7-1-2023

Glycomedicine: The current state of the art

Wei Wang
Edith Cowan University
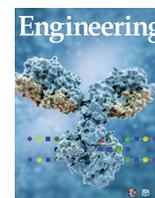
Views & Comments

Glycomedicine: The Current State of the Art

Wei Wang [a,b,c,d]

[a] Centre for Precision Health, Edith Cowan University, Perth, WA 6027, Australia
[b] Beijing Municipal Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing 100069, China
[c] School of Public Health, Shandong First Medical University & Shandong Academy of Medical Sciences, Tai'an 271016, China
[d] The First Affiliated Hospital of Shantou University Medical College, Shantou 515041, China

Life requires more than nucleic acids and proteins; sweet sugar molecules could be another life code beyond the central dogma of molecular biology.

There are four equally important major building blocks of life: nucleic acids (DNA and RNA), proteins, carbohydrates (glycans), and lipids. The first two are also known as the first and second alphabets of biology, following the principle of the "central dogma" of transcription (DNA to RNA) and translation (RNA to protein). However, the latter two crucial components, glycans and lipids, are missing from biology's central dogma. Regarding the communication between glycans and lipids, there may be a yet-to-be-discovered law: Does a paracentral dogma exist? This commentary focuses on glycans, the third alphabet of life, and their role in the sociomateriality of the cell, which provides a novel dimension of medical science—glycomedicine. This is an allied new discipline that employs glycomics approaches with the aim of better targeting disease diagnostics, as well as drug discovery, prescription choice, and dosing based on individual glycomics profiles to enable preventive, predictive, and precision medicine.

Glycans have broad physiological significance as signaling molecules, in addition to serving as a source of energy, a component of the nucleic acid backbone, and the biological cement of the cell wall structures (i.e., the glycocalyx). These executive molecules fine-tune strategic intra- and inter-cell communication, coordinate biological networks, and recognize the host versus foreign cells such as infectious agents (virus and bacteria) or tissues from another organism (in the case of organ transplantation) [1–3]. Therefore, glycans play key roles in health, suboptimal health, and illness, as they mediate temporal–spatial informed responses to hostile or benevolent changes in the intercellular space as well as sensing and interacting with the broader microcellular environment (Fig. 1) [4].

Glycosylation, the covalent attachment of sugar moieties to proteins, is a significant process in the endoplasmic reticulum and the Golgi apparatus [1]. More than 50% of all proteins (> 85% of secretory proteins) within the cell undergo such post-translation modifications (PTMs) [1,5,6]. When bound to proteins, sugar moieties affect their structure, function, stability, folding, half-life, trafficking, solubility, and interactions with other proteins. Glycans vary in terms of glycosylic linkages, the position of the hydroxyl group on the anomeric carbon, the number and type of constituent monosaccharides, and the degree of branching. Subsequently, their attachment to a protein increases the complexity of the proteome. Based on how glycans are bound to proteins, glycosylation can be classified into ① N-glycosylation, ② O-glycosylation, ③ C-glycosylation, ④ Glypiation, and ⑤ Phosphoglycosylation [1,5,6]. Each of these glycan structures can be identified by its core structure, its consensus sequence, and how the constituent monosaccharides are branched or spatially arranged on the protein. *N*-glycans bind to asparagine residues, *O*-glycans bind to the hydroxyl groups of serine (Ser) and threonine (Thr), and *C*-glycans bind to tryptophan residues. Glypiation occurs when glycans bind to phospholipids, and phosphoglycosylation takes place when glycans are bound to Ser through phosphodiester bonds [1,5,6]. Proteins are also glycosylated through non-enzymatic glycation, in which glucose (in its aldehyde form) reacts with the arginine and lysine residues in proteins and undergoes further changes. This eventually results in advanced glycation end products such as glycated hemoglobin (HbA1c), thus playing important functions in biological aging and the pathogenesis of disease (e.g., diabetes) [7].

