Received: 30 January 2021 | Revised: 10 May 2021 | Accepted: 13 May 2021
DOI: 10.1002/mas.21708

REVIEW ARTICLE

WILEY

Systematic characterization of extracellular glycoproteins using mass spectrometry

Fangxu Sun | **Suttipong Suttapitugsakul** | **Ronghu Wu**

School of Chemistry and Biochemistry and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia, USA

Correspondence
Ronghu Wu, School of Chemistry and Biochemistry and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332, USA.
Email: ronghu.wu@chemistry.gatech.edu

Abstract

Surface and secreted glycoproteins are essential to cells and regulate many extracellular events. Because of the diversity of glycans, the low abundance of many glycoproteins, and the complexity of biological samples, a system-wide investigation of extracellular glycoproteins is a daunting task. With the development of modern mass spectrometry (MS)-based proteomics, comprehensive analysis of different protein modifications including glycosylation has advanced dramatically. This review focuses on the investigation of extracellular glycoproteins using MS-based proteomics. We first discuss the methods for selectively enriching surface glycoproteins and investigating protein interactions on the cell surface, followed by the application of MS-based proteomics for surface glycoprotein dynamics analysis and biomarker discovery. We then summarize the methods to comprehensively study secreted glycoproteins by integrating various enrichment approaches with MS-based proteomics and their applications for global analysis of secreted glycoproteins in different biological samples. Collectively, MS significantly expands our knowledge of extracellular glycoproteins and enables us to identify extracellular glycoproteins as potential biomarkers for disease detection and drug targets for disease treatment.

KEYWORDS
biomarker, cell surface glycoproteins, comprehensive analysis, enrichment methods, MS-based proteomics, secreted glycoproteins

1 | INTRODUCTION

The majority of extracellular proteins are modified with glycosylation, including *N*- and *O*-linked glycosylation where glycans are covalently attached to the asparagine or serine/threonine residues, respectively (Schjoldager et al., 2020; Spiro, 2002). Unlike other biomolecules including nucleic acids and proteins, the synthesis of glycans is non-template-driven. Extracellular glycoproteins are normally secreted through the classical secretory pathway, in which glycans are added or trimmed by various glycosyltransferases and glycosidases in the endoplasmic reticulum and the Golgi. These glycoproteins are eventually transported to the plasma membrane or the extracellular space (Burda & Aebi, 1999; Moremen et al., 2012; Reily et al., 2019). Extracellular glycoproteins are involved in many important cellular events such as cell–cell communications, cell migration, and immune response (Haltiwanger & Lowe, 2004; Varki, 2016). Aberrant glycosylation of surface and secreted proteins is associated with different diseases including cancers, immune disorders, and Alzheimer's disease, and thus these glycoproteins can be used as biomarkers for disease diagnosis, especially in the early stage (Nie et al., 2014, Vajaria & Patel, 2017; Veillon et al., 2018).

With the development of modern instrumentation and data analysis tools, mass spectrometry (MS)-based proteomics has become very powerful to globally and site-specifically analyze protein modifications including phosphorylation, ubiquitination, acetylation, and glycosylation, which greatly advances our understanding of the functions of protein modifications in biological systems (Chen et al., 2018; Riley et al., 2019; Suttapitugsakul et al., 2020; Witze et al., 2007; Wu et al., 2011a; Wu et al., 2011b; Xiao et al., 2019; Xu et al., 2020; Yu et al., 2018; Zhu et al., 2019). However, it is still very challenging to comprehensively analyze extracellular glycoproteins because of their low abundance, the glycan heterogeneity, and the extracellular specificity. The abundances of many surface and secreted glycoproteins are very low and could be several orders of magnitude lower than the unmodified ones. For cells grown in serum-containing media, secreted proteins (in the ng/ml range) are far less abundant compared with the background consisting of many serum proteins (in the mg/ml range) (Eichelbaum et al., 2012). The occupancy of glycan on a specific glycosylation site and the different glycans located on the same site significantly increase the number of glycoforms of a protein, which further complicate the analysis of extracellular glycoproteins (Caval et al., 2020). The extracellular specificity is also critical because highly abundant intracellular proteins could dominate MS spectra and they cannot be distinguished from extracellular ones during the MS analysis. Therefore, it is very necessary to selectively enrich extracellular glycoproteins before MS analysis to minimize their interference.

In this review, we highlight the development in MS-based methods for a comprehensive investigation of extracellular glycoproteins. In the first part, we discuss methods to specifically analyze surface glycoproteins, followed by the description of surface glycoprotein interactions and the discussion of glycoprotein dynamics and biomarker discovery. We then summarize the methods for global analysis of secreted glycoproteins, including the method development and the investigation of secreted glycoproteins in different samples.

2 | COMPREHENSIVE ANALYSIS OF GLYCOPROTEINS ON THE CELL SURFACE USING MS-BASED GLYCOPROTEOMICS

2.1 | Enrichment methods for profiling the surface glycoproteome

Enrichment of surface glycoproteins is of great importance to achieve their comprehensive analysis by MS because of the sub-stoichiometry of glycosylation and the low abundance of many glycoproteins compared with the highly abundant intracellular proteins. Furthermore, MS itself cannot distinguish between surface and intracellular glycoproteins. Different enrichment methods have been developed to facilitate the investigation of surface glycoproteins (Table 1).

2.1.1 | Methods for enriching cell surface proteins

Due to their unique properties, various methods were developed to target surface proteins. Ultracentrifugation is a classical method to separate the plasma membrane from the cells (Lebowitz et al., 2002; Liang et al., 2006).

TABLE 1 A summary of some enrichment methods for global analysis of surface glycoproteins[a]

Method	Specificity	Advantages	Potential limitations
Traditional methods (ultracentrifugation, *N*-hydroxysuccinimide labeling, etc.)	Membrane proteins	Sample preparation is convenient; reagents and instrument are easily accessible	Not specific for cell surface glycoproteins
Cell-surface capture	Surface glycoproteins	Widely used for studying surface glycoproteins from various sample types; high labeling efficiency and enrichment specificity	Relatively harsh labeling conditions due to periodate oxidation; requirement of large sample amount
Enzyme-based methods	Surface glycoproteins	High enzymatic labeling efficiency and mild reaction conditions	Limited accessibility of glycosyltransferases and sugar analogs
Metabolic labeling	Surface glycoproteins	Very useful method to site-specifically investigate surface glycoproteins	Varied metabolic labeling efficiencies in different cell lines

[a]Detailed information and cited papers are included in the main text.

To increase the purity of the plasma membrane fraction, density-gradient centrifugation, particularly through the sucrose gradient medium, was employed to enrich cell surface proteins (Blonder et al., 2006; Lund et al., 2009). However, ultracentrifugation cannot completely eliminate soluble or cytosolic proteins in the plasma membrane fraction. Furthermore, protein biotinylation was employed to specifically tag the extracellular regions of proteins located on the cell surface, followed by affinity enrichment (Elia, 2008; Li et al., 2019b; Qiu & Wang, 2008; Zhao et al., 2004). Cationic colloidal silica-bead coating based on the anionic nature of the plasma membrane and enzymatic shaving relying on enzymes to remove the exposed extracellular regions of cell surface proteins have also been reported for surface protein enrichment (Prior et al., 2011; Solis & Cordwell, 2011).

However, the aforementioned methods aim to separate all plasma membrane proteins and do not specifically target cell surface glycoproteins. Nevertheless, these methods were recently integrated with other glycoprotein/glycopeptide-enrichment strategies such as hydrophilic interaction liquid chromatography (HILIC) and lectin-based methods to investigate cell surface glycoproteins. After ultracentrifugation and protein digestion, the resulting glycopeptides were enriched with HILIC for surface glycoprotein analysis (Park et al., 2018). The combination of sulfo-*N*-hydroxysuccinimide (NHS)-SS-biotin labeling and peanut agglutinin (PNA)-lectin purification enabled a comprehensive analysis of surface proteins containing the Thomsen–Friedenreich (TF) antigen (Li et al., 2017).

2.1.2 | Cell surface capture (CSC)

Carbohydrates on glycoproteins have different chemical reactivities compared with proteins, which offers the possibility of developing methods to target the carbohydrate moiety for glycoprotein enrichment. Several decades ago, it was found that the *cis*-diols on carbohydrates can be oxidized into aldehyde groups in the presence of sodium periodate (Bobbitt, 1956; Gahmberg & Andersson, 1977). Integrating periodate oxidation and hydrazide chemistry, Zhang et al. (2003) first developed a beautiful method, i.e., the hydrazine chemistry-based method, to enrich glycoproteins, which has been applied to a wide range of samples including sera, cells, and tissues (Sun et al., 2019b).

On the basis of the oxidation of carbohydrates, Wollscheid et al. (2009) introduced the CSC technology for large-scale analysis of surface glycoproteins. They first systematically optimized parameters for oxidizing cell surface glycoproteins including the concentration of periodate, and the reaction conditions to maintain the cell integrity and minimize the side reactions. Then, surface glycoproteins containing the aldehyde groups were covalently labeled with biocytin hydrazide (BH) on live cells. After cell lysis and protein digestion, the biotinylated glycopeptides were enriched with streptavidin beads, followed by glycopeptide elution through the PNGase F treatment. CSC can achieve the site-specific analysis of cell surface glycoproteins, which significantly decreased the false positive rate for surface glycoprotein identification. CSC has been used to investigate surface glycoproteins from different sample types including mouse and human cell lines, and primary cells (Gundry et al., 2012; Haverland et al., 2017). Bausch-Fluck et al. (2015) applied CSC to analyze surface glycoproteins on 41 human and 31 mouse cell lines, and identified 1492 and 1296 glycoproteins, respectively. They then built an MS-derived Cell Surface Protein Atlas (CSPA), which provides valuable information for the classification of different cell types and can also be used as a platform for screening drug targets on the cell surface. Using CSPA as a training data set for machine learning, the same group also developed a tool called SURFY to predict 2886 proteins on the cell surface with an accuracy of over 93% (Bausch-Fluck et al., 2018).

To improve its performance, different versions of CSC have been reported, which mainly focus on optimizing the conditions for tagging the aldehyde groups such as the addition of a catalyst, the utilization of different biotinylation reagents, and the enrichment at the peptide or protein level. Kalxdorf et al. (2017) used alkoxyamine-PEG$_4$-biotin instead of BH to covalently label the aldehyde groups on the cell surface. Additionally, aniline was used as a catalyst to increase the labeling rate (Kalxdorf et al., 2017). They enriched the biotinylated surface glycoproteins at the protein level instead of the peptide level, and utilized the non-glycosylated peptides directly derived from on-bead digestion for surface glycoprotein identification and quantification. The optimized protocol allowed for the identification of more plasma membrane proteins because it eliminated the deamination reaction during glycopeptide elution with the PNGase F treatment. However, without the site-specific analysis of glycoproteins, the experimental evidence for glycopeptide identification was lost, and the enrichment specificity of surface glycoproteins was compromised (~70% vs. ~95% in CSC).

Although the CSC method is very effective to globally analyze cell surface glycoproteins and offers information on the topology of proteins in the extracellular region based on the identified glycosylation site, it is often applied to study cultured cells but not clinical samples due to the limited accessibility of the reagents to cell surface glycoproteins from tissues and the large amount of starting materials normally required (more than 10^7 cells for each sample). Recently, a new version of CSC termed autoCSC was introduced by incorporating miniaturization and automation into the workflow (van Oostrum et al., 2019). After oxidative labeling, cell lysis, and protein digestion, the biotinylated peptides can be automatically processed by a robot including streptavidin resin binding, beads washing, and peptide elution, which can minimize the sample loss during the manual sample preparation and increase the sensitivity of CSC (Figure 1). With autoCSC, they achieved a fivefold increase in sensitivity and a decreased coefficient of variation compared with manual sample preparation. The autoCSC method enabled the quantification of 248 unique glycosylation sites located on 147 proteins from only 1×10^6 developing B cells per sample. The Gundry group recently developed an automated bioinformatic workflow to rapidly analyze the data from CSC and facilitate the discovery of cell-type-specific surface markers by integrating proteomic and transcriptomic data (Waas et al., 2020). Complete automation of sample preparation and bioinformatic analysis is expected to further improve the efficiency of CSC in the future.

2.1.3 | Enzyme-based methods

Compared with the oxidation by periodate, the conditions of enzymatic labeling of surface glycoproteins are milder and the reaction could be more efficient. Galactose oxidase (GAO) is a copper-dependent enzyme that can specifically oxidize the hydroxyl group at the C6 position on galactose/*N*-acetylgalactosamine (Gal/GalNAc) to an aldehyde group (Parikka et al., 2015; Zheng et al., 2017). The enzyme can be used to tag glycoproteins on the cell surface for their large-scale analysis (Ramya et al., 2013). By systematically optimizing the conditions, our group dramatically improved the performance of the GAO-based method (Sun et al., 2019a). Horseradish peroxidase was incorporated to increase the glycan oxidation efficiency by GAO. After on-bead digestion, methoxylamine was used to elute glycopeptides from the hydrazide beads, instead of the commonly used PNGase F treatment, to increase the recovery rate of glycopeptides. When sialidase was employed to remove sialic acid at the termini of glycans, the method became even more effective. Integrating the GAO-based method and stable isotope labeling by amino acids in cell culture (SILAC), the systematic analysis of surface glycoprotein trafficking was achieved, which revealed that most surface glycoproteins were transported to the plasma membrane through the classical secretory pathway.

Chemoenzymatic labeling is another very popular strategy to investigate glycoproteins on the cell surface. Here, a sugar analog with a chemical reporter from a nucleotide sugar donor is transferred to a specific glycan acceptor on the cell surface by a recombinant glycosyltransferase (Lopez Aguilar et al., 2017). Yu et al. (2016) developed a two-step selective

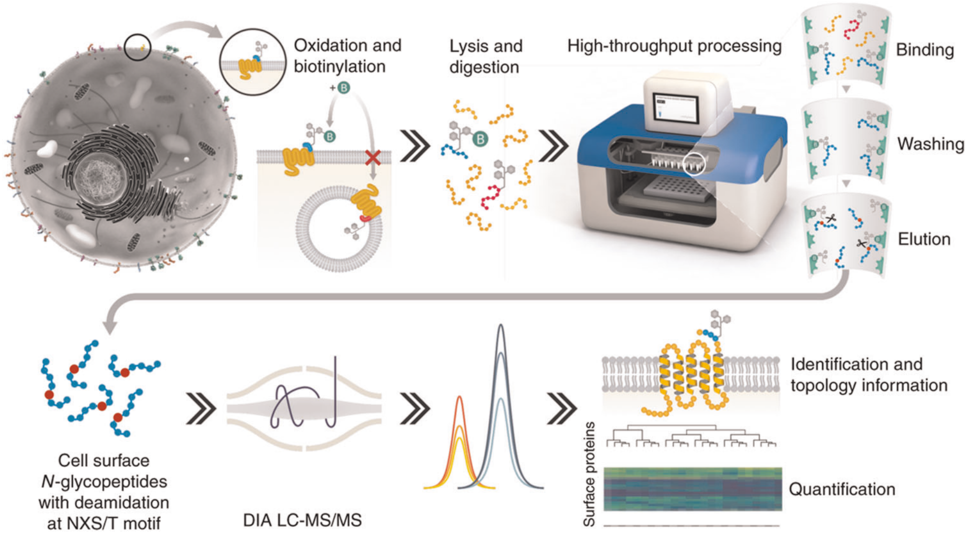

FIGURE 1 Workflow of autocell-surface capture. The steps include glycan oxidation, cell lysis and protein digestion, automatic process of enrichment, and data-independent acquisition (DIA)-mass spectrometry (MS) analysis. Reproduced with permission from van Oostrum et al. (2019). Copyright retained by the authors under a CC BY 4.0 license [Color figure can be viewed at wileyonlinelibrary.com]

exo-enzymatic labeling (SEEL) method to study surface glycoprotein changes in the differentiation of human erythroleukemia cells, in which recombinant α-(2,6)-sialyltransferase (ST6Gal1) or α-(2,3)-sialyltransferase (ST3Gal1) and a sialic acid nucleotide analog CMP-Neu5Ac modified by an azide group were employed to label cell surface glycans, followed by biotinylation of the tagged surface glycoproteins through click chemistry, enrichment, and MS identification. Neuraminidase treatment before the chemoenzymatic labeling can markedly facilitate the identification of surface glycoproteins. To minimize the adverse effects of click chemistry in the two-step SEEL method such as side reactions and low labeling efficiency, Sun et al. (2016) synthesized CMP-Neu5Ac modified by a biotin moiety for direct labeling of surface glycoproteins with recombinant ST6Gal1 (Figure 2A). The one-step SEEL displayed a higher labeling efficiency compared with the two-step SEEL approach and enabled the identification of more surface glycoproteins. They also used this method to comprehensively study how surface glycoproteins responded to the lysosomal dysfunction and found that the internalization of glycoproteins such as IGF2R and EPHA2 within the endolysosomal system for degradation was reduced with the lysosomal disruption. As the glycosyltransferases in chemoenzymatic labeling are able to recognize specific glycans, the method can also be utilized to analyze surface glycoproteins containing certain glycans (Wen et al., 2016). Sialyltransferase ST6GalNAc-IV specifically recognizing the sialylated TF (Sialyl-T) antigen can transfer a sialic acid analog containing a biotin moiety to the Sialyl-T antigens on the cell surface (Figure 2B) (Wen et al., 2018). The tagged proteins with the Sialyl-T antigen were pulled down using avidin resins, followed by digestion on beads and MS analysis, which enabled the identification of 78 and 43 potential glycoproteins containing the Sialyl-T antigen on the cell surface from MCF7 and HT29 cells, respectively. Among the identified proteins, those related to binding activity and catalytic activity were highly enriched in both cell lines based on the bioinformatic analysis. Although chemoenzymatic labeling is effective and specific to analyze surface glycoproteins, the limited availability of the recombinant glycosyltransferases and the modified nucleotide sugar donors restricts its wide applications.

