Mapping Sites of *O*-GlcNAc Modification Using Affinity Tags for Serine and Threonine Post-translational Modifications*

Lance Wells‡§, Keith Vosseller‡§, Robert N. Cole‡, Janet M. Cronshaw¶, Michael J. Matunis¶, and Gerald W. Hart‡‖

Identifying sites of post-translational modifications on proteins is a major challenge in proteomics. *O*-Linked β-*N*-acetylglucosamine (*O*-GlcNAc) is a dynamic nucleocytoplasmic modification more analogous to phosphorylation than to classical complex *O*-glycosylation. We describe a mass spectrometry-based method for the identification of sites modified by *O*-GlcNAc that relies on mild β-elimination followed by Michael addition with dithiothreitol (BEMAD). Using synthetic peptides, we also show that biotin pentylamine can replace dithiothreitol as the nucleophile. The modified peptides can be efficiently enriched by affinity chromatography, and the sites can be mapped using tandem mass spectrometry. This same methodology can be applied to mapping sites of serine and threonine phosphorylation, and we provide a strategy that uses modification-specific antibodies and enzymes to discriminate between the two post-translational modifications. The BEMAD methodology was validated by mapping three previously identified *O*-GlcNAc sites, as well as three novel sites, on Synapsin I purified from rat brain. BEMAD was then used on a purified nuclear pore complex preparation to map novel sites of *O*-GlcNAc modification on the Lamin B receptor and the nucleoporin Nup155. This method is amenable for performing quantitative mass spectrometry and can also be adapted to quantify cysteine residues. In addition, our studies emphasize the importance of distinguishing between *O*-phosphate *versus* *O*-GlcNAc when mapping sites of serine and threonine post-translational modification using β-elimination/Michael addition methods. ***Molecular & Cellular Proteomics 1:791–804, 2002.***

The rapid identification of proteins by mass spectrometry has become commonplace in the postgenomic era (1). However, one major challenge that remains is the identification of post-translational modifications on these proteins. More than 25 years ago, Finn Wold and colleagues (2) recognized the abundance of naturally occurring modified forms of the genetically encoded 21 amino acids. In addition to phosphorylation, a variety of post-translational modifications, including acetylation (3), methylation (4), and *O*-linked β-*N*-acetylglucosamine (*O*-GlcNAc)[1] (5–7), are now recognized to regulate protein functions in cellular processes. Therefore, identification of proteins along with their post-translational modifications, which has been referred to as "functional proteomics," is an important step in the characterization of proteomes. *O*-GlcNAc is a dynamic post-translational modification occurring on a variety of nucleocytoplasmic proteins and, in several instances, *O*-GlcNAc maps to the same or adjacent sites as phosphorylation (8, 9). Diverse classes of proteins are modified including cytoskeletal proteins, transcription factors, signaling adapter molecules, hormone receptors, nuclear pore complex (NPC) proteins, and kinases (10). The nucleocytoplasmic enzymes for the addition (*O*-GlcNAc transferase) and removal (neutral β-*N*-acetylglucosaminidase, *O*-GlcNAcase) of this modification have been cloned and characterized (11–16) and may act analogously to kinases and phosphatases (17). The functional importance of the *O*-GlcNAc modification has been illustrated by several groups. For instance, *O*-GlcNAc transferase is recruited to mSin3A transcriptional repression complexes where its activity is necessary for optimal repression of transcription (18). Further, elevated *O*-GlcNAc levels via pharmacological or genetic means attenuate insulin signaling, leading to insulin resistance in cultured adipocytes and in whole animals (19, 20). Also, deletion of the *O*-GlcNAc transferase gene is lethal in mouse embryonic stem cells (21). However, given the widespread occurrence of this modification, relatively few studies addressing its functional significance have been undertaken. The identification of new *O*-GlcNAc-modified proteins and the sites of modification would facilitate more global studies of the regulatory role of this post-translational modification.

From the ‡Department of Biological Chemistry, Johns Hopkins University School of Medicine and the ¶Department of Biochemistry and Molecular Biology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21218

Received, August 20, 2002, and in revised form, September 23, 2002

Published, MCP Papers in Press, September 26, 2002, DOI 10.1074/mcp.M200048-MCP200

[1] The abbreviations used are: *O*-GlcNAc, *O*-linked β-*N*-acetylglucosamine; DTT, dithiothreitol; BEMAD, β-elimination followed by Michael addition with DTT; NPC, nuclear pore complex; *m/z*, mass to charge ratio; BAP, biotin pentylamine; MS, mass spectrometry; MS/MS, tandem MS; LC-MS/MS, reverse-phase liquid chromatography to nanospray MS/MS; BPP, basic phosphoprotein; PBS, phosphate-buffered saline; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight.

© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

While the mapping and subsequent mutagenesis of phosphorylation sites has been a valuable approach for determining site-specific functions of phosphate, less than 50 sites of *O*-GlcNAc modification have been published (10). The technique used to map most of the known sites of *O*-GlcNAc modification was based on the traditional enzymatic tagging of *O*-GlcNAc with radiolabeled galactose (22). This method lacks sensitivity, necessitating purification of relatively large amounts of protein. The technique is also labor intensive, usually involving purification of the radiolabeled peptides and Edman sequencing. Direct observation of *O*-GlcNAc in mass spectrometry during collision-induced dissociation has proved difficult as the glycosidic linkage is labile and easily cleaved, resulting in very little peptide fragmentation (23–25). Aebersold and Haynes (24) enriched galactose-tagged *O*-GlcNAc-modified proteins but still found that the modification was labile. Greis and coworkers (23) performed MS/MS analysis to map sites on peptides by identifying dehydroamino acids resulting from prior β-elimination of *O*-GlcNAc. However, under their conditions, significant peptide degradation was observed, and no enrichment of modified peptides was possible. In addition, none of these methods provides a basis for doing comparative quantitation.

Post-translational modifications are substoichiometric and often labile, making their identification difficult. Method development for identification of phosphorylation sites has attempted to address these issues (26–29). Techniques include precursor ion scanning (30), enrichment of phosphopeptides via immobilized metal affinity chromatography (31, 32) or antibodies (33), and chemical modification of the phosphoamino acid for stabilization and enrichment (34–38). β-Elimination of phosphate from serine or threonine followed by attack of the resulting α,β-unsaturated carbonyl with a nucleophile that allows for enrichment and also results in a "tag" conferring a unique molecular weight to the modified amino acid has been used successfully to map phosphorylation sites (34–36, 38). However, any modification of serines and threonines that is susceptible to β-elimination may be targeted by this method as well as cysteines that have been alkylated or methionines (39). *O*-GlcNAc is more susceptible to alkali-induced β-elimination than is *O*-phosphate (23, 39–41). Therefore, we tested the potential use of β-elimination followed by Michael addition chemistry for mapping *O*-GlcNAc-modified serines and threonines.

In this report, we describe an adapted method that uses mild β-elimination followed by Michael addition of dithiothreitol (Cleland's reagent, DTT) (BEMAD) or biotin pentylamine (BAP) to tag *O*-GlcNAc sites (as well as phosphorylation sites). The tag allows for enrichment via affinity chromatography and is stable during collision-induced dissociation, allowing for site identification by LC-MS/MS. An immunoaffinity and enzymatic strategy is provided to discriminate between *O*-GlcNAc and phosphorylation sites with the use of BEMAD. We show that synthetic glyco- and phosphopeptides can be derivatized and enriched, and the site of modification can be identified using the BEMAD technique followed by affinity chromatography and LC-MS/MS. As validation of the method, we used BEMAD to map several previously known sites of *O*-GlcNAc modification on Synapsin I. Finally, we use this methodology to identify sites of *O*-GlcNAc modification on Nup155 and the Lamin B receptor from a purified NPC preparation.

EXPERIMENTAL PROCEDURES

Reagents—All reagents were of the highest grade commercially available. Anti-IgM-agarose, dimethyl pimelimidate, and DTT were purchased from Sigma. BAP was obtained from Pierce. Activated thiol-Sepharose columns and monoavidin columns were from Amersham Biosciences and Applied Biosystems, respectively. The 110.6 (*O*-GlcNAc-specific) antibody was from ascites (commercially available from Covance Research Products) (42). The synthesis of *O*-GlcNAc peptides has been previously reported (23) as well as site mapping of the BPP peptide (PSVPVS(*O*-GlcNAc)GSAPGR) (43). The phosphopeptide (KHFPQFS(P)YSAS) we used was derived from AKT and was obtained commercially (Upstate Biotechnology).

Rat Brain Extract Preparation—1 g of rat brain (Pel-Freeze)/7 ml of buffer (15 mM Tris/HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, protease inhibitor mixture, and 10 μM PUGNAc (*O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate)) was mechanically lysed with a Polytron (2 $\times$ 30-s pulse) and clarified at 35,000 $\times$ *g* for 30 min, and the supernatant was passed through a 1-μm glass fiber filter and stored at -80 °C until used.

Immunopurification of Rat Brain Extract—Lysates were precleared for 1 h at 4 °C with anti-IgM-agarose. Samples were then batch bound overnight at 4 °C with the *O*-GlcNAc-specific antibody 110.6 ascites (1 mg) covalently coupled with dimethyl pimelimidate to anti-IgM-agarose or anti-IgM-agarose alone. The affinity columns were washed five times with 10 column volumes of RIPA (1% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS in TBS (15 mM Tris/HCl, pH 7.5, 150 mM NaCl)), 5 column volumes of TBS, and then eluted with 1 M GlcNAc in TBS. The samples were desalted by buffer exchange into 40 mM ammonium bicarbonate using a spin column (10-kDa cutoff, Millipore).

Synapsin and NPC Preparation—Purification of Synapsin I from rat brain and the NPC preparation from rat liver and unit definition were performed exactly as described previously (44, 45).

SDS-PAGE, Coomassie Staining, and Western Blotting—SDS-PAGE was performed under reducing conditions on precast 10% Criterion gels (Bio-Rad). Coomassie G-250 staining and Western blotting of one-dimensional gels with 110.6 antibody was as described previously (42).

In-gel Digestion—Reduction, alkylation, and digestion were performed essentially as described previously (46). Briefly, Coomassie G-250-stained bands were excised, dehydrated with acetonitrile, and reswelled in 40 mM ammonium bicarbonate. This was repeated, and then gel pieces were reduced with 10 mM DTT for 1 h at 56 °C and then carboxyamidomethylated with 55 mM iodoacetamide in the dark for 45 min. The gel was dehydrated and reswollen in 40 mM ammonium bicarbonate with 10 ng/μl trypsin (Promega) on ice for 45 min. After excess trypsin was replaced with 40 mM ammonium bicarbonate, digestion was allowed to proceed overnight at 37 °C. The peptides were extracted three times for 20 min in 5% formic acid, 50% acetonitrile and dried down in a Speed Vac.

