

CHEMISTRY

Advancements in mass spectrometry-based glycoproteomics and glycomics

Haojie Lu^{1,2,3}, Ying Zhang^{1,2} and Pengyuan Yang^{1,2,3,*}

ABSTRACT

Protein *N*-glycosylation plays a crucial role in a considerable number of important biological processes. Research studies on glycoproteomes and glycomes have already characterized many glycoproteins and glycans associated with cell development, life cycle, and disease progression. Mass spectrometry (MS) is the most powerful tool for identifying biomolecules including glycoproteins and glycans, however, utilizing MS-based approaches to identify glycoproteomes and glycomes is challenging due to the technical difficulties associated with glycosylation analysis. In this review, we summarize the most recent developments in MS-based glycoproteomics and glycomics, including a discussion on the development of analytical methodologies and strategies used to explore the glycoproteome and glycome, as well as noteworthy biological discoveries made in glycoproteome and glycome research. This review places special emphasis on China, where scientists have made sizeable contributions to the literature, as advancements in glycoproteomics and glycomics are occurring quite rapidly.

Keywords: mass spectrometry, glycoproteomics, glycomics

INTRODUCTION

Glycosylation is one of the most important post-translational modifications (PTMs) of proteins. It is estimated that more than 50% of all proteins in mammals can be glycosylated, and, most importantly, the majority of the FDA-approved tumor biomarkers currently available are glycoproteins. Therefore, intensive efforts have been devoted to the development of glycoproteome and glycome identification techniques.

Glycoproteomic analysis uses proteomics strategies, in general, to identify proteins and their glycosylation sites. Glycans detached before analysis are not considered in glycoproteomics. Conversely, glycomic study focuses on analyzing the whole glycan profile from samples, without considering the glycosylation sites. In past years, glycoproteomics and glycomics were closely related research areas; now, they are relatively separate disciplines. Currently, alongside the emergence of mass spectrometry (MS) spectrometric techniques for the analysis of intact glycopeptides, glycoproteomics practices typically identify glycosylation sites and glycan changes on a specific site as well—glycomics

practices are increasingly developed to identify glycans on a specific site.

Though there has been considerable (and rapid) progress made in these two overlapping areas of study in recent years, glycoproteome and glycome researchers still face significant challenges. For example, although glycosylation is prevalent in the human body, the amount of glycopeptides or glycans is relatively low. As opposed to the protein cycle, glycosylation is a non-template-driven process, which generates enormous diversity in their structures (which differ in composition, branching, and linkage.) The attached glycan further complicates the glycoproteome, as each glycosylation site on a certain glycoprotein contains various glycans. Additionally, the ionization efficiency of glycopeptide or glycan is much lower than that of non-modified peptides. Therefore, the detection of glycopeptide and glycan is much more difficult than the analysis of ordinary peptides.

Further, analysis of glycosylation sites and glycan composition requires specialized bioinformatics software to interpret complex MS spectra. For this reason, research on the glycome is far more challenging than on the proteome. In effort to achieve

¹Department of Systems Biology for Medicine, School of Basic Medicine and Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China; ²Key Lab of Glycoconjugate of Ministry of Health and Birth Control, Fudan University, Shanghai 200032, China and ³Department of Chemistry, Fudan University, Shanghai 200433, China

*Corresponding author. E-mail: pyyang@fudan.edu.cn

Received 15 September 2015;

Revised 3 December 2015; Accepted 11 December 2015

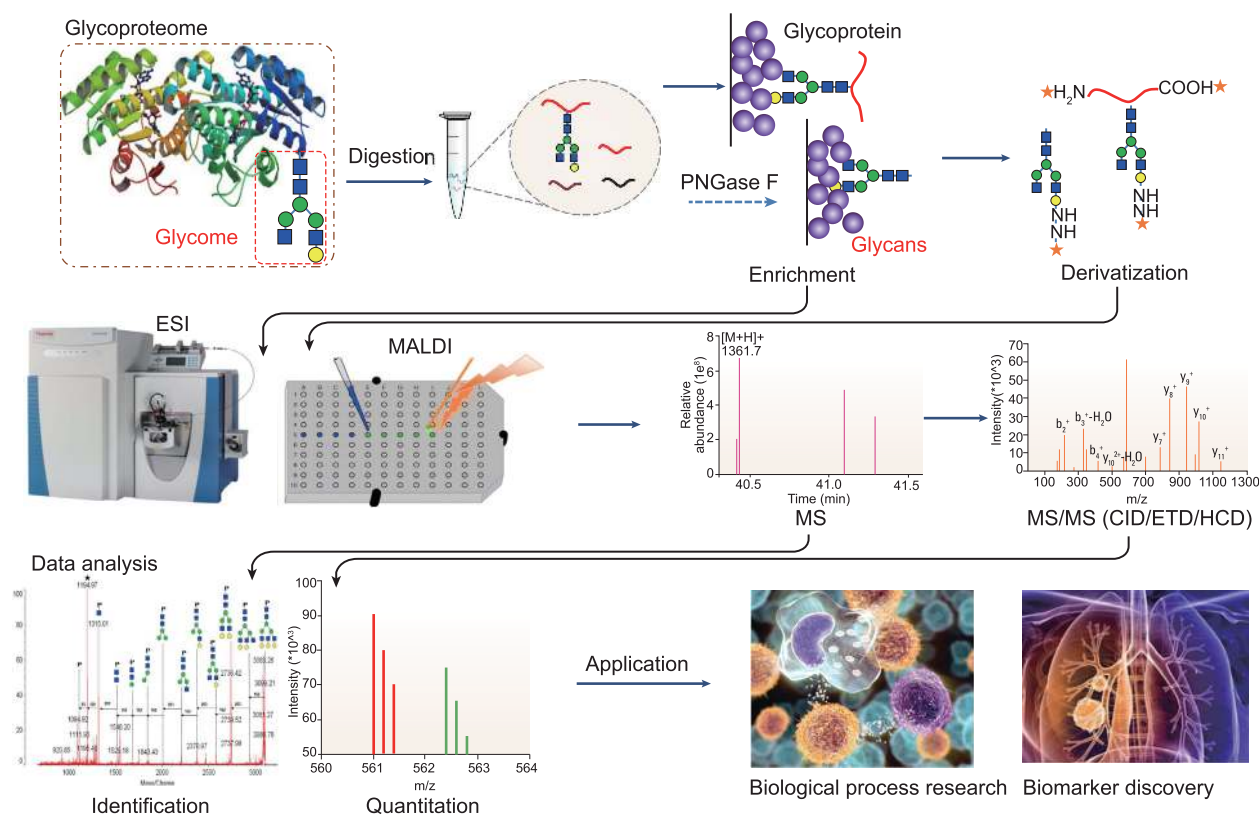


Figure 1. General strategy of MS-based glycoproteomics and glycomics.

the specific and sensitive analysis of glycoproteomes and glycomes, researchers have achieved tremendous developments in bioanalytical chemistry that have improved glycoprotein characterization, glycosylation site identification, understanding of the composition and linkage of glycans, and the quantitation of glycoproteins and glycans. These technical advances and their applications in biological process research and biomarker discovery are reviewed in this study. Fig. 1 illustrates general background information relevant to MC-based glycoproteomics and glycomics.

SAMPLE PREPARATION AND ENRICHMENT

Sample preparation is a prerequisite of successful analysis of glycoproteomes or glycomes. In general, due to low abundance of glycoproteins, the glycoproteome or glycome must be pre-separated prior to MS analysis.

Glycoproteome enrichment

The glycoproteome can be selectively enriched through lectin enrichment [1], hydrophilic

interaction chromatography HILIC [2], boronic chemistry [3], hydrazide chemistry (HC) [4], reductive amination chemistry [5], oxime click chemistry [6], or non-reductive amination chemistry [7]. These separation techniques were reviewed comprehensively in our recently published paper [8].

Glycopeptidome enrichment

The glycopeptidome, as opposed to the aforementioned glycoproteome, refers to endogenous glycopeptides. The glycopeptidome is an important subset of the peptidome, but has been researched less frequently. Peptidomes, the endogenous peptides in biological samples, have garnered increasing attention since 2001. Due to the low abundance and low molecular weight of the peptidome, peptidome analysis requires special sample preparation and enrichment techniques. These low molecular weight peptides can be obtained using size-exclusion methods that effectively deplete the high-molecular-weight proteins. Glycopeptidomes have been relatively infrequently researched since about 2007, while the phosphopeptidome, the other subset of the peptidome, has been studied rather

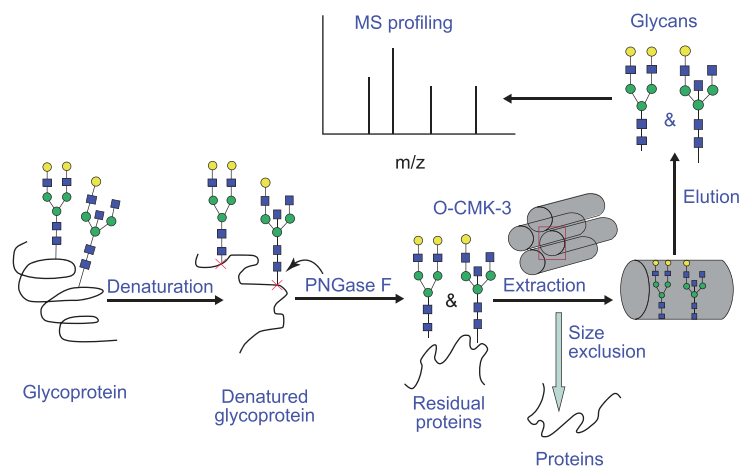


Figure 2. Scheme for size-selective enrichment of glycans by O-CMK-3 [9].

intensively. We profiled the rat serum glycopeptidome in a study conducted in our own laboratory by first synthesizing a boronic acid-functionalized, highly ordered mesoporous nanomaterial (MCM-41-APTES-CPB) with glycopeptide-suitable pore size and glycopeptide-specific selectivity using a post-synthesis grafting method [3]. The combined advantages of size-exclusive effect of mesopores against protein, and glyco-selective effect of boronic acid toward glycopeptide, allowed a successful profiling of the glycopeptidome of rat serum for the first time.

Glycome enrichment

Generally, glycomes are obtained by detaching the glycan from the glycoprotein by De-*N*-glycosylation using peptide *N*-glycosidase F (PNGase F). Upon liberation, the glycome is further purified from the protein mixture through size exclusion, HILIC chromatography, graphitized carbon chromatography, or tagging-assisted strategies. Graphitized carbon is the most commonly used stationary phase for enriching glycans, because glycans are typically more hydrophilic than peptides [9]. In addition to the commercially available PGC columns and tips, new solid phases like cotton wool have been developed for the preparation of glycans from protein mixtures. Cotton wool SPE microtips, for example, allow the removal of salts, most non-glycosylated peptides, and detergents, such as SDS from glycoconjugate samples [10].

The molecular weights of glycans are small compared to those of proteins. Therefore, glycan can be easily separated from a protein mixture through size

exclusion. For example, Zou *et al.* developed a simple method to enrich the N-linked glycan using oxidized ordered mesoporous carbon. By taking advantage of the size-exclusive effect of mesopores against proteins, as well as the interaction between glycan and carbon, N-linked glycans from serum samples were effectively purified (Fig. 2) [11].

The same group of researchers also used silica-carbon composite nanoparticles with uniform shapes and highly ordered mesoporous structures to enrich N-linked glycan from complex biological samples [12]. Recently, Deng *et al.* designed a carbon-functionalized, ordered graphene/mesoporous silica composite (C-graphene@mSiO₂) for glycome enrichment, where the newly synthesized material possessed a uniform pore size. By taking advantage of a special interaction between carbon and glycan, as well as size-exclusion ability, glycans were efficiently separated from proteins. After enrichment with the C-graphene@mSiO₂ composites, 48 N-linked glycans with sufficient peak intensities (signal to noise, S/N > 10) were successfully identified by MS from 400 nL human serum [13]. The same group further developed a magnetic nanoporous carbon (NPC) material for glycome enrichment [14].

Introducing an enrichment tag on glycan enables the glycan to be specifically removed from protein mixtures due to the strong interaction between the tag and tailor-made solid phases. Very recently, we reported a new strategy for the sensitive and selective MS analysis of glycans based on fluororous chemistry. With this strategy, the reducing end of the glycan is derivatized with a hydrophobic fluorinated carbon tag, and then the fluororous solid-phase extraction (FSPE) is used to specifically enrich the glycan-containing fluorinated carbon tag. Results showed that the *N*-glycome of human serum was efficiently enriched from highly contaminated solutions and protein mixtures [15].

