

Conformational Analysis of Oligoarabinofuranosides: Overcoming Torsional Barriers with Umbrella Sampling

Shahidul M. Islam,[†] Michele R. Richards,[‡] Hashem A. Taha,[‡] Simon C. Byrns,[‡] Todd L. Lowary,[‡] and Pierre-Nicholas Roy^{*,†}

[†]Department of Chemistry, University of Waterloo, Waterloo, ON, Canada N2L 3G1

[‡]Department of Chemistry and Alberta Ingenuity Centre for Carbohydrate Science, University of Alberta, Edmonton, AB, Canada T6G 2G2

 Supporting Information

ABSTRACT: In this report, the conformations of a series of mono- and oligoarabinofuranosides were probed through the use of umbrella sampling simulations with the AMBER force field and the GLYCAM carbohydrate parameter set. The rotamer population distribution about the exocyclic C4–C5 bonds and the puckering distributions of the rings obtained from these umbrella sampling simulations were found to be in excellent agreement with those obtained from conventional long MD simulations for small monosaccharide fragments. For larger systems, the conventional MD approach becomes impractical, and we propose the use of umbrella sampling to circumvent poor sampling of certain conformations. The same umbrella sampling simulations were used to calculate the distributions about the vicinal protons and ensemble-averaged vicinal proton–proton coupling constants ($^3J_{\text{H,H}}$). The distributions about the vicinal protons of a monomer, methyl- α -L-arabinofuranoside (**1**), were found to be very similar to those obtained from direct umbrella sampling simulations about the vicinal protons. We calculated $^3J_{\text{H,H}}$ based on DFT-based Karplus-like relationships for L-arabinofuranosides. The $^3J_{\text{H,H}}$ values were found to be very similar to those obtained with the conventional MD simulations. For **1**, the $^3J_{\text{H,H}}$ values obtained with the DFT-based Karplus equations agree very well with experimental results; the agreement is, however, not as good for the larger oligomers. An approach to determine the experimental rotamer populations from the simulations is also discussed.

INTRODUCTION

In nature, the monosaccharide arabinose is found in both possible cyclized ring forms (pyranose and furanose) and absolute stereochemistries (D and L).¹ The furanose forms are more prevalent than the pyranose forms, and the distribution of D- vs L-arabinofuranose is species-specific. D-Arabinofuranose is found predominantly in cell wall polysaccharides in mycobacteria (including the human pathogen *Mycobacterium tuberculosis*) and other members of the actinomycetes family of bacteria.² On the other hand, L-arabinofuranose is abundant in plant cell walls.^{3,4} Regardless of their source, arabinofuranose-containing glycans play important roles in the organisms that produce them and in interactions with their environment. For example, the D-arabinofuranose-containing polysaccharides present in mycobacteria are essential for viability.^{2,4} In addition, plant glycoproteins containing L-arabinofuranose moieties are believed to be essential in diverse functions such as cell division and plant–microbe interactions,⁵ and these glycoconjugates have also been implicated in the response to some plant allergens.⁶

In previous studies, we have reported conformational investigations on molecules containing D-arabinofuranose rings using a combination of NMR spectroscopy and *ab initio*/density functional theory (DFT) or molecular dynamics calculations.^{7–17} These studies were carried out with the expectation that a better understanding of the conformational preferences of D-arabinofuranose-containing polysaccharides (D-arabinans) would facilitate the design of molecules that would interfere with their binding to proteins. For example, such molecules targeted to enzymes involved in the

biosynthesis of D-arabinans would be anticipated to be lead compounds for the treatment of mycobacterial diseases including tuberculosis and leprosy.

Having developed an understanding of the conformational preferences of D-arabinans in solution, we endeavored to study their interactions with proteins using computational methods. Access to X-ray crystal structures of proteins in complex with molecules containing D-arabinofuranose rings would greatly facilitate such investigations. Structural information of this type was, however, unavailable at the time this study was initiated.¹⁸ We therefore turned our attention to L-arabinans, for which a larger amount of crystal structure data is available,^{19–22} and focused on the structure of a mutant arabinofuranosidase, which was obtained in complex with an L-arabinofuranose trisaccharide.²³

The present study has been carried out to better understand the conformational preferences of larger oligoarabinofuranosides (Figure 1) and relies on techniques that provide enhanced sampling of conformational space. Furanosides are highly flexible compared to their pyranoside counterparts,²⁴ and the rings can occupy several envelope (E) and twist (T) forms with low energy barriers.¹¹ According to the Altona–Sundaralingam notation,^{25,26} each ring conformation can be described by a phase angle of pseudorotation (P) and a puckering amplitude (ϕ_m). Figure 2 shows the pseudorotational itinerary for the L-arabinofuranose ring.

Received: May 17, 2011

Published: August 04, 2011

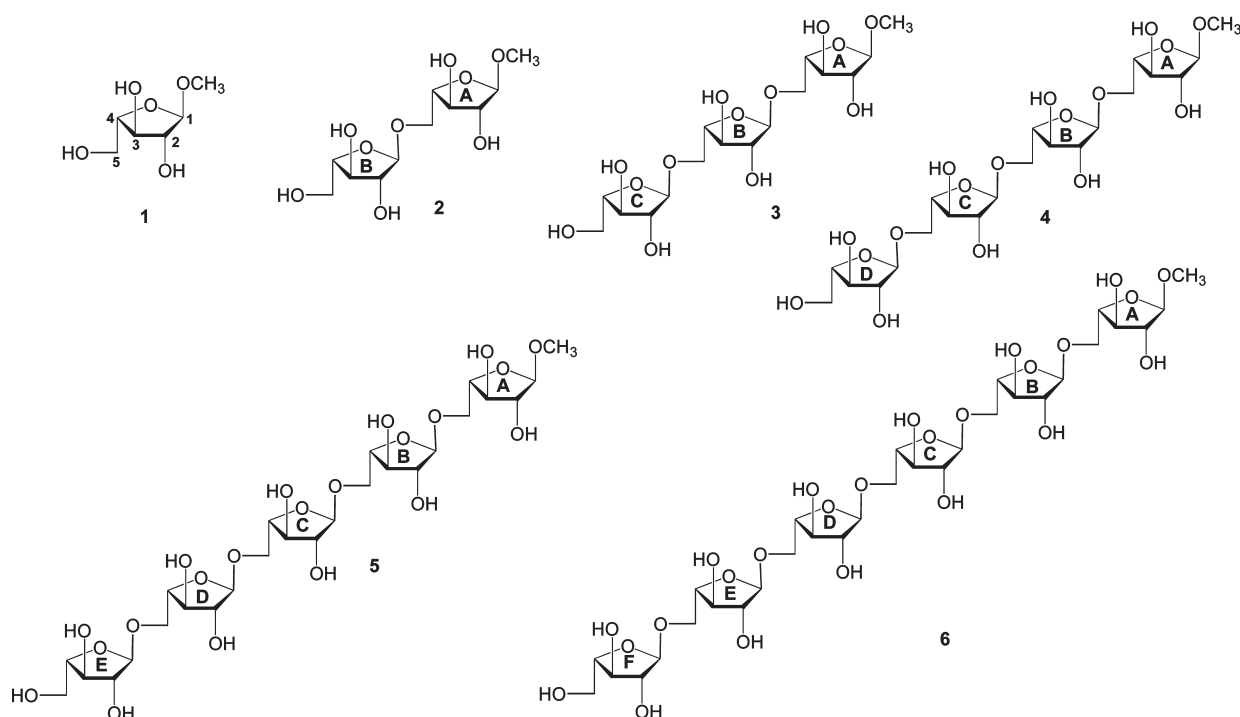


Figure 1. L-Arabinofuranosyl oligosaccharides studied in this paper.

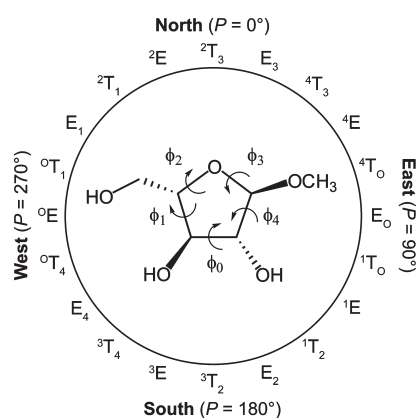


Figure 2. Pseudorotational itinerary for the L-arabinofuranose ring.

The angle P of a given conformer can be calculated from the five endocyclic torsion angles ϕ_0 – ϕ_4 as defined below

$$\tan P = \frac{(\phi_2 + \phi_4) - (\phi_1 + \phi_3)}{3.077\phi_0} \quad (1)$$

The puckering amplitude, ϕ_m , which measures the maximum displacement from the planar ring form, is related to P and ϕ_0 via the relation

$$\phi_m = \frac{\phi_0}{\cos P} \quad (2)$$

Other important features that should be considered during the conformational analysis of furanositides are rotamer populations about the exocyclic C–C and C–O bonds. In the present study, we investigated rotamer populations about the C4–C5 bond, which are influenced by a combination of steric and stereoelectronic

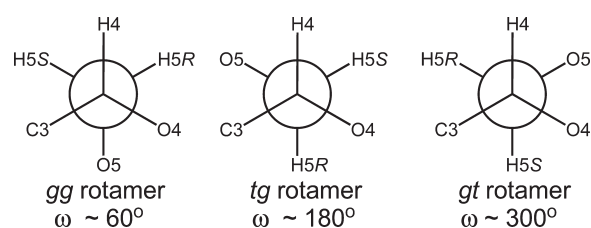


Figure 3. Definitions of the three staggered rotamers about the C4–C5 bond in L-arabinofuranosides, as shown looking down the bond from C4. The angle ω is defined as the O4–C4–C5–O5 torsion angle; see 1 in Figure 1 for the atom numbering scheme.

