Glycoconj J (2012) 29:249–258
DOI 10.1007/s10719-012-9398-x

Recent progress in quantitative glycoproteomics

Ying Zhang · Hongrui Yin · Haojie Lu

Received: 15 March 2012 / Revised: 17 May 2012 / Accepted: 21 May 2012 / Published online: 15 June 2012

Abstract Protein glycosylation is acknowledged as one of the major posttranslational modifications that elicit significant effects on protein folding, conformation, distribution, stability, and activity. The changes in glycoprotein abundance, glycosylation degree, and glycan structure are associated with a variety of diseases. Therefore, the quantitative study of glycoproteomics has become a new and popular research topic, and is quickly emerging as an important technique for biomarker discovery. Mass spectrometry-based protein quantification technologies provide a powerful tool for the systematic and quantitative assessment of the quantitative differences in the protein profiles of different samples. Combined with various glycoprotein/glycopeptide enrichment strategies and other glycoprotein analysis methods, these techniques have been further developed for application in quantitative glycoproteomics. A comprehensive quantitative analysis of the glycoproteome in a complex biological sample remains challenging because of the enormous complexity of biological samples, intrinsic characteristics of glycoproteins, and lack of universal quantitative technology. In this review, recently developed technologies in quantitative glycoproteome, especially those focused on two of the most common types of glycosylation (N-linked and O-linked glycoproteome), were summarized. The strengths and weaknesses of the various approaches were also discussed.

Y. Zhang · H. Yin · H. Lu (✉)
Department of Chemistry and Institutes of Biomedical Sciences,
Fudan University,
Shanghai 200032, China
e-mail: luhaojie@fudan.edu.cn

Keywords Glycosylation · Proteomics · Quantitation · Mass Spectrometry · Enrichment

Glycosylation is one of the most frequent and complex protein posttranslational modifications. More than 50 % of proteins are predicted to be glycosylated in mammals [1]. The two most common types of glycosylation are N-linked and O-linked. In N-linked glycoproteins, the carbohydrates are linked to asparagine residues that fall into the N-X-S/T sequence motif (X denotes any amino acid except proline). In O-linked glycoproteins, the carbohydrates are linked to serine or threonine residues. Some other forms of glycosylation such as C-linked [2] and S-linked have also been reported [3, 4]. In this review, we will limit our discussion to the two most common types of glycosylation, N- and O-glycosylation.

Protein glycosylation plays an important role in many biological processes such as embryonic development, cell adhesion and signal transduction, molecular trafficking, immune response, and so on [5]. In the past years, aberrant glycosylation has been reported in many diseases. In addition, many glycoproteins have been identified as cancer biomarkers, such as prostate-specific antigen (PSA) in prostate cancer [6], Cancer antigen 125 (CA-125) in ovarian cancer [7], human epidermal growth factor receptor-2 (Her-2/neu) in breast cancer [8], and α-fetoprotein (AFP) in liver cancer [9]. The changes in glycoprotein abundance, glycosylation degree, and glycan structure have been shown to be closely associated with diseases. Therefore, quantitative glycoproteomics has become a popular research area. Monitoring the changes in glycosylation is important for the diagnosis, prognosis, and understanding molecular mechanisms involved in pathogenesis [10]. Techniques based on mass spectrometry (MS) have been widely used in glycoproteomics

research [11]. However, previous studies mainly focused on the enrichment and isolation of glycoproteins/glycopeptides, identification of glycoproteins and their glycosylation sites, and glycan structure analysis. To date, the methodology for the systematic quantitation of glycoproteins and their glycosylation site occupancy remains in its infancy.

In theory, the quantitative techniques used in proteomics are also applicable to glycoproteomics. However, several new difficulties due to the inherent characteristics of glycoproteins need to be addressed. First, glycopeptides have relatively low abundance (2 % to 5 %) and ionization efficiency is much lower compared with nonglycopeptides. Consequently, nonglycopeptides always significantly suppress the mass spectrometric response of glycopeptides. Therefore, highly efficient enrichment strategies are required to isolate glycoproteins/glycopeptides prior to MS analysis. Second, complexity arises from the microheterogeneity of glycans, making glycoproteins typically difficult to be characterized through MS than the mass spectrometric analysis of proteins. Third, glycosylation degree is not always consistent with glycoprotein abundance. Quantitative methods should be capable of quantifying glycoprotein expression and glycosylation levels simultaneously.

The quantitative techniques used in glycoproteomics can be grouped into two main categories: relative quantitation and absolute quantitation of target proteins (Fig. 1). The former may be with or without stable isotope labeling. This review mainly discusses the recent developments in quantitative glycoproteomics, particularly addressing research on N-linked and O-linked glycoproteins. Table 1 shows the comparison between the advantages and disadvantages of each method.

Relative quantitation

MS-based relative quantitation can be obtained through label-based or label-free approaches (Fig. 1). The former is based on the introduction of a stable isotopic label into proteins or peptides. The labels change the mass of a protein or peptide without affecting the analytical or biochemical properties. Through MS analysis, the protein level can be obtained by comparing the abundance of the differently labeled peptides. The isotopic label can be introduced into the sample at different stages of sample preparation in various manners, including isotopic labeling through chemical reactions, metabolic incorporation of stable isotope labels, and introduction of isotope labels via enzymatic reactions. In label-free quantitation strategies, protein profiling comparisons are based on either mass spectrometric signal intensities of peptides or the number of MS/MS spectra matched to peptides and proteins (spectral counting). Changes in the glycopeptide ratio among different samples can reflect changes in the glycosylation occupancy of individual glycosylation sites. Changes in the nonglycopeptide ratio can reflect changes in glycoprotein abundance at the protein level.

