

Performance of the SCC-DFTB Model for Description of Five-Membered Ring Carbohydrate Conformations: Comparison to Force Fields, High-Level Electronic Structure Methods, and Experiment

Shahidul M. Islam and Pierre-Nicholas Roy*

Department of Chemistry, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Supporting Information

ABSTRACT: The solution conformations of biologically important mono- and di- α -D-arabinofuranosides were investigated using the dispersion-corrected self-consistent charge density functional tight binding (SCC-DFTB) and the AMBER/GLYCAM models. Simulations were performed using both long dynamics and umbrella sampling simulations. Angular distributions about the exocyclic C–C bonds and puckering distributions of the rings obtained from the SCC-DFTB model were quite different from those obtained with the AMBER/GLYCAM approach. The joint probability distribution of rotamer and ring puckering parameters reveals further discrepancies, while both methods predict weak correlation between exocyclic torsions and ring puckering. To assess the reliability of the simulations, ensemble-averaged vicinal proton–proton coupling constants ($^3J_{H,H}$) were calculated and compared directly to experimental NMR coupling constants. It is found that the $^3J_{H,H}$ values obtained from the AMBER/GLYCAM simulations agree with experiment, while those obtained from the SCC-DFTB method, in most cases, differ from experimental $^3J_{H,H}$ values. Potential energy surfaces (PES) along the exocyclic torsion obtained from ab initio and DFT calculations differ significantly from those obtained with the SCC-DFTB method. This study establishes that a high-quality all-atom force field would be more suitable than one of the superior semiempirical methods, SCC-DFTB, for investigation of rotamer populations and ring puckering in floppy ring systems such as furanosides.

INTRODUCTION

Furanosides play an important role in the structure and dynamics of numerous glycoconjugates, nucleic acids, and antibiotics. Furanoside rings are highly flexible compared to their pyranoside counterpart in which bond lengths, angles, and dihedral angles can be maintained close to their minimum energy chair conformation under normal conditions. It is very important to identify the best theoretical method that is able to detect the most probable conformation of furanosides in both their monomeric and their oligomeric states and also when they are bound to a protein or are a structural component in DNA/RNA. Most well-known examples of furanose conjugates are the nucleic acids, which contain either D-ribofuranose or 2-deoxy-D-erythro-pentofuranose (2-deoxy-D-ribose).¹ Furanosides are also involved in the mechanism of various infections and diseases such as tuberculosis, leprosy, and *C. difficile*. The bacteria that cause these infections and diseases biosynthesize furanose-containing polysaccharides which are typically found on the cell surface and thus believed to play important roles in the survival of the organisms and their interaction with the environment.^{2–4} Therefore, a better understanding of the conformation of furanosides is important to develop new and better drugs.

Experimental and computational methods have been used to study the conformational preference of furanosides.^{5–16} These included the analysis of vicinal coupling constants and ab initio and density functional theory to study the conformational preferences of methyl- α -D-arabinofuranoside, methyl- β -D-arabinofuranoside, and other analogs. In solution, furanosides can adopt various twist and envelope conformations contrary to their pyranoside counterpart, which typically adopt single well-

defined chair conformations. The activation energy required to interchange between various conformations in pyranosides is higher than those in furanosides.^{9,17–20} Using an ab initio metadynamics method, Parrinello et al. studied the conformational free energy landscape of β -D-glucopyranose as a function of the puckering coordinates.¹⁸ The 4C_1 chair conformation was found as the most stable conformation and the system reached the second most stable conformation, $B_{3,0}$, by overcoming a free energy barrier of about 8 kcal/mol. On the contrary, Houseknecht et al. showed, using high-level ab initio and DFT calculations,⁹ that different conformations of furanosides can be visited with low barriers of only a few kcal/mol. The conformations of various arabinofuranosides were recently studied with long molecular dynamics simulations using the force field models.^{12–15} From those studies it became apparent that the results obtained from the force field models needed to be validated by an alternative method. Some conformational properties such as the ring puckering obtained from force field models did not quite agree with experiment. It was argued that the two-state model used to calculate the experimental ring puckering parameters might not account for the actual phenomenology of ring dynamics. In the present study, the conformational preference of furanose systems is investigated with a semiempirical method, the so-called self-consistent charge-density functional tight binding (SCC-DFTB),²¹ that emerged as a relatively reliable method in the past few years. SCC-DFTB is not only reasonably accurate but also quite efficient with a speed comparable to other semiempirical

Received: November 5, 2011

Published: June 4, 2012



methods such as AM1²² and PM3.^{23,24} SCC-DFTB is also proposed to reproduce the sugar conformation along with other biological systems.^{25–28} The present study has been motivated by that proposal, and we became interested to see how SCC-DFTB calculations perform on floppy carbohydrates such as the furanosides. It will be interesting to compare the results obtained from SCC-DFTB to those obtained from the AMBER/GLCYAM^{29–31} approach.

The self-consistent charge density functional tight binding (SCC-DFTB) is an approximate density functional theory (DFT), which represents the DFT total energy with a second-order expansion in terms of the charge density variation with respect to a reference density. It represents an improvement of the original density functional tight binding (DFTB)^{21,32} method because of a self-consistent procedure, which iteratively relaxes the atomic Mulliken charges. The total energy of the system in SCC-DFTB is given by

$$E^{\text{SCC-DFTB}} = \sum_i \langle \phi_i | \hat{H}_0 | \phi_i \rangle + \frac{1}{2} \sum_{A,B} \gamma_{AB} \Delta q_A \Delta q_B + E_{\text{rep}} \quad (1)$$

where ϕ_i are the Kohn–Sham orbitals, \hat{H}_0 is the effective Kohn–Sham Hamiltonian, and Δq_A and Δq_B are the Mulliken charge fluctuation of atoms A and B relative to the number of valence electrons in their neutral state. The parameter γ_{AB} approximates the second derivative with respect to charge fluctuation, and E_{rep} is the pairwise repulsive potential. The latter is fitted to high-level DFT results for reference systems. The computational cost of the SCC-DFTB model is similar to widely used semiempirical methods such as AM1 and PM3. It is about 2–3 orders of magnitude faster than HF and DFT methods with small to medium basis sets. The SCC-DFTB method has been extensively tested and applied successfully to many chemical and biochemical systems with impressive results.^{25,26,33–46} SCC-DFTB is found to better represent geometries and energetics than the widely used semiempirical AM1 and PM3 approaches for many molecular systems. Extensive comparison of molecular geometries and energetics has shown that the SCC-DFTB method typically reproduces the geometries and energetics obtained from DFT calculation with small to moderate basis sets.^{44–46} SCC-DFTB is currently extensively used to study enzymatic reactions within the QM/MM approach.^{35–37,47–51} Overall, the SCC-DFTB method has been shown to provide reliable information regarding molecular geometries,^{33,52} vibrational frequencies,⁵³ and reaction energies^{33,34,52} of many enzymatic processes.

The present study has been carried out to better understand the conformational preferences of arabinofuranosides described with the SCC-DFTB and AMBER/GLYCAMS models and relies on techniques that provide enhanced sampling of conformational space. The furanosides used here are presented in Figure 1.

Given their inherent flexibility, the conformational analysis of furanosides is more complicated than comparable studies with pyranosides. In solution, furanoside rings equilibrate or interconvert between two low-energy intermediates. However, they do not equilibrate through formation of a planar intermediate because such a structure is too high in energy. Rather, they interconvert by a process known as pseudorotation. The ring proceeds along this pseudorotational itinerary to go from one conformation to another, adopting a number of envelope and twist conformations along the way. Those are

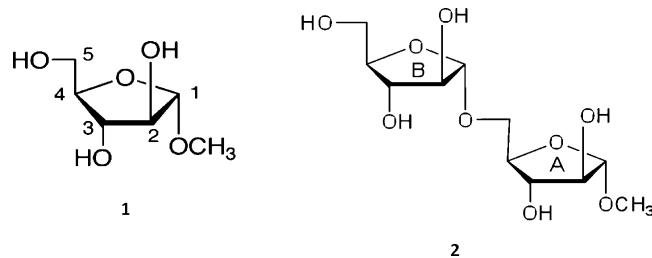


Figure 1. Arabinofuranosides studied in this work.

separated by low-energy barriers. Figure 2 shows the pseudorotational itinerary for the D-arabinofuranose ring.