(gauche) effects.^{27–31} Figure 3 shows the three staggered rotamers about the C4–C5 bond in L-arabinofuranose rings.

Because both experimental rotamer and puckering distributions can be obtained from the analysis of NMR ^1H – ^1H vicinal coupling constants, $^3J_{\text{H,H}}$, a comparison with theoretically derived vicinal coupling constants allows one to assess the reliability of the simulation results. A common approach for assessing rotamer populations is to assume that a set of three discrete values of ϕ angles contains all conformational possibilities. In this case, the average coupling constant $\langle J \rangle$ is

$$\langle J \rangle = \sum_{i=1}^3 X_i J(\phi_i), \text{ with } \sum_{i=1}^3 X_i = 1 \quad (3)$$

where X_i ($i = 1, 2, 3$) are the unknown populations of the discrete rotamers and $J(\phi_i)$ are the vicinal coupling constants between H4 and H5R or H5S for the three values of ϕ_i . The values of $J(\phi_i)$ are calculated on the basis of ϕ_i , the dihedral angle between the two coupled protons. The ϕ_i values can either represent ideally staggered rotamers (60° , 180° , and 300°) or dihedral angles having high probability obtained from MD simulations. In a

previous study,¹⁶ we showed that the discrete approach as described above does not provide good average coupling constants when compared to experimentally obtained coupling constants or to those from a continuous approach. According to the continuous approach, $\langle J \rangle$ is measured as an average over the entire conformational space of the molecules using³²

$$\langle J \rangle = \int_0^{360} J(\phi) \rho(\phi) d\phi \quad (4)$$

where $J(\phi)$ is a Karplus relation that correlates the vicinal nuclear spin–spin coupling constants to the dihedral angle ϕ between the coupled spins, and $\rho(\phi)$ is the probability distribution of the dihedral angles about a particular bond. This approach provides coupling constants in good agreement with experimental results, suggesting the necessity of using a method that considers all angles across the range of 0 to 360°. An improved approach is the continuous probability distribution (CUPID) method,³² where $\rho(\phi)$ is represented as a Fourier series. In one study,³³ the CUPID method has been modified for five-member ring systems (CUPID-5). The Fourier series must be truncated, and the coefficients of the Karplus relationships must be accurate. This limits the use of this method for conformationally flexible molecules such as furanositides.

Various Karplus relationships are available in the literature to describe the dependence of the coupling constant on the dihedral angles between the coupled spins in arabinofuranositides.^{7,16,17,34} Recently developed DFT-derived relationships have been found to provide the best $^3J_{\text{H,H}}$ values when compared to experimental results:¹⁶

$$^3J_{1,2} = 4.62 + 3.16\cos(\phi) + 4.57\cos(2\phi) \quad (5)$$

$$^3J_{2,3} = 8.04 + 8.07\cos(\phi) + 7.24\cos(2\phi) \quad (6)$$

$$^3J_{3,4} = 4.44 + 0.50\cos(\phi) + 4.25\cos(2\phi) \quad (7)$$

$$^3J_{4,5S} = 4.95 - 0.42\cos(\phi) + 4.03\cos(2\phi) \quad (8)$$

$$^3J_{4,5R} = 5.23 + 0.02\cos(\phi + 15.1^\circ) + 4.67\cos(2\phi + 30.2^\circ) \quad (9)$$

We present here studies directed ultimately at understanding the conformation of L-arabinofuranose-containing oligosaccharides and their interaction with proteins. In particular, we explore the use of umbrella sampling in carrying out conformational searches of these molecules, and we compare our results with experimentally obtained $^3J_{\text{H,H}}$ values and conventional MD approaches, where applicable. The umbrella sampling approach was chosen because it allows one to overcome high free energy barriers along specified reaction coordinates. The advantages of the approach are expected to become increasingly important as the size of the oligofuranositide increases. In addition, we make an effort to develop a new approach to predict experimental rotamer populations that properly takes into account thermal fluctuations and is independent of the need to experimentally measure $^3J_{\text{H,H}}$ values. Such an approach could prove valuable in large molecules where spectral overlap prohibits the measurement of these parameters.

METHODS

All molecular dynamics simulations were carried out with the AMBER 10³⁵ suite of programs. The AMBER ff99SB force field³⁶ with the GLYCAM (version 04f) parameter set for carbohydrates^{37,38} was employed for the description of **1–6**. The additivity principle^{16,39,40} was used to build the topology of the oligosaccharides. We chose this particular version of the GLYCAM parameter set for consistency with our previous simulations on these ring systems.^{14–16} Note that similar C4–C5 rotamer populations and ring puckering distribution were obtained from a recently developed GLYCAM06 parameter set⁴¹ for the monomer, methyl α -L-arabinofuranositide **1** (data not shown). For the solution simulations, **1** was solvated by 288 TIP3P⁴² water molecules with a total box size of 25.816 \times 24.997 \times 24.007 Å. The box sizes and number of waters were then gradually increased for oligosaccharides up to 35 \times 35 \times 35 Å and 1325 TIP3P waters for hexasaccharide (**6**).

Long MD Simulations. Prior to long MD simulations and umbrella sampling simulations, the systems were minimized and equilibrated. First, the water molecules were minimized, keeping the geometries of the saccharides constrained. The entire system was then minimized. In both minimization steps, a steepest descent energy minimization was carried out for 50 cycles. The conjugate gradient algorithm was then used for 950 cycles. A total of 100 ps of annealing was then carried out with 50 ps each for temperature heating (5 to 300 K) and cooling (300 to 5 K). This was then followed by a short equilibration run of 250 ps. During this run, the temperature of the systems was gradually increased from 5 K to 300 K (150 ps) and then kept constant (100 ps). The production solution simulations of all compounds were performed under NPT conditions where the temperature was kept at 300 K and the pressure at 1 atm to remain consistent with experimental conditions. To assess any effect the thermostat may have on the dynamics of our systems, separate MD simulations were performed using the Langevin⁴³ or Berendsen⁴⁴ thermostat to control the temperature of the simulation box. A collision frequency γ of 2.0 ps^{−1} was used for the Langevin thermostat. MD simulations of 200 ns in length were carried out for compounds **1–3**. In all of the simulations, a 1 fs integration time step was used. Periodic boundary conditions were used, and a cutoff of 8 Å was set for nonbonded interactions. The SHAKE⁴⁵ algorithm was used to fix all hydrogen-containing bonds to their equilibrium values. Long-range electrostatic behavior was controlled with the particle mesh Ewald (PME) method.^{46,47} Gas phase simulations were also performed to observe the effect of solvation. A cutoff of 18 Å was set for nonbonded interactions in the gas phase. Periodic boundary conditions and the PME method were not used in the gas phase simulations. Other parameters remained unchanged from those employed in the solution simulations.

Umbrella Sampling Simulations. The average distribution function, $\rho(\chi)$ (eq 10), along some reaction coordinate χ is defined as the Boltzmann weighted average:⁴⁸

$$\langle \rho(\chi) \rangle = \frac{\int dq \delta(\chi'(q) - \chi) e^{-V(q)/k_B T}}{\int dq e^{-V(q)/k_B T}} \quad (10)$$

where $V(q)$ is the total energy of the system as a function of the conformation q and $\chi'(q)$ is the functional dependence of the reaction coordinate on the conformation. The potential of mean force (PMF) $W(\chi)$, or the change in free energy along the

coordinate χ , can be defined as^{49–51}

$$W(\chi) = -k_B T \ln \langle \rho(\chi) \rangle \quad (11)$$

The accurate sampling of conformational space can, however, be hindered by the presence of large energy barriers along χ . Therefore, the calculation of $\rho(\chi)$ and the PMF by conventional MD simulations is unrealistic for large systems, especially if they require long equilibration periods for convergence. Umbrella sampling, originally proposed by Valleau and Torrie,^{52,53} is a useful tool to obtain information on processes that require extremely long simulation times. In umbrella sampling, several simulations along the chosen coordinate are carried out. A biasing potential, $V_b(\chi)$, is added to the total energy to enhance the sampling of certain regions of conformational space. A harmonic form is often chosen for the biasing potential and is defined as^{49,50}

$$V_b(\chi)_i = \frac{1}{2} k (\chi - \chi_i)^2 \quad (12)$$

where k is the force constant and χ_i is the target position. The separate simulations are then combined to obtain the unbiased $\rho(\chi)$ and its associated PMF. Among the various approaches to combine the simulation results,^{52–56} the weighted histogram analysis method (WHAM) proposed by Kumar et al.⁵⁷ uses all of the information present in the umbrella sampling without discarding the overlapping regions. The WHAM method is a practical approach to obtaining average $\rho(\chi)$ and the PMF.

