RSC Advances

REVIEW

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2023, **13**, 264

Recent development of analytical methods for disease-specific protein *O*-GlcNAcylation

Wenhua Hu,[a] Guolin Zhang,[b] Yu Zhou,[c] Jun Xia,[c] Peng Zhang,[d] Wenjin Xiao,[e] Man Xue,[b] Zhaohui Lu*[f] and Shuang Yang *[a]

The enzymatic modification of protein serine or threonine residues by *N*-acetylglucosamine, namely *O*-GlcNAcylation, is a ubiquitous post-translational modification that frequently occurs in the nucleus and cytoplasm. *O*-GlcNAcylation is dynamically regulated by two enzymes, *O*-GlcNAc transferase and *O*-GlcNAcase, and regulates nearly all cellular processes in epigenetics, transcription, translation, cell division, metabolism, signal transduction and stress. Aberrant *O*-GlcNAcylation has been shown in a variety of diseases, including diabetes, neurodegenerative diseases and cancers. Deciphering *O*-GlcNAcylation remains a challenge due to its low abundance, low stoichiometry and extreme lability in most tandem mass spectrometry. Separation or enrichment of *O*-GlcNAc proteins or peptides from complex mixtures has been of great interest because quantitative analysis of protein *O*-GlcNAcylation can elucidate their functions and regulatory mechanisms in disease. However, valid and specific analytical methods are still lacking, and efforts are needed to further advance this direction. Here, we provide an overview of recent advances in various analytical methods, focusing on chemical oxidation, affinity of antibodies and lectins, hydrophilic interaction, and enzymatic addition of monosaccharides in conjugation with these methods. *O*-GlcNAcylation quantification has been described in detail using mass-spectrometric or non-mass-spectrometric techniques. We briefly summarized dysregulated changes in *O*-GlcNAcylation in disease.

Received 12th November 2022
Accepted 13th December 2022

DOI: 10.1039/d2ra07184c

rsc.li/rsc-advances

1. *O*-GlcNAcylation and its analytical challenges

Glycosylation is the most abundant and diverse post-translational modification (PTM) on proteins, lipids and small RNAs.[1] *O*-GlcNAcylation occurs in the nucleus, cytoplasm and mitochondria,[2,3] and is a reversible glycosylation that is regulated only by the OGT (*O*-GlcNAc transferase) and OGA (*O*-GlcNAcase) enzymes. These enzymes transfer or hydrolyze GlcNAc from serine (Ser) or threonine (Thr) residues (Fig. 1), and GlcNAc is generally not extended by additional carbohydrates.[4,5] *O*-GlcNAcylation acts as a stress and nutrient receptor, 2–5% of glucose uptake produces UDP-GlcNAc *via* the hexosamine biosynthesis pathway (HBP), which is further enzymatically transferred to protein substrate *via* OGT catalysis.[6,7] *O*-GlcNAcylation regulates multiple important biological processes and physiological events, including epigenetics and transcription,[8,9] translation,[10] cell division,[11] metabolism,[12] signal transduction,[13] stress[14] and daily rhythms.[15] Thus, aberrant *O*-GlcNAcylation is often involved in the development of many chronic metabolic diseases.

O-GlcNAcylation remains a challenge due to its limited analytical tools and specificity. Although *O*-GlcNAc (*O*-linked β-D-*N*-acetylglucosamine) is a simple monosaccharide, there is a lack of reliable methods to enrich, quantify and localize *O*-GlcNAc proteins or peptides. This is mainly due to its low abundance, site ambiguity, low sensitivity and fragility. *O*-GlcNAcylation is sub-stoichiometric at each *O*-GlcNAc site and requires enrichment. Since GalNAc and GlcNAc are isomer, it is difficult to distinguish them by mass spectrometry (MS) alone. Furthermore, *O*-GlcNAc is labile when fragmented by collision-induced dissociation (CID), resulting in loss of *O*-GlcNAc from Ser or Thr.[16,17] To overcome these challenges, techniques have been developed in the past including antibodies,[18–21] chemoenzymatic tools[22–24] and other types of MS fragmentation techniques.[25,26] Nonetheless, methods with high specificity and sensitivity for clinical studies are limited, and there is an urgent need to develop a better assay for *O*-GlcNAcylation.

[a]*Center for Clinical Mass Spectrometry, College of Pharmaceutical Sciences, Soochow University, Suzhou, Jiangsu, 215123, China. E-mail: yangs2020@suda.edu.cn*
[b]*Suzhou Institute for Drug Control, Suzhou, Jiangsu, 215104, China*
[c]*Laboratory Medicine Center, Department of Clinical Laboratory, Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College, Hangzhou, Zhejiang, 310014, China*
[d]*Department of Orthopedics, The Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu, 215004, China*
[e]*Department of Endocrinology, The Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu, 215004, China*
[f]*Health Examination Center, The Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu, 215004, China. E-mail: luzhdoc@163.com*

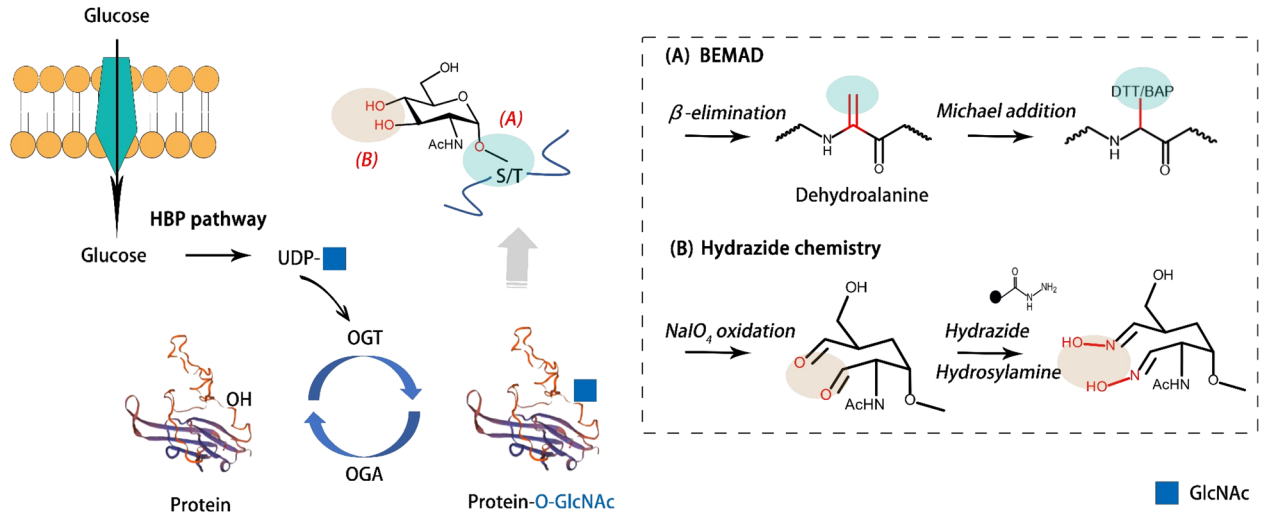

Fig. 1 Metabolic pathway for *O*-GlcNAcylation biosynthesis and analytical methods for enriching *O*-GlcNAc proteins. (A) BEMAD: after formation of dehydroalanine by β-elimination, alkene functionality is introduced into *O*-GlcNAcylated Ser/Thr residues by Michael addition. (B) Hydrazide chemistry: the *trans*-dihydroxyl group of GlcNAc is oxidized to aldehyde by sodium periodate at higher concentrations (20 mM).

Here, we mainly focus on two aspects: (1) updating existing methods for evaluating protein *O*-GlcNAcylation using chemical, physical, enzymatic and chemoenzymatic strategies; and (2) state-of-the-art of techniques for deciphering disease-associated *O*-GlcNAcylation. We hope to provide insights and future directions for *O*-GlcNAcylation analysis in the search for drug targets and biomarker discovery.

2. Analytical methods for studying protein *O*-GlcNAcylation

2.1 Direct chemical modification of GlcNAc

The existing methods for directly chemically modifying *O*-GlcNAc usually introduce alkene or aldehyde functional groups, and then utilize the properties of the introduced functional groups for subsequent enrichment, purification or enrichment. The approaches used for *O*-GlcNAc enrichment were summarized in Table 1.

2.1.1 β-Elimination and Michael addition of DTT (BEMAD). BEMAD can generate alkene on Ser or Thr *via* mild β-elimination (Fig. 1A), which is similar to the identification of phosphosites by elimination of phosphates in Ser/Thr and nucleophile attack on unsaturated carbonyl groups.[27–30] Compared to *O*-phosphate, *O*-GlcNAc is easier and more sensitive to the loss of β-elimination and is therefore suitable for *O*-GlcNAcylation analysis.[31,32] The formerly *O*-GlcNAc site on Ser/Thr is converted to 2-aminoacrylic acid by β-elimination, followed by addition of dithiothreitol (DTT) *via* Michael addition or biotin amylamine (BAP). These modifications are stable in CID tandem-MS fragmentation and allow site-specific identification by MS/MS or biotin affinity enrichment. However, phosphate, GalNAc,[33] glycosaminoglycan (GAG)[34] and cysteine residue[35] can also undergo β-elimination, resulting in low specificity of BEMAD.

2.1.2 *O*-GlcNAcylation enriched by hydrazide chemistry. Hydrazide chemistry has been widely used for the enrichment of *N*-glycopeptides.[36–38] While glycoproteins do not have aldehydes to react with hydrazides, glycans on proteins can be oxidized to form aldehydes by chemical or enzyme reagents such as sodium periodate (Fig. 1B). Periodate oxidation was most efficient on exocyclic and *cis*-diols; thus, *N*-glycopeptides are easily oxidized by sodium periodate *via* different monosaccharides (sialic acid, fucose, mannose, galactose, GalNAc and mannose).[39] Because GlcNAc contains *trans*-diol rather than *cis*-diol, the oxidation condition is more severe than for monosaccharide containing *cis*-diol. Oxidation of GlcNAc *trans*-diol at C3 and C4 requires higher temperatures and longer incubation times. Oxidized glycopeptides are typically captured by hydrazide resins and further released by hydroxylamine cleavage.[40] However, severe oxidation conditions result in N-terminal serine and threonine residues also being oxidized to their respective acetaldehyde derivatives, which are captured by hydrazide resins and released altogether during chemical cleavage of the glycosidic bond, thereby reducing selectivity. Notably, sodium periodate can also oxidize glycans with cis or *trans*-diols, such as cellulose,[41] glucose and cyclodextrins,[42,43] thereby reducing specificity to some extent. Therefore, when choosing this method to enrich *O*-GlcNAc glycopeptides, certain pretreatment is required, such as removing *N*-glycan, which can greatly improve the enrichment efficiency and accuracy.

2.2 Affinity enrichment of *O*-GlcNAcylation

2.2.1 Lectin affinity enrichment of *O*-GlcNAcylation. Lectins are non-immune polysaccharide-binding proteins with a wide range of biological activities at the cellular, tissue and organism levels. Wheat germ agglutinin (WGA) has a strong affinity for *O*-GlcNAc and is commonly used to detect *O*-GlcNAcylation (Fig. 2A). WGA has 4 functional binding sites, mainly through hydrogen bond interactions.[44] The WGA

Table 1 Comparison of analytical methods for enrichment of protein *O*-GlcNAcylation. There are many methods for enriching *O*-GlcNAcylated proteins/peptides, which are mainly divided into three categories: direct GlcNAc modification, affinity enrichment, selective capture and release (including carbohydrate metabolic labeling, enzymatic addition)

Methods		Pros	Cons	Ref.
Direct modification of GlcNAc	BEMAD	Mild reaction conditions; low cost	Low specificity	31 and 32
	Hydrazide chemistry	Solid-phase separation; simple sample preparation	Harsh reaction conditions; insufficient specificity; sample loss	40
Affinity enrichment	Lectin enrichment	Enrichment at the protein level; simple sample preparation; non-destructive	Large sample starting volume; insufficient specificity (but stronger than the above two methods); high cost	47 and 49–54
	Antibody enrichment	Enrichment at the protein level; visualization is possible; higher specificity and sensitivity; non-destructive	Protein sequence dependent; high cost; low affinity	19, 55, 56, 59 and 60
	Hydrophilic enrichment	High specificity and efficiency; non-destructive; suitable for larger size samples	Specially formulated hydrophilic enrichment material; high cost	63 and 64
Carbohydrate metabolic labeling		Suitable for live-cell analysis	Incomplete labeling	69–72 and 75–79
Enzymatic addition		High specificity and efficiency; flexible and changeable; high compatibility with other methods	Complex process; high cost	22, 23, 74, 82–87, 91–93, 96–98, 101, 104, 105 and 180

binding sites consist of amino acids of different polypeptide chains involved in ligand binding, in addition to GlcNAc, also recognize *N*-acetylneuraminic acid (Neu5Ac)[45] and GalNAc.[46] Therefore, WGA can enrich for Neu5Ac and GalNAc peptides. If the oligosaccharide has multiple GlcNAcs instead of a single GlcNAc, the multiple binding sites of WGA make it more