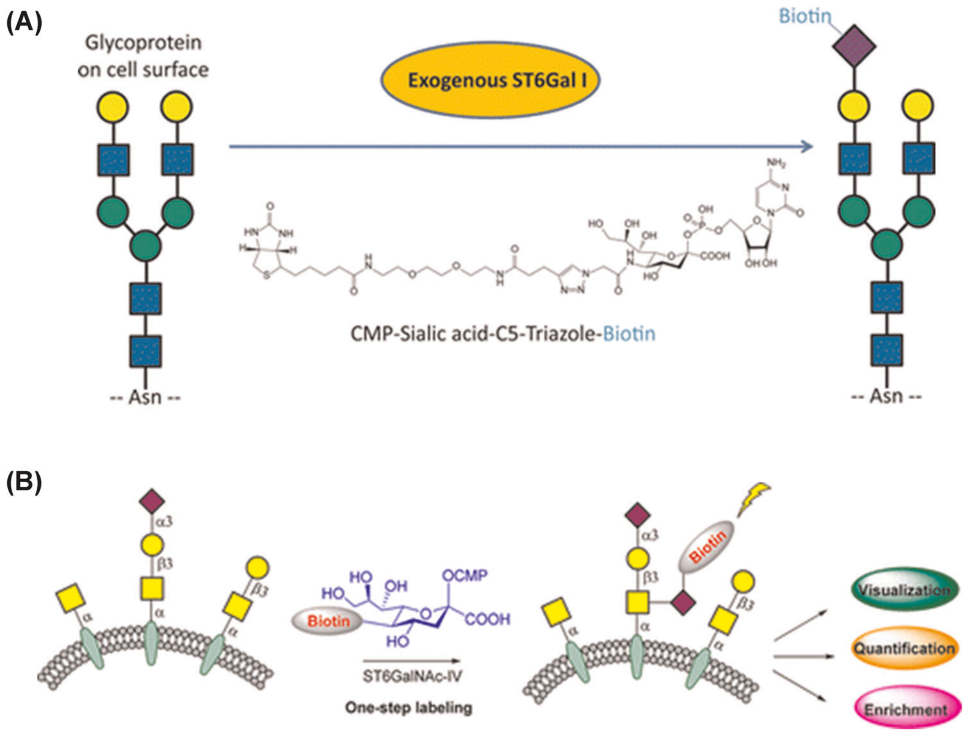

FIGURE 2 (A) One-step selective exo-enzymatic labeling (SEEL) for cell surface glycoprotein profiling. Reprinted with permission from Sun et al. (2016). Copyright retained by American Chemical Society. (B) Workflow for the analysis of Sialyl-T antigen-bearing surface proteins with chemoenzymatic labeling. Reprinted with permission from Wen et al. (2018). Copyright retained by American Chemical Society. GalNAc (yellow circle), sialic acid (purple diamond), and sialylated Thomsen–Friedenreich (purple diamond-yellow circle-yellow square) [Color figure can be viewed at wileyonlinelibrary.com]

2.1.4 | Integrating metabolic labeling and bioorthogonal chemistry to target surface glycoproteins

With metabolic labeling, a sugar analog modified with a chemical reporter can be taken up by cells and then incorporated into glycoproteins, including those located on the cell surface, allowing for the specific analysis of the surface glycoproteome (Wang & Mooney, 2020). The glycan biosynthetic pathway in cells can tolerate different unnatural sugar analogs with chemical reporters including *N*-azidoacetylmannosamine (ManNAz), *N*-azidoacetylgalactosamine (GalNAz), and *N*-azidoacetylglucosamine (GlcNAz) (Dube & Bertozzi, 2003; Mahal et al., 1997). The chemical reporters in the sugar analogs are critical for the downstream enrichment of surface glycoproteins. They cannot be present in biological systems or react with other chemical functionalities in living cells, which ensures that the tagging of glycoproteins is specific and selective (Prescher & Bertozzi, 2005). Subsequently, a bioorthogonal reaction that proceeds without disrupting the native biological system is performed to install an affinity group such as biotin for the enrichment of surface glycoproteins. On the basis of the chemical reporters in sugar analogs, different bioorthogonal reactions have been developed for labeling surface glycoproteins, such as Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAC), and strain-promoted azide–alkyne cycloaddition (SPAAC) (Lang & Chin, 2014; Sletten & Bertozzi, 2009). The tagged surface glycoproteins can be selectively enriched for MS identifications and quantifications.

With the development of MS-based proteomics, site-specific analysis of surface glycoproteins has been growing rapidly. It not only provides valuable information about the glycosylation sites but also offers direct experimental evidence for the identification of glycopeptides and glycoproteins (Sun & Wu, 2019; Xiao & Wu, 2017c; Xiao et al., 2018c; Xu et al., 2021). Our group has developed an effective method to comprehensively and site-specifically investigate glycoproteins located on the cell surface by combining metabolic labeling, bioorthogonal chemistry, and MS-based proteomics (Figure 3) (Chen et al., 2015; Smeekens et al., 2015; Xiao et al., 2016b). After metabolic labeling using a sugar analog containing the azide group (GalNAz, GlcNAz, or ManNAz), surface glycoproteins can be tagged with dibenzocyclooctyne (DBCO)-sulfo-biotin through SPAAC under physiological conditions, which maintains the viability of

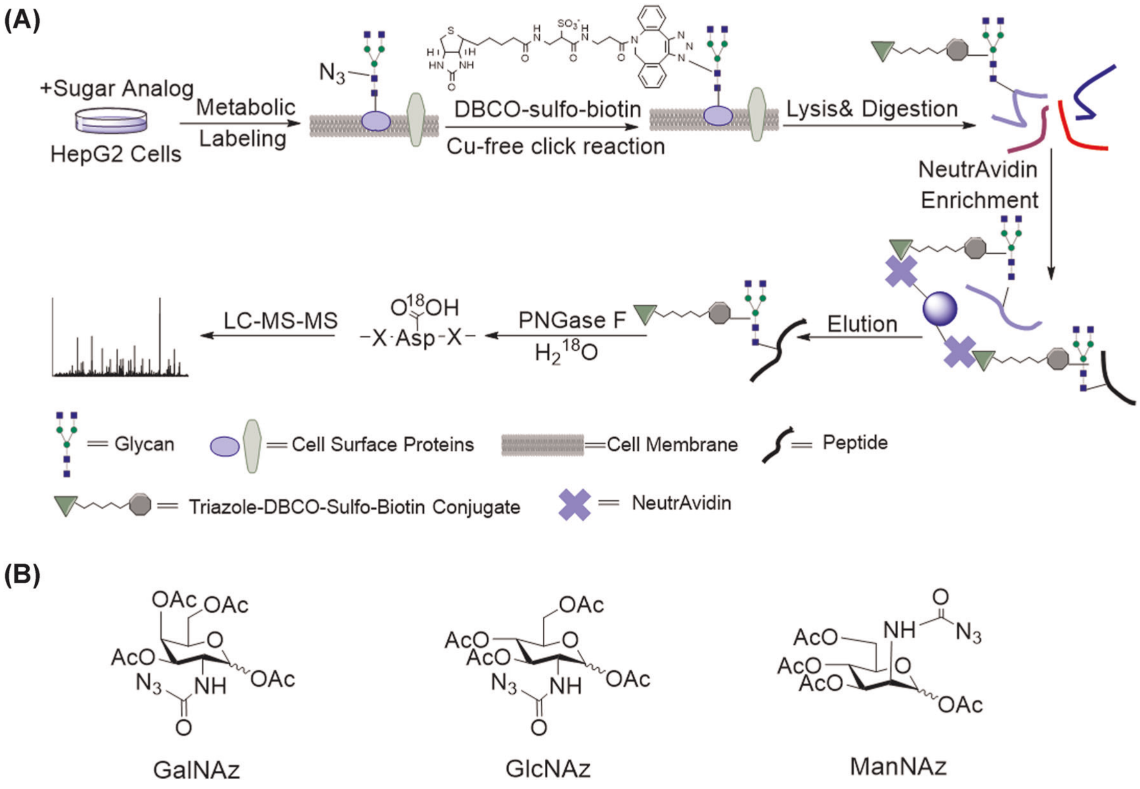

FIGURE 3 (A) Experimental procedure for comprehensively investigating surface glycoproteins by integrating metabolic labeling, bioorthogonal chemistry, and MS-based proteomics. (B) Structures of the three major sugar analogs containing the azide group. Adapted with permission from Xiao et al. (2016b). Copyright retained by American Chemical Society [Color figure can be viewed at wileyonlinelibrary.com]

cells. DBCO-sulfo-biotin is very hydrophilic, therefore it cannot penetrate the plasma membrane of cells and can only label surface glycoproteins. After cell lysis and protein digestion, NeutrAvidin beads were used to enrich the biotinylated glycopeptides, followed by site-specific analysis by MS. Enrichment at the peptide level minimizes the nonspecific binding compared with that at the protein level. Comparison of the three sugar analogs for metabolic labeling revealed that GalNAz resulted in the highest coverage of the surface *N*-glycoproteins (Xiao et al., 2016b).

Cell-surface glycoproteins play extremely important roles in the cellular immune response. Recently, coupling selective enrichment with multiplexed proteomics, we comprehensively and site-specifically quantified the dynamics of surface glycoproteins in monocytes and macrophages in response to lipopolysaccharide (LPS) (Suttapitugsakul et al., 2021b). The time-resolved quantification results unraveled the gradual or transient changes of cell-surface glycoprotein in monocytes and macrophages during the response, and the responses were different for both monocytes and macrophages. The further quantification of surface glycoproteins in the differentiation experiment from monocyte to macrophage revealed that the different responses were at least partly attributed to the priming of monocytes during the differentiation. Besides the well-documented glycoprotein changes in response to LPS, we also identified some new surface glycoproteins participating in the immune response such as APMAP, TSPAN3, and IGSF8. Furthermore, this study provides site-specific information regarding protein glycosylation changes during the LPS treatment, for instance, the site N229 of CSF2RA but not N195 and N223 in monocytes. Systematic investigation of the dynamics of surface glycoproteins results in a better understanding of glycoprotein functions and cellular immune responses.

Metabolic labeling is normally restricted to cultured cells and difficult to be applied to clinical samples. However, Spiciarich et al. (2017) cultured the sliced tissues in a medium containing ManNAz to metabolically label sialylated proteins on the cell surface. Surface glycoproteins on the tissues can then be identified through click chemistry-based biotinylation, enrichment, and MS analysis. The wide application of metabolic labeling to investigate surface glycoproteins still needs to address some limitations. For example, the labeling efficiency for surface glycoproteins in different cell lines and glycoproteins may not be the same, which would affect the quantification (Chang et al., 2009a).

2.2 | MS-based characterization of surface glycoprotein interactions

Cell surface glycoproteins often function through interacting with molecules in the extracellular matrix and proteins nearby, creating a highly interactive network on the plasma membrane (Martinez-Martin et al., 2016; Wright et al., 2010). The interaction network adapts to the ever-changing extracellular environment, and, therefore, aberrant glycoprotein interactions are closely correlated with various diseases including cancers (Pinho & Reis, 2015). Analysis of cell surface glycoprotein interactions will deepen our understanding of the collaborations among proteins located on the cell surface to regulate cellular events. With the development of MS-based proteomics and different methods to capture the interactions, identification of surface glycoprotein interactors was achieved (Bausch-Fluck et al., 2019; Paek et al., 2017). Although different strategies coupled with MS were developed for studying protein interactions, including affinity purification, chemical crosslinking, size-exclusion chromatography (SEC), and native MS (Heck, 2008; Smits & Vermeulen, 2016; Swietlik et al., 2020), it is still challenging to apply them to investigate the interactions only located on the cell surface. This section focuses on the method development for studying the interaction events on the cell surface especially those in the extracellular region.

2.2.1 | Ligand-based receptor capture (LRC)

LRC is first reported by Frei et al. (2012) to identify the glycosylated receptors on the cell surface for a specific ligand. This method is based on a trifunctional reagent TRICEPS consisting of an NHS ester for conjugating the ligand of interest, a trifluoroacetyl-protected hydrazine group for reacting with the aldehyde group on surface glycoproteins generated by the periodate oxidation, and a biotin group for the affinity enrichment. The ligand-conjugated TRICEPS was incubated with intact cells after the periodate oxidation, allowing for the ligand-specific recognition of glycosylated protein receptors. The covalent hydrazone bond formed between hydrazine and the aldehyde groups can stabilize the interactions between ligand and receptors. After cell lysis, protein digestion, and glycopeptide enrichment, the receptors can be identified through the glycosylated peptide abundance changes compared with the control group without the ligand. The LRC technology was then applied to discover glycosylated protein receptors on the cell surface of growth factors, therapeutic antibodies, and viruses. Later on, the Hill group described another type of trifunctional probe with an aldehyde-reactive aminooxy group, a sulfhydryl group for ligand conjugation via a disulfide bond, and a biotin group (Tremblay & Hill, 2017). They identified receptors through nonglycopeptides directly from on-bead digestion after the protein-level enrichment instead of glycopeptides in the TRICEPS method. With a similar quantification strategy and a modified trifunctional reagent called HATRIC, Sobotzki et al. (2018) reported a second generation of LRC (Figure 4). Acetone was used as the protection reagent for hydrazine to replace the original trifluoroacetyl-protection group, which increased the labeling efficiency of hydrazone formation. An azide group was used for the pull-down experiment instead of biotin, which tolerated more harsh washing conditions and thus decreased the contamination from the streptavidin beads. The hydrazone formation can take place at pH 7.4 by utilizing 5-methoxyanthranilic acid instead of aniline as the catalyst, making the ligand interact with surface receptors under more physiological conditions. Because of the increased sensitivity in the second generation of LRC, they can achieve the identification of surface receptors using only one million cells.

2.2.2 | Proximity labeling-based methods

Proximity-based biotinylation is frequently used to investigate the interactions of proteins including those located on the cell surface (Branon et al., 2018; Samavarchi-Tehrani et al., 2020). Enzymes such as ligases and peroxidases fused to specific proteins (baits) allow for the covalent biotinylation of proteins (preys) interacting with the baits directly or indirectly with the addition of enzyme substrates. With appropriate controls and quantitative proteomics, the proteins (preys) close to the baits can be identified. When the enzymes are conjugated to the extracellular domains of proteins located on the cell surface through genetic engineering or attached to antibodies that can specifically recognize surface proteins, proximity-based biotinylation methods including EMARS, SPPLAT, and PUP-IT are able to detect the extracellular surface protein interactions (Jiang et al., 2011; Liu et al., 2018; Rees et al., 2017). Recently, a photocatalyst-based proximity labeling strategy with carbenes as the intermediates was developed to decrease the labeling radius, displaying higher spatial labeling accuracy compared with that of classical proximity-based labeling methods (Geri et al., 2020). Although these proximity-based biotinylation approaches play critical roles in deciphering the surface protein interactions, they

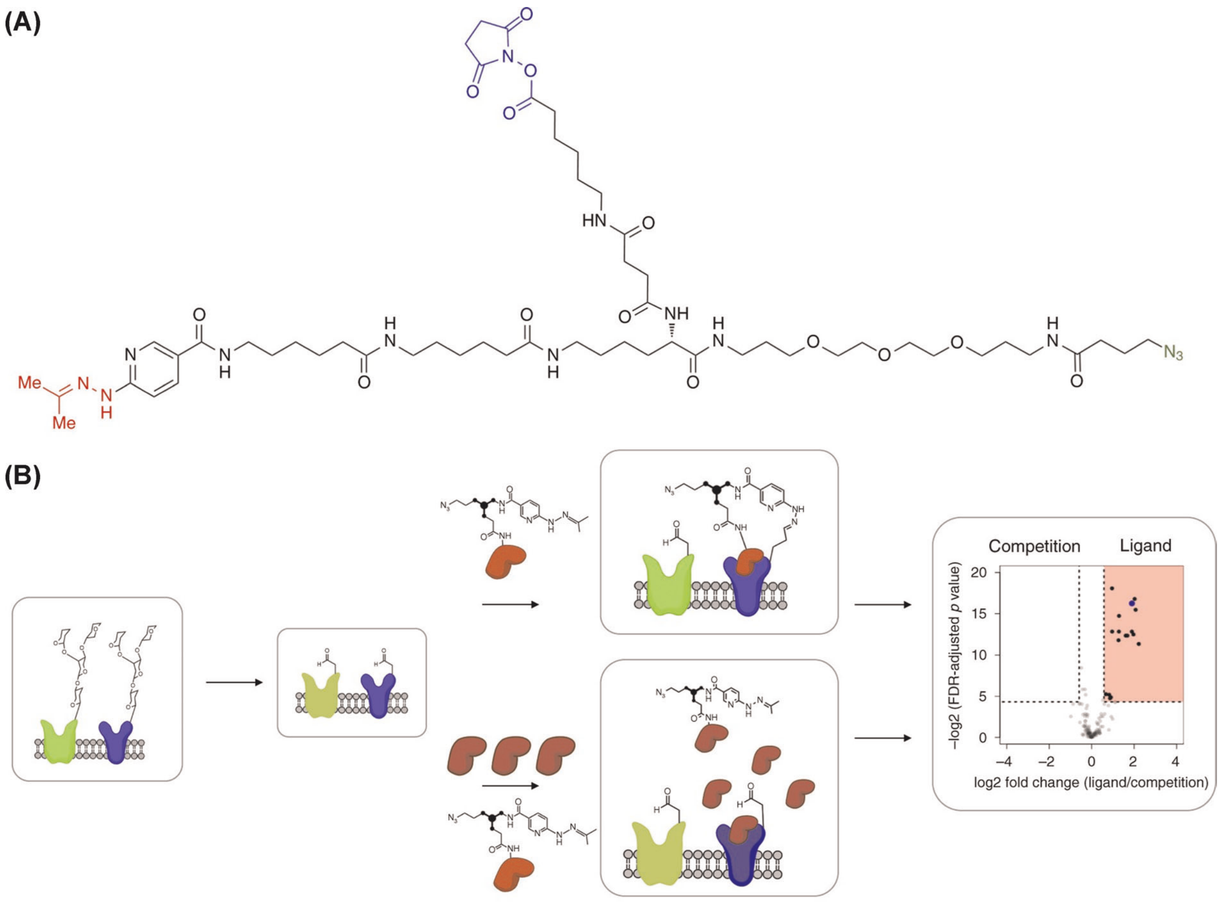

FIGURE 4 (A) Structure of the tri-functional compound HATRIC including protected hydrazide, azide, and *N*-hydroxysuccinimide ester. (B) Experimental procedure of identifying the surface receptors of one specific ligand using cell-surface capture and mass spectrometry-based proteomics. Reproduced with permission from Sobotzki et al. (2018). Copyright owned by the authors under a CC BY 4.0 license. Ligand of interest (maroon protein), surface receptor (blue protein), and other surface glycoproteins (green) [Color figure can be viewed at wileyonlinelibrary.com]

do not specifically target the glycoforms of proteins because antibodies or genetic fusion recognize all the forms of their targets.

