Capture and Dimethyl Labeling of Glycopeptides on Hydrazide Beads for Quantitative Glycoproteomics Analysis

Zhen Sun, Hongqiang Qin, Fangjun Wang, Kai Cheng, Mingming Dong, Mingliang Ye, and Hanfa Zou*

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

Supporting Information

ABSTRACT: Incorporation of isotopic tag onto peptides via chemical labeling is a popular approach for quantitative proteomics. Chemical labeling via solution based methods usually lead to a tedious process and sample loss because several sample preparation steps including buffer exchange and desalting are performed. In this study, a solid phase based labeling approach by integration of glycopeptide enrichment and stable isotope labeling on hydrazide beads was developed for relative quantification of protein glycosylation, by which enrichment, washing, labeling, and release of the glycopeptides were all performed on the hydrazide beads sequentially. This approach was proved to be accurate in quantitative glycoproteome analysis and have good linearity range with 2 orders of magnitude for quantification of glycopeptides. Compared with dimethyl labeling conventionally performed in solution, the developed approach has better enrichment recovery (10−330% improvement) and high detection sensitivity in which 42% of annotated glycosites (vs 26%) still can be quantified using only 10 μg of four standard glycoprotein mixtures and 400 μg of bovine serum albumin interference as starting sample. The applicability of the approach for quantitative glycopeptide profiling was also explored by differential analysis of glycoproteome between human normal serum and liver cancer serum.

Q uantitative proteome analysis using stable isotope labeling has emerged as a valuable tool to screen the disease-specific biomarkers in clinic samples. Chemical labeling through the specific and complete reactions with terminal or side groups is widely used at either the protein or peptide level, including isotope-coded affinity tags (ICAT),1 isotope tags for relative and absolute quantification (iTRAQ),2 and so on. Chemical labeling is typically performed in solution,3−4 which contains several procedures including buffer exchange and desalting. These additional steps may cause sample loss, which will influence the quantification accuracy and decrease the detection sensitivity, but solid phase based chemical labeling can facilitate these steps by simple centrifugation. Combining the capture of post-translational modification (PTM) peptides with chemical labeling will simplify the sample preparation procedure and has great potential in quantification of protein or peptide modification. Qin et al.5 have developed an in situ enrichment and isotope labeling strategy for relative quantification of endogenous phosphopeptides. As the enrichment, washing, labeling, and eluting of the endogenous phosphopeptides were all performed on the same affinity beads sequentially, this strategy was demonstrated to have improved recovery and sensitivity. Glycosylation is one of the most important post-translational modifications, which plays an important role in many molecular and cellular processes and particularly was involved in cancer progression and immune responses.6,7 Many clinical protein biomarkers used are glycoproteins, e.g., prostate-specific antigen (PSA) in prostate cancer, cancer antigen (CA) 125 in ovarian cancer, and Her2/neu in breast cancer.8 Thus, discovery of a candidate as a new biomarker by quantitative glycoproteomics analysis is very important.

Early in 2003, glycoprotein capture and stable isotope labeling by succinic anhydride were sequentially carried out on the hydrazide beads by Zhang et al.8 Then, this method was further improved by performing hydrazide chemistry on the peptide level.9 Succinic anhydride, the labeling reagent, targets the peptide N-terminus and the epsilon-amino group of lysine residues, which surpass the ICAT method in quantifying PTM peptides. The ICAT method is obviously not suitable for quantifying the significant number of peptides that do not contain any cysteine residues, so it is of limited use for analysis of post-translational modifications and splice isoforms. However, the succinic anhydride method requires incubation of the sample with succinic anhydride solution repeatedly to achieve the completeness of labeling the reaction. Additionally, the analysis may be complicated because the hydroxyl group on serine can be esterified by succinic anhydride, and phenolate and imidazolyl groups will also be partially acylated, unless treatment of hydroxylamine was performed.10 The dimethyl labeling method using formaldehyde can also globally label the peptide N-terminus and the epsilon-amino group of Lys through reductive amination.11 Since hydrazide chemistry can isolate glycopeptides with specificity of more than 90%,12 here, we performed the solid phase enrichment and dimethyl labeling...
of glycopeptides sequentially on the hydrazide beads for quantitative glycoproteomics analysis. It was found that this solid phase based labeling approach has much better enrichment recovery and high detection sensitivity than the dimethyl labeling conventionally performed in solution.

