We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,600
Open access books available

137,000
International authors and editors

170M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Improving the Study of Protein Glycosylation with New Tools for Glycopeptide Enrichment

Minyong Chen, Steven J. Dupard, Colleen M. McClung, Cristian I. Ruse, Mehul B. Ganatra, Saulius Vainauskas, Christopher H. Taron and James C. Samuelson

Abstract

High confidence methods are needed for determining the glycosylation profiles of complex biological samples as well as recombinant therapeutic proteins. A common glycan analysis workflow involves liberation of N-glycans from glycoproteins with PNGase F or O-glycans by hydrazinolysis prior to their analysis. This method is limited in that it does not permit determination of glycan attachment sites. Alternative proteomics-based workflows are emerging that utilize site-specific proteolysis to generate peptide mixtures followed by selective enrichment strategies to isolate glycopeptides. Methods designed for the analysis of complex samples can yield a comprehensive snapshot of individual glycans species, the site of attachment of each individual glycan and the identity of the respective protein in many cases. This chapter will highlight advancements in enzymes that digest glycoproteins into distinct fragments and new strategies to enrich specific glycopeptides.

Keywords: glycoproteomics, glycopeptide enrichment, lectin, Fbs1, BGL, O-glycoprotease, alpha-lytic protease

1. Introduction

Protein glycosylation is a post-translational carbohydrate ('glycan') modification of eukaryotic proteins that may affect their folding, stability, localization and biological function. Glycan profiles differ from cell type to cell type and are known to be altered in carcinogenesis [1, 2], inflammation [3], and Alzheimer's disease [4] etc. Importantly, circulating plasma proteins may serve as biomarkers as altered glycosylation profiles may signal specific types of disease. Glycosylation may be assessed on a global level by isolating total protein from tissue, cells or serum followed by liberation of glycans and analysis by reliable methods such as ultra performance liquid chromatography (UPLC) coupled to mass spectrometry (MS). Profiling liberated glycans is useful in some cases, but site of attachment (glycosite) and the glycan structure at each glycosite is more valuable. This "total picture" is attainable when performing liquid chromatography (LC) with tandem mass spectrometry (LC–MS/MS) at the glycopeptide level. Furthermore, it is critically important to be able to strictly characterize therapeutic proteins to ensure reproducible glycosylation,

safety and efficacy as the absence, presence or type of glycan is known to dictate the efficacy of some therapeutic molecules [5].

Glycans are attached to certain asparagine residues (N-linked glycans) or serine/ threonine residues (O-linked glycans). N-linked glycosylation occurs at the Asn-X-Ser/Thr/Cys (where X is not proline) consensus sequence on proteins that pass through the eukaryotic secretory pathway. There are three structural classes of N-glycan that share a common trimannosyl chitobiose core motif (Man3GlcNAc2) (**Figure 1a**). This core may be further variably decorated with mannose, fucose, galactose, sialic acid, N-acetylgalactosamine (GalNAc) and N-acetylglucosamine (GlcNAc). In contrast to N-glycans, O-glycans are appended to the hydroxyl oxygen of Ser or Thr residues with no strong consensus sequence defining a glycosite. There are eight structural classes of O-glycans that are defined by core di- or tri-saccharides that occupy a glycosite (**Figure 1b**). Each of these cores can be further elaborated with other sugars yielding a large variety of possible O-glycan structures. Over 10% of secreted human proteins carry some form of O-glycan modification. A second form of O-glycosylation occurs on nuclear and cytoplasmic proteins, where a single β-linked GlcNAc is attached to Ser or Thr residues. β-O-GlcNAcylation is an essential, dynamic modification that is important in cell signaling and differentiation [6]. Finally, chemical groups (*e.g.*, sulfate, phosphate, acetate, methyl, etc.) may also occur at various positions on certain N- and O-glycan sugars [7].
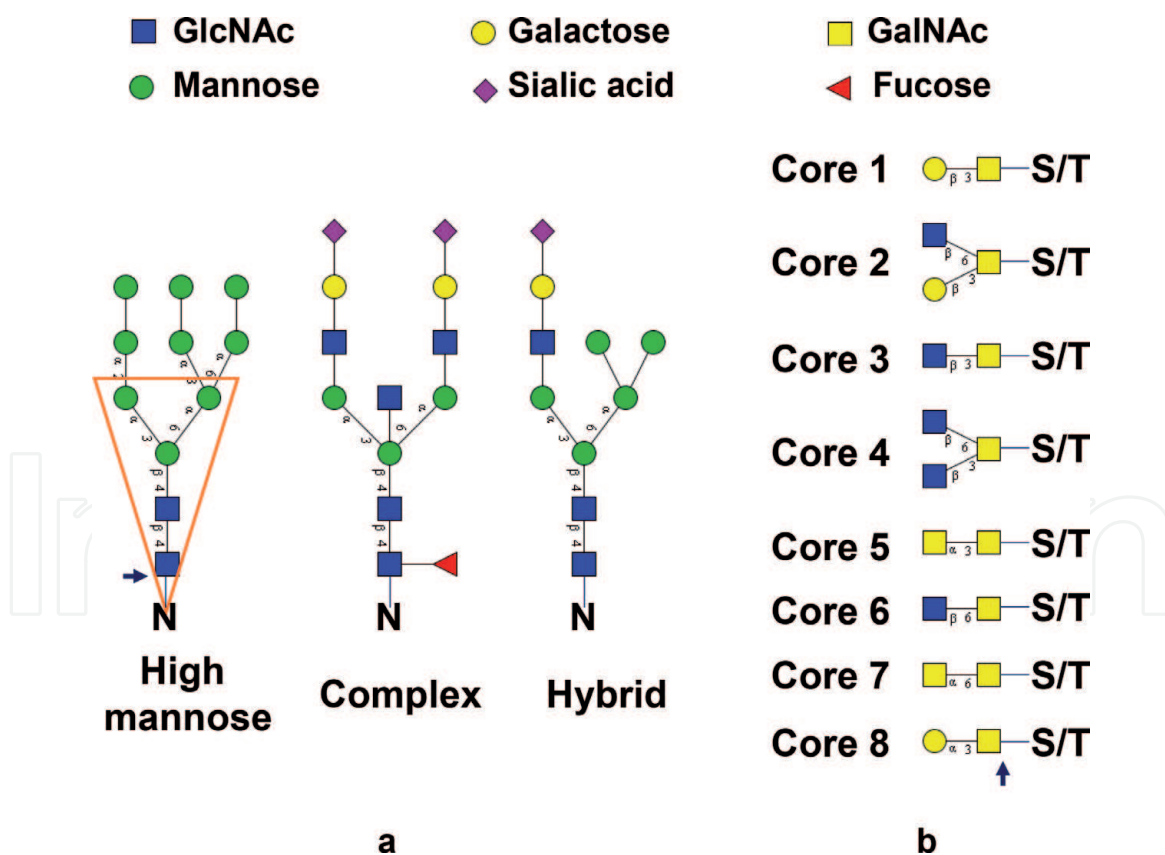

Figure 1.
Basic structures of N-glycans (a) and core structures of O-glycans (b). a, N-glycans can be categorized into three basic types: High mannose, Complex and Hybrid. The core structure (Man3GlcNAc2) of N-glycans is indicated by the orange triangle. A GlcNAc residue can attach to a β-mannose of the N-linked glycan core, resulting in a bisecting N-glycan (illustrated in Complex N-glycan). The reducing end GlcNAc (indicated by an arrow) of N-glycans can also be modified with a fucose (illustrated in Complex N-glycan). A N-glycan modifies a peptide via its reducing end GlcNAc attaching to Asparagine (N) within the peptide. b, eight core structures of O-glycans. O-glycan starts with a GalNAc (reducing end, indicated by an arrow), and further modifications can be added to the non-reducing end of the core structures. In O-glycopeptides, O-glycans are attached to the hydroxyl group of Serine (S) or Threonine (T) via the reducing end GalNAc.

