High-Throughput Determination of the Site-Specific N-Sialoglycan Occupancy Rates by Differential Oxidation of Glycoproteins Followed with Quantitative Glycoproteomics Analysis

Zhang Zhang,†‡ Zhen Sun,†‡ Jun Zhu,†‡ Jing Liu,†‡ Guang Huang,†‡ Mingliang Ye,*† and Hanfa Zou*†

†CAS Key Laboratory of Separation Sciences for Analytical Chemistry, National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China
‡University of Chinese Academy of Sciences, Beijing 100049, China

Supporting Information

ABSTRACT: Sialylated glycoproteins, which play important roles in tumor progression, have been extensively analyzed for the discovery of potential biomarkers for cancer diagnosis and prognosis. The site-specific N-sialoglycan occupancy rates of glycoproteins reflect the activities of glycosyltransferases and glycosidases in vivo and could be novel disease biomarkers. However, a high-throughput method to determine the N-sialoglycan occupancy rates is not available. On the basis of the fact that dihydroxy of sialic acid of glycan chains in glycoproteins can be specifically oxidized to aldehyde in mild periodate concentration while all types of glycan chains can be oxidized in high periodate concentration, we developed a modified protein-level hydrazide chemistry method for the determination of the N-sialoglycan occupancy rates. This method was first applied to determine the N-sialoglycan occupancy rates of two glycosites on human transferrin. These two sites were found to be fully sialylated and the N-sialoglycan occupancy rates were found to under significant decrease after the neuraminidase treatment. This method was then applied to analyze N-sialoglycan occupancy rates in proteome samples. We determined 496 and 632 site-specific N-sialoglycan occupancy rates on 334 and 394 proteins from hepatocellular carcinoma (HCC) and normal human liver tissues, respectively. By comparing the N-sialoglycan occupancy rates between the above two samples, we determined 76 N-sialoglycosites with more than a 2-fold change. This method was demonstrated to be an effective and high-throughput method for the analysis of the N-sialoglycan occupancy rates.

Many current used disease biomarkers, like α-fetoprotein (AFP) for hepatocellular carcinoma (HCC), cancer antigen 125, and prostate-specific antigen, are glycoproteins. However, they are still lacking specificity and sensitivity for the early diagnosis of cancers. It is of urgent need to find more useful biomarkers. The alteration of glycoproteins during the genesis and progress of a disease may happen in one of the following three ways or the combinations: (i) the abundance of glycoproteins reflect the activities of glycosyltransferases and glycosidases in vivo and could be novel disease biomarkers. However, a high-throughput method to determine the N-sialoglycan occupancy rates is not available. On the basis of the fact that dihydroxy of sialic acid of glycan chains in glycoproteins can be specifically oxidized to aldehyde in mild periodate concentration while all types of glycan chains can be oxidized in high periodate concentration, we developed a modified protein-level hydrazide chemistry method for the determination of the N-sialoglycan occupancy rates. This method was first applied to determine the N-sialoglycan occupancy rates of two glycosites on human transferrin. These two sites were found to be fully sialylated and the N-sialoglycan occupancy rates were found to under significant decrease after the neuraminidase treatment. This method was then applied to analyze N-sialoglycan occupancy rates in proteome samples. We determined 496 and 632 site-specific N-sialoglycan occupancy rates on 334 and 394 proteins from hepatocellular carcinoma (HCC) and normal human liver tissues, respectively. By comparing the N-sialoglycan occupancy rates between the above two samples, we determined 76 N-sialoglycosites with more than a 2-fold change. This method was demonstrated to be an effective and high-throughput method for the analysis of the N-sialoglycan occupancy rates.

Received: July 2, 2014
Accepted: September 10, 2014
Published: September 10, 2014

© 2014 American Chemical Society
glycan occupancy rates may be also used to infer disease state. As cancer has profound effects on gene expression, including various types of glycosyltransferases and glycosidases, the occupancy rates of different types of glycan chains in glycosites may change during disease genesis and progression. However, a proteomic approach to analyze the N-glycan occupancy rates for a specific type of glycan is still lacking.