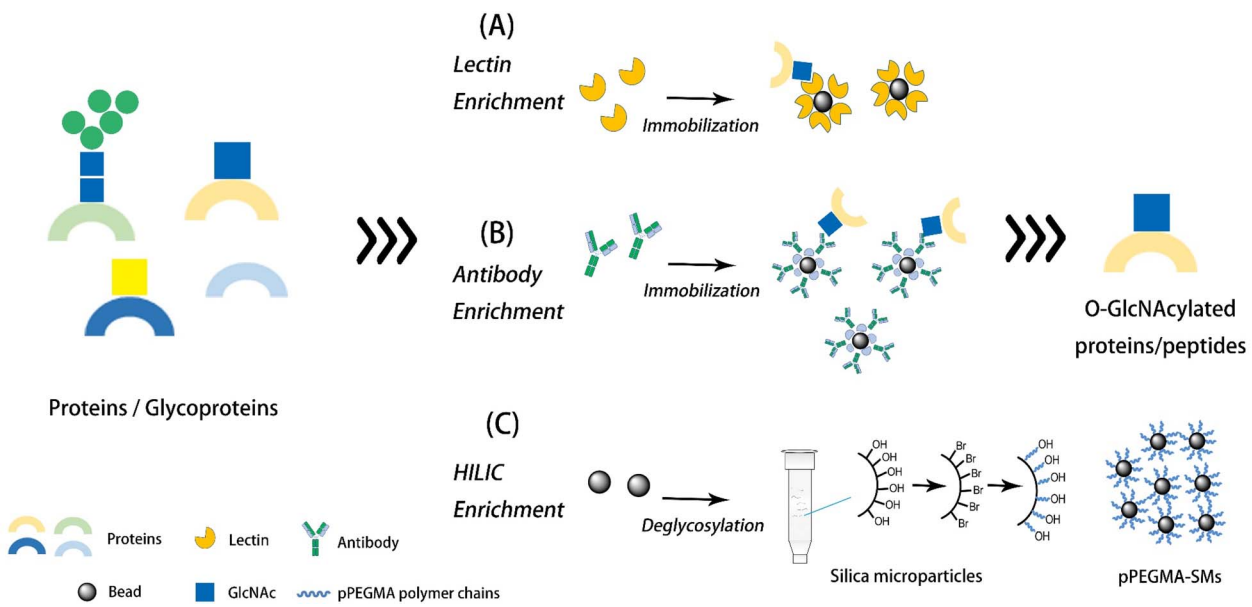

Fig. 2 Different strategies for targeted analysis of *O*-GlcNAc proteins or peptides. (A) Lectin enrichment: WGA captures *O*-GlcNAc and removes other types of glycosylation. (B) Antibody enrichment: *O*-GlcNAc proteins or peptides are enriched by *O*-GlcNAc antibody. (C) HILIC method: *O*-GlcNAc peptides can be enriched by polymer-modified hydrophilic stationary phase upon removal of *N*-glycans, Tn antigens and mucin-type *O*-glycans *via* enzymatic deglycosylation. pPEGMA-SMs: hydrophilic polymer modified-silica microparticles.

active.[47] Due to the negative charge of the lectin, succinylated WGA (sWGA) hardly binds to sialoglycoproteins,[48] thereby improving the specificity of *O*-GlcNAc, although its affinity is not as good as that of WGA.[49]

Lectins can be used for separation of *O*-GlcNAc peptides in liquid chromatography (LC). Vosseller *et al.* demonstrated that lectin (WGA) weak affinity chromatography (LWAC) successfully separated *O*-GlcNAc peptides, which eluted later than peptides.[50] Due to the different chromatographic properties between sialic acid, GalNAc and GlcNAc, LWAC can also separate these different glycoforms. *Griffonia simplicifolia* lectin II (GSL II) recognizes *O*-GlcNAc with lower affinity and preferentially recognizes highly branched *N*-glycans.[51] On the other hand, *Psathyrella velutina* lectin (PVL) is specific for antenna GlcNAc, and recombinant *Psathyrella velutina* lectin (rPVL) has a stronger affinity for GlcNAc than sialic acid.[52] AANL (*Agrocybe aegerita* lectin 2) has a similar affinity to rPVL and primarily interacts with *N*-glycans and terminal GlcNAc.[53] Rambaruth *et al.* also found that helix pomatia agglutinin (HPA) can bind to a range of carbohydrate moieties, including GlcNAc and GalNAc.[54]

2.2.2 *O*-GlcNAcylation-specific antibody enrichment. Antibodies that detect *O*-GlcNAc can be used in several applications, including western blot (WB), immunoprecipitation and flow cytometry (Fig. 2B). Pan-*O*-GlcNAc includes IgM monoclonal antibody (mAb) (CTD110.6)[55] and IgG mAb (RL2).[19] CTD110.6 and RL2 can recognize *O*-GlcNAc proteins, but they have lower affinity and can only detect proteins with multiple *O*-GlcNAcylation sites or higher abundance.[56] CTD110.6 can also bind to diacetylchitobiose[57] and GlcNAcβ1,4Man-Thr/Ser on α-dystroglycan.[58] Four other antibodies have been reported, IgG mAbs (18B10.C7(3), 9D1.E4(10), 1F5.D6(14), HGAC85) and IgM mAb (10D8). Tashima *et al.* used several antibodies to detect *O*-GlcNAc modifications on the extracellular domains of membrane or secreted glycoproteins. They found that CTD110.6, 18B10.C7(3) and 9D1.E4(10) can detect *O*-GlcNAc modified proteins on the cell surface.[59] Additional novel antibodies are being investigated to improve the sensitivity and specificity of *O*-GlcNAcylation. Burt *et al.*[60] evaluated an anti-*O*-GlcNAc antibody with 4 novel rabbit mAbs *via* protein-A beads, and they identified over 1300 *O*-GlcNAc peptides and over 1000 sites in murine synapse. Compared with other metabolic labeling strategies, anti-*O*-GlcNAc Abs can precipitate native *O*-GlcNAc-modified polypeptides from tissues and cells with higher sensitivity and specificity.

2.2.3 Hydrophilic enrichment of *O*-GlcNAc. Hydrophilic Interaction Chromatography (HILIC) can efficiently retain and separate glycopeptides through hydrophilic interactions between the glycan moiety and the stationary phase (Fig. 2C).[61,62] *N*-Glycopeptides and mucin-type *O*-glycopeptides are generally more abundant and hydrophilic, and can be bound by HILIC more efficiently than *O*-GlcNAc peptides. Shen *et al.* used selective enzymatic digestion and hydrophilic microparticles to identify *O*-GlcNAc proteins.[63] To improve the enrichment performance, the method includes two steps: (1) removal of *N*-glycans and *O*-glycans by PNGase F, sialidase and *O*-glycosidase, (2) novel HILIC materials hydrophilic polymer modified-silica microparticles (pPEGMA-SMs) synthesized by surface-initiated atom transfer radical polymerization (SI-ATRP) for highly specific binding to the *O*-GlcNAc peptide. This material mainly uses poly(ethylene glycol) methacrylate (PEGMA) to synthesize hydrophilic stationary phases, and the dense accumulation of hydroxyl groups along the PEGMA polymer brushes on the surface provides much higher hydrophilicity and strong hydrogen bond interaction compared to traditional HILIC material. Analysis of human urinary *O*-GlcNAc proteins by this approach revealed that 474 *O*-GlcNAc peptides were identified from 457 proteins. While this method is promising, it fails to distinguish *O*-GalNAc (*O*-linked β-*N*-acetylgalactosamine) peptides that are highly abundant in tumor tissue. Recently, nitroxide-grafted nanospheres showed substantially enhanced hydrogen-bonding interactions in *O*-GlcNAc analysis. Site-specific analysis of *O*-GlcNAc peptides from PNAC-1 cells identified an additional 197 *O*-GlcNAc sites from 183 glycoproteins, indicating high-throughput identification of *O*-GlcNAcylation.[64]

2.3 Selective capture and release of protein *O*-GlcNAcylation by metabolic labeling and cleavage

2.3.1 Carbohydrate metabolic labeling. Metabolic oligosaccharide engineering (MOE) can introduce subtle modifications into the monosaccharide residues of cellular glycans. Through cellular biosynthetic pathways, synthesized monosaccharide analogs containing biorthogonal functions (alkynes or azides) are converted into activated nucleotide sugars, which are transferred to integrate the synthetic monosaccharide into the structure.[65–67] Vocadlo *et al.* pioneered a method to study *O*-GlcNAc proteomics in living cells using peroxyacetylated *N*-azidoacetylglucosamine (Ac4GlcNAz). Ac4GlcNAz has excellent cell membrane penetration efficiency, and the enzymes of the GlcNAc salvage pathway are tolerant to its azide moiety and can be metabolized into nucleoproteins by OGTase.[68] Nandi *et al.* used a tagging-*via*-substrate (TAS) approach to identify 199 *O*-GlcNAc proteins from HeLa cells by enhanced Staudinger ligation between the azide group of Ac4GlcNAz and the phosphine probe.[69] Since the pyrophosphorylase step of the GlcNAc salvage pathway is rate-limiting for UDP-GlcNAc biosynthesis and GlcNAz may be involved in *N*-glycan synthesis, this may lead to low GlcNAc labeling and poor selectivity.[70] The mammalian enzyme UDP-galactose-4′-epimerase can convert UDP-GlcNAc to its isoform UDP-GalNAc, exploiting this metabolic crosstalk, Ac4GalNAz probe was developed to label *O*-GlcNAc acylation protein.[71] However, interconversion between GlcNAz and GalNAz in cells results in indistinguishability from mucin glycoforms.[72,73] The alkynyl-modified GlcNAc analog (GlcNAlk) does not interconvert to GalNAlk is therefore more specific to *O*-GlcNAc labeling.[74]

Other analogs have been investigated to improve *O*-GlcNAc metabolic labeling. 6-Azido-6-deoxy-*N*-acetylglucosamine (6AzGlcNAc) is directly phosphorylated by phosphoacetylglucosamine mutase (AGM1), which bypasses the canonical salvage pathway and is a specific metabolic chemical reporter (MCR) for *O*-GlcNAc proteins. It is dynamically integrated into

the *O*-GlcNAc protein, but its removal from the protein depends on the activity of OGA, thus diminishing metabolic labeling efficiency.[75] 1,3,6-Tri-*O*-acetyl-2-azidoacetamido-2,4-dideoxy-D-glucopyranose (Ac34dGlcNAz) is resistant to OGA enzymatic hydrolysis due to the absence of a hydroxyl group at C4, leading to significantly improved metabolic labeling.[76] 4-Deoxy-4-fluoro-GalNAz (4FGalNAz) replaces the hydroxyl with fluorine at C4, which is more selective without interfering with the endogenous glycosylation metabolic pathway.[77] The analog (GlcNDAz) introduces a diazine-based photo-crosslinker into the *O*-GlcNAc modification in cells to form a covalent bond with its binding counterpart.[78] Although MOE cannot fully label *O*-GlcNAc due to competition from natural sugars, it is still an excellent method to analyze *O*-GlcNAc proteins in living cells.

2.3.2 Enrichment of *O*-GlcNAcylation by enzymatic addition

Galactosyltransferase-catalyzed addition

UDP-[^{3}H]-galactose labeling. Galactosyltransferase (GalT) is the most widely studied glycosyltransferase in mammals, and β1,4-GalT catalyzes the transfer of UDP-[^{3}H]-Gal to β1,4-GlcNAc (Fig. 3A-a), forming a Gal–GlcNAc.[79,80] The radiolabeled Gal is added to the GlcNAc residue of the target protein by β1,4-GalT and further detected by autoradiography. After β-elimination, the released disaccharide can be analyzed to determine whether the substrate carries a single GlcNAc. This strategy facilitates the detection of *O*-GlcNAc proteins and has been used to probe cell surface *O*-GlcNAcylation, and identify GlcNAcylation.[2,81–83] However, tritium labeling is not as sensitive as other radioactive elements, and the handling of radioactive elements is expensive; the process is time-consuming, as autoradiography can take days to months to detect a signal.[22]

Galactose oxidase for hydrazide conjugation. Addition of a *cis*-diol monosaccharide to GlcNAc allows identification of *O*-GlcNAc proteins using hydrazide chemistry for identification of *O*-GlcNAc proteins (Fig. 3A-b). Following the transfer of Gal to *O*-GlcNAc *via* GalT, Nishikaze *et al.* used a triethylbenzaldehyde probe to form aromatic aldehyde derivatives *via* a copper-catalyzed Huisgen 1,3-cycloaddition "click reaction" in which *O*-GlcNAc peptides were enriched by hydrazide resin with negligible side reactions.[84] Gal also can be oxidized to form two aldehydes that can react to hydrazide. The Gal C6-OH can be selectively oxidized by highly specific galactose oxidase (GAO).[23,85] However, the activity of GAO is unstable and side reactions frequently occur, which significantly reduce its labeling efficiency. In contrast, variant F2 of the GAO enzyme is active on GlcNAc and can introduce aldehydes directly into GlcNAc.[86] The *cis*-diol of Gal can also be oxidized by sodium periodate at a concentration of 1 mM,[87] allowing enrichment by hydrazide chemistry.