■ EXPERIMENTAL SECTION

Sample Preparation. Albumin from chicken egg white (Ovalbumin), fetuin from fetal calf serum (two isoforms, fetuin-A and fetuin-B), transferrin from human blood plasma (Transferrin), and α1-acid glycoprotein from human plasma (two isoforms, AGP1 and AGP2) were four standard glycosylated proteins, while bovine serum albumin (BSA) was a standard nonglycosylated protein. Human serum from healthy volunteers (n = 16) and liver cancer patients (n = 11) was provided by Second Affiliated Hospital of Dalian Medical University using the standard clinical protocols and stored at −80 °C before analysis. These standard proteins were mixed first according to the experimental design. Then, the human serum or the standard protein mixture was digested with trypsin as previously reported by Wang et al.13 The tryptic peptides were desalted with a homemade C18 solid phase extraction column and then dried down in a Speed Vac (Thermo, CA).

Solid Phase Based Labeling Method. The dried tryptic peptides were reconstituted in oxidation buffer (100 mM NaAc, 150 mM NaCl, pH = 5.5), and a final concentration of 10 mM sodium periodate was added. The reaction was kept in the dark for 1 h and quenched by adding 20 mM sodium thiosulfate. One milligram of the oxidized glycosylated peptides were incubated with 50 μL of Affi-Gel HZ hydrazide beads (slurry volume) (Bio-Rad, USA) overnight at room temperature. After washing the hydrazide beads with 1.5 M NaCl, 80% acetonitrile (ACN), and 100 mM NH4HCO3 sequentially, 100 μL of 100 mM triethyl ammonium bicarbonate (TEAB) was added to the hydrazide beads. Then, 8 μL of 4% CH3O/CD2O was added to the sample to be light/heavy labeled, respectively, and 8 μL of 0.6 M NaBH3CN was added to both samples. The dimethyl labeling reaction was carried out at room temperature for 2 h with gentle shaking. Compared with the solution based labeling method, the amount of label reagents was doubled to accelerate the labeling reaction and the reaction time was lengthened to make the labeling reaction efficient. After rinsing the resulting hydrazide beads twice with deionized water, the glycopeptides were released by adding 500 unit PNGase F (New England Biolabs, USA) in 10 mM NH4HCO3 to the resin and incubating at 37 °C overnight with gentle shaking. The released deglycosylated peptides were carefully collected by gentle centrifugation.

Solution Based Labeling Method. No matter if the stable isotope labeling was performed before or after glycopeptide enrichment, the prepared peptides were labeled with dimethyl labeling reagents according to the published protocol.14 In brief, the prelabeled peptides were resuspended in 100 μL of 100 mM TEAB. Then, 4 μL of 4% CH3O/CD2O was added to the sample to be light/heavy labeled, respectively, and 4 μL of 0.6 M NaBH3CN was added to both samples. The labeling reaction was carried out at room temperature for 1 h with intense vortexing. After 16 μL of 1% ammonia was added to quench the reaction, the sample was desalted with the C18 SPE column. For more details, see Supplemental Experimental Procedures in the Supporting Information.

Mass Spectrometry Data Acquisition and Identification. The labeled deglycosylated peptides were analyzed on the Triple TOF 5600 mass spectrometer (AB SCIEX, USA) coupled to the Waters NanoACQUITY UPLC system via a nanospray source. Identification of peptides and proteins and relative quantification of their abundance were performed using the Paragon Algorithm15 in Protein Pilot software (AB SCIEX, USA) with trypsin specificity in “Thorough ID” mode. See Supplemental Experimental Procedures in the Supporting Information for further information.

■ RESULTS AND DISCUSSION

Quantification of Protein Glycosylation by Solid Phase Based Labeling Method. As presented in Figure 1, the oxidized glycopeptides were captured onto hydrazide beads and labeled with light/heavy dimethyl labels on the beads, and then, the labeled glycopeptides were released by PNGase F. Finally, the released light and heavy labeled peptides are pooled together for LC-MS/MS analysis. It is obvious that labeling glycopeptides on hydrazide beads has many fewer steps than that by solution based labeling methods. Because the changing of reaction buffers can be performed easily by removing supernatant after simple centrifugation, this approach could decrease sample loss and increase detection sensitivity greatly.

