Published in final edited form as:

Anal Chem. 2019 March 19; 91(6): 4195–4203. doi:10.1021/acs.analchem.9b00441.

Enzymatic Tagging of Glycoproteins on the Cell Surface for Their Global and Site-Specific Analysis with Mass Spectrometry

Fangxu Sun, Suttipong Suttapitugsakul, 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

Abstract

The cell surface is normally covered with sugars that are bound to lipids or proteins. Surface glycoproteins play critically important roles in many cellular events, including cell–cell communications, cell–matrix interactions, and response to environmental cues. Aberrant protein glycosylation on the cell surface is often a hallmark of human diseases such as cancer and infectious diseases. Global analysis of surface glycoproteins will result in a better understanding of glycoprotein functions and the molecular mechanisms of diseases and the discovery of surface glycoproteins as biomarkers and drug targets. Here, an enzyme is exploited to tag surface glycoproteins, generating a chemical handle for their selective enrichment prior to mass spectrometric (MS) analysis. The enzymatic reaction is very efficient, and the reaction conditions are mild, which are well-suited for surface glycoprotein tagging. For biologically triplicate experiments, on average 953 N-glycosylation sites on 393 surface glycoproteins per experiment were identified in MCF7 cells. Integrating chemical and enzymatic reactions with MS-based proteomics, the current method is highly effective to globally and site-specifically analyze glycoproteins only located on the cell surface. Considering the importance of surface glycoproteins, this method is expected to have extensive applications to advance glycoscience.

Graphical Abstract

Surface glycoproteins are essential for cells, and they regulate many cellular events, including cell–cell communication, cell signaling, and immune defense.^{1,2} Aberrant glycosylation is often correlated with human diseases.^{3,4} Besides its impacts on cell adhesion and migration, changes in glycosylation of receptors on the cell surface can alter

*Corresponding Author: Phone: 404-385-1515. Fax: 404-894-7452, ronghu.wu@chemistry.gatech.edu.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.analchem.9b00441](https://doi.org/10.1021/acs.analchem.9b00441).

The authors declare no competing financial interest.

cell signaling and help tumor cells escape from immune surveillance.^{2,5} Therefore, surface glycoproteins can serve as effective biomarkers for cancer diagnosis.^{6,7} Furthermore, glycoproteins located on the cell surface may serve as drug targets for disease treatments, especially when macromolecules such as antibodies or enzymes are developed as drugs in promising immunotherapy.^{8,9}

Despite the importance of surface glycoproteins, the analysis of the surface glycoproteome is underrepresented compared to the whole proteome analysis because it is extraordinarily challenging to comprehensively analyze glycoproteins only located on the cell surface. Besides the low abundance of many glycoproteins and the heterogeneity of glycans, it is also difficult to specifically target surface glycoproteins. Antibody-based analysis has provided valuable information for cell surface glycoproteins.¹⁰ However, the low throughput and high cost have limited their wide applications. Technological advancements in mass spectrometry (MS) provide an opportunity to globally identify and quantify proteins and their modifications.^{11–21} In combination with MS, subcellular fractionation for the plasma membrane has been employed to analyze cell surface proteins,²² but it is not specific and the contaminants from highly abundant intracellular proteins are a serious issue. The combination of lectin enrichment and MS has provided another powerful method for glycoprotein analysis,^{23,24} but unfortunately it cannot be applied for surface glycoprotein analysis.

Several years ago, Wollscheid et al. developed an elegant method, termed cell surface capturing (CSC) technology, to systematically analyze surface N-glycoproteins.²⁵ This technique uses sodium periodate (NaIO₄) to generate a chemical reporter on the extracellular glycan moieties which can be used for surface glycoprotein enrichment prior to MS analysis. Recently, bioorthogonal chemistry offers another excellent opportunity to analyze cell surface glycoproteins,^{26–28} in which a sugar analogue with a chemical reporter is incorporated into glycoproteins through the biosynthetic machinery of a cell.²⁹ Previously, we combined metabolic labeling, copper-free click chemistry, enzymatic reaction, and MS-based proteomics to globally and site-specifically analyze cell surface N-glycoproteins.^{30–34} Strain-promoted alkyne–azide cycloaddition (SPAAC),³⁵ which avoids cytotoxic copper ions, is very mild and suitable for tagging cell surface glycoproteins. However, this method relies on metabolic labeling, and thus its applications to tissue or clinical samples are restricted. Enzymatic-based methods have great potential to tag surface glycoproteins because normally the reaction conditions are mild and the reactions are highly efficient.³⁶

