UCSF
UC San Francisco Previously Published Works

Title

N- and O-Glycosylation in the Murine Synaptosome*

Permalink

https://escholarship.org/uc/item/02p3532v

Journal

Molecular & Cellular Proteomics, 12(12)

ISSN

1535-9476

Authors

Trinidad, Jonathan C
Schoepfer, Ralf
Burlingame, Alma L
et al.

Publication Date

2013-12-01

DOI

10.1074/mcp.m113.030007

Peer reviewed

N- and *O*-Glycosylation in the Murine Synaptosome*⃞ₛ

Jonathan C. Trinidad‡¶, Ralf Schoepfer§, Alma L. Burlingame‡, and Katalin F. Medzihradszky‡‖

We present the first large scale study characterizing both *N*- and *O*-linked glycosylation in a site-specific manner on hundreds of proteins. We demonstrate that a lectin-affinity fractionation step using wheat germ agglutinin enriches not only peptides carrying intracellular *O*-GlcNAc, but also those bearing ER/Golgi-derived *N*- and *O*-linked carbohydrate structures. Liquid chromatography-MS (LC/MS) analysis with high accuracy precursor mass measurements and high sensitivity ion trap electron-transfer dissociation (ETD) were utilized for structural characterization of glycopeptides. Our results reveal both the identity of the precise sites of glycosylation and information on the oligosaccharide structures possible on these proteins. We report a novel iterative approach that allowed us to interpret the ETD data set directly without making prior assumptions about the nature and distribution of oligosaccharides present in our glycopeptide mixture. Over 2500 unique *N*- and O-linked glycopeptides were identified on 453 proteins. The extent of microheterogeneity varied extensively, and up to 19 different oligosaccharides were attached at a given site. We describe the presence of the well-known mucin-type structures for *O*-glycosylation, an EGF-domain-specific fucosylation and a rare *O*-mannosylation on the transmembrane phosphatase Ptprz1. Finally, we identified three examples of *O*-glycosylation on tyrosine residues.   *Molecular & Cellular Proteomics 12: 10.1074/mcp.M113.030007, 3474–3488, 2013.*

Post-translational modifications (PTMs)[1] regulate a wide range of protein functions and cellular processes (1). Numerous large-scale studies have been conducted on intracellular, regulatory modifications, such as phosphorylation (2), GlcNAcylation (3), ubiquitination (4), and acetylation (5). Although ER/Golgi-derived glycosylation is one of the most frequently occurring classes of PTMs (6), these extracellular modifications continue to be a major challenge because of technical difficulties both in their structural characterization and in deciphering their biological roles. However, there is a renewed interest in the characterization of glycosylation driven by the production of protein pharmaceuticals as well as by the search for biomarkers and by new findings about the function(s) of glycosylation (7–12).

Many large-scale mass spectrometry studies on extracellular glycosylation have been conducted in recent years, mainly targeting *N*-glycosylation (13–17). *N*-glycosylation is more experimentally tractable because of the established consensus motif (NXS/T/C, where X cannot be Pro), common core structure and the existence of a nearly universal glycan-releasing enzyme (Peptide *N*-Glycosidase F, also known as PNGase F). A number of *N*-glycopeptide enrichment strategies exist, including: HILIC (13); ERLIC (14); TiO_2 (15); lectin-affinity-chromatography (16); or periodate-oxidation/hydrazide capture (17). Each of these approaches relies on PNGase F for the removal of the carbohydrate structures. Hydrolysis of the carbohydrate-linked asparagine residues converts them to aspartic acids. Deglycosylated peptides are analyzed by liquid chromatography-tandem MS (LC/MS/MS), and the site of modification is determined by the existence of an aspartic acid residue where the protein sequence calls for an asparagine. One potential shortcoming of this approach is that non-modified asparagine residues can undergo hydrolysis during sample preparation, a process that is especially promoted when they are amino-terminal to glycine residues (14, 18, 19). Importantly, the use of PNGase F in these approaches results in the loss of all information regarding the oligosaccharide(s) present on specific glycosylation sites.

Characterization of *O*-glycosylation represents a greater challenge, in part because this term covers multiple distinct modifications (20). A series of different sugars may link directly to Ser, Thr, and as recently reported, Tyr residues (21, 22). These core units may then be elongated with simple or more elaborate carbohydrate structures. Multiple different enzymes are responsible for depositing the core structures on the peptides. Thus, not surprisingly, no consensus motif has been established for *O*-glycosylation and there is no universal enzyme to remove the *O*-linked oligosaccharides from proteins. Chemical methods exist for cleavage of *O*-glycans from

[1] The abbreviations used are: PTM, post-translational modifications; CID, collision-induced dissociation; ETD, electron transfer dissociation; HCD, higher energy C-trap dissociation; WGA, wheat germ agglutinin; FPR, false positive rate.

protein attachment sites, however they have significant limitations (20). Some mucin-type *O*-glycopeptides can be enriched by lectin-affinity chromatography (23). Other enrichment strategies applied to *N*-linked glycopeptides, such as HILIC and ERLIC fractionation as well as TiO_2-based enrichment of sialylated structures can also be applied to *O*-linked glycopeptides (13–15, 24). Studies of native *O*-glycosylation remain challenging. In the case of the most extensive *O*-glycosylation study to date, an engineered cell line was used to prevent elongation of the core GalNAc moiety, significantly simplifying the analysis (25). No large-scale studies have been conducted on the less prevalent *O*-linked structures.

There have been a limited number of studies examining *N*- and/or *O*-linked glycopeptides isolated from complex biological samples. The most extensive list of glycosylation sites and carbohydrate structures was reported by Halim *et al.* (26). Their study focused on sialylated glycopeptides that were captured via limited periodate oxidation, followed by their release using mild acid hydrolysis. Thus, no information on neutral glycans or sialylation heterogeneity could be obtained.

The paucity of studies at the intact glycopeptide level may be due to the fact that collision-induced dissociation (CID) analysis of such molecules rarely yields sufficient structural information to answer all our questions. Full characterization of a glycopeptide would include: the sequence of the modified peptide; the site of modification; the identity of the glycan structures attached. Unfortunately, glycosidic bonds dissociate more readily than peptide bonds during the CID process. Thus, the resulting spectra are dominated by carbohydrate fragments that represent both the reducing and nonreducing ends (oxonium ions). In most cases, the degree of peptide backbone cleavages is limited (20, 27). In addition, interpretation of such spectra has to be carried out manually, because of the lack of software capable of deciphering the oligosaccharide structures from these spectra. Furthermore, no search engine can efficiently interpret glycopeptide spectra that contain extensive carbohydrate fragmentation. Finally, MS/MS search algorithms typically require that potential modifications be specified beforehand.

Software exists that can address some of these issues in the case of *N*-glycosylation (28, 29). Site assignment for *N*-linked glycopeptides is aided by the fact that these peptides typically will have a single asparagine residue that matches the NXS/T/C motif. In contrast, *O*-linked glycopeptides (and peptides in general) often contain several serine, threonine and/or tyrosine residues. Analysis of *O*-glycosylation is also more difficult because it does not occur on a strict amino acid consensus motif and can occur as multiple core structures. MS/MS spectra resulting from collisional activation are more likely to feature sufficient peptide backbone fragmentation to allow the peptide to be identified if the *O*-linked glycan structure is small (on the order of 1–3 sugar units). Unfortunately, the sugars are usually completely eliminated via gas-phase rearrangement in such a fashion that the site(s) of modification cannot be determined when there are multiple possibilities (20, 23). However, the mass spectrometric and proteomic search engines Protein Prospector (30) and ByOnic (31) can handle neutral losses and thus, identify *O*-glycopeptides. Both software tools are also able to assign the site(s) of modification(s) in those cases where fragment ions that retain the sugar are detected.

In contrast to the forgoing issues inherent in spectra generated by collisional activation processes, electron-transfer dissociation (ETD) possesses characteristics that reveal peptide sequence directly from glycopeptide components while displaying significantly reduced glycosidic bond cleavage (32, 33). However, to search these spectra in an automated fashion, it is still necessary to specify the mass value(s) of the potential oligosaccharide structure(s) present in advance.

In this work, we present results from the structural characterization of both *N*- and *O*-linked glycosylation occurring in murine synaptosomes. It includes a description of an iterative data analysis approach that enabled our determination of the most prevalent carbohydrate compositions, without having to specify *a priori* which structures were present in the sample. This approach is based entirely on analysis of ETD spectra of the glycopeptides, and required neither CID nor higher energy C-trap dissociation (HCD) analysis. It must be emphasized that this approach provides only the monoisotopic masses of the oligosaccharides, which modify the peptides. These masses can be translated into the most likely carbohydrate compositions. However, no information is obtained about the identity of the sugar units and their linkages. We present here the first large-scale study where both the sites of modifications were mapped and information about the site-specific glycan structures was obtained for 453 gene products.

EXPERIMENTAL PROCEDURES

Preparation of Mouse Synaptic Membranes—Sample preparation has been described in detail in (34). In summary, mouse synaptic membranes were purified at 4 °C in the presence of the O-GlcNAcase inhibitor PUGNAc (Toronto Research Chemicals, North York, ON, Canada) and a mixture of protease and phosphatase inhibitors. The membranous fraction obtained from several animals was layered on a sucrose density gradient and fractionated by centrifugation. The synaptosome fraction was collected at the 1.0–1.2 M sucrose interface and subsequently pelleted by centrifugation.

Tryptic Digestion of Synaptosome Samples—Thirty milligrams of synaptosome were resuspended in 1 ml buffer containing 50 mM ammonium bicarbonate, 6 M guanidine hydrochloride, 6× Phosphatase Inhibitor Cocktails I and II (Roche), and 20 μM PUGNAc (Tocris, Ellisville, MO). Disulfide-bridges were reduced with Tris(2-carboxyethyl) phosphine hydrochloride and then alkylated with iodoacetamide. Tryptic digestion was performed in the sixfold diluted mixture for 12 h at 37 °C, 2% (w/w) modified trypsin (Promega, Madison, WI) was added. The peptide digest was desalted and lyophilized to dryness.

Enrichment of Glycopeptides Using a Wheat Germ Agglutinin (WGA) Column—Peptides were resuspended in 50 μl buffer A (100 mM Tris *p*H 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 5% acetonitrile). Approximately 5 mg of sample were loaded onto a (250 × 2 mm) Poros-WGA column (34) at a flow rate of 125 μl/min.

GlcNAcylated peptides eluted as an unresolved smear at the end of the flow through peak. After ~10 min, 100 μl of 20 mM GlcNAc in buffer A was injected to elute any remaining peptides. A fraction enriched in glycopeptides was collected between ~10 and 30 min. For subsequent rounds of enrichment, the pooled fractions were run together in a similar fashion.

