Recent advances in glycoprotein production for structural biology: toward tailored design of glycoforms

Yukiko Kamiya^{1,2,6}, Tadashi Satoh^{2,3}, and Koichi Kato^{1,2,4,5}

¹Okazaki Institute for Integrative Bioscience and Institute for Molecular Science, National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji, Okazaki, Aichi 444-8787, Japan,

²Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603, Japan, ³JST, PRESTO, 3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603, Japan, ⁴The Glycoscience Institute, Ochanomizu University, 2-1-1 Ohtsuka, Bunkyo-ku, Tokyo 112-8610, Japan, ⁵GLYENCE Co., Ltd., 2-22-8 Chikusa, Chikusa-ku, Nagoya 464-0858, Japan

⁶Present address: EcoTopia Science Institute and Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan.

Corresponding author:

Koichi Kato

Okazaki Institute for Integrative Bioscience and Institute for Molecular Science, National Institutes of Natural Sciences

5-1 Higashiyama, Myodaiji, Okazaki, Aichi 444-8787, Japan

Fax: 81-564-59-5224

Tel: 81-564-59-5225

Email: kkatonmr@ims.ac.jp

Short title: Tailored glycoprotein production for structural biology

Abstract

Because of the complexity, heterogeneity, and flexibility of the glycans, the structural analysis of glycoproteins has been eschewed until recently, with a few prominent exceptions. This aversion may have branded structural biologists as glycophobics. However, recent technological advancements in glycoprotein expression systems, employing genetically engineered production vehicles derived from mammalian, insect, yeast, and even bacterial cells, have yielded encouraging breakthroughs. The major advance is the active control of glycoform expression of target glycoproteins based on the genetic manipulation of glycan biogenetic pathways, which was previously overlooked, abolished, or considered unmanageable. Moreover, synthetic and/or chemoenzymatic approaches now enable the preparation of glycoproteins with uniform glycoforms designed in a tailored fashion.

Introduction

Glycosylation is one of the most common post-translational protein modifications; it is estimated that over half of proteins in nature are modified by carbohydrate moieties. Protein glycosylation is now under investigation not only as a fundamental concept of biology but also as an industrially important subject, because the efficacies of biopharmaceuticals critically depend on their glycoforms [1,2]. Additionally, carbohydrate–protein interactions are potential drug targets for the prevention of viral infections, cancer metastasis, and even neurodegenerative disorders [3,4]. Recent advances in glycoscience have yielded technical tools for structural and functional glycomics. Most genes encoding enzymes involved in glycobioynthesis have been identified and cloned, offering a technological platform [5]. Approaches utilizing mass spectrometry and chromatography have enabled comprehensive profiling of glycosylation at the molecular, cellular, and tissue levels and the consequent discovery of potential diagnostic biomarkers [6,7]. Various sugar libraries are currently available for the systematic analysis of carbohydrate–protein interactions [8–10[•]].

Given this situation, what is the current status of the collection of three-dimensional (3D)-structural data on glycoproteins? To date, the atomic coordinates of *N*-linked glycoproteins determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy represent only 3.5% (3,467 entries) of the total entries (97,789 entries; February 11, 2014) deposited in the Protein Data Bank (PDB; Figure 1a). Most of these glycoprotein structures (99.5%) were determined by X-ray crystallography, whereas the remainder (0.5%) was derived by NMR spectroscopy. Although the number of glycoprotein entries has grown significantly, there is a large gulf between PDB and the nature in terms of the proportions of glycosylated and non-glycosylated proteins. Moreover, the majority (95%) of carbohydrates attached to proteins in PDB comprise fewer than six residues (Figure 1b), although statistical analysis using the GLYCOSCIENCES.de database [11] indicates that the average chain size of mammalian oligosaccharides is eight residues [12]. These discrepancies suggest that the determination of glycoprotein structures to yield information about the functions of glycans remains challenging, despite technical advancements in glycomics.

In this review, we describe cutting-edge developments in the determination of glycoprotein structure and state-of-the-art technologies for the production and detailed structural analysis of glycoproteins. By taking advantage of progress in glycobiology and glycochemistry, our trend is now moving toward tailored protein glycosylation.

Recent progresses in glycoprotein X-ray crystallography

Needless to say, difficulties in the crystallographic analysis of glycoproteins are primarily

attributed to the heterogeneous and flexible properties of the outer branches of their glycans. Variability of the outer parts of *N*-glycans depends on the eukaryotic species they are produced in, because distinct glycosylation pathways exist in their Golgi apparatus (Figure 2). Insect cells [typified by SF9, SF21, BTI-Tn-5B1-4 (High Five™ cells), and *Drosophila melanogaster* S2 cells] have been widely used for the expression of glycoproteins for crystallographic analyses, because they produce sufficient quantities of glycoproteins containing relatively simple, homogeneous *N*-glycans, such as Man₃GlcNAc₂ and Man₃GlcNAc₂Fuc [13,14]. Their use prompted recent remarkable achievements in glycoprotein X-ray crystallography, including determination of the structures of the Wnt8/Frizzled-8 complex (Figure 3a) [15[•]], the brassinosteroid-insensitive 1 (BRI1)/somatic embryogenesis receptor-like kinase 1 (SERK1) complex [16[•]], and Toll-like receptor (TLR) 8 (Figure 3b) [17[•]]. In the case of TLR8, the presence of the endoplasmic reticulum (ER) α -mannosidase I inhibitor kifunensine in the culture medium resulted in recombinant proteins being exclusively modified with high-mannose-type glycans, which could then be cleaved using endoglycosidase H to facilitate crystallization.

