A Novel Quantitative Mass Spectrometry Platform for Determining Protein O-GlcNAcylation Dynamics*§

Xiaoshi Wang‡, Zuo-Fei Yuan‡, Jing Fan§, Kelly R. Karch‡, Lauren E. Ball¶, John M. Denu§, and Benjamin A. Garcia‡

Over the past decades, protein O-GlcNAcylation has been found to play a fundamental role in cell cycle control, metabolism, transcriptional regulation, and cellular signaling. Nevertheless, quantitative approaches to determine in vivo GlcNAc dynamics at a large-scale are still not readily available. Here, we have developed an approach to isotopically label O-GlcNAc modifications on proteins by producing $^{13}$C-labeled UDP-GlcNAc from $^{13}$C$_6$-glucose via the hexosamine biosynthetic pathway. This metabolic labeling was combined with quantitative mass spectrometry-based proteomics to determine protein O-GlcNAcylation turnover rates. First, an efficient enrichment method for O-GlcNAc peptides was developed with the use of phenylboronic acid solid-phase extraction and anhydrous DMSO. The near stoichiometry reaction between the diol of GlcNAc and boronic acid dramatically improved the enrichment efficiency. Additionally, our kinetic model for turnover rates integrates both metabolomic and proteomic data, which increase the accuracy of the turnover rate estimation. Other advantages of this metabolic labeling method include in vivo application, direct labeling of the O-GlcNAc sites and higher confidence for site identification. Concentrating only on nuclear localized GlcNAc modified proteins, we are able to identify 105 O-GlcNAc peptides on 42 proteins and determine turnover rates of 20 O-GlcNAc peptides from 14 proteins extracted from HeLa nuclei. In general, we found O-GlcNAcylation turnover rates are slower than those published for phosphorylation or acetylation. Nevertheless, the rates widely varied depending on both the protein and the residue modified. We believe this methodology can be broadly applied to reveal turnovers/dynamics of protein O-GlcNAcylation from different biological states and will provide more information on the significance of O-GlcNAcylation, enabling us to study the temporal dynamics of this critical modification for the first time. Molecular & Cellular Proteomics 15: 10.1074/mcp.O115.049627, 2462–2475, 2016.

Glycosylation is one of the most abundant types of protein modification in biology. More than 50% of all proteins have been found to be glycosylated in cells according to the SwissProt database (1). Among the large diversity of protein glycosylation, enzyme-catalyzed O-GlcNAcylation is a type of post-translational modification (PTM)$^1$ with the attachment of a β-D-N-acetylglucosamine moiety to the hydroxyl group of serine or threonine residues. O-GlcNAcylation has been mostly found within the cytoplasm, but it is also found in the nucleoplasm (2). So far, more than 500 proteins have been identified to be O-GlcNAcylated, and these modifications have fundamental roles in cellular structure, cell cycle control, metabolism, and transcriptional regulation (3, 4) Additionally, the mutually exclusive occurrence of O-GlcNAc and protein phosphorylation on the same residues (serine and threonine) indicates competition between these two modifications, which may be responsible for their antagonistic roles in regulation of cellular signaling (5, 6). More importantly, alterations in O-GlcNAcylation have been found to be associated with different kinds of diseases such as cancer and diabetes (7, 8). Therefore, the identification of O-GlcNAc modified proteins and the dynamics of modification will lead to the better understanding of the functions of protein O-GlcNAcylation in human physiology and disease.

O-GlcNAc modification of proteins is also a dynamic process. The addition and removal of the modification is regulated by O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), respectively (9). OGT activity is dependent on the presence of $^1$ The abbreviations used are: PTM, Post-translational modification; OGT, O-GlcNAc transferase; CID, Collision-induced dissociation; ETD, Electron-transfer dissociation; GlcNAc, β-D-N-acetylglucosamine; GalNAc, α-D-N-acetylgalactosamine; OGA, O-GlcNAcase; HBP, Hexosamine Biosynthetic Pathway; PBA, Phenylboronic acid; HCF-1, Host Cell Factor 1; AUC, Area Under Curve; LWAC, Lectin Weak Affinity Chromatography.
its cofactor, UDP-GlcNAc, which is synthesized through the hexosamine biosynthetic pathway (HBP). HBP is a branch of glucose metabolism (Fig. 1A), but is also affected by various other metabolic pathways in the cell (10). About 2–5% of total cellular glucose is used to generate UDP-GlcNAc (2, 11). Therefore, O-GlcNAcylation is highly regulated by cellular glucose availability. For instance, hyperglycemia was found to induce O-GlcNAc modification of CaMKII at Ser279 and enhance CaMKII autonomous activation in cardiac myocytes (12). Not only is O-GlcNAcylation regulated by levels of HBP metabolites, it can also be dynamically controlled by the inhibition of OGT and OGA (13). Inhibitors of either OGT or OGA have been developed to alter protein GlcNAcylation and might have potential clinical applications (14).

Identifying proteins and sites of O-GlcNAc modification has been challenging. Several enrichment methods have been developed over the past years to tackle this challenge. Lectin-proteins have a weak affinity with the O-GlcNAc modification group and have been used to separate and retain O-GlcNAc peptides from unmodified peptides through lectin-agarose packed columns (15, 16). With this method, named lectin weak affinity chromatography (LWAC), a recent study detected 142 O-GlcNAc sites on 62 mouse ESC nuclear proteins (17). Besides the lectin-O-GlcNAc interaction, pan-specific antibodies have also been produced to immunoprecipitate O-GlcNAc containing proteins (18). Two-hundred and eighty-four GlcNAc-modified chromatin proteins were identified with the use of the RL2 antibody and mass spectrometry based proteomics. Chemical methods have also been designed and successfully applied in the enrichment of O-GlcNAc (19).

β-elimination of the O-GlcNAc group followed by Michael addition of a dithiothreitol or biotin tag, or named BEMAD, was developed and applied in many studies. This strategy was further modified by using d₀ or d₆ DTT to discriminate the derivatization between O-phosphate and O-GlcNAc under specific conditions (20). For example, multiple sites on histones H2A, H2B, and H4 were identified to be O-GlcNAcylated with this method (21). Later, a mutant enzyme β-1,4-galactosyltransferase (GaT) was developed to specifically label O-GlcNAc residues with a keto- or azido-derivative of UDP-Galactose. The keto- or azido- group could then be used as a tag to chemically label the modified site with a biotin or fluorescent probe (22, 23). With the utilization of this chemoenzymatic labeling method coupled with mass spectrometry based proteomics, 141 O-GlcNAc-modified sites on 64 proteins were identified from less than 15 μg of HeLa spindle and midbody samples (24). Metabolic labeling to introduce a chemical tag on GlcNAc has also been used in enrichment. These tags include diazirine, azide, or alkyne modification groups added to GlcNAc (25, 26). Upon click reaction or photo-crosslinking, GlcNAc peptides were enriched with beads bearing chemical groups that react with these tags (27). Another chemical method utilized the reversible reaction between cis-1,3-diol in O-GlcNAc and boronic acid to form boronate esters (28).

