Glycobiology vol. 24 no. 1 pp. 17–25, 2014
doi:10.1093/glycob/cwt083
Advance Access publication on September 20, 2013

Presentation, presentation, presentation! Molecular-level insight into linker effects on glycan array screening data

Oliver C Grant[2], Hannah MK Smith[2], Daria Firsova[2], Elisa Fadda[2,3,4], and Robert J Woods[1,2,3]

[2]School of Chemistry, National University of Ireland, Galway, University Road, Galway, Ireland and [3]Complex Carbohydrate Research Centre, University of Georgia, 315 Riverbend Road, Athens, GA, 30602, USA

Received on July 30, 2013; revised on September 14, 2013; accepted on September 15, 2013

Changes in cell-surface glycan patterns are markers of the presence of many different disease and cancer types, offering a relatively untapped niche for glycan-targeting reagents and therapeutics in diagnosis and treatment. Of paramount importance for the success of any glycan-targeting reagent is the ability to specifically recognize the target among the plethora of different glycans that exist in the human body. The preeminent technique for defining specificity is glycan array screening, in which a glycan-binding protein (GBP) can be simultaneously screened against multiple glycans. Glycan array screening has provided unparalleled insight into GBP specificity, but data interpretation suffers from difficulties in identifying false-negative binding arising from altered glycan presentation, associated with the linker used to conjugate the glycan to the surface. In this work, we model the structure and dynamics of the linkers employed in the glycan arrays developed by the Consortium for Functional Glycomics. The modeling takes into account the physical presence and surface polarity of the array, and provides a structure-based rationalization of false-negative results arising from the so-called "linker effect." The results also serve as a guide for interpreting glycan array screening data in a biological context; in particular, we show that attempts to employ natural amino acids as linkers may be prone to unexpected artifacts compromising glycan recognition.

Keywords: glycan array screening / glycan-binding protein / glycan specificity / linker effects / molecular dynamics

Introduction

Glycans are involved in almost all essential processes in a cell's life cycle and are also key players in the onset and progression of many diseases, such as viral and bacterial infection, acquired and inherited diseases and many types of cancer (Dennis et al. 1999; Rudd et al. 2001; Varki et al. 2009a, b). The biological function of glycans depends on their interactions with glycan-binding proteins (GBPs), which include enzymes, lectins and antibodies. The knowledge of the specificity and affinity of GBPs is essential for illuminating the inner workings fundamental to glycobiology, and in the design of glycan-targeting reagents and therapeutics. Currently, the predominant technique used to determine GBP specificity is glycan array screening (Paulson et al. 2006; Oyelaran and Gildersleeve 2009). Glycan arrays consist of a series of glycans covalently conjugated, or non-covalently adhered, to discrete locations on a solid support surface. In covalent arrays, the glycans are often linked to the surface via a synthetic linker. Screening typically consists of incubating a labeled GBP on the array, followed by one or more wash steps, with binding subsequently quantified using fluorescence detection. The types of support surfaces, as well as immobilization techniques, and binding detection methods, vary among glycan arrays (Paulson et al. 2006), as do the structures of the glycans that are immobilized on the surface. It has been observed (Lewallen et al. 2009; Padler-Karavani et al. 2012; Tessier et al. 2013) that the type of linker used to immobilize the glycan on the surface may affect the ability of a glycan to be recognized by its cognate GBP. Linker-dependent false-negative binding has been attributed to the influence of the linker on the presentation of the glycan relative to the support surface. Ideally, the linker would present the glycan in the same way as it is found in natural conditions, but this is impossible with nonnatural linkers and surfaces. While some linkers may exhibit less presentation-related issues, these are difficult to anticipate, and thus in an attempt to circumvent this problem, it is common to use a variety of linker types to conjugate the same glycan to the array surface.

One such array has been made publically available through the Consortium for Functional Glycomics (CFG) (Blixt et al. 2004; Alvarez and Blixt 2006). In the latest version of the CFG array, 21 different types of linkers are present, comprising amino acids, peptides and nonnatural molecules, creating over 600 unique glycan-linker combinations.

In this work, we employ molecular dynamics (MD) simulations of surface-conjugated glycans to gain insight into the role of linker chemistry on glycan presentation. The exact chemical

[1]To whom correspondence should be addressed: Tel: +1-706-542-4454; e-mail: rwoods@ccrc.uga.edu
[4]Present address: Department of Chemistry, National University of Ireland Maynooth, County Kildare, Ireland.

compositions of the silica-based support surfaces employed in glycan arrays are poorly defined, in part because of a lack of physical characterization data, compounded by the fact that many surfaces have been chemically modified, for example by PEGylation, prior to glycan conjugation (Alvarez and Blixt 2006). In order to avoid these issues, and yet provide useful information, MD simulations were performed with the linkers attached to idealized hydrophilic and hydrophobic surfaces. Both surface types were modeled using a flat layer of methanol (MeOH) molecules in a 2D hexagonal packing arrangement. One face of this surface is hydrophilic, while the other is hydrophobic (Figure 1). The shapes adopted by the linkers during these simulations serve as the basis for determining whether they are able to present a glycan for recognition by a GBP.

Linker-dependent false-negative binding has been recently noted in data from glycan array screening of the antitumour antibody JAA-F11 (Tessier et al. 2013), which was raised against the Thomsen–Friedenreich (TF) antigen (Rittenhouse-Diakun et al. 1998). Thus, we selected this system for a more detailed analysis of linker effects. The "TF-antigen" or "T tumour antigen" (Galβ1-3GalNAcα-Ser/Thr), which is also the mucin core 1 precursor in O-linked glycan biosynthesis, is normally hidden from the immune system by extension at the 3- or 6-OH of the Galβ residue and/or at the 6-OH of the α-galactosamine (GalNAcα) residue. This process is disrupted during cancer, and the TF-antigen has been shown to be expressed on the cell surfaces of pancreatic, colon and breast cancers (Heimburg-Molinaro et al. 2011). The TF antigen has been described as a pan-carcinoma antigen after being found to be present in over 90% of human carcinomas where its expression is associated with, and has a functional role in, cancer metastasis (Yu 2007). JAA-F11 is a murine monoclonal antibody that was raised against the TF-antigen and has been shown to target and slow the progression of metastatic tumors in a mouse model (Heimburg et al. 2006).

When screened against the CFG glycan array v4.0, JAA-F11 showed binding only to TF-containing glycans on the CFG array that were linked via a simple alkyl spacer (Sp8). The mAb was unable to recognize the same TF-containing glycans if they were attached to the array via threonine (Sp14) or a phenylene-containing linker (Sp16). The primary objective of the present work is to provide a structure-based rationalization of the false-negative binding that can arise from these "linker-effects" in glycan array screening. Our approach is based on the hypothesis that the physical properties of the linker, the glycans and the array surface are responsible for glycan recognition and that these properties can be determined from molecular simulation.

