MCP Papers in Press. Published on September 26, 2002 as Manuscript M200048-MCP200

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 *§

Johns Hopkins University School of Medicine, Department of Biological Chemistry* and School of Hygiene and Public Health, Department of Biochemistry and Molecular Biology†, Baltimore, MD 21218

Running Title: Mapping O-GlcNAc Sites

# LW and KV contributed equally to this work

§ To whom reprint requests should be addressed: gwhart@jhmi.edu

ABSTRACT

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 DTT as the nucleophile. The modified peptides can be efficiently enriched by affinity chromatography, and the sites 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 (NPC) 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.

The abbreviations used are: O-linked β-N-acetylglucosamine (O-GlcNAc), β-elimination followed by Michael addition with DTT (BEMAD), Nuclear Pore Complex (NPC), O-GlcNAc Transferase (OGT), Immobolized Metal Affinity Chromatography (IMAC), Collision-Induced Dissociation (CID), Isotope-Coded Affinity Tags (ICAT), Mass to Charge Ratio (m/z), Biotin Pentylamine (BAP), Reverse-phase Liquid Chromatography to Nanospray Tandem Mass Spectrometry (LC-MS/MS)

INTRODUCTION

The rapid identification of proteins by mass spectrometry has become commonplace in the post-genomic 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 recognized the abundance of naturally occurring modified forms of the genetically encoded twenty-one amino acids (2). In addition to phosphorylation, a variety of post-translational modifications, including acetylation (3), methylation (4), and O-linked β-N-acetylglucosamine (O-GlcNAc) (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, 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, OGT) 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, OGT 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 OGT 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.

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 enriched galactose-tagged O-GlcNAc modified proteins, but still found that the modification was labile (24). Greis and coworkers performed MS/MS analysis to map sites on peptides by identifying dehydroamino acids resulting from prior β-elimination of O-GlcNAc (23). 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 substoichometric and often labile, making their identification difficult. Methods development for identification of phosphorylation sites has attempted to address these issues (26-29). Techniques include precursor ion scanning (30), enrichment of phosphopeptides via IMAC (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 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 (Clelands 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, enriched, and the site of modification 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. AntiIgM-Agarose, dimethyl pimelimidate, and DTT were purchased from Sigma . Biotin pentylamine (BAP) was obtained from Pierce. Activated thiol sepharose columns and monoavidin columns were from Amersham Pharmacia 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 (P-S-V-P-V-S(O-GlcNAc)-G-S-A-P-G-R) (43). The phosphopeptide (K-H-F-P-Q-F-S(P)-Y-S-A-S) we used was derived from AKT and was obtained commercially (Upstate Biotechnology).

Rat Brain Extract Preparation: 1 gram of rat brain (Pel-Freeze) per 7ml of buffer (15 mM Tris,HCl pH 7.5, 150 mM NaCl, 1mM EDTA, 1mM DTT, protease inhibitor cocktail, and 10 uM PUGNAc) was mechanically lysed with a polytron (2 X 30 sec. pulse) and clarified at 35,000 X G for 30 min. and the supernatant was passed through a 1uM glass fiber filter and stored at -80C until used.

Immunopurification of rat brain extract: Lysates were pre-cleared for 1 hour at 4C with anti-IgM agarose. Samples were then batch bound overnight at 4C 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 5 times with 10 column volumes of RIPA (1% NP-40, 0.5% Deoxycholate, 0.1% SDS, in TBS (15mM Tris, HCl pH 7.5, 150mM NaCl)), 5 column volume of TBS, and then eluted with 1M 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 previously

described (44,45).

SDS-PAGE, Coomassie Staining, and Western blotting: SDS-PAGE was performed under reducing conditions on precast 10% Criterion gels (Biorad). Coomassie G-250 staining and Western blotting of 1D gels with 110.6 antibody was as previously described (42).

In-Gel Digestion: Reduction, alkylation, and digestion were performed essentially as previously described (46). Briefly, Coomassie G-250 stained bands were excised, dehydrated with acetonitrile and reswelled in 40mM ammonium bicarbonate. This was repeated and then gel pieces were reduced with 10mM DTT for one hour at 56C and then carboxyamidomethylated with 55mM iodoacetamide in the dark for 45 minutes. The gel was dehydrated and reswollen in 40mM ammonium bicarbonate with 10ng/ul trypsin (Promega) on ice for 45 minutes. After excess trypsin was replaced with 40 mM ammonium bicarbonate, digestion was allowed to proceed overnight at 37C. The peptides were extracted 3 times for 20 minutes in 5% formic acid, 50% acetonitrile and dried down in a speed-vac.

Phosphatase and N-acetylglucosaminidase Treatment: Where indicated in the results, peptides were spiked with O-GlcNAc BPP and/or phosphorylated AKT peptides and then subjected to dephosphorylation with alkaline phosphatase (1U per 10ul, Promega) for 3 hrs at 37C in the presence of added 1 mM $MgCl_2$ or β-N-acetlyglucosaminidase (1U per 20ul, New England Biolabs) for 16 hrs at 37C after acidification with TFA 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 37C 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 one hour on ice and then dried down. The

proteins were resuspended in 40 mM ammonium bicarbonate and digested overnight at 37C by the addition of 1:10 (w/w) trypsin. The peptides were dried down, resuspended in 40mM ammonium bicarbonate, 1mM magnesium chloride and then spiked with 100 pMoles of phosphoAkt peptide. 25 units of alkaline phosphatase was added and the peptides were incubated at 37C for 4 hours 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 50C for various amounts of time (see results, biological samples 2.5 hrs) and the reaction was quenched with TFA (final 1%). The peptides were cleaned up via reverse phase C18 spin columns (The Nest Group) and eluted in 0.1% TFA, 70% acetonitrile and dried in a speed-vac.

Affinity Chromatography: DTT modified peptides were purified over activated thiol sepharose (Thiol column) from Pharmacia Biotech. Resin was swelled in degassed PBS containing 1 mM EDTA (PBS/EDTA) and dried peptides suspended in same buffer were bound with 1 hour incubation in 200 ul of 50% slurry. Column was washed with 15 ml PBS/EDTA and eluted three times sequentially with 150 ul of PBS/EDTA containing 20 mM free DTT. BAP derivitized peptides were enriched by sequential cation exchange and monoavidin columns provided in ICAT kits from Applied Biosystems according to manufacturers protocol. Peptides eluted from thiol or avidin affinity columns were acidified (brought to 1% TFA) and were desalted with reverse phase C18 spin columns (eluted in 70% acetonitrile, 0.1% TFA) and dried for subsequent analysis.