Mass spectrometry (MS) plays an increasingly important role in quantitative proteomics because of the improvements in accuracy and sensitivity of MS related technologies (29). MS methods have been applied to identify proteins, map modification sites, and quantify changes in peptide abundance. Collision-induced dissociation (CID) fragmentation is the most time efficient MS method to identify large numbers of peptides in global proteomics experiments. However, just as with phosphorylation (30), O-GlcNAc is a labile PTM group. CID fragmentation of O-GlcNAc containing peptides causes partial loss of the GlcNAc moiety, thus losing the important site-specific localization information. Electron-transfer dissociation (ETD), on the other hand, is a different fragmentation method that does not break the bond between the GlcNAc moiety and the amino acid (31). In general, ETD fragmentation of parental ions with a charge state of 3 or higher provide better spectra than those with charge state of 2. Therefore, the combination of CID and ETD has been found advantageous to provide complete information for PTM identification (32, 33).

Different labeling strategies and quantification methods have been developed to characterize the turnover rates and kinetics of protein modifications, such as methylation (34), acetylation (35, 36), and phosphorylation (37). For example, histone methylations associated with gene activation have faster turnover rates than those associated with silent genes (34). Additionally, measurement of PTM turnover rates has become a strategy for determining the efficiency a potential epigenetic drug function (38). For example, in the study of EPZ-5676, a potent and selective aminonucleoside inhibitor of DOT1L histone methyltransferase activity reduced the cellular H3K79me2 levels with a half-time of 1.0 day, which was similar to the turnover rate of methylated H3K79 previously identified (39). For protein O-GlcNAcylation, the dynamic properties and significance have been widely recognized, however, turnover rates of protein O-GlcNAcylation have never been measured by any method before.

Here, we have developed a metabolic labeling method to isotopically label protein O-GlcNAc modifications in vivo by feeding cells with 13C₆-glucose, which produces 13C-labeled UDP-GlcNAc through the hexosamine biosynthetic pathway. 13C-glucose is involved in many cellular pathways, and is metabolized by cells through glycolysis, the Krebs cycle, the pentose phosphate pathway, the pyrimidine biosynthetic pathway, and the hexosamine biosynthetic pathway to incorporate 13C into different portions of the UDP-GlcNAc molecule (40). For example, the GlcNAc moiety in UDP-GlcNAc can be partially labeled as 13C(6)-GlcNAc or 13C(8)-GlcNAc. The six 13C labels on the glucosamine ring are biosynthesized directly from 13C-glucose. The additional two 13C labels are on the acetyl group in GlcNAc moiety. The acetyl group is transferred from acetyl-CoA. Cellular acetyl-CoA becomes
heavily labeled from glucose metabolism, gets shuttled into cytosol and reacts with GlcNH2-6-P. (Fig. 1A) Glucose is also a major source of amino acids, primarily Ala, Asp, and Glu, via glycolysis and Krebs cycle.

In our methods, we used this $^{13}$C$_6$-glucose labeling approach to determine the time course of $^{13}$C incorporation from $^{13}$C$_6$-glucose into UDP-GlcNAc, amino acids and O-GlcNAcylated peptides in HeLa cells. An optimized strategy using phenylboronic acid solid phase extraction cartridge was utilized to enrich O-GlcNAc peptides. This enrichment method improves our ability to detect modified peptides and provides increased confidence in the identification of GlcNAc sites. Next, a mathematic model was developed to determine turnover of GlcNAc modifications on proteins. Overall, our metabolic labeling approach provides a useful quantitative platform for understanding the roles of dynamic protein O-GlcNAcylation in biology.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Isotope-labeling Experiments**—HeLa cells were cultured in DMEM media in which normal glucose was replaced with 25 mM $^{13}$C$_6$-glucose and supplemented with 10% dialyzed FBS and penicillin-streptomycin antibiotics. For protein turnover measurement, HeLa cells were cultured in DMEM media in which normal lysine and arginine were replaced with same amount of lysine($^{13}$C$_6$-$^{15}$N$_2$) and arginine($^{13}$C$_6$-$^{15}$N$_2$) and supplemented with 10% dialyzed FBS and penicillin-streptomycin antibiotics. Cells were incubated at 37 °C, 5% CO$_2$, and harvested after 6, 12, and 24 h after switching and incubating in heavy media. For pulse-labeling experiment, cells were harvested after 0.5, 2, 5, 9, 24, 32, and 48 h incubation. HeLa cells which were grown in 25 cm normal glucose DMEM media were used as zero start time point.

**Cell Fragmentation and Protein Digestion**—Frozen cell pellets were kept on ice to bring up to 4 °C. Cells pellets were resuspended in 10× volume of nuclear isolation buffer (15 mM Tris-HCL pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl$_2$, 1 mM CaCl$_2$, 250 mM Sucrose, 1 mM DTT 1 mM, 1X HALT phosphase inhibitor, 20 mM Thiamet G, 10 mM sodium butyrate, and 0.3% Nonidet P-40), then gently pipetted up and down and incubated on ice for 5 min to break the cell membranes. Trypan blue showed more than 95% staining, indicating that most cells were lysed to nuclei. Fractioned cells were centrifuged 228 × g for 5 min at 4 °C to pellet nuclei. Whole nuclei were isolated and washed with nuclear isolation buffer (15 mM Tris-HCL pH 7.5, 60 mM KCl, 20 mM HEPES, 1X HALT phosphase inhibitor, 20 mM Thiamet G, 10 mM sodium butyrate, 1 mM DTT, and 0.3% Nonidet P-40), then gently pipetted up and down and incubated for 5 min at room temperature overnight. Proteins for turnover measurement were digested overnight with endoproteinase Lys-C (1:50 ratio) instead. Digested peptides were desalted with Oasis HLB cartridges.

**Enrichment of O-GlcNAc Peptides**—PNGased F was used to remove N-glycosylation in nondenaturing conditions. After Lys-C digestion, samples were desalted and dried using a speed-vac. Dried peptides (1 mg) at different time points were resuspended in 50 μL of 1× G7 reaction buffer, pH 7.5. Then, 5.0 μL of PNGaseF (NEB 500 U/μL) was added, and the samples were incubated at 37 °C overnight. Then, O-HexNac peptide enrichment was done by solid state extraction with PBA resin (1:50 ratio) instead. Digested peptides were desalted with Oasis HLB cartridges.

Lastly, bound peptides were eluted with 0.1% TFA solution after incubation at room temperature for 1 h. For the enrichment of O-GlcNAc peptide from a BSA peptide mixture, different solvents as indicated in the text were used.

**Nano Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry**—Desalted O-GlcNAc peptides were analyzed on an Orbitrap Fusion™ Tribrid™ Mass Spectrometer attached to an EasyLC system run at 300 nL/min. Peptides were eluted with a 135 min gradient from 2% to 30% ACN at 120 K and a maximum injection time was 100 ms and AGC of 500,000. Then, a series of MS2 scans were acquired for the most abundant ions from the MS1 scan (mass collection cycle time: 3 s) Ions were filtered with charge 2–7. An isolation window of 2.0 m/z was used with quadruple isolation mode. Ions were fragmented using collision induced dissociation (CID) with collision energy of 27%. Orbitrap detection was used with resolution of 15 K, maximum injection time of 120 ms and AGC 50000. If one or more than one product ions among 204.0867, 186.0761, 168.0555, 144.0565, 138.0550, or 126.0550 were detected, an additional ETD fragmentation on the same parent ions was applied. ETD energy was charge dependent (automatically scaled). Orbitrap detection was used with a resolution of 15 K. The maximum injection time was 120 ms and the AGC was set to 50,000. For protein turnover measurement, a similar LC-MS method was used without the product ion trigger ETD fragmentation.