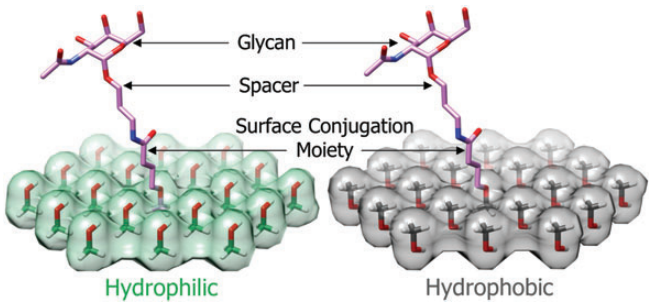

Fig. 1. Glycan-linker-surface models. Left: GalNAcα-Sp8 (stick, pink atom coloring) on a hydrophilic surface (green, transparent) made from methanol (stick, green atom coloring). Right: A hydrophobic surface (dark gray, transparent) generated by inverting the methanol (stick, dark gray atom coloring).

Methods

System generation

The chemical sequences of the spacers used on the CFG array were taken from the CFG website: www.functionalglycomics.org. The 3D structure of each spacer (terminated with a methyl group) was built using Maestro, a package contained within the Schrodinger Software Suite (Suite 2011). In order to obtain energy-optimized ϕ and ψ torsion angles, each spacer was also built with a terminating GalNAcα and/or β-glucosamine (GlcNAcβ) where necessary (Table I). During array fabrication, the spacers are immobilized onto the surface via formation of an amide bond with the n-hydroxysuccinimade-activated surface (Blixt et al. 2004). A 3D structure for this "surface conjugation moiety" attached to a methanol was also generated (Table I). The combination of surface conjugation moiety and spacer constitutes a "linker."

The resulting 3D structures were initially geometry optimized using Argus Labs (Thompson) at the MNDO level of theory using a BFGS search with a 10^{-3} kcal/mol/Å gradient convergence criterion. These MNDO minima were further geometry optimized at the HF/6-31++g(2d,2p) level in Gaussian09 (Frisch et al. 2009). Atomic partial charges were derived for the methyl terminating spacer structures by RESP fitting (Bayly et al. 1993) to the quantum mechanical (QM) molecular electrostatic potential computed with the CHELPG (Breneman and Wiberg 1990) sampling method at the HF/cc-pVTZ level, in accordance with standard AMBER protocols (Kirschner et al. 2008).

The reducing terminal carbohydrates most commonly found on the CFG glycan array, GalNAcα and GlcNAcβ were built using GLYCAM-Web (Woods Group 2005–2013) and attached to the linkers using the tLEaP module of AMBER12 (Case et al. 2012). Linkers that are present on the array attached to glycans via both α- and β-linkages, or whose anomeric configuration was not specified, were modeled with both carbohydrates. Some linkers were found attached to glycans via only one linkage type, and so only one model was necessary (Table I). The glycosidic ϕ and ψ torsion angles were set according to the geometry-optimized values. All of these structures were then exported from tLEaP as PDB files for attachment to the model array surfaces.

Both a hydrophobic and a hydrophilic surface were modeled using a flat layer of methanol (MeOH) molecules. The surfaces were created with a computer script written in the C programing language using the Glylib library of functions (Foley et al.). The C–O bonds of the MeOH molecules were aligned parallel to each other with all of the hydroxyl groups forming one face and all of the methyl groups the other. The molecules were located on a 2D hexagonal grid with an intermolecular spacing of 5 Å. The number of MeOH molecules in the surface ranged from a minimum of 19 to a maximum of 127, corresponding to distances of from 10 to 30 Å between the central MeOH and

Table I. Details of the linkers used on the CFG glycan arrays

Spacer	Formula[a]	Anomer[b]	Surface radius (Å)
Sp0	–NCCO–	α/β	20
Sp8	–NCCCO–	α/β	20
Sp9	–NCCCCCO–	α/β	20
Sp10	–NCC(=O)N–	α/β	10
Sp11	–NCC(=O)Nc1ccc(cc1)CO–	α[c]	30
Sp12 (Asn)[d]	–NC(C(=O)O)CC(=O)N–	α/β	20
Sp13 (Gly)[e]	–NCC(=O)O–	β	10
Sp14 (Thr)[d]	–NC(C(=O)O)C(C)O–	α/β	10
Sp15 (Ser)[d]	–NC(C(=O)O)CO–	α/β	10
Sp16	–Nc1ccc(cc1)O–	α	20
Sp17	–Nc1ccc(cc1)CO–	α[c]	20
Sp18	–NCCCCCC(=O)NCCCO–	β	30
Sp19[f]	–Glu–**Asn** or –**Asn**–Lys	α/β	20
Sp20	–Gly–Glu–**Asn**–Arg	β	30
Sp21	–NCCON(C)–	α/β	20
Sp22	–**Asn**–Ser–Thr	β	20
Sp23	–NCCOCCOCCOCCOCCOCCO–	β	30
Sp24	–Lys–Val–Ala–**Asn**–Lys–Thr	α/β	30
Sp25	–Val–Ala–**Asn**–Lys	α/β	30
Surface conjugation moiety[g]	CO–CCCC(=O)–	–	–

[a]Interpreted from the formulae reported on the CFG website (see Supplementary data, Table SI). The chemical and amino acid type spacers are written as SMILES strings with the restriction that the leftmost dash indicates where the spacer attaches to the surface conjugation moiety and the rightmost dash indicates where the spacer attaches to the glycan. The residues of the peptide type spacers are reported using amino acid three letter codes and in each case the glycan is N-linked via the asparagine (in boldface) side-chain. Each peptide's C terminal is a free carboxylate and the spacer is attached to the surface conjugation moiety via the N terminus of the leftmost amino acid.
[b]If glycans on the array were attached to the spacer via both anomeric configurations or if the anomeric configuration was undefined by the CFG then models for both α and β were generated.
[c]Reported configuration is for Neu5Ac; the inverted configuration was used when modeling with GlcNAc or GalNAc.
[d]Glycan is attached to the spacer via the side-chain. The C terminal carboxylate is free.
[e]Glycan is attached to the spacer via the backbone C terminal carboxylate.
[f]Sp19 is modeled as both EN (Sp19) and NK (Sp19a).
[g]The surface conjugation moiety is connected to the central surface methanol (left side) and to the amino group present in each of the spacers. The atoms of surface conjugation moiety and spacer constitute a linker.

the furthest vertex. Surface size was selected so as to ensure that at its largest extension the glycan-linker did not protrude beyond the edge of the surface.

The linker was attached to the hydroxyl oxygen atom of the central MeOH. In the case of the hydrophobic surface, this linking MeOH was inverted relative to the MeOH molecules comprising the surface (Figure 1). Thus, in the hydrophilic surface the glycan-linker could interact directly with the polar hydroxyl groups, whereas in the hydrophobic case, it could interact with the apolar methyl groups.

