Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 340 (2005) 1007–1018
Carbohydrate RESEARCH

# Structural elucidation of type III group B *Streptococcus* capsular polysaccharide using molecular dynamics simulations: the role of sialic acid

Jorge González-Outeiriño, Renuka Kadirvelraj and Robert J. Woods*

*Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA*

Received 30 September 2004; accepted 16 December 2004

Dedicated to Professor David A. Brant

**Abstract**—The conformational properties of the capsular polysaccharide (CPS) from group B *Streptococcus* serotype III (GBS III) are derived from 50 ns explicitly solvated molecular dynamics simulations of a 25-residue fragment of the CPS. The results from the simulations are shown to be consistent with experimental NMR homo- and heteronuclear *J*-coupling and NOE data for both the sialylated native CPS and for the chemically desialylated polysaccharide. A helical structure is predicted with a diameter of 29.3 Å and a pitch 89.5 Å, in which the sialylated side chains are arrayed on the exterior surface of the helix. The results provide an explanation for the observation that CPS antigenicity varies with carbohydrate chain length up to approximately 4 pentasaccharide repeat units. The conformation of the immunodominant region is established and shown to be independent of the presence of sialic acid. The data provide an explanation for the observation that the specificity of the determinant, associated with the major population of antibodies raised upon immunization of rabbits with GBS III, is dependent on the presence of sialic acid. In the sialylated native CPS, the antibody response is largely directed against the immunodominant core of the helix. From simulations of the desialylated CPS, a model emerges which suggests that the minor population of antibodies, whose determinant is not sialic acid dependent, recognizes the same immunodominant region, but that in the disordered CPS this region is not presented in a regular repeating motif.


*Keywords:* AMBER; Capsular polysaccharide; Conformational epitope; GLYCAM; Group B *Streptococcus*; Helix; Molecular dynamics simulation

## 1. Introduction

Group B *Streptococcus* (GBS) is the leading cause of bacterial infection in pregnant women and newborns.[1] Expression of the disease may occur within the first week (early onset disease) or up to several months post-partum (late-onset disease). Typical presentations of the disease are sepsis, pneumonia and meningitis. GBS bacteraemia is becoming an increasingly significant disease in the elderly, particularly in those with underlying diseases.[2] In non-pregnant adults, incidence of GBS disease is on the rise and commonly appears as skin, soft tissue and osteoarticular infections, and pneumonia, while in the elderly, meningitis and endocarditis are associated with high morbidity and mortality.[3]

GBS bacteria are encapsulated by complex branched polysaccharides, and variations in these sequences correspond to strain classifications.[4] Each capsular polysaccharide (CPS) is derived generally from the same set of glycosyl residues; structural and immunological diversity arising from differences in linkage position and anomeric configuration.[4] At present, nine serotypes of GBS (Ia, Ib, II–VIII) have been identified and their CPS sequences determined. The CPS sequence has been suggested as the main cause of the organism's virulence.[5] Serotype III, in particular, accounts for the majority of cases of both early and late onset meningitis.[6,7]

* Corresponding author. Tel.: +1 706 542 4454; fax: +1 706 542 4412; e-mail: rwoods@ccrc.uga.edu

0008-6215/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2004.12.034

Each polymeric repeat unit in the CPS of GBS III contains three backbone residues, →4)-β-D-Glc-(1→6)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→, with a further two residues, α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)–, attached as a side chain to the GlcNAc residue (Fig. 1). It is notable that all CPS serotypes include a terminal sialic acid in the side chain.

The relationship between epitope structure and the presence of the sialic acid has been the focus of considerable attention. The presence of the negatively charged carboxylate moiety has been suggested to play a role in the pathogenicity of the organism.[5,8–10] Cleavage of the exocyclic C7–C8 or C8–C9 bonds had no effect on the affinity of the modified polymer for CPS-specific antibody, in contrast, reduction of the carboxylic acid moiety had a profound effect on the antigenicity.[10] Removal of the entire side chain from the CPS diminished antigenicity, and led to modest changes in $^{13}C$ NMR chemical shifts, relative to the native CPS, which were interpreted as indicating a conformational change in the backbone.[10] Thus, the sialic acid residues exert conformational control over the CPS, and the associated immunological determinant, apparently without participating directly in the antigen binding.

In the quest to establish the size of the epitope, it has been determined that 3–4 pentasaccharide repeat units are required to fully inhibit native CPS from binding to CPS-specific monoclonal antibodies.[11] The antigenicity of a single repeat unit is low, but increases with increasing carbohydrate chain length from 2 to 7 repeat units.[11,12] These studies also suggested that the β-D-Gal-(1→4)-β-D-Glc sequence in the backbone was critical to antibody binding.[10,12] Taken together, the inhibition and binding data indicate an epitope of 3–4 repeat units, or 9–12 backbone residues, in length. However, the inconsistency of this straightforward interpretation with the early work of Kabat on anti-dextran antibodies,[13] which suggested a maximum size of six residues for carbohydrate epitopes, has been noted.[12] An alternative interpretation is that the oligosaccharide undergoes a conformational change as the chain length increases from 2 to 4 repeat units, which introduces a conformational epitope that is equivalent to that present in the native CPS.[10,11] Based on extrapolation of the NMR and simulational data for a single repeat unit, an earlier study predicted that the GBS III CPS could potentially adopt any one of several helical structures.[14] To the extent that the CPS is helical in structure, many questions remain. What is the role of sialic acid in maintaining this structure; how stable is the helical form relative to the desialylated structure; is the helix the preponderant conformation of the polysaccharide prior to antibody binding; does antibody binding induce a change in the CPS conformation and, or, lead to propagation of the helical conformation throughout the CPS?[11,15] Here we examine the first three of these questions, employing extensive molecular dynamics simulations on models of the native and desialylated CPS. These models consist of five polymeric repeat units each, and so may be expected to unambiguously display the conformational epitope and dynamic properties. The ability of explicitly solvated MD simulations to reproduce the conformational properties of oligosaccharides is well established.[16–20] However, to adequately sample the conformational properties of 1→6 linkages, as present in the GBS III CPS, it is necessary to perform the simulations for extensive periods. In the case of mono- and disaccharides, we have shown that simulation times of 10–100 ns are required to achieve agreement with experimental NMR data for the rotamer populations of the C-5–C-6 bond.[20] Despite the computational demand associated with the present systems, which contain approximately 18,000 atoms each, including waters and counter ions, the simulations were performed for 50 ns each.

