

Preferred Conformations of Lipooligosaccharides and Oligosaccharides of *Moraxella catarrhalis*

Ya Gao^{1,2}, Jumin Lee², Göran Widmalm^{3,*} and Wonpil Im^{2,4,*}

¹School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science,
Shanghai 201620, China

²Departments of Biological Sciences and Bioengineering, Lehigh University, Bethlehem, PA 18015,
USA

³Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-10691
Stockholm, Sweden

⁴School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Republic of
Korea

*Correspondence: wonpil@lehigh.edu and goran.widmalm@su.se

Abstract

Moraxella catarrhalis (*M. catarrhalis*) is a pathogenic gram-negative bacterium that causes otitis media and sinusitis in children. Three major serotypes A, B, and C are identified to account for approximately 95% of the clinical isolates. Understanding the conformational properties of different serotypes of *M. catarrhalis* provides insights into antigenic determinants. In this work, all-atom molecular dynamics simulations were conducted for *M. catarrhalis* lipooligosaccharides (LOS) bilayer systems and oligosaccharides (OS) in water solution to investigate the conformational similarities and differences of three serotypes. For up to ten neutral monosaccharides in the core part, the conformational ensembles described by the pairwise root mean square deviation distributions are similar among the three serotypes of either the LOS or OS. At the central β -(1 $\rightarrow$ 4)-linkage, anti- ψ conformation in conjunction with the gauche-gauche (g^-) conformation of the central trisubstituted glucosyl residue is observed as the dominant conformation to sustain the structural characteristics of *M. catarrhalis* three types, which is further supported by calculated transglycosidic $^3J_{C,H}(\psi_H)$ of serotype A in comparison to experimental data. Interestingly, the conformational variability of three serotypes is more restricted for the OS in water solution than that in the LOS bilayer systems. The LOS-LOS interactions in the bilayer systems are responsible for the increased conformational diversity despite of tight packing. Solvent-accessible surface area analysis suggests that a trisaccharide attached to the β -(1 $\rightarrow$ 6)-linked sugar in all three serotypes of LOS could be the common epitope and have the possibility to interact with antibodies.

Introduction

Moraxella catarrhalis (*M. catarrhalis*), a gram-negative bacterium, is an important human pathogen causing otitis media in children and lower respiratory tract infections in adults, in particular those with chronic obstructive pulmonary disease (Hays 2009), exemplified by an outbreak of *M. catarrhalis* in a respiratory unit (Richards et al. 1993). The carrier rate in healthy children is more than 50% where it forms a reservoir of repeated infections in childhood. While most gram-negative species carry lipopolysaccharides (LPS) with long O-antigen polysaccharides, the outer leaflet of the outer membrane of *M. catarrhalis* instead has short-chain LPS referred to as lipooligosaccharides (LOS) (Holme et al. 1999).

Most LOS of *M. catarrhalis* are divided into three serotypes A, B, and C with a conventional lipid A structure (Masoud et al. 1994a), two 2-keto-3-deoxyoctulosonate (Kdo) residues, and up to ten neutral sugar residues, but the structures are devoid of heptose residues that are commonly observed in the core region of LPS. These major serotypes A, B, and C account for approximately 95% of the clinical isolates (Vaneechoutte et al. 1990). Serotype A contains eight neutral hexose residues (Edebrink et al. 1994; Masoud et al. 1994b), serotype B has six glycoforms (Edebrink et al. 1996), and serotype C has been demonstrated to contain four different oligosaccharides (Edebrink et al. 1995), with the largest structures of each serotype schematically shown in **Figure 1**. These LOS have experimentally been shown to present unusual carbohydrate three-dimensional structures (Masoud et al. 1994b; Lycknert et al. 2004; Frank et al. 2015). A proposed conserved common epitope, yet to be elucidated in detail, from the inner oligosaccharide region is presently being pursued as a means to make a vaccine covering all three serotypes (Oishi et al. 1996; Cox et al. 2011; Ren et al. 2011; Blakeway et al. 2017).

In this work, to gain insight into the LOS conformational similarities and differences of the three serotypes and also to discern antigenic determinants, we have performed molecular dynamics (MD) simulations of LOS bilayer systems and oligosaccharides (OS) in solution (**Figure 2**).

Methods

Simulation Details

M. catarrhalis lipooligosaccharides (LOS) and oligosaccharides (OS) simulations

For MD simulations of LOS, symmetric bilayer systems of three serotypes, in which each leaflet contains 50 LOS, were generated using *Membrane Builder* (Jo et al. 2007; Jo et al.

2009; Wu et al. 2014; Lee et al. 2016; Lee et al. 2019) in CHARMM-GUI (<http://charmm-gui.org>) (Jo et al. 2008). During the building procedure, Ca^{2+} ions were added to the LOS region (above lipid A) to neutralize the system and 150 mM KCl was also added to the bulk region to mimic the bulk ion solution. Three independent replicas with different random seed numbers were generated for each serotype to improve sampling and to check the simulation convergence. The force field parameters for each system were assigned from the CHARMM36 force field for LPS (Jo et al. 2015; Kim et al. 2016), lipids (Klauda et al. 2010), and carbohydrates (Guvench et al. 2008; Guvench et al. 2009; Hatcher et al. 2009; Guvench et al. 2011). Equilibration runs followed by production simulations were then conducted. During the equilibration, NVT dynamics was first used and subsequently followed by NPT simulations. Following the *Membrane Builder* equilibration protocol, the decreasing restraints in 6 steps were applied to the LOS and water molecules to assure gradual equilibration of the assembled systems. After equilibration, 1- μs production simulation with 2-fs time step was conducted for each replica and each serotype. All bonds to hydrogen atoms were fixed using the SHAKE algorithm (Ryckaert et al. 1977). The van der Waals interactions were smoothly switched off at 10-12 Å by a force-switching function (Steinbach and Brooks 1994) and the long-range electrostatic interactions were calculated using the particle-mesh Ewald method (Essmann et al. 1995). Langevin dynamics was used for the temperature coupling with a collision frequency of 1 ps^{-1} . A semi-isotropic Monte Carlo (MC) barostat method with pressure coupling frequency of 100 steps was used to model the pressure (Chow and Ferguson 1995; Aqvist et al. 2004). The temperature was maintained at 310.15 K and the pressure was set to 1.0 bar.

For *M. catarrhalis* OS simulations, initial structures for each serotype were generated based on the clustering analysis of LOS bilayer simulations (see below). Each representative structure of each serotype was modeled from the top 10 clusters. Three independent replicas for cluster 1 and one replica from each remaining 9 clusters were generated and simulated to verify the convergence of simulations. In the OS simulations, lipid A and two Kdo residues were excluded, and the trisubstituted Glc was modified to an *O*-methyl glucoside. The initial OS structure was immersed in the center of a rectangular box of TIP3P water molecules (Jorgensen et al. 1983), and all of the OS atoms were at least 10 Å away from the boundary of the water box; 150 mM KCl was added to mimic the ion concentration. Before the production simulations, 5,000 steps of energy minimization and a 25-ps equilibration with a time step of 1 fs in the NVT ensemble were conducted to minimize and relax the system, for which weak positional and dihedral restraints were applied to the OS structure. After equilibration, a 1- μs

NPT production simulation was performed for each replica and each serotype with a temperature of 303 K and a pressure at 1.0 bar. All the LOS bilayer and OS solution simulations were conducted utilizing OpenMM (Eastman et al. 2013).

