Glycoconj J (2012) 29:433–443
DOI 10.1007/s10719-012-9404-3

Enhanced N-glycosylation site exploitation of sialoglycopeptides by peptide IPG-IEF assisted TiO_2 chromatography

Weiqian Cao · Jing Cao · Jiangming Huang · Lei Zhang · Jun Yao · Haoqi Xu · Pengyuan Yang

Received: 14 April 2012 / Revised: 22 May 2012 / Accepted: 23 May 2012 / Published online: 8 July 2012
© Springer Science+Business Media, LLC 2012

Abstract Playing an important role in a broad range of biological and pathological processes, sialylation has been drawing wide interest. The efficient sialoglycopeptides enrichment methods are therefore attracting considerable attention. In this paper, we first compared two conventional enrichment methods, lectin and TiO_2, and analyzed their characteristics. Furthermore, considering the highly negatively charged nature of sialic acids, we developed a new strategy, peptide immobilized pH gradient isoelectric focusing (IPG-IEF) assisted TiO_2 chromatography (PIAT), for the highly efficient enrichment of sialoglycopeptides. In this method, peptides were first separated into 24 fractions using peptide IPG-IEF. Sialoglycopeptides were relatively concentrated in low-pH fractions of the immobilized pH strips and were captured using TiO_2 chromatography. As a result, 614 N-glycosylation sites were identified in 582 sialoglycopeptides within 322 sialoglycoproteins from rat liver using PIAT. To our knowledge, this work represents one of the most comprehensive sialoglycoproteomic analyses in general and exhibits the largest database of sialoglycoproteome in rat liver currently. So the new strategy introduced here exhibits high efficiency and universality in the sialoglycopeptide enrichment, and is a powerful tool for sialoglycoproteome exploration.

Keywords Sialoglycopeptides · N-glycosylation site · Peptide IPG-IEF · Lectin · TiO_2 chromatography · LC-ESI-MS/MS

Weiqian Cao and Jing Cao contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10719-012-9404-3) contains supplementary material, which is available to authorized users.

W. Cao · J. Cao · J. Huang · L. Zhang · J. Yao · H. Xu · P. Yang
Institutes of Biomedical Sciences and Department of Chemistry, Fudan University,
Shanghai 200433, China

P. Yang (✉)
Institutes of Biomedical Sciences and Department of Chemistry, Fudan University,
Shanghai 200032, China
e-mail: pyyang@fudan.edu.cn

Abbreviations

PIAT	Peptide IPG-IEF assistant TiO_2 chromatography
SAs	Sialic acids
MW	Molecular weight
pI	Isoelectric point
LC	Liquid chromatography
MS/MS	Tandem mass spectrometry
PNGase F	N-glycosidase F

Introduction

Protein glycosylation is one of the most common and important post-translation modifications in nature. The glycan moieties of proteins are composed of a series of monosaccharides with prominent micro heterogeneity of carbohydrate chains based on branching patterns and monosaccharides composition. The sialic acids (SAs) are typically located on the non-reducing end of glycans [1]. The SAs are a family of nonulosonic acids with a shared nine-carbon back-bone and more than 50 structural variations known in nature [2]. Since the first discovery of SAs [3] in bovine submaxillary mucin in 1936, many studies have shown the outstanding role of this biomolecule. The SAs have been demonstrated to be involved in a broad range of biological processes, such as cell signaling,

microbial attachment [4, 5] and host-pathogen interaction [6–8], especially for flu viruses [9–11]. The over-representation of SAs on the surface of glycoproteins is thought to be associated with a plethora of diseases, particularly cancer [12–14], such as cancer invasion and metastasis [15–17]. For example, increasing sialylation can attenuate EGFR-mediated invasion of lung cancer cells [18]. In addition, different branching patterns of SAs are also linked with cancer aggravation. For example, the α-(2-3) linked SA on prostate-specific antigen can potentially discriminate malignant from benign prostate conditions [19].

Therefore, various methods of isolating sialoglycoproteins/sialoglycopeptides from complex mixtures have been developed to analyze the sialylation status in organisms. The main methodological approaches include chemical labeling, metabolic labeling, affinity binding and titanium dioxide (TiO_2) multidentate binding. The metabolic labeling method generally utilizes the SA biosynthetic pathway to incorporate unnatural monosaccharides into cells or organs. These incorporated monosaccharides bearing bioorthogonal functional handles can be captured through Staudinger ligation [20] or Huisgen cycloaddition [21] or biotinylated alkyne capture reagent [22]. Although metabolic labeling has been used as an efficient tool for capturing SA-containing proteins, it can only be applied to live cells or organisms. The chemical labeling approach, typically called modified hydrazide chemistry has recently been developed. In this method, the sialoglycoproteins are selectively periodate-oxidized, captured on hydrazide beads and released via the acid hydrolysis of SA. With this method, a total of 36 N-linked and 44 O-linked glycosylation sites on glycoproteins were identified from human cerebrospinal fluid [23]. A similar method that used mild periodate oxidation to introduce aldehydes onto the cell surface of SAs for subsequent ligation with aminooxy-biotin has also been reported [24]. Recently, modified hydrazide chemistry has been used to explore altered expressions of sialoglycoproteins in breast cancer with the identification of 90 N-sialoglycoproteins [25]. Although this method has been successfully used as an efficient tool obtaining many important findings, the true selectivity of this method can't be addressed, because the SAs can't be retained on the glycopeptides.

Lectin-affinity chromatography and TiO_2 enrichment are two other methods widely used in sialoglycoproteins exploration. Sambucus Nigra Lectin (SNA) and Maackia Amurensis Lectin II (MAL II) are usually combined or used separately to recognize terminal SAs with α-(2-6) and α-(2-3) linked to Gal/GalNAc based on their affinities to SAs [26, 27]. For example, serial lectin affinity chromatography has been developed to analyze SA-containing complex N–glycoproteins in human serum [28]. Recently, a "filter aided sample preparation"-based method in which glycopeptides are enriched by binding to lectins on the top of a filter has been developed [29]. Using this method, an unprecedented depth of glycoproteome coverage has been shown [30]. In this method, lectins do not need to be coupled with a solid support or filled in a chromatography column. Therefore any lectin or lectin mixtures can be employed, including SNA and MAL II. In many studies, TiO_2, which has a high affinity toward negatively charged molecules, has been successfully used to capture sialoglycopeptides/sialoglycoproteins [31, 32] after the alkaline phosphatase treatment of samples. With these methods, the enriched sialoglycopeptides can be directly analyzed to determine the glycosylation sites and the glycan structure without glycan cleavage or they can be subjected to deglycosylation to map the glycosylation sites. For the large-scale elucidation of N-glycosylation sites, the glycans of sialoglycopeptides were generally removed by PNGase F, which can convert asparagines to aspartic with a 0.98 Da mass shift detectable by mass spectrometry (MS). Thus, N-glycosylation sites are determined. For the high coverage of N-glycosylation sites of sialoglycopeptides, these methods are often used with some chromatographic separation technology [33], such as strong cation exchange [34] and HILIC chromatography [35], followed by LC-MS/MS analysis. However, because of the complexity and heterogeneity of glycosylation, none of these methods can capture all the sialoglycoproteins/sialoglycopeptides from complex biological mixtures. Therefore, more efficient methods of isolating sialoglycopeptides from mixture are arousing attention.

