analytical chemistry

Article

pubs.acs.org/ac

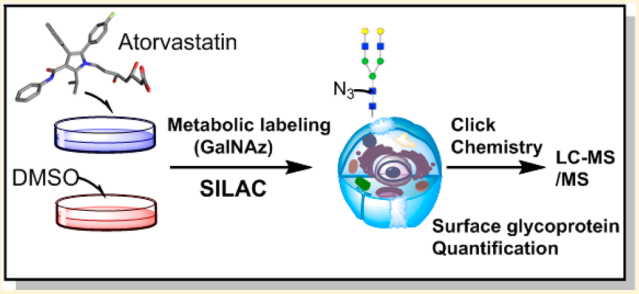
Site-Specific Quantification of Surface N-Glycoproteins in Statin-Treated Liver Cells

Haopeng Xiao, George X. Tang, and Ronghu Wu*

School of Chemistry and Biochemistry and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Supporting Information

ABSTRACT: The frequent modification of cell-surface proteins by N-linked glycans is known to be correlated with many biological processes. Aberrant glycosylation on surface proteins is associated with different cellular statuses and disease progression. However, it is extraordinarily challenging to comprehensively and site-specifically analyze glycoproteins located only on the cell surface. Currently mass spectrometry (MS)-based proteomics provides the possibility to analyze the N-glycoproteome, but effective separation and enrichment methods are required for the analysis of surface glycoproteins prior to MS measurement. The introduction of bio-orthogonal groups into proteins accelerates research in the robust visualization, identification, and quantification of proteins. Here we have comprehensively evaluated different sugar analogs in the analysis of cell-surface N-glycoproteins by combining copper-free click chemistry and MS-based proteomics. Comparison of three sugar analogs, N-azidoacetylgalactosamine (GalNAz), N-azidoacetylglucosamine (GlcNAz), and N-azidoacetylmannosamine (ManNAz), showed that metabolic labeling with GalNAz resulted in the greatest number of glycoproteins and glycosylation sites in biological duplicate experiments. GalNAz was then employed for the quantification experiment in statin-treated HepG2 liver cells, and 280 unique N-glycosylated sites were quantified from 168 surface proteins. The quantification results demonstrated that many glycosylation sites on surface proteins were down-regulated in statin-treated cells compared to untreated cells because statin prevents the synthesis of dolichol, which is essential for the formation of dolichol-linked precursor oligosaccharides. Several glycosylation sites in proteins that participate in the Alzheimer's disease pathway were down-regulated. This method can be extensively applied for the global analysis of the cell-surface N-glycoproteome.



Glycosylation is one of the most important protein modifications and is essential for cell survival.[1] There are two major types of protein glycosylation: N-linked glycosylation and O-linked glycosylation. For O-glycosylation, glycans are bound to the side chains of serine and threonine, while N-glycosylation involves glycans covalently attached to the side chain of asparagine. In eukaryotic cells, there is machinery in the endoplasmic reticulum (ER) responsible for attaching the initial glycan block ($(GlcNAc)_2(Mannose)_9$-$(Glucose)_3$) to nascent peptides. It is well-known that N-glycosylation plays determinant roles in protein folding and trafficking, and N-glycosylated proteins are especially important in regulating extracellular activities, including cell−cell communication and cell-matrix interactions.[2] Aberrant protein N-glycosylation is frequently related to human disease,[3] including Alzheimer's disease (AD),[4] cancer,[5] and infectious diseases.[6]

Half a century ago, early mammalian cell morphology studies discovered abundant carbohydrates on the external surface of the cell membrane.[7,8] To date, numerous research results have indicated that the majority of cell surface proteins are glycosylated.[9,10] Despite the number of glycoproteins located on the cell surface and their importance in biological functions, the global analysis of surface glycoproteins is extraordinarily challenging.[11,12] Modern mass spectrometry (MS)-based proteomics techniques provide the capacity to comprehensively analyze protein modifications.[13−24] These methods can be employed to systematically identify modified proteins, localize the modification sites, and quantify their abundance changes.[25−29] However, the heterogeneity of glycans and low abundance of many glycoproteins make the global analysis of glycoproteins very difficult.[30] It is even more challenging to specifically and comprehensively analyze N-glycoproteins only on the cell surface because it requires the selective separation and enrichment of surface N-glycoproteins.

Statins have been widely used to lower cholesterol levels in patients by inhibiting 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGCR),[31] an enzyme in the upstream portion of the mevalonate pathway. Besides cholesterol, the synthesis of many intermediate and end products in this pathway, including ubiquinone and dolichol, is significantly affected by the

Received: December 23, 2015
Accepted: February 19, 2016
Published: February 19, 2016

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

inhibition of this enzyme.[32] Dolichol plays an essential role in protein N-glycosylation and functions as a membrane anchor for the formation of a precursor oligosaccharide.[33] The effect of statin on surface protein glycosylation is still unknown which may contribute to the pleiotropic effects of statins. The systematic and quantitative analysis of surface glycoproteins in statin-treated cells will potentially shed light on the molecular mechanisms behind the pleiotropic effects of statins, which will allow patients to receive the full benefits of this medicine.

In this work, we systematically evaluated metabolic labeling with three sugar analogs, i.e. GalNAz, GlcNAz, and ManNAz, for the identification of cell surface N-glycoproteins by combining copper-free click chemistry and MS-based proteomics. The parallel experiments showed that GalNAz labeling resulted in the greatest number of protein N-glycosylation sites identified, while GlcNAz resulted in the smallest number of protein N-glycosylation sites. Thus, GalNAz labeling was employed for the global quantification of surface glycoproteins in HepG2 liver cells treated with statin. Systematic and quantitative analysis of surface proteins in statin-treated cells clearly demonstrated that many glycosylation sites were down-regulated compared to untreated cells. This method offers a means to globally, site-specifically, and quantitatively study protein N-glycosylation on the cell surface.

EXPERIMENTAL SECTION

Cell Culture and Metabolic Labeling. HepG2 (C3A) cells (Hep G2 [HEPG2] (ATCC HB8065)) were grown in a humidified incubator at 37 °C and 5.0% CO_2 in Dulbecco's modified eagle's medium (DMEM) (Sigma-Aldrich) containing low glucose and 10% fetal bovine serum (FBS) (Thermo). For the glycoprotein identification experiments, when cells reached about 40% confluency, 100 μM GalNAz, GlcNAz, or ManNAz (Click Chemistry Tools) was added to the media, respectively. Cells were further cultured for 24 h. In each case, duplicate biological experiments were performed.

For the quantification experiment, "heavy" and "light" stable isotope labeling by amino acids in cell culture (SILAC) (Sigma-Aldrich) media were freshly prepared by adding 0.146 g/L $^{13}C_6{}^{15}N_2$ L-lysine (Lys-8) and 0.84 g/L $^{13}C_6$ L-arginine (Arg-6) (Cambridge Isotopes Inc.) or the corresponding nonlabeled lysine (Lys-0) and arginine (Arg-0) to DMEM and supplemented with 10% dialyzed FBS (Corning). Cells were cultured for about six generations before the atorvastatin treatment. 40 mM atorvastatin (Cayman Chemical) stock solution was prepared in DMSO (Sigma-Aldrich). About 2 × 10^7 cells were treated with 15 μM atorvastatin in serum-free heavy medium for 48 h. A similar number of light cells were treated by DMSO in serum-free light medium as a control. 100 μM GalNAz was added in after 24 h of atorvastatin or DMSO treatment.