Unlike proteins, whose formation follows the processes of the transcription and translation of genes, glycans are formed without requiring a template, and their synthesis involves multiple enzymes that respectively add monosaccharide units (glycosyltransferases) to proteins or subtract monosaccharide units (glycosidases) from proteins. The structures and catalytic mechanisms of glycosyltransferases involve common structural scaffolds with distinct acceptor substrate specificities, which are partly conferred by variable loop regions extending from the core catalytic unit. Furthermore, the diversified functions of glycosyltransferases involve mutations in the common core sugar nucleotide-binding region and varying loop regions, which drive the divergence in donor sugars and acceptor substrate recognition, respectively. Glycosyltransferases that catalyze the transfer of activated sugars, also termed donor substrates, have high nucleotide specificity, although they may have some flexibility for the donor glycans, and generally form only one type of glycosidic linkage structure.

The final outcome of glycosylation is influenced by many other factors—including the availability of substrates and glycans,
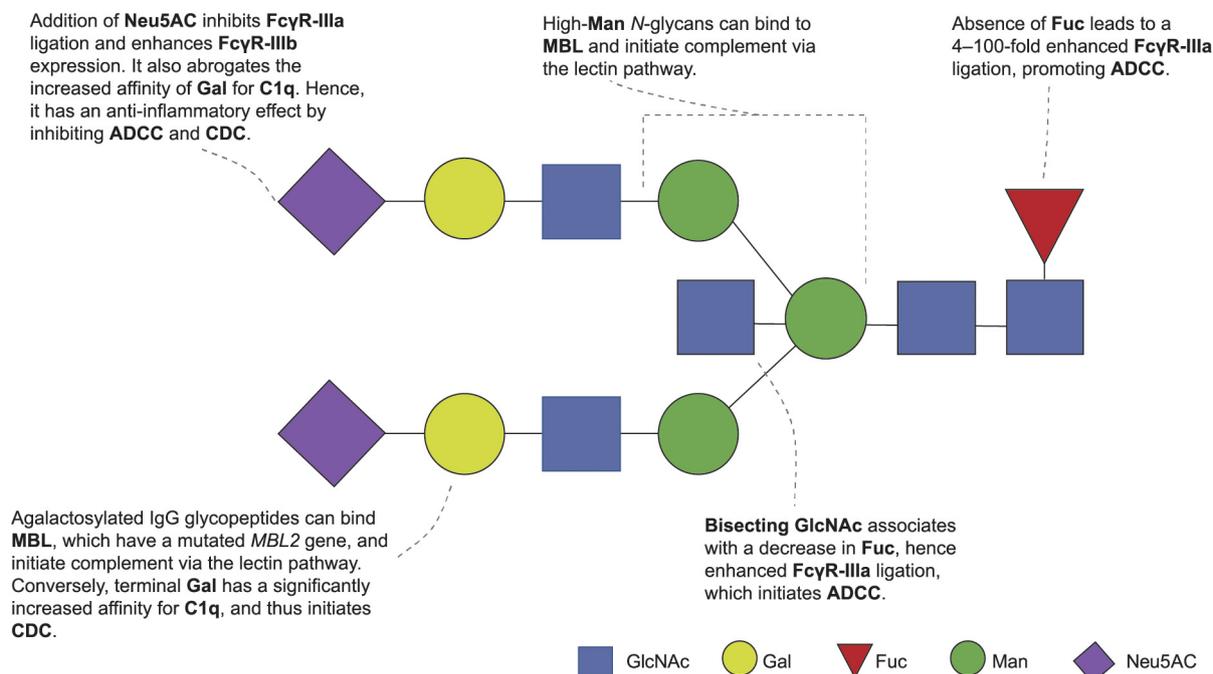
Fig. 1. An illustrated summary of altered immunoglobulin G (IgG) glycosylation and its downstream effects. GlcNAc: *N*-acetylglucosamine; Gal: galactose; Fuc: core fucose; Man: mannose; Neu5Ac: *N*-acetylneuraminic acid (sialic acid); ADCC: antibody-dependent cellular cytotoxicity; CDC: complement-dependent cytotoxicity; C1q: complement 1q; MBL: mannose-binding lectin; FcγR: fragment crystallizable region γ receptor for immunoglobulin. Reproduced from Ref. [4] with permission of Springer, © 2021.

