

Article

# Conformation and Cross-Protection in Group B Streptococcus Serotype III and *Streptococcus pneumoniae* Serotype 14: A Molecular Modeling Study

Michelle M. Kuttel <sup>1,\*</sup>  and Neil Ravenscroft <sup>2</sup><sup>1</sup> Department of Computer Science, University of Cape Town, Cape Town 7701, South Africa<sup>2</sup> Department of Chemistry, University of Cape Town, Cape Town 7701, South Africa;  
neil.ravenscroft@uct.ac.za

\* Correspondence: mkuttel@cs.uct.ac.za; Tel.: +27-21-6505107

Received: 19 January 2019; Accepted: 9 February 2019; Published: 13 February 2019



**Abstract:** Although the branched capsular polysaccharides of *Streptococcus agalactiae* serotype III (GBSIII PS) and *Streptococcus pneumoniae* serotype 14 (Pn14 PS) differ only in the addition of a terminal sialic acid on the GBSIII PS side chains, these very similar polysaccharides are immunogenically distinct. Our simulations of GBSIII PS, Pn14 PS and the unbranched backbone polysaccharide provide a conformational rationale for the different antigenic epitopes identified for these PS. We find that side chains stabilize the proximal  $\beta$ DGlc(1 $\rightarrow$ 6) $\beta$ DGlcNAc backbone linkage, restricting rotation and creating a well-defined conformational epitope at the branch point. This agrees with the glycotope structure recognized by an anti-GBSIII PS functional monoclonal antibody. We find the same dominant solution conformation for GBSIII and Pn14 PS: aside from the branch point, the backbone is very flexible with a “zig-zag” conformational habit, rather than the helix previously proposed for GBSIII PS. This suggests a common strategy for bacterial evasion of the host immune system: a flexible backbone that is less perceptible to the immune system, combined with conformationally-defined branch points presenting human-mimic epitopes. This work demonstrates how small structural features such as side chains can alter the conformation of a polysaccharide by restricting rotation around backbone linkages.

**Keywords:** capsular polysaccharide; carbohydrate antigen; molecular modeling; Group B Streptococcus; *Streptococcus pneumoniae*; conjugate vaccines

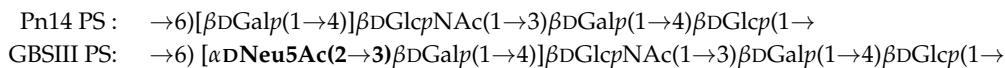
## 1. Introduction

The bacterium *Streptococcus agalactiae*, usually termed Group B Streptococcus, is a primary cause of neonatal sepsis and meningitis, particularly in infants born to carriers of the pathogen. The *Streptococcus pneumoniae* bacterium is another common cause of serious infections in young infants, including meningitis and pneumonia. Ten serotypes of Group B Streptococcus have been characterized, of which serotype III (GBSIII) is currently the most prevalent [1]. Over 90 serotypes of *Streptococcus pneumoniae* have been identified, with serotype 14 (Pn14) being the most common cause of invasive pneumococcal disease in children prior to the introduction of conjugate vaccines [2].

Both of these gram-positive bacteria are encapsulated by polysaccharides that vary in structure according to bacterial serotype and are essential for bacterial virulence: vaccination with carbohydrate-protein conjugates can provide effective serotype-specific protection. The Pn14 capsular polysaccharide (PS) is a component of all licensed conjugate vaccines since the introduction of the

7-valent Prevenar vaccine. GBSIII PS is present in a trivalent conjugate vaccine targeting serotypes Ia, Ib, and III that has completed phase-2 trials [3,4] and a hexavalent vaccine currently in clinical trials.

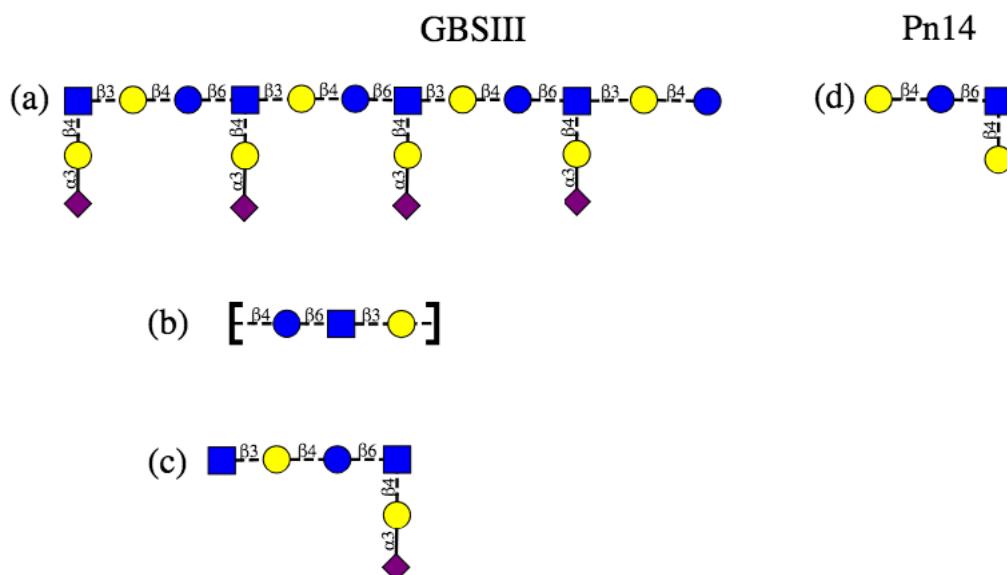
The similarity of the branched GBSIII PS and Pn14 PS has long been of interest [5]: they are identical except that the GBSIII PS carries a terminal  $\alpha$ (2→3)-linked sialic acid ( $\alpha$ DNeu5Ac) on the galactose side chain. The Pn14 PS thus has a four-residue repeat unit (RU) and the GBSIII PS a five-residue RU, as follows.



This structural similarity of the GBSIII PS and Pn14 PS raises the possibility that type-specific antibodies induced by one of these organisms might protect against disease caused by the other, a phenomenon referred to as cross-protection. Indeed, vaccination with GBSIII PS has been shown to raise two types of anti-carbohydrate antibody: a major population recognising the native PS but not the desialylated PS (equivalent to Pn14 PS) and a minor population that cross-reacts with Pn14 PS. However, the converse has not been found to be true: antibodies elicited by Pn14 PS are not protective against GBSIII bacteria, although desialylation of the GBSIII PS significantly increases the cross-reactivity with Pn14 antibodies [6,7]. Evidence for serotype cross-protection is necessarily indirect and complicated by the fact that cross-reaction of a PS with antibody raised by a different PS does not reliably predict cross-protection: vaccination raises families of antibodies against various PS epitopes, not all of which are of high avidity and thus effective opsonophagocytic (killing) antibodies. Indeed, effective cross-protection between GBSIII and Pn14 has not been demonstrated: antibodies elicited by Pn14 PS (desialylated GBSIII PS) are not protective against GBSIII bacteria and there is considerable evidence that the presence of the terminal sialic acid residue is essential for the elicitation of protective antibodies against GBSIII PS [6–9].

The native PS produced by GBSIII and Pn14 contains between 50 and 300 RU, which is far longer than the epitope bound by an antibody. There has been some effort expended into identification of the minimal epitope for both GBSIII PS and Pn14 PS, with some conflicting results (Figure 1). Originally, on the basis of NMR measurements and molecular modeling, a long 3–4 RU helical conformational epitope for GBSIII PS was proposed (Figure 1a), with the Pn14 PS being comparatively flexible and disordered. The antigenic differences between GBSIII and Pn14 were thus originally attributed to significantly different PS conformations (and hence conformational epitopes), rather than direct interaction of the antibody with the sialic acid side chain in GBSIII PS [9–12]. The hypothesis was that anti-GBSIII antibodies bind the helical GBSIII PS backbone (stabilized by the sialic acid residues on the exterior surface of the helix) and not the sialylated side chain, with a 3 to 4 RU epitope necessary for raising protective antibodies.

However, this hypothesis is challenged by later work that provides evidence that a helical conformational epitope of GBSIII PS is not required for antigen recognition and that the sialic acid participates directly in antibody recognition of GBSIII. Safari et al. investigated the epitope specificity of GBSIII and showed that human anti-GBSIII PS antibodies recognized the linear backbone epitope common to Pn14 PS and GBSIII PS: -Glc-GlcNAc-Gal- (Figure 1b) [7]. However, although conjugates of linear oligosaccharides of GBSIII PS (such as Gal-Glc-GlcNAc, Glc-GlcNAc-Gal, and GlcNAc-Gal-Glc) did evoke specific oligosaccharide antibodies in mice, these antibodies bound neither native nor desialylated GBSIII PS. Therefore, it was assumed that they are too small or flexible to raise antibody and longer chain lengths exhibiting this epitope are required. Furthermore, the recent elucidation of the crystal structure of a short six-sugar GBSIII epitope in complex with a functional antibody showed a branched hexasaccharide functional epitope with sialic acid participating directly in antigen binding [13]. Identification of a short epitope for GBSIII is valuable information for the development of synthetic vaccines, due to the difficulty and cost associated with synthesis of longer oligosaccharides.



**Figure 1.** Schematic representation of the protective epitopes previously identified for GBS PS (left column) and Pn14 PS (right column). GBSIII: (a) 3–4 RU helical conformational epitope postulated for GBSIII PS [10,11]; (b) the linear backbone epitope identified from fragment binding to GBSIII antibodies [14] and (c) a 6-residue epitope identified from a DP2-Fab crystal structure [13]. Pn14: (d) the tetrasaccharide epitope first identified by Safari et al. [7,14] and then by Kurbatova et al. [15] from antibody studies. Structures are depicted using the ESN symbol set [16] with yellow circle: Gal, blue circle: Glc, blue square: GlcNAc, purple diamond: Neu5Ac.

