Carbohydrate Solution Simulations: Producing a Force Field with Experimentally Consistent Primary Alcohol Rotational Frequencies and Populations

MICHELLE KUTTEL,[1] J. W. BRADY,[2] KEVIN J. NAIDOO[1]
[1] *Department of Chemistry, University of Cape Town, Rondebosch 7701, Cape Town, South Africa*
[2] *Department of Food Science, Cornell University, Ithaca, New York 48109*

Received 24 September 2001; Accepted 10 April 2002

Abstract: We present a CHARMM Carbohydrate Solution Force Field (CSFF) suitable for nanosecond molecular dynamics computer simulations. The force field was derived from a recently published sugar parameter set.[1] Dihedral angle parameters for the primary alcohol as well as the secondary hydroxyl groups were adjusted. Free energy profiles of the hydroxymethyl group for two monosaccharides (β-D-glucose and β-D-galactose) were calculated using the new parameter set and compared with similar force fields. Equilibrium rotamer populations obtained from the CSFF are in excellent agreement with NMR data (glucose $gg{:}gt{:}tg \approx 66{:}33{:}1$ and galactose $gg{:}gt{:}tg \approx 4{:}75{:}21$). In addition, the primary alcohol rotational frequency is on the nanosecond time scale, which conforms to experimental observations. Equilibrium population distributions of the primary alcohol conformers for glucose and galactose are reached within 10 nanoseconds of molecular dynamics simulations. In addition, gas phase vibrational frequencies computed for β-D-glucose using this force field compare well with experimental frequencies. Carbohydrate parameter sets that produce both conformational energies and rotational frequencies for the pyranose primary alcohol group that are consistent with experimental observations should allow for increased accuracy in modeling the flexibility of biologically important (1-6)-linked saccharides in solution.

 J Comput Chem 23: 1236–1243, 2002

Key words: carbohydrates; force field; hydroxyl rotation; free energy; saccharide solutions; hydroxymethyl; CHARMM; WHAM

Introduction

Polysaccharides are, in general, highly flexible molecules, particularly in solution. Computer models can explain the complex role of sugar flexibility in biology and thus complement experimental measurements. For a saccharide computer model to successfully interpret conformational function in condensed phase systems it must predict the thermodynamic and transport properties of saccharides correctly. Therefore, accurate calculations require a model that produces solution conformational sampling and dynamics that are consistent with experimental observations.

Because pyranose rings in solution are relatively rigid, existing chiefly in the $^4C_1$ chair conformation,[2] the principal degrees of freedom explored at room temperature in solution for these systems are rotations about the glycosidic linkages and the hydroxyl dihedrals, in particular the 6-hydroxymethyl dihedral (Fig. 1). Rotation about the hydroxymethyl C_5–C_6 bond may be defined by either of two equivalent dihedral angles: O_5–C_5–C_6–O_6 (ω) or C_4–C_5–C_6–O_6 (ω_2). Herein we use the first definition. Typically, only three minimum energy staggered conformations of the hydroxymethyl group are considered. They are termed *gg*, *gt*, and *tg*, respectively, where the first letter refers to the *trans* or *gauche* orientation of the first dihedral (ω) and the second letter the orientation of the second dihedral (ω_2). The hydroxymethyl orientations and their corresponding ω values are shown in Figure 2.

The rotameric distribution of the hydroxymethyl group for different monosaccharides has proven difficult to establish experimentally. Conformational preferences have been determined

Correspondence to: K. J. Naidoo; e-mail: knaidoo@science.uct.ac.za

Contract/grant sponsor: USDA-ARS; contract/grant number: 58-4012-5-F120

Contract/grant sponsor: National Research Foundation (NRF Pretoria)

This article contains Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/23/1236 or http://www.interscience.wiley.com/jpages/0192-8651/v23.1236.html

chiefly from X-ray crystallography and NMR studies, and to some extent from optical rotation experiments.[3] Although there is variation in the reported experimental rotameric distributions, the overall *gg*:*gt*:*tg* population rankings are estimated to be approximately 6:4:0 for glucopyranosides and 2:6:2 for galactopyranosides.[3–5] In particular, Bock et al. report β-D-glucose to have a distribution of *gg*:*gt*:*tg* = 52:41:7[6] and α-D-galactose a distribution of *gg*:*gt*:*tg* = 12:56:32.[3] Earlier studies by Nishida et al. give a ratio for β-D-glucose of *gg*:*gt*:*tg* = 53:45:2[4] and β-D-galactose of *gg*:*gt*:*tg* = 18:61:21.[5]