Sialylation of glycoproteins is one of the most common glycosylations, which regulates essential biological functions such as intermolecular interactions, the formation of cellular characteristics, cell metastasis, and so forth. Because aberrant sialylation of glycoproteins is relevant to many different cancers such as breast cancer, colon cancer, acute myeloid cancer, leukemia cancer, pancreatic cancer, and gynecologic cancers, researchers have paid more and more attention to this type of glycosylation. Currently, various glycoproteomics methods were applied for the identification and quantification of sialylated glycoproteins. Kubota et al. combined lectin affinity chromatography and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) to determine the glycan structure of glycoproteins and used Sambucus nigra agglutinin (SNA) to specifically isolate Sia2–6 glycopeptides of transferrin. The sialylation chemistry method that was introduced into glycoproteomics analysis by Zhang and coworkers has been used broadly for the identification and quantification of N-glycoproteomics in the past decade.

Moreover, the modified hydrazide chemistry, using mild periodate concentration to selectively oxidize the sialylated glycoproteins in the protein level, had been applied to the analysis of the sialoglycoproteins of cell surface and human cerebrospinal fluid. The mild periodate oxidation can also be performed at peptide level after the sialoglycoproteins are digested. For example, the sialoglycoproteins in breast cancer and paired noncancer tissues were quantified successfully by using this method. All these methods directly compare the modiﬁed sialoglycan occupancy rates of different types of glycan chains in glycosites.

The sialylglycan occupancy rate can be deﬁned as the molar ratio of sialylglycan chains to all glycan chains on the same glycosite of a protein. On the basis of the fact that dihydroxy of sialic acid of glycan chains in glycoproteins can be speciﬁcally oxidized to aldehyde by 1 mM NaIO₄ while all types of glycan chains can be oxidized by 10 mM NaIO₄, we developed a modiﬁed hydrazide chemistry method for the determination of the N-sialylglycan occupancy rates by comparing the abundances of the deglycosylated peptides generated by differential oxidation of the same proteome sample. This method was ﬁrst validated by determining the N-sialylglycan occupancy rates of two N-glycosites on human transferrin. This method was then applied to determine the N-sialylglycan occupancy rates of glycoproteins in HCC and normal human liver tissues. By comparing the N-sialylglycan occupancy rates between above two samples, we determined 76 N-sialylglycosites with more than a 2-fold change in N-sialylglycan occupancy rates. This method was demonstrated to be an effective and high-throughput method for the analysis of the N-sialylglycan occupancy rates.

■ EXPERIMENTAL SECTION

Materials and Reagents. Transferrin (from human blood plasma), sodium periodate, protease inhibitor cocktail, trypsin, and 2,5-dihydroxy benzoin acid (DHB) were purchased from Sigma (St. Louis, MO, U.S.A.). Hydrazide sepharose resin was purchased from Bio-Rad (Hercules, CA, U.S.A.). PNGase F and neuraminidase were purchased from New England Biolabs (Ipswich, MA, U.S.A.). BCA assay kit was purchased from Pierce (Rockford, IL, U.S.A.). Chemical reagents of IAA (iodacetamide), DTT (di-thiothreitol), and TEAB (triethylammonium bicarbonate buffer) were obtained from Sigma. Formic acid (FA) was obtained from Fluka (Buchs, Germany). ACN (acetonitrile, HPLC grade) was obtained from Merck (Darmstadt, Germany). Ammonium bicarbonate and urea were obtained from Bio Basic Inc. (Ontario, Canada). Zeba Spin desalting columns were purchased from Thermo Scientiﬁc (Rockford, IL, U.S.A.). Pure water used in all of the experiments was puriﬁed with a Milli-Q system (Millipore, Milford, MA, U.S.A.). Other chemicals used were either of analytical grade or of better grade.

Preparation of Asialo Human Transferrin. An amount of 1 mg of human transferrin was dissolved in 90 μL of denaturing buffer (8 M urea and 100 mM NH₄HCO₃, pH = 8.2) and incubated in boiling water for 10 min. Then, it was desalted by Thermo Scientiﬁc Zeba Spin desalting columns (0.5 mL); then 10 μL of 100 mM ammonium acetate (pH = 6.0) with 20 units of neuraminidase was added to the eluted solution for an incubation at 37 ºC for 24 h to remove sialic acids from the glycan chains.