Effective short, branched epitopes have also been identified for the Pn14 PS: Safari et al. demonstrated that one RU of the Pn14 PS (Figure 1c) is essential and sufficient for inducing protective Pn14-specific antibodies: the presence of the trisaccharide branch point in an epitope is crucial for anti-Pn14 antibody recognition and the extra galactose contributes to the immunogenicity of the epitope [14]. Kurbatova et al. recently identified a similar branched tetrasaccharide as the most effective epitope for Pn14 PS [15]. In contrast, linear Pn14 PS fragments were found to be completely ineffective: none of a range of short linear epitopes of the PS backbone were recognized by Pn14 antibodies [7]. Only branched fragments containing the Gal-Glc-(Gal)-GlcNAc moiety were found to provide significant protection.

Identification of the capsular polysaccharide epitopes recognized by protective antibodies (or glycotopes) and their conformation is crucial for understanding the affinity and specificity of carbohydrate-antibody interactions and, ultimately the cross-protection mechanisms. Furthermore, identification of the conformational effect of side chains on PS conformation may usefully inform optimal antigen design and the development of effective conjugate vaccines. However, as direct experimental evidence of the key PS epitopes is difficult to obtain, systematic molecular modeling protocols have been developed to provide a theoretical estimate of carbohydrate conformation and dynamics [17]. In the last decade, a considerable improvement in both carbohydrate force fields [18–21] and computer hardware has facilitated far larger, longer and more accurate computer simulations of polysaccharides than was possible when the previous simulations of GBSIII/Pn14 PS were performed. Therefore, we considered it timely to embark on more extensive modeling of these capsular polysaccharides. Our aim is to shed light on the remaining unanswered questions on the conformations of the GBSIII and Pn14 PS, including the following. Does the GBSIII PS have a helical conformation? Are the conformation and dynamics of the GBSIII PS significantly different to Pn14 PS? What are the likely minimal epitopes for these two PS? Why are some fragments more effective antigens than others?

To answer these questions, we compare the solution conformation and dynamics of the GBSIII and Pn14 PS, as well as the corresponding unbranched saccharide, to determine the effects of the side chains on the polysaccharide backbone conformation. We ran long simulations of 1  $\mu$ s, an order of magnitude more than the previous simulations of 50 ns.

We do not consider O-acetylation of the terminal  $\alpha$ DNeu5Ac sialic acid residues [22], as O-acetylation was not found to be necessary for elicitation of functional antibodies against GBSIII [23]. We find that contrary to previous simulations, the polysaccharide backbone dynamics is almost identical in the GBSIII and Pn14 polysaccharides. The backbone is not helical, but rather has a highly flexible zig-zag conformation, whereas the branch points are relatively inflexible with well-defined conformational epitopes. In contrast, the unbranched PS is highly flexible and conformationally varied. Our results are supported by NMR NOESY experiments performed on the GBSIII polysaccharide.

## 2. Materials and Methods

Our established systematic approach to modeling of polysaccharide antigens involves first determining the preferred conformations of each of the glycosidic linkages in the polysaccharide by calculation of the  $\phi$ ,  $\psi$  potential of mean force (PMF) for the corresponding disaccharides and then progresses to molecular dynamics simulations of three- and six-RU oligosaccharides in aqueous solution to establish the preferred conformations and dynamics of the carbohydrate chains [24–26].

### 2.1. Disaccharide PMF Calculations

We identified the preferred conformations of the each of the glycosidic linkages in isolation by calculation of the potential of mean force (PMF) for rotation about the  $\phi$  and  $\psi$  dihedral angles. PMFs were calculated using the metadynamics [27] routine incorporated into NAMD [28] with the glycosidic linkage torsion angles used as collective variables. For the three-bond (1→6) linkages, we calculated a two-dimensional PMF as a function of  $\phi$  and  $\psi$  only, allowing the  $\omega$  dihedral to rotate freely. All PMF surfaces were calculated in gas-phase, except for the charged  $\alpha$ DNeu5NAc(2→3) $\beta$ DGalp disaccharide, which required simulation in explicit aqueous solution with a neutralizing counter-ion. Gas phase PMFs for uncharged disaccharides have been demonstrated to be a reasonable approximation to solution PMF in a polysaccharide [25,29,30].

### 2.2. Molecular Dynamics Simulations

All simulations were performed with the NAMD molecular dynamics program [28] version 2.12 (employing NAMD CUDA extensions for calculation of long-range electrostatics and nonbonded forces on graphics processing units [31]). Carbohydrates were modeled with the CHARMM36 additive force field for carbohydrates [19,32] and water was simulated with the TIP3P model [33].

Initial configurations of three-repeat (3 RU) and six-repeat unit (6 RU) oligosaccharides for GBSIII and Pn14 PS were built using our in-house CarbBuilder software [34,35] which employs the psfgen tool to create “protein structure” (psf) files for modeling with the CHARMM force field and the NAMD molecular dynamics program. These initial oligosaccharide structures were optimized through 20,000 steps of standard NAMD minimization in vacuum and then solvated (using the *solvate* plugin to the Visual Molecular Dynamics (VMD) [36] analysis package) in a periodic cubic unit cell with randomly distributed sodium ions to electrostatically neutralize the system.

All MD simulations were preceded by a 30,000 step minimization phase, with a temperature control and equilibration regime involving 10 K temperature reassignments from 10 K culminating in a maximum temperature of 300 K. Equations of motion were integrated using a Leap-Frog Verlet integrator with a step size of 1 fs and periodic boundary conditions. Simulations were performed under isothermal-isobaric (nPT) conditions at 300 K maintained using a Langevin piston barostat [37] and a Nose-Hoover [38,39] thermostat.

Long-range electrostatic interactions were treated using particle mesh Ewald (PME) summation, with  $\kappa = 0.20 \text{ \AA}^{-1}$  and 1  $\text{\AA}$  PME grid spacing. Non-bonded interactions were truncated with a

switching function applied between 12.0 and 15.0 Å to groups with integer charge. The 1–4 interactions were not scaled, in accordance with the CHARMM force field recommendations.

Each metadynamics simulation comprised a 1500 ns MD simulation, with a Gaussian hill height of 0.5 and width of 2.5 degrees. Structures were collected at intervals of 250 ps for analysis. For the solution simulation, the  $\alpha$ DNeu5NAc(2→3) $\beta$ DGalp disaccharide was placed in the center of a cubic box with sides of 30 Å. The box was filled with approximately 2500 water molecules and a single Na<sup>+</sup> counter ion.

The 3RU strands were placed in the center of a cubic water box with sides of 60 Å, while the 6RU strands employed a box of length 80 Å. The Pn14 strands were solvated with 6810 (3RU) and 16,313 (6 RU) water molecules, while the GBSIII strands used 6790 (3 RU) and 16,279 (6 RU) water molecules. The GBSIII systems were neutralized with 3 (3 RU) and 6 (6 RU) Na<sup>+</sup> ions. In each case, the system was equilibrated 0.03 ns with a cycled temperature increase from 0 K to 300 K in 10 K increments with each cycle commencing with a 10,000 step energy minimization followed by a 0.001 ns MD simulation at the specified temperature until 300 K. The 3 RU MD simulations ran for 250 ns and the 6 RU simulations ran for 1 μs.

### 2.3. Data Analysis

In this work, two-bond (1→X) glycosidic linkages are defined by the torsion angles  $\phi = H1'-C1'-OX-CX$  and  $\psi = C1'-OX-CX-HX$ . The (2→3) glycosidic linkages use  $\phi = C1'-C2'-O3-C3$  and  $\psi = C2'-O3-C3-H3$ . These definitions for  $\phi$  and  $\psi$  are analogous to  $\phi_H$  and  $\psi_H$  in IUPAC convention. For the (1→6) glycosidic linkage, the three dihedral angles are defined as  $\phi = H1'-C1'-O6-C6$ ,  $\psi = C1'-O6-C6-C5$  and  $\omega = O6-C6-C5-O5$ .

Analysis of the simulations used time series frames 25 ps apart, discarding the first 100 ns as equilibration. Molecular conformations extracted from the MD simulations were depicted with VMD, where necessary using the PaperChain visualization algorithm for carbohydrates [40] to highlight the hexose rings. Dihedral angles from the simulations and the DP2-Fab complex crystal structure (PDB ID code 5M63) were extracted using VMD's Tcl scripting interface and statistical values calculated with in-house Python scripts. The DP2 fragment comprises RU2 with a terminal 2,5-anhydro-D-Man. The GBSIII 6RU simulation conformations and the DP2-Fab complex crystal structure were aligned for comparison on the ring atoms of the  $\alpha$ DNeu5NAc(2→3) $\beta$ DGalp(1→4) $\beta$ DGlcNAc branch point.

Conformations from both 3RU trajectories were clustered using VMD's internal *measure cluster* command to calculate clusters according to the quality threshold algorithm [41]. Frames were aligned on the GlcNAc residues in RU1 to RU5 of the 6RU chains. Clustering was then performed with a cut-off of 7 Å on an RMSD fit to the atoms in the backbone residues of RU1 to RU5.