In contrast to insect cells, the outer carbohydrate structures of mammalian *N*-glycans are considerably more variable and heterogeneous (Figure 2). Hence, minimizing heterogeneity is a key issue in glycoprotein crystallography [13,14]. Human embryonic kidney 293S GnTI [18] and Chinese hamster ovary (CHO)-Lec 3.2.8.1 [19] cells have been used to produce glycoproteins for crystallization, because they are deficient in *N*-acetylglucosaminyltransferase I (GnTI) activity and therefore capable of producing a homogeneously truncated *N*-linked glycoform (Man₅GlcNAc₂). Recent achievements in glycoprotein crystallography attained by exploiting these mammalian cell lines include determination of the structure of the α 5 β 1 integrin ectodomain (Figure 3c) [20], the human immunodeficiency virus (HIV)-1 gp120 V1-V2 region (Figure 3d) [21], and ectonucleotide pyrophosphatase/phosphodiesterase-1 [22].

Glycoform manipulation by genetically engineered mammalian production vehicles has been undertaken to improve the functional activities of glycoproteins, as exemplified by immunoglobulin G (IgG). For example, defucosylation of Fc glycans dramatically enhances their antibody-dependent cellular cytotoxicity by increasing their affinity for the Fc γ receptor III (Fc γ RIII) [23,24]. Afucosyl IgG can be produced in cells by knocking out the gene encoding α -1,6-fucosyltransferase (Fut8) [25] or by overexpressing β 1,4-*N*-acetylglucosaminyltransferase III, which catalyzes the addition of a bisecting GlcNAc residue to the core β -mannose [26], precluding the fucosylation catalyzed by Fut8. Use of these engineered cells permitted the determination of the crystal structure of afucosyl IgG-Fc in complexes with the extracellular domain of Fc γ RIII modified with two *N*-glycans (Figure 3e)

[27,28]. The crystal structure revealed that the *N*-glycan moieties of both Fc and Fc γ RIII mediated their interaction, which may explain the improved affinity of IgG glycoproteins for defucosylated Fc glycans. Recently reported crystal structures of α -L-iduronidase (produced in CHO cells or *Arabidopsis thaliana*) show that its *N*-glycan contributes to substrate recognition (Figure 3f) [29[•],30[•]]. Collectively, these studies demonstrate the importance of the integrity of carbohydrate moieties for elucidating the structural basis of protein function. Notably, mutational deglycosylation can induce an unexpected post-translational modification of glycoproteins, as exemplified by tyrosine sulfation of aglycosyl IgG-Fc [31]. Glycosyltransferase knock-in and knock-out approaches have also been used to produce glycoproteins with mammalian-type *N*-glycans in insect and yeast cells [32,33].

An alternative approach for glycoform remodeling is the enzymatic treatment of glycoproteins *in vitro*. For example, α 2,6-sialylated IgG, which is endowed with anti-inflammatory activity, was prepared by a two-step reaction with β 1,4-galactosyltransferase and α -2,6-sialyltransferase *in vitro* [34–36] and successfully subjected to crystal structure determination [37]. *In vitro* glycoform remodeling is also useful for the NMR analysis of *N*-glycans in conjunction with stable isotope-labeling techniques [38–40]. A more tailored approach to glycan remodeling has been proposed on the basis of enzymatic transglycosylation (*vide infra*).

Production of isotopically labeled glycoproteins for NMR analyses

In addition to high glycoprotein yield, the utility of stable isotope-labeling techniques is key to the NMR analysis of glycoproteins. Several review articles [40–43] have outlined methodologies and applications of uniform and selective isotope labeling of glycoproteins, based on metabolic labeling within eukaryotic cells or enzymatic attachment of isotope-labeled monosaccharides *in vitro*. As mentioned earlier, genetic engineering has been used to control the glycoforms of glycoproteins produced by insect and yeast cells: this, in conjunction with metabolic labeling, offers a promising tool for the preparation of isotope-labeled glycoproteins for NMR experiments. Protocols for uniform and amino acid-selective labeling have been developed for use in insect cells [44–46].

From a cost- and time-saving perspective, yeast cells are useful for the expression of isotopically labeled glycoproteins. More importantly, genetic manipulation protocols are well established in yeast. In wild type yeast, extensive attachment of mannose residues to high-mannose-type oligosaccharides is a hallmark of *N*-glycans (Figure 2). Such hypermannosylation could be prevented by knocking out the *OCH1* and *MNN1* genes, which encode α 1,6-mannosyltransferase and α 1,3-mannosyltransferase, respectively, produces the Man₈GlcNAc₂

