



*Teaser This review provides an integrated roadmap for evaluating the impact of glycosylation in the development and production of biopharmaceuticals, and the associated challenges in glycosylation analysis and process control.*



# Challenges of glycosylation analysis and control: an integrated approach to producing optimal and consistent therapeutic drugs

Peiqing Zhang<sup>1,2</sup>, Susanto Woen<sup>1</sup>, Tianhua Wang<sup>1</sup>, Brian Liao<sup>1</sup>, Sophie Zhao<sup>1</sup>, Chen Chen<sup>1</sup>, Yuansheng Yang<sup>1</sup>, Zhiwei Song<sup>1,2</sup>, Mark R. Wormald<sup>3</sup>, Chuanfei Yu<sup>4</sup> and Pauline M. Rudd<sup>1,5</sup>

<sup>1</sup> Bioprocessing Technology Institute, 20 Biopolis Way, #06-01 Centros, Singapore

<sup>2</sup> National University of Singapore, 21 Lower Kent Ridge Road, Singapore

<sup>3</sup> Oxford Glycobiology Institute, South Parks Road, Oxford OX1 3QU, UK

<sup>4</sup> National Institutes for Food and Drug Control, No. 2, Tiantan Xili, Dongcheng District, Beijing, China

<sup>5</sup> NIBRT GlycoScience Group, NIBRT – The National Institute for Bioprocessing Research and Training, Fosters Avenue, Mount Merrion, Blackrock, Co. Dublin, Ireland

**Glycosylation of therapeutic proteins has a profound impact on their safety and efficacy. Many factors shape the glycosylation of biotherapeutics, ranging from expression systems and cell culture processes to downstream purification strategies. Various analytical technologies have been developed to address questions concerning different aspects of glycosylation. Informatics tools are also crucial for a systematic understanding of the glycosylation processes. Hence, an integrated approach is required to harness glycosylation for the production of optimal and consistent glycoprotein-based therapeutic drugs. Here, we review the latest developments and challenges in glycosylation analysis and control in the context of bioprocessing monoclonal antibodies.**

## Introduction

Glycosylation is one of the most common post-translational modifications (PTMs) of proteins, present on more than 50% of the eukaryotic proteome [1]. Glycans impart a wide range of functions on their protein carriers, ranging from folding and quality control in the endoplasmic reticulum [2], intracellular and extracellular targeting [3], peptide loading onto MHC class I [4], mediation of virus binding to host cells [5] to immune modulation of Fc receptor interactions [6]. Glycosylation of biopharmaceutical drugs has a pivotal role in their safety and efficacy by modulating a wide range of drug properties, including immunogenicity [7], *in vivo* circulatory half-life [8] and effector functions [9]. A survey of the top 20 best-selling biopharmaceutical drugs up to 2013 shows that 11 of them are glycoproteins, out of which eight are monoclonal antibodies (mAbs) and antibody Fc-based fusion proteins [10]. Therefore, this review will focus on mAbs because they represent the major category of glycoprotein-based therapeutic drugs.

Corresponding author: Rudd, P.M. (pauline.rudd@nibr.ie)

**Pauline M. Rudd** Professor

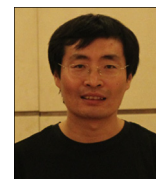
Pauline M. Rudd BSc, LRIC, MA (Oxon), PhD is the Adjunct Professor of Glycobiology at University College, Dublin. She heads the GlycoSciences Research Group at the National Institute for BioProcessing



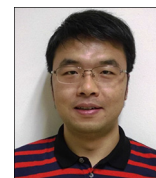
Research and Training in Ireland (NIBRT). She has many links with academic groups, regulators and pharmaceutical companies across the world because GlycoScience is a major area of specialized expertise required to probe the pathways in many diseases and to ensure the safety and efficacy of biotherapeutic drugs, such as monoclonal antibodies for cancer and autoimmune disorders. Professor Rudd obtained a BSc in chemistry at the University of London and a PhD in glycobiology at the Open University, UK. She was a Founding Scientist of Wessex Biochemicals (later Sigma London), Visiting Research Associate at The Scripps Research Institute, CA, Visiting Professor of Biochemistry at Shanghai Medical University PRC, Visiting Scientist at Ben Gurion University of the Negev, Israel, and Erskine Visiting Fellow, Canterbury University, Christchurch, New Zealand. She is a Fellow of the Royal Society of Medicine, London, and a Visiting Professor at St George's Hospital, London, and an Adjunct Professor at North Eastern University, Boston, NUI Galway and Trinity College Dublin. She is a Visiting Investigator at the Bioprocessing Technology Institute (BTI) in A-Star, Singapore.

**Chuanfei Yu** is Associate

Professor at National Institutes for Food and Drug Control (NIFDC), affiliated to China Food and Drug Administration (CFDA). He received his PhD in immunology from Peking University Health Science Center in 2008. From 2008 to 2010, he was a Postdoctoral Research Fellow at the Genetics Department of Harvard Medical School. He is now engaged in the quality control of recombinant antibody-based biotherapeutics. Now his current research interests focus on the impact of different glycosylation of monoclonal antibody products on their immunogenicity and bioactivity, among others.



**Zhang Peiqing** is a Research Scientist at the Bioprocessing Technology Institute, A\*STAR Singapore. He leads the Analytics Group focusing on biopharmaceutical glycosylation analysis and is the co-program lead of

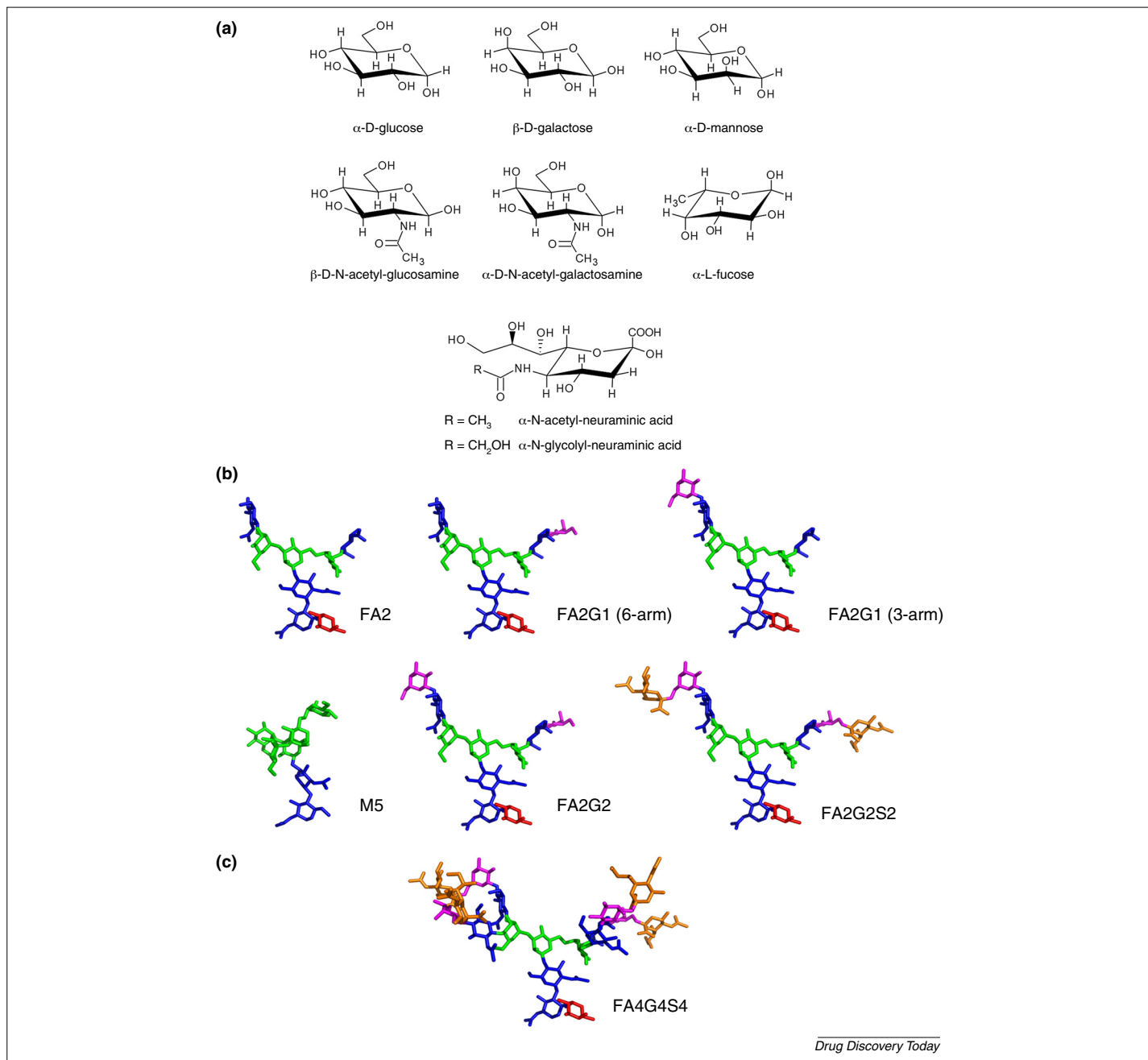


GlycoSing, a strategic program for building a center for glycobiotechnology and glycomics in Singapore. He also manages HighGlycoART, a Visiting Investigator Program for Professor Pauline Rudd funded by A\*STAR. Zhang Peiqing obtained his PhD in biochemistry from the National University of Singapore (NUS). He is currently an Adjunct Assistant Professor at NUS School of Medicine.

Since the approval of the first mAb product in 1986 (OKT3, which is a murine mAb against CD3 for kidney transplant rejection therapy), 49 mAb-based biotherapeutics have been approved and marketed in the USA and Europe [10]. As the growth engine of modern translational biotechnology, global sales of therapeutic mAbs totaled nearly US\$75 billion in 2013, accounting for approximately half of the revenues accrued from all biopharmaceutical products

[11,12]. Except for three biopharmaceuticals produced in *Escherichia coli*, all other products have been glycoproteins expressed in mammalian cells, highlighting the importance of understanding mammalian glycosylation and its regulation [13].

In contrast to nucleic acids and proteins, the biosynthesis of glycans is not directly template-driven but, rather, is a result of a complex network of metabolic and enzymatic reactions that are



**FIGURE 1**

Representative monosaccharide and oligosaccharide species in mammalian *N*-glycosylation. Glycans are oligomers of monosaccharide species that are connected together by various glycosidic bonds. **(a)** Structures of monosaccharides commonly found in mammalian *N*-glycans, such as hexoses (Hex) that include glucose (Glc), galactose (Gal), mannose (Man), as well as *N*-acetylhexosamines (HexNAc) that include *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc), and sialic acids such as *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc). Neu5Gc is not synthesized in humans and is considered an immunogenic glycan epitope that can be found on glycoproteins expressed by non-human systems. **(b)** Structure models of major *N*-glycan species found on recombinant IgG antibody drugs, including FA2 (also called G0F), FA2G1 (or G1F), FA2G2 (or G2F), FA2G2S2 (or A2F) and Man5. Note the positional isomers of FA2G1, in which the terminal Gal residue can be attached to the 6-arm or the 3-arm of the core structure. **(c)** Structure model of a tetra-antennary, fully sialylated *N*-glycan (FA4G4S4) that has been reported as a major oligosaccharide species on erythropoietin (EPO). Here, we only depict the conformation in which all four Neu5Ac residues are linked by  $\alpha$ 2,3-bond to the Gal residues. Other stereoisomers can arise from  $\alpha$ 2,6-linkage. This also shows the complexity in detailed glycan analysis.

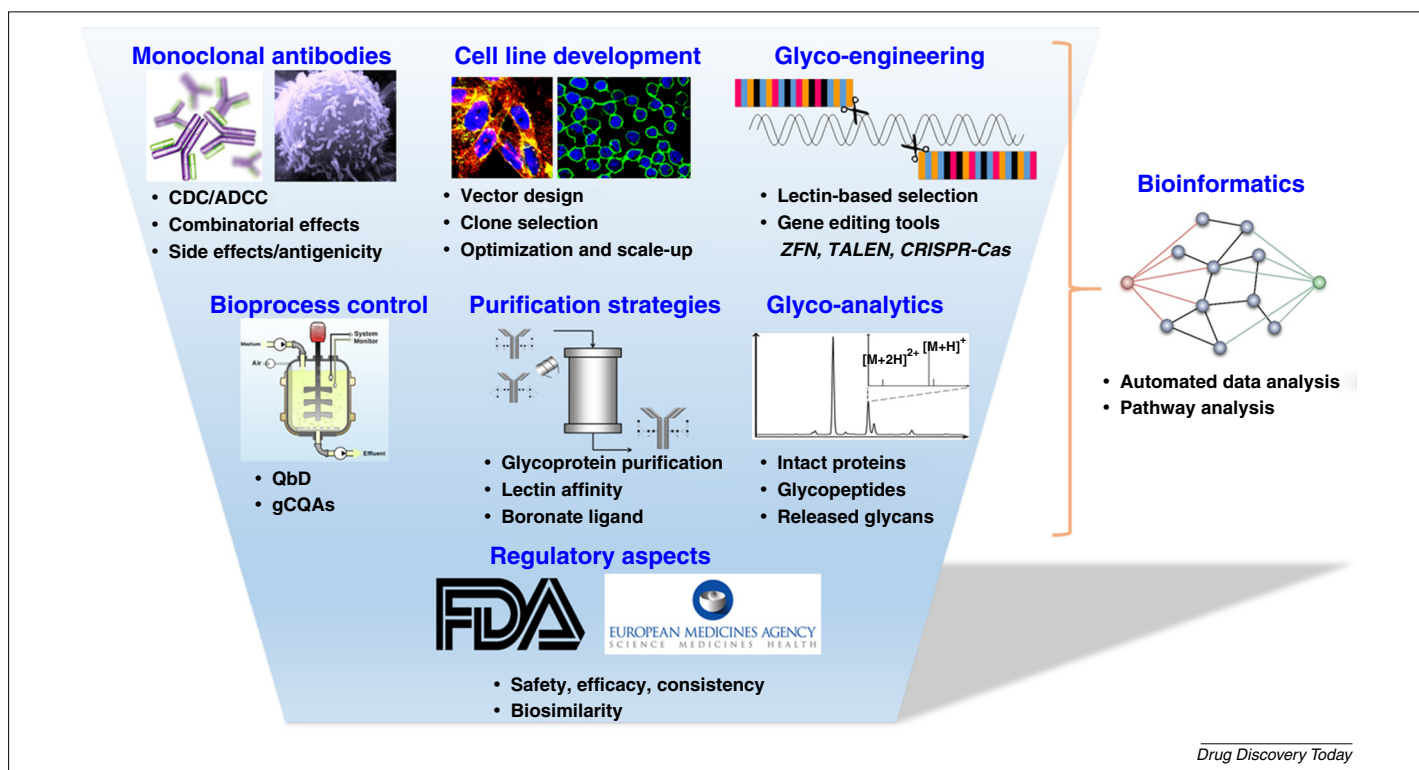
influenced by many factors, including the genetic profile of the cells in which the glycoconjugates are expressed [14], epigenetics [15] and the extracellular environment [16]. Consequently, glycoproteins, including biopharmaceutical products, always display a heterogeneous set of glycans that can be influenced by the host cell line as well as the upstream and downstream bioprocesses. Fig. 1 shows chemical structures of common monosaccharide species that form *N*-glycans, as well as structural models of oligosaccharides typically found on antibody drugs and erythropoietin (EPO). From a regulatory perspective, human-compatible and consistent glycosylation is required for a safe and efficacious drug product. This requires the drug developers to analyze the glycosylation systematically throughout the drug development and manufacturing processes. In addition, controlling glycosylation has been a long-standing challenge that requires a detailed understanding of the glycosylation pathways in cell culture processes.

This review aims to update our knowledge of the role of glycosylation on drug properties, with a focus on IgG antibodies and regulator requirements for glycosylation analysis. We will then walk through the process of producing mAbs, ranging from upstream and downstream bioprocesses to analytics and informatics tools. This will cover the cell line development, and tools for glyco-engineering host cell lines. Various upstream process parameters have been shown to influence glycosylation and we will provide a summary in the context of quality-by-design (QbD) as well as a review of alternative downstream practices on glycoform selection. State-of-the-art analytical technologies for

structural analysis of glycosylation will be discussed. Finally, a major bottleneck in glycosylation analysis and control lies with the lack of informatics tools available for efficient data processing and pathway modeling; therefore we will review the recent progress in these two areas. Fig. 2 provides an overview of the manuscript.

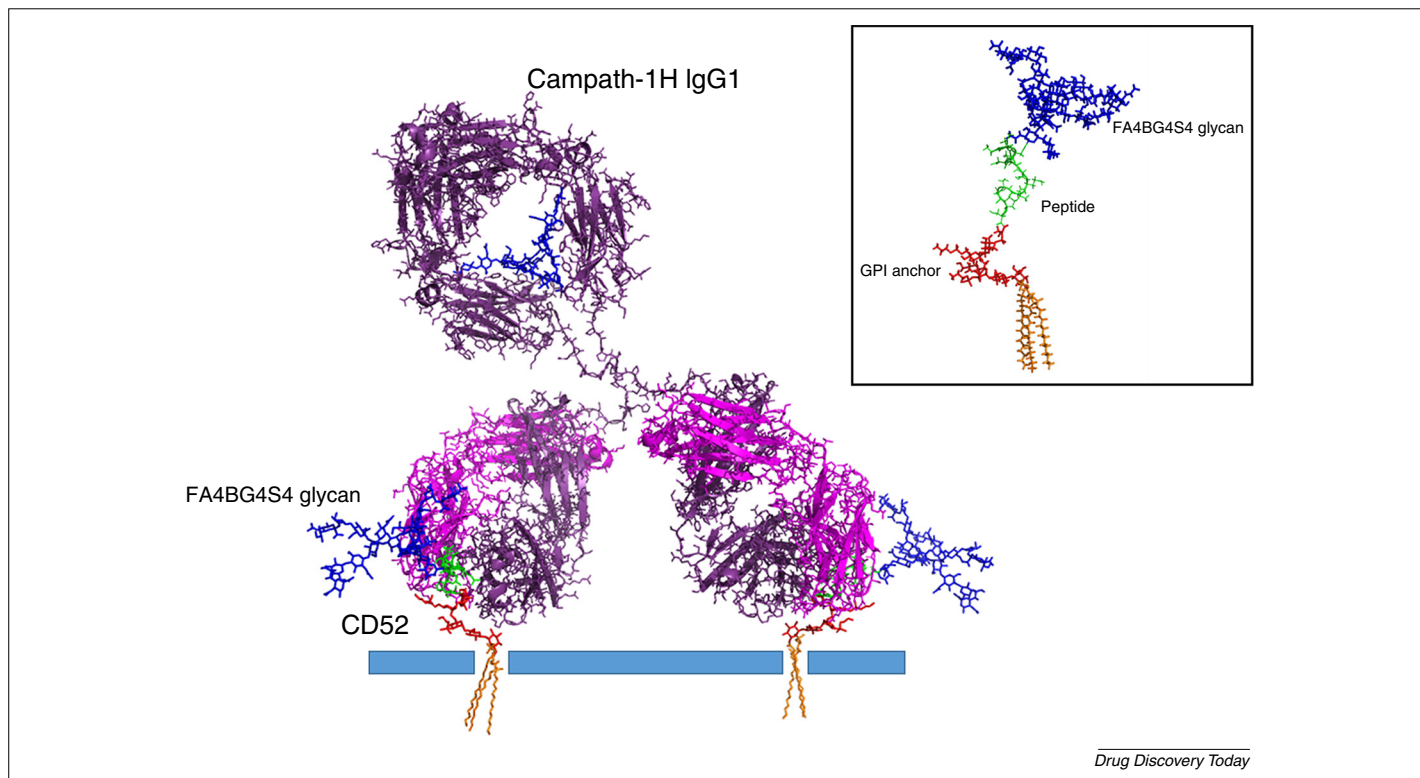
### Impact of glycosylation on antibody characteristics

IgG is the most abundant among the five classes of serum immunoglobulins. It makes up about 15–20% of serum glycoproteins. In humans, there are four subclasses of IgG (IgG1–4), which have differences in structure, antigen-binding characteristics as well as in a range of interactions with cellular receptors through the Fc region [17]. To date, most approved therapeutic antibodies are of the IgG1 subclass which effectively initiates antibody-dependent cell-mediated cytotoxicity (ADCC). We can learn from studies of natural human antibodies and apply the knowledge to the development of recombinant mAbs. The Fc region of IgG1 contains a conserved *N*-glycosylation site at Asn297 in each CH<sub>2</sub> domain. Fig. 3 illustrates the binding of alemtuzumab (Campath<sup>®</sup>), a humanized IgG1κ mAb used to treat B cell chronic lymphocytic leukemia (B-CLL), to its target CD52 expressed on the B cell surface. About 20% of the circulating IgG molecules carry additional *N*-glycans located in their Fab region [18–21]. With the single exception of cetuximab, which has an *N*-glycosylation site in the variable region of each heavy chain (Fig. 4), all approved therapeutic mAbs are only *N*-glycosylated in the Fc region



**FIGURE 2**

Developments and challenges of glycosylation analysis and control in bioprocessing. Bioprocessing for recombinant drug production involves a complex pipeline spanning the structure–function study of the drug molecule, cell line development, upstream and downstream processes, analytics and informatics, until regulatory approval and marketing. Each of the sectors presents significant challenges and opportunities for glycosylation analysis and control. This roadmap provides a summary of this review.