GalT Y289L combined with click chemistry. A new chemoenzymatic approach has been developed to detect *O*-GlcNAc proteins (Fig. 3A-c). GalT Y289L forms a cavity in the catalytic pocket, thereby enhancing its catalytic activity towards GalNAc with similar specificity.[88] β1,4-GalT is enzymatically active when its C2-, C3- or C6-OH is substituted with an azido or amide

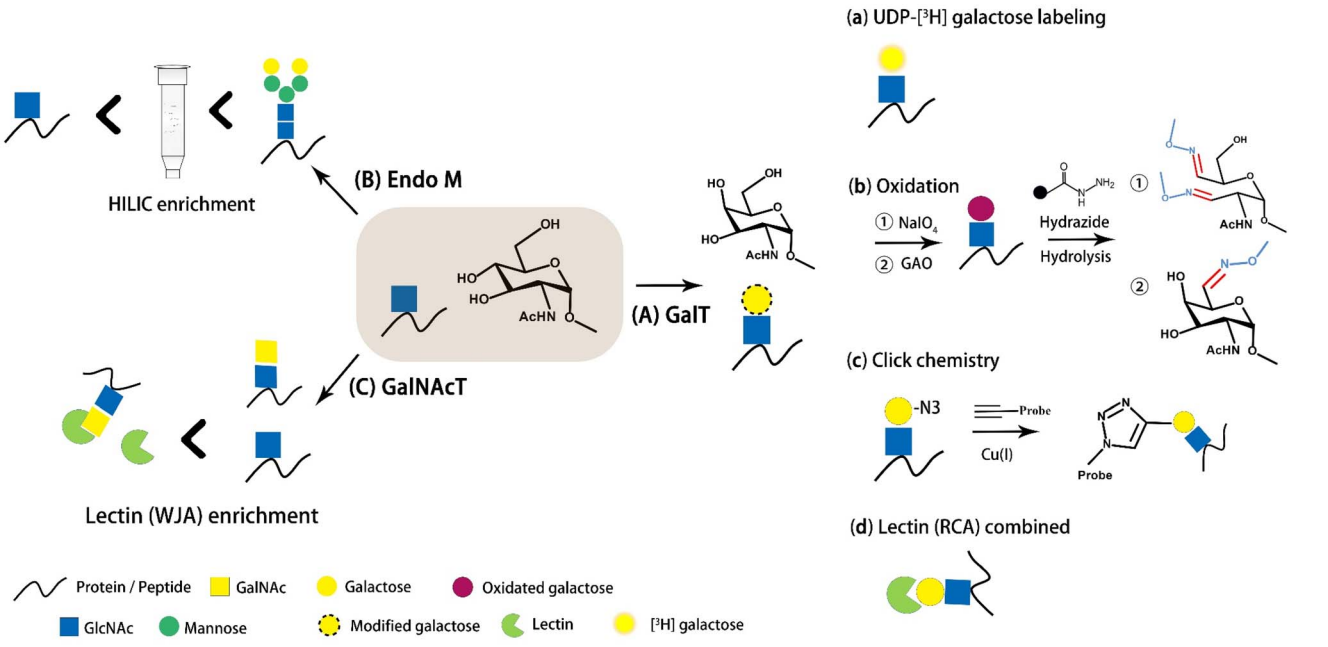

Fig. 3 Combination of enzymatic modification of *O*-GlcNAc with other approaches for specific capture-release of *O*-GlcNAcylation. (A) GalT (galactosyltransferase addition). The subsequent enrichment is mainly through galactose. (a) Isotope ^{3}H-labeled galactose is incorporated into *O*-GlcNAc and detected by autoradiography. (b) The *cis*-diol and 6-OH of galactose are oxidized by sodium periodate ① and galactosidase (GAO) ②, respectively. The aldehydes are produced and reacted to hydrazide resin for *O*-GalNAcylated peptide enrichment. (c) The azido-modified galactose is bound to the probe through click chemistry. (d) The Gal–GlcNAc disaccharide is enriched by RCA lectins. (B) Endo-M modification: extended glycan chain of *O*-GlcNAc can be enriched by HILIC, and the extended glycan is then cleaved by wild-type Endo-M/S. (C) GalNAcT (GalNAc transferase modification): the GalNAc–GlcNAc structure can be enriched by RCA lectins.

derivative.[89,90] UDP-Gal analogs were designed by replacing a keto[22] or azide[91] as a linker, which can be conjugated to the *O*-GlcNAc peptide *via* an alkyne-derivatized probe (biotin). After biotinylation, *O*-GlcNAc peptides were detected by chemiluminescence using streptavidin and horseradish peroxidase (HRP). Using this approach, Khidekel *et al.* identified 34 unique *O*-GlcNAc peptides among 25 glycoproteins from the mouse brain.[92] Wu *et al.* developed a technique for the detection of *O*-GlcNAc using GalT Y289L in conjunction with ST6Gal1. The *O*-GlcNAc moiety was first galactosylated and then sialylated with a sialic acid for highly sensitive fluorescence detection. This method requires removal of *N*-glycans and mucin type *O*-glycans.[93,94]

To improve the recovery of enriched *O*-GlcNAc peptides, cleavable probes have been developed in *O*-GlcNAcylation assays. Zaro *et al.* used chemically cleavable biotin-azide probes and identified 374 proteins without site-specific information.[95] Due to its increased sensitivity, a photochemically cleavable biotin probe was designed to label *O*-GlcNAc peptides using the biotinylated reagent PC-PEG-biotin-alkene. Glycopeptides were captured by affinity chromatography and then released after photocleavage by exposure to UV (365 nm). In this way, Wang *et al.* discovered 8 *O*-GlcNAc sites in rat brain[96] and Alfaro *et al.* detected 274 *O*-GlcNAc proteins in mouse cerebral cortical brain tissue.[97] Ma *et al.* further simplified the method by combining chemoenzymatic labeling with copper-free click chemistry.[98]

Lectin combined (Ricinus communis agglutinin). *Ricinus communis* agglutinin (RCA) is widely used as a tool to study cell surfaces and purify glycans with high affinity and specificity for glycan conjugates containing terminal β-D-Gal residues[99,100] (Fig. 3A-d). Haynes *et al.* used GalT to introduce terminal Gal into *O*-GlcNAc and selectively enriched labeled glycopeptides by delayed elution through an RCAI affinity columns, showing good selectivity, sensitivity and accuracy.[101]

Endo-M based trans-glycosylation. GalT1 is the only glycosyltransferase used to label *O*-GlcNAc proteins, whereas Endo-M N175Q is the second chemoenzyme to label *O*-GlcNAc glycopeptides other than GalT1 mutant (Fig. 3B). GalT-based method may introduce cumbersome tags into glycopeptides, complicating MS/MS spectra and compromising identification efficiency. To prevent this obstacle, Chen *et al.* developed a reversible label-free method for glycopeptide analysis using Endo-M mutants-HILIC.[24] Endo-M transfers *N*-glycan oligosaccharides to receptors, whereas its N175Q mutant has significantly enhanced activity of sugar synthesis enzymes due to reduced hydrolysis when the catalytic site of N175 is replaced by glutamine (Gln).[102] Endo-M N175Q has *trans*-glycosidase-like activity on the innermost GlcNAc residue of *N*-glycans through which released oligosaccharide was transferred to *O*-GlcNAc.[103] Transferring an oligosaccharide to *O*-GlcNAc significantly increases its hydrophilicity, allowing enrichment by HILIC. *O*-GlcNAc peptides are identified after removing oligosaccharide by wild-type (WT). Using this strategy, 657 *O*-GlcNAc proteins were identified using only 400 μg of HeLa nucleoproteins.

Enhancement of lectin affinity by adding GalNAc to O-GlcNAc. Several lectins show excellent affinity for GalNAc, so adding GalNAc to GlcNAc is beneficial (Fig. 3C). β4-GalNAc transferase (GalNAc-T) catalyzes UDP-GalNAc to *O*-GlcNAc substrate. Abo *et al.* converted GlcNAc to GalNAc-GlcNAc using soluble GalNAc-T expressed in the nucleus or cytoplasm and captured the GalNAc moiety using wisteria lectin (WJA).[104] They found a new *O*-GlcNAc site at S263 of the transcription factor Sox2. GalNAc–GlcNAc binds to WJA with a stronger K_a, $\sim 1.4 \times 10^5\ M^{-1}$, and its affinity is more sensitive than the WGA lectin.[105] WJA affinity is independent on the amino acid sequence of the *O*-GlcNAc modification site and thus enriches GalNAc–GlcNAc peptides better than amino acid-dependent methods (antibody assay). Notably, GalNAc–GlcNAc is only present in pituitary glycoprotein hormones and is not detected in cells and tissues, thereby minimizing false-positive identifications.[106,107] Since the GalNAc-T labeling must be expressed in cells, it cannot be used to study pathological specimens. WJA lectins have affinity for other glycan structures containing terminal GalNAc residues. Addressing these limitations could further improve its specificity.

3. Quantitative analysis of protein *O*-GlcNAcylation

3.1 Mass spectrum quantification of *O*-GlcNAcylation

3.1.1 Stable isotope labeling. The MS and non-MS approaches used for *O*-GlcNAcylation quantification are summarized in Table 2. Stable isotope labeling is the most widely used method in quantitative proteomics. Relative quantitative proteomics can be achieved through *in cellulo* metabolic labeling and *in vitro* chemical labeling (Fig. 4A-①). Stable isotope labeling of amino acids in cell culture (SILAC) is a cellular approach used to study protein synthesis and turnover.[108] SILAC labels amino acids on lysine or arginine in the medium by C13/N15, followed by quantification of cultured cells or organisms by tandem mass spectrometry. Zhong *et al.* assessed and quantified the global phosphorylated proteome of OGT-WT and null cells, demonstrating that OGT depletion resulted in *O*-GlcNAc-induced changes in the phosphorylated proteome, which are primarily associated with DNA damage response and cell cycle regulation.[109] Combining SILAC with a photocleavable probe, Wang *et al.* identified 141 new *O*-GlcNAc sites in HeLa cells and quantified *O*-GlcNAc proteins in OGT-overexpressing HeLa cells.[11] Overall, the SILAC method is quantitatively accurate and simple in sample preparation, and convenient for use in combination with other methods. However, SILAC is only applicable to cultured cells or organisms and cannot be applied to tissues and biofluids.

In vitro quantification (IVQ) can overcome the limitations of the SILAC method. IVQ uses several functional groups such as lysine, carboxyl or sulfhydryl to introduce stable isotope tags (Fig. 4A-②). The most common tags include iTRAQ[110,111] (Multiplexed Isobaric Tagging Technology for Relative Quantitation), TMT (Tandem mass tags)[93] or non-isobaric dimethyl labeling.[112,113] Quantitative isotope and chemoenzymatic tagging (QUIC-Tag) combine selective chemoenzymatic reagents with an isotopic labeling strategy for monitoring and

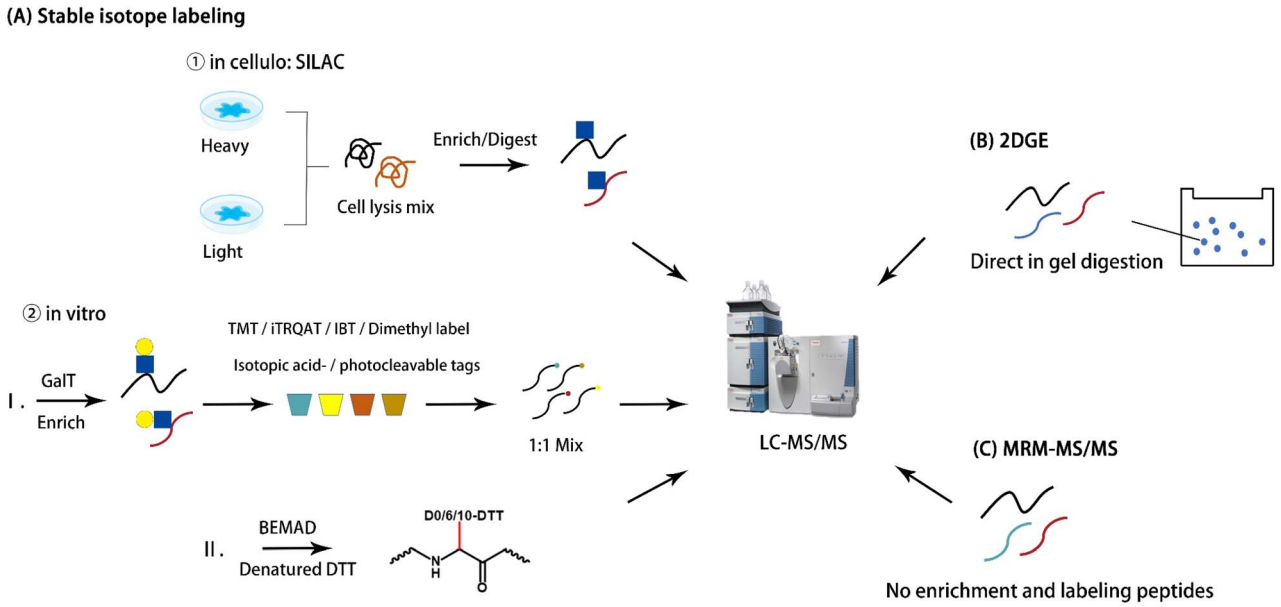

Fig. 4 Different methods for quantitative analysis of *O*-GlcNAcylation using mass spectrometry-based proteomics. (A) Stable isotopic labeling ① in cellulo and ② *in vitro* labeling: SILAC labels *O*-GlcNAc proteins in cellulo and are further enriched for LC-MS/MS quantification; *in vitro* quantification is achieved with isobaric tags (TMT, iTRAQ, IBT *etc.*); *O*-GlcNAc peptides can be labeled by stable-isotope with deuterium. (B) 2D-LC: the *O*-GlcNAc proteins separated by two-dimensional (2D) gel electrophoresis are directly digested in the gel for LC-MS quantification. (C) MRM-MS/MS: multiple reaction monitoring is widely used for quantification on known *O*-GlcNAc peptide.