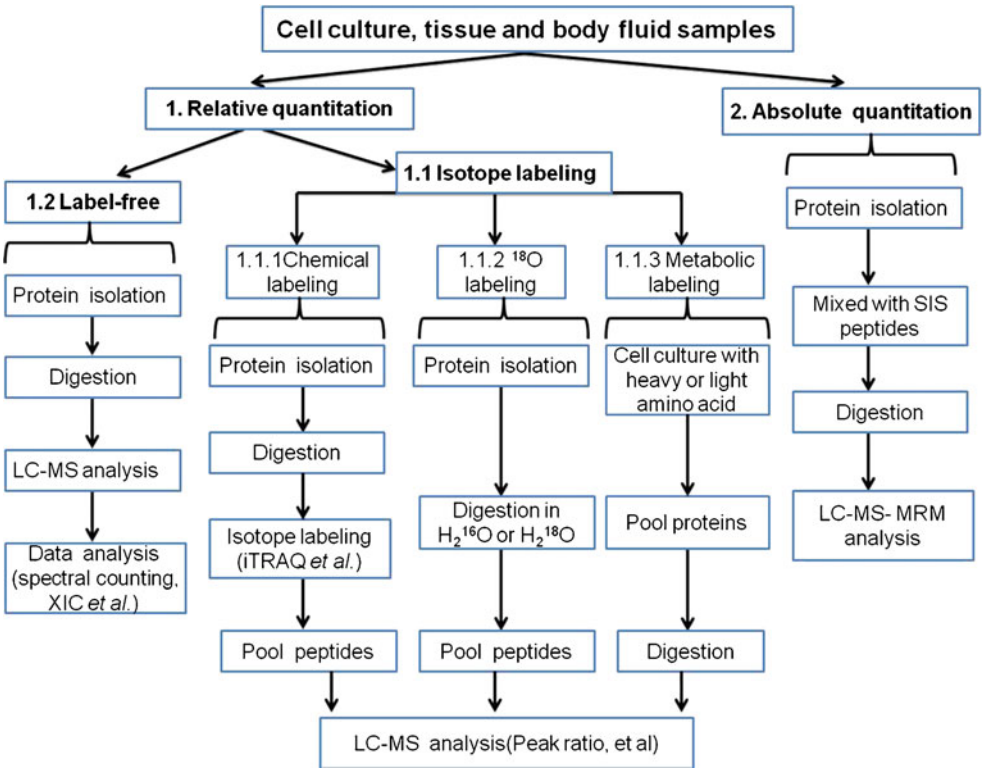

Fig. 1 The most commonly used strategies for quantitative analysis of glycoproteomics

Table 1 Comparison of different methods used in quantitative analysis of glycoproteomics.

Method			Key points	Advantages	Disadvantages
Relative Quantitation	Isotope labeling	Chemical isotope labeling	Labeling isotopic tags on the active groups(such as, -COOH,-NH_2 and so on) of proteins or peptides.	Can be applied to cells in culture, tissues or fluid sample. Many kinds of Labeling reagents can be selected.	Labeling reagents are always expensive. Incomplete labeling
		Proteolytic ^{18}O labeling	During the enzymatic digestion, ^{18}O was incorporated into the cleavage sites.	Residue-specific catalytic method. Reagents are cost-effective. The reaction conditions are mild. Compatible with all peptide-level fractionation methods.	Incomplete labeling Back exchange Can not used for the quantitation of O-linked glycoproteins, because of lacking suitable enzymes for releasing O-glycans.
		Metabolic labeling	Culturing cells with isotope labeled amino acids.	Labeling in the cell culture, reducing the uneven losses of protein or peptide during sample preparation. The labeling can be achieved to 100 %.	Expensive Restricted to cell cultures
	Label-free		Comparing the mass signals directly using spectral counting or extracted ion chromatogram (XIC).	Expensive isotope labeling reagents and extra sample preparation steps are unnecessary. No limitation to the number of samples for comparision.	Run-to-run variability can affect the accuracy of the result, needing normalization techniques. Only reflect the glycoproteins abundance in protein expression level, not the glycosylation site occupancy.
Absolute quantitation			Incorporating stable isotope-labeled standard (SIS) peptides into the sample mixture.	The absolute concentration can be known.	SIS peptides need to be selected and synthesized, which is expensive.

Isotopic label-based relative quantitation

Chemical labeling

In 2003, Zhang *et al*. [12] reported a method for large-scale quantitative glycoprotein profiling. To enrich the glycoproteins, hydrazide chemistry was first used to conjugate the glycoproteins to a solid support. The captured glycoproteins were proteolyzed, and the N-termini of glycopeptides were labeled with succinic anhydride carrying either d0 or d4. The N-linked glycosylated peptides were released from the solid phase via peptide-*N*-glycosidase F (PNGase F) treatment. The recovered peptides were then identified and quantified via MS. In a proof-of-principle experiment, the observed ratios were generally close to the expected ratios, and the differences between the observed and expected ratio ranged between 0 % and 29 %, with a mean of 8 %. This result indicated that the glycoprotein capture method combined with stable isotope tagging could allow reasonable quantification of the glycoproteome. The high glycoprotein selectivity of this method reduced the sample complexity, and a 4 Da mass difference avoided the overlapping of light and heavy isotopic envelops. Hence, highly selective enrichment and accurate quantitation are two of this method's strong features. However, this method was only focused on the analysis of N-linked glycosylation sites. The changes in glycoprotein expression level were not shown. In addition, the identification of glycosylation site was based on a 0.98 Da mass change when asparagine residues changed into aspartic acid residues via the PNGase F treatment, which was not sufficiently accurate. Zhou *et al*. [13] also employed hydrazide chemistry-based solid extraction to enrich the glycoproteins. The isobaric tags for relative and absolute quantitation (iTRAQ) techniques were used to quantify the glycoproteins. This method was used to investigate the tear fluid of patients with climatic droplet keratopathy in comparison to that of normal controls. Five N-linked glycoproteins were found to have significant changes in the N-glycosylation levels, without changes in the overall level of these glycoproteins.