Ovalbumin, fetuin, transferrin, and AGP are glycosylated proteins with 19 N-linked glycosites annotated by Swiss-Prot (Table S1, Supporting Information). The mixture of these four glycoproteins was used to evaluate the performance of the solid phase based labeling method. An equal amount of the mixture was, respectively, labeled with light/heavy dimethyl labels and then pooled together for analysis. As shown in Figure S1,
Supporting Information, 16 glycosites can be quantified by the solid phase based labeling method in three technical replicates around H/L ratio = 1 with RSD < 50% with only one exception of FETUA-N99. Furthermore, the feasibility of this solid phase based labeling method was further investigated by analysis of an equal amount of human normal serum. 121 glycosites can be successfully quantified in three technical replicates around H/L ratio = 1 with RSD < 50% but with 3 exceptions (Figure S2, Supporting Information). In other words, 97% of the observed H/L ratios of glycosites were all distributed between 0.50 and 1.50 (theoretically 1.00, 3 from 121 glycosites out of this range). These results indicate that quantification by the solid phase based labeling method on hydrazide beads is accurate.

The Recovery of Quantification by Solid Phase Based Labeling Method. For the solid phase based labeling method, the enrichment of glycopeptides onto the hydrazide beads takes place before labeling of glycopeptides on-beads. For the solution based labeling method, the isotopic labeling could be performed after the glycopeptides are released from the beads (enrichment before labeling, method A) or before the capture of glycopeptides onto the hydrazide beads (enrichment after labeling), and for the solution based labeling method (enrichment after labeling), the oxidation could be performed either before or after labeling. Thus, this method could be further classified into two types: one is with the order of oxidation, labeling, and enrichment (method B) and another one is with the order of labeling, oxidation, and enrichment (method C). The recovery of glycopeptides for the solid phase based labeling method was compared with those solution based labeling methods.

Four standard glycoprotein mixtures were also used as the sample to evaluate these methods. As shown in Table S2, Supporting Information, it can be seen that a similar number of glycosites can be quantified by the solid phase based labeling method and solution based labeling method A. However, much fewer glycosites were quantified by solution based labeling methods B and C. To facilitate the evaluation of recovery, the samples were labeled with heavy dimethyl reagent by the solid phase based method and labeled with light dimethyl reagent, respectively, by three solution based methods. The light peptides by three solution based methods were combined, respectively, with the heavy peptides with equal amount of starting sample for LC-MS/MS analysis. It was found that the intensity of glycopeptides by the solid phase based labeling method was always slightly higher (1.1−4.3-fold, 2.1-fold on average) than that obtained by the solution based labeling method A, which was much more higher than that obtained by solution based labeling methods B (2.0−10.8-fold, 5.0-fold on average) and C (3.9−74.9-fold, 27.9-fold on average) (Figure 2). On the whole, the solid phase based labeling method had the least sample loss, while the most sample lost by solution based labeling method C had labeling performed before the enrichment.

Linearity Range of Quantification by Solid Phase Based Labeling Method. Additionally, the linearity range for quantification of glycopeptides was investigated using two mixtures (A and B). Mixture A labeled with light dimethyl reagent was composed of different molar ratios of AGP versus fetuin from 10:1 to 1:10, while the ratio of AGP to fetuin in mixture B which was heavy labeled was just the opposite. In other words, two glycoproteins of AGP and fetuin with H/L ratio from 10:1 to 1:10 were used to investigate the linearity range of the solid phase based labeling method. Also, a control experiment with a 1:1 AGP/fetuin mixture captured by hydrazide beads without stable isotope labeling was carried out. As annotated by Swiss-Prot, AGP and fetuin have 16 N-linked glycosites. In the control experiment, a total of 14 glycosites were identified, while 13 glycosites were quantified when the AGP/fetuin ratios of mixtures A and B were both 1:1. Therefore, performing dimethyl labeling on the hydrazide beads has no significant influence on the identification of glycopeptides. In total, 13 glycosites of AGP and fetuin were quantified in at least one technical replicate of analysis of different AGP/fetuin ratios. Most of the quantified glycosites yield great calibration curves with a slope close to 1 and R² value more than 0.9 (Table S3, Supporting Information).
general, the good linearity range of our method for quantification of glycopeptides can be extended to 2 orders of magnitude.

Sensitivity of Quantification by Solid Phase Based Labeling Method. Furthermore, the sensitivity of the solid phase based labeling method was compared with solution based labeling method A. Different amounts of four standard glycoprotein mixtures (FSGP for short) and certain amounts of BSA interference were mixed as the sample to evaluate the sensitivity. As described above, the solid phase based labeling method and solution based labeling method A quantified a similar number of glycosites when 1 mg of starting sample was used. In this experiment of sensitivity comparison, the same result was obtained when 1 mg of FSGP was used even if 0.4 mg of BSA was added as interference. The number of glycosites quantified by the solid phase based labeling method dropped to 12 ± 1 or 8 ± 0 when the amount of FSGP decreased to 0.1 or 0.01 mg, but the number dropped much faster by solution based labeling method A (Table S4, Supporting Information). The intensity of quantified glycosites by both labeling methods was also compared. As observed above, the intensity of glycosites quantified by the solid phase labeling method was always higher (7.6-fold on average) than that quantified by solution based labeling method A when 1 mg of FSGP with 0.4 mg of BSA interference was included in the sample (Table S5, Supporting Information). The difference between the intensity of quantified glycosites by the solid phase based labeling method and by solution based labeling method A became more significant when less FSGP was used (42.6−48.1-fold improvement on average) (Tables S6 and S7, Supporting Information). Thus, it can be concluded that the solid phase based method always has higher detection sensitivity than solution based labeling method A, and this advantage is more obvious when the starting sample is less (Figure 3).