In this work, global and site-specific analysis of N-glycoproteins on the cell surface was achieved through an effective method integrating enzymatic and chemical reactions with MS-based proteomics. Galactose oxidase (GAO), which specifically converts the hydroxyl group at C6 of Gal/GalNAc to an aldehyde group, was employed to tag cell surface glycoproteins. In order to make the oxidation more efficient, horseradish peroxidase (HRP) was added, which shifts the reaction toward completion by consuming one of the products, i.e., H₂O₂, and activates the deactivated form of GAO. The enzymatic reaction conditions are mild, which can minimize the stress posed on cells, and thus maintain the intact cellular state. After cell lysis, surface glycoproteins containing the aldehyde group were enriched with the hydrazide beads, followed by on-bead digestion and the removal of non-

glycopeptides. Glycopeptides were released through hydrazine–oxime exchange, and the enriched surface glycopeptides were treated with peptide N-glycosidase F (PNGase F) in heavy-oxygen water. The release of glycopeptides from the hydrazide beads eliminates the possible steric hindrance during the PNGase F treatment. Finally, a unique tag was generated on the N-glycosylation sites for MS analysis. When neuraminidase was utilized to remove the terminal sialic acids on glycans, more surface glycoproteins were identified. The experimental results demonstrated that the current method is highly effective to cover cell surface glycoproteins and can be applied for comprehensive and site-specific analysis of surface glycoproteins.

EXPERIMENTAL SECTION

Cell Culture.

MCF7 cells (from American type culture collection (ATCC)) were grown in Dulbecco's modified eagle's medium (DMEM) (Sigma-Aldrich) containing high glucose and 10% fetal bovine serum (FBS) (Thermo) inside a humidified incubator with 5.0% CO₂ at 37 °C. Jurkat cells (ATCC) were grown in Roswell Park Memorial Institute (RPMI) 1640 media (Sigma) with 10% FBS. For the quantification experiment, cells were grown in DMEM with 10% dialyzed FBS (Corning) supplemented with 0.146 g/L ¹³C₆¹⁵N₂ L-lysine (Lys-8) and 0.84 g/L ¹³C₆ L-arginine (Arg-6) (Cambridge Isotopes Inc.) or lysine (Lys-0) and arginine (Arg-0) with the same concentrations. After cells were cultured for at least six generations, heavy cells were treated with BFA (5 μg/mL) for 18 h, and the light cells were treated with DMSO as a control.

Oxidation of Glycans with GAO and Removal of Sialic Acid with Neuraminidase.

When the confluency of cells reached about 80%, cells were washed twice with phosphate-buffered saline (PBS, Sigma) and harvested by scraping. The cells were pelleted by centrifugation at 300 *g* for 3 min and washed once with cold PBS. Then cells were oxidized with a PBS solution containing galactose oxidase (50 U/mL, Innovative Research), HRP (40 U/mL, Sigma), and 5% FBS for 1 h at 37 °C. For the removal of sialic acids, cells were treated with neuraminidase (250 mU/mL, Sigma) in PBS (pH = 6.7) at 37 °C for 30 min before harvest, followed by the oxidation of surface glycoproteins with GAO.

Enrichment of Surface Glycoproteins and Treatment with PNGase F.

After cell lysis, the cell lysate was centrifuged, and the resulting supernatant was transferred to a new tube. Hydrazide beads (100 μL, Thermo Scientific) were washed three times with deionized water and transferred to the tube with the supernatant from cell lysis. After aniline (10 mM) was added, the mixture was incubated at 4 °C through end-over-end rotation for 24 h. On the next day, the beads were washed three times with a buffer containing 8 M urea, 0.4 M ammonium carbonate (NH₄HCO₃), and 0.1% sodium dodecyl sulfate (SDS). In this case, only tagged surface glycoproteins were enriched with the hydrazide beads.

After on-bead digestion with trypsin, glycopeptides were eluted twice by incubating with the elution buffer (0.2 M methoxylamine hydrochloride, 1.5 M NaCl, and 0.1 M aniline in 0.1 M sodium acetate solution, pH = 4.5) at 37 °C for 30 and 60 min, respectively. Purified and

completely dried peptides were then dissolved in 40 μ L of buffer containing 50 mM NH_4HCO_3 (pH = 9) in heavy-oxygen water (H_2^{18}O) and deglycosylated with peptide-N-glycosidase F (PNGase F, Sigma-Aldrich) for 3 h. The detailed information about protein alkylation and digestion and the treatment with PNGase F is given in the Supporting Information.

LC-MS/MS Analysis, Database Search, and Data Filtering.

The glycopeptide samples were dissolved in loading buffer consisting of 5% ACN and 4% FA and loaded onto a microcapillary column packed with C18 beads. Peptides were separated by HPLC and detected in a hybrid dual-cell quadrupole linear ion trap-Orbitrap MS. LC-MS analysis is described in more details in the Supporting Information.

SEQUEST (version 28) was used to search glycopeptides after converting the raw files to an mzXML format.³⁷ Spectra were matched against a database consisting of sequences from all human proteins (*Homo sapiens*) downloaded from the UniProt. The following parameters were employed for the database search: 20 ppm precursor mass tolerance; 1.0 Da product ion mass tolerance; fully digested with trypsin; up to three missed cleavages; variable modifications: oxidation of methionine (+15.9949 Da), ^{18}O tag of Asn (glycosylation site) (+2.9883 Da), heavy lysine (+8.0142 Da), and heavy arginine (+6.0201 Da); fixed modifications: carbamidomethylation of cysteine (+57.0214 Da).