MALDI Analysis: Peptide samples were resuspended in matrix (10mg/ml α-cyano-4-hydroxycinnamic acid in 0.3% TFA, 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 by 0.075 mM 5u C18 column using positive N_2 pressure, desalted with 1% acetic acid and then separated via a 75 minute linear gradient of increasing acetonitrile at a flow rate of approximately 200 nl/min directly into the source (Finnigan LCQ, (47)). In some cases following BEMAD treatment, to prevent disulfide formation, the samples were loaded under mild reducing conditions (200uM DTT). The LCQ was run in automatic mode collecting a MS scan (2 X 500ms) followed by two MS/MS scans (3 X 750ms) 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 (BAP) modified peptides, a mass increase of 136.2 (310.5) daltons 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 derivitized 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 database (downloaded 4/22/02 from NCI at Frederick) search with a Xcorr>2.5 (unless otherwise stated in the results or legend) 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 2D-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 minute run by LC-MS/MS (Fig. 1A). We also performed a control experiment using anti-IgM agarose that had not been coupled to the 110.6 antibody and subtracted the resulting non-specific proteins (actin, tubulin, serum albumin, and human keratins) from the 110.6 results. In figure 1, we show the MS profile at 35.94 minutes of the run (Fig. 1B) and the resulting MS/MS fragmentation of the m/z 542.5 peptide (Fig. 1C) that corresponds to a peptide from the small G protein Ran when analyzed (Table 1). Analysis of the resulting data identified thirty proteins (with coverage ranging from 2 to 12 sequenced peptides that matched the listed protein as the best hit in the database, Table 1). Five of these proteins had previously been shown to be modified by O-GlcNAc (the nucleoporins p54 and p62 (49), OGT (11), the heat shock protein HSP90 (S. Iyer and G.W. Hart, unpublished data) and the kinase CKII (13, 14)) suggesting the method was valid. The other 25 proteins could be subdivided into multiple functional classes (Table 1). 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. In order to stabilize sites of O-GlcNAc modification, enrich these peptides, and map the sites of modification and identify the protein, we developed the BEMAD method (Fig. 2A) followed by affinity chromatography and LC-MS/MS. Initially, we tested this method on synthetic O-GlcNAc and O-phosphate modified peptides (Fig. 2-5). We tried two different nucleophiles (DTT and BAP) for the Michael addition, both of which reacted equally well with the dehydroamino acids (Figure 2C, 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 (Figure 3) since O-glycosidic linkages are more easily eliminated than phospholinkages (23,39-41). As can be seen in figure 3, at two hours all of the serine-O-GlcNAc peptide had been converted while a significant portion of the serine-O-phosphate peptide remained unmodified after four hours 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 hours) 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. 4A) or BAP (Fig. 4B) 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 database (containing over 500,000 proteins) to identify the peptide, protein, and site of modification (Fig. 5). While 3 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 7th best hit in the non-redundant database. 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 (d10) 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. 6A). Synapsin I was subjected to alkylation, digestion, alkaline phosphatase treatment (to dephosphorylate peptides), and subsequent BEMAD modification. Alkaline phosphatase treatment was successful based on internal controls (data not shown and Fig. 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 cleanup on reverse-phase C18 spin columns, the peptides were analyzed on LC-MS/MS and the resulting fragmentation data analyzed, allowing for modification of serine and threonine by 136.2 daltons. Prior to thiol chromatography, 8 unmodified Synapsin I peptides and 1 DTT modified Synapsin I peptide were identified (Table 2). After thiol column enrichment, each of the top 9 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 2 and Fig. 6 B,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 2 sites not previously known to be O-GlcNAc modified (LS(S*)T and G(S*)HS) (Table 2). 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 3 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 7 previously mapped O-GlcNAc sites on Synapsin I, 3 of these were identified using BEMAD. In addition, 3 novel sites were mapped on this protein. Whether these differences in mapping reflect bias in the two site mapping methods or are from differences resulting from various deglycosylation 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 approximately 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 Western blotted with the anti-O-GlcNAc 110.6 antibody (Figure 7A). Band 7, corresponding to an O-GlcNAc modified protein by Western blot and running between 55-60 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 a β-N-acetylglucosamindase to remove O-GlcNAc. The resulting peptides were modified by the BEMAD method and 95% of the sample purified over a thiol column. We analyzed 5% of the dephosphorylated sample on LC-MS/MS before thiol chromatography and identified three proteins (Lamin A, Nup58, and the Lamin B

receptor, Fig. 7B). The dephosphorylated form of the spiked phospho-AKT peptide was identified (Fig 7B, 8E), and LC-MS/MS of the β-N-acetylglucosamindase treated sample prior to thiol enrichment lead to identification of the deglycosylated form of the spiked O-GlcNAc modified peptide(Fig. 8C), confirming that dephosphorylation and deglycosylation treatments were successful (Fig 8). The purified thiol containing peptides were subjected to LC-MS/MS and the data analyzed using Turbosequest software and manual inspection allowing for serine and threonines to be modified by 136.2 daltons. We observed several MS/MS spectra which appeared of high quality but could not be matched to any peptides in the non-redundant database. 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 7B). Additionally, from the thiol 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-acetylglucosamindase 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-acetylglucosamindase treatment (Fig 7B, 8A-C). The O-GlcNAc modified serine (Ser96) is in the N-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 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. 9A, 45). We were able to map a novel site of O-GlcNAc modification on Nup155 at Ser525 (Fig. 9) and we were not able to find the DTT modified AKT peptide. Thiol enrichment of the sample treated with β-N-acetylglucosamindase 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 independent 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.

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 5 previously identified O-GlcNAc proteins, in a single LC-MS/MS run (Table 1). 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 fifty 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 experimental 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 1, and (6,10,26,27,35)). 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 phoshospecific 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 which 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. BEMAD method also can be used as an alternative method to ICAT (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 derivitized cysteine (+120.2 daltons, Figure 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.

ACKNOWLEDGEMENTS

We would like to thank members of the Hart laboratory for helpful discussions and critical reading of the manuscript. We performed MALDI analysis 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 a national center for research resources shared instrumentation grant (1S10-RR14702), to GWH. 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. This work was supported by national institute of health grants (CA43486 and DK61671) to GWH, a grant from the american cancer society (RSG-01-064-01-CSM) to MJM, and national research service award fellowships to LW (CA83261) and KV(GM20528).

FIGURE LEGENDS

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 analyzed by nanospray LC-MS/MS. (B) A single full MS scan at 35.94 minutes 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 database analysis.

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, D) MALDI-TOF analysis of a synthetic O-GlcNAc modified peptide that was (B) untreated, or incubated at 50^{o} C for 2 hours in 1% TEA, 0.1% NaOH in the presence of (C) 10 mM DTT, or (D) 20 mM BAP. 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) A synthetic O-GlcNAc modified peptide or (B) a synthetic O-phosphorylated modified peptide were untreated (0 hour) or incubated at 50^{o} C for 2 or 4 hours in 1% TEA, 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.

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. The ion at 624.1 [M+2H]/2 was selected for CID (MS/MS) shown in, all theoretical b and y ions are indicated by dashed lines (A). (B) Interpretation of MS/MS data in (A) by Turbosequest search against the Owl database 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.

Fig. 6. BEMAD on Purified Rat Brain Synapsin I Identifies Previously Known Sites of O-GlcNAc Modification. (A) 20 ug 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 which 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 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 while all theoretical b and y ions are indicated by dashed lines in the spectrum.

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 was 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 pmols of phoshoAKT 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.

Fig. 8. Controls for O-GlcNAc Specificity in the Identification of Sites by BEMAD 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-acetylglucosamindase 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.

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 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 CID of the modified Nup155 peptide, all theoretical b and y ions are displayed as dashed lines, and (C) the resulting DTT modified sequence from Nup155 with b and y ion fragments correctly interpreted are shown in bold.

REFERENCES

1. Haynes, P. A., Gygi, S. P., Figeys, D., and Aebersold, R. (1998) *Electrophoresis* **19,** 1862-71.

2. Uy, R., and Wold, F. (1977) *Science* **198,** 890-6.

3. Kouzarides, T. (2000) *Embo J* **19,** 1176-9.

4. McBride, A. E., and Silver, P. A. (2001) *Cell* **106,** 5-8.

5. Hart, G. W. (1997) *Annu. Rev. Biochem.* **66,** 315-35

6. Wells, L., Vosseller, K., and Hart, G. W. (2001) *Science* **291,** 2376-8.

7. Hanover, J. A. (2001) *Faseb J* **15,** 1865-76.

8. Comer, F. I., and Hart, G. W. (2000) *J Biol Chem* **275,** 29179-82.

9. Zachara, N. E., and Hart, G. W. (2002) *Chem Rev* **102,** 431-8.

10. Vosseller, K., Wells, L., and Hart, G. W. (2001) *Biochimie* **83,** 575-81.

11. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) *J Biol Chem* **272,** 9308- 15.