The glycan-linker-surface structure was solvated by TIP3P (Jorgensen et al. 1983) water, using the tLEaP module of AMBER12 (Case et al. 2012), with a 10 Å buffer and a contact spacing of 0.5 Å between the solvent and solute van der Waals' surfaces. Charged systems were neutralized by adding either Na^+ or Cl^- counter ions as appropriate, again using tLEaP.

Molecular dynamics

All MD simulations were performed using the CUDA (Gotz et al. 2012) implementation of PMEMD in the Amber12 software suite (Case et al. 2012). The carbohydrate parameters were taken from GLYCAM06h (Kirschner et al. 2008), while the amber99SB force field (Hornak et al. 2006) was used to represent the surface and linker portions. Any parameter missing from the amber99SB force field was obtained from the General Amber force field (Wang et al. 2004). Cutoffs of 10.0 Å for van der Waals and of 8.0 Å for real-space electrostatics were employed. Initial energy minimization (10,000 steps of steepest descent, followed by 10,000 steps of conjugate gradient) was performed with 100 kcal/mol/Å^2 Cartesian restraints on all solute heavy atoms to optimize the water molecules positions and orientations. This minimization step was followed by heating phase (nPT) where the system was brought from 5 to 300 K over 100 ps, followed by 400 ps of equilibration at 300 K. Equilibration was followed by a 100 ns production run, except for the simulations of Sp20, Sp23, Sp24 and Sp25, in which production was performed for 500 ns. Throughout all simulations, the heavy atoms of the surface were restrained with 100 kcal/mol/Å^2 Cartesian restraints.

Creation of 3D linker library

Post-processing of the trajectories was performed with the ptraj module of AMBER12 (Case et al. 2012). For each glycan-linker system, 50 uncorrelated linker shapes were selected at regular intervals and combined from both the hydrophilic and hydrophobic surface simulations.

Root mean squared fluctuation plots

The atomicfluct command in the cpptraj module of AMBER12 (Case et al. 2012) was used to generate root mean squared fluctuation (RMSF) plots from 5,000 snapshots taken at regularly

spaced intervals from the MD simulation. The RMSF data were calculated for each atom in the spacer and surface conjugation moiety that connect the glycan to the surface (i.e., the linker backbone atoms) as well as the C1 and C4 atoms of the glycan. The data were averaged for the atoms of the surface conjugation moiety, the spacer and the glycan.

Isodensity surface generation

The grid command in ptraj was used to generate atomic spatial distribution functions (SDFs) from 5,000 snapshots taken at regularly spaced intervals from the MD simulation. A 50 Å cubic grid was built at 0.5 Å spacing for binning the atomic positions. The resulting SDFs were visualized as isodensity surfaces using chimera (Pettersen et al. 2004). Individual isodensity surfaces of the C1 and C4 ring atoms of the reducing terminal saccharide attached to each linker were generated and combined to give an indication of the glycan presentation relative to the surface.

Grafting of linkers into JAA-F11

The 3D linker libraries of each of the Sp8, Sp14, Sp15 and Sp16 and were grafted onto TF-antigen's reducing terminal residue, using a 3D structure of TF-antigen in complex with JAA-F11 (Tessier et al. 2013). Any vdW overlaps between the atoms in JAA-F11 and the linker were quantified and normalized relative to the overlap of a buried carbon atom.

Results and discussion

Quantifying glycan presentation

The data from the MD simulations were employed to derive average properties for the surface-linker-glycan combinations, of relevance to glycan presentation. Specifically, the effective linker length was computed, as the distance from the surface methanolic oxygen atom, to which the linker is attached, to the C1 atom of the glycan (Table II). This distance enables a direct comparison to be made between linker chemical structure and observed length. For example, the addition of a single methylene group to Sp0 to form Sp8 lengthens the linker by on average 1.3 Å, whereas the addition of three methylenes to form Sp9 lengthens it by only 2.7 Å, far less than that expected from simple linear chain extension (3.9 Å). As the linker becomes longer and more flexible, this deviation from the fully extended idealized conformation becomes more profound. For example, in the case of Sp23, the additional five ethylene glycol units, relative to Sp0, extends the linker by less than half (6.5 Å) of the potential extension (16.7 Å), in accord with the extensive disorder observed experimentally in polyethylene glycol fragments conjugated to proteins (Pai et al. 2011).

Notably, the average linker lengths appear to be largely insensitive to either the monosaccharide or to the polarity of the surface; only in the case of the longest linkers (Sp24 and Sp25) are the differences in length greater than 1 Å.

Differences in the elevation height can be observed as a function of surface type, with the GalNAcα giving rise to average values of 7.9 and 8.6 Å, for hydrophilic and hydrophobic

Table II. Effective linker length and saccharide elevation height of GalNAcα and GlcNAcβ conjugated to linkers on both hydrophilic and hydrophobic surfaces

Linker	GalNAcα				GlcNAcβ			
	Hydrophilic		Hydrophobic		Hydrophilic		Hydrophobic	
	<Length>[a,b]	<Height>[c]	<Length>	<Height>	<Length>	<Height>	<Length>	<Height>
Sp0	8.3 (1.0)[d]	6.2 (1.7)	8.3 (1.1)	7.5 (1.6)	8.4 (1.1)	6.8 (2.0)	8.8 (0.8)	7.8 (1.5)
Sp8	10.0 (1.2)	8.1 (2.1)	9.6 (1.1)	8.5 (1.8)	9.7 (1.0)	7.2 (2.6)	9.8 (1.2)	8.8 (2.5)
Sp9	11.0 (1.7)	7.7 (2.8)	11.2 (1.5)	8.8 (2.6)	11.1 (1.5)	7.3 (3.2)	11.4 (1.4)	8.9 (3.0)
Sp10	8.5 (1.0)	7.0 (2.0)	9.0 (0.7)	8.5 (1.7)	8.7 (0.8)	7.7 (1.6)	8.6 (0.6)	8.2 (1.2)
Sp11	–	–	–	–	12.9 (1.9)	8.5 (3.8)	13.5 (1.4)	8.5 (3.3)
Sp12	10.7 (0.7)	8.3 (2.1)	10.6 (0.9)	9.2 (1.9)	10.8 (0.7)	9.9 (2.0)	10.7 (0.7)	10.3 (2.0)
Sp13	–	–	–	–	8.6 (1.0)	7.7 (1.8)	8.6 (0.7)	8.2 (1.3)
Sp14	8.1 (1.1)	5.9 (1.5)	8.1 (1.1)	7.1 (1.2)	9.0 (0.8)	7.7 (1.4)	8.8 (0.6)	7.7 (1.1)
Sp15	8.4 (1.1)	6.5 (1.9)	8.3 (1.1)	7.1 (1.7)	–	–	–	–
Sp16	11.3 (0.6)	9.3 (2.5)	10.5 (0.8)	6.6 (3.2)	–	–	–	–
Sp17	12.0 (1.1)	7.5 (3.0)	11.7 (1.1)	9.1 (2.3)	–	–	–	–
Sp18	–	–	–	–	14.3 (2.6)	9.1 (4.0)	14.7 (2.2)	9.9 (3.9)
Sp19	10.8 (0.7)	8.3 (2.1)	9.7 (1.3)	8.4 (2.0)	11.0 (0.7)	9.8 (2.0)	10.0 (1.1)	8.8 (2.3)
Sp19a	12.4 (1.6)	9.0 (3.0)	12.6 (1.2)	9.0 (3.3)	12.3 (1.6)	9.0 (3.2)	12.1 (1.4)	9.7 (2.7)
Sp20	–	–	–	–	14.1 (2.6)	11.2 (3.8)	13.8 (2.2)	10.3 (3.7)
Sp21	9.0 (1.2)	6.6 (1.7)	9.0 (1.0)	6.7 (1.6)	8.5 (1.4)	6.6 (2.2)	9.0 (1.3)	7.7 (2.1)
Sp22	–	–	–	–	10.5 (1.1)	8.7 (2.5)	11.1 (0.6)	10.7 (1.6)
Sp23	–	–	–	–	14.6 (4.3)	8.8 (4.1)	15.3 (4.3)	10.5 (4.5)
Sp24	16.5 (3.1)	11.1 (4.4)	16.4 (3.2)	12.7 (4.6)	15.3 (3.6)	12.6 (4.4)	16.9 (2.6)	13.1 (4.5)
Sp25	14.2 (2.3)	9.6 (3.5)	13.7 (2.7)	10.8 (3.6)	13.7 (3.0)	10.1 (3.9)	14.3 (2.4)	11.0 (3.8)
Range	8.1–16.5	5.9–11.1	8.1–16.4	6.7–12.7	8.5–15.3	6.6–12.6	8.6–16.9	7.7–13.1