competing glycosylation reactions, co-factors (e.g., Mn^{2+}), intracellular transport, pH changes, chaperone protein and glycosidase action, stress, and other general factors—which may affect the normal cellular state. For example, the *N*-acetylgalactosamine (GalNAc)-type O-glycosylation of Ser/Thr is initiated in the Golgi apparatus by the involvement of up to 20 polypeptide GalNAc transferase isomers with distinct and partly overlapping specificities [8,9]. It leads to the generation of a simple GalNAcα1–*O*-Ser/Thr monosaccharide structure known as the cancer-associated Tn antigen [8].

Although the human genome contains about 700 genes encoding the enzymes, transporters, and chaperones required for the cellular glycosylation machinery, glycan modifications and their degradation correspond to around 4% of the genome [10,11]. In fact, the glycosylation process may be dictated by another principle—namely, the paracentral dogma—acting in parallel to the existing central dogma [2]. For example, small RNAs are modified with *N*-glycans and displayed on the surface of living cells [12]. Although this concept and process may seem cumbersome and complex due to the multiple enzymes involved, it is a highly ordered process, and each enzyme is encoded by glycogenes. Hence, a disruption of glycogenes or a deficiency of any of the related enzymes can lead to a condition commonly termed as congenital disorders of glycosylation (CDG) [9].

Apart from CDG, evidence from the literature shows that glycans are highly dynamic, and their structures change in response to biological and environmental triggers as well as disease presence. For example, changes in glycan structure play a key role in maintaining the balance between pro- and anti-inflammation, which has been linked with the pathogenesis of biological aging, suboptimal health, fatigue, cancer, cognitive impairment, Alzheimer's disease, Parkinson's disease, diabetes, metabolic syndrome, dyslipidemia, hypertension, inflammatory bowel disease, ulcerative colitis, Crohn's disease, systemic lupus erythematosus, rheumatoid arthritis, sclerosis, stroke, acquired immune deficiency syndromes (AIDs), and coronavirus disease 2019 (COVID-19) [1,2,4,7,13].

Immunoglobulin G (IgG), the most common type of antibody, is an ideal model for studying protein glycosylation due to its clear functional domains and the highly conserved glycosylation sites at asparagine-297 of its heavy chains (Fig. 2). In regard to the abovementioned genetic alterations, the cellular environment is associated with aberrant glycosylation, which strongly influences inflammatory properties. For example, IgG glycome is malleable, as it is reliant on the expression levels of enzymes and the abundance of sugar nucleotide donors, which in turn are epigenetically regulated within the producing B/plasma cells (Fig. 1). Furthermore, the IgG N-glycome is considered to be a link between the genetic code of cells and the cellular environment. Therefore, in theory, it is possible to change one's IgG *N*-glycan composition by modifying one's lifestyle choices, such as by participating in certain activities (e.g., reduced/no smoking or alcohol, and regular physical activity) and eating a healthy diet. Aside from the presence of disease, altered plasma protein glycosylation has been linked to gender, age, smoking status, body mass index, plasma lipids, total cholesterol and triglyceride levels, blood pressure, fasting blood glucose, certain medications, and diet [1,4,7].

