THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Conformational Diversity of O-Antigen Polysaccharides of the Gram-Negative Bacterium *Shigella flexneri* Serotype Y

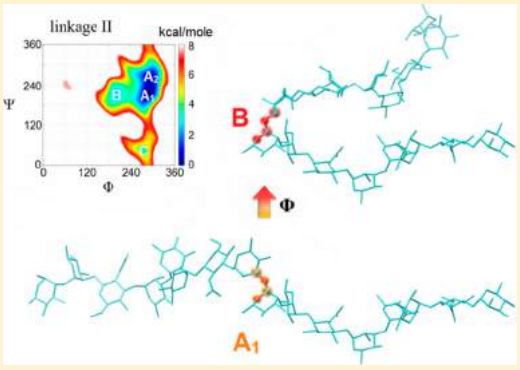
Yu Kang,[†] Stefanie Barbirz,[‡] Reinhard Lipowsky,[†] and Mark Santer*[,†]

[†]Max Planck Institute of Colloids and Interfaces, 14424 Potsdam, Germany

[‡]Physikalische Biochemie, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Supporting Information

ABSTRACT: O-Antigen polysaccharides constitute the outer protective layer of most Gram-negative bacteria, important for the bacterium's survival and adaption within its host. Although important for many functions, the three-dimensional structure of the dense polysaccharide coat remains to be elucidated. In this study, we present a systematic numerical investigation of O-antigen polysaccharide chains of *Shigella flexneri* serotype Y composed of one up to four tetrasaccharide repeat units. To bridge the gap between atomistic and coarse-grained levels of description, we employ a genuine multiscale modeling approach. It reveals that even for a few repeat units polymer-like flexibility emerges, which is furthermore complemented by extreme, hairpin-like conformations. These can facilitate the formation of metastable compact states, but this conclusion depends sensitively on the force field used to model the carbohydrates. Thus, our computational analysis represents an essential prerequisite for developing reliable coarse-grained models that may help visualizing changes in O-antigen coat morphology upon variations in chain length distribution or chemical composition of the polysaccharide characterizing a certain serotype.



INTRODUCTION

Lipopolysaccharide (LPS) is the outer membrane building block of Gram-negative bacteria.[1] A single LPS molecule consists of a fatty acid tail (Lipid A), an oligosaccharide core region, and a polysaccharide built from repeating oligosaccharide units (repeat units, RUs), protruding into extracellular space. The resulting dense layer of polysaccharides forms a protective coat against the host's innate immune system or against bacteriophage attack. The oligosaccharide repeat units are highly variable in composition even in the same bacterial species,[2−4] whereas the number of RUs and hence the chain length distributions of the polysaccharide are tightly regulated.[5] The polysaccharide can be targeted by antibodies and therefore is referred to as O-antigen. In human pathogens like *E. coli*, *Salmonella*, or *Shigella*, O-antigen is an important carbohydrate vaccine target to fight enterohemorrhagic strains that cause severe disease in humans.[6] The O-antigen also fine-tunes the interaction with host cells, in general establishing a balance between a sufficient level of self-protection and the bacterium's ability to colonize or invade tissue.[7,8]

Despite its significance, relatively little is known about the three-dimensional structure of the O-antigen chains in the LPS coat. Direct experimental visualization, as has been accomplished with the O-antigen coat of single *E. coli* bacteria *in vitro* by AFM imaging,[9] is in general very difficult. Although a rugged, heterogeneous morphology is suggested, it remains unclear how this morphology would look like under physiological conditions. Results from experiments on the adhesion of *E. coli* to quartz surfaces allow for an indirect interpretation of long O-antigen chains shielding the charged core regions,[10] but a concise picture of O-antigen chain organization in LPS coats is lacking, in particular with respect to variations in morphology resulting from carbohydrate compositions across different serotypes.

With the all-atomistic modeling of (bilayer-) membrane patches consisting of short LPS molecules containing the lipid A and core region, it has been recognized that computer simulations can provide useful information of the inner structure of LPS layers complementary to and corroborating experimental results.[11−13] In these computational studies, situations are addressed where only the lipid A and the inner core are kept.[14,15] By contrast, to take into account the multimodal chain length distributions naturally occurring in O-antigen, a fully atomistic description is not feasible any more. For instance, *Shigella flexneri* (*S. flexneri*) serotype 2a possesses a bimodal distribution of so-called very short (11−17 RUs) and very long (∼90 RUs) chains; it has been suggested that this composition significantly enhances the virulence of that strain as compared to a variety of mutations.[16] In *S. flexneri* serotype 5a it has been found that the presence of a glucose unit as a side chain leads to expression of significantly reduced O-antigen chain lengths, whereas the protection against penetration of

Received: November 13, 2013
Revised: February 4, 2014
Published: February 21, 2014