Isoelectric focusing (IEF) has been conventionally used as a first dimension of two-dimensional electrophoresis (2-DE) in proteomics. Recently, peptide immobilized pH gradient IEF (IPG-IEF) has become increasingly popular [36–42] because of its high reproducibility and resolution (~0.1 pH) [39]. The samples separated using peptide IPG-IEF are recovered in liquid form, which is compatible with downstream analysis, usually LC-MS. Compared with the native, non-modified sequence, peptide modifications that can shift peptide pI will influence the movement of peptides on IPG strips [43]. A previous study used peptide IEF to calculate the pI of phosphorylated peptides and found that the peptide pI decreased when phosphoric acid was attached to peptides [44]. Similarly, the pI of sialoglycopeptides with highly negatively charged SAs will also decrease. The sialoglycopeptides with low pI would move to the anode of the IPG strips and could be relatively concentrated in low-pH fractions of the IPG strips at the end of IEF. Consequently, sialoglycopeptides could be already quite centralized before enrichment. Therefore, peptide IPG-IEF was utilized in the current study of sialoglycopeptides based on the nature of SAs and the potency of the peptide IPG-IEF method.

To enhance N-glycosylation site exploitation of sialoglycopeptides, here we first employed and compared the two

mostly used methods, lectin affinity and TiO_2 chromatography. Base on the result, TiO_2 chromatography was chosen for further study. Thus, for the first time, a novel strategy, peptide IPG-IEF assistant TiO_2 Chromatography (PIAT) was developed for large-scale exploration of N-glycosylation sites of sialoglycopeptides. In this strategy, peptide IPG-IEF was used as the first separation/ enrichment step, then the resulting fractions were further enriched with TiO_2, and followed by LC-ESI-MS/MS analysis. As a result, a more comprehensive, less biased sialoglycoproteome was acquired.

Materials and methods

Chemicals and reagents

Biotinylated SNA, unconjugated SNA, and unconjugated MAL II were purchased from Vector Laboratories (Burlingame, CA). Sequencing grade trypsin and alkaline phosphatase were obtained from Promega (Madison, WI, USA). PNGase F (glycerol free) was obtained from New England Biolabs (Ipswich, MA). Complete protease inhibitor mixture tablets were purchased from Roche Applied Science (Basel, Switzerland). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Titansphere TiO_2 (5-μm chromatographic material) was purchased from GL Sciences (Tokyo, Japan). Microcon devices YM-30 were purchased from Millipore (Bedford, MA). Sep-Pak C18 columns were purchased from Waters (Milford, MA). Rat liver tissues were obtained from healthy adult rats. Water used in all experiments was produced by a Milli-Q Plus system from Millipore (Bedford, MA).

Sample preparation

The total proteins were extracted from rat liver tissues using liquid nitrogen grinding method. Briefly, the liver tissues were rinsed and diced into small pieces in ice-cold PBS buffer. Then, the tissues were ground in liquid nitrogen, and added to a lysis buffer (8 M urea, 2 M thiourea, 1 mM PMSF and cocktail). After being incubated on ice for 2 h, centrifuged at 15,000 g for 45 min, the supernatant was collected and stored at -80 °C. The protein concentration was determined using the Bradford method.

Lectin blot analysis

The proteins of each lane (20 μg) were separated by SDS-PAGE and then were transferred to PVDF membranes. The membranes were blocked in 5 % BSA in TBST (0.1 % Tween 20 in TBS) for 1 h, and then incubated with biotinylated SNA dissolved in BSA/TBST (3 μg/mL) for 30 min at room temperature. The membranes were washed subsequently with TBST, incubated with strepavidin-HRP for 30 min at room temperature. Finally, the membranes were washed with TBST and visualized using an ECL plus detection system (GE, Amersham).

Protein digestion and alkaline phosphatase treatment

The proteins were in-solution digested using the method previously described [45] with minor modification. The total proteins were reduced by 10 mM dithiothreitol at 37 °C for 1 h, then alkylated by 20 mM Iodoacetamide in the dark, at room temperature for 30 min. After being diluted 10-fold with 50 mM NH_4HCO_3 buffer, to the proteins was added trypsin (1:50, w/w) and incubated at 37 °C overnight. The tryptic peptides were desalted by passing through a C18 column. Then, the peptide solution was dried in a vacuum centrifuge.

The tryptic peptides were dephosphorylated using alkaline phosphatase in 50 mM NH_4HCO_3 buffer at 37 °C for 12 h. After that, Trifluoroacetic acid (TFA) was added to stop the reaction and the peptide mixture was dried in a vacuum centrifuge.

SNA/MAL II -affinity enrichment of N-sialoglycopeptides

The N-sialoglycopeptides were enriched using filter aided capture and elution (FACE) method previously described [30] with minor modification. Briefly, the digested peptides were mixed with lectin solution containing either SNA or MAL II (1:2, w/w) in 2×binding buffer (2×BB, 40 mM Tris/HCl pH 7.6, 2 mM $MnCl_2$, 2 mM $CaCl_2$, 1 M NaCl). The mixtures were transferred to YM-30 filter units. After being incubated at 4 °C overnight, the unbounded peptides were removed via centrifugation at 14,000 g at 18 °C for 10 min. The captured peptides were washed four times with 1×BB and twice with 50 mM NH_4HCO_3 buffer. Then, PNGase F in 50 mM NH_4HCO_3 buffer was added to the filter unities and the samples were incubated at 37 °C overnight. Finally, the deglycosylated peptides were collected through centrifugation and dried in a vacuum centrifuge.

Enrichment of N-sialoglycopeptides using TiO_2

The N-sialoglycopeptides were enriched using the method previously described [35] with a little modification. The dephosphorylated peptides were mixed with TiO_2 beads (1:6, w/w) in loading buffer (1 M glycolic acid in 80 % (V/V) acetonitrile, 5 % (V/V) TFA).The mixtures were then incubated for 30 min at room temperature with gentle agitation. The beads were precipitated through centrifugation at 1,000 g for 1 min and washed with loading buffer, washing buffer 1 (80 % (V/V) acetonitrile, 1 % (V/V) TFA) for 1 min

and 15 s, respectively. Washing buffer 2 (20 % (V/V) acetonitrile, 0.1 % (V/V) TFA) was added to the beads. Then the beads were dried in a vacuum centrifuge. The sialoglycopeptides were eluted by incubating the beads with eluting buffer (25 % NH_4OH in ultrapure water, pH 11.3) at room temperature for 20 min. The supernatant containing sialoglycopeptides were collected through centrifugation at 1,000 g for 5 min and then dried in a vacuum centrifuge. The dried peptides were resuspended in 50 mM NH_4HCO_3 buffer and deglycosylated with PNGase F. Finally, the deglycosylated peptides were dried in a vacuum centrifuge.

