distinct glycopeptide identifications from three separate enrichments from 1 mg aliquots of digested mESC WCL digests.

obtained with the Abs to the use of WGA, a lectin with affinity for GlcNAc and a diverse array of glycosylation states (28). The extracted ion chromatograms (XICs) of the WGA-based enrichment runs showed a high prevalence of diagnostic ions for HexNAc (204.087 m/z), HexHexNAc (366.140 m/z), and sialic acid minus water (SA, 274.092 m/z) (Fig. 2C). The diagnostic ions of m/z 366 and 274 are derived from extended glycan structures (*e.g.*, biantennary and more highly branched and modified oligosaccharides) on peptides and likely not from an OGT substrate (44). Whereas the observation of the

204 m/z peak in the absence of 366 or 274 is suggestive of an O-linked HexNAc and not extended carbohydrate structures. In contrast to the WGA results, XICs from the Ab-based enrichment showed a high number of m/z 204 HexNAc peaks with few HexHexNAc or SA diagnostic ions (Fig. 2D). Additional evidence of specificity of the Abs for O-GlcNAc was obtained by quantifying the number of MS2 scans that exhibited the respective diagnostic HexNAc fragment ions (Fig. 2E). While the WGA-based enrichments had twice as many HexNAc-containing spectra than the Ab-based
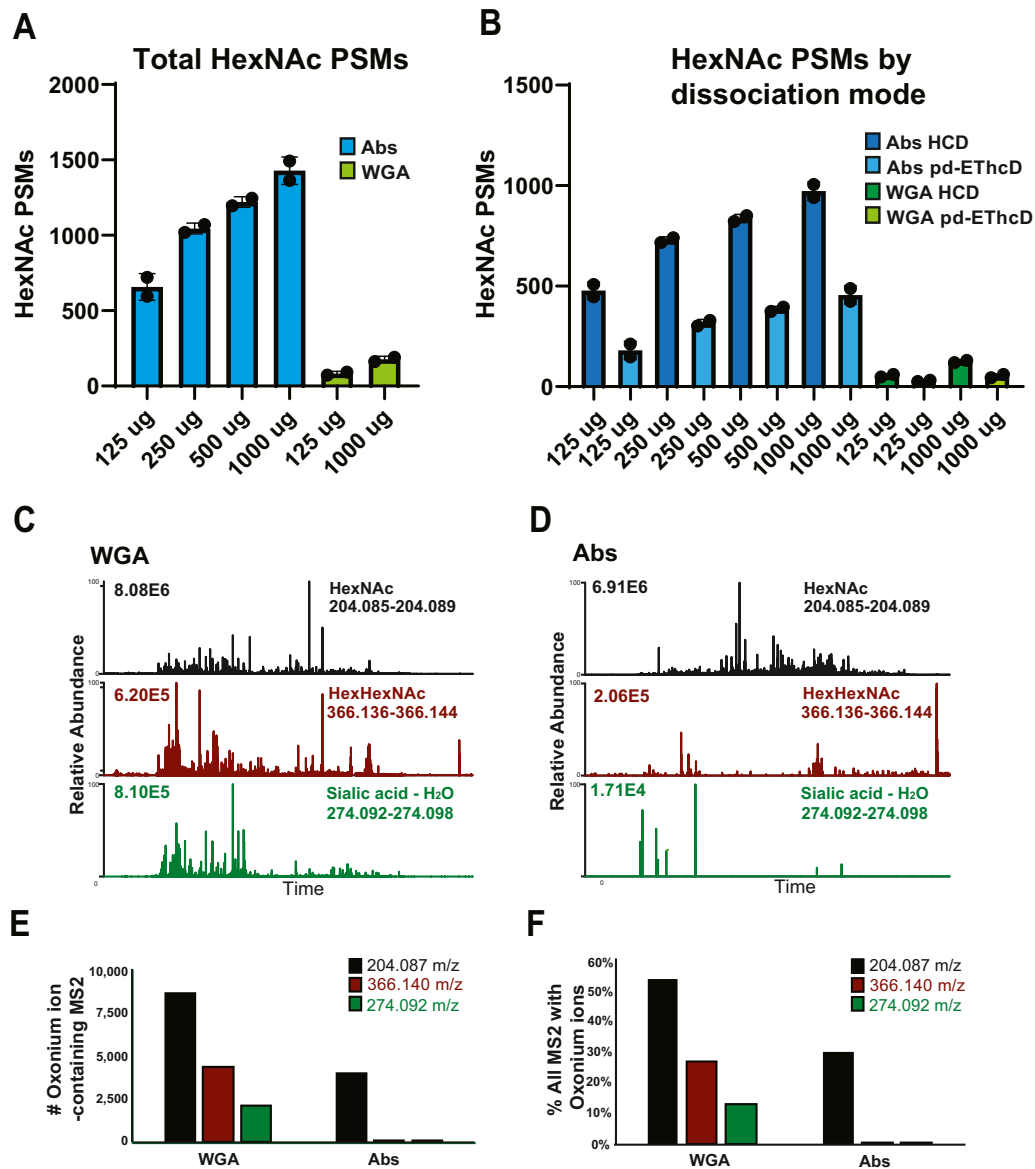
FIG. 2. **Characterization of sensitivity and specificity of anti-O-GlcNAc antibodies and their comparison with WGA.** *A*, total number of HexNAc-containing PSMs obtained by enrichment with Ab (*blue bars*) or WGA (*green bars*) across a range of input levels of mESC WCL digest; the optimal amount of Ab determined in Figure 1 was used. *B*, numbers of HexNAc-containing PSMs obtained at varying input levels and separated by dissociation method. "pd-EThcD" indicates the EThcD scan was triggered by detection of HexNAc fragment ions in the preceding HCD scan. *C*, extracted ion chromatograms (XICs) from the WGA-based glycopeptide enrichment experiments of HexNAc fragment ions (+/– 20 ppm of calculated masses) diagnostic for respective carbohydrates; 204.087 m/z represents HexNAc (*black*), 366.140 m/z for HexHexNAc (*burgundy*), and 274.092 m/z for sialic acid minus water (*green*). Base peak intensities are shown on the *upper left*. *D*, same as in (*C*), above, but for the anti-O-GlcNAc Ab-based glycopeptide enrichments. *E*, number of MS2 spectra containing one of the aforementioned HexNAc fragment ions from the respective enrichment strategies from 1 mg of peptide inputs. *F*, percentage of all MS2 scans with aforementioned HexNAc fragment ions from the respective enrichment strategies from 1 mg.