12. Lubas, W. A., Frank, D. W., Krause, M., and Hanover, J. A. (1997) *J Biol Chem* **272,** 9316-24.

13. Kreppel, L. K., and Hart, G. W. (1999) *J Biol Chem* **274,** 32015-22.

14. Lubas, W. A., and Hanover, J. A. (2000) *J Biol Chem* **275,** 10983-8.

15. Gao, Y., Wells, L., Comer, F. I., Parker, G. J., and Hart, G. W. (2001) *J Biol Chem* **276,** 9838-45.

16. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) *J Biol Chem* **277,** 1755-61.

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., and et al. (1995) *Adv Exp Med Biol* **376,** 115-23

18. Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110,** 69-80.

19. Vosseller, K., Wells, L., Lane, M. D., and Hart, G. W. (2002) *Proc Natl Acad Sci U S A* **99,** 5313-8.

20. McClain, D. A., Lubas, W. A., Cooksey, R. C., Hazel, M., Parker, G. J., Love, D. C., and Hanover, J. A. (2002) *Proc Natl Acad Sci U S A* **99**, 10695-9.

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) *Proc Natl Acad Sci U S A* **97,** 5735-9.

22. Roquemore, E. P., Chou, T. Y., and Hart, G. W. (1994) *Methods Enzymol* **230,** 443-60

23. Greis, K. D., Hayes, B. K., Comer, F. I., Kirk, M., Barnes, S., Lowary, T. L., and Hart, G. W. (1996) *Anal Biochem* **234,** 38-49.

24. Haynes, P. A., and Aebersold, R. (2000) *Anal Chem* **72,** 5402-10.

25. Chalkley, R. J., and Burlingame, A. L. (2001) *J Am Soc Mass Spectrom* **12,** 1106- 13.

26. Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., and Pandey, A. (2002) *Trends Biotechnol* **20,** 261-8.

27. Pandey, A., Andersen, J. S., and Mann, M. (2000) *Sci STKE* **2000,** L1.

28. Conrads, T. P., Issaq, H. J., and Veenstra, T. D. (2002) *Biochem Biophys Res Commun* **290,** 885-90.

29. Loughrey Chen, S., Huddleston, M. J., Shou, W., Deshaies, R. J., Annan, R. S., and Carr, S. A. (2002) *Mol Cell Proteomics* **1,** 186-96.

30. Zappacosta, F., Huddleston, M. J., Karcher, R. L., Gelfand, V. I., Carr, S. A., and Annan, R. S. (2002) *Anal Chem* **74,** 3221-31.

31. Li, J., LeRiche, T., Tremblay, T. L., Wang, C., Bonneil, E., Harrison, D. J., and Thibault, P. (2002) *Mol Cell Proteomics* **1,** 157-68.

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) *Nat Biotechnol* **20,** 301-5.

33. Gronberg, M., Kristiansen, T. Z., Stensballe, A., Andersen, J. S., Ohara, O.,

Mann, M., Jensen, O. E., and Pandey, A. (2002) *Mol. Cell. Proteomics* **published online ahead of print July 22, 2002**

34. Adamczyk, M., Gebler, J. C., and Wu, J. (2001) *Rapid Commun Mass Spectrom* **15,** 1481-8

35. Goshe, M. B., Conrads, T. P., Panisko, E. A., Angell, N. H., Veenstra, T. D., and Smith, R. D. (2001) *Anal Chem* **73,** 2578-86.

36. Weckwerth, W., Willmitzer, L., and Fiehn, O. (2000) *Rapid Commun Mass Spectrom* **14,** 1677-81

37. Zhou, H., Watts, J. D., and Aebersold, R. (2001) *Nat Biotechnol* **19,** 375-8.

38. Oda, Y., Nagasu, T., and Chait, B. T. (2001) *Nat Biotechnol* **19,** 379-82.

39. Meyer, H. E., Hoffmann-Posorske, E., and Heilmeyer, L. M., Jr. (1991) *Methods Enzymol* **201,** 169-85

40. Torres, C. R., and Hart, G. W. (1984) *J Biol Chem* **259,** 3308-17.

41. Czeszak, X., Ricart, G., Tetaert, D., Michalski, J. C., and Lemoine, J. (2002) *Rapid Commun Mass Spectrom* **16,** 27-34

42. Comer, F. I., Vosseller, K., Wells, L., Accavitti, M. A., and Hart, G. W. (2001) *Anal Biochem* **293,** 169-77.

43. Greis, K. D., Gibson, W., and Hart, G. W. (1994) *J Virol* **68,** 8339-49.

44. Cole, R. N., and Hart, G. W. (1999) *J Neurochem* **73,** 418-28.

45. Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., and Matunis, M.J. (2002) J. Cell Biol. **in press**.

46. Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S., and Mische, S. M. (1999) *Electrophoresis* **20,** 601-5.

47. Arnott, D., Shabanowitz, J., and Hunt, D. F. (1993) *Clin Chem* **39,** 2005-10.

48. Gatlin, C. L., Eng, J. K., Cross, S. T., Detter, J. C., and Yates, J. R., 3rd. (2000) *Anal Chem* **72,** 757-63.

49. Lubas, W. A., Smith, M., Starr, C. M., and Hanover, J. A. (1995) *Biochemistry*

34, 1686-94.

50. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) *Nat Biotechnol* **17,** 994-9.

51. D'Onofrio, M., Starr, C. M., Park, M. K., Holt, G. D., Haltiwanger, R. S., Hart, G. W., and Hanover, J. A. (1988) *Proc Natl Acad Sci U S A* **85,** 9595-9.

52. Takano, M., Takeuchi, M., Ito, H., Furukawa, K., Sugimoto, K., Omata, S., and Horigome, T. (2002) *Eur J Biochem* **269,** 943-53.

53. Kamemura, K., Hayes, B. K., Comer, F. I., and Hart, G. W. (2002) *J Biol Chem* **277,** 19229-35.

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) *Genomics* **57,** 144-51.

Table 1. 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—homol. 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-alpha
40S Ribosomal Protein s24

Signal Transduction:
Casein kinase II
O-GlcNAc Transferase (OGT)
Phosphatase 2A Inhibitor (i2pp2a)
Rho GDP-dissociation inhibitor I (rho-gdi alpha)

Table 2. 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 database 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).

prior to enrichment by thiol chromatography:

z	predicted MH+	observed MH+	Xcorr	matched Ions	protein	peptide sequence
2	1726.5	1725.9	3.78	21/30	syn1 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:

z	predicted MH+	observed MH+	Xcorr	matched Ions	protein	peptide sequence
2	2189.0	2188.5	4.05	25/44	syn 1 rat	Q(T#)TAAAAATFSEQVGGGSGGAGR
2	1699.3	1698.3	3.10	22/30	syn 1 rat	G(S*)HSQTPSPGALPLGR
2	1591.3	1590.3	2.50	12/24	syn 1 rat	EMLS(S*)TTYPVVVK
1	1590.3	1590.3	2.42	14/24	syn 1 rat	EMLS(S*)TTYPVVVK
2	2189.0	2188.5	2.49	14/44	syn 1 rat	QTTAAAAATF(S*)EQVGGGSGGAGR
2	2972.1	2972.0	2.41	26/64	syn 1 rat	A(S*)TAAPVASPAAPSPGSSGGGGFFSSLSNAVK
2	2189.0	2188.4	1.97	14/44	syn 1 rat	QTTAAAAATFSEQVGGG(S*)GGAGR
2	2547.2	2546.2	1.88	12/46	syn 1 rat	LPSPTAAPQQ(S*)ASQATPMTQGQGR