In-Flask Copper-Free Click Reaction, Cell Lysis, and Protein Digestion. Cells were washed twice with phosphate buffered saline (PBS) before 100 μM dibenzocyclooctyne (DBCO)-sulfo-biotin in PBS was added into the cell culture flasks. Surface glycoproteins were tagged with biotin through the specific click reaction between DBCO and the azido group in the sugar analogs under physiological conditions.[34−36] Cells were incubated for 1 h with gentle agitation at 37 °C and then harvested by scraping in PBS. For the quantification experiments, heavy and light cells were equally combined based on the protein ratio of 1:1 from a trial run. The cell mixtures were pelleted by centrifugation at 500 g for 3 min and washed twice with cold PBS. Cytosol proteins were removed by incubating in a buffer containing 150 mM NaCl, 50 mM HEPES pH = 7.6, 25 μg/mL digitonin, and 1 tablet/10 mL protease inhibitor (complete mini, EDTA-free, Roche) on ice for 10 min. After incubation, samples were centrifuged at 2000 g for another 10 min. Cell pellets were washed with the previous buffer and then lysed through end-over-end rotation at 4 °C for 45 min in lysis buffer (50 mM HEPES pH = 7.6, 150 mM NaCl, 0.5% SDC, 10 units/mL benzonase, and 1 tablet/10 mL protease inhibitor). Lysates were centrifuged, and the resulting supernatant was transferred to new tubes. Proteins were subjected to disulfide reduction with 5 mM DTT (56 °C, 25 min) and alkylation with 14 mM iodoacetamide (RT, 20 min in the dark). Detergent was removed by methanol-chloroform protein precipitation. The purified proteins were digested with 10 ng/μL Lys-C (Wako) in 50 mM HEPES pH 8.6, 1.6 M urea, 5% ACN at 31 °C for 16 h, followed by further digestion with 8 ng/uL Trypsin (Promega) at 37 °C for 4 h.

Glycopeptide Separation and Enrichment. Digestion mixtures were acidified by the addition of trifluoroacetic acid (TFA) to a final concentration of 0.1%, clarified by centrifugation, and desalted using a tC18 Sep-Pak cartridge (Waters). Purified peptides were dried and then enriched with NeutrAvidin beads (Thermo) at 37 °C for 30 min. The samples were transferred to spin columns, followed by thoroughly washing according to the manufacturer's protocol. Peptides were then eluted from the beads three times by 2 min incubations with 200 μL of 8 M guanidine-HCl (pH = 1.5) at 56 °C. Eluates were combined, desalted using tC18 Sep-Pak cartridge, and lyophilized overnight. Completely dry peptides were deglycosylated with eight units of peptide-*N*-glycosidase F (PNGase F, Sigma-Aldrich) in 40 μL buffer containing 50 mM NH_4HCO_3 (pH = 9) in heavy-oxygen water ($H_2{}^{18}O$) for 3 h at 37 °C. The reaction was quenched by adding formic acid (FA) to a final concentration of 1%. Peptides were further purified via stage tip and separated into 3 fractions using 20%, 50%, and 80% ACN containing 1% HOAc.

LC-MS/MS Analysis. Purified and dried peptide samples were each dissolved in 10 μL of 5% ACN and 4% FA, and 4 μL of the resulting solutions was loaded onto a microcapillary column packed with C18 beads (Magic C18AQ, 3 μm, 200 Å, 100 μm × 16 cm, Michrom Bioresources) by a Dionex WPS-3000TPLRS autosampler (UltiMate 3000 thermostated Rapid Separation Pulled Loop Wellplate Sampler). Peptides were separated by reversed-phase chromatography using an UltiMate 3000 binary pump with a 110 min gradient of 3−25%, 8−38%, and 10−50% ACN (with 0.125% FA) for the three fractions, respectively. Peptides were detected with a data-dependent Top20 method[37] in a hybrid dual-cell quadrupole linear ion trap − Orbitrap mass spectrometer (LTQ Orbitrap Elite, ThermoFisher, with Xcalibur 3.0.63 software). For each cycle, one full MS scan (resolution: 60,000) in the Orbitrap at 10^6 AGC target was followed by up to 20 MS/MS in the LTQ for the most intense ions. The selected ions were excluded from further analysis for 90 s. Ions with singly or unassigned charge were not sequenced. Maximum ion accumulation times were 1000 ms for each full MS scan and 50 ms for MS/MS scans.

Database Search and Data Filtering. All MS^2 spectra were converted into an mzXML format and then searched using the SEQUEST algorithm (version 28).[38] Spectra were matched against a database containing sequences of all proteins in the UniProt Human (*Homo sapiens*) database. The following parameters were used during the search: 10 ppm precursor