The conformational preferences of the primary alcohol group of pyranose rings in solution are assumed to be determined by (1) stereoelectronic effects (the 1,3-biaxial interaction and *gauche* effect), (2) steric 1–4 interactions, (3) hydrogen bonding, and (4) solvent effects.[7] However, there has as yet been no consensus on the relative importance of these effects, with opinion divided on which influence predominates. Experimental studies tend to discount the effects of hydrogen bonding. Rockwell et al. investigated the solvent dependence of hydroxymethyl rotamer populations by analyzing NMR $^3J_{H5,H6R}$ and $^3J_{H5,H6S}$ coupling constants of fully methylated glucose and glucose methylated at all positions except for the primary alcohol in solvents of varying polarity.[8] As the aqueous rotameric populations of the two molecules were identical, they concluded that the hydroxymethyl conformational preferences in water are due to a polarization effect on the intrinsic stability of the molecule,[9] rather than an explicit hydrogen bond effect. A similar study of galactose derivatives by de Vries and Buck concluded that the differences in rotamer populations were due to stereoelectronic effects.[10]

In contrast, some theoretical studies have emphasized the importance of hydrogen bonding. Cramer and Truhlar concluded from semiempirical calculations of D-glucopyranose conformers in a continuum solvent that the hydroxymethyl conformational equilibrium is dominated by solute–solvent hydrogen bonding interactions.[7] Molecular orbital calculations of carbohydrate model compounds by Tvaroška et al. indicate that *gauche* preference of the hydroxymethyl group is due to the presence of hydrogen bonding and not the *gauche* effect.[11] However, a recent theoretical study by Kirschner and Woods[12] suggests that water plays a central role by disrupting the intramolecular hydrogen bonding in both glucose and galactose, allowing the rotamer populations to be determined by internal electronic interactions and steric repulsions.

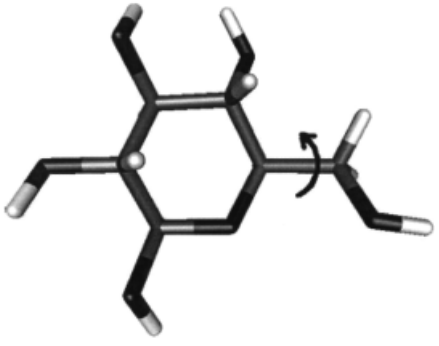

Figure 1. β-D-Glucose molecule, showing rotation about the primary alcohol ω dihedral.

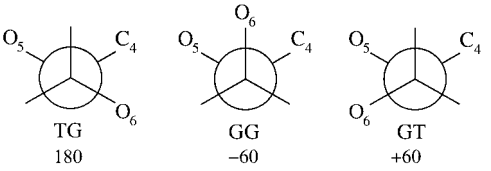

Figure 2. Definitions of the three minimum-energy primary alcohol positions for a pyranose ring. Values for the primary alcohol dihedral ω are shown below each conformation.

Therefore, the relative degree of importance of the various influences on the primary alcohol population in solution remains undecided. It is beyond the scope of any work using a classical force field approach to resolve this issue, as it depends on complex electronic effects. The most a classical force field can hope to achieve is a satisfactory reproduction of experimental conformational preferences.

Several parameter sets for molecular dynamics simulations of carbohydrates have been developed, including those for the CHARMM,[1,13,14] AMBER,[15–18] and the GROMOS force fields.[19] In these force fields, X-ray structures and *ab initio* results have traditionally been used for parameterization and validation of the dihedral angles, with the emphasis on obtaining correct relative conformational energies. However, comparison of the relative energies of static primary alcohol conformations in vacuum does not take into account the contributions of entropy and solvation to the equilibrium conformer distribution. In addition, little attention has been paid to the frequency or time scale of transition between the primary alcohol conformations, which is determined by the height of the free energy barriers between conformational minima. An accurate force field for molecular dynamics calculations in solution should reproduce not only the experimental equilibrium conformational distribution, but also the experimental time scale for the primary and secondary alcohol group rotations in solution.

Experimental evidence indicates that hydroxymethyl rotational motion is fairly restricted. Exchange between the different conformations of the hydroxymethyl groups has been estimated to occur on a time scale of between 10^{-10} and 10^{-3} s,[2,6] with recent studies placing rotational isomerization more specifically on the nanosecond time scale.[20,21] As hydroxymethyl transitions are fairly infrequent on the time scale of molecular dynamics simulations, it is difficult to judge from a nanosecond MD simulation whether the primary alcohol rotameric populations are correctly represented by a given force field. Difficulties with achieving a representative sample for α(1–6) linked oligosaccharides using a CHARMM parameter set have been recently reported.[22] In addition, there have been several reports of MD simulations with various force fields exhibiting very few primary alcohol rotational transitions.[2,23]

The original CHARMM parameter set for carbohydrates,[13] here termed HGFB after the authors' initials, is known to produce incorrect relative populations of the primary alcohol rotamers in most simulations, as it favors the *tg* conformation over *gt* and *gg* in both vacuum and solution.[24,25] Palma et al. produced a new CHARMM parameter set,[1] here termed PHLB, which aimed to correct the relative energies of the primary alcohol, as well as the unrealistic flexibility of the HGFB carbohydrate model. PHLB

replaces the general dihedral angle terms used in HGFB with specific dihedral angle terms, which were parameterized to reproduce experimental vibrational frequency data and small molecule *ab initio* dihedral angle rotational energy profiles, such as ethylene glycol.[26] Adiabatic map calculations for PHLB indicate correct relative energies for the three conformations ($gg < gt < tg$). However, the PHLB parameter set has been found to exhibit very infrequent conformational transitions of the primary alcohol, to the extent that equilibration is not achieved in nanosecond MD simulations.