### 2.4. NMR Analysis

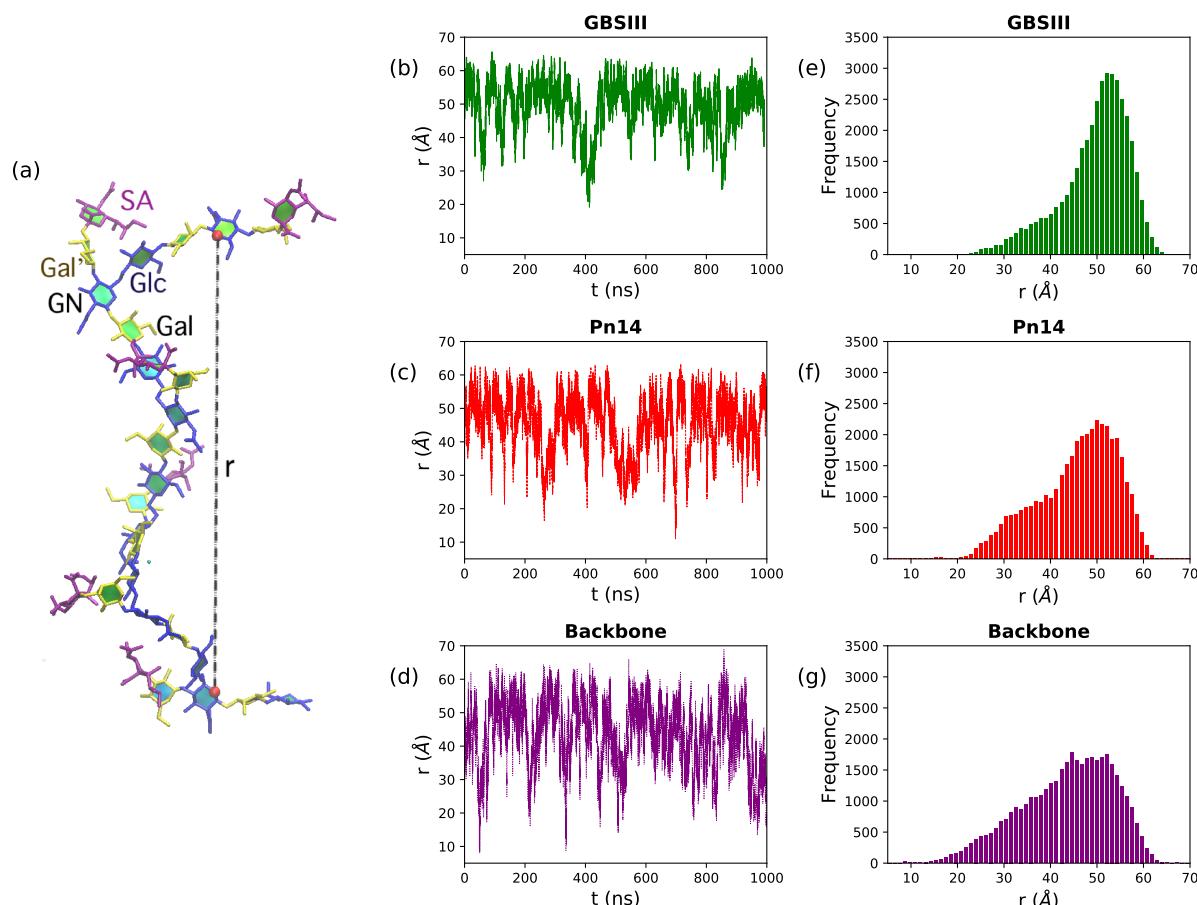
The GBSIII polysaccharide sample (10 mg) was lyophilized and exchanged twice with 99.9% D<sub>2</sub>O (Sigma Aldrich, Pty. Ltd., Johannesburg, South Africa), then dissolved in 600 μL of D<sub>2</sub>O and introduced into a 5 mm NMR tube for data acquisition. 1D <sup>1</sup>H and <sup>13</sup>C and 2D, COSY, TOCSY, NOESY, HSQC, HMBC and hybrid HSQC-TOCSY and HSQC-NOESY spectra were obtained using a Bruker Avance III 600 MHz NMR spectrometer (Bruker BioSpin AG, Fällanden, Switzerland) equipped with a BBO Prodigy cryoprobe and processed using standard Bruker software (Topspin 3.2). The probe temperature was set at 343 K. The 2D TOCSY experiment was performed using a mixing time of 180 ms and the 1D variants using a mixing time of 200 ms. The 2D NOESY experiment was performed using a mixing times of 300 and 500 ms and the 1D variants using mixing times of 300, 400 and 500 ms. The <sup>1</sup>H-<sup>13</sup>C HSQC and HMBC experiments were optimized for J = 145 Hz and 8 Hz, and the HSQC-TOCSY and HSQC-NOESY experiments were recorded using mixing times of 120 and 250 ms, respectively. 2D experiments were recorded using non-uniform sampling: 50% for homonuclear and 25% for heteronuclear experiments. Spectra were referenced relative to the H3ax/C3 signal of terminal sialic acid: <sup>1</sup>H at 1.79 ppm and <sup>13</sup>C at 40.68 ppm [42].

### 3. Results

Our 1  $\mu$ s simulations of 6 RU show that polysaccharides with the  $\rightarrow 6)\beta\text{D}\text{GlcNAc}(1\rightarrow 3)\beta\text{D}$   $\text{Galp}(1\rightarrow 4)\beta\text{D}\text{Glc}(1\rightarrow$  backbone are all extremely flexible, with none of the well-defined conformations we have found in other, linear bacterial polysaccharides [24,25]. However, a closer inspection of the range of motion for GBSIII PS, Pn14 PS and the unbranched backbone PS reveal a common dominant conformational habit, as well as significant differences in flexibility.

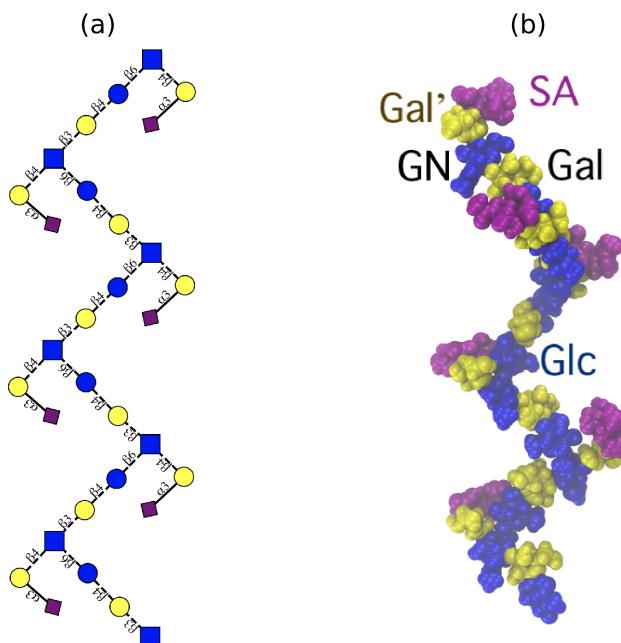
#### 3.1. PS Chain Conformations

In Figure 2, we quantify range of motion for these very flexible polysaccharides with the simple end-to-end distance,  $r$ , measured between the GlcNAc O5 atoms in repeat units RU1 and RU6 (illustrated on a sample conformation in Figure 2a). Comparison of the time series (Figure 2, center column) and distribution (right column) of  $r$  shows that GBSIII PS is the most conformationally defined, and the unbranched polysaccharide backbone the most flexible, of the three molecules. Only 12% of the GBSIII PS conformations have  $r < 40 \text{ \AA}$ , as compared to 25% for Pn14 and 33% for the unbranched backbone. Overall, the unbranched backbone is less extended (more globular) than the branched polysaccharides and shows more rapid conformational transitions (compare time series plots in Figure 2b–d). The corresponding mean squared end-to-end distances for the simulations reflect this trend: 2505  $\text{\AA}$  (GBSIII), 2157  $\text{\AA}$  (Pn14) and 1993  $\text{\AA}$  (backbone).



**Figure 2.** Polysaccharide end-to-end distance distributions for the 6RU PS. (a) The molecular end-to-end distance,  $r$ , is defined as the distance between O5 atoms in GlcNAc residues in RU1 and RU6, here shown on the GBSIII PS. The center column shows the  $r$  time series for (b) GBSIII PS, (c) Pn14 PS and (d) unbranched backbone PS. The corresponding distance distributions are in the rightmost column for (e) GBSIII PS, (f) Pn14 PS and (g) the unbranched backbone PS.

Although flexible, the GBSIII PS and Pn14 PS have a similar dominant overall conformation with  $48 < r < 58 \text{ \AA}$ . Here the PS backbone has an overall “zig-zag” arrangement, bending at the 1→6 linkage, with the side chains exposed at the branch points, as shown in the GBSIII schematic in Figure 3a and the sample conformation in Figure 3b.

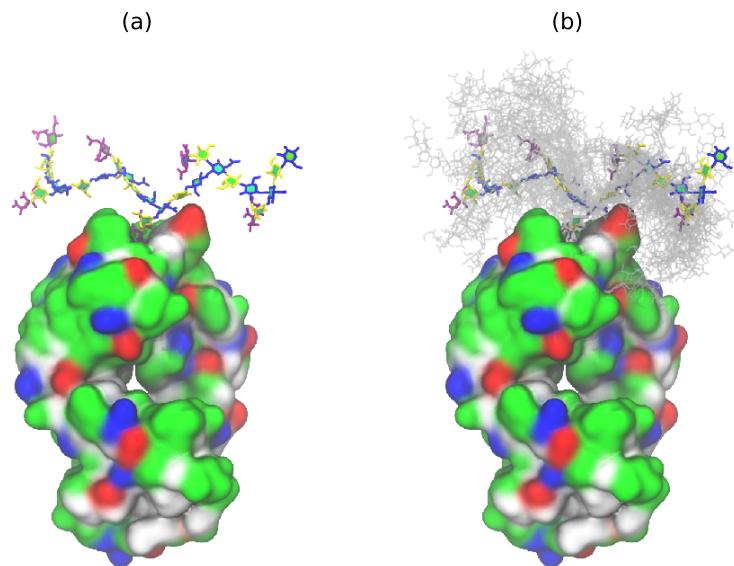


**Figure 3.** The ‘zig-zag’ conformational habit of GBSIII and Pn14 PS. (a) Schematic representation with the ESN symbol set [16] (yellow circle: Gal, blue circle: Glc, blue square: GlcNAc, purple diamond: Neu5Ac) and (b) a representative 6RU GBSIII PS simulation snapshot. Residues are highlighted as follows: Glc and GlcNAc: blue; Gal: yellow; sialic acid: purple.