enrichments, they also contained 91- and 264-fold more HexHexNAc- and SA-containing spectra, respectively, compared with the Ab-based enrichments. The percentage of total MS2 spectra with the respective diagnostic ions showed similar trends (Fig. 2*F*). These data suggest that the anti-O-GlcNAc Abs have a high degree of specificity for non-extended O-HexNAc-modified peptides and do not

enrich HexNAc-containing complex glycans to a significant extent.

While the anti-O-GlcNAc monoclonal Abs were developed against synthetic O-GlcNAc-modified peptides with degenerative amino acid sequences, we questioned whether the HexNAc signal we observed was coming from non-specific enrichment of O-linked N-acetylgalactosamine

(O-GalNAc)-modified peptides. Singly O-GalNAc-modified serines and threonines (a.k.a. Tn antigen) are the base of the Core 1 O-glycosylation pathway in the Golgi (Reviewed in Schjoldager *et al.* (45)). GlcNAc and GalNAc are isomeric, displaying the same HexNAc fragment ions, and thus cannot be differentiated by mass alone. In work analyzing synthetic glycopeptides by HCD-MS/MS, other groups have suggested that O-GlcNAc and O-GalNAc undergo different dissociation pathways and the ratio of the resultant HexNAc fragment ions can be used to discriminate between these epimers (21, 46). We developed a script to calculate the ratio of select HexNAc fragment ions for each individual HCD MS2 spectrum, focusing on the GlcNAc/GalNAc ratio suggested by Halim *et al.*, [(138 + 168)/(126 + 144)] (46). To test the ability of our

script to differentiate between GlcNAc and GalNAc HexNAc fragment ion ratio distributions, we turned to proteins with well-characterized glycosylation states for digestion and analysis by HCD-MS/MS. For a model that is predominantly O-GalNAcylated, we used the mucin SPN (a.k.a. CD43 or Leukosialin), which has 60 documented O-GalNAc sites, and a single annotated N-linked glycosylation site (47). For a protein containing large numbers of both O-GalNAc and N-linked glycans, we used the SEA domain of human MUC16. Plotting the GlcNAc/GalNAc ratios for these two proteins, we found that SPN, the highly O-GalNAc-modified protein, showed a tight distribution of ratios with a mean of 0.8. MUC16 had a mean of 3.3 and a much broader distribution, likely due to its diverse range of glycosylation types (Fig. 3*A*). Consistent with
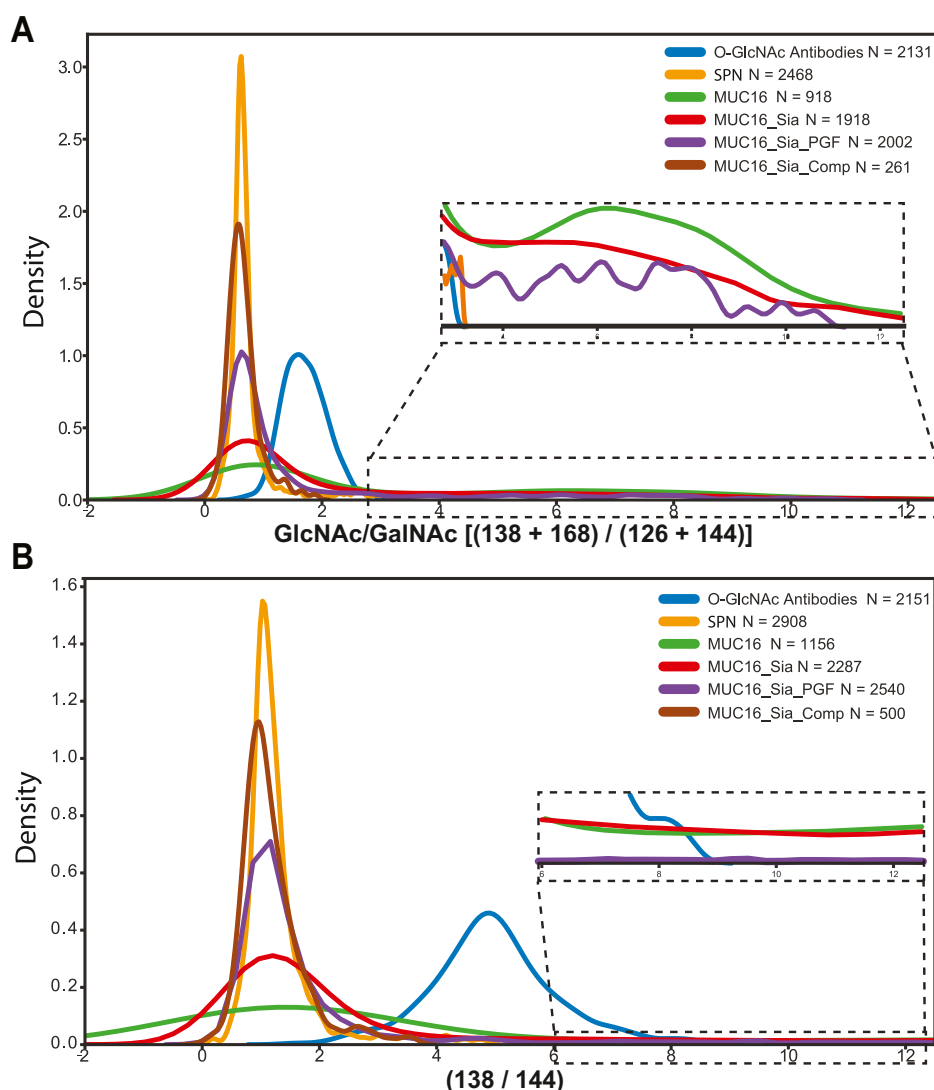