**FIGURE 3**

Structure model depicting the binding of Campath-1H to CD52. The model is based on 1ce1 (PDB ID), which shows the complex of Campath-1H Fab and CD52 [259]. Color scheme is as follows: Campath-1H heavy chain (dark purple), Campath-1H light chain (light purple), glycans on Campath-1H and CD52 (blue), CD52 peptide (green), glycosylphosphatidylinositol (GPI) anchor (red), CD52 lipid carrier (orange), B cell membrane (light blue box). The glycan on CD52 is represented by FA4BG4S4, which is a core fucosylated, tetra-antennary, fully sialylated *N*-glycan with a bisecting GlcNAc. Inset shows the topology of CD52. Note the relatively small portion of the peptide in relation to the carbohydrate moiety, the GPI anchor and the lipid tail.

[7,22,23]. Fig. 5 shows typical glycan profiles of human serum IgG and trastuzumab, which is a recombinant humanized IgG1 produced in Chinese hamster ovary (CHO) cells used as a targeted therapy for HER2-positive metastatic breast cancer.

#### Effect of Fc region glycosylation on therapeutic antibody characteristics

In contrast to the Fab region which recognizes antigenic epitopes, the Fc region has limited structural diversity. The Fc region determines the function that will ensue after antigen binding. It can recruit molecules in the innate immune system, such as C1q, as well as cytotoxic and antigen-presenting cells via binding interactions with Fc $\gamma$  receptors. IgGFc contains two conserved *N*-glycosylation sites at Asn297, one on each heavy chain [24]. Variations in the structure of these glycans result in subtle changes in structure that have a significant influence on the interaction of IgG with the immune system (Table 1).

One of the crucial functions of Fc glycans, and indeed of serum glycoprotein *N*-glycans in general [25], is to regulate protein turnover and renewal. Circulating exoglycosidases progressively remodel glycoprotein glycans, rendering them vulnerable to clearance by the immune system [25]. For example, it has been shown that the G0 glycoform of IgG, which contains no galactose and therefore presents terminal *N*-acetylglucosamine (GlcNAc), is bound by a C-type lectin, the mannose receptor, and is cleared by dendritic cells and macrophages [26,27]. Similarly, IgGs bearing high mannose-type glycans

exhibit reduced serum half-life [28]. These findings have strong implications for therapeutic antibody quality control, suggesting that such glycoforms should be carefully controlled at a low level. IgG, modified by additional sugars in the CH3 domains, enables more-efficient binding to FcRn, the neonatal receptor, extending the half life [29]. Fc glycans also play a part in IgG transport, as shown by the increase in galactosylation during pregnancy [30].

Fc region glycans can directly influence the affinity of IgGs to Fc $\gamma$  receptors, either by changing the conformation of the Fc region [31,32] or through glycan–glycan interactions [33], thus strongly influencing their ability to recruit immune effector cells. Fc region glycosylation helps to maintain the stability of the molecule, and it has been shown that IgGs that have enzymatically truncated or removed Fc glycans show reduced thermal stability, complement binding [34,35] and impaired cell-dependent cytotoxic effects [36]. Furthermore, it has also been shown that the presence of glycans containing core fucose [37] or terminal sialic acid [38] at Asn297 reduces the affinity of IgGs to Fc $\gamma$  receptors. In the case of core fucose, it has been shown that either knockdown of the gene encoding  $\alpha$ 1,6-fucosyltransferase (*Fut8*) or overexpression of the gene encoding GnT-III transferase (*Mgat3*) results in a lack of core fucosylation at Asn297, vastly increasing the affinity of IgG for Fc $\gamma$ RIIIa more than 100 times [37] and thus increasing ADCC. Subsequent analysis of crystal structures of the Fc receptor in complex with either the fucosylated or afucosylated Fc regions showed that the lack of fucose allows carbohydrate–carbohydrate



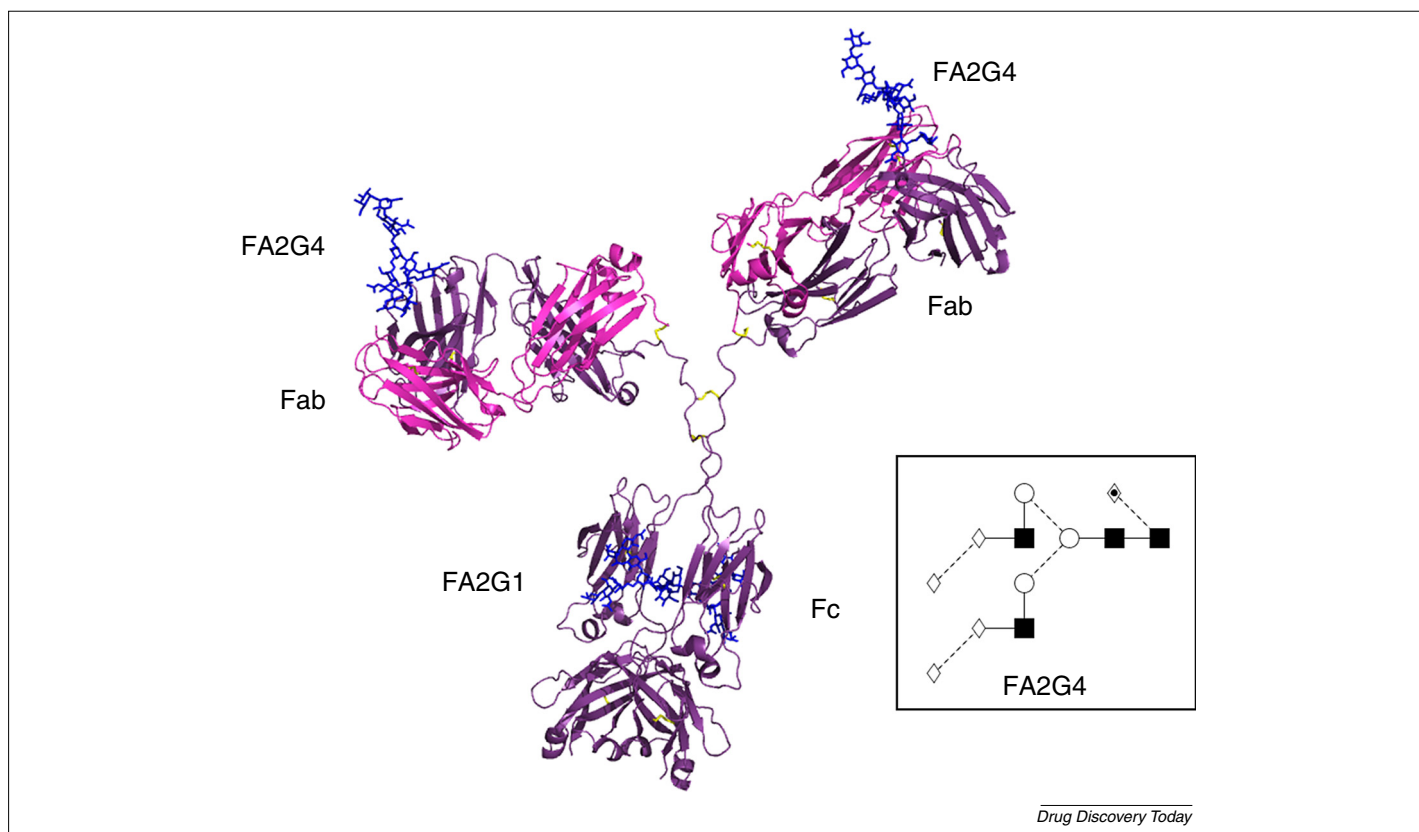


FIGURE 4

Structure model of cetuximab and the glycans attached to its Fab and Fc regions. Cetuximab is glycosylated in the Fab and Fc regions. This figure shows a model of cetuximab with an FA2G4 glycan attached to each of the heavy chain variable regions ( $V_H$ ), and an FA2G1 glycan attached to each of the Fc regions. The model is built upon 1YY8 (PDB ID), which modeled the Fab region of cetuximab. Inset shows the structure model of FA2G4, which terminates with an  $\alpha$ Gal immunogenic epitope on each arm. Color scheme is as follows: cetuximab heavy chain (dark purple), cetuximab light chain (light purple), glycans (blue).

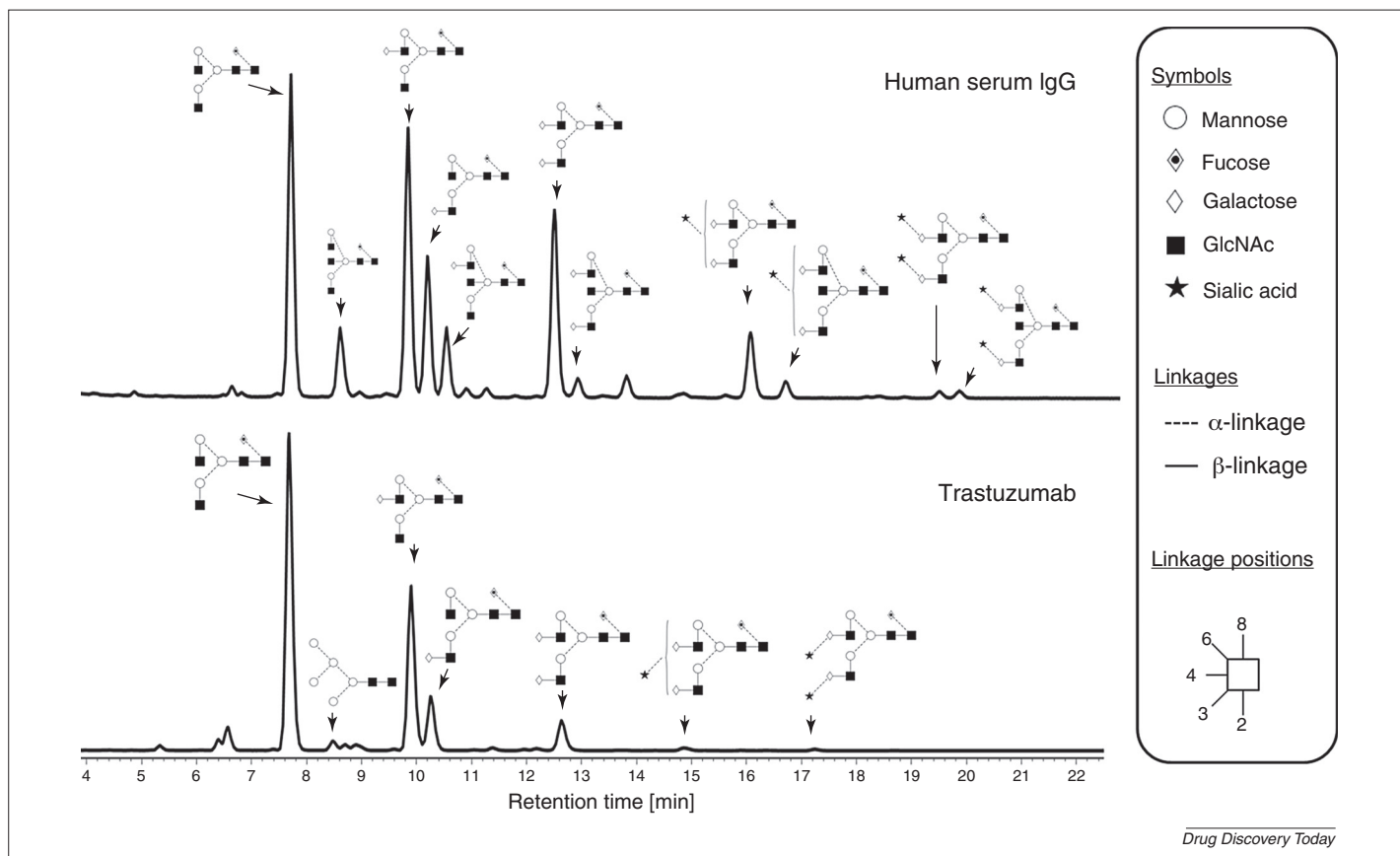
interactions to form between the glycans of the Fc receptor and those at Asn297, enhancing their binding affinity [33].

Fc glycans are infrequently sialylated owing to their relative inaccessibility and their interactions with the protein surface [39], and sophisticated methods are required to achieve a high degree of sialylation *in vitro* [40]. Several studies suggest that the presence of sialylated glycans on the Fc region reduces the ability of IgGs to recruit effector cells via Fc $\gamma$  receptor binding [38,41], and that Fc glycan sialylation is highly correlated with the presence of C-terminal lysine [42]. Fc glycan sialylation has also been strongly associated with anti-inflammatory properties. These studies were initially carried out on intravenous immunoglobulin (IVIG), which is a therapeutic product consisting of IgGs extracted from the pooled plasma of 30 000 - 60 000 healthy donors per batch [43]. IVIG has profound yet poorly understood anti-inflammatory effects, and as such has been used for over 30 years to treat a wide range of acute and chronic autoimmune and inflammatory diseases. It was shown that sialic-acid-enriched fractions of IVIG had up to tenfold higher anti-inflammatory activity compared with unfractionated IVIG [43]. Moreover, subsequent reports showed that the Fab regions were dispensable for this effect, and that Fc regions were solely responsible [44]. Unsialylated IVIG was found to interact primarily with type I Fc receptors, which includes all Fc $\gamma$  receptors, whereas sialylated IVIG interacted primarily with type II Fc $\gamma$  receptors, which includes lectin receptors such as dendritic-cell-specific intercellular adhesion molecule-3-grabbing

non-integrin (DC-SIGN) [38]. This suggests that differential engagement of Fc receptors on effector cells was responsible for the observed differences in activity. These findings have not been reproducible under all experimental conditions – a number of researchers maintain that the anti-inflammatory properties of IVIG are not dependent on sialylation [45–47]. Moreover, the precise anti-inflammatory mechanism is still very much in doubt. It has been proposed that IVIG potentiates the activation of basophils [48] or regulatory T cells [49,50], but several studies have shown that this is not necessarily the case [45,51].

#### Effect of Fab region glycosylation on antigen binding

Although the majority of IgG antibodies are modulated by glycosylation in the Fc region, some 20% also contain nonconserved N-glycosylation sites (Asn-Xaa-Ser/Thr) in the variable region. Indeed, analyses of human and animal antibodies have shown that variable region (mainly  $V_H$ -region) carbohydrates are frequently present [18–21]. A survey of human cDNA sequences indicated that 9% of variable regions had a potential N-linked glycosylation site [52], and other studies indicate that about 20% of polyclonal IgGs from human serum have V-region glycosylation [53,54]. V-region glycosylation can modulate the affinity of the antibody for its antigen, potentially either enhancing binding strength or abolishing it completely. The net effect of V-region glycosylation is exquisitely sensitive to the location of the carbohydrate chain on the complementarity-determining



**FIGURE 5**

Glycan profiles of human serum IgG and trastuzumab (Herceptin<sup>®</sup>) analyzed by HILIC-UPLC. Glycans on human serum IgG and trastuzumab were released by PNGase F treatment and labeled with 2-AB. The labeled glycans were analyzed by HILIC-UPLC with fluorescence detection. In both IgG samples, the majority of the glycans are complex-type, neutral glycans, such as FA2 (G0F), two isomers of FA2G1 (G1F) and FA2G2 (G2F). Other truncated complex-type glycans and afucosylated structures are also found, but at lower abundance. Oligomannose structures, most notably Man5, are typically found in recombinant IgG samples. Their abundance in human serum IgG is much lower as a result of mannose-receptor-mediated clearance. Sialylated glycans exist in much higher abundance in human serum IgG compared with trastuzumab, because of clones bearing Fab glycans. Other significant differences in the two IgG samples are bisecting GlcNAc and Neu5Ac connected in  $\alpha$ 2,6-linkage, which are present in human serum IgG but not CHO-derived recombinant IgG, including trastuzumab.

region (CDR) loops which determine recognition and binding. For example, a study of murine antibodies of V<sub>H</sub> subgroup IIIb indicated that glycosylation at Asn58 in the second CDR of V<sub>H</sub> enhanced the affinity of an antidextran antibody for its antigen by approximately tenfold [21,55]. However, when nonglycosylated, antidextran antibodies were engineered to express N-glycosylation sites, it was found that glycosylation at Asn60 increased overall affinity only threefold, whereas glycosylation at Asn54 inhibited binding to dextran. Thus, a shift in glycosylation site by only a few amino acids can turn an enhancing glycan into an inhibitory one. To avoid any complications, during the screening of mAbs, clones bearing Fab glycans are typically discarded or the glycosylation sites mutated. In certain cases, however, Fab glycosylation can be exploited to improve drug properties. V-region glycosylation has been proposed as a way to increase the neutralizing ability of antibodies. A recent study of an anti-factor-VIII neutralizing antibody showed that V-region glycosylation enhances the ability of the antibody to neutralize factor VIII, but does not increase its binding affinity [56]. Song *et al.* engineered an N-glycan into the light chain of ibalizumab, a neutralizing antibody against HIV-1 that binds to the surface envelope glycoprotein gp120 [57]. The Fab N-glycan refilled a

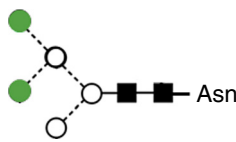
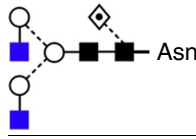
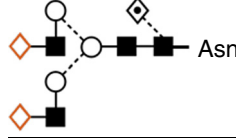
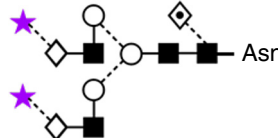
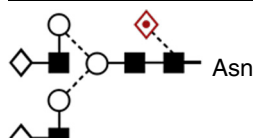
cavity caused by a mutation in gp120. *In vitro* data show the glyco-engineered ibalizumab binds to a broad range of HIV-1 strains, including the ibalizumab-resistant mutants.

Fab glycosylation has also been proposed as a mechanism to block self-binding sites during immune system development. Autoantibodies on anergic B cells contained V-region glycans that blocked antigen-binding sites capable of recognizing foreign- and self-antigens. Reactivation of anergic B cells by foreign antigen challenge was characterized by an increase in the germinal cell progeny of IgM<sup>low</sup> IgD<sup>+</sup> anergic B cells that initially had 72% of their antigen-binding sites blocked by V-region glycans. Following somatic hypermutation, these binding sites were cleared of blocking glycans, increasing their binding affinities 100-fold [58].

In contrast to the conserved glycans in the Fc, most V-region glycosylation sites are located on solvent-exposed loops and are readily accessible to endogenous lectins. They tend to contain fully processed, sialylated glycans, with a high incidence of bisecting GlcNAc residues [59,60]. Intriguingly, in some cases oligomannose glycans have been found within the CDR antigen-binding regions [55], particularly in the case of follicular lymphoma patients [61]. These glycans are accessible to C-type lectin-binding domains in carbohydrate recognition domains such as those

TABLE 1

Examples of Fc glycan on antibody properties and functions.