dx.doi.org/10.1021/jp4111713 | *J. Phys. Chem. B* 2014, 118, 2523−2534

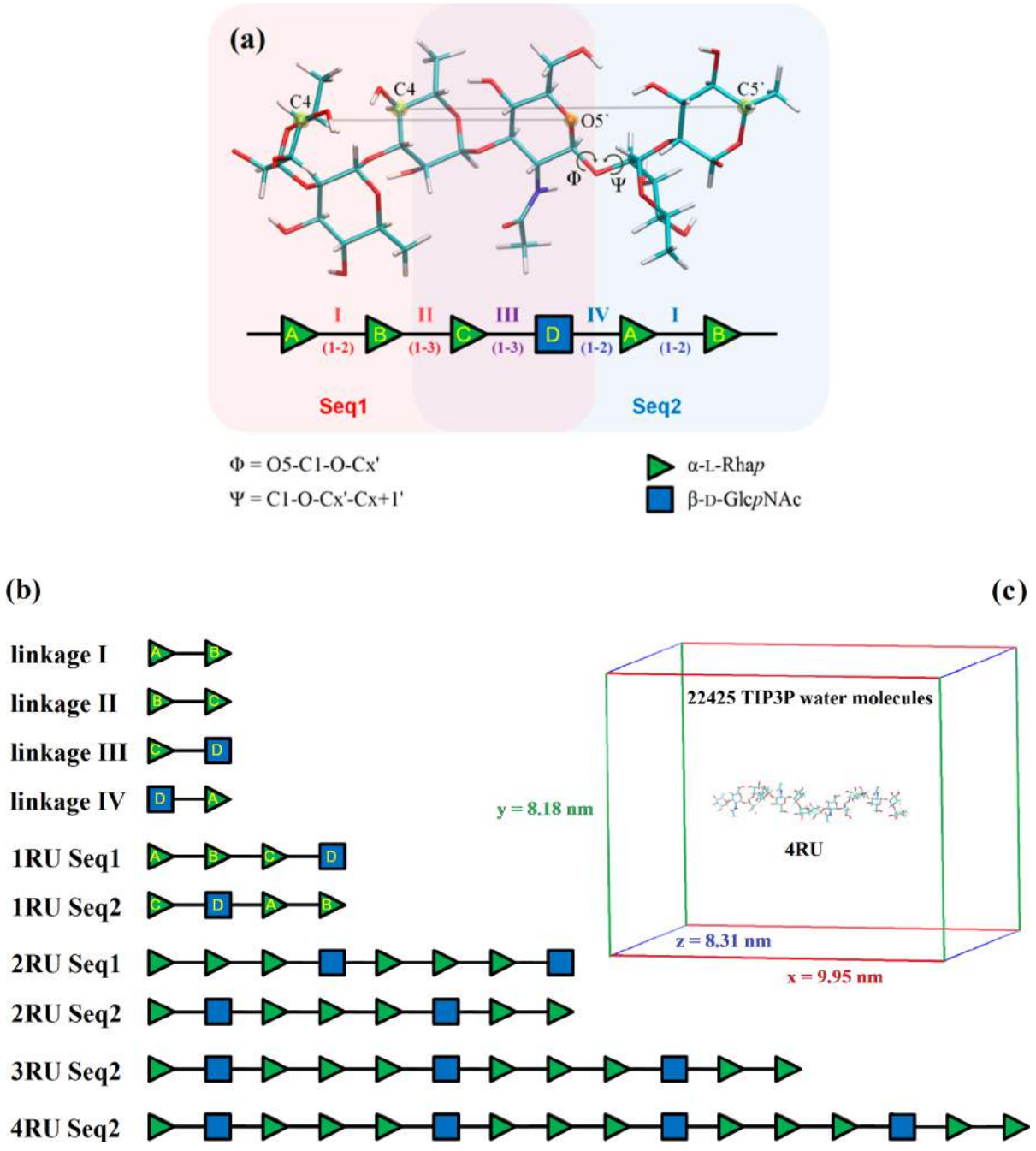


Figure 1. (a) Hexamer fragment of the *S. flexneri* serotype Y O-antigen comprising two different repeat units denoted as Seq1 and Seq2. Seq1 consists of a −2)-*α*-L-Rha*p*-(1−2)-*α*-L-Rha*p*-(1−3)-*α*-L-Rha*p*-(1−3)-*β*-D-Glc*p*NAc-(1− tetrasaccharide denoted as ABCD from left to right (red background). The other tetrasaccharide unit investigated in this work is the tetrasaccharide −3)-*α*-L-Rha*p*-(1−3)-*β*-D-Glc*p*NAc-(1−2)-*α*-L-Rha*p*-(1−2)-*α*-L-Rha*p*-(1−, denoted as CDAB (blue background). The structure is also represented below with the CFG nomenclature (http://www.functionalglycomics.org) where Rha*p* is represented as a green triangle and Glc*p*NAc as a blue square. The four different glycosidic linkages occurring are denoted as linkage I: -*α*-L-Rha*p*-(1−2)-*α*-L-Rha*p*- (AB); linkage II: -*α*-L-Rha*p*-(1−3)-*α*-L-Rha*p*- (BC); linkage III: -*α*-L-Rha*p*-(1−3)-*β*-D-Glc*p*NAc- (CD); and linkage IV: -*β*-D-Glc*p*NAc-(1−2)-*α*-L-Rha*p*- (DA). The glycosidic torsion angles Φ and Ψ are indicated along with linkage IV, but all Φ and Ψ angles are defined according to the IUPAC convention for X-ray crystal structures as O5−C1−O−Cx′ (Φ) and C1−O−Cx′−Cx+1′ (Ψ). (b) The CFG representation for all of the O-antigen fragments we used in this work for simulation, from di-, tetra- (1RU), octa (2RU), dodeca (3RU), and hexadecamer (4RU) O-antigen oligosaccharides. (c) Simulation box with a snapshot of the 4RU chain, fully hydrated using 22425 water molecules, which are not displayed for clarity.

antibacterial agents is not affected.[17,18] The numerical setups required to capture such phenomena would require excessive simulation times to achieve adequate sampling, especially when explicit solvent models must be used.[19−23] Coarse-grained computational models are in principle suitable to reach sufficiently long time scales. Approaches where (O-antigen) polysaccharides are represented as a simple, semiflexible chains of beads (each representing one carbohydrate ring) have been pursued,[24,25] also incorporating the variation in chain length,[26] to mimic the penetration of antimicrobial peptides toward the core region (on a ms time scale), or the interaction with charged nanoparticles.[27] However, the dependence of O-

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

antigen coat morphology certainly requires more molecular detail. Taking the significant variation in chain length of some bacteria as the starting point one may ask under which conditions, at which length the transition to a simplified polymer-like description is adequate, and how much the atomistic description may be coarse-grained.

In the present work we shall address this question by a systematic study of the conformational diversity of serotype Y O-antigen fragments of *S. flexneri* starting from one tetrasaccharide RU and going up to four RUs. We find that, up to two RUs (octasaccharides), the conformations of the O-antigen reflect statistically independent disaccharide fragments. Assessing larger O-antigen chains composed of up to four RUs is facilitated by the use of implicit solvent models, which we can show to be an accurate approximation. In addition to the polymer-like behavior that emerges early, the special geometrical arrangement of glycosidic linkages also allows for extreme, hairpin-like conformations. Depending on the force field used, these are either rare events that may be discarded or facilitate the formation of metastable compact states that drastically change the conformational ensemble. In the next section, we shall first give an overview of the computational strategies used. In the further exposition, O-antigen chain fragments of increasing length are studied in a bottom-up fashion. We conclude with a discussion on the implication of our results on the development of reliable coarse-grained models.

COMPUTATIONAL STRATEGY AND METHODOLOGY

Serotype Y O-antigen of *S. flexneri* has the basic tetrasaccharide repeat unit −2)-α-L-Rha*p*-(1−2)-α-L-Rha*p*-(1−3)-α-L-Rha*p*-(1−3)-β-D-Glc*p*NAc-(1− (Figure 1a: Seq1, ABCD).[4] All other serotypes are defined by glucosyl and O-acetyl modifications of this tetrasaccharide.[28] Apart from this sequence,[29] we also consider a shifted repeat unit, −3)-α-L-Rha*p*-(1−3)-β-D-Glc*p*NAc-(1−2)-α-L-Rha*p*-(1−2)-α-L-Rha*p*-(1− (Seq2, CDAB in Figure 1a), important in O-antigen cleavage by tail spike proteins.[30] For simulations we use the GLYCAM06(g)[31] force field for the major part of this study, a considerable number of comparative simulations have been carried out with CHARMM36, a version specifically dedicated to carbohydrates.[32] Molecular Dynamics (MD) simulations were conducted with two simulation packages, Gromacs, v. 4.5.5[33,34] and Amber 11.[35] Both comply with the same explicit water model TIP3P.[36] All details on computational methods and simulations protocols can be found in the Supporting Information, Section A.

In order to systematically investigate the conformational preferences of longer oligosaccharides composed of a few RUs, we first start with MD simulations of all possible disaccharide fragments contained in Seq1 and Seq2 in explicit solvent, as shown in Figure 1b, and characterize their conformation by Ramachandran-like plots with glycosidic dihedral angles Φ and Ψ as reaction coordinates, as defined in Figure 1a. To sample the conformational space of these torsions more thoroughly, adaptively biased MD (ABMD)[37] was performed to obtain more extensive free energy landscapes, see Figure 2 for a survey of all four different linkages. Our use of the ABMD algorithm is explained in detail in the Supporting Information, Section B. The biased simulations allow the space of glycosidic angles to be decomposed into certain states (Figure 2), and each state contains one local minimum of the free energy landscape, also

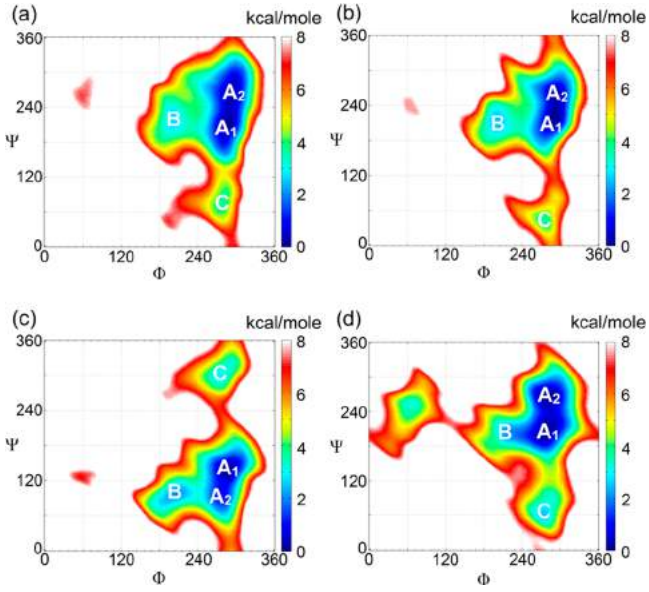


Figure 2. Free energy landscapes of the glycosidic dihedral angles Φ and Ψ for all disaccharide linkages of *S. flexneri* serotype Y O-antigens. (a) Linkage I: α-L-Rha*p*-(1−2)-α-L- Rha*p*−OH; (b) linkage II: α-L-Rha*p*-(1−3)-α-L-Rha*p*−OH; (c) linakge III: α-L-Rha*p*-(1−3)-β-D-Glc*p*NAc−OH; and (d) linkage IV: β-D-Glc*p*NAc-(1−2)-α-L-Rha*p*−OH. Free energies were determined with an ABMD calculation with explicit solvent as explained in the Supporting Information, Section B. Values up to 8 kcal/mol are color coded. Labels of the important (meta-) stable conformational states, denoted as A_1, A_2, B and C have been placed in their respective minima. For linkages I, II and IV with the α-L-Rhamnose unit at the reducing end, we define A_1, A_2, B, and C as follows. A_1: $120° < \Psi < 240°$ and $(\Psi - 240°) < 6(\Phi - 240°)$; A_2: $\Psi > 240°$ and $(\Psi - 240°) > (240° - \Phi)$; state B: $120° < \Phi < 240°$ and $(\Psi - 240°) < (240° - \Phi)$ and $(\Psi - 240°) > 6(\Phi - 240°)$; state C: $\Psi < 120°$. For Linkage III with the β-D unit at the reducing end the definitions are: state A_1: $\Psi > 120°$ and $(\Psi - 120°) > (240° - \Phi)$; state A_2: $\Psi < 120°$ and $(\Psi - 120°) < 6(\Phi - 240°)$; state B: $120° < \Phi < 240°$ and $(\Psi - 120°) < (240° - \Phi)$ and $(\Psi - 120°) > 6(\Phi - 240°)$; state C: $\Psi > 240°$.

labeled as A_1, A_2, B, and C. In the following, we shall classify conformational characteristics in terms of the occupation or population of these states.

The populations of the different states inferred from MD and from the ABMD technique are in very good agreement, as can be inferred from the first column of Table 1. We therefore use the ABMD free energy landscapes to construct a reduced model, by which longer oligosaccharide sequences can be represented as a linear backbone consisting of a sequence of single bonds as described in ref 23, see also the Supporting Information, Section C. In this description, bond lengths and angles remain fixed, the only remaining degrees of freedom are torsions around glycosidic bonds. Using the free energy landscapes of Figure 2 to parametrize glycosidic dihedral torsions, this model can conveniently be explored with Monte Carlo (MC) sampling. In investigating O-antigen chains with an increasing number of RUs, simulations with the reduced model will serve as a control in order to identify changes in conformation that arise from intramolecular interactions beyond the disaccharide level, i.e., from chain "self-avoidance".