Several factors have been further explored in association with the IgG glycosylation profile that could drastically affect the affinity of IgG fragment cystallizable region γ (Fcγ) for the aforementioned Fcγ receptors (FcγRs) and complementary factors (Fig. 2). One of the most profound factors associated with IgG glycosylation—particularly in terms of increasing agalactosylation—is aging. The IgG glycome explains between 23.3%–58.0% of the variance in age [14,15]. "Glycan-age" concept studies have been able to explain age in different populations [14,15]. They have the potential to not only inform individuals of their "biological age" but also provide an incentive to improve overall health. Although the concept of aging can be the culmination of unfavorable levels of multiple factors, the translation of glycomics (i.e., the system-wide study of the relative abundance of glycan moieties) for use in predictive, preventive, and personalized or precision medicine is becoming a reality [1,2,13–16]. Sex and hormone levels are also
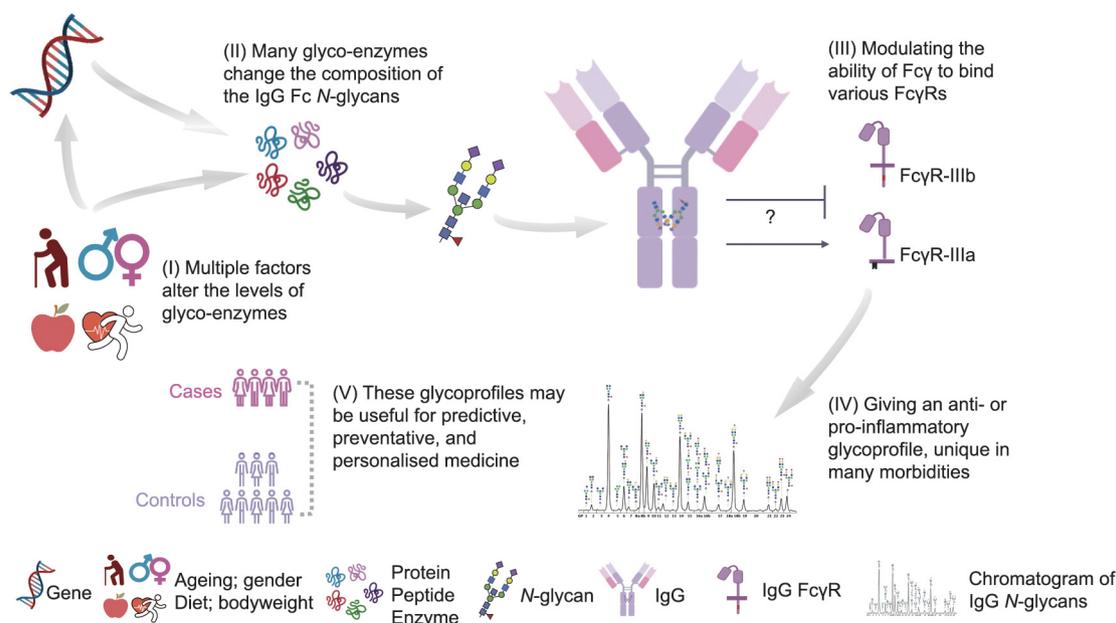
Fig. 2. Glycomedicine: glycans and their potential roles in precision medicine. Fcγ: fragment cystallisable region γ. Reproduced from Ref. [4] with permission of Springer, © 2021.

associated with notable changes to the IgG Fc glycome. In particular, these factors affect IgG Fc galactosylation and sialylation, with evidence of cyclical changes, such as during the menstrual cycle. IgG antigen-binding fragment (Fab) glycosylation is also associated with altered patterns of hormones during pregnancy, suggesting that estrogens may be responsible for modulating IgG Fc galactosylation in both women and men, with the estradiol aromatized from testosterone being responsible for these cyclic changes.

Aside from hormones, many other blood factors (e.g., extracellular glucose) are associated with *in vitro* changes to IgG galactosylation and sialylation, with the increased availability of Gal sugar nucleotide donors having been proposed as a mechanism. The association of fasting blood glucose with IgG glycosylation has also been seen *in vivo* in multiple populations. Other clinical traits found to be associated with IgG glycosylation, after correcting for age, include lipid profiles, blood pressure, insulin, alanine aminotransferase (ALT) and aspartate aminotransferase (AST), uric acid and urea, and fibrinogen, calcium, and HbA1c [4,7].

An increase in body fat has been found to correlate with an increased pro-inflammatory potential of IgG, and an increase in body mass index is associated with an increase in IgG agalactosylation. Furthermore, waist circumference, waist-to-hip and waist-to-height ratios, and dual-energy X-ray absorptiometry body fat parameters are associated with altered IgG glycosylation, with the latter explaining the most variation in the IgG glycome. The importance of these findings should be validated in longitudinal follow-up studies if it is found that reducing body fat via exercise, diet, or medication leads to positive changes in the IgG glycome [4].