The target-decoy method, in which each sequence from a protein was listed in both forward and reverse orders, was used to evaluate the false discovery rates (FDR) of glycopeptide and glycoprotein identifications.³⁸ The quality of glycopeptide and glycoprotein identification was evaluated and controlled by linear discriminant analysis (LDA) employing multiple parameters such as XCorr, charge state, and precursor mass error.³⁹ Peptides with fewer than seven amino acids were deleted, and peptide spectral matches were filtered to <1% FDR. The obtained data set was restricted to only glycopeptides when determining FDRs. In addition, a protein-level filter was further performed in each data set to reduce the protein-level FDRs to be <1% for glycoproteins.

Glycosylation Site Localization and Quantification.

A probabilistic algorithm was used to calculate the possibility of glycosylation localization based on fragmentation ions.⁴⁰ The ModScore indicates the possibility of the glycosylation site, and sites with a ModScore > 19 ($P < 0.01$) were considered to be confidently localized. For peptide quantification, signal-to-noise (S/N) ratios for both heavy and light peptides were required to be larger than 3. If the S/N ratio of heavy peptide was lower than 3, S/N of the light version of the peptide was required to be larger than 10 and vice versa. If the same peptides were quantified several times, the median ratio was calculated to represent the abundance change. For the glycosylation site quantification, the quantified glycopeptide must contain only one glycosylation site, and the ModScore has to be larger than 19 to make sure the glycosylation site was well-localized.

RESULTS

Specifically Tagging Cell Surface Glycoproteins with an Enzymatic Reaction.

In order to specifically analyze glycoproteins on the cell surface, they need to be selectively separated and enriched prior to MS analysis. Otherwise, glycoproteins on the cell surface cannot be distinguished from those inside of cells by MS. Enzymatic-based methods have great potential to tag surface glycoproteins,⁴¹ and previously galactose oxidase was employed to oxidize glycans on the cell surface.⁴² This enzymatic method is highly promising because while cells are alive, the enzyme as a macromolecule cannot penetrate the cell plasma membrane and thus only extracellular glycans are oxidized. In addition, the enzymatic reaction conditions are mild, which do not affect cell viability and growth during surface glycoprotein tagging. Therefore, this method is well-suited for surface glycoprotein analysis. To demonstrate the selective tagging of surface glycoproteins, we performed the cell imaging experiments, and the detailed information and the results are shown in Supporting Information.

GAO oxidizes the hydroxyl group at C6 position of the Gal/GalNAc residues and converts it to an aldehyde group. The aldehyde groups allow us to selectively separate and enrich surface glycoproteins. Based on the reports in the literature and our previous work,^{36,43} the addition of HRP is able to facilitate the oxidation reaction because it can consume one of the products, i.e., H₂O₂, which pushes the oxidation reaction toward completion. During the oxidation reaction, GAO is deactivated, which can be activated by HRP. In our parallel experiments, with the addition of HRP, the number of identified glycoproteins was increased by about 25% (Figure S1).

After oxidation, cells were lysed and proteins were extracted. The hydrazide beads were used to enrich surface glycoproteins containing the aldehyde group. On-bead digestion was performed with trypsin. Surface glycopeptides remaining on the beads were eluted with methoxylamine (CH₃ONH₂) after stringent wash. Finally, eluted glycopeptides were deglycosylated with PNGase F in heavy-oxygen water (H₂¹⁸O), which converts Asn to Asp, leaving a unique tag containing the heavy oxygen at the glycosylation sites for MS analysis (Figure 1).⁴⁴ Here surface glycopeptides were released from the beads before the cleavage of N-glycans with PNGase F. Without the release of N-glycopeptides from the beads, the steric hindrance caused by the beads may prevent the effective removal of N-glycans with PNGase F.

N-Glycoproteins and Glycosylation Sites Identified on the Surface of MCF7 Cells.

As described above, surface glycopeptides were enriched and then analyzed by LC-MS/MS. An example of glycopeptide identification is shown in Figure 2A.

CEAAN#GTEACVCGGAFVGPR (# represents the glycosylation site) was confidently identified with XCorr of 5.6 and mass accuracy of -0.47 ppm. The glycosylation site was well-localized at N41 with a ModScore of 1000. This peptide is from neurogenic locus notch homologue protein 1 (NOTCH1), which is a cell surface receptor that regulates cell-fate decisions including cell differentiation, proliferation, and apoptosis. In this work, we also identified two other glycosylation sites, N1241 and N1587, in NOTCH1.