Protein glycosylation is remarkable in its structural complexity. This trait reflects the way in which glycans are synthesized and transferred to proteins. Glycans are assembled by complex biosynthetic pathways consisting of many different enzymes. Individual monosaccharides become linked together by glycosyltransferases that each have sugar and stereochemical specificity. For example, the elaborate mammalian 14 sugar N-glycan precursor consists of only three types of monosaccharides (Glc3Man9GlcNAc2), yet its assembly requires the coordinated action of 13 different glycosyltransferases. There are over 200 different glycosyltransferases that affect glycan structures in the mammalian glycome [8]. Gene expression of some of these enzymes varies by tissue, cell type, and epigenetic regulation resulting in significant structural variation of glycans. A glycoform is a single protein isoform having a defined glycan present at each glycosite. As such, proteins naturally exist as collections of glycoforms. Additionally, some protein isoforms periodically lack glycan occupancy at a potential glycosite. The complexity of these attributes of glycoproteins underscores the technical challenges associated with deconvoluting any given glycome.

Analysis of glycan structure has been performed several different ways. However, the most common approaches typically utilize one of two strategies: (i) analysis of glycans that have been released from glycoproteins or ii) bottom-up proteomics analysis of peptide/glycopeptide mixtures. Standard N-glycan profiling methods begin with liberation of N-glycans from a glycoprotein with the enzyme PNGase F. Typically, they are then labelled at their reducing ends with a fluorophore, and separated via high/ultra-performance liquid chromatography (H/UPLC) or capillary electrophoresis (CE) with fluorescence detection and optional inline mass detection [9]. Glycan structures are assigned to observed peaks by comparing mobility and mass data to glycan reference databases [10]. Exoglycosidases with precise specificities can be used to further confirm structural assignments [11, 12]. For O-glycans, no enzyme that releases a broad range of elaborated O-glycan structures has been identified. Chemical release of O-glycans via hydrazinolysis can be achieved, but this can damage some released glycans [13, 14]. In addition, released N- and O-glycans may be permethylated and analyzed directly by LC–MS/MS or MALDI-MS [15]. Finally, for both N- and O-glycans, profiling of released glycans provides a catalog of the range of structures present in a sample, but it does not provide information regarding their point of attachment in a protein.

A more data-rich method of glycoprotein analysis uses bottom-up proteomics to analyze peptide/glycopeptide mixtures. In this approach, a glycoprotein is treated with a protease (*e.g.*, trypsin) to generate a pool of peptides that are then analyzed by mass spectrometry (typically LC–MS/MS). Data are processed by computer algorithms with the help of protein and glycan mass reference databases (*e.g.*, Byonic software and O-Pair Search) to generate a peptide map and identify appended glycans. Advantages of this method are that the same workflow can yield information about both N- and O-glycans (and other protein modifications), it identifies glycosites, it can determine both glycan occupancy and the range of glycan structures at each glycosite, and it can be quantitative. This approach (termed the 'multi-attribute method', MAM) is gaining traction in the pharmaceutical industry for monitoring the purity of biologic drugs and is expected to become the industry standard for final product characterization [16]. Despite its benefits, there are still technical challenges facing glycoproteomics analyses. For example, existing proteases (*e.g.*, trypsin) used in proteomics often generate large peptides that may have multiple glycans (especially for O-glycans that tend to be clustered within proteins). These generated glycopeptides can be either too large to detect by MS or it can be difficult to assign glycosites on such peptides with high confidence. Therefore, better approaches are needed to generate glycopeptides. Additionally,

glycopeptides represent a small portion of a peptide mixture and often do not ionize well. New methods that address sample complexity through enrichment of specific glycopeptides are emerging.

The field has begun to address these issues through development of new reagents that aid in glycoproteomics. Novel proteases, including those that have specificity for O-glycans, have recently been characterized and validated in glycoproteomics workflows. Additionally, reagents and methods that permit selective enrichment of glycopeptides have been applied to reduce sample complexity. In this chapter, we review advances in glycopeptide generation and enrichment methods that are helping to improve glycopeptide analysis. Additionally, we present an example case study illustrating N-glycopeptide enrichment to address glycan heterogeneity in Wnt signaling.

2. A workflow for intact glycopeptide identification

A common strategy to determine protein glycosylation is shown in **Figure 2**. This process generally involves: (i) protease treatment of protein(s) to generate a peptide/glycopeptide mixture, (ii) glycopeptide enrichment, (iii) analysis of isolated glycopeptides by LC–MS/MS, (iv) computational analysis of mass data

Protein and glycoprotein mixture

Proteolysis

Peptide and glycopeptide mixture

Enrichment
(e.g. HILIC, Fbs1-GYR, BGL, ExoO)

Enriched glycopeptides

LC-MS/MS

MS/MS spectra

Software tools
(e.g. Byonic, O-Pair Search)

**Intact glycopeptide identification
(peptide sequence and glycan composition)**

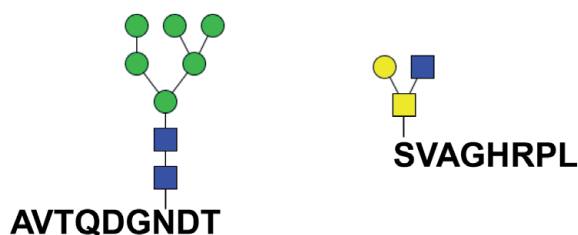**AVTQDGNDT**

SVAGHRPL

Figure 2.
Basic workflow of intact glycopeptide identification.

against proteome and glycan reference databases to yield both the peptide sequence and possible glycan structure for each peptide. Here we review technical challenges and recent advances for each step of glycopeptide analysis.

2.1 Peptide generation

To generate peptides for proteomics analyses, a protein sample is first digested with a protease. The specificity of the protease used can significantly impact the protein coverage obtained by the method. Trypsin, a protease that cleaves after lysine and arginine residues, has been the workhorse of the proteomics field for over two decades. Trypsin generally produces peptides of sufficient length to ionize efficiently in mass spectrometry. However, protein-specific challenges can occur with trypsin, especially with glycoproteins. For example, some proteins naturally lack lysine or arginine residues, have these residues disparately positioned, or have bulky glycans in close proximity that sterically hinder proteolysis. Each of these factors can produce larger peptides that typically do not ionize as well. As such, other proteases with cleavage specificities orthogonal to trypsin are often used to increase proteolytic peptide coverage (**Table 1**) [17]. For example, **Figure 3** shows that α-Lytic Protease can be used alone or in combination with other proteases to yield increased sequence coverage.