Fɪɢ. 3. **Discrimination between GlcNAc and GalNAc through HexNAc fragment ion ratio distributions.** *A*, kernel density plots for the distribution of GlcNAc/GalNAc ratios [(138 + 168)/(126 + 144)], as determined by Halim *et al* (46) for control proteins and conditions. Inset is a zoom of the higher ratios for all values highlighted. Comp, computational removal of MS2 spectra containing diagnostic ions for 366 and 274; N, the number of MS2 spectra used to derive ratios; PGF, PNGase F treatment; Sia, sialidase treatment. *B*, kernel density plots for the distribution of 138/144 peaks across the same samples in Figure 3*A*. Inset is a zoom of the higher ratios. Abbreviations are the same from Figure 3*A*.

Halim *et al.*, the large GlcNAc/GalNAc ratios of MUC16 were likely due to N-linked glycans containing GlcNAc residues. To further test this hypothesis, we removed specific sugars from MUC16 in multiple ways. First, we treated MUC16 with sialidase, which removes terminal SA, and replotted the GlcNAc/GalNAc ratio. Sialidase treatment increased the number of lower ratios but kept the same wide distribution as untreated MUC16. We next removed the N-linked glycans from MUC16 with PNGase F, which cleaves glycans from Asn but not Ser or Thr. The GlcNAc/GalNAc ratios from PNGase F-treated MUC16 greatly increased the number of low ratios and decreased the number of high ratios, creating a distribution resembling SPN's glycopeptides. Finally, we computationally removed MS2 spectra from MUC16 that contained either 366.140 or 274.092 m/z to simulate a perfect enzymatic reaction that should leave only O-GalNAc-modified peptides. The GlcNAc/GalNAc distribution of this "computational glycosylation removal" showed a similar pattern to that of the PNGase F-treated MUC16 and the heavily O-GalNAc-modified SPN. These results show our ratio distribution analysis can discriminate between O-HexNAc and hybrid, complex glycans.

To test whether our GlcNAc/GalNAc distribution analysis could differentiate between O-GalNAc and our putative O-GlcNAc-modified peptides, we plotted the GlcNAc/GalNAc ratios for MS2 spectra generated from the anti-O-GlcNAc Ab-based enrichments. We found that the GlcNAc/GalNAc ratios from these samples were distinctly higher (mean = 1.7) than those from SPN and had a tighter distribution than those derived from MUC16 (Fig. 3*A*). Consistent with Halim *et al.*, our putative O-GlcNAc-containing spectra had intermediate ratios compared with O-GalNAc and more complex glycans derived from MUC16, suggesting we are indeed enriching O-GlcNAc-modified peptides. To better discriminate between O-GlcNAc and O-GalNAc, we examined alternate HexNAc fragment ion ratio distributions. Our data showed that using a simpler ratio (138 m/z/144 m/z) provided stronger separation between O-GalNAc-only spectra and the data from our Ab-based enrichments (Fig. 3*B*). These results demonstrate that the ratio of the 138/144 HexNAc fragment ions can differentiate between O-GlcNAc and O-GalNAc. Together, these data suggest that the anti-O-GlcNAc Abs show strong specificity toward O-GlcNAc and not O-GalNAc or GlcNAc residing in more complex glycan-modified peptides.

Anti-O-GlcNAc Antibodies Allow for Deep O-GlcNAc Proteome Profiling from Tissues

An advantage of native O-GlcNAc-modified peptide enrichment over alternative strategies such as metabolic labeling is the ability to enrich glycopeptides from tissues. To test the capacity of the anti-O-GlcNAc Abs to enrich from tissues, we turned to synaptosome preparations (30), a classic model used

for O-GlcNAc-modified peptide enrichment (12, 24, 43). Murine synaptosomes were solubilized in 3% SDS, subjected to suspension trapping (31, 32), and digested with trypsin. Four milligrams of the synaptosome digest was fractionated with basic reversed-phase (bRP) fractionation into 96 fractions, then concatenated back to 12 fractions (34). Each final fraction containing approximately 330 µg of peptide was enriched for O-GlcNAc-modified peptides using the anti-O-GlcNAc Abs, and the data was acquired using either the HCD-pd-EThcD method or EThcD alone without a preceding HCD scan.