Fc glycan feature	Impact on antibody properties and functions	Refs
	High mannose glycan increases serum clearance of IgG	[28]
	IgG G0 interacts with mannose binding protein to (i) activate complement and (ii) facilitate serum clearance	[26,27]
	IgG G2 glycoform is increased in pregnant women and umbilical cords	[30]
	Desialylation of intravenous immunoglobulin (IVIg) abrogates anti-inflammatory properties in KN mice	[42]
	Loss of core alpha(1,6) fucose on IgG results in enhanced antibody-dependent cell-mediated cytotoxicity (ADCC) activity	[33,37]

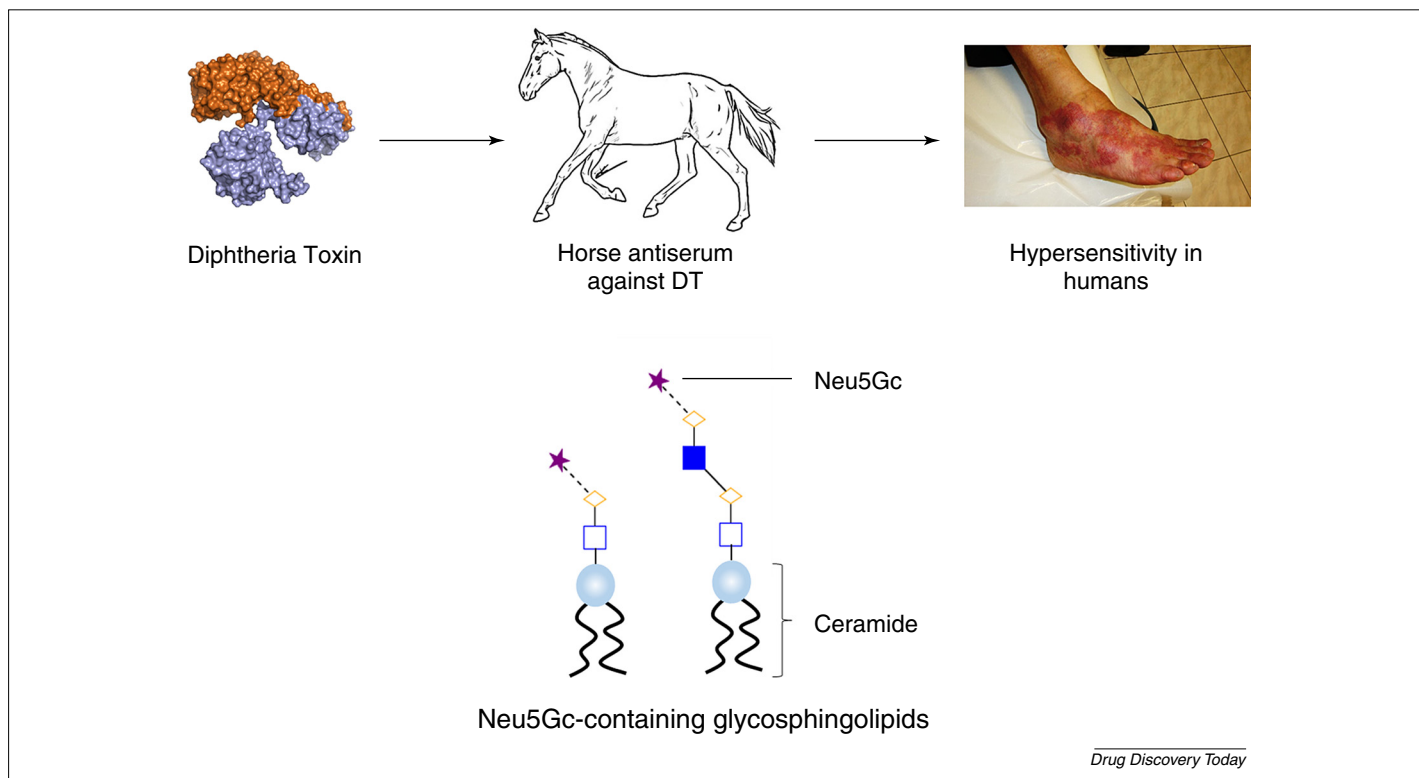


FIGURE 6

The role of non-human sialic acid, Neu5Gc, in serum sickness. Injection of horse antiserum against diphtheria toxin led to hypersensitivity reactions in humans. Antibodies were directed against H-D antigen, and were capable of agglutinating erythrocytes from a variety of mammals, including cows. Later it was found that antibodies against H-D antigen react only with Neu5Gc-containing glycosphingolipids.

presented by the mannose-binding lectin. The mannose glycans were not the result of impaired glycosylation pathways because Fc region glycosylation remained complex. V-region oligomannose glycans do not always inhibit antigen binding [55]; instead, recent evidence suggests that they contribute to pathological processes by binding to lectin domains of the innate immune system such as the mannose receptor and DC-SIGN, blocking their signaling functions [62].

#### Immune reactions to antibody glycans

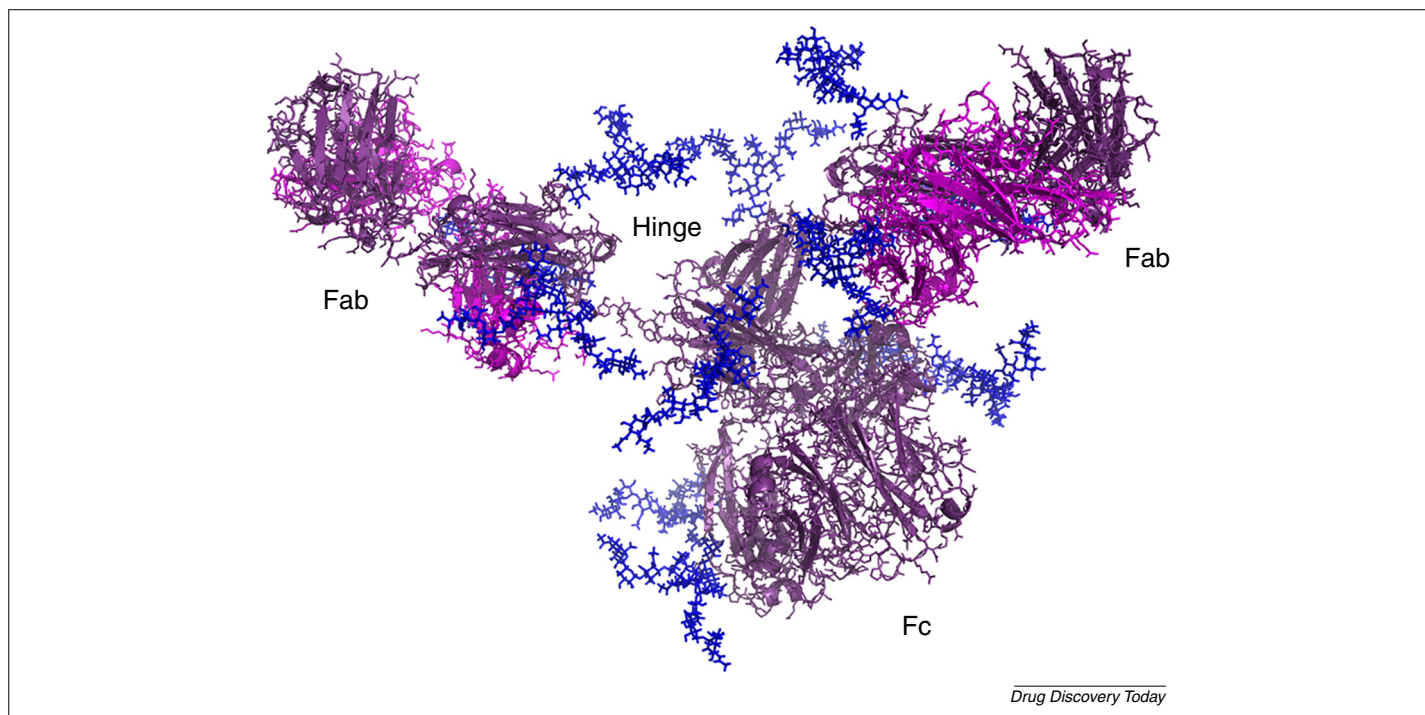
Humans have baseline levels of antibodies against certain non-human glycan motifs including *N*-glycolylneuraminic acid (Neu5Gc) [63] (Fig. 1) and terminal  $\alpha$ -1,3-linked galactose ( $\alpha$ -gal) (Fig. 4) [7]. Their potential inclusion as PTMs on biologics thus presents an immunogenic risk to patients. Excluding these glycan motifs from recombinant glycoproteins is not a matter that can be simply solved by switching production cell lines – human cells cultured *in vitro* retain the ability to incorporate Neu5Gc through the salvage pathway when supplied with exogenous precursors to Neu5Gc in the medium [63,64] and CHO cells, which are the workhorses of the biopharmaceutical industry, can incorporate Neu5Gc [65] and  $\alpha$ -gal [66] into recombinant glycoproteins, thus emphasizing the need for the optimization of culture conditions, as well as stringent quality control measures.

The production of heterophilic antibodies upon injection with foreign serum was first described by Hanganutziu [67] and Deicher [68] in the 1920s as part of a severe immune response termed ‘serum sickness’ (Fig. 6). These antibodies caused higher levels of serum hemagglutination of horse, calf, sheep and rabbit erythrocytes, as well as complete adsorption to guinea-pig kidney

sediment. However, their binding targets remained undetermined until the 1970s. Then, some early experiments with biologic drugs were performed when anti-human thymocyte immunoglobulins were raised in goats and injected into patients with venous immunologically mediated or lymphoproliferative diseases [69]. Although none of the patients experienced severe side effects, it was found that 91% of them developed the ‘H-D antibodies’ originally described by Haganutziu and Deicher. Researchers subsequently discovered that the immunological agents giving rise to H-D antibodies were Neu5Gc-containing glycosphingolipids [70], thus demonstrating that glycans on biologic therapeutics can provoke systemic immune responses (Fig. 6).

Nevertheless, hypersensitivity reactions toward Neu5Gc in biologics are rare. IVIG Fc fractionated with *Sambucus nigra* agglutinin (SNA), a sialic-acid-binding lectin, was the primary anti-inflammatory component of IVIG [71], and only Fc containing two or more sialic acids was capable of binding SNA owing to the spatial constraints of the CH2 domains and sialic acid exposure on the surface of the antibody [72,73]. Using ELISA and BIAcore analysis, combined with thorough characterization of glycan profiles at the intact antibody level with mass spectrometry (MS), our data [41] demonstrate that the binding reactivity of the mAbs to the anti-Neu5Gc antibody resided in the SNA-bound portion containing two or more Neu5Gc, whereas the SNA-non-bound portion harboring only one Neu5Gc showed no reactivity. The fact that most Neu5Gc epitopes are distributed singly on the Fc of mAbs suggests the low potential immunogenicity of Fc Neu5Gc.

Despite increasingly sophisticated methods for biologics production, hypersensitivity reactions still occur in 1–3% of treated patients [7,74]. Adverse immune responses to poorly controlled



**FIGURE 7**

Structure model of IgE with the attached glycans. IgE has three conserved *N*-glycosylation sites at asparagine residues at positions 140, 168 and 218 located in the Fab region of the heavy chain as well as positions 265, 371, 383 and 394 located in the Fc region. Color scheme is as follows: IgE heavy chain (dark purple), IgE light chain (light purple), glycans (blue).



biologic glycosylation can be as minor as local irritation or as serious as cardiovascular failure [75]. Recent evidence indicates that hypersensitivity to glycan immunogens is strongly mediated by IgE antibodies [7,74,76] (Fig. 7). In a landmark study, Chung *et al.* found that a high prevalence of hypersensitivity reactions to cetuximab (Fig. 4), reported in some parts of the USA, were due to IgE antibodies against  $\alpha$ -gal [7], with symptoms that ranged from transient flushing or rash to full-blown anaphylaxis. Not all humans have sufficiently high levels of anti- $\alpha$ -gal IgE to elicit hypersensitivity responses – researchers observed that the incidence of either cetuximab or diet-associated  $\alpha$ -gal hypersensitivity were higher in South-Eastern USA relative to the northern states [74,76]. It has been hypothesized that environmental stimuli such as tick bites, which are prevalent in the southern states but rarer in the north, could be responsible for regional differences in circulating anti- $\alpha$ -gal IgE levels, accounting for differences in sensitivity to cetuximab [77]. It was found that tick bites increase the level of anti- $\alpha$ -gal IgE antibodies 20-fold [77], providing a neat explanation for the observed phenomenon. Interestingly, it has been found that the glycans in cetuximab responsible for provoking hypersensitivity responses lie in the Fab region as opposed to the Fc region because they are more accessible to IgE binding [22]. Other approved biologics such as infliximab and palivizumab, which have been produced in murine SP2/0 and NS0 lines, also contain appreciable amounts of  $\alpha$ -gal, but do not provoke similar rates of hypersensitivity reactions because they are not glycosylated in the Fab region.

### Regulatory perspectives on glycosylation of therapeutic antibodies

Appropriate glycosylation is one of the critical quality attributes (CQAs) that must be demonstrated to ensure the safety and potency of commercial mAbs before regulatory approval. WHO guidelines on biotherapeutics [78] and the International Conference on Harmonization (ICH) Q6B [79] mandate state that PTMs such as glycosylation should be identified and adequately characterized. Thorough characterization of mAb glycan profiles and their potential impacts on safety and therapeutic efficacy are therefore an essential part of quality control strategy. Glycan profile characterization consists of two main steps: first, complete characterization of reference and/or conformance lots of a mAb; and, second, abbreviated tests of subsequent batches for lot release [80]. Because antibody glycosylation can be influenced even by subtle changes in the manufacturing process, the abbreviated tests are often necessary to demonstrate manufacturing consistency. However, the abbreviated tests might not be included in the lot release assay if the sponsors manage to prove that tight process control strategies can ensure the consistency of the manufacturing process by adopting QbD approaches. WHO guidelines on biotherapeutics suggest that the selection of tests to be included in routine control programs should be product-specific and should take into account the necessary quality attributes; so, for mAb products in which glycosylation has been identified as a CQA, sponsors are encouraged to include the abbreviated test in the lot release assays. Because glycosylation is so sensitive to environmental perturbations, therapeutic products that are subject to manufacturing process changes must undergo comparability studies as suggested by ICH Q5E. As part of these tests, glycan profile analysis should be thoroughly exercised to ensure the

comparability of mAb products between, before and after process changes have been implemented.

Biosimilars are drawing increasing attention because of the impending 'patent cliff'. Regulatory approval for a biosimilar product is based on its comparability to an originator, whereas a thorough quality comparability exercise is crucial for demonstrating biosimilarity [81]. WHO guidelines [78] recommend the use of the same host cell type for glycoproteins in most cases, because glycosylation patterns vary significantly between different host cell types [82]. European Medicines Agency (EMA) guidelines [79] require a detailed comparison of carbohydrate structures including the overall glycan profile, site-specific glycosylation patterns and site occupancy. The EMA also points out that the presence of glycosylation structures or variants not observed in the reference medicinal product can raise concerns and would require appropriate justification, with particular attention to non-human structures (non-human linkages, sequences or sugars). FDA guidelines [83] also require the detailed comparison of glycan profiles for biosimilars.

### Cell line development

Mammalian cell lines used for manufacturing recombinant therapeutic proteins are derived from single cell clones producing high amounts of product with consistent quality to meet regulatory requirements. The process of cell line development for recombinant protein production generally starts with transfecting the mammalian host cells with plasmid vectors carrying the gene of interest and a selection marker gene. Antibiotic-based selection is next performed to isolate a pool of stably transfected cells with the plasmid integrated into their genome by killing the untransfected cells. Owing to random integration and host cell heterogeneity, the stably transfected pool contains cells that exhibit variations in productivity, stability and other product characteristics [84,85]. Cells with high productivity, long-term stable production and good product quality, all of which are required for a production cell line, are rare within a stably transfected pool. To identify cell lines suitable for the final production, hundreds to thousands of clones are screened through multiple stages of evaluation. Typically, a primary screen is first carried out on a small scale, for example 96- and 24-well plates are used for productivity assessment [86]. High-producing clones are identified and scaled up to larger volumes, using shake flasks and micro bioreactors [87,88]. Although micro reactors can better mimic the subsequent, larger-scale bioprocesses, many of the available systems are costly and can increase the development cost significantly. Using a larger culture volume at this stage provides sufficient material for more-detailed characterization of cell growth, metabolism, volumetric productivity (titer) and some product quality attributes, such as glycosylation, in either batch or simple fed-batch cultures. The short-list of the top-performing clones is finally characterized again for productivity and comprehensive product quality attributes in laboratory-scale bioreactors with well-controlled culture environments to mimic the final large-scale manufacturing. In parallel, these top-performing clones are evaluated for long-term production stability in shake-flask cultures. Crucial product quality attributes include glycosylation profiles, charge variants, aggregate levels and protein sequence variants [89,90]. Assays used for early-stage cell line development, such as ELISA, are chosen for

their high throughput and requirement of small amounts of samples, whereas those used at a late stage are more complex and are able to provide more-detailed characterizations, such as HPLC and MS. By considering productivity, production stability and quality requirement, a final production clone and a backup clone are chosen for future large-scale manufacturing.

#### *Impacts of expression vectors on glycosylation*

During the past decade, R&D has focused on improving the efficiency of generating cell lines with high productivity [91]. With optimized expression vectors, high-throughput clonal selection methods and improved media formulation, it is now possible to generate cell lines with titers of more than 2 g/l on a routine basis. The timeline required to generate a high-producing cell line has also been reduced from over a year to a few months [92]. There is now a shift from increasing productivity toward improving product quality motivated by the improved analytical tools and the rise of biosimilars for existing blockbusters with expiring patents. Besides obtaining high titers to reduce cost-of-goods, regulatory approval of biosimilar products requires demonstration of comparable product quality as the innovator drugs [93,94]. Within the list of product characteristics, glycosylation stands out as one of the most crucial quality attributes and the most difficult parameter to control owing to its complex structures and sensitivity to the manufacturing process. Many steps involved during the cell line development process are crucial to the final product glycosylation profile. Known factors that contribute include the host cells, the protein itself and various cell culture environment variables [16,95]. Glycosylation differs between different mammalian host cells and even within a population of the same cells [96]. The choice of host cells is crucial to determining the glycosylation profile of a product. Few studies have looked at the impact of expression vectors on glycosylation. Different vector designs have a direct impact on the relative amounts of light chain and heavy chain for mAbs, affecting assembly and the glycosylation of the product [97,98].

Promoters and other DNA regulatory elements for enhancing recombinant protein expression could also have an impact on glycosylation because reports have shown direct and inverse correlations between protein synthesis rates and glycosylation for different proteins [95]. Antibiotic resistance genes and metabolic selectable genes, such as dihydrofolate reductase (DHFR) and glutamine synthetase (GS) have been commonly used for the generation of stably transfected cell lines [99]. Different selection markers seem to have little effect on obtaining product with the desired glycosylation despite their different mechanisms of action (Yang *et al.*, unpublished). Many studies have looked at the impact of culture conditions and media components on glycosylation. One conclusion drawn from these studies is that the impact of each culture variable is cell line and protein specific [95].

#### *Controlling glycosylation by targeted integration*

As a consequence of our incomplete understanding of cell culture parameters that affect glycosylation, the generation of cell lines with desired glycosylation still relies on empirical clone selection, product characterization and process development. Glycosylation varies from clone to clone in a stably transfected pool [84]. Identifying a cell line with a desired glycan profile is challenging using the current cell line development process. Detailed product

characterization for a large number of cell lines at early stages of cell line development is limited by the low throughput of the currently available analytical tools. Moreover, the glycosylation behavior of a cell line at an early stage might not be reflective of that at a late stage because of changes in the cell culture media, process and vessel types. Identifying cell lines with a desired glycoform profile at an early stage also requires the development of low-cost micro bioreactors to represent better the final production process. Another attractive concept is to reuse cell lines with well-characterized glycosylation pathways and product characteristics. The idea involves generating a panel of master cell lines using plasmid vectors containing homologous recognition sequences. Besides having high productivity and stable production, these master cell lines are chosen for their different glycosylation potentials and would thus be able to produce products with different glycan profiles. A specific master cell line able to produce the required glycosylation could be selected for expressing a specific protein through targeted integration and cassette exchange by using either recombinase- or dsDNA-break-based technologies [92].

#### **Glyco-engineering host cell lines**

Apart from using vector design and clonal selection to arrive at a clone capable of producing drugs with pre-targeted glycosylation patterns, various tools can be utilized to engineer the glycosylation machinery of the host cell line. Alternatively, the glycans of the protein products can be chemoenzymatically remodeled. *In vitro* glyco-engineering of proteins has been demonstrated in several cases, such as the use of a series of exoglycosidases to remove terminal monosaccharides from  $\beta$ -glucocerebrosidase for the production of Cerezyme<sup>®</sup> [100] and, recently, the use of glycosyltransferases for the remodeling of IgG glycoforms to regulate the levels of terminal galactose and sialic acid [101]. In a different approach, Mimura *et al.* [102] demonstrated that, by substituting a phenylalanine residue at position 243 of the human IgG1 Fc region with alanine, the level of sialylated IgG glycoforms can be increased by 5–15-fold. The concept of host cell line glyco-engineering presents several advantages: (i) the glyco-engineered cell lines can have restricted glycosylation capabilities, hence reducing the overall glycan heterogeneity of their recombinant products, this could prevent the synthesis of unfavorable glycan structures or structures not observed on the reference materials; (ii) it is possible for glyco-engineered cell lines to produce optimized glycoforms, which would represent a small fraction of the total glycoforms in the wild-type cells. This would be highly desirable for the production of biobetter versions of therapeutic proteins, such as afucosylated mAbs.

#### *Lectin-based glycosylation mutant isolation*

Glyco-engineering of host cell lines can be done through untargeted screening [103,104] or targeted, gene–genome editing tools [105,106]. In the former approach, a common methodology is to use cytotoxic lectins for negative selection of glycosylation mutants [103]. Lectins can be rationally selected based on the targeted glycosylation feature. *Maaackia amurensis* agglutinin (MAA) is specific for sialic acid ( $\alpha$ 2,3)Gal linkage [107,108], which is expressed on CHO cells [109]. We previously reported the use of MAA for the isolation of a CHO glycosylation mutant (R11, now

renamed as CHO-gmt1) with a nonsense mutation in the gene *SLC35A1* encoding the CMP-sialic-acid transporter [110]. CHO-gmt1 was shown to contain only a basal level of sialic acids and no significant amount of sialylated glycans was found on the recombinant interferon- $\gamma$  produced from this cell line [110]. Theoretically, MAA can also be used to isolate glycosylation mutants with loss-of-function genetic lesions in the nonredundant steps before sialylation along the glycan-processing pathway. Indeed, we have isolated a CHO mutant of UDP-galactose transporter using MAA [104] (Song *et al.*, unpublished). Therefore, it is possible to aim for different asialo glycan structures using MAA. One intriguing discovery by Song and co-workers is that *Ricinus communis* agglutinin (RCA-I) only negatively selects for mutants of *N*-acetylglucosaminyl transferase I (GnT-I) [111]. It was shown that a panel of CHO mutants selected by RCA-I treatment all harbor mutations in GnT-I, despite different types of genetic lesions that range from insertion, deletion, frameshift, missense to nonsense mutations [111]. Therefore, RCA-I-based mutant isolation strategies can effectively function as a GnT-I-targeted gene editing (i.e. knockout) tool. Consequent to the loss-of-function mutation in GnT-I, the mutants accumulate *N*-glycans at the Man5 stage, making them desired hosts for the production of oligomannose-type glycoforms such as  $\beta$ -glucocerebrosidase. Interestingly, complementing such mutants with a functional GnT-I results in not only the restoration of the *N*-glycosylation pathway but also an increase in the sialylation of *N*-glycans [111]. This concept was validated in a perfusion cell culture process and demonstrated its capability of enhancing the sialylation of recombinant erythropoietin [112].