Another method used to enrich glycoproteins is lectin affinity chromatography. For example, Qiu *et al*. [14] used Sambucus nigra agglutinin (SNA) to enrich glycoproteins containing sialic acid. First,the peptides from the control and experimental samples were acetylated at the N-terminus with N-acetoxysuccinamide and deuterated *N*-acetoxysuccinamide, and then mixed. Glycopeptides were selected via lectin affinity chromatography and deglycosylated via PNGase F, followed by electrospray ionization-MS (ESI-MS) analysis. The concentration changes in the sialic acid containing glycoproteins between the control and experimental sample were obtained by comparing their isotope ratios. In addition, they used concanavalin A (Con A) coupled to SNA to capture sialic acid-containing glycopeptides. Subsequently,

they used Con A alone to capture glycopeptides without sialic acid but with complex biantennary N-linked, hybrid, and high-mannose glycans. The isotope ratio between these two groups, which indicated the relative degree of sialylation of one glycopeptide, was obtained using MS. In this method, they used glycopeptides for quantitation, which just reflected the changes in the degree of glycosylation, whereas the glycoprotein changes in the protein expression level could not be obtained. In addition, this group described a similar method based on serial lectin affinity chromatography for fractionation and comparison of glycan site heterogeneity on glycoproteins in human serum [15]. The serum sample was digested, and the sialic acid-containing glycopeptides were selected using an SNA affinity column. After elution from the SNA affinity column, the sialylated glycopeptides with biantennary complex-type glycans were captured through a Con A column, whereas the glycopeptides with tri- and tetraantennary complex-type oligosaccharides passed the Con A column. The glycopeptides with the two different glycan forms were collected individually and labeled with the light (d0) and heavy (d3) forms of *N*-acetoxysuccinimide, mixed and deglycosylated through PNGase F, and finally analyzed via liquid chromatography (LC)-MS. The isotope ratio indicated the relative abundance of the different glycan forms of one peptide. Ueda *et al*. [16] established a system containing the lectin enrichment step to analyze lung cancer and healthy control serum glycoproteins. The serum samples were first immunodepleted of six high-abundance proteins and then divided into two equal parts. One part was processed by Lens culinaris (LCA) lectin column chromatography to enrich glycoproteins containing α(1-6)-fucosylation. Quantitative analysis was performed via $^{12}C_6$- or $^{13}C_6$-2-nitrobenzenesulfenyl (NBS) stable isotope labeling followed by matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight mass (MALDI-QIT) spectrometric analysis. The obtained MS intensity ratio was designated as $^{13}C/^{12}C_{LCA}$. Another part was also directly used for quantitative analysis via $^{12}C_6$-or $^{13}C_6$-NBS stable isotope labeling, and the MS intensity ratio was designated as $^{13}C/^{12}C_{pre}$. The alteration of α-1,6-fucosylation level was thereby calculated as $(^{13}C/^{12}C_{LCA})/(^{13}C/^{12}C_{Pre})$. Approximately 34 serum glycoproteins were identified to reveal a significant difference in the α-1,6-fucosylation level between lung cancer and the healthy control. In addition, the abundance of glycoproteins in the protein expression level could be obtained from the ratio of $^{13}C/^{12}C_{pre}$. Ueda *et al*. [17] also reported a new approach for the identification of carbohydrate-targeting serum biomarkers. They enriched the glycopeptides through lectin column chromatography and site-directed labeled the N-glycosylation sites in $H_2{}^{18}O$ during the elution with N-glycosidase. They termed this method isotopic glycosidase elution and labeling on lectin-column chromatography (IGEL). Ueda *et al*. combined IGEL with eight-plex iTRAQ labeling for relative quantification of glycopeptides and applied them to search for carbohydrate-targeting serum biomarkers in lung cancer patient sera. In this approach, nonglycopeptides were used to quantify the glycoproteins in the protein expression level, whereas glycopeptides enriched by lectins were used to quantify the glycosylation degree. The most significant advantage of this technology is that it allows labeling of up to eight different samples within a single experiment. This method is useful for quantifying proteins from multiplex samples, such as those in a time course study. In addition, ^{18}O was incorporated into the glycosylation site, which was beneficial for the identification of the glycosylation site. Shetty *et al*. [18] developed a lectin-directed tandem labeling quantitative proteomics strategy. The sialylated glycopeptides were enriched using SNA, and their N-termini were labeled using acetic anhydride reagents. The different labeled peptides were mixed and deglycosylated with PNGase F in $H_2{}^{18}O$ and finally analyzed through LC-MS/MS. Using this method, the sialylation changes in prostate cancer serum samples were compared with those in healthy controls. A total of 45 sialylated glycopeptides were found, and most of the glycoproteins in prostate cancer serum samples showed an increase in sialylation level. The mass difference between differentially labeled peptides was only 3 Da. Hence, the overlapping of light and heavy isotopic envelops could introduce small errors in the quantitation. Therefore, they considered that sialylation was increased in a peptide only if the heavy/light ratio of that peptide was≥2 fold.

Kurogochi *et al*. [19] devised a method called reverse glycoblotting for the selective enrichment of sialylated glycopeptides from a tryptic peptide mixture using glycoblotting technology in combination with site-specific oxidation of terminal sialic acid residues. Chen *et al*. [20] applied this method using reverse glycoblotting to enrich sialylated glycopeptides. The glycopeptides were cleaved using endoglycosidase F3. Finally, the deglycosylated peptides were labeled with dimethyl reagents and analyzed via MS with both neutral-loss-triggered MS^3 in collision-induced dissociation and electron transfer dissociation. Using reverse glycoblotting, N-glycosylation sites with both terminal sialylation and core fucosylation were specifically enriched. Up to 69 aberrant N-glycosylation sites were identified in serum samples from hepatocellular carcinoma (HCC) patients. Chen *et al*. also compared HCC and liver cirrhosis sera and found six N-glycosites in three glycoproteins with increased aberrancy.

