HHS Public Access

Author manuscript

Anal Chem. Author manuscript; available in PMC 2021 January 07.

Published in final edited form as:

Anal Chem. 2020 January 07; 92(1): 267–291. doi:10.1021/acs.analchem.9b04651.

Recent Advances in Glycoproteomic Analysis by Mass Spectrometry

Suttipong Suttapitugsakul[†], **Fangxu Sun**[†], **Ronghu Wu**[*]

School of Chemistry and Biochemistry and the Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Glycosylation is one of the most common protein modifications and is essential for cells. This modification is exceptionally complex because glycans are highly diverse and can be covalently bound to several amino acid residues in proteins through various configurations. There are two major types of protein glycosylation, i.e., N-linked glycosylation in which glycans are attached to the side chain of asparagine and O-linked glycosylation referring to glycans being bound to the side chains of serine and threonine.[1,2] Glycosylation plays vital roles in cells, including determination of protein folding, trafficking and stability, and regulation of nearly every extracellular activity such as cell–cell communication and cell–matrix interactions.[3,4] Aberrant protein glycosylation is directly related to multiple diseases, including cancer, neurodegenerative disorders, pulmonary diseases, blood disorders, and genetic diseases.[5,6] Due to the importance and complexity of protein glycosylation in biological systems, there is a longstanding interest to develop innovative methods to study glycoproteins and apply them for biomedical research. Investigation of protein glycosylation has become more popular with the development of modern instrumentation and computational methods. According to a PubMed search using the keyword “glycosylation”, 16 publications were listed during 1960–1970 while over 20 000 studies were reported in the past 10 years. With the growing interests in protein glycosylation, this trend is expected to continue in the next decades.

Mass spectrometry (MS)-based proteomics provides an excellent opportunity to globally analyze proteins and their modifications.[7–19] Nonetheless, it is still extremely challenging to comprehensively analyze protein glycosylation.[20] Unlike many other modifications with a fixed structure for the modified group, such as phosphorylation, the diversity of glycans makes it more challenging to employ the commonly used database searching methods such as SEQUEST and Mascot to identify glycopeptides in bottom-up proteomics. Low-abundance glycoproteins in complex biological samples are also hindered for detection by many high-abundance nonglycoproteins. Furthermore, glycans can interfere with the fragmentation of the peptide backbone.[20,21] Innovative and effective methods are critical to overcome these hurdles and to allow for comprehensive analysis of glycoproteins using MS.

The authors declare no competing financial interest.

[*]**Corresponding Author**: ronghu.wu@chemistry.gatech.edu. Phone: 404-3851515. Fax: 404-894-7452.

[†]Author Contributions

S.S. and F.S. contributed equally.

In this Review, after a brief introduction of protein glycosylation, we highlight several enrichment methods to analyze different types of protein glycosylation and describe the analysis of glycoproteins with a particular glycan. Recent advances in instrumentation and bioinformatics for intact glycopeptide analysis are also covered. Lastly, we summarize some recent glycoproteomic applications, including studies of glycoprotein functions, biomarker discovery, and analysis of glycoproteins from different biological sources, with a focus on their applications in clinical and biomedical research. For general information on proteomic analysis, there are many excellent review papers published previously,[22,23] including some for glycoproteomic analysis.[24,25]

OVERVIEW OF PROTEIN GLYCOSYLATION

Glycosylation is a common and essential modification where glycans are covalently attached to proteins. Among several types of protein glycosylation, N- and O-linked glycosylation are the two major ones. For N-linked glycosylation, the glycan precursor $GlcNAc_2Man_9Glc_3$ is transferred *en bloc* to nascent peptides immediately after they are being translated by the ribosome. Thus, glycosylation is also called a cotranslational modification. N-Glycans normally contain a common $GlcNAc_2Man_3$ core that is further modified with various monosaccharides by enzymes in the endoplasmic reticulum (ER) and Golgi apparatus, including the terminal sialic acid and core fucose residues. N-Glycosylation sites typically have a canonical motif, i.e., N-X-S/T where X can be any amino acid residue except proline. However, the presence of this motif does not guarantee that a particular copy of the same protein would be glycosylated, which may be referred to as the macroheterogeneity of protein glycosylation. Glycans may contain several types of monosaccharides linked together through different modes of connections even at the same site on different copies of the same protein, which is referred to as the microheterogeneity of protein glycosylation, further increasing their complexity. For O-linked glycosylation, glycans are attached to the side chains of serine and threonine. Mucin-type O-glycosylation is the most common protein O-glycosylation where, typically, monosaccharides are sequentially added to proteins by various glycosyltransferases, instead of the *en bloc* glycan transfer in N-glycosylation. O-GlcNAcylation, where the GlcNAc group is dynamically added or removed from serine and threonine similar to phosphorylation, belongs to this category as well. Unlike N-linked glycosylation, there is no canonical motif for O-linked glycosylation.

In a typical glycoproteomic study, proteins are first extracted from cells or tissues before glycoproteins or glycopeptides are enriched or separated from the complex samples. MS analysis can then be performed at different levels with different glycospecies, including glycoproteins, intact or derivatized glycopeptides, deglycosylated peptides, as well as released glycans, which provide different information from glycosylation sites to glycan structures. As described in the introduction, glycosylation is essential for cells, and glycoproteins contain much valuable information regarding the statuses of cellular development and disease. Therefore, it is critically important to identify glycosylation sites on glycoproteins and elucidate glycan structures.

ENRICHMENT METHODS FOR GLOBAL ANALYSIS OF GLYCOPROTEINS/GLYCOPEPTIDES BY MS

Enrichment of glycoproteins/glycopeptides from complex biological samples is imperative to achieve their global analysis by MS because of the low abundance of many glycoproteins. The heterogeneity of glycans further makes the enrichment more challenging.[24–26] A variety of enrichment methods have been developed that dramatically improve the coverage of glycoproteins/glycopeptides in different samples.

Lectin.

Lectins have been widely used to enrich N- or O-glycopeptides and glycoproteins.[27] Concanavalin A (Con A) and wheat germ agglutinin (WGA) are frequently employed for N-glycopeptide enrichment.[28] Con A recognizes α-linked mannose, while WGA binds to *N*-acetylglucosamine and sialic acid. Jacalin and *Vicia villosa* agglutinin (VVA), which are often exploited for mucin-type O-glycopeptide enrichment, bind to O-linked galactosyl (β-1,3) *N*-acetylgalactosamine and α- or β-linked terminal *N*-acetylgalactosamine, respectively.[29,30] Lectins are normally immobilized onto solid supports, such as agarose or magnetic beads. Recently, magnetic nanoparticles (MNPs) have become an attractive solid support because of the increased lectin density on the particle surface, and the easy preparation and separation.[31] Waniwan et al. identified 2290 and 2767 glycopeptides from EGFR-TKI sensitive PC9 cells and resistant PC9-IR cells, respectively, by using MNPs conjugated with Con A, *Aleuria aurantia* lectin (AAL), and *Sambucus nigra* agglutinin (SNA).[32]

The selective affinity of a lectin toward specific carbohydrates makes lectin-based enrichment methods biased for glycoproteins with certain glycans. A combination of different lectins, which is normally called multilectin affinity chromatography (M-LAC), can be used to improve the coverage of glycopeptides.[33] However, the combination of multiple lectins complicates the preparation of lectin-based chromatography and does not fully solve the problem of biased recognition. Coupling lectin enrichment with other separation methods can further improve the enrichment efficiency. Zhou et al. found that the tandem use of lectin followed by MAX (a commercial hydrophilic-based strong anion exchange column) enrichment resulted in a better performance for fucosylated glycopeptide identification. They employed this tandem enrichment method to identify and quantify 973 intact fucosylated glycopeptides from 252 proteins in nonaggressive and aggressive prostate cancer cell lines.[34] In order to minimize the biased issue of lectin-based methods, an engineered Fbs1 carbohydrate binding protein was recently reported and showed its promising application for N-glycopeptide enrichment. Fbs1 is a component of the E3 ubiquitin ligase complex that can recognize N-glycans of misfolded glycoproteins and thus mediates the removal of misfolded N-glycoproteins through the endoplasmic reticulum-associated degradation (ERAD) pathway.[35] However, wild-type Fbs1 prefers to recognize high-mannose N-glycans. Through mutagenesis and plasmid display selection, it was revealed that the Fbsl GYR mutant, which displayed a higher binding affinity toward diverse types of N-glycans, had great potential for N-glycopeptide enrichment. Over 2500 intact

glycopeptides were identified through the Fbs1 GYR enrichment, while the lectin enrichment resulted in the identification of only 1172 glycopeptides in the same work.[36]

HILIC.

Hydrophilic interaction liquid chromatography (HILIC), which relies on the hydrophilicity difference between glycopeptides and nonglycopeptides, has been widely employed in glycoproteomic analysis.[37–39] One large disadvantage of HILIC is the low enrichment specificity caused by coelution of hydrophilic nonglycopeptides. Recent developments of HILIC focus on synthesis of novel solid materials conjugated with various hydrophilic groups to increase the enrichment specificity. Different materials were developed as the solid matrices including metal–aorganic frameworks (MOF)[40] and magnetic particles.[41] Wang et al. synthesized a MOF-based solid matrix by conjugating Zn–MOF onto the surface of magnetic graphene for the enrichment of glycopeptides. The newly functionalized MOF enabled the identification of 517 N-glycopeptides from 151 glycoproteins in 1 μL of human serum.[42] Various types of hydrophilic functional groups were also immobilized onto the solid matrices. Zwitterionic stationary phase-based HILIC (ZIC-HILIC), which carries both positive and negative charges on the surface, displayed higher enrichment specificity compared with normal HILIC.[43] Cao et al. introduced zwitterionic groups onto poly(amidoamine) dendrimer (PAMAM) and identified 48 glycosylation sites from 28 glycoproteins in only 0.1 μL of human serum.[44] Even though different types of solid supports and HILIC materials were reported in the literature, comprehensive evaluation of their performance on glycopeptide enrichment for large-scale analysis and systematic comparison with other commonly used enrichment methods remain to be performed in order to evaluate their practical utility for glycoproteomic analysis.

Hydrazide Chemistry-Based Enrichment Methods.

Hydrazide chemistry-based methods rely on the covalent bond formation between the hydrazide groups and the aldehyde groups on oxidized glycans to enrich glycopeptides/glycoproteins. With the help of peptide N-glycosidase F (PNGase F) to remove N-glycans, hydrazide chemistry-based methods showed a great capacity for N-glycosylation site mapping.[8,45] Recently, a responsive polymer-based platform was developed to achieve homogeneous hydrazide chemistry-based enrichment, which is expected to have higher enrichment efficiency compared with the traditional solid–liquid heterogeneous enrichment. The hydrazide groups were conjugated to a polymeric backbone, i.e., poly(acrylic acid) (PAA) that is pH-responsive[46] or poly(*N*-isopropylacrylamide) (PNIPAM) that is thermal-responsive.[47] The obtained polymers can be dissolved in aqueous solution under certain conditions, which facilitate the reaction between the hydrazide groups and the aldehyde groups on oxidized glycopeptides. By changing the pH or temperature, the polymers were precipitated, which allowed for the enrichment of glycopeptides. Based on this strategy, 1317 N-glycopeptides from 458 N-glycoproteins were characterized in mouse brain samples and 329 N-glycosylation sites on 180 N-glycoproteins were identified from the plasma exosomes.

Boronic Acid-Based Enrichment Methods.

Boronic acids can form cyclic boronate esters with cis-diols on glycans under basic conditions, and the resulting boronate esters can reversibly undergo hydrolysis under acidic conditions, bringing back the boronic acids and cis-diols without affecting the glycan structures. Because glycans carry multiple hydroxyl groups, boronic acid-based methods hold great potential to globally enrich glycopeptides.[48–54] However, the weak interactions between boronic acids and glycans restrict its wide applications in glycoproteomic analysis.

Boronic acid was immobilized onto different types of solid supports, including MOF,[55] magnetic nanoparticles,[56] and graphene,[57] to improve their capacities for glycopeptide and glycoprotein enrichment. Recently, a dendrimer platform was employed to further facilitate the enrichment.[58] Since different types of boronic acids display distinct binding affinities to cis-diols, evaluation of several types of boronic acids revealed that benzoboroxole, which was reported to possess high binding affinity toward cis-diols,[59] allowed for the identification of the largest number of N-glycopeptides (Figure 1A,B). The Dendrimer conjugated with the Boronic Acid derivative (DBA) can increase the density of benzoboroxole molecules on the surface of the dendrimer beads, and thus several benzoboroxole molecules can interact with a single glycan simultaneously (Figure 1C,D). The synergistic effect between glycans and multiple benzoboroxole molecules in the dendrimer platform dramatically enhances their interactions, and therefore, the affinity toward glycans is increased. The optimization of dendrimer size revealed that the number of identified glycopeptides increased significantly compared with the same magnetic beads without a dendrimer (Figure 1E). The enrichment is fast and can be completed in only 10 min (Figure 1F). The newly developed method was employed to enrich glycopeptides from different samples. For N-linked glycosylation, over 1000 sites on 501 glycoproteins and 4195 sites on 1608 glycoproteins were identified from yeast cells and mouse brain tissues, respectively. From three human cell lines (MCF7, HEK 293T, and Jurkat), 4691 sites were identified on 1906 glycoproteins. The method was also applied to enrich O-glycosylated peptides. Over 230 O-mannosylated proteins were identified from yeast, and more than 200 O-GlcNAcylated proteins were characterized from human cells even though the modified group (GlcNAc) is small and does not have cis-diol, which further demonstrates that the DBA method is highly effective for the enrichment of glycopeptides.[58]

Methods Integrating Chemical and Enzymatic Reactions.

Chemical and enzymatic reactions have been frequently employed to investigate protein glycosylation. Sun et al. designed a beautiful method called solid phase extraction of N-linked glycans and glycosite-containing peptides (NGAG) for global and simultaneous analysis of N-linked glycans, glycosylation sites, and intact glycopeptides from complex samples (Figure 2).[60] In this approach, peptides were first covalently conjugated to beads, followed by a series of chemical and enzymatic reactions to release N-glycans and deglycosylated peptides (Figure 2A). The identification of N-glycans and deglycosylated peptides enabled the construction of a library of all possible intact glycopeptides. Then intact N-glycopeptides can be identified by assigning the oxonium ions-containing spectra generated from enriched glycopeptides based on the constructed library (Figure 2B). NGAG is an excellent example showing an integration of various chemical and enzymatic reactions

for glycoproteomic analysis. Recently, the same lab reported another innovative method named extraction of O-linked glycopeptides (EXoO) for large-scale analysis of mucin-type O-glycosylation.[61] In this method, peptides were conjugated to a solid support, and O-glycopeptides were specifically released by a protease called OpeRATOR, which cleaves N-terminally the Ser and Thr sites containing O-linked glycans. Using this method, they identified 3055 O-linked glycosylation sites from 1060 glycoproteins in different samples including kidney tissues, T cells, and sera.

The introduction of bioorthogonal functional groups, including azides, alkynes, or ketones, to glycoproteins through metabolic or chemoenzymatic labeling followed by affinity enrichment represents a class of useful strategies for comprehensive analysis of protein glycosylation.[62–65] Integrating metabolic labeling, copper-free click chemistry, and MS-based proteomics, the global and site-specific analysis of cell-surface glycoproteins have been systematically studied, and their analysis is described further in the Applications of MS-Based Glycoproteomics in Biological Systems. Cell-surface glycoproteins can also be analyzed by chemoenzymatic labeling using recombinant glycosyltransferases and a nucleotide sugar functionalized with biotin, followed by enrichment and MS identifications.[66] Isotope-targeted glycoproteomics (IsoTaG), which integrated metabolic labeling, chemical enrichment, and isotopic coding of glycopeptides, is an innovative approach for intact N- and O-glycopeptide analysis.[67] The key for this method is to use a probe with two bromine atoms to tag metabolically labeled glycoproteins, which allowed for targeted assignment of labeled glycopeptides during MS analysis.

COMPREHENSIVE ANALYSIS OF GLYCOPROTEINS BEARING A SPECIFIC GLYCAN

For glycoproteins, both the protein and the glycan components contain valuable information about the disease and developmental statuses of cells.[68,69] While the majority of glycoproteomic studies mainly focus on analyzing glycosylation sites and glycans separately due to the complexity of protein glycosylation, the study of proteins containing a particular glycan still remains largely unexplored. The development of modern MS-based proteomics and effective separation methods provides an opportunity to systematically analyze glycoproteins containing a particular and important glycan.

Truncated Mucin-Type O-Glycans.

Aberrant glycosylation is often correlated with different diseases.[70,71] The abnormal changes include altered expression of certain types of glycans or expression of new glycans.[72] Previous studies revealed that the overexpression of certain glycans is the hallmark of cancer. Among those, truncated mucin-type O-glycans including the Tn, Sialyl-Tn (STn), T, and Sialyl-T (ST) antigens are often highly expressed in human cancers.[73,74] Therefore, it is of great importance to detect tumor-associated glycans in order to study the physiological and pathological processes that they participate in. Lectins and antibodies have been used to study truncated mucin-type O-glycans for decades, but they often suffer from weak binding affinity and low specificity.[75,76] The combination of effective separation and MS-based

proteomics provides a feasible approach to globally characterize glycoproteins with these truncated mucin-type O-glycans.

Tn Antigen.—Lectins such as VVA with relatively high affinity toward the Tn antigen can be used to separate Tn-bearing glycoproteins. Hoja-Łukowicz et al. employed VVA-based chromatography to enrich glycoproteins bearing the Tn antigen, followed by digestion and LC–MS/MS analysis.[77] They identified 146 Tn-bearing glycoproteins, which represented the first large-scale proteomic analysis of glycoproteins with the Tn antigen in human melanoma cells. Jacalin was also used to enrich Tn-bearing glycopeptides prior to MS analysis for site-specific identifications.[78] Recently, a method integrating enzymatic and chemical reactions was developed to enrich glycoproteins with the Tn antigen.[79] Galactose oxidase (GAO) was employed to specifically convert the hydroxyl group at the C6 position on the Tn antigen to an aldehyde group, followed by enrichment with hydrazide beads. Glycopeptides with the Tn antigen were then released from the beads with methoxylamine for MS analysis (Figure 3A). The method enabled the identification of 96 glycoproteins bearing the Tn antigen in Jurkat cells (Figure 3B), and the overlap of the glycoproteins with the Tn antigen identified from biologically triplicate experiments demonstrated the reasonably high reproducibility of the method (Figure 3C).

STn Antigen.—To simplify the identification of STn-bearing proteins, neuraminidase is often utilized to remove the terminal sialic acid residue so the resulting Tn-bearing glycopeptides or glycoproteins can be enriched by lectins that are commercially available for the Tn antigen recognition, followed by LC–MS/MS analysis to achieve protein or site-specific identifications. Campos et al. used this methodology to analyze glycoproteins with the STn antigen in sera from gastric carcinoma patients.[80] They identified 37 O-glycoproteins with 49 O-glycosylation sites. A similar strategy was also applied to identify STn-bearing proteins from hypoxic T24, 5637, and HT1376 cells.[81]

This approach, however, cannot distinguish glycoproteins with the STn antigen from those with the endogenous Tn antigen. Further experiments including immunohistochemistry and a proximity ligation assay are normally required to validate the STn-bearing glycoproteins. The hydrazide chemistry-based method can minimize the limitations of the above method. In one study, sialic acid was selectively oxidized by periodate followed by hydrazide chemistry-based enrichment. After mild acid hydrolysis to remove the sialic acid residue, the identification of Tn-bearing peptides, which is the desialylated form of STn, was performed through MS analysis.[82] However, this method cannot reveal the degree of sialylation.

T Antigen.—Chemoenzymatic labeling is very promising to study glycoproteins with a specific glycan, in which a recombinant glycosyltransferase is employed to transfer a sugar analogue with a chemical handle from a nucleotide sugar donor to a specific glycan acceptor.[83] Li et al. reported a novel tandem chemoenzymatic strategy to selectively detect the T antigen (also called the Thomsen–Friedenreich (TF) antigen).[84] They first employed fucosyltransferase WbwK, a glycosyltransferase from bacteria that can specifically recognize the T antigen, to convert the T antigen to Fucα1,2-Galβ1,3-GalNAc-α. Another glycosyltransferase (*N*-acetylgalactosaminyltransferase BgtA) is then used to transfer a GalNAc analogue with an azido group, UDP-*N*-azidoacetylgalactosamine (UDP-GalNAz),

to the Fuc-containing trisaccharide. Therefore, the tandem utilization of the two glycosyltransferases converted the T antigen into a tetrasaccharide with an azido group, which can be tagged by a probe through click chemistry for visualization. MS-based proteomics was also performed to identify T antigen-bearing glycoproteins in MCF7 cells. The cell lysates were treated with the tandem chemoenzymatic labeling, and the glycoproteins were captured by alkyne resins, followed by trypsin digestion and LC–MS/MS analysis.

Lectins, including *Agaricus bisporus* lectin (ABL) and peanut agglutinin (PNA), were also utilized to isolate T antigen-containing glycoproteins. A tandem enrichment strategy was developed to identify T antigen-bearing glycoproteins on the surface of metastatic prostate cancer cells. Cell surface proteins were first labeled by sulfo-NHS-SS-biotin and then separated with streptavidin resins. The isolated surface proteins were further enriched with PNA-based chromatography, followed by MS analysis for protein identifications.[85] However, without site-specific information, the results may contain false positive identifications due to nonspecific binding. With Jacalin-based lectin weak affinity chromatography (LWAC), Valoskova et al. site-specifically characterized T antigen-bearing glycoproteins in *Drosophila* embryo and identified 219 T antigen-containing glycosylation sites on 106 glycoproteins.[78]

ST Antigen.—Recently, Wen described an elegant chemoenzymatic strategy to detect ST antigen-bearing glycoproteins.[86] They found that human sialyltransferase ST6GalNAc-IV was capable of specifically recognizing the ST antigen. At the same time, it had unrestricted donor specificity, which means that ST6GalNAc-IV can tolerate a sialic acid analog containing a bioorthogonal functional group. They labeled the ST antigen on the cell surface by transferring a biotin-containing Neu5Ac analogue using ST6GalNAc-IV. The labeled ST antigen on the cell surface was visualized by the fluorophore–streptavidin conjugate. In addition, proteins with the labeled ST antigen were pulled down by avidin resins (Figure 4A). After on-bead digestion and MS analysis, they identified 78 and 43 potential cell-surface glycoproteins containing the ST antigen from MCF7 and HT29 cells, respectively (Figure 4B). Based on PANTHER classification, proteins with binding activity and catalytic activity were highly enriched in both cell lines (Figure 4C). However, no site-specific information was available for ST-bearing glycoproteins. In another study, the ST antigen was converted to the T antigen through the removal of sialic acid with the neuraminidase treatment, followed by enrichment with PNA-based chromatography and MS analysis.[80] Although this method enabled site-specific analysis, the degree of sialylation was lost, and endogenous T antigen-containing glycoproteins may result in false positive identifications.