#### Targeted gene editing tools

Although screening-based glycosylation mutant selection using lectins represents a highly efficient approach, it is challenging to select for glycosylation mutants where there exist multiple copies of the same glyco gene or functional redundancy. With the advent of gene editing tools, such as zinc finger nuclease (ZFN), transcription-activator-like effector nuclease (TALEN) and CRISPR-Cas9, it is now possible to create any nonlethal knockout by targeting a single or a set of genes for cell engineering.

#### ZFNs

ZFNs are unnatural nucleases consisting of arrays of zinc-finger-binding proteins linked to the cleavage domain of an endonuclease, typically FokI. The zinc-finger motif specifically binds the target site on the chromosomal DNA. A pair of ZFNs binds to complementary strands of the DNA, bringing the FokI domains to dimerization, which in turn generates a double strand break. As the nonhomologous end-joining mechanism is activated in the cell to repair the DNA, insertion or deletion mutations (indels) can be introduced, which often results in functional knockout of the gene. The key to the success rate of ZFN-mediated gene knockouts is the target site and the corresponding zinc-finger sequence. An open-access pathway for zinc-finger design is the modular assembly method [113], in which each finger is designed to target a specific DNA triplicate. The fingers that bind the 5'-GNN-3' triplets are the best studied and strong DNA-binding fingers [113–115]. Three-fingered ZFNs or four-fingered ZFNs are commonly used. For four-fingered ZFNs, the sequence should match the ideal target sequence which is: 5'-NNCNCNCNCNCNxxxxxxGNNGNNGNNGNN-3'. The web-based ZiFiT program provided by the Zinc Finger Consortium [ZiFiT:

software for engineering zinc finger proteins (V3.0)] (<http://bindr.gdcb.iastate.edu/ZiFiT/>) [116] can help to identify potential sites for ZFNs in the genes. Using ZFNs designed through the modular assembly method, we previously demonstrated the knockout of the Golgi GDP-fucose transporter in CHO cells [105], which were then shown to produce fucose-free glycoproteins.

#### TALENs

TALENs are customized transcription-activator-like effectors (TALEs) linked to FokI nuclease [117,118]. TALEs contain a central DNA-binding domain with several tandem repeats of 34 amino acids. Each repeat differs mainly in two amino acid residues at the 12th and 13th positions, the repeat variable di-residues (RVDs). Each TALE repeat binds one target nucleotide base as determined by the RVDs, with the following specificity: NI = A, HD = C, NG = T, NN = G or A [119,120]. The online tool TALE Effector Nucleotide Targeter 2.0 is available (<https://tale-nt.cac.cornell.edu/node/add/talen>) [117], to assist the identification of the targeting site. Binding sites of TALEs always begin with a T (thymine). Different scaffolds are available to generate TALENs with repeats typically ranging from 15 to 20 RVDs [117,118] with spacer length of 12–21 bp. For the assembly of TALENs, Golden Gate methods [121] or high-throughput platforms [122] can be applied.

#### CRISPR-Cas

The CRISPR-Cas system was first discovered in bacteria and archaea to direct the degradation of complementary sequences present within invading viral and plasmid DNA [123,124]. Components of CRISPR include two noncoding RNAs: crRNA (CRISPR RNA) fused to tracrRNA and CRISPR-associated nuclease (Cas9). crRNA and tracrRNA together are referred to as sgRNA (single guide RNA) [123,124]. Requirements for CRISPR-Cas9-based gene targeting include an initiating G, which facilitates RNA Pol III transcription at U6 promoter region, an –NGG PAM (protospacer-adjacent motif), followed by a 20 bp crRNA target. In theory, sgRNAs can be designed to target any genomic sequence that has the consensus sequence GN<sub>19</sub>NGG to induce a double-strand break [124]. This has allowed the targeting of multiple genes simultaneously. Researchers from the Technical University of Denmark and KAIST reported the use of CRISPR-Cas9 to knockout Fut8, Bak and Bax in CHO cells [125]. This has opened the door to fine-tuning glycosylation and improving cell line characteristics (e.g. high productivity, resistance to cell death and viral infection) in a rapid and cost-efficient manner. Web-based tools are available for designing CRISPR-Cas-targeting constructs, for example CHOPCHOP (<https://chopchop.rc.fas.harvard.edu/>) [126]. However, off-target effects are a major concern because there is sequence homology throughout the genome [127,128]. Off-target effects that potentially impair the cell culture performance, for example cell growth, recombinant gene expression and quality control, should be eliminated by carefully selecting the gene target loci and optimizing the targeting constructs.

#### QbD for glycosylation control in upstream bioprocesses

QbD is a conceptual framework that should be adopted in manufacturing activity so as to build the product quality into the process via detailed designing in the product development stage [129]. With respect to the pharmaceutical industry, the full implementation of QbD on biologics manufacturing is more challenging compared with that of small molecule drugs because of the complexity of the biotechnology-derived products [130]. It requires a

thorough understanding of the process and aims at ensuring end-product quality through working within carefully outlined design space for crucial process conditions. To apply a QbD framework, first the desired clinical performance of the drug must be identified to establish the quality target product profile (QTPP). Based on this profile, related CQAs can be determined together with the process parameters that can impact these CQAs [131]. By contrast, it is also necessary to monitor the process parameters in a timely fashion via process analytical technology (PAT) to ensure the previously defined specifications for CQAs are met [130].

#### Glycosylation-related CQAs

Earlier, we mentioned that, where CDC is concerned, the extent of galactosylation [132] of a mAb is considered a glycosylation-related CQA (gCQA), whereas removal of the core fucose residue can increase ADCC dramatically (100-fold) [9,133]. Apart from impacting the therapeutic efficacy of the drug, the presence of other glycan motifs, such as  $\alpha$ 1,3-linked terminal galactose, posed a problem for the safety profile of the drug cetuximab [7]. All these above-mentioned glycan structures can be used to construct an array of gCQAs under the QbD framework.

Pivotal to the implementation of QbD is the pinpointing of the gCQAs followed by monitoring and controlling them. Ludger has proposed a systematic approach to identifying and ranking the gCQAs via the safety and efficacy (SE) profiling and the construction of impact maps [134]. SE profiling is a visualization tool aimed at revealing those gCQAs that alter the clinical performance of the drug. Impact maps are mathematical graphs showing the effect of these gCQAs on patient safety and clinical efficacy.

#### Upstream bioprocess parameters that affect glycosylation

To exert control over the predefined gCQAs, it is important to understand the impact of manufacturing conditions on the glycan distribution in the product. Apart from host cell line selection and engineering, the conditions used in running the bioreactor have been shown to affect the glycan profile on the biotherapies in a product-specific manner. Among the many process parameters, dissolved oxygen (DO) levels, culture temperatures of bioreactor and nutrient or supplement availability are commonly controlled and their effects on glycosylation have been studied extensively. Some of the findings are summarized in Table 2.

#### High-throughput and online-ready technologies for QbD implementation

Implementation of QbD to biologics manufacturing requires detailed characterization of the process and thorough exploration of the various conditions that can affect the gCQAs as described in Table 2. Besides being able to run many conditions in micro bioreactor systems such as Ambr<sup>TM</sup> workstation, Biolector<sup>®</sup> and Micro-24 MicroReactor System, a high-throughput glycan-profiling platform is also needed to assess the impact of process variables on the glycan profiles quickly in a cost-effective way. GlykoPrep [135] is such a commercial kit that allows glycan profiling in 96-well plates using proprietary reagents and micro-cartridges. Alternative platforms for high-throughput glycan analysis have also been reported. Doherty *et al.* used a protein A immobilization strategy to achieve efficient glycan release by PNGase F together with sample purification and concentration

[136]. Beads functionalized with hydrazide groups can conjugate with aldehyde groups from sugars enabling the purification of glycans [137,138]. Utilizing this chemistry, Stockmann *et al.* reported a novel profiling workflow with full automation [139] and used it to profile glycans from IgG in serum samples. Instead of analyzing the released glycan, Reusch *et al.* presented a workflow in which the IgG samples were subjected to trypsin digestion and the resulting glycopeptides analyzed by electrospray MS [140]. Although high-throughput technology can increase the data points one can explore simultaneously, it is also advisable to employ design of experiment (DoE) techniques to maximize the information obtained from the screening data. Its application in bioprocessing has been discussed in the review by Kumar *et al.*, with case studies of applications for chromatography and refolding steps [141].

Developing online methods for quick product characterization, including glycan profiling, is also crucial for QbD implementation because it enables real-time control over the bioreactor to inform manufacturing operations. At Genentech, an online mAb characterization strategy was proposed in which ion exchange or size exclusion separation was followed by trapping and desalting on reverse-phase cartridges with subsequent quadrupole time of flight (QToF) MS analysis [142]. Major glycoforms can be identified via MS analysis at the intact antibody level. Primack *et al.* from Amgen used a microchip-based CE method to profile the glycans from antibody cell culture [143]. The glycans were released from the purified mAbs and sent for CE analysis on the Caliper Labchip GXII Microchip-CE platform. M5, G0, G0F, G1F and G2F can be identified from the electropherogram. Agilent's multilayered microfluidic chip provided an integrated system for glycan release, clean-up, PGC-based separation and integrated electrospray into MS [144]. Only 100 ng of mAb was needed for the assay, and the whole analysis including sample preparation can be done within 10 min.

#### Quantitative models for predicting N-glycosylation in cell cultures

Although real-time monitoring of the detailed glycoforms during the biomanufacturing process still remains a challenge, now that the CHO cell genome has been sequenced [145] a systematic understanding of the pathways can offer some help in our ability to understand and eventually predict the effects of process changes on glycan profiles of the final product. Initial computational models [146–148] approximated the Golgi apparatus as a series of continuously stirred tank reactors (CSTRs) where reactions are centered on glycosyltransferases to generate more than 7500 potential glycan structures. Hossler *et al.* started to adopt the Golgi maturation model and used plug flow reactors in series to mimic the change in enzyme concentrations along the Golgi apparatus [149]. Follow-up models were developed to look into the microheterogeneity of antibody glycosylation specifically, and particular attention was paid to the nucleotide sugar donor (NSD) level to link glycosylation to cellular metabolism [150,151].

#### Downstream bioprocessing of glycoproteins

In common with other biopharmaceuticals, therapeutic glycoproteins typically require exceptional purity and constant potency, making effective and efficient downstream processing a crucial



TABLE 2

## Effects of cell culture conditions of on glycosylation.

Conditions in the bioreactor	Effect on the bioprocess or product	Mechanism of action
<b>Dissolved oxygen tension (DOT)</b>	For erythropoietin (EPO)-Fc produced in CHO, the sialylation fluctuated between 11 and 14 mol sialic acid/mol EPO-Fc for 10–90% DOT, but dropped drastically at 100% DOT	Mechanism not defined [106]
	For human follicle-stimulating hormone produced in CHO, maximum sialylation at 100% DOT For IgG1 produced in murine hybridoma cells, reduced DOT resulted in reduced galactosylation	Mechanism not defined [107] Not clear. One suggested reason is that it is due to decreased UDP-Gal transport from the cytosol to Golgi [108]
<b>pH/[NH<sub>4</sub><sup>+</sup>]</b>	For tumor necrosis factor receptor (TNFR)-IgG produced in CHO, as [NH <sub>4</sub> <sup>+</sup> ] increased from 1 to 15 mM terminal galactosylation and sialylation decreased up to 40%	It was hypothesized that [NH <sub>4</sub> <sup>+</sup> ] inhibits SiaT and GalT by increasing the pH of the trans-Golgi compartment [109]
	For EPO produced in CHO, at 10 mM NH <sub>4</sub> Cl, a reduction in sialylation was observed for both <i>N</i> - and <i>O</i> -linked glycans	It might be ascribed to an increase in intra-Golgi pH. Another possibility of increase in extracellular sialidase activity was ruled out [110]
	For IgG3 produced in mouse hybridoma cell line, total sialic acid content is in the order of pH 7.2 > HEPES > pH 7.4 > pH 6.9; highest relative amount of agalacto structures were found in pH 7.2 and pH 6.9 culture For interleukin (IL)-2 produced in BHK-21, at 15 mM NH <sub>4</sub> Cl there is a significant increase of tri- and tetra-antennary glycans	Mechanism not defined [111] The increase in complexity is considered to be brought by the increased levels of intracellular UDP-GlcNAc/GalNAc in the presence of high NH <sub>3</sub> /NH <sub>4</sub> <sup>+</sup> [112]
<b>Temperature</b>	For EPO produced in CHO, temperature range of 25–37 °C was used. The proportion of tetra-antennary structures reached maximum at 32 °C. The proportion of tetrasialylated glycans stayed roughly the same between 37 and 32 °C, but decreased by 24.4% when culture temperature dropped to 25 °C	The author found that high specific EPO productivity at high temperature is probably not the cause instead the observed difference might be attributed to enzyme activity reaching optimum between 32 and 37 °C [113]
	For EPO-Fc produced in CHO, when culture temperature was shifted from 37 to 30 °C there was a reduction of sialylation by 40%	Mechanism not defined [114]
<b>Nutrient supplement</b>	For interferon (IFN)-γ produced in CHO, lipoprotein-supplemented medium helps to maintain the extent of glycosylation with time in batch culture	Mechanism not defined [115]
	For monoclonal antibodies (mAbs) produced in murine hybridoma, higher glycan microheterogeneity was found for culture supplemented with fetal bovine serum (FBS); G0 content was higher with FBS supplementation; with FBS addition, increase in sialylation and decrease in fucosylation were also observed	It was mentioned that extracellular β-galactosidase activity increases threefold for serum-free medium [116]
	For IFN-γ produced in CHO, glucose limitation and glutamine limitation can reduce the site occupancy of the product	Nucleotide triphosphate depletion was thought to be the cause of decrease site occupancy during glucose limitation. Whereas glutamine limitation influenced glycosylation by reducing amino sugar formation [117]
	For chimeric heavy chain antibody (EG2) produced in CHO, deprivation of glucose resulted in a 45% nonglycosylated mAb fraction	When glucose is not sufficient, it will be preferentially used for energy metabolism rather than replenishing the nucleotide sugar pool that serves as the precursor to glycosylation [118]
	For IFN-γ produced in CHO, very low glutamine (<0.1 mM) or glucose (<0.7 mM) can lead to decrease of sialylation and increase of hybrid and high-mannose glycans For human chorionic gonadotrophin (HCG) produced in CHO, in continuous culture setting, glutamine supplement concentration was alternating between 0 mM and 8 mM for 90 days. 15 glycan structures from HCG were profiled at various time points. Reduction in sialylation and antennarity of HCG was observed with a decrease in glutamine coupled with low glycolytic flux	Glutamine limitation can limit the formation of UDP-GlcNAc by limiting amino sugar formation. This will in turn decrease the formation of CMP-NeuAc [119] From metabolic flux analysis and monitoring of nucleotide sugar metabolism, it was suggested that high glutamine concentration contributes to the accumulation of UDP-GlcNAc and UDP-GalNAc which are required by glycotransferases when adding new glycan chains to biantennary structures. UDP-GlcNAc is also a necessary precursor for the synthesis of CMP-NeuAc and hence impacts sialylation [120]
For IgG produced in human cell line rF2N78, aglycosylated antibody was produced when glucose was depleted from the medium. Alycosylated antibody was not observed with glucose feeding in fed-batch culture	Glucose capping on oligosaccharide is a key step in the oligosaccharide recognition by the oligosaccharyl transferase, which eventually leads to glycosylation at the asparagine site [121]	

step in the overall production process [152]. The generic process for the recovery and purification of glycoproteins produced by animal cell culture consists of successive capture, purification and polishing steps, each comprising one or more procedures [153]. Protein A chromatography is the industry standard for commercial-scale purification of mAb products. In the research environment, however, the capture step might employ lectin or boronate affinity chromatography by selectively capturing the glycan component. Purification and polishing are conducted with ion exchange and hydrophobic interaction chromatography to remove aggregates, endotoxins, host cell proteins and DNA [154]. In addition, some intermediate, nonchromatographic purifications are also employed such as ultra- and dia-filtration as well as clarification techniques [155–157].

#### Lectin-based purification

Lectins are carbohydrate-binding proteins and have been widely used to identify and isolate glycoproteins, glycopeptides, glycolipids and oligosaccharides [158]. Lectins are found in plants and animals, and have specificity for certain types of carbohydrate residues. The most common lectins currently being used are concanavalin A (Con A), which has specificity for  $\alpha$ -D-mannose,  $\alpha$ -D-glucose residues and bi-antennary *N*-glycans [159], wheat-germ agglutinin (WGA) with specificity for *N*-acetylglucosamine residues [160], jacalin with specificity for galactose and mannose residues [161], MAA and SNA, with specificities for Sia  $\alpha$ 2,3- and  $\alpha$ 2,3-linked Gal, respectively [108]. Additional lectins are available for binding to less common glycans. EPO, influenza vaccine and other therapeutic glycoproteins can be purified using lectin affinity chromatography [162,163]. However, as a class of biological receptors, lectins have significant disadvantages including multiple specificities, expensive purification costs, degradation during standard sterilization and subsequent contamination of the end-product as a result of ligand leaching.

#### Boronate-based purification

Since 1970, boronate columns have been exploited as less-expensive and more-stable alternatives to lectins for the separation of glycoproteins. The most commonly used boronate adsorbent is *m*-aminophenylboronic acid agarose, which has been investigated for the specific capture of glycoproteins from protein mixtures; successful applications include the selective capture of IgG from CHO cell supernatant [164], EPO from mammalian cell cultures [165], ovalbumin from egg white [166], ovalbumin and ribonuclease B (RNase B) from BSA [167], horseradish peroxidase (HRP) from BSA, RNase B from myoglobin [168] and glycopeptides from non-glycopeptides of digested HRP [169], among others. However, to form stable cyclic esters between *m*-aminophenylboronic acid and *cis*-diol compounds, alkaline conditions are required ( $\text{pH} \geq 8.5$ ) [170]. It is not desirable to perform glycoprotein separation at such high pH, because it can lead to the degradation of some labile glycoproteins. To address the challenge, new boronate ligands that can work at physiological conditions have been studied and applied to the enrichment of glycoproteins [171,172].

#### Future needs for cost-efficient glycoform separation technologies

Unlike protein synthesis, which is under genetic control, glycan synthesis is not directly template-mediated and thus glycoproteins

are typically expressed and marketed as populations of glycoforms with potentially different biological properties [173]. The different properties associated with each glycoform within the resulting heterogeneous mixture present regulatory difficulties for therapeutic glycoproteins and also problems in determining exact structure–activity relationships [173]. It is therefore important to the biopharmaceutical industry to be able to produce products with homogeneous glycoforms that can confer significant and predictable therapeutic advantages. However, protein production costs remain extremely high, with downstream processing constituting a substantial proportion (50–80%) of the overall cost [174]. It is expected that the development of new purification techniques exploiting inexpensive, specific, effective, robust and easy-to-use methods and materials will shape the future of glycoprotein purification [175].