A lectin that recognizes glycans on the cell surface can facilitate the investigation of the interactions between lectin and glycosylated ligands on the cell surface with proximity-based biotinylation. Recently, a comprehensive analysis of galectin-3 interactome in live human HSCs was achieved by fusing APEX2 to galectin-3 (PX-Gal3) that can bind to β-galactosides (Joeh et al., 2020). PX-Gal3 was used to label proteins with the addition of H_2O_2 and biotin-phenol, followed by cell lysis, enrichment, and on-bead digestion (Figure 5). With tandem mass tag (TMT)-based multiplexed quantitative proteomics, the authors selectively identified proteins that are highly enriched in the labeling group compared with the control one, allowing them to construct a data set for the potential interactors of galectin-3. The data set included various glycoproteins located on the cell surface such as vasorin, basigin, neuroplastin, CD9 in the tetraspanin family, EPHB1 in the ephrin family. The interactions between these proteins and galectin-3 were further verified by an in vitro binding assay. A similar strategy was also developed for the identification of Siglec ligands on the cell surface (Chang et al., 2017). Besides enzymes, chemical probes can also be conjugated to lectins. Xie et al. used iron (S)−1-(p-bromoacetamidobenzyl) EDTA as a catalyst to generate radicals in the presence of H_2O_2 for the oxidation of proteins that interact with a lectin (Xie et al., 2020). With this method, they investigated surface glycoproteins that potentially interact with different types of lectins such as wheat germ agglutinin (WGA), *Sambucus nigra* (SNA), and *Maackia amurensis* lectin, each of which can recognize specific glycan motifs.

2.2.3 | Photo-crosslinking strategies

A sugar analog with a photo-reactive functionality can be incorporated into glycans located on the cell surface through metabolic labeling (Wu & Kohler, 2019). After UV crosslinking, the glycan-protein interactions for a specific protein target can be pulled down through affinity enrichment such as immunoprecipitation, followed by protein digestion and MS identification. To remove nonspecific binding proteins, a control group without UV radiation is usually included. Different sugar analogs with a photoreactive group have been developed to investigate glycoprotein interactions on the cell surface. The Paulson group reported the use of a sialic acid analog containing an aryl azide at C9 (9AAzSia) to study the surface interactors of CD22, which binds specifically to the terminal *N*-acetylneuraminic acid α(2-6) galactose (NeuAc-α(2-6)-Gal) on glycoproteins and plays an important role in regulating B-cell receptor signaling (Han et al., 2005; Ramya et al., 2010). They found that CD22 formed homo-multimeric complexes by binding to nearby CD22 as *cis* ligands on the same B-cell, whereas the B-cell receptor IgM on another cell was a major *trans* ligand of CD22. Using a sialic acid analog containing a diazirine group (ManNDAz), the Kohler group systematically investigated the Cholera toxin B subunit (CTB)-binding glycoproteins on the cell surface (Figure 6) (Wands et al., 2015). After metabolic

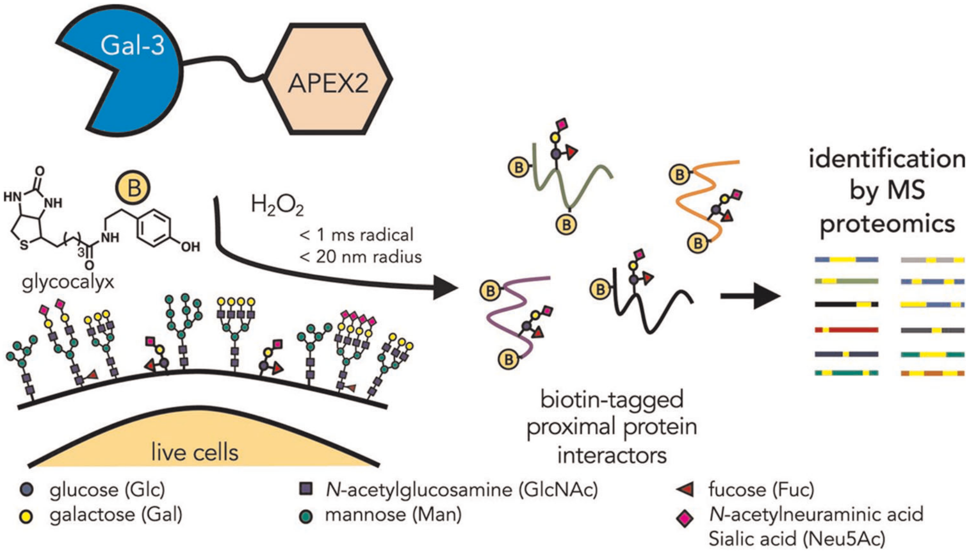

FIGURE 5 Identification of proteins that interact with galectin-3 (Gal-3) on live cells using proximity labeling and mass spectrometry (MS)-based proteomics. Reproduced with permission from Joeh et al. (2020). Copyright retained by the authors under CC BY-NC-ND [Color figure can be viewed at wileyonlinelibrary.com]

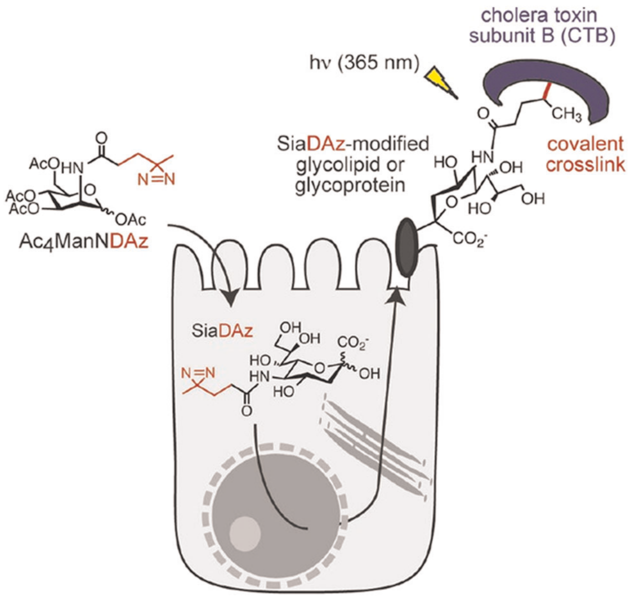

FIGURE 6 Identification of sialylated surface proteins that interact with cholera toxin B subunit (CTB) with metabolic labeling and UV crosslinking. Crosslinked complexes can be visualized by immunoblot or identified by liquid chromatography–mass spectrometry (MS)/MS analysis. Reproduced with permission from Wands et al. (2015). Copyright retained by the authors under a CC-BY license [Color figure can be viewed at wileyonlinelibrary.com]

labeling, a biotinylated CTB was incubated with the cells, followed by UV crosslinking, streptavidin enrichment, and MS analysis. CEACAM5 was found to be an interactor of CTB. This interaction was dependent on the glycan instead of the protein.

The photocrosslinking sugar analogs normally enable the identification of binding molecules for one protein of interest. Feng et al. (2013) synthesized a sialic acid analog with two functional groups, a diazine group and an azide group, to metabolically label surface glycans. The bifunctional sugar analog can achieve UV crosslinking and pull-down purification simultaneously, holding great potential to comprehensively identify sialic-acid binding proteins on the cell surface by MS. Recently, Li et al. (2019a) reported a method integrating metabolic and chemical proximity labeling to globally identify sialic acid-binding proteins on the cell surface. Instead of using a photocrosslinking sugar analog, ManNAz with an azide group was employed to label sialic acids on the cell surface, followed by conjugation of a catalyst probe through SPAAC that can oxidize nearby proteins under the H_2O_2 treatment. The oxidation sites can be identified and quantified using MS-based proteomics, allowing for global identifications of sialic-acid binding proteins on the cell surface. Integrating chemical crosslinking, GAO labeling, and TMT-based multiplexed quantitative proteomics, our group recently achieved comprehensive analysis of surface glycoprotein interactions (Sun et al., 2021). With the design of appropriate probes, it will be possible to identify the binding sites among the interactions involved with surface glycoproteins (Li et al., 2020; Udeshi et al., 2017).

2.3 | Applications of MS-based methods to study surface glycoproteins

The advancement of MS-based methods for investigating glycoproteins located on the cell surface enables systematic analysis of their properties and functional roles in different biological systems. In this section, we reviewed the applications of MS-based methods to study surface glycoproteins including protein dynamics and biomarker discovery.

2.3.1 | Surface glycoprotein dynamics

Glycoproteins located on the cell surface are highly dynamic to adapt to the ever-changing environment. Glycosylation is reversible with enzymes responsible for the addition and removal of glycans. Glycans are of great importance for proteins to maintain their stability and proper functions on the cell surface. However, comprehensive analysis of surface protein dynamics is still challenging. Schiess et al. (2009) used CSC and label-free quantification to study the dynamic distribution of surface glycoproteins, such as internalization without degradation and intracellular redistribution in response to intracellular signaling network changes. A similar strategy was employed to determine surface glycoprotein dynamics during monocyte to macrophage differentiation of THP-1 cells, which revealed the dynamic remodeling of the surface proteome during differentiation, including rapid translocation of surface proteins, quick internalization of unneeded proteins, and delayed translocation of newly synthesized macrophage markers (Kalxdorf et al., 2017). Recently, the Wollscheid group achieved a systematic investigation of surface glycoprotein dynamics of neurons with the newly developed autoCSC method (van Oostrum et al., 2020). To study the dynamics of surface glycoproteins during neuronal differentiation, they performed in vitro primary neuronal cultures and tagged the surface glycoproteins using the periodate oxidation every 2 days from 2 to 20 days. After sample preparation using autoCSC, glycopeptides were analyzed by DIA-MS. Over 1700 glycosylation sites on over 1000 proteins were identified and quantified in the time-resolved experiment. They found different clusters of surface glycoprotein dynamics that are consistent with the developmental stage-specific patterns. Compared with the whole proteome, the surface glycoproteome displayed a higher level of dynamics in a limited time scale and widespread regulation on surface trafficking during development and homeostatic synaptic scaling.

Recently, our group integrated metabolic labeling and TMT-based multiplexed proteomics to study surface glycoprotein dynamics (Xiao & Wu, 2017). The cells were first cultured in a medium containing GalNAz to metabolically label surface glycoproteins. After biotinylation with copper-free click chemistry, the cells were changed to another medium without the sugar analog and harvested at different time points (Figure 7). The half-lives of surface glycoproteins were determined by the abundance changes of glycoproteins in the time-resolved experiment. We found that glycoproteins with catalytic activities were more stable than those with binding and receptor activities. Combining metabolic labeling with label-free quantitative MS, Herber et al. studied the dynamics of neuronal surface glycoproteins in cells with BACE1 inhibition. BACE1 is an important protease that cleaves the membrane protein APP to generate amyloid β peptide (Herber et al., 2018). They found that several substrates of BACE1 such as APP, APLP1, and SEZ6 were upregulated, and these protein changes on the cell surface were not strongly correlated with their changes in the secretome.

2.3.2 | Biomarker discovery on the cell surface

The presence of specific surface glycoproteins enables the targeted selection of cell population from plenty of cells with various identities (Meyfour et al., 2021). Aberrant glycosylation has been reported to be closely correlated with multiple diseases such as cancers (Schjoldager et al., 2020). In this section, we focused on the study of surface glycoproteins used as biomarkers for the classification of cell types and disease detection.

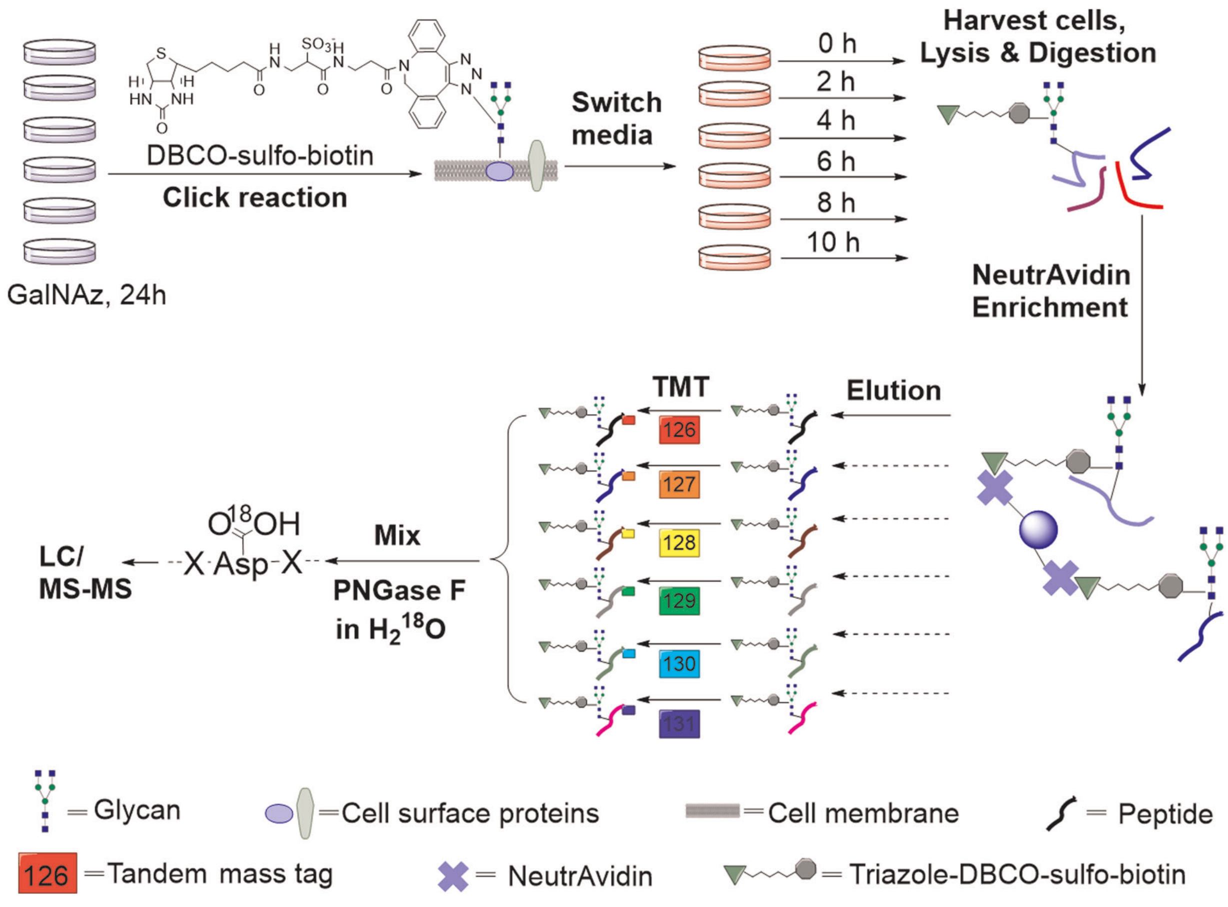

FIGURE 7 Experimental procedure for the investigation of surface glycoprotein dynamics by integrating pulse-chase labeling, metabolic labeling, and tandem mass tag (TMT)-based multiplexed proteomics. Reproduced with permission from Xiao & Wu (2017a). Copyright retained by the authors under a CC-BY license. DBCO, dibenzocyclooctyne; LC, liquid chromatography; MS, mass spectrometry [Color figure can be viewed at wileyonlinelibrary.com]

A comprehensive label-free quantification of CSPA consisting of identified surface glycoproteins from 41 human and 31 mouse cell types revealed that the cell-type-specific differences mainly come from quantitative differences of surface glycoproteins, rather than from protein identities such as CD proteins frequently used as cell-type classifications (Bausch-Fluck et al., 2015). Boheler et al. (2014) systematically analyzed the surface glycoproteomes of human pluripotent stem cells (hPSCs), human embryonic stem cells (hESCs), and human fibroblasts to investigate specific pluripotency markers, and found the unique expression of eight surface glycoproteins (ADCYAP1R1, EFNA3, FAM216A, FGFR3, HTR2C, IL17RD, NPR1, and OPCML) on hPSC that can be used as biomarkers, which were further validated by immunostaining. The cell surface markers for hPSC derivatives were also studied. Mallanna et al. (2017) profiled surface glycoproteins of hPSC-derived hepatic endoderm cells. By comparing the data with CSPA, 15 proteins including ABCC2, SLC2A8, and TMUB1 were overlapped between hPSC-derived hepatic endoderm and primary human hepatocytes. Using CSC, Haverland et al. (2017) assessed the surface glycoproteomes of four human lymphocyte cell lines. After label-free quantification and bioinformatic analysis, the proteins specific for each cell type were revealed. By combining CSC and TMT-based multiplexed proteomics, Ravenhill et al. (2019) compared the expression difference of the surface proteomes on red blood cells (RBCs) from two genetically distinct populations (United Kingdom and Senegal), and identified potential Plasmodium receptors. They also used the same approach to investigate the surface markers of classical, intermediate, and nonclassical monocytes (Ravenhill et al., 2020). Our group used metabolic labeling strategy and label-free quantification to analyze the surface glycoproteins from eight types of human cells and further emphasized the importance of unique and differential expression of surface glycoproteins for the cell type determination (Suttapitugsakul et al., 2019).