For both data acquisition methods, the number of HexNAc-containing PSMs was fairly evenly spaced across the 12 fractions with the exception of fraction one (Fig. 4*A*). Using the HCD-pd-EThcD acquisition method, we identified 6872 high-confidence HexNAc-containing PSMs, 4427 of which were identified by HCD and 2445 by pd-EThcD (Fig. 4*B*). After collapsing redundant HexNAc PSMs, 1328 distinct glycopeptides remained by HCD and 1289 distinct glycopeptides by pd-EThcD. Of the distinct glycopeptide forms identified by pd-EThcD, 714 had a Delta Mod Score of 10 or greater, the suggested lower limit of confidence for site assignment by Byonic (35). Using this metric, we identified 1040 O-HexNAc sites through pd-EThcD (supplemental Table S2). While HCD of O-linked glycopeptides generally does not retain site localization information due to the prominent cleavage of the O-glycosidic bond, we found 190 distinct glycopeptide forms with Delta Mods scores of 10 or greater, leading to 272 O-HexNAc sites. Site assignments with HCD were enabled largely by having fewer S/T residues within the peptides (Fig. 4*C*). With EThcD acquisition alone, we identified 1532 high-confidence HexNAc-containing PSMs containing 1316 localized HexNAc PSMs (Fig. 4*B*). After collapsing redundant glycoPSMs, EThcD alone identified 919 distinct glycopeptides with 791 O-HexNAc sites with a Delta Mod Score of 10 or greater (supplemental Table S3). Although HCD provides little O-HexNAc site information, its use led to twice as many distinct glycoprotein identifications compared with EThcD alone (Fig. 4*D*). These data show that HCD provides complementary information on glycopeptide and glycoprotein identifications to EThcD alone, without sacrificing the number of EThcD-based identifications.

To better understand the properties of O-HexNAc-modified peptides enriched using the anti-O-GlcNAc Abs, we analyzed the characteristics of the identified glycopeptides. We first looked at the extent of glycosylation per peptide. Of all HexNAc PSMs across both datasets, 62% had a single HexNAc per glycopeptide (Fig. 4*E*). Interestingly, 38% of HexNAc PSMs were multiply O-HexNAcylated, roughly twice as many as seen previously (24). We next explored the frequency of amino acid occurrences in the sequence surrounding the O-HexNAc site. We used two common motif analysis tools, Weblogo (48) and PhosphositePlus (49), as they both use different background corrections (see Experimental
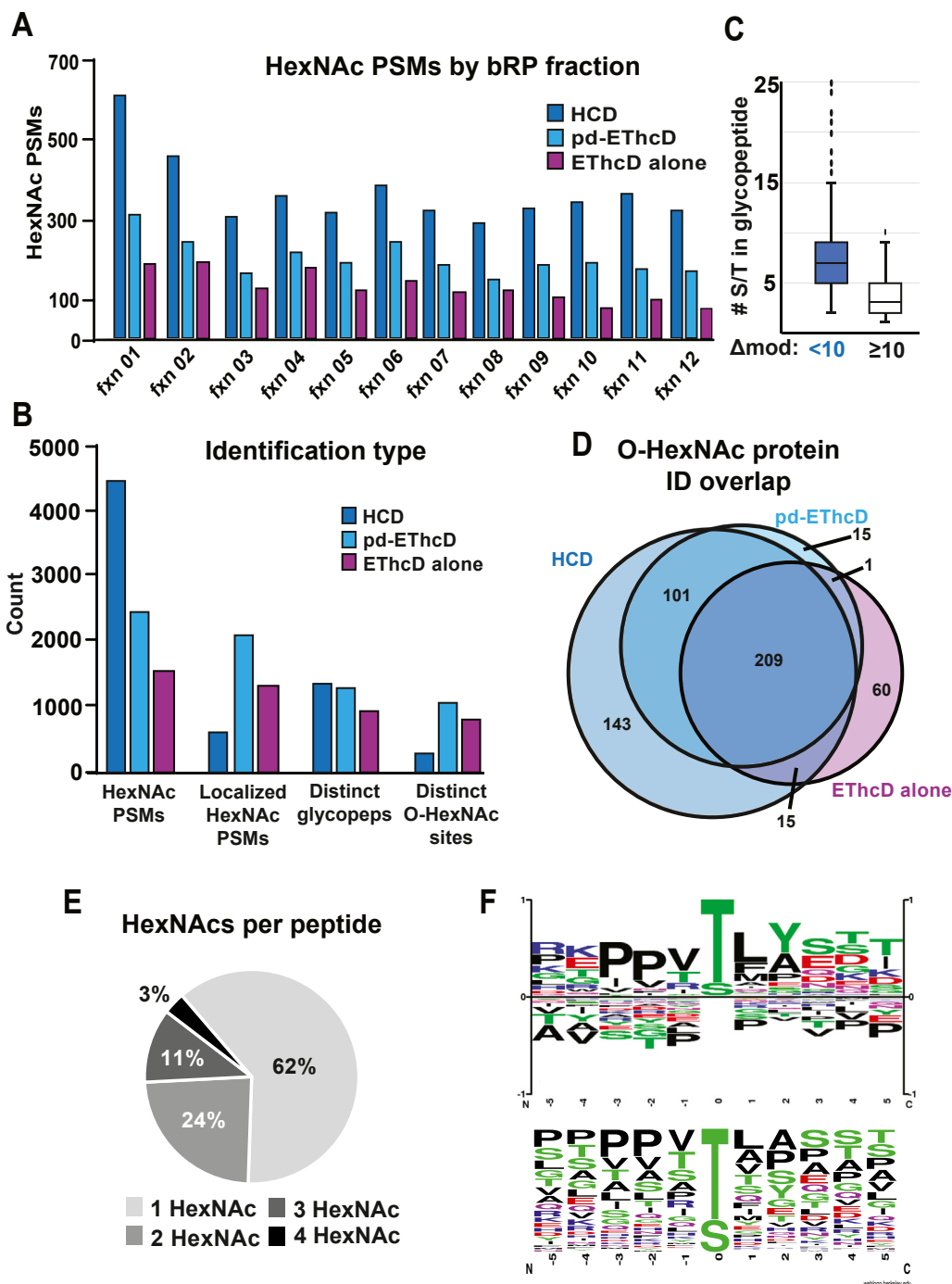
Fig. 4. **Benchmarking anti-O-GlcNAc antibodies with postsynaptic density preparations across different data acquisition methods.** *A*, HexNAc-containing PSMs across basic reversed-phase fractions. *Blue bars* indicate the number of HexNAc-containing PSMs identified by HCD or the subsequent pd-EThcD scan. *Purple bars* indicate the number of HexNAc-containing PSMs from EThcD alone. *B*, breakdown of various identifications from the cumulative analysis analyzing O-GlcNAc peptides from mouse synaptosome fractions. "HexNAc PSMs" refer to PSMs that contain at least one HexNAc modification, regardless of localization. "Localized HexNAc PSMs" refer to all glycosylation sites within HexNAc-containing PSMs that had a Delta Mod Score of 10 or greater. "Distinct Glycopeps" refers to nonredundant glycopeptides identified by the respective dissociation mode. However, HCD-based identifications were treated as if site localization could not be used (see Experimental Procedures). "Distinct O-HexNAc sites" refer to the count of HexNAc site assignments from all sites within nonredundant HexNAc-containing PSMs with Delta Mod Score of 10 or greater. *C*, number of S/T residues within HCD-based glycopeptide identifications with a Delta Mod Score of less than 10 (*blue box*) or 10 or greater (*white box*) to symbolize the trend HCD mostly identifies O-HexNAc sites in peptide with few amino acid possibilities for modification. *D*, overlap of proteins identified as O-HexNAc modified by the respective data acquisition methods. Only a single HexNAc-containing PSM was required to be called an O-GlcNAc-modified protein. "pd-EThcD" indicates the EThcD scan triggered from