glycoform [47], which could be further trimmed to $\text{Man}_5\text{GlcNAc}_2$ by co-expressing mannosidase from *Aspergillus saitoi* in engineered yeast [48]. As an alternative strategy to suppress hypermannosylation, $\text{Man}_5\text{GlcNAc}_2$ is exclusively produced as an *N*-glycan precursor in the dolichol pathway [49]. Engineered *Saccharomyces cerevisiae* cells lacking the genes encoding Mnn1p, Mnn4p, and Och1p produce glycoproteins with the homogeneous glycoform $\text{Man}_8\text{GlcNAc}_2$ [47,50]. Furthermore, additional deletion of the gene encoding ER mannosidase resulted in exclusive accumulation of glycoproteins with the $\text{Man}_9\text{GlcNAc}_2$ glycoform [51,52]. Indeed, sufficient quantities of ^{13}C -labeled homogeneous oligosaccharides have been isolated from glycoprotein extracts of cells for NMR characterization of their dynamic conformations (Figure 2b). Conversely, erythropoietin, IgG, and other glycoproteins carrying a homogeneous complex-type oligosaccharide ($\text{Sia}_2\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ or $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$) have been successfully produced in engineered yeast cells that harbor genes encoding Golgi glycosyltransferases [53,54]. Thus, engineered yeast represent promising vehicles for the overexpression of isotopically labeled glycoproteins with uniform glycoforms.

Bacterial expression of eukaryotic glycoproteins

Glycoprotein expression in prokaryotic cells is at an advanced stage. A key enzyme in this system is the *Campylobacter jejuni* oligosaccharyltransferase known as protein glycosylation B (PglB), which alone is capable of catalyzing *N*-glycosylation with $\text{Glc-GalNAc}_5\text{-Bac-PP-undecaprenyl}$ [55,56]. The *N*-glycosylation system of *C. jejuni* was successfully placed in *Escherichia coli* by genetic engineering to replace the reducing terminal bacillosamine with GlcNAc [57]. The bacterial glycans of the resultant recombinant glycoproteins were trimmed *in vitro* to generate a GlcNAc-Asn linkage as an acceptor site for enzymatic transglycosylation, allowing the production of eukaryotic *N*-glycoproteins, such as antibody fragments, with homogeneous glycoforms. It was also demonstrated that a cell-free system could be made from such genetically engineered *E. coli* for the production of *N*-linked glycoproteins [58[•]].

Further progress has been achieved through the creation of *E. coli* cells equipped with an engineered eukaryotic *N*-glycosylation pathway, comprising yeast uridine diphosphate-*N*-acetylglucosamine transferases and mannosyltransferases and *C. jejuni* PglB (Figure 4a) [59^{••}]. Using this engineered *E. coli*, $\text{Man}_3\text{GlcNAc}_2$ was successfully transferred to recombinant eukaryotic proteins, such as human growth hormone. Because the amino acid consensus sequence of the glycosylation site recognized by PglB (D-Q-N-A-T) is distinct from that recognized by mammalian oligosaccharyltransferase (e.g., N-X-S/T), mutational modifications are necessary at the

N-glycosylation site. This possible drawback would be overcome by engineering PglB so as to recognize the N-X-S/T sequon.

Synthetic approaches to creating homogeneous glycoproteins

Another emerging approach is the use of synthetic technologies to design and produce artificial glycoproteins. Procedures for the chemical synthesis of high-mannose- and complex-type oligosaccharides have become established during the last decades [60–62], permitting the chemical and chemoenzymatic creation of homogeneous glycoproteins. One conventional method for protein conjugation with synthetic oligosaccharides is disulfide bridge formation between iodo/bromoacetamide or methanethiosulfonate group at the reducing terminus and a recombinant protein harboring a cysteine substitution at the glycosylation site [63–66]. Artificial glycan acceptor sites could also be created by cell-free incorporation of keto group-containing unnatural amino acids using the amber codon suppression technique or via bacterial incorporation of azide- or alkene-containing methionine analogs using a Met (–) auxotrophic *E. coli* strain [67–69].

One of the most promising approaches for creating a natural glycan–protein junction is enzymatic transglycosylation at specific GlcNAc-Asn residues. This could be performed by engineered *E. coli* combined with *in vitro* glycosidase treatment (*vide supra*; Figure 4b). Several endo- β -*N*-acetylglucosaminidases (ENGases) have been employed for this purpose because of their efficient glycosylation of GlcNAc-containing proteins with synthetic *N*-glycans. To improve production yield, a synthetic *N*-glycan is chemically converted to an oxazoline analog, which serves as an activated donor substrate. Through the combined use of oxazoline donors and engineered ENGases with mutational removal of catalytic amino acid residues involved in product hydrolysis, homogeneous complex-type oligosaccharides were successfully attached to IgG with a sufficient yield [70–72[•]]. This approach may be useful for protein conjugation with synthetic oligosaccharides harboring structural biology probes, such as stable isotopes or selenium, at specific positions.

Kajihara et al. [65,73^{••},74^{••}] succeeded in creating an erythropoietin analogue as neo-glycoprotein and the third domain of ovomucoid using full chemical synthesis. Their approach involved the chemical ligation of peptides and a glycopeptide synthesized by the Fmoc-based solid-phase method. Although their folding yields require improvement for structural biology studies, they are proven excellent probes for characterizing enzymes involved in the quality control system in the ER [73^{••},74^{••}].