**Proteomics MS Data Analysis Including Peptide Identification and Quantification**—RAW files were converted to MS1 and MS2 files by xtract 1.0.8 in pFind Studio 2.8 (pfind.ict.ac.cn) and converted to MGF files by pParseto correlate the precursor isotopic mass to the monoisotopic mass and export coeluted precursors, which could increase the identification rate (41). The MS1 and MS2 files were used for quantification and the MGF files were used for database searching and site localization. The Uniprot human database was used for database searching (the version of May 2014 with 88,949 entries). Search parameters were set with the database search engine pFind 2.8 (42), including precursor mass tolerance of 5 ppm, fragment mass tolerance of 0.02 Da for HCD and ETD, Lys-C cleavage and up to 2 miscleavage. Cysteine carboxamidomethylation was set as fixed modification, while methionine oxidation, asparagine and glutamine deamidation, serine, threonine, and asparagine (Asn) N-acetylhexosamine (HexNAc) were set as variable modifications. Asn(N) was used considering the low activity of PNGase F in de-N-glycosylation (43).

For the protein turnover experiment, the following parameters were used: precursor mass tolerance of 5 ppm, fragment mass tolerance of 0.02 Da, trypsin cleavage, and up to 2 miscleavages. Cysteine carboxamidomethylation was set as fixed modification while methionine oxidation was set as a variable modification. The target-decoy approach was used to filter the search results, in which the false discovery rate was less than 1% at the spectral level (44). Before the site localization, the MGF files were preprocessed, including removing the entire precursor isotopic clusters, removing the isotopic peaks of fragment ions, keeping the monoisotopic peaks, and keeping top-6 peaks in every 100 m/z window of each MS/MS spectrum. After the preprocessing, Ascore for HCD spectra and their corresponding ETD spectra were calculated, respectively. Meanwhile, HCD and ETD spectral pairs were combined to calculate Cscore. Annotated MS2 including both HCD and ETD spectra for all identified HexNAc modified peptides are reported in supplemental Table S1 and supporting supplemental Fig. S5.

Then, the peptide and site identification list with GlcNAc modification was used for quantification. The unique peptides with GlcNAc were obtained with precursor m/z, charge state, retention time,
isolation of the desired proteome. We first purified total cell lysates containing each GlcNAc peptide, an isoelectric cluster was obtained, including the monoisotopic peak (noted as M0), the M±6 peak, and the M±8 peak. At different time points, these three peaks were extracted with a close retention time. The area under curve (AUC) for each peak was calculated. Both the isotopic influence of M0 on M±6 and the isotopic influence of M±6 on M±8 were subtracted. Finally, the sum of M±6 AUC and M±8 AUC divided by total M AUC at 0 h was calculated as the [G!] /[G] ratios or relative distributions at each time point.

Metabolite Extraction and MS Analysis of Metabolites—HeLa cells were labeled in DMEM media containing 13C6-glucose. At various time-points after switching to 13C6-glucose, media was rapidly removed and metabolism was quenched by immediate addition of cold extraction solvent (80:20 methanol/water, −80°C). The plate was kept at −80°C freezer for 15 min before the cells were scraped off and collected. The solution was centrifuged at maximum speed at 4°C for 5 min and the supernatant was collected. The pellet was re-extracted twice with 80:20 methanol/water, and supernatant was combined. Samples were dried down under N2 flow, resuspended in water (LC-MS grade, Sigma), and analyzed by a Thermo Q-exactive Orbitrap mass spectrometer coupled to a UPLC (Dionex 3000). Metabolites were separated with an ACQUITY UPLC® BEH C18 2.1 × 100 mm column, 1.7 μm particle size, with a gradient of solvent A (95% H2O, 5% methanol, 10 mM tributylamine, 9 mM acetate, pH = 8.2) and solvent B (100% methanol) at 0.2 ml/min flow rate. The gradient is: 0 min, 5% B; 2.5 min, 5% B; 5 min, 20% B; 7.5 min, 20% B; 13 min, 55% B; 15.5 min, 95% B; 18.5 min, 95% B; 19 min, 5% B; 25 min, 5% B. Data was collected on full scan negative mode at a resolution of 17,500, AGC of 5e5, a maximum injection time of 40 ms and AGC of 1E6. B. Data was collected on negative mode at a resolution of 55% B; 15.5 min, 95% B; 18.5 min, 95% B; 19 min, 5% B; 25 min, 5% B.

Visualization Engine (MAVEN) (45).

Kinetics Modeling—Relative distributions of the labeled UDP-GlcNAc, unlabeled unmodified peptides and O-GlcNAc-peptides were used to determine the turnover rates of the formation of the labeled GlcNAc peptides. The data were simulated according to the kinetic model as shown in Fig. 7 and fitted to differential equations (Eq. 1 and Eq. 2) using MATLAB by least square fitting. The k1 and k−1 in the derivative equation Eq. 2 were fitted using “fmincon” function in Matlab. For each k1 and k−1, sampled by “fmincon” function, the Y values at different time point t were simulated with following equation: $Y(t) = Y(0) + \Delta t(k1[X(t)](U(t)) - k−1[Y(t)])$, where the step time $\Delta t = t_{n+1} - t_n = 0.01$ h, the initial point $Y(t = 0) = 0$, $[X(t)]$ was calculated based on $X(t) = 0.97 - 0.97*exp(-0.082*t)$ and $U(t)$ was calculated based on $U(t) = exp(-t/\lambda_d)$. After obtaining Y(t) curves until $t_n = 24$, the simulated Y value $Y(t_{24}) = 6$, $Y(t_{121}) = 12$ and $Y(t_{2401}) = 24$ were compared with the measured Y values at 6, 12, 24 h, and the root of mean square was calculated to choose the k1 and k−1, which can minimize the difference between simulated Y(t) curves and our measured Y(t) values. Simulation results with k1, k−1, and $R^2$ for all GlcNAc-peptides quantified in this study were shown in supporting supplemental Fig. S4.

RESULTS

Metabolic O-GlcNAc Labeling—To quantitatively determine the rates of protein O-GlcNAcylation, we utilized the hexosamine biosynthetic pathway (HBP) that metabolically produces UDP-GlcNAc, the cofactor of the O-GlcNAc modification reaction as shown in Fig. 1A and 1B. About 2–5% of total glucose (the carbon source of UDP-GlcNAc) is utilized in HBP. By growing cells in 13C6-glucose, heavily labeled UDP-GlcNAc metabolites were produced and accumulated over time. A time course of heavily labeled UDP-GlcNAc production was determined by using two different quantification methods. Supporting supplemental Table S3 shows the relative abundance of all isotopic UDP-GlcNAc species (from M(0) to M(+17)) at different measured time points at the full mass level (MS1). As shown in Fig. 2, the GlcNAc unlabeled fraction, named UDP-12C-GlcNAc-MS1 was calculated from percentages of unlabeled M(0) and those labeled forms with less or equal M(+5) over total UDP-GlcNAc. The use of M(+5) as the cut-off is because the six 13C labels on the glucosamine ring are biosynthesized directly from 13C6-glucose. The GlcNAc labeled fraction, named UDP-13C-GlcNAc-MS1, was calculated from percentages of those labeled forms with equal or more than M(+6) over total UDP-GlcNAc.