Clustering of conformational states using glycosidic torsion angles

Clustering based on the glycosidic torsion angle was utilized to explore the conformational ensemble and to obtain representative structures. The clustering method has already been employed in our previous work (Jo et al. 2016) and we therefore give only a brief description of the methodology. For each glycosidic linkage, several well-defined basins were roughly identified by examining the torsion angle distributions. Then, the basins were refined by assigning the torsion angles observed during the simulations to the nearest basin using the *k*-medoid algorithm (Park and Jun 2009). We denoted the torsional states of each glycosidic linkage based on the size of the basins, i.e., “A” refers to the largest basin, “B” the second largest basin, and so on. For the ω torsion angle, the basins were denoted according to the staggered rotameric states of the ω torsion angle along the C6-C5 bond: g^- (gauche-gauche), g^+ (gauche-trans), and t (trans-gauche). By combining the torsion angle states, the conformation of each serotype can be described with a series of letter notation. Note that conformational states of an ω torsion angle are given in lower case letters (after the glycosidic linkage conformational state of the sugar residue) and the other residues for each linkage are denoted by upper case letters, beginning with the letter A, followed by B and so on; e.g., one of the states in serotype B is described by $Ag^-Ag^+ABg^-AAAAAAAAA$, i.e., 13 glycosidic linkages in total (cf. Fig. S4).

Results and Discussion

Lipooligosaccharide conformational distribution and preferred conformational states

To evaluate if each replica has reached its equilibration, the X or Y length of each bilayer system and hydrophobic thickness were measured. The hydrophobic thickness is defined by the average Z values between carbon atoms in the upper and lower leaflets; **Figure S1** shows these carbons in red circles. After about 500 ns, both the X or Y length and hydrophobic thickness of each replica of the three serotypes (**Figure S2**) have stabilized, indicating simulations have reached its equilibrium. The equilibrated average X or Y lengths for serotypes B and C are both larger than that for serotype A because there are more sugar residues in serotypes B and C, i.e., a volume effect. Furthermore, serotype C has a wider X or

Y dimension than serotype B, indicating more flexible and diverse conformations sampled for serotype C. For the hydrophobic thickness, there is no obvious difference after about 500 ns among three serotypes, indicating that all three serotypes did not undergo dramatic conformational change in lipid A region and systems reached their stabilization.

Pair-wise root mean square deviations (RMSD), measuring the conformational variability among the sampled conformations (without lipid A and the two Kdo units), were calculated for three serotypes using the last 500-ns simulation trajectories. As shown in **Figure S3A**, the LOS conformations of all three serotypes in the bilayer systems are dynamic and flexible during simulations with broad pair-wise RMSD distributions (up to 6 Å for serotype A and up to 8 Å for serotypes B and C). Despite of minor differences in the peak's locations of each replica's RMSD distribution, similar conformational ensembles were sampled among the three replicas. In the following, three replicas for each serotype are combined to have 3- μ s ensemble sampling for further analysis. To gain further insight into the structural characteristics of the three serotypes, especially the structural similarities and differences, pair-wise RMSDs for the common segment (residue 4 to 11 in **Figure 1**) of three serotypes were calculated. Overall, as shown in **Figure 3A**, serotypes A, B, and C share similar distributions with pair-wise RMSD up to 6 Å. The RMSD peak is mainly located at about 4.0 Å for serotype A although there is also a smaller peak at about 1.8 Å. For serotypes B and C, the peak is located at about 2.4 Å and 3.8 Å, respectively, in a bimodal kind of distribution. The pair-wise RMSD analysis indicates that structures of the common segment of three serotypes in the LOS bilayer systems are flexible and diverse.

Next, a clustering method based on the glycosidic torsion angle was employed to investigate the conformational features and representative structures of LOS from three serotypes; see Methods for details. Different from clustering based on RMSD, this clustering method could provide the preference for each glycosidic torsion angle and can be used for comparison among the LOS conformations of the three serotypes. The torsion angle (ϕ and ψ) distributions with a set of conformational basins for the three serotypes were classified (**Figure S4**) utilizing the *k*-medoid algorithm. Each torsional angle populated a similar region among all three serotypes. The top 10 conformational states and their population in each serotype from LOS bilayer simulations are shown in **Table S1**. Due to the β -(1 $\rightarrow$ 6) glycosidic linkage in lipid A, as well as the α -(2 $\rightarrow$ 6) glycosidic linkage between residue 2 and lipid A, conformations for each serotype are very flexible as indicated by small population of each cluster. The accumulated populations of top 10 clusters for serotypes B (23%) and C (16%) are

smaller than that for serotype A (37%) due to the difference in the number of sugar residues. Thus, similar to the pair-wise RMSD analysis, clustering for the common segment (residue 4 to 11) of the three serotypes was also conducted.

As shown in **Table 1**, a conformational ensemble of the common part for serotypes A and B is quite similar despite the fact that the galactosyl disaccharide is linked to a glucosyl residue (#11), corresponding to an N-acetyl-D-glucosamine residue in the former serotype. The top 10 clusters cover a large portion of the conformational ensemble with a population of 61% (serotype A) and 60% (serotype B), respectively. However, for serotype C, the accumulated population for top 10 clusters is 45%, which is smaller compared to serotypes A and B, showing that the conformational ensemble sampled for serotype C is more diverse than those for serotypes A and B. By further inspection, such a diverse conformational ensemble in serotype C appears to stem from the sampling of replica 2, which is shown in **Figure S5A**. For residue 11 in serotype C, a certain amount of anti- ψ conformations are sampled, which is not observed in replicas 1 and 3; this sampling occurs to a small extent for serotypes A and B. With the formation of this anti- ψ conformation, additional hydrogen bonds, such as HO2@Glc7-O6@GlcNAc11, HO6@GlcNAc11-O3@Glc4, and HO2@Gal12-O3@Glc6 with a population 24%, 19%, and 19%, respectively, are formed to stabilize this structure (**Figure S5C**). Thus, residue 11 in serotype C makes a more diverse conformational ensemble compared to serotypes A and B.

A previous study shows that a minimum of two sugar residues (10 and 11) extending from O4 of the central residue as well as two residues (6 and 7) extending from O6 of the same residue are needed to maintain the three-dimensional scaffold characteristic of *M. catarrhalis* oligosaccharides in solution (Lycknert et al. 2004). Also the fact that the central glucosyl residue in the *M. catarrhalis* LOS is additionally substituted at O3 by a glucosyl group leads at O3 and O4 to a vicinal di-substitution pattern (Soderman et al. 1998), an arrangement that already at the trisaccharide level results in highly flexible structures where the anti- ψ conformers become accessible on the potential energy surface (Yang et al. 2016). In our simulations, an antiperiplanar conformation of ψ is dominant for all three serotypes (cf. population distribution of glycosidic torsions for residue 6 in **Figure S4**). In addition, in residue 4 the distribution of the ω torsion angle in **Figure 4** shows an almost exclusive gauche-gauche (g^-) conformations for all three serotypes of LOS. For the β -(1 $\rightarrow$ 4)-linkage joining residues 4 and 10, two-dimensional distributions of ϕ and ψ torsion angles are plotted in **Figure 5A**. The torsion angle ϕ predominantly populates the exo-syn conformation. For the ψ torsion

angle, the antiperiplanar conformation is observed as the dominant one and in addition there is a small portion of the populated state between -120° and 0° . In all three serotypes, the two linkages share the similar torsional angle distribution, which maintains the characteristic oligosaccharide structure of *M. catarrhalis* LOS. The representative structures for each serotype from residue 4 (Bg⁻AAAAAAA for serotype A and Ag⁻AAAAAAAAA for serotype B and C) are shown in **Figure 3C**. It is noteworthy that the dominant anti- ψ conformation (Lycknert et al. 2004) and the gauche-gauche (g⁻) conformation of the β -(1 $\rightarrow$ 4)- and β -(1 $\rightarrow$ 6)-linkages, respectively, is consistent with long-range NOEs in NMR experiments, i.e., residue 11 folds back to sustain a compact conformation (Frank et al. 2015). Hydrogen bond analysis for the three serotypes (**Figure S5B**) reveal hydrogen bonds in the two branches, for example, HO3@GlcNAc11-O5@Gal12, HO3@Glc7-O5@Gal8, HO3@Gal12-O5@Gal13, and HO3@Gal8-O5@Gal9, occupy the dominate population. To some extent, hydrogen bonds exist in-between different ‘arms’ of the serotype-specific structures.