To allow for sufficient sampling with the atomistic model for O-antigens longer than two repeat units (octasaccharide), we employ implicit solvent models,[38−40] the validity of which has been carefully checked, see also the Supporting Information,

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

Table 1. Distributions (%) of the Low Energy Conformational States A and B[a]

			tetramer		octamer (CDABC′D′A′B′)	
		dimer	ABCD	CDAB	1st RU	2nd RU
linkage I	A_1	61.4(0.8)/62.7/64.8	64.5(0.8)	63.4(0.9)	64.2(0.8)	60.8(0.8)
	A_2	37.4(0.8)/35.7/35.1	34.8(0.8)	34.2(0.8)	33.7(0.7)	37.5(0.7)
	B	1.2(0.2)/1.4/0.2	0.7(0.1)	2.4(0.3)	2.1(0.2)	1.7(0.2)
linkage II	A_1	57.6(0.8)/57.9/30.8	56.8(0.9)		48.0(1.1)	
	A_2	40.4(0.8)/40.0/68.9	41.3(0.9)		47.6(2.3)	
	B	2.0(0.2)/2.0/0.4	1.9(0.2)		4.4(1.8)	
linkage III	A_1	68.4(0.7)/68.1/36.8	67.9(0.7)	76.2(0.6)	65.1(0.8)	57.0(0.8)
	A_2	26.7(0.8)/26.8/60.6	27.0(0.8)	20.0(0.8)	28.4(0.9)	39.0(0.9)
	B	4.9(0.3)/4.1/2.7	5.2(0.3)	3.8(0.3)	6.5(0.4)	4.0(0.3)
linkage IV	A_1	55.1(0.9)/53.4/57.9		63.0(0.8)	57.9(0.8)	56.6(0.8)
	A_2	43.1(0.8)/43.7/42.0		34.5(0.8)	40.0(0.8)	41.0(0.8)
	B	1.8(0.2)/2.1/0.1		2.5(0.2)	2.1(0.2)	2.4(0.2)

[a]Standard deviations for populations inferred from plain trajectory data are given in parentheses. They have been obtained by bootstrap resampling with 2000 replications on a set of 4000 statistically independent data points along the MD trajectory. The populations for disaccharides calculated from ABMD free energy profiles are shown as second entry in column "dimers", and they were obtained from Boltzmann inversion of the free energy landscapes in Figure 2. The third entry in this column is the occupation obtained with the force field CHARMM36, using the same run length (200 ns) and the same filtering (4000 statistically independent data points).

Section C. The multiscale strategy as outlined above will now be used to systematically characterize di-, tetra- (1RU), octa (2RU), dodeca (3RU) and hexadecamer O-antigen oligosaccharides. We shall start with discussing disaccharide linkages.

RESULTS

Disaccharide Linkages and Dependence on the Force Field. For each disaccharide two major minima in free energy (states A_1, A_2) are found (Figure 2). In each case the angle Φ is centered around 283°, reproducing the expected exoanomeric effect.[41] The secondary minima along the Φ direction we refer to as nonexoanomeric states (labeled as B) and the anti-Ψ configuration (labeled C). In the following, we shall also use indices to indicate the linkage and write B_I, B_{II}, etc. The population in the space of glycosidic angles is in general agreement with results found in databases such as GlycoMapsDB;[42] in particular, recently obtained simulation results on three-RU oligosaccharides of serotype Y[28] with GLYCAM06(g) fully comply with the results reported here. A study specifically devoted to rhamnose linkages has previously been carried out using the PARM22/SU01 force field.[43] In the following, we shall compare the corresponding results with ours, including data from the CHARMM36 force field.[32] Table 1 contains the compilation of populations for the states of each linkage.

Linkage I. The α-L-Rha*p*-(1−2)-α-L-Rha*p*−OH disaccharide shows two similarly populated minima A_1 and A_2 at a ratio 61:37 (Figure 2a). In the PARM22/SU01 force field, state A_1 is more emphasized (by about 10%).[43] State B is populated by less than 2%. As a result of frequent exchange between A_1 and A_2, Ψ is centered at 229° and thus mainly in a syn-conformation, with H1 and H2′ on the same side of the plane orthogonal to the C1−H1 bond vector. State C corresponds to an anticonformation and is barely visited during a typical MD run at 310 K and 200 ns duration.

Linkage II. The α-L-Rha*p*-(1−3)-α-L-Rha*p* glycosidic linkage shows a similar distribution in Ψ as compared to the α-1−2 linkage I (Figure 2b). The states A_1 and A_2 with Ψ < 240° and Ψ > 240° are equally populated, but, compared to linkage I, the Ψ angle is restricted to a narrower region from 180° to 300°. State B is populated by ∼2% (inferred from MD trajectories/ABMD data). In a study using the PARM22/SU01 (CHARMM) force field, the distributions in Φ and between the two syn-Ψ-states A_1 and A_2, agree well with our results, but deviations with respect to state B are apparent, for which occupations of 4% in a disaccharide and 3−5% in a hexamer are found.[43] Using CHARMM36 results in merely 0.4% (Table 1).

Linkage III. The α-L-Rha*p*-(1−3)-β-D-Glc*p*NAc−OH disaccharide (Figure 2c) shows a pronounced asymmetry in the population of A_1 and A_2 (68:27). This difference is due to the 6N-acetyl fragment at the ring carbon C2 pointing parallel to the glycosidic bond (in the direction from C1 to C3′) rather than perpendicular to it. During a 200 ns MD production run, the C4 hydroxyl group of GlcNAc temporarily forms hydrogen bonds to O5 of α-L-Rha at the nonreducing end with occupancy of around 19% (Figure 3a). The occurrence of this HO4′-O5 hydrogen bond is, for geometrical reasons, correlated with the Ψ visiting state A_2 (Figure 3b).

Linkage IV. In the β-D-Glc*p*NAc-(1−2)-α-L-Rha*p*−OH disaccharide, the carbonyl oxygen of GlcNAc forms a relatively loose hydrogen bond with the hydroxyl group at C3 of α-L-Rha*p* (Figure 2d). The occupancy of this HO3′-O2N hydrogen bond is around 4% within MD production runs of 200 ns duration (Figure 3c). This hydrogen bond can be observed with Φ and Ψ in both states, A_1 and A_2, respectively (Figure 3d), the occupation of which is about equal and in agreement with CHARMM36.

It is useful to summarize and point out the variations in results across the different force fields considered. Discrepancies arising from the different parametrizations between the GLYCAM and CHARMM force fields have been reported before.[44] Compared to our reference calculations the B states with CHARMM36 are generally only weakly occupied; most of them except B_{III} by less than 1%. Furthermore, the ratio A_1:A_2 neither fully agrees with CHARMM36 nor with results obtained with PARMM22/SU1. Our analysis compiled for disaccharides in Table 1 suggests that we can indeed distinguish between the different parametrizations numerically; experimentally, at the level of small oligosaccharides this is much more difficult. To illustrate this point, we list in Table 2 the calculated $^3J_{COCH}$ coupling constants of glycosidic linkage II torsions for the force fields considered here, compared to

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

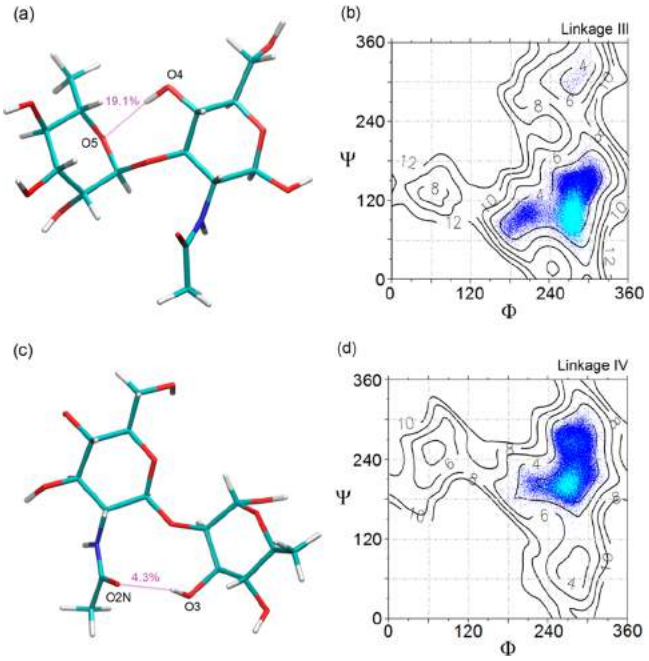


Figure 3. Hydrogen bridges as monitored in explicit solvent. (a) An intramolecular hydrogen bond observed in the disaccharide α-L-Rha*p*-(1−3)-β-D-Glc*p*NAc−OH (linkage III) with an occupancy of 19.1%. (b) The corresponding Φ−Ψ distribution. Dots in dark blue represent all points of the Φ−Ψ trajectory within the MD production run, light blue dots highlight those with the HO4′-O5 hydrogen bond formed. The contours are taken from the ABMD simulation of the disaccharide in Figure 2. (c and d) The analogous representation for the hydrogen bridge HO3′-O2N in the disaccharide β-D-Glc*p*NAc-(1−2)-α-L-Rha*p*−OH (linkage IV).