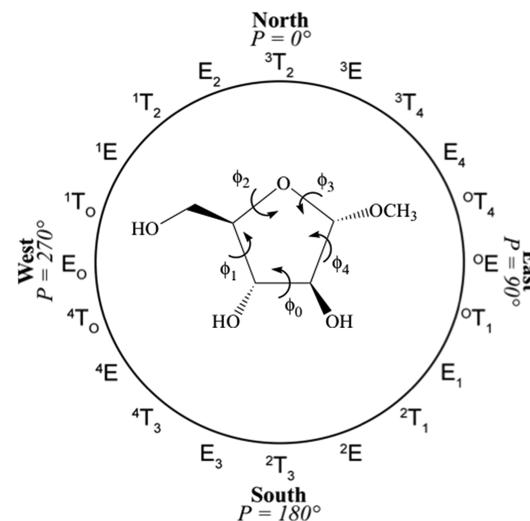


Figure 2. Pseudorotational itinerary for D-arabinofuranosides.

The conformers that belong to the northern hemisphere of the pseudorotational wheel are termed northern conformers, and those that belong to the southern hemisphere of the pseudorotational wheel are termed southern conformers. According to the Altona–Sundaralingam notation,^{54,55} each furanose ring conformation can be described by a phase angle of pseudorotation (P) and a puckering amplitude (ϕ_m). The P angle of a given conformer can be calculated from the five endocyclic torsion angles ϕ_0 – ϕ_4 as

$$\tan P = \frac{(\phi_2 + \phi_4) - (\phi_1 + \phi_3)}{3.077\phi_0} \quad (2)$$

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 below

$$\phi_m = \frac{\phi_0}{\cos P} \quad (3)$$

Other important features that should be considered during conformational analysis of furanosides are rotamer populations about the exocyclic C–C and C–O bonds. In the present study, we investigated rotamer populations about the C–C bond, which are influenced by a combination of steric and stereoelectronic (gauche) effects.^{56–60} Figure 3 shows the three staggered rotamers about the C4–C5 bond in α -D-arabinofuranose rings.

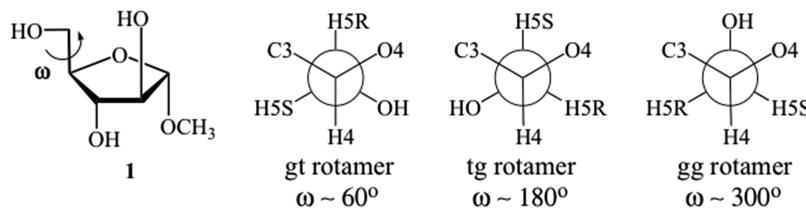


Figure 3. Definitions of the three staggered rotamers about the C4–C5 bond in 1. Angle ω is defined as the O5–C5–C4–O4 torsion angle.

From our previous studies it became apparent that convergence of rotamer population about the exocyclic torsion of monoarabinofuranoside fragments was achieved after several hundred nanoseconds of conventional MD simulation. Therefore, treatment of larger oligomers of arabinofuranosides with conventional long MD is impractical due to the limit of current computational power. Moreover, SCC-DFTB calculations are generally more expensive than force field methods and, therefore, would require more computational time to achieve convergence of rotamer populations. Considering all these factors it became important to investigate the conformation of furanosides with special sampling techniques so that less computation would provide reliable results. In this study we utilized umbrella sampling simulation techniques to obtain rotamer population along the exocyclic torsion and compared the results obtained from such simulations to those obtained from long MD simulations for both SCC-DFTB and AMBER/GLYCAM. An effective sampling of all the configurational space along the exocyclic torsion is possible by the umbrella sampling technique.⁶¹ The results obtained from such simulations should be more reliable than the long conventional MD simulation and would therefore be more effective during comparison of SCC-DFTB and AMBER/GLYCAM.

METHODS

MD Simulation Protocol. Both the SCCC-DFTB model²¹ and the AMBER ff99SB force field²⁹ with the GLYCAM (version 04f) parameter set for carbohydrates^{30,31} were employed for description of the arabinofuranosides. We chose this particular version of the GLYCAM parameter set for consistency with our previous simulations on furanosides.^{12–16} Note that we tested a recently developed GLYCAM06 parameter set⁶² for α -D-arabinofuranoside (**1**), and the results were found to be very similar to those obtained with the present force field. The additivity principle^{63,64} was used to build the topology of the disaccharide in AMBER/GLYCAM. Once the furanosyl system was constructed it was solvated by a pre-equilibrated sphere of TIP3P waters.⁶⁵ All simulations were performed with the AMBER10 suite of molecular simulation codes.⁶⁶ First, the water molecules were minimized keeping the geometries of the arabinofuranosyl molecules constrained. Then the whole system was minimized. In both minimization steps a steepest decent energy minimization was carried out for 50 cycles to remove bad contacts. The conjugate gradient algorithm was then used for 950 cycles. A total of 100 ps annealing was then carried out with 50 ps each for temperature heating and cooling. This was then followed by a short equilibration run of 250 ps. During this run the temperature was slowly increased from 0 K to room temperature (300 K), and the system was allowed to equilibrate for 100 ps. Finally, a production 200 ns long MD simulation was performed on monosaccharide **1**. To enhance the sampling of the configurational space, umbrella sampling simulations of both **1** and **2** and 1.5 ns simulations for each window were carried out to

were also performed. The umbrella sampling simulation protocol is discussed in the following sections. 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. All simulations were performed with the Langevin⁶⁷ thermostat to control the temperature of the simulation box during simulation. A collision frequency γ of 2.0 ps⁻¹ was used for the Langevin thermostat. The 1–4 electrostatic and van der Waals interactions in the AMBER/GLYCAM simulations were scaled by the recommended values for carbohydrates (SCEE = 1.0, SCNB = 1.0). In all 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 SHAKE⁶⁸ 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.^{69,70} SCC-DFTB/TIP3P electrostatic interactions were also controlled by PME with a cutoff of 8 Å.

Calculation of rotamer population and ring puckering 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,^{71,72} 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 in order to enhance the sampling of certain regions of configuration space. A harmonic form is often chosen for the biasing potential and defined as

$$V_b(\chi)_i = \frac{1}{2}k(\chi - \chi_i)^2 \quad (4)$$

where k is the force constant and χ_i is the target position. The separate simulations are then combined to obtain the unbiased distribution, $\rho(\chi)$. The free energy change or PMF along the P can be calculated by integrating the puckering distributions and using the equation^{73,74}

$$W(\chi) = -k_B T \ln(\rho(\chi)) \quad (5)$$

Among the various approaches to combine the simulation results,^{71,72,75–77} the weighted histogram analysis method (WHAM) proposed by Kumar et al.⁷⁸ uses all the information present in the umbrella sampling without discarding the overlapping regions. The WHAM method is a practical approach to obtain average $\rho(\chi)$ and the PMF.

In this study, dihedral angles about the exocyclic C–C bond were taken as the reaction coordinate χ . All umbrella sampling simulations were performed using the final structure obtained from a short equilibration run (250 ps) of the systems. A total of 36 windows with a window width of 10° were used to cover the entire dihedral angle range from 0° to 360°. For **1** and **2**, 3 and 1.5 ns simulations for each window were carried out to

yield a total simulation time of 108 and 54 ns, respectively. 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. All other simulation parameters used in the umbrella sampling simulations were consistent with the long MD simulations discussed above. Once the simulations were completed, 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.⁷⁹ The bootstrap error analysis method^{79,80} was utilized to obtain relative errors of the distributions of rotamer population.

Ring puckering distributions were readily obtained from the long MD simulations and the combined umbrella sampling simulation trajectories. Note that it has been previously found by both experiment and computational study that the ring puckering occurs on a much faster time scale than the exocyclic torsion.^{9,81} To study the correlation of rotamers and ring puckering, a joint probability distribution was also constructed from analysis of umbrella sampling simulations. For each window trajectory, values of P were obtained. The 2D-WHAM program was then used to construct the joint probability distribution and 2D-potential of mean force. Biased dihedral values were unbiased using the same force constant, 30 kcal/mol, as was used to restrain the dihedrals in the umbrella sampling. Since no biasing was applied on P , a force constant of zero was used during the 2D-WHAM analysis.