Conclusion and perspective

In typical studies on structural biology, once the 3D structure of a protein–protein complex is determined, one is prompted to confirm or evaluate the contributions of the observed atomic contacts to the affinity and specificity of the interactions, usually through site-directed mutations, such as alanine substitutions. However, glycoprotein interactions cannot always be characterized in this way, as recently reported crystal structures of glycoprotein–glycoprotein complexes have revealed that protein–protein interaction as well as protein–carbohydrate and even carbohydrate–carbohydrate interactions are significantly involved in complex formation [27,28]. To quantitatively characterize such interactions based on structural data, a systematic methodology is necessary, enabling tailored modifications of glycoprotein glycoforms.

Furthermore, dynamic conformational ensembles of glycans should also be considered in quantitative evaluations of the energetics of glycoprotein interactions. Because oligosaccharides have intrinsically flexible properties in their uncomplexed states, interactions involving their glycans considerably restrict their conformational space. NMR methods combined with simulations of molecular dynamics have now been developed to characterize the dynamic conformations of free oligosaccharides [75[•]]. At the next stage, glycoform-defined glycoproteins with isotopically labeled glycans would be subjected to detailed analyses of molecular dynamics.

In conclusion, the liaison between chemistry-based approaches and genetic engineering of cellular vehicles has paved the way for tailored design and production of natural and artificial glycoproteins for systematic structural glycobiology.

Acknowledgements

The authors thank Dr. Hirokazu Yagi (Nagoya City University) for valuable discussion and comments on this manuscript. This work was supported in part by the Okazaki ORION project, the Nanotechnology Network Project, and Grants-in-Aid for Scientific Research (Grant Numbers 24770102, 25121730 to T.S., and 25102008, 24249002 to K.K.) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by the CREST and PRESTO projects from the Japan Science and Technology Agency.

Figure legends

Figure 1. Current statistics regarding Protein Data Bank entries on *N*-linked glycoprotein structures (data were extracted on February 11, 2014). (a) Number and percentage of X-ray crystallography and NMR structures of *N*-linked glycoproteins. (b) Size distribution of *N*-glycans among glycoprotein structures revealed by X-ray crystallography (top) and NMR spectroscopy (bottom).

Figure 2. *N*-glycan processing pathways as the basis of glycoform-engineering strategies. (a) Schematic drawing of the *N*-glycan processing pathway in the ER common to yeast, insect, and mammalian cells, and the distinct pathways in the Golgi complex. In yeast, mannosyltransferases and mannosylphosphate transferases are involved in the formation of hypermannosyl oligosaccharides. In insect cells, high-mannose-type glycans are further trimmed by mannosidases in the Golgi complex to produce Man₅GlcNAc₂, which is then modified by adding a GlcNAc residue. Subsequent processing includes further mannose trimming, core fucosylation, and cleavage of the outer GlcNAc residue. Mammalian cells lack *N*-acetylglucosaminidase activity, and therefore can produce various complex-type glycans. (b) Metabolic ¹³C-labeling of the homogeneous high-mannose-type oligosaccharide Man₉GlcNAc₂ overexpressed in genetically engineered *S. cerevisiae* lacking the genes encoding Mnl1p, Och1p, Mnn1p, and Mnn4p [51]. The six ¹H-¹³C HSQC spectra were superimposed for the pyridylamino derivative of Man₉GlcNAc₂ metabolically labeled with D-[1-¹³C]glucose (black), D-[2-¹³C]glucose (blue), D-[3-¹³C]glucose (green), D-[4-¹³C]glucose (magenta), D-[5-¹³C]glucose (orange), or D-[6-¹³C]glucose (cyan). The ¹³C-labeled positions in the metabolic precursors (glucose isotopomers) are shown as circles in the same color as the corresponding spectrum. Adapted from reference no. 51 with permission from MDPI.

Figure 3. Recent prominent examples of glycoprotein structures revealed by X-ray crystallography. (a) *Xenopus laevis* Wnt8/mouse Frizzled-8 cysteine-rich domain complex (PDB code: 4F0A) [15[•]]. (b) Human toll-like receptor 8 (PDB code: 3W3J) [17[•]]. (c) Human α5β1 integrin–SG/19 Fab fragment complex (PDB code: 3VI3) [20]. (d) Human CD4–PGT135–17b Fab fragments–HIV-1 gp120 (PDB code: 4JM2) [21]. (e) Human α-L-iduronidase (PDB code: 4KGJ) [29[•],30[•]]. (f) Human IgG1-Fc–FcγRIII (PDB code: 3AY4) [27,28]. *N*-acetylglucosamine, mannose, galactose, and fucose residues are shown as blue, green, orange, and red spheres, respectively.

Figure 4. State-of-the-art approaches for the production of glycoproteins.