Another type of enrichment method with high specificity relies on the formation of reversible bonds between glycans and solid phases. For example, Qian *et al.* used 1-pyrenebutyryl chloride-functionalized, free graphene oxide (PCGO) to capture glycan through the reversible covalent bond formation between the hydroxyl groups of glycan and the acyl chloride groups on graphene oxide (GO). The captured glycans were then released through reversible covalent bond dissociation for MS analysis. When the glycans were captured, the free PCGO sheets aggregated within 30 s; therefore, this method allows simple and efficient visual monitoring of the enrichment process [16].

DERIVATIZATION STRATEGIES FOR GLYCOMES AND GLYCOPROTEOMES

Because the analysis of native glycans or glycopeptides by MS is often hindered by their poor ionization efficiency, various chemical derivatization strategies with different labels have been developed. These derivatization techniques can improve the ionization efficiency of glycans and glycolpeptides in MS by increasing their hydrophobicity; derivatization can also influence their tandem MS fragmentation patterns, thereby facilitating successful structural analysis. The following provides an overview of the techniques available for glycomic and glycoproteomic derivatization, including reductive amination, hydrazide labeling, permethylation, and amidation.

Derivatization strategies for glycoproteomes

Glycopeptide analysis is a challenging endeavor due to the poor ionization efficiency of glycopeptides and interference from other highly abundant peptides on the MS signal. Derivation can be used to address this problem. Junko *et al.*, for example, proposed 1-pyrenyldiazomethane as a unique derivatization reagent for enhanced and selective detection of glycopeptide, which is derivatized at the C-terminus of the peptide without releasing glycan. Notably, underivatized forms of glycopeptide on the spectra were observed due to the dissociation of pyrene groups from glycopeptide, while ionization of other peptides was significantly reduced [17]. The same group of researchers also developed a method for structural analysis of N-linked glycopeptide, where carboxyl groups on the peptide moiety (i.e. Asp, Glu, and C-terminus) are completely derivatized via methylamidation, and then glycopeptide-produced informative glycan fragment ions reflect detailed glycan structural features, as well as the sialylated glycopeptide [18]. Very recently, Wuhrer *et al.* proposed a novel technique using dimethylamidation and lactonization for matrix-assisted laser desorption/ionization (MALDI)-TOF-MS profiling of IgG glycopeptide through which both glycan and peptide moieties are stabilized, allowing fragmentation spectra information to be obtained for profiling glycosylation and peptide sequences [19].

Derivatization strategies for glycomes

There are more derivatization strategies for glycome than for glycoproteome. Based on the chemical structures of glycan, labeling sites can fall into different categories: the reducing terminal, the hy-

droxyl groups on the glycan, and the carboxyl group of sialylated glycan.

Reducing terminal derivatization

To date, the reductive amination reaction labeled at the terminal aldehyde-group of glycan is the most often utilized derivatization approach. In this type of reaction, the aldehyde group of glycan reacts with a small molecule containing a primary amine group, resulting in an imine or Schiff base after a condensation reaction, and then reduced by a reducing agent to obtain a secondary amine. Its high yields and the ready availability of amine reagents make this technique attractive for glycomic derivatization, and as such it is commonly employed in the field; 2-aminobenzoic acid (2-AA), 2-aminobenzamide (2-AB), and 2-aminopyridine (2-AP) are the labels most often applied to improve the ionization response of oligosaccharides. Further, experimental details for the application of these reductive amination derivatives were provided in an earlier review [20].

The disadvantage of reductive amination reaction-based derivatization is that it requires removing of excess reducing reagents (for example, sodium cyanoborohydride [NaBH₃CN]), which involves tedious cleanup steps that may cause sample loss. Therefore, non-reductive amination reactions, which no longer require the reduction of oligosaccharides with NaBH₃CN, are a more common research object. In fact, we investigated a new derivatization reagent, aminopyrazine, which can also act as a co-matrix to improve the detection of oligosaccharides by MALDI-TOF MS. This method eliminates the purification step. We found that the S/N ratio of glycan increased about 2–6-fold compared to the control, with good signal reproducibility [21]. Furthermore, reductive amination can be used for introducing new tags in glycans to enhance their ionization efficiency. As mentioned above, we proposed a technique for the sensitive and selective MS analysis of glycan based on fluororous chemistry (Fig. 3), through which glycan-reducing terminals are derivatized with a hydrophobic fluorinated carbon tag based on the reductive amination reaction. Results indicated that the fluororous tag significantly increased glycan hydrophobicity and glycan ionization efficiency during MS, by more than one order of magnitude [15].

Hydrazide labeling is another derivatization strategy based on a reaction of glycan ends with labels containing a hydrazide end group. Because hydrazone formation does not require reductive conditions, its advantage is the corresponding

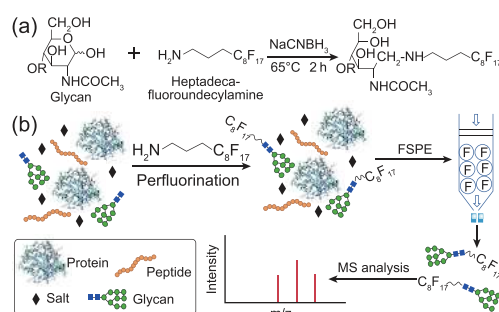


Figure 3. (a) Illustration of fluorous derivatization of glycans with heptadecafluoroundecylamine through reductive amination reaction; (b) schematic diagram of fluorous amine-based glycan derivatization, FSPE enrichment, and MS analysis [15].

omission of purification steps. Phenylhydrazine is the most often applied hydrazine derivatization reagent. (An earlier review discussed the significance of hydrazine derivatives in more detail [22].) Very recently, we proposed hydrazinonicotinic acid as a novel derivatized reagent for enhanced and selective detection of oligosaccharides, and found that the derivatives imparted improved MS sensitivity; in positive mode, S/N of the maltoheptaose was improved by more than one order of magnitude, and in negative mode, a 15-fold enhancement of S/N in the detection for sialylated glycan was observed compared to the control. In addition, through using different acid reagents as the catalyst, derivatized product signals corresponding to $[M + Na]^+$ or $[M + H]^+$ were obtained. This discovery represents complementary fragmentation patterns for the structural elucidation of oligosaccharides [23].

Derivatives of other sites

Permethylation has long been applied in glycomic analysis. In this strategy, hydrogen on hydroxyl groups, amine groups, and carboxyl groups are converted by methyl groups, producing a hydrophobic glycan derivative. The advantage of permethylation is that it enhances glycan signal strength in both ESI and MALDI-MS [24]. Detailed structural analysis of detached glycan often exploits the permethylation that stabilizes glycan; acidic structures as well as neutral ones can be detected in a positive ion mode [25]. Tandem MS of sodium adducts of permethylated glycan facilitates spectral interpretation by providing detailed information on linkage positions [26]. Permethylated glycan can also be applied to analyze sulfated glycan. Khoo *et al.* developed a novel technique for target MS analysis of permethylated glycan in sulfoglycomics which works by enriching the sulfated glycan, and has had considerable impact on the field of

MS-based glycomics [27]. Recently, the same group of researchers applied reverse phase-based nanoLC-MS/MS for sulfoglycomics analysis of permethylated glycan in negative ion mode. Results showed that this method can obtain complementary sets of diagnostic fragment ions which facilitate identification of various fucosylated, sialylated, and sulfated glycotopes [28].

Sialic-acid moieties of various sialylated glycans have been reported to be easily lost during MS analysis due to the instability of glycosidic bonds between sialic-acid and glycan residue during ionization. To solve this problem, several different derivatization methods have been developed in effort to stabilize the sialic acid group (e.g. methylester formation, pyrenylester derivatives, and amide derivatization.) Further, details regarding derivatization methods for sialic acid glycomics were presented in an earlier review [29]. Liu *et al.*, in one such study, described a facile derivatization in the presence of methylamine and (7-azabenzotriazol-1-yloxy) trispyrrolidinophosphonium hexafluorophosphate (PyAOP) for sialoglycomics by MALDI-MS. Results showed that the method effectively avoided the dissociation of sialic acid moieties by MALDI-MS, and both 3'- and 6'-sialyllactose derivatives achieve quantitative conversions regardless of their linkage types [30]. Xin *et al.* also tried to employ this derivatization for isomer-specific sialylated glycan profiling by nanoLC-MS, and found that the detection sensitivity of sialylated glycan increased by 2–80-fold compared to previous ESI-MS methods. Up to 19 tetrasialylated glycan species and 293 glycan species have been identified in human serum. Representative XIC of derivatized glycan and corresponding MS/MS spectra are shown in Fig. 4 [31].

DEVELOPMENT OF NEW MATRIX FOR MALDI-MS ANALYSIS

Alongside the development of soft ionization technology, MALDI-MS has become a very valuable research topic in the life sciences field. Due to lack of basic groups for protonation and the hydrophilic nature of glycan moieties, glycopeptides or glycans are more difficult to ionize than peptides or proteins. Glycoproteins/glycopeptides/glycans are inherently suppressed by non-glycosylated species. To develop successful MALDI analysis techniques, then, is to improve the ionization efficiency of glycopeptides and glycans. Scientists have used chemical derivation methods to improve the ionization efficiency of glycopeptides/glycans, in addition to developing new matrices. Compared to these

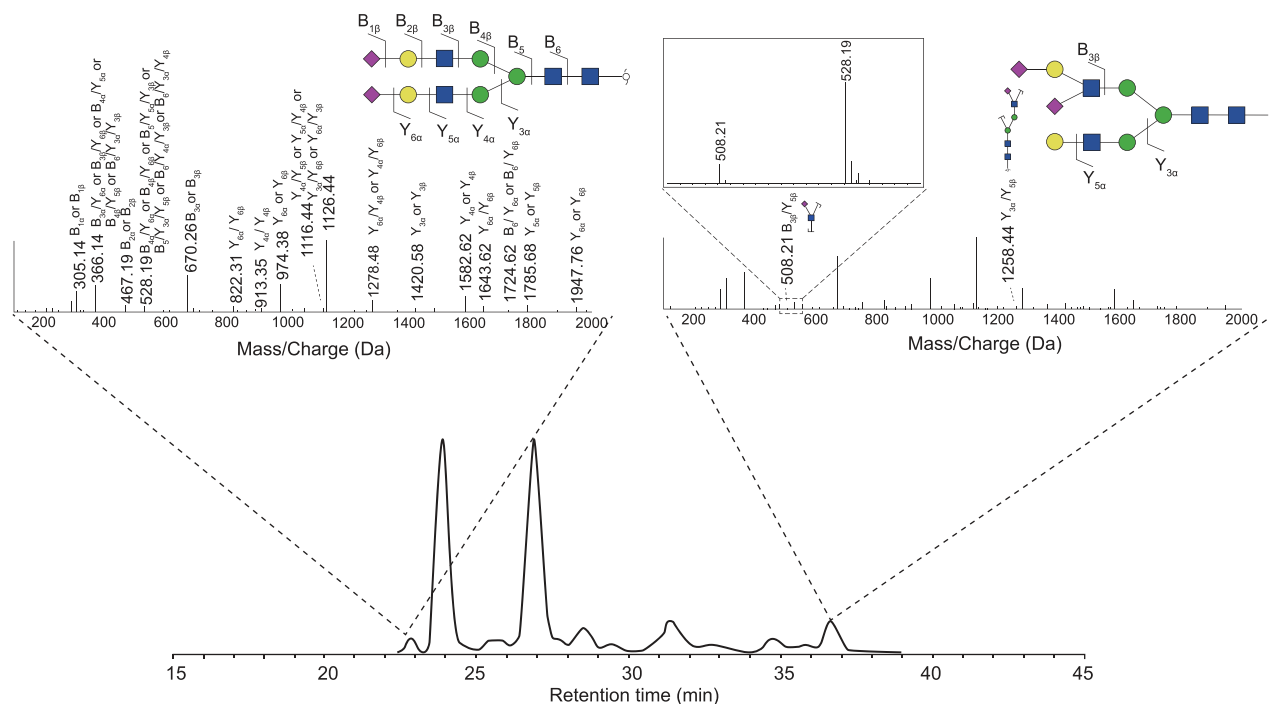


Figure 4. Representative XIC of derivatized glycans and corresponding MS/MS spectra [31].

derivatization methods, selecting a suitable matrix to facilitate the effective ionization of glycopeptides/oligosaccharides is simple and straightforward in order to identify signals in their native states and reduce the complexity of spectra. The most commonly used matrix is 2,5-Dihydroxybenzoic acid (DHB), but its weak points are notable; for one, inhomogeneous needle-shaped crystals are often formed around the rim of the MALDI spot after drying. As a result, an ideal spectrum is excessively time-consuming to produce, and a reproducibility of detection is not satisfactory.

