Review

Function and 3D Structure of the *N*-Glycans on Glycoproteins

Masamichi Nagae and Yoshiki Yamaguchi *

Structural Glycobiology Team, RIKEN, Advanced Science Institute, 2-1 Hirosawa, Wako-City, Saitama 351-0198, Japan; E-Mail: mnagae@riken.jp

* Author to whom correspondence should be addressed; E-Mail: yyoshiki@riken.jp; Tel.: +81-48-467-9619; Fax: +81-48-467-9620.

Abstract: Glycosylation is one of the most common post-translational modifications in eukaryotic cells and plays important roles in many biological processes, such as the immune response and protein quality control systems. It has been notoriously difficult to study glycoproteins by X-ray crystallography since the glycan moieties usually have a heterogeneous chemical structure and conformation, and are often mobile. Nonetheless, recent technical advances in glycoprotein crystallography have accelerated the accumulation of 3D structural information. Statistical analysis of "snapshots" of glycoproteins can provide clues to understanding their structural and dynamic aspects. In this review, we provide an overview of crystallographic analyses of glycoproteins, in which electron density of the glycan moiety is clearly observed. These well-defined *N*-glycan structures are in most cases attributed to carbohydrate-protein and/or carbohydrate-carbohydrate interactions and may function as "molecular glue" to help stabilize inter- and intra-molecular interactions. However, the more mobile *N*-glycans on cell surface receptors, the electron density of which is usually missing on X-ray crystallography, seem to guide the partner ligand to its binding site and prevent irregular protein aggregation by covering oligomerization sites away from the ligand-binding site.

Keywords: glycoprotein; glycoform; *N*-glycan; protein crystallography

Abbreviations: GlcNAc, *N*-acetyl-D-glucosamine; Man, D-mannose; Gal, D-galactose; Sia, sialic acid; Glc, D-glucose; Fuc, L-fucose; GnT, *N*-acetylglucosaminyltransferase; CHO, Chinese hamster ovary; HEK, human embryonic kidney; IgG, immunoglobulin G; Fc,

crystallizable fragment; ADCC, antibody-dependent cellular cytotoxicity; FcRn, neonatal Fc receptor; NA, neuraminidase; APA, *Antheraea pernyi* arylphorin; Tr-β-gal, *Trichoderma reesei* β-galactosidase; TLR, Toll-like receptor; ICAM, intracellular adhesion molecule

1. Introduction

Glycosylation is a common and highly diverse modification of proteins and occurs during or after protein synthesis [1]. More than 50% of eukaryotic proteins are glycosylated [2]. Glycosylation profoundly alters the behavior of proteins, making them more soluble, protecting them from proteolysis, covering antigenic sites, and altering the orientation of proteins on cell surfaces. The importance of protein glycosylation is becoming widely recognized through studies on protein localization and trafficking, biological half-life as well as investigations of cell-cell interactions. In almost all glycoproteins, the carbohydrate units are attached to the protein backbone either by *N*- or *O*-glycosidic bonds or both. *N*-glycans are covalently attached to proteins at the amide of asparagine (Asn) residues, forming an *N*-glycosidic bond. The consensus sequence for *N*-glycosylation, called a sequon, is Asn-X-Ser/Thr, where X can be any amino acid except proline. In *O*-glycosylation, the glycan is attached to the side chains of serine or threonine residues. Unlike *N*-linked glycosylation, no consensus sequence defining an *O*-linked glycosylation site has been reported.

Figure 1. (**a**) Representative chemical structures of high-mannose and complex-type *N*-glycans. (**b**) *N*-glycan processing pathways in mammalian cells. The enzymes and structures of intermediate *N*-glycans are shown. Glc I, α-glucosidase I; Glc II, α-glucosidase II; ER Man, ER α-mannosidase; α-ManI, α-mannosidase I; GnT I, β-*N*-acetylglucosaminyltransferase I; α-ManII, α-mannosidase II; GnT II, β-*N*-acetylglucosaminyltransferase II; β4GalT, β-1,4-galactosyltransferase; SiaT, sialyltransferase; GnT III, β-*N*-acetylglucosaminyltransferase III; GnT V, β-*N*-acetylglucosaminyltransferase V; and α1,6-fucosyltransferase, Fut8.
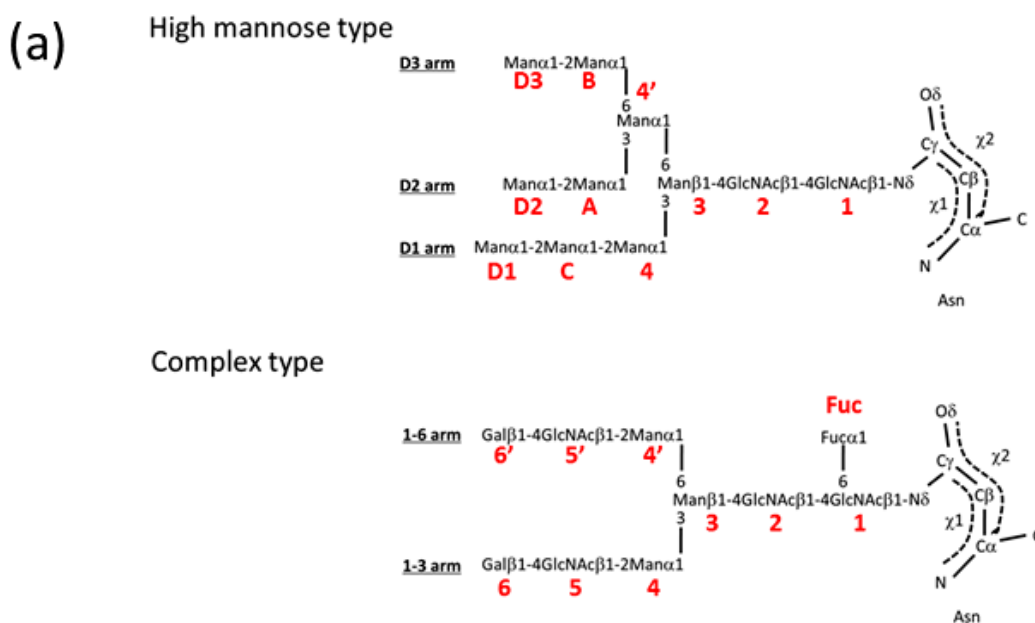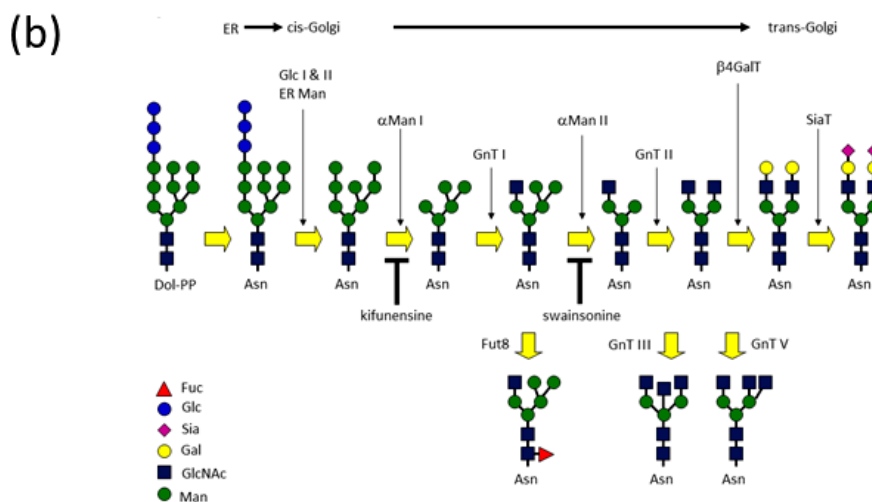
Figure 1. *Cont.*

Human *N*-glycan is typically composed of *N*-acetyl-D-glucosamine (GlcNAc), D-mannose (Man), D-galactose (Gal), sialic acid (general name for *N*-acetylneuraminic acid, which can be abbreviated to Sia or Neu5Ac), D-glucose (Glc) and L-fucose (Fuc) residues. *N*-glycan is classified into three groups: high-mannose type, hybrid type, and complex type. Representative chemical structures of high-mannose (Man$_9$GlcNAc$_2$) and biantennary complex-type (Gal$_2$GlcNAc$_2$Man$_3$GlcNAc$_2$) glycans are shown in Figure 1a. Both types contain Man$_3$GlcNAc$_2$ core structures. In high-mannose type glycan, Manα1-2Manα1-2Man, Manα1-2Manα1-3Man, and Manα1-2Manα1-6Man chains are designated the D1, D2 and D3 arms, respectively. In biantennary complex-type glycan, the α1-3, and α1-6 branched oligosaccharide chains (Galβ1-4GlcNAcβ1-2Man) are termed α1-3 and α1-6 arms, respectively. *N*-glycan precursor (Glc$_3$Man$_9$GlcNAc$_2$) is assembled on a lipid carrier, dolichylpyrophosphate (Dol-PP), and is transferred onto a polypeptide chain by oligosaccharyltransferase (OST) in the endoplasmic reticulum (ER) lumen (Figure 1b). The attached *N*-glycan is sequentially processed by various glycosyl-hydrolases and -transferases. Three glucose and one mannose residues are immediately removed by glucosidase I, II and ER α-mannosidase. Removal of glucose residues is closely related to the folding of glycoproteins. Further processing of *N*-glycans occurs along the secretory pathway as the properly folded glycoprotein moves through the Golgi apparatus to its final destination. The initial stage of the pathway consists of several trimming steps by α-mannosidase I that generate a key intermediate, Man$_5$GlcNAc$_2$. In the Golgi apparatus, the conversion of high-mannose type to hybrid-type glycan is initiated by the action of *N*-acetylglucosaminyl transferase I (GnT I), which transfers GlcNAc, in a β1-2 linkage, to the α1-3 arm mannose residue of Man$_5$GlcNAc$_2$ substrate. Hybrid-type glycan is further remodeled by a series of enzymes to yield complex-type *N*-glycans. There are several pathways which increase the heterogeneity of *N*-glycans, including *N*-acetylglucosaminyltransferase III (GnT III), V (GnT V), and α1,6-fucosyltransferase (Fut8) (Figure 1b). GnT III catalyzes the addition of GlcNAc via a β1-4 linkage to the β-Man of the mannosyl core of *N*-glycans. This GlcNAc is designated as "bisecting GlcNAc" and is involved in the suppression of cancer metastasis [3]. GnT V catalyzes the addition of a β1-6 linked GlcNAc unit to α1-6 linked Man of the trimannosyl core of *N*-linked glycans to form tri- or tetra-antenary branches [4]. Fut8 introduces a fucose residue onto the innermost GlcNAc of the *N*-linked biantennary complex-type

oligosaccharides via an α1-6 linkage [5] (Figure 1b). This reaction is denoted as "core-fucosylation" and the Fuc residue is termed "core-fucose". Each glycosyltransferase has unique substrate specificity that determines the *N*-glycan structure. For example, the action of GnT V and Fut8 is inhibited by the presence of a bisecting GlcNAc residue. Through these modifications, complex-type glycan is often branched and contains a trimannosyl core, several GlcNAc, Gal, sialic acid and Fuc residues, resulting in a high-level of complexity (Figure 1b).

3D structural information on glycoproteins helps in the understanding of *N*-glycan function. However, the inherent flexibility of glycan hampers X-ray crystallographic analysis. Actually, diffraction quality crystals of glycoproteins are normally only obtained following either the abolition of the glycosylation site by site-directed mutagenesis or enzymatic deglycosylation treatment [6–8]. The number of glycoproteins resolved by X-ray or NMR represents less than 3% of the total number of reported 3D structures in the Protein Data Bank (PDB) [9]. However, significant improvements in experimental methods have led to an increase in the number of solved glycoprotein structures with well-defined carbohydrate residues. Various methodologies using mammalian expression systems have been developed for the production of large amounts of homogeneous glycoproteins. Chinese hamster ovary (CHO)-lec 3.2.8.1 cells [10] and human embryonic kidney (HEK) 293S GnT I deficient cells [11] are suitable for producing glycoproteins for crystallization. Since both cell lines lack GnT I activity, uniform glycoprotein can be produced with the truncated *N*-linked oligosaccharide, $Man_5GlcNAc_2$ (Figure 1b). Moreover, co-cultivation of mammalian cells with a *N*-glycosylation inhibitor, kifunensine or swainsonine, reduces the chemical heterogeneity of the product glycoform and can be readily deglycosylated with endoglycosidase (Figure 1b, [12]). Other eukaryotic expression systems, such as yeast and insect cells, are also available for producing homogeneous protein for crystallography. However, in fungal and insect expression systems, mature *N*-linked glycan is generally of the heterogeneous high-mannose type, whereas human *N*-glycans are mainly hybrid or complex types. Since glycan structure depends on the expression cell type, the relationship between glycoform and its physiological function needs careful inspection.

Previous statistical analyses of the available X-ray diffraction data on oligosaccharides [9,13,14] identified the energetically-favorable conformations for individual sugar linkages. GLYCOSCIENCES.de web portal [15] and the Glycoconjugate Data Bank [16] offer convenient ways to search for carbohydrate structures in the PDB. In this review, we introduce several examples in which the glycans affect intra- and inter-molecular interactions. The glycans are often stabilized to assume highly ordered structures via extensive protein-carbohydrate or carbohydrate-carbohydrate interactions. We also introduce several cell-surface glycoprotein structures where the glycans seem to assist in proper ligand recognition and to prevent protein aggregation, although many of these glycans have disordered structures.