In the present study, we investigate thoroughly the primary alcohol behavior of the PHLB and HGFB parameter sets. In each case, the relative energies of the three primary alcohol conformers, the barriers to rotation, and the equilibrium rotamer distributions were determined by calculation of the potential of mean force for rotation about the primary alcohol dihedral, ω, for both glucose and galactose in vacuum and solution. Based on the results obtained, we modified the PHLB parameter set to alter the relative free energies of the gg, tg, and gt primary alcohol conformations and lower the primary alcohol rotational energy barriers. We refer to this modified force field as the Carbohydrate Solution Force Field (CSFF). Calculated potential of mean force profiles for β-D-glucose and β-D-galactose in solution using this new parameter set confirm that it will exhibit hydroxymethyl conformer equilibrium distributions and rotational frequencies in agreement with experimental observations. The validity of the force field was further established by a comparison of the calculated and observed normal modes of vibration of α-D-glucose.

Methods

Potential of Mean Force (PMF) Calculations

A complete description of the energy barriers to primary alcohol rotation in solution is available only through calculation of the free energy change as a function of the rotation, or the potential of mean force. The free energy incorporates contributions from entropy and solvent–solute interactions. Given the ω dihedral angle probability distribution, $P(\omega)$, it is possible to calculate the potential of mean force, $W(\omega)$, from the relation

$$W(\omega) = -kT \ln P(\omega) \tag{1}$$

where k is the Boltzmann constant and T is the temperature.

In principle, $P(\omega)$ can be obtained from a single molecular dynamics simulation. However, conventional MD simulations do not sample the high-energy barrier regions very frequently, and impracticably long simulations are required to provide a satisfactory estimate of $P(\omega)$. To increase the sampling in the high-energy regions of ω, several free-energy simulation techniques have been developed, one of which is adaptive umbrella sampling.[27,28] This method uses an applied "umbrella potential" that adapts over a series of simulations until uniform sampling over the entire range of the degree(s) of freedom chosen is achieved. At this point, the umbrella potential has converged to the inverse of the potential of mean force. We have previously reported details of this method as applied to carbohydrates.[29]

In this study, we implemented adaptive umbrella sampling using a "flat" umbrella potential for the first simulation and successively applying the Weighted Histogram Analysis Method (WHAM)[30–32] to arrive at subsequent estimates. WHAM is a technique whereby biased probability distributions from several simulations are combined using computed weighting factors for each of the distribution histograms to obtain the best estimate for the actual unbiased probability distribution. Briefly, the method is as follows. The degree of freedom of interest (in this case, ω) is divided into k bins. The biased probability distribution for simulation i is computed as an unnormalized histogram, n_{ij}, of the values of ω occurring during the simulation. After j simulations, the optimal unbiased probability histogram, p_k, is found by iterating the following WHAM equations:

$$p_k = \frac{\sum_j n_{j,k}}{\sum_j N_j f_j c_{j,k}} \tag{2}$$

and

$$f_j = \frac{1}{\sum_k c_{j,k} p_k} \tag{3}$$

with

$$c_{j,k} = e^{-\beta U_j(\omega_k)} \tag{4}$$

$$N_j = \sum_k n_{j,k} \tag{5}$$

Here, f_j, are the free energy weighting factors, N_j is the total number of configurations stored during the jth simulation and $U_j(\omega_k)$ is the umbrella potential of the kth bin in the jth simulation. The implementation of WHAM and its application to adaptive umbrella sampling in both Monte Carlo and Molecular Dynamics simulations has been detailed in several publications.[32–37]

The ω dihedral PMF profiles in both vacuum and solution were computed for β-D-glucose using the parameter sets HGFB, PHLB, and CSFF and the profiles for β-D-galactose using the PHLB and CSFF parameter sets.