This conformation is in very good agreement with the new antibody binding model recently proposed for GBSIII PS (see Figure 5 in Carboni et al. [13]). Figure 4 shows the GBSIII PS conformations from our 6RU simulation superimposed on the solved crystal structure for DP2-Fab. The zig-zag conformation of the backbone exposes the sialic acid side chain for binding (Figure 4a) with the flexible backbone held well away from the antibody over the course of the simulation (Figure 4b). Conversely, we find no evidence in our simulations to support the helical model for GBSIII PS previously proposed [10,11]. Further, in the zig-zag conformation, the polysaccharide backbone is relatively inaccessible to antibody binding.

However, although the 6RU backbone shows the same dominant conformation for GBSIII PS, Pn14 PS and the unbranched backbone PS, they have a significant difference in flexibility. Clustering of the simulation conformations shows that, while GBSIII PS is in this general zig-zag conformational family for 85% of the simulation, Pn14 PS is more flexible and is in a zig-zag for 78% of the simulations, moving occasionally into alternative, more bent conformations (with smaller  $r$ ). The unbranched backbone polysaccharide is the most disorganized, with the zig-zag appearing for 66% of the simulation and showing significant sub-populations of globular conformations (with  $r < 25 \text{ \AA}$ ).

The source of the conformational differences between the polysaccharides is a conformational constraint on the  $\beta\text{DGlc}(1\rightarrow6)[\beta\text{DGal}(1\rightarrow4)]\beta\text{DGlcNAc}$  linkage in GBSIII PS and, to a lesser extent, in Pn14 PS. The conformation and flexibility of the other backbone linkages ( $\beta\text{DGlcNAc}(1\rightarrow3)\beta\text{DGalp}$  and  $\beta\text{DGalp}(1\rightarrow4)\beta\text{DGlcP}$ ) is very similar in all three polysaccharides—see a detailed analysis in Appendix A, particularly Figure A1. However, for the  $\beta\text{DGlcP}(1\rightarrow6)\beta\text{DGlcNAc}$  linkage, proximity of the side chains impose a restriction on rotation, as follows.

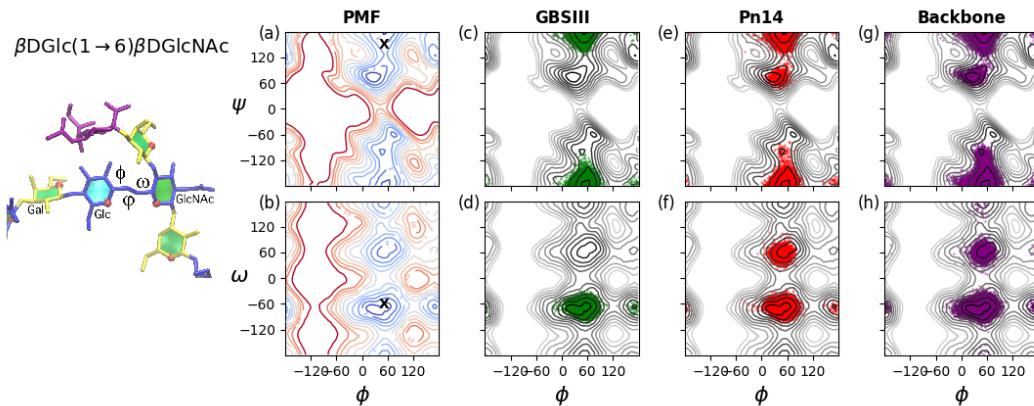


**Figure 4.** GBSIII 6RU structures superimposed on the bound branch point of the DP2-Fab crystal structure from Carboni et al. [13] (PDB ID code 5M63): (a) a single representative structure and (b) superimposed conformational snapshots at 12.5 ns intervals (with the first 125 ns discarded as equilibration).

### 3.2. Conformations of the 1→6 Linkage

Figure 5 provides a comparison of the flexibility of the three-bond glycosidic linkage in a  $\beta$ DGlc(1→6) $\beta$ DGlcNAc disaccharide and the GBSIII PS, Pn14 PS and unbranched backbone PS. The three dihedral angles describing rotation about the three bonds in the (1→6) glycosidic linkage are here defined as  $\phi = H1'-C1'-O6-C6$ ,  $\psi = C1'-O6-C6-C5$  and  $\omega = O6-C6-C5-O5$ . These dihedrals are labeled on the GBSIII fragment in Figure 5, left. A PMF energy surface for the (1→6) glycosidic linkage has, therefore, three dimensions. However, for ease of visualization and comprehension, we show only 2D projections of the 3D volume: the  $\phi, \psi$  PMF (Figure 5a) and the  $\phi, \omega$  PMF surface (Figure 5b) for a  $\beta$ DGlc(1→6) $\beta$ DGlcNAc disaccharide in the gas phase. These PMF surfaces illustrate the range of motion possible for an unrestrained linkage. The  $\phi, \psi$  PMF in Figure 5a reveals that the central bond in the linkage described by the  $\psi$  dihedral is relatively flexible, with multiple minima within the 2 kcal·mol<sup>-1</sup> contours: a global minimum conformation at  $\psi = 71^\circ$ , a secondary *anti* minimum conformation at  $\psi = -179^\circ$  and a tertiary minimum in the  $\psi = -60$  region. In contrast, rotation about the  $\phi$  dihedral is much more constrained, with a broad global minimum around  $0^\circ < \phi < 75^\circ$  and a narrow secondary well at  $\phi = 160^\circ$ . The  $\phi, \omega$  PMF surface in Figure 5b confirms these minima for  $\phi$  and shows the expected three minima for the  $\omega$  dihedral, which are by convention termed *gg* ( $\omega \approx -60^\circ$ ), *gt* ( $\omega \approx 60^\circ$ ) and *tg* ( $\omega \approx 180^\circ$ ). Glucopyranosides are expected to be primarily in the *gg* conformation, with the *gg:gt:tg* population ratios approximately 6:4:0 [43].

As a comparison to the unrestrained disaccharide, the range of motion actually explored during the simulations of the 6RU oligosaccharides is revealed by the corresponding dihedral time series scatter plots superimposed on the disaccharide PMFs. The restriction of the  $\beta$ DGlc(1→6) $\beta$ DGlcNAc linkage in GBSIII is apparent in the time series scatterplots for the 6RU GBSIII oligosaccharide strand (Figure 5, second column). These confirm a single dominant conformer for GBSIII, with average torsion angle values  $\phi, \psi, \omega = 51^\circ, -175^\circ, -67^\circ$  (standard deviations = 10, 11, 8 respectively). These angles are compatible with the GBSIII DP2-Fab complex crystal structure where this linkage has  $\phi, \psi, \omega = 51^\circ, 140^\circ, -70^\circ$  (indicated by 'X' on the PMF plots in Figure 5a,b). The central 1→6 linkage in GBSIII PS is in this dominant conformation for 100% of the simulation and the  $\omega$  dihedrals remain in the low-energy *gg* conformation for the duration.



**Figure 5.** Rotation of the three-bond  $\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$  linkages in GBSIII, Pn14 and the unbranched backbone PS. The  $\phi$ ,  $\psi$  and  $\omega$  angles describing the orientation of this linkage are labeled on the GBSIII fragment shown on the left. Contoured 2D (a)  $\phi$ ,  $\psi$  PMF and (b) the  $\phi$ ,  $\omega$  PMF surfaces for a  $\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$  disaccharide in the gas phase illustrate the range of motion possible for an unrestrained linkage. Contours are drawn at intervals of 1  $\text{kcal}\cdot\text{mol}^{-1}$  to a maximum of 12  $\text{kcal}\cdot\text{mol}^{-1}$  and 'X' markers indicate dihedral angle values from the six-sugar GBSIII epitope identified in the DP2-Fab crystal structure [13]. The range of motion actually explored by the 6RU oligosaccharides is demonstrated with time series scatter plots of the two central  $\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$  linkages in the 6RU strands superimposed on the PMFs as follows: (c) GBSIII  $\phi$ ,  $\psi$ ; (d) GBSIII  $\phi$ ,  $\omega$ ; (e) Pn14  $\phi$ ,  $\psi$ ; (f) Pn14  $\phi$ ,  $\omega$ ; (g) unbranched backbone  $\phi$ ,  $\psi$ ; and (f) unbranched backbone  $\phi$ ,  $\omega$ .

The average torsion angle values for Pn14 PS are  $\phi$ ,  $\psi$ ,  $\omega = 50^\circ$ ,  $180^\circ$ ,  $-62^\circ$  (standard deviations = 12, 18, 26 respectively). The central  $\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$  linkage is in the dominant conformation (equivalent to the GBSIII PS) for 91% of the simulation. The increase in standard deviation for the  $\phi$  and  $\omega$  dihedrals relative to GBSIII PS is indicative of increased flexibility and alternative conformations of the  $\omega$  dihedral, which shows a ratio of approximate 10:1 of gg to gt conformations, with no tg conformations appearing. Therefore, the 1 $\rightarrow$ 6 linkage shows a slight increase in flexibility relative to GBSIII PS, which is cumulative with increasing chain length. This reflects the increased constraint on the backbone imposed by the terminal sialic in the GBSIII side chain.

The 1 $\rightarrow$ 6 linkage is markedly more flexible in the unbranched backbone saccharide and the zig-zag conformation drops to 86% of the simulation. The average torsion angle values for the unbranched backbone are  $\phi$ ,  $\psi$ ,  $\omega = 50^\circ$ ,  $-179^\circ$ ,  $-69^\circ$  (standard deviations = 15, 22, 30 respectively). In particular, the  $\omega$  dihedral has an 8% population of gt conformations and a small population of tg conformations. In addition, both the  $\phi$  and  $\psi$  dihedrals show a broader range of rotation.

Interestingly, we saw no effect on conformation and dynamics of increasing chain length. Our simulations of 3RU for GBSIII and Pn14 PS showed the same dihedral populations as the middle linkages in the 6RU simulations (data not shown). Further, we performed a 250 ns simulation of the effective branch epitope previously identified (Figure 1d) [14,15]. This molecule shows the same conformation of the branch point and other dihedrals as Pn14 PS (see Appendix B, Figure A2) and is thus a faithful representation of the branch point in the Pn14 PS.