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

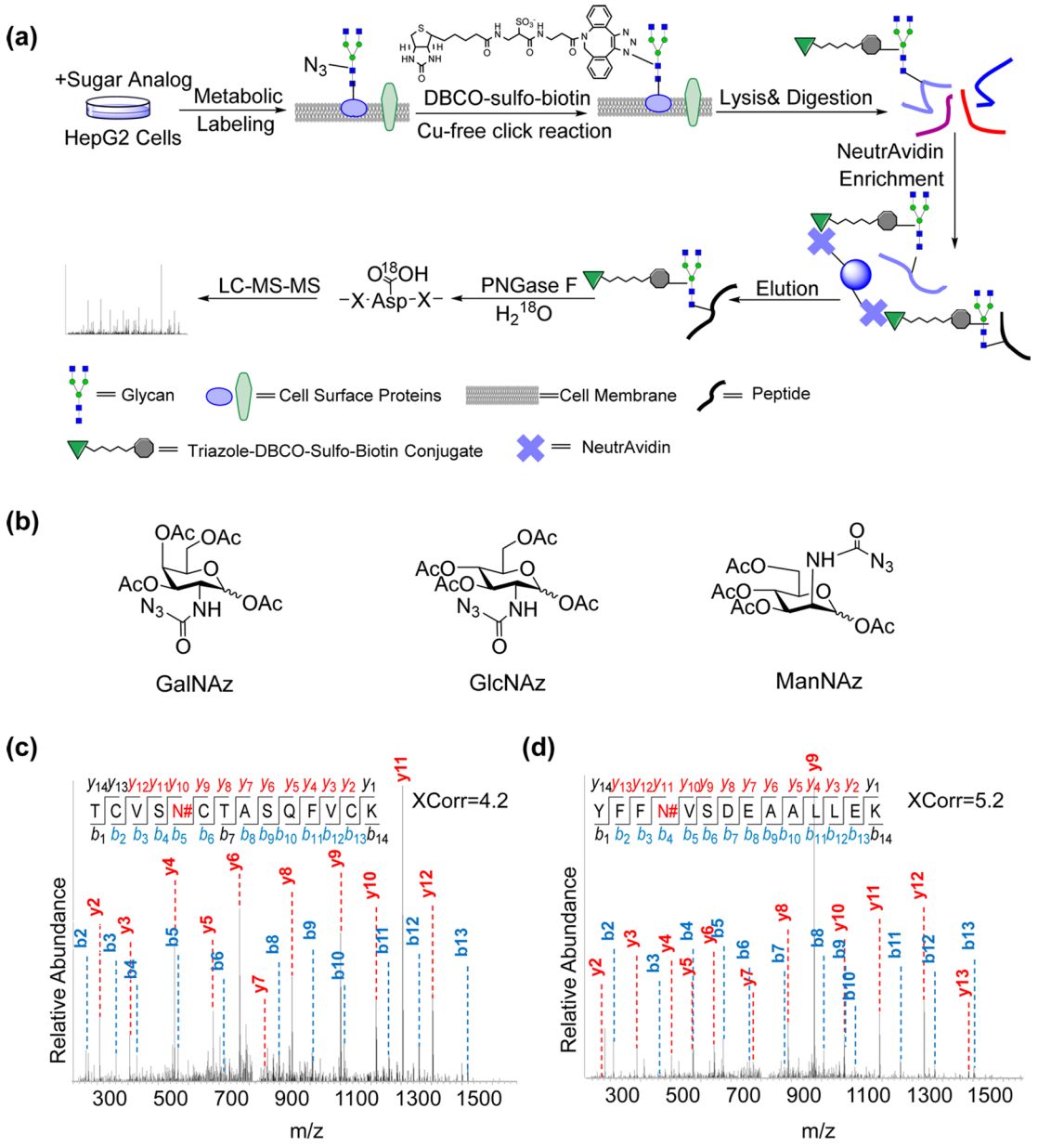


Figure 1. (a) Experimental procedure for the global analysis of the N-glycoproteome on the cell surface. (b) The structures of three sugar analogs used: GalNAz, GlcNAz, and ManNAz. (c) A sample tandem mass spectrum of the peptide TCVSN#CTASQFVCK from LRP1. (d) Another sample MS^2 of YFFN#VSDEAALLEK from ITGA2 (# denotes the glycosylation site).

mass tolerance; 1.0 Da product ion mass tolerance; fully digested with trypsin; up to three missed cleavages; fixed modifications: carbamidomethylation of cysteine (+57.0214); variable modifications: oxidation of methionine (+15.9949), O^{18} tag of asparagine (+2.9883), heavy lysine (+8.0142), and heavy arginine (+6.0201). False discovery rates (FDR) of peptide and protein identifications were evaluated and controlled by the target-decoy method.[39] Each protein sequence was listed in both forward and reversed orders. Linear discriminant analysis (LDA), which is similar to other methods in the literature,[40] was used to control the quality of peptide identifications using parameters such as XCorr, precursor mass error, and charge state.[41] Peptides fewer than seven amino acid residues in length were deleted. Furthermore, peptide spectral matches were filtered to a 1% FDR. The data set was restricted to glycopeptides when determining FDRs for glycopeptide identification.[42] Furthermore, an additional protein-level filter was applied in each data set to reduce the protein-level FDRs (<1%) for glycoproteins. Consequently the FDRs at the peptide level were much less than 1%.

Glycosylation Site Localization and Peptide Quantification. We assigned and measured the confidence of glycosylation site localizations by calculating their ModScores, which applies a probabilistic algorithm[42] that considers all possible glycosylation sites in a peptide and uses the presence of experimental fragment ions unique to each site. Sites with a ModScore >13 (P < 0.05) were considered confidently localized. For peptide quantification, we required an S/N value >3 for both heavy and light peptides. If the S/N value of a certain heavy peptide was less than 3, then that of the corresponding light peptide was required to be greater than 5 and *vice versa*. If the same glycopeptide was quantified several times, the median value was used as the glycopeptide abundance change. Glycosylation site quantification had the

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

following criteria: first, the quantified glycopeptide contain only a single glycosylation site; second, the site be well-localized with a ModScore >13. If multiple unique singly glycosylated peptides containing the same glycosylation site were identified, the ratio of the glycosylation sites was the median value of these glycopeptide ratios.

RESULTS AND DISCUSSION

Metabolic Labeling, Surface Glycoprotein Enrichment, and MS Analysis. In recent years, unnatural sugars containing a bio-orthogonal group have been used to label glycosylated proteins.[9,43] Glycoproteins bearing the biologically inert azido or alkyne group can be bound to a fluorescence probe to visualize them. This metabolic labeling can also be used to selectively enrich glycoproteins based on the unique bio-orthogonal group. Several sugar analogs, including GlcNAz, GalNAz, and ManNaz (structures are in Figure 1b), have been reported to label cells.[9,44,45] Here, in parallel experiments, we labeled cells using each of the three sugar analogs and evaluated their effectiveness for the global and site-specific analysis of N-glycoproteins on the cell surface in combination with MS-based proteomics.

After cells were cultured in low glucose media containing each of these three sugar analogs, surface glycoproteins containing the azido functional group on living cells were selectively bound to DBCO-sulfo-biotin via in-flask copper-free click chemistry under physiological conditions for 1 h, as shown in Figure 1a. Because the hydrophilic DBCO-sulfo-biotin cannot penetrate the cell plasma membrane, only glycoproteins located on the cell surface were tagged under mild conditions. After cell lysis and protein digestion, surface glycoproteins tagged with a biotin group allowed further enrichment to be performed based on strong and specific interactions between biotin in glycopeptides and NeutrAvidin, which was conjugated to beads. Nonmodified peptides were removed by washing the beads several times. Nonspecific binding is a drawback of the streptavidin enrichment method; however, the enrichment took place at the peptide level, which increased specificity compared to protein level enrichment.

In order to generate a common tag for MS analysis, enriched peptides were treated with PNGase F in $H_2{}^{18}O$ to remove N-glycans, which converted asparagine (Asn) to heavy-oxygen aspartic acid (Asp) and created a mass shift of +2.9883 Da for glycosylation site identification by MS.[46,47] This strategy was similar to a previous method for proteolytic stable isotope labeling, in which heavy-oxygen water was used to label digested peptides with trypsin and individual proteins from two proteome samples were quantitatively analyzed.[48] In this case, heavy oxygen on Asp enabled us to distinguish authentic N-glycosylation sites from those caused by deamidation on nonglycosylated asparagines *in vitro* and *in vivo*. Deamidation on nonglycosylated asparagines could also occur during PNGase F treatment, which may result in false positive identifications, which is why we ran the reaction for 3 h to minimize false positive identification. Control experiments showed that the effect of uncontrolled deamidation within the 3 h PNGase F treatment was nearly negligible, which is described in more detail below. In addition, after glycopeptide enrichment, the presence of nonglycosylated peptides was significantly decreased; therefore, the chance of any deamidation from nonglycosylated peptides was dramatically reduced. Overall, the PNGase F treatment in heavy oxygen water increased the confidence of glycopeptide identification.