The length of the simulation notwithstanding, the potential for incomplete conformational sampling and or weaknesses in the computational force field require that careful attention be paid to validation of the simulational data. For carbohydrates in solution, NMR data

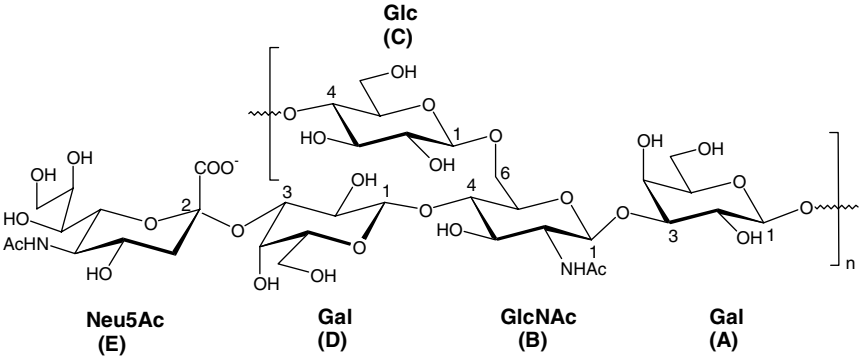


**Figure 1.** Monomeric repeat unit of CPS from type III group B *Streptococcus*. The backbone sequence is →4)-β-D-Glc-(1→6)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→, while the side chain sequence is: α-Neu5Ac-(2→3)-β-D-Gal-(1→4)–.

such as *J*-couplings, nuclear Overhauser effect (NOE) intensities, residual dipolar couplings and relaxation data are frequently employed for validation. In the case of GBS III, considerable NMR data have been reported, and the assessment of the current simulations was carried out in light of that data.[14]

## 2. Materials and methods

A fully solvated 50 ns MD simulation was performed on a 25-residue fragment of the GBS III CPS. The initial conformation of the polymer was generated from the experimentally consistent structure determined for the single pentasaccharide repeat unit.[14] The anionic system was neutralized by the addition of five sodium counter ions. The solute was then placed in a theoretical box of 5934 TIP3P water molecules. Concurrently, a 50 ns MD simulation of the desialylated CPS fragment was performed, beginning from the same initial backbone conformation. The MD simulations were performed under periodic boundary conditions at constant pressure (1 atm), and constant temperature (300 K), with a particle mesh Ewald treatment of long-range electrostatic interactions, using the SANDER module of the AMBER 7 suite programs.[21] The GLYCAM_2000 parameters[20,22] were implemented for the oligosaccharide simulations. Energy minimizations and MD simulations were performed with a dielectric constant of unity, and a cut-off value for non-bonded interactions of 8 Å. The 1–4 electrostatic and van der Waals interactions were scaled by the standard values (SCEE = 1.2, SCNB = 2.0).

For each system, initial solvent configurations were subjected to 15,000 cycles of energy minimization, composed of 1000 steepest descent cycles and 14,000 cycles of conjugate gradients minimization, with restrained solute coordinates. Subsequently, energy minimization was carried out on the entire system for 30,000 cycles (1000 steepest descent followed by 29,000 conjugate gradient), allowing all the atoms to relax. Energy minimization was followed by a simulated annealing of the solvent, during which the waters were heated from 5 to 300 K in 50 ps, held at 300 K for 200 ps and then cooled to 5 K in 50 ps. The entire system was then subjected to a final minimization for 30,000 cycles. Prior to the production MD run, the system was warmed to 300 K over a period of 50 ps, with initial velocities assigned from a Boltzmann distribution at 5 K. A 2 fs time step was employed in the integration of the equations of motion. Configurational snapshots were collected every ps.

Post-processing of the trajectory (torsion, distance and positional RMSD analysis) was performed using the CARNAL module of AMBER 7. NMR scalar couplings were computed using Karplus-type relationships for homonuclear[23] and heteronuclear[24,25] 3-bond interactions. NOE intensities were computed using a full relaxation matrix analysis employing the CORMA program.[26] The 2D NOESY intensities were computed for an ensemble of 3000 snapshots, extracted at approximately 16 ps intervals from the trajectory, using the experimental correlation time of 2 ns and a mixing time of 100 ms.

## 3. Results and discussion

The modeled CPS fragment was formed from five polymeric repeat units, each containing five glycosyl residues, labeled historically as A, B, C for the backbone residues (β-D-Gal, A; β-D-GlcNAc, B; β-D-Glc, C) and D and E for the side chain residues (β-D-Gal, D; αNeu5Ac, E). A sequence of this length is longer than the minimum number of units required to inhibit the intact CPS from binding to a monoclonal antibody.[10,11] The model for the desialylated structure also contained five polymeric repeat units, each consisting only of residues A–D. For both the native and desialylated CPS, initial evaluations of the conformational fluctuations during the MD simulations were obtained from 1D plots of the root-mean-square deviation (RMSD) of the coordinate positions of the non-hydrogen atoms, with respect to the corresponding starting conformation (Fig. 2a and b). To focus on the behavior of the polymer, and to avoid including artifactual fluctuations associated with the terminal repeat units, the RMSD analysis included only the three central repeat units.

### 3.1. RMSD analysis

The 1D RMSD of the native CPS fragment showed an average value of 2.98 Å with a standard deviation of 1.0 Å, while the desialylated CPS average value was 3.86 Å with a 1.2 Å standard deviation. These are remarkably small values for such a large fragment, and although the desialylated CPS fragment showed larger overall fluctuations, a 1D analysis did not provide sufficient insight into the conformational flexibilities of these two molecules. A 2D RMSD analysis (Fig. 2c and d) showed that the native CPS underwent only one brief 7 ns excursion from an essentially constant structure. During this period the $\Phi$ and $\Psi$-angle of the β-D-Glc-(1→6)-β-D-GlcNAc torsion angle adopted values of −30° and 92°, respectively, which caused the oligosaccharide to transiently adopt a less regular structure. The desialylated derivative, however, starting from the same initial structure as the native CPS, quickly decayed to conformations that showed much less similarity to each other. That is, although the 1D RMSD values suggested a slightly larger deviation for the desialylated case, the 2D analysis indicated that there was a fundamental difference in the internal