Peptide IPG-IEF assisted TiO_2 chromatography (PIAT)

IEF assisted TiO_2 separation of N-sialoglycopeptides

The total dephosphorylated peptides of 1 mg proteins were separated using Agilent 3100 OFFGEL fractionator (Agilent, G3100AA). An OFFGEL High Res Kit, pH 3–10 (Agilent, 5188–6425) was used. The peptides were separated according to the procedure provided in the manual. Briefly, the frames and 24 cm IPG strips were placed on the trays. The strips were rehydrated for 15 min with 40 μL/well of 1×peptide OFFGEL stock solution (5 % glycerol, 2 % OFFGEL Buffer, pH 3–10). The peptides were resuspended in 1×peptide OFFGEL stock solution and then added into each well (150 μL/well). The cover seal was set into place, and the cover fluid (mineral oil) was pipetted onto the gel strip ends. The fractionations were focused for 50 kV·h with a maximum current of 50 μA and power of 200 mW. Then, the peptides in each well were collected and desalted through C18 columns. Each of 24 peptide portions obtained from IEF was used for TiO_2 enrichment of N-sialoglycopeptide using the method mentioned above (in Section 2.6).

Lectin dot blot analysis

The PVDF membrane was soaked in methanol for several seconds and in TBS for 5 min. Each of the 24 peptide fractions was resuspended in TBS and spotted onto the membrane. After the membrane dried, it was treated according to the method mentioned (in Section 2.3). The relative grayscale of each dot was calculated using software integrated with Las3000.

Nano-LC-ESI -MS/MS analysis

Nano-LC-ESI-MS/MS analysis was performed as the method previously described [46] with minor modification. The enriched deglycosylated peptides were suspended in 5 % ACN containing 0.1 % formic acid (phase A), separated by a 15-cm reverse phase column (100 μm i.d., MICHROM Bioresources, Inc., Auburn, CA). The peptides were analyzed using a LC-20AB system (Shimadzu, Tokyo, Japan) connected to an LTQ Orbitrap mass spectrometer (Thermo Electron, Bremen, Germany) equipped with an online nano-electrospray ion source (Michrom Bioresources, Auburn, CA).

The peptide mixtures were injected onto the trap-column at a flow of 20 μL/min and eluted directly into an ESI source of MS with a gradient of 5 %–45 % phase B (95 % ACN with 0.1 % formic acid) over 130 min at a constant column-tip flow rate of 500 nL/min. The electrospray voltage was used at 1.7 kV. The eluted peptides were analyzed by MS and data-dependent MS/MS acquisition, selecting the 8 most abundant precursor ions for MS/MS with dynamic exclusion durations of 90 s. The scan range was set from m/z 350 to m/z 1800. The samples enriched by lectin or TiO_2 were performed under the same LC-MS/MS condition.

Data exploration

The MS/MS data was searched against Swiss-Prot database by SEQUEST. The parameters were set as follows: enzyme, trypsin (partially enzymatic); maximum missed cleavages (MCs), two; fixed modification, carboxyamidomethylation (C, + 57.02); variable modifications, oxidation (M, + 15.99), deamidation (N, + 0.98); peptide tolerance, 10 ppm; and fragment tolerance, ± 1.0 Da.

The database search results were further analyzed statistically using PeptideProphet [47]. A minimum peptideprophet probability score (P) filter of 0.9 was selected to remove low-probability result. Here only those peptides that passed the peptide probability threshold of 0.9 and the proteins that passed the protein probability threshold of 0.95 were accepted for further data interpretation. The N-linked glycosylation that did not occur at a consensus N-X-S/T motif (X ≠ P) [48] was also removed to reduce false positive rates of the identified N-glycosylation sites of sialoglycopeptides.

Result and discussion

The complementarity of lectin affinity and TiO_2 chromatography for sialoglycopeptide enrichment

Three LTQ-Orbitrap experiments of two biological replicates (6 LTQ-Orbitrap runs in total) were performed for each enrichment method in this study. As a result, a total of 110 N-glycosylation sites from 73 N-sialoglycoproteins were identified by TiO_2 enrichment, on the other hand, 118 sites from 65 sialoglycoproteins were identified by lectin enrichment, among which, 114 sites from 61 sialoglycoproteins were identified by SNA, and 11 sites from 8 sialoglycoproteins were identified by MAL II. To compare the two

methods, the reproducibility of the biological replicates of the two methods was first and foremost examined. It was found that the reproducibility of TiO_2 chromatography (reached up to 93.33 % in the two replicates) was much better than that of lectin affinity (Fig. S1). To illustrate, the two enrichment methods have different mechanisms for enriching sialoglycopeptides: SNA/MAL II has been used to recover SA-containing glycopeptides and neutral glycopeptides through lectin-glycan interactions [49, 50] which are relatively weak [51], while TiO_2 interacts with highly negatively charged SAs via a multipoint binding, similar to a multidentate binding [31], which is stronger interaction [52]. So, in the process of lectin enrichment, captured sialoglycopeptides were relative easily washed away randomly. And the random events of detachment decreased the reproducibility. After further analysis, it was found that among a total of 168 nonredundant N-glycosylation sites from 100 sialoglycoproteins, 60 sites from 38 sialoglycoproteins were identified by both lectin and TiO_2 enrichment methods, showing a 35.71 % overlap for N-glycosylation sites and a 38 % overlap for sialoglycoproteins (Fig. S2). Although some overlaps were observed, a significant number of N-glycosylation sites and sialoglycoproteins were detected by only one method that 50 and 58 N-glycosylation sites, 35 and 27 sialoglycoproteins were observed dependent on the methods of TiO_2 and lectin, respectively. Thus, it indicates that the two methods are complementary, though TiO_2 theoretically can capture more sialoglycopeptides with little bias.