Enrichment of Phosphorylated Peptides Using Titanium Dioxide (Removal of Sialyl Glycopeptides)—Digests were resuspended in 250 μl buffer B1 (1% trifluoroacetic acid, 20% acetonitrile). Peptides were loaded onto a 62 μl titanium dioxide column at 80 μl/min in buffer B1. The column was rinsed with water, then eluted with 3 × 250 μl saturated KH_2PO_4 followed by 3 × 250 μl 5% phosphoric acid. Collected phosphopeptide mixtures were desalted and dried.

High pH Reverse Phase Chromatography—High *p*H reversed phase fractionation was performed using an ÄKTA Purifier (GE Healthcare, Piscataway, NJ) equipped with a 1 × 100 mm Gemini 3μ C18 column (Phenomenex, Torrance, CA). Buffer A was 20 mM HCOONH$_4$, *p*H 10. Buffer B was 20 mM HCOONH$_4$ in 50% acetonitrile. A gradient was applied increasing the solvent B concentration from 1% to 21% B in 14 min, then to 62% in 67 min, and eventually rinsing the column with 100% B. The flow rate was 80 μl/min. Twenty fractions were collected from the glycopeptide mixtures.

Mass Spectrometry—Experiments were performed on a linear trap-Orbitrap hybrid mass spectrometer (LTQ Velos, Thermo Fischer Scientific) coupled with a NanoACQUITY HPLC system (Waters, Milford, MA). The peptide mixtures were fractionated on a reverse-phase column with a 90 min gradient elution. Fractions desalted after LWAC-enrichment were subjected to LC/MS analysis with ETD activation. Each fraction was analyzed twice. In the first, only doubly charged ions were selected for MS/MS. In the second, only ions with a charge state of three or higher were selected for MS/MS. All other parameters were the same: mass measurements were performed in the Orbitrap, and a mass range of *m/z* 350–1400 was monitored. The 10 most abundant ions meeting the precursor selection criteria, *i.e.* displaying the permitted charge(s) and having a minimal intensity of 2000 counts, were subjected to ETD analysis. ETD activation and spectral acquisition were performed in the linear trap. The AGC settings were 10^4 and 2×10^5 for the peptide and fluoranthene ions, respectively. Reaction time was 100 msec for (2+) ions and for higher charge states it was automatically adjusted. The isolation window was set to three Th. Supplemental activation was enabled at 20% CE. Dynamic exclusion was enabled for 30 s.

Data Processing—Peaklists were generated with PAVA, an in-house software (35). Database searches were performed using Protein Prospector (version 5.9.0). Data was searched against *Mus musculus* sequences in Uniprot (downloaded 07/06/2011) along with a randomized version of each entry, for a total of 72,930 entries. Only tryptic peptides were considered, with one missed cleavage permitted. Carbamidomethylation of cysteine residues was set as a fixed modification. Variable modifications allowed were: protein *N*-terminal acetylation; *N*-terminal Gln cyclization; oxidation of Met. Modification of serine, threonine, tyrosine, or asparagine residues with an individual glycan was permitted for the ETD data. A total of three variable modifications per peptide were considered. The mass accuracy requirements were 10 ppm for precursor ions and 0.6 Da for fragment ions. When multiple MS/MS were matched to the exact same peptide, the best-matching spectrum was evaluated. The strength of modification site assignments was calculated in Protein Prospector using SLIP scoring (36). The overall false discovery rate was ~1% before manual confirmation of the glycopeptide spectra.

Acceptance Criteria for the Glycopeptides—

N-Glycosylation—Searches with oligosaccharide structures that may occur at Asn residues were merged. A list of oligosaccharide compositions considered can be found in supplemental Data S2. To achieve a false positive rate (FPR) of <1% for *N*-glycosylated peptides, the final filtering criteria were: minimum score = 22, max E = 0.01 and 0.05, for proteins and peptides, respectively. Site assignment reported is based on SLIP score as well as on the presence of consensus *N*-glycosylation motifs. If the same ETD spectra could be matched to multiple different glycoforms, all matches meeting the acceptance criteria were included.

O-Linked Glycosylation—Searches with oligosaccharide structures that may occur on Ser, Thr, and Tyr residues were merged. A list of oligosaccharide compositions considered can be found in supplemental Data S2. To achieve an FPR of < 1% for *O*-glycopeptides, the final filtering criteria were as follows: minimum score = 22, max E = 0.005, max mass error = 5 ppm. Site assignments listed in the Tables are based on SLIP scores (36) that sometimes were overwritten based on manual evaluation. Some assignments were deleted from the list after careful inspection, while a few glycopeptides that did not meet the original acceptance criteria were reinstated when the same site was confidently identified with a different sugar structure. If the same ETD spectra could be matched to multiple different glycoforms all matches meeting the acceptance criteria were included.

Domain-specific O-Glycosylation (Ser/Thr-Fucosylation) was Handled Separately—Search results permitting fucose or HexNAcFuc on any Ser and Thr residues were merged using a series of filters: minimum score = 22, max E = 0.001, and the mass error should not be more than 5 ppm. Of the 31 identifications that meet these criteria, only one survived the manual inspection (see "Results"). In addition, a peptide modified by HexHexNAcFuc was identified in the original unspecified modification search, and manually validated.

O-Mannosylation—Peptides modified with masses corresponding to *O*-mannosylation were found in the second iteration of the unspecified modification search, and the data were manually validated.

Glycopeptides meeting the acceptance criteria may represent multiple UniProt entries. Whenever there was a SwissProt entry among the equally well fitting proteins that was chosen to be included in our list. If a non-SwissProt (*i.e.* TrEMBL) entry featured more matching peptides, that was selected. Data representing all PTMs listed in supplemental Data S3 can be viewed using the viewer file available at prospector.ucsf.edu; search key = **tbguwfn09u**. The table displays the same columns as supplemental Data S3 , except the peptide column is not manually curated, and the last column with the references was not included. This viewer file was constructed permitting a 1Da mass error for fragment ions. The appropriate Table and peaklists were also uploaded to the journal's website as supplemental Data S8 and S9.

Characterizing the Distribution of Unknown Modifications—To discover what oligosaccharide structures could be present on the peptides enriched by LWAC, a database search was performed against 345 proteins that were identified as HexNAc-modified, or annotated as secreted or transmembrane based on Gene Ontology annotation (www.geneontology.org). Unknown modifications up to 3000 Da were permitted on Ser, Thr, and Asn residues. Peptides bearing a mass modification were used if their expectation value was less than or equal to 0.01. If a given peptide sequence and nominal mass combination was identified by more than one MS/MS spectrum, this was counted only as a single instance. The frequency of each nominal mass was then calculated. From this analysis, 22 masses corresponding to potential glycan structures were selected and each potential glycan was searched as an explicit modification (thus specifying the atomic formula and taking advantage of the high mass accuracy of the precursor ions).This unknown modification approach was repeated with the 423 proteins identified confidently in the first 22 database searches.
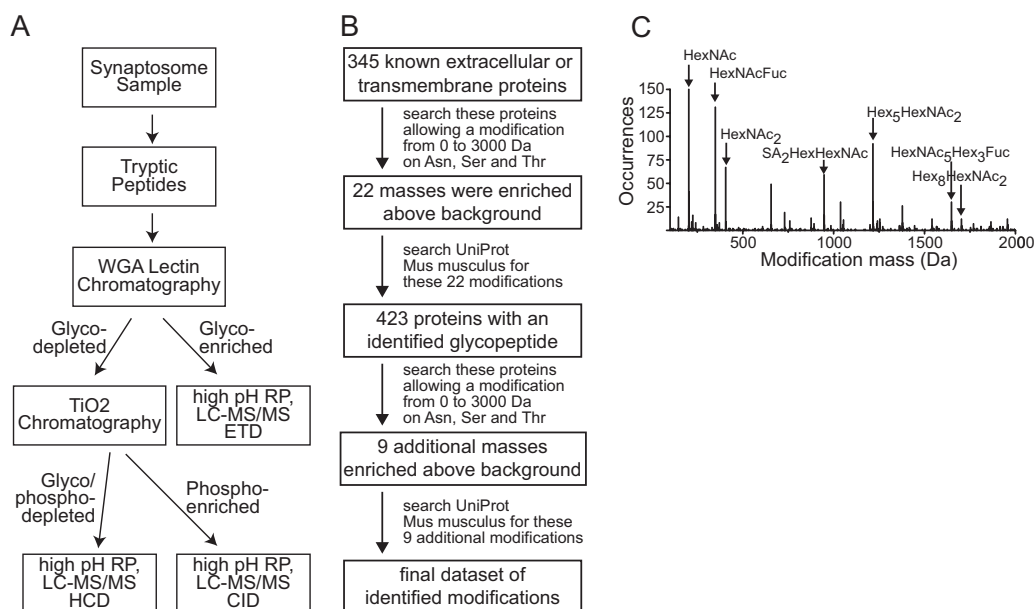
Fig. 1. *A, Schematic illustrating the serial analysis of glycosylation, phosphorylation, and protein content.* Synaptosomes were digested with trypsin. Glycopeptides were enriched using WGA affinity chromatography. Phosphopeptides were enriched using TiO_2 chromatography. In the first round of the analysis, the WGA enrichment was conducted first, followed by the TiO_2 chromatography. In the second round, the order of the WGA and TiO_2 enrichments were reversed. The final glyco/phospho depleted fraction as well as the individual PTM-enriched fractions were each fractionated using high *p*H reverse phase chromatography. *B*, Schematic illustrating the iterative approach to characterizing the glycans present in our sample. Because hundreds of potential glycans exist, it is computationally challenging to directly search for all of them in MS/MS data if one is searching an entire organism's proteome. *C*, The frequency of mass modifications identified on Asn, Ser, and Thr residues in the glycopeptide mixture studied. The *y* axis shows the number of unique peptide sequences that were found modified by a given mass. See detailed list in supplemental Data S1. Data was binned in one Dalton increments.

RESULTS

We have previously developed a workflow suitable for the concurrent characterization of phosphorylation and GlcNAcylation, key signaling and regulatory PTMs within the cell (Fig. 1*A*). Our WGA affinity chromatography was aimed primarily at the enrichment of intracellular *O*-GlcNAcylated peptides. However, we were aware that numerous membrane-bound proteins were present in the synaptosome (34, 37). Therefore, we wanted to investigate the extent to which our WGA enrichment approach was able to isolate peptides bearing carbohydrates other than just single *O*-GlcNAc.