quantifying *O*-GlcNAc proteins in neuronal cells.[113] QUIC-Tags transfers ketone-containing Gal analogs to GlcNAc residues *via* GalT Y289L and purifies tagged biotinylated proteins by avidin chromatography. They found that *O*-GlcNAcylation was reversible and cycled rapidly within neurons. Compared to dimethyl labeling, iTRAQ/TMT can process multiple samples simultaneously and improve identification, although it may suffer from background interference due to co-eluting peaks of similar sized precursor ions and reporter impurities.[114]

As another unique method for incorporating stable isotopes, β-elimination and Michael addition of DTT (BEMAD) can easily acquire the site information of *O*-GlcNAcylation. Deuterated DTT (D0-DTT, D6-DTT, D10-DTT) was used to introduce isotope labeling on *O*-GlcNAcylated Ser/Thr to compare *O*-GlcNAc site specificity and occupancy across different samples.[111,115] In contrast to the aforementioned methods for quantifying *O*-GlcNAc at the peptide level, Qin *et al.* developed an isotope-tagged cleavable linker (isoTCL)[116] that incorporates isotope tags directly into the *O*-GlcNAc moiety for quantification and identified more than 100 *O*-GlcNAcylated sites in male and female mouse placentas. In an optimized strategy termed isotopic photocleavable tagging for *O*-GlcNAc profiling (isoPTOP) developed by Liu *et al.*,[117] they identified and quantified approximately 1000 *O*-GlcNAc sites in colorectal cancer cell lines (SW480, SW620), of which many are involved in tumor progression and metastasis. In conclusion, *in vitro* chemical labelling quantification methods are flexible, applicable to a variety of samples, and have high accuracy, but are less effective and specific than *in cellulo* metabolic labeling.

3.1.2 Two-dimensional gel electrophoresis (2DGE). 2DGE-MS is a classic method for quantitative analysis of protein and can separate *O*-GlcNAc proteins for subsequent MS analysis (Fig. 4B). Typically, gel-bands of interest are selected for bottom-up proteomics. 2DGE-MS has some limitations on analysis of low or large molecular weight of proteins, or highly acidic/basic proteins due to their poor solubility. This can be overcome by narrowing the linear dynamic range and pH range, but it does not fully quantify low-abundant proteins with satisfactory performance.[118]

3.1.3 Label-free *O*-GlcNAc peptide quantification. Isotopically or isobarically labelled MS quantification is relatively time-consuming, costly, and incompletely labeled, which has led to the development of label-free quantification (LFQ) for easy and quick quantification (Fig. 4C). LFQ includes peptide peak intensities and multiple reaction monitoring (MRM).[119] Because of the low stoichiometry of *O*-GlcNAc peptide, it can be quantified by intensity but is labor-intensive. MRM uses a MS/MS scanning in a triple quadrupole,[120] and Yuzwa *et al.* used this method to detect three *O*-GlcNAc sites on Tau.[121] However, the properties of MRM-MS for the detection of *O*-GlcNAc peptides are not exhaustive. Maury *et al.* were the first to demonstrate that MRM-MS can be used to analyze native *O*-GlcNAc peptides in complex mixtures.[122] They discovered a novel peptide GSK-3β with three potential *O*-GlcNAc sites in human embryonic stem cells. The detection limit of MRM-MS for standard *O*-GlcNAc peptides is as low as 3 fmol, and the detection of complex samples does not require labeling and enrichment. However, since the retained fragment ions of *O*-GlcNAc cannot be distinguished from background noise, MRM cannot directly

determine the exact sites of *O*-GlcNAc modification. It is difficult to quantify *O*-GlcNAc using MRM-MS due to lack of peptide sequence information or unknown samples.[123]

3.2 Non-MS quantitative methods

3.2.1 Flow cytometry. Flow cytometry can sort cells of interest and serve as a quantitative analysis tool for intracellular *O*-GlcNAc signaling (Fig. 5A). It characterizes changes in neutrophil *O*-GlcNAcylation after cells are treated with agonist-inducing agents.[124] The drug-treated cells were fixed with 3% formaldehyde/PBS, permeabilized with 90% methanol and blocked with 1% casein/PBS, followed by incubation with anti-*O*-GlcNAc antibody (CTD110.6). Finally, the cell suspension was prepared by incubating with Alexa Fluor 488-labeled fluorescent secondary antibody. The signal was acquired by a FACSCalibur to record the fluorescence intensity of *O*-GlcNAc cells. The results showed that an increase in *O*-GlcNAc proteins was detected within hours after stimulation of neutrophils with 100 nM fMLF (fMet–Leu–Phe) or 10 nM PMA (phorbol myristate acetate). Compared with immunoblot or immunofluorescence microscopy, flow cytometry greatly reduces the time for staining, data acquisition and analysis, can analyze large numbers of cells, and is a fast semi-quantitative method.

3.2.2 GalT in combination with metabolic labeling

PEG tag. Polyethylene glycol (PEG) tag can visualize a global subset of *O*-GlcNAc proteins in complex samples. It is biologically inert as a biotag and has a defined molecular weight. The PEG tag is selectively incorporated into the *O*-GlcNAc moiety of the protein by chemoenzymatic and biorthogonal reactions, resulting in a slower migration rate of the targeted protein in sodium dodecyl sulfate-polyacrylamide (SDS-PAGE), which can be further quantified with WB (Fig. 5B). UDP-2-ketogalactose can label the GlcNAc moiety *via* GalT (Y289L), which is then reacted with an aminooxy-functionalized PEG mass tag.[125,126] Darabedian *et al.* developed an optimization strategy based on copper-free click chemistry (SPAAC) using *N*-azidoacetylgalactosamine (GalNAz) instead of keto-galactose.[127] PEG-based quantification does not require advanced instrumentation, and radioactive tags are inexpensive without protein purification. Importantly, PEG tag can also determine the relative occupancy of protein *O*-GlcNAcylation, which cannot be determined by MS/MS analysis.

Glyco-Seek. Glyco-Seek is an *O*-GlcNAc protein quantification method developed by Robinson *et al.*[128] It combines UDP-

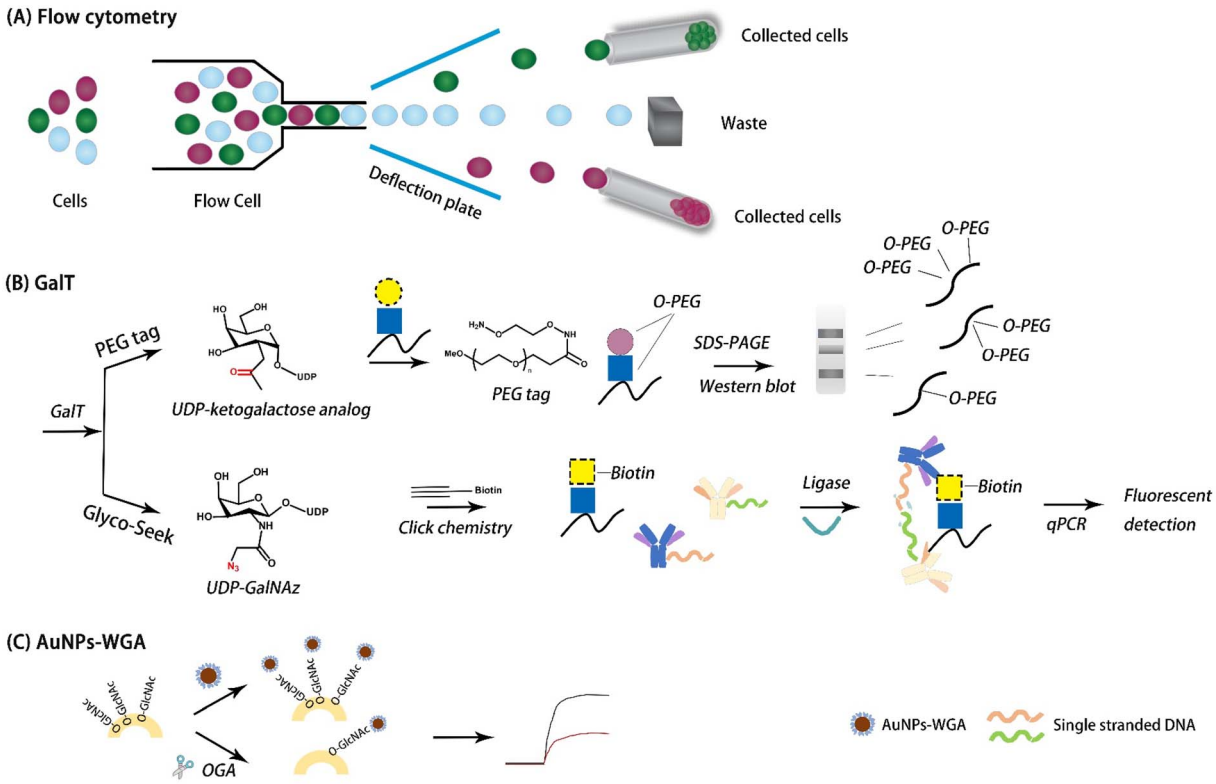

Fig. 5 *O*-GlcNAcylation quantification by non-mass spectrometry methods. (A) Flow cytometry: the cells in the *O*-GlcNAc induction and control groups are sorted by flow cytometry, and the intracellular *O*-GlcNAc abundance is evaluated by the acquired density dot map. (B) GalT: PEG labeling – UDP-Ketogalactose analog is enzymatically transferred to *O*-GlcNAc proteins that are further conjugated with PEG tag. The western blotting (WB) can directly perform semi-quantification of *O*-GlcNAc proteins. Glyco-Seek – GalNAz is incorporated into *O*-GlcNAc for biotinylation and incubate with antibody–DNA conjugates targeting biotin and the protein of interest, respectively. Two single-stranded complementary DNA hybrids are detected by qPCR. (C) Two sets of *O*-GlcNAc proteins treated with/without OGA digestion are enriched with AuNPs-WGA and thereby the signal difference is used for *O*-GlcNAcylated protein quantification.

GalNAz labeling, proximity ligation assay (PLA) and quantitative polymerase chain reaction (qPCR). First, GalNAz was transferred to *O*-GlcNAc protein by chemoenzymatic reaction and labeled with alkynyl biotin by click chemistry. Second, the PCR amplicon is divided into two halves of ssDNA, which are covalently linked to antibodies that bind the target protein and biotinylated *O*-GlcNAc, respectively, thereby binding the target protein to the PCR amplicon produced by PLA. Glyco-Seek utilizes the exponential amplification of nucleic acids to greatly improve the detection sensitivity of low-abundance *O*-GlcNAc, which is 3 orders of magnitude higher than WB. The method is non-destructive compared to MS, which relies on protein digestion and MS/MS fragmentation. It has good sensitivity and low sample consumption for the detection of low-abundance glycoproteins. Unfortunately, Glyco-Seek is impractical for *de novo* discovery of previously unknown proteins, but can be used as a complement to other methods to provide more sensitive analysis of target proteins.

3.2.3 AuNPs/WGA. Gold nanoparticles (AuNPs) have good biocompatibility and high catalytic activity, which can be used for signal amplification, thereby improving detection sensitivity (Fig. 5C).[129,130] Liu *et al.* developed a colorimetric strategy for non-enzymatic signal amplification based on the recognition between WGA and *O*-GlcNAc and copper deposition catalyzed by AuNPs.[131] The biotinylated lectin was used as a probe to bind *O*-GlcNAc proteins and incubated with AuNP-streptavidin tags to catalyze the deposition of copper on the lectin surface, followed by oxidation by $FeCl_3$ to generate a color reaction. Thus, the level of *O*-GlcNAcylation can be qualitatively assessed by visual inspection or quantitatively detected by absorbance measurement. The method does not use sensitive and expensive natural enzymes, and has the advantages of low cost, simplicity and stability. However, since WGA also recognizes other GlcNAc terminal carbohydrates, its specificity for *O*-GlcNAcylation recognition is reduced.[103] Gao *et al.* designed a sensitive surface plasmon resonance (SPR) biosensor based on AuNPs/WGA and OGA.[132] SPR recorded the interaction between AuNPs/WGA and *O*-GlcNAc, which was detected by the difference in signal between OGA-treated and untreated samples. Accuracy is greatly improved by avoiding interference from other glycans and non-specific adsorption of AuNPs.