Ab Initio, DFT, AM1, PM3, and PM6 Calculations. To verify the results obtained from the SCC-DFTB and AMBER/GLYCAM approaches outlined above, high-level ab initio and DFT potential energy surface scans along the exocyclic torsion angle ω were also performed. All of the electronic structure calculations were carried out with the Gaussian 09 software package.⁸² The potential energy surface (PES) scans were performed with the MP2 and B3LYP levels of theory using the 6-31+G(d,p) basis set. In addition to gas-phase calculations, a potential energy surface scan in H₂O was also performed with the B3LYP/6-31+G(d,p) level of theory with the conductor polarizable continuum model (CPCM)^{83,84} as implemented in Gaussian 09. To assess the effects of level of theory and basis sets, CCSD(T)/6-31+G(d,p) and B3LYP/aug-cc-pVTZ single-point energies were also calculated for the three lowest energy minima and three highest energy maxima obtained from the potential energy surface scan with the B3LYP/6-31+G(d,p) level of theory. In addition to gas-phase calculations, a potential energy surface scan in H₂O was also performed with the B3LYP/6-31+G(d,p) level of theory with the conductor polarizable continuum model (CPCM)^{83,84} as implemented in Gaussian 09. By default, the CPCM model builds up the cavity using the UFF model, i.e., putting a sphere around each atom including the hydrogens. The PES scans were also performed with several widely used semiempirical methods, AM1,²² PM3,^{23,24} and PM6,⁸⁵ including SCC-DFTB.²¹

RESULTS AND DISCUSSION

Rotamer Populations. In a first step, the length of the umbrella sampling and conventional MD simulations required to achieve convergence of the rotamer population was determined (see Supporting Information). For SCC-DFTB and AMBER/GLYCAM, 500 ps umbrella sampling simulations for each window and 200 ns conventional long MD simulations were required to obtain a converged rotamer populations of all rotamers with uncertainties of less than a few percent units. This study illustrates that a total of 18 ns of umbrella sampling

simulations is capable of providing reliable rotamer populations, which is significantly lower than the conventional 200 ns MD simulation.

The rotamer population distribution about the C4–C5 bond of **1** obtained from the SCC-DFTB and AMBER/GLYCAM approaches is shown in Figure 4. Note that very similar rotamer

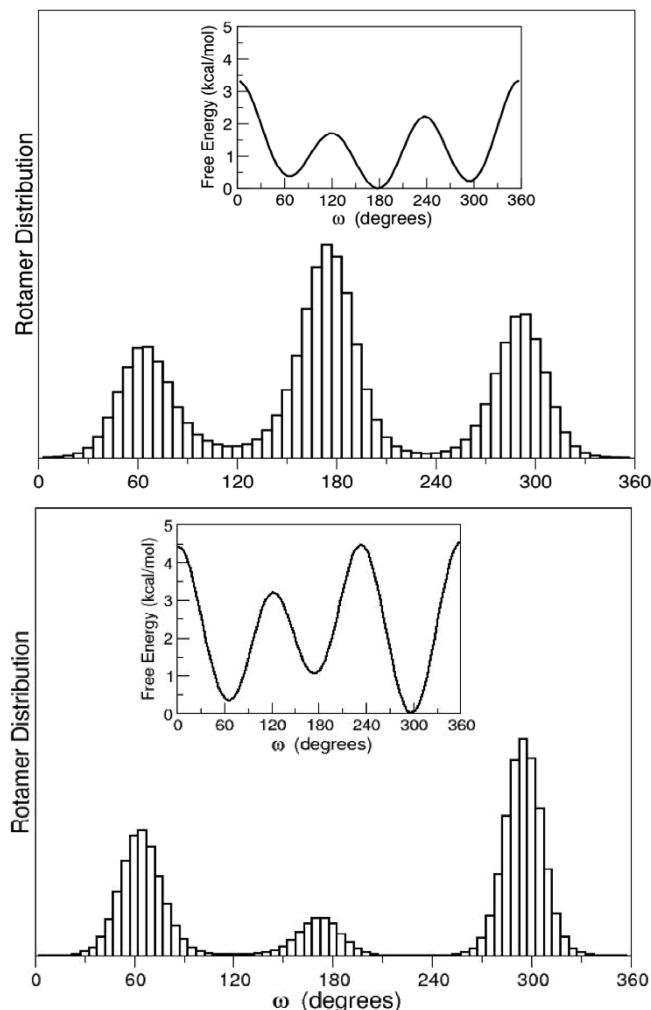


Figure 4. Rotamer population distribution of **1** in solution obtained from SCC-DFTB (top) and AMBER/GLYCAM (bottom) using the umbrella sampling simulation. (Inset) Corresponding PMF along the dihedral angle (ω).

populations are obtained with the long MD and umbrella sampling simulations. It is clear from Figure 4 that the rotamer population distribution obtained from the SCC-DFTB approach is very different from that of AMBER/GLYCAM. The rotamer populations decrease from tg > gg > gt with SCC-DFTB, while the trend is gg > gt > tg with AMBER/GLYCAM. The change in free energy along the C4–C5 bond obtained from the umbrella sampling simulation provides information on the relative stability of C4–C5 rotamers. It is clear from the free energy surface, as shown in the inset of Figure 4, that all three rotamers are populated with SCC-DFTB since the free energy barriers to go from the tg rotamer (180°), which is the most stable rotamer, to the gt (60°) and gg (300°) rotamers are only about 1.7 and 2.3 kcal/mol, respectively. On the other hand, with the AMBER/GLYCAM model, the free energy barriers from the most stable rotamer (gg) to the tg and gt

rotamers are about 4.4 and 4.5 kcal/mol, respectively. Overall, free energy barriers to go to different rotameric states are higher with the AMBER/GLYCAM model than with the SCC-DFTB description. Note that rotamer populations and ring puckering distributions for **1** obtained with and without dispersion correction to the SCC-DFTB energy are found to be very similar.

The rotamer population distribution about the glycosidic C–C bonds involved in the α -(1 \rightarrow 5) linkage for the rings of **2** obtained from SCC-DFTB are shown in Figure 5. The

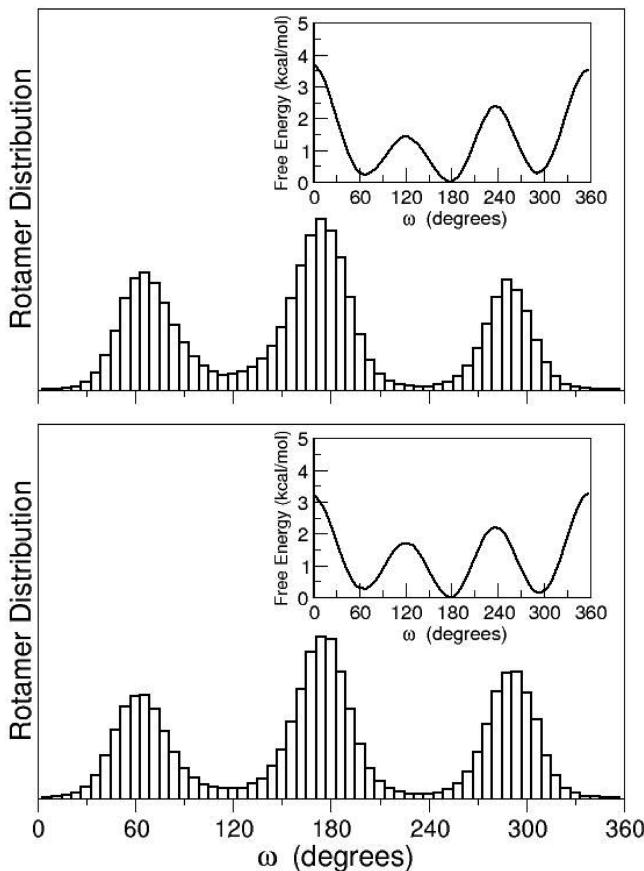


Figure 5. Rotamer population distribution of the internal (top) and external (bottom) dihedral angles of **2** in solution obtained from SCC-DFTB using the umbrella sampling simulation. (Insets) Corresponding PMF along the dihedral angles (ω).

histograms show that the rotamer populations of the internal dihedral angle of **2** have a trend of tg > gt > gg, while the terminal rotamer populations follow a trend very similar to **1** with tg > gg > gt. As in the case of **1**, the free energy barriers to go from the tg rotamer to the gt and gg rotamers are very small, ranging from 1.3 to 2.5 kcal/mol. Although the trend in rotamer population is different in internal and external rotamers, it is likely that all rotameric states can be visited with almost equal probability. The trend in rotamer populations and free energy barriers of **2** obtained from AMBER/GLYCAM is also very different than those calculated with the SCC-DFTB model.