2. Conserved Glycan Mediates Inter-Subunit Interactions of Proteins (Fc Fragment and Influenza Neuraminidase)

2.1. N-*Glycan on Fc Fragment Affects Intra- and Inter-Molecular Interactions*

2.1.1. Overview of *N*-Glycan of Fc Fragment

Immunoglobulins are Y-shaped glycoproteins that participate in the adaptive component of the immune system which is directed against extracellular pathogens. The structural diversity of their antigen binding sites allows them to bind specifically to millions of structurally unique molecules. Immunoglobulin G (IgG) consists of two light and two heavy chains and comprises three independent parts connected through a flexible linker or hinge (Figure 2a). Two of these, the Fab fragments, are identical in structure, each with an antigen-specific binding site. The third one, the Fc fragment, has a highly conserved structure even between different isotypes [17–20]. It has antibody effector functions such as antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC) through its interaction either with lymphocyte receptors (Fc receptors, FcγRs) on effector cells such as natural killer cells or with the C1q component of complement. The Fc fragment is a dimer and displays a horseshoe-like arrangement of two antiparallel β-sandwich domains, named C_H2 and C_H3, connected by a short flexible linker (Figure 2a). Interaction between subunits is mainly through the pair of C_H3 domains. These two domains form a compact, non-covalent dimer with a buried surface area of ~1,000 Å^2. A biantennary complex-type glycan is attached onto each Asn297 in the C_H2 domains [17]. Characteristics of Fc *N*-glycosylation include a low incidence of monosialylation, no disialylation, little bisecting GlcNAc, a high incidence of core fucose and heterogeneity of galactose residues [21]. The particular glycoform of the Fc fragment impacts on its physiological function. Removal of fucose is known to enhance ADCC activity both *in vitro* [22] and *in vivo* [23]. Moreover, the galactose content of human IgG-Fc correlates inversely with disease progression in rheumatoid arthritis and other auto-immune diseases [24]. The anti-inflammatory activity of intravenous Ig (IVIG) can be recapitulated with a fully recombinant preparation of appropriately sialylated IgG Fc fragments [25]. Thus, manipulation of Asn297 glycan structures has emerged as a strategy to modulate effector functions of therapeutic antibodies [26,27].

Pioneering X-ray crystallographic studies of the isolated human IgG1 Fc domain [17] and rabbit IgG [28] have shown that the two conserved *N*-glycan chains of the Fc are well defined, with the oligosaccharide bound to the surfaces of the C_H2 domains. Since then, our structural knowledge of Fc fragments and their glycans have grown substantially. As of March 2012, there are more than 30 PDB entries. The structures containing glycosylated Fc fragments are summarized in Table 1. A representative structure of human IgG1 Fc and its *N*-glycan is shown in Figure 2b. The α1-6 arm of the *N*-glycan of Asn297 contacts two hydrophobic residues, Phe241 and Phe243, and makes several hydrogen bonds with Lys246, Asp265, and Arg301 (Figure 2b). A core Fuc is located in the vicinity of Tyr296 and indirectly affects the hydration mode of Tyr296 [29]. In contrast, Tyr313, which corresponds to Tyr296 in human IgG1, makes direct hydrogen bonds with a core Fuc residue in the mouse IgG2a structure [19].

Figure 2. (**a**) Overall structure of immunoglobulin G (PDB code; 1igt) is shown in a ribbon model. One light and two heavy chains are shown in beige, blue and cyan, respectively. Carbohydrate residues attached on the Fc region are shown in sphere models. (**b**) Close-up view of Asn297 attached glycan of human IgG1 Fc (PDB code; 2dts). Carbohydrate moiety and amino acid residues which interact with *N*-glycan are shown in the rod model. Hydrogen bonds between protein and carbohydrate are shown as red dotted lines.
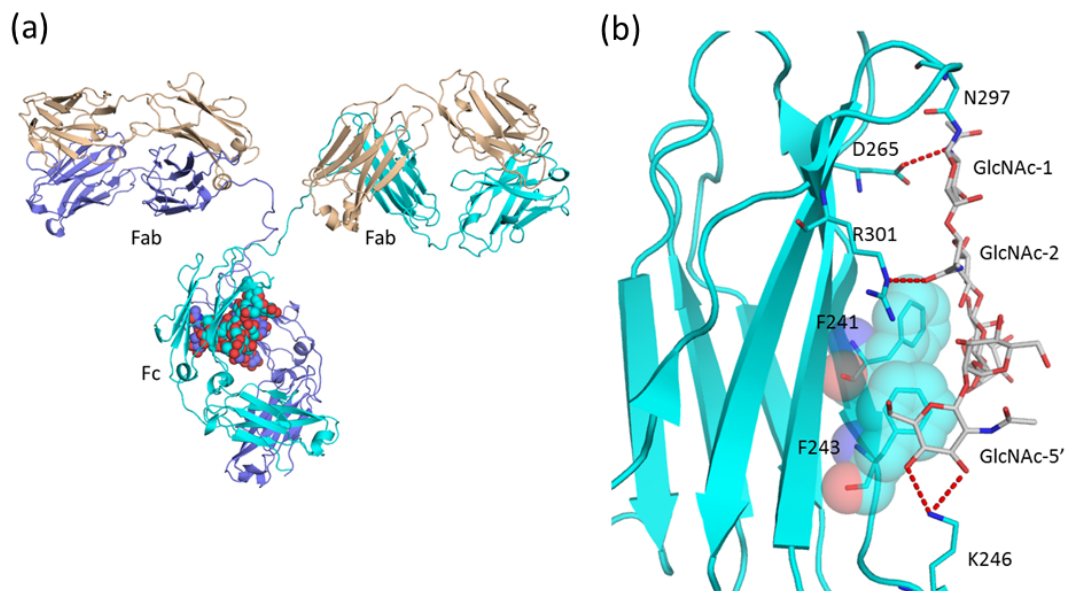

Table 1. Summary of Fc fragment structures.

PDB ID	Glycan structure *	Resolution	Reference
Human IgG1 Fc fragment produced by HEK293T cell + kifunensine			
2wah	M_7GN_2 (chain-A)	2.51	[30]
	M_1GN_2 (chain-B)	-	-
Human IgG1 Fc fragment by enzymatic treatment			
1h3t	M_1GN_2F	2.4	[31]
1h3u	M_3GN_2F	2.4	enzymatic
1h3x	$GN_2M_3GN_2F$	2.44	treatment
1h3v	$G_2GN_2M_3GN_2F$	3.1	[32])
1h3w	$G_2GN_2M_3GN_2F$	2.82	-
1h3y	$G_2GN_2M_3GN_2F$	4.1	-
Human IgG1 Fc produced by CHO or Fut8$^{-/-}$ CHO cells			
3ave (2dtq)	$GN_2M_3GN_2F$	2.0	[29]
2dts	$GN_1M_3GN_2$ (chain-A)	2.2	-
	$GN_2M_3GN_2$ (chain-B)	-	-
Human IgG1 Fc fragment triple mutant (M252Y/S254T/T256E)			
3fjt	$GN_2M_3GN_2F$	2.5	[33]
Human IgG1 Fc fragment triple mutant (L234F/L235E/P331S)			
3c2s	$G_1GN_2M_3GN_2F$	2.3	[34]
Protein is produced by HEK293 cells.			

Table 1. *Cont.*

PDB ID	Glycan structure *	Resolution	Reference
Human IgG1 Fc fragment triple mutant (S239D/A330L/I332E)			
2ql1	$G_1GN_2M_3GN_2F$	2.5	[35]
Human IgG1 Fc fragment + 13 residues peptide			
1dn2	$G_1GN_1M_3GN_2F$ (chain-A)	2.7	[36]
	$G_1GN_2M_3GN_2F$ (chain-B)	-	-
Human IgG1 (Rituxan) Fc fragment + *Staphylococcus aureus* Protein A domain B			
1l6x	$G_2GN_2M_3GN_2F$	1.65	[37]
Human IgG1 Fc fragment			
3do3	$G_1GN_2M_3GN_2F$	2.5	[38]
Human IgG1			
1hzh	$G_2GN_2M_3GN_2F$ (chain-H)	2.7	[39]
	$G_1GN_2M_3GN_2F$ (chain-K)	-	-
Mouse IgG2a			
1igt	$G_1GN_2M_3GN_2F$	2.8	[19]
Human IgG1 Fc fragment + Protein-A mimetic peptide dendrimer.			
3d6g	$GN_2M_3GN_2F$	2.30	[40]
Mouse IgG2b Fc fragment			
2rgs	$GN_2M_3GN_2F$	2.1	[41]
Rabbit IgG Fc fragment			
2vuo	$G_1GN_2M_3GN_2$	1.95	[42]
Human IgG1 Fc fragment + minimized protein A			
1oqo	$GN_2M_3GN_2F$	2.3	[43]
1oqx	M_3GN_2F	2.6	-
Human IgG Fc fragment			
1fc1	$G_1GN_2M_3GN_2F$	2.9	[17]
Rat IgG2a Fc fragment			
1i1c	$GN_2M_3GN_2F$	2.7	[18]
Human IgG1 Fc fragment + human Fc receptor (FcγRIIIb)			
1t83 (1its)	$G_1GN_1M_3GN_2F$ (chain-A)	3.0	[44]
	$GN_2M_3GN_2F$ (chain-B)	-	-
1t89 (1iix)	$G_1GN_1M_3GN_2F$ (chain-A)	3.5	-
	$GN_2M_3GN_2F$ (chain-B)	-	-
Human IgG1 Fc fragment + human Fc receptor (FcγRIIIb)			
1e4k	$G_1GN_2M_3GN_2F$	3.2	[45]
Human IgG1 Fc fragment + human Fc receptor (FcγRIIIa)			
3sgj	$GN_2M_3GN_2F$	2.2	[46]
3sgk	$GN_3M_3GN_2$	2.4	-
Human IgG1 Fc fragment + human Fc receptor (FcγRIIIa)			
3ay4	$G_1GN_2M_3GN_2$	2.2	[47]
Human heterodimeric Fc + human neonatal FcR (FnRn)			
1i1a	$GN_2M_3GN_2F$	2.8	[18]

  * Glycan structures deposited in each PDB coordinate file are shown. G: D-galactose, GN: *N*-acetyl-D-glucosamine, M: D-mannose, F: L-fucose.

Previous comparative conformational analysis has demonstrated that both glycosylated and non-glycosylated Asn side chains exhibit $\chi 1$ (N-Cα-Cβ-Cγ) torsion angles of $-60°$, $60°$, $180°$, corresponding to the $g-$, $g+$ and t conformers, respectively. The $\chi 2$ (Cα-Cβ-Cγ-Nδ) torsion angle shows a wide range centered at about $0°$, but is nevertheless limited by glycosylation. In the case of glycosylated Asn side chains, the $g-$ conformer is most preferred and the g^+ conformer rarest in both glycosylated and non-glycosylated Asn residues [9,48]. The conformation of 56 Fc *N*-glycans is statistically analyzed in Figure 3, from the 29 different Fc structures listed in Table 1. The side chain torsion angles, $\chi 1$ and $\chi 2$, of Asn297 are plotted in Figure 3a. In the case of Fc structures, the $\chi 1$ torsion angle of Asn297 exhibits a marked preference for $\sim 60°$, corresponding to the g^+ conformer. The torsion angles, ϕ and ψ, for all glycosidic linkages are plotted in Figure 3b. All the dihedral angles of the glycosidic linkages are within an acceptable region compared with other glycoproteins [48]. The dihedral angle distribution of the $\alpha 1$-6 arms appears to be more restricted than that of the $\alpha 1$-3 arms (Man-4-Man-3, Man-4'-Man-3, GlcNAc-5-Man-4, and GlcNAc-5'-Man-4' in Figure 3b). This difference derives from the fact that the $\alpha 1$-6 arms of the glycan interact with the first two strands of the C$_H$2 domains. Solution NMR analyses of the Fc indicate that the $\alpha 1$-6 arm is completely immobilized through interactions with the protein surface, whereas the $\alpha 1$-3 mannose terminus is more dynamic [49,50]. On the other hand, spin relaxation NMR studies indicate that Fc glycan could exhibit a dynamic conformational state as well as the fixed state as seen in the crystal structure [51].

Figure 3. (**a**) The side-chain torsions of Asn297 of the Fc fragment. Torsion angles of the Asn297 side chain are measured by MolProbity [52]. We excluded Fc with high-mannose type glycan (PDB code 2wah) from this inspection, since this structure contains high-mannose type glycans. (**b**) Comparison of glycosidic torsions of *N*-glycan attached on Asn297 of Fc fragment. Dihedral angles of each linkage are calculated with CARP [15]. The vertical and horizontal axes indicate φ and ϕ angles, respectively. The residues with errors are carefully excluded from this analysis. In many cases, a $\beta 1$-6 linkage is erroneously used between core Fuc and GlcNAc instead of an $\alpha 1$-6 bond [53]. Eight entries are plotted in Fuc-GlcNAc-1 (PDB code; 1h3w, 3ave, 3d6g, 2rgs, chain-A in 1e4k, and chain-B in 3sgj).
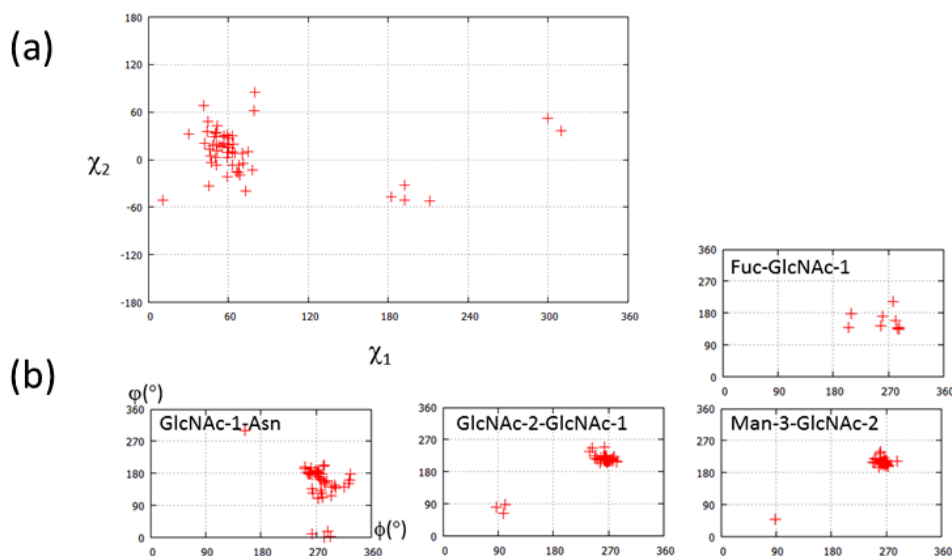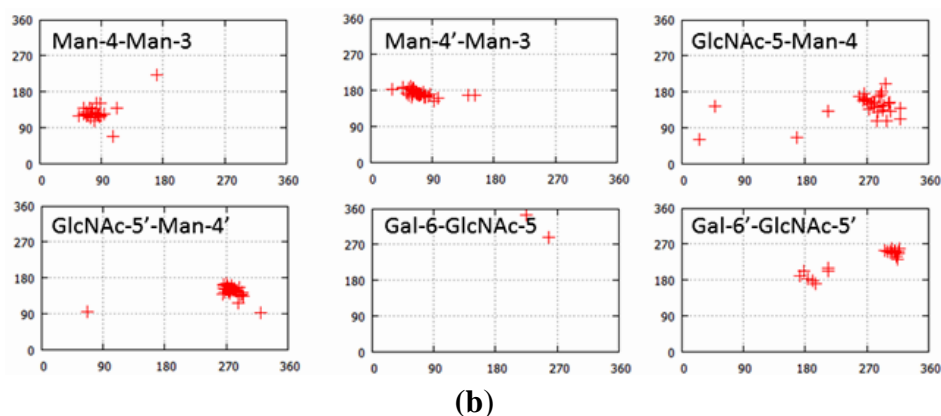
Figure 3. *Cont.*

(b)

2.1.2. Glycoform Affects the Relative Interdomain Angles of the Fc Fragment

Glycan structure can potentially affect the overall structure of a glycoprotein. The influence of glycoform on the conformation of the Fc fragment has been extensively investigated. Two papers report on the relationship between glycoform and the interdomain angles of the C_H2-C_H3 domains.