An Orbitrap mass spectrometer with high resolution and mass accuracy provides the capability to confidently identify glycopeptides. For example, two tandem mass spectra for two glycopeptides in a cluster of differentiation (CD) proteins, which are very important for differentiation and classification of cells, are shown in Figure 1c and d. The glycopeptide TCVSN#CTASQFVCK (# represents the glycosylation site) from LRP1 (CD91), prolow-density lipoprotein receptor-related protein 1, was identified with an XCorr of 4.2 and a mass accuracy of −0.47 ppm. LRP1 is a single-pass type I membrane protein and is involved in endocytosis and in phagocytosis of apoptotic cells. It may modulate cellular events, such as kinase-dependent intracellular signaling, neuronal calcium signaling, and neurotransmission. As shown in Figure 1d, YFFN#VSDEAALLEK was also very confidently identified with an XCorr of 4.6 and a mass accuracy of 0.13 ppm. This peptide is from the protein ITGA2, integrin alpha-2, which is a receptor for laminin, collagen, collagen C-propeptides, fibronectin, and E-cadherin. It is an extremely important surface protein that regulates cell adhesion, cell-matrix interactions, and host−virus interactions. In this work, we have confidently identified several glycosylation sites, i.e., N105, N112, N343, N432, N1057, and N1074 in ITGA2.

Evaluation of Glycopeptides and Glycosylation Sites Identified in Cells Labeled with Different Sugar Analogs. The procedure for our parallel experiments differed only in the sugar analogs, which allowed us to objectively evaluate the effectiveness of three sugar analogs for global surface glycoprotein analysis. The number of identified glycosylation sites and glycoproteins and their overlap between biological duplicate experiments using each of the three sugar analogs are displayed in Figure 2. Theoretically, each sugar analog labels a different group of glycoproteins based on glycan structure and the enzymes responsible for glycan synthesis; therefore, the results from these three labeling experiments are expected to be different. Overall, 590 glycosylation sites were identified on 274 proteins in the GalNAz labeling experiments (Table S1), including 261 proteins (95.3%) which were either secreted, located on the cell membrane, or exported by extracellular vesicular exosomes based on the information from Uniprot (www.uniprot.org). Meanwhile, 446 glycosylation sites on 219 proteins and 117 sites on 91 proteins were identified in the duplicate ManNAz (Table S2) and GlcNAz (Table S3) labeling experiments, respectively. The GlcNAz labeling covered the fewest number of glycosylation sites and glycoproteins, which corresponds very well to previous work showing better incorporation of GalNAz or ManNAz over GlcNAz.[49] Among 434 and 467 N-glycosylation sites identified in each GalNAz labeling experiment, 311 sites were common to both experiments. At the glycoprotein level, as expected, the overlap was even higher. In the two experiments, 217 and 231 glycoproteins were identified, and the number of overlapped proteins was 174. Considering the large-scale analysis, this level of overlap is within a reasonable range. Since we ran biological duplicate experiments for each sugar analog, the inconsistencies between duplicates could be due to the sample differences, the dynamic nature of protein glycosylation, sample preparation (sample loss), or false positive identifications. The comparison of surface glycosylation sites and glycoproteins identified in the three parallel experiments using different sugar analogs is displayed in Figure 3. The majority of the glycosylation sites and glycoproteins identified in ManNAz labeling experiments (335 of 446 sites, 176 of 219 proteins) and GlcNAz labeling

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

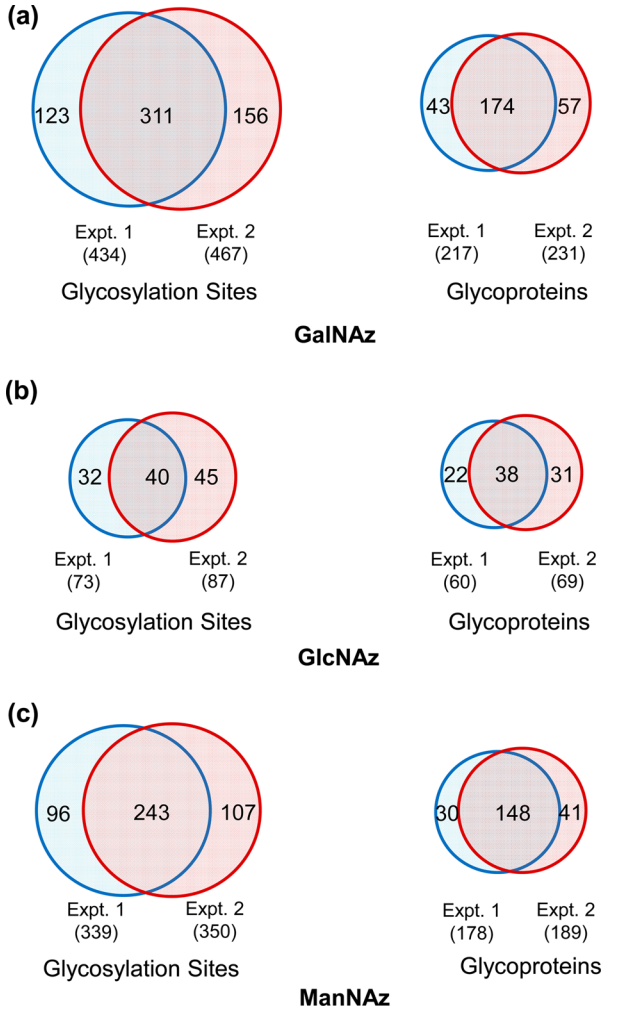


Figure 2. Reproducibility assessment in duplicate labeling experiments of (a) GalNAz, (b) GlcNAz, and (c) ManNAz.

experiments (90 of 117 sites, 70 of 91 proteins) were also identified in the GalNAz experiments, demonstrating the highest coverage of surface glycosylation sites and glycoproteins. Based on these results, GalNAz was employed for the quantification of surface proteins in statin-treated cells in order to obtain higher glycoprotein coverage, as described below.

Overall, 725 cell-surface glycosylation sites on 337 glycoproteins were identified combining all these experimental results.

In order to verify that the results were reliable, we designed two controls to run in parallel with the GalNAz labeling experiment. The experimental procedure for the first control was consistent with the GalNAz labeled experiment, except the click reaction was omitted; the second control omitted the PNGase F deglycosylation reaction. We identified 886 unique glycopeptides in the GalNAz labeling experiment and only 20 glycopeptides in the first control. These 20 unique glycopeptides may have resulted from nonspecific binding of the NeutrAvidin enrichment, nonglycosylated Asn deamidation in heavy-oxygen water during PNGase F treatment, or false positive identification of glycopeptides. However, any nonglycosylated Asn deamidation before or after the PNGase F treatment would result in a mass difference of 0.9840 Da, not 2.9883 Da, so they would be easily distinguishable during data analysis. Only nonglycosylated Asn deamidation within the 3 h PNGase F treatment in heavy-oxygen water would contribute to false positive identification. In the second control experiment without PNGase F, only 7 unique glycopeptides were identified, likely due to the deamination of free Asn. This is less than 1% compared to the 886 unique glycopeptides identified in the parallel GalNAz labeling experiment, which indicates that the effect of nonglycosylated Asn deamidation within the 3 h reaction was nearly negligible. These control experiments clearly verified the reliability of the current results.