Extraction of Proteins from HCC and Normal Human Liver Tissues. The utilization of human tissues was stood to guidelines of the Ethics Committee of Second Affiliated Hospital of Dalian Medical University (Dalian, China), and the HCC and normal human liver tissues were provided by the hospital. The normal human liver tissues were the non-cancerous liver tissues ≥2 cm outside the hepatic cancer nodules sliced by surgical operation after being approved by the patients. The normal liver tissues had been verified by histopathological examination which excluded the presence of invading or microscopic metastatic cancer cells. The HCC tissues were acquired from the HCC patients of advanced stage by surgical operation. About 1 cm × 1 cm of the normal liver tissues from five patients with hepatic cancer were pooled together for all the experiments to reduce the biological variation, and so were the HCC tissues.

First, the isolated human liver tissues were cut into pieces and washed several times with PBS buffer (pH = 7.4) to remove the existing blood. Second, the tissues were placed in an ice-cold homogenization buffer I consisting of 10 mM HEPES, 1.5 mM MgCl₂, 5 mM KCl, 0.1 mM EDTA, 2% protease inhibitor cocktail at pH 7.4, followed by homogenization using an IKA Ultra Turbax blender. After that, the tissues were homogenized in a Potter-Elvehjem homogenizer with a Teflon piston for a second time on ice, and then the homogenates were centrifuged at 1000g for 5 min to pellet the nuclei and debris. Third, the supernatant was collected, and 5 vol of buffer II (0.1 mM Tris–HCl, 4% SDS, 1% Triton, pH = 7.4) was added for sonication using an ultrasonic cell disrupter (3 s with 3 s intervals for 180 times at 400 W) in ice water bath and centrifuged at 20 000g for 15 min. The supernatant was collected, and the concentration of proteins was determined by the BCA method.
Differential Oxidation and Dimethyl Labeling. Each sample to be analyzed was divided into two aliquots. One part was used for the isolation of total glycopeptides and the other part for the selective capture of sialoglycopeptides. For the total glycopeptide enrichment, glycoproteins with all types of glycan chains were captured onto hydrazide resins by a solid-phase extraction of glycoproteins (SPEG) approach with slight modifications.26 Extracted proteins solution (1 mg) from either HCC or normal liver tissues were boiled in water bath for 10 min and then desalted by Thermo Scientific Zeba Spin desalting columns. After dilution by coupling buffer (100 mM NaAc and 150 mM NaCl, pH = 5.5) to 400 μL, the samples were oxidized by adding 100 μL of 50 mM NaIO4 (i.e., the final NaIO4 concentration was 10 mM) in dark at room temperature for 1 h. Then the oxidation reaction was quenched by adding 125 μL of 100 mM Na2S2O4 for 20 min at room temperature. Finally, the oxidized samples were added into the tube with 100 μL of prewashed hydrazide resins in coupling buffer. The coupling reaction was carried out at room temperature with gentle shaking overnight. Non-glycopeptides were then removed by washing the resins three times with an equal volume of 100 mM NH4HCO3, (pH = 8.2). Then the resins were suspended in 490 μL of denaturing buffer (8 M urea, 100 mM NH4HCO3), followed by sequential adding of 10 μL of 1 M DTT for 2 h at 37 °C and 3.7 mg of iodoacetamide for 1 h at room temperature in dark for reduction and alkylation of captured glycoproteins. After washing successively with 80% acetonitrile and 100 mM NH4HCO3, three times, trypsin was added with an enzyme-to-protein ratio of 1/25 in 500 μL of 100 mM NH4HCO3 for digestion at 37 °C overnight with gentle shaking. Non-glycopeptides were removed by washing the resins successively with 1.5 M NaCl, 80% acetonitrile, 100 mM NH4HCO3, and water. Finally, only the glycopeptides (AllGlycan-peptides) were still captured on hydrazide resins.

The procedure for the capture of sialylated glycopeptides (SialoGlycan-peptides) from another aliquot of the sample is the same as that for the above capture workflow except the oxidation condition. The sample was oxidized by 1 mM sodium periodate in PBS (pH = 7.4) at 4 °C for 15 min to selectively oxidize sialic acids on glycans according to the previous report.28 In this way, only SialoGlycan-peptides (AllGlycan-peptides) were still captured on hydrazide resins.