A recent advance has been the use of O-glycan-specific proteases (O-endoproteases) for generating O-glycopeptides for analysis. These enzymes recognize and bind to mucin-type O-glycans, then cleave the peptide bond immediately N-terminal to the glycosylated serine or threonine. Used either alone or in series with other proteases like trypsin, glycopeptides are generated that have an O-glycan on their amino-terminal amino acid following cleavage. The first commercial enzyme of this class was the O-endoprotease from *Akkermansia muciniphila* (sold under the trade name OpeRATOR, Genovis AB, Sweden). This enzyme recognizes mammalian O-glycans but it is inhibited by the presence of terminal sialic acids. Accordingly, sialidase treatment is required for efficient performance which results in loss of glycan structural information. Recently, chemical modification of sialic acids has also been shown to improve OpeRATOR function [18]. In contrast, the O-glycoprotease newly available from New England Biolabs, is not inhibited by the presence of sialic acids and it also exhibits a broad specificity towards proteins with mammalian

	Protease	Specificity
N-terminal cleavage	AspN	D (E)
	LysN	K
	LysargiNase	R, K
	O-endoprotease	S/T with O-glycan
C-terminal cleavage	ArgC	R (K)
	GluC	E (D)
	LysC	K
	Trypsin	K, R
	chymotrypsin	F, Y, L, W, M
	Pepsin	Y, F, W
	α-Lytic Protease	T, A, S, V (C, L)

Table 1.
Proteases used in proteomics. Protease specificities are indicated using single letter codes for amino acid residues. Recognition sites that are cleaved at a lower rate are indicated by amino acids bracketed by parentheses.
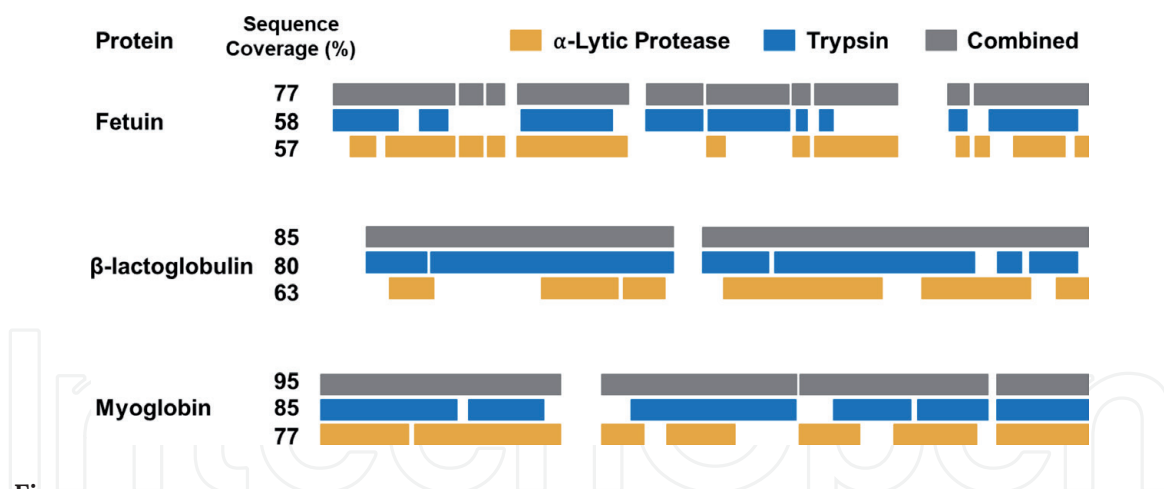
Figure 3.
α-Lytic Protease can be used alone or in combination with other proteases to yield increased sequence coverage. Comparison of sequence coverage for three protein standards after parallel digestion using Trypsin (blue) or a-Lytic Protease (gold). The combined data set (grey) results in overlapping peptides and increased sequence coverage. (Reprinted by permission from New England Biolabs. https://www.neb.com/products/ p8113-a-lytic-protease#Product%20Information).

O-glycans. This enzyme recognizes O-glycans ranging in size from a minimal GalNAc-α-Ser/Thr structure to larger mucin-type O-glycans bearing branches and sialic acids. This specificity negates the need for sialidase treatment or chemical modification prior to O-glycopeptide generation. Resulting O-glycopeptides can be mapped to identify the protein of origin, the position of O-glycosites, and the range of O-glycan structures present at any given glycosite in a single experiment.

2.2 Glycopeptide enrichment methods

Glycopeptides are typically in low abundance compared to aglycosylated peptides in a peptide mixture. Additionally, it is well-established that ionization of glycopeptides is often weaker compared to aglycosylated peptides during MS analyses [19]. This results in aglycosylated peptide signals often dominating MS experiments. Therefore, enrichment of glycopeptides prior to sample analysis has been a growing trend to improve intact glycopeptide identification. Several enrichment schemes that vary in their rationales have been described. These approaches range from general methods (enrichment of both N- and O-linked glycopeptides) to newer glycan class-specific approaches that selectively enrich for either N- or O-linked glycopeptides. Several approaches are summarized here.

2.2.1 Hydrophilic interaction liquid chromatography (HILIC)

HILIC has been widely used for glycopeptide enrichment. It is based on the interaction between the hydrophilic glycan moiety of a glycopeptide and the polar stationary phase in the non-polar mobile phase (typically acetonitrile). Many HILIC materials have been developed, however zwitterionic HILIC (ZIC-HILIC) enrichment is generally the most useful due to higher loading capacity and broader specificity. HILIC does not discriminate between O-linked and N-linked glycopeptides and hydrophilic non-glycosylated peptides may co-elute [20]. Thus, for more complete glycopeptide enrichment, HILIC may require a complementary chromatography fractionation step [21].

2.2.2 Boronic acid

One method utilizes boronic acid presented on a solid support to react with cis-diol-containing saccharides or polyols to form five- or six-membered cyclic

esters. This property has been used to capture glycoproteins and glycopeptides [22]. Importantly, the covalent linkage is reversible at acidic pH which results in release of intact glycopeptides [20]. The interaction between boronic acid and sugars is relatively weak but newly characterized derivatives show promise for enrichment of low-abundance glycopeptides [19]. A final consideration is that boronic acid enrichment does not discriminate between N- and O-linked glycopeptides.

2.2.3 Metal affinity chromatography

This method exploits the ability of negatively charged sialylated glycans to coordinate with titanium, zirconium or silver [23]. However, metal ion affinity chromatography is not strictly selective for sialylated glycans as negatively charged phosphopeptides or acidic peptides may compete for binding. Additionally, the method does not discriminate between N- and O-linked glycopeptides.

2.2.4 Hydrazide chemistry

Hydrazide chemistry has been widely used for glycosite characterization. Cis-diols within glycans of glycopeptides may be oxidized to aldehydes (using periodate oxidation) forming a non-reversible covalent bond with hydrazide immobilized on a bead. PNGase F is then used to release the formerly N-linked glycosylated peptides to enable N-glycosite determination using MS [24]. Although more commonly used for N-glycosite determination, this chemistry can also be used to enrich glycopeptides having sialylated glycans. In this method, mild periodate treatment selectively oxidizes sialic acids thus enabling capture of sialyated N- and O-glycopeptides on hydrazide beads. The intact glycopeptides can then be selectively released by acid hydrolysis and analyzed by MS [25].

2.2.5 Enzyme-mediated O-glycopeptide enrichment

O-glycopeptides may be enriched using an enzyme-based workflow termed "EXoO" (extraction of O-linked glycopeptides) [26]. This method is enabled by the availability of O-endoproteases (described above). The workflow (**Figure 4**) involves digestion of a protein/biological sample with a standard protease such as trypsin to generate a peptide mixture. The peptides are conjugated to a solid support via the terminal NH_2 group on each peptide (*e.g.*, Aminolink™ beads, ThermoFisher). An O-endoprotease is used to specifically release O-glycopeptides from the beads. Efficiency of the method is dependent on the specificity of the O-endoprotease. This approach may be practiced with OpeRATOR (Genovis) following chemical modification of sialic acids [18] or with O-glycoprotease (New England Biolabs) which cleaves without pre-treatment to remove or modify sialic acids.