[a]Averaging is indicated by angled brackets and is based on 5,000 evenly spaced snapshots from the MD simulation.
[b]Distance between the oxygen atom of central surface methanol and the C1 atom of the glycan, in Å.
[c]Distance between the array surface and the geometric centroid of the glycan ring atoms, in Å.
[d]Standard deviations are presented in parentheses. With 5,000 snapshots, and using maximum standard deviations of 4.3 and 4.6 for length and height respectively, two values can be deemed statistically different ($P = 0.001$) if their difference is larger than 0.3 Å for either length or height.

surfaces, respectively. These may be compared with values of 8.7 and 9.4 Å for GlcNAcβ. As in the case of linker length, linker elevation was also relatively insensitive to the monosaccharide type. The subtle increase in elevation in the case of GlcNAcβ may be accounted for by the equatorial configuration at C1, which tends to place the monosaccharide ring further from the surface, compared with the axial (α) configuration. Overall, in the hydrophilic surface simulations the monosaccharide is closer to the surface by 0.7 Å. The aromatic linker, Sp16, is a notable exception to this trend, displaying a strong tendency to be closer to the hydrophobic surface (by 2.7 Å).

As expected, the degree of motion of each glycan-linker structure varies significantly with the length of the linker. For example, a short linker, such as Sp0, displays an RMSF of 3–4 Å, which can increase to approximately 8 Å for the longest linkers (see Figure 2 and Supplementary data, Figures S1–S5). It is notable that the extent of the linker motion is not exclusively dependent on the number of atoms in the linker connecting the glycan to the surface (i.e., the linker backbone atoms). Rather, linker length and motional properties also depend on the chemical structure of the linker, its interaction with solvent, and the surface polarity.

In order to more effectively illustrate the presentation properties of the surface-linker-glycan combinations, the positions of the C1 and C4 atoms were monitored over the course of the MD simulation and plotted in three dimensions as isodensity surfaces. The isodensity surfaces for the C1 and C4 atoms were overlaid for each system, generating a total of 62 plots (Figure 3 and Supplementary data, Figures S6–S10). When presented in this manner, the relative positions of the C1 and C4 atoms provide an indication of the orientation of the glycan ring throughout the simulations and complement the distance data in Table II. Notably, the anomeric configuration at the reducing terminus of the glycan has a pronounced effect on glycan presentation, which was not captured fully in the data in Table II. This feature is exemplified by comparing the simulations of GalNAcα and GlcNAcβ linked to the surface via Sp0. A β-linkage with Sp0 frequently places the C4 atom distal to a hydrophilic surface, whereas the α-configuration leads to a bimodal distribution in which the C4 atom is more often proximal to the surface (Figure 3).

The type of surface can also have a pronounced effect on the relative orientation of the glycan. This effect can be observed in the GalNAcα-Sp0 plots where the C4 atom is frequently seen to be closer to the surface than the C1 atom, indicating that this glycan appears to interact much more with the hydrophilic surface than with the hydrophobic surface. This surface polarity effect is particularly apparent in the Sp16 simulation (Figure 3), where the phenylene ring in the linker forms stacking interactions with the hydrophobic surface, whereas the glycan is more often orientated away from the surface in the hydrophilic surface simulation. This phenomenon is also observed with the simulation of GlcNAcβ-Sp11 and GalNAcα-Sp17, which both contain phenylene rings but are longer than Sp16. With some linkers the length and flexibility were so great that the distributions appear to be much less influenced by the surface polarity, as observed for GlcNAcβ-Sp18.

Four amino acid linkers are present in the current CFG array (Sp12, asparagine, Sp13, glycine, Sp14, threonine and Sp15 serine), which are linked to the surface via their N-termini and are all conjugated to the glycan via their side-chains, with the exception of glycine. In the case of glycine, the glycan is connected via the carboxylate. Because the amino acids project approximately perpendicularly from the array surface, their side-chains are presented parallel to the surface. For residues with short side-chains (Ser and Thr), this results in glycan-presentation that is

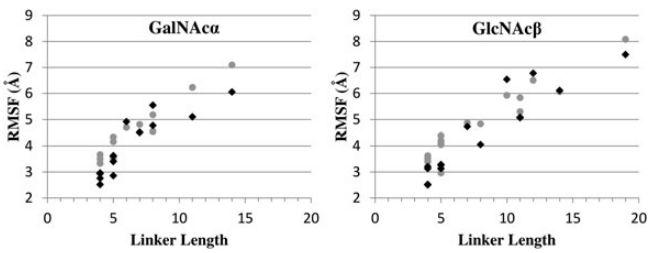

Fig. 2. Linker motion (RMSF) as a function of the number of atoms in the spacer portion of the linker calculated for linkers attached to both hydrophilic (gray circles) and hydrophobic (black diamonds) surfaces.