Phosphatase and N-Acetylglucosaminidase Treatment—Where indicated under "Results," peptides were spiked with *O*-GlcNAc BPP and/or phosphorylated AKT peptides and then subjected to dephosphorylation with alkaline phosphatase (1 unit/10 μl, Promega) for 3 h

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

at 37 °C in the presence of added 1 mM $MgCl_2$ or β-*N*-acetylglucosaminidase (1 unit/20 μl, New England Biolabs) for 16 h at 37 °C after acidification with trifluoroacetic acid to pH 4.5.

In-solution Digest—Both the 110.6-enriched proteins from rat brain and the nuclear pore preparation were buffer-exchanged using spin columns (10-kDa cutoff, Millipore) with 40 mM ammonium bicarbonate. 110.6-enriched samples were digested overnight at 37 °C by the addition of 1:10 (w/w) trypsin (Promega). Peptides were dried down in a Speed Vac. For the nuclear pore preparation, proteins were oxidized with performic acid (5% hydrogen peroxide, 45% formic acid (v/v)) for 1 h on ice and then dried down. The proteins were resuspended in 40 mM ammonium bicarbonate and digested overnight at 37 °C by the addition of 1:10 (w/w) trypsin. The peptides were dried down, resuspended in 40 mM ammonium bicarbonate, 1 mM magnesium chloride, and then spiked with 100 pmol of phospho-AKT peptide. 25 units of alkaline phosphatase were added, and the peptides were incubated at 37 °C for 4 h and then dried in a Speed Vac.

BEMAD—Dried down peptides were β-eliminated and subjected to Michael addition with DTT or BAP via resuspension in 1% triethylamine, 0.1% NaOH, 0–20% ethanol (peptides 20%, proteins 0%), and 10 mM DTT (or 25 mM BAP). The final pH was adjusted with triethylamine to 12.0–12.5 if necessary. The reaction was incubated at 50 °C for various amounts of time (see "Results," biological samples, 2.5 h), and the reaction was quenched with trifluoroacetic acid (final concentration, 1%). The peptides were cleaned up via reverse-phase C18 spin columns (The Nest Group), eluted in 0.1% trifluoroacetic acid, 70% acetonitrile, and dried in a Speed Vac.

Affinity Chromatography—DTT-modified peptides were purified over activated thiol-Sepharose (thiol column) from Amersham Biosciences. Resin was swelled in degassed PBS containing 1 mM EDTA (PBS/EDTA), and dried peptides suspended in the same buffer were bound with a 1-h incubation in 200 μl of 50% slurry. The column was washed with 15 ml of PBS/EDTA and eluted three times sequentially with 150 μl of PBS/EDTA containing 20 mM free DTT. BAP-derivatized peptides were enriched by sequential cation exchange and monoavidin columns provided in isotope-coded affinity tag kits from Applied Biosystems according to the manufacturer's protocol. Peptides eluted from thiol or avidin affinity columns were acidified (brought to 1% trifluoroacetic acid), desalted with reverse-phase C18 spin columns (eluted in 70% acetonitrile, 0.1% trifluoroacetic acid), and dried for subsequent analysis.

MALDI Analysis—Peptide samples were resuspended in matrix (10 mg/ml α-cyano-4-hydroxycinnamic acid in 0.3% trifluoroacetic acid, 60% acetonitrile) and analyzed in reflector or linear mode on an Applied Biosystems Voyager DE STR MALDI instrument.

LC-MS/MS Analysis—Peptides were resuspended in 1% acetic acid and loaded on a 10-cm × 0.075-mm column packed with 5-μm-diameter beads of C18 using positive N_2 pressure, desalted with 1% acetic acid, and then separated via a 75-min linear gradient of increasing acetonitrile at a flow rate of ~200 nl/min directly into the source (Finnigan LCQ, Ref. 47). In some cases following BEMAD treatment, to prevent disulfide formation, the samples were loaded under mild reducing conditions (200 μM DTT). The LCQ was run in automatic mode collecting a MS scan (2 × 500 ms) followed by two MS/MS scans (3 × 750 ms) of the two highest intensity peptides with a dynamic exclusion set at 2 with a mass gate of 2.0 daltons.

Data Analysis—Turbosequest software was used to interpret all MS/MS data (48). For DTT- or BAP-modified peptides, a mass increase of 136.2 or 310.5 daltons, respectively, was allowed on serines and threonines. When the samples were subjected to alkylation with iodoacetamide an increase of 57.052 daltons was allowed for cysteines. We also allowed for a mass increase of 120.2 daltons for cysteine when we found that our alkylated cysteines became derivatized using DTT. Samples that were treated with performic acid allowed for oxidation of cysteine (48.0 daltons), tryptophan (32.0 daltons), and methionine (32.0 daltons). All MS/MS spectra identifying proteins or peptides reported were the best hit in a non-redundant FASTA data base (downloaded April 22, 2002 from NCI, National Institutes of Health at Frederick, MD) search with an Xcorr > 2.5 (unless otherwise stated under "Results" or in figure legends) and were manually inspected for accuracy.

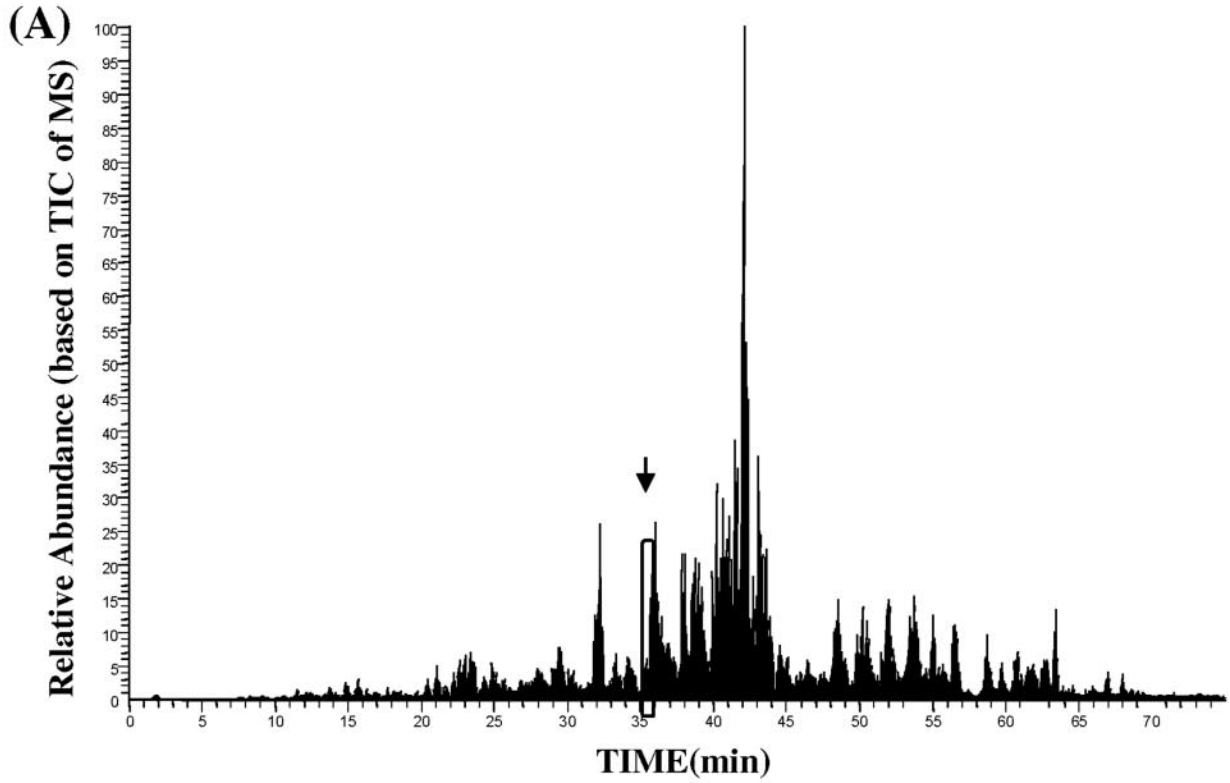
RESULTS

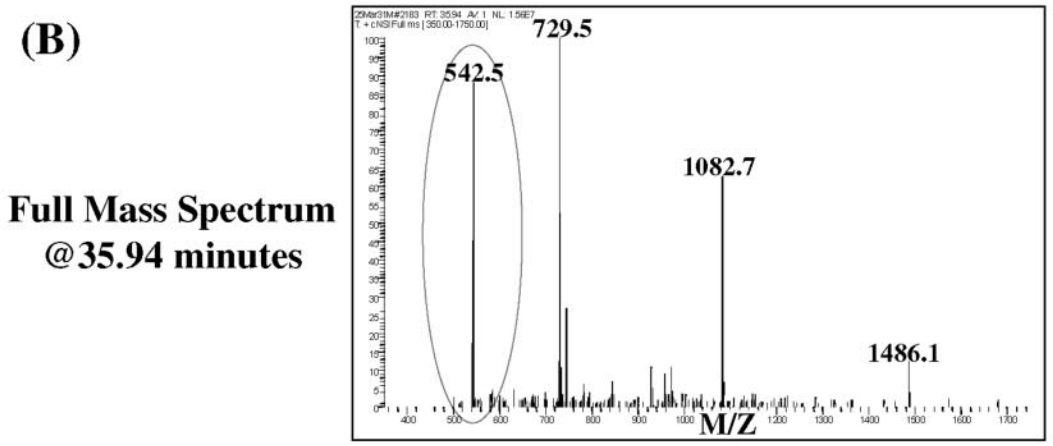
Identification of Multiple O-GlcNAc-modified Proteins—We previously demonstrated that *O*-GlcNAc occurs on a multitude of proteins using both lectin (succinylated wheat germ agglutinin) chromatography and immunoaffinity chromatography (with the *O*-GlcNAc-specific antibody 110.6) followed by two-dimensional gels and silver staining (6, 10). Since phosphorylated proteins have been identified using immunoaffinity purification (33), we sought to enrich and identify *O*-GlcNAc-modified proteins using an antibody column. We performed 110.6 immunoaffinity chromatography on a rat brain extract that had been precleared with anti-IgM-agarose. Samples that were specifically eluted with free GlcNAc were digested with trypsin, and the complex crude mixtures were analyzed by a 75-min run by LC-MS/MS (Fig. 1*A*). We also performed a control experiment using anti-IgM-agarose that had not been coupled to the 110.6 antibody and subtracted the resulting nonspecific proteins (actin, tubulin, serum albumin, and human keratins) from the 110.6 results. In Fig. 1, we show the MS profile at 35.94 min of the run (Fig. 1*B*) and the resulting MS/MS fragmentation of the *m/z* 542.5 peptide (Fig. 1*C*) that corresponds to a peptide from the small G protein Ran when analyzed (Table I). Analysis of the resulting data identified 30 proteins (with coverage ranging from 2 to 12 sequenced peptides that matched the listed protein as the best hit in the data base, Table I). Five of these proteins had previously been shown to be modified by *O*-GlcNAc (the nucleoporins p54 and p62 (49), *O*-GlcNAc transferase (11), the heat shock protein HSP90,[2] and the kinase casein kinase II (13, 14)) suggesting the method was valid. The other 25 proteins could be subdivided into multiple functional classes (Table I). Potentially, some of the identified proteins may be false positives, having co-purified with an *O*-GlcNAc-modified binding partner. Additionally, this strategy only identifies the most abundant proteins in the complex mixture and does not provide any information about sites of *O*-GlcNAc modification.