In this study, the dihedral angle about one of the exocyclic C4–C5 bonds (i.e., O4–C4–C5–O5) was taken as the reaction coordinate χ . All of the umbrella sampling simulations were performed using the final structure obtained from a short equilibration run (150 ps) of the systems. A total of 72 windows with a window width of 5° were used to cover the entire dihedral angle range from 0° to 360°. We also increased the width to 10° (with 36 windows in total) to study convergence. For monosaccharide **1**, 200 ps simulations for each window were carried out to yield a total simulation time of 14.4 ns. The simulation time for each window was increased to 1 ns for oligosaccharides **2** and **3** and 2 ns for the larger oligosaccharides **4–6**, with a total simulation time of 72 and 144 ns, respectively. For oligosaccharides, only one torsional sampling over 5° increments has been performed in a given window. Therefore, 72 simulations were performed for a single torsional angle, and a total of 432 (72 × 6) simulations have been performed for **6**. A harmonic biasing potential was chosen as the biasing potential energy function, and the force constant k was set to 30 kJ mol^{−1} rad^{−2}. The Langevin⁴³ thermostat was used to regulate the temperature of the umbrella sampling simulations. Umbrella sampling simulations were also performed with a Berendsen thermostat for monomers to compare the behavior of the two thermostats. All of the other simulation parameters used in the umbrella sampling simulations were identical to the long MD simulations discussed above. Once the simulations were complete, the PMF and $\rho(\chi)$ were calculated as a function of dihedral coordinates about the C4–C5 bond, using the WHAM software package by Grossfield.⁵⁸ The bootstrap error analysis method⁵⁹ was utilized to obtain relative errors of the distribution of rotamer populations. During a bootstrapping procedure, some data points were randomly removed from the total ensemble; however, the total number of data points was kept constant by adding data points that were already in the ensemble. A number of such bootstrapping cycles were performed by randomly removing and duplicating the data points from the

ensemble. The distributions obtained from this bootstrapping procedure were then compared to the original rotamer population distribution obtained from the umbrella sampling simulation to calculate the relative error in the original distribution. Vicinal proton–proton coupling constants ($^3J_{H,H}$) and ring puckering of the compounds were also studied using the conformational ensembles obtained from the umbrella sampling simulations. Errors in $^3J_{H4,H5R}$ and $^3J_{H4,H5S}$ were calculated from the errors in distribution obtained from the bootstrap error analysis. Errors in $\langle ^3J_{H1,H2} \rangle$, $\langle ^3J_{H1,H2} \rangle$, and $\langle ^3J_{H1,H2} \rangle$ were calculated from the standard deviations of the simulation data points along the ring proton–proton dihedrals.

Hydrogen Bonding Analysis. An intramolecular hydrogen bonding analysis of arabinofuranosides was performed using the ptraj module of the AMBER suite. All hydroxyl hydrogen atoms were assigned as potential hydrogen bond acceptors, and all oxygen atoms were assigned as potential hydrogen bond donors. Hydrogen bonding was evaluated with a heavy atom cutoff distance of 4 Å and an angle cutoff of 120°. Percent occupancies of the hydrogen bonds throughout the simulations were also calculated.

RESULTS AND DISCUSSION

NMR Spectroscopy. The spectra for **1**, **3**, and **6** were acquired in D₂O at 300 K on a 600 MHz spectrometer. The $^3J_{H,H}$ values for **1** were determined from a 1D ¹H NMR spectrum and are listed in Table 2. To overcome the spectral overlap present in **3** and **6**, the variable time version of the 1D gradient-enhanced chemical shift selective filtering (ge-CSSF) TOCSY spectra^{60,61} was used to provide the $^3J_{H,H}$ values (Table 3). The spectra for **3** and **6** were simulated using the program WinDNMR⁶² in order to confirm the coupling constants. The spectra, as well as a table of the chemical shifts, can be found in the Supporting Information. As would be expected, the coupling constant data for all of the compounds are within ±0.1 Hz.⁷ The synthesis of **3** and **6** was carried out as described previously,^{7,63,64} and the details are provided in the Supporting Information.

Simulations. The length of the long MD and umbrella sampling simulations required that convergence of the rotamer populations of **1** and **3** first be estimated. For monosaccharide **1**, it was found that a 200 ns MD simulation was required to obtain converged populations of all of the rotamers with uncertainties of less than a few percentage units, while only 200 ps simulations per window were required in the case of umbrella sampling. Simulations of less than 50 ns for long MD and 120 ps for umbrella sampling produced significantly different rotamer populations with larger variances from those obtained after 200 ns and 200 ps simulations, respectively. Because a total of 72 windows were used in the umbrella sampling, the total simulation time for umbrella sampling is 14.4 ns, which is significantly lower than the 200 ns value required in the long MD simulation. We also performed umbrella sampling simulations with 36 windows with 10° intervals. These simulations required a total simulation time of 7.2 ns. The rotamer populations were found to be very similar to those obtained with the 14.4 ns simulation. Convergence studies of the rotamer populations of **3** show that a 200 ns MD simulation still provides converged rotamer populations with errors of a few percent in all three rotamers. However, simulations of less than 150 ns produce very different rotamer populations. In the case of umbrella sampling, only 1 ns for each window, which equals a total of 72 ns of simulation, was required to obtain the converged result for all rotamers. Results obtained

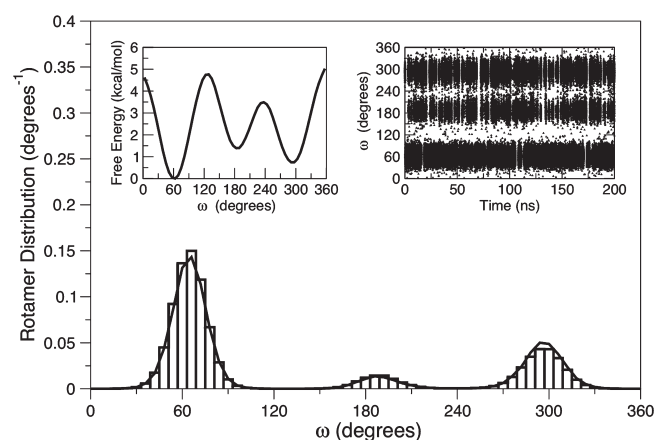


Figure 4. Comparison of the histograms obtained from the conventional MD (shown as a solid line) and umbrella sampling simulations (shown as a bar graph) of **1**. The left inset shows the PMF along the dihedral angle (ω) obtained from umbrella sampling, and the right inset shows the time dependence of ω during the long MD simulation.

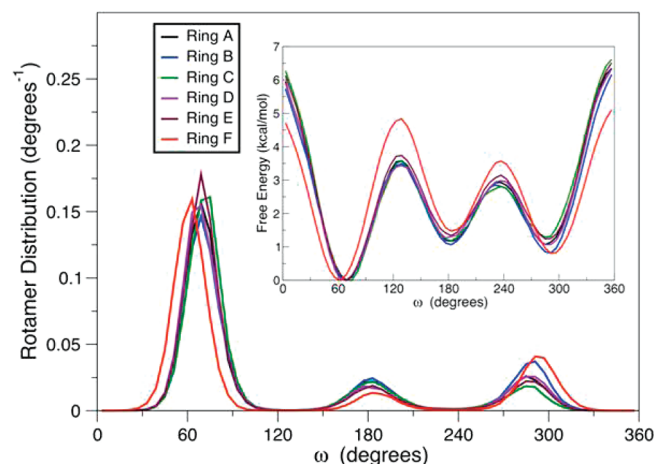


Figure 5. Comparison of the histograms of the dihedral angles (ω) obtained from the MD simulations of **6**. The inset shows the PMF along the dihedral angles (ω) obtained from umbrella sampling simulation.

from the umbrella sampling simulation are also expected to be superior because of the enhanced nature of the sampling. For larger oligomers **4–6**, only umbrella sampling simulations were performed with 2 ns simulations per window.

The rotamer population distribution about the C4–C5 bond of **1** obtained from both 200 ns MD simulations and 200 ps umbrella sampling simulations per window are shown in Figure 4. It is clear that both conventional MD and umbrella sampling simulations show similar distributions of rotamer populations with $gg > gt > tg$. The PMF obtained from the umbrella sampling simulations showed a $0.7 \text{ kcal mol}^{-1}$ difference in free energy between the *gg* and *gt* rotamers. This result underscores the importance of having accurate force field torsional parameters for these systems.