From the biologically triplicate experiments, we identified 878, 1013, and 969 glycosylation sites from 381, 409, and 388 N-glycoproteins, respectively (Figure 2B and 2C). The identified glycosylation sites are listed in Table S1. Three hundred and seventeen glycoproteins were identified in all three experiments, which demonstrated high reproducibility of the current method. Overall, 1199 glycosylation sites from 469 N-glycoproteins were identified (Table S1). Among the N-glycoproteins identified in those experiments, nearly half of them are single-pass type-I membrane proteins (230 proteins), which extend their N-termini to the extracellular space. Only 24 proteins belong to single-pass type-II membrane proteins whose C-termini are located in the extracellular region. About 30% of identified N-glycoproteins are multipass membrane proteins, and only 7 proteins are peripheral membrane proteins (Figure 3A). For 69 proteins without cell surface localization information from UniProt (<http://www.uniprot.org>), we performed computation analysis using Phobius to predict the transmembrane domain and signal peptides of those proteins.⁴⁵ The sequences of those undefined proteins were downloaded from Uniprot and then uploaded onto the Phobius Web site (<http://phobius.sbc.su.se/>). All default parameters set on the Web site were used for the prediction of transmembrane and signal peptides. More than half of them (37) contain a signal peptide, and thus they may be secreted through the classical secretory pathway and then resided on the cell surface. Three proteins contain only the transmembrane domain, while 29 proteins have both the transmembrane domain and signal peptides. Overall, the vast majority of the identified glycoproteins were membrane proteins.

Gene Ontology analysis using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (v6.8) was performed to determine biological processes and molecular functions for the identified glycoproteins.⁴⁶ Surface glycoproteins were uploaded onto the DAVID Web site (<https://david.ncifcrf.gov/home.jsp>). *Homo sapiens* was chosen as the background, and the *P* value for each enriched group of proteins was calculated against the human proteome. Proteins related to cell adhesion were the most highly enriched with an extremely low *P* value of 1.2×10^{-52} , and proteins belonging to the categories of cell surface receptor signaling (3.2×10^{-42}) and cell mobility (3.1×10^{-29}) were also highly enriched (Figure 3B). Protein clustering based on molecular function revealed that the most enriched categories include molecular transducer activity, signaling receptor activity, growth factor binding, and integrin binding, which are well-known functions of cell surface proteins (Figure S2). These results further demonstrated the effectiveness and specificity of the current method.

Improvement of N-Glycoprotein Identification through Neuraminidase Treatment.

Sialic acids are a family of nine-carbon acidic sugars found at the termini of many glycans.⁴⁷ Protein sialylation is commonly formed with the transfer of CMP-Neu5Ac to Gal/GalNAc through α -2,3- or α -2,6-linkage catalyzed by sialyltransferases,⁴⁸ which may affect the oxidation of Gal/GalNAc with GAO. Neuraminidases (sialidases) can remove sialic acid from glycans, and here neuraminidase from *Clostridium perfringens* was used to cleave sialic acids, which was expected to improve the coverage of surface glycoprotein identifications.

For MCF7 cells, 393 and 406 N-glycoproteins were identified without and with the neuraminidase treatment, respectively. The number of identified N-glycoproteins did not increase dramatically with the neuraminidase treatment, which could be due to the following reasons: First, MCF7 cells were derived from an early stage breast cancer patient and are minimally invasive, and compared to highly invasive cells, sialylation of surface proteins is less frequent. Second, for a protein, even some glycosylation sites are sialylated, but nonsialylated glycosylation sites still allow us to identify it. Similarly, for a glycan, the branches without sialylation still enable us to identify the glycoprotein.

Jurkat cells were also used to study the influence of the neuraminidase treatment on N-glycoprotein identifications. In the parallel experiments without or with the neuraminidase treatment, 366 and 588 unique glycosylation sites from 178 and 270 N-glycoproteins were identified, respectively (Figure 4). The results demonstrated that 61% more glycosylation sites and 52% more glycoproteins were identified with the neuraminidase treatment. The N-glycoproteins identified in the neuraminidase treatment experiment covered the majority of glycoproteins identified in the experiment without the neuraminidase treatment (96.6%). Depending on surface protein sialylation, the neuraminidase treatment may notably increase the coverage of the cell surface glycoproteome.

Sialic acid carries a negative charge under physiological conditions, which may increase electrostatic repulsion among cells and promote cell migration and metastasis. Based on the protein clustering using DAVID, 26 of 98 proteins only identified in the neuraminidase treatment experiment from Jurkat cells were involved in cell migration. For example, fibroblast growth factor receptor 1 (FGFR1) is a tyrosine-protein kinase that functions as a cell-surface receptor and plays an essential role in the regulation of cell proliferation, differentiation, and migration. Two glycosylation sites, N264 and N317, were identified on FGFR1 in this experiment. The two sites are located in the Ig-like domain, which is involved in cell adhesion and cell surface recognition.⁴⁹ It was reported that N-glycans of FGFR1 with bi- and triantennary corefucosylated complex structures were heavily sialylated.⁵⁰ In combination with neuraminidase, the current method can be employed to comprehensively identify cell surface glycoproteins.

Quantification of Surface N-Glycoproteins in MCF7 Cells with BFA Treatment.