Simulation Details

The CHARMM[38] program was used for all simulations. The simulations used for the PMF calculations incorporated local modifications to the USERE routine to implement adaptive umbrella sampling using a continuous potential rather than the windowing procedure implemented in CHARMM.[29]

Both the β-D-glucose and β-D-galactose simulations comprised a single monosaccharide molecule surrounded by 499 TIP3P water molecules in a cube of length 24.79 Å with periodic boundary conditions. The nonbonded interactions were truncated using a

Table 1. Modified Dihedral Force Constants ($E_\theta = \Sigma k_n(1 + \cos(n\theta - \theta_0))$, $n = 1, 2, 3$) for the CSFF Parameter Set

Dihedral	Force constant (k_n)	Periodicity (n)	Phase angle (θ_0)
OES–CTS–CPS–HAS	0.2086 (0.1618)	3	0.0
OES–CTS–CPS–OHS	−4.0193 (−3.7993)	1	0.0
	−1.2688 (0.5686)	2	0.0
	−0.9704 (0.4202)	3	0.0
CTS–CTS–CPS–OHS	−1.7139 (−1.9139)	1	0.0
	−1.0239 (−0.3739)	2	0.0
	−0.0340 (−0.0340)	3	0.0
HOS–OHS–CTS–HAS	0.0677 (0.1686)	3	0.0
HOS–OHS–CBS–HAS	0.0677 (0.1686)	3	0.0
HOS–OHS–CPS–HAS	0.0677 (0.1686)	3	0.0

Units are kcal/mol. The PHLB values appear in parentheses.

switching function applied on a neutral group basis between 10.0 and 12.0 Å. Molecular dynamics trajectories were integrated as a microcanonical ensemble using a Verlet integrator with a 1 fs time step. The SHAKE algorithm[39] was used to fix the water molecule geometry and covalent bonds between hydrogen and the heavy atoms.

Force Field Parameters

To adjust the primary alcohol rotation without affecting the ring parameterization, a new atom type for the C6 carbon (CPS) was introduced into the CSFF parameter set. This atom type can be considered a generic nonring carbon. Force constants for all the H—O—C—H dihedrals were reduced to 40% of their original value to increase the slow rate of rotation of the secondary hydroxyls. Alterations of other dihedral force constants were restricted to those affecting the O_5–C_5–C_6–O_6, O_5–C_5–C_6–H_6 and C_4–C_5–C_6–O_6 dihedrals. This ensures that the ethylene glycol rotational profile remains parameterized to closely approximate the profile obtained from *ab initio* calculations.[26] The new dihedral term parameters are listed in Table 1. The k_θ values for all other primary alcohol dihedral angles (see supplementary material) remain the same as the CTS atom type in the PHLB force field. Except the O—H bond force constant, which was lowered to 460.5 kcal $Å^{-2}$ mol^{-1} to better match the experimentally measured O—H stretching frequency all other parameters were left unchanged from the PHLB set.

Normal Mode Analysis

The normal modes for α-D-glucose were calculated for the CSFF and PHLB parameter sets by diagonalizing the mass-weighted Hessian. The α-D-glucose structure for the normal mode calculation was taken from the neutron-diffraction data of Brown and Levey,[40] with subsequent minimization. Three stages of Newton-Raphson minimization were employed: the first with all the hydroxyl and the primary alcohol dihedrals constrained to their experimental configuration, the second with constraints only on the secondary hydroxyl dihedrals and the final with all constraints released. The resultant conformation was closer to that reported in the crystal structure than one obtained using constraint-free minimization. Vibrational frequency assignments were made on the basis of the potential energy distributions for each mode and used to match the CSFF and PHLB modes. Our calculated frequencies were compared with the experimental frequencies of Dauchez et al.[41] based on the vibrational assignments for their calculated normal modes.

Results and Discussion

The ω PMF profiles calculated for β-D-glucose using the HGFB, PHLB, and CSFF parameter sets in both vacuum and solution are shown in Figure 5. The ω solution PMF for the HGFB force field clearly indicates that the hydroxymethyl rotational profile for this force field is not in agreement with experimental observations: the *tg* conformation is lowest in energy in both vacuum and solution. Further, the equilibrium conformer distribution calculated from the

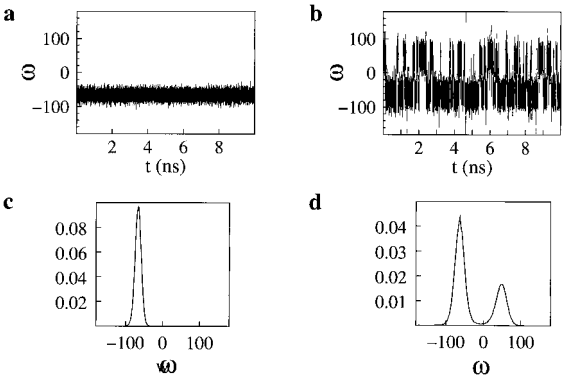

Figure 3. β-D-Glucose: times series and statistics from 10 ns MD simulations using the PHLB (left column) and CSFF (right column) parameter sets. Top: primary alcohol ω dihedral time series plots [PHLB (a) and CSFF (b)]. Bottom: probability distribution profiles for the ω dihedral [PHLB (c) and CSFF (d)].