The umbrella sampling-derived rotamer distribution about the C4–C5 bonds involved in the α -(1 \rightarrow 5) linkages for the rings of the largest oligomer **6** are shown in Figure 5. Similar rotamer population distributions were also observed for oligomers **2–5**. The histograms show that the (1 \rightarrow 5)-linked rotamer populations all follow the same

Table 1. Rotamer Population (%) of **1–6** in Solution (TIP3P Water) Using US Simulations^a

structure	population	ring A	ring B	ring C	ring D	ring E	ring F
1	X_{gt}	34(2)					
	X_{tg}	9(1)					
	X_{gg}	55(2)					
2	X_{gt}	14(1)	24(1)				
	X_{tg}	15(1)	8(1)				
	X_{gg}	72(3)	67(2)				
3	X_{gt}	15(2)	11(1)	23(2)			
	X_{tg}	13(1)	15(1)	8(1)			
	X_{gg}	72(3)	74(3)	69(3)			
4	X_{gt}	11(1)	14(1)	16(1)	22(1)		
	X_{tg}	12(1)	15(1)	16(1)	7(1)		
	X_{gg}	77(3)	72(2)	68(2)	71(2)		
5	X_{gt}	10(1)	13(1)	14(1)	12(1)	23(1)	
	X_{tg}	12(1)	14(1)	14(1)	11(1)	9(1)	
	X_{gg}	78(2)	73(2)	72(2)	77(2)	69(2)	
6	X_{gt}	14(1)	20(1)	11(1)	15(1)	13(1)	22(1)
	X_{tg}	14(1)	15(1)	14(1)	12(1)	12(1)	8(1)
	X_{gg}	72(2)	66(2)	75(2)	73(2)	76(2)	70(2)

^a Values in parentheses are relative errors obtained from the bootstrap error analysis of the distributions obtained from the umbrella sampling simulation along the O4–C4–C5–O5 dihedral angle.

trend ($X_{gg} > X_{gt} \approx X_{tg}$), while the terminal rotamer populations follow the trend $X_{gg} > X_{gt} > X_{tg}$.

The inset of Figure 5 contains the PMFs of the rotation about the torsion angle ω . It is interesting to note that the two lowest free energy barriers along ω are actually highest in the case of the terminal ring (ring F). This suggests that the internal rings more readily explore their different rotameric states (i.e., are more flexible) than the unsubstituted ring. Quantifying the simulation rotamer populations was done by integrating the distributions of Figure 5 (Table 1).

From Table 1, it is clear that the C4–C5 rotamer population (%) decreases in the order $X_{gg} > X_{gt} > X_{tg}$ for methyl- α -L-arabinofuranoside (**1**) and all of the terminal rings of **2–6**, while the trend for internal rings is $X_{gg} > X_{gt} \approx X_{tg}$. Similar results were also observed for **1** and **3** with conventional MD simulations. With the umbrella sampling simulations, the conformation of larger α -L-arabinofuranosides and their interactions with large biological systems can be studied. Note that we also found that the rotamer distribution about the C4–C5 bond in **1** obtained with the Langevin thermostat agrees well with that obtained using the Berendsen thermostat.

Spin–Spin Coupling Constants. Having determined the rotamer distributions from simulation, we next compared our results to experimental results. The experimental rotamer distributions can be determined from eq 3. This approach, however, assumes discrete rotamers and, as mentioned above, can lead to errors. A more appropriate approach to check the reliability of our simulation is to compare the $\langle {}^3J_{H,H} \rangle$ obtained from eq 4 directly with the experimental coupling constants. The DFT-derived Karplus relationship for α -D-arabinofuranosides, eqs 5–9, can be used for α -L-arabinofuranosides.⁶⁵ However, it should be noted that for the α -L-arabinofuranoside enantiomers, the ${}^3J_{H4,H5R}$ and ${}^3J_{H4,H5S}$ functions are exchanged. Moreover, eq 9 has a phase factor, which should be adjusted accordingly for

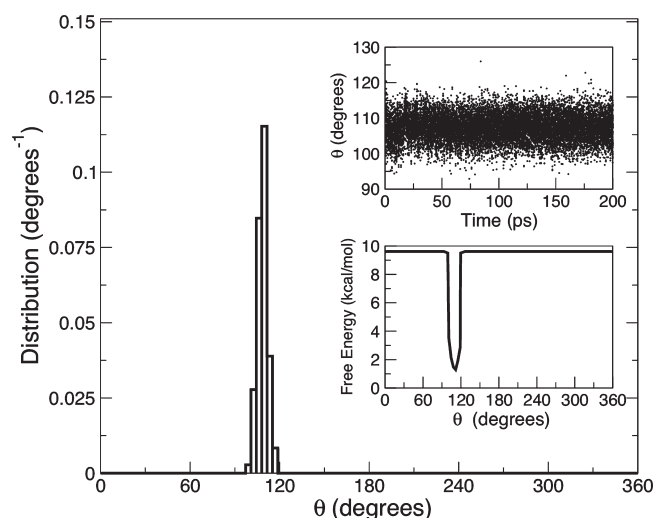


Figure 6. The distribution of the H4–C4–O4 angle (θ) obtained from the umbrella sampling simulations of **1**. The upper and lower insets show the time dependence and the PMF along the angle of θ , respectively.

$^3J_{\text{H4,H5S}}$ in α -L-arabinofuranosides. From our analysis (see Supporting Information), we found that the DFT-based Karplus relationships for $^3J_{\text{H4,H5R}}$ and $^3J_{\text{H4,H5S}}$ in α -L-arabinofuranosides are

$$^3J_{\text{H4,H5R}} = 4.95 - 0.42\cos(\phi) + 4.03\cos(2\phi) \quad (13)$$

$$^3J_{\text{H4,H5S}} = 5.23 + 0.02\cos(\phi - 15.1^\circ) + 4.67\cos(2\phi - 30.2^\circ) \quad (14)$$

Rotamer distributions along the C4–C5 bond ω were calculated using the conformational ensembles obtained from the umbrella sampling simulations. The calculation of spin–spin coupling constants requires the knowledge of rotamer distributions along the proton–proton dihedrals, i.e., the H4–C4–C5–H5R and H4–C4–C5–H5S angles, which can be obtained either from a long MD simulation or from an umbrella sampling simulation. Long MD simulations are undesirable for larger systems, and umbrella sampling simulations along each proton–proton dihedral would require additional calculations. As mentioned earlier, during the umbrella sampling simulations along the C4–C5 bond, a biasing window potential was applied at the reaction coordinate ω (i.e. the O4–C4–C5–O5 angle) and later unbiased using the WHAM approach. Because the simulations were done under biased conditions and the proton–proton dihedrals are on the same C4–C5 bond, a direct calculation of the proton–proton distributions from the combined umbrella sampling simulation trajectories does not provide a reliable rotamer distribution along the proton–proton dihedral. (See the Supporting Information for the proton–proton distributions obtained from combined umbrella sampling simulation trajectories.)

A theoretically better approach is to calculate these distributions from the unbiased rotamer distributions along the C4–C5 bond obtained from the umbrella sampling simulation. This approach is feasible if significant correlations exist between the O4–C4–C5–O5, H4–C4–C5–H5R, and H4–C4–C5–H5S dihedral angles. As can be seen in Figure 3, the rotamer populations along H4–C4–C5–H5R and H4–C4–C5–H5S are related to the rotamer

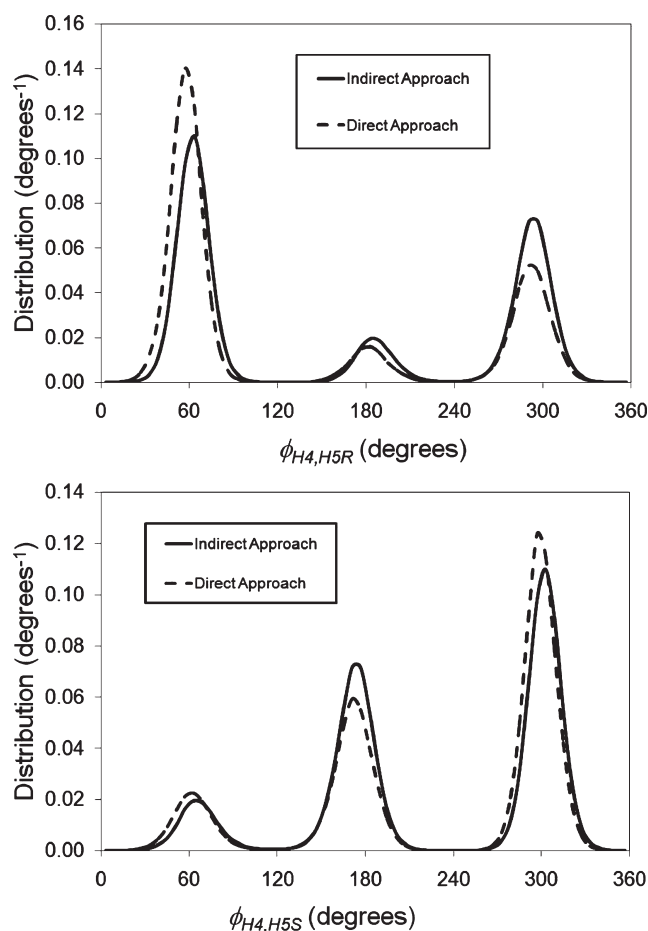


Figure 7. Comparison of the histograms obtained from indirect and direct umbrella sampling simulations about the proton–proton dihedral angles ($\phi_{\text{H4,H5R}}$ and $\phi_{\text{H4,H5S}}$) of **1**.

distribution of O4–C4–C5–O5 such that $\rho(\phi_{\text{H4,H5R}}) = \rho(\phi_{\text{O4,O5}})$ for all three rotamers and $\rho(\phi_{\text{H4,H5S}}) = \rho(\phi_{\text{O4,O5}} - 120)$, except in the case of the gg rotamer where $\rho(\phi_{\text{H4,H5S}}) = \rho(\phi_{\text{O4,O5}} - 120 + 360)$ for periodicity. We term the use of this assumption the “indirect approach”. Such relations are expected to hold because of the relative stiffness of the bond angles involving the various protons. To support this assumption, the H4–C4–O4 bond angle distribution obtained from the umbrella sampling simulations is shown in Figure 6. It is clear that the use of a biasing potential along the dihedral angle has little effect on the H4–C4–O4 bond angle. A long molecular dynamics simulation of **1** also provides a similar population distribution along the H4–C4–O4 bond angle. We also found that the bond angles in the oligomers were rigid as expected. In fact, the three similar bond angles in **3** are found to have the same average value ($\sim 110^\circ$). (See the Supporting Information for the distributions of bond angles in **3**.) Of course, this finding is apparent as the C4 atom forms a covalent bond with four other atoms through sp^3 hybridization. Nevertheless, our study confirms that the bond angles are more rigid than the dihedral angles along the C4–C5 bonds, and our indirect approach is feasible.