O-GlcNAc.

The O-linked β-*N*-acetylglucosamine (O-GlcNAc) modification on the serine and threonine residues is another important type of protein glycosylation.[87] O-GlcNAcylation is ubiquitous in cells and involved in many biological processes including cell signaling and transcriptional regulation.[88,89] Abnormal O-GlcNAcylation is correlated with various diseases including neurodegenerative disorders, diabetes, and cancers.[90] The developments

of modern MS-based proteomics and novel enrichment methods have greatly facilitated our understanding of protein O-GlcNAcylation.

Both lectin- or antibody-based affinity purification and chemical methods including β-elimination and hydrazide chemistry were reported for selective enrichment of O-GlcNAcylated proteins/peptides, which provided a wealth of valuable information about protein O-GlcNAcylation.[91,92] However, these methods suffer from the low specificity and weak binding affinity. Moreover, chemical methods based on β-elimination that normally take place under basic conditions cause the damage of the peptide backbone, thus affecting glycoprotein identification. In addition, the incompletion of the reaction and β-elimination from other types of modifications such as phosphorylation may also be an issue. Chemoenzymatic and metabolic labeling with a sugar analogue bearing a bioorthogonal functional group followed by click chemistry-based tagging with affinity probes has become powerful for O-GlcNAcylation profiling.[93–95] Photocleavable (PC) linker, such as alkyne-PC-biotin, was employed in a chemoenzymatic strategy for site-specific analysis of O-GlcNAcylation.[96] A new type of biotin cleavable linker containing 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl (DDE) group was found to have higher cleavage efficiency compared with the photocleavable linker.[97] The combination of chemoenzymatic labeling and polyethylene glycol (PEG) tags, followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) separation, enabled the measurement of O-GlcNAcylation stoichiometry (Figure 5A).[98] More recently, this method was further improved by employing strain-promoted 1,3-dipolar cycloaddition (SPAAC) to tag O-GlcNAcylated proteins (Figure 5B).[99]

Integrating metabolic labeling with an isotope-tagged cleavable linker (isoTCL), Qin et al. site-specifically identified and quantified protein O-GlcNAcylation.[94] The characteristic mass difference from the isotopic cleavable alkyne-biotin probe allowed for O-GlcNAcylation site identification and quantification because of the unique isotopic pattern in MS^1. They also applied this method to quantify O-GlcNAcylation in female and male mouse placentas and found that around 40% O-GlcNAcylation sites in the female placentas are displayed at a higher level compared with the male ones. Recently, a method based on a new linker, triarylphosphine functionalized trimethylpiperidine (TFT), for one-step selective enrichment of protein O-GlcNAcylation was reported.[100] The triarylphosphine functional group reacted with azide-containing O-GlcNAcylated proteins through the copper-free Staudinger ligation, and the trimethylpiperidine was specifically recognized by the anti-TMT antibody conjugated to agarose beads. The new reagent enabled the identification of 1706 O-GlcNAcylation sites in HeLa cells.

O-Mannose.

O-Mannosylation is the only known type of protein O-glycosylation in yeast. Using the DBA method, Xiao et al. site-specifically identified 234 O-mannosylated proteins with glycan information in yeast cells. Protein clustering revealed that O-glycoproteins located on the cell wall ($P = 4.25 \times 10^{-32}$) were the most highly enriched.[58] Neubert et al. used a yeast mutant lacking the O-glycan elongation enzyme to simplify the complexity of O-mannosylation, followed by LWAC enrichment and MS analysis.[101] They identified more

than 2300 O-mannosylation sites on over 500 proteins in yeast cells and found that half of the O-glycoproteins were the targets of protein O-mannosyltransferases localized in the ER.

Conversely, the knowledge on O-mannosylation in mammalian cells is still very limited. Compared with yeast, comprehensive analysis of O-mannosylation in mammalian cells is more difficult due to the greater structural complexity of O-mannosylation.[58,102] Vester-Christensen et al. used a SimpleCell strategy to simplify the structure of O-mannose glycans in human MDA-MB-231 breast cancer cells.[103] After the knockout of POMGnT1, which controls the first step in the elongation of O-mannose glycans, the peptides were treated with PNGase F to remove N-glycans and then O-mannosylated peptides were enriched by Con A lectin chromatography. They identified 235 O-mannosylation sites on 51 O-mannosylated proteins and found that O-mannosylation sites were mainly distributed on cadherins and plexins. While these large-scale studies of protein O-mannosylation reveal their functions in mammalian cells, comprehensive analysis of O-mannosylation without trimming the glycans in mammalian cells still remains challenging.

Mannose-6-Phosphate.

In mammalian cells, the mannose-6-phosphate (M6P) modification plays a critical role in the transportation of soluble lysosomal proteins. Most newly synthesized lysosomal proteins are delivered to the lysosome through the M6P pathway, in which the proteins undergo specific phosphorylation on mannose residues of high-mannose N-glycans for an efficient transportation from the Golgi to the lysosomes.[104] Comprehensive analysis of proteins with the M6P modification provides valuable information on the enzymes located in the lysosomes and establishes the correlation between malfunctions of lysosomal acid hydrolases and different human diseases. M6P-modified glycoproteins were purified by affinity chromatography based on mannose 6-phosphate receptors (MPRs) or ultracentrifugation of the lysosomes, followed by digestion and MS analysis.[105,106] Although these methods were used to analyze M6P-modified glycoproteins from a wide range of samples, M6P modification sites were not directly identified, which may result in false positive identifications. Recently, Fe^{3+}-IMAC chromatography was optimized to enrich M6P modified glycopeptides.[107] A specific diagnostic fragment ion of M6P generated during higher energy C-trap dissociation (HCD) was utilized to trigger electron-transfer/higher-energy collision dissociation (EThcD). With that method, 46 M6P glycosylation sites from 35 glycoproteins were identified in HeLa cells. The method was also applied to analyze M6P-bearing glycoproteins in CHO cells with a double knockout of the Acp2 and Acp5 phosphatases that are responsible for the removal of the phosphate group on M6P-modified glycans, and 160 unique glycopeptides containing M6P were identified, which was increased by 4-fold compared with those identified in wild type CHO cells.

COMPREHENSIVE ANALYSIS OF INTACT GLYCOPEPTIDES

In a typical bottom-up glycoproteomic study, especially for N-glycosylation analysis, glycopeptides are often deglycosylated to generate a common mass tag on the glycosylation sites before MS analysis.[8,9,108] Although the deglycosylation-centric strategies greatly simplify the MS identification and tremendously expand the knowledge of glycoproteins, the

removal of glycans located on glycopeptides results in the loss of the glycan information. Intact glycopeptide analysis obtains the information on both the glycosylation sites and glycan structures. However, it is challenging to achieve system-wide characterization of intact glycopeptides due to the heterogeneity of glycans and the low ionization efficiencies of glycopeptides compared with the nonmodified ones. The exploration of comprehensive analysis of intact glycopeptides was recently reported mainly because of the advances of effective fragmentation methods and novel bioinformatics tools.[109]

Complete characterization of intact glycopeptides requires sufficient fragments from both the peptide backbone and the attached glycans. However, the common fragmentation methods alone cannot simultaneously provide substantial details on the two components. Collision-induced dissociation (CID) and HCD mainly yield B- and Y-ions from the glycans attached to glycopeptides, while electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) primarily fragment the peptide backbone to generate c- and z-ions and leave the glycans intact.[110] Hybrid fragmentation, including a combination of different fragmentation methods for the same precursor ions such as EThcD,[111,112] and product-ion triggered fragmentation strategies such as HCD product-ion dependent EThcD (HCD-pd-EThcD),[113] were shown to be beneficial to characterize intact glycopeptides. In an EThcD spectrum (Figure 6B), more information on the peptide backbone and the glycan was obtained compared with CID or ETD (Figure 6A).[114]

Interpretation of complex tandem mass spectra generated from intact glycopeptides requires the development of effective search tools. Until now, a variety of software has been developed to help identify intact glycopeptides, including Byonic,[115] GPQuest,[116] pGlyco,[117] SugarQb,[118] and GPSeeker.[119] Developments of instrumentation and bioinformatics were reviewed previously.[109,120–122] To avoid the overlap with previous reports, we emphasize the large-scale characterization of intact glycopeptides with advanced fragmentation methods and software.

Extensive efforts have been made to comprehensively analyze the N-glycoproteome with the glycan structure information.[123] The development of more powerful instrumentation and search tools has facilitated the increase of the N-glycoproteome coverage in recent years. An integrated workflow, including an optimized MS/MS acquisition method and a new search tool for quality control of intact glycopeptide identifications, was reported to globally characterize intact N-glycopeptides.[117] Glycopeptides were enriched by ZIC-HILIC and then analyzed by MS with stepped collision energies (SCE)-HCD. The use of stepped collision energies generated more abundant information from both the glycan and the glycopeptide backbone within one scan. The authors also developed a search tool called pGlyco 2.0, which enabled the estimation of the false discovery rates (FDRs) of glycans, peptides, and glycopeptides at the same time. They identified 1988 glycosylation sites from 955 glycoproteins in five mouse tissues.

More recently, pGlyco 2.0 was applied to study intact N-glycopeptides from the APP/PS1 mouse model for Alzheimer’s disease (AD) and wild type mouse, and a total of 3524 intact N-glycopeptides were characterized.[124] Integrating the identification of intact N-glycopeptides, quantification of the whole proteome and the N-glycoproteome, and lectin

microarray for glycan epitopes, the authors provided a multilayer N-glycoproteome analysis in the APP/PS1 versus wild type mice. It was found that oligo-mannose and fucosylated N-glycans were highly expressed in the brains of both types of mice. N-Glycosylation of most membrane proteins including glutamate receptors were down-regulated in the APP/PS1 mouse, which indicated that the dysfunctions of N-glycoproteins may affect the development of AD.

Multilectin weak affinity chromatography (M-LWAC) and different fragmentation methods were combined to comprehensively analyze intact N-glycopeptides in human serum and brain tissue samples.[125] To minimize the bias caused by lectin-based enrichment methods, an M-LWAC column immobilized with six types of lectins was employed to enrich glycopeptides. Then glycopeptides were analyzed by HCD and EThcD independently, which provided complementary information on the glycopeptide backbone and glycan structures. pGlyco and Protein Prospector were utilized to interpret tandem mass spectra generated by HCD and EThcD, respectively. They identified 7503 intact N-glycopeptides from 666 glycoproteins in human serum and brain samples. Comparing the N-linked glycoproteomes in the two types of tissues, they found different distributions of high-mannose N-glycans, in which ~35% of the identified glycopeptides from the brain samples possessed high-mannose glycans, while only ~15% occurred in the serum. Nonetheless, glycan sialylation appeared more frequently in the serum. They also found that the number of glycans identified at one site was proportional to the number of glycans at the remaining sites on one protein, which indicated that protein structures may control the glycan microheterogeneity at each glycosylation site. The investigation of high-mannose glycan structures processed through the canonical glycan biosynthesis pathway revealed many differences between the two types of tissues, which demonstrated that intact glycopeptide analysis can be potentially applied to study the glycan processing pathways and measure glycan metabolic changes during disease development.

Infrared (IR) photoactivation coupled with electron transfer dissociation (AI-ETD), which uses IR photoactivation concurrent with ion-ion reaction, has emerged as a highly effective fragmentation method for proteomic applications.[126] AI-ETD was employed for an in-depth study of intact N-glycopeptides from the mouse brain tissue.[127] The authors first evaluated the performance of AI-ETD on the characterization of glycopeptides and demonstrated that the integration of vibrational and electron-driven dissociation methods was capable of providing abundant information on both the glycan and the glycopeptide backbone. A total of 5662 unique N-glycopeptides with 1545 sites on 771 glycoproteins were identified using Byonic as the search tool. Compared with EThcD, AI-ETD was faster because of the concurrent vibrational activation and electron-driven dissociation. Therefore, it enabled more scans and resulted in more glycopeptide identifications. With the large data set of intact N-glycopeptides, they systematically investigated the site-specific microheterogeneity of protein N-glycosylation. High-mannose glycans were the most prevalent, while fucosylated, paucimannose, and sialylated glycans preferred to be on proteins with multiple glycosylation sites (Figure 7A). High-mannose glycans appeared to be on the same glycosylation site simultaneously and also co-occurred with complex/hybrid, fucosylated, and sialylated glycans (Figure 7B). Protein clustering based on cellular component revealed that the diversity of glycans on the plasma membrane, other membranes, and extracellular

proteins was increased substantially. Proteins located in the lysosome also showed a high occurrence of M6P, which is consistent with the M6P synthesis pathway for lysosomal protein trafficking.[128]

Compared with N-glycosylation, system-wide characterization of intact O-glycopeptides is less frequently reported. Studies of O-glycosylation mainly focus on simplifying the glycan structures to facilitate the enrichment of glycopeptides and MS identifications.[129] Neuraminidase and various glycosidases were employed to trim O-glycans before enrichment and characterization. The SimpleCell technology was proven to be very useful for mucin-type O-glycosylation analysis by blocking O-glycan elongation.[130] Although O-glycan simplification strategies help the identification of glycosylation sites, the loss of the information on glycan structures is the cost. Methods for intact O-glycopeptide analysis are urgently needed to decipher the functions of protein O-glycosylation.

Qin et al. reported an integrated method to comprehensively analyze intact mucin-type O-glycopeptides by using HILIC enrichment, beam-type CID on a Q-TOF MS, and an *in silico* deglycosylation approach to interpret the recorded spectra.[131] The *in silico* deglycosylation approach enabled a “deglycosylated” peptide search by removing the peaks from the spectra generated by glycosylated peptide fragments (Y ions). The combination of the mass difference between a precursor and the Y0 ion (the precursor ion with the loss of glycan) and the presence of oxonium ions allowed for the identification of glycoforms. They identified 407 intact O-glycopeptides from 93 glycoproteins with various glycan compositions, and over 80% of the glycans were sialylated, which revealed the high sialylation of proteins in human serum and showed the heterogeneity of O-glycans.

Recently, a similar search tool, named O-Search, was developed to identify intact O-glycopeptides from the spectra generated by HCD on an Orbitrap MS.[132] Compared with N-glycans, the glycosidic bond between O-glycans and peptides is relatively easily broken during MS fragmentation, which makes the localization of O-glycosylation even more difficult. The development of EThcD provides an opportunity to identify O-glycopeptides site-specifically. In one study, Zhang et al. presented global and site-specific analysis of intact mucin-type O-glycopeptides.[133] They noticed that EThcD provided more abundant fragments of the glycopeptide backbone, which increased the confidence of site localization and offered more information on the glycan structures. Combining EThcD and multiple-enzyme digestion with multidimensional separation including lectin, size-exclusion chromatography, and HILIC, they identified 499 unique intact O-glycopeptides containing six types of glycan compositions with 173 O-glycosylation sites from 49 O-glycoproteins in human serum.

Ultraviolet photodissociation (UVPD) utilizing UV photons with high energy to excite the precursor ions for dissociation is another promising method for the characterization of protein N- and O-glycosylation.[134] Although UVPD showed the capacity to decipher the O-glycopeptide backbone and the attached glycan structures at the same time, its performance on large-scale characterization of intact O-glycopeptides still needs to be systematically evaluated. Furthermore, automated data interpretation of intact O-glycopeptide spectra generated by the hybrid fragmentation methods is quite challenging. The performance of

two suites of popular search software for O-glycopeptide identification, Byonic and Protein Prospector, was systematically evaluated for the characterization of intact O-glycopeptides from urinary samples.[135] The results revealed that the identification rate was low, and thus it is necessary to develop new search tools to more effectively interpret the spectra of intact O-glycopeptides. It is also imperative to increase the sequence coverage of proteins because O-glycosylation sites are usually clustered on glycoproteins, especially for mucin-type O-glycosylation, which prevents efficient digestion by the commonly used proteases such as trypsin. The newly reported O-proteases, including secreted protease of C1 esterase inhibitor (StcE)[136] and OpeRATOR[61] that cleaves N-terminally the O-glycosylated serine/threonine residues, offer new possibilities for intact O-glycopeptide analysis.

Although bottom-up glycoproteomics allows for in-depth glycosylation analysis, it could not provide a whole picture of individual glycoproteins. Therefore, intact glycoprotein analysis has some advantages over the glycopeptide-centric approaches because it is more powerful to identify different proteoforms of glycoproteins, which enables the analysis of one or multiple PTMs simultaneously and provides unique information that cannot be readily obtained by the glycopeptide-centric approaches. In one study, a combination of middle-down proteomics and native MS allowed for comprehensive characterization of different proteoforms of human erythropoietin and human plasma properdin.[137] For the middle-down analysis, proteins were digested with enzymes, followed by HCD-pd-CID and HCD-pd-EThcD MS analysis, to identify glycosylation sites and glycoforms. The middle-down proteomics data was then used to interpret the complex native MS spectra and helped identify three new C-glycosylation sites on properdin. In another study, a lectin purification method coupled with native MS was employed to characterize the interactions between lectin and specific glycoproteoforms at the intact protein level. Multifucosylation was found to attenuate the interactions between *Phaseolus vulgaris* leucoagglutinin (PHA-L) and haptoglobin but did not affect α1-acid glycoprotein-PHA-L binding.[138] The approach for intact glycoprotein analysis represents a promising direction to comprehensively characterize complex glycoproteins and decipher their biological functions.

APPLICATIONS OF MS-BASED GLYCOPROTEOMICS IN BIOLOGICAL SYSTEMS

Innovative enrichment methods, advanced instrumentation, and computational tools have been reported to facilitate the investigation of the glycoproteome by MS. These developments allow for effective analysis of glycoproteins in different biological systems considering that glycans and glycoproteins have emerged as important players in clinics and health science. In this section, we review some recent applications of glycoproteomic analysis, starting with the use of MS to understand the functions and properties of glycoproteins in biological contexts, followed by their applications for studying glycoproteins from different biological sources.

Deciphering Glycoprotein Functions and Properties.

Glycosylation has profound impacts on protein functions and properties. With the technological advances in glycoprotein analysis, their functions, properties, and behaviors

have been studied in various biological settings. Glycosylation regulates the trafficking of proteins from the ER to the Golgi apparatus and other subcellular compartments.[139] N-Glycosylation was found to serve as a possible lock mechanism for folded protein structures and conformations and protect proteins from degradation.[140] Mucin-type O-glycosylation is involved in protein secretion and stability and plays crucial roles in cellular development and organ homeostasis.[141,142] Some O-glycans participate in the immune system and cellular recognition, such as the A, B, or O antigen present on red blood cells.[143] For both N-linked and O-linked glycosylation, while they are well-studied in eukaryotes, these modifications in prokaryotes were discovered not a long time ago. Their functions, however, have some differences from those in eukaryotes, which are mainly involved in the cell integrity, particularly in the S-layer of archaea, response to environmental stress, and host-microbe interactions, as reviewed recently by Eichler and Koomey.[144]

The O-GlcNAc modification was not studied extensively in the past due to the low abundance and the dynamic nature of the modification. Early studies unraveled its roles in cell signaling and the crosstalk with phosphorylation.[145] O-GlcNAcylation was found to participate in protein homeostasis and be involved in neurodegenerative diseases such as AD.[146] Combining MS and molecular biology techniques such as RT-qPCR and ChIP, Machacek et al. identified the possible link between O-GlcNAcylation and inflammation, with the elevated O-GlcNAcylation level resulting in the increased production of IL-17A.[147] O-GlcNAcylation was also linked to cancer metastasis and progression. In a study from Lo et al., O-GlcNAcylation was found to contribute to the stability of histone methyltransferase EZH2. Mutation of some O-GlcNAcylation sites on this protein decreased its stability, especially near the C-terminus that reduced its histone di- and trimethylation and may inhibit tumor progression.[148] O-GlcNAcylation can also influence the resistance or sensitivity of cancer cells in both cell culture and xenograft models to proteasome inhibitors by activating the binding of the transcription factor NRF1, and therefore stabilizing it, to enhance or suppress the production of the proteasome subunits.[149]

The addition of the sialic acid or fucose residues is frequently found in some glycans. Sialylated glycans are involved in cell fate decision, embryo development, and disease progression.[150,151] Terminal sialic acid modification is known to be enriched on the cell surface of highly metastatic cancer cells.[152] Its cellular metabolism is also significantly up-regulated in invasive breast tumors.[153] The protein identity, cell type and status, and cell culture conditions may affect the branches of glycans. Systematic changes of glycan structures based on factors such as diets and cell culture environments were observed, including hypersialylation when cells were grown in acidic conditions, and increases in high-mannose glycans when short chain fatty acids were used in the medium or when extracellular fructose, galactose, and glutamine concentrations were elevated.[154] These modifications by sialic acid or fucose may also affect the binding of proteins to other substances such as proteins and drugs.[155] Toonstra et al. employed PNGase F and sialidase to remodel the glycosylation profiles of platelet proteins before affinity purification with a collagen-immobilized column. Collagen is known to activate platelets, and its interactions with platelet proteins were altered, suggesting the importance of glycans in protein–protein interactions.[156] Core fucosylation is also interesting, and its elevation was related to diseases such as mucinous adenocarcinoma.[157]

The subtle difference in the glycan structures is considered as the microheterogeneity of protein glycosylation. Qin et al. developed a "virtual MS^3" method where the same intact glycopeptides and deglycosylated peptides were analyzed in the same run. After retention time adjustment, the site-specific glycoforms were determined as well as the quantification at both the glycosite and site-specific glycoform levels.[158] Researchers have also investigated the site occupancy of protein glycosylation. In one study comparing the proteomes and glycoproteomes of benign and malignant ovarian cancer cells, many glycosylation site occupancies varied while the expression of many parent proteins did not change.[159] Yang et al. employed the data-independent acquisition (DIA) approach via sequential window acquisition to determine the stoichiometry of N-linked glycosylation from cultured cells.[160] In this approach, glycopeptides were enriched using a lectin-affinity method and subsequently deglycosylated with PNGase F. Peptides from both groups were analyzed with SWATH MS and the stoichiometries of 2274 glycosylation sites were determined from two cell types. While the stoichiometries varied from 0 to 100%, the majority of the sites had low stoichiometry values.