In search of the virulence factors of *M. catarrhalis*, attention has been paid to surface exposed antigens such as LOS, outer membrane proteins, and pili (Murphy and Bartos 1989; Fomsgaard et al. 1991; Rikitomi et al. 1991; Holme et al. 1999). In this work, solvent-accessible surface area (SASA) (Lee and Richards 1971) of each residue in each serotype was calculated to determine the potential glycan residues that may be interacting with antibodies. As shown in **Figure S6**, the central trisubstituted residue 4 has the smallest SASA value for all three serotypes, indicating that it is buried in the LOS molecule and not accessible to antibodies. Residues 5, 6, and 10, which all substitute residue 4 and compose the inner part of three serotypes, have similar SASA with about 42 Å², 39 Å², and 22 Å² at the peak, respectively. For the outer part, residues 7, 8 in three serotypes and 11, 12 in serotypes B and C are more exposed to water with larger SASA values than the inner part. Besides, due to the lack of residues 12 and 13 in serotype A, residues 7 and 8 in serotype A are more exposed to solvent compared to serotypes B and C. For terminal residues, 9 in three serotypes, 11 in serotype A, and 13 in serotypes B and C all have the largest SASA values at the peak, indicating that these terminal residues are exposed to solvent extensively. Analysis from SASA calculations indicates that one ‘arm’ (residue 7 to 9) is more exposed to solvent compared to the inner part in all three serotypes and could serve as a common epitope interacting with antibodies for three serotypes. Conversely, residues 10 - 13 in the middle ‘arm’ extending from the central residue 4 are proposed to lead to different and unique epitopes being the basis for the three serotypes in *M. catarrhalis*.

Oligosaccharide conformational distribution and preferred conformational states

NMR spectroscopy for serotype A was recorded in D₂O solution at 25 °C (Masoud et al. 1994b). To compare with NMR experiment, and also explore different structural characteristics between LOS in a bilayer and OS in water solution, simulations of the three serotype OS in water solution were performed. A representative snapshot in water solution is shown in **Figure 2B**. Pair-wise RMSD distribution in each independent simulation for each serotype was calculated (**Figure S3B**), and comparisons between simulation 1 (3 replicas starting from LOS cluster 1) and simulations 2-10 (1 replica each starting from LOS cluster 2-10) are shown in **Figure S7**. Clearly, the pair-wise RMSD distributions among different replicas and between simulation 1 and simulations 2-10 are quite similar, indicating the convergence of the OS simulations in water solution. In the following, trajectories from 12 replicas (12- μ s in total) are combined for analysis.

Clustering based on the glycosidic torsion angle was also conducted for three types of OS. To have a consistent comparison between OS and LOS, population was also calculated for the common segment (residue 4 to 11), and the population for top 10 clusters is shown in **Table 2**. The definition of each state is the same as in the bilayer simulations. The most representative structure for each serotype is consistent with that in LOS bilayer simulations, g⁻AAAAAAA for serotypes A, B, and C. The anti- ψ and gauche-gauche (g⁻) conformations for the β -(1 $\rightarrow$ 4)- and β -(1 $\rightarrow$ 6)-linkages, respectively, are still observed as a dominate conformation in all three serotypes shown in **Figure 4** and **Figure 5B**. However, it should be noted that a decreased sampling in the g⁻ region and an increased sampling in the g⁺ region is detected in the OS simulations for ω of the β -(1 $\rightarrow$ 6)-linkage. However, for the β -(1 $\rightarrow$ 4)-linkage ψ is nearly exclusively populated in the anti-conformation compared to that in the bilayer simulations.

Interestingly, the conformational variation of OS in water solution is more restricted than that of LOS in bilayer simulations despite of the tight packing in LOS bilayer systems. For example, the top 10 accumulated populations (ω of residue 4 and (ϕ , ψ) of residues 5 to 11) are 83% (serotypes A OS), 73% (B), and 84% (C) in solution, respectively, while they are 70% (serotypes A LOS), 69% (B), and 58% (C) in bilayers. Comparison of the population for each cluster between OS and LOS is also shown in **Figure 6**, indicating a restricted conformational ensemble in water solution. Our previous study also shows a more restricted conformational variability for N-glycan pentasaccharide free in solution than for the N-glycan on the protein surface (Jo et al. 2016). The interaction among LOS is responsible for the

increased conformational diversity compared to OS in water solution. For serotypes A, B, and C bilayer systems each containing 100 LOS molecules, the average number of inter-LOS hydrogen bonds is 91.3 (serotype A), 129.0 (B), and 135.0 (C), respectively. The intermolecular interaction causes a more flexible conformational ensemble compared to the isolate in water solution.

For serotype A OS, the transglycosidic $^3J_{C,H}(\psi_H)$ from C1 in residue 10 to H4 in residue 4 was determined to be 7.4 Hz with NMR spectroscopy (Masoud et al. 1994b). The uncertainty from NMR experiment is estimated to be 0.25 Hz. $^3J_{C,H}(\psi_H)$ was calculated for snapshots extracted from the 1- μ s trajectories in each replica utilizing the Karplus equation with an estimated uncertainty of 0.50 Hz (Sawen et al. 2010).

$$^3J_{C,H}(\psi_H) = 6.54\cos^2(\psi_H) - 0.62\cos(\psi_H) + 0.33 \\ + 0.6\exp(k\cos(\phi_{O5'} - 180))/\exp(\kappa)$$

where κ is set to 8. The calculated $^3J_{C,H}(\psi_H)$ is 6.70 ± 0.32 Hz, which is within the uncertainty (0.75 Hz) of experimental value, indicating the consistency in the structure ensemble with experiment. However, there still exists some deviation from the experimental value. According to the uncertainty from NMR experiment and Karplus equation, the J coupling distribution for all trajectories (120,000 snapshots for 12 replicas) is divided into three regions. As shown in **Figure 7A**, in which the red and green regions mean a deviation within 0.25 Hz and 0.25 - 0.75 Hz, respectively. The cyan region means a deviation greater than 0.75 Hz. The corresponding torsion angle (ϕ and ψ) distribution for the β -(1 $\rightarrow$ 4)-linkage and representative structures are also shown in **Figure 7B**. The population for structures within 0.75 Hz uncertainty is 67%, with the averaged ψ torsion angles of 174° (red), 164° and -163° (green), respectively. For the remaining 33% structures larger than 0.75 Hz uncertainty, despite of similar structures with the red and green ones (with RMSD of ~ 2 Å), it exhibits a more diverse and flexible conformational ensemble with the averaged ψ torsion angles of -146° and 155° . The $^3J_{C,H}(\psi_H)$ coupling constant analysis supports the anti- ψ conformation (near 180°) in the OS and that it is highly populated in solution.