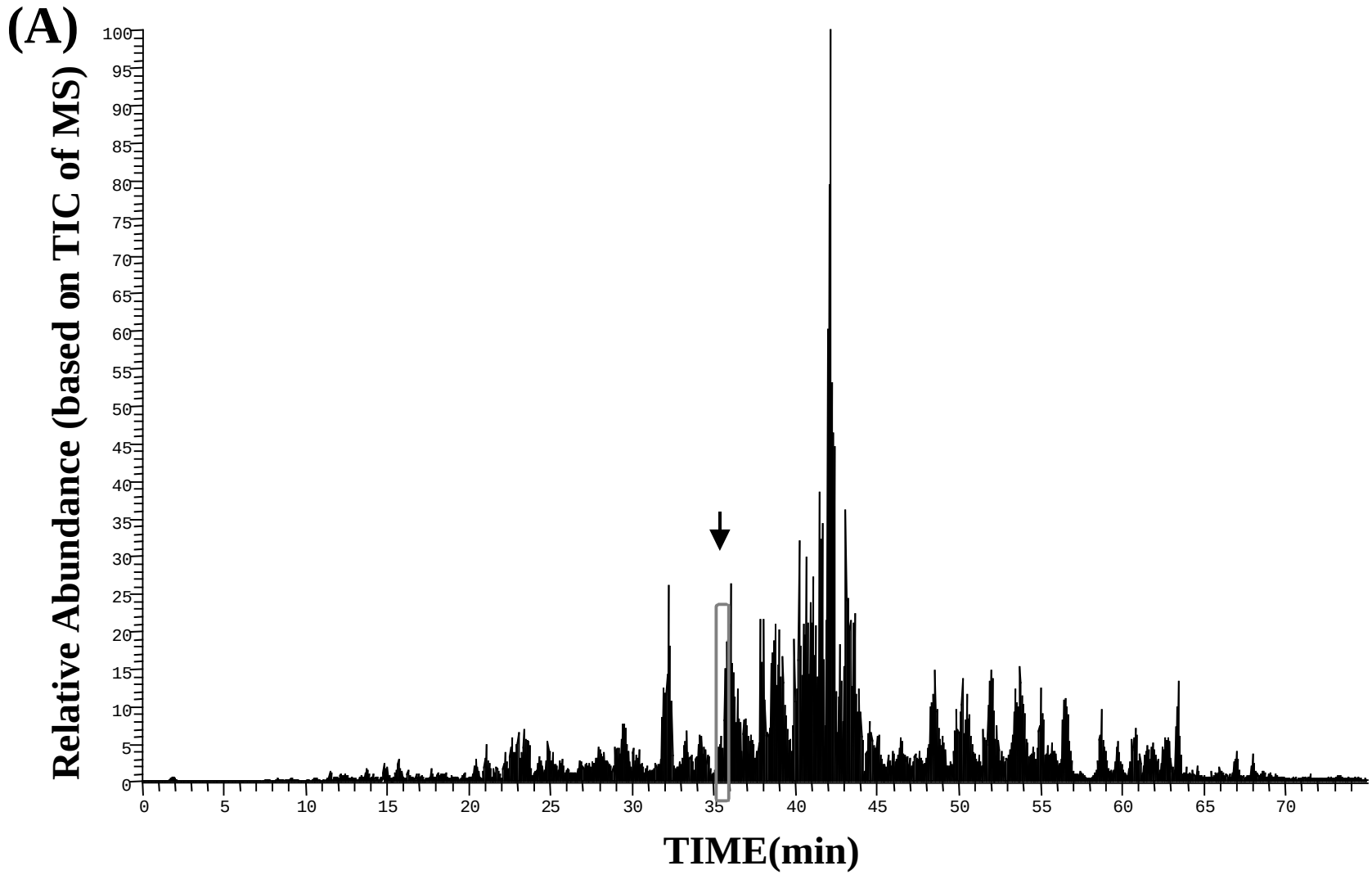
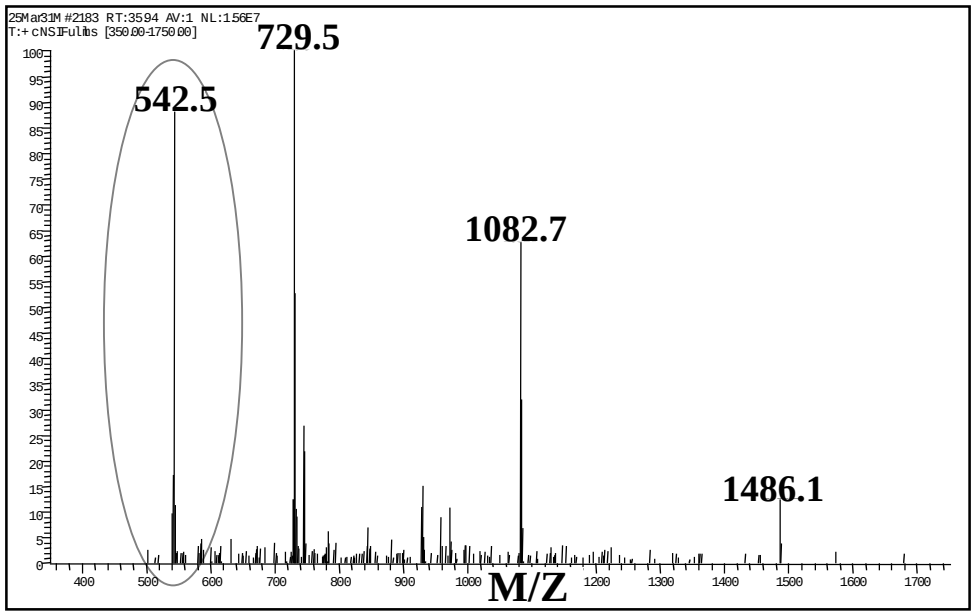


Fig. 1

(B)

Full Mass Spectrum @35.94 minutes



(C)