To further assess the reliability of the above assumptions, umbrella sampling simulations along the H4–C4–C5–H5R and H4–C4–C5–H5S dihedrals were also carried out for **1** (termed “direct approach”). The distributions about the

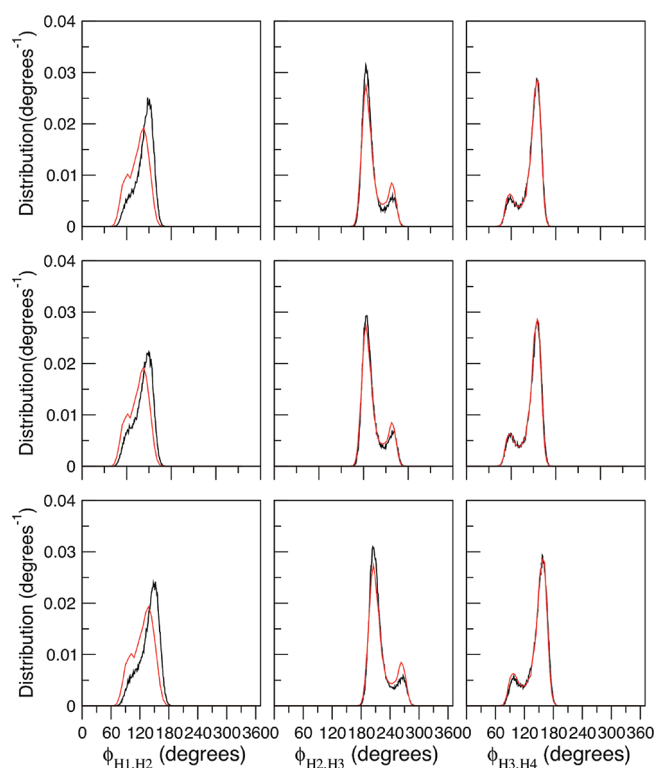


Figure 8. Comparison of the histograms obtained from umbrella sampling simulations about the proton–proton dihedral angles of rings A (upper panel), B (middle panel), and C (lower panel) of **3** (shown in black colored lines) and the ring proton–proton dihedral angles of **1** (shown in red colored lines). Due to the differences in H1–H2 dihedral angles, a ~ 1 Hz difference in coupling constants compared to experimental results can be observed.

Table 2. Comparison of Experimental and Theoretical ($\langle^3J_{H,H}\rangle$ Values (in Hz) for Monomer **1** Obtained from Umbrella Sampling Simulations

coupling	umbrella sampling	experimental
$\langle^3J_{H1,H2}\rangle$	2.1 (0.02)	1.7
$\langle^3J_{H2,H3}\rangle$	3.6 (0.02)	3.4
$\langle^3J_{H3,H4}\rangle$	5.4 (0.04)	5.8
$\langle^3J_{H4,H5R}\rangle$	3.2 (0.01)	3.3
	3.2 (0.01) ^a	
$\langle^3J_{H4,H5S}\rangle$	4.5 (0.02)	5.8
	4.1 (0.02) ^a	

^aValues obtained from the umbrella sampling along the dihedrals H4–C4–C5–H5R and H4–C4–C5–H5S.

H4–C4–C5–H5R and H4–C4–C5–H5S dihedrals obtained from the indirect and direct umbrella sampling approaches are shown in Figure 7. As can be seen in Figure 7, the overall distributions are quite similar. Using these distributions, $\langle^3J_{H4,H5R}\rangle$ and $\langle^3J_{H4,H5S}\rangle$ values were calculated and compared to the corresponding experimental values (Table 2). The $\langle^3J_{H4,H5R}\rangle$ obtained with the indirect approach are in excellent agreement with those obtained by the direct approach and experimentation. Similarly, the $\langle^3J_{H4,H5S}\rangle$ obtained with the indirect approach are also in good agreement with those obtained with the direct approach; although the magnitudes of $\langle^3J_{H4,H5S}\rangle$ calculated by both

Table 3. Comparison of Experimental and Theoretical ($\langle^3J_{H,H}\rangle$ Values (in Hz) for Each Ring in Oligomers **3** and **6** Obtained from Umbrella Sampling Simulations^a

coupling	US		exptl		US		exptl		US		exptl		
	ring A		ring B		ring C								
3	$\langle^3J_{H1,H2}\rangle$	3.0	1.7	2.7	1.5	2.9	1.6						
	$\langle^3J_{H2,H3}\rangle$	3.9	3.3	3.8	3.2	3.9	3.3						
	$\langle^3J_{H3,H4}\rangle$	5.4	5.8	5.4	5.9	5.5	6.0						
	$\langle^3J_{H4,H5R}\rangle$	2.9	3.2	3.0	3.2	3.1	3.3						
	$\langle^3J_{H4,H5S}\rangle$	3.4	5.8	3.4	5.8	3.7	5.9						
	ring A		ring B		ring C		ring D		ring E		ring F		
6	$\langle^3J_{H1,H2}\rangle$	3.0	1.7	3.1	1.6 ^b	2.6	1.6 ^b	2.9	1.6 ^b	2.9	1.6 ^b	2.8	1.5
	$\langle^3J_{H2,H3}\rangle$	4.0	3.2	3.9	3.2 ^b	3.7	3.2 ^b	4.0	3.2 ^b	3.8	3.2 ^b	3.8	3.3
	$\langle^3J_{H3,H4}\rangle$	5.5	5.8	5.4	6.0 ^b	5.3	6.0 ^b	5.3	6.0 ^b	5.3	6.0 ^b	5.3	6.0
	$\langle^3J_{H4,H5R}\rangle$	2.9	3.2	2.9	3.3 ^b	2.8	3.3 ^b	2.8	3.3 ^b	2.8	3.3 ^b	3.1	3.4
	$\langle^3J_{H4,H5S}\rangle$	3.4	5.8	3.8	5.8 ^b	3.4	5.8 ^b	3.4	5.8 ^b	3.3	5.8 ^b	3.8	5.9

^aexptl = experimental, US = umbrella sampling simulation. Errors in theoretical coupling constants are very small within the range of 0.02–0.05 Hz. (See the Supporting Information for errors in $\langle^3J_{H,H}\rangle$ values.) ^bThe experimental $\langle^3J_{H,H}\rangle$ values for residues B–E could not be measured individually due to spectral overlap. The average values for all four residues are shown.

approaches are lower than experimental results. It is possible that the DFT-based Karplus equation for $\langle^3J_{H4,H5S}\rangle$ in α -L-arabinofuranosides, eq 14, provides underestimated $\langle^3J_{H4,H5S}\rangle$ values. A more likely rationale for this discrepancy stems from undersampling of the *gt* rotamer, the largest contributor to $\langle^3J_{H4,H5S}\rangle$ in **1**, in the MD simulations. A similar effect was observed on $\langle^3J_{H4,H5R}\rangle$ in D-arabinofuranosides where MD simulations also underestimated the population of the *gt* rotamer and subsequently predicted a lower overall $\langle^3J_{H4,H5R}\rangle$ value.^{16,17}

Ring proton coupling constants were also calculated from the same umbrella sampling simulations. In this case, the dihedral angles between the coupled spins were calculated from the combined umbrella sampling trajectories. The distributions obtained from the dihedral angles were then used to calculate the average ring proton–proton coupling constants. We found that the distributions are in excellent agreement with those obtained from conventional MD, although an additional population in the distribution is found to be more pronounced in the umbrella sampling simulations (see Figure 8). Table 2 shows the comparison of experimental and theoretical (umbrella sampling) $\langle^3J_{H1,H2}\rangle$, $\langle^3J_{H2,H3}\rangle$, and $\langle^3J_{H3,H4}\rangle$ for **1**. All three theoretically obtained coupling constants are found to be in good agreement with experimental results. The $\langle^3J_{H1,H2}\rangle$, $\langle^3J_{H2,H3}\rangle$, and $\langle^3J_{H3,H4}\rangle$ obtained from the MD simulations of **1** (2.3, 3.8, and 5.7 Hz, respectively) are also very similar to that obtained from the umbrella sampling simulation. This suggests that, although this second populated region in the ring proton–proton population distributions is more pronounced in the umbrella sampling simulations, it has a minimal effect on the magnitude of the coupling constants.