This workflow was applied to studies of murine synaptosomes (34). The tryptic digest of the synaptosome proteins was first subjected to lectin-affinity chromatography using WGA. Next phosphopeptides were isolated from the glyco-peptide-depleted mixture using TiO_2. We also performed the enrichments in the reverse order, *i.e.* then the phosphopeptide-depleted mixture was subjected to lectin-affinity enrichment. The PTM-enriched and final flow-through fractions were further fractionated off-line by high pH reversed phase chromatography. Each of these fractions was analyzed by LC/MS/MS on an LTQ-Orbitrap Velos. The precursor ions were measured in the Orbitrap analyzer at high resolution with high mass accuracy, whereas the glycopeptides were subjected to ETD analysis.

Determining the Distribution of Detectible Glycan Structures—We developed an iterative approach to screen for glycans in our sample (Fig. 1*B*), which took advantage of the fact that we had previously characterized synaptosomes in great detail and identified 6621 proteins present in the preparation (34). For an initial screen, we selected a subset of these proteins (345 in total) that were likely transmembrane or secreted proteins based largely on Gene Ontology annotation (www. geneontology.org). We then used Protein Prospector to search these proteins, allowing for modification of serine, threonine and asparagine residues with mass values between 100 and 3000 Da. The results are shown as a frequency histogram in Fig. 1*C* and listed in supplemental Data S1. It is readily apparent that certain masses occur much more frequently than others. It was striking that practically all of the most frequently occurring modifications were masses that corresponded to potential oligosaccharide structures, thereby demonstrating the feasibility of this strategy. Based on sugar subunit composition differences between *N*- and *O*-linked glycosylation, it was clear that both classes of glycosylation were enriched during our WGA affinity chromatography step. Considering the most likely carbohydrate compositions for the mass differences detected, neither GlcNAc nor sialic acid was strictly required for glycopeptide enrichment using WGA (see below and supplemental Data S2).

We did not design this first pass filter to identify all the carbohydrates present in the sample. Rather, we initially wanted to identify the most common carbohydrate modifications. We then selected 22 modification mass values (203 Da included) that showed a clear increase in occurrence above background (listed in supplemental Data S2). We were able to match these masses to those of common sugar structures. Since the structure of the *N*-linked core structure is known: $GlcNAc_2Man_3$, (*i.e.* $HexNAc_2Hex_3$ with or without a Fuc residue) in general, the sugar composition directly implied whether the linkage was *N*- or *O*-, but in some cases (such as $HexNAc_2Hex_2$), the linkage type was ambiguous. Based upon the potential linkages, we then searched the Uniprot *Mus musculus* database, allowing these *N*-linked structures on asparagine residues and the *O*-linked structures on serine, threonine and tyrosine residues; for the ambiguous structures, *i.e.* when the sugar composition indicated truncated *N*-linked or *O*-linked oligosaccharides, all four residues were chosen as potential sites (supplemental Data S2). We identified a total of 423 unique gene products that appeared modified in this search. This subset represented a list of proteins in our synaptosome preparation that we could confidently conclude were glycosylated with carbohydrates larger than a single HexNAc. Using this high-confidence subset, we repeated the unknown modification search, again allowing modifications from mass 100 to 3000 Da on Asn, Ser, and Thr residues. This more extensive second search allowed us to identify an additional nine masses, the frequencies of which were clearly above background. These masses matched to potential *N*-linked carbohydrate structures (listed in supplemental Data S2). We then explicitly searched the entire Uniprot *Mus musculus* database allowing for these modifications as well. The search results for *N*-glycosylation, GalNAc-core *O*-glycosylation, and domain-specific *O*-fucosylation were individually evaluated. In each instance, a concatenated decoy database was used to evaluate the false discovery rate. A total of 453 proteins were found glycosylated (supplemental Data S3). 375 proteins were found to be *N*-glycosylated, while only 122 proteins were found to be *O*-glycosylated. Because their enrichment and detection efficiencies are not identical, these numbers may not reflect the relative distribution of these PTMs in the initial synaptosome material. The two enrichment strategies (WGA either before or after TiO_2) yielded a similar number of glycopeptide identifications, although there were ~10% more glycopeptides identified when the glycopeptide-enrichment was performed after TiO_2 chromatography (supplemental Data S3). Of note, the number of sialylated structures identified was ~50% lower when the TiO_2 chromatography was performed first.

Analysis of Peptide N-Glycosylation—Almost 2100 *N*-linked glycopeptides were identified, representing 678 glycosylation sites on 375 proteins (supplemental Data S3). We detected 463 peptides modified by single HexNAc moiety located on extracellular regions of each given protein (this excludes res-idues that are likely intracellular *O*-GlcNAc sites based on Uniprot-assigned transmembrane topology). Approximately 80% of these single HexNAcs were attached to asparagine residues, and therefore represent *N*-GlcNAc. The remaining 20% occurred on serine, threonine, and tyrosine residues, and we suggest that these represent *O*-GalNAc moieties, although some site assignments are ambiguous. The single GlcNAc-bearing *N*-linked glycopeptides represented 18% of all the characterized *N*-linked glycopeptides. Similar truncation products were also observed from the core fucosylated structures. Overall, these truncated core structures represent ~32% of the all *N*-linked glycans detected (a summary on carbohydrate distribution has been included into Supplement 2). The complete *N*-linked core, $GlcNAc_2Man_3$, with or without fucosylation represents ~6.5% of *N*-linked structures we identified. Among the extended glycans, oligomannose structures appear to be dominant (~39% of the structures detected). Approximately 41% of the detected structures feature a fucose unit, among these we detected a glycan mass of ~1362.5 Da on 35 peptides (23 unique sites), consistent with the presence of fucosylated oligomannose structures, $GlcNAc_2Man_5Fuc$. The most likely interpretation is that this corresponds to core-fucosylation. Interestingly, 13% of these structures were di-fucosylated. Since the mass difference between neuraminic acid and two fucose units is one Da, we originally suspected that some of these assignments resulted from mis-assignment of the glycopeptide mono-isotopic peak by the mass spectrometry software (*i.e.* faulty peak-picking). However, manual inspection of the raw spectra in these instances confirmed the original assignments. Misidentification was found in only one case and this glycopeptide was actually sialylated (in this case, $HexNAc_4Hex_4SA$ on Gria1, supplemental Data S3). All together only nine sialylated structures were identified in our study (supplemental Data S3).

It is important to note that our assignments are only based on the mass values of each glycan observed, so we do not have direct information regarding the identity of the sugar units or their linkage positions. In addition, the same carbohydrate structural assignments may be counted multiple times, if the peptide involved was detected in multiple forms, *i.e.* with missed cleavages, or with other possible PTMs.

N-Glycosylation on Nonconsensus Asparagine Residues—Our search parameters for *N*-glycosylation allowed this modification to occur on any asparagine residue, not just those matching the NXS/T/C sequon. If the initial search results indicated asparagine modification at a nonconsensus motif site, manual validation generally revealed that a nearby serine or threonine residue was in fact *O*-glycosylated (or the database spectral match was of low quality and rejected as a false positive). However, five examples of high quality spectral matches for noncanonical *N*-glycosylation were observed: Asn-222 on Gabrb3 (**N**VV); Asn-311 on Itgb8 (**N**NV), Asn-972 on Nrxn3 (**N**VV); Asn-652 on Lphn3 (**N**AG), and Asn-1018 on

Cntnap2 (**N**FQ). In each of these noncanonical examples, the particular asparagine residue was found modified by Man5 (although it should be noted that Man5 was the most common structure in the overall data set). In addition, a fucosylated Man5 structure, a Man9 oligosaccharide, and a complex fucosylated structure (supplemental Data S3) was observed on the integrin beta chain.

N-Glycosylation of Proteins Predicted to be Intracellular—Although the majority of the glycoproteins that we identified were annotated as secreted or transmembrane in the Gene Ontology, there were two *N*-glycoproteins that were annotated as cytosolic. The first was pre-B-cell leukemia transcription factor-interacting protein 1 (Q3TVI8, Asn-455), which bore a Man5 structure. The second was deleted in bladder cancer protein 1 homolog (Q920P3, Asn-156), which featured only a fucosylated GlcNAc. A related annotation conflict was observed for Solute carrier family 12 members 5 and 6. In these cases, we identified *N*-linked glycosylation on regions of transmembrane proteins that were annotated as intracellular (namely Asn-333, -351 and -362 on member 5 and Asn-398 on member 6). Only a single GlcNAc, which was sometimes fucosylated, was found at each site (supplemental Data S3).

Global Overview of Microheterogeneity—We observed many instances where individual protein residues displayed extensive glycan microheterogeneity, whereas at other sites only one or two glycoforms were found. To get an overall perspective on the distribution of microheterogeneity, we plotted out the number of glycoforms identified at each *N*-linked site in our data set (Fig. 2*A*). For almost half of the modified residues (314 cases), only a single glycan was identified. In 152 cases, either the single glycan was the core GlcNAc alone or as a fucosylated analog (∼65% of the time). The next three most common glycans were $HexNAc_2Hex_5$, $HexNAc_5Hex_3Fuc$, and $HexNAc_2Hex_3Fuc$, found as the only glycan identified on a given residue 30%, 4%, and 4% of the time, respectively. The relative proportion of $HexNAc_2Hex_9$ was higher among the "single" glycans than in the general population (5% *versus* 3% overall). Approximately 1/5th of the glycosylation sites for which we identified two glycoforms bore the core GlcNAc, either alone or fucosylated. Otherwise again the "usual suspects": $HexNAc_2Hex_3Fuc$, $HexNAc_2Hex_5$, and $HexNAc_5Hex_3Fuc$ were represented with higher numbers: 15, 59, and 23 instances, respectively.

Identification of Additional Glycoforms Using LC-MS Data—It has previously been observed that the individual glycoforms (of any particular glycopeptide bearing a range of carbohydrate structures) generally co-elute on standard RP C18 chromatography (29). Therefore, it is potentially useful to examine the distribution of precursor mass values in the corresponding MS survey scans to assess the extent of microheterogeneity for a given modified peptide sequence. Even for the case where all glycoforms of a glycopeptide were not selected for MS/MS analysis, it is possible to estimate their distribution.

Our approach for in-depth characterization relies on off-line high *p*H fractionation. Frequently, glycoforms of a given glycopeptide were detected in multiple fractions. As a result, their relative intensities within an LC-MS run varied as a function of high *p*H fraction. Thus, we did not try to deduce the relative amounts of different glycoforms present. However, we investigated whether any additional ones could be identified based on similar charge distribution and retention time as of the glycoforms confidently assigned from ETD data.