Most of the matrices developed to improve the ionization efficiency of glycopeptides/glycans are ionic liquid matrices (ILMs), which consist of a conventional solid matrix such as 2,5-dihydroxybenzoic acid (DHB), α -cyano-4-hydroxycinnamic acid (CHCA), or trihydroxyacetophenone (THAP), and an organic base like tributylamine or pyridine, enabling a relative state of 'liquidity' under vacuum conditions. Due to the liquidity state, ILMs provide high homogeneity of matrix/analyte mixtures, enabling high reproducibility and high throughput analysis. Tanaka *et al.* developed a direct glycan labeling method with 3-minoquinoline (3-AQ) on a MALDI target using a liquid matrix 3-AQ/CHCA, and achieved high-sensitivity detection [32]. By removing the labeling process, Tanaka *et al.* introduced a similar liquid matrix 3-Aminoquinoline/p-Coumaric Acid (3-AQ/CA)

for glycopeptides and carbohydrates. Compared to 3-AQ/CHCA or DHB, 3-AQ/CA showed higher sensitivity (maximum of 1000-fold), suppressed the dissociation of sialic acids or phosphate groups, and improved the fragmentation of neutral carbohydrates [33].

Negative-ion mode analysis of glycopeptides/glycans has proven an effective complementary tool to positive-ion mode analysis, but the sensitivity of negative-ion mode is much lower than that of positive-ion mode. Wührer *et al.* used 4-chloro- α -cyanocinnamic acid (Cl-CCA) as a matrix to analyze glycosylated peptides containing sialic acid and released *N*- and *O*-glycans in negative-ion mode. Compared to DHB, the Cl-CCA crystal was smaller and homogeneous, improving the reproducibility and accuracy [34]. For neutral *N*-glycans, Tanaka *et al.* used anion-doped liquid matrix (G3CA), *p*-coumaric acid, and 1,1,3,3-tetramethylguanidine to detect neutral *N*-glycans in negative-ion mode, and found that the ILMs suppressed dissociation of sulfate groups or sialic acids of carbohydrates and the detection limits of anion-adducted *N*-glycan were 1 fmol/well for NO_3^- in negative-ion mode. In short, Tanaka achieved a very sensitive analysis method, especially for MS^2 measurement of neutral *N*-glycans in negative-ion MALDI- MS^n [35].

In addition to improving the ionization efficiency of glycopeptides and glycans, certain matrices can

also suppress peptide ionization, thereby allowing the glycopeptide or glycan to be detected directly in the peptide mixture. We proposed a new matrix, hydrazinonicotinic acid (HYNIC), which realized highly selective and sensitive analysis of oligosaccharides without pre-purification. The detection limit of maltoheptaose provided by HYNIC was as low as 1 amol. Furthermore, compared to the traditional matrix DHB, HYNIC exhibited several advantages: higher homogeneous crystallization and better salt tolerance, for example, and adequate fragmentation in tandem mass spectrometry, which provided useful information for structural elucidation of the glycan [36]. In a similar study, Yisheng *et al.* used diamond nanoparticles (DNPs) to enhance the sensitivity of carbohydrate analysis in MALDI. A trilayer was formed using the matrix plus diamond nanoparticles and analyte solutions. Owing to the high thermal conductivity of DNPs, low extinction coefficient in the UV-vis spectral range, and stable chemical properties, their proposed technique improved sensitivity of dextran about 79- and 7-fold compared to the control [37]. Cassidy *et al.* analyzed glycans using a GSH-capped iron oxide NP matrix, which produces a clean mass spectral background in the low m/z region and abundant ISD fragmentation. This matrix shows very favorable potential as far as glycan sequencing [38].

We developed a novel class of ILMs based on THAP for glycopeptide analysis. Compared to DHB, the S/N ratios of the proposed matrix were enhanced more than 10-fold. Moreover, glycopeptide derived from 10 ng tryptic digests of horseradish peroxidase was detectable with high sensitivity when using 2,3,4-THAP/DMA as matrix, while no signals were obtained when solid matrix 2,3,4-THAP was used [39].

ION DISSOCIATION METHODS

In addition to the sample preparation and ionization techniques discussed above, gas-phase fragmentation of glycopeptide or glycan ions during MS/MS experiments also plays an important role, as it provides a considerable degree of structural information that is difficult to achieve by any other means. Analyzing glycopeptides/glycans with single-stage MS (without dissociation) can provide valuable information such as measurements, but fall short of providing unambiguous glycopeptide or glycan assignments, as there are rather immediate limitations in the context of glycomics research. The MS/MS fragmentation patterns of glycopeptide or glycan ions can readily furnish much of the missing information, however [40]. Here,

we classify ion dissociation methods related to glycoproteomics or glycomics into four categories: ion/neutral, ion/electron, ion/ion, and ion/photon interactions [41].

Ion/neutral interactions

Ion/neutral interaction is the most common means of ion dissociation, typically involving collision of a precursor ion with a gaseous target to convert translational energy to internal energy; this process is called 'collision-induced dissociation' (CID). CID is the most commonly used ion dissociation method in glycoproteomics and glycomics. Combined with MALDI-TOF/TOF MS instruments, CID is regarded as a powerful tool for analysis of carbohydrates and glycoconjugates [42]. It can also be incorporated into other MS instruments. For example, Both *et al.* used ion-mobility MS to separate epimeric O-linked glycopeptide and enabled characterization of the attached glycan based on the drift time of the monosaccharide product ions generated after CID [43]. For analysis of glycomics, CID is usually incorporated with ESI-MS due to its compatibility with separation techniques such as LC [44] or capillary electrophoresis (CE) [45].

In glycoproteomics research, especially N-linked glycoproteomics, CID has garnered a great deal of attention either for analysis of deglycosylated peptides or glycopeptides. However, classical ion-trap CID MS cannot be directly used for N-linked glycopeptide analysis, because it results in the formation of oxonium ions of HexNAc (m/z 204) and HexHexNAc (m/z 336) in addition to other fragments resulting from cleavage within the glycan, which are not detected in ion-trap CID due to the low m/z cutoff. As a result, HCD (higher energy collision dissociation) is preferred for analyzing unabridged glycopeptides.

Around the time it was first proposed, HCD was accomplished via C-trap coupled to an Orbitrap analyzer (LTQ Orbitrap, Thermo Fisher Scientific) as a collision chamber to enable triple quadrupole-like fragmentation in the Orbitrap. The fragmentation of HCD yields similar patterns as CID, except that the HCD spectrum contains more ions in a low-mass region of the spectrum [46]. Yehia *et al.* first considered its application to glycosylation analysis and found that HCD generated distinct Y1 ions (peptide + GlcNAc) that were useful in N-glycosylation sites assignment. The common glycan oxonium ions in addition to the Y1 ion are also detectable, because HCD overcomes the 1/3 m/z cutoff associated with ion trap CID [47]. There are indeed numerous

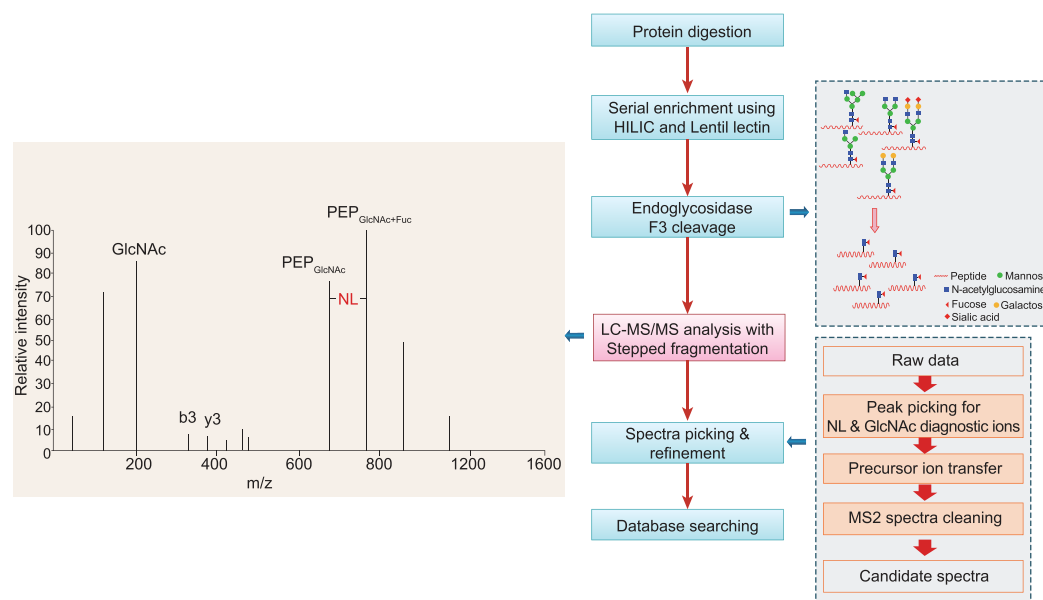


Figure 5. The workflow of strategy for CF glycoproteomics analysis with stepped CID fragmentation [48].

advantages to HCD in glycoproteomics research compared to classical ion-trap CID.

By coupling low-energy CID and HCD, Qian *et al.* introduced a novel strategy for the precise and large-scale identification of core fucosylated (CF) glycopeptide based on stepped CID fragmentation. The workflow of this strategy (Fig. 5) consists of three key steps. (i) Glycopeptides in the peptide digests are obtained through sequential enrichment by HILIC and lentil lectin. (ii) Endoglycosidase F3 is used to simplify the CF glycopeptide, leaving the innermost GlcNAc and fucose α 1–6 linked to the GlcNAc (GlcNAcFuc) on the CF glycopeptide. (iii) The simplified CF glycopeptides are analyzed by LC-MS/MS with stepped (MS^2) fragmentation. The stepped fragmentation can be run using the ‘Stepped NCE’ (normalized collision energy) function provided by the Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo, USA). When the NCE value is set at a low region, the glycosidic bond between GlcNAc and GlcNAc-Fuc breaks, resulting in mass shift of 146 Da which contributes to the highest peaks in the MS^2 spectrum. At a moderate NCE energy, the simplified glycopeptide is fragmented into a series of b/y-type ions or a GlcNAc residue attached with b/y-type ions, and when the NCE energy rises again, the spectrum of simplified glycopeptide is dominated by fragment ions of GlcNAc residue. All this information can be presented in one spectrum using the ‘Stepped NCE’ function. The method was successfully applied to analyze CF glycoproteomes of mouse liver tissue and HeLa cell samples spiked with standard CF glycoprotein [48].

Ion/electron interactions, ion/ion interactions

Electron capture dissociation (ECD) was first demonstrated by Zubarev, Kelleher and McLafferty [49]. Several years later, Hunt *et al.* introduced a similar approach, electron transfer dissociation (ETD), involving electron transfer to multiply protonated precursor ions from negatively charged and radical reagent ions [50]. In both cases, liberation of the recombination energy is accompanied by radical-driven fragmentation pathways that are not observed in CID. Instead of b- and y-ions produced by CID, mass spectra of peptides generated by ECD or ETD are usually occupied by c- and z-ions [40]. These methods (collectively called ExD) show significant potential for the characterization of glycopeptides, because the modification is largely affected by the fragmentation process, allowing assessment of attachment sites and information on the multiple glycans associated with each site. However, because ExD provides only short peptide sequences, it is usually combined with other dissociation methods such as CID or HCD in glycoproteomics research.

Developments in ExD techniques benefit N-linked glycoproteomics research. By incorporating CID and ETD, Zou *et al.* developed a chemical and enzymatic strategy as an alternative to mass spectrometric analysis of aberrant glycosylation. First, glycopeptides bearing terminal sialylation were enriched with reverse glycoblotting followed by the release of deglycosylated peptides. Mass spectrometric analysis of both CID and ETD demonstrated a

specific pattern of aberrant glycosylation, which carried both terminal sialylation and core fucosylation. The application of this method to a study on serum from hepatocellular carcinoma (HCC) patients revealed 69 aberrant sites [51]. In a similar study, Cooper *et al.* coupled HCD with ETD and developed a strategy for the targeted analysis of N-linked glycopeptides in complex mixtures; their technique does not require prior knowledge of the glycan structures or pre-enrichment of the glycopeptide. The researchers applied online ZIC-HILIC liquid chromatography HCD production-triggered ETD, and glycopeptides (which were pre-purified by ZIC-HILIC and HCD) were used to generate GlcNAc oxonium ions as the trigger for ETD. The ETD spectrum provided information on glycosylation sites and peptide sequences, while the corresponding HCD spectrum provided information on glycan structures. This product ion-triggered ETD approach overcomes glycan heterogeneity and allows target analysis of glycopeptides [52].