CSC was applied to identify surface glycoproteins from five different thyroid cancer cell lines (Arcinas et al., 2009). Besides the known biomarkers for thyroid cancer, such as CD44, galectin 3, and metalloproteinase inhibitor 1, some surface glycoproteins were found uniquely in specific thyroid cancer cell lines, which provided a set of glycoprotein biomarker candidates. Mirkowska et al. (2013) investigated the surface glycoproteome from leukemia xenografts derived from 19 B-cell precursor ALL (BCP-ALL) patients using CSC. After data filtering based on gene expression information from normal hematopoietic cells, they used flow cytometry to verify nine candidate markers (CD18, CD63, CD31, CD97, CD102, CD157, CD217, CD305, and CD317) in patient samples. CD97, CD157, CD63, and CD305 displayed the most significant difference, which may be considered as the biomarkers for leukemia. Oldham et al. (2020) combined CSC and a targeted MS assay to detect the significantly up-regulated surface glycoproteins in multiple myeloma cells compared with healthy controls, followed by flow cytometry validation. Recently, Leung et al. (2020) quantified the surface proteome in isogenic breast epithelial cell lines transformed with one of the oncogenes, including receptor tyrosine kinases (EGFR and HER2) and the corresponding signaling partners (KRAS, BRAF, MEK, and AKT). The surface proteome remodeling for each oncogene was correlated with different biological processes such as alteration of glycosylation and carbohydrate metabolism, and down-regulation of differentiation and adhesion. Some example surface glycoproteins as potential biomarkers are listed in Table 2.

With metabolic labeling using ManNAz to label sialoglycoproteins, our group compared the expression of surface *N*-sialoglycoproteins between MCF7 and MDA-MB-231 breast cancer cells with different levels of invasiveness, and revealed that surface protein sialylation was closely correlated with cell migration and metastasis. Using a similar strategy, Pan et al. (2012) unraveled the significantly increased expression of the neuronal cell adhesion molecule in non-small-cell lung adenocarcinoma compared with nearby noncancerous tissues, which could potentially be served as a novel target and biomarker for treatment and diagnosis of this cancer. With the development of methods for surface glycoprotein profiling and MS-based proteomics, novel biomarkers and drug targets will be able to be discovered for different diseases.

3 | GLOBAL ANALYSIS OF SECRETED GLYCOPROTEINS USING MS-BASED PROTEOMICS

Cells secrete proteins into the extracellular environment, and most of them are glycosylated through the classical secretory pathway (Arigoni-Affolter et al., 2019; Seiradake et al., 2015; Yagi et al., 2020). These glycoproteins participate in many extracellular events including cellular signaling and cell–cell communication. Combining the results from antibody detection with those from MS analysis, Uhlen et al. (2019) determined the data set of human secreted proteins based on protein characteristics such as the presence of signal peptides or the transmembrane region, and showed that the number of potential genes encoding secreted proteins was 2641, which are about 13% of all human protein-coding genes. Secreted glycoproteins in bodily fluids can serve as excellent noninvasive sources of biomarkers for disease detection (Ideo et al., 2020; Kirwan et al., 2015; Liu et al., 2010; Zhao

TABLE 2 Selected examples of biomarker discovery using mass spectrometry-based glycoproteomics

Cell type	Method	Selected potential biomarkers	Purpose	Reference
Human pluripotent stem cells	Cell surface capture (CSC)	ADCYAP1R1, EFNA3, FAM216A, FGFR3, HTR2C, IL17RD, NPR1, and OPCML	Cell type determination	Boheler et al. (2014)
Red blood cells	CSC	CD44, CD55, CD59, CD151, and SLC44A2	Cell type determination	Ravenhill et al. (2019)
Thyroid cancer cell lines	CSC	CD44, galectin 3, and metalloproteinase inhibitor 1	Disease detection	Arcinas et al. (2009)
Non-small-cell lung adenocarcinoma	Metabolic labeling	Neuronal cell adhesion molecule	Disease detection	Pan et al. (2012)

et al., 2007; Zhu et al., 2020). Particularly, glycosylated proteins have enhanced stability, increasing the possibility of their detection (Boersema et al., 2013). Several molecular biology approaches, especially antibody-based methods, have been employed to study secreted glycoproteins (Frenette et al., 1988; Gal et al., 1985). Nevertheless, these methods do not allow for large-scale analysis of secreted glycoproteins nor reveal the information of the glycosylation site or the glycans. MS-based proteomics is powerful for secreted glycoprotein analysis. However, there are still some challenges that must be overcome to effectively identify and quantify secreted glycoproteins on a large scale. In this section, we review approaches to study secreted glycoproteins from different biological samples using MS and their applications in the biological and biomedical fields.

3.1 | Extensive and multidimensional fractionation

One major obstacle in the global analysis of secreted glycoprotein is the extremely high dynamic range of proteins present in the culture medium, blood, or plasma (Hortin & Sviridov, 2010; Pare et al., 2016). In serum, the concentration of albumin could be in the tens of grams per liter (Weaving et al., 2016), whereas others may be present in the micrograms per liter or less (Merrell et al., 2004). Because MS is biased for abundant proteins/peptides, especially with the popular data-dependent acquisition approach, those highly abundant proteins such as albumin are typically dominant in secreted protein analysis. Therefore, many important proteins including those released from diseased cells normally with a very low abundance are extraordinarily challenging to be detected using MS-based proteomics (Kim et al., 2011). The simplest method to decrease the complexity of proteins/peptides is to perform extensive and multidimensional fractionation. After protein digestion, peptides are fractionated through orthogonal separation methods before MS analysis. The fractionation decreases the complexity and improves the chance for the detection of low-abundance peptides by MS. Though extensive fractionation greatly improves the coverage of secreted proteins, the analysis is time-consuming, as seen by the long analysis time from the study by Dey et al. (2019) (>500 h) to cover ~5000 proteins.

3.2 | Depletion of highly abundant proteins

As high-abundance proteins can affect the coverage of detected proteins during MS analysis, another approach to reduce the complexity is through depletion columns. Several types of depletion columns are commercially available. These columns are based on antibodies that can capture the proteins with high abundance in samples. In one example, CaptureSelect HumanPlasma 14 affinity resin was used to remove proteins including albumin, immunoglobulins, transferrin, and others from plasma samples before multilectin affinity chromatography for glycoprotein enrichment (Totten et al., 2018). The authors identified glycoproteins that are differently secreted from prostate cancer and benign prostatic hyperplasia patients. Similarly, Berven et al. (2010) combined protein depletion with solid phase extraction of N-linked glycoproteins (SPEG), which is based on a reaction between oxidized glycans and hydrazine resins, to enrich glycoproteins from human plasma. Compared with the experiment without the depletion, the number of identified glycoproteins increased by 24%.

Though depletion columns can effectively remove high-abundance proteins from serum samples, studies have shown that these columns may also result in sample loss, which affects the quantification of proteins and the reproducibility (Jankovska et al., 2019; Tu et al., 2010). One of the proteins with the highest abundance is albumin, which can bind to several secreted proteins (Merlot et al., 2014). Therefore, the removal of albumin caused the loss of other secreted proteins (Granger et al., 2005). The capacity of these columns is normally low, and thus the approach is not very cost-effective. Additionally, nonspecific binding of proteins to the materials used in these columns may occur (Polaskova et al., 2010). de Jesus et al. (2017) compared three depletion methods to remove the most abundant proteins from cells, including magnetic nanoparticles, sequential application of dithiothreitol and acetonitrile, and the commercial ProteoMiner apparatus based on immunoaffinity. The results revealed that the magnetic nanoparticle depletion was more versatile, reproducible, and removed proteins with high molecular weights (>80 kDa), resulting in the highest number of identified proteins. For the chemical method, though the depletion of proteins was also observed, the number of identified proteins was lower. The ProteoMiner apparatus that was not cost-effective showed similar results to the chemical method.

3.3 | Serum starvation

In cell culture models, serum such as fetal bovine serum is often used in the growth media as it contains proteins and other small molecules required for in vitro cell growth (Zheng et al., 2006). The concentration of proteins, especially albumin, is very high in the final culture media, which affects the detection of low-abundance, secreted glycoproteins. One solution to this problem is to grow cells in a serum-free medium for a short period of time (Chang et al., 2009b; Stastna et al., 2012). Before switching to the serum-free medium, the cells must also be washed thoroughly to remove residual serum proteins. Rogers et al. (2014) showed that three washes are optimal to remove serum proteins, whereas more washes can decrease cellular viability. Additionally, washing cells with serum-free medium leads to the contamination of proteins from lysed cells.

One major problem with this approach is the effects of cellular starvation on cell growth and proliferation, and thus the alteration of protein secretion from cells. Cell death may also occur, resulting in the leakage of intracellular proteins into the extracellular space. This complicates and affects secreted glycoproteomic analysis. One study focused on enriching newly synthesized, secreted proteins from cells by combining azidohomoalanine (AHA) labeling with pulsed SILAC (Eichelbaum et al., 2012). AHA is a methionine analog that can be conjugated by aminoacyl tRNA (transfer RNA) synthetase during protein synthesis instead of methionine. Proteins labeled with AHA containing the azide group can be selectively enriched through click chemistry. Over 1000 secreted proteins were identified and quantified including some extracellular matrix proteins and cytokines. The authors further applied this method for a time-resolved analysis of secreted proteins from mouse macrophages (RAW264.7) treated with LPS after 0, 6, and 17 h. Several inflammatory cytokines were identified and quantified within a 2-h labeling period at each time point. The authors also showed that the secretion of proteins could be affected by incubating cells under serum-free conditions only after 3 h, raising the concern that the results from secreted protein analysis under such conditions should be further examined and validated by other methods.

3.4 | Lectin-based enrichment

Lectins are proteins that bind specifically to glycans and have been extensively used for glycoproteome analysis (Abbott & Pierce, 2010; Chen et al., 2014a; Feng et al., 2009; Mechref et al., 2008; Wei et al., 2010). They were also employed to enrich secreted glycoproteins in the extracellular compartment (Jung et al., 2009; Plavina et al., 2007; Qiu & Regnier, 2005; Yang & Hancock, 2004; Zhang et al., 2012). In a study by Boersema et al. (2013), the authors profiled secreted glycoproteins from 11 breast cancer cell lines in different stages (Figure 8). Two lectins, concanavalin A (Con A) and WGA, were combined with filter-aided

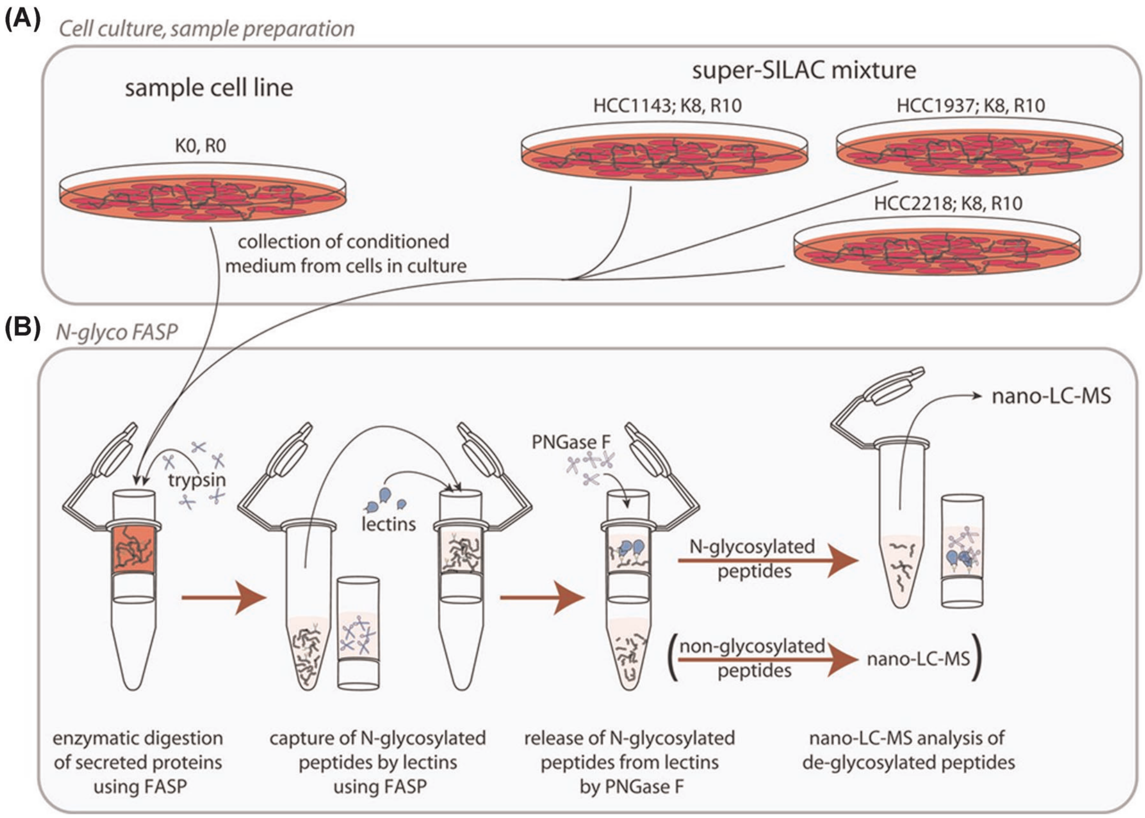

FIGURE 8 Workflow combining filter-aided sample preparation (FASP) with lectins to enrich secreted *N*-glycoproteins from cell culture media. Reproduced from Boersema et al. (2013). Copyright retained by the American Society for Biochemistry and Molecular Biology under the CC-BY license [Color figure can be viewed at wileyonlinelibrary.com]

sample preparation (FASP) to enrich secreted *N*-glycoproteins under the serum-free conditions. They identified and quantified peptides with 1398 *N*-glycosylation sites from 701 proteins. The results also demonstrated differential protein secretions from cancer cells in different stages. Though the authors showed that the serum-free conditions did not affect the viability of the cells, the biological interpretation of the quantitation results from cells under such conditions might need to be carefully considered. Due to their glycan specificity, multiple lectins are often required to capture glycoproteins with diverse glycans in the secretome. Li et al. (2013) and colleagues developed a method termed lectin affinity capture followed by solid-phase extraction of glycosite-containing peptides (LecSPEG). *Aleuria aurantia* lectin-, SNA-lectin, and WGA-conjugated beads were used to enrich glycoproteins from serum samples. The eluted proteins were digested, and the resulting glycopeptides were further oxidized with sodium periodate for their enrichment with hydrazide beads. With this approach, the number of unique glycosites increased from two with lectin enrichment alone to 46 with the LecSPEG method, and the specificity increased from 2% to 82%.

3.5 | HILIC

In HILIC, hydrophilic peptides, including those with glycans, are retained in the column and eluted by an increasing amount of water in the mobile phase (Alpert, 2008). HILIC has been reported for the study of secreted glycoproteins in various samples (Shu et al., 2020). ZIC-HILIC, where the stationary phase contains both positive and negative charges, was also employed for secreted glycoprotein analysis (Alagesan et al., 2017; Cao et al., 2016). Other work has combined different separation modes to increase the specificity for glycopeptide enrichment (Adav et al., 2015). In the work from R. Chen et al. (2019), phosphopeptides and glycopeptides were simultaneously enriched through a method called mode switchable solid phase extraction (MS-SPE). A solid-phase extraction column packed with PolyWAX was first used to separate both phosphopeptides and glycopeptides in the electrostatic repulsion hydrophilic interaction chromatography (ERLIC) mode. The column was then switched to the HILIC mode by using a high concentration of organic solvent where phosphopeptides were eluted and glycopeptides were retained. Eventually, glycopeptides were eluted with water. With this approach, 136 phosphorylation sites and 283 *N*-glycosylation sites were identified from the secretome of A549 cells under serum-free conditions. Nevertheless, nonglycosylated peptides may also be hydrophilic and retained in these columns, which affects the enrichment specificity.