After the last wash with water, the resins were dispersed in 400 μL of 100 mM TEAB (pH = 8.0). Then, 16 μL of 4% CD3O/CH3O was added into the two samples oxidized by 1 mM/10 mM NaIO4 to be heavy/light labeled, respectively, and 16 μL of 0.6 M NaBH3CN was added to all samples. The labeling reaction was carried out at room temperature for 3 h with gentle shaking. After washing the hydrazide resins with deionized water three times, the N-glycopeptides were released by adding 1 μL (500 units) of PNGase F in 200 μL of 10 mM NH4HCO3 and incubating at 37 °C overnight with gentle shaking. After centrifugation, the released deglycosylated peptides were collected for MS analysis.

Mass Spectrometry Analysis. For the analysis of glycopeptides from transferrin, an AB Sciex 5800 MALDI-TOF/TOF mass spectrometer (AB Sciex, CA) equipped with a pulsed Nd:YAG laser at 355 nm was used. Matrix DHB (25 mg/mL) was dissolved in 70% ACN (v/v) containing 1% H3PO4. A 0.5 μL mixture of the deglycosylated peptides (1 μg/μL) and 0.5 μL of DHB matrix were sequentially dropped onto the MALDI plate for MS analysis.

After pooling the light and heavy dimethyl-labeled deglycosylated peptides originated from the protein extract (1 mg) of either HCC or normal liver tissues, LC–MS/MS analysis was performed by an LTQ-Orbitrap Velos with an Accela 600 HPLC system (Thermo, San Jose, CA, U.S.A.). The system included a 5 cm C18 capillary trap column (200 μm id) and a 15 cm C18 capillary analysis column (75 μm id). The trap column was packed with C18 AQ beads (5 μm, 120 Å), and the separation column with a homemade spray tip was packed with C18 AQ beads (3 μm, 120 Å) as we reported previously.33,34 For the generation of linear gradient for reversed-phase liquid chromatography (RPLC) separation, formic acid (0.1%) water solution (buffer A) and pure acetonitrile (ACN) with 0.1% formic acid (buffer B) were used. Then, 1/10 of the mixed sample was loaded onto the C18 trap column using an autosampler for each analysis. In all of the gradient separations, the flow rate after splitting was adjusted to about 200 nL/min. The RP gradient was developed as follows: from 0% to 5% buffer B (ACN/0.1% FA) for 5 min, from 5% to 35% buffer B for 120 min, and from 35% to 80% buffer B for 5 min. After running with 80% buffer B for 10 min, the separation system was equilibrated with buffer A for 10 min. A spray voltage of 2.2 kV was applied between the spray tip and the MS interface. The temperature of the ion transfer capillary was 250 °C. The mass spectrometer was set that one full MS scan was followed by 20 MS/MS scans on the 20 most intense ions by collision-induced dissociation (CID). A value of 300 was set as the minimum signal threshold for MS/MS scan. The normalized collision energy was set at 35.0%, and the activation time was 10 ms. The mass resolution was set at 60 000 for full MS. The dynamic exclusion was set as follows: repeat count, 1; duration, 30 s; exclusion list size, 500; exclusion duration, 90 s. The scan range was set from m/z 400 to 2000. System control and data collection were carried out by Xcalibur software version 2.1.

Database Search. All the MS/MS raw data were searched by MaxQuant version 1.3.0.5 against a composite database including original and reversed human protein database of International Protein Index (IPI human v3.80 fasta, including 86 719 entries, ftp://ftp.ebi.ac.uk/pub/databases/IPI/last_release/current/). For parent ions and 0.5 Da for fragment ions. For the reverse database component. The FDR values of peptide identification rate (FDR) was determined by the equation of FDR = [2 × FP/(FP + TP)] × 100%, where TP (true positive) is the number of peptides that were identified based on sequences in the forward database component and FP (false positive) is the number of peptides that were identified based on sequences in the reverse database component. The FDR values of peptide identifications were controlled less than 1%.