Clustering of Surface N-Glycoproteins Identified in GalNAz Labeling Experiments. Most of the identified glycoproteins contain a single glycosylation site. There were also some proteins with more than ten sites; for example, 21 N-glycosylation sites were identified on LRP1. The clinical importance of LRP1 in Alzheimer's disease and cardiovascular disease also brings extensive attention to this protein. Glycosylation may stabilize this receptor-related protein and also differentiates the protein's functions in different tissues.[50]

In order to further evaluate the specificity of our method, the identified glycoproteins in the GalNAz labeling experiments were clustered according to molecular function and biological process using the Database for Annotation, Visualization, and Integrated Discovery 6.7 (DAVID 6.7).[51] We investigated the molecular functions of the identified glycoproteins and the biological processes they are involved in. The molecular functions with the highest level of enrichment were receptor

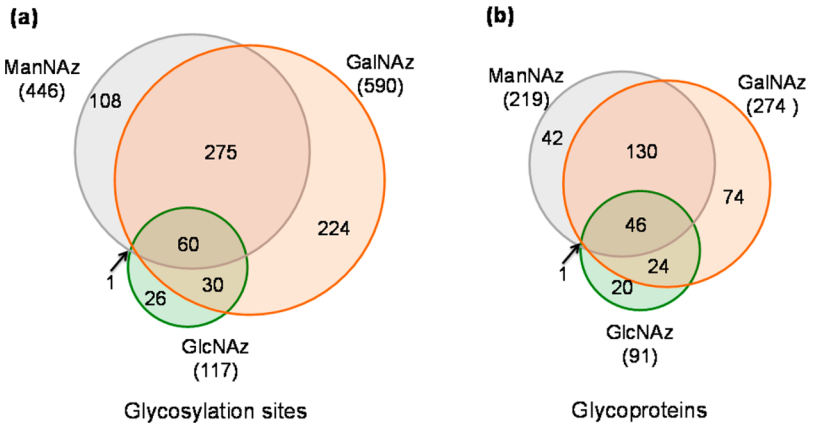


Figure 3. Comparison of (a) surface N-glycosylation sites and (b) N-glycoproteins identified in GalNAz, GlcNAz, and ManNAz labeling experiments.

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

activity, signal transducer activity, and binding, with remarkably low P values (Figure S1). Among biological processes, cell adhesion was prominently enriched with a P value of 3.6×10^{-20} and 48 proteins involved. Integrin-mediated signaling pathway and cell motion were also notably enriched with P values of 1.8×10^{-10} and 4.1×10^{-10}, respectively. These are consistent with the well-known molecular functions and biological processes of surface proteins, which demonstrated that the current method for surface glycoprotein identification is effective.

Quantification of Surface Protein N-Glycosylation Changes in Atorvastatin-Treated HepG2 Cells. Surface glycoproteins play critical roles in cell−cell and cell-matrix interactions. Systematic and quantitative analysis of surface proteins can help us better understand surface protein functions and cellular activities, which will lead to a better understanding of the molecular mechanisms of disease, the discovery of biomarkers, and elucidating the side effects of drugs. Statins are the most popular and effective drugs for lowering patients' cholesterol. As effective cholesterol-lowering HMGCR inhibitors, statins inhibit the rate-limiting step of the cholesterol biosynthesis pathway, known as the mevalonate pathway. It has been extensively documented that these drugs have pleiotropic effects,[52] but their molecular mechanisms remain to be explored. The inhibition of HMGCR also prevents the synthesis of other products in this pathway, including ubiquinone, dolichol, and farnesyl-pyrophosphate (farnesyl-PP). Dolichol is essential to protein N-glycosylation in the form of dolichyl phosphate (Dol-P). Dol-P serves as the carrier in pyrophosphate-linked oligosaccharide assembly as well as acting as the acceptor in the synthesis of the sugar donors Dol-P-Man and Dol-P-Glc from GDP-Man and UDP-Glc, respectively. Upon the inhibition of dolichol, protein N-glycosylation is expected to be dramatically impacted due to the inability to process lipid-linked oligosaccharide biosynthesis and transportation. However, systematic and quantitative analysis of surface N-glycoproteins in statin-treated cells has yet to be reported.

Based on the optimized sugar analog labeling method discussed above, surface protein N-glycosylation changes in atorvastatin-treated cells were analyzed by combining GalNAz labeling and a quantitative proteomics method. Since the primary organ target of statins is the liver, HepG2, a human liver carcinoma cell line was used in this work. Stable isotope labeling by amino acids in cell culture (SILAC)[53] was employed to evaluate the surface N-glycoprotein changes between statin-treated and untreated cells. Cells were treated by atorvastatin for 24 h to inhibit dolichol synthesis and then labeled with GalNAz for another 24 h in the presence of atorvastatin. An in-flask copper-free click reaction with DBCO-sulfo-biotin was then performed to specifically tag surface N-glycoproteins (Figure 4a). Subsequent cell lysis, protein digestion, and enrichment of surface glycopeptides with NeutrAvidin beads were performed as described above. The selectively enriched surface N-glycopeptide samples were analyzed by LC-MS. An example of peptide quantification is shown in Figure 4, and the full MS and elution profile of a peptide (VASVININPN#TTHSTGSCR, where # is the glycosylation site) from LAMP2 (CD107) are shown in Figure 4b and c, respectively. LAMP2 is a single-pass type I membrane protein that regulates cell adhesion and inter/intracellular signal transduction when expressed on the cell surface. Based on the elution profiles, we can very accurately quantify the abundance changes of this