In order to distinguish if the 13C label was on the GlcNAc moiety or on the UDP moiety in UDP-GlcNAc molecule, we then fragmented every possible isotopic form of UDP-GlcNAc from 606.07 m/z to 623.13 m/z with 13C(0)- 13C(17) labels. MS2 of UDP-GlcNAc has a distinct fragment ion indicating the UDP moiety, for example, the 384.984 m/z for [UDP-H2O-H]− in the unlabeled UDP-GlcNAc. The number of 13C on the UDP moiety was determined by the mass difference between labeled [UDP-H2O-H]− ion and unlabeled [UDP-H2O-H]− ion at 384.984 m/z. The number of 13C on the GlcNAc moiety was then determined by subtracting the 13C number on UDP moiety from the total 13C number on UDP-GlcNAc molecule. All the possible combinations of parent masses were scanned and fragment masses were adjusted with isotopic abundance and quantified as shown in supplemental Table S3. We did not detect any 13C(3), 13C(4), and 13C(5) labelings on the GlcNAc moiety after isotopic correction. The fraction of labeled UDP-GlcNAc was quantified as the sum of all labeled forms of GlcNAc moiety in UDP-GlcNAc molecule. As shown in supplemental Fig. S1, UDP-13C(6)-GlcNAc and UDP-13C(8)-GlcNAc are the two major species at all times. At the same time of the increasing fraction of cellular 13C(2)-acetyl-CoA over time, 13C(8)-GlcNAc fraction increases and becomes the major metabolite after 3 h.

Finally, we found that MS1 quantification on UDP-GlcNAc and MS2 quantification on GlcNAc moiety gave a very similar result as shown in Fig. 2. The difference is less than 4%. Fitting the MS2 quantification data to an exponential equation $(X(t) = 0.97 - 0.97*exp(-0.082*t))$ allowed the relative abundance of heavy UDP-GlcNAc, annotated as X(t), to be determined at any specific time. At 8.9 h, 50% of the total UDP-GlcNAc became labeled at the GlcNAc moiety. This useful information was used as an input for the determination on protein O-GlcNAcylation turnover rates.

Enrichment and Identification of O-GlcNAc Peptides—Cells harvested at different time points were subfractionated to obtain nuclear fractions. The O-linkage between the sugar
modification and Ser/Thr is labile under peptide CID fragmentation and the GlcNAc residue is generally fragmented. For example, it produces a distinct product ion (204.0867 m/z), and several sequential fragmented product ions, 186.0761 m/z, 168.0655 m/z, 144.0655 m/z, and 126.0550 m/z as seen in Fig. 3, and a neutral loss of the parental ion (-203.0794 m/z). On the other hand, fragmentation by electron transfer dissociation, the radical-based fragmentation approach, leaves the O-GlcNAc group attached to the amino acid residue during peptide backbone fragmentation. Because CID and ETD are complementary, we implemented an approach for O-GlcNAc site localization by the combined use of MS/MS data obtained using both CID and ETD. All precursors were first fragmented with CID, and ETD fragmentation was triggered only when O-GlcNAc related product ions were detected in the initial CID spectra. Both CID and ETD MS/MS spectra were used to identify peptides and O-GlcNAc sites using the peptide spectral matching. CID spectra in combination with ETD spectra boosted the site assignment of the GlcNAc modification. To validate the identified modification sites, we calculated an Ascore, widely used in the phosphoproteomics field, which is calculated based on the information of the O-GlcNAc site determining ions in both CID and ETD spectra. In addition to Ascore, another scoring algorithm, Cscore, was also used to validate the GlcNAc modification sites. Finally, the GlcNAc site assignments were manually checked and ambiguous assignments caused by missing diagnostic c/z/b/y ions were also corrected and reported.

The increased information obtained by using two complementary

**Fig. 1.** A, Hexosamine biosynthetic pathway (HBP) and the metabolic labeling strategy of UDP-GlcNAc from $^{13}$C-glucose. Species that are heavy labeled are marked with an asterisk. B, Modification cycle of O-GlcNAcylation. Protein O-GlcNAcylation reaction involves a UDP-GlcNAc cofactor and two enzymes that are O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) that add and remove the O-GlcNAc moiety, respectively. C, Workflow for the O-GlcNAc peptide analysis, including isotopic labeling, sample preparation, mass spectrometry, and quantification.

**Fig. 2.** The production of $^{13}$C labeled UDP-GlcNAc metabolite from $^{13}$C$_6$-glucose and the reduction of normal light UDP-GlcNAc in HeLa cells over time. Relative abundances of unlabeled and labeled UDP-GlcNAc were quantified by using two different methods. Black and red traces were quantified from MS1 of all isotopic forms of UDP-GlcNAc. Blue and orange traces were quantified from MS/MS fragmentation ions of all isotopic forms of UDP-GlcNAc. Supporting supplemental Table S3 showed all values of the quantification. By fitting the relative abundances of heavily labeled UDP-GlcNAc over time to an exponential equation, a half-time rate of 8.9 h was determined for the production of 50% of GlcNAc labeled UDP-GlcNAc.
tary fragmentation techniques improved the confidence in the O-GlcNAc site-localization.

We enriched O-GlcNAc modified peptides before MS analysis to facilitate identification. It is well known that boronic acids react with cis-1,2- or 1,3 -diols in glycans of glycopeptides, saccharides and many other molecules. In our study, we used PNGase F to remove N-glycosylation, as a large amount of N-glycan might compete with the reaction between O-GlcNAc and phenylboronic acid. (Fig. 1C) However, deamination may occur at this step, so variable modifications have to be considered for downstream peptide identification (46).

O-GlcNAc enrichment was done with PBA solid phase extraction. The condensation between the 1,3-diol and phenylboronic acid generates a water molecule, and removing the water molecule could facilitate condensation and therefore binding, as shown in Fig. 4A. Therefore, instead of using aqueous solution (55 mM H2O), we tested nonaqueous solvents. A GlcNAc-standard peptide TAPT(GlcNAc)STIAPG was mixed with BSA digested peptides. Four different solvents were used in the enrichment including anhydrous DMSO, anhydrous DMF, anhydrous ACN, and pH 8.5, 200 mM ammonium acetate buffer. Peptide mixtures were incubated with PBA at 35 °C for 2 h to allow covalent binding between boronic acid and GlcNAc-peptides, and then the columns were washed to remove unbound peptides. GlcNAc-peptides were eluted by using 0.1% TFA solution. Results indicated that DMSO as the solvent had the highest degree of enrichment among the four solvents we tested. Fig. 4B shows the TIC of the enriched O-GlcNAc peptide from a peptide mixture and compared with the TIC of the input mixture. The O-GlcNAc peptide is the dominant peak in the eluted fraction with 46% retention. The O-GlcNAc standard peptide was enriched 39-fold by using DMSO, whereas it was only enriched fourfold by DMF and twofold by using pH 8.5 buffer. (Fig. 4C) Using ACN as the solvent had no obvious enrichment. We chose seven BSA peptides which have different hydrophobicity and retention times on the C18 column to calculate folds of enrichment. Their peptide identification, charges, m/z, and retention times are listed in the supplemental Table S4. The ratios of O-GlcNAc peptide/BSA peptides were calculated for input, flow-through, and elution fraction for the four different conditions (DMSO, H2O, ACN, and DMF) from the AUC of the extracted ion chromatogram. Folds of enrichments were then calculated and normalized with the ratio of the input and plotted in Fig. 4C. From these three ratios, we calculated the percentage of GlcNAc peptide retention with the column (bound GlcNAc versus total GlcNAc peptide). When using DMSO as the solvent, ~ 46% of GlcNAc peptides are retained and ~ 99% BSA peptides were found in the flow through. However, when water was used as the solvent, almost all of the GlcNAc peptides were retained, but
there is a lot of nonspecific binding of BSA peptides (~42%), which caused a lower fold of enrichment than using DMSO.