Procedures). We found both sequence logo generators showed threonine to be more often O-HexNAc modified than serine (Fig. 4F), a departure from previous studies (12, 24, 50, 51). Consistent with previous work, we found a preference for proline residues at position -2 and -3, some hydrophobic character at position +1 and +2, and an enrichment for hydroxyl-containing residues C-terminal to the distinct O-HexNAc sites. Together these data suggest that the anti-O-GlcNAc Abs efficiently and specifically enrich for O-GlcNAc-modified peptides possessing the weak primary sequence preference of OGT.

DISCUSSION

Here we report the in-depth characterization and application of new anti-O-GlcNAc Abs to enrich for native O-GlcNAc-modified peptides from cells and tissues. We also describe the benefits and drawbacks to combining HCD with EThcD for the analysis of a complex mixture of O-GlcNAc peptides.

Benefits and Drawbacks of HCD Product-Triggering EThcD

While the necessity of EThcD for O-linked glycopeptide analysis is well established (13–15), the value of combining it with HCD for global O-GlcNAc proteome analysis proved manifold. The HCD-derived HexNAc-PSMs provided additional, complementary evidence for peptide O-GlcNAcylation, though site localization was minimal. The speed of HCD also provided a substantial increase in O-GlcNAcylated peptide/protein identifications compared with EThcD alone, though not all HCD scans with the 204 m/z triggered a subsequent EThcD scan (41,157 HCD scans containing 204 m/z and only 12,409 EThcD scans). We saw little to no decrease in O-GlcNAc site information when using the HCD "scouting" scan to trigger EThcD. In fact, we found that increasing HCD collision energies led to higher scoring O-HexNAc-modified PSMs in both HCD and EThcD. The improved scoring observed in EThcD may be due to the greater charge state distribution of PSMs when acquired with higher HCD CE, promoting EThcD fragmentation on higher charge state precursors, improving EThcD spectral quality. This suggests that HCD "scouting" scans to trigger EThcD will be compatible with isobaric labeling for quantitation (iTRAQ, TMT) of O-GlcNAcylation (9, 15, 27) or to leverage combined information from both dissociation events (52).