There are two reasons for the complementarity of these two methods: first, these two methods are based on the different mechanisms in binding the SAs as mentioned above; second, as we know, glycosylation is really complex due to the microheterogeneity of the glyco-chains and the highly dynamic glycoproteome. Thus, for large-scale sialoglycoproteomic studies, single step purification by affinity materials was not effective enough to isolate all sialoglycopeptides, and peptide IPG-IEF as a prefractionation technique could be introduced prior to sialoglycopeptide enrichment using either lectin or TiO_2 chromatography. In this study, TiO_2 was selected to be linked with

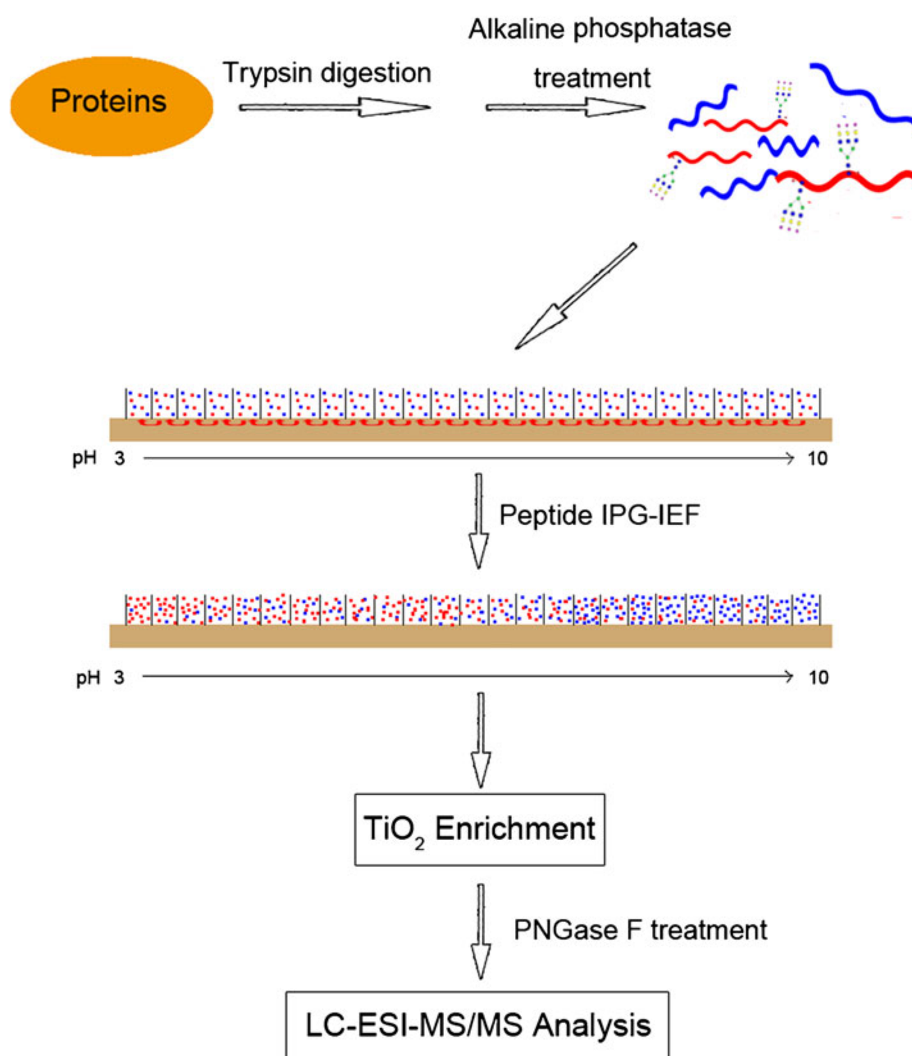

Fig. 1 Workflow of the peptide IPG-IEF assistant TiO_2 enrichment (PIAT) method. The proteins were treated with trypsin and alkaline phosphatase sequentially. Then, the treated peptides were separated into 24 fractions by IPG-IEF. After separation, peptide mixture in each fraction was enriched by TiO_2. Then the enriched sialic-containing peptides were deglycosylated with PNGase F, and analyzed by mass spectrometry

peptide IPG-IEF, due to its better reproducibility and less bias.

A new sialoglycopeptide enrichment strategy based on peptide IPG-IEF assisted TiO_2 chromatography (PIAT)

Peptide IPG-IEF assisted TiO_2 (PIAT) was developed to maximize the profile of sialoglycopeptides further. Figure 1 depicts this new strategy: first, proteins extracted from tissues were digested with trypsin and subsequently treated with alkaline phosphatase. Then the dephosphorylated peptides were separated by peptide IPG-IEF into 24 fractions. Each fraction of peptide solution was subjected to TiO_2 enrichment. Finally, the enriched sialoglycopeptides were treated with PNGase F and then analyzed by LC-ESI-MS/MS. In this study, three biological replicates were performed. As a result, a total of 614 N-glycosylation sites mapping to 582 sialoglycopeptides from 322 sialoglycoproteins were identified. The reproducibility of PIAT in this study was also examined. About 72.1 % of sialoglycopeptides were identified at least in two independent experiments (Fig. S3). More detailed, as for peptide IPG-IEF step, the identified sialoglycopeptides of each fraction in three replicates also showed a big overlap. Thus, these results evidently reveal the good reproducibility and stability of PIAT.

Worth mentioning are the lectin dot blot step and its results. As we referred in the introduction, due to highly negatively charged SAs, sialoglycopeptides could be relatively concentrated in low pH fractions of the IPG strips after IEF steps. To investigate this, the 24 peptide solution fractions from IEF were collected, respectively, and then applied to lectin dot blot analysis. Figure 2 showing the result from lectin dot blot suggested that the sialylation level of proteins in each solution fraction was approximately negatively correlated with the strip pH. Interestingly known from the identification result, the sialoglycopeptides were apparently focused in the top 16 low pH solutions. Figure 3 shows the sialoglycopeptide proportions in different fractions. Among the 582 identified sialoglycopeptides, 480 sialoglycopeptides (82.47 %) were enriched from fractions 1 to 6. Fractions 7 to 12 covered 71 out of the other 102 sialoglyopeptides (69.61 %). Meanwhile, fractions 12 to 16 and 17 to 24 only covered 26 (25.49 %) and 5 (4.90 %) out of the other 102 sialoglycopeptides. That is, 99.1 % of sialoglycopeptides were enriched from fractions 1 to 16, which is consistent with the lectin dot blot analysis results. By PIAT, the 582 sialoglycopeptides were not only identified but also indicated that the pI of peptides can be shifted by the introduced SAs, some of which were greatly different from their calculated pI. For example, the pI of peptide (SIIGFN#TTRPVPK or LRN#ATITQALTNK) is calculated as 11. Theoretically, they should be in the fraction with high pH. However they were both found in fraction 1. Moreover, we found that the identified 26 sialoglycopeptides with multiple N-glycosylation sites were all captured from the top 16 low pH fractions and among them 11 sialoglycopeptides were enriched only from fraction 1, which had the lowest pH. Thus, our analysis and hypothesis of the efficiency of Peptide IPG-IEF for sialoglycopeptides has been validated. And the Peptide IPG-IEF has been proved to be not only a good separation method but also a useful enrichment method for sialoglycopeptides.