Proteins residing on the cell surface participate in many important cellular events, such as the regulation of bioactive molecule transportation and the mediation of signaling pathways in response to external stimuli. Nearly all proteins on the plasma membrane are glycosylated. The classical secretory pathway is the major route for the trafficking of proteins on the cell surface and in the secretome.⁵¹ In addition, multiple nonclassical secretion pathways have been reported.⁵² For protein N-glycosylation, it is highly conserved from yeast to human cells. After a peptide is synthesized, the initial N-glycan consisting of 14 sugar residues (Glc₃Man₉GlcNAc₂) is “en bloc” transferred to the nascent polypeptide in the lumen of the ER membrane. The initial trimming of the precursor glycan occurs in the ER. After leaving from the ER, glycoproteins enter the Golgi apparatus, where glycans are further modified by various enzymes. Followed by transporting via secretory vesicles, some glycoproteins are inserted into the plasma membrane.

Brefeldin A (BFA) is a potent inhibitor to block protein secretion through the classical pathway because it has dramatic influence on the structure and function of the Golgi apparatus and impairs intracellular protein transport and secretion.⁵³ The current method is highly effective to globally identify surface N-glycoproteins, and together with quantitative proteomics, it can be used for quantification of surface glycoproteins. Quantification of surface glycoproteins may further demonstrate the effectiveness of the method.

In combination with stable isotope labeling by amino acids in cell culture (SILAC), we quantified the abundant changes of cell surface N-glycoproteins with BFA treatment. The quantification results will aid us in a better understanding of surface glycoprotein trafficking. The experimental procedure is displayed in Figure 5A. MCF7 cells labeled with heavy lysine and arginine were treated with BFA while cells in light media were treated with DMSO as a control. Surface glycoprotein oxidation and enrichment were performed as described in the Experimental Section. In this experiment, 909 unique N-glycopeptides from 334 N-glycoproteins were quantified, and 65% of them were down-regulated by over twofold (Figure 5B) in cells with BFA treatment. For example, EPHB4 is a tyrosine kinase receptor and one of the known proteins secreted through the classical ER-Golgi pathway.⁵⁴ It was quantified to be down-regulated by 4.8-fold under BFA treatment.

Proteins with a signal peptide are normally transported to the extracellular region through the classical secretory pathway.⁵⁵ Based on the prediction from SignalP 4.1,⁵⁶ 242 quantified proteins have a signal peptide, and 170 (about 70%) of them were down-regulated by over twofold. The down-regulated surface glycoproteins are normally transported through the classical pathway. For type I transmembrane proteins, the N-terminal domains target the ER lumen after the signal sequence was recognized by signal-recognition particle (SRP) during biosynthesis.⁵⁷ According to the information from UniProt, 168 quantified proteins were type I membrane proteins, and 121 of them were down-regulated by more than twofold. The abundances of a small number of glycoproteins did not change much, which may be due to the following reasons. First, they were very stable during the treatment; second, their transportations were through nonclassical secretory pathways, which were not affected by the inhibition of the classical secretion pathway, i.e., BFA treatment. These results strongly suggested that most of surface proteins were transported through the classical secretory pathway. The experiment also demonstrates that in combination with quantitative proteomics, the current method can be used to quantify surface glycoprotein abundance changes.

DISCUSSION

Almost all proteins on the cell surface are glycosylated, and these surface glycoproteins regulate nearly every extracellular activity, from cell–cell interactions to the initiation of cell signaling. Global and site-specific analysis of surface glycoproteins can help us have a better understanding of the functions of many important proteins, such as receptors, transporters, and enzymes, on the cell surface, which will lead to discover surface glycoproteins as biomarkers and drug targets. The site-specific analysis of surface glycoproteins using MS provides not only valuable information about protein glycosylation site but also solid experimental evidence for glycopeptide identifications.

Cell surface receptors are extremely important proteins that sense and respond to extracellular nutrition and signal changes. Extracellular ligands (cytokines, growth factors, and hormones, etc.) bind to cell surface receptors, activating signaling pathways, including pathways required for cell survival, differentiation, and proliferation.⁵⁸ A total of 255 glycosylation sites from 93 cell surface receptors were quantified in this experiment (Table S4). For example, IGF1R, the insulin-like growth factor 1 receptor, is a well-known receptor located on the plasma membrane and is transported through the classical secretory pathway. Three quantified glycosylation sites on IGF1R (N51, N747, and N913) were down-regulated by over twofold. Of all the quantified receptors, over 78% were down-regulated. For the proteins that remained unchanged (down-regulated by less than twofold), we reasoned that these proteins might be long-lived or secreted through nonclassical pathways. For example, ERBB2 is a very important protein tyrosine kinase that is part of several cell surface receptor complexes. Its half-life was longer than 200 h according to a previous report,³³ and its abundance showed nearly no change under BFA treatment.

As described above, the CSC method was highly innovative despite that the oxidation conditions were a bit harsh to live cells. Using the CSC method, surface glycoproteins in many different types of cells were analyzed, and around 288 glycoproteins were identified for each cell line on average.⁵⁹ Previously GAO was used to tag surface glycoproteins, followed by the enrichment and MS analysis. In that work, 112 N-glycosylation sites on 68 glycoproteins were identified on the cell surface of K20 B-cells.⁴² Here, among the biologically triplicate experiments, on average 953 N-glycosylation sites on 393 glycoproteins per experiment were identified on the surface of MCF7 cells. Of note, besides the method differences, different MS and cell types may also contribute to different results.