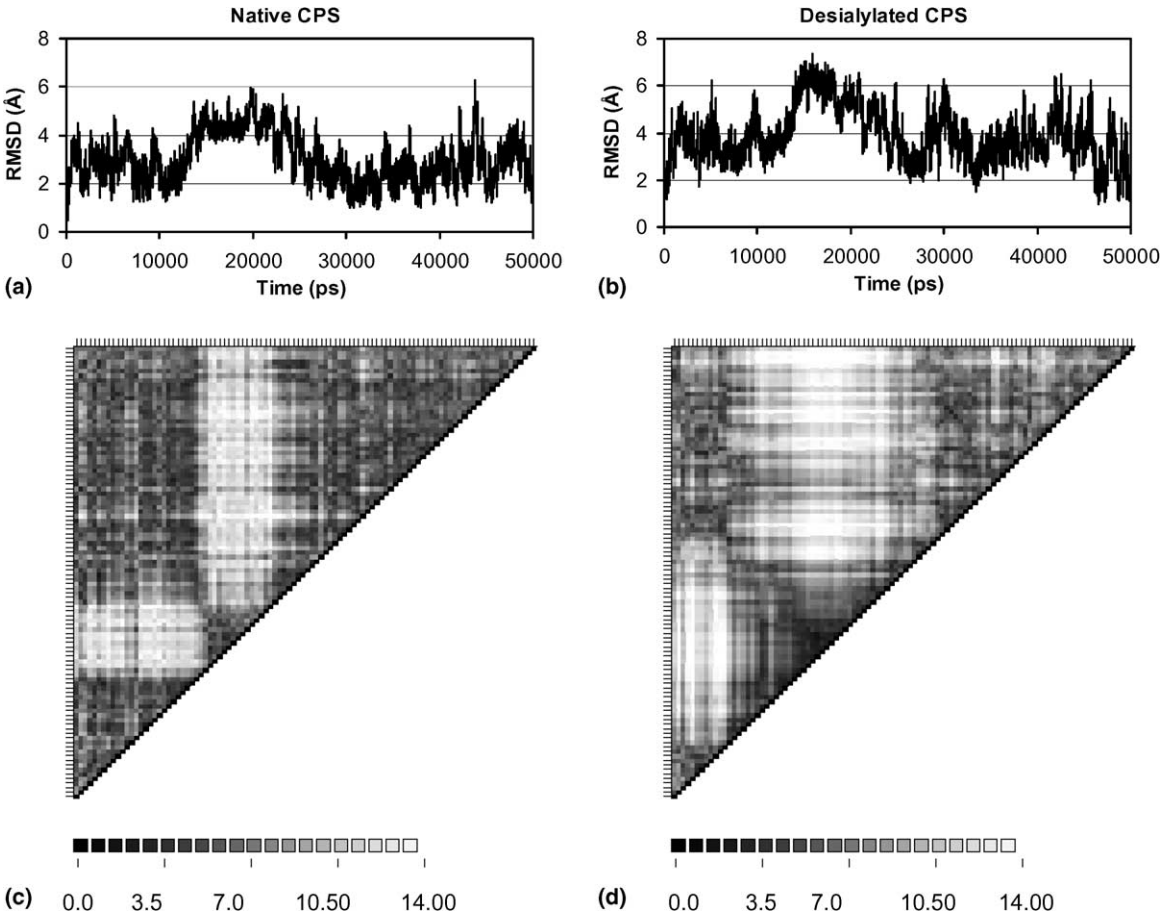


**Figure 2.** 1D and 2D RMSD plots (Å) of the fully solvated MD trajectories for the native (panels a and c) and desialylated CPS fragments (panels b and d).

motions of the two oligosaccharides. To visualize the impact of removal of the sialic acid on the CPS structure, snapshots selected at 10 ns intervals for each simulation are presented in Figure 3. These images highlight the helical character of the native CPS, which is maintained throughout the simulation, and indicate that loss of sialic acid led to a less regular overall structure.

### 3.2. Glycosidic torsion angles

While RMSD data can provide insight into overall conformational properties, internal motions are well characterized by the values for the glycosidic torsion angles. Presented in Table 1 are the average torsion values computed from 50,000 snapshots extracted from the 50 ns simulations for the native and desialylated CPS fragments. Included for comparison are the values reported earlier from a 1 ns simulations of a single repeat unit of both the native and desialylated structures.[14]

For each respective CPS structure, the 1 and 50 ns data are in qualitative agreement. An examination of the standard deviations indicates that the larger fragment displays considerably less internal motion. It should be noted that data shown in Table 1 for the pentameric oligosaccharides represent the average of the values for each corresponding torsion angle over the whole oligosaccharide fragment. Unlike the case of the RMSD analysis, this data does include contributions from the terminal repeat units. The average values of each of the glycosidic torsion angles for the pentameric unit fall into expected regions of $\Phi$, $\Psi$-space.[27] However, there are two linkages whose conformations differ noticeably between the single and pentameric fragments. In the case of the native CPS, the monomeric and pentameric fragments differ in the predicted values of the $\Phi$- and $\Psi$-angles of the α-Neu5Ac-(2→3)-β-D-Gal (residues D–E) linkage. In linkages of this type, the $\Phi$-angle is known to populate predominantly only two conformations, namely, approximately −60° (*gauche*) and 180° (*anti*),[28,29] which are the two values predicted by the MD simulations.

An examination of the $\Phi$, $\Psi$ plots (Fig. 4) for the α-Neu5Ac-(2→3)-β-D-Gal linkages in the oligosaccharide shows clearly that the $\Phi$-angle samples both conformations during the simulation, but a detailed analysis results in a population ratio for the *anti*/*gauche* states of approximately 96:4. With such a high ratio, the *gauche* conformers make a negligible impact on the average

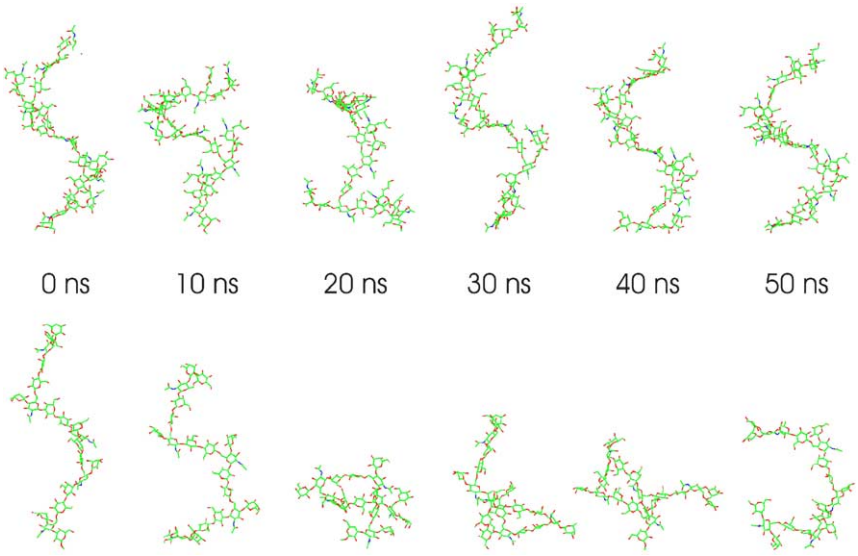


**Figure 3.** Snapshots from the 50 ns MD simulation of the native and desialylated CPS extracted at 10ns intervals. The upper trajectory corresponds to the native CPS and the lower trajectory to the desialylated derivative.

**Table 1.** Average glycosidic torsion angles[a] from MD trajectories for native and desialylated CPS fragments

| Linkage | Torsion angle | Native CPS | | Desialylated CPS | |
|---|---|---|---|---|---|
| | | MD (1 ns) $[ac(ed)ba]_1$[14] | MD (50 ns) $[c(ed)ba]_5$ | MD (1 ns) $[ac(ed)ba]_1$[14] | MD (50 ns) $[c(ed)ba]_5$ |
| Gal(1→4)Glc | $\Phi$ | 37 (20) | 47 (12) | 45 (15) | 43 (12) |
| (A–C) | $\Psi$ | −12 (30) | 8 (12) | −7 (39) | −7 (16) |
| GlcNAc(1→3)Gal | $\Phi$ | 40 (20) | 42 (13) | 48 (15) | 41 (12) |
| (B–A) | $\Psi$ | 30 (28) | 2 (18) | 40 (17) | −1 (19) |
| Glc(1→6)GlcNAc | $\Phi$ | 48 (14) | 41 (10) | 42 (21) | 40 (14) |
| (B–C) | $\Psi$[b] | −170 (18) | −159 (22) | −182 (17) | −167 (31) |
| | $\omega$ | −65 (14), 60 (17)[c] | −66.6 (8.7) | −59 (10), 51 (13)[d] | −65 (10), 58 (13)[e] |
| Gal(1→4)GlcNAc | $\Phi$ | 52 (13) | 39 (10) | 50 (15) | 43 (11) |
| (D–B) | $\Psi$ | −12 (21) | 5 (10) | −12 (26) | 5 (11) |
| NeuAc(3→2)Gal | $\Phi$ | −43 (18) | −173 (14) | — | — |
| (E–D) | $\Psi$ | −116 (35) | −11 (11) | — | — |