It is worth noting that O-GalNAc modification on serine and threonine can also be enriched with PBA strategy as the GalNAc structure bears a 1,3-diol as well. Therefore, we distinguished peptides with O-GalNAc modification from O-GlcNAc modification by observing the oxonium ion fragmentation profile of every identified HexNAc modified peptide and compared with results from a prior study (47). Oxonium ions of identified O-HexNAc modified peptides were extracted in the supplemental Table S1. We calculated ratios of the 138.055/144.066, 186.076/168.066, and 186.076/204.087 m/z ions, respectively. As shown in supplemental Fig. S2, which is a 3D scatter plot of log2 of oxonium ion ratios, HexNAc modified peptides are grouped to two distinct clusters. In our nuclear samples, nine out of 114 peptides (7%) are O-GalNAc modified, which have both positive values of log2 (138.055/144.066) and log2 (204.087/186.076).

Large-scale Incorporation of 13C-GlcNAc Into Peptides as Assessed by Mass Spectrometry—Next, we performed a PBA enrichment of O-GlcNAc peptides from a much more complex peptide mixture extracted from nuclear HeLa cells. In total, we identified 105 O-GlcNAc peptides on 42 proteins from HeLa nuclei. The sequences of the peptide with modification sites identified and the description of the proteins are summarized in supplemental Table S1. Among the O-GlcNAc peptides, one of the most remarkable and heavily O-GlcNAc modified proteins is host cell factor C1 (HCF-1). HCF-1 is a transcriptional coregulator involved in regulating cell cycle progression. HCF-1 undergoes an unusual proteolytic maturation, and the basic region of N-terminal HCF-1 (HCF-1N) is O-GlcNAcylated at multiple sites. O-GlcNAc modification is necessary for HCF-1 maturation and might also be involved in regulating HCF-1 function after its proteolytic cleavage (48, 49). Several modification sites have been previously reported.
Mass Spectrometry Method Determines Protein O-GlcNAcylation Rates

with other enrichment methods, such as LWAC and enzymatic-chemical derivatization enrichment methods (48, 50). In addition to HCF-1, nuclear pore complex proteins (NUPs) were highly modified, including NUP153, NUP214, and NUP98. Indeed, NUPs are observed to be one of the most heavily O-GlcNAc-modified proteins and were also some of the first O-GlcNAc-modified proteins identified in prior studies (51–53). Many transcription regulators were also detected as being O-GlcNAc modified, including protein EMSY, Polyhomeotic-like protein 3 (PHC3), Mxs2-interacting protein, transcription factor MaF, transcription initiation factor TFIID, Protein SON, SAP30-binding protein. O-GlcNAc modifications on transcription-regulating proteins can influence DNA and RNA binding, transcriptional activity, localization, stability, and interactions with other proteins (5). Finally, we found that no obvious consensus sequences of O-GlcNAc modification were identified by motif analysis; and this may be because of the fact that only one enzyme (OGT) is able to modify serine and threonine with GlcNAc.

Interestingly, about 20% of total identified O-GlcNAc peptides have more than one GlcNAc modification site on the same peptides. supplemental Fig. S3 shows one example peptide, KGTTTNTSTTTTTASTADAQK_5,T(HexNAc[T]) on T3 or T4 from SAP30-binding protein with the incorporation of two labeled O-GlcNAc PTMs. The monoisotopic mass at 914.786 m/z is the peptide containing one light and one heavy labeled GlcNAc. The monoisotopic mass at 917.461 m/z is the peptide containing two heavy labeled O-GlcNAc.

We were able to confirm by nanoLC-MS/MS analysis that O-GlcNAc peptides were metabolically labeled with 13C-GlcNAc. As shown in Fig. 5B, the MS1 between pre-existing light O-GlcNAc peptides and the heavy 13C-labeled O-GlcNAc has a mass difference of 6.020 Da or 8.024 Da, corresponding to 13C(6)-GlcNAc or 13C(8)-GlcNAc modified peptide, respectively. The cause of 13C labeling is the metabolic production of 13C-acetyl-CoA from 13C-glucose. Glucose is converted via glycolysis in the cytosol to pyruvate, and then to acetyl-CoA in the mitochondrion. Citrate is produced from excess acetyl-CoA and exported into the cytosol to give rise to cytosolic and nuclear acetly-CoA, and acetyl-CoA provides the acetyl group for the GlcNAc moiety, adding two heavy carbons to the heavy hexose group for a total of 8 13C carbons (35) (Fig. 1A). Similar to the UDP-GlcNAc labeling, O-GlcNAc modification on the peptide can be either 13C(6)-GlcNAc or 13C(8)-GlcNAc. Therefore, when we quantified the fractions of heavily labeled GlcNAc-peptide, we used the sum of both 13C(6)-GlcNAc-peptide and 13C(8)-GlcNAc-peptide.

Pairs of heavy and light labeled O-GlcNAc peptides coeluted at the same retention time as shown in Fig. 5A. This coelution is helpful for the identification of the heavy O-GlcNAc peptides as well as the quantification of their relative abundance. Additionally, we confirmed the 13C-GlcNAc labeling by comparing tandem mass spectra between the unlabeled peptide and its labeled counterpart. Fig. 5C shows that the O-GlcNAc labeled ions ([M+2H]2+) from the MS2 of the heavy peptide have an 8 Da shift in mass relative to the MS2 spectrum of the light species. O-GlcNAc is a labile modification under CID fragmentation, and the ions, such as [M+2H-GlcNAc]2+, [b-GlcNAc]− and [y-GlcNAc]− are observed with a neutral loss of O-GlcNAc. Those fragment ions have the same masses as the corresponding light fragment ions.

In addition, oxonium ions also have signature product ions at the low m/z region of the HCD MS2 spectra. In Fig. 5C, the oxonium ion peak at 204.09 m/z in the spectrum of light peptide compared with the heavy oxonium ion peak at 212.11 m/z in the spectrum of labeled peptide shows an 8-Da mass shift. This information clearly indicates the 13C labeling happens at the GlcNAc moiety. Fig. 3 shows all GlcNAc signature product ions that were detected in the low region of paired light and heavy labeled MS2. One fragmentation route of the GlcNAc oxonium ion proceeds by a stepwise elimination of two water molecules, first generating an ion with a mass of 186.0761 m/z, then an aromatic ring species with a mass of 168.0655 m/z. It is followed by the loss of a formaldehyde molecule (originally from the hydroxymethyl group) to produce an ion with a mass of 138.0550 m/z. Another route is to eliminate a ketene (originating from the acetyl group) from the 186.0761 m/z ion to form a species with a mass of 144.0655 m/z, followed by a loss of a water molecule, resulting in another aromatic species with a mass of 126.0550 m/z. For the 13C labeled GlcNAc oxonium ions, the fragmentation pathways are the same. The relative intensities of different products ions are very similar to the light ones. The only differences are the expected mass shifts of product ions. The labeled ones are shifted toward the heavier molecular weight based on the number of carbons in the molecule. For example, there are 8 carbons in the GlcNAc oxonium ion, the mass shifts from 204.086 m/z to 212.113 m/z. But there are only seven carbons in 3-acetamidopyruvylion (138.0550 m/z). Therefore, a peak at 145.078 m/z was detected in the MS2.