As mentioned above, ExD also shows more advantages in researching O-linked glycoproteomics, which suffer the lability of O-glycosylation in CID or HCD, because it can induce peptide bond cleavage with minimal loss of PTM [53]. Clausen *et al.*, for example, applied Zinc-finger nuclease (ZFN) targeting to truncate the O-glycan elongation pathway in human cells. They were able to generate stable 'simple cell' lines with homogenous O-glycosylation, which only expressed truncated GalnAc α or NeuAc α 2-6GalnAc α O-glycans, allowing straightforward isolation and sequencing of GalNAc O-glycopeptides from total cell lysates using lectin chromatography and nanoflow liquid chromatography–MS with HCD and ETD fragmentation. By this means, more than 100 O-glycoproteins with more than 350 O-glycan sites were identified including a GalNAc O-glycan linkage to a tyrosine residue. The vast majority of these O-linked glycopeptides and O-glycan sites were previously unidentified [54]. In another example provided by Nilsson *et al.*, human cerebrospinal fluid (CSF) was pretreated with PNGase F to remove N-glycan and enriched O-glycopeptides with reversible HC, facilitating the selective characterization of O-glycopeptides. What is more, they enabled the use of an automated CID–MS²/MS³ search protocol for glycopeptide identification. ECD and ETD were then used to pinpoint the glycosylation sites, identified as predominantly core-1-like HexHexNAc-O-structure attached to one–four Ser/Thr residues [55].

Ion/photon interactions

Irradiation of precursor ions with photons offers an alternative to dissociation through ion–neutral collisions or electron capture/transfer. Ion dissociation by means of IR irradiation is typically performed, using a continuous wave CO₂ laser at a wavelength of 10.6 μ m. This process is called 'infrared multiphoton dissociation' (IRMPD) in the literature, because tens to hundreds of IR photons must be absorbed to accumulate sufficient vibrational energy to access dissociation. This method is an alternative to CID for glycomics research [56]. Lebrilla *et al.* applied Fourier transform ion cyclotron resonance (FTICR)-IRMPD MS for analysis of O-linked glycans of jelly coat glycoproteins in *X. borealis* eggs, and successfully preliminarily determined 29 neutral glycans of *X. borealis* [57]. Similar studies evaluated the IRMPD behavior of N-glycopeptides [58] and O-glycopeptides [59], as well. This technique has not been used in proteome-wide analysis of glycoproteins, however, due to instrument-related and sample complexity limitations. Another dissociation method of ion/photon interactions, ultraviolet photo-dissociation (UVPD), has also been used to analyze O-glycopeptides but not glycoproteomics [60]. It may be some time before these strategies become applicable for glycoproteomics.

SOFTWARE FOR INTERPRETATION OF MS DATA FROM GLYCANS AND GLYCOPEPTIDES

Unlike protein transcription or DNA replication, glycosylation is a non-template driven process—the glycan may have a large degree of variability, be long or short, branched or linear, or linked in a variety of ways [61,62]. The inherent complexity of glycosylation poses a large challenge to the bioinformatics of glycoproteome. Compared to automated proteomics software solutions, the software for interpretation of MS data from glycans and glycopeptides usually require a great deal of manual intervention and can only focus on a specific area.

Glycan analysis

The most common way that protein glycan populations are identified is by enzymatically cleaving the glycan substituents and analyzing the monosaccharide residues directly [61,62]. The released glycan can then be analyzed by MS (usually MALDI-MS) or MS/MS. The most often used program for glycan spectrum annotation is GlycoWorkbench, which

evaluates glycan compositions (which are proposed user-by-user) by searching the spectral peak list of user-input MS data for matches between calculated theoretical glycan masses and corresponding m/z values [63]. GlycoWorkbench also has a particularly user-friendly graphic interface for users to set up/modify glycan structures. Similar software for glycan spectral interpretation includes Cartoonist [64] and GlycoSpectrumScan [65].

The annotation of *N*-glycan MS spectra with glycan compositions is generally fairly simple [66]. Software coupled with specifically designed databases have been developed to resolve isomeric compounds thereby allowing complete structures to be readily identified [67]. One such project was recently launched: building a database containing both *N*- and *O*-glycan structures with their tandem MS spectra and chromatographic retention times [68].

Although all the above software provide potential glycan assignment in a high-throughput manner, none of them report global false positive rate (FPR). Lack of proven quality control is one of the weak points of glyco-bioinformatics. Unlike proteins, which have a linear structure database derived from DNA/RNA analysis, conventional FPR evaluation methods such as the target-decoy database approach are not particularly applicable to glycan spectral analysis.

Protein glycosylation site analysis

After the release of glycan from glycoprotein, the glycosylation site on the protein can be elucidated by routine proteomics techniques. PNGaseF releases most types of *N*-glycans on proteins, and adds a mass tag (0.98402) to ASN. This tag can be used for the identification of glycosylation sites. Over 6000 *N*-glycosylation sites from five mouse tissues have been reported [69]. Other enzymes or chemical methods could also release *N/O*-glycans and add mass tags to the original amino acids [62]. Interpretation of glycosylation sites is usually performed through proteomics research, and does not require glycan-based bioinformatics knowledge.

Intact glycopeptide analysis

Confident characterization of the microheterogeneity of protein glycosylation through identification of intact glycopeptides remains one of the toughest analytical challenges for glycoproteomics researchers. Glycopeptides are inherently more diverse and have less ionization efficiency than unmodified peptides due to the enormous complexity of glycosylation

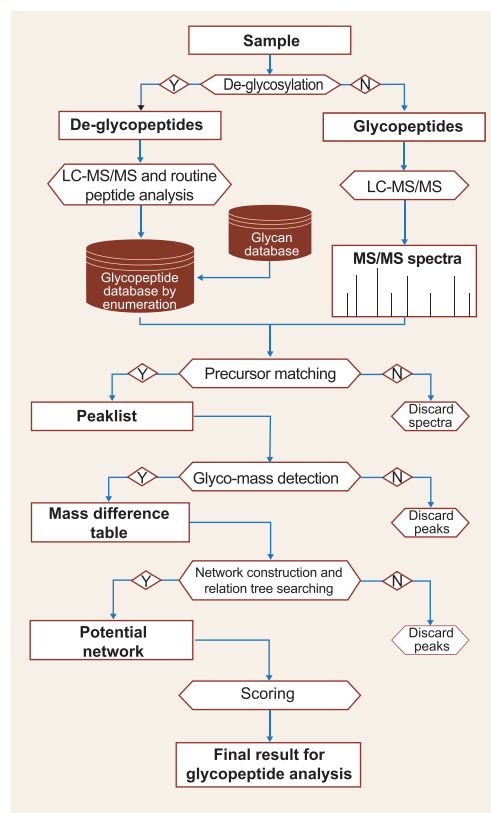


Figure 6. Overview of proposed strategy for automated glycopeptide analysis. Sample is prepared in two parallel experiments to generate glycopeptide database and collect glycopeptide spectra with and without deglycosylation; homemade software elucidates glycopeptide spectra in several steps as shown by the decision tree.

and the nature of glycan. In recent years, many researchers have developed different pipelines for glycopeptide interpretation, however.

The most widely used MS configuration for the interpretation of intact glycopeptides is CID-MS/MS coupled with ETD-MS/MS. Generally, in a CID-MS/MS spectrum, sufficient Y ions can be observed to deduce the glycan of a glycopeptide, while some peptide backbone cleavages can be interpreted in ETD-MS/MS. A few software tools have been developed to identify intact glycopeptides by integrating complementary information from CID- and ETD-MS/MS, such as GlypID [70,71] and GlycoPep Grader [72,73]. However, the sensitivity and the applicable scope of ETD-MS/MS are arguably limited compared to HCD- and CID-MS/MS in current-generation MS instruments—only a few hundred glycopeptides can be identified in complex samples.

To increase the overall throughput, we developed software using de-glycopeptide database coupled with CID-MS/MS (Fig. 6). First, parallel

experiments were performed on the same sample to collect de-glycopeptides and intact glycopeptides spectra; after that, an integrated software package generated the de-glycopeptide database, and interpreted intact glycopeptide spectra using both de-glycopeptide and glycan databases [74]. Zou *et al.* explored a similar method that included retention time in the de-glycopeptide database [75]. These approaches allow about a thousand glycopeptides in complex samples to be identified.

Recent studies have investigated the use of search engines to detect peptide backbone signals from CID-MS/MS with higher collision energy. For example, the commercial direct protein database search engine Byonic for glycopeptide MS/MS data, which considers the glycan as a normal variable modification attached on the glycosylation site [76]. Unfortunately, this strategy can result in identifications with high FPR even if the peptide-spectrum match (PSM) score is high, because the FDR control is applied at the peptide level only and not at the glycan level [77].

For core-fucosylated glycopeptides, the glycans on proteins can be partially released by enzymes such as Endo series. After Endo treatment, there are usually only one or two glycan(s) still attached to the peptide/protein, which significantly lowers the complexity of glycopeptide analysis. Qian *et al.* developed an MS-based approach for interpreting core-fucosylated glycopeptides derived from Endo treatment. They first used neutral-loss-triggered MS3 to collect high-quality spectra of the core-fucosylated glycopeptides, and developed a corresponding spectral analysis pipeline. Diagnostic ions from core glycans and neutral-loss from fucose were then integrated into the database [78]. They then developed a stepped energy approach specifically for core-fucosylated glycopeptides, in which complementary information from different energies was integrated; the overall performance of the system increased about 10-fold [48].

Due to space limitations and lack of any standard glycopeptide spectral dataset, a fair comparison between different software is not fully possible. A recent review focused on automated glycopeptide analysis provided useful information regarding future developments in software design and data handling [79], and another review listed free available *N*- and *O*-glycopeptide analysis tools [61]. Glycomics and glycoproteomic research would benefit significantly from access to the automated tools described here. Compared to search engines in proteomics fields that focus primarily on database searching or generating spectral libraries, glycoproteomics software packages are much more diverse (and less developed.) Ideally, the software for glycan analysis

should be able to perform automated spectral interpretation of many different kinds of glycans without sacrificing rigid quality control or structural annotation; the software for glycopeptide analysis should be able to perform systematic analysis of *N*- and *O*-glycopeptides in complex samples without sacrificing quantitiveness.

GLYCOPROTEOME AND GLYCOME QUANTITATION

To better understand the function of glycosylation, it is crucial to develop quantitative glycoproteome and glycome analysis strategies among different biological samples and physiological conditions. As of now, many research efforts have been made to establish quantitative analysis methods for glycoproteins/glycopeptides and glycans [80]. We will focus this review mainly on two aspects: quantitation of glycoproteins/glycopeptides and quantitation of glycans.

Glycoproteome quantitation

Various strategies have been developed to quantify the abundance of glycoprotein while identifying glycosylation sites. These strategies can be split into two basic categories: relative quantitation and absolute quantitation. Relative quantitation can be further divided into metabolic labeling, chemical labeling, and enzymatic ¹⁸O-labeling methods.

Metabolic labeling

The metabolic labeling strategy for glycoproteins/glycopeptides is based on incorporating stable isotope tags into biomolecules via biosynthetic pathways. Stable isotope labeling by amino acid-coded tagging (AACT, also called SILAC) in cell culture is a well-accepted quantitation method [81,82]. Mann *et al.*, for example, developed a method using a super-SILAC mix from several labeled breast cancer cell lines as a cross-sample internal standard for accurate quantitation [83]. In total, 1398 unique *N*-glycosylation sites from 11 cell lines that are representative of different stages of breast cancer were identified and quantified. It is worth mentioning that metabolic labeling approaches remain limited due to the difficulties of culturing cells, as well as relatively high cost.