In the first report, the influence of glycoform on the structure and function of IgG Fc was assessed by sequential exo-glycosidase treatment [31]. Krapp *et al.* solved the crystal structures of human IgG1 Fc of four glycoforms bearing consecutively truncated oligosaccharides (PDB code; 1h3t, 1h3u, 1h3x, 1h3v and 1h3w). Removal of the terminal GlcNAc as well as the mannose residues causes the largest conformational change in both the oligosaccharide and in the polypeptide loop containing the *N*-glycosylation site. The conformational change in the C_H2 domain affects the interface between the IgG-Fc fragments and the FcγRs. Moreover, removal of the sugar residues permits the mutual approach of the C_H2 domains and the generation of a closed conformation. This contrasts with the open conformation of fully galactosylated IgG Fc, which may be optimal for FcγR binding. Solution NMR analysis of a series of Fc glycoforms also indicates that the carbohydrate moieties are required for maintaining the structural integrity of the FcγR binding site [54].

In the second report, Crispin and co-workers solved the crystal structure of a recombinant human IgG1 Fc fragment designed to possess high-mannose type glycan at Asn297 [30]. Recombinant Fc was transiently expressed using HEK 293T cells in the presence of the α-mannosidase I inhibitor, kifunensine (Figure 1b). The glycan structure was confirmed to be $Man_9GlcNAc_2$ by MALDI-TOF-MS and the crystal structure of the glycoform was solved. The electron density map of the high-mannose glycan is quite asymmetric. Extensive branched density is observed in chain A, assigned as $Man_7GlcNAc_2$. On the other hand, poor electron density is found in chain B corresponding to three reducing terminal saccharide residues ($Man_1GlcNAc_2$). The overall structure of the high-mannose glycan is similar to that of a complex type. In Figure 4a, there is a structural comparison between high-mannose type (chain A of 2wah) and complex-type glycans (PDB code; 2dts). The $\chi1$ and $\chi2$ angles of the side chain of Asn297 are essentially identical to those of other Fc fragments possessing complex-type glycans. Moreover, the dihedral angles of the glycosidic linkages in a pentasaccharide $Man_3GlcNAc_2$ core of both glycans are also similar (Figure 4a). However, the α1-6 branches (D2 and D3 arms) of the high-mannose type occupy different positions compared with complex-type glycans.

This is due to the difference between the α1-6 and β1-2 glycosidic bonds. Structural superposition of two Fc fragments reveals that the attachment of high-mannose type *N*-glycans opens the interdomain cavity between the C_H2 domains (Figure 4b). This structural difference between the two glycoforms might explain the evidence that recombinant monoclonal antibody with human oligomannose-type glycans display enhanced ADCC, together with reduced complement activation through C1q binding [55]. It should be noted that the expression construct of the recombinant Fc fragment lacks the cysteine residues which form a disulfide bond in the hinge region. Further studies are needed to fully reveal the relationship between glycoform and structure.

Figure 4. (**a**) Structural comparison between high-mannose glycan (PDB code; 2wah, green) and complex-type glycan (PDB code; 2dts, cyan). (**b**) Structural superposition between high-mannose type Fc fragment (green) and complex-type Fc fragment (cyan). Protein molecules and carbohydrate chains are shown in wire and stick models, respectively. The positions of Asn297 are indicated by red asterisks. Structural superposition of crystal structures were performed by the program SUPERPOSE [56].
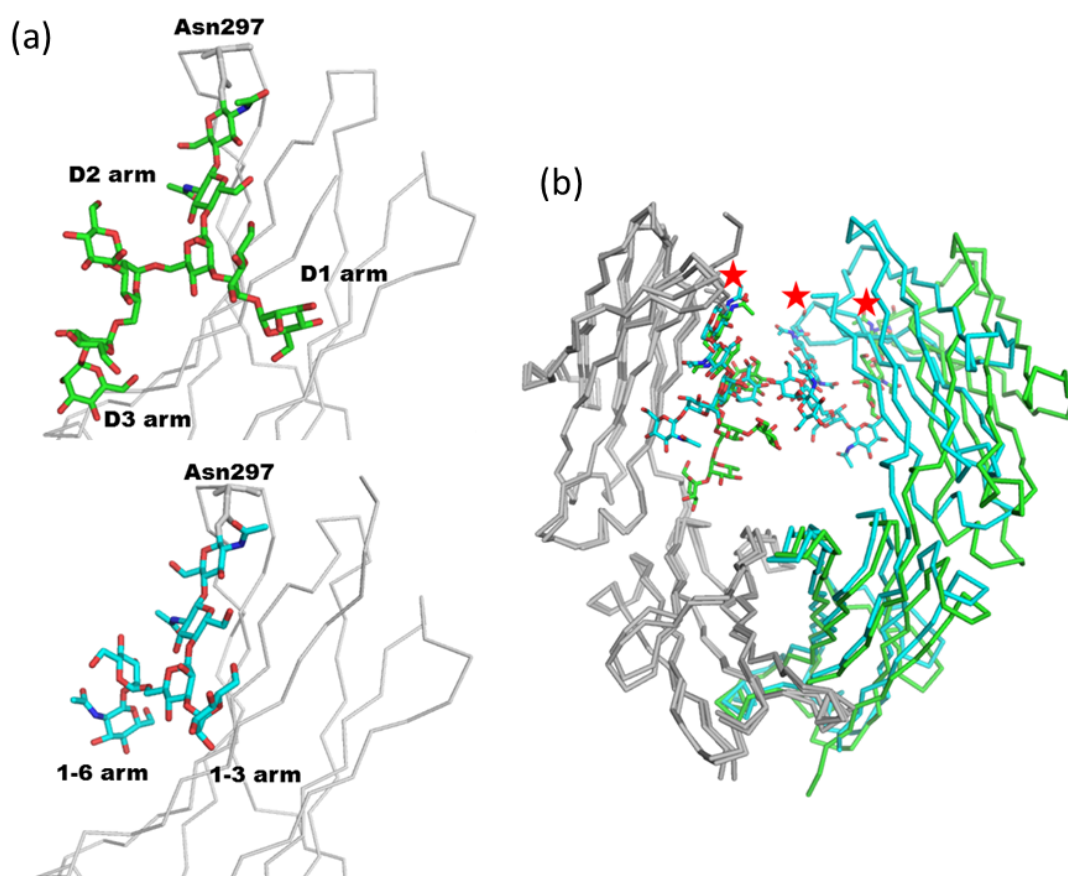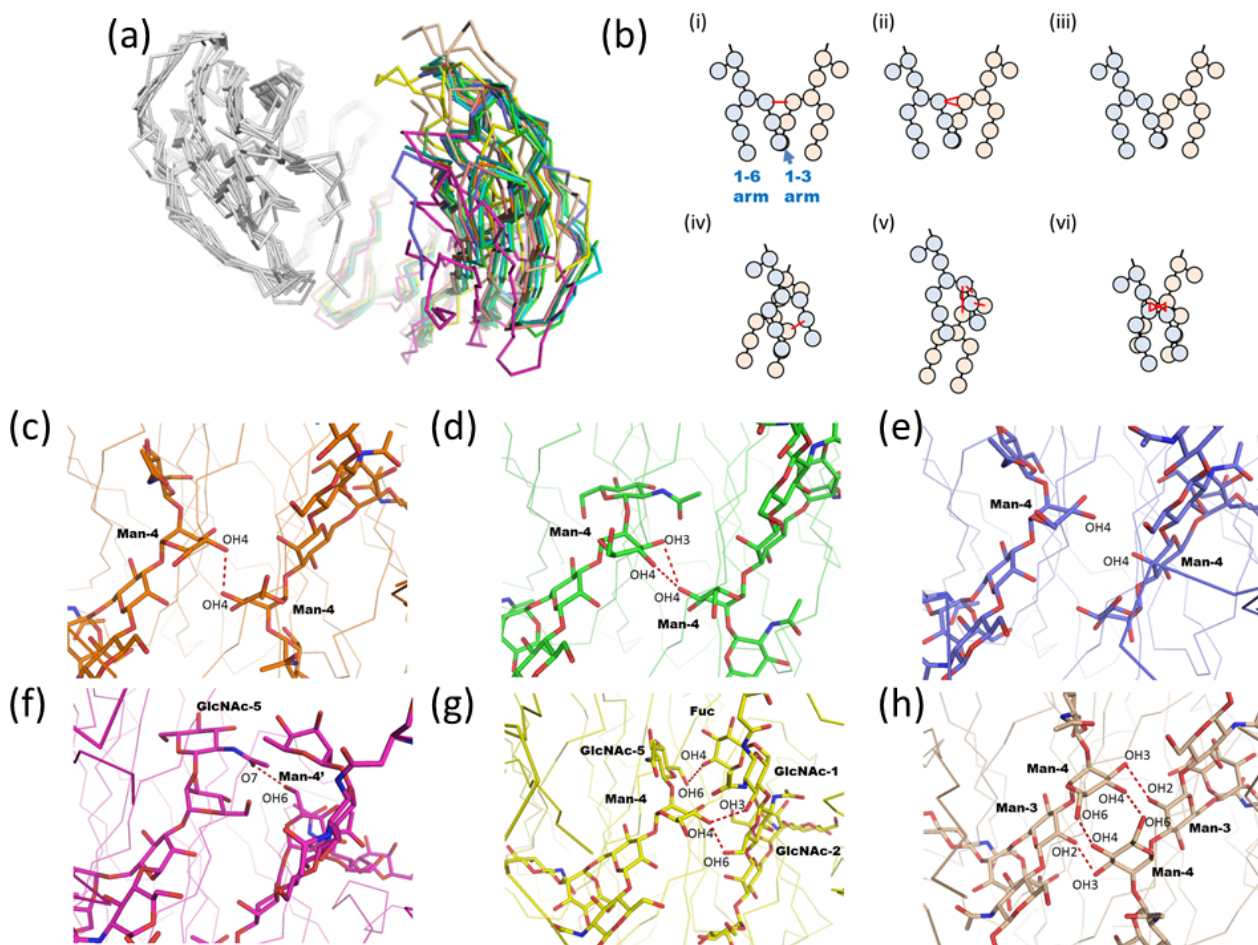

2.1.3. Intra-molecular Carbohydrate-Carbohydrate Interaction

Direct C_H2-C_H2 interaction is accomplished mainly through the pair of *N*-glycans at Asn297. To analyze the intramolecular carbohydrate-carbohydrate interaction modes of Fc fragment, we compared 10 uncomplexed Fc fragment structures from Table 1; wild type human IgG1 Fc (PDB code; 1fc1, 1h3v, 1h3y, 2dts, 3ave, and 3do3), mutated human IgG1 Fc (PDB code; 3fjt), rat IgG2a Fc (PDB code; 1i1c), mouse IgG2b Fc (PDB code; 2rgs), and rabbit IgG Fc (PDB code; 2vuo). Structural

superimposition among these 10 Fc structures reveals that the interdomain angles of the C_H2-C_H3 domains are highly variable (Figure 5a). The highly mobile domain angles result in a variety of carbohydrate-carbohydrate interaction modes as described below. The modes are classified into six types (i~vi), and a schematic representation is shown in Figure 5b. (i) In four of 10 Fc fragments (PDB code; 2dts, 3fjt, 3ave, and 3do3), only one hydrogen bond is found between the OH4 hydroxyl group of each α1-3 linked Man residue (Man-4) (Figure 5c). (ii) In the first reported human IgG1 Fc fragment structure (PDB code; 1fc1), the carbohydrate-carbohydrate interaction mode is slightly different. The OH4 hydroxyl group of the α1-3 linked Man (Man-4) asymmetrically interacts with both OH3 and OH4 of the counterpart α1-3 linked Man (Man-4) (Figure 5d). (iii) In two structures of human IgG1 Fc fragment (PDB code; 1h3v and 2vuo), the distances between the two OH4 hydroxyl groups of the two α1-3 linked Man (Man-4) residues are rather long for a direct hydrogen bond (Figure 5e). In (i) ~ (iii), the Fc structures superimpose well, whereas the remaining three Fc structures (PDB code; 1h3y, 1i1c, and 2rgs) are unique. (iv) The structure of the human IgG1 Fc fragment (PDB code; 1h3y) is completely different compared with other human wild type IgG1 Fc fragments (see magenta in Figure 5a). In this structure, one weak hydrogen bond is observed between O7 of GlcNAc in the α1-3 arm (GlcNAc-5) and OH6 hydroxyl group of the α1-6 linked Man (Man-4') at a distance of 3.2 Å (Figure 5f). The human IgG1 Fc fragment (1h3y) was crystallized in a high salt concentration (2.0 M NaCl) while the other human wild type IgG1 Fc fragments were crystallized under lower ionic strength conditions (1fc1; 30 mM sodium chloride, 1h3v and 2dts; distilled water, 3ave; 20% butanediol, and 3do3; 0.2 M sodium chloride and 20% polyethylene glycol 3,350). The range of ionic strength conditions may contribute to the conformational differences (v) The interaction mode observed in rat IgG2a Fc (PDB code; 1i1c) is asymmetric. The α1-3 arm of one side contacts the core Fuc and the chitobiose of the other (Figure 5g). The O5 and OH6 of GlcNAc in the α1-3 arm (GlcNAc-5) interact with OH4 of the Fuc residue, whereas OH4 of the α1-3 linked Man (Man-4) makes hydrogen bonds with the hydroxyl groups of the chitobiose. (vi) In the crystal structure of mouse IgG2b Fc fragment (PDB code; 2rgs, [40]), four symmetric hydrogen bonds are found between the two oligosaccharide chains. The OH3 of the α1-3 linked Man (Man-4) makes hydrogen bonds with OH2 of its counterpart β-Man (Man-3), and likewise the OH4 with OH6 (Figure 5h).