Conclusions

M. catarrhalis is now increasingly recognized as an important human pathogen causing otitis media in children and lower respiratory tract infections in adults. There are three serological types (A, B, and C) with differences in the composition of glycan residues and

lengths of branches. To gain insights into the conformational similarities and differences of three serotypes and also to provide hints to the antigenic determinants stimulating antibody formation in humans, MD simulations of LOS bilayer systems and OS in water solution were conducted to investigate the conformational properties.

For LOS in the bilayer systems, the conformational ensemble distributions for the three serotypes are quite similar despite the slightly more diverse structures observed for serotype C. Based on the clustering analysis of torsional angles of each linkage, anti- ψ in the β -(1 $\rightarrow$ 4)-linkage (between residue 4 and 10) and gauche-gauche (g^-) of ω in the β -(1 $\rightarrow$ 6)-linkage (between residue 4 and 6) were observed for all three serotypes as a dominant structure to maintain the characteristics of *M. catarrhalis* LOS. Hydrogen bonds between sugars in the branches linked to position 4 and to position 6 of the central residue for three serotypes are also identified. Based on the SASA for each serotype the branch containing residues 7 to 9 in three serotypes of LOS could be the common epitope and have the possibility to interact with antibodies.

To compare with NMR experiment, and also to explore different structural characteristics between LOS in bilayer systems and OS in water solution, simulations of three serotype OS in water solution were performed. Notably, the conformational variation of OS in water solution is more restricted than that of LOS in bilayer simulations, in spite of the tight packing in LOS bilayer systems. Interestingly, it is consistent with our previous study showing more restricted conformational variability for N-glycan pentasaccharide free in solution than the N-glycan on the protein surface due to protein-N-glycan interactions. For OS in water solution, the significance of the anti- ψ and gauche-gauche (g^-) conformational states lead to the dominant conformation for the β -(1 $\rightarrow$ 4)- and β -(1 $\rightarrow$ 6)-linkage at the central multibranched residue, resulting in folded structures in all three serotypes, which is similar with that in LOS. This combination of anti- ψ and g^- states is necessary to maintain the in LOS and OS three-dimensional structures of *M. catarrhalis* three serotypes. However, the ψ torsion at the β -(1 $\rightarrow$ 4)-linkage is almost exclusively populated in OS compared to that in LOS bilayer systems. Furthermore, J coupling analysis for the OS conformational ensembles also supported that the anti- ψ conformation (near 180°) is preferred in OS structures, which gives credence to the existence of this state in LOS as identified by the MD simulations.

Acknowledgement

We thank to Hongjing Ma for initial setup and simulations. This work was supported in part by grants from the NSF MCB-1727508, NSF MCB-1810695, XSEDE MCB070009, Friedrich Wilhelm Bessel Research Award from the Humboldt foundation (to WI), and the Swedish Research Council 2017-03703 (GW).

References

- Aqvist J, Wennerstrom P, Nervall M, Bjelic S, Brandsdal BO. 2004. Molecular dynamics simulations of water and biomolecules with a monte carlo constant pressure algorithm. *Chem Phys Lett* **384**: 288-294.
- Blakeway LV, Tan A, Peak IRA, Seib KL. 2017. Virulence determinants of moraxella catarrhalis: Distribution and considerations for vaccine development. *Microbiology-Sgm* **163**: 1371-1384.
- Chow KH, Ferguson DM. 1995. Isothermal isobaric molecular-dynamics simulations with monte-carlo volume sampling. *Comput Phys Commun* **91**: 283-289.
- Cox AD, Michael FS, Cairns CM, Lacelle S, Filion AL, Neelamegan D, Wenzel CQ, Horan H, Richards JC. 2011. Investigating the potential of conserved inner core oligosaccharide regions of moraxella catarrhalis lipopolysaccharide as vaccine antigens: Accessibility and functional activity of monoclonal antibodies and glycoconjugate derived sera. *Glycoconj J* **28**: 165-182.
- Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, Beauchamp KA, Lane TJ, Wang LP, Shukla D et al. 2013. Openmm 4: A reusable, extensible, hardware independent library for high performance molecular simulation. *J Chem Theory Comput* **9**: 461-469.
- Edebrink P, Jansson PE, Rahman MM, Widmalm G, Holme T, Rahman M. 1995. Structural studies of the o-antigen oligosaccharides from 2 strains of moraxella-catarrhalis serotype-c. *Carbohydr Res* **266**: 237-261.
- Edebrink P, Jansson PE, Rahman MM, Widmalm G, Holme T, Rahman M, Weintraub A. 1994. Structural studies of the o-polysaccharide from the lipopolysaccharide of moraxella (branhamella) catarrhalis serotype-a (strain atcc-25238). *Carbohydr Res* **257**: 269-284.
- Edebrink P, Jansson PE, Widmalm G, Holme T, Rahman M. 1996. The structures of oligosaccharides isolated from the lipopolysaccharide of moraxella catarrhalis serotype b, strain ccug 3292. *Carbohydr Res* **295**: 127-146.
- Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh ewald method. *J Chem Phys* **103**: 8577-8593.
- Fomsgaard JS, Fomsgaard A, Hoiby N, Bruun B, Galanos C. 1991. Comparative immunochemistry of lipopolysaccharides from branhamella-catarrhalis strains. *Infect Immun* **59**: 3346-3349.
- Frank M, Collins PM, Peak IR, Grice ID, Wilson JC. 2015. An unusual carbohydrate conformation is evident in moraxella catarrhalis oligosaccharides. *Molecules* **20**: 14234-14253.
- Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, Mackerell AD. 2008. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* **29**: 2543-2564.
- Guvench O, Hatcher E, Venable RM, Pastor RW, MacKerell AD. 2009. Charmm additive all-atom force field for glycosidic linkages between hexopyranoses. *J Chem Theory Comput* **5**: 2353-2370.
- Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, Jamison FW, MacKerell AD. 2011. Charmm additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **7**: 3162-3180.
- Hatcher E, Guvench O, MacKerell AD. 2009. Charmm additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* **113**: 12466-12476.
- Hays JP. 2009. Moraxella catarrhalis: A mini review. *J Pediatric Infect Dis* **04**: 211-220.