[a] Values in degrees, standard deviations in parentheses.
[b] $\Psi$ = C-1–O-6–C-6–C-5.
[c] Ratio of *gg* (−60):*gt* (60) = 80:20.
[d] Ratio of *gg* (−60):*gt* (60) unreported.
[e] Ratio of *gg* (−60):*gt* (60) = 56:44.

values presented in Table 1. In the single repeat unit the $\Psi$-angle was notably flexible, as evidenced by a large standard deviation, however, the average value was not consistent with the expected eclipsed conformation.[30] In the larger structure, the $\Psi$-angle was both less flexible and more consistent with previous data.[28,29] It should be noted that the single repeat unit is incapable of displaying the effects of inter-repeat unit interactions, which were observed in the larger structure, vide infra. The second discrepancy between the previously reported data and our calculations lies in the average value of the $\Psi$-glycosidic angle for the β-D-GalNAc-(1→3)-β-D-Gal sequence. While all data sets place this angle within ±40° of the expected eclipsed value,[30] the current results are in overall better agreement with experiment. Taken together, the data for the single repeat unit indicate a more flexible structure than present in the oligosaccharide, which is not unexpected given the increased inertial and frictional resistance associated with motions of the larger structure. It is important to note that the backbone conformations are essentially equivalent in the monomeric and pentameric fragments. This observation is particularly relevant to the characterization of the conformation of the immunodominant epitope associated with the β-D-Gal-(1→4)-β-D-Glc sequence in either the native or desialylated CPS. In terms of the glycosidic torsion angles, the structure of this region in the oligosaccharide is independent of sialylation and is indistinguishable from that in the single repeat unit. However, the environment around the epitope depends on the overall conformation of the CPS, which is profoundly influenced by the presence or absence of sialic acid in the

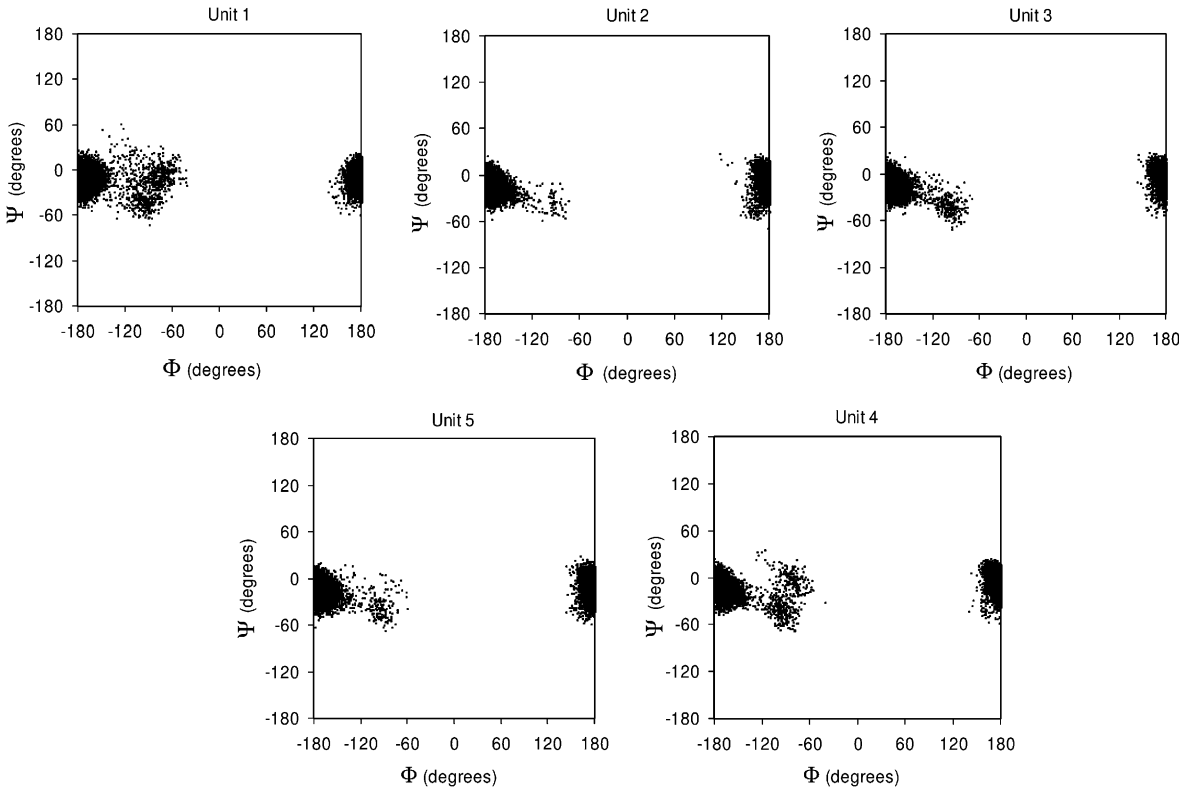


**Figure 4.** $\Phi$, $\Psi$ plot for the α-NeuAc-(2→3)-β-D-Gal sequences in each repeat unit in the pentameric model of the native CPS.

side chain. The strength of the immune response depends on the presence of multiple repeat units. The extent to which this response requires that multiple repeat units make simultaneous contact with the antibody surface would be strongly dependant on the relative orientation of each immunodominant sequence in the CPS.

In the case of the β-D-Glc-(1→6)-β-D-GlcNAc linkages, the potential exists for more than one rotamer of the $\omega$-angle to be populated. As seen in Table 1 and Figure 5, the desialylated form showed a nearly 50:50 mixture of the *gg* to *gt* rotamers, whereas the native form remained almost constantly in the *gg* conformation. Only the terminal repeat units in the native oligosaccharide exhibited any significant presence of internal motion. In contrast, the removal of the sialic acid led to rotamer interconversions within the internal as well as terminal repeat units. These transitions were extremely infrequent, with lifetimes of 20–40 ns, and so were not sampled in a statistically accurate manner. It is primarily the internal rotations associated with the 1→6 linkage in the central repeat unit that led to the apparently disordered overall structure of the desialylated CPS (Fig. 5).

### 3.3. Scalar $^3J$-couplings

The heteronuclear $^3J$-values for the glycosidic linkages in the backbone of the CPS were computed from snapshots extracted at 1 ps intervals from the MD trajectories (Table 1). A Karplus-type equation optimized for heteronuclear *J*-couplings across glycosidic linkages was used to determine the $^3J$ values for the $\Phi$-(H-1–C-1–O-*x*–C-*x*) and $\Psi$-(C-1–O-*x*–C-*x*–H-*x*) angles (Table 2).[24,25] In order to distinguish between any differences that might arise from the particular parameterization of the Karplus curves, from those arising from conformational properties, we have employed two Karplus-type relationships; one was derived by fitting to experimental data,[25] while the other was derived by fitting to theoretical data.[24]

The maximum absolute unsigned error in the computed *J*-values, compared to the experimental data, is 1.0 Hz. The majority of the experimental values for both the native and desialylated structures fall within the standard deviations of the simulational data for the pentameric oligosaccharides, regardless of the Karplus-type equation employed to compute the *J*-values. Overall the agreement between the experimental data for a single repeat unit and the simulated data for the large oligosaccharide is superior to that obtained earlier from the simulation of the single repeat units. However, neither the single repeat unit nor the pentamer completely reproduce the experimental values associated with the 1→6 linkage. The experimental values for the $^3J_{CH}$ couplings for the $\Psi$ (4.2 Hz) and the $\Psi'$-angles (3.8 Hz) were reported only for the single repeat unit. Average values of approximately 4 Hz for each torsion angle can only