It is possible to quantify rotamer populations of **1** and **2** by integrating the distributions obtained from the SCC-DFTB and AMBER/GLYCAM simulations. Such rotamer populations about the torsion ω of **1** and **2** are given in Table 1. Rotamer populations (in percent) obtained from SCC-DFTB decrease

Table 1. Rotamer Population (%) of **1** and **2** in Solution Obtained from SCC-DFTB and AMBER/GLYCAM Force Field^a

	SCC-DFTB			AMBER/GLYCAM		
	2			2		
	1	A	B	1	A	B
X _{tg}	25(2)	31(2)	27(1)	35(2)	14(1)	24(1)
X _g	46(2)	45(3)	42(3)	11(1)	15(1)	8(1)
X _{gg}	29(2)	24(1)	31(2)	54(2)	72(3)	67(3)

^aValues in parentheses are standard errors obtained from the Bootstrap error analysis of the distributions obtained from the umbrella sampling simulation about the dihedral angle ω .

from $X_{tg} > X_{gg} > X_{gt}$ for **1** with values of 46%, 29%, and 25%, respectively; while rotamer populations obtained from AMBER/GLYCAM approach decrease from $X_{gg} > X_{gt} > X_{tg}$ for **1** with values of 54%, 35%, and 11%, respectively. From Table 1, it is also evident that the rotamer populations of the internal dihedrals of **2** obtained from SCC-DFTB decrease from $X_{tg} > X_{gt} > X_{gg}$ while the external rotamer population decrease from $X_{tg} > X_{gg} > X_{gt}$. On the contrary, the trend in rotamer populations for internal and external rotamers decreases from $X_{gg} > X_{gt} \approx X_{tg}$ and $X_{gg} > X_{gt} > X_{tg}$, respectively, with the AMBER/GLYCAM approach. Our study clearly shows that the rotamer populations and free energy barriers obtained from the SCC-DFTB and AMBER/GLYCAM approaches are very different for both mono- and disaccharides.

Distribution of Ring Conformers. Having determined the rotamer populations along the exocyclic C–C bond, ring conformations of **1** and **2** have been studied. In addition to long MD simulation, the umbrella sampling simulation trajectories obtained from the restrained exocyclic torsions were utilized for calculation of ring conformations of **1**. All umbrella sampling simulations corresponding to a ring were combined, and then the distributions of puckering angles P and amplitudes ϕ_m were calculated. Since ring puckering occurs on a much faster time scale than the exocyclic torsion, the above approach to calculate the P and ϕ_m distributions from umbrella sampling trajectories is justifiable. To confirm that the ring puckering truly occurs faster than the exocyclic torsion for furanosides, ring puckering distributions were calculated for each of the umbrella sampling simulation trajectories with SCC-DFTB and AMBER/GLYCAM. Figure 6 represents the ring puckering distributions obtained from three different umbrella sampling simulations restrained at 0°, 120°, and 240° with SCC-DFTB. From Figure 6 it is clear that the distribution of P in **1** is very similar for each window. AMBER/GLYCAM also provides very similar distributions for each umbrella sampling trajectory.

The distributions of P for **1** obtained using the SCC-DFTB and AMBER/GLYCAM models are shown in Figure 7. It has been found that a broad ring puckering distribution is obtained with SCC-DFTB, while two distinct regions in the ring puckering distribution are observed with the AMBER/GLYCAM approach. With SCC-DFTB, the conformers centered at about $P = 80^\circ$ predominate over all other values of P . On the other hand, with the AMBER/GLYCAM approach, conformations in the northern hemisphere of the pseudorotational itinerary are clearly favored, although a small fraction of the conformers is also present in the southern hemisphere. Free energy profiles with respect to P (presented in the insets of Figure 7) clearly show that the global minima are at the northern hemisphere of the pseudorotational

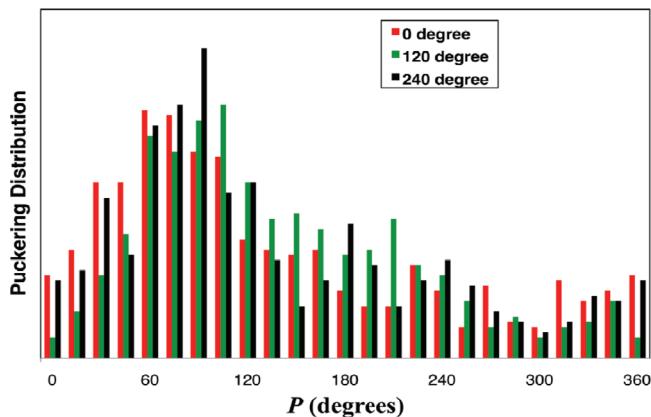


Figure 6. Puckering distribution obtained from SCC-DFTB using three different umbrella sampling simulations restrained at 0°, 120°, and 240°.

AMBER/GLYCAM correspond well with previous ab initio and density functional theory calculations.^{6–10} The ab initio and DFT calculations of **1** revealed a global minimum at the southern hemisphere of the pseudorotational itinerary. Comparison of the free energy profiles along P obtained from the SCC-DFTB and AMBER/GLYCAM methods reveal that the maximum free energy barrier is only 1.3 kcal/mol with SCC-DFTB, while the maximum free energy barrier is about 2.5 kcal/mol with AMBER/GLYCAM. Moreover, two distinct minima centered at $P = 40^\circ$ and 180° are visible from the AMBER/GLYCAM approach, while with SCC-DFTB only one minimum is noticeable at $P = 80^\circ$. The distribution in puckering amplitudes, ϕ_m , obtained from both SCC-DFTB and AMBER/GLYCAM are found to be very similar, as shown in Figure 8. ϕ_m is centered about 35° in **1**, which agrees very well with those obtained from the crystal structure⁸⁶ and from ab initio and DFT calculations.^{6–10}

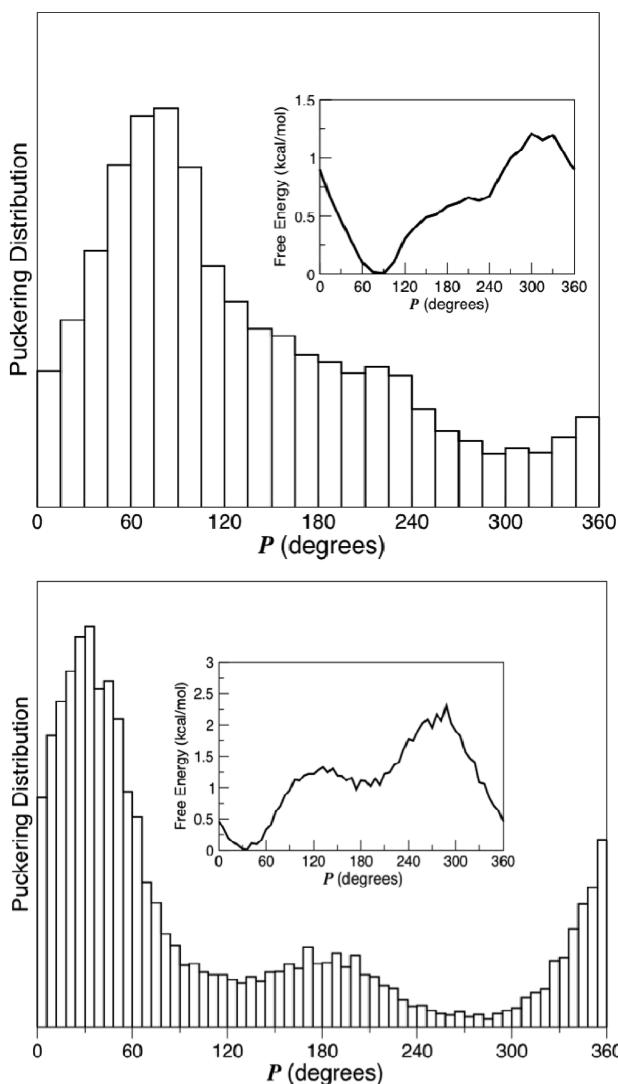


Figure 7. Distribution of the pseudorotational phase angle (P) for **1** in solution obtained from SCC-DFTB (top) and AMBER/GLYCAM (bottom) using the umbrella sampling simulations. (Insets) Corresponding PMF along the P .

itinerary for both SCC-DFTB and AMBER/GLYCAM. The global minima obtained from neither SCC-DFTB nor

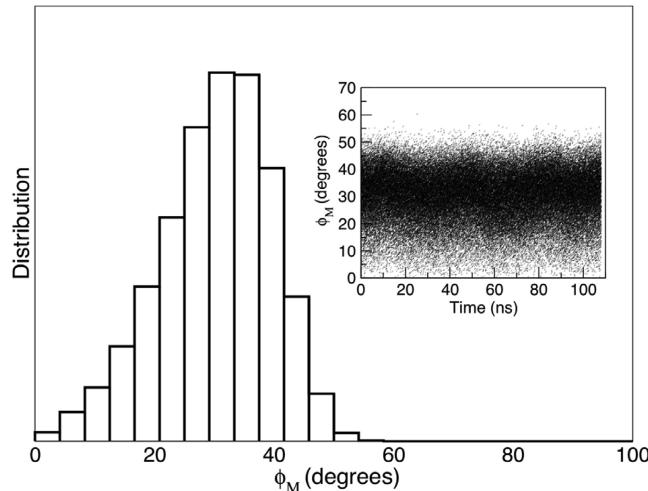


Figure 8. Distribution of the puckering amplitude (ϕ_m) for **1** in solution obtained from SCC-DFTB. (Inset) Time dependence of the ϕ_m angle.