- Holme T, Rahman M, Jansson PE, Widmalm G. 1999. The lipopolysaccharide of moraxella catarrhalis - structural relationships and antigenic properties. *Eur J Biochem* **265**: 524-529.
- Jo S, Kim T, Im W. 2007. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *Plos One* **2**: e880.
- Jo S, Kim T, Iyer VG, Im W. 2008. Software news and updates - charnm-gui: A web-based graphical user interface for charmm. *J Comput Chem* **29**: 1859-1865.
- Jo S, Lim JB, Klauda JB, Im W. 2009. Charmm-gui membrane builder for mixed bilayers and its application to yeast membranes. *Biophys J* **97**: 50-58.
- Jo S, Qi YF, Im W. 2016. Preferred conformations of n-glycan core pentasaccharide in solution and in glycoproteins. *Glycobiology* **26**: 19-29.
- Jo S, Wu EL, Stuhlsatz D, Klauda JB, MacKerell AD, Widmalm G, Im W. 2015. Lipopolysaccharide membrane building and simulation. In *Glycoinformatics*, (ed. T Lütke, M Frank), pp. 391-406. Springer New York, New York, NY.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**: 926-935.
- Kim S, Patel DS, Park S, Slusky J, Klauda JB, Widmalm G, Im W. 2016. Bilayer properties of lipid a from various gram-negative bacteria. *Biophys J* **111**: 1750-1760.
- Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, Pastor RW. 2010. Update of the charmm all-atom additive force field for lipids: Validation on six lipid types. *J Phys Chem B* **114**: 7830-7843.
- Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* **55**: 379-IN374.
- Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi YF et al. 2016. Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *J Chem Theory Comput* **12**: 405-413.
- Lee J, Patel DS, Stahle J, Park SJ, Kern NR, Kim S, Lee J, Cheng X, Valvano MA, Holst O et al. 2019. Charmm-gui membrane builder for complex biological membrane simulations with glycolipids and lipoglycans. *J Chem Theory Comput* **15**: 775-786.
- Lycknert K, Edebrink P, Widmalm G. 2004. A conformational carbohydrate scaffold is present in the short-chain lipopolysaccharides of moraxella catarrhalis. *Angew Chem Int Ed* **43**: 2288-2290.
- Masoud H, Ferry MB, Richards JC. 1994a. Characterization of the lipopolysaccharide of moraxella-catarrhalis structural-analysis of the lipid-a from m catarrhalis serotype-a lipopolysaccharide. *Eur J Biochem* **220**: 209-216.
- Masoud H, Perry MB, Brisson JR, Uhrin D, Richards JC. 1994b. Structural elucidation of the backbone oligosaccharide from the lipopolysaccharide of moraxella-catarrhalis serotype-a. *Can J Chem* **72**: 1466-1477.
- Murphy TF, Bartos LC. 1989. Surface-exposed and antigenically conserved determinants of outer-membrane proteins of branhamella-catarrhalis. *Infect Immun* **57**: 2938-2941.
- Oishi K, Tanaka H, Sonoda F, Borann S, Ahmed K, Utsunomiya Y, Watanabe K, Nagatake T, Vanechoutte M, Verschraegen G et al. 1996. A monoclonal antibody reactive with a common epitope of moraxella (branhamella) catarrhalis lipopolysaccharides. *Clin Diagn Lab Immunol* **3**: 351-354.
- Park HS, Jun CH. 2009. A simple and fast algorithm for k-medoids clustering. *Expert Syst Appl* **36**: 3336-3341.
- Ren DB, Yu SQ, Gao S, Peng DX, Petralia RS, Muszynski A, Carlson RW, Robbins JB, Tsai CM, Lim DJ et al. 2011. Mutant lipooligosaccharide-based conjugate vaccine

- demonstrates a broad-spectrum effectiveness against *Moraxella catarrhalis*. *Vaccine* **29**: 4210-4217.
- Richards SJ, Greening AP, Enright MC, Morgan MG, McKenzie H. 1993. Outbreak of *Moraxella catarrhalis* in a respiratory unit. *Thorax* **48**: 91-92.
- Rikitomi N, Andersson B, Matsumoto K, Lindstedt R, Svanborg C. 1991. Mechanism of adherence of *Moraxella (Branhamella) catarrhalis*. *Scand J Infect Dis* **23**: 559-567.
- Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *J Comput Phys* **23**: 327-341.
- Sawen E, Massad T, Landersjö C, Damberg P, Widmalm G. 2010. Population distribution of flexible molecules from maximum entropy analysis using different priors as background information: Application to the ϕ , ψ -conformational space of the α -(1 $\rightarrow$ 2)-linked mannose disaccharide present in n- and o-linked glycoproteins. *Org Biomol Chem* **8**: 3684-3695.
- Söderman P, Jansson PE, Widmalm G. 1998. Synthesis, NMR spectroscopy and conformational studies of the four anomeric methyl glycosides of the trisaccharide d-glcp-(1 $\rightarrow$ 3) d-glcp-(1 $\rightarrow$ 4) - α -d-glcp. *J Chem Soc, Perkin Trans 2*: 639-648.
- Steinbach PJ, Brooks BR. 1994. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem* **15**: 667-683.
- Vaneechoutte M, Verschraegen G, Claeys G, Vandenabeele AM. 1990. Serological typing of *Branhamella catarrhalis* strains on the basis of lipopolysaccharide antigens. *J Clin Microbiol* **28**: 182-187.
- Wu EL, Cheng X, Jo S, Rui H, Song KC, Davila-Contreras EM, Qi YF, Lee JM, Monje-Galvan V, Venable RM et al. 2014. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J Comput Chem* **35**: 1997-2004.
- Yang MJ, d'Ortoli TA, Sawen E, Jana M, Widmalm G, MacKerell AD. 2016. Delineating the conformational flexibility of trisaccharides from NMR spectroscopy experiments and computer simulations. *Phys Chem Chem Phys* **18**: 18776-18794.

Table 1. Top 10 conformational states of each serotype (residue 4 to 11) in LOS bilayer systems.

#	serotype A (%)		serotype B (%)		serotype C (%)	
1	Bg ⁻ AAAAAAA	14.94	Ag ⁻ AAAAAAA	16.88	Ag ⁻ AAAAAAA	11.51
2	Ag ⁻ AAAAAAA	12.71	Ag ⁻ AAAABAA	11.93	Ag ⁻ AAAABAA	8.47
3	Bg ⁻ AAAABAA	10.62	Ag ⁻ BAAAAAA	6.53	Bg ⁻ AAAAAAA	5.49
4	Ag ⁻ AAAABAA	8.95	Bg ⁻ AAAAAAA	5.13	Bg ⁻ AAAABAA	4.47
5	Ag ⁻ CAAAAAA	4.92	Ag ⁻ BAAABAA	4.85	Ag ⁻ ABAAAAB	2.89
6	Ag ⁻ CAAABAA	3.50	Bg ⁻ AAAABAA	3.94	Ag ⁻ ABAAAAA	2.87
7	Ag ⁻ BAAAAAA	1.74	Bg ⁻ BAAAAAA	3.34	Ag ⁻ ABBAAAA	2.60
8	Ag ⁻ ABAAAAA	1.61	Ag ⁻ ABBAAAA	3.03	Ag ⁻ ABAABAB	2.51
9	Ag ⁻ ABAABAA	1.28	Bg ⁻ BAAABAA	2.55	Bg ⁻ ABAAAAA	2.39
10	Ag ⁻ BAAABAA	1.19	Ag ⁻ AABAAAA	2.04	Ag ⁻ AAAAAAC	2.13
		61.46		60.22		45.33

Table 2. Top 10 conformational states of each serotype (residue 4 to 11)^a in OS water solution.