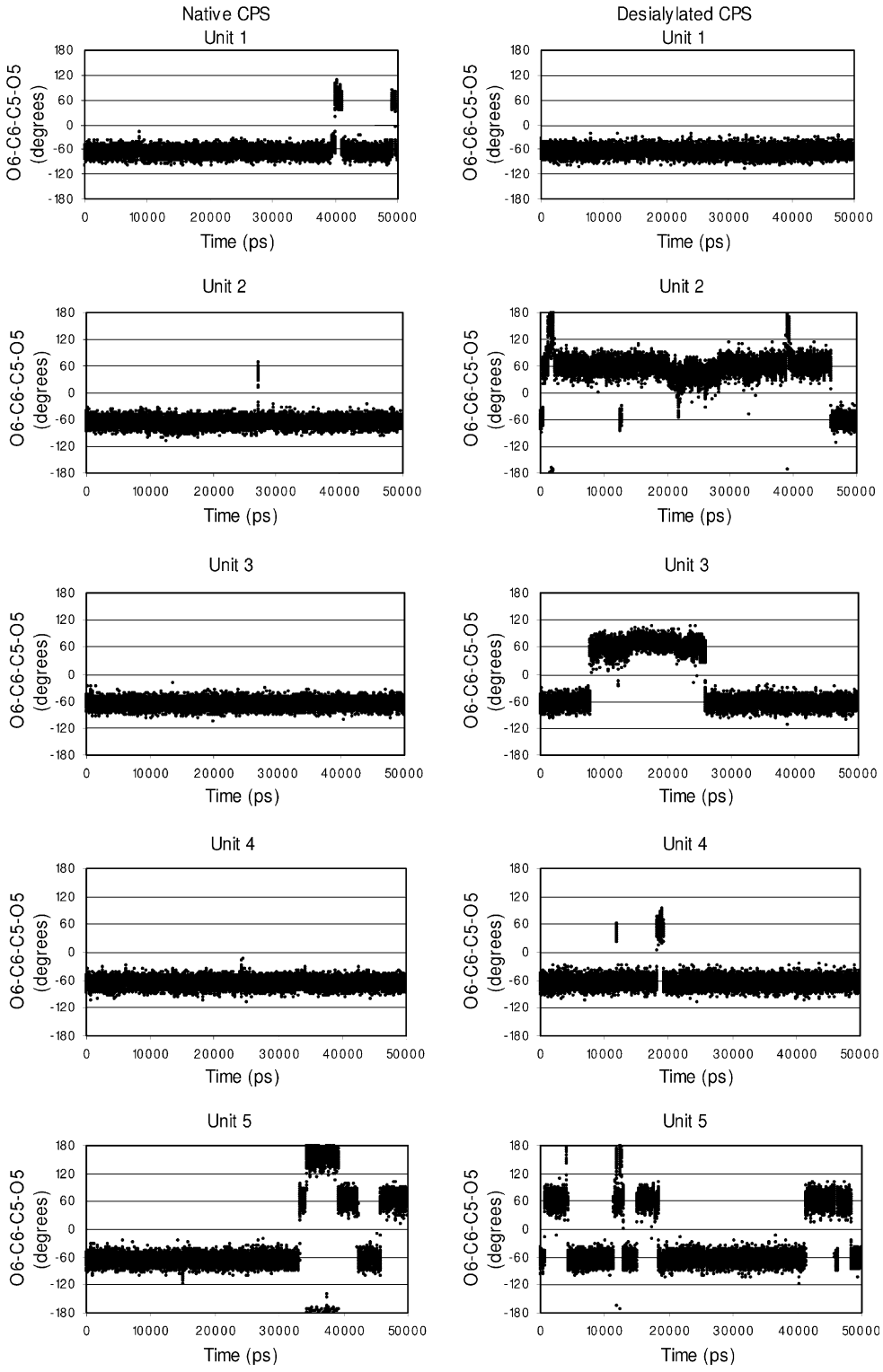


**Figure 5.** Trajectories for $\omega$-torsion angle for the β-D-Glc-(1→6)-β-D-GlcNAc sequence in the native and desialylated pentameric CPS fragments.

arise from essentially free rotation about the C-1–O-6–C-6–C-5 sequence. While this may be the case for the single repeat unit, it is not predicted or expected to be the case in the large oligosaccharide. An examination of the trajectories for the $\Psi$- and $\Psi'$-angles of the native CPS fragment indicated that the only significant rotational transition associated with a C-1–O-6–C-6–C-5 torsion angle occurred within the second repeat unit (Fig. 6). For approximately 7 ns the $\Psi$-angle of the β-D-Glc-(1→6)-β-D-GlcNAc linkage in the second repeat

**Table 2.** Experimental and computed $^3J_{CH}$ coupling constants[a] for the glycosidic linkages for the CPS fragments

| Linkage | Torsion angle | [ac(ed)ba]$_1$[14] | | [c(ed)ba]$_5$ | |
|---|---|---|---|---|---|
| | | $^3J$ Exptl. | $^3J$ computed[25] from 1 ns MD | $^3J$ computed[24] from 50 ns MD | $^3J$ computed[25] from 50 ns MD |
| *Native CPS* | | | | | |
| Gal(1→4)Glc | $\Phi$ | 2.2 | 3.4 (1.3) | 2.9 (1.2) | 2.7 (0.8) |
| | $\Psi$ | — | — | 6.4 (0.5) | 5.4 (0.4) |
| GlcNAc(1→3)Gal | $\Phi$ | 4.6 | 3.2 (1.5) | 4.0 (1.2) | 3.5 (0.9) |
| | $\Psi$ | 3.7 | 3.5 (1.3) | 5.5 (0.9) | 4.7 (0.6) |
| Glc(1→6)GlcNAc | $\Phi$ | 4.3 | 3.3 (1.3) | 3.4 (1.2) | 3.0 (0.9) |
| | $\Psi$ | 4.2 | 1.9 (1.0) | 4.2 (1.7) | 3.7 (1.3) |
| | $\Psi'$ | 3.8 | 1.9 (1.0) | 1.0 (1.3) | 1.2 (1.0) |
| Gal(1→4)GlcNAc | $\Phi$ | 4.0 | 2.3 (1.2) | 3.8 (1.0) | 3.4 (0.8) |
| | $\Psi$ | 5.3 | 5.0 (1.2) | 6.4 (0.3) | 5.4 (0.3) |
| | | [ac(d)ba]$_1$[14] | | [c(d)ba]$_5$ | |
| | | $^3J$ Exptl. | $^3J$ computed[25] from 1 ns MD | $^3J$ computed[24] from 50 ns MD | $^3J$ computed[25] from 50 ns MD |
| *Desialylated CPS* | | | | | |
| Gal(1→4)Glc | $\Phi$ | 4.3 | 2.5 (1.1) | 3.4 (1.1) | 3.0 (0.9) |
| | $\Psi$ | 4.6 | 4.9 (1.5) | 6.3 (0.4) | 5.3 (0.4) |
| GlcNAc(1→3)Gal | $\Phi$ | 4.9 | 2.7 (1.2) | 4.1 (1.2) | 3.6 (0.9) |
| | $\Psi$ | — | — | 5.5 (1.0) | 4.7 (0.8) |
| Glc(1→6)GlcNAc | $\Phi$ | 4.3 | 3.0 (1.2) | 3.7 (1.3) | 3.3 (1.0) |
| | $\Psi$ | — | — | 3.3 (1.9) | 2.9 (1.6) |
| | $\Psi'$ | — | — | 2.1 (2.0) | 2.0 (1.5) |