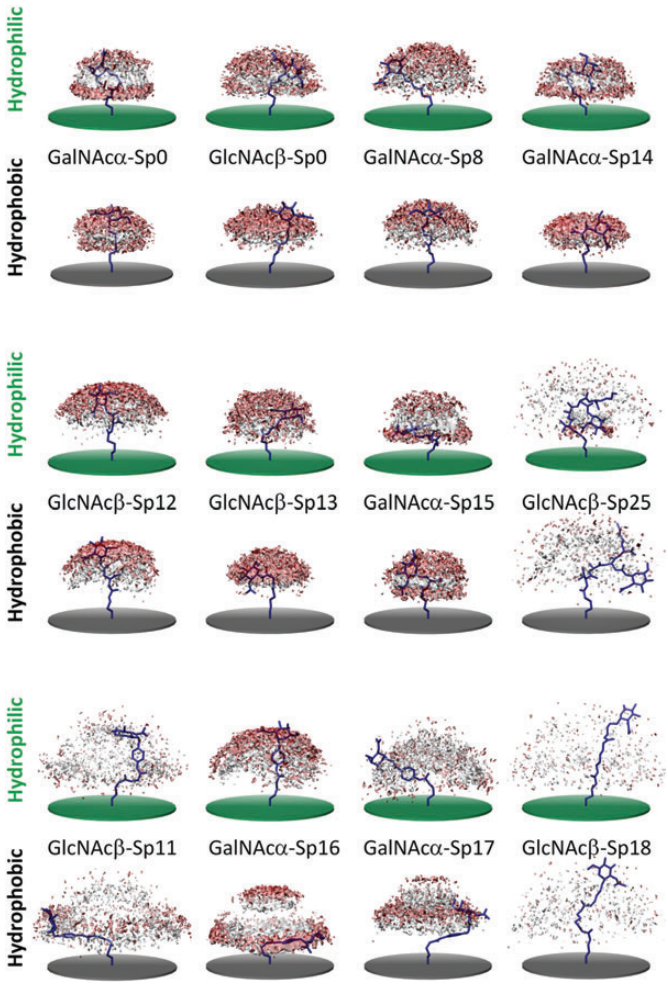

Fig. 3. Isodensity surfaces depicting the spatial distributions of the C1 (light gray) and C4 (red) glycan atoms from 5,000 snapshots extracted from the MD simulations. A single 3D structure of the glycan and linker, extracted from each MD simulation, is shown (stick, navy blue).

also approximately parallel to the surface. This orientation is in contrast to what would be expected in a glycoprotein, in which the glycosylated side-chains would be presented more or less perpendicularly from the protein surface. In the case of the asparagine linker, its slightly greater length (relative to Ser and Thr) as well as its amido functionality results in glycan that is approximately perpendicular. The glycine linker also tends to present the glycan in a perpendicular orientation, presumably as a result of the nonnatural conjugation via the terminal carboxylate. While some of these presentational properties might be anticipated, the present analysis provides a unique method to both predict and quantify their existence.

The peptide linkers used to represent N-linked glycosylation sites on the CFG array are highly flexible and adopt a multitude of different shapes (Figure 3). In a biological context, this may be a good model for N-linked glycosylation occurring at flexible N or C termini, but does not necessarily reflect the properties of N-linked glycosylation sites. A peculiar example is Sp25, which in the hydrophilic simulation appears to prefer a shape where the glycan is interacting with both the surface and the peptide linker. Thus, care should be taken when interpreting data from glycans linked to specific peptide sequences. An alternative to this approach is to attach glycans to carrier proteins, forming neoglycoproteins, which can be printed onto the array surface (Oyelaran and Gildersleeve 2009) or to neoglycolipids (Feizi et al. 1994). In principle, employing arrays of neoglycolipids and neoglycoproteins for glycan presentation offers some benefits over the use of flat polymeric surfaces; particularly, in the sense that the polarity of the lipid or protein carrier is biologically relevant. However, the use of nonnatural linkers between the glycan and the biomolecule may again introduce recognition artifacts. Further, the topology (convex, concave etc.) of the site of glycosylation, in the case of a neoglycoprotein, may also impact on glycan presentation (Petrescu et al. 2004).

Assessing array binding

When considering array data, binding may be abrogated by several properties of the system. A true binder is a ligand that can be accommodated in the receptor protein, with a measurable affinity. A true nonbinder is a glycan which contains a change to the structure of the glycan that prevents it from being bound (Tessier et al. 2013). In contrast, false-negative binding may occur if the linker shape cannot fit into the binding site (linker-collision false-negative binding), or if the preferred presentation of the linker relative to the array surface is such that the glycan is effectively inaccessible to the protein (presentation-dependent false-negative binding). Presented in Table III are binding data for the JAA-F11 system, which illustrate such linker-related false-negative binding. Thus, in order for a linker to facilitate binding, there are two key issues: First, there must not be any large steric overlaps between the linker and the receptor-binding site and, secondly, the linker must orient the glycan relative to the surface so that the receptor can bind the ligand in the correct way, without being inhibited by the array surface (Figure 4).

Collisions between the protein and the array surface, which may arise depending on linker conformation, were accounted for if any protein Cα atom was predicted to penetrate the array surface. These putative collisions were normalized as the percentage of Cα atoms predicted to be below the plane of the array surface. In the case of a hydrophilic array, the surface was defined by the plane containing the oxygen atoms; a similar definition based on the carbon atoms of the methyl groups was employed for the hydrophobic surface. It should be noted that

Table III. Comparison of experimental specificity data for mAb JAA-F11 with TF-containing glycans[a] demonstrating a linker effect on the CFG v4.0 glycan array

CFG array v4.0 ID			Glycan sequence	Experimental RFU[b]		
Sp8	Sp14	Sp16		Sp8	Sp14	Sp16
131	278	132	**Galβ1-3GalNAcα**	98	0	0
127	128	–[c]	Neu5Acα2-6(**Galβ1-3**)**GalNAcα**	78	0	–
159	158[d]	–	Galβ1-4GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**	52	0	–
125	126[d]	–	GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**	51	0	–

[a]The TF-disaccharide portion of each glycan is in boldface.
[b]Normalized RFUs averaged over all protein concentrations (0.1, 5 and 200 μg/mL) and over multiple values for the same glycan, when present on the CFG array.
[c]Dashes indicate that glycan-linker combination is not present on the glycan array.
[d]Reducing anomeric configuration undefined on the CFG array, α-configuration assumed for the CCG analysis.