2.2.6 Native lectin-mediated glycopeptide enrichment

Lectins are non-catalytic proteins that bind to carbohydrates. Lectins have been used in a variety of glycan, glycoprotein and glycopeptide enrichment strategies. A common approach utilizes broad-specificity bead-immobilized lectins to capture a wide spectrum of glycopeptides. For example, the lectins Concanavalin A (ConA) and wheat germ agglutinin (WGA) bind to high mannose structures and GlcNAc or sialic acid residues, respectively. Each has been used to isolate N-glycopeptides from peptide mixtures [27]. However, WGA does not exclusively bind to N-glycopeptides as it also binds O-β-GlcNAc found on intracellular proteins [28]. Similar strategies
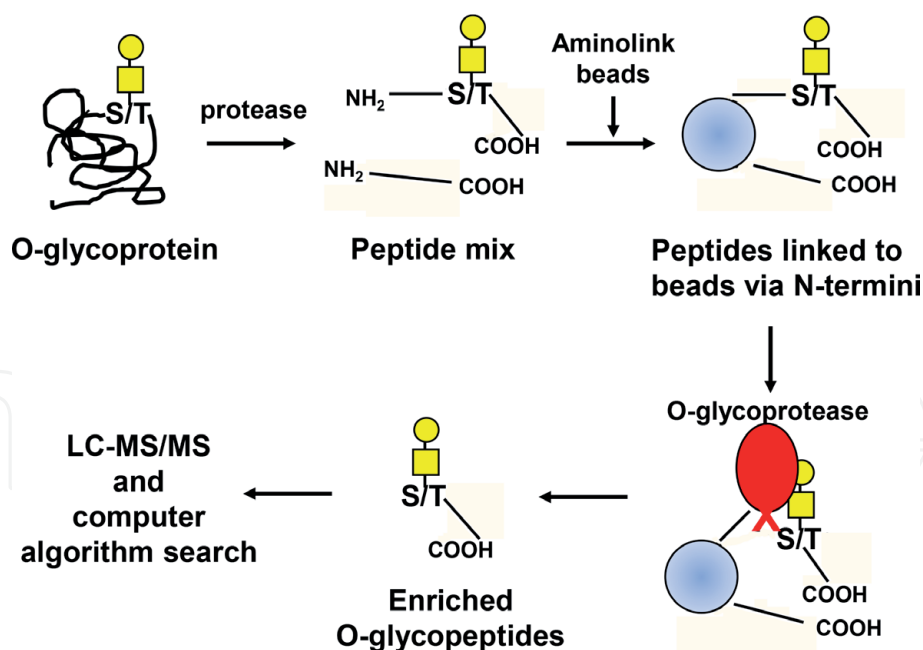
Figure 4.
Basic workflow of O-glycopeptide enrichment by O-glycoprotease.

have been applied to O-glycopeptide enrichment. For example, the lectins Jacalin and *Vicia villosa* agglutinin (VVA) bind to O-linked Gal(β-1,3)GalNAc and α- or β- linked terminal N-acetylgalactosamine, respectively [29, 30].

Lectin-based enrichment strategies have some limitations due to their natural properties. First, most lectins bind their substrates rather weakly (Kd of ~10 mM to 1 μM) [31]. Additionally, limitations in a lectin's specificity can introduce bias into an enrichment scheme. Strategies employing multiple lectins (multi-lectin affinity chromatography, M-LAC) have successfully increased glycopeptide recovery and coverage but do not completely solve the problem of lectin specificity bias [32]. To improve the performance of lectins in glycopeptide enrichment strategies, today's advanced capabilities for cloning and recombinant expression of lectins allows for mutagenesis and selection of lectins with improved binding properties.

2.2.7 Engineered lectins for N- and O-glycopeptide enrichment

The use of structure-guided protein engineering techniques has been used to create lectins with enhanced utility for glycopeptide enrichment. One area of interest has been to engineer binding proteins that can stratify a peptide mixture into different classes of glycopeptides (*e.g.*, N-glycopeptides or O-glycopeptides). Here we summarize recent progress in creating such reagents.

An ideal lectin for N-glycopeptide enrichment would bind to a structurally invariable portion of the N-glycan structure. A common trimannosyl chitobiose (Man3GlcNAc2) core glycan is a common feature of all N-glycans (**Figure 1a**). The human Fbs1 protein specifically recognizes this core motif [33, 34]. Fbs1 participates in glycoprotein quality control within the endoplasmic-reticulum-associated degradation (ERAD) system by binding to misfolded glycoproteins that have been retrotranslocated into the cytosol for degradation [35]. As part of the E3 ubiquitin complex, Fbs1 mediates ubiquitination and degradation of glycoproteins by the proteosome [33, 34]. Wild-type (wt) Fbs1 preferentially binds to high mannose N-glycans with sub-micromolar binding affinity (Kd of 0.1–0.2 μM) and only weakly binds to complex N-glycans having terminal sialic acids [36]. To adapt Fbs1 for use as a universal N-glycan/N-glycopeptide binding reagent, Fbs1 variants with greater tolerance for the presence of sialic acids were engineered using a novel

plasmid display strategy where library variants were enriched for their ability to bind immobilized fetuin [37]. An Fbs1 variant (termed Fbs1-GYR) containing S155G, F173Y and E174R substitutions was identified that efficiently binds to both high mannose N-glycans and complex N-glycans (**Figure 5**). Fbs1-GYR is unhindered by sialic acid and core fucose substitution, but does not bind to N-glycans bearing bisecting GlcNAc.

Fbs1-GYR is an efficient and substantially unbiased N-glycopeptide enrichment reagent. It enabled a deep characterization of the human serum N-glycoproteome [37] where Fbs1-GYR enrichment outperformed enrichment by the native lectin mixture of WGA, ConA and RCA$_{120}$ (WCR). Fbs1-GYR enrichment enabled identification of 2.2-fold more N-glycopeptides: an average of 2,142 N-glycopeptide spectra with Fbs1-GYR whereas enrichment with the WCR lectin mixture yielded an average of 965 N-glycopeptide spectra when the same amount of sample was analyzed by MS [37]. Fbs1-GYR mediated enrichment may be performed by using the N-glyco FASP method [32] or by using Fbs1-GYR immobilized beads. In the latter case, Fbs1-GYR has been expressed as a fusion to a SNAP-tag which permits covalent conjugation to benzyl-guanine beads [37–39].

A lectin (termed 'BGL') from the North American Kurokawa mushroom (*Boletopsis grisea*) was recently shown to have a specificity suitable for enrichment of a broad range of O-glycan and O-glycopeptide structures [40]. BGL is a member of the fungal fruit body lectins (Pfam PF07367) that possess two ligand binding sites, as verified by x-ray crystallography [41, 42]. One site binds to N-glycans possessing outer-arm terminal GlcNAc and the other to O-glycans bearing the TF-antigen disaccharide Galβ1,3GalNAc [40]. Ganatra *et al*. used structure-guided mutagenesis to generate single ligand binding site BGL variants [40]. One mutant BGL protein (R103Y) lost the ability to bind N-glycans with a terminal GlcNAc but retained the ability to bind O-glycans bearing the Galβ1,3GalNAc epitope. Both the R103Y BGL variant and *wt* BGL were shown to specifically isolate O-glycopeptides from proteolyzed fetuin, a peptide mixture that contains N-, O- and aglycosylated peptides [40]. As the R103Y BGL variant does not bind to N-glycans, it shows promise as a selective O-glycan/O-glycopeptide enrichment reagent (**Figure 6**). It is plausible that BGL (R103Y) and Fbs1-GYR could be used in tandem to stratify glycopeptide mixtures into enriched pools of O- or N-glycopeptides, respectively.
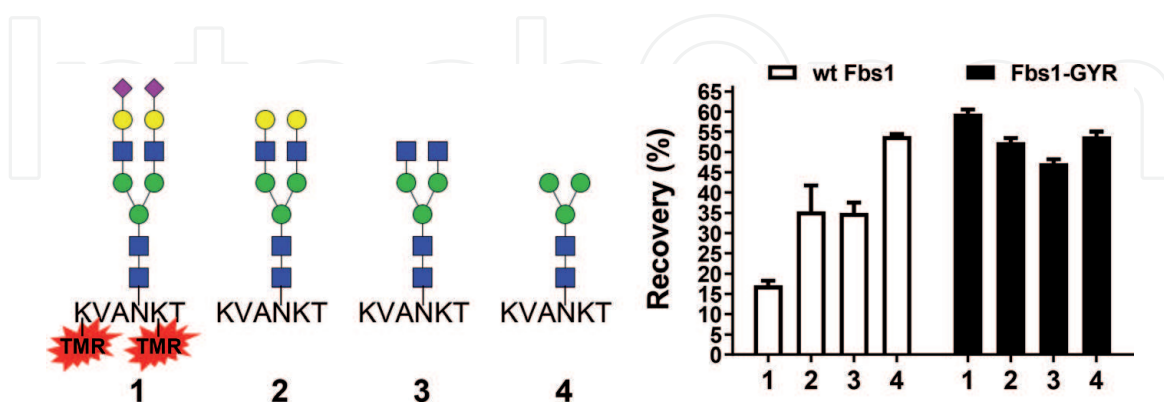