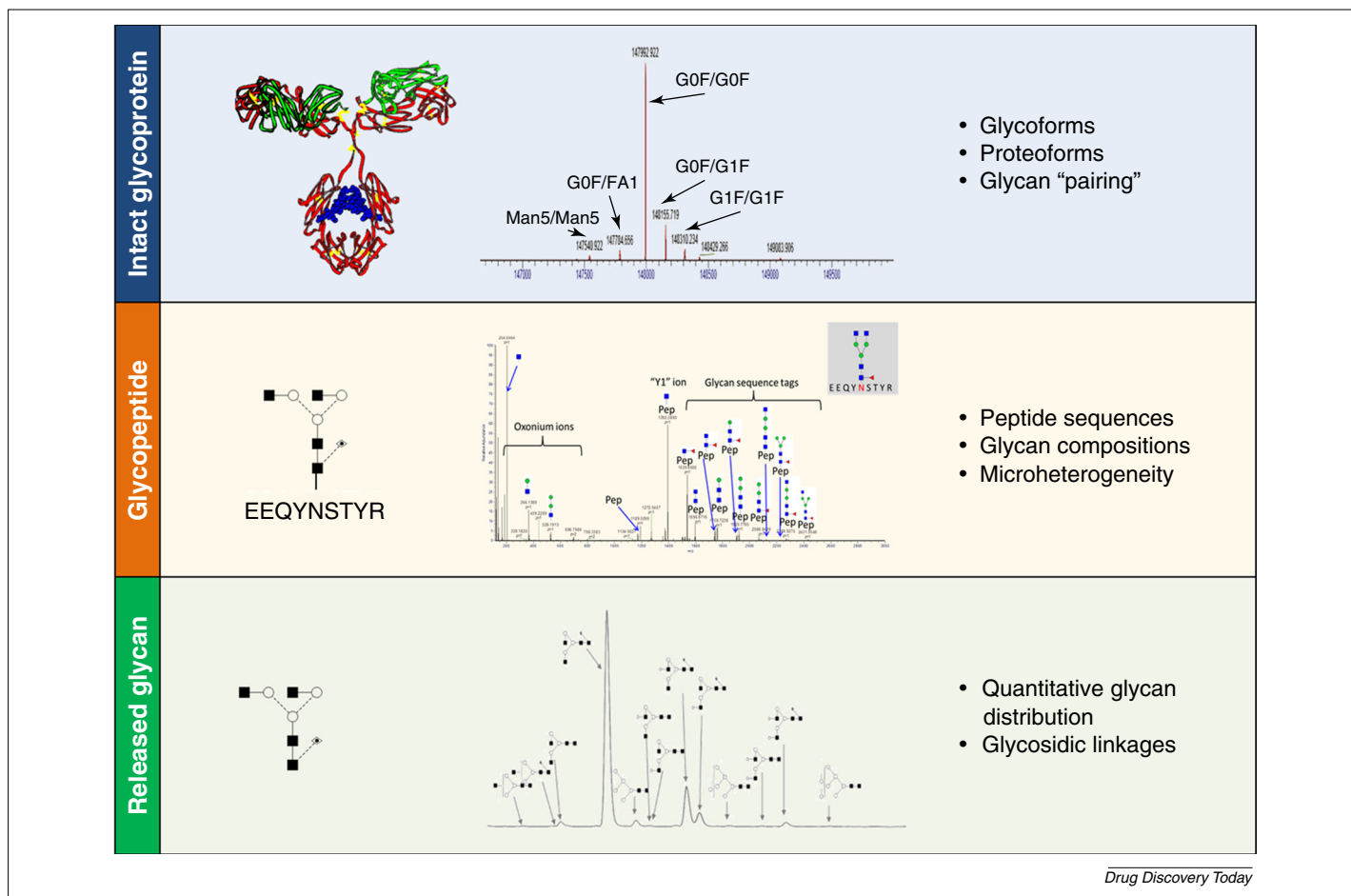
#### Technologies for glycosylation analysis

The absence of a direct genetic blueprint and the nonlinear structures of glycans makes glycosylation analysis one of the most challenging physicochemical characterizations of biopharmaceutical products. The past two decades have seen tremendous developments in analytical instruments and knowledge of protein glycosylation. It has now become a consensus that, in glycosylation analysis, the question dictates the exact approach [176]. Broadly speaking, glycosylation analysis of therapeutic proteins can be performed at three levels, namely, intact protein, glycopeptide and released glycans, as illustrated in Fig. 8.

#### Intact glycoprotein analysis

There are a variety of analytical methods for characterization at the intact protein level, which include chromatography, electrophoresis, MS and other spectroscopic techniques. Owing to the superior resolution, MS has become a crucial analytical tool to characterize mAbs in many phases of the developmental process. Through top-down approaches that deal with intact protein samples, information such as molecular weight, and major modifications such as glycosylation, can be obtained. In top-down analysis, high-resolution mass analyzers are required, including ToF, Orbitrap and Fourier transform ion cyclotron resonance (FT-ICR) instruments to resolve small mass differences [177]. Valeja *et al.* showed that it is possible to achieve isotopic resolution through the use of FT-ICR for the analysis of intact IgG1 [178]. However, even with the use of FT-ICR, it is still challenging to resolve small mass changes such as deamidation, dehydration and oxidation [179]. In addition, FT-ICR is not a very commonly used instrument in the biopharmaceutical environment because of the tedious maintenance of the superconducting magnets and its high price.

Intact protein samples for top-level analysis usually undergo an online or offline separation step (e.g. LC, CE) before ionization by electrospray (ESI) or matrix-assisted laser desorption (MALDI) MS [180]. Although direct infusion of the intact mAbs is also possible, a desalting step is often required to obtain high-quality spectra with fewer salt adducts and better ionization efficiency. Using an ESI source, intact IgG samples can form quasi-molecular ions with high charge state ( $z$  in the range of 20–60) and the resulting  $m/z$  values in the range of 2000–5000 amu, which can be measured by many types of mass analyzers. By contrast, an intact IgG sample

**FIGURE 8**

The top, middle and bottom levels of glycosylation analysis. Depending on the exact objective, glycosylation analysis can be performed at different levels. This picture depicts the top, middle and bottom levels of glycosylation analysis in analogy to protein analysis, using monoclonal antibody drugs as an example. In top-level analysis, the intact glycoprotein is analyzed by mass spectrometry either through direct infusion or preceded by a separation system such as a liquid chromatograph. The high-resolution accurate mass, (HR/AM) mass spectrometer measures, to high precision, a series of mass:charge ratios ( $m/z$ ) of the IgG under question, which arise from a consecutive series of charge states. Using a computer-assisted deconvolution algorithm, neutral mass (molecular weight) of the IgG can be derived. Typically for monoclonal IgG molecules, several neutral masses are revealed after deconvolution that correspond to different glycoforms and/or proteoforms. By examining the mass difference among the different forms of the intact IgG, and comparing with the mass(es) of the deglycosylated IgG, each neutral mass can be assigned to a particular glycoform or proteoform. Here, we show our data on a monoclonal antibody product, showing mainly G0F/G0F glycoform. The top-level analysis offers a unique insight into the pairing of glycans on the two heavy chains of an IgG molecule. We refer intact glycopeptide analysis as middle-level analysis. This typically involves proteolytic digestion of the glycoprotein, followed by LC-MS/MS analysis. Fragmentation of the glycan moiety can produce rich information on the composition and arrangement of the monosaccharide building blocks on the glycan. Middle-level analysis serves to answer the question of glycosylation microheterogeneity because it supports the elucidation of the carrier peptide sequence in addition to glycan compositional analysis. This is often necessary in the case of multiple glycosylation sites on a protein. Peptide-specific fragmentation is now better supported with the development of new ion fragmentation techniques, such as electron-transfer dissociation (ETD). In bottom-level analysis of glycosylation, glycans are cleaved off the carrier protein and analyzed as a pool. Popular techniques for released glycan analysis involve tagging the glycans with a fluorescent dye and separating the labeled glycans by LC-capillary electrophoresis (CE), with or without introducing them to a mass spectrometer. Away from the carrier proteins or peptides that often dominate the overall physicochemical properties of the glycoconjugate, bottom-level analysis of released glycans shows high sensitivity and quantitation power, which supports in-depth characterization of glycans including their compositions, sequences and glycosidic linkages. However, once detached from the protein backbone, the information on the local context, such as site of attachment and pairing, is lost.

typically carries a maximum of four charges under MALDI conditions, which diminishes the resolution and mass accuracy [181].

High mass resolution required in the analysis of top-down analysis of intact monoclonal IgG antibodies without the use of FT-ICR is also achievable using hybrid QToF [182] or hybrid linear quadrupole ion trap-Orbitrap MS [183]. Through the use of the Orbitrap analyzer, Bondarenko *et al.* employed three different delivery methods: (i) direct nano-ESI infusion; (ii) step-elution; and (iii) online reversed-phase HPLC, to determine major glycoforms with different numbers of terminal galactose ( $\pm 162$  Da)

residues accurately in the intact IgG from the deconvoluted mass spectra [184]. The analysis of intact mAbs represents a fast and robust way to determine batch-to-batch consistency (analysis time is a few minutes), in which minimal sample preparation is required. However, the detection of minor glycoforms remains a challenge. In a similar fashion, antibody subunits or fragments can be analyzed by middle-down strategies based on high-resolution-accurate MS platforms coupled with appropriate separation fronts. Different chemical and enzymatic approaches, such as reduction of inter-chain disulfide bonds by reducing agents and specific

proteases (especially IdeS), have been demonstrated to support simultaneous elucidation of glycoforms and other PTMs located in the heavy chain of antibody molecules [185].

### Glycopeptide analysis

The analysis of glycosylation microheterogeneity is essential to understanding the site-specific contribution of glycan structures on the safety and efficacy of drug products. Typical of bottom-up approaches, a glycoprotein sample is first digested by proteolytic enzymes, such as trypsin, then analyzed by separation techniques (e.g. HPLC or CE) or MS. MALDI or ESI are often used as the soft ionization methods to preserve the structural integrity of glycopeptides. The applicability of MALDI for high-throughput analysis has been recently demonstrated in a study profiling the *N*-glycans of the Fc region in IgG through glycopeptide analysis [186]. Such methods could be implemented when analyzing large batches of samples such as during process development or batch-to-batch monitoring. Competitive ion suppression (glycopeptides vs peptides) is often a major hindrance resulting from the complexity of enzymatic digestions of the glycoprotein. Therefore, various methods for glycopeptide enrichment have been developed, such as the use of size exclusion chromatography [187] and hydrophilic interaction chromatography with cellulose or sepharose [188]. Alternatively, an online hydrophilic liquid chromatography (HILIC) separation that is coupled to nano-ESI-MS or UPLC-MS could also be employed based on the hydrophilicity of the glycopeptides [189–192].

Tandem mass spectrometry (MS/MS) is required to elucidate the sequences of the carrier peptide, the glycan composition and the site of glycan attachment. Collision-induced dissociation (CID), such as high-energy C-trap dissociation (HCD) in Orbitrap instruments, generates distinct product ions of peptide with the innermost GlcNAc residue of an *N*-glycan (i.e. the  $y_1$  ion), which can be used to determine the *N*-glycosylation site [193]. Using database searching for bottom-up proteomic measurements with the core GlcNAc as a variable modification, the identity of the glycopeptide together with the *N*-glycosylation site can be determined. The presence of glycan oxonium ions (e.g. *b* ions at  $m/z = 204$  for HexNAc, and 366 for HexNAc + Hex, and 276 for dehydrated Neu5Ac) at low mass range are also used as a marker for glycosylation. CID fragmentation of glycopeptides almost exclusively results in the cleavage of glycans from the peptide backbone to generate a series of *y*-fragment ions related to glycosidic cleavages, at higher mass range [193,194]. Thus, CID is a suitable technique for determining glycan compositions and even structures based on the glycan sequence tag. The use of QToF-MS/MS to obtain CID fragment-type ions has been widely used in the analysis of sialo- and asialo-complex-type *N*-glycopeptides to obtain information on the glycan moiety [195]. In more recent papers, information on peptide sequence, glycan moiety and glycosylation site were obtained within one analytical run with the use of alternating high and low collision energy ( $MS^E$ ) [182]. Precautionary measures have to be taken when analyzing the glycan moiety from the CID spectra of glycopeptides, because of the possibility of fucose rearrangement, which could be mistaken for *b*- and *y*-ions [196,197].

Newer fragmentation techniques, such as electron-activated dissociation (ExD), electron capture dissociation (ECD) and electron transfer dissociation (ETD), preserve labile PTMs such

as glycosylation [198–200] but selectively fragment the peptide backbone, allowing the determination of peptide sequence and *N*-glycan attachment sites through the generation of peptide fragment ions and a mass-shift resulting from the attached glycan [201]. ExD complements CID and the two techniques can be combined to obtain information about the glycopeptide. Analysis of the glycopeptide can be done on MALDI-MS/MS to provide information of the peptide and glycan moiety. Depending on the configuration of the instruments, various distinct fragment ions can be observed. However, ExD only works efficiently for ions of high charge states and in the low  $m/z$  range; one way to circumvent this would be the addition of *m*-nitrobenzyl alcohol into the mobile phase to increase the charge-state of glycopeptides [202]. With the use of MALDI-ToF/ToF, cleavage of the side-chain amide bond of glycosylated asparagine usually occurs, giving a diagnostic peak together with *b*- and *y*-ions from the peptide backbone. Advances in instrumentation, particularly fragmentation techniques, now relocate the major bottleneck in glycopeptide analysis to the software solutions that are needed for high-throughput, high-confidence MS interpretation. This point will be discussed below.

### Analysis of released glycans

The determination of monosaccharide composition, glycosidic linkage and relative abundance of a glycan in a mixture (Fig. 1) is highly relevant to understanding the safety, efficacy and batch-to-batch consistency of glycoprotein therapeutics. Depending on the exact analytical requirement, glycans released from a protein carrier can be profiled, characterized and/or analyzed in-depth. Glycan analysis poses significant technical challenges owing to the lack of a chromophore in the natural carbohydrate molecules and structural heterogeneity caused by anomericity, linkage positions and configurations. Therefore, released glycans are often fluorescently labeled for sensitive detection [203] or derivatized to improve the ionization efficiency and minimize in-source or post-source decay in MS analysis [204]. Over the past two decades, high-performance tools have been developed for structural analysis of released glycans, mainly based on high-resolution separations (e.g. HPLC/UPLC, CE) and MS [176].

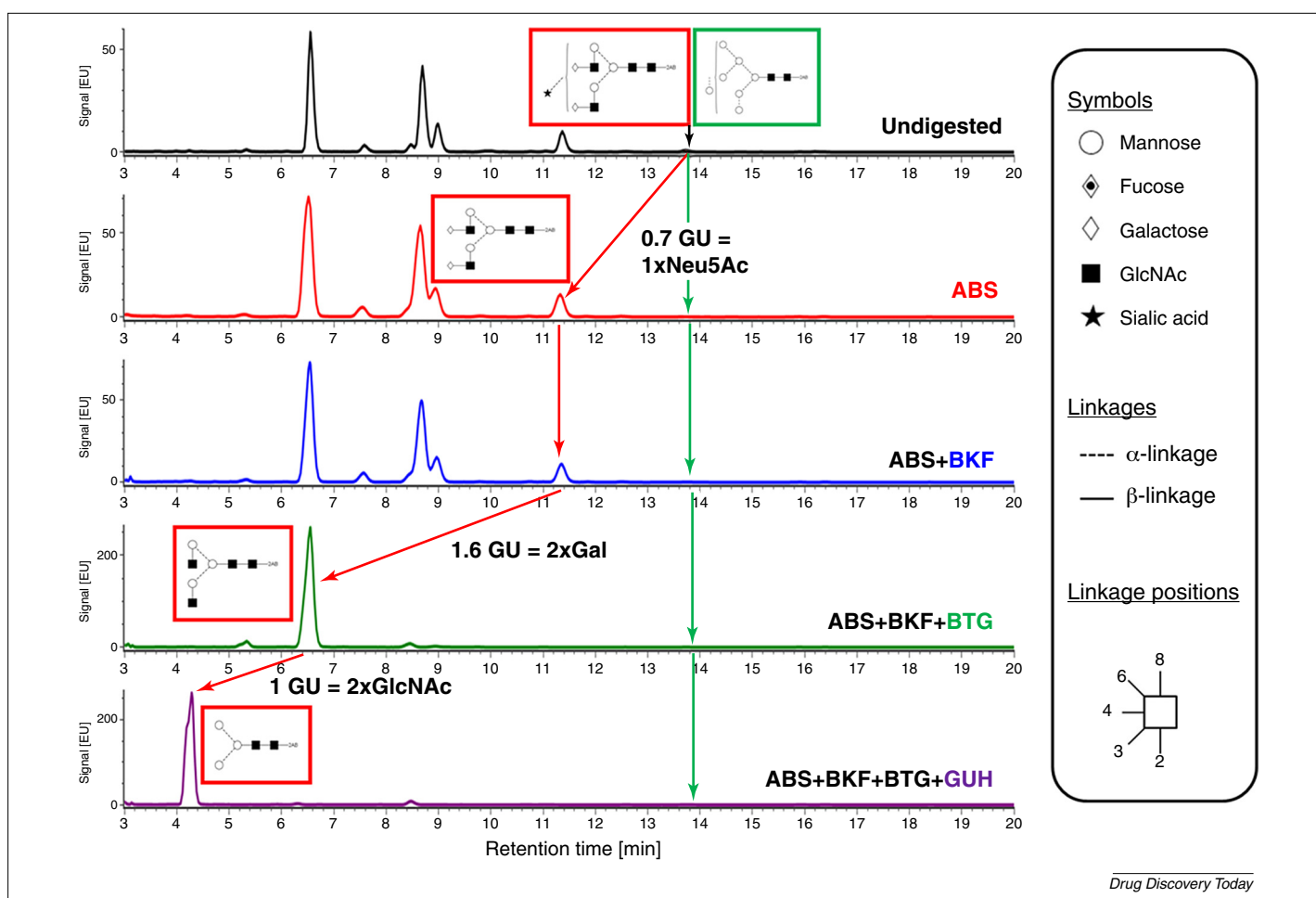
### Liquid chromatography

HPLC-based glycan analytical workflows typically involve fluorescent labeling of glycan species. Several fluorescent tags have been attached to *N*-glycans, such as 2-aminobenzamide (2-AB) and 2-aminobenzoic acid (2-AA) [203]. Recently, two new fluorescent labels for *N*-glycans have been reported, namely RapiFluor-MS™ (RFMS) [205] and 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AQC) [206], both based on the rapid reaction of glycosylamine with *N*-hydroxysuccinimide-activated carbamate. The use of UPLC has greatly reduced the amount of time required for data collection and gives good resolution because it can withstand higher pressures as a result of the smaller particle size (often less than 2  $\mu\text{m}$ ), in comparison to HPLC. UPLC has become the method of choice for efficient separation owing to shorter separation time, better resolution, particularly of arm-specific isomers, and relatively easy coupling to mass spectrometers [207,208]. The use of HILIC as a separation mode and a fluorescently labeled dextran ladder as an external calibration standard has enabled fast and highly reproducible glycan analyses across



different systems and operators [209,210]. A dedicated experimental database (i.e. GlycoBase) contains normalized retention time values (GU values) for a wide collection of glycan species analyzed by HILIC–HPLC/UPLC, in addition to data collected from RP–HPLC and CE–LIF platforms [211]. Coupled with exoglycosidase array digestions (Fig. 9), this has enabled relatively straightforward structural assignment [211,212]. By specifically removing terminal monosaccharides of particular linkages, exoglycosidase array digestion also helps identify co-eluting glycan species (Fig. 9). Recently, a robotized and low-cost *N*-glycan analysis platform to profile IgG was developed. The platform includes features from rapid IgG purification from cell culture through affinity binding to the labeling of released glycans in a 96- or 384-well format with excellent reproducibility based on UPLC analysis [139]. The development of high-throughput platforms for the analysis

of IgGs will prove to be invaluable for clonal selection of mAb drugs, development of biosimilar candidates and genome-wide association studies [14,213]. An added advantage in using UPLC is its ease of coupling with MS technologies, which can be used as an orthogonal method for glycan analysis. One commercially available LC–MS technology for detailed glycan analysis is the UniFi<sup>®</sup> biopharmaceutical platform. The UniFi<sup>®</sup> platform incorporates HILIC–UPLC-based glycan separation, fluorescence detection and online high-resolution mass confirmation, as well as integrated informatics for glycan identification based on normalized retention time (expressed as GU values) and accurate mass. The HILIC–UPLC–MS approach is especially powerful in light of a new *N*-glycan labeling reagent, RFMS, which gives up to 160-times higher MS sensitivity to *N*-glycans compared with the traditional 2-AB label [205].



**FIGURE 9**

Determining monosaccharide compositions and linkages using exoglycosidase array digestion of *N*-glycans. In case of ambiguity caused by co-elution of glycan species, an array of exoglycosidases can be used to remove monosaccharide building blocks sequentially from the nonreducing (outer) to the reducing (inner) end. The undigested and digested glycan samples are analyzed by HILIC–UPLC. By following the movement of chromatographic peaks in response to each of the enzymes, one can deduce the sequence and linkages of monosaccharides of the glycan in question. In this figure, we show two IgG-derived *N*-glycans that showed close retention times, namely A2G2S1 (boxed in red) and Man7 (boxed in green), in the undigested chromatogram. The array of enzymes used in this example include: *Arthrobacter ureafaciens* sialidase (ABS), bovine testes β-galactosidase (BTG), bovine kidney alpha-fucosidase (BKF), β-*N*-acetylglucosaminidase (GUH). Depending on the specificity of the exoglycosidase, co-eluting glycans can be separated apart. In this case, treatment of A2G2S1 removed the sialic acid and resulted in forward shift of 0.7 GU (indicated by the red arrow) which is characteristic of a Neu5Ac residue. BKF treatment on top of ABS did not change the glycan profile, thus confirming the absence of fucosylated glycans in this sample. A further digestion using BTG shifted the A2G1 structure forward by 1.6 GU to the position of A2, which suggested the removal of two Gal residues. Finally, a further digestion by GUH shifted the A2 forward by 1 GU to the position for Man3, which suggested the removal of two GlcNAc residues. Throughout the digestion, Man7 remained at the original elution envelope (indicated by the green arrow). Taken together, the exoglycosidase array enables the deconvolution of co-elution structures and in-depth analysis of glycans in a highly quantitative and linkage-specific manner.