Improved N-sialoglycopeptides coverage by PIAT

Combining the results from PIAT, lectin and TiO_2 methods, we totally identified 623 N-glycosylation sites in 595 sialoglycopeptides within the 328 sialoglycoproteins in this study. As shown in Table 1 almost all sialoglycopeptides identified by either lectin or TiO_2 were captured by PIAT. Compared with TiO_2 by which only 49.56 % of sialoglycopeptides captured by lectin were identified, PIAT covers 92.92 % of sialoglycopeptides detected by lectin (Fig. S4). These results show a significantly improved ability of identification of sialoglycopeptides with PIAT compared with TiO_2 and lectin methods. By the way, we appreciate a recent research by Palmisano *et al* which proved that a common pitfall of chemical deamidation exists in large-scale identification of N-glycosylation sites [53]. To reduce false positive, all the 113 deamidated Asn (N+0.98 Da) matches that did not occur in a consensus N-X-S/T (X ≠ P) motif were not counted in the final result.

We next investigated the isoelectric point (pI), molecular mass (MW) and subcellular distribution of the

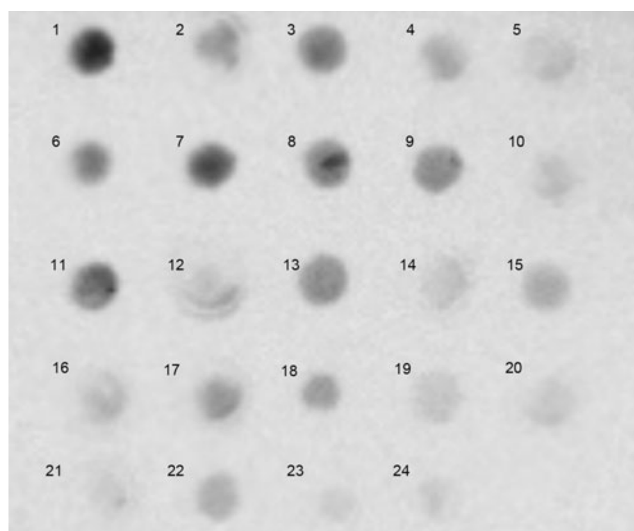

Fig. 2 Lectin dot blot analysis of sialoglycopeptides from each fraction of peptide IPG-IEF. The peptides in each fraction separated by IPG-IEF were collected and subjected to lectin dot blot analysis. The grayscale indicates the sialylation level of each fraction and the number 1 to 24 indicates the 24 fractions from pH 3-10

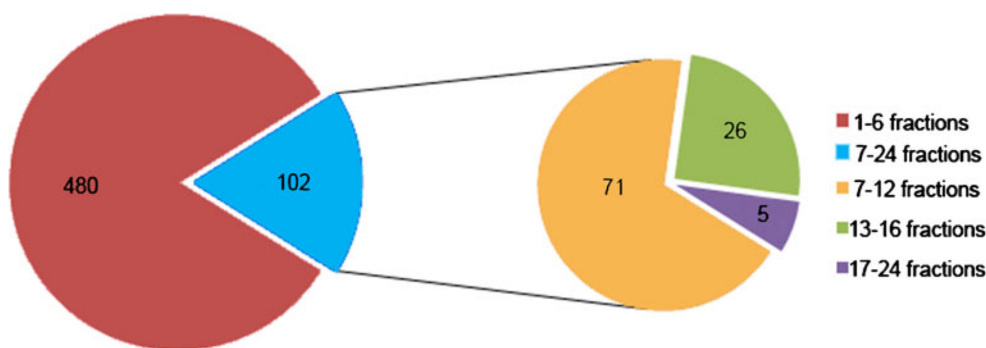

Fig. 3 The proportions of sialoglycopeptides in different fractions resulted by IPG-IEF

sialoglycopeptides/sialoglycoproteins to examine the improved universality by PIAT. As a result, PIAT exhibits better enrichment capacity of sialoglycopeptides with different pI than lectin and TiO_2 methods (Fig. S5a), even including very basic peptides (up to pI 12). The 582 sialoglycopeptides enriched by PIAT showed a wider MW distribution than lectin and TiO_2 (Fig. S6a). Among them, 23 sialoglycopeptides were identified with MW over 3 kDa. While only 5 sialoglycopeptides with MW over than 3 kDa were identified by lectin and none by the TiO_2 method. The sialoglycoproteins identified by PIAT also display a wider pI and MW distribution compared with lectin and TiO_2 methods (Fig. S5b, Fig. S6b). The biggest MW of sialoglycoprotein from PIAT is 536 kDa. The subcellular localization of the identified sialoglycopeptides was further investigated. As shown in Table 2, sialoglycoproteins identified by PIAT performs more diverse subcellular distribution than that from lectin and TiO_2. For example, sialoglycoprotein located in nucleus was detected by neither lectin nor TiO_2, while 8 nuclear sialoglycoproteins were detected by PIAT. Moreover, 38.71 % of sialoglycoproteins were identified as membrane proteins which are thought being of higher sialylation level compared with other localizations of proteins. These results show the high efficiency and universality of PIAT and underscore the value of PIAT in capturing sialoglycopeptides.

Interestingly, we found that sialoglycopeptides with multiple N-glycosylation sites were enriched to a higher extent with PIAT compared with lectin and TiO_2 methods. A total of 26 sialoglycopeptides were identified to be with multiple N-glycosylation sites in this study. All of them can be captured by PIAT. In addition, some sialoglycopeptides were identified to be with one N-glycosylation site by lectin or TiO_2 methods, but were detected to be with multiple N-glycosylation sites by PIAT. For example, a singly N-glycosylated sialoglycopeptide (SVGTGTNMVFQNCSCIQSSGN#SSAVLGLCNK) from solute carrier organic anion transporter family member 1A4 was identified by SNA to be with one N-glycosylation site of N-491. However, the same peptide sequence (SVGTGTNMVFQN#CSCIQSSGN#SSAVLGLCNK) was identified to be with two N-glycosylation sites of N-482 and N-491 by PIAT. It indicates that PIAT is efficient to enrich not only singly N-glycosylated sialoglycopeptides but also multiply N-glycosylated sialoglycopeptides. Thus, all these examined aspects evidently reveal the high efficiency and universality of PIAT for exploiting sialoglycosylation at proteome-wide scale.

Table 1 The number of N-glycosylation sites, sialoglycopeptides, sialoglycoproteins identified by each method

Method	Sites[a]	Peptides[b]	Proteins[c]
SNA	114	109	61
MAL II	11	8	8
TiO_2	110	105	73
PIAT	614	582	322
Total	623	595	328

a: N-glycosylation sites b: sialoglycopeptides c: sialoglycoproteins

Analysis of N-sialoglycosylation in rat liver

In this study, 623 N-glycosylation sites were identified in 595 sialoglycopeptides within the 328 sialoglycoproteins. Among them, only one protein Gamma-glutamyltranspeptidase 1 (P07154) was recorded of containing SA in the Swiss-Prot database (February 2012 Version 46). However, the exact glycosylation site of this protein was not provided. Moreover, among the identified 623 N-glycosylation sites, only 12 sites

Table 2 The number of the identified sialoglycoproteins with different subcellular location