4. Regulation of abnormal protein *O*-GlcNAcylation in diseases

4.1 Diabetes-associated *O*-GlcNAcylation

Diabetes is a metabolic disorder that includes type I (insufficient insulin secretion) and type II (insulin insensitivity).[133,134] Diabetes is characterized by hyperglycemia in the blood,[135] and glucose metabolism through the hexosamine synthesis pathway (HBP) is associated with many adverse effects of hyperglycemia (Table 3). UDP-GlcNAc is the final product of the HBP pathway and its level is proportional to cellular glucose flux.[6] Excess glucose flux through HBP leads to elevated *O*-GlcNAcylation, which mediates multiple biological processes, including insulin signaling, β-cell regulation, glycogen synthase, and tissue damage. For example, under hyperglycemic conditions, increased *O*-GlcNAc modification of insulin receptor substrate (IRS)-1/2 and Akt reduces their phosphorylation, leading to desensitization of insulin signaling to induce insulin resistance and decrease insulin-dependent glucose uptake.[136–138] Pancreatic β-cells regulate insulin storage and secretion, while HBP affects its development and differentiation. Hyperglycemia increases *O*-GlcNAcylation of glycogen synthase, thereby inhibiting its activity, leading to blood glucose retention and the formation of a glucotoxic cycle.[139,140] Meanwhile, high *O*-GlcNAcylation reduces muscle sensitivity to calcium, thus modulating diabetes-related muscle atrophy.[141] *O*-GlcNAc may impair angiogenesis in endothelial cells by inhibiting Akt signaling, leading to complications of retinal vasculopathy.[142–144]

4.2 *O*-GlcNAcylation involved in neurodegenerative diseases

O-GlcNAc modifications are most abundant in the brain and play key roles in regulating various neurological functions,[145] their levels varying with pathophysiological states. Alzheimer's disease (AD) patients have 22–50% lower levels of *O*-GlcNAcylation than healthy individuals.[146] *O*-GlcNAc-modified tau protein and amyloid-β (Aβ) plaques are two pathological hallmarks of AD.[147] Hyperphosphorylated tau protein with less *O*-GlcNAcylation loses its microtubule-binding ability, exhibits prion-like properties, and aggregates to form neurofibrillary tangles (NFTs). Conversely, increased *O*-GlcNAcylation prevents its aggregation by competing with phosphorylation sites on tau protein.[148,149] Amyloid precursor protein (APP) is transferred to endosomes by endocytosis and cleaved by γ-secretase to form insoluble β-amyloid, which further aggregates to form Aβ plaques.[150,151] APP *O*-GlcNAcylation at T576 prevents endocytosis, promotes non-amyloid formation and reduces Aβ production.[152] These results suggest a neuroprotective effect of *O*-GlcNAcylation on the brain.

Parkinson's disease (PD) is the second most common neurodegenerative disease, and the main hallmark of PD is Lewy bodies (LBs) formed by abnormal aggregation of proteins. α-Synuclein is a major component of LBs and studies have shown that it is modified by *O*-GlcNAc.[153–156] *O*-GlcNAcylation of α-synuclein inhibits its phosphorylation-induced pathological aggregation and reduces neuronal toxicity and death.[157–159] In PD model mice, elevated *O*-GlcNAcylation levels significantly attenuated PD pathology in dopamine neurons.[160] However, studies have shown that postmortem temporal cortex samples from PD patients have higher levels of *O*-GlcNAcylation than normal individuals. Excessive *O*-GlcNAcylation might inhibit autophagic and induce α-synuclein aggregation.[161] In conclusion, the regulatory mechanism of *O*-GlcNAc in PD has not been fully characterized, and targeting *O*-GlcNAcylation may lead to find therapy for PD treatment, but more comprehensive *O*-GlcNAc proteomic studies are required to achieve this goal.

4.3 Cancer-induced *O*-GlcNAcylation dysregulation

Metabolic alterations and energy dysregulation are hallmarks of cancers, and the nutrient-driven *O*-GlcNAcylation is highly

variable in cancer.[162] *O*-GlcNAcylation is involved in a variety of biological processes, including metabolic reprogramming, excessive proliferation, invasion, metastasis, angiogenesis, and epigenetic gene regulation. The Warburg effect is a cancer-related phenomenon characterized by excess aerobic glycolysis, in which cancer cells adapt to hypoxic conditions by switching cellular energy from the oxidative phosphorylation pathway to glycolysis.[163,164] Glucose is also shunted into metabolic pathways as glucose uptake increases from branches of the glycolysis, HBP, or oxidative pentose phosphate pathways (PPP).[165] Importantly, the survival and growth of cancer cells *in vivo* depend on high concentrations of glutamine, and a high rate of glutamine uptake may also lead to an increase in the HBP pathway, ultimately leading to high *O*-GlcNAcylation.[165,166]

O-GlcNAc modification alters regulator activity as an enzyme or catalyst. Glycolytic enzymes such as phosphofructokinase 1 (PFK1) catalyze the rate-limiting step of glycolysis, and S529 *O*-GlcNAcylation inhibits PFK1 activity and interferes with its oligomerization by blocking binding of PFK1 to fructose-6-phosphate (F2,6BP). PFK1 inhibition results in accumulation of glycolytic intermediates and redirection of metabolic flux to the PPP pathway, thereby providing cancer cells with pentose sugars for nucleotide and nucleic acid synthesis. *O*-GlcNAc can also modify key transcription factors and metabolic sensing signaling pathways to regulate metabolic reprogramming.[166] High *O*-GlcNAcylation is the result of altered metabolism in cancer and is also involved in regulating the process of metabolic reprogramming, forming a vicious cycle of cancer metabolism.

O-GlcNAcylation is also involved in cancer cell proliferation, invasion, metastasis and angiogenesis. Oncogenic proliferation of cells mainly revolves around the cell cycle. *O*-GlcNAc maintains cell cycle progression under hypoxia, and high *O*-GlcNAcylation maintains more hypoxia-induced nuclear 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFKFB3) and higher cells split rate.[167] Decreased *O*-GlcNAcylation in breast and prostate cancer cells resulted in decreased protein levels of the oncogenic transcription factor FoxM1 and its target matrix metalloproteinases (MMPs), reduced cell cycle progression and cancer cell invasion, and inhibited tumor growth.[168,169] Besides, malignant cells induce angiogenesis, which transports oxygen and nutrients to the interior of tumors for survival and

Table 2 Comparison of analytical methods for quantification of protein *O*-GlcNAcylation. Mass spectrometry (MS) and non-MS methods have been widely used for *O*-GlcNAcylation quantification, including multi-dimensional liquid chromatography (LC), stable isotope labeling, label-free, flow cytometry, PEG labeling, Glyco-seek, AuNP/WGA lectin affinity. SILAC = stable isotope labeling by amino acids in cell culture; MRM = multiple reaction monitoring; PEG = polyethylene glycol

Methods			Sample type	Advantages	Limitations	Ref.
MS	2D-LC		Viruses; cell; tissue	Simple operation, low cost, visualization, and detection of modified proteins for which there is no effective enrichment methods	Inefficient analysis of extremely small/large and highly acidic/basic proteins; narrow linear dynamic range; insensitive detection of low-abundance proteins	91, 118, 181 and 182
	Stable isotope labeling	*In cellulo* metabolic labeling (SILAC)	Cell	High quantification accuracy	Not practical for biological samples (tissues or body fluids)	11, 109 and 183
		In vitro chemical reaction labeling	Cell; tissue	Simultaneous analysis of multiple samples, high quantitative accuracy, conducive to in-depth identification	High cost using isobaric tags; not as effective and specific as *in cellulo* metabolic labeling; complex signal	93, 110 , 111, 113–117 and 184–186
	Label-free	MRM-MS	Cell	No enrichment and labeling required, avoiding the breakage of *O*-glycosidic bonds, simplifying experimental steps and data analysis process	Peptide sequence needed, difficult to quantify unknown samples	121 and 122
Non-MS	Flow cytometry		Cell	Time-saving for intracellular signaling quantification	Semi-quantification	124
	PEG labeling		Cell; tissue	No need for purification, provide overall *O*-GlcNAc site occupancy information	Semi-quantification	125–127
	Glyco-seek		Cell	Non-destructive, ultrasensitive and low sample consumption	Not particularly well to *de novo* discovery of previously unknown glycoproteins; probably affected by antibody specificity	128
	AuNP/WGA		Cell; serum	Signal amplification, linear range increased tenfold	High cost and relatively complex procedure (when need to use OGA to improve specificity)	131 and 132

Table 3 Aberrant changes of protein *O*-GlcNAcylation are associated with a variety of diseases. *O*-GlcNAcylation is heavily involved in the initiation and progression of diabetes. Neurodegenerative diseases, including Alzheimer's disease (AD) and Parkinson's disease (PD). A wide range of cancers are reported to have dysregulated *O*-GlcNAcylation

Disease	Sample type	Sample size[a]	Method	*O*-GlcNAcylated	HL *vs.* DS[b]	Ref.
Diabetes	Blood	10 : 10	WB (CTD110.6) chemoenzymatic label (GalT) PEG labeling	25/35/—	↑ N/A N/A	111
AD	Brain tissue (frontal cerebral cortices; cerebellar cortices)	7 : 7	WB (RL2)	—	↓ N/A N/A	187
	Brain tissue (hippocampal; brain cortical)	8 : 11	WB (RL2)	—	↓ N/A N/A	188
	Brain tissue (IPL; cerebellum)	13 : 13	WB (CTD110.6)	—	↑ ↓ -	189
	Brain tissue (grey matter of the mid frontal gyru)	10 : 10	Chemoenzymatic label (GalT) click chemistry	530/1850/1094	↑ - -	184
PD	Brain tissue (temporal cortical)	12 : 29	WB (CTD110.6)	—	↑ N/A N/A	161
Lung cancer	Tissue	31 : 31	IHC	—	↑ - ↑	190
Colorectal cancer	Tissue	7 : 7	2D-gel, WB (CTD110.6)	16/—/—	↑ - ↑	191
	Tissue	31 : 31	IHC	—	↑ - ↑	190
	Serum	7 : 7	WGA enrichment WB (RL2) BEMAD	11/—/—	↑ N/A N/A	192
Prostate cancer	Tissue	29 : 55	IHC, WB (CTD110.6)	—	↑ N/A N/A	193
Hepatocellular carcinoma	Tissue	10 : 40	IHC, WB (RL2)	—	↑ ↓ -	194
Breast cancer	Tissue	39 : 43	TMA, IHC	—	↑ N/A N/A	195
	Tissue	30 : 30	sWGA enrichment, WB (CTD110.6)	42/—/—	↑ N/A ↑	196
	Tissue	12 : 12	WB (RL2)	—	↓ ↑ -	197
Laryngeal cancer	Tissue	73 : 106	WB (RL2)	—	↑ ↑ ↑	198

Table 3 (*Contd.*)

Disease	Sample type	Sample size[a]	Method	*O*-GlcNAcylated	HL *vs.* DS[b]	Ref.
Ovarian cancer	Tissue	9 : 48	TMA, IHC	—	↓ ↓ ↑	199
Thyroid cancer	Tissue	18 : 31	WB (RL2/CTD110.6)	—	↓ ↑	200
Gastric cancer	Tissue	90 : 90	IHC	—	↑ N/A N/A	201

[a] Ratio of normal and patient samples used in the studies. [b] Changes of *O*-GlcNAcylation enzymes OGT (*O*-GlcNAc transferase) and OGA (*O*-GlcNAcase). HL = healthy; DS = disease; "↑" "↓" "—" represents the change status of the disease group compared with the normal group increased, decreased and unchanged, respectively (Black: *O*-GlcNAcylation level; Orange: OGA level; Blue: OGT level); CTD110.6 = *O*-GlcNAc mouse antibody; RL2 = *O*-GlcNAc monoclonal antibody; GalT = galactosyltransferase; WGA = wheat germ agglutinin; BEMAD = β-elimination followed by Michael addition of dithiothreitol; TMA = tissue microarray; IHC = immunohistochemistry; WB = western blotting.

proliferation, and helps cancer cells invade and spread to distant organs.[170,171] *O*-GlcNAcylation promotes angiogenesis by upregulating VEGF, FoxM1, MMP[169] and FGF signaling.[172] *O*-GlcNAc is involved in the interaction and modification of many epigenetic factors, and plays a non-negligible role in the regulation of cancer epigenetic genes. Loss of ubiquitinated histone H2B caused by glucose deprivation has been observed in a variety of cancer cells such as breast, colon and lung cancer.[173] The *O*-GlcNAc group of H2B S112 acts as an anchor for the ubiquitin ligase, facilitating its K120 mono-ubiquitination.[174] Hyperphosphorylation of histone H3 S10 promotes oncogene-mediated transformation, and *O*-GlcNAcylation at this site is negatively correlated with phosphorylation.[174,175] Increased *O*-GlcNAcylation in tumor microenvironment, especially in macrophages, reduces inflammatory cytokine production, inhibits p38 MAPK and promotes tumor proliferation.[176]

Many clinical data show that *O*-GlcNAcylation is abnormal in various cancers, most of which show a significant up-regulation trend, such as colon, lung, prostate and breast cancer.[177] However, downregulation of *O*-GlcNAc has also been observed in a few cancers such as ovarian or thyroid cancer. The vast majority of these clinical data come from relevant tissue samples, and studies on biological fluids such as serum, urine or saliva are rare.[178,179] On the other hand, as shown in Table 2, immunohistochemistry (IHC) and WB are mainly used to quantify the level of *O*-GlcNAcylated protein, which can only be assessed from the overall level of the protein. There are still large research gaps for specific proteins, including protein species and modification sites, that are altered by *O*-GlcNAcylation in certain diseases. If the existing methods can be properly selected for clinical sample research and the corresponding protein information and quantitative levels can be obtained, it will be of great benefit to understand the role of *O*-GlcNAc in related diseases.