For the method based on the combination of metabolic labeling and click-chemistry, 274 surface N-sialoglycoproteins were identified in HEK 293T cells in our previous work.³¹ We compared the cell surface protein labeling with different sugar analogues, i.e., N-azidoacetylgalactosamine (GalNAz), N-azidoacetylglucosamine (GlcNAz), and N-azidoacetylmannosamine (ManNAz). The results demonstrated that GalNAz was the most effective for surface N-glycoprotein analysis, and on average 224 N-glycoproteins per experiment were identified in HepG2 cells. For that work, we used the same MS and parameters as the current one.

GAO was previously used to tag surface glycoproteins, but the current method has several differences. First, HRP was used to facilitate the oxidation of Gal/GalNAc. HRP can push the oxidation reaction toward completion by reacting with one of the products (H_2O_2), and at the same time, it activates GAO that may be deactivated during the reaction. Second, instead of using PNGase F to directly release glycopeptides from the beads, we used methoxylamine to elute the enriched glycopeptides, followed by the enzymatic treatment. PNGase F is an enzyme, and thus due to the steric hindrance it is less effective to cleave N-glycans and release glycopeptides from the beads directly. Third, we performed the PNGase F treatment in heavy-oxygen water, which can allow us to distinguish the bona fide N-glycosylation sites from those due to the spontaneous asparagine deamination. Asparagine deamination may happen in cells and during sample preparation. In addition, another potential advantage is that the tagged surface glycopeptides were released from the beads

using methoxylamine, and thus we may perform intact glycopeptides analysis from the cell surface. Furthermore, sialic acids at the termini of glycans were removed to make surface glycoproteins more effectively tagged. Sialic acid at the glycan terminus may prevent its oxidation with GAO and, using neuraminidase, sialic acid can be removed before the oxidation.

Integrating enzymatic and chemical reactions and MS-based proteomics, global and site-specific analysis of surface glycoproteins can be achieved. The results demonstrated that the current method is highly effective. In combination with quantitative proteomics, surface glycoprotein abundance changes can be quantified. Because of the importance of glycoproteins, this method will have extensive applications in the biological and biomedical research fields.

CONCLUSIONS

Glycoproteins on the cell surface regulate many extracellular events, and they are often reflective of the developmental and diseased status of cells. Despite the importance of surface glycoproteins, it is extraordinarily challenging to globally analyze glycoproteins only located on the cell surface because of the low abundance of many glycoproteins, the heterogeneity of glycans, and the difficulty to selectively separate surface glycoproteins. Effective methods for the global analysis of surface glycoproteins are urgently needed. The current method integrating enzymatic and chemical reactions with MS-based proteomics enables comprehensive and site-specific analysis of cell surface glycoproteins. The reaction conditions for tagging surface glycoproteins are mild, which is excellently suitable for live cell analysis. With the addition of HRP, the oxidation reaction is more efficient. The resulting aldehyde groups on surface glycoproteins serve as a chemical handle for their selective separation and enrichment. The enriched surface glycoproteins through hydrazide chemistry were digested on the beads. The remaining glycopeptides were released with methoxylamine before the PNGase F treatment in heavy-oxygen water, which eliminated the potential steric hindrance. Among the biologically triplicate experiments, on average 953 N-glycosylation sites on 393 surface glycoproteins per experiment were identified in MCF7 cells. In combination with quantitative proteomics, we performed systematic and quantitative analysis of the surface glycoproteome in MCF7 cells treated with BFA, and most glycoproteins were down-regulated under BFA treatment. Many down-regulated glycoproteins have a signal peptide, which indicates that these proteins are transported through the classical secretory pathway. The current method is highly effective for large-scale analysis of surface glycoproteins and can be extensively applied for global analysis of surface glycoproteins in the biological and biomedical research fields.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM118803.