Recently, we [21] have developed a technique termed iTRAQ plus ^{18}O combined with antibody purification to quantify the target glycoprotein. Using this technique, iTRAQ Reagent 114- and iTRAQ Reagent 116-labeled samples were mixed and treated with PNGase F in $H_2{}^{16}O$, whereas 115-labeled and 116-labeled samples were mixed and treated with PNGase F in $H_2{}^{18}O$. All samples were then mixed together and analyzed through LC-MS/MS. The

nonglycopeptides of glycoproteins were labeled only with iTRAQ reagents, and glycopeptides were both labeled with iTRAQ and $^{16}O/^{18}O$. Therefore, the peak area ratio of reporter ions from nonglycopeptides reflected the changes in glycoprotein expression level. The peak areas ratio of reporter ions from glycopeptides was indicative of the relative amount of glycopeptides. The glycosite ratio was thereby calculated by dividing the abundance ratio of glycopeptides by the abundance ratio of protein. For glycopeptides with one N-glycosylation sites, ^{16}O and ^{18}O labeling introduced a mass difference of 2 Da, which could support the identification of glycosylated sites. Therefore, our method afforded elevated specificity in glycopeptide identification in parallel with quantitation analysis. This method was applied to analyze serum haptoglobin β in healthy individuals and patients with hepatitis B virus (HBV), liver cirrhosis (LC), and HCC. The results showed that the glycosite ratios of VVLHPN#YSQVDIGLIK significantly changed in HCC patients compared with LC and HBV patients.

Except for these various glycoprotein/glycopeptide enrichment methods, Lee *et al*. [22] described a simple and efficient strategy for the selective identification and quantitation of N-linked glycoproteins without extensive enrichment steps prior to MS/MS analysis. Briefly, trypsin-digested samples were subjected to 1D LC MS/MS without prefractionation to construct an exclude peak list (EPL). All peptides in the exclusion list were nonglycosylated. Human plasma samples were then digested with trypsin, deglycosylated with PNGase F treatment, and subjected to 1D LC MS/MS with the EPL. Thus, the nonglycopeptides identified in previous runs were excluded from the MS/MS analysis, permitting the available data acquisition time to be used only for the peptides deglycosylated through PNGase F. These procedures allow N-linked glycopeptides present in complex human plasma to be detected using a 1D LC MS/MS run without prefractionation. This method was combined with iTRAQ labeling to quantify the N-linked glycoprotein differences between the plasma of healthy and HCC patients. Up to 14 higher expression N-linked glycopeptides were identified in the plasma of HCC patients. This method demonstrates an apparent advantage because it does not require enrichment steps, thereby reducing the risk of uneven losses of protein or peptide during enrichment.

The aforementioned methods are mainly applied in N-linked glycoprotein analysis because no enzymes are suitable to release the O-glycans. The strategies used for the quantitative studies of O-linked glycoproteins are limited. Vosseller *et al*. [23] described an approach based on β-elimination/Michael addition (BEMAD) for the identification and comparative quantitation of O-linked *N*-acetylglucosamine (GlcNAc)-modified peptides. The peptides were isotope labeled via Michael addition with normal DTT (d0) or deuterated DTT (d6) and then enriched through thiol chromatography. The changes in the glycosylation site occupancy could be obtained from the changes in the MS signal intensity ratios of d6/d0. In addition, they used isotope-coded affinity tags (ICAT) to compare with BEMAD. In the ICAT analysis, the isotope ratios of nonglycopeptides could indicate the glycoprotein changes in protein expression level. However, extensive internal controls are required to determine which O-linked species is being quantified because the reaction of β-elimination can occur in any modified O-linked serine or threonine residues. Khidekel *et al*. [24] described a new strategy termed quantitative isotopic and chemoenzymatic tagging for monitoring the dynamics of O-GlcNAc glycosylation. First, the O-GlcNAc-glycosylated proteins were chemoenzymatically tagged with a ketogalactose sugar. The ketone functionality was then reacted with an aminooxy biotin derivative to biotinylate the proteins. Thus, the O-GlcNAc-modified species were isolated through avidin chromatography. Finally, the O-GlcNAc proteins were proteolyzed and labeled with formaldehyde/$NaCNBH_3$ or deuterated formaldehyde/$NaCNBD_3$, and the labeled peptides were quantified through LC-MS. In addition, the tagged peptides could produce special ion signals, such as ketogalactose-biotin moiety (515.3 Da) and GlcNAc-ketogalactose-biotin moiety (718.4 Da), allowing for unambiguous detection of O-GlcNAc peptides. Using this method, they proved for the first time that O-GlcNAc glycosylation is dynamically modulated by excitatory stimulation of the brain *in vivo*. Rexach *et al*. [25] described a different method to visualize the O-GlcNAc–modified protein subpopulation using resolvable polyethylene glycol (PEG) mass tags. O-GlcNAc–modified proteins were chemoenzymatically labeled using the uridine diphosphate-ketogalactose analog and an engineered GalT enzyme and then reacted with an aminooxy-functionalized PEG mass tag. The O-GlcNAc–modified proteins labeled with the mass tag were not analyzed through MS but were resolved using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized in parallel by immunoblotting with antibodies against proteins of interest. Simple inspection of the mass-shifted bands would establish whether a protein was singly, doubly, or multiply glycosylated, and *in vivo* glycosylation stoichiometries could be determined by quantifying the relative intensities of each band. This approach enables rapid quantification without the need for protein purification, advanced instrumentation, or expensive radiolabels.

Proteolytic ^{18}O labeling

In 2002, Reynolds *et al*. [26] for the first time described a sequential double-labeling strategy to characterize N-linked glycopeptides. Using this approach, the protein mixtures were first digested by Glu-C in $H_2{}^{16}O$ or $H_2{}^{18}O$, and the