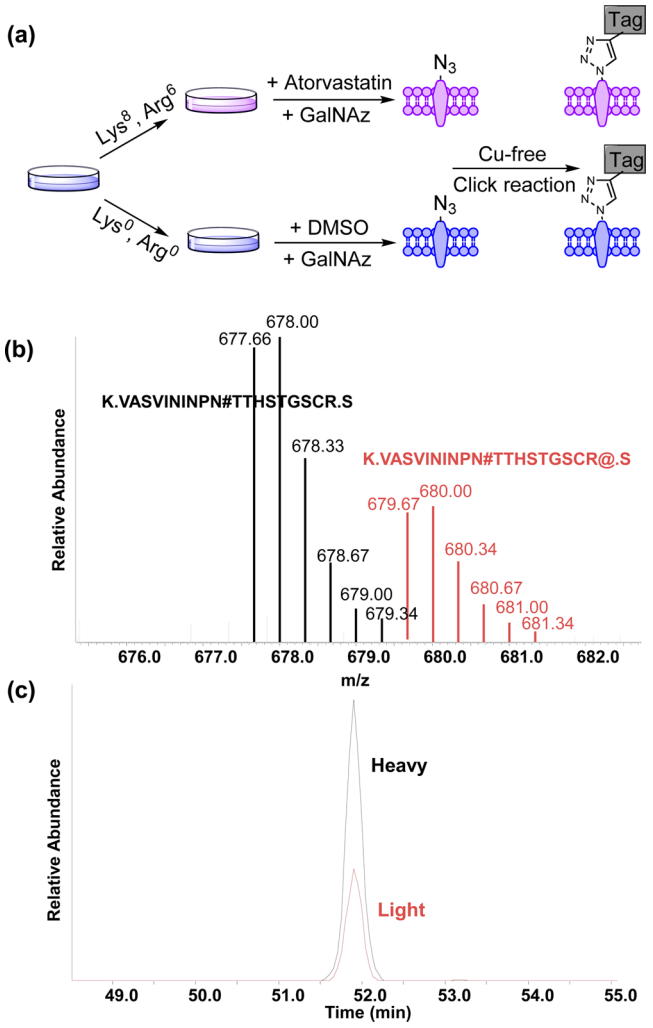


Figure 4. (a) Overview of labeling and tagging workflow in quantification experiments and (b, c) the quantification of the heavy and light versions of an example glycopeptide from LAMP2: (b) full MS (# represents glycosylation site and @ represents heavy arginine) and (c) extracted elution profiles for both versions of the peptides.

peptide in statin-treated cells vs untreated cells, *i.e.* the ratio of the areas under the curves for heavy and light versions of the peptide (H/L = 0.38).

The combination of GalNAz labeling and SILAC led to the quantification of 360 unique N-glycopeptides from 178 cell-surface glycoproteins (Table S4). Among quantified unique glycopeptides, the majority only contained a single N-glycosylation site, while only 20 contained two sites, as shown in Figure 5a. The distribution of 360 quantified unique glycopeptides is shown in Figure 5b. Based on the two criteria described above, 280 singly glycosylated sites (Table S5) were quantified with a similar distribution, as shown in Figure S2. In this quantification experiment, we identified significantly more down-regulated glycopeptides and glycosylation sites (103 sites) than up-regulated glycopeptides and glycosylation sites (37 sites) in atorvastatin-treated cells. Although dolichol biosynthesis was inhibited by atorvastatin for 1 day before GalNAz labeling, dolichol can be recycled in cells after sugar transportation is completed.[54] Therefore, statin treatment for a short time can impact but not entirely prevent protein N-glycosylation. Another possible explanation for site up-regulation could be due to the up-regulation of the

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

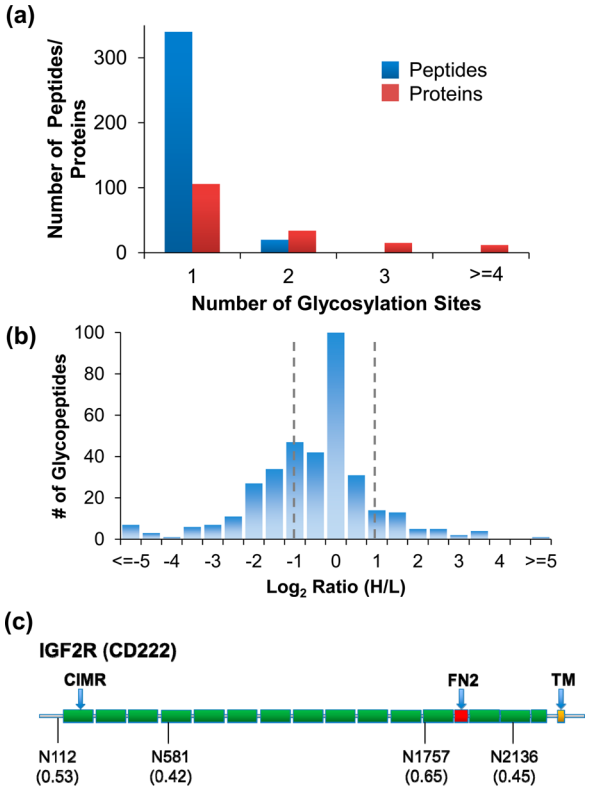


Figure 5. (a) Distribution of the quantified glycosylation sites on each peptide and protein, (b) ratio distribution of quantified unique glycopeptides, and (c) domain analysis of IGF2R and quantified N-glycosylated sites (ratio is shown below each site).

corresponding parent protein. Namely, if a protein is dramatically up-regulated in treated cells while N-glycosylation sites from this protein are largely unaffected or even slightly down-regulated, we could still find these sites up-regulated. For instance, the abundance of the N1523 site on APOB changed by 1.6-fold in the treated cells, whereas the protein ratio was found to be up-regulated by 2.0-fold, as reported previously.[55] Similar effects have been found in protein phosphorylation studies reported in the literature.[56]

Analysis of Down-Regulated Surface N-Glycosylation Sites in Atorvastatin-Treated Cells. Among 84 surface proteins bearing 103 down-regulated glycosylation sites, we performed protein clustering using DAVID 6.7.[51] Glycoproteins in the Alzheimer's disease pathway were highly enriched, and glycoproteins and sites in this pathway are listed in Table 1. Previous studies have shown inconsistent effects of statin use in AD.[57] Some studies found beneficial effects,[58] but others did not.[59] Glycosylation defects in amyloid precursor protein (APP), tau, nicastrin, and other proteins in AD were reported previously,[60] and defective glycosylation may be important in AD pathogenesis. A prior study found that N-glycosylation of human nicastrin was required to interact with lectins, including calnexin and ERGIC-53.[61] In this study, N417 on nicastrin was quantified to be down-regulated by 5.4-fold. Furthermore, proteins participating in immune response processes, such as response to wounding, external stimulus, etc., were also enriched.

We also performed domain analysis to correlate the localization of glycosylation sites and functional domains of proteins.[62] Domains on proteins carry out a wide variety of functions or interactions. Investigating glycosylation site regulations within domains may provide useful information in biological events. For example, CD222, also called IGF2R, is a transporter of phosphorylated lysosomal enzymes from the Golgi complex and the cell surface to the lysosome and has 15 repeating cation-independent mannose-6-phosphate receptor domains (CIMR) (shown in green in Figure 5c). These domains specifically bind the phosphomannosyl residues on lysosomal enzymes. IGF2R also has a fibronectin type II domain (FN2) (shown in red) which serves as the binding site

Table 1. Down-Regulated Glycosylation Sites Quantified from Proteins in the Alzheimer's Disease Pathway (P = 0.027)[a]