The interdomain angle of the C_H2-C_H3 domains could be affected by the crystallization conditions, especially ionic strength [31]. Crystallographic analysis only provides static snapshots and gives no direct information on the dynamics of a glycosidic linkage or of polypeptide fluctuations. Thus, it is likely that the single static linkage conformation observed in a crystal does not directly reflect the average solution conformation. Nonetheless, the average conformation for a given linkage within a large set of static structures is likely to correspond well to the average solution conformation, and the distribution of static structures will give an indication of the flexibility of the linkage, as long as crystal packing forces do not impose systematic changes.

Figure 5. (**a**) Structural superposition of 10 Fc fragment structures (1fc1; green, 1h3v; cyan, 1h3y; magenta, 1i1c; yellow, 2dts; pink, 1i1c; yellow, 2rgs; wheat, 2vuo; slate, 3ave; orange, 3do3; lime, 3fjt; deep teal). For four entries (PDB code; 1h3w, 3c2s, 2ql1, and 1l6x), the asymmetric units of these crystals contain only one heavy chain. Thus, the symmetry-related neighboring heavy chains were compensated for in this analysis. Structural superposition was performed by SUPERPOSE. (**b**) Schematic representation of six types of carbohydrate-carbohydrate interaction modes. *N*-glycans of two chains are shown in blue and pink. Hydrogen bonds are shown as red lines. (**c**)–(**h**) Close-up views of the interfaces of carbohydrate-carbohydrate interactions. Carbohydrate moiety is shown in the rod model. Hydrogen bonds between carbohydrates are shown as red dotted lines. The structural superimposition of four structures which have only one carbohydrate-carbohydrate interaction is shown in (c). Human IgG1 Fc fragment (PDB code; 1fc1) is shown in (d). The superimposition of two structures which have no interaction between glycans is shown in (e). Human IgG1 Fc in high salt condition (PDB code; 1h3y), rat IgG2a (PDB code; 1i1c), and mouse IgG2b Fc fragment (PDB code; 2rgs) are shown in (f), (g), and (h), respectively.

2.1.4. Carbohydrate-Assisted Intermolecular Interaction (Neonatal Fc Receptor and Fcγ Receptor IIIa)

Carbohydrates attached to glycoproteins are often required for tight binding to partner proteins. In this section, two topics are introduced in which *N*-linked carbohydrates on receptors contribute to full Fc binding activity.

The neonatal Fc receptor (FcRn) transports maternal immunoglobulin G (IgG) across the neonatal intestine in rodents and across the placenta in humans, thereby conferring humoral immunity to the fetus or newborn against antigens encountered by the mother. FcRn binds IgG with nanomolar affinity at acidic pH (≤6.5) in the intracellular transport vesicles and releases IgG upon encountering the basic pH of the bloodstream (7.4). FcRn is a heterodimer and is composed of a soluble light chain β2-microglobulin and a membrane-bound heavy chain that includes three extracellular domains (α1~3), a single pass transmembrane domain, and a short cytoplasmic domain. FcRn interacts with the C_H2-C_H3 domain interface on each chain of the Fc homodimer [57]. The 2:2 interaction mode creates higher ordered structures, called "oligomeric ribbons", and prohibits the growth of well-ordered co-crystals. To improve the crystal quality, Martin and co-workers designed a heterodimeric version of Fc that cannot bridge between FcRn molecules, since it contains only a single FcRn binding site [58], and solved the crystal structure of a FcRn ectodomain in complex with the heterodimeric Fc fragment at 2.8-Å resolution [18]. The heavy chain (α2) and β2-microglobulin domains of FcRn interact with the Fc C_H2-C_H3 interface (Figure 6a). The binding interface between FcRn and Fc spans a large surface area (buried surface area up to 1,870 $Å^2$) and is highly complementary. The complex is stabilized by extensive hydrophobic and electrostatic interactions. Interestingly, the complex-type *N*-linked glycan attached on Asn128 of FcRn has a highly ordered structure and contributes 10–15% of the total buried surface area in the interface. The core Fuc residue and GlcNAc of the α1-3 arm (GlcNAc-5) tightly interact with the complementary surface of the binding partner (Figure 6b). This suggests that complex-type glycan, but not high-mannose type glycan, on FcRn is required for maximal Fc binding affinity. Actually, differential glycosylation of mouse FcRn could affect the receptor/ligand stoichiometry under non-equilibrium conditions [59].

The crystal structure of the Fc-FcRn complex reveals how the *N*-glycan on FcRn affects the binding affinity. In contrast, the *N*-glycan of IgG Fc strongly influences the interaction between IgG Fc and FcγR, and therapeutic antibodies may be modulated by selecting the appropriate glycoform. For example, removal of Fc core fucose selectively and significantly increases binding affinity to FcγRIII, thereby leading to enhanced cellular immune effector functions, such as ADCC. This may be especially relevant with respect to therapeutic anticancer antibodies [60,61], and has been the focus of research during the past decade. A crystal structure of glycosylated human IgG1 Fc fragment in complex with the unglycosylated extracellular domain of FcγRIII was first reported by Sondermann and co-workers [45]. One FcγRIII binds to the two halves of the Fc fragment and contacts residues in the C_H2 domains and the hinge region. Complex formation significantly increases the angle between the two soluble FcγRIII domains and the Fc fragment is asymmetrically open. Recently, crystal structures of afucosylated human Fc in complex with glycosylated FcγRIIIa ectodomain were solved by two independent groups (Figure 6c) [46,47]. Although both groups produced the glycosylated FcγRIIIa ectodomain using mammalian expression systems, the glycan structures attached on Asn162 differ. Ferrara and colleagues report that the *N*-glycan structure is high-mannose type, while

Mizushima *et al.* find asialylated complex type. The overall fold of the Fc-FcγRIIIa complexes where both proteins are glycosylated is very similar to that of the complexes where only the Fc protein is glycosylated. Clear electron density was obtained for both the Asn162-linked glycan of the receptor and the glycans linked to the Fc fragment. The carbohydrate attached on Asn162 shares a large interaction surface area (approximately 12% of the total interface area —145 Å^2—in the case of PDB code; 3ay4) with the Fc formed by various polar, van der Waals, and hydrogen bond interactions. The receptor Asn162-carbohydrate interactions center on the Asn297-carbohydrate core of Fc chain A and its immediate vicinity (Figure 6d). Overall, a combination of direct or water-mediated carbohydrate-carbohydrate and carbohydrate-protein contacts are observed as part of the newly formed interaction between afucosylated Fc and the Asn162-glycosylated receptor.

Ferrara and colleagues also solved the crystal structure of fucosylated Fc in complex with glycosylated FcγRIIIa ectodomain. The core fucose linked to Fc is oriented towards the second GlcNAc (GlcNAc-2) of the chitobiose connected to Asn162 of FcγRIIIa and has to be accommodated in the interface between the interacting glycan chains. This steric rearrangement causes the movement of the whole oligosaccharide attached on Asn162 up to a maximum distance of 2.6 Å while almost no movement is observed in the case of afucosylated Fc. This rearrangement of the interaction network reduces the enthalpy contribution in the fucosylated Fc complex. It is noteworthy that even such subtle displacement of carbohydrate chains affects physiological activity, such as in ADCC [46].

Figure 6. (**a**) Overall structure of neonatal Fc receptor (FcRn) in complex with heterodimeric Fc (hdFc) (PDB code; 1i1a). Heavy chain and soluble light chain β2-microglobulin (β2m) of FcRn are shown in slate and cyan, respectively. Proximal and distal Fc fragments of hdFc are shown in pink and white, respectively. The region delineated in black dotted lines is magnified in (b). (**b**) Close-up view of FcRn-hdFc complex. *N*-glycan attached at Asn128 of FcRn is shown in rod and semitransparent sphere model. (**c**) Overall structure of human Fc-glycosylated human Fcγ receptor IIIa (FcγRIIIa) complex (PDB code; 3sgk). Two chains of Fc fragment and FcγRIIIa are shown in green, cyan, and yellow, respectively. The region delineated in black dotted lines is magnified in (d). (**d**) Close-up view of carbohydrate-carbohydrate interaction in Fc-FcγRIIIa. Hydrogen bonds are shown as red dotted lines.
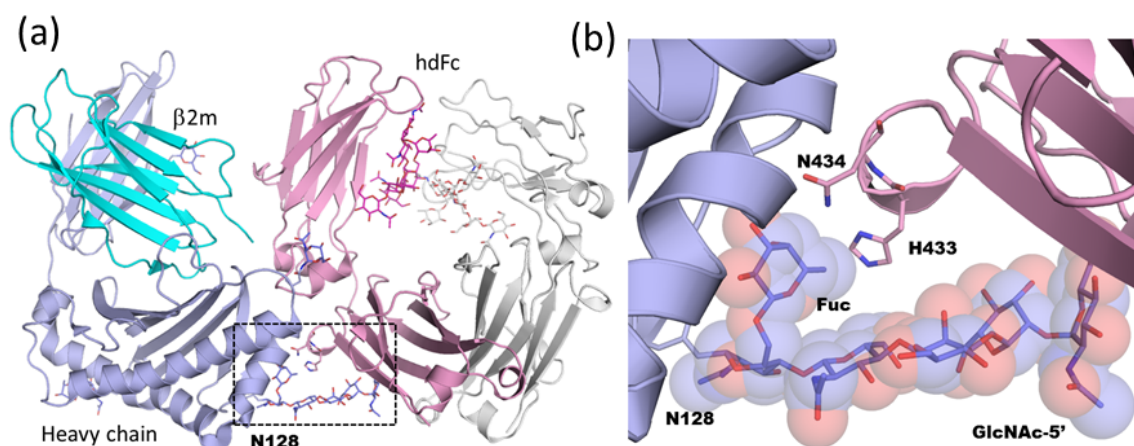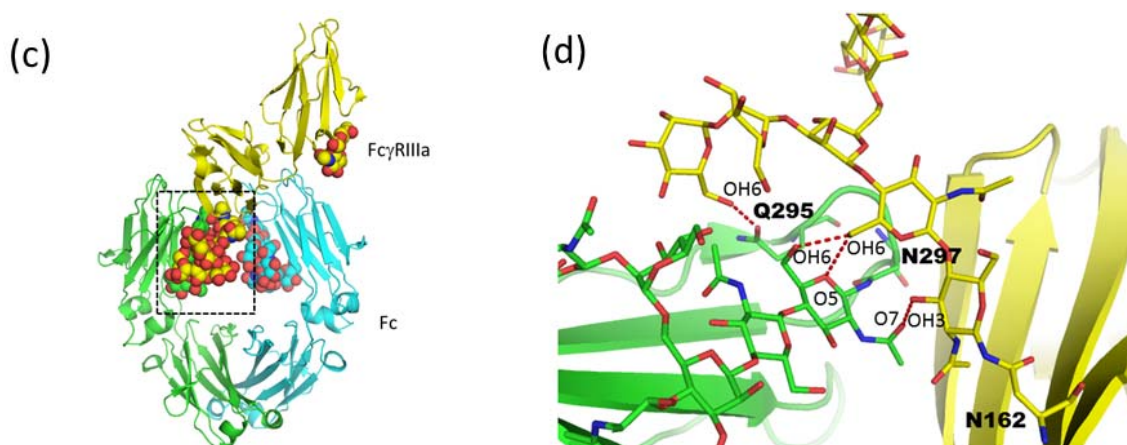
Figure 6. *Cont.*

2.2. High-Mannose Type Glycan on Group 2 Influenza Virus Neuraminidase

Influenza virus infection has been a major threat to public health throughout the world for centuries. Influenza types A and B are enveloped RNA viruses carrying two glycoproteins on their surface, hemagglutinin (HA) and neuraminidase (NA, acylneuraminyl hydrolase, EC 3.2.1.18). Influenza NA removes terminal α2-3 or α2-6 linked sialic acid residues from carbohydrate moieties on cell surface glycoconjugates and is thought to thereby facilitate virus release and infection of another cell. Inhibition of NA delays the release of progeny virions from the surface of infected cells [62], suppressing the viral population, thus allowing time for the host immune system to eliminate the virus.

Antigenic differences are used to classify influenza type A viruses into nine NA (N1–N9) subtypes [63]. Phylogenetically, there are two groups of NAs: group 1 contains N1, N4, N5 and N8, and group 2 contains N2, N3, N6, N7 and N9 [64].

In both influenza A and B, functional NA is a tetramer of identical subunits with four-fold rotational symmetry. The NA tetramer forms a box-like head on top of a long stalk domain and is anchored in the viral membrane by a hydrophobic sequence near the N-terminus [65,66]. The surface of an influenza virus typically has about 50 tetrameric NA spikes [67]. NA has been targeted in structure-based enzyme inhibitor design programs that have resulted in the production of two drugs, zanamivir (Relenza) [68] and oseltamivir (Tamiflu) [69], that mimic the transition state of the normal enzyme reaction.