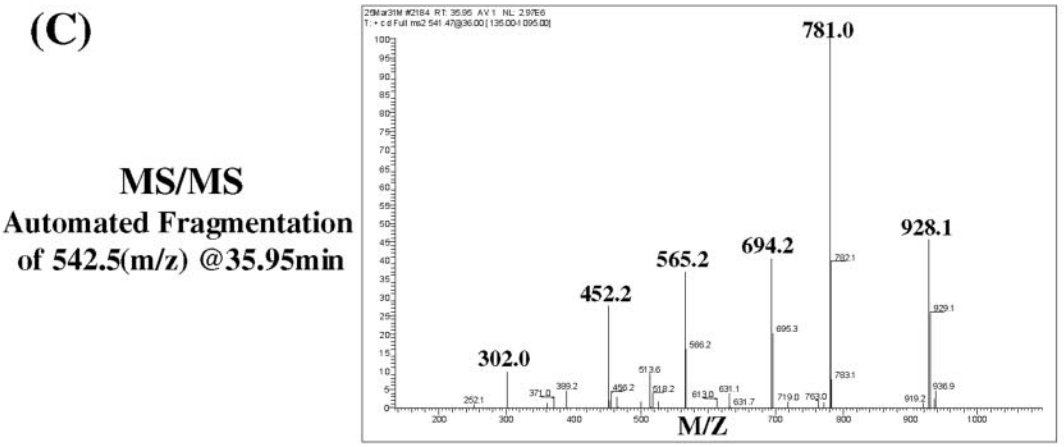
Modification, Enrichment, and Site Mapping of Glyco- and Phosphopeptides—β-Elimination followed by Michael addition of an affinity tag has been used by others to map sites of phosphorylation (34–36, 38). Thus, we adapted this strategy to map sites of *O*-GlcNAc modification. To stabilize sites of *O*-GlcNAc modification, enrich these peptides, map the sites of modification, and identify the protein, we developed the BEMAD method (Fig. 2*A*) followed by affinity chromatography and LC-MS/MS. Initially, we tested this method on synthetic

[2] S. Iyer and G. W. Hart, unpublished data.

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019







Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

TABLE I
Anti-O-GlcNAc immunopurified proteins

Glycolysis
Glyceraldehyde-3-phosphate dehydrogenase
Phosphoglycerate kinase
Phosphopyruvate hydratase (enolase)
Pyruvate kinase
Glycogen synthesis
UDP-glucose pyrophosphorylase
Cytoskeleton/vesicle trafficking
Protein Transport Protein SEC23
Annexin I (calpactin II)
Presynaptic cytomatrix protein Piccolo
Cytoplasmic Dynein light chain 1
Cofilin
Nuclear pore/transport
Nucleoporin p62
Nucleoporin p54
Ran (GTP-binding nuclear protein Ran)
Protein folding/stability/degradation
HSC70
HSP90
Proteasome component C2
Nucleophosmin (nucleolar phosphoprotein b23)
Peptidylprolyl isomerase
Q04323, homologous to ubiquitin carboxyl-terminal hydrolase
Transcription/translation
Human C1 (HCF) transcription factor
KIAA0144, Oct 1 transcription factor homologue
Enhancer factor 2D
Ewing sarcoma RNA-binding protein (EWS)
Eukaryotic initiation factor 4a-1 (EIF-4A-1)
Elongation factor 1-α
40 S ribosomal protein s24
Signal transduction
Casein kinase II
O-GlcNAc transferase (OGT)
Phosphatase 2A inhibitor (i2pp2a)
Rho GDP-dissociation inhibitor I (rho-gdi α)

O-GlcNAc- and *O*-phosphate-modified peptides (Figs. 2–5). We tried two different nucleophiles (DTT and BAP) for the Michael addition, both of which reacted equally well with the dehydroamino acids (Fig. 2, *C* and *D*). We successfully derivatized several other *O*-GlcNAc-modified peptides ((*a*) YSPTS(*O*-GlcNAc)PSK, (*b*) SPVVS(*O*-GlcNAc)GDTSPR, (*c*) QAGPPQAT(*O*-GlcNAc)RQASISGPAPPKVS(*O*-GlcNAc)GASPGGQQR, and (*d*) KKFELLPT(*O*-GlcNAc)PPLLSPSRR), finding that a peptide with an *O*-GlcNAc-modified threonine followed by a proline residue (*d*) was the most difficult to convert, requiring longer incubation times or an increase in pH for complete conversion (data not shown).

Next we compared the modification rates of glyco- and phosphopeptides. As expected, we found that under the mild β-elimination conditions used the *O*-GlcNAc-modified peptides converted to the BEMAD product much faster (Fig. 3) since *O*-glycosidic linkages are more easily eliminated than phospholinkages (23, 39–41). As can be seen in Fig. 3, at 2 h all of the serine-*O*-GlcNAc peptide had been converted, while a significant portion of the serine-*O*-phosphate peptide remained unmodified after 4 h of treatment. We also performed time courses on several other *O*-GlcNAc-modified peptides with similar results (data not shown). For subsequent work, we chose short incubation times for the BEMAD reaction (2.5 h) that appeared to maximally convert *O*-GlcNAc sites without completely modifying phosphorylation sites.

We next tested the ability of thiol columns and monomeric avidin columns to enrich the DTT and BAP peptides, respectively. We mixed together unmodified, *O*-GlcNAc-modified, and DTT- (Fig. 4*A*) or BAP (Fig. 4*B*)-modified peptides and then performed affinity chromatography. The unmodified and *O*-GlcNAc-modified peptides did not bind to the columns, and the BEMAD products bound and were specifically eluted (Fig. 4), demonstrating the utility of these columns for enrichment.

Next we tested the method for its ability to map sites on the modified peptides. MS/MS analysis of the DTT-modified peptide, allowing for modification of serines or threonines by 136.2 daltons (DTT-modified), generated fragment data that could be analyzed against a non-redundant data base (containing over 500,000 proteins) to identify the peptide, protein, and site of modification (Fig. 5). While three potential sites of modification exist in this peptide (three serines), the correct DTT-modified serine was unambiguously identified. MS/MS analysis of the same peptide modified with BAP was more difficult to interpret (data not shown). The BAP moiety carried a charge making the peptide triply charged, and the BAP also partially fragmented upon MS/MS analysis. Allowing for addition of 310.5 daltons (BAP-modified) to serines and threonines, Turbosequest identified the peptide and site of modification correctly but reported this correct hit as the seventh best hit in the non-redundant data base. Thus, DTT performed much better than BAP in the MS/MS analysis. DTT and the thiol column are also less expensive than the BAP and avidin columns. Furthermore, deuterated (d_{10}) DTT is commercially available (Isotec) making the method easily adaptable for performing quantitative mass spectrometry. Thus, subsequent site mapping was performed using DTT as the nucleophile.

Site Mapping on Purified Synapsin—For confirmation of this method, we wanted to perform site analysis on a known *O*-GlcNAc-modified protein that had previously had sites mapped. Previous work by our group had mapped seven sites of *O*-GlcNAc modification on Synapsin I isolated from rat brain (44). Thus, we purified Synapsin I from rat brain to homogeneity and confirmed the presence of *O*-GlcNAc via 110.6 Western blotting (Fig. 6*A*). Synapsin I was subjected to alkylation, digestion, alkaline phosphatase treatment (to de-

FIG. 1. **Identification of proteins immunopurified with an *O*-GlcNAc-specific antibody.** *A*, proteins enriched and specifically eluted from an anti-*O*-GlcNAc antibody column were digested with trypsin, and the resulting peptides were analyzed by nanospray LC-MS/MS. *B*, a single full MS scan at 35.94 min contains several peptides. *C*, MS/MS profile of the *m/z* 542.5 peptide that along with other MS/MS profiles allows for the identification of the small G protein Ran upon data base analysis. *TIC*, total ion current.