Chemical labeling

Chemical labeling is commonly applied in proteomics research due to its flexibility and convenience. Zou *et al.* developed a solid-phase-based

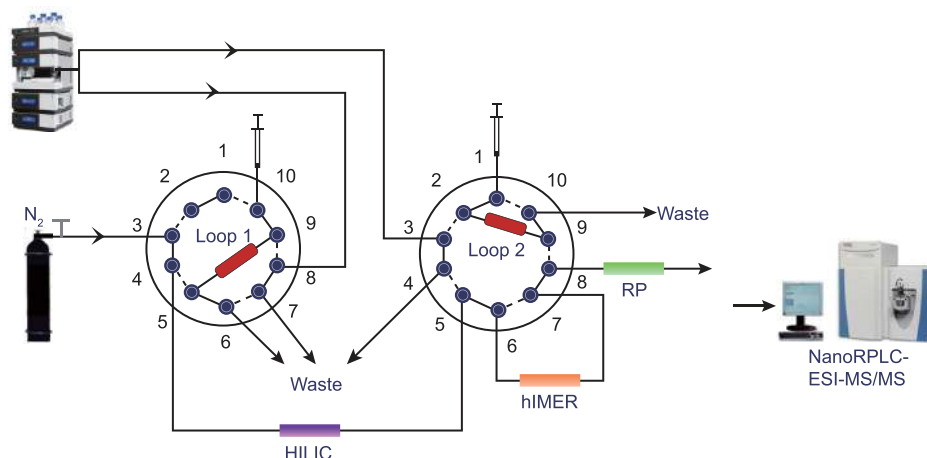


Figure 7. An integrated sample pretreatment platform for quantitative *N*-glycoproteome analysis.

labeling approach by integrating glycopeptide enrichment and stable isotope labeling on hydrazide beads [84]. The method showed high enrichment recovery (10%–330% improvement) and high detection sensitivity, where 42% of the annotated glycosites were quantifiable using only 10 μg of four standard glycoprotein mixtures and 400 μg of bovine serum albumin interference as a starting sample. The conventional protocol for quantitative analysis of glycoproteomes is usually off-line, resulting unfortunately in lengthy sample preparation time and risk of sample loss. Zhang *et al.* presented an integrated platform for online quantitative *N*-glycoproteome analysis, a combination of HILIC enrichment, deglycosylation, and dimethyl labeling (Fig. 7) [85]. Using this platform, 43 upregulated and 30 downregulated (Hca-F/P) *N*-glycosylation sites, and 11 significantly changed *N*-glycoproteins from two types of hepatocarcinoma ascites syngeneic cell lines were successfully quantified. There are common drawbacks to chemical labeling, however, such as the additional reaction steps necessary and low labeling efficiency.

Enzymatic ^{18}O labeling

Enzymatic ^{18}O labeling has been extensively studied by proteomic analysis researchers. Proteases that have high specificities on C-terminal residues, such as trypsin, Glu-C, and Lys-C, can stably incorporate two atoms of ^{18}O into a newly generated C-terminal carboxyl of a peptide during proteolytic digestion, displaying a 4 Da mass shift for peptides. We developed a tandem ^{18}O stable isotope labeling (TOSIL) method to precisely quantify the changes in glycoprotein expression as well as changes in individual *N*-glycosylation site occupancy [86]. In this work, a unique mass shift of 6 Da was identified

for *N*-glycosylated peptides with single glycosylation sites, which can be distinguished from non-glycosite peptide pairs (provided they have a mass difference of 4 Da,) thus enabling the simultaneous quantitation of glycoprotein and glycosylation sites. This labeling method is efficient, simple, and low cost. Based on this work, we recently proposed another comprehensive strategy, PNGase F-catalyzed glycan ^{18}O labeling (PCGOL) [87]. This method can be used for comprehensive *N*-glycosylation quantification, achieving simultaneous quantification of glycans, glycopeptides, and glycoproteins in a single workflow. The method showed good linearity and high reproducibility within at least two orders of magnitude in the dynamic range, however, due to the lack of specific enzymes for *O*-glycan release, enzymatic ^{18}O labeling may not be applicable for *O*-glycoprotein/*O*-glycopeptide quantitation.

Utilizing the above stable isotope labeling method, combination with selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) assays may provide unparalleled sensitivity and selectivity for absolute quantification of the glycoproteome. Domon *et al.*, for example, provided an MS-based absolute quantitation method for *N*-glycoproteins [88]. Taking advantage of HC enrichment and SRM, this strategy was proven to efficiently and precisely facilitate glycoproteomic quantitation.

Protein glycosylation stoichiometry quantitation

Many studies have been conducted in effort to accurately quantify various glycosylation stoichiometries, which is undoubtedly essential to further understanding the complex functions of

glycosylation. Hsieh-Wilson *et al.* used an engineered glycan transferase and a well-established chemical reaction to incorporate several distinct mass tags on O-GlcNAcylated proteins *in vivo* [89], and were able to determine whether a protein was singly, doubly, or multiply glycosylated by simple inspection of the mass-shifted bands on a Western blot. In addition, *in vivo* glycosylation stoichiometries were determinable by quantifying the relative intensities of each band. Sun and Zhang developed another method that enables large-scale determination of absolute glycosylation stoichiometry using three independent relative ratios [90]. The stoichiometry of a given glycoform can be determined by comparing the changes in relative abundance between the glycoform-containing and non-glycoform-containing forms of the glycosite using quantitative proteomics and glycoproteomics. These methods are useful especially in studies on disease-associated changes in protein glycosylation occupancy and different glycoforms.

Glycome quantitation

Glycan quantitation can provide valuable information on glycan abundance alterations and glycan structure aberrance under various physiological and pathological conditions, thus representing the diagnostic biomarker for a variety of vital human diseases. Similar to glycoproteomics quantitation, MS-based quantitation of glycans benefits from the incorporation of stable isotopes in the process of glycan generation or sample preparation. To date, three main strategies have been developed for glycomic quantitation: metabolic labeling, chemical labeling, and enzymatic assisted labeling.

Metabolic labeling

The metabolic labeling approach applies labels during the cell culture process. In a related study, Orlando *et al.* developed the isotopic detection of aminosugars with glutamine (IDAWG) method [91]. Based on the fact that the amide side chain of Gln is the sole source of nitrogen for the biosynthesis of GlcNAc, GalNAc, and sialic acid via the hexosamine biosynthetic pathway, ^{15}N -Gln media was used to label mouse embryonic stem cells N- and O-glycans within a 72-h labeling period. By introducing metabolic labeling strategies, differentially labeled cells can be mixed together at the beginning of the analytic procedure to minimize sample variations. Drawbacks to this method include the fact that it can only quantify cell-culture samples and is inapplicable to sera, urine, saliva, or tissues, and that the uncertain amount of heavy labeled aminosugars complicates the mass spectra interpretation.

Chemical labeling

Chemical labeling strategies incorporate stable isotopes into glycan via chemical reaction. Permethylolation labeling and glycan-reducing end amination are the main chemical labeling methods. Permethylolation labeling utilizes chemical derivatization of hydroxyl groups with iodomethane to introduce mass tags and quantitation via comparison of the peak areas or intensities of different samples. Hu, Desantos-Garcia and Mechref developed a permethylolation labeling strategy using iodomethane (light) and iodomethane- d_1 or d_3 (heavy) reagents [92], and applied the method to explore glycomic variance in blood serum samples. Glycan-reducing end amination or glycan-reductive amination modified the reduced end of N-glycans, and could be used in relative quantification. Like permethylolation, derivatizations can usually improve the MS performance of labile glycan. Glycan-reducing end amination modified by the reduced end of N-glycan with isotope-coded labels also allows for relative quantification. For example, a deuterium reagent, 1-(d_5) phenyl-3-methyl-5-pyrazolone (d_5 -PMP), was synthesized and used for relative quantitative analysis of oligosaccharides by MS using d_0/d_5 -PMP stable isotopic labeling [93]. Orlando *et al.* [91] introduced another strategy of methylation-based isotope labeling that uses $^{13}\text{CH}_3\text{I}$ and CH_2DI as labeling reagents. The mass difference between the two reagents is merely 0.002922 Da, however, the multiple methylation sites enable quantitation between two samples using a high-resolution mass spectrometer. Moreover, with this method, heavy and light N-glycans have identical chromatographic retention times and as such can outperform deuterium-encoded labeling reagents [94].

Because most homemade isotope-coded reagents are not commercially available, we developed a glycan-reductive, isotope-coded amino acid labeling method (GRIAL) for N-glycomic quantitation that successfully labeled the reducing end of the N-glycan with isotope-coded arginine through reductive amination. The chemical labeling strategy offered a large mass difference of 6 Da, avoiding inaccurate glycan quantitation due to isotope interference [95]. This method has been successfully applied to quantify serum N-glycans from healthy control and CRC patients, and showed favorable quantitation results. Additionally, compared to homemade isotopic tags, the GRIAL strategy used commercially available isotope-coded amino acid as label material—this feature makes this novel strategy easily accessible for a variety of laboratories. The method also is applicable to higher-plex labeling via a well-designed combination of labeling reagents

(like Arg0, Arg6, and Arg10,) thus broadening the range of its application.

Enzymatic ^{18}O labeling

Enzymatic ^{18}O -labeling strategies incorporate heavy or light atoms via the enzymatic process. Enzymatic ^{18}O labeling shows advantages for glycomics similar to those posed by protease-catalyzed ^{18}O labeling for proteomics, like low cost, easy operation, and high efficiency. We developed a glycan-reducing end ^{18}O -labeling method (GREOL) using endoglycosidase to release N-linked glycans from glycoproteins and to incorporate an ^{18}O atom into the N-glycan-reducing end in the presence of H_2^{18}O [96]. This method provides good linearity with high reproducibility within two orders of magnitude in the dynamic range, and was also used to analyze changes in human serum N-glycans associated with HCC, proving it to be an effective tool for MS-based glycan quantitation. However, the isotope overlap caused by only a 2 Da mass difference between 16O- and 18O-labeled glycans requires an extra deconvolution step to achieve accurate quantitation. In addition, endoglycosidases are not as comprehensive as peptide-N-glycosidase F (PNGase F) for glycan release, making the method less desirable. In effort to solve this problem, we recently devised an improved method called glycan-reducing end dual isotopic labeling (GREDIL) for mass spectrometry-based, quantitative N-glycomics [97]. This work added an additional reduction step with $\text{NaBH}_4/\text{NaBD}_4$, which not only stabilized the enzymatic ^{18}O -labeling glycan but also increased the mass difference between the two samples to 3 Da, making the quantitation results more accurate and reliable.

Despite the powerful quantitation ability of MRM strategies for glycoproteomes, MS-based absolute quantitation for glycomes remains challenging due to a lack of glycan internal standards and generally poor ionization efficiency of glycans. There have still been many valuable contributions to the literature regarding absolute quantitation of glycans, however. Kim *et al.* [88], for example, proposed a relative and absolute quantitative strategy using carboxymethyl trimethylammonium hydrazide derivatization to generate a permanent cationic charge at the reducing end of neutral oligosaccharides. The treated glycans were later analyzed with internal standards of dextran ladders (using MALDI-TOF MS) and verified by comparison with those performed by conventional normal-phase (NP)-HPLC profiling. This method shows promise for neutral glycan analysis in the future, especially quantifying minute glycan samples with undetectable levels using HPLC [98].

NOTABLE FINDINGS IN GLYCOPROTEOMICS AND GLYCOMICS

Protein glycosylation is implicated in virtually all cellular processes (e.g. protein folding, molecular recognition, gene expression, signal transduction, protein turnover, cell cycle control, and stress protection.) Although protein glycosylation is crucially involved in many diseases, this review focuses solely on its relationship with cancer.

O-GlcNAcylation is a type of post-translational modification through which only the β -N-acetylglucosamine (O-GlcNAc) is added to serine and threonine residues of proteins. In addition to its important role in biomarker discovery, O-GlcNAc modification is also closely related to nervous system function maintenance.