Method Location[a]	SNA	MAL2	TiO_2	PIAT	Total
Membrane	30	7	24	120	125
Extracellular	21	0	19	64	66
Cytosol	7	0	16	55	57
ER/Golgi	2	0	5	54	54
Mitochondria	0	1	8	20	23
Nucleus	0	0	0	8	8
Unknown	0	0	1	6	6

[a] The subcellular localization was determined by the annotations from UniProtKB

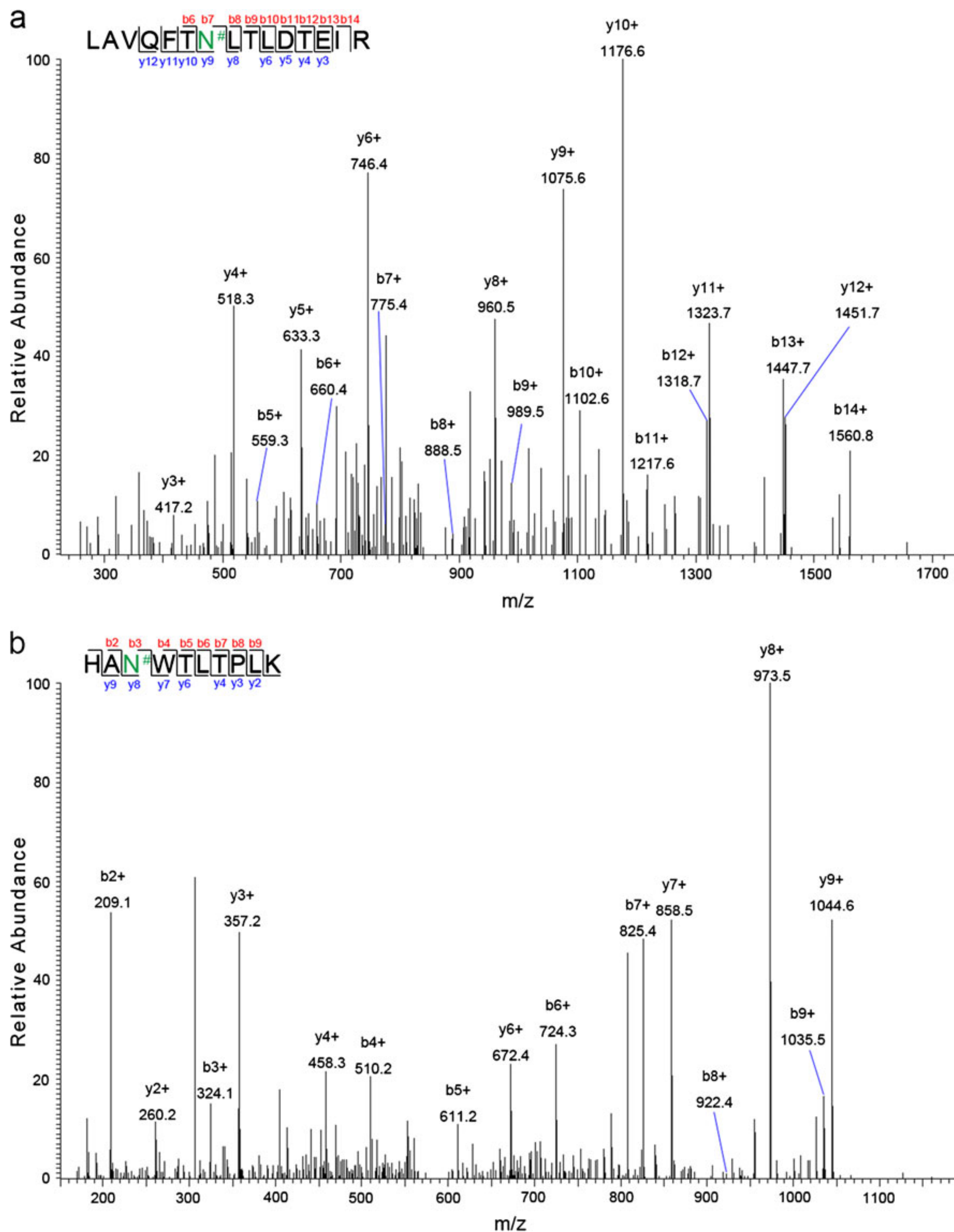

Fig. 4 Nano-LC-MS/MS mass spectra of doubly charged N-sialoglycopeptide LAVQFTN#LTLDTEIR [(M+2 H)$^{2+}$ at m/z 867.97] from potassium-transporting ATPase subunit beta-1 (a), and HAN#WTLTPLK [(M+2 H)$^{2+}$ at m/z 591.32] from Serum paraoxonase/arylesterase 1 (b)

were presently documented in the Swiss-Prot database as N-glycosylation sites based on the experimental findings. Meanwhile 498 sites were annotated as potential (476 sites), probable (8 sites), or by similarity (14 sites), which are three types of qualifiers indicating the nonexperimental evidence. The other 113 sites (18.14 %) were totally unknown in the Swiss-Prot database. These data suggest that the N-linked sialylation in rat is largely unknown and should be investigated more. To our knowledge, this study not only shows a large-scale assignment of N-glycosylation sites of sialoglycoproteins in rat liver

for the first time but also provides experimental evidence for numerous potential N-glycosylation sites as well as a number of new N-glycosylation sites for the Swiss-Prot database.

The SAs mainly attach to the terminal glycans through α(-2-3) and α(-6) glycosidic linkages. Moreover, different branching patterns of SAs were reported to be associated with pathological process. In this study, 114 N-glycosylation sites were determined as having α(2-6) linked SAs by SNA, whereas only 11 N-glycosylation sites were identified to have α(2-3) linked SAs by MALII. This result indicates that the level of sialylation with α(2-3) glycosidic linkage is much lower than that of sialylation with α(2-6) glycosidic linkage in rat liver, which is consistent with the lectin blot result (Fig. S7). Recently, 11 N-glycosylation sites with core-fucosylation were assigned by Endo H in rat liver by our lab [54]. Among them, the residue 197 of the peptide (N#SIFLADINQER) from palmitoyl-protein thioesterase 1 was identified as an N-glycosylation site with core-fucose, which was also identified as an N-glycosylation site with sialic acid by PIAT in this study. Figure 4a shows the MS/MS spectra of a deglycosylated, formerly SA-containing peptide [$(M+2\ H)^{2+}$ at m/z 867.97] from potassium-transporting ATPase subunit beta-1 as an example. The location of the N-glycosylation site was determined by a mass increase of 0.98 Da after PNGase F treatment and MS identification. The b- and y-series display a mass shift from the conversion of N to D on the N-X-S/T ($X \neq P$) motif. Figure 4b shows another example of a deglycosylated, former sialoglycopeptide [$(M+2\ H)^{2+}$ at m/z 591.32] from Serum paraoxonase/arylesterase 1. The N-glycosylation site on this peptide was identified on the N-253. All the sialoglycopeptides and their N-glycosylation sites identified in this study are shown in supplemental Table S1.