The above analysis was also carried out for larger oligomers. The same Karplus-type equations were used for **3** and **6** as for **1**, even though glycosidic linkages are present. Karplus-like relationships derived for the internal and external C4–C5 bond of the α -(1 \rightarrow 5)-linked disaccharide of D-arabinofuranose showed little difference in earlier studies (see the Supporting Information

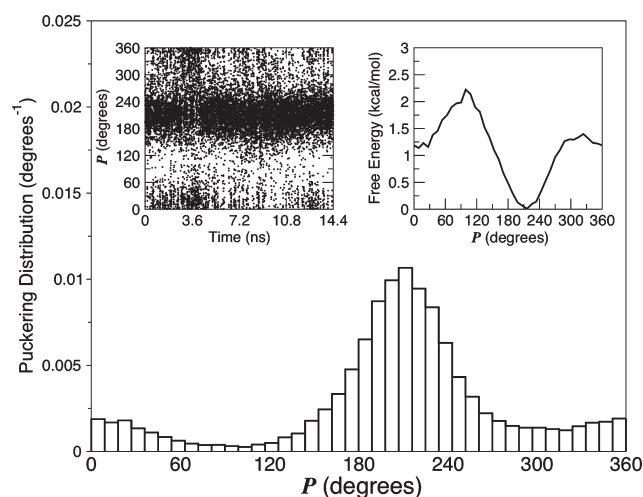


Figure 9. The distribution of the pseudorotational phase angle (P) for **1** in solution obtained from the umbrella sampling simulation. The left inset represents the time dependence of the P angle, and the right inset shows the PMF along the P of **1**.

for details). The simulation-derived $^3J_{H,H}$ values for each ring in **3** and **6** are shown in Table 3 and compared to experimental results.⁶⁶ The results demonstrate that the $^3J_{H4,H5R}$ and $^3J_{H3,H4}$ values are still in very good agreement with experimental results; on the contrary, $^3J_{H4,H5S}$, $^3J_{H1,H2}$, and $^3J_{H2,H3}$ deviate slightly when compared to those observed in the monomer (Table 2). The distributions along the H1–H2 dihedrals of **3**, as shown in Figure 8, are found to be slightly different than the corresponding H1–H2 distribution of **1**, which might be the reason for the ~ 1 Hz difference in simulation $^3J_{H1,H2}$ compared to experimental results. On the other hand, we suspect that a better Karplus-like relationship is necessary to obtain better values of $^3J_{H4,H5S}$ and $^3J_{H2,H3}$ for oligosaccharides. It should be noted that the above approach is only applicable if the system explores its ring conformers relatively quickly compared to the time scale of the exocyclic torsions and if there is little correlation between the rotamers and the ring conformers. Such a correlation analysis is presented in a forthcoming section.

Distribution of Ring Conformers. The same umbrella sampling simulations were also used for the analysis of ring puckering of **1**–**6**. For each umbrella sampling trajectory, the ring puckering, P , has been calculated. The distribution of P in **1** is found to be very similar for each window (see the Supporting Information). This indicates that changes in ring puckering occur on a much faster time scale than rotation about exocyclic torsions; earlier *ab initio* and DFT calculations on methyl α -D-arabinofuranoside provided results consistent with this.⁸ All of the umbrella sampling simulations corresponding to a ring were combined, and the distributions of puckering angle P and amplitude ϕ_m were then calculated. The variation of P for **1** obtained by umbrella sampling simulations is shown in Figure 9. Conventional MD also provides similar distribution of P for **1**. The umbrella sampling simulations of **1** predict that 77–84% of the conformations exist in the southern hemisphere of the pseudorotational itinerary with the area of the conformational space centered around 210–215° (3T_4), although a small fraction of the conformer is also present in the northern hemisphere (Figure 9). The agreement between experimental and theoretical ring coupling constants (Table 2) suggests that the ring conformations obtained

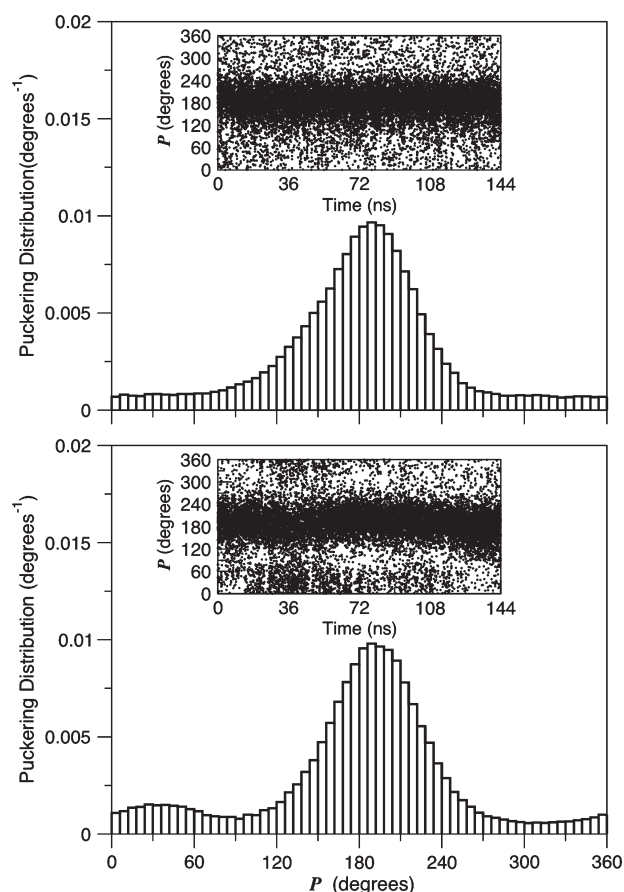


Figure 10. The distribution of the pseudorotational phase angle (P) for ring A (upper panel) and ring F (lower panel) of **6** in solution obtained from the umbrella sampling simulations. The insets represent the time dependence of the respective P angles of **6**. A minor secondary peak is also observed for the terminal ring of other oligomers with both umbrella sampling and long MD simulations.

from the simulation are accurate. The maximum free energy barrier along the P angle is about 2.2 kcal mol^{−1}, which justifies the relative ease of exploring the P compared to the ω conformational space. The distribution in ϕ_m is centered at about 36°, which agrees very well with those obtained from the crystal structure⁶⁷ and from the *ab initio* and DFT calculations.^{8–12,14}

Having successfully determined the distribution of ring conformations of **1** using umbrella sampling simulations, we focused our attention on the oligomers. Figure 10 represents the distribution of puckering for rings A and F of **6** in solution obtained from the umbrella sampling simulations. It is clear that the puckering distribution in ring A has a predominantly southern conformer at $P_S = 190$ – 200° (81–86%). Rings B, C, D, and E also have similar values for P and ϕ_m . On the contrary, ring F, which contains the terminal rotamer, has a slightly different distribution of P ; a second region of conformational space in the northern hemisphere, centered at $P_N = 30^\circ$, of the pseudorotational wheel is populated. The ring conformations of oligomers **2**–**5** are found to be very similar to those observed in the rings of **6**. The puckering amplitude (ϕ_m) remains the same for all of the rings in **1**–**6** with ϕ_m in the range of 35–40°.

Correlation Study of Rotamers and Ring Puckering. Having successfully determined the rotamer populations about the C4–C5 bond and ring puckering from umbrella sampling simulations, we

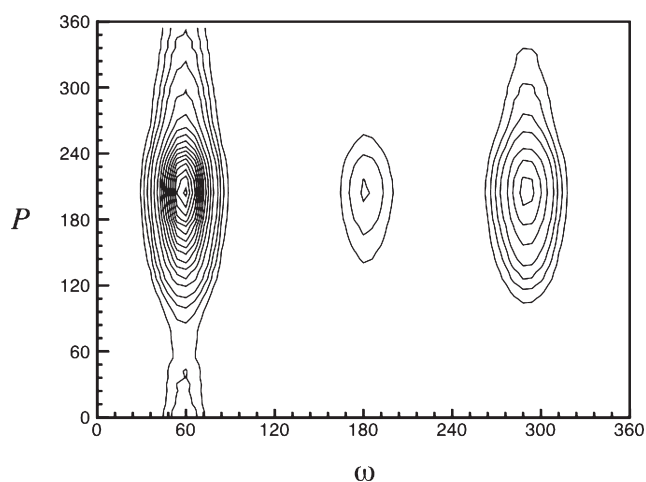


Figure 11. Joint probability distribution of the dihedral angle (ω) and the puckering angle (P) of **1** in the solution phase obtained from umbrella sampling simulation. The units of the angles ω and P are in degrees.

studied the correlation between them. The joint probability distribution ($\rho_{\omega,P}^{2D}$) of the C4–C5 torsion (ω) and puckering angle (P) for **1** and **3** calculated from umbrella sampling simulations are shown in Figures 11 and 12. Very little correlation between ω and P for **1** and **3** is observed; this is consistent with previous studies on α -D-arabinofuranoside using conventional MD simulations.^{8,14} Figures 11 and 12 provide a qualitative indication of the uncorrelated behavior of ω and P for compounds **1** and **3**, respectively, but the extent of their correlation is absent. Therefore, we next endeavored to obtain a quantitative measure of this correlation.