(a) Scheme of the *E. coli* expression system of a recombinant protein with eukaryotic *N*-glycans. This system employs PglB, the *C. jejuni* oligosaccharyltransferase, which transfers a GalNAc₂GlcGalNAc₃Bac oligosaccharide to protein, and *S. cerevisiae* asparagine-linked glycoprotein (Alg 1, Alg2, Alg13, and Alg14, which promote the expression of the trimannosyl glycan precursor. The *pgl* locus in *E. coli* was modified by deleting the genes involved in the biosynthesis and transfer of bacillosamine; this promotes the formation of a GlcNAc-containing donor substrate, which is primed by endogenous WecA. The *N*-glycosylation consensus sequence in the target protein expressed in *E. coli* was substituted with D-Q-N-A-T for *en bloc* *N*-glycan transfer catalyzed by PglB. (b) Chemoenzymatic synthesis of a glycoprotein containing a homogeneous glycan structure. Folding of chemically synthesized glycoproteins or enzymatic trimming of *N*-glycan residues on heterogeneous glycoforms on the protein provides an exposed GlcNAc-bearing protein. In the endoglycosidase-catalyzed transglycosylation reaction, the glycosidic oxygen is protonated, giving rise to a presumed sugar oxazolinium intermediate. Synthetic sugar oxazoline analogs corresponding to the outer *N*-glycan are excellent substrates for transglycosylation catalyzed by recombinant ENGases with a specific mutation in their active site (N175A, N171A, or D233A), which suppresses the hydrolysis of the product, thereby enabling *en bloc* transfer of well-defined glycans on to the protein. The glycan specificity of the reaction depends on the enzymes employed: high-mannose-type oligosaccharides and complex-type oligosaccharides are substrates of EndoA and EndoS, respectively, whereas high-mannose-type and bi-antennary-type oligosaccharides are substrates of EndoM.

References

1. Jiang XR, Song A, Bergelson S, Arroll T, Parekh B, May K, Chung S, Strouse R, Mire-Sluis A, Schenerman M: **Advances in the assessment and control of the effector functions of therapeutic antibodies.** *Nat Rev Drug Discov* 2011, **10**:101-111.
 2. Berkowitz SA, Engen JR, Mazzeo JR, Jones GB: **Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars.** *Nat Rev Drug Discov* 2012, **11**:527-540.
 3. Ernst B, Magnani JL: **From carbohydrate leads to glycomimetic drugs.** *Nat Rev Drug Discov* 2009, **8**:661-677.
 4. Kamiya Y, Yagi-Utsumi M, Yagi H, Kato K: **Structural and molecular basis of carbohydrate-protein interaction systems as potential therapeutic targets.** *Curr Pharm Des* 2011, **17**:1672-1684.
 5. Artemenko NV, McDonald AG, Davey GP, Rudd PM: **Databases and tools in glycobiology.** *Methods Mol Biol* 2012, **899**:325-350.
 6. Narimatsu H, Sawaki H, Kuno A, Kaji H, Ito H, Ikehara Y: **A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics.** *FEBS J* 2010, **277**:95-105.
 7. Adamczyk B, Tharmalingam T, Rudd PM: **Glycans as cancer biomarkers.** *Biochim Biophys Acta* 2012, **1820**:1347-1353.
 8. Liu Y, Palma AS, Feizi T: **Carbohydrate microarrays: key developments in glycobiology.** *Biol Chem* 2009, **390**:647-656.
 9. Rillahan CD, Paulson JC: **Glycan microarrays for decoding the glycome.** *Annu Rev Biochem* 2011, **80**:797-823.
 10. Kamiya Y, Satoh T, Kato K: **Molecular and structural basis for N-glycan-dependent determination of glycoprotein fates in cells.** *Biochim Biophys Acta* 2012, **1820**:1327-1337.
- Review article on the molecular and structural basis for the underlying mechanisms of glycoprotein-fate determination in cells that are achieved through the interactions with variety of intracellular lectins. Sugar binding specificities are comprehensively documented for the lectins involved in folding in the ER, vesicular transport, or ER-associated degradation. The authors also summarize the structural insights into how these lectins recognize specific glycotopes.
11. Lütteke T, Böhne-Lang A, Loss A, Goetz T, Frank M, von der Lieth CW: **GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research.** *Glycobiology* 2006, **16**:71R-81R.

12. Werz DB, Ranzinger R, Herget S, Adibekian A, von der Lieth CW, Seeberger PH: **Exploring the structural diversity of mammalian carbohydrates ("glycospace") by statistical databank analysis.** *ACS Chem Biol* 2007, **2**:685-691.
 13. Nettleship JE, Assenberg R, Diprose JM, Rahman-Huq N, Owens RJ: **Recent advances in the production of proteins in insect and mammalian cells for structural biology.** *J Struct Biol* 2010, **172**:55-65.
 14. Aricescu AR, Owens RJ: **Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology.** *Curr Opin Struct Biol* 2013, **23**:345-356.
 15. Janda CY, Waghray D, Levin AM, Thomas C, Garcia KC: **Structural basis of Wnt recognition by Frizzled.** *Science* 2012, **337**:59-64.
- This study provides important insights into how Wnt ligands interact with their receptors: the Frizzled family. Co-expression of an almost full-length *X. laevis* Wnt8 and the cysteine-rich domain of mouse Frizzled8 with an Fc affinity tag in *D. melanogaster* S2 cells enabled preparation of a glycosylated and lipidated Wnt-Frizzled complex.
16. Santiago J, Henzler C, Hothorn M: **Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases.** *Science* 2013, **341**:889-892.
- Interaction between brassinosteroid-insensitive 1 (BRI1) and somatic embryogenesis receptor-like kinase 1 (SERK1) triggers plant growth and differentiation. The authors describe the structural basis of the steroid hormone- (brassinosteroid-) dependent interaction between the BRI1 and SERK1 leucine-rich repeat domains. Intact BRI1 glycoprotein and SERK1 with mutations of two of the five N-glycosylation sites were expressed in High FiveTM insect cells.
17. Tanji H, Ohto U, Shibata T, Miyake K, Shimizu T: **Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands.** *Science* 2013, **339**:1426-1429.
- The authors show the crystal structures of the TLR 8 ectodomain in the absence of ligand and in complexes with three agonists. Ligand binding triggers conformational changes of a TLR8 dimer that juxtaposes the C-terminal domains, potentially initiating downstream signaling processes. TLR8 was expressed in *D. melanogaster* S2 cells in the presence of kifunensine, which exclusively produces endoglycosidase H-sensitive high-mannose-type glycans.
18. Reeves PJ, Callewaert N, Contreras R, Khorana HG: **Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line.** *Proc Natl Acad Sci U S A* 2002, **99**:13419-13424.