#	serotype A (%)		serotype B (%)		serotype C (%)	
1	g ⁻ AAAAAAA	27.04	g ⁻ AAAAAAA	27.81	g ⁻ AAAAAAA	31.91
2	g ⁻ AAAABAA	18.90	g ⁻ AAAABAA	19.34	g ⁻ AAAABAA	22.57
3	g ⁺ AAAAAAA	16.84	g ⁺ AAAAAAA	8.54	g ⁺ AAAAAAA	9.72
4	g ⁺ AAAABAA	11.82	g ⁺ AAAABAA	5.90	g ⁺ AAAABAA	7.06
5	g ⁻ ABAAAAA	1.70	g ⁻ AAAAABA	2.43	g ⁺ AAAAAAB	3.51
6	g ⁻ AAABAAA	1.65	g ⁺ BAAAAAA	2.11	g ⁺ AAAABAB	2.38
7	g ⁺ AABAAAA	1.37	g ⁺ AAAAAAB	2.06	g ⁻ AAAAAAB	2.08
8	g ⁺ AAABAAA	1.36	g ⁻ AAAAAAB	1.96	g ⁻ ABAAAAA	1.54
9	g ⁻ ABBAAAA	1.25	g ⁻ BAAAAAA	1.64	g ⁻ AAAABAB	1.51
10	g ⁺ AACAAAA	1.11	g ⁻ AAAABBA	1.58	g ⁻ AAABAAA	1.24
		83.04		73.37		83.52

^aSince residue 4 is modified to an *O*-methyl glucoside in OS systems, only the torsion angle ω is considered for this residue.

Figure Captions

Figure 1. Structures of oligosaccharides for *M. catarrhalis* serotypes A, B and C. Lipid A molecular structure is depicted in Figure S1.

Figure 2. Molecular graphics snapshots of *M. catarrhalis* serotype B lipooligosaccharides (LOS) bilayer and oligosaccharides (OS) solution systems.

Figure 3. Pair-wise RMSD distribution calculated for serotypes A, B and C in (A) LOS bilayer simulations and (B) OS solution simulations. To eliminate the size difference, RMSDs for the common segment from glycan residue 4 to 11 were calculated. (C) The most populated structures for serotypes A, B, and C from 1- μ s LOS bilayer simulations. Residue 4 was modified to an *O*-methyl glucoside to perform OS simulations.

Figure 4. Distribution of the ω torsion angle of residue 4 for LOS (black) and OS (red) systems. The torsion angle ω is defined as O6-C6-C5-O5 where g^+ denotes gauche-trans, g^- denotes gauche-gauche, and t denotes trans-gauche. A β -(1 $\rightarrow$ 6)-linkage, having glycosidic torsion angle definitions H1'-C1'-O6-C6 (ϕ) and C1'-O6-C6-C5 (ψ), connects residues number 6 and 4.

Figure 5. Two-dimensional distribution of glycosidic torsion angles ϕ and ψ for β -(1 $\rightarrow$ 4)-linkage between residues 10 and 4 in (A) LOS and (B) OS. The glycosidic torsion angle definitions are H1'-C1'-Ox-Cx (ϕ) and C1'-Ox-Cx-Hx (ψ). Density 0 is plotted as white, 0.1 as blue, 0.3 as green, 0.7 as yellow, and 1 as red.

Figure 6. Comparison of each cluster population for serotypes A, B, and C (ω of residue 4 and (ϕ , ψ) of residues 5 to 11) in LOS bilayer simulations and OS water solution simulations.

Figure 7. (A) Calculated $^3J_{C,H}(\psi_H)$ with deviations from experiment within 0.25 Hz, 0.25 - 0.75 Hz, and greater than 0.75 Hz, respectively, are shown in red, green and cyan regions; (B) the corresponding torsion angle (ϕ , ψ) distribution for the β -(1 $\rightarrow$ 4)-linkage between residues 10 and 4 of OS structures using 12 replicas for the simulations. A representative structure for each region and the population of structures in the corresponding region are also indicated.