[a] Values in Hz, standard deviations in parentheses.

unit adopted a 92° conformation. While the *trans* orientation is generally preferred for the $\Psi$-angle in 1→6 linkages, the 90° is energetically accessible.[31] It is this transition that is primarily responsible for the conformational change noted in the RMSD analysis. The homonuclear $^3J_{H5-H6}$ coupling constants reported earlier (of $^3J_{5,6} = 3.9$ Hz and $^3J_{5,6'} = 1.8$ Hz),[14] may be used to establish the conformational behavior of the $\omega$-angle in the 1→6 linkages. Based on the data for the pentameric fragment of the native CPS, the homonuclear *J*-couplings were computed[23] to be $^3J_{5,6} = 5.2$ (1.6) Hz and 1.2 (1.3) Hz for $H_{6S}$ and $H_{6R}$, respectively. Unlike the case of the $\Psi$-angles, the data for the $\omega$-angles in the native structure are consistent with a single rotameric state, namely the *gg*, as observed in the MD simulation. In the case of the desialylated CPS fragment, the rotational interconversions led to average *J*-values for the $\omega$-angle of $^3J_{5,6} = 4.5$ (1.8) Hz for the $H_{6S}$ and $^3J_{5,6'} = 3.6$ (2.8) Hz for the $H_{6R}$. These values arise from the 56:44 ratio of *gg* to *gt* rotamers observed during the simulation. These values are in poor agreement with the experimental data reported for the monomeric unit ($^3J_{5,6} = 4.2$ Hz, $^3J_{5,6'} = 1.2$ Hz, for $H_{6S}$ and $H_{6R}$, respectively), which are again consistent with a single rotameric state. It is unlikely that the $\omega$-angles in the polymeric form would truly be more flexible than in the monomeric subunit, thus this discrepancy in the *J*-values suggests that the simulation of the larger structure had not fully equilibrated. It should be noted that the frequency of the $\omega$-angle transitions was too low to obtain accurate populations for the 1→6 linkages, and the current data may not represent the ultimate conformational distribution for the desialylated CPS.

### 3.4. Nuclear Overhauser effects

An extensive collection of NOE intensities has been reported for the single repeat unit of the native CPS.[14]

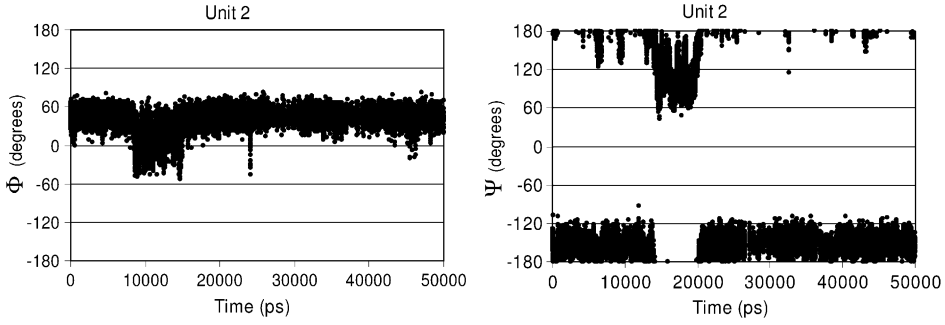


**Figure 6.** Trajectories for $\Phi$- and $\Psi$-torsion angles for the β-D-Glc-(1→6)-β-D-GlcNAc sequence in the second repeat unit in the native CPS fragment.

However, the majority of the values are for intra-residue interactions, which provide little insight into the overall 3D structure of the CPS. A limited number of inter-residue NOEs were identified, and it is these which are most relevant. A complete relaxation matrix analysis using the CORMA[26] approach was employed to derive theoretical NOE intensities from the MD data. The results from the NOE analysis of the native CPS fragments are shown in Table 3.

The 50 ns trajectory of the native CPS fragment was able to reproduce the experimental NOEs with essentially the same average error (1.4%) as that obtained from the simulation of the single repeat unit. Particularly important NOEs are those which span the glycosidic linkages, or those that arise from longer-range contacts. Given the ambiguity of the experimental *J*-coupling data associated with the $\Psi$-angle in the 1→6 linkages, it is noteworthy that the experimental NOE intensities for the H-1c–H-6b (3.8%) and for the H-1c–H-6′b (2.2%) interactions are extremely well reproduced from the simulation of the pentameric fragment (3.7 and 2.0%, respectively). *trans*-Glycosidic NOEs associated with the β-D-GlcNAc-(1→3)-β-D-Gal and β-D-Gal-(1→4)-β-D-GlcNAc sequences are also well reproduced. The NOEs computed for the α-Neu5Ac-(2→3)-β-D-Gal in the pentamer are predicted to be stronger than observed experimentally in the monomer, which reflects the increased flexibility for this linkage in the monomeric fragment.[14] A similar situation arises in the case of the H-1c–H-1d NOE, which is predicted to be larger in the pentamer than observed in the monomer. The magnitude of this NOE is enhanced by the lack of flexibility of the 1→6 linkage in the larger CPS fragment.

In the course of computing the NOEs, four previously unreported intensities were observed, corresponding to interactions between the methyl group of the *N*-acetyl moiety in the sialic acid residues with H-2a in the galactosyl residues (0.8%); between H-3c–H-$3_{ax}$e (1.4%); between H-5c–H-$3_{ax}$e (1.0%); and between H-3c–H-5e (3.3%) (Fig. 7a). These NOEs confirm the proximity of the sialic acid to the polymeric backbone.

### 3.5. Hydrogen-bond analysis

To establish the origin of the forces responsible for orienting the sialylated side chains in close proximity to the backbone, an analysis of the inter-residue hydrogen bonds was undertaken. The most relevant results are displayed in Figure 7b.