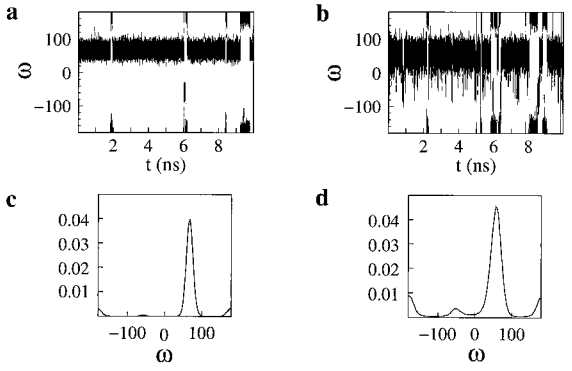

Figure 4. β-D-Galactose: times series and statistics from 10-ns MD simulations using the PHLB (left column) and CSFF (right column) parameter sets. Top: primary alcohol dihedral ω time series plots [PHLB (a) and CSFF (b)]. Bottom: probability distribution profiles for the ω dihedral [PHLB (c) and CSFF (d)].

solution PMF curve [using eq. (1)] is $gg{:}gt{:}tg \approx 24{:}27{:}49$, which is not compatible with experimental populations.

The free-energy ordering obtained for the PHLB parameter set, $gg < gt \ll tg$, is an improvement on the HGFB set. Here, the equilibrium distribution for β-glucose calculated from the PMF

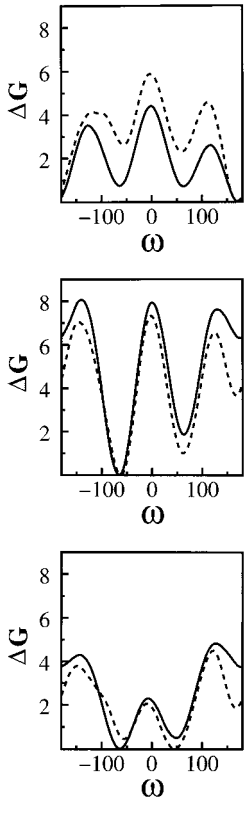

Figure 5. β-D-Glucose: vacuum (- - -) and TIP3P (—) solution PMF profiles (kcal mol^{-1}) obtained for the primary alcohol ω dihedral using the three-parameter sets: HGFB (top), PHLB (middle) and CSFF (bottom).

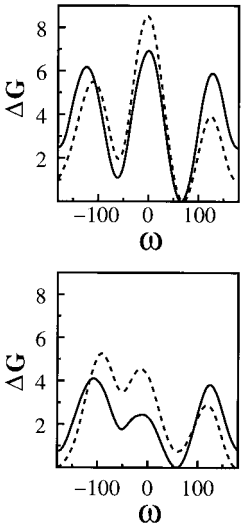

Figure 6. β-Galactose: vacuum (- - -) and TIP3P (—) solution PMF profiles (kcal mol^{-1}) obtained for the primary alcohol ω dihedral using the PHLB (top) and modified CSFF (bottom) parameter sets.

profile is $gg{:}gt{:}tg = 85{:}15{:}0$, which is in better agreement with experiment, although it favors the gg conformation somewhat more than is generally predicted ($gg{:}gt{:}tg = 52{:}41{:}7$,[6] $53{:}45{:}2$[4]). However, the energy barriers between conformations (7 to 8 kcal mol^{-1}) are more than 10 kT, making it improbable for transitions between the different conformations to occur during nanosecond room temperature (298 K) molecular dynamics simulations. This is illustrated by the ω time series for a β-D-glucose TIP3P solution simulation (Fig. 3): a 10-ns molecular dynamics simulation started from gg maintained this conformation throughout and equilibrium was not achieved.

The CSFF ω PMF's are significantly altered from the corresponding PHLB profiles. In the case of glucose (Fig. 5), the gg conformation is still favored in solution, although the relative energies of the gt and tg conformations have been lowered by about 1.5 kcal mol^{-1} and 2.0 kcal mol^{-1}, respectively. The calculated solution equilibrium conformational distribution is $gg{:}gt{:}tg \approx 66{:}33{:}1$, which is an improvement on the conformational distribution for the PHLB parameter set and in good agreement with NMR experimental estimates. Further, the barriers to rotation about the primary alcohol dihedral have been significantly lowered compared to those from the PHLB parameter set. This lowering is reflected in the increased frequency of transition seen in Figure 3b. Rotation takes place on a nanosecond time scale, in agreement with recent experimental estimates.[20,21] The rotameric distribution of β-D-glucose for a 10 ns simulation (Fig. 3) is $gg{:}gt{:}tg \approx 69{:}31{:}0$, indicating that hydroxymethyl equilibrium populations have been achieved.