5. Conclusions and perspectives

O-GlcNAc modification is closely related to the occurrence and development of various diseases. Therefore, in the past few decades, *O*-GlcNAc modification of proteins has attracted much attention, and research methods have been constantly updated, breakthrough and developed. At the same time, a few methods have been practiced in clinical research, and there is still a large gap in the research of *O*-GlcNAc proteomics in diseases. Due to cost and operational constraints, there is currently a lack of a universally applicable method that can selectively enrich all *O*-GlcNAc peptides and is suitable for large-scale clinical sample analysis. Identification of *O*-GlcNAc peptides and sites can also be aided by databases and site mapping software. In the future, based on the existing methods and development trends, more representative methods will be developed, and the database of *O*-GlcNAc proteins related to diseases will be gradually improved to realize automatic software evaluation.

The ability to implement automated assessments depends on three things. (1) The platform has a very powerful and comprehensive *O*-GlcNAc database, covering multiple species and sample types, and is constantly updated with the deepening of research. At present, OVS laboratory and Ma laboratory have established such a public resource database (the *O*-GlcNAc database v1.2; the *O*-GlcNAcAtalas 2.0). The database provides comprehensive information for each corresponding protein, including species, sample type, modified peptide sequence, modification site and score, references, *etc.* However, information on *O*-GlcNAcylated proteins in clinical samples still needs to be developed and supplemented. (2) There are various enrichment analysis methods for *O*-GlcNAc, and different sample processing or labeling methods can be selected in the software for database retrieval. (3) Enzyme enrichment by OGA is the most specific method for analyzing *O*-GlcNAcylation. However, the existing methods are insufficient to uniquely label the modified sites after OGA digestion, making data search

difficult. When the database is comprehensive enough, OGA-processed and enriched samples can be identified and verified by comparing the site information of the database.

Abbreviation

PTM	Post-translational modification
O-GlcNAc	*O*-linked β-D-*N*-acetylglucosamine
OGT	*O*-GlcNAc transferase
OGA	*O*-GlcNAcase
O-GalNAc	*O*-linked β-*N*-acetylgalactosamine
GAG	*O*-linked-glycosaminoglycan
CID	Collision-induced dissociation
HBP	Hexosamine biosynthesis pathway
BEMAD	β-elimination followed by Michael addition of dithiothreitol
WGA	Wheat germ agglutinin
sWGA	Succinylated WGA
LWAC	Lectin weak affinity chromatography
HPA	Helix pomatia agglutinin
HILIC	Hydrophilic interaction liquid chromatography
MOE	Metabolic oligosaccharide engineering
RCA	*Ricinus communis* agglutinin
GAO	Galactose oxidase
SILAC	Stable isotope labeling with amino acids in cell culture
iTRAQ	Isobaric tag for relative and absolute quantitation
TMT	Tandem mass tags
MRM	Multiple reaction monitoring
GalNAz	*N*-Azidoacetylgalactosamine-tetraacylated
AD	Alzheimer's disease
PD	Parkinson's disease
PLA	Proximity ligation assay
SPR	Surface plasmon resonance

Conflicts of interest

The authors declare no competing financial interests.

Acknowledgements

This work was supported by the Soochow University Start-up Fund. We thank the Priority Academic Program Development of the Jiangsu Higher Education Institutes (PAPD), Jiangsu Science and Technology Plan Funding (BX2022023) and Jiangsu Shuangchuang Boshi Funding (JSSCBS20210697).

References

1 R. Apweiler, H. Hermjakob and N. Sharon, *Biochim. Biophys. Acta*, 1999, **1473**, 4–8.
2 C. R. Torres and G. W. Hart, *J. Biol. Chem.*, 1984, **259**, 3308–3317.
3 G. W. Hart, *Annu. Rev. Biochem.*, 1997, **66**, 315–335.
4 R. S. Haltiwanger, G. D. Holt and G. W. Hart, *J. Biol. Chem.*, 1990, **265**, 2563–2568.
5 D. L. Dong and G. W. Hart, *J. Biol. Chem.*, 1994, **269**, 19321–19330.
6 S. Marshall, V. Bacote and R. R. Traxinger, *J. Biol. Chem.*, 1991, **266**, 4706–4712.
7 Y. Liu, R. Z. Yao, S. Lian, P. Liu, Y. J. Hu, H. Z. Shi, H. M. Lv, Y. Y. Yang, B. Xu and S. Z. Li, *Cell Stress Chaperones*, 2021, **26**, 297–309.
8 J. A. Hanover, M. W. Krause and D. C. Love, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 312–321.
9 Y. F. Chen, G. C. Shao, J. Li, A. Q. Yang, J. Li and X. S. Ye, *Mol. Cancer Res.*, 2022, **20**, 650–660.
10 X. Li, Q. Zhu, X. Shi, Y. Cheng, X. Li, H. Xu, X. Duan, L. C. Hsieh-Wilson, J. Chu, J. Pelletier, M. Ni, Z. Zheng, S. Li and W. Yi, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 7857–7866.
11 Z. Wang, N. D. Udeshi, C. Slawson, P. D. Compton, K. Sakabe, W. D. Cheung, J. Shabanowitz, D. F. Hunt and G. W. Hart, *Sci. Signaling*, 2010, **3**, ra2.
12 E. P. Tan, S. R. McGreal, S. Graw, R. Tessman, S. J. Koppel, P. Dhakal, Z. Zhang, M. Machacek, N. E. Zachara, D. C. Koestler, K. R. Peterson, J. P. Thyfault, R. H. Swerdlow, P. Krishnamurthy, L. DiTacchio, U. Apte and C. Slawson, *J. Biol. Chem.*, 2017, **292**, 14940–14962.
13 V. Makwana, S. Rudrawar and S. Anoopkumar-Dukie, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2021, **1867**, 166129.
14 C. Han, Y. Gu, H. Shan, W. Mi, J. Sun, M. Shi, X. Zhang, X. Lu, F. Han, Q. Gong and W. Yu, *Nat. Commun.*, 2017, **8**, 1491.
15 X. Liu, I. Blaženović, A. J. Contreras, T. M. Pham, C. A. Tabuloc, Y. H. Li, J. Ji, O. Fiehn and J. C. Chiu, *Nat. Commun.*, 2021, **12**, 4173.
16 G. W. Hart, M. P. Housley and C. Slawson, *Nature*, 2007, **446**, 1017–1022.
17 E. J. Kim, *Molecules*, 2011, **16**, 1987–2022.
18 F. I. Comer, K. Vosseller, L. Wells, M. A. Accavitti and G. W. Hart, *Anal. Biochem.*, 2001, **293**, 169–177.
19 C. M. Snow, A. Senior and L. Gerace, *J. Cell Biol.*, 1987, **104**, 1143–1156.
20 J. R. Turner, A. M. Tartakoff and N. S. Greenspan, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 5608–5612.
21 N. Yoshida, R. A. Mortara, M. F. Araguth, J. C. Gonzalez and M. Russo, *Infect. Immun.*, 1989, **57**, 1663–1667.
22 N. Khidekel, S. Arndt, N. Lamarre-Vincent, A. Lippert, K. G. Poulin-Kerstien, B. Ramakrishnan, P. K. Qasba and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2003, **125**, 16162–16163.
23 S. Xu, F. Sun and R. Wu, *Anal. Chem.*, 2020, **92**, 9807–9814.
24 Y. Chen, F. Tang, H. Qin, X. Yue, Y. Nie, W. Huang and M. Ye, *Angew. Chem., Int. Ed. Engl.*, 2022, e202117849; *Angew. Chem.*, 2022, **134**, e202117849.
25 R. J. Chalkley and A. L. Burlingame, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 1106–1113.
26 S. A. Myers, S. Daou, B. Affar el and A. Burlingame, *Proteomics*, 2013, **13**, 982–991.

27 M. Adamczyk, J. C. Gebler and J. Wu, *Rapid Commun. Mass Spectrom.*, 2001, **15**, 1481–1488.
28 W. Weckwerth, L. Willmitzer and O. Fiehn, *Rapid Commun. Mass Spectrom.*, 2000, **14**, 1677–1681.
29 Y. Oda, T. Nagasu and B. T. Chait, *Nat. Biotechnol.*, 2001, **19**, 379–382.
30 M. B. Goshe, T. P. Conrads, E. A. Panisko, N. H. Angell, T. D. Veenstra and R. D. Smith, *Anal. Chem.*, 2001, **73**, 2578–2586.
31 K. D. Greis, B. K. Hayes, F. I. Comer, M. Kirk, S. Barnes, T. L. Lowary and G. W. Hart, *Anal. Biochem.*, 1996, **234**, 38–49.
32 L. Wells, K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis and G. W. Hart, *Mol. Cell. Proteomics*, 2002, **1**, 791–804.
33 M. Fukuda, *Curr. Protoc. Mol. Biol.*, 2001, ch. 17, Unit17.15B.
34 H. E. Conrad, *Curr. Protoc. Mol. Biol.*, 2001, ch. 17, Unit17.15A.
35 H. E. Meyer, E. Hoffmann-Posorske and L. M. Heilmeyer Jr, *Methods Enzymol.*, 1991, **201**, 169–185.
36 H. Zhang, X. J. Li, D. B. Martin and R. Aebersold, *Nat. Biotechnol.*, 2003, **21**, 660–666.
37 C. Leteux, R. A. Childs, W. Chai, M. S. Stoll, H. Kogelberg and T. Feizi, *Glycobiology*, 1998, **8**, 227–236.
38 S. Yang and H. Zhang, *Proteomics: Clin. Appl.*, 2012, **6**, 596–608.
39 J. Li, J. Zhang, M. Xu, Z. Yang, S. Yue, W. Zhou, C. Gui, H. Zhang, S. Li, P. G. Wang and S. Yang, *J. Sep. Sci.*, 2022, **45**, 3169–3186.
40 E. Klement, Z. Lipinszki, Z. Kupihár, A. Udvardy and K. F. Medzihradszky, *J. Proteome Res.*, 2010, **9**, 2200–2206.
41 B. Sun, Q. Hou, Z. Liu and Y. Ni, *Cellulose*, 2015, **22**, 1135–1146.
42 M. Kobayashi, T. Urayama, I. Suzawa, S. Takagi, K. Matsuda and E. Ichishima, *Agric. Biol. Chem.*, 1988, **52**, 2695–2702.
43 M. Pumera, I. Jelínek, J. Jindrich, P. Coufal and J. Horský, *J. Chromatogr. A*, 2000, **891**, 201–206.
44 T. K. Dam and C. F. Brewer, in *Comprehensive Glycoscience*, ed. H. Kamerling, Elsevier, Oxford, 2007, pp. 397–452, DOI: 10.1016/B978-044451967-2/00058-1.
45 C. S. Wright, *J. Biol. Chem.*, 1992, **267**, 14345–14352.
46 A. M. Wu, J. H. Wu, S. C. Song, M. S. Tsai and A. Herp, *FEBS Lett.*, 1998, **440**, 315–319.
47 J. T. Gallagher, A. Morris and T. M. Dexter, *Biochem. J.*, 1985, **231**, 115–122.
48 M. Monsigny, C. Sene, A. Obrenovitch, A. C. Roche, F. Delmotte and E. Boschetti, *Eur. J. Biochem.*, 1979, **98**, 39–45.
49 K. Fahie, B. Narayanan, F. Zahra, R. Reeves, S. M. Fernandes, G. W. Hart and N. E. Zachara, *Curr. Protoc.*, 2021, **1**, e129.
50 K. Vosseller, J. C. Trinidad, R. J. Chalkley, C. G. Specht, A. Thalhammer, A. J. Lynn, J. O. Snedecor, S. Guan, K. F. Medzihradszky, D. A. Maltby, R. Schoepfer and A. L. Burlingame, *Mol. Cell. Proteomics*, 2006, **5**, 923–934.
51 S. Nakamura-Tsuruta, J. Kominami, M. Kamei, Y. Koyama, T. Suzuki, M. Isemura and J. Hirabayashi, *J. Biochem.*, 2006, **140**, 285–291.
52 O. Machon, S. Baldini, J. P. Ribeiro, A. Steenackers, A. Varrot, T. Lefebvre and A. J. G. Imberty, *Glycobiology*, 2017, **27**, 123–128.
53 W. Liu, G. Han, Y. Yin, S. Jiang, G. Yu, Q. Yang, W. Yu, X. Ye, Y. Su, Y. Yang, G. W. Hart and H. Sun, *Glycobiology*, 2018, **28**, 363–373.
54 N. D. Rambaruth, M. V. Greenwell P Fau - Dwek and M. V. Dwek, *Glycobiology*, 2012, **22**, 839–848.
55 F. I. Comer, L. Vosseller, K. Fau-Wells, M. A. W. L. Fau-Accavitti, G. W. Accavitti Ma Fau-Hart and G. W. Hart, *Anal. Biochem.*, 2001, **293**, 169–177.
56 C. F. Teo, S. Ingale, M. A. Wolfert, G. A. Elsayed, L. G. Nöt, J. C. Chatham, L. Wells and G. J. Boons, *Nat. Chem. Biol.*, 2010, **6**, 338–343.
57 T. Isono, *PLoS One*, 2011, **6**, e18959.
58 M. Ogawa, N. Nakamura, Y. Nakayama, A. Kurosaka, H. Manya, M. Kanagawa, T. Endo, K. Furukawa and T. Okajima, *Biochem. Biophys. Res. Commun.*, 2013, **440**, 88–93.
59 Y. Tashima and P. Stanley, *J. Biol. Chem.*, 2014, **289**, 11132–11142.
60 R. A. Burt, B. Dejanovic, H. J. Peckham, K. A. Lee, X. Li, J. R. Ounadjela, A. Rao, S. A. Malaker, S. A. Carr and S. A. Myers, *Mol. Cell. Proteomics*, 2021, **20**, 100167.
61 H. Qin, Y. Chen, J. Mao, K. Cheng, D. Sun, M. Dong, L. Wang, L. Wang and M. Ye, *Anal. Chim. Acta*, 2019, **1070**, 60–68.
62 G. Shen, F. Zhang, B. Yang, C. Chu and X. Liang, *Talanta*, 2013, **115**, 129–132.
63 B. Shen, W. Zhang, Z. Shi, F. Tian, Y. Deng, C. Sun, G. Wang, W. Qin and X. Qian, *Talanta*, 2017, **169**, 195–202.
64 C. Wu, S. Shi, C. Hou, Y. Luo, S. Byers and J. Ma, *ACS Appl. Mater. Interfaces*, 2022, **14**, 47482–47490.
65 A. Cioce, G. Bineva-Todd, A. J. Agbay, J. Choi, T. M. Wood, M. F. Debets, W. M. Browne, H. L. Douglas, C. Roustan, O. Y. Tastan, S. Kjaer, J. T. Bush, C. R. Bertozzi and B. Schumann, *ACS Chem. Biol.*, 2021, **16**, 1961–1967.
66 C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu and P. Wu, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 8051–8056.
67 R. T. Almaraz, U. Aich, H. S. Khanna, E. Tan, R. Bhattacharya, S. Shah and K. J. Yarema, *Biotechnol. Bioeng.*, 2012, **109**, 992–1006.
68 D. J. Vocadlo, H. C. Hang, E. J. Kim, J. A. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9116–9121.
69 A. Nandi, R. Sprung, D. K. Barma, Y. Zhao, S. C. Kim, J. R. Falck and Y. Zhao, *Anal. Chem.*, 2006, **78**, 452–458.
70 Y. Zhu, J. Wu and X. Chen, *Angew. Chem., Int. Ed. Engl.*, 2016, **55**, 9301–9305.
71 M. Boyce, I. S. Carrico, A. S. Ganguli, S. H. Yu, M. J. Hangauer, S. C. Hubbard, J. J. Kohler and