While there is no sequence motif for O-glycosylation sites, the well-known and canonical motif for N-glycosylation sites is N-X-S/T. In one of the earlier studies by Zielinska et al., several other motifs were also observed including N-X-C, NG, and N-X-V.[33] Interestingly, the reverse N-glycosylation consensus motif S/T-X-N was also reported.[161] So far, only a few of these motifs have been confirmed by other orthogonal methods such as the glycosylation on the N-X-C motif observed with crystallography.[162]

Advances in MS-based protein quantification also benefit the field of glycoproteomics. Large-scale and relative quantification of proteins using MS can be achieved via label-free or label-based techniques.[163] In particular, the development of isobaric tagging reagents allows for simultaneous quantification of proteins from multiple samples, which can increase the throughput and quantification accuracy and enables a time-course analysis of glycoproteins under particular conditions. Higher-order multiplexing may be achieved by combining TMT (MS^2-based quantification) with SILAC (MS^1-based quantification).[164–166] Stadlmann et al. separated glycopeptides by HILIC from different sister clones of mouse embryonic stem cells before TMT labeling to compare the abundance changes of intact glycopeptides. The study resulted in the discovery of glycoproteins that participated in the cellular resistance to ricin toxicity.[167] Xiao et al. combined pulse-chase labeling and multiplexed proteomics using TMT with glycoprotein enrichment to simultaneously determine the synthesis and degradation rates of over 700 glycoproteins from MCF7 cells.[168] In the near future, with the development of higher-plexed reagents and approaches, simultaneous analysis of more samples will be achievable.

Identification of Glycoproteins as Potential Biomarkers.

There are generally three main steps when developing new protein biomarkers, i.e., discovery, verification, and validation.[169] Discovery-based proteomics by MS is powerful to screen potential candidates in the early phase, followed by further verification and validation using orthogonal methods in the later steps. This may be accomplished by comparing cells or tissues in different states to pinpoint differentially regulated proteins. With the advances

in glycoproteomic analysis, biomarker discovery has shifted toward identifying glycoproteins or glycosylation events as biomarkers for early disease detection or monitoring disease treatment.[170–175] Several FDA-approved cancer biomarkers are glycoproteins, such as thyroglobulin (TG), human epidermal growth factor receptor 2 (HER2/neu), alpha-fetoprotein (AFP), human epididymis protein 4 (HE4), cancer antigen 15-3 (CA15-3, MUC1), prostate-specific antigen (PSA), cancer antigen 125 (CA125, MUC 16), and carcinoembryonic antigen (CEA).[176] Even though they have been approved by the FDA, some current biomarkers have the selectivity and specificity issues for disease detection, especially at early disease stages.[177] If the glycoforms are considered, the results will be more reliable as shown in the case of PSA.[178,179]

In this section, we focus on glycoproteomic analysis in biomarker discovery from whole tissues or cell culture models. According to the recent estimation by the American Cancer Society, lung and bronchial cancers have the highest mortality rates among cancer patients in the United States, followed by prostate, colon, and rectum cancers.[180] Using iTRAQ, Yang et al. quantified proteins and glycoproteins in 18 primary lung squamous cell carcinoma, adenocarcinoma, and healthy controls. Several glycoproteins enriched by hydrazide chemistry were differentially regulated among the tumors, such as procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2), calumenin (CALU), and periostin (POSTN).[181] Waniwan et al. used lectin-conjugated magnetic nanoparticles to enrich glycopeptides from sensitive and drug-resistant nonsmall cell lung cancer cells, and the higher frequency of protein fucosylation in the drug-resistant cancer cells was found.[32] Jia et al. recently reviewed and highlighted increased fucosylation in lung cancer progression.[182] The elevation of fucosylation was also reported in other types of prostate cancer. In one study where the proteomic and glycoproteomic profiles of LNCap and PC3 prostate cancer cell lines were compared, the site occupancy differences among the two types of cells were observed.[183]

In colon cancer, mucin-type O-glycosylation was reported to be altered, which was supported by a recent study showing the up-regulation of *N*-acetylgalactosaminyltransferase 6 (GalNAc-T6) in colon adenocarcinoma but not in adjacent normal tissue.[184] Sinha et al. enriched glycoproteins from tissues and sera of patient-derived xenografts with high-grade serous ovarian carcinoma and identified potential glycoprotein biomarkers that were further verified by parallel reaction monitoring (PRM).[185] In melanoma, the overexpression of *N*-acetylglucosaminyltransferase III (GnT-III) in metastatic cells caused the increase of N-glycans containing bisecting GlcNAc.[186] The up-regulation of fucosyltransferase FUT8 resulted in the elevation of core fucosylation, which was involved in metastatic melanoma.[187] Differential expressions of Tn antigen-bearing glycoproteins were also observed in cutaneous primary and metastatic melanoma cells.[188]

Aberrant glycosylation was found in several neurodegenerative diseases as well. In AD, aberrant glycosylation, particularly O-GlcNAcylation, was linked to disease progression.[189] By combining TMT labeling with a chemoenzymatic approach to enrich O-GlcNAcylated peptides from post-mortem brain tissues, Wang et al. detected 1094 O-GlcNAcylation sites, with 131 peptides from 81 proteins being differentially regulated in patients with AD.[190] Lamoureux et al. employed lectin affinity enrichment of glycoproteins from the brain tissues

of mice with transmissible spongiform encephalopathies, known as the mad cow disease in cows, and identified four glycoproteins that were regulated compared with the aged-matched controls.[191]

Biological and clinical samples are often preserved in the formalin-fixed paraffin-embedded (FFPE) platform, which may affect the glycoprotein stability. Earlier, Tian et al. used hydrazide chemistry to enrich N-linked glycoproteins from mouse FFPE and frozen lung tissues, with an extra deparafinization step performed for the FFPE tissue. While the abundance of glycopeptides from the FFPE tissue was lower, the identification was not much affected.[192] In another study, the proteome, phosphoproteome, and N-glycoproteome were found to be quantitatively preserved.[193] The glycomes of the FFPE and frozen tissues were also compared and only minor differences were found.[194] This increases the variety of samples available, especially in clinics, for MS-based glycoproteomic analysis.

Cell-Surface Glycoprotein Analysis.

The surface of eukaryotic cells is typically covered with sugars that are attached to various embedded proteins and lipids. This thick layer of glycoproteins and glycolipids, including N-linked and mucin-type O-glycoproteins, glycosylphosphatidylinositol (GPI)-anchored proteins, proteoglycans, and glycosphingolipids, surrounding the cell is called the glycocalyx.[195] Surface glycoproteins are generally synthesized through the classical secretory pathway and modified into the mature glycoforms by many enzymes in the ER and Golgi apparatus that add or remove sugars in the glycan moeity.[143,196] These cell-surface glycoproteins participate in many intra- and extracellular activities, including cell–matrix adhesion, cell–cell interaction, cell migration and motility, signal reception and activation of intracellular signaling pathways, vesicle-mediated transport, and molecule transportation across the plasma membrane.[197,198] The presence of specific surface glycoproteins can be used as a marker for the classification of cell types.[199] Aberrant protein glycosylation or abundance changes of surface glycoproteins can reflect the cellular statuses.[200,201]

Surface glycoproteins are conventionally studied by employing antibodies to specifically target proteins of interest. This is normally coupled with methods such as fluorescence microscopy or flow cytometry in immunophenotyping.[202,203] Besides the low throughput of these methods, the availability and specificity of antibodies could be an issue. MS can be exploited to alleviate these problems because it can detect thousands of proteins with high confidence in one experiment. Furthermore, antibodies are not required, and thus no prior knowledge is needed for surface protein detection. However, many surface glycoproteins have low abundance, and it is challenging to distinguish surface glycoproteins from those inside cells. Therefore, selective separation and enrichment of surface glycoproteins are critical prior to MS analysis.

One powerful method to enrich surface glycoproteins is through hydrazide chemistry. In 2003, Zhang et al. developed a highly innovative MS method for surface glycoproteomic analysis based on this approach.[8] They identified 104 unique peptides from 64 surface glycoproteins in LNCaP cells. Later on, the Cell Surface Capture (CSC) method was reported in 2009, where glycans were first oxidized with sodium periodate and a biocytin-hydrazide tag was employed to generate a chemical handle for the following enrichment

with streptavidin beads.[204] About 100 surface glycoproteins were identified and quantified in each experiment. The CSC method was also applied to study surface glycoproteins in 41 human and 31 mouse cell types, and on average, 284 surface glycoproteins were identified from each cell type.[205] The data were compiled for the Cell Surface Protein Atlas (CSPA) database. CSC has been modified into an automated system with smaller number of cells required, termed autoCSC.[206]

Chemical oxidization of cis-diol groups using the oxidants such as sodium periodate in hydrazide chemistry-based methods may affect the cell viability. Minimizing cell death is critical for surface glycoprotein analysis because many highly abundant intracellular proteins could leak out and interfere with the tagging and enrichment of surface glycoproteins. Recently, a milder approach using galactose oxidase was reported. Ramya et al. employed this method, together with aminooxy-biotin tagging and streptavidin enrichment, to identify 68 glycoproteins on the cell surface.[207] We optimized this method and coupled the enzymatic reaction with hydrazide chemistry-based enrichment (Figure 8). The approach was further improved by the pretreatment of cells with neuraminidase to remove the terminal sialic acid residues. Moreover, the addition of horseradish peroxidase during the oxidation reaction, which consumes hydrogen peroxide (one of the oxidation products), pushes the reaction to completion.[45] Using this approach, we identified, on average, 953 N-glycosylation sites from 393 surface glycoproteins per experiment in MCF7 cells. Combining with SILAC, the approach enabled us to quantify 909 unique N-glycopeptides from 334 surface glycoproteins with 65% being down-regulated by over 2-fold in cells treated with brefeldin A, which inhibits protein secretion through the classical secretory pathway.

Another approach to analyze cell-surface glycoproteins is through metabolic labeling with sugar analogues.[62,208] This method results from the development of bioorthogonal chemistry, where a specific chemical reaction in biological systems is achieved through functional groups that do not naturally exist or interfere with normal biological activities.[209,210] It allows for selective studies of biomolecules including glycoproteins on the cell surface.[211,212] The sugar analogue is usually peracetylated to increase its passive diffusion rate into cells, such as N-azidoacetylmannosaminetetraacylated (Ac_4ManNAz). The acetyl groups are deacetylated by intracellular esterases into azidosialic acid, which is incorporated into glycoproteins including the surface ones, through the Roseman–Warren biosynthetic pathway.[213] ManNAz was successfully applied to identify sialylated surface glycoproteins in both cultured cells and model animals such as the mouse.[211,214] Notably, metabolic labeling with sugar analogues is very compatible with cultured cell because cells can be grown in a medium containing these sugar analogues and incorporate them into their surface glycoproteins. Furthermore, Spiciarich et al. applied this method to study human prostate cancer tissue *ex vivo* through a tissue slice culture, in which the tissues stayed metabolically active for days.[215] The tissues were cultured in the presence of Ac_4ManNAz, and surface glycoproteins were tagged with the biotin-alkyne reagent before enrichment with avidin resins. Over 900 proteins were detected from the normal and cancerous prostate tissues. Among those, 68% were membrane or secreted proteins and 45% were known glycoproteins.

Metabolic labeling with Ac_4ManNAz may be applied to determine the glycan–protein interactions. Li et al. developed a beautiful method called protein oxidation of sialic acid environments (POSE) to study surface protein interactions (Figure 9).[216] Cells were first metabolically labeled and tagged with dibenzocyclooctyne (DBCO)-(S)-1-(*p*-bromoacetamidobenzyl)ethylenediaminetetraacetate (FeBABE), followed by hydrogen peroxide treatment and quenched with methionine amide hydrochloride. The treatment with hydrogen peroxide generated radical species that can oxidize proteins nearby. The study identified 150–200 proteins that were oxidized from each cell line. Interestingly, they also evaluated the incorporation efficiency of azidosialic acid into glycans. While the incorporation rate in PNT2 cells was as high as 87%, it was only 18% in Caco-2 cells. The conjugation efficiency between DBCO and azide was estimated to be over 86%. The variation of the incorporation efficiency needs to be considered in metabolic labeling with the sugar analogue among different cell types in order to increase the quantification accuracy.

Another concern with Ac_4ManNAz labeling is its effects on cell viability and proliferation. In two recent studies from Han et al., cells were treated with different concentrations of Ac_4ManNAz, Ac_4GlcNAz, and Ac_4GalNAz from 10 to 50 μM for 72 h.[217,218] Higher concentrations of Ac_4ManNAz were found to affect biological processes such as the MAPK activity, apoptotic process, and immune and inflammatory response according to transcriptomic analysis. The authors suggested that labeling with 10 μM Ac_4ManNAz should be sufficient. However, the labeling time and the cell type need to be further considered. Our lab previously compared the three sugar analogues in a surface glycoprotein identification experiment and found that labeling with 100 μM Ac_4GalNAz resulted in the highest coverage of surface glycoproteins from HepG2 cells.[219] We also compared the labeling with 10–250 μM Ac_4GalNAz, 100 μM Ac_4GalNAc, and a vehicle control group in A549 cells. Within the incubation time of 24 h, nearly all quantified proteins in these tested conditions were not affected by Ac_4GalNAz treatment (unpublished data).

Ac_4GalNAz has been used in several applications for surface glycoprotein analysis.[208] For example, Xiao et al. combined the method with pulse-chase labeling to study the dynamics of cell-surface glycoproteins and measure their half-lives.[220] Cells were first labeled with Ac_4GalNAz and then tagged with DBCO-biotin before being switched to the medium without the sugar analogue. The cells were collected at different time points, and the tagged surface glycoproteins were separated. After digestion and enrichment, the glycopeptides were labeled with the TMT reagents. It was found that the half-lives of surface glycoproteins were generally longer than those of newly synthesized proteins, which may be due to the presence of glycans that can protect proteins from being degraded.

Other methods target all proteins on the cell surface even though the majority, if not all, of surface proteins are glycosylated. For membrane separation by ultracentrifugation, contamination of membrane proteins from other cellular compartments could be an issue.[221,222] Alternatively, surface proteins may be directly biotinylated for further separation. In a study from Hormann et al., three surface protein purification methods were compared, including sulfo-NHS-SS-biotinylation, aminooxy-biotinylation, and surface coating with silica beads. Sulfo-NHS-SS-biotinylation outperformed the other two in identifying surface

proteins, with the localization of over 90% of the proteins on the plasma membrane.[223] In the same study, instead of a typical avidin denaturation elution by SDS, elution through biotin competition further increased the coverage.

The development for cell-surface glycoprotein analysis also benefits other research fields, especially biomedical areas. For example, using CSC, surface glycoproteins were identified from four human lymphocyte cell lines and human induced pluripotent stem cells, respectively, and the results may be used in cell type classification and drug discovery.[224,225] Matta et al. employed a method similar to CSC, but the aminooxy-biotin tag was used, to compare the surfaceomes of chondrogenic progenitor cells and bone marrow mesenchymal stem cells. Even though the two types of cells are very similar, distinct populations of surface glycoproteins were observed.[226] Kalxdorf et al. successfully combined chemical oxidation of glycans with alkoxylamine-PEG_4-biotin tagging and multiplexed proteomics to analyze changes of the surfaceome during macrophage differentiation over 72 , as well as the effects of a kinase inhibitor on the cell differentiation.[227] Zarif et al. isolated CD14+ monocytes from human blood and differentiated them *in vitro* into pro-inflammatory and anti-inflammatory macrophages. Using solid-phase extraction of N-linked glycopeptides (SPEG) by glycan oxidation and glycopeptide enrichment with hydrazide beads, they specifically identified glycoproteins on the surface of anti-inflammatory macrophages that may be involved in prostate cancer proliferation and metastasis, as opposed to the pro-inflammatory ones.[228] Another field that benefits from the advances in surface glycoprotein analysis is immunotherapy. It is well-known that the expression of specific surface proteins is crucial for the effectiveness of the immunotherapy,[229] and thus global analysis of surface glycoproteins may lead to the discovery of new drug targets for disease treatment.

Although we can currently detect a few hundred surface glycoproteins in a single experiment, the total number of glycoproteins present on the cell surface still remains elusive. Bausch-Fluck and Goldman et al. employed machine learning to predict if a glycoprotein can be localized on the surface based on the experimentally derived surface glycoproteins from the CSPA database and those from inside of cells in other subcellular compartments.[230] The predictor, called SURFY, scores proteins based on factors determined from a random forest classifier, including the frequencies of the N-X-S/T motif and cysteine residues, the presence of C-glycosylation site, and the length and number of transmembrane domains. This approach predicted a total of 2886 proteins that could be present on the surface of human cells with the accuracy of 93.5%, corresponding to 14.3% proteins in the human proteome that may be present on the cell surface.[230] When the results were matched with transcriptomic data from an RNA-seq experiment, only a subset of these proteins was identified in a specific cell type, such as 507 proteins from HeLa cells, corresponding well with the number of surface glycoproteins typically identified from experiments.

Secreted Glycoprotein Characterization.

Cells secrete many biomolecules including proteins, lipids, and other small molecules for cell–cell communication, cell–matrix interactions, and manipulation of the surrounding microenvironment.[231,232] Many proteins that enter the classical secretory pathway may be released outside of the cells. Secreted glycoproteins containing much valuable information

on their origin cells and thus can serve as noninvasive biomarkers for disease detection, diagnosis, and cell or tissue type classification.[233]

For *in vitro* experiments, secreted glycoproteins are analyzed from the cell culture medium. The major hurdle for global analysis of secreted glycoproteins is their low abundance compared with the background proteins such as those in the serum required for the cell growth. Typically, cells are initially grown in a serum-containing medium before being switched to the serum-free one. This approach was applied to analyze secreted proteins from cultured mouse macrophages treated with lipopolysaccharides (LPS), where 775 secreted proteins were identified.[234] However, serum-free media may affect cell growth and proliferation, which could result in cell death and release of highly abundant, intracellular proteins. In an early study by Eichelbaum et al., total secreted proteins were analyzed using a combination of bioorthogonal noncanonical amino acid tagging (BONCAT) and pulsed-SILAC in the serum-containing medium. The authors quantified 684 secreted proteins and also noted differential secretion of proteins after only 3 h when cells were grown in the serum-free medium.[235] In a recent study, cells were grown in the serum-free or serum-containing medium, respectively, and secreted proteins were compared using a similar approach reported by Eichelbaum et al. Combining with *in silico* prediction, the approach revealed that the abundance of truly secreted proteins detected from the serum-containing medium was generally higher than that from the serum-free medium. Notably, U87MG cells secreted 77 proteins exclusively in the serum-containing conditions and 93.5% of these proteins were predicted to be truly secreted proteins, while only 39.7% of 58 proteins exclusively from the serum-free medium were predicted to be truly secreted.[236] This highlighted the importance of preserving the optimal growth conditions for secretome analysis. The volume of the medium required for cell growth is also generally large in order to detect low-abundance secreted glycoproteins. In a recent secretome study from induced pluripotent stem cells, the growth conditions and sample preparation were optimized, and over 500 secreted proteins were detected from only 500 μL of medium.[237]

Apart from the cell culture model, secretome analysis may be performed *ex vivo* or *in vivo*. Kristensen et al. quantified glycoprotein changes in the secretome during *ex vivo* osteoblast differentiation of human stromal stem cells over 14 days using SILAC. One of the glycoproteins detected was stanniocalcin 2 (STC2), which was found to enhance the differentiation as an autocrine/paracrine factor.[238] A study from Kuljanin et al. expanded donor-derived human multipotent stromal cells and collected the conditioned media *ex vivo*. Almost all 16 secreted proteins identified that could be used to predict the regenerative potency of the cells are glycoproteins.[239] For *in vivo* studies, proteins may be collected from several types of bodily fluids, such as blood, plasma, urine, saliva, nipple aspirate fluid, and cerebrospinal fluids.[240–242] Reduction of the protein complexity by a trap column or immunoaffinity depletion to remove abundant proteins before further analysis may be performed to increase the coverage of secreted glycoproteins.[243] Multidimensional fractionation is another effective way to increase the coverage of secreted proteins. In a study from Peng's lab, 2-dimensional fractionation together with TMT labeling was performed to achieve one of the deepest proteome analysis from nondepleted sera of normal and AD patients. After being labeled with the TMT reagents, peptides were fractionated into 180 fractions using high-pH HPLC. Each fraction was then analyzed with a 3-h gradient

during LC–MS/MS analysis. They identified 4826 proteins, among which 30 proteins may be related to the disease progression.[244]