RESULTS AND DISCUSSION

Workflow and Validation. The N-sialoglycan occupancy rate is defined as the molar ratio of the sialoglycan chains to all glycan chains on the same N-glycosite of a glycoprotein. In this study, we presented a proteomics approach to determine the site-specific N-sialoglycan occupancy rates for complex samples. It is well-known that terminal sialic acids can be selectively
oxidized at low sodium periodate concentration while all glycan chains can be oxidized at high sodium periodate concentration. On the basis of this fact, we developed the workflow in Figure 1 for the determination of N-sialoglycan occupancy rates. The proteome sample to be analyzed is divided into two aliquots. One part is treated with 1 mM NaIO₄ which will result in the oxidation of sialic acids on glycan. The treated sample is then incubated with hydrazide resins. The sialoglycoproteins will be captured onto the beads via the covalent link formed between the aldehyde groups on the sialic acids of glycans and the hydrazide groups on the beads. Thereafter, trypsin was added for on-beads digestion. After digestion, only glycopeptides with sialic acids are still captured on the beads. The reagents for stable isotope dimethyl labeling are added to label the captured peptides. After labeling, the N-terminus and the side chain of lysine residues were converted to dimethylamines. The isotope-labeled glycopeptides were finally released from hydrazide resins by PNGase F treatment.

Figure 1. Schematic diagram of the workflow for the determination of the site-specific N-sialoglycan occupancy rates.

A. HT-10mM_L-and-1mM_H

Asn₆₃⁰

(L 2544.0254) (H 2548.0442)

B. AHT-10mM_L-and-1mM_H

Asn₄₃²

(L 2544.0400) (H 2548.0285)

Figure 2. MALDI TOF/TOF 5800 mass spectrogram validation of sialylation-specific identification of human transferrin (HT) and asialo human transferrin (AHT). The equal amount of proteins were used in two isolations. HT (A) and AHT (B) were processed by the workflow shown in Figure 1.
Table 1. Determination of N-Sialoglycan Occupancy Rates on Human Transferrin before and after the Neuraminidase Treatment

<table>
<thead>
<tr>
<th>process</th>
<th>peptides</th>
<th>site</th>
<th>intensity (H)</th>
<th>intensity (L)</th>
<th>occupancy rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>before treatment</td>
<td>R.QQQHLFGSN#VTDCSGNFCLFR.S</td>
<td>Asn360</td>
<td>23084</td>
<td>18348</td>
<td>125.81</td>
</tr>
<tr>
<td></td>
<td>K.CGLVPVLAEYN#KS</td>
<td>Asn362</td>
<td>4362</td>
<td>4133</td>
<td>105.54</td>
</tr>
<tr>
<td>after treatment</td>
<td>R.QQQHLFGSN#VTDCSGNFCLFR.S</td>
<td>Asn360</td>
<td>4289</td>
<td>11516</td>
<td>37.24</td>
</tr>
<tr>
<td></td>
<td>K.CGLVPVLAEYN#KS</td>
<td>Asn362</td>
<td>601</td>
<td>2452</td>
<td>24.51</td>
</tr>
</tbody>
</table>

Thus, the labeled deglycosylated peptides originated from the sialylglycan-linked peptide is obtained. We call these peptides SialoGlycan-peptides. The other aliquot is treated with 10 mM NaIO₄ to oxidize all glycans on glycoproteins. The other steps are the same as above. Thus, all the glycopeptides are captured onto the beads and are isotopically labeled for this case. The deglycosylated peptides released from the hydrazide resins are termed as AllGlycan-peptides. These two types of peptides are then pooled together for MS analysis. If the SialoGlycan-peptides, and AllGlycan-peptides are labeled with heavy (H) and light (L) isotope tags, respectively, the sialylglycan occupancy rate is simply determined as the ratio of these two isotope peaks (H/L). In this study, the isolation of glycopeptides was achieved by the using hydrazide chemistry method at the protein level, which allowed all the processes including glycoprotein coupling, protease digestion, and stable isotope labeling to be performed in one reaction vessel.