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

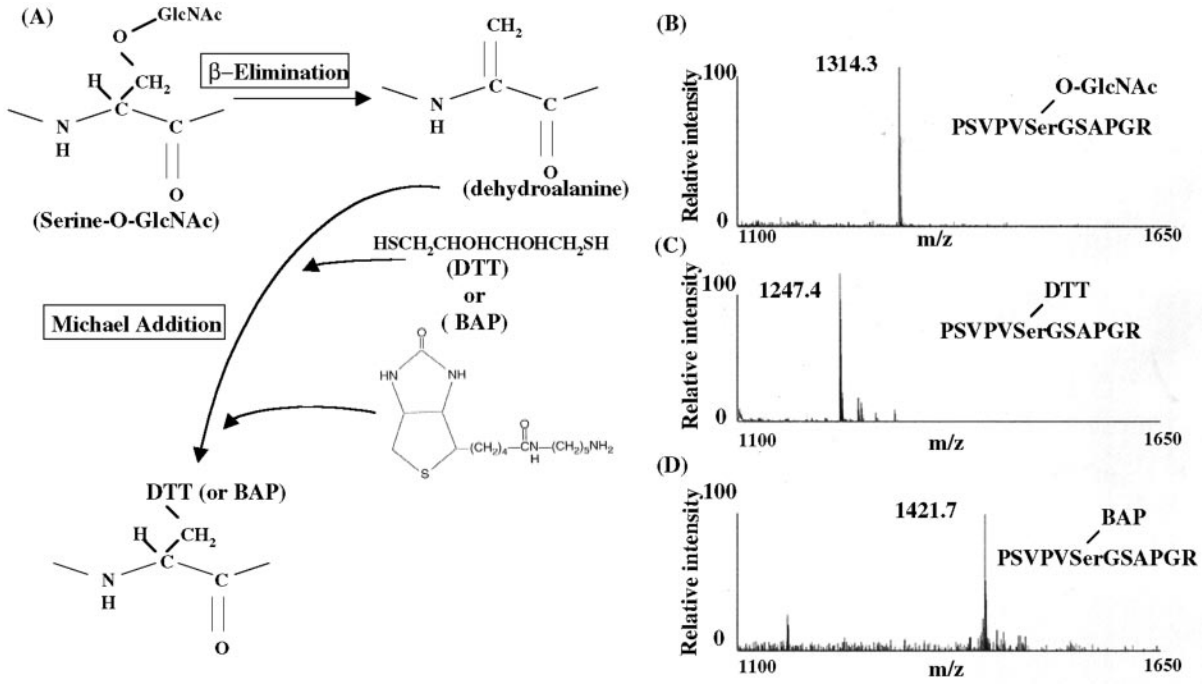


FIG. 2. **β-Elimination of *O*-GlcNAc and replacement with DTT (BEMAD) or BAP through Michael addition chemistry.** *A*, strategy for replacement of serine- or threonine-linked *O*-GlcNAc with the stable affinity tags DTT or BAP after β-elimination. *B*, *C*, and *D*, MALDI-TOF analysis of a synthetic *O*-GlcNAc-modified peptide that was untreated (*B*) or incubated at 50 °C for 2 h in 1% triethylamine, 0.1% NaOH in the presence of 10 mM DTT (*C*) or 20 mM BAP (*D*). Mass shifts in *C* and *D* correspond to loss of *O*-GlcNAc (203 daltons) and addition of DTT (136.2) and BAP (310.5), respectively.

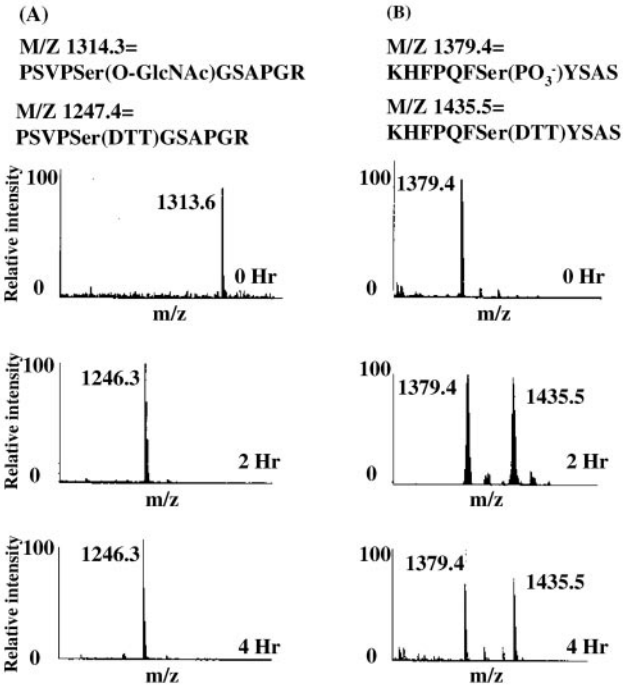


FIG. 3. **BEMAD performed on *O*-GlcNAc occurs under mild β-elimination conditions and more readily than on *O*-phosphate.** A synthetic *O*-GlcNAc-modified peptide (*A*) or a synthetic *O*-phosphorylated modified peptide (*B*) were untreated (0 h) or incubated at 50 °C for 2 or 4 h in 1% triethylamine, 0.1% NaOH in the presence of 10 mM DTT and analyzed by MALDI-TOF.

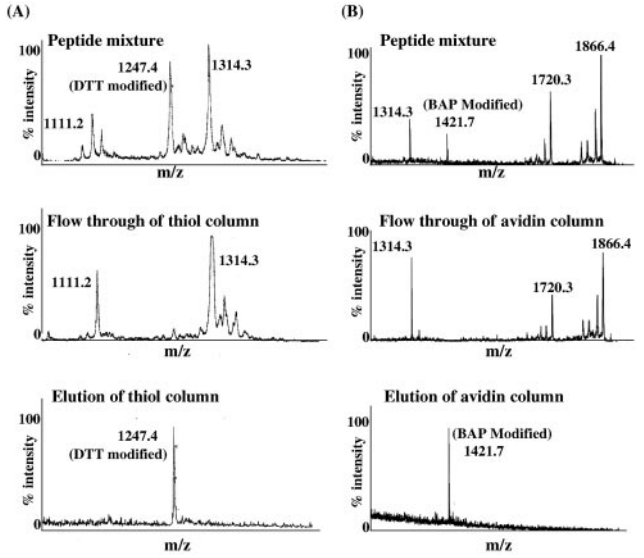


FIG. 4. **DTT- or BAP-modified peptides can be affinity-enriched from a mixture of peptides.** *A*, a peptide mixture containing the DTT-modified peptide PSVPVS(DTT)GSAPGR at [M + 1H] 1246.6 was bound to a thiol column (which forms mixed disulfides with free sulfhydryl groups) and eluted with 10 mM free DTT. The peptide mixture, the flow-through of the thiol column, and the eluted fraction were analyzed by MALDI-TOF. *B*, a peptide mixture containing the BAP-modified peptide PSVPVS(BAP)GSAPGR at *m/z* 1421.7 was bound to a tetrameric avidin column and eluted with 20 mM free biotin. The peptide mixture, the flow-through of the avidin column, and the eluted fraction were analyzed by MALDI-TOF.

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

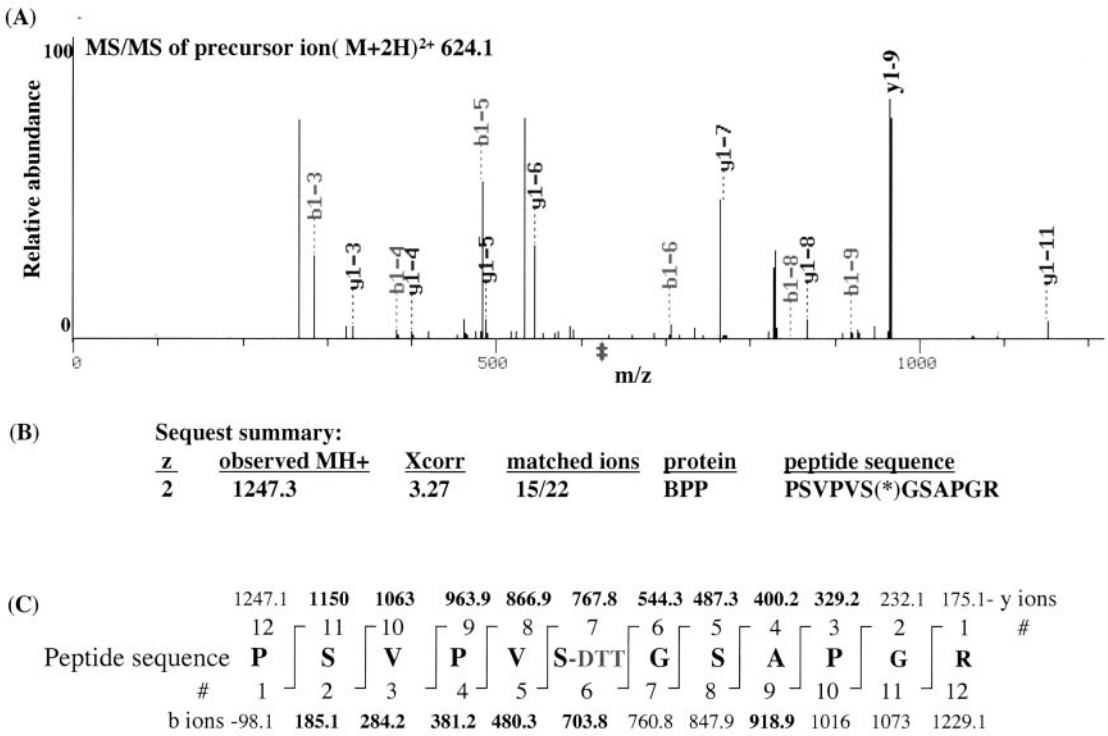


FIG. 5. **DTT replacement of *O*-GlcNAc through BEMAD is stable during tandem mass spectrometry, allowing for identification of the peptide and the DTT-modified residue.** BEMAD was performed on the peptide PSVPVS(*O*-GlcNAc)GSAPGR, and the sample was analyzed by nanospray LC-MS/MS. *A*, MS/MS spectrum from collision-induced dissociation of a precursor ion selected at 624.1 $[M + 2H]^{2+}$. All theoretical b and y ions are indicated by *dashed lines*. *B*, interpretation of MS/MS data in *A* by Turbosequest search against the Owl data base allowing for addition of 136.2 daltons to serine (*) or threonine (#) correctly identifies the peptide PSVPVS(DTT)GSAPGR. *C*, b and y ion fragments correctly interpreted are shown in *bold*. Both the b and y ions ending at the DTT-modified serine are present, making assignment of the site of modification unambiguous.