Turnover Rate Determination—Identified O-GlcNAc containing peptides were aligned across all the time points based on the m/z of the precursor ions within a narrowed retention time window as shown in Fig. 6A. The heavy GlcNAc-peptide fraction, [G’]/[G]z, was determined by the sum of AUC(M/z-6)/AUCM and AUC(M-6)/AUCM after adjustment of natural abundances. AUC(A-8) and AUC(A-6) were determined at different labeling time points. AUCM was determined from time zero before labeling which is the total amount of O-GlcNAc peptide in the sample. Samples at different time points were prepared with same conditions all the time and injected the same amount. Because of the complexity of multiple mass shifts, we did not quantify turnover rates of those O-GlcNAc-peptides with more than one O-GlcNAc sites.

Fig. 6B is an example showing the relative abundance of heavily labeled GlcNAc peptides ([G’]/[G]z) increasing over time. This peptide is from NUP153 and the sequence is FGVSSSSSGPQRSTGFK with 1 O-GlcNAc modification.
at Ser/Thr/Asn of the peptide. Before metabolic labeling, no heavy labeled GlcNAc peptide could be detected. After six hours, about 27% of this peptide became labeled, reaching 40% after 24 h in heavy glucose media. Among the 105 O-GlcNAc peptides, we were able to quantify 20 paired heavily labeled GlcNAc sites in 14 proteins identified across all time points, and for each of these, the abundances of heavy GlcNAc peptides increase over time. The turnover rates and relative abundances of heavy labeled species are all summarized in supplemental Table S2.

\[ \frac{[G^*]}{[G]} \] fractions at different time points were fitted to a model representing isotopic labeling under nonsteady state conditions (Fig. 7). As expressed in Eq. 1, the concentration change of labeled O-GlcNAc peptide \([G^*]\) can occur in two ways:

\[ [G^*] \rightarrow [G] \]

\[ [G] \rightarrow [G^*] \]

The possible GlcNAc site is colored red in the sequence on top of panels. B, MS1 spectra for light and heavy labeled peptide containing GlcNAc. Heavy labeled peptides had mass shifts of 6.020 Da and 8.024 Da compared with light peptides, corresponding to the theoretic mass differences of 6.018 Da and 8.024 Da including 6 and 8 \(^{13}\)C atoms, respectively. The isotopic distribution of heavy peptides suggests that both species exist. C, Tandem mass spectra for the heavy in green and light in red labeled peptides with GlcNAc modification. Fragment ions with the GlcNAc loss have the same mass as the corresponding light one. Depending on the position of the GlcNAc, the b and y ions might also be b/y ions with glucan loss. Therefore, we label them as b’ or y’ representing either b/y ion peaks or b/y with glucan neutral loss ion peaks. An 8-Da shift is present in GlcNAc containing ions such as [M+2H]^{2+} and GlcNAc oxonium product ion.

Fig. 5. A, Extracted ion chromatograms showed that light and heavy O-GlcNAc peptides from NUP153 with the sequence of FGVSSTSSPSQTLSTGNFK, 1GlcNAc on Ser/Thr/Asn were eluted at the same retention time. The possible GlcNAc site is colored red in the sequence on top of panels. B, MS1 spectra for light and heavy labeled peptide containing GlcNAc. Heavy labeled peptides had mass shifts of 6.020 Da and 8.024 Da compared with light peptides, corresponding to the theoretic mass differences of 6.018 Da and 8.024 Da including 6 and 8 \(^{13}\)C atoms, respectively. The isotopic distribution of heavy peptides suggests that both species exist. C, Tandem mass spectra for the heavy in green and light in red labeled peptides with GlcNAc modification. Fragment ions with the GlcNAc loss have the same mass as the corresponding light one. Depending on the position of the GlcNAc, the b and y ions might also be b/y ions with glucan loss. Therefore, we label them as b’ or y’ representing either b/y ion peaks or b/y with glucan neutral loss ion peaks. An 8-Da shift is present in GlcNAc containing ions such as [M+2H]^{2+} and GlcNAc oxonium product ion.

Mass Spectrometry Method Determines Protein O-GlcNAcylation Rates

Molecular & Cellular Proteomics 15.7
ways: (1) the formation of new labeled O-GlcNAc peptide by O-GlcNAcylation using labeled UDP-GlcNAc*; and (2) the decay of old labeled GlcNAc peptide by protein degradation or de-GlcNAcylation. Flux is the rate of a molecule concentration change through a biochemical pathway. Here, in Eq. 1, \( F_{OGT}^{*} \) is the formation flux of O-GlcNAc peptide catalyzed by OGT. \( F_{OGA}^{*} \) is the de-GlcNAcylation flux of the GlcNAc peptide catalyzed by OGA. \( F_{Deg}^{*} \) is the degradation flux of the O-GlcNAc peptide through protein degradation.

\[
\frac{d[G^*]}{dt} = \frac{F_{OGT}^{*} [UDP - GlcNAc^*]}{[UDP - GlcNAc]^*} \times \left[ \frac{[U]}{[U]^*} - \frac{F_{OGA}^{*} + F_{Deg}^{*}}{[G]^*} \right]
\]

(Eq. 1)

In our labeling experiment, the building blocks of newly synthesized proteins are partially from labeled amino acids that are metabolized from \(^{13}\)C\(_6\)-glucose. We also measured the \(^{13}\)C-labeling of amino acids from \(^{13}\)C\(_6\)-glucose by extracting metabolites from cells at different time points and using LC-MS based metabolomics. Data are summarized in supplemental Table S3. We found the relative abundances of unlabeled amino acid decay exponentially over time. By fitting the relative abundances over time, the easiest amino acid to label is alanine with a \( t_{1/2} \) of 0.6 h. Because of the fast turnover and nearly complete labeling of alanine, if there is at least one alanine in the peptide sequence, we assume the newly synthesized peptide is labeled. Because the quantification of turnover rates only considers old peptide, \( F_{OGT}^{*} \) is deconvoluted from the synthesis of new peptide with a modification factor of \( \frac{[U]}{[U]^*} \). This factor was determined from a pulse-labeling experiment. In the experiment, HeLa cells were incubated with DMEM replaced with lysine(\(^{13}\)C\(_6\), \(^{15}\)N\(_4\)) and arginine (\(^{13}\)C\(_6\), \(^{15}\)N\(_4\)). Cells were harvested at 0.5, 2, 5, 9, 24, 32, and 48 h. Nuclear proteins were isolated and digested. The relative abundances of old protein were determined at different time points. The factor \( \frac{[U]}{[U]^*} \) was then determined by fitting relative abundances of old protein over time with an exponential equation. All turnover rates were summarized in supplemental Table S2. We did not determine the O-GlcNAc turnover rates if the peptide turnover was not available from the pulse-labeling experiment. Other amino acids were only partially labeled with much slower rates than Ala, for example, Asp has a \( t_{1/2} \) of 1.6 h with 15% heavy, Glu has a \( t_{1/2} \) of 1.2 with 12% heavy, and Ser has a \( t_{1/2} \) of 1.3 with 11% heavy.
Mass Spectrometry Method Determines Protein O-GlcNAcyltion Rates

Additionally, Asn, Pro, and Gly were also labeled, but the labeling rates were even slower. Therefore, if no Ala in the peptide sequence, we assume \([\text{U}/\text{UT}] = 1\). For example, the peptide VIIVTSPSTFVPLNSK\(_1\)HexNAc on any Ser/Thr/Asn has no Ala. Although there is no protein EMSY measured, we are still able to determine the GlcNAc turnover rates by considering \((\text{U}/\text{UT}) = 1\).

