Enrichment and Site-Mapping of O-Linked N-Acetylglucosamine by a Combination of Chemical/Enzymatic Tagging, Photochemical Cleavage, and Electron Transfer Dissociation (ETD) Mass Spectrometry

Zihao Wang¹, Namrata D. Udeshi²*, Meaghan O’Malley¹*, Jeffrey Shabanowitz², Donald F. Hunt²,³ and Gerald W. Hart¹$""'

¹Johns Hopkins School of Medicine, Department of Biological Chemistry, Baltimore, MD; University of Virginia, Departments of ²Chemistry and ³Pathology, Charlottesville, VA 22904

*These authors contributed equally.

$To whom correspondence should be addressed.

Email: gwhart@jhmi.edu

Running Title: Novel Approach for O-GlcNAc Enrichment and Site-Mapping
Abbreviations

O-GlcNAc: O-linked N-acetylglicosamine
ETD: Electron transfer dissociation
CAD: collision activated dissociation
GalT1: mutant galactosyltransferase
SCX: strong cation exchange
OMSSA: Open Mass Spectrometry Search Algorithm
UDP-GalNAz: UDP-N-azidoacetylgalactosamine
PC: photocleavable
MeCP2: methyl-CpG-binding protein
SUMMARY

Numerous cellular processes are regulated by the reversible addition of either phosphate or N-acetylglucosamine (O-GlcNAc) to nuclear and cytoplasmic proteins. While sensitive methods exist for the enrichment and identification of protein phosphorylation sites, those for the enrichment of O-GlcNAc containing peptides are lacking. Reported here is highly efficient methodology for enrichment and characterization of O-GlcNAc sites from complex samples. In this method, O-GlcNAc modified peptides are tagged with a novel biotinylation reagent, enriched by affinity chromatography, released from the solid support by photochemical cleavage, and analyzed by electron transfer dissociation (ETD) mass spectrometry. Using this strategy, eight O-GlcNAc sites were mapped from a tau-enriched sample from rat brain. Sites of GlcNAcylation were characterized on important neuronal proteins such as tau, synucleins, and the methyl CpG binding protein 2.

INTRODUCTION
Numerous cytoplasmic and nuclear proteins are post-translationally modified with O-linked β-N-acetylglucosamine (O-GlcNAc). GlcNAcylation is involved in almost all aspects of cellular metabolism (1) and is highly dependent on the nutrient status of the cell (2). The O-GlcNAc modification rivals phosphorylation in both abundance and protein distribution. Recent studies indicate that signaling pathways can be regulated by the interplay of these two modifications at the same or proximal sites on numerous protein substrates (3).

Current understanding of the functions of O-GlcNAc and of the function of O-GlcNAcylation, and its relationship to phosphorylation are severely hampered by the difficulties in detecting this labile monosaccharide modification. Problems associated with the identification of O-GlcNAc sites include: (a) O-GlcNAc is quickly removed by hydrolases during cell lysis, (b) like phosphorylation, O-GlcNAc is usually present in less than stoichiometric amounts at given sites on protein substrates, (c) O-GlcNAc is readily lost as an oxonium ion during conventional peptide sequence analysis by collision activated dissociation (CAD) (Supplementary Fig. 1 online), (d) modified and unmodified forms of the peptide often co-elute during reverse phase HPLC (Supplementary Fig. 2a online), and the preferential ionization of the unmodified peptide suppresses the signal observed for the corresponding O-GlcNAc modified peptide (Supplementary Fig. 2b-c online).

Several attempts have been made to enrich samples for O-GlcNAc modified proteins and peptides. Immunoaffinity purification of O-GlcNAc modified peptides with an antibody (CTD 110.6) has been largely unsuccessful because of low binding avidity (4). Long, wheat germ agglutinin (WGA) lectin columns (~39 ft) provide some enrichment, but also bind strongly to complex glycans (5). A mutant galactosyltransferase (GalT1) has been used to label GlcNAcylated proteins with a ketone-containing galactose analog (6). Following proteolytic digestion, O-GlcNAc modified peptides were biotinylated with hydrazine chemistry, isolated on a column packed with avidin beads, eluted with free biotin, and sequenced by ETD mass spectrometry. Failure to elute peptides with high efficiency from the
avidin column and an inability to direct the fragmentation to the peptide backbone, limits the usefulness of this approach. Reported here is an enrichment methodology that: (a) is highly specific for O-GlcNAc modified peptides, (b) provides for efficient release of the captured peptides from an affinity support, and (c) facilitates complete characterization of the released peptides by ETD mass spectrometry.