phosphorylate peptides), and subsequent BEMAD modification. Alkaline phosphatase treatment was successful based on internal controls (data not shown and Figs. 7–9). 5% of the peptide mixture was analyzed directly via LC-MS/MS. The remaining peptides were purified over a thiol column and specifically eluted with free DTT. After clean-up on reverse-phase C18 spin columns, the peptides were analyzed by LC-MS/MS, and the resulting fragmentation data were analyzed, allowing for modification of serine and threonine by 136.2 daltons. Prior to thiol chromatography, eight unmodified Synapsin I peptides and one DTT-modified Synapsin I peptide were identified (Table II). After thiol column enrichment, each of the top nine peptides reported were from Synapsin I and contained a DTT-modified residue, indicating that the thiol column facilitated the identification of these low abundance peptides of interest. After thiol column enrichment, the best hit reported by Turbosequest based on MS/MS of the doubly charged precursor $[M + 2H]^{2+}$ at *m/z* 1095.1 was a Synapsin I peptide DTT-modified at a site previously determined to be *O*-GlcNAc-modified (Table II and Fig. 6, *B* and *C*). Two additional previously mapped *O*-GlcNAc sites (A(S*)TA and QQ(S*)A) were also identified after thiol column enrichment. However, after manual inspection of the MS/MS, the spectra for one these peptides (LPSPTAAPQQ(S*)ASQATPMTQGQGR) was not of sufficient quality that we would have reported modification at this site independently. Additionally, high quality MS/MS data indicated DTT modification of two sites not previously known to be *O*-GlcNAc-modified (LS(S*)T and G(S*)HS) (Table II). Based on the original Synapsin I site mapping data, at least two sites of *O*-GlcNAc modification on Synapsin I had not been identified (44). In addition to the previously known site of modification on the peptide Q(T#)TAAAAATFSEQVGGGSGGAGR, two novel sites of modification on the same peptide were reported by Turbosequest after thiol column enrichment. Ions leading to the identification of peptides containing these three distinct sites eluted at different times in the gradient for LC-MS/MS. Upon visual inspection of the MS/MS spectra, one of these sites (GG(S*)GG) was not of sufficient quality to unambiguously identify it as a site of modification. However, MS/MS spectra of the other site (TF(S*)EQ) unambiguously identifies it as a site of modification. Of the seven previously mapped *O*-GlcNAc sites on Synapsin I, three of these were identified using BEMAD. In addition, three novel sites were mapped on this protein. Whether these differences in mapping reflect bias in the two-site mapping methods or are from differences result-

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

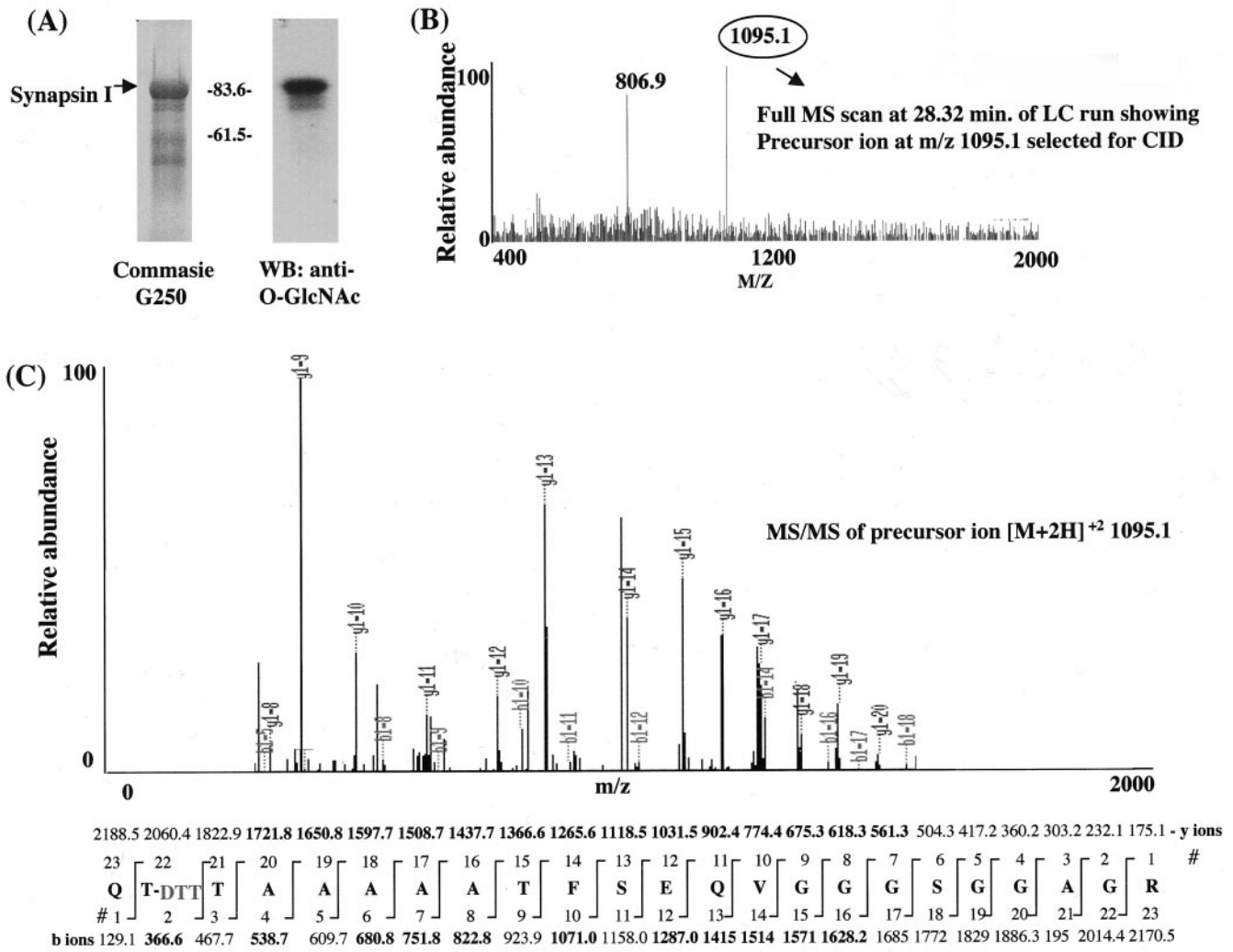
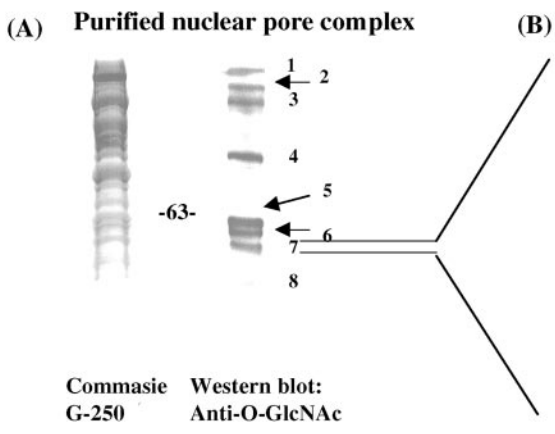


FIG. 6. **BEMAD on purified rat brain Synapsin I identifies previously known sites of *O*-GlcNAc modification.** *A*, 20 μg of purified rat brain Synapsin I separated by SDS-PAGE was detected by Coomassie G250 staining or anti-*O*-GlcNAc 110.6 Western blotting. BEMAD was performed on in-gel tryptic digests of Synapsin that had been alkaline phosphatase-treated, and the sample was analyzed by nanospray LC-MS/MS. An MS spectrum from this analysis showing a precursor ion selected for collision-induced dissociation (*CID*) (*B*) and the resulting MS/MS (*C*) leading to identification of a DTT-modified residue in Synapsin I previously shown to be modified by *O*-GlcNAc is shown. b and y ions correctly interpreted from MS/MS by Turbosequest are shown in *bold.*

ing from various deglycosylations occurring in purification and/or storage are unclear.

Site Mapping from a NPC Preparation—Having validated the BEMAD method for mapping sites of *O*-GlcNAc modification on a known *in vivo O*-GlcNAc-modified protein, we wanted to apply this technique to identify novel *O*-GlcNAc sites. While the NPC has been shown to be heavily glycosylated, only two sites on nucleoporin p62 have been mapped (49, 51). Recently, proteomic analysis of purified NPC from rat liver has identified ~40 abundant proteins (45). A NPC preparation from rat liver nuclei was separated on two SDS-PAGE gels. One gel was stained with Coomassie G-250, and the other was Western blotted with the anti-*O*-GlcNAc 110.6 antibody (Fig. 7*A*). Band 7, corresponding to an *O*-GlcNAc-modified protein by Western blot and running between 55 and 60 kDa, was excised from the gel, reduced, alkylated, and digested with trypsin. The extracted peptides were spiked with both an *O*-GlcNAc-modified (BPP) and phosphorylated (AKT) synthetic peptide. Peptides were then treated with alkaline phosphatase to remove any phosphate sites, and half of the sample was treated with β-*N*-acetylglucosaminidase to remove *O*-GlcNAc. The resulting peptides were modified by the BEMAD method, and 95% of the sample was purified over a thiol column. We analyzed 5% of the dephosphorylated sample by LC-MS/MS before thiol chromatography and identified three proteins (Lamin A, Nup58, and the Lamin B receptor; Fig. 7*B*). The dephosphorylated form of the spiked phospho-AKT peptide was identified (Figs. 7*B* and 8*E*), and LC-MS/MS of the β-*N*-acetylglucosaminidase-treated sample prior to thiol enrichment led to identification of the deglyco-

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019



(B) *Sequest results of LC-MS/MS of band 7 tryptic peptides after BEMAD*

BEFORE THIOL CHROMATOGRAPHY

Protein Identified	peptides sequenced&	matched ions	Xcorr
Lamin A	LQEKEDLQELNDR	21/24	4.62
	LEAALGEAK	10/16	1.87
Nucleoporin p58	LKLETAQELK	21/36	2.98
	QVQEEISR	13/14	2.55
Lamin B receptor	GPVPLGTFQVTTPQR	20/28	3.28
	VVEGTPLVDGR	17/20	3.09
AKT	KHFPQFSYSAS	12/20	2.35

AFTER THIOL CROMATOGRAPHY

Protein Identified	peptides sequenced&	matched ions	Xcorr
Lamin B receptor	SPGRAPKG(S*)R	13/18	2.04
Nucleoporin p58	ALKDENLPPVI(C#)QDVENLQK	30/76	3.53
BPP	PSVPV(S*)SAPGR	14/22	3.17

FIG. 7. **BEMAD identification of a novel *O*-GlcNAc-modified site on the Lamin B receptor at serine 96.** *A*, 200 units of a purified NPC fraction from rat liver were separated by SDS-PAGE and stained with Coomassie G250 or immunoblotted with the anti-*O*-GlcNAc 110.6 antibody. Eight major *O*-GlcNAc-immunoreactive bands are indicated. Tryptic peptides extracted from a Coomassie G250 gel piece corresponding to *O*-GlcNAc-immunoreactive band 7 were spiked with 100 pmol of phospho-AKT peptide and *O*-GlcNAc-modified BPP peptide. The sample was left untreated or treated with alkaline phosphatase and/or β-*N*-acetylglucosaminidase, and BEMAD was then performed followed by nanospray LC-MS/MS analysis. *B*, a table showing results of band 7 LC-MS/MS before and after thiol column enrichment. &, only the top two peptides from each protein identified are shown. The search allowed for addition to threonine (*) or serine (*) of 136.2 daltons and addition to cysteine (#) of 120.2 daltons, indicating DTT modification.