peptides were then deglycosylated in $H_2{}^{16}O$ or $H_2{}^{18}O$ with PNGase F treatment. Thus, a 6 Da difference occurred for glycopeptides having a single N-glycosylation site. The isotope-labeled peptides were pooled and measured via MS. The changes in the glycosylation site occupancy were obtained from the changes in signal intensity ratios of $^{18}O/^{16}O$. This strategy has its unique advantages: the mass shift of 6 Da can avoid the overlapping of light and heavy isotopic envelops, and the ^{18}O incorporation during the deglycosylation process increases the accuracy in identifying the glycosylation sites. However, no enrichment step is present, and the low-abundance glycoproteins are difficult to quantify. Moreover, Glu C is expensive and less efficient in proteomics analysis compared with trypsin. In 2010, we [27] developed a tandem ^{18}O stable isotope labeling (TOSIL) method to quantify accurately the changes in glycoprotein expression and changes in individual N-glycosylation site occupancy. As shown in Fig. 2, in our method, we used trypsin instead of Glu C and added the enrichment step (*i.e.,* enrichment of glycopeptides through hydrophilic affinity extraction) between the two labeling steps. The C-terminal of peptides were labeled with two ^{18}O in $H_2{}^{18}O$ catalyzed by trypsin, and the glycans were released with PNGase F hydrolysis in $H_2{}^{18}O$. For glycopeptides with a single glycosylation site, a mass shift of 6 Da could be shown. The relative quantities of N-glycosylated site occupancy and its parent protein level were obtained simultaneously by measuring the $^{18}O/^{16}O$ intensity ratios of glycopeptides and non-glycopeptides, respectively. The newly developed TOSIL approach yielded good linearity in quantitative response within a 10-fold dynamic range with a correlation coefficient $r^2 > 0.99$. This TOSIL method was successfully applied in the analysis of the serum glycoproteome of patients with ovarian cancer and healthy controls. A total of 86 N-glycosylation sites were quantified, and N-glycosylation levels of 56 glycopeptides showed significant changes. Most changes in N-glycosylation at specific sites have the same trends as those of protein expression levels. Shakey *et al.* [28] developed a similar strategy for relative quantitation of glycoproteins in complex biological mixtures. In this strategy, ^{18}O stable isotope labeling was also catalyzed by both trypsin and PNGase F, a different glycopeptide enrichment method was used, the trypsin digests were first oxidated, and the glycopeptides were enriched via hydrazide-coupled magnetic beads. Using this approach, they identified and quantitated 224 N-glycopeptides representing 130 unique glycoproteins from 20 μL of undepleted mouse serum samples.

In summary, these enzyme-catalyzed methods have the same features. First, the ^{18}O labeling is a residue-specific catalytic method, and no by-product is produced. Second, reagents are cost-effective and readily available. Third, the reaction conditions are mild, and the physical and chemical properties of peptides are not changed. Fourth, the method is compatible with all peptide-level fractionation methods and provides uniform labeling of all peptides in each protein. This proteolytic labeling has its own disadvantages and limitations. The method is limited to binary comparisons or series thereof. Incomplete labeling and back exchange should be considered and carefully controlled in experiments. In addition, these enzyme-catalyzed ^{18}O labeling methods have not been used for the quantitation of O-linked glycoproteins owing to the lack of suitable enzymes for releasing O-glycans.

Metabolic labeling

Metabolic labeling of cultured cells was introduced by Oda *et al.* [29] in 1999. Currently, stable isotope labeling with amino acids in cell culture (SILAC) has become a highly popular technique for the quantitation of proteins from cell cultures. It is also used in quantitative glycoproteomics studies.

Wollscheid *et al.* [30] showed a method to analyze cell-surface glycoproteins. They used the SILAC method to "pre-label" the proteins for quantitative analysis. The cell-surface glycoproteins were then captured using a

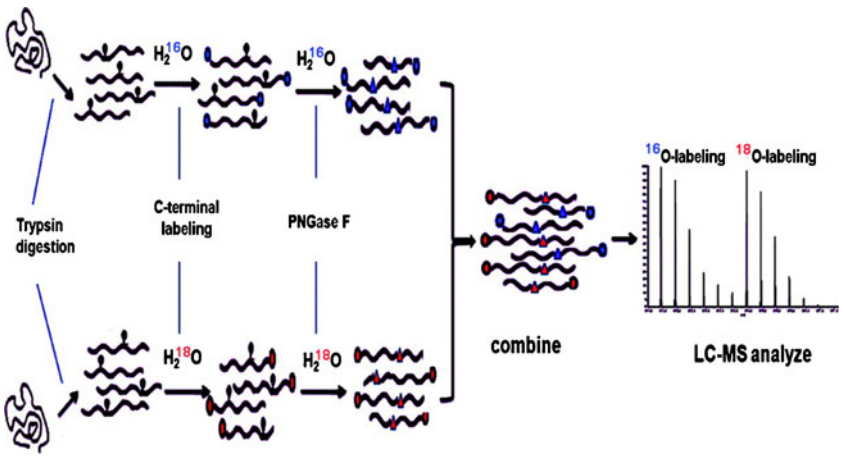

Fig. 2 Analytical strategy of Tandem ^{18}O Stable Isotope Labeling (TOSIL) in quantitative proteomics. (Reproduced from J. Proteome. Res. 2010, 9, 227–236 with permission from the American Chemical Society. ©2010 American Chemical Society)

bifunctional linker (one side with a hydrazide functional group to affix covalently to the glycans of cell surface glycoproteins, and the other side with a biotin tag that attaches to streptavidin beads for glycoprotein purification). After tryptic digestion, the glycopeptides were released with PNGase F. Protein concentration was inferred based on the differences in abundance of the various isotopically labeled peptides. Using this method, they compared the changes in cell surface glycosylation between T and B cells during neuronal activation. Metabolic labeling has many advantages. It reduces the risk of uneven losses of protein or peptide material during sample preparation. The stable isotopic tags are incorporated into the early stages of sample preparation, and SILAC labeling incorporation is virtually 100 %. As a result, the variation between the samples is reduced, and highly accurate quantification is achieved. However, metabolic labeling is relatively expensive and is not practical for analyzing biological samples (*e.g.,* tissues or body fluids) that cannot be grown in culture.