Table 2. $^3J_{COCH}$ Coupling Constants (Hz) of Linkage II in Aqueous Solution at 310 K[a]

torsion angle	Φ_H (H1−C1−O−C3′)	Ψ_H (C1−O−C3′−H3′)
a	3.7	5.0
b	3.7	4.7
c	3.7	5.2
d	4.1	5.1

[a] $^3J_{COCH}$ coupling constants of linkage II taken from (a) a disaccharide in this work using GLYCAM06(g) force field; (b) a disaccharide in ref 43using PARM22/SU01 force field; (c) an octasaccharide in this work using GLYCAM06(g) force field; (d) experiment.[52] The equations used here to calculate the $^3J_{COCH}$ coupling constants were taken from ref 53.

experimental values for the two couplings characterizing the conformation of Linkage II. It is evident that when comparing *J* couplings as inferred from NMR experiments, no clear distinction can be made between the three parametrizations at the disaccharide level.

Tetrasaccharides Inherit Glycosidic Bond Conformations from Their Disaccharide Fragments. Two tetrasaccharides with sequence ABCD and CDAB respectively, were used for simulations of 800 ns duration each (explicit solvent). In general, the distribution of the four types of linkages in both tetramers agrees well with the respective distributions in the disaccharides (Figure 4 and Tables 1 and 3). There is, however, evidence for some residual correlation across linkages, especially notable for the sequence CDAB: the HO4′-O5 hydrogen bond between residues C and D (Figure 5c) within

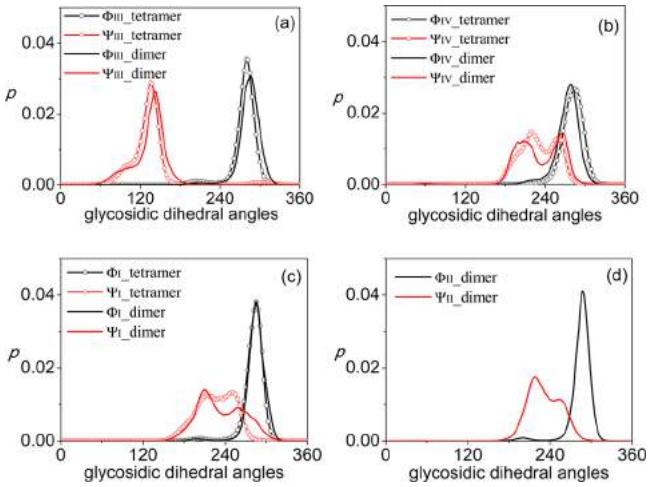


Figure 4. Probability distributions projected onto the angles Φ and Ψ of (a) linkage III, (b) linkage IV, (c) linkage I, and (d) linkage II in disaccharides compared with the corresponding linkages in the tetrasaccharide CDAB. Simulations with explicit solvent.

linkage III is still the one with highest occupancy among all of the intramolecular hydrogen bonds formed in the tetrasaccharide, but compared to the corresponding disaccharide (or in the tetramer ABCD), the occupancy drops from ∼19.1% (or ∼20.4%) to only ∼10.6%. This correlates with a shift in the Ψ distribution from state A_2 to A_1 as shown in Table 1, with A_1 increasing from 68% up to 76% for linkage III (the average Ψ increases). The hydrogen bond found for linkage IV drops in occupancy from 4.6% down to 1%, with a corresponding shift in Ψ (population of A_1 increasing from 55% to 63%, average Ψ decreases). In addition, a new hydrogen bond between HO6-group in the GlcNAc unit (D) and the O3 atom in the Rha unit (B) as next-nearest-neighbor residues emerges, compare Figure 5c, and now ranks second in occupancy (around 4.8%). The occurrence of this new hydrogen bond correlates with the conformational shift of linkage IV toward A_1 (cf. Figure 5b), which simply reduces the distance between the hydrogen HO6 in monomer D and O3 in B.

Irrespective of the fine details described above, the gross conformational behavior can still be well described by assuming independent glycosidic linkages with preferences inherited from disaccharides. To characterize this, we consider end-to-end distances r_{ete} between distal atoms. For the two sequences shown in Figure 5a, r_{ete} for the tetrasaccharide ABCD was defined as the distance from C4 at the nonreducing end to O5′ at the reducing end, and for CDAB from C4 to C5′ (cf. Figure 1). End-to-end distances in both tetrasaccharides comprise one major peak. For Seq1, it is distributed widely around 1.2 nm. By contrast, the peak for Seq2 is much narrower and centered around 1.4 nm. The apparent difference between the two sequences is solely due to the special geometrical arrangement of the participating glycosidic linkages. With the reduced backbone model (Supporting Information, Section D), we obtain the corresponding distributions for the two sequences as dashed lines in Figure 5a. Both maxima in the distribution roughly correspond to all glycosidic torsions assuming conformational states of type A. The small deviation in shape of the curves is due to the fact that only a single initial snapshot of the saccharide (taken from an MD run) was used in the MC simulation.

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

Table 3. Average Torsion Angles (°) of Each Glycosidic Linkage in All Serotype Y O-Antigenoligosaccharides up to Octamers[a]

		linkage I		linkage II		linkage III		linkage IV	
		Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ
dimer		285(14)	229(31)	286(16)	233(27)	278(21)	128(24)	278(14)	239(31)
tetramer (ABCD)	1	285(12)	228(26)	283(15)	234(25)	276(23)	128(29)		
tetramer (CDAB)	1	283(17)	226(27)			279(18)	132(25)	282(18)	229(26)
octamer (CDABC′D′AB′)	1	284(16)	223(33)	284(20)	237(24)	278(24)	126(28)	276(15)	230(34)
	2	284(15)	228(29)			277(21)	121(26)	276(16)	232(30)

[a]Standard deviations computed with statistically independent data points are given in parentheses. Simulations with GLYCAM06(g) and explicit solvent.

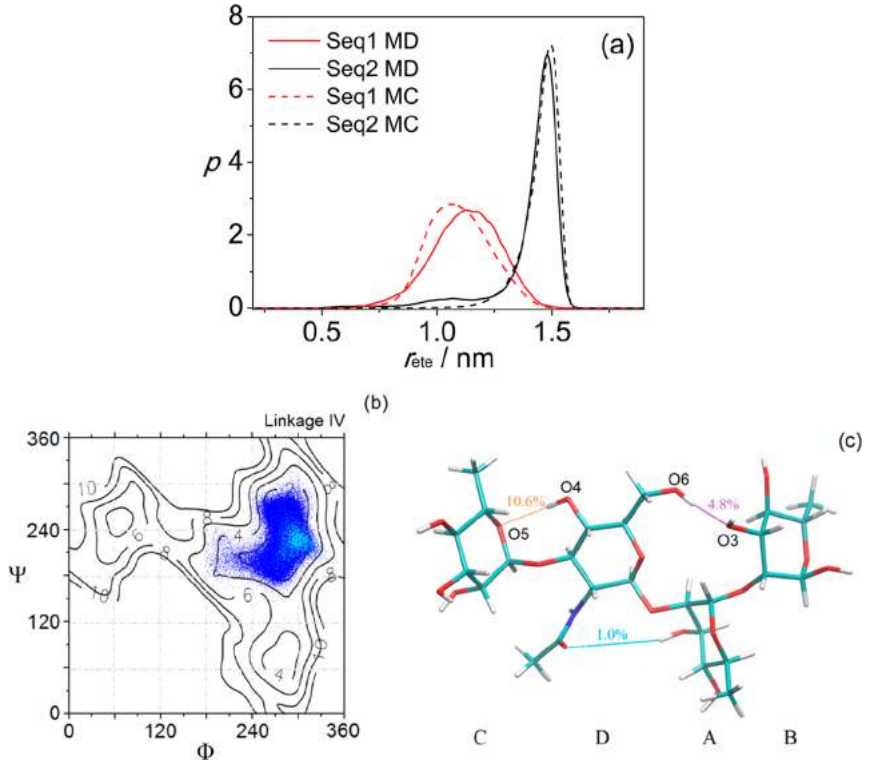


Figure 5. (a) Probability distribution $p(r_{ete})$ for the tetrasaccharide units ABCD (Seq 1, red) and CDAB (Seq 2, black) inferred from MD data with explicit solvent (solid lines) and MC simulations with the reduced model (dashed lines, 10^6 MC steps). (b) The Φ−Ψ distribution of linkage IV in the tetrasaccharide CDAB. The blue dots represent the trajectory of Φ and Ψ in the MD production runs. Light blue dots signify the subset when the HO6-O3 hydrogen bond is formed. The contour lines are taken from Figure 2. (c) CPK representation of intramolecular hydrogen bridges observed in the tetrasaccharide CDAB with occupancies annotated.

Conformation of Octasaccharides and Hairpin-Like Configurations. Four separate MD runs with explicit solvent of 200 ns duration each were performed to sample the conformation of the Seq1 octasaccharide (ABCDA′B′C′D′) and the Seq2 octasaccharide (CDABC′D′A′B′) in aqueous solution. Again, to a very good approximation, glycosidic linkages in the octasaccharide units preserve the characteristics of the corresponding disaccharides. Moreover, the differences in occupancy found for the HO4′-O5 hydrogen bond of tetramer CDAB between residues C and D (linkage III) disappear. This finding indicates that the slight changes of the conformational characteristics in the tetramer most certainly average out as longer saccharide chains with increasing number of repeat units are considered. This is also documented by the average glycosidic angles measured across all fragments a certain linkage is part of (Table 3). They all agree within error bars.