TABLE II

BEMAD identifies previously mapped sites of O-GlcNAc on Synapsin I

Rat brain Synapsin I tryptic peptides identified from LC-MS/MS after BEMAD. Turbosequest software was used to search MS/MS spectrums obtained on an LCQ ion trap mass spectrometer against a non-redundant fasta data base containing over 500,000 protein entries. Shown are the top ranked peptides derived from analysis before and after enrichment by thiol chromatography. # or * following serine or threonine denotes mass addition at that site of 136.2 daltons, indicating modification by DTT. Underlined peptides contain DTT-modified sites previously identified as *O*-GlcNAc modified. z, charge state of ion; Xcorr, sequest confidence score for peptide identification.

z	Predicted MH^+	Observed MH^+	Xcorr	Matched ions	Protein	Peptide sequence
			Prior to enrichment by thiol chromatography			
2	1726.5	1725.9	3.78	21/30	Syn I rat	TNTGSAMLEQIAMSDR
3	1562.9	1562.7	3.15	29/60	Syn I rat	GSHSQTPSPGALPLGR
2	1053.7	1053.2	3.12	14/20	Syn I rat	QASISGPAPPK
2	2151.2	2150.4	3.03	15/34	Syn I rat	LGTEEFPLIDQTFYPNHK
2	816.8	817.0	2.90	7/12	Syn I rat	MTQALPR
2	879.7	878.9	2.82	11/14	Syn I rat	TSVSGNWK
1	924.7	925.0	2.60	8/14	Syn I rat	IHGEIDIK
2	2189.0	2188.5	2.56	16/44	Syn I rat	QTTAAAAATFSEQVGGG(S*)GGAGR
3	3223.5	3223.1	2.41	20/120	Syn I rat	QGPPLQQRPPPQGQQHLSGLGPPAGSPLPQR
			After enrichment by thiol chromatography			
2	2189.0	2188.5	4.05	25/44	Syn I rat	Q(T#)TAAAAATFSEQVGGGSGGAGR
2	1699.3	1698.3	3.10	22/30	Syn I rat	G(S*)HSQTPSPGALPLGR
2	1591.3	1590.3	2.50	12/24	Syn I rat	EMLS(S*)TTYPVVVK
1	1590.3	1590.3	2.42	14/24	Syn I rat	EMLS(S*)TTYPVVVK
2	2189.0	2188.5	2.49	14/44	Syn I rat	QTTAAAAATF(S*)EQVGGGSGGAGR
2	2972.1	2972.0	2.41	26/64	Syn I rat	A(S*)TAAPVASPAAPSPGSSGGGGFFSSLSNAVK
2	2189.0	2188.4	1.97	14/44	Syn I rat	QTTAAAAATFSEQVGGG(S*)GGAGR
2	2547.2	2546.2	1.88	12/46	Syn I rat	LPSPTAAPQQ(S*)ASQATPMTQGQGR

sylated form of the spiked *O*-GlcNAc-modified peptide (Fig. 8*C*), confirming that dephosphorylation and deglycosylation treatments were successful (Fig. 8). The purified thiol-containing peptides were subjected to LC-MS/MS, and the data were analyzed using Turbosequest software and manual inspection allowing for serines and threonines to be modified by 136.2 daltons. We observed several MS/MS spectra that appeared to be of high quality but could not be matched to any peptides in the non-redundant data base. Since β-elimination of alkylated cysteines has been reported (39), we searched the MS/MS files of the thiol-enriched sample allowing for modification of cysteines by 120.2 daltons (mass of DTT addition). A DTT-modified cysteine-containing peptide belonging to Nup58 was identified (Fig. 7*B*). Additionally, from the thiol-

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

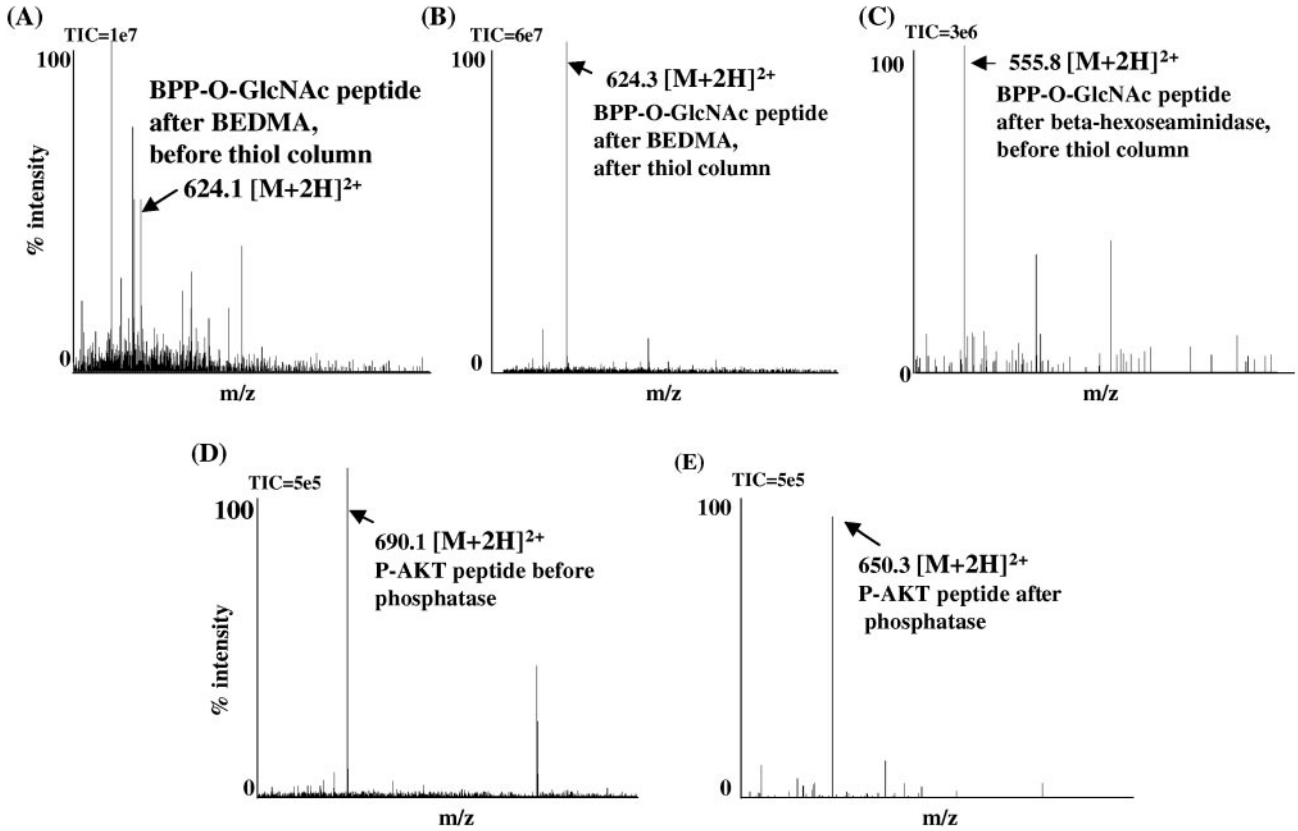


FIG. 8. **Controls for *O*-GlcNAc specificity in the identification of sites by BEMAD.** Shown are band 7 LC-MS spectra showing ions representing internal controls for replacement of *O*-GlcNAc with DTT in band 7 (Fig. 7). *A*, BPP DTT-modified peptide before thiol column. *B*, BPP DTT-modified peptide enriched by thiol chromatography. *C*, BPP deglycosylated peptide that did not become modified by DTT after β-*N*-acetylglucosaminidase treatment. *D*, phosphorylated AKT peptide before phosphatase treatment. *E*, dephosphorylated AKT peptide that did not become modified by DTT after alkaline phosphatase treatment. In each case the expected mass shift of the $[M + 2H]^{2+}$ is observed. *TIC*, total ion current; *P-AKT*, phospho-AKT.

enriched sample, we identified a DTT-modified serine in a peptide from the Lamin B receptor (Fig. 7*B*). This peptide could not be observed in the β-*N*-acetylglucosaminidase-treated sample. The spiked phospho-AKT peptide was not found in its DTT form after phosphatase treatment, BEMAD, and thiol enrichment, while the spiked *O*-GlcNAc peptide was converted, enriched, and correctly identified only in the absence of β-*N*-acetylglucosaminidase treatment (Figs. 7*B* and 8, *A–C*). The *O*-GlcNAc-modified serine (Ser-96) is in the amino-terminal nucleoplasmic tail of the Lamin B receptor in a domain known to interact with chromatin (52).

Finally, we performed an in-solution digest on 400 units of the NPC preparation. Following buffer exchange, to remove protease inhibitors, the resulting soluble proteins were trypsin-digested and oxidized with performic acid to avoid DTT modification of cysteines during subsequent BEMAD treatment. The resulting peptides were spiked with 100 pmol of phospho-AKT peptide, and then the mixture was treated with alkaline phosphatase. The sample was split, and half was subjected to β-*N*-acetylglucosaminidase treatment. Both samples were then modified by the BEMAD procedure. 5% of the resulting peptides were analyzed by LC-MS/MS directly. The remaining peptides were subjected to thiol chromatography, and the DTT-eluted peptides were analyzed by LC-MS/MS. Our coverage from in-solution digest of the nuclear pore preparation was very low due to the small amount of protein that stayed in solution during the spin filtering; however, we did identify the dephosphorylated AKT peptide as well as a few proteins known to be associated with the NPC including GP210, Nup155, and Ran-GAP1 (Fig. 9*A*, Ref. 45). We were able to map a novel site of *O*-GlcNAc modification on Nup155 at Ser-525 (Fig. 9), and we were not able to find the DTT-modified AKT peptide. Thiol enrichment of the sample treated with β-*N*-acetylglucosaminidase prior to BEMAD did not lead to identification of the DTT-modified Nup155 peptide. Taken together these data indicate that serine 525 of Nup155 is *O*-GlcNAc-modified *in vivo* and that the BEMAD method can be used to map *O*-GlcNAc sites independently of phosphorylation. Improvements to NPC sample preparation and treatment should make it possible to map other sites of *O*-GlcNAc modification and are currently being pursued.

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

(A)

BEMAD of soluble nuclear pore complex fraction.