Human transferrin (HT) is well-known to have sialylated biantennary complex-type oligosaccharides residing at Asn630 and Asn432. Neuraminidase is a glycoside hydrolase enzyme which can cleave sialic acid residues from glycan chains. If HT is treated with neuraminidase, then the sialylglycan occupancy rates on the above two sites should decrease dramatically. In this study, HT and asialo human transferrin (AHT), which was the HT after being treated with neuraminidase for removal of sialic acid residues, were used to evaluate the feasibility of this method for the determination of N-sialylglycan occupancy rates. The experiments were performed by following the workflow in Figure 1. In brief, two aliquots of HT were oxidized with 1 mM and 10 mM NaIO₄, respectively, then captured by hydrazide resins and digested with trypsin on beads. After washing away the non-glycopeptides, the 1 mM NaIO₄ and 10 mM NaIO₄ oxidized samples were labeled with CD₂O (heavy) or CH₂O (light) separately on hydrazide resins. Next, the newly N-linked glycopeptides were released by PNGase F from hydrazide resins. The two labeled samples, corresponding with two different oxidation conditions, were mixed together and analyzed by MALDI TOF/TOF 5800. The two tryptic deglycopeptides of HT are R.QQQHLFGSN#VTDCSGNFCLFR.S (2515 Da, Asn630) and K.CGLVPVLAEYN#KS (1476 Da, Asn432), respectively (where # denotes the deamidated asparagine and the theoretical mass of peptides is after cysteine carboxamidomethylation). The two peptides’ mass shifted into 2548/2544 Da and 1541/1533 Da ([M + H]⁺, respectively), after dimethyl heavy/light labeling (+32/28 Da and +64/56 Da). It can be seen from the MALDI mass spectra, both the deglycopeptides were detected in heavy and light forms with similar intensity (Figure 2A). The heavy-labeled peptides represent the SialoGlycan-peptides, whereas the light-labeled peptides represent the AllGlycan-peptides. The similar intensities of the two isotope-labeled peptide pairs indicate these two sites are fully occupied by sialylglycan. The N-sialylglycan occupancy rates for these two sites, Asn630 and Asn432, were determined to be 125.81% and 105.54%, respectively, by the ratio of intensity (H/L) (Table 1). The above results indicated these two sites were fully sialylated, which was consistent with the facts that sialylated biantennary complex-type oligosaccharides resided at Asn630 and Asn432 of HT. To investigate if the proposed method is able to quantitatively reflect the change in sialylglycan occupancy rates, the HT was treated with neuraminidase for removal of sialic acid residues (this sample was termed as AHT). This sample was processed as that of HT. It was observed that the heavy-labeled peptides, representing the SialoGlycan-peptides, decreased significantly after neuraminidase treatment for the both deglycopeptides (Figure 2B). The N-sialylglycan occupancy rates for these two sites, Asn630 and Asn432, were determined to be 37.24% and 24.51% (H/L), respectively (Table 1). Thus, the neuraminidase treatment resulted in significant decrease of N-sialylglycan occupancy rates on Asn630 and Asn432. Clearly, the proposed method is able to accurately determine N-sialylglycan occupancy rates.

**Determination of N-Sialylglycan Occupancy Rates in Human Liver Tissues.** This approach was then applied to analyze the N-sialylglycan occupancy rates of glycoproteins in the proteomics samples. For the determination of the N-sialylglycan occupancy rates in HCC human liver tissue, we enriched and labeled the SialoGlycan-peptides with CD₂O (heavy) and AllGlycan-peptides with CH₂O (light) following the workflow shown in Figure 1. The mixture of the isotopically labeled sample was analyzed by RPLC–MS/MS. Target/decoy search was performed to control the confidence of peptide identifications. Statistically, multiple measurements are essential to improve the quantification accuracy. Thus, three instrument replicates were performed to analyze the mixture of the isotopically labeled sample. The quantified N-sialylglycan occupancy rates could be classified into two groups: (1) the rates quantified from only one analysis, which are most likely inaccurate quantifications and will be removed; (2) the sites quantified from more than one analysis, which relative standard deviation (RSD) could be determined among the ratios. In quantitative proteomics, RSD < 50.00% are often applied to filter the quantified data to improve the quantification accuracy. Thus, the sites quantified from more than one analysis were filtered with the RSD criterion. Finally, 496 site-specific N-sialylglycan occupancy rates on 334 glycoproteins were reliably quantified with RSD < 50.00% for HCC sample (Table S1, Supporting Information). The occupancy rates were calculated by averaging the H/L ratios from at least two analyses. For instance, the N-sialylglycan occupancy rate on the N-glycosite (N93) of α-1-acid glycoprotein-2 in HCC sample was determined by a glycopeptide QNQCFYN#SSYINVQR (N# as asparagines residue detected with deamidation). The site was quantified by three replicate MS runs with ratio H/L equal to 4.23%, 4.56%, and 7.57%. Accordingly, the N-sialylglycan occupancy rate of this site was determined to be 5.45% by averaging above three ratios. Theoretically, the sites quantified by all the three replicate MS runs were the most reliable.