MS/MS
Automated Fragmentation of 542.5(m/z) @35.95min

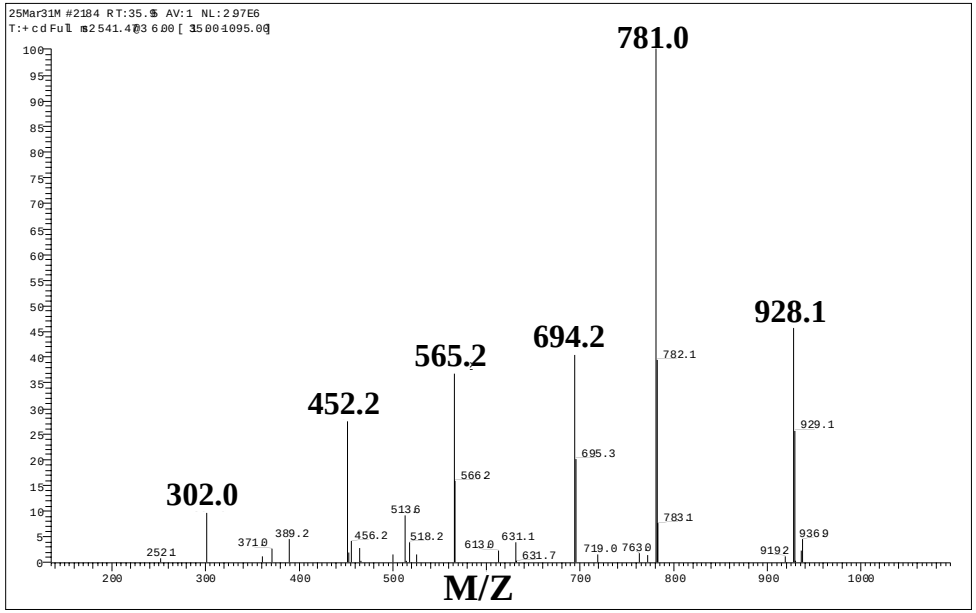


Fig. 1

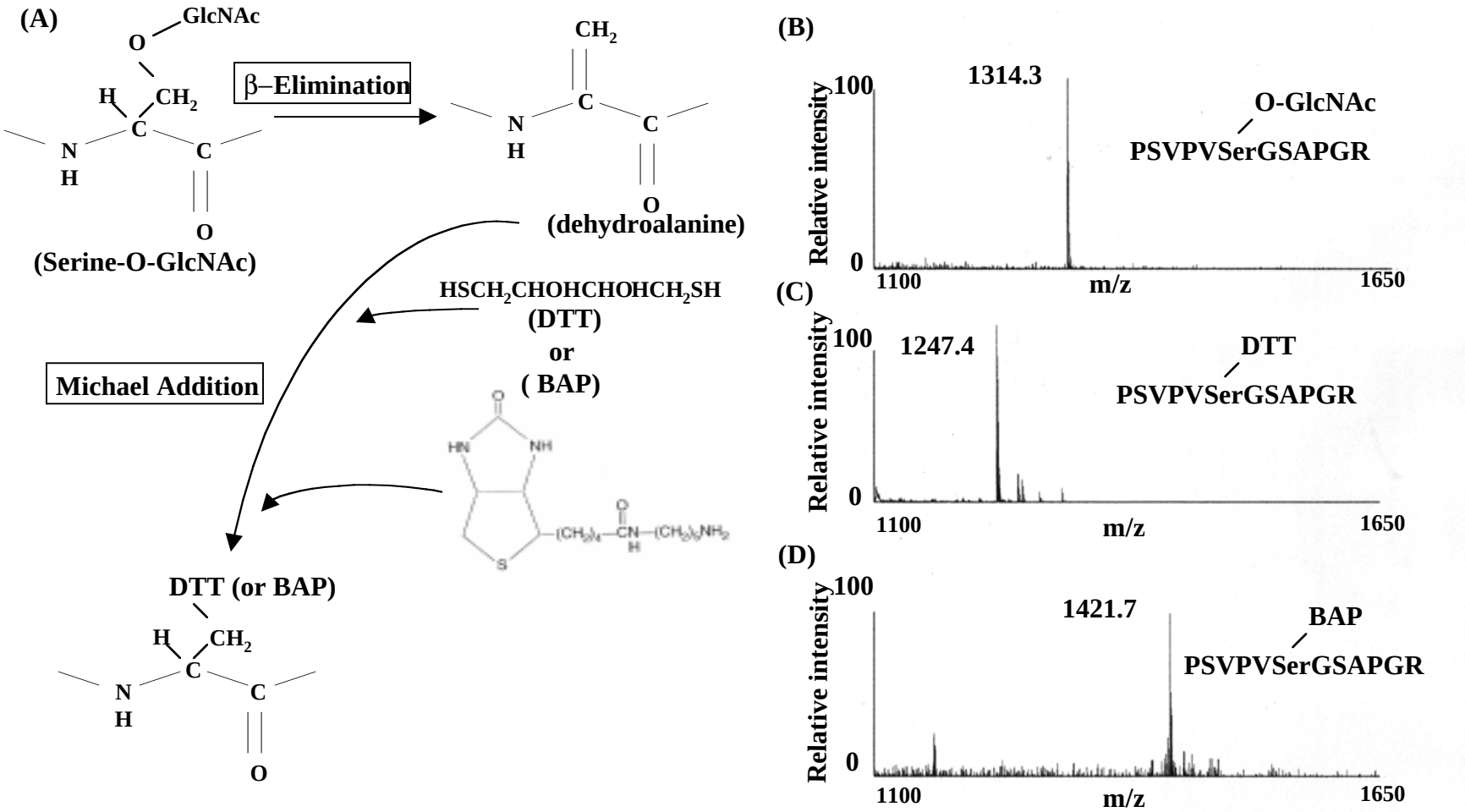


Fig. 2

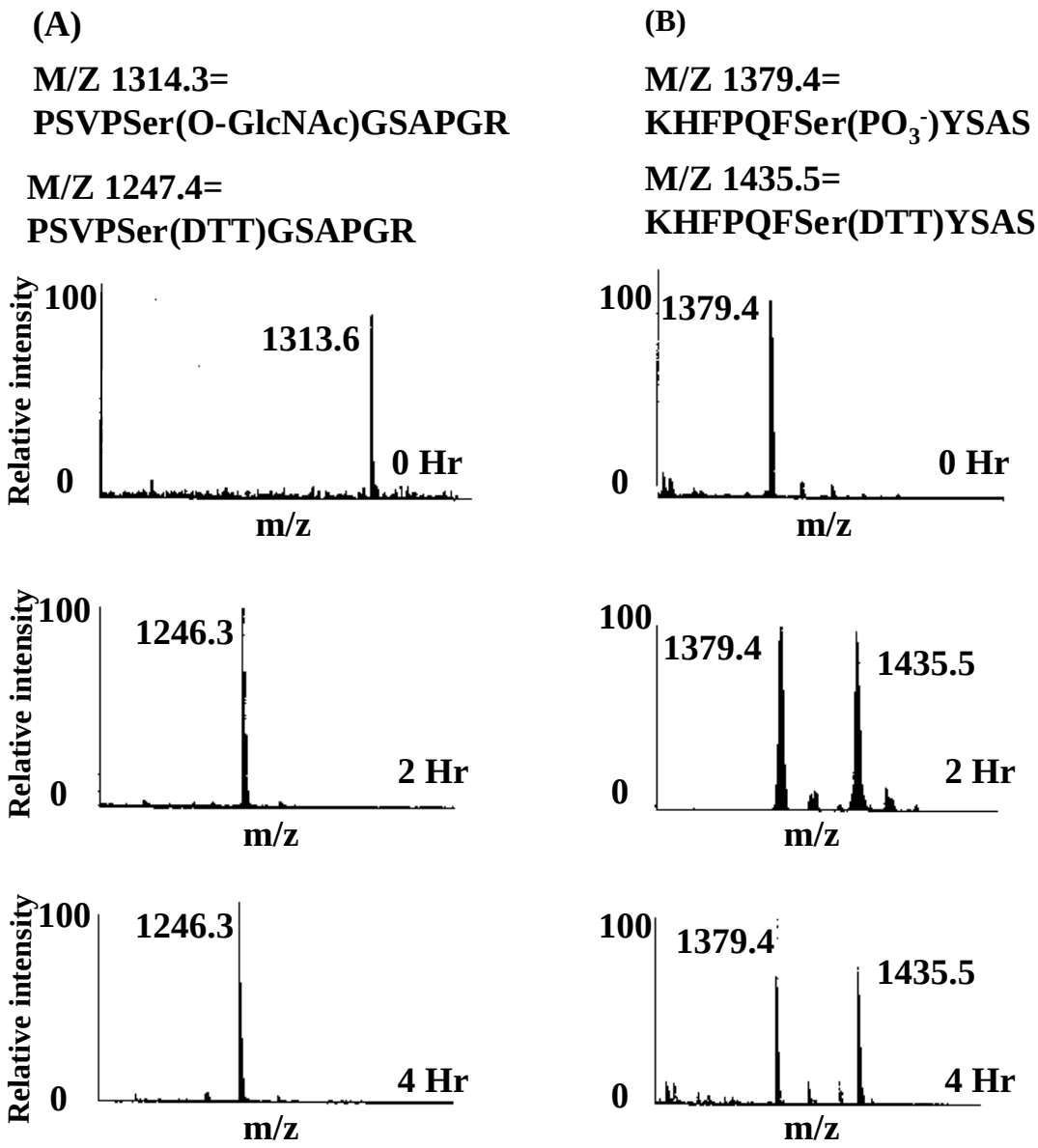


Fig. 3

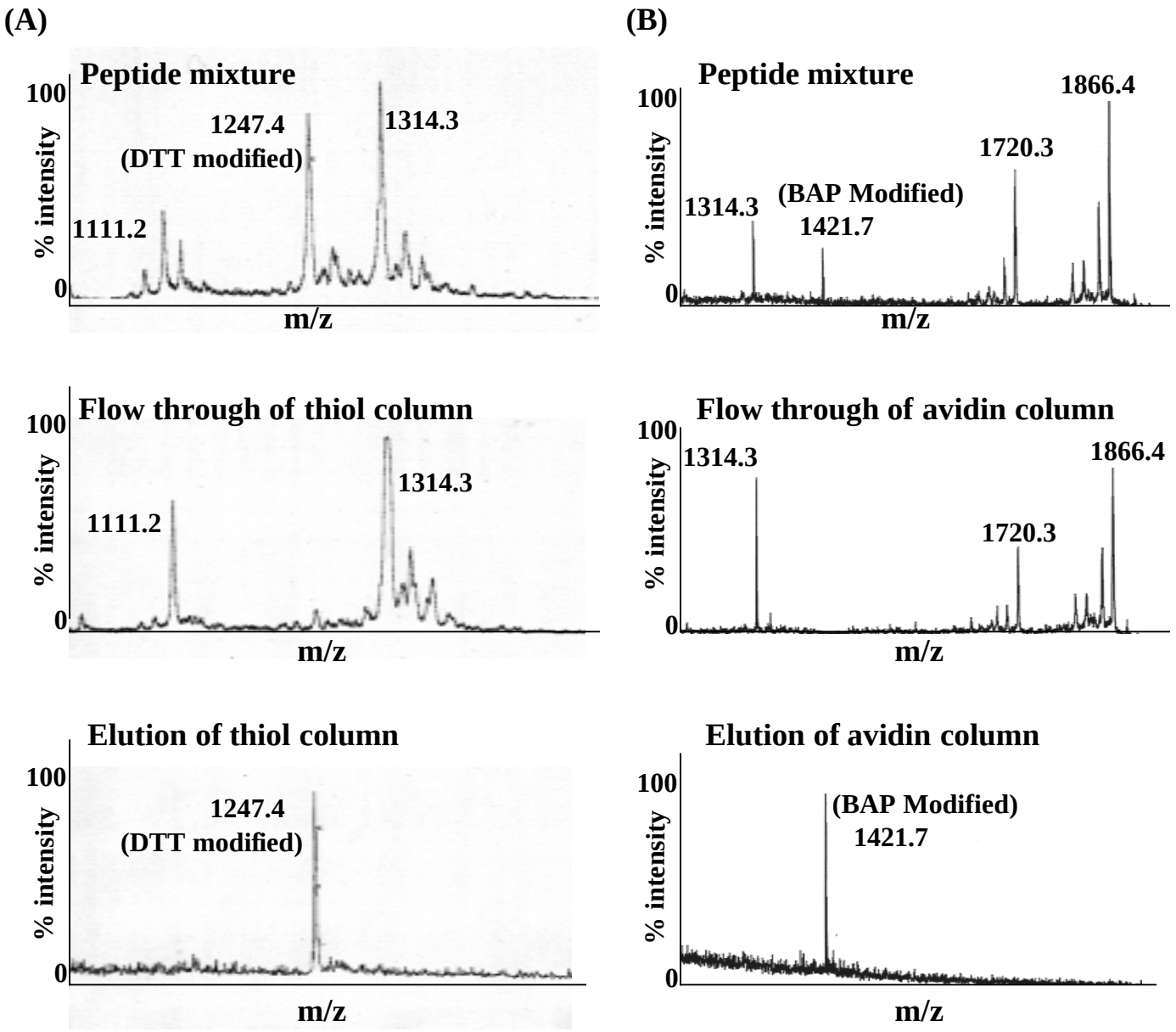