Figure 5.
Fbs1-GYR variant binding to a diverse set of N-glycopeptides is substantially unbiased. Sialylglycopeptide (SGP), an Fbs1 binding substrate, was fluorescently labeled with Tetramethylrhodamine (TMR) at the epsilon-amino group of lysine. For simplicity, TMR is only shown in N-glycopeptide structure 1. N-glycans of SGP-TMR (1) were trimmed with different combinations of exoglycosidases to produce asialo-SGP-TMR (2), SGP-TMR without sialic acids and galactose (3) and SGP-TMR without sialic acids, galactose and GlcNAc (4). The trimmed glycopeptides were then added to binding assays with wt Fbs1 or Fbs1-GYR beads in 50 mM ammonium acetate pH 7.5. The relative binding affinity to wt Fbs1 or Fbs1-GYR is reported as the recovery percentage (TMR fluorescence on beads/input TMR fluorescence). Results represent the mean ± s.e.m. of three replicates. (This figure was originally published within Nature Communications, Volume 8, Article number: 15487 (2017)).
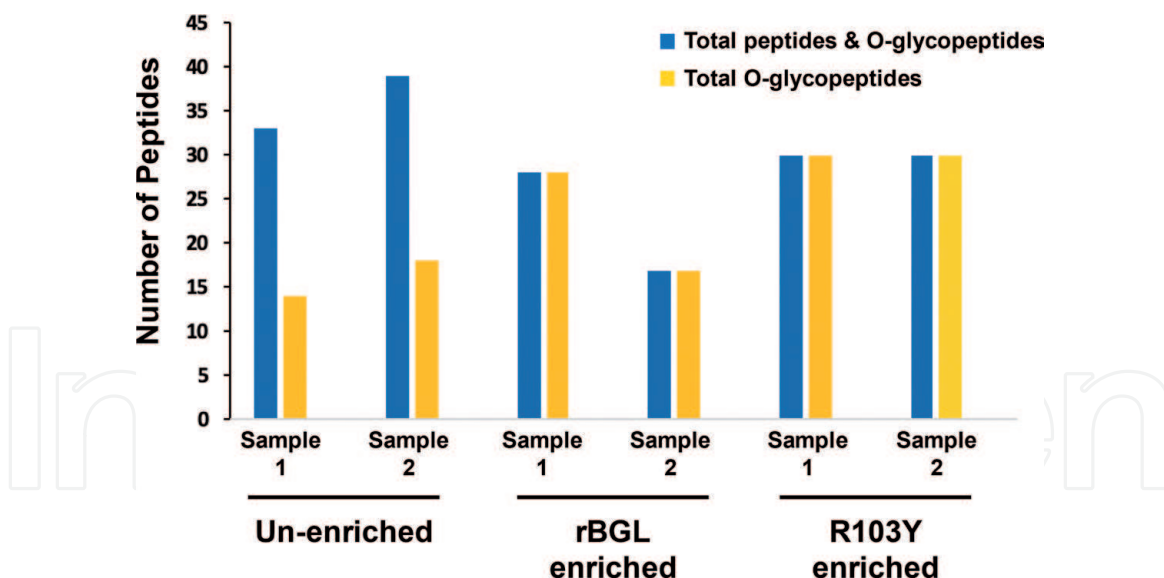
Figure 6.
Enrichment of O-glycosylated peptides/peptiforms from Pronase digested bovine fetuin before or after enrichment with BGL or BGL variant R103Y. Sample 1 and 2 represent replicate samples that were each separately digested with Pronase and subjected to lectin enrichment. Blue bars represent the total number of peptides identified (unglycosylated peptides and O-glycopeptides). Yellow bars represent the number of unique O-glycopeptides/peptiforms identified in each sample. (This figure was originally published within Scientific Reports, Volume 11: Article number: 160 (2021)).

2.3 LC-MS/MS and computer algorithms to search glycopeptides

2.3.1 LC-MS/MS

To identify intact glycopeptides, information of both the peptide backbone and the appended glycan is required. There are four major MS/MS fragmentation methods: collision induced dissociation (CID), electron-capture dissociation (ECD), electron transfer dissociation (ETD), and higher energy collisional dissociation (HCD). CID mainly fragments the peptide backbone, while ECD/ETD is more specific for glycan fragmentation. HCD can fragment both peptide backbone and glycan, and is widely used in intact glycopeptide MS/MS. A combination of different fragmentation methods can improve intact glycopeptide identification.

One recent study reported analysis of more than 5,600 glycopeptides and 1545 N-glycosites [43]. This report implemented a new type of tandem MS fragmentation: activated-ion electron transfer (AI-ETD). The analysis illustrated one of the first studies of glycoproteome profiling with AI-ETD on a quadrupole-Orbitrap-linear ion trap MS system (Orbitrap Fusion Lumos) [44]. Through specialized ion scanning routines, the authors acquired glycopeptide spectra with a higher-energy collision dissociation-product dependent-activated ion electron transfer dissociation (HCD-pd-AI-ETD). This strategy borrows from an established approach in N-glycopeptide analysis, HCD-product ion-triggered-ETD activation where abundant oxonium ions (*m/z* 204.087, HexNAc) in HCD MS/MS initiate subsequent ETD of the selected precursors [45, 46]. The new method HCD-pd-AI-ETD showed a median of peptide backbone sequence coverage of 89% and a median 78% glycan sequence coverage [44]. These parameters were derived from informatics tools with multiple filtering steps post-analysis. Overall, the filtering strategy aimed to attain no decoy peptide hits within the constraints of below 1% FDR estimations for both AI-ETD and HCD spectra.

2.3.2 Computer algorithms for intact glycopeptide identification

The complicated structure of intact glycopeptides makes the MS/MS spectra extremely complex. Therefore, special computer algorithms have been developed

to match the MS/MS spectra to both the peptide sequences and the attached glycan compositions. The algorithms include Byonic [47], GPQuest [48], pGlyco 2.0 [49], and O-pair Search [50]. Amongst these programs, the Byonic search engine provides high sensitivity identification of glycopeptides and allows the use of customized databases for both glycans and proteins. Byonic software identifies glycopeptides to the level of glycan composition and peptide sequence, and it is suitable for both N-glycopeptide and O-glycopeptide searches. A newly published computer algorithm, called O-Pair Search, is specific for O-glycopeptide searches [50]. The authors claim that O-Pair Search can not only greatly reduce search times (up to more than 2,000-fold) compared to a Byonic search, but it can also generate more O-glycopeptide identifications.

3. A case study: application of the Fbs1-GYR enrichment method to study N-glycan heterogeneity in Wnt signaling

The Wnt signaling pathway plays important roles in normal development and in cancer progression [51]. Several enzymes (such as DPAGT1) which are involved in N-glycan biogenesis are regulated by Wnt3a ligand stimulation [52, 53]. Therefore, an N-glycosylation study was performed to reveal potential biomarkers for the detection of Wnt-related cancers. We applied Fbs1-GYR enrichment technology to investigate whether protein N-glycosylation heterogeneity changes upon Wnt3a stimulation in mammalian cells.