To describe how assessment of the site heterogeneity was tackled, Asn-1107 of Neural cell adhesion molecule L1 was chosen to illustrate the additional information that can be manually extracted from our data set as well as some of the difficulties of data-mining (Fig. 2*B*). Seven glycoforms were identified by MS/MS for this particular site, ranging from a fucosylated GlcNAc to oligomannose, complex and hybrid structures. All the "neutral" structures, except the much smaller FucGlcNAc-derivative were identified from a single high *p*H fraction. The acidic (Neu5Ac-containing) glycoform eluted earlier during the off-line fractionation, and therefore this glycoform was not detected in the same LC-MS run as the neutral oligosaccharide structures. In addition, we observed a series of glycoforms in both high *p*H fractions, for which we did not have confident MS/MS-based identification (Figs. 2*C*–2*F*). Four additional acidic glycoforms were noted in the early high *p*H fraction, whereas nine additional neutral glycoforms were observed in the later eluting high *p*H fraction. Selected ion chromatograms for each glycoform illustrate that related structures have similar, but not identical retention times. Although acidic glycoforms eluted earlier in the high *p*H reversed phase chromatography, when formic acid is used as the ion-pairing reagent in the subsequent low *p*H LC-MS analysis, they elute later than neutral glycoforms (27). Our investigation revealed that ETD spectra were recorded for each of these additional glycoforms, however they were not identified either because the MS/MS spectra did not meet the acceptance criteria outlined in the Methods, or the modifying oligosaccharide was not included in our targeted database searches. In the cases where the ETD spectra were of insufficient quality for automated database interpretation, manual inspection clearly established that the peptides in question were related to the glycopeptide family with which they co-eluted (supplemental Data S4). Similar findings were obtained for Asn-59 of Myelin-oligodendrocyte glycoprotein: seven additional monosialo structures were identified (supplemental Data S6).

Analysis of Peptide O-Glycosylation—We identified 463 *O*-glycopeptides, corresponding to modifications on 122 proteins. For these peptides, 190 sites of modification were determined unambiguously, whereas 152 are listed as ambiguous (some sites were identified with multiple peptides due to missed cleavage or oxidized versions). (supplemental Data S3). In 33 instances, we identified peptides that were simul-
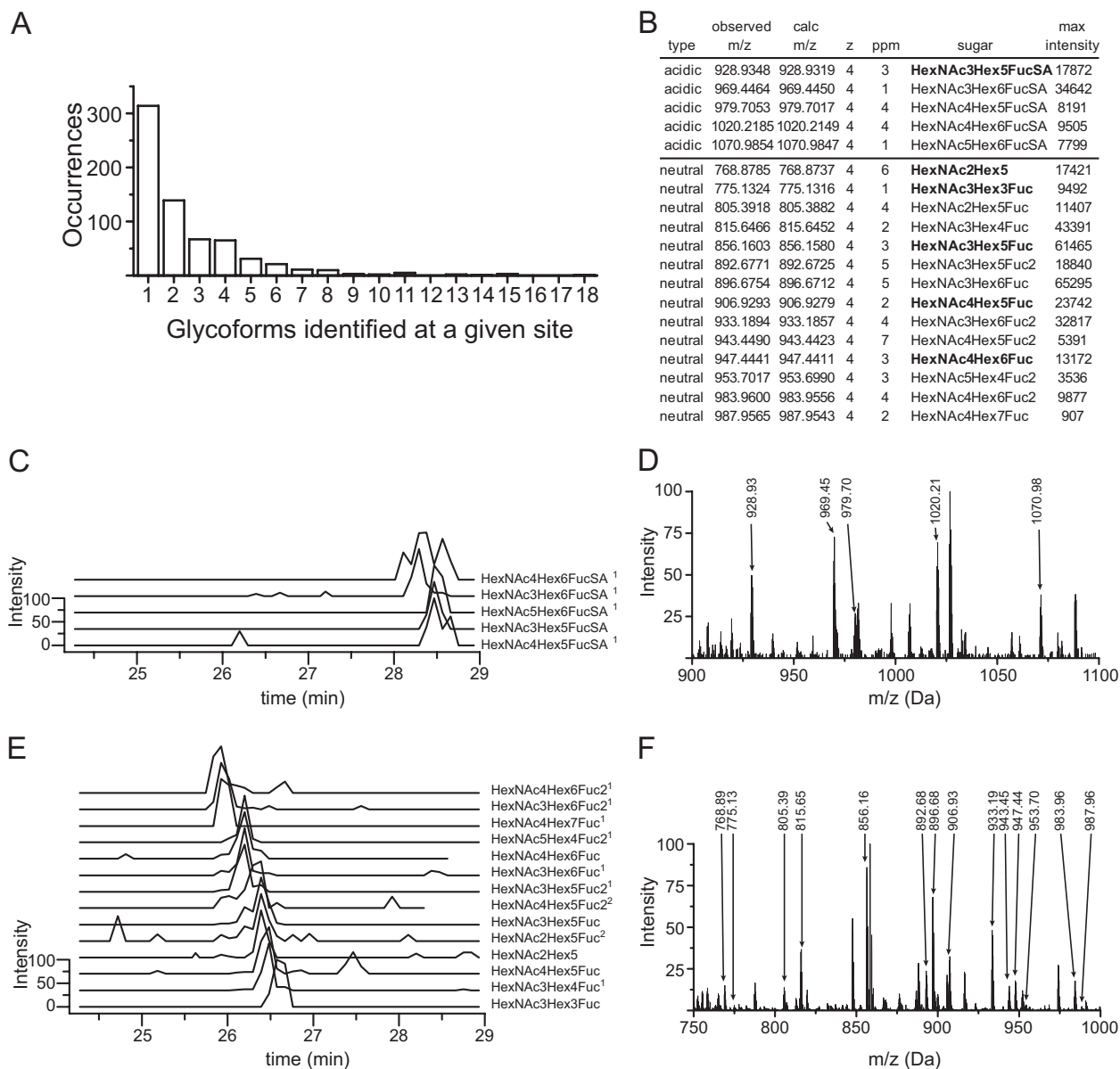
A

B

type	observed m/z	calc m/z	z	ppm	sugar	max intensity
acidic	928.9348	928.9319	4	3	**HexNAc3Hex5FucSA**	17872
acidic	969.4464	969.4450	4	1	HexNAc3Hex6FucSA	34642
acidic	979.7053	979.7017	4	4	HexNAc4Hex5FucSA	8191
acidic	1020.2185	1020.2149	4	4	HexNAc4Hex6FucSA	9505
acidic	1070.9854	1070.9847	4	1	HexNAc5Hex6FucSA	7799
neutral	768.8785	768.8737	4	6	**HexNAc2Hex5**	17421
neutral	775.1324	775.1316	4	1	**HexNAc3Hex3Fuc**	9492
neutral	805.3918	805.3882	4	4	HexNAc2Hex5Fuc	11407
neutral	815.6466	815.6452	4	2	HexNAc3Hex4Fuc	43391
neutral	856.1603	856.1580	4	3	**HexNAc3Hex5Fuc**	61465
neutral	892.6771	892.6725	4	5	HexNAc3Hex5Fuc2	18840
neutral	896.6754	896.6712	4	5	HexNAc3Hex6Fuc	65295
neutral	906.9293	906.9279	4	2	**HexNAc4Hex5Fuc**	23742
neutral	933.1894	933.1857	4	4	HexNAc3Hex6Fuc2	32817
neutral	943.4490	943.4423	4	7	HexNAc4Hex5Fuc2	5391
neutral	947.4441	947.4411	4	3	**HexNAc4Hex6Fuc**	13172
neutral	953.7017	953.6990	4	3	HexNAc5Hex4Fuc2	3536
neutral	983.9600	983.9556	4	4	HexNAc4Hex6Fuc2	9877
neutral	987.9565	987.9543	4	2	HexNAc4Hex7Fuc	907

C

D

E

F

FIG. 2. *A*, **For each site of *N*-linked glycosylation identified, the number of unique glycans identified was determined.** This was plotted as a histogram showing the number of sites for which a given number of glycans were identified. The most common occurrence was that only a single glycan was identified at a given modification site. *B*, Glycoforms of the peptide VLLHHLDVKTN*GTGPVR from Neural cell adhesion molecule L1 (P11627). Glycopeptides bearing acidic glycans eluted in fraction 14 of the high *p*H RP gradient, whereas those bearing neutral glycans eluted later (in fraction 16). Structures in bold were identified in the nonspecified modification search (Fig. 1*C*) and explicitly included in the subsequent database searches. *C* and *E*, Overlaid extracted ion chromatograms of the acidic (*C*) and neutral (*E*) glycopeptides in B. Intensities were extracted with a mass window of ± 0.2 Da. Peptides denoted with a superscript "1" indicate oligosaccharide compositions not included in targeted database searches. Peptides denoted with a superscript "2" indicate glycopeptides for which MS/MS were acquired, but the spectra matches did not meet acceptance criteria. Peptides were ordered by increasing retention time within a given high *p*H fraction. *D*, Summed MS survey scan of the acidic elution window from 27.77 to 28.14 min in high *p*H fraction 14. Mass labels correspond to peptides listed in Fig 2*B*. *F*, Summed MS survey scan of the neutral glycopeptide elution window from 25.83 to 26.50 min in high *p*H fraction 16. Mass labels correspond to peptides listed in Fig 2*B*.

taneously glycosylated at two residues. When an individual glycosylation site was modified by a single HexNAc as well as modified with elongated structures, parsimony would argue that the single sugar unit was GalNAc. On the other occasions, we relied on predicted extracellular regions to decide the assignment of the single sugar unit as GalNAc. These results confirm that our WGA affinity chromatography enriched for both GlcNAc and GalNAc.