Figure 7. High-mannose type glycan of influenza neuraminidase assists tetramer formation. (**a**) Overall structure of monomeric influenza N2 neuraminidase (PDB code; 1nn2). Protein, carbohydrate, and calcium ion are shown in ribbon, stick, and sphere models, respectively. (**b**) Tetrameric structure of influenza N2 neuraminidase (PDB code; 1nn2). *N*-linked glycans at Asn200 are shown in sphere models. The region delineated in black dotted lines is magnified in (c). (**c**) Close-up view of *N*-glycan at Asn200 and symmetry related molecule. Hydrogen bonds are shown in red dotted lines. (**d**) The side-chain torsion angles of Asn200 of N2, Asn207 of N6, and Asn200 of N9 NA (Asn201 in PDB code; 2b8h). (**e**) Amino acid sequence alignment of group 1 and 2 influenza neuraminidase around Asn200 glycosylation sites. Putative *N*-linked glycosylation sites in group 2 are highlighted.
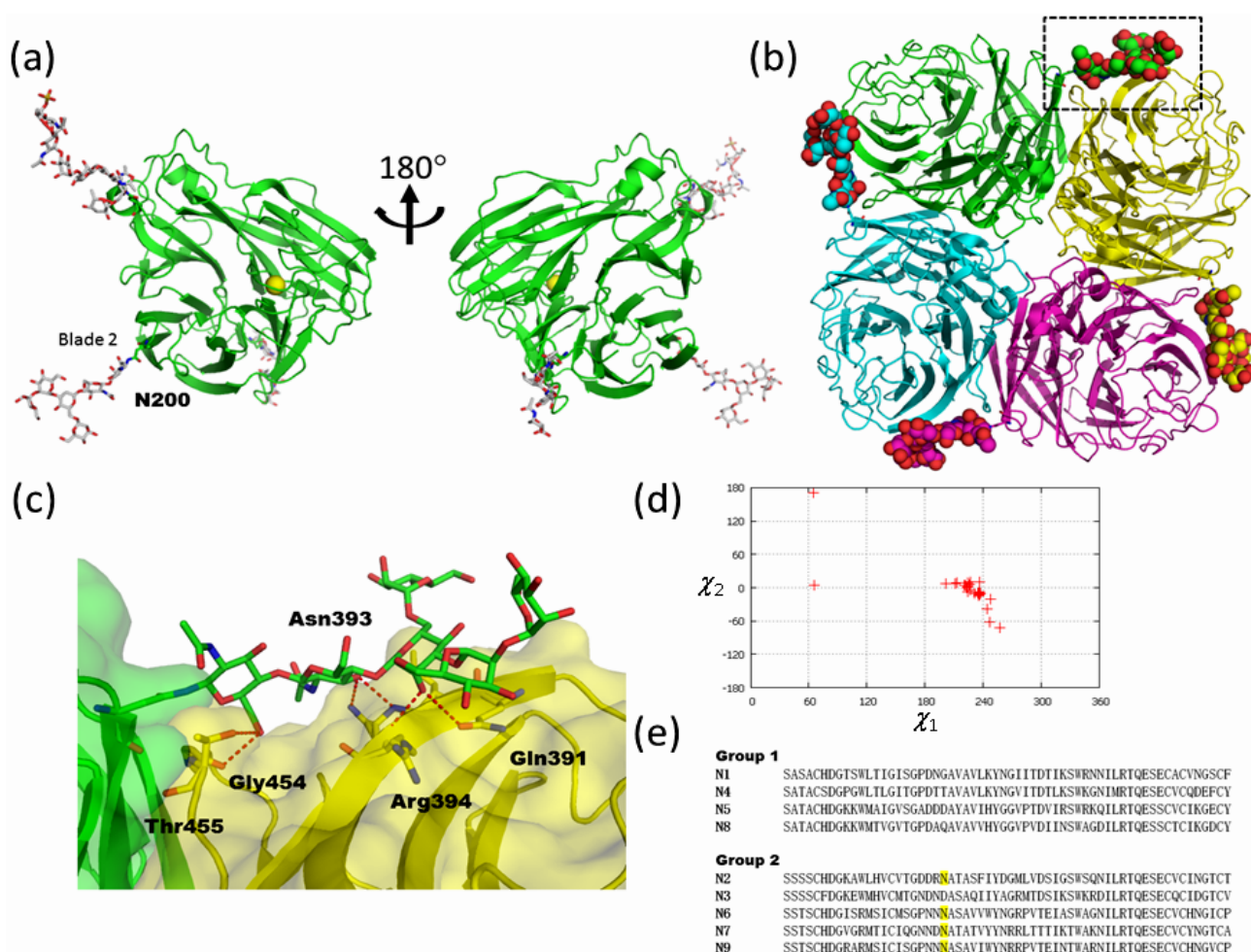
Crystallographic analysis of influenza NA has a history of almost 30 years. The first crystal structure of influenza N2 NA was reported in 1983 [70], and the first N9 in 1987 [71]. The NA monomer consists of a symmetric six-bladed β-propeller arrangement stabilized in part by calcium ions bound on the symmetry axis (Figure 7a). There are five *N*-glycosylation sequons exposed on the protein surface, namely Asn86, Asn146, Asn200, Asn234, and Asn402, the first four of which are assigned in the model. The two *N*-glycans at Asn146 and Asn200 are highly ordered. The glycan at Asn146 is of the core-fucosylated complex type, whereas that at Asn200 is of the high-mannose type. The catalytic active site is in the middle of the β-propeller. In the tetrameric enzyme, each active site is directed

sideways rather than upwards, an orientation consistent with the enzyme having to cleave off sialic acid from nearby membrane proteins to avoid virus trapping (Figure 7b). The tetramer is approximately $100 \times 100 \times 60$ Å in size with a large hole underneath around the 4-fold axis. The high-mannose type glycan at Asn200 is located in the rim of blade 2 and bridges the right-handed neighboring subunit, contributing to the intersubunit interactions in the tetramer (Figure 7b). A close up view of the *N*-glycan attached at Asn200 and neighboring molecules is shown in Figure 7c. The *N*-glycan lies over blade 5 of a neighboring molecule. A β-Man (Man-3) residue of this glycan is buried in the neighboring molecule. The accessible surface area of this residue is calculated as ~19 Å^2 by AREAIMOL [72]. The crystal structures of N2, N6 and N9 NA in group 2 have been determined and are summarized in Table 2. The *N*-glycan attached at Asn200 in N2 corresponds to those of Asn207 in N6 and Asn200 in N9 and their structures are well superimposed on each other. The distribution of the torsion angles of the side chain of Asn200 is shown in Figure 7d. The $\chi1$ torsion angles mainly assume ~240°, which is normally rare in both glycosylated and non-glycosylated Asn side chain conformers [48] (Figure 7d). Strong interaction between the *N*-glycan and the neighboring subunit may stabilize such a conformation of Asn200. Amino acid sequence alignment reveals the sequon at this site to be conserved among group 2 NAs except for N3 (Figure 7e). Structural comparison between group 2 and other NAs (N1, N4, N8 and type B NAs) reveal no obvious common structural feature [64,73]. Viral proteins are synthesized and secreted by host cells. Thus, the glycans attached on viral proteins are also processed by host glycosyl-hydrolases and -transferases. Since glycan structures of Asn200 are of the high-mannose type, group 2 NA likely forms a tetrameric structure before glycan processing.

Table 2. Summary of group 2 influenza neuraminidase structure.

PDB ID	Glycan structure *	Resolution	Reference
N2 (A/Tokyo/3/1967)			
1nn2	M_4GN_2	2.20	[74]
N2 (A/Tokyo/3/1967)			
1inw	M_3GN_2	2.40	[75]
1inx	M_3GN_2	2.40	-
N6 (A/swine/KU/2/2001)			
1v0z	$M_{3-5}GN_2$	1.84	[76]
1w1x	$M_{1-5}GN_2$	2.00	-
1w20	$M_{1-6}GN_2$	2.08	-
1w21	$M_{1-6}GN_2$	2.08	-
2cml	M_6GN_2	2.15	-
N9 (A/Tern/Australia/G70C/75)			
1iny	M_5GN_2	2.40	[75]
N9 (A/Tern/Australia/G70C/1975 (H11N9))			
1f8b	M_5GN_2	1.80	[77]
1f8c	M_5GN_2	1.70	-
1f8d	M_5GN_2	1.40	-
1f8e	M_5GN_2	1.40	-

Table 2. *Cont.*

PDB ID	Glycan structure *	Resolution	Reference
N9 (A/Tern/Australia/G70C/1975) in complex with single chain Fv fragment			
1a14	M_5GN_2	2.50	[78]
N9 (A/NWS/whale/Maine/1/84)			
2b8h	$M_{7\text{-}8}GN_2$	2.20	[79]

* Glycan structures deposited in each PDB coordinate file are shown. M: D-mannose, GN: *N*-acetyl-D-glucosamine.

3. Immature High-Mannose Type Glycans Contribute to Inter-Subunit and Inter-Domain Interactions

Initial processing of glycoproteins takes the form of deglucosylation of high-mannose type glycan in the ER, and is conserved among all eukaryotes [80]. It is generally considered to be the major event to signal the completion of protein folding. Mono- or di-glucosylated *N*-glycans are rarely observed in mature glycoproteins. However, glucosylated *N*-glycans have been detected in several secreted proteins. Here, we introduce two structures which possess immature glucosylated high-mannose type *N*-glycans on their surfaces. In both cases, the immature glycan extensively interacts with the surface of the protein. Investigation of the co-existence of both unprocessed and processed glycans on a single polypeptide may help to unravel the relationship between protein folding and glycan maturation.

3.1. Monoglucosylated High-Mannose Type Glycan Stabilizes Hexamer Formation of Arylphorin from Antheraea pernyi

Our first example is arylphorin from the Chinese oak silkworm, *Antheraea pernyi* (abbreviated as APA hereafter), which is a hexameric protein of 688 amino acid residues per subunit [81]. It is a hexamerin, a group of proteins belonging to a superfamily that includes arthropod tyrosinase, arthropod hemocyanin, and dipteran arylphorin receptor [82]. The hexamerins show clear structural similarities with the hemocyanins, but have lost the ability to bind copper ions and transport oxygen. They are synthesized in the fat body of a wide range of lepidopteran and dipteran larvae, among other insect orders. These proteins accumulate to high concentrations in the hemolymph. Hexamerins appear to serve as a storage form of amino acids, a resource required for complete development of the adult, since insect pupae do not feed during metamorphosis. In addition to being a storage protein, hexamerins appear to play other important roles during the lifespan of insects. There are at least two types of hexamerins in Lepidoptera: arylphorin and a methionine-rich storage protein [83].

APA has two *N*-glycans attached at Asn196 and Asn344, although APA possesses four possible candidate sequons. Glycosylation of Asn344 is critical for the folding process, whereas glycosylation of Asn196 is not. Mass spectrometric analysis revealed that N344-glycan is a trimmed high-mannose type ($Man_{5\text{-}6}GlcNAc_2$), whereas the N196-glycan remains in a monoglucosylated *N*-glycan ($Glc_1Man_9GlcNAc_2$) state and is resistant to peptide *N*-glycosidase F (PNGaseF) treatment [84]. Although the recombinant N344Q mutant protein is not secreted in culture medium, the Asn196Gln mutant protein is expressed as in wild type and has the same ecdysone-binding activity as wild-type. The crystal structure of APA was solved at 2.42 Å resolution. The overall structure of APA is similar

to that of lobster hemocyanin and is composed of an N-terminal all α-helical fold and a C-terminal β-sandwich like fold (Figure 8a). The *N*-glycan at Asn344 is exposed to solvent and only the chitobiose portion is assigned. In contrast, the *N*-glycan at Asn196 has a clear electron density map and is assigned as a monoglucosylated structure. The asymmetric unit of the APA crystal contains six monomers (one hexamer) as shown in Figure 8b. While all *N*-glycan chains at Asn344 are completely exposed to solvent in the hexamer, the *N*-glycans at Asn196 are buried inside the hexamer and well organized in the deep cleft of the subunit interface. A comparison between monomeric and hexameric APA revealed that the D1 arms of the *N*-glycans are buried inside during hexamer formation. Actually, the accessible surface areas of a D1 arm in monomeric and hexameric APA are 370, and 195 $Å^2$, respectively. The *N*-glycan forms about 20 direct water-mediated hydrogen bonds with adjacent amino acid residues, which are located in the same or different subunits. Typical dihedral ϕ and ψ angles at Manα1-2Man are ~60° and ~150°, respectively. In contrast, the dihedral ϕ and ψ angles at Manα1-2Man in the D2 arm (Man-D2-Man-A) are 279° and 130°, respectively. The D2 arm of the glycan adopts a unique conformation to accommodate the curvature formed by the blue and yellow molecules (Figure 8c). Indeed, the accessible surface of area of α1-2 linked Man in the D2 arm (Man-D2) is dramatically different in monomeric (217 $Å^2$) and hexameric (86 $Å^2$) APA structures. In addition to the inter-subunit disulfide bond between Cys73 and Cys649, extensive intermolecular interactions between *N*-glycan and APA also stabilize the overall trimer-trimer interaction by enhancing interaction between each top and bottom dimer (Figure 8b). Analytical ultracentrifugation and guanidinium chloride unfolding experiments revealed that the presence of the N196-glycan is important for stabilizing the hexameric state and overall stability of APA.

Figure 8. Immature monoglucosylated *N*-glycan on *Antheraea pernyi* arylphorin (PDB code; 3gwj) (**a**) Overall structure of monomeric APA. Protein and carbohydrate are shown in ribbon and rod models, respectively. (**b**) Hexameric structure of APA. Each monomer is shown in surface model. The attached *N*-linked glycans at Asn196 are shown in spheres. (**c**) Close-up view of Asn196-attached *N*-glycan in hexameric APA.
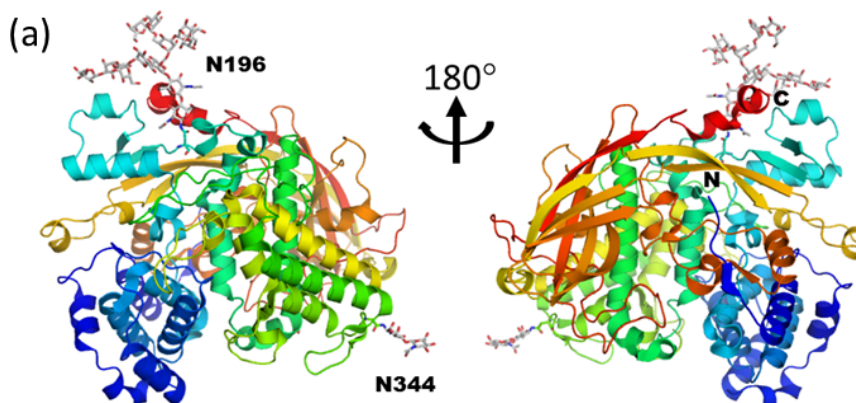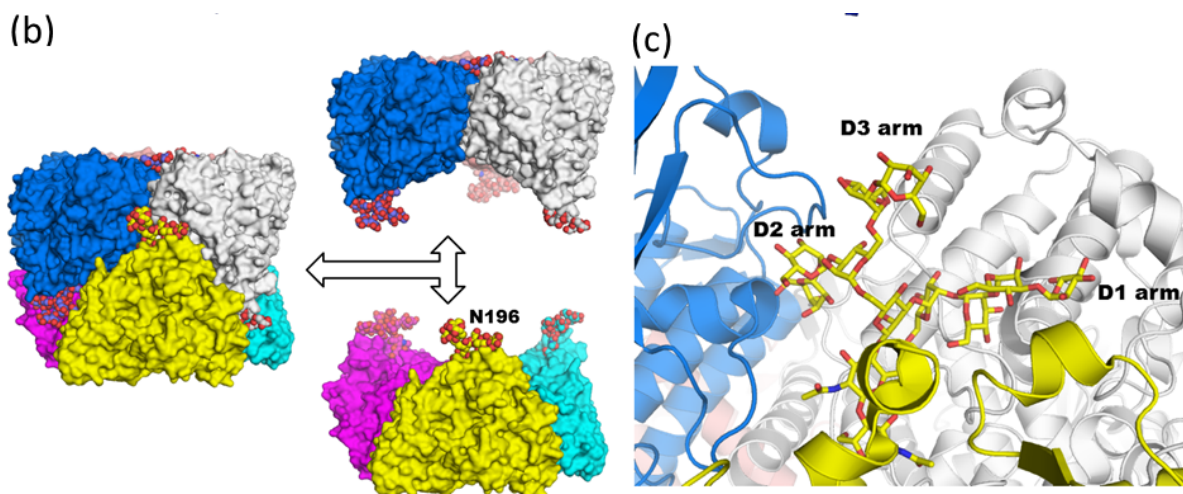
Figure 8. *Cont.*

3.2. Diglucosylated N-*Glycan Stabilizes Inter-Domain Interaction of β-Dalactosidase from* Trichoderma reesei

β-galactosidase is an enzyme (E.C. 3.2.1.23) that catalyzes the hydrolysis of β1-3 or β1-4 linked Gal residues in oligo- and disaccharides, such as lactose, galactobiose, aryl- and alkyl-β-D-galactosides. This enzyme also has the ability to catalyze the reverse reaction of the hydrolysis called transglycosylation. β-galactosidases have been isolated from various sources, such as animals, plants, bacteria, yeasts and fungi. They have many important applications in the industrial and biotechnological fields. In the CAZy database [85], the β-galactosidases occur in the GH 1, 2, 35, and 42 subfamilies.