Glycoproteomics and cancer

Cancer is one of the most dangerous and harmful diseases that threaten human beings. Although there have been notable advancements in cancer treatment in recent decades, cancer still claims hundreds of thousands of lives each year across the globe. More and more studies show that protein glycosylation is essential in cancer development and progression [99]; for example, the core-fucosylated glycoform of AFP, which is also known as AFP-L3, can be used for specific diagnosis of hepatocellular carcinoma (HCC). It has been approved by the FDA as a supplemental test in patients with elevated total AFP. Recently, Liu *et al.* developed an approach that utilizes lectin arrays combined with an LC-MS/MS-based quantitative glycoproteomic platform to identify glycosylated-protein biomarkers in early stages of HCC. They found that fucosylated glycoprotein is elevated in HCC patient serum, and that C3, CE, HRG, CD14, and HGF are candidate biomarkers for distinguishing early HCC [100]. Using a similar strategy, we combined the hydrophilic affinity (HA) and HC enrichment of glycopeptides with nano-LC-ESI-MS/MS analysis and found 300 different glycosylation sites within 194 unique glycoproteins, in which 172 glycosites had not previously been determined experimentally. Results showed that N-glycosylated alpha-fetoprotein, CD44, and laminin are implicated in HCC development and metastasis [101]. Recently, Zhang *et al.* found 26 differentially expressed serum glycoproteins derived from defined stages in an orthotopic xenograft tumor model by using isobaric tags for relative and absolute quantitation (iTRAQ)-based quantitative N-glycoproteomic analysis. Among them, the lower N-glycosylated sEGFR was a potential candidate for metastasis-associated HCC biomarkers [102].

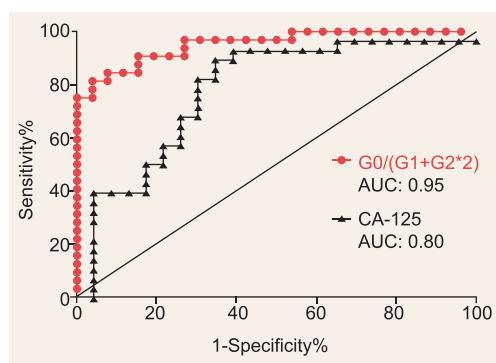


Figure 8. A ROC curve for ratios to differentiate ovarian cancer from benign gynecological conditions; the ROC curve for CA-125 alone. IgG galactosylation measured from relative intensities of IgG digalactosyl (G2), monogalactosyl (G1), and agalactosylated (G0) *N*-glycans according to the formula $G0/(G1 + G2 \cdot 2)$ [103].

Ovarian cancer can cause multiple tumors in patients, and cannot be detected until large-scale and remote metastasis occurs. CA-125, the most often used biomarker for ovarian cancer detection, cannot provide fully accurate diagnosis due to its poor specificity (i.e. it may also mark many benign gynecological conditions.) Therefore, researchers are urgently tasked with discovering metastasis-related biomarkers for the detection of ovarian cancer in its occult metastasis stage. Gu *et al.* quantitatively analyzed serum IgG galactosylation in ovarian cancer patients with similar elevated CA-125 levels, and found that combining quantitative alteration of IgG galactosylation with CA-125 may represent a robust approach to the differential diagnosis of ovarian cancer (Fig. 8) [103].

Breast cancer is a slowly developing disease that involves the accumulation of genomic aberrations including amplifications, deletions, and rearrangements. In effort to elucidate the steps of breast cancer progression, Macher *et al.* identified more than 40 glycoproteins that are differentially expressed in either premalignant (MCF10AT) or fully malignant (MCF10CA1a) cell lines using a two-dimensional liquid chromatography-tandem MS system. They found that the collagen alpha-1 (XII) chain is expressed at dramatically higher (near 10-fold) levels in breast cancer patients, and as such, can be used as an identification marker [104]. O'Connor-McCourt *et al.* similarly found that 90 glyco-proteins are more highly expressed and 86 proteins are underexpressed in triple-negative breast tumors relative to the luminal tumors, and can be identified using a label-free LC-MS based approach [105].

Core fucosylation (CF) is a special glycosylation pattern of proteins that is strongly related to cancer. Qian *et al.* [103] used an innovative strategy to

identify CF glycopeptides at a large scale, integrating the stepped fragmentation function, one novel feature of quadrupole-orbitrap mass spectrometry, with 'glycan diagnostic ion'-based spectrum optimization. By using stepped fragmentation, 1364 and 856 CF glycopeptides belonging to 702 and 449 CF glycoproteins were identified in HCC [48]. Recently, Lubman *et al.* used LCA enrichment to improve the identification of core-fucosylation sites in pancreatic cancer patient serum; they identified 630 core-fucosylation sites from 322 core-fucosylation proteins in the serum, and eight core-fucosylation peptides exhibited a significant difference between pancreatic cancer and other controls, suggesting valuable potential diagnostic biomarkers for pancreatic cancer [106]. Studies on core-fucosylated glycopeptides are a practical example demonstrating the significant potential of this proposed method for glycoproteome analysis. We anticipate this flexible new method will be further applied in other research fields, as well.

In addition to large-scale study of the glycoproteome, specific protein glycosylation studies have also made notable achievements related to identifying cancer development and metastasis. Bone morphogenetic protein-2 (BMP-2), a glycosylated protein, has been demonstrated to play a key role in TGF- β -mediated cancer metastasis, for example. Shen *et al.* found that BMP-2 contains three potential *N*-glycosylation sites (N135, N200, and N338) and that N135 glycosylation promotes BMP-2 secretion and mediates breast cancer cell metastasis [107].

The membrane glycoprotein CD133 is a commonly used marker for cancer stem cells, as it contributes to cancer initiation and invasion in a number of tumor types. Jiang *et al.* identified eight potential *N*-glycosylation sites on the CD133 by mass spectrometry, and Asn548 glycosylation as involved in hepatoma cell growth via affecting the β -catenin-signaling pathway, which may have implications for cancer-targeting therapies [108].

E-cadherin is a central molecule in the process of gastric carcinogenesis. Pinho *et al.* found four potential *N*-glycosylation sites of E-cadherin, and found that Asn-554 is the key site; selectively modified with β -1, 6-GlcNAc-branched *N*-glycans are catalyzed by GnT-V. This aberrant glycan modification on this specific asparagine site of E-cadherin affects critical functions in gastric cancer cells by affecting E-cadherin cellular localization, cis-dimer formation, molecular assembly and stability of the adherens junctions, and cell-cell aggregation, and was further observed in human gastric carcinomas [109].

Glycomics and cancer

In addition to protein sites, oligosaccharide structures may also be altered during the initiation or progression of cancer [99]. Gu *et al.* employed a quantitative glycomics approach based on metabolically stable isotope labeling to compare the differential *N*-glycosylation of secretome between an ovarian cancer cell line, SKOV3, and its high metastatic derivative, SKOV3-ip. They identified 17 *N*-glycans altogether, and found that the *N*-glycans with bisecting GlcNAc marked ovarian cancer in the early peritoneal metastasis stage, which may accordingly improve prognoses for ovarian cancer patients [110].

In vitro models are frequently used to study cancer, but little is known about the differences that may exist between these model cell systems and tumor tissue. Packer compared the membrane protein glycosylation of five colorectal cancer cell lines (SW1116, SW480, SW620, SW837, and LS174T) with epithelial cells from colorectal tumors using liquid chromatography tandem mass spectrometry, and found five abundant *O*-glycans in the tumor cells that were undetected in the low-mucin-producing cell lines; the *O*-glycans included the well-known glycan cancer marker, sialyl-Tn, which is associated with mucins [111].

In addition to the large-scale observation of glycan changes, specific protein glycan changes are important in protein function. A C-type, lectin-like, oxidized low-density lipoprotein receptor-1 (LOX-1) was identified as the receptor for Hsp60, for example. Inhibition of Hsp60 by LOX-1 decreases Hsp60-mediated cross-presentation of OVA and specific CTL responses and protective tumor immunity *in vivo* [112]. Gu *et al.* paired nano-LC with MALDI-QIT-TOF MS and identified the Asn 134 glycosylation site in high-molecular-weight CLEC-2, which contained complex types of bi-antennary, tri-antennary, tetra-antennary, N-linked, and fucosylated glycans [113]. Ag recognition and Ab production in B cells are major components responding the humoral immune system. Li *et al.* found that the CF of IgG-BCR mediates Ag recognition and, concomitantly, cell signal transduction via BCR and Ab production [114].

O-GlcNAc in cancer and nerve system disease

O-GlcNAc is a post-translational addition of β -*N*-acetylglucosamine to serine and threonine residues of nuclear and cytoplasmic proteins. According to Hart's 1984 study, O-GlcNAc glycosylation is chemically and enzymatically labile, often substoichiometric, and subject to complex cellular regu-

lation, including affected key transcription factors, metabolic enzymes, and major oncogenic-signaling pathways [115]. A shunt of glycolysis known as the hexosamine biosynthesis pathway (HBP) links cellular signaling and gene expression to glucose metabolism. The HBP generates a metabolically expensive moiety used for glycosylation, uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc). The β -*N*-acetylglucosaminyltransferase (OGT) enzyme uses UDP-GlcNAc to covalently attach a single O-linked β -*N*-acetylglucosamine (O-GlcNAc) moiety to serine or threonine (S/T) residues within target proteins. During adipocyte differentiation, Tang *et al.* found that CCAAT enhancer-binding protein (C/EBP) beta is modified by O-GlcNAcylation at Ser180 and Ser181, which decreases the relative phosphorylation and DNA binding activity of C/EBPbeta and then delays adipocyte differentiation [116].

Cancer cells exhibit heightened uptake of glucose and glutamine, and rewire the metabolic flux toward anabolic pathways important for cell growth and proliferation. Understanding of how this altered metabolism is regulated has recently emerged as an intense research topic in cancer biology. Yi *et al.* found that phosphofructokinase 1 (PFK1) serine 529 is induced by O-GlcNAcylation in response to hypoxia; glycosylation inhibits PFK1 activity and redirects glucose flux through the pentose phosphate pathway, conferring a selective growth advantage on cancer. This finding reveals a previously uncharacterized mechanism for the regulation of metabolic pathways in cancer patients, and a possible target for therapeutic intervention [117].

The PI3K-Akt signal pathway is essential to cancer development. Geng *et al.* found that O-GlcNAcylation at Akt Thr 305 and Thr 312 inhibit Akt phosphorylation at Thr 308 via disrupting the interaction between Akt and PDK1, which suppresses cell proliferation and migration capabilities [118]. They also found that actin-binding protein cofilin is O-GlcNAcylated at Ser 108. Cofilin O-GlcNAc modification is required for proper localization and breast cancer metastasis [119].

Dysfunctions in Wnt signaling increase β -catenin stability and are associated with several cancers, including colorectal cancer. Lefebvre *et al.* identified 4 O-GlcNAcylation sites at the N-terminus of β -catenin (S23/T40/T41/T112) using ETD-MS/MS. They also found that elevated O-GlcNAcylation in human colon cell lines drastically reduced phosphorylation at T41, which decreased β -catenin/ α -catenin interactions and slowed the development of cancer [120].

Apart from cancer cells, O-GlcNAcylation is present mainly in the brain. Evidence has been

mounting in recent years that connect defects in glucose metabolism in the brain with neurodegenerative diseases [121]. The amount of O-GlcNAcylation of Neurofilament-M and protein phosphatase 2A are decreased in Alzheimer's disease patients, which show compromised neuron pathological function [122,123]. In a recent study, Shen *et al.* found that nitric oxide synthase adaptor (NOS1AP), which plays an important part in glutamate-induced neuronal apoptosis, promotes the modification of O-GlcNAc at Ser 47, Ser 183, Ser 204, Ser 269, and Ser 271. Higher O-GlcNAc of NOS1AP enhances its binding with neuronal nitric oxide synthases and causes glutamate-induced neuronal apoptosis during ischemia [124].

PERSPECTIVES

The past decades have witnessed an explosive growth in advanced techniques for the analysis of glycoproteins and glycans. That said, techniques for successfully analyzing intact glycopeptides are in high demand, but have traditionally been quite difficult to actualize. Gradual developments such as the modern mass spectrometer [125–127] and advanced techniques [128] now allow intact glycopeptide analysis, but also represent other new challenges—glycan attachment changes the properties of glycopeptide considerably, making analysis rather difficult. To this effect, techniques for enriching the glycopeptide and enhancing the ionization of glycopeptides are urgently necessary. In addition to ion dissociation methods for generating glycopeptide fragments, software that can effectively manage mass spectra information and methods of accurately quantifying glycopeptides must also be developed. As research continues on MS-based intact glycopeptide analysis, we believe that clear understanding the mechanisms of important biological processes will be achieved in the near future.

FUNDING

This work received partial financial support from the National Basic Research Program of China (2013CB911201, 2013CB910502, 2011CB910600, 2014CBA020-00), the National Natural Science Foundation of China (31570825, 21227805), and National High-tech R&D Program of China (2012AA020203).