C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 3141–3146.
72 P. S. Banerjee, P. Ostapchuk, P. Hearing and I. Carrico, *J. Am. Chem. Soc.*, 2010, **132**, 13615–13617.
73 K. Qin, H. Zhang, Z. Zhao and X. Chen, *J. Am. Chem. Soc.*, 2020, **142**, 9382–9388.
74 B. W. Zaro, Y.-Y. Yang, H. C. Hang and M. R. Pratt, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 8146–8151.
75 K. N. Chuh, B. W. Zaro, F. Piller, V. Piller and M. R. Pratt, *J. Am. Chem. Soc.*, 2014, **136**, 12283–12295.
76 J. Li, J. Wang, L. Wen, H. Zhu, S. Li, K. Huang, K. Jiang, X. Li, C. Ma, J. Qu, A. Parameswaran, J. Song, W. Zhao and P. G. Wang, *ACS Chem. Biol.*, 2016, **11**, 3002–3006.
77 E. G. Jackson, G. Cutolo, B. Yang, N. Yarravarapu, M. W. N. Burns, G. Bineva-Todd, C. Roustan, J. B. Thoden, H. M. Lin-Jones, T. H. van Kuppevelt, H. M. Holden, B. Schumann, J. J. Kohler, C. M. Woo and M. R. Pratt, *ACS Chem. Biol.*, 2022, **17**, 159–170.
78 S.-H. Yu, M. Boyce, A. M. Wands, M. R. Bond, C. R. Bertozzi and J. J. Kohler, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 4834–4839.
79 M. Amado, R. Almeida, T. Schwientek and H. Clausen, *Biochim. Biophys. Acta*, 1999, **1473**, 35–53.
80 K. Furukawa and H. Clausen, in *Handbook of Glycosyltransferases and Related Genes*, ed. N. Taniguchi, K. Honke, M. Fukuda, H. Clausen, K. Furukawa, G. W. Hart, R. Kannagi, T. Kawasaki, T. Kinoshita, T. Muramatsu, M. Saito, J. H. Shaper, K. Sugahara, L. A. Tabak, D. H. Van den Eijnden, M. Yanagishita, J. W. Dennis, K. Furukawa, Y. Hirabayashi, M. Kawakita, K. Kimata, U. Lindahl, H. Narimatsu, H. Schachter, P. Stanley, A. Suzuki, S. Tsuji and K. Yamashita, Springer Japan, Tokyo, 2002, pp. 20–26, DOI: 10.1007/978-4-431-67877-9_3.
81 G. D. Holt and G. W. Hart, *J. Biol. Chem.*, 1986, **261**, 8049–8057.
82 E. P. Roquemore, T. Y. Chou and G. W. Hart, *Methods Enzymol.*, 1994, **230**, 443–460.
83 B. Fang and M. W. Miller, *Exp. Cell Res.*, 2001, **263**, 243–253.
84 T. Nishikaze, S. Kawabata, S. Iwamoto and K. Tanaka, *Analyst*, 2013, **138**, 7224–7232.
85 J. Zheng, H. Xiao and R. Wu, *Angew. Chem., Int. Ed. Engl.*, 2017, **56**, 7107–7111.
86 J. B. Rannes, A. Ioannou, S. C. Willies, G. Grogan, C. Behrens, S. L. Flitsch and N. J. Turner, *J. Am. Chem. Soc.*, 2011, **133**, 8436–8439.
87 Y. Chen, H. Qin, X. Yue, J. Zhou, L. Liu, Y. Nie and M. Ye, *Anal. Chem.*, 2021, **93**, 16618–16627.
88 B. Ramakrishnan and P. K. Qasba, *J. Biol. Chem.*, 2002, **277**, 20833–20839.
89 X. Qian, K. Sujino, M. M. Palcic and R. M. Ratcliffe, *J. Carbohydr. Chem.*, 2002, **21**, 911–942.
90 X. Qian, K. Sujino, M. M. Palcic, B. Ernst, G. W. Hart and P. J. C. Sinaÿ, *J. Cheminf.*, 2008, **32**, 685–703.
91 P. M. Clark, J. F. Dweck, D. E. Mason, C. R. Hart, S. B. Buck, E. C. Peters, B. J. Agnew and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2008, **130**, 11576–11577.
92 N. Khidekel, S. B. Ficarro, E. C. Peters and L. C. Hsieh-Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13132–13137.
93 Z. L. Wu, A. Luo, A. Grill, T. Lao, Y. Zou and Y. Chen, *Bioconjugate Chem.*, 2020, **31**, 2098–2102.
94 X. Sun, L. Tao, L. Yi, Y. Ouyang, N. Xu, D. Li, R. J. Linhardt and Z. Zhang, *J. Pharm. Anal.*, 2017, **7**, 87–94.
95 B. W. Zaro, Y.-Y. Yang, H. C. Hang and M. R. Pratt, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 8146–8151.
96 Z. Wang, N. D. Udeshi, M. O'Malley, J. Shabanowitz, D. F. Hunt and G. W. Hart, *Mol. Cell. Proteomics*, 2010, **9**, 153–160.
97 J. F. Alfaro, C.-X. Gong, M. E. Monroe, J. T. Aldrich, T. R. W. Clauss, S. O. Purvine, Z. Wang, D. G. Camp, J. Shabanowitz, P. Stanley, G. W. Hart, D. F. Hunt, F. Yang and R. D. Smith, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7280–7285.
98 J. Ma, W. H. Wang, Z. Li, J. Shabanowitz, D. F. Hunt and G. W. Hart, *Anal. Chem.*, 2019, **91**, 2620–2625.
99 J. U. Baenziger and D. Fiete, *J. Biol. Chem.*, 1979, **254**, 9795–9799.
100 J. P. Frénoy, A. T. Tran and R. Bourrillon, *Biochem. J.*, 1986, **240**, 227–231.
101 P. A. Haynes and R. Aebersold, *Anal. Chem.*, 2000, **72**, 5402–5410.
102 M. Umekawa, C. Li, T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto and L.-X. Wang, *J. Biol. Chem.*, 2010, **285**, 511–521.
103 K. Yamamoto, T. Tsuji, I. Matsumoto and T. Osawa, *Biochemistry*, 1981, **20**, 5894–5899.
104 H. Abo, M. Kume, F. Pecori, T. Miura, N. Matsumoto, S. Nishihara and K. Yamamoto, *PLoS One*, 2022, **17**, e0267804.
105 K. Soga, F. Teruya, H. Tateno, J. Hirabayashi and K. Yamamoto, *PLoS One*, 2013, **8**, e83886.
106 J. U. Baenziger and E. D. Green, *Biochim. Biophys. Acta*, 1988, **947**, 287–306.
107 M. Gotoh, T. Sato, K. Kiyohara, A. Kameyama, N. Kikuchi, Y. D. Kwon, Y. Ishizuka, T. Iwai, H. Nakanishi and H. Narimatsu, *FEBS Lett.*, 2004, **562**, 134–140.
108 S.-E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann, *Mol. Cell. Proteomics*, 2002, **1**, 376–386.
109 J. Zhong, M. Martinez, S. Sengupta, A. Lee, X. Wu, R. Chaerkady, A. Chatterjee, R. N. O'Meally, R. N. Cole, A. Pandey and N. E. Zachara, *Proteomics*, 2015, **15**, 591–607.
110 Y. V. Skorobogatko, J. Deuso, J. Adolf-Bryfogle, M. G. Nowak, Y. Gong, C. F. Lippa and K. Vosseller, *Amino Acids*, 2011, **40**, 765–779.
111 Z. Wang, K. Park, F. Comer, L. C. Hsieh-Wilson, C. D. Saudek and G. W. Hart, *Diabetes*, 2009, **58**, 309–317.
112 J. Liu, X. Shao, W. Qin, Y. Zhang, F. Dang, Q. Yang, X. Yu, Y. X. Li, X. Chen, C. Wang and Y. L. Wang, *Cell Chem. Biol.*, 2021, **28**, 788–801.e5.
113 N. Khidekel, S. B. Ficarro, P. M. Clark, M. C. Bryan, D. L. Swaney, J. E. Rexach, Y. E. Sun, J. J. Coon,