Ring puckering distributions of disaccharide **2** obtained using the SCC-DFTB model also differ greatly from the AMBER/GLYCAM results. The distributions are very similar to those of **1** (shown in Figure 7). A broad distribution of the P angle is found for both ring A and ring B of **2** with the SCC-DFTB model, while two distinct and narrow regions of the P angle predominate in the AMBER/GLYCAM description. Both SCC-DFTB and AMBER/GLYCAM were unable to distinguish between the two rings in **2**. The distribution in ϕ_m is found to be very similar with SCC-DFTB and AMBER/GLYCAM for both ring A and ring B of **2**. ϕ_m for ring A and ring B are maximum at 32° with both SCC-DFTB and AMBER/GLYCAM.

Correlation Study of Rotamers and Ring Puckering.

Having successfully determined the rotamer populations and ring puckering with SCC-DFTB and AMBER/GLYCAM, we 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** obtained from SCC-DFTB and AMBER/GLYCAM are shown in Figure 9. It is clear that $\rho_{\omega,P}^{2D}$ obtained from the SCC-DFTB and AMBER/GLYCAM methods are very different, with the tg rotamer dominating over most of the values of P at SCC-DFTB while the gg rotamer is found to be

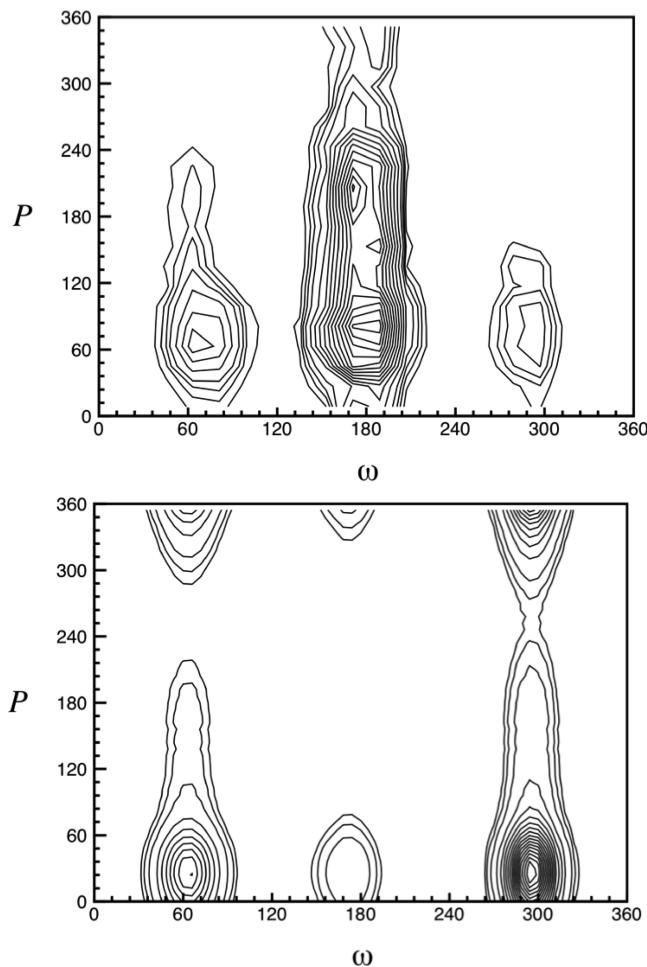


Figure 9. Joint probability distribution of the dihedral angle (ω) and puckering angle (P) of **1** in solution obtained from SCC-DFTB (top) and AMBER/GLYCAM (bottom). Units of the angles ω and P are degrees.

predominant over most of the values of P with the AMBER/GLYCAM approach. Very little correlation between ω and P for **1** is observed with both methods. Figure 9 provides a qualitative indication of the uncorrelated behavior of ω and P , but the extent of their correlation is absent. Therefore, we looked at the quantitative measure of this correlation. If there is a correlation between ω and P , the following equation represents the magnitude of that correlation (see Supporting Information for the derivation of this equation)

$$\chi^2 = 2((1 - \langle fg \rangle)^2) \quad (6)$$

where f represents the square root or amplitude of the true 2D distribution and g denotes the square root of the products of 1D distributions of each degree of freedom. The value of χ^2 can range from 0 to 2. The smaller the value of χ^2 , the less will be the correlation, and the larger the value of χ^2 , the larger will be the correlation between ω and P . χ^2 was found to be only 0.03 and 0.02 with the SCC-DFTB and AMBER/GLYCAM, respectively, suggesting almost no correlation between ω and P with the two methods. The 2D-PMF calculated from 2D-WHAM analysis clearly shows that 3 kcal/mol free energy is adequate to explore all values of ω in the range from 0° and 360° , while only 1.5 kcal/mol free energy is necessary to explore all values of P with SCC-DFTB. Conversely, energies of

5 and 3 kcal/mol are, respectively, required to explore all values of ω and P with AMBER/GLYCAM. The correlation study of **2** with SCC-DFTB and AMBER/GLYCAM also shows no correlation between ω and P ; however, the joint probability distributions obtained from the SCC-DFTB method are again found to be different from the AMBER/GLYCAM ones.

Comparison with Experiment. Since the rotamer populations and ring puckering distribution obtained from SCC-DFTB and AMBER/GLYCAM were found to be very different, it is important to identify which method best reproduces the conformation of furanosides. For several decades it has been customary to obtain experimental rotamer and puckering distributions from analysis of NMR vicinal coupling constants, ${}^3J_{H,H}$. Therefore, an appropriate approach to assess the reliability of the simulation results is to compare the theoretically obtained average coupling constants, $\langle J \rangle$, directly with the experimental NMR coupling constants.¹⁶ The $\langle J \rangle$ value is measured as an average over the entire conformational space of the molecule using the relation⁸⁷

$$\langle J \rangle = \int_0^{360} J(\phi) \rho(\phi) d\phi \quad (7)$$

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. The following five DFT-derived Karplus relationships corresponding to five three-bond H,H coupling pathways in α -D-arabinofuranosyl molecule ((H1–C1–C2–H2, H2–C2–C3–H3, H3–C3–C4–H4, H4–C4–C5–HSR, H4–C4–C5–HSS) were used to correlate the spin–spin coupling constants to the dihedral angles¹⁴

$${}^3J_{1,2} = 4.62 + 3.16 \cos(\phi) + 4.57 \cos(2\phi) \quad (8)$$

$${}^3J_{2,3} = 8.04 + 8.07 \cos(\phi) + 7.24 \cos(2\phi) \quad (9)$$

$${}^3J_{3,4} = 4.44 + 0.50 \cos(\phi) + 4.25 \cos(2\phi) \quad (10)$$

$${}^3J_{4,SS} = 4.95 - 0.42 \cos(\phi) + 4.03 \cos(2\phi) \quad (11)$$

$${}^3J_{4,SR} = 5.23 + 0.02 \cos(\phi + 15.1^\circ) + 4.67 \cos(2\phi + 30.2^\circ) \quad (12)$$

The $\rho(\phi)$ distribution can be obtained from the long or umbrella sampling simulations performed with the SCC-DFTB and AMBER/GLYCAM models. The procedure utilized to obtain $\rho(\phi)$ from umbrella sampling simulation data is discussed in one of our recent papers⁸¹ (relevant section is provided in the Supporting Information). The basic assumption is that the rotamer population distribution along O4–C4–C5–O5 can be used to obtain the population distribution along the H4–C4–C5–HSR and H4–C4–C5–HSS dihedrals, which are then used to calculate the vicinal proton–proton coupling constants. The dihedral angles between the coupled spins in the ring, used to calculate the ring proton–proton coupling constants, are obtained from the combined umbrella sampling trajectories. All distributions along the proton–proton dihedrals obtained from the SCC-DFTB model are found to be very different from those obtained with the AMBER/GLYCAM model (see Supporting Information for the distributions). Table 2 shows the comparison of experimental