### Capillary electrophoresis

Capillary electrophoresis with laser-induced fluorescence (CE-LIF) is another powerful technique for analyzing released *N*-glycans that are typically labeled with 8-aminopyrene-1,3,6-trisulfonate (APTS). The separation of the labeled glycans is based on the charge:hydrodynamic-volume ratio. Similar to HILIC-HPLC/UPLC, there is an experimental database for normalized retention, expressed in GU values, of glycan species, for CE-LIF-based analytical platforms (<http://glycobase.nibr.t.ie/>) [214]. CE-LIF can be used alone or in conjunction with UPLC as a 2D method [214]. Like UPLC, CE-LIF can be coupled with MS. However, it still remains a challenge and less explored in the field of glycomics. A similar technique to CE-LIF is capillary gel electrophoresis with laser-induced fluorescence (CGE-LIF) based on a DNA sequencer for APTS-labeled-glycan analysis [215]. This technique is capable of automated high-throughput glycoprofiling. However, owing to the lack of a publically available database, assignment of the glycan structures remains a challenge

### Mass spectrometry

The most widely used ionization methods for glycan analysis by MS are MALDI and ESI. MALDI is a fast and robust method to characterize and profile neutral glycans. However, MALDI-MS leads to dissociation of labile glycosidic bonds such as sialic acids, thus the need for the glycans to be permethylated or the sialic acid residues derivatized [216]. ESI is a milder ionization method that can be coupled to online separation methods such as LC. Although ESI is capable of ionizing native glycans, it still suffers from poor ionization of the hydrophilic glycans. In many cases, single-stage mass spectrometry (MS1) cannot confidently assign structures to glycan-related ions because of ambiguity arising from isobaric and isomeric structures. MS/MS, using common fragmentation techniques such as CID, can resolve some of the issues but can be limited in solving some of the structural isomers. However, it has been shown that the use of low-energy CID on the protonated adduct of fucosylated *N*-glycans could lead to misleading fragment ions, caused by the rearrangement of fucose residues from other parts of the molecule [196]. Recently, Harvey *et al.* have shown the use of negative ion fragmentation provides more structural information (e.g. cross-ring fragmentation) as compared with the positive ion mode without the need to derivatize *N*-glycans [217,218]. A more recent fragmentation technique, electronic excitation dissociation (EED), has been implemented to give a more sensitive and higher-throughput analysis. Yu *et al.* have demonstrated that permethylation and labeling of the reducing end, together with the high resolving power of MS, enabled them to process accurately the complex yet rich structural information from EED spectra [219]. The use of negative electron transfer dissociation (NETD) as a fragmentation technique is also very useful for the analysis of glycosaminoglycans (GAGs), this results in a lower level of sulfate loss [220]. Advances in nanoflow LC-MS has led to the development of chip-based technologies (nanoHPLC chip-MS) that promise faster and more-sensitive glycoprofiling strategies for the analysis of glycans from monoclonal antibodies [144,221]. Recent publications demonstrate the use of ion mobility separation as an added dimension to MS (IMS-MS) for separating isobaric *N*-glycans from a mixture. The glycans separate on the basis of mass, charge and shape, determining their unique collisional cross-section (CCS) [218]. This provides complementary information to

the existing analytical methods. The potential of coupled technologies, such as LC-IM-MS, in the separation and structural analysis of complex carbohydrates has already been demonstrated, where the CCS of glycans and their fragments can be estimated from a set of calibration data (e.g. a dextran ladder or negatively charged *N*-glycans and their fragments) [218,222]. Yamaguchi *et al.* have demonstrated the use of 3D methods, for example UPLC-IMS-MS, to separate pyridylaminated biantennary isomeric *N*-glycans successfully [223], which could also prove to be a valuable tool in future glycomics studies.

### Informatics for glycosylation analysis and pathway modeling

The development of powerful analytical tools such as CE, HPLC and MS significantly promoted glycan data accumulation [204,224]. In response, plenty of software and databases have emerged to facilitate glycan data analysis. For example, Glycobase (<http://glycobase.nibr.t.ie/>) documents the experimental analytical data for a large collection of *N*- and *O*-glycans generated from liquid chromatography and capillary electrophoresis [211]. UniCarbDB curates data for glycans analyzed by LC-MS/MS [225]. There are several excellent reviews on currently available software and databases [212,226–228]. This section does not aim to review these databases but rather to focus on automated methods for MS-based glycan and glycopeptide analyses, the latter expected to gain more importance as a result of new antibody-based fusion protein drugs that contain multiple *N*- and *O*-glycosylation sites [10]. This will be followed by a review of current methods for glycosylation pathway modeling and analysis.

#### Data interpretation of MS-based glycosylation analysis

A multi-institutional study comparing glycan-profiling methods shows that MS is an efficient method for the characterization of glycan [229]. However, the automation of MS spectra interpretation is still a bottleneck in MS-based glycomics projects [230]. Several software platforms have been developed in the past few years serving this purpose. Here, we summarize some of the currently available tools on glycan compositional analysis through MS spectral data.

#### Database search

Table 3 summarizes the key search engines for glycosylation data processing based on library matching. GlycoMod is a web tool designed to find all possible glycan compositions by experiment mass [231]. It is based on a pre-generated *in silico* library of glycan masses. By comparing experiment mass and theoretical mass stored in the library, GlycoMod returns all compositions under the same experiment mass. It supports parameter restrictions such as type of glycan, derivatization and number range for each monosaccharide and also supports queries on glycopeptides when the peptide sequence is known. Input experiment masses are confined to singly charged ions or neutral species and because of lacking biological filters implausible structures are also included in the hit list. To avoid giving results with nonbiologically relevant carbohydrate compositions, GlycosidIQ [232] matches experimental data with theoretical glycofragment mass generated from GlycoSuiteDB (which is now integrated into UniCarbKB). Meanwhile, GlycoFragment and GlycoSearchMS, a set of web tools interpreting MS/MS spectra [233,234], has also been developed

TABLE 3

## Tools for mass spectrometry (MS)-based glycan analytical data processing based on a library search.

Name	Input MS stage	Databases used for the search	Supported oligosaccharide types	Link
GlycoMod	MS	In silico glycan database	N- and O-linked	<a href="http://web.expasy.org/glycomod/">http://web.expasy.org/glycomod/</a>
GlycosidiQ	MS <sup>2</sup>	A library of theoretical glycofragment mass generated from GlycoSuiteDB	N- and O-linked	Not maintained
GlycoSearchMS	MS <sup>2</sup>	A library of all possible product ions for MS/MS data generated from SweetDB	N- and O-linked	<a href="http://www.glycosciences.de/sweetdb/start.php?action=form_ms_search">http://www.glycosciences.de/sweetdb/start.php?action=form_ms_search</a>
GlycoPepDB	MS <sup>n</sup>	A library of 316 N-linked glycans reported in the literature	N-linked	<a href="http://hexose.chem.ku.edu/sugar.php">http://hexose.chem.ku.edu/sugar.php</a>
GlycoSpectrumScan	MS	A library of all potential oligosaccharide compositions and a library of in silico digested peptide mass of the sample	N- and O-linked	Not maintained
Cartoonist	MS	A library of previously assigned known structures	N-linked	No open access
CartoonistTwo	MS <sup>n</sup>	A library of previously assigned known structures	O-linked	No open access
Peptonist	MS <sup>1-2</sup>	A library of N-glycans generated by applying same rule as in Cartoonist and a library of all possible glycopeptides from sample	N-linked	No open access
GlycoPep ID	MS <sup>2</sup>	A library of theoretical <i>m/z</i> values of all plausible product ions	N-linked	<a href="http://hexose.chem.ku.edu/predictiontable.php">http://hexose.chem.ku.edu/predictiontable.php</a>
GlycoPep Grader (GPG)	MS <sup>2</sup>	A pool of [peptide + core component] ions calculated from a list of user input glycopeptide candidate compositions	N-linked	<a href="http://glycopro.chem.ku.edu/GPGHome.php">http://glycopro.chem.ku.edu/GPGHome.php</a> (cannot access now)
GlycoPeptide Search (GPS)	MS <sup>2</sup>	GlycomeDB and a library of in silico digested peptide mass of user input protein sequences	N-linked	<a href="https://edwardslab.bmcb.georgetown.edu/trac/GlycoPeptideSearch/">https://edwardslab.bmcb.georgetown.edu/trac/GlycoPeptideSearch/</a>
GlycoPepDetector	MS <sup>2</sup>	A library of <i>m/z</i> values of c-, z- and y-ions generated from user input glycopeptide candidate list	N-linked	<a href="http://glycopro.chem.ku.edu/ZZKHome.php">http://glycopro.chem.ku.edu/ZZKHome.php</a> (cannot access now)
GlycoMaster DB	MS <sup>2</sup>	A library of N-linked glycans extracted from GlycomeDB (user can append their own glycan data) and a library of <i>in silico</i> digested peptide mass from user specified protein sequences	N-linked	<a href="http://www.novo.cs.uwaterloo.ca:8080/GlycoMasterDB/">http://www.novo.cs.uwaterloo.ca:8080/GlycoMasterDB/</a>

employing mass fingerprinting and matches experiment mass with a library of theoretical spectra extracted from SweetDB, which comprises N-linked and O-linked glycans. A similar tool that searches against carbohydrate databases with glycans reported in the literature is GlycoPep DB [235]. Together with a library of potential peptide sequence masses, GlycoSpectrumScan [236] finds out glycoprotein compositions by analyzing single MS data. The user needs to input a list of possible oligosaccharide compositions of the sample and a list of peptide masses after *in silico* digestion of the sample protein. This software supports analysis of multiple charged ions, identification of N- and O-linked glycoforms and glycan assignment for proteins with multiple site glycosylation.

Cartoonist is an algorithm that helps to label MALDI-MS peaks with cartoons for permethylated N-linked oligosaccharides released from glycoproteins [237]. It matches peaks to potential

glycans by comparing a table of glycans with theoretical mass and abundance of isotopes. Rather than arbitrarily cutting off the low intensity peaks, Cartoonist uses isotope envelopes as a filter for mass matching. Different species have different databases in response to restrictions on sample source. Further modifications are added to enable O-linked oligosaccharide analysis from MS/MS data [238]. Apart from release glycan analysis, N-glycopeptide identification is achieved by combing single MS-characterized N-glycans with MS/MS-identified glycosylated peptides [239]. These programs help to specify oligosaccharide composition but do not give information concerning the type of bond or the isomeric oligosaccharide.

Web tools matching against theoretical product ions include GlycoPep ID, GlycoPep Grader, GlycoPeptideSearch (GPS), GlycoPep Detector and GlycoMaster DB. GlycoPep ID is specialized in identifying peptide portions from glycoproteins when nonspecific

enzymes are employed [240]. It calculates the theoretical  $m/z$  of all plausible product ions for input glycoprotein and compares the result with CID for a match. No scoring algorithm is included in this method to screen out the most likely glycopeptide candidate. GlycoPep Grader [241], another free web tool that helps to determine glycopeptide composition using MS/MS data, scores glycopeptide candidates using CID data in a charge-state-free manner. GPS is also a program for glycopeptide characterization using CID data [242]. MS/MS data were screened for glycosylation oxonium ions at  $m/z$  204 and 366 followed by matching of intact-peptide fragment ions with an *in silico* digested peptide list. Glycan structures are deduced by matching glycan compositions in GlycomeDB with the precursor masses after subtraction of previously identified peptide masses. GlycoPep Detector [243] is a web tool that targets ETD data analysis. It first generates a library of all possible fragmentation ions for known glycopeptides and then assigns a core for each input composition after searching their  $m/z$  value of c-, z- and y-ions against the library. Each time only one MS/MS spectrum is processed. GlycoMaster DB [244] is software for processing HCD and ETD spectra when interpreting MS/MS data of N-linked glycopeptide. It simultaneously searches a library of glycan structures from GlycomeDB and a library of protein sequences specified by the user. Besides searching for N-linked glycopeptide diagnostic peaks, GlycoMaster DB incorporates ion ladder checks to confirm the glycopeptide spectrum. If ETD data are provided, separate lists of potential glycans and peptides are computed independently. Glycopeptides are finally identified by scoring all pairs of glycan and peptide.

Methods relying on glycan databases are limited by the lack of comprehensive collections of glycan structures and are insufficient to discover new glycan structures that have not yet been reported. Another limitation of performing database searches for glycoprotein identification is the need-to-know potential protein sequences.

### De novo sequencing

To discover glycan structures that are not registered in any existing glycan databases, while avoiding too many implausible structures being assigned, approaches independent of database search have been invented to facilitate automatic interpretation of MS data for released glycans and glycoproteins (Table 4).

STAT is a pioneer software that uses MS/MS data for oligosaccharide composition analysis [245]. It determines sequence information for glycans of up to ten monosaccharide residues. After the

user specifies the monosaccharide component, charge carrier, precursor and product ion mass, STAT generates all possible topologies for potential compositions. A rating system is used to list the most likely structures based on number of bond cleavages. The limitation of this method is that the user needs to input information at each step and the determination of large N-glycans is not supported. N-glycans containing bisecting GlcNAc are also beyond its analytical scope.

GLYCH (GlycanCharacterization) is an algorithm for automated MS/MS glycan identification and characterization [246]. It utilizes cross-ring fragmentation patterns to derive linkage information between monosaccharide residues and distinguishes isomeric oligosaccharides from mass spectra. Low-intensity peaks are removed before processing the spectrum in GLYCH, which might lead to neglect of the right peaks for glycan structure.

Glyco-Peakfinder is a web application that determines glycan compositions by *de novo* annotation of glycan MS spectra [247]. It enables users to specify possible cross-ring fragments, potential modifications and charge state. Similar to GlycoFragment, theoretically possible fragments of a carbohydrate structure are enumerated. The difference is that Glyco-Peakfinder does not depend on glycan databases with theoretically computed masses or prior experimentally verified structures. All possible compositions are computed based on user-defined constraints. Multiple annotations for a single peak would need the user to judge the most probable structure because no scoring function is applied. Automatic peak detection is not covered so the user has to specify the monoisotopic  $m/z$  values. *De novo* sequencing tools could be used to find new structures not reported previously but expert knowledge is required in picking a proper solution to match experimental data.

### Glycosylation pathway analysis

With the growing number of databases accumulating experimental data of glycan structures, there is a need to link these structures to our existing knowledge of genes and proteins to understand complex cellular processes. To associate chemical attributes of glycans with genomic information, several groups have developed glycan-related pathway-analysis tools. This section will describe these resources.

### GlycoVis

The main biosynthetic pathways of N-linked glycans have been well studied in mammalian cells. A relatively small number of

TABLE 4

#### Methods for processing mass spectrometry (MS)-based glycan analytical data based on *de novo* sequencing.

Name	Input MS stage	Principles	Limitations
STAT	MS <sup>n</sup>	Enumerate all possible topologies for potential compositions calculated after monosaccharide component, charge carrier and precursor $m/z$ specified by the user	Confined to glycans with no more than ten monosaccharide residues; bisecting N-glycans are not supported
GLYCH	MS <sup>2</sup>	Characterize glycan structures by cross-ring cleavage ions	The scoring function is biased toward linear structure rather than branching structure. Double cleavage ions are not considered in the algorithm
GlycoPeakfinder	MS <sup>n</sup>	Generate all possible compositions for a user-specified peak list by considering all types of fragmentation	No scoring function is applied. User needs to determine the most probable structure when single peak gets multiple annotations



enzymes and nucleotide sugars are involved but single enzymes sometimes utilize multiple substrates so that the biosynthetic pathway is more like a network of 800 or so proteins [248]. Hossler *et al.* created a program, called GlycoVis, to visualize the relevant reaction pathways leading to a distinct N-linked glycan structure. Distribution of glycan structures on reaction pathways is also displayed in different colors [249].

### **Glyco-Net**

Glyco-Net is a dynamic functional presentation of relationships among various biological molecules and biological events. It is part of Glycoconjugate Data Bank [250]. Besides glycans, biological molecules like carbohydrate-binding proteins, genes, inhibitors, lipids and glycosphingolipids are all included. Diseases as well as biosynthesis are all considered into biological events. External linkages are also enabled to obtain detailed information of biological molecules. Glyco-Net serves as an interface between different biological databases and dynamically outputs the functional network among biological molecules.

### **KEGG GLYCAN**

KEGG GLYCAN is a nonredundant glycan database that stores carbohydrate structures that were originally derived from CarbBank [251]. It is a comprehensive glycome informatics resource integrating knowledge-base of protein networks, genomic information and chemical information [252]. This resource includes: (i) GLYCAN, a database of glycan structures; (ii) glycan-related pathways; and (iii) composite structure map (CSM), a tool that illustrates all possible variations of carbohydrate structures.

KEGG GLYCAN provides a structural blast tool called KCaM. It aligns two glycan tree structures and finds the maximum common branches. It cleans redundant entries obtained from CarbBank and merges to single entries. Another useful tool in this resource is KegDraw, it enables querying glycan structures by drawings.

Carbohydrates are synthesized by different types of glycosyltransferases. Each enzyme recognizes certain substrates and catalyzes specific biosynthetic reactions. Once the repertoire of glycosyltransferases is known, the repertoire of glycan structures can be predicted theoretically (and vice versa). Based on this, CSM converts genomic glycosyltransferase information into variations of glycan structures [253].

### **GlycoVault**

GlycoVault is a bioinformatics infrastructure that facilitates visualization, analysis and modeling of glycan pathways [254]. It supports storage of a variety of data formats using relational databases and ontologies. Two glycomics analysis tools are included: GlycoBrowser and GlyMpse.

GlycoBrowser is a biological pathway visualization tool that employs graphical representation of glycan structures in the biosynthetic pathways. Like Glycan Builder [255], GlycoBrowser allows users to build graphic glycan molecules. The difference lies in the restrictions it enforces using ontology with graphical editing. If experimental data or relevant information are present, GlycoBrowser will overlay experimental data with the displayed pathway.

GlyMpse is an ontology-driven simulation tool that models biochemical pathways. Biochemical entities such as glycans and enzymes are represented as nodes and enzymatic reactions as edges. Simulation datasets comprise metabolic pathways and enzyme kinetics datasets. Pathway simulation is based on firing delays and enzyme concentrations.

### **Glycan Pathway Predictor (RINGS)**

As a part of the resource for informatics of glycomes at Soka (RINGS), Glycan Pathway Predictor (GPP) is a web-based tool that models specified glycans with substrate specificity rules [256]. It implements the mathematical model of N-glycosylation developed by Krambeck *et al.* [146,147]. GPP can be used to generate potential glycans catalyzed by a select set of glyco-enzymes together with detailed pathways and corresponding enzymes. Taking any glycan structure as an entry point, GPP will retrieve a possible glycan synthesis pathway map with a selected glyco-enzyme set.

GPP utilizes KEGG Chemical Function (KCF) format as the input for glycan structures but enables multiple output formats by using LinerCode<sup>®</sup>. Free from the requirements of detailed concentration information of substrates, GPP does not support enzyme kinetic simulations. If the enzyme set is not well picked or the input format is wrong, GPP might fail to return any result.

### **Experimental implementation of glycosylation pathway analysis**

Glycosylation pathway analysis is based on information of glycan structures, glycosyltransferases, glycosidases and other glycan-binding molecules. New experimental technologies accelerate the accumulation of such fundamental knowledge. A medium-throughput quantitative, real-time reverse-transcriptase PCR (qRT-PCR) platform now allows the verification of glycomic pathway models that associate changes in glycan abundance and gene expression level for corresponding biosynthesis enzymes [248]. Recent development of high-throughput HPLC/UPLC technologies facilitates straightforward glycan analysis and enriches the knowledge-base of glycan structures [210]. Advances in genotyping technologies now permit identifying and validating millions of new single nucleotide polymorphisms (SNPs), which further enables hypothesis-free genome-wide association studies (GWAS) [257]. A GWAS study that combines high-throughput glycomics analysis discovered that HNF1 and its downstream target HNF4 play an important part in key fucosyltransferase and fucose biosynthetic gene regulation [213]. A more recent GWAS study identified five genes that have no prior report on their participation in protein glycosylation to be strongly associated with IgG N-glycosylation and autoimmune diseases [14]. With more pieces of information added into our knowledge-base, glycomic pathway analysis would be more informative and would give better guidance in new discoveries and provide better ways of controlling the cellular pathways in bioprocessing.