The Uniprot-designated subcellular distribution of the total identified sialoglycoproteins was also depicted in Table 2. Most of the sialoglycoproteins are located in the membrane (38.11 %). 20.12 % of sialoglycoproteins are classified as extracellular proteins, which is consistent with a previous report that SAs typically attach to the terminal glycans of the cell surface and secreted glycoproteins [1]. 17.38 % and 16.46 % of the identified sialoglycoproteins are classified as being present in the cytoplasm and ER/Golgi. Only a small number of sialoglycoproteins are located in the mitochondria and nucleus (7.01 % and 2.44 %, respectively.). The localization of 6 sialoglycoproteins is still unknown.

Conclusion

In this study, by a combination of three strategies, lectin, TiO_2 and PIAT methods, a total of 623 N-glycosylation sites were identified in 595 sialoglycopeptides within the 328 sialoglycoproteins, which provided an in-depth identification of N-glycosylation sites of sialoglycopeptides in rat liver for the first time. The majority of the identification results were contributed by PIAT, *i.e.*, 614 N-glycosylation sites mapping to 582 sialoglycopeptides from 322 sialoglycoproteins were identified by PIAT. In PIAT method, sialoglycopeptides were separated by peptide IPG-IEF followed by TiO_2 enrichment and LC-MS/MS analysis. The sialoglycopeptides concentrated by PIAT exhibit better enrichment capacity of sialoglycopeptides with wide molecular weight and different pI values compared with lectin and TiO_2 methods. The sialoglycoproteins identified by this method also performs diverse subcellular localization. PIAT displays high ability and universality for sialoglycopeptides enrichment and is a powerful tool for sialylation exploration.

Acknowledgments This work was supported in part by National 973/S973/863 and NSF projects (S973-2011CB910600, NFS-31100590 and 20975024, S973-2010CB912700, 863-2012AA020200), and Shanghai Municipal Natural Science Foundation (11ZR1403000).

References

1. Varki, A.: Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. Nature **446**, 1023–1029 (2007)
2. Angata, T., Varki, A.: Chemical diversity in the sialic acids and related α-keto acids: an evolutionary perspective. Chem. Rev. **102**, 439–470 (2002)
3. Gunnar, B.: The carbohydrate groups of the submaxillary mucin. Z. Physiol. Chem. **240**, 43 (1936)
4. Rempel, H., Calosing, C., Sun, B., Pulliam, L.: Sialoadhesin expressed on IFN-induced monocytes binds HIV-1 and enhances infectivity. PLoS One **3**, e1967 (2008)
5. Haselhorst, T., Fleming, F.E., Dyason, J.C., Hartnell, R.D., Yu, X., Holloway, G., Santegoets, K., Kiefel, M.J., Blanchard, H., Coulson, B.S., Itzstein, M.V.: Sialic acid dependence in rotavirus host cell invasion. Nat. Chem. Biol. **5**, 91–93 (2008)
6. Varki, A.: Nothing in glycobiology makes sense, except in the light of evolution. Cell **126**, 841–845 (2006)
7. Bishop, J., Gagneux, P.: Evolution of carbohydrate antigens — microbial forces shaping host glycomes? Glycobiology **17**, 23R–34R (2007)
8. Varki, A., Angata, T.: Siglecs — the major subfamily of I-type lectins. Glycobiology **16**, 1R–27R (2006)
9. Shinya, K., Ebina, M., Yamada, S., Ono, M., Noriyuki, K., Kawaoka, Y.: Avian flu: Influenza virus receptors in the human airway. Nature **440**, 435–436 (2006)
10. Hidari, K.I., Shimada, S., Suzuki, Y., Suzuki, T.: Binding kinetics of influenza viruses to sialic acid-containing carbohydrates. Glycoconj. J. **24**, 583–590 (2007)
11. Hidari, K.I.P.J., Suzuki, T.: Glycan receptor for influenza virus. The Open Antimicrobial Agents Journal **2**, 26–33 (2010)
12. Shetty, V., Nickens, Z., Shah, P., Sinnathamby, G., Semmes, O.J., Philip, R.: Investigation of sialylation aberration in N-linked glycopeptides by Lectin and Tandem Labeling (LTL) quantitative proteomics. Anal. Chem. **82**, 9201–9210 (2010)
13. Scanlin, T.F., Glick, M.C.: Terminal glycosylation and disease: influence on cancer and cystic fibrosis. Glycoconj. J. **17**, 617–626 (2000)