Targeted proteomics such as selected reaction monitoring (SRM) and multiple reaction monitoring (MRM) has also been employed to analyze glycoproteins of interest.[245,246] Lim et al. employed PRM to quantify the abundance of tryptic fucosylated peptides from human hepatocellular carcinoma sera. The level of several fucosylated peptides increased in patients with cirrhosis compared with the healthy controls, showing their potential use as biomarkers.[247] N-Glycans were analyzed in the interstitial fluids collected from breast cancer patients and the healthy controls, and distinct patterns of N-glycans were clearly observed, i.e., the tumor interstitial fluids contained high levels of high-mannose glycans, low bisecting N-glycans, and low complex N-glycans, which may be used for biomarker development.[248] Interestingly, N-linked glycosylation has been linked to neuropsychiatric disorders. Glycan analysis of sera from two mouse groups, one exposed to chronic unpredictable mild stress to induce depressive-like behaviors and the other serving as a control group, showed different N-glycan profiles that were correlated with their behaviors.[249]

For the analysis of secreted glycoproteins in clinical and biomedical research, several types of bodily fluids can be easily collected as mentioned previously. Belczacka et al. collected urine samples from 969 patients with bladder, prostate, and pancreatic cancers, cholangiocarcinoma, and renal cell carcinoma and identified three O-glycopeptides and five N-glycopeptides that were differentially regulated in those different types of cancer.[250] Sajic et al. analyzed 284 blood samples from patients with five types of localized-stage carcinomas, including colorectal, pancreatic, lung, prostate, and ovarian cancers and healthy controls (Figure 10A). Glycoproteins from blood were first conjugated to a solid support by hydrazide chemistry, and PNGase F was employed to deglycosylate glycopeptides before MS analysis. About 200 N-linked glycoproteins were detected, and the presence of potentially specific biomarkers, such as polymeric immunoglobulin receptor (PIGR) for pancreatic cancer, and common biomarkers for the tumors were identified (Figure 10B).[242] In a study by Togayachi et al., while secretogranin III was found in the secretome of all types of lung cancer, the low molecular weight, short form secretogranin III with two fucosylation sites was only found in small cell lung carcinoma and may be used as a biomarker.[251]

Another well-known secreted glycoprotein used in prostate cancer diagnosis is prostate-specific antigen (PSA), which is secreted by the prostate gland, and the secretion is typically tightly regulated. Normal prostatic cells can also secrete PSA but the secretion in prostate cancer patients may be ~10^5 times higher and could be even higher in patients with the metastatic stage compared with the localized one. PSA is commonly used as a marker for early detection and monitoring of prostate cancer and may be served as a predictor for disease recurrence. There have been concerns on the sensitivity and selectivity of PSA (termed “grey area concentration” in the range of 4–10 ng/mL) and the concentration variations with other conditions such as age.[177,252] Since PSA is a secreted glycoprotein, researchers have moved toward analyzing its glycosylation forms instead of only detecting the protein concentration in the bodily fluids. In general, PSA from cancer patients have higher levels of sialylation and core fucosylation compared to those with benign prostatic

hyperplasia (BPH).[179,253] In one study, the glycosylation profiles of PSA were investigated in urine samples from prostate cancer and BPH patients by anti-PSA antibody purification before MS analysis. While there is no common distinction between the glycosylation profiles of the two groups, an increase of the glycans of FA2 and FM5A2G2S1 was found in prostate cancer patients. The increased level of FA2, which contains a core fucosylated biantennary glycan, was also observed in ovarian cancer and may be used as a distinction marker between low- and intermediate-risk patients from high-risk ones.[254] Haga et al. immunoprecipitated PSA and subsequently enriched its glycosylated forms before MS analysis. The authors found the increase in multisialylated LacdiNAc from sera of prostate cancer patients compared with BPH patients. Based on these results, the PSA G-index was created as a secondary method to exclude false positive diagnosis in PSA screening.[255]

Glycoprotein Analysis in Extracellular Vesicles.

Cells release several types of membrane-bound, extracellular vesicle (EV), including exosomes and microvesicles. Exosomes are intraluminal vesicles contained within the multivesicular body that, upon fusion with the plasma membrane, release their contents outside of the cells. Release of exosomes was once thought to be a method for waste disposal, but later studies showed that it can mediate intercellular communication through transferring molecules such as proteins, lipids, and RNAs to acceptor cells.[256,257] Exosome analysis has gained increased interest in recent years due to its roles in the immune system and the potential to serve as disease biomarkers and targeted drug delivery vehicles for therapeutics.[258]

MS analysis can identify exosomal proteins and their PTMs, including phosphorylation, ubiquitination, palmitoylation, sumoylation, and glycosylation, that may unveil the mechanisms of exosome formation and their biological significance.[259,260] Apart from the well-known proteins found in the exosomes such as tetraspanin CD63, certain glycoproteins, including MUC1, were highly enriched according to a recent data-independent MS analysis of the exosomes from nonsmall cell lung cancer carcinoma.[261] In a study by Chauhan et al., surface glycoproteins from the exosomes of myeloid-derived suppressor cells were oxidized and biotinylated for their enrichment with streptavidin beads, and the majority of the proteins identified can also be found on the cell surface of the parent cells.[262] Sialoglycoproteins were found to be enriched in the exosomes of ovarian carcinoma cells. Several types of glycans were identified, including di-, tri-, and tetra-antennary N-glycans, as well as the core fucose modification and high-mannose glycans. In diantennary glycans, bisecting GlcNAc was detected.[263] Sharma et al. compared the exosomes that were released from cancer cell lines with different invasiveness. It was found that the exosomes from the highly invasive cells, which contained several unique glycoproteins involved in cell migration, were correlated with greater cell migration than those from the less invasive ones.[264] In cells, glycosylation can regulate the trafficking of proteins to the cell surface. Similarly, glycosylation affects the trafficking of proteins to the exosomes. For protein EWI-2, the abundance was decreased when its complex N-linked glycan synthesis was limited to only the high-mannose ones[260].

The exosomes contain specific molecules from their parent cells, and therefore, glycoproteins detected may serve as biomarkers for disease detection. Glypican-1, a cell-surface glycoprotein on the exosomes, was identified as a potential biomarker for early detection of pancreatic cancer using MS.[265] In patients with benign and metastatic pulmonary nodules, conventional diagnosis is difficult to distinguish them apart. The exosomes were isolated from the plasma of these patients, and label-free quantification by MS revealed that glycoproteins, such as fibrinogen beta chain (FGB) and fibrinogen gamma chain (FGG), were up-regulated in the exosomes from patients with metastatic pulmonary nodules.[266] In patients with nonsmall cell lung cancer, alpha-2-HS-glycoprotein (AHSG) and extracellular matrix protein 1 (ECM1) were upregulated compared with normal patients.[267]

Global analysis of glycoproteins in the exosomes may suffer from the high dynamic range of proteins, and thus effective enrichment methods can increase the coverage especially for the low-abundance ones. For example, Bai et al. recently developed a hydrazide-based polymer that can homogeneously enrich N-glycoproteins from the exosomes, and the polymer can be recovered by raising the temperature.[47] Some protein markers such as aquaporin-2 and others related to renal diseases and blood pressure regulation were identified from the exosomes separated from urine.[268] The glycoproteins tetraspanins-1 and hemopexin were recently proposed as early biomarkers for T cell-mediated kidney transplant rejection.[269] Because of the amount of the exosomes needed for proteomic analysis, the large volume of urine may hinder the analysis. In one glycomic study, through a combination of miniaturized sample preparation and prefractionation, the volume of urine was decreased dramatically while comprehensive analysis of glycans was not compromised. The study confirmed the presence of sialylated glycans and several core fucosylations, as well as 16 mucin-type O-glycans and paucimannosidic glycans that were only reported in invertebrates and plants previously.[270]

Recently, a new type of nonmembranous, extracellular nanoparticles termed exomeres was discovered trough asymmetric flow field-flow fractionation (AF4).[271] The exomeres are smaller and stiffer than the exosomes. Unique sets of proteins were found, including N-glycoproteins and those associated with the extracellular matrix, ER, mitochondrion, and cytoskeleton. Proteins involved in glycan processing and control of glycan-mediated protein folding, and sialoglycoproteins were both enriched in the exomeres, suggesting the possible roles in specific glycan recognition and glycosylation modulation of the recipient cells. Analysis with MALDI-TOF MS and LC–MS/MS revealed different N-glycan profiles between the exosomes and exomeres. Glycans from the exomeres contain α-2,3-linked and α-2,6-linked sialic acids while α-2,3-linked sialic acids were found exclusively in glycans from small exosome vesicles. A later study showed that the enzyme β-galactoside α-2,6-sialyltransferase 1 (ST6Gal-I) was present on the membrane of the exomeres and can be transferred to the acceptor cells, causing hypersialylation on the cells.[272]

There are some major differences between different EVs, such as the size, density, biogenesis pathways, and protein markers. Ultracentrifugation or sucrose-gradient centrifugation is typically employed to isolate the exosomes based on their physical properties. This may lead to a mixed population of EVs due to their similar properties such

as the overlapping size and density, which could cause different glycosylation profiles from the same sample.[273] Ideally, the identity or the presence of specific markers on different EVs needs to be confirmed before further characterization. The International Society for Extracellular Vesicles (ISEV) has updated the guidelines for exosome studies in 2018 (Minimal Information for Studies of Extracellular Vesicles, MISEV2018) and the term "exosome" may not be appropriate unless the origin of the vesicles can be verified. Otherwise, the term should be addressed as small EVs.[274] Databases such as ExoCarta contain information about proteins, RNAs, DNAs, and lipids that have been found in the exosomes and meet the minimum requirements from ISEV.[275,276] Other databases including information from extracellular vesicles (i.e., microvesicles, apoptotic bodies, and membrane blebs) are Vesiclepedia and EVpedia.[277,278]

SUMMARY AND FUTURE OUTLOOK

The development of novel methods in glycoproteomics has tremendously expanded the knowledge of protein glycosylation and facilitated our understanding of glycoprotein functions. Here, we review the recent advances in glycoproteomic analysis by MS. Because of the low abundance of many glycoproteins and the complexity of biological samples, it is imperative to enrich glycopeptides/glycoproteins prior to MS analysis. We summarize several enrichment methods to analyze different types of protein glycosylation, including glycoproteins with a particular and important glycan. Due to the high diversity of glycans and their effects on the ionization efficiencies of glycopeptides, it is very challenging to perform global analysis of protein glycosylation. The deglycosylation approaches, especially for N-glycopeptides, are commonly used since a mass tag that can facilitate MS analysis is generated simultaneously. This allows us to confidently identify glycopeptides and localize glycosylation sites, but the glycan information is lost. To avoid such a problem, glycoproteomic analysis of intact glycopeptides has become more popular in the past few years together with the developments in instrumentation and bioinformatics. Additionally, some recent glycoproteomic applications are discussed, including studying glycoprotein functions and analyzing glycoproteins from different biological sources for clinical and biomedical research.

Considering the importance of glycoproteins and the complexity of glycosylation, it is expected that the field of glycoproteomics will continue to thrive in the future. Global analysis of protein glycosylation requires further development of effective and innovative methods. Glycoproteins with a specific glycan represent a rich source to study the roles of abnormal expression of glycoproteins in diseases such as cancer, and innovative methods are urgently needed to separate these glycoproteins for MS identifications and quantifications. New instrumentation with better sensitivity, resolution, and speed and bioinformatic tools with better performance for interpreting MS spectra will not only advance the identification of glycoproteins but also improve the throughput and the quantification accuracy. Furthermore, it will allow us to detect glycoproteins with extremely low abundance that often carry highly valuable information. Future glycoproteomics will further increase the depth of glycosylation analysis and reveal the roles of glycoproteins in physiological and pathological processes, which will facilitate our understanding of the molecular mechanisms

of human diseases and the discovery of glycoproteins as novel and effective biomarkers for disease detection and surveillance.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (CAREER Award, Grant CHE-1454501) and the National Institutes of Health (Grant R01GM118803).

Biographies

Suttipong Suttapitugsakul is a Ph.D. candidate in the School of Chemistry and Biochemistry at the Georgia Institute of Technology. He received his B.S. in Chemistry from the Rensselaer Polytechnic Institute in 2016. He has been working with Prof. Ronghu Wu on systematic investigation of cell-surface glycoproteins using MS since then.

Fangxu Sun is a Ph.D. candidate in the School of Chemistry and Biochemistry at the Georgia Institute of Technology. He received his B S. in Chemistry from the East China University of Science and Technology in 2013 and M.S. in Organic Chemistry from the Shanghai Institute of Organic Chemistry in 2016. He is currently working with Prof. Wu on developing novel MS-based chemical and enzymatic methods for glycoprotein analysis.

Ronghu Wu is an Associate Professor in the School of Chemistry and Biochemistry at the Georgia Institute of Technology. His research group has worked on MS-based proteomics and developed effective methods to globally analyze protein modifications (especially glycosylation) and dynamics.

REFERENCES

(1). Spiro RG Glycobiology 2002, 12 (4), 43R–56R.
(2). Moremen KW; Tiemeyer M; Nairn AV Nat. Rev. Mol. Cell Biol 2012, 13 (7), 448–62. [PubMed: 22722607]
(3). Shental-Bechor D; Levy Y Proc. Natl. Acad. Sci. U. S. A 2008, 105 (24), 8256–61. [PubMed: 18550810]
(4). Helenius A; Aebi M Science 2001, 291 (5512), 2364–9. [PubMed: 11269317]
(5). Munkley J; Elliott DJ Oncotarget 2016, 7 (23), 35478. [PubMed: 27007155]
(6). Ohtsubo K; Marth JD Cell 2006, 126 (5), 855–67. [PubMed: 16959566]
(7). Carr SA; Huddleston MJ; Bean MF Protein Sci. 1993, 2 (2) , 183–96. [PubMed: 7680267]
(8). Zhang H; Li X.-j.; Martin DB; Aebersold R. Nat. Biotechnol 2003, 21 (6), 660–666. [PubMed: 12754519]
(9). Kaji H; Saito H; Yamauchi Y; Shinkawa T; Taoka M; Hirabayashi J; Kasai K; Takahashi N; Isobe T Nat. Biotechnol 2003, 21 (6), 667–72. [PubMed: 12754521]
(10). Wu RH; Haas W; Dephoure N; Huttlin EL; Zhai B; Sowa ME; Gygi SP Nat. Methods 2011, 8 (8), 677–683. [PubMed: 21725298]
(11). Zhang YJ; Wang JQ; Ding M; Yu YH Nat. Methods 2013, 10 (10), 981–984. [PubMed: 23955771]
(12). Desaire H Mol. Cell. Proteomics 2013, 12 (4), 893–901. [PubMed: 23389047]
(13). Reddy RJ; Gajadhar AS; Swenson EJ; Rothenberg DA; Curran TG; White FM Proc. Natl. Acad. Sci. U. S. A 2016, 113 (11), 3114–3119. [PubMed: 26929352]
(14). Chen WX; Smeekens JM; Wu RH Chem. Sci 2016, 7 (2), 1393–1400. [PubMed: 29910897]

(15). Simithy J; Sidoli S; Yuan ZF; Coradin M; Bhanu NV; Marchione DM; Klein BJ; Bazilevsky GA; McCullough CE; Magin RS; Kutateladze TG; Snyder NW; Marmorstein R; Garcia BA Nat. Commun 2017, 8, 1141. [PubMed: 29070843]

(16). Lin ZQ; Wei LM; Cai WX; Zhu YL; Tucholski T; Mitchell SD; Guo W; Ford SP; Diffee GM; Ge Y Mol. Cell. Proteomics 2019, 18 (3), 594–605. [PubMed: 30591534]

(17). Qu YY; Sun LL; Zhu GJ; Zhang ZB; Peuchen EH; Dovichi NJ Talanta 2018, 179, 22–27. [PubMed: 29310225]

(18). Chandler KB; Mehta N; Leon DR; Suscovich TJ; Alter G; Costello CE Mol. Cell. Proteomics 2019, 18 (4), 686–703. [PubMed: 30659065]

(19). Cao LW; Diedrich JK; Kulp DW; Pauthner M; He L; Park SKR; Sok D; Su CY; Delahunty CM; Menis S; Andrabi R; Guenaga J; Georgeson E; Kubitz M; Adachi Y; Burton DR; Schief WR; Yates JR; Paulson JC Nat. Commun 2017, 8, 14954. [PubMed: 28348411]

(20). Yang Y; Franc V; Heck AJR Trends Biotechnol. 2017, 35 (7), 598–609. [PubMed: 28527536]

(21). Yuan W; Wei RHZ; Goldman R; Sanda M Anal. Chem 2019, 91 (14), 9206–9212. [PubMed: 31268672]

(22). Zhang Y; Fonslow BR; Shan B; Baek MC; Yates JR 3rd Chem. Rev 2013, 113 (4), 2343–94. [PubMed: 23438204]

(23). Aebersold R; Mann M Nature 2016, 537 (7620), 347–55. [PubMed: 27629641]

(24). Xiao H; Sun F; Suttapitugsakul S; Wu R Mass Spectrom. Rev 2019, 38 (4–5), 356–379. [PubMed: 30605224]

(25). Thaysen-Andersen M; Packer NH Biochim. Biophys. Acta, Proteins Proteomics 2014, 1844 (9), 1437–1452.

(26). Zhang Y; Kuang M; Zhang LJ; Yang PY; Lu HJ Anal. Chem 2013, 85 (11), 5535–5541. [PubMed: 23659689]

(27). Song E; Zhu R; Hammoud ZT; Mechref Y J. Proteome Res 2014, 13 (11), 4808–4820. [PubMed: 25134008]

(28). Ruiz-May E; Catalá C; Rose JKC N-Glycoprotein Enrichment by Lectin Affinity Chromatography In Plant Proteomics: Methods and Protocols; Jorrin-Novo JV, Komatsu S, Weckwerth W, Wienkoop S, Eds.; Humana Press: Totowa, NJ, 2014; pp 633–643.

(29). Sankaranarayanan R; Sekar K; Banerjee R; Sharma V; Surolia A; Vijayan M Nat. Struct. Biol 1996, 3 (7), 596–603. [PubMed: 8673603]

(30). Wang K; Peng ED; Huang AS; Xia D; Vermont SJ; Lentini G; Lebrun M; Wastling JM; Bradley PJ PLoS One 2016, 11 (3), e0150561. [PubMed: 26950937]

(31). Dong L; Feng S; Li S; Song P; Wang J Anal. Chem 2015, 87 (13), 6849–6853. [PubMed: 26066908]

(32). Waniwan JT; Chen Y-J; Capangpangan R; Weng S-H; Chen Y-J J. Proteome Res 2018, 17 (11), 3761–3773. [PubMed: 30261726]

(33). Zielinska DF; Gnad F; Wisniewski JR; Mann M Cell 2010, 141 (5), 897–907. [PubMed: 20510933]

(34). Zhou J; Yang W; Hu Y; Höti N; Liu Y; Shah P; Sun S; Clark D; Thomas S; Zhang H Anal. Chem 2017, 89 (14), 7623–7630. [PubMed: 28627880]

(35). Yoshida Y; Adachi E; Fukiya K; Iwai K; Tanaka K EMBO Rep. 2005, 6 (3), 239–244. [PubMed: 15723043]

(36). Chen M; Shi X; Duke RM; Ruse CI; Dai N; Taron CH; Samuelson JC Nat. Commun 2017, 8, 15487. [PubMed: 28534482]

(37). Totten SM; Feasley CL; Bermudez A; Pitteri SJ J. Proteome Res 2017, 16 (3), 1249–1260. [PubMed: 28199111]

(38). Zacharias LG; Hartmann AK; Song EH; Zhao JF; Zhu R; Mirzaei P; Mechref Y J. Proteome Res 2016, 15 (10), 3624–3634. [PubMed: 27533485]

(39). Kozlik P; Sanda M; Goldman R J. Chromatogr. A 2017, 1519, 152–155. [PubMed: 28888681]

(40). Wu YL; Liu QJ; Deng CH Anal. Chim. Acta 2019, 1061, 110–121. [PubMed: 30926029]

(41). Li YL; Wang JW; Sun NR; Deng CH Anal. Chem 2017, 89 (20), 11151–11158. [PubMed: 28945362]

(42). Wang J; Li J; Wang Y; Gao M; Zhang X; Yang P ACS Appl. Mater. Interfaces 2016, 8 (41), 27482–27489. [PubMed: 27681085]

(43). Xia CS; Jiao FL; Gao FY; Wang HP; Lv YY; Shen YH; Zhang YJ; Qian XH Anal. Chem 2018, 90 (11), 6651–6659. [PubMed: 29742898]

(44). Cao W; Huang J; Jiang B; Gao X; Yang P Sci. Rep 2016, 6, 29776. [PubMed: 27412817]

(45). Sun F; Suttapitugsakul S; Wu R Anal. Chem 2019, 91 (6), 4195–4203. [PubMed: 30794380]

(46). Bai H; Fan C; Zhang W; Pan Y; Ma L; Ying W; Wang J; Deng Y; Qian X; Qin W Chem. Sci 2015, 6 (7), 4234–4241. [PubMed: 29218189]

(47). Bai H; Pan Y; Qi L; Liu L; Zhao X; Dong H; Cheng X; Qin W; Wang X Talanta 2018, 186, 513–520. [PubMed: 29784395]

(48). Chen W; Smeekens JM; Wu R Mol. Cell. Proteomics 2014, 13 (6), 1563–1572. [PubMed: 24692641]

(49). Zhang LJ; Xu YW; Yao HL; Xie LQ; Yao J; Lu HJ; Yang PY Chem. - Eur. J 2009, 15 (39), 10158–10166. [PubMed: 19658139]

(50). Xiao HP; Hwang JE; Wu RH Int. J. Mass Spectrom 2018, 429, 66–75. [PubMed: 30147434]

(51). Li DJ; Chen Y; Liu Z Chem. Soc. Rev 2015, 44 (22), 8097–8123. [PubMed: 26377373]

(52). Bie ZJ; Chen Y; Li HY; Wu RH; Liu Z Anal. Chim. Acta 2014, 834, 1–8. [PubMed: 24928239]

(53). Qu YY; Liu JX; Yang KG; Liang Z; Zhang LH; Zhang YK Chem. - Eur. J 2012, 18 (29), 9056–9062. [PubMed: 22707097]

(54). Xiao HP; Smeekens JM; Wu RH Analyst 2016, 141 (12), 3737–3745. [PubMed: 27007503]

(55). Xie YQ; Liu QJ; Li Y; Deng CH J. Chromatogr. A 2018, 1540, 87–93. [PubMed: 29429745]

(56). Zhang X; Wang J; He X; Chen L; Zhang Y ACS Appl. Mater. Interfaces 2015, 7 (44), 24576–24584. [PubMed: 26479332]

(57). Wang J; Wang Y; Gao M; Zhang X; Yang P ACS Appl. Mater. Interfaces 2015, 7 (29), 16011–16017. [PubMed: 26161682]

(58). Xiao H; Chen W; Smeekens JM; Wu R Nat. Commun 2018, 9 (1), 1692. [PubMed: 29703890]

(59). Dowlut M; Hall DG J. Am. Chem. Soc 2006, 128 (13), 4226–4227. [PubMed: 16568987]

(60). Sun S; Shah P; Eshghi ST; Yang W; Trikannad N; Yang S; Chen L; Aiyetan P; Hoti N; Zhang Z; Chan DW; Zhang H Nat. Biotechnol 2016, 34 (1), 84–8. [PubMed: 26571101]

(61). Yang W; Ao M; Hu Y; Li QK; Zhang H Mol. Syst. Biol 2018, 14 (11), e8486. [PubMed: 30459171]

(62). Hubbard SC; Boyce M; McVaugh CT; Peehl DM; Bertozzi CR Bioorg. Med. Chem. Lett 2011, 21 (17), 4945–4950. [PubMed: 21798741]

(63). Smeekens JM; Chen WX; Wu RH J. Am. Soc. Mass Spectrom 2015, 26 (4), 604–614. [PubMed: 25425172]

(64). Xiao HP; Wu RH Anal. Chem 2017, 89 (6), 3656–3663. [PubMed: 28234450]

(65). Zhu YT; Wu J; Chen X Angew. Chem., Int. Ed 2016, 55 (32), 9301–9305.