### **Concluding remarks**

The biopharmaceutical industry has experienced tremendous growth during the past two decades [12]. mAbs provide a major driving force in this development. The industry will be further propelled by the rise of biosimilars [10]. The cost of goods for therapeutic mAbs has dropped significantly over the years because of advancements in the cell culture processes [91]. In addition, guidelines for regulating the biosimilars have placed an emphasis on glycosylation analysis and comparability [79,83]. Therefore, one major challenge in successfully developing therapeutic mAbs lies in controlling glycosylation. Glycosylation will play an even more important part in the light of biobetter drugs [258]. Producing optimal and consistent glycosylation requires an integrated

approach. At present, efforts toward this aim are mainly concentrated on empirical adjustments of upstream cell culture parameters. With the accumulation of knowledge and development of new tools, we see opportunities in the following areas: (i) custom-making glycoforms by editing the glycosylation machinery of the host cells; (ii) development of online and atline analytical technologies to monitor and adjust the process continuously; (iii) the effective application of QbD and PATs in real-time monitoring and controlling of glycosylation; (iv) glycoform-specific downstream purification strategies; (v) integrated informatics platforms for data processing and model-driven glycosylation predictions; and (vi) systems biology of molecular and cellular events in the cell culture. An integrated approach combining the

capabilities of the above areas will truly enable the industry to harness glycosylation for successful biopharmaceutical drug development.

### Acknowledgments

This work is supported by Biomedical Research Council and Joint Council Office, Agency for Science, Technology and Research (A\*STAR), Singapore. The authors would like to thank Corrine Wan and Ce Huang Poo for the technical assistance. Regulatory views in this article are presented by the authors for discussion and considerations. They do not reflect the policy of China Food and Drug Administration (CFDA) or official views of the Chinese Government.

### References

- Apweiler, R. *et al.* (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8
- Parodi, A.J. (2000) Protein glycosylation and its role in protein folding. *Annu. Rev. Biochem.* 69, 69–93
- Moremen, K.W. *et al.* (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 13, 448–462
- Wearsch, P.A. *et al.* (2011) Essential glycan-dependent interactions optimize MHC class I peptide loading. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4950–4955
- Stencel-Baerenwald, J.E. *et al.* (2014) The sweet spot: defining virus-sialic acid interactions. *Nat. Rev. Microbiol.* 12, 739–749
- Hayes, J.M. *et al.* (2014) Glycosylation and Fc receptors. *Curr. Top. Microbiol. Immunol.* 382, 165–199
- Chung, C.H. *et al.* (2008) Cetuximab-induced anaphylaxis and IgE specific for galactose- $\alpha$ -1,3-galactose. *N. Engl. J. Med.* 358, 1109–1117
- Sinclair, A.M. (2013) Erythropoiesis stimulating agents: approaches to modulate activity. *Biologics* 7, 161–174
- Shinkawa, T. *et al.* (2003) 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.* 278, 3466–3473
- Walsh, G. (2014) Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* 32, 992–1000
- Ecker, D.M. *et al.* (2015) The therapeutic monoclonal antibody market. *MAbs* 7, 9–14
- Aggarwal, R.S. (2014) What's fueling the biotech engine—2012 to 2013. *Nat. Biotechnol.* 32, 32–39
- Arnold, J.N. *et al.* (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* 25, 21–50
- Lauc, G. *et al.* (2013) Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.* 9, e1003225
- Saldova, R. *et al.* (2011) 5-AZA-2'-deoxycytidine induced demethylation influences N-glycosylation of secreted glycoproteins in ovarian cancer. *Epigenetics* 6, 1362–1372
- Hossler, P. *et al.* (2009) Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19, 936–949
- Daeron, M. (2014) Fc receptors as adaptive immunoreceptors. *Curr. Top. Microbiol. Immunol.* 382, 131–164
- Milstein, C.P. and Deverson, E.V. (1974) Primary structure of  $\kappa$  light chain of a human myeloma protein. *Eur. J. Biochem.* 49, 377–391
- Arakawa, F. *et al.* (1993) Construction and expression of two mouse–human chimeric antibodies with high specificity and affinity for carcinoembryonic antigen. *Hybridoma* 12, 365–379
- Kusakabe, K. *et al.* (1994) Accumulation enhancement of human monoclonal antibody HB4C5 to lung tumor xenografts by N-deglycosylation. *J. Nucl. Med.* 35, 289–295
- Wallick, S.C. *et al.* (1988) Glycosylation of a VH residue of a monoclonal antibody against alpha (1–6) dextran increases its affinity for antigen. *J. Exp. Med.* 168, 1099–1109
- van Bueren, J.J.L. *et al.* (2011) Anti-galactose- $\alpha$ -1,3-galactose IgE from allergic patients does not bind  $\alpha$ -galactosylated glycans on intact therapeutic antibody Fc domains. *Nat. Biotechnol.* 29, 574–576
- Commins, S.P. and Platts-Mills, T.A.E. (2010) Allergenicity of carbohydrates and their role in anaphylactic events. *Curr. Allergy Asthma Rep.* 10, 29–33
- Rudd, P.M. (2001) Glycosylation and the immune system. *Science* 291, 2370–2376
- Yang, W.H. *et al.* (2015) An intrinsic mechanism of secreted protein aging and turnover. *Proc. Natl. Acad. Sci. U. S. A.* 112, 13657–13662
- Dong, X. *et al.* (1999) Binding and uptake of agalactosyl IgG by mannose receptor on macrophages and dendritic cells. *J. Immunol.* 163, 5427–5434
- Malhotra, R. *et al.* (1995) Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat. Med.* 1, 237–243
- Goetze, A.M. *et al.* (2011) High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* 21, 949–959
- Ishino, T. *et al.* (2013) Engineering a monomeric Fc domain modality by N-glycosylation for the half-life extension of biotherapeutics. *J. Biol. Chem.* 288, 16529–16537
- Kibe, T. *et al.* (1996) Glycosylation and placental transport of immunoglobulin G. *J. Clin. Biochem. Nutr.* 21, 57–63
- Krapp, S. *et al.* (2003) Structural analysis of human IgG-Fc glycoforms reveals correlation between glycosylation and structural integrity. *J. Mol. Biol.* 325, 979–989
- Mimura, Y. *et al.* (2001) Role of oligosaccharide residues of IgG1-Fc in FcRIIb binding. *J. Biol. Chem.* 276, 45539–45547
- Ferrara, C. *et al.* (2011) Unique carbohydrate–carbohydrate interactions are required for high affinity binding between Fc(R)III and antibodies lacking core fucose. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12669–12674
- van Schravendijk, M.R. and Dwek, R.A. (1981) The use of protein A and concanavalin A to examine the possible role of the carbohydrate of IgG in the binding of Clq. *Mol. Immunol.* 18, 1079–1085
- Burton, D.R. *et al.* (1980) The Clq receptor site on immunoglobulin G. *Nature* 288, 338–344
- Mimura, Y. *et al.* (2000) The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol. Immunol.* 37, 697–706
- Shields, R.L. *et al.* (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc(R)III and antibody-dependent cellular toxicity. *J. Biol. Chem.* 277, 26733–26740
- Pincetic, A. *et al.* (2014) Type I and type II Fc receptors regulate innate and adaptive immunity. *Nat. Immunol.* 15, 707–716
- Padlan, E.A. (1994) Anatomy of the antibody molecule. *Mol. Immunol.* 31, 169–217
- Washburn, N. *et al.* (2015) Controlled tetra-Fc sialylation of IVIg results in a drug candidate with consistent enhanced anti-inflammatory activity. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1297–E1306
- Yu, C. *et al.* (2016) At least two Fc Neu5Gc residues of monoclonal antibodies are required for binding to anti-Neu5Gc antibody. *Sci. Rep.* 7 <http://dx.doi.org/10.1038/srep20029> (article no. 20029)
- Yang, J.-M. *et al.* (2014) Investigation of the correlation between charge and glycosylation of IgG1 variants by liquid chromatography–mass spectrometry. *Anal. Biochem.* 448, 82–91
- Kaneko, Y. *et al.* (2006) Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* 313, 670–673
- Samuelsson, A. *et al.* (2001) Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. *Science* 291, 484–486

- 45 Campbell, I.K. *et al.* (2014) Therapeutic effect of IVIG on inflammatory arthritis in mice is dependent on the Fc portion and independent of sialylation or basophils. *J. Immunol.* 192, 5031–5038
- 46 Issekutz, A.C. *et al.* (2015) Intravenous IgG (IVIG) and subcutaneous IgG (SCIG) preparations have comparable inhibitory effect on T cell activation, which is not dependent on IgG sialylation, monocytes or B cells. *Clin. Immunol.* 160, 123–132
- 47 Nagelkerke, S.Q. *et al.* (2014) Inhibition of Fc(R-mediated phagocytosis by IVIG is independent of IgG-Fc sialylation and Fc(RIIb) in human macrophages. *Blood* 124, 3709–3718
- 48 Anthony, R.M. *et al.* (2011) Intravenous gammaglobulin suppresses inflammation through a novel T(H)2 pathway. *Nature* 475, 110–113
- 49 Bayry, J. *et al.* (2012) Intravenous immunoglobulin expands regulatory T cells in autoimmune rheumatic disease. *J. Rheumatol.* 39, 450–451
- 50 Kessel, A. *et al.* (2007) Intravenous immunoglobulin therapy affects T regulatory cells by increasing their suppressive function. *J. Immunol.* 179, 5571–5575
- 51 Ramakrishna, C. and Cantin, E.M. (2014) Fc-sialylated IgGs in intravenous immunoglobulins are not responsible for induction of regulatory T cells. *J. Allergy Clin. Immunol.* 134, 1469
- 52 Zhu, D. *et al.* (2002) Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood* 99, 2562–2568
- 53 Scherer, H.U. *et al.* (2009) Immunoglobulin 1 (IgG1) Fc-glycosylation profiling of anti-citrullinated peptide antibodies from human serum. *Proteomics Clin. Appl.* 3, 106–115
- 54 Kabat, E.A. *et al.* eds (1992) *Sequences of Proteins of Immunological Interest*, DIANE Publishing
- 55 Wright, A. *et al.* (1991) Antibody variable region glycosylation: position effects on antigen binding and carbohydrate structure. *EMBO J.* 10, 2717
- 56 Jacquemin, M. (2009) Variable region heavy chain glycosylation determines the anticoagulant activity of a factor VIII antibody. *Haemophilia* 16, 16–19
- 57 Song, R. *et al.* (2013) Strategic addition of an N-linked glycan to a monoclonal antibody improves its HIV-1-neutralizing activity. *Nat. Biotechnol.* 31, 1047–1052
- 58 Sabouri, Z. *et al.* (2014) Redemption of autoantibodies on anergic B cells by variable-region glycosylation and mutation away from self-reactivity. *Proc. Natl. Acad. Sci. U. S. A.* 111, E2567–E2575
- 59 Kinoshita, N. *et al.* (1991) Glycosylation at the Fab portion of myeloma immunoglobulin G and increased fucosylated biantennary sugar chains: structural analysis by high-performance liquid chromatography and antibody-lectin enzyme immunoassay using *Lens culinaris* agglutinin. *Cancer Res.* 51, 5888–5892
- 60 Youings, A. *et al.* (1996) Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. *Biochem. J.* 314, 621–630
- 61 Radcliffe, C.M. *et al.* (2007) Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J. Biol. Chem.* 282, 7405–7415
- 62 Coelho, V. *et al.* (2010) Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins. *Proc. Natl. Acad. Sci. U. S. A.* 107, 18587–18592
- 63 Tangvoranuntakul, P. *et al.* (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12045–12050
- 64 Ghaderi, D. *et al.* (2010) Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat. Biotechnol.* 28, 863–867
- 65 Noguchi, A. *et al.* (1995) Immunogenicity of N-glycolylneuraminic acid-containing carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells. *J. Biochem.* 117, 59–62
- 66 Bosques, C.J. *et al.* (2010) Chinese hamster ovary cells can produce galactose- $\alpha$ -1,3-galactose antigens on proteins. *Nat. Biotechnol.* 28, 1153–1156
- 67 Hanganutziu, M. (1924) Hémagglutinines hétérogénéiques après injection de sérum de cheval. *CR Séances Soc. Biol.* 91, 1457–1459 (in French)
- 68 Deicher, D.H. (1926) Über die erzeugung heterospezifischer hämagglutinine durch injektion artfremden serums. *Zeitschrift Hygiene Infektionskrankheiten* 106, 561–579 (in German)
- 69 Pirofsky, B. *et al.* (1973) Foreign serum heterophile antibodies in patients receiving antithymocyte antisera. *Blood* 42, 385–393
- 70 Higashi, H. *et al.* (1977) Antigen of serum sickness type of heterophile antibodies in human sera: identification as gangliosides with N-glycolylneuraminic acid. *Biochem. Biophys. Res. Commun.* 79, 388–395
- 71 Anthony, R.M. *et al.* (2008) Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science* 320, 373–376
- 72 Stadlmann, J. *et al.* (2009) A close look at human IgG sialylation and subclass distribution after lectin fractionation. *Proteomics* 9, 4143–4153
- 73 Stadlmann, J. *et al.* (2010) Analytical and functional aspects of antibody sialylation. *J. Clin. Immunol.* 30, 15–19
- 74 O'Neil, B.H. *et al.* (2007) High incidence of cetuximab-related infusion reactions in Tennessee and North Carolina and the association with atopic history. *J. Clin. Oncol.* 25, 3644–3648
- 75 Casadevall, N. *et al.* (2002) Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. *N. Engl. J. Med.* 346, 469–475
- 76 Commins, S.P. *et al.* (2009) Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose- $\alpha$ -1,3-galactose. *J. Allergy Clin. Immunol.* 123, 426–433 e422
- 77 Commins, S.P. *et al.* (2011) The relevance of tick bites to the production of IgE antibodies to the mammalian oligosaccharide galactose- $\alpha$ -1,3-galactose. *J. Allergy Clin. Immunol.* 127, 1286–1293 e1286
- 78 World Health Organization (2009) *Guidelines on Evaluation of Similar Biotechnological Products (SBPs)*. Available at: <http://apps.who.int/medicinedocs/documents/s19941en/s19941en.pdf>
- 79 European Medicines Agency (2013) *Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Active Substance: Non-clinical and Clinical Issues*. Available at: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003920.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003920.pdf)
- 80 Read, E.K. *et al.* (2011) Industry and regulatory experience of the glycosylation of monoclonal antibodies. *Biotechnol. Appl. Biochem.* 58, 213–219
- 81 Schiestl, M. *et al.* (2011) Acceptable changes in quality attributes of glycosylated biopharmaceuticals. *Nat. Biotechnol.* 29, 310–312
- 82 Schiestl, M. (2011) A biosimilar industry view on the implementation of the WHO guidelines on evaluating similar biotechnological products. *Biologics* 39, 297–299
- 83 U.S. Department of Health and Human Services (2012) *Guidance for Industry: Quality Considerations in Demonstrating Biosimilarity to a Reference Protein Product*. Available at: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm291134.pdf>
- 84 Van Berkel, P.H.C. *et al.* (2009) N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG. *Biotechnol. Prog.* 25, 244–251
- 85 Chusainow, J. *et al.* (2009) A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? *Biotechnol. Bioeng.* 102, 1182–1196
- 86 Rouiller, Y. *et al.* (2014) Modulation of mAb quality attributes using microliter scale fed-batch cultures. *Biotechnol. Prog.* 30, 571–583
- 87 Bareither, R. *et al.* (2013) Automated disposable small scale reactor for high throughput bioprocess development: a proof of concept study. *Biotechnol. Bioeng.* 110, 3126–3138
- 88 Barrett, T.A. *et al.* (2010) Microwell engineering characterization for mammalian cell culture process development. *Biotechnol. Bioeng.* 105, 260–275
- 89 Eon-Duval, A. *et al.* (2012) Quality attributes of recombinant therapeutic proteins: an assessment of impact on safety and efficacy as part of a quality by design development approach. *Biotechnol. Prog.* 28, 608–622
- 90 Gramer, M.J. (2014) Product quality considerations for mammalian cell culture process development and manufacturing. In *Mammalian Cell Cultures for Biologics Manufacturing*, (Vol. 139) (Zhou, W. and Kantardjiev, A., eds) pp. 123–166, Springer-Verlag, Berlin
- 91 Kelley, B. (2009) Industrialization of mAb production technology: the bioprocessing industry at a crossroads. *MAbs* 1, 443–452
- 92 Ho, S.C.L. *et al.* (2013) Generation of monoclonal antibody-producing mammalian cell lines. *Pharm. Bioprocess.* 1, 71–87
- 93 Wadhwa, M. and Thorpe, R. (2013) European perspective on biosimilars. *Bioanalysis* 5, 521–524
- 94 Chow, S-C. *et al.* (2013) Comments on the FDA draft guidance on biosimilar products. *Statistics Med.* 32, 364–369
- 95 Hossler, P. (2012) Protein glycosylation control in mammalian cell culture: past precedents and contemporary prospects. *Adv. Biochem. Eng. Biotechnol.* 127, 187–219
- 96 Davies, S.L. *et al.* (2013) Functional heterogeneity and heritability in CHO cell populations. *Biotechnol. Bioeng.* 110, 260–274
- 97 Ho, S.C.L. *et al.* (2012) IRES-mediated Tricistronic vectors for enhancing generation of high monoclonal antibody expressing CHO cell lines. *J. Biotechnol.* 157, 130–139
- 98 Ho, S.C.L. *et al.* (2013) Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability. *J. Biotechnol.* 165, 157–166
- 99 Li, F. *et al.* (2010) Cell culture processes for monoclonal antibody production. *MAbs* 2, 466–479
- 100 Pastores, G.M. (2010) Recombinant glucocerebrosidase (imiglucerase) as a therapy for Gaucher disease. *BioDrugs* 24, 41–47
- 101 Thomann, M. *et al.* (2015) In vitro glycoengineering of IgG1 and its effect on Fc receptor binding and ADCC activity. *PLoS One* 10, e0134949