Label-free relative quantitation

Label-free methods have been developed because of the high cost of labeling reagents and some other limitative factors. Mann *et al*. [31] developed an approach for label-free quantitative glycoproteomics. They employed two lectins, Aleuria aurantia lectin (AAL) and Lotus tetragonolobus agglutinin (LTA), to enrich the fucosylated glycoproteins. The proteins were then fractionated with reversed-phase LC. Finally, the proteins were digested with trypsin and analyzed through LC-MS/MS. The proteins were quantified by summing the peak areas for extracted ion chromatograms of the identified peptides. They investigated the sera from two different esophagus disease states, high-grade dysplasia (HGD) and esophageal adenocarcinoma (EAC), and disease-free conditions. They identified 136 putative fucosylated glycoproteins with very high confidence level and found that Fetuin-B and EMILIN-2 were upregulated in HGD and EAC. During the workflow, a bicinchoninic acid assay (BCA) was introduced before LC fractionation to ensure that the same amount of protein was subjected to the reversed-phase fractionation, which could increase the accuracy of quantitative results. Goo *et al*. [32] also used a label-free method to analyze the secreted glycoproteins from cultured normal prostate and bladder stromal mesenchyme cells. After enrichment using hydrazide resin, the glycopeptides were analyzed via MS. They used calculated peptide ion current area to quantify the difference in protein abundance between the prostate and bladder. Up to 116 prostate- and 84 bladder-secreted glycoproteins were identified in their analysis. In addition, some stromal proteins were upregulated in the prostate, such as cathepsin L, follistatin-related protein, and so on.

Ishihara *et al*. [33] reported a quantitative label-free 2D LC-MALDI MS system for large-scale N-glycoproteomic analysis. In the method, glycoproteins were enriched using hydrazide resin. After trypsin digestion, the glycopeptides were deglycosylated by PNGase F and analyzed by 2D LC (strong cation exchange+reverse phase high-performance LC)-MALDI MS. A peptide internal control was used to normalize the signal intensities produced by 2D LC-MALDI MS, and the values linked to the LC data were treated using the DeView^{TM} software. They applied this method to analyze plasma samples from HCC patients and identified several biomarker candidates, including ceruloplasmin, alpha-1 antichymotrypsin, and multimerin-1. Using this method, an internal standard peptide was mixed into the samples for normalization, which increased the quantitation accuracy.

These label-free methods offer advantages, such as simpler workflow, because additional reagents and extra sample preparation steps are unnecessary. However, these methods always use nonglycosylated peptides of glycoproteins to quantify the relative abundance between different samples. Therefore, they reflect the abundance of glycoproteins in the protein expression level, but the glycosylation site occupancy is not shown.

Absolute quantitation

In recent years, absolute quantitation of target proteins has become a burgeoning topic in quantitative proteomics. However, studies on the absolute quantitation of glycoproteins are very limited. In general, stable isotope-labeled standard (SIS) peptides related to the target glycoproteins were mixed into the sample. Selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) techniques were then employed to accomplish absolute quantitation.

Ahn *et al*. [34] presented a sensitive method for the quantitative analysis of aberrant GlcNAcylated tissue inhibitor of metalloproteinase 1 (TIMP1) in the serum of colorectal cancer patients. Using this approach, the glycoproteins containing an N-linked glycan terminating with β(1-6)-GlcNAc were first enriched by phytohemagglutinin-L4 (L-PHA) and then digested with trypsin. Using a monoclonal anti-peptide TIMP1 antibody linked covalently to magnetic beads, a unique target peptide (antigen) of TIMP1 was immunoenriched from the L-PHA-enriched tryptic digests. Two peptides were selected from TIMP1, and their stable isotope-coded counterpart was prepared. With these peptides, the TIMP1 was quantified via the MRM technique at a serum level of 0.8 ng/mL. Thereafter, the same group [35] used a method consisting of multiplex lectin-fractionation, mass profiling of the lectin-captured fractions, and MRM-based quantitative analysis to monitor quantitatively the protein glycosylation diversity of human serum. They used different

lectins to capture the glycoproteins, such as AAL, L-PHA, Con A, and Datura stramonium agglutinin (DSA). Each lectin-captured serological glycoprotein was digested with trypsin and analyzed via 1D LC-MS/MS. Finally, the abundance of certain serological glycoproteins profiled in each lectin-captured fraction was further verified via the MRM-based quantitation method. The diversity of serological protein glycosylation in a human serum sample was quantitatively monitored because different lectins could recognize different carbohydrate structures. Using lectins and antibodies to enrich glycoproteins decreases the complexities of samples and increases the identification efficiency of the glycoproteins and accuracy of quantitation. However, during the enrichment process, some glycoproteins may be lost, thereby decreasing the accuracy of the quantitation results. In addition, the method reflects the glycoprotein abundance in protein expression, but does not show the degree of glycosylation.

In addition, some techniques were developed to quantify glycoproteins containing specific glycan structures. Recently, Ahn *et al*. [36] developed a lectin-coupled MRM-based mass spectrometric method to quantify fucosylated glycoproteins from serum samples. They used AAL immobilized on magnetic beads to capture fucosylated glycoproteins, and the captured proteins were digested by trypsin and profiled by tandem mass spectrometry, then target glycoproteins were selected from the proteomic profiling data and quantitatively analyzed by MRM-based analysis. Targeted glycoproteins were absolutely quantified, for example, the abundance of AGP was measured with wide ranges, from 5 fmol to 438 fmol, obtained from eight human plasma samples consisting of four healthy individuals and four HCC patients. This technique has also been extended to analyze other glycoproteins with specific glycan structures. For instance, Zhao *et al*. [37] used lectin enrichment (core fucosylation glycoproteins captured by LCH sepharose 4B) combined with MRM technique to quantify core fucosylated glycoproteins. Kurogochi *et al*. [38] used reverse glycoblotting combined with MRM assay to quantify sialylated glycoproteins.

Li *et al*. [39] developed a method combining the glycopeptide solid-phase extraction strategy and SRM to analyze glycosylated and sialylated PSA in prostate cancer and non-cancer tissues. Through this approach, proteins were first digested with trypsin, and glycopeptides were isolated from complex mixtures with solid-phase extraction. The N-glycopeptides were then released by PNGase F and subsequently identified via LC-MS/MS. Clinical immunoassay was used to quantify the total PSA protein. The enriched N-linked glycopeptides from target glycoprotein combined with SRM were employed to quantify the glycosylated PSA, whereas enriched sialylated glycopeptides from target glycoprotein coupled with SRM were used to quantify sialylated PSA. The data showed that the total PSA and the total glycosylated PSA from cancer and non-cancer tissues were not significantly different. However, the sialylated glycopeptide level was elevated in cancer tissues compared with that in non-cancer tissues. The relative abundance of glycosylated PSA was not correlated with the total PSA protein level, and sialylated PSA was not correlated with either total PSA or glycosylated PSA. In this study, the amounts of total PSA, glycosylated PSA, and sialylated PSA were obtained.