In a similar fashion to *N*-linked peptides, ~35% of the *O*-linked glycans were HexNAc or HexNAc$_2$ structures, sug-
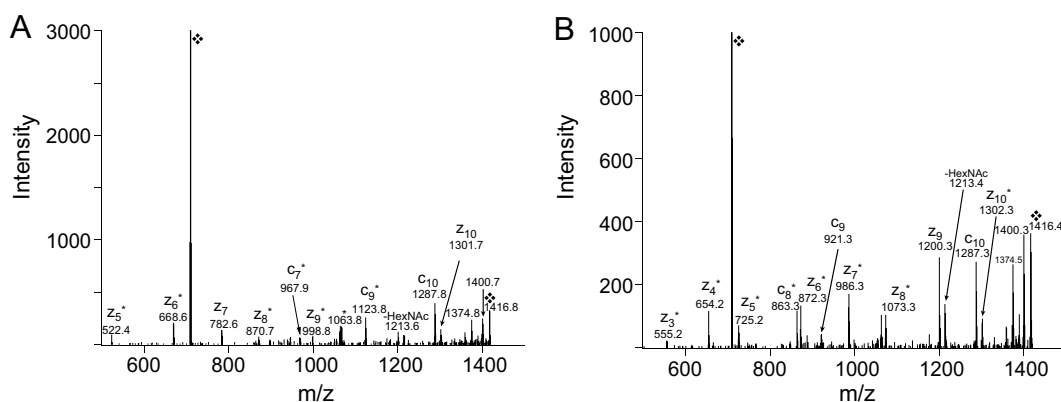
Fig. 3. ***A*, ETD spectrum of [177]VT(HexNAc)QSNFAVGYK[187] of Voltage-dependent anion-selective channel protein 1.** (Precursor ion: m/z 708.8543 (2+); within 1 ppm of the calculated mass). The diamond symbols indicate the precursor or charge-reduced ions; z+1 and c-1· fragments are labeled with asterisks. Base peak intensity was ~6000. ***B*, ETD spectrum of [177]VTQSNFAVGY(HexNAc)K[187] of Voltage-dependent anion-selective channel protein 1.** (Precursor ion: *m/z* 708.8541 (2+); within 1 ppm of the calculated mass). The diamond symbols indicate the precursor or charge-reduced ions; z+1 and c-1· fragments are labeled with asterisks. Base peak intensity was ~2000.

gesting that these sites were glycosidically processed (either endogenously or during sample preparation). Roughly half of the truncated structures represent HexNAc-dimers (supplemental Data S2 and S3). The most common core structures from which these could be derived are mucin-type core-2 or core-3 oligosaccharides (GlcNAcGalNAc). Interestingly, the majority of the "complete" structures were sialylated, ~55% disialo-GalGalNAc (assuming that indeed the most frequently occurring mammalian structure, mucin-type core 1, was enriched), and 41% Neu5AcGalGalNAc. Asialo GalGalNAc was identified only in four instances, whereas asialo tetrasaccharide was identified seven times, this structure could be a core-3 structure modified with a Gal on the GlcNAc, as has been described for bovine fetuin (38). As mentioned previously, the major caveat is the lack of direct evidence for the identity of the sugar units and their linkages.

Interestingly, some of the *O*-glycosylation sites identified were located close to or within the *N*-glycosylation sequon. For example, on Low-density lipoprotein receptor-related protein 11, glycosylation occurs at Ser-386 and/or Thr-387 (whereas Asn-403 was found unmodified in a potential *N*-glycosylation motif). Other examples include oligodendrocyte-myelin glycoprotein (Ser-391 and Asn-389); matrix metalloproteinase-15 (Thr-355 and Asn-364); Poliovirus receptor-related protein 1 (Thr-334 or 2 other potential sites and Asn-332) and fractalkine (Thr-111 and Asn-109). In cell adhesion molecule 2, Thr-213 is *O*-glycosylated with a HexNAc unit, while Asn-211 has been detected bearing a series of different glycans. We cannot exclude the possibility, that *O*- and *N*-glycosylation may occur concurrently in these proteins. To address this issue, we also conducted database searches considering all *N*- and *O*-linked structures as variable modifications, and permitting two per peptide on confirmed glycoproteins. However, as discussed below, we found that "double" glycosylation usually featured an identical mass to a single larger oligosaccharide. In addition, using these search parameters, *i.e.* considering the entire

"glycan set" and permitting the combination of any two of them modifying the same peptide, resulted in an unacceptably high rate of false positive identifications. (Results not shown).

Evidence for O-Glycosylation on Tyrosine Residues—Three tyrosine residues were identified unambiguously as *O*-glycosylation sites (supplemental Data S3): Tyr-418 of ATP synthase subunit beta, in peptide IMDPNIVGNEHYDVAR; Tyr-96 of aspartate aminotransferase, in peptide NLDKEYLPIGGLAEFCK; and Tyr-186 of voltage-dependent anion-selective channel protein, in peptide VTQSNFAVGYK. These proteins share subcellular localization, all localize to the mitochondria, and in each instance, the tyrosine was found modified by a single HexNAc unit. Two of the proteins, Atp5b, and Vdac1 featured an additional site of HexNAc-modification, however, no elongated structures were found. We do not have information to establish whether the modification is GlcNAc or GalNAc. Figs. 3*A*, 3*B* show the glycosylation of the same Vdac1 peptide, but at different residues. We aligned three tyrosine glycosylation sequences identified here with the two previously reported instances (21, 22). Using this limited set of sequences, no clear consensus motif was evident (supplemental Data S7).

O-Fucosylation—When fucosylation (+146 Da) was allowed as a variable modification, in 28 instances we observed potential fucosylation with very high scoring MS/MS spectra (supplemental Data S5). These sequences all contained Cys-residues, and examination of the actual MS/MS spectra, with one exception, revealed incomplete carbamidomethylation of a HexNAc-modified peptide (rather than fucosylation). Unfortunately, the mass addition as well as the elemental composition for these two options are the same: $-C_2H_3NO + C_8H_{13}NO_5 = C_6H_{10}O_4$.

However, the assignment of the fucose residue at Thr-3103 of Versican core protein (Q62059) was correct (supplemental Data S5). The peptide ([3100]Asn-Arg[3112]) was detected with

three different glycans: Fuc, HexNAc-Fuc, and Hex-HexNAc-Fuc (supplemental Data S3). The ETD spectrum of the trisaccharide-bearing glycopeptide contains multiple ions that result from carbohydrate-losses (Fig. 4). Similar glycan fragmentation was detected in numerous spectra of our dataset (For example, Figs. 5 and 6, and supplemental Data S4).
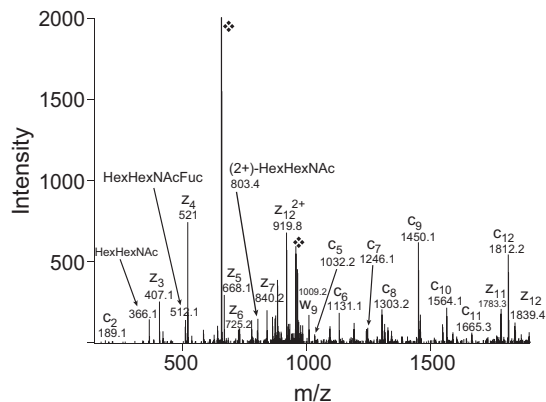

FIG. 4. **ETD spectrum of [3100]NGAT(GalGlcNAcFuc)C(Carbamidomethyl)VDGFNTFR[3112] of Versican core protein (Q62059).** The precursor ion was *m/z* 657.2828 (3+), within 1 ppm from the calculated value. Base peak intensity was ~4000. The carbohydrate structure was assigned from references reporting EGF-domain-specific glycosylation (60). The diamond symbol indicates the precursor ion.
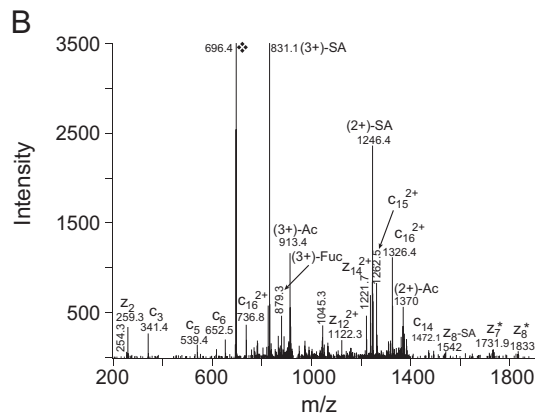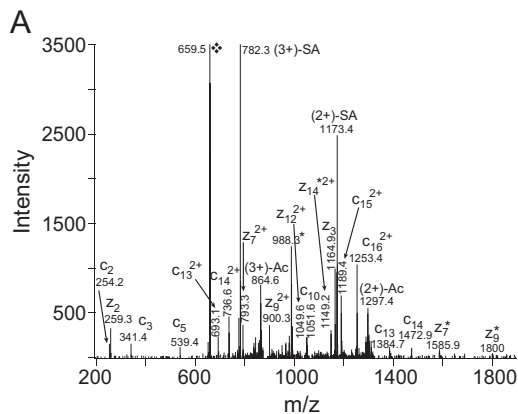
O-Mannosylation—In addition to the structures discussed above, we also detected mannosylation on protein tyrosine phosphatase, receptor type Z, polypeptide 1 (Ptprz1). The mass modifications 818 Da and 964 Da were detected on Ser-1273 with good scores and relatively high confidence (supplemental Data S3, Figs. 5*A*, 5*B*). These mass increments correspond to $HexNAcHex_2Neu5Ac$ and $HexNAcHex_2Neu5AcFuc$ compositions, respectively. Such compositions do not correspond to any mucin-type *O*-linked structures, but may represent *O*-mannosylation, where the smaller structure is SA-Gal-GlcNAc-Man, and the presence of additional fucose has been reported (39). Interestingly, we also detected the new mannosylation site, Ser-1506, with the more common mucin-type glycosylation (supplemental Data S3, Fig. 6).

DISCUSSION

Enrichment Strategy—Initially WGA was reported as having specificity for oligosaccharide structures containing either beta-linked GlcNAc or neuraminic acid (40). Because of this affinity toward GlcNAc-modified residues, it has been used to characterize intracellular protein *O*-GlcNAcylation (37, 41). However, a recent study by Chang *et al.* indicated that WGA has a significantly broader sugar residue binding specificity, including toward oligosaccharides containing either terminal

FIG. 5. **ETD spectra of VHSTPLQGLTVPHSS*KK of Protein tyrosine phosphatase, receptor type Z, polypeptide 1, featuring oligosaccharides of 818 Da (*A*) and 964 Da (*B*) that were tentatively assigned based as having carbohydrate compositions of Neu5Ac-Gal-GlcNAc-Man and (Fuc)Neu5Ac-Gal-GlcNAc-Man, respectively (39).** The diamond symbol indicates the precursor ions; z+1 and c-1 fragment ions are labeled with asterisks.
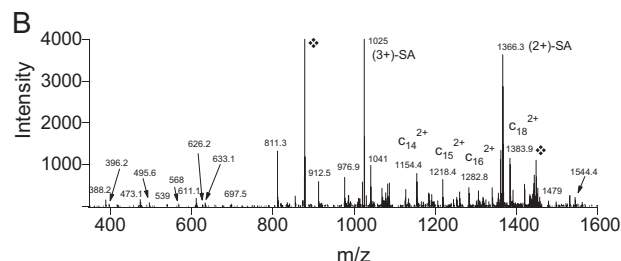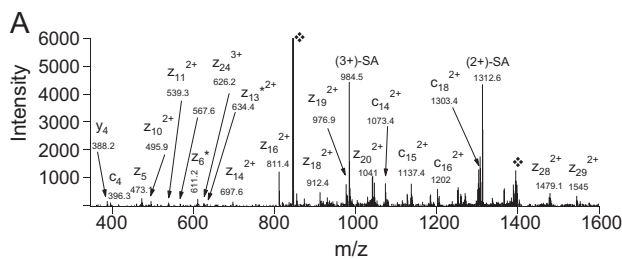

FIG. 6. **ETD spectra of SPPPS*MLPQNHNDGKEDSDIQMGSAVLPHTPGSK of Protein tyrosine phosphatase, receptor type Z, polypeptide 1, featuring Neu5AcGalGalNAc oligosaccharide (*A*) and Neu5AcGalGlcNAcMan (*B*) panel.** The diamond symbol indicates the precursor ions or charge-reduced ions.