Differential Glycoproteome Analysis of Human Normal Serum and Liver Cancer Serum. Hepatocellular carcinoma (HCC) is the fifth in cancer incidence worldwide and the third leading cause of cancer death. Most cases of HCC are chronic hepatitis B (HBV) or HCV infection, and long latency exists between infection and disease onset of HCC, so it is necessary to screen for biomarkers of disease in the high-risk populations for early detection well before the onset of advanced disease. Thus, our solid phase based labeling method coupled with glycopeptide enrichment was applied to investigate the alterations of glycosylation in the human serum with and without liver cancer. In total, 117 glycosites were quantified with normal serum light labels and cancer serum heavy labels. As described above, the H/L ratios around 1 with RSD < 50% were considered without significant change. Thus, the ratios of H/L out of the range of 0.50−1.50 could be considered as significantly changed for real sample analysis with the same filtering criterion. It was found that 25 glycosites were quantified as significant expression changes with the H/L ratio >1.5 or <0.5 (Table S8, Supporting Information). Galectin-3-binding protein was quantified with five glycosites, i.e., N69, N125, N192, N398, and N551, which all exhibited significant up-regulated expression. The glycosite of N551 was significantly up-regulated compared with the other four glycosites. Thus, the occupancy rate of glycosites of N551 may be different from the other four glycosites for those two states. The galectin-3 binding protein is a 90 kDa tumor antigen and is present at elevated levels in the serum of subpopulations of patients with various types of cancer. Correale et al. demonstrated that 90K sensitivity was higher than AFP in all their tested groups of hepatic patients. The statistical analysis further made by Lacovazzi et al. in 2003 showed a highly significant influence on overall survival of age below 70 years and 90K serum levels below the cutoff of 14 ng/mL. Serum AFP had positive prognostic value only when it was associated with 90K levels. Therefore, the galectin-3 binding protein may be proposed as a clinical useful tumor marker, but further investigations are still needed.

■ CONCLUSIONS

The cost-effective isotopic formaldehyde and fast and simple derivatization procedure are the great advantages of the dimethyl labeling method. In addition, the charge state is not changed significantly by dimethyl modification, and the ionization efficiency of the fragment is more likely to be conserved. Therefore, it is worthwhile to develop a solid phase based labeling approach by integration of glycopeptide enrichment and dimethyl labeling on hydrazide beads for relative quantification of protein glycosylation. This integrated approach was proved to be accurate in the quantitative glycoproteomics analysis, and the good linearity range for quantification of glycopeptides by this approach can be extended to 2 orders of magnitude. Compared with the solution based labeling method, the developed approach has better enrichment recovery (10−330% improvement) and high detection sensitivity in which 42% of annotated glycosites (vs 26%) still can be quantified using only 10 μg of four standard glycoprotein mixtures and 400 μg of bovine serum album interference as starting sample. The galectin-3 binding protein was proposed as a potential clinical useful tumor marker by analyzing the difference of protein glycosylation between human normal serum and liver cancer serum with this approach. On the whole, this is a solid phase labeling approach with convenient and fast-processing, which could have great potential application in high-throughput analyses of biological samples for screening and discovery of disease-specific
biomarkers, investigating the roles of important proteins within different physiological processes, and so on.

**ASSOCIATED CONTENT**

<table>
<thead>
<tr>
<th>Supporting Information</th>
</tr>
</thead>
</table>
| Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

**Corresponding Author**

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

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

Financial support of this work is gratefully acknowledged from the Creative Research Group Project by NSFC (21021004), Scientific Research Equipment Development Projects of CAS (P020090520), the China State Key Basic Research Program Grant (212CB910101, 212CB910601), National Key Special Program on Infection diseases (212ZX10002009-011), the Analytical Method Innovation Program of MOST (2009IM31800, 2010IM30500), and the Knowledge Innovation program of DICP to H.F.Z.

**REFERENCES**