Gene Symbol	Peptide	PPM	XCorr	Glycosylation Site	Mod Score	Site Ratio	Annotation
ADAM 17	EQQLESCACN#ETDNSCK	1.89	4.69	594	76.7	0.03	Disintegrin and metalloproteinase domain-containing protein 17
	KCQEAIN#ATCK	-0.99	2.41	539	1000	0.21	
NCSTN	RPN#QSQPLPPSSLQR	4.69	2.61	417	1000	0.18	Nicastrin, Essential subunit of the gamma-secretase complex
LRP1	CIPEHWTCDGDNDCGDYSDETHAN#CTNQATRPPGGCHTDEFQCR	2.01	3.17	1050	5.3	0.24	Prolow-density lipoprotein receptor-related protein 1
	QSGDVTCN#CTDGR	-2.61	1.43	4364	1000	0.27	
	CTQQVCAGYCAN#NSTCTVNQGNQPQCR	0.55	5.40	4278	6.6	0.32	
	CTQQVCAGYCANN#STCTVNQGNQPQCR@	-1.68	4.82	4279	6.6	0.32	
FAS	CKPNFFCN#STVCEHCDPCTK	0.1	4.27	136	28.5	0.19	Tumor necrosis factor receptor superfamily member 6
ITPR1	VESGEN#CSSPAPR@	0.13	2.37	2512	1000	0.13	Inositol 1,4,5-trisphosphate receptor type 1

[a]# - glycosylation site; @ - heavy arginine.

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

for collagens. All these domains are located in the extracellular space, as shown on the left of the transmembrane domain (TM), which is integrated into the cell plasma membrane. We have quantified four glycosylation sites (N112, N581, N1757, and N2136), among which three are localized within the CIMR domains with the other is located at the N-terminal tail. These site abundances decreased to 53%, 42%, 65%, and 45%, respectively, under the statin treatment, which may affect the interactions between this CD and its interactors.

By combining GalNAz labeling, click chemistry tagging, and MS-based proteomics, we found that many glycosylation sites on surface proteins were down-regulated in atorvastatin-treated HepG2 cells. Patients are typically prescribed the drug long-term (months to years). Here, we found that many surface protein glycosylation sites were down-regulated when cells were treated for only 2 days. Further studies, including time-course experiments and animal model experiments, will help us better understand the protein glycosylation changes caused by statin and the molecular mechanisms of its pleiotropic effects.

CONCLUSIONS

Glycosylation changes on cell-surface proteins are hallmarks of many diseases, but global and site-specific analysis of cell-surface N-glycoproteins is extraordinarily challenging. In-depth analyses of the surface glycoproteome changes will potentially lead to clinical applications, such as the identification of diagnostic and therapeutic targets. In this work, we compared labeling with three sugar analogs (GalNAz, ManNAz, and GlcNAz) for the global analysis of surface glycoproteins, in combination with click chemistry tagging, selective enrichment, and MS analysis. The results clearly demonstrated that more protein glycosylation sites on the cell surface were identified with GalNAz labeling compared to GlcNAz or ManNAz. By using GalNAz labeling, surface protein N-glycosylation changes between statin-treated and untreated cells were comprehensively and site-specifically analyzed in combination with quantitative proteomics. Many glycopeptides were down-regulated in statin-treated HepG2 cells compared to untreated cells because statin prevents the synthesis of dolichol, which is essential for the formation of dolichol-linked precursor oligosaccharides. Several N-glycosylation sites on surface proteins related to Alzheimer's disease were found to be down-regulated. Site-specific information regarding surface proteins will provide insight into protein functions and also lead to a better understanding of the molecular mechanisms of statin's pleiotropic effects.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b04871.

Clustering of identified surface N-glycoproteins based on molecular function and biological process (Figure S1); distribution of quantified unique glycosylation sites (Figure S2) (PDF)

Cell-surface *N*-glycosylation sites identified in GalNAz labeling experiments (Table S1) (XLSX)

Cell-surface *N*-glycosylation sites identified in ManNAz labeling experiments (Table S2) (XLSX)

Cell-surface *N*-glycosylation sites identified in GlcNAz labeling experiments (Table S3) (XLSX)

Cell-surface N-glycosylation peptides and sites quantified in atorvastatin-treated cells (Table S4) (XLSX)

Cell-surface N-glycosylation peptides and sites quantified in atorvastatin-treated cells (Table S5) (XLSX)

AUTHOR INFORMATION

Corresponding Author

*Phone: 404-385-1515. Fax: 404-894-7452. E-mail: ronghu.wu@chemistry.gatech.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Weixuan Chen for his valuable input and Johanna Smeekens for critical reading. This work is supported by the National Science Foundation (CAREER Award, CHE-1454501).

REFERENCES

(1) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. *Essentials of glycobiology*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2008.

(2) Gabius, H. J.; Andre, S.; Kaltner, H.; Siebert, H. C. *Biochim. Biophys. Acta, Gen. Subj.* **2002**, *1572*, 165−177.

(3) Dennis, J. W.; Granovsky, M.; Warren, C. E. *BioEssays* **1999**, *21*, 412−421.

(4) Wang, J. Z.; GrundkeIqbal, I.; Iqbal, K. *Nat. Med.* **1996**, *2*, 871−875.

(5) Christiansen, M. N.; Chik, J.; Lee, L.; Anugraham, M.; Abrahams, J. L.; Packer, N. H. *Proteomics* **2014**, *14*, 525−546.

(6) Vigerust, D. J. *Cent. Eur. J. Biol.* **2011**, *6*, 802−816.

(7) Rambourg, A.; Neutra, M.; Leblond, C. P. *Anat. Rec.* **1966**, *154*, 41−79.

(8) Gahmberg, C. G.; Tolvanen, M. *Trends Biochem. Sci.* **1996**, *21*, 308−311.

(9) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125−1128.

(10) Lau, K. S.; Partridge, E. A.; Grigorian, A.; Silvescu, C. I.; Reinhold, V. N.; Demetriou, M.; Dennis, J. W. *Cell* **2007**, *129*, 123−134.

(11) Smeekens, J. M.; Chen, W. X.; Wu, R. H. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 604−614.

(12) Tian, Y. A.; Kelly-Spratt, K. S.; Kemp, C. J.; Zhang, H. *J. Proteome Res.* **2010**, *9*, 5837−5847.

(13) Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D. *Anal. Chem.* **1995**, *67*, 1426−1436.

(14) Wollscheid, B.; Bausch-Fluck, D.; Henderson, C.; O'Brien, R.; Bibel, M.; Schiess, R.; Aebersold, R.; Watts, J. D. *Nat. Biotechnol.* **2009**, *27*, 378−386.

(15) Cannon, J.; Nakasone, M.; Fushman, D.; Fenselau, C. *Anal. Chem.* **2012**, *84*, 10121−10128.

(16) Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. *Nat. Biotechnol.* **2003**, *21*, 660−666.

(17) Sidoli, S.; Lin, S.; Xiong, L.; Bhanu, N. V.; Karch, K. R.; Johansen, E.; Hunter, C.; Mollah, S.; Garcia, B. A. *Mol. Cell. Proteomics* **2015**, *14*, 2420−2428.