- 102 Mimura, Y. *et al.* (2015) Enhanced sialylation of a human chimeric IgG1 variant produced in human and rodent cell lines. *J. Immunol. Methods* <http://dx.doi.org/10.1016/j.jim.2015.11.009>
- 103 Patnaik, S.K. and Stanley, P. (2006) Lectin-resistant CHO glycosylation mutants. *Methods Enzymol.* 416, 159–182
- 104 Zhang, P. *et al.* (2013) CHO glycosylation mutants as potential host cells to produce therapeutic proteins with enhanced efficacy. *Adv. Biochem. Eng. Biotechnol.* 131, 63–87
- 105 Zhang, P. *et al.* (2012) Identification of functional elements of the GDP-fucose transporter SLC35C1 using a novel Chinese hamster ovary mutant. *Glycobiology* 22, 897–911
- 106 Steentoft, C. *et al.* (2014) Precision genome editing: a small revolution for glycobiology. *Glycobiology* 24, 663–680
- 107 Wang, W.C. and Cummings, R.D. (1988) The immobilized leucoagglutinin from the seeds of *Maackia amurensis* binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked alpha-2,3 to penultimate galactose residues. *J. Biol. Chem.* 263, 4576–4585
- 108 Knibbs, R.N. *et al.* (1991) Characterization of the carbohydrate binding specificity of the leucoagglutinating lectin from *Maackia amurensis*. Comparison with other sialic acid-specific lectins. *J. Biol. Chem.* 266, 83–88
- 109 Ghaderi, D. *et al.* (2012) Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnol. Genet. Eng. Rev.* 28, 147–175
- 110 Lim, S.F. *et al.* (2008) The Golgi CMP-sialic acid transporter: a new CHO mutant provides functional insights. *Glycobiology* 18, 851–860
- 111 Goh, J.S. *et al.* (2010) RCA-I-resistant CHO mutant cells have dysfunctional GnT I and expression of normal GnT I in these mutants enhances sialylation of recombinant erythropoietin. *Metab. Eng.* 12, 360–368
- 112 Goh, J.S. *et al.* (2014) Highly sialylated recombinant human erythropoietin production in large-scale perfusion bioreactor utilizing CHO-gmt4 (JW152) with restored GnT I function. *Biotechnol. J.* 9, 100–109
- 113 Segal, D.J. *et al.* (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2758–2763
- 114 Dreier, B. *et al.* (2000) Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains. *J. Mol. Biol.* 303, 489–502
- 115 Liu, Q. *et al.* (2002) Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.* 277, 3850–3856
- 116 Sander, J.D. *et al.* (2007) Zinc finger targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res.* 35, W599–W605
- 117 Cermak, T. *et al.* (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82
- 118 Miller, J.C. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148
- 119 Boch, J. *et al.* (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512
- 120 Moscou, M.J. and Bogdanove, A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501
- 121 Sanjana, N.E. *et al.* (2012) A transcription activator-like effector toolbox for genome engineering. *Nat. Protocols* 7, 171–192
- 122 Reyon, D. *et al.* (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* 30, 460–465
- 123 Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
- 124 Mali, P. *et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826
- 125 Grav, L.M. *et al.* (2015) One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* 10, 1446–1456
- 126 Montague, T.G. *et al.* (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 42, W401–W407
- 127 Wu, X. *et al.* (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676
- 128 Kescu, C. *et al.* (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683
- 129 Jedrzejewski, P.M.J. *et al.* (2013) Applying quality by design to glycoprotein therapeutics: experimental and computational efforts of process control. *Pharm. Bioprocess.* 1, 51–69
- 130 Rathore, A.S. (2009) Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends Biotechnol.* 27, 546–553
- 131 del Val, I.J. *et al.* (2010) Towards the implementation of quality by design to the production of therapeutic monoclonal antibodies with desired glycosylation patterns. *Biotechnol. Prog.* 26, 1505–1527
- 132 Hodoniczky, J. *et al.* (2005) Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling *in vitro*. *Biotechnol. Prog.* 21, 1644–1652
- 133 Shields, R.L. *et al.* (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human FcγRIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* 277, 26733–26740
- 134 Fernandes, D., ed. (2011) *High-Throughput Multi-Product Liquid Chromatography for Characterization of Monoclonal Antibodies*, BioPharm International
- 135 Prozyme® GlykoPrep™ Guidebook. Available at: <http://www.prozyme.com/documents/TNGP100.pdf>.
- 136 Doherty, M. *et al.* (2013) An automated robotic platform for rapid profiling oligosaccharide analysis of monoclonal antibodies directly from cell culture. *Anal. Biochem.* 442, 10–18
- 137 Nishimura, S. *et al.* (2004) High-throughput protein glycomics: combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem. Int. Ed. Engl.* 44, 91–96
- 138 Zhang, H. *et al.* (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* 21, 660–666
- 139 Stockmann, H. *et al.* (2013) Automated, high-throughput IgG-antibody glycoproteomic platform. *Anal. Chem.* 85, 8841–8849
- 140 Reusch, D. *et al.* (2013) High-throughput work flow for IgG Fc-glycosylation analysis of biotechnological samples. *Anal. Biochem.* 432, 82–89
- 141 Kumar, V. *et al.* (2014) Design of experiments applications in bioprocessing: concepts and approach. *Biotechnol. Prog.* 30, 86–99
- 142 Alvarez, M. *et al.* (2011) On-line characterization of monoclonal antibody variants by liquid chromatography–mass spectrometry operating in a two-dimensional format. *Anal. Biochem.* 419, 17–25
- 143 Primack, J. *et al.* (2011) A high-throughput microchip-based glycan screening assay for antibody cell culture samples. *Electrophoresis* 32, 1129–1132
- 144 Bynum, M.A. *et al.* (2009) Characterization of IgG N-glycans employing a microfluidic chip that integrates glycan cleavage, sample purification, LC separation, and MS detection. *Anal. Chem.* 81, 8818–8825
- 145 Xu, X. *et al.* (2011) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 29, 735–741
- 146 Krambeck, F.J. *et al.* (2009) A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology* 19, 1163–1175
- 147 Krambeck, F.J. and Betenbaugh, M.J. (2005) A mathematical model of N-linked glycosylation. *Biotechnol. Bioeng.* 92, 711–728
- 148 Umama, P. and Bailey, J.E. (1997) A mathematical model of N-linked glycoform biosynthesis. *Biotechnol. Bioeng.* 55, 890–908
- 149 Hossler, P. *et al.* (2007) Systems analysis of N-glycan processing in mammalian cells. *PLoS One* 2, e713
- 150 Jedrzejewski, P.M. *et al.* (2014) Towards controlling the glycoform: a model framework linking extracellular metabolites to antibody glycosylation. *Int. J. Mol. Sci.* 15, 4492–4522
- 151 Jimenez del Val, I. *et al.* (2011) A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnol. Prog.* 27, 1730–1743
- 152 Carta, G. and Jungbauer, A., eds (2010) *Protein Chromatography: Process Development and Scale-up*, John Wiley & Sons
- 153 El Khoury, G. *et al.* (2015) Bespoke affinity ligands for the purification of therapeutic proteins. *Pharm. Bioprocess* 3, 139–152
- 154 Liu, H.F. *et al.* (2010) Recovery and purification process development for monoclonal antibody production. *MAbs* 2, 480–499
- 155 Gagnon, P. *et al.* (2014) High productivity purification of immunoglobulin G monoclonal antibodies on starch-coated magnetic nanoparticles by steric exclusion of polyethylene glycol. *J. Chromatogr. A* 1324, 171–180
- 156 Gagnon, P. *et al.* (2014) Chromatin-mediated depression of fractionation performance on electronegative multimodal chromatography media, its prevention, and ramifications for purification of immunoglobulin G. *J. Chromatogr. A* 1374, 145–155
- 157 Gagnon, P. (2012) Technology trends in antibody purification. *J. Chromatogr. A* 1221, 57–70
- 158 Fanayan, S. *et al.* (2012) Using lectins to harvest the plasma/serum glycoproteome. *Electrophoresis* 33, 1746–1754
- 159 Weatherman, R.V. *et al.* (1996) Specificity of C-glycoside complexation by mannose/glucose specific lectins. *Biochemistry* 35, 3619–3624
- 160 Pusztai, A. *et al.* (1993) Antinutritive effects of wheat-germ agglutinin and other N-acetylglucosamine-specific lectins. *Br. J. Nutr.* 70, 313–321
- 161 Bourne, Y. *et al.* (2002) Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose. *Biochemical J* 364, 173–180



- 162 Opitz, L. *et al.* (2008) Capture of cell culture-derived influenza virus by lectins: strain independent, but host cell dependent. *J. Virol. Methods* 154, 61–68
- 163 Opitz, L. *et al.* (2007) Lectin-affinity chromatography for downstream processing of MDCK cell culture derived human influenza A viruses. *Vaccine* 25, 939–947
- 164 dos Santos, R. *et al.* (2014) Phenylboronic acid as a multi-modal ligand for the capture of monoclonal antibodies: development and optimization of a washing step. *J. Chromatogr. A* 1355, 115–124
- 165 Zanette, D. *et al.* (2003) Evaluation of phenylboronate agarose for industrial-scale purification of erythropoietin from mammalian cell cultures. *J. Biotechnol.* 101, 275–287
- 166 Lin, Z.-A. *et al.* (2011) Synthesis of magnetic nanoparticles with immobilized aminophenylboronic acid for selective capture of glycoproteins. *J. Mater. Chem.* 21, 518–524
- 167 Yeap, W.S. *et al.* (2008) Using detonation nanodiamond for the specific capture of glycoproteins. *Anal. Chem.* 80, 4659–4665
- 168 Zhou, W. *et al.* (2008) Facile synthesis of aminophenylboronic acid-functionalized magnetic nanoparticles for selective separation of glycopeptides and glycoproteins. *Chem. Commun.* 43, 5577–5579
- 169 Ma, R. *et al.* (2014) Facile synthesis of boronic acid-functionalized magnetic carbon nanotubes for highly specific enrichment of glycopeptides. *Nanoscale* 6, 3150–3156
- 170 Li, H. *et al.* (2011) A Wulff-type boronate for boronate affinity capture of cis-diol compounds at medium acidic pH condition. *Chem. Commun.* 47, 8169–8171
- 171 Zhang, Y. *et al.* (2014) Benzoboroxole-functionalized magnetic core/shell microspheres for highly specific enrichment of glycoproteins under physiological conditions. *Small* 10, 1379–1386
- 172 Schumacher, S. *et al.* (2011) Label-free detection of enhanced saccharide binding at pH 7.4 to nanoparticulate benzoboroxole based receptor units. *J. Mol. Recog.* 24, 953–959
- 173 Gamblin, D.P. *et al.* (2009) Glycoprotein synthesis: an update. *Chem. Rev.* 109, 131–163
- 174 Roque, A.C. *et al.* (2004) Antibodies and genetically engineered related molecules: production and purification. *Biotechnol. Prog.* 20, 639–654
- 175 Chen, C. *et al.* (2014) Affinity ligands for glycoprotein purification based on the multi-component Ugi reaction. *J. Chromatogr. B* 969, 171–180
- 176 Marino, K. *et al.* (2010) A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* 6, 713–723
- 177 Garcia, B.A. (2010) What does the future hold for top down mass spectrometry? *J. Am. Soc. Mass Spectrom.* 21, 193–202
- 178 Valeja, S.G. *et al.* (2011) Unit mass baseline resolution for an intact 148 kDa therapeutic monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 83, 8391–8395
- 179 Zhang, Z. *et al.* (2009) Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* 28, 147–176
- 180 Breuker, K. *et al.* (2008) Top-down identification and characterization of biomolecules by mass spectrometry. *J. Am. Soc. Mass Spectrom.* 19, 1045–1053
- 181 Beck, A. *et al.* (2005) Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NS0 cells. *J. Chromatogr. B* 819, 203–218
- 182 Xie, H. *et al.* (2010) Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. *MAbs* 2, 379–394
- 183 Zhang, Z. and Shah, B. (2007) Characterization of variable regions of monoclonal antibodies by top-down mass spectrometry. *Anal. Chem.* 79, 5723–5729
- 184 Bondarenko, P.V. *et al.* (2009) Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-Orbitrap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 20, 1415–1424
- 185 Beck, A. *et al.* (2013) Characterization of therapeutic antibodies and related products. *Anal. Chem.* 85, 715–736
- 186 Bakovic, M.P. *et al.* (2013) High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides. *J. Proteome Res.* 12, 821–831
- 187 Alvarez-Manilla, G. *et al.* (2006) Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J. Proteome Res.* 5, 701–708
- 188 Wada, Y. *et al.* (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* 76, 6560–6565
- 189 Hagglund, P. *et al.* (2004) A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* 3, 556–566
- 190 Takegawa, Y. *et al.* (2006) Separation of isomeric 2-aminopyridine derivatized N-glycans and N-glycopeptides of human serum immunoglobulin G by using a zwitterionic type of hydrophilic-interaction chromatography. *J. Chromatogr. A* 1113, 177–181
- 191 Wuhrer, M. *et al.* (2005) Protein glycosylation analysis by liquid chromatography–mass spectrometry. *J. Chromatogr. B* 825, 124–133
- 192 Wuhrer, M. *et al.* (2005) Protein glycosylation analyzed by normal-phase nano-liquid chromatography–mass spectrometry of glycopeptides. *Anal. Chem.* 77, 886–894
- 193 Segu, Z.M. and Mechref, Y. (2010) Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun. Mass Spectrom.* 24, 1217–1225
- 194 Wuhrer, M. *et al.* (2007) Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr. B* 849, 115–128
- 195 Hakansson, K. *et al.* (2001) Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptic to yield complementary sequence information. *Anal. Chem.* 73, 4530–4536
- 196 Harvey, D.J. *et al.* (2002) Internal residue loss: rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus. *Anal. Chem.* 74, 734–740
- 197 Wuhrer, M. *et al.* (2006) Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* 20, 1747–1754
- 198 Syka, J.E. *et al.* (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9528–9533
- 199 Adamson, J.T. and Hakansson, K. (2006) Infrared multiphoton dissociation and electron capture dissociation of high-mannose type glycopeptides. *J. Proteome Res.* 5, 493–501
- 200 Hakansson, K. *et al.* (2003) Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.* 75, 3256–3262
- 201 Harazono, A. *et al.* (2006) Site-specific N-glycosylation analysis of human plasma ceruloplasmin using liquid chromatography with electrospray ionization tandem mass spectrometry. *Anal. Biochem.* 348, 259–268
- 202 Iavarone, A.T. and Williams, E.R. (2003) Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* 125, 2319–2327
- 203 Ruhaak, L.R. *et al.* (2010) Glycan labeling strategies and their use in identification and quantification. *Anal. Bioanal. Chem.* 397, 3457–3481
- 204 Zaia, J. (2008) Mass spectrometry and the emerging field of glycomics. *Chem. Biol.* 15, 881–892
- 205 Lauber, M.A. *et al.* (2015) Rapid preparation of released N-glycans for HILIC analysis using a labeling reagent that facilitates sensitive fluorescence and ESI-MS detection. *Anal. Chem.* 87, 5401–5409
- 206 Stockmann, H. *et al.* (2015) Ultrahigh throughput, ultrafiltration-based N-glycomics platform for ultra-performance liquid chromatography (ULTRA(3)). *Anal. Chem.* 87, 8316–8322
- 207 Guillaume, D. *et al.* (2010) New trends in fast and high-resolution liquid chromatography: a critical comparison of existing approaches. *Anal. Bioanal. Chem.* 397, 1069–1082
- 208 Novakova, L. *et al.* (2006) Advantages of application of UPLC in pharmaceutical analysis. *Talanta* 68, 908–918
- 209 Ahn, J. *et al.* (2010) Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7 microm sorbent. *J. Chromatogr. B* 878, 403–408
- 210 Pucic, M. *et al.* (2011) High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol. Cell. Proteomics* 10 <http://dx.doi.org/10.1074/mcp.M111.010090>
- 211 Campbell, M.P. *et al.* (2015) GlycoBase and autoGU: resources for interpreting HPLC-glycan data. *Methods Mol. Biol.* 1273, 17–28
- 212 Duffy, F.J. and Rudd, P.M. (2015) GlycoProfileAssigner: automated structural assignment with error estimation for glycan LC data. *Bioinformatics* 31, 2220–2221
- 213 Lauc, G. *et al.* (2010) Genomics meets glycomics—the first GWAS study of human N-glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *PLoS Genet.* 6, e1001256
- 214 Mittermayr, S. *et al.* (2011) Multiplexed analytical glycomics: rapid and confident IgG N-glycan structural elucidation. *J. Proteome Res.* 10, 3820–3829
- 215 Reusch, D. *et al.* (2014) High-throughput glycosylation analysis of therapeutic immunoglobulin G by capillary gel electrophoresis using a DNA analyzer. *MAbs* 6, 185–196
- 216 Leymarie, N. and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem.* 84, 3040–3048
- 217 Harvey, D.J. *et al.* (2014) Fragmentation of negative ions from N-linked carbohydrates: part 6. Glycans containing one N-acetylglucosamine in the core. *Rapid Commun. Mass Spectrom.* 28, 2008–2018
- 218 Harvey, D.J. *et al.* (2013) Travelling wave ion mobility and negative ion fragmentation for the structural determination of N-linked glycans. *Electrophoresis* 34, 2368–2378

- 219 Yu, X. *et al.* (2013) Detailed glycan structural characterization by electronic excitation dissociation. *Anal. Chem.* 85, 10017–10021
- 220 Wolff, J.J. *et al.* (2010) Negative electron transfer dissociation of glycosaminoglycans. *Anal. Chem.* 82, 3460–3466
- 221 Oh, M.J. *et al.* (2013) Analytical platform for glycomic characterization of recombinant erythropoietin biotherapeutics and biosimilars by MS. *Bioanalysis* 5, 545–559
- 222 Hofmann, J. *et al.* (2014) Estimating collision cross sections of negatively charged N-glycans using traveling wave ion mobility-mass spectrometry. *Anal. Chem.* 86, 10789–10795
- 223 Yamaguchi, Y. *et al.* (2012) Confident identification of isomeric N-glycan structures by combined ion mobility mass spectrometry and hydrophilic interaction liquid chromatography. *Rapid Commun. Mass Spectrom.* 26, 2877–2884
- 224 Domann, P.J. *et al.* (2007) Separation-based glycoprofiling approaches using fluorescent labels. *Proteomics* 7 (Suppl. 1), 70–76
- 225 Hayes, C.A. *et al.* (2011) UniCarb-DB: a database resource for glycomic discovery. *Bioinformatics* 27, 1343–1344
- 226 Artemenko, N.V. *et al.* (2012) Databases and tools in glycobiology. *Methods Mol. Biol.* 899, 325–350
- 227 Karlsson, N.G. and Packer, N.H. (2009) Glycomic mass spectrometric analysis and data interpretation tools. In *Bioinformatics for Glycobiology and Glycomics* (von der Lieth, C.-W. *et al.* eds), pp. 223–256, John Wiley & Sons
- 228 Campbell, M.P. *et al.* (2014) Toolboxes for a standardised and systematic study of glycans. *BMC Bioinformatics* 15 (Suppl. 1), S9
- 229 Wada, Y. *et al.* (2007) Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* 17, 411–422
- 230 von der Lieth, C.W. *et al.* (2006) The role of informatics in glycobiology research with special emphasis on automatic interpretation of MS spectra. *Biochim. Biophys. Acta* 1760, 568–577
- 231 Cooper, C.A. *et al.* (2001) GlycoMod – a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 1, 340–349
- 232 Joshi, H.J. *et al.* (2004) Development of a mass fingerprinting tool for automated interpretation of oligosaccharide fragmentation data. *Proteomics* 4, 1650–1664
- 233 Lohmann, K.K. and von der Lieth, C.W. (2003) GLYCO-FRAGMENT: a web tool to support the interpretation of mass spectra of complex carbohydrates. *Proteomics* 3, 2028–2035
- 234 Lohmann, K.K. and von der Lieth, C.W. (2004) GlycoFragment and GlycoSearchMS: web tools to support the interpretation of mass spectra of complex carbohydrates. *Nucleic Acids Res.* 32, W261–W266
- 235 Go, E.P. *et al.* (2007) GlycoPep DB: a tool for glycopeptide analysis using a Smart Search. *Anal. Chem.* 79, 1708–1713
- 236 Deshpande, N. *et al.* (2010) GlycoSpectrumScan: fishing glycopeptides from MS spectra of protease digests of human colostrum sIgA. *J. Proteome Res.* 9, 1063–1075
- 237 Goldberg, D. *et al.* (2005) Automatic annotation of matrix-assisted laser desorption/ionization N-glycan spectra. *Proteomics* 5, 865–875
- 238 Goldberg, D. *et al.* (2006) Automatic determination of O-glycan structure from fragmentation spectra. *J. Proteome Res.* 5, 1429–1434
- 239 Goldberg, D. *et al.* (2007) Automated N-glycopeptide identification using a combination of single- and tandem-MS. *J. Proteome Res.* 6, 3995–4005
- 240 Irungu, J. *et al.* (2007) Simplification of mass spectral analysis of acidic glycopeptides using GlycoPep ID. *Anal. Chem.* 79, 3065–3074
- 241 Woodin, C.L. *et al.* (2012) GlycoPep grader: a web-based utility for assigning the composition of N-linked glycopeptides. *Anal. Chem.* 84, 4821–4829
- 242 Chandler, K.B. *et al.* (2013) Exploring site-specific N-glycosylation microheterogeneity of haptoglobin using glycopeptide CID tandem mass spectra and glycan database search. *J. Proteome Res.* 12, 3652–3666
- 243 Zhu, Z. *et al.* (2013) GlycoPep Detector: a tool for assigning mass spectrometry data of N-linked glycopeptides on the basis of their electron transfer dissociation spectra. *Anal. Chem.* 85, 5023–5032
- 244 He, L. *et al.* (2014) GlycoMaster DB: software to assist the automated identification of N-linked glycopeptides by tandem mass spectrometry. *J. Proteome Res.* 13, 3881–3895
- 245 Gaucher, S.P. *et al.* (2000) STAT: a saccharide topology analysis tool used in combination with tandem mass spectrometry. *Anal. Chem.* 72, 2331–2336
- 246 Tang, H. *et al.* (2005) Automated interpretation of MS/MS spectra of oligosaccharides. *Bioinformatics* 21 (Suppl. 1), 431–439
- 247 Maass, K. *et al.* (2007) Glyco-peakfinder – *de novo* composition analysis of glycoconjugates. *Proteomics* 7, 4435–4444
- 248 Nairn, A.V. *et al.* (2008) Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J. Biol. Chem.* 283, 17298–17313
- 249 Hossler, P. *et al.* (2006) GlycoVis: visualizing glycan distribution in the protein N-glycosylation pathway in mammalian cells. *Biotechnol. Bioeng.* 95, 946–960
- 250 Nakahara, T. *et al.* (2008) Glycoconjugate Data Bank: structures – an annotated glycan structure database and N-glycan primary structure verification service. *Nucleic Acids Res.* 36, D368–D371
- 251 Aoki-kinoshita, K.F. and Kanehisa, M. (2009) Bioinformatics analysis of glycan structures from a genomic perspective. In *Bioinformatics for Glycobiology and Glycomics* (Luetteke, T. and Frank, M., eds), pp. 125–141, John Wiley & Sons
- 252 Hashimoto, K. *et al.* (2006) KEGG as a glycome informatics resource. *Glycobiology* 16, 63–70R
- 253 Kanehisa, M. *et al.* (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* 34, D354–D357
- 254 Nimmagadda, S. *et al.* (2008) GlycoVault: a bioinformatics infrastructure for glycan pathway visualization, analysis and modeling. *Information Technology: New Generations. Fifth International Conference* pp. 692–697
- 255 Ceroni, A. *et al.* (2007) The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures. *Source Code Biol. Med.* 2, 3
- 256 Aoki-Kinoshita, K.F. (2015) Analyzing glycan structure synthesis with the Glycan Pathway Predictor (GPP) tool. *Methods Mol. Biol.* 1273, 139–147
- 257 Lauc, G. *et al.* (2010) Complex genetic regulation of protein glycosylation. *Mol. Biosyst.* 6, 329–335
- 258 Mossner, E. *et al.* (2010) Increasing the efficacy of CD20 antibody therapy through the engineering of a new type II anti-CD20 antibody with enhanced direct and immune effector cell-mediated B-cell cytotoxicity. *Blood* 115, 4393–4402
- 259 James, L.C. *et al.* (1999) 1.9 A structure of the therapeutic antibody CAMPATH-1H Fab in complex with a synthetic peptide antigen. *J. Mol. Biol.* 289, 293–301