19. Stanley P: **Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity.** *Mol Cell Biol* 1989, **9**:377-383.
20. Nagae M, Re S, Mihara E, Nogi T, Sugita Y, Takagi J: **Crystal structure of $\alpha 5 \beta 1$ integrin ectodomain: atomic details of the fibronectin receptor.** *J Cell Biol* 2012, **197**:131-140.
21. Pancera M, Shahzad-Ul-Hussan S, Doria-Rose NA, McLellan JS, Bailer RT, Dai K, Loesgen S, Louder MK, Staupe RP, Yang Y, et al.: **Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16.** *Nat Struct Mol Biol* 2013, **20**:804-813.
22. Jansen S, Perrakis A, Ulens C, Winkler C, Andries M, Joosten RP, Van Acker M, Luyten FP, Moolenaar WH, Bollen M: **Structure of NPP1, an ectonucleotide pyrophosphatase/phosphodiesterase involved in tissue calcification.** *Structure* 2012, **20**:1948-1959.
23. Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, Weikert SH, Presta LG: **Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc γ RIII and antibody-dependent cellular toxicity.** *J Biol Chem* 2002, **277**:26733-26740.
24. Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, Uchida K, Anazawa H, Satoh M, Yamasaki M, et al.: **The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity.** *J Biol Chem* 2003, **278**:3466-3473.
25. Matsumiya S, Yamaguchi Y, Saito J, Nagano M, Sasakawa H, Otaki S, Satoh M, Shitara K, Kato K: **Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1.** *J Mol Biol* 2007, **368**:767-779.
26. Ferrara C, Brunker P, Suter T, Moser S, Puntener U, Umana P: **Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous $\beta 1,4$ -N-acetylglucosaminyltransferase III and Golgi α -mannosidase II.** *Biotechnol Bioeng* 2006, **93**:851-861.
27. Ferrara C, Grau S, Jäger C, Sondermann P, Brünker P, Waldhauer I, Hennig M, Ruf A, Rufer AC, Stihle M, et al.: **Unique carbohydrate-carbohydrate interactions are required for high affinity binding between Fc γ RIII and antibodies lacking core fucose.** *Proc Natl Acad Sci U S A* 2011, **108**:12669-12674.
28. Mizushima T, Yagi H, Takemoto E, Shibata-Koyama M, Isoda Y, Iida S, Masuda K, Satoh M,

- Kato K: **Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycans.** *Genes Cells* 2011, **16**:1071-1080.
29. Bie H, Yin J, He X, Kermode AR, Goddard-Borger ED, Withers SG, James MN: **Insights into mucopolysaccharidosis I from the structure and action of α -L-iduronidase.** *Nat Chem Biol* 2013, **9**:746.
- See reference [30[•]].
30. Maita N, Tsukimura T, Taniguchi T, Saito S, Ohno K, Taniguchi H, Sakuraba H: **Human α -L-iduronidase uses its own N-glycan as a substrate-binding and catalytic module.** *Proc Natl Acad Sci U S A* 2013, **110**:14628-14633.
- Together with reference [29[•]], this work demonstrates a very unique carbohydrate-carbohydrate interaction; human α -L-iduronidase uses its own N-glycan (ManD3 residue in the high-mannose-type glycan) for substrate recognition. Therefore, integrity of the N-glycan of this glycoprotein is crucial for its function. Two independent groups utilized different sources, *A. thaliana* [29[•]] or CHO cells [30[•]], for the crystallization.
31. Masuda K, Yamaguchi Y, Takahashi N, Jefferis R, Kato K: **Mutational deglycosylation of the Fc portion of immunoglobulin G causes O-sulfation of tyrosine adjacently preceding the originally glycosylated site.** *FEBS Lett* 2010, **584**:3474-3479.
32. Hamilton SR, Gerngross TU: **Glycosylation engineering in yeast: the advent of fully humanized yeast.** *Curr Opin Biotechnol* 2007, **18**:387-392.
33. Shi X, Jarvis DL: **Protein N-glycosylation in the baculovirus-insect cell system.** *Curr Drug Targets* 2007, **8**:1116-1125.
34. Kaneko Y, Nimmerjahn F, Ravetch JV: **Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation.** *Science* 2006, **313**:670-673.
35. Sondermann P, Pincetic A, Maamary J, Lammens K, Ravetch JV: **General mechanism for modulating immunoglobulin effector function.** *Proc Natl Acad Sci U S A* 2013, **110**:9868-9872.
36. Yu X, Vasiljevic S, Mitchell DA, Crispin M, Scanlan CN: **Dissecting the molecular mechanism of IVIg therapy: the interaction between serum IgG and DC-SIGN is independent of antibody glycoform or Fc domain.** *J Mol Biol* 2013, **425**:1253-1258.
37. Crispin M, Yu X, Bowden TA: **Crystal structure of sialylated IgG Fc: implications for the mechanism of intravenous immunoglobulin therapy.** *Proc Natl Acad Sci U S A* 2013, **110**:E3544-3546.
38. Barb AW, Prestegard JH: **NMR analysis demonstrates immunoglobulin G N-glycans are**