Murine recombinant Wnt3a ligand and the Wnt Protein Stabilizer (AMS. bWps) (ASMBio, Cambridge, MA) in combination were able to stimulate canonical Wnt signaling in HEK293 SuperTopFlash STF cells (ATCC CRL-3249) in serum free media. Controls cells (non-Wnt3a stimulated cells) were treated in the same manner but without addition of Wnt3a. A 103.6 ± 3.92 (n = 3) fold change in TopFlash reporter gene expression was observed after 24-hour stimulation with 50 ng/ml Wnt3a and 50 μg/ml of Stabilizer. Note that serum free media was necessary to prevent possible glycoprotein contamination from bovine serum.

Control cells or Wnt3a-stimulated cells were harvested, and total protein was digested with trypsin. N-glycopeptides were enriched with 50 μg Fbs1-GYR purified protein using the N-glyco-FASP method [32] from 200 μg of tryptic peptides prepared from either control cells or Wnt3a-stimulated cells. The enriched N-glycopeptide samples were subjected to LC–MS/MS analysis and N-glycopeptide searching by Byonic as described [37]. 1556 and 1233 N-glycopeptide spectrum matches (N-glyco PSM) were obtained from Wnt3a-stimulated cells and control cells, respectively. (The complete dataset is available upon request from the corresponding author). The numbers of peptide spectrum matches (PSM) are suggestive of the relative abundance of the peptides [26]. Thus, the value of N-glyco PSM is used to evaluate and compare protein N-glycosylation in the Wnt3a-stimulated cells and the control cells. Using criteria of at least a two-fold change and a minimum 10 PSM difference between Wnt3a-stimulated and control cells, 17 proteins were identified exhibiting significant changes in N-glycosylation (**Table 2**). Among them, N-glycosylation of 11 proteins (MPRI, AN32B, PON2, SAP, NOMO3, TMED4, FKBP9, ATRN1, LMNB2, ZMAT4, and BASI) showed a significant increase upon Wnt3a stimulation, while N-glycosylation of 6 proteins (MA2B1, PLOD1, MPRD, MOMO1, HEAT1, and ABCAD) was significantly reduced with Wnt3a stimulation. MPRI (Cation-independent M6P receptor) and MPRD (Cation-dependent M6P receptor) are both mannose-6-phosphate (M6P) receptors. However, they display an opposite response to Wnt3a stimulation with regard to N-glycosylation (**Table 2**, **Table 3** highlighted within the dark blue box). The detected N-glycosylation of

	|Uniprot ID| N-glycoprotein identity	Spectral Counts		ratio
		Wnt3a	control	Wnt3a/Control
1	|P11717|MPRI_HUMAN Cation-independent mannose-6-phosphate receptor OS=Homo sapiens GN=IGF2R PE = 1 SV = 3	35	3	11.7
2	|Q92688|AN32B_HUMAN Acidic leucine-rich nuclear phosphoprotein 32 family member B OS=Homo sapiens GN = ANP32B PE = 1 SV = 1	20	6	3.3
3	|Q15165|PON2_HUMAN Serum paraoxonase/ arylesterase 2 OS=Homo sapiens GN=PON2 PE = 1 SV = 3	19	9	2.1
4	|O00754|MA2B1_HUMAN Lysosomal alpha-mannosidase OS=Homo sapiens GN = MAN2B1 PE = 1 SV = 3	3	15	0.2
5	|Q02809|PLOD1_HUMAN Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 OS=Homo sapiens GN=PLOD1 PE = 1 SV = 2	4	17	0.2
6	|P20645|MPRD_HUMAN Cation-dependent mannose-6-phosphate receptor OS=Homo sapiens GN = M6PR PE = 1 SV = 1	7	22	0.3
7	|P07602|SAP_HUMAN Prosaposin OS=Homo sapiens GN=PSAP PE = 1 SV = 2	63		
8	|P69849|NOMO3_HUMAN Nodal modulator 3 OS=Homo sapiens GN=NOMO3 PE = 2 SV = 2	18		
9	|Q7Z7H5|TMED4_HUMAN Transmembrane emp24 domain-containing protein 4 OS=Homo sapiens GN = TMED4 PE = 1 SV = 1	17		
10	|O95302–3|FKBP9_HUMAN Isoform 3 of Peptidyl-prolyl cis-trans isomerase FKBP9 OS=Homo sapiens GN=FKBP9	13		
11	|Q5VV63|ATRN1_HUMAN Attractin-like protein 1 OS=Homo sapiens GN = ATRNL1 PE = 2 SV = 2	11		
12	|Q03252|LMNB2_HUMAN Lamin-B2 OS=Homo sapiens GN = LMNB2 PE = 1 SV = 3	10		
13	|Q9H898|ZMAT4_HUMAN Zinc finger matrin-type protein 4 OS=Homo sapiens GN = ZMAT4 PE = 2 SV = 1	10		
14	|P35613|BASI_HUMAN Basigin OS=Homo sapiens GN=BSG PE = 1 SV = 2	10		
15	|Q15155|NOMO1_HUMAN Nodal modulator 1 OS=Homo sapiens GN=NOMO1 PE = 1 SV = 5		26	
16	|Q9H583|HEAT1_HUMAN HEAT repeat-containing protein 1 OS=Homo sapiens GN=HEATR1 PE = 1 SV = 3		17	
17	|Q86UQ4|ABCAD_HUMAN ATP-binding cassette sub-family A member 13 OS=Homo sapiens GN = ABCA13 PE = 2 SV = 3		14	

Table 2.
List of 17 proteins with significant differences in overall N-glycosylation upon Wnt3a stimulation.
N-glycosylation is evaluated by spectral counting label-free quantification. The scoring criteria was a minimum
of 10 PSM difference and a fold change minimum of 2 between Wnt3a-stimulated cells and the control cells.