Trichoderma reesei β-galactosidase (Tr-β-gal) belongs to the GH 35 subfamily. The sequence and the enzymatic properties of this industrially useful enzyme have previously been reported [86,87]. Crystal structures of *Trichoderma reesei* β-galactosidase (Tr-β-gal) in unliganded and ligand complexes were solved at 1.2~1.75 Å resolutions (PDB code; 3og2, 3ogr, 3ogs and 3ogv [88]). The overall structure of Tr-β-gal consists of six domains (Figure 9a). The N-terminal domain forms an eight-stranded α/β barrel structure and is responsible for the catalytic reaction. The subsequent five domains form anti-parallel β-sandwich structures. Tr-β-gal possesses 11 putative *N*-linked glycosylation sites on the surface of the protein [87]. Ten of the 11 sites are exposed (Asn810 is buried). Electron density maps corresponding to the carbohydrate moieties of the *N*-glycans are observed at five of the positions. Two of these (Asn627 and Asn930) contain oligosaccharide chains that represent high-mannose glycan forms. One of these positions (Asn930) has diglucosylated high-mannose type glycan of the form $Glc_2Man_8GlcNAc_2$ (PDB code; 3og2). The glycan at Asn930 is located between three different domains (first, fifth, and sixth domains) and makes a number of hydrogen bonds with the protein surface, stabilizing the structure of Tr-β-gal (Figure 9b). The carbonyl oxygen from Ile955 interacts with the chitobiose core. The side chain of Asp776 bridges β-Man (Man-3) and Glc residues. The glycan at Asn930 is also connected to the catalytic domain because OD1 and OD2 from Asp265 tightly interact with an α1-2 linked Man (Man-C). This glycan covers several hydrophobic and aromatic amino acid residues and may protect the structure from proteolysis. The β-Man (Man-3) is the most buried residue and reaches to the protein cavity (accessible surface area only ~36 Å^2). The

glucose units of Tr-β-gal reach very close to the catalytic site. Thus, it is plausible that glycosylation affects the catalytic properties of the enzyme.

Recombinant Tr-β-gal was overexpressed by a *Trichoderma reesei* expression system [89]. In fungal expression systems, Glc residues are expected to be trimmed by glucosidases I (GLS-I) and II (GLS-II). Nevertheless the *N*-glycan attached at Asn930 in Tr-β-gal is di-glucosylated and interacts with residues far apart in the primary sequence. It suggests that the protein folding of Tr-β-gal might be finished before GLS-II encounters the terminal Glc residues on the glycan. The tight association between the D1 arm and the protein surface likely prevents access of the glucose units to the catalytic site of GLS-II, inhibiting deglucosylation and further conversion to complex-type glycans.

Figure 9. Immature diglucosylated *N*-glycan on β-galactosidase from *Trichoderma reesei* (**a**) Overall structure of Tr-β-gal (PDB code; 3ogv). *N*-linked glycan at Asn627 and Asn930 are shown in sphere models. The region delineated in black dotted lines is magnified in (b). (**b**) Close-up view of Asn930 attached glycan. Hydrogen bonds are shown as red dotted lines.
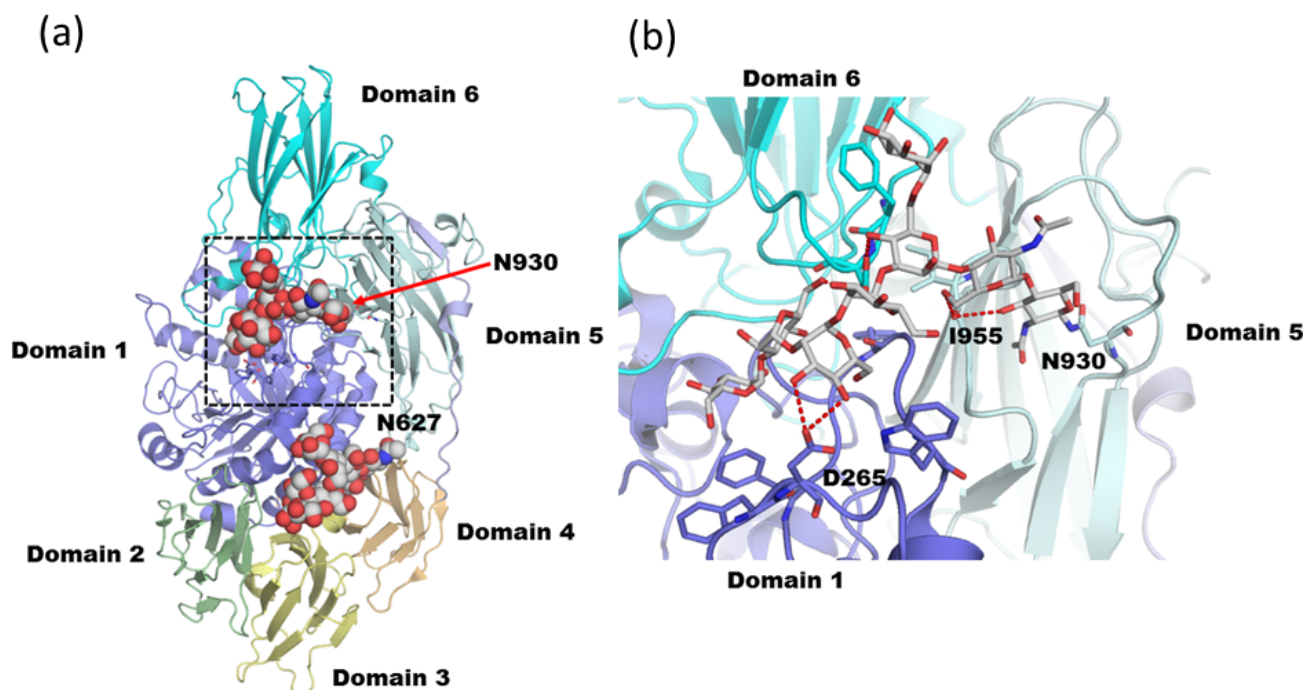
4. What Is the Function of Mobile/Disordered *N*-Glycans?

Cell surface membrane proteins, such as cell surface receptors, are often glycosylated. Crystal structures of these cell surface glycoproteins revealed that their *N*-glycans are usually mobile and most of the sugars disordered. Although surface *N*-glycans are sometimes immobilized by symmetry-related molecules in crystal packing, only the first one or two GlcNAc residues are usually ordered enough to be traced in the electron density map. The alteration of glycoform is associated with various physiological and pathological events, including tumor invasion [90]. These disordered or "missing" glycans are therefore considered to assume certain physiological functions as with the highly-ordered *N*-glycans. In this chapter, we introduce several structures of cell surface glycoproteins which play a role in the immune system and cell-cell adhesion.

Infectious diseases caused by various pathogens account for about one-third of all human deaths in the world, more than all forms of cancer combined [91]. To fight against these powerful pathogens, vertebrates use two types of immune defense carried out by specialized proteins and cells; the innate immune response and the adaptive immune response. Innate immunity is based on an ancient and ubiquitous system of cells and molecules that defend the host against infection. This system can recognize virtually all microbes using a limited repertoire of germ-line-encoded receptors that recognize broadly conserved components of bacterial and fungal cell walls or genetic material, such as double-stranded viral RNA (dsRNA) [92,93]. Toll-like receptors (TLRs) are the most important sensors in the innate immune system [94]. Ten human TLRs (TLR1~10) which specifically recognize pathogen-associated molecules have been identified. Human TLR3 is activated by dsRNA associated with viral infection, endogenous cellular mRNA, and sequence-independent small interfering RNAs. The human TLR3 ectodomain is a large horseshoe-shaped solenoid-like structure assembled from 23 leucine-rich repeats (PDB code; 1ziw, [95]). Human TLR3 ectodomain possesses 15 potential *N*-glycosylation sites. Due to the poor electron density of the carbohydrate moieties, only one or two GlcNAc residues are assigned at eight of the *N*-glycosylation sites. A putative fully glycosylated model structure reveals that almost all the surface of the TLR3 molecule is covered by carbohydrate, but one face is glycosylation-free. The structure of mouse TLR3 ectodomain in complex with dsRNA demonstrates that dsRNA contacts occur through residues on the glycosylation-free surface (PDB code; 3ciy, [96]). The glycosylation sites of both human and mouse TLR3 are almost identical. Compared with human TLR3, the electron density of carbohydrate is clearly observed in mouse TLR3. *N*-glycan at Asn413, located in the vicinity of dsRNA, is assigned as $Man_3GlcNAc_2$ and an α1-6 linked Man residue interacts with the sugar-phosphate backbone of dsRNA. As in the case of human TLR3, the location of *N*-glycan is thought to be useful in order to restrict the preferential orientation of ligand (Figure 10a).

As for adaptive immune systems, most of the cell surface receptors which are involved in antigen recognition by T cells and in the orchestration of the subsequent cell signaling events are glycosylated. Rudd *et al.* postulated the *N*-glycans on the protein surface play a wide range of roles, such as controlling the assembly and stabilization of the protein complexes in the adaptive immune system [97]. *N*-linked glycans attached on the membrane proximal domains of CD2 (PDB code; 1hnf, [98]) and CD48 (PDB code; 2dru, [99]) are distributed so as to provide a scaffold to orient the binding faces, which leads to increased apparent affinity (Figure 10b). Moreover, the glycans on T-cell receptors (TCR) are located over the protein surface in such a way that they could prevent non-specific aggregation. Another important point to be emphasized is that *N*-glycans limit the possible geometry and spacing of TCR/major histocompatibility complex (MHC) clusters which precede cell signaling.

Cells produce, organize and degrade extracellular matrix. The matrix in turn exerts a powerful influence on the cells, mainly through "matrix receptors". Integrins are the principle matrix receptors on animal cells and transmit bidirectional signals across the plasma membrane, thereby linking the extracellular environment to the internal actin cytoskeleton [100]. All integrins are non-covalently linked heterodimeric molecules consisting of one α and one β subunit, which together create a binding site for specific extracellular ligands on that part of the protein furthest from the membrane. α5β1 integrin is a major cellular receptor for the extracellular matrix protein fibronectin and plays a fundamental role during mammalian development. Fibronectin is a principle component of the

extracellular matrix and is a modular protein composed of homologous repeats of small domains with an elongated shape arranged as "beads on a string". α5β1 integrin interacts with fibronectin through Arg-Gly-Asp (RGD) sequences present in a flexible loop region in the middle of the protein. A crystal structure of the α5β1 integrin ectodomain shows the fibronectin-binding pocket surrounded by four *N*-glycans (two in α5 and the other two in β1), forming a trench-like exposed surface along the subunit interface. This topography and location of the *N*-glycans presumably limits the choice of docking orientations when the elongated fibronectin molecule tries to make close contact (Docking model based on α5β1 integrin ectodomain (PDB code; 3vi4) and fibronectin fragment (PDB code; 2mfn and 1fnf), [101,102] (Figure 10c).

Figure 10. Highly flexible *N*-glycans on cell surface receptors. Complex-type *N*-glycans ($GlcNAc_2Man_3GlcNAc_2Fuc$) are superimposed, based on the position of chitobiose or sequons by using LSQKAB [103]. (**a**) Fully glycosylated Toll-like receptor-3 (TLR3) ectodomain in complex with dsRNA (PDB code; 3ciy). Protein molecules are shown as green and cyan surface models. dsRNA is shown as a gray sphere. (**b**) Extracellular domains of CD2 (PDB code; 1hnf) and CD48 (PDB code; 2dru). (**c**) Crystal structure of α5β1 integrin ectodomain (PDB code; 3vi4) and fibronectin FN7-10 fragment (PDB code; 1fnf). In the fibronectin structure, the amino acid residues which interact with α5β1 integrin are shown in red stick model. Dashed lines on α5β1 integrin outline the shallow groove formed by *N*-glycans. (**d**) Intercellular cell adhesion molecule (ICAM)-2 ectodomains (PDB code; 1zxq).
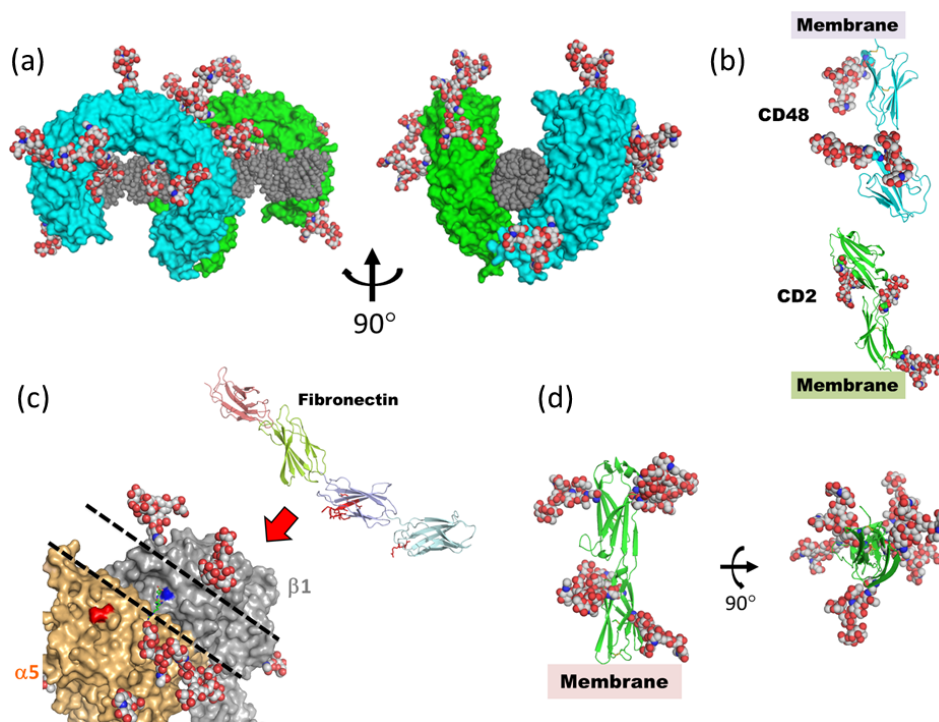
The intercellular adhesion molecules 1~3 (ICAM-1~3) are members of the immunoglobulin superfamily (IgSF) and are known ligands for integrins on the surface of cells. ICAM-2, a ligand for αLβ2 integrin, is composed of two *N*-terminal extracellular Ig domains (which share 35% sequence homology with ICAM-1), a transmembrane domain, and a short cytoplasmic tail. The Ig domain, the

characteristic building block of IgSF, consists of two antiparallel β-sheets packed tightly against each other and linked by a disulfide bond. The crystal structure of the two extracellular Ig domains of ICAM-2 (PDB code: 1zxq) shows that they adopt a hockey stick shape. Six *N*-glycans are exposed to solvent and assigned as chitobiose portions. When one looks down the axis of the two Ig domains, these *N*-glycans are seen as uniformly distributed around the perimeter of the domains [104] (Figure 10d), and are well placed to prevent non-specific protein aggregation.