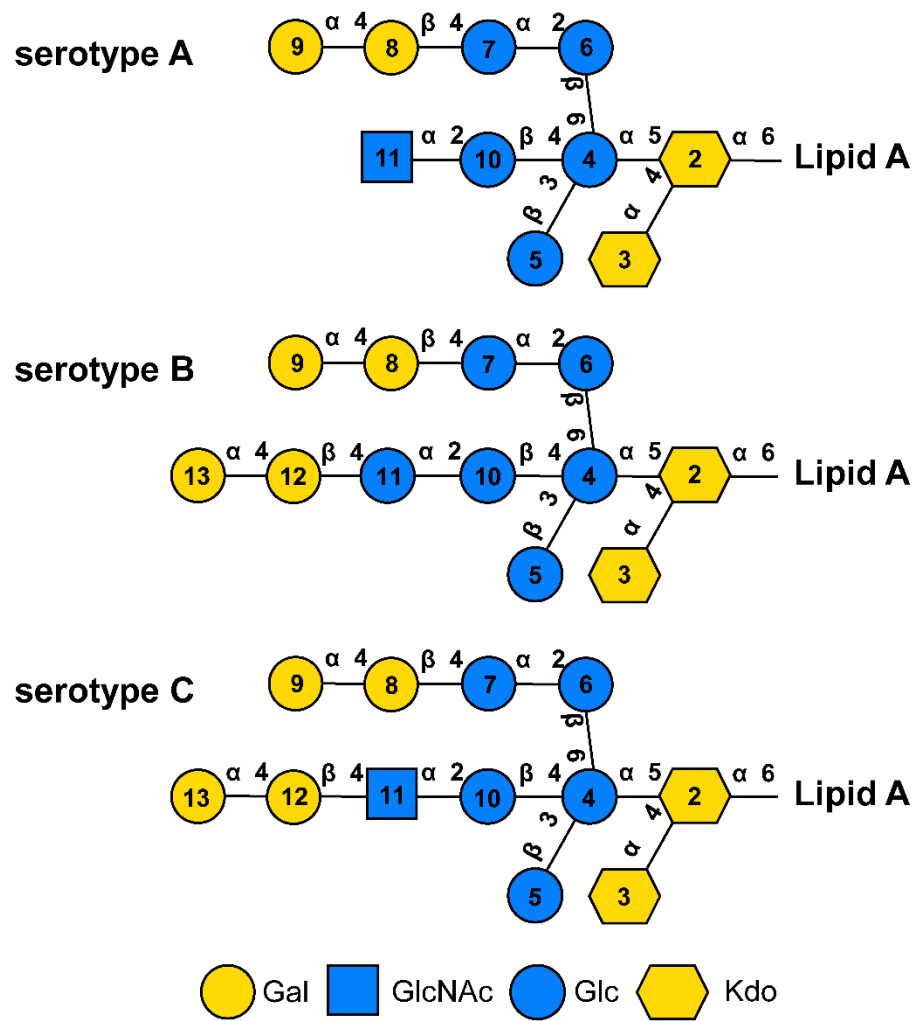


Figure 1

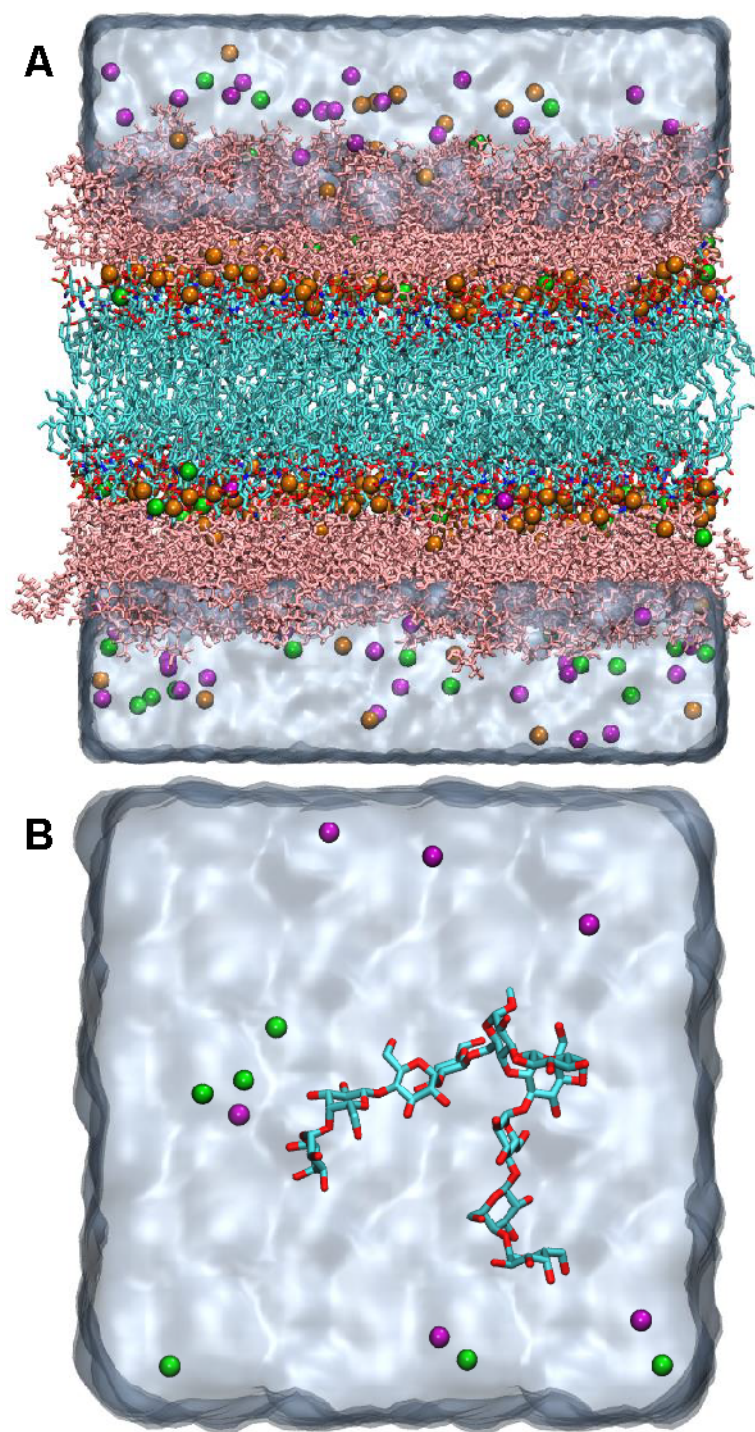


Figure 2

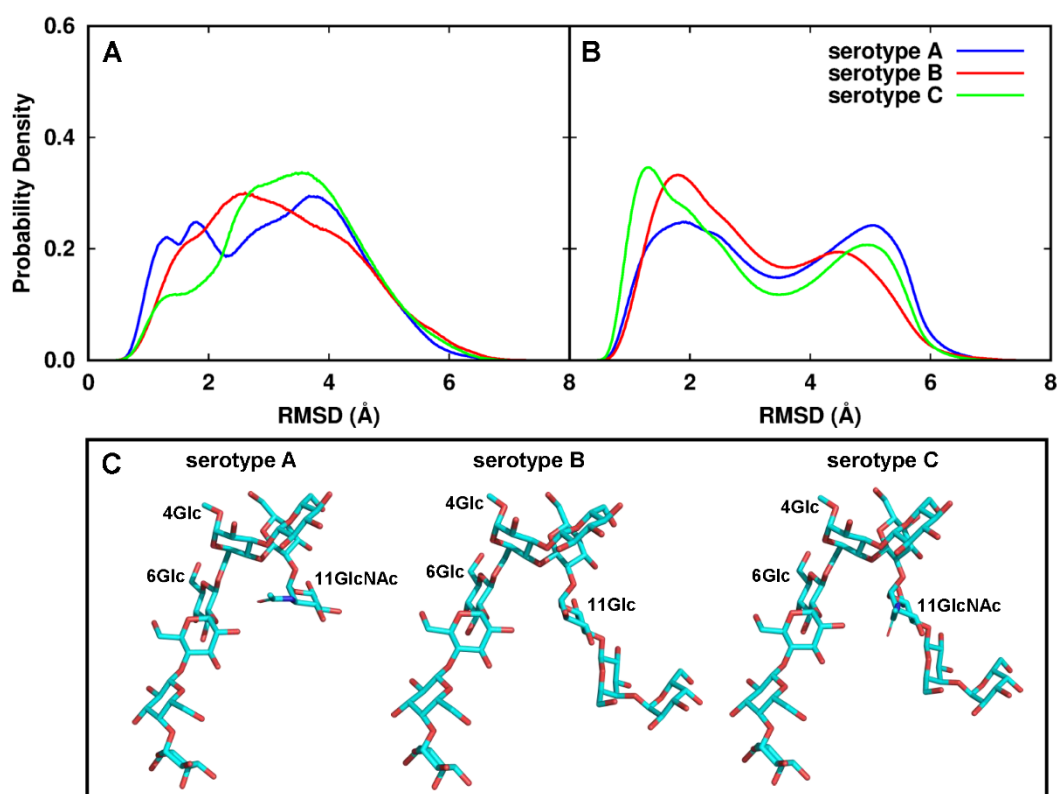


Figure 3

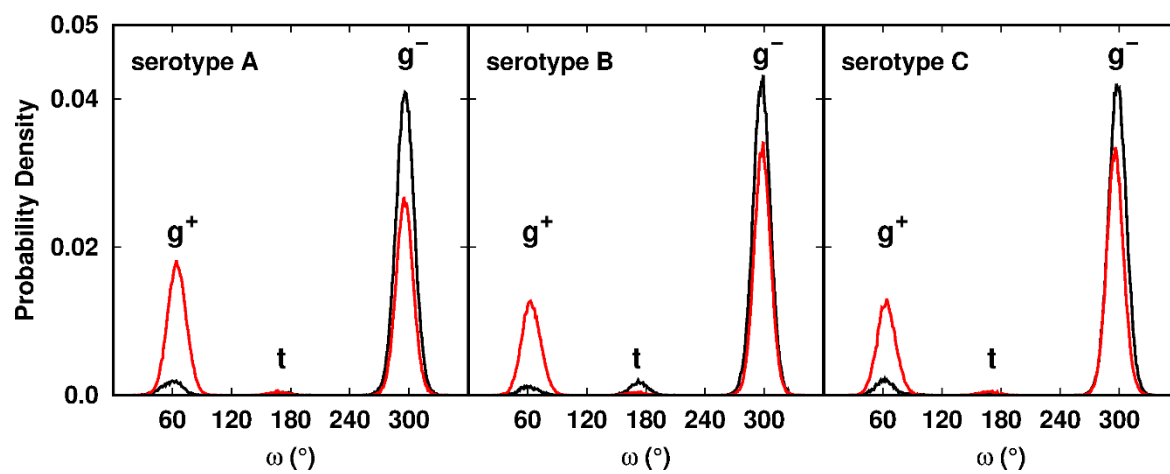