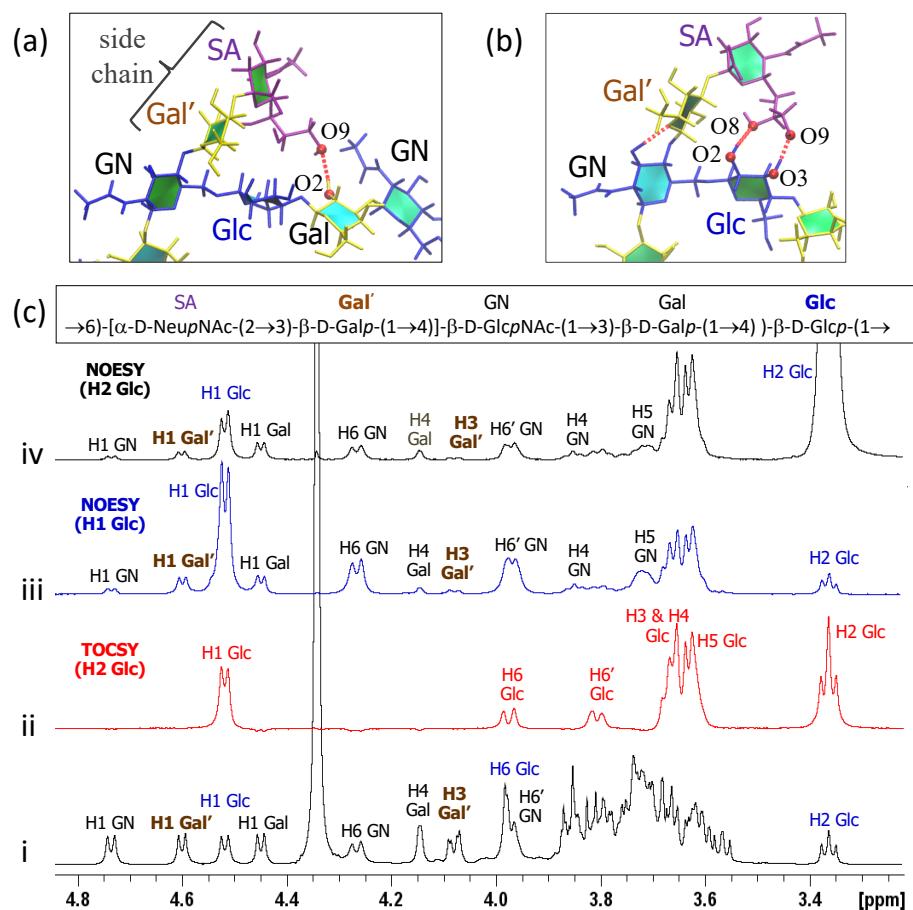
### 3.3. Inter-Residue Atomic Contacts in GBSIII

In GBSIII PS, interactions between the side chain and the backbone residues stabilize the conformation of the 1 $\rightarrow$ 6 linkage. We compare the simulation data with NMR NOESY experiments on the GBSIII PS, focussing on inter-residue NOEs that are diagnostic for close proximity of the side chain to the backbone.

In GBSIII PS, hydrogen bonding interactions between the side chain sialic acid (SA) and side chain Gal (Gal') and the backbone Glc and Gal stabilize the conformation of the 1 $\rightarrow$ 6 linkage, as follows. Transient hydrogen bonds occur between the bound SA O9 and backbone Gal O2

(shown in Figure 6a) as well as SA O9 – Glc O3 and SA O8 – Glc O2 hydrogen bonds (occasionally simultaneous, as in Figure 6b). The crystal structure for the GBSIII PS fragment in DP2-Fab shows a close hydrogen bond between SA O9 and the Glc O6, with distance of 2.8 Å [13]. We do not find this hydrogen bond in our simulations, although the atoms come within 4 Å of each other. This is due to the fact that the sialic acid is in an alternative (higher energy) conformation in the crystal structure—possibly stabilized by interactions with the antibody. The conformation of this sialyl linkage is known to have no fixed standard conformation and to be heavily dependent on the molecular environment [44]. See Appendix A and Figure A1e for more detailed analysis of the sialic acid conformation. Unfortunately, considerable overlap between the signals of the NAc group of SA and GlcNAc precluded investigation of contact between the sialic acid side chain and the backbone in GBSIII, as previously reported [13].

Additional hydrogen bonds occur between the Gal' O2 and backbone Glc O3 in both the GBSIII PS and Pn14 PS simulations. In this case, NMR signals are better resolved: a series of 1D and 2D NOESY experiments gave key H1 Glc crosspeaks to H4, H5 and both H6s of the neighboring GlcNAc (consistent with the  $\beta$ DGlc(1→6)[ $\beta$ DGal(1→4)] $\beta$ DGlcNAc linkage) as well as small additional correlations to the main chain repeating unit (Figure 6c(iii)). In particular, peaks for H1 Glc to H1 Gal' as well as H1 Glc to H3 Gal' provide evidence for close proximity of the side chain to the backbone and are consistent with our simulations. This is corroborated by 2D NOESY cross peaks, as well as peaks from the well-resolved H2 of Glc at 3.36 ppm to H1 and H3 of Gal' (Figure 6c(iv)).



**Figure 6.** Inter-residue contacts in GBSIII PS. (a) Example of SA O9–Gal O2 hydrogen bond between the side chain and the backbone residues. (b) Example of simultaneous SA O9–Glc O3 and SA O8–Glc O2 hydrogen bonds. (c) Overlay of 1D NMR experiments of GBSIII PS showing (i) 1D proton spectrum; (ii) 1D TOCSY (200 ms) with irradiation of H2 Glc; (iii) 1D NOESY (500 ms) with irradiation of H1 Glc; and (iv) 1D NOESY (500 ms) with irradiation of H2 Glc.

#### 4. Discussion

This modeling study shows that the polysaccharide backbone conformation is very similar in the GBSIII and Pn14 polysaccharides, albeit with increased flexibility for Pn14 PS. Both GBSIII and Pn14 PS exhibit a constrained branch epitope and a flexible zig-zag backbone. The zig-zag conformation is in remarkable agreement with the binding model for GBSIII PS proposed by Carboni et al. [13]. The  $\beta$ DGlc(1→6)[ $\beta$ DGal(1→4)] $\beta$ DGlcNAc branch point is a relatively rigid and exposed component of the flexible backbone, and thus a likely site for antibody binding. This stationary branch point is in agreement with preclinical evidence that has indicated that the Pn14 branching element  $\beta$ DGalp(1→4)] $\beta$ DGlcNAc is necessary and sufficient for induction of an effective antibody response [14]. Further, the lack of accessibility of the backbone in GBSIII and Pn14 PS to antibody binding together with increased flexibility and alternative conformations of the unbranched backbone provides a rational for why antibodies to linear oligosaccharide fragments of GBSIII PS (such as Gal-Glc-GlcNAc, Glc-GlcNAc-Gal, and GlcNAc-Gal-Glc) bound neither GBSIII PS nor Pn14 PS [7]. In addition, the lack of extended conformational epitopes supports experimental data that finds short epitopes to be effective for GBSIII and Pn14, epitopes comprising solely their respective branch points and side chain (i.e.,  $\beta$ DGlc(1→6)[ $\alpha$ DNeu5Ac(2→3) $\beta$ DGal(1→4)] $\beta$ DGlcNAc and  $\beta$ DGlc(1→6)[ $\beta$ DGal(1→4)] $\beta$ DGlcNAc, respectively). This is consistent with the DP2-Fab crystal structure elucidated by Carboni et al.: the binding involves the branch and does not require a conformational epitope [13].

Kurbatova et al. identified a  $\beta$ DGalp(1→4) $\beta$ DGlc(1→6)[ $\beta$ DGalp(1→4)] $\beta$ DGlcNAc tetrasaccharide as the most effective epitope for Pn14 PS, as compared to hexa- and octasaccharides in a mouse model [15]. Our work provides a possible explanation for this: vaccination with this primary epitope could raise a single class of effective antibody, as opposed to a family of antibodies raised by the more flexible hexa- and octasaccharides, thus making this short chain more protective. In contrast, the hexa- and octasaccharides could potentially present epitopes not present in the polysaccharide, as seen in the unbranched backbone.

The flexible zig-zag conformation and stationary branch point were not identified for the Pn14 PS in the earlier study. In contrast to the small 8% population of the *gt* conformation in Pn14 that we see in these simulations, in the 5RU simulations with the AMBER force field by González-Outeiriño et al. the Pn14 oligosaccharide showed a nearly 50:50 mixture of the *gg* to *gt* rotamers, whereas the GBSIII PS remained almost constantly in the *gg* conformation. This discrepancy was attributed to an anomaly arising from insufficient equilibration in the 50 ns simulations [10]. Therefore, the extreme flexibility of Pn14 PS and disordered structure as compared to GBSIII PS suggested by this prior work could be a consequence of the increased flexibility of this dihedral in the AMBER force field as compared to the CHARMM force field. In the presence of competing computational models, experimental evidence is key. The DP2-Fab crystal structure provides evidence that a short strand can effectively bind to an antibody and that a helical conformation of the backbone is not necessary for antigen binding. In addition, we observed inter-residue NMR NOEs for the GBSIII PS that are consistent with our simulations.

The branches of the GBSIII and Pn14 PS both present common terminal glycan epitopes: 3-Sialyl-N-acetyllactosamine is present in human biofluids and is a common sequence terminating *N*- and *O*-glycans on the surface of all mammalian cells, while the shorter  $\beta$ DGalp(1→4)] $\beta$ DGlcNAc constitutes the ubiquitous LacNAc building block in mammalian *N*-linked protein glycans. In general, mimics of mammalian cell surface residues can subvert the immune system.