Table 2. Comparison of Experimental and Theoretical $\langle^3J_{H,H} \rangle$ Values (in Hertz) for **1 and **2** in Solution Obtained from SCC-DFTB and AMBER/GLYCAM^a**

	1			2					
				ring A			ring B		
	SCC-DFTB	AMBER/GLYCAM	EXP	SCC-DFTB	AMBER/GLYCAM	EXP	SCC-DFTB	AMBER/GLYCAM	EXP
$\langle^3J_{H_4,HSR} \rangle$	5.6	4.5	5.8	5.4	3.4	5.5	5.4	3.7	5.5
$\langle^3J_{H_4,HSS} \rangle$	5.3	3.2	3.3	5.1	2.9	3.4	4.9	3.1	3.4
$\langle^3J_{H_1,H_2} \rangle$	0.9	2.1	1.7	0.8	2.4	1.6	0.9	2.5	1.6
$\langle^3J_{H_2,H_3} \rangle$	1.6	3.6	3.4	1.6	3.8	3.1	1.7	3.8	3.1
$\langle^3J_{H_3,H_4} \rangle$	3.8	5.4	5.8	4.1	5.5	5.8	3.9	5.6	6.1

^aEXP = experimental values. Errors in theoretical coupling constants are very small within the range of 0.02–0.05 Hz. (See the Supporting Information for errors in $^3J_{H,H}$ values.)

and theoretical coupling constants for **1** and **2** obtained from the SCC-DFTB and AMBER/GLYCAM methods. Note that both umbrella and long MD simulations show similar coupling constant values. The coupling constants obtained from AMBER/GLYCAM agree better with the experimental coupling constants except in the case of $\langle^3J_{H_4,HSR} \rangle$, where SCC-DFTB performs better. Note that we also calculated coupling constants with a generalized Karplus relationship.¹³ These coupling constants are not in better agreement with experiment than those obtained from the DFT-based Karplus relationships. However, comparison of AMBER/GLYCAM and SCC-DFTB coupling constants obtained from the generalized Karplus relationship reveals better agreement of AMBER/GLYCAM coupling constants to those obtained from experiment (see Table S3 in the Supporting Information). This clearly suggests that SCC-DFTB may not be able to yield correct rotamer populations and ring puckering for arabinofuranosides. A well-developed force field such as AMBER/GLYCAM provides better results. An alternative and a better approach would be to use a high-level theory such as ab initio or DFT on the fly. However, the size of the system and computational time limit such simulations. We nevertheless explore below important structural and energetic features in the context of higher level methods via the use of potential energy surface (PES) scans.

Potential Energy Surface Scans along Exocyclic Torsions.

The rotamer and ring puckering distributions obtained from AMBER/GLYCAM are very different from those with the SCC-DFTB model. We also observed that the coupling constants obtained from the SCC-DFTB approach did not agree well with experiment. The SCC-DFTB energetic properties along the exocyclic torsions are now compared to high-level ab initio and DFT calculations. PES scans along the C4–C5 bond of **1** are performed using the SCC-DFTB, B3LYP/6-31+G(d,p), and MP2/6-31+G(d,p) levels of theory. The geometries obtained from B3LYP/6-31+G(d,p) were used to calculate single-point energies at the CCSD(T)/6-31+G(d,p) level of theory. PES scans were performed by allowing the torsional dihedral angle to vary by 10° increments, while most of the other structural parameters were allowed to relax. A first PES scan along ω was performed by freezing C4–C5–O5–H5 and all ring dihedrals. It was necessary to freeze the C4–C5–O5–H5 dihedral angle to avoid formation of extraneous minimum energy structures due to H bonding and therefore allow a full PES scan in the range from 0° to 360°. The corresponding PES is shown in Figure 10. All higher levels of theory provide similar PESs with the energies of the rotamers decreasing from tg (~180°) > gg (~300°) > gt (~60°).

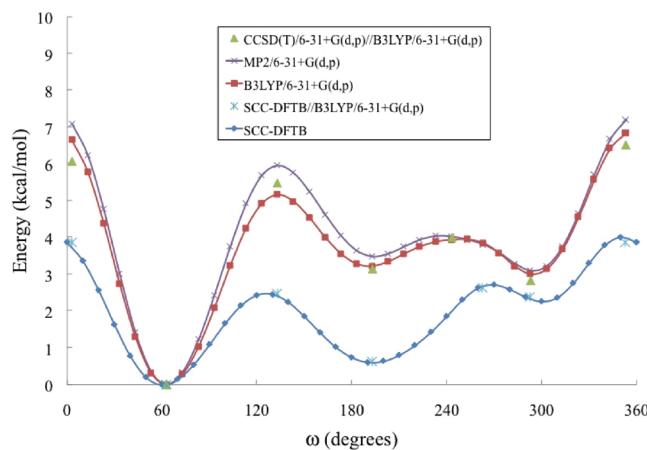


Figure 10. Potential energy along ω in **1** by freezing C4–C5–O5–H5 and all ring dihedrals (levels of theory are shown in the inset).

However, the PES obtained from the SCC-DFTB model is very different from the energies of the rotamers decreasing from gg (~300°) > tg (~180°) > gt (~60°). It is especially important to note here that the geometries are stabilized more at around 180°, which corresponds to the tg rotamer, with the SCC-DFTB model compared to the higher level theories. Very similar PESs about the ω angle are also found by freezing the ring dihedral angles and dihedrals about the C–O bonds attached to the ring.

Freezing an extra degree of freedom, the C4–C3–O3–H3 dihedral angle, along with the dihedrals that were already frozen, as discussed above, provides a PES where the energy of the rotamers decreases from tg (~180°) > gt (~60°) > gg (~300°) with the high-level ab initio and DFT calculations, as shown in Figure 11. However, the PES obtained with the SCC-DFTB model is again found to be very different from the high-level theories, and the energy of the rotamers decreases from tg (~180°) > gt (~60°) > gg (~300°). With the addition of an extra constraint along the C4–C3–O3–H3 dihedral angle, the PES for all levels of theory changed from those observed in Figure 10 with the global minimum found at around 300° with all levels of theory. The geometries obtained from SCC-DFTB are stabilized more at around 180° when compared to those obtained from the high-level theories. Single-point calculations with CCSD(T)/6-31+G(d,p)//B3LYP/6-31+G(d,p) also provide energies similar to those of B3LYP/6-31+G(d,p) and MP2/6-31+G(d,p) levels of theory. To study the effect of geometry optimization on SCC-DFTB and higher level theories, single-point energy calculations were performed with SCC-DFTB using six minimum and maximum energy geo-

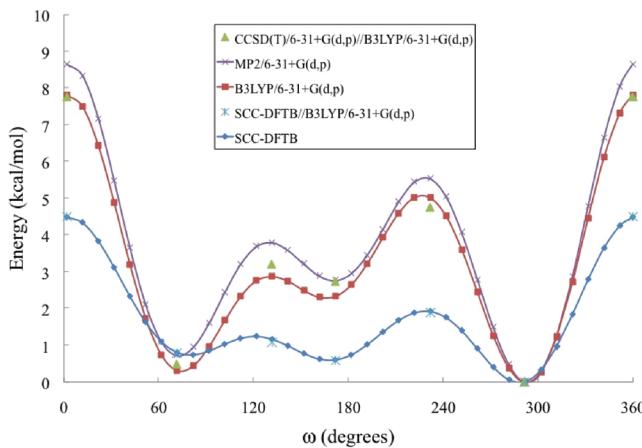


Figure 11. Potential energy along ω in **1** by freezing C4–C5–O5–H5, C4–C3–O3–H3, and all ring dihedrals as determined by the levels of theory shown in the caption.

metries obtained from B3LYP/6-31+G(d,p). These single-point energies are found to be almost identical to those obtained from the PES scan with the SCC-DFTB model (Figures 10 and 11). This study demonstrates that similar structures provide very different energetics with SCC-DFTB and higher level ab initio and DFT calculations.