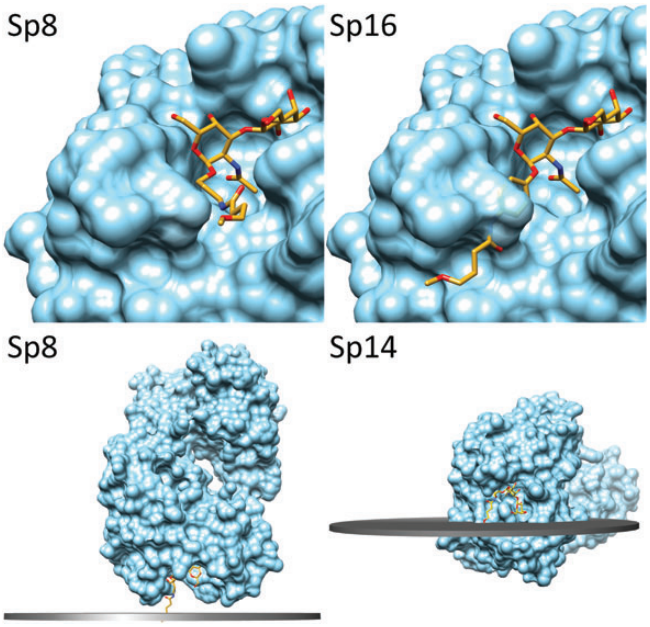

Fig. 4. Top left: The linker portion of TF-Sp8 (stick, goldenrod atom coloring) is tolerated in the binding site of the glycan-binding protein, JAA-F11 (surface, cyan). Top right: The linker portion of TF-Sp16 is not tolerated in the binding site of JAA-F11 (surface, cyan) due to overlaps with the side-chain of Tyr 34 (surface, transparent cyan). Bottom left: The linker portion of Sp8-TF (stick, goldenrod atom coloring) presents the TF-antigen relative to the array surface (disc, dark gray) so that JAA-F11 (surface, cyan) is able to bind. Bottom-right: The linker portion of Sp14-TF (stick, goldenrod atom coloring) presents the TF-antigen relative to the array surface (disc, dark gray) so that JAA-F11 (surface, cyan) is unable to access the glycan.

any induced fit in the glycan or the receptor upon binding and any effect of larger oligosaccharides on the dynamics of the linker are neglected by this model.

Identification of linker-related false-negative binding

The data in Table III indicate that mAb JAA-F11 binds only to TF-glycans that are linked to the array via Sp8. On examination, several TF-containing glycans that do bind when immobilized via Sp8 do not bind when conjugated via other linkers, for example, the antibody does not bind to the TF-disaccharide if the glycan is linked to the array via Sp14 or Sp16. These are clearly evidence of false-negative linker-dependent effects. Sp14 and Sp16 are, therefore, incompatible with binding, either because they are not tolerated in the binding site or because they present the glycan in such a way that it is inaccessible to the protein. To assess the ability of any given linker shape to fit into the protein-binding site, the reducing terminal residue in the bound glycan was superimposed onto the corresponding monosaccharide in the monosaccharide-linker structures obtained from the MD simulations. The protein–ligand complex was then evaluated for the presence of any steric overlaps between the linker and the protein surface (see Table IV and Figure 4, top right).

This analysis indicated that the rigid phenylene structure of Sp16 does not adopt any conformations that could sterically fit into the mAb-binding site. Similarly, the majority (~90%) of the conformations of Sp14 (Thr) are incapable of fitting into the binding site. This is also predicted to be the case for Sp15 (Ser), although there are no relevant TF-containing glycans linked via serine on the CFG array (v4.0). Notably, the simulations predict that even for the noninterfering linker Sp8, only ~37% of the conformations are compatible with binding. Thus, it is not necessary for all linker conformations to be acceptable.

However, the main factor responsible for the lack of binding of TF-linked via Sp14 appears to arise from unfavorable glycan presentation. This is evident from the fact that only 2% of the shapes of Sp14 permit the protein to bind without collisions with the array surface (Figure 4, Bottom right). Sp15 appears to be more acceptable than Sp14 in general; however, Sp8 is predicted to be a superior linker in each category.

The comparison between the most populated pose from the hydrophilic simulations of Sp14 and Sp8 clearly demonstrates the difference in glycan presentation (Figure 5). The curved geometry of Sp14 presents the glycan toward the surface, whereas Sp8 is perpendicular to, and orients the glycan away from, the surface. O-linked mucin type glycans are commonly found on intrinsically disordered protein regions (Nishikawa et al. 2010). The orientation of an O-linked glycan on a threonine residue is, therefore, possibly better represented by the Sp8 linker than the Sp14 "threonine" linker (Figure 5).

Predictive reinterpretation of CFG v4.0 array screening of JAA-F11

The glycan-binding preferences of JAA-F11 have been recently interpreted from a structural perspective by employing a theoretical method termed computational carbohydrate grafting (CCG) (Tessier et al. 2013). Using CCG, a library of 3D glycan shapes for the TF-containing glycans on the CFG array (v4.0) was created and searched for putative binding and nonbinding glycans, based on their ability to fit into the binding site. The screening was carried out by grafting the branches of the TF-glycans onto a structure of the bound TF-disaccharide. If a branch formed excessive overlaps with the protein-binding site, it was predicted to be a non-binder. The CCG results were in complete agreement with the glycan array screening data; however, some inconsistencies in the experimental data were noted as a function of linker type. Fortunately, several of the JAA-F11 binding partners are conjugated to the CFG array via multiple linker types (Table V, glycans **1**–**9**) allowing linker effects to be inferred from the data. In this work, we have provided a structure-based rationale for the observed linker effects, using a modification of the CCG protocol, allowing reinterpretation of the JAA-F11 glycan array data (Table V). All of the false-negative data points in the CFG data for this system were predicted to arise either from glycan presentation defects or from physical incompatibility of the linker with the binding site topology of the mAb. This analysis enabled us to predict that a further two glycans (**14** and **21**), which are found conjugated solely by Sp14, are probable binding partners for JAA-F11.

The ability to predict false negatives in glycan array data is essential in order to fully define specificities and binding motifs. In order to aid in the detection of linker-dependent false-negative binding, array developers employ multiple linker types for each glycan. However, given the high number of linkers in use, it is likely unfeasible to include all permutations of linker chemistry with any given glycan. An independent

Table IV. Results from grafting TF-linkers into the JAA-F11-binding site

Linker	Tolerated[a] shapes (hydrophilic, hydrophobic)	Correct[b] glycan orientation (hydrophilic, hydrophobic)	**Permit[c] Binding (hydrophilic, hydrophobic)**
Sp8	37 (16/21)	36 (15/21)	**23 (8/14)**
Sp14	11 (5/6)	2 (0/2)	**1 (0/1)**
Sp15	12 (4/8)	13 (5/8)	**4 (1/3)**
Sp16	0	31 (23/8)	**0**

[a]Percentage of linker shapes from the MD simulation that have less than one carbon atom equivalent buried in the JAA-F11 surface, when superimposed onto the reducing terminal of the co-complexed TF-antigen.
[b]Percentage of linker shapes from the MD simulation that present the glycan so that JAA-F11 can bind without colliding with the array surface.
[c]Percentage of linker shapes that simultaneously satisfy both conditions in [a] and [b].