**EXPERIMENTAL PROCEDURES**

**Reagents and Chemical Synthesis.** All chemicals were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise noted. PC-PEG-Biotin-Alkyne reagent (Reagent 1 in Fig. 1a) containing a photocleavable 1,2-(nitrophenyl) ethyl moiety (PC-PEG-Biotin, synthesized according to Olejnik et al. (9) or obtained from Ambergen, Watertown, MA) was prepared by treating the corresponding N-hydroxysuccinimidyl ester with 10-fold excess of propargylamine (Sigma, St. Louis, MO) in dry methanol at room temperature for 4 hrs in the dark. The total reaction volume was 20 µl containing 5 mg PC-PEG-Biotin (6 umol) and 4 µl propargylamine (60 umol). Product (PC-PEG-Biotin-Alkyne) was purified by thin layer chromatography (250 microns silica gel plate, Analtech, Newark, DE) with methanol/chloroform (1:9, v/v) as the mobile phase. The product is located on the thin layer plate by brief exposure (< 1 second) to 254 nm U.V. light and extracted by scraping the thin-layer zone into dry methanol. Silica gel is removed by centrifugation. Identity of the product is confirmed by mass spectrometry (Supplement Figure 4). The yield is greater than 98%. The purified product was stored in methanol at -20 °C until use.

**Enzymatic and Chemical Derivatization of O-GlcNAc Modified Peptides; Enrichment, and Photochemical Release of Tagged Peptides.** Proteins were digested with trypsin (50:1 substrate/enzyme) in 100 mM ammonium bicarbonate (pH 8) overnight at 37°C. Trypsin was removed with a 10 kDa molecular weight cut-off membrane (Millipore, Billerica, MA). Solvent was removed under vacuum and the residue was then re-suspended in 120 µL of 10 mM HEPES (pH 7.9) containing 5
mM MnCl₂, UDP-N-azidoacetylgalactosamine (UDP-GalNAz) and 10 units of GalT1. The reaction mixture was incubated overnight at 4°C, treated with 10 units calf intestine phosphatase (CIP) (New England Biolabs, Ipswich, MA), incubated for an additional 2 hr at room temperature, and then passed through a C18 spin column (Nestgroup, Southborough, MA). UDP-GalNAz and derivatized peptides were eluted in 0.1% TFA/1% acetonitrile and 0.1% TFA/80% acetonitrile, respectively. Solvent was removed under vacuum and the sample was reconstituted in 20 ul of methanol containing 0.05 µmol PC-PEG-Biotin-Alkyne, 10 mM sodium ascorbate, 1 mM Tris [(1-benzyl-1H-1,2,3-triazol-4-yl) methyl] amine (TBTA, in 4:1 t-butanol:DMSO), and 2 mM CuSO₄. The reaction mixture was incubated overnight at room temperature with gentle agitation. To remove excess PC-PEG-Biotin-Alkyne, the sample was diluted into strong cation exchange (SCX) loading buffer (5 mM KH₂PO₄, 25% acetonitrile, pH 3.0), and then passed through a SCX spin column (Nestgroup, Southborough, MA). The column was washed with several column volumes of loading buffer, and the retained peptides were eluted with high salt buffer (5 mM KH₂PO₄, 400 mM KCl, 25% acetonitrile, pH 3.0). Eluant was adjusted to pH 7 with ammonium hydroxide and then allowed to interact with high capacity avidin beads (Pierce, Rockford, IL) for 2 hr at room temperature. Avidin beads were washed 10 times with PBS solution, twice with 20% methanol/water, and resuspended in 70% methanol/ water. The suspension was transferred to a thin-walled PCR tube, irradiated with 365 nm UV light (2mW/cm²) (Spectroline ENF-240C, Westbury, NY) for 25 min at a distance of 10 cm with rotation, and the supernatant was then dried under vacuum, and stored at -20° C. The overall process is also demonstrated in Fig. 1a