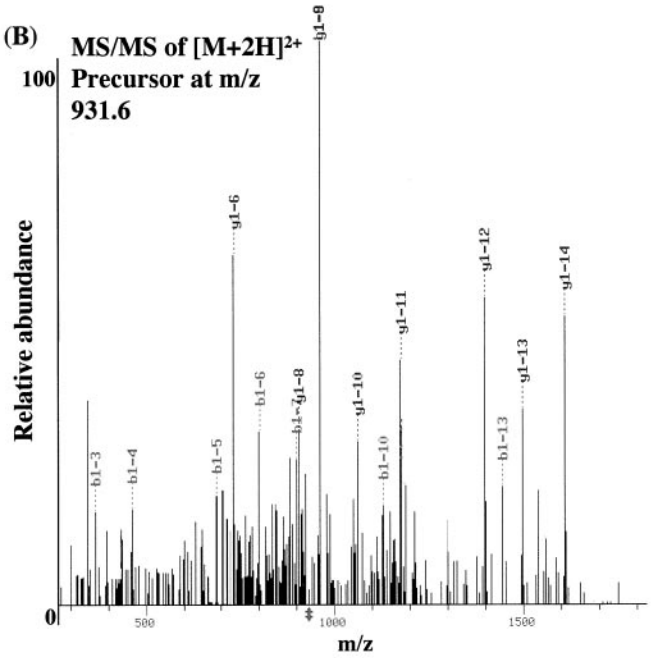
BEFORE THIOL CHROMATOGRAPHY

Protein Identified	peptides sequenced&	matched ions	Xcorr
GP210	NPLLDLGAYDQQGR	18/26	3.75
	DTEANGFSDSHNALR	12/28	2.32
NUP155	HLLVSNVGGDGEEIER	22/30	4.08
	KFHEAQLSEK	17/36	2.34
RAN-GAP1	VSVLIVQQTDTSDPEK	20/30	3.67
	VINLNDNTFTEK	17/22	2.75
AKT	KHFPQFSYSAS	11/20	2.10

AFTER THIOL CROMATOGRAPHY

Protein Identified	peptides sequenced&	matched ions	Xcorr
NUP155	HLLV(S*)NVGGDGEEIER	17/30	3.44

(B)



(C)

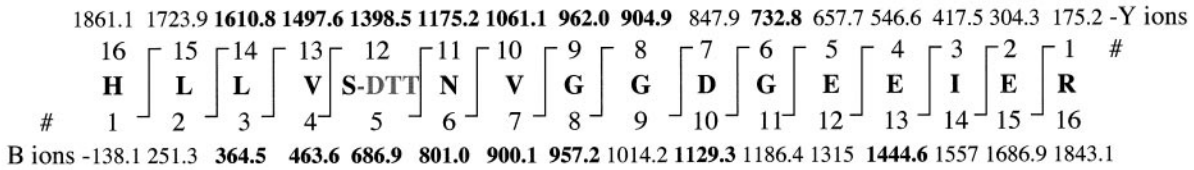


FIG. 9. **BEMAD performed on a soluble subset of NPC proteins identifies a novel *O*-GlcNAc site on Nup155 at serine 525.** *A*, in-solution digest after BEMAD modification identifies peptides from the NPC. Before thiol purification, GP210, Nup155, and Ran-GAP1 were identified (&, only the top two peptides from each are shown). The dephosphorylated AKT peptide was also identified before thiol chromatography and was not identified after chromatography. After thiol chromatography, serine 525 of Nup155 was identified as being modified by *O*-GlcNAc. *B*, MS/MS scan following collision-induced dissociation of the modified Nup155 peptide. *C*, the resulting DTT-modified sequence from Nup155 with b and y ion fragments correctly interpreted shown in *bold*.

DISCUSSION

Identification of types and sites of post-translational modification is an important step in performing functional proteomics. Site mapping allows the function of post-translational modifications on a specific protein to be tested by site-directed mutagenesis. Additionally, generation of modification site-specific antibodies, such as those recognizing the phosphorylated or *O*-GlcNAc-modified form of Thr-58 in c-Myc (53), are useful for studying dynamic changes of specific post-translational modifications. Given the widespread occurrence of *O*-GlcNAc, the list of identified *O*-GlcNAc-modified proteins is still quite small. Based upon their specific elution from an anti-*O*-GlcNAc antibody column, we identified 25 candidate *O*-GlcNAc-modified proteins as well as five previously identified *O*-GlcNAc proteins in a single LC-MS/MS run (Table I). This enrichment strategy, similar to strategies that have been used for identifying phosphorylated proteins (26), however, did not allow for us to map sites, perform quantitative mass spectrometry, or distinguish between *O*-GlcNAc-modified proteins *versus* proteins tightly associating with *O*-GlcNAc-modified proteins. Thus, we sought a better method for identifying *O*-GlcNAc-modified proteins and their sites of modification.

While hundreds, if not thousands, of phosphorylation sites have been mapped, less than 50 *O*-GlcNAc sites have been published. β-Elimination/Michael addition strategies with biotinylated nucleophiles or ethanedithiol followed by biotin tagging has been used to successfully map phosphorylation sites (34–36, 38). However, based on our results, we predict that the protocols used in those studies would modify and map *O*-GlcNAc sites as well. We have adapted the β-elimination/Michael addition strategy to specifically map *O*-GlcNAc *versus* *O*-phosphorylation sites. We have verified the BEMAD method using synthetic peptides and, importantly, *in vivo* modified biological samples. The mapping of *O*-GlcNAc sites on proteins purified from rat brain (Synapsin I) and rat liver (Nup155 and the Lamin B receptor) using BEMAD highlights the need for researchers applying β-elimination/Michael addition strategies to discriminate between *O*-GlcNAc and *O*-phosphorylation. Several steps in our method provide ex-

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

perimental tools to map *O*-GlcNAc *versus* *O*-phosphorylation sites. We and others have demonstrated the use of modification-specific antibodies for enrichment of proteins containing the post-translational modification of interest (Fig. 1, Table I, and Refs. 6, 10, 26, 27, and 35). We describe mild β-elimination conditions that preferentially eliminate *O*-GlcNAc. It should be noted that some *O*-GlcNAc-modified residues more resistant to β-elimination (*e.g.* *O*-GlcNAc-Thr followed by a Pro) are not likely to be detected by this method as they would be only partially converted. For specificity of mapping *O*-GlcNAc sites, peptides can be dephosphorylated enzymatically prior to BEMAD, and specific loss of mapped *O*-GlcNAc sites due to β-*N*-acetylglucosaminidase treatment prior to modification adds another level of specificity. A similar strategy could be applied to mapping phosphorylation sites using phospho-specific antibodies for enrichment and β-*N*-acetylglucosaminidase treatment to remove *O*-GlcNAc prior to BEMAD treatment.

In mapping novel *in vivo* *O*-GlcNAc sites on the Lamin B receptor and Nup155, we spiked samples with known *O*-GlcNAc-modified and *O*-phosphorylated peptides as internal controls for specificity of the BEMAD method. Lamin B is an integral inner nuclear membrane protein that binds lamins and plays a role in nuclear envelope interactions with chromatin (52). The *O*-GlcNAc-modified serine 96, which is located in the nucleoplasmic amino-terminal tail, may play a role in protein-protein interactions and/or NPC assembly and function. Ser-525 was also found to be *O*-GlcNAc-modified on Nup155, a recently identified member of the mammalian NPC whose function remains to be elucidated (54). We are currently using the BEMAD method to perform a more global analysis of sites of *O*-GlcNAc modification on the NPC and other subproteomes.

The utility of the BEMAD method is demonstrated by our ability to use automated data-dependent scanning for MS/MS and Turbosequest searching against non-redundant databases to identify sites of *O*-GlcNAc modification. Thus, the method has potential for high throughput and automation. Since deuterated DTT is commercially available, the next logical step in this procedure is performing comparative quantitative mass spectrometry (50) by differential isotopic DTT labeling. This will allow us to determine what sites on proteins are changing in response to various cellular treatments and may provide insight into the regulatory role of *O*-GlcNAc at specific sites. The BEMAD method also can be used as an alternative method to isotope-coded affinity tags (50) since alkylated cysteines are susceptible to β-elimination (39). In fact, from band 7 of the NPC preparation we enriched and sequenced a peptide containing a DTT-derivatized cysteine (+120.2 daltons, Fig. 7). Also, further work has demonstrated that alkylated cysteines are readily modified by the BEMAD procedure (data not shown). Thus, it appears that this method can enrich for both the cysteine-containing peptides as well as serines and threonines that are post-translationally modified. Differential isotopic labeling with DTT (light) and deuterated DTT (heavy) could be used to quantify changes in both protein level (by labeling of cysteines) as well as at specific sites of post-translational modification on serines and threonines in the same LC-MS/MS experiment and thus reveal dynamic post-translational modifications relative to protein levels. This type of methodology is currently being developed by our group.

While phosphorylation is a well established and recognized modification in altering protein function, *O*-GlcNAc is now beginning to also emerge as an important intracellular post-translational modification (9). Mapping of *O*-GlcNAc sites should help to facilitate our understanding of the role of this enigmatic modification.

Acknowledgment—We thank members of the Hart laboratory for helpful discussions and critical reading of the manuscript.

* This work was supported by National Institutes of Health Grants CA43486 and DK61671 (to G. W. H.), American Cancer Society Grant RSG-01-064-01-CSM (to M. J. M.), and National Research Service Award Fellowships CA83261 (to L. W.) and GM20528 (to K. V.). MALDI analysis was performed in the Johns Hopkins School of Medicine, Applied Biosystems Mass Spectrometry facility, which is funded in part by the Johns Hopkins Fund for Medical Discovery, the Institute for Cell Engineering, and National Center for Research Resources Shared Instrumentation Grant 1S10-RR14702 (to G. W. H.). Under a licensing agreement between Covance Research Products, Hoffman LaRoche, and The Johns Hopkins University, Dr. Hart receives a share of royalty received by the university on sales of the CTD 110.6 antibody. The terms of this arrangement are being managed by The Johns Hopkins University in accordance with its conflict of interest policies. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Both authors contributed equally to this work.

‖ To whom reprint requests should be addressed. E-mail: gwhart@jhmi.edu.