3.6 | Boronic acid-based enrichment

Boronic acid forms reversible covalent bonds with the diol groups on glycans, allowing for the enrichment of glycoproteins in biological samples (Chen et al., 2014b; Liu & He, 2017; Liu et al., 2012; Qu et al., 2012; Xiao & Wu, 2017b; Xiao et al., 2018a; Xiao et al., 2018b; Xiao et al., 2016a; Zhang et al., 2009). Boronic acid was compared with lectin-based enrichment (ConA, *Lens culinaris* agglutinin, or WGA) for glycoprotein analysis in human serum samples, and it resulted in the highest number of identified glycoproteins possibly due to the wider array of glycans enriched (Sparbier et al., 2007). For *Saccharomyces cerevisiae*, boronic acid-conjugated magnetic beads were employed to enrich secreted glycopeptides from the growth medium (Smeekens et al., 2017). The analysis requires the careful medium collection step to avoid cytosolic protein leakage, which can interfere with the detection of secreted glycoproteins. In this study, the cells were also treated with tunicamycin to inhibit protein *N*-glycosylation. Many secreted *N*-glycoproteins were downregulated compared with the control group, showing the effects of *N*-glycosylation on protein secretion.

3.7 | Hydrazide chemistry-based enrichment

Hydrazide chemistry-based methods have also been employed to enrich secreted glycoproteins (Goo et al., 2009; Wang et al., 2012). In one example, combining MS analysis with hydrazide chemistry and ZIC-HILIC, Li et al. (2013) identified 1213 unique N-glycosylation sites from 611 glycoproteins from two cell lines with low or high metastatic potential. Differential regulation of several secreted glycoproteins was discovered based on the metastasis capability and could be used as biomarkers. Note that this was done under serum-free conditions.

Whereas the DDA method is biased for abundant peptides, DIA records all the peptides by spacing the measurement length into smaller sections. This minimizes the bias of MS in detecting low-abundance secreted glycoproteins. In one study, Sajic et al. (2018) employed Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS) to detect secreted glycoproteins in the blood from several carcinomas. Glycoproteins were enriched from the blood of patients with localized lung, pancreas, ovary, prostate, and colorectal cancer, or their control matches using hydrazine resins and subsequently analyzed by MS. Over 1000 glycopeptides from >200 glycoproteins were quantified. The author identified proteins that are generally altered in patients with cancer and other glycoproteins that are unique and specific for each cancer type.

3.8 | Metabolic labeling-based enrichment

Metabolic labeling with a sugar analog is powerful to analyze glycoproteins on the cell surface, as discussed above. It has also been used in secreted glycoprotein analysis. For instance, Roper et al. (2013) employed ManNAz to label secreted glycoproteins from two stromal cell lines. The azide-labeled glycoproteins are then enriched with alkyne beads through the CuAAC reaction. From the cells being serum-starved for 48 h, 75 secreted glycoproteins were detected, whereas in 1% serum-containing media, 100 secreted glycoproteins were detected. Kuhn et al. (2012) developed the secretome protein enrichment with click sugars (SPECS) method for shed and secreted protein analysis of primary cells in the presence of serum proteins. In this study, glycoproteins were first labeled using ManNAz. The media were then collected, and the glycoproteins were tagged with DBCO-PEG12-biotin before streptavidin enrichment and MS analysis. Later, this method was modified in the high-performance SPEC (hiSPECS), which decreased the number of cells required for secretome studies (Figure 9) (Tushaus et al., 2020). The authors also labeled glycoproteins with ManNAz and enriched secreted glycoproteins using ConA and DBCO beads. Secreted glycoproteins from different systems were identified, including those in brain slices upon LPS-induced neuroinflammation and primary astrocytes, microglia, neurons, and oligodendrocytes in a cell-specific manner. The results need to be carefully considered as nonspecific binding may occur. Additionally, the incorporation efficiency of these sugar analogs in different cell types must be considered.

3.9 | Signal boosting approach

The development of the TMT reagents allowed for the simultaneous identification and quantification of proteins in multiple samples. With this approach, peptides from different samples are labeled with different

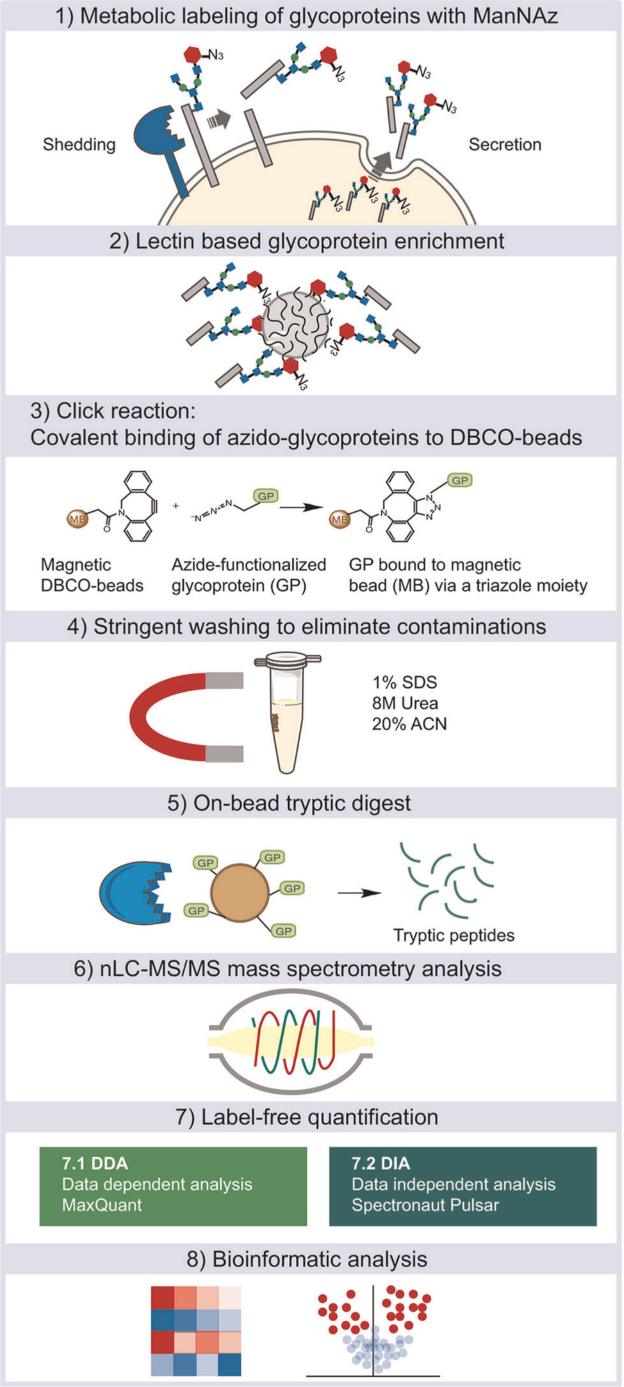

FIGURE 9 hiSPEC (high-performance secretome protein enrichment with click sugars) workflow for shed glycoprotein analysis. The cells are first labeled with ManNAz, and the resulting glycoproteins are then enriched with lectins and tagged through the click chemistry reaction. Eventually, the glycopeptides are analyzed with liquid chromatography–mass spectrometry (LC-MS)/MS. Reproduced from Tushaus et al. (2020). Copyright retained by authors under the CC BY 4.0 license. ManNAz (red hexagon), and magnetic dibenzocyclooctyne (DBCO)-beads (gray sphere) [Color figure can be viewed at wileyonlinelibrary.com]

channels of the TMT reagents. The labeled peptides are then combined and analyzed with MS. The peptides from different samples appear as a single peak in MS1 spectra due to the same m/z. Once they are selected for MS2 analysis, the intensities of the reporter ions can be used for accurate quantification of peptides in different samples. This multiplexed approach improves the reproducibility and shortens the analysis time. Furthermore, recent studies showed that the TMT labeling can improve the peptide signal intensity in MS1, facilitating the selection of low-abundance peptides for MS2 analysis. Notably, Budnik et al. (2018) employed this method for single-cell proteomics analysis (Single Cell ProtEomics by Mass Spectrometry [SCoPE-MS]). Another approach termed Boosting to Amplify Signal with Isobaric Labeling (BASIL) employed a similar approach for protein phosphorylation analysis (Yi et al., 2019). In these approaches, a protein carrier sample, for example, the greater number of cells in SCoPE-MS or the combination of proteins from the quantitation channels in BASIL, with a higher protein abundance compared with the quantitation samples was employed and labeled with one of the TMT channels (typically the last channel to avoid overlapping of the isotopic envelope).

Using the boosting approach, secreted proteins in clinical samples were studied (Russell et al., 2017). Our lab recently combined the boosting approach with metabolic labeling using GalNAz to label secreted glycoproteins from cultured cells without serum starvation (Figure 10) (Suttapitugsakul et al., 2021a). We compared different glycoprotein sources used for the boosting channel and evaluated the effect of the boosting-to-sample ratios. With this approach, over 200 glycoproteins were detected in a site-specific manner from different cell types, including monocytes, macrophages, and Hep G2 cells under serum-containing conditions.

3.10 | Methods to study glycoproteins from extracellular vesicles

Another group of proteins is secreted through vesicle trafficking (Doyle & Wang, 2019). Extracellular vesicles can be classified into different groups based on their size and properties, including microvesicles, exosomes, and apoptotic bodies (van der Pol et al., 2012). A recent study showed that extracellular vesicles mediated the communication between cells through their lipid, protein, or

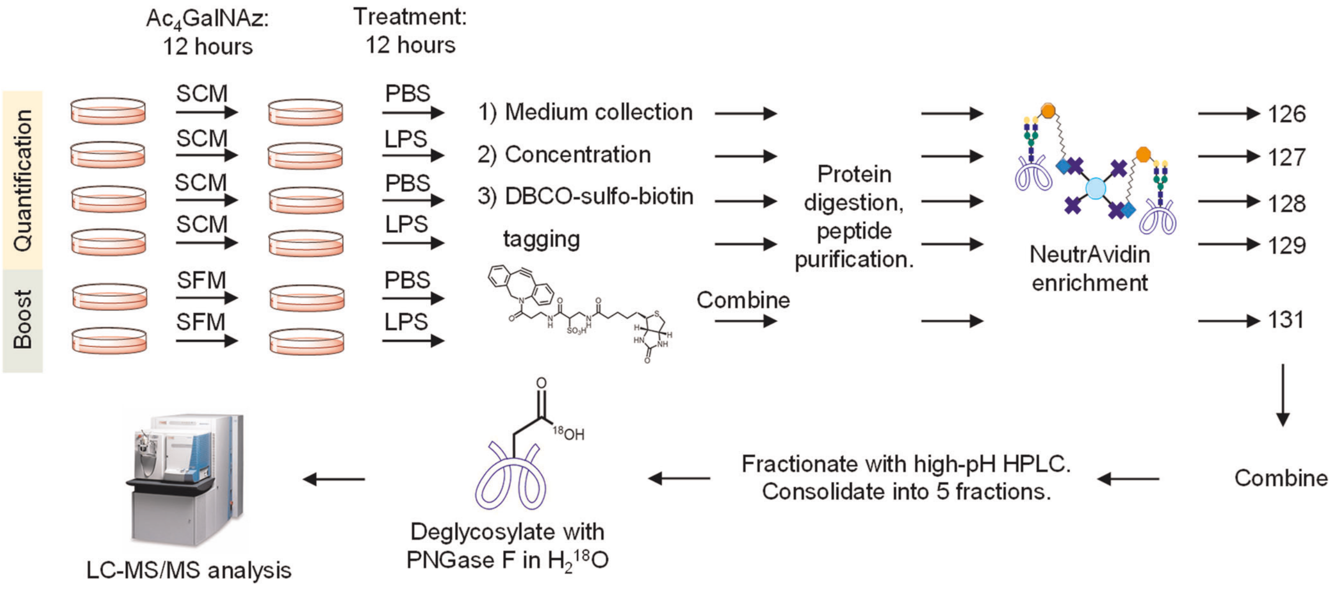

FIGURE 10 Workflow to identify and quantify secreted glycoproteins from cultured cells without serum starvation. Reprinted with permission from Suttapitugsakul et al. (2021a). Copyright retained by American Chemical Society. HPLC, high-performance liquid chromatography; LC, liquid chromatography; LPS, lipopolysaccharide; MS, mass spectrometry; PBS, phosphate-buffered saline; SCM, serum-containing medium; SFM, serum-free medium [Color figure can be viewed at wileyonlinelibrary.com]

nucleic acid contents (Maacha et al., 2019). These vesicles are specific and can serve as a molecular fingerprint for the cell state (Dickhout & Koenen, 2018; Lim et al., 2020). Glycoproteins were also found in extracellular vesicles, and a review on glycosylation in extracellular vesicles was published previously (Williams et al., 2018).

Typically, methods to study glycoproteins in extracellular vesicles are coupled with methods to separate these vesicles, which are reviewed extensively in the literature (Carnino et al., 2019; Konoshenko et al., 2018). Although several methods are available for their isolation, a study demonstrated that different methods could lead to the different glycoprotein profiles (Freitas et al., 2019). Similar protein and glycan profiles were found by ultracentrifugation, OptiPrep density gradient, and SEC, with the last two also enhancing the glycoprotein detection. On the contrary, total exosome isolation, which refers to a technique from Invitrogen that separates less soluble components including exosomes from samples, resulting in a different and unique population. The authors also demonstrated the importance of the serum used for cell growth, which can affect the quality of the extracellular vesicles obtained because some proteins such as albumin can bind to these vesicles.

Exosomes are one type of extracellular vesicles that can be secreted through the formation of multivesicular endosomes within cells. Exosomes are also defined by their size to be less than 150 nm in diameter compared with other larger vesicles (Hessvik & Llorente, 2018). Several biological species have been reported to present in these vesicles, including proteins and glycoproteins, lipids, and nucleic acids (Mashouri et al., 2019). Although initially thought to function in the removal of proteins from cells, exosomes were found to be involved in the cell–cell communication and cellular signaling (de Toro et al., 2015) such as their roles in T-cell mediated immune responses (Anel et al., 2019), thus attracting further studies on their roles, functions, biogenesis, and composition.

Generally, exosomes are separated by ultracentrifugation approaches. Yet, the volume of biological fluids required may be high. The contamination from other proteins or other types of extracellular vesicles could also exist. Alternatively, immuno-capture of exosomes can be performed as they are formed by the multivesicular body formation in the presence of CD9 and CD63 (Akers et al., 2013). Exosomes may contain a high abundance of glycoproteins such as MUC1 compared with the parent cells (Pan et al., 2019). Commonly used methods for exosome protein analysis include ultracentrifugation, antibody pull-down, click chemistry-based methods (Bu et al., 2019; Patel et al., 2019).

Lectin has been used to enrich glycoproteins/glycopeptides from exosomes. In urine, lectin affinity chromatography was employed to identify *N*-glycopeptides from the exosome, in which the sample collection is easy and could reveal the physiological states of the parent epithelial cells lining the urinary tract (Figure 11) (Saraswat et al., 2015). The authors found 126 *N*-glycopeptides from 37 glycoproteins, including 66 *N*-glycans and 13 sulfated or phosphorylated glycans. In ovarian carcinoma SKOV3 cells, lectin was used to

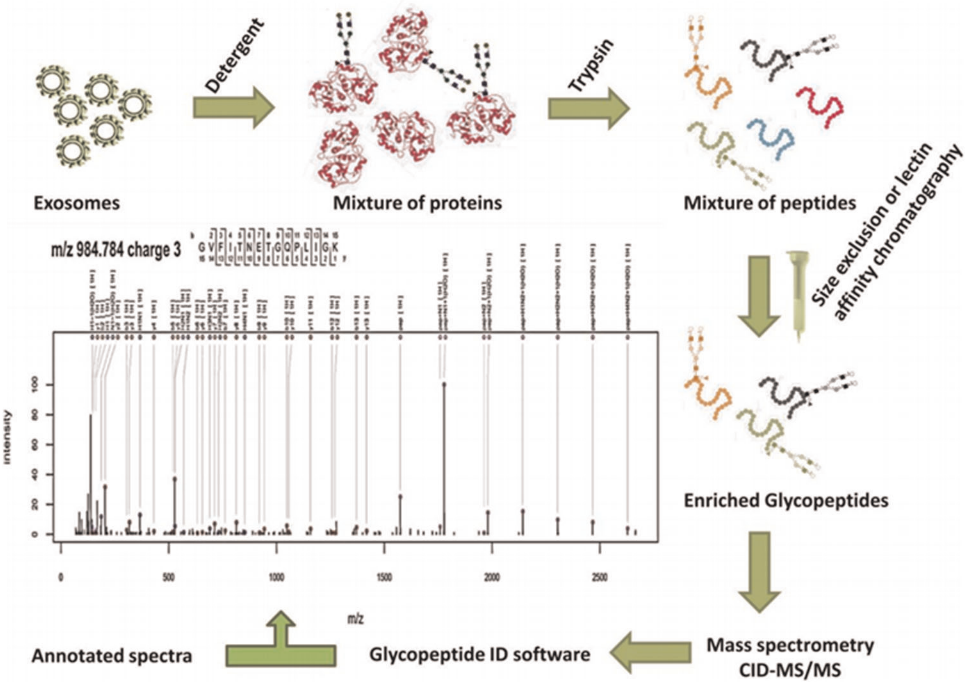

FIGURE 11 The method implementing size exclusion or lectin affinity chromatography to study urinary exosome. Reproduced from Saraswat et al. (2015). Copyright retained by the American Society for Biochemistry and Molecular Biology [Color figure can be viewed at wileyonlinelibrary.com]

determine the glycan composition of the exosomes, where a high amount of sialylation was found (Escrevente et al., 2013). The authors also found that the removal of the sialic acid residues can cause a small but nonsignificant uptake of the exosomes by cells (Escrevente et al., 2011). Chauhan et al. (2017) employed the surface glycan oxidation with sodium periodate and tagging with the aminooxy-biotin method to target surface glycoproteins on myeloid-derived suppressor cells. The authors identified 21 surface glycoproteins on the exosomes and showed that the majority of glycoproteins (>80%) on the surface of the exosomes resembled surface glycoproteins of cells. These glycoproteins facilitated the chemotaxis and migration of the cells. HILIC-based methods have also been employed to study glycoproteins from the exosomes. In a study from the Qian lab, molybdenum disulfide (MoS_2)/gold nanoparticle (Au-NP)-L-cysteine nanocomposite was developed for glycopeptide enrichment (Xia et al., 2018). Compared with ZIC-HILIC, the novel material was effective and they identified 1920 glycopeptides from 775 glycoproteins using 50 μg proteins from HeLa exosomes. Later, the same group developed a magnetic hydrophilic material with MoS_2-Fe_3O_4 nanoparticle-gold nanowires (Au/NWs)-glutathione and identified 1250 glycopeptides from human urine exosomes (Zhang et al., 2020). Alternatively, azido sugars have also been used to label proteins on the surface for exosome tracking and imaging (Lee et al., 2018; Wang et al., 2015).