REFERENCES

- (1). Boscher C; Dennis JW; Nabi IR *Curr. Opin. Cell Biol* 2011, 23 (4), 383–392. [PubMed: 21616652]
- (2). Rabinovich GA; Toscano MA *Nat. Rev. Immunol* 2009, 9, 338. [PubMed: 19365409]
- (3). Stowell SR; Ju T; Cummings RD *Annu. Rev. Pathol.: Mech. Dis* 2015, 10, 473–510.
- (4). Holst S; Belo AI; Giovannetti E; van Die I; Wührer M *Sci. Rep* 2017, 7, 15. [PubMed: 28148958]
- (5). Macauley MS; Crocker PR; Paulson JC *Nat. Rev. Immunol* 2014, 14 (10), 653–666. [PubMed: 25234143]
- (6). Liu FT; Rabinovich GA *Nat. Rev. Cancer* 2005, 5 (1), 29–41. [PubMed: 15630413]
- (7). Leth-Larsen R; Lund RR; Ditzel HJ *Mol. Cell. Proteomics* 2010, 9 (7), 1369–1382. [PubMed: 20382631]
- (8). Scott AM; Wolchok JD; Old LJ *Nat. Rev. Cancer* 2012, 12 (4), 278–287. [PubMed: 22437872]
- (9). Sliwkowski MX; Mellman I *Science* 2013, 341 (6151), 1192–1198. [PubMed: 24031011]
- (10). Belov L; de la Vega O; dos Remedios CG; Mulligan SP; Christopherson RI *Cancer Res* 2001, 61 (11), 4483. [PubMed: 11389079]
- (11). Zhang H; Li X.-j.; Martin DB; Aebersold R *Nat. Biotechnol* 2003, 21 (6), 660–666. [PubMed: 12754519]
- (12). Wu R; Haas W; Dephoure N; Huttlin EL; Zhai B; Sowa ME; Gygi SP *Nat. Methods* 2011, 8 (8), 677–683. [PubMed: 21725298]
- (13). Rodrigues SP; Alvarez S; Werth EG; Slade WO; Gau B; Cahoon EB; Hicks LM *Anal. Methods* 2015, 7 (17), 7336–7344.
- (14). McShane E; Sin C; Zaubert H; Wells JN; Donnelly N; Wang X; Hou J; Chen W; Storchova Z; Marsh JA; Valleriani A; Selbach M *Cell* 2016, 167 (3), 803–815. [PubMed: 27720452]
- (15). Xiao HP; Chen WX; Smeekens JM; Wu RH *Nat. Commun* 2018, 9, 1692. [PubMed: 29703890]
- (16). Zhu Z; Desaire H *Annu. Rev. Anal. Chem* 2015, 8 (1), 463–483.
- (17). Schunter AJ; Yue XS; Hummon AB *Anal. Bioanal. Chem* 2017, 409 (7), 1749–1763. [PubMed: 27987026]
- (18). Zacharias LG; Hartmann AK; Song EH; Zhao JF; Zhu R; Mirzaei P; Mechref Y J. *Proteome Res* 2016, 15 (10), 3624–3634. [PubMed: 27533485]
- (19). Shi TJ; Song EW; Nie S; Rodland KD; Liu T; Qian WJ; Smith RD *Proteomics* 2016, 16 (15–16), 2160–2182. [PubMed: 27302376]
- (20). King CD; Singh D; Holden K; Govan AB; Keith SA; Ghazi A; Robinson RAS J. *Proteomics* 2018, 181, 92–103. [PubMed: 29656019]
- (21). Cai R; Huang M; Wang YS *Anal. Chem* 2018, 90 (24), 14339–14346. [PubMed: 30433760]
- (22). He JT; Liu YS; He SZ; Wang QS; Pu H; Tong YP; Ji JG *Proteomics: Clin. Appl* 2007, 1 (2), 231–241. [PubMed: 21136673]
- (23). Chen WX; Smeekens JM; Wu RH *Mol. Cell. Proteomics* 2014, 13 (6), 1563–1572. [PubMed: 24692641]
- (24). Song EW; Zhu R; Hammoud ZT; Mechref Y J. *Proteome Res* 2014, 13 (11), 4808–4820. [PubMed: 25134008]
- (25). Wollscheid B; Bausch-Fluck D; Henderson C; O'Brien R; Bibel M; Schiess R; Aebersold R; Watts JD *Nat. Biotechnol* 2009, 27 (4), 378–386. [PubMed: 19349973]
- (26). Mahal LK; Yarema KJ; Bertozzi CR *Science* 1997, 276 (5315), 1125. [PubMed: 9173543]
- (27). Hubbard SC; Boyce M; McVaugh CT; Peehl DM; Bertozzi CR *Bioorg. Med. Chem. Lett* 2011, 21 (17), 4945–4950. [PubMed: 21798741]
- (28). Grammel M; Hang HC *Nat. Chem. Biol* 2013, 9 (8), 475–484. [PubMed: 23868317]
- (29). Hsu TL; Hanson SR; Kishikawa K; Wang SK; Sawa M; Wong CH *Proc. Natl. Acad. Sci. U. S. A* 2007, 104 (8), 2614–2619. [PubMed: 17296930]
- (30). Xiao H; Tang GX; Wu R *Anal. Chem* 2016, 88 (6), 3324–3332. [PubMed: 26894747]
- (31). Chen W; Smeekens JM; Wu R *Chem. Sci* 2015, 6 (8), 4681–4689. [PubMed: 29142707]