Over the period of the trajectory the hydrogen-bond analysis suggested interactions between HO-4 of the glucosyl residue (C) and the carbonyl-oxygen of the *N*-acetyl moiety (3.3 Å, 45% occupancy). Also observed were hydrogen bonds between the HO-2 of the backbone galactosyl residue (A + 1) and HO-4 of the sialic acid (E) (3.4 Å, 37% occupancy), and between HO-3 of the

**Table 3.** Experimental and computed[a] NOE intensities for the native CPS derivatives

| Linkage | Protons | [ac(ed)ba]$_1$[14] | | [c(ed)ba]$_5$ |
|---|---|---|---|---|
| | | Exptl. | Computed from 1 ns MD | Computed[26] from 50 ns MD |
| Glc(1→6)GlcNAc | H-1c–H-2c | 1.2 | 2.1 | 1.2 |
| | H-1c–H-1d | 1.4 | 1.7 | 3.7 |
| | H-1c–(H-3c + H-5c) | 9 | 13 | 7.3 |
| | H-1c–H-6′b | 2.2 | 3.4 | 2.0 |
| | H-1c–H-6b | 3.8 | 5.4 | 3.7 |
| GlcNAc(1→3)Gal | H-1b–H-2b | 1.1 | 2.4 | 1.2 |
| | H-1b–H-4a | 0.4 | 0.7 | 0.5 |
| | H-1c–(H-3a + H-3b + H-5b) | 16 | 19 | 13 |
| | H-6′b–H-6b | 16 | 18 | 21 |
| | H-6′b–H-4b | 1.7 | 1.0 | 1.0 |
| | H-6′b–H-5b | 4.5 | 5.9 | 4.4 |
| Gal(1→4)GlcNAc | H-1d–H-2d | 1.5 | 2.3 | 1.2 |
| | H-1d–H-3d | 4.8 | 5.3 | 3.8 |
| | H-1d–H-5d | 7.6 | 6 | 4.5 |
| | H-1d–H-4b | 8.8 | 7 | 6.4 |
| | H-1d–H-6′b | 1 | 0.8 | 0.8 |
| | H-1d–(H-6b + H4d) | 2.7 | 2.7 | 2.5 |
| Neu5Ac(2→3)Gal | H-$3_{ax}$e-H-$3_{eq}$e | 16 | 20 | 21 |
| | H-$3_{ax}$e–H-4e | 2.9 | 3.1 | 1.6 |
| | H-$3_{ax}$e–H-5e | 3.3 | 5.1 | 2.9 |
| | H-$3_{ax}$e–H-3d | 2.5 | 4 | 6.3 |
| | H-$3_{eq}$e–H-4e | 5.5 | 5.8 | 3.6 |
| | H-$3_{eq}$e–H-5e | 0.8 | 2.3 | 0.8 |
| | H-$3_{eq}$e–H-3d | 0.8 | 1.9 | 1.6 |

[a] Computed from the MD trajectory with the experimentally derived mixing time (100 ms) and correlation time (2 ns).[14]

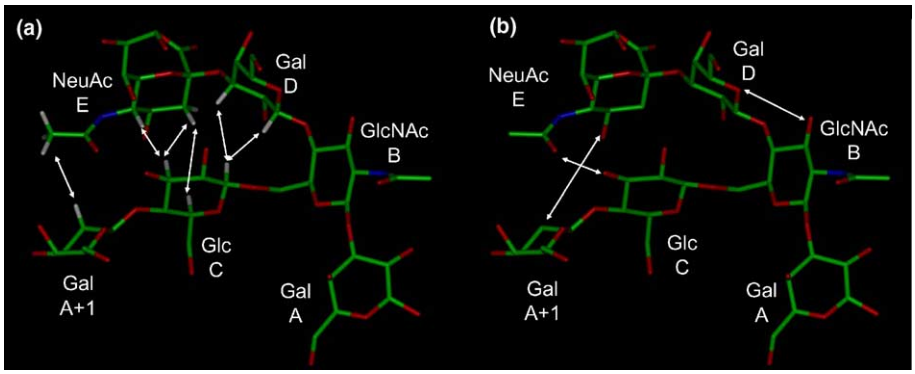


**Figure 7.** Illustration of the central repeat unit in the native CPS fragment indicating the inter-residue NOEs (panel a) and the hydrogen bond interactions (panel b).

GlcNAc residue (B) and the ring oxygen of the side chain galactosyl residue (D) (3.1 Å, 88% occupancy). Although these interactions were relatively weak, as indicated by their lengths, and frequently mediated by solvent molecules, as suggested by the occupancies, this network of interactions stabilizes the orientation of the side chain relative to the backbone in the native CPS. Perhaps more relevant to the relationship between sialylation and epitope structure is the corollary that the conformation of the backbone is stabilized by the presence of interactions with the side chains. The observation that chemical cleavage of the glyceryl side chain has no effect on the immunological properties of the CPS is consistent with the fact that the hydrogen-bond interactions between the sialic acid and the backbone do not involve the glyceryl side chain. No direct interaction with the carboxylate group and the backbone was observed during the simulation. Thus, the mechanism through which removal of the charge in the sialic acid, by chemical reduction of the carboxylate to a primary alcohol, influences antigenicity remains undetermined. It may be that the loss of the charge removes long-range electrostatic interactions in the CPS and thereby induces a conformational change in the epitope, or it may be that it is a loss of electrostatic interactions with the antibody surface that reduces the affinity for the antibody.

### 3.6. Generation of a CPS model

The explicit solvent simulation of the pentamer provided an approximate model for the behavior of the polysaccharide, to the extent that the fluctuations of the terminal residues could be ignored. However, to gain a better understanding of the long-range structure of the CPS, it was necessary to adopt an alternative approach. Under the assumption that the central repeat unit most accurately represented the conformation that would be present in the intact polysaccharide, an average structure for this unit was computed from the 50 ns trajectory. A model for the intact CPS was generated by replicating the resultant average central repeat unit, and repeatedly superimposing the coordinates of the terminal galactosyl residue of one unit with those of another. In this manner it was possible to generate a perfectly regular polymeric structure for the CPS (Fig. 8). Helical conformations of polysaccharides are often suggested by theoretical studies[14,32,33] and, as in the case of the GBS III CPS, they may be consistent with experimental NMR data, however helical polysaccharide conformations are generally extremely difficult to confirm unambiguously experimentally.[34] The helical model for the native CPS has a diameter of 29.3 Å and a pitch 89.5 Å, which encompasses 3 pseudo-helical repeats for a total of 34 backbone residues. These values may be compared, for reference, to a diameter of 20 Å and a pitch of 34 Å for B-DNA.[35] The CPS model predicts a rise of 2.6 Å per residue and a twist of 10.6° per residue. The side chains are placed around the periphery of the helix with the immunodominant region forming the core of the helix. Thus, our model infers that anti-GBS III antibodies recognize the core and not the surface of the helix. This mode of interaction is in contrast to that suggested for the O-specific polysaccharide from type 1 *Shigella*

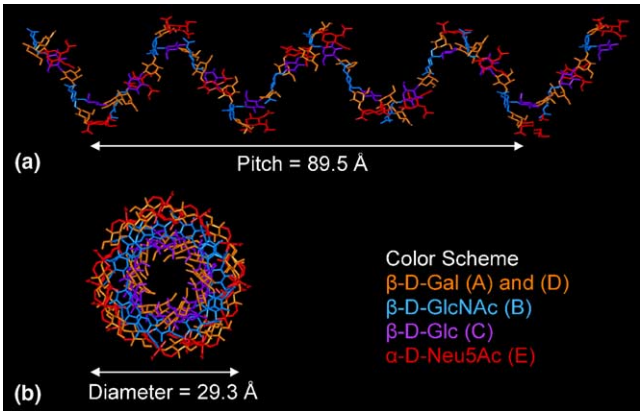


**Figure 8.** Native CPS model viewed along the principal helix axis (panel a) and down the helical axis (panel b).

*dysenteriae*, in which the antigen is exposed on the helical surface.[33] This difference is not surprising in light of the fact that the sialylated helical surface in the GBS III CPS consists of α-Neu5Ac-(2→3)-β-D-Gal side chains. This side chain sequence is also present on mammalian cell surfaces and is incapable of raising a strong immune response. While our model does not fully address helix stability or flexibility, it should provide a basis for future theoretical and experimental analysis of the CPS and of CPS–antibody interactions.