Because what we measured in the experiment are the relative abundance of heavy UDP-GlcNAc, annotated as \(X(t)\), the relative abundance of heavy GlcNAc peptide \([\text{G}\,^1]/[\text{G}\,^0]\), also annotated as \(Y(t)\), and the relative abundance of old peptide, annotated as \(U(t)\), Eq. 1 is converted to Eq. 2 by dividing the total GlcNAc peptide concentration \([G]\). Simulation results in O-GlcNAc formation rate \(k_1\) and decay rate \(k_{-1}\), where \(k_1 = F_{\text{UGA}}^T/[\text{G}\,^0], k_{-1} = (F_{\text{OGA}}^T + F_{\text{DEG}}^T)/[\text{G}\,^0].\) Supplemented Fig. S4 shows the simulation results with turnovers for all GlcNAc-peptides quantified in this study. And supporting Eq. 1–7 illustrates the derivation of Eq. 1 and Eq. 2.

\[
\frac{dY(t)}{dt} = k_1[X(t)\,U(t)] - k_{-1}[Y(t)] \quad (\text{Eq. 2})
\]

The simulation fits the observation well for most measured peptides as shown in Fig. 8 and in supplemented Fig. S4, suggesting the model is sufficient for explaining the turnover of GlcNAc PTMs. We determined the \(k_1\) for the detectable O-GlcNAc peptides within a range from 0.02 h\(^{-1}\) to 1.6 h\(^{-1}\). Turnover rates are different for proteins or even different modified sites on the same protein. For example, turnover rates of the dynamic HCF-1 O-GlcNAc sites are different from one peptide to another as shown in Fig. 8.

DISCUSSION

In this study, we first optimized the enrichment of GlcNAc peptides using boronic acid solid phase extraction with nonaqueous DMSO as the solvent. Boronic acids are known to bind with diol groups through boronate ester formation. Two different forms of boronic acid exist and react with diols. One is the trigonal neutral form and the other is the tetrahedral negative charged form, and both forms have affinity toward GlcNAc peptides. The tetrahedral form has a larger rate constant of ester formation than the neutral form according to literature (54). Additionally, both the generation of the tetrahedral negative charged boronic acid and incubation with samples require basic conditions. However, GlcNAc modifications are not stable in basic conditions. Therefore, we optimized PBA enrichment method using nonaqueous solvent. The optimized PBA enrichment method is able to keep intact HexNAc PTM on the peptide, allowing us to study the HexNAc isomerization from ratios of oxonium ions. The optimized PBA enrichment using neutral DMSO as the solvent is also more efficient than PBA enrichment by using basic aqueous solution. This result could be explained by the fact that nonspecific binding caused by trigonal neutral boronate was less than the tetrahedral negative charged boronate anions within a mixture of peptides.

Next, we demonstrated that our novel \(^{13}\text{C}_6\)-glucose metabolic labeling method can be used to label endogenous O-GlcNAc sites on proteins in complex cell lysates. To our knowledge, this is the first large-scale attempt to specifically label O-GlcNAc modifications on proteins and to determine the protein O-GlcNAcylation rates. It is interesting to note that the most dominant fraction of total proteins identified with dynamic O-GlcNAc sites are involved in biological process that regulate transcription.

We compared the turnover rates of protein O-GlcNAc modification with other types of modifications, such as serine/threonine phosphorylation and lysine acetylation. In general, turnovers of GlcNAc in proteins are slow \((k_1\, \text{from} 0.02 \text{ h}^{-1} \text{ to } 1.6 \text{ h}^{-1})\), compared with turnover rates obtained from a large scale quantification experiment on protein phosphorylation \((0.06 \text{ h}^{-1} \text{ to } 2.40 \text{ h}^{-1})\) (37) and histone acetylation \((0.3–1 \text{ h}^{-1})\) (35, 36). It is also worth noting that we observed that the production of heavy UDP-GlcNAc is much slower than the production of heavy acetyl CoA. The half time of generating heavy acetyl CoA is less than 1 h. As the substrate for O-GlcNAc modification is the metabolite (UDP-GlcNAc) directly produced from \(^{13}\text{C}\)-glucose source, the rates of protein GlcNAc labeling can be heavily dependent on the rates of

![Fig. 8](image-url)
metabolite labeling. Therefore, in order to accurately quantify the rate of protein PTM turnover from metabolic labeling, it is important to consider the contribution of the labeling kinetics of the direct metabolite precursor. This is especially critical when the metabolism labeling process is slow, such as with UDP-GlcNAc. Sophisticated compartment kinetic models based on metabolic isotopic labeling for protein turnover studies have been reported before, which led to a more satisfactory fitting of experimental data (56). To our knowledge, our study is the first to determine protein PTM turnover rates with an additional integration of metabolite kinetics.

It is also worth noting that the factor of $\frac{[U]}{[UT]}$ in the model from $^{13}$C$_6$-glucose labeling experiment is theoretically not the same as fractions of the old peptide determined from the pulse-labeling experiment in which lysine and arginine were labeled. But from our metabolomics measurement, we found the labeling of Ala from $^{13}$C$_6$-glucose is really fast ($t_{1/2} = 0.6$ h). At our first experimental time points which was 6 h, 70% Ala were already heavy labeled. So we made this assumption that if there is at least one Ala in the peptide sequence, newly synthesized peptides are labeled. A similar analysis was used as in a study that determined histone acetyl PTM turnovers labeled from $^{13}$C$_6$-glucose (35). With this assumption, we used the turnover rates determined from lysine and arginine pulse-labeling experiment to calculate the factor of $\frac{[U]}{[UT]}$.

Additionally, for the term $\frac{[G^*]}{[G]}$, which represents the fraction of old peptide labeled at the GlcNAc modification, the kinetic model was not further adjusted with amino acid labeling. We can statistically calculate the chance of M+6 or M+8 caused by amino acids labeling. The easiest scenario to cause M+6 labeling in a given peptide is to label two Ala. With an average peptide length of 15 residues and 5 amino acids that can be labeled from glucose, M+6 labeling is resulted from fully labeling of two Ala and nonlabeling of the rest 13 residues. From the metabolomics results on amino acids, at steady state of amino acid labeling, labeled Ala had a percentage of ~80%. The averaged percentage of nonlabeled nonessential amino acids was 80%. Therefore, the chance of M+6 labeling in newly synthesized peptides was $0.8^2 \times 0.8^5$, which is around 20%. Because our method is enriched for O-GlcNAc peptides and the identified O-GlcNAc-proteins have slow turnovers, the contribution of labeled peptide with labeling on amino acids instead of GlcNAc moiety to total M+6 signals is presumably even lower. If more than two Ala in a sequence, only two out of the total Ala labeled, the rest nonlabeled Ala decreases the chance of M+6 labeling. Last, Lys-C digestion usually results in long peptides, a longer peptide and more nonessential amino acids in the sequence decrease the chance too. Although the small error caused by Ala labeling with the method, this LC-MS/MS metabolomics and proteomics-based $^{13}$C$_6$-glucose metabolic labeling method is still good for most of the O-GlcNAc-peptides and is able to provide us novel information on O-GlcNAc-protein turnover rates.