Anti-O-GlcNAc Abs Provide a Simple, Efficient, and Specific Enrichment Strategy for Analysis of Native O-GlcNAc Peptides from Cells and Tissues

We found the novel anti-O-GlcNAc Abs coupled to Protein A-coated agarose beads to be a strong enrichment reagent to analyze O-GlcNAc-modified peptides from cells and tissues. A major benefit of the anti-O-GlcNAc Abs is their specificity toward O-GlcNAc versus O-GalNAc or extended, GlcNAc-containing glycans. We saw little evidence of HexHexNAc—or sialic acid—containing peptides immuno-precipitated by the Abs, suggesting they do not recognize GlcNAc within glycan chains. This specificity was reflected in the higher number of HexNAc-containing PSMs compared with WGA, where our search parameters did not allow for any glycan mass additions. We also found evidence for a strong specificity of the anti-O-GlcNAc Abs toward O-GlcNAc and not its epimer O-GalNAc. Leveraging prior studies (21, 22, 46), we confirmed that the distributions of the ratios of HexNAc fragment ions were distinct between a heavily O-GalNAc-modified protein and our putative O-GlcNAc-modified peptides. We then modified the previous GlcNAc/GalNAc ratio to utilize only two fragment ions, 138/144, which provided a clearer separation between the two glycopeptide forms. We suggest this ratio should be used in future studies to help differentiate between these isomeric molecules. Future work will be needed to understand how to use both these ratios, and potentially other HexNAc fragment ions, to estimate the relative sugar composition of more complex glycans, especially when examining the higher ratios.

The simplicity of use and sensitivity of the anti-O-GlcNAc Abs are also major advances for global O-GlcNAc analysis. Compared with other landmark studies (12, 24, 43), we achieved similar numbers of distinct glycopeptides and O-GlcNAc sites, using nearly tenfold less material, and a significantly smaller fraction of labor and instrument time. The sample preparation time is dramatically decreased compared with other O-GlcNAc enrichment protocols and can be done with little to no special equipment or chemistries. As this enrichment strategy does not require incubating live cells with chemically modified sugars, it is compatible with enrichment from tissues, demonstrated by our analysis of synaptosomes. We expect that the enrichment of O-GlcNAc-modified peptides using these Abs can be easily coupled to serial PTM enrichment workflows routinely used in our laboratory (34).

In summary, we provide an in-depth characterization of novel anti-O-GlcNAc Abs and their application to the O-GlcNAc-modified proteome analysis of cells and tissues. These Abs are efficient and specific toward O-GlcNAc, but not other glycosylation states. This strategy is simple and sensitive, allowing the rapid enrichment of large numbers of O-GlcNAc-modified peptides from complex biological samples. We believe these Abs will provide a significant technical advancement in our ability to analyze and understand O-GlcNAc signaling.

---

the detection of HexNAc fragment ions in the antecedent HCD scan. *E*, distribution of the extent of HexNAc modifications of HexNAc-containing PSMs across all high-confidence HexNAc-containing PSMs. *F*, sequence logo of distinct O-HexNAc sites according to two different analysis methods (see Experimental Procedures). The top logo was derived from Phosphosite.org; the bottom from Weblogo (Berkeley).

REFERENCES

1. O'Donnell, N., Zachara, N. E., Hart, G. W., and Marth, J. D. (2004) Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability. *Mol. Cell. Biol.* **24**, 1680–1690
2. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5735–5739
3. Liu, B., Salgado, O. C., Singh, S., Hippen, K. L., Maynard, J. C., Burlingame, A. L., Ball, L. E., Blazar, B. R., Farrar, M. A., Hogquist, K. A., and Ruan, H.-B. (2019) The lineage stability and suppressive program of regulatory T cells require protein O-GlcNAcylation. *Nat. Commun.* **10**, 354
4. Yang, Y., Li, X., Luan, H. H., Zhang, B., Zhang, K., Nam, J. H., Li, Z., Fu, M., Munk, A., Zhang, D., Wang, S., Liu, Y., Albuquerque, J. P., Ong, Q., Li, R., *et al* (2020) OGT suppresses S6K1-mediated macrophage inflammation and metabolic disturbance. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 16616–16625
5. Howerton, C. L., and Bale, T. L. (2014) Targeted placental deletion of OGT recapitulates the prenatal stress phenotype including hypothalamic mitochondrial dysfunction. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9639–9644
6. Keembiyehetty, C., Love, D. C., Harwood, K. R., Gavrilova, O., Comly, M. E., and Hanover, J. A. (2015) Conditional knock-out reveals a requirement for O-linked N-acetylglucosaminase (O-GlcNAcase) in metabolic homeostasis. *J. Biol. Chem.* **290**, 7097–7113
7. Wang, A. C., Jensen, E. H., Rexach, J. E., Vinters, H. V., and Hsieh-Wilson, L. C. (2016) Loss of O-GlcNAc glycosylation in forebrain excitatory neurons induces neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 15120–15125
8. Zachara, N., Akimoto, Y., and Hart, G. W. (2015) The O-GlcNAc modification. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., eds. *Essentials of Glycobiology*, 3rd Ed., Cold Spring Harbor, New York, NY: 239–251
9. Zhao, P., Viner, R., Teo, C. F., Boons, G.-J., Horn, D., and Wells, L. (2011) Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. *J. Proteome Res.* **10**, 4088–4104
10. Chalkley, R. J., and Burlingame, A. L. (2001) Identification of GlcNAcylation sites of peptides and alpha-crystallin using Q-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **12**, 1106–1113
11. Mirgorodskaya, E., Roepstorff, P., and Zubarev, R. A. (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem.* **71**, 4431–4436
12. Chalkley, R. J., Thalhammer, A., Schoepfer, R., and Burlingame, A. L. (2009) Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8894–8899
13. Myers, S. A., Daou, S., Affar, E. B., and Burlingame, A. (2013) Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments-a study of the O-GlcNAcylated protein host cell factor C1. *Proteomics* **13**, 982–991
14. Wang, Z., Udeshi, N. D., Slawson, C., Compton, P. D., Sakabe, K., Cheung, W. D., Shabanowitz, J., Hunt, D. F., and Hart, G. W. (2010) Extensive

crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Sci. Signal.* **3**, ra2

15. Riley, N. M., Malaker, S. A., Driessen, M. D., and Bertozzi, C. R. (2020) Optimal dissociation methods differ for N- and O-glycopeptides. *J. Proteome Res.* **19**, 3286–3301

16. Maynard, J. C., and Chalkley, R. J. (2021) Methods for enrichment and assignment of N-acetylglucosamine modification sites. *Mol. Cell. Proteomics* **20**, 100031

17. Qin, W., Lv, P., Fan, X., Quan, B., Zhu, Y., Qin, K., Chen, Y., Wang, C., and Chen, X. (2017) Quantitative time-resolved chemoproteomics reveals that stable O-GlcNAc regulates box C/D snoRNP biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6749–E6758

18. Boyce, M., Carrico, I. S., Ganguli, A. S., Yu, S.-H., Hangauer, M. J., Hubbard, S. C., Kohler, J. J., and Bertozzi, C. R. (2011) Metabolic cross-talk allows labeling of O-linked β-N-acetylglucosamine-modified proteins *via* the N-acetylgalactosamine salvage pathway. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3141–3146

19. Pedowitz, N. J., Zaro, B. W., and Pratt, M. R. (2020) Metabolic labeling for the visualization and identification of potentially O-GlcNAc-modified proteins. *Curr. Protoc. Chem. Biol.* **12**, e81

20. Khidekel, N., Ficarro, S. B., Peters, E. C., and Hsieh-Wilson, L. C. (2004) Exploring the O-GlcNAc proteome: Direct identification of O-GlcNAc-modified proteins from the brain. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13132–13137

21. Malaker, S. A., Penny, S. A., Steadman, L. G., Myers, P. T., Loke, J. C., Raghavan, M., Bai, D. L., Shabanowitz, J., Hunt, D. F., and Cobbold, M. (2017) Identification of glycopeptides as posttranslationally modified neoantigens in leukemia. *Cancer Immunol. Res.* **5**, 376–384

22. Wang, X., Yuan, Z.-F., Fan, J., Karch, K. R., Ball, L. E., Denu, J. M., and Garcia, B. A. (2016) A novel quantitative mass spectrometry platform for determining protein O-GlcNAcylation dynamics. *Mol. Cell. Proteomics* **15**, 2462–2475

23. Xu, S.-L., Chalkley, R. J., Maynard, J. C., Wang, W., Ni, W., Jiang, X., Shin, K., Cheng, L., Savage, D., Hühmer, A. F. R., Burlingame, A. L., and Wang, Z.-Y. (2017) Proteomic analysis reveals O-GlcNAc modification on proteins with key regulatory functions in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1536–E1543

24. Trinidad, J. C., Barkan, D. T., Gulledge, B. F., Thalhammer, A., Sali, A., Schoepfer, R., and Burlingame, A. L. (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell. Proteomics* **11**, 215–229

25. Myers, S. A., Panning, B., and Burlingame, A. L. (2011) Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9490–9495

26. Maynard, J. C., Fujihira, H., Dolgonos, G. E., Suzuki, T., and Burlingame, A. L. (2020) Cytosolic N-GlcNAc proteins are formed by the action of endo-β-N-acetylglucosaminidase. *Biochem. Biophys. Res. Commun.* **530**, 719–724

27. White, C. W., 3rd, Fan, X., Maynard, J. C., Wheatley, E. G., Bieri, G., Couthouis, J., Burlingame, A. L., and Villeda, S. A. (2020) Age-related loss of neural stem cell O-GlcNAc promotes a glial fate switch through STAT3 activation. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22214–22224

28. Trinidad, J. C., Schoepfer, R., Burlingame, A. L., and Medzihradszky, K. F. (2013) N- and O-glycosylation in the murine synaptosome. *Mol. Cell. Proteomics* **12**, 3474–3488

29. Maynard, J. C., Burlingame, A. L., and Medzihradszky, K. F. (2016) Cysteine S-linked N-acetylglucosamine (S-GlcNAcylation), a new post-translational modification in mammals. *Mol. Cell. Proteomics* **15**, 3405–3411

30. Dejanovic, B., Huntley, M. A., De Mazière, A., Meilandt, W. J., Wu, T., Srinivasan, K., Jiang, Z., Gandham, V., Friedman, B. A., Ngu, H., Foreman, O., Carano, R. A. D., Chih, B., Klumperman, J., Bakalarski, C., *et al.* (2018) Changes in the synaptic proteome in tauopathy and rescue of tau-induced synapse loss by C1q antibodies. *Neuron* **100**, 1322–1336.e7