The gas-phase PESs may be very different from those in solutions due to potentially large solvent effects. Full free energy surface scans along the exocyclic torsion were performed at the B3LYP/6-31+G(d,P) level of theory in water using the CPCM model.^{83,84} Note that in solution we could obtain the full potential energy surface without constraining any degrees of freedom. Vibrational frequencies were calculated for three minima and three maxima along the PES to ensure the absence of imaginary frequencies in the minima and the presence of only one imaginary frequency in the transition states. As expected, three of the minimum energy structures have no imaginary frequency and three of the highest energy maxima have one imaginary frequency. Free energies of the six structures calculated within the harmonic approximation at the B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ levels were compared to the free energy profile of AMBER/GLYCAM and SCC-DFTB obtained from the umbrella sampling simulations (Figure 12).

From Figure 12 it is clear that the PES obtained from the B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ levels are very similar, although there is a difference of 1 and 0.7 kcal/mol in the free energies at 130° and 180° , respectively. B3LYP and AMBER/GLYCAM were found to provide very similar trends with the highest energy conformation being the tg rotamer ($\sim 180^\circ$) while the gt ($\sim 60^\circ$) and gg ($\sim 300^\circ$) rotamers have very similar energies. However, the free energy obtained from B3LYP at 0° or 360° is about 3 kcal/mol higher than AMBER/GLYCAM. The implicit solvent model in B3LYP may not be able to capture some of the interactions with water at around 0° or 360° . However, the other five structures obtained at B3LYP have very similar energies when compared to AMBER/GLYCAM. On the other hand, the free energy profile obtained from the SCC-DFTB model is very different from the lowest energy structure, being the tg rotamer, which is actually the highest in B3LYP and AMBER/GLYCAM. The difference in tg rotamer obtained at B3LYP/aug-cc-pVTZ and SCC-DFTB is about 2 kcal/mol. Therefore, the conformational properties of

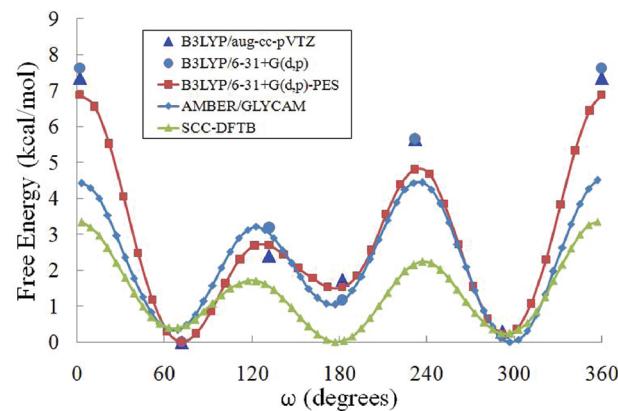


Figure 12. Free energy profile along ω in **1** obtained from the levels of theory shown in the caption. B3LYP calculations were performed using the CPCM model. AMBER/GLYCAM and SCC-DFTB free energy profiles are obtained from umbrella sampling simulations with explicit TIP3P waters.

furanosides, especially the rotamer populations obtained from the AMBER/GLYCAM, should be more reliable than those of the SCC-DFTB model.

The potential energy surfaces for the internal and external dihedrals of **2** obtained from the SCC-DFTB model are also found to be very different from those obtained with B3LYP/6-31+G(d,p). The potential energy surfaces obtained from SCC-DFTB and B3LYP/6-31+G(d,p) are very similar to those shown in Figures 10 and 11, and therefore, they are not shown here. In all cases, the geometries obtained from the SCC-DFTB model favor the tg rotamer.

Since the potential energy surfaces obtained at SCC-DFTB are very different from those obtained from high-level ab initio and DFT calculations, few other semiempirical methods, such as the AM1,²² PM3,^{23,24} and PM6,⁸⁵ are also investigated. Potential energy surfaces obtained at PM6 and B3LYP/6-31+G(d,p) are shown in Figure 13. AM1 and PM3 also provide

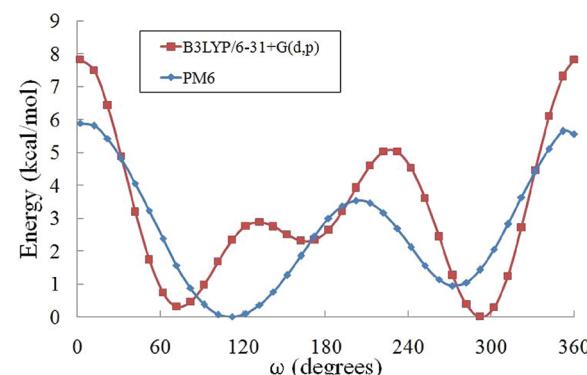


Figure 13. Potential energy along ω in **1** by freezing C4–C5–O5–H5, C4–C3–O3–H3, and all dihedrals in the ring as determined by the B3LYP/6-31+G(d,p) and PM6 levels of theory. AM1 and PM3 also provide potential energy surface very similar to PM6.

potential energy surfaces very similar to PM6. It is clear from the figure that the potential energy surface obtained with the PM6 method does not reproduce the PES obtained with B3LYP/6-31+G(d,p). In fact, only two minima have been found with these semiempirical methods. Therefore, the SCC-DFTB method is clearly better than AM1, PM3, and PM6 methods for furanosides.

CONCLUSIONS

A conformational study of mono- and diarabinofuranosides has been carried out to assess the reliability of the SCC-DFTB model for description of floppy five-membered ring systems. A carefully constructed force field, the AMBER/GLYCAM model, was used as a basis for comparison. Both conventional long MD and umbrella sampling simulations were performed to effectively sample the conformational space of the systems. The rotamer populations along the exocyclic torsion and the ring puckering obtained from the SCC-DFTB model were found to be qualitatively different from those obtained from the AMBER/GLYCAM approach. Proton–proton coupling constants were calculated from these simulations using both DFT-based Karplus relationships and the generalized Karplus relationships. The force field coupling constants obtained from both Karplus relationships were shown to agree well with the experimental NMR coupling constants, while the SCC-DFTB simulations led to poor agreement with experiment. Addition of dispersion corrections to DFTB does not appear to improve the situation. We conclude that in its current state of parametrization the SCC-DFTB model fails to capture the conformational preferences of arabinofuranosyl systems.

Additional analysis reveals that energetic properties obtained from the SCC-DFTB model are quite different from those calculated with high-level ab initio and DFT electronic structure approaches. The potential energy surface about the exocyclic torsion obtained from all ab initio and DFT methods were very similar in the gas phase; however, the potential energy surface obtained from SCC-DFTB was found to be very different. In solution, free energy profiles obtained from the AMBER/GLYCAM approach agree very well with those obtained from ab initio and DFT methods. It may be worth mentioning here that the high-level calculations in solutions performed with the CPCM implicit solvent model may not be able to describe hydrogen-bond interactions between waters and the OH groups in furanosides, and inclusion of explicit waters in the simulation may be necessary. The energetic barriers of exocyclic torsion and furanose puckering are too low in the semiempirical SCC-DFTB model compared to AMBER/GLYCAM. This suggests that the method may not account for subtle changes in structure. It is crucial to accurately describe both intermolecular (water–furanose) and intramolecular (OH groups in furanosides) hydrogen bonding in furanosides. Since SCC-DFTB was shown to underestimate the strength of hydrogen-bonding interactions,^{88,89} the method may not be used to study furanosides and similar systems. Therefore, a well-parametrized high-quality all-atom force field should be used to investigate statistically converged solution conformational properties of furanosides, and this may be true for other floppy systems. Our proposal is consistent with the findings of Roitberg et al.,⁹⁰ who also found from a thermodynamic study on L-alanine dipeptide that the results obtained from the classical ff99SB force field are in better agreement with experiment than most of the semiempirical methods such as the MNDO, AM1, PM3, RM1, PDDG/MNDO, PDDG/PM3, and SCC-DFTB. Therefore, the presumption that a more expensive method would yield better accuracy can be deceiving in the present case. Further developments of the SCC-DFTB model are necessary to obtain trustworthy conformational properties of furanoside systems. It should be pointed out here that the best semiempirical method found in this study is the SCC-DFTB

model when compared to AM1, PM3, and PM6. The SCC-DFTB method is still a good choice to study chemical reactions in large biological systems where the free energy barriers are large.