In Seq1 octasaccharides (ABCDA′B′C′D′) end-to-end distances are shorter on average than in Seq2 octasaccharides (CDABC′D′A′B′), still indicating the influence of a specific geometric arrangement of glycosidic bonds. Accordingly, preferences of glycosidic torsions do not depend on their position within the fragment (Figure 6a,b). In contrast to the tetrasaccharides, however, the octasaccharides bear asymmetric distributions of lengths (Figure 6c). The distributions calculated from the MC results show again good agreement with the MD data, and inspection of all trajectories reveals that the dynamics of all linkages considered within A-states only is sufficient to generate these distributions.

Interestingly, we can now also observe rather extreme, hairpin-like conformations (Figure 7a). In Seq2 octasaccharides they are most pronounced due to linkage II connecting the two RUs. The linkage II is then found in state B (B_{II}) with Φ < 240° (cf. Figure 2, and also Figure 7b). Although state B_{II} is not significant for the overall end-to-end distribution for the octasaccharide in solution, it is special in the sense that when it is visited, the conformational change in the O-antigen chain is most pronounced in terms of r_{ete}, see Figure 8. Starting from a

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

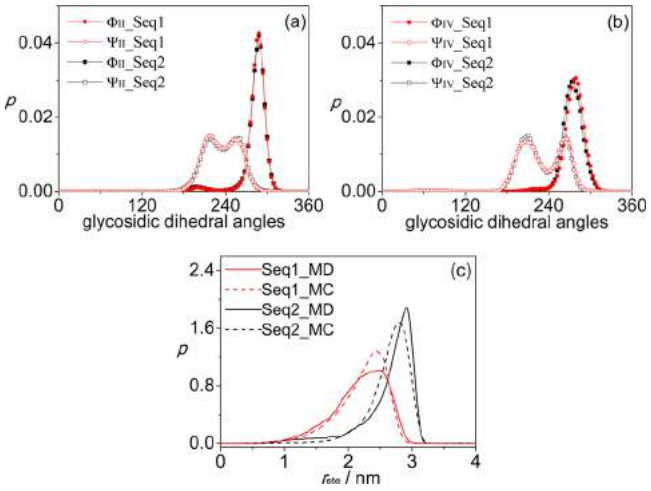


Figure 6. Distribution of the torsion angles in (a) linkage II and (b) linkage IV as part of different tetrasaccharides (Seq1 or Seq2), as well as (c) the distribution of the end-to-end distance $p(r_{ete})$ for the two octasaccharides with sequences ABCDA′B′C′D′ (red) and CDABC′-D′A′B′ (black). The MD data are taken from explicit solvent simulations. The dashed lines in panel c represent MC results of the reduce backbone model.

dodecamer where each linkage has been put in its global minimum (one of the A states), we put each linkage separately into the minimum of its corresponding B state (B_I–B_{IV} corresponding to a–d), with B_{II} and B_{III} facilitating rather sharp turns within the chain. Ultimately, B_{II} is more significant for hairpin-like conformations; these cannot be formed with B_{III} due to the steric hindrance of the NAc group. These are geometrical/topological aspects of the Serotype Y O-antigen polysaccharide quite independent of the force field.

With a hairpin-like conformation larger portions of an oligo- or polysaccharide are brought into close proximity, and one might expect nonbonded intramolecular interactions to take effect. Although the increased population of B_{II} in the octasaccharide case (Table 1) might indicate this, an ABMD simulation with r_{ete} as the reaction coordinate suggests only a rather limited stability of conformations with small (<1.5 nm) r_{ete} in the Seq2 octasaccharide with B_{II} as the central linkage. The corresponding free energy in Figure 7c at best exhibits a plateau for this region at roughly 3 kcal/mol above the global minimum. However, in the gas phase the hairpin state is indeed found stabilized and B_{II} is populated by more than 90% (Table 4). This is true whether the GLYCAM or the CHARMM36 force field is employed. In vacuum, the hairpin conformation can receive enhanced stabilization by intramolecular hydrogen bonds between the first and second RU (mainly originating from the GlcNAc units), whereas in solution they are effectively screened.[45]

Note that the biased simulations have been carried out with implicit solvent, as with explicit solvation the biased MD has been found to take excessive runtimes to converge. Nevertheless, the distribution $p(r_{ete})$ calculated from the free energy profile excellently fits the one obtained from MD production runs (shown is the result combined from 4 separate runs of 200 ns each), indicating a solid consistency with the implicit solvent representation. We have mainly used the generalized-Born HCT model,[38,39] with the dielectric constant set to 80. Comparative simulations with the OBC scheme[40] led to almost identical results. A detailed account of the use of implicit solvent models in this work is given in the Supporting Information, Section C.

Validity of Extrapolating the Behavior of O-Antigens to a Polymer Model Depends on the Force Field. Although the hairpin-like conformation for Seq2 octasacchar-

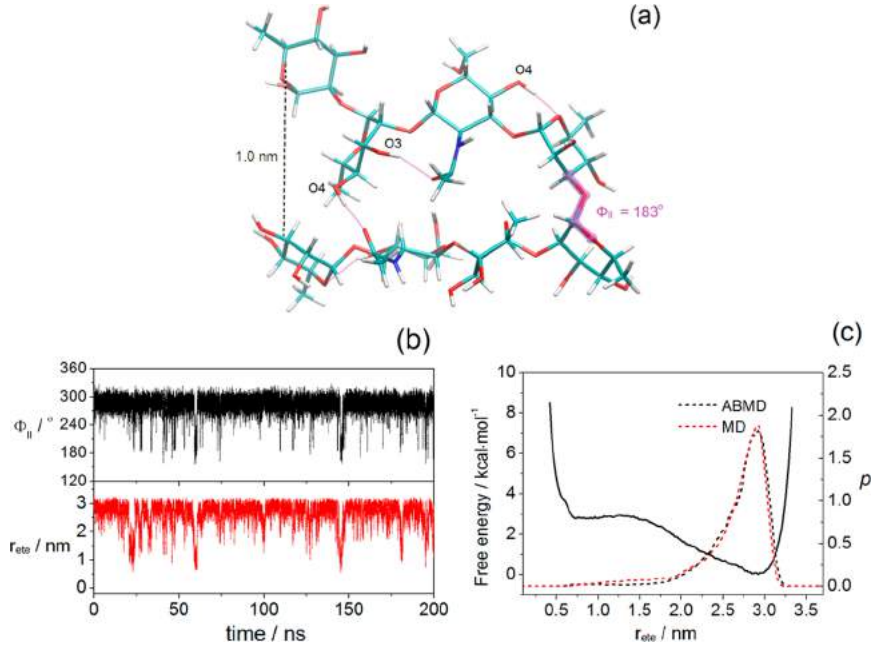


Figure 7. (a) Snapshot of a hairpin conformation of the octasaccharide fragment CDABC′D′A′B′ taken from a MD trajectory with explicit solvent. The torsion angle Φ of linkage II is highlighted (here the state B_{II} is occupied). (b) The angle Φ and the end-to-end distance r_{ete} of the octasaccharide as a function of time (explicit solvent). (c) Free energy profile as a function of the r_{ete} obtained by an ABMD simulation with implicit solvent (black solid line) together with the distributions $p(r_{ete})$ calculated from that profile (black dashed line) and the MD production run with explicit solvent (red).