(18) Marino, F.; Bern, M.; Mommen, G. P. M.; Leney, A. C.; van Gaans-van den Brink, J. A. M.; Bonvin, A.; Becker, C.; van Els, C.; Heck, A. J. R. *J. Am. Chem. Soc.* **2015**, *137*, 10922−10925.

(19) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. *Cell* **2006**, *127*, 635−648.

(20) Burke, M. C.; Oei, M. S.; Edwards, N. J.; Ostrand-Rosenberg, S.; Fenselau, C. *J. Proteome Res.* **2014**, *13*, 5965−5972.

(21) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G. *Nat. Methods* **2007**, *4*, 798−806.

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332

(22) Huang, H.; Lin, S.; Garcia, B. A.; Zhao, Y. M. *Chem. Rev.* **2015**, *115*, 2376−2418.
(23) Khidekel, N.; Ficarro, S. B.; Clark, P. M.; Bryan, M. C.; Swaney, D. L.; Rexach, J. E.; Sun, Y. E.; Coon, J. J.; Peters, E. C.; Hsieh-Wilson, L. C. *Nat. Chem. Biol.* **2007**, *3*, 339−348.
(24) Jin, M.; Bateup, H.; Padovan, J. C.; Greengard, P.; Nairn, A. C.; Chait, B. T. *Anal. Chem.* **2005**, *77*, 7845−7851.
(25) Chen, W. X.; Smeekens, J. M.; Wu, R. H. *Mol. Cell. Proteomics* **2014**, *13*, 1563−1572.
(26) Chen, W. X.; Smeekens, J. M.; Wu, R. H. *Chemical Science* **2015**, *6*, 4681−4689.
(27) Woo, E. M.; Fenyo, D.; Kwok, B. H.; Funabiki, H.; Chait, B. T. *Anal. Chem.* **2008**, *80*, 2419−2425.
(28) Sidoli, S.; Lin, S.; Karch, K. R.; Garcia, B. A. *Anal. Chem.* **2015**, *87*, 3129−3133.
(29) Phanstiel, D.; Brumbaugh, J.; Berggren, W. T.; Conard, K.; Feng, X.; Levenstein, M. E.; McAlister, G. C.; Thomson, J. A.; Coon, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4093−4098.
(30) Liu, T.; Qian, W. J.; Gritsenko, M. A.; Camp, D. G.; Monroe, M. E.; Moore, R. J.; Smith, R. D. *J. Proteome Res.* **2005**, *4*, 2070−2080.
(31) Liao, J. K. *Hypertension* **2004**, *43*, 1171−1172.
(32) Forbes, K.; Shah, V. K.; Siddals, K.; Gibson, J. M.; Aplin, J. D.; Westwood, M. *Mol. Hum. Reprod.* **2015**, *21*, 105−114.
(33) Burda, P.; Aebi, M. *Biochim. Biophys. Acta, Gen. Subj.* **1999**, *1426*, 239−257.
(34) Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2010**, *21*, 1912−1916.
(35) Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. H. *Chem. Commun.* **2011**, *47*, 6257−6259.
(36) Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F.; van Hest, J. C. M.; van Delft, F. L. *Chem. Commun.* **2010**, *46*, 97−99.
(37) Chen, W. X.; Smeekens, J. M.; Wu, R. H. *J. Proteome Res.* **2014**, *13*, 1466−1473.
(38) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(39) Elias, J. E.; Gygi, S. P. *Nat. Methods* **2007**, *4*, 207−214.
(40) Kall, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. *Nat. Methods* **2007**, *4*, 923−925.
(41) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausolei, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. *Cell* **2010**, *143*, 1174−1189.
(42) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. *Nat. Biotechnol.* **2006**, *24*, 1285−1292.
(43) Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14846−14851.
(44) Mercer, N.; Ramakrishnan, B.; Boeggeman, E.; Verdi, L.; Qasba, P. K. *Bioconjugate Chem.* **2013**, *24*, 144−152.
(45) Nandi, A.; Sprung, R.; Barma, D. K.; Zhao, Y. X.; Kim, S. C.; Falck, J. R.; Zhao, Y. M. *Anal. Chem.* **2006**, *78*, 452−458.
(46) Kaji, H.; Saito, H.; Yamauchi, Y.; Shinkawa, T.; Taoka, M.; Hirabayashi, J.; Kasai, K.; Takahashi, N.; Isobe, T. *Nat. Biotechnol.* **2003**, *21*, 667−672.
(47) Asperger, A.; Marx, K.; Albers, C.; Molin, L.; Pinato, O. *J. Proteome Res.* **2015**, *14*, 2633−2641.
(48) Yao, X. D.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 2836−2842.
(49) Boyce, M.; Carrico, I. S.; Ganguli, A. S.; Yu, S. H.; Hangauer, M. J.; Hubbard, S. C.; Kohler, J. J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3141−3146.
(50) May, P.; Bock, H. H.; Nimpf, J.; Herz, J. *J. Biol. Chem.* **2003**, *278*, 37386−37392.
(51) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. *Nucleic Acids Res.* **2009**, *37*, 1−13.
(52) Liao, J. K.; Laufs, U. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 89−118.
(53) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.
(54) Cantagrel, V.; Lefeber, D. J. *J. Inherited Metab. Dis.* **2011**, *34*, 859−867.
(55) Xiao, H. P.; Chen, W. X.; Tang, G. X.; Smeekens, J. M.; Wu, R. H. *J. Proteome Res.* **2015**, *14*, 1600−1611.
(56) Wu, R. H.; Haas, W.; Dephoure, N.; Huttlin, E. L.; Zhai, B.; Sowa, M. E.; Gygi, S. P. *Nat. Methods* **2011**, *8*, 677−683.
(57) Serrano-Pozo, A.; Vega, G. L.; Lutjohann, D.; Locascio, J. J.; Tennis, M. K.; Deng, A.; Atri, A.; Hyman, B. T.; Irizarry, M. C.; Growdon, J. H. *Alzheimer Dis. Assoc. Disord.* **2010**, *24*, 220−226.
(58) Wong, W. B.; Lin, V. W.; Boudreau, D.; Devine, E. B. *Pharmacoepidemiol. Drug Saf.* **2013**, *22*, 345−358.
(59) Ellul, J.; Archer, N.; Foy, C. M. L.; Poppe, M.; Boothby, H.; Nicholas, H.; Brown, R. G.; Lovestone, S. *J. Neurol., Neurosurg. Psychiatry* **2007**, *78*, 233−239.
(60) Schedin-Weiss, S.; Winblad, B.; Tjernberg, L. O. *FEBS J.* **2014**, *281*, 46−62.
(61) Morais, V. A.; Brito, C.; Pijak, D. S.; Crystal, A. S.; Fortna, R. R.; Li, T.; Wong, P. C.; Doms, R. W.; Costa, J. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2006**, *1762*, 802−810.
(62) Finn, R. D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E. L. L.; Tate, J.; Punta, M. *Nucleic Acids Res.* **2014**, *42*, D222−D230.

DOI: 10.1021/acs.analchem.5b04871
Anal. Chem. 2016, 88, 3324−3332