---

**Table 1. Determination of N-Sialoglycan Occupancy Rates on Human Transferrin before and after the Neuraminidase Treatment**

<table>
<thead>
<tr>
<th>process</th>
<th>peptides</th>
<th>site</th>
<th>intensity (H)</th>
<th>intensity (L)</th>
<th>occupancy rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>before treatment</td>
<td>R.QQQHLFGSN#VTDCSGNFCLFR.S</td>
<td>Asn360</td>
<td>23084</td>
<td>18348</td>
<td>125.81</td>
</tr>
<tr>
<td></td>
<td>K.CGLVPVLAEYN#KS</td>
<td>Asn362</td>
<td>4362</td>
<td>4133</td>
<td>105.54</td>
</tr>
<tr>
<td>after treatment</td>
<td>R.QQQHLFGSN#VTDCSGNFCLFR.S</td>
<td>Asn360</td>
<td>4289</td>
<td>11516</td>
<td>37.24</td>
</tr>
<tr>
<td></td>
<td>K.CGLVPVLAEYN#KS</td>
<td>Asn362</td>
<td>601</td>
<td>2452</td>
<td>24.51</td>
</tr>
</tbody>
</table>
However, for HCC sample, only 331 site-specific N-sialoglycan occupancy rates on 236 glycoproteins were quantified with RSD < 50.00% (Table S2, Supporting Information). To balance the sensitivity and reliability, the site-specific N-sialoglycan occupancy rates determined from more than one analysis were used. The same experiment workflow as described above was also carried out for the analysis of the N-sialoglycan occupancy rates of normal human liver tissues. At last, there were 632 site-specific N-sialoglycan occupancy rates in 394 glycoproteins quantified from more than one instrument replicate (Table S3, Supporting Information). Clearly the proposed method is able to quantify site-specific N-sialoglycan occupancy rates in a high-throughput way.

The distributions of Log2 N-sialoglycan occupancy rates determined from HCC and normal human liver tissues are given in Figure 3A. The N-sialoglycan occupancy rates for normal human liver tissues were relative higher than those for HCC. The occupancy rates on specific sites between the two samples may have significant differences, which will be compared in the following section.

Comparison of N-Sialoglycan Occupancy Rates between Human HCC and Normal Liver Tissues. The modified hydrazide chemistry, using mild periodate concentration to selectively oxidize the sialylated glycoproteins, had been applied to the analysis of the sialoglycoproteins in different proteome samples. In these studies the different abundances of sialylated glycopeptides between samples were measured. Such differences could be due to the differences either in protein abundance or in the degree of sialylation. While in this study, the site-specific N-sialoglycan occupancy rates were determined. As these rates were the part of sites occupied by sialoglycan, they did not depend on protein abundances. Thus, these rates reflected only the activities of glycosyltransferases and glycosidases in vivo. Comparing the occupancy rates on specific sites between normal and disease states could lead to discovery of new disease biomarkers. Thus, it is of interest to determine the change of occupancy rates between different samples. If the site-specific N-sialoglycan occupancy rates were determined for two samples, the relative change of N-sialoglycan (RN) occupancy rate of a specific site could be represented by their ratio:

\[ \text{RN} = \frac{\text{OR}_2}{\text{OR}_1} \]

where OR1 and OR2 are the N-sialoglycan occupancy rate of sample 1 and sample 2, respectively. For example, the RNs of the Asn630 and Asn432 on human transferrin in the above example could be determined as 0.296 (37.25%/125.81%) and 0.232 (24.51%/105.53%) after the neuraminidase treatment, indicating significant decrease. Clearly the determined RNs accurately reflected the change of occupancy rate during this process.