31. Zougman, A., Selby, P. J., and Banks, R. E. (2014) Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006-0

32. HaileMariam, M., Eguez, R. V., Singh, H., Bekele, S., Ameni, G., Pieper, R., and Yu, Y. (2018) S-Trap, an ultrafast sample-preparation approach for shotgun proteomics. *J. Proteome Res.* **17**, 2917–2924

33. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906

34. Mertins, P., Tang, L. C., Krug, K., Clark, D. J., Gritsenko, M. A., Chen, L., Clauser, K. R., Clauss, T. R., Shah, P., Gillette, M. A., Petyuk, V. A., Thomas, S. N., Mani, D. R., Mundt, F., Moore, R. J., *et al.* (2018) Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat. Protoc.* **13**, 1632–1661

35. Bern, M., Kil, Y. J., and Becker, C. (2012) Byonic: Advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics*

36. Malaker, S. A., Pedram, K., Ferracane, M. J., Bensing, B. A., Krishnan, V., Pett, C., Yu, J., Woods, E. C., Kramer, J. R., Westerlind, U., Dorigo, O., and Bertozzi, C. R. (2019) The mucin-selective protease StcE enables molecular and functional analysis of human cancer-associated mucins. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7278–7287

37. Shon, D. J., Malaker, S. A., Pedram, K., Yang, E., Krishnan, V., Dorigo, O., and Bertozzi, C. R. (2020) An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 21299–21307

38. Wu, P., Sun, P., Nie, K., Zhu, Y., Shi, M., Xiao, C., Liu, H., Liu, Q., Zhao, T., Chen, X., Zhou, H., Wang, P., and Cheng, G. (2019) A gut commensal bacterium promotes mosquito permissiveness to arboviruses. *Cell Host Microbe* **25**, 101–112.e5

39. Myers, S. A., Peddada, S., Chatterjee, N., Friedrich, T., Tomoda, K., Krings, G., Thomas, S., Maynard, J., Broeker, M., Thomson, M., Pollard, K., Yamanaka, S., Burlingame, A. L., and Panning, B. (2016) SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells. *Elife* **5**, e10647

40. Svinkina, T., Gu, H., Silva, J. C., Mertins, P., Qiao, J., Fereshetian, S., Jaffe, J. D., Kuhn, E., Udeshi, N. D., and Carr, S. A. (2015) Deep, quantitative coverage of the lysine acetylome using novel anti-acetyl-lysine antibodies and an optimized proteomic workflow. *Mol. Cell. Proteomics* **14**, 2429–2440

41. Udeshi, N. D., Mertins, P., Svinkina, T., and Carr, S. A. (2013) Large-scale identification of ubiquitination sites by mass spectrometry. *Nat. Protoc.* **8**, 1950–1960

42. Labots, M., van der Mijn, J. C., Beekhof, R., Piersma, S. R., de Goeij-de Haas, R. R., Pham, T. V., Knol, J. C., Dekker, H., van Grieken, N. C. T., Verheul, H. M. W., and Jiménez, C. R. (2017) Phosphotyrosine-based-phosphoproteomics scaled-down to biopsy level for analysis of individual tumor biology and treatment selection. *J. Proteomics* **162**, 99–107

43. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. O., Guan, S., Medzihradszky, K. F., Maltby, D. A., Schoepfer, R., and Burlingame, A. L. (2006) O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5**, 923–934

44. Toghi Eshghi, S., Yang, W., Hu, Y., Shah, P., Sun, S., Li, X., and Zhang, H. (2016) Classification of tandem mass spectra for identification of N- and O-linked glycopeptides. *Sci. Rep.* **6**, 37189

45. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J., and Clausen, H. (2020) Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749

46. Halim, A., Westerlind, U., Pett, C., Schorlemer, M., Rüetschi, U., Brinkmalm, G., Sihlbom, C., Lengqvist, J., Larson, G., and Nilsson, J. (2014) Assignment of saccharide identities through analysis of oxonium ion fragmentation profiles in LC-MS/MS of glycopeptides. *J. Proteome Res.* **13**, 6024–6032

47. Wisnovsky, S., Möckl, L., Malaker, S. A., Pedram, K., Hess, G. T., Riley, N. M., Gray, M. A., Smith, B. A. H., Bassik, M. C., Moerner, W. E., and Bertozzi, C. R. (2021) Genome-wide CRISPR screens reveal a specific ligand for the glycan-binding immune checkpoint receptor Siglec-7. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2015024118

48. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190

49. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015) PhosphoSitePlus, 2014:

Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520

50. Toleman, C. A., Schumacher, M. A., Yu, S.-H., Zeng, W., Cox, N. J., Smith, T. J., Soderblom, E. J., Wands, A. M., Kohler, J. J., and Boyce, M. (2018) Structural basis of O-GlcNAc recognition by mammalian 14-3-3 proteins. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 5956–5961

51. Wu, H.-Y., Lu, C.-T., Kao, H.-J., Chen, Y.-J., Chen, Y.-J., and Lee, T.-Y. (2014) Characterization and identification of protein O-GlcNAcylation sites with substrate specificity. *BMC Bioinformatics* **15 Suppl 16**, S1

52. Lu, L., Riley, N. M., Shortreed, M. R., Bertozzi, C. R., and Smith, L. M. (2020) O-pair search with MetaMorpheus for O-glycopeptide characterization. *Nat. Methods* **17**, 1133–1138