**Western blotting with HRP-conjugated avidin.** α-Crystallin (Invitrogen, Carlsbad, CA) was subjected to chemoenzymatic tagging as described above but without SCX and avidin enrichment. The protein was resolved by SDS-PAGE and transferred to nitrocellulose membrane. The membrane was blocked with 10% (w/w) bovine serum albumin in tris buffered saline (50 mM Tris.HCl, 150 mM NaCl, pH 7.4),
incubated with HRP-strepavidin (1:20000 dilution in TBS with 0.1% Tween-20, Pierce, Rockford, IL) and visualized by chemiluminescence.

**Preparation of Tau Containing Protein Fractions from Rat Brain.** Rat brain (3.4 g) was homogenized in 12.5 mL of ice cold 1% perchloric acid (Polytron Homogenizer, Glen Mills, Clifton, NJ.). The resulting suspension was incubated on ice for 20 min and centrifuged at 20,000xg for 20 min. Supernatant was passed through a 1 μm filter and concentrated with a 10 kDa molecular weight cut-off membrane with simultaneous buffer exchange to 10 mM HEPES, pH 7.5. Proteins (1.9 mg) were fractionated on Suparose 12 PC 3.2/30 gel filtration column (GE Healthcare) by using a buffer containing 20 mM HEPES and 50 mM NaCl, pH 7.5. Tau containing fractions were combined. Total protein amount was estimated by absorbance at UV 280 nm. An aliquot of this material (about 1.3 μg) was used for O-GlcNAc site mapping.

**LC-MS Analysis of Tagged and Enriched O-GlcNAc Modified Peptides.** CAD spectra were recorded on a LTQ-Orbitrap XL (Thermo Fisher Scientific, Bremen, Germany) interfaced to an Eksigent nano-LC system (Dublin, CA). The HPLC gradient was 5-40% solvent B (A = 0.1% formic acid; B = 90% acetonitrile, 0.1% formic acid) in 40 min at a flow rate of 300 nL/min. CAD MS/MS spectra were recorded with the LTQ operated in the data dependent mode on the five most intense ions observed in MS¹ scans recorded with the Fourier transform analyzer set at a resolution of 60,000 at m/z 400. Parameters for acquiring MS/MS spectra in the ion trap were as follows; activation time = 30 ms; activation Q = 0.25; dynamic exclusion=enabled with a repeat count of 2 and exclusion duration = 60 sec. For ETD analysis, an aliquot of sample reconstituted in 0.1% acetic acid was pressure loaded onto a 360 μm o.d. x 75 μm i.d. microcapillary pre-column packed with C18 (5-20 μm diameter, 120 Angstrom), and then washed with 0.1% acetic acid as previously described (7). The pre-column was connected to a 360 μm o.d. x 50 μm i.d. microcapillary analytical column packed with C18 (5 μm diameter, 120 Angstrom) and equipped with an integrated electrospray emitter tip (7). Peptides were
gradient-eluted into the mass spectrometer at a flow rate of 60 nL/min. ETD MS/MS spectra were acquired using the following parameters (reaction time=100 milliseconds, reagent AGC target=4E5 ion counts, full AGC target= 2E4 ion counts, MSn AGC target=2E4 ion counts, isolation window= 3 m/z, repeat count=2, repeat duration=20 seconds, exclusion duration=30 seconds).

**Analysis of Mass Spectrometry Data.** Peak lists were generated from raw data files using Bioworks software (version 3.3.1 sp1). In-house developed software was used to remove charge reduction species. Open Mass Spectrometry Search Algorithm (OMSSA) (version 2.1.1) was utilized to search c- and z-type fragment ions present in ETD MS/MS spectra against all rat and mouse proteins (mouse and rat, 193,424 entries) in the NCBI non-redundant NR database (release date 11/2008). Database searches were performed with the following fixed parameters; precursor mass tolerance of +/- 1.5 Da, product ion mass tolerance +/- 0.35 Da, and 3 missed cleavages, in addition to the following variable modifications; methionine in the oxidized and non-oxidized forms, and serine and threonine residues with and without the tagged O-GlcNAc group. All database assignments of O-GlcNAc modified peptides were confirmed by manual interpretation of the corresponding ETD MS/MS spectra.