>sp|Uniprot ID|N-glycoprotein name N-glycopeptide; N@denotes N-glycosite N-glycan attaching to the above N-glycopeptide	PSM in Control samples	PSM in Wnt3a samples
>sp|O00754|MA2B1_HUMAN Lysosomal alpha-mannosidase OS=Homo sapiens GN=MAN2B1 PE=1 SV=3	15	3
K.AN@LTWSVK.H	1	
HexNAc(2)Hex(5)	1	
K.LN@QTEPVAGNYYPVNTR.I	2	1
HexNAc(2)Hex(3)Fuc(1)	2	1
R.N@LSAPVTLNLR.D	1	
HexNAc(2)Hex(6)	1	
R.WWHQQTN@ATQEVVR.D	11	1
HexNAc(2)Hex(3)Fuc(1)	11	1
R.YYRTN@HTVMTMGSDFQYENANMWFK.N		1
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)Na(1)		1
>sp|P11717|MPRI_HUMAN Cation-independent mannose-6-phosphate receptor OS=Homo sapiens GN=IGF2R PE=1 SV=3	3	35
K.MN@FTGGDTCHK.V		13
HexNAc(2)Hex(6)		3
HexNAc(2)Hex(7)		5
HexNAc(2)Hex(8)		4
HexNAc(2)Hex(9)		1
K.TN@ITLVCKPGDLESAPVLR.T	1	20
HexNAc(2)Hex(5)		7
HexNAc(2)Hex(5)Fuc(1)		4
HexNAc(2)Hex(6)	1	9
R.N@GSSIVDLSPLIHR.T	2	2
HexNAc(2)Hex(9)	2	2
>sp|P20645|MPRD_HUMAN Cation-dependent mannose-6-phosphate receptor OS=Homo sapiens GN=M6PR PE=1 SV=1	22	7
R.LKPLFN@K.S	22	7
HexNAc(2)Hex(5)	6	
HexNAc(2)Hex(6)	10	2
HexNAc(2)Hex(7)	3	3
HexNAc(2)Hex(8)	3	2
>sp|Q9Y4L1|HYOU1_HUMAN Hypoxia up-regulated protein 1 OS=Homo sapiens GN=HYOU1 PE=1 SV=1	71	119
K.EN@GTDTVQEEEESPAEGSK.D	7	2
HexNAc(2)Hex(5)	4	1
HexNAc(2)Hex(6)	3	1
K.EN@GTDTVQEEEESPAEGSKDEPGEQVELK.E	1	
HexNAc(2)Hex(5)	1	
K.EN@GTDTVQEEEESPAEGSKDEPGEQVELKEEAEAPVEDGSQPPPPEPK.G	7	8
HexNAc(2)Hex(5)	4	3
HexNAc(2)Hex(5)Fuc(1)		1
HexNAc(2)Hex(6)	3	4
K.N@GTR.A	2	9
HexNAc(4)Hex(4)Fuc(1)	1	4
HexNAc(4)Hex(5)Fuc(1)	1	2
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)Na(1)		3
K.VIN@ETWAWK.N	14	20
HexNAc(2)Hex(5)	2	4
HexNAc(2)Hex(6)	6	8
HexNAc(2)Hex(7)	3	3
HexNAc(2)Hex(8)	1	1
HexNAc(2)Hex(9)	2	4
R.AEPPLN@ASASDQGEK.V	19	57
HexNAc(2)Hex(5)	7	11
HexNAc(2)Hex(6)	11	46
HexNAc(2)Hex(8)	1	
R.LSALDNLLN@HSSMFLK.G	2	4
HexNAc(2)Hex(6)	2	3
HexNAc(2)Hex(8)		1
R.VFGSQN@LTTVK.L	19	19
HexNAc(2)Hex(5)	2	
HexNAc(2)Hex(6)	3	4
HexNAc(2)Hex(7)	1	2
HexNAc(2)Hex(8)	5	5
HexNAc(2)Hex(9)	8	8

Table 3.
Comparison of detected N-glycosylation in MA2B1, MPRI, MPRD, and HYOU1 in Wnt3a-stimulated cells and the control cells. The light green rows indicate N-glycoprotein identity. Beneath the protein identity row, N-glycosites are listed in light blue. Beneath each N-glycosite, the respective N-glycan composition is listed. N@ indicates the asparagine with N-glycan modification. PSM numbers of individual N-glycosylation modifications are listed in columns on the right side.

MPRI increases 11.7 fold, while N-glycosylation of MPRD decreases approximately 3-fold after Wnt3a stimulation. The observed N-glycosylation changes may be due to the changes of protein expression level, which deserves further investigation.

The N-glyco PSM of lysosomal alpha-mannosidase (MA2B1(O00754)), is greatly reduced (5-fold) with Wnt3a stimulation. Interestingly, the N-glycosylation change is mainly due to differential glycosylation of N133 of MA2B1(O00754). **Table 3** shows 11 PSM with a fucosylated N-glycan (HexNAc(2)Hex(3)Fuc(1)) were found at position N133 attached to this mannosidase in control cells, but only one PSM was detected in the Wnt3a-stimulated cells (highlighted in the red box). Thus, we speculate that reduced N-glycosylation may affect stability of this mannosidase in the lysosome resulting in altered N-glycosylation of substrate proteins. The extent of N-glycosylation of some enzymes did not differ significantly between Wnt3a-stimulated cells and the control cells. However, modification of a specific N-glycosite did differ significantly. For example, there was no significant fold change with regard to the total numbers of N-glyco PSM of HYOU1, Hypoxia up-regulated protein 1, which were 71 and 119 in control and Wnt3a-stimulated cells, respectively (**Table 3**). However, a 3-fold increase in N-glycosylation at position N931 of HYOU1 was found upon Wnt3a stimulation (57 PSM in Wnt3a-stimulated cells vs. 19 PSM in control cells, **Table 3**, highlighted in the green box). Overall, this study demonstrates that the Fbs1-GYR enrichment method allows for the examination of glycosite heterogeneity of individual cellular proteins, and this study has revealed candidate biomarkers for Wnt-related cancers.

4. Conclusion

Due to the inherent complexity of protein glycosylation, better reagents and workflows are required in order to thoroughly and accurately characterize the glycosylation profile within a sample of interest. Intact glycopeptide identification (glycosite and glycan composition) has emerged as a more effective means to study heterogeneity, to investigate disease biomarkers and to characterize therapeutic proteins. Fortunately, several notable advances have arisen in the last few years. These advances include chemical enrichment strategies, engineered lectins with improved specificity, a greater selection of site-specific proteases, more sophisticated mass spectrometry methods/instruments and finally the development of computer algorithms designed for deconvolution of glycopeptide fragmentation spectra. Although challenges remain, these advances have certainly simplified the study of protein glycosylation.

Acknowledgements

Author details

Minyong Chen, Steven J. Dupard, Colleen M. McClung, Cristian I. Ruse,
Mehul B. Ganatra, Saulius Vainauskas, Christopher H. Taron
and James C. Samuelson*
New England Biolabs, Inc., Ipswich, MA, USA

*Address all correspondence to: samuelson@neb.com

IntechOpen

References

[1] Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. Cell surface protein glycosylation in cancer. Proteomics. 2014;14(4-5):525-46.

[2] Stuchlová Horynová M, Raška M, Clausen H, Novak J. Aberrant O-glycosylation and anti-glycan antibodies in an autoimmune disease IgA nephropathy and breast adenocarcinoma. Cell Mol Life Sci. 2013;70(5):829-39.

[3] Scott DW, Patel RP. Endothelial heterogeneity and adhesion molecules N-glycosylation: implications in leukocyte trafficking in inflammation. Glycobiology. 2013;23(6):622-33.

[4] Schedin-Weiss S, Winblad B, Tjernberg LO. The role of protein glycosylation in Alzheimer disease. Febs j. 2014;281(1):46-62.

[5] Mizushima T, Yagi H, Takemoto E, Shibata-Koyama M, Isoda Y, Iida S, et al. Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycans. Genes Cells. 2011;16(11):1071-80.

[6] Slawson C, Hart GW. O-GlcNAc signalling: implications for cancer cell biology. Nat Rev Cancer. 2011;11(9):678-84.

[7] Muthana SM, Campbell CT, Gildersleeve JC. Modifications of glycans: biological significance and therapeutic opportunities. ACS Chem Biol. 2012;7(1):31-43.

[8] Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 2012;13(7):448-62.

[9] Duke R, Taron CH. N-Glycan Composition Profiling for Quality Testing of Biotherapeutics. BioPharm International. 2015;28(12):59-64.

[10] Zhao S, Walsh I, Abrahams JL, Royle L, Nguyen-Khuong T, Spencer D, et al. GlycoStore: a database of retention properties for glycan analysis. Bioinformatics. 2018;34(18):3231-2.

[11] Rudd PM, Shi X, Taron CH, Walsh I. Recent Advances in the Use of Exoglycosidases to Improve Structural Profiling of N-glycans from Biologic Drugs. BioPharm International 2018;31 (10):16-23.

[12] Walsh I, Nguyen-Khuong T, Wongtrakul-Kish K, Tay SJ, Chew D, José T, et al. GlycanAnalyzer: software for automated interpretation of N-glycan profiles after exoglycosidase digestions. Bioinformatics. 2019;35(4):688-90.

[13] Kozak RP, Royle L, Gardner RA, Fernandes DL, Wuhrer M. Suppression of peeling during the release of O-glycans by hydrazinolysis. Anal Biochem. 2012;423(1):119-28.