4 | CONCLUSIONS

MS-based proteomics has significantly facilitated the comprehensive study of extracellular glycoproteins, increasing our understanding of their critical roles in the regulation of cellular activities. We review the recent advances in extracellular glycoprotein analysis using MS-based proteomics. To minimize the adverse effects from highly abundant intracellular and background proteins, chemical and enzymatic methods were developed to specifically target glycoproteins on the cell surface or in the secretome, enabling their selective enrichment before MS analysis. Strategies for analyzing surface glycoprotein interactions promote the studies of protein interaction events on the cell surface such as ligand-receptor binding and lectin-glycoprotein recognition. MS-based proteomics also expands our knowledge of secreted glycoproteins and aids in the discovery of extracellular glycoproteins as biomarkers.

During the analysis of extracellular glycoproteins, the contamination from highly abundant intracellular proteins may potentially pose a question on the specificity of separated glycoproteins only on the cell surface or in the secretome. Therefore, novel and effective separation methods to increase the enrichment specificity are critical to minimize the interference of glycoproteins from other cellular compartments. The identification and quantification of extracellular glycoproteins will be prompted with the development of next-generation instrumentation with higher sensitivity and more effective enrichment methods, especially for those with extremely low abundances that often carry more valuable information. The improvement of sensitivity will also decrease the starting materials required and facilitate the application of MS-based methods for clinical samples. The glycan information on the glycosylation site is often lost in the extracellular glycoprotein analysis. With the development of MS-based glycoproteomics including instrumentation and bioinformatic tools, large-scale analysis of intact glycopeptides has achieved promising results (Hu et al., 2017; Liu et al., 2017; Riley et al., 2019; Yu et al., 2017) promoting future studies on intact glycopeptides and then providing more valuable information of glycoproteins and how glycans regulate glycoprotein functions and interactions in the extracellular region. Recently, mucins and some other types of glycosylation such as sulfated proteins in the extracellular space have attracted much attention due to their important functions (Griffin et al., 2021; Shon et al., 2020). However, it still requires the development of novel methods to tackle the difficulty to specifically analyze these glycoproteins using MS. It is expected that future MS-based methods will not only expand the coverage of extracellular glycoproteins, but also lead to the discovery of glycoproteins as drug targets and biomarkers.

ACKNOWLEDGMENTS

The funding from the National Institute of General Medical Sciences of the National Institutes of Health (R01GM127711) and the National Science Foundation (2003597) was greatly acknowledged for supporting this study.

REFERENCES

Abbott KL, Pierce JM. 2010. Lectin-based glycoproteomic techniques for the enrichment and identification of potential biomarkers. *Methods Enzymol. 480*:461-476.

Adav SS, Hwa HH, de Kleijn D, Sze SK. 2015. Improving blood plasma glycoproteome coverage by coupling ultracentrifugation fractionation to electrostatic repulsion-hydrophilic interaction chromatography enrichment. *J. Proteome Res. 14*:2828-2838.

Akers JC, Gonda D, Kim R, Carter BS, Chen CC. 2013. Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies. *J Neurooncol. 113*:1-11.

Alagesan K, Khilji SK, Kolarich D. 2017. It is all about the solvent: on the importance of the mobile phase for ZIC-HILIC glycopeptide enrichment. *Anal. Bioanal. Chem. 409*:529-538.

Alpert AJ. 2008. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal. Chem. 80*:62-76.

Anel A, Gallego-Lleyda A, de Miguel D, Naval J, Martinez-Lostao L. 2019. Role of exosomes in the regulation of T-cell mediated immune responses and in autoimmune disease. *Cells 8*:154.

Arcinas A, Yen TY, Kebebew E, Macher BA. 2009. Cell surface and secreted protein profiles of human thyroid cancer cell lines reveal distinct glycoprotein patterns. *J. Proteome Res. 8*:3958-3968.

Arigoni-Affolter I, Scibona E, Lin CW, Bruhlmann D, Souquet J, Broly H, Aebi M. 2019. Mechanistic reconstruction of glycoprotein secretion through monitoring of intracellular N-glycan processing. *Sci. Adv. 5*:eaax8930.

Bausch-Fluck D, Goldmann U, Müller S, van Oostrum M, Müller M, Schubert OT, Wollscheid B. 2018. The in silico human surfaceome. *Proc. Natl. Acad. Sci. USA. 115*:E10988-E10997.

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. 2015. A mass spectrometric-derived cell surface protein atlas. *PLOS One 10*:e0121314.

Bausch-Fluck D, Milani ES, Wollscheid B. 2019. Surfaceome nanoscale organization and extracellular interaction networks. *Curr. Opin. Chem. Biol. 48*:26-33.

Berven FS, Ahmad R, Clauser KR, Carr SA. 2010. Optimizing performance of glycopeptide capture for plasma proteomics. *J Proteome Res. 9*:1706-1715.

Blonder J, Chan KC, Issaq HJ, Veenstra TD. 2006. Identification of membrane proteins from mammalian cell/tissue using methanol-facilitated solubilization and tryptic digestion coupled with 2D-LC-MS/MS. *Nat. Protoc. 1*:2784-2790.

Bobbitt JM. 1956. Periodate oxidation of carbohydrates. In: Wolfrom ML, Tipson RS, Editors. *Advances in Carbohydrate Chemistry*. Academic Press. p 1-41.

Boersema PJ, Geiger T, Wisniewski JR, Mann M. 2013. Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. *Mol. Cell. Proteomics 12*:158-171.

Boheler KR, Bhattacharya S, Kropp EM, Chuppa S, Riordon DR, Bausch-Fluck D, Burridge PW, Wu JC, Wersto RP, Chan GCF, Rao S, Wollscheid B, Gundry RL. 2014. A human pluripotent stem cell surface N-glycoproteome resource reveals markers, extracellular epitopes, and drug targets. *Stem Cell Rep. 3*:185-203.

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY. 2018. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol. 36*:880-887.

Bu H, He D, He X, Wang K. 2019. Exosomes: isolation, analysis, and applications in cancer detection and therapy. *ChemBioChem 20*:451-461.

Budnik B, Levy E, Harmange G, Slavov N. 2018. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol. 19*:161.

Burda P, Aebi M. 1999. The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta, Gen. Subj. 1426*:239-257.

Cao W, Huang J, Jiang B, Gao X, Yang P. 2016. Highly selective enrichment of glycopeptides based on zwitterionically functionalized soluble nanopolymers. *Sci. Rep. 6*:29776.

Carnino JM, Lee H, Jin Y. 2019. Isolation and characterization of extracellular vesicles from Broncho-alveolar lavage fluid: a review and comparison of different methods. *Respir. Res. 20*:240.

Caval T, Heck AJR, Reiding KR. 2020. Meta-heterogeneity: evaluating and describing the diversity in glycosylation between sites on the same glycoprotein. *Mol. Cell. Proteomics 20*:100010.

Chang L, Chen Y-J, Fan C-Y, Tang C-J, Chen Y-H, Low P-Y, Ventura A, Lin C-C, Chen Y-J, Angata T. 2017. Identification of Siglec ligands using a proximity labeling method. *J. Proteome Res. 16*:3929-3941.

Chang PV, Chen X, Smyrniotis C, Xenakis A, Hu T, Bertozzi CR, Wu P. 2009a. Metabolic labeling of sialic acids in living animals with alkynyl sugars. *Angew. Chem. Int. Ed. 48*: 4030-4033.

Chang YH, Wu CC, Chang KP, Yu JS, Chang YC, Liao PC. 2009b. Cell secretome analysis using hollow fiber culture system leads to the discovery of CLIC1 protein as a novel plasma marker for nasopharyngeal carcinoma. *J. Proteome Res. 8*:5465-5474.

Chauhan S, Danielson S, Clements V, Edwards N, Ostrand-Rosenberg S, Fenselau C. 2017. Surface glycoproteins of exosomes shed by myeloid-derived suppressor cells contribute to function. *J. Proteome Res. 16*:238-246.

Chen R, Williamson S, Fulton KM, Twine SM, Li J. 2019. Simultaneous analysis of phosphopeptides and intact glycopeptides from secretome with mode switchable solid phase extraction. *Anal. Methods 11*:5243-5249.

Chen W, Smeekens J, Wu R. 2014a. Comprehensive analysis of protein N-glycosylation sites by combining chemical deglycosylation with LC-MS. *J. Proteome Res. 13*:1466-1473.

Chen W, Smeekens J, Wu R. 2014b. A universal chemical enrichment method for mapping the yeast N-glycoproteome by mass spectrometry (MS). *Mol. Cell. Proteomics 13*: 1563-1572.

Chen W, Smeekens JM, Wu R. 2015. Systematic and site-specific analysis of N-sialoglycosylated proteins on the cell surface by integrating click chemistry and MS-based proteomics. *Chem. Sci. 6*:4681-4689.

Chen Z, Huang J, Li L. 2018. Recent advances in mass spectrometry (MS)-based glycoproteomics in complex biological samples. *TrAC, Trends Anal. Chem. 118*:880-892.

de Jesus JR, da Silva Fernandes R, de Souza Pessoa G, Raimundo IM, Arruda MAZ. 2017. Depleting high-abundant and enriching low-abundant proteins in human serum: an evaluation of sample preparation methods using magnetic nanoparticle, chemical depletion and immunoaffinity techniques. *Talanta 170*:199-209.

de Toro J, Herschlik L, Waldner C, Mongini C. 2015. Emerging roles of exosomes in normal and pathological conditions: new insights for diagnosis and therapeutic applications. *Front Immunol. 6*:203.

Dey KK, Wang H, Niu M, Bai B, Wang X, Li Y, Cho J-H, Tan H, Mishra A, High AA, Chen P-C, Wu Z, Beach TG, Peng J. 2019. Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. *Clin. Proteom. 16*.

Dickhout A, Koenen RR. 2018. Extracellular vesicles as biomarkers in cardiovascular disease; chances and risks. *Front. Cardiovasc. Med. 5*:113.

Doyle LM, Wang MZ. 2019. Overview of extracellular vesicles, their origin, composition, purpose, and methods for exosome isolation and analysis. *Cells 8*:727.

Dube DH, Bertozzi CR. 2003. Metabolic oligosaccharide engineering as a tool for glycobiology. *Curr. Opin. Chem. Biol. 7*:616-625.

Eichelbaum K, Winter M, Berriel Diaz M, Herzig S, Krijgsveld J. 2012. Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat. Biotechnol. 30*:984-990.

Elia G. 2008. Biotinylation reagents for the study of cell surface proteins. *Proteomics 8*:4012-4024.

Escrevente C, Grammel N, Kandzia S, Zeiser J, Tranfield EM, Conradt HS, Costa J. 2013. Sialoglycoproteins and N-glycans from secreted exosomes of ovarian carcinoma cells. *PLOS One 8*:e78631.

Escrevente C, Keller S, Altevogt P, Costa J. 2011. Interaction and uptake of exosomes by ovarian cancer cells. *BMC Cancer 11*: 108.

Feng L, Hong S, Rong J, You Q, Dai P, Huang R, Tan Y, Hong W, Xie C, Zhao J, Chen X. 2013. Bifunctional unnatural sialic acids for dual metabolic labeling of cell-surface sialylated glycans. *J. Am. Chem. Soc. 135*:9244-9247.

Feng S, Yang N, Pennathur S, Goodison S, Lubman DM. 2009. Enrichment of glycoproteins using nanoscale chelating concanavalin A monolithic capillary chromatography. *Anal. Chem. 81*:3776-3783.

Frei AP, Jeon O-Y, Kilcher S, Moest H, Henning LM, Jost C, Pluckthun A, Mercer J, Aebersold R, Carreira EM, Wollscheid B. 2012. Direct identification of ligand-receptor interactions on living cells and tissues. *Nat. Biotechnol. 30*: 997-1001.

Freitas D, Balmana M, Pocas J, Campos D, Osorio H, Konstantinidi A, Vakhrushev SY, Magalhaes A, Reis CA. 2019. Different isolation approaches lead to diverse glycosylated extracellular vesicle populations. *J. Extracell. Vesicles 8*: 1621131.

Frenette GP, Carey TE, Varani J, Schwartz DR, Fligiel SE, Ruddon RW, Peters BP. 1988. Biosynthesis and secretion of laminin and laminin-associated glycoproteins by nonmalignant and malignant human keratinocytes: comparison of cell lines from primary and secondary tumors in the same patient. *Cancer Res. 48*:5193-5202.

Gahmberg CG, Andersson LC. 1977. Selective radioactive labeling of cell surface sialoglycoproteins by periodate-tritiated borohydride. *J. Biol. Chem. 252*:5888-5894.

Gal S, Willingham MC, Gottesman MM. 1985. Processing and lysosomal localization of a glycoprotein whose secretion is transformation stimulated. *J. Cell Biol. 100*:535-544.

Geri JB, Oakley JV, Reyes-Robles T, Wang T, McCarver SJ, White CH, Rodriguez-Rivera FP, Parker DL, Hett EC, Fadeyi OO, Oslund RC, MacMillan DWC. 2020. Microenvironment mapping via Dexter energy transfer on immune cells. *Science 367*:1091.

Goo YA, Liu AY, Ryu S, Shaffer SA, Malmstrom L, Page L, Nguyen LT, Doneanu CE, Goodlett DR. 2009. Identification of secreted glycoproteins of human prostate and bladder stromal cells by comparative quantitative proteomics. *Prostate 69*: 49-61.

Granger J, Siddiqui J, Copeland S, Remick D. 2005. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics 5*:4713-4718.

Griffin ME, Sorum AW, Miller GM, Goddard WA, Hsieh-Wilson LC. 2021. Sulfated glycans engage the Ang–Tie pathway to regulate vascular development. *Nat. Chem. Biol. 17*:178–186.

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 H-J, Marshall S, Duncan SA, Wollscheid B, Wersto RP, Rao S, van Eyk JE, Boheler KR. 2012. A cell surfaceome map for immunophenotyping and sorting pluripotent stem cells. *Mol. Cell. Proteomics 11*:303-316.

Haltiwanger RS, Lowe JB. 2004. Role of glycosylation in development. *Annu. Rev. Biochem. 73*:491-537.

Han S, Collins BE, Bengtson P, Paulson JC. 2005. Homomultimeric complexes of CD22 in B cells revealed by protein-glycan cross-linking. *Nat. Chem. Biol. 1*:93-97.

Haverland NA, Waas M, Ntai I, Keppel T, Gundry RL, Kelleher NL. 2017. Cell surface proteomics of N-linked glycoproteins for typing of human lymphocytes. *Proteomics 17*:1700156.

Heck AJR. 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods 5*:927-933.

Herber J, Njavro J, Feederle R, Schepers U, Muller UC, Brase S, Muller SA, Lichtenthaler SF. 2018. Click chemistry-mediated biotinylation reveals a function for the protease BACE1 in modulating the neuronal surface glycoproteome. *Mol. Cell. Proteomics 17*:1487-1501.

Hessvik NP, Llorente A. 2018. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci. 75*:193-208.

Hortin GL, Sviridov D. 2010. The dynamic range problem in the analysis of the plasma proteome. *J. Proteomics 73*:629-636.

Hu H, Khatri K, Zaia J. 2017. Algorithms and design strategies towards automated glycoproteomics analysis. *Mass Spectrom. Rev. 36*:475-498.

Ideo H, Kondo J, Nomura T, Nonomura N, Inoue M, Amano J. 2020. Study of glycosylation of prostate-specific antigen secreted by cancer tissue-originated spheroids reveals new candidates for prostate cancer detection. *Sci. Rep. 10*:2708.

Jankovska E, Svitek M, Holada K, Petrak J. 2019. Affinity depletion versus relative protein enrichment: a side-by-side comparison of two major strategies for increasing human cerebrospinal fluid proteome coverage. *Clin. Proteomics 16*:9.

Jiang S, Kotani N, Ohnishi T, Miyagawa-Yamguchi A, Tsuda M, Yamashita R, Ishiura Y, Honke K. 2011. A proteomics approach to the cell-surface interactome using the enzyme-mediated activation of radical sources reaction. *Proteomics 12*: 54-62.