Gal and GalNAc moieties (42). Thus, it was important to carry out a comprehensive analysis of all glycopeptides present in an enrichment using this lectin.

Indeed, our present results establish the diversity of glyco-peptide structures obtained using WGA-based affinity chromatography. We show that it is effective in the enrichment of both *N*-linked and *O*-linked glycopeptides. While we have noted in our previous reports that *N*-linked glycosylation was present in WGA enrichments (37, 41), it is clear at this juncture that this affinity enrichment strategy does not chromatographically resolve (intracellular) *O*-GlcNAcylated peptides from (extracellular) *O*-GalNAcylated sequences. Therefore, in this study we have assigned the identity of single HexNAc residues based primarily on the cellular localization of the protein and/or domain identified as annotated in Gene Ontology.

Our original sample separation protocol was aimed at the enrichment of both phospho- and glycopeptides from proteolytic digestions (34), and glycopeptide isolation has been carried out from both the unaltered and the phosphopeptide-depleted tryptic digests. So these experiments provided two major glycopeptide-containing fractions for study; namely, a fraction enriched with all glycopeptides present in the total digest and a fraction having the sialo-glycopeptides depleted by TiO_2 treatment of the digest (15). The further fractionated gly-copeptide mixtures were analyzed by the most efficient method for GlcNAcylated peptide site assignment, ETD MS (37).

Interpretation of ETD Data of Glycopeptides—Because ETD does not induce side-chain fragmentation usually, structural information about the attached glycan(s) is not obtained. At the same time, traditional database searching requires prior knowledge of potential glycan structures. These structural possibilities can be obtained by consultation of multiple databases (*e.g.* www.functionalglycomics.org and http://www.glycome-db.org). In addition, of direct relevance to our present consideration of the synaptosome, carbohydrate pools isolated from brain lysate and synaptosomal fractions have been reported from rat as well as chicken (43–46). Because we wanted to conduct an unbiased search of actual synaptosome constituents, we took advantage of the ability of the search engine Protein Prospector to allow for modifications of an arbitrary mass range. These searches enabled us to determine which glycans were most frequently detected in our sample. We then performed a set of glycan-specific searches against the entire *Mus musculus* proteome using this subset. As discussed above (Fig. 2), we also obtained evidence for the presence of additional glycans that were not considered as potential modifications in our final search. Although we could explicitly search for these additional modifications, the initial variable mass modification search results (Fig. 1*C*) indicate that they likely yielded good quality ETD data only for a very small subset of peptides.

Our findings suggest that given the limitations of the present methodology, an alternative approach could be more successful in the identification of glycoforms of higher mass.

We suggest it may be more efficient to utilize additional information such as chromatographic co-elution and related MS/MS fragmentation patterns of the (lower mass) glycopeptides already identified in a given study (see further discussion below) when a high mass glycoform of a given peptide is present in the sample together with additional (lower mass) glycoforms.

Database searching of glycopeptide spectra is complicated by the fact that because of the isomeric nature of particular sugar units, the same mass values may correspond to a series of different structures (not to speak of the linkage variations). In addition, the combined masses of two or more smaller structures frequently correspond to a single larger carbohydrate, which may also be present in the mixture. Because our database search approach only considered a single oligosaccharide structure in each iteration, we had to handle such cases only with respect to *O*-glycosylation analysis, in which usually multiple potential glycan linkage sites occur within a given modified peptide. Thus, sometimes we had to decide whether a peptide bore single or double modification (three variable modifications were permitted per peptide). If such a decision could not be made convincingly, *i.e.* two different assignments met the acceptance criteria for the very same precursor ion, we have chosen to list both among our search results.

A similar issue exists when a peptide that has a Trp or Met residue is *N*-glycosylated by a fucose-containing glycan. If the Trp or Met residue becomes oxidized (and is in close proximity to the glycosylation site), it becomes difficult to distinguish between this situation and the case where the Trp/Met is not oxidized, but there is a hexose present instead of the fucose (supplemental Data S3). Of course, carbohydrate fragmentation using CID or HCD can be used to confirm the presence of the fucose, and may help in such instances.

The above example also highlights an additional problem that arises from the combination of incomplete peptide backbone fragmentation, the proximity of potential modifications sites and the "repetitive" mass values of sugar units. For example, database searches with the combination of *N*- and *O*-linked structures are prone to yield ambiguous results, because the majority of *N*-glycosylation motifs contain Ser/Thr residues, which are potential *O*-glycosylation sites. Therefore, if a peptide is modified by a large *N*-linked structure, the search engine may have a difficult time differentiating this from the case where there is a smaller *N*-linked sugar with the corresponding "missing" sugar units linked to the neighboring Ser/Thr residue.

In general, characterization of complex glycosylation patterns remains extremely challenging (47). CID analysis primarily provides oligosaccharide composition information, with some degree of information on glycan linkages. ETD analysis provides site assignment information, and usually indicates the presence of sialic acids in the form of neutral losses (See

Figs. 5 and 6, and supplemental Data S4). MS[3] analysis in combination with both ETD and HCD, as well as specific exo- or endoglycosidase treatments (48) may be necessary to decipher glycopeptide structures.

Reliability of Assignments—In general, having a large number of potential different modifications to deal with presents a unique challenge to currently available automated algorithms aimed at interpretation of glycopeptide mass spectra. In our experience, the results of "automated interpretation" still need to be manually evaluated and verified, even if very stringent acceptance thresholds are applied to the search results. For example, the case of fucosylation *versus* HexNAc modification combined with incomplete carbamidomethylation of Cys residues presented above illustrates beautifully one of the difficulties to be addressed during the automated assignment of covalent modifications.

The spectral interpretation and reliable assignment of diverse modifications encountered in large-scale PTM studies present several specific, difficult challenges. They include establishing the correct acceptance criteria for the spectral assignments in an automated fashion, while minimizing the FPR. Presently this is usually accomplished by applying a universal acceptance threshold (for example, universal minimal score or maximum E values) to unmodified and modified sequences alike. However, this approach is unacceptable and has to be changed. Modified sequences have to be handled separately. Even if just a single variable modification is considered, the number of potential sequences increases in a nonlinear fashion. The situation becomes worse when an individual modification is allowed to occur on multiple residues, and when multiple modifications per peptide are permitted. Thus, acceptance criteria have to be tailored to the specific variable modifications considered. This phenomenon has been discussed previously (49). However, its implications have not been adopted by the community. In our study, we had to use different acceptance thresholds for *N*- and *O*-glycosylation in order to maximize the number of likely true positives while keeping each FPR rate for unique structural assignments below 1%. Even so, some possible multiply modified glycopeptides that met the acceptance criteria were removed from our list when manual inspection revealed that they were likely miss-assignments. We believe the automated evaluation of multiply modified peptides, especially when a series of different variable modifications are permitted is a problem that has not been fully solved. This represents a serious hurdle, which limits our ability to decipher the full extent of protein post-translation modifications.

N-Glycosylation—Our findings with respect to neutral oligosaccharides show good agreement with previous studies on the rat brain *N*-linked carbohydrate pool (43, 44). For example, those reports described the occurrence of a relatively high percentage of oligomannose structures, found evidence for truncated and bisecting complex carbohydrate structures, and noted the high occurrence of the Lewis[X] determinant. The

presence of this epitope on the major *N*-glycan also has been reported in the analysis of synaptic plasma membrane proteins of chickens (46). In the rat brain studies, the relative proportions found among the oligomannose structures were approximately Man5:Man6:Man7:Man8:Man9 = 4:2:1:1:2. Our results show a higher relative prevalence of Man5 structures with proportions of $\sim$6:2:1:1:1. However, our numbers may reflect the higher efficiency of the ETD fragmentation process expected for such smaller structures, rather than the actual underlying glycoform distribution. Chen *et al.* (43) have also reported that three other glycans: $HexNAc_4Hex_3$ Fuc, $HexNAc_5Hex_3Fuc$, and $HexNAc_5Hex_4Fuc_2$ (present approximately at the same level) were the most prevalent in addition to the oligomannose structures. Interestingly, we detected significantly more of the larger structures from these three glycans (14, 110, and 56 unique sites modified, respectively), which may indicate that their distribution differs between whole brain and synaptosomes and/or between rat and mouse. Such differences may result from dendrite specific glycoprotein synthesis (50) or synaptic localization via oligomannose-binding proteins such as NCAM or Synapsin I (51).

We also detected fucosylated Man5 structures. It has been a common assumption that oligomannose structures cannot be fucosylated, however it has been reported recently that the structure we observed indeed does occur (52). As far as we are aware, our findings are the first confirmation of the presence of such structures in the brain.

We detected *N*-glycosylation occurring on nonconsensus asparagine residues. The sites on Gabrb3, Lphn3, and Cntnap2 were previously identified after PNGase F treatment, which is largely orthogonal to our present approach with respect to errors in site assignment (53). The site on Gabrb3 features the reverse glycosylation motif SR**N**VV, with a serine residue in the P2 rather than P2′ position. A reverse glycosylation motif has been shown to direct glycosylation in both recombinant and endogenously expressed proteins (54).

We identified a high degree of microheterogeneity at many of the sites. At this stage, it is difficult to determine whether this finding reflects the endogenous distribution of glycosylation, or products of incomplete degradation by glycosidase activity during sample preparation. However, the fact that we identified the plurality of glycosylation sites with only a single glycan argues that glycosidase activity was not a major factor in our preparation. While a wide range of glycosidase inhibitors have been characterized (55), the large number of potential glycosidases would complicate any attempts to completely inhibit their activity. To the best of our knowledge, none of the glycosylation studies we have referenced herein employed glycosidase inhibitors during the isolation of the proteins or glycans studied.

We identified an unexpectedly high number of *N*-linked single GlcNAc units. The presence of such *N*-linked glycopeptides has been reported in *O*-GlcNAc studies following different enrichment strategies (40, 56, 57). Single GlcNAc

molecules *N*-linked to asparagine are not cleaved efficiently by PNGase F, which would explain why they have not been detected in previous glycan-level analyses (58).