Among the site-specific N-sialoglycan occupancy rates determined from HCC and normal samples, 284 were obtained on the same N-glycosites in both samples (Figure 3B and Table S5 in the Supporting Information). Thus, the RN values could be determined for these sites. Figure 4 gives the Log2 values distribution of the determined RNs. It was found that 76 N-glycosites had significant N-sialoglycan occupancy rate change (RN < 0.50 or > 2.00, Table S6 in the Supporting Information). Among them, 40 site-specific N-sialoglycan occupancy rates of glycosites on 37 glycoproteins were significantly down-regulated (RN < 0.50), which indicated sialylation level decreased at least 2-fold in HCC. In the other hand, 36 site-specific N-sialoglycan occupancy rates on 31 glycoproteins were up-regulated 2-fold (RN > 2.00) in HCC.

It is reported that the MET, a transmembrane tyrosine kinase receptor for hepatocyte growth factor (HGF), plays important Figure 3. N-Sialoglycan occupancy rates determined from HCC and normal human liver tissues. (A) The distribution of Log2 N-sialoglycan occupancy rates. (B) The overlap of N-glycosites quantified with N-sialoglycan occupancy rates.

Figure 4. Log2 RNs distribution of the changes in N-sialoglycan occupancy rates between HCC and normal human liver tissues.
roles in various types of human cancers including HCC. The N-sialoglycan occupancy rate at N607 of MET was quantified using our developed method, and it is down-regulated (RN = 0.02) in HCC tissues. For another glycoprotein cathepsin L1 (CTSL), which was associated with prostate and breast cancer cells, the up-regulated N-sialoglycan occupancy rate at N221 (RN = 2.66) may perform a crucial role in the dissemination of tumor cells in HCC development. It is interesting that the change of occupancy rates for different sites on the same protein could be either down- or up-regulated. For example, for the five quantified N-glycosites on apolipoprotein B-100 (ApoB-100), two sites (N3465 and N2982) were significantly down-regulated with RNs less than 0.50, one site (N185) was significantly up-regulated with RNs greater than 2.00, and the left two sites (N1523 and N3411) were only moderate regulated with RNs in the range of 0.50–2.00. The above example indicated that the glycosylation was regulated in a site-specific way. The full list of the site-specific RNs quantified with significant change between HCC and normal tissues may be potential biomarkers for future exploration.

CONCLUSIONS

In summary, a modified hydrazide chemistry method has been successfully developed to quantify the site-specific N-sialoglycan occupancy rates on glycoproteins. The method was validated by analysis of a standard glycoprotein, human transferrin. It was then applied to analyze of the proteins extracted from HCC and normal human liver tissues. Significant changes were observed for 76 N-sialoglycan occupancy rates between HCC and normal liver tissues. In conventional biomarker discovery studies, the change in glycoprotein abundance is typically used as the criteria to screen biomarkers. In this study, we found numerous N-glycosites with significant change in N-sialoglycan occupancy rates between disease and normal tissue. Thus, the change in N-sialoglycan occupancy rates, which reflects the activities of glycosyltransferases and glycosidases in vivo, could be used as the new criteria to screen new type of biomarkers. This study presented an effective platform to quantify the N-sialoglycan occupancy rates.

ASSOCIATED CONTENT

Supporting Information
Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors
*Phone: +86-411-84379620. Fax: +86-411-84379620. E-mail: mingliang@dicp.ac.cn.
*Phone: +86-411-84379610. Fax: +86-411-84379620. E-mail: hanfazou@dicp.ac.cn.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by China State Key Basic Research Program Grants (2013CB911202, 2012CB910101, and 2012CB910604), the Creative Research Group Project of NSFC (21321064), the National Natural Science Foundation of China (21235006, 81161120540, and 81361128015), and Chinese National Key Project (2012ZX10002009-011).

REFERENCES

(9) Hakomori, S. Cancer Res. 1996, 56, 5309−5318.