Joeh E, O'Leary T, Li W, Hawkins R, Hung JR, Parker CG, Huang ML. 2020. Mapping glycan-mediated galectin-3 interactions by live cell proximity labeling. *Proc. Natl. Acad. Sci. USA. 117*:27329.

Jung K, Cho W, Regnier FE. 2009. Glycoproteomics of plasma based on narrow selectivity lectin affinity chromatography. *J. Proteome Res. 8*:643-650.

Kalxdorf M, Gade S, Eberl HC, Bantscheff M. 2017. Monitoring cell-surface N-glycoproteome dynamics by quantitative proteomics reveals mechanistic insights into macrophage differentiation. *Mol. Cell. Proteomics 16*:770-785.

Kim HO, Kim HS, Youn JC, Shin EC, Park S. 2011. Serum cytokine profiles in healthy young and elderly population assessed using multiplexed bead-based immunoassays. *J. Transl. Med. 9*:113.

Kirwan A, Utratna M, O'Dwyer ME, Joshi L, Kilcoyne M. 2015. Glycosylation-based serum biomarkers for cancer diagnostics and prognostics. *Biomed. Res. Int. 2015*:490531.

Konoshenko MY, Lekchnov EA, Vlassov AV, Laktionov PP. 2018. Isolation of extracellular vesicles: general methodologies and latest trends. *Biomed. Res. Int. 2018*:8545347.

Kuhn PH, Koroniak K, Hogl S, Colombo A, Zeitschel U, Willem M, Volbracht C, Schepers U, Imhof A, Hoffmeister A, Haass C, Rossner S, Brase S, Lichtenthaler SF. 2012. Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons. *EMBO J. 31*:3157-3168.

Lang K, Chin JW. 2014. Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol. 9*:16-20.

Lebowitz J, Lewis MS, Schuck P. 2002. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci. 11*:2067-2079.

Lee TS, Kim Y, Zhang W, Song IH, Tung CH. 2018. Facile metabolic glycan labeling strategy for exosome tracking. *Biochim. Biophys. Acta. Gen. Subj. 1862*:1091-1100.

Leung KK, Wilson GM, Kirkemo LL, Riley NM, Coon JJ, Wells JA. 2020. Broad and thematic remodeling of the surfaceome and glycoproteome on isogenic cells transformed with driving proliferative oncogenes. *Proc. Natl. Acad. Sci. USA. 117*: 7764-7775.

Li F, Glinskii OV, Mooney BP, Rittenhouse-Olson K, Pienta KJ, Glinsky VV. 2017. Cell surface Thomsen-Friedenreich proteome profiling of metastatic prostate cancer cells reveals potential link with cancer stem cell-like phenotype. *Oncotarget 8*:98598-98608.

Li Q, Xie Y, Xu G, Lebrilla CB. 2019a. Identification of potential sialic acid binding proteins on cell membranes by proximity chemical labeling. *Chem. Sci. 10*:6199-6209.

Li X, Jiang J, Zhao X, Wang J, Han H, Zhao Y, Peng B, Zhong R, Ying W, Qian X. 2013. N-glycoproteome analysis of the secretome of human metastatic hepatocellular carcinoma cell lines combining hydrazide chemistry, HILIC enrichment and mass spectrometry. *PLOS One 8*:e81921.

Li Y, Wang Y, Mao J, Yao Y, Wang K, Qiao Q, Fang Z, Ye M. 2019b. Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *J. Proteomics 196*:33-41.

Li Y, Tian C, Liu K, Zhou Y, Yang J, Zou P. 2020. A clickable APEX probe for proximity-dependent proteomic profiling in yeast. *Cell Chem. Biol. 27*:858-865.

Liang X, Zhao J, Hajivandi M, Wu R, Tao J, Amshey JW, Pope RM. 2006. Quantification of membrane and membrane-bound proteins in normal and malignant breast cancer cells isolated from the same patient with primary breast carcinoma. *J. Proteome Res. 5*:2632-2641.

Lim CZJ, Zhang L, Zhang Y, Sundah NR, Shao H. 2020. New sensors for extracellular vesicles: insights on constituent and associated biomarkers. *ACS Sens. 5*:4-12.

Liu LT, Zhang Y, Zhang L, Yan GQ, Yao J, Yang PY, Lu HJ. 2012. Highly specific revelation of rat serum glycopeptidome by boronic acid-functionalized mesoporous silica. *Anal. Chim. Acta 753*:64-72.

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. 2017. pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nat. Commun. 8*:438.

Liu Q, Zheng J, Sun W, Huo Y, Zhang L, Hao P, Wang H, Zhuang M. 2018. A proximity-tagging system to identify membrane protein–protein interactions. *Nat. Methods 15*:715-722.

Liu YS, He JT, Li C, Benitez R, Fu S, Marrero J, Lubman DM. 2010. Identification and confirmation of biomarkers using an integrated platform for quantitative analysis of glycoproteins and their glycosylations. *J. Proteome Res. 9*:798-805.

Liu Z, He H. 2017. Synthesis and applications of boronate affinity materials: from class selectivity to biomimetic specificity. *Acc. Chem. Res. 50*:2185-2193.

Lopez Aguilar A, Briard JG, Yang L, Ovryn B, Macauley MS, Wu P. 2017. Tools for studying glycans: recent advances in chemoenzymatic glycan labeling. *ACS Chem. Biol. 12*:611-621.

Lund R, Leth-Larsen R, Jensen ON, Ditzel HJ. 2009. Efficient isolation and quantitative proteomic analysis of cancer cell plasma membrane proteins for identification of metastasis-associated cell surface markers. *J. Proteome Res. 8*:3078-3090.

Maacha S, Bhat AA, Jimenez L, Raza A, Haris M, Uddin S, Grivel JC. 2019. Extracellular vesicles-mediated intercellular communication: roles in the tumor microenvironment and anti-cancer drug resistance. *Mol. Cancer 18*:55.

Mahal LK, Yarema KJ, Bertozzi CR. 1997. Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis. *Science 276*:1125.

Mallanna SK, Waas M, Duncan SA, Gundry RL. 2017. N-glycoprotein surfaceome of human induced pluripotent stem cell derived hepatic endoderm. *Proteomics 17*:1600397.

Martinez-Martin N, Ramani SR, Hackney JA, Tom I, Wranik BJ, Chan M, Wu J, Paluch MT, Takeda K, Hass PE, Clark H, Gonzalez LC. 2016. The extracellular interactome of the human adenovirus family reveals diverse strategies for immunomodulation. *Nat. Commun. 7*:11473.

Mashouri L, Yousefi H, Aref AR, Ahadi AM, Molaei F, Alahari SK. 2019. Exosomes: composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. *Mol. Cancer 18*:75.

Mechref Y, Madera M, Novotny MV. 2008. Glycoprotein enrichment through lectin affinity techniques. *Methods Mol. Biol. 424*:373-396.

Merlot AM, Kalinowski DS, Richardson DR. 2014. Unraveling the mysteries of serum albumin-more than just a serum protein. *Front. Physiol. 5*:299.

Merrell K, Southwick K, Graves SW, Esplin MS, Lewis NE, Thulin CD. 2004. Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. *J. Biomol. Tech. 15*:238-248.

Meyfour A, Pahlavan S, Mirzaei M, Krijgsveld J, Baharvand H, Salekdeh GH. 2021. The quest of cell surface markers for stem cell therapy. *Cell. Mol. Life Sci. 78*:469-495.

Mirkowska P, Hofmann A, Sedek L, Slamova L, Mejstrikova E, Szczepanski T, Schmitz M, Cario G, Stanulla M, Schrappe M, van der Velden VHJ, Bornhauser BC, Wollscheid B, Bourquin J-P. 2013. Leukemia surfaceome analysis reveals new disease-associated features. *Blood 121*:e149-e159.

Moremen KW, Tiemeyer M, Nairn AV. 2012. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol. 13*:448.

Nie S, Lo A, Wu J, Zhu JH, Tan ZJ, Simeone DM, Anderson MA, Shedden KA, Ruffin MT, Lubman DM. 2014. Glycoprotein biomarker panel for pancreatic cancer discovered by quantitative proteomics analysis. *J. Proteome Res. 13*: 1873-1884.

Oldham RAA, Faber ML, Keppel TR, Buchberger AR, Waas M, Hari P, Gundry RL, Medin JA. 2020. Discovery and validation of surface N-glycoproteins in MM cell lines and patient samples uncovers immunotherapy targets. *J. Immunother. Cancer 8*:e000915.

Paek J, Kalocsay M, Staus DP, Wingler L, Pascolutti R, Paulo JA, Gygi SP, Kruse AC. 2017. Multidimensional tracking of GPCR signaling via peroxidase-catalyzed proximity labeling. *Cell 169*: 338-349.

Pan D, Chen J, Feng C, Wu W, Wang Y, Tong J, Zhou D. 2019. Preferential localization of MUC1 glycoprotein in exosomes secreted by non-small cell lung carcinoma cells. *Int. J. Mol. Sci. 20*:323.

Pan P, Zhang Q, Hou J, Liu Z, Bai F, Cao M, Sun T, Bai G. 2012. Cell surface glycoprotein profiling of cancer cells based on bioorthogonal chemistry. *Anal. Bioanal. Chem. 403*:1661-1670.

Pare B, Deschenes LT, Pouliot R, Dupre N, Gros-Louis F. 2016. An optimized approach to recover secreted proteins from fibroblast conditioned-media for secretomic analysis. *Front. Cell Neurosci. 10*:70.

Parikka K, Master E, Tenkanen M. 2015. Oxidation with galactose oxidase: multifunctional enzymatic catalysis. *J. Mol. Catal. B: Enzym. 120*:47-59.

Park DD, Xu G, Wong M, Phoomak C, Liu M, Haigh NE, Wongkham S, Yang P, Maverakis E, Lebrilla CB. 2018. Membrane glycomics reveal heterogeneity and quantitative distribution of cell surface sialylation. *Chem. Sci. 9*:6271-6285.

Patel GK, Khan MA, Zubair H, Srivastava SK, Khushman M, Singh S, Singh AP. 2019. Comparative analysis of exosome isolation methods using culture supernatant for optimum yield, purity and downstream applications. *Sci. Rep. 9*:5335.

Pinho SS, Reis CA. 2015. Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer 15*:540-555.

Plavina T, Wakshull E, Hancock WS, Hincapie M. 2007. Combination of abundant protein depletion and multi-lectin affinity chromatography (M-LAC) for plasma protein biomarker discovery. *J. Proteome Res. 6*:662-671.

Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS. 2010. High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. *Electrophoresis 31*: 471-482.

Prescher JA, Bertozzi CR. 2005. Chemistry in living systems. *Nat. Chem. Biol. 1*:13-21.

Prior MJ, Larance M, Lawrence RT, Soul J, Humphrey S, Burchfield J, Kistler C, Davey JR, La-Borde PJ, Buckley M, Kanazawa H, Parton RG, Guilhaus M, James DE. 2011. Quantitative proteomic analysis of the adipocyte plasma membrane. *J. Proteome Res. 10*:4970-4982.

Qiu H, Wang Y. 2008. Quantitative analysis of surface plasma membrane proteins of primary and metastatic melanoma cells. *J. Proteome Res*. *7*:1904-1915.

Qiu R, Regnier FE. 2005. Use of multidimensional lectin affinity chromatography in differential glycoproteomics. *Anal. Chem*. *77*:2802-2809.

Qu YY, Liu JX, Yang KG, Liang Z, Zhang LH, Zhang YK. 2012. Boronic acid functionalized core-shell polymer nanoparticles prepared by distillation precipitation polymerization for glycopeptide enrichment. *Chem.-Eur. J*. *18*:9056-9062.

Ramya TN, Weerapana E, Liao L, Zeng Y, Tateno H, Liao L, Yates JR, Cravatt BF, Paulson JC. 2010. In situ trans ligands of CD22 identified by glycan-protein photocross-linking-enabled proteomics. *Mol. Cell. Proteomics 9*:1339-1351.

Ramya TN, Weerapana E, Cravatt BF, Paulson JC. 2013. Glycoproteomics enabled by tagging sialic acid- or galactose-terminated glycans. *Glycobiology 23*:211-221.

Ravenhill BJ, Kanjee U, Ahouidi A, Nobre L, Williamson J, Goldberg JM, Antrobus R, Dieye T, Duraisingh MT, Weekes MP. 2019. Quantitative comparative analysis of human erythrocyte surface proteins between individuals from two genetically distinct populations. *Commun. Biol*. *2*:350.

Ravenhill BJ, Soday L, Houghton J, Antrobus R, Weekes MP. 2020. Comprehensive cell surface proteomics defines markers of classical, intermediate and non-classical monocytes. *Sci. Rep*. *10*:4560.

Rees JS, Li X-W, Perrett S, Lilley KS, Jackson AP. 2017. Selective proteomic proximity labeling assay using tyramide (SPPLAT): a quantitative method for the proteomic analysis of localized membrane-bound protein clusters. *Curr. Protoc. Protein Sci*. *88*:19.27.1-19.27.18.

Reily C, Stewart TJ, Renfrow MB, Novak J. 2019. Glycosylation in health and disease. *Nat. Rev. Nephrol*. *15*:346-366.

Riley NM, Hebert AS, Westphall MS, Coon JJ. 2019. Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. *Nat. Commun*. *10*:1311.

Rogers LJ, Suchowerska N, Khan A, Polaskova V, McKenzie DR. 2014. Profiling of the secretome of human cancer cells: preparation of supernatant for proteomic analysis. *Electrophoresis 35*:2626-2633.

Roper SM, Zemskova M, Neely BA, Martin A, Gao P, Ellen Jones E, Kraft AS, Drake RR. 2013. Targeted glycoprotein enrichment and identification in stromal cell secretomes using azido sugar metabolic labeling. *Proteomics Clin. Appl. 7*:367-371.

Russell CL, Heslegrave A, Mitra V, Zetterberg H, Pocock JM, Ward MA, Pike I. 2017. Combined tissue and fluid proteomics with tandem mass tags to identify low-abundance protein biomarkers of disease in peripheral body fluid: an Alzheimer's Disease case study. *Rapid Commun. Mass Spectrom*. *31*:153-159.

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. 2018. Similarities and differences of blood n-glycoproteins in five solid carcinomas at localized clinical stage analyzed by SWATH-MS. *Cell Rep*. *23*:2819-2831.

Samavarchi-Tehrani P, Samson R, Gingras AC. 2020. Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches. *Mol. Cell. Proteomics 19*:757-773.

Saraswat M, Joenvaara S, Musante L, Peltoniemi H, Holthofer H, Renkonen R. 2015. N-linked (N-) glycoproteomics of urinary exosomes. *Mol. Cell. Proteomics 14*:263.

Schiess R, Mueller LN, Schmidt A, Mueller M, Wollscheid B, Aebersold R. 2009. Analysis of cell surface proteome changes via label-free, quantitative mass spectrometry. *Mol. Cell. Proteomics 8*:624-638.

Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. 2020. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol*. *21*:729-749.

Seiradake E, Zhao Y, Lu W, Aricescu AR, Jones EY. 2015. Production of cell surface and secreted glycoproteins in mammalian cells. *Methods Mol. Biol*. *1261*:115-127.

Shon DJ, Malaker SA, Pedram K, Yang E, Krishnan V, Dorigo O, Bertozzi CR. 2020. An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins. *Proc. Natl. Acad. Sci. USA*. *117*:21299-21307.

Shu Q, Li M, Shu L, An Z, Wang J, Lv H, Yang M, Cai T, Hu T, Fu Y, Yang F. 2020. Large-scale identification of N-linked intact glycopeptides in human serum using HILIC enrichment and spectral library search. *Mol. Cell. Proteomics 19*:672-689.

Sletten EM, Bertozzi CR. 2009. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed*. *48*:6974-6998.

Smeekens JM, Chen W, Wu R. 2015. Mass spectrometric analysis of the cell surface N-glycoproteome by combining metabolic labeling and click chemistry. *J. Am. Soc. Mass Spectrom*. *26*:604-614.

Smeekens JM, Xiao H, Wu R. 2017. Global analysis of secreted proteins and glycoproteins in saccharomyces cerevisiae. *J. Proteome Res*. *16*:1039-1049.

Smits AH, Vermeulen M. 2016. Characterizing protein–protein interactions using mass spectrometry: challenges and opportunities. *Trends Biotechnol*. *34*:825-834.

Sobotzki N, Schafroth MA, Rudnicka A, Koetemann A, Marty F, Goetze S, Yamauchi Y, Carreira EM, Wollscheid B. 2018. HATRIC-based identification of receptors for orphan ligands. *Nat. Commun*. *9*:1519.

Solis N, Cordwell SJ. 2011. Current methodologies for proteomics of bacterial surface-exposed and cell envelope proteins. *Proteomics 11*:3169-3189.

Sparbier K, Asperger A, Resemann A, Kessler I, Koch S, Wenzel T, Stein G, Vorwerg L, Suckau D, Kostrzewa M. 2007. Analysis of glycoproteins in human serum by means of glycospecific magnetic bead separation and LC-MALDI-TOF/TOF analysis with automated glycopeptide detection. *J. Biomol. Tech*. *18*:252-258.

Spiciarich DR, Nolley R, Maund SL, Purcell SC, Herschel J, Iavarone AT, Peehl DM, Bertozzi CR. 2017. Bioorthogonal labeling of human prostate cancer tissue slice cultures for glycoproteomics. *Angew. Chem. Int. Ed*. *56*:8992-8997.

Spiro RG. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology 12*:43R-56R.

Stastna M, Van Eyk JE. 2012. Investigating the secretome: lessons about the cells that comprise the heart. *Circ. Cardiovasc. Genet*. *5*:o8-o18.