(66). Sun T; Yu SH; Zhao P; Meng L; Moremen KW; Wells L; Steet R; Boons GJ J. Am. Chem. Soc 2016, 138 (36), 11575–11582. [PubMed: 27541995]

(67). Woo CM; Iavarone AT; Spiciarich DR; Palaniappan KK; Bertozzi CR Nat. Methods 2015, 12 (6), 561–567. [PubMed: 25894945]

(68). Dube DH; Bertozzi CR Nat. Rev. Drug Discovery 2005, 4, 477–488. [PubMed: 15931257]

(69). Pinho SS; Reis CA Nat. Rev. Cancer 2015, 15 (9), 540–55. [PubMed: 26289314]

(70). Vajaria B; Patel P Glycoconjugate J. 2017, 34 (2), 147–156.

(71). Veillon L; Fakih C; Abou-El-Hassan H; Kobeissy F; Mechref Y ACS Chem. Neurosci 2018, 9 (1), 51–72. [PubMed: 28982002]

(72). Maverakis E; Kim K; Shimoda M; Gershwin ME; Patel F; Wilken R; Raychaudhuri S; Ruhaak LR; Lebrilla CB J. Autoimmun 2015, 57, 1–13. [PubMed: 25578468]

(73). Kudelka MR; Ju TZ; Heimburg-Molinaro J; Cummings RD. Adv. Cancer Res 2015, 126, 53–135. [PubMed: 25727146]

(74). Chia J; Goh G; Bard F Biochim. Biophys. Acta, Gen. Subj 2016, 1860 (8), 1623–1639.

(75). Ju TZ; Lanneau GS; Gautam T; Wang YC; Xia BY; Stowell SR; Willard MT; Wang WY; Xia JY; Zuna RE; Laszik Z; Benbrook DM; Hanigan MH; Cummings RD Cancer Res. 2008, 68 (6), 1636–1646. [PubMed: 18339842]

(76). Ghazarian H; Idoni B; Oppenheimer SB Acta Histochem. 2011, 113 (3), 236–247. [PubMed: 20199800]

(77). Hoja-Łukowicz D; Szwed S; Laidler P; Lityńska A Biochimie 2018, 151, 14–26. [PubMed: 29802864]

(78). Valoskova K; Biebl J; Roblek M; Emtenani S; Gyoergy A; Misova M; Ratheesh A; Reis-Rodrigues P; Shkarina K; Larsen ISB; Vakhrushev SY; Clausen H; Siekhaus DE eLife 2019, 8, e41801. [PubMed: 30910009]

(79). Zheng J; Xiao H; Wu R Angew. Chem., Int. Ed 2017, 56 (25), 7107–7111.

(80). Campos D; Freitas D; Gomes J; Magalhães A; Steentoft C; Gomes C; Vester-Christensen MB; Ferreira JA; Afonso LP; Santos LL; Pinto de Sousa J; Mandel U; Clausen H; Vakhrushev SY; Reis CA Mol. Cell. Proteomics 2015, 14 (6), 1616. [PubMed: 25813380]

(81). Peixoto A; Fernandes E; Gaiteiro C; Lima L; Azevedo R; Soares J; Cotton S; Parreira B; Neves M; Amaro T; Tavares A; Teixeira F; Palmeira C; Rangel M; Silva AMN; Reis CA; Santos LL; Oliveira MJ; Ferreira JA Oncotarget 2016, 7 (39), 63138–63157. [PubMed: 27542232]

(82). Halim A; Nilsson J; Ruetschi U; Hesse C; Larson G Mol. Cell. Proteomics 2012, 11 (4), M111.013649.

(83). Lopez Aguilar A; Briard JG; Yang L; Ovryn B; Macauley MS; Wu P ACS Chem. Biol 2017, 12 (3), 611–621. [PubMed: 28301937]

(84). Li Q; Li Z; Duan X; Yi WJ Am. Chem. Soc 2014, 136 (36), 12536–12539.

(85). Li F; Glinskii OV; Mooney BP; Rittenhouse-Olson K; Pienta KJ; Glinsky VV Oncotarget 2017, 8 (58), 98598–98608. [PubMed: 29228713]

(86). Wen L; Liu D; Zheng Y; Huang K; Cao X; Song J; Wang PG ACS Cent. Sci 2018, 4 (4), 451–457. [PubMed: 29721527]

(87). Wells L; Vosseller K; Hart GW Science 2001, 291 (5512), 2376–2378. [PubMed: 11269319]

(88). Yang XY; Qian KV Nat. Rev. Mol. Cell Biol 2017, 18 (7), 452–465. [PubMed: 28488703]

(89). Wang XS; Yuan ZF; Fan J; Karch KR; Ball LE; Denu JM; Garcia BA Mol. Cell. Proteomics 2016, 15 (7), 2462–2475. [PubMed: 27114449]

(90). Slawson C; Hart GW Nat. Rev. Cancer 2011, 11 (9), 678–684. [PubMed: 21850036]

(91). Ma JF; Hart GW Clin. Proteomics 2014, 11, 8. [PubMed: 24593906]

(92). You X; Qin H; Ye M J. Sep. Sci 2018, 41 (1), 248–261. [PubMed: 28988430]

(93). Khidekel N; Ficarro SB; Peters EC; Hsieh-Wilson LC Proc. Natl. Acad. Sci. U. S. A 2004, 101 (36), 13132–7. [PubMed: 15340146]

(94). Qin K; Zhu Y; Qin W; Gao J; Shao X; Wang Y.-l.; Zhou W; Wang C; Chen X ACS Chem. Biol 2018, 13 (8), 1983–1989. [PubMed: 30059200]

(95). Hao Y; Fan XQ; Shi YJ; Zhang C; Sun DE; Qin K; Qin W; Zhou W; Chen X Nat. Commun 2019, 10, 4065. [PubMed: 31492838]

(96). Alfaro JF; Gong CX; Monroe ME; Aldrich JT; Clauss TRW; Purvine SO; Wang ZH; Camp DG; Shabanowitz J; Stanley P; Hart GW; Hunt DF; Yang F; Smith RD Proc. Natl. Acad. Sci. U. S. A 2012, 109 (19), 7280–7285. [PubMed: 22517741]

(97). Griffin ME; Jensen EH; Mason DE; Jenkins CL; Stone SE; Peters EC; Hsieh-Wilson LC Mol. BioSyst 2016, 12 (6), 1756–1759. [PubMed: 27063346]

(98). Rexach JE; Rogers CJ; Yu SH; Tao JF; Sun YE; Hsieh-Wilson LC Nat. Chem. Biol 2010, 6 (9), 645–651. [PubMed: 20657584]

(99). Darabedian N; Thompson JW; Chuh KN; Hsieh-Wilson LC; Pratt MR Biochemistry 2018, 57 (40), 5769–5774. [PubMed: 30169966]

(100). Huo B; Zhang W; Zhao X; Dong H; Yu Y; Wang J; Qian X; Qin W Chem. Commun 2018, 54 (98), 13790–13793.

(101). Neubert P; Halim A; Zauser M; Essig A; Joshi HJ; Zatorska E; Larsen ISB; Loibl M; Castells-Ballester J; Aebi M; Clausen H; Strahl S Mol. Cell. Proteomics 2016, 15 (4), 1323–1337. [PubMed: 26764011]

(102). Lommel M; Strahl S Glycobiology 2009, 19 (8), 816–28. [PubMed: 19429925]

(103). Vester-Christensen MB; Halim A; Joshi HJ; Steentoft C; Bennett EP; Levery SB; Vakhrushev SY; Clausen H Proc. Natl. Acad. Sci. U. S. A 2013, 110 (52), 21018. [PubMed: 24101494]

(104). Dahms NM; Olson LJ; Kim JJP Glycobiology 2008, 18 (9), 664–678. [PubMed: 18621992]

(105). Sleat DE; Sun P; Wiseman JA; Huang L; El-Banna M; Zheng H; Moore DF; Lobel P Mol. Cell. Proteomics 2013, 12 (7), 1806. [PubMed: 23478313]

(106). Markmann S; Krambeck S; Hughes CJ; Mirzaian M; Aerts JMFG; Saftig P; Schweizer M; Vissers JPC; Braulke T; Damme M Mol. Cell. Proteomics 2017, 16 (3), 438. [PubMed: 28062798]

(107). Čaval T; Zhu J; Tian W; Remmelzwaal S; Yang Z; Clausen H; Heck AJR Mol. Cell. Proteomics 2019, 18 (1), 16. [PubMed: 30237200]

(108). Chen WX; Smeekens JM; Wu RH Mol. Cell. Proteomics 2014, 13 (6), 1563–1572. [PubMed: 24692641]

(109). Thaysen-Andersen M; Packer NH; Schulz BL Mol. Cell. Proteomics 2016, 15 (6), 1773–1790. [PubMed: 26929216]

(110). Zhu ZK; Su XM; Clark DF; Go EP; Desaire H Anal. Chem 2013, 85 (17), 8403–8411. [PubMed: 23909558]

(111). Frese CK; Zhou H; Taus T; Altelaar AFM; Mechtler K; Heck AJR; Mohammed SJ. Proteome Res 2013, 12 (3), 1520–1525. [PubMed: 23347405]

(112). Zhu J; Chen Z; Zhang J; An M; Wu J; Yu Q; Skilton, St J; Bern M; Ilker Sen K; Li L; Lubman DM J. Proteome Res 2018, 18 (1), 359–371. [PubMed: 30370771]

(113). Yu Q; Canales A; Glover MS; Das R; Shi XD; Liu Y; Keller MP; Attie AD; Li LJ Anal. Chem 2017, 89 (17), 9184–9191. [PubMed: 28726377]

(114). Yu Q; Wang B; Chen Z; Urabe G; Glover MS; Shi X; Guo LW; Kent KC; Li LJ. Am. Soc. Mass Spectrom 2017, 28 (9) , 1751–1764. [PubMed: 28695533]

(115). Bern M; Kil YJ; Becker C Byonic: Advanced Peptide and Protein Identification Software In Current Protocols in Bioinformatics, Vol. 40; John Wiley & Sons, Inc.: Hoboken, NJ, 2012; pp 13.20.1–13.20.14,

(116). Toghi Eshghi S; Shah P; Yang W; Li X; Zhang H Anal. Chem 2015, 87 (10), 5181–5188. [PubMed: 25945896]

(117). Liu M-Q; Zeng W-F; Fang P; Cao W-Q; Liu C; Yan G-Q; Zhang Y; Peng C; Wu J-Q; Zhang X-J; Tu H-J; Chi H; Sun R-X; Cao Y; Dong M-Q; Jiang B-Y; Huang J-M; Shen H-L; Wong CCL; He S-M; Yang P-Y Nat. Commun 2017, 8 (1), 438. [PubMed: 28874712]

(118). Stadlmann J; Taubenschmid J; Wenzel D; Gattinger A; Dürnberger G; Dusberger F; Elling U; Mach L; Mechtler K; Penninger JM Nature 2017, 549 (7673), 538–542. [PubMed: 28959962]

(119). Xiao K; Tian ZJ Proteome Res. 2019, 18 (7), 2885–2895.

(120). Reiding KR; Bondt A; Franc V; Heck AJR TrAC, Trends Anal. Chem 2018, 108, 260–268.

(121). Hu H; Khatri K; Zaia J Mass Spectrom. Rev 2017, 36 (4), 475–498. [PubMed: 26728195]

(122). Woodin CL; Maxon M; Desaire H Analyst 2013, 138 (10) , 2793–2803. [PubMed: 23293784]

(123). Parker BL; Thaysen-Andersen M; Fazakerley DJ; Holliday M; Packer NH; James DE Mol. Cell. Proteomics 2016, 15 (1), 141. [PubMed: 26537798]

(124). Fang P; Xie J-J; Sang S-M; Zhang L; Liu M-Q; Yang L-J; Xu Y-T; Yan G-Q; Yao J; Gao X; Qian W-J; Wang Z-F; Zhang Y; Yang P-Y; Shen H-L bioRxiv 2019, 615989.

(125). Brown CJ; Grassmyer KT; MacDonald ML; Clemmer DE; Trinidad JC bioRxiv 2019, 647081.

(126). Riley NM; Hebert AS; Durnberger G; Stanek F; Mechtler K; Westphall MS; Coon JJ Anal. Chem 2017, 89 (12), 6367–6376. [PubMed: 28383256]

(127). Riley NM; Hebert AS; Westphall MS; Coon JJ Nat. Commun 2019, 10, 1311. [PubMed: 30899004]

(128). Braulke T; Bonifacino JS Biochim. Biophys. Acta, Mol. Cell Res 2009, 1793 (4), 605–614.

(129). King SL; Joshi HJ; Schjoldager KT; Halim A; Madsen TD; Dziegiel MH; Woetmann A; Vakhrushev SY; Wandall HH Blood Adv. 2017, 1 (7), 429. [PubMed: 29296658]

(130). Steentoft C; Vakhrushev SY; Vester-Christensen MB; Schjoldager KTBG; Kong Y; Bennett EP; Mandel U; Wandall H; Levery SB; Clausen H Nat. Methods 2011, 8, 977–982. [PubMed: 21983924]

(131). Qin H; Cheng K; Zhu J; Mao J; Wang F; Dong M; Chen R; Guo Z; Liang X; Ye M; Zou H Anal. Chem 2017, 89 (3) , 1469–1476. [PubMed: 28035807]

(132). Mao JW; You X; Qin HQ; Wang CY; Wang LM; Ye ML Anal. Chem 2019, 91 (6), 3852–3859. [PubMed: 30802037]

(133). Zhang Y; Xie X; Zhao X; Tian F; Lv J; Ying W; Qian X J. Proteomics 2018, 170, 14–27. [PubMed: 28970103]

(134). Madsen JA; Ko BJ; Xu H; Iwashkiw JA; Robotham SA; Shaw JB; Feldman MF; Brodbelt JS Anal. Chem 2013, 85 (19), 9253–9261. [PubMed: 24006841]

(135). Pap A; Klement E; Hunyadi-Gulyas E; Darula Z; Medzihradszky KF J. Am. Soc. Mass Spectrom 2018, 29 (6), 1210–1220. [PubMed: 29730764]

(136). Malaker SA; Pedram K; Ferracane MJ; Bensing BA; Krishnan V; Pett C; Yu J; Woods EC; Kramer JR; Westerlind U; Dorigo O; Bertozzi CR Proc. Natl. Acad. Sci. U. S. A 2019, 116, 7278. [PubMed: 30910957]

(137). Yang Y; Liu F; Franc V; Halim LA; Schellekens H; Heck AJ Nat. Commun 2016, 7, 13397. [PubMed: 27824045]

(138). Wu D; Li J; Struwe WB; Robinson CV Chem. Sci 2019, 10 (19), 5146–5155. [PubMed: 31183067]

(139). Vagin O; Kraut JA; Sachs G Am. J. Physiol. Renal Physiol 2009, 296 (3), F459–69. [PubMed: 18971212]

(140). Lee HS; Qi Y; Im W Sci. Rep 2015, 5, 8926. [PubMed: 25748215]

(141). Tran DT; Ten Hagen KG J. Biol. Chem 2013, 288 (10), 6921–9. [PubMed: 23329828]

(142). Bergstrom KS; Xia L Glycobiology 2013, 23 (9), 1026–37. [PubMed: 23752712]

(143). Brockhausen I; Stanley P O-GalNAc Glycans In Essentials of Glycobiology; Varki A, Cummings RD, Esko JD; Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2015; pp 113–123.

(144). Eichler J; Koomey M Trends Microbiol. 2017, 25 (8), 662–672. [PubMed: 28341406]

(145). Hart GW; Slawson C; Ramirez-Correa G; Lagerlof O Annu. Rev. Biochem 2011, 80, 825–58. [PubMed: 21391816]

(146). Akan I; Olivier-Van Stichelen S; Bond MR; Hanover JA J. Neurochem 2018, 144 (1), 7–34. [PubMed: 29049853]

(147). Machacek M; Saunders H; Zhang Z; Tan EP; Li J; Li T; Villar MT; Artigues A; Lydic T; Cork G; Slawson C; Fields PE J. Biol. Chem 2019, 294 (22), 8973–8990. [PubMed: 31010828]

(148). Lo PW; Shie JJ; Chen CH; Wu CY; Hsu TL; Wong CH Proc. Natl. Acad. Sci. U. S. A 2018, 115 (28), 7302–7307. [PubMed: 29941599]

(149). Sekine H; Okazaki K; Kato K; Alam MM; Shima H; Katsuoka F; Tsujita T; Suzuki N; Kobayashi A; Igarashi K; Yamamoto M; Motohashi H Mol. Cell. Biol 2018, 38 (17), e00252–18. [PubMed: 29941490]

(150). Li F; Ding J Protein Cell 2019, 10 (8), 550–565. [PubMed: 30478534]

(151). Zhang QW; Li Z; Wang YW; Zheng Q; Li JJ Mass Spectrom. Rev 2018, 37 (5), 652–680. [PubMed: 29228471]

(152). Yogeeswaran G; Salk PL Science 1981, 212 (4502), 1514–6. [PubMed: 7233237]

(153). Teoh ST; Ogrodzinski MP; Ross C; Hunter KW; Lunt SY Front. Oncol 2018, 8, 174. [PubMed: 29892572]

(154). Park D; Xu G; Barboza M; Shah IM; Wong M; Raybould H; Mills DA; Lebrilla CB Glycobiology 2017, 27 (9), 847–860. [PubMed: 28486580]

(155). Wu D; Struwe WB; Harvey DJ; Ferguson MAJ; Robinson CV Proc. Natl. Acad. Sci. U. S. A 2018, 115 (35), 8763–8768. [PubMed: 30111543]

(156). Toonstra C; Hu Y; Zhang H J. Proteome Res 2019, 18 (6), 2467–2477. [PubMed: 31055923]

(157). Saarinen L; Nummela P; Leinonen H; Heiskanen A; Thiel A; Haglund C; Lepisto A; Satomaa T; Hautaniemi S; Ristimaki A Mol. Cell. Proteomics 2018, 17 (11), 2107–2118. [PubMed: 30072579]

(158). Qin H; Chen Y; Mao J; Cheng K; Sun D; Dong M; Wang L; Wang L; Ye M Anal. Chim. Acta 2019, 1070, 60–68. [PubMed: 31103168]

(159). Li QK; Shah P; Tian Y; Hu Y; Roden RBS; Zhang H; Chan DW Clin. Proteomics 2017, 14, 16. [PubMed: 28491011]

(160). Yang X; Wang Z; Guo L; Zhu Z; Zhang Y J. Proteome Res 2017, 16 (10), 3830–3840. [PubMed: 28845672]

(161). Valliere-Douglass JF; Eakin CM; Wallace A; Ketchem RR; Wang W; Treuheit MJ; Balland A J. Biol. Chem 2010, 285 (21), 16012–22. [PubMed: 20233717]

(162). Zajonc DM; Striegl H; Dascher CC; Wilson IA Proc. Natl. Acad. Sci. U. S. A 2008, 105 (46), 17925–30. [PubMed: 19004781]

(163). Bantscheff M; Lemeer S; Savitski MM; Kuster B Anal. Bioanal. Chem 2012, 404 (4), 939–65. [PubMed: 22772140]

(164). Dephoure N; Gygi SP Sci. Signaling 2012, 5 (217), rs2.

(165). Evans AR; Robinson RAS. Proteomics 2013, 13 (22), 3267–3272. [PubMed: 24124127]

(166). Arul AB; Robinson RAS. Anal. Chem 2019, 91 (1), 178–189. [PubMed: 30525468]

(167). Stadlmann J; Taubenschmid J; Wenzel D; Gattinger A; Durnberger G; Dusberger F; Elling U; Mach L; Mechtler K; Penninger JM Nature 2017, 549 (7673), 538–542. [PubMed: 28959962]

(168). Xiao H; Wu R Anal. Chem 2017, 89 (19), 10361–10367. [PubMed: 28850217]

(169). Parker CE; Borchers CH Mol. Oncol 2014, 8 (4), 840–58. [PubMed: 24713096]

(170). Tan Z; Yin H; Nie S; Lin Z; Zhu J; Ruffin MT; Anderson MA; Simeone DM; Lubman DM J. Proteome Res 2015, 14 (4), 1968–1978. [PubMed: 25732060]

(171). Zhu JH; Warner E; Parikh ND; Lubman DM Mass Spectrom. Rev 2019, 38 (3), 265–290. [PubMed: 30472795]

(172). Kammeijer GSM; Nouta J; de la Rosette J; de Reijke TM; Wuhrer M Anal. Chem 2018, 90 (7), 4414–4421. [PubMed: 29502397]

(173). Holst S; Belo AI; Giovannetti E; van Die I; Wuhrer M Sci. Rep 2017, 7, 16623. [PubMed: 29192278]

(174). Zhang Y; Jiao J; Yang PY; Lu HJ Clin. Proteom 2014, 11, 18.