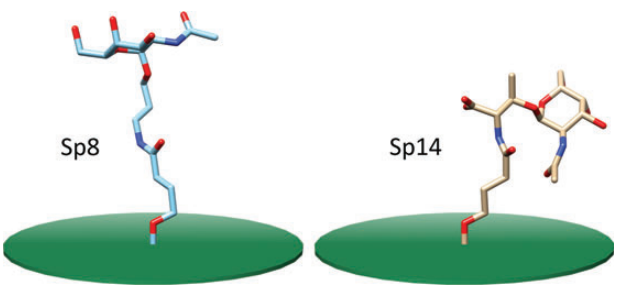

Fig. 5. Orientation of the most populated shape from the hydrophilic surface of GalNAcα conjugated via Sp8 (left side, stick model) and Sp14 (right side, stick model).

Table V. Reinterpreted glycan array data for mAb JAA-F11 highlighting predicted false-negatives

ID	Array glycan	Experimental binding[a]
1	**Galβ1-3GalNAcα**-Sp8	Yes
2	**Galβ1-3GalNAcα**-Sp16	No[d]
3	**Galβ1-3GalNAcα**-Sp14	No[d]
4	Neu5Acα2-6(**Galβ1-3**)**GalNAcα**-Sp8	Yes
5	Neu5Acα2-6(**Galβ1-3**)**GalNAcα**-Sp14	No[d]
6	GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp8	Yes
7	GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp14	No[d]
8	Galβ1-4GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp8	Yes
9	Galβ1-4GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp14	No[d]
10	Neu5Acβ2-6(**Galβ1-3**)**GalNAcα**-Sp8	Yes
11	KDNα2-3**Galβ1-3GalNAcα**-Sp14	No
12	Neu5Acα2-6(Neu5Acα2-3**Galβ1-3**)**GalNAcα**-Sp14	No
13	Neu5Acα2-3Galβ1-4GlcNAcβ1-6(Neu5Acα2-3**Galβ1-3**)**GalNAcα**-Sp14	No
14	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp14	No[d]
15	Fucα1-2**Galβ1-3GalNAcα**1-3(Fucα1-2)Galβ1-4Glcβ-Sp0	No
16	GlcNAcβ1-2**Galβ1-3GalNAcα**-Sp8	No
17	Neu5Acα2-3**Galβ1-3GalNAcα**-Sp8	No
18	Fucα1-2**Galβ1-3GalNAcα**-Sp14	No
19	Neu5Acα2-3**Galβ1-3GalNAcα**-Sp14	No
20	Neu5Acα2-6(Neu5Acα2-3**Galβ1-3**)**GalNAcα**-Sp8	No
21	Neu5Acα2-3Galβ1-4GlcNAcβ1-6(**Galβ1-3**)**GalNAcα**-Sp14	No[d]
22	(3S)**Galβ1-3GalNAcα**-Sp8	No
23	Fucα1-2**Galβ1-3GalNAcα**-Sp8	No
24	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-6(Neu5Acα2-3**Galβ1-3**)**GalNAc**[b]-Sp14	No
25	Fucα1-2**Galβ1-3GalNAcα**1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp0	No
26	GalNAcα1-3(Fucα1-2)**Galβ1-3GalNAcα**1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp0	No
27	GlcNAcβ1-3**Galβ1-3GalNAcα**-Sp8	No
28	**Galβ1-3GalNAcα**1-3(Fucα1-2)Galβ1-4Glc[c]-Sp0	No
29	GlcNAcα1-4**Galβ1-3GalNAc**[b]-Sp14	No
30	**Galβ1-3GalNAcα**1-3(Fucα1-2)Galβ1-4GlcNAc[c]-Sp0	No
31	6S(Neu5Acα2-3**Galβ1-3**)**GalNAcα**-Sp8	No

[a]Binders defined as having greater than 10% of maximal RFUs at each protein concentrations (0.1, 5 and 200 μg/mL).
[b]Reducing anomeric configuration undefined on the CFG array, α-configuration assumed for the CCG analysis.
[c]Reducing anomeric configuration undefined on the CFG array, both configurations tested in the CCG analysis.
[d]Highlighted in grey are the predicted false negatives from the v4.0 CFG array screening of JAA-F11.

analysis that can predict false negatives can assist experimentalists by focusing effort on re-evaluating only a subset of key glycans.

Conclusions

Glycan array screening has emerged as one of the preeminent methods for defining the specificity of carbohydrate-binding proteins. These arrays are frequently generated by immobilizing glycans through synthetic or natural linkers to a surface, such as glass or silica. The binding of proteins to immobilized glycans has been shown to be sensitive to linker choice (Lewallen et al. 2009; Padler-Karavani et al. 2012; Tessier et al. 2013), presumably as a result of differences in the 3D presentation of the glycan.

Here, we have employed computational modeling to examine the effect of linker structure and properties on glycan presentation, relative to both a hydrophilic and a hydrophobic surface. The modeling provides a structural basis for interpreting and predicting the impact of linker choice on glycan recognition. We have focused on the binding properties reported recently (Tessier et al. 2013) for the anticarbohydrate mAb JAA-F11, against the CFG glycan array (v4.0); however, the observations are relevant to array design and screening data interpretation in general. The impact of linker effects is likely accentuated in the case of small glycans, such as TF, but false-negative binding has also been attributed to linker effects in the binding of proteins to recognition motifs on the outer arms of larger glycans (Lewallen et al. 2009; Padler-Karavani et al. 2012).

The degree to which linker chemistry affects glycan presentation depends both on the length of the linker (shorter being more likely to restrict access to the glycan), and on the orientation of the linker relative to the array surface. Presentation can be influenced by nonspecific interactions between the linker and the surface. This is evident in the case of hydrophobic linkers (Sp11, Sp16 and Sp17) attached to hydrophobic surfaces, which prefer orientations parallel to the array surface, thereby maximizing hydrophobic contacts. Glycan presentation can also be altered by the 3D structure of the linker itself. For example, linkers Sp14 (Thr) and Sp15 (Ser) are predicted to prefer to orient the TF-glycan parallel to the array surface, making it inaccessible to the JAA-F11 antibody. Further, the results illustrate that certain relatively rigid linkers, such as Sp16, can prevent binding because they make unfavorable interactions with the protein surface. Finally, each glycan-binding protein will have a unique sensitivity to presentation effects by virtue of the variations in protein shape and binding site topology. It is hoped that the present work provides both the technique and motivation for array designers to perform a linker assessment during the early stages of linker selection. We have concluded with regard to the CFG linkers that Sp8, Sp9, Sp12, Sp18, Sp20, Sp23 and Sp24, are highly flexible and/or have good glycan presentation properties. In terms of surface polarity, we have considered two extremes; a purely hydrophilic and a purely hydrophobic surface. Neither is directly equivalent to a real array surface, the details of which are poorly defined. Nevertheless, the data permit us to conclude that increasing surface hydrophobicity is likely to increase the presence of nonspecific interactions between the linker and or the glycan, leading to glycan recognition artifacts.