[14] Merry AH, Neville DC, Royle L, Matthews B, Harvey DJ, Dwek RA, et al. Recovery of intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis. Anal Biochem. 2002;304(1):91-9.

[15] Wilkinson H, Saldova R. Current Methods for the Characterization of O-Glycans. Journal of Proteome Research. 2020;19(10):3890-905.

[16] Rogers RS, Abernathy M, Richardson DD, Rouse JC, Sperry JB, Swann P, et al. A View on the Importance of "Multi-Attribute Method" for Measuring Purity of Biopharmaceuticals and Improving Overall Control Strategy. AAPS J. 2017;20(1):7.

[17] Giansanti P, Tsiatsiani L, Low TY, Heck AJR. Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. Nature Protocols. 2016;11(5):993-1006.

[18] Yang S, Wu WW, Shen R, Sjogren J, Parsons L, Cipollo JF. Optimization of O-GIG for O-Glycopeptide Characterization with Sialic Acid Linkage Determination. Anal Chem. 2020;92(16): 10946-51.

[19] Suttapitugsakul S, Sun F, Wu R. Recent Advances in Glycoproteomic Analysis by Mass Spectrometry. Anal Chem. 2020;92(1):267-91.

[20] Chen CC, Su WC, Huang BY, Chen YJ, Tai HC, Obena RP. Interaction modes and approaches to glycopeptide and glycoprotein enrichment. Analyst. 2014;139(4):688-704.

[21] Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, et al. Site-Specific Glycan-Peptide Analysis for Determination of N-Glycoproteome Heterogeneity. Journal of Proteome Research. 2013; 12(12):5791-800.

[22] Xu Y, Wu Z, Zhang L, Lu H, Yang P, Webley PA, et al. Highly specific enrichment of glycopeptides using boronic acid-functionalized mesoporous silica. Anal Chem. 2009;81(1):503-8.

[23] Larsen MR, Jensen SS, Jakobsen LA, Heegaard NH. Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. Mol Cell Proteomics. 2007;6(10):1778-87.

[24] Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol. 2003;21(6):660-6.

[25] Nilsson J, Rüetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm G, et al. Enrichment of glycopeptides for glycan structure and attachment site identification. Nature Methods. 2009; 6(11):809-11.

[26] Yang W, Ao M, Hu Y, Li QK, Zhang H. Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). Mol Syst Biol. 2018;14(11):e8486.

[27] Ruiz-May E, Catalá C, Rose JK. N-glycoprotein enrichment by lectin affinity chromatography. Methods Mol Biol. 2014;1072:633-43.

[28] Ma J, Hart GW. O-GlcNAc profiling: from proteins to proteomes. Clin Proteomics. 2014;11(1):8.

[29] Sankaranarayanan R, Sekar K, Banerjee R, Sharma V, Surolia A, Vijayan M. A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a beta-prism fold. Nat Struct Biol. 1996;3(7):596-603.

[30] Wang K, Peng ED, Huang AS, Xia D, Vermont SJ, Lentini G, et al. Identification of Novel O-Linked Glycosylated Toxoplasma Proteins by *Vicia villosa* Lectin Chromatography. PloS one. 2016;11(3):e0150561-e.

[31] Fanayan S, Hincapie M, Hancock WS. Using lectins to harvest the plasma/serum glycoproteome. Electrophoresis. 2012;33(12):1746-54.

[32] Zielinska DF, Gnad F, Wiśniewski JR, Mann M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell. 2010;141(5):897-907.

[33] Mizushima T, Hirao T, Yoshida Y, Lee SJ, Chiba T, Iwai K, et al. Structural basis of sugar-recognizing ubiquitin ligase. Nat Struct Mol Biol. 2004;11(4): 365-70.

[34] Mizushima T, Yoshida Y, Kumanomidou T, Hasegawa Y, Suzuki A, Yamane T, et al. Structural basis for the selection of glycosylated substrates by SCF(Fbs1) ubiquitin ligase. Proc Natl Acad Sci U S A. 2007;104(14):5777-81.

[35] Yoshida Y, Mizushima T, Tanaka K. Sugar-Recognizing Ubiquitin Ligases:

Action Mechanisms and Physiology. Front Physiol. 2019;10:104.

[36] Hagihara S, Totani K, Matsuo I, Ito Y. Thermodynamic Analysis of Interactions between N-Linked Sugar Chains and F-Box Protein Fbs1. Journal of Medicinal Chemistry. 2005;48(9):3126-9.

[37] Chen M, Shi X, Duke RM, Ruse CI, Dai N, Taron CH, et al. An engineered high affinity Fbs1 carbohydrate binding protein for selective capture of N-glycans and N-glycopeptides. Nat Commun. 2017;8:15487.

[38] Juillerat A, Gronemeyer T, Keppler A, Gendreizig S, Pick H, Vogel H, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. Chem Biol. 2003; 10(4):313-7.

[39] Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat Biotechnol. 2003;21(1):86-9.

[40] Ganatra MB, Potapov V, Vainauskas S, Francis AZ, McClung CM, Ruse CI, et al. A bi-specific lectin from the mushroom Boletopsis grisea and its application in glycoanalytical workflows. Sci Rep. 2021;11(1):160.

[41] Carrizo ME, Capaldi S, Perduca M, Irazoqui FJ, Nores GA, Monaco HL. The Antineoplastic Lectin of the Common Edible Mushroom (Agaricus bisporus) Has Two Binding Sites, Each Specific for a Different Configuration at a Single Epimeric Hydroxyl*. Journal of Biological Chemistry. 2005;280(11):10614-23.

[42] Leonidas DD, Swamy BM, Hatzopoulos GN, Gonchigar SJ, Chachadi VB, Inamdar SR, et al. Structural Basis for the Carbohydrate Recognition of the Sclerotium rolfsii Lectin. Journal of Molecular Biology. 2007;368(4):1145-61.

[43] Riley NM, Hebert AS, Westphall MS, Coon JJ. Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. Nat Commun. 2019;10(1):1311.

[44] Riley NM, Westphall MS, Hebert AS, Coon JJ. Implementation of Activated Ion Electron Transfer Dissociation on a Quadrupole-Orbitrap-Linear Ion Trap Hybrid Mass Spectrometer. Anal Chem. 2017;89(12):6358-66.

[45] Saba J, Dutta S, Hemenway E, Viner R. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. Int J Proteomics. 2012;2012:560391.

[46] Singh C, Zampronio CG, Creese AJ, Cooper HJ. Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins. J Proteome Res. 2012;11(9):4517-25.

[47] Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. Curr Protoc Bioinformatics. 2012;Chapter 13: Unit13.20.

[48] Toghi Eshghi S, Shah P, Yang W, Li X, Zhang H. GPQuest: A Spectral Library Matching Algorithm for Site-Specific Assignment of Tandem Mass Spectra to Intact N-glycopeptides. Analytical Chemistry. 2015;87(10):5181-8.

[49] Liu M-Q, Zeng W-F, Fang P, Cao W-Q, Liu C, Yan G-Q, et al. pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. Nature Communications. 2017;8(1):438.

[50] Lu L, Riley NM, Shortreed MR, Bertozzi CR, Smith LM. O-Pair Search with MetaMorpheus for O-glycopeptide characterization. Nat Methods. 2020;17(11):1133-8.

[51] Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. Oncogene. 2017;36(11):1461-73.

[52] Sengupta PK, Bouchie MP, Kukuruzinska MA. N-glycosylation gene DPAGT1 is a target of the Wnt/beta-catenin signaling pathway. J Biol Chem. 2010;285(41):31164-73.

[53] Jamal B, Sengupta PK, Gao ZN, Nita-Lazar M, Amin B, Jalisi S, et al. Aberrant amplification of the crosstalk between canonical Wnt signaling and N-glycosylation gene DPAGT1 promotes oral cancer. Oral Oncol. 2012;48(6):523-9.