REFERENCES

1. Li Y, Shah P and De Marzo A *et al.* Identification of glycoproteins containing specific glycans using a lectin-chemical method. *Anal Chem* 2015; **87**: 4683–7.

2. Maurice S, Mandi H and Andre D *et al.* Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* 2011; **83**: 2492–9.
3. Liu LT, Zhang Y and Zhang L *et al.* Highly specific revelation of rat serum glycopeptidome by boronic acid-functionalized mesoporous silica. *Anal Chim Acta* 2012; **753**: 64–72.
4. Zhang H, Li XJ and Martin DB. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003; **21**: 660–6.
5. Zhang Y, Kuang M and Zhang LJ *et al.* An accessible protocol for solid-phase extraction of N-linked glycopeptides through reductive amination by amine-functionalized magnetic nanoparticles. *Anal Chem* 2013; **85**: 5535–41.
6. Zhang Y, Yu M and Zhang C *et al.* Highly selective and ultrafast solid-phase extraction of N-glycoproteome by oxime click chemistry using aminoxy-functionalized magnetic nanoparticles. *Anal Chem* 2014; **86**: 7920–4.
7. Zhang Y, Yu M and Zhang C *et al.* Highly specific enrichment of N-glycoproteome through a nonreductive amination reaction using Fe₃O₄@SiO₂-aniline nanoparticles. *Chem Commun* 2015; **51**: 5982–5.
8. Zhang Y, Zhang C and Jiang HC *et al.* Fishing the PTM proteome with chemical approaches using functional solid phases. *Chem Soc Rev* 2015, DOI: 10.1039/C4CS00529E.
9. Hecht ES, McCord JP and Muddiman DC. Definitive screening design optimization of mass spectrometry parameters for sensitive comparison of filter and solid phase extraction purified, INLIGHT plasma N-glycans. *Anal Chem* 2015; **87**: 7305–12.
10. Selman MHJ, Hemayatkar M and Deelder AM *et al.* Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* 2011; **83**: 2492–9.
11. Qin HQ, Zhao L and Li RB *et al.* Size-selective enrichment of N-linked glycans using highly ordered mesoporous carbon material and detection by MALDI-TOF MS. *Anal Chem* 2011; **83**: 7721–8.
12. Qin HQ, Hu ZY and Wang FJ *et al.* Facile preparation of ordered mesoporous composite nanoparticles for glycan enrichment. *Chem Commun* 2013; **49**: 5162–4.
13. Sun NR, Deng CH and Li Y *et al.* Highly selective enrichment of N-linked glycan by carbon-functionalized ordered graphene/mesoporous silica composites. *Anal Chem* 2014; **86**: 2246–50.
14. Sun NR, Zhang XM and Deng CH. Designed synthesis of MOF-derived magnetic nanoporous carbon materials for selective enrichment of glycans for glycomics analysis. *Nanoscale* 2015; **7**: 6487–91.
15. Li LL, Jiao J and Cai Y *et al.* Fluorinated carbon tag derivatization combined with fluorosolid-phase extraction: a new method for the highly sensitive and selective mass spectrometric analysis of glycans. *Anal Chem* 2015; **87**: 5125–31.
16. Zhang WJ, Han HH and Bai HH *et al.* A highly efficient and visualized method for glycan enrichment by self-assembling pyrene derivative functionalized free graphene oxide. *Anal Chem* 2013; **85**: 2703–9.

17. Amano J, Nishikaze T and Tougasaki F *et al.* Derivatization with 1-pyrenyldiazomethane enhances ionization of glycopeptides but not peptides in matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 2010; **82**: 8738–43.
18. Nishikaze T, Kawabata S and Tanaka K. In-depth structural characterization of N-linked glycopeptides using complete derivatization for carboxyl groups followed by positive- and negative-ion tandem mass spectrometry. *Anal Chem* 2014; **86**: 5360–9.
19. Haan N, Reiding KR and Habberger M *et al.* Linkage-specific sialic acid derivatization for MALDI-TOF-MS profiling of IgG glycopeptides. *Anal Chem* 2015, DOI: 10.1021/acs.analchem.5b02426.
20. Ruhaak LR, Zauner G and Huhn C *et al.* Glycan labeling strategies and their use in identification and quantification. *Anal Bioanal Chem* 2010; **397**: 3457–81.
21. Cai Y, Zhang Y and Yang P *et al.* Improved analysis of oligosaccharides for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using aminopyrazine as a derivatization reagent and a co-matrix. *Analyst* 2013; **138**: 6270–6.
22. Lattová E and Perreault H. The usefulness of hydrazine derivatives for mass spectrometric analysis of carbohydrates. *Mass Spectrom Rev* 2013; **32**: 366–85.
23. Jiao J, Yang L and Zhang Y *et al.* Hydrazinonicotinic acid derivatization for selective ionization and improved glycan structure characterization by MALDI-MS. *Analyst* 2015; **140**: 5475–80.
24. Zaia J. Mass spectrometry of oligosaccharides. *Mass Spectrom Rev* 2004; **23**: 161–227.
25. Guillard M, Gloerich J and Wessels HJ *et al.* Automated measurement of permethylated serum N-glycans by MALDI-linear ion trap mass spectrometry. *Carbohydr Res* 2009; **344**: 1550–7.
26. Prien JM, Ashline DJ and Lapadula AJ *et al.* The high mannose glycans from bovine ribonuclease B isomer characterization by ion trap MS. *J Am Soc Mass Spectrom* 2009; **20**: 539–56.
27. Cheng PF, Snovidá S and Ho MY *et al.* Increasing the depth of mass spectrometry-based glycomic coverage by additional dimensions of sulfoglycomics and target analysis of permethylated glycans. *Anal Bioanal Chem* 2013; **405**: 6683–95.
28. Cheng CW, Chou CC and Hsieh HW *et al.* Efficient mapping of sulfated glycotopes by negative ion mode nanoLC-MS/MS-based sulfoglycomic analysis of permethylated glycans. *Anal Chem* 2015; **87**: 6380–8.
29. Nie H, Li Y and Sun XL. Recent advances in sialic acid-focused glycomics. *J Proteomics* 2012; **75**: 3098–112.
30. Liu X, Qiu H and Lee RK *et al.* Methylamidation for sialoglycomics by MALDI-MS: a facile derivatization strategy for both α 2, 3- and α 2, 6-linked sialic acids. *Anal Chem* 2010; **82**: 8300–6.
31. Zhang Q, Feng X and Li H *et al.* Methylamidation for isomeric profiling of sialylated glycans by NanoLC-MS. *Anal Chem* 2014; **86**: 7913–9.
32. Kaneshiro K, Fukuyama Y and Iwamoto S *et al.* Highly sensitive MALDI analyses of glycans by a new aminoquinoline-labeling method using 3-aminoquinoline/ α -cyano-4-hydroxycinnamic acid liquid matrix. *Anal Chem* 2011; **83**: 3663–7.
33. Fukuyama Y, Funakoshi N and Takeyama K *et al.* 3-Aminoquinoline/p-coumaric acid as a MALDI matrix for glycopeptides, carbohydrates, and phosphopeptides. *Anal Chem* 2014; **86**: 1937–42.
34. Selman MH, Hoffmann M and Zauner G *et al.* MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro- α -cyanocinnamic acid matrix. *Proteomics* 2012; **12**: 1337–48.
35. Nishikaze T, Fukuyama Y and Kawabata S *et al.* Sensitive analyses of neutral N-glycans using anion-doped liquid matrix G3CA by negative-ion matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 2012; **84**: 6097–103.
36. Jiao J, Zhang Y and Yang PY *et al.* Hydrazinonicotinic acid as a novel matrix for highly sensitive and selective MALDI-MS analysis of oligosaccharides. *Analyst* 2015; **140**: 156–61.
37. Wu CL, Wang CC and Lai YH *et al.* Selective enhancement of carbohydrate ion abundances by diamond nanoparticles for mass spectrometric analysis. *Anal Chem* 2013; **85**: 3836–41.
38. Liang Q, Macher T and Xu Y *et al.* MALDI MS in-source decay of glycans using a glutathione-capped iron oxide nanoparticle matrix. *Anal Chem* 2014; **86**: 8496–503.
39. Kuang M, Zhang Y and Yang PY *et al.* Novel ionic matrices for enhanced ionization of oligosaccharides/glycopeptides during MALDI-MS analysis. *Acta Chim Sinica* 2013; **71**: 1007–10.
40. Dodds ED. Gas-phase dissociation of glycosylated peptide ions. *Mass Spectrom Rev* 2012; **31**: 666–82.
41. McLuckey SA and Mentinova M. Ion/neutral, ion/electron, ion/photon, and ion/ion interactions in tandem mass spectrometry: do we need them all? Are they enough? *J Am Soc Mass Spectrom* 2011; **22**: 3–12.
42. Harvey DJ. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2009–2010. *Mass Spectrom Rev* 2015; **34**: 268–422.
43. Both P, Green AP and Gray CJ *et al.* Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. *Nat Chem* 2014; **6**: 65–74.
44. Novotny MV and Alley WJ. Recent trends in analytical and structural glyco-biology. *Curr Opin Chem Biol* 2013; **17**: 832–40.
45. Zhong X, Chen Z and Snovidá S *et al.* Capillary electrophoresis-electrospray ionization-mass spectrometry for quantitative analysis of glycans labeled with multiplex carbonyl-reactive tandem mass tags. *Anal Chem* 2015; **87**: 6527–34.
46. Olsen JV, Macek B and Lange O *et al.* Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 2007; **4**: 709–12.
47. Segu ZM and Mechref Y. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun Mass Spectrom* 2010; **24**: 1217–25.
48. Cao Q, Zhao X and Zhao Q *et al.* Strategy integrating stepped fragmentation and glycan diagnostic ion-based spectrum refinement for the identification of core fucosylated glycoproteome using mass spectrometry. *Anal Chem* 2014; **86**: 6804–11.
49. Zubarev RA, Kelleher NL and McLafferty FW. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J Am Chem Soc* 1998; **120**: 3265–6.
50. Coon JJ, Syka J and Schwartz JC *et al.* Anion dependence in the partitioning between proton and electron transfer in ion/ion reactions. *Int J Mass Spectrom* 2004; **236**: 33–42.
51. Chen R, Wang F and Tan Y *et al.* Development of a combined chemical and enzymatic approach for the mass spectrometric identification and quantification of aberrant N-glycosylation. *J Proteomics* 2012; **75**: 1666–74.
52. Singh C, Zampronio CG and Creese AJ *et al.* Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins. *J Proteome Res* 2012; **11**: 4517–25.
53. Lavery SB, Steentoft C and Halim A *et al.* Advances in mass spectrometry driven O-glycoproteomics. *Biochim Biophys Acta* 2015; **1850**: 33–42.