REFERENCES

1. Haynes, P. A., Gygi, S. P., Figeys, D., and Aebersold, R. (1998) Proteome analysis: biological assay or data archive? *Electrophoresis* **19,** 1862–1871
2. Uy, R., and Wold, F. (1977) Posttranslational covalent modification of proteins. *Science* **198,** 890–896
3. Kouzarides, T. (2000) Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* **19,** 1176–1179
4. McBride, A. E., and Silver, P. A. (2001) State of the arg: protein methylation at arginine comes of age. *Cell* **106,** 5–8
5. Hart, G. W. (1997) Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu. Rev. Biochem.* **66,** 315–335
6. Wells, L., Vosseller, K., and Hart, G. W. (2001) Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science* **291,** 2376–2378
7. Hanover, J. A. (2001) Glycan-dependent signaling: O-linked N-acetylglucosamine. *FASEB J.* **15,** 1865–1876
8. Comer, F. I., and Hart, G. W. (2000) *O*-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between *O*-GlcNAc and *O*-phosphate. *J. Biol. Chem.* **275,** 29179–29182
9. Zachara, N. E., and Hart, G. W. (2002) The emerging significance of O-GlcNAc in cellular regulation. *Chem. Rev.* **102,** 431–438

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

10. Vosseller, K., Wells, L., and Hart, G. W. (2001) Nucleocytoplasmic O-glycosylation: O-GlcNAc and functional proteomics. *Biochimie* (*Paris*) **83,** 575–581
11. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique *O*-GlcNAc transferase with multiple tetratricopeptide repeats. *J. Biol. Chem.* **272,** 9308–9315
12. Lubas, W. A., Frank, D. W., Krause, M., and Hanover, J. A. (1997) *O*-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J. Biol. Chem.* **272,** 9316–9324
13. Kreppel, L. K., and Hart, G. W. (1999) Regulation of a cytosolic and nuclear *O*-GlcNAc transferase. Role of the tetratricopeptide repeats. *J. Biol. Chem.* **274,** 32015–32022
14. Lubas, W. A., and Hanover, J. A. (2000) Functional expression of *O*-linked GlcNAc transferase. Domain structure and substrate specificity. *J. Biol. Chem.* **275,** 10983–10988
15. Gao, Y., Wells, L., Comer, F. I., Parker, G. J., and Hart, G. W. (2001) Dynamic *O*-glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a neutral, cytosolic β-*N*-acetylglucosaminidase from human brain. *J. Biol. Chem.* **276,** 9838–9845
16. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) Dynamic *O*-glycosylation of nuclear and cytosolic proteins. Further characterization of the nucleocytoplasmic β-*N*-acetylglucosaminidase, *O*-GlcNAcase. *J. Biol. Chem.* **277,** 1755–1761
17. Hart, G. W., Greis, K. D., Dong, L. Y., Blomberg, M. A., Chou, T. Y., Jiang, M. S., Roquemore, E. P., Snow, D. M., Kreppel, L. K., Cole, R. N., *et al.* (1995) O-linked N-acetylglucosamine: the "yin-yang" of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation. *Adv. Exp. Med. Biol.* **376,** 115–123
18. Yang, X., Zhang, F., and Kudlow, J. E. (2002) Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A. Coupling protein O-GlcNAcylation to transcriptional repression. *Cell* **110,** 69–80
19. Vosseller, K., Wells, L., Lane, M. D., and Hart, G. W. (2002) Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 5313–5318
20. McClain, D. A., Lubas, W. A., Cooksey, R. C., Hazel, M., Parker, G. J., Love, D. C., and Hanover, J. A. (2002) Altered glycan-dependent signaling induces insulin resistance and hyperleptinemia. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 10695–10699
21. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5735–5739
22. Roquemore, E. P., Chou, T. Y., and Hart, G. W. (1994) Detection of O-linked N-acetylglucosamine (O-GlcNAc) on cytoplasmic and nuclear proteins. *Methods Enzymol.* **230,** 443–460
23. Greis, K. D., Hayes, B. K., Comer, F. I., Kirk, M., Barnes, S., Lowary, T. L., and Hart, G. W. (1996) Selective detection and site-analysis of O-GlcNAc-modified glycopeptides by β-elimination and tandem electrospray mass spectrometry. *Anal. Biochem.* **234,** 38–49
24. Haynes, P. A., and Aebersold, R. (2000) Simultaneous detection and identification of O-GlcNAc-modified glycoproteins using liquid chromatography-tandem mass spectrometry. *Anal. Chem.* **72,** 5402–5410
25. Chalkley, R. J., and Burlingame, A. L. (2001) Identification of GlcNAcylation sites of peptides and α-crystallin using Q-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **12,** 1106–1113
26. Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., and Pandey, A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol.* **20,** 261–268
27. Pandey, A., Andersen, J. S., and Mann, M. (2000) Use of mass spectrometry to study signaling pathways. *Science's STKE* http://stke.sciencemag.org/cgi/content/full/OC_sigtrans;2000/37/pl1
28. Conrads, T. P., Issaq, H. J., and Veenstra, T. D. (2002) New tools for quantitative phosphoproteome analysis. *Biochem. Biophys. Res. Commun.* **290,** 885–890
29. Loughrey Chen, S., Huddleston, M. J., Shou, W., Deshaies, R. J., Annan, R. S., and Carr, S. A. (2002) Mass spectrometry-based methods for phosphorylation site mapping of hyperphosphorylated proteins applied to Net1, a regulator of exit from mitosis in yeast. *Mol. Cell. Proteomics* **1,** 186–196
30. Zappacosta, F., Huddleston, M. J., Karcher, R. L., Gelfand, V. I., Carr, S. A., and Annan, R. S. (2002) Improved sensitivity for phosphopeptide mapping using capillary column HPLC and microionspray mass spectrometry: comparative phosphorylation site mapping from gel-derived proteins. *Anal. Chem.* **74,** 3221–3231
31. Li, J., LeRiche, T., Tremblay, T. L., Wang, C., Bonneil, E., Harrison, D. J., and Thibault, P. (2002) Application of microfluidic devices to proteomics research. Identification of trace-level protein digests and affinity capture of target peptides. *Mol. Cell. Proteomics* **1,** 157–168
32. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **20,** 301–305
33. Grønborg, M., Kristiansen, T. Z., Stensballe, A., Andersen, J. S., Ohara, O., Mann, M., Jensen, O. E., and Pandey, A. (2002) A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies. Identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol. Cell. Proteomics* **1,** 517–527
34. Adamczyk, M., Gebler, J. C., and Wu, J. (2001) Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry. *Rapid Commun. Mass Spectrom.* **15,** 1481–1488
35. Goshe, M. B., Conrads, T. P., Panisko, E. A., Angell, N. H., Veenstra, T. D., and Smith, R. D. (2001) Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.* **73,** 2578–2586
36. Weckwerth, W., Willmitzer, L., and Fiehn, O. (2000) Comparative quantification and identification of phosphoproteins using stable isotope labeling and liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **14,** 1677–1681
37. Zhou, H., Watts, J. D., and Aebersold, R. (2001) A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19,** 375–378
38. Oda, Y., Nagasu, T., and Chait, B. T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **19,** 379–382
39. Meyer, H. E., Hoffmann-Posorske, E., and Heilmeyer, L. M., Jr. (1991) Determination and location of phosphoserine in proteins and peptides by conversion to S-ethylcysteine. *Methods Enzymol.* **201,** 169–185
40. Torres, C. R., and Hart, G. W. (1984) Topography and polypeptide distribution of terminal *N*-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for *O*-linked GlcNAc. *J. Biol. Chem.* **259,** 3308–3317
41. Czeszak, X., Ricart, G., Tetaert, D., Michalski, J. C., and Lemoine, J. (2002) Identification of substituted sites on MUC5AC mucin motif peptides after enzymatic O-glycosylation combining β-elimination and fixed-charge derivatization. *Rapid Commun. Mass Spectrom.* **16,** 27–34
42. Comer, F. I., Vosseller, K., Wells, L., Accavitti, M. A., and Hart, G. W. (2001) Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. *Anal. Biochem.* **293,** 169–177
43. Greis, K. D., Gibson, W., and Hart, G. W. (1994) Site-specific glycosylation of the human cytomegalovirus tegument basic phosphoprotein (UL32) at serine 921 and serine 952. *J. Virol.* **68,** 8339–8349
44. Cole, R. N., and Hart, G. W. (1999) Glycosylation sites flank phosphorylation sites on synapsin I: O-linked N-acetylglucosamine residues are localized within domains mediating synapsin I interactions. *J. Neurochem.* **73,** 418–428
45. Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., and Matunis, M. J. (2002) Proteomic analysis of the mammalian nuclear pore complex. *J. Cell Biol.* **158,** 915–927
46. Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S., and Mische, S. M. (1999) Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* **20,** 601–605
47. Arnott, D., Shabanowitz, J., and Hunt, D. F. (1993) Mass spectrometry of proteins and peptides: sensitive and accurate mass measurement and sequence analysis. *Clin. Chem.* **39,** 2005–2010
48. Gatlin, C. L., Eng, J. K., Cross, S. T., Detter, J. C., and Yates, J. R., III (2000) Automated identification of amino acid sequence variations in proteins by HPLC/microspray tandem mass spectrometry. *Anal. Chem.* **72,** 757–763

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019

49. Lubas, W. A., Smith, M., Starr, C. M., and Hanover, J. A. (1995) Analysis of nuclear pore protein p62 glycosylation. *Biochemistry* **34,** 1686–1694
50. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999
51. D'Onofrio, M., Starr, C. M., Park, M. K., Holt, G. D., Haltiwanger, R. S., Hart, G. W., and Hanover, J. A. (1988) Partial cDNA sequence encoding a nuclear pore protein modified by O-linked N-acetylglucosamine. *Proc. Natl. Acad. Sci. U. S. A.* **85,** 9595–9599
52. Takano, M., Takeuchi, M., Ito, H., Furukawa, K., Sugimoto, K., Omata, S., and Horigome, T. (2002) The binding of lamin B receptor to chromatin is regulated by phosphorylation in the RS region. *Eur. J. Biochem.* **269,** 943–953
53. Kamemura, K., Hayes, B. K., Comer, F. I., and Hart, G. W. (2002) Dynamic interplay between *O*-glycosylation and *O*-phosphorylation of nucleocytoplasmic proteins. Alternative glycosylation/phosphorylation of Thr-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens. *J. Biol. Chem.* **277,** 19229–19235
54. Zhang, X., Yang, H., Corydon, M. J., Pedersen, S., Korenberg, J. R., Chen, X. N., Laporte, J., Gregersen, N., Niebuhr, E., Liu, G., and Bolund, L. (1999) Localization of a human nucleoporin 155 gene (NUP155) to the 5p13 region and cloning of its cDNA. *Genomics* **57,** 144–151

Downloaded from http://www.mcponline.org/ by guest on April 22, 2019