In summary, we have developed an approach to isotopically label O-GlcNAc on proteins by producing $^{13}$C-labeled UDP-GlcNAc through the hexosamine biosynthetic pathway from $^{13}$C$_6$-glucose. This metabolic labeling was combined with quantitative mass spectrometry-based proteomics to determine protein O-GlcNAcylation turnover rates. By integrating metabolite isotopic labeling with protein O-GlcNAc labeling, we determined protein O-GlcNAcylation kinetics in vivo for nuclear proteins. In total, we identified kinetics of 20 O-GlcNAc peptides in from 14 proteins extracted from HeLa nuclei with turnover rates varying from 0.02 h$^{-1}$ to 1.6 h$^{-1}$. We expect that this labeling methodology can be broadly applied to reveal O-GlcNAcylation dynamics from different biological states, adding important temporal aspects to these types of biological studies.

* BAG acknowledges funding of NIH Innovator grant (DP2OD007447) from the Office of the Director, NIH grant R01GM11017, and the National Science Foundation (NSF) Early Faculty Career award. JMD gratefully acknowledges funding from NIH grant (GM059785–15/P250VA). LEB acknowledges funding from NIH/NIDCR grant (R01 DE020925). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains supplemental material.

To whom correspondence should be addressed: Epigenetics Prog. Program, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Room 9-124, 3400 Civic Center Blvd, Bldg 421, Philadelphia, PA 19104. Tel.: 1-215-573-9423; Fax: 215-573-4764; E-mail: bgarci@mail.med.upenn.edu.

REFERENCES

1. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta 1473, 4–8
2. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. Nature 446, 1017–1022
3. Bond, M. R., and Hanover, J. A. (2013) O-GlcNAc cycling: A link between metabolism and chronic disease. Annu. Rev. Nutr. 33, 205–229
4. Wells, L., Vosseller, K., and Hart, G. W. (2001) Glycosylation of nucleocytoplasmic proteins: Signal transduction and O-GlcNAc. Science 291, 2376–2378
5. Slawson, C., and Hart, G. W. (2011) O-GlcNAc signalling: implications for cancer cell biology. Nat. Rev. Cancer 11, 678–684
6. Hu, P., Shimoji, S., and Hart, G. W. (2010) Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. FEBS Lett. 584, 2526–2538
7. Ma, Z., and Vosseller, K. (2013) O-GlcNAc in cancer biology. Amino Acids 45, 719–733
8. Ma, J. F., and Hart, G. W. (2013) Protein O-GlcNAcylation in diabetes and diabetic complications. Expert Rev. Proteomics 10, 365–380
9. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate. Nature 469, 564–567
10. Ying, H., Kimmelkamp, A. C., Lyssiotis, C. A., Hua, S., Chu, G. C., Fletcher-Sanokone, E., Locasale, J. W., Son, J., Zhang, H., Coloff, J. L., Yan, H., Wang, W., Chen, S., Viale, A., Zheng, H., Paik, J. H., Lim, C., Guimarães, A. R., Martin, E. S., Chang, J., Hezel, A. F., Perry, S. R., Hu, J., Gan, B., Xiao, Y., Asara, J. M., Weissleder, R., Wang, Y. A., Chin, L., Cantley, L., and DePinho, R. A. (2012) Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. Cell 149, 656–670
29. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007)
36. Yu, S. H., Boyce, M., Wands, A. M., Bond, M. R., Bertozzi, C. R., and
35. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R.
33. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
32. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
31. McClain, D.A., and Crook, E. D. (1996) Hexosamines and insulin resistance. 
28. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
27. Hahne, H., Sobotzki, N., Nyberg, T., Helm, D., Borodkin, V. S., van Aalten,
26. Yu, S. H., Boyce, M., Wands, A. M., Bond, M. R., Bertozzi, C. R., and
25. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R.
24. Gross, B. J., Kraybill, B. C., and Walker, S. (2005) Discovery of O-GlcNac 
23. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
22. Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S.,
21. Myers, S. A., Panning, B., and Burlingame, A. L. (2011) Polycomb repres- 
20. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
19. Banerjee, P. S., Hart, G. W., and Cho, J. W. (2013) Chemical approaches to 
18. Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S.,
17. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
16. Ma, Z. Y., Skorobogatko, Y., and Vosseller, K. (2013) Tandem lectin weak 
15. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
14. Gross, B. J., Kraybill, B. C., and Walker, S. (2005) Discovery of O-GlcNac 
13. Gloster, T. M., and Vocadlo, D. J. (2010) Mechanism, structure, and inhi- 
12. Erickson, J. R., Pereira, L., Wang, L. G., Han, G. H., Ferguson, A., Dao, K.
11. McClain, D.A., and Crook, E. D. (1996) Hexosamines and insulin resistance. 
10. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
9. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
8. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
7. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A.
6. Yu, S. H., Boyce, M., Wands, A. M., Bond, M. R., Bertozzi, C. R., and
5. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
4. Gross, B. J., Kraybill, B. C., and Walker, S. (2005) Discovery of O-GlcNac 
3. Clark, P. M., Dweck, J. F., Mason, D. E., Hart, C. R., Buck, S. B., Peters,
2. Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S.,
1. McClain, D.A., and Crook, E. D. (1996) Hexosamines and insulin resistance. 

Molecular & Cellular Proteomics 15.7

2474
50. Myers, S. A., Daou, S., Affar, el, B., and Burlingame, A. (2013) Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments—a study of the O-GlcNAcylated protein Host Cell Factor C1. *Proteomics* **13**, 982–991

51. Davis, L.I., and Blobel, G. (1987) Nuclear-pore complex contains a family of glycoproteins that includes P62 - glycosylation through a previously unidentified cellular pathway. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7552–7556

52. Hanover, J. A., Cohen, C. K., Willingham, M. C., and Park, M. K. (1987) O-linked N-acetylglucosamine is attached to proteins of the nuclear-pore - evidence for cytoplasmic and nucleoplasmic glycoproteins. *J. Biol. Chem.* **262**, 9887–9894

53. Snow, C. M., Senior, A., and Gerace, L. (1987) Monoclonal-antibodies identify a group of nuclear-pore complex glycoproteins. *J. Cell Biol.* **104**, 1143–1156

54. Springsteen, G., and Wang, B. H. (2002) A detailed examination of boronic acid-diol complexation. *Tetrahedron* **58**, 5291–5300

55. Rogers, L. D., and Foster, L. J. (2009) Phosphoproteomics-finally fulfilling the promise? *Mol. Biosyst.* **5**, 1122–1129

56. Guan, S., Price, J. C., Ghaemmaghami, S., Prusiner, S. B., and Burlingame, A. L. (2012) Compartment modeling for mammalian protein turnover studies by stable isotope metabolic labeling. *Anal. Chem.* **84**, 4014–4021