This work thus suggests a strategy for bacterial evasion of the host immune system: expression of a very flexible backbone that is shielded from the immune system by both its zig-zag conformation and hyper-mobility, combined with exposed, inflexible branches that present human-mimic epitopes. This strategy should be considered with other polysaccharides with similar backbones and human-like branch epitopes, where it is likely that the branch points will be key to immunogenicity.

In summary, this work demonstrates how small structural features such as side chains can alter the conformation of a polysaccharide by restricting rotation around backbone linkages. It also highlights the explanatory power of simulation for optimal antigen design and the development of effective conjugate vaccines.

**Author Contributions:** Conceptualization, M.M.K. and N.R.; Funding acquisition, M.M.K. and N.R.; Investigation, M.M.K. and N.R.; Methodology, M.M.K. and N.R.; Project administration, M.M.K.; Resources, N.R.; Software, M.M.K.; Validation, N.R.; Visualization, M.M.K.; Writing—original draft, M.M.K.; Writing—review & editing, M.M.K. and N.R.

**Funding:** This research was funded by The South African National Research Foundation grant number 103805 and grant number 86038 (NMR equipment).

**Acknowledgments:** Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: <http://hpc.uct.ac.za>. The authors thank PATH and the Biovac Institute for providing the GBSIII polysaccharide sample.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

NMR	Nuclear Magnetic Resonance
PS	polysaccharide
RU	repeat unit

## Appendix A. Analysis of Glycosidic Linkages

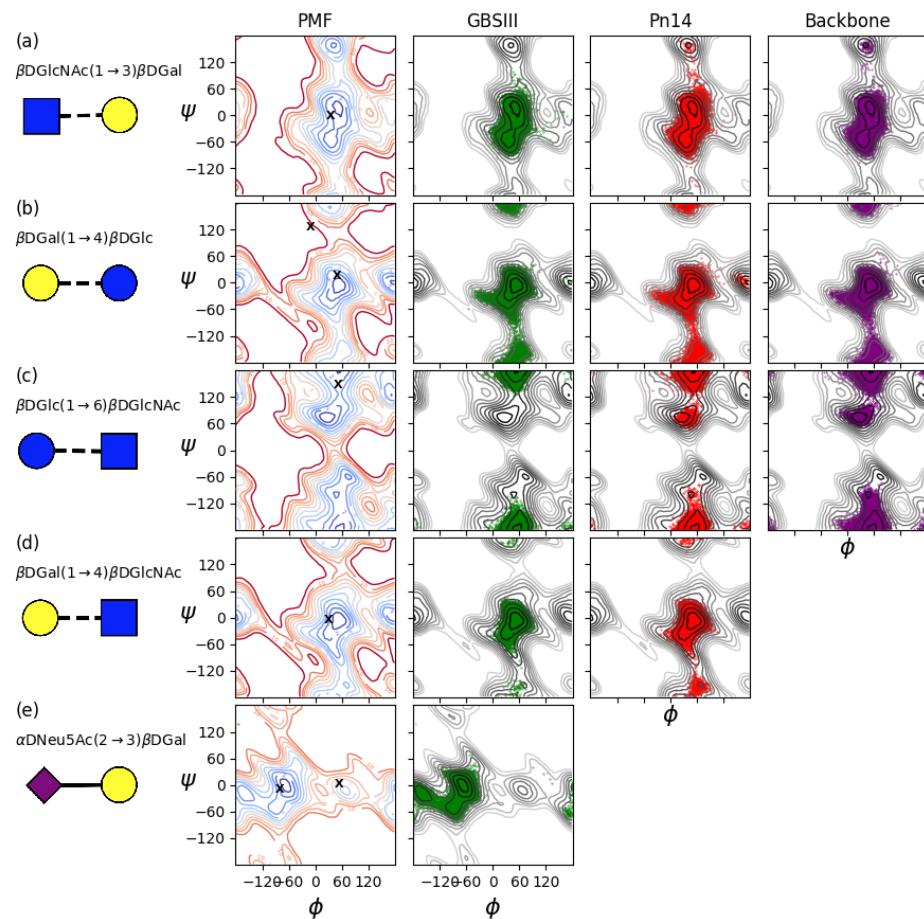
The GBSIII glycosidic linkages can be conveniently classified into the three linkages which make up the common polymer backbone and the two glycosidic linkages which comprise the side chains. The contoured disaccharide  $\phi, \psi$  PMF energy surfaces for disaccharides representing all the constituent linkages in GBSIII (Pn14 and the backbone being subsets) are shown in the left column of Figure A1. The three linkages which make up the common polymer backbone are shown in the top three rows and the two glycosidic linkages which comprise the side chains in the bottom two rows.

### Appendix A.1. $\beta$ DGlcNAc(1→3) $\beta$ DGal

The  $\beta$ DGlcNAc(1→3) $\beta$ DGal backbone linkage is flexible: the  $\phi, \psi$  PMF (Figure A1a, left) shows a broad, shallow central well, encompassing the global energy minimum at  $\phi, \psi = 54^\circ, 16^\circ$ , and a secondary syn-syn minimum at  $\phi, \psi = 44^\circ, -54^\circ$  with  $\Delta G = 1 \text{ kcal}\cdot\text{mol}^{-1}$ . This is compatible with the GBSIII DP2-Fab complex crystal structure [13], where this linkage has angles equivalent to  $\phi, \psi = 35^\circ, -9^\circ$  (indicated by 'X' on the PMF plot). In addition, the tertiary anti- $\psi$  minimum is also relatively low in energy ( $\Delta G = 1.5 \text{ kcal}\cdot\text{mol}^{-1}$ ), albeit with a high energy barrier to rotation of  $>6 \text{ kcal}\cdot\text{mol}^{-1}$ .

The time series scatter plots for this linkage in GBSIII, Pn14 and the unbranched backbone show that all 6RU oligosaccharides explored the full range of linkage conformations in the central well, with the average  $\phi, \psi$  value showing a shift to favor the secondary well in solution (Table A1). This is in agreement with the reported behaviour of this linkage in the sialyl Lewis X pentasaccharide (sLe<sup>X</sup>-5) for a 500 ns MD simulation at 37 °C in explicit water [45]. However, transitions to the anti- $\psi$  conformation occurred rarely throughout the duration simulations. This is in contrast to sLe<sup>X</sup>-5, where the unconstrained terminal  $\beta$ DGal showed a significant population of the anti- $\psi$  conformation. This difference is likely due to the more constrained environment in the oligosaccharides, as our  $\beta$ DGlcNAc(1→3) $\beta$ DGal vacuum PMF is very similar to the energy map calculated for this linkage in sLe<sup>X</sup>-5. Indeed, the previous 5RU simulations of GBSIII and Pn14 by González-Outeiriño et al. showed an even smaller range of motion for the  $\beta$ DGlcNAc(1→3) $\beta$ DGal linkage, encompassing only the vacuum global energy minimum (Table A1). This can be attributed to limited conformational

sampling in the earlier 50 ns simulations, as transitions between the two minima in the central well occurred on a 20- to 50 ns time scale in our oligosaccharide simulations.



**Figure A1.** The left column shows contoured disaccharide  $\phi$ ,  $\psi$  PMF surfaces for (a)  $\beta\text{D}\text{GlcNAc}(1 \rightarrow 3)\beta\text{D}\text{Gal}$ , (b)  $\beta\text{D}\text{Gal}(1 \rightarrow 4)\beta\text{D}\text{Glc}$ , (c)  $\beta\text{D}\text{Glc}(1 \rightarrow 6)\beta\text{D}\text{GlcNAc}$ , (d)  $\beta\text{D}\text{Gal}(1 \rightarrow 4)\beta\text{D}\text{GlcNAc}$  and (e)  $\alpha\text{D}\text{Neu5Ac}(2 \rightarrow 3)\beta\text{D}\text{Gal}$ . The PMF for the charged  $\alpha\text{D}\text{Neu5Ac}(2 \rightarrow 3)\beta\text{D}\text{Gal}$  disaccharide is in solution, all others are gas-phase. Contours are drawn at intervals of 1  $\text{kcal} \cdot \text{mol}^{-1}$  to a maximum of 12  $\text{kcal} \cdot \text{mol}^{-1}$  and 'X' markers indicate dihedral angle values from the six-sugar GBSIII epitope in complex with Fab crystal structure [13]. Scatter plots of the corresponding two central linkages in the 6RU oligosaccharide strands for the last 900 ns of simulation time are superimposed on the PMFs for 6RU of GBSIII (column 2, green), Pn14 (column 3, red) and the unbranched backbone (right column, purple).

**Table A1.** Average values of the glycosidic dihedrals in degrees for the two middle repeating units of 6 RU of GBSIII and Pn14 and the unbranched backbone, including comparison with González-Outeiriño et al. [10] (standard deviations in parenthesis). \* Note that the 1→6 linkage comprises three diehdral.