## 4. Conclusions

The results of 50 ns explicitly solvated MD simulations of fragments of native and desialylated type III group B *Streptococcus* CPS are presented. The theoretical results are in excellent agreement with NMR $^3J$ scalar coupling and NOE data reported earlier for single repeat units of the CPS fragments.[14] The conformation of the backbone of the monomeric units, whether native or desialylated, was indistinguishable from that present in the pentameric structure. This suggests that there is no conformational change associated with increasing the number of repeat units from 1 to 5. Rather, the immunological epitope arises from the formation of a larger structure with a helical conformation. While the structure of the CPS helix appears to be quite large, a single pseudo-repeat is predicted to consist of four pentasaccharide repeat units. This is significant in light of the experimental data for CPS binding that indicates a minimum size of three to four repeat units is required to achieve reasonable antigen binding and immunogenicity.

As evidenced by both the overall RMSD analysis and the torsional analysis of the glycosidic angles, removal of sialic acid causes the helical structure of the native CPS to decay to a more flexible and disordered state. The long lifetimes (20–40 ns) of the rotational states of the backbone $\omega$-angles in the desialylated derivative make accurate statistical sampling of the conformations challenging.

An analysis of experimental NOE intensities and theoretical hydrogen-bond interactions leads to the conclusion that specific interactions exist between the side chain sialic acid and the backbone glucosyl and galactosyl residues. These interactions both influence the orientation of the side chain and stabilize the conformation of the backbone. In the absence of these interactions the overall helical structure collapses, due to rotations of the 1→6 linkages, while the conformations of the remaining backbone linkages, which constitute the immunodominant region, are unaffected. These results provide an explanation for the observation that immunization with type III GBS induces two distinct antibody populations, which nevertheless share a common determinant in the backbone of the CPS.[10] In addition, the data from the simulation of the desialylated CPS confirm the role of the sialic acid in stabilizing the helical form. It appears likely that the major sialic acid dependant antibody population recognizes the immunodominant sequence as presented in the core of the helix, while the minor population, which is not dependant on sialic acid,[10] recognizes the same determinant in an alternative presentation, potentially that associated with random coil regions of the polysaccharide.

## Acknowledgements

We thank the National Institutes of Health (RR05357 and GM55230) for financial support, and R.J.W. wishes to thank Dr. H. J. Jennings for useful discussions.

## Supplementary data

Supplementary data in the form of mpeg video clips of the MD trajectories, as well as other data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2004.12.034.

## References

1. Schuchat, A. *Lancet* **1999**, *353*, 51–56.
2. Trivalle, C.; Martin, E.; Martel, P.; Jacque, B.; Menard, J. F.; Lemeland, J. F. *J. Med. Microbiol.* **1998**, *47*, 649–652.
3. Farley, M. M. *Clin. Infect. Dis.* **2001**, *33*, 556–561.
4. Wessels, M. R. *J. Appl. Microbiol., Symp. Suppl.* **1997**, *83*, 20S–31S.
5. Rubens, C. E.; Wessels, M. R.; Kuypers, J. M.; Kasper, D. L.; Weiser, J. N. *Semin. Perinatol.* **1990**, *14*, 22–29.
6. Baker, C. J.; Barrett, F. F. *J. Am. Med. Assoc.* **1974**, *230*, 1158–1160.
7. Liao, C. H.; Huang, L. M.; Lu, C. Y.; Lee, C. Y.; Hsueh, P. R.; Tsao, P. N.; Hsieh, W. S.; Tsou, K. I. *Acta Paediatr. Taiwan* **2002**, *43*, 326–329.
8. Wessels, M. R.; Rubens, C. E.; Benedi, V. J.; Kasper, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8983–8987.
9. Yeung, N. K.; Mattingly, S. J. *Infect. Immun.* **1983**, *42*, 141–151.
10. Jennings, H. J.; Lugowski, C.; Kasper, D. L. *Biochemistry* **1981**, *20*, 4511–4518.
11. Zou, W.; Mackenzie, R.; Thèrien, L.; Hirama, T.; Yang, Q.; Gidney, M. A.; Jennings, H. J. *J. Immunol.* **1999**, *163*(2), 820–825.
12. Wessels, M. R.; Pozsgay, V.; Kasper, D. L.; Jennings, H. J. *J. Biol. Chem.* **1987**, *262*, 8262–8267.
13. Kabat, E. A. *Structural concepts in immunology and immunochemistry*, 2nd ed.; Holt, Rinehart and Winston: New York, 1976.
14. Brisson, J.-R.; Uhrínova, S.; Woods, R. J.; van der Zwan, M.; Jarrell, H. C.; Paoletti, L. C.; Kasper, D. L.; Jennings, H. J. *Biochemistry* **1997**, *36*, 3278–3292.
15. Wessels, M. R.; Muñoz, A.; Kasper, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 9170–9174.

16. Naidoo, K. J.; Denysyk, D.; Brady, J. W. *Protein Eng.* **1997**, *10*(11), 1249–1261.
17. Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredson, C. H.; Duus, J. O. *J. Am. Chem. Soc.* **2001**, *123*, 4792–4802.
18. Eklund, R.; Widwalm, G. *Carbohydr. Res.* **2003**, *338*, 393–398.
19. Corzana, F.; Motawia, M. S.; Herve du Penhoat, C.; Perez, S.; Tschampel, S.; Woods, R. J.; Engelsen, S. B. *J. Comp. Chem.* **2004**, *25*(4), 573–586.
20. Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*(19), 10541–10545.
21. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002.
22. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832–3846.
23. Stenutz, R.; Carmichael, I.; Widmalm, G.; Serianni, A. S. *J. Org. Chem.* **2002**, *67*, 949–958.
24. Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **2001**, *123*, 4781–4791.
25. Tvaroska, I.; Hricovini, H.; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359–362.
26. James, T. L.; Leu, H. E. *CORMA*; University of California: San Francisco, 1999.
27. Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformations of Carbohydrates*; Harwood Academic: Amsterdam, 1998.
28. Poppe, L.; Dabrowski, J.; von der Lieth, C. W.; Koike, K.; Ogawa, T. *Eur. J. Biochem.* **1990**, *189*, 313–325.
29. Rutherford, T. J.; Spackman, D. G.; Simpson, P. J.; Homans, S. W. *Glycobiology* **1994**, *4*, 59–68.
30. Imberty, A.; Delage, M.-M.; Bourne, Y.; Cambillau, C.; Pérez, S. *Glyconj. J.* **1991**, *8*, 456–483.
31. Bernardi, A.; Colombo, A.; Sanchez-Medina, I. *Carbohydr. Res.* **2004**, *339*, 967–973.
32. Perez, S.; Kouwijzer, M.; Mazeau, K.; Engelsen, S. B. *J. Mol. Graph.* **1996**, *14*, 307–321.
33. Nyholm, P. G.; Mulard, L. A.; Miller, C. E.; Lew, T.; Olin, R.; Glaudermans, C. P. *Glycobiology* **2001**, *11*, 945–955.
34. McIntire, T. M.; Penner, R. M.; Brant, D. A. *Macromolecules* **1995**, *28*, 6375–6377.
35. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.