(175). Scott DA; Norris-Caneda K; Spruill L; Bruner E; Kono Y; Angel PM; Mehta AS; Drake RR Int. J. Mass Spectrom 2019, 437, 69–76. [PubMed: 31031563]

(176). Kirwan A; Utratna M; O'Dwyer ME; Joshi L; Kilcoyne M BioMed Res. Int 2015, 2015, 490531. [PubMed: 26509158]

(177). Lilja H; Ulmert D; Vickers AJ Nat. Rev. Cancer 2008, 8 (4), 268–78. [PubMed: 18337732]

(178). Llop E; Ferrer-Batalle M; Barrabes S; Guerrero PE; Ramirez M; Saldova R; Rudd PM; Aleixandre RN; Comet J; de Llorens R; Peracaula R Theranostics 2016, 6 (8), 1190–204. [PubMed: 27279911]

(179). Song E; Mayampurath A; Yu CY; Tang HX; Mechref Y J. Proteome Res 2014, 13 (12), 5570–5580. [PubMed: 25327667]

(180). Siegel RL; Miller KD; Jemal A Ca-Cancer J. Clin 2019, 69 (1), 7–34. [PubMed: 30620402]

(181). Yang S; Chen L; Chan DW; Li QK; Zhang H Clin. Proteomics 2017, 14, 31. [PubMed: 28814946]

(182). Jia L; Zhang J; Ma T; Guo Y; Yu Y; Cui J Front. Oncol 2018, 8, 565. [PubMed: 30619732]

(183). Shah P; Wang X; Yang W; Toghi Eshghi S; Sun S; Hoti N; Chen L; Yang S; Pasay J; Rubin A; Zhang H Mol. Cell. Proteomics 2015, 14 (10), 2753–63. [PubMed: 26256267]

(184). Lavrsen K; Dabelsteen S; Vakhrushev SY; Levann AMR; Haue AD; Dylander A; Mandel U; Hansen L; Frodin M; Bennett EP; Wandall HH J. Biol. Chem 2018, 293 (4), 1298–1314. [PubMed: 29187600]

(185). Sinha A; Hussain A; Ignatchenko V; Ignatchenko A; Tang KH; Ho VWH; Neel BG; Clarke B; Bernardini MQ; Ailles L; Kislinger T Cell Syst. 2019, 8 (4), 345–351. [PubMed: 30981729]

(186). Link-Lenczowski P; Bubka M; Balog CIA; Koeleman CAM; Butters TD; Wuhrer M; Litynska A Glycoconjugate J. 2018, 35 (2), 217–231.

(187). Agrawal P; Fontanals-Cirera B; Sokolova E; Jacob S; Vaiana CA; Argibay D; Davalos V; McDermott M; Nayak S; Darvishian F; Castillo M; Ueberheide B; Osman I; Fenyo D; Mahal LK; Hernando E Cancer Cell 2017, 31 (6), 804–819. [PubMed: 28609658]

(188). Hoja-Lukowicz D; Szwed S; Laidler P; Litynska A Biochimie 2018, 151, 14–26. [PubMed: 29802864]

(189). Zhu Y; Shan X; Yuzwa SA; Vocadlo DJ J. Biol. Chem 2014, 289 (50), 34472–81. [PubMed: 25336656]

(190). Wang S; Yang F; Petyuk VA; Shukla AK; Monroe ME; Gritsenko MA; Rodland KD; Smith RD; Qian WJ; Gong CX; Liu T J. Pathol 2017, 243 (1), 78–88. [PubMed: 28657654]

(191). Lamoureux L; Simon SLR; Waitt B; Knox JD Proteomics 2018, 18 (1), 1700296.

(192). Tian Y; Gurley K; Meany DL; Kemp CJ; Zhang H J. Proteome Res. 2009, 8 (4), 1657–62. [PubMed: 19714870]

(193). Ostasiewicz P; Zielinska DF; Mann M; Wisniewski JR J. Proteome Res 2010, 9 (7), 3688–700. [PubMed: 20469934]

(194). Hinneburg H; Korac P; Schirmeister F; Gasparov S; Seeberger PH; Zoldos V; Kolarich D Mol. Cell. Proteomics 2017, 16 (4), 524–536. [PubMed: 28122943]

(195). Varki A Essentials of Glycobiology, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2017.

(196). Stanley P; Taniguchi N; Aebi M N-Glycans In Essentials of Glycobiology, 3rd ed.; Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2015; pp 99–111.

(197). Salmi M; Jalkanen S Nat. Rev. Immunol 2005, 5 (10), 760–71. [PubMed: 16200079]

(198). Shankaran H; Resat H; Wiley HS PLoS Comput. Biol 2007, 3 (6), e101. [PubMed: 17542642]

(199). Engel P; Boumsell L; Balderas R; Bensussan A; Gattei V; Horejsi V; Jin BQ; Malavasi F; Mortari F; Schwartz-Albiez R; Stockinger H; van Zelm MC; Zola H; Clark G J. Immunol 2015, 195 (10), 4555–63. [PubMed: 26546687]

(200). DeVeale B; Bausch-Fluck D; Seaberg R; Runciman S; Akbarian V; Karpowicz P; Yoon C; Song H; Leeder R; Zandstra PW; Wollscheid B; van der Kooy D Stem Cells 2014, 32 (1), 258–268. [PubMed: 24023036]

(201). Hofmann A; Thiesler T; Gerrits B; Behnke S; Sobotzki N; Omasits U; Bausch-Fluck D; Bock T; Aebersold R; Moch H; Tinguely M; Wollscheid B Proteomics: Clin. Appl 2015, 9 (7–8), 661–670. [PubMed: 26076441]

(202). Gundry RL; Riordon DR; Tarasova Y; Chuppa S; Bhattacharya S; Juhasz O; Wiedemeier O; Milanovich S; Noto FK; Tchernyshyov I; Raginski K; Bausch-Fluck D; Tae HJ; Marshall S; Duncan SA; Wollscheid B; Wersto RP; Rao S; Van Eyk JE; Boheler KR Mol. Cell. Proteomics 2012, 11 (8), 303–16. [PubMed: 22493178]

(203). Vernay A; Cosson P BMC Res. Notes 2013, 6, 317. [PubMed: 23937900]

(204). Wollscheid B; Bausch-Fluck D; Henderson C; O'Brien R; Bibel M; Schiess R; Aebersold R; Watts JD Nat. Biotechnol 2009, 27 (4), 378–86. [PubMed: 19349973]

(205). Bausch-Fluck D; Hofmann A; Bock T; Frei AP; Cerciello F; Jacobs A; Moest H; Omasits U; Gundry RL; Yoon C; Schiess R; Schmidt A; Mirkowska P; Hartlova A; Van Eyk JE; Bourquin JP; Aebersold R; Boheler KR; Zandstra P; Wollscheid B PLoS One 2015, 10 (3), e0121314. [PubMed: 25894527]

(206). van Oostrum M; Müller M; Klein F; Bruderer R; Zhang H; Pedrioli PGA; Reiter L; Tsapogas P; Rolink A; Wollscheid B bioRxiv 2019, 620344.

(207). Ramya TN; Weerapana E; Cravatt BF; Paulson JC Glycobiology 2013, 23 (2), 211–21. [PubMed: 23070960]

(208). Suttapitugsakul S; Ulmer LD; Jiang CD; Sun FX; Wu RH Anal. Chem 2019, 91 (10), 6934–6942. [PubMed: 31025852]

(209). Sletten EM; Bertozzi CR Angew. Chem., Int. Ed 2009, 48 (38), 6974–98.

(210). Devaraj NK ACS Cent. Sci 2018, 4 (8), 952–959. [PubMed: 30159392]

(211). Chen W; Smeekens JM; Wu R Chem. Sci 2015, 6 (8), 4681–4689. [PubMed: 29142707]

(212). Xiao HP; Suttapitugsakul S; Sun FX; Wu RH Acc. Chem. Res 2018, 51 (8), 1796–1806. [PubMed: 30011186]

(213). Laughlin ST; Bertozzi CR Nat. Protoc 2007, 2 (11), 2930–44. [PubMed: 18007630]

(214). Chang PV; Chen X; Smyrniotis C; Xenakis A; Hu T; Bertozzi CR; Wu P Angew. Chem., Int. Ed 2009, 48 (22), 4030–3.

(215). Spiciarich DR; Nolley R; Maund SL; Purcell SC; Herschel J; Iavarone AT; Peehl DM; Bertozzi CR Angew. Chem., Int. Ed 2017, 56 (31), 8992–8997.

(216). Li Q; Xie Y; Xu G; Lebrilla CB Chem. Sci 2019, 10 (24), 6199–6209. [PubMed: 31360427]

(217). Han SS; Lee DE; Shim HE; Lee S; Jung T; Oh JH; Lee HA; Moon SH; Jeon J; Yoon S; Kim K; Kang SW Theranostics 2017, 7 (5), 1164–1176. [PubMed: 28435456]

(218). Han SS; Shim HE; Park SJ; Kim BC; Lee DE; Chung HM; Moon SH; Kang SW Sci. Rep 2018, 8 (1), 13212. [PubMed: 30181604]

(219). Xiao H; Tang GX; Wu R Anal. Chem 2016, 88 (6), 3324–32. [PubMed: 26894747]

(220). Xiao H; Wu R Chem. Sci 2017, 8 (1), 268–277. [PubMed: 28616130]

(221). Lund R; Leth-Larsen R; Jensen ON; Ditzel HJ J. Proteome Res 2009, 8 (6), 3078–90. [PubMed: 19341246]

(222). Fleming KG Applications of Analytical Ultracentrifugation to Membrane Proteins In Analytical Ultracentrifugation: Instrumentation, Software, and Applications; Uchiyama S, Arisaka F, Stafford WF,, Laue T, Eds.; Springer Japan: Tokyo, Japan, 2016; pp 311–327.

(223). Hormann K; Stukalov A; Muller AC; Heinz LX; Superti-Furga G; Colinge J; Bennett KL J. Proteome Res 2016, 15 (2), 647–58. [PubMed: 26699813]

(224). Haverland NA; Waas M; Ntai I; Keppel T; Gundry RL; Kelleher NL Proteomics 2017, 17 (19), 1700156.

(225). Mallanna SK; Waas M; Duncan SA; Gundry RL Proteomics 2017, 17 (5), 1600397.

(226). Matta C; Boocock DJ; Fellows CR; Miosge N; Dixon JE; Liddell S; Smith J; Mobasheri A Sci. Rep 2019, 9 (1), 9018. [PubMed: 31227739]

(227). Kalxdorf M; Gade S; Eberl HC; Bantscheff M Mol. Cell. Proteomics 2017, 16 (5), 770–785. [PubMed: 28336715]

(228). Zarif JC; Yang W; Hernandez JR; Zhang H; Pienta KJ Mol. Cell. Proteomics 2017, 16 (6), 1029–1037. [PubMed: 28348171]

(229). Piccaluga PP; Arpinati M; Candoni A; Laterza C; Paolini S; Gazzola A; Sabattini E; Visani G; Pileri SA Leuk. Lymphoma 2011, 52 (2), 325–7. [PubMed: 21077738]

(230). Bausch-Fluck D; Goldmann U; Muller S; van Oostrum M; Muller M; Schubert OT; Wollscheid B Proc. Natl. Acad. Sci. U. S. A 2018, 115 (46), E10988–E10997. [PubMed: 30373828]

(231). Smeekens JM; Xiao H; Wu R J. Proteome Res 2017, 16 (2), 1039–1049. [PubMed: 27933904]

(232). Uhlen M; Tegel H; Sivertsson Å; Kuo C-C; Gutierrez JM; Lewis NE; Forsström B; Dannemeyer M; Fagerberg L; Malm M; Vunk H; Edfors F; Hober A; Sjöstedt E; Kotol D; Mulder J; Mardinoglu A; Schwenk JM; Nilsson P; Zwahlen M; Takanen JO; von Feilitzen K; Stadler C; Lindskog C; Ponten F; Nielsen J; Palsson BO; Volk A-L; Lundqvist M; Berling A; Svensson A-S; Kanje S; Enstedt H; Afshari D; Ekblad S; Scheffel J; Katona B; Vuu J; Lindström E; Xu L; Mihai R; Bremer L; Westin M; Muse M; Mayr LM; Knight S; Göpel S; Davies R; Varley P; Hatton D; Fields R; Voldborg BG; Rockberg J; Schiavone LH; Hober S bioRxiv 2018, 465815.

(233). Stastna M; Van Eyk JE Proteomics 2012, 12 (4–5), 722–35. [PubMed: 22247067]

(234). Meissner F; Scheltema RA; Mollenkopf HJ; Mann M Science 2013, 340 (6131), 475–8. [PubMed: 23620052]

(235). Eichelbaum K; Winter M; Diaz MB; Herzig S; Krijgsveld J Nat. Biotechnol 2012, 30 (10), 984. [PubMed: 23000932]

(236). Shin J; Rhim J; Kwon Y; Choi SY; Shin S; Ha CW; Lee C Sci Rep. 2019, 9, 3096. [PubMed: 30816242]

(237). Baud A; Little D; Wen TQ; Heywood WE; Gissen P; Mills K J. Proteome Res 2019, 18 (3), 1198–1207. [PubMed: 30562036]

(238). Kristensen LP; Chen L; Nielsen MO; Qanie DW; Kratchmarova I; Kassem M; Andersen JS Mol. Cell. Proteomics 2012, 11 (10), 989–1007. [PubMed: 22801418]

(239). Kuljanin M; Elgamal RM; Bell GI; Xenocostas D; Xenocostas A; Hess DA; Lajoie GA Cell Rep. 2018, 25 (9), 2524–2536. [PubMed: 30485817]

(240). Kawahara R; Ortega F; Rosa-Fernandes L; Guimaraes V; Quina D; Nahas W; Schwammle V; Srougi M; Leite KRM; Thaysen-Andersen M; Larsen MR; Palmisano G Oncotarget 2018, 9 (69), 33077–33097. [PubMed: 30237853]

(241). Estelius J; Lengqvist J; Ossipova E; Idborg H; Le Maitre E; Andersson MLA; Brundin L; Khademi M; Svenungsson E; Jakobsson PJ; Lampa J Arthritis Res. Ther 2019, 21, 60. [PubMed: 30770760]

(242). Sajic T; Liu Y; Arvaniti E; Surinova S; Williams EG; Schiess R; Huttenhain R; Sethi A; Pan S; Brentnall TA; Chen R; Blattmann P; Friedrich B; Nimeus E; Malander S; Omlin A; Gillessen S; Claassen M; Aebersold R Cell Rep. 2018, 23 (9), 2819–2831. [PubMed: 29847809]

(243). Tu C; Rudnick PA; Martinez MY; Cheek KL; Stein SE; Slebos RJ; Liebler DC J. Proteome Res 2010, 9 (10), 4982–91. [PubMed: 20677825]

(244). Dey KK; Wang H; Niu M; Bai B; Wang X; Li Y; Cho JH; Tan H; Mishra A; High AA; Chen PC; Wu Z; Beach TG; Peng J Clin. Proteomics 2019, 16, 16. [PubMed: 31019427]

(245). Pappa KI; Kontostathi G; Makridakis M; Lygirou V; Zoidakis J; Daskalakis G; Anagnou NP Cancer Genomics Proteomics 2017, 14 (6), 507–521. [PubMed: 29109100]

(246). Shi TJ; Fillmore TL; Sun XF; Zhao R; Schepmoes AA; Hossain M; Xie F; Wu S; Kim JS; Jones N; Moore RJ; Pasa-Tolic L; Kagan J; Rodland KD; Liu T; Tang KQ; Camp DG; Smith RD; Qian WJ Proc. Natl. Acad. Sci. U. S. A 2012, 109 (38), 15395–15400. [PubMed: 22949669]

(247). Kim KH; Park GW; Jeong JE; Ji ES; An HJ; Kim JY; Yoo JS Anal. Bioanal. Chem 2019, 411 (14), 3009–3019. [PubMed: 31076819]

(248). Terkelsen T; Haakensen VD; Saldova R; Gromov P; Hansen MK; Stockmann H; Lingjaerde OC; Borresen-Dale AL; Papaleo E; Helland A; Rudd PM; Gromova I Mol. Oncol 2018, 12 (6), 972–990. [PubMed: 29698574]

(249). Mahmoud ME; Rehan IF; El-Dawy Ahmed K; Abdelrahman A; Mohammadi S; Abou-Elnaga AF; Youssef M; Diab HM; Salman D; Elnagar A; Mohammed HH; Shanab O; Ibrahim RM; Ahmed EKH; Hesham AE; Gupta A Mol. Biol. Rep 2019, 46 (3), 2733–2748. [PubMed: 30915686]

(250). Belczacka I; Pejchinovski M; Krochmal M; Magalhaes P; Frantzi M; Mullen W; Vlahou A; Mischak H; Jankowski V Proteomics: Clin. Appl 2019, 13 (3), 1800111.

(251). Togayachi A; Iwaki J; Kaji H; Matsuzaki H; Kuno A; Hirao Y; Nomura M; Noguchi M; Ikehara Y; Narimatsu H J. Proteome Res 2017, 16 (12), 4495–4505. [PubMed: 28949141]

(252). Vermassen T; Speeckaert MM; Lumen N; Rottey S; Delanghe JR Clin. Chim. Acta 2012, 413 (19–20), 1500–5. [PubMed: 22722018]

(253). Kosanovic MM; Jankovic MM J. BUON 2005, 10 (2), 247. [PubMed: 17343337]

(254). Jia G; Dong Z; Sun C; Wen F; Wang H; Guo H; Gao X; Xu C; Yang C; Sun Y Oncotarget 2017, 8 (44), 76987–76999. [PubMed: 29100363]

(255). Haga Y; Uemura M; Baba S; Inamura K; Takeuchi K; Nonomura N; Ueda K Anal. Chem 2019, 91 (3), 2247–2254. [PubMed: 30669833]

(256). Nagarajah S J. Circ. Biomark 2016, 5, 7. [PubMed: 28936255]

(257). Colombo M; Raposo G; Thery C Annu. Rev. Cell Dev. Biol 2014, 30, 255–89. [PubMed: 25288114]

(258). Hessvik NP; Llorente A Cell. Mol. Life Sci 2018, 75 (2), 193–208. [PubMed: 28733901]

(259). Chen IH; Xue L; Hsu CC; Paez JS; Pan L; Andaluz H; Wendt MK; Iliuk AB; Zhu JK; Tao WA Proc. Natl. Acad. Sci. U. S. A 2017, 114 (12), 3175–3180. [PubMed: 28270605]

(260). Liang Y; Eng WS; Colquhoun DR; Dinglasan RR; Graham DR; Mahal LK J. Biol. Chem 2014, 289 (47), 32526–37. [PubMed: 25261472]

(261). Pan D; Chen J; Feng C; Wu W; Wang Y; Tong J; Zhou D Int. J. Mol. Sci 2019, 20 (2), 323.