**RESULTS AND DISCUSSION**

**Approach for highly specific enrichment and site-mapping of GlcNAcylation**

For this enrichment approach (Figure 1a), proteins were first proteolytically digested. Following this, GlcNAcylated sites were labeled with an azido sugar using the substrate UDP-N-azidoacetylgalactosamine (UDP-GalNAz) the galactosyltransferase GalT1 (6). Since the UDP product formed in this reaction potently feedback inhibits GalT1, alkaline phosphatase was added to the reaction mixture to rapidly degrade UDP and to ensure complete derivitazation. ([Supplementary Fig. 3 online](#)).

Peptide: N-Glycosidase F (PNGase F), which removes N-linked glycans, was also added to the reaction mixture to ensure N-glycans with terminal GlcNAc residues were not labeled with GalNAz. Excess
UDP-GalNAz was removed using a C18 spin column. GalNAz groups were subsequently biotinylated through a copper-catalyzed 1,3 dipolar cycloaddition reaction of the free azide on GalNAz to an alkyne containing a terminal biotin group and a photocleavable linker (8). After removing excess biotinylation reagent by strong cation exchange (SCX), enrichment of the tagged O-GlcNAc modified peptides was performed by avidin affinity chromatography. A major issue with using avidin chromatography is the difficulty in elution of biotin-containing peptides because binding of biotin to avidin is essentially irreversible \((K_a = 10^{15} \text{ M}^{-1})\). The elution efficiency is also low when using monomeric avidin. Therefore harsh conditions are generally used for elution from avidin beads; however, this can easily damage O-GlcNAc. Reduction-cleavable biotin is not suitable for this application because the dipolar addition reaction requires strong reducing conditions. Acid-cleavable biotin typically requires treatment with 95% trifluoroacetic acid, which we also find is strong enough to partially hydrolyze O-GlcNAc.

Olejnik et al. (9) previously reported a photocleavable biotin reagent. Based on their findings, we synthesized a photocleavable biotin-alkyne reagent to tag GlcNAc modified peptides (Fig. 1a, inset). This photocleavable biotin-linked alkyne contains a terminal biotin group, which facilitates enrichment of GlcNAc modified peptides. In addition, after enrichment with avidin chromatography, the tagged GlcNAc modified peptides are efficiently released upon brief exposure to UV light (365 nm) (Supplemental Figure 4) and the released peptides carry a basic aminomethyltriazolylacetyl-galactosamine group. This overall approach is illustrated in Figure 1a. Shown in Figure 1b is a CAD MS/MS spectrum recorded on \([M+3H]^{+3}\) ions (m/z 457) corresponding to the tagged, GlcNAclylated tryptic peptide having the sequence, YSPTgSPSK. Note that the dominant fragmentation pathway occurs at the glycosidic linkage to produce oxonium ions at m/z 300.1 and 503.2. These signature ions are diagnostic for tagged GlcNAc residues and can be used to detect GlcNAcylated peptides in complex mixtures. Figure 1c displays the ETD MS/MS spectrum recorded on the same \([M+3H]^{+3}\) ions. Because the tag added to the O-GlcNAc residue contains a basic group, all GlcNAcylated tryptic peptides should
exist in charge state of +3 or higher and thus fragment well along the peptide backbone under ETD conditions. Cleavage at the glycosidic linkage is not observed.