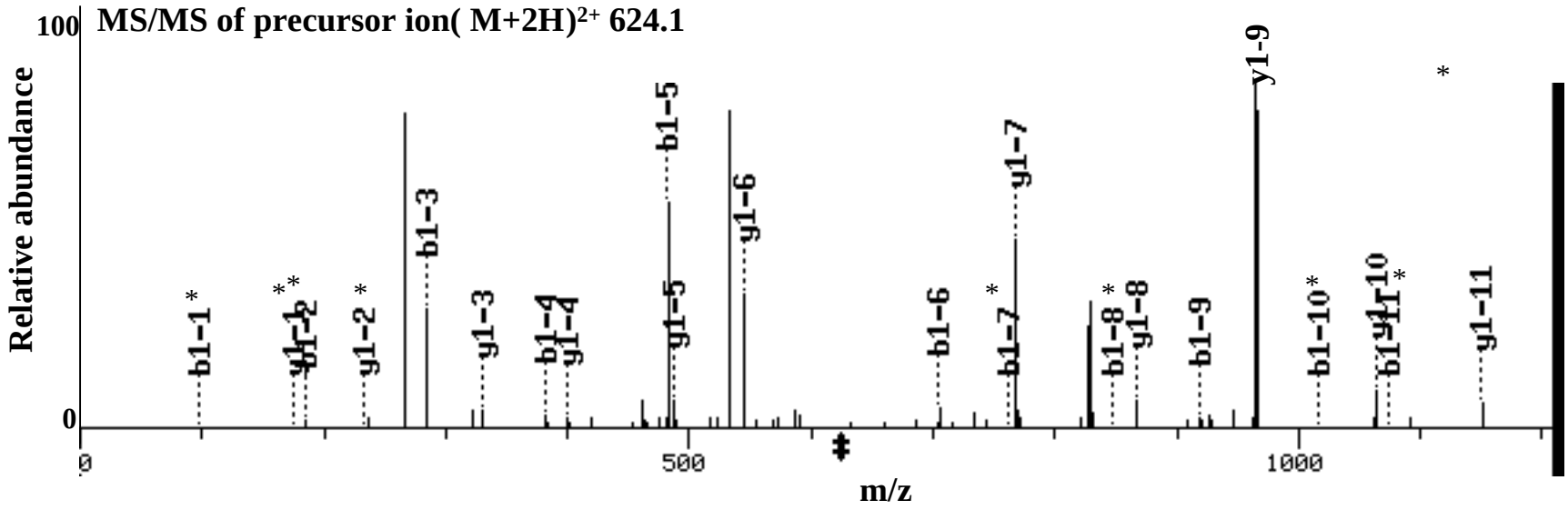


Fig. 4

(A)



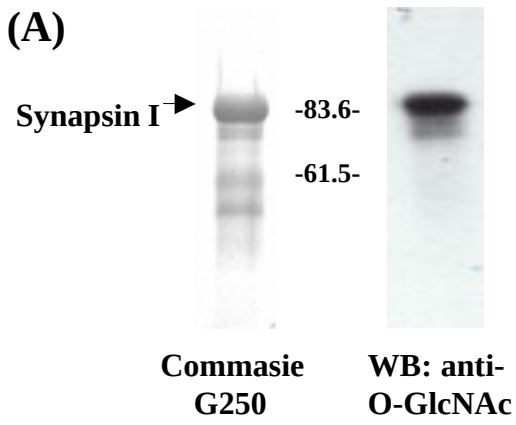
(B) **Sequest summary:**

z	observed MH+	Xcorr	matched ions	protein	peptide sequence
2	1247.3	3.27	15/22	BPP	PSVPVS(*)GSAPGR