Similarly, calculation of the ω free-energy profiles for galactose in water and vacuum (Fig. 6) illustrates the decrease in CSFF ω rotational barriers compared with those in the PHLB force field. As with glucose, the lowered barrier heights in CSFF facilitate rotation about the primary alcohol and, hence, more rapid equilibration (Fig. 4). Calculated population distributions for the galactose PMF curves are PHLB: $gg{:}gt{:}tg = 19{:}69{:}8$ and CSFF: $gg{:}gt{:}tg =$

Table 2. Observed (Raman and Infrared) Frequencies (from Dauchez et al.[41]), Computed Vibrational Frequencies for the CSFF and PHLB Force Fields and an Approximate Description of Each Mode of Vibration for α-D-glucose

Observed frequencies (cm^{-1})		Calculated frequencies (cm^{-1})		
Raman	IR	CSFF	PHLB	Approximate potential energy distribution
41				
49	48			
59	60			
70	70			
76	77	74		τC_5C_6 + CCO + CCC + CC
85	88		87	τC_5C_6 + τCC_{cyc} + CCC + OCC
92				
101	98	103	117	τC_5C_6 + CCO + CCC + CC
112	112			
	126			
134	136	126	139	CCC + CC + CCO + COC + τCC_{cyc} + τC_5C_6
142				
	152			
155				
184	188	186	186	$C_4C_5C_6$ + $C_1O_5C_5$ + $C_5C_6O_6$ + CCO + CC_{cyc}
234	237	241	286	τC_6O + $O_5C_5C_6$
	255			
277	270			
292	290	294	309	τC_3O_3 + τC_5C_6 + $O_5C_5C_6$ + $C_4C_5C_6$ + COH + CO + CC
312		309	326	OCC + τC_3O_3 + τC_5C_6
		318	327	τCO + CCO
		335	337	CCO + τCO + τCC
	350	355	357	CCO + CC + CO + τCO
369		370	379	τCO + CCC + OCC + CC
400	399	395	407	τCO + OCO_5
411	412	408	413	CCO + CC + CO + C_6C_5O
426	430	436	445	CCO + CC + CO + τCO
443	442	444	449	CCO + OC_1O_5 + CC + τCO + CO
		469	472	CCC + CO + CCC + CCO + τCO
497		513	519	CCO + CCC + CC + CO + τCO
542	549	552	555	CCC + CO + CC + CO + τCO (C_1)
559				
580	572	574	572	CCO + CO + CC + CCC
615	618	610	615	$O_1C_1O_5$ + O_5CC + CO_5C + CCO + CCC + CO + CC
657	645	682	683	$O_1C_1O_5$ + O_5CC + CO_5C + CCO + CCC + CO + CC
773	775	734	736	OCC + CC
		802	804	O_5CC + CO + CCC + CCO
842	840	845	857	C_5C_6H + HC_6O + CO
916	917			
		930	935	CO + CC + CCC + COH
1004	995			
1023		1033	1033	CO (C_6) + C_6OH + HC_6O + HC_6H
1055	1045	1063	1065	CC + CO + HCC + HCO
1070		1072	1073	CO + CC + HCC + CCO
1078	1078	1077	1081	CO + HCO (c3)
1105	1105	1100	1101	CCH + HCO + CO + CC
1113	1113	1109	1111	HCO (c1) + HCC + CO + CC + CCO
1123				
1133	1145	1142	1146	HCC + CC
1153		1156	1163	CCO + CCC + CC + CO + HCO + COH + HCC
1204	1202	1186	1188	HCC + OCC + HCO + CO
1226	1222	1217	1218	HCO + C_1O_5 + C_5C_6 + CO + CCH + CCC
		1248	1246	C_1O_5 + $c5O_5$ + HC_6O + c5c6H + CCH + HCO
		1264	1265	CCH + COH

(*continued*)

Table 2. (Continued)

Observed frequencies (cm^{-1})		Calculated frequencies (cm^{-1})		
Raman	IR	CSFF	PHLB	Approximate potential energy distribution
1279	1275	1279	1282	COH + O_5C_1H + CCH + CCO + C_1O + C_1O_5 + C_5C_6H
		1291	1293	COH + O_5C_1H + CCH + CCO + C_1O + C_1O_5 + C_5C_6H
1296	1295	1297	1301	COH + O_5C_1H + CCH + CCO + C_1O
		1324	1324	COH + C_1OH + O_5C_1H + CCH + CCO
1332	1340	1331	1333	COH + O_5C_1H + CCH + CCO
		1338	1341	COH + CCH + HCO
1347		1351	1353	C_6OH + HC_6H
1372	1375			
1407	1407	1389	1391	COH + HCO + CO
1440	1440	1422	1427	C_5C_6H + HC_6H + C_6OH
1460	1460	1495	1497	HC_6H + HC_6O + C_6O_6 + C_5C_6
		1515	1515	CO + CC + HCO + COH
		1539	1539	CC + HCO + CO + CCH
		1592	1592	CC + CO + HCO + HCC + CCC
		1631	1628	CO + CC + HCO + OCC + CCC
		1641	1641	CC + CO + HCO + HCC + OCC
		1739	1738	CO + CC + CCC + HCO + CCH
		1768	1766	CCC + CO + CC + HCO + OCC
2877	2885	2899	2899	C_6—H
2890	2898	2926	2926	C—H
2914	2920	2931	2931	C—H
		2934	2934	C—H
		2938	2939	C—H
2946	2947	2945	2945	C—H
2959		2958	2959	C—H
	3100	3383	3685	O—H
3385	broad	3385	3688	O—H
		3387	3689	O—H
3405	band	3389	3691	O—H
3415	3500	3393	3694	O—H