**Detecting GlcNAcylation in α-crystallin**

To test our enrichment strategy, we analyzed a mixture containing tryptic peptides from α-crystallin (50 fmol), which has two known O-GlcNAc sites, and bovine serum albumin (2.5 pmol). GlcNAc peptides were tagged with GalNAz followed by photocleavable-biotin, enriched and subsequently released from an avidin resin, and analyzed by CAD using a LTQ-Orbitrap mass spectrometer and by ETD using a LTQ XL mass spectrometer. Shown in **Figure 2a** is the CAD spectrum recorded on \([M+3H]^+\) ions at m/z 715 from the tagged-C-terminal tryptic peptide (AIPVgSREEKPSSAPSS) of α-crystallin. Note that the signals labeled as Peak 1 and Peak 2 correspond to the diagnostic oxonium ions, discussed above. Shown in **Figure 2b** is the ETD MS/MS spectrum recorded on \([M+4H]^+\) ions for the same peptide. Predicted fragment ions of type c and z are shown above and below the insert in **Figure 2b**. Those ions that were observed are underlined. Note that the mass difference between both c5 and c4 and z12 and z11 is 589 Da rather than the 87 Da expected for an unmodified Ser residue. We conclude that the first Ser in the peptide carries the tagged O-GlcNAc moiety. Tagged EEKPAVgTAAPK from the β-chain of α-crystallin was also detected (data not shown). Unmodified BSA peptides were not detected in the analysis indicating that the enrichment proceeded with high selectivity. The level of O-GlcNAc on α-crystallin is approximately 10% (11). Knowing this, we estimate that recovery of the two O-GlcNAc modified peptides from α-crystallin is over 90%.

Next, we asked if this tagging strategy could be employed to detect O-GlcNAc at the intact protein level without using mass spectrometry. As shown in **Figure 2c**, tagged α-crystallin was readily detectable by avidin-HRP blotting. In contrast, the signal disappeared when the tagged protein was illuminated by UV before the sample was resolved by SDS-PAGE. We conclude that tagged proteins
and proteins bound non-specifically to the avidin beads can be readily distinguished by blotting samples before and after the photochemical cleavage.

**Detecting and site-mapping GlcNAcylation in complex biological sample**

The described protocol was used to analyze tau-containing protein fractions obtained from rat brain. These fractions were isolated by a combination of perchloric acid extraction and gel filtration chromatography. Eight O-GlcNAc sites on 7 different proteins were identified (*Table 1 and Supplementary Figure 5 online*). Shown in **Figure 3a** is the ETD mass spectrum recorded on [M+3H]+3 ions from residues 709-717 of the tau microtubule associated protein. Note that the mass difference between both c3 and c2 and z7 and z6 is 589 Da rather than the 87 Da expected for an unmodified Ser residue. We conclude that Thr711 (more commonly known as Thr400 according to the most common splice variant) in the carboxyl terminal region of tau is O-GlcNAc modified. This same site can also be phosphorylated by glycogen synthase kinase-3β (GSK-3β)(12). These data support our hypothesis that one function of O-GlcNAc on tau is to prevent hyperphosphorylation in normal brain (13, 14). Since we observed limited digestion of tau with trypsin, we anticipate that additional O-GlcNAc sites will be detected when multiple proteases are employed to extend the sequence coverage.

From the same tau containing protein fractions, O-GlcNAc sites on α-, β-, and γ-synuclein (*Table 1*) were also characterized. α-, β-, and γ-synuclein are three small proteins that are expressed primarily in neural tissues (15). Aggregated α-synuclein is a component of filamentous inclusions associated with neurodegenerative conditions, such as Alzheimer’s and Parkinson’s disease (16). Here we report that α-synuclein is O-GlcNAc modified on Thr72, a residue located within a 35 amino acid stretch that forms the protease-resistant core of α-synuclein fibrils (15). This region is prone to self-aggregation and is able to seed the formation of amyloid fibrils (17). It seems likely that GlcNAcylation on this peptide could prevent fibril formation. Interestingly, Shimura et. al. (18) reported that the E3 ubiquitin protein ligase, parkin, binds and ubiquinates glycosylated α-synuclein, but not the non-
glycosylated form in human brain, although the identity and site of the glycosylation were unknown at that time.

Two O-GlcNAc were also identified on the methyl-CpG-binding protein (MeCP2). One site was unambiguously identified at Thr434 (Fig. 3b) while the second site was localized to either Thr443 or Thr444. This protein is also of interest since it binds to methylated DNA and mediates transcriptional repression through interactions with a histone deacetylase and corepressor SIN3A (19). Mutations in MeCP2 cause Rett syndrome, a developmental disorder characterized by mental retardation, motor dysfunction, and autistic behavior (20). Brain specific phosphorylation of Ser421 in MeCP2 by a CaMKII-dependent mechanism is thought to regulate a program of gene expression that mediates nervous system maturation (21). Whether or not GlcNAcylation on the nearby residues (T434 and either T442 or T443) alters this process will require further investigation.