Conclusion and future view

Recent research has implied that monitoring cancer-specific glycosylation states may be a promising approach for detecting cancer with high specificity and sensitivity. Therefore, a high-throughput and highly sensitive method that can accurately quantify glycosylation site occupancy and analyze the relationship between the changes in glycosylation site occupancy and protein expression is urgently required for more effective biomarker development. However, much remains to be accomplished. First, more effective and higher selective methods must be developed to enrich glycoproteins/glycopeptides with high recovery rate, reduce sample complexity, and enhance dynamic ranges. Second, for stable isotope labeling methods, the efficiency of labeling should be increased. For example, in enzyme-catalyzed labeling, immobilized enzymes can be used to increase labeling efficiency and decrease back-exchange. Third, in the label-free method, fastidious attention should be given to all steps during the sample preparation, and more robust and reproducible LC-MS/MS method should be developed to increase quantitation accuracy. In addition, software and bioinformatics methods need to be developed to improve quantitative analysis. With the development in glycoproteomics research and MS techniques, we hope to see growth in the development of novel methods for high-throughput quantitative glycoproteomics in the near future.

Acknowledgements The work was supported by NST (2012CB910602 and 2012AA020203), NSF (21005020, 31070732 and 21025519), Shanghai Projects (11XD1400800, Shuguang, Eastern Scholar and B109)

References

1. Apweiler, R., Hermjakob, H., Sharon, N.: On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta **1473**, 4–8 (1999)
2. Hofsteenge, J., Muller, D.R., Beer, T., Loffler, A., Richter, W.J., Vliegenthart, J.F.G.: New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase. Biochemistry **33**, 13524–13530 (1994)