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523–2534

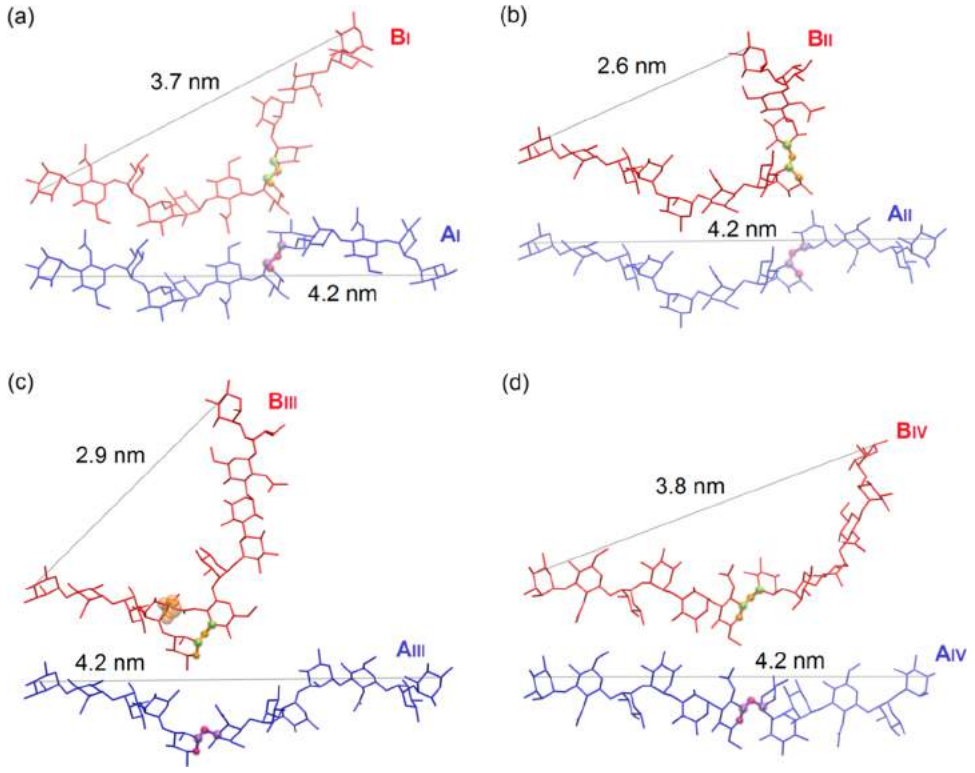


Figure 8. Change of the global structure as well as the end-to-end distance of dodecasaccharides with Seq 2 (CDAB) as the RU when the glycosidic angle Φ is turned by 90° (roughly from 300° to 210°) in order to emulate a transition to a B state of linkages I–IV. Blue: all glycosidic angles are in the minima of A states (GLYCAM06). Red: the A state of (a) linkage I, (b) linakge II, (c) linkage III, or (d) linkage IV defined by the highlighted sequence of atoms (CPK representation) has been moved from the A to the B state. The (1–2) linkages I and IV do not affect the end-to-end distance as much as the (1–3) linkages (linkages II and III) do. The transition to B_{II} or B_{III} (with other torsions fixed) can in principle lead to similar decrease in end-to-end distance; however, in panel c the formation of a more compact conformation is hindered by steric interaction with the bulky acetylamino group (corresponding atoms in VdW representation).

Table 4. Distributions (%) of the Low Energy Conformational States A and B in an Octasaccharide CDABC′D′A′B′ Obtained with Different Solvent Conditions[a]

		implicit solvent[b]		vacuum	
linkage	state	1st RU	2nd RU	1st RU	2nd RU
I	A_1	66.5(0.8)	60.4(0.8)	88.9(0.5)	3.0(0.3)
	A_2	32.5(0.8)	37.8(0.8)	1.1(0.2)	96.8(0.3)
	B	1.0(0.2)	1.8(0.2)	10.0(0.4)	0.2(0.1)
II	A_1	55.6(0.8)		1.6(0.2)	
	A_2	40.7(0.8)		0.1(0.1)	
	B	3.7(0.3)		98.3(0.2)	
III	A_1	75.1(0.7)	68.7(0.7)	3.2(0.3)	61.7(0.8)
	A_2	22.2(0.7)	26.0(0.8)	96.6(0.3)	28.3(0.9)
	B	2.7(0.2)	5.3(0.3)	0.2(0.1)	10.0(0.5)
IV	A_1	59.4(0.8)	59.7(0.8)	89.9(0.5)	80.4(0.6)
	A_2	40.4(0.8)	39.2(0.7)	8.5(0.5)	19.6(0.6)
	B	0.2(0.1)	1.1(0.2)	1.6(0.2)	0.0(0.0)

[a]Standard deviations given in parentheses have been calculated with bootstrap resampling as in Table 1. [b]An improper torsion restraint was applied on each carbohydrate ring (C5–O5–C4-C2 for rhamnose and C2–C3–C1-C5 for N-acetyl-glucosamine) to prevent unexpected ring flips, see the discussion in Supporting Information, Section C.

ides suggests the importance of intramolecular interactions, there was little impact on the distributions of the end-to-end distance. Consistently, the reduced backbone model relying on mutually independent glycosidic linkages and the implicit/explicit MD results show excellent agreement. In Figure 9 we now show the distribution of r_{ete} for three (a) and four (b) repeat units (Seq2), calculated from MD with implicit solvent (GLYCAM06(g)) using the HCT model. The results are compared to the backbone model, which again exhibits only one major peak when all glycosidic torsions are close to the minima of the A states and a long tail reflecting the combined dynamics within these states; conformations with B_{II} states involved are rare, as expected. The MD results deviate from this behavior, a slight enhancement in $p(r_{ete})$ for values of r_{ete} smaller than 2 nm and a sudden drop at 0.75 nm are observed (Figure 9a).

The cutoff in r_{ete} is due to steric repulsion of reducing and the nonreducing end, for conformations that smoothly bend back on itself, compare snapshot **1** in Figure 9a. In the region between 1 and 2 nm showing the enhancement, many congested structures such as the two examples displayed in 2 are found, comprising the presence of at least one B_{II} state. It is instructive to represent $p(r_{ete})$ with all conformations excluded containing one or more B_{II} states (dashed line in panel a). The enhancement disappears, and the distribution is turned into a smooth tail reminiscent of the MC results. For four RUs, the difference between MD and MC becomes much more pronounced (Figure 9b). In the MD simulation, the stretched conformation becomes less dominant while compact conformations gain importance, shown by the peak at 1 nm that comprises a total weight of about 30%. A similar analysis as in the case of three repeat units again reveals the importance of B_{II} for the formation of these states (dashed line in panel b). It

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523–2534

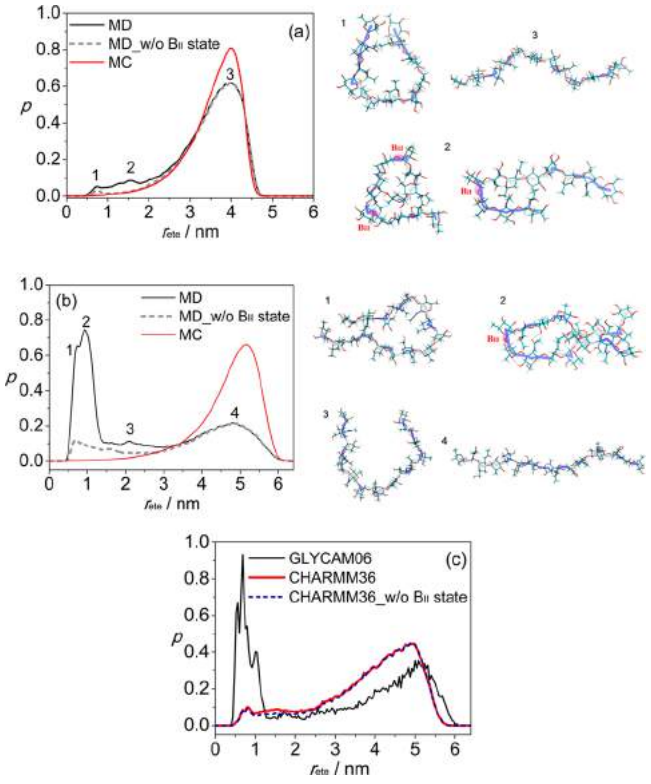


Figure 9. Distribution $p(r_{ete})$ of dodeca- and hexadecasaccharides with Seq 2(CDAB) as the repeating unit. The distributions for (a) 3RU and (b) 4RU calculated from implicit solvent MD (black) and MC (red) are shown. The dashed lines represent modified (not normalized) distributions where all conformations containing any B_{II} state in the MD simulations have been excluded. The conformational snapshots correspond to the end-to-end distances labeled in the graphs. In (a), for distance **2**, two structures are shown with two and one B_{II} state visited, respectively, where naturally the latter represents the more frequent conformation. **1** and **3** are structures with A states only. In panel b, only in snapshot **2** a B_{II} state is occupied. (c) $p(r_{ete})$ acquired from 200 ns MD trajectories with the GLYCAM06(g) (black solid line) and CHARMM36 (red solid line) force fields for the 4RU case with explicit solvent.

should be noted that the enhancement a low r_{ete} in both cases is not simply due to kinetically trapped conformations, i.e., insufficient sampling. Between the compact states around 1 nm and the major extended conformation at 5 nm there are frequent transitions on the time scale of the simulation (4 μs, see the Supporting Information, Figure S6).

We have to conclude that the formation of relatively stable compact states must be a result of partly attractive intramolecular interactions. While hairpin-like conformations are not a prerequisite for reaching vary small end-to-end distances (these can be realized within A states only, with the polysaccharide smoothly bending back on itself), the possibility to visit the B_{II} state facilitates that intramolecular interactions can become effective. We emphasize that the special role attributed to the state B_{II} has been verified by performing a similar analysis as described above by excluding states B_I, B_{III}, and B_{IV} from the analysis, with state B_{II} clearly showing the most pronounced influence on the global conformations.

As a reference, we have carried out simulations in explicit solvent using GLYCAM06(g) and CHARMM36 for the case of 4RUs (Figure 9c). Although the sampling cannot be considered converged, the simulations largely support the results obtained with implicit solvent, for both force fields: similar to the GLYCAM case, we have performed corresponding tests with CHARMM36 and results obtained with implicit and explicit solvent are fully consistent. All distributions in Figure 9c comprise data from simulation runs of 200 ns. For the CHARMM force field, however, the enhancement at small r_{ete} is almost absent. Excluding the B_{II} states from the analysis only slightly depletes the distribution there (dashed line), although we note that all of the conformations displayed along with Figure 9a,b occur but do not lead to stabilization. One might well argue that owing to the very low population of B states the intramolecular interactions cannot compensate the penalty in free energy to be paid to increase the population of B_{II}, which has been demonstrated to be a prerequisite to maintain a metastable compact conformation. Indeed, converting the populations of linkage II found in Table 1 for the disaccharide into free energies by Boltzmann inversion, one finds that the difference in free energy for B_{II} between both force fields is about 1 kcal/mol. Although notable, this value is rather small, and the possible differences in the strength of nonbonded interactions between the two force fields could easily account for it. However, we should point out that the difference in distributions among A-states between the force fields in a compact conformation contributes to an energetic/entropic penalty as well. Disentangling the various contributions that decide over the formation of compact states will be subject of a forthcoming publication.