In summary, we describe here an efficient enrichment protocol that is highly specific for O-GlcNAc modified proteins and peptides, and that facilitates site mapping of O-GlcNAc modified amino acids at the low femtomole level by ETD mass spectrometry. The procedure employs a novel photocleavable biotin tag that allows for efficient release of the enriched O-GlcNAc modified peptides from the solid affinity support. In addition, the photocleavage reaction leaves a basic aminomethyltriazole tag at the site of the O-GlcNAc modification. As a result, all modified, tryptic peptides exist in a charge state of +3 or higher, and thus fragment efficiently along the peptide backbone when subjected to ETD. For site-specific O-GlcNAc quantification, a heavy isotope labeled photocleavable biotin alkyne is currently being synthesized.

Because the enrichment procedure is highly specific, the flow-through from the avidin-chromatography can be further enriched for phosphopeptides or for peptides with other post-translational modifications. Note, in the case, the optional use of a phosphatase during the labeling process must be omitted. A flow chart diagramming the sequential enrichment of GlcNAcylated
peptides and phosphorylated peptides is shown in Supplemental Figure 6. In a recent large-scale study, we applied this serial enrichment protocol (leaving out the alkaline phosphatase) to investigate the interplay of phosphorylation and GlcNAcylation in the regulation of cytokinesis. By combining the above serial enrichment protocol with SILAC (Stable Isotope Labeling by Amino acid in Cell culture), we have mapped, quantified, and compared relative site occupancy for over 120 specific O-GlcNAc modified residues and over 350 phosphorylated residues by tandem mass spectrometry using only 15 µg of sample from a spindle/midbody preparation. This confirms the robustness of our protocol for analysis of complex mixtures (Wang et al., submitted). Since the sensitivity of the above approach is already on par with that used to study protein phosphorylation, we expect that the method will be a valuable tool for deciphering the “O-GlcNAcome”. Together with other methods (6, 22), we expect a significant increase of known O-GlcNAc sites in the very near future, which will highly facilitate investigation of the function of GlcNAcylation.

ACKNOWLEDGEMENTS

We thank Ms. Kyoungsook Park and Ms. Shino Shimoji for providing synthetic O-GlcNAc peptides. We thank Dr. Robert J. Cotter for providing access to Orbitrap mass spectrometer (supported in part by S10RR023025-01 to R.J.C. from High End Instrumentation Program of NIH). This work is supported by United States Public Health Service, NIH Grants, DK61671, CA42486 and NHLBI Contract N01-HV-28180 to G.W.H. and by GM37537 to D.F.H.

AUTHOR CONTRIBUTIONS

Z.W. conceived the idea and developed the methods. Z.W. performed the photocleavable biotin-alkyne synthesis, chemoenzymatic tagging, and CAD mass spectrometric analysis. N.U., J.S., and D.F.H. performed ETD mass spectrometric analysis and ETD data analysis. M.O. prepared tau samples and
assisted in chemoenzymatic tagging. G.W.H contributed to the idea, conception, and overall experimental design. Z.W., N.U., M.O., D.F.H., and G.W.H. contributed to the writing of the manuscript.

REFERENCES


FIGURE LEGENDS

Figure 1. Enrichment of O-GlcNAc Modified peptides by a combination of enzymatic labeling with UDP-GalNAz and chemical derivatization with the photocleavable reagent, PC-PEG-Biotin-Alkyne. (a) Flow chart showing the overall strategy. Inset; structure of the photocleavable, PC-PEG-Biotin-alkyne reagent. CIP: calf intestine phosphatase; PNGase F: Peptide: N-Glycosidase F. (b) CAD mass spectrum of a tagged synthetic O-GlcNAc modified peptide, YSPT\textsubscript{gS}SPSK (gS = O-GlcNAcylated Ser) showing signature ions at m/z 300.2 and 503.1 that result from cleavage at the two sugar ketal linkages and that confirm the presence of the tagged O-GlcNAc moiety. (c) ETD mass spectrum recorded on [M+3H]\textsuperscript{+3} (m/z 457.2) ions from the same peptide, YSPT\textsubscript{gS}SPSK. Predicted m/z values for ions of type c’ and z’ (monoisotopic and average masses for singly- and doubly-charge ions, respectively) are shown above and below the peptide sequence. Observed product ions are underlined and also labeled in the spectrum. Ions in the precursor isolation window are labeled with a triangle (▼). Brackets enclose ions that correspond to charge-reduced species and fragments derived from them by loss of small, neutral molecules. Product ions that result from loss of an aminomethyltriazole radical are labeled with a circle (○). M represents the tagged peptide.