3. Floyd, N., Vijayakrishnan, B., Koeppe, J.R., Davis, B.G.: Thiyl glycosylation of olefinic proteins: S-linked glycoconjugate synthesis. Angew. Chem. Int. Ed. **48**, 7798–7802 (2009)
4. Thayer, D.A., Yu, H.N., Galan, M.C., Wong, C.H.: A general strategy toward S-linked glycopeptides. Angew. Chem. Int. Ed. **44**, 4596–4599 (2005)
5. Ohtsubo, K., Marth, J.D.: Glycosylation in cellular mechanisms of health and disease. Cell **126**, 855–867 (2006)
6. Catalona, W.J., Smith, D.S., Ratliff, T.L., Dodds, K.M., Coplen, D.E., Yuan, J.J.J., Petros, J.A., Andriole, G.L.: Measurement of prostate-specific antigen in serum as a screening-test for prostate-cancer. New. Eng. J. Med. **324**, 1156–1161 (1991)
7. Zhang, Z., Bast, R.C., Yu, Y.H., Li, J., Sokoll, L.J., Rai, A.J., Rosenzweig, J.M., Cameron, B., Wang, Y.Y., Meng, X.Y., Berchuck, A., Haaften-Day, C.V., Hacker, N.F., Bruijn, H.W.A., Zee, A.G.J., Jacobs, I.J., Fung, E.T., Chan, D.W.: Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res. **64**, 5882–5890 (2004)
8. Clark, G.M., Mcguire, W.L.: Follow-up-study of her2/neu amplification in primary breast-cancer. Cancer Res. **51**, 944–948 (1991)
9. Sato, Y., Nakata, K., Kato, Y., Shima, M., Ishii, N., Koji, T., Taketa, K., Endo, Y., Nagataki, S.: Early recognition of hepatocellular-carcinoma based on altered profiles of alpha-fetoprotein. N. Engl. J. Med. **328**, 1802–1806 (1993)
10. Wei, X., Li, L.: Comparative glycoproteomics: approaches and applications. Brief. Funct. Genomic. Proteomic. **8**, 104–113 (2009)
11. Pan, S., Chen, R., Aebersold, R., Brentnall, T.A.: Mass spectrometry based glycoproteomics—from a proteomics perspective. Mol. Cell. Proteomics. **10**, 1–14 (2011)
12. Zhang, H., Li, X.J., Martin, D.B., Aebersold, R.: Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat. Biotech. **21**, 660–666 (2003)
13. Zhou, L., Beuerman, R.W., Chew, A.P., Koh, S.K., Cafaro, T.A., Urrets-Zavalia, E.A., Urrets-Zavalia, J.A., Li, S.F.Y., Serra, H.M.: Quantitative analysis of N-linked glycoproteins in tear fluid of climatic droplet keratopathy by glycopeptide capture and iTRAQ. J. Proteome. Res. **8**, 1992–2003 (2009)
14. Qiu, R.Q., Regnier, F.E.: Use of Multidimensional lectin affinity chromatography in differential glycoproteomics. Anal. Chem. **77**, 2802–2809 (2005)
15. Qiu, R.Q., Regnier, F.E.: Comparative glycoproteomics of N-linked complex-type glycoforms containing sialic acid in human serum. Anal. Chem. **77**, 7225–7231 (2005)
16. Ueda, K., Katagiri, T., Shimada, T., Irie, S., Sato, T.A., Nakamura, Y., Daigo, Y.: Comparative profiling of serum glycoproteome by sequential purification of glycoproteins and 2-nitrobenzenesulfenyl (NBS) stable isotope labeling: A new approach for the novel biomarker discovery for cancer. J. Proteome. Res. **6**, 3475–3483 (2007)
17. Ueda, K., Takami, S., Saichi, N., Daigo, Y., Ishikawa, N., Kohno, N., Katsumata, M., Yamane, A., Ota, M., Sato, T.A., Nakamura, Y., Nakagawa, H.: Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. Mol. Cell Proteomics. **9**, 1819–1828 (2010)
18. Shetty, V., Nickens, Z., Shah, P., Sinnathamby, G., Semmes, O.J., Philip, R.: Investigation of sialylation aberration in n-linked glycopeptides by lectin and tandem labeling(LTL) quantitative proteomics. Anal. Chem. **82**, 9201–9210 (2010)
19. Kurogochi, M., Amano, M., Fumoto, M., Takimoto, A., Kondo, H., Nishimura, S.I.: Reverse Glycoblotting Allows rapid-enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers. Angew. Chem. Int. Ed. **46**, 8808–8813 (2007)
20. Chen, R., Wang, F.J., Tan, Y.X., Sun, Z., Song, C.X., Ye, M.L., Wang, H.Y., Zou, H.F.: Development of a combined chemical and enzymatic approach for the mass spectrometric identification and quantification of aberrant N-glycosylation. J. Proteomics. **75**, 1666–1674 (2012)
21. Zhang, S., Liu, X.H., Kang, X.N., Sun, C., Lu, H.J., Yang, P.Y., Liu, Y.K.: iTRAQ plus ^{18}O: A new technique for target glycoprotein analysis. Talanta **91**, 122–127 (2012)
22. Lee, H.J., Na, K., Choi, E.Y., Kim, K.S., Kim, H., Paik, Y.K.: Simple method for quantitative analysis of N-linked glycoproteins in hepatocellular carcinoma specimens. J. Proteome Res. **9**, 308–318 (2010)
23. Vosseller, K., Hansen, K.C., Chalkley, R.J., Trinidad, J.C., Wells, L., Hart, G.W., Burlingame, A.L.: Quantitative analysis of both protein expression and serine/threonine post-translational modifications through stable isotope labeling with dithiothreitol. Proteomics **5**, 388–398 (2005)
24. Khidekel, N., Ficarro, S.B., Clark, P.M., Bryan, M.C., Swaney, D.L., Rexach, J.E., Sun, Y.E., Coon, J.J., Peters, E.C., Hsieh-Wilson, L.C.: Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. Nat. Chem Bio. **3**, 339–348 (2007)
25. Rexach, J.E., Rogers, C.J., Yu, S.H., Tao, J.F., Sun, Y.E., Hsieh-Wilson, L.C.: Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags. Nat. Chem Bio. **6**, 645–651 (2010)
26. Reynolds, K.J., Yao, X.D., Fenselau, C.: Proteolytic ^{18}O labeling for comparative proteomics: evaluation of endoprotease Glu-c as the catalytic agent. J. Proteome. Res. **1**, 27–33 (2002)
27. Liu, Z., Cao, J., He, Y.F., Qiao, L., Xu, C.J., Lu, H.J., Yang, P.Y.: Tandem 18O stable isotope labeling for quantification of N-glycoproteome. J. Proteome. Res. **9**, 227–236 (2010)
28. Shakey, Q., Bates, B., Wu, J.: An approach to quantifying N-linked glycoproteins by enzyme-catalyzed $^{18}O_3$-labeling of solid-phase enriched glycopeptides. Anal. Chem. **82**, 7722–7728 (2010)
29. Oda, Y., Huang, K., Cross, F.R., Cowburn, D., Chait, B.T.: Accurate quantitation of protein expression and site-specific phosphorylation. Proc. Natl. Acad. Sci. **96**, 6591–6596 (1999)
30. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., Watts, J.D.: Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat. Biotech. **27**, 378–386 (2009)
31. Mann, B., Madera, M., Klouckova, I., Mechref, Y., Dobrolecki, L.E., Hickey, R., Hammoud, Z.T., Novotny, M.: A quantitative investigation of fucosylated serum glycoproteins with application to esophageal adenocarcinoma. Electrophoresis **31**, 1833–1841 (2010)
32. Goo, Y.A., Liu, A.Y., Ryu, S., Shaffer, S.A., Malmstrom, L., Page, L., Nguyen, L.T., Doneanu, C.E., Goodlett, D.R.: Identification of secreted glycoproteins of human prostate and bladder stromal cells by comparative quantitative proteomics. Prostate **69**, 49–61 (2009)
33. Ishihara, T., Fukuda, I., Morita, A., Takinami, Y., Okamoto, H., Nishimura, S.I., Numata, Y.: Development of quantitative plasma N-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma. J. Proteomics. **74**, 2159–2168 (2011)
34. Ahn, Y.H., Lee, J.Y., Lee, J.Y., Kim, Y.S., Ko, J.H., Yoo, J.S.: Quantitative analysis of an aberrant glycoform of timp1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. J. Proteome. Res. **8**, 4216–4224 (2009)
35. Ahn, Y.H., Ji, E.S., Shin, P.M., Kim, K.H., Kim, Y.S., Kod, J.H., Yoo, J.S.: A multiplex lectin-channel monitoring method for human serum glycoproteins by quantitative mass spectrometry. Analyst **137**, 691–703 (2012)
36. Ahn, Y.H., Shin, P.M., Ji, E.S., Kim, K., Yoo, J.S.: A lectin-coupled, multiple reaction monitoring based quantitative analysis

of human plasma glycoproteins by mass spectrometry. Anal. Bioanal. Chem. **402**, 2101–2112 (2012)
37. Zhao, Y., Jia, W., Wang, J.F., Ying, W.T., Zhang, Y.J., Qian, X.H.: Fragmentation and site-specific quantification of core fucosylated glycoprotein by multiple reaction monitoring-mass spectrometry. Anal. Chem. **83**, 8802–8809 (2011)
38. Kurogochi, M., Matsushista, T., Amano, M., Furukawa, J., Shinohara, Y., Aoshima, M., Nishimura, S.I.: Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay. Mol. Cell. Proteomics. **9**, 2354–2368 (2010)
39. Li, Y., Tian, Y., Rezai, T., Prakash, A., Lopez, M.F., Chan, D.W., Zhang, H.: Simultaneous analysis of glycosylated and sialylated PSA reveals differential distribution of glycosylated PSA isoforms in prostate cancer tissues. Anal. Chem. **83**, 240–245 (2011)