These examples illustrate the important role of glycan in assisting the docking of large ligands, both by orienting the ligand and by spacing the receptor through inhibiting aggregation. It has been considered that a principle role of the mobile carbohydrate attached on cell surface receptors is to increase protein conformational stability [105]. Thus solution analysis of RNaseB revealed that the glycan has a significant stabilizing effect on the protein structure by decreasing the flexibility of the protein backbone both near to and distant from the glycan attachment site [106]. These secondary functions of glycans, supplementary to actual binding interactions, seem to be fundamental to the role of the carbohydrates in glycoproteins on cell surfaces.

5. Future Perspective

Several structural and functional aspects of glycosylation, in terms of intra- and inter- molecular interactions, are provided by available crystal structures. An emphasis on glycoprotein-oriented structural biology will further the understanding of glycan function. Crystallographic analysis of glycoproteins will require a more thorough investigation of the glycoprotein expression system and of glycan structure. Although methodologies for glycoprotein production are advancing [107,108], improvements are required in the expression and selection of a particular glycoform of a target protein. The glycan sequence at each glycosylation site can be analyzed in advance by mass spectrometry, coupled with liquid chromatography. However, even with this sequence information, the electron density of the glycan must be interpreted with great caution. Backbone flexibility dictates that structural information on the glycan is largely missing in X-ray diffraction data. Other techniques, such as molecular dynamic simulation and NMR, are needed to help join the discontinuous snapshots derived from X-ray studies and to evaluate the contribution of the glycan moieties to molecular fluctuations.

The glycoform at a particular site on a protein is often closely related to physiological function and is tightly regulated [18,46,47,109]. However, in many cases the structural relationship between glycoform and function still remains unclear. Recent technical advances in the chemical and enzymatic syntheses of homogeneous glycoproteins are going to make further valuable contributions to the study of glycoform-function relationships [110–112]. A survey of current PDB data indicates that the glycan structures of available glycoproteins are mainly of the high-mannose type, and may be biased due to the expression system used. Thus the relationship between glycoform function and its 3D structure needs careful investigation. Structural biology focusing on glycoform-specific functions is the challenge for the near future.

References

1. Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science* **2001**, *291*, 2364–2369.

2. Apweiler, R.; Hermjakob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1999**, *1473*, 4–8.

3. Nishikawa, A.; Ihara, Y.; Hatakeyama, M.; Kangawa, K.; Taniguchi, N. Purification, cDNA cloning, and expression of UDP-*N*-acetylglucosamine:β-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase III from rat kidney. *J. Biol. Chem.* **1992**, *267*, 18199–18204.

4. Cummings, R.D.; Trowbridge, I.S.; Kornfeld, S. A mouse lymphoma cell line resistant to the leukoagglutinating lectin from *Phaseolus vulgaris* is deficient in UDP-GlcNAc:α-D-mannoside-β-1,6-*N*-acetylglucosaminyltransferase. *J. Biol. Chem.* **1982**, *257*, 13421–13427.

5. Uozumi, N.; Yanagidani, S.; Miyoshi, E.; Ihara, Y.; Sakuma, T.; Gao, C.X.; Teshima, T.; Fujii, S.; Shiba, T.; Taniguchi, N. Purification and cDNA cloning of porcine brain GDP-L-Fuc:*N*-acetyl-β-D-glucosaminide α1→6 fucosyltransferase. *J. Biol. Chem.* **1996**, *271*, 27810–27817.

6. Wyss, D.F.; Wagner, G. The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.* **1996**, *7*, 409–416.

7. Kwong, P.D.; Wyatt, R.; Desjardins, E.; Robinson, J.; Culp, J.S.; Hellmig, B.D.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A. Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). *J. Biol. Chem.* **1999**, *274*, 4115–4123.

8. Nakata, Z.; Nagae, M.; Yasui, N.; Bujo, H.; Nogi, T.; Takagi, J. Crystallization and preliminary crystallographic analysis of human LR11 Vps10p domain. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2011**, *67*, 129–132.

9. Petrescu, A.J.; Milac, A.L.; Petrescu, S.M.; Dwek, R.A.; Wormald, M.R. Statistical analysis of the protein environment of *N*-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology* **2004**, *14*, 103–114.

10. Stanley, P. Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell. Biol.* **1989**, *9*, 377–383.

11. Reeves, P.J.; Callewaert, N.; Contreras, R.; Khorana, H.G. Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous *N*-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13419–13424.

12. Chang, V.T.; Crispin, M.; Aricescu, A.R.; Harvey, D.J.; Nettleship, J.E.; Fennelly, J.A.; Yu, C.; Boles, K.S.; Evans, E.J.; Stuart, D.I.; *et al.* Glycoprotein structural genomics: Solving the glycosylation problem. *Structure* **2007**, *15*, 267–273.

13. Petrescu, A.J.; Petrescu, S.M.; Dwek, R.A.; Wormald, M.R. A statistical analysis of N- and O-glycan linkage conformations from crystallographic data. *Glycobiology* **1999**, *9*, 343–352.

14. Wormald, M.R.; Petrescu, A.J.; Pao, Y.L.; Glithero, A.; Elliott, T.; Dwek, R.A. Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem. Rev.* **2002**, *102*, 371–386.

15. Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C.W. GLYCOSCIENCES.de: An Internet portal to support glycomics and glycobiology research. *Glycobiology* **2006**, *16*, 71R–81R.

16. Nakahara, T.; Hashimoto, R.; Nakagawa, H.; Monde, K.; Miura, N.; Nishimura, S. Glycoconjugate Data Bank: Structures—An annotated glycan structure database and *N*-glycan primary structure verification service. *Nucleic Acids Res.* **2008**, *36*, D368–371.

17. Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. *Biochemistry* **1981**, *20*, 2361–2370.

18. Martin, W.L.; West, A.P., Jr.; Gan, L.; Bjorkman, P.J. Crystal structure at 2.8 Å of an FcRn/heterodimeric Fc complex: Mechanism of pH-dependent binding. *Mol. Cell* **2001**, *7*, 867–877.

19. Harris, L.J.; Larson, S.B.; Hasel, K.W.; McPherson, A. Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **1997**, *36*, 1581–1597.

20. Saphire, E.O.; Stanfield, R.L.; Crispin, M.D.; Parren, P.W.; Rudd, P.M.; Dwek, R.A.; Burton, D.R.; Wilson, I.A. Contrasting IgG structures reveal extreme asymmetry and flexibility. *J. Mol. Biol.* **2002**, *319*, 9–18.

21. Dwek, R.A.; Lellouch, A.C.; Wormald, M.R. Glycobiology: "The function of sugar in the IgG molecule". *J. Anat.* **1995**, *187 (Pt 2)*, 279–292.

22. Shields, R.L.; Lai, J.; Keck, R.; O'Connell, L.Y.; Hong, K.; Meng, Y.G.; Weikert, S.H.; Presta, L.G. Lack of fucose on human IgG1 *N*-linked oligosaccharide improves binding to human Fcγ RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **2002**, *277*, 26733–26740.

23. Nimmerjahn, F.; Ravetch, J.V. Divergent immunoglobulin G subclass activity through selective Fc receptor binding. *Science* **2005**, *310*, 1510–1512.

24. Parekh, R.B.; Dwek, R.A.; Sutton, B.J.; Fernandes, D.L.; Leung, A.; Stanworth, D.; Rademacher, T.W.; Mizuochi, T.; Taniguchi, T.; Matsuta, K.; *et al.* Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* **1985**, *316*, 452–457.

25. Samuelsson, A.; Towers, T.L.; Ravetch, J.V. Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. *Science* **2001**, *291*, 484–486.

26. Umaña, P.; Jean-Mairet, J.; Moudry, R.; Amstutz, H.; Bailey, J.E. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat. Biotechnol.* **1999**, *17*, 176–180.

27. Shinkawa, T.; Nakamura, K.; Yamane, N.; Shoji-Hosaka, E.; Kanda, Y.; Sakurada, M.; Uchida, K.; Anazawa, H.; Satoh, M.; Yamasaki, M.; *et al.* The absence of fucose but not the presence of galactose or bisecting *N*-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* **2003**, *278*, 3466–3473.

28. Sutton, B.J.; Phillips, D.C. The three-dimensional structure of the carbohydrate within the Fc fragment of immunoglobulin G. *Biochem. Soc. Trans.* **1983**, *11*, 130–132.

29. Matsumiya, S.; Yamaguchi, Y.; Saito, J.; Nagano, M.; Sasakawa, H.; Otaki, S.; Satoh, M.; Shitara, K.; Kato, K. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J. Mol. Biol.* **2007**, *368*, 767–779.

30. Crispin, M.; Bowden, T.A.; Coles, C.H.; Harlos, K.; Aricescu, A.R.; Harvey, D.J.; Stuart, D.I.; Jones, E.Y. Carbohydrate and domain architecture of an immature antibody glycoform exhibiting enhanced effector functions. *J. Mol. Biol.* **2009**, *387*, 1061–1066.

31. Krapp, S.; Mimura, Y.; Jefferis, R.; Huber, R.; Sondermann, P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* **2003**, *325*, 979–989.

32. Mimura, Y.; Sondermann, P.; Ghirlando, R.; Lund, J.; Young, S.P.; Goodall, M.; Jefferis, R. Role of oligosaccharide residues of IgG1-Fc in FcγRIIb binding. *J. Biol. Chem.* **2001**, *276*, 45539–45547.

33. Oganesyan, V.; Damschroder, M.M.; Woods, R.M.; Cook, K.E.; Wu, H.; Dall'acqua, W.F. Structural characterization of a human Fc fragment engineered for extended serum half-life. *Mol. Immunol.* **2009**, *46*, 1750–1755.

34. Oganesyan, V.; Gao, C.; Shirinian, L.; Wu, H.; Dall'Acqua, W.F. Structural characterization of a human Fc fragment engineered for lack of effector functions. *Acta Crystallogr. D Biol. Crystallogr.* **2008**, *64*, 700–704.

35. Oganesyan, V.; Damschroder, M.M.; Leach, W.; Wu, H.; Dall'Acqua, W.F. Structural characterization of a mutated, ADCC-enhanced human Fc fragment. *Mol. Immunol.* **2008**, *45*, 1872–1882.

36. DeLano, W.L.; Ultsch, M.H.; de Vos, A.M.; Wells, J.A. Convergent solutions to binding at a protein-protein interface. *Science* **2000**, *287*, 1279–1283.

37. Idusogie, E.E.; Presta, L.G.; Gazzano-Santoro, H.; Totpal, K.; Wong, P.Y.; Ultsch, M.; Meng, Y.G.; Mulkerrin, M.G. Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc. *J. Immunol.* **2000**, *164*, 4178–4184.

38. Branden, B.C. Department of Natural Sciences, Bowie State University, Bowie, MD, USA. Unpublished Work, 2009.

39. Saphire, E.O.; Parren, P.W.; Pantophlet, R.; Zwick, M.B.; Morris, G.M.; Rudd, P.M.; Dwek, R.A.; Stanfield, R.L.; Burton, D.R.; Wilson, I.A. Crystal structure of a neutralizing human IgG against HIV-1: A template for vaccine design. *Science* **2001**, *293*, 1155–1159.

40. Moiani, D.; Salvalaglio, M.; Cavallotti, C.; Bujacz, A.; Redzynia, I.; Bujacz, G.; Dinon, F.; Pengo, P.; Fassina, G. Structural characterization of a Protein A mimetic peptide dendrimer bound to human IgG. *J. Phys. Chem. B* **2009**, *113*, 16268–16275.

41. Kolenko, P.; Dohnálek, J.; Dušková, J.; Skálová, T.; Collard, R.; Hašek, J. New insights into intra- and intermolecular interactions of immunoglobulins: Crystal structure of mouse IgG2b-Fc at 2.1-Å resolution. *Immunology* **2009**, *126*, 378–385.

42. Girardi, E.; Holdom, M.D.; Davies, A.M.; Sutton, B.J.; Beavil, A.J. The crystal structure of rabbit IgG-Fc. *Biochem. J.* **2009**, *417*, 77–83.

43. Raju, T.S.; Mulkerrin, M.G.; Parker, M.; de Vos, A.M.; Gazzano-Santoro, H.; Totpal, K.; Ultsch, M.H. Department of Protein Engineering, Genentech Inc. South San Francisco, CA, USA. Unpublished Work, 2003.

44. Radaev, S.; Motyka, S.; Fridman, W.H.; Sautes-Fridman, C.; Sun, P.D. The structure of a human type III Fcγ receptor in complex with Fc. *J. Biol. Chem.* **2001**, *276*, 16469–16477.

45. Sondermann, P.; Huber, R.; Oosthuizen, V.; Jacob, U. The 3.2-Å crystal structure of the human IgG1 Fc fragment-FcγRIII complex. *Nature* **2000**, *406*, 267–273.