Figure 4

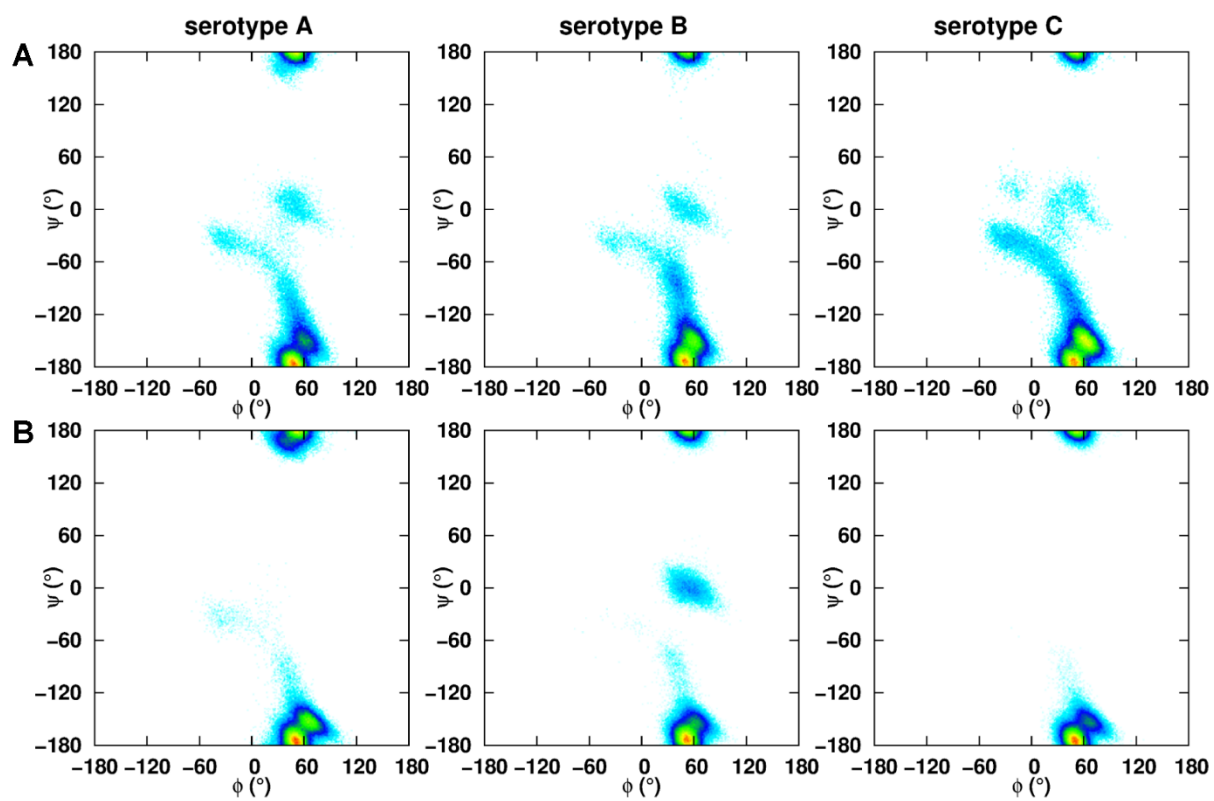


Figure 5

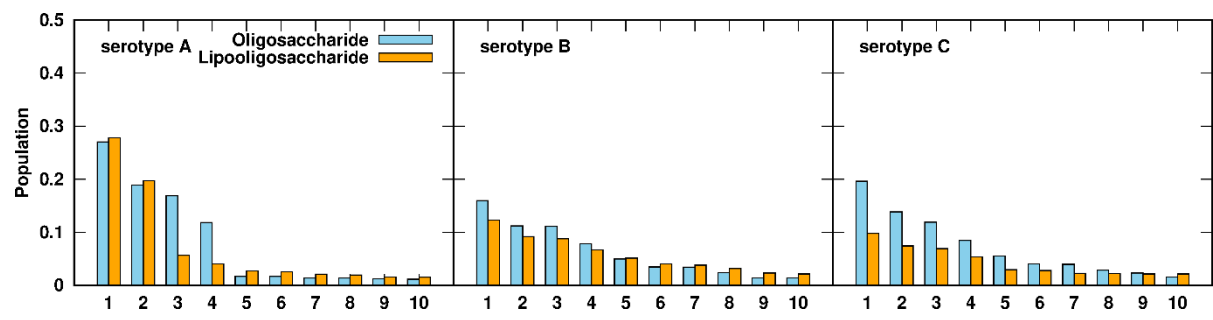


Figure 6

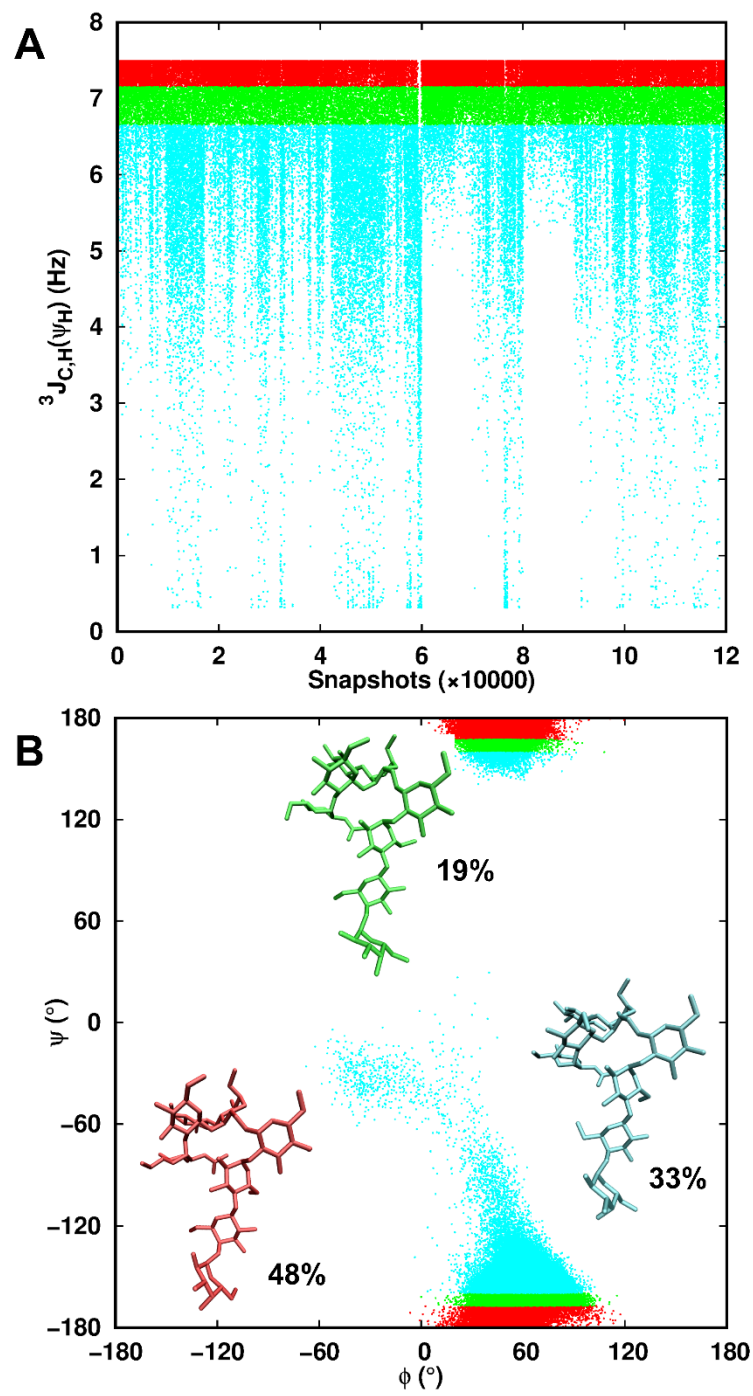


Figure 7