(262). Chauhan S; Danielson S; Clements V; Edwards N; Ostrand-Rosenberg S; Fenselau C J. Proteome Res 2017, 16 (1), 238–246. [PubMed: 27728760]

(263). Escrevente C; Grammel N; Kandzia S; Zeiser J; Tranfield EM; Conradt HS; Costa J. PLoS One 2013, 8 (10), e78631. [PubMed: 24302979]

(264). Sharma S; Alharbi M; Kobayashi M; Lai A; Guanzon D; Zuniga F; Ormazabal V; Palma C; Scholz-Romero K; Rice GE; Hooper JD; Salomon C Clin. Sci 2018, 132 (18), 2029–2044. [PubMed: 30219799]

(265). Melo SA; Luecke LB; Kahlert C; Fernandez AF; Gammon ST; Kaye J; LeBleu VS; Mittendorf EA; Weitz J; Rahbari N; Reissfelder C; Pilarsky C; Fraga MF; Piwnica-Worms D; Kalluri R Nature 2015, 523 (7559), 177–82. [PubMed: 26106858]

(266). Kuang M; Tao X; Peng Y; Zhang W; Pan Y; Cheng L; Yuan C; Zhao Y; Mao H; Zhuge L; Zhou Z; Chen H; Sun Y Clin. Proteomics 2019, 16, 5. [PubMed: 30733650]

(267). Niu L; Song X; Wang N; Xue L; Song X; Xie L Cancer Sci. 2019, 110 (1), 433–442. [PubMed: 30407700]

(268). Pisitkun T; Shen RF; Knepper MA Proc. Natl. Acad. Sci. U. S. A 2004, 101 (36), 13368–73. [PubMed: 15326289]

(269). Lim JH; Lee CH; Kim KY; Jung HY; Choi JY; Cho JH; Park SH; Kim YL; Baek MC; Park JB; Kim YH; Chung BH; Lee SH; Kim CD PLoS One 2018, 13 (9), e0204204. [PubMed: 30226858]

(270). Zou G; Benktander JD; Gizaw ST; Gaunitz S; Novotny MV Anal. Chem 2017, 89 (10), 5364–5372. [PubMed: 28402650]

(271). Zhang H; Freitas D; Kim HS; Fabijanic K; Li Z; Chen H; Mark MT; Molina H; Martin AB; Bojmar L; Fang J; Rampersaud S; Hoshino A; Matei I; Kenific CM; Nakajima M; Mutvei AP; Sansone P; Buehring W; Wang H; Jimenez JP; Cohen-Gould L; Paknejad N; Brendel M; Manova-Todorova K; Magalhaes A; Ferreira JA; Osorio H; Silva AM; Massey A; Cubillos-Ruiz JR; Galletti G; Giannakakou P; Cuervo AM; Blenis J; Schwartz R; Brady MS; Peinado H; Bromberg J; Matsui H; Reis CA; Lyden D Nat. Cell Biol 2018, 20 (3), 332–343. [PubMed: 29459780]

(272). Zhang Q; Higginbotham JN; Jeppesen DK; Yang YP; Li W; McKinley ET; Graves-Deal R; Ping J; Britain CM; Dorsett KA; Hartman CL; Ford DA; Allen RM; Vickers KC; Liu Q; Franklin JL; Bellis SL; Coffey RJ Cell Rep. 2019, 27 (3), 940–954. [PubMed: 30956133]

(273). Freitas D; Balmaña M; Poças J; Campos D; Osório H; Konstantinidi A; Vakhrushev SY; Magalhães A; Reis CA J. Extracell. Vesicles 2019, 8 (1), 1621131. [PubMed: 31236201]

(274). Thery C; Witwer KW; Aikawa E; Alcaraz MJ; Anderson JD; et al. J. Extracell. Vesicles 2018, 7 (1), 1535750. [PubMed: 30637094]

(275). Mathivanan S; Fahner CJ; Reid GE; Simpson RJ Nucleic Acids Res. 2012, 40 (D1), D1241. [PubMed: 21989406]

(276). Keerthikumar S; Chisanga D; Ariyaratne D; Al Saffar H; Anand S; Zhao K; Samuel M; Pathan M; Jois M; Chilamkurti N; Gangoda L; Mathivanan S J. Mol. Biol 2016, 428 (4), 688–692. [PubMed: 26434508]

(277). Pathan M; Fonseka P; Chitti SV; Kang T; Sanwlani R; Van Deun J; Hendrix A; Mathivanan S Nucleic Acids Res. 2019, 47 (D1), D516–D519. [PubMed: 30395310]

(278). Kim DK; Lee J; Simpson RJ; Lotvall J; Gho YS Semin. Cell Dev. Biol 2015, 40, 4–7. [PubMed: 25704310]

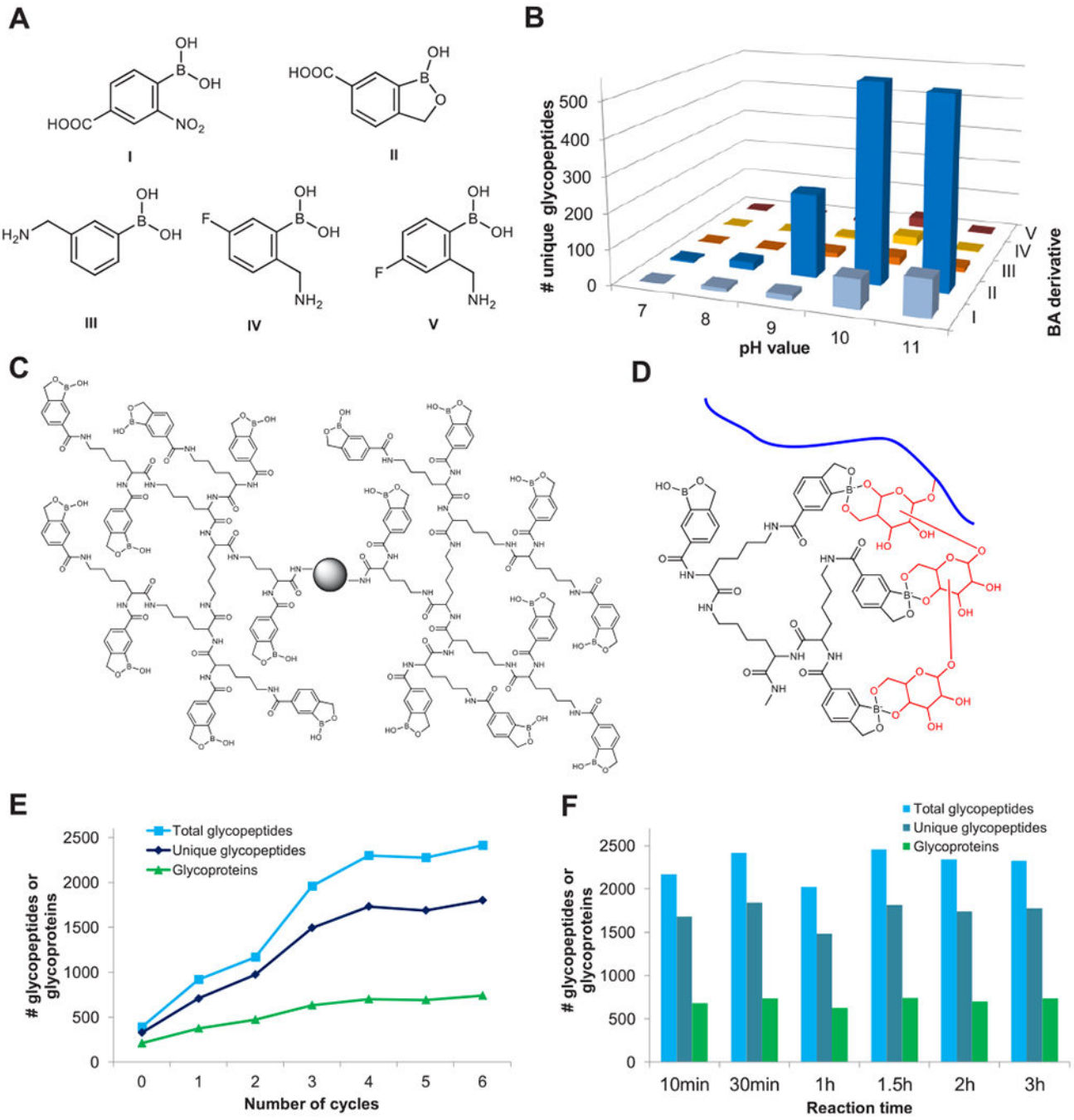


Figure 1.

Dendrimer conjugated with the boronic acid derivative (DBA) platform for glycopeptide enrichment. (A) Structures of boronic acids tested including benzoboroxole (II). (B) Number of glycopeptides identified with each type of boronic acid at different pH values from parallel experiments. (C) Structure of DBA. (D) Synergistic interactions between one glycan of a glycopeptide and multiple benzoboroxole molecules on the DBA bead. (E) Relationship between the number of synthesis cycles (dendrimer size) and the number of identified glycopeptides. (F) Effect of the reaction time on the identification of N-glycopeptides and

N-glycoproteins from parallel experiments. Reproduced from Xiao, H.; Chen, W.; Smeekens, J. M.; Wu, R., An enrichment method based on synergistic and reversible covalent interactions for large-scale analysis of glycoproteins, *Nat. Commun.* **2018**, *9*, 1692 (ref 58) under the Creative Commons Attribution 4.0 license (http://creativecommons.org/licenses/by/4.0/).

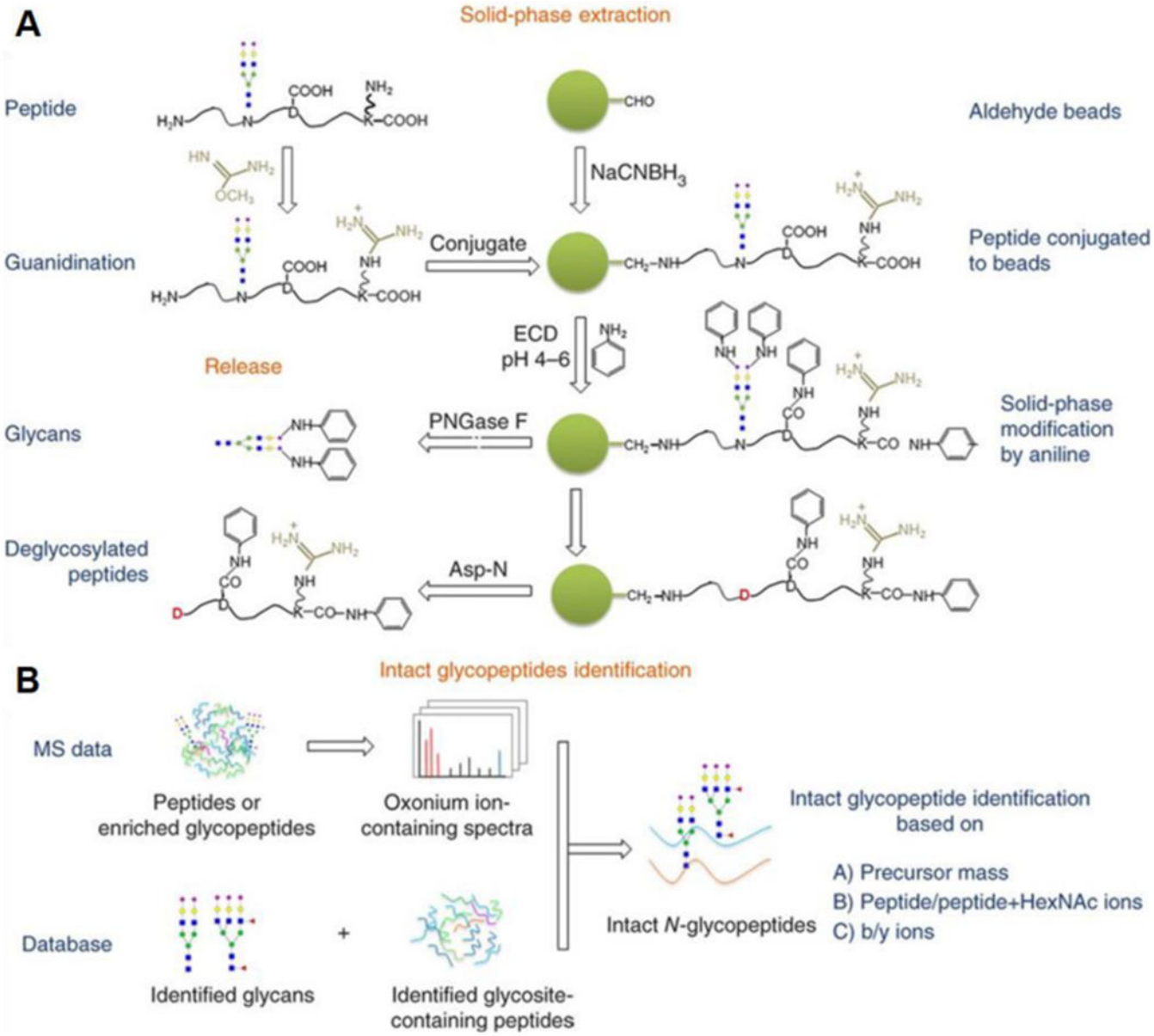


Figure 2.

(A) Workflow of the NGAG method to separate N-glycans and deglycosylated peptides. (B) Principles of NGAG to identify intact N-glycopeptides. Reprinted by permission from Springer Nature Customer Service Center GmbH: Springer Nature. Nature Biotechnology, Comprehensive analysis of protein glycosylation by solid-phase extraction of N-linked glycans and glycosite-containing peptides, Sun, S.; Shah, P.; Eshghi, S. T.; Yang, W.; Trikannad, N.; Yang, S.; Chen, L.; Aiyetan, P.; Hoti, N.; Zhang, Z.; Chan, D. W.; Zhang, H., Comprehensive analysis of protein glycosylation by solid-phase extraction of N-linked glycans and glycosite-containing peptides, *Nat. Biotechnol.* **2016**, *34* (1), 84–88 (ref 60). Copyright 2016.

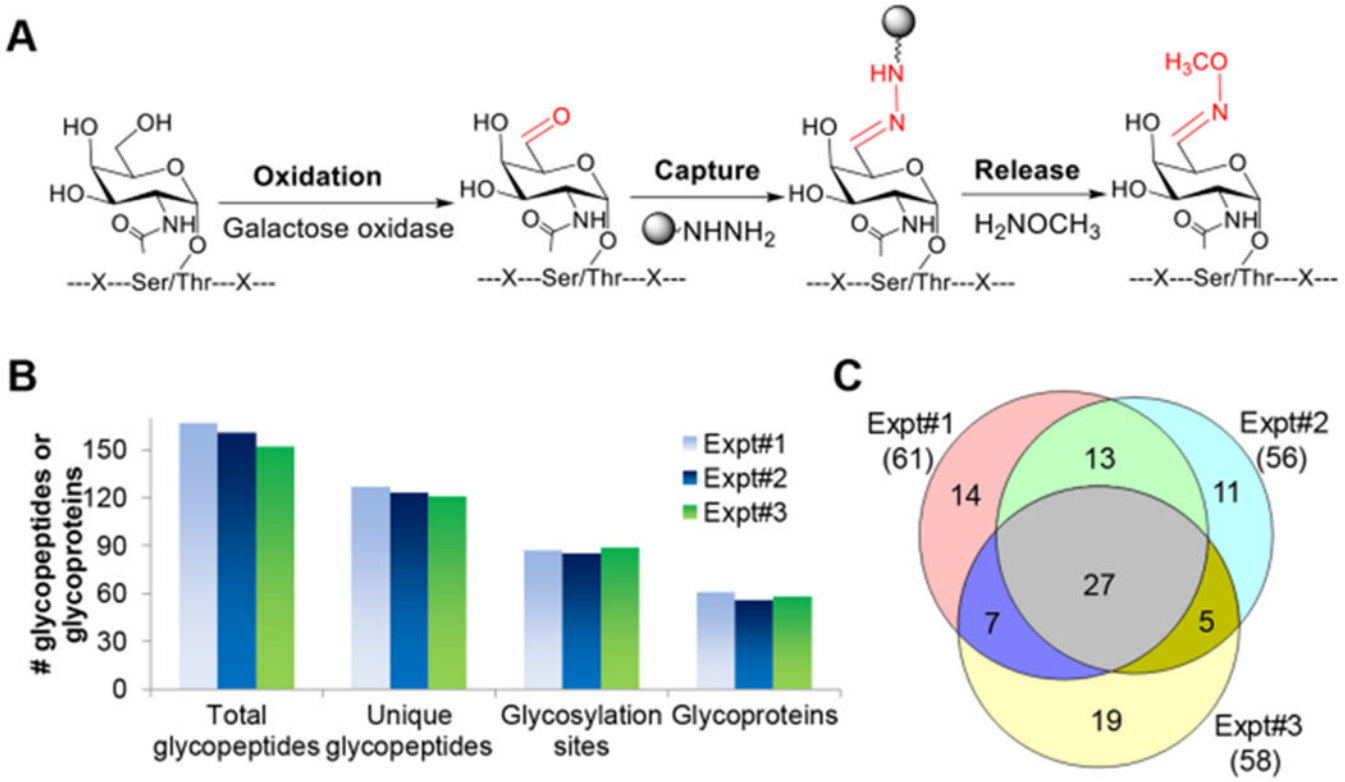
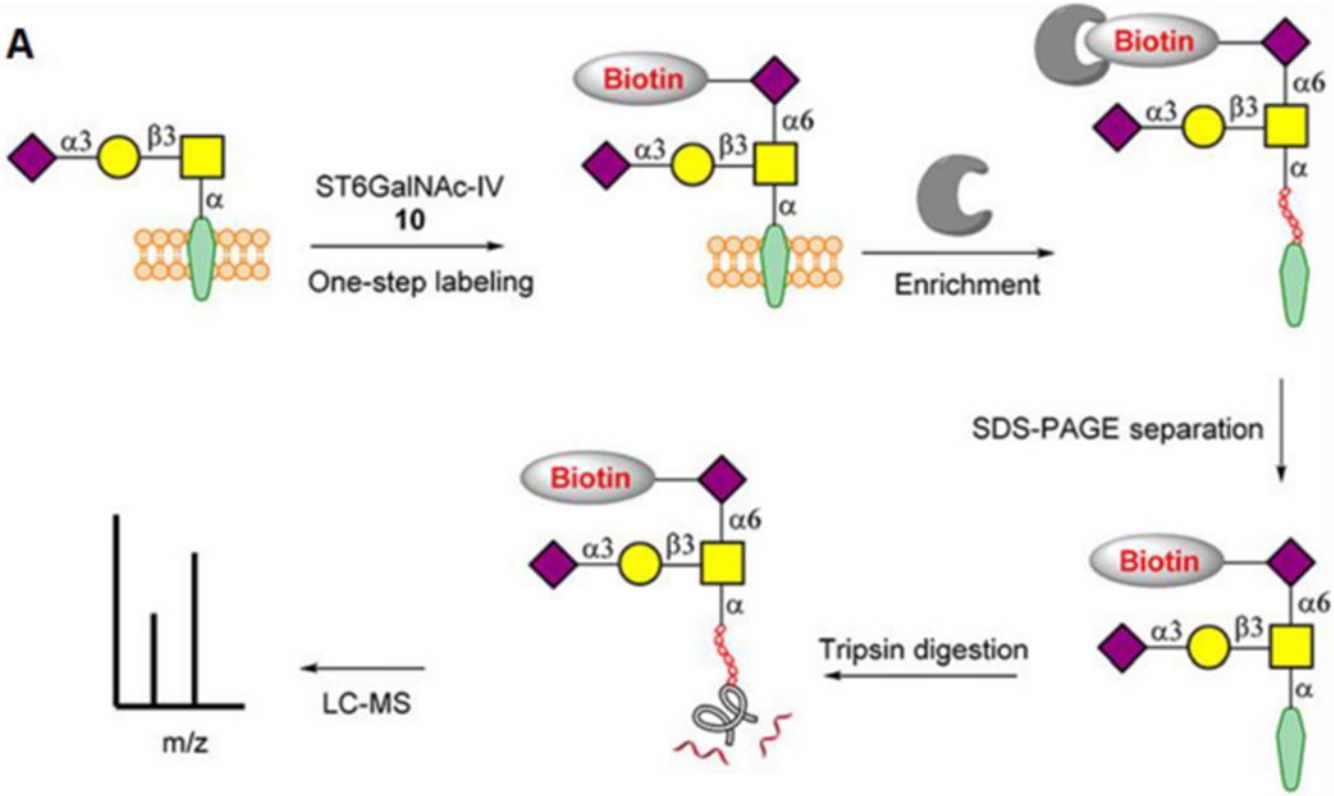
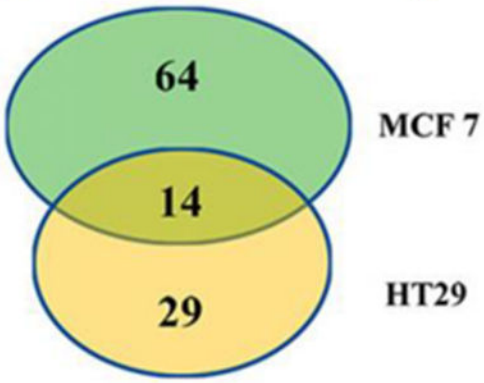


Figure 3.

(A) Principle for selective enrichment of Tn antigen-bearing glycopeptides by integrating enzymatic and chemical reactions. (B) Identification of total and unique glycopeptides, glycosylation sites, and glycoproteins with the Tn antigen from biologically triplicate experiments in Jurkat cells. (C) The overlap of glycoproteins with the Tn antigen identified from biologically triplicate experiments in Jurkat cells. Reproduced with permission from Zheng, J.; Xiao, H.; Wu, R., Specific identification of glycoproteins bearing the Tn antigen in human cells, *Angew Chem. Int. Ed.* **2017**, *56* (25), 7107–7111 (ref 79). Copyright 2017 John Wiley and Sons.



B



C

Classification	MCF7 (%)	HT29 (%)
Catalytic activity	37.5	34.1
Binding activity	37.5	29.5
Structural molecule activity	11.1	20.5
Receptor activity	5.6	9.1
Transporter activity	8.3	6.8

Figure 4.

(A) Workflow for global profiling of ST-bearing glycoproteins with the one-step chemoenzymatic labeling strategy. (B) Identification of potential ST-bearing glycoproteins from MCF7 and HT29 cells. (C) Protein classification using PANTHER. Reproduced with permission from Wen, L.; Liu, D.; Zheng, Y.; Huang, K.; Cao, X.; Song, J.; Wang, P. G., A one-step chemoenzymatic labeling strategy for probing sialylated Thomsen–Friedenreich antigen, *ACS Cent. Sci.* **2018**, *4* (4), 451–457 (ref 86). Further permissions related to the material excerpted should be directed to the ACS.

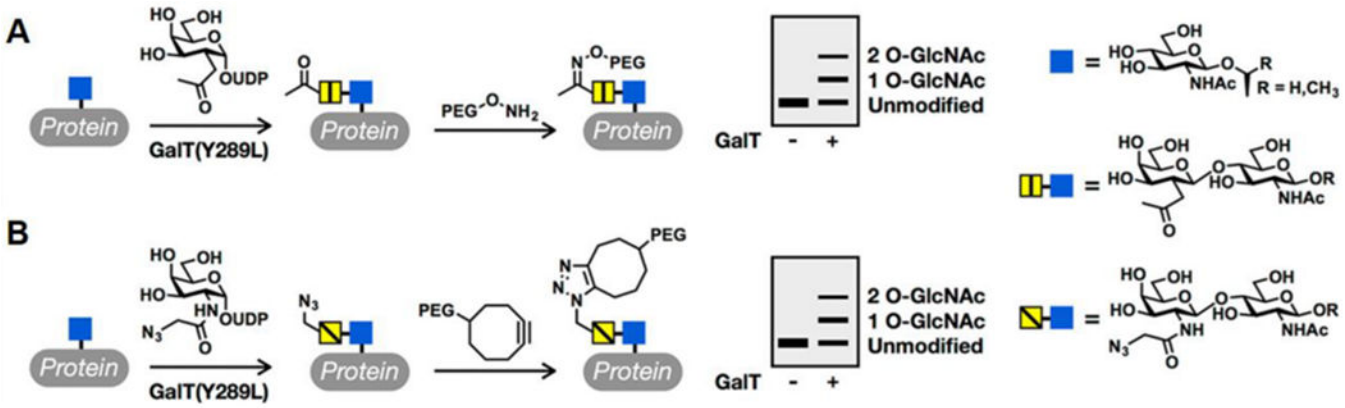


Figure 5.