4:75:21. These are both comparable with experimental results (*gg*:*gt*:*tg* = 12:56:32[3] and 18:61:21[5]). Although the CSFF parameter set favors the *gt* conformation more than is predicted, the relative preference for the *tg* conformation over *gg* compares well with experimental observations. The β-D-galactose distributions calculated from 10 ns solution simulations (Fig. 4) are PHLB: *gg*:*gt*:*tg* ≈ 1:91:8 and CSFF:*gg*:*gt*:*tg* ≈ 6:80:14, indicating that CSFF, in contrast with PHLB, has, in fact, achieved equilibrium.

Comparisons of the vacuum and solution PMF profiles reveal the effect that solvent has on the relative sampling of the three different rotamers. For each of the parameter sets, the effect of the TIP3P solvent is to raise the energies of the *gt* and *tg* conformations relative to *gg*, by approximately 1 kcal mol^{-1} and 2.5 kcal mol^{-1}, respectively, for the PHLB and CSFF parameter sets and 0.5 kcal mol^{-1} and 2 kcal mol^{-1} for the HGBF case. Thus, solvation favors the *gg* and *gt* conformations over *tg*. This accords with quantum mechanical studies of glucose in aqueous solution by Cramer and Truhlar, where unfavorable solvation of *tg* was found to destabilize this conformation relative to *gt* and *gg*.[7] Furthermore, the PMF profiles show that solvation favors the *gg* conformation over *gt*.

The normal modes of α-D-glucose for both the CSFF and PHLB parameter sets were calculated. These are reported in Table 2, along with the experimental frequencies[41] and an approximate description of each mode in terms of bond stretches, angle bends and torsion angle rotations. Comparison with observed frequencies is good, particularly considering that the PHLB force field was not parameterized to fit saccharide vibrational frequencies. The fit is particularly close in the mid- to low-frequency range. These are the most important modes for the sugar force field to reproduce, as it is the region of torsional vibrations and complex motions of the ring. The normal modes calculated for PHLB and CSFF are similar and are generally with in a few wave numbers of each other. In the very high-frequency region associated with O—H and C—H stretches, the CSFF force field better matches the experimental frequencies. This was achieved by lowering the O—H bond force constant, as mentioned earlier. Aside from the O—H stretching region, the parameter sets differ most in the modes associated with the C5–C6 and hydroxyl dihedrals, as is to be expected. The CSFF

modes in these regions compare better with experiment than do the PHLB predicted modes.

Both CSFF and PHLB exhibit seven modes in the 1500 to 1800 cm^{-1} region that have no counterparts in the experimental spectra. These modes involve bond stretches and angle bends for the ring atoms, and are a reflection of the high bond and angle force constants for the ring atoms in the CSFF and PHLB parameter sets. The CHARMM force field used here does not make use of crossterms, and so it was necessary to increase these force constants to reduce the unusual ring flexibility observed in the HGFB force field.[42] NMR relaxation experiments indicate that the pyranose sugar rings conserve the $^4C_1$ conformation in solution. Pyranose ring C^{13} T_1 relaxation times have been previously calculated using the PHLB model, and were shown to be in excellent agreement with experiment.[22]

Conclusions

The design of carbohydrate force fields has previously emphasized agreement with *ab initio* static conformational calculations. The accuracy of the solution conformational dynamics of essential saccharide variables such as the hydroxymethyl group has, therefore, largely been ignored in the parameterization process. In this work we have modified a recently produced force field (PHLB) with the express aim of ensuring that the rotational frequency of the primary alcohol group is in agreement with recent experimental observations. The calculated potential of mean force profiles for the Carbohydrate Solution Force Field (CSFF) confirm that it will produce hydroxymethyl rotameric distributions and frequencies in solution for β-D-glucose and β-D-galactose that are consistent with experimental evidence. The CSFF parameter set exhibits primary alcohol rotational transitions on a nanosecond time scale for molecular dynamics solution simulations of β-D-glucose and β-D-galactose. This is in agreement with the estimated experimental rotational frequencies, and is a marked improvement on the behavior of the PHLB parameter set. As a result, equilibrium populations for the *gg*, *gt*, and *tg* primary alcohol conformations are now achievable with nanosecond simulations. Saccharide simulations in the nanosecond regime, although longer than has been generally the case, are attainable with current computing power. In addition, normal coordinate analysis of α-D-glucose in vacuum compares favorably with experimental data, producing vibrational spectra similar to those of contemporary force fields. The CSFF parameter set thus allows for conformationally representative simulations of (1–6)-linked saccharides in solution on a computationally feasible time scale.