DISCUSSION AND CONCLUSIONS

In summary, we have pursued a systematic investigation of di-, tetra-, octa-, dodeca-, and hexadecamer fragments (four RUs) of *S. flexneri* O-antigen serotype Y by MD and ABMD simulations. MC simulations of reduced backbone models have been carried out in parallel as a control, in order to identify those cases where the simple, intuitive picture of independent glycosidic linkages breaks down. To compare the reduced with the fully atomistic model, an efficient implicit solvent model had to be employed, which by testing against explicit solvent simulations was found an excellent approximation. Exploiting the studies with implicit solvent, it was demonstrated that conformations in the cases of three and four RUs are influenced by intramolecular interactions beyond the disaccharide level. These are facilitated by extensive conformational changes due to the nonexoanomeric states B, by which larger portions of the O-antigen polysaccharide can be folded back onto itself. This was seen to be a purely geometrical aspect, independent of the force field. In turn, state B of linkage II was identified to be crucial for the formation of relatively stable compact states within the GLYCAM06(g) force field. These were absent within CHARMM36, where the state B_{II} is higher in energy. We have to conclude that the subtle interplay between the geometrical nature of conformations, the populations of all relevant angular states and the strength of nonbonded interactions can have a drastic impact on the appearance of polysaccharides such as the O-antigen studied in this work. This might also be relevant for experimental investigations. We have demonstrated that the force fields considered here may be undistinguishable in a typical NMR experiment by analysis of J couplings, at least on the level of disaccharides. The distinction between compact states and elongated polysaccharide chains, in contrast, might be possible using NMR.

Nevertheless, some general conclusions can be drawn from the simulation results, irrespective of the force field. The

dx.doi.org/10.1021/jp4111713 | J. Phys. Chem. B 2014, 118, 2523−2534

possibility of O-antigen chains of only a few RUs to assume rather diverse conformations is striking. The occurrence of rather extreme, hairpin-like motifs, for example, is interesting. Even though the statistical weight for these hairpins depends, for a single chain in solution, on the underlying force field, these conformations could just as well be induced by the interactions with cell membranes or protein scaffolds, and could be another important aspect for understanding O-antigen recognition by antibodies[28] or bacteriophages,[30] in the light of conformational selection. Conformational diversity should also be taken into account when conceptualizing structure of heterogeneous LPS layers with a distinct multimodal distribution of chain lengths. For instance, the morphology of the O-antigen coat might not resemble a brush but rather a web of entangled chains. Indeed, in a recent computational study,[46] a significant increase in conformational freedom of O-antigen chains has been observed when going from a homogeneous (constant chain length 10 RUs) to a heterogeneous (mixture of 10 and 5 RU). This also raises the question about the mobility of LPS molecules as a whole; that is, whether inhomogeneities in LPS composition across the bacterium's surface are maintained or how quickly they are removed. With respect to the diffusivity of several types of phospholipid, most force fields lead to values very similar to experiment,[47] of the order of 10^{-7} cm^2/s at ambient or slightly elevated temperatures. In a computational study of the outer membrane of *E. coli*, the diffusivity within the leaflet with the LPS fraction (only lipid A and inner core part) is one order of magnitude smaller than that of the other one with phospholipids.[15] It is, however, unclear whether the diffusivity is influenced by the amount and length of the O-antigen fragments, in particular by the global conformations of the polysaccharide chains.

Clearly, to explore these questions further, more efficient computational models are required in order to access both larger length and time scales. In future studies, we are going to utilize recent coarse-graining (CG) strategies employing united or superatoms, where a small group of atoms is combined into a single interaction site (bead). A common approach that has evolved for carbohydrates is to substitute one six-ring by three CG beads.[48−51] The molecular topology is replaced by effective bonds, angles, and torsions between the CG beads. Usually, the glycosidic linkage is then represented by a single bond, and the relative orientation of two consecutive carbohydrate rings is determined by additional angle bending potentials. The extensive and accurate free energy landscapes displayed in Figure 2 can then be applied to these reduced degrees of freedom. However, our analysis of extreme but rare conformations involving the nonexoanomeric states B suggests that it might be useful to preserve the atomistic topology as given by the reduced backbone model and use the landscapes in Figure 2 directly in order to truthfully retain all details of glycosidic conformations. This is viable as the excellent agreement between implicit and explicit solvent models suggests that we might start directly with matching energies between implicitly solvated all-atom MD and coarse-grained energies,[48] without the need to include an explicit coarse-grained water model. In addition, the discrepancies found between GLYCAM06(g) and CHARMM36 with respect to the conformational ensembles for longer O-antigen chains might serve as a crucial benchmark about how accurately the corresponding coarse graining for each force field has been carried out.

■ ASSOCIATED CONTENT

Supporting Information

Simulation parameters (Section A), discussion on the use of adaptively biased molecular dynamics in this work (Section B), assessment of the validity of implicit solvent models (Section C), and use and scope of the reduced backbone model (Section D). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.santer@mpikg.mpg.de. Phone: +49-(0)331-567-9610.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank M. Wehle for technical assistance and G. Widmalm for useful and stimulating discussions.

■ ABBREVIATIONS

ABMD: adaptively biased moleculear dynamics; AFM: atomic force microscopy; CG: coarse grained; LPS: lipopolysaccharide; MD: molecular dynamics; MC: monte carlo; RU: repeat unit

■ REFERENCES

(1) Raetz, C. R.; Whitfield, C. Lipopolysaccharide Endotoxins. *Annu. Rev. Biochem.* **2002**, *71*, 635−700.

(2) Kingsley, R. A.; Bäumler, A. J. Host Adaptation and the Emergence of Infectious Disease: the *Salmonella* Paradigm. *Mol. Microbiol.* **2000**, *36*, 1006−1014.

(3) Stenutz, R.; Weintraub, A.; Widmalm, G. The Structures of *Escherichia coli* O-polysaccharide Antigens. *FEMS Microbiol. Rev.* **2006**, *30*, 382−403.

(4) Allison, G. E.; Verma, N. K. Serotype-converting Bacteriophages and O-antigen Modification in *Shigella flexneri*. *Trends Microbiol.* **2000**, *8*, 17−23.

(5) Kalynych, S.; Valvano, M. A.; Cygler, M. Polysaccharide Co-polymerases: the Enigmatic Conductors of the O-antigen Assembly Orchestra. *Protein Eng. Des. Sel.* **2012**, *25*, 797−802.

(6) Barry, E. M.; Pasetti, M. F.; Sztein, M. B.; Fasano, A.; Kotloff, K. L.; Levine, M. M. Progress and Pitfalls in *Shigella* Vaccine Research. *Nat. Rev. Gastroenterol. Hepatol.* **2013**, *10*, 245−255.

(7) Murray, G. L.; Attridge, S. R.; Morona, R. Altering the Length of the Lipopolysaccharide O antigen Has an Impact on the Interaction of *Salmonella enterica* Serovar Typhimurium with Macrophages and Complement. *J. Bacteriol.* **2006**, *188*, 2735−2739.

(8) Hölzer, S. U.; Schlumberger, M. C.; Jäckel, D.; Hensel, M. Effect of the O-antigen Length of Lipopolysaccharide on the Functions of Type III Secretion Systems in *Salmonella enterica*. *Infect. Immun.* **2009**, *77*, 5458−5470.

(9) Amro, N. A.; Kotra, L. P.; Wadu-Mesthrige, K.; Bulchevy, A.; Mobashery, S.; Liu, G.-Y. High-resolution Atomic Force Microscopy Studies of the *Escherichia coli* Outer Membrane: Structural Basis for Permeability. *Langmuir* **2000**, *16*, 2789−2796.

(10) Walker, S. L.; Redman, J. A.; Elimelech, M. Role of Cell Surface Lipopolysaccharides in *Eschcherichia coli* K12 Adhesion and Transport. *Langmuir* **2004**, *20*, 7736−7746.

(11) Shroll, R. M.; Straatsma, T. P. Molecular Structure of the Outer Bacterial Membrane of Pseudomonas Aeruginosa via Classical Simulation. *Biopolymers* **2002**, *65*, 395−407.