We finally remark that the AMBER/GLYCAM model is currently more appropriate than the SCC-DFTB model to study ensemble average properties of furanosides and probably their complexes with larger biological molecules such as proteins. The discrepancies between SCC-DFTB and different ab initio and DFT methods suggest the need for more accurate semiempirical methods. It is worthwhile to mention here that a recently developed next generation DFTB method called DFTB3 might prove useful to study furanose systems since DFTB3 describes the hydrogen-bonding interactions better than SCC-DFTB.⁸⁹

ASSOCIATED CONTENT

Supporting Information

Coupling constant data (experimental and theoretical), time dependence of dihedral angle and ring puckering, C4–C5 rotameric distributions and convergence plots, ring puckering distributions, H4–C4–O4 angular distributions, proton–proton dihedral angle distributions, ring pucker and dihedral angle correlation details and their 2D PMF, potential energy surfaces calculated from ab initio, DFT, SCC-DFTB and semiempirical methods for 1 and 2. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pnroy@uwaterloo.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to the Natural Sciences and Engineering Council of Canada (NSERC) for financial support as well as the SHARCNET for dedicated computational resources. We also acknowledge Q. Timerghazin, M. Nooijen, and T. Lowary for useful discussion.

REFERENCES

- (1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984; pp 1–556.
- (2) Lowary, T. L. *Curr. Opin. Chem. Biol.* **2003**, *7*, 749.
- (3) Houseknecht, J. B.; Lowary, T. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 677.
- (4) Minnikin, D. E. *The Biology of the Mycobacteria*; Ratledge, C., Stanford, J., Eds.; Academic Press: London, 1982; Vol. 1, pp 95–184.
- (5) D’Souza, F. W.; Ayers, J. D.; McCarren, P. R.; Lowary, T. L. *J. Am. Chem. Soc.* **2000**, *122*, 1251.
- (6) McCarren, P. R.; Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2001**, *105*, 5911.
- (7) Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Org. Chem.* **2000**, *65*, 4954.
- (8) Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.* **1999**, *121*, 9682.
- (9) Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2003**, *107*, 5763.
- (10) Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M. *J. Phys. Chem. A* **2003**, *107*, 372.
- (11) Houseknecht, J. B.; McCarren, P. R.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.* **2001**, *123*, 8811.

- (12) 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.
- (13) Taha, H. A.; Castillo, N.; Roy, P. N.; Lowary, T. L. *J. Chem. Theory Comput.* **2009**, *5*, 430.
- (14) Taha, H. A.; Castillo, N.; Sears, D. N.; Wasylshen, R. E.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theory Comput.* **2010**, *6*, 212.
- (15) Taha, H. A.; Roy, P.-N.; Lowary, T. L. *J. Chem. Theory Comput.* **2011**, *7*, 420.
- (16) Taha, H. A. Conformational Analysis of D-Arabinofuranosides using NMR Spectroscopy and Computational Chemistry. Unpublished Ph.D. Dissertation. University of Alberta, Edmonton, Alberta, Canada, 2010.
- (17) Angyal, S. J. *Adv. Carbohydr. Chem. Biochem.* **1984**, *42*, 15.
- (18) Biarnes, X.; Ardevol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686.
- (19) Spiwok, V.; Kralova, B.; Tvaroska, I. *Carbohydr. Res.* **2010**, *345*, 530.
- (20) Spiwok, V.; Tvaroska, I. *Carbohydr. Res.* **2009**, *344*, 1575.
- (21) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
- (22) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (23) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
- (24) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221.
- (25) Elstner, M. *Theor. Chim. Acta* **2006**, *116*, 316.
- (26) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2010**, *114*, 17142.
- (27) Sattelle, B. M.; Almond, A. *J. Comput. Chem.* **2010**, *31*, 2932.
- (28) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2010**, *132*, 12800.
- (29) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
- (30) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraserreid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
- (31) 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.
- (32) Porezag, D.; Frauenheim, T.; Kohler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947.
- (33) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.
- (34) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
- (35) Cui, Q.; Elstner, M.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 2721.
- (36) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203.
- (37) Zhang, X. D.; Harrison, D. H. T.; Cui, Q. *J. Am. Chem. Soc.* **2002**, *124*, 14871.
- (38) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9032.
- (39) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2000**, *256*, 15.
- (40) Elstner, M.; Frauenheim, T.; Kaxiras, E.; Seifert, G.; Suhai, S. *Phys. Status Solidi B* **2000**, *217*, 357.
- (41) Li, G. H.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125*, 15028.
- (42) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565.
- (43) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715.
- (44) Bohr, H. G.; Jalkanen, K. J.; Elstner, M.; Frimand, K.; Suhai, S. *Chem. Phys.* **1999**, *246*, 13.
- (45) Shishkin, O. V.; Gorb, L.; Luzanov, A. V.; Elstner, M.; Suhai, S.; Leszczynski, J. *J. Mol. Struct. (THEOCHEM)* **2003**, *625*, 295.
- (46) Zhou, H. Y.; Tajkhorshid, E.; Frauenheim, T.; Suhai, S.; Elstner, M. *Chem. Phys.* **2002**, *277*, 91.
- (47) Guo, H. B.; Rao, N.; Xu, Q.; Guo, H. *J. Am. Chem. Soc.* **2005**, *127*, 3191.
- (48) Gu, W.; Zhu, J.; Liu, H. *J. Theor. Comput. Chem.* **2002**, *1*, 62.
- (49) Xu, D.; Zhou, Y.; Xie, D.; Guo, H. *J. Med. Chem.* **2005**, *48*, 6679.
- (50) Xu, D.; Xie, D.; Guo, H. *J. Biol. Chem.* **2006**, *281*, 8740.
- (51) Xu, D. G.; Guo, H.; Cui, Q. *J. Phys. Chem. A* **2007**, *111*, 5630.
- (52) Pu, J.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 5454.
- (53) Witek, H. A.; Morokuma, K. *J. Comput. Chem.* **2004**, *25*, 1858.
- (54) Altona, C.; Sundaralal, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
- (55) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333.
- (56) Oleary, D. J.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 6629.
- (57) Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102.
- (58) Devries, N. K.; Buck, H. M. *Carbohydr. Res.* **1987**, *165*, 1.
- (59) Bock, K.; Duus, J. O. *J. Carbohydr. Chem.* **1994**, *13*, 513.
- (60) Tvaroska, I.; Carver, J. P. *J. Phys. Chem. A* **1997**, *101*, 2992.
- (61) Islam, S. M.; Richards, M. R.; Taha, H. A.; Byrns, S. C.; Lowary, T. L.; Roy, P.-N. *J. Chem. Theory Comput.* **2011**, *7*, 2989.
- (62) 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.
- (63) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357.
- (64) 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.
- (65) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (66) 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.; Kolossváry, 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.
- (67) Adelman, S. A.; Doll, J. D. *J. Chem. Phys.* **1976**, *64*, 2375.
- (68) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
- (69) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (70) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
- (71) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578.
- (72) Valleau, J. P.; Torrie, G. M. *A guide for Monte Carlo for statistical mechanics in: Statistical Mechanics Part A*; Plenum Press: New York, 1977; pp 169–194.
- (73) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959.
- (74) Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press Inc.: New York, 1987; pp 1–274.
- (75) Shen, J.; Mccammon, J. A. *Chem. Phys. Lett.* **1991**, *158*, 191.
- (76) Woolf, T. B.; Roux, B. *J. Am. Chem. Soc.* **1994**, *116*, 5916.
- (77) Haydock, C.; Sharp, J. C.; Prendergast, F. G. *Biophys. J.* **1990**, *57*, 1269.
- (78) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.
- (79) Grossfield, A. *WHAM: the weighted histogram analysis method*, version 2.0.2; University of Rochester Medical Center: Rochester, NY, 2008; <http://membrane.urmc.rochester.edu/content/wham>.
- (80) Efron, B.; Tibshirani, R. J. *An introduction to the bootstrap*; Chapman and Hall: New York, 1994; pp 1–426.
- (81) van Halbeek, H. *Carbohydrates and Glycoconjugates. Encyclopedia of Magnetic Resonance*; John Wiley & Sons, Ltd.: New York, 2007.
- (82) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.;

- Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (83) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.
- (84) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.
- (85) Stewart, J. J. *P. J. Mol. Model.* **2007**, *13*, 1173.
- (86) Evdokimov, A. G.; Kalb, A. J.; Koetzle, T. F.; Klooster, W. T.; Martin, J. M. *L. J. Phys. Chem. A* **1999**, *103*, 744.
- (87) Dzakula, Z.; Westler, W. M.; Edison, A. S.; Markley, J. L. *J. Am. Chem. Soc.* **1992**, *114*, 6195.
- (88) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861.
- (89) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931.
- (90) Seabra, G. D.; Walker, R. C.; Roitberg, A. E. *J. Phys. Chem. A* **2009**, *113*, 11938.