(C)

	1247.1	**1150**	**1063**	**963.9**	**866.9**	**767.8**	**544.3**	**487.3**	**400.2**	**329.2**	232.1	175.1- y ions	
	12	11	10	9	8	7	6	5	4	3	2	1	#
Peptide sequence	**P**	**S**	**V**	**P**	**V**	**S-DTT**	**G**	**S**	**A**	**P**	**G**	**R**	
#	1	2	3	4	5	6	7	8	9	10	11	12	
b ions -98.1		**185.1**	**284.2**	**381.2**	**480.3**	**703.8**	760.8	847.9	**918.9**	1016	1073	1229.1	

Fig. 5



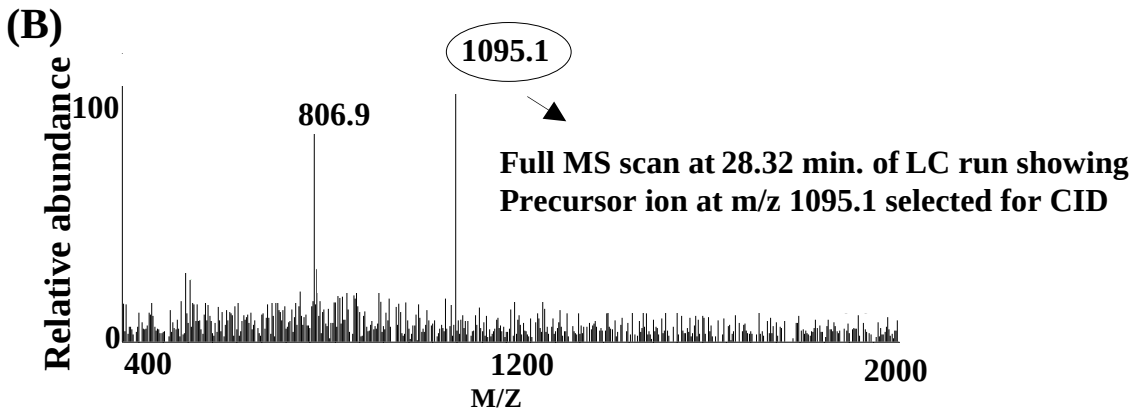


Fig. 6

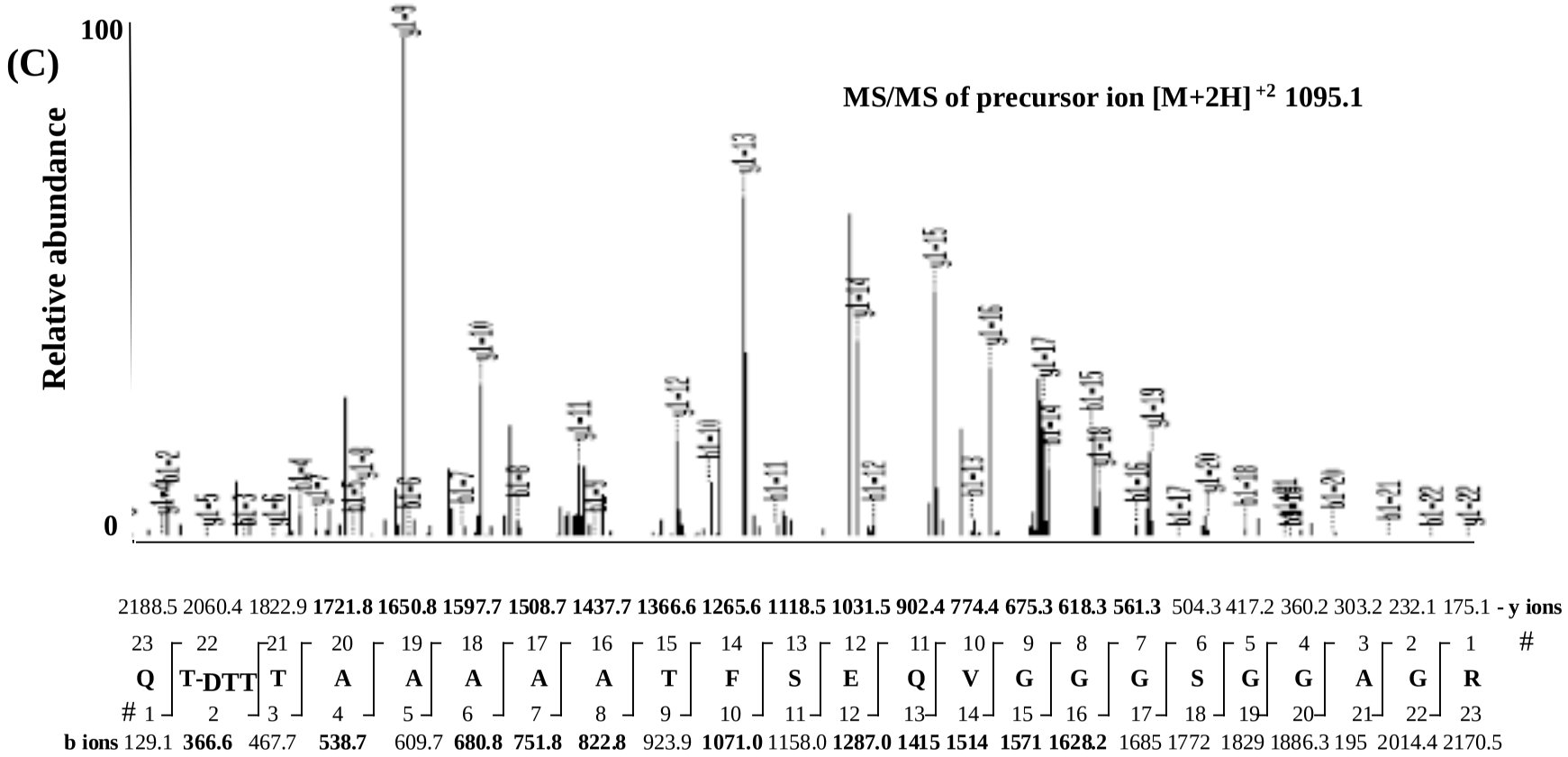
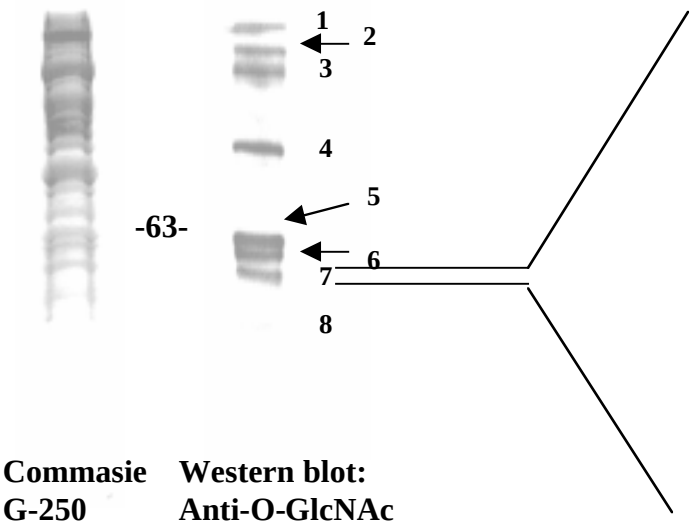