Medications are associated with overall plasma glycosylation and IgG-specific glycosylation. Moreover, the effect of statin use has been associated with IgG glycosylation [4]. Although studies implicate medications as affecting the relative abundance of certain IgG *N*-glycan moieties, inconclusive results have been presented, suggesting that this effect might be so small that it does not have a significant effect on the IgG glycome.

Overall, these associations may directly influence the activity of the producing B cell or alter the expression of a number of glycogenes that encode glycosyltransferases and glycosyl hydrolases.

In addition to these identified factors associated with glycosylation, we expect that the plasma contains other biomarkers, which are yet to be explored in terms of their effect on the glycome. Thus, although there has been a considerable increase in our knowledge of the endogenous and exogenous factors associated with dynamic changes of glycosylation, further investigation is still warranted.

In 2012, a milestone in progress was reported, which paved the way for the glycan remodeling of intact therapeutic IgG antibodies to obtain new glycoforms with natural or selectively modified Fcγ glycans [16]. The novel glycosynthase mutants Endo-S-D233A and Endo-S-D233Q from Endo-S, a GH18 endoglycosidase from *Streptococcus pyogenes*, were identified, and aspartic acid-233 (Asp-233) was recognized as an essential residue that promotes the formation of the sugar oxazolinium ion intermediate during endoglycosidase-catalyzed hydrolysis. This Asp-233 residue site-specific mutation of Endo-S renders the enzyme incapable of catalyzing hydrolysis, but the mutant can still use the synthetic glycan oxazoline as a transition state mimic for the glycosylation of an acceptor. Both glycosynthase mutants, Endo-S-D233A and Endo-S-D233Q, can glycosylate the GlcNAc- or core-fucosylated GlcNAc moiety of a deglycosylated antibody. For example, the glycan remodeling of rituximab, an anticancer monoclonal antibody, leads to the efficient generation of a fully sialylated glycoform and a nonfucosylated glycoform of rituximab in high yield. These nonfucosylated G2 glycoforms showed a more than 20-fold enhanced affinity for the FcγR-IIIa in comparison with that of commercial antibodies. Furthermore, an azido tag could be selectively introduced at the Fcγ glycan, creating an opportunity for further chemoselective modification. Another example of glycoengineered cetuximab demonstrated increased affinity for the FcγR-IIIa and showed significantly enhanced antibody-dependent cellular cytotoxicity (ADCC) activity (Figs. 1 and 2). It is worth emphasizing that these glycoengineering techniques will provide a general platform for the production of diverse homogeneous antibody glycoforms, which are valuable tools for glycomedicine-focused studies and for the development of efficient antibody-based therapeutics. Such a platform is likely to make use of biosynthetic-pathway glycoengineering to produce low-fucose and/or nonfucosylated antibodies, to produce antibodies with increased galactose or sialic acid,

and to achieve the chemoenzymatic synthesis of homogeneous IgG Fc glycoforms. Recently, it has been reported that the engineered sialylation of pathogenic antibodies *in vivo* attenuates autoimmune disease [17].

In conclusion, it has been established that genetic and other factors influence glycosylation, which in turn can affect whether glycoproteins will elicit an anti-inflammatory or pro-inflammatory response. It is important to underscore these processes when considering the use of glycoprotein moieties as an indication of disease presence, progress, or response to therapeutics, as well as when considering the therapy itself. In addition, glycomedicine provides a clinical-translational platform for glycomic studies towards predictive, preventive, and personalized or precision medicine.

Although glycomedicine elicits considerable interest, it is not devoid of obstacles. Most glycomics analytical tools are unable to detect glycan concentration on a microscale level. Heterogeneity and the complexity of glycan structures make glycome analysis difficult, warranting the need for a new glycan analytical platform and automated glycobioinformatics resources. Moreover, only a few laboratories with advanced tools and expertise are able to analyze glycan structures, which poses another challenge for clinical application.

Despite these challenges, the scope of glycomedicine is broadening. The lessons learned from the unraveling of the sugar code, the innovative advances highlighted here, and the sociomateriality of the cell and biomolecules are instructive in charting a robust future for glycomedicine and its application potential in medicine.