Sun F, Suttapitugsakul S, Wu R. 2019a. Enzymatic tagging of glycoproteins on the cell surface for their global and site-specific analysis with mass spectrometry. *Anal. Chem. 91*: 4195-4203.

Sun F, Suttapitugsakul S, Wu R. 2021. Unraveling the surface glycoprotein interaction network by integrating chemical crosslinking with MS-based proteomics. *Chem. Sci. 12*: 2146-2155.

Sun F, Wu R. 2019. Systematic and site-specific analysis of N-glycoproteins on the cell surface by integrating bioorthogonal chemistry and MS-based proteomics. *Methods Enzymol. 626*: 223-247.

Sun S, Hu Y, Ao M, Shah P, Chen J, Yang W, Jia X, Tian Y, Thomas S, Zhang H. 2019b. *N*-GlycositeAtlas: a database resource for mass spectrometry-based human N-linked glycoprotein and glycosylation site mapping. *Clin. Proteomics 16*:35.

Sun T, Yu SH, Zhao P, Meng L, Moremen KW, Wells L, Steet R, Boons GJ. 2016. One-step selective exoenzymatic labeling (SEEL) strategy for the biotinylation and identification of glycoproteins of living cells. *J. Am. Chem. Soc. 138*: 11575-11582.

Suttapitugsakul S, Ulmer LD, Jiang C, Sun F, Wu R. 2019. Surface glycoproteomic analysis reveals that both unique and differential expression of surface glycoproteins determine the cell type. *Anal. Chem. 91*:6934-6942.

Suttapitugsakul S, Sun F, Wu R. 2020. Recent advances in glycoproteomic analysis by mass spectrometry. *Anal. Chem. 92*:267-291.

Suttapitugsakul S, Tong M, Sun F, Wu R. 2021a. Enhancing comprehensive analysis of secreted glycoproteins from cultured cells without serum starvation. *Anal. Chem. 93*: 2694-2705.

Suttapitugsakul S, Tong M, Wu R. 2021b. Time-resolved and comprehensive analysis of surface glycoproteins reveals distinct responses of monocytes and macrophages to bacterial infection.*Angew. Chem. Int. Ed. 60*:11494-11503.

Swietlik JJ, Sinha A, Meissner F. 2020. Dissecting intercellular signaling with mass spectrometry–based proteomics. *Curr. Opin. Cell Biol. 63*:20-30.

Totten SM, Adusumilli R, Kullolli M, Tanimoto C, Brooks JD, Mallick P, Pitteri SJ. 2018. Multi-lectin affinity chromatography and quantitative proteomic analysis reveal differential glycoform levels between prostate cancer and benign prostatic hyperplasia sera. *Sci. Rep. 8*:6509.

Tremblay T-L, Hill JJ. 2017. Biotin-transfer from a trifunctional crosslinker for identification of cell surface receptors of soluble protein ligands. *Sci. Rep. 7*:46574.

Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJ, Liebler DC. 2010. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J. Proteome Res. 9*:4982-4991.

Tushaus J, Muller SA, Kataka ES, Zaucha J, Sebastian Monasor L, Su M, Guner G, Jocher G, Tahirovic S, Frishman D, Simons M, Lichtenthaler SF. 2020. An optimized quantitative proteomics method establishes the cell type-resolved mouse brain secretome. *EMBO J. 39*:e105693.

Udeshi ND, Pedram K, Svinkina T, Fereshetian S, Myers SA, Aygun O, Krug K, Clauser K, Ryan D, Ast T, Mootha VK, Ting AY, Carr SA. 2017. Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Methods 14*:1167-1170.

Uhlen M, Karlsson MJ, Hober A, Svensson AS, Scheffel J, Kotol D, Zhong W, Tebani A, Strandberg L, Edfors F, Sjostedt E, Mulder J, Mardinoglu A, Berling A, Ekblad S, Dannemeyer M, Kanje S, Rockberg J, Lundqvist M, Malm M, Volk AL, Nilsson P, Manberg A, Dodig-Crnkovic T, Pin E, Zwahlen M, Oksvold P, von Feilitzen K, Haussler RS, Hong MG, Lindskog C, Ponten F, Katona B, Vuu J, Lindstrom E, Nielsen J, Robinson J, Ayoglu B, Mahdessian D, Sullivan D, Thul P, Danielsson F, Stadler C, Lundberg E, Bergstrom G, Gummesson A, Voldborg BG, Tegel H, Hober S, Forsstrom B, Schwenk JM, Fagerberg L, Sivertsson A. 2019. The human secretome. *Sci. Signal. 12*:eaaz0274.

Vajaria B, Patel P. 2017. Glycosylation: a hallmark of cancer? *Glycoconj. J. 34*:147-156.

van der Pol E, Boing AN, Harrison P, Sturk A, Nieuwland R. 2012. Classification, functions, and clinical relevance of extracellular vesicles. *Pharmacol. Rev. 64*:676-705.

van Oostrum M, Müller M, Klein F, Bruderer R, Zhang H, Pedrioli PGA, Reiter L, Tsapogas P, Rolink A, Wollscheid B. 2019. Classification of mouse B cell types using surfaceome proteotype maps. *Nat. Commun. 10*:5734.

van Oostrum M, Campbell B, Seng C, Müller M, tom Dieck S, Hammer J, Pedrioli PGA, Földy C, Tyagarajan SK, Wollscheid B. 2020. Surfaceome dynamics reveal proteostasis-independent reorganization of neuronal surface proteins during development and synaptic plasticity. *Nat. Commun. 11*: 4990.

Varki A. 2016. Biological roles of glycans. *Glycobiology 27*:3-49.

Veillon L, Fakih C, Abou-El-Hassan H, Kobeissy F, Mechref Y. 2018. Glycosylation changes in brain cancer. *ACS Chem. Neurosci. 9*:51-72.

Waas M, Snarrenberg ST, Littrell J, Jones Lipinski RA, Hansen PA, Corbett JA, Gundry RL. 2020. SurfaceGenie: a web-based application for prioritizing cell-type-specific marker candidates. *Bioinformatics 36*:3447-3456.

Wands AM, Fujita A, McCombs JE, Cervin J, Dedic B, Rodriguez AC, Nischan N, Bond MR, Mettlen M, Trudgian DC, Lemoff A, Quiding-Järbrink M, Gustavsson B, Steentoft C, Clausen H, Mirzaei H, Teneberg S, Yrlid U, Kohler JJ. 2015. Fucosylation and protein glycosylation create functional receptors for cholera toxin. *eLife 4*:e09545.

Wang H, Mooney DJ. 2020. Metabolic glycan labelling for cancer-targeted therapy. *Nat. Chem. 12*:1102-1114.

Wang L, Aryal UK, Dai Z, Mason AC, Monroe ME, Tian ZX, Zhou JY, Su D, Weitz KK, Liu T, Camp DG, Smith RD, Baker SE, Qian WJ. 2012. Mapping N-linked glycosylation sites in the secretome and whole cells of Aspergillus niger using hydrazide chemistry and mass spectrometry. *J. Proteome Res. 11*:143-156.

Wang M, Altinoglu S, Takeda YS, Xu Q. 2015. Integrating protein engineering and bioorthogonal click conjugation for extracellular vesicle modulation and intracellular delivery. *PLOS One 10*:e0141860.

Weaving G, Batstone GF, Jones RG. 2016. Age and sex variation in serum albumin concentration: an observational study. *Ann. Clin. Biochem. 53*:106-111.

Wei X, Dulberger C, Li LJ. 2010. Characterization of murine brain membrane glycoproteins by detergent assisted lectin affinity chromatography. *Anal. Chem. 82*:6329-6333.

Wen L, Zheng Y, Jiang K, Zhang M, Kondengaden SM, Li S, Huang K, Li J, Song J, Wang PG. 2016. Two-step chemoenzymatic detection of N-acetylneuraminic acid−α(2-3)-galactose glycans. *J. Am. Chem. Soc. 138*:11473-11476.

Wen L, Liu D, Zheng Y, Huang K, Cao X, Song J, Wang PG. 2018. A one-step chemoenzymatic labeling strategy for probing sialylated Thomsen–Friedenreich antigen. *ACS Cent. Sci. 4*: 451-457.

Williams C, Royo F, Aizpurua-Olaizola O, Pazos R, Boons GJ, Reichardt NC, Falcon-Perez JM. 2018. Glycosylation of extracellular vesicles: current knowledge, tools and clinical perspectives. *J. Extracell. Vesicles 7*:1442985.

Witze ES, Old WM, Resing KA, Ahn NG. 2007. Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods 4*:798-806.

Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibel M, Schiess R, Aebersold R, Watts JD. 2009. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol. 27*:378-386.

Wright GJ, Martin S, Bushell KM, Söllner C. 2010. High-throughput identification of transient extracellular protein interactions. *Biochem. Soc. Trans. 38*:919-922.

Wu H, Kohler J. 2019. Photocrosslinking probes for capture of carbohydrate interactions. *Curr. Opin. Chem. Biol. 53*:173-182.

Wu R, Dephoure N, Haas W, Huttlin EL, Zhai B, Sowa ME, Gygi SP. 2011a. Correct interpretation of comprehensive phosphorylation dynamics requires normalization by protein expression changes. *Mol. Cell. Proteomics 10*:M111.009654.

Wu R, Haas W, Dephoure N, Huttlin EL, Zhai B, Sowa ME, Gygi SP. 2011b. A large-scale method to measure absolute protein phosphorylation stoichiometries. *Nat. Methods 8*: 677-683.

Xia C, Jiao F, Gao F, Wang H, Lv Y, Shen Y, Zhang Y, Qian X. 2018. Two-dimensional MoS2-based zwitterionic hydrophilic interaction liquid chromatography material for the specific enrichment of glycopeptides. *Anal. Chem. 90*:6651-6659.

Xiao H, Chen W, Smeekens JM, Wu R. 2018a. An enrichment method based on synergistic and reversible covalent interactions for large-scale analysis of glycoproteins. *Nat. Commun. 9*:1692.

Xiao H, Hwang J, Wu R. 2018b. Mass spectrometric analysis of the N-glycoproteome in statin-treated liver cells with two lectin-independent chemical enrichment methods. *Int. J. Mass Spectrom. 429*:66-75.

Xiao H, Smeekens J, Wu R. 2016a. Quantification of tunicamycin-induced protein expression and N-glycosylation changes in yeast. *Analyst 141*:3737-3745.

Xiao H, Sun F, Suttapitugsakul S, Wu R. 2019. Global and site-specific analysis of protein glycosylation in complex biological systems with mass spectrometry. *Mass Spectrom. Rev. 38*: 356-379.

Xiao H, Suttapitugsakul S, Sun F, Wu R. 2018c. Mass spectrometry-based chemical and enzymatic methods for global analysis of protein glycosylation. *Acc. Chem. Res. 51*:1796-1806.

Xiao H, Tang GX, Wu R. 2016b. Site-specific quantification of surface N-glycoproteins in statin-treated liver cells. *Anal. Chem. 88*:3324-3332.

Xiao H, Wu R. 2017a. Quantitative investigation of human cell surface N-glycoprotein dynamics. *Chem. Sci. 8*:268-277.

Xiao H, Wu R. 2017b. Simultaneous quantitation of glycoprotein degradation and synthesis rates by integrating isotope labeling, chemical enrichment, and multiplexed proteomics. *Anal. Chem. 89*:10361-10367.

Xiao H, Wu R. 2017c. Global and site-specific analysis revealing unexpected and extensive protein S-GlcNAcylation in human cells. *Anal. Chem. 89*:3656-3663.

Xie Y, Sheng Y, Li Q, Ju S, Reyes J, Lebrilla CB. 2020. Determination of the glycoprotein specificity of lectins on cell membranes through oxidative proteomics. *Chem. Sci. 11*: 9501-9512.

Xu S, Sun F, Tong M, Wu R. 2021. MS-based proteomics for comprehensive investigation of protein O-GlcNAcylation. *Mol. Omics 17*:186-196.

Xu S, Sun F, Wu R. 2020. A chemoenzymatic method based on easily accessible enzymes for profiling protein O-GlcNAcylation. *Anal. Chem. 92*:9807-9814.

Yagi H, Yagi-Utsumi M, Honda R, Ohta Y, Saito T, Nishio M, Ninagawa S, Suzuki K, Anzai T, Kamiya Y, Aoki K, Nakanishi M, Satoh T, Kato K. 2020. Improved secretion of glycoproteins using an N-glycan-restricted passport sequence tag recognized by cargo receptor. *Nat. Commun. 11*:1368.

Yang Z, Hancock WS. 2004. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J. Chromatogr. 1053*:79-88.

Yi L, Tsai CF, Dirice E, Swensen AC, Chen J, Shi T, Gritsenko MA, Chu RK, Piehowski PD, Smith RD, Rodland KD, Atkinson MA, Mathews CE, Kulkarni RN, Liu T, Qian WJ. 2019. Boosting to amplify signal with isobaric labeling (BASIL) strategy for comprehensive quantitative phosphoproteomic characterization of small populations of cells. *Anal. Chem. 91*:5794-5801.

Yu AY, Zhao JF, Peng WJ, Banazadeh A, Williamson SD, Goli M, Huang YF, Mechref Y. 2018. Advances in mass spectrometry-based glycoproteomics. *Electrophoresis 39*:3104-3122.

Yu Q, Wang BW, Chen ZW, Urabe G, Glover MS, Shi XD, Guo LW, Kent KC, Li LJ. 2017. Electron-transfer/higher-energy collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization. *J. Am. Soc. Mass Spectrom. 28*:1751-1764.

Yu S-H, Zhao P, Sun T, Gao Z, Moremen KW, Boons G-J, Wells L, Steet R. 2016. Selective exo-enzymatic labeling detects increased cell surface sialoglycoprotein expression upon megakaryocytic differentiation. *J. Biol. Chem. 291*:3982-3989.

Zhang H, Li X-J, Martin DB, Aebersold R. 2003. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol. 21*:660-666.

Zhang H, Lv Y, Du J, Shao W, Jiao F, Xia C, Gao F, Yu Q, Liu Y, Zhang W, Zhang Y, Qin W, Qian X. 2020. A GSH functionalized magnetic ultra-thin 2D-MoS_2 nanocomposite for HILIC-based enrichment of N-glycopeptides from urine exosome and serum proteins. *Anal. Chim. Acta 1098*:181-189.

Zhang LJ, Xu YW, Yao HL, Xie LQ, Yao J, Lu HJ, Yang PY. 2009. Boronic acid functionalized core-satellite composite nanoparticles for advanced enrichment of glycopeptides and glycoproteins. *Chem.-Eur. J. 15*:10158-10166.

Zhang Y, Tang X, Yao L, Chen K, Jia W, Hu X, Xu LX. 2012. Lectin capture strategy for effective analysis of cell secretome. *Proteomics 12*:32-36.

Zhao J, Qiu WL, Simeone DM, Lubman DM. 2007. N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *J. Proteome Res. 6*:1126-1138.

Zhao Y, Zhang W, Kho Y, Zhao Y. 2004. Proteomic analysis of integral plasma membrane proteins. *Anal. Chem. 76*: 1817-1823.

Zheng J, Xiao H, Wu R. 2017. Specific identification of glycoproteins bearing the Tn antigen in human cells. *Angew. Chem. Int. Ed. 56*:7107-7111.

Zheng X, Baker H, Hancock WS, Fawaz F, McCaman M, Pungor E 2006. Proteomic analysis for the assessment of different lots of fetal bovine serum as a raw material for cell culture. Part IV. Application of proteomics to the manufacture of biological drugs. *Biotechnol. Prog. 22*:1294-1300.

Zhu JH, Huang JF, Zhang J, Chen ZW, Lin Y, Grigorean G, Li LJ, Liu SY, Singal AG, Parikh ND, Lubman DM. 2020. Glycopeptide biomarkers in serum haptoglobin for hepatocellular carcinoma detection in patients with nonalcoholic steatohepatitis. *J. Proteome Res. 19*:3452-3466.

Zhu JH, Warner E, Parikh ND, Lubman DM. 2019. Glycoproteomic markers of hepatocellular carcinoma-mass spectrometry based approaches. *Mass Spectrom. Rev. 38*:265-290.

AUTHOR BIOGRAPHIES

Fangxu Sun obtained his Master's degree from the Shanghai Institute of Organic Chemistry (SIOC) in 2016 with a focus on synthesis of well-defined graft polymers. After graduation, he switched to the field of mass spectrometry and started to work with Prof. Ronghu Wu at the Georgia Institute of Technology. His current research mainly focuses on developing novel MS-based methods for glycoprotein analysis.

Suttipong Suttapitugsakul received his Bachelor's degree from the Rensselaer Polytechnic Institute in 2016. Intrigued by the beauty of MS-based proteomics, he decided to pursue his Ph.D with Prof. Ronghu Wu in the School of Chemistry and Biochemistry at the Georgia Institute of Technology. He currently works on applying MS-based proteomics to investigate glycoproteins from the secretome and the cell surface, and decipher their biological significance in cellular systems.

Ronghu Wu is an associate professor in the School of Chemistry and Biochemistry at the Georgia Institute of Technology. After his postdoc at the Harvard Medical School, he established his own lab in 2012. The research in his lab mainly focuses on global analysis of protein dynamics and modifications, especially glycosylation, by using MS-based proteomics.

How to cite this article: Sun F, Suttapitugsakul S, Wu R. 2021. Systematic characterization of extracellular glycoproteins using mass spectrometry. *Mass Spectrom Rev*. 1-27. https://doi.org/10.1002/mas.21708