Fig. 6

(A) **Purified nuclear pore complex**



Commasie G-250 **Western blot: Anti-O-GlcNAc**

(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

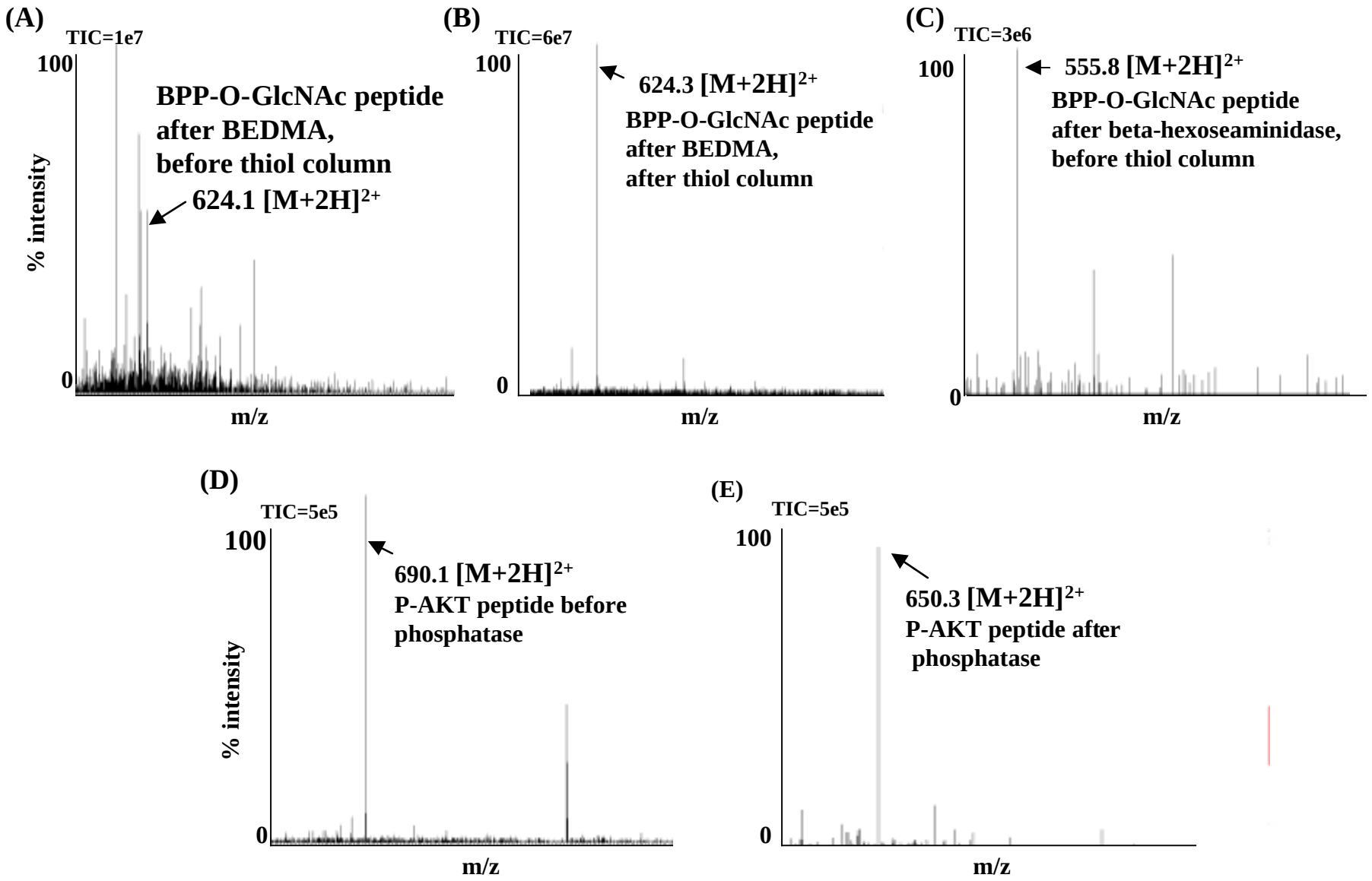


Fig. 8

(A)

BEMAD of soluble nuclear pore complex fraction.

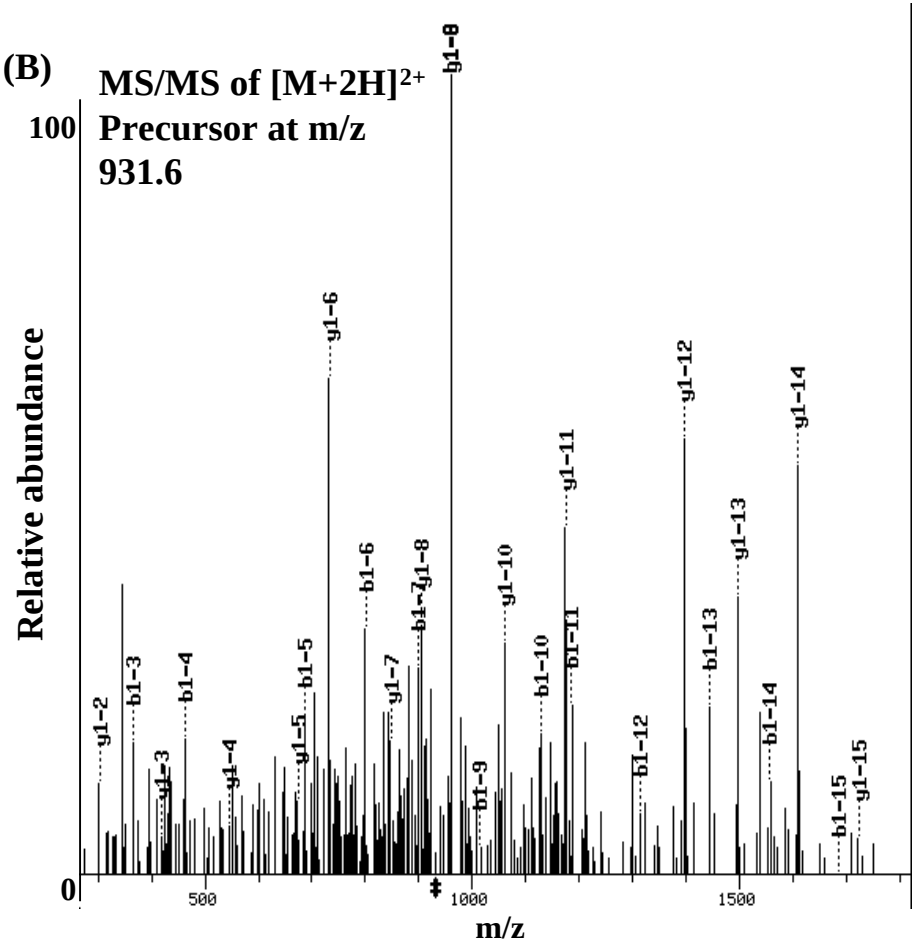
BEFORE THIOL CHROMATOGRAPHY

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)

1861.1 1723.9 **1610.8** **1497.6** **1398.5** **1175.2** **1061.1** **962.0** **904.9** 847.9 **732.8** 657.7 546.6 417.5 304.3 175.2 -Y ions

	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	#
	H	**L**	**L**	**V**	**S-DTT**	**N**	**V**	**G**	**G**	**D**	**G**	**E**	**E**	**I**	**E**	**R**	
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

B ions -138.1 251.3 **364.5** **463.6** **686.9** **801.0** **900.1** **957.2** 1014.2 **1129.3** 1186.4 1315 **1444.6** 1557 1686.9 1843.1

Fig. 9