Quantification of protein O-GlcNAcylation stoichiometry by integrating chemoenzymatic labeling and (A) oxime chemistry or (B) strain-promoted cycloaddition (SPAAC) chemistry. Reproduced with permission from Darabedian, N.; Thompson, J. W.; Chuh, K. N.; Hsieh-Wilson, L. C.; Pratt, M. R., Optimization of chemoenzymatic mass tagging by strain-promoted cycloaddition (SPAAC) for the determination of O-GlcNAc stoichiometry by Western blotting, *Biochemistry* **2018**, *57* (40), 5769–5774 (ref 99). Copyright 2018 American Chemical Society.

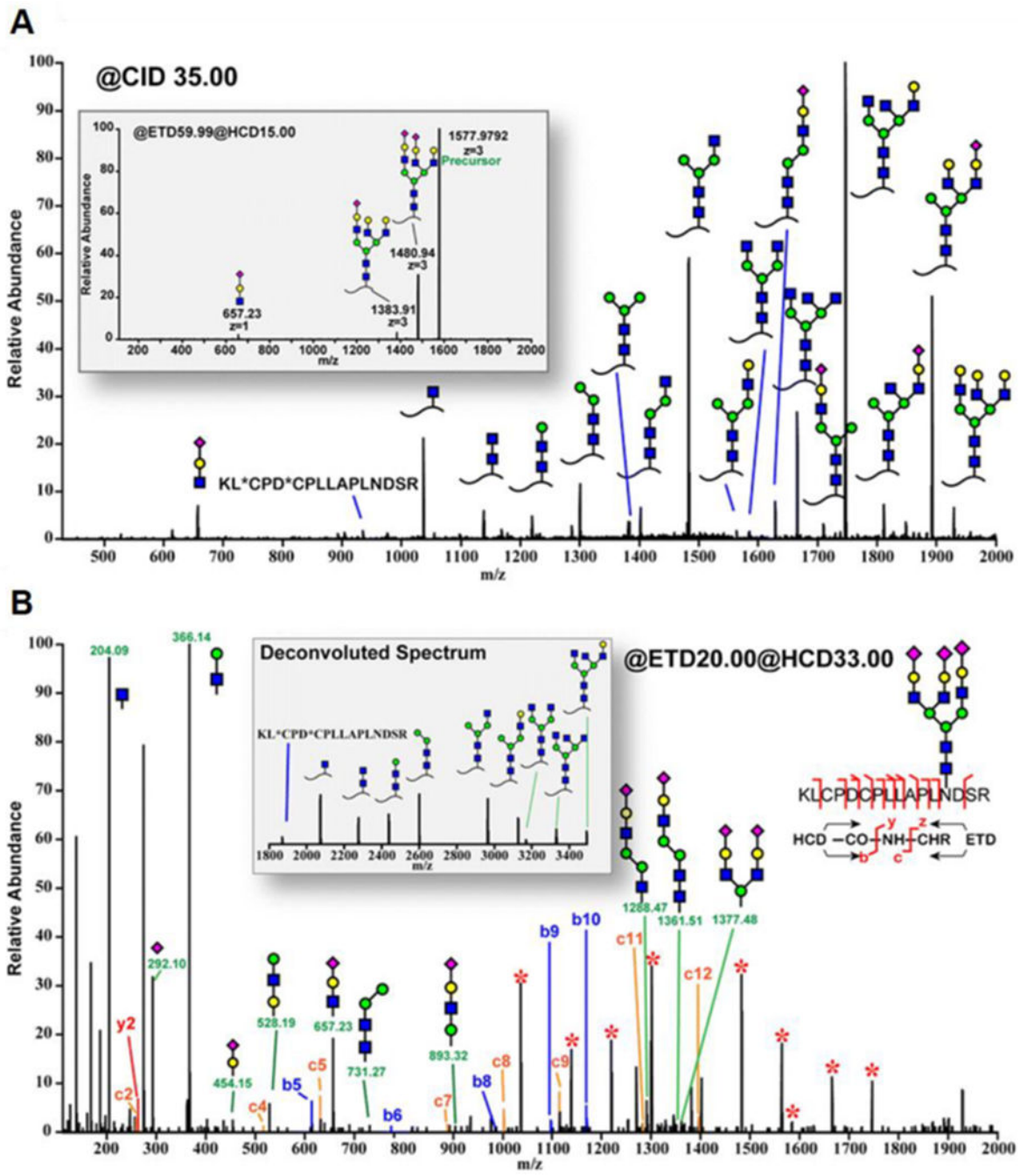


Figure 6.

MS/MS of the N-glycopeptide KL*CPD*CPLLAPLN#DSR from bovine fetuin. (A) CID and ETD spectra (inset), and (B) EThcD spectrum. # and * represent glycosylation and carbamidomethylation sites, respectively. Reprinted by permission from Springer Nature Customer Service Center GmbH: Springer Nature. Journal of The American Society for Mass Spectrometry, Electron-transfer/higher-energy collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization, Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L., Electron-transfer/higher-energy

collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization, *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1751–1764 (ref 114). Copyright 2017.

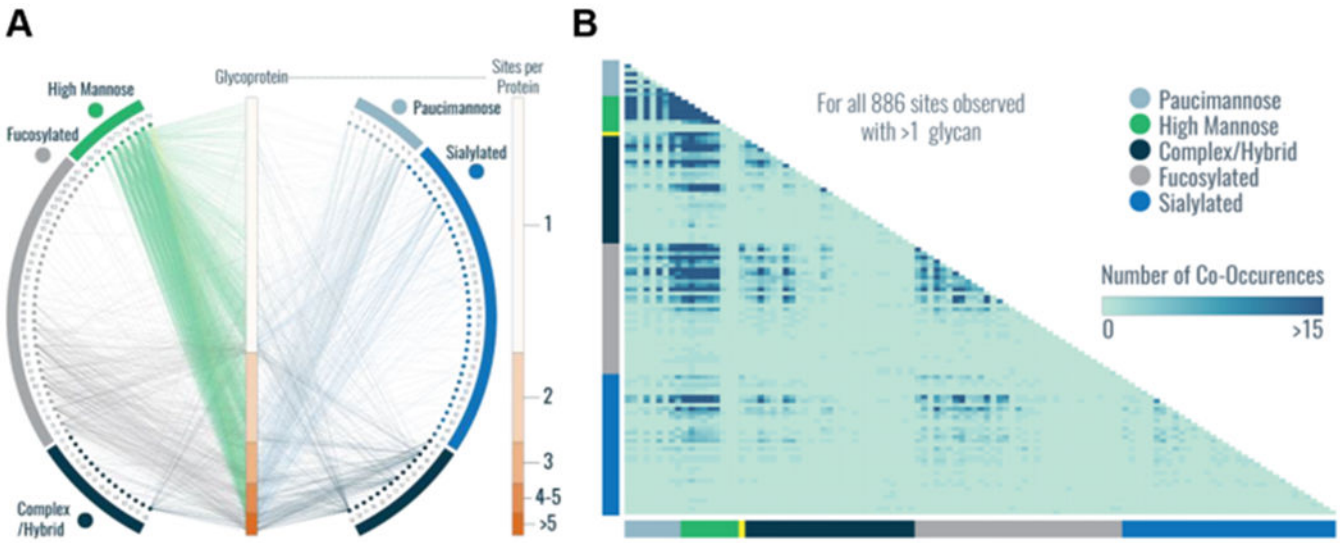


Figure 7.

(A) Network map between glycoproteins and glycans. The edges between different glycans and glycoproteins are colored by the original glycan node, except that the edges from mannose-6-phosphate are yellow. (B) Glycan co-occurrence heat map, which displays the frequency of glycan pairs appeared together at the same glycosylation site. Reproduced from Riley, N. M.; Hebert, A. S.; Westphall, M. S.; Coon, J. J., Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis, *Nat. Commun.* **2019**, *10*, 1311 (ref 127) under the Creative Commons Attribution 4.0 license (http://creativecommons.org/licenses/by/4.0/).

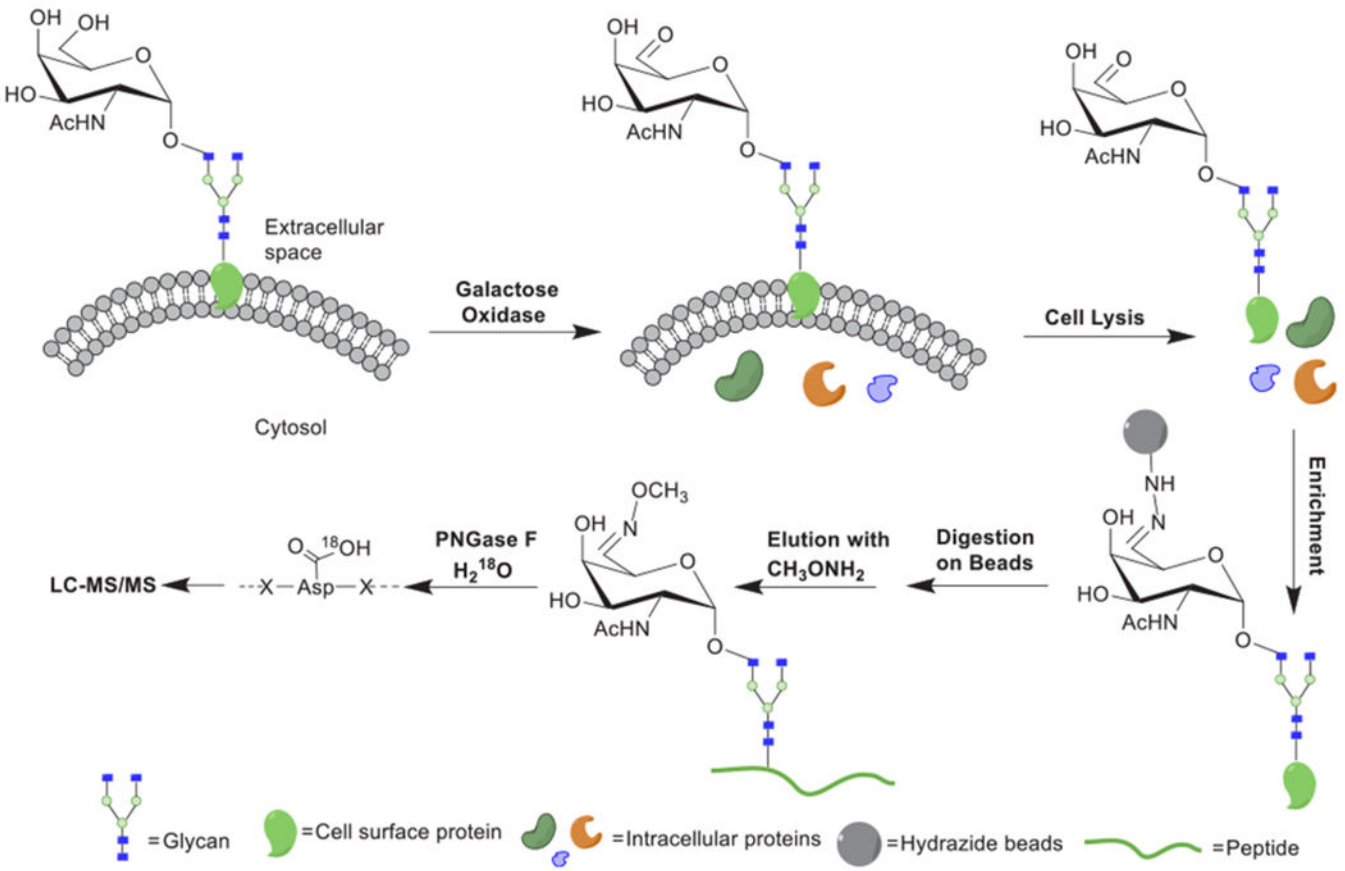


Figure 8.

Enzymatic approach to identify cell-surface glycoproteins. Glycans on the cell surface were first oxidized with galactose oxidase (GAO) and immobilized on hydrazide beads after cell lysis. Surface glycopeptides were eluted from the beads with methoxylamine before PNGase F treatment and LC–MS/MS analysis. Reproduced with permission from Sun, F.; Suttapitugsakul, S.; Wu, R., Enzymatic Tagging of Glycoproteins on the Cell Surface for Their Global and Site-Specific Analysis with Mass Spectrometry, *Anal. Chem.* **2019**, *91* (6), 4195–4203 (ref 45). Copyright 2019 American Chemical Society.

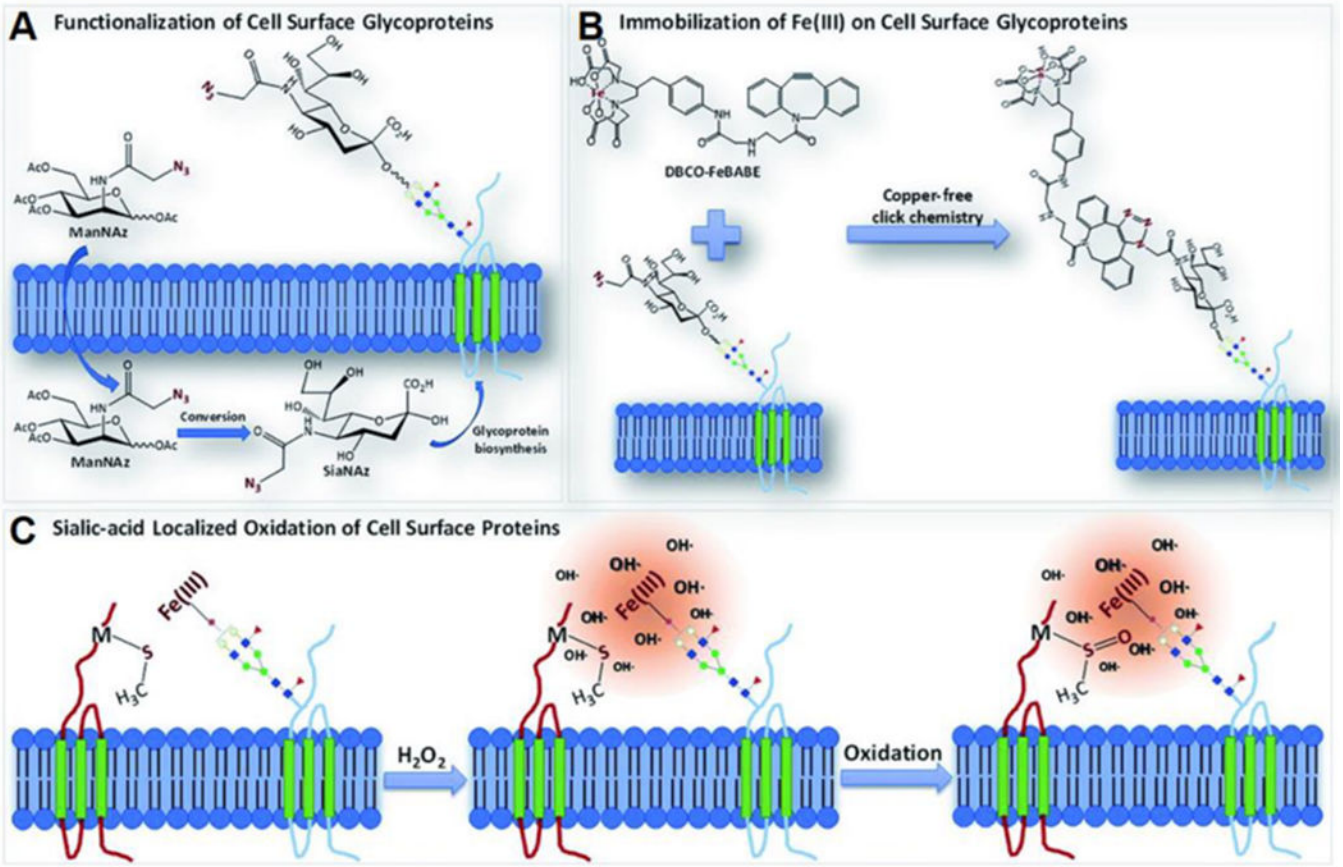


Figure 9.

Protein oxidation of sialic acid environments (POSE) approach to map protein environment of sialic acid. Cells were first labeled with ManNAz (A) and tagged with DBCO-FeBABE (B) before hydrogen peroxide treatment (C). Amino acids in the vicinity were oxidized by radical species and proteins were subsequently analyzed with LC–MS/MS. Reproduced from Li, Q.; Xie, Y.; Xu, G.; Lebrilla, C. B., Identification of potential sialic acid binding proteins on cell membranes by proximity chemical labeling, *Chem. Sci.* **2019**, *10*, 6199–6209 (ref 216). Copyright 2019, with permission from the Royal Society of Chemistry.

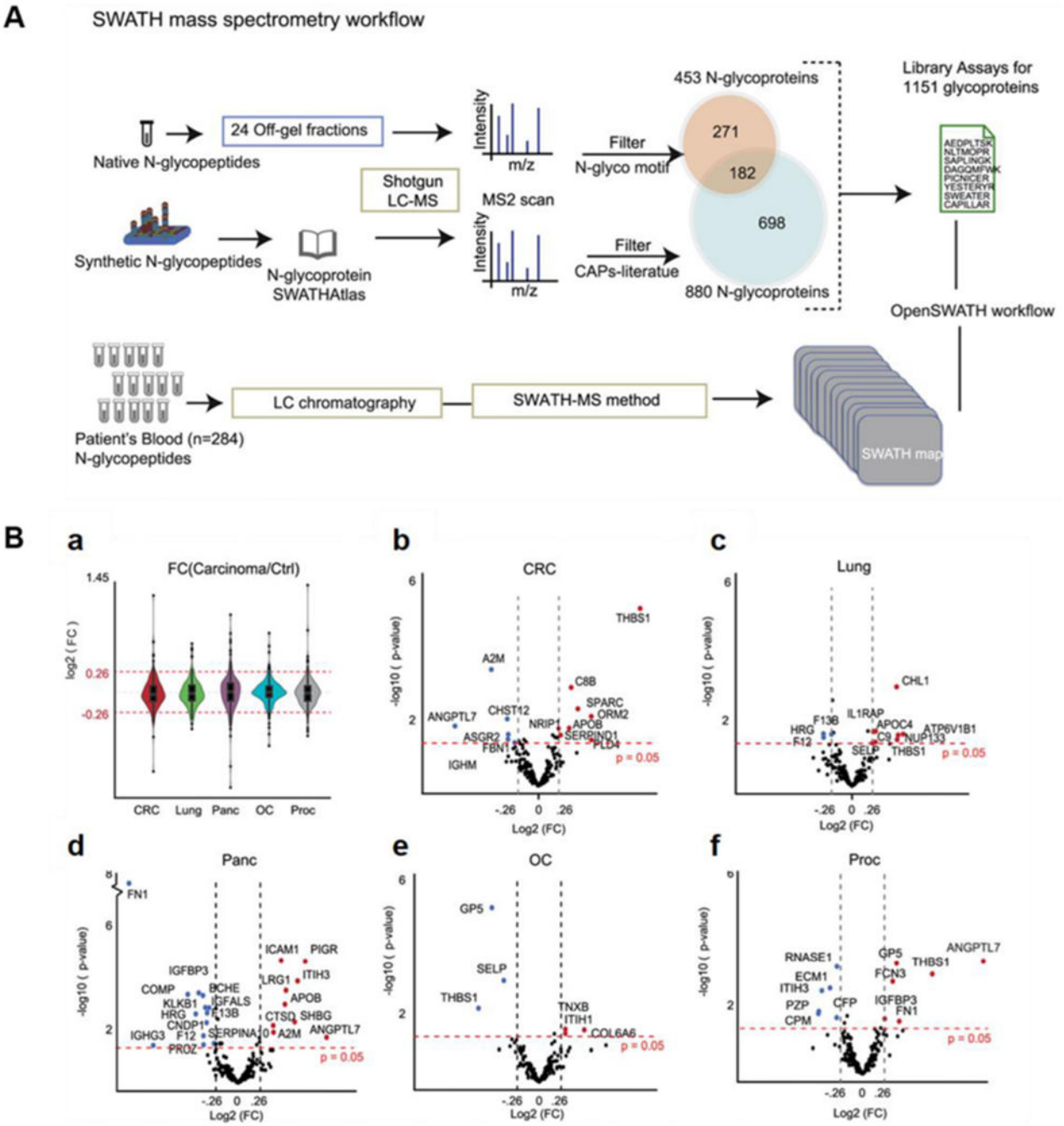


Figure 10.

(A) SWATH-MS workflow for the identification of glycoproteins from blood of patients with colorectal cancer (CRC), lung cancer, pancreatic cancer (Panc), ovary cancer (OC), and prostate cancer (Proc). The spectral library was built from a data-dependent acquisition and extensive fractionation of native N-glycopeptides from natural blood samples and synthetic peptides. Blood samples were then analyzed using SWATH-MS and the OpenSWATH tool. (B) Glycoprotein expression across different types of cancer. (a) Violin plots showing the distributions of fold changes of cancer/control samples. (b–f) Volcano plots showing

differential regulations of glycoproteins from the five samples. FC means fold change. Reprinted from Sajic, T.; Liu, Y.; Arvaniti, E.; Surinova, S.; Williams, E. G.; Schiess, R.; Huttenhain, R.; Sethi, A.; Pan, S.; Brentnall, T. A.; Chen, R.; Blattmann, P.; Friedrich, B.; Nimeus, E.; Malander, S.; Omlin, A.; Gillessen, S.; Claassen, M.; Aebersold, R., Similarities and differences of blood N-glycoproteins in five solid carcinomas at localized clinical stage analyzed by SWATH-MS, *Cell Rep.* **2018**, *23* (9), 2819–2831 (ref 242). Copyright 2018, with permission from Elsevier.