We did not conclusively identify any glycopeptides bearing carbohydrate structures greater than 1955 Da (for example, disialo biantennary complex structures, their monoisotopic mass addition would be 2204.77 Da). In contrast, a study on the sialylated *N*-glycans of rat brain indicated extensive sialylation and the presence of large complex structures (44). It is clear that our methodology is currently undersampling these type of glycopeptides. We believe such glycopeptides may not yield high quality ETD spectra (59), making their characterization difficult. At the same time, our investigation on co-eluting glycoforms establishes that peptides modified with larger sugar structures are in fact present. However, we could not confidently identify these glycopeptides based upon ETD analyses alone, because the spectra did not meet the acceptance criteria employed. That being stated, the interpretation of such ETD spectra in connection with confidently identified co-eluting glycoforms may in the future allow for improved characterization, just as the additional glycoforms presented for Asn-59 of Myelin-oligodendrocyte glycoprotein above (supplemental Data S6). Five of these manually identified glycoforms have been described in brain by Zamze *et al.*, (44) and we also observed mass values consistent with two larger hybrid structures. Based on these observations, we believe our raw MS data still contains a significant amount of additional potentially interpretable information. This situation highlights the need for continued software development that can generate "glycopeptide-families" based on confidently identified ETD spectra, using MS/MS data, accurate mass differences, as well as retention time information.

O-Glycosylation—Not surprisingly, the majority of *O*-linked carbohydrates detected represented mucin-type structures. Significant site heterogeneity was detected for certain proteins. During the process of identification of additional *N*-linked oligosaccharides we could rely on the co-elution patterns of different glycoforms. However, this approach is inadequate for assignment of *O*-linked glycopeptides found in such a complex mixture. One of the reasons is that *O*-glycosylation frequently occurs in "clusters," thus, the co-eluting glycopeptide may represent a different modification site. In addition, *O*-glycans are typically of smaller size than *N*-glycans. Therefore, the addition of a few sugar units represents more significant changes in hydrophilicity, which can result in larger retention time differences across glycoforms of a given peptide (relative to those observed for *N*-linked glycoforms). In summary, we had to rely on ETD based identifications, though we are convinced that more *O*-linked glycoforms are "hiding" among the data, waiting for the appropriate mining tool.

We also detected two rarer forms of *O*-glycosylation: EGF-domain-specific *O*-fucosylation and more interestingly, *O*-mannosylation of a new protein.

O-fucosylation elongated with a Neu5Ac-Gal-GlcNAc-structure has been reported for EGF domains, with a consensus sequence of CXXGG**T/S**C (60). The glycosylation of the EGF-like domain that we detected in Versican core protein occurred on sequence ³⁰⁹⁸CRNGA**T**C³¹⁰⁴, where there is an Ala residue instead of a Gly in the P1 position of the consensus sequence.

Mammalian *O*-mannosylation, *i.e.* the SA-Gal-GlcNAc-Man oligosaccharide was first described from peripheral nerves (61), and the prevalence of *O*-linked mannosyl structures in brain has also been established (62). The augmented structure, bearing an additional fucose, also has been reported, but its linkage position has not been determined (39).

Although this type of glycosylation is most frequently discussed in connection with alpha-dystroglycan, other proteins such as CD24, neurofascin, and receptor tyrosine phosphatase beta have also been shown to contain *O*-mannosyl structures (63–65). However, the exact site of such modifications has not been identified for these proteins. Thus, our study is the first that has established the actual *O*-mannosylation sites on a previously uncharacterized protein. In addition, this is the first instance when the occurrence of mucin-type and Man-core structures has been shown to occur at the same site.

Tyrosine glycosylation with GalNAc-core structures has been reported previously on two secreted proteins: nucleobindin-2 and amyloid precursor protein (21, 22). We also identified three tyrosine-*O*-glycosylated mitochondrial proteins. However, the UniProt database for each protein indicates the potential localization in the plasma membrane. Thus, it is not clear whether this finding represents novel GlcNAcylation of tyrosine residues or additional examples of Tyr-GalNAc-modified proteins (21, 22).

Correlations Derived From Gene Ontology Database—When elongated glycan structures were identified for specific *O*-linked sites, each was located on domains predicted to be extracellular. However, in the few instances listed in the following, our *N*-glycosylation results contradict localization predictions for intracellular proteins pre-B-cell leukemia transcription factor-interacting protein 1, Deleted in bladder cancer protein 1 homolog (Q3TVI8 & Q920P3) and intracellular domains of solute carrier family 12 members 5 and 6. This observation strongly suggests that the first two proteins are not exclusively intracellular (if at all), while the domain localization for the solute carrier proteins must be different. All sites mentioned above have also been identified in PNGaseF-treated samples (53). The site-specific location of newly detected PTMs is a valuable tool to provide support for models of transmembrane topology, and on rare occasion will trigger a re-evaluation of the prevalent model for an individual protein (66).

Final Conclusions—In summary, our study reveals the complex challenge remaining in deriving a comprehensive grasp on the site specific glycosylation even in a relatively "simple" system, such as a model biological pseudo-organelle,

the synaptosome. It highlights the serious lack of effective affinity or chromatographic methodologies available for sample fractionation compared with the current level of sophistication of ETD and HCD methods on Orbitrap based hardware platforms.

It also proves that our iterative data interpretation can be successful for the analysis of mass spectral data of glycopeptides without prior knowledge or assumptions regarding the distribution of glycans. Although our approach provides only the sugar-unit composition of the glycans modifying the proteins, without revealing the identity of those units and their linkages; important conclusions can be made using this knowledge alone or combined with detailed structural information obtained from other studies. This novel approach permitted the assignment of both *N*- and *O*-glycosylation including some rare modifications, such as *O*-mannosylation, EGF-domain specific *O*-fucosylation, and tyrosine *O*-glycosylation. We envision this type of approach will prove important as large-scale ETD-based glycopeptide analysis becomes applied to a wide range of systems and organisms. We demonstrated that the lectin, wheat germ agglutinin, can be used to efficiently enrich glycopeptides bearing a broad range of glycan structures. These studies of intact glycopeptides have enabled us to look globally at site-specific microheterogeneity. The number of distinct glycan structures at a given site was found to vary extensively. Lastly, this investigation represents the first large-scale analysis of *O*- and *N*-glycosylation from *in vivo* samples where the sites of modification as well as glycan structures have been established, to the level of confident sugar unit composition assignment of the glycans attached to this complex mixture of glycopeptides. We have shown that these techniques are able to profile site-specific *N*- and *O*-linked glycosylation, and thus provide detailed insight into the global role of extracellular glycosylation at the synapse.

[S] This article contains supplemental Data S1 to S9, and Legends for the Supplementary Material.

¶ Present address: Department of Chemistry, Indiana University, Simon Hall 120B; 212 S. Hawthorne Drive, Bloomington, IN 47405, trinidad@indiana.edu.

‖ To whom correspondence should be addressed: Department of Pharmaceutical Chemistry, 600 16th Street, Genentech Hall, suite N472A, Box 2240, University of California, San Francisco, San Francisco, CA 94158-2517. Tel.: 415-476-5160; Fax: 415-502-1655; E-mail: folkl@cgl.ucsf.edu.

REFERENCES

1. Walsh, C. (2006) *Posttranslational Modification Of Proteins: Expanding Nature's Inventory*, Roberts and Company Publishers, Greenwood Village, Colorado
2. St-Denis, N., and Gingras, A. C. (2012) Mass spectrometric tools for systematic analysis of protein phosphorylation. *Prog. Mol. Biol. Transl. Sci.* **106**, 3–32
3. Zachara, N. E. (2009) Detecting the "O-GlcNAc-ome"; detection, purification, and analysis of O-GlcNAc modified proteins. *Methods Mol. Biol.* **534**, 251–279
4. Bustos, D., Bakalarski, C. E., Yang, Y., Peng, J., and Kirkpatrick, D. S. (2012) Characterizing Ubiquitination Sites by Peptide-based Immunoaffinity Enrichment. *Mol. Cell Proteomics* **11**, 1529–1540
5. Zee, B. M., and Garcia, B. A. (2012) Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem.* **52**, 147–163
6. Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462
7. Li, H., and d'Anjou, M. (2009) Pharmacological significance of glycosylation in therapeutic proteins. *Curr. Opin. Biotechnol.* **20**, 678–684
8. Arnold, J. N., Saldova, R., Hamid, U. M., and Rudd, P. M. (2008) Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics* **8**, 3284–3293
9. An, H. J., Kronewitter, S. R., de Leoz, M. L., and Lebrilla, C. B. (2009) Glycomics and disease markers. *Curr. Opin. Chem. Biol.* **13**, 601–607
10. Parker, R. B., and Kohler, J. J. (2010) Regulation of intracellular signaling by extracellular glycan remodeling. *ACS Chem. Biol.* **5**, 35–46
11. Freire-de-Lima, L., Gelfenbeyn, K., Ding, Y., Mandel, U., Clausen, H., Handa, K., and Hakomori, S. I. (2011) Involvement of O-glycosylation defining oncofetal fibronectin in epithelial-mesenchymal transition process. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17690–17695
12. Schjoldager, K. T., Vester-Christensen, M. B., Bennett, E. P., Levery, S. B., Schwientek, T., Yin, W., Blixt, O., and Clausen, H. (2010) O-glycosylation modulates proprotein convertase activation of angiopoietin-like protein 3: possible role of polypeptide GalNAc-transferase-2 in regulation of concentrations of plasma lipids. *J. Biol. Chem.* **285**, 36293–36303
13. Picariello, G., Ferranti, P., Mamone, G., Roepstorff, P., and Addeo, F. (2008) Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **8**, 3833–3847
14. Hao, P., Ren, Y., Alpert, A. J., and Sze, S. K. (2011) Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Mol. Cell Proteomics* **10**, O111.009381
15. Palmisano, G., Lendal, S. E., Engholm-Keller, K., Leth-Larsen, R., Parker, B. L., and Larsen, M. R. (2010) Selective enrichment of sialic acid-containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. *Nat. Protoc.* **5**, 1974–1982
16. Drake, P. M., Schilling, B., Niles, R. K., Braten, M., Johansen, E., Liu, H., Lerch, M., Sorensen, D. J., Li, B., Allen, S., Hall, S. C., Witkowska, H. E., Regnier, F. E., Gibson, B. W., and Fisher, S. J. (2011) A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. *Anal. Biochem.* **408**, 71–85
17. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21**, 660–666
18. Robinson, N. E., and Robinson, A. B. (2001) Molecular clocks. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 944–949
19. Palmisano, G., Melo-Braga, M. N., Engholm-Keller, K., Parker, B. L., and Larsen, M. R. (2012) Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses. *J. Proteome Res.* **11**, 1949–1957
20. Peter-Katalinić, J. (2005) Methods in enzymology: O-glycosylation of proteins. *Methods Enzymol.* **405**, 139–171
21. Steentoft, C., Vakhrushev, S. Y., Vester-Christensen, M. B., Schjoldager, K. T., Kong, Y., Bennett, E. P., Mandel, U., Wandall, H., Levery, S. B., and Clausen, H. (2011) Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nat. Methods* **8**, 977–982
22. Halim, A., Brinkmalm, G., Rüetschi, U., Westman-Brinkmalm, A., Portelius, E., Zetterberg, H., Blennow, K., Larson, G., and Nilsson, J. (2011) Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11848–11853
23. Darula, Z., and Medzihradszky, K. F. (2009) Affinity enrichment and characterization of mucin core-1 type glycopeptides from bovine serum. *Mol*