46. Ferrara, C.; Grau, S.; Jäger, C.; Sondermann, P.; Brünker, P.; Waldhauer, I.; Hennig, M.; Ruf, A.; Rufer, A.C.; Stihle, M.; *et al.* Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcγRIII and antibodies lacking core fucose. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12669–12674.

47. Mizushima, T.; Yagi, H.; Takemoto, E.; Shibata-Koyama, M.; Isoda, Y.; Iida, S.; Masuda, K.; Satoh, M.; Kato, K. Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycans. *Genes Cells* **2011**, *16*, 1071–1080.

48. Lütteke, T. Analysis and validation of carbohydrate three-dimensional structures. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 156–168.

49. Wormald, M.R.; Rudd, P.M.; Harvey, D.J.; Chang, S.C.; Scragg, I.G.; Dwek, R.A. Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* **1997**, *36*, 1370–1380.

50. Yamaguchi, Y.; Kato, K.; Shindo, M.; Aoki, S.; Furusho, K.; Koga, K.; Takahashi, N.; Arata, Y.; Shimada, I. Dynamics of the carbohydrate chains attached to the Fc portion of immunoglobulin G as studied by NMR spectroscopy assisted by selective ^{13}C labeling of the glycans. *J. Biomol. NMR* **1998**, *12*, 385–394.

51. Barb, A.W.; Prestegard, J.H. NMR analysis demonstrates immunoglobulin G *N*-glycans are accessible and dynamic. *Nat. Chem. Biol.* **2011**, *7*, 147–153.

52. Lovell, S.C.; Davis, I.W.; Arendall, W.B., III; de Bakker, P.I.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Cα geometry: φ, φ and Cβ deviation. *Proteins* **2003**, *50*, 437–450.

53. Kolenko, P.; Skálová, T.; Dohnálek, J.; Hašek, J. L-fucose in crystal structures of IgG-Fc: Reinterpretation of experimental data. *Collect. Czechoslov. Chem. Commun.* **2008**, *73*, 608–615.

54. Yamaguchi, Y.; Nishimura, M.; Nagano, M.; Yagi, H.; Sasakawa, H.; Uchida, K.; Shitara, K.; Kato, K. Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. *Biochim. Biophys. Acta* **2006**, *1760*, 693–700.

55. Kanda, Y.; Yamada, T.; Mori, K.; Okazaki, A.; Inoue, M.; Kitajima-Miyama, K.; Kuni-Kamochi, R.; Nakano, R.; Yano, K.; Kakita, S.; *et al.* Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different *N*-linked Fc oligosaccharides: The high-mannose, hybrid, and complex types. *Glycobiology* **2007**, *17*, 104–118.

56. Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2256–2268.

57. Burmeister, W.P.; Huber, A.H.; Bjorkman, P.J. Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* **1994**, *372*, 379–383.

58. Martin, W.L.; Bjorkman, P.J. Characterization of the 2:1 complex between the class I MHC-related Fc receptor and its Fc ligand in solution. *Biochemistry* **1999**, *38*, 12639–12647.

Int. J. Mol. Sci. **2012**, *13*   **8426**

59. Sánchez, L.M.; Penny, D.M.; Bjorkman, P.J. Stoichiometry of the interaction between the major histocompatibility complex-related Fc receptor and its Fc ligand. *Biochemistry* **1999**, *38*, 9471–9476.

60. Clynes, R.A.; Towers, T.L.; Presta, L.G.; Ravetch, J.V. Inhibitory Fc receptors modulate *in vivo* cytotoxicity against tumor targets. *Nat. Med.* **2000**, *6*, 443–446.

61. Cartron, G.; Dacheux, L.; Salles, G.; Solal-Celigny, P.; Bardos, P.; Colombat, P.; Watier, H. Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcγRIIIa gene. *Blood* **2002**, *99*, 754–758.

62. Liu, C.; Eichelberger, M.C.; Compans, R.W.; Air, G.M. Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. *J. Virol.* **1995**, *69*, 1099–1106.

63. A revision of the system of nomenclature for influenza viruses: A WHO memorandum. *Bull. World Health Organ.* **1980**, *58*, 585–591.

64. Russell, R.J.; Haire, L.F.; Stevens, D.J.; Collins, P.J.; Lin, Y.P.; Blackburn, G.M.; Hay, A.J.; Gamblin, S.J.; Skehel, J.J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* **2006**, *443*, 45–49.

65. Blok, J.; Air, G.M.; Laver, W.G.; Ward, C.W.; Lilley, G.G.; Woods, E.F.; Roxburgh, C.M.; Inglis, A.S. Studies on the size, chemical composition, and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N-terminal region of the NA is not processed and serves to anchor the NA in the viral membrane. *Virology* **1982**, *119*, 109–121.

66. Air, G.M.; Laver, W.G. The neuraminidase of influenza virus. *Proteins* **1989**, *6*, 341–356.

67. Murti, K.G.; Webster, R.G. Distribution of hemagglutinin and neuraminidase on influenza virions as revealed by immunoelectron microscopy. *Virology* **1986**, *149*, 36–43.

68. Von Itzstein, M.; Wu, W.Y.; Kok, G.B.; Pegg, M.S.; Dyason, J.C.; Jin, B.; van Phan, T.; Smythe, M.L.; White, H.F.; Oliver, S.W.; *et al.* Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **1993**, *363*, 418–423.

69. Kim, C.U.; Lew, W.; Williams, M.A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M.S.; Mendel, D.B.; Tai, C.Y.; *et al.* Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681–690.

70. Varghese, J.N.; Laver, W.G.; Colman, P.M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **1983**, *303*, 35–40.

71. Baker, A.T.; Varghese, J.N.; Laver, W.G.; Air, G.M.; Colman, P.M. Three-dimensional structure of neuraminidase of subtype N9 from an avian influenza virus. *Proteins* **1987**, *2*, 111–117.

72. Lee, B.; Richards, F.M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379–400.

73. Burmeister, W.P.; Ruigrok, R.W.; Cusack, S. The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid. *EMBO J.* **1992**, *11*, 49–56.

74. Varghese, J.N.; Colman, P.M. Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. *J. Mol. Biol.* **1991**, *221*, 473–486.

75. White, C.L.; Janakiraman, M.N.; Laver, W.G.; Philippon, C.; Vasella, A.; Air, G.M.; Luo, M. A sialic acid-derived phosphonate analog inhibits different strains of influenza virus neuraminidase with different efficiencies. *J. Mol. Biol.* **1995**, *245*, 623–634.

76. Rudino-Pinera, E.; Tunnah, P.; Crennell, S.J.; Webster, R.G.; Laver, W.G.; Garman, E.F. Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK. Unpublished Work, 2006.

77. Smith, B.J.; Colman, P.M.; von Itzstein, M.; Danylec, B.; Varghese, J.N. Analysis of inhibitor binding in influenza virus neuraminidase. *Protein Sci.* **2001**, *10*, 689–696.

78. Malby, R.L.; McCoy, A.J.; Kortt, A.A.; Hudson, P.J.; Colman, P.M. Three-dimensional structures of single-chain Fv-neuraminidase complexes. *J. Mol. Biol.* **1998**, *279*, 901–910.

79. Smith, B.J.; Huyton, T.; Joosten, R.P.; McKimm-Breschkin, J.L.; Zhang, J.G.; Luo, C.S.; Lou, M.Z.; Labrou, N.E.; Garrett, T.P. Structure of a calcium-deficient form of influenza virus neuraminidase: Implications for substrate binding. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 947–952.

80. Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E. *Essentials of Glycobiology*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2009.

81. Ryu, K.S.; Lee, J.O.; Kwon, T.H.; Choi, H.H.; Park, H.S.; Hwang, S.K.; Lee, Z.W.; Lee, K.B.; Han, Y.H.; Choi, Y.S.; *et al.* The presence of monoglucosylated N196-glycan is important for the structural stability of storage protein, arylphorin. *Biochem. J.* **2009**, *421*, 87–96.

82. Burmester, T.; Scheller, K. Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and dipteran arylphorin receptor. *J. Mol. Evol.* **1996**, *42*, 713–728.

83. Telfer, W.H.; Kunkel, J.G. The function and evolution of insect storage hexamers. *Annu. Rev. Entomol.* **1991**, *36*, 205–228.

84. Jung, H.I.; Kim, Y.H.; Kim, S. Structural basis for the presence of a monoglucosylated oligosaccharide in mature glycoproteins. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 100–106.

85. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–238.

86. Seiboth, B.; Hartl, L.; Salovuori, N.; Lanthaler, K.; Robson, G.D.; Vehmaanperä, J.; Penttilä, M.E.; Kubicek, C.P. Role of the *bga1*-encoded extracellular β-galactosidase of *Hypocrea jecorina* in cellulase induction by lactose. *Appl. Environ. Microbiol.* **2005**, *71*, 851–857.

87. Gamauf, C.; Marchetti, M.; Kallio, J.; Puranen, T.; Vehmaanperä, J.; Allmaier, G.; Kubicek, C.P.; Seiboth, B. Characterization of the *bga1*-encoded glycoside hydrolase family 35 β-galactosidase of *Hypocrea jecorina* with galacto-β-D-galactanase activity. *FEBS J.* **2007**, *274*, 1691–1700.

88. Maksimainen, M.; Hakulinen, N.; Kallio, J.M.; Timoharju, T.; Turunen, O.; Rouvinen, J. Crystal structures of *Trichoderma reesei* β-galactosidase reveal conformational changes in the active site. *J. Struct. Biol.* **2011**, *174*, 156–163.

89. Maksimainen, M.; Timoharju, T.; Kallio, J.M.; Hakulinen, N.; Turunen, O.; Rouvinen, J. Crystallization and preliminary diffraction analysis of a β-galactosidase from *Trichoderma reesei*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2009**, *65*, 767–769.

90. Hakomori, S. Glycosylation defining cancer malignancy: New wine in an old bottle. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10231–10233.

91. *Infectious Disease Movement in a Borderless World: Workshop Summary*, 2010/10/15 ed.; National Academies Press (US): Washington, DC, USA, 2010.

92. Janeway, C.A., Jr.; Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **2002**, *20*, 197–216.

93. Beutler, B. Innate immunity: An overview. *Mol. Immunol.* **2004**, *40*, 845–859.

94. Takeda, K.; Akira, S. Toll-like receptors in innate immunity. *Int. Immunol.* **2005**, *17*, 1–14.

95. Choe, J.; Kelker, M.S.; Wilson, I.A. Crystal structure of human Toll-like receptor 3 (TLR3) ectodomain. *Science* **2005**, *309*, 581–585.

96. Liu, L.; Botos, I.; Wang, Y.; Leonard, J.N.; Shiloach, J.; Segal, D.M.; Davies, D.R. Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **2008**, *320*, 379–381.

97. Rudd, P.M.; Wormald, M.R.; Stanfield, R.L.; Huang, M.; Mattsson, N.; Speir, J.A.; DiGennaro, J.A.; Fetrow, J.S.; Dwek, R.A.; Wilson, I.A. Roles for glycosylation of cell surface receptors involved in cellular immune recognition. *J. Mol. Biol.* **1999**, *293*, 351–366.

98. Bodian, D.L.; Jones, E.Y.; Harlos, K.; Stuart, D.I.; Davis, S.J. Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 Å resolution. *Structure* **1994**, *2*, 755–766.

99. Evans, E.J.; Castro, M.A.; O'Brien, R.; Kearney, A.; Walsh, H.; Sparks, L.M.; Tucknott, M.G.; Davies, E.A.; Carmo, A.M.; van der Merwe, P.A.; *et al.* Crystal structure and binding properties of the CD2 and CD244 (2B4)-binding protein, CD48. *J. Biol. Chem.* **2006**, *281*, 29309–29320.

100. Hynes, R.O. Integrins: Bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673–687.

101. Nagae, M.; Re, S.; Mihara, E.; Nogi, T.; Sugita, Y.; Takagi, J. Crystal structure of α5β1 integrin ectodomain: Atomic details of the fibronectin receptor. *J. Cell. Biol.* **2012**, *197*, 131–140.

102. Leahy, D.J.; Aukhil, I.; Erickson, H.P. 2.0 Å crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell* **1996**, *84*, 155–164.

103. Kabsch, W. Solution for Best Rotation to Relate 2 Sets of Vectors. *Acta Crystallogr. Sect. A* **1976**, *32*, 922–923.

104. Casasnovas, J.M.; Springer, T.A.; Liu, J.H.; Harrison, S.C.; Wang, J.H. Crystal structure of ICAM-2 reveals a distinctive integrin recognition surface. *Nature* **1997**, *387*, 312–315.

105. Wormald, M.R.; Dwek, R.A. Glycoproteins: Glycan presentation and protein-fold stability. *Structure* **1999**, *7*, R155–160.

106. Joao, H.C.; Scragg, I.G.; Dwek, R.A. Effects of glycosylation on protein conformation and amide proton exchange rates in RNase B. *FEBS Lett.* **1992**, *307*, 343–346.

107. Valderrama-Rincon, J.D.; Fisher, A.C.; Merritt, J.H.; Fan, Y.Y.; Reading, C.A.; Chhiba, K.; Heiss, C.; Azadi, P.; Aebi, M.; Delisa, M.P. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat. Chem. Biol.* **2012**, *8*, 434–436.

108. Mikami, S.; Kobayashi, T.; Yokoyama, S.; Imataka, H. A hybridoma-based *in vitro* translation system that efficiently synthesizes glycoproteins. *J. Biotechnol.* **2006**, *127*, 65–78.

109. Sanchez, E.J.; Lewis, K.M.; Munske, G.R.; Nissen, M.S.; Kang, C. Glycosylation of skeletal calsequestrin: Implications for its function. *J. Biol. Chem.* **2012**, *287*, 3042–3050.

110. Zou, G.; Ochiai, H.; Huang, W.; Yang, Q.; Li, C.; Wang, L.X. Chemoenzymatic synthesis and Fcγ receptor binding of homogeneous glycoforms of antibody Fc domain. Presence of a bisecting sugar moiety enhances the affinity of Fc to FcγIIIa receptor. *J. Am. Chem. Soc.* **2011**, *133*, 18975–18991.

111. Murakami, M.; Okamoto, R.; Izumi, M.; Kajihara, Y. Chemical synthesis of an erythropoietin glycoform containing a complex-type disialyloligosaccharide. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 3567–3572.

112. Aussedat, B.; Fasching, B.; Johnston, E.; Sane, N.; Nagorny, P.; Danishefsky, S.J. Total synthesis of the α-subunit of human glycoprotein hormones: Toward fully synthetic homogeneous human follicle-stimulating hormone. *J. Am. Chem. Soc.* **2012**, *134*, 3532–3541.