- accessible and dynamic. *Nat Chem Biol* 2011, **7**:147-153.
39. Barb AW, Meng L, Gao Z, Johnson RW, Moremen KW, Prestegard JH: **NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation.** *Biochemistry* 2012, **51**:4618-4626.
 40. Kato K, Yamaguchi Y, Arata Y: **Stable-isotope-assisted NMR approaches to glycoproteins using immunoglobulin G as a model system.** *Progress in Nuclear Magnetic Resonance Spectroscopy* 2010, **56**:346-359.
 41. Skrisovska L, Schubert M, Allain FH: **Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins.** *J Biomol NMR* 2010, **46**:51-65.
 42. Takahashi H, Shimada I: **Production of isotopically labeled heterologous proteins in non-*E. coli* prokaryotic and eukaryotic cells.** *J Biomol NMR* 2010, **46**:3-10.
 43. Kato K, Yamaguchi Y: **Glycoproteins and antibodies: Solution NMR studies.** In *Encyclopedia of Magnetic Resonance*. Edited by Harris RK, Wasylishen RE; 2012. vol 3, 1779-1790.
 44. Walton WJ, Kasprzak AJ, Hare JT, Logan TM: **An economic approach to isotopic enrichment of glycoproteins expressed from Sf9 insect cells.** *J Biomol NMR* 2006, **36**:225-233.
 45. Strauss A, Bitsch F, Fendrich G, Graff P, Knecht R, Meyhack B, Jahnke W: **Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells.** *J Biomol NMR* 2005, **31**:343-349.
 46. Gossert AD, Hinniger A, Gutmann S, Jahnke W, Strauss A, Fernández C: **A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility.** *J Biomol NMR* 2011, **51**:449-456.
 47. Nakanishi-Shindo Y, Nakayama K, Tanaka A, Toda Y, Jigami Y: **Structure of the N-linked oligosaccharides that show the complete loss of α -1,6-polymannose outer chain from *och1*, *och1 mnn1*, and *och1 mnn1 alg3* mutants of *Saccharomyces cerevisiae*.** *J Biol Chem* 1993, **268**:26338-26345.
 48. Chiba Y, Suzuki M, Yoshida S, Yoshida A, Ikenaga H, Takeuchi M, Jigami Y, Ichishima E: **Production of human compatible high mannose-type (Man₅GlcNAc₂) sugar chains in *Saccharomyces cerevisiae*.** *J Biol Chem* 1998, **273**:26298-26304.
 49. Davidson RC, Nett JH, Renfer E, Li H, Stadheim TA, Miller BJ, Miele RG, Hamilton SR, Choi BK, Mitchell TI, et al.: **Functional analysis of the *ALG3* gene encoding the**

- Dol-P-Man: Man₅GlcNAc₂-PP-Dol** mannosyltransferase enzyme of *P. pastoris*. *Glycobiology* 2004, **14**:399-407.
50. Kamiya Y, Yamamoto S, Chiba Y, Jigami Y, Kato K: **Overexpression of a homogeneous oligosaccharide with ¹³C labeling by genetically engineered yeast strain.** *J Biomol NMR* 2011, **50**:397-401.
 51. Kamiya Y, Yanagi K, Kitajima T, Yamaguchi T, Chiba Y, Kato K: **Application of metabolic ¹³C labeling in conjunction with high-field nuclear magnetic resonance spectroscopy for comparative conformational analysis of high mannose-type oligosaccharides.** *Biomolecules* 2013, **3**:108-123.
 52. Yamaguchi T, Kamiya Y, Choo Y-M, Yamamoto S, Kato K: **Terminal spin labeling of a high-mannose-type oligosaccharide for quantitative NMR analysis of its dynamic conformation.** *Chem. Lett.* 2013, **42**:544-546.
 53. Li H, Sethuraman N, Stadheim TA, Zha D, Prinz B, Ballew N, Bobrowicz P, Choi BK, Cook WJ, Cukan M, et al.: **Optimization of humanized IgGs in glycoengineered *Pichia pastoris*.** *Nat Biotechnol* 2006, **24**:210-215.
 54. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, Bobrowicz P, Stadheim TA, Li H, Choi BK, et al.: **Humanization of yeast to produce complex terminally sialylated glycoproteins.** *Science* 2006, **313**:1441-1443.
 55. Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I, Aebi M: **N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase.** *Science* 2006, **314**:1148-1150.
 56. Slynko V, Schubert M, Numao S, Kowarik M, Aebi M, Allain FH: **NMR structure determination of a segmentally labeled glycoprotein using in vitro glycosylation.** *J Am Chem Soc* 2009, **131**:1274-1281.
 57. Schwarz F, Huang W, Li C, Schulz BL, Lizak C, Palumbo A, Numao S, Neri D, Aebi M, Wang LX: **A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation.** *Nat Chem Biol* 2010, **6**:264-266.
 58. Guarino C, DeLisa MP: **A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins.** *Glycobiology* 2012, **22**:596-601.
- This article illustrates the cell-free production of N-linked glycoproteins accomplished by use of lysates of the engineered *E. coli* that express PglB from *C. jejuni*.
59. Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi P, Aebi M, DeLisa MP: **An engineered eukaryotic protein glycosylation pathway in**

Escherichia coli. *Nat Chem Biol* 2012, **8**:434-436.