- (32). Smeekens JM; Chen W; Wu R J. Am. Soc. Mass Spectrom 2015, 26 (4), 604–614. [PubMed: 25425172]
- (33). Zheng J; Xiao H; Wu R Angew. Chem., Int. Ed 2017, 56 (25), 7107–7111.
- (34). Xiao H; Suttapitugsakul S; Sun F; Wu R Acc. Chem. Res 2018, 51 (8), 1796–1806. [PubMed: 30011186]
- (35). Baskin JM; Prescher JA; Laughlin ST; Agard NJ; Chang PV; Miller IA; Lo A; Codelli JA; Bertozzi CR Proc. Natl. Acad. Sci. U. S. A 2007, 104 (43), 16793–16797. [PubMed: 17942682]
- (36). Parikka K; Master E; Tenkanen M J. Mol. Catal. B: Enzym 2015, 120, 47–59.
- (37). Eng JK; McCormack AL; Yates JR J. Am. Soc. Mass Spectrom 1994, 5 (11), 976–989. [PubMed: 24226387]
- (38). Elias JE; Gygi SP Nat. Methods 2007, 4 (3), 207–214. [PubMed: 17327847]
- (39). Huttlin EL; Jedrychowski MP; Elias JE; Goswami T; Rad R; Beausoleil SA; Villen J; Haas W; Sowa ME; Gygi SP Cell 2010, 143 (7), 1174–1189. [PubMed: 21183079]
- (40). Beausoleil SA; Villen J; Gerber SA; Rush J; Gygi SP Nat. Biotechnol 2006, 24 (10), 1285–1292. [PubMed: 16964243]
- (41). Sun TT; Yu SH; Zhao P; Meng L; Moremen KW; Wells L; Steet R; Boons GJ J. Am. Chem. Soc 2016, 138 (36), 11575–11582. [PubMed: 27541995]
- (42). Ramya TN; Weerapana E; Cravatt BF; Paulson JC Glycobiology 2013, 23 (2), 211–221. [PubMed: 23070960]
- (43). Zheng JN; Xiao HP; Wu RH Angew. Chem., Int. Ed 2017, 56 (25), 7107–7111.
- (44). Küster B; Mann M Anal. Chem 1999, 71 (7), 1431–1440. [PubMed: 10204042]
- (45). Kall L; Krogh A; Sonnhammer EL Nucleic Acids Res 2007, 35 (Web Server issue), W429–W432. [PubMed: 17483518]
- (46). Huang DW; Sherman BT; Lempicki RA Nat. Protoc 2009, 4, 44. [PubMed: 19131956]
- (47). Thaysen-Andersen M; Larsen MR; Packer NH; Palmisano G RSC Adv 2013, 3 (45), 22683.
- (48). Tanner ME Bioorg. Chem 2005, 33 (3), 216–228. [PubMed: 15888312]
- (49). Martin M; Romero X; de la Fuente MA; Tovar V; Zapater N; Esplugues E; Pizcueta P; Bosch J; Engel P J. Immunol 2001, 167 (7), 3668–3676. [PubMed: 11564780]
- (50). Duchesne L; Tissot B; Rudd TR; Dell A; Fernig DG J. Biol. Chem 2006, 281 (37), 27178–27189. [PubMed: 16829530]
- (51). Brandizzi F; Barlowe C Nat. Rev. Mol. Cell Biol 2013, 14 (6), 382–392. [PubMed: 23698585]
- (52). Nickel W Traffic 2005, 6 (8), 607–614. [PubMed: 15998317]
- (53). Nebenfuhr A; Ritzenthaler C; Robinson DG Plant Physiol 2002, 130 (3), 1102–1108. [PubMed: 12427977]
- (54). Pasquale EB Nat. Rev. Cancer 2010, 10, 165. [PubMed: 20179713]
- (55). von Heijne G J. Membr. Biol 1990, 115 (3), 195–201. [PubMed: 2197415]
- (56). Petersen TN; Brunak S; von Heijne G; Nielsen H Nat. Methods 2011, 8 (10), 785–786. [PubMed: 21959131]
- (57). Ott CM; Lingappa VR J. Cell Sci 2002, 115 (10), 2003. [PubMed: 11973342]
- (58). Deller MC; Yvonne Jones E Curr. Opin. Struct. Biol 2000, 10 (2), 213–219. [PubMed: 10753817]
- (59). Bausch-Fluck D; Hofmann A; Bock T; Frei AP; Cerciello F; Jacobs A; Moest H; Omasits U; Gundry RL; Yoon C; Schiess R; Schmidt A; Mirkowska P; Hartlova A; Van Eyk JE; Bourquin JP; Aebersold R; Boheler KR; Zandstra P; Wollscheid B PLoS One 2015, 10 (3), No. e0121314. [PubMed: 25894527]

Figure 1. Principle of site-specific and comprehensive analysis of N-glycoproteins on the cell surface by integrating enzymatic and chemical reactions with MS-based proteomics.

Figure 2.

(A) Tandem mass spectrum of CEAAN#GTEACVCGGAFVGPR (# represents the glycosylation site). (B) Comparison of N-glycosylation sites and (C) N-glycoproteins identified on the cell surface of MCF7 cells in biologically triplicate experiments.

Figure 3. Classification of cell surface N-glycoproteins identified in MCF7 cells (A) and clustering of cell surface N-glycoproteins identified in MCF7 cells based on biological process (B).

Figure 4. Identification of N-glycosylation sites (A) and N-glycoproteins (B) on the cell surface of Jurkat cells with and without the neuraminidase treatment.

Figure 5. Procedure of quantification experiment with BFA treatment (A) and quantification results of surface N-glycoproteins (B) and N-glycopeptides (C) in MCF7 cells with BFA treatment.