In the case of mAb JAA-F11, the CFG array data could lead to the initial conclusion that this mAb does not recognize TF-antigen whenever the antigen is attached to threonine. Given that this mAb does localize in tumor tissues in vivo, the implication would be that it must be binding to serine-linked TF. However, our analysis suggests that the lack of binding to threonine depends at least in part on the nature of the presentation of TF-Thr on the array, and that if TF-Ser were present, it would also exhibit attenuated binding. Thus, this mAb may bind to more cancer-associated TF-linkages than indicated by the glycan array data. Also noteworthy is the fact that JAA-F11 was raised by immunization with TF-O-CH_2CH_2-S-CH_2-CH_2-NH-$COCH_2$

CH_2-BSA (Rittenhouse-Diakun et al. 1998), whose "TF-O-CH_2 CH_2" portion is analogous to Sp8, and to the side-chain of serine.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by the National Institutes for Health (GM094919 (EUREKA)), as well as the Science Foundation of Ireland (08/IN.1/B2070) and the European Research Development Fund.

Acknowledgements

The authors thank Ryan McBride and James Paulson at the Department of Molecular Biology and Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA, and David Smith at the Department of Biochemistry, Emory University, Atlanta, GA, USA for clarifying aspects of the linker chemistry used on the CFG glycan array.

Conflict of interest

None declared.

Abbreviations

CCG, computational carbohydrate grafting; CFG, Consortium for Functional Glycomics; GalNAcα, α-galactosamine; GBP, glycan-binding protein; GlcNAcβ, β-glucosamine; MD, Molecular Dynamics; QM, quantum mechanical; RMSF, root mean squared fluctuation; SDFs, spatial distribution functions; TF Thomsen-Friedenreich.

References

Alvarez RA, Blixt O. 2006. Identification of ligand specificities for glycan-binding proteins using glycan arrays. *Methods Enzymol*. 415:292–310.

Bayly CI, Cieplak P, Cornell W, Kollman PA. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J Phys Chem*. 97:10269–10280.

Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, et al. 2004. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA*. 101:17033–17038.

Breneman CM, Wiberg KB. 1990. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J Comput Chem*. 11:361–373.

Case DA, Darden TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, et al. 2012. *AMBER 12*. San Francisco, CA: University of California.

Dennis JW, Granovsky M, Warren CE. 1999. Glycoprotein glycosylation and cancer progression. *Biochim Biophys Acta*. 1473:21–34.

Feizi T, Stoll MS, Yuen CT, Chai W, Lawson AM. 1994. Neoglycolipids: Probes of oligosaccharide structure, antigenicity, and function. *Methods Enzymol*. 230:484–519.

Foley BL, Tessier MN, Makeneni S, Nivedha A. *GLYLIB Programmers' Libraries*. Complex Carbohydrate Research Center, The University of Georgia. http://glycam.ccrc.uga.edu/glylib.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, et al. *Gaussian 09, Revision A.1*. Wallingford, CT: Gaussian, Inc.

Gotz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. 2012. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput*. 8:1542–1555.

Heimburg J, Yan J, Morey S, Glinskii OV, Huxley VH, Wild L, Klick R, Roy R, Glinsky VV, Rittenhouse-Olson K. 2006. Inhibition of spontaneous breast cancer metastasis by anti-Thomsen-Friedenreich antigen monoclonal antibody JAA-F11. *Neoplasia*. 8:939–948.

Heimburg-Molinaro J, Lum M, Vijay G, Jain M, Almogren A, Rittenhouse-Olson K. 2011. Cancer vaccines and carbohydrate epitopes. *Vaccine*. 29:8802–8826.

Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*. 65:712–725.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 79:926–935.

Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J Comput Chem*. 29:622–655.

Lewallen DM, Siler D, Iyer SS. 2009. Factors affecting protein-glycan specificity: Effect of spacers and incubation time. *Chembiochem*. 10:1486–1489.

Nishikawa I, Nakajima Y, Ito M, Fukuchi S, Homma K, Nishikawa K. 2010. Computational prediction of O-linked glycosylation sites that preferentially map on intrinsically disordered regions of extracellular proteins. *Int J Mol Sci*. 11:4991–5008.

Oyelaran O, Gildersleeve JC. 2009. Glycan arrays: Recent advances and future challenges. *Curr Opin Chem Biol*. 13:406–413.

Padler-Karavani V, Song X, Yu H, Hurtado-Ziola N, Huang S, Muthana S, Chokhawala HA, Cheng J, Verhagen A, Langereis MA, et al. 2012. Cross-comparison of protein recognition of sialic acid diversity on two novel sialoglycan microarrays. *J Biol Chem*. 287:22593–22608.

Pai SS, Hammouda B, Hong K, Pozzo DC, Przybycien TM, Tilton RD. 2011. The conformation of the poly(ethylene glycol) chain in mono-PEGylated lysozyme and mono-PEGylated human growth hormone. *Bioconjug Chem*. 22:2317–2323.

Paulson JC, Blixt O, Collins BE. 2006. Sweet spots in functional glycomics. *Nat Chem Biol*. 2:238–248.

Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. 2004. Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology*. 14:103–114.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem*. 25:1605–1612.

Rittenhouse-Diakun K, Xia Z, Pickhardt D, Morey S, Baek MG, Roy R. 1998. Development and characterization of monoclonal antibody to T-antigen: (gal beta1–3GalNAc-alpha-O). *Hybridoma*. 17:165–173.

Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. 2001. Glycosylation and the immune system. *Science*. 291:2370–2376.

Suite. 2011. *Maestro, version 9.2*. New York, NY: Schrödinger, LLC.

Tessier MB, Grant OC, Heimburg-Molinaro J, Smith D, Jadey S, Gulick AM, Glushka J, Deutscher SL, Rittenhouse-Olson K, Woods RJ. 2013. Computational screening of the human TF-glycome provides a structural definition for the specificity of anti-tumor antibody JAA-F11. *PLoS One*. 8: e54874.

Thompson MA. *ArgusLab 4.0.1*. Seattle, WA: Planaria Software LLC.

Varki A, Cummings RD, Esko JD, H.H. F, Stanley P, Bertozzi CR, Hart GW, Etzler ME. 2009a. *Glycans in Acquired Human Diseases.* Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Available from http://www.ncbi.nlm.nih.gov/books/NBK1908/.

Varki A, Cummings RD, Esko JD, H.H. F, Stanley P, Bertozzi CR, Hart GW, Etzler ME. 2009b. *Glycans in Physiology and Disease.* Essentials of Glycobiology. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, Available from http://www.ncbi.nlm.nih.gov/books/NBK1908/.

Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *J Comput Chem*. 25:1157–1174.

Woods Group. (2005–2013). *GLYCAM Web*. Athens, GA: Complex Carbohydrate Research Center, University of Georgia (http://www.glycam.com).

Yu LG. 2007. The oncofetal Thomsen-Friedenreich carbohydrate antigen in cancer progression. *Glycoconj J*. 24:411–420.