We first define the square root or amplitude of the true 2D distribution, $\rho_{\omega,P}^{2D}$, as

$$f = \sqrt{\rho_{\omega,P}^{2D}} \quad (15)$$

and its uncorrelated counterpart as the product of the 1D distributions for each degree of freedom:

$$g = \sqrt{\rho_{\omega}^{1D} \rho_P^{1D}} \quad (16)$$

To study the overlap between the correlated and uncorrelated distributions, we define the inner product:

$$\int dP d\omega f(P, \omega) g(P, \omega) = \langle f|g \rangle \quad (17)$$

Because both distributions are normalized, the above inner product would equal unity in the case of the absence of correlation in the $\rho_{\omega,P}^{2D}$ distribution. If there is a correlation between ω and P , the following quantity, χ , represents the magnitude of that correlation

$$\int d\tau (f - g)^2 = \chi^2 \quad (18)$$

where χ^2 ranges from 0 to 2. It can be readily shown that $\langle f|g \rangle = 1 - \chi^2/2$. Note that if $\langle f|g \rangle = 1$, χ^2 equals 0 and there is no correlation between ω and P , while if $\langle f|g \rangle = 0$, χ^2 equals 2 and there will be high correlation between ω and P . Intermediate values of $\langle f|g \rangle$ and χ^2 will indicate some correlation between ω and P . In this study, the $\langle f|g \rangle$ and χ^2 are found to be 0.99 and 0.02, respectively, for **1**, suggesting almost no correlation between ω

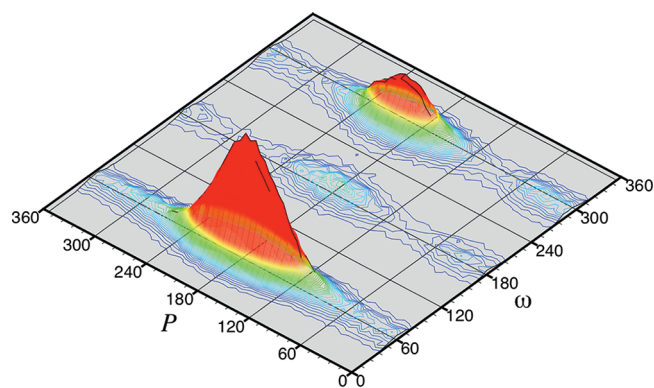


Figure 12. Joint probability distribution of the puckering angle (P) and the torsion angle (ω) for **3** in solution. All three rings have similar joint probability distributions. The units of the angles ω and P are in degrees.

and P . Correlation studies on oligomers also reveal that there is very little correlation between the rotamers and the respective ring conformations. The rotamer distributions show similar trends for all P values with $gg > gt > tg$ for **1–6**.

Determination of Experimental Rotamer Populations. We have seen that the average coupling constants for **1** obtained by using umbrella sampling simulations agree very well with experimental results (Table 2). This average coupling constant was obtained from all of the $J(\phi)$ and $\rho(\phi)$ values in the range of 0 to 360° and accounts for thermal fluctuations. When using eq 3 to determine rotamer populations from experimental results, only three discrete values of $J(\phi_i)$ and ϕ_i are used, and one, therefore, neglects these effects and assumes that each rotamer is “frozen” at some value of the dihedral angle. We propose here a simple approach to properly account for thermal fluctuations of the dihedral angle when determining experimental rotamer populations.

According to our approach, the average coupling constant $\langle J \rangle$ is given by the following sum of integrals:

$$\begin{aligned} \langle J \rangle = & \int J(\phi) \rho_{gg}(\phi) d\phi + \int J(\phi) \rho_{gt}(\phi) d\phi \\ & + \int J(\phi) \rho_{tg}(\phi) d\phi \end{aligned} \quad (19)$$

where $\rho_i(\phi)$ is the local distribution function corresponding to rotamer i . Umbrella sampling simulations along the H4–C4–C5–H5R and H4–C4–C5–H5S dihedrals reveal three distinct Gaussian type distributions in the range of 0 to 360°. Each of these Gaussian distributions corresponds to a particular rotamer. In our approach, the mean and standard deviations were calculated from the three populations of the O4–C4–C5–O5 dihedral angle obtained from the umbrella sampling simulation. As shown above, the H4–C4–C5–H5R and H4–C4–C5–H5S dihedral angle distributions can be obtained by analysis of the O4–C4–C5–O5 dihedral angle. Therefore, the mean and standard deviations can be used to construct thermal Gaussian distributions of H4–C4–C5–H5R and H4–C4–C5–H5S of the form

$$g_i(\phi) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-(\phi - \phi_i)^2/2\sigma^2} \quad (20)$$

for each rotamer i , where σ is the standard deviation and ϕ_i is the mean of the distribution. The experimental rotamer population

Table 4. Experimental Rotameric Distribution (%) about the C4–C5 Bond of 1, 3, and 6 in Solution Obtained Using the Gaussian Distribution Approach^a

	1			3			6					
		A	B	C	A	B	C	D	E	F		
X_{gt}	56	58	55	58	58	58	60	59	60	59		
X_{tg}	14	19	19	12	19	20	21	21	20	11		
X_{gg}	30	23	25	30	23	22	19	20	20	30		

^a The values obtained from the ideal and most probable dihedrals are very similar to those obtained from the Gaussian approach, and these are provided in the Supporting Information.

($X_{i=gg,gt,tg}$) is obtained by using expressions for ${}^3J_{H4,H5R}$ and ${}^3J_{H4,H5S}$ of the form below

$$\langle J \rangle = X_{gt} \int_0^{120} J(\phi) g_{gt}(\phi) d\phi + X_{tg} \int_{120}^{240} J(\phi) g_{tg}(\phi) d\phi + X_{gg} \int_{240}^{360} J(\phi) g_{gg}(\phi) d\phi \quad (21)$$

along with the restriction that the populations should add up to unity.

The experimental rotamer populations obtained for 1, 3, and 6 using our proposed approach are summarized in Table 4. The experimental rotamer populations for 1 and all of the terminal rotamers in 3 and 6 decrease from $X_{gt} > X_{gg} > X_{tg}$. In the case of internal rotamers in 3 and 6, the rotamer population decreases from $X_{gt} > X_{gg} \approx X_{tg}$. The differences between X_{gg} and X_{tg} are larger for the terminal rotamers than the internal rotamers in 3 and 6. We also determined mean and standard deviations of the populations of H4–C4–C5–H5R and H4–C4–C5–H5S dihedral angles of 1 using the conventional MD and umbrella sampling simulations. The mean and standard deviations were used to construct the Gaussian distributions about the H4–C4–C5–H5R and H4–C4–C5–H5S dihedrals, which were then used to calculate the experimental rotamer populations of 1. The experimental rotamer population obtained from the MD ($X_{gt} = 58\%$, $X_{tg} = 13\%$, $X_{gg} = 29\%$) and the umbrella sampling parameters ($X_{gt} = 57\%$, $X_{tg} = 12\%$, $X_{gg} = 31\%$) were found to be very similar to those obtained from the O4–C4–C5–O5 dihedral angle (Table 4). In our opinion, the thermal Gaussian distribution approach should be used because it properly accounts for thermal fluctuations in the dihedral angles unlike earlier studies^{14–17} where only either ideal dihedral angles (60°, 180°, and 300°) or the most probable dihedral values from conventional MD simulations were used to calculate the experimental rotamer population. In this study, the experimental rotamer population of 1 and 3 obtained from ideal and most probable dihedral angles were found to differ by a few percentage units when compared to the thermal Gaussian distribution approach.

A comparison of theoretical and experimental rotamer populations (Tables 1 and 4) shows that the rotamer populations obtained from simulation decrease from $X_{gg} > X_{gt} > X_{tg}$, while the experimental values decrease from $X_{gt} > X_{gg} > X_{tg}$. The experimental rotamer populations of course greatly depend on the quality of the J function. One explanation for the difference between experimental and theoretical rotamer populations is that the DFT-based Karplus relationship used for ${}^3J_{H4,H5S}$ led to an

average coupling constant that differed by about 1.3 Hz from experimental results (see Table 2). One could adjust the ${}^3J_{H4,H5S}$ function to obtain average couplings that agree better with experimental results and subsequently use this adjusted ${}^3J_{H4,H5S}$ function to repeat the calculation of experimental rotamer populations. We do not attempt this sort of adjustment here.

CONCLUSIONS

Umbrella sampling simulations were carried out to accurately determine the C4–C5 rotamer populations and the distribution of ring puckering in oligoarabinofuranosides 1–6. A comparison of the rotamer populations calculated with umbrella sampling and conventional MD simulations for monomer 1 reveals that umbrella sampling reproduces the results obtained from the MD simulation. For larger oligomers (e.g., 3), converged rotamer populations were obtained after several hundred nanoseconds of conventional MD simulations, which limits the use of conventional MD simulations for even larger oligomers (e.g., 6). On the other hand, convergence of the rotamer populations was achieved more efficiently with the umbrella sampling simulations. In general, an umbrella sampling simulation only allows one to calculate properties associated with the reaction coordinate being restrained. We have shown that for these oligofuranosides, results based on a C4–C5 dihedral reaction coordinate can also be used to obtain accurate proton–proton dihedral distributions. Vicinal proton–proton coupling constants (${}^3J_{H,H}$) could therefore be calculated from the information obtained from umbrella sampling simulations and using the DFT-based Karplus relationship for methyl α -L-arabinofuranoside. For 1, the ${}^3J_{H,H}$ values agree very well with those obtained directly from experimental results, while for oligomers 3 and 6 the ${}^3J_{H,H}$ values obtained using the umbrella sampling simulations are found to be very similar to those obtained from conventional MD simulations conducted on 3. Ring puckering distributions and amplitudes were also calculated from the umbrella sampling simulations, and the results agree very well with those obtained from the MD simulations. We have established that the umbrella sampling simulations along a particular bond can be used to determine ring conformations, as the energy barriers for changes in the ring geometries are much lower. A correlation study showed that there is no correlation between the ring conformation and rotamer population in α -L-arabinofuranosides in solution.