●●In this article, the authors successfully demonstrated the *E. coli* expression system of glycoproteins with eukaryotic trimannosyl chitobiose glycans. This intriguing achievement was made by installing the yeast *N*-acetylglucosamine transferases Alg13 and Alg14 and mannosyltransferases Alg1 and Alg2 and the *C. jejuni* oligosaccharyltransferase general glycosylation pathway protein into *E. coli*.

60. Wang Z, Chinoy ZS, Ambre SG, Peng W, McBride R, de Vries RP, Glushka J, Paulson JC, Boons GJ: **A general strategy for the chemoenzymatic synthesis of asymmetrically branched *N*-glycans**. *Science* 2013, **341**:379-383.
61. Lepenies B, Yin J, Seeberger PH: **Applications of synthetic carbohydrates to chemical biology**. *Curr Opin Chem Biol* 2010, **14**:404-411.
62. Takeda Y, Totani K, Matsuo I, Ito Y: **Chemical approaches toward understanding glycan-mediated protein quality control**. *Curr Opin Chem Biol* 2009, **13**:582-591.
63. Macmillan D, Bill RM, Sage KA, Fern D, Flitsch SL: **Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin**. *Chem Biol* 2001, **8**:133-145.
64. Watt GM, Lund J, Levens M, Kolli VS, Jefferis R, Boons GJ: **Site-specific glycosylation of an aglycosylated human IgG1-Fc antibody protein generates neoglycoproteins with enhanced function**. *Chem Biol* 2003, **10**:807-814.
65. Hirano K, Macmillan D, Tezuka K, Tsuji T, Kajihara Y: **Design and synthesis of a homogeneous erythropoietin analogue with two human complex-type sialyloligosaccharides: combined use of chemical and bacterial protein expression methods**. *Angew Chem Int Ed Engl* 2009, **48**:9557-9560.
66. Fernández-González M, Boutureira O, Bernardes GJL, Chalker JM, Young MA, Errey JC, Davis BG: **Site-selective chemoenzymatic construction of synthetic glycoproteins using endoglycosidases**. *Chemical Science* 2010, **1**:709-715.
67. Liu H, Wang L, Brock A, Wong CH, Schultz PG: **A method for the generation of glycoprotein mimetics**. *J Am Chem Soc* 2003, **125**:1702-1703.
68. van Kasteren SI, Kramer HB, Jensen HH, Campbell SJ, Kirkpatrick J, Oldham NJ, Anthony DC, Davis BG: **Expanding the diversity of chemical protein modification allows post-translational mimicry**. *Nature* 2007, **446**:1105-1109.
69. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR: **Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation**. *Proc Natl Acad Sci U*

SA 2002, **99**:19-24.

70. Wang LX, Lomino JV: **Emerging technologies for making glycan-defined glycoproteins.** *ACS Chem Biol* 2012, **7**:110-122.

71. Fan SQ, Huang W, Wang LX: **Remarkable transglycosylation activity of glycosynthase mutants of endo-D, an endo- β -N-acetylglucosaminidase from *Streptococcus pneumoniae*.** *J Biol Chem* 2012, **287**:11272-11281.

72. Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX: **Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions.** *J Am Chem Soc* 2012, **134**:12308-12318.

●The authors demonstrate that mutated ENGases from *Streptococcus pyogenes* (EndoS) are capable of transferring sialylated and asialylated complex-type *N*-glycans to fucosylated GlcNAc-expressing proteins.

73. Kajihara Y, Tanabe Y, Sasaoka S, Okamoto R: **Homogeneous human complex-type oligosaccharides in correctly folded intact glycoproteins: evaluation of oligosaccharide influence on protein folding, stability, and conformational properties.** *Chemistry* 2012, **18**:5944-5953.

●●See reference [74^{●●}].

74. Izumi M, Makimura Y, Dedola S, Seko A, Kanamori A, Sakono M, Ito Y, Kajihara Y: **Chemical synthesis of intentionally misfolded homogeneous glycoprotein: a unique approach for the study of glycoprotein quality control.** *J Am Chem Soc* 2012, **134**:7238-7241.

●●Together with reference [73^{●●}], this article describes the demonstration of fully chemical synthesis of glycoproteins with homogeneous glycans. Using their pioneering approach, the authors produced folded and unfolded glycoproteins and characterized the substrate recognition properties of the ER enzyme UGGT, which senses the folding states of glycoproteins in the quality control system in the ER.

75. Zhang Y, Yamaguchi T, Kato K: **New NMR tools for characterizing the dynamic conformations and interactions of oligosaccharides.** *Chem. Lett.* 2012, **42**: 1455-1462.

●Review article providing overview of recently developed NMR methods for characterizing dynamic conformations and interactions of oligosaccharides.

Figures

(a)

(b)

Figure 1

Figure 2

Figure 3

Figure 4