(12) Lins, R. D.; Straatsma, T. P. Computer Simulations of the Rough Lipopolysaccharide Membrane of Pseudomonas Aeruginosa. *Biophys. J.* **2001**, *81*, 1037−1046.
(13) Kotra, L. P.; Golemi, D.; Amro, N. A.; Liu, G.-Y.; Mobashery, S. Dynamics of the Lipopolysaccharide Assembly on the Surface of *Escherichia coli*. *J. Am. Chem. Soc.* **1999**, *121*, 8707−8711.
(14) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Conformational Dynamics and Membrane Interactions of the *E. coli* Outer Membrane Protein FecA: A Molecular Dynamics Simulation Study. *Biochim. Biophys. Acta* **2013**, *1828*, 284−293.
(15) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Electroporation of the *E. coli* and *S. aureus* Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes. *J. Phys. Chem. B* **2011**, *115*, 13381−13388.
(16) Morona, R.; Daniels, C.; Van Den Bosch, L. Genetic Modulation of *Shigella flexneri* 2a Lipopolysaccharide O-antigen Modal Chain Length Reveals that it has been Optimized for Virulence. *Microbiology* **2003**, *149*, 925−939.
(17) West, N. P.; et al. Optimization of Virulence Functions through Glucosylation of *Shigella* LPS. *Science* **2005**, *307*, 1313−1317.
(18) Clément, M.-J.; Imberty, A.; Phalipon, A.; Pérez, S.; Simenel, C.; Mullard, L. A.; Delepierre, M. Conformational Studies of the O-specific Polysaccharide of *Shigella flexeneri* 5a and of Four Related Synthetic Pentasaccharide Fragments Using NMR and Molecular Modeling. *J. Biol. Chem.* **2003**, *278*, 47928−47936.
(19) Kirschner, K. N.; Lins, R. D.; Maass, A.; Soares, T. A. A Glycam-based Force Field for Simulations of Lipopolysaccharide Membranes: Parametrization and Validation. *J. Chem. Theory Comput.* **2012**, *8*, 4719−4731.
(20) DeMarco, M. L.; Woods, R. J. Structural Glycobiology: A Game of Snakes and Ladders. *Glycobiology.* **2008**, *18*, 426−440.
(21) Shen, T.; Langan, P.; French, A. D.; Johnson, G. P.; Gnanakaran, S. Conformational Flexibility of Soluble Cellulose Oligomers: Chain Length and Temperature Dependence. *J. Am. Chem. Soc.* **2009**, *131*, 14786−14794.
(22) Landström, J.; Widmalm, G. Glycan Flexibility: Insights into Nanosecond Dynamics from a Microsecond Molecular Dynamics Simulation Explaining an Unusual Nuclear Overhauser Effect. *Carbohydr. Res.* **2010**, *345*, 330−333.
(23) Wehle, M.; Vilotijevic, I.; Lipowsky, R.; Seeberger, P. H.; Silva, D. V.; Santer, M. Mechanical Compressibility of the Glycosylphosphatidylinositol (GPI) Anchor Backbone by Independent Glycosidic Linkages. *J. Am. Chem. Soc.* **2012**, *134*, 18964−18972.
(24) Oliveira, R. G.; Schneck, E.; Quinn, B. E.; Konovalov, O. V.; Brandenburg, K.; Seydel, U.; Gill, T.; Hanna, C. B.; Pink, D. A.; Tanaka, M. Physical Mechanisms of Bacterial Survival Revealed by Combined Grazing-incidence X-ray Scattering and Monte Carlo Simulation. *C.R. Chim.* **2009**, *12*, 209−217.
(25) Schneck, E.; Schubert, T.; Konovalov, O. V.; Quinn, B. E.; Gutsmann, T.; Brandenburg, K.; Oliveira, R. G.; Pink, D. A.; Tanaka, M. Quantitative Determination of Ion Distributions in Bacterial Lipopolysaccharide Membranes by Grazing-incidence X-ray Fluorescence. *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 9147−9151.
(26) Pink, D. A.; Truelstrup Hansen, L.; Gill, T. A.; Quinn, B. E.; Jericho, M. H.; Beveridge, T. J. Divalent Calcium Ions Inhibit the Penetration of Protamine through the Polysaccharide Brush of the Outer Membrane of Gram-negative Bacteria. *Langmuir* **2003**, *19*, 8852−8858.
(27) Feris, K.; et al. Electrostatic Interactions Affect Nanoparticle-mediated Toxicity to Gram-negative Bacterium Pseudomonas Aeruginosa PAO1. *Langmuir* **2010**, *26*, 4429−4436.
(28) Theillet, F.-X.; Simenel, C.; Guerreiro, C.; Phalipon, A.; Mulard, L. A.; Delepierre, M. Effects of Backbone Substitutions on the Conformational Behaviour of *Shigella flexneri* O-antigens: Implications for Vaccine Strategy. *Glycobiology* **2011**, *21*, 109−121.
(29) Carlin, N. I. A.; Lindberg, A. A.; Bock, K.; Bundle, D. R. The *Shigella flexneri* O-antigenic Polysaccharide Chain − Nature of the Biological Repeating Unit. *Eur. J. Biochem.* **1984**, *139*, 189−194.
(30) Müller, J. J.; Barbirz, S.; Heinle, K.; Freiberg, A.; Seckler, R.; Heinemann, U. An Intersubunit Active Site between Supercoiled Parallel β Helices in the Trimeric Tailspike Endorhamnosidase of *Shigella flexneri* Phage Sf6. *Structure* **2008**, *16*, 766−775.
(31) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622−655.
(32) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. CHARMM Additive All Atom Force Field for Glycosidic Linkages between Hexapyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.
(33) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(34) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(35) Case, D. A.; et al. *AMBER11*; University of California: San Francisco, CA, 2010.
(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(37) Babin, V.; Roland, C.; Sagui, C. Adaptively Biased Molecular Dynamics for Free Energy Calculations. *J. Chem. Phys.* **2008**, *128*, 134101.
(38) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824−19839.
(39) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(40) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383−394.
(41) Allinger, N. L. *Molecular structure: Understanding steric and electronic effects from molecular mechanics*; John Wiley & Sons, Inc.: Hoboken, NJ, 2010; pp 189−211.
(42) Frank, M.; Lütteke, T.; Von Der Lieth, C.-W. GlycoMapsDB: A Database of the Accessible Conformational Space of Glycosidic Linkages. *Nucleic Acids Res.* **2007**, *35*, 287−290.
(43) Jonsson, K. H. M.; Säwén, E.; Widmalm, G. Studies on the Conformational Flexibility of α-L-rhamnose-containing Oligosaccharides Using ^{13}C-site-specific Labeling, NMR Spectroscopy and Molecular Simulations: Implications for the Three-dimensional Structure of Bacterial Rhamnan Polysaccharides. *Org. Biomol. Chem.* **2012**, *10*, 2453−2463.
(44) Zaccheus, M.; Pendrill, R.; Jackson, T. A.; Wang, A.; Auzanneau, F.-I.; Widmalm, G. Conformational Dynamics of a Central Trisaccharide Fragment of the Le^{a}Le^x Tumor Associated Antigen Studied by NMR Spectroscopy and Molecular Dynamics Simulations. *Eur. J. Org. Chem.* **2012**, *25*, 4705−4715.
(45) Kirschner, K. N.; Woods, R. J. Solvent Interactions Determine Carbohydrate Conformation. *Proc. Nat. Acad. Sci. U.S.A.* **2001**, *98*, 10541−10545.
(46) Wu, E. L.; Engström, O.; Jo, S.; Stuhlsatz, D.; Yeom, M. S.; Klauda, J. B.; Widmalm, G. Molecular Dynamics and NMR Spectroscopy Studies of *E. coli* Lipopolysaccharide Structure and Dynamics. *Biophys. J.* **2013**, *105*, 1444−1455.
(47) Piggot, T. J.; Piñeiro, Á.; Khalid, S. Molecular Dynamics Simulations of Phosphatidyl-choline Membranes: A Comparative Force Field Study. *J. Chem. Theory Comput.* **2012**, *8*, 4593−4609.
(48) Molinero, V.; Goddard, W. A., III 3B: A Coarse Grain Force Field for Molecular Simulations of Malto-oligosaccharides and their Water Mixtures. *J. Phys. Chem. B* **2004**, *108*, 1414−1427.
(49) Liu, P.; Izvekov, S.; Voth, G. A. Multiscale Coarse-Graining of Monosaccharides. *J. Phys. Chem. B* **2007**, *111*, 11566−11575.

(50) Lopez, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. Martini Coarse-grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195−3210.
(51) Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. Coarse-Grain Model for Glucose, Cellobiose, and Cellotetraose in Water. *J. Chem. Theory Comput.* **2011**, *7*, 2137−2150.
(52) Hardy, B. J.; Bystricky, S.; Kovac, P.; Widmalm, G. Conformational Analysis and Molecular Dynamics Simulation of α-(1→2) and α-(1→3) Linked Rhamnose Oligosaccharides: Reconciliation with Optical Rotation and NMR Experiments. *Biopolymers* **1997**, *41*, 83−96.
(53) Säwén, E.; Massad, T.; Landersjö, C.; Damberg, P.; Widmalm, G. 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→2)-linked Mannose Disaccharide Present in *N*- and *O*-linked Glycoproteins. *Org. Biomol. Chem.* **2010**, *8*, 3684−3695.

NOTE ADDED IN PROOF

The distribution functions $p(r_{ete})$ of the end-to-end distance as obtained here resemble the corresponding distribution functions for the worm-like chain model. A detailed comparison of our simulation data with the latter model, which leads to an estimate for the bending stiffness and the persistence length of the oligosaccharide chains, will be presented in a forthcoming publication.