Nomenclature

Glycan/carbohydrate/saccharide/sugar	Generic terms used interchangeably in this context; include monosaccharides, oligosaccharides, polysaccharides, and derivatives of these compounds
Glycoproteins	Proteins containing oligosaccharide chains (glycans) covalently attached to amino acid sidechains
Proteoglycans	A subclass of glycoproteins in which the carbohydrate units are polysaccharides that contain amino sugars; such polysaccharides are also known as glycosaminoglycans
Glycome	The entire glycan library of an organism/tissue/cell/protein, as systematically studied by glycomics
Glycomics	The systematic study of all glycan structures and sequences of a given cell type or organism
Glycosylation	The covalent attachment of sugar moieties (glycans) to proteins; a significant process in the endoplasmic reticulum and the Golgi apparatus
Glycation	Arginine and lysine residues in a protein attached to a glucose molecule (aldehyde form) via non-enzymatic reactions
Glycolipids	Lipids with a carbohydrate attached by a glycosidic (covalent) bond
Glycoside hydrolases	Hydrolases that catalyze the breakage of glycosidic bonds
Glycosyltransferases	Enzymes that establish natural glycosidic linkages
Glycocalyx	A fuzzy "sugar coating" often found on cell surfaces

References

[1] Wang Y, Adua E, Russell A, Roberts P, Ge S, Zeng Q, et al. Glycomics and its application potential in precision medicine. In: Precision medicine in China. Washington, DC: American Association for the Advancement of Science; 2016. p. 36–9.

[2] Özdemir V, Arga KY, Aziz RK, Bayram M, Conley SN, Dandara C, et al. Digging deeper into precision/personalized medicine: cracking the sugar code, the third alphabet of life, and sociomateriality of the cell. OMICS 2020;24(2):62–80.

[3] Hou H, Yang H, Liu P, Huang C, Wang M, Li Y, et al. Profile of immunoglobulin G N-glycome in COVID-19 patients: a case-control study. Front Immunol 2021;12:748566.

[4] Russell A, Wang W. The rapidly expanding nexus of immunoglobulin G N-glycomics, suboptimal health status, and precision medicine. Exp Suppl 2021;112:545–64.

[5] Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTBG, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 2013;32(10):1478–88.

[6] 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.

[7] Fournet M, Bonté F, Desmoulière A. Glycation damage: a possible hub for major pathophysiological disorders and aging. Aging Dis 2018;9(5):880–900.

[8] Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology 2012;22(6):736–56.

[9] Ondruskova N, Cechova A, Hansikova H, Honzik T, Jaeken J. Congenital disorders of glycosylation: still "hot" in 2020. BBA-Gen Subjects 2021;1865(1):129751.

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

[11] Hansen L, Husein DM, Gericke B, Hansen T, Pedersen O, Tambe MA, et al. A mutation map for human glycoside hydrolase genes. Glycobiology 2020;30(8):500–15.

[12] Flynn RA, Pedram K, Malaker SA, Batista PJ, Smith BAH, Johnson AG, et al. Small RNAs are modified with *N*-glycans and displayed on the surface of living cells. Cell 2021;184(12):3109–24.e22.

[13] Štambuk J, Nakić N, Vučković F, Pučić-Baković M, Razdorov G, Trbojević-Akmačić I, et al. Global variability of the human IgG glycome. Aging 2020;12(15):15222–59.

[14] Yu X, Wang Y, Kristic J, Dong J, Chu Xi, Ge S, et al. Profiling IgG *N*-glycans as potential biomarker of chronological and biological ages: a community-based study in a Han Chinese population. Medicine 2016;95(28):e4112.

[15] Vučković F, Krištić J, Gudelj I, Teruel M, Keser T, Pezer M, et al. Association of systemic lupus erythematosus with decreased immunosuppressive potential of the IgG glycome. Arthritis Rheumatol 2015;67(11):2978–89.

[16] 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(29):12308–18.

[17] Pagan JD, Kitaoka M, Anthony RM. Engineered sialylation of pathogenic antibodies *in vivo* attenuates autoimmune disease. Cell 2018;172(3):564–77.e13.