Acknowledgments

The authors would like to thank Dr. J. L. Willet and Dr. Frank Momany (USDA-ARS, Peoria IL laboratory) for helpful discussions.

References

1. Palma, R.; Himmel, M. E.; Liang, G.; Brady, J. W. In ACS Symposium Series: Glycosyl Hydrolases in Biomass Conversion; Himmel, M. E., Ed.; American Chemical Society: Washington, DC, 2001; p 112.
2. Hajduk, P. J.; Horita, D. A.; Lerner, L. E. J Am Chem Soc 1993, 115, 9196.
3. Bock, K.; Duus, J. Ø. J Carbohydr Chem 1994, 13, 513.
4. Nishida, Y.; Ohrui, H.; Meguro, H. Tetrahedron Lett 1984, 25, 1575.
5. Nishida, Y.; Ohrui, H.; Meguro, H. J Carbohydr Res 1988, 7, 239.
6. Bock, K.; Guzman, J. B. F.; Ogawa, S. Carbohydr Res 1988, 174, 354.
7. Cramer, C. J.; Truhlar, D. G. J Am Chem Soc 1993, 115, 5745.
8. Rockwell, G. D.; Grindley, T. B. J Am Chem Soc 1998, 120, 10953.
9. Ma, B.; Schaefer, H. F.; Allinger, N. L. J Am Chem Soc 1998, 120, 3411.
10. de Vries, N. K.; Buck, H. M. Carbohydr Res 1987, 165, 1.
11. Tvaroska, I.; Carver, J. P. J Phys Chem B 1997, 101, 2992.
12. Kirschner, K. N.; Woods, R. J. Proc Natl Acad Sci USA 2001, 98, 10541.
13. Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. Carbohydr Res 1988, 180, 207.
14. Reiling, S.; Schlenkrich, M.; Brickmann, J. J Comput Chem 1996, 17, 450.
15. Glennon, T. M.; Zheng, Y.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M. J Comput Chem 1994, 15, 1019.
16. Woods, R. J.; Dwek, R. A.; Edge, C. J. J Phys Chem 1995, 99, 3832.
17. Senderowitz, H.; Parish, C.; Still, W. C. J Am Chem Soc 1996, 118, 2078.
18. Momany, F. A.; Willett, J. L. Carbohydr Res 2000, 326, 194.
19. Ott, K.; Meyer, B. J Comput Chem 1996, 17, 1068.
20. Behrends, R.; Cowman, M. C.; Eggers, F.; Eyring, E. M.; Kaatze, U.; Majewski, J.; Petrucci, S.; Richmann, K.-H.; Riech, M. J Am Chem Soc 1997, 119, 2182.
21. Stenger, J.; Cowman, M. C.; Eggers, F.; Eyring, E. M.; Kaatze, U.; Petrucci, S. J Phys Chem B 2000, 104, 4782.
22. Best, R. E.; Jackson, G. E.; Naidoo, K. J. J Phys Chem B 2001, 105, 4742.
23. Simmerling, C.; Fox, T.; Kollman, P. A. J Am Chem Soc 1998, 120, 5771.
24. Brady, J. W. J Am Chem Soc 1989, 111, 5155.
25. Schmidt, R. K.; Brady, J. W.; Teo, B. J Phys Chem 1995, 99, 11339.
26. Murcko, M. A.; DiPaolo, R. A. J Am Chem Soc 1992, 114, 10010.
27. Mezei, M. J Comp Phys 1987, 68, 237.
28. Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. J Chem Phys 1992, 97, 3639.
29. Naidoo, K. J.; Brady, J. W. J Am Chem Soc 1999, 121, 2244.
30. Ferrenberg, A. M.; Swendsen, R. H. Phys Rev Lett 1988, 61, 2635.
31. Ferrenberg, A. M.; Swendsen, R. H. Phys Rev Lett 1989, 63, 1195.
32. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. J Comput Chem 1992, 13, 1011.
33. Crouzy, S.; Baudry, J.; Smith, J. C.; Roux, B. J Comput Chem 1999, 20, 1644.
34. Kumar, S.; Payne, P. W.; Vasquez, M. J Comput Chem 1997, 17, 1269.
35. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. J Comput Chem 1995, 16, 1339.
36. Bartels, C.; Karplus, M. J Comput Chem 1997, 18, 1450.
37. Bartels, C.; Karplus, M. J Phys Chem B 1998, 102, 865.
38. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187.
39. van Gunsteren, W. F.; Berendsen, H. J. C. Mol Phys 1977, 34, 1311.
40. Brown, G. M.; Levey, H. A. Science 1965, 147, 1038.
41. Dauchez, M.; Derreumaux, P.; Vergoten, G. J Comput Chem 1992, 14, 263.
42. Brady, J. W. Adv Biophys Chem 1990, 1, 155.