	PMF $\phi, \psi, \omega$	GBSIII $\phi, \psi, \omega$	Pn14 $\phi, \psi, \omega$	Backbone
$\beta\text{D}\text{GlcNAc}(1\rightarrow 3)\beta\text{D}\text{Gal}$ González-Outeiriño et al.	54, 16	45 (12), -22 (25) 42 (13), 2 (18)	45 (13), -20 (27) 41 (12), 1 (19)	44 (12), -23 (26)
$\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{Glc}$ González-Outeiriño et al.	56, -1	47 (14), -14 (40) 47 (12), 8 (12)	49 (15), -14 (37) 43 (12), -7 (16)	49 (14), -17 (46)
$\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$ * González-Outeiriño et al.	26, 71	51 (10), -175 (11), -67 (8) 41 (10), -159 (22), -67 (9)	50 (12), -180 (18), -62 (26) 40 (14), -167 (31), -65 (10)	50 (15), -179 (22), -69 (30)
$\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{GlcNAc}$ González-Outeiriño et al.	49, 1	50 (10), 0 (14) 39 (10), 5 (10)	50 (11), -2 (18) 43 (11), 5 (11)	-
$\alpha\text{D}\text{Neu5Ac}(2\rightarrow 3)\beta\text{D}\text{Gal}$ González-Outeiriño et al.	-66, 1	-70 (21), -8 (19) -173 (14), -11 (11)	-	-

#### Appendix A.1.1. $\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{Glc}$

The  $\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{Glc}$  linkage is also flexible. The  $\phi, \psi$  PMF (Figure A1b, right) has a single central syn-syn well with the global energy minimum at  $\phi, \psi = 56^\circ, -1^\circ$  and both a narrow anti- $\phi$  minimum ( $\Delta G = 0.5 \text{ kcal}\cdot\text{mol}^{-1}$ ) and an anti- $\psi$  minimum ( $\Delta G = 3.6 \text{ kcal}\cdot\text{mol}^{-1}$ ). In the crystal structure of the six-sugar GBSIII epitope in complex with Fab, this linkage has angles equivalent to  $\phi, \psi = -14^\circ, 6^\circ$  and  $\phi, \psi = 45^\circ, -6^\circ$  (indicated by 'X's on the PMF plot). The 6RU simulations of GBSIII and Pn14 show sub-populations of anti- $\psi$  conformations: transitions to this conformation are infrequent and require long simulation times to become apparent. Indeed, anti- $\psi$  conformations of this linkage were not reported for the previous 50 ns simulation, which accounts for the smaller  $\psi$  standard deviation in the average value reported by González-Outeiriño et al. (Table A1).

#### Appendix A.1.2. $\beta\text{D}\text{Glc}(1\rightarrow 6)\beta\text{D}\text{GlcNAc}$

This linkage is discussed in detail in the main text.

#### Appendix A.2. $\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{GlcNAc}$

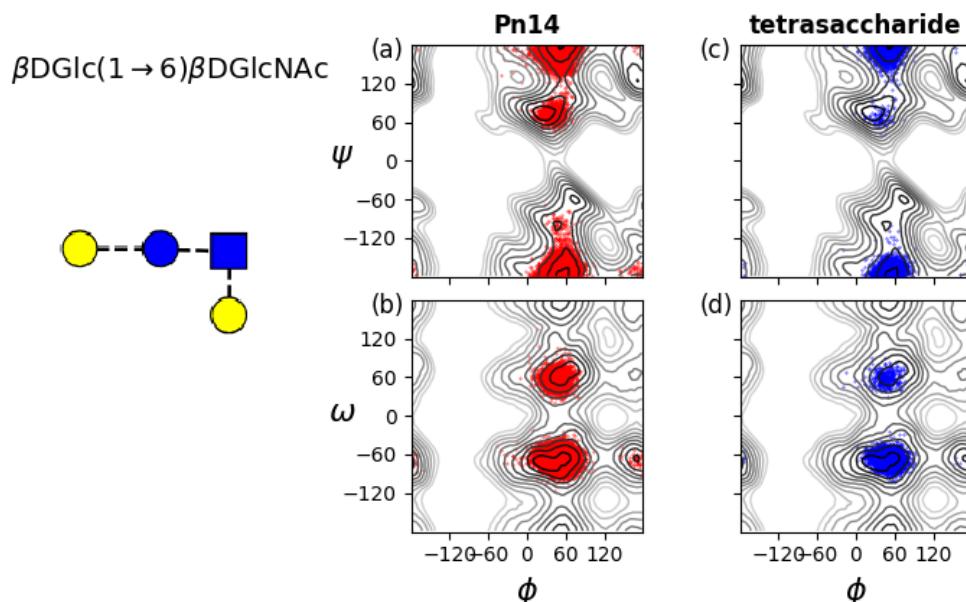
Comparison of the PMF maps reveals that addition of a N-Acetyl group restricts the range of motion of the  $\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{GlcNAc}$  linkage (Figure A1d, left) somewhat as compared to the  $\beta\text{D}\text{Gal}(1\rightarrow 4)\beta\text{D}\text{Glc}$  linkage (Figure A1b, left). The energy of the anti- $\psi$  minimum is slightly lowered and the anti- $\phi$  minimum raised. This is reflected in the narrower range of rotation in the simulations of the GBSIII and Pn14 6RU oligosaccharides (Figure A1d). Rotation to the anti- $\psi$  orientation of this linkage is prevented by the proximity of the backbone  $\beta\text{D}\text{Glc}$  residue. In Pn14, the last RU which is not in contact with a branch has a significant population of the anti- $\psi$  conformer. This is in agreement with simulations of both sLe<sup>X</sup> [20] and sLe<sup>X-5</sup> [45], where the same linkage is constrained by the close proximity of the branching  $\alpha\text{L}\text{Fuc}$ .

#### Appendix A.3. $\alpha\text{D}\text{Neu5Ac}(2\rightarrow 3)\beta\text{D}\text{Gal}$

The PMF for the charged  $\alpha\text{D}\text{Neu5Ac}(2\rightarrow 3)\beta\text{D}\text{Gal}$  disaccharide PMF (Figure A1c, left) was computed in water and shows a very flexible linkage with multiple minima separated by low energy barriers. The conformation of this sialyl linkage is known to have no fixed standard conformation and to be heavily dependent on the molecular environment [44]. Indeed, in contrast to the flexibility of the disaccharide, in our simulation of GBSIII 6RU, a *gauche* conformation predominates for this branching residue, with an average value of  $\phi, \psi = -67^\circ, -4^\circ$ . The angle distribution for this linkage is consistent with previous MD simulations of sLe<sup>X</sup> [20] and sLe<sup>X-5</sup> [45]. In contrast, González-Outeiriño et al. found that the ( $\phi \approx -180^\circ$ ) *anti*-phi conformations predominated with the mean of  $\phi, \psi = -173^\circ, -11^\circ$  and a population ratio for the anti/gauche states of approximately 96:4 [10].

In the GBSIII DP2-Fab complex crystal structure [13], the unbound branch has angles equivalent to  $\phi, \psi = -80^\circ, -17^\circ$  (close to our values) the bound branch has  $\phi, \psi = 54^\circ, -4^\circ$  (indicated by 'X's on the PMF plot).

## Appendix B. Branch Point in Tetrasaccharide



**Figure A2.** Rotation of the three-bond  $\beta\text{DGlC}(1\rightarrow 6)\beta\text{DGlCNAC}$  linkages in Pn14 and the effective tetrasaccharide epitope previously identified (left) [14]. Time series scatter plots are superimposed on the  $\phi, \psi$  PMF (top row) and the  $\phi, \omega$  PMF (bottom row) for the two central  $\beta\text{DGlC}(1\rightarrow 6)\beta\text{DGlCNAC}$  linkages for: (a) Pn14  $\phi, \psi$ ; (b) Pn14  $\phi, \omega$ ; (c) tetrasaccharide  $\phi, \psi$ ; and (d) tetrasaccharide  $\phi, \omega$ .

## References

- Edmond, K.M.; Kortsalioudaki, C.; Scott, S.; Schrag, S.J.; Zaidi, A.K.M.; Cousens, S.; Heath, P.T. Group B streptococcal disease in infants aged younger than 3 months: Systematic review and meta-analysis. *Lancet* **2012**, *379*, 547–556. [[CrossRef](#)]
- Johnson, H.L.; Deloria-Knoll, M.; Levine, O.S.; Stoszek, S.K.; Hance, L.F.; Reithinger, R.; Muenz, L.; O'Brien, K.L. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: The pneumococcal global serotype project. *PLoS Med.* **2010**, *7*, e1000348. [[CrossRef](#)] [[PubMed](#)]
- Donders, G.; Halperin, S.; Devlieger, R.; Baker, S.; Forte, P.; Wittke, F.; Slobod, K.; Dull, P. Maternal immunization with an investigational trivalent group B Streptococcal vaccine: A randomized controlled trial. *Obstet. Gynecol.* **2016**, *127*, 213–221. [[CrossRef](#)] [[PubMed](#)]
- Madhi, S.; Koen, A.; Cutland, C.; Jose, L.; Govender, N.; Wittke, F.; Olugbosi, M.; Sobanjo-ter Meulen, A.; Baker, S.; Dull, P.; Narasimhan, V. Antibody Kinetics and Response to Routine Vaccinations in Infants Born to Women Who Received an Investigational Trivalent Group B Streptococcus Polysaccharide CRM197-Conjugate Vaccine During Pregnancy. *Clin. Infect. Dis.* **2017**, *65*, 1897–1904. [[CrossRef](#)] [[PubMed](#)]
- Crumrine, M.; Fischer, G.; Balk, M. Immunochemical cross-reactions between type III group B streptococcus and type 14 *Streptococcus pneumoniae*. *Infect. Immun.* **1979**, *25*, 960–963. [[PubMed](#)]
- Guttormsen, H.K.; Baker, C.J.; Nahm, M.H.; Zughaiher, L.C.P.S.M.; Edwards, M.S.; Kasper, D.L. Type III Group B Streptococcal Polysaccharide Induces Antibodies That Cross-React with *Streptococcus pneumoniae* Type 14. *Infect. Immun.* **2002**, *70*, 1724–1738. [[CrossRef](#)] [[PubMed](#)]
- Safari, D.; Dekker, H.A.T.; Rijkers, G.T.; van der Ende, A.; Kamerling, J.P.; Snippe, H. The immune response to group B streptococcus type III capsular polysaccharide is directed to the -Glc-GlcNAc-Gal- backbone epitope. *Glycoconj. J.* **2011**, *28*, 557–562. [[CrossRef](#)] [[PubMed](#)]