14. Gorelik, E., Galili, U., Raz, A.: On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. Cancer Metastasis Rev. **20**, 245–277 (2001)
15. Varki, N.M., Varki, A.: Heparin inhibition of selectin-mediated interactions during the hematogenous phase of carcinoma metastasis: rationale for clinical studies in humans. Semin. Thromb. Hemost. **28**, 53–66 (2002)
16. Borsig, L., Wong, R., Hynes, R.O., Varki, N.M., Varki, A.: Synergistic effects of L- and P-selectin in facilitating tumor metastasis can involve non-mucin ligands and implicate leukocytes as enhancers of metastasis. Proc. Natl. Acad. Sci. U. S. A. **99**, 2193–2198 (2002)
17. Laubli, H., Stevenson, J.L., Varki, A., Varki, N.M., Borsig, L.: L-selectin facilitation of metastasis involves temporal induction of Fut7-dependent ligands at sites of tumor cell arrest. Cancer Res. **66**, 1536–1542 (2006)
18. Ying-Chih, L., Hsin-Yung, Y., Chien-Yu, C., Chein-Hung, C., Ping-Fu, C., Yi-Hsiu, J., Chung-Hsuan, C., Kay-Hooi, K., Chong-Jen, Y., Pan-Chyr, Y., Tsui-Ling, H., Chi-Huey, W.: Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. PNAS **108**, 11332–11337 (2011)
19. Tajiri, M., Ohyama, C., Wada, Y.: Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: a glycopeptide approach. Glycobiol. **18**, 2–8 (2008)
20. Bond, M.R., Kohler, J.J.: Chemical methods for glycoprotein discovery. Curr. Opin. Chem. Biol. **11**, 52–58 (2007)
21. Laughlin, S.T., Baskin, J.M., Amacher, S.L., Bertozzi, C.R.: In vivo imaging of membrane- associated glycans in developing zebrafish. Science **320**, 664–667 (2008)
22. Yang, L., Nyalwidhe, J.O., Guo, S., Drake, R.R., Semmes, O.J.: Targeted identification of metastasis-associated cell-surface sialoglycoproteins in prostate cancer. Mol. Cell. Proteomics (2011). doi:10.1074/mcp.M110.007294
23. Nilsson, J., Rüetschi, U., Halim, A., Hesse, C., Carlsohn, E., Brinkmalm, G., Larson, G.: Enrichment of glycopeptides for glycan structure and attachment site identification. Nat. Methods **6**, 809–811 (2009)
24. Zeng, Y., Ramya, T.N., Dirksen, A., Dawson, P.E., Paulson, J.C.: High-efficiency labeling of sialylated glycoproteins on living cells. Nat. Methods **6**, 207–209 (2009)
25. Tian, Y., Esteva, F.J., Song, J., Zhang, H.: Altered expression of sialylated glycoproteins in breast cancer using hydrazide chemistry and mass spectrometry. Mol. Cell. Proteomics (2012). doi:10.1074/mcp.M111.011403
26. Drake, P.M., Schilling, B., Niles, R.K., Braten, M., Johansen, E., Liu, H., Lerch, M., Sorensen, D.J., Li, B., Allen, S., Hall, S.C., Witkowska, H.E., Regnier, F.E., Gibson, B.W., Fisher, S.J.: A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. Anal. Biochem. **408**, 71–85 (2011)
27. Zhao, J., Simeone, D.M., Heidt, D., Anderson, M.A., Lubman, D.M.: Comparative serum glycoproteomics using lectin selected sialicacid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. J. Proteome Res. **5**, 1792–1802 (2006)
28. Qiu, R., Regnier, F.E.: Comparative glycoproteomics of N-linked complex-type glycoforms containing sialic acid in human serum. Anal. Chem. **77**, 7225–7231 (2005)
29. Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M.: Universal sample preparation method for proteome analysis. Nat. Methods **6**, 359–362 (2009)
30. Zielinska, D.F., Gnad, F., Wiśniewski, J.R., Mann, M.: Precision mapping of an in vivo N-Glycoproteome reveals rigid topological and sequence constraints. Cell **141**, 897–907 (2010)
31. Larsen, M.R., Jensen, S.S., Jakobsen, L.A., Heegaard, N.H.: Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. Mol. Cell. Proteomics **6**, 1778–1787 (2007)
32. Palmisano, G., Lendal, S.E., Larsen, M.R.: Titanium dioxide enrichment of sialic acid-containing glycopeptides. Methods Mol. Biol. **753**, 309–322 (2011)
33. Zhang, B., Sheng, Q., Li, X., Liang, Q., Yan, J., Liang, X.: Selective enrichment of glycopeptides for mass spectrometry analysis using C18 fractionation and titanium dioxide chromatography. J. Sep. Sci. **34**, 2745–2750 (2011)
34. Lewandrowski, U., Zahedi, R.P., Moebius, J., Walter, U., Sickmann, A.: Enhanced N-glycosylation site analysis of sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. Mol. Cell. Proteomics **6**, 1933–1941 (2007)
35. Palmisano, G., Lendal, S.E., Engholm-Keller, K., Leth-Larsen, R., Parker, B.L., Larsen, M.R.: Selective enrichment of sialic acid–containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. Nat. Protoc. **5**, 1974–1982 (2010)
36. Hoerth, P., Miller, C.A., Preckel, T., Wenz, C.: Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. Mol. Cell. Proteomics **5**, 1968–1974 (2006)
37. Picotti, P., Aebersold, R., Domon, B.: The implications of proteolytic background for shotgun proteomics. Mol. Cell. Proteomics **6**, 1589–1598 (2007)
38. Cox, J., Mann, M.: MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. **26**, 1367–1372 (2008)
39. Godoy, L.M., Olsen, J.V., Cox, J., Nielsen, M.L., Hubner, N.C., Frohlich, F., Walther, T.C., Mann, M.: Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature **455**, 1251–1254 (2008)
40. Chick, J.M., Haynes, P.A., Molloy, M.P., Bjellqvist, B., Baker, M.S., Len, A.C.: Characterization of the rat liver membrane proteome using peptide immobilized pH gradient isoelectric focusing. J. Proteome Res. **7**, 1036–1045 (2008)
41. Pan, C., Kumar, C., Bohl, S., Klingmueller, U., Mann, M.: Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific. Mol. Cell. Proteomics **8**, 443–450 (2009)
42. Mulvenna, J., Hamilton, B., Nagaraj, S.H., Smyth, D., Loukas, A., Gorman, J.J.: Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the Hookworm, Ancylostoma caninum. Mol. Cell. Proteomics **8**, 109–121 (2009)
43. Lengqvist, J., Eriksson, H., Gry, M., Uhlén, K., Björklund, C., Bjellqvist, B., Jakobsson, P.J., Lehtiö, J.: Observed peptide pI and retention time shifts as a result of post-translational modifications in multidimensional separations using narrow-range IPG-IEF. Amino Acids **40**, 697–711 (2011)
44. Gauci, S., Breukelen, B.V., Lemeer, S.M., Krijgsveld, J., Heck, A.J.: A versatile peptide pI calculator for phosphorylated and N-terminal acetylated peptides experimentally tested using peptide isoelectric focusing. Proteomics **8**, 4898–4906 (2008)
45. Cao, J., Hu, Y., Shen, C.P., Yao, J., Wei, L.M., Yang, F.Y., Nie, A.Y., Wang, H., Shen, H., Liu, Y.K., Zhang, Y., Tang, Y., Yang, P.Y.: Zeolite LTL nanocrystal-driving high efficient enrichment of secretory proteins in human hepatocellular carcinoma cells. Proteomics **9**, 4881–4888 (2009)
46. Cao, J., Shen, C.P., Wang, H., Shen, H.L., Chen, Y.H., Nie, A.Y., Lu, H.J., Liu, Y.K., Yang, P.Y.: Identification of N-glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach. J. Proteome Res. **8**, 662–672 (2009)
47. Elias, J.E., Gygi, S.P.: Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods **4**, 207–214 (2007)
48. Bause, E.: Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. Biochem. J. **209**, 331–336 (1983)

49. Cummings, R.D., Kornfeld, S.: Fractionation of asparagine-linked oligosaccharides by serial lectin-Agarose affinity chromatography. A rapid, sensitive, and specific technique. J. Biol. Chem. **257**, 11235–11240 (1982)
50. Yang, Z., Hancock, W.S.: Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. J. Chromatogr. A. **1070**, 57–64 (2005)
51. Uchiyama, N., Kuno, A., Koseki-Kuno, S., Ebe, Y., Horio, K., Yamada, M., Hirabayashi, J.: Development of a lectin microarray based on an evanescent-field fluorescence principle. Methods Enzymol. **415**, 341–351 (2006)
52. Engholm-Keller, K., Larsen, M.R.: Titanium dioxide as chemo-affinity chromatographic sorbent of biomolecular compounds–applications in acidic modification-specific proteomics. J. Proteomics **75**, 317–328 (2011)
53. Palmisano, G., Melo-Braga, M.N., Engholm-Keller, K., Parker, B.L., Larsen, M.R.: Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses. J. Proteome Res. **11**, 1949–1957 (2012)
54. Zhang, W., Wang, H., Zhang, L., Yao, J., Yang, P.Y.: Large-scale assignment of N-glycosylation sites using complementary enzymatic deglycosylation. Talanta **85**, 499–505 (2011)