E. C. Peters and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2007, **3**, 339–348.
114 Y. Zhou, Y. Shan, L. Zhang and Y. Zhang, *J. Chromatogr. A*, 2014, **1365**, 1–11.
115 P. J. Lund, J. E. Elias and M. M. Davis, *J. Immunol.*, 2016, **197**, 3086–3098.
116 K. Qin, Y. Zhu, W. Qin, J. Gao, X. Shao, Y. L. Wang, W. Zhou, C. Wang and X. Chen, *ACS Chem. Biol.*, 2018, **13**, 1983–1989.
117 J. Liu, Y. Hao, C. Wang, Y. Jin, Y. Yang, J. Gu and X. Chen, *ACS Chem. Biol.*, 2022, **17**, 513–520.
118 A. Rogowska-Wrzesinska, M.-C. Le Bihan, M. Thaysen-Andersen and P. Roepstorff, *J. Proteomics*, 2013, **88**, 4–13.
119 W. Zhu, J. W. Smith and C. M. Huang, *J. Biomed. Biotechnol.*, 2010, **2010**, 840518.
120 D. S. Kirkpatrick, S. A. Gerber and S. P. Gygi, *Methods*, 2005, **35**, 265–273.
121 S. A. Yuzwa, A. K. Yadav, Y. Skorobogatko, T. Clark, K. Vosseller and D. J. Vocadlo, *Amino Acids*, 2011, **40**, 857–868.
122 J. J. Maury, D. Ng, X. Bi, M. Bardor and A. B. Choo, *Anal. Chem.*, 2014, **86**, 395–402.
123 S. Cecioni and D. J. Vocadlo, *Curr. Opin. Chem. Biol.*, 2013, **17**, 719–728.
124 S. A. Madsen-Bouterse, Y. Xu, H. R. Petty and R. Romero, *Cytometry, Part A*, 2008, **73**, 667–672.
125 J. E. Rexach, C. J. Rogers, S. H. Yu, J. Tao, Y. E. Sun and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2010, **6**, 645–651.
126 J. E. Rexach, P. M. Clark, D. E. Mason, R. L. Neve, E. C. Peters and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2012, **8**, 253–261.
127 N. Darabedian, J. W. Thompson, K. N. Chuh, L. C. Hsieh-Wilson and M. R. Pratt, *Biochemistry*, 2018, **57**, 5769–5774.
128 P. V. Robinson, C. T. Tsai, A. E. de Groot, J. L. McKechnie and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2016, **138**, 10722–10725.
129 V. Pavlov, Y. Xiao, B. Shlyahovsky and I. Willner, *J. Am. Chem. Soc.*, 2004, **126**, 11768–11769.
130 Y. Zhao, Y. Zheng, C. Zhao, J. You and F. Qu, *Biosens. Bioelectron.*, 2015, **71**, 200–206.
131 Y. Liu, J. Xie, Z. Zhang and Z. Lu, *RSC Adv.*, 2016, **6**, 89484–89491.
132 L. Gao, R. Zhao, Y. Wang, M. Lu, D. Yang, M. Fa and X. Yao, *Anal. Chim. Acta*, 2018, **1040**, 90–98.
133 D. Daneman, *Lancet*, 2006, **367**, 847–858.
134 M. Stumvoll, B. J. Goldstein and T. W. van Haeften, *Lancet*, 2005, **365**, 1333–1346.
135 F. Giacco and M. Brownlee, *Circ. Res.*, 2010, **107**, 1058–1070.
136 M. Federici, R. Menghini, A. Mauriello, M. L. Hribal, F. Ferrelli, D. Lauro, P. Sbraccia, L. G. Spagnoli, G. Sesti and R. Lauro, *Circulation*, 2002, **106**, 466–472.
137 S. A. Whelan, W. B. Dias, L. Thiruneelakantapillai, M. D. Lane and G. W. Hart, *J. Biol. Chem.*, 2010, **285**, 5204–5211.
138 A. Kaleem, S. Javed, N. Rehman, R. Abdullah, M. Iqtedar, M. N. Aftab, D. C. Hoessli and I. U. Haq, *Protein Pept. Lett.*, 2021, **28**, 333–339.
139 G. J. Parker, K. C. Lund, R. P. Taylor and D. A. McClain, *J. Biol. Chem.*, 2003, **278**, 10022–10027.
140 Z. Wang, A. Pandey and G. W. Hart, *Mol. Cell. Proteomics*, 2007, **6**, 1365–1379.
141 J. Hedou, C. Cieniewski-Bernard, Y. Leroy, J. C. Michalski, Y. Mounier and B. Bastide, *J. Biol. Chem.*, 2007, **282**, 10360–10369.
142 B. Luo, Y. Soesanto and D. A. McClain, *Arterioscler., Thromb., Vasc. Biol.*, 2008, **28**, 651–657.
143 Z. Gurel, B. W. Zaro, M. R. Pratt and N. Sheibani, *PLoS One*, 2014, **9**, e95561.
144 Z. Gurel and N. Sheibani, *Clin. Sci.*, 2018, **132**, 185–198.
145 E. Wulff-Fuentes, R. R. Berendt, L. Massman, L. Danner, F. Malard, J. Vora, R. Kahsay and S. Olivier-Van Stichelen, *Sci. Data*, 2021, **8**, 25.
146 B. E. Lee, P. G. Suh and J. I. Kim, *Exp. Mol. Med.*, 2021, **53**, 1674–1682.
147 J. Cheng, X. Xia, Y. Rui, Z. Zhang, L. Qin, S. Han and Z. Wan, *Food Chem. Toxicol.*, 2016, **93**, 32–40.
148 S. A. Yuzwa, X. Shan, M. S. Macauley, T. Clark, Y. Skorobogatko, K. Vosseller and D. J. Vocadlo, *Nat. Chem. Biol.*, 2012, **8**, 393–399.
149 F. Kametani and M. Hasegawa, *Front. Neurosci.*, 2018, **12**, 25.
150 H. N. Woo, S. H. Baik, J. S. Park, A. R. Gwon, S. Yang, Y. K. Yun and D. G. Jo, *Biochem. Biophys. Res. Commun.*, 2011, **404**, 10–15.
151 G. S. Bloom, *JAMA Neurol.*, 2014, **71**, 505–508.
152 Y. S. Chun, O. H. Kwon and S. Chung, *Biochem. Biophys. Res. Commun.*, 2017, **490**, 486–491.
153 V. M. Lee and J. Q. Trojanowski, *Neuron*, 2006, **52**, 33–38.
154 W. S. Kim, K. Kågedal and G. M. Halliday, *Alzheimer's Res. Ther.*, 2014, **6**, 73.
155 Z. Wang, N. D. Udeshi, M. O'Malley, J. Shabanowitz, D. F. Hunt and G. W. Hart, *Mol. Cell. Proteomics*, 2010, **9**, 153–160.
156 J. Zhang, X. Li and J. D. Li, *Front. Neurosci.*, 2019, **13**, 381.
157 N. P. Marotta, Y. H. Lin, Y. E. Lewis, M. R. Ambroso, B. W. Zaro, M. T. Roth, D. B. Arnold, R. Langen and M. R. Pratt, *Nat. Chem.*, 2015, **7**, 913–920.
158 P. M. Levine, A. Galesic, A. T. Balana, A. L. Mahul-Mellier, M. X. Navarro, C. A. De Leon, H. A. Lashuel and M. R. Pratt, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 1511–1519.
159 N. P. Marotta, C. A. Cherwien, T. Abeywardana and M. R. Pratt, *Chembiochem*, 2012, **13**, 2665–2670.
160 B. E. Lee, H. Y. Kim, H. J. Kim, H. Jeong, B. G. Kim, H. E. Lee, J. Lee, H. B. Kim, S. E. Lee, Y. R. Yang, E. C. Yi, J. A. Hanover, K. Myung, P. G. Suh, T. Kwon and J. I. Kim, *Brain*, 2020, **143**, 3699–3716.
161 W. Y. Wani, X. Ouyang, G. A. Benavides, M. Redmann, S. S. Cofield, J. J. Shacka, J. C. Chatham, V. Darley-Usmar and J. Zhang, *Mol. Brain*, 2017, **10**, 32.
162 D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646–674.
163 O. Warburg, F. Wind and E. Negelein, *J. Gen. Physiol.*, 1927, **8**, 519–530.
164 O. Warburg, *Science*, 1956, **124**, 269–270.

165 J. W. Sinclair, D. R. Hoying, E. Bresciani, D. D. Nogare, C. D. Needle, A. Berger, W. Wu, K. Bishop, A. G. Elkahloun, A. Chitnis, P. Liu and S. M. Burgess, *npj Regener. Med.*, 2021, **6**, 55.
166 Z. Li and W. Yi, *Glycoconjugate J.*, 2014, **31**, 185–191.
167 Y. Lei, T. Chen, Y. Li, M. Shang, Y. Zhang, Y. Jin, Q. Yu, F. Guo and T. Wang, *Oncogenesis*, 2020, **9**, 21.
168 S. A. Caldwell, S. R. Jackson, K. S. Shahriari, T. P. Lynch, G. Sethi, S. Walker, K. Vosseller and M. J. Reginato, *Oncogene*, 2010, **29**, 2831–2842.
169 T. P. Lynch, C. M. Ferrer, S. R. Jackson, K. S. Shahriari, K. Vosseller and M. J. Reginato, *J. Biol. Chem.*, 2012, **287**, 11070–11081.
170 D. Hanahan and J. Folkman, *Cell*, 1996, **86**, 353–364.
171 R. Lugano, M. Ramachandran and A. Dimberg, *Cell. Mol. Life Sci.*, 2020, **77**, 1745–1770.
172 D. Mariappa, K. Sauert, K. Mariño, D. Turnock, R. Webster, D. M. F. van Aalten, M. A. J. Ferguson and H.-A. J. Müller, *Sci. Signaling*, 2011, **4**, ra89.
173 Y. Urasaki, L. Heath and C. W. Xu, *PLoS One*, 2012, **7**, e36775.
174 R. Fujiki, W. Hashiba, H. Sekine, A. Yokoyama, T. Chikanishi, S. Ito, Y. Imai, J. Kim, H. H. He, K. Igarashi, J. Kanno, F. Ohtake, H. Kitagawa, R. G. Roeder, M. Brown and S. Kato, *Nature*, 2011, **480**, 557–560.
175 I. S. Strelkov and J. R. Davie, *Cancer Res.*, 2002, **62**, 75–78.
176 K. Moriwaki and M. Asahi, *Mol. Cancer Res.*, 2017, **15**, 1287–1298.
177 M. Xu, A. Yang, J. Xia, J. Jiang, C.-F. Liu, Z. Ye, J. Ma and S. Yang, *Transl. Res.*, 2022, 1–17.
178 Z. Gao, Z. Wu, Y. Han, X. Zhang, P. Hao, M. Xu, S. Huang, S. Li, J. Xia, J. Jiang and S. Yang, *ACS Omega*, 2022, **7**, 17894–17906.
179 H. Iqbal, T. Yang, T. Li, M. Zhang, H. Ke, D. Ding, Y. Deng and H. Chen, *J. Controlled Release*, 2021, **329**, 997–1022.
180 Y. Chen, F. Tang, H. Qin, X. Yue, Y. Nie, W. Huang and M. Ye, *Angew. Chem., Int. Ed. Engl.*, 2022, **61**, e202117849.
181 J. Park, H. Kwon, Y. Kang and Y. Kim, *J. Biochem. Mol. Biol.*, 2007, **40**, 1058–1068.
182 L. Drougat, S. Olivier-Van Stichelen, M. Mortuaire, F. Foulquier, A.-S. Lacoste, J.-C. Michalski, T. Lefebvre and A.-S. Vercoutter-Edouart, *Biochim. Biophys. Acta, Gen. Subj.*, 2012, **1820**, 1839–1848.
183 A. Lee, D. Miller, R. Henry, V. D. Paruchuri, R. N. O'Meally, T. Boronina, R. N. Cole and N. E. Zachara, *J. Proteome Res.*, 2016, **15**, 4318–4336.
184 S. Wang, F. Yang, V. A. Petyuk, A. K. Shukla, M. E. Monroe, M. A. Gritsenko, K. D. Rodland, R. D. Smith, W. J. Qian, C. X. Gong and T. Liu, *J. Pathol.*, 2017, **243**, 78–88.
185 H. Tsumoto, Y. Akimoto, T. Endo and Y. Miura, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 5022–5026.
186 G. A. Ramirez-Correa, J. Ma, C. Slawson, Q. Zeidan, N. S. Lugo-Fagundo, M. Xu, X. Shen, W. D. Gao, V. Caceres, K. Chakir, L. DeVine, R. N. Cole, L. Marchionni, N. Paolocci, G. W. Hart and A. M. Murphy, *Diabetes*, 2015, **64**, 3573–3587.
187 F. Liu, J. Shi, H. Tanimukai, J. Gu, J. Gu, I. Grundke-Iqbal, K. Iqbal and C. X. Gong, *Brain*, 2009, **132**, 1820–1832.
188 T. S. Pinho, S. C. Correia, G. Perry, A. F. Ambrósio and P. I. Moreira, *Biochim. Biophys. Acta, Gen. Subj.*, 2019, **1865**, 2048–2059.
189 S. Förster, A. S. Welleford, J. C. Triplett, R. Sultana, B. Schmitz and D. A. Butterfield, *Biochim. Biophys. Acta*, 2014, **1842**, 1333–1339.
190 W. Mi, Y. Gu, C. Han, H. Liu, Q. Fan, X. Zhang, Q. Cong and W. Yu, *Biochim. Biophys. Acta*, 2011, **1812**, 514–519.
191 T. Phueaouan, P. Chaiyawat, P. Netsirisawan, D. Chokchaichamnankit, P. Punyarit, C. Srisomsap, J. Svasti and V. Champattanachai, *Oncol. Rep.*, 2013, **30**, 2929–2936.
192 C. Verathamjamras, T. E. Sriwitool, P. Netsirisawan, P. Chaiyawat, D. Chokchaichamnankit, N. Prasongsook, C. Srisomsap, J. Svasti and V. Champattanachai, *Glycoconjugate J.*, 2021, **38**, 55–65.
193 Y. Gu, J. Gao, C. Han, X. Zhang, H. Liu, L. Ma, X. Sun and W. Yu, *Mol. Med. Rep.*, 2014, **10**, 897–904.
194 Q. Zhu, L. Zhou, Z. Yang, M. Lai, H. Xie, L. Wu, C. Xing, F. Zhang and S. Zheng, *Med. Oncol.*, 2012, **29**, 985–993.
195 G. M. Trinca, M. L. Goodman, E. K. Papachristou, C. S. D'Santos, P. Chalise, R. Madan, C. Slawson and C. R. Hagan, *Horm. Cancer*, 2018, **9**, 12–21.
196 V. Champattanachai, P. Netsirisawan, P. Chaiyawat, T. Phueaouan, R. Charoenwattanasatien, D. Chokchaichamnankit, P. Punyarit, C. Srisomsap and J. Svasti, *Proteomics*, 2013, **13**, 2088–2099.
197 C. Slawson, J. Pidala and R. Potter, *Biochim. Biophys. Acta*, 2001, **1537**, 147–157.
198 K. Starska, E. Forma, E. Brzezińska-Błaszczyk, I. Lewy-Trenda, M. Bryś, P. Jóźwiak and A. Krześlak, *Clin. Exp. Med.*, 2015, **15**, 455–468.
199 R. M. de Queiroz, R. Madan, J. Chien, W. B. Dias and C. Slawson, *J. Biol. Chem.*, 2016, **291**, 18897–18914.
200 A. Krzeslak, L. Pomorski and A. Lipinska, *Int. J. Mol. Med.*, 2010, **25**, 643–648.
201 M. Jiang, Z. Qiu, S. Zhang, X. Fan, X. Cai, B. Xu, X. Li, J. Zhou, X. Zhang, Y. Chu, W. Wang, J. Liang, T. Horvath, X. Yang, K. Wu, Y. Nie and D. Fan, *Oncotarget*, 2016, **7**, 61390–61402.