8. Kasper, D.L.; Baker, C.; Baltimore, R.; Crabb, J.; Schiffman, G.; Jennings, H. Immunodeterminant specificity of human immunity to type III group B streptococcus. *J. Exp. Med.* **1979**, *149*, 327–339. [[CrossRef](#)] [[PubMed](#)]
9. Jennings, H.J.; Lugowski, C.; Kasper, D.L. Conformational aspects critical to the immunospecificity of the type III group B streptococcal polysaccharide. *Biochemistry* **1981**, *20*, 4511–4518. [[CrossRef](#)] [[PubMed](#)]
10. González-Outeiriño, J.; Kadirvelraj, R.; Woods, R.J. Structural elucidation of type III group B Streptococcus capsular polysaccharide using molecular dynamics simulations: The role of sialic acid. *Carbohydr. Res.* **2005**, *340*, 1007–1018. [[CrossRef](#)] [[PubMed](#)]
11. Kadirvelraj, R.; González-Outeiriño, J.; Foley, L.; Beckham, M.L.; Jennings, H.J.; Foote, S.; Ford, M.G.; Woods, R.J. Understanding the bacterial polysaccharide antigenicity of *Streptococcus agalactiae* versus *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8149–8154. [[CrossRef](#)] [[PubMed](#)]
12. Jennings, H.J.; Katzenellenbogen, E.; Lugowski, C.; Michon, F.; Roy, R.; Kasper, L.D. Structure, conformation and immunology of sialic acid containing polysaccharides of human pathogenic bacteria. *Pure Appl. Chem.* **1984**, *56*, 893–905. [[CrossRef](#)]
13. Carboni, F.; Adamo, R.; Fabbrini, M.; Ricco, R.D.; Cattaneo, V.; Brogioni, B.; Veggi, D.; Pinto, V.; Passalacqua, I.; Oldrini, D.; Rappuoli, R.; Malito, E.; Margarita, I.R.; Berti, F. Structure of a protective epitope of group B Streptococcus type III capsular polysaccharide. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5017–5022. [[CrossRef](#)] [[PubMed](#)]
14. Safari, D.; Dekker, H.A.T.; Joosten, J.A.; Michalik, D.; de Souza, A.C.; Adamo, R.; Lahmann, M.; Sundgren, A.; Oscarson, S.; Kamerling, J.P.; Snippe, H. Identification of the smallest structure capable of evoking opsonophagocytic antibodies against *Streptococcus pneumoniae* type 14. *Infect. Immun.* **2008**, *76*, 4615–4623. [[CrossRef](#)] [[PubMed](#)]
15. Kurbatova, E.A.; Akhmatova, N.K.; Akhmatova, E.A.; Egorova, N.B.; Yastrebova, N.E.; Sukhova, E.V.; Yashunsky, D.V.; Tsvetkov, Y.E.; Gening, M.L.; Nifantiev, N.E. Neoglycoconjugate of Tetrasaccharide Representing One Repeating Unit of the *Streptococcus pneumoniae* Type 14 Capsular Polysaccharide Induces the Production of Opsonizing IgG1 Antibodies and Possesses the Highest Protective Activity As Compared to Hexa- and Octasaccharide Conjugates. *Front. Immunol.* **2017**, *8*, 659.
16. Varki, A.; Freeze, H.H.; Manzi, A.E. Overview of Glycoconjugate Analysis. *Curr. Protoc. Protein Sci.* **2009**, *57*, 12.
17. Kuttel, M.M.; Ravenscroft, N. The Role of Molecular Modeling in Predicting Carbohydrate Antigen Conformation and Understanding Vaccine Immunogenicity. In *Carbohydrate-Based Vaccines: From Concept to Clinic*; ACS Symposium Series; American Chemical Society: Washington, DC, USA, 2018; Volume 1290, Chapter 7, pp. 139–173.
18. Kirschner, K.; Yongye, A.; Tschampel, S.; Gonzalez-Outeirino, J.; Daniels, C.; Foley, B.; Woods, R. GLYCAMS06: A Generalizable Biomolecular Force Field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622–655. [[CrossRef](#)]
19. Guvench, O.; Hatcher, E.; Venable, R.M.; Pastor, R.W.; Alexander, D.; MacKerell, J. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370. [[CrossRef](#)]
20. Guvench, O.; Mallajosyula, S.S.; Raman, E.P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T.J.; Jamison, F.W., II; MacKerell, A.D., Jr. CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and CarbohydrateProtein Modeling. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180. [[CrossRef](#)]
21. Mallajosyula, S.S.; Guvench, O.; Hatcher, E.; MacKerell, A.D., Jr. CHARMM Additive All-Atom Force Field for Phosphate and Sulfate linked to carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 759–776. [[CrossRef](#)]
22. Lewis, A.L.; Nizet, V.; Varki, A. Discovery and characterization of sialic acid O-acetylation in group B Streptococcus. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11123–11128. [[CrossRef](#)] [[PubMed](#)]
23. Pannaraj, P.S.; Edwards, M.S.; Ewing, K.T.; Lewis, A.L.; Rencha, M.A.; Baker, C.J. Group B streptococcal conjugate vaccines elicit functional antibodies independent of strain O-acetylation. *Vaccine* **2009**, *27*, 4452–4456. [[CrossRef](#)] [[PubMed](#)]
24. Kuttel, M.M.; Jackson, G.E.; Mafata, M.; Ravenscroft, N. Capsular polysaccharide conformations in pneumococcal serotypes 19F and 19A. *Carbohydr. Res.* **2015**, *406*, 27–33. [[CrossRef](#)] [[PubMed](#)]
25. Kuttel, M.M.; Timol, Z.; Ravenscroft, N. Cross-protection in *Neisseria meningitidis* serogroups Y and W polysaccharides: A comparative conformational analysis. *Carbohydr. Res.* **2017**, *446–447*, 40–47. [[CrossRef](#)] [[PubMed](#)]

26. Hlozek, J.; Kuttel, M.M.; Ravenscroft, N. Conformations of *Neisseria meningitidis* serogroup A and X polysaccharides: The effects of chain length and O-acetylation. *Carbohydr. Res.* **2018**, *465*, 44–51. [CrossRef] [PubMed]
27. Laio, A.; Parrinello, M. Escaping free energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12565. [CrossRef] [PubMed]
28. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef]
29. Yang, M.; MacKerell, A.D. Conformational Sampling of Oligosaccharides Using Hamiltonian Replica Exchange with Two-Dimensional Dihedral Biasing Potentials and the Weighted Histogram Analysis Method (WHAM). *J. Chem. Theory Comput.* **2015**, *11*, 788–799. [CrossRef]
30. Kuttel, M.M. Conformational free energy maps for globobiose ( $\alpha$ -D-Gal-(1-4)- $\beta$ -D-Gal) in implicit and explicit aqueous solution. *Carbohydr. Res.* **2008**, *343*, 1091–1098. [CrossRef]
31. Stone, J.E.; Phillips, J.C.; Freddolino, P.L.; Hardy, J.; Trabuco, L.G.; Schulten, K. Accelerating Molecular Modeling Applications with Graphics Processors. *J. Comput. Chem.* **2007**, *28*, 2618–2639. [CrossRef]
32. Guvench, O.; Greene, S.N.; Kamath, G.; Brady, J.W.; Venable, R.M.; Pastor, R.W.; MacKerell, A.D. Additive Empirical Force Field for Hexopyranose Monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543–2564. [CrossRef] [PubMed]
33. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulations of liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
34. Kuttel, M.; Ravenscroft, N.; Foschiatti, M.; Cescutti, P.; Rizzo, R. Conformational properties of two exopolysaccharides produced by *Inquilinus limosus*, a cystic fibrosis lung pathogen. *Carbohydr. Res.* **2012**, *350*, 40–48. [CrossRef] [PubMed]
35. Kuttel, M.M.; Stähle, J.; Widmalm, G. CarbBuilder: Software for Building Molecular Models of Complex Oligo- and Polysaccharide Structures. *J. Comput. Chem.* **2016**, *37*, 2098–2105. [CrossRef] [PubMed]
36. Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]
37. Feller, S.E.; Zhang, Y.; Pastor, R.W.; Brooks, B.R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621. [CrossRef]
38. Nose, S.; Lein, M.L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055–1076. [CrossRef]
39. Hoover, W.G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*. [CrossRef]
40. Cross, S.; Kuttel, M.M.; Stone, J.E.; Gain, J.E. Visualisation of Cyclic and Multi-Branched Molecules with VMD. *J. Mol. Graph. Model.* **2009**, *28*, 131–139. [CrossRef] [PubMed]
41. Heyer, L.J.; Kruglyak, S.; Yoosoph, S. Exploring Expression Data: Identification and Analysis of Coexpressed Genes. *Genome Res.* **1999**, *9*, 1106–1115. [CrossRef] [PubMed]
42. Lundborg, M.; Widmalm, G. Structure Analysis of Glycans by NMR Chemical Shift Prediction. *Anal. Chem.* **2011**, *83*, 1514–1517. [CrossRef] [PubMed]
43. Nishida, Y.; Hori, H.; Ohru, H.; Meguro, H.  $^1\text{H}$ -NMR analyses of rotameric distributions of C5-C6 bonds of D-glucopyranoses in solution. *J. Carbohydr. Chem.* **1988**, *7*, 239–250. [CrossRef]
44. Miyazaki, T.; Sato, H.; Sakakibara, T.; Kajihara, Y. An Approach to the Precise Chemoenzymatic Synthesis of  $^{13}\text{C}$ -Labeled Sialyloligosaccharide on an Intact Glycoprotein: A Novel One-Pot [3- $^{13}\text{C}$ ]-Labeling Method for Sialic Acid Analogues by Control of the Reversible Aldolase Reaction, Enzymatic Synthesis of [3- $^{13}\text{C}$ ]-NeuAc- $\alpha$ -(2→3)-[U- $^{13}\text{C}$ ]-Gal- $\beta$ -(1→4)-GlcNAc- $\beta$ - Sequence onto Glycoprotein, and Its Conformational Analysis by Developed NMR Techniques. *J. Am. Chem. Soc.* **2000**, *122*, 5678–5694.
45. Battistel, M.D.; Azurmendi, H.F.; Frank, M.; Freedberg, D.I. Uncovering Nonconventional and Conventional Hydrogen Bonds in Oligosaccharides through NMR Experiments and Molecular Modeling: Application to Sialyl Lewis X. *J. Am. Chem. Soc.* **2015**, *137*, 13444–13447. [CrossRef] [PubMed]