Because short umbrella sampling simulations provide reliable rotamer populations and ring puckering distributions in 1–6, it is a desirable and more efficient alternative to long MD simulations for the conformational study of larger oligofuranosides. The efficiency of the umbrella sampling approach also makes it possible to envisage simulations with more accurate descriptions of the electronic structure in order to capture polarization effects. The simulation of larger systems nevertheless remains a great challenge, and approaches to further improve sampling efficiency will continue to be developed.

ASSOCIATED CONTENT

S Supporting Information. Synthetic details, NMR spectra, and coupling constant data (experimental and theoretical); ${}^3J_{H4,H5}$ coupling profiles; C4–C5 rotameric distributions and convergence plots; dihedral angle and puckering distributions; and H-bond occupancies. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pnroy@uwaterloo.ca.

ACKNOWLEDGMENT

We are grateful to the Natural Sciences and Engineering Council of Canada (NSERC), the Canada Foundation for Innovation (CFI), and the Alberta Ingenuity Centre for Carbohydrate Science for financial support as well as the SHARCNET for dedicated computational resources. M.R.R. is supported by an Alberta Innovates Health Solutions Ph.D. studentship and H. A.T by a Queen Elizabeth II scholarship.

REFERENCES

- Lowary, T. L. *J. Carbohydr. Chem.* **2002**, *21*, 691.
- Brennan, P. J.; Nikaido, H. *Annu. Rev. Biochem.* **1995**, *64*, 29.
- Doblin, M. S.; Pettolino, F.; Bacic, A. *Funct. Plant Biol.* **2010**, *37*, 357.
- Kaur, D.; Guerin, M. E.; Skovierova, H.; Brennan, P. J.; Jackson, M. *Adv. Appl. Microbiol.* **2009**, *69*, 23.
- Seifert, G. J.; Roberts, K. *Annu. Rev. Plant Biol.* **2007**, *58*, 137.
- Leonard, R.; Petersen, B. O.; Himly, M.; Kaar, W.; Wopfner, N.; Kolarich, D.; van Ree, R.; Ebner, C.; Duus, J. O.; Ferreira, F.; Altmann, F. *J. Biol. Chem.* **2005**, *280*, 7932.
- D'Souza, F. W.; Ayers, J. D.; McCarren, P. R.; Lowary, T. L. *J. Am. Chem. Soc.* **2000**, *122*, 1251.
- McCarren, P. R.; Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2001**, *105*, 5911.
- Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Org. Chem.* **2000**, *65*, 4954.
- Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.* **1999**, *121*, 9682.
- Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2003**, *107*, 5763.
- Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2003**, *107*, 372.
- Houseknecht, J. B.; McCarren, P. R.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.* **2001**, *123*, 8811.
- Seo, M.; Castillo, N.; Ganzynkowicz, R.; Daniels, C. R.; Woods, R. J.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theory Comput.* **2007**, *4*, 184.
- Taha, H. A.; Castillo, N.; Roy, P. N.; Lowary, T. L. *J. Chem. Theory Comput.* **2009**, *5*, 430.
- Taha, H. A.; Castillo, N.; Sears, D. N.; Wasylshen, R. E.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theory Comput.* **2010**, *6*, 212.
- Taha, H. A.; Roy, P.-N.; Lowary, T. L. *J. Chem. Theory Comput.* **2011**, *7*, 420.
- It should be noted that structures of three related D-arabinofuranose-containing oligosaccharides in complex with an antibody Fab fragment were published recently: Murase, T.; Zheng, R. B.; Joe, M.; Bai, Y.; Marcus, S. L.; Lowary, T. L.; Ng, K. K. S. *J. Mol. Biol.* **2009**, *392*, 381.
- Taylor, E. J.; Smith, N. L.; Turkenburg, J. P.; D'Souza, S.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2006**, *395*, 31.
- Miyana, A.; Koseki, T.; Matsuzawa, H.; Wakagi, T.; Shoun, H.; Fushinobu, S. *J. Biol. Chem.* **2004**, *279*, 44907.
- Yamaguchi, A.; Tada, T.; Wada, K.; Nakaniwa, T.; Kitatani, T.; Sogabe, Y.; Takao, M.; Sakai, T.; Nishimura, K. *J. Biochem.* **2005**, *137*, 587.
- Paes, G.; Skov, L. K.; O'Donohue, M. J.; Remond, C.; Kastrup, J. S.; Gajhede, M.; Mirza, O. *Biochemistry* **2008**, *47*, 7441.
- Nurizzo, D.; Turkenburg, J. P.; Charnock, S. J.; Roberts, S. M.; Dodson, E. J.; McKie, V. A.; Taylor, E. J.; Gilbert, H. J.; Davies, G. J. *Nat. Struct. Biol.* **2002**, *9*, 665.
- Angyal, S. J. *Adv. Carbohydr. Chem. Biochem.* **1984**, *42*, 15.
- Altona, C.; Sundaral, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
- Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333.
- Oleary, D. J.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 6629.
- Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102.
- Devries, N. K.; Buck, H. M. *Carbohydr. Res.* **1987**, *165*, 1.
- Bock, K.; Duus, J. O. *J. Carbohydr. Chem.* **1994**, *13*, 513.
- Tvaroska, L.; Carver, J. P. *J. Phys. Chem. A* **1997**, *101*, 2992.
- Dzakula, Z.; Westler, W. M.; Edison, A. S.; Markley, J. L. *J. Am. Chem. Soc.* **1992**, *114*, 6195.
- Dzakula, Z.; DeRider, M. L.; Markley, J. L. *J. Am. Chem. Soc.* **1996**, *118*, 12796.
- Altona, C.; Francke, R.; de Haan, R.; Ippel, J. H.; Daalmans, G. J.; Westra Hoekzema, A. J. A.; Wijk, V. J. *Magn. Reson. Chem.* **1994**, *32*, 670.
- Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, 2008.
- Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
- Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraserreid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
- Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
- Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- Adelman, S. A.; Doll, J. D. *J. Chem. Phys.* **1976**, *64*, 2375.
- Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
- Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
- Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275.
- Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959.
- Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press Inc.: New York, 1987.
- Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press Inc.: New York, 1987.
- Torrie, G. M.; Valleau, J. P. *J. Chem. Phys. Lett.* **1974**, *28*, 578.
- Valleau, J. P.; Torrie, G. M. *A Guide for Monte Carlo for Statistical Mechanics. In Statistical Mechanics Part A*; Plenum Press: New York, 1977.
- Shen, J.; Mccammon, J. A. *J. Chem. Phys. Lett.* **1991**, *158*, 191.
- Woolf, T. B.; Roux, B. *J. Am. Chem. Soc.* **1994**, *116*, 5916.
- Haydock, C.; Sharp, J. C.; Prendergast, F. G. *Biophys. J.* **1990**, *57*, 1269.
- Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.
- Grossfield, A. *WHAM*, version 2.0.2.; 2011.
- Efron, B.; Tibshirani, R. J. *An Introduction to the Bootstrap*; Chapman and Hall, CRC: Boca Raton, FL, 1994.
- Liu, C.; Richards, M. R.; Lowary, T. L. *Org. Biomol. Chem.* **2011**, *9*, 165.
- Robinson, P. T.; Pham, T. N.; Uhrin, D. *J. Magn. Reson.* **2004**, *170*, 97.

- (62) Reich, H. J. In *J. Chem. Educ. Software*; 3D2: Madison, WI, 1996.
- (63) Yin, H.; D'Souza, F. W.; Lowary, T. L. *J. Org. Chem.* **2002**, *67*, 892.
- (64) Carapito, R.; Imbert, A.; Jeltsch, J. M.; Byrns, S. C.; Tam, P. H.; Lowary, T. L.; Varrot, A.; Phalip, V. *J. Biol. Chem.* **2009**, *284*, 12285.
- (65) For compounds 2–6, we used the Karplus equations derived for the monosaccharide **1**. The calculated coupling constants using Karplus equations developed specifically for (1→5)-linked disaccharide showed little difference when compared with those for the monosaccharide (Supporting Information).
- (66) Given the close similarity of the coupling constants for trisaccharide **3** and hexasaccharide **6**, analogous NMR studies on oligosaccharides **2**, **4**, and **5** were not carried out.
- (67) Evdokimov, A. G.; Kalb, A. J.; Koetzle, T. F.; Klooster, W. T.; Martin, J. M. L. *J. Phys. Chem. A* **1999**, *103*, 744.