Cell Proteomics **8,** 2515–2526

24. Darula, Z., Sherman, J., and Medzihradszky, K. F. (2012) How to dig deeper? Improved enrichment methods for mucin core-1 type glycopeptides. *Mol. Cell Proteomics* **11,** O111.016774

25. Vakhrushev, S. Y., Steentoft, C., Vester-Christensen, M. B., Bennett, E. P., Clausen, H., and Levery, S. B. (2013) Enhanced mass spectrometric mapping of the human GalNAc-type O-glycoproteome with SimpleCells. *Mol. Cell Proteomics* **12,** 932–944

26. Halim, A., Nilsson, J., Rüetschi, U., Hesse, C., and Larson, G. (2012) Human urinary glycoproteomics; attachment site specific analysis of N- and O-linked glycosylations by CID and ECD. *Mol Cell Proteomics* **11,** M111.013649

27. Medzihradszky, K. F. (2005) Characterization of protein N-glycosylation. *Methods Enzymol.* **405,** 116–138

28. Ozohanics, O., Krenyacz, J., Ludányi, K., Pollreisz, F., Vekey, K., and Drahos, L. (2008) GlycoMiner: a new software tool to elucidate glycopeptide composition. *Rapid Commun. Mass Spectrom* **22,** 3245–3254

29. Goldberg, D., Bern, M., Parry, S., Sutton-Smith, M., Panico, M., Morris, H. R., and Dell, A. (2007) Automated N-glycopeptide identification using a combination of single- and tandem-MS. *J. Proteome Res.* **6,** 3995–4005

30. Darula, Z., Chalkley, R. J., Baker, P., Burlingame, A. L., and Medzihradszky, K. F. (2010) Mass spectrometric analysis, automated identification and complete annotation of O-linked glycopeptides. *Eur. J. Mass Spectrom (Chichester, Eng)* **16,** 421–428

31. Bern, M., Cai, Y., and Goldberg, D. (2007) Lookup peaks: a hybrid of de novo sequencing and database search for protein identification by tandem mass spectrometry. *Anal. Chem.* **79,** 1393–1400

32. Wuhrer, M., Stam, J. C., van de Geijn, F. E., Koeleman, C. A., Verrips, C. T., Dolhain, R. J., Hokke, C. H., and Deelder, A. M. (2007) Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* **7,** 4070–4081

33. Perdivara, I., Petrovich, R., Allinquant, B., Deterding, L. J., Tomer, K. B., and Przybylski, M. (2009) Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation. *J. Proteome Res.* **8,** 631–642

34. 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

35. Guan, S., Price, J. C., Prusiner, S. B., Ghaemmaghami, S., and Burlingame, A. L. (2011) A data processing pipeline for mammalian proteome dynamics studies using stable isotope metabolic labeling. *Mol. Cell Proteomics* **10,** M111.010728

36. Baker, P. R., Trinidad, J. C., and Chalkley, R. J. (2011) Modification site localization scoring integrated into a search engine. *Mol. Cell Proteomics* **10,** M111.008078

37. 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

38. Edge, A. S., and Spiro, R. G. (1987) Presence of an O-glycosidically linked hexasaccharide in fetuin. *J. Biol. Chem.* **262,** 16135–16141

39. Gomez Toledo, A., Raducu, M., Cruces, J., Nilsson, J., Halim, A., Larson, G., Rüetschi, U., and Grahn, A. (2012) O-Mannose and O-N-acetyl galactosamine glycosylation of mammalian alpha-dystroglycan is conserved in a region-specific manner. *Glycobiology* **22,** 1413–1423

40. Gallagher, J. T., Morris, A., and Dexter, T. M. (1985) Identification of two binding sites for wheat-germ agglutinin on polylactosamine-type oligosaccharides. *Biochem. J.* **231,** 115–122

41. 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

42. Chang, C. F., Pan, J. F., Lin, C. N., Wu, I. L., Wong, C. H., and Lin, C. H. (2011) Rapid characterization of sugar-binding specificity by in-solution proximity binding with photosensitizers. *Glycobiology* **21,** 895–902

43. Chen, Y. J., Wing, D. R., Guile, G. R., Dwek, R. A., Harvey, D. J., and Zamze, S. (1998) Neutral N-glycans in adult rat brain tissue–complete charac-

terisation reveals fucosylated hybrid and complex structures. *Eur. J. Biochem.* **251,** 691–703

44. Zamze, S., Harvey, D. J., Chen, Y. J., Guile, G. R., Dwek, R. A., and Wing, D. R. (1998) Sialylated N-glycans in adult rat brain tissue–a widespread distribution of disialylated antennae in complex and hybrid structures. *Eur. J. Biochem.* **258,** 243–270

45. Clark, R. A., Gurd, J. W., Bissoon, N., Tricaud, N., Molnar, E., Zamze, S. E., Dwek, R. A., McIlhinney, R. A., and Wing, D. R. (1998) Identification of lectin-purified neural glycoproteins, GPs 180, 116, and 110, with NMDA and AMPA receptor subunits: conservation of glycosylation at the synapse. *J. Neurochem.* **70,** 2594–2605

46. Koles, K., McDowell, W., Mileusnic, R., and Rose, S. P. (2005) Glycan analysis of the chicken synaptic plasma membrane glycoproteins–a major synaptic N-glycan carries the LewisX determinant. *Int. J. Biol. Sci.* **1,** 126–134

47. Christiansen, M. N., Kolarich, D., Nevalainen, H., Packer, N. H., and Jensen, P. H. (2010) Challenges of determining O-glycopeptide heterogeneity: a fungal glucanase model system. *Anal. Chem.* **82,** 3500–3509

48. Stimson, E., Hope, J., Chong, A., and Burlingame, A. L. (1999) Site-specific characterization of the N-linked glycans of murine prion protein by high-performance liquid chromatography/electrospray mass spectrometry and exoglycosidase digestions. *Biochemistry* **38,** 4885–4895

49. Baker, P. R., Medzihradszky, K. F., and Chalkley, R. J. (2010) Improving software performance for peptide electron transfer dissociation data analysis by implementation of charge state- and sequence-dependent scoring. *Mol. Cell Proteomics* **9,** 1795–1803

50. Villanueva, S., and Steward, O. (2001) Glycoprotein synthesis at the synapse: fractionation of polypeptides synthesized within isolated dendritic fragments by concanavalin A affinity chromatography. *Brain Res. Mol. Brain Res.* **91,** 137–147

51. Wang, S., Cesca, F., Loers, G., Schweizer, M., Buck, F., Benfenati, F., Schachner, M., and Kleene, R. (2011) Synapsin I is an oligomannose-carrying glycoprotein, acts as an oligomannose-binding lectin, and promotes neurite outgrowth and neuronal survival when released via glia-derived exosomes. *J. Neurosci.* **31,** 7275–7290

52. Crispin, M., Harvey, D. J., Chang, V. T., Yu, C., Aricescu, A. R., Jones, E. Y., Davis, S. J., Dwek, R. A., and Rudd, P. M. (2006) Inhibition of hybrid- and complex-type glycosylation reveals the presence of the GlcNAc transferase I-independent fucosylation pathway. *Glycobiology* **16,** 748–756

53. Zielinska, D. F., Gnad, F., Wiśniewski, J. R., and Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141,** 897–907

54. Valliere-Douglass, J. F., Eakin, C. M., Wallace, A., Ketchem, R. R., Wang, W., Treuheit, M. J., and Balland, A. (2010) Glutamine-linked and non-consensus asparagine-linked oligosaccharides present in human recombinant antibodies define novel protein glycosylation motifs. *J. Biol. Chem.* **285,** 16012–16022

55. Asano, N. (2003) Glycosidase inhibitors: update and perspectives on practical use. *Glycobiology* **13,** 93R-104R

56. Klement, E., Lipinszki, Z., Kupihár, Z., Udvardy, A., and Medzihradszky, K. F. (2010) Enrichment of O-GlcNAc modified proteins by the periodate oxidation-hydrazide resin capture approach. *J. Proteome Res.* **9,** 2200–2206

57. Alfaro, J. F., Gong, C. X., Monroe, M. E., Aldrich, J. T., Clauss, T. R., Purvine, S. O., Wang, Z., Camp, D. G., 2nd, Shabanowitz, J., Stanley, P., Hart, G. W., Hunt, D. F., Yang, F., and Smith, R. D. (2012) Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 7280–7285

58. Chu, F. K. (1986) Requirements of cleavage of high mannose oligosaccharides in glycoproteins by peptide N-glycosidase F. *J. Biol. Chem.* **261,** 172–177

59. Alley, W. R., Jr., Mechref, Y., and Novotny, M. V. (2009) Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data. *Rapid Commun. Mass Spectrom* **23,** 161–170

60. Harris, R. J., and Spellman, M. W. (1993) O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology* **3,** 219–224

61. Chiba, A., Matsumura, K., Yamada, H., Inazu, T., Shimizu, T., Kusunoki, S., Kanazawa, I., Kobata, A., and Endo, T. (1997) Structures of sialylated

O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J. Biol. Chem.* **272,** 2156–2162

62. Chai, W., Yuen, C. T., Kogelberg, H., Carruthers, R. A., Margolis, R. U., Feizi, T., and Lawson, A. M. (1999) High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur. J. Biochem.* **263,** 879–888

63. Bleckmann, C., Geyer, H., Lieberoth, A., Splittstoesser, F., Liu, Y., Feizi, T., Schachner, M., Kleene, R., Reinhold, V., and Geyer, R. (2009) O-glycosylation pattern of CD24 from mouse brain. *Biol. Chem.* **390,** 627–645

64. Pacharra, S., Hanisch, F. G., and Breloy, I. (2012) Neurofascin 186 is O-mannosylated within and outside of the mucin domain. *J. Proteome Res.* **11,** 3955–3964

65. Abbott, K. L., Matthews, R. T., and Pierce, M. (2008) Receptor tyrosine phosphatase beta (RPTPbeta) activity and signaling are attenuated by glycosylation and subsequent cell surface galectin-1 binding. *J. Biol. Chem.* **283,** 33026–33035

66. Thalhammer, A., Trinidad, J. C., Burlingame, A. L., and Schoepfer, R. (2009) Densin-180: revised membrane topology, domain structure and phosphorylation status. *J. Neurochem.* **109,** 297–302