54. Steentoft C, Vakhrushev SY and Vester-Christensen MB *et al.* Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nat Methods* 2011; **8**: 977–82.
55. Halim A, Ruetschi U and Larson G *et al.* LC-MS/MS characterization of O-glycosylation sites and glycan structures of human cerebrospinal fluid glycoproteins. *J Proteome Res* 2013; **12**: 573–84.
56. An HJ and Lebrilla CB. Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry. *Mass Spectrom Rev* 2011; **30**: 560–78.
57. Li B, Russell SC and Zhang J *et al.* Structure determination by MALDI-IRMPD mass spectrometry and exoglycosidase digestions of O-linked oligosaccharides from *Xenopus borealis* egg jelly. *Glycobiology* 2011; **21**: 877–94.
58. Seipert RR, Dodds ED and Clowers BH *et al.* Factors that influence fragmentation behavior of N-linked glycopeptide ions. *Anal Chem* 2008; **80**: 3684–92.
59. Fukui K and Takahashi K. Infrared multiple photon dissociation spectroscopy and computational studies of O-glycosylated peptides. *Anal Chem* 2012; **84**: 2188–94.
60. Madsen JA, Ko BJ and Xu H *et al.* Concurrent automated sequencing of the glycan and peptide portions of O-linked glycopeptide anions by ultraviolet photodissociation mass spectrometry. *Anal Chem* 2013; **85**: 9253–61.
61. Woodin CL, Maxon M and Desaire H *et al.* Software for automated interpretation of mass spectrometry data from glycans and glycopeptides. *Analyst* 2013; **138**: 2793–803.
62. Alley WR, Jr, Mann BF and Novotny MV. High-sensitivity analytical approaches for the structural characterization of glycoproteins. *Chem Rev* 2013; **113**: 2668–732.
63. Ceroni A, Maass K and Geyer H *et al.* GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 2008; **7**: 1650–9.
64. Goldberg D, Bern M and Parry S *et al.* Automated N-glycopeptide identification using a combination of single- and tandem-ms. *J Proteome Res* 2007; **6**: 3995–4005.
65. Deshpande N, Jensen PH and Packer NH *et al.* GlycoSpectrumScan fishing glycopeptides from MS spectra of protease digests of human colostrum slgA. *J Proteome Res* 2010; **9**: 1063–75.
66. Kailemia MJ, Ruhaak L and Lebrilla CB *et al.* Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Anal Chem* 2014; **86**: 196–212.
67. Wu S, Salcedo J and Tang N *et al.* Employment of tandem mass spectrometry for the accurate and specific identification of oligosaccharide structures. *Anal Chem* 2012; **84**: 7456–62.
68. Hayes CA, Karlsson NG and StruWe WB *et al.* UniCarb-DB: a database resource for glycomic discovery. *Bioinformatics* 2011; **27**: 1343–4.
69. Zielinska DF, Gnad F and Wiśniewski JR *et al.* Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 2010; **141**: 897–907.
70. Wu Y, Mechref Y and Klouckova I *et al.* Mapping site-specific protein N-glycosylations through liquid chromatography/mass spectrometry and targeted tandem mass spectrometry. *Rapid Commun Mass Spectrom* 2010; **24**: 965–72.
71. Mayampurath A, Yu CY and Song E *et al.* Computational framework for identification of intact glycopeptides in complex samples. *Anal Chem* 2014; **86**: 453–63.
72. Woodin CL, Hua D and Maxon M *et al.* GlycoPep grader: a web-based utility for assigning the composition of N-linked glycopeptides. *Anal Chem* 2012; **84**: 4821–9.
73. Zhu Z, Hua D and Clark DF *et al.* GlycoPep detector: a tool for assigning mass spectrometry data of N-linked glycopeptides on the basis of their electron transfer dissociation spectra. *Anal Chem* 2013; **85**: 5023–32.
74. Liu M, Zhang Y and Chen Y *et al.* Efficient and accurate glycopeptide identification pipeline for high-throughput site-specific N-glycosylation analysis. *J Proteome Res* 2014; **13**: 3121–9.
75. Cheng K, Chen R and Seebun D *et al.* Large-scale characterization of intact N-glycopeptides using an automated glycoproteomic method. *J Proteomics* 2014; **110**: 145–54.
76. He L, Xin L and Shan B *et al.* GlycoMaster DB: software to assist the automated identification of N-linked glycopeptides by tandem mass spectrometry. *J Proteome Res* 2014; **13**: 3881–95.
77. Wu SW, Pu TH and Viner R *et al.* Novel LC-MS2 product dependent parallel data acquisition function and data analysis workflow for sequencing and identification of intact glycopeptides. *Anal Chem* 2014; **86**: 5478–86.
78. Jia W, Lu Z and Fu Y *et al.* A strategy for precise and large scale identification of core fucosylated glycoproteins. *Mol Cell Proteomics* 2009; **8**: 913–23.
79. Dallas DC, Martin WF and Hua S *et al.* Automated glycopeptide analysis—review of current state and future directions. *Brief Bioinform* 2013; **14**: 361–74.
80. Zhang Y, Yin HR and Lu HJ. Recent progress in quantitative glycoproteomics. *Glycoconj J* 2012; **29**: 249–58.
81. Chen X, Sun L and Yu Y *et al.* Amino acid-coded tagging approaches in quantitative proteomics. *Expert Rev Proteomics* 2007; **4**: 25–37.
82. Ong SE, Blagoev B and Kratchmarova I *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002; **5**: 376–86.
83. Boersema PJ, Geiger T and Wisniewski JR *et al.* Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. *Mol Cell Proteomics* 2013; **12**: 158–71.
84. Sun Z, Qin H and Wang F *et al.* Capture and dimethyl labeling of glycopeptides on hydrazide beads for quantitative glycoproteomics analysis. *Anal Chem* 2012; **84**: 8452–6.
85. Weng Y, Qu Y and Jiang H. An integrated sample pretreatment platform for quantitative N-glycoproteome analysis with combination of on-line glycopeptide enrichment, deglycosylation and dimethyl labeling. *Anal Chim Acta* 2014; **833**: 1–8.
86. Liu Z, Cao J and He Y *et al.* Tandem ¹⁸O stable isotope labeling for quantification of N-glycoproteome. *J Proteome Res* 2010; **9**: 227–36.
87. Zhang W, Cao WQ and Huang JM *et al.* PNGase F-mediated incorporation of ¹⁸O into glycans for relative glycan quantitation. *Analyst* 2015; **140**: 1082–9.
88. Kim YJ, Zaidi-Ainouch Z and Gallien S *et al.* Mass spectrometry-based detection and quantification of plasma glycoproteins using selective reaction monitoring. *Nat Protoc* 2012; **7**: 859–71.
89. Rexach JE, Rogers CJ and Yu SH *et al.* Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags. *Nat Chem Biol* 2010; **6**: 645–51.
90. Sun S and Zhang H. Large-scale measurement of absolute protein glycosylation stoichiometry. *Anal Chem* 2015; **87**: 6479–82.
91. Orlando R, Lim JM and Atwood JA, III *et al.* IDAWG: metabolic incorporation of stable isotope labels for quantitative glycomics of cultured cells. *J Proteome Res* 2009; **8**: 3816–23.
92. Hu Y, Desantos-Garcia JL and Mechref Y. Comparative glycomic profiling of isotopically permethylated N-glycans by liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom* 2013; **27**: 865–77.

93. Zhang P, Zhang Y and Xue XD *et al.* Relative quantitation of glycans using stable isotopic labels 1-(d0/d5) phenyl-3-methyl-5-pyrazolone by mass spectrometry. *Anal Biochem* 2011; **418**: 1–9.
94. Atwood JA, Cheng L and Alvarez-Manilla G *et al.* Quantitation by isobaric labeling: applications to glycomics. *J Proteome Res* 2008; **7**: 367–74.
95. Cai Y, Jiao J and Bin ZC *et al.* Glycan reductive isotope-coded amino acid labeling (GRIAL) for mass spectrometry-based quantitative N-glycomics. *Chem Commun (Camb)* 2015; **51**: 772–5.
96. Zhang W, Wang H and Tang H *et al.* Endoglycosidase-mediated incorporation of ¹⁸O into glycans for relative glycan quantitation. *Anal Chem* 2011; **83**: 4975–81.
97. Cao WQ, Zhang W and Huang JM *et al.* Glycan reducing end dual isotopic labeling (GREDIL) for mass spectrometry-based quantitative N-glycomics. *Chem Commun* 2015, DOI: 10.1039/C5CC05365J.
98. Gil GC, Kim YG and Kim BG. A relative and absolute quantification of neutral N-linked oligosaccharides using modification with carboxymethyl trimethylammonium hydrazide and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal Biochem* 2008; **379**: 45–59.
99. Tian Y and Zhang H. Glycoproteomics and clinical applications. *Proteomics Clin Appl* 2010; **4**: 124–32.
100. Liu Y, He J and Li C *et al.* Identification and confirmation of biomarkers using an integrated platform for quantitative analysis of glycoproteins and their glycosylations. *J Proteome Res* 2009; **1265**: 798–805.
101. Cao J, Shen C and Wang H *et al.* Identification of N-glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach. *J Proteome Res* 2009; **8**: 662–72.
102. Hu H, Gao L and Wang C *et al.* Lower serum soluble-EGFR is a potential biomarker for metastasis of HCC demonstrated by N-glycoproteomic analysis. *Discov Med* 2015; **19**: 333–41.
103. Qian Y, Wang Y and Zhang X *et al.* Quantitative analysis of serum IgG galactosylation assists differential diagnosis of ovarian cancer. *J Proteome Res* 2013; **12**: 4046–55.
104. Yen TY, Haste N and Timpe LC. Using a cell line breast cancer progression system to identify biomarker candidates. *J Proteomics* 2014; **96**: 173–83.
105. Hill JJ, Tremblay TL and Fauteux F *et al.* Glycoproteomic comparison of clinical triple-negative and luminal breast tumors. *J Proteome Res* 2015; **14**: 1376–88.
106. Tan Z, Yin H and Nie S *et al.* Large-scale identification of core-fucosylated glycopeptide sites in pancreatic cancer serum using mass spectrometry. *J Proteome Res* 2015; **14**: 1968–78.
107. Hang Q, Zhou Y and Hou S *et al.* Asparagine-linked glycosylation of bone morphogenetic protein-2 is required for secretion and osteoblast differentiation. *Glycobiology* 2014; **24**: 292–304.
108. Liu Y, Ren S and Xie L *et al.* Mutation of N-linked glycosylation at Asn548 in CD133 decreases its ability to promote hepatoma cell growth. *Oncotarget* 2015, PMID: 26029999.
109. Carvalho S, Catarino TA and Dias AM *et al.* Preventing E-cadherin aberrant N-glycosylation at Asn-554 improves its critical function in gastric cancer. *Oncogene* 2015, DOI: 10.1038/onc.2015.225.
110. Zhang X, Wang Y and Qian Y *et al.* Discovery of specific metastasis-related N-glycan alterations in epithelial ovarian cancer based on quantitative glycomics. *PLoS One* 2014; **9**: e87978.
111. Chik JH, Zhou J and Moh ES *et al.* Comprehensive glycomics comparison between colon cancer cell cultures and tumours: implications for biomarker studies. *J Proteomics* 2014; **108**: 146–62.
112. Xie J, Zhu H and Guo L *et al.* Lectin-like oxidized low-density lipoprotein receptor-1 delivers heat shock protein 60-fused antigen into the MHC class I presentation pathway. *J Immunol* 2010; **185**: 2306–13.
113. Zhou L, Qian Y and Zhang X *et al.* Elucidation of differences in N-glycosylation between different molecular weight forms of recombinant CLEC-2 by LC MALDI tandem MS. *Carbohydr Res* 2015; **402**: 180–8.
114. Li W, Yu R and Ma B *et al.* Core fucosylation of IgG B cell receptor is required for antigen recognition and antibody production. *J Immunol* 2015; **194**: 2596–606.
115. Hart GW, Slawson C and Ramirez-Correa G *et al.* Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem* 2011; **80**: 825–58.
116. Li X, Molina H and Huang H *et al.* O-linked N-acetylglucosamine modification on CCAAT enhancer-binding protein beta: role during adipocyte differentiation. *J Biol Chem* 2009; **284**: 19248–54.
117. Yi W, Clark PM and Mason DE *et al.* Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. *Science* 2012; **3**: 975–80.
118. Wang S, Huang X and Sun D *et al.* Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates Akt signaling. *PLoS One* 2012; **7**: e37427.
119. Huang X, Pan Q and Sun D *et al.* O-GlcNAcylation of cofilin promotes breast cancer cell invasion. *J Biol Chem* 2013; **288**: 36418–25.
120. Olivier-Van Stichelen S, Dehennaut V and Buzy A *et al.* O-GlcNAcylation stabilizes β -catenin through direct competition with phosphorylation at threonine 41. *FASEB J* 2014; **28**: 3325–38.
121. Dias WB and Hart GW. O-GlcNAc modification in diabetes and Alzheimer's disease. *Mol Biosyst* 2007; **3**: 766–72.
122. Deng Y, Li B and Liu F *et al.* Regulation between O-GlcNAcylation and phosphorylation of neurofilament-M and their dysregulation in Alzheimer disease. *FASEB J* 2008; **22**: 138–45.
123. Liu F, Shi J and Tanimukai H *et al.* Reduced O-GlcNAcylation links lower brain glucose metabolism and tau pathology in Alzheimer's disease. *Brain* 2009; **132**: 1820–32.
124. Zhu L, Tao T and Zhang D *et al.* NOS1AP O-GlcNAc modification involved in neuron apoptosis induced by excitotoxicity. *Int J Mol Sci* 2015; **16**: 16560–75.
125. Makarov A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* 2000; **72**: 1156–62.
126. Dodonov AF, Kozlovski VI and Soulimenkov IV *et al.* High-resolution electrospray ionization orthogonal-injection time-of-flight mass spectrometer. *Eur J Mass Spectrom* 2000; **6**: 481–90.
127. Douglas DJ, Frank AJ and Mao DM. Liner ion traps in mass spectrometry. *Mass Spectrom Rev* 2005; **24**: 1–29.
128. Nilsson T, Mann M and Aebersold R *et al.* Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods* 2010; **7**: 681–5.