Figure 2. Characterization of an O-GlcNAc modified peptide from \(\alpha\)-crystallin. (a) CAD mass spectrum showing signature ions characteristic of the derivatized O-GlcNAc moiety. M represents the tagged peptide. (b) ETD mass spectrum recorded on [M+4H]\textsuperscript{+4} ions (m/z 537.10) from the same peptide, AIPV\textsubscript{gS}REEKPSSAPSS. Predicted m/z values for ions of type c’ and z’ (monoisotopic masses and average masses for singly- and doubly-charge ions, respectively) are shown above and below the peptide sequence. Observed product ions are underlined and also labeled in the spectrum. Ions in the precursor isolation window are labeled with a triangle (▼). Brackets enclose ions that correspond to
charge-reduced species and fragments derived from them by loss of small, neutral molecules. Product ions that result from loss of an aminomethyltriazole radical are labeled with a circle (o). (c) The tagging approach enables detection of O-GlcNAc modified proteins by avidin-HRP blotting. This figure shows that tagged α-crystalline provides a strong signal after blotting with avidin-HRP blotting. Upon UV illumination no signal is observed indicating photocleavage and loss of the biotin tag from α-crystallin.

**Figure 3.** Enriched and tagged O-GlcNAc modified peptides from tau-containing, protein fractions isolated from rat brain. (a) ETD MS/MS spectrum recorded on [M+3H]$^{+3}$ ions (m/z 474.2) for the derivatized peptide, VVgSDTSPR, from the tau protein. (b) ETD MS/MS spectrum recorded on [M+3H]$^{+3}$ ions (m/z 748.1) for the derivatized peptide, VAAAAgTTTTTTTTVVAEK, from the methyl-CpG-binding protein, MeCP2. Spectra are labeled as described above in Figures 1c and 2b.
Table 1. O-GlcNAc Sites Identified using ETD Mass Spectrometry.

<table>
<thead>
<tr>
<th>Protein</th>
<th>GI number</th>
<th>Peptide Sequence</th>
<th>O-GlcNAc Site(s)</th>
<th>Novel?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tau</td>
<td>517394</td>
<td>V331VgSGDTSPR339</td>
<td>S711</td>
<td>Yes</td>
</tr>
<tr>
<td>Alpha-synuclein</td>
<td>149037058</td>
<td>61EQVTNVGGAVVG\text{\textit{gT}}GVTAVAQK80</td>
<td>T72</td>
<td>Yes</td>
</tr>
<tr>
<td>Gamma-synuclein</td>
<td>122066261</td>
<td>59TKEQANAVgSEAVVSSVNTVATK80</td>
<td>S67</td>
<td>Yes</td>
</tr>
<tr>
<td>Beta-synuclein *</td>
<td>77404215</td>
<td>24QGVgTEAAEK32</td>
<td>T27</td>
<td>Yes</td>
</tr>
<tr>
<td>Host cell factor C1</td>
<td>213385315</td>
<td>1147lTvPGALER1156</td>
<td>T1148</td>
<td>Yes</td>
</tr>
<tr>
<td>Methyl CpG binding protein 2</td>
<td>149029883</td>
<td>434VAAAAGTTTTTTTTTTTTVAEK455</td>
<td>T443</td>
<td>Yes</td>
</tr>
<tr>
<td>Spectrin beta 2</td>
<td>149044856</td>
<td>2310HDTASTQSTPASgSR2324</td>
<td>S2323</td>
<td>No5</td>
</tr>
</tbody>
</table>

*, the same peptide is also in gamma-synuclein.
Figure 1

a.

b.
c.
<table>
<thead>
<tr>
<th></th>
<th>Crystallin</th>
<th>PC-Biotin-Alkyne</th>
<th>UDP-GalNAz</th>
<th>GaT1</th>
<th>UV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>+</td>
